Fast quantum tomography with optimal error bounds

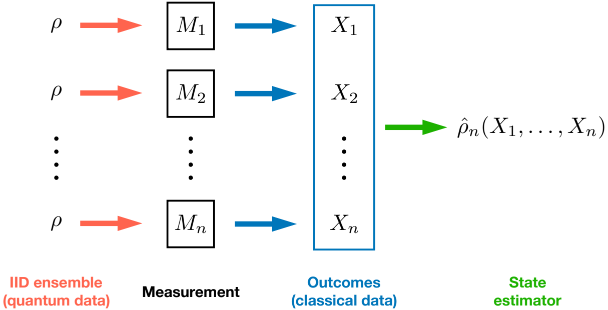
Mădălin Guță

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., J. Phys. A (2019), arXiv:1901.07991

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

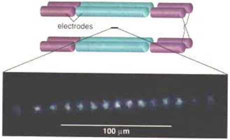


Key questions:

- measurement design: which measurements to perform?
- estimation method: which estimators are accurate and fast?
- confidence regions: how to define computationally feasible error bars ?
- statistical model: how to deal with large dimensional systems?

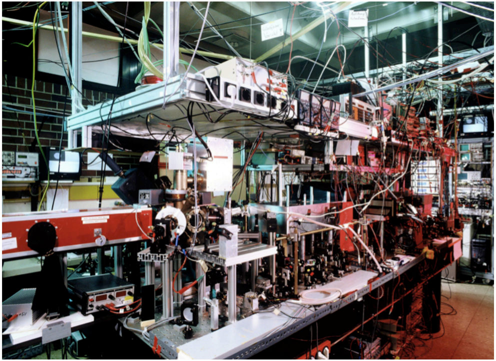
Tomography in Quantum Engineering

- **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)

Outline

- States, measurements, data
- Least Squares estimator
- Projected Least Squares estimator
- Asymptotic rates and Wigner's semicircle law
- Compressed sensing with random bases measurements

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)
  - $\mathrm{Tr}(\rho) = 1$ (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

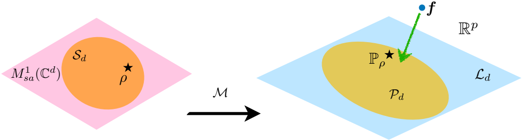
- States of low rank $r \ll d$: manifold of dimension $r(2 \cdot d - r) - 1$

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

- Natural distances: $\tau := \rho_1 - \rho_2$

$$\|\tau\| := \lambda_{\max}(|\tau|), \quad \|\tau\|_2^2 := \mathrm{Tr}(|\tau|^2), \quad \|\tau\|_1 := \mathrm{Tr}(|\tau|), \quad d_b(\rho_1, \rho_2) := 2 - 2 \cdot \mathrm{Tr}\left(\sqrt{\sqrt{\rho_1}\,\rho_2\sqrt{\rho_1}}\right)$$

Measurements, empirical frequencies and estimators

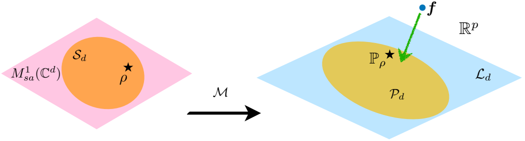


- Measurement with outcomes $\{1,\ldots,p\}$ as a direct map: $\mathcal{M} : M_{sa}(\mathbb{C}^d) \to \mathbb{R}^p$

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

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

Measurements, empirical frequencies and estimators



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

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

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

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

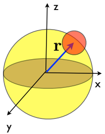
$$f(a) = \frac{\sharp\{i : X_i = a\}}{n}$$

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

Example: spin / two-level ion / qubit tomography

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

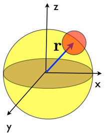
$$\rho_{\mathbf{r}} = \frac{1}{2} \begin{pmatrix} 1 + r_z & r_x - i r_y \\ r_x + i r_y & 1 - r_z \end{pmatrix}$$



Example: spin / two-level ion / qubit tomography

- 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

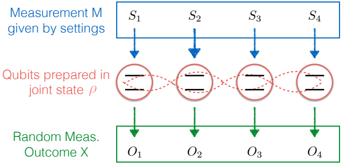
$$\underbrace{|e_x^{\pm}\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ \pm 1\end{pmatrix}}_{s=x} \qquad \underbrace{|e_y^{\pm}\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\ \pm i\end{pmatrix}}_{s=y} \qquad \underbrace{|e_z^{+}\rangle = \begin{pmatrix}1\\ 0\end{pmatrix}, \quad |e_z^{-}\rangle = \begin{pmatrix}0\\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

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

Pauli measurements of (correlated) multiple qubits states



- 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 $\boldsymbol{f} \in \mathbb{R}^{2^k \times 3^k}$

$$p_\rho\left(\{N(\mathbf{o}|\mathbf{s}) : \mathbf{o} \in \{1, -1\}^k\}, \mathbf{s} \in \{x, y, z\}^k\right) = \prod_{\mathbf{s}} \underbrace{\frac{n!}{\prod_{\mathbf{o}} N(\mathbf{o}|\mathbf{s})!} \prod_{X=\mathbf{o}} \mathbb{P}_\rho(\mathbf{o}|\mathbf{s})^{N(\mathbf{o}|\mathbf{s})}}_{\text{multinomial distribution}}$$

Measurement data

- 3^{k} columns of length 2^{k}
- one column for each measurement setting
- each column contains the counts totalling $n = 100$, of the $2^{\mathrm{k}} = 16$ 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

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

where $\mu(d\psi)$ is the Haar measure on the unit ball of $\mathbb{C}^d$

- Symmetrically informationally complete measurement:

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

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

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

- Covariant measurement: continuous set of projections

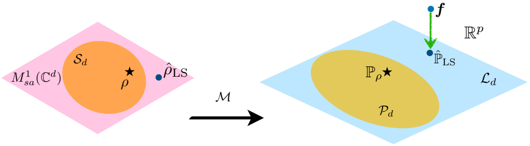
$$M(dv) = d\,|v\rangle\langle v|\,\mu(dv)$$

- stabiliser states....

Outline

- States, measurements, data
- Least Squares estimator
- Projected Least Squares estimator
- Asymptotic rates and Wigner's semicircle law
- Compressed sensing with random bases measurements

Least squares estimator



- Tomography as 'linear regression' problem

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

- Least squares estimator

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

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

Theorem

The following concentration bound holds

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

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

[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

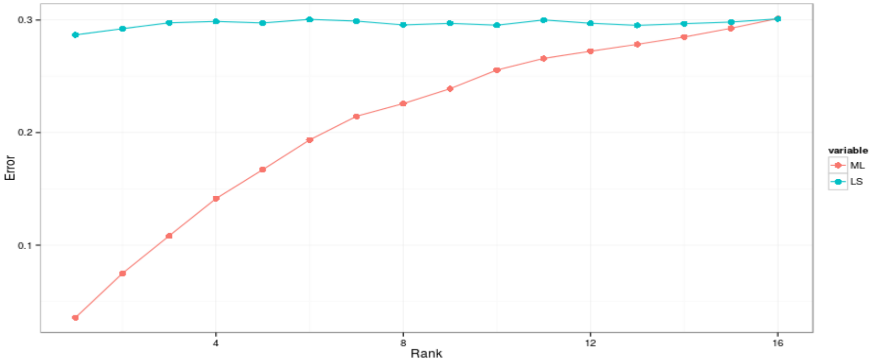
- Use matrix Bernstein inequality[3] for i.i.d. $d \times d$ Hermitian matrices

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

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

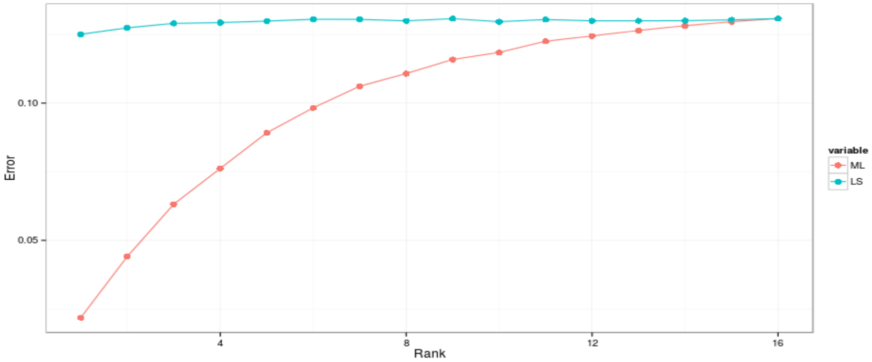
[3] J. A. Tropp, *Found Comput Math* **12** 389-434 (2012)

Simulation: trace-norm error of LS vs ML for Pauli measurements



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

Simulation: trace-norm error of LS vs ML for random bases measurements

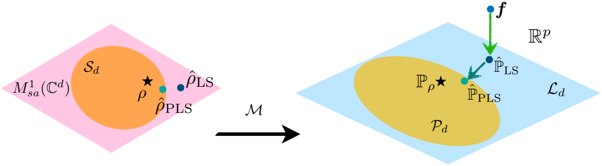


Trace-norm mean errors of LS (green) and ML (red) for rank-r states of 4 atoms, with 200 random bases measurement

Outline

- States, measurements, data
- Least Squares estimator
- Projected Least Squares estimator
- Asymptotic rates and Wigner's semicircle law
- Compressed sensing with random bases measurements

Projected Least squares estimator



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

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

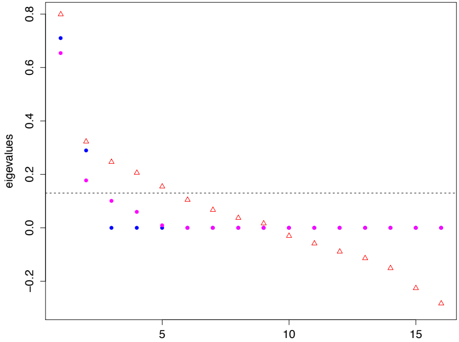
- Solution $\hat{\rho}_{\text{PLS}}$ has the same eigenvectors as $\hat{\rho}_{\text{LS}}$

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

- Simple algorithm for 'adjusting' the eigenvalues

Projection algorithm

- Repeat until all eigenvalues are positive:
  - truncate negative eigenvalues: $\hat{\lambda}_i < 0 \longrightarrow \hat{\lambda}_i = 0$
  - shift remaining eigenvalues so that $\sum_i \hat{\lambda}_i = 1$: $\hat{\lambda}_i \geq 0 \longrightarrow \hat{\lambda}_i - c$



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

Projected Least Squares: concentration bounds[4]

Theorem

Let ρ be a state of unknown rank r.

The following concentration bound holds

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

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

[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

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

3) Concentration bound for LS

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

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

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

PLS + covariant measurement = optimal

Theorem

Let ρ be a state of unknown rank r.

The following concentration bound holds for covariant measurements

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

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]

$$\mathbb{E}\|\rho - \hat{\rho}\|_1 \geq C \frac{r\sqrt{d}}{\sqrt{n}} \quad \Longrightarrow \quad \text{PLS is optimal (up to constant)}$$

[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 $\mathcal{N} = \{|y_1\rangle, \dots |y_m\rangle\}$, $|\mathcal{N}| \approx 3^{2d}$

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

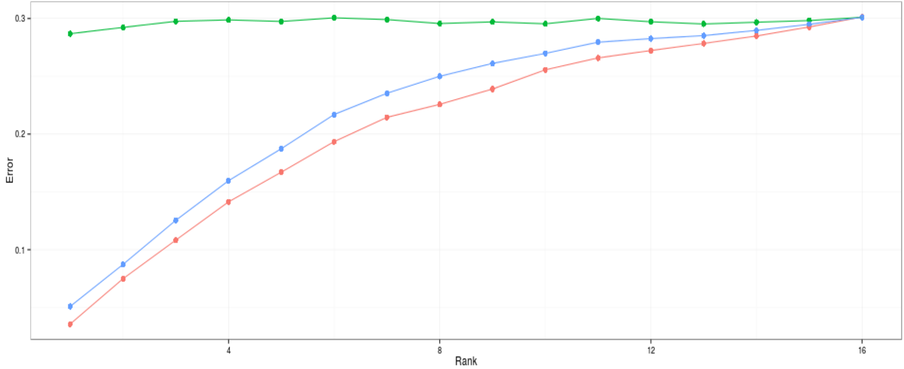
- Classical concentration bound for $\langle y|\hat{\rho}_{\mathrm{LS}} - \rho|y\rangle = \frac{1}{n}\sum_{i=1}^{n}\langle y|A_i|y\rangle$

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

- Union bound

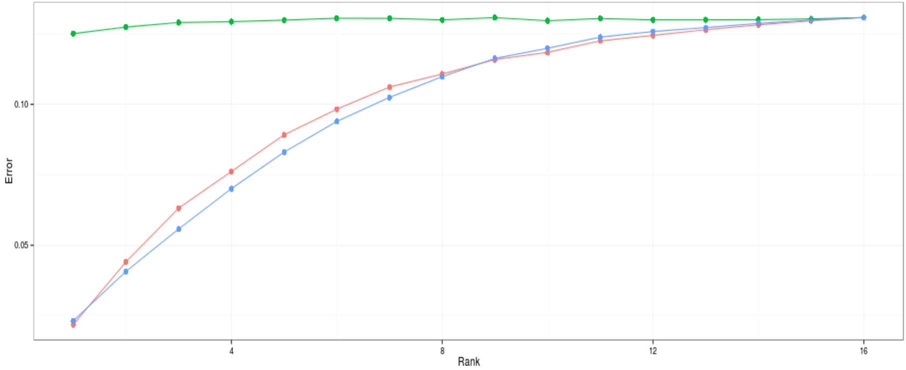
$$\mathbb{P}(\|\hat{\rho}_{\mathrm{LS}} - \rho\| \geq \epsilon) \leq \mathbb{P}\left(\max_i |\langle y_i|\hat{\rho}_{\mathrm{LS}} - \rho|y_i\rangle| \geq \frac{\epsilon}{2}\right) \leq 2 \cdot 3^{2d} \cdot e^{-\frac{n\epsilon^2}{480}}$$

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 error of LS vs PLS vs ML for random bases measurements



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

Outline

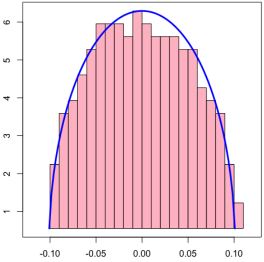
- States, measurements, data
- Least Squares estimator
- Projected Least Squares estimator
- Asymptotic rates and Wigner's semicircle law
- Compressed sensing with random bases measurements

LS asymptotics for covariant measurements

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

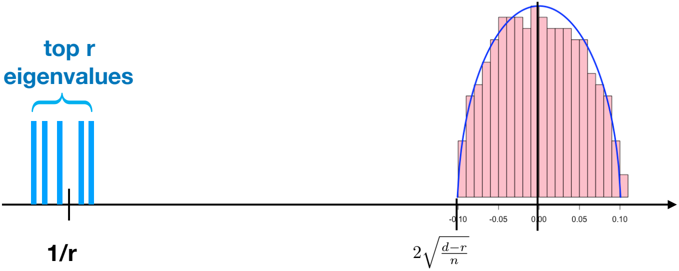
$$\hat{\rho}_{\mathrm{LS}} = \rho + \frac{1}{\sqrt{n}}\Delta = \frac{1}{r}\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} + \frac{1}{\sqrt{n}}\begin{pmatrix} A & B \\ B^* & C \end{pmatrix}$$

- 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



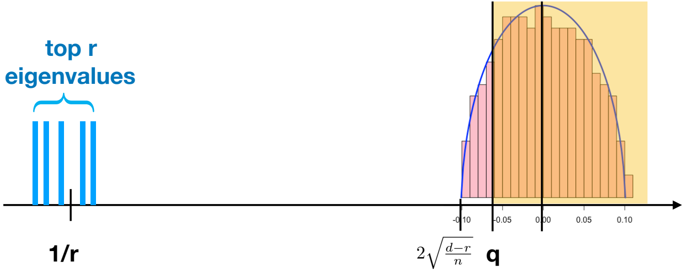
- Asymptotic rates: $\mathbb{E}\|\hat{\rho}_{\mathrm{LS}} - \rho\|_2^2 \approx \frac{d^2}{n}$, $\mathbb{E}\|\hat{\rho}_{\mathrm{LS}} - \rho\|_1 \approx 2\frac{\sqrt{d}}{\sqrt{n}}$

PLS asymptotics



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

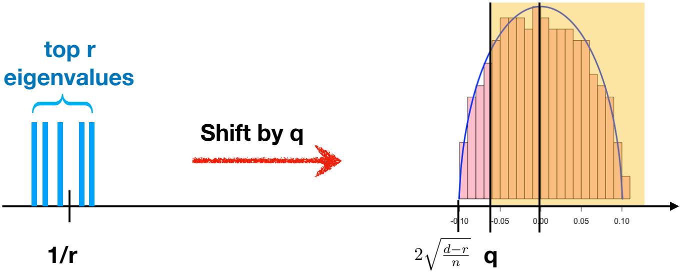
PLS asymptotics



- For $n \gg d \gg 1 \quad \Longrightarrow \quad$ LS has 2 well separated groups of eigenvalues
- Thresholding cut-off point $q = q(r, d)$ can be computed deterministically

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

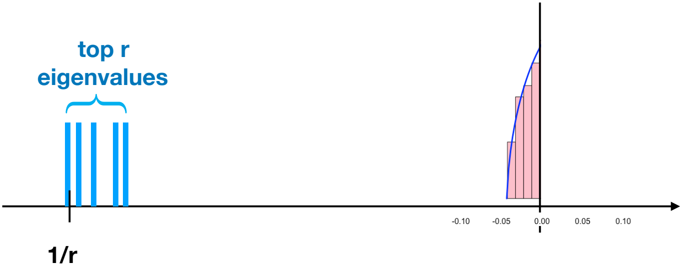
PLS asymptotics



- For $n \gg d \gg 1 \quad \Longrightarrow \quad$ LS has 2 well separated groups of eigenvalues
- Thresholding cut-off point $q = q(r, d)$ can be computed deterministically

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

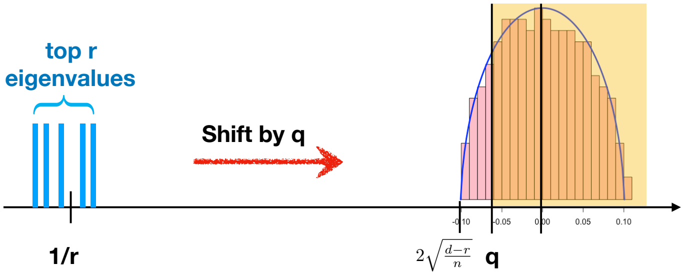
PLS asymptotics



- For $n \gg d \gg 1 \quad \Longrightarrow \quad$ LS has 2 well separated groups of eigenvalues
- Thresholding cut-off point $q = q(r, d)$ can be computed deterministically

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

PLS asymptotics

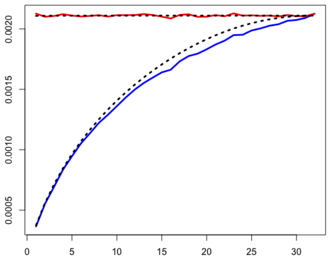


- For $n \gg d \gg 1 \quad \Longrightarrow \quad$ LS has 2 well separated groups of eigenvalues
- Thresholding cut-off point $q = q(r,d)$ can be computed deterministically

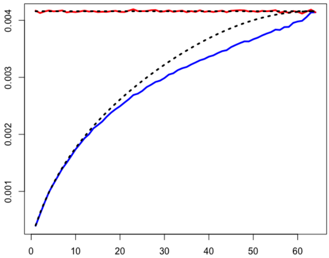
$$rq = \int_q^{2\sqrt{\frac{d-r}{n}}} (x-q)\frac{n}{2\pi}\sqrt{\frac{4(d-r)}{n} - x^2}\, dx$$

- Asymptotic rates: $\mathbb{E}\|\hat{\rho}_{\mathrm{PLS}} - \rho\|_2^2 \approx \frac{6\cdot r\cdot d}{n}, \qquad \mathbb{E}d_B(\hat{\rho}_{\mathrm{PLS}}, \rho) \approx \frac{r\cdot q(r,d)}{\sqrt{n}}$

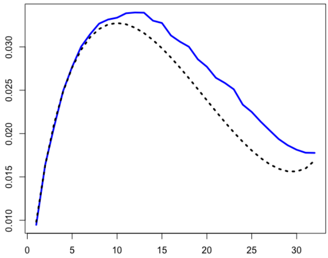
Simulations: Frobenius and Bures errors vs asymptotic expressions



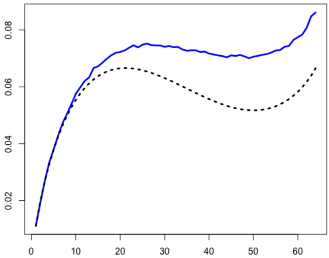
Frobenius error for LS (red) PLS (blue) 5 atoms



Frobenius error for LS (red) PLS (blue) 6 atoms



Bures error for LS (red) PLS (blue) 5 atoms

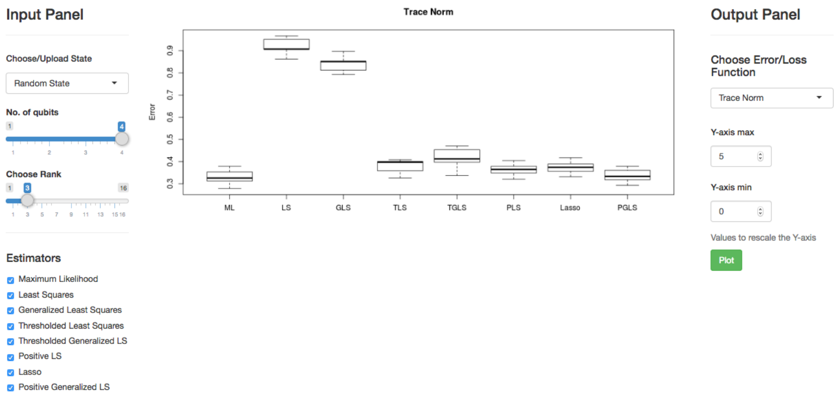


Bures error for LS (red) PLS (blue) 6 atoms

Online quantum tomography tools

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

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

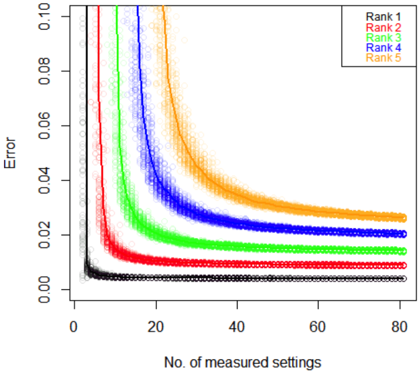


Outline

- States, measurements, data
- Least Squares estimator
- Projected Least Squares estimator
- Asymptotic rates and Wigner's semicircle law
- Compressed sensing with random bases measurements

Can we estimate low rank states with reduced measurement settings ?[6]

- Counting parameters: rank r state $\longrightarrow r \cdot d$ parameters $\longrightarrow \approx r$ settings ($\ll 3^k$)
- Random measurement design:
  choose m random settings $\mathcal{S} := \{\mathbf{s}_1, \ldots, \mathbf{s}_m\}$ and measure each setting $n = \frac{N}{m}$ times
- Mean square error of MLE is stable for a large range of number of settings m



Mean square error $\mathbb{E}\|\hat{\rho}^{(ml)} - \rho\|_2^2$ for 4 ions states of ranks 1-5 and randomly chosen settings

[6] similar to “compressed sensing" D. Gross, *et al*, Phys. Rev. Lett. (2010) but uses “raw" rather than “coarse grained" data

Concentration for Fisher information matrix[7]

- More randomness helps: consider measurements w.r.t. random bases (Haar measure)
- Asymptotics: for large n mean square error of ML estimator scales as in Cramér-Rao bound

$$\|\hat{\rho}^{(ml)} - \rho\|_2^2 \approx \frac{1}{N}\mathrm{Tr}(I(\rho|\mathcal{S})^{-1}G(\rho))$$

- Fisher information matrix (per setting) converges to average

$$I(\rho|\mathcal{S}) = \frac{1}{m}\sum_{i=1}^{m} I(\rho|\mathbf{s}_i) \longrightarrow \bar{I}(\rho) = \int I(\rho|\mathbf{s})d\mathbf{s}$$

Theorem (Fisher info & MSE concentrate with $r \cdot \log rd$ settings)

Let ρ be rank r state with spectrum $(1/r, \dots, 1/r, 0, \dots, 0)$.

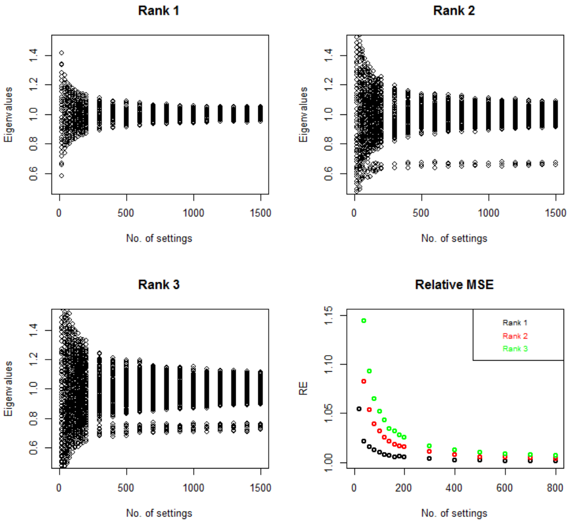
If $m = C(r+1)\log(2(2rd - r^2 - 1)/\delta\epsilon^2)$ then the bounds hold with probability $1-\delta$

$$(1-\epsilon)\bar{I}(\rho) \leq I(\rho|\mathcal{S}) \leq (1+\epsilon)\bar{I}(\rho)$$

$$(1-\epsilon)\mathrm{Tr}\left[\bar{I}(\rho)^{-1}G(\rho)\right] \leq \mathrm{Tr}\left[I(\rho|\mathcal{S})^{-1}G(\rho)\right] \leq (1+\epsilon)\mathrm{Tr}\left[\bar{I}(\rho)^{-1}G(\rho)\right]$$

[7] A. Acharya, T. Kypraios, M.G., New Journal of Physics (2016)

Eigenvalues and MSE concentration



Concentration of the eigenvalues of Fisher information matrix and the MSE for 4 ions states of ranks 1,2,3

Proof

- Matrix Chernoff bound[8]

$$(1-\epsilon)\overline{I}(\rho) \leq I(\rho|\mathcal{S}) \leq (1+\epsilon)\overline{I}(\rho)$$

- Number of settings required (up to log factors)

$$m \approx \frac{\lambda_{\max}}{\lambda_{\min}} := \frac{\max_{\mathbf{s}} \lambda_{\max} I(\rho|\mathbf{s})}{\lambda_{\min}(\bar{I})}$$

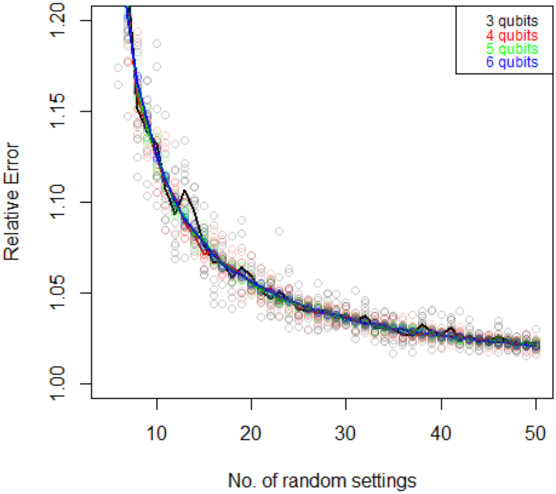
- $\bar{I}$ can be computed explicitly $\longrightarrow \lambda_{\min}(\bar{I}) = r/(r+1)$

- Quantum Cramér-Rao bound

$$I(\rho|\mathbf{s}) \leq F(\rho) \quad \longrightarrow \quad \lambda_{\max} I(\rho|\mathbf{s}) \leq \lambda_{\max} F(\rho) = 2r$$

[8] Ahlswede R. and Winter A., IEEE Transactions Information Theory **48** 569-579 (2002)

Log factors may not be necessary



Relative error w.r.t. asymptotic MSE for random settings, and pure states of 3-6 qubits.

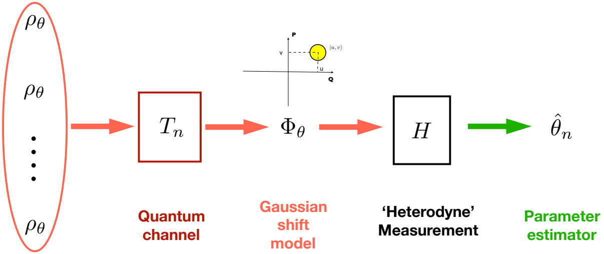
Outlook

- PLS is faster than ML and has (almost) optimal rates for 2-design measuremens
- PLS achieves fundamental lower bound for covariant measurements
- Wigner's semicircle used for computing asymptotic risk of PLS
- $r \log d$ random basis suffice to optimally estimate rank-r states

Future work

- Confidence regions for PLS / boundary problem / LAN not valid
- Can PLS be extended to other statistical models (eg MPS)
- Find minimax risk for Pauli measurements / which states are 'hard' to estimate
- Adaptive and 'compressed sensing' measurements

Tomorrow I will talk about quantum Fisher information & quantum LAN



L. Le Cam

- Quantum Fisher information & quantum Cramér-Rao bound(s) [9]
- LAN: sequence of IID models converges to simpler Gaussian shift model[10] [11] [12] [13]

$$\lim_{n\to\infty} \sup_{\|u\|\leq n^{\epsilon}} \left\| T_n\left(\rho^{\otimes n}_{\theta_0+u/\sqrt{n}}\right) - G(u, V_0)\right\|_1 = 0$$

- LAN is used to derive minimax rates and optimal measurements

[9] foundation work of C. W. Helstrom, A. S. Holevo, V. P. Belavkin, C. M. Caves, R. D. Gill ...

[10] J. Kahn, M.G., *Commun. Math. Phys.* (2009), M.G., B. Janssens and J.Kahn, *Commun. Math. Phys.* (2008)

[11] R.D. Gill, M.G., *I.M.S. Collections* (2012)

[12] C. Butucea, M.G. and M. Nussbaum *Ann. Statist.* (2018)

[13] M.G. J. Kiukas, *J. Math. Phys.* (2017)