Title: Computation in generalised probabilistic theories

Date: Feb 25, 2015 04:00 PM

URL: http://pirsa.org/15020107

Abstract: From the general difficulty of simulating quantum systems using classical systems, and in particular the existence of an efficient quantum algorithm for factoring, it is likely that quantum computation is intrinsically more powerful than classical computation. At present, the best upper bound known for the power of quantum computation is that BQP is in AWPP, where AWPP is a classical complexity class (known to be included in PP, hence PSPACE). This work investigates limits on computational power that are imposed by simple physical, or information theoretic, principles. To this end, we define a circuit-based model of computation in a class of operationally-defined theories more general than quantum theory, and ask: what is the minimal set of physical assumptions under which the above inclusions still hold? We show that given only an assumption of tomographic locality (roughly, that multipartite states and transformations can be characterised by local measurements), efficient computations are contained in AWPP. This inclusion still holds even without assuming a basic notion of causality (where the notion is, roughly, that probabilities for outcomes cannot depend on future measurement choices). Then, following Aaronson, we extend the computational model by allowing post-selection on measurement outcomes. Aaronson showed that the corresponding quantum complexity class, PostBQP, is equal to PP. Given only the assumption of tomographic locality, the inclusion in PP still holds for post-selected computation in general theories. Hence in a world with post-selection, quantum theory is optimal for computation in the space of all operational theories. We then consider whether one can obtain relativised complexity results for general theories. It is not obvious how to define a sensible notion of a computational oracle in the general framework that reduces to the standard notion in the quantum case. Nevertheless, it is possible to define computation relative to a `classical oracle'. Then, we show there exists a classical oracle relative to which efficient computation in any theory satisfying the causality assumption does not include NP. This provides some degree of evidence that NP-complete problems cannot be solved efficiently in any theory satisfying tomographic locality and causality. Based on arXiv:1412.8671. Joint work with Jon Barrett.

Computation in generalised probabilistic theories

Ciarán Lee

Joint work with Jon Barrett

arXiv:1412.8671

Outline

- Motivation
- Introduce problem
- Framework for generalised probabilistic theories
- Computational model and results.

Motivation

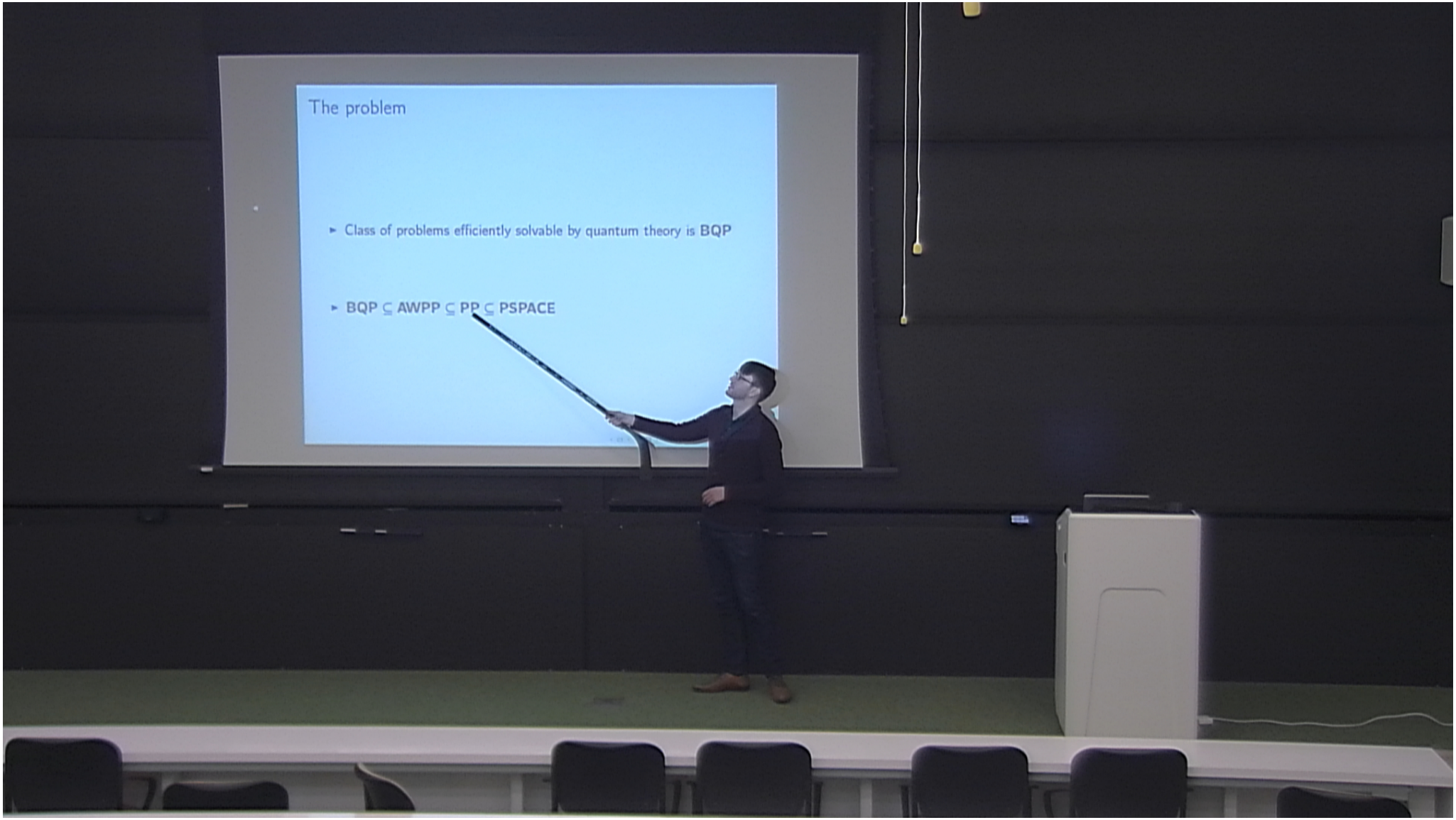
- Quantum theory offers dramatic new advantages for various information theoretic tasks
- What broad relationships exist between physical principles and information theoretic advantages?

Motivation

- Much progress has already been made in understanding connections between physical principles and some tasks
- Insights resulted in device independent cryptography, connection between non-locality and communication complexity, etc...

The problem

- Class of problems efficiently solvable by quantum theory is **BQP**
- **BQP** $\subseteq$ **AWPP** $\subseteq$ **PP** $\subseteq$ **PSPACE**


The problem
Class of problems efficiently solvable by quantum theory is BQP
BQP ⊆ AWPP ⊆ PP ⊆ PSPACE

The problem

- Class of problems efficiently solvable by quantum theory is **BQP**
- **BQP** $\subseteq$ **AWPP** $\subseteq$ **PP** $\subseteq$ **PSPACE**

The problem

- **PP** contains all problems that can be solved by a classical random computer that must get the answer right with probability $> 1/2$

- **PSPACE** contains all problems that can be solved by a classical computer using a polynomial amount of memory

The problem

Problem : What is the minimal set of physical principles such that efficient computation in an operational theory satisfies this inclusion?

Operational part

- *Tests* and *systems* are the primitive notions of operational theories

Operational part

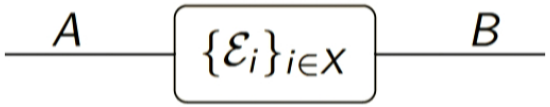
1. Tests represent one use of a physical device with input/output ports and a classical pointer

2. When a physical device is used, the pointer ends up in one of a (finite) number of classical outcomes $i \in X$. This tells us some *event* has occurred. A test is a collection of events $\{\mathcal{E}_i\}_{i \in X}$.

3. Systems label the input/output of physical devices. Systems come in different types $A, B, C, \ldots$

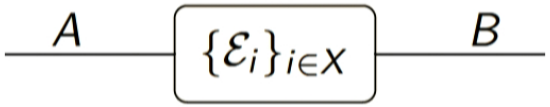
Diagrams

We can represent a test diagrammatically as follows:



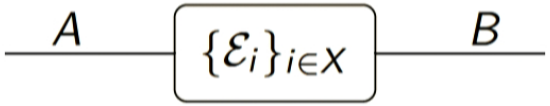
Diagrams

We can represent a test diagrammatically as follows:



Diagrams

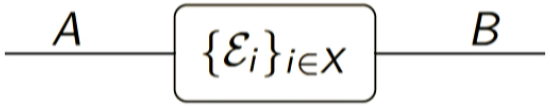
We can represent a test diagrammatically as follows:



We represent a specific event diagrammatically as:

Diagrams

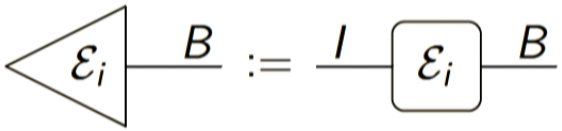
We can represent a test diagrammatically as follows:



We represent a specific event diagrammatically as:

Diagrams

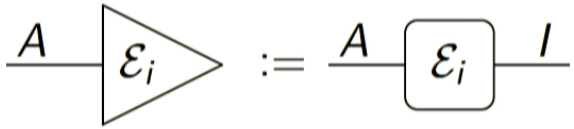
Preparation events have the trivial system as input.



Algebraically, we use 'Dirac-like' notation $|\mathcal{E}_i)_B$ to represent a preparation event.

Diagrams

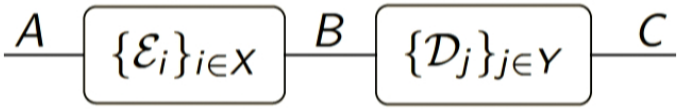
Observation events have the trivial system as output.

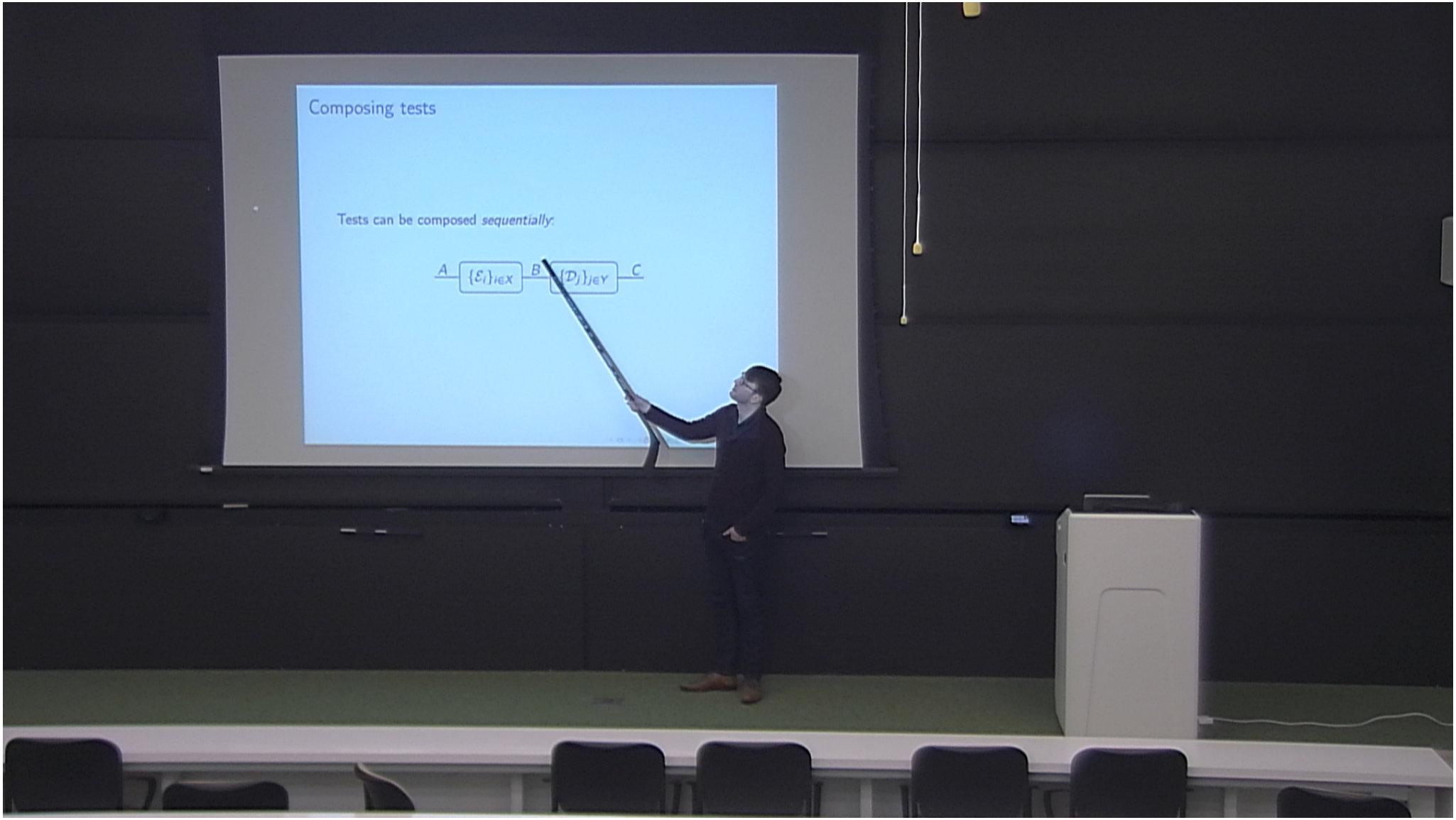


Algebraically, we use 'Dirac-like' notation ($\mathcal{E}_i|_A$ to represent a observation event.

Composing tests

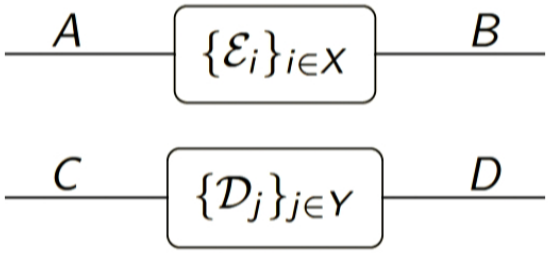
Tests can be composed *sequentially*:




Composing tests
Tests can be composed *sequentially*:
A
$\{\mathcal{E}_i\}_{i\in X}$
B
$\{\mathcal{D}_j\}_{j\in Y}$
C

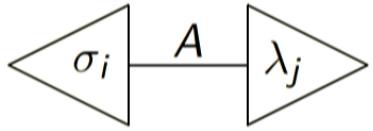
Composing tests

and in *parallel*:



Probabilistic part

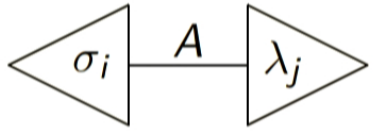
We demand that closed circuits give probabilities:

$P(i,j) :=$ 

This induces a linear structure

Probabilistic part

We demand that closed circuits give probabilities:

$P(i,j) :=$ 

This induces a linear structure

Probabilistic part

- Events $\mathcal{E}_0$ and $\mathcal{E}_1$ are *equivalent* if replacing $\mathcal{E}_0$ with $\mathcal{E}_1$ in every closed circuit does not change the probabilities.

- Results in $\mathbf{Transf}(\mathbf{A}, \mathbf{B})$, $\mathbf{St}(\mathbf{B})$ and $\mathbf{Eff}(\mathbf{A})$

Probabilistic part

- Events $\mathcal{E}_0$ and $\mathcal{E}_1$ are *equivalent* if replacing $\mathcal{E}_0$ with $\mathcal{E}_1$ in every closed circuit does not change the probabilities.

- Results in $\mathbf{Transf}(\mathbf{A}, \mathbf{B})$, $\mathbf{St}(\mathbf{B})$ and $\mathbf{Eff}(\mathbf{A})$

Probabilistic part

- $T = T_0 + rT_1 \iff P(T) = P(T_0) + rP(T_1)$, $\forall$ closed circuits

- Results in **real** vector spaces $\mathbf{Transf}_{\mathbb{R}}(\mathbf{A}, \mathbf{B})$, $\mathbf{St}_{\mathbb{R}}(\mathbf{B})$ and $\mathbf{Eff}_{\mathbb{R}}(\mathbf{A})$

Probabilistic part

- $T = T_0 + rT_1 \iff P(T) = P(T_0) + rP(T_1)$, $\forall$ closed circuits

- Results in **real** vector spaces $\mathbf{Transf}_{\mathbb{R}}(\mathbf{A}, \mathbf{B})$, $\mathbf{St}_{\mathbb{R}}(\mathbf{B})$ and $\mathbf{Eff}_{\mathbb{R}}(\mathbf{A})$

Probabilistic part

- $T = T_0 + rT_1 \iff P(T) = P(T_0) + rP(T_1)$, $\forall$ closed circuits

- Results in **real** vector spaces $\mathbf{Transf}_{\mathbb{R}}(\mathbf{A}, \mathbf{B})$, $\mathbf{St}_{\mathbb{R}}(\mathbf{B})$ and $\mathbf{Eff}_{\mathbb{R}}(\mathbf{A})$

Probabilistic part

- $T = T_0 + rT_1 \iff P(T) = P(T_0) + rP(T_1)$, $\forall$ closed circuits

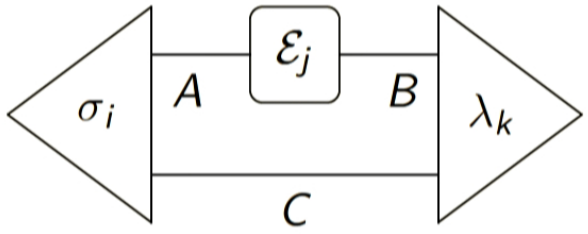
- Results in **real** vector spaces $\mathbf{Transf}_{\mathbb{R}}(\mathbf{A}, \mathbf{B})$, $\mathbf{St}_{\mathbb{R}}(\mathbf{B})$ and $\mathbf{Eff}_{\mathbb{R}}(\mathbf{A})$

Probabilistic part

Assumption: All vector spaces are finite dimensional

Joint probabilities for circuits

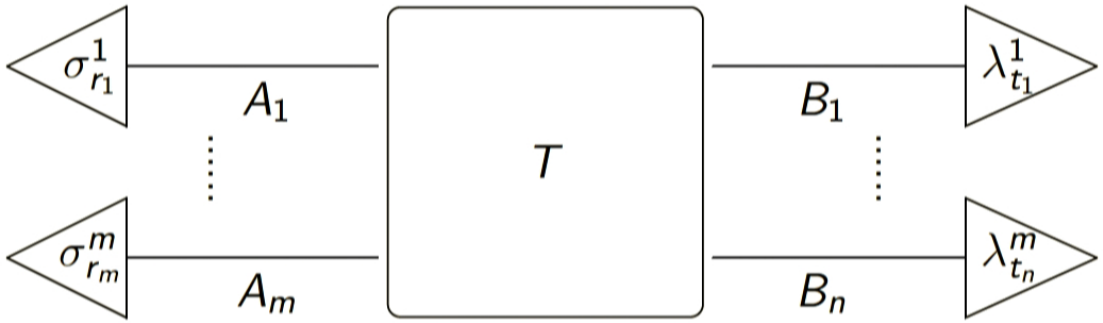
We can now associate joint probabilities to events occurring in an experiment:

$P(i,j,k) =$ 

Algebraically, we write $P(i,j,k) = (\lambda_k|_{BC}(\mathcal{E}_j \otimes \mathcal{I}_C)|\sigma_i)_{AC}$.

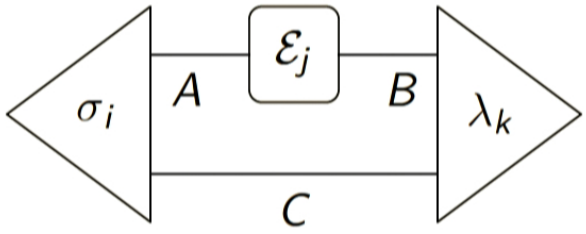
Tomographic locality

A theory satisfies **tomographic locality** if every transformation can be fully characterised by local process tomography



Tomographic locality

Vector space tensor product:



$$M_k^3.(M_j^2 \otimes I_C).M_i^1 = M_k^3.(M'^2_j).M_i^1$$

Causality

- A generalised probabilistic theory is **causal** if the probability of a preparation event is independent of the choice of which observation test follows the preparation

- For all $\{(\lambda_j|\}_j$ and $\{(\theta_k|\}_k$ we have

$$\sum_j (\lambda_j|\sigma_i) = \sum_k (\theta_k|\sigma_i)$$

Causality

We will **not** assume all theories are casual

Example

Quantum theory:

1. Systems are finite dimensional complex Hilbert spaces.
2. $St_{\mathbb{R}}(\mathbb{C}^d)$ is real vector space of hermitian operators on $\mathbb{C}^d$, i.e density matrices.
3. Tests are collections of CP trace non-increasing maps, i.e quantum instruments.

Computation

- Can draw circuits of experimental set-up and the specific events that took place in runs of the experiment.
- What do we need for these circuits to be a meaningful model of computation?
- Need to define *uniform family of circuits* for general probabilistic theories.

Uniform circuits

- Given the matrix M representing (a particular outcome of) a gate in $\mathcal{G}$, a Turing machine can output a matrix $\widetilde{M}$ with rational entries, such that $|(M - \widetilde{M})_{ij}| \leq \epsilon$, in time polynomial in $\log(1/\epsilon)$

- There is a Turing machine that, acting on input x, outputs a classical description of C_x in poly-time

Uniform circuits

- Given the matrix M representing (a particular outcome of) a gate in $\mathcal{G}$, a Turing machine can output a matrix $\widetilde{M}$ with rational entries, such that $|(M - \widetilde{M})_{ij}| \leq \epsilon$, in time polynomial in $\log(1/\epsilon)$

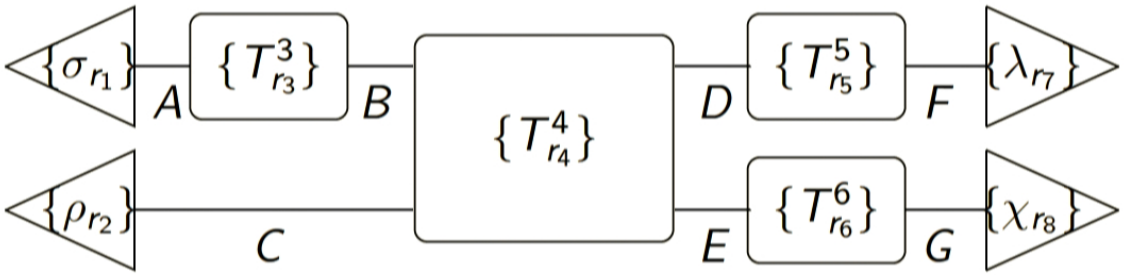
- There is a Turing machine that, acting on input x, outputs a classical description of C_x in poly-time

Acceptance criterion

- In quantum computation, all gates are deterministic and all measurements can be postponed until end
- Accepts an input if measurement of first outcome qubit is $|0\rangle$
- In an arbitrary generalised probabilistic theory this may not be the case, need more general acceptance criterion

Acceptance criterion

Construct circuit:

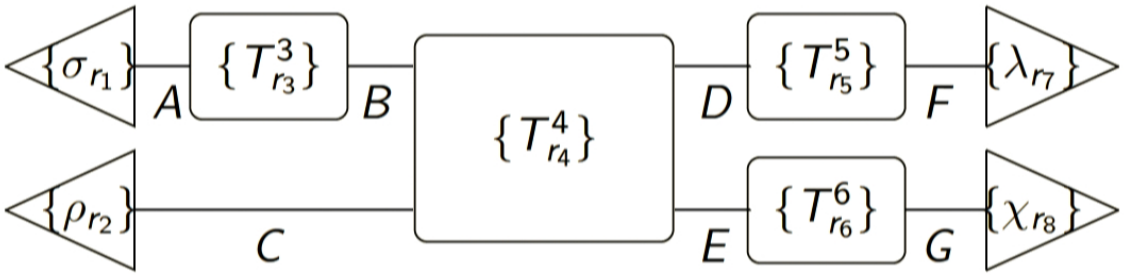


Outcome:

$$P(r_1,\ldots,r_8) = (\chi_{r_8}|(\lambda_{r_7}|\left(T^6_{r_6}\otimes T^5_{r_5}\right)T^4_{r_4}\left(T^3_{r_3}\otimes I_C\right)|\rho_{r_2})|\sigma_{r_1}).$$

Acceptance criterion

Construct circuit:



Outcome:

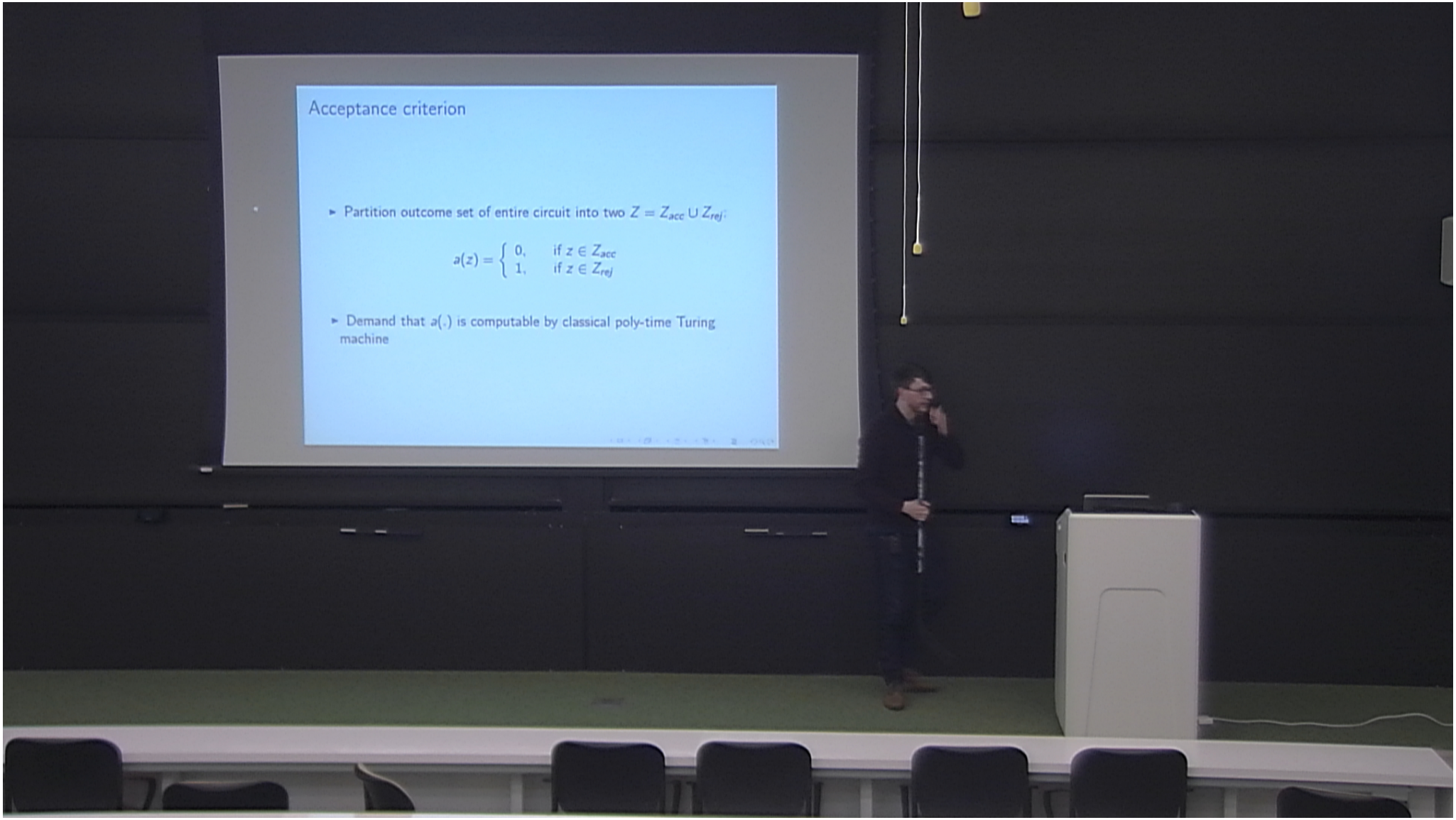
$$P(r_1,\ldots,r_8) = (\chi_{r_8}|(\lambda_{r_7}|(T^6_{r_6}\otimes T^5_{r_5})\,T^4_{r_4}(T^3_{r_3}\otimes I_C)|\rho_{r_2})|\sigma_{r_1}).$$

Acceptance criterion

- Partition outcome set of entire circuit into two $Z = Z_{acc} \cup Z_{rej}$:

$$a(z) = \begin{cases} 0, & \text{if } z \in Z_{acc} \\ 1, & \text{if } z \in Z_{rej} \end{cases}$$

- Demand that $a(.)$ is computable by classical poly-time Turing machine


Acceptance criterion
Partition outcome set of entire circuit into two $Z = Z_{acc} \cup Z_{rej}$:
$$a(z) = \begin{cases} 0, & \text{if } z \in Z_{acc} \\ 1, & \text{if } z \in Z_{rej} \end{cases}$$
Demand that $a(.)$ is computable by classical poly-time Turing machine

Acceptance criterion

Probability to accept input x is then

$$P_x(\text{accept}) = \sum_{z|a(z)=0} P(z),$$

sum ranges over all possible outcome strings of the circuit C_x

BGP

For a generalised probabilistic theory **G**, a language $\mathcal{L}$ is in the class **BGP** if there exists a poly-sized uniform family of circuits in **G**, and an efficient acceptance criterion, such that

1. $x \in \mathcal{L}$ is accepted with probability at least $\frac{2}{3}$.

2. $x \notin \mathcal{L}$ is accepted with probability at most $\frac{1}{3}$.

Upper bounds

For any generalised probabilistic theory **G**, we have

$$\mathbf{BGP} \subseteq \mathbf{AWPP} \subseteq \mathbf{PP} \subseteq \mathbf{PSPACE}$$

Upper bounds

For any generalised probabilistic theory **G**, we have

$$\mathbf{BGP} \subseteq \mathbf{AWPP} \subseteq \mathbf{PP} \subseteq \mathbf{PSPACE}$$

Proof sketch

- Consider a general circuit C_x, with $q(|x|)$ gates from $\mathcal{G}$
- Tensoring these gates with identity transformations on systems on which they do not act, and padding them with rows and columns of zeros, results in a sequence of square matrices $M^{r_q,q}, \ldots, M^{r_1,1}$
- $M^{r_n,n}$ is the matrix representing the r_n^{th} outcome of the n^{th} gate

Proof sketch

- Consider a general circuit C_x, with $q(|x|)$ gates from $\mathcal{G}$
- Tensoring these gates with identity transformations on systems on which they do not act, and padding them with rows and columns of zeros, results in a sequence of square matrices $M^{r_q,q}, \ldots, M^{r_1,1}$
- $M^{r_n,n}$ is the matrix representing the r_n^{th} outcome of the n^{th} gate

Proof sketch

- The matrix entries of gates from $\mathcal{G}$ can be efficiently computed
- Tomographic locality implies that entries of $M^{r_n,n}$ can also be efficiently computed

Proof sketch

- The matrix entries of gates from $\mathcal{G}$ can be efficiently computed

- Tomographic locality implies that entries of $M^{r_n,n}$ can also be efficiently computed

Proof sketch

The probability for outcome $z = r_1 \dots r_q$, is given by

$$b^T . M^{r_q,q} \dots M^{r_2,2} M^{r_1,1} . b = \sum_{\{i_1,\dots,i_{q-1}\}} M^{r_q,q}_{1 i_{q-1}} \dots M^{r_2,2}_{i_2 i_1} M^{r_1,1}_{i_1 1}$$

where b is the vector $b = (1, 0, \dots, 0)$

Proof sketch

- Exponentially long sum, but each entry is a product of polynomially many terms.
- Each term in sum can be efficiently calculated
- Entire sum can be calculated in polynomial space, as individual terms can be erased after being added to running total.

Role of assumptions

- Relies on the ability to decompose the acceptance probability of the computation in a form reminiscent of a (discrete) Feynman path integral.

- Linear structure of generalised probabilistic theories arises from the requirement that a physical theory should be able to give probabilistic predictions about the occurrence of possible outcomes.

Role of assumptions

- Requires ability to compute efficiently the entries in the matrices representing the transformations applied in parallel in a specific circuit.

- If a transformation from **A** to **B** acts on one half of a system **AC**, there may be no simple way to relate the linear map $\mathrm{St}(\mathbf{AC}) \to \mathrm{St}(\mathbf{BC})$ to the action of the transformation when it is applied to a system **A** on its own

Role of assumptions

- Requires ability to compute efficiently the entries in the matrices representing the transformations applied in parallel in a specific circuit.

- If a transformation from **A** to **B** acts on one half of a system **AC**, there may be no simple way to relate the linear map $\mathrm{St}(\mathbf{AC}) \to \mathrm{St}(\mathbf{BC})$ to the action of the transformation when it is applied to a system **A** on its own

Post-selection

- $\mathbf{Post-BGP} \subseteq \mathbf{PP} = \mathbf{Post-BQP}$

- In a world with post-selection, quantum theory is optimal for computation in the space of all generalised theories

Post-selection

- $\mathbf{Post-BGP} \subseteq \mathbf{PP} = \mathbf{Post-BQP}$

- In a world with post-selection, quantum theory is optimal for computation in the space of all generalised theories

Post-selection

- Under reasonable assumptions $\mathbf{DQC} \subsetneq \mathbf{BQP}$
- But $\mathbf{Post} - \mathbf{DQP} = \mathbf{Post} - \mathbf{BQP}$

Further results

- Non-trivial reversible transformations imply $\mathbf{BPP} \subseteq \mathbf{BGP}$
- Generalised probabilistic oracle hard to define, but can define 'classical oracle' in causal theory
- 'Classical oracle' separation result: $\exists \mathbf{A}$ such that, for all causal theories, $\mathbf{NP}^{\mathbf{A}} \not\subseteq \mathbf{BGP}^{\mathbf{A}}_{cl}$

Conclusion

- Defined the class of problems that can be efficiently solved by an arbitrary general probabilistic theory
- Theories satisfying tomographic locality satisfy the best known quantum bounds
- In a world with post-selection, quantum theory is optimal for computation in the space of all theories satisfying tomographic locality

Outlook

- Even though we have not assumed the causality principle, the gates in our circuits appear in a fixed structure

- Investigate the computational power of theories in which there is no definite structure?

Further results

- Non-trivial reversible transformations imply $\mathbf{BPP} \subseteq \mathbf{BGP}$
- Generalised probabilistic oracle hard to define, but can define 'classical oracle' in causal theory
- 'Classical oracle' separation result: $\exists \mathbf{A}$ such that, for all causal theories, $\mathbf{NP}^{\mathbf{A}} \not\subseteq \mathbf{BGP}^{\mathbf{A}}_{cl}$