Computation in generalised probabilistic theories

Ciarán Lee

Joint work with Jon Barrett

arXiv:1412.8671

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...

Motivation

- Relatively little has been learned about the connection between physical principles and computation
- We consider computation in a framework suitable for describing arbitrary operational theories

Motivation

- An operational theory specifies a set of laboratory devices that can be connected together in different ways, and assigns probabilities to experimental outcomes.

Outline

- Introduce problem
- Framework for operationally-defined theories
- Computational model and results

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?

Introduction to framework

We work in the circuit framework developed by Hardy and Chiribella, D'Ariano and Perinotti.

Introduction to framework

- *Tests* are the primitive notions of operational theories
- Represent one use of a physical device with input/output ports and a classical pointer

Introduction to framework

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

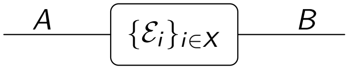
Introductions to framework

- Physical systems can be thought of as passing through the input and output ports of tests

- Systems labelled by $A, B, C, \ldots$

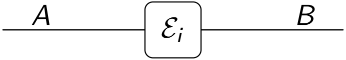
Diagrams

We can represent a test diagrammatically as follows:



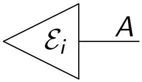
Diagrams

We represent a specific event diagrammatically as:



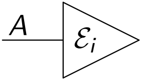
Diagrams

Test with no inputs *prepares* a system



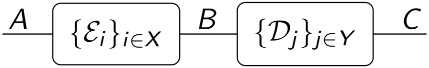
Diagrams

Tests with no outputs *measures* a system



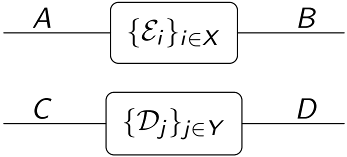
Composing tests

Tests can be composed *sequentially*:



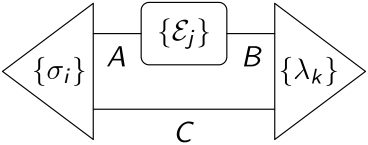
Composing tests

and in *parallel*:



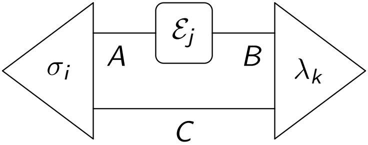
Circuits

In an operational theory, one can draw circuits representing the connections of physical devices in an experiment:



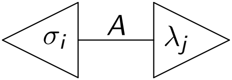
Circuits

and circuit outcomes representing which specific events took place in said experiment:



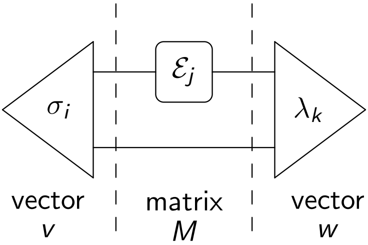
Probabilistic part

We demand that closed circuits give probabilities:

$P(i, j) \;:=$ 

Linear structure

Probabilistic structure imposes linear structure:



$$P(i, j, k) = v.M.w$$

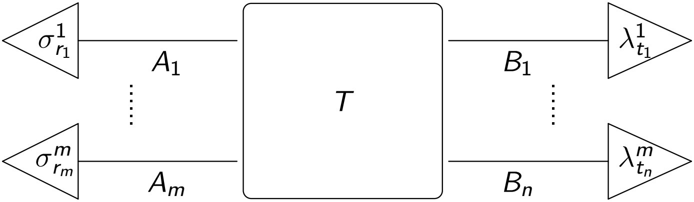
Tomographic locality

- Every transformation from **A** to **B** induces a linear map between the corresponding vectors

- 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 $\mathbf{AC} \to \mathbf{BC}$ to the action of the transformation when it is applied to a system **A** on its own

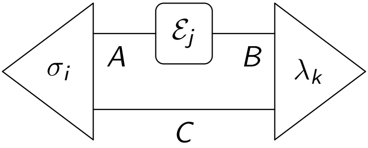
Tomographic locality

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



Tomographic locality

Vector space tensor product:



$$v.M.w = v.(G \otimes \mathbb{I}).w$$

Tomographic locality

Assumption: Tomographic locality is satisfied

Causality

- An operational theory is **causal** if the probability of a preparation is independent of the choice of which measurement 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

- **Causal** = 'no signalling from the future'

- Nothing obviously pathological about theories *without* causality

Causality

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

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 operational theories.

Uniform circuits

- In quantum/classical circuit model, a circuit family $\{C_n\}$ is indexed by input system size n.
- Each C_n built by composing a polynomial number of gates.
- 'Classical description' of C_n can be efficiently computed

Uniform circuits

- In generalised circuit model, the entire circuit encodes the problem instance
- Circuit family $\{C_x\}$, for x a classical string
- Each circuit is build with a polynomial number of gates from a (finite) gate set $\mathcal{G}$

Uniform circuits

- Given the matrix M representing (a particular outcome of) a gate in $\mathcal{G}$, can approximate its entries efficiently

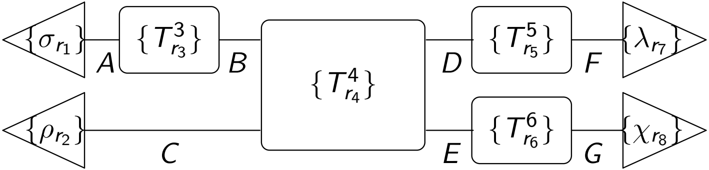
- Classical description of C_x can be computed efficiently

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 operational 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

- 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

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

sum ranges over all possible outcome strings of the circuit C_x

BGP

For an operational 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

Theorem

For any operational theory **G** *that satisfies tomographic locality, we have*

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

Role of assumptions

1. **Linear structure:** arises from the requirement that a physical theory should be able to give probabilistic predictions about the occurrence of possible outcomes

2. **Tomographic locality:** ability to efficiently compute the entries of matrices representing transformations applied in parallel

A question

- Best upper bounds on **BQP** follow from very mild assumptions and don't exploit any uniquely quantum features (don't even need a notion of causality!)

- Can we do better?

Further questions

- Can quantum theory simulate computation in any any operationally-defined theory? If so, could provide explanation of quantum speed-up

- Certain situations in which quantum theory is provably optimal for computational in this landscape of operationally-defined theories

Post-selection

- Aaronson has introduced the notion of *post-selected* quantum circuits

- Quantum circuits with a ‘post-selected’ register. Only those runs of the computation for which a measurement of the post-selected qubit yields 0 are considered.

Post-selection

- Aaronson has shown that $\mathbf{PostBQP} = \mathbf{PP}$

- Thus a quantum computer with post-selection can simulate computation in any other generalised probabilistic theory

Post-selection

- Can also define generalised circuits with post-selection
- Here we can post-select on any (efficiently computable) subset of the circuit outcomes

Post-selection

Theorem

For any tomographically local theory **G***, we have*

$$\mathbf{PostBGP} \subseteq \mathbf{PP} = \mathbf{PostBQP}$$

In a world with post-selection, quantum theory is optimal for computation in the space of all (tomographically local) operational theories

Conclusion

- Defined the class of problems that can be efficiently solved by an arbitrary operationally-defined 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 questions

- Does there exist an operationally-defined theory that can simulate quantum computation?

- If so, could compare to quantum theory in the hope of learning why quantum theory **isn't** that way

Thank you!

Post-selection

- Can we view **PostBGP** $\subseteq$ **PostBQP** as evidence that quantum theory *on its own* is optimal (or at least powerful) for computation in the space of general theories?
- Caution is needed

Post-selection

- Consider the ‘one clean qubit model’ **DQC**

- Restricted form of quantum computation where input to circuit is one pure qubit with as many maximally mixed qubits as desired

Post-selection

- Under reasonable assumptions $\mathbf{DQC} \subsetneq \mathbf{BQP}$
- But $\mathbf{PostDQP} = \mathbf{PostBQP}$

Post-selection

- So while **PostBQP** $\subseteq$ **PostDQP**
- Under reasonable assumptions it is not the case that **BQP** $\subseteq$ **DQP**

Further results

- Non-trivial reversible transformations imply $\mathbf{BPP} \subseteq \mathbf{BGP}$ for non-classical $\mathbf{G}$
- 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, $\mathbf{NP^A} \not\subseteq \mathbf{BGP}^{\mathbf{A}}_{cl}$

Proof sketch of **PSPACE**

- 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 of **PSPACE**

- 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 of **PSPACE**

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

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

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

Proof sketch of **PSPACE**

- 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.