Research Project – Report: The Batch Tomography Problem

CR04 – Quantum Information and Computation

Jean-Yves Franceschi

January 3, 2017

Abstract

This report presents the results of my work about the Batch Tomography Problem that has been done in collaboration with Omar Fawzi. The Batch Tomography Problem is an open problem raised by Scott Aaronson in [Aar16] questioning the possibility of estimating for an unknown quantum state ρ the probability of each outcome of a set of two-outcome POVMs of size N using only a number of copies of ρ that would be polynomial in $\log N$.

This document essentially reports the main attempts we made to advance towards a possible solution to this problem. Most of the presented ideas were given to me by Omar Fawzi, whom I would like to thank for his help and and the discussions I could have with him.

Contents

1 Problem Statement

In the proof of an unpublished theorem ("Secret Acceptor Lemma") stating the non-existence of a secure quantum money scheme with limited storage [Aar16, p. 73], Aaronson provides a method which, given a mixed state ρ and a set $\{E_1, \dots, E_N\}$ of two-outcome POVMs[1], provided that $\operatorname{Tr}(E_{i^*}\rho) \geq p$ for some i^*, is able to find E_i such that[2] $\operatorname{Tr}(E_i\rho) \geq p - \epsilon$ using a polylog N number of copies[3] of ρ.

Therefore, finding E_i with large $\operatorname{Tr}(E_i\rho)$ can be done in sublinear time in the number of candidate POVMs. As a side note, Aaronson raises the possibility of a generalization of this result: would it be possible to estimate with any fixed precision each one of the $\operatorname{Tr}(E_i\rho)$'s using polylog N copies of ρ?

Question 1 (The Batch Tomography Problem, [Aar16, p. 77])**.** *Let ρ be an unknown state, $\{E_1, \dots, E_N\}$ a set of two-outcome POVMs, and suppose we have access to k copies $\rho^{\otimes k}$ of state ρ. Is it possible to estimate with fixed minimum probability p, for all $i \in [\![1, N]\!]$, the probability* $\operatorname{Tr}(E_i\rho)$ *of outcome i, with fixed precision ϵ, using only a polylog N number k of copies of ρ?*

Remark. A direct method to solve this problem, regardless of the constraint on the number of state copies, would be determine exactly the $d \times d$ matrix ρ – which is initially unknown – using a *full quantum state tomography* [AJK04], and subsequently compute $\operatorname{Tr}(E_i\rho)$ for all i, giving an exact estimation with probability 1. This can be performed using $\mathcal{O}\left(d^2\right)$ independent measurement on ρ, *i.e.*, using $\mathcal{O}\left(d^2\right)$ copies[4] of ρ. Note that this is a very high cost: as $d = 2^n$, where n is the number of qubits characterized by ρ, this method requires a number of copies that is exponential in n.

The following sections present most of our attempts we made to advance towards a solution to this problem. Section 2 exposes a classical version of this problem that induces a quasi-linear solution. Section 3 presents some methods based on the Quantum Union Bound and Chernoff bound that turned out to be ineffective. Section 4 develops another method based on the extension of the classical method of Section 2, which turned out to be inefficient. Section 5 finally finds a polylog solution to a variant of the problem under specific conditions.

2 Classical Case

A similar classical problem that can serve as an analogy to the Batch Tomography Problem is the following.

Question 2. *Given an unknown random variable X over a set $\mathcal{X}$ and $\mathcal{X}_1, \dots, \mathcal{X}_N \subseteq \mathcal{X}$, how many independent copies $X_1, \dots, X_k$ of X suffice to estimate with fixed minimum probability p, for all $i \in [\![1, N]\!]$,* $\Pr[X \in \mathcal{X}_i]$*, with fixed precision ϵ?*

In this classical version, the quantum state ρ is replaced by a random variable, for which each POVM corresponds to a test of membership to a given set of possible values. The following simple lemma shows that a logarithmic number of independent copies is sufficient to solve this problem.

Lemma 3. *The following estimator solves Question 2 with $k = \mathcal{O}(\log N)$: for every i,* $\Pr[X \in \mathcal{X}_i]$ *is estimated by*

$$Y_i = \frac{|\{j \in [\![1, k]\!] \mid X_j \in \mathcal{X}_i\}|}{k} = \frac{1}{k}\sum_{j=1}^{k} \mathbb{1}_{X_j \in \mathcal{X}_i}.$$

Proof. Let us bound the error probability P_{err} of the estimator.

$$\begin{aligned} P_{\text{err}} &= \Pr\left[\exists i \in [\![1, N]\!] \mid |Y_i - \Pr[X \in \mathcal{X}_i]| \geq \epsilon\right] \\ &\leq \sum_{i=1}^{N} \Pr\left[\left|\frac{1}{k}\sum_{j=1}^{k} \mathbb{1}_{X_j \in \mathcal{X}_i} - \Pr[X \in \mathcal{X}_i]\right| \geq \epsilon\right] \end{aligned}$$

[1] *I.e.*, $(E_i, I - E_i)$ is a POVM: after measurement, E_i accepts ρ with probability $\operatorname{Tr}(E_i\rho)$, and rejects it with probability $1 - \operatorname{Tr}(E_i\rho)$.

[2] With probability at least $1 - \frac{1}{N}$.

[3] *I.e.*, polynomial in $\log N$.

[4] This bound can be improved to $\mathcal{O}\left(3^{\log d}\right)$.

Besides, $\left(\mathbb{1}_{X_j \in \mathcal{X}_i}\right)_j$ is a family of mutually independent random variables over $\{0,1\}$ following the same distribution, and $\mathbb{E}\left[\mathbb{1}_{X_j \in \mathcal{X}_i}\right] = \Pr\left[X \in \mathcal{X}_i\right]$. Therefore, using Chernoff bound [Hoe63]:

$$P_{\text{err}} \leq \sum_{i=1}^{N} 2e^{-2k\epsilon^2} = 2Ne^{-2k\epsilon^2}$$

Since we would want to have P_{err} bounded by a constant, it follows that $k = \mathcal{O}\left(\log N\right)$. Therefore, a logarithmic number of copies of X suffices for this estimator to solve the problem. □

This method, however, can not be transposed to the quantum case; the computations are actually generalizable, but it is impossible to transpose the number of states because a measurement in quantum mechanics may perturb the measured state, thus altering the following measures. In the classical case, once x has been chosen randomly following X_j, we can test if $x \in \mathcal{X}_i$ for all i without perturbing the result, whereas, in the quantum case, once we have measured whether E_i accepts ρ, the resulting state $\widetilde{\rho}$ may be very different from ρ, thus measuring whether E_{i+1} accepts $\widetilde{\rho}$ cannot give much information about the probability of E_{i+1} accepting ρ.

One way to avoid this problem of damaging measurements is to grant $\mathcal{O}\left(\log N\right)$ states per POVM E_i, thus giving to each measurement a fresh state ρ. This gives the following lemma.

Lemma 4 ([Aar16, p. 76])**.** *$\mathcal{O}\left(N \log N\right)$ states are sufficient to solve the Batch Tomography Problem.*

We present in the following sections some attempts to improve this result.

3 Using the Quantum Union Bound

This section exposes some propositions that are based on a result that helps to minimize the impact of damaging states.

3.1 Preliminaries

One way to overcome the problem of damaging measurements is to make good use of the Almost As Good As New Lemma [Aar04], stating intuitively that a two-outcome measurement with a very likely outcome damages slightly the measured state. This result can be generalized to a set of sequentially applied POVMs, each of these having a very probable outcome, as the following Quantum Union Bound states.

Theorem 5 (Quantum Union Bound, [Aar06])**.** *Let $\Lambda_1, \ldots, \Lambda_m$ be two-outcome POVMs such that, for all $i \in [\![1, m]\!]$, $\operatorname{Tr}\left(\Lambda_i \rho\right) \geq 1 - \epsilon$. Then, after applying $\Lambda_1, \ldots, \Lambda_m$ sequentially to ρ, the probability that any of the Λ_i's rejects is at most $m\sqrt{\epsilon}$.*

This allows to solve the Batch Tomography Problem with precision $\epsilon = \frac{1}{2}$ with $\mathcal{O}\left(\log N\right)$ states when the probabilities to estimate are bounded away from $\frac{1}{2}$.

Lemma 6 ([Aar16, p. 76])**.** *Suppose that for all i, $\operatorname{Tr}\left(E_i \rho\right) \in \left[0, \frac{1}{2} - \frac{\delta}{2}\right] \cup \left[\frac{1}{2} + \frac{\delta}{2}, 1\right]$, with $\delta > 0$. Then $\mathcal{O}\left(\log N\right)$ fresh states suffice to solve the problem with precision $\frac{1}{2}$.*

Proof. Let us recall that, in this case with precision $\frac{1}{2}$, a valid estimation should output 1 if $\operatorname{Tr}\left(E_i \rho\right) \in \left[\frac{1}{2} + \frac{\delta}{2}, 1\right]$, and 0 otherwise.

Let us define, for any $i \in [\![1, N]\!]$, the following two-outcome POVM:

$$\Pi_i = \sum_{\substack{J \subseteq [\![1,k]\!] \\ |J| \geq \frac{1}{2}}} \bigotimes_{j=1}^{k} \bar{E}_i^{j,J}, \text{ where } \bar{E}_i^{j,J} = \begin{cases} E_i & \text{if } j \in J \\ I - E_i & \text{otherwise} \end{cases}.$$

Π_i is designed to accept $\rho^{\otimes k}$ with probability that is equal to the one of E_i accepting at least half of the states $\rho^{\otimes k}$. The procedure is the following: for i from 1 to N, we estimate $\operatorname{Tr}\left(E_i \rho\right)$ to 1 if Π_i accepts the current state (initially, $\rho^{\otimes k}$); otherwise, the outputted estimation is 0. Since, with k high enough, each

Π_i accepts $\rho^{\otimes k}$ with a probability tending to either 0 or 1, this method avoids the problem of damaging measurements, as the following shows.

Suppose we observe the outcome of Π_i on $\rho^{\otimes k}$; let us assume without loss of generality that $\mathrm{Tr}\left(E_i\rho\right) \in \left[0, \frac{1}{2} - \frac{\delta}{2}\right]$ (the other case is symmetric). As $\Pr\left[E_i \text{ accepts } \rho\right]$ is bounded away from $\frac{1}{2}$, using Chernoff bound with distance to the mean $\frac{\delta}{2}$, we get that, by definition of Π_i:

$$\mathrm{Tr}\left(\Pi_i\rho^{\otimes k}\right) \leq e^{-2k\left(\frac{\delta}{2}\right)^2},$$

which bounds the probability of failure for an isolated estimation. Therefore, using the Quantum Union Bound, the probability P_{err} of any of the estimations to fail can be bounded:

$$P_{\mathrm{err}} \leq N\sqrt{e^{-2k\left(\frac{\delta}{2}\right)^2}} = Ne^{-k\left(\frac{\delta}{2}\right)^2},$$

which immediately gives $k = \mathcal{O}\left(\log N\right)$, for P_{err} to be bounded by a constant. □

What makes this proof fail when the probabilities to estimate are no longer bounded away from $\frac{1}{2}$ is that we cannot predict what will be the outcome of measuring $\rho^{\otimes k}$ using Π_i if $\mathrm{Tr}\left(E_i\rho\right)$ is too close to $\frac{1}{2}$ – *e.g.*, if $\mathrm{Tr}\left(E_i\rho\right) = \frac{1}{2}$, then Chernoff bound does not provide the “amplification” making the failure probability tend to 0. Thus, the states are likely to be damaged after any measure.

3.2 Propositions and Deadlock

The previous proof shows how it is possible to decide in which interval a probability $\mathrm{Tr}\left(E_i\rho\right)$ lies, provided that the value to estimate is bounded away from the separation of these intervals.

An idea that could help to overcome this problem of an “unknown behavior” when approaching this separation is to choose the pivot value – the fraction to which we compare the proportion of accepted states by E_i ($\frac{1}{2}$ in the proof of Lemma 6) – at random, so that the measured states would be damaged with low probability.

Therefore, a first simple idea to estimate $\mathrm{Tr}\left(E_i\rho\right)$ is to perform a dichotomy using random pivot values: a random pivot value t is chosen uniformly at random in $\left[\frac{1}{4}, \frac{3}{4}\right]$; if more than t of the states are accepted by E_1, we iterate on the new interval $[t, 1]$ – otherwise, on $[0, t]$ –, until a sufficient precision is reached. Unfortunately, as the following analysis shows, the number of states required by this method is at least quadratic.

Let us focus only on the first step of each dichotomy. In order for the Quantum Union Bound to be used, the chosen pivot t should not be too close to $\mathrm{Tr}\left(E_i\rho\right)$, which happens with probability at most 2δ, where δ is the deviation to the mean used in Chernoff bound. Furthermore, if t is well chosen, then the choice of the interval must be correct. This gives the following bound on the probability that all dichotomy succeed:

$$\begin{aligned} P_{\mathrm{success}} &\geq \overbrace{\left(1-2\delta\right)^N}^{\text{first pivot of estimation } i \text{ avoid } \mathrm{Tr}\left(E_i\rho\right)\text{'s neighborhood}} \cdot \overbrace{\left(1 - N\Pr\left[\text{a test fails} \mid \text{pivot is well chosen}\right]\right)}^{\text{Quantum Union Bound}} \\ &\geq \left(1-2\delta\right)^N \underbrace{\left(1 - Ne^{-k\delta^2}\right)}_{\text{given by Chernoff bound}} \end{aligned}$$

If we want P_{success} to be lower-bounded by some constant, we must have:

$$k = \mathcal{O}\left(\frac{1}{\delta^2}\log N - \log\left(1 - \frac{1}{\left(1-2\delta\right)^N}\right)\right) = \mathcal{O}\left(\frac{1}{\delta^2}\log N + \frac{1}{\left(1-2\delta\right)^N}\right),$$

which would be exponential in N if δ were constant. Even if δ is set to $\frac{1}{N}$ to prevent the second term from growing exponentially, it makes the first term explode and finally gives a more than quadratic number of states.

This example shows how methods based on a sequential estimation – several measurements being done to the set of states to produce the ith estimation, followed by the same procedure on the resulting set

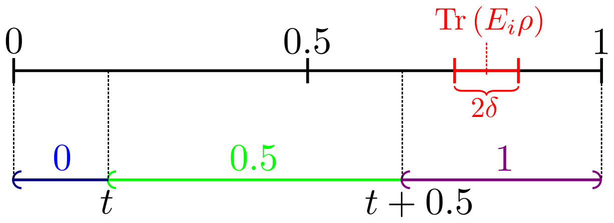

Figure 1: Illustration of discretization using intervals.

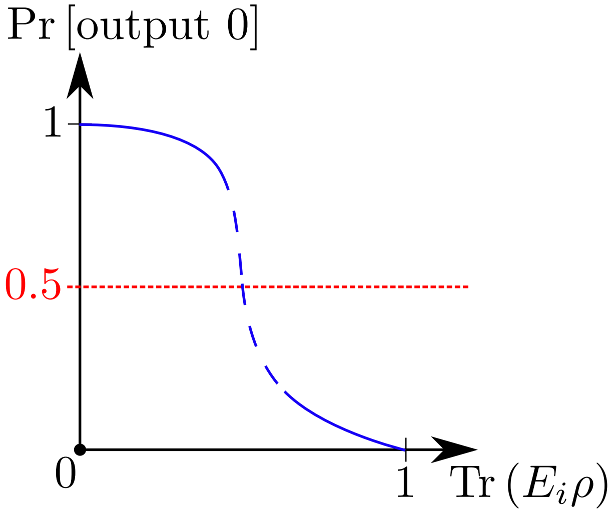

Figure 2: Schematic representation of the reason why aggregation techniques fails.

of states for the $(i+1)$th one – involving randomization to prevent prohibitive damage to the set states can fail when there is a constant probability for each estimation that the set of states is irremediably damaged.

We tried several ideas to overcome this obstacle, mostly based on "hiding" the pivot in the measuring step: instead of measuring whether a given proportion of states are accepted, we tried to design an experiment performing a pivot test on several sets of states, but that does not measure the results right away: it would rather output a value that depends on the results of these tests, value which will be the result of the measure. This technique allows to hide the low probability of getting a bad pivot – which, repeated for all $i \in [\![1, N]\!]$, makes the number of copies explode – inside the measurement, and to leave to the output function the responsibility to output in all cases a very probable value – which can depend on the value to estimate, but whose probability should be able to tend to 1 as the number of states grows –, so that the Quantum Union Bound can be applied. In other words, instead of directly testing one set of states, we test several sets of states and only measure an aggregation of the given results, which will hopefully be designed so that the Quantum Union Bound can be applied.

For instance, a first idea would be the following, for precision $\epsilon = \frac{1}{2}$: choose uniformly at random $t \in [0, 0.5]$, and perform two pivot tests on $\rho^{\otimes m}$ using t and $t+0.5$, which is equivalent to guess in which interval among $\{[0,t), [t, t+0.5), [t+0.5, 1]\}$ the probability to estimate lies. We associate to each possible result a value among $\{0, 0.5, 1\}$, as Figure 1 shows. The experiment is repeated several times – say, n times – on fresh states – *i.e.*, $k = nm$ –, and we only measure the value (0, 0.5 or 1) that appeared the most during those tests.

However, this idea shows that such strategies are very likely to fail. Indeed, if $\mathrm{Tr}(E_i\rho)$ is close to 0.25, then the outcomes 0 and 0.5 will have equivalent probabilities to appear, damaging the whole set of states. We tried several techniques in order to avoid this problem, but it appeared that it could not be solved – the following presents an intuition of why these techniques are bound to fail. Let us consider the simplified case with precision $\frac{1}{2}$, knowing that $\mathrm{Tr}(E_i\rho) \neq \frac{1}{2}$ – *i.e.*, we should output 0 if $\mathrm{Tr}(E_i\rho) < 0.5$, and 1 if $\mathrm{Tr}(E_i\rho) > 0.5$. Techniques based on aggregation would have the following scheme: if $\operatorname{Tr}(E_i\rho) = 0$, then the probability to output 0 will be close to 1, and if $\operatorname{Tr}(E_i\rho) = 1$, the probability to output 0 will be close to 0. Since all aggregations that we tried induced either the continuity of $\Pr[\text{output } 0]$ with respect to $\operatorname{Tr}(E_i\rho)$, or an undetermined behavior in some part of $[0,1]$, $\Pr[\text{output } 0]$ could not be bounded away from $\frac{1}{2}$ (see Figure 2). Thus, we could not prove that any of our ideas did not induce prohibitive damage to the set of states.

4 Extending the Classical Procedure

This section exposes a technique based on an extension of the proof for the classical case developed in Section 2.

4.1 Quantum Equivalent of the Classical Case

An equivalent case of the classical problem where the same proof skeleton can be used is the case where the E_i's are mutually codiagonalizable, i.e. diagonalizable in the same basis.

Lemma 7. *If the E_i's are mutually codiagonalizable, then $\mathcal{O}(\log N)$ suffice to solve the Batch Tomography Problem.*

Proof. Let us assume without loss of generality that the E_i's are diagonal.

Then $E_i = \sum_{j=1}^{d} E_i(j,j) |j\rangle\langle j|$, where $\langle j|$ is the row vector with value 1 at position j and 0 elsewhere, and:

$$\operatorname{Tr}(E_i\rho) = \sum_{j=1}^{d} E_i(j,j) \operatorname{Tr}(|j\rangle\langle j|\rho)$$

Therefore, estimating $\operatorname{Tr}(E_i\rho)$ can be done by estimating $\operatorname{Tr}(|j\rangle\langle j|\rho)$ for all j. Since these values are independent from E_i, they can be estimated only once and used for all E_i's.

Remark. The correspondence with the classical case can be seen here: the previous formula expresses the law of total probability, with $\operatorname{Tr}(|j\rangle\langle j|\rho)$ being the probability to get j from ρ, and $E_i(j,j)$ the probability that E_i accepts ρ knowing that it outputted j.

$\mathcal{B} = (|j\rangle)_{j\in[\![1,d]\!]}$ is a basis of $\mathbb{C}^d$, so we can measure ρ in $\mathcal{B}$. Hence, the procedure is the following: measure each state of $\rho^{\otimes k}$ in $\mathcal{B}$, and store for all $j \in [\![1,d]\!]$ the number N_j of states for which the measure outputted j. Hence, the estimation for $\operatorname{Tr}(E_i\rho)$ is $\sum_{j=1}^{d} E_i(j,j)\frac{N_j}{k}$.

Let P_{err} be the error probability of the whole process, and O_l be the outcome of the lth measurement.

$$\begin{aligned} P_{\text{err}} &= \Pr\left[\exists i \in [\![1,N]\!] \mid \left|\frac{1}{k}\sum_{j=1}^{d} E_i(j,j) N_j - \operatorname{Tr}(E_i\rho)\right| \geq \epsilon\right] \\ &\leq \sum_{i=1}^{N} \Pr\left[\left|\frac{1}{k}\sum_{j=1}^{d} E_i(j,j) N_j - \operatorname{Tr}(E_i\rho)\right| \geq \epsilon\right] \\ &= \sum_{i=1}^{N} \Pr\left[\left|\frac{1}{k}\sum_{l=1}^{k}\left(\sum_{j=1}^{d} E_i(j,j)\mathbb{1}_{O_l=j}\right) - \operatorname{Tr}(E_i\rho)\right| \geq \epsilon\right] \end{aligned}$$

Besides, $\mathbb{E}\left[\sum_{j=1}^{d} E_i(j,j)\mathbb{1}_{O_l=j}\right] = \sum_{j=1}^{d} E_i(j,j)\operatorname{Tr}(|j\rangle\langle j|\rho) = \operatorname{Tr}(E_i\rho)$, and $\sum_{j=1}^{d} E_i(j,j)\mathbb{1}_{O_l=j}$ is real in $[0,1]$, since E_i is hermitian positive semi-definite with eigenvalues in $[0,1]$. So, using Hoeffding's inequality [Hoe63]:

$$P_{\text{err}} \leq \sum_{i=1}^{N} 2e^{-2k\epsilon^2} = 2Ne^{-2k\epsilon^2},$$

which indeed induces $k = \mathcal{O}(\log N)$. □

We discuss in the following of how this method can be generalized to any set of POVMs E_i's.

4.2 Extension

We try in this section to generalize the previous method without any prior hypothesis on the E_i's.

In the general case, $E_i = \sum_{j,j' \in [\![1,d]\!]} E_i(j,j') \left|j\right\rangle \left\langle j'\right|$. However, as $\left(\left|j\right\rangle \left\langle j'\right|\right)_{j,j' \in [\![1,d]\!]}$ is not a POVM (the elements are not all hermitian, for instance), we cannot measure ρ using this family and get the same result as in the previous section. Instead, we can use the family $\left(D_j\right)_j \cup \left(M_{j,j'}\right)_{j \neq j'} \cup \left(M'_{j,j'}\right)_{j \neq j'}$, where:

$$D_j = \frac{1}{2d-1} \left|j\right\rangle \left\langle j\right|, \; M_{j,j'} = \frac{1}{2d-1} \left(\left|j\right\rangle + \left|j'\right\rangle\right) \left(\left\langle j\right| + \left\langle j'\right|\right) \text{ and } M'_{j,j'} \frac{1}{2d-1} \left(\left|j\right\rangle - \left|j'\right\rangle\right) \left(\left\langle j\right| - \left\langle j'\right|\right),$$

which indeed sum to the identity and are hermitian, positive and semi-definite. The outcome of the corresponding measurements will be denoted in the following as, respectively, $\mathfrak{D}_j$, $\mathfrak{M}_{j,j'}$ and $\mathfrak{M}'_{j,j'}$.

Using this measurement basis:

$$\begin{aligned}
\operatorname{Tr}\left(E_i \rho\right) &= \sum_{j,j' \in [\![1,d]\!]} E_i(j,j') \operatorname{Tr}\left(\left|j\right\rangle \left\langle j'\right| \rho\right) \\
&\overset{(a)}{=} \sum_{j \leq j'} E_i(j,j') \operatorname{Tr}\left(\left|j\right\rangle \left\langle j'\right| \rho\right) + \sum_{j > j'} \overline{E_i(j',j)} \operatorname{Tr}\left(\left|j\right\rangle \left\langle j'\right| \rho\right) \\
&\overset{(b)}{=} \sum_{j \leq j'} E_i(j,j') \operatorname{Tr}\left(\left|j\right\rangle \left\langle j'\right| \rho\right) + \sum_{j > j'} \overline{E_i(j',j)} \operatorname{Tr}\left(\left|j'\right\rangle \left\langle j\right| \rho\right) \\
&= \sum_{j} E_i(j,j) \operatorname{Tr}\left(\left|j\right\rangle \left\langle j\right| \rho\right) + 2 \sum_{j < j'} \Re\left(E_i(j,j')\right) \operatorname{Tr}\left(\left|j\right\rangle \left\langle j'\right| \rho\right) \\
&\overset{(c)}{=} \sum_{j} E_i(j,j) \operatorname{Tr}\left(\left|j\right\rangle \left\langle j\right| \rho\right) + \sum_{j < j'} \Re\left(E_i(j,j')\right) \left(\operatorname{Tr}\left(\left(\left|j\right\rangle \left\langle j'\right| + \left|j'\right\rangle \left\langle j\right|\right) \rho\right)\right) \\
\operatorname{Tr}\left(E_i \rho\right) &= (2d-1) \left(\sum_{j} E_i(j,j) \operatorname{Tr}\left(D_j \rho\right) + \sum_{j < j'} \Re\left(E_i(j,j')\right) \operatorname{Tr}\left(\left(M_{j,j'} - D_j - D_{j'}\right) \rho\right) \right),
\end{aligned}$$

where:

- (a) and (b) are obtained by noticing that E_i and ρ are hermitian;
- (c) uses $\operatorname{Tr}\left(\left|j\right\rangle \left\langle j'\right| \rho\right) = \operatorname{Tr}\left(\left|j'\right\rangle \left\langle j\right| \rho\right)$, obtained from (b).

Therefore, if O_l is the outcome of the lth measure and $N_{\mathfrak{D}_j}$ and $N_{\mathfrak{M}_{j,j'}}$ are the number of outcomes, respectively, $\mathfrak{D}_j$ and $\mathfrak{M}_{j,j'}$, the estimation for $\operatorname{Tr}\left(E_i \rho\right)$ is:

$$\frac{2d-1}{k} \left(\sum_{j} E_i(j,j) N_{\mathfrak{D}_j} + \sum_{j < j'} \Re\left(E_i(j,j')\right) \left(N_{\mathfrak{M}_{j,j'}} - N_{\mathfrak{D}_j} - N_{\mathfrak{D}_{j'}} \right) \right),$$

which gives the following bounds on P_{err}:

$$P_{\text{err}} \leq \sum_{i=1}^{N} \Pr\left[\left| \frac{1}{k} \sum_{l=1}^{k} Y_l - \operatorname{Tr}\left(E_i \rho\right) \right| \geq \epsilon \right],$$

where:

$$Y_l = (2d-1) \left(\sum_{j} E_i(j,j) \mathbb{1}_{O_l = \mathfrak{D}_j} + \sum_{j < j'} \Re\left(E_i(j,j')\right) \left(\mathbb{1}_{O_l = \mathfrak{M}_{j,j'}} - \mathbb{1}_{O_l = \mathfrak{D}_j} - \mathbb{1}_{O_l = \mathfrak{D}_{j'}} \right) \right).$$

Hoeffding's inequality can finally be applied, knowing that $E_i(j,j) \in [0,1]$ and $\left|\sum_j \Re\left(E_i(j,j')\right)\right| \leq \sqrt{\text{maximum eigenvalue of } E_i E_i^{\dagger}} \leq 1$ since E_i is hermitian positive semi-definite and POVM; thus, $Y_l \in \left[-2(2d-1), 2(2d-1)\right]$, and:

$$P_{\text{err}} \leq 2Ne^{-\frac{2k\epsilon^2}{(2(2d-1))^2}},$$

which finally gives $k = \mathcal{O}\left(d^2 \log N\right)$.

4.3 Discussion

We extended in the previous subsection the method used for mutually diagonalizable E_i's, but the resulting method requires $\mathcal{O}\left(d^2 \log N\right)$ states, which is worse than the naive method based on the full tomography of ρ (see Section 1).

This problem, however, does not necessarily disqualify this method. Indeed, the d^2 factor results from the choice of the measurement basis $\left(D_j, M_{j,j'}, M'_{,j,j'}\right)$, as it can be observed in the previous computations (it is related to the constraint that any POVM basis should sum to the identity). Therefore, there may be another basis which could lead to a smaller number of states. We tried[5] to find a better basis, but unfortunately failed to do so. Even though it might not be possible to get a pure polylog N number of states (without d appearing in the final result, but it seems likely to appear due to the previously mentioned constraint), there still exists the possibility of an improvement for this method – like, for instance $\mathcal{O}(d \log N)$ instead of $\mathcal{O}\left(d^2 \log N\right)$, result which would be a good progress.

It would also be interesting, if no better measurement basis can be found, to know if this methodology could work provided some specific properties on the E_i's – but more general than having them codiagonalizable: for instance, $\mathcal{O}(\log N)$ states suffice to solve the problem if there are a constant number of non-null off-diagonal elements in each of the E_i's (using a very similar proof)), giving possibly good results for sparse matrices. The goal would be to find wider classes of E_i's that would also give a small number of states.

5 Restricting to Small Inner Products

This section presents a method solving a variant of the Batch Tomography problem under a new condition on the E_i's.

5.1 Result

As all our previous attempts failed to provide an improvement of the $\mathcal{O}(N \log N)$ result mentioned in Section 2, we studied particular cases for which the Batch Tomography Problem seemed difficult to solve, in order to provide possible lower bounds on the number of required states. When studying the case where the E_i's were pure states with little inner product between two different E_i, E_j, we realized that in this case, a variant of the problem – for which the previously presented techniques were ineffective – could be solved using a simple pivot technique.

5.1.1 Preliminaries

Let us assume that for all i and j with $i \neq j$, E_i represents a pure state $|\psi_i\rangle\langle\psi_i|$, and $\mathrm{Tr}(E_i E_j) \leq \eta < 1$ (small inner products between $|\psi_i\rangle$ and $|\psi_j\rangle$). Then the following lemma holds.

Lemma 8. *The size of the set* $\{i \mid \mathrm{Tr}(E_i \rho) > q\}$ *is upper-bounded by* 1 *if* $q > \frac{1}{2-\eta}$.

Proof. Since the E_i's commute, it is possible to bound $\mathrm{Tr}(E_i\rho)$ using $\mathrm{Tr}(E_j\rho)$, by using the results presented in [dMJS79][6]:

$$\begin{aligned}
\mathrm{Tr}(E_i\rho) &= \overbrace{\mathrm{Tr}(E_iE_j)}^{\text{probability of getting } i \text{ from } j} \cdot \overbrace{\mathrm{Tr}(E_j\rho)}^{\text{probability of getting } j \text{ from } \rho} + \overbrace{\mathrm{Tr}(E_i(I-E_j))\,\mathrm{Tr}((I-E_j)\rho)}^{\text{symmetric case where } E_j \text{ does not accept } \rho} \\
&\leq \eta\,\mathrm{Tr}(E_j\rho) + \mathrm{Tr}(E_i)(\mathrm{Tr}(\rho) - \mathrm{Tr}(E_j\rho)) \\
&= \eta\,\mathrm{Tr}(E_j\rho) + 1 - \mathrm{Tr}(E_j\rho) \\
\mathrm{Tr}(E_i\rho) &\leq 1 - (1-\eta)\,\mathrm{Tr}(E_j\rho).
\end{aligned}$$

This gives the following bound on the discrepancy between $\mathrm{Tr}(E_j\rho)$ and $\mathrm{Tr}(E_i\rho)$:

$$\begin{aligned}
\mathrm{Tr}(E_j\rho) - \mathrm{Tr}(E_i\rho) &\geq (2-\eta)\,\mathrm{Tr}(E_j\rho) - 1 \\
&> (2-\eta)\,q - 1 > 0,
\end{aligned}$$

[5] I would like to acknowledge Titouan Carette for his help regarding these attempts.

[6] The main idea is that, if E_i and E_j commute, then there exists a joint probability distribution for the outcome of both POVMs E_i and E_j.

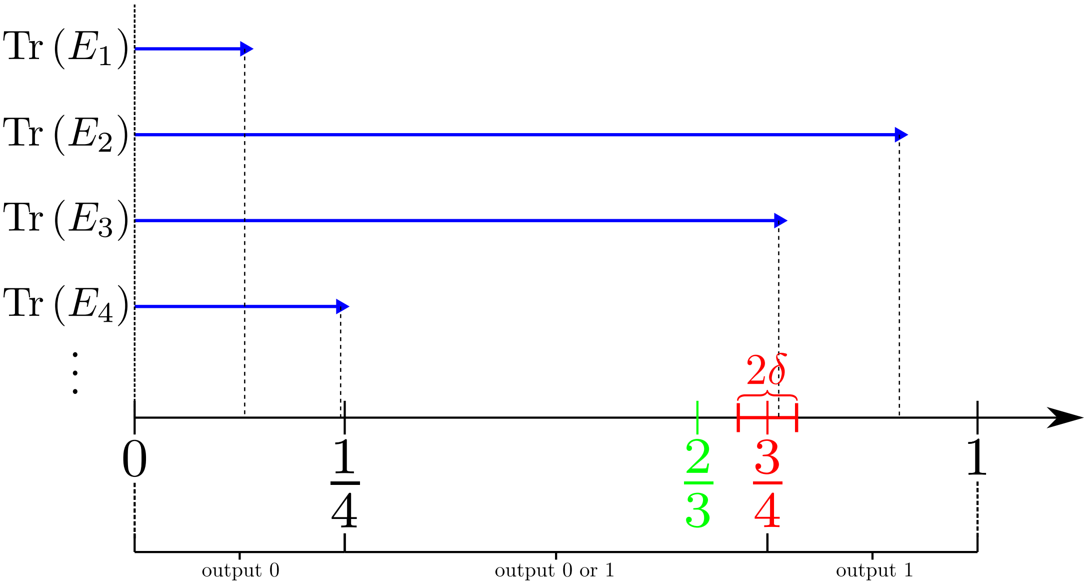

Figure 3: Illustration of the variant of the problem and the settings of the solution.

if $\mathrm{Tr}\,(E_j\rho) > q > \frac{1}{2-\eta}$. □

In the following, we'll more generally assume that there exists $\frac{1}{2} < q < \frac{2}{3}$ such that $|\{i \mid \mathrm{Tr}\,(E_i\rho) > q\}|$ is bounded by a constant (which is true for η small enough in the previous result).

5.1.2 Problem Variant and Method

Let us consider the following problem: given the E_i's and ρ, we would want to estimate $\mathrm{Tr}\,(E_i\rho)$ for all i with precision $\frac{3}{4}$, using the following rules: if $\mathrm{Tr}\,(E_i\rho) < \frac{1}{4}$ (resp. $\mathrm{Tr}\,(E_i\rho) > \frac{3}{4}$), the corresponding estimation should be 0 (resp. 1); otherwise, it can be either 0 or 1 (see Figure 3). This variant corresponds to a version of the problem highlighting the failure of aggregation methods (see Section 3), and its usage is justified by the fact that the Batch Tomography Problem is trivial for precision $\frac{3}{4}$ (one can output $\frac{1}{2}$ for all estimations).

The previous assumption will help us to design a method solving this variant using polylog N states, *i.e.*, that can handle the problem of damaging measurements without having to use $\log N$ states per estimation. Indeed, as the number of E_i's such that $\mathrm{Tr}\,(E_i\rho) > \frac{2}{3}$ is bounded by a constant, it is possible to perform successive pivot tests with pivot value $\frac{3}{4}$, if we are able to detect if the states have been damaged so that we can replace them (which happens a constant number of times) after the measure, since the number of E_i's such that $\mathrm{Tr}\,(E_i\rho)$ falls into $\left[\frac{3}{4}-\delta, \frac{3}{4}+\delta\right]$ is constant (see Figure 3 for an illustration).

Therefore, we need to be able to decide if a state has been damaged, once the measure is done. A reliable method could be the following: instead on manipulating a single set of states of size k, the procedure uses m sets of states, each of those being of size n; the procedure for each estimation consists in performing a pivot test with pivot value $\frac{3}{4}$ separately on each of these sets, and, knowing the results of these measures:

- if all the results of the pivot tests are the same (all 0's, or all 1's), then the procedure outputs this value – this case corresponds to the case where $\mathrm{Tr}\,(E_i\rho) \notin \left[\frac{3}{4}-\delta, \frac{3}{4}+\delta\right]$, i.e. Chernoff bound can be applied;
- otherwise, it outputs 1 – this case corresponds to $\mathrm{Tr}\,(E_i\rho) \in \left[\frac{3}{4}-\delta, \frac{3}{4}+\delta\right]$.

If the estimation that is given for a particular i is 1, then all sets of states are replaced by fresh ones, since it means there could be a high probability that the sets were damaged.

Let us perform an analysis of this method. Let us first study the case where $\mathrm{Tr}\left(E_i\rho\right) \in \left[\frac{3}{4}-\delta, \frac{3}{4}+\delta\right]$, to know if a constant (with respect to N) number of sets of states (with $k = \mathcal{O}\left(\log N\right)$) is sufficient for the correct output value to be very probable if this case.

Chernoff bound – lower bounding the probability for an estimation to lie in a neighborhood of its mean value–, associated with the fact that the binomial probability mass function approaches symmetry when the number of sample grows, allows to state that, for n sufficiently high to use the near-symmetry of the binomial probability mass function and to get a sufficiently large lower bound on the previously mentioned probability, the probability of outputting 1 when $\mathrm{Tr}\left(E_i\rho\right) \in \left[\frac{3}{4}-\delta, \frac{3}{4}+\delta\right]$ becomes high enough if the number of tests m is high enough. Notice that these lower bounds on n and m do not depend on N. Coupled with a constant-bounded number of state replacements and the usage of the Quantum Union Bound as we did in Section 3, this allows us to get that this method works for $n = \mathcal{O}\left(\log N\right)$, *i.e.* $k = \mathcal{O}\left(\log N\right)$.

5.2 Discussion

This result is interesting in the sense that it is a first step towards a possible method working for the Batch Tomography problem using pivot tests, assisted by the Quantum Union Bound. The problem at this point is to be able to extend this method in order to use it for precision $\frac{1}{2}$, which could help for the Batch Tomography problem. As we did not have the time to study the problem further, we propose in the following some ideas that could help to find this extension.

The previous result can be extended to the case where the bound on $\left|\{i \mid \mathrm{Tr}\left(E_i\rho\right) > q\}\right|$ depends on N: the resulting number of states would be $\mathcal{O}\left(\text{bound} \cdot \log N\right)$. If the bound is simply N, then we get the result mentioned in Section 2 (replacing all states for each estimation); but, if we can find an equivalent version of the Batch Tomography Problem with a bound of, for instance, $\mathcal{O}\left(\log N\right)$, then the desired polylog N number of states can be achieved.

We can notice that this method also works symmetrically if the number of E_i's with $\mathrm{Tr}\left(E_i\rho\right) < \frac{1}{3}$ is bounded by a constant. Thus, if we could reduce the Batch Tomography problem – possibly with precision $\frac{1}{2}$ – to the resolution of two problems that are similar to the previous one, with the first one having a bound on the number of E_i's with low probability, and the second one a bound on the number of E_i's with high probability, then we could solve it efficiently.

Another possibility which might extend the previous method is to artificially lower the inner products between the E_i's, by for instance increasing the dimension of the space (this could work since the previous result does not depend on d).

Finally, an analogy with the classical case could be used in order to find new ideas. Indeed, we can see the inner products between the E_i's in this case as a characterization of set coverings between sets $\mathcal{X}_1, \ldots, \mathcal{X}_N$ in the classical case. As for the method presented in Section 4, studying the classical case can help to design new methods for the quantum problem.

6 Conclusion

We tried during this project to provide non-trivial results regarding the Batch Tomography Problem, using successively the Quantum Union bound to prevent prohibitive damage to the states, an extension of the method that is used to solve the classical version of the problem, and a restriction to the case where the E_i's are pure states with small mutual inner products. These attempts were eventually not successful, but provided some possible guidelines for future attempts.

We also worked on other ideas, but the main ones were presented in this report. We also considered the possibility that the answer to the question was negative, but as we began to think about this possibility, we ended up finding that subcases that we thought were hard were actually solvable.

My opinion on this problem would be that the answer to this question is negative, if we want the number of states to be independent from d: it seems, seeing the computations of Section 4, that the quantum case depends a lot on the dimension. I think, however, that a polylog N result, considering d as a constant, could be achieved, or at least a sublinear number of states.

Finally, I would like to thank Omar Fawzi for suggesting this problem to me, and for his help and the discussions I could have with him.

References

[Aar04] S. Aaronson. Limitations of quantum advice and one-way communication. In *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.*, pages 320–332, June 2004. `doi:10.1109/CCC.2004.1313854`.

[Aar06] S. Aaronson. Qma/qpoly /spl sube/ pspace/poly: de-merlinizing quantum protocols. In *21st Annual IEEE Conference on Computational Complexity (CCC'06)*, pages 13 pp.–273, 2006. `doi:10.1109/CCC.2006.36`.

[Aar16] Scott Aaronson. The complexity of quantum states and transformations: From quantum money to black holes, 2016. `arXiv:1607.05256v1`.

[AJK04] Joseph B. Altepeter, Daniel F.V. James, and Paul G. Kwiat. Qubit quantum state tomography. In Matteo Paris and Jaroslav Řeháček, editors, *Quantum State Estimation*, pages 113–145. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004. `doi:10.1007/978-3-540-44481-7_4`.

[dMJS79] Willem M. de Muynck, Peter A. E. M. Janssen, and Alexander Santman. Simultaneous measurement and joint probability distributions in quantum mechanics. *Foundations of Physics*, 9(1):71–122, 1979. `doi:10.1007/BF00715052`.

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association*, 58(301):13–30, 1963.