What is generalized noncontextuality?

David Schmid

Solstice of Foundations 2022

What it is

Why it is our best notion of classical-explainability

How to study it (systematically!)

We all believe that quantum theory is weird and can't be explained by "any classical theory".

But what is the rigorous proof of this?
What particular phenomena resist classical explanations?
In what rigorous sense?

We need a principled way of dividing phenomena into those which can be "explained classically", and those which are rigorous proofs of nonclassicality.

What we want in a notion of nonclassicality

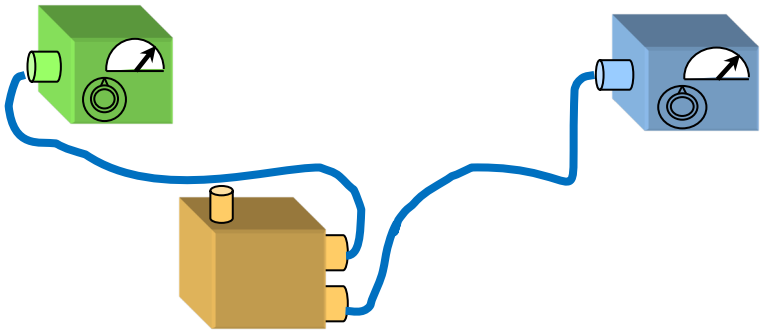
Subject to direct experimental test	Constitutes a resource	Applicable to a broad range of physical scenarios

What we want in a notion of nonclassicality

	Subject to direct experimental test	Constitutes a resource	Applicable to a broad range of physical scenarios
Nonlocality	✓	✓	✗

What is needed to witness nonlocality

particular causal structure
two or more systems
entanglement
incompatible mmts
freedom of choice
highly efficient detectors
space-like separation



What is needed to witness contextuality

none of the above are needed

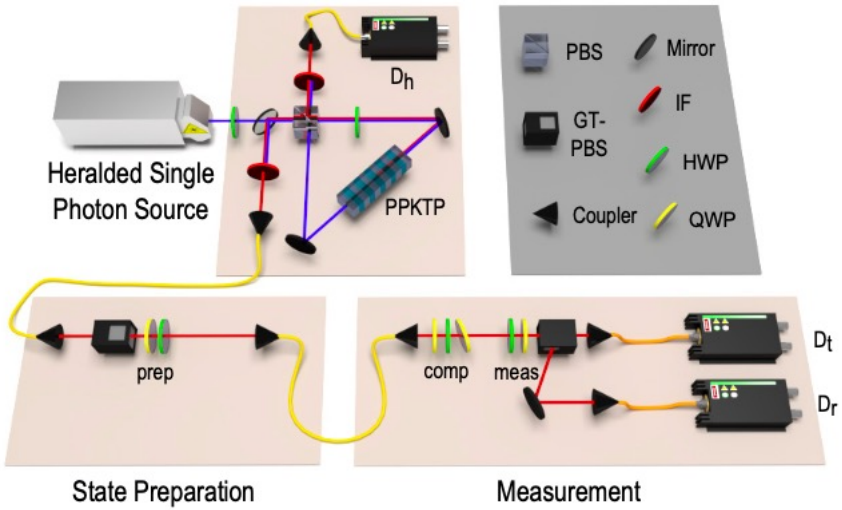
What we want in a notion of nonclassicality

	Subject to direct experimental test	Constitutes a resource	Applicable to a broad range of physical scenarios
Nonlocality	✓	✓	✗
Contextuality	✓	✓	✓

computation
communication
cryptography

metrology
cloning
state discrimination

arxiv:1505.06244



Don't need:

any particular causal structure
multiple systems
entanglement
incompatible measurements
freedom of choice
highly efficient detectors
space-like separation

Like in a Bell test, one doesn't need:

validity of quantum theory
determinism
pure states
projective mmts

What we want in a notion of nonclassicality

	Subject to direct experimental test	Constitutes a resource	Applicable to a broad range of physical scenarios
Nonlocality	✓	✓	✗
Contextuality	✓	✓	✓

What we want in a notion of nonclassicality

	Subject to direct experimental test	Constitutes a resource	Applicable to a broad range of physical scenarios
Nonlocality	✓	✓	✗
Contextuality	✓	✓	✓
Kochen-Specker contextuality	?	✓	~

Noncontextuality

Observed facts

Noncontextuality

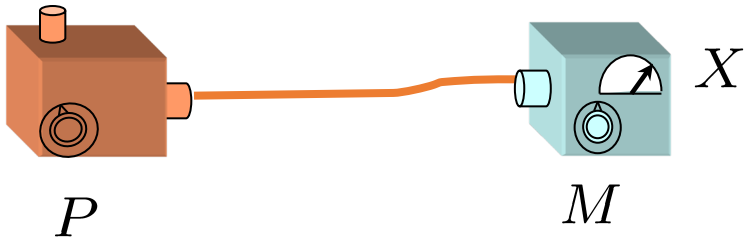
Hypothesized explanations

arXiv:0406166

Observed facts

Operational theory

operational procedure: detailed list of laboratory instructions

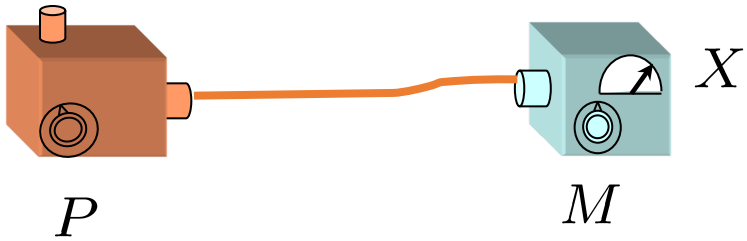


$$p(X|M,P)$$

observed probabilities

Observed facts

Operational theory



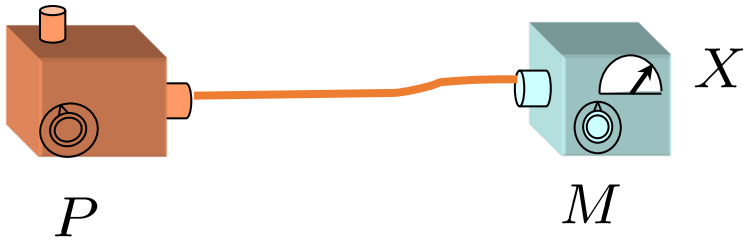
$$p(X|M,P)$$

Ontological model of an operational theory

Hypothesized explanations

Observed facts

Operational theory



$$p(X|M,P)$$

Ontological model of an operational theory

Hypothesized explanations

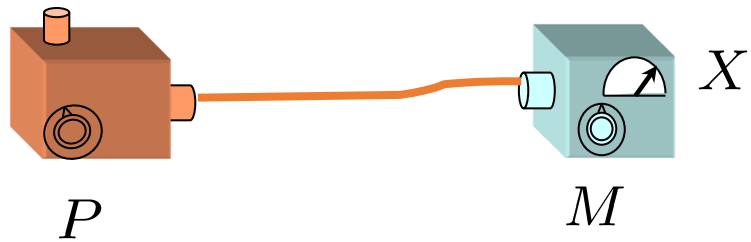
$\lambda \in \Lambda$ Ontic state space

This is the set of physical attributes a system may possess
-position, momentum, mass, charge, color…

Outcomes of measurements just reveal something about these properties
ƛ causally mediates between P and M

Observed facts

Operational theory



$$p(X|M,P)$$

Ontological model of an operational theory

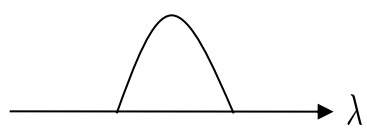
Hypothesized explanations

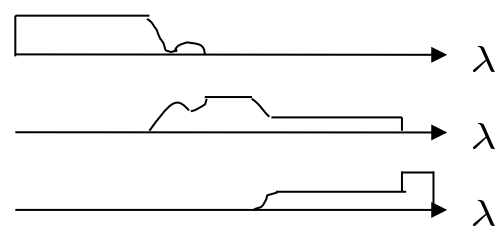
$\lambda \in \Lambda$ Ontic state space

$\mathsf{P} \leftrightarrow \mu(\lambda|\mathsf{P})$

$\mathsf{M} \leftrightarrow \xi(X|\mathsf{M},\lambda)$

epistemic state





response functions

$$p(X|M,P) = \sum_\lambda \xi(X|M,\lambda)\,\mu(\lambda|P)$$

arXiv:0706.2661

An ontological model of an operational theory is noncontextual if

experimental procedures which always lead to the same observational data $\longrightarrow$ identical representations in the ontological model

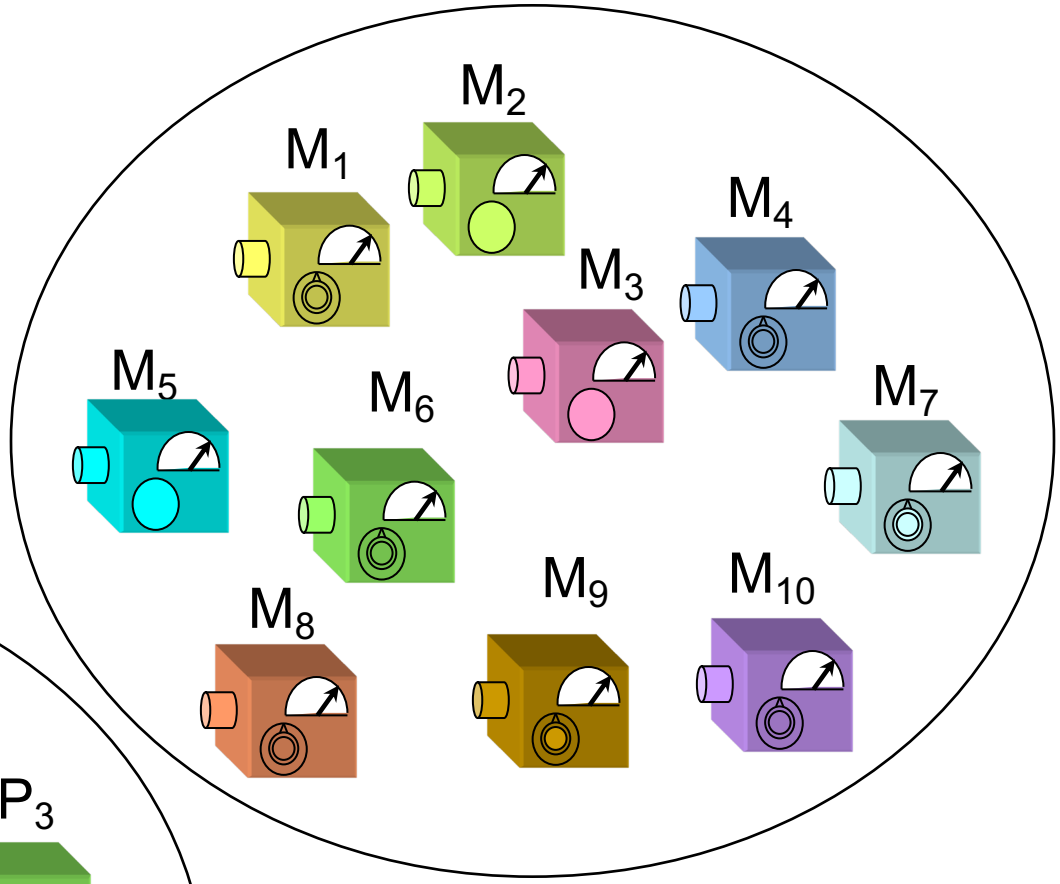
arXiv:0406166

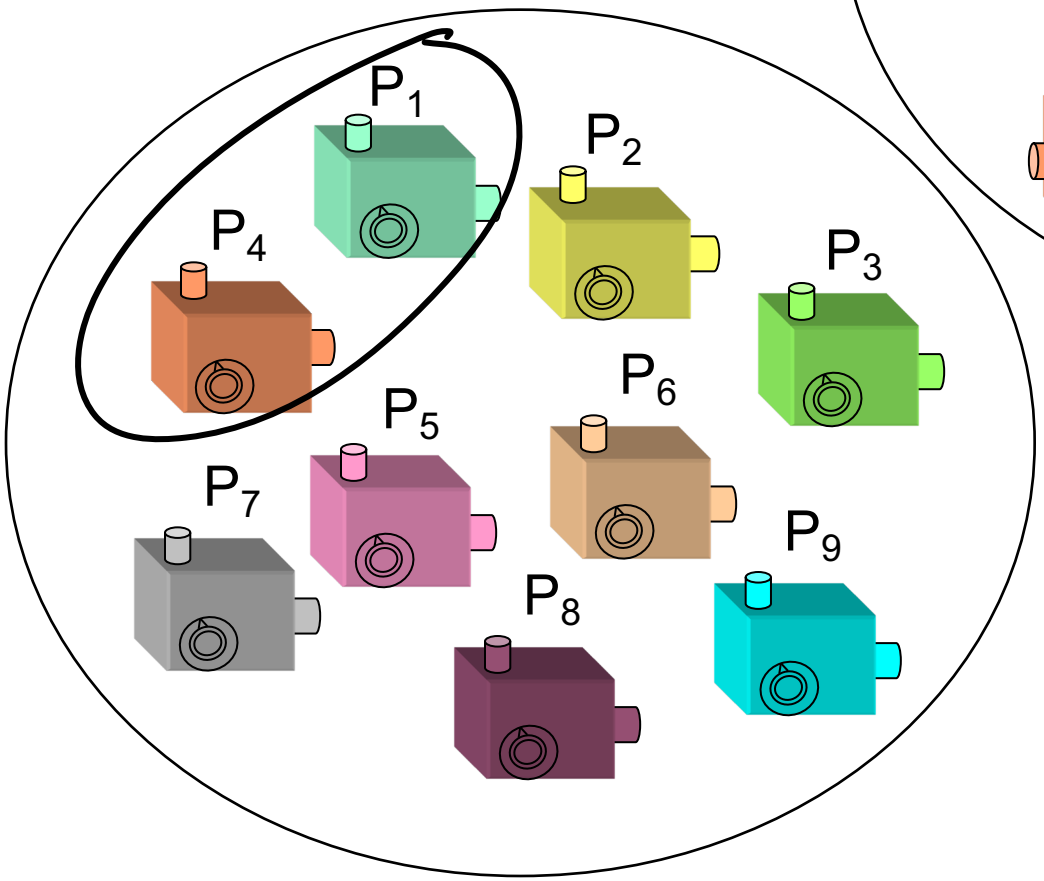
Noncontextuality for preparations

Operational equivalence

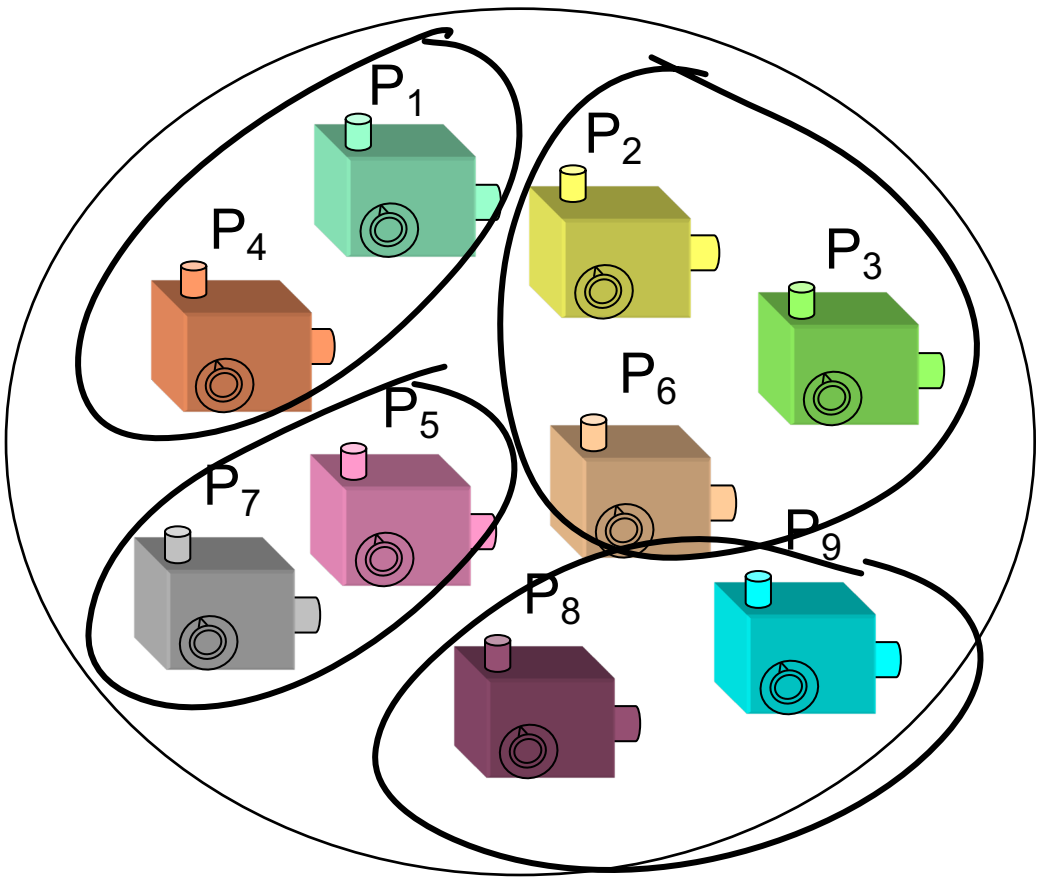
$$\mathsf{P}_1 \simeq \mathsf{P}_4$$

$$\forall \mathsf{M} : p(X|\mathsf{P}_1, \mathsf{M}) = p(X|\mathsf{P}_4, \mathsf{M})$$



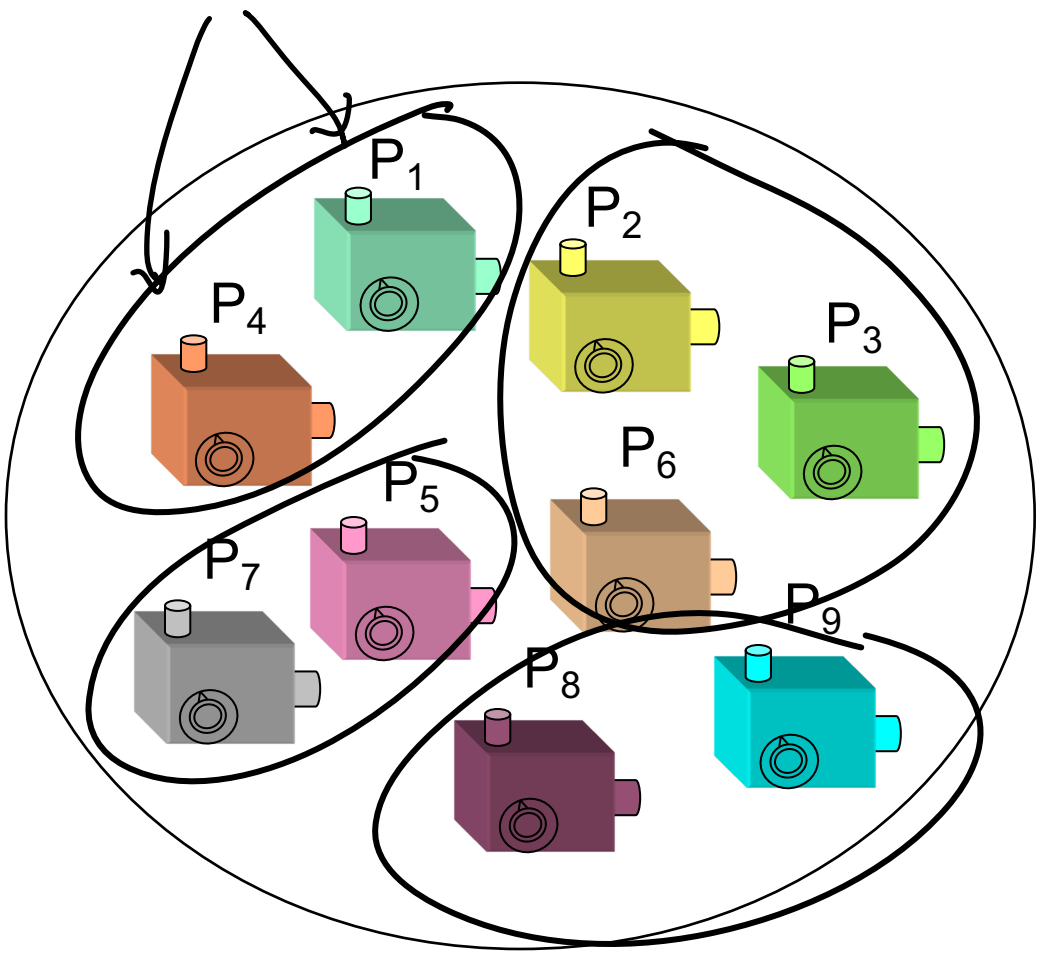


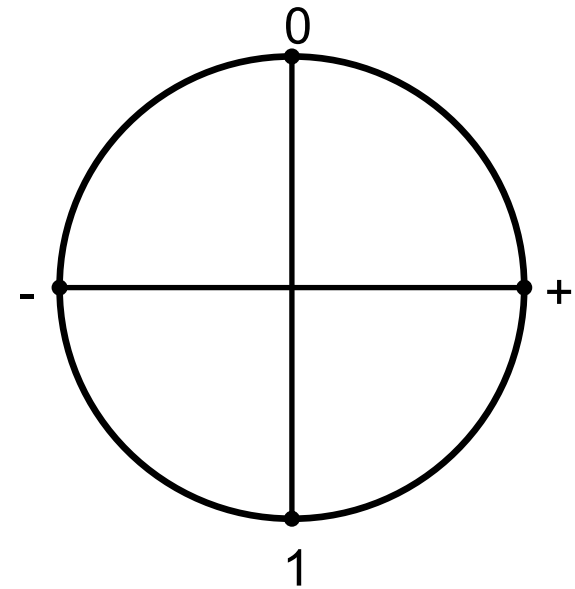
We can group the preparations into operational equivalence classes:



Quantum example

Difference of “context”



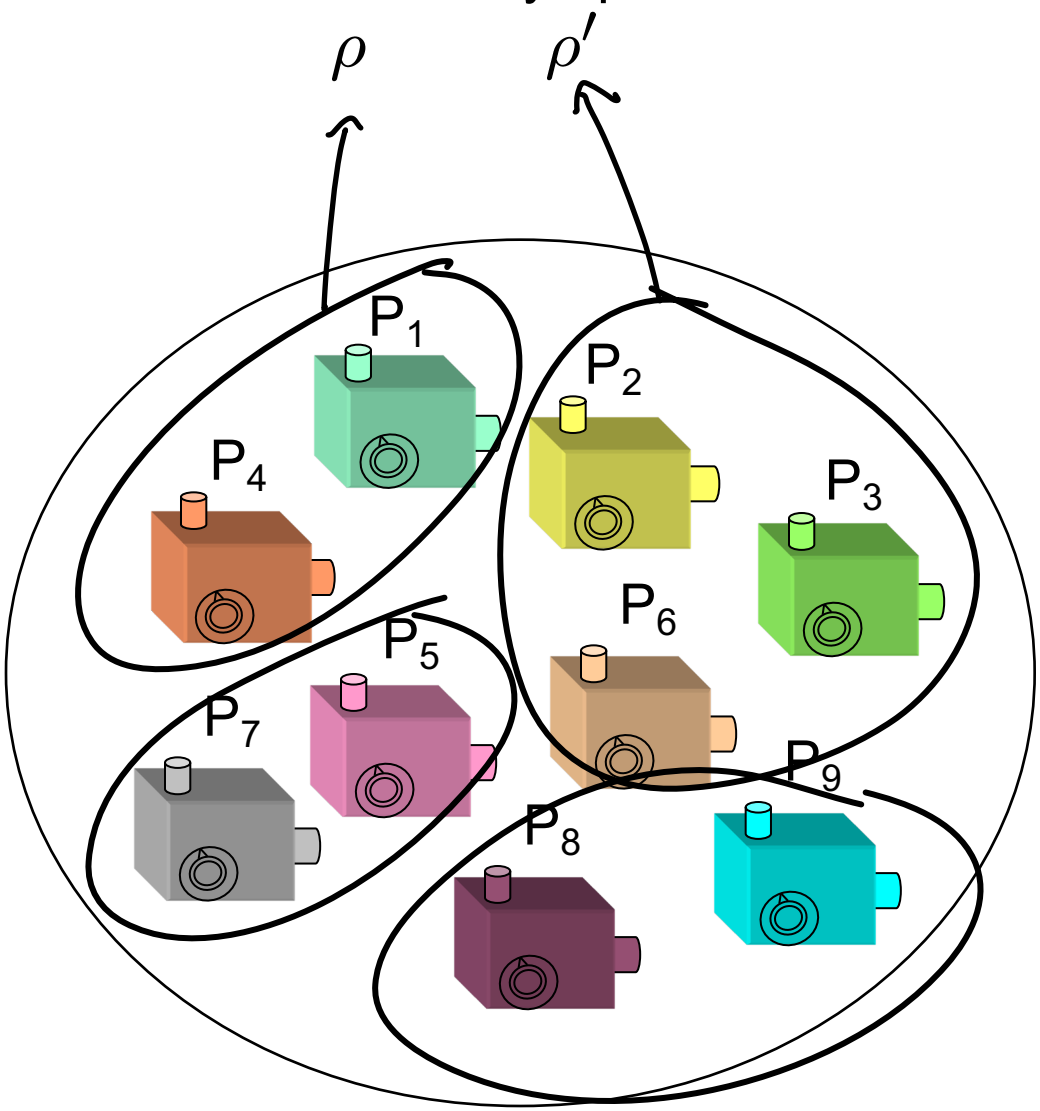


$$\frac{1}{2}I = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$$

$$= \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|$$

Quantum example

Different density operators

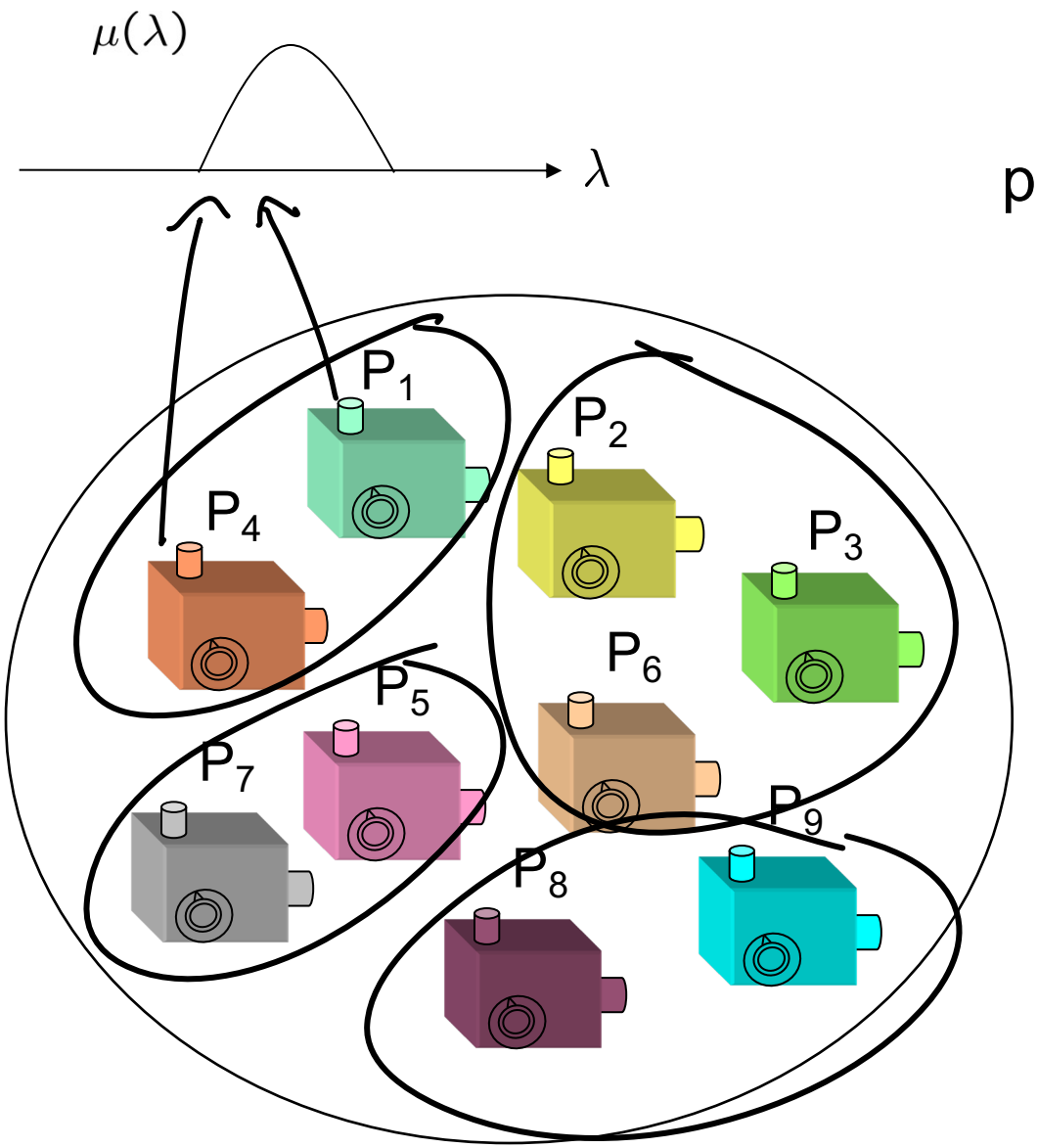


operational equivalence

$\Longleftrightarrow$

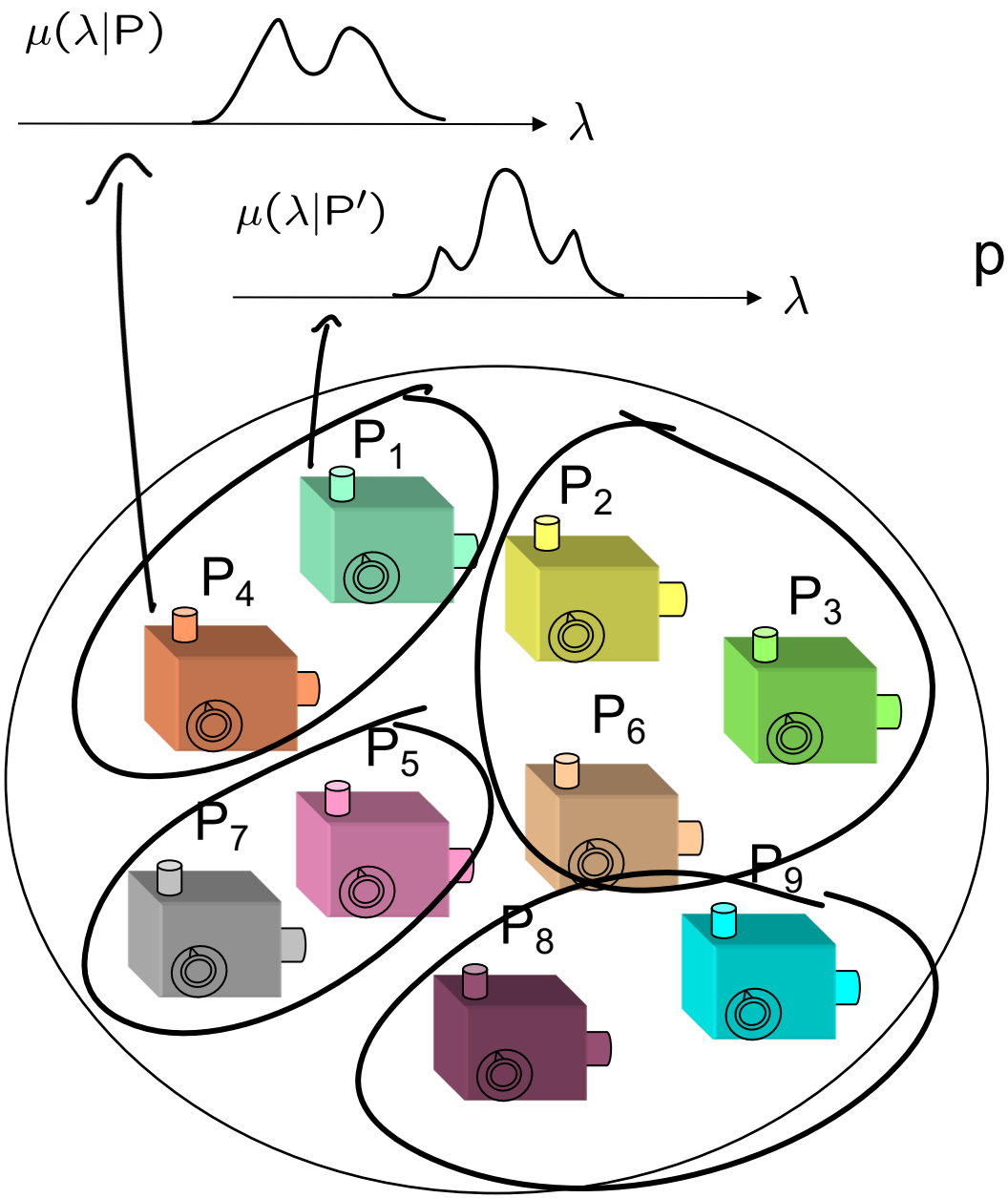
same density operator

Preparation noncontextual ontological model



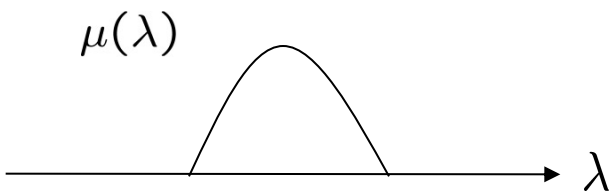
representation of each preparation does not depend on the context

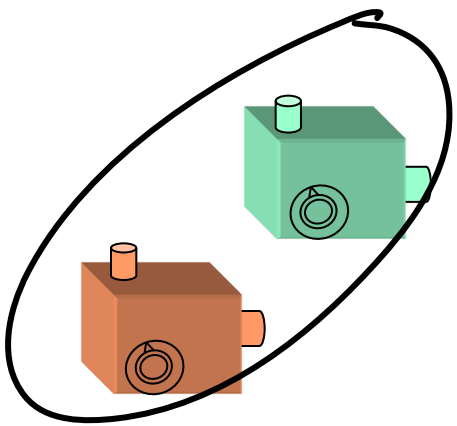
Preparation contextual ontological model



representation of each preparation **does** depend on the context

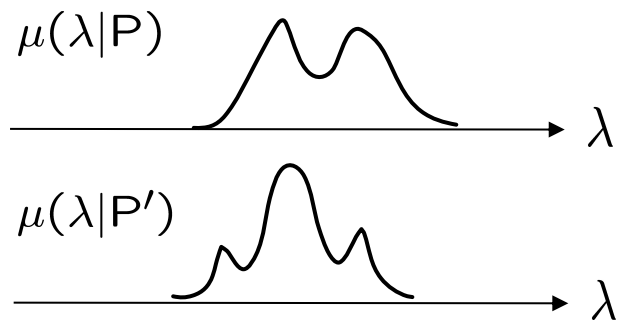
Preparation noncontextual model





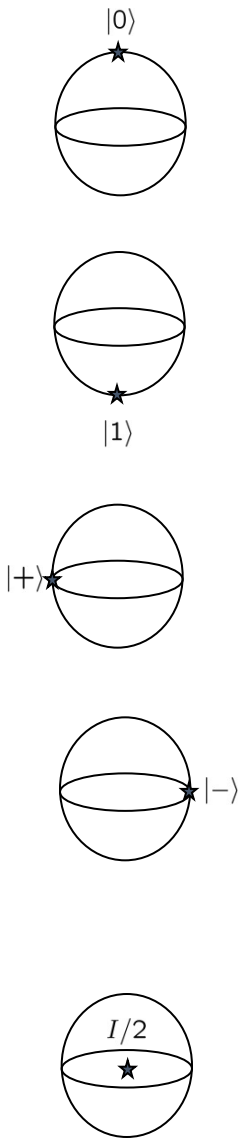
PNC *explains* why two preparations are in the same equivalence class

Preparation contextual model



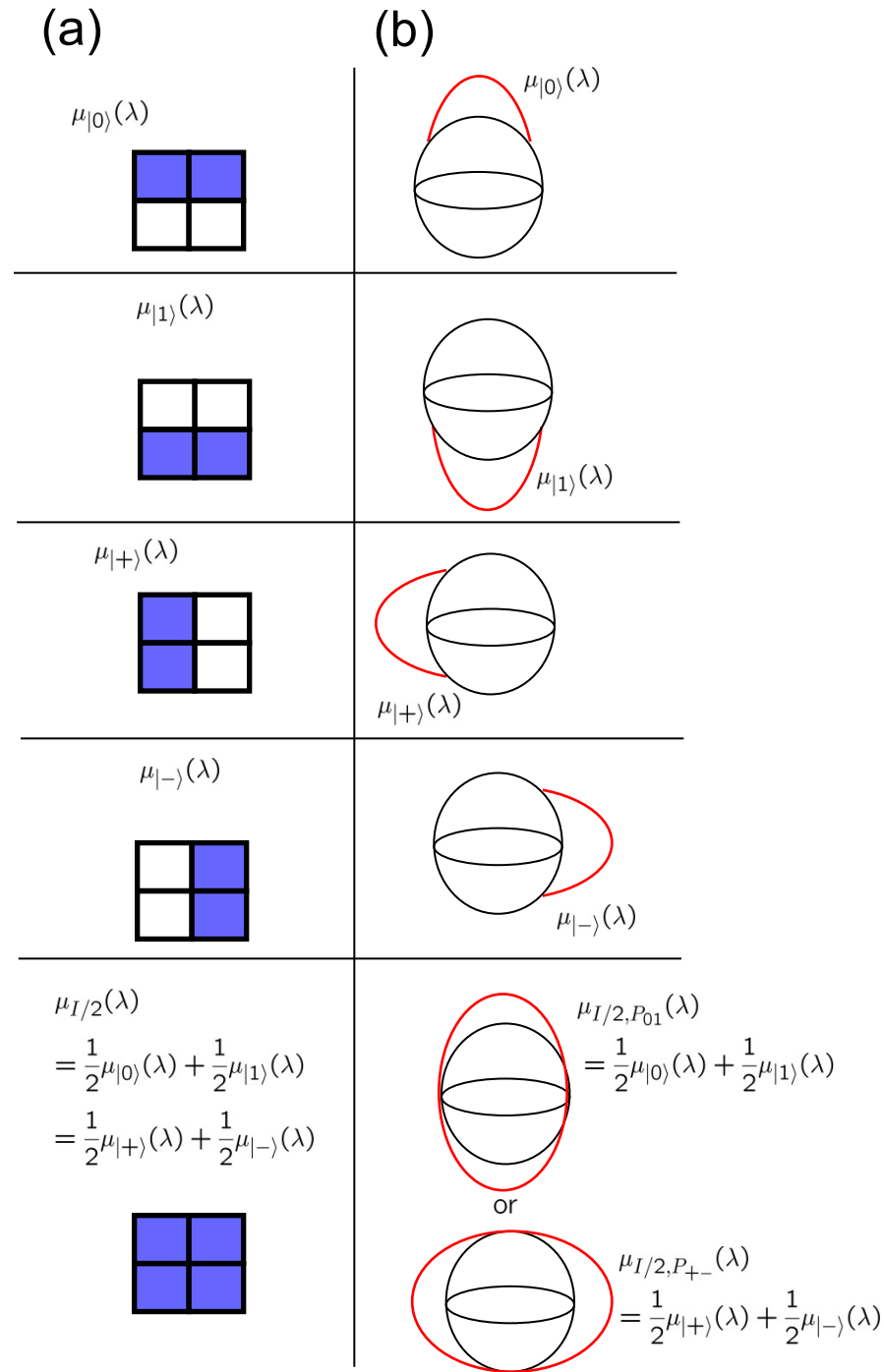
contextual models require a conspiratorial ‘fine-tuning’ of response functions

Five quantum states of a qubit

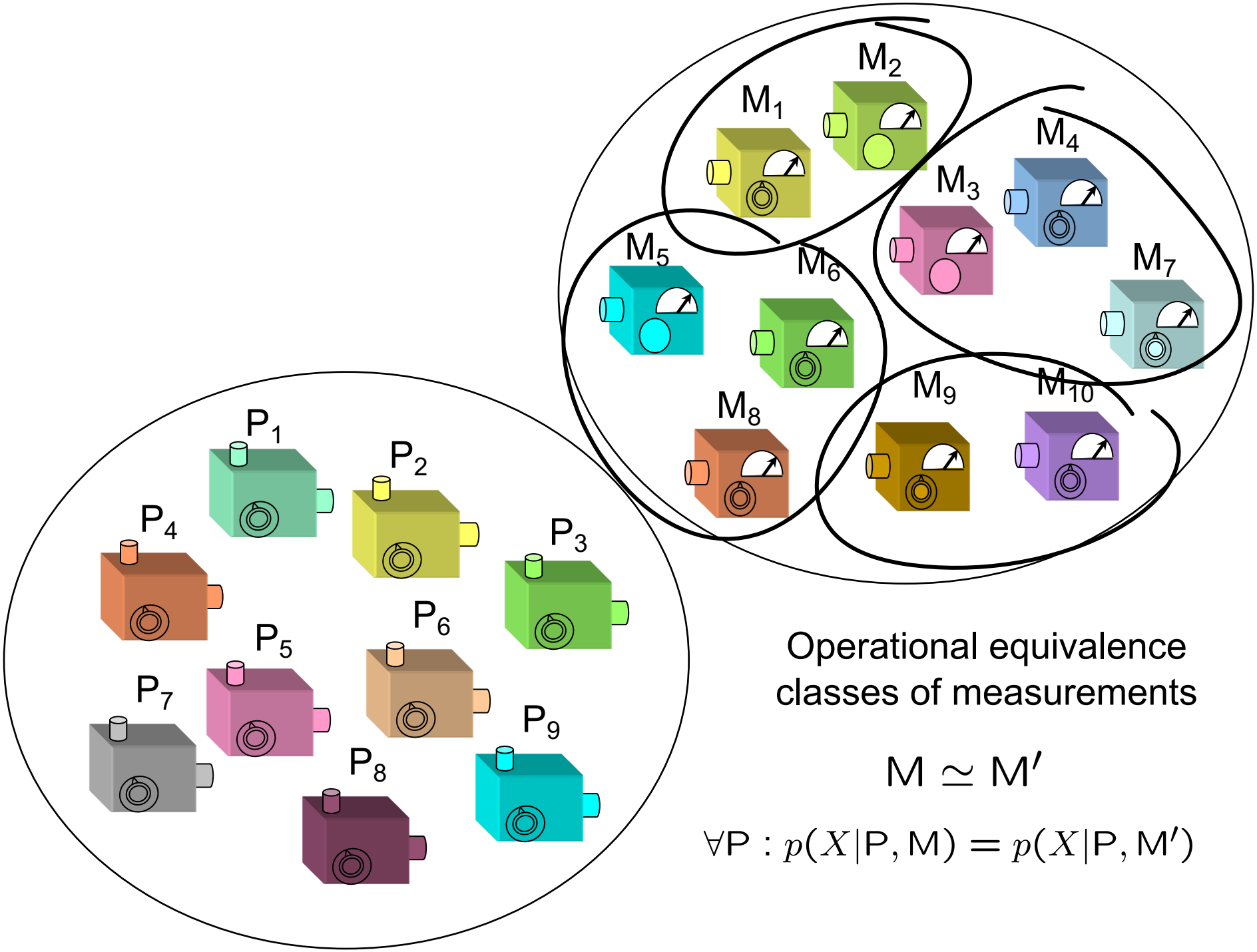


(a) A preparation noncontextual model of these (Spekkens toy theory)

(b) A preparation contextual model of these (Kochen-Specker model)



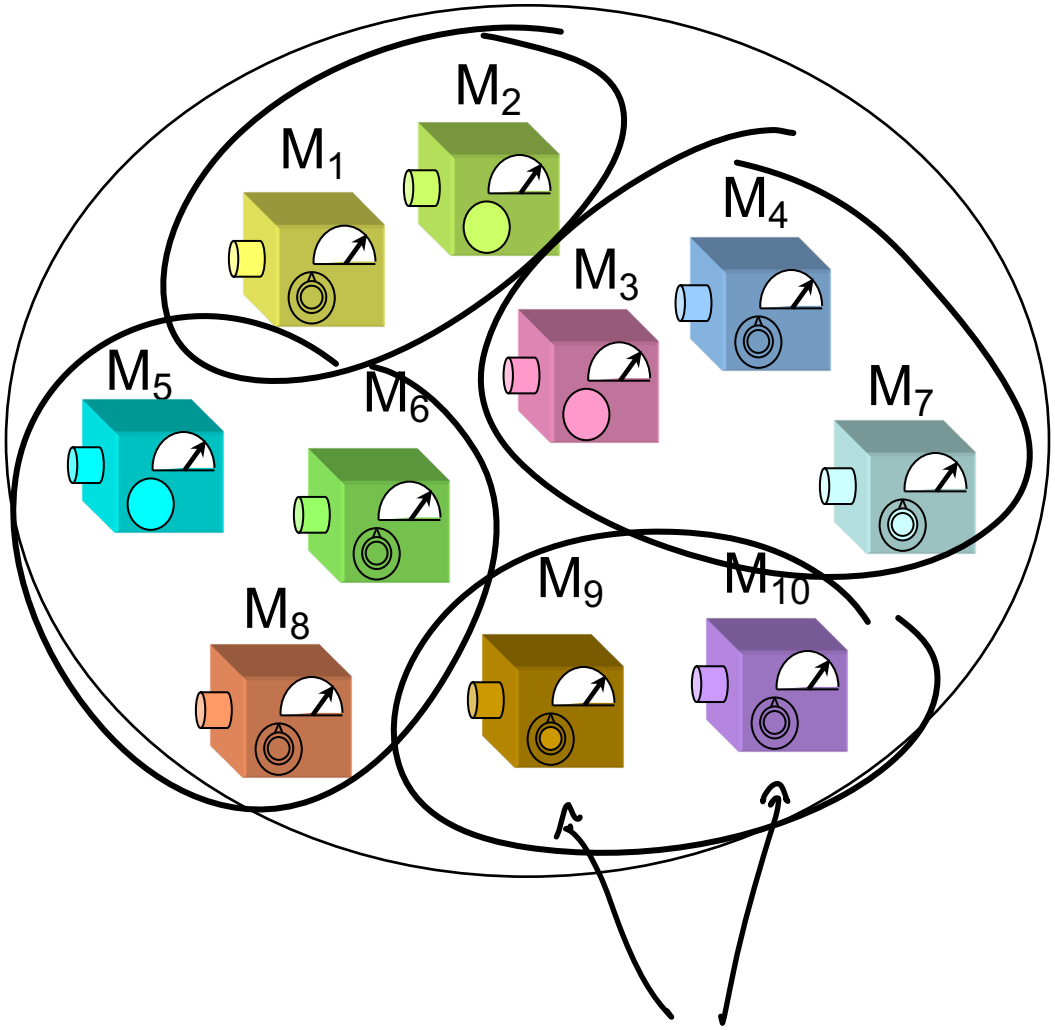
Noncontextuality for measurements



Operational equivalence classes of measurements

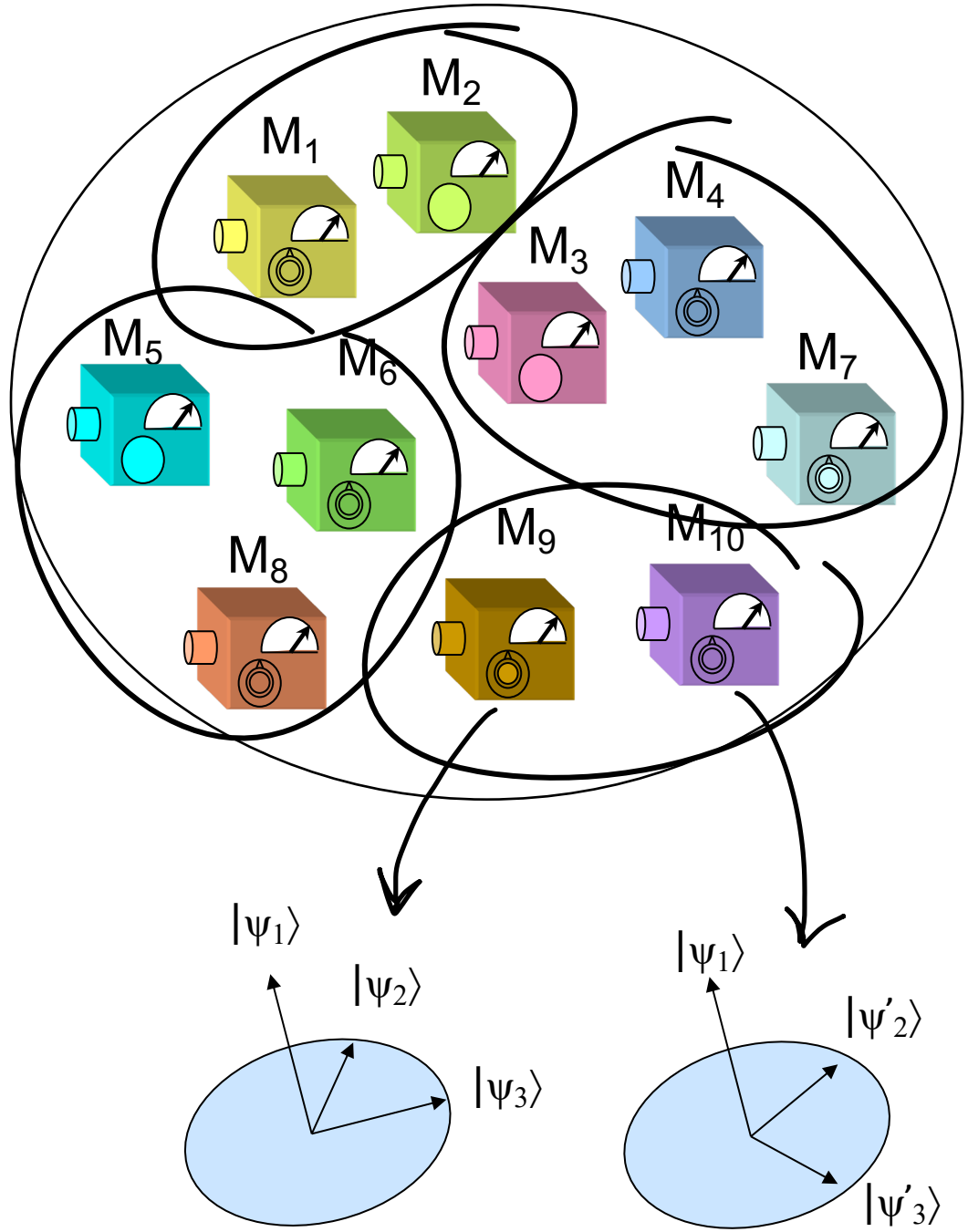
$$\mathsf{M} \simeq \mathsf{M}'$$

$$\forall \mathsf{P} : p(X|\mathsf{P},\mathsf{M}) = p(X|\mathsf{P},\mathsf{M}')$$



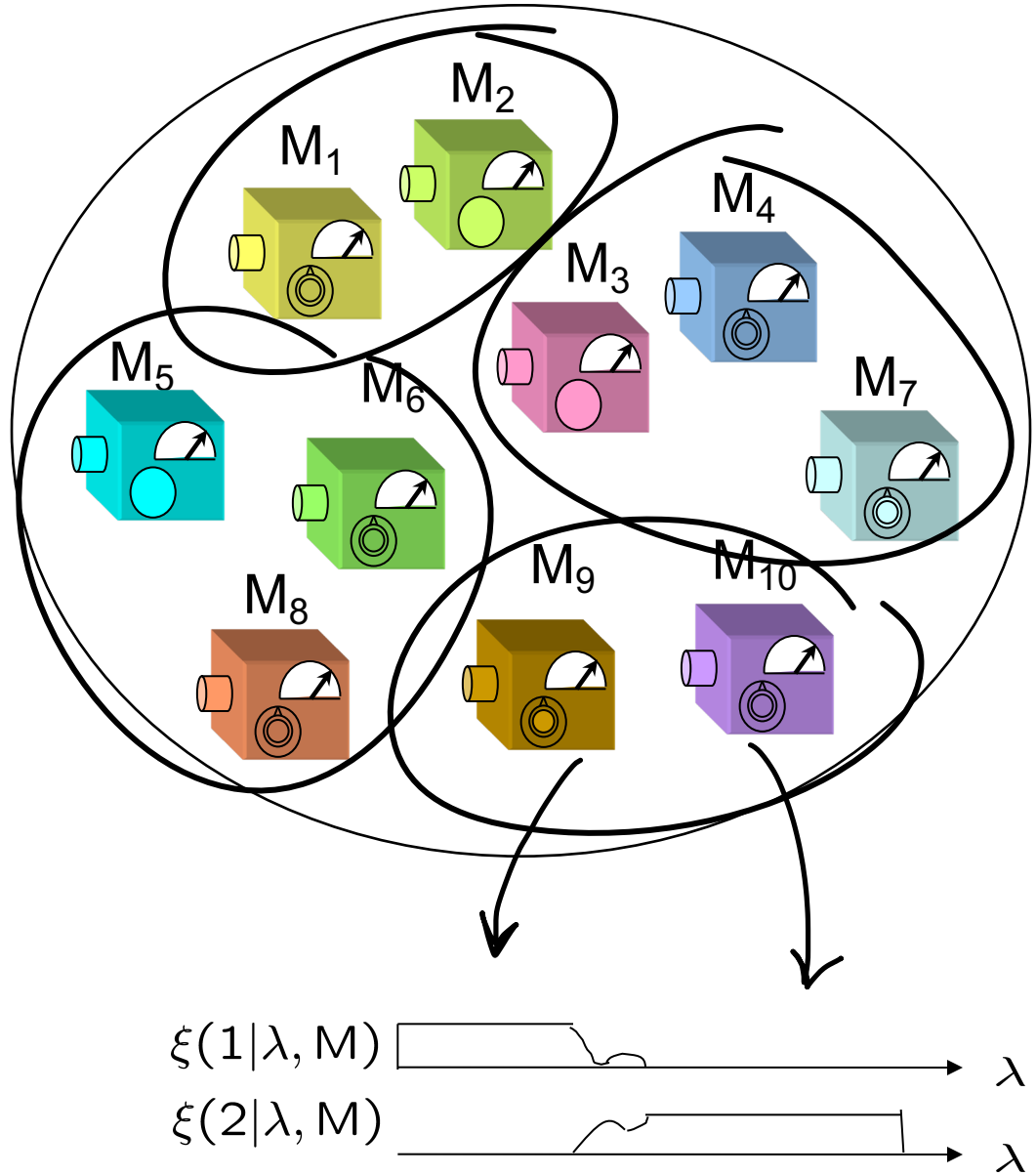
Difference of context

Example from quantum theory

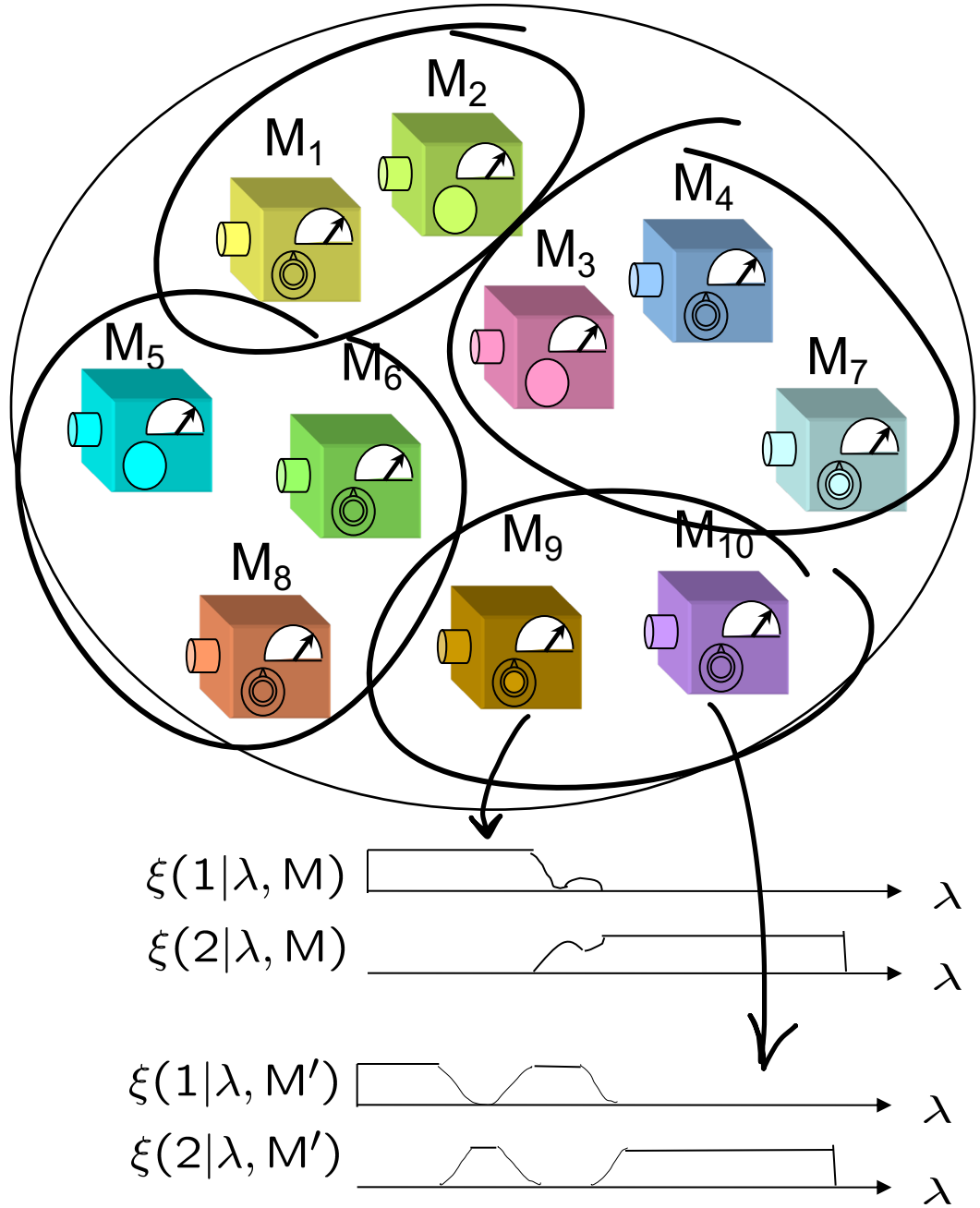


$$|\psi_2\rangle\langle\psi_2| + |\psi_3\rangle\langle\psi_3| = |\psi_2'\rangle\langle\psi_2'| + |\psi_3'\rangle\langle\psi_3'|$$

$$\{|\psi_1\rangle\langle\psi_1|, I - |\psi_1\rangle\langle\psi_1|\}$$



Measurement
noncontextual model



Measurement contextual model

Noncontextuality for transformations

(similarly)

Observed facts $\longrightarrow$ **Hypothesized explanations**

$$\mathsf{P} \simeq \mathsf{P}'$$

$$\forall \mathsf{M} : p(X|\mathsf{P},\mathsf{M}) = p(X|\mathsf{P}',\mathsf{M}) \xrightarrow{\text{noncontextuality}} \mu(\lambda|\mathsf{P}) = \mu(\lambda|\mathsf{P}')$$

$$\mathsf{M} \simeq \mathsf{M}'$$

$$\forall \mathsf{P} : p(X|\mathsf{P},\mathsf{M}) = p(X|\mathsf{P},\mathsf{M}') \xrightarrow{\text{noncontextuality}} \xi(X|\lambda,\mathsf{M}) = \xi(X|\lambda,\mathsf{M}')$$

Can define a similar notion for any type of process (transformations, circuit fragments, etc)

This is a rigorous way of dividing phenomena into
those which can be “explained classically”
and
those which are rigorous proofs of nonclassicality:

An experiment/theory/phenomena is classically-explainable if you can reproduce the operational predictions in some NCOM.

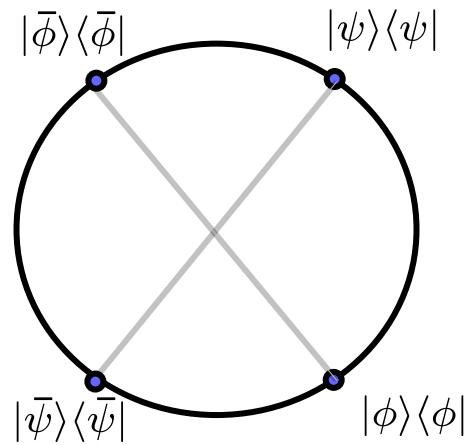
Alternatively, if you can prove that <u>no possible</u> NCOM reproduces the operational predictions, then it is a proof of nonclassicality.

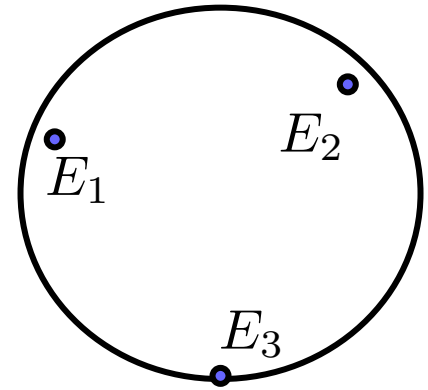
Typical construction of a
noncontextuality no-go theorem

(explicit example in my third talk)

consider a specific circuit

…and the states/effects/transformations on it





Find the operational equivalences these satisfy:

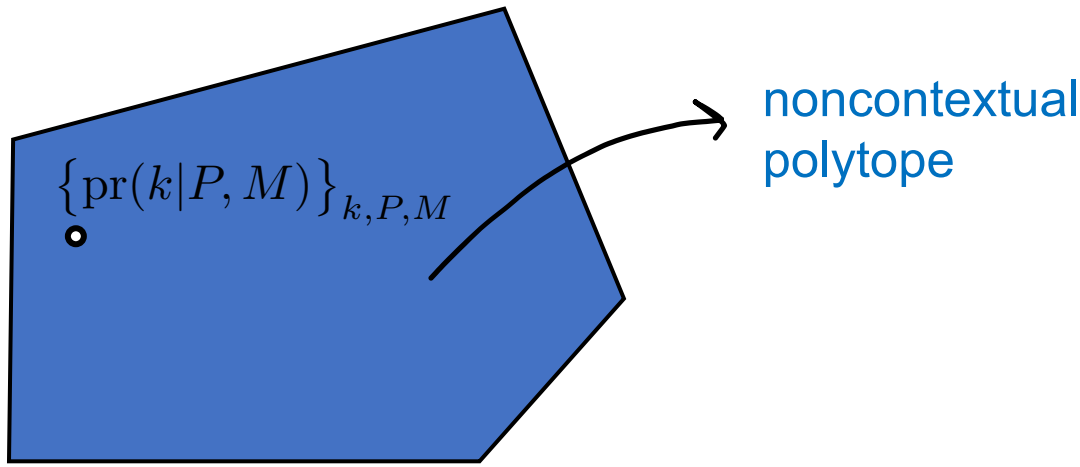
$$\frac{1}{2}\left|\phi\right\rangle\left\langle\phi\right| + \frac{1}{2}\left|\bar{\phi}\right\rangle\left\langle\bar{\phi}\right| = \frac{1}{2}\left|\psi\right\rangle\left\langle\psi\right| + \frac{1}{2}\left|\bar{\psi}\right\rangle\left\langle\bar{\psi}\right|$$

These imply constraints on any NC representation:

$$\frac{1}{2}\mu_\phi(\lambda) + \frac{1}{2}\mu_{\bar{\phi}}(\lambda) = \frac{1}{2}\mu_\psi(\lambda) + \frac{1}{2}\mu_{\bar{\psi}}(\lambda)$$

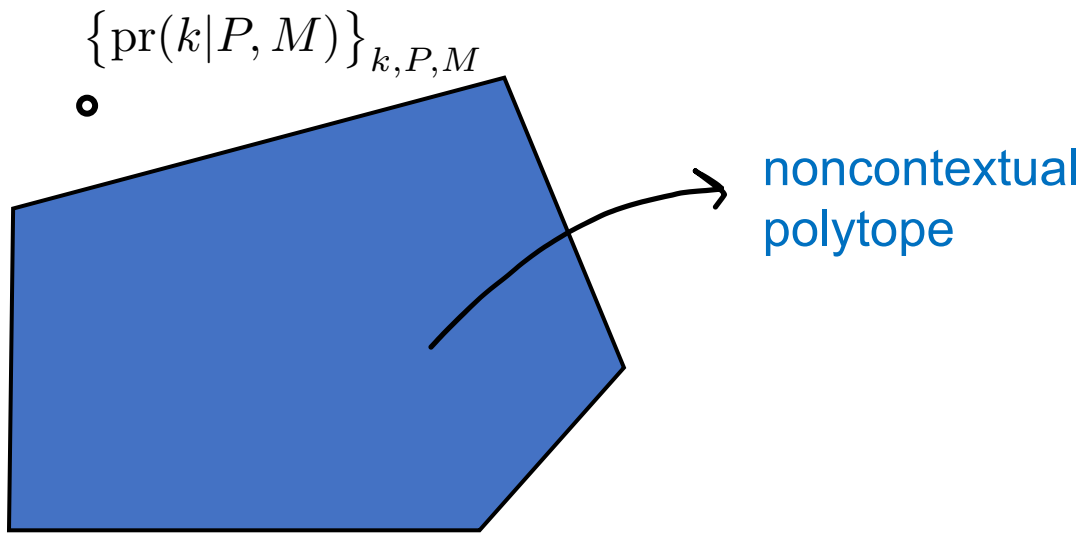
…which imply bounds on the observable data — noncontextuality inequalities!

$$\mathrm{pr}(k|P,M) \qquad \sum_{P,k,M} \alpha_{P,k,M}\ \mathrm{pr}(k|P,M) \leq r$$



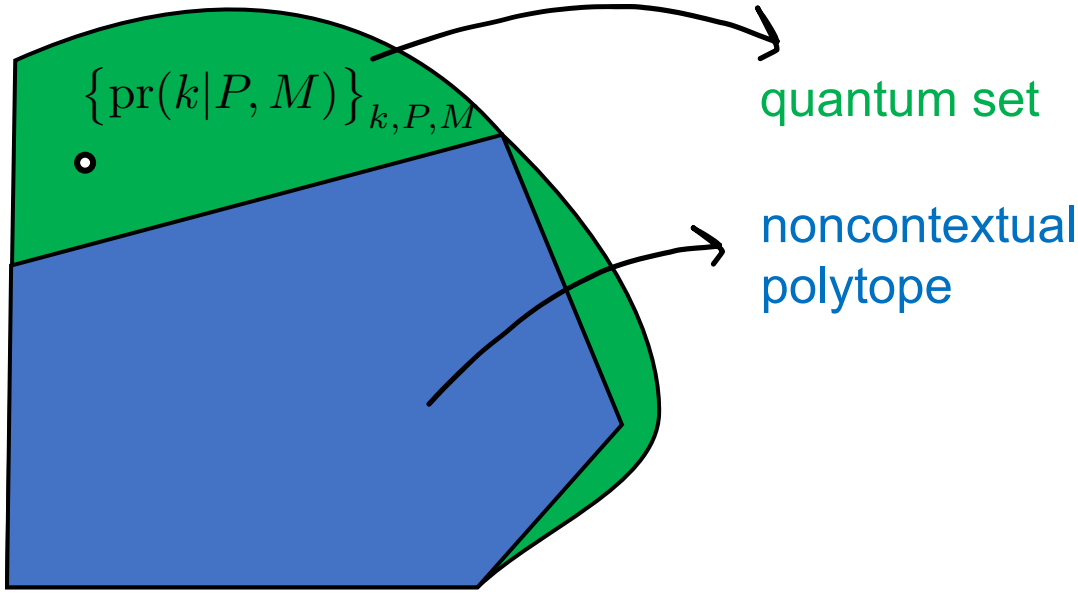
…which imply bounds on the observable data — noncontextuality inequalities!

$$\mathrm{pr}(k|P,M) \qquad \sum_{P,k,M} \alpha_{P,k,M}\ \mathrm{pr}(k|P,M) \leq r$$



Observing data which violates any noncontextuality inequality is a proof of nonclassicality.

Such proofs don't rely on the correctness of quantum theory

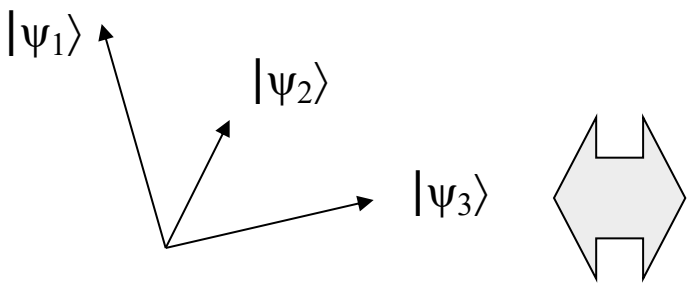


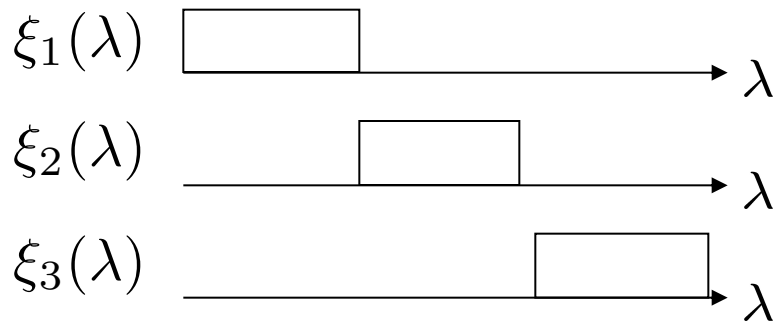
Observing data which violates any noncontextuality inequality is a proof of nonclassicality.

Such proofs don't rely on the correctness of quantum theory

Relationship with Kochen-Specker noncontextuality

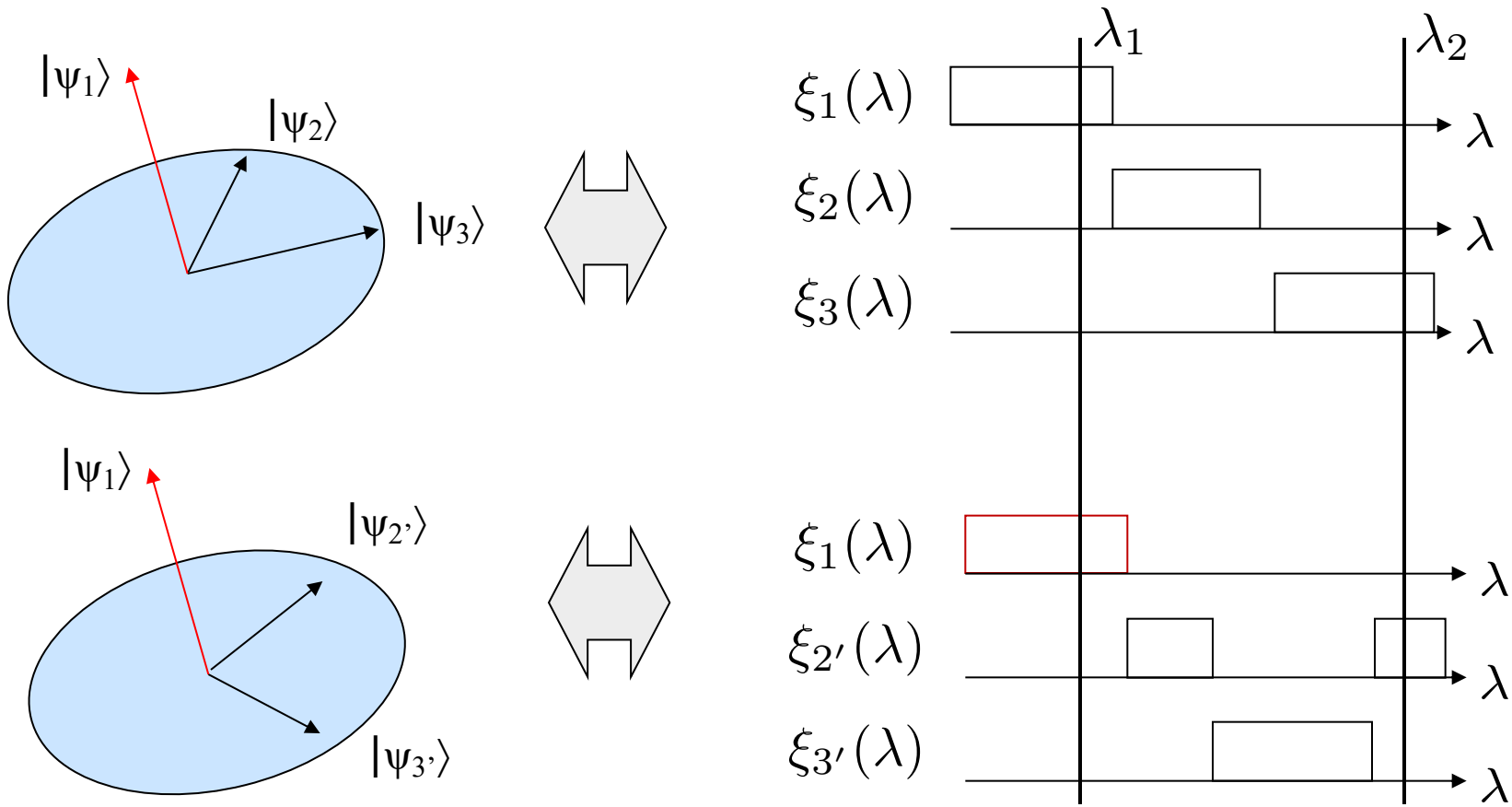
Kochen-Specker noncontextuality = MNC + outcome determinism (and it only applies to projective quantum measurements)





0/1-valued!

A given projector may appear in many different measurements

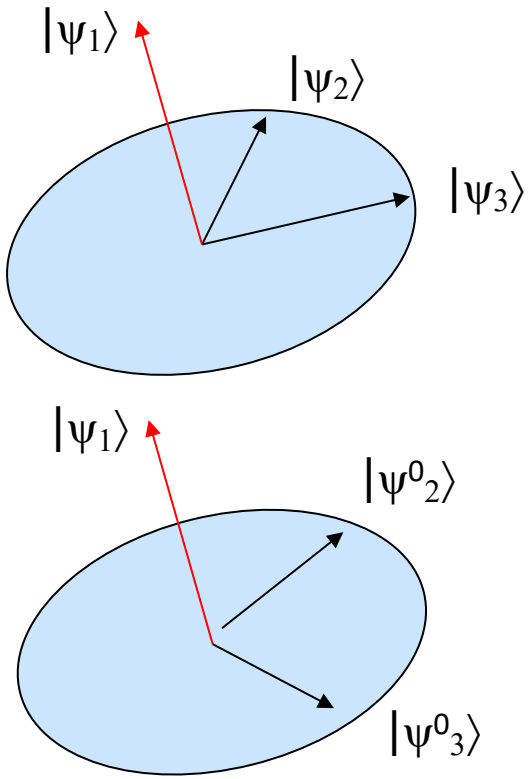


KSNC:
Every projector is associated with the same $\chi(\lambda)$ regardless of how it is measured (i.e. the context)

Bell-Kochen-Specker theorem: No hidden variable model that is noncontextual in this sense can reproduce the quantum predictions for Hilbert spaces of dimension 3 or greater.

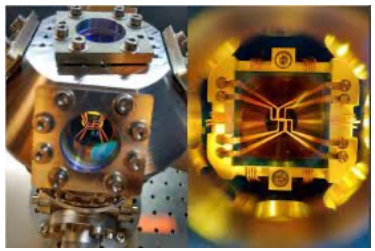
Is noncontextuality a reasonable assumption?

Incompatible mmts!

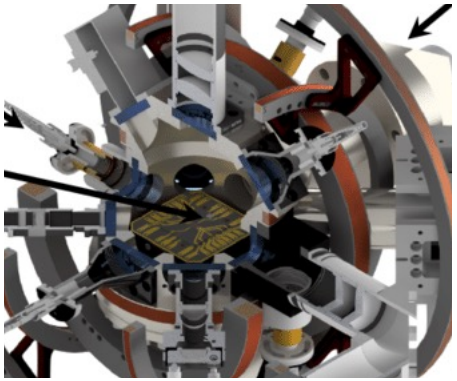


necessarily distinct experimental devices

M_1



M_2



So *why* should one expect the value assignments be the same?

"These different possibilities require different experimental arrangements; there is no a priori reason to believe that the results... should be the same. The result of an observation may reasonably depend not only on the state of the system (including hidden variables) but also on the complete disposition of the apparatus"

–Bell 1966

So *why* should one expect the value assignments be the same?

Rob Spekkens: because this is the best explanation for why the different measurement devices give rise to exactly the same possible statistics, for *all possible preparations*.

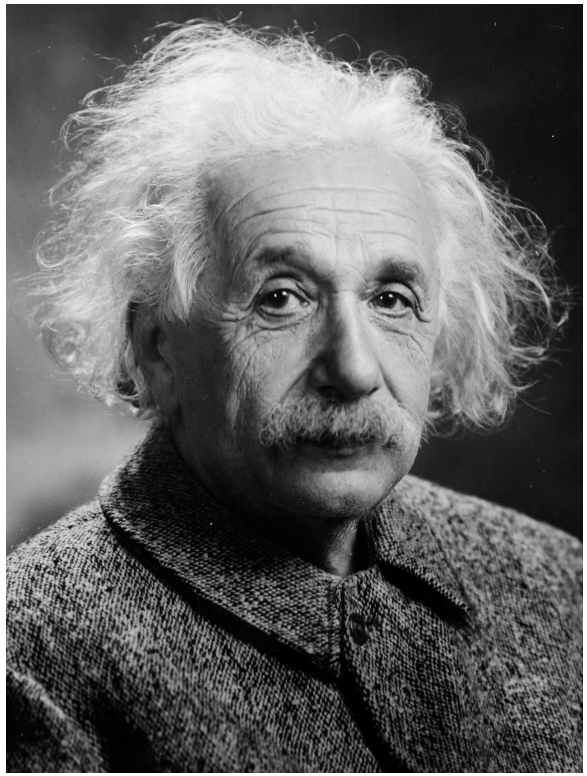
Leibniz's principle of the identity of indiscernibles—
if a difference in set-up is not distinguished in the observable phenomena then it should not be distinguished in the ontological picture either

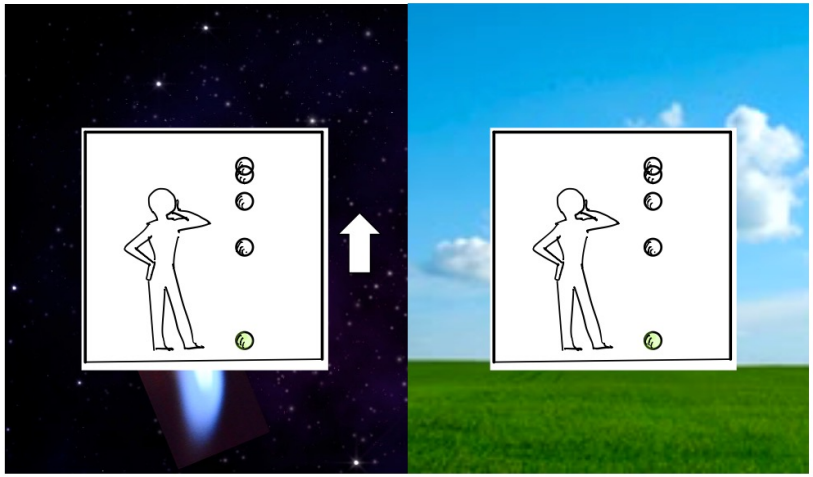
This is a methodological principle which guides us in constructing good physical theories

arxiv:1909.04628

Leibniz's principle in action

Einstein's arguments against the ether





Einstein's strong equivalence principle

arxiv:1909.04628

The justifications one can give for KSNC also imply the generalized notion of noncontextuality, which has a number of advantages:

-applies to noisy measurements
-applies preparations, transformations, etc
-does not assume determinism
-has an unparalleled pedigree as a notion of classicality

Next lecture: why noncontextuality is our best notion of classical explainability!

Generalized noncontextuality from the perspective of GPTs

David Schmid

Solstice of Foundations 2022

Why generalized noncontextuality is our best notion of classically-explainability:

a priori motivations

-equivalent to classical-explainability of GPTs
-equivalent to existence of a positive quasiprobability representation
-Leibniz's principle (identity of indiscernibles)

subsumes other notions

-Kochen-Specker contextuality
-Bell inequality violations
-anomalous weak values

key features

-applies to any experiment/scenario
-experimentally testable
-doesn't assume correctness of quantum theory
-classicality emerges under sufficient noise or coarse-graining or decoherence

contextuality is a resource

-computation
-communication
-cryptography
-state discrimination
-cloning
-metrology

In the landscape of generalized probabilistic theories (GPTs), which are “classical”?

(focus on P-M)

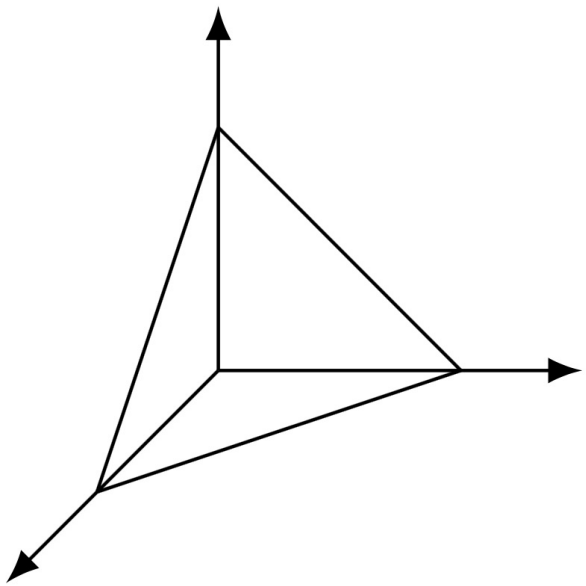
arXiv:1911.10386

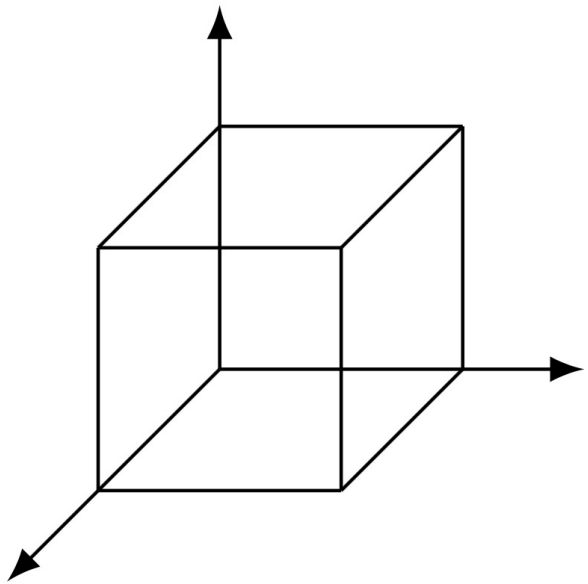
Traditionally, a GPT has been considered "classical" if it is *simplicial*:

state space: simplex

effect space: dual of simplex

d=3



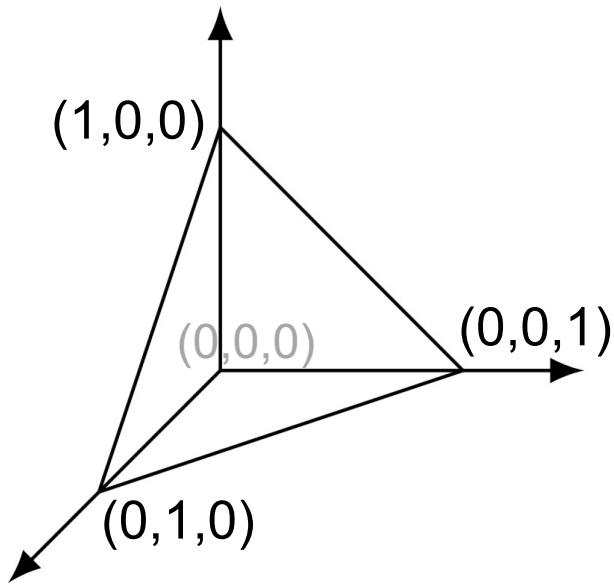


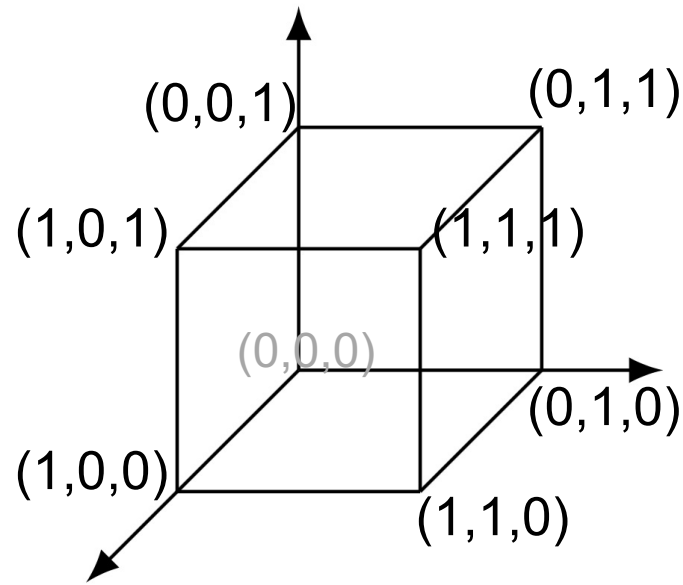
Traditionally, a GPT has been considered "classical" if it is *simplicial*:

state space: simplex

effect space: dual of simplex

d=3

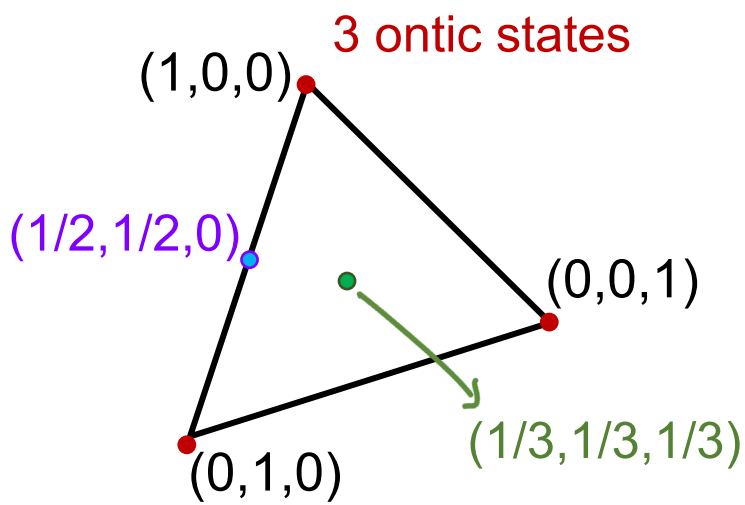


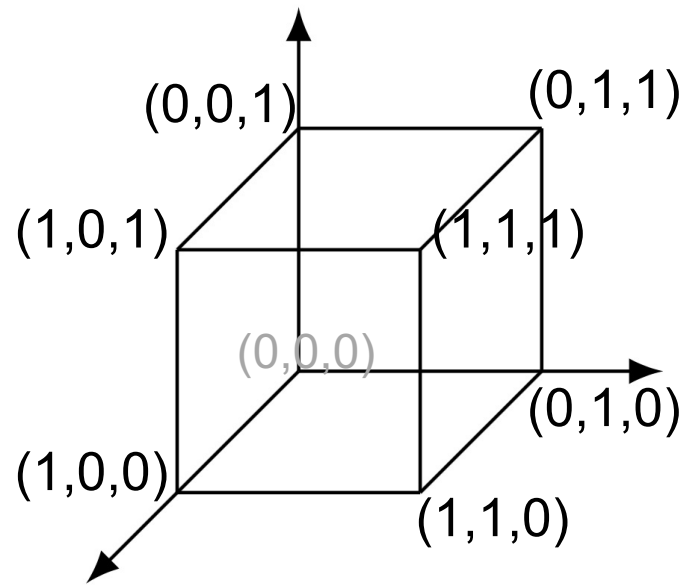


normalized states

effects

d=3



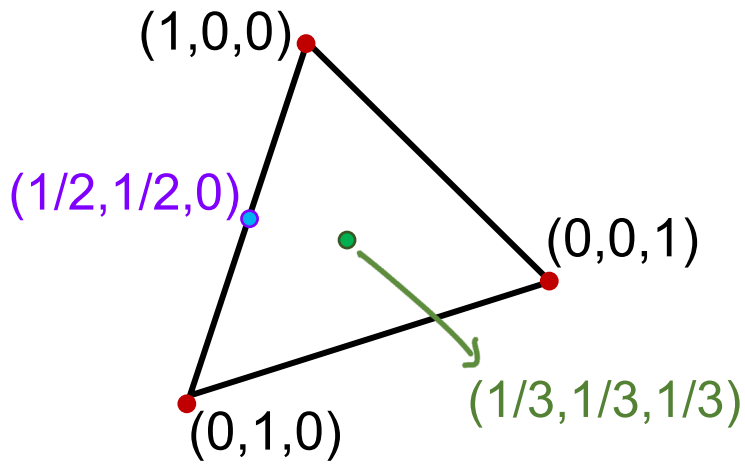


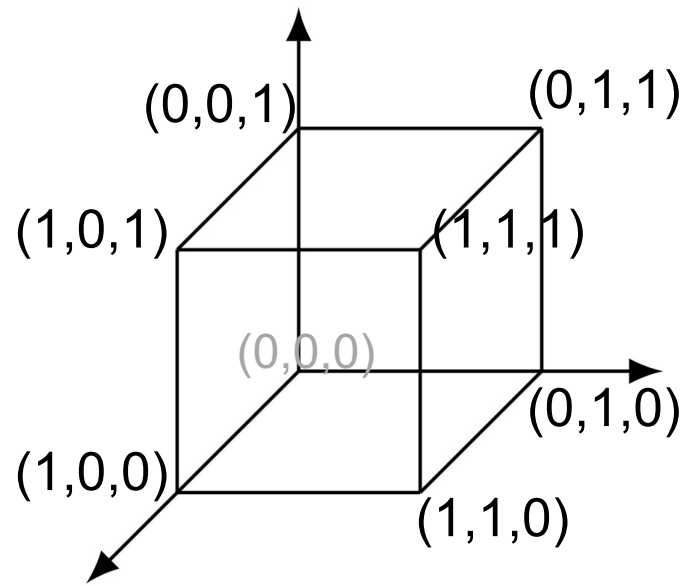
Classical statistical theory: probability distributions over a set of classical states

normalized states

effects

d=3



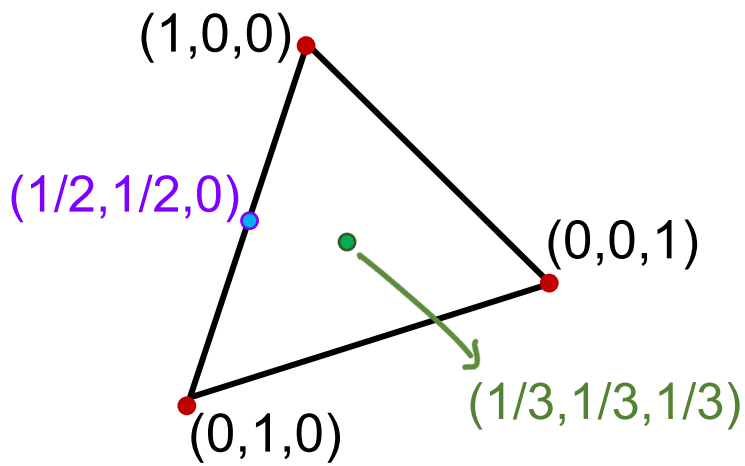


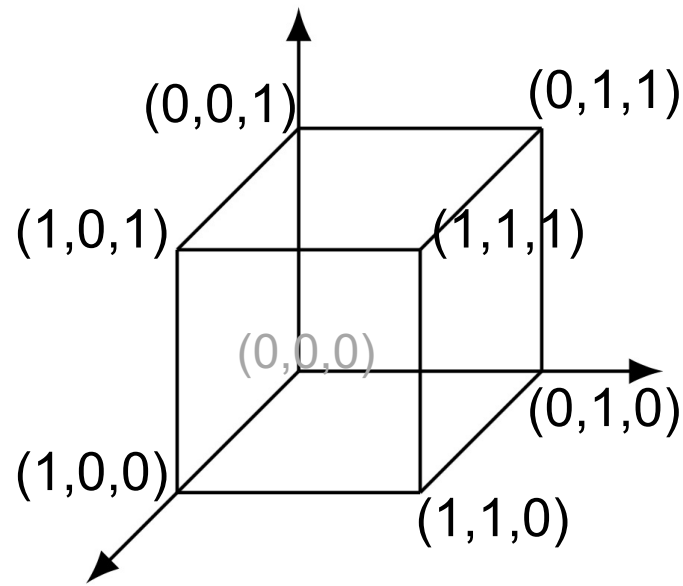
Every mixed state decomposes into pure states in a unique way
⇒ One can always imagine that there is a true state of the system, and any mixed state can be *uniquely* interpreted as uncertainty about the true state.

normalized states

effects

d=3



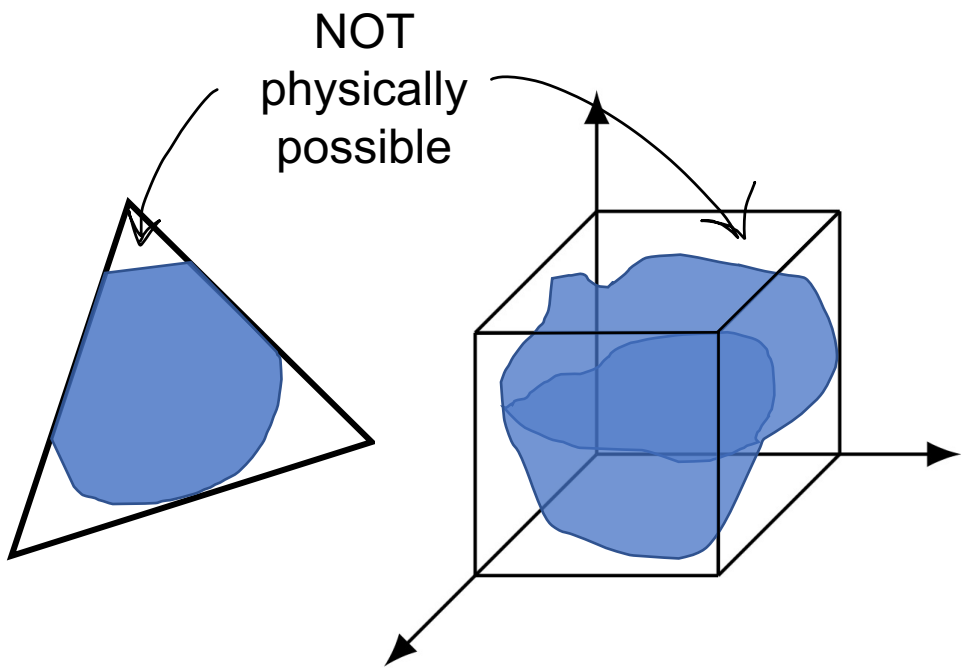


All logically possible measurements are *physically possible* and *compatible*.
⇒ One can determine the exact state of the system in a single measurement.

$$\text{simplicial} \Longrightarrow \text{classical}$$

But what about subtheories of a simplicial theory?

that is, GPTs whose states and effects are a strict subset of those from a simplicial GPT



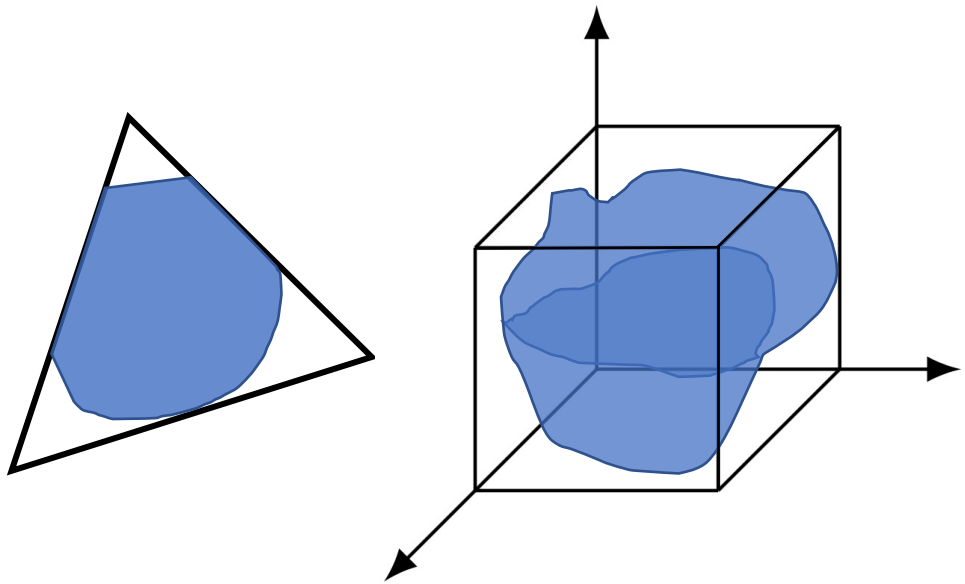
intuition: these are classically-explainable as well

-one can still imagine that the vertices correspond to ontic states
-so the ontology is the same as in a simplicial theory
-it is just that one cannot perfectly prepare/measure/know the ontic state

Formal definition

A GPT is simplex-embeddable if there exists

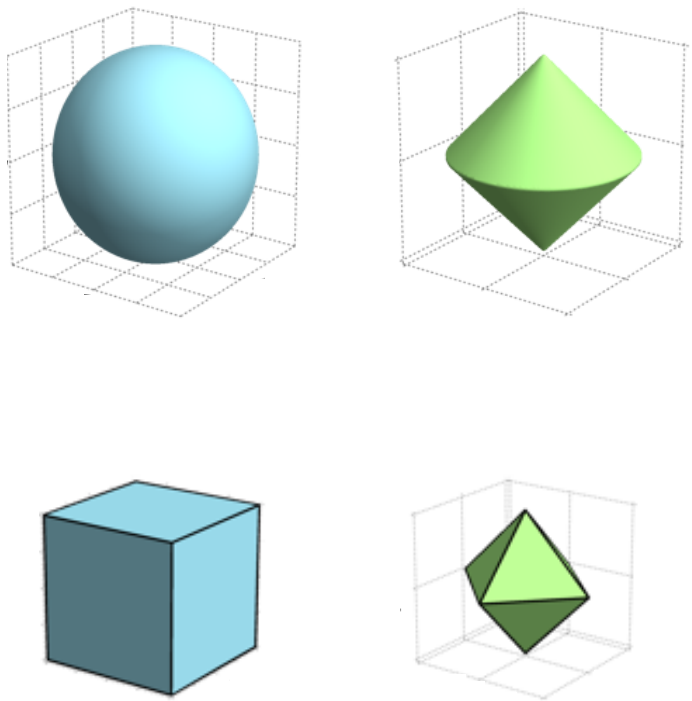
1) a linear map taking states into a simplex, and
2) a linear map taking effects into the dual to that simplex, such that
3) probabilities are preserved



$$\text{simplicial} \Longrightarrow \text{classical}$$

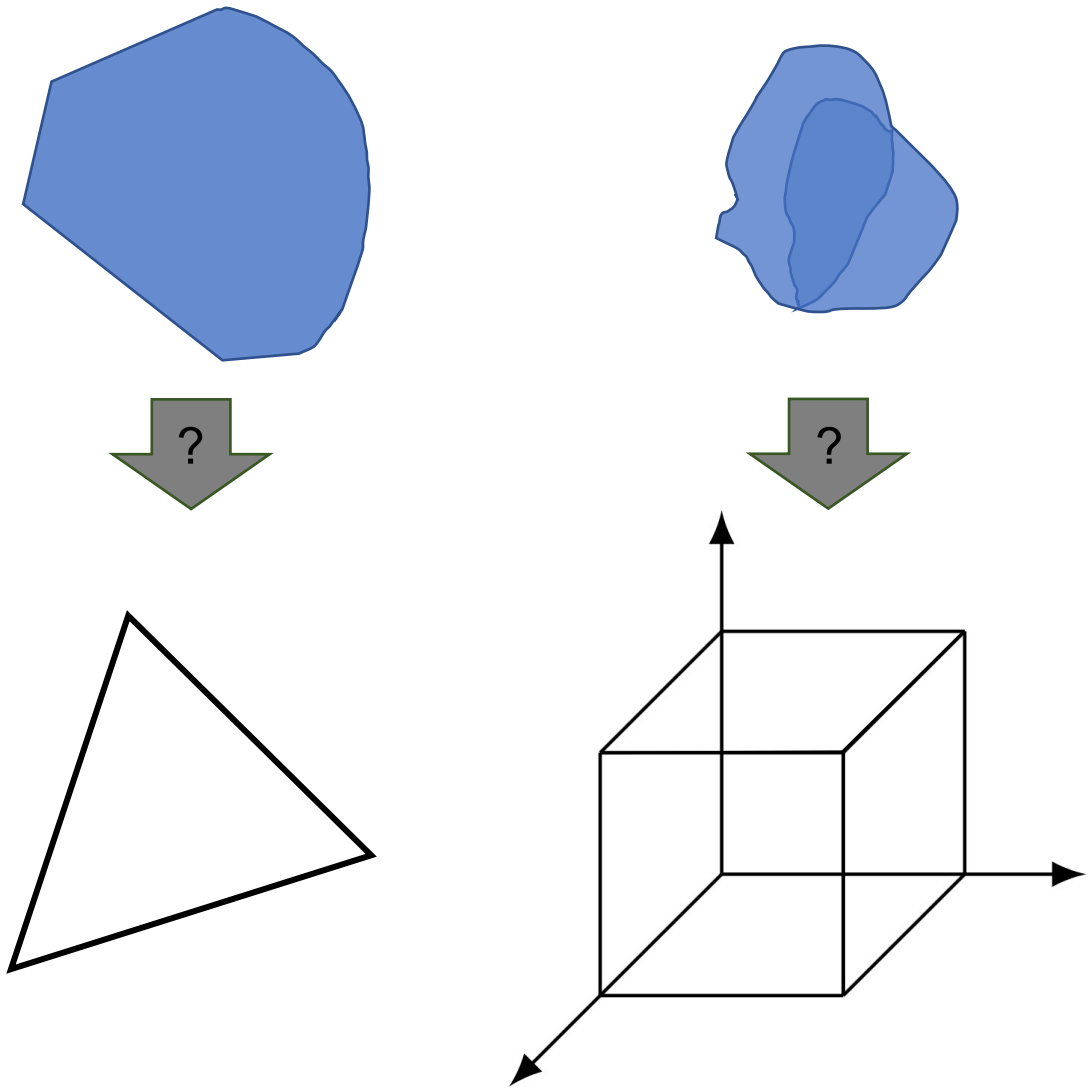
$$\text{simplex-embeddable} \Longleftrightarrow \text{classically explainable}$$

Which GPTs are simplex-embeddable?



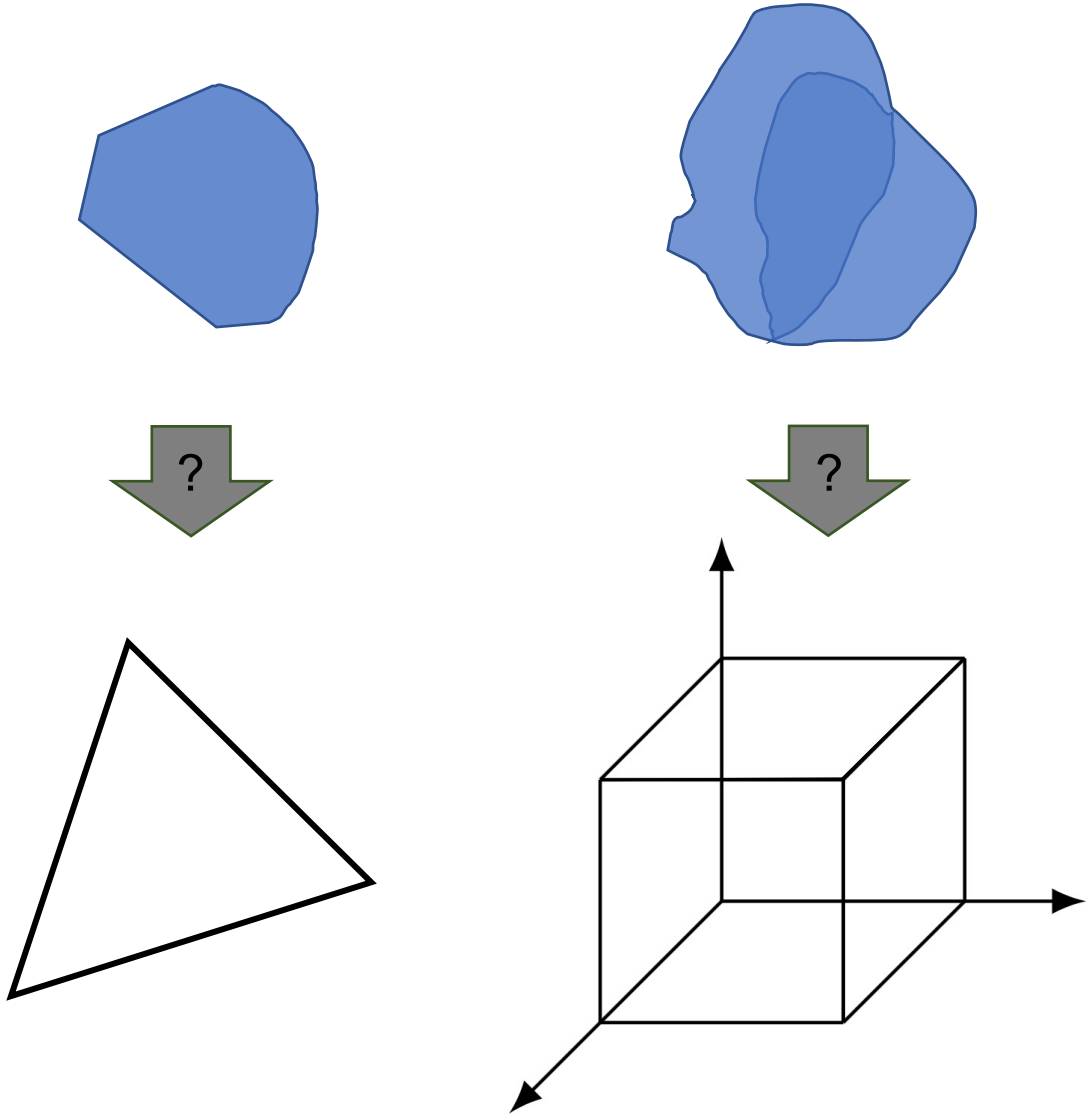
No-go theorems for classical-explainability of GPTs!

Which GPTs are simplex-embeddable?



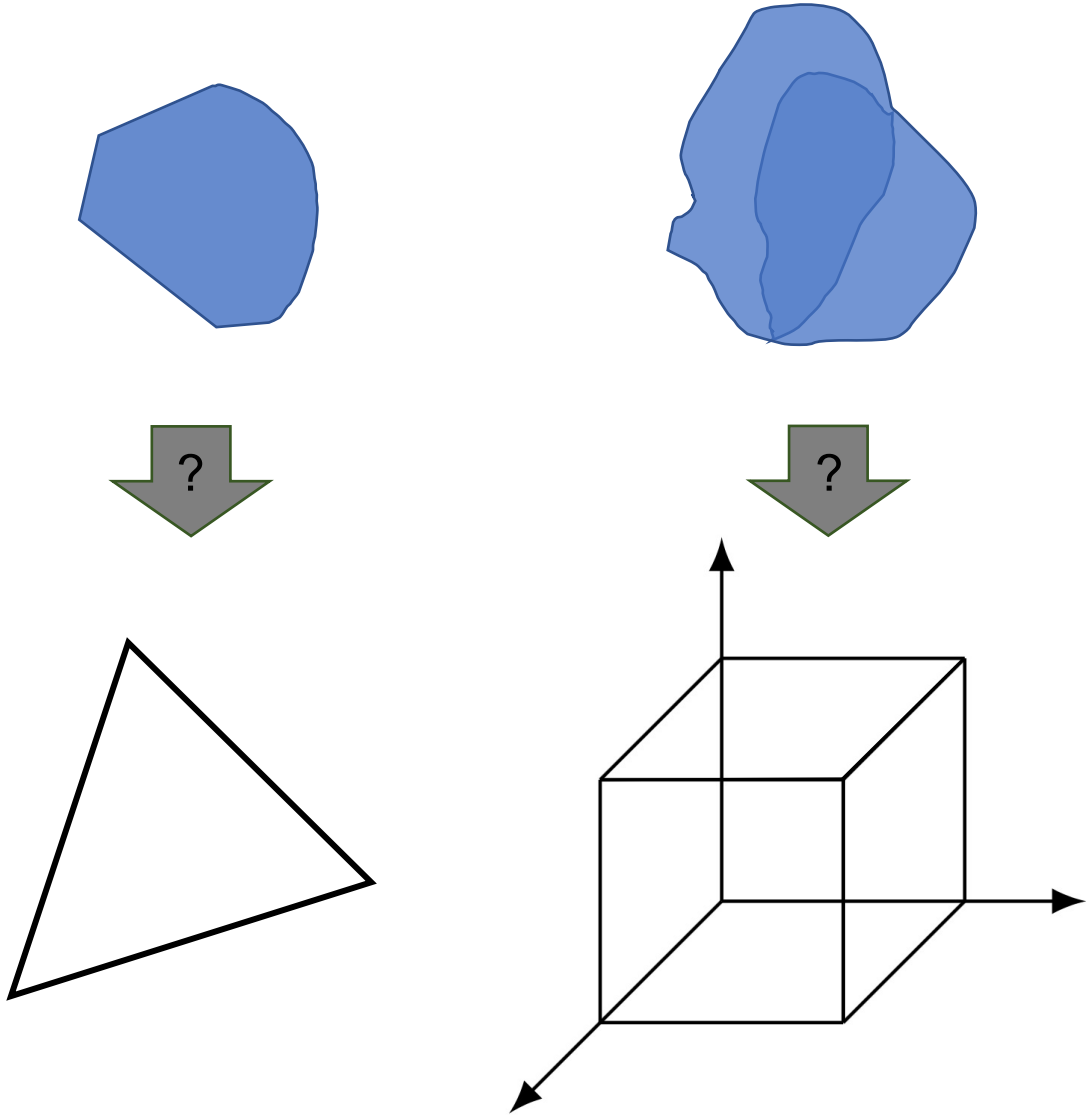
-may need to jointly rescale states and effects

Which GPTs are simplex-embeddable?



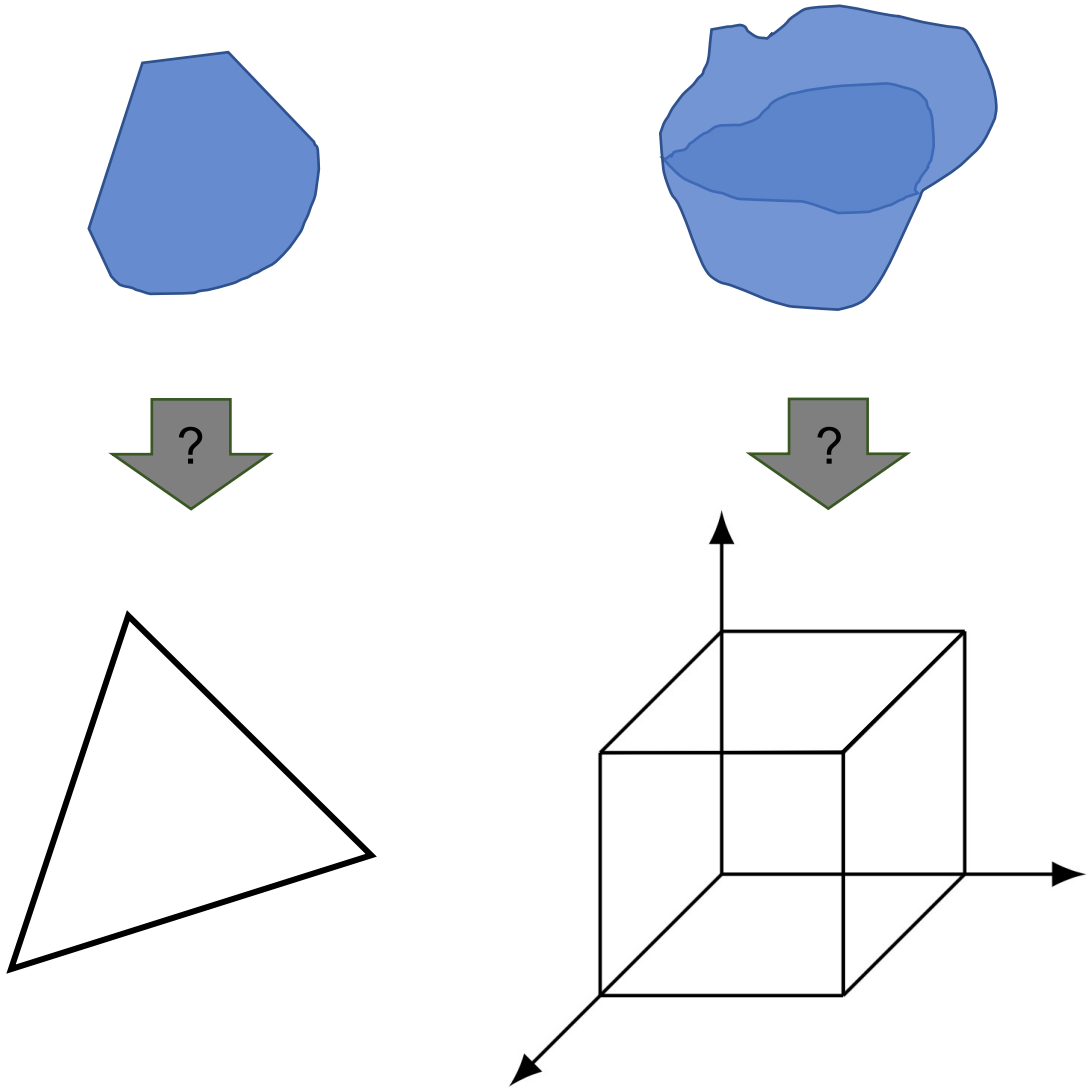
-may need to jointly rescale states and effects

Which GPTs are simplex-embeddable?



-may need to jointly reorient states and effects relative to the simplex

Which GPTs are simplex-embeddable?



-may need to jointly reorient states and effects relative to the simplex

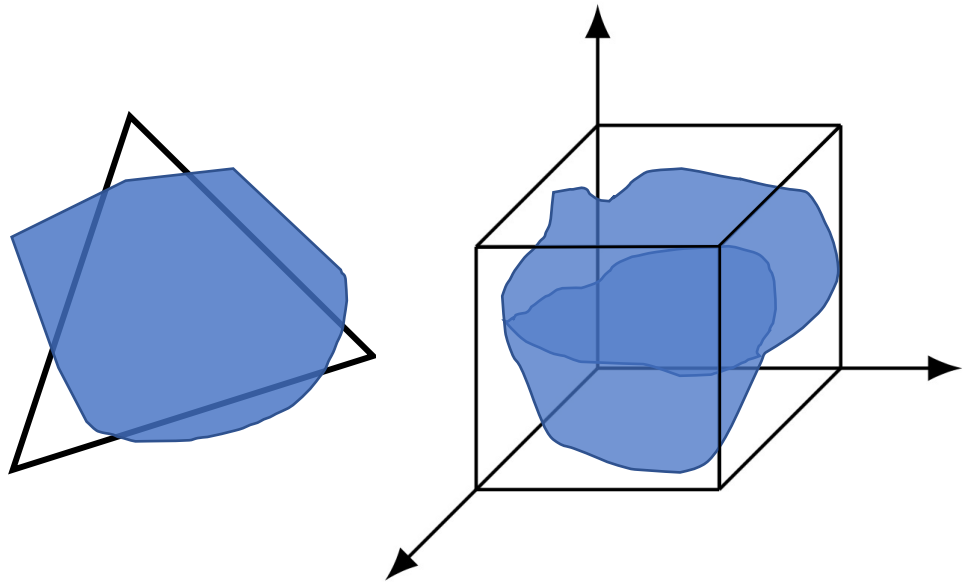
(in fact any linear transformation can be applied to the states if the inverse is applied to the effects)

-dimension of simplex may be greater than GPT dimension!
-…but not greater than GPT dimension squared

Deciding if a GPT is simplex-embeddable is just a linear program!

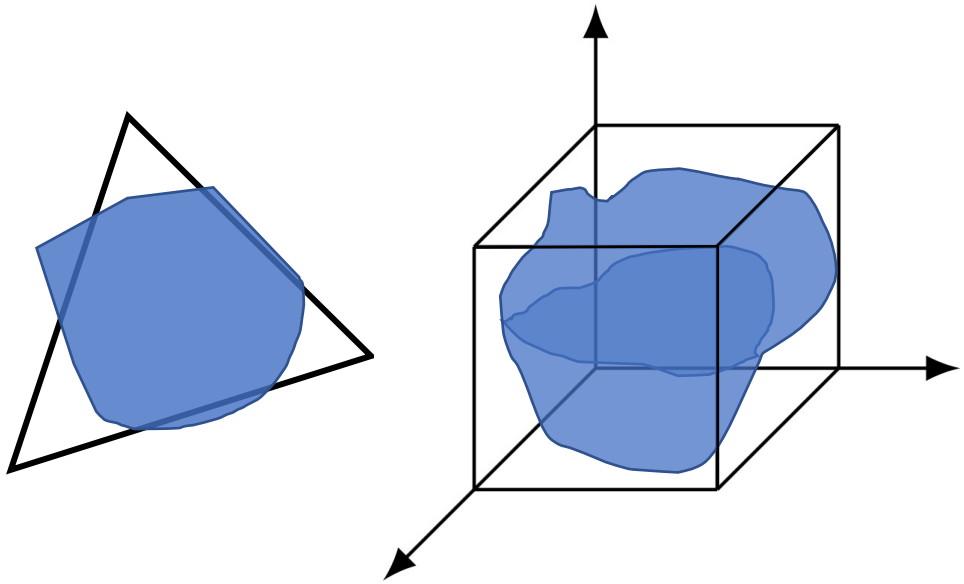
arXiv:2204.11905

It is clear geometrically that every GPT becomes simplex-embeddable under sufficient depolarizing noise.



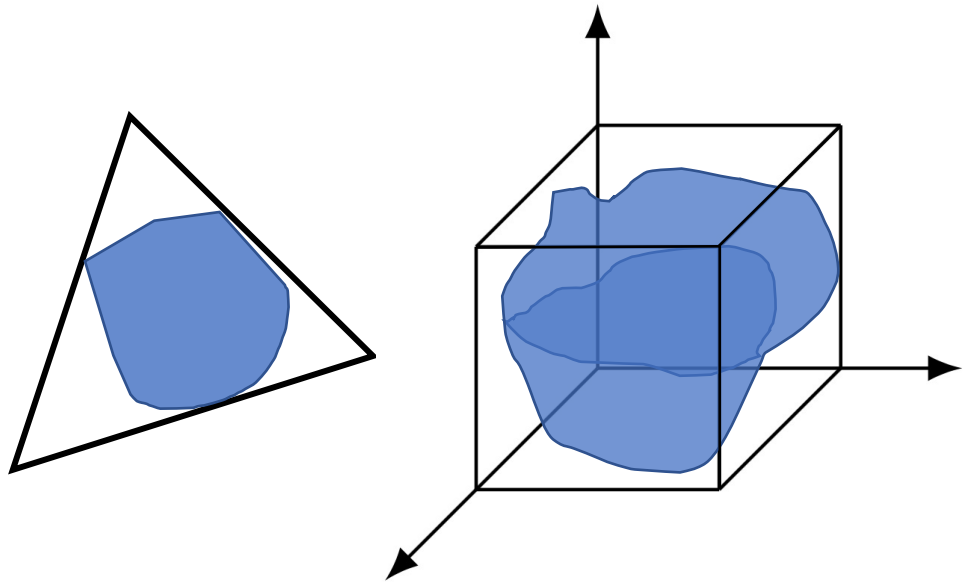
arXiv:2204.11905

It is clear geometrically that every GPT becomes simplex-embeddable under sufficient depolarizing noise.



arXiv:2204.11905

It is clear geometrically that every GPT becomes simplex-embeddable under sufficient depolarizing noise.



A measure of nonclassicality of a GPT: how much depolarization can it undergo before it becomes classical?

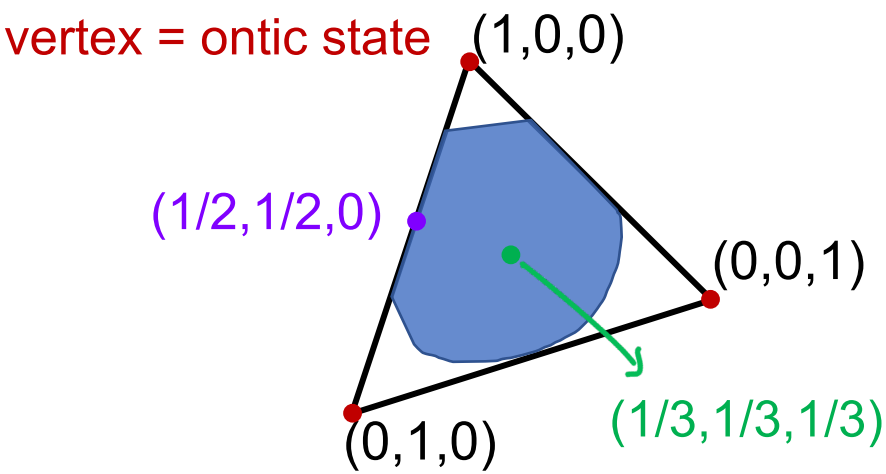
arXiv:2204.11905

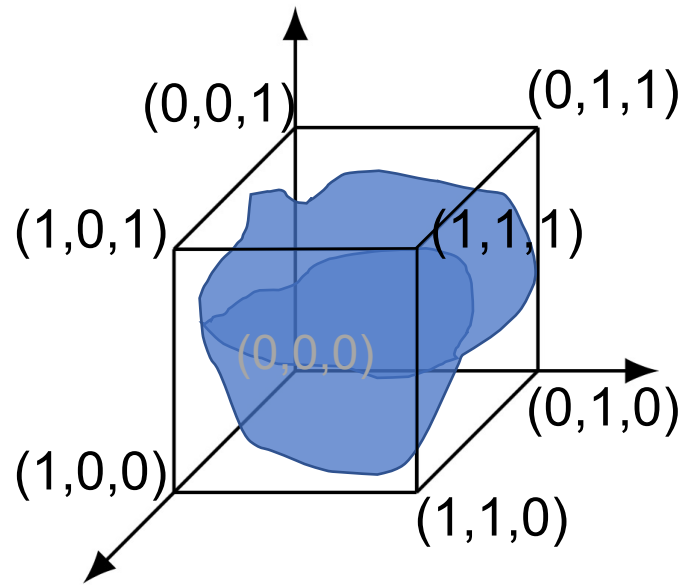
Equivalence to noncontextuality

arXiv:1911.10386

A simplex-embedding of a GPT is just an ontological model for it!

A simplex-embedding of a GPT is just an ontological model for it!





the epistemic state for a given GPT state is just the vector of its coordinates

the response function for a given GPT effect is just the vector of coordinates

and the predictions are reproduced as expected: $e \cdot s = \sum_{\lambda} \mu(\lambda)\xi(\lambda)$

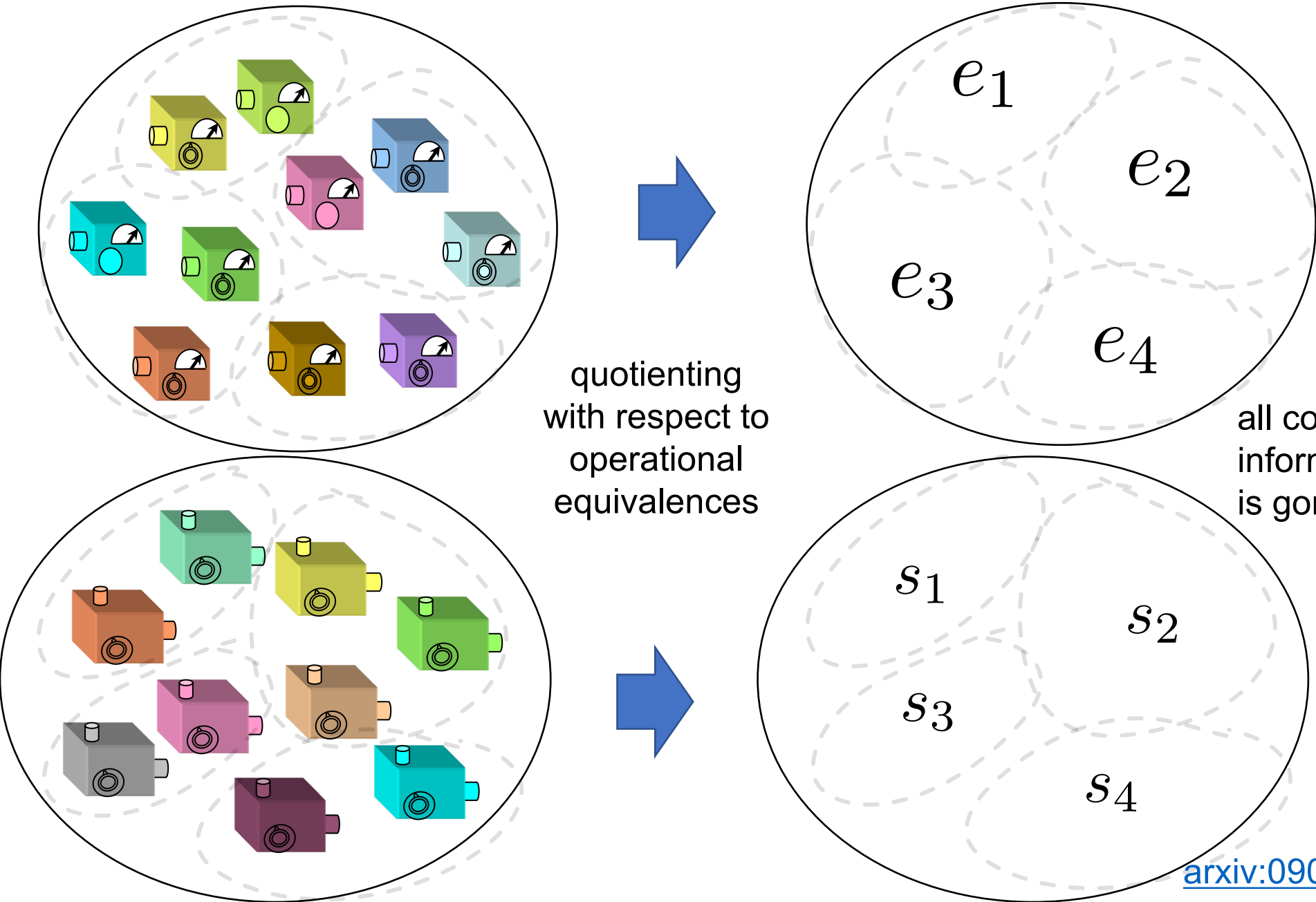
So a GPT is simplex-embeddable IFF it admits of **an ontological model.**

A GPT admits of an ontological model if and only if the operational theory which is described by the GPT admits of a noncontextual ontological model.

(This is clear once one understands the relationship between operational theories and GPTs.)

(unquotiented) operational theory

quotiented operational theory = GPT



all context
information
is gone!

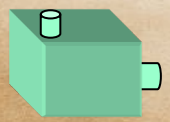
arxiv:0908.1583

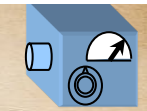
Quantum example

Unquotiented description:

-detailed lists of lab instructions
-includes context info

1. flip a coin
2. prepare 0 if heads, 1 if tails
3. forget the coin flip outcome





1. pass through a polarizer
2. check if photocell clicks

Quotiented description:

-density ops, POVMs
-no context info

quotient

ontological representation

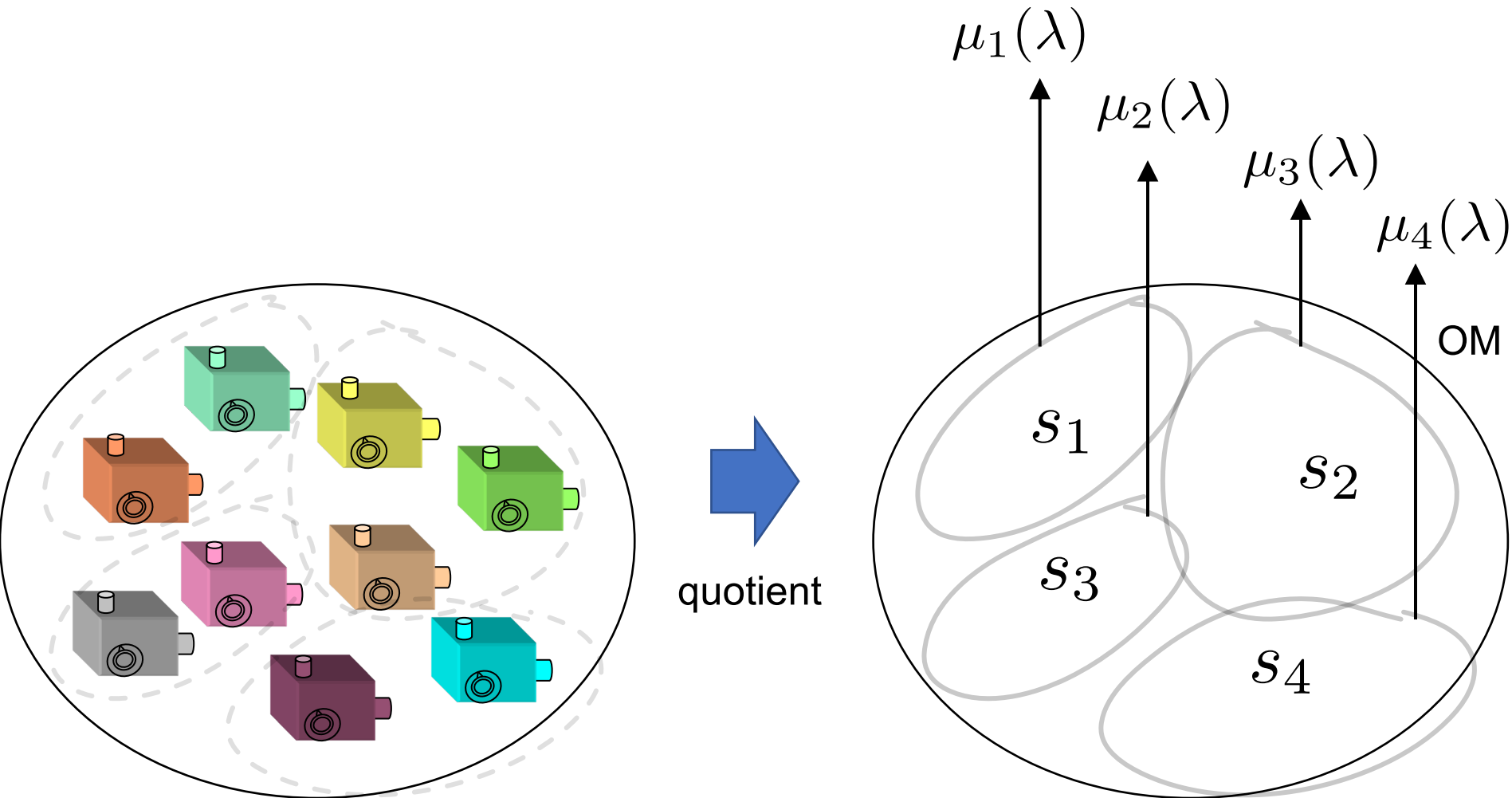
$$\rho \quad \mu(\lambda)$$

$$E \quad \xi(\lambda)$$

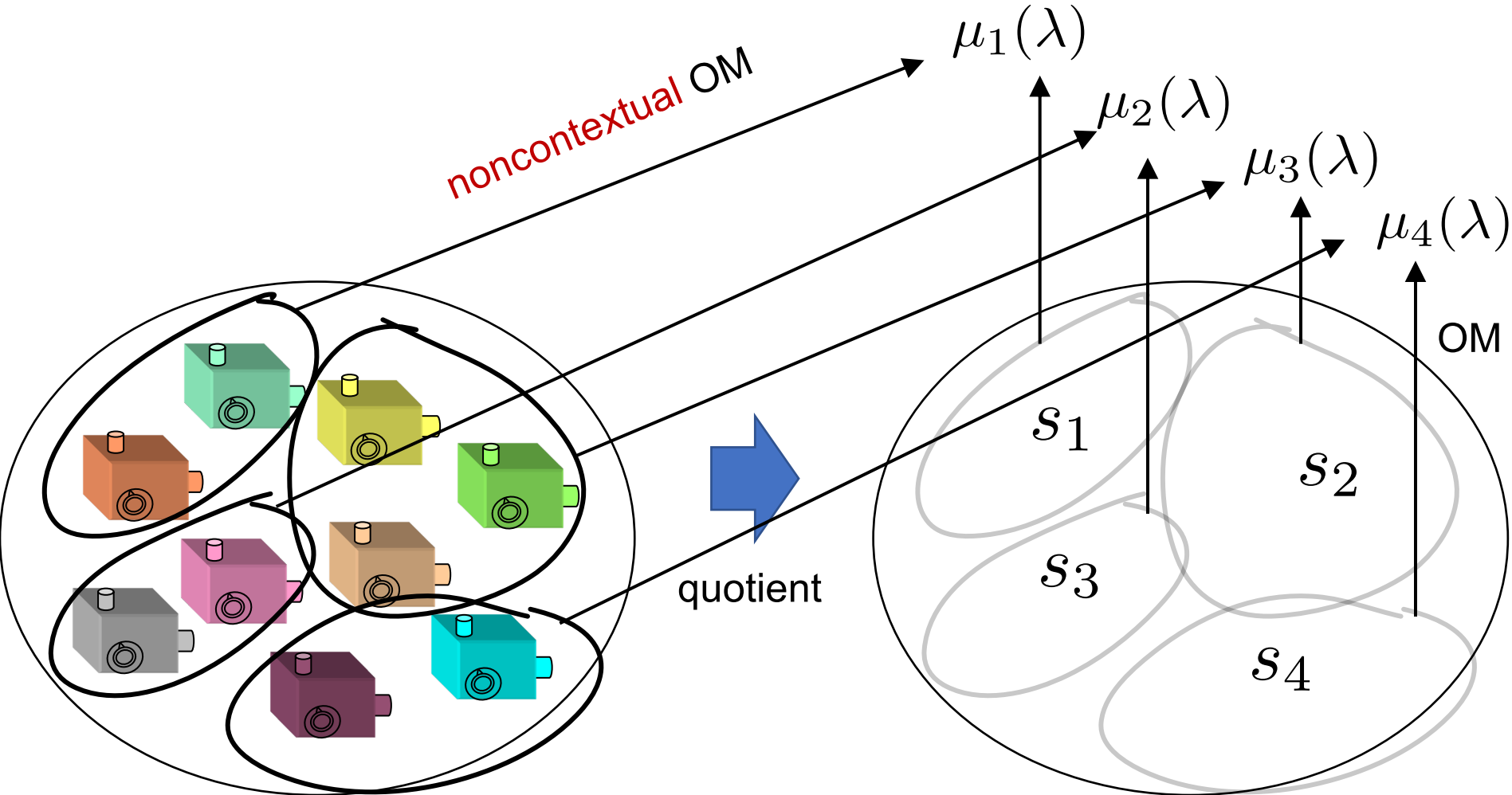
quotient plus ontological repn

necessarily defines a NONCONTEXTUAL model!

Any ontological model μ for a GPT gives a <u>noncontextual</u> ontological model for the unquotiented theory which it came from



Any ontological model μ for a GPT gives a <u>noncontextual</u> ontological model for the unquotiented theory which it came from



In summary: consider an operational theory and the GPT associated to it (by quotienting with respect to operational equivalences)

The operational scenario is classically-explainable if and only if

1. it admits of a noncontextual ontological model
2. its GPT admits of any ontological model
3. its GPT admits of a simplex-embedding

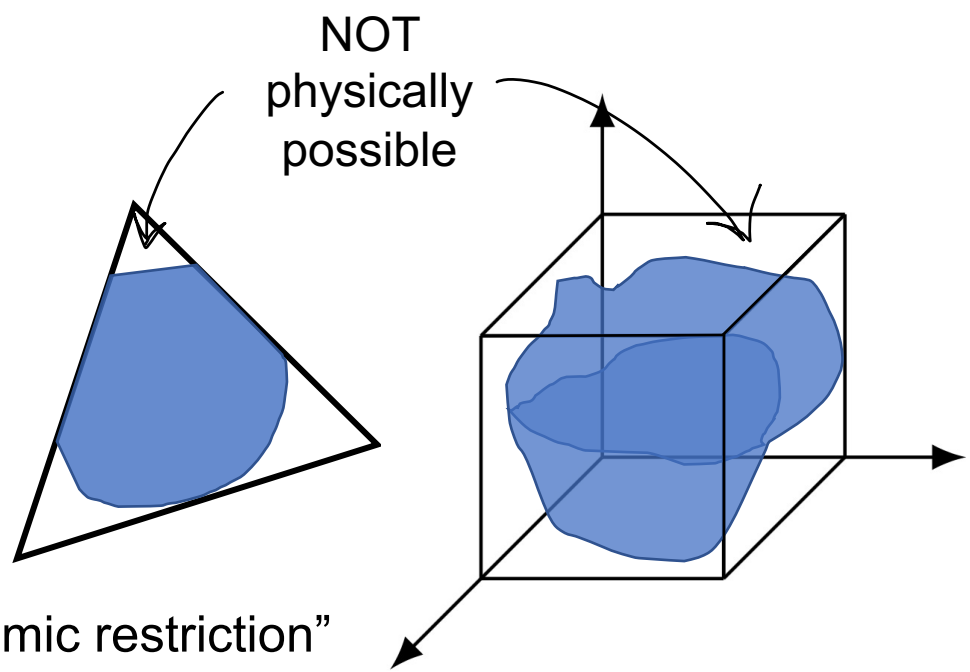
These are equivalent notions of classicality!

arXiv:1911.10386
arxiv:2005.07161

Simplex-embeddable GPT

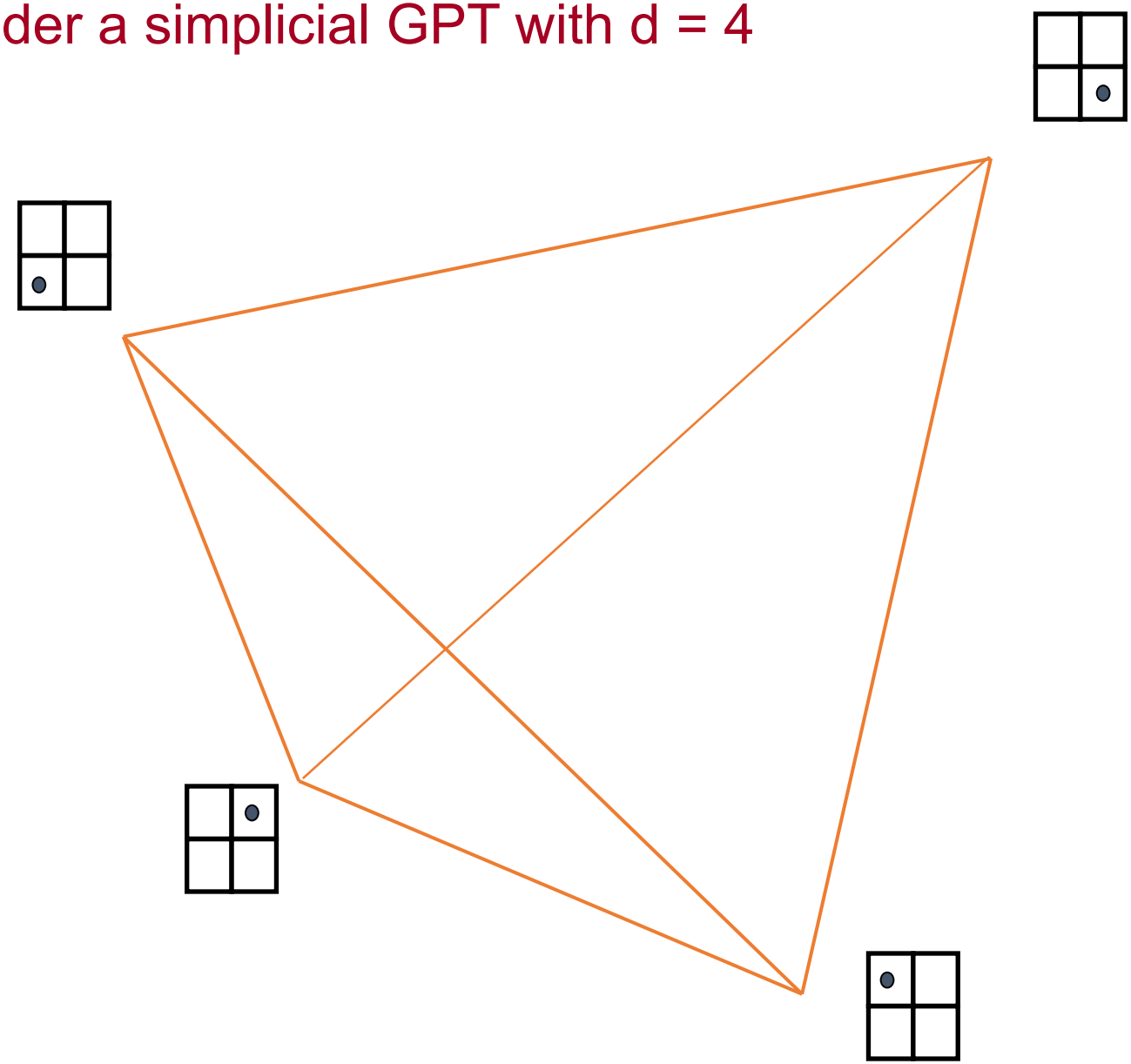
=

Epistemically restricted classical theory



ruled out by an “epistemic restriction”

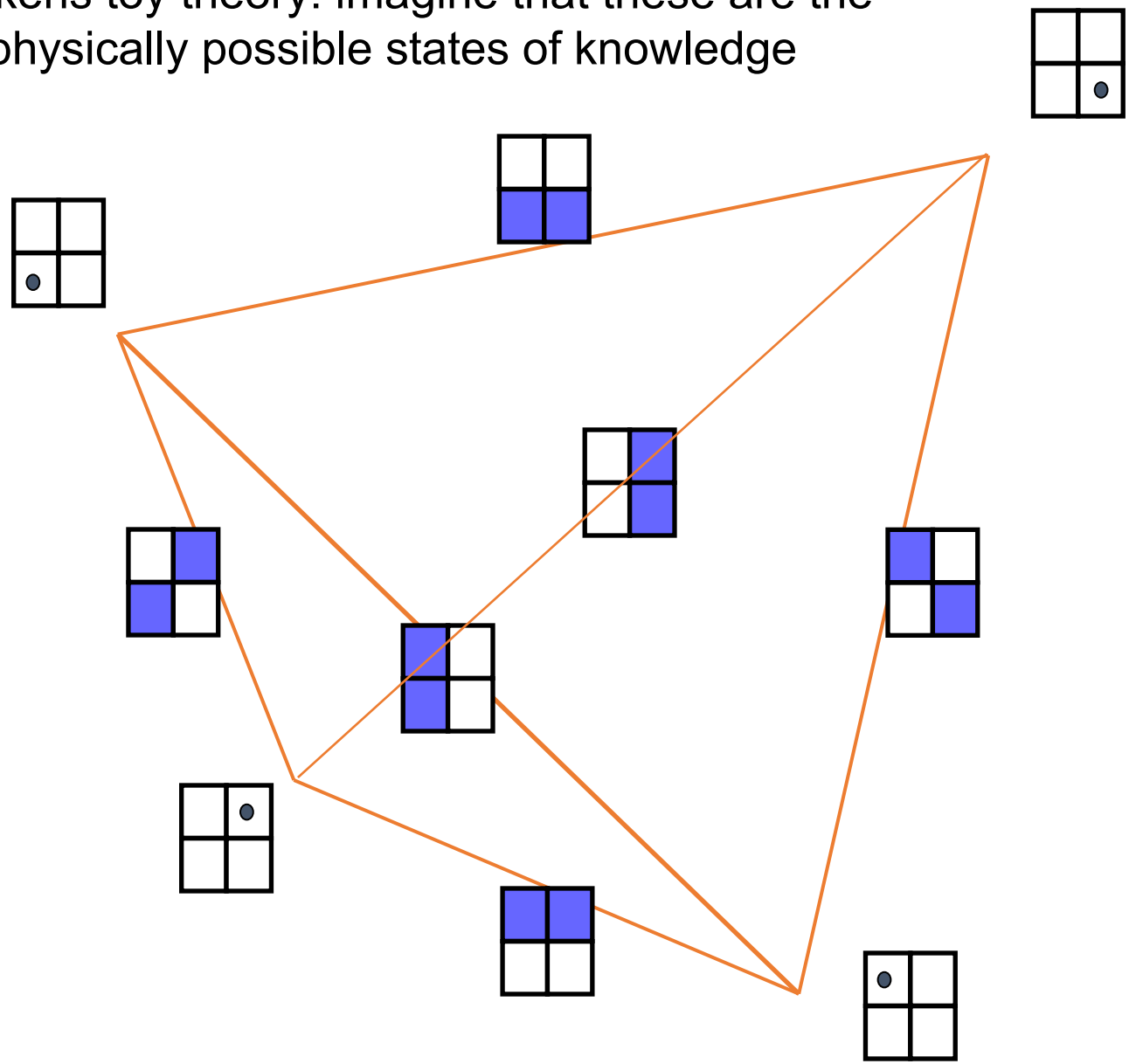
Consider a simplicial GPT with d = 4



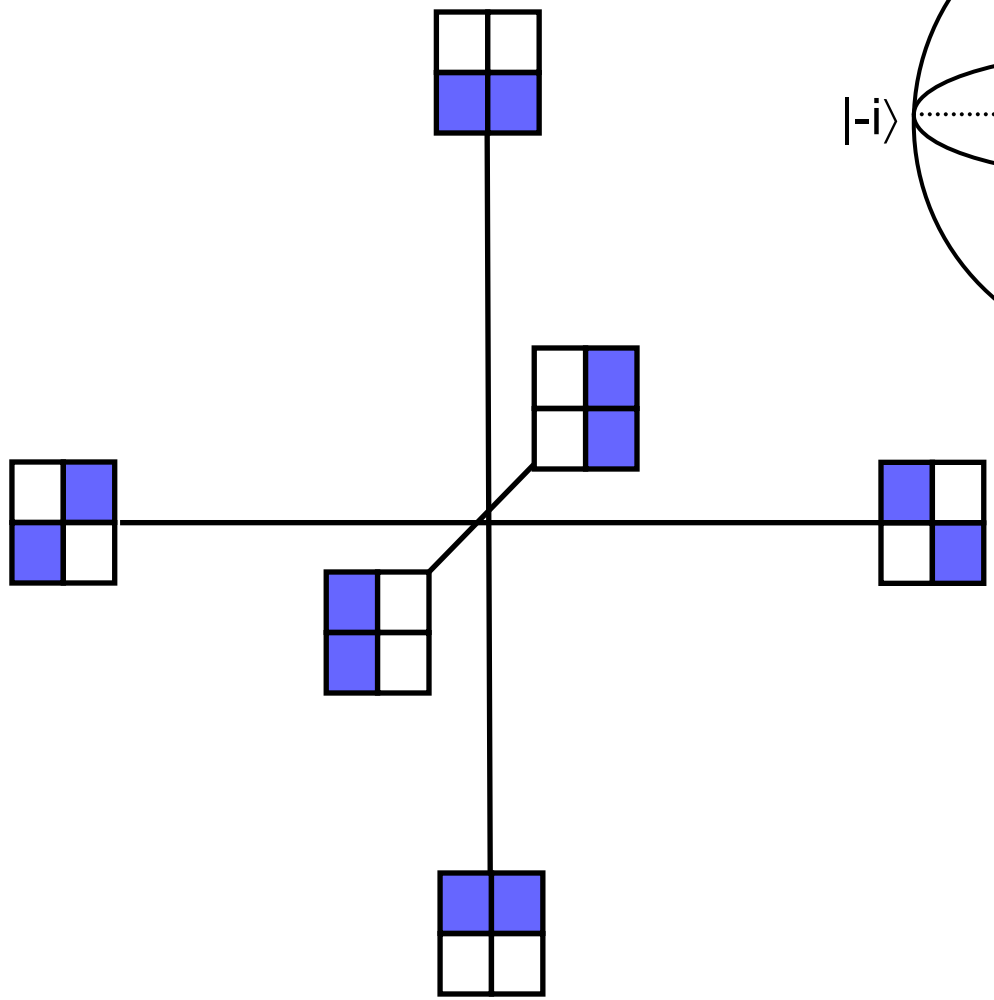
points in the simplex correspond to valid probability distributions

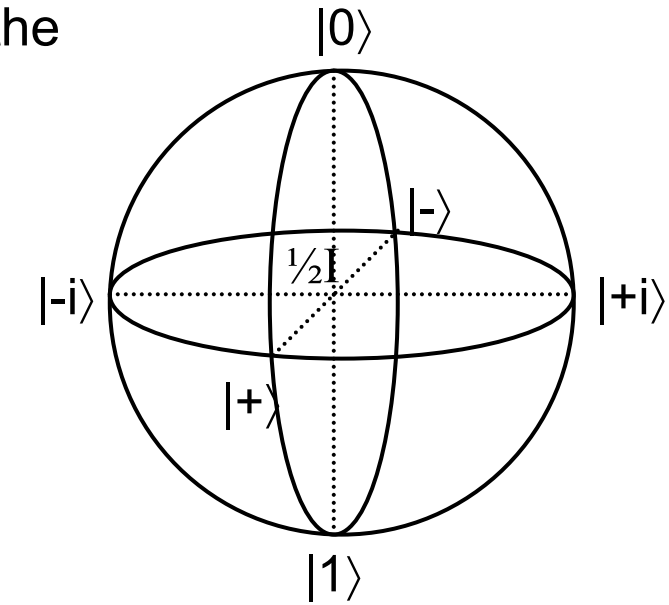
arXiv:0401052

Spekkens toy theory: imagine that these are the only physically possible states of knowledge

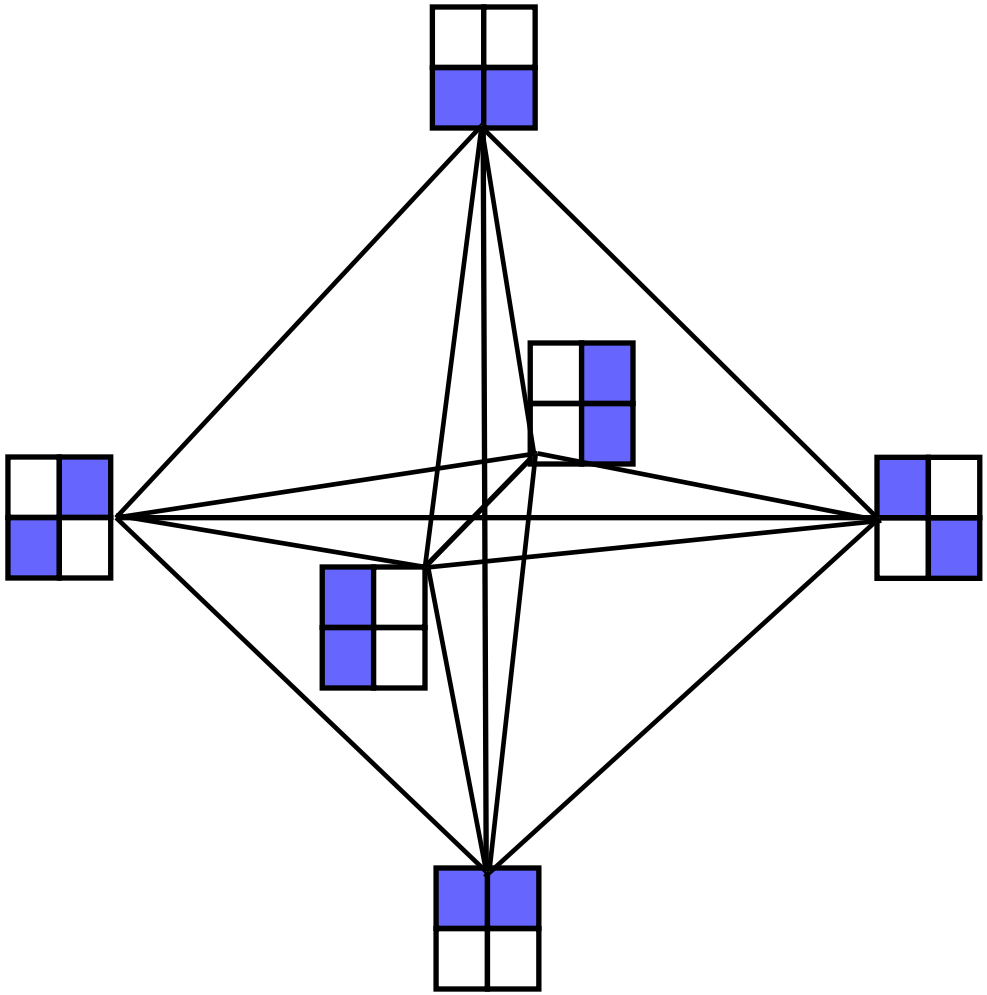


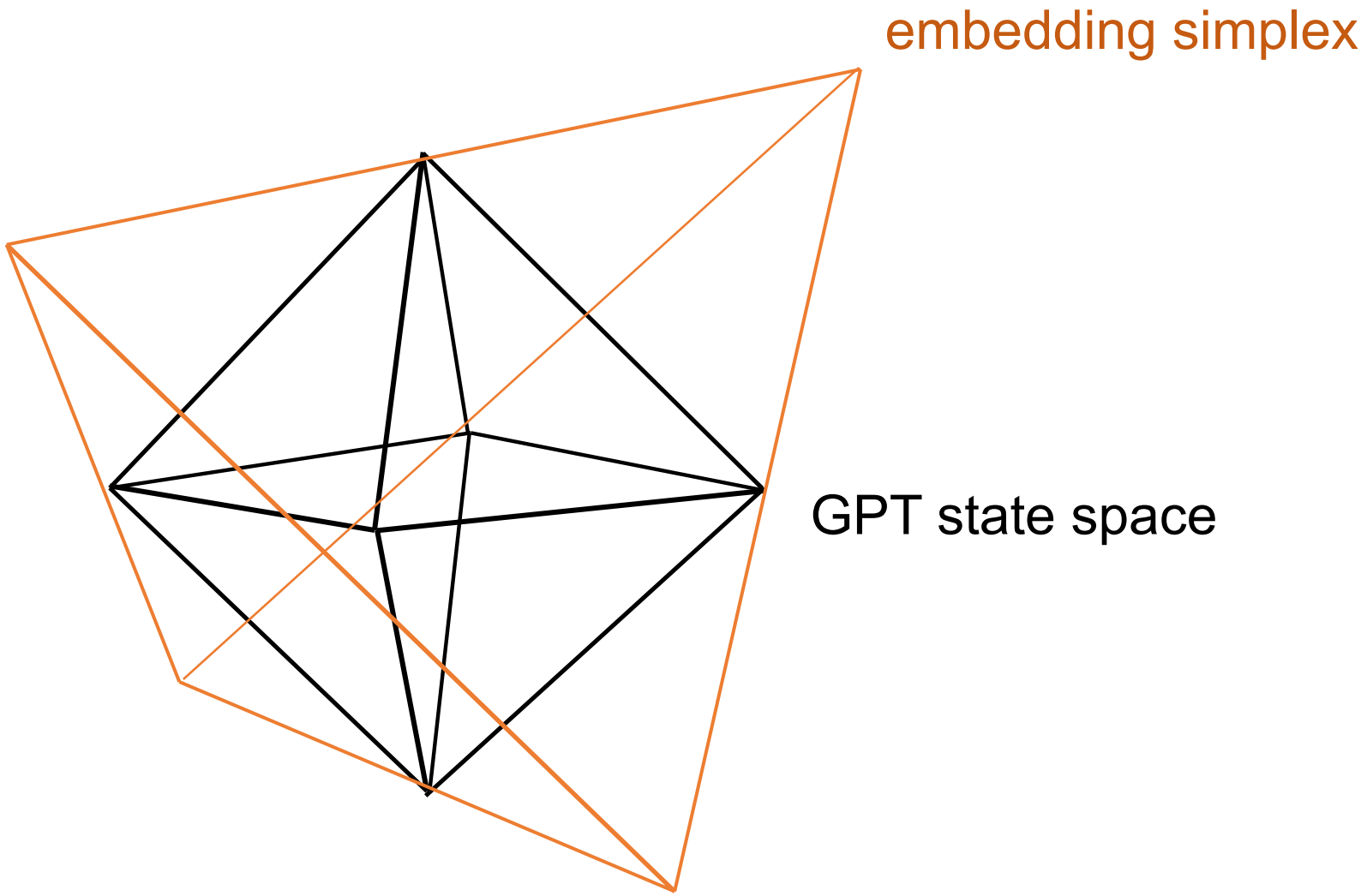
Spekkens toy theory: imagine that these are the only physically possible states of knowledge





As a GPT, we also include all mixtures of these:

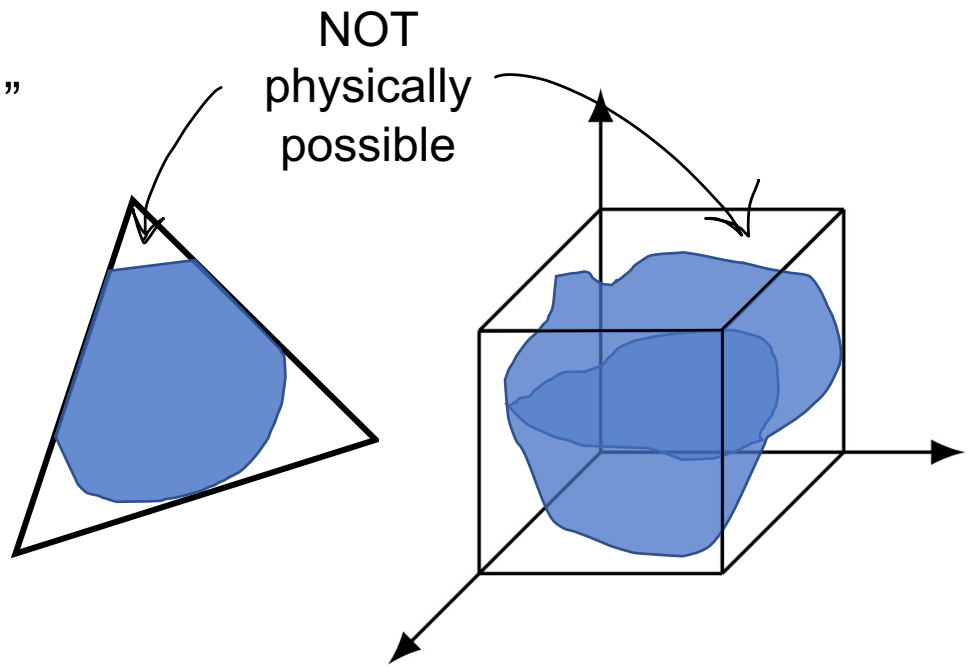




(the effect space is constructed similarly)

Any simplex-embeddable GPT can be viewed similarly:

ruled out by an “epistemic restriction”



What is the scope of possible Spekkens-like toy theories?
Answer: all and only noncontextual models

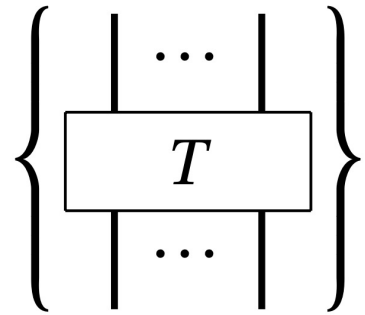
Simplicial GPTs	Simplex-embeddable GPTs	Non-embeddable GPTs
simplex + dual	simplex + dual + restriction	all other GPTs
“noncontextual”	“noncontextual”	“contextual”
strictly classical	classically-explainable	nonclassical
-all mmts compatible -unique decomposition of mixed states	-noncommutativity -complementarity -interference -no-cloning -teleportation -dense coding -entanglement -remote steering -quantum eraser -mmts must disturb -ambiguity of mixtures -no perfect state discr. …	-contextuality -computational speedups -nonlocality -specific *quantitative* advantages (see next lecture!)

Noncontextuality beyond PM scenarios

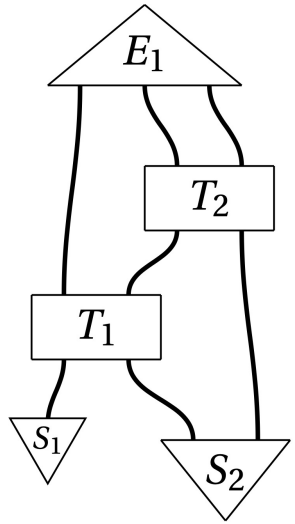
arxiv:2005.07161

Operational Theory = process theory + probability rule

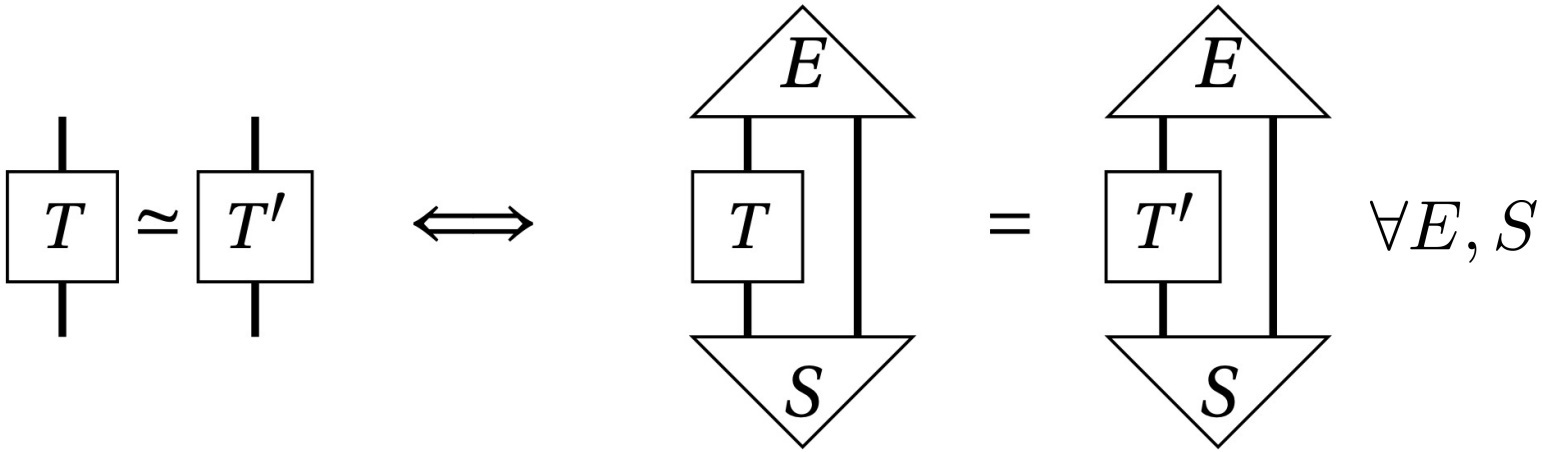
-systems

-laboratory procedures applied to systems 

-probability rule

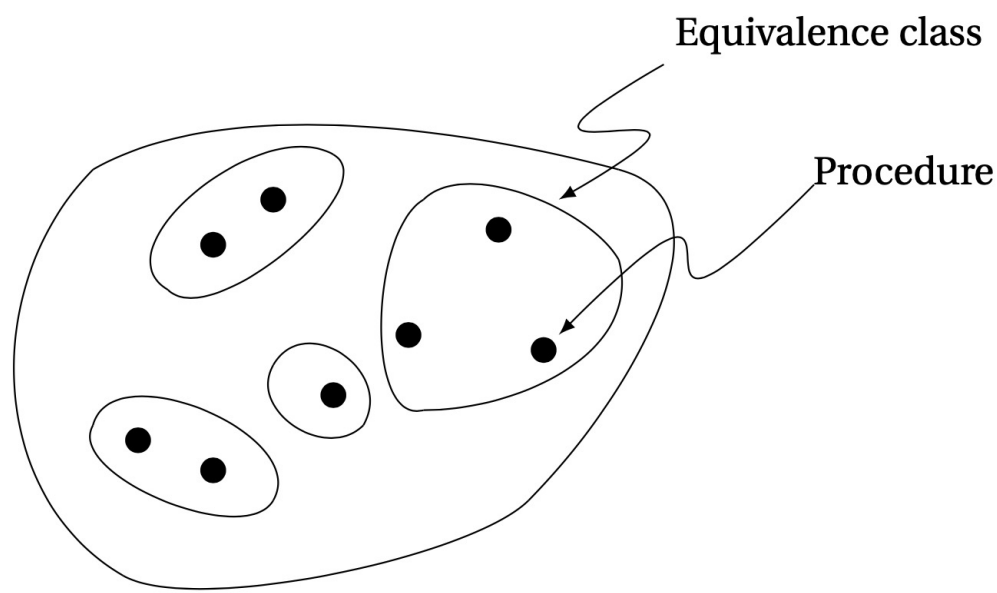
pr(E_1) = 

Operational Equivalence of Two Procedures



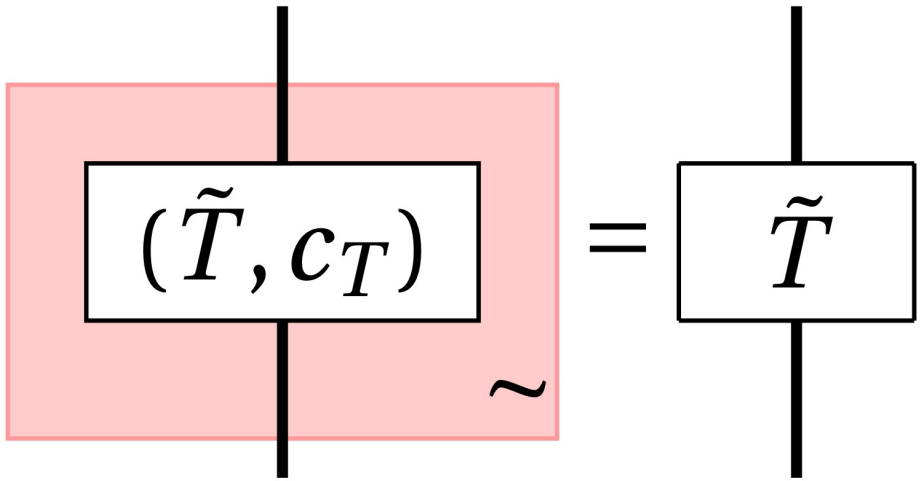
$$(\tilde{T}, c_T)$$

Equivalence Class Contexts



Quotienting with respect to Operational Equivalence

define a quotienting map $\sim :: (\tilde{T}, c_T) \longmapsto \tilde{T}$



Operational theory ➡ GPT

Recall that there is a NCOM of an operational scenario IFF there is an OM of the GPT.

so henceforth we will just talk about OMs of GPTs!

Ontological representations for arbitrary circuits in GPTs

arxiv:2005.07161

Ontological model of a GPT

$$\Gamma : GPT \rightarrow SubStoch$$

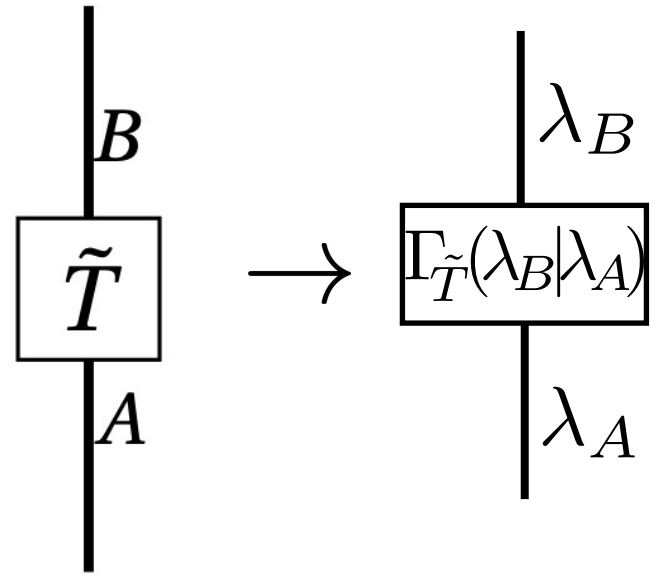
which:

1) reproduces the predictions

2) represents ignoring appropriately

3) is diagram-preserving

4) is linear



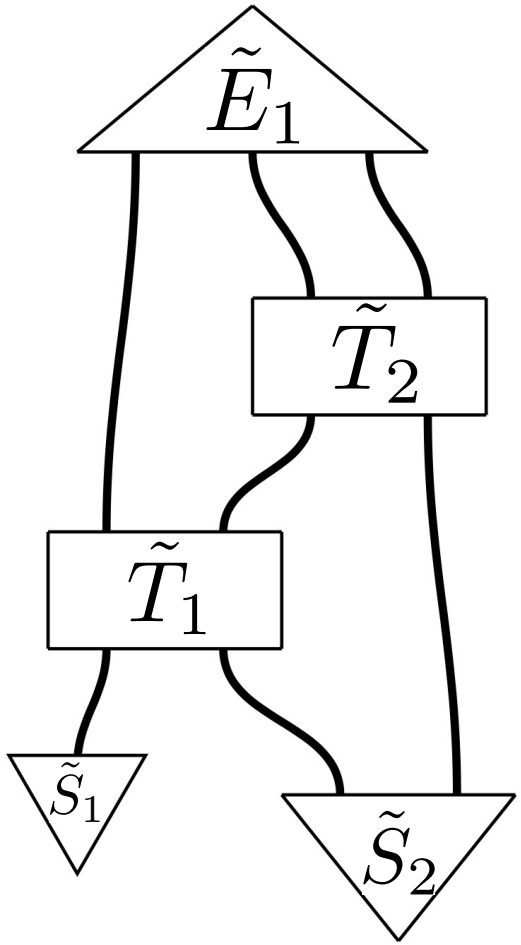
GPT transformation

subStochastic map

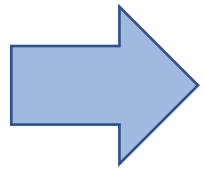
special cases: states, effects

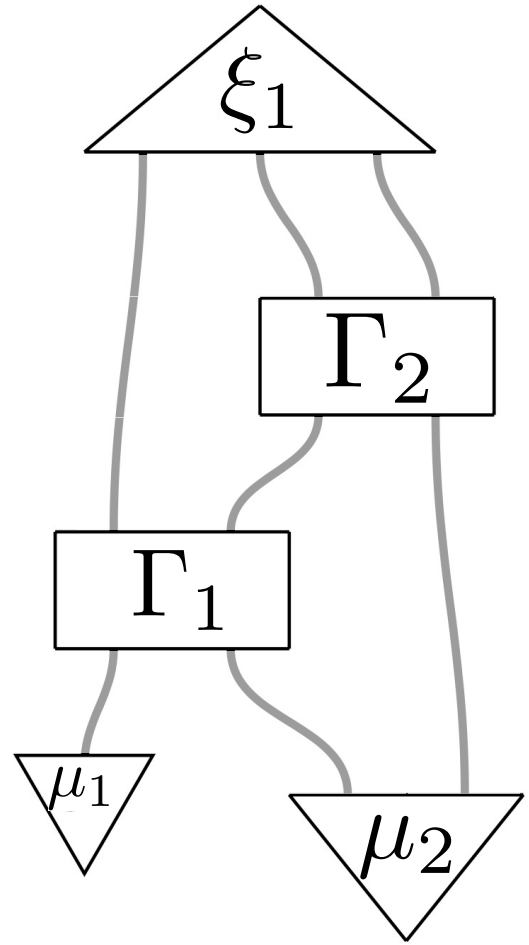
GPT processes

(sub)Stochastic processes



ontological representation





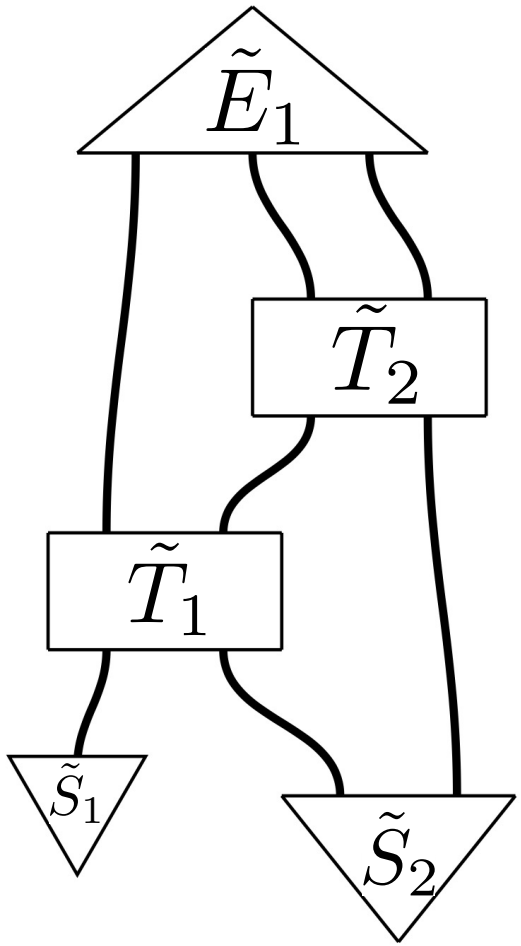
GPT systems

classical systems
(random variables)

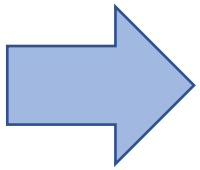
This incredibly naïve notion of classical explanation
is equivalent to noncontextuality!

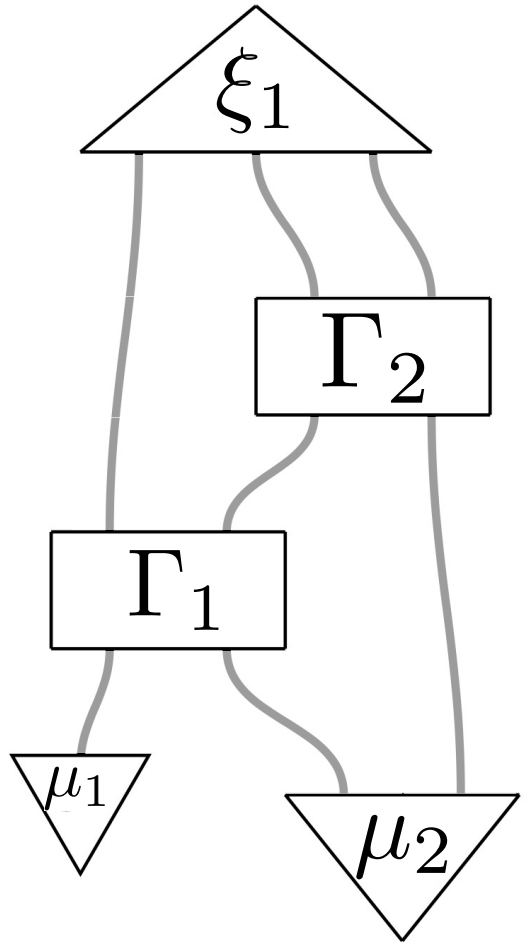
GPT processes

Quasi(sub)Stochastic processes



ontological representation





GPT systems

classical systems
(random variables)

Such representations ARE always possible if you allow “the probabilities to go negative”.

Quasiprobability representation: a map from density operators and POVM elements to real functions over some space of classical states

e.g. the Wigner function:

$$W(x,p) = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty} \langle x+y|\hat{\rho}|x-y\rangle e^{-2ipy/\hbar}\, dy$$

$$g(x,p) \equiv \int_{-\infty}^{\infty} dy \left\langle x - \frac{y}{2}\middle| \hat{G} \middle| x + \frac{y}{2}\right\rangle e^{ipy/\hbar}$$

Hence a *nonnegative* quasiprobability representation is exactly the same thing as an OM of a GPT

arxiv:2005.07161

In summary: consider an operational theory and the GPT associated to it (by quotienting with respect to operational equivalences)

The operational scenario is classically-explainable if and only if

1. it admits of a noncontextual ontological model
2. its GPT admits of any ontological model
3. its GPT admits of a simplex-embedding
4. its GPT admits of a nonnegative quasiprobability representation

These are equivalent notions of classicality!

arXiv:1911.10386
arxiv:2005.07161

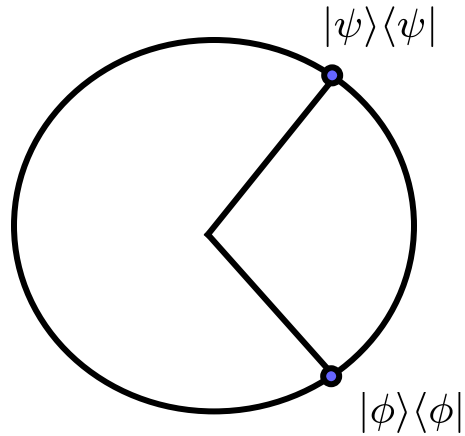
Systematically deriving noncontextuality inequalities

David Schmid

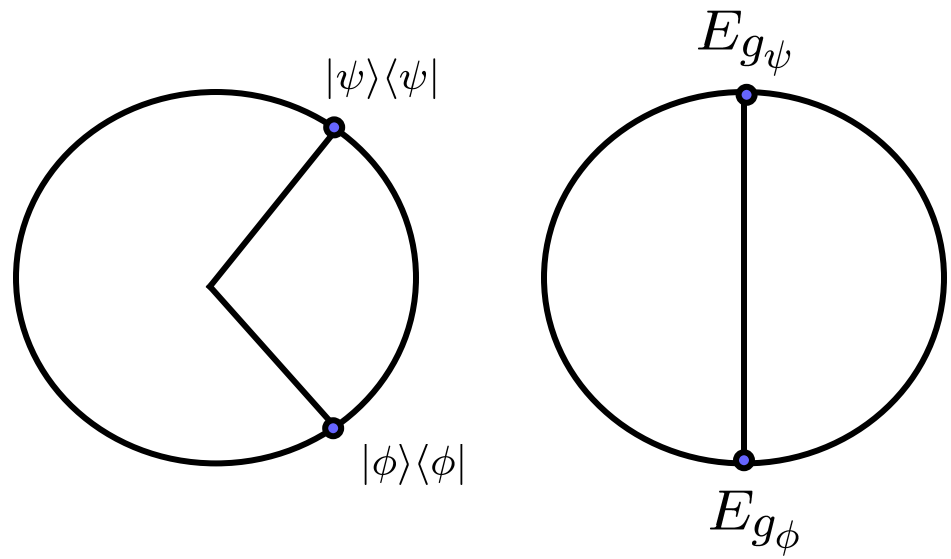
Solstice of Foundations 2022

Let’s consider a *particular* prepare-measure scenario, and let’s derive noncontextuality inequalities for it.

example: state discrimination

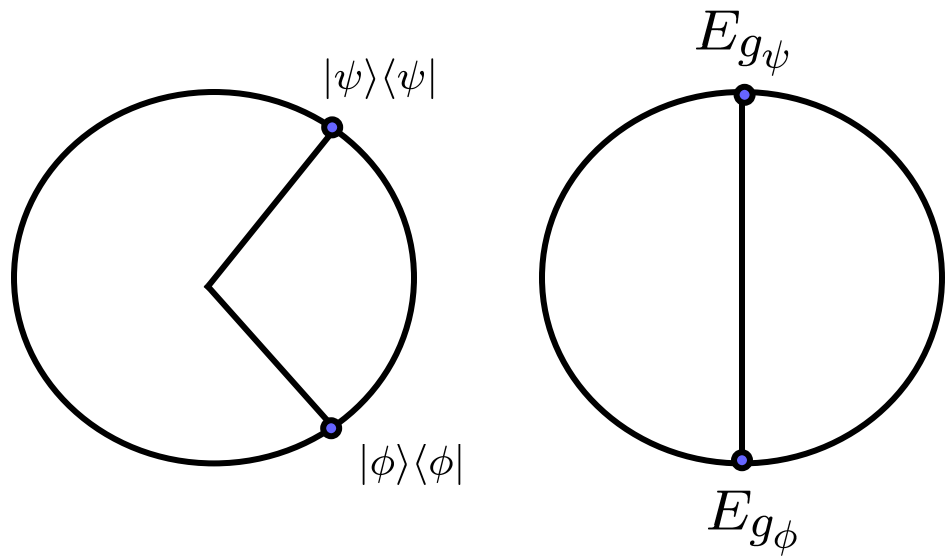


example: state discrimination



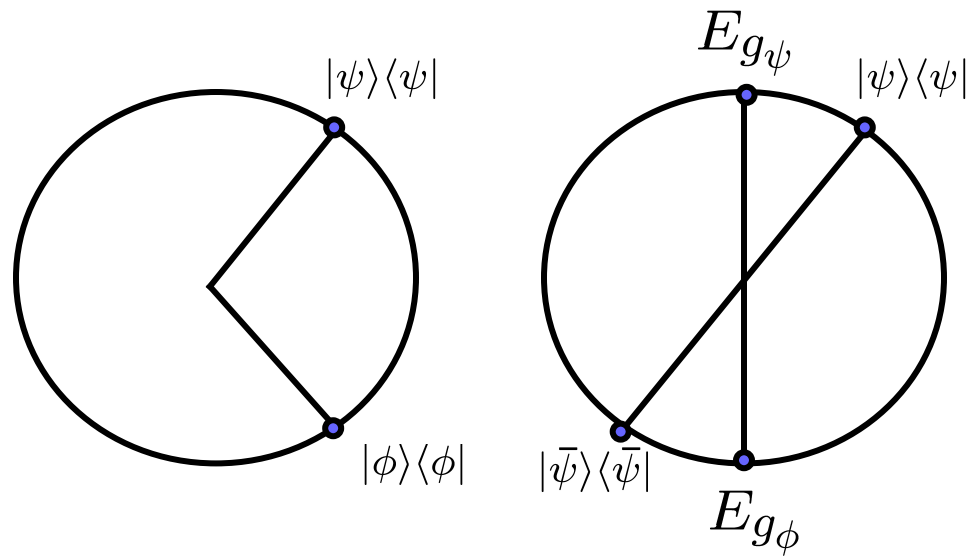
Prob. of successful discrimination in QT

$$s = \frac{1}{2}(1 + \sqrt{1 - |\langle\phi|\psi\rangle|^2})$$



Prob. of successful discrimination in QT

$$s = \frac{1}{2}(1 + \sqrt{1 - |\langle\phi|\psi\rangle|^2})$$

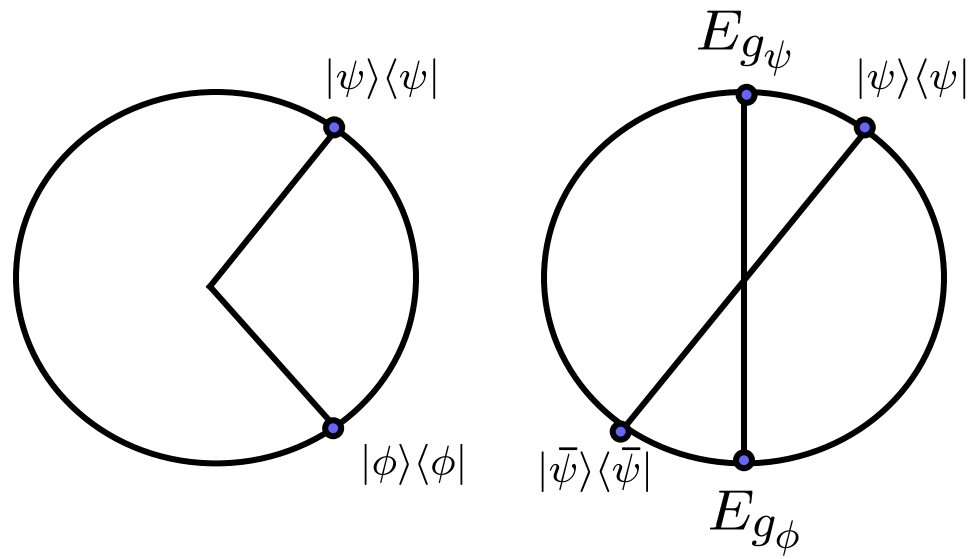


Prob. of successful discrimination in QT

$$s = \frac{1}{2}(1 + \sqrt{1 - |\langle\phi|\psi\rangle|^2})$$

Confusability

$$c = |\langle\phi|\psi\rangle|^2$$

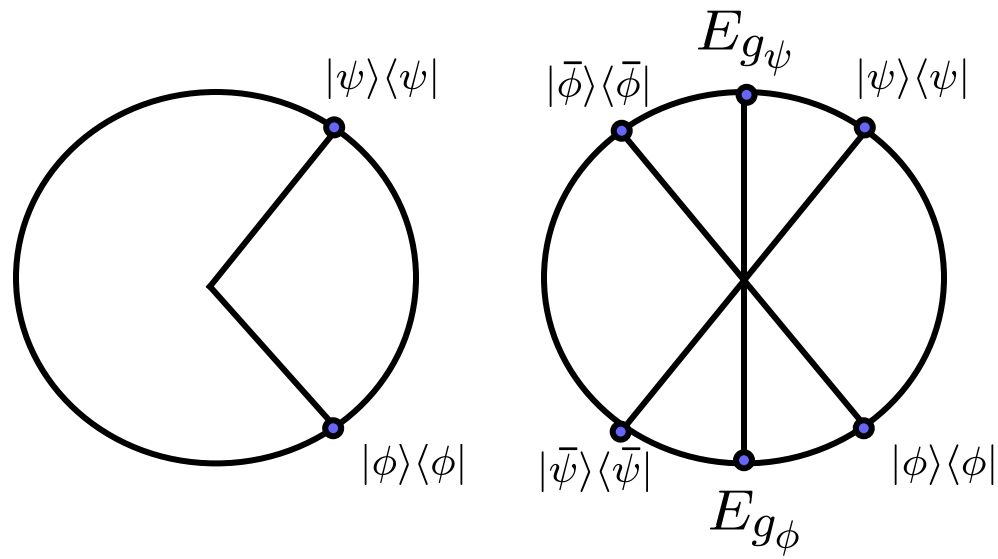


Prob. of successful discrimination in QT

$$s = \frac{1}{2}(1 + \sqrt{1 - |\langle\phi|\psi\rangle|^2})$$

Confusability

$$c = |\langle\phi|\psi\rangle|^2$$

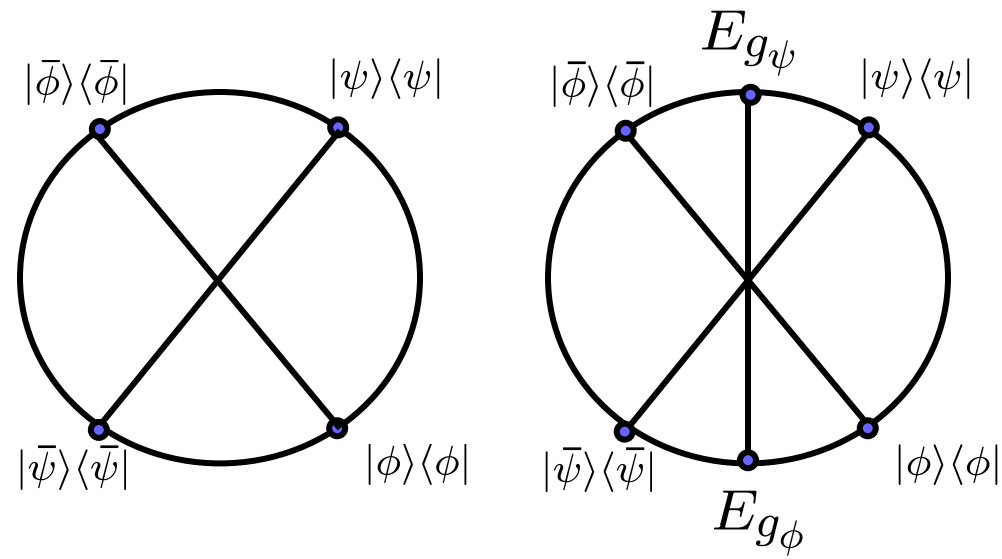


Prob. of successful discrimination in QT

$$s = \frac{1}{2}(1 + \sqrt{1 - |\langle\phi|\psi\rangle|^2})$$

Confusability

$$c = |\langle\phi|\psi\rangle|^2$$



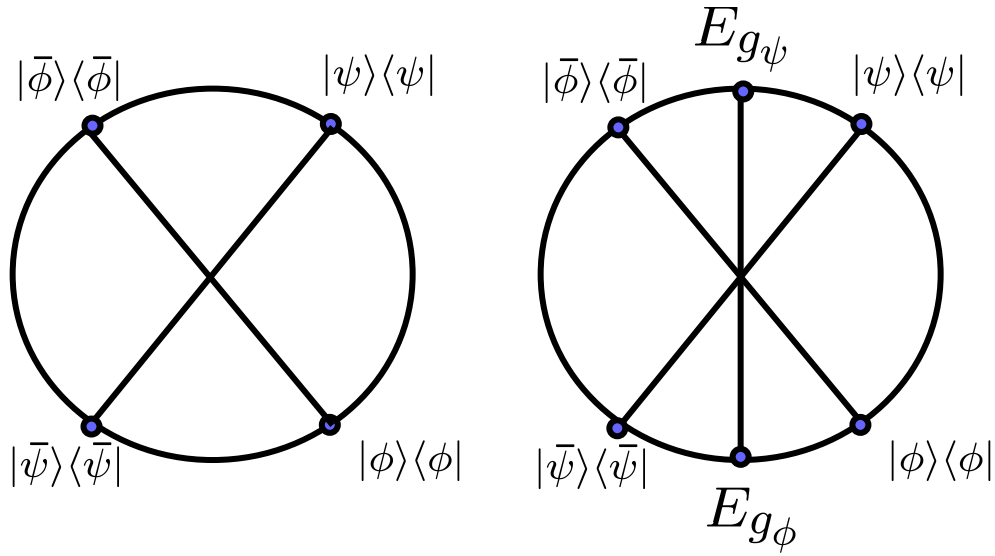
Operational equivalence:

$$\frac{1}{2}|\phi\rangle\langle\phi| + \frac{1}{2}|\bar{\phi}\rangle\langle\bar{\phi}| = \frac{1}{2}|\psi\rangle\langle\psi| + \frac{1}{2}|\bar{\psi}\rangle\langle\bar{\psi}| = \frac{\mathbb{1}}{2}$$

Noncontextuality:

$$\frac{1}{2}\mu_\phi(\lambda) + \frac{1}{2}\mu_{\bar{\phi}}(\lambda) = \frac{1}{2}\mu_\psi(\lambda) + \frac{1}{2}\mu_{\bar{\psi}}(\lambda)$$

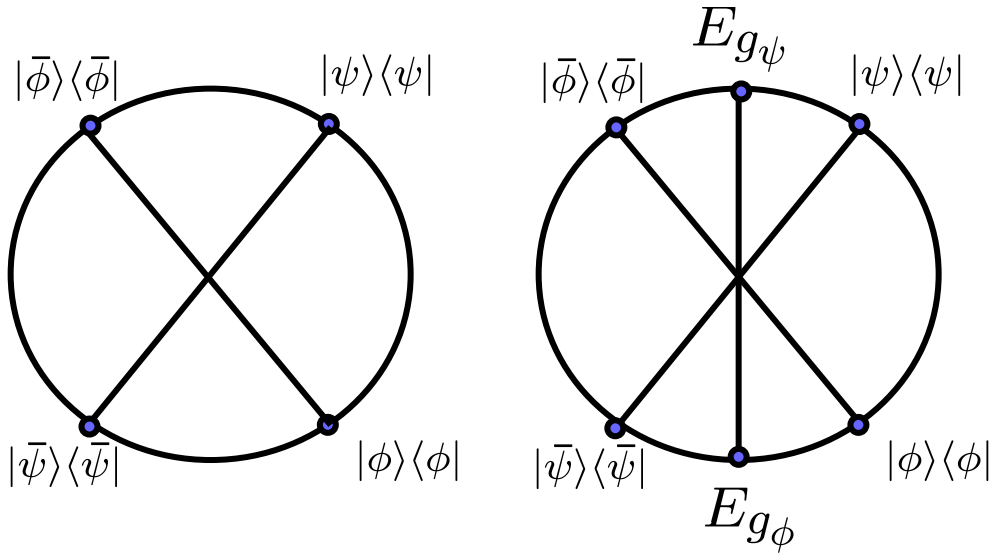
What are the consequences of this constraint?



Empirical Quantities

	P_ϕ	P_ψ	$P_{\bar{\phi}}$	$P_{\bar{\psi}}$
$\phi\|M_\phi$	1	c	0	$1-c$
$\psi\|M_\psi$	c	1	$1-c$	0
$g_\phi\|M_d$	s	$1-s$	$1-s$	s

What are the consequences of this constraint?



Noncontextuality implies

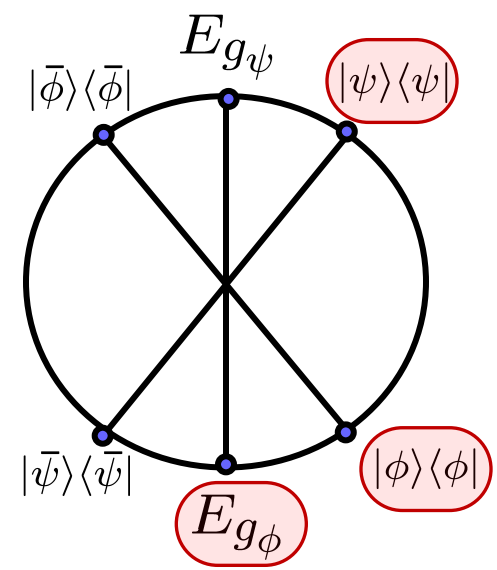
$$s \leq 1 - \frac{c}{2}$$

What are the consequences of this constraint?

Derivation of this NC inequality

Each ontic state must (at a minimum) specify the probability of each outcome for the measurements under consideration:

$$\xi_{\phi|B_\phi}(\lambda) \qquad \xi_{\psi|B_\psi}(\lambda) \qquad \xi_{g_\phi|B_d}(\lambda)$$

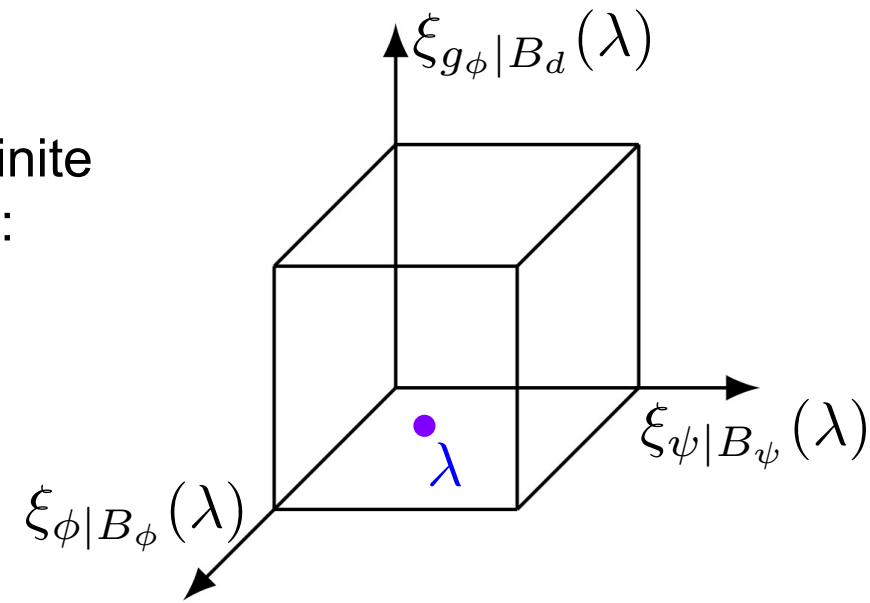


Each ontic state must (at a minimum) specify the probability of each outcome for the measurements under consideration:

$$\left(\xi_{\phi|B_\phi}(\lambda), \xi_{\psi|B_\psi}(\lambda), \xi_{g_\phi|B_d}(\lambda)\right)$$ (probabilistic) measurement assignment

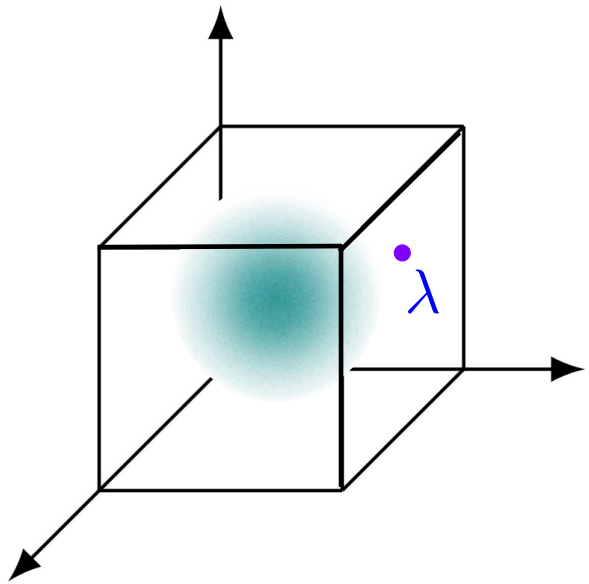
For our scenario, these are the only features of an ontic state that matter

…so we can assume wlog a continuously infinite ontic state space isomorphic to the unit cube:

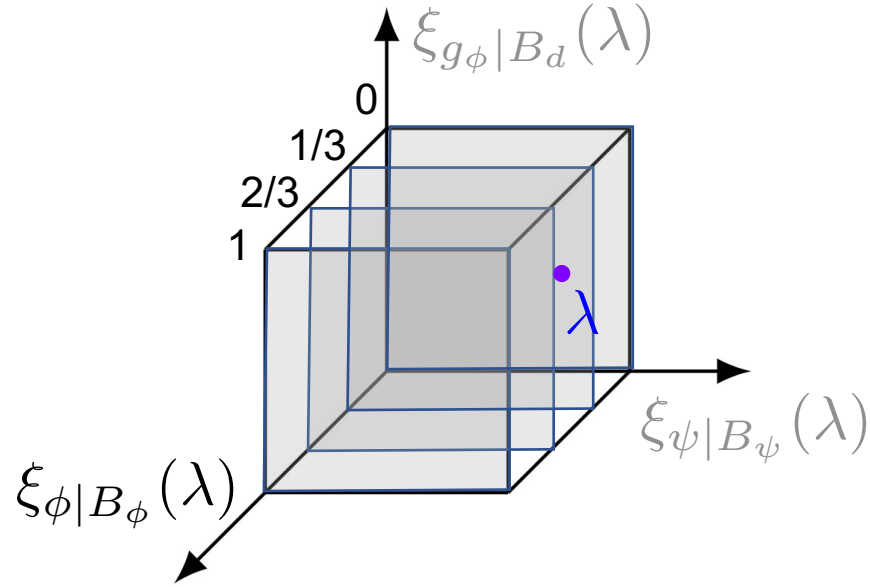


each point in the cube is an ontic state; each ontic state is fully specified by its coordinates: its three probability assignments to the three effects

Epistemic states are arbitrary probability distributions over the set of points in the cube:



response function for Bϕ:



The predictions made by the model are $p(k|M,P) = \int_\Lambda \xi_{k|M}(\lambda)\mu_P(\lambda)d\lambda$

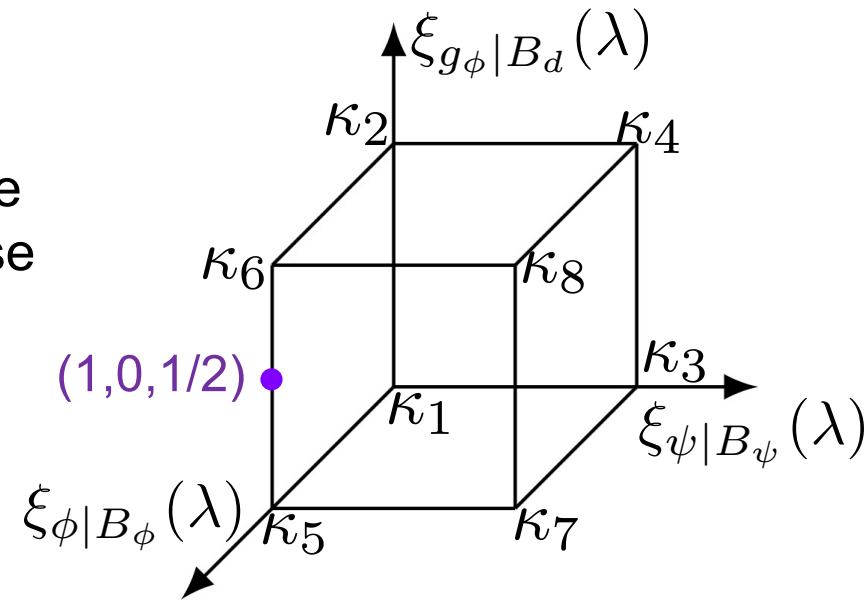
We won't actually work with any such model, though, because we can *simplify it* wlog.

We can decompose any measurement assignment (point inside the cube) as a convex mixture of deterministic assignments (vertices of the cube).

Label the vertices κ_1 to κ_8 and define a p(κ|λ) such that

$$\xi_{k|M}(\lambda) = \sum_{\kappa} \xi_{k|M}(\kappa) p(\kappa|\lambda)$$

e.g. this value assignment would be written as a uniform mixture of those made by κ_5 (1,0,0) and κ_6 (1,0,1)



Define a new model with just 8 ontic states (the κ's)

$$\xi_{k|M}(\lambda) = \sum_{\kappa} \xi_{k|M}(\kappa) p(\kappa|\lambda)$$

$$p(k|M,P) = \int_{\Lambda} \xi_{k|M}(\lambda)\mu_P(\lambda)d\lambda = \int_{\Lambda} \sum_{\kappa} \xi_{k|M}(\kappa)\underbrace{p(\kappa|\lambda)\mu_P(\lambda)d\lambda}$$

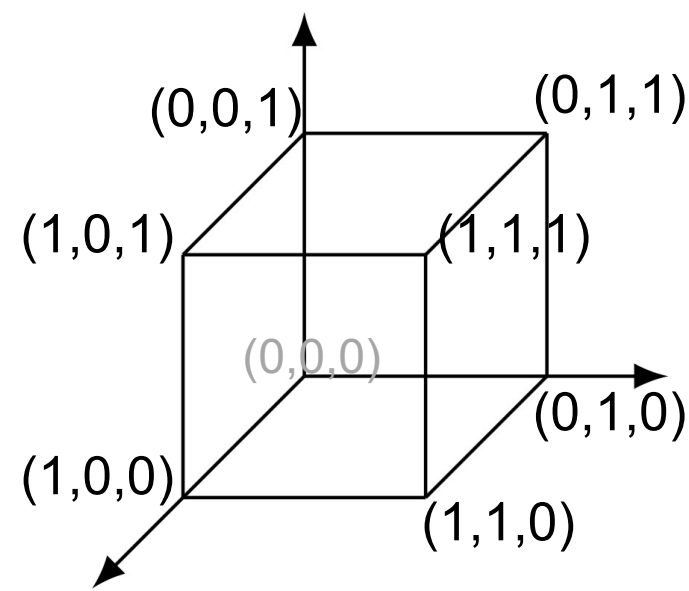
$$\mu_P(\kappa) \equiv \int_{\Lambda} d\lambda \; p(\kappa|\lambda)\mu_P(\lambda)$$

this “moves uncertainty from the mmts to the prepns (wlog)”

$$p(k|M,P) = \sum_{\kappa} \xi_{k|M}(\kappa)\mu_P(\kappa)$$

Wlog, we have converted an outcome-indeterministic model over a continuum of ontic states (the unit cube) to an outcome-deterministic model over just 8 ontic states (its vertices)

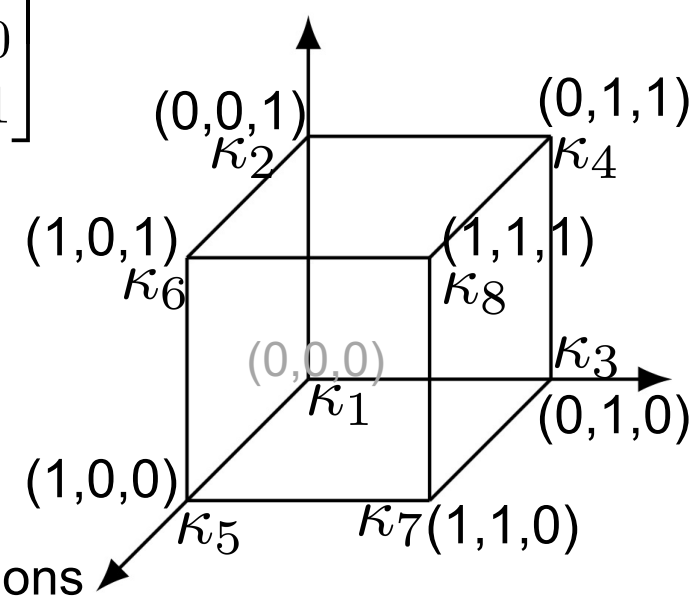
$$p(k|M,P) = \sum_{\kappa} \xi_{k|M}(\kappa)\mu_P(\kappa)$$



3 deterministic response functions that make them:

$$\vec{\xi}_{\phi|B_\phi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{\psi|B_\psi} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{g_\phi|B_d} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

8 (deterministic)
mmt assignments



each ontic state makes assignments for 3 response functions
each response function makes assignments for 8 ontic states

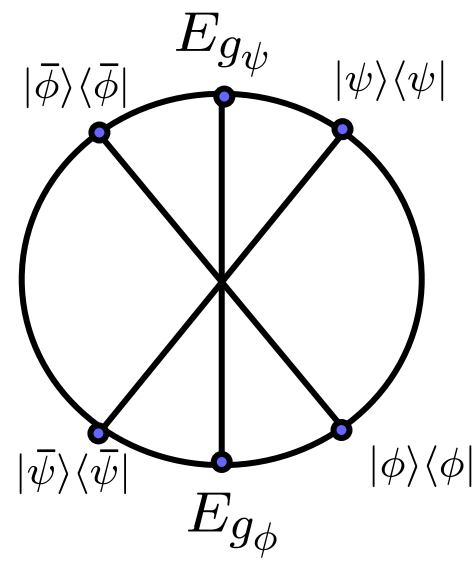
In fact, if the OM is optimal, then the fourth and fifth value assignments never occur

$$\vec{\xi}_{\phi|B_\phi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \boxed{1} \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{\psi|B_\psi} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ \boxed{0} \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{g_\phi|B_d} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ \boxed{0} \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

state was not $\bar{\phi}$

state was not ψ

so one should always guess g_ϕ



In fact, if the OM is optimal, then the fourth and fifth value assignments never occur

$$\vec{\xi}_{\phi|B_\phi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cancel{0} \\ \cancel{1} \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{\psi|B_\psi} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \cancel{1} \\ \cancel{0} \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{g_\phi|B_d} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \cancel{1} \\ \cancel{0} \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

So an *optimal* model can be taken to have just 6 ontic states, wlog

$$\vec{\xi}_{\phi|B_\phi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{\psi|B_\psi} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{g_\phi|B_d} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Now, let's parametrize the epistemic states as arbitrary probability distributions.

$$\vec{\xi}_{\phi|B_\phi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{\psi|B_\psi} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{g_\phi|B_d} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\vec{\mu}_\phi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix}$$

$$\vec{\xi}_{\phi|B_\phi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{\psi|B_\psi} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{g_\phi|B_d} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\vec{\mu}_\phi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix}, \vec{\mu}_{\bar{\phi}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\vec{\xi}_{\phi|B_\phi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{\psi|B_\psi} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad \vec{\xi}_{g_\phi|B_d} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\vec{\mu}_\phi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix}, \vec{\mu}_{\bar{\phi}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \vec{\mu}_\psi = \begin{bmatrix} 0 \\ 0 \\ b_3 \\ 0 \\ b_5 \\ b_6 \end{bmatrix}, \vec{\mu}_{\bar{\psi}} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \\ b_4 \\ 0 \\ 0 \end{bmatrix}$$

(by the perfect predictability of the Bψ and Bϕ mmts)

must be normalized: $a_4 + a_5 + a_6 = 1$

(etc)

$$\vec{\mu}_\phi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix}, \vec{\mu}_{\bar{\phi}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \vec{\mu}_\psi = \begin{bmatrix} 0 \\ 0 \\ b_3 \\ 0 \\ b_5 \\ b_6 \end{bmatrix}, \vec{\mu}_{\bar{\psi}} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \\ b_4 \\ 0 \\ 0 \end{bmatrix}$$

$$\vec{\mu}_\phi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix}, \vec{\mu}_{\bar{\phi}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \vec{\mu}_\psi = \begin{bmatrix} 0 \\ 0 \\ b_3 \\ 0 \\ b_5 \\ b_6 \end{bmatrix}, \vec{\mu}_{\bar{\psi}} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \\ b_4 \\ 0 \\ 0 \end{bmatrix} \qquad \vec{\xi}_{\phi|B_\phi} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \vec{\xi}_{\psi|B_\psi} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \vec{\xi}_{g_\phi|B_d} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

reproducing the empirical predictions:

$$c = \vec{\mu}_\psi \cdot \vec{\xi}_{\phi|B_\phi} = \vec{\mu}_\phi \cdot \vec{\xi}_{\psi|B_\psi} = 1 - \vec{\mu}_{\bar{\psi}} \cdot \vec{\xi}_{\phi|B_\phi} = 1 - \vec{\mu}_{\bar{\phi}} \cdot \vec{\xi}_{\psi|B_\psi}$$

$$= 1 - b_3 = 1 - a_4 = 1 - b_4 = 1 - a_3$$

hence $b_3 = a_4 = b_4 = a_3$

$$s = \vec{\mu}_\phi \cdot \vec{\xi}_{g_\phi|B_d} = 1 - \vec{\mu}_\psi \cdot \vec{\xi}_{g_\phi|B_d} = 1 - \vec{\mu}_{\bar{\phi}} \cdot \vec{\xi}_{g_\phi|B_d} = \vec{\mu}_{\bar{\psi}} \cdot \vec{\xi}_{g_\phi|B_d}$$

$$= 1 - a_5 = 1 - b_6 = 1 - a_2 = 1 - b_1$$

hence $a_5 = b_6 = a_2 = b_1$

$$s = 1 - a_2$$

$$c = 1 - a_3$$

Without further constraints, a_2 and a_3 can range independently from 0 to 1, and so s and c are independent

But preparation noncontextuality gives us further constraints!

operational equivalence: $$\frac{1}{2}\left|\phi\right\rangle\left\langle\phi\right| + \frac{1}{2}\left|\bar{\phi}\right\rangle\left\langle\bar{\phi}\right| = \frac{1}{2}\left|\psi\right\rangle\left\langle\psi\right| + \frac{1}{2}\left|\bar{\psi}\right\rangle\left\langle\bar{\psi}\right|$$

noncontextuality: $$\frac{1}{2}\mu_\phi(\lambda) + \frac{1}{2}\mu_{\bar{\phi}}(\lambda) = \frac{1}{2}\mu_\psi(\lambda) + \frac{1}{2}\mu_{\bar{\psi}}(\lambda)$$

$$\vec{\mu}_\phi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix}, \vec{\mu}_{\bar{\phi}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \vec{\mu}_\psi = \begin{bmatrix} 0 \\ 0 \\ b_3 \\ 0 \\ b_5 \\ b_6 \end{bmatrix}, \vec{\mu}_{\bar{\psi}} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \\ b_4 \\ 0 \\ 0 \end{bmatrix}$$

therefore $$\frac{1}{2}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{bmatrix}$$

and so $$a_i = b_i$$

$$a_1 + a_2 + a_3 = 1$$ normalization

$$a_1 = b_1$$ PNC

$$b_1 = a_2$$ empirical predictions

$$2a_2 + a_3 = 1$$ substitution

$$c = 1 - a_3 = 2a_2$$ empirical predictions and substitution

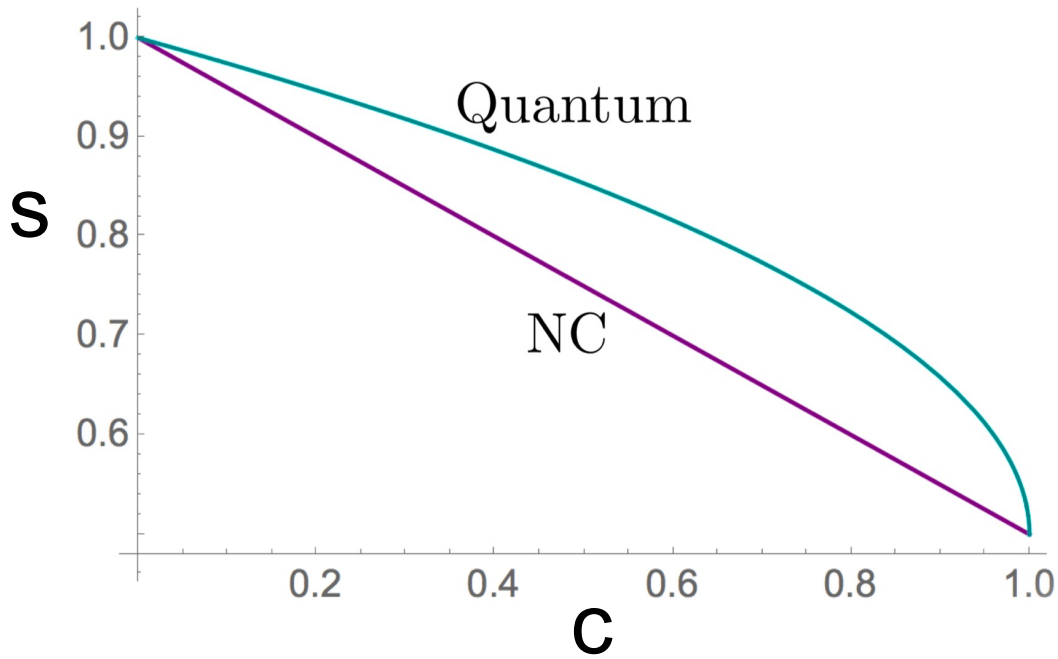
$$s = 1 - a_2 = 1 - \frac{c}{2}$$ empirical predictions and substitution

So the *optimal* NC model achieves $$s = 1 - \frac{c}{2}$$

Every NC model satisfies $$s \leq 1 - \frac{c}{2}$$

This is a noncontextuality inequality!

Probability
of success



confusability

So contextuality provides an advantage
for state discrimination

So, what is nonclassical about quantum state discrimination?

Traditional answer: isn't it weird that we can't perfectly discriminate every pair of quantum states?

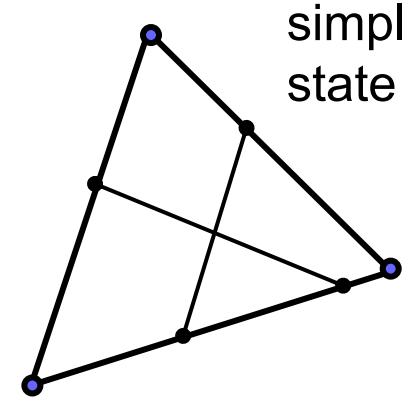
> Perhaps… but this is also a feature of a huge range of GPTs, including all simplex-embeddable but nonsimplicial GPTs!

New answer: for states *satisfying the relevant operational equivalence*, quantum theory allows us to discriminate them <u>better than linearly</u>!

This completely reverses the traditional intuitions about what is nonclassical about state discrimination!

(similarly for no-cloning, uncertainty relations, etc)

Traditional intuitions come from thinking about the extremal states in a strictly classical theory



These don't satisfy *any* operational equivalences, so NC inequalities <u>simply don't apply to them</u>.

States in a simplicial theory that *do* satisfy the operational equivalences (which must be mixed) *do* satisfy the NC inequality.

elaborate on this?
but mixed states in a simplicial GPT do satisfy OEs,
and so must satisfy the NC inequality

maybe I should put this at the start

Our arguments so far assumed some idealizations

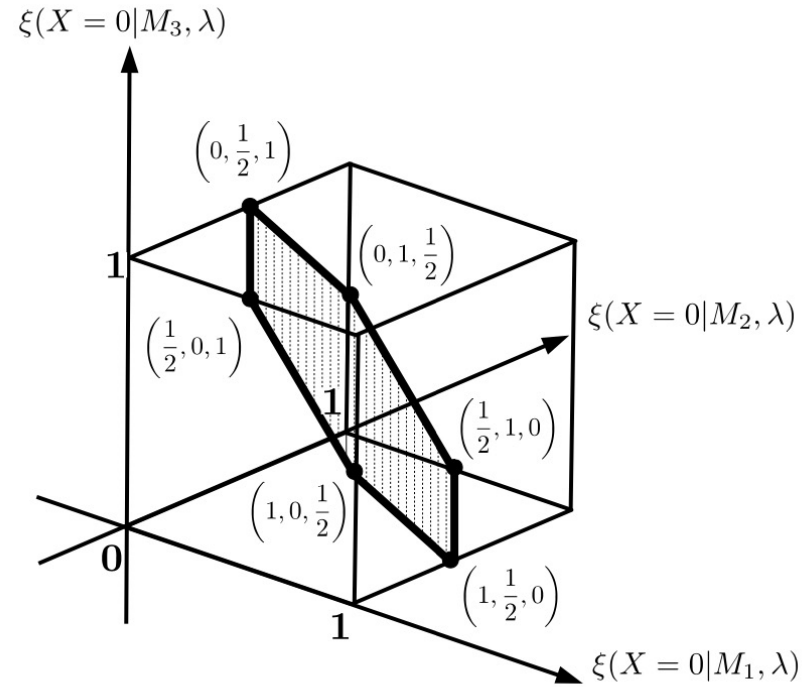
	P_ϕ	P_ψ	$P_{\bar\phi}$	$P_{\bar\psi}$
$\phi \vert M_\phi$	1	c	0	$1-c$
$\psi \vert M_\psi$	c	1	$1-c$	0
$g_\phi \vert M_d$	s	$1-s$	$1-s$	s

but these techniques work just as well for the general scenario:

	P_ϕ	P_ψ	$P_{\bar\phi}$	$P_{\bar\psi}$
$\phi \vert M_\phi$	$1-\epsilon_\phi$	c_ψ	$\epsilon_{\bar\phi}$	$1-\epsilon_\phi+\epsilon_{\bar\phi}-c_\psi$
$\psi \vert M_\psi$	c_ϕ	$1-\epsilon_\psi$	$1-c_{\bar\phi}$	$c_\phi-c_{\bar\phi}+\epsilon_\psi$
$g_\phi \vert M_d$	s_ϕ	$1-s_\psi$	$1-s_{\bar\phi}$	$s_\phi-s_{\bar\phi}-s_\psi$

Generalizing this procedure to arbitrary scenarios

What about measurement noncontextuality?



MNC + operational equivalences among effects give linear constraints on the hypercube of response functions:

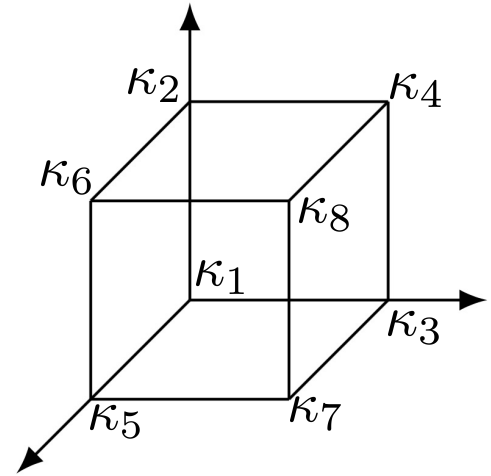
$$\xi_{b|M_1}(\lambda) + \xi_{b|M_2}(\lambda) + \xi_{b|M_3}(\lambda) = \frac{1}{2}$$

Outcome deterministic response functions generally violate MNC.

write converse here?

However, we can explicitly compute the finite set of extremal MNC assignments. We then assign one ontic state to each of these, and decompose all MNC assignments in terms of the extremal assignments.

Deriving all the NC inequalities for a PM scenario:



1. find the extremal mmt assignments consistent with the constraints from MNC
2. associate one ontic state to each of these
3. parametrize the epistemic states as vectors of unknown probabilities
4. write empirical quantities as the dot product of epistemic states with extremal response functions

$$\vec{\mu}_\phi = \begin{bmatrix} 0 \\ 0 \\ 0 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix}, \vec{\mu}_{\bar{\phi}} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \vec{\mu}_\psi = \begin{bmatrix} 0 \\ 0 \\ b_3 \\ 0 \\ b_5 \\ b_6 \end{bmatrix}, \vec{\mu}_{\bar{\psi}} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \\ b_4 \\ 0 \\ 0 \end{bmatrix}$$

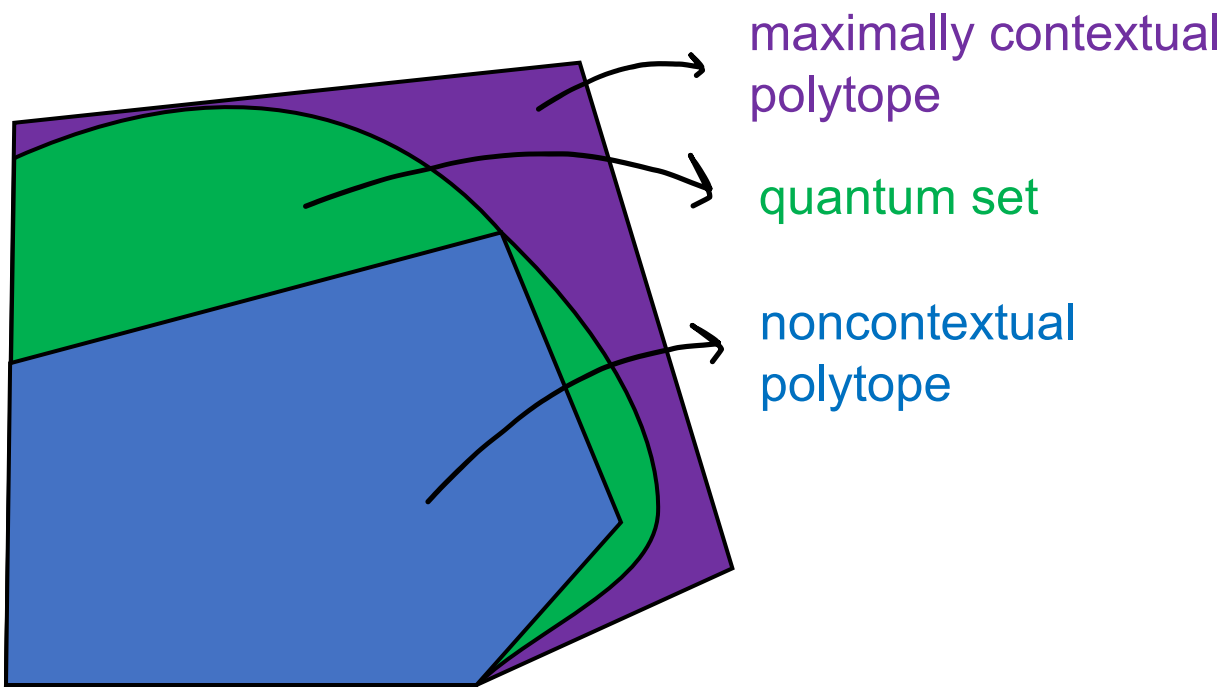
At this stage, one has the empirical quantities written as (linear!) functions of some unknown parameters that are between 0 and 1.

$$\begin{aligned} c &= \vec{\mu}_\psi \cdot \vec{\xi}_{\phi|B_\phi} \\ &= 1 - b_3 \end{aligned}$$

5. write down the (linear!) constraints these unknown parameters satisfy as a consequence of normalization and preparation noncontextuality
6. solve for the empirical quantities (can be done via a linear program—
“quantifier elimination”)

$$\sum_i p_i = 1 \qquad \frac{1}{2}\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{bmatrix}$$

The result is always a set of inequalities on the empirical quantities, whose satisfaction is necessary and sufficient for a set of data to be noncontextual.



arXiv:1710.08434

arXiv:2010.05853

Systematic tools for studying noncontextuality:

finding all the NC inequalities:
arXiv:1710.08434

determining simplex-embeddability and (some) measures of nonclassicality:
arXiv:2204.11905

computing quantum or maximally-contextual bounds: (see Vicky's talk)
arXiv:2010.05853

basics of simplex-embeddability:
arXiv:1911.10386

definitions and results regarding NC in compositional theories:
arxiv:2005.07161

Noncontextuality: odds and ends

David Schmid

Solstice of Foundations 2022

The linearity assumption

GPTs have a vector space structure
We assume that ontological representations are linear on this structure.

$$\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$$

ontological representation: $\mu_{\frac{1}{2}|0\rangle\langle 0|+\frac{1}{2}|1\rangle\langle 1|}(\lambda)$

using linearity: $\frac{1}{2}\mu_{|0\rangle\langle 0|}(\lambda) + \frac{1}{2}\mu_{|1\rangle\langle 1|}(\lambda)$

note: operational theories do *not* have vector space structure!
(unless formalized in a more sophisticated manner)

WHY must an ontological model be linear?

-we assume the correctness of classical probability theory
-we imagine that the convex mixtures arise simply because one is uncertain which pure state was prepared

$$\left\{ \left(\frac{1}{2}, |0\rangle\right), \left(\frac{1}{2}, |1\rangle\right) \right\}$$

knowledge of preparation

$$p(\lambda) = p(\lambda|0)p(0) + p(\lambda|1)p(1)$$
$$= \frac{1}{2}\mu_{|0\rangle\langle 0|}(\lambda) + \frac{1}{2}\mu_{|1\rangle\langle 1|}(\lambda)$$

law of total probability

$$P(A) = \sum_n P(A \mid B_n)P(B_n)$$

The tomographic completeness assumption

Every assessment of contextuality or noncontextuality starts from *some* operational equivalence.

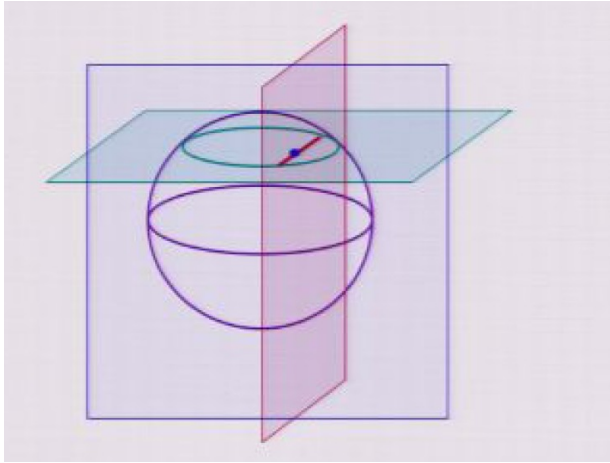
$$\mathsf{P} \simeq \mathsf{P}' \xrightarrow{\text{NC}} \mu(\lambda|\mathsf{P}) = \mu(\lambda|\mathsf{P}')$$

$$\boxed{\forall \mathsf{M}}: p(X|\mathsf{P},\mathsf{M}) = p(X|\mathsf{P}',\mathsf{M})$$

"For all measurements" = all measurements on a system that are possible within the fundamental physical theory

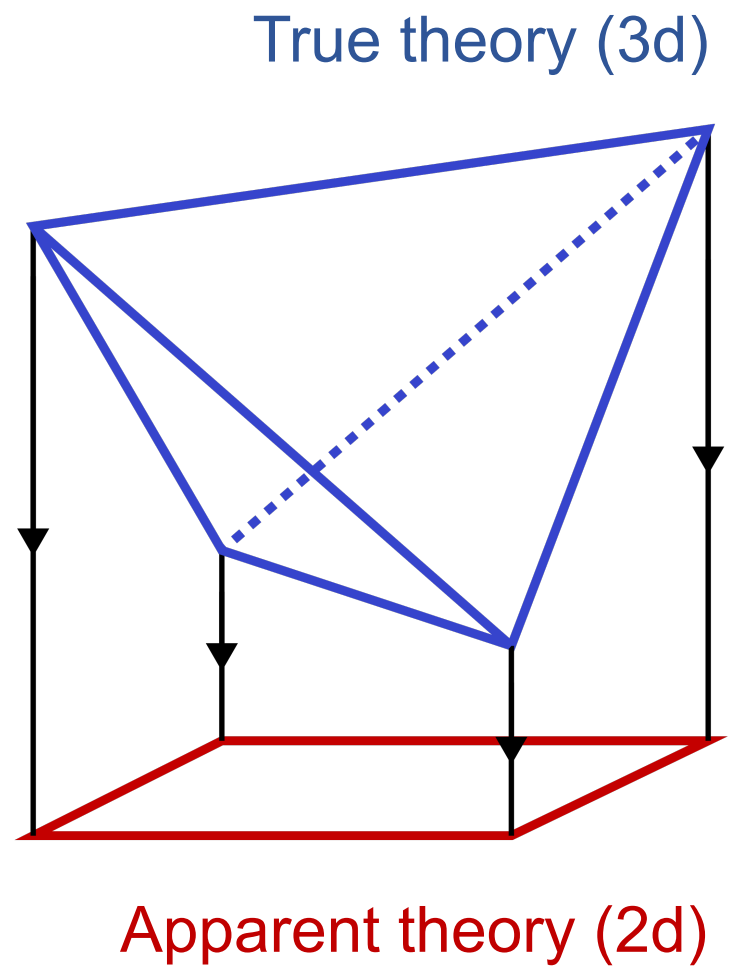
It is sufficient (and necessary) to check if operational equivalence holds for a tomographically complete set of measurements

For a qubit, for example, one can just check the X, Y, and Z measurements



But what if one is *mistaken* about tomographic completeness?

Mistaken assumption of tomographic completeness
can lead to
incorrect operational equivalences
and hence to
mistaken assessments of (non)classicality.



Imagine you happen to *not* know about measurements along the Z axis. Then the 3-d simplicial (classical!) GPT will give statistics that look like those of the 2-d Boxworld, which is highly nonclassical.

So how can we be confident in our assessments of (non)classicality?

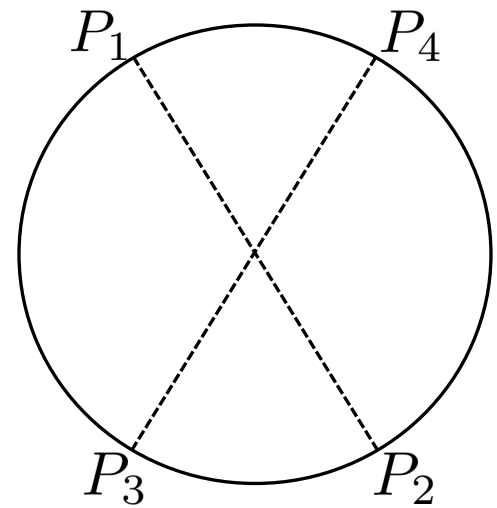
One can never *verify* the claim that a given set of procedures is tomographically complete, but one can accumulate more and more evidence for it by trying one's best to falsify the claim and failing to do so.

This is *exactly* the same kind of evidence that tells us it is sufficient to use density operators to describe laboratory preparations.

Dealing with imperfect operational equivalences

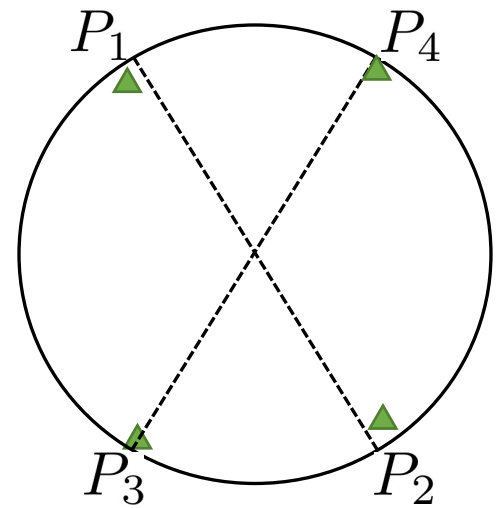
Target preparations ➡ satisfy some operational equivalences

$$\frac{1}{2}P_1 + \frac{1}{2}P_2 \simeq \frac{1}{2}P_3 + \frac{1}{2}P_4$$



Target preparations ➡ satisfy some operational equivalences

$$\frac{1}{2}P_1 + \frac{1}{2}P_2 \simeq \frac{1}{2}P_3 + \frac{1}{2}P_4$$



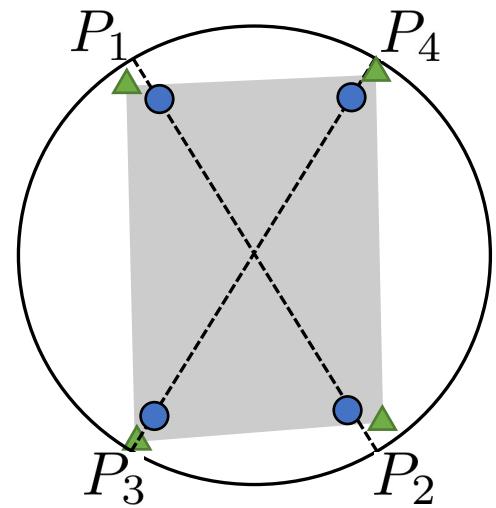
Real preparations ➡ noisy; don't exactly satisfy the OEs!

$$\frac{1}{2}P_1 + \frac{1}{2}P_2 \not\simeq \frac{1}{2}P_3 + \frac{1}{2}P_4$$

So the four realized states do *not* allow one to test NC inequalities based on this operational equivalence!

Target preparations ➡ satisfy some operational equivalences

$$\frac{1}{2}P_1 + \frac{1}{2}P_2 \simeq \frac{1}{2}P_3 + \frac{1}{2}P_4$$



Real preparations ➡ noisy; don't exactly satisfy the OEs!

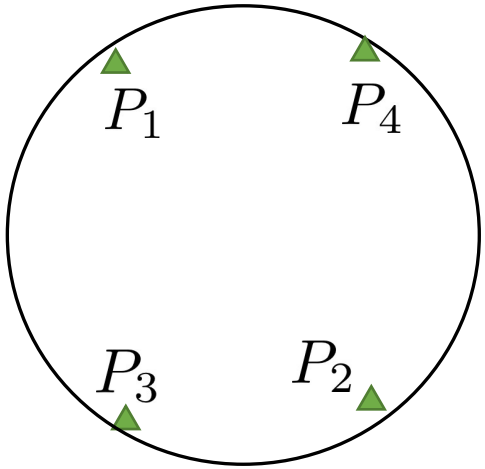
$$\frac{1}{2}P_1 + \frac{1}{2}P_2 \not\simeq \frac{1}{2}P_3 + \frac{1}{2}P_4$$

But one can infer the statistics of all states in the convex hull of the realized ones!

And some of these "inferred states" *do* exactly satisfy the OEs!

$$\frac{1}{2}P_1' + \frac{1}{2}P_2' \simeq \frac{1}{2}P_3' + \frac{1}{2}P_4'$$

Second approach



1. Do whatever procedures one likes (no need to even target anything!)

2. Characterize their GPT representations

3. Compute the *actual* operational equivalences which hold among the realized procedures
and derive NC inequalities for these

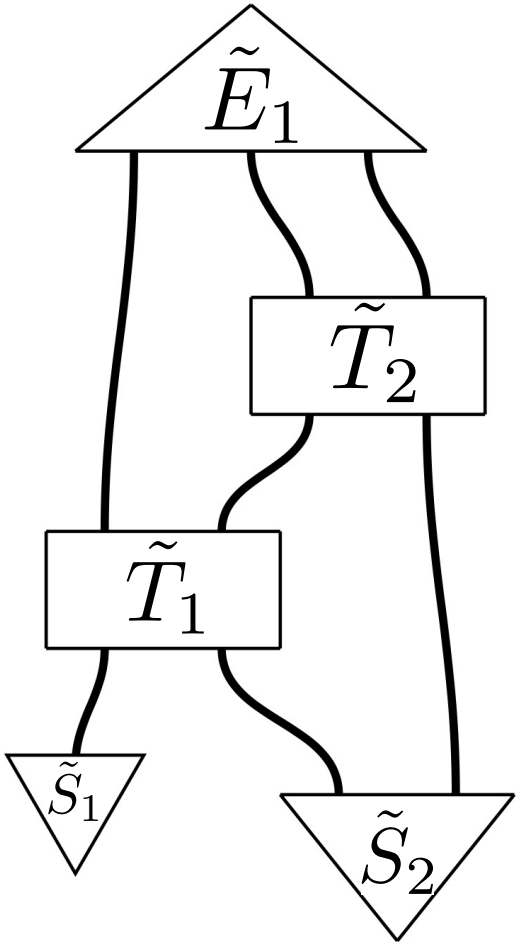
$$\alpha P_1 + (1-\alpha)P_2 = \beta P_3 + (1-\beta)P_4$$

OR

3. Determine if the GPT is simplex-embeddable

Noncontextuality beyond prepare-measure scenarios (again)

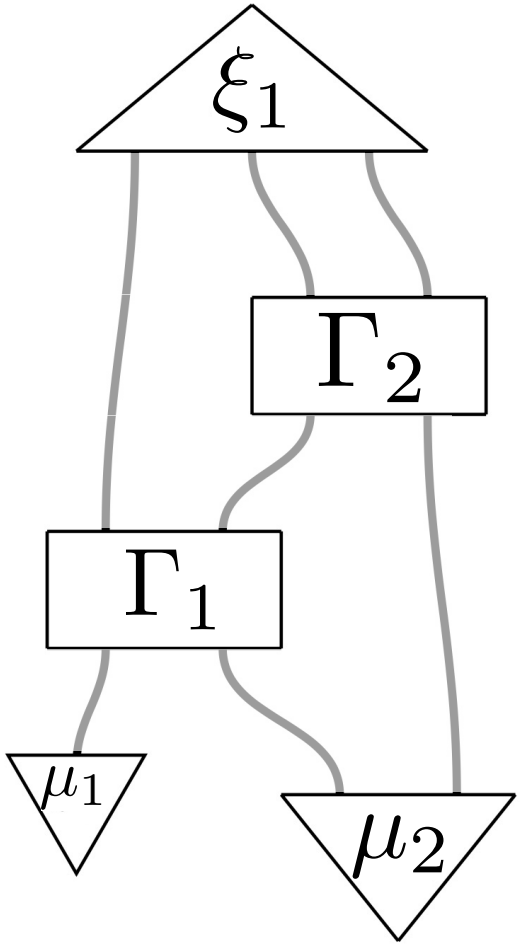
arxiv:2005.07161



GPT processes

GPT systems

quasiprobability representation



QuasiSubStochastic processes

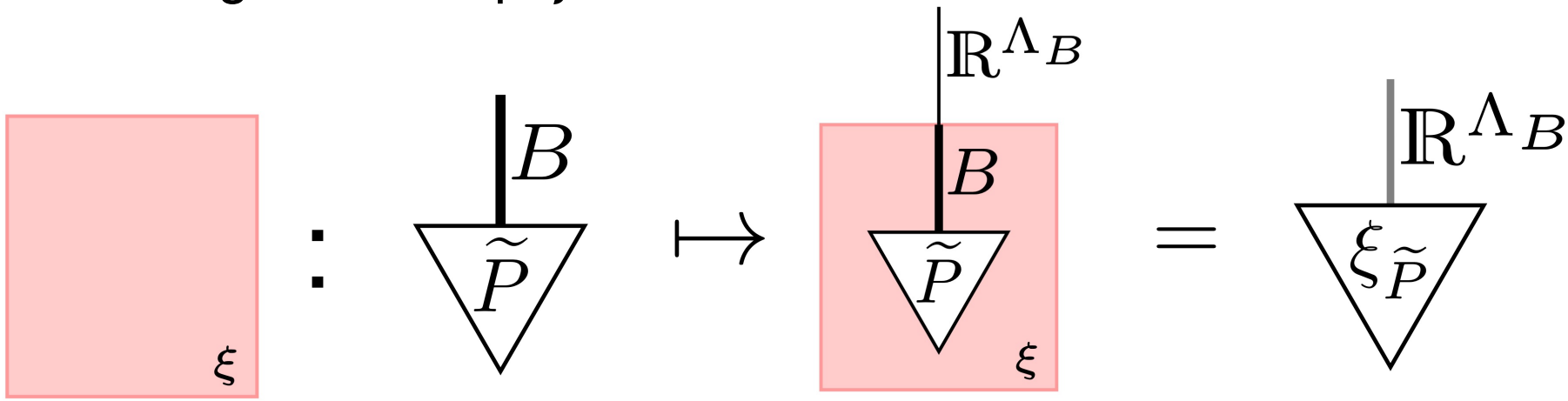
classical systems
(random variables)

If the image of the map is positive, this is an OM
(and hence a classical explanation for the GPT)

All quasiprobability representations of GPTs with transformations have a simple structure

(and hence so do all noncontextual ontological models of operational theories)

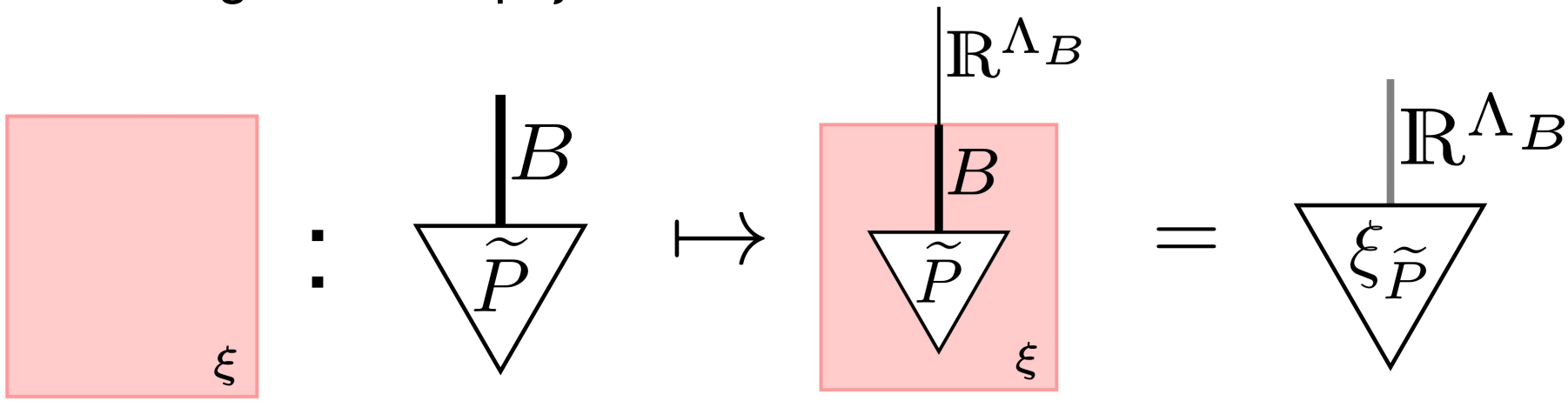
Recall that a quasiprobability representation is a diagram-preserving linear map ξ from a GPT to QuasiSubStoch:



examples: $\frac{\mathbb{I}}{2}$, $|0\rangle\langle 0|$ $\begin{bmatrix} 0.25 \\ 0.25 \\ 0 \\ 0.5 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

$\mathbb{R}^{\Lambda_B}$ is the vector space of real-valued functions over the state space Λ_B

Recall that a quasiprobability representation is a diagram-preserving linear map ξ from a GPT to QuasiSubStoch:

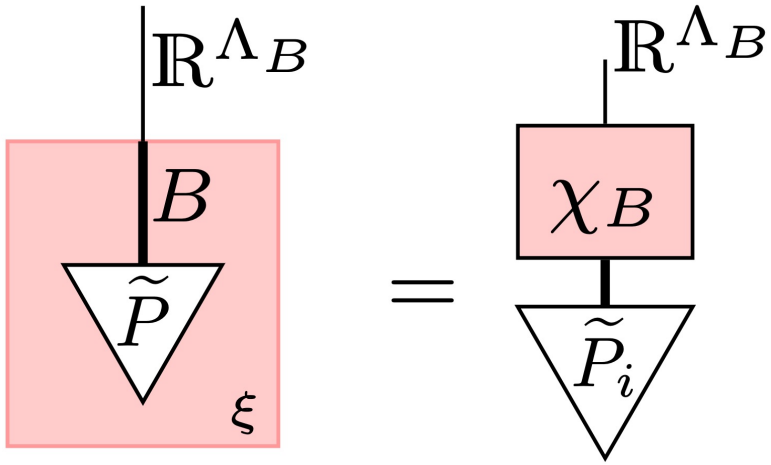


ξ is just a linear map from one real vector space to a different one

We can define a larger process theory RLinear which contains all real vector spaces as system types, and all linear maps on these as processes

Then, we can represent this DP map between process theories in a different way: *as a process within a larger process theory* (RLinear)

(e.g., χ can be a simplex embedding map as in lecture 2!)

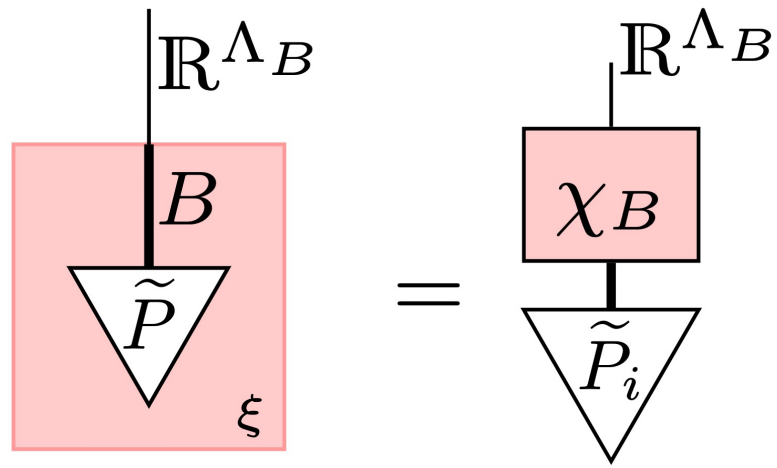


ξ is just a linear map from one real vector space to a different one

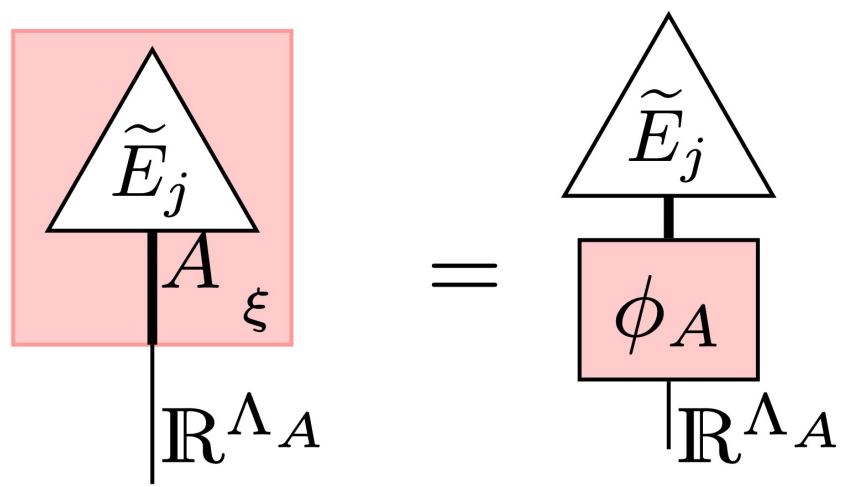
We can define a larger process theory RLinear which contains all real vector spaces as system types, and all linear maps on these as processes

Then, we can represent this DP map between process theories in a different way: *as a process within a larger process theory* (RLinear)

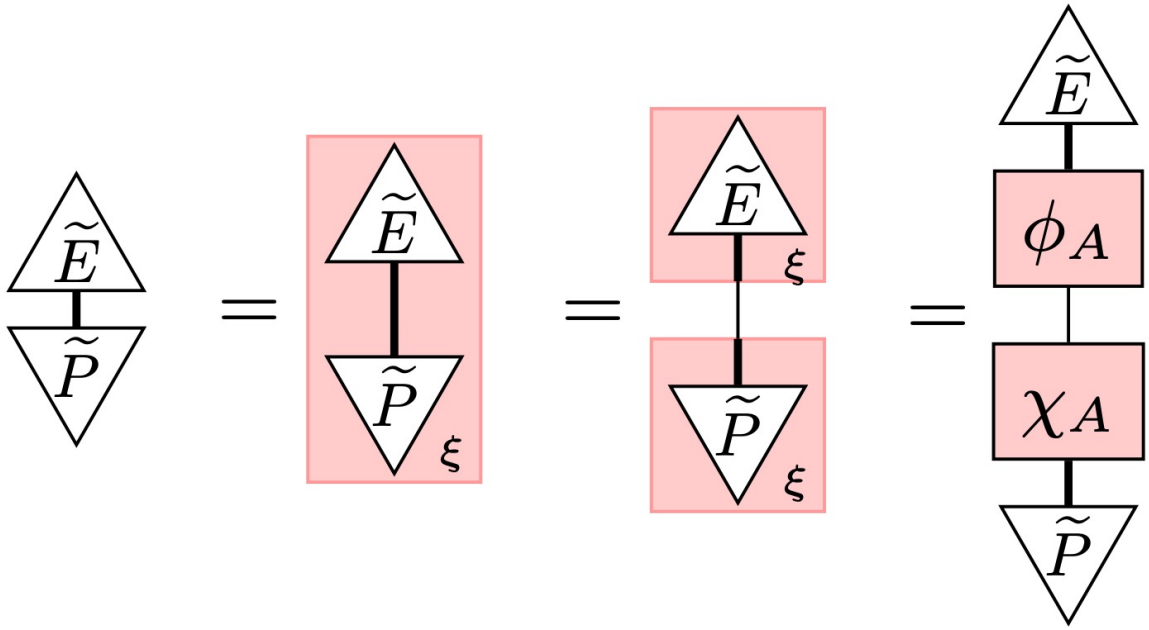
for states:



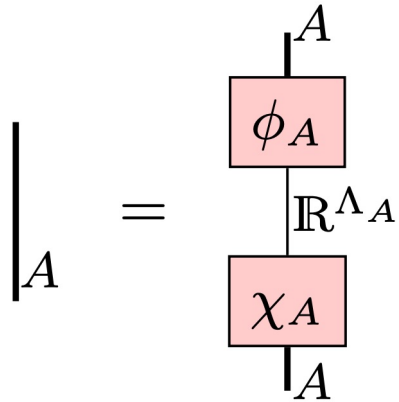
similarly for effects:



ϕ is the left inverse of χ

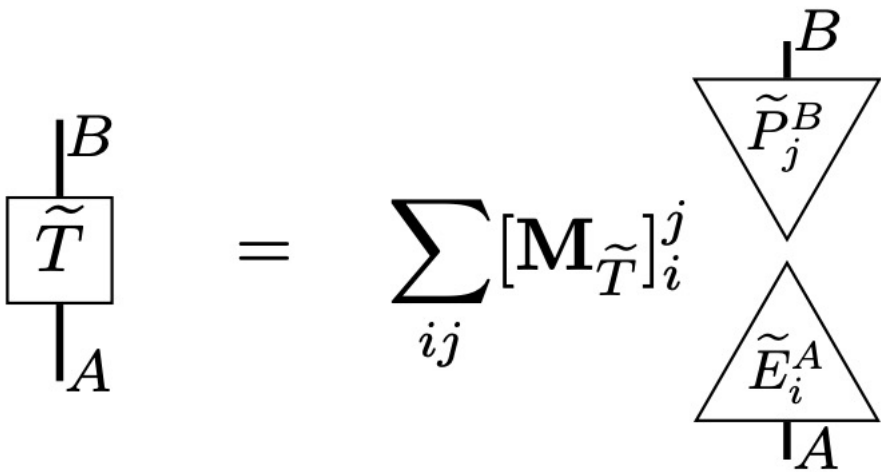


this holds for all E and P, so by tomography it follows that



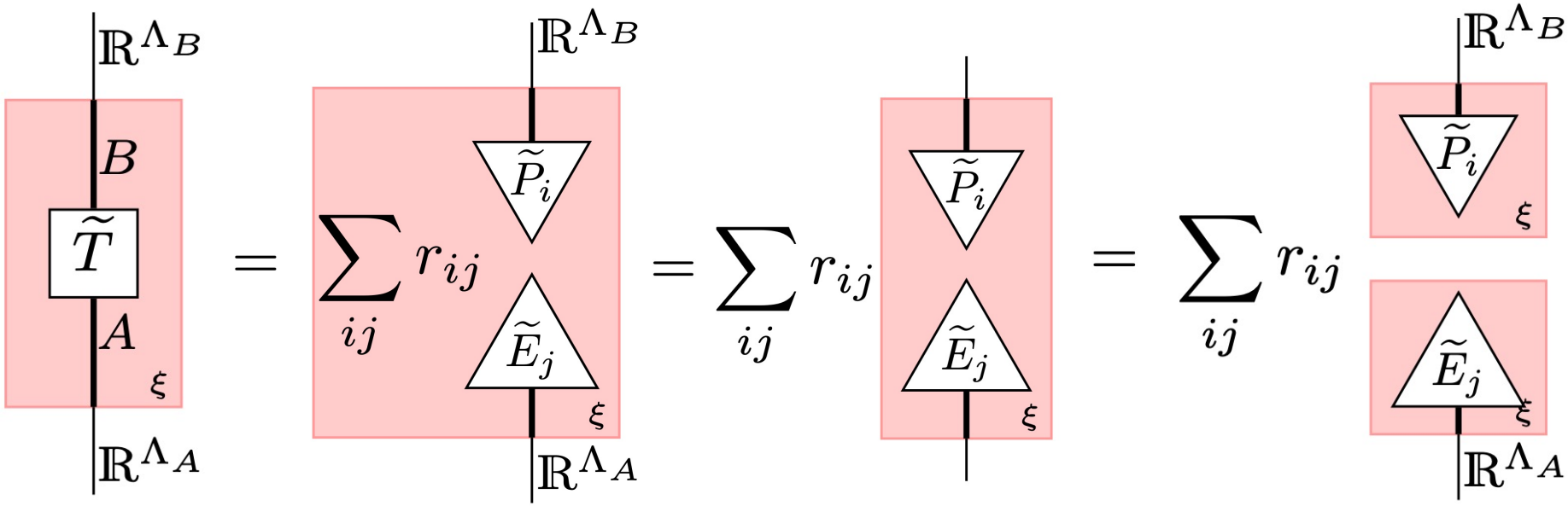
What about the representation of transformations?

For any GPT satisfying the principle of local tomography, one can decompose any transformation into a linear sum of states and effects in the GPT:



This is analogous to decomposing the identity operator on some Hilbert space…
-but the matrix M is not just the identity, since there is no orthonormal basis of density operators and POVM elements (e.g.)

arxiv:2005.07161

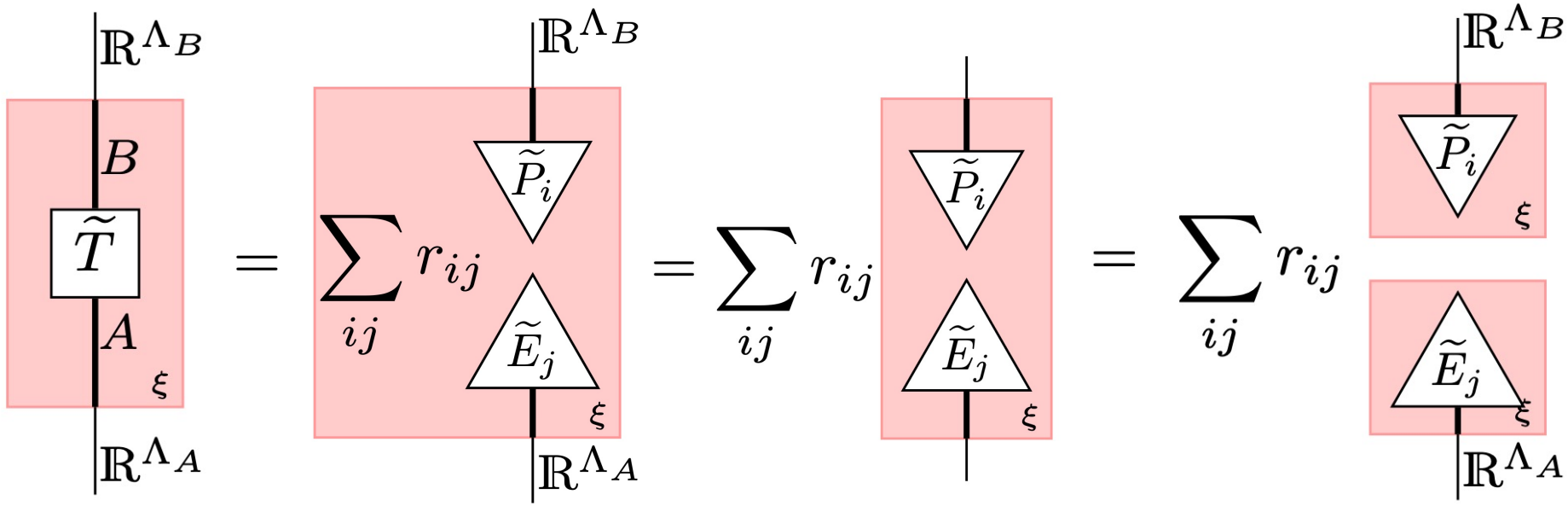


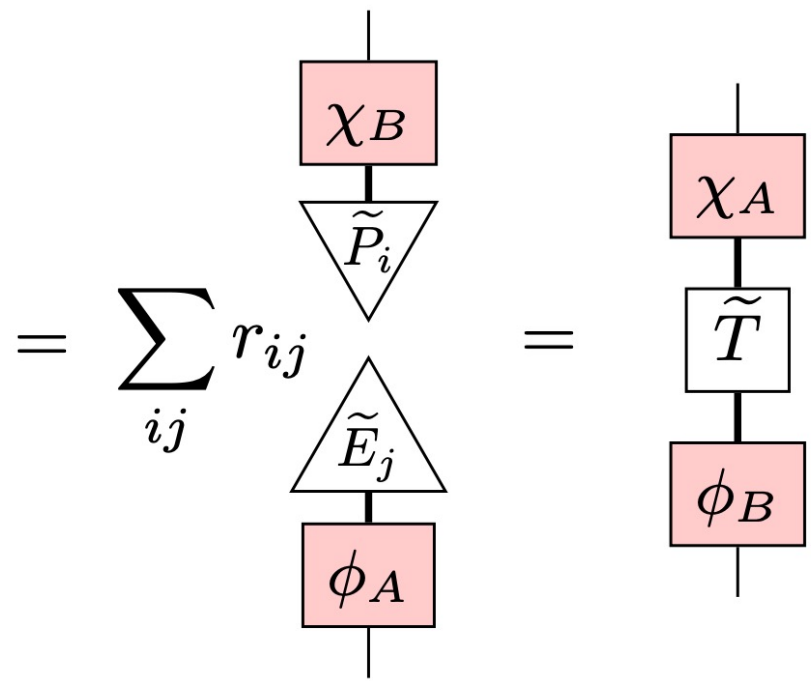
decomposition of T

linearity

diagram preservation

So the representation of transformations can be fully constrained by the representation of states and effects!

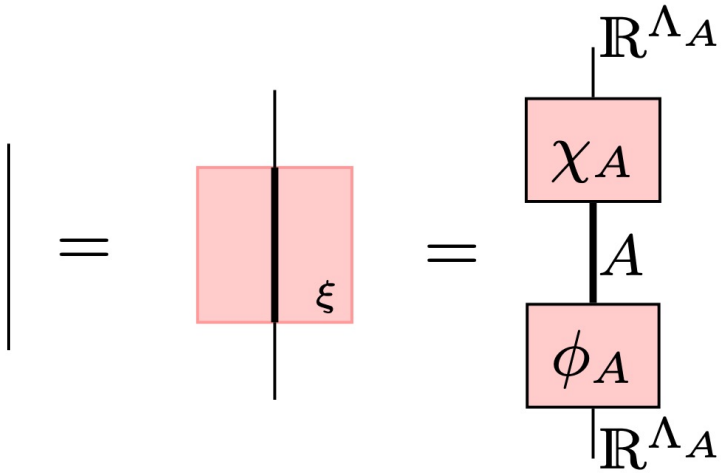




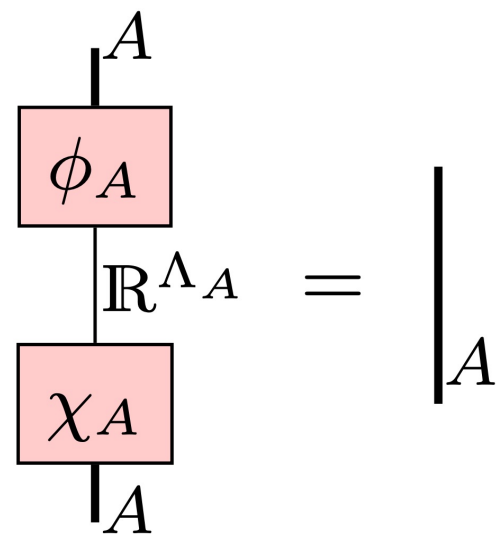
representation of DP map within RLinear

un-decomposition of T

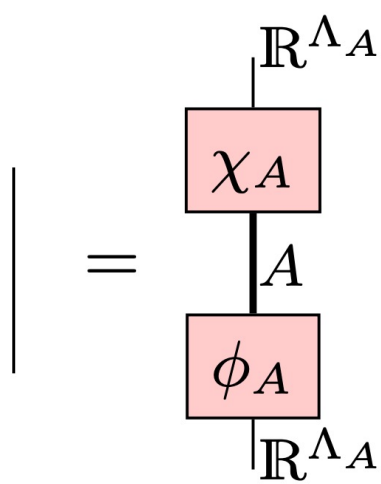
representation of identity:



and recall that



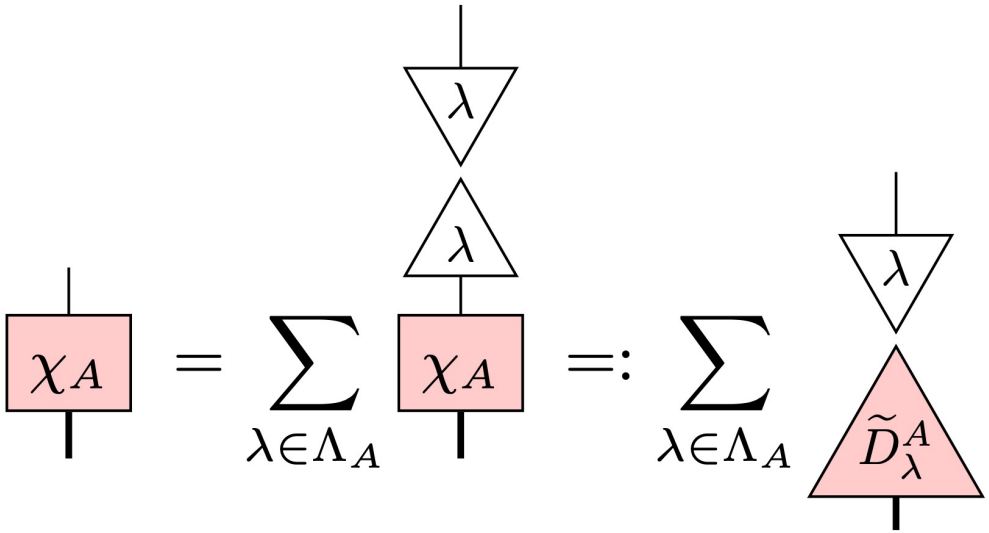
so

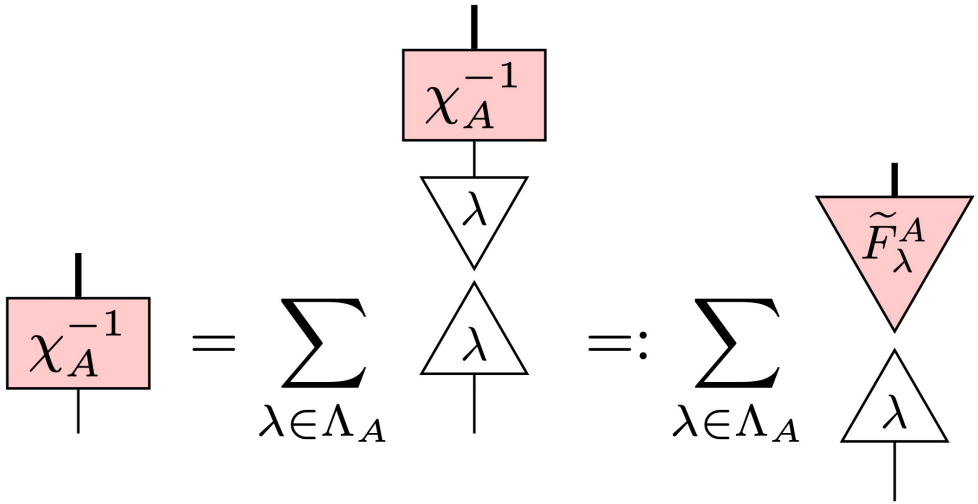


so χ and ϕ are each other's inverses

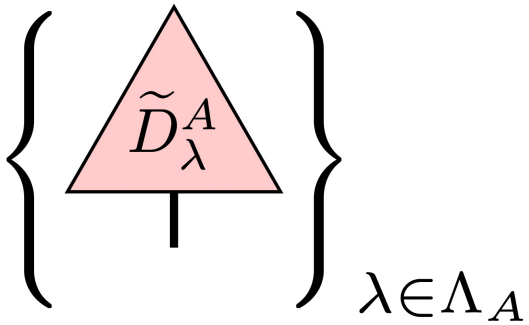
Finding an explicit form for χ and χ^{-1} in terms of a basis

insert an identity:

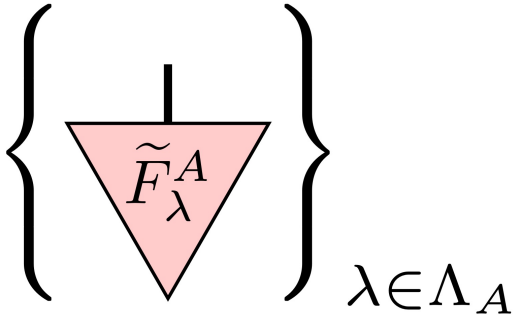




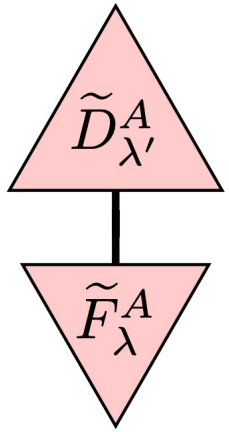
can check that the D's form a basis for the vector space of effects



Similarly for some set of F's that form a basis for the vector space of states

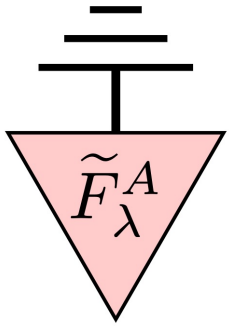


invertibility



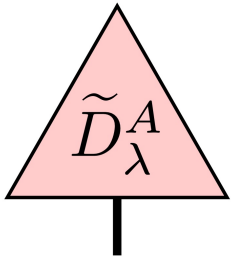
$$= \delta_{\lambda\lambda'}$$

representing ignoring correctly

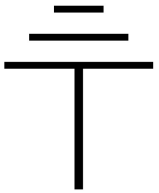


$$= 1$$

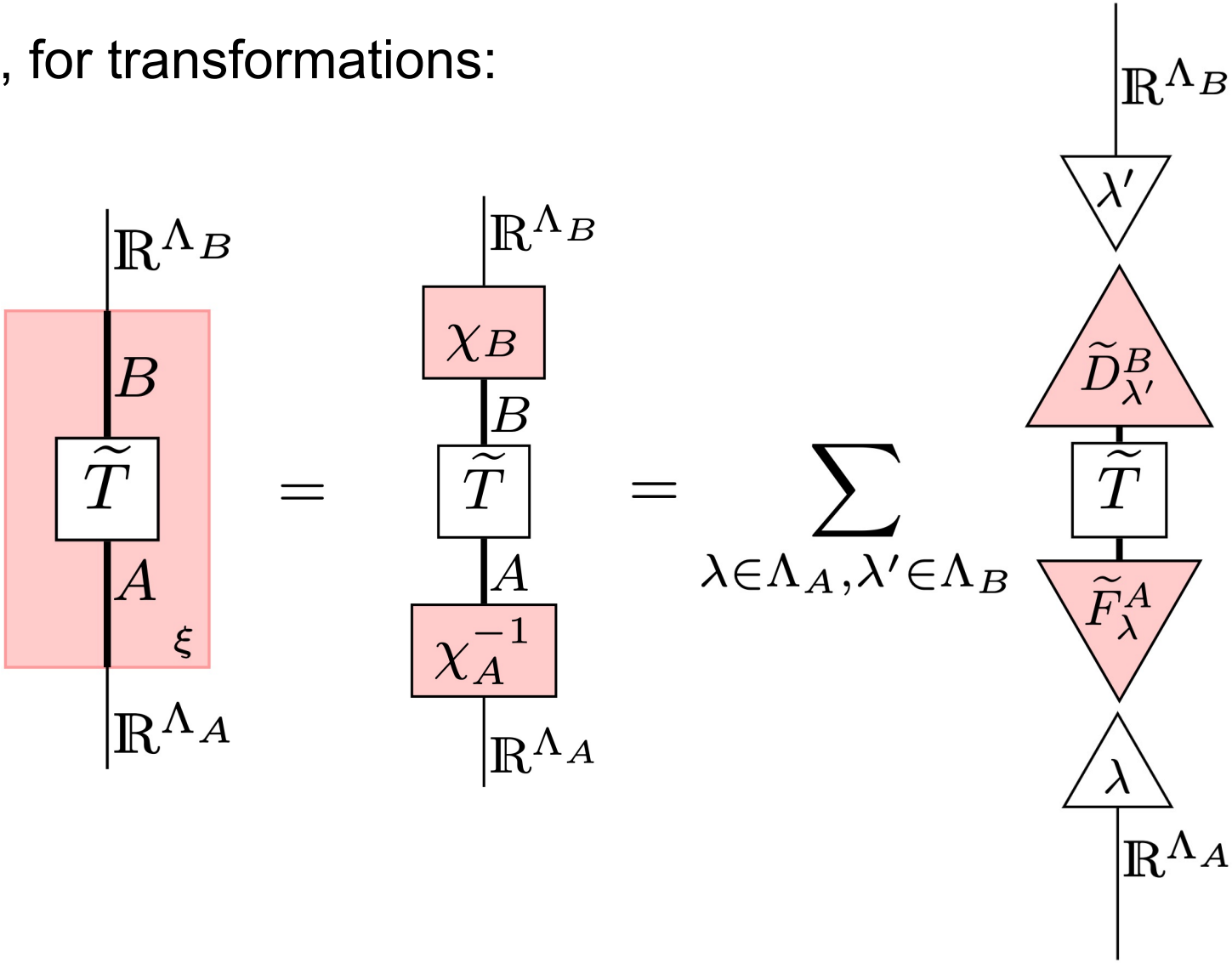
$$\sum_{\lambda \in \Lambda_A}$$



$$=$$



hence, for transformations:



Every diagram-preserving quasiprobability representation has this form

Quasiprobability representation of quantum theory:

pick any basis of trace-one Hermitian operators $\{F_\lambda\}_\lambda$

compute the unique dual basis satisfying $\mathrm{tr}[D_{\lambda'} F_\lambda] = \delta_{\lambda,\lambda'}$

states $$\xi_\rho(\lambda) = \mathrm{tr}[D_\lambda \rho]$$

effects $$\xi_E(\lambda) = \mathrm{tr}[F_\lambda E]$$

transformations $$\xi_{\mathcal{E}}(\lambda'|\lambda) = \mathrm{tr}[D_{\lambda'} \mathcal{E}(F_\lambda)]$$

“complete frame representation”

If the processes representing a given quantum scenario are all nonnegative, this constitutes an OM for it.

If the processes go negative, this constitutes a quasiprobability representation that is *not* an OM.

If there is no choice of basis for which the resulting processes are nonnegative, then the scenario does not admit of an OM, and it is not classically-explainable.

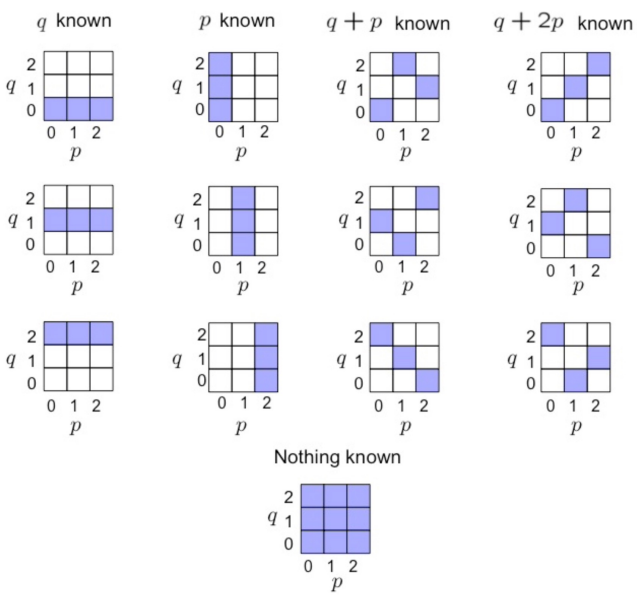
All noncontextual models of operational theories with transformations have this structure!

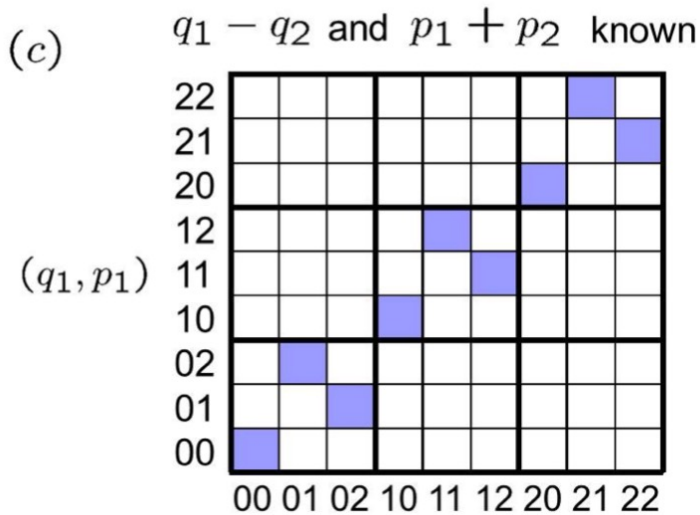
Example 1: Wigner representation

Example 2: Spekkens toy theory (for odd-dimensional stabilizer subtheories) has exactly this form, with basis:

$$\{A_{p,q}\}_{p,q} := \left\{ \frac{1}{d} \sum_{p,q} \exp(pq - p'q' + i\pi/d) W_{p,q}{}^{\dagger} \right\}$$

$$\xi_{\rho_{\mathrm{stab}}}(p,q) = \mathrm{tr}[A_{p,q}\rho_{\mathrm{stab}}]$$



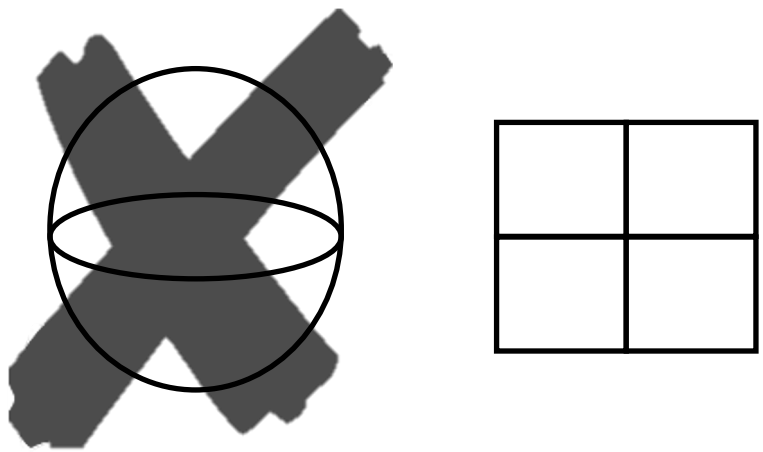


arXiv:2101.06263

The basis $\{F_\lambda\}$ is a set of linearly independent vectors spanning the GPT vector space

dimension of GPT = cardinality of ontic state space

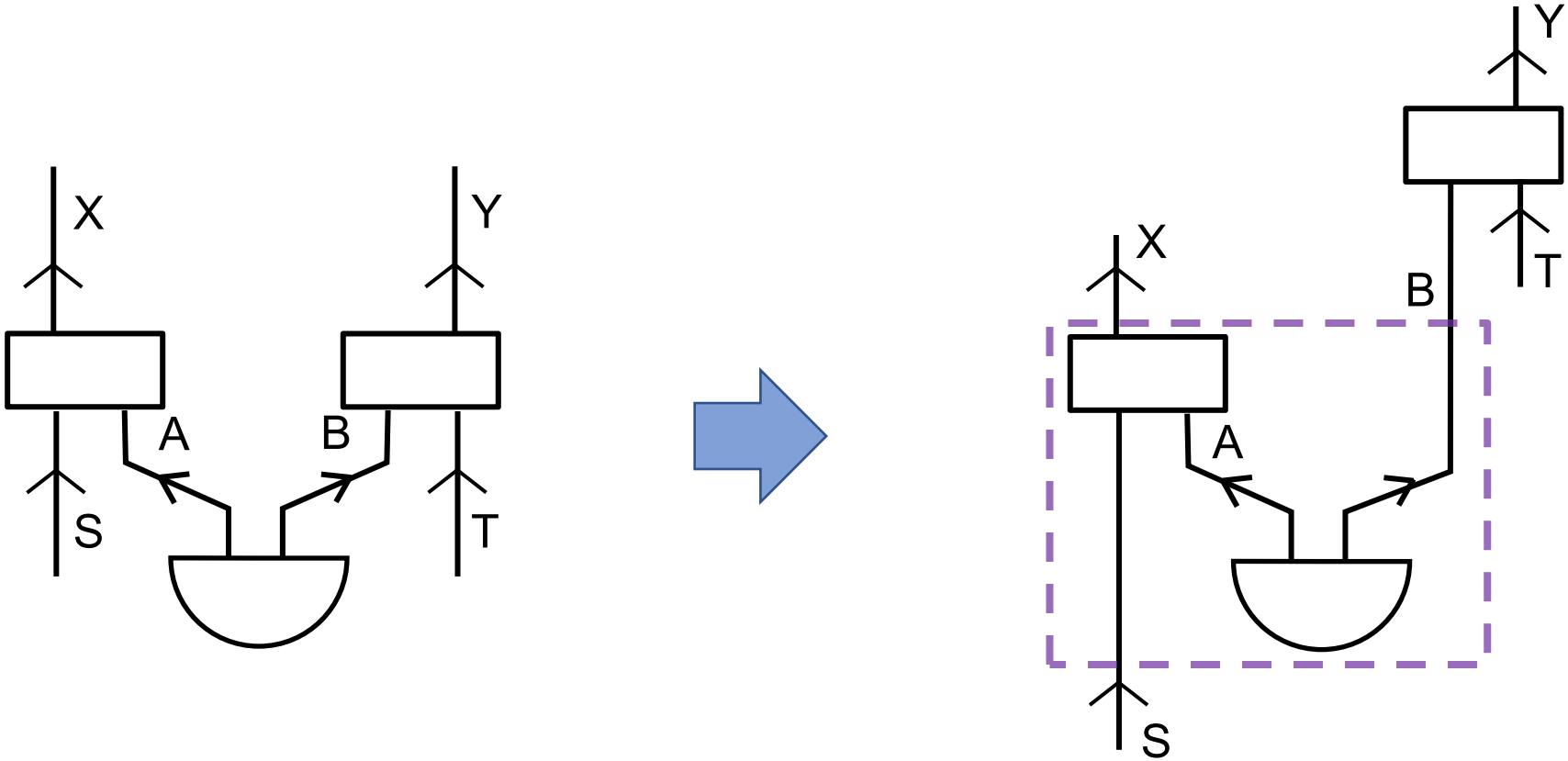
E.g. for a qubit, d = 4:



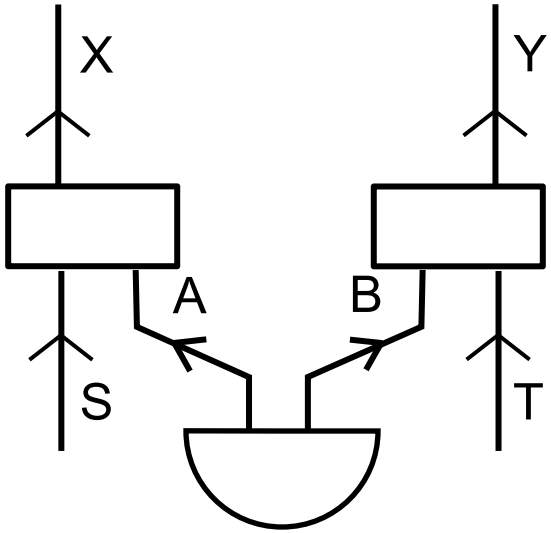
note: in prepare-measure scenarios, the best known bound on the number of ontic states is (GPT dim)^2

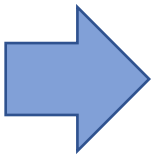
The relationship between noncontextuality and locality

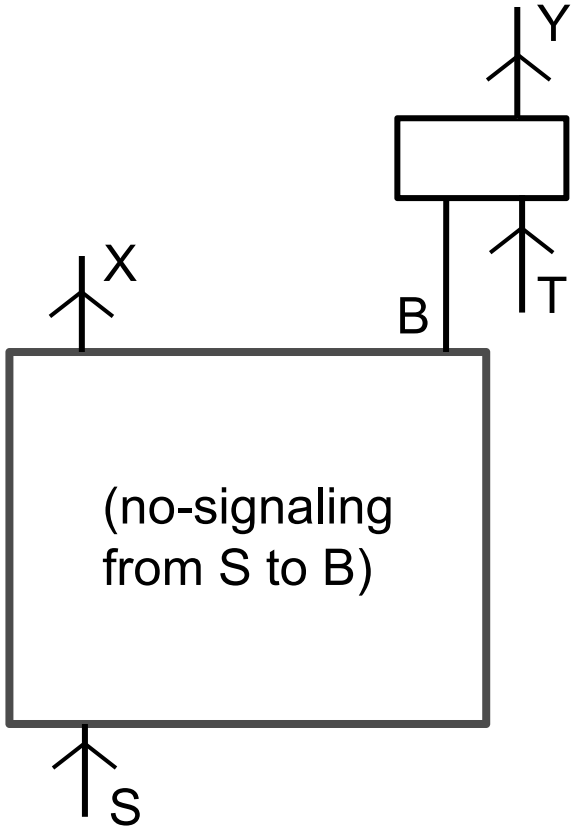
View the Bell scenario as a prepare-measure scenario:



View the Bell scenario as a prepare-measure scenario:



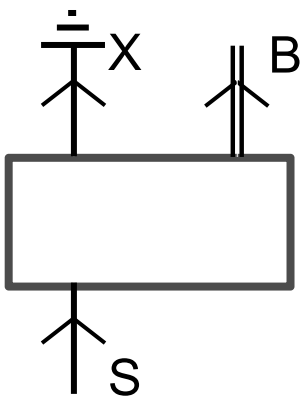




Locality implies an instance of PNC

PNC for this scenario with respect to the no-signaling operational equivalence implies P(XY|ST) satisfy Bell inequalities

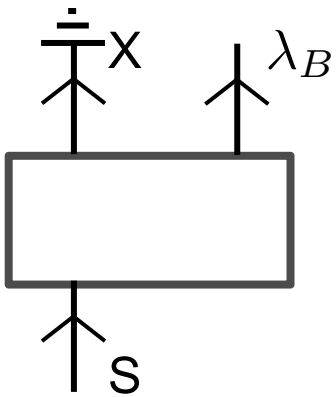
PNC implies Locality



The relevant operational equivalence:

$$\rho^{B}_{S=s} = \rho^{B}_{S=s'}$$

PNC implies Locality



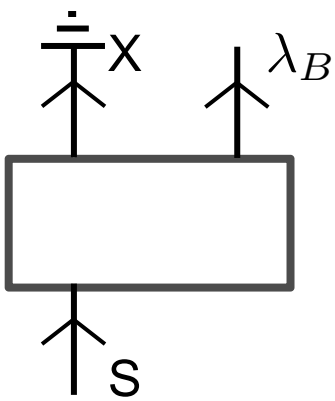
The relevant operational equivalence:

$$\rho^B_{S=s} = \rho^B_{S=s'}$$

PNC:

$$p(\lambda_B|S=s) = p(\lambda_B|S=s')$$

PNC implies Locality



The relevant operational equivalence:

$$\sum_X \rho^B_{X|S=s} p(X|S=s) = \sum_X \rho^B_{X|S=s'} p(X|S=s')$$

NC then implies:

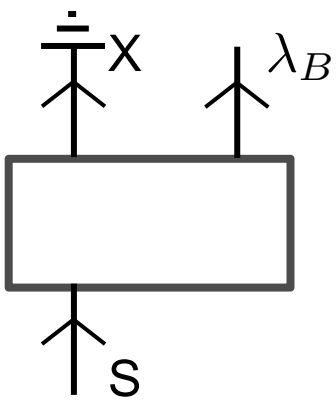
$$\sum_X \mu_{X|S=s}(\lambda_B) p(X|S=s) = \sum_X \mu_{X|S=s'}(\lambda_B) p(X|S=s')$$

$$\sum_X p(\lambda_B|X, S=s) p(X|S=s) = \sum_X p(\lambda_B|X, S=s') p(X|S=s')$$

$$\sum_X p(\lambda_B, X|S=s) = \sum_X p(\lambda_B, X|S=s')$$

$$p(\lambda_B|S=s) = p(\lambda_B|S=s') \qquad \forall s, s'$$

PNC implies Locality



The relevant operational equivalence:

$$\sum_X \rho^B_{X|S=s} p(X|S=s) = \sum_X \rho^B_{X|S=s'} p(X|S=s')$$

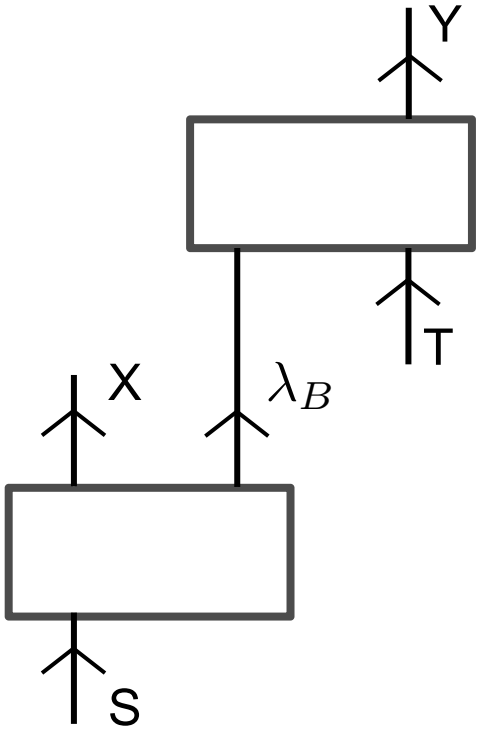
NC then implies:

$$\sum_X \mu_{X|S=s}(\lambda_B) p(X|S=s) = \sum_X \mu_{X|S=s'}(\lambda_B) p(X|S=s')$$

$$\sum_X p(\lambda_B|X, S=s) p(X|S=s) = \sum_X p(\lambda_B|X, S=s') p(X|S=s')$$

$$\sum_X p(\lambda_B, X|S=s) = \sum_X p(\lambda_B, X|S=s')$$

PNC: $\boxed{p(\lambda_B|S) = p(\lambda_B)}$



PNC implies Locality

Now write the empirical predictions as

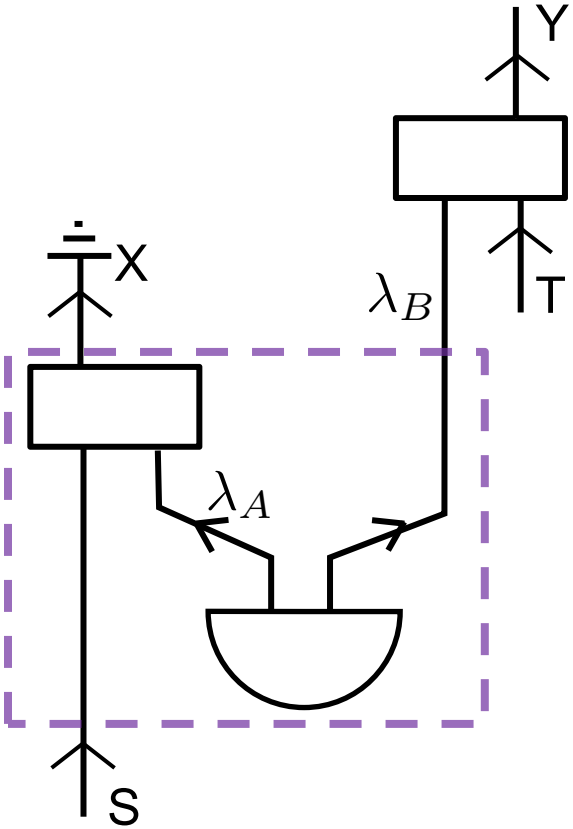
$$p(XY|ST) = \sum_{\lambda_B} p(Y|T,\lambda_B)p(\lambda_B, X|S)$$

$$= \sum_{\lambda_B} p(Y|T,\lambda_B)p(X|\lambda_B, S)p(\lambda_B|S)$$

$$= \sum_{\lambda_B} p(Y|T,\lambda_B)p(X|\lambda_B, S)p(\lambda_B)$$

$\Rightarrow$ Bell inequalities!!

PNC: $\boxed{p(\lambda_B|S) = p(\lambda_B)}$



Locality implies PNC (wrt the no-signaling OE)

$$p(X, \lambda_B|S) = \sum_{\lambda_A} p(X|S, \lambda_A)p(\lambda_A, \lambda_B)$$

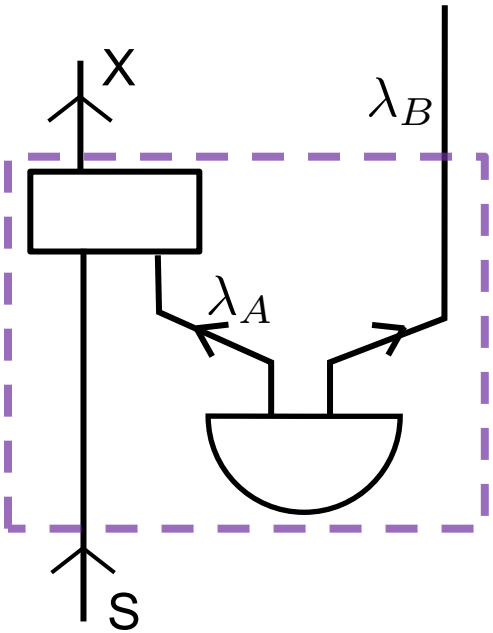
$$\sum_X p(X, \lambda_B|S) = \sum_X \sum_{\lambda_A} p(X|S, \lambda_A)p(\lambda_A, \lambda_B)$$

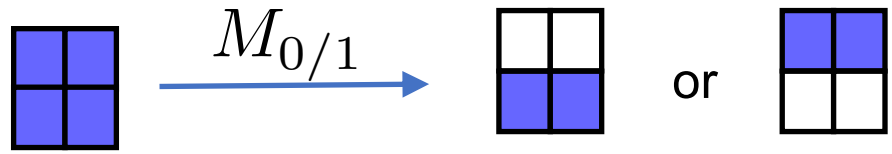
$$p(\lambda_B|S) = \sum_{\lambda_A} p(\lambda_A, \lambda_B)$$

$$= p(\lambda_B)$$

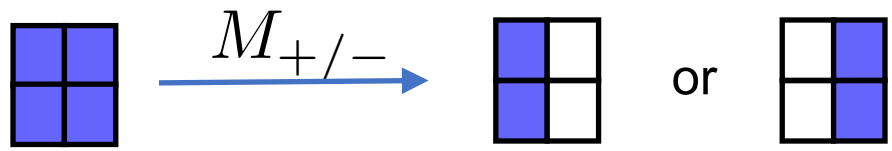
PNC: $\boxed{p(\lambda_B|S) = p(\lambda_B)}$

Spekkens toy theory example:





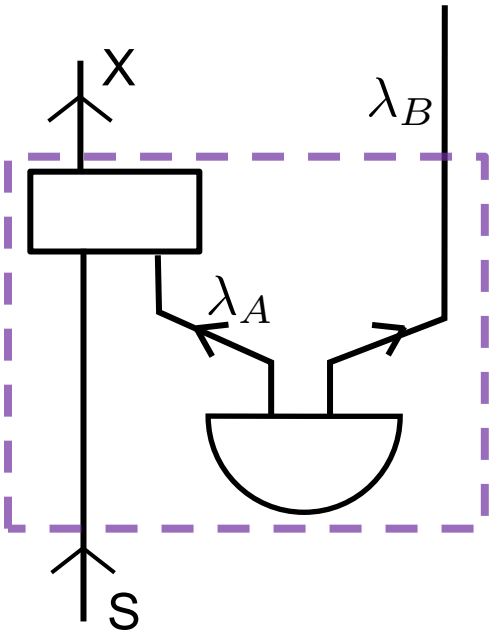
marginal state …conditioned on X

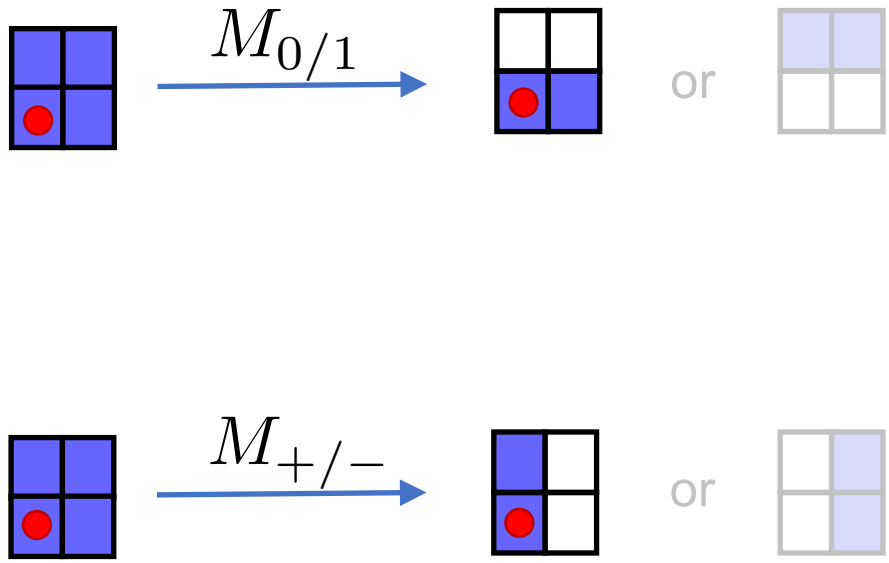


different measurements just give different ways of resolving one's uncertainty about A… and also about B, since the two are perfectly correlated

-the true ontic state of B doesn't change at all

Spekkens toy theory example:

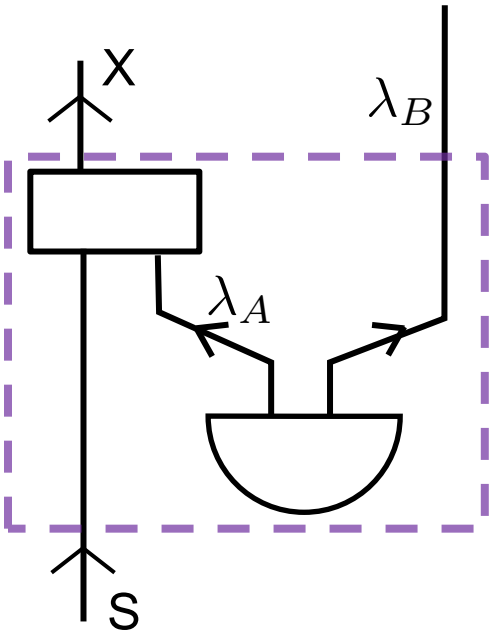




different measurements just give different ways of resolving one's uncertainty about A… and also about B, since the two are perfectly correlated

-the true ontic state of B doesn't change at all
-only one's knowledge changes
-inference, not influence!

compare to quantum theory:



$$\frac{\mathbb{I}}{2} \xrightarrow{M_{0/1}} |0\rangle \text{ or } |1\rangle$$

where $\frac{\mathbb{I}}{2} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|$

$$\frac{\mathbb{I}}{2} \xrightarrow{M_{+/-}} |+\rangle \text{ or } |-\rangle$$

where $\frac{\mathbb{I}}{2} = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|$

-inference, not influence?
-this suggests that the wavefunction is epistemic, since it is updated in a manner analogous to Bayesian updating of probability distributions

So, no noncontextual ontological model describes our world. Now what?

https://pirsa.org/20110051

arXiv:2009.03297

Abandon Leibniz's principle? No.

Give up on realism? No.

We need to revise our notion of realism!
(i.e., the ontological models framework)

So let me say a few words about our nascent approach to constructing a new kind of realist interpretation of quantum theory

https://pirsa.org/20110051

arXiv:2009.03297

Guiding principles:

locality
noncontextuality
unscrambling causation and inference

These two distinguish our approach from other nonclassical realist approaches:

-Many worlds
-Consistent histories
-QBism
-etc

arXiv:2009.03297

Classical causal models ⇨ **Classical realist theory** (ontological models) ⇦ **Classical probability theory**

Nonclassical theory of causation ⇨ **Nonclassical realist theory** ⇦ **Nonclassical theory of inference**

arXiv:2009.03297

$$\frac{\text{Nonclassical notions of Causation and Inference}}{\text{Classical notions of Causation and Inference}} \sim \frac{\text{Relativistic notions of Space and Time}}{\text{PreRelativistic notions of Space and Time}}$$

The notion of realism we seek to salvage:

Statistical correlations have causal explanations

The problem: Bell and noncontextuality no-go theorems

The idea:

By allowing nonclassical notions of causation and inference, one can provide satisfactory causal explanations while salvaging the spirit of locality and noncontextuality

arXiv:2009.03297