Quantum χ^2 goodness of fit testing

K. Temme and Frank Verstraete

Massachusetts Institute of Technology
Vienna Center for Quantum Science and Technology

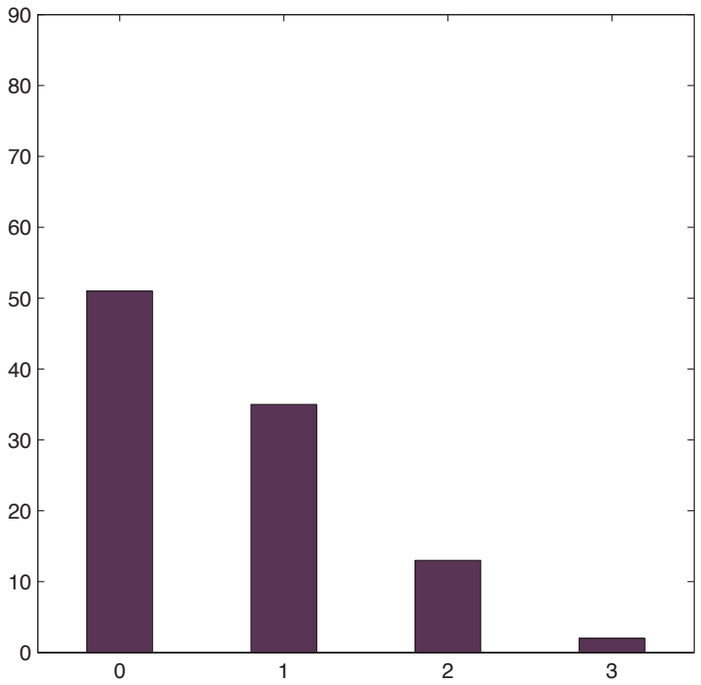
ArXiv: 1112.6343

Nottingham June 26th

A guy walks into a casino ...

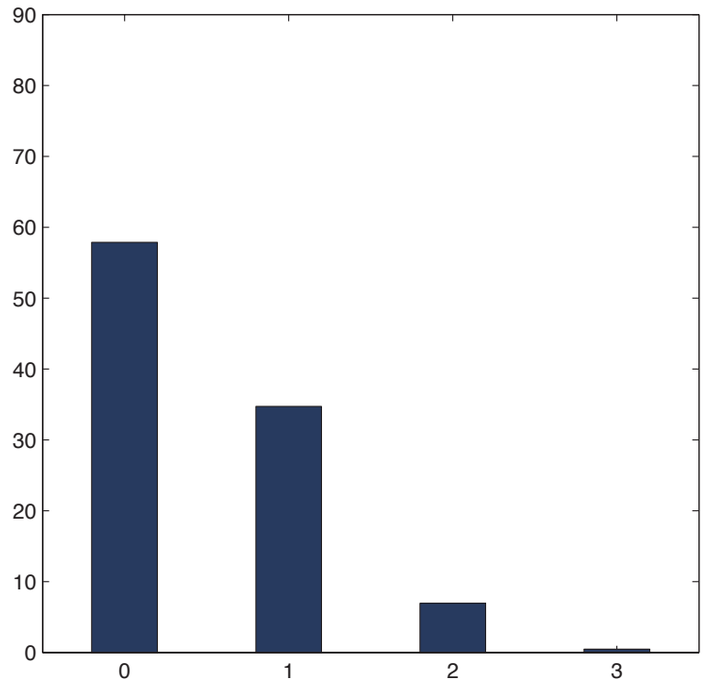
100 times	0	1	2	3
observed	51	35	13	2
expected	58	34.5	7	0.5

observed



$$c^2 = 12.7$$

expected



$$\mathbb{E}(c^2) = 3$$

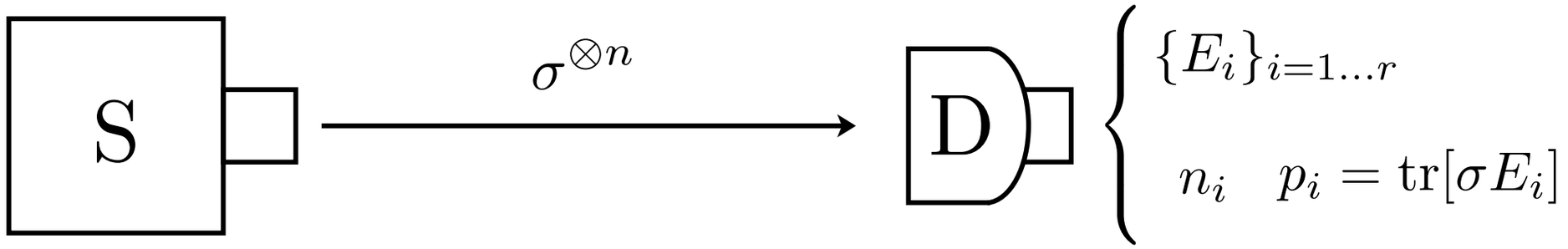
With **error**:

$$\alpha = 0.01 \implies \chi^2_\alpha = 11.35$$

reject:

$$c^2 \geq \chi^2_\alpha$$

Set-up



- An experimentalist claims he can **prepare the state** σ (null Hypothesis).
- We only allow for **single copy measurements** (no entangled MS). We have to choose a POVM with **r elements**.
- The **only** thing we have **access to** are the **frequencies** n_i which are distributed according to $p_i = \mathrm{tr}\left[\sigma E_i\right]$.
- We want to have **confidence** when we choose to **reject** the Hypothesis

Hypothesis testing

- Hypothesis testing:
  - We want to test a hypothesis in the setting where you can repeat the measurement many times
  - The central task is to derive “optimal” tests.
  - Goodness of fit testing is different from state discrimination (Nyman - Pearson).
    We only have one hypothesis (H_0), we assume we have some given state.

- The goodness of test
  - Given the observations we need to test whether our hypothesis is compatible with the Hypothesis otherwise we need to reject
  - We need to define what “compatible” means. (confidence intervals)

The χ^2 goodness of fit test

We only have access to the frequencies n_i

Individual outcomes are distributed according to: $p_i = \mathrm{tr}\left[\sigma E_i\right]$

Multinomial**:**

$$P(n_1, \ldots, n_r) = \frac{n!}{n_1! \ldots n_r!} p_1^{n_1} \ldots p_r^{n_r} \approx_{n\to\infty} \exp\left(-\frac{1}{2}\sum_i \frac{(n_i - np_i)^2}{np_i}\right)$$

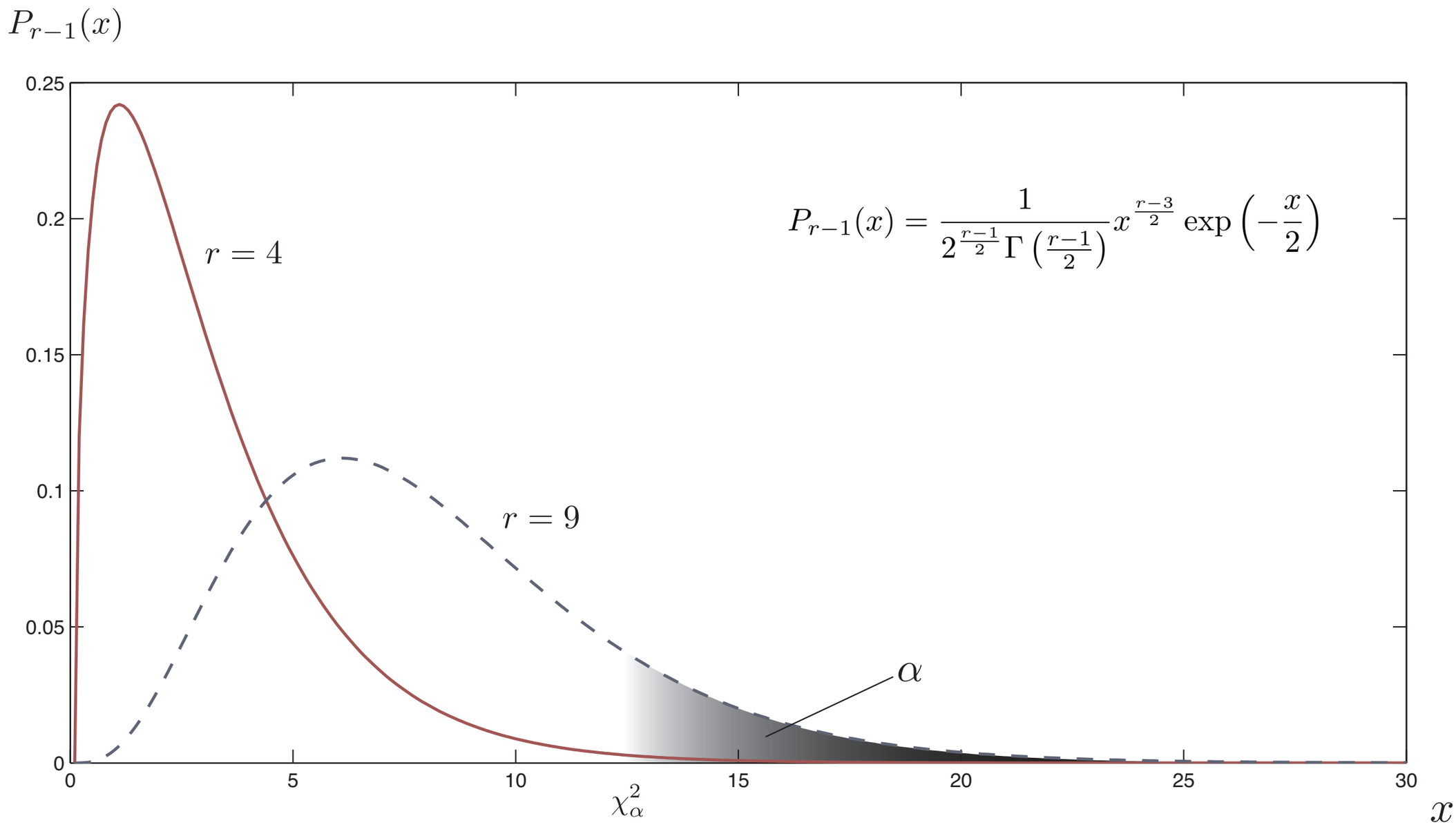
Test statistic:

(Random variable) $$c^2 = \sum_{i=1}^{r} \frac{(n_i - p_i n)^2}{np_i}$$

Crucial properties: $\mathbb{E}\left[c^2\right]_p = r - 1$ n and p independent

$$\mathbb{E}\left[(c^2 - \mathbb{E}[c^2])^2\right]_p \to_{n\to\infty} 2(r-1)$$

The χ^2 distribution



a) **reject** Hypothesis if: $c^2 \geq \chi^2_\alpha$ with error probability α

b) Hypothesis is **consistent** if: $c^2 < \chi^2_\alpha$

Assume: The hypothesis is false

Suppose we don't sample from $\rho \neq \sigma$

For some $\|\rho - \sigma\| \geq \epsilon$

We now compute the average with respect to $q_i = \mathrm{tr}\left[E_i \rho\right]$

$$\begin{aligned}\mathbb{E}[c^2]_q &= \sum_i \frac{n^2 q_i^2 + n q_i (1 - q_i)}{n p_i} - n \\ &= (n-1)\chi^2(p,q) + \mathrm{const}\end{aligned}$$

grows linearly in n : with prefactor

χ^2-divergence $\qquad \chi^2(p,q) = \sum_i \frac{(q_i - p_i)^2}{p_i}$

The questions are:

- We want a test for quantum systems that allows to reject the null Hypothesis with confidence.

- We only have access to the classical frequencies (“clicks”).

- What is the “optimal” strategy to measure?

- What does “optimal” mean?

Quantum χ^2 testing

- Definition: the **divergence rate** is given by

$$\xi(\sigma) = \frac{1}{\epsilon^2} \max_{\{E_i\}} \min_{\|\rho-\sigma\|_F \geq \epsilon} \chi^2(p,q)$$

$$\chi^2(p,q) = \sum_i \frac{(q_i - p_i)^2}{p_i} \qquad q_i = \mathrm{tr}\left[E_i \rho\right] \qquad p_i = \mathrm{tr}\left[\sigma E_i\right]$$

- Motivation: The POVM that maximizes this expression will ensure that the expectation value of the test statistic diverges as fast as possible even for the worst ρ

$$(r-1) + (n-1)\xi(\sigma) \approx \chi^2_\alpha$$

on average: $$n = \mathcal{O}\left(\frac{D}{\epsilon^2 \xi(\sigma)}\right)$$

Optimization Problem

- Constrained EV problem

$$\xi(\sigma) = \max_{\{E_i\}} \min_{\|X\|=1} (X| \underbrace{\left(\sum_i |E_i) \frac{1}{(E_i|\sigma)} (E_i| \right)}_{M} |X)$$

Maximize the smallest eigenvalue of the matrix M , on the space of hermitian traceless $X = \frac{1}{\epsilon^2}(\rho - \sigma)$

- Optimal POVM

1) The optimal POVM has to be **informationally complete**

2) The POVM elements are **pure** $E_i = p_i |\psi_i\rangle\langle\psi_i|$

Main Result

- Theorem: Let $\sigma = \sum_{\alpha} \lambda_\alpha |\alpha\rangle\langle\alpha|$

1) then the **divergence rate** is equal to: $\xi(\sigma) = \dfrac{1}{1+\mu_1(S)}$

where μ_1 is the largest eigenvalue of: $S = \sum_{\alpha} \lambda_\alpha |\alpha\rangle\langle\alpha| - \sum_{\alpha\beta} \lambda_\alpha \lambda_\beta |\alpha\rangle\langle\beta|$

2) The **optimal POVM** is given by:

$$1 \leq i \leq D \qquad E_i = (1-\xi)|i\rangle\langle i|$$

$$j \geq D \qquad E_j \propto \xi |\chi_j\rangle\langle\chi_j|$$

with: $|\chi_j\rangle = \dfrac{1}{\sqrt{D}} \sum_{\alpha} e^{i\theta_j^\alpha} |\alpha\rangle$

Comments

1) The divergence rate is bounded

$$\frac{2}{3} \leq \xi(\sigma) \leq 1$$

2) The optimal POVM

i) The optimal POVM is covariant!

ii) The optimal POVM is **not unique**!

iii) SIC-POVM optimal for maximally mixed states

3) Proof sketch:

i) First proof an **upper bound**

ii) Then proof **lower bound**

iii) Show that they **match**

Proof: The upper bound

Simple symmetry: $$tr[E_i\sigma] = tr[E_i D(-\theta)\sigma D(\theta)]$$

$$D(\theta) = \sum_\alpha e^{i\theta_\alpha}|\alpha\rangle\langle\alpha|$$

Twirling with respect to θ does not decrease the smallest eigenvalue of M

$$\xi \leq \max_{E_i}\min_X \int d\theta (X|D(\theta)\otimes D(-\theta)\left(\sum_i |E_i)\frac{1}{(E_i|\sigma)}(E_i|\right) D(\theta)\otimes D(-\theta)|X)$$

$$\xi \leq \max_{e^i}\min_X \sum_i p_i \frac{\sum_{\alpha\beta}\langle e^i|\alpha\rangle\langle\beta|e^i\rangle(|X_{\alpha\beta}|^2(1-\delta_{\alpha\beta})+X_{\alpha\alpha}X_{\beta\beta})}{\langle e^i|\lambda\rangle}$$

Proof: The upper bound

Biggest Problem the **denominator** $\langle e^i|\lambda\rangle$. However, one can get rid of it if one chooses an X so that

$$|X_{\alpha\beta}|^2(1-\delta_{\alpha\beta}) + X_{\alpha\alpha}X_{\beta\beta} = \lambda_\alpha t_\beta^2 + \lambda_\beta t_\alpha^2$$

Introduce **auxiliary variables t**, so that

$$|X_{\alpha\beta}|^2 = \left(\sqrt{\lambda_\alpha}t_\beta - \sqrt{\lambda_\beta}t_\alpha\right)^2 \qquad X_{\alpha\alpha} = \sqrt{2\lambda_\alpha}t_\alpha$$

One can find an upper bound as:

$$\xi \leq 2\sum_i p_i \langle e^i|t^2\rangle \frac{\langle\lambda|e^i\rangle}{\langle\lambda|e^i\rangle} = 2\sum_\alpha |t_\alpha|^2$$

Proof: The upper bound

Together with the constraints this leads to the optimization problem:

$$\text{minimize} \quad \|t\|_2^2$$

$$\text{subject to:} \quad \sum_\alpha t_\alpha \sqrt{\lambda_\alpha} = 0$$

$$\sum_\alpha (1+\lambda_\alpha) t_\alpha^2 = 1$$

The solution is given by the second smallest eigenvalue of

$$P_\lambda \left(\sum_\alpha \frac{1}{1+\lambda_\alpha} |\alpha\rangle\langle\alpha| \right) P_\lambda$$

where: P_λ is the projectors onto the complement of: $\sum_\alpha \frac{\sqrt{\lambda_\alpha}}{1+\lambda_\alpha} |\alpha\rangle$

Proof: The lower bound

A lower bound can be found by choosing a 1 - parameter POVM :

$$\begin{aligned} 1 \leq i \leq D: \quad E_i &= (1-p)|i\rangle\langle i| \\ j > D: \quad E_j &= p\, c|\chi_j\rangle\langle\chi_j| \end{aligned}$$

With this POVM the optimization turns into:

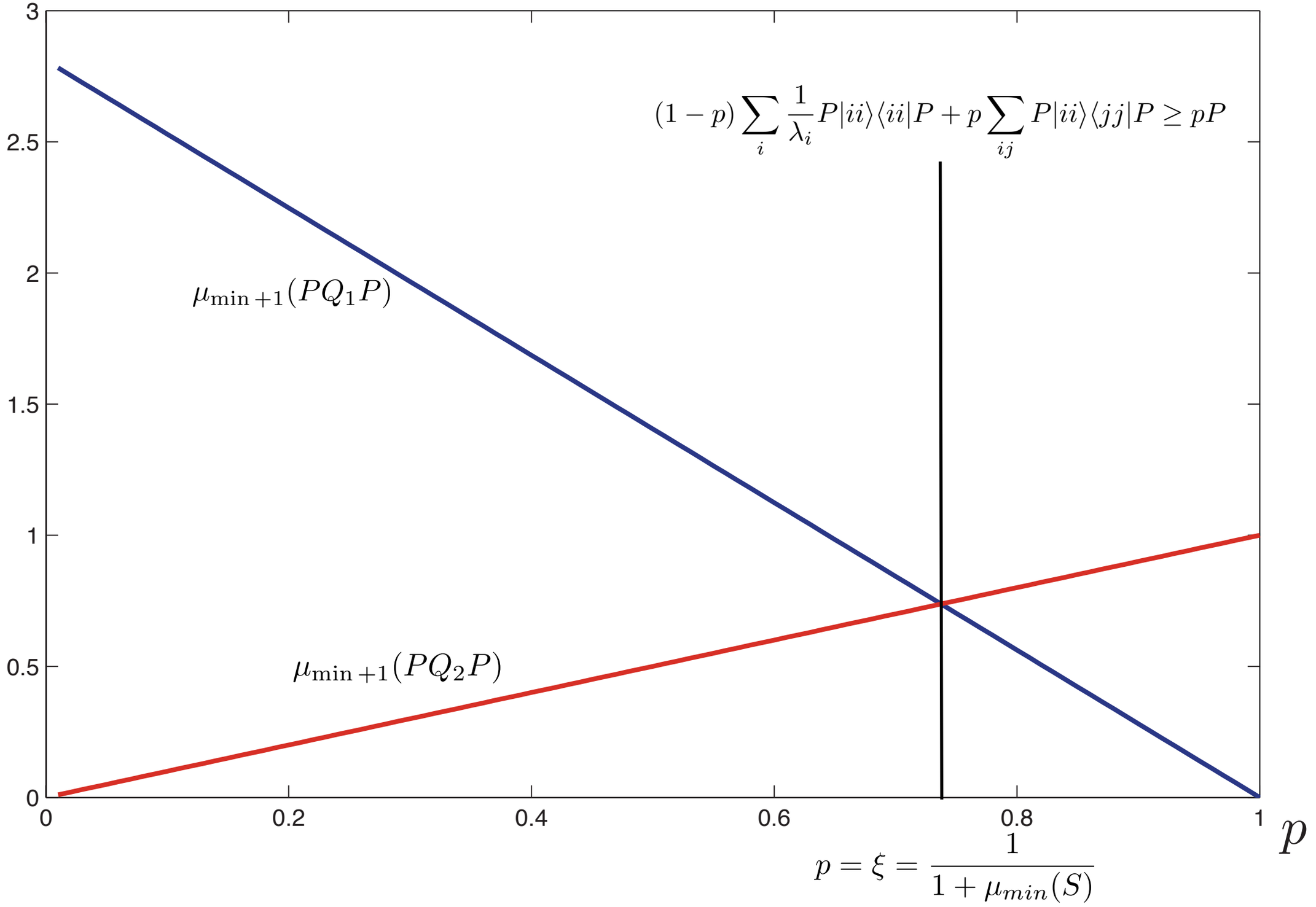
$$\xi \geq \max_p \min_X \langle X|Q_1 \oplus Q_2|X\rangle$$

where:

$$Q_2 = p\sum_{i\neq j}^{D} |ij\rangle\langle ij| = p\mathbb{I}_2$$

$$Q_1 = (1-p)\sum_{i=1}^{D} \frac{1}{\lambda_i}|ii\rangle\langle ii| + p\sum_{i,j}^{D} |ii\rangle\langle jj|$$

Proof: The lower bound



Proof: upper and lower bound

1) This characterizes the divergence rate $\xi(\sigma) = \dfrac{1}{1+\mu_1(S)}$ and $S = \sum_\alpha \lambda_\alpha |\alpha\rangle\langle\alpha| - \sum_{\alpha\beta} \lambda_\alpha \lambda_\beta |\alpha\rangle\langle\beta|$

and fixes the POVM with $\xi(\sigma) = p$

2) Recall: the **upper bound** was given by second smallest EV of

$$R = P_\lambda \left(\sum_\alpha \frac{1}{1+\lambda_\alpha} |\alpha\rangle\langle\alpha| \right) P_\lambda$$

3) Furthermore: Note that S is the pseudoinverse of $P\sigma P$ where, $P = \mathbb{I} - |\hat{1}\rangle\langle\hat{1}|$

We have that the pseudoinverse of $P + S$ is similar to R

with similarity transformation $D = \mathrm{diag}(\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_N})$

Hence, upper and lower bound match

The protocol

A goodness of fit test protocol for the state σ is then given as follows:

1. Choose the POVM r element $\{E_i^*\}$ that optimizes ξ as given in definition
2. Measure $\{E_i^*\}$ on n independent samples of the state ρ and record the frequencies n_i of the i'th outcome.
3. Compute the test statistic $c^2 = \sum_{i=1}^{r} \frac{(n_i - p_i n)^2}{p_i n}$, where $p_i = \mathrm{Tr}[E_i \sigma]$ corresponds to the hypothesis H.
4. Reject the Hypothesis with error probability α if $c^2 \geq \chi_\alpha^2$, where the constant χ_α^2 is determined via

$$\alpha = \int_{\chi_\alpha^2}^{\infty} P_{r-1}(x) dx$$

5. If the test statistic c^2 is smaller than χ_α^2, we state that the observed data is consistent with the hypothesis H up to a statistical error α.

Example systems

1) For pure states: $\xi(|\psi\rangle\langle\psi|) \to 1$

2) For the maximally mixed state: $\xi(\mathbb{I}/D) = \dfrac{D}{1+D}$

3) For qubits we have: $\xi(\sigma) = \dfrac{1}{1+2\lambda_1\lambda_2}$

The optimal POVM is given by:

$$(1-\xi)|0\rangle\langle 0|, (1-\xi)|1\rangle\langle 1|,$$
$$\frac{\xi}{2}|+\rangle\langle +|, \frac{\xi}{2}|-\rangle\langle -|,$$
$$\frac{\xi}{2}|+i\rangle\langle +i|, \frac{\xi}{2}|-i\rangle\langle -i|$$

Conclusions and discussion

- Goodness of fit testing provides a complementary approach to tomography.
  - gives a statistical definition of error (confidence intervals)

- The chi-squared test is one of the most widely used test in statistics the quantum version involves an extra optimization over the POVM

- The optimal POVM can be found through a max-min optimization. The solution shows that one can always find an efficient test

- We only focused on what is actually measurable! The “clicks” (the frequencies) are what is actually relevant