Generalized probabilistic theories

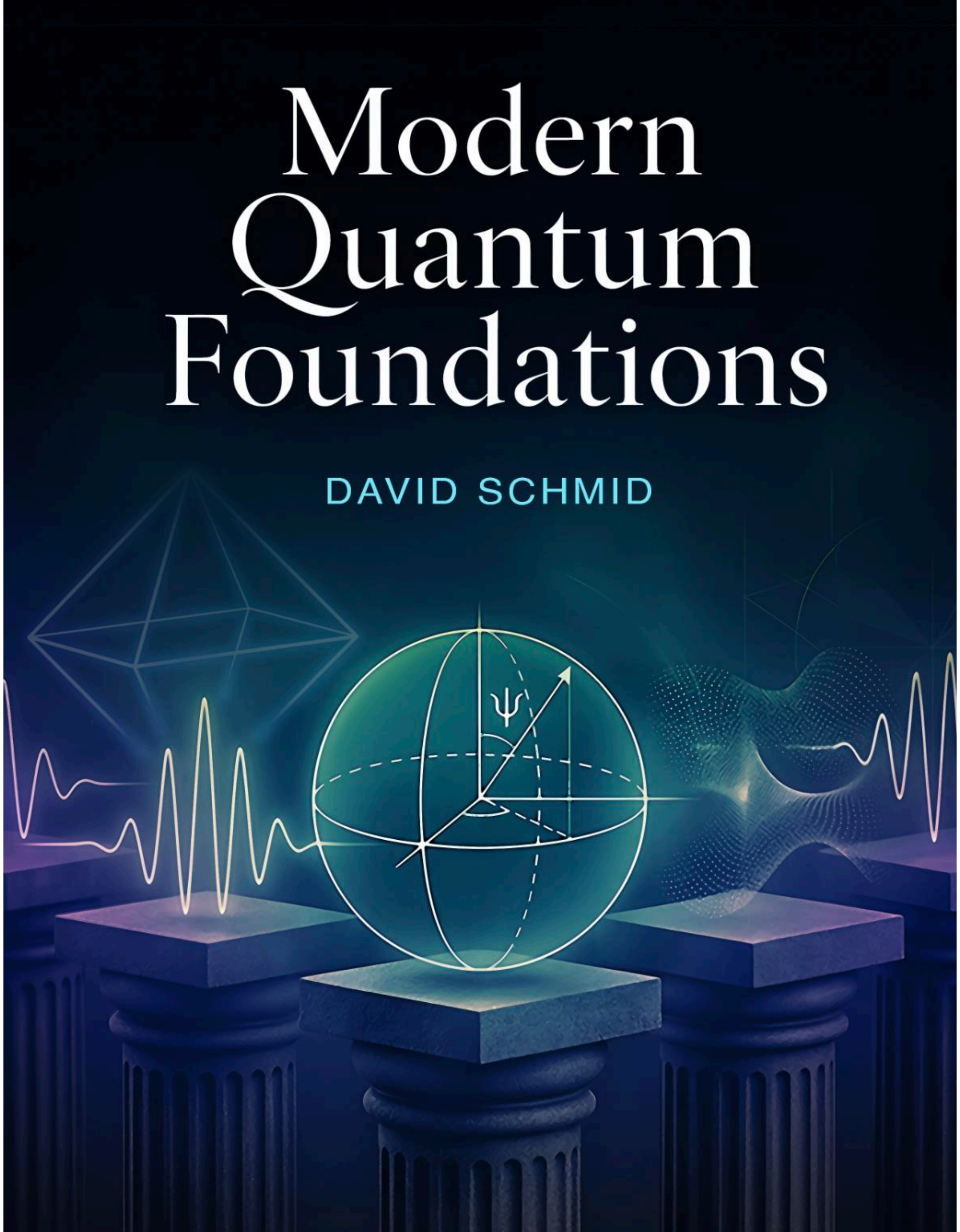
David Schmid

dschmid1@perimeterinstitute.ca

PERIMETER INSTITUTE

Textbook coming in ~1 month!

If you want me to email you a copy, message me at dschmid1@perimeterinstitute.ca



Background: the Bloch ball

(a representation of the qubit as real-valued vectors)

-density operators/POVM elements are Hermitian operators

-Hermitian operators form a real Euclidean vector space

$A = A^\dagger \qquad B = B^\dagger$

$aA + bB$ is Hermitian

Inner product = Tr[AB]

ON basis: $\{\sigma_i/\sqrt{2}\}$ = $\left\{\frac{1}{\sqrt{2}}\mathbb{I},\ \frac{1}{\sqrt{2}}X,\ \frac{1}{\sqrt{2}}Y,\ \frac{1}{\sqrt{2}}Z\right\}$

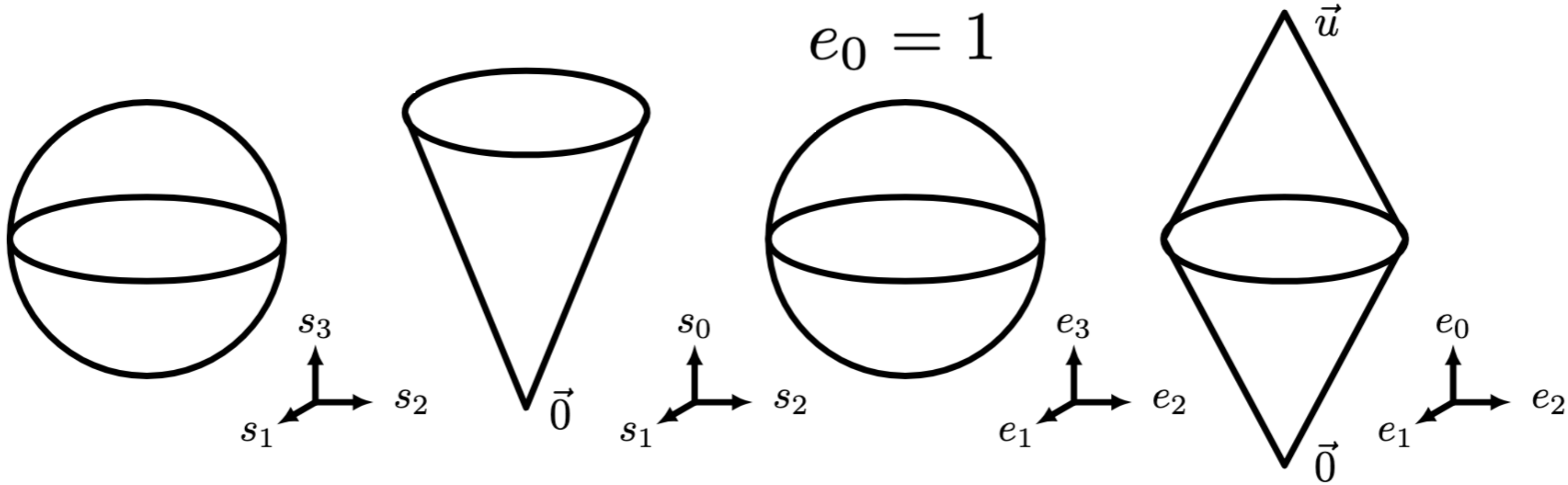
$$\rho = \sum_{i=0}^{3}\left\langle \rho, \frac{1}{\sqrt{2}}\sigma_i \right\rangle \frac{1}{\sqrt{2}}\sigma_i$$
$$= \frac{1}{2}\sum_{i=0}^{3}\mathrm{Tr}\left[\rho\,\sigma_i\right]\,\sigma_i$$
$$= \frac{1}{2}\left(s_0\mathbb{I} + s_1X + s_2Y + s_3Z\right)$$

$$E = \sum_{i=0}^{3}\left\langle E, \frac{1}{\sqrt{2}}\sigma_i \right\rangle \frac{1}{\sqrt{2}}\sigma_i$$
$$= \frac{1}{2}\sum_{i=0}^{3}\mathrm{Tr}\left[E\,\sigma_i\right]\,\sigma_i$$
$$= \frac{1}{2}\left(e_0\mathbb{I} + e_1X + e_2Y + e_3Z\right)$$

$$\mathrm{Tr}(\rho E) = \frac{1}{2}\left(s_0e_0 + s_1e_1 + s_2e_2 + s_3e_3\right) = \frac{1}{2}\vec{s}\cdot\vec{e}$$

$\mathrm{Tr}[\rho] = 1 \qquad \rho \geq 0 \qquad 0 \leq E \leq \mathbb{I}$

$s_0 = 1 \qquad |\vec{s}| \leq 1 \qquad |\vec{e}| \leq \min\{e_0, 1 - e_0\} \quad \text{with } 0 \leq e_0 \leq 1$



normalized states

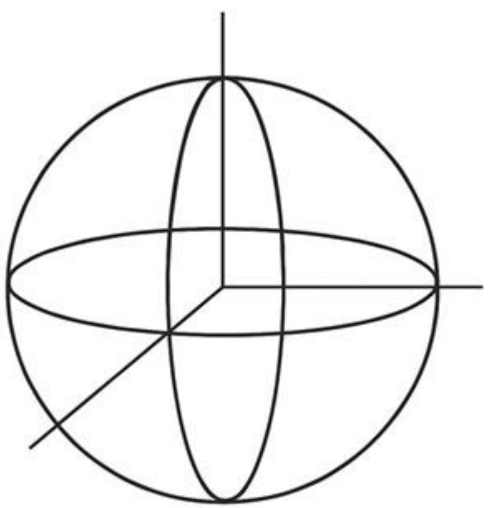
subnormalized states

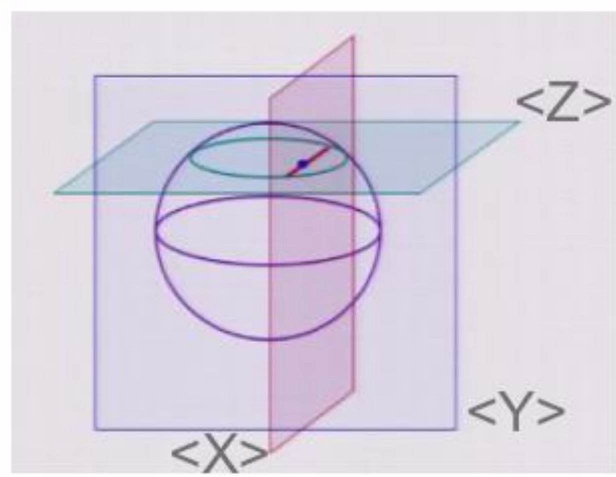
rank-1 effects

also other effects

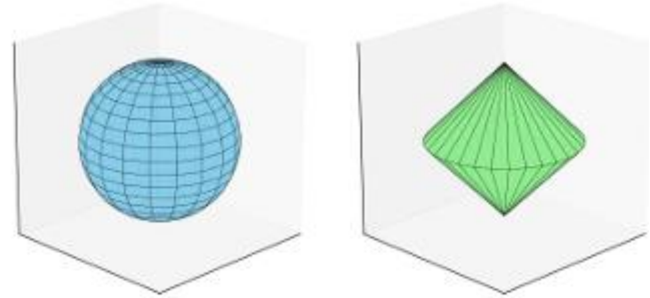
Many operational features of QT are encoded in the geometry

- -all the probabilities
- -uncertainty relations $\langle X\rangle^2 + \langle Y\rangle^2 + \langle Z\rangle^2 \leq 1,$
- -number of pure states
- -possibility of continuous reversible transformations
- -ambiguity of mixtures
- -tomographic completeness of Pauli measurements

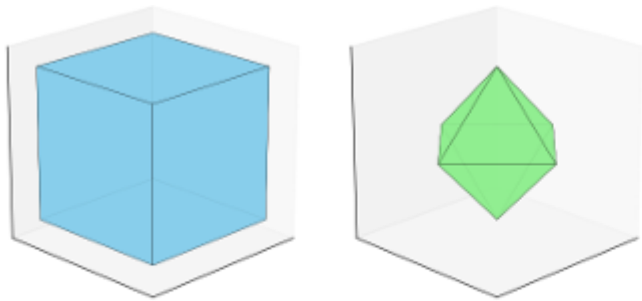




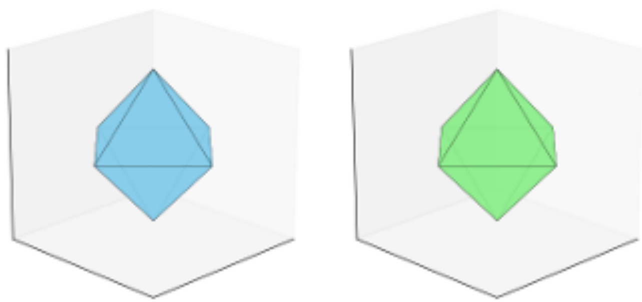
Qubit



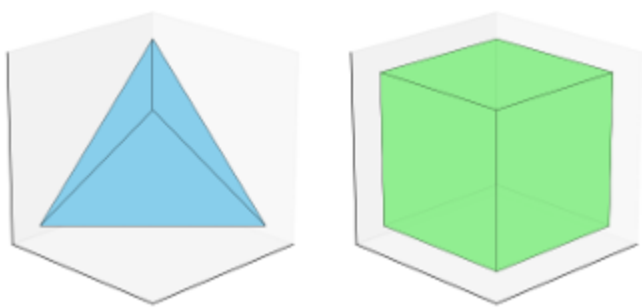
Boxworld



Spekkens toy theory

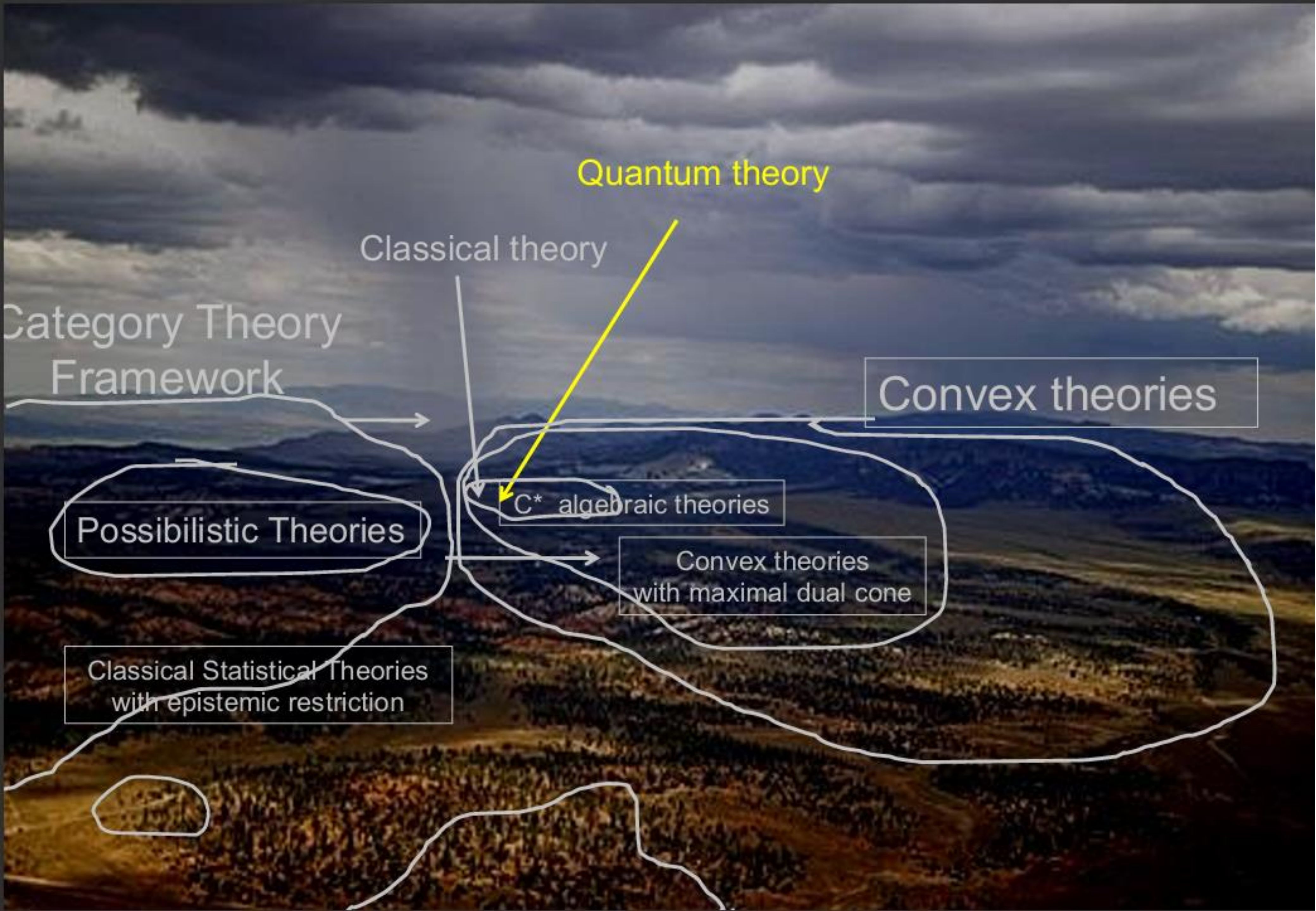


Classical theory (d=4)



Changing the geometry changes all of these operational features of a theory!

Generalized Probabilistic Theories


Quantum theory
Classical theory
Category Theory
Framework
Convex theories
C* algebraic theories
Possibilistic Theories
Convex theories
with maximal dual cone
Classical Statistical Theories
with epistemic restriction

Methodology of foil theories:
“understanding quantum theory by contrast to other (hypothetical) theories”

- -what features do all theories share
- -what features are quantum-specific
- -what features of a given theory logically imply what other features
- -reducing the number of postulates
- -finding axioms that refer to operational facts
  (rather than mathematics or ontology)
- -laying out possibilities for postquantum theories
- -what parts of a given theory are classically explainable

a GPT = an informational/predictive/operational description of a possible world/theory.

- -electron spin and photon polarization will look identical
- -no claim about how to associate GPT states to properties
- -etc

Very few metaphysical and ontological commitments.
-although *some* such commitments are unavoidable:

- -carving up the world into systems
- -scope of possible prepns/mmts
- -etc

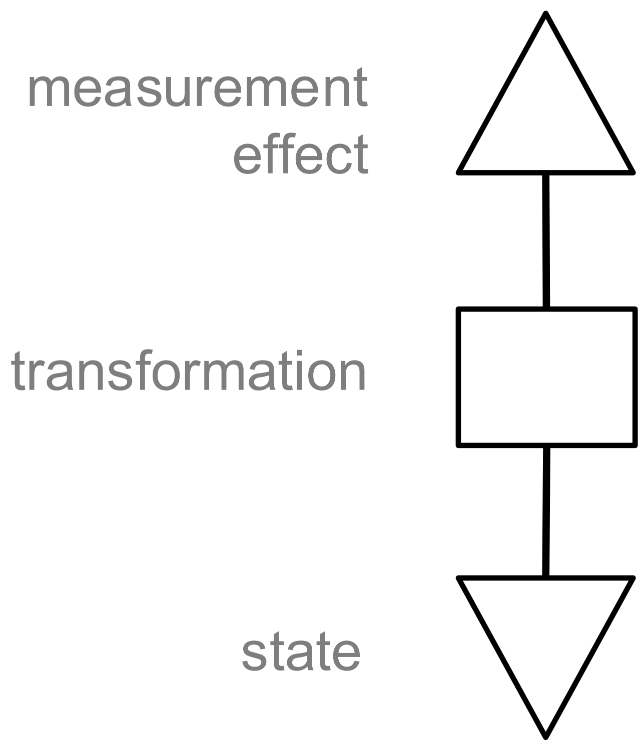
Generalized Probabilistic Theories

collection of systems

collection of processes

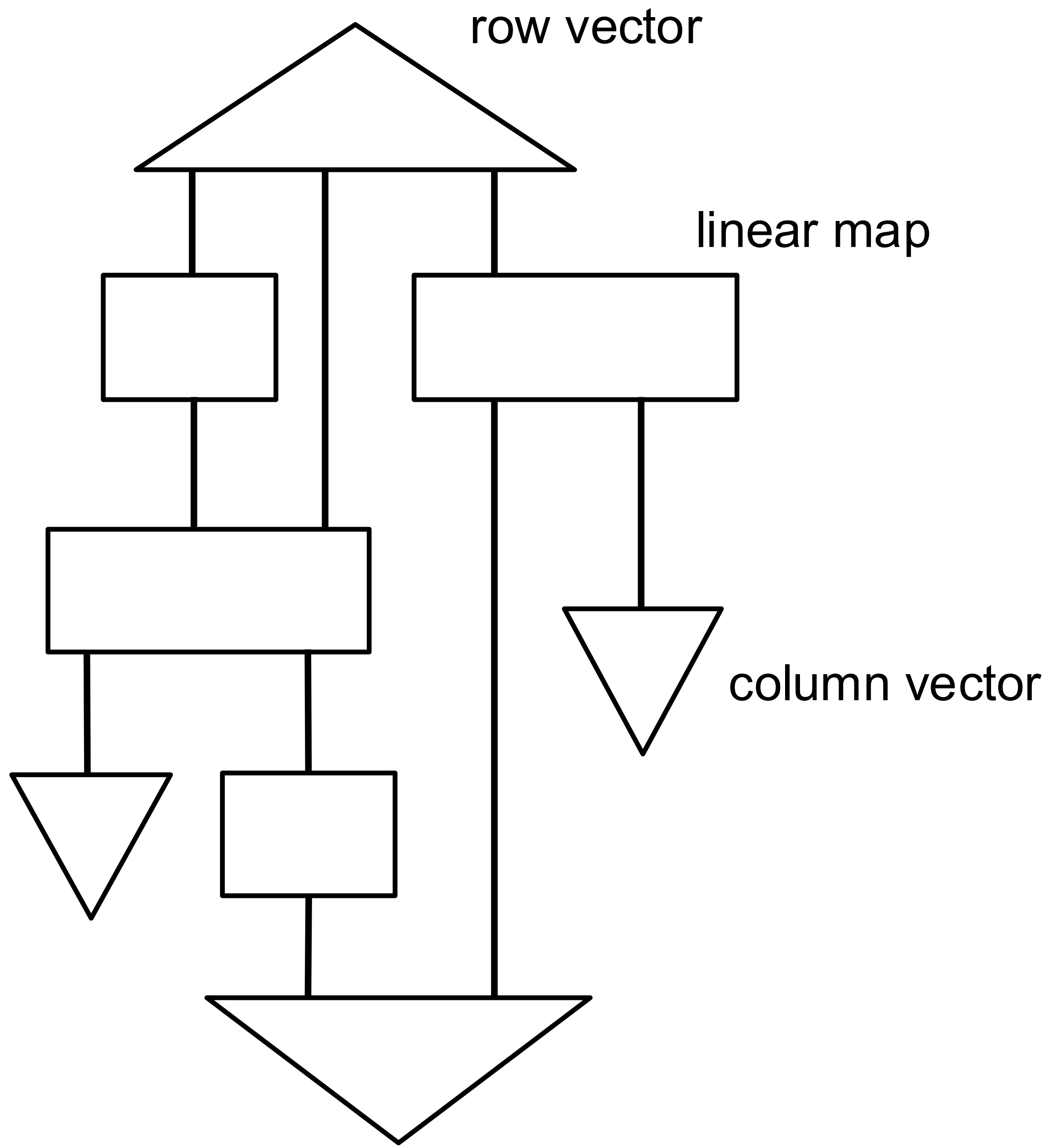
composition rule

⇒ observable probabilities



= prob(effect|transformed state)

Can generate arbitrary circuits/experiments by composition:



(But we will focus mostly on states and effects, and their geometry)

Different GPTs are defined by their:

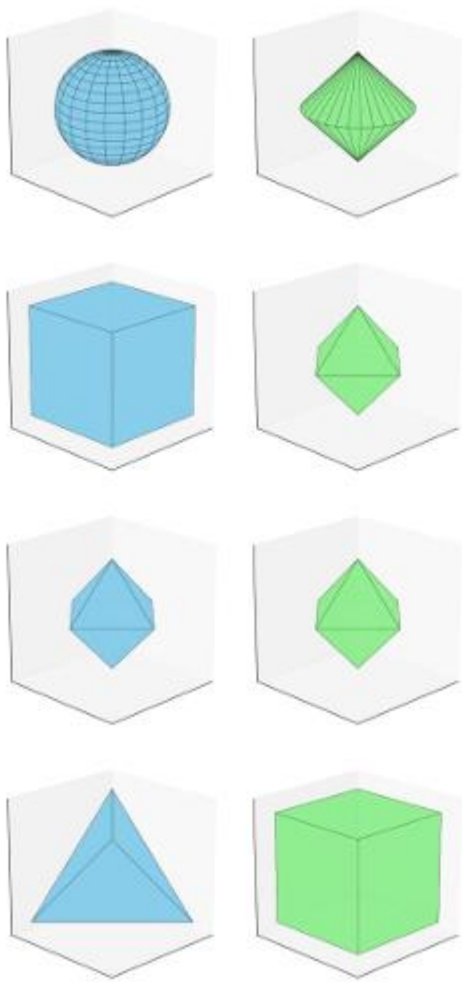
1. Convex geometry

qubit

Boxworld

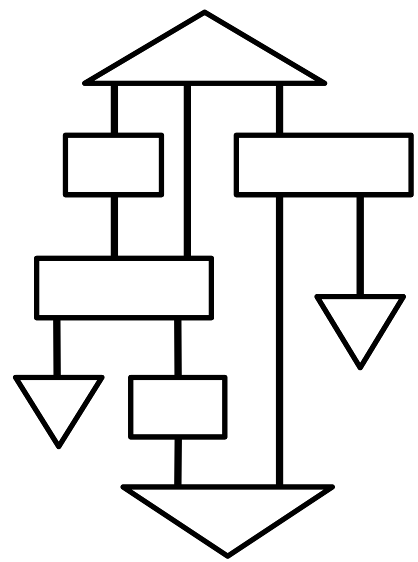
Spekkens toy theory

classical theory

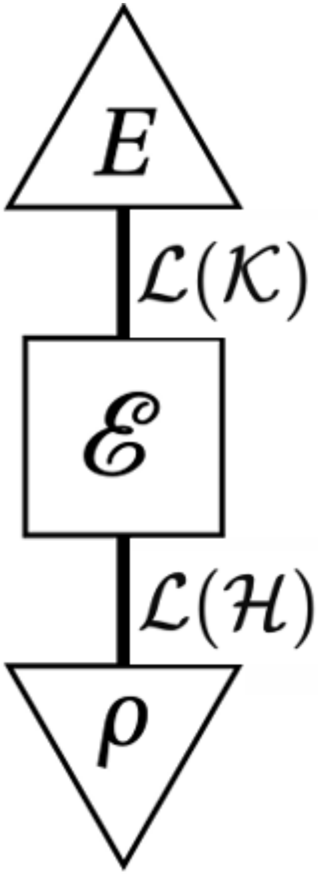


2. Compositional structure

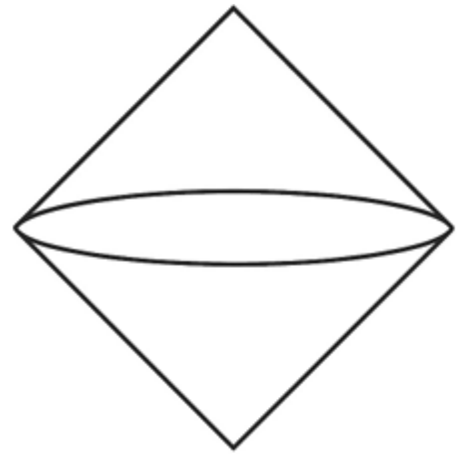
-multipartite states
-multipartite effects
-allowed transformations
-probabilities = inner product
-etc



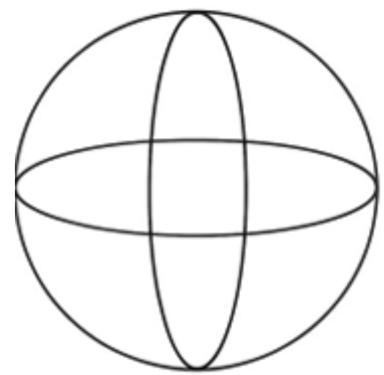
Quantum theory as a GPT



$$= \; e_E \cdot T_{\mathcal{E}}(s_\rho)$$



e_E



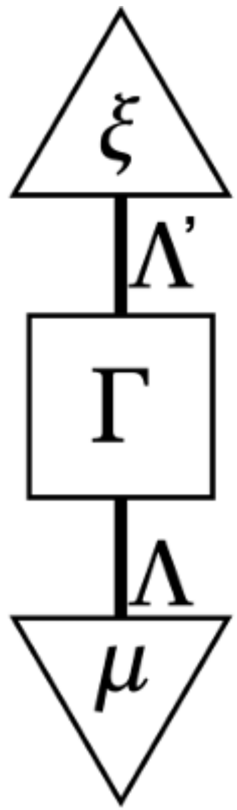
s_ρ

The classical theory as a GPT

response function

stochastic map

probability distribution



$$= \sum_{\lambda,\lambda'} \xi(\lambda')\Gamma(\lambda'|\lambda)\mu(\lambda)$$

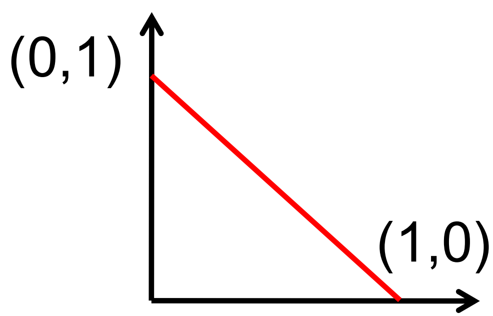
Sets (random variables)

s can be any probability distribution

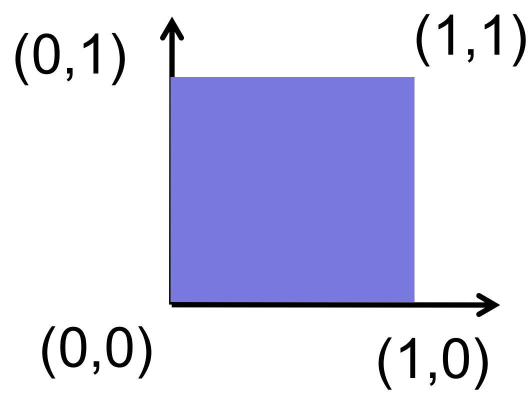
e can be any vector of conditional probabilities (a vector with elements in [0,1])

d=2

$$s = (p(1),\, p(2))$$



$$e = \big(p(e|1),\, p(e|2)\big)$$

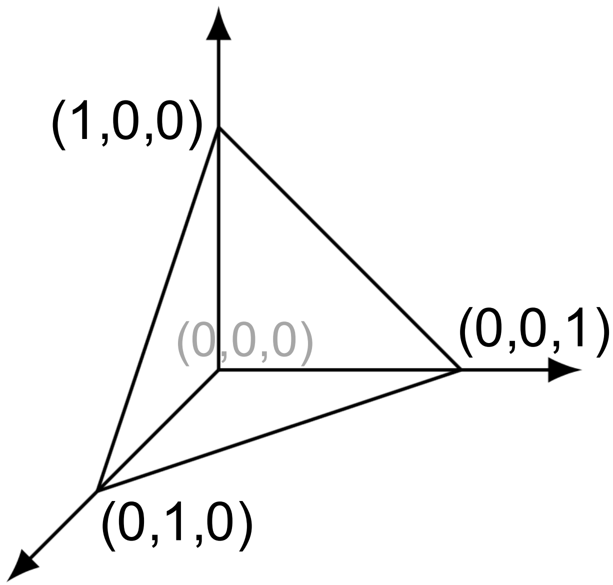


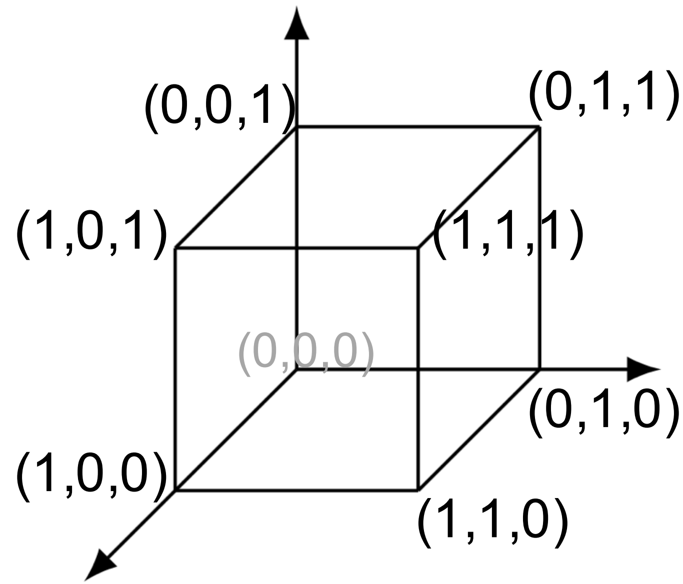
$$p(e) = \sum_{i=1}^{2} p(e|i)\, p(i) = e \cdot s$$

state space: simplex

effect space: dual of simplex

d=3

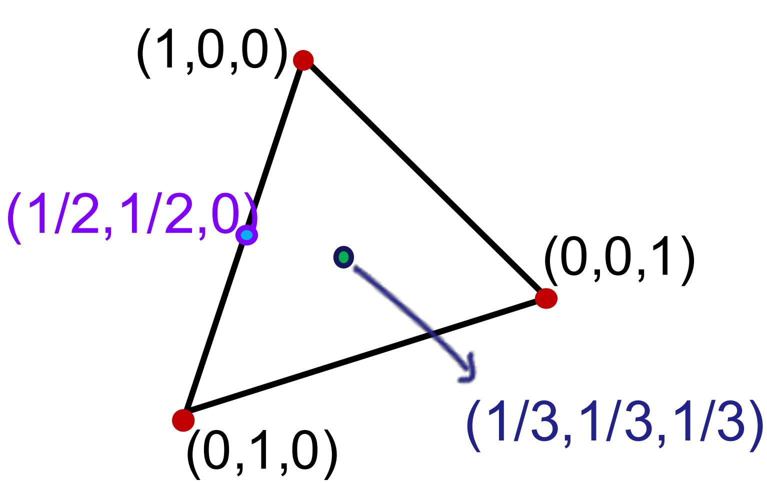


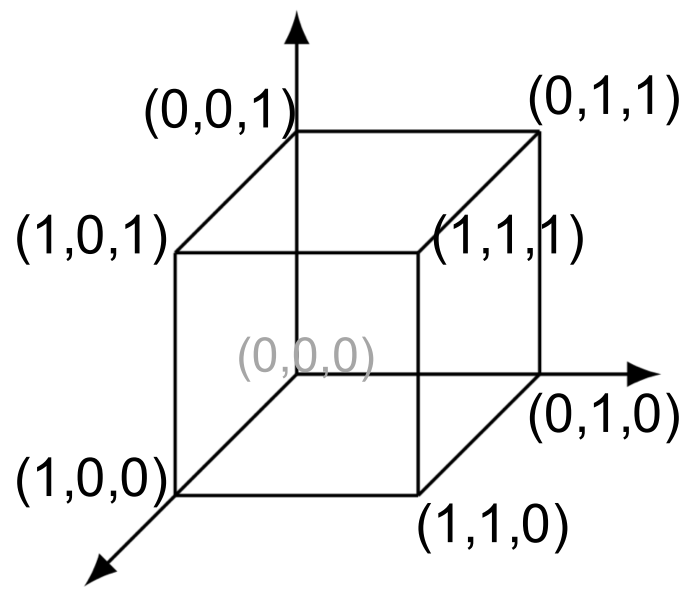


normalized states

effects

d=3



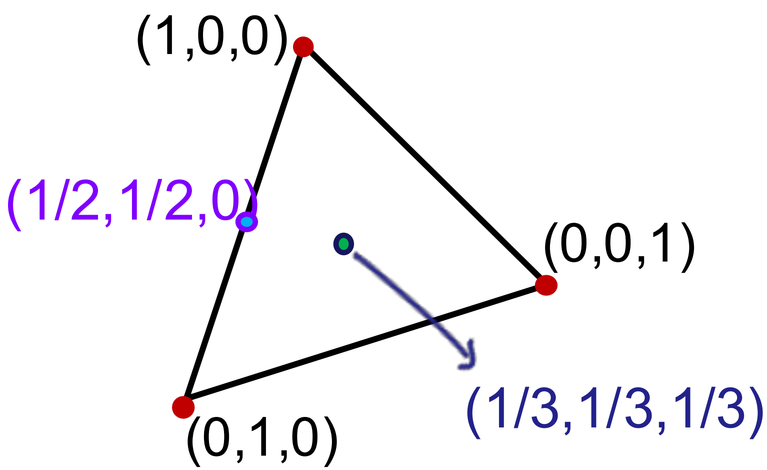


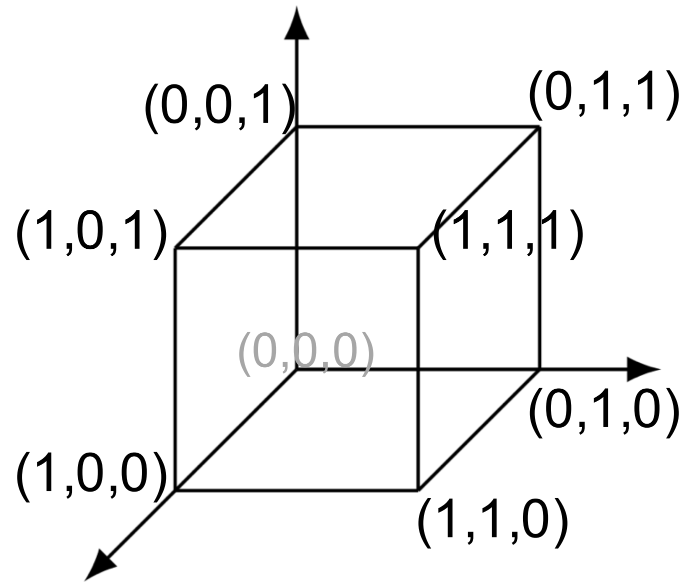
Classical statistical theory: probability distributions over a set of classical states

normalized states

effects

d=3

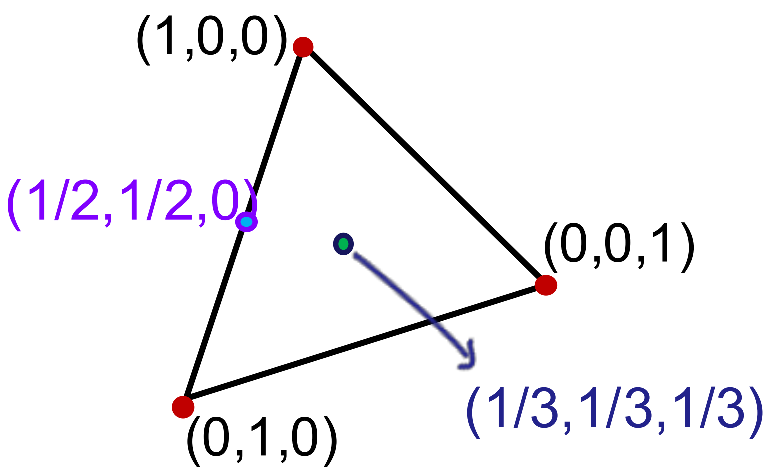




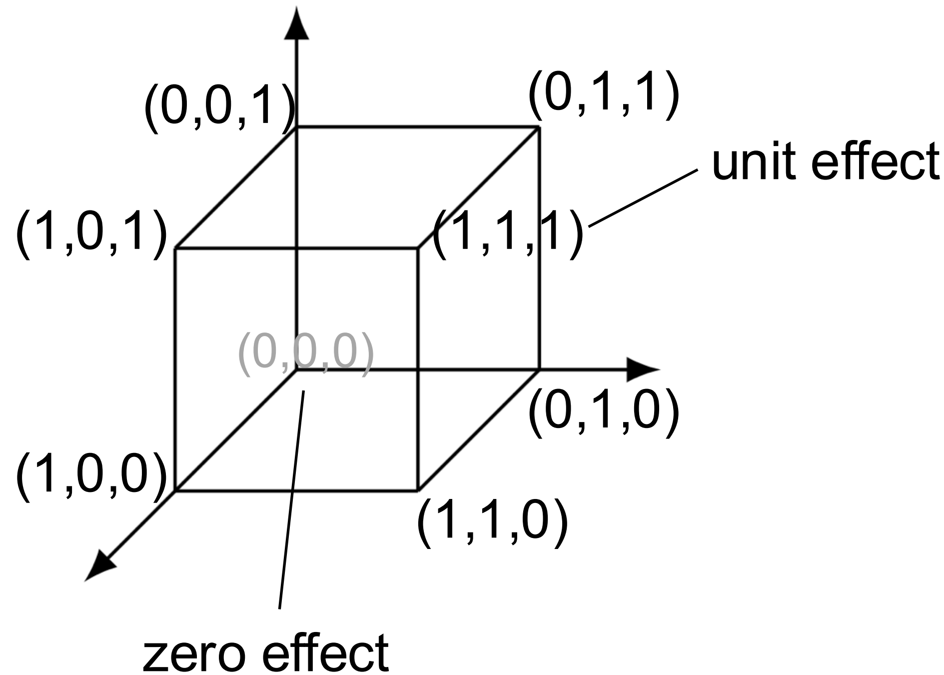
Every mixed state decomposes into pure states in a unique way
⇒ One can always imagine that there is a true state of the system, and any mixed state can be *uniquely* interpreted as uncertainty about the true state.

d=3

normalized states

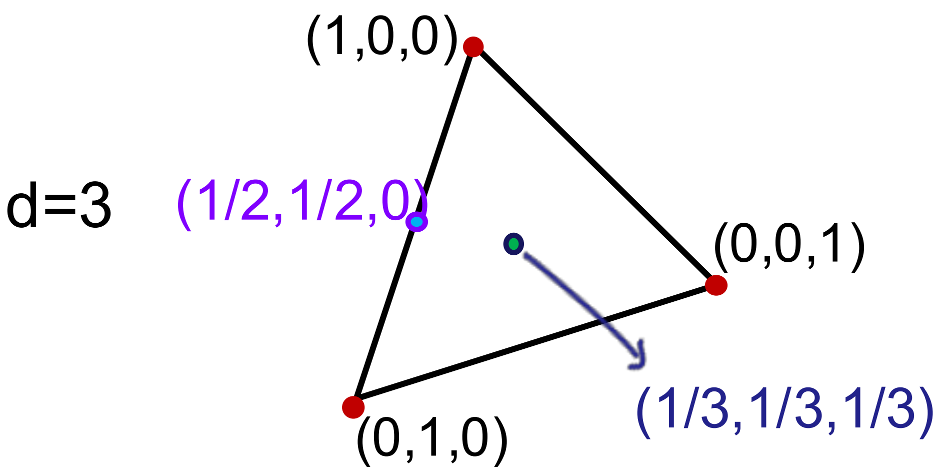


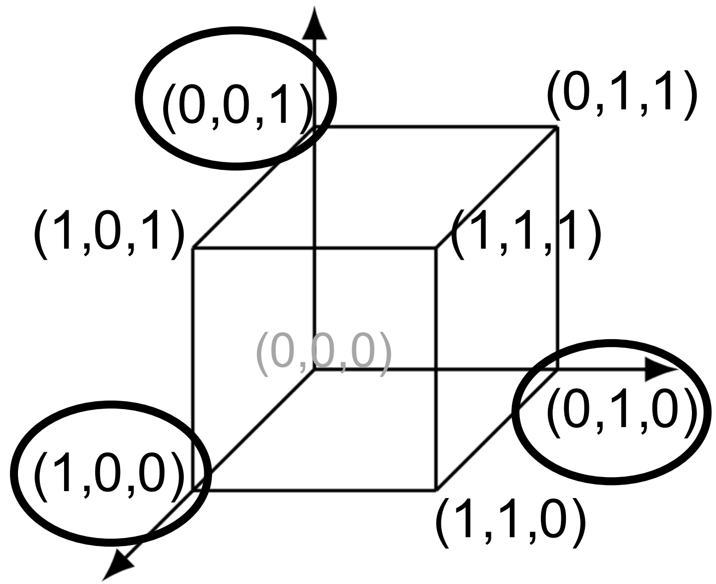
effects



normalized states

effects





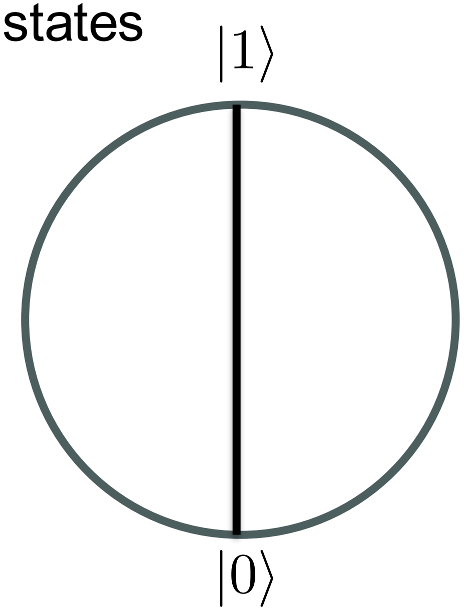
One can determine the exact state of the system in a single measurement.
All other measurements can be obtained by post-processing this measurement.
⇒ All measurements are compatible.
All logically possible effects are physically allowed.

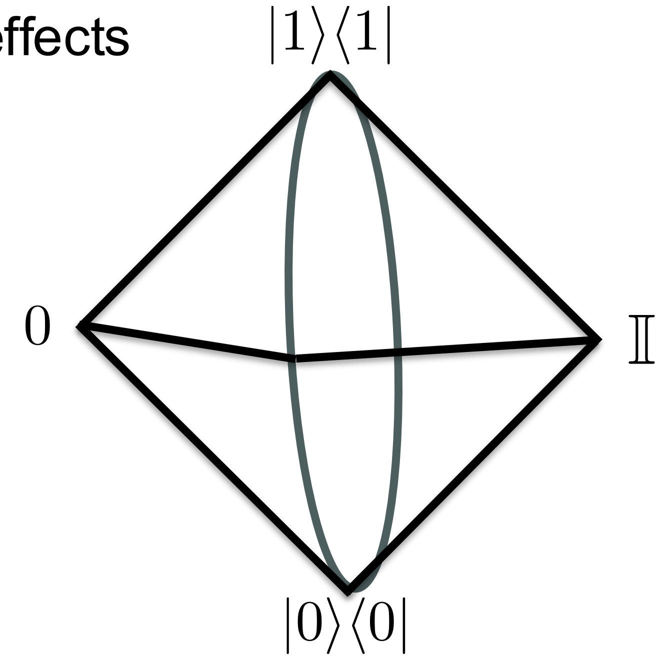
note: every simplicial system embeds inside quantum theory

states

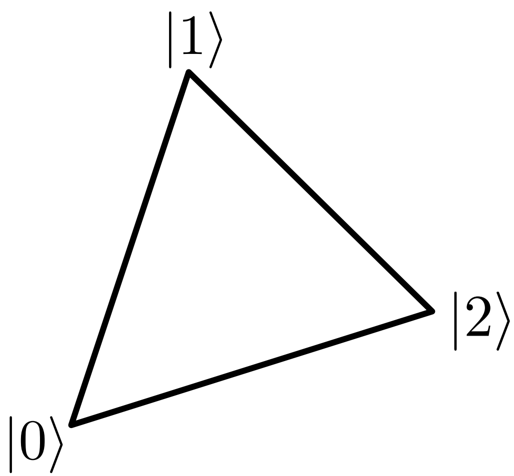
effects

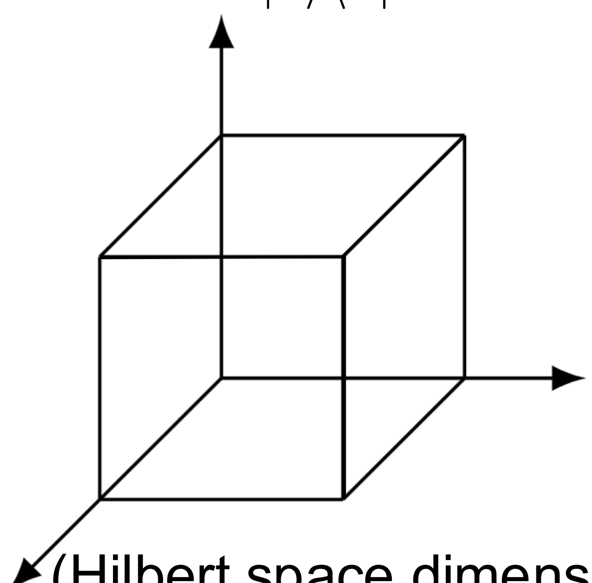
d = 2





d = 3





(Hilbert space dimension 3+ required)

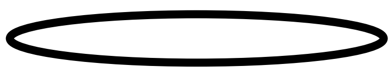
Real quantum theory

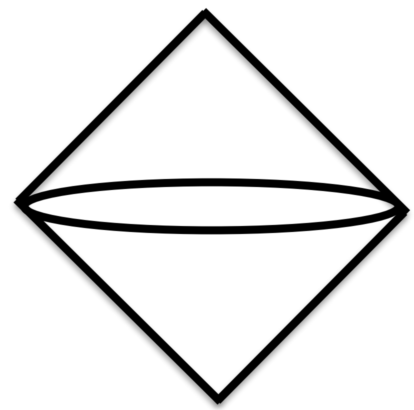
Like quantum theory, but defined over a real Hilbert space (rather than a complex Hilbert space)

$$\cancel{\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}}$$

states and effects are **real-valued**
positive operators

rebit:





$$\rho = \frac{1}{2}\left(\mathbb{1} + a_x\,\sigma_x + a_z\,\sigma_z\right)$$

$$a_x^2 + a_z^2 \leq 1$$

$$E = \frac{1}{2}\Big(b_0\,\mathbb{1} + b_x\sigma_x + b_z\sigma_z\Big)$$

This theory looks a lot like quantum theory at first glance. But it has some striking differences!

state space of a pair of qubits:
each has dimension 4-1. composite has dimension 16-1

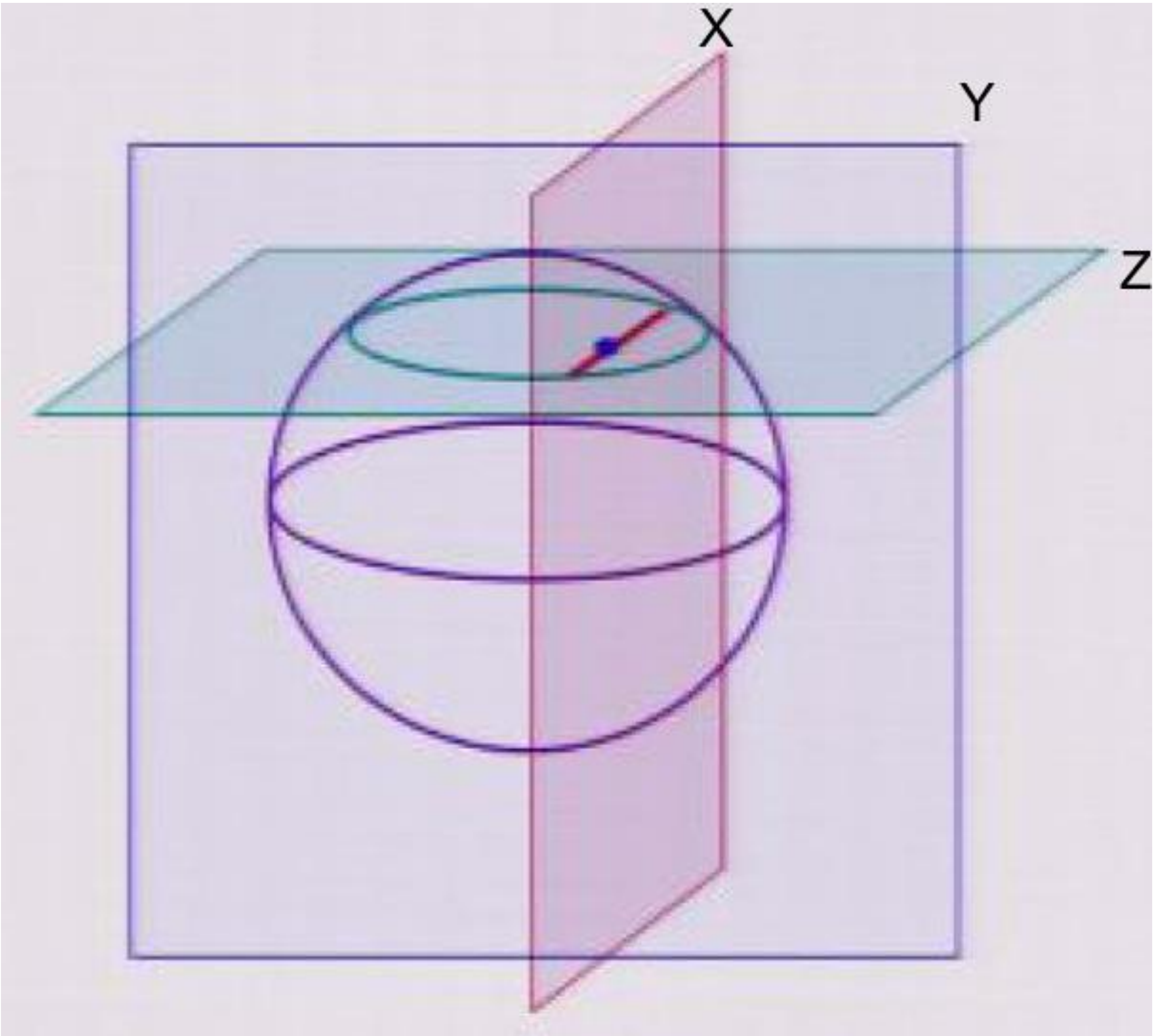
$$\sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

state space for a pair of rebits:
each has dimension 3-1. composite has dimension 9-1+1!

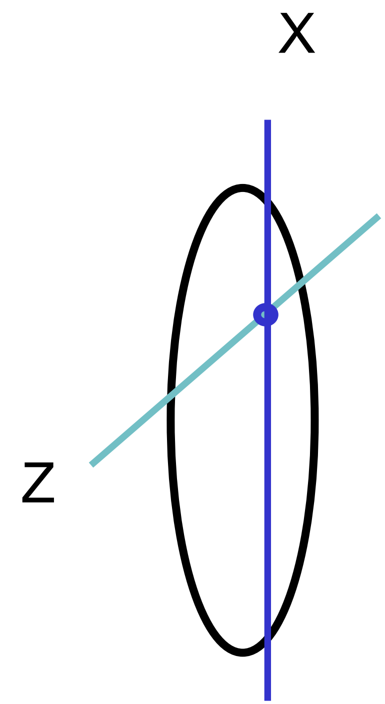
$$\rho_{AB} = \frac{1}{4}\Big(\mathbb{1}\otimes\mathbb{1} + \sum_{i\in\{x,z\}} a_i\,\sigma_i\otimes\mathbb{1} + \sum_{j\in\{x,z\}} b_j\,\mathbb{1}\otimes\sigma_j + \sum_{i,j\in\{x,z\}} T_{ij}\,\sigma_i\otimes\sigma_j\Big) + t\,(\sigma_y\otimes\sigma_y),$$

since $\sigma_y \otimes \sigma_y = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$ is real and symmetric

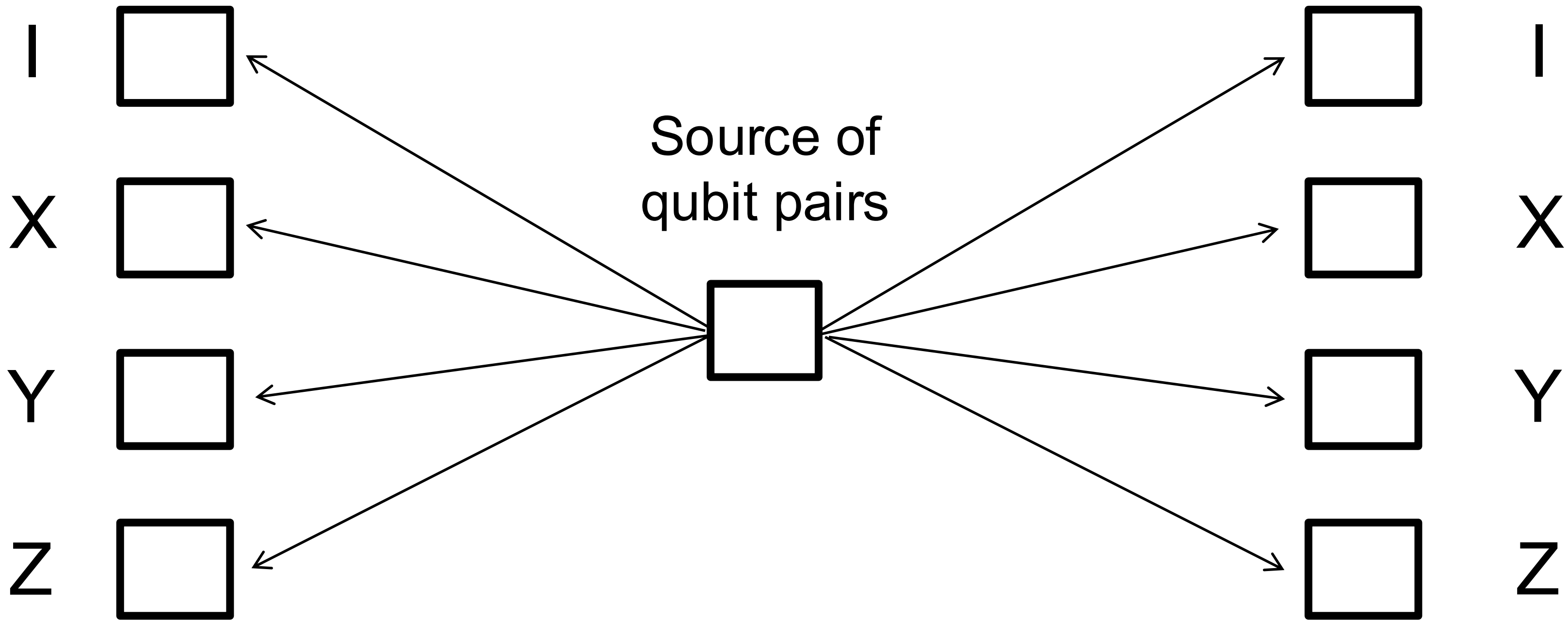
Tomography for a qubit



Tomography for a rebit



Tomography for two qubits $\quad 4^2 - 1 = 15$ parameters

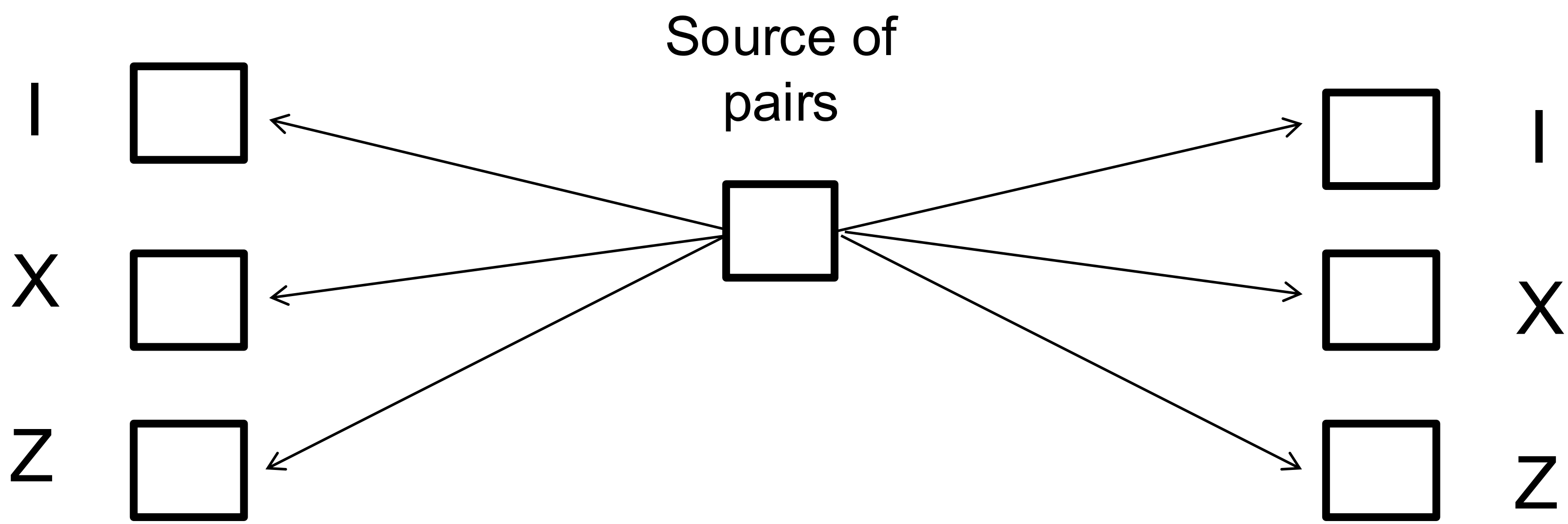


We obtain $4^2 - 1 = 15$ parameters

The state of two qubits can be determined from local measurements!

"Tomographic Locality"

Tomography for two rebits 9 parameters



We obtain $3^2 - 1 = 8$ parameters

$Y \otimes Y$ cannot be determined by local measurements!

The state of two real-amplitude qubits
cannot be determined from local measurements -- a kind of holism

Boxworld

Imagine a theory with two extremal, incompatible binary outcome mmts, labeled x=0,1, with outcomes a=0,1

any preparation of the system is completely summarized by the probabilities $p(a \mid x)$ for these two choices.

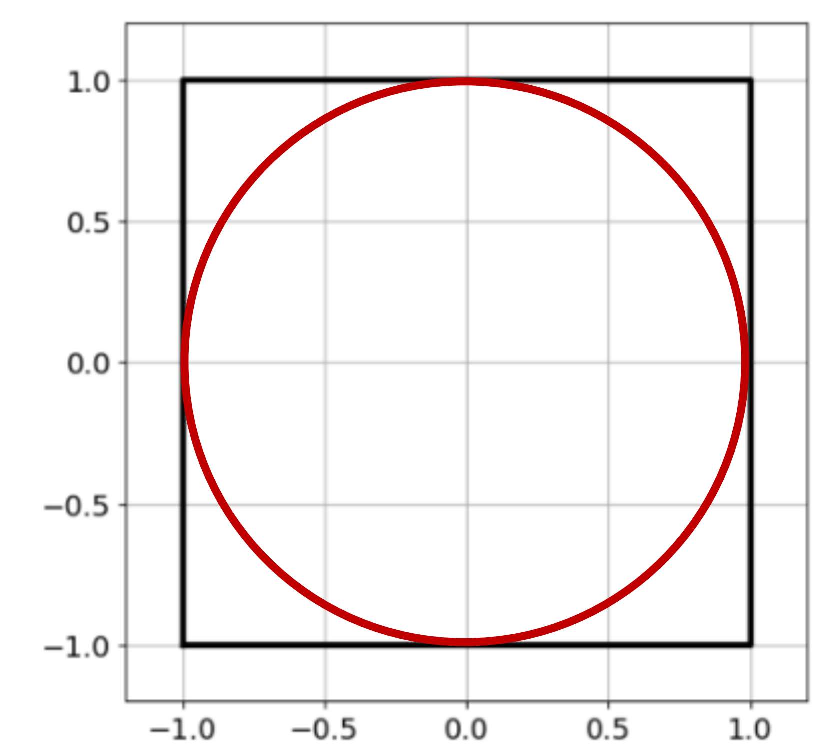
Since there are only two outcomes for each mmt, it is elegant to work with the bias $r_x = p(0|x) - p(1|x)$ for each mmt

$$r_x \in [-1, 1]$$

Imagine that in this theory, all logical possibilities are physically allowed

state space:

$$s = (1, r_1, r_2)$$



Joint biases allowed by QT.
Here we have “more states”

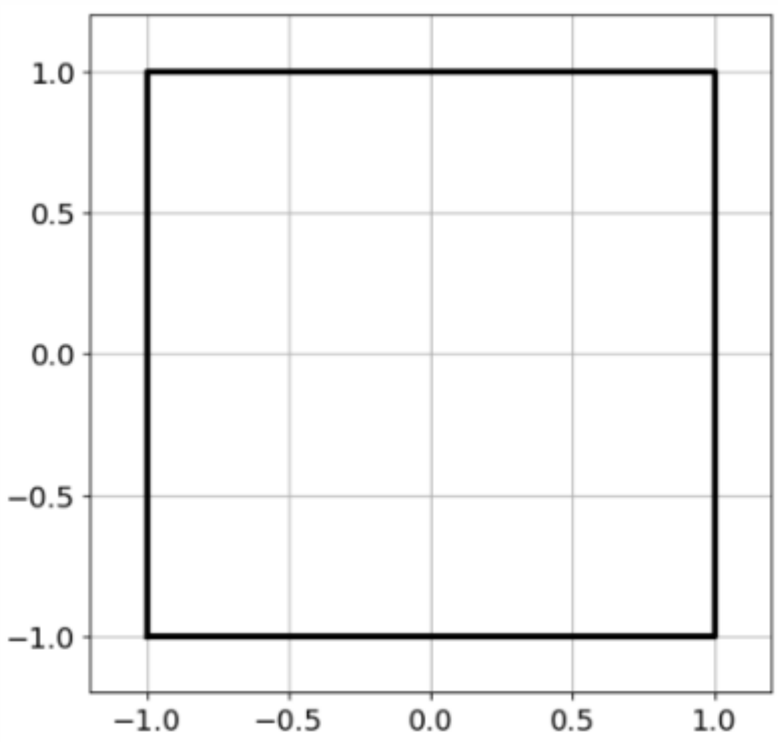
No uncertainty principle!
-extremal states make both extremal measurements perfectly predictable

Imagine that in this theory, all logical possibilities are physically allowed

effect = linear functional taking states to $p \in [0,1]$

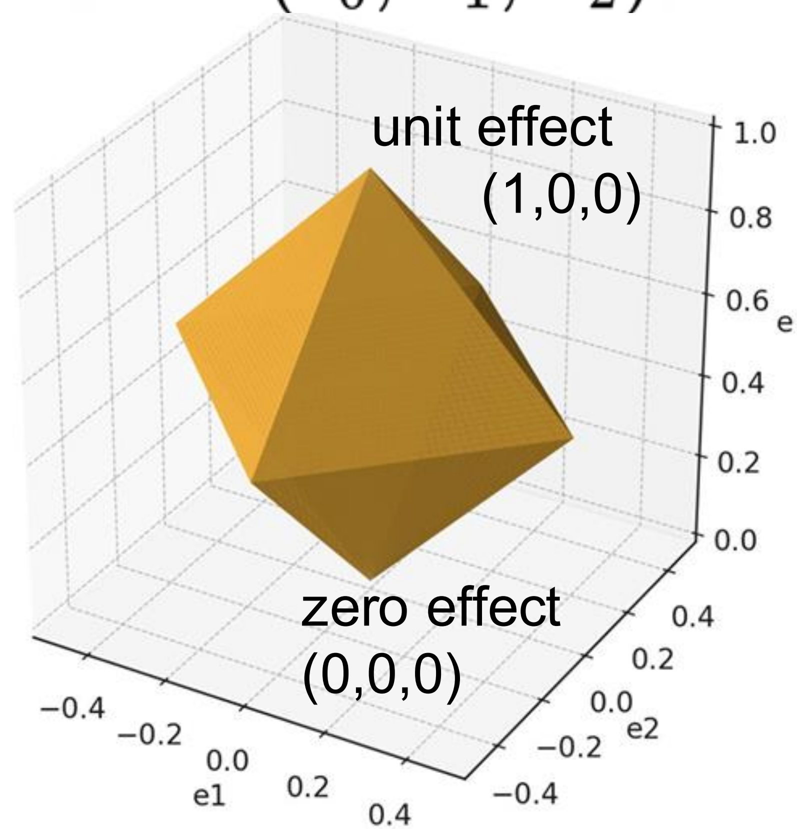
state space:

$$s = (1,\, r_1,\, r_2)$$



effect space:

$$e = (e_0,\, e_1,\, e_2)$$



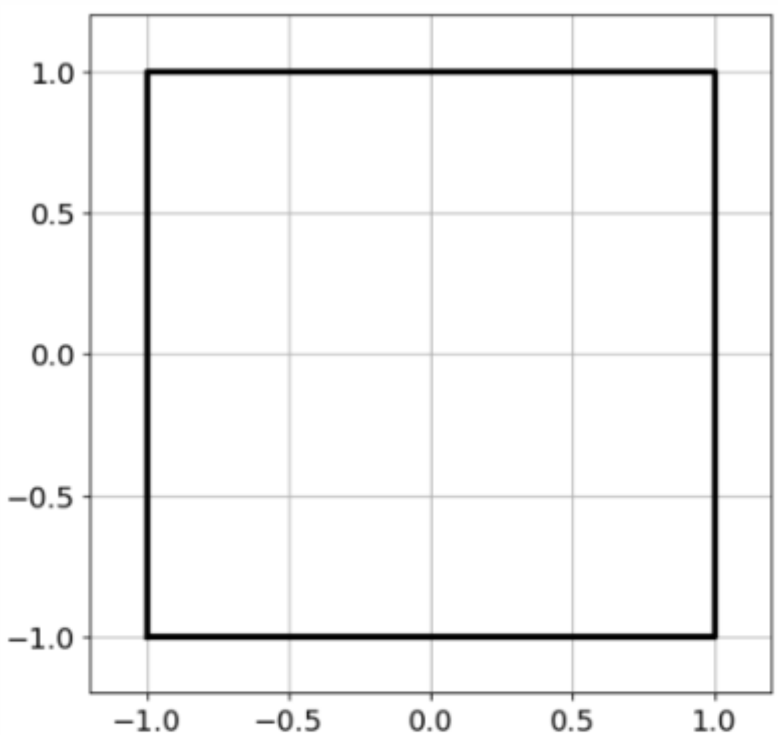
$0 \le e0 \le 1,\ |e1| + |e2| \le \min(e0,\, 1 - e0)$

Imagine that in this theory, all logical possibilities are physically allowed

effect = linear functional taking states to p ∈ [0,1]

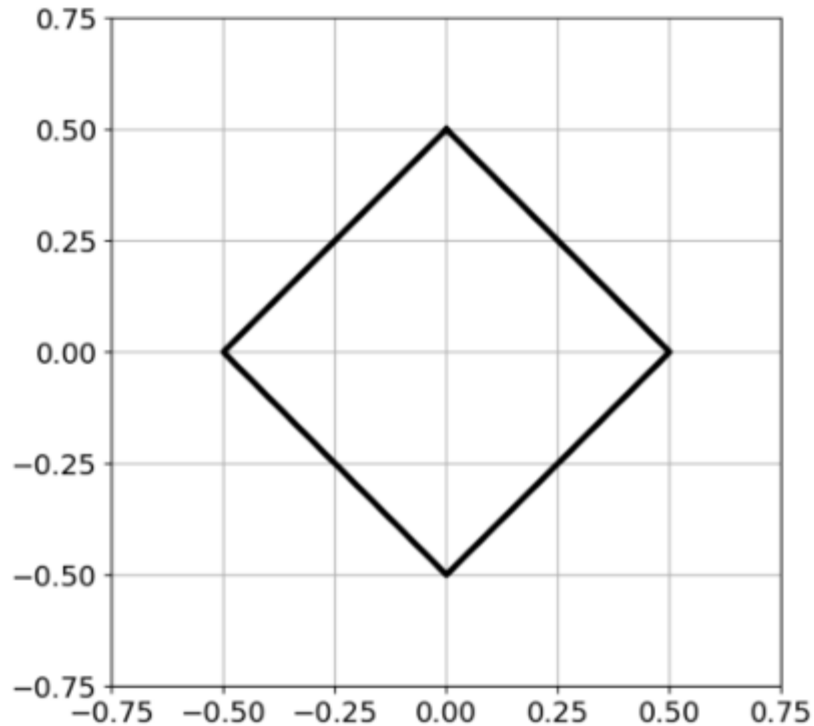
state space:

$$s = (1,\ r_1,\ r_2)$$



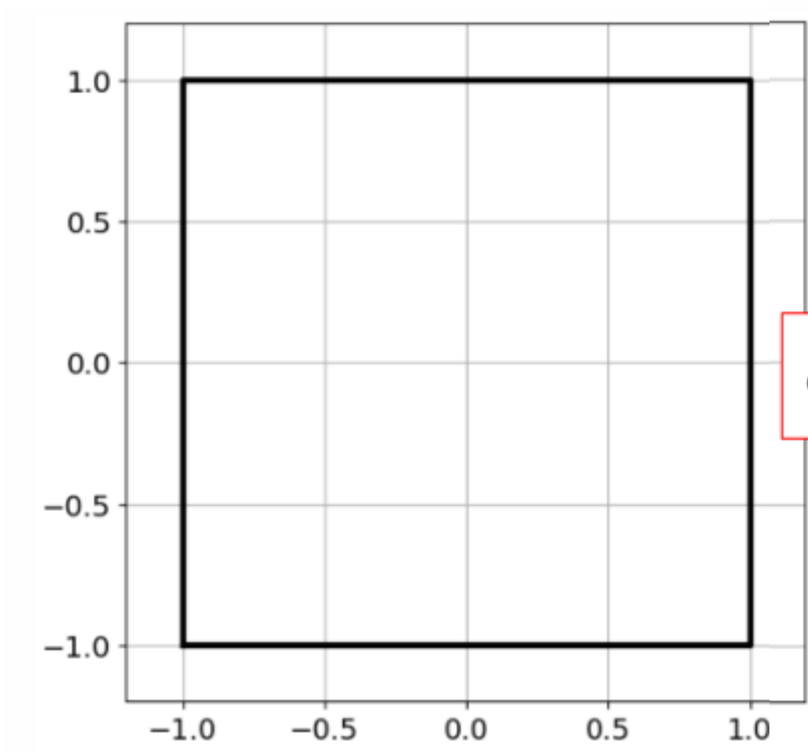
effect space:

$$e = (e_0,\ e_1,\ e_2) \qquad e_0 = \frac{1}{2}$$



state space:

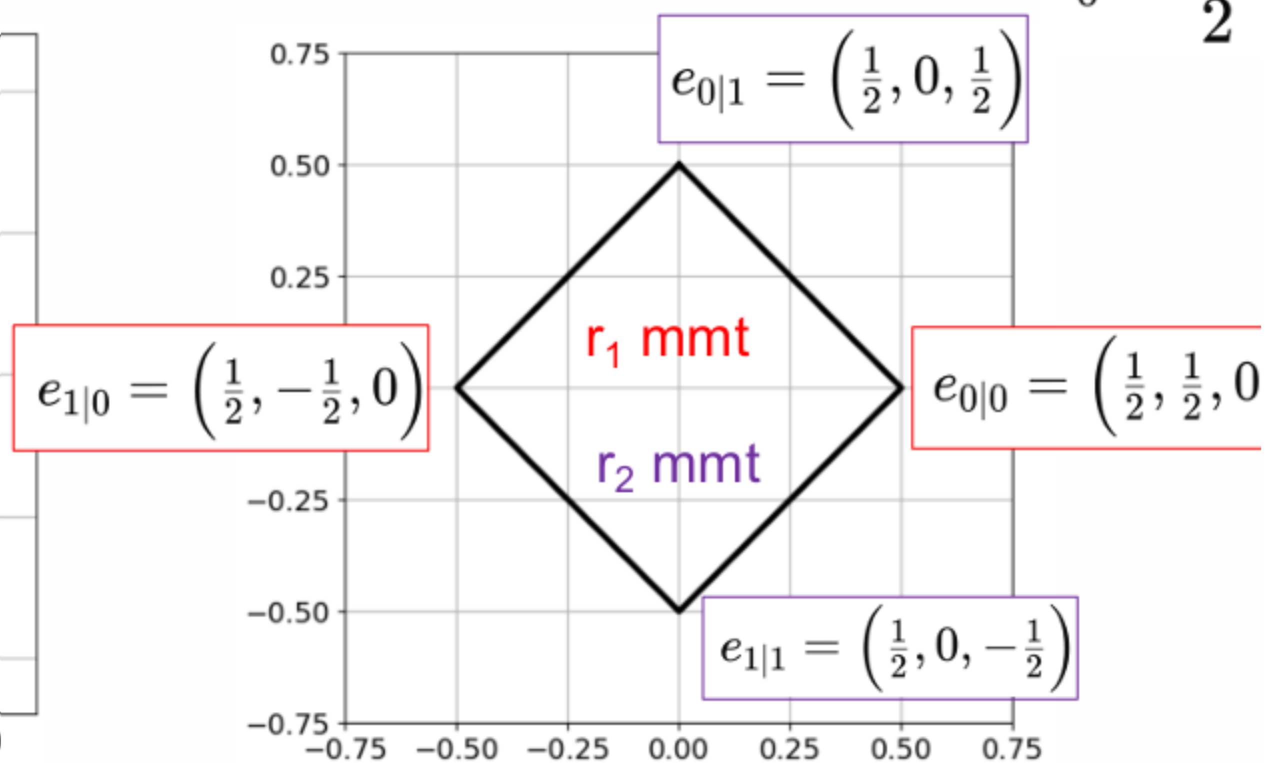
$$s = (1,\ r_1,\ r_2)$$



effect space:

$$e = (e_0,\ e_1,\ e_2)$$

$$e_0 = \frac{1}{2}$$

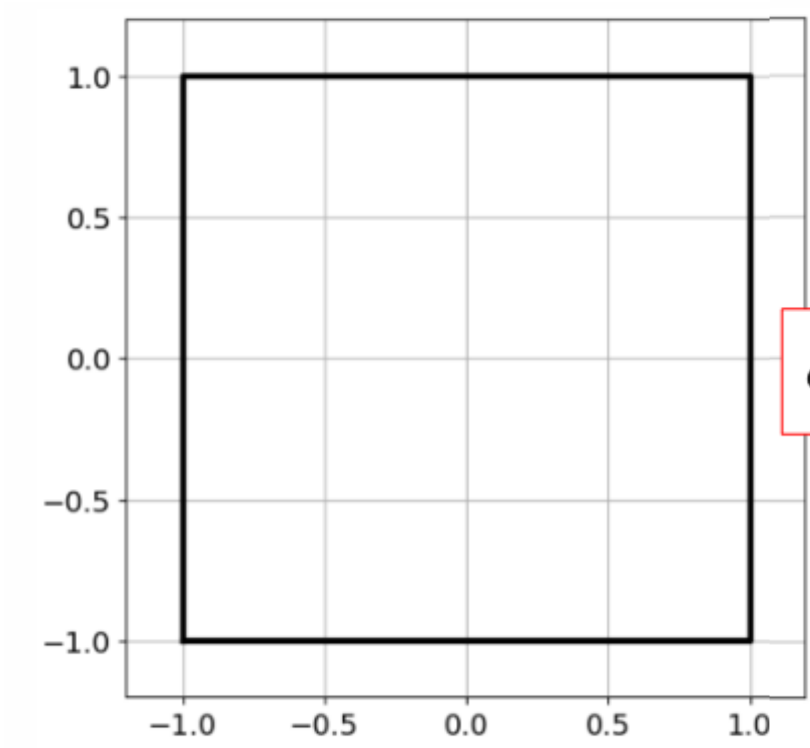


Has incompatible measurements.
(the two extremal mmts)

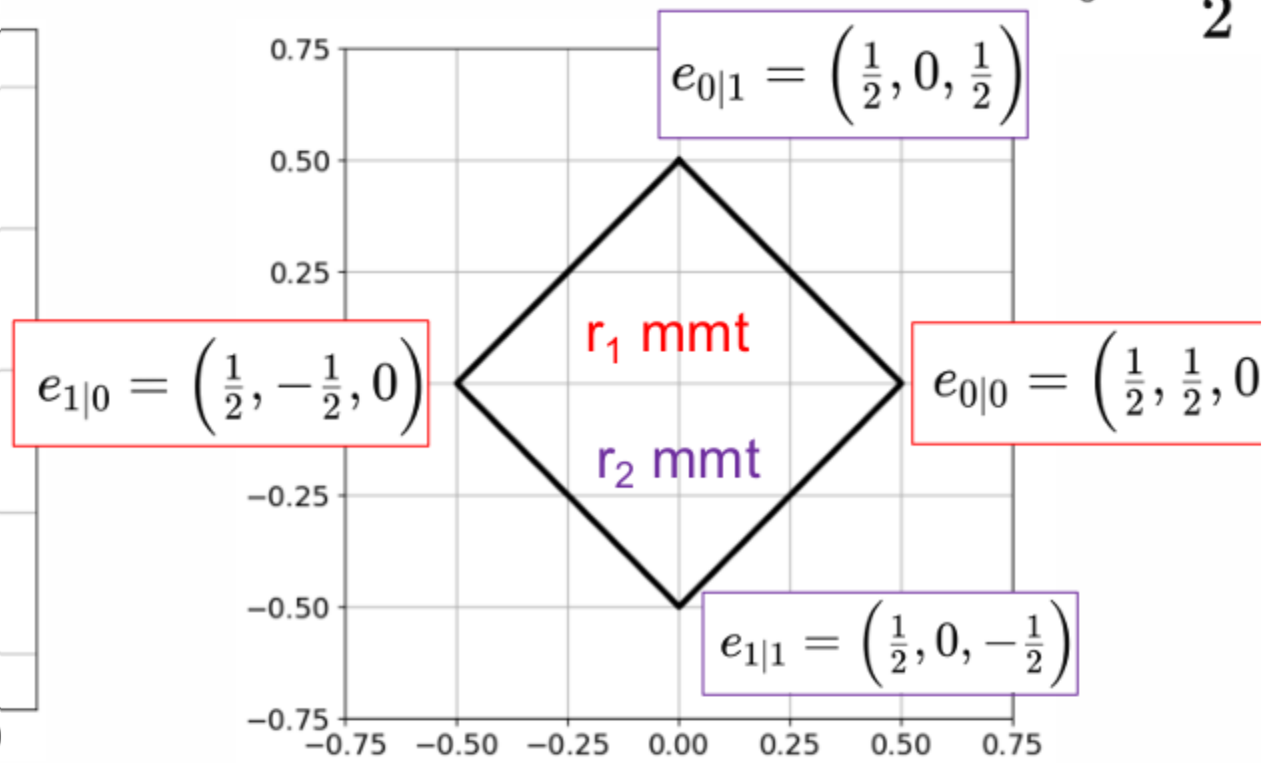
no measurement contains both

$$e_{0|1} = \left(\frac{1}{2}, 0, \frac{1}{2}\right) \qquad e_{0|0} = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$$

$$s = (1,\, r_1,\, r_2)$$



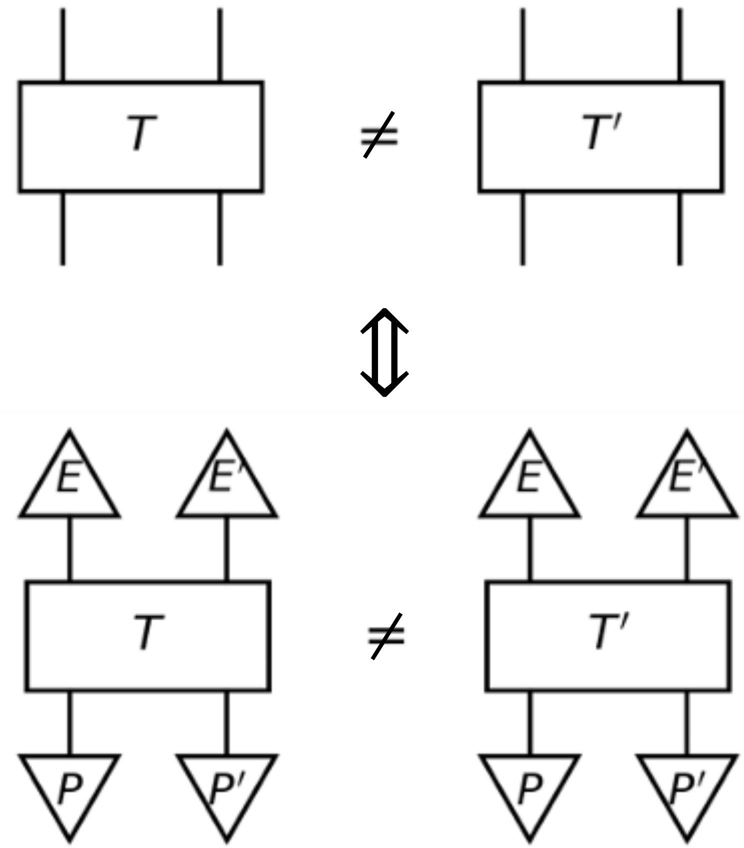
$$e = (e_0,\, e_1,\, e_2) \qquad e_0 = \frac{1}{2}$$



Operational Principles/Axioms

The principle of Tomographic Locality

Any two distinct processes in the theory can be differentiated by the statistics they give on some localized states/effects



$\exists E, E', P, P'$

Quantum Theory ✓

Real Quantum Theory ✗

Classical Theory ✓

Boxworld ✓

two ways this could fail:

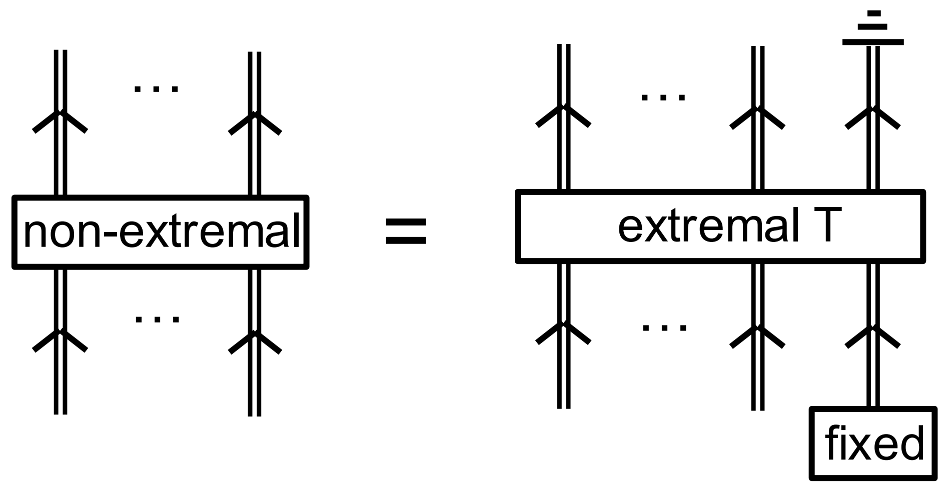
1) need correlated states/effects
2) need side channel

The Purification principle

every nonextremal process in the GPT can be obtained from ignorance about an extremal process in the GPT on some larger systems

(a unique process, up to reversible transformations)

∃ extremal T s.t.



Quantum Theory ✓

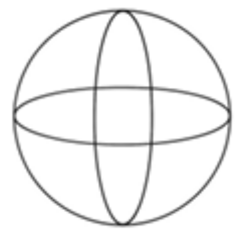
Real Quantum Theory ✓

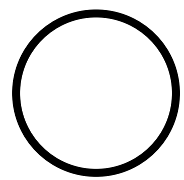
Classical Theory ✗

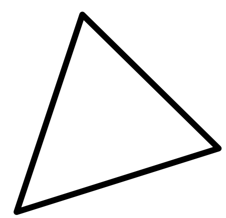
Boxworld ✗

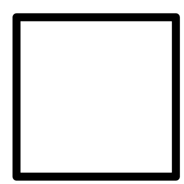
Continuous Transitivity

there is a continuous reversible transformation between any two pure states

Quantum Theory ✓ 

Real Quantum Theory ✓ 

Classical Theory ✗ 

Boxworld ✗ 

The no-restriction hypothesis

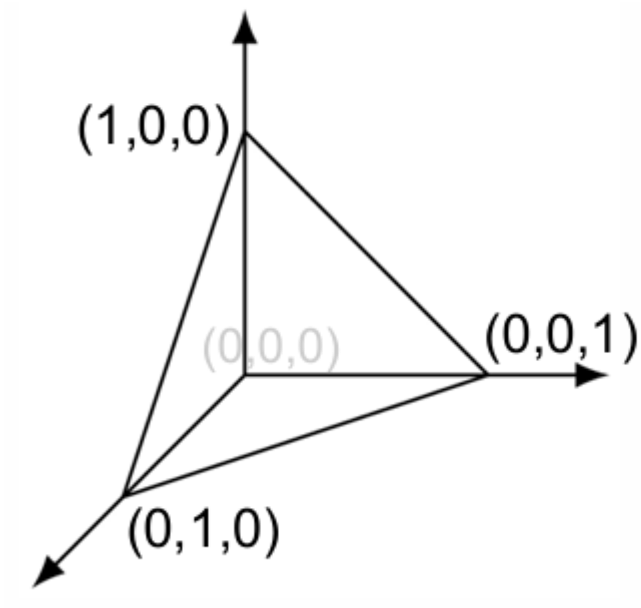
logically possible process ⇒ physical possible process

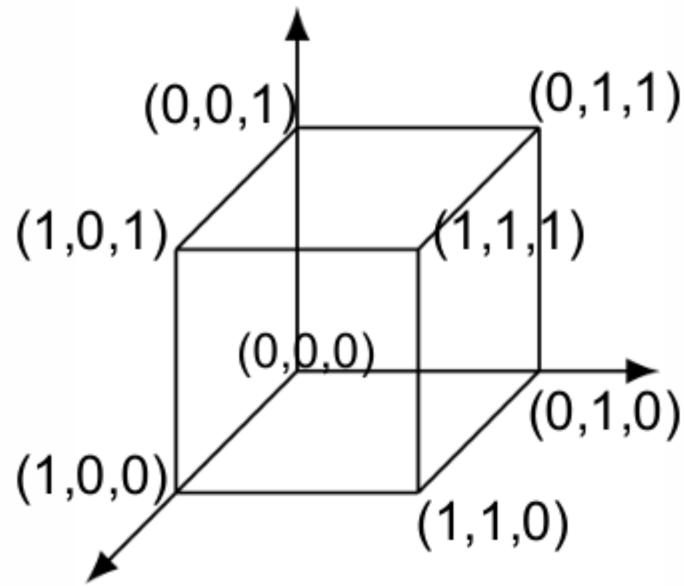
A GPT satisfies the No-Restriction Hypothesis <u>for effects</u> iff the set of physical effects is as large as is logically possible, given the state space.

A GPT satisfies the No-Restriction Hypothesis <u>for states</u> iff the set of physical states is as large as is logically possible, given the effect space.

(similarly for measurements, transformations, etc)

Classical theory satisfies the NRH





Quantum theory satisfies the NRH

in QT, the set of states = the set of density matrices

Any logically possible effect is a map from the state space to R.

$$f(\rho) : \mathrm{Herm}(\mathcal{H}) \longrightarrow \mathbb{R}$$

f must be linear (to respect classical uncertainty)

Riesz representation theorem: any linear functional on a vector space can be written as an inner product with some fixed vector in the space.

$$f(\rho) : \mathrm{Herm}(\mathcal{H}) \longrightarrow \mathbb{R}$$

$$\Rightarrow \exists\, E \in \mathrm{Herm}(\mathcal{H}) \quad \text{s.t. } f(\rho) = \langle \rho, E \rangle = \mathrm{Tr}[\rho E]$$

Moreover, these E operators must be POVM elements, since:

$$\mathrm{Tr}(\rho E_k) \geq 0 \quad \forall \rho \in \mathcal{S}(\mathcal{H})$$

$$\Rightarrow E_k \text{ is a positive operator}$$

$$\sum_k \mathrm{Tr}(\rho E_k) = 1 \quad \forall \rho \in \mathcal{S}(\mathcal{H})$$

$$\Rightarrow \sum_k E_k = I$$

If you assume states are density operators, then
the set of logically possible effects is *exactly* the set of POVMs.

⇒ NRH for effects

The no-restriction hypothesis

Quantum Theory ✓

Real Quantum Theory ✓

Classical Theory ✓

Boxworld ✓

Spekkens Toy Theory ✗

Finding better postulates for quantum theory

Recall these postulates:

1. Preparations are density operators.

2. Every measurement is a POVM, and the probability of getting outcome k is $\mathrm{Tr}(\rho E_k)$

But given Postulate 1, we showed that an assumption that <u>logically possible measurements are all physically possible</u> *implies* Postulate 2!

Recall these postulates:

1. Preparations are density operators.

~~2. Every measurement is a POVM, and the probability of getting outcome k is~~ $\cancel{\mathrm{Tr}(\rho E_k)}$

2. No-restriction hypothesis

But given Postulate 1, we showed that an assumption that <u>logically possible measurements are all physically possible</u> *implies* Postulate 2!

The following set of postulates that implies quantum theory:

1. All logically possible measurements are physically possible
2. States can be identified through local measurements
3. Evolution is continuous, reversible, and connects every pair of extremal states
4. Subsystems of higher-level systems are simply lower-level systems

Compare this to:

Every preparation P is associated with a trace-one Hermitian operator.

Every measurement M is associated with a positive operator-valued measure $\{E_k\}$. The probability of M yielding outcome k given a preparation P is $Pr(k|\mathsf{P},\mathsf{M}) = \mathrm{Tr}(\rho E_k)$

Every transformation is associated with a trace-preserving completely-positive linear map $\rho \to \rho' = \mathcal{T}(\rho)$

Every measurement outcome k is associated with a completely-positive trace-nonincreasing linear map T_k such that

$$\rho \to \rho_k = \frac{\mathcal{T}_k(\rho)}{\mathrm{Tr}[\mathcal{T}_k(\rho)]} \quad \text{where} \quad \mathcal{T}_k^\dagger(I) = E_k$$

The following set of postulates that implies quantum theory:

1. All logically possible measurements are physically possible
2. States can be identified through local measurements
3. Evolution is continuous, reversible, and connects every pair of extremal states
4. Subsystems of higher-level systems are simply lower-level systems

Many other axiomatizations of quantum theory are known.

Typically, these have a similar operational and informational flavor.

These axioms still don’t constitute a realist theory, and they still refer to approximate/emergent concepts like measurements. But surely this is progress relative to the usual axioms!

Compositional structure

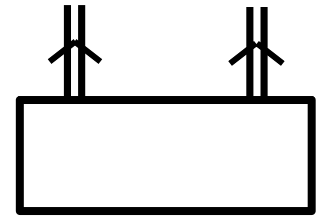
A full definition of a theory should specify all possible transformations on all system types.

Quantum and classical theory satisfy the no-restriction hypothesis for transformations.

But in general, a theory must specify exactly which transformations are possible.

Theories with larger state spaces must have smaller effect spaces, and often smaller transformation spaces.
-E.g., boxworld has more states than a qubit, as we saw, but fewer effects, and nontrivial reversible interactions between pairs of such systems are logically impossible.

Bipartite state and effect spaces



The state space of a composite system is not fully specified by the state space of the components.

Min tensor product (for state spaces):
Allow only product states and mixtures of them.

(separable only)

Max tensor product (for state spaces):
Allow all states that give valid probabilities on product effects.

(typically includes entangled states)

Min and max tensor product for effect spaces are defined analogously

-Max tensor product for one of these implies min tensor product for the other.
Intuitively: allowing more states/effects restricts the possibilities for the other.
-Classical theory is min = max tensor product for both.
-Quantum theory is between the two extremes.

If bipartite states can have holism, then there can be *even larger* state spaces.

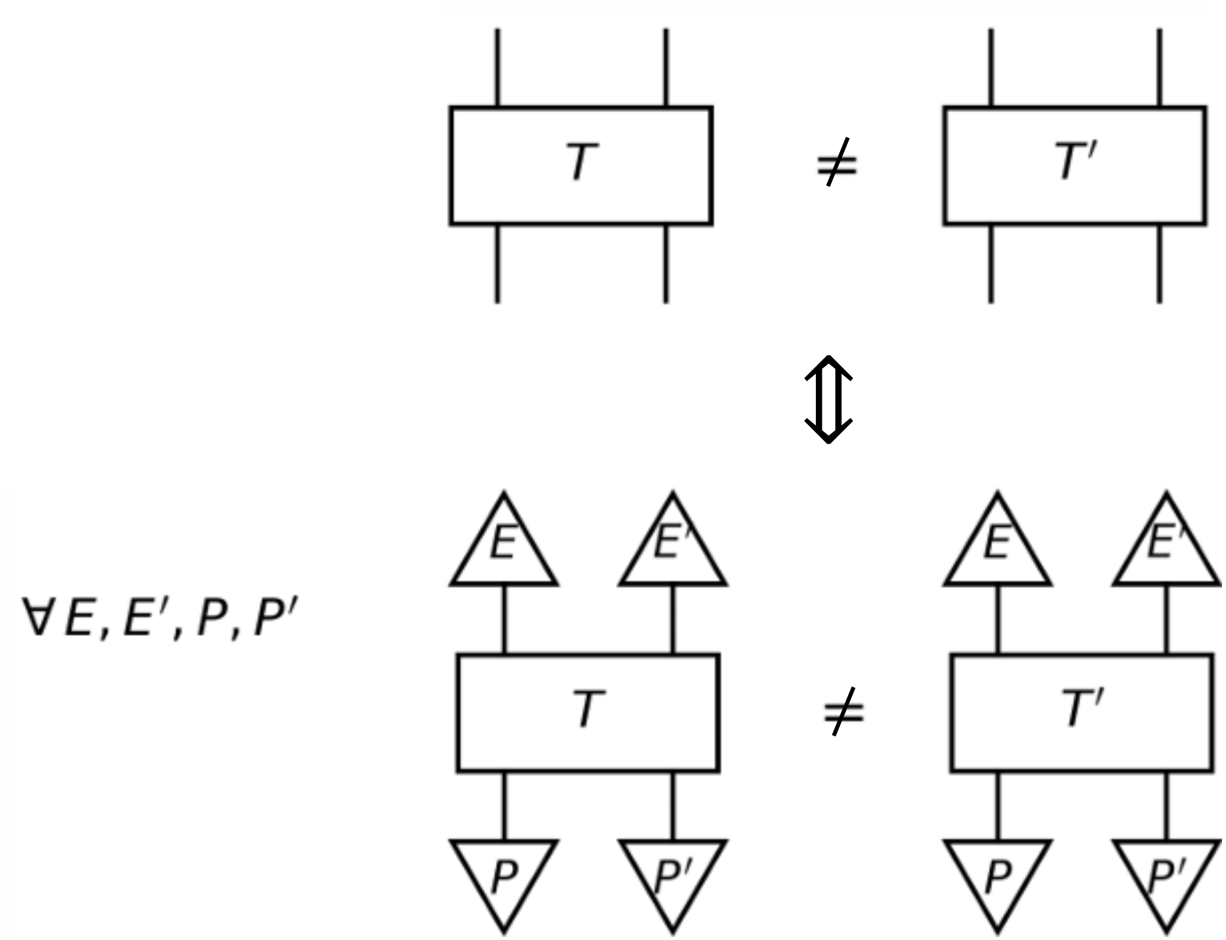
e.g., recall rebit states:

$$\rho_{AB} = \frac{1}{4}\Big(\mathbb{1}\otimes\mathbb{1} + \sum_{i\in\{x,z\}} a_i\,\sigma_i\otimes\mathbb{1} + \sum_{j\in\{x,z\}} b_j\,\mathbb{1}\otimes\sigma_j + \sum_{i,j\in\{x,z\}} T_{ij}\,\sigma_i\otimes\sigma_j\Big)$$
$$+\,t\,(\sigma_y\otimes\sigma_y),$$

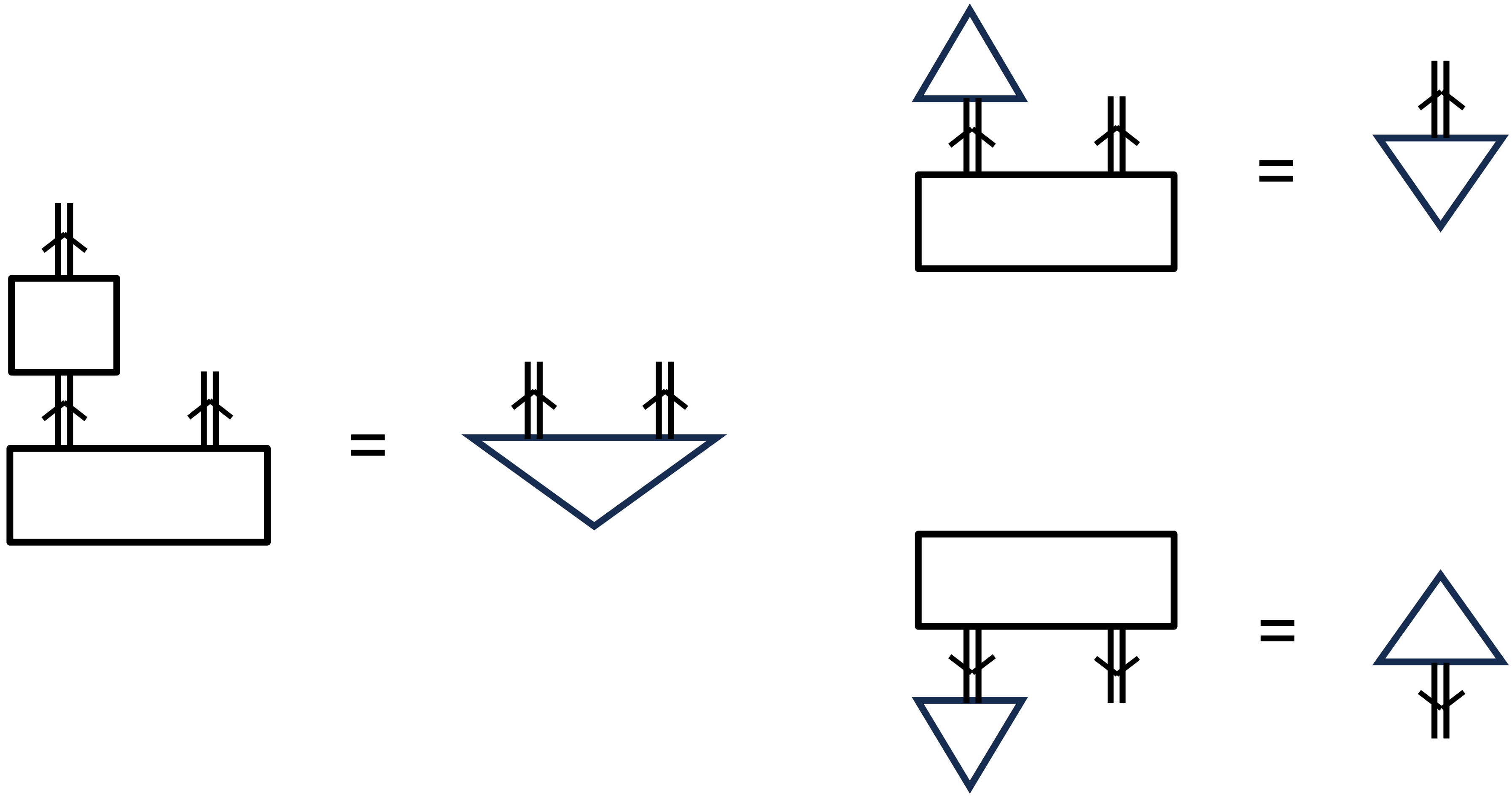
Products of real-valued effects live in a 9-dimensional space, with no Y⊗Y term. So states could have arbitrarily large component in that axis, without contributing to the probabilities!

The concepts of min and max tensor product implicitly assume that all bipartite states and effects are in the linear span of the local states and effects.

Tomographic locality:
Any two distinct processes in the theory can be differentiated
by the statistics they give on some localized states/effects



Every GPT is closed under arbitrary (type-matched) compositions



When constructing a theory, this is a key condition to check.

Defining and witnessing nonclassicality

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

rigorously, what does that mean?

Wave-particle duality?

Entanglement?

Remote steering?

No-cloning?

Nonlocality?

Teleportation?

Quantum interference?

Coherent superposition?

etc…

Contextuality

Uncertainty relations?

But most of these are consistent with simple realist explanations!
(e.g. Spekkens toy theory)

~~Wave-particle duality?~~

~~Entanglement?~~

~~Remote steering?~~

~~No-cloning?~~

Nonlocality?

~~Teleportation?~~

~~Quantum interference?~~

~~Coherent superposition?~~

etc…

Contextuality?

~~Uncertainty relations?~~

But most of these are consistent with simple realist explanations!
(e.g. Spekkens toy theory)

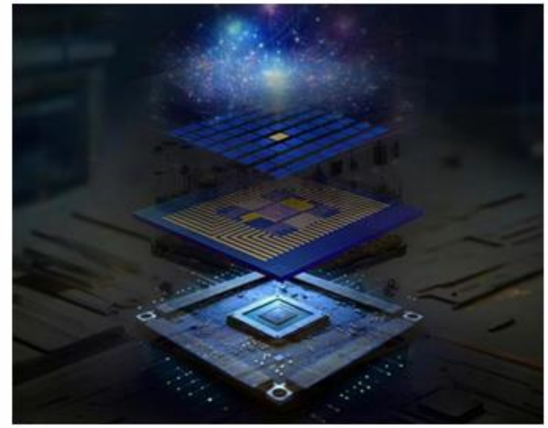
We need a principled way of dividing phenomena into
those which can be “explained classically”, and
those which are rigorous proofs of nonclassicality.

Why bother?

Intrinsic interest

settling century-old debates about if and when quantum phenomena force us to reject various physical principles

Resources and benchmarks for QIP



Influencing how we interpret and extend quantum theory

quantum gravity?
quantum causal modeling?
quantum machine learning?
quantum thermodynamics?

The things that we want to explain are phenomena/experiments/theories

(not, e.g., quantum states)

describing a theory:	describing an experiment
GPT	GPT fragment
all possible systems, processes, and circuits	subset of systems, processes, and circuits

When is a GPT classically explainable?

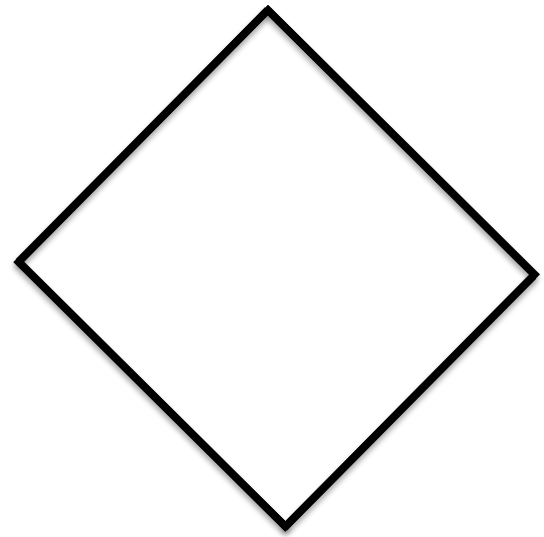
Classical explainability
(prepare-measure)

classical systems are consistent with (explainable by) quantum theory:

states

effects

d = 2

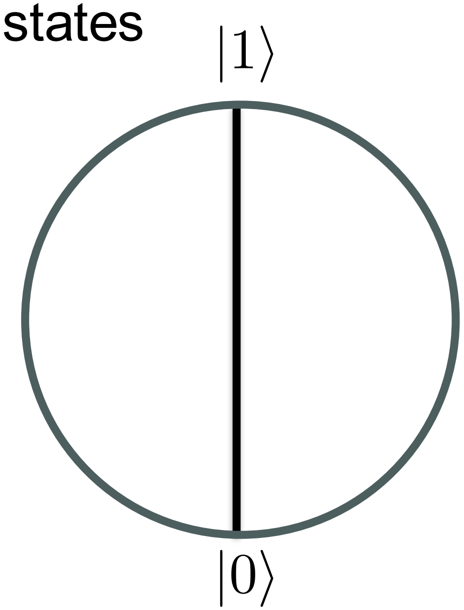


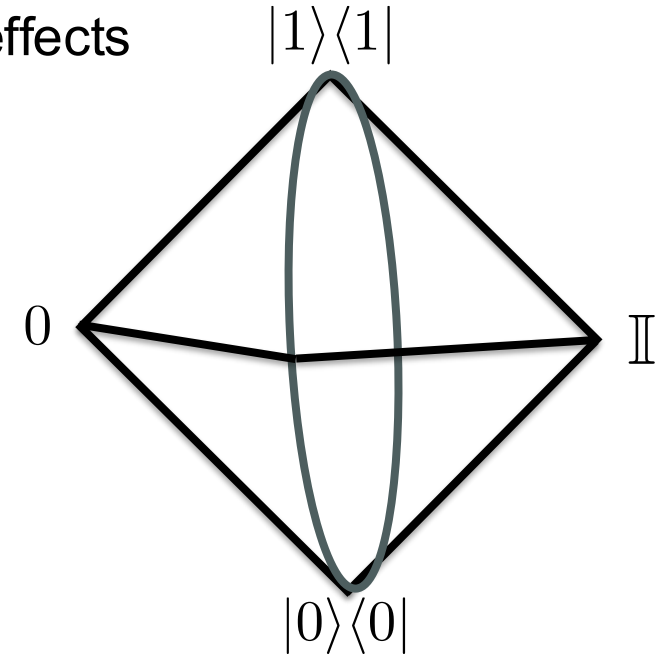
classical systems are consistent with (explainable by) quantum theory:

states

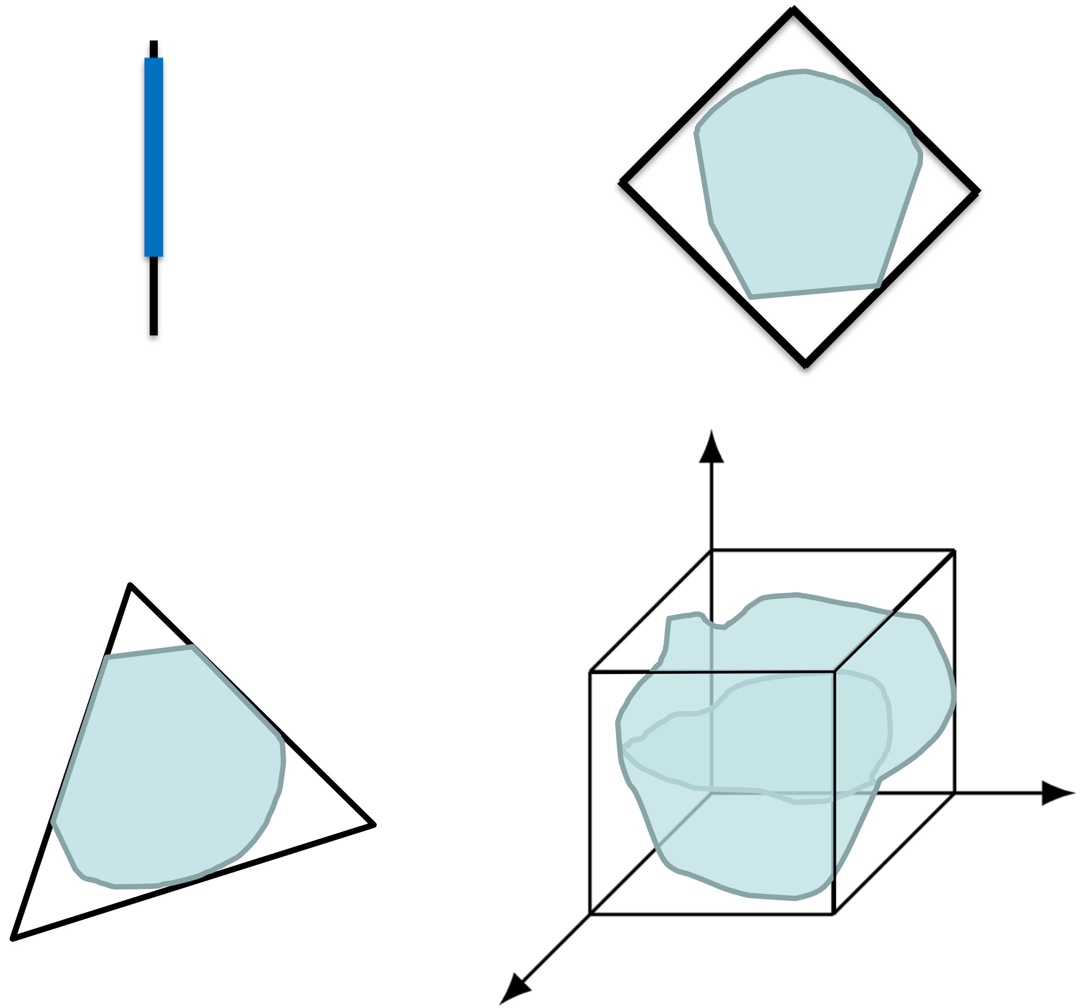
effects

d = 2



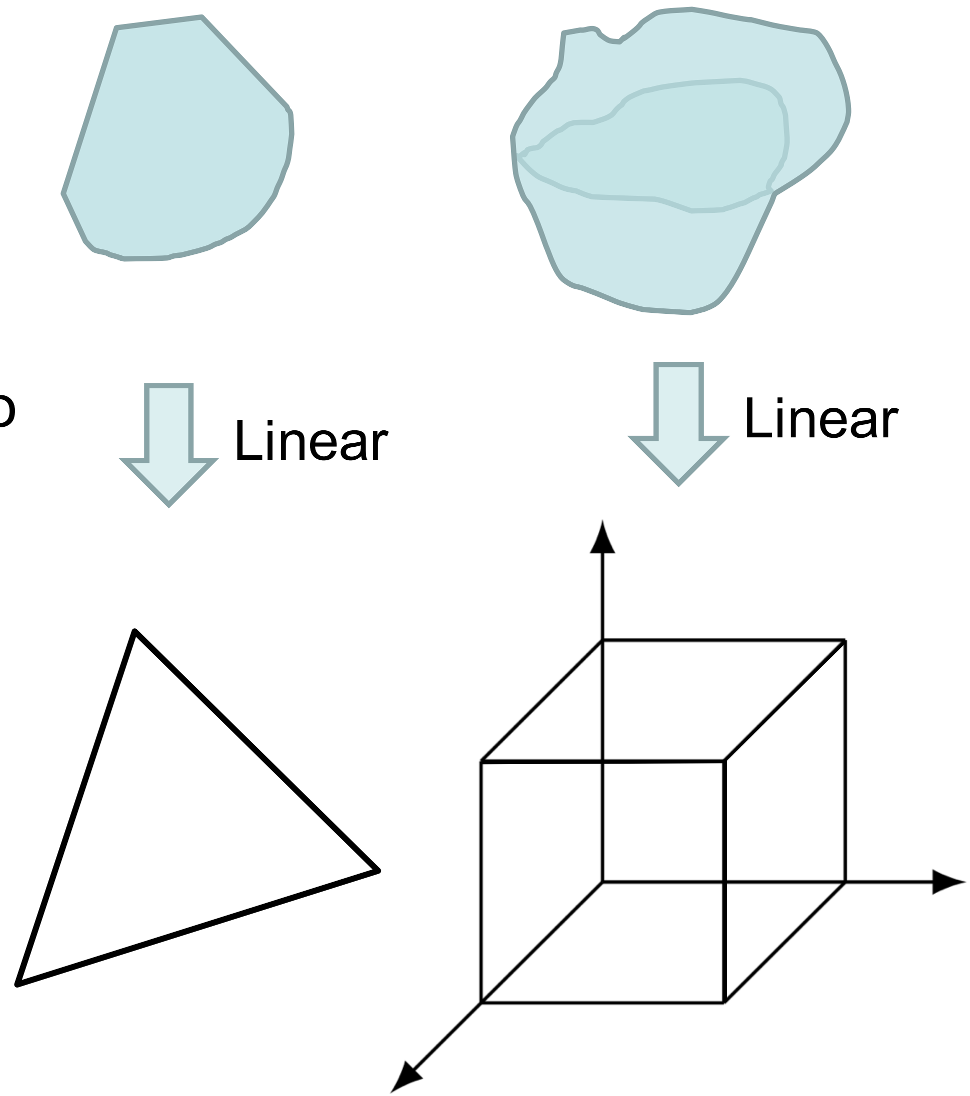


Similarly, any theory/fragment that embeds into the classical GPT is classically explainable



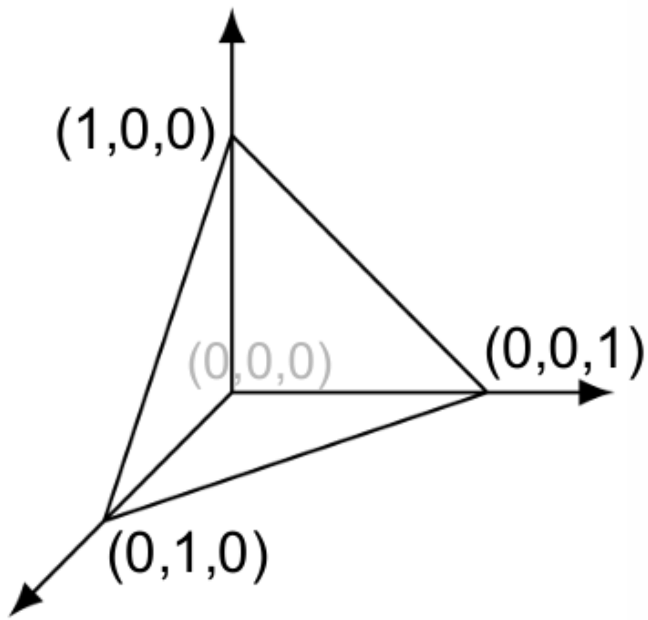
A prepare-measure GPT (fragment) is classically explainable iff there exists

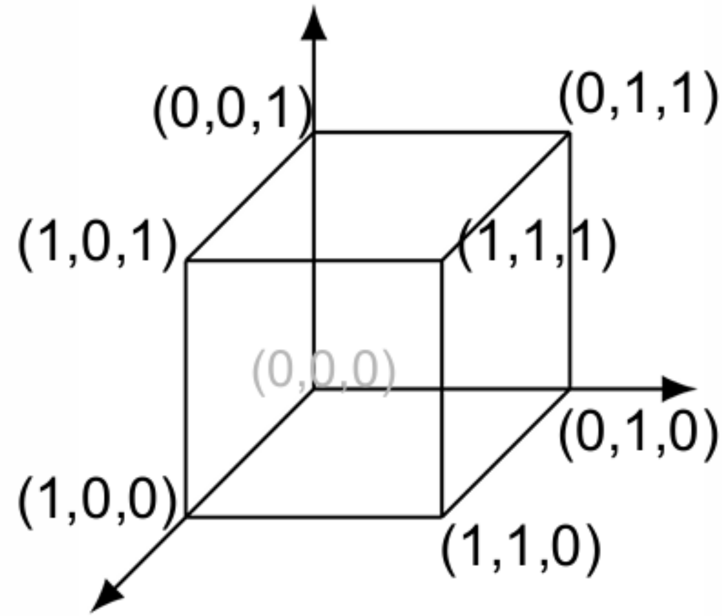
1) a linear map taking its states into a simplex, and
2) a linear map taking its effects into the dual to that simplex, such that
3) probabilities are preserved



"simplex embedding"

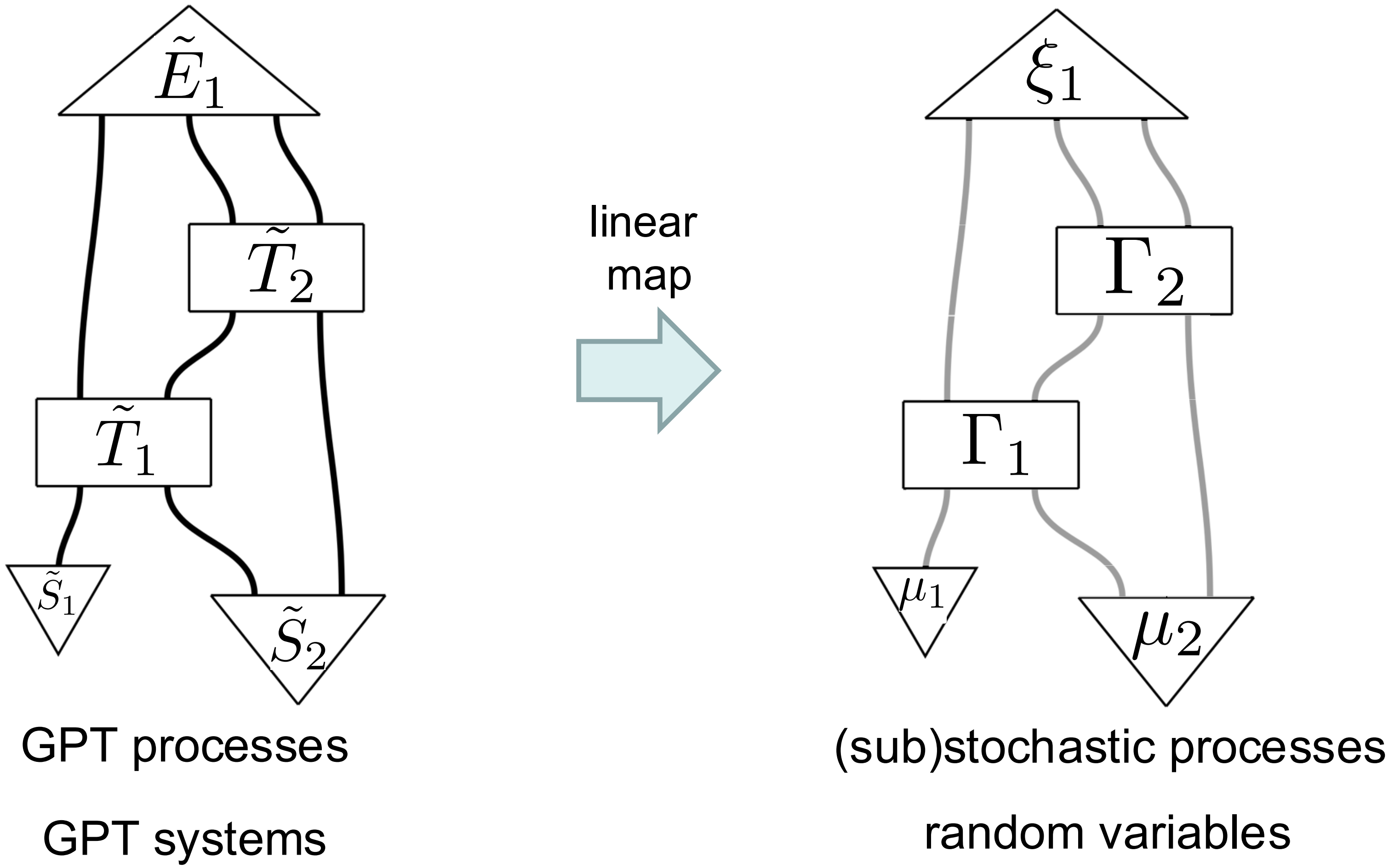
d=3





Any such mapping gives an <u>ontological model</u> for the GPT!

Beyond prepare-measure



This defines <u>classical-explainability</u> for a GPT/fragment.

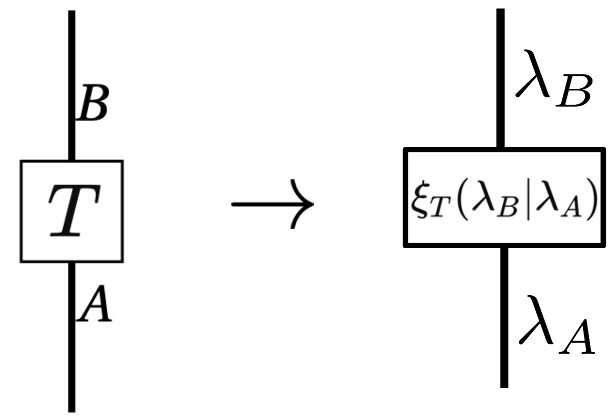
If *no* such mapping exists, the GPT/frag is *not* classically explainable.

Ontological model of a GPT:

ξ : GPT $\rightarrow$ Classical GPT

which:

1) preserves the predictions
2) is linear
3) is diagram-preserving



GPT transformation

(sub)stochastic map

Linearity

preservation of
convex geometry

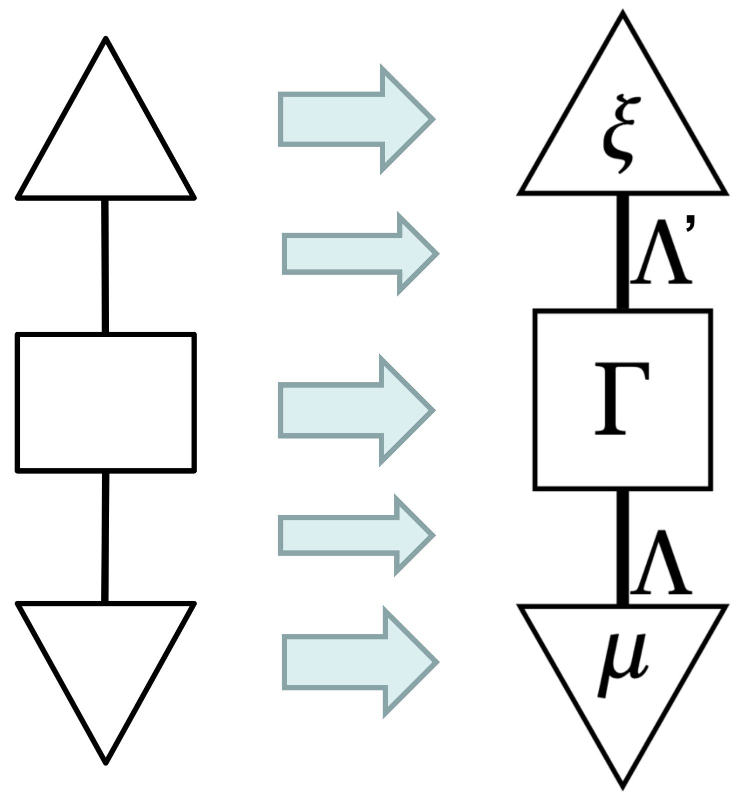
$$\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2}|+\rangle\langle +| + \frac{1}{2}|-\rangle\langle -|$$

$$\Downarrow$$

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

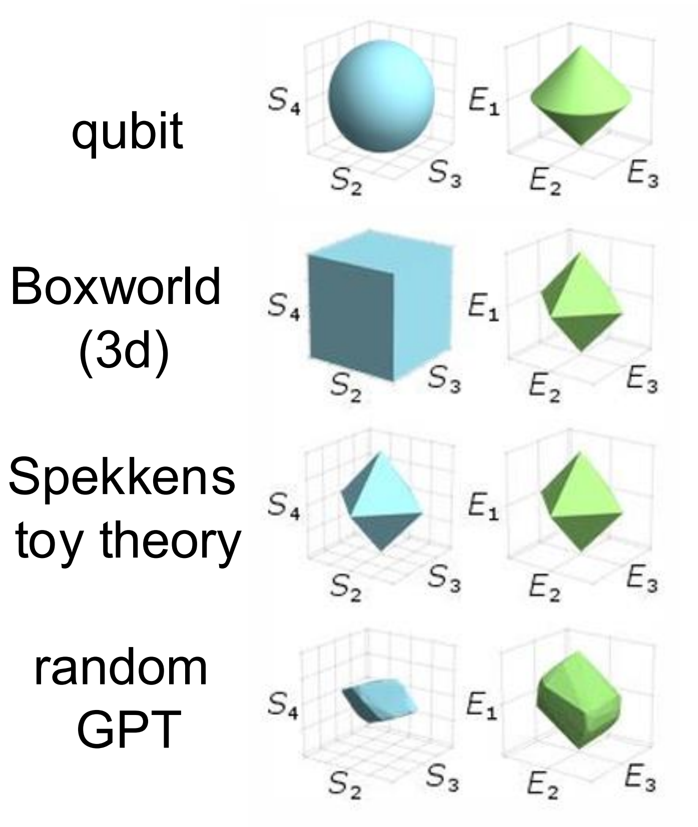
Diagram-preservation

preservation of
compositional structure



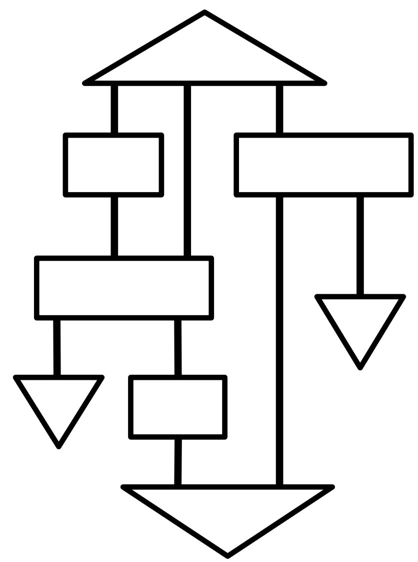
Recall that these are the two basic structures of a GPT!

1. Convex geometry



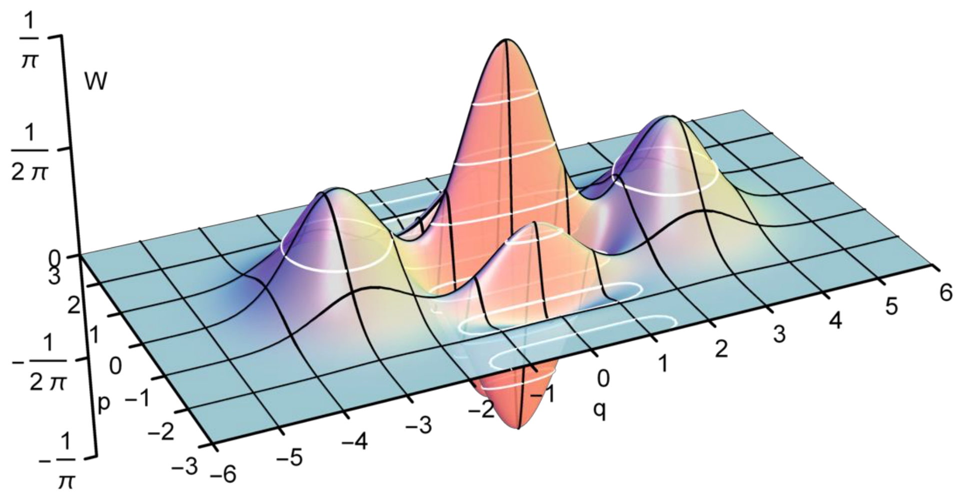
2. Compositional structure

-multipartite states
-multipartite effects
-allowed transformations
-probabilities = inner product
-etc



compare to the Wigner function: linear map from density matrices to real-valued vectors

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



We simply generalize this to <u>arbitrary</u> linear functions over <u>arbitrary</u> classical variables

and to channels/mmts/circuits instead of just states

Then we ask whether <u>**any**</u> such mapping exists that is everywhere positive.

This is exactly equivalent to noncontextuality!

Consider an unquotiented operational theory and the GPT associated to it (by quotienting with respect to operational equivalences).

The operational theory is classically-explainable if and only if

1. it admits of a noncontextual ontological model
2. its GPT admits of <u>any</u> ontological model

2’. its GPT admits of any simplex-embedding

2’’. its GPT admits of any positive quasiprobability representation

These are equivalent notions of classicality!

Coming this month!

If you want me to email you a copy, message me at dschmid1@perimeterinstitute.ca