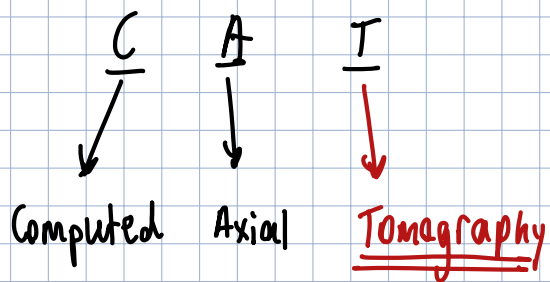


① Tomography

- You have probably heard of (maybe even had!) a CT/CAT scan.



- Also in geophysics $\rightarrow$ Seismology (earthquakes)
- In general: tomography refers to any form of imaging or measurements used to reconstruct a description of an object via the imaging/measurements.

② Quantum State Tomography

- Object of interest is a quantum system — suppose it consists of n qubits
- We want to determine/estimate the state of the system.
- Suppose the (unknown) state is ρ ($\rho \geq 0$, $\rho^\dagger = \rho$, $\text{Tr}[\rho] = 1$).

We want an estimate $\hat{\rho}$ satisfying:

(1) $\hat{\rho}^\dagger = \hat{\rho}$ (Hermitian)

(2) $\hat{\rho} \geq 0$ (positive semi-definite) $\rightarrow$ This requirement is sometimes relaxed!

(3) $\text{Tr}[\hat{\rho}] = 1$

(4) $\frac{1}{2} \|\rho - \hat{\rho}\|_1 \leq \varepsilon$ (Estimate should be close in trace norm).
 (with high probability).

→ Recall trace norm! $\|M\|_1 \rightarrow$ sum of singular values.

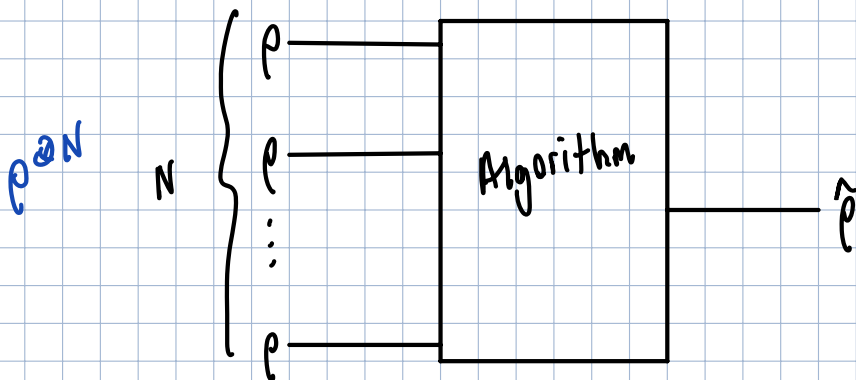
Theorem: $\frac{1}{2} \|\rho_1 - \rho_2\|_1 = \max_{\{M_x\}_x} \frac{1}{2} \sum_x |\text{Tr}[M_x(\rho_1 - \rho_2)]|$

optimize over
all measurements
(POVMs).

⊛ If $\frac{1}{2} \|\rho - \hat{\rho}\|_1 \leq \varepsilon$, then for all measurements / "tests" $\{M_x\}_x$
 we obtain $\frac{1}{2} \sum_x |\text{Tr}[M_x \rho] - \text{Tr}[M_x \hat{\rho}]| \leq \varepsilon$.

- Note: In general, we need multiple copies of ρ to get an estimate.
 (B/c measuring the state destroys it.)

⊛ We also want to minimize the "sample complexity" N



⊛ What is the optimal trade-off b/w: n (# of qubits), N (# of samples),
 ε (accuracy)?

(a) Pauli-based tomography algorithm

Recall: Every ρ can be written as

$$\rho = \frac{1}{2^n} \left(\mathbb{1} + \sum_{\substack{P \in \mathcal{P}_n \\ P \neq \mathbb{1}}} c_P P \right), \quad \mathcal{P}_n = \{\mathbb{1}, X, Y, Z\}^{\otimes n}, \quad c_P = \text{Tr}[P\rho]$$

Pauli expectation values!

Estimator: $\hat{\rho} = \frac{1}{2^n} \left(\mathbb{1} + \sum_{\substack{P \in \mathcal{P}_n \\ P \neq \mathbb{1}}} \hat{c}_P P \right)$

⊛ Recall algorithm from before!

For $i = 1, 2, \dots, N$: ($N \equiv \#$ of samples).

1. Prepare the state ρ

2. Measure each qubit in the basis defined by P .

3. Collect the outcome $\vec{x}_i \in \{0, 1\}^n$

Take the sample mean: $\hat{c}_P(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) := \frac{1}{N} \sum_{i=1}^N \lambda_{\vec{x}_i}$

⊛ Basis defined by P : $P = P_1 \otimes P_2 \otimes \dots \otimes P_n$, $P_i \in \{\mathbb{1}, X, Y, Z\}$

• If $P_i = X \rightarrow$ measure in $\{|+\rangle, |-\rangle\}$ basis $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$\lambda_{x_i} = +1$ if "+" outcome ($x_i = 0$)

$\lambda_{x_i} = -1$ if "-" outcome ($x_i = 1$)

• If $P_i = Y \rightarrow$ measure in $\{|+i\rangle, |-i\rangle\}$ basis

$\lambda_{x_i} = +1$ if "+i" outcome ($x_i = 0$)

$\lambda_{x_i} = -1$ if "-i" outcome ($x_i = 1$)

• If $P_i = Z \rightarrow$ measure in $\{|0\rangle, |1\rangle\}$ basis

$\lambda_{x_i} = +1$ if "0" outcome ($x_i = 0$).

$\lambda_{x_i} = -1$ if "1" outcome ($x_i = 1$)

Analysis: How good is our estimate? (accuracy) What is the sample complexity?

$E[\hat{C}_p] = \text{Tr}(P\rho)$ Recall that $\hat{C}_p$ is an unbiased estimator of $\text{Tr}(P\rho)$:

$$E[\hat{C}_p] = \frac{1}{N} \sum_{i=1}^N E[\hat{\lambda}_i]$$

↑
Estimator for N samples (sample mean)

↑
Estimator for a single sample

⊛ The estimators $\hat{\lambda}_i$ are iid! (Independent and identically distributed)

$$\Rightarrow E[\hat{\lambda}_i] \equiv E[\hat{\lambda}] \quad \forall i \in \{1, 2, \dots, N\}$$

$$E[\hat{\lambda}] = \sum_{\vec{x} \in \{0,1\}^n} P[\hat{\lambda} = \tilde{\lambda}_{\vec{x}}] \tilde{\lambda}_{\vec{x}} = \sum_{\vec{x} \in \{0,1\}^n} \text{Tr}(|v_{\vec{x}}\rangle\langle v_{\vec{x}}| \rho) \tilde{\lambda}_{\vec{x}}$$

$\tilde{\lambda}_{\vec{x}} = \lambda_{x_1} \lambda_{x_2} \dots \lambda_{x_n}$

$\rightarrow = \text{Tr}(|v_{\vec{x}}\rangle\langle v_{\vec{x}}| \rho) \rightarrow |v_{\vec{x}}\rangle \equiv \text{eigenvector of } P \text{ with eigenvalue } \tilde{\lambda}_{\vec{x}}.$

$$= \text{Tr} \left[\left(\sum_{\vec{x} \in \{0,1\}^n} \tilde{\lambda}_{\vec{x}} |v_{\vec{x}}\rangle\langle v_{\vec{x}}| \right) \rho \right]$$

$\underbrace{\sum_{\vec{x} \in \{0,1\}^n} \tilde{\lambda}_{\vec{x}} |v_{\vec{x}}\rangle\langle v_{\vec{x}}|}_{= P}$

$$P = P_1 \otimes P_2 \otimes \dots \otimes P_n$$

$$|v_{\vec{x}}\rangle = |v_{x_1}\rangle \otimes |v_{x_2}\rangle \otimes \dots \otimes |v_{x_n}\rangle$$

$$= \text{Tr}(P\rho)$$

$$\Rightarrow E[\hat{C}_p] = \frac{1}{N} \sum_{i=1}^N E[\hat{\lambda}_i] = \frac{1}{N} \sum_{i=1}^N \text{Tr}(P\rho) = \text{Tr}(P\rho) \quad \checkmark$$

• Now recall the Hoeffding Inequality

PROBABILITY INEQUALITIES FOR SUMS OF BOUNDED RANDOM VARIABLES¹

WASSILY Hoeffding
University of North Carolina

Author(s): Wassily Hoeffding

Source: *Journal of the American Statistical Association*, Vol. 58, No. 301 (Mar., 1963), pp. 13-30

Published by: Taylor & Francis, Ltd. on behalf of the American Statistical Association

$\{X_i\}_{i=1}^N$ independent, bounded random variables, $a_i \leq X_i \leq b_i \quad \forall i \in \{1, 2, \dots, N\}$.

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i, \quad \mu = \mathbb{E}[\bar{X}] \rightarrow \Pr[|\bar{X} - \mu| \geq t] \leq 2 \exp\left(\frac{-2N^2 t^2}{\sum_{i=1}^N (b_i - a_i)^2}\right), \quad t > 0$$

In our case: $a_i = -1 \quad \forall i, \quad b_i = +1 \quad \forall i$ $\bar{X} = \hat{c}_p, \quad \mu = \text{Tr}[P\rho]$.

$$\Rightarrow \Pr[|\hat{c}_p - \underbrace{\text{Tr}[P\rho]}_{=c_p}| \geq t] \leq 2 \exp\left(\frac{-2N^2 t^2}{\sum_{i=1}^N (1)^2}\right) = 2 \exp\left(\frac{-2N^2 t^2}{4N}\right) = 2e^{-\frac{Nt^2}{2}}$$

$$\Rightarrow \Pr[|\hat{c}_p - \text{Tr}[P\rho]| \geq t] \leq 2e^{-\frac{Nt^2}{2}} \quad \forall P \in \mathcal{P}_n \setminus \{\mathbb{1}\}.$$

⊛ We want to make a statement like: $\Pr[\|\hat{\rho} - \rho\|_1 \leq \varepsilon] \geq (1 - e^{-\dots})$

Consider first the 2-norm: $\|\hat{\rho} - \rho\|_2^2 = \text{Tr}[(\hat{\rho} - \rho)(\hat{\rho} - \rho)]$

$$\begin{aligned} \hat{\rho} - \rho &= \frac{1}{2^n} \left(\mathbb{1} + \sum_{\substack{P \in \mathcal{P}_n \\ P \neq \mathbb{1}}} \hat{c}_p P \right) - \frac{1}{2^n} \left(\mathbb{1} + \sum_{\substack{P \in \mathcal{P}_n \\ P \neq \mathbb{1}}} c_p P \right) \\ &\downarrow \\ &= \frac{1}{2^n} \sum_{\substack{P \in \mathcal{P}_n \\ P \neq \mathbb{1}}} (\hat{c}_p - c_p) P \end{aligned}$$

$$\Rightarrow (\hat{\rho} - \rho)^2 = \left(\frac{1}{2^n}\right)^2 \sum_{\substack{P, P' \in \mathcal{P}_n \\ P, P' \neq \mathbb{1}}} (\hat{c}_p - c_p)(\hat{c}_{p'} - c_{p'}) P P'$$

$$\Rightarrow \text{Tr}[(\hat{\rho} - \rho)^2] = \left(\frac{1}{2^n}\right)^2 \sum_{\substack{P, P' \in \mathcal{P}_n \\ P, P' \neq \mathbb{1}}} (\hat{c}_p - c_p)(\hat{c}_{p'} - c_{p'}) \underbrace{\text{Tr}[P P']}_{=2^n \delta_{p,p'}}$$

$$= \frac{1}{2^n} \sum_{\substack{p \in \mathcal{P}_n \\ p \neq 1}} (\hat{c}_p - c_p)^2$$

$$\Rightarrow \|\hat{p} - p\|_2 = \left(\frac{1}{2^n} \sum_{\substack{p \in \mathcal{P}_n \\ p \neq 1}} (\hat{c}_p - c_p)^2 \right)^{1/2}$$

$$\text{Suppose } |\hat{c}_p - c_p| \leq t \equiv \left(\frac{2^n}{4^n - 1} \right)^{1/2} \varepsilon' \quad \forall p \in \mathcal{P}_n \setminus \{1\}.$$

$$\begin{aligned} \Rightarrow \|\hat{p} - p\|_2 &= \left(\frac{1}{2^n} \sum_{\substack{p \in \mathcal{P}_n \\ p \neq 1}} (\hat{c}_p - c_p)^2 \right)^{1/2} \\ &\leq \frac{1}{2^n} \frac{2^n}{4^n - 1} \cdot (4^n - 1) \varepsilon'^2 \quad \forall p \in \mathcal{P}_n \setminus \{1\} \\ &= \varepsilon' \end{aligned}$$

Therefore:

$$\begin{aligned} \Pr[\|\hat{p} - p\|_2 \leq \varepsilon'] &\geq \Pr\left[\bigcap_{\substack{p \in \mathcal{P}_n \\ p \neq 1}} |\hat{c}_p - c_p| \leq t\right] \\ &= 1 - \Pr\left[\bigcup_{\substack{p \in \mathcal{P}_n \\ p \neq 1}} |\hat{c}_p - c_p| \geq t\right] \\ &\geq 1 - \sum_{\substack{p \in \mathcal{P}_n \\ p \neq 1}} \Pr[|\hat{c}_p - c_p| \geq t] \\ &\geq 1 - 2(4^n - 1) e^{-\frac{Nt^2}{2}} \quad \text{from above} \end{aligned}$$

* For events E_1, E_2

$$\Pr[E_1 \cap E_2] = 1 - \Pr[\overline{E_1 \cap E_2}]$$

$$\Pr[\overline{E_1 \cap E_2}] = \Pr[E_1 \cup E_2]$$

* Union bound:

$$\Pr[E_1 \cup E_2] \leq \Pr[E_1] + \Pr[E_2]$$

$$\rightarrow \Pr\left[\bigcup_{i=1}^N E_i\right] \leq \sum_{i=1}^N \Pr[E_i].$$

$$\downarrow$$

$$\approx 1 - 2 \cdot 4^n e^{-\frac{N}{2} \left(\frac{1}{4^n} \varepsilon'\right)^2} \quad t = \left(\frac{2^n}{4^n - 1}\right)^{1/2} \varepsilon' \approx \left(\frac{2^n}{4^n}\right)^{1/2} \varepsilon' = \left(\frac{1}{2^n}\right)^{1/2} \varepsilon'$$

Now, let $d \equiv 2^n$; $N_{tot} = (4^n - 1)N \approx 4^n N = d^2 N \Rightarrow N = \frac{1}{d^2} N_{tot}$

$$\Rightarrow \Pr[\|\hat{\rho} - \rho\|_2 \leq \varepsilon'] \gg 1 - 2 \cdot d^2 e^{-\frac{N_{tot}}{2d^2} \cdot \frac{1}{d} (\varepsilon')^2}$$

We want
trace norm!

$$\downarrow$$

$$= 1 - 2d^2 e^{-\frac{N_{tot}}{2d^3} (\varepsilon')^2}$$

Lemma: For any $M \in L(\mathbb{C}^d)$, $\|M\|_1 \leq \sqrt{d} \|M\|_2$.

Proof: $\|M\|_1 = \text{Tr}[(M^\dagger M)^{1/2}] \leq \left(\text{Tr}[(M^\dagger M)^{1/2} (M^\dagger M)^{1/2}]\right)^{1/2} \left(\text{Tr}[\mathbb{1}]\right)^{1/2}$

* Cauchy-Schwarz inequality: $|\text{Tr}[X^\dagger Y]|^2 \leq \text{Tr}[X^\dagger X] \cdot \text{Tr}[Y^\dagger Y]$

Set $X = (M^\dagger M)^{1/2}$, $Y = \mathbb{1}$

$$\Rightarrow \|M\|_1 \leq \underbrace{\left(\text{Tr}[M^\dagger M]\right)^{1/2}}_{\|M\|_2} \cdot d^{1/2} = \sqrt{d} \|M\|_2. \quad \square$$

Now, set $\varepsilon' = \frac{2\varepsilon}{\sqrt{d}} \Rightarrow \frac{1}{2} \|\hat{\rho} - \rho\|_1 \leq \frac{1}{2} \cdot \sqrt{d} \|\hat{\rho} - \rho\|_2 \leq \frac{\sqrt{d}}{2} \varepsilon' = \frac{\sqrt{d}}{2} \frac{2\varepsilon}{\sqrt{d}} = \varepsilon$

$$\Rightarrow \Pr\left[\frac{1}{2} \|\hat{\rho} - \rho\|_1 \leq \varepsilon\right] \gg 1 - 2d^2 e^{-\frac{N_{tot}}{2d^3} \left(\frac{2\varepsilon}{\sqrt{d}}\right)^2} = 1 - 2d^2 e^{-\frac{2N_{tot}}{d^4} \varepsilon^2}$$

$$\Pr\left[\frac{1}{2} \|\hat{\rho} - \rho\|_1 \leq \varepsilon\right] \gg 1 - 2d^2 e^{-\frac{2N_{tot}}{d^4} \varepsilon^2}$$

⊛ One more thing! We need to ensure our $\hat{\rho}$ is PSD!



use a projection: $\Pi(\hat{\rho}) = \operatorname{argmin} \{ \|\hat{\rho} - \sigma\|_1 : \sigma \geq 0, \operatorname{Tr}(\sigma) = 1 \}$.

can also use 2-norm here.

The map Π takes any input and finds the closest density operator to it.

Observe: if ω is a density operator already, then

$$\begin{aligned} \Pi(\omega) = \omega &\Rightarrow \text{For any } M, \Pi^2(M) = \Pi \circ \underbrace{\Pi(M)}_{\text{density operator}} = \Pi(M) \\ &\Rightarrow \Pi^2 = \Pi \quad (\text{i.e., projection!}) \end{aligned}$$

$$\begin{aligned} \text{If } \frac{1}{2} \|\hat{\rho} - \rho\|_1 \leq \varepsilon, \text{ then } \frac{1}{2} \|\hat{\rho} - \Pi(\hat{\rho})\|_1 &= \min \{ \frac{1}{2} \|\hat{\rho} - \sigma\|_1 : \sigma \geq 0, \operatorname{Tr}(\sigma) = 1 \}. \\ &\downarrow \rightarrow \text{pick } \sigma = \rho \\ &\leq \frac{1}{2} \|\hat{\rho} - \rho\|_1 \\ &\leq \varepsilon \end{aligned}$$

$$\begin{aligned} \Rightarrow \frac{1}{2} \|\Pi(\hat{\rho}) - \rho\|_1 &\leq \underbrace{\frac{1}{2} \|\Pi(\hat{\rho}) - \hat{\rho}\|_1}_{\leq \varepsilon} + \underbrace{\frac{1}{2} \|\hat{\rho} - \rho\|_1}_{\leq \varepsilon} \leq 2\varepsilon \\ &\uparrow \\ &\text{Triangle inequality} \end{aligned}$$

So change $\varepsilon \rightarrow \varepsilon/2$ in above inequality

$$\Rightarrow \Pr \left[\frac{1}{2} \|\Pi(\hat{\rho}) - \rho\|_1 \leq \varepsilon \right] \geq 1 - 2d^2 e^{-\frac{N_{\text{tr}}}{2d^4} \varepsilon^2}$$

- Sample complexity: How many samples do we need to get error ϵ with prob. at least $1-\delta$?

$$\Pr\left[\frac{1}{2}\|\pi(\hat{\rho}) - \rho\|_1 \leq \epsilon\right] \geq 1 - 2d^2 e^{-\frac{N_{\text{tot}}}{2d^4} \epsilon^2} \geq 1 - \delta$$

$$\Rightarrow \delta \geq 2d^2 e^{-\frac{N_{\text{tot}}}{2d^4} \epsilon^2}$$

$$\Rightarrow \log\left(\frac{\delta}{2d^2}\right) \geq -\frac{N_{\text{tot}}}{2d^4} \epsilon^2$$

$$\Rightarrow \frac{N_{\text{tot}}}{2d^4} \epsilon^2 \geq -\log\left(\frac{\delta}{2d^2}\right) = \log\left(\frac{2d^2}{\delta}\right)$$

$$\Rightarrow N_{\text{tot}} \geq \frac{2d^4}{\epsilon^2} \log\left(\frac{2d^2}{\delta}\right)$$

Sample-optimal tomography of quantum states

Jeongwan Haah,^{1,2} Aram W. Harrow,² Zhengfeng Ji,^{3,4,5} Xiaodi Wu,⁶ and Nengkun Yu^{4,3,7}

¹Station Q Quantum Architectures and Computation,
Microsoft Research, Redmond, Washington, USA

²Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

³Centre for Quantum Computation & Intelligent Systems,
Faculty of Engineering and Information Technology,
University of Technology, Sydney, NSW 2007, Australia

⁴Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario, Canada

⁵State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, Beijing, China.

⁶Department of Computer and Information Science, University of Oregon, Eugene, Oregon, USA

⁷Department of Mathematics & Statistics, University of Guelph, Guelph, Ontario, Canada

It is a fundamental problem to decide how many copies of an unknown mixed quantum state are necessary and sufficient to determine the state. Previously, it was known only that estimating states to error ϵ in trace distance required $O(dr^2/\epsilon^2)$ copies for a d -dimensional density matrix of rank r . Here, we give a theoretical measurement scheme (POVM) that requires $O(dr/\delta) \ln(d/\delta)$ copies of ρ to error δ in infidelity, and a matching lower bound up to logarithmic factors. This implies $O((dr/\epsilon^2) \ln(d/\epsilon))$ copies suffice to achieve error ϵ in trace distance. We also prove that for independent (product) measurements, $\Omega(dr^2/\delta^2)/\ln(1/\delta)$ copies are necessary in order to achieve error δ in infidelity. For fixed d , our measurement can be implemented on a quantum computer in time polynomial in n .

→ For the optimal value, we need to go beyond single-copy measurements!

Efficient quantum tomography

Ryan O'Donnell*

John Wright*

September 15, 2015

Abstract

In the quantum state tomography problem, one wishes to estimate an unknown d -dimensional mixed quantum state ρ , given few copies. We show that $O(d/\epsilon)$ copies suffice to obtain an estimate $\hat{\rho}$ that satisfies $\|\hat{\rho} - \rho\|_2^2 \leq \epsilon$ (with high probability). An immediate consequence is that $O(\text{rank}(\rho) \cdot d/\epsilon^2) \leq O(d^2/\epsilon^2)$ copies suffice to obtain an ϵ -accurate estimate in the standard trace distance. This improves on the best known prior result of $O(d^3/\epsilon^2)$ copies for full tomography, and even on the best known prior result of $O(d^2 \log(d/\epsilon)/\epsilon^2)$ copies for spectrum estimation. Our result is the first to show that nontrivial tomography can be obtained using a number of copies that is just *linear* in the dimension.

Next, we generalize these results to show that one can perform efficient principal component analysis on ρ . Our main result is that $O(kd/\epsilon^2)$ copies suffice to output a rank- k approximation $\hat{\rho}$ whose trace distance error is at most ϵ more than that of the best rank- k approximator to ρ . This subsumes our above trace distance tomography result and generalizes it to the case when ρ is not guaranteed to be of low rank. A key part of the proof is the analogous generalization of our spectrum-learning results: we show that the largest k eigenvalues of ρ can be estimated to trace-distance error ϵ using $O(k^2/\epsilon^2)$ copies. In turn, this result relies on a new coupling theorem concerning the Robinson-Schensted-Knuth algorithm that should be of independent combinatorial interest.

(b) Other estimation procedures:

- Least squares (with projection)
- Maximum-likelihood (no projection needed).
- Maximum-entropy principle / mirror descent. (no projection needed).

OPEN ACCESS
IOP Publishing

Made open access 9 October 2020
Journal of Physics A: Mathematical and Theoretical
<https://doi.org/10.1088/1751-8121/ab8111>

J. Phys. A: Math. Theor. 53 (2020) 204001 (28pp)

Fast state tomography with optimal error bounds

M Guţă^{1,4}, J Kahn², R Kueng³ and J A Tropp³

¹ School of Mathematical Sciences, University of Nottingham, United Kingdom

² Institut de Mathématiques de Toulouse, Toulouse, France

³ California Institute of Technology, Pasadena, United States of America

→ Projected least squares

Fast and robust quantum state tomography from few basis measurements

Daniel Stilck França¹, Fernando G.S L. Brandão^{2,3}, and Richard Kueng⁴

¹QMATH, Department of Mathematical Sciences, University of Copenhagen, DK 2100

²AWS Center for Quantum Computing, Pasadena, CA 91125

³Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, CA 91125

⁴Institute for Integrated Circuits, Johannes Kepler University Linz, AT 4040

Maximum-entropy principle / mirror descent.

Learning quantum many-body systems from a few copies

Cambyse Rouzé^{1,*} and Daniel Stilck França^{2,†}

¹Department of Mathematics, Technische Universität München, 85748 Garching, Germany

²QMATH, Department of Mathematical Sciences, University of Copenhagen, Denmark

(Dated: August 25, 2021)

Optimal, reliable estimation of quantum states

Robin Blume-Kohout¹

Institute for Quantum Information, Caltech 107-81, Pasadena, CA 91125, USA
E-mail: robin@blumekohout.com

New Journal of Physics 12 (2010) 043034 (25pp)

Received 7 December 2009

Published 20 April 2010

Online at <http://www.njp.org/>

doi:10.1088/1367-2630/12/4/043034

Maximum-likelihood