CS286.2 Lecture 14: Quantum de Finetti Theorems II

Scribe: Maria Okounkova

1 Statement of the theorem

Recall the last statement of the quantum de Finetti theorem from the previous lecture.

Theorem 1 (Quantum de Finetti)**.** *Let $\rho \in \text{Dens}\left((\mathbb{C}^2)^{\otimes n}\right)$, $k \leq n$ and $t \leq n-k$. Then there exists an $m \leq t$ such that if for any $\vec{i} \in [n]^k$, $\vec{j} \in [n]^m$ and $\vec{x} \in [d^8]^m$, $\rho_{\vec{i}}^{\vec{j},\vec{x}}$ is the post-measurement state on qudits $i_1 \dots i_k$ obtained after measuring $j_i \dots j_m$ using Λ and obtaining outcomes $x_1 \dots x_m$, we have*

$$\mathbb{E}_{\vec{i}}\,\mathbb{E}_{\vec{j},\vec{x}}\left\|\rho_{\vec{i}}^{\vec{j},\vec{x}} - \rho_{i_1}^{\vec{j},\vec{x}} \otimes \dots \otimes \rho_{i_k}^{\vec{j},\vec{x}}\right\|_1^2 \leq \frac{4\ln(d)(18d)^k k^2}{t}. \tag{1}$$

This statement is different from the previous statement of the quantum de Finetti theorem as ρ is not assumed to be permutation invariant (this is replaced by the fact that the indices $\vec{j}$ that are measured, and the indices $\vec{i}$ that are kept, are distributed uniformly in $[n]$ subject to being distinct indices), the expectation over $\vec{x}$ is taken outside (if we move it inside, we obtain a convex combination of products as in the previous de Finetti, but taking the expectation value outside makes the statement stronger), and because the bound has a larger dependence on the dimension (d^k instead of the previous d). This last point is the main drawback of the theorem and is directly due to the use of the informationally complete measurement Λ (it is an open problem whether this can be avoided).

Theorem 1 has a classical analogue, with the same definitions of n, m, k, and t:

Theorem 2 (Classical Analogue)**.** *Let P be a distribution on S^n for some set S of size d. Let $\vec{i} = i_1 \dots i_k$, $\vec{j} = j_1 \dots j_m$, and $\vec{x} = x_1 \dots x_m$ be sampled according to the marginal $P_{\vec{j}}$. Let $P_{\vec{i}}^{\vec{j},\vec{x}}$ be the marginal on $\vec{i}$ conditioned on the variables indexed by $\vec{j}$ being equal to the values specified by $\vec{x}$. Then*

$$\mathbb{E}_{\vec{i}}\mathbb{E}_{\vec{j},\vec{x}}\left\|P_{\vec{i}}^{\vec{j},\vec{x}} - P_{i_1}^{\vec{j},\vec{x}} \otimes \dots P_{i_k}^{\vec{j},\vec{x}}\right\|_1^2 \leq \frac{2k^2 \ln d}{t}. \tag{2}$$

2 Classical $\Longrightarrow$ Quantum

Our path now is to prove that the quantum statement follows from the classical statement, and then prove the classical statement, and finally to apply this to QPCP.

Proof. Consider the distribution P obtained from measuring all qudits in ρ using Λ (we assume that Λ is informationally complete with distortion $\leq (18d)^{k/2}$, as before, and is measured as a tensor product across all quits). By the classical statement of the theorem, we have,

$$\mathbb{E}_{\vec{i}}\mathbb{E}_{\vec{j},\vec{x}}\left\|P_{\vec{i}}^{\vec{j},\vec{x}} - P_{i_1}^{\vec{j},\vec{x}} \otimes \ldots P_{i_k}^{\vec{j},\vec{x}}\right\|_1^2 \leq \frac{2k^2\ln(d^8)}{t},$$

since P is a distribution on the set $[d^8]^n$ of all possible outcomes of Λ. Now observe that by definition $P_{\vec{i}}^{\vec{j},\vec{x}} = \Lambda^{\otimes k}(\rho_{\vec{i}}^{\vec{j},\vec{x}})$ and $P_{i_\ell}^{\vec{j},\vec{x}} = \Lambda(\rho_{i_\ell}^{\vec{j},\vec{x}})$ for $\ell = 1, \ldots, k$. Hence from the above we get

$$\mathbb{E}_{\vec{i}}\mathbb{E}_{\vec{j},\vec{x}}\left\|\Lambda^{\otimes k}(\rho_{\vec{i}}^{\vec{j},\vec{x}}) - \Lambda(\rho_{i_1}^{\vec{j},\vec{x}}) \ldots \otimes \Lambda(\rho_{i_k}^{\vec{j},\vec{x}})\right\|_1^2 \leq \frac{2k^2\ln(d^8)}{t}$$

Using the fact that $\Lambda^{\otimes k}$ is informationally complete with distorsion $(18d)^{k/2}$ yields the quantum theorem.

□

3 Proof of Classical Statement

Now our aim is to prove the classical analogue of the theorem, which we will do using an information theoretic proof. First, we will define various information theoretic quantities (both classical and quantum) to help us in the effort.

Let $\rho_{AB} \in \mathrm{Dens}(\mathbb{C}^d \otimes \mathbb{C}^d)$ be a density matrix, with reduced states ρ_A and ρ_B.

Definition 3. *The entropy of a state ρ_A is defined as*

$$S(\rho_A) = -\mathrm{Tr}(\rho_A \ln \rho_A) = S(\lambda(\rho_A)), \tag{3}$$

where $\lambda(\rho_A)$ are the eigenvalues of ρ_A (these form a distribution since ρ_A is a density matrix).

Note that if $\rho_{AB} = |\psi\rangle\langle\psi|$ is pure then the eigenvalues of ρ_A are the squares of Schmidt coefficients in the decomposition of $|\psi\rangle$ across the $A : B$ partition.

Definition 4. *The mutual information $I(A : B)_\rho$ is defined as*

$$I(A : B)_\rho = S(\rho_A) + S(\rho_B) - S(\rho_{AB}). \tag{4}$$

We will also use conditional mutual information. If ρ_{ABX} is a qqc state, i.e. $\rho_{ABX} = \sum_x p_x \rho_{AB}^x \otimes |x\rangle\langle x|$ for some distribution p_x and density matrices ρ_{AB}^x, then $I(A : B|X)_{\rho_{ABX}} = \sum_x p_x I(A : B)_{\rho_{AB}^x} = S(A|X) + S(B|X) - S(AB|X)$ where $S(A|X) = S(AX) - S(X)$.

Example 5. *If $\rho = \rho_A \otimes \rho_B$, then $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$ and so $I(A : B)_\rho = 0$. In other words, in a product state, neither subsystem carries information about the other.*

Example 6. *On the other hand, if $\rho = |\psi\rangle\langle\psi|$ is pure, then $S(\rho) = 0$ and $I(A : B) = 2S(\rho_A)$. In particular, is $|\psi\rangle = \frac{1}{\sqrt{d}} \sum_i |i\rangle|i\rangle$ is the maximally entangled state, then $I(A : B) = 2\ln d$.*

Claim 7. *In fact these two extreme examples saturate the following inequalities, which always hold:*

$$0 \leq I(A : B) \leq 2\min(\ln d, \ln d') \tag{5}$$

One key property of mutual information is given by Pinsker's inequality.

Claim 8. *For any bipartite state ρ_{AB},*

$$I(A:B)_\rho \geq \frac{1}{2}\|\rho_{AB} - \rho_A \otimes \rho_B\|_1^2. \tag{6}$$

Note that, if $I(A:B) = 0$, then Pinsker's inequalities implies that the state must be a tensor product.

Mutual information has several useful properties.

Proposition 9. *Mutual information satisfies the chain rule*

$$I(A:BX) = I(A:X) + I(A:B|X) \tag{7}$$

and monotonicity under measurements M_A on A,

$$I(A:B)_{M_A \otimes \mathrm{Id}_B(\rho_{AB})} \leq I(A:B)_{\rho_{AB}} \tag{8}$$

as well as tracing out:

$$I(A:B)_{\mathrm{Tr}_C(\rho_{ABC})} \leq I(AC:B)_{\rho_{ABC}}. \tag{9}$$

Monotonicity is intuitive: given a system A, we cannot increase the information that A contains about B by performing an operation on A alone. In the classical case it is easy to prove, but in the quantum case it requires much more work! We will take these properties for granted here.

We are ready to prove the classical statement of the de Finetti theorem. For simplicity we will do the proof for $k = 2$, but the proof for general k is virtually the same. We show the following:

Claim 10. *Under the same notation as Theorem 2,*

$$\mathbb{E}_{0\leq m<t}\mathbb{E}_{\vec{i}=(i_1,i_2),\vec{j}=(j_1\dots j_m)} I(X_1 : X_2|X_{j_1}\dots X_{j_m}) \leq \frac{1}{2}\mathbb{E}_i I(X_i : X_{\bar{i}}) \tag{10}$$

where $(X_i \dots X_n)$ are random variables with distribution P and $\bar{i}$ is shorthand for $[n]\backslash\{i\}$.

First let's show that the claim leads to the classical theorem.

Proof of Theorem 2. Since X_i ranges over a set of size d we have $I(X_i : X_{\bar{i}}) \leq 2\ln d$. Applying Pinsker's inequality,

$$\|P^{\vec{j}\vec{x}}_{i_1,i_2} - P^{\vec{j}\vec{x}}_{i_1} \otimes P^{\vec{j}\vec{x}}_{i_2}\|_1^2 \leq 2I(X_{i_1} : X_{i_1}|X_{j_1}\dots X_{j_m}).$$

So we obtain

$$\mathbb{E}_m\mathbb{E}_{\vec{i}\vec{j}}\mathbb{E}_x\left\|P^{\vec{j}\vec{x}}_{i_1,i_2} - P^{\vec{j}\vec{x}}_{i_1} \otimes P^{\vec{j}\vec{x}}_{i_2}\right\|_1^2 \leq \frac{2\ln d}{t},$$

as desired. □

Now let's prove the claim.

Proof of Claim 10. By monotonicity, and then by the chain rule, we have

$$\begin{aligned}
\frac{1}{t}\mathbb{E}_i I(X_i : X_{\bar{i}}) &\geq \frac{1}{t}\mathbb{E}_{i,j_1,\dots j_t} I(X_i : X_{j_1}\dots X_{j_t}) \\
&= \frac{1}{t}\mathbb{E}_{i,j_1,\dots j_t}\sum_{m=0}^{t-1} I(X_i : X_{j_{m+1}}|X_{j_1}\dots X_{j_m}) \\
&= \mathbb{E}_{0\leq m<t}\mathbb{E}_{\vec{i}=(i_1,i_2),\vec{j}=(j_1\dots j_m)} I(X_1 : X_2|X_{j_1}\dots X_{j_m}),
\end{aligned}$$

where the last line is a simple relabeling of coordinates. □

4 Application to QPCP

Now we apply this statement of the quantum de Finetti theorem to the quantum PCP conjecture to show the following.

Theorem 11. *Let H be a 2-local Hamiltonian on qudits such that the interaction graph has degree exactly D. Then there exists a product state $|\psi\rangle = |\psi_1\rangle \otimes \ldots \otimes |\psi_n\rangle$ such that*

$$\langle\psi|H|\psi\rangle \leq \lambda_0(H) + O\left(\frac{d^2 \ln d}{D}\right)^{1/3} \cdot m, \tag{11}$$

where $m = \dfrac{nD}{2}$ is the number of local terms in H.

This theorem can be seen as a strong "no-go" for QPCP. Indeed, it shows that QPCP is *false* on D-regular graphs for any value of the gap parameter $\gamma = \Omega(d^2 \ln d/D)^{1/3}$ (assuming QMA$\neq$NP). Indeed, the theorem shows that, for these values of γ there is always a product state that has energy less than $b = \lambda_0(H) + \gamma$ and can serve as a *classical* witness for this fact. Since QPCP asserts hardness for constant γ, we see that the conjecture can only be true on graphs of constant degree D (unless the local dimension d is allowed to grow faster than D). In an earlier lecture we also saw (as an exercise) that QPCP required graphs of high dimension in order to be true. Together these two statements narrow down the kind of local Hamiltonians for which the conjecture may be true: their interaction degree should have high dimension (it cannot be embedded on any lattice of dimension $O(\log n)$), but constant degree.

Proof. Let $|\psi\rangle$ be the ground state of H, and the Λ be the IC measurement, defined as before. Let $\rho = |\psi\rangle\langle\psi|$. $\forall \vec{j}, \vec{x}$, define

$$\sigma^{\vec{j},\vec{x}} = \otimes_{i \notin \vec{j}} \rho_i^{\vec{j},\vec{x}} \otimes \frac{1}{d^m} \mathcal{I}_{j_1 \cdots j_m} \tag{12}$$

In other words, this state is $\rho_i^{\vec{j},\vec{x}}$ if $i \notin \vec{j}$ and $\mathcal{I}_d/d$ otherwise. By the quantum de Finetti theorem applied to ρ with $k = 2$ and t to be determined later, we have

$$\mathbb{E}_{i_1,i_2} \mathbb{E}_{\vec{j},\vec{x}} \|\rho_{\vec{i}}^{\vec{j},\vec{x}} - \rho_{i_1}^{\vec{j},\vec{x}} \otimes \rho_{i_2}^{\vec{j},\vec{x}}\|_1^2 = O\left(\frac{d^2 \ln d}{t}\right) \tag{13}$$

For the state we defined, we have

$$\mathbb{E}_{\vec{i},\vec{j},\vec{x}} \|\rho_{\vec{i}}^{\vec{j},\vec{x}} - \sigma_{i_1}^{\vec{j},\vec{x}} \otimes \sigma_{i_2}^{\vec{j},\vec{x}}\|_1^2 \leq O\left(\frac{d^2 \ln d}{t}\right) + \frac{2t}{n} \tag{14}$$

where the last quantity comes from the rare cases when $i \in \vec{j}$. Restricting the expectation on $\vec{i} = (i_1, i_2)$ to the $Dn/2$ pairs that correspond to edges of the constraint graph of the local Hamiltonian and applying Jensen's inequality we obtain

$$\mathbb{E}_{\vec{i}=(i_1,i_2)\in E} \mathbb{E}_{\vec{j},\vec{x}} \|\rho_{\vec{i}}^{\vec{j},\vec{x}} - \sigma_{i_1}^{\vec{j},\vec{x}} \otimes \sigma_{i_2}^{\vec{j},\vec{x}}\|_1 = O\left(\sqrt{\left(\frac{d^2 \ln d}{t} + \frac{t}{n}\right)\frac{n}{D}}\right) \tag{15}$$

By definition of the trace norm and using that for any $\vec{j}$ the reduced density of $\mathbb{E}_{\vec{x}}\rho^{\vec{j},\vec{x}}$ on the $n-t$ qudits that were not measured is the same as that of ρ, we obtain

$$\mathbb{E}_{\vec{j}}\left|\langle\psi|H|\psi\rangle - \langle\psi|\mathbb{E}_{\vec{x}}\sigma^{\vec{j},\vec{x}}|\psi\rangle\right| \leq O\left(\sqrt{\left(\frac{d^2\ln d}{t}+\frac{t}{n}\right)\frac{n}{D}}\right) nD + Dt, \tag{16}$$

where here the multiplicative nD comes from switching the expectation over edges $\vec{i}$ to a sum over edges (since H is given as a sum of local terms), and the additional Dt corresponds to imposing a maximum energy penalty of 1 for each edge involving one of the t measured qubits. Choosing the t that minimizes the right-hand side, $t \sim nd^{2/3}\ln d^{1/3}/D^{1/3}$, we obtain the desired result: there will exist at least one choice of $\vec{j}$ such that the separable state $\mathbb{E}_{\vec{x}}\sigma^{\vec{j},\vec{x}}$ has low energy; using linearity of the Hamiltonian there is at least one pure product state in the support of $\mathbb{E}_{\vec{x}}\sigma^{\vec{j},\vec{x}}$ that has at most the same energy. $\square$