Proof of preparation contextuality
(a preparation noncontextual hidden variable model is impossible)
Important features of hidden variable models

Let $P \leftrightarrow \mu(\lambda)$

$P' \leftrightarrow \mu'(\lambda)$

Representing one-shot distinguishability:

If $P$ and $P'$ are distinguishable with certainty, then $\mu(\lambda) \mu'(\lambda) = 0$
Important features of hidden variable models

Let $P \leftrightarrow \mu(\lambda)$
$P' \leftrightarrow \mu'(\lambda)$

Representing one-shot distinguishability:
If $P$ and $P'$ are distinguishable with certainty
then $\mu(\lambda) \mu'(\lambda) = 0$

Representing convex combination:
If $P'' = P$ with prob. $p$ and $P'$ with prob. $1 - p$
Then $\mu''(\lambda) = p \mu(\lambda) + (1 - p) \mu'(\lambda)$
Proof based on finite construction in 2d

\[ P_a \leftrightarrow \psi_a = (1, 0) \]
\[ P_A \leftrightarrow \psi_A = (0, 1) \]
\[ P_b \leftrightarrow \psi_b = (1/2, \sqrt{3}/2) \]
\[ P_B \leftrightarrow \psi_B = (\sqrt{3}/2, -1/2) \]
\[ P_c \leftrightarrow \psi_c = (1/2, -\sqrt{3}/2) \]
\[ P_C \leftrightarrow \psi_C = (\sqrt{3}/2, 1/2) \]
Proof based on finite construction in 2d

\[
P_a \leftrightarrow \sigma_a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}
\]

\[
P_A \leftrightarrow \sigma_A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}
\]

\[
P_b \leftrightarrow \sigma_b = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} \sqrt{3} \\ \frac{1}{4} \sqrt{3} & \frac{3}{4} \end{pmatrix}
\]

\[
P_B \leftrightarrow \sigma_B = \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} \sqrt{3} \\ -\frac{1}{4} \sqrt{3} & \frac{1}{4} \end{pmatrix}
\]

\[
P_c \leftrightarrow \sigma_c = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} \sqrt{3} \\ -\frac{1}{4} \sqrt{3} & \frac{3}{4} \end{pmatrix}
\]

\[
P_C \leftrightarrow \sigma_c = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \sqrt{3} \\ \frac{1}{4} \sqrt{3} & \frac{1}{4} \end{pmatrix}
\]

\[
\sigma_a \sigma_A = 0
\]

\[
\sigma_b \sigma_B = 0
\]

\[
\sigma_c \sigma_C = 0
\]
Proof based on finite construction in 2d

\[ \sigma_a = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \]
\[ \sigma_A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \]
\[ \sigma_b = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \sqrt{3} \\ \frac{1}{4} \sqrt{3} & \frac{3}{4} \end{pmatrix} \]
\[ \sigma_B = \begin{pmatrix} \frac{3}{4} & -\frac{1}{4} \sqrt{3} \\ -\frac{1}{4} \sqrt{3} & \frac{3}{4} \end{pmatrix} \]
\[ \sigma_c = \begin{pmatrix} \frac{1}{4} \sqrt{3} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix} \]
\[ \sigma_C = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \sqrt{3} \\ \frac{1}{4} \sqrt{3} & \frac{3}{4} \end{pmatrix} \]

\[ \sigma_a \sigma_A = 0 \]
\[ \sigma_b \sigma_B = 0 \]
\[ \sigma_c \sigma_C = 0 \]

\( P_a \) and \( P_A \) are distinguishable with certainty
\( P_b \) and \( P_B \) are distinguishable with certainty
\( P_c \) and \( P_C \) are distinguishable with certainty

\[ \mu_a(\lambda) \mu_A(\lambda) = 0 \]
\[ \mu_b(\lambda) \mu_B(\lambda) = 0 \]
\[ \mu_c(\lambda) \mu_C(\lambda) = 0 \]
\[ P_{aA} \equiv P_a \text{ and } P_A \text{ with prob. } 1/2 \text{ each} \]
\[ P_{bB} \equiv P_b \text{ and } P_B \text{ with prob. } 1/2 \text{ each} \]
\[ P_{cC} \equiv P_c \text{ and } P_C \text{ with prob. } 1/2 \text{ each} \]
\[ P_{abc} \equiv P_a, P_b \text{ and } P_c \text{ with prob. } 1/3 \text{ each} \]
\[ P_{ABC} \equiv P_A, P_B \text{ and } P_C \text{ with prob. } 1/3 \text{ each} \]
\[ P_{aA} \equiv P_a \text{ and } P_A \text{ with prob. } 1/2 \text{ each} \]
\[ P_{bB} \equiv P_b \text{ and } P_B \text{ with prob. } 1/2 \text{ each} \]
\[ P_{cC} \equiv P_c \text{ and } P_C \text{ with prob. } 1/2 \text{ each} \]
\[ P_{abc} \equiv P_a, P_b \text{ and } P_c \text{ with prob. } 1/3 \text{ each} \]
\[ P_{ABC} \equiv P_A, P_B \text{ and } P_C \text{ with prob. } 1/3 \text{ each} \]

\[ \mu_{aA}(\lambda) = \frac{1}{2} \mu_a(\lambda) + \frac{1}{2} \mu_A(\lambda) \]
\[ \mu_{bB}(\lambda) = \frac{1}{2} \mu_b(\lambda) + \frac{1}{2} \mu_B(\lambda) \]
\[ \mu_{cC}(\lambda) = \frac{1}{2} \mu_c(\lambda) + \frac{1}{2} \mu_C(\lambda) \]
\[ \mu_{abc}(\lambda) = \frac{1}{3} \mu_a(\lambda) + \frac{1}{3} \mu_b(\lambda) + \frac{1}{3} \mu_c(\lambda) \]
\[ \mu_{ABC}(\lambda) = \frac{1}{3} \mu_A(\lambda) + \frac{1}{3} \mu_B(\lambda) + \frac{1}{3} \mu_C(\lambda) \]
\[
I/2 = \frac{1}{2}\sigma_a + \frac{1}{2}\sigma_A \\
= \frac{1}{2}\sigma_b + \frac{1}{2}\sigma_B \\
= \frac{1}{2}\sigma_c + \frac{1}{2}\sigma_C \\
= \frac{1}{3}\sigma_a + \frac{1}{3}\sigma_b + \frac{1}{3}\sigma_c \\
= \frac{1}{3}\sigma_A + \frac{1}{3}\sigma_B + \frac{1}{3}\sigma_C.
\]
\[ I/2 = \frac{1}{2} \sigma_a + \frac{1}{2} \sigma_A \]
\[ \quad = \frac{1}{2} \sigma_b + \frac{1}{2} \sigma_B \]
\[ \quad = \frac{1}{2} \sigma_c + \frac{1}{2} \sigma_C \]
\[ \quad = \frac{1}{3} \sigma_a + \frac{1}{3} \sigma_b + \frac{1}{3} \sigma_c \]
\[ \quad = \frac{1}{3} \sigma_A + \frac{1}{3} \sigma_B + \frac{1}{3} \sigma_C. \]

\[ P_{aA} \sim P_{bB} \sim P_{cC} \]
\[ \sim P_{abc} \sim P_{ABC} \]
\[
I/2 = \frac{1}{2} \sigma_a + \frac{1}{2} \sigma_A \\
= \frac{1}{2} \sigma_b + \frac{1}{2} \sigma_B \\
= \frac{1}{2} \sigma_c + \frac{1}{2} \sigma_C \\
= \frac{1}{3} \sigma_a + \frac{1}{3} \sigma_b + \frac{1}{3} \sigma_c \\
= \frac{1}{3} \sigma_A + \frac{1}{3} \sigma_B + \frac{1}{3} \sigma_C.
\]

\[
P_{aA} \simeq P_{bB} \simeq P_{cC} \\
\simeq P_{abc} \simeq P_{ABC}
\]

By **preparation noncontextuality**

\[
\mu_{aA}(\lambda) = \mu_{bB}(\lambda) = \mu_{cC}(\lambda) \\
= \mu_{abc}(\lambda) = \mu_{ABC}(\lambda) \\
\equiv \nu(\lambda)
\]
\[ I/2 = \frac{1}{2} \sigma_a + \frac{1}{2} \sigma_A \]
\[ = \frac{1}{2} \sigma_b + \frac{1}{2} \sigma_B \]
\[ = \frac{1}{2} \sigma_c + \frac{1}{2} \sigma_C \]
\[ = \frac{1}{3} \sigma_a + \frac{1}{3} \sigma_b + \frac{1}{3} \sigma_c \]
\[ = \frac{1}{3} \sigma_A + \frac{1}{3} \sigma_B + \frac{1}{3} \sigma_C. \]

\[ P_{aA} \simeq P_{bB} \simeq P_{cC} \]
\[ \simeq P_{abc} \simeq P_{ABC} \]

By preparation noncontextuality

\[ \mu_{aA}(\lambda) = \mu_{bB}(\lambda) = \mu_{cC}(\lambda) \]
\[ = \mu_{abc}(\lambda) = \mu_{ABC}(\lambda) \]
\[ \equiv \nu(\lambda) \]

\[ \nu(\lambda) = \frac{1}{2} \mu_a(\lambda) + \frac{1}{2} \mu_A(\lambda) \]
\[ = \frac{1}{2} \mu_b(\lambda) + \frac{1}{2} \mu_B(\lambda) \]
\[ = \frac{1}{2} \mu_c(\lambda) + \frac{1}{2} \mu_C(\lambda) \]
\[ = \frac{1}{3} \mu_a(\lambda) + \frac{1}{3} \mu_b(\lambda) + \frac{1}{3} \mu_c(\lambda) \]
\[ = \frac{1}{3} \mu_A(\lambda) + \frac{1}{3} \mu_B(\lambda) + \frac{1}{3} \mu_C(\lambda). \]
\( P_{aA} \equiv P_a \text{ and } P_A \text{ with prob. } 1/2 \text{ each} \)

\( P_{bB} \equiv P_b \text{ and } P_B \text{ with prob. } 1/2 \text{ each} \)

\( P_{cC} \equiv P_c \text{ and } P_C \text{ with prob. } 1/2 \text{ each} \)

\( P_{abc} \equiv P_a, P_b \text{ and } P_c \text{ with prob. } 1/3 \text{ each} \)

\( P_{ABC} \equiv P_A, P_B \text{ and } P_C \text{ with prob. } 1/3 \text{ each} \)

\[ \mu_{aA}(\lambda) = \frac{1}{2} \mu_a(\lambda) + \frac{1}{2} \mu_A(\lambda) \]

\[ \mu_{bB}(\lambda) = \frac{1}{2} \mu_b(\lambda) + \frac{1}{2} \mu_B(\lambda) \]

\[ \mu_{cC}(\lambda) = \frac{1}{2} \mu_c(\lambda) + \frac{1}{2} \mu_C(\lambda) \]

\[ \mu_{abc}(\lambda) = \frac{1}{3} \mu_a(\lambda) + \frac{1}{3} \mu_b(\lambda) + \frac{1}{3} \mu_c(\lambda) \]

\[ \mu_{ABC}(\lambda) = \frac{1}{3} \mu_A(\lambda) + \frac{1}{3} \mu_B(\lambda) + \frac{1}{3} \mu_C(\lambda) \]
\[ I/2 = \frac{1}{2} \sigma_a + \frac{1}{2} \sigma_A \]
\[ = \frac{1}{2} \sigma_b + \frac{1}{2} \sigma_B \]
\[ = \frac{1}{2} \sigma_c + \frac{1}{2} \sigma_C \]
\[ = \frac{1}{3} \sigma_a + \frac{1}{3} \sigma_b + \frac{1}{3} \sigma_c \]
\[ = \frac{1}{3} \sigma_A + \frac{1}{3} \sigma_B + \frac{1}{3} \sigma_C. \]

\[ P_{aA} \simeq P_{bB} \simeq P_{cC} \]
\[ \simeq P_{abc} \simeq P_{ABC} \]

By preparation noncontextuality

\[ \mu_{aA}(\lambda) = \mu_{bB}(\lambda) = \mu_{cC}(\lambda) \]
\[ = \mu_{abc}(\lambda) = \mu_{ABC}(\lambda) \]
\[ \equiv \nu(\lambda) \]

\[ \nu(\lambda) = \frac{1}{2} \mu_a(\lambda) + \frac{1}{2} \mu_A(\lambda) \]
\[ = \frac{1}{2} \mu_b(\lambda) + \frac{1}{2} \mu_B(\lambda) \]
\[ = \frac{1}{2} \mu_c(\lambda) + \frac{1}{2} \mu_C(\lambda) \]
\[ = \frac{1}{3} \mu_a(\lambda) + \frac{1}{3} \mu_b(\lambda) + \frac{1}{3} \mu_c(\lambda) \]
\[ = \frac{1}{3} \mu_A(\lambda) + \frac{1}{3} \mu_B(\lambda) + \frac{1}{3} \mu_C(\lambda). \]
Our task is to find
\[ \mu_a(\lambda), \mu_A(\lambda), \mu_b(\lambda), \]
\[ \mu_B(\lambda), \mu_c(\lambda), \mu_C(\lambda), \]
and \( \nu(\lambda) \) such that

\[ \mu_a(\lambda) \mu_A(\lambda) = 0 \]
\[ \mu_b(\lambda) \mu_B(\lambda) = 0 \]
\[ \mu_c(\lambda) \mu_C(\lambda) = 0 \]

\[ \nu(\lambda) = \frac{1}{2} \mu_a(\lambda) + \frac{1}{2} \mu_A(\lambda) \]
\[ = \frac{1}{2} \mu_b(\lambda) + \frac{1}{2} \mu_B(\lambda) \]
\[ = \frac{1}{2} \mu_c(\lambda) + \frac{1}{2} \mu_C(\lambda) \]
\[ = \frac{1}{3} \mu_a(\lambda) + \frac{1}{3} \mu_b(\lambda) + \frac{1}{3} \mu_c(\lambda) \]
\[ = \frac{1}{3} \mu_A(\lambda) + \frac{1}{3} \mu_B(\lambda) + \frac{1}{3} \mu_C(\lambda). \]
Our task is to find
\[ \mu_a(\lambda), \mu_A(\lambda), \mu_b(\lambda), \]
\[ \mu_B(\lambda), \mu_c(\lambda), \mu_C(\lambda), \]
and \( \nu(\lambda) \) such that
\[
\mu_a(\lambda) \mu_A(\lambda) = 0 \\
\mu_b(\lambda) \mu_B(\lambda) = 0 \\
\mu_c(\lambda) \mu_C(\lambda) = 0
\]

\[ \nu(\lambda) = \frac{1}{2} \mu_a(\lambda) + \frac{1}{2} \mu_A(\lambda) \]
\[ = \frac{1}{2} \mu_b(\lambda) + \frac{1}{2} \mu_B(\lambda) \]
\[ = \frac{1}{2} \mu_c(\lambda) + \frac{1}{2} \mu_C(\lambda) \]
\[ = \frac{1}{3} \mu_a(\lambda) + \frac{1}{3} \mu_b(\lambda) + \frac{1}{3} \mu_c(\lambda) \]
\[ = \frac{1}{3} \mu_A(\lambda) + \frac{1}{3} \mu_B(\lambda) + \frac{1}{3} \mu_C(\lambda). \]

i.e., paralleling the quantum structure:
\[
\sigma_a \sigma_A = 0 \\
\sigma_b \sigma_B = 0 \\
\sigma_c \sigma_C = 0
\]

\[ I/2 = \frac{1}{2} \sigma_a + \frac{1}{2} \sigma_A \]
\[ = \frac{1}{2} \sigma_b + \frac{1}{2} \sigma_B \]
\[ = \frac{1}{2} \sigma_c + \frac{1}{2} \sigma_C \]
\[ = \frac{1}{3} \sigma_a + \frac{1}{3} \sigma_b + \frac{1}{3} \sigma_c \]
\[ = \frac{1}{3} \sigma_A + \frac{1}{3} \sigma_B + \frac{1}{3} \sigma_C. \]
Our task is to find

\[ \mu_a(\lambda), \mu_A(\lambda), \mu_b(\lambda), \]
\[ \mu_B(\lambda), \mu_c(\lambda), \mu_C(\lambda), \]
and \( \nu(\lambda) \) such that

\[ \mu_a(\lambda) \mu_A(\lambda) = 0 \]
\[ \mu_b(\lambda) \mu_B(\lambda) = 0 \]
\[ \mu_c(\lambda) \mu_C(\lambda) = 0 \]

\[ \nu(\lambda) = \frac{1}{2} \mu_a(\lambda) + \frac{1}{2} \mu_A(\lambda) \]
\[ = \frac{1}{2} \mu_b(\lambda) + \frac{1}{2} \mu_B(\lambda) \]
\[ = \frac{1}{2} \mu_c(\lambda) + \frac{1}{2} \mu_C(\lambda) \]
\[ = \frac{1}{3} \mu_a(\lambda) + \frac{1}{3} \mu_b(\lambda) + \frac{1}{3} \mu_c(\lambda) \]
\[ = \frac{1}{3} \mu_A(\lambda) + \frac{1}{3} \mu_B(\lambda) + \frac{1}{3} \mu_C(\lambda) \]
Our task is to find

\[ \mu_a(\lambda), \mu_A(\lambda), \mu_b(\lambda), \mu_B(\lambda), \mu_c(\lambda), \mu_C(\lambda), \]

and \( \nu(\lambda) \) such that

\[ \begin{align*}
\mu_a(\lambda) \mu_A(\lambda) &= 0 \\
\mu_b(\lambda) \mu_B(\lambda) &= 0 \\
\mu_c(\lambda) \mu_C(\lambda) &= 0
\end{align*} \]

From decompositions (1)-(3), for \( \lambda = \lambda' \)

\[ \begin{align*}
\mu_a(\lambda') &= 0 \text{ or } 2\nu(\lambda') \\
\mu_b(\lambda') &= 0 \text{ or } 2\nu(\lambda') \\
\mu_c(\lambda') &= 0 \text{ or } 2\nu(\lambda')
\end{align*} \]

\[ \nu(\lambda) = \frac{1}{2} \mu_a(\lambda) + \frac{1}{2} \mu_A(\lambda) \]
\[ = \frac{1}{2} \mu_b(\lambda) + \frac{1}{2} \mu_B(\lambda) \]
\[ = \frac{1}{2} \mu_c(\lambda) + \frac{1}{2} \mu_C(\lambda) \]
\[ = \frac{1}{3} \mu_a(\lambda) + \frac{1}{3} \mu_b(\lambda) + \frac{1}{3} \mu_c(\lambda) \]
\[ = \frac{1}{3} \mu_A(\lambda) + \frac{1}{3} \mu_B(\lambda) + \frac{1}{3} \mu_C(\lambda) \]
Our task is to find
\( \mu_a(\lambda), \mu_A(\lambda), \mu_b(\lambda), \mu_B(\lambda), \mu_c(\lambda), \mu_C(\lambda), \)
and \( \nu(\lambda) \) such that

\[
\begin{align*}
\mu_a(\lambda) \mu_A(\lambda) &= 0 \\
\mu_b(\lambda) \mu_B(\lambda) &= 0 \\
\mu_c(\lambda) \mu_C(\lambda) &= 0
\end{align*}
\]

\[
\nu(\lambda) = \frac{1}{2} \mu_a(\lambda) + \frac{1}{2} \mu_A(\lambda)
\]
\[
= \frac{1}{2} \mu_b(\lambda) + \frac{1}{2} \mu_B(\lambda)
\]
\[
= \frac{1}{2} \mu_c(\lambda) + \frac{1}{2} \mu_C(\lambda)
\]
\[
= \frac{1}{3} \mu_a(\lambda) + \frac{1}{3} \mu_b(\lambda) + \frac{1}{3} \mu_c(\lambda)
\]
\[
= \frac{1}{3} \mu_A(\lambda) + \frac{1}{3} \mu_B(\lambda) + \frac{1}{3} \mu_C(\lambda)
\]

From decompositions (1)-(3), for \( \lambda = \lambda' \)

\[
\begin{align*}
\mu_a(\lambda') &= 0 \text{ or } 2\nu(\lambda') \\
\mu_b(\lambda') &= 0 \text{ or } 2\nu(\lambda') \\
\mu_c(\lambda') &= 0 \text{ or } 2\nu(\lambda')
\end{align*}
\]

But then the RHS of decomposition (4) is

\[
0, \frac{2}{3} \nu(\lambda'), \frac{4}{3} \nu(\lambda'), 2\nu(\lambda') \\
\neq \nu(\lambda')
\]

for \( \lambda' \) such that \( \nu(\lambda') \neq 0 \)

**CONTRADICTION**
Measurement noncontextuality
new definition versus old
Another feature of a hidden variable model

Let \( M \leftrightarrow \{ \chi_k(\lambda) \} \)

\( M' \leftrightarrow \{ \chi'_j(\lambda) \} \)

Representing coarse-graining of measurement outcomes:
Suppose the outcomes \( k \) of \( M \) are sorted into subsets \( S_j \). Suppose \( M' \equiv \) implement \( M \) and upon obtaining outcome \( k \), record the \( j \) such that \( k \in S_j \).

Then \( \chi'_j(\lambda) = \sum_{k \in S_j} \chi_k(\lambda) \)

\[
\begin{align*}
\chi_1(\lambda) & \quad \chi_1(\lambda) \\
\chi_2(\lambda) & \quad \chi_{-1}(\lambda) \\
\chi_3(\lambda) & \quad \chi_{-1}(\lambda)
\end{align*}
\]
Recall the traditional notion of noncontextuality:

\[ |\psi_1\rangle \quad |\psi_2\rangle \quad |\psi_3\rangle \quad \iff \quad \chi_1(\lambda) \quad \chi_2(\lambda) \quad \chi_3(\lambda) \]

\[ |\psi_1'\rangle \quad |\psi_2'\rangle \quad |\psi_3'\rangle \quad \iff \quad \chi_1(\lambda) \quad \chi_2'(\lambda) \quad \chi_3'(\lambda) \]

\( \chi_1(\lambda) \) is the same in the two cases.
This is equivalent to assuming:

\[ \chi_1(\lambda) \]  
\[ \chi_{-1}(\lambda) \]

measure \[ |\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle \]  
coarse-grain \[ |\psi'_2\rangle \text{ and } |\psi'_3\rangle \]

\[ \{ |\psi_1\rangle\langle\psi_1|, I - |\psi_1\rangle\langle\psi_1| \} \]

\[ \chi_1(\lambda) \]  
\[ \chi_{-1}(\lambda) \]
Recall the traditional notion of noncontextuality:

\[|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle \quad \leftrightarrow \quad \chi_1(\lambda), \chi_2(\lambda), \chi_3(\lambda)\]

\[|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle \quad \leftrightarrow \quad \chi_1(\lambda), \chi'_2(\lambda), \chi'_3(\lambda)\]

\[\chi_1(\lambda) \text{ is the same in the two cases}\]
This is equivalent to assuming:

\[ \psi_1 \langle \psi |, I - |\psi_1 \rangle \langle \psi_1 | \] for measure and coarse-grain of \( |\psi_2 \rangle \) and \( |\psi_3 \rangle \)

\[ \psi_1 \langle \psi |, I - |\psi_1 \rangle \langle \psi_1 | \] for measure and coarse-grain of \( |\psi'_2 \rangle \) and \( |\psi'_3 \rangle \)
Recall the traditional notion of noncontextuality:

\[ |\psi_1\rangle \quad |\psi_2\rangle \quad |\psi_3\rangle \\
\leftrightarrow \\
\chi_1(\lambda) \quad \chi_2(\lambda) \quad \chi_3(\lambda) \\
\lambda \quad \lambda \quad \lambda

\[ |\psi_1\rangle \quad |\psi_2\rangle \quad |\psi_3\rangle \\
\leftrightarrow \\
\chi_1(\lambda) \quad \chi_2'(\lambda) \quad \chi_3'(\lambda) \\
\lambda \quad \lambda \quad \lambda

\chi_1(\lambda) \text{ is the same in the two cases}
This is equivalent to assuming:

\[ \{ |\psi_1\rangle\langle\psi_1|, I - |\psi_1\rangle\langle\psi_1| \} \]

\[ \{ |\psi_1\rangle\langle\psi_1|, I - |\psi_1\rangle\langle\psi_1| \} \]
But recall that the most general representation was

$$\{ P_k \} \quad \overset{M}{\leftrightarrow} \quad \xi_{P_1}(\lambda), \quad \xi_{P_2}(\lambda), \quad \xi_{P_3}(\lambda)$$
This is equivalent to assuming:

\[ \{|\psi_1\rangle\langle\psi_1|, I - |\psi_1\rangle\langle\psi_1|\} \]

measure \[|\psi_2\rangle \text{ and } |\psi_3\rangle \]

coarse-grain

\[ \chi_1(\lambda) \]

\[ \chi_{-1}(\lambda) \]

\[ M \]

\[ M' \]

\[ \{|\psi_1\rangle\langle\psi_1|, I - |\psi_1\rangle\langle\psi_1|\} \]

measure \[|\psi_2\rangle \text{ and } |\psi_3\rangle \]

coarse-grain
But recall that the most general representation was

\[ \{ P_k \} \quad \xrightarrow{M} \quad \xi_{P_1}(\lambda) \xrightarrow{} \lambda \]

\[ \xi_{P_2}(\lambda) \xrightarrow{} \lambda \]

\[ \xi_{P_3}(\lambda) \xrightarrow{} \lambda \]
But recall that the most general representation was

\[ \{ P_k \} \quad \xrightarrow[\text{M}]{\quad} \quad \xi_{P_1}(\lambda) \quad \rightarrow \lambda \]
\[ \xi_{P_2}(\lambda) \quad \rightarrow \lambda \]
\[ \xi_{P_3}(\lambda) \quad \rightarrow \lambda \]

Therefore:

\text{traditional notion of noncontextuality} \quad = \quad \text{revised notion of noncontextuality for sharp measurements and outcome determinism for sharp measurements}
So, the proposed definition of noncontextuality is not simply a generalization of the traditional notion.

For sharp measurements, it is a revision of the traditional notion.
So, the proposed definition of noncontextuality is not simply a generalization of the traditional notion.

For sharp measurements, it is a revision of the traditional notion.

Noncontextuality and determinism are separate issues!
But recall that the most general representation was

\[ \{ P_k \} \rightarrow M \]

\[ \xi_{P_1}(\lambda) \rightarrow \lambda \]

\[ \xi_{P_2}(\lambda) \rightarrow \lambda \]

\[ \xi_{P_3}(\lambda) \rightarrow \lambda \]

Therefore:

traditional notion of noncontextuality = revised notion of noncontextuality for sharp measurements and outcome determinism for sharp measurements
So, the proposed definition of noncontextuality is not simply a generalization of the traditional notion. For sharp measurements, it is a revision of the traditional notion.
traditional notion of noncontextuality = revised notion of noncontextuality for sharp measurements and outcome determinism for sharp measurements

No-go theorems for previous notion are not necessarily no-go theorems for the new notion!

In face of contradiction, could give up ODSM
But

preparation noncontextuality → outcome determinism for sharp measurements
But

preparation noncontextuality $\rightarrow$ outcome determinism for sharp measurements

Therefore:

measurement noncontextuality and preparation noncontextuality $\rightarrow$ noncontextuality for sharp measurements and outcome determinism for sharp measurements
But

preparation noncontextuality \rightarrow \text{outcome determinism for sharp measurements}

Therefore:

measurement noncontextuality and preparation noncontextuality \rightarrow \text{Traditional notion of noncontextuality}
But

preparation noncontextuality \hspace{1cm} \rightarrow \hspace{1cm} \text{outcome determinism for sharp measurements}

Therefore:

measurement noncontextuality \hspace{1cm} \rightarrow \hspace{1cm} \text{Traditional notion of noncontextuality}

and

preparation noncontextuality

no-go theorems for the traditional notion of noncontextuality can be salvaged as no-go theorems for the generalized notion

... and there are many new proofs, even in 2d
Is contextuality mysterious?
Is contextuality mysterious?

I would say YES.
Is contextuality mysterious?

I would say YES.

• There is a tension between the dependence of representation on certain details of the experimental procedure and the independence of outcome statistics on those details of the experimental procedure.
Phenomena that are a form of generalized contextuality

- all variants of the Bell-Kochen-Specker theorem (algebraic, state-specific, statistical, continuous, discrete)

- all variants of Bell’s theorem

- all the novel no-go theorems, including the 2d ones (see RS, PRA 71, 052108)

- Aspects of pre- and post-selected “paradoxes” (joint work with M. Leifer, PRL 95, 200405)

- The necessity of having negativity in quasi-probability representations of quantum theory

- all variants of von Neumann’s no-go theorem

- Quantum improvements in certain IP tasks
Phenomena that are a form of generalized contextuality

- all variants of the Bell-Kochen-Specker theorem (algebraic, state-specific, statistical, continuous, discrete)

- all variants of Bell’s theorem

- all the novel no-go theorems, including the 2d ones (see RS, PRA 71, 052108)

- Aspects of pre- and post-selected “paradoxes” (joint work with M. Leifer, PRL 95, 200405)

- The necessity of having negativity in quasi-probability representations of quantum theory

- all variants of von Neumann’s no-go theorem

- Quantum improvements in certain IP tasks
Conclusions about contextuality

The notion of contextuality can and should be separated from that of outcome indeterminism.

It can be extended to preparations and unsharp measurements.

It can be made operational and thus subject to experimental test.

It powers better-than-classical performance of certain information-processing tasks.

The generalized notion is seen to be an umbrella for many notions of nonclassicality.
Open questions

What other notions of nonclassicality might be instances of contextuality? Fermionic statistics?

What other information-processing tasks might be powered by contextuality? Quantum computation?

Can we quantify contextuality as a resource?

Why isn’t the world more contextual? For instance, why can’t we implement perfect parity-oblivious 2-to-1 random access code?

What physical principle relieves the tension between the context-dependence at the hidden variable level and the lack of context-dependence at the operational level?

Do quantum states describe reality or our knowledge of reality?
Do quantum states describe reality or our knowledge of reality?
“But our present QM formalism is not purely epistemological; it is a peculiar mixture describing in part realities of Nature, in part incomplete human information about Nature --- all scrambled up by Heisenberg and Bohr into an omelette that nobody has seen how to unscramble. Yet we think that the unscrambling is a prerequisite for any further advance in basic physical theory. For, if we cannot separate the subjective and objective aspects of the formalism, we cannot know what we are talking about; it is just that simple.”

--E.T. Jaynes
$\psi$-complete vs. $\psi$-incomplete

$\psi$-ontic vs. $\psi$-epistemic

$\psi$-complete

---

Complete state is $\psi$
$\psi$-complete vs. $\psi$-incomplete

$\psi$-ontic vs. $\psi$-epistemic

$\psi$-complete

Complete state is $\psi$

$\psi$-incomplete

Complete state is $(\psi, \lambda)$

Complete state is $\lambda$
**ψ-complete vs. ψ-incomplete**

**ψ-ontic vs. ψ-epistemic**

**ψ-complete**

Complete state is $\psi$

**ψ-incomplete**

Complete state is $(\psi, \lambda)$

**ψ-ontic**

**ψ-epistemic**

Complete state is $\lambda$
$\psi$-complete vs. $\psi$-incomplete

$\psi$-ontic vs. $\psi$-epistemic

$\psi$-complete

Complete state is $\psi$

$\psi$-incomplete

Complete state is $(\psi, \omega)$

$\psi$-ontic

$\psi$-epistemic

Complete state is $\lambda$
ψ-complete model:

Space of physical states = space of rays in Hilbert space

\[ \lambda = \psi \]

ψ-ontic model:

For preparation procedures \( P_{|\psi_1\rangle}, P_{|\psi_2\rangle} \) with \( |\psi_1\rangle \neq |\psi_2\rangle \)

\[ \mu(\lambda|P_{|\psi_1\rangle})\mu(\lambda|P_{|\psi_2\rangle}) = 0 \text{ for all } \lambda \]

ψ-epistemic model:

\[ \mu(\lambda|P_{|\psi_1\rangle})\mu(\lambda|P_{|\psi_2\rangle}) \neq 0 \text{ for some } \lambda \]
$\psi$-complete vs. $\psi$-incomplete

$\psi$-ontic vs. $\psi$-epistemic

$\psi$-complete

Complete state is $\psi$

$\psi$-ontic

$\psi$-incomplete

Complete state is $(\psi, \omega)$

$\psi$-ontic

$\psi$-epistemic

Complete state is $\lambda$
ψ-complete model:
Space of physical states = space of rays in Hilbert space
\[ \lambda = \psi \]

ψ-ontic model:
For preparation procedures \( P_{\psi_1}, P_{\psi_2} \) with \( |\psi_1\rangle \neq |\psi_2\rangle \)
\[ \mu(\lambda |P_{\psi_1}\rangle)\mu(\lambda |P_{\psi_2}\rangle) = 0 \text{ for all } \lambda \]

ψ-epistemic model:
\[ \mu(\lambda |P_{\psi_1}\rangle)\mu(\lambda |P_{\psi_2}\rangle) \neq 0 \text{ for some } \lambda \]
$\psi$-complete vs. $\psi$-incomplete

$\psi$-ontic vs. $\psi$-epistemic

$\psi$-complete

Complete state is $\psi$

$\psi$-incomplete

Complete state is $(\psi, \omega)$

$\psi$-ontic

Complete state is $\psi$

$\psi$-epistemic

Complete state is $\lambda$
ψ-complete model:

Space of physical states = space of rays in Hilbert space
\[ \lambda = \psi \]

ψ-ontic model:

For preparation procedures \[ P_{\psi_1}, P_{\psi_2} \] with \[ |\psi_1\rangle \neq |\psi_2\rangle \]
\[ \mu(\lambda|P_{\psi_1})\mu(\lambda|P_{\psi_2}) = 0 \text{ for all } \lambda \]

ψ-epistemic model:

\[ \mu(\lambda|P_{\psi_1})\mu(\lambda|P_{\psi_2}) \neq 0 \text{ for some } \lambda \]
$\psi$-complete vs. $\psi$-incomplete

$\psi$-ontic vs. $\psi$-epistemic

$\psi$-complete

$\psi$-ontic

Complete state is $\psi$

$\psi$-incomplete

$\psi$-ontic

Complete state is $(\psi, \omega)$

$\psi$-epistemic

Complete state is $\lambda$
ψ-complete vs. ψ-incomplete

ψ-ontic vs. ψ-epistemic

ψ-complete

Complete state is ψ

ψ-ontic

ψ-incomplete

Complete state is (ψ, ω)

ψ-ontic

Complete state is λ

ψ-epistemic