

## Measurements, symmetric subspace, pure state estimation

Lecture 2

Michael Walter, Stanford University

These lecture notes are not proof-read and are offered for your convenience only. They include additional detail and references to supplementary reading material. I would be grateful if you email me about any mistakes and typos that you find.

Today we will talk about measurements in quantum mechanics and discuss the problem of estimating an unknown pure state.

## 2.1 Generalized measurements

From your quantum mechanics class you know that observable in quantum mechanics are modeled by Hermitian operators  $X$ . Let  $X = \sum_{x \in \Omega} x P_x$  denote the spectral decomposition of an observable, i.e.,  $P_x$  denotes the projector onto the eigenspace corresponding to an eigenvalue  $x \in \Omega$ . Thus we can repackage  $X$  in terms the collection of projections  $P_x$ , labeled by the possible measurement outcomes  $x \in \Omega$ . This is convenient for two reasons: First, the probability of outcome  $x$  in state  $|\psi\rangle$  is given by the *Born rule*:

$$\Pr(\text{outcome } x) = \langle \psi | P_x | \psi \rangle, \quad (2.1)$$

which is naturally expressed in terms of the projections  $P_x$ . Second, this formalism allows us to consider more general sets of outcomes  $\Omega$  that are not necessarily real numbers. Instead of using observables, we will therefore often prefer to work with the collection of operators  $\{P_x\}_{x \in \Omega}$ . We call  $\{P_x\}_{x \in \Omega}$  a *projective measurement*. Mathematically, it is specified by operators  $P_x$  such that (i)  $P_x \geq 0$ , (ii)  $\sum_x P_x = \mathbb{1}$ , and (iii)  $P_x P_y = \delta_{xy} P_x$ .

Can we think of more general measurement schemes? Suppose we couple our system  $A$  to an auxiliary system  $B$  that is initialized in a fixed state:

$$|\psi\rangle \mapsto |\psi\rangle_A \otimes |0\rangle_B$$

We then apply an arbitrary projective measurement on the joint system, modelled by some  $\{P_{AB,x}\}$ . The subscript  $AB$  reminds us that we are applying a projective measurement on the full system. See fig. 2 for illustration. Then the Born rule eq. (2.1) says that

$$\Pr(\text{outcome } x) = (\langle \psi |_A \otimes \langle 0 |_B) P_{AB,x} (|\psi\rangle_A \otimes |0\rangle_B) = \langle \psi |_A \left( \underbrace{(\mathbb{1}_A \otimes \langle 0 |_B) P_{AB,x} (\mathbb{1}_A \otimes |0\rangle_B)}_{=: Q_x} \right) |\psi\rangle_A,$$

where we have introduced new operators  $Q_x$  on  $\mathcal{H}_A$ . These operators have the property that (i)  $Q_x \geq 0$  and (ii)  $\sum_x Q_x = \mathbb{1}_A$ .

We say call any collection of operators  $\{Q_x\}$  satisfying (i) and (ii) a *generalized measurement* or a *POVM measurement* (POVM is short for positive-operator valued measure). The  $Q_x$  are called *POVM elements*. As we saw above, the Born rule for POVM measurements takes the familiar form

$$\Pr(\text{outcome } x) = \langle \psi | Q_x | \psi \rangle. \quad (2.2)$$



Figure 2: A generalized measurement implemented by coupling the system  $A$  to an auxiliary system  $B$  initialized in a fixed state  $|0\rangle_B$  and performing a projective measurement on the joint system.

A *binary POVM measurement*, i.e., one that has precisely two outcomes, has the form  $\{Q, \mathbb{1} - Q\}$  and is therefore specified by a single POVM element  $0 \leq Q \leq \mathbb{1}$ .

**Remark.** In problem 1.4, you will show any POVM can be implemented in the fashion described above. An alternative way of thinking about a POVM measurement is the following: After coupling to an auxiliary system  $B$ , we apply a unitary  $U_{AB}$  and then perform a projective measurement on the auxiliary system. This fits nicely with our intuitive model of measuring a quantum system – we couple it to an apparatus  $B$ , apply an interacting unitary time evolution, and read off the result at the apparatus.

While eqs. (2.1) and (2.2) look identical, POVM measurements are truly more general than projective measurements. This is because while the projections  $P_x$  are necessarily orthogonal,  $P_x P_y = \delta_{xy} P_x$ , this does not need to be the case for the  $Q_x$ .

**Example.** The four operators  $\frac{1}{2}|0\rangle\langle 0|$ ,  $\frac{1}{2}|1\rangle\langle 1|$ ,  $\frac{1}{2}|+\rangle\langle +|$ ,  $\frac{1}{2}|-\rangle\langle -|$  make up a POVM with four possible outcomes. It can be thought of performing either a projective measurement in the basis  $|0\rangle, |1\rangle$  or in the basis  $|+\rangle, |-\rangle$ , with 50% probability each.

**Example 2.1.** Another example is the POVM that consists of the three (mutually non-orthogonal) operators  $\{\frac{2}{3}|0\rangle\langle 0|, \frac{2}{3}|\alpha^+\rangle\langle \alpha^+|, \frac{2}{3}|\alpha^-\rangle\langle \alpha^-|\}$ , where  $|\alpha^\pm\rangle = \frac{1}{2}|0\rangle \pm \frac{\sqrt{3}}{2}|1\rangle$ . Indeed, it is easily verified that

$$\frac{2}{3}|0\rangle\langle 0| + \frac{2}{3}|\alpha^+\rangle\langle \alpha^+| + \frac{2}{3}|\alpha^-\rangle\langle \alpha^-| = \mathbb{1}.$$

Unlike the previous example, this POVM cannot be decomposed in an interesting way.

In problem 1.3 you will study a state discrimination scenario where POVM measurements outperform projective measurements.

## Continuous POVMs

How can we generalize the concept of a POVM measurement to an infinite set of outcomes  $\Omega$  (e.g., the set of all real numbers  $\mathbb{R}$ , the set of all quantum states, ...)? Let us assume that the space of outcomes  $\Omega$  carries some measure  $dx$ . Then the conditions on  $\{Q_x\}_{x \in \Omega}$  to be a POVM measurement are as follows, (i)  $Q_x \geq 0$ , as before, and (ii)  $\int_{\Omega} dx Q_x = \mathbb{1}$ , and Born's rule now states that

$$p(x) = \langle \psi | Q_x | \psi \rangle$$

is now the *probability density* of the outcome distribution. In other words, probabilities and expectation values can be computed as follows:

$$\begin{aligned}\Pr(\text{outcome } x \in S) &= \int_S dx \langle \psi | Q_x | \psi \rangle, \\ E[f(x)] &= \int dx \langle \psi | Q_x | \psi \rangle f(x).\end{aligned}\tag{2.3}$$

We sometimes say that  $\{Q_x\}$  is a *continuous POVM*.

**Remark.** *This is the most general kind of POVM measurement on a finite-dimensional Hilbert space. In infinite dimensions, one needs a more mathematically sophisticated concept – positive operator-valued measures – which is where the term “POVM” originated (e.g., Holevo, 2011).*

*You might be concerned whether we need an infinite-dimensional auxiliary Hilbert space in order to implement a POVMs with infinitely many outcomes. Interestingly, any continuous POVM on a finite-dimensional Hilbert space can be implemented by performing a discrete POVM chosen at random from a continuous probability distribution (Chiribella et al., 2007). This paper could make for a good course project.*

## Today’s goal: State estimation

Suppose we are given a quantum system and we would like to learn about the underlying quantum state  $|\psi\rangle$ . Is there a measurement that gives us a classical description “ $\psi$ ” of the state  $|\psi\rangle$ ? Clearly, this cannot be done perfectly – since otherwise we could first perform this measurement and then prepare the state from its classical description multiple times, thereby achieving the impossible task of *cloning*:

$$|\psi\rangle \mapsto \text{“}\psi\text{”} \mapsto |\psi\rangle \otimes |\psi\rangle.$$

On the other hand, suppose that we are not given just one copy of a state, but in fact many copies  $|\psi\rangle^{\otimes n}$ . Note that  $\langle \psi^{\otimes n} | \phi^{\otimes n} \rangle = \langle \psi | \phi \rangle^n$ , so if two states are not equal then they rapidly become orthogonal as  $n$  becomes large – suggesting that we can distinguish them arbitrarily well. Of course, since  $\langle \psi | \phi \rangle$  can be arbitrarily close to one this is not yet a completely rigorous argument. But note that in this case the states are essentially the same, and so we make only a small error by conflating them. Thus it seems plausible that we can achieve the following task, known as *pure state estimation*:

We want to design a continuous POVM  $\{Q_{\hat{\psi}}\}$  on  $(\mathbb{C}^d)^{\otimes n}$ , labeled by the pure states on  $\mathbb{C}^d$ , such when we measure on  $|\psi\rangle^{\otimes n}$  we obtain an outcome  $\hat{\psi}$  that is “close” to  $\psi$  (on average, or even with high probability).

To solve this problem and come up with a good measurement for estimating pure states, we need to talk about the symmetries inherent in this problem: If  $|\psi\rangle \in \mathbb{C}^d$  then not only is  $|\psi\rangle^{\otimes n} \in (\mathbb{C}^d)^{\otimes n}$ , but  $|\psi\rangle^{\otimes n}$  is invariant under permuting the subsystems. Let’s make this a bit more precise.

## 2.2 Symmetric subspace

Let  $S_n$  denote the *symmetric group* on  $n$  symbols. Its elements are permutations  $\pi: \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ . Thus,  $S_n$  has  $n!$  elements. This is a *group*, meaning that products and inverses are again contained in  $S_n$ . For any  $\pi \in S_n$ , we can define an operator  $R_\pi$  on the  $n$ -fold tensor power  $(\mathbb{C}^d)^{\otimes n}$  in the following way:

$$R_\pi |\psi_1\rangle \otimes \dots \otimes |\psi_n\rangle = |\psi_{\pi^{-1}(1)}\rangle \otimes \dots \otimes |\psi_{\pi^{-1}(n)}\rangle$$

It is clear that

$$R_1 = \mathbb{1}, \quad R_\tau R_\pi = R_{\tau\pi} \quad (2.4)$$

Indeed, the latter is guaranteed by our judicious use of inverses:

$$\begin{aligned} R_\tau R_\pi |\psi_1\rangle \otimes \dots \otimes |\psi_n\rangle &= R_\tau |\psi_{\pi^{-1}(1)}\rangle \otimes \dots \otimes |\psi_{\pi^{-1}(n)}\rangle \\ &= R_\tau |\psi_{\pi^{-1}(1)}\rangle \otimes \dots \otimes |\psi_{\pi^{-1}(n)}\rangle \\ &= |\psi_{\pi^{-1}(\tau^{-1}(1))}\rangle \otimes \dots \otimes |\psi_{\pi^{-1}(\tau^{-1}(n))}\rangle \\ &= |\psi_{(\tau\pi)^{-1}(1)}\rangle \otimes \dots \otimes |\psi_{(\tau\pi)^{-1}(n)}\rangle \\ &= R_{\tau\pi} |\psi_1\rangle \otimes \dots \otimes |\psi_n\rangle. \end{aligned}$$

Equation (2.4) says that the map  $\pi \mapsto R_\pi$  turns  $(\mathbb{C}^d)^{\otimes n}$  into a *representation* of the symmetric group  $S_n$ .

Let us return to the vectors  $|\psi\rangle^{\otimes n}$ . Clearly, they have the property that  $R_\pi |\psi\rangle^{\otimes n} = |\psi\rangle^{\otimes n}$  for all  $\pi$ . That is,  $|\psi\rangle^{\otimes n}$  are elements of the *symmetric subspace*

$$\text{Sym}^n(\mathbb{C}^d) = \{|\Phi\rangle \in (\mathbb{C}^d)^{\otimes n} : R_\pi |\Phi\rangle = |\Phi\rangle\}.$$

The symmetric subspace is also known as the  $n$ -particle sector of the bosonic Fock space for  $d$  modes.

Given an arbitrary vector  $|\Phi\rangle \in (\mathbb{C}^d)^{\otimes n}$ , we can always symmetrize it to obtain a vector in the symmetric subspace. Indeed, let us define the *symmetrizer*

$$\Pi_n = \frac{1}{n!} \sum_{\pi \in S_n} R_\pi$$

This operator is the projector on the symmetric subspace. Let's verify this: (i) If  $|\Phi\rangle$  is in the symmetric subspace then  $\Pi_n |\Phi\rangle = |\Phi\rangle$ :

$$\Pi_n |\Phi\rangle = \frac{1}{n!} \sum_{\pi \in S_n} R_\pi |\Phi\rangle = \frac{1}{n!} \sum_{\pi \in S_n} |\Phi\rangle = |\Phi\rangle.$$

(ii) For any vector  $|\Phi\rangle \in (\mathbb{C}^d)^{\otimes n}$ , the vector  $|\tilde{\Phi}\rangle = \Pi_n |\Phi\rangle$  is in the symmetric subspace:

$$R_\tau |\tilde{\Phi}\rangle = R_\tau (\Pi_n |\Phi\rangle) = R_\tau \frac{1}{n!} \sum_{\pi \in S_n} R_\pi |\Phi\rangle = \frac{1}{n!} \sum_{\pi \in S_n} R_{\tau\pi} |\Phi\rangle = \frac{1}{n!} \sum_{\pi' \in S_n} R_{\pi'} |\Phi\rangle = \Pi_n |\Phi\rangle = |\tilde{\Phi}\rangle.$$

Here, we used that as  $\pi$  ranges over all permutations, so does  $\pi' = \tau\pi$  (indeed, we obtain any  $\pi'$  exactly from  $\pi = \tau^{-1}\pi'$ ).

In particular, we can obtain a basis of the symmetric subspace by taking a basis  $|i\rangle$  of  $\mathbb{C}^d$ , considering a tensor product basis element  $|i_1, \dots, i_n\rangle$ , and symmetrizing. The result does not depend on the order of the elements, but only on the number of times  $t_i = \#\{i_k = i\}$ . Thus  $\text{Sym}^n(\mathbb{C}^d)$  has the *occupation number basis*

$$|t_1, \dots, t_d\rangle \propto \Pi_n (|1\rangle^{\otimes t_1} \otimes \dots \otimes |d\rangle^{\otimes t_d}), \quad (2.5)$$

where  $t_i \geq 0$  and  $\sum_i t_i = n$ .

**Example** ( $n=2, d=2$ ). A basis of  $\text{Sym}^2(\mathbb{C}^2)$  is given by

$$|2, 0\rangle = |00\rangle, \quad |1, 1\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |01\rangle), \quad |0, 2\rangle = |11\rangle.$$

Note that we can complete this to a basis of  $\mathbb{C}^2 \otimes \mathbb{C}^2$  by adding the antisymmetric singlet state  $(|10\rangle - |01\rangle)/\sqrt{2}$ . It is true more generally that  $(\mathbb{C}^d)^{\otimes 2} = \text{Sym}^2(\mathbb{C}^d) \oplus \Lambda^2(\mathbb{C}^d)$ .

In general, there are  $\binom{n+d-1}{n}$  such basis vectors and therefore

$$\dim \text{Sym}^n(\mathbb{C}^d) = \text{tr} \Pi_n = \binom{n+d-1}{n} = \frac{(n+d-1)!}{n!(d-1)!}.$$

## A resolution of the identity for the symmetric subspace

The reason why we studied the symmetric subspace is that it contains the states  $|\psi\rangle^{\otimes n}$  that arise in our estimation problem. Not every vector in  $\text{Sym}^n(\mathbb{C}^d)$  is of this form – for example,  $\frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$  isn't. Moreover, the  $|\psi\rangle^{\otimes n}$  are not orthogonal. Nevertheless, we have the following alternative formula for the projection onto the symmetric subspace:

$$\Pi'_n = \binom{n+d-1}{n} \int d\psi |\psi\rangle^{\otimes n} \langle \psi|^{\otimes n}. \quad (2.6)$$

The integral requires some explanation: We integrate over all unit vectors  $|\psi\rangle \in \mathbb{C}^d$ , and the measure  $d\psi$  is the unique probability measure that is invariant under the unitary group  $U(d)$ . That is, expectation values do not change when we substitute  $|\psi\rangle \mapsto U|\psi\rangle$ , where  $U$  is a unitary  $d \times d$  matrix. Sometimes this measure is called the *Haar measure*. (Concretely, we can think of the  $|\psi\rangle$  as unit vectors in  $S^{2d-1}$  and the Haar measure can be realized as the unique rotation invariant measure on that sphere.) (Mathematically speaking, I am somewhat conflating the vectors  $|\psi\rangle$  and the pure states  $|\psi\rangle\langle\psi|$  – but if this concerns you then you know how to fix it!) For example, the invariance property immediately implies the following:

$$\Pi'_n = U^{\otimes n} \Pi'_n U^{\dagger, \otimes n}, \quad \text{or} \quad U^{\otimes n} \Pi'_n = \Pi'_n U^{\otimes n} \quad (2.7)$$

One way of interpreting eq. (2.6) is that the vectors  $|\psi\rangle^{\otimes n}$  form an “overcomplete basis” of the symmetric subspace. Indeed, if  $|\Phi\rangle$  is an arbitrary vector then

$$|\Phi\rangle = \Pi_n |\Phi\rangle = \binom{d+n-1}{n} \int d\psi |\psi\rangle^{\otimes n} \langle \psi^{\otimes n} | \Phi \rangle = \int d\psi c_\psi(\Psi) |\psi\rangle^{\otimes n},$$

where  $c_\psi(\Psi) = \binom{d+n-1}{n} \langle \psi^{\otimes n} | \Psi \rangle$ . This means that we can write  $|\Phi\rangle$  as a linear combination of the states  $|\psi\rangle^{\otimes n}$ .

Another way to interpret eq. (2.6), though, is that it shows that

$$Q_{\hat{\psi}} = \binom{d+n-1}{n} |\hat{\psi}\rangle^{\otimes n} \langle \hat{\psi}|^{\otimes n} \quad (2.8)$$

defines a continuous POVM  $\{Q_{\hat{\psi}}\}$  on the symmetric subspace! It is this so-called *uniform POVM* that we will use to solve our estimation problem!

## 2.3 Pure state estimation

We will now solve the problem of pure state estimation (cf. Chiribella, 2010, Brandao et al., 2016, Harrow, 2013). Recall that we are given  $n$  copies of some  $|\psi\rangle^{\otimes n}$ . To obtain a good estimate, we want to measure the uniform POVM (2.8).

How do we quantify the goodness of this strategy? There are several options, but the one that is most natural in the present context is to consider the overlap squared,  $|\langle\psi|\hat{\psi}\rangle|^2$ , between estimate and true state. We will in fact look at a slightly more general figure of merit, namely  $|\langle\psi|\hat{\psi}\rangle|^{2k}$  for some fixed  $k > 0$ , since this is just as easy and we will use it in Tuesday's lecture.

**Remark.** *If  $k > 1$  then this is a more stringent figure of merit since unequal states become more orthogonal in this way:  $|\langle\psi|\hat{\psi}\rangle|^{2k} < |\langle\psi|\hat{\psi}\rangle|^2$ .*

**Remark.** *The overlap has a good operational meaning: In problem 1.2, you will show that two quantum states with overlap close to one are indeed almost indistinguishable by any possible measurement.*

Let us compute the expected value of  $|\langle\psi|\hat{\psi}\rangle|^{2k}$  (the average is over the measurement outcome  $\hat{\psi}$ ):

$$\begin{aligned}
 E [|\langle\psi|\hat{\psi}\rangle|^{2k}] &= \int d\hat{\psi} \langle\psi^{\otimes n}|Q_{\hat{\psi}}|\psi^{\otimes n}\rangle |\langle\psi|\hat{\psi}\rangle|^{2k} \\
 &= \binom{n+d-1}{n} \int d\hat{\psi} |\langle\psi|\hat{\psi}\rangle|^{2(k+n)} \\
 &= \binom{n+d-1}{n} \langle\psi^{\otimes(k+n)}|\left(\int d\hat{\psi} |\hat{\psi}\rangle^{\otimes(k+n)} \langle\hat{\psi}|^{\otimes(k+n)}\right)|\psi^{\otimes(k+n)}\rangle \\
 &= \binom{n+d-1}{n} \binom{n+k+d-1}{n+k}^{-1} \langle\psi^{\otimes(k+n)}|\Pi_{n+k}|\psi^{\otimes(k+n)}\rangle \\
 &= \binom{n+d-1}{n} \binom{n+k+d-1}{n+k}^{-1} \\
 &= \frac{(n+d-1)!}{n!} \frac{(n+k)!}{(n+k+d-1)!} = \frac{(n+d-1)\dots(n+1)}{(n+k+d-1)\dots(n+k+1)} \\
 &\geq \left(\frac{n+1}{n+k+1}\right)^{d-1} = \left(1 - \frac{k}{n+k+1}\right)^{d-1} \\
 &\geq 1 - \frac{k(d-1)}{n+k+1} \geq 1 - \frac{kd}{n}.
 \end{aligned} \tag{2.9}$$

The first equality holds because  $\langle\psi^{\otimes n}|Q_{\hat{\psi}}|\psi^{\otimes n}\rangle$  is the probability density of the measurement outcome  $\hat{\psi}$ , as we know from eq. (2.3). For the second equality, we plugged in the definition of the POVM element eq. (2.8). The third is just some simple manipulation using linearity of the integral, and the fourth follows by plugging in the formula for the projector onto the symmetric subspace  $\text{Sym}^{n+k}(\mathbb{C}^d)$ . The rest are some simple inequalities that I explained in class.

Success! We have shown that the uniform POVM (2.8) gives us a very good estimate of  $|\psi\rangle$  as soon as  $n \gg d$  (if we measure its goodness by the overlap squared, corresponding to  $k = 1$ ).

**Remark.** *Later in this course we will learn how to go beyond the symmetric subspace and solve the state estimation problem for general, not necessarily pure quantum states (lecture 7).*

## Bibliography

Alexander S Holevo. *Probabilistic and statistical aspects of quantum theory*, volume 1. Springer Science & Business Media, 2011.

Giulio Chiribella, Giacomo Mauro D'Ariano, and Dirk Schlingemann. How continuous quantum measurements in finite dimensions are actually discrete, *Physical Review Letters*, 98(19):190403, 2007. arXiv:quant-ph/0702068.

Giulio Chiribella. On quantum estimation, quantum cloning and finite quantum de finetti theorems. In *Conference on Quantum Computation, Communication, and Cryptography*, pages 9–25. Springer, 2010.

Fernando GSL Brandao, Matthias Christandl, Aram W Harrow, and Michael Walter. The Mathematics of Entanglement. 2016. arXiv:1604.01790.

Aram W Harrow. The church of the symmetric subspace. 2013. arXiv:1308.6595.

