Quantum Foundations Lecture 18

April 11, 2018

Dr. Matthew Leifer
leifer@chapman.edu
HSC112

Announcements

- Adam Becker is returning to Chapman:
  - Book event and signing at 1888 center: Monday April 16. RSVP required https://bit.ly/AdamBecker
  - Extra Credit will be added to Hwk 3!
- Assignments
  - First Draft due on Blackboard April 11.
  - Peer review until April 16.
  - Discussion in class April 16.
  - Final Version due May 2.
- Homework 3 due April 11.

9.v) Contextuality

- We follow an approach to contextuality that is due to Rob Spekkens – Phys. Rev. A 71, 052108 (2005).
- The basic philosophy is based on Leibniz Principle of the Identity of Indiscernables:
  - No two distinct things exactly resemble each other.
- This principle is arguably very successful in physics:
  - e.g. Principle of relativity, Einstein's equivalence principle.
- The principle can also be thought of as a no fine tuning argument.
  - e.g. suppose objects A and B have some distinct physical property, but there is absolutely no measurement we can do to tell A and B apart. Then, our measurements must only reveal coarse-grained information that is fine-tuned in just such a way so as not to reveal the difference.
- Not all apparent fine tunings are evil, but they do require explanation.

Preparation Contextuality

- Define an equivalence relation on preparations in an operational theory:

$$P \sim Q \quad \Leftrightarrow \quad \mathrm{Prob}(k|P,M) = \mathrm{Prob}(k|Q,M)$$ for all measurement-outcome pairs (M,k).

- In particular, if $\rho_P = \rho_Q$ then $P \sim Q$.
- An ontological model is preparation noncontextual if,

$$P \sim Q \quad \Rightarrow \quad \Pr(\lambda|P) = \Pr(\lambda|Q).$$

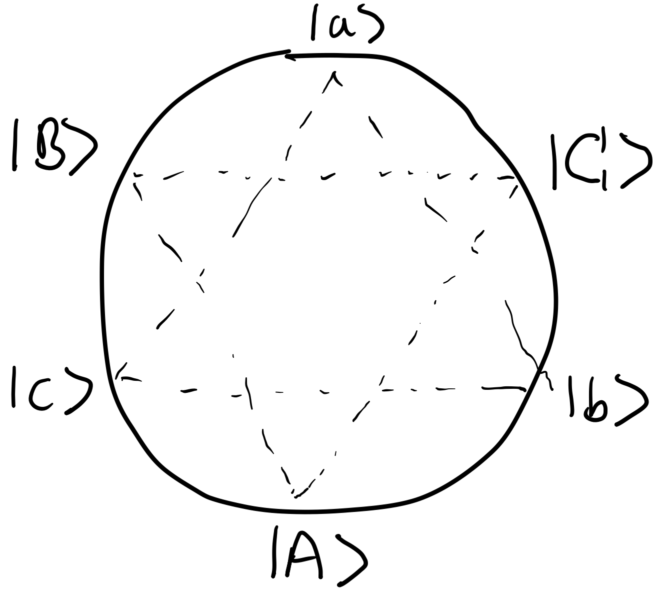
- In words, whenever there is no observable distinction between two preparations, they are represented by the same epistemic state in the ontological model.
- A model that is not preparation noncontextual is called preparation contextual.

Mixing Preparations

- If an operational theory contains preparations P and Q then we can construct a mixed preparation $pP + (1-p)Q$.
  - Physically this means, toss a coin with $p(\text{heads}) = p$, do P if it lands heads or Q if it lands tails, then forget the coin toss outcome.
- We will assume that the ontological model preserves mixtures:
$$\Pr(\lambda|pP + (1-p)Q) = p\Pr(\lambda|P) + (1-p)\Pr(\lambda|Q)$$
- This is actually an instance of preparation noncontextuality applied to the joint coin-system system. Conditioning on the outcome of the coin yields a preparation equivalent to P or Q.

Proof of Preparation Contextuality

○ Consider the following 6 states on the equator of the Bloch sphere



We have $\langle a|A\rangle = \langle b|B\rangle = \langle c|C\rangle = 0$

This implies $\Lambda_a \cap \Lambda_A = \Lambda_b \cap \Lambda_B = \Lambda_c \cap \Lambda_C = \phi$

where $\Lambda_\psi = \{\lambda \mid Pr(\lambda|\psi) > 0\}$

Why?

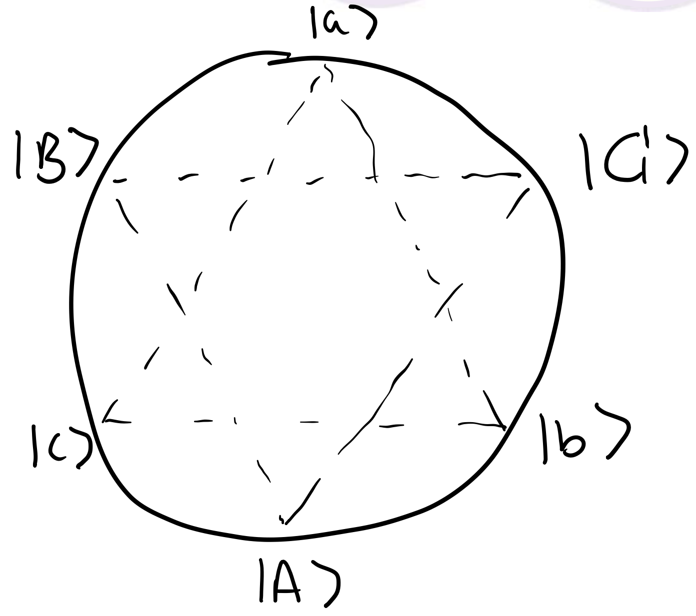
By lemma $\Lambda_\psi \subseteq \Gamma_\psi^M$ where $\Gamma_\psi^M = \{\lambda \mid Pr(\psi|M,\lambda) = 1\}$

But $Pr(a|M,\lambda) + Pr(A|M,\lambda) = 1$ for all $\lambda \in \Lambda$

so $Pr(a|M,\lambda) = 1 \Rightarrow Pr(A|M,\lambda) = 0$ and vice versa

$\therefore \Gamma_a^M \cap \Gamma_A^M = \phi \Rightarrow \Lambda_a \cap \Lambda_A = \phi$

Proof of Preparation Contextuality



We also have:

$$
\begin{aligned}
\frac{I}{2} &= \frac{1}{2}\left(|a\rangle\langle a| + |A\rangle\langle A|\right) \\
&= \frac{1}{2}\left(|b\rangle\langle b| + |B\rangle\langle B|\right) \\
&= \frac{1}{2}\left(|c\rangle\langle c| + |C'\rangle\langle C'|\right) \\
&= \frac{1}{3}\left(|a\rangle\langle a| + |b\rangle\langle b| + |c\rangle\langle c|\right) \\
&= \frac{1}{3}\left(|A\rangle\langle A| + |B\rangle\langle B| + |C'\rangle\langle C'|\right)
\end{aligned}
$$

So by preparation noncontextuality:

$$
\begin{aligned}
\Pr(\lambda|\tfrac{I}{2}) &= \frac{1}{2}\left(\Pr(\lambda|a) + \Pr(\lambda|A)\right) \\
&= \frac{1}{2}\left(\Pr(\lambda|b) + \Pr(\lambda|B)\right) \\
&= \frac{1}{2}\left(\Pr(\lambda|c) + \Pr(\lambda|C)\right)
\end{aligned}
$$

$$
\begin{aligned}
\Pr(\lambda|\tfrac{I}{2}) &= \frac{1}{3}\left(\Pr(\lambda|a) + \Pr(\lambda|b) + \Pr(\lambda|c)\right) \\
&= \frac{1}{3}\left(\Pr(\lambda|A) + \Pr(\lambda|B) + \Pr(\lambda|C')\right)
\end{aligned}
$$

Proof of Preparation Contextuality

$$\Pr(\lambda|\tfrac{I}{2}) = \tfrac{1}{2}\left(\Pr(\lambda|a) + \Pr(\lambda|A)\right)$$
$$= \tfrac{1}{2}\left(\Pr(\lambda|b) + \Pr(\lambda|B)\right)$$
$$= \tfrac{1}{2}\left(\Pr(\lambda|c) + \Pr(\lambda|C)\right)$$

$$\Pr(\lambda|\tfrac{I}{2}) = \tfrac{1}{3}\left(\Pr(\lambda|a) + \Pr(\lambda|b) + \Pr(\lambda|c)\right)$$
$$= \tfrac{1}{3}\left(\Pr(\lambda|A) + \Pr(\lambda|B) + \Pr(\lambda|C)\right)$$

Now, any given λ can only be in at most one of Λ_a or Λ_A, Λ_b or Λ_B, Λ_c or Λ_C.

Let's choose a λ that is not in Λ_a, not in Λ_b, and not in Λ_C. Then

$$\Pr(\lambda|\tfrac{I}{2}) = \tfrac{1}{2}\Pr(\lambda|c) \qquad \Pr(\lambda|\tfrac{I}{2}) = \tfrac{1}{3}\Pr(\lambda|c)$$

$\Rightarrow 2\Pr(\lambda|\tfrac{I}{2}) = 3\Pr(\lambda|\tfrac{I}{2}) \Rightarrow \Pr(\lambda|\tfrac{I}{2}) = 0$ for this particular λ

We get a similar result for every choice of not in $\Lambda_{a/A}$, $\Lambda_{b/B}$, $\Lambda_{c/C}$

This exhausts $\Lambda \Rightarrow \Pr(\lambda|\tfrac{I}{2}) = 0$ everywhere, but this cannot be true for a probability distribution.

Measurement Contextuality

- Define an equivalence relation on measurement-outcome pairs in an operational theory:

  $(\mathrm{M},\mathrm{k})\sim(N,l) \quad \Leftrightarrow \quad \mathrm{Prob}(k|P,M)=\mathrm{Prob}(l|P,N)$ for all preparations P.
- In particular, if $E_k^M = E_l^N$ then $(\mathrm{M},\mathrm{k})\sim(M,l)$.
- An ontological model is measurement noncontextual if,

$$(M,k)\sim(N,l) \quad \Rightarrow \quad \mathrm{Pr}(k|M,\lambda)=\mathrm{Pr}(l|N,\lambda).$$

- In words, whenever there is no observable distinction between two measurement-outcome pairs, they are represented by the same response function in the ontological model.
- A model that is not measurement noncontextual is called measurement contextual.

Kochen-Specker Contextuality

- Measurement noncontextual models exist:
  - e.g. Beltrametti-Bugajski: $\mathrm{Pr}(k|M,\lambda) = \mathrm{Tr}(\, E_k^M\ |\lambda\rangle\langle\lambda|)$.
- A Kochen-Specker (KS) noncontextual model is:
  - A model that only contains projective measurements.
  - Measurement noncontextual.
  - Outcome deterministic: $\mathrm{Pr}(\Pi|\lambda) = 0 \text{ or } 1$ for all λ.
- We will prove in a later lecture that:

  $$\text{KS contextual} \quad \Rightarrow \quad \text{maximally } \psi\text{−epistemic} \quad \Longrightarrow \quad \text{preparation contextual}$$

  so KS contextuality is still worth proving.
- KS contextuality can only be proved in $d \geq 3$.
- By applying KS noncontextuality for projective measurements and measurement noncontextuality for POVMs, Spekkens obtained a proof in $d = 2$. We will focus on traditional KS proofs.

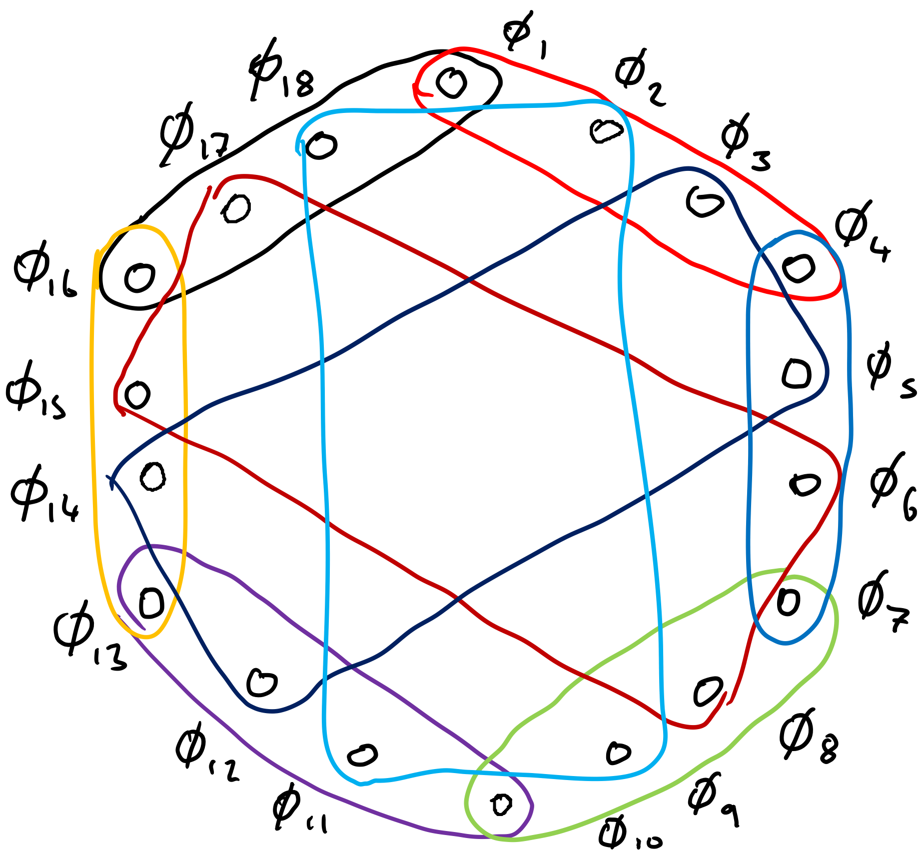
KS Contextuality and value assignments

- Due to the outcome determinism assumption, each λ determines a value function v_λ that assigns a value 0 or 1 to each projector.

$$v_\lambda(\Pi) = \Pr(\Pi|\lambda)$$

- Since probabilities must sum to 1, in each projective measurement $\{\Pi_k\}$, exactly one of the projectors must get value 1, the others getting value 0.
- Measurement noncontextuality then implies that the value assigned to Π_k does not depend on which measurement it is a part of.
- In particular, this applies to an orthonormal basis. For each basis $\{|\phi_k\rangle\}$, exactly one vector gets the value 1, the rest 0, and this value is the same for every basis that $|\phi_k\rangle$ appears in.
- In proving Kochen-Specker contextuality, we can focus on whether such a value function exists.

The 18-Ray Proof



- A. Cabello, J. Estebaranz, G. Garcia-Alcaine, Phys. Lett. A 212:183 (1996).
- In 4-dimensional quantum mechanics, we can find 18 states with the (test space) structure depicted.
- Each test is an orthonormal basis.

ϕ_1	$(1,0,0,0)$	ϕ_{10}	$(0,1,0,-1)$
ϕ_2	$(0,1,0,0)$	ϕ_{11}	$(1,0,1,0)$
ϕ_3	$(0,0,1,1)$	ϕ_{12}	$(1,1,-1,1)$
ϕ_4	$(0,0,1,-1)$	ϕ_{13}	$(-1,1,1,1)$
ϕ_5	$(1,-1,0,0)$	ϕ_{14}	$(1,1,1-1)$
ϕ_6	$(1,1,-1,-1)$	ϕ_{15}	$(1,0,0,1)$
ϕ_7	$(1,1,1,1)$	ϕ_{16}	$(0,1,-1,0)$
ϕ_8	$(1,-1,1,-1)$	ϕ_{17}	$(0,1,1,0)$
ϕ_9	$(1,0,-1,0)$	ϕ_{18}	$(0,0,0,1)$

The 18-Ray Proof

Red	Blue	Green	Purple	Yellow	Black	Light Blue	Navy	Burgundy
ϕ_1	ϕ_4	ϕ_7	ϕ_{10}	ϕ_{13}	ϕ_{16}	ϕ_2	ϕ_3	ϕ_6
ϕ_2	ϕ_5	ϕ_8	ϕ_{11}	ϕ_{14}	ϕ_{17}	ϕ_9	ϕ_5	ϕ_8
ϕ_3	ϕ_6	ϕ_9	ϕ_{12}	ϕ_{15}	ϕ_{18}	ϕ_{11}	ϕ_{12}	ϕ_{15}
ϕ_4	ϕ_7	ϕ_{10}	ϕ_{13}	ϕ_{16}	ϕ_1	ϕ_{18}	ϕ_{14}	ϕ_{17}

- There are nine bases, and in each one, one of the $\phi_j's$ has to receive the value 1, the rest 0. So there will be 9 rays assigned the value 1 in total.
- However, each ϕ_j appears exactly two times in the table, so whichever of them are assigned the value 1, there will always be an even number of 1's in total. Contradiction!

KS Contextuality and value assignments

- We can also think of the value functions as assigning definite values to observables (self-adjoint operators) via

$$v(M) = \sum_j m_j \, v(\Pi_j)$$

- Now, if two observables M and N commute then they have a joint eigendecomposition.

$$M = \sum_j m_j \Pi_j \qquad N = \sum_j n_j \, \Pi_j$$

- And we will have:

$$MN = \sum_j m_j n_j \, \Pi_j \qquad M + N = \sum_j (m_j + n_j) \Pi_j$$

KS Contextuality and value assignments

- Since, in all of these decompositions, the same projector will get the value 1, whenever $[M, N] = 0$, the value functions will obey

$$v(MN) = v(M)v(N) \qquad\qquad v(M+N) = v(M) + v(N)$$

- If we define functions of operators by power series, this implies that whenever $M_1, M_2, \dots$ all mutually commute then

$$v\big(f(M_1, M_2, \dots)\big) = f(v(M_1), v(M_2), \dots)$$

- So another way of defining KS noncontextuality is: there exists a value function that assigns eigenvalues to observables that obeys $v\big(f(M_1, M_2, \dots)\big) = f(v(M_1), v(M_2), \dots)$ for mutually commuting observables.

The Peres-Mermin Square

- Consider the following table of 9 two qubit observables:

$\sigma_1 \otimes \sigma_1$	$\sigma_1 \otimes I$	$I \otimes \sigma_1$	I
$\sigma_3 \otimes \sigma_3$	$I \otimes \sigma_3$	$\sigma_3 \otimes I$	I
$\sigma_2 \otimes \sigma_2$	$\sigma_1 \otimes \sigma_3$	$\sigma_3 \otimes \sigma_1$	$-I$
I	I	I	

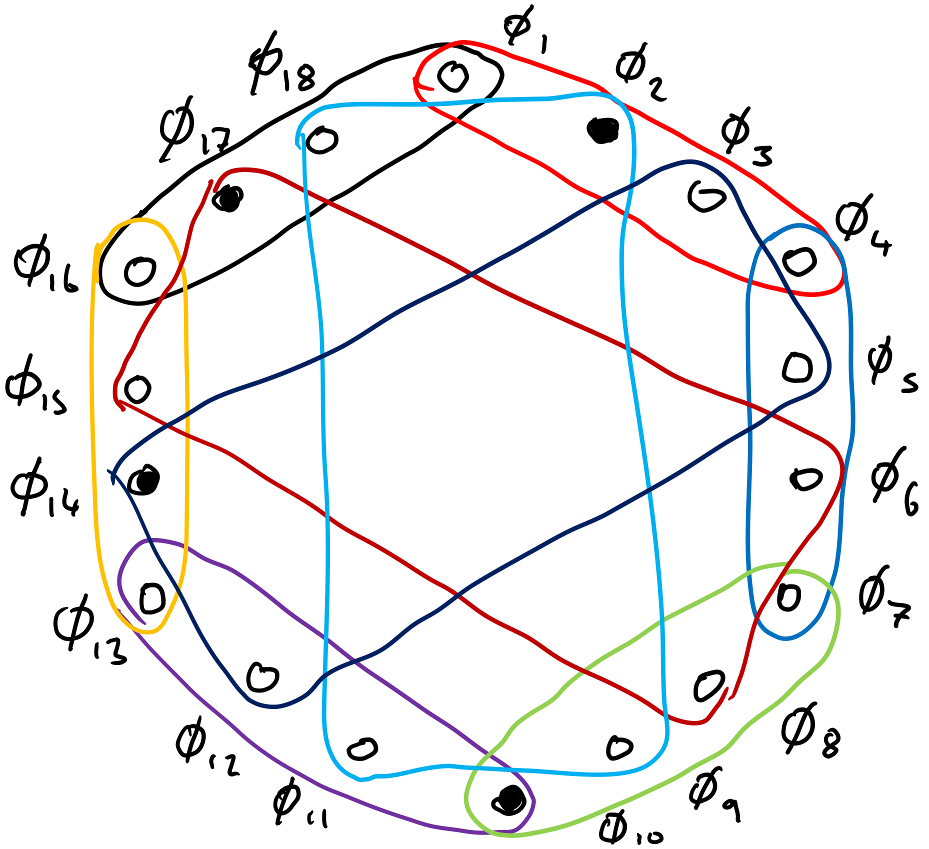
Row products

Column products

- Each observable has eigenvalues ± 1, so receives values ± 1.
- Each row and column consists of mutually commuting observables.
- The column products are all $+I$, which has value $+1$, so there must be an even number of -1's in each column, so an even number in total
- However, one of the row products is $-I$, so there must be an odd number of -1's in that row, and an odd number in total $\Rightarrow$ contradiction.

Noncontextuality Inequalities

- People sometimes want to detect contextuality using inequalities like we do for nonlocality in Bell's theorem.
- Example: 18 ray proof.



- Each grouped set of vertices is a basis, one vector should get value 1, the rest 0.
- But however you try to do this there is always one basis left over that cannot be filled.
- People then say that, in a noncontextual theory

$$\sum_j \Pr(\phi_j|\lambda) \leq 4$$

- However, this is wrong: any theory must predict a valid probability distribution for every measurement $\Rightarrow$ no noncontextual model exists.

Noncontextual Sets

- We can make sense of noncontextuality inequalities in the following way.
- Let $\mathcal{M} = \{M_1, M_2, \ldots, M_n\}$ be a finite set of orthonormal bases.
- If $|\phi\rangle$ is an outcome in $M \in \mathcal{M}$, define

$$\Gamma_\phi^M = \{\lambda \,|\, \Pr(\phi|M,\lambda) = 1\}$$

- Define the noncontextual set for $|\phi\rangle$ as

$$\Gamma_\phi = \bigcap_{\{M \in \mathcal{M} \,|\, |\phi\rangle \in M\}} \Gamma_\phi^M$$

This is the set of ontic states that always assign $|\phi\rangle\langle\phi|$ probability 1 regardless of the basis it appears in

i.e. the set of ontic states that give the outcome $|\phi\rangle$ noncontextually.

Noncontextual Sets

- In a KS noncontextual model, we would have

$$|\langle\phi|\psi\rangle|^2 = \int_\Lambda d\lambda \, Pr(\phi|M,\lambda) Pr(\lambda|\psi) = \int_{\Gamma_\phi} d\lambda \, Pr(\phi|M,\lambda) Pr(\lambda|\psi)$$

$$= \int_{\Gamma_\phi} d\lambda \, Pr(\lambda|\psi) \underset{\text{def}}{\equiv} Pr(\Gamma_\phi|\psi)$$

- This is actually equivalent to KS noncontextuality (up to measure-zero issues)
- In a KS contextual model $Pr(\Gamma_\phi|\psi) \leq |\langle\phi|\psi\rangle|^2$

  but $Pr(\Gamma_\phi|\psi)$ still makes sense.

  It measures the proportion of the probability of obtaining outcome $|\phi\rangle$ that is accounted for by ontic states that are noncontextual for $|\phi\rangle$.

Noncontextuality Inequalities Revisited

- Now if $\langle \phi_1 | \phi_2 \rangle = 0$ then Γ_{ϕ_1} and Γ_{ϕ_2} are disjoint.

  Why? $\exists$ a basis M that includes $|\phi_1\rangle$ and $|\phi_2\rangle$

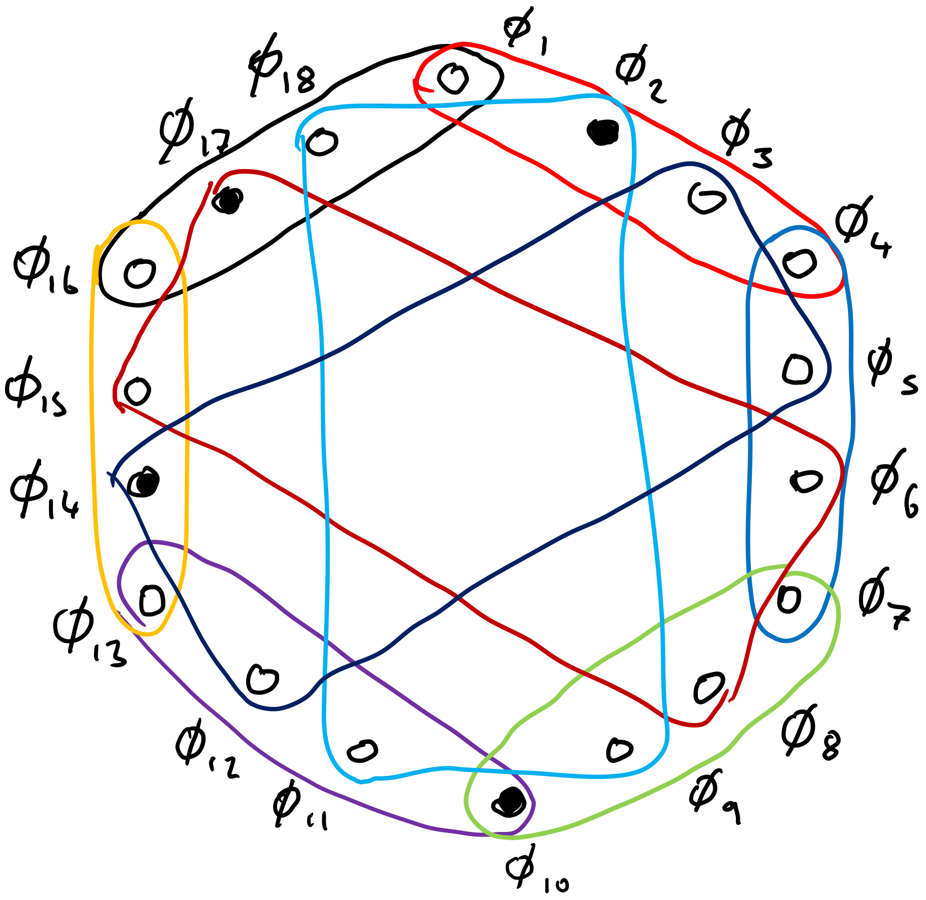
  $\Gamma_{\phi_1}^{M}$ and $\Gamma_{\phi_2}^{M}$ are disjoint because $Pr(\phi_1 | M, \lambda) + Pr(\phi_2 | M, \lambda) \leq 1$

  But $\Gamma_{\phi_1} \subseteq \Gamma_{\phi_1}^{M}$ and $\Gamma_{\phi_2} \subseteq \Gamma_{\phi_2}^{M}$

- If $M = \{|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_d\rangle\}$ is an orthonormal basis then any λ can be in at most 1 of $\Gamma_{\phi_1}, \Gamma_{\phi_2}, \ldots, \Gamma_{\phi_d}$

- But it doesn't have to be in any of them. It could be in a nondeterministic or contextual state instead.

Noncontextuality Inequalities Revisited



We can now make sense of the inequality for the 18-ray proof

$$\sum_{j=1}^{18} \Pr(\Gamma_{\phi_j} | \lambda) \leq 4 \quad \text{for all } \lambda$$

$$\therefore \sum_{j=1}^{18} \int_{\Lambda} d\lambda \, \Pr(\Gamma_{\phi_j} | \lambda) \Pr(\lambda | \psi) \leq 4$$

$$\Rightarrow \sum_{j=1}^{18} \Pr(\Gamma_{\phi_j} | \psi) \leq 4$$

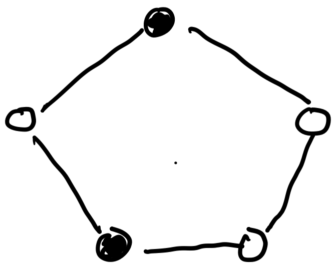
And if $\sum_{j=1}^{18} |\langle \phi_j | \psi \rangle|^2 > 4$ we have detected contextuality.

CSW Noncontextuality Inequalities

- Cabello, Severini and Winter (CSW) introduced a class of noncontextuality inequalities based on graph theory. Phys. Rev. Lett. 112: 040401 (2014).
- Consider a graph $G = (V, E)$

e.g.

$\alpha(G) = 2$



- To each vertex $v \in V$ we assign a pure state $|\phi_v\rangle$
- If the vertices are connected by an edge $(v, v') \in E$ then we demand $\langle \phi_v | \phi_{v'} \rangle = 0$

- The independence number $\alpha(G)$ is the size of the largest set of vertices such that no two vertices are connected by an edge.
- Since Γ_{ϕ_v} and $\Gamma_{\phi_{v'}}$ are disjoint for orthogonal states, a KS noncontextual model satisfies

$$\sum_{v \in V} \Pr(\Gamma_{\phi_v} | \psi) \leq \alpha(G) \quad \text{for any state } |\psi\rangle.$$

CSW Noncontextuality Inequalities

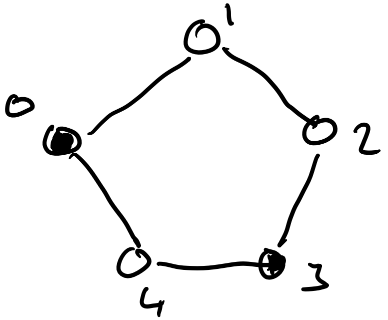
- To determine the maximum possible quantum violation we want to optimize

$$\max_{\{|\phi_v\rangle\}, |\psi\rangle} \sum_{v \in V} |\langle \phi_v | \psi \rangle|^2 = \theta(G)$$ subject to the orthogonality constraints.

- It turns out that $\theta(G)$ had been studied in graph theory for other reasons. It is called the Lovasz theta function. In particular, it can be efficiently computed numerically.
- So finding CSW contextuality proofs is equivalent to finding graphs with $\theta(G) > \alpha(G)$.

Example: Klyatchko Inequality

- A previously known example is the Klyatchko inequality, based on a 5-cycle



$$\alpha(G) = 2 \quad \text{so} \quad \sum_{j=0}^{4} \Pr(\Gamma_{\phi_j} | \psi) \leq 2$$

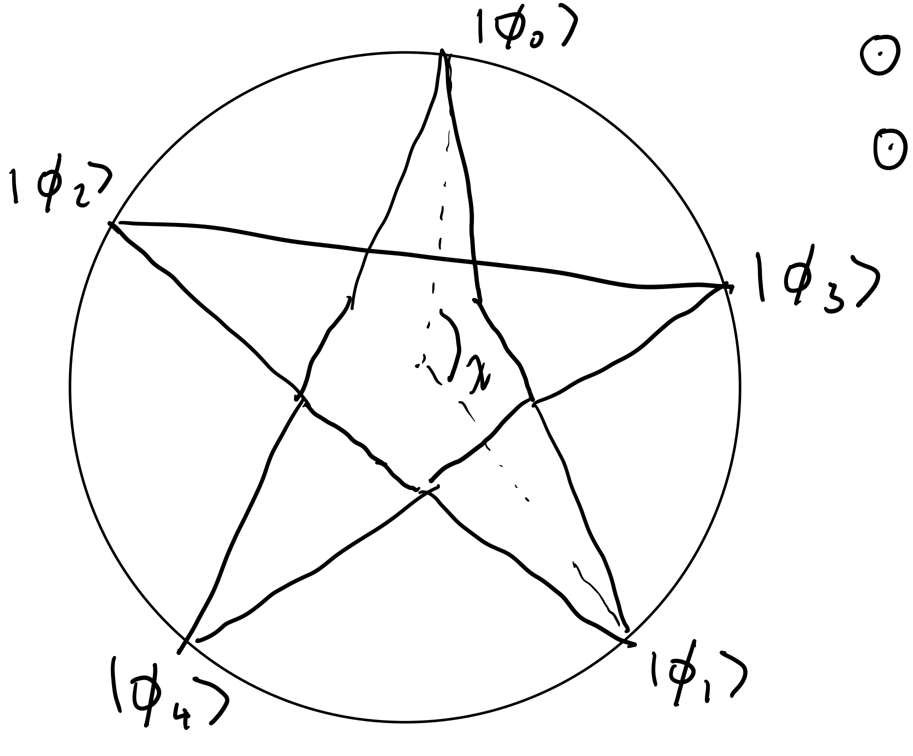
- The maximum quantum violation is found in a 3-d real Hilbert space

$$|\phi_j\rangle = \begin{pmatrix} \sin\chi \cos\eta_j \\ \sin\chi \sin\eta_j \\ \cos\chi \end{pmatrix} \quad |\psi\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{with } \eta_j = \frac{4\pi j}{5} \quad \cos\chi = \frac{1}{\sqrt[4]{5}}$$

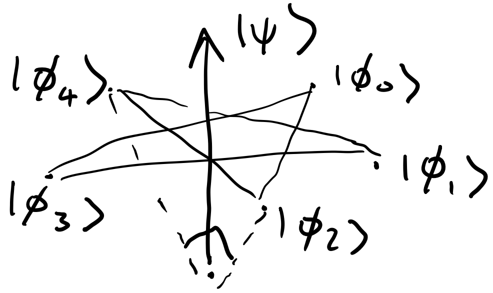
$$\Theta(G) = \max \sum_{j=0}^{4} |\langle \phi_j | \psi \rangle|^2 = 5\cos^2\chi = \frac{5}{\sqrt{5}} = \sqrt{5} \simeq 2.24$$

Example: Klyatchko Inequality

- Geometrically we can understand the states $|\phi_j\rangle$ as follows
- Consider the following 5 states on the equator of the unit sphere



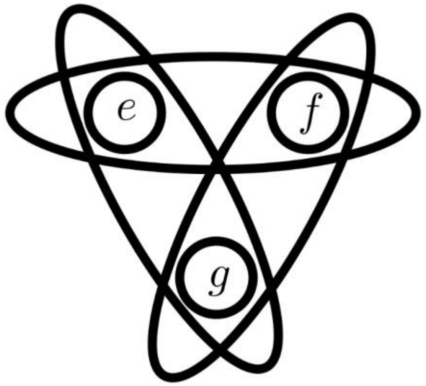
- The angle $\chi > 90^\circ$ so these states are not orthogonal.
- However if we raise the pentagram up the surface of the sphere, keeping it parallel to the equator then eventually the angle will hit 90°



KS Contextuality in Test Spaces

- The 18-ray proof is based on a test space. We can generalize this approach to arbitrary test spaces.
- Recall that a *finite test space* (X, Σ) consists of
  - A finite set X of *outcomes*.
  - A finite set Σ of *tests*.
  - Each test E is a finite subset of X, interpreted as the set of outcomes for a measurement that can be performed on the system.
- Example: Specker's Triangle

$$(\{e, f, g\}, \{\{e, f\}, \{f, g\}, \{g, e\}\})$$



KS Contextuality in Test Spaces

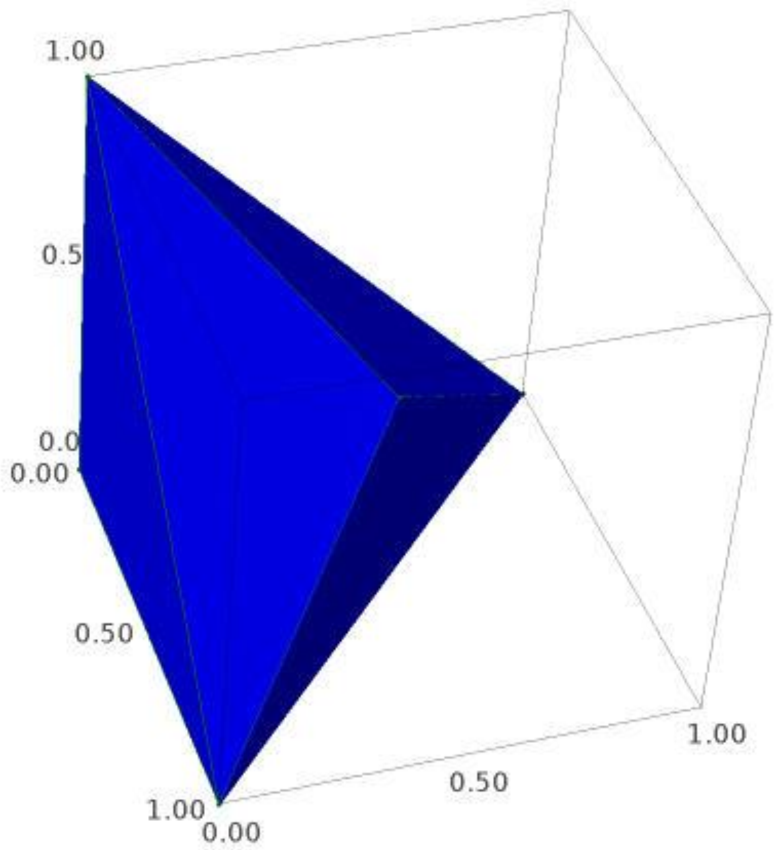
- A *state* on a test space is a function $\omega: X \to [0,1]$ such that
$$\forall E \in \Sigma, \qquad \sum_{e \in \mathrm{E}} \omega(e) = 1$$
- Let $\mathcal{S}(X,\Sigma)$ be the set of states on (X,E). For a finite test space this is a polytope.
- An *unnormalized state* on a test space is a function $\omega: X \to [0,1]$ such that
$$\forall E \in \Sigma, \qquad \sum_{e \in \mathrm{E}} \omega(e) \leq 1$$
- Let $\mathcal{S}_u(X,\Sigma)$ be the set of unnormalized states on (X,Σ). For a finite test space this is also a polytope.
- The advantage is that not all test spaces have states, but they do all have unnormalized states.
- Interpretation. We let our measurements sometimes fail, and not register an outcome. The probability of this happening can depend on which test we are measuring.

Example: Specker Triangle

- We proved previously that the only normalized state on a Specker triangle is
$$\omega(e) = \omega(f) = \omega(g) = \frac{1}{2}$$
- Unnormalized states just have to satisfy the inequalities

$$\omega(e) \geq 0, \qquad \omega(f) \geq 0, \qquad \omega(g) \geq 0$$

$$\begin{gathered}\omega(e) + \omega(f) \leq 1 \\ \omega(f) + \omega(g) \leq 1 \\ \omega(g) + \omega(e) \leq 1\end{gathered}$$



Example Klyachko

- By a similar argument to Specker, the only normalized state is

$$\omega_j = \tfrac{1}{2} \text{ for } j = 0,1,2,3,4$$

- For unnormalized states we have

$$\omega_j \geq 0$$
$$\omega_j + \omega_{j+1} \ (\mathrm{mod}\ 5) \leq 1$$

- From this, we can derive

$$\omega_0 + \omega_1 + \omega_2 + \omega_3 + \omega_4 \leq \frac{5}{2}$$

which is saturated by the normalized state.