Lecture 22: Introduction to Quantum Tomography

November 23, 2015
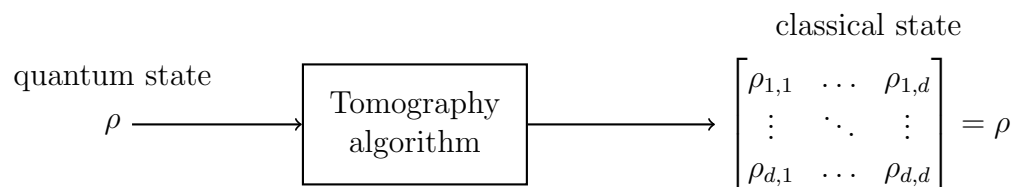
Lecturer: John Wright   *Scribe: Tom Tseng*

1 Motivation

When conducting research in quantum physics, it is essential that, once we create or are otherwise in possession of a quantum state, we can determine what the state is with high accuracy. For instance, researchers have conducted various experiments that support that the quantum teleportation procedure, which we discussed in lecture 3, works. These experiments involve sending quantum states over long distances. But how does the receiving party verify that the state that they receive is the same as the one that was sent by the other party? This is a problem with significant practical importance and is precisely what tomography aims to solve.

2 The problem

2.1 Definition

Roughly speaking, tomography is the problem where we are given an unknown mixed state $\rho \in \mathbb{C}^{d \times d}$ and where our goal is to "learn" what ρ is. We would like to take the input ρ and put it through some tomography algorithm that will output the classical $d \times d$ state describing ρ.

$$
\text{quantum state} \quad \rho \longrightarrow \boxed{\begin{array}{c} \text{Tomography} \\ \text{algorithm} \end{array}} \longrightarrow \overset{\text{classical state}}{\begin{bmatrix} \rho_{1,1} & \cdots & \rho_{1,d} \\ \vdots & \ddots & \vdots \\ \rho_{d,1} & \cdots & \rho_{d,d} \end{bmatrix}} = \rho
$$

This is similar to hypothesis testing, which we have been discussing during the past few lectures. The difference is that we have no prior information as to what ρ is.

Actually, there is a problem with the above description: the goal is impossible to achieve. If we really had an algorithm that could completely learn everything about ρ, then we could replicate ρ, contradicting the no-cloning theorem.

What *can* we say about the state ρ, then? Not much. Suppose we measured ρ with respect to some basis $\{|v_1\rangle, |v_2\rangle, \ldots, |v_m\rangle\}$ and observed $|v_i\rangle$. All we learn is that $\langle v_i| \rho |v_i\rangle$,

the probability of observing $|v_i\rangle$, is non-zero, and maybe we can further guess that $\langle v_i| \rho |v_i\rangle$ is somewhat large. This tells us little useful information about ρ.

Although no perfect tomography algorithm exists, we can make some adjustments to the conditions of the problem to make tomography more tractable. We saw last lecture that when using the Pretty Good Measurement for the hidden subgroup problem, we could generate several copies of the coset state in order to obtain measurement probabilities that were more favorable. That gives our first adjustment: we can stipulate that we are given $n \in \mathbb{N}$ many copies of ρ. In addition, the requirement that we have to output exactly ρ is too restrictive, so our second adjustment will be that we only have to output some $\widetilde{\rho} \in \mathbb{C}^{d \times d}$ that approximates ρ with high probability.
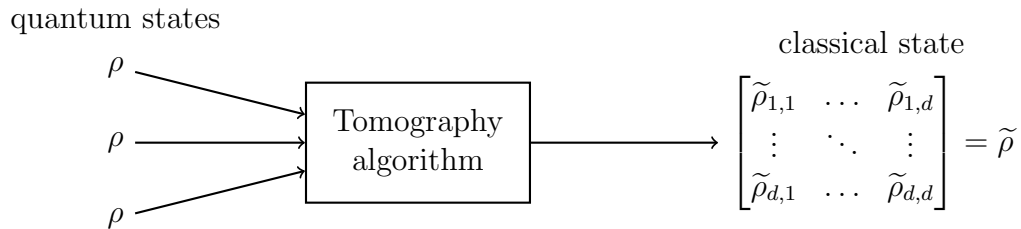
This situation is more sensible. With n copies of ρ, perhaps we can learn enough information to make a fairly precise guess as to what ρ is.

Experiments that test quantum teleportation are verified in this way. In 2012, a group of researchers sent a quantum state between two islands that were 143 miles apart. The result was analyzed by applying quantum tomography to the state, which was sent many times over [MHS$^+$12].

The above discussion motivates the following definition of tomography.

Definition 2.1. In the problem of *quantum state tomography*, the goal is to estimate an unknown d-dimensional mixed state ρ when given $\rho^{\otimes n}$, where $n, d \in \mathbb{N}$. Explicitly, we want a process with the following input and output.

Input: a tensor power $\rho^{\otimes n}$ of the unknown mixed state $\rho \in \mathbb{C}^{d \times d}$, and an error margin $\varepsilon > 0$.

Output: a state $\widetilde{\rho} \in \mathbb{C}^{d \times d}$ such that
$$d_{\mathrm{tr}}(\rho, \widetilde{\rho}) \leq \varepsilon$$
with high probability.

2.2 Optimizing the input

If ρ were given to us in the form $\rho^{\otimes n}$ where n is unimaginably large, we could estimate ρ with great accuracy. However, in reality, preparing so many copies of ρ is expensive. The main question, then, is, given the constraint on the output, *how small can n be?* Here are a couple of results that shed some light on this question.

Fact 2.2. *To solve the tomography problem, it is necessary to have* $n \in \Omega\left(d^2/\varepsilon^2\right)$ *and sufficient to have* $n \in \Theta\left(d^2/\varepsilon^2\right)$.

Fact 2.3. *If ρ is pure, then it is sufficient to have* $n \in \Theta\left(d/\varepsilon^2\right)$.

If ρ is pure, then it is likely also necessary to have $n \in \Omega\left(d/\varepsilon^2\right)$. However, the best known bound is $n \in \Omega\left(d/\left(\varepsilon^2 \log(d/\varepsilon)\right)\right)$ [HHJ+15].

It makes sense that n need not be as great if ρ is pure, since ρ is then a rank-1 matrix that can be fully described as the outer product of a d-dimensional vector.

At first glance, this looks promising. Quadratic and linear complexity with respect to d seems efficient. However an m-qubit state will have dimension $d = 2^m$. Therefore, in the case of fact 2.2, we would need $n \in \Theta\left(4^m/\varepsilon^2\right)$. This fast exponential growth in the size of the input is the biggest obstacle to quantum tomography at present.

In the remainder of this lecture, we will focus on pure state tomography and prove fact 2.3. The general bound from fact 2.2 is considerably more difficult to prove, so we will not discuss it. For more detail, consult the recent paper written by the instructors of this course [OW15].

3 Pure state tomography

3.1 Relation to hypothesis testing

In pure state tomography, we are given $|\psi\rangle \in \mathbb{C}^d$ in the form $|\psi\rangle^{\otimes n}$, and we want to output $|\widetilde{\psi}\rangle \in \mathbb{C}^d$ such that

$$d_{\mathrm{tr}}\left(|\psi\rangle\langle\psi|, |\widetilde{\psi}\rangle\langle\widetilde{\psi}|\right) \le \varepsilon.$$

When looking at pure states and their trace distance, we have the following fact.

Fact 3.1. *For $|\psi\rangle, |\widetilde{\psi}\rangle \in \mathbb{C}^d$, the trace distance of $|\psi\rangle$ and $|\widetilde{\psi}\rangle$ satisfies the following equality:*

$$d_{\mathrm{tr}}\left(|\psi\rangle\langle\psi|, |\widetilde{\psi}\rangle\langle\widetilde{\psi}|\right) = \sqrt{1 - \left|\langle\psi|\widetilde{\psi}\rangle\right|^2}.$$

Thus the goal is to make our guess $|\widetilde{\psi}\rangle$ have large inner product with $|\psi\rangle$.

How do we achieve that? This problem resembles worst-case hypothesis testing. In worst-case hypothesis testing, we are given some state $|\psi\rangle^{\otimes n}$ from some set $\{|v\rangle^{\otimes n}\}_{|v\rangle}$, and we try to guess at $|\psi\rangle$. The difference in tomography, of course, is that the set of possibilities is infinitely large and that we have some margin of error.

This problem is also similar to average-case hypothesis testing, which, as we saw last lecture, was related to worst-case hypothesis testing through the minimax theorem. In tomography, we are not given a probability distribution, so it is similar to the situation where we try to guess $|\psi\rangle$ given a uniformly random $|\psi\rangle^{\otimes n} \sim \{|v\rangle^{\otimes n}\}_{\mathrm{unit}\ |v\rangle}$.

We saw that if we have an optimal algorithm for average-case hypothesis testing, then it is closely related to the optimal algorithm for worst-case hypothesis testing. Thus, to devise the algorithm for tomography, we can get away with just solving the average-case situation on the hardest distribution – the uniform distribution.

3.2 Infinitesimals

The strategy for tomography is that we will merely pretend we are given a probability distribution. With input $|\psi\rangle^{\otimes n}$, we will treat $|\psi\rangle$ as a uniformly random unit vector in $\mathbb{C}^d$.

What does it mean to have a uniformly random unit vector? Since the state is uniformly random, the probability of measuring any particular $|\psi\rangle \in \mathbb{C}^d$ is infinitesimally small. We will write this as

$$\mathbf{Pr}[|\psi\rangle] = d\psi.$$

We need to choose a POVM suitable for tackling this probability distribution. Last lecture, we saw that a good way to deal with average-case hypothesis testing was to run the Pretty Good Measurement, in which, given an ensemble $\{(\sigma_i, p_i)\}_i$ of density matrices and probabilities, we would choose a POVM whose elements resembled the states we knew were in the ensemble. Namely, we took each POVM element E_i to be $p_i \sigma_i$ adjusted by a couple of $S^{1/2}$ factors.

We will reappropriate that technique for our situation. The version of the Pretty Good Measurement for tomography is a POVM with elements resembling

$$E_{|\psi\rangle} \approx d\psi \cdot |\psi\rangle^{\otimes n} \langle\psi|^{\otimes n}.$$

This gives an infinite number of POVM elements, which means that the typical constraint that $\sum_i E_i = I$ needs to be rethought. We instead consider a new constraint that

$$\int_{|\psi\rangle} E_{|\psi\rangle} = I$$

where I is the $n \times n$ identity matrix.

This talk of infinitesimals and integrals might seem non-rigorous, but it is really just informal shorthand that will be notationally convenient for us. A more rigorous analysis could be achieved by instead taking finite-sized POVMs that approximate this infinite scheme and then considering the limit as you make the finite POVMs more and more accurate (and consequently larger and larger in the number of elements).

(A natural question to ask is, "how would we actually implement measurement with a POVM of infinitely many elements?" We would have to take a finite-sized POVM that approximates the infinite POVM. This is not a big sacrifice; we do a similar discretization of randomness in classical algorithms whenever we implement a randomized algorithm that calls for a real number from the uniform distribution over $[0, 1]$.)

3.3 The algorithm

Now we are ready to present the algorithm for pure state tomography that will prove fact 2.3.

Pure State Tomography Algorithm

Input: We are given $|\psi\rangle^{\otimes n}$, where $|\psi\rangle \in \mathbb{C}^d$. We know n and d, which are natural numbers.

Strategy: Measure with the POVM consisting of $\{E_{|v\rangle}\}_{\text{unit vectors } |v\rangle \in \mathbb{C}^d}$, where

$$E_{|v\rangle} = \binom{n+d-1}{d-1} |v\rangle^{\otimes n} \langle v|^{\otimes n} \, dv.$$

Our output: The measurement will give some vector $|\widetilde{\psi}\rangle \in \mathbb{C}^d$, which is what we will output.

Let's do a quick sanity check by making sure our output is sensible. We have

$$\mathbf{Pr}[|\widetilde{\psi}\rangle] = \mathrm{tr}\left(E_{|\widetilde{\psi}\rangle} |\psi\rangle^{\otimes n} \langle\psi|^{\otimes n} \right) = \binom{n+d-1}{d-1} \mathrm{tr}\left(|\widetilde{\psi}\rangle^{\otimes n} \langle\widetilde{\psi}|^{\otimes n} |\psi\rangle^{\otimes n} \langle\psi|^{\otimes n} \right) d\widetilde{\psi}$$

$$= \binom{n+d-1}{d-1} \left| \langle\psi|\widetilde{\psi}\rangle \right|^{2n} d\widetilde{\psi} \quad (1)$$

where the last equality follows from the cyclic property of the trace operator. Notice that $\left| \langle\psi|\widetilde{\psi}\rangle \right|$ is 1 if $|\widetilde{\psi}\rangle$ is equal to $|\psi\rangle$ and less than 1 otherwise. This signals that the probability of outputting a particular vector is greater if that output vector is close in dot product to the input state, which is what we want.

Before the analyze the accuracy of the algorithm, we need to first verify that $\int_{|\psi\rangle} E_{|\psi\rangle} = I$.

3.4 The integral constraint

The presence of that binomial coefficient in the POVM elements might seem strange, but it is merely a scaling factor that makes the constraint $\int_{|\psi\rangle} E_{|\psi\rangle} = I$ hold. Why does this particular coefficient appear, though?

Notice that the input is guaranteed to be of a certain form, namely, a tensor power with specific dimensions. Naturally, we want to understand the space in which these inputs live.

Definition 3.2. The vector space $\mathrm{Sym}(n, d)$ is $\mathrm{span}\{||v\rangle^{\otimes n}\rangle : |v\rangle \in \mathbb{C}^d\}$.

The notation stands for "symmetric" since vectors of the form $|v\rangle^{\otimes n}$ are, in a way, symmetric as a tensor product. As an abbreviation, we will simply refer to this space as Sym with the n and d parameters being implied.

The binomial coefficient comes from the dimension of this space.

Lemma 3.3. *The dimension of* $\mathrm{Sym}(n, d)$ *is* $\binom{n+d-1}{d-1}$.

We will not prove this lemma. (It might show up as a homework problem.)

We promised that this would make our POVM satisfy the integral constraint, so of course we must now check that

$$\int_{|v\rangle} E_{|v\rangle} = \binom{n+d-1}{d-1} \int_{|v\rangle} |v\rangle^{\otimes n} \langle v|^{\otimes n} \, dv \stackrel{?}{=} I.$$

holds. (Here, the integral is only over unit vectors, and this will be true of all the integrals we consider in the remainder of this lecture.) In order to show this, we will use the following proposition.

Proposition 3.4. *Let*

$$M = \int_{|v\rangle} |v\rangle^{\otimes n} \langle v|^{\otimes n} \, dv = \mathop{\mathbf{E}}_{\text{unit } |v\rangle} \left[|v\rangle^{\otimes n} \langle v|^{\otimes n} \right]$$

and let

$$\Pi = \binom{n+d-1}{d-1} M.$$

Then Π projects onto Sym.

Proof. Let P_{Sym} be the projector onto Sym. We will start by showing that $M = C P_{\text{Sym}}$ for some constant C.

To show this, we only need to show that M zeroes out any vector perpendicular to Sym and that M scales vectors in Sym by a constant factor. This is sufficient since it is easy to see that M is linear.

The first part is easy. Because $|v\rangle^{\otimes n} \in$ Sym, we have that for any $|u\rangle \in \text{Sym}^\perp$,

$$M |u\rangle = \mathop{\mathbf{E}}_{\text{unit}|v\rangle} \left[|v\rangle^{\otimes n} \underbrace{\langle v|^{\otimes n} |u\rangle}_{0} \right] = 0$$

as desired.

For the second part, take any $|u\rangle \in$ Sym and write it as $|u\rangle = |w\rangle^{\otimes n}$ for some $|w\rangle$ in $\mathbb{C}^d$. We will illustrate the case where $n = 1$. (The proof for the general case where $n > 1$ is about the same, except more annoying. For that reason, we will skip it.) Since M is linear, we can assume that $|w\rangle$ is a unit vector without loss of generality. We have

$$M |w\rangle = \mathop{\mathbf{E}}_{\text{unit } |v\rangle} \left[|v\rangle \langle v| \, |w\rangle \right]$$

For arbitrary $|v\rangle$, decompose $|v\rangle$ into

$$|v\rangle = \alpha |w\rangle + |w^\perp\rangle$$

where $|w^\perp\rangle$ is some vector that is perpendicular to $|w\rangle$. Because $|v\rangle$ is random, the coefficient α is random as well. In addition, because $|v\rangle$ is uniformly random, $|v\rangle$ appears with the same

probability as $\alpha\,|w\rangle - |w^{\perp}\rangle$. We use this to break up the expectation before combining it again with linearity of expectation, getting

$$M\,|w\rangle = \mathop{\mathbf{E}}_{\text{unit }|v\rangle}\big[|v\rangle\,\langle v|\,|w\rangle\big] = \mathop{\mathbf{E}}_{\alpha}\big[(\alpha\,|w\rangle + |w^{\perp}\rangle)(\alpha^{\dagger}\,\langle w| + \langle w^{\perp}|)\,|w\rangle\big]$$

$$= \frac{1}{2}\mathop{\mathbf{E}}_{\alpha}\big[(\alpha\,|w\rangle + |w^{\perp}\rangle)(\alpha^{\dagger}\,\langle w| + \underbrace{\langle w^{\perp}|)\,|w\rangle}_{0}\big] + \frac{1}{2}\mathop{\mathbf{E}}_{\alpha}\big[(\alpha\,|w\rangle - |w^{\perp}\rangle)(\alpha^{\dagger}\,\langle w| - \underbrace{\langle w^{\perp}|)\,|w\rangle}_{0}\big]$$

$$= \frac{1}{2}\mathop{\mathbf{E}}_{\alpha}\big[(\alpha\,|w\rangle + |w^{\perp}\rangle)\alpha^{\dagger}\,\langle w|w\rangle\big] + \frac{1}{2}\mathop{\mathbf{E}}_{\alpha}\big[(\alpha\,|w\rangle - |w^{\perp}\rangle)\alpha^{\dagger}\,\langle w|w\rangle\big]$$

$$= \frac{1}{2}\mathop{\mathbf{E}}_{\alpha}\big[(\alpha\,|w\rangle + |w^{\perp}\rangle)\alpha^{\dagger}\,\langle w|w\rangle + (\alpha\,|w\rangle - |w^{\perp}\rangle)\alpha^{\dagger}\,\langle w|w\rangle\big]$$

$$= \mathop{\mathbf{E}}_{\alpha}\big[\alpha\alpha^{\dagger}\,|w\rangle\,\underbrace{\langle w|w\rangle}_{1}\big] = \mathop{\mathbf{E}}_{\alpha}\big[|\alpha|^{2}\big]\,|w\rangle$$

It is not hard to see that $\mathbf{E}_{\alpha}\big[|\alpha|^{2}\big]$ is independent of $|w\rangle$, i.e. is a constant, thanks to the symmetry induced by the uniform randomness of $|v\rangle$. Thus $M\,|w\rangle = C\,|w\rangle$ where we let C be $\mathbf{E}_{\alpha}\big[|\alpha|^{2}\big]$.

All this implies that $M = CP_{\text{Sym}}$ for some constant C. We can find the actual value of the constant by taking the trace of both sides of the equality and performing some manipulations:

$$C\binom{n+d-1}{d-1} = C\dim\text{Sym} = C\operatorname{tr}(P_{\text{Sym}}) = \operatorname{tr}(CP_{\text{Sym}})$$

$$= \operatorname{tr}(M) = \mathop{\mathbf{E}}_{\text{unit }|v\rangle}\operatorname{tr}\big(|v\rangle^{\otimes n}\,\langle v|^{\otimes n}\big) = \mathop{\mathbf{E}}_{\text{unit }|v\rangle}\langle v|v\rangle^{n} = 1$$

where the first equality uses lemma 3.3, the penultimate equality uses the cyclic property of the trace operator, and the last equality uses the observation that $|v\rangle$ is a unit vector. Dividing through by $\binom{n+d-1}{d-1}$ gives

$$C = \frac{1}{\binom{n+d-1}{d-1}},$$

and therefore

$$\Pi = \frac{1}{C}M = P_{\text{Sym}}.$$

This proves that Π indeed projects onto Sym. $\qquad\square$

This proposition tells us that $\int_{|v\rangle} E_{|v\rangle} = \binom{n+d-1}{d-1}\int_{|v\rangle}|v\rangle^{\otimes n}\,\langle v|^{\otimes n}\,dv = \Pi$ projects onto Sym, and consequently it is the identity within Sym. Since Sym is the only subspace our POVM operates upon, this shows $\int_{|v\rangle} E_{|v\rangle} = I$ as desired.

3.5 Algorithm analysis

We have proven that the POVM in our algorithm is valid, and now we will show that our algorithm is accurate.

The previous proposition gives the following corollary.

7

Corollary 3.5. *For any unit $|\psi\rangle \in \mathbb{C}^d$, the equality*

$$\mathop{\mathbf{E}}_{unit\ |v\rangle} \left[|\langle v|\psi\rangle|^{2n} \right] = \frac{1}{\binom{n+d-1}{d-1}}.$$

holds.

Proof. We know that Π projects onto Sym and that $|\psi\rangle^{\otimes n} \in$ Sym. Thus $\Pi |\psi\rangle^{\otimes n} = |\psi\rangle^{\otimes n}$, which gives

$$\mathrm{tr}\left(\Pi |\psi\rangle^{\otimes n} \langle\psi|^{\otimes n} \right) = \mathrm{tr}\left(|\psi\rangle^{\otimes n} \langle\psi|^{\otimes n} \right) = \mathrm{tr}(\langle\psi|\psi\rangle^n) = 1.$$

We also have

$$\mathrm{tr}\left(\Pi |\psi\rangle^{\otimes n} \langle\psi|^{\otimes n} \right) = \mathrm{tr}\left(\langle\psi|^{\otimes n} \Pi |\psi\rangle^{\otimes n} \right) = \langle\psi|^{\otimes n} \Pi |\psi\rangle^{\otimes n}$$

$$= \binom{n+d-1}{d-1} \mathop{\mathbf{E}}_{unit\ |v\rangle} \left[\langle\psi|^{\otimes n} |v\rangle^{\otimes n} \langle v|^{\otimes n} |\psi\rangle^{\otimes n} \right] = \binom{n+d-1}{d-1} \mathop{\mathbf{E}}_{unit\ |v\rangle} \left[|\langle v|\psi\rangle|^{2n} \right].$$

Putting these two equalities together, we get

$$\binom{n+d-1}{d-1} \mathop{\mathbf{E}}_{unit\ |v\rangle} \left[|\langle v|\psi\rangle|^{2n} \right] = 1.$$

Dividing both sides by $\binom{n+d-1}{d-1}$ gives the result. $\qquad\square$

The next theorem gives an upper bound on the error of our algorithm.

Theorem 3.6. *Given an input $|\psi\rangle^{\otimes n}$, the pure state tomography algorithm outputs $|\widetilde{\psi}\rangle$ with the property that*

$$\mathop{\mathbf{E}}_{|\widetilde{\psi}\rangle} \left[d_{\mathrm{tr}}\left(|\psi\rangle, |\widetilde{\psi}\rangle \right) \right] \leq \sqrt{\frac{d-1}{n+d}}.$$

Proof. We have

$$\mathop{\mathbf{E}}_{|\widetilde{\psi}\rangle} \left[d_{\mathrm{tr}}\left(|\psi\rangle, |\widetilde{\psi}\rangle \right) \right] = \mathop{\mathbf{E}}_{|\widetilde{\psi}\rangle} \left[\sqrt{1 - \left| \langle\psi|\widetilde{\psi}\rangle \right|^2} \right] \leq \sqrt{1 - \mathop{\mathbf{E}}_{|\widetilde{\psi}\rangle} \left[\left| \langle\psi|\widetilde{\psi}\rangle \right|^2 \right]}.$$

The first equality uses fact 3.1 from earlier in the lecture. The second equality first uses Jensen's inequality on the concavity of the square root operator and then uses linearity of expectation.

Looking only at the inner term, we have

$$\mathop{\mathbf{E}}_{|\widetilde{\psi}\rangle} \left[\left| \langle\psi|\widetilde{\psi}\rangle \right|^2 \right] = \int_{|\widetilde{\psi}\rangle} \left| \langle\psi|\widetilde{\psi}\rangle \right|^2 \mathbf{Pr}\left[|\widetilde{\psi}\rangle \right]$$

$$= \int_{|\widetilde{\psi}\rangle} \left| \langle\psi|\widetilde{\psi}\rangle \right|^2 \binom{n+d-1}{d-1} \left| \langle\psi|\widetilde{\psi}\rangle \right|^{2n} d\widetilde{\psi}$$

$$= \binom{n+d-1}{d-1} \int_{|\widetilde{\psi}\rangle} \left| \langle\psi|\widetilde{\psi}\rangle \right|^{2(n+1)} d\widetilde{\psi}$$

$$= \binom{n+d-1}{d-1} \frac{1}{\binom{n+d}{d-1}} = 1 - \frac{d-1}{n+d}$$

where the second line uses equation (1) from early in the lecture and where the last line uses corollary 3.5.

This finally gives

$$\mathop{\mathbf{E}}_{|\widetilde{\psi}\rangle}\left[d_{\mathrm{tr}}\left(|\psi\rangle,|\widetilde{\psi}\rangle\right)\right] \leq \sqrt{1 - \mathop{\mathbf{E}}_{|\widetilde{\psi}\rangle}\left[\left|\langle\psi|\widetilde{\psi}\rangle\right|^2\right]} = \sqrt{1 - \left(1 - \frac{d-1}{n+d}\right)}$$

which is the desired result. $\qquad\square$

In particular, if we choose $n \in \Theta\left(d/\varepsilon^2\right)$ as in fact 2.3, we get

$$\mathop{\mathbf{E}}_{|\widetilde{\psi}\rangle}\left[d_{\mathrm{tr}}\left(|\psi\rangle,|\widetilde{\psi}\rangle\right)\right] \leq \sqrt{\frac{d-1}{n+d}} \leq \sqrt{d/n} \in \Theta(\varepsilon).$$

This proves that the bound given in fact 2.3, $n \in \Theta\left(d/\varepsilon^2\right)$, is indeed sufficient.

References

[HHJ+15] Jeongwan Haah, Aram W. Harrow, Zhengfeng Ji, Xiaodi Wu, and Nengkun Yu. Sample-optimal tomography of quantum states. 2015.

[MHS+12] Xiao-Song Ma, Thomas Herbst, Thomast Scheidl, Daqing Wang, Sebastian Kropatschek, William Naylor, Alexandra Mech, Bernhard Wittmann, Johannes Kofler, Elena Anisimova, Vadim Makarov, Thomas Jennewein, Rupert Ursin, and Anton Zeilinger. Quantum teleportation using active feed-forward between two canary islands. *Nature*, 489(7415):269–723, 2012.

[OW15] Ryan O'Donnell and John Wright. Efficient quantum tomography. 2015.