Contents

Chapter 1

Bayesian probability theory & classical causality

On this account all the sciences would only be unconscious applications of the calculus of probabilities; to condemn this calculus would be to condemn science entirely.
— H. Poincaré [1]

1.1 Probability space and probability distributions

As a mathematical theory, probability theory requires us to define three elements: the *sample space*, the *algebra of events*, and a *measure* on this algebra.

The sample space is essentially a set Ω (for this school will always be considered finite), from which we will draw the events. Events are simply subsets $X \subset \Omega$, and will be associated with measurement outcomes of experiments. In particular, a single measurement outcome is a subset with a single element in the event space, $x \in \Omega$.

The *algebra of events* is a collection $\mathcal{X}$ of the subsets of Ω relevant for the theory. In particular, this collection needs to satisfy some properties:

1. $\emptyset \in \mathcal{X}$ and $\Omega \in \mathcal{X}$. This means that the set of "no events" and the set of "all events" are both relevant events and must be in the algebra;

2. Given X and X' both in $\mathcal{X}$, the sets $X \cup X'$, $X \cap X'$, and X/X' are also in $\mathcal{X}$. This means that we can make logical combinations between events (respectively OR, AND, and NOT). As a consequence, $\bigcup_{i=1}^{N} X_i \in \mathcal{X}$, where N is the total number of subsets in the algebra.

From now on, our events will be subsets $X \in \mathcal{X}$. This way, we are ensuring that there is a structure allowing us to make logical combinations of events without incurring any abnormalities.

Finally, a *measure* over this algebra is simply a map $p : \mathcal{X} \to \mathbb{R}$ such that $X \mapsto p(X) \in \mathbb{R}$. A *probability distribution* is a measure over an algebra of events satisfying the following properties:

1. $p(X) \geq 0$, $\forall X \in \mathcal{X}$;
2. $p(\Omega) = 1$, which is called *normalisation*;
3. For any collection of disjoint events $X_1, ..., X_N$ with $X_i \cap X_j = \emptyset$, $\forall i \neq j = 1, ..., N$, then

$$p\left(\cup_{i=1}^{N} X_i\right) = \sum_{i=1}^{N} p(X_i). \tag{1.1}$$

These properties, called *Kolmogorov axioms*, ensure many of the usual features of probabilities we will be using along the subject. For instance, you can derive from these axioms and basic set theory the following equation

$$p(X \cup X') = p(X) + p(X') - p(X \cap X'), \quad \forall X, X' \in \mathcal{X}. \tag{1.2}$$

Another relevant concept for this course is the one of *conditional probability*. It is meant to capture the likelihood with which an event will occur given that another event has occurred. It consists of a map $p(\cdot|X) : \mathcal{X} \to \mathbb{R}$, defined for all $X \in \mathcal{X}$ such that $p(X) > 0$, and such that

$$X' \mapsto p(X'|X) := \frac{p(X' \cap X)}{p(X)}. \tag{1.3}$$

The demand that $p(X)$ is nonzero is natural since we want to condition the event X' to something that has occurred. Conditional probabilities are important because it is from them that we construct the notion of *independence*: two events X, X' are said to be *independent* or *uncorrelated* when

$$p(X'|X) = p(X'), \tag{1.4}$$

or from Eq. 1.3,

$$p(X' \cap X) = p(X')p(X). \tag{1.5}$$

Also from Eq. 1.3, we can find that

$$p(X' \cap X) = p(X'|X)p(X), \tag{1.6}$$

where we just multiplied both sides of Eq. 1.3 by $p(X)$. Therefore

$$\begin{aligned} p(X|X') &= \frac{p(X \cap X')}{p(X')} && (1.7) \\ &= \frac{p(X' \cap X)}{p(X')} && (1.8) \\ &= p(X'|X)\frac{p(X)}{p(X')}, && (1.9) \end{aligned}$$

which is the so called *Bayes' rule*. It simply states that one can invert the conditioning between two variables by multiplying it by the ratio between the individual probabilities.

Example 1 – Tossing a coin

Consider the following sample space associated with the tossing of a single coin: $\Omega = \{H, T\}$, where H is a shortcut for the string *"the outcome of the tossing is heads"*, and similar for T and tails. The algebra $\mathcal{X}$ of this sample space is given by

$$\mathcal{X} = \{\{\emptyset\}, \{H\}, \{T\}, \{H, T\}\}. \tag{1.10}$$

Notice that it satisfies the properties of the algebra of events: it contains the empty set as well as the whole sample space, and any logical combination of subsets is in the algebra. To see that, notice for instance that

$$H \cup T = \{H, T\} \in \mathcal{X}; \quad H \cap T = \emptyset \in \mathcal{X}; \quad H/T = \{H\} \in \mathcal{X}. \tag{1.11}$$

Feel free to try with any other combination of two subsets in $\mathcal{X}$! Consider now the probability measure $p : \mathcal{X} \to \mathbb{R}$ such that

$$p(H) = p(T) = \frac{1}{2}. \tag{1.12}$$

Notice that this measure satisfies the Kolmogorov axioms: it is non-negative and normalised, so the probability of $H \cup T$ is just the sum of the individual probabilities.

Example 2 – Tossing two coins

Consider now the sample space associated with the tossing of two distinguishable coins, i.e., always two coin outcomes are produced. The possible outcomes will be $\Omega = \{HH, HT, TH, TT\}$, where HH is the shortening for *"the outcome of the tossing is heads for coin 1 and heads for coin 2"*, and similar for the other events.

The algebra of this set will contain all possible partitions of the set, such that

$$\begin{aligned} \mathcal{X} = \{\{\emptyset\}, \{HH\}, \{HT\}, \{TH\}, \{TT\}, \{HH, HT\}, \{HH, TH\}, \{HH, TT\}, \\ \{HT, TH\}, \{HT, TT\}, \{TH, TT\}, \{HH, HT, TH\}, \{HH, HT, TT\}, \\ \{HH, TH, TT\}, \{HT, TH, TT\}, \{HH, HT, TH, TT\}\}. \end{aligned} \tag{1.13}$$

Finally, the map

$$p(HH) = p(HT) = p(TH) = p(TT) = \frac{1}{4} \tag{1.14}$$

forms a probability measure.

1.2 Random variables and expectation values

Although the previous definition of the sample space Ω algebra of events $\mathcal{X}$ is sufficient to construct probabilities and the most important rules to manipulate them, the events X are still abstract concepts. They could be, for instance, sentences in English describing how the outcome of a measurement is perceived in a lab. It is convenient instead to treat events as *numbers*, which we are much more used to manipulating.

A random variable is a function $A : \Omega \to \mathbb{R}$ mapping each event x to a real number $a = A(x)$, such that it has a reasonably defined inverse map. This means that $X = A^{-1}(B)$ is in the algebra of events for any upper bounded set $B \in \mathbb{R}$. This demand ensures that we can attribute probabilities to the random variables the same way as we attribute to the events themselves, such that

$$p(A \in B) := p(A^{-1}(B)). \tag{1.15}$$

Usually, we will want to estimate *expectation values* of some random variable. This is defined as

$$\langle A \rangle := \sum_{a \in Im(A)} a\, p(A = a), \tag{1.16}$$

and this definition can be generalised for any function $f : \mathbb{R} \to \mathbb{R}$ over the numbers a. That is,

$$\langle f(A) \rangle := \sum_{a \in Im(A)} f(a)\, p(A = a). \tag{1.17}$$

In particular, expectation values of functions that are just of the form $f(a) = a^m$ for some $m \in \mathbb{N}^*$ receive the special name of *moments*, and are relevant for the majority of statistical analyses. This however goes a bit beyond the scope of this course and thus will not be commented on.

Example 3 – Correlators

Consider again the case where a coin is tossed. Let us consider the map $A : \Omega \to \mathbb{R}$ such that

$$A(H) = 1, \quad A(T) = -1. \tag{1.18}$$

In this case, the expectation value of A is simply

$$\langle A \rangle = A(H)p(H) + A(T)p(T) = p(H) - p(T). \tag{1.19}$$

In the case of two coin tosses, we can define a new map A' as a function of A, such that

$$A'(HH) = A(H) \cdot A(H), \qquad A'(HT) = A(H) \cdot A(T), \tag{1.20}$$

$$A'(TH) = A(T) \cdot A(H), \qquad A'(TT) = A(T) \cdot A(T). \tag{1.21}$$

The expectation value of A' will thus be

$$\begin{aligned} \langle A' \rangle &= A(H)A(H)p(HH) + A(H)A(T)[p(HT) + p(TH)] + A(T)A(T)p(TT) \quad (1.22) \\ &= p(HH) - p(HT) - p(TH) + p(TT). \quad (1.23) \end{aligned}$$

This quantity often receives the name of *correlator*, and can be ubiquitously found in the literature of Bell nonlocality.

1.3 A short disclaimer

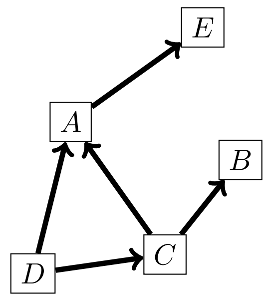
In this school, we take an operationalist approach to probabilities, attaching to them some substantial meaning about experiments. In fact, probabilities in this school will not represent "*how many times this particular event happens among all possible events when we repeat this same experiment to infinity*". Instead, we will interpret probabilities as the degree of likelihood that a rational agent is willing to attribute to the occurrence of an event.

Despite the bad impression that this called *bayesian* bias might give out, it is still as realist as the *frequentist* one in the sense that infinite repetitions of an experiment are never possible. Therefore, you can think of the frequentist interpretation as being subjective too, since it is up to the experimenter to determine when a sufficient number of runs of the experiment have been carried out. It is crucial to emphasize that everything discussed in this school can also be easily imported into the frequentist interpretation. The subject will however not dive deep into this topic, but students are free to search for literature on the topic if they are interested.

1.4 Classical causality

We are now acquainted with bayesian probability theory, but how can we tell that a correlation between random variables $p(A_1, ..., A_n)$ is classical? One way of doing so is by assuming that there is an underlying *classically causal model* to the correlations.

Let's begin by defining a notion of a classical causal model. These are described by a collection of random variables and a collection of arrows between these representing when one random has a direct causal influence on another. We can graphically denote these by *directed acyclic graphs* (DAGs), for example:

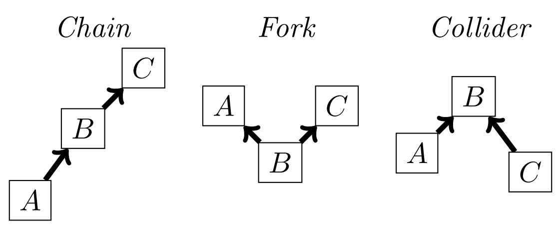


(1.24)

where here the random variables are $A, ..., E$, and there are five arrows representing direct causal influence.

The fact that this is a directed graph means that the connections are represented by arrows, which captures the fact that, for example, A is the cause of E rather than vice versa. The fact that it is acyclic means that we don't get any causal loops. For example, we don't find a situation where $A \to B \to C \to A$ which would be paradoxical as it now says that A is the cause of itself. We then want the sorts of correlations that we can have between these random variables to always come from a causal connection.

Definition 1.4.1 *(Reichenbach's principle) A correlation between two random variables satisfies Reichenbach's principle if (i) there is a direct causal influence from one to the other; (ii) there is some common cause that influences them both; or (iii) there may be some common future that one has conditioned on which induces the correlation. In other words, two correlated random variables A and C must be causally explained by one of the following DAGs:*



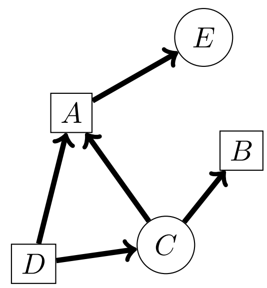
(1.25)

This implies that all correlations between the random variables $A_1, ..., A_n$ satisfying Reichenbach's principle will have the form

$$p(A_1, ..., A_n) = \prod_{i=1}^{n} p(A_i|\mathsf{Pa}(A_i)), \tag{1.26}$$

where $\mathsf{Pa}(A_i)$ is the set of *parent nodes* of A_i, i.e., the set of all variables $A_{j \neq i}$ with an arrow pointing towards A_i. Evidently, $p(A_i|A_j)$ are all valid probability distributions, satisfying all the necessary axioms.

Now, the framework as we have developed it so far assumes that all of the nodes are, in principle, observed. We may marginalise over one variable to find a particular distribution, but we can also condition over all of them taking particular values. However, there will be many experiments in which a particular variable is not observable. We denote these as circular nodes, for example:



(1.27)

In this case, the correlations between the observed random variables $A_1, ..., A_n$ when there are $O_1, ..., O_m$ unobserved ones will have the form

$$p(A_1, ..., A_N) = \sum_{o_1,...,o_m} \prod_{i=1}^{n} p(A_i|\mathsf{Pa}(A_i)) \prod_{j=1}^{m} p(O_j = o_j|\mathsf{Pa}(O_j)), \tag{1.28}$$

where all that is going on here is that we're imagining that there is some underlying description in which all nodes are observed and we're just marginalising over those that we have decided are unobserved. Despite these probabilities being unconditioned, we can always apply Bayes' rule (as long as the conditioners are nonzero).

References

1. H. Poincaré, *La science et l'hypothèse*, Paris, pp. 171,217 (1906).

2. H. P. Brauer & F. Petruccione, *The Theory of Open Quantum Systems.* Oxford University Press, Oxford (2002).

3. D. Frauchiger, *A Non-Probabilistic Framework for Scientific Theories.* PhD Thesis. ETH Zürich, Zürich (2016).

4. J. Pearl, *Causality.* Cambridge University Press (2009).

5. J. H. Selby, *Signatures of Non-Classicality.* Tutorials for the Quantum Information Technology Masters program of the University of Gdańsk, Gdańsk (2022).

Chapter 2

Operational Probabilistic Theories

Without an interpretation probabilities are a purely mathematical concept and thus, they cannot explain anything about nature. What makes this issue difficult is that our understanding of probabilities is overloaded from everyday experience. For instance, we may say that it is very probable that it rains tomorrow or that the probability that the universe came out the way it did is 0.0000034. But what exactly do we mean by such statements on an operational level?
— D. Frauchiger [1]

2.1 Operational language & compositional rules

Before introducing the formalism of Operational Probabilistic Theories (OPTs), you might be asking yourself what it means for a theory to be operational. Essentially, this framework wants to drop off any assumptions about the underlying reality of an experiment and focus exclusively on objective elements that are intrinsic to any experiment. An experiment here is understood as a chain of actions that can be implemented in order to investigate a physical system[1], in particular, instructions on how to prepare the system to be investigated, which questions can you make about it, and which answers you can get for each question.

An operational language (OL) Θ is the formalisation of these concepts. It consists of a tuple $\Theta := (\mathsf{Sys}, \mathsf{Test}, \mathsf{Out}, \mathsf{Event})$. The elements $S, A, B, C \in \mathsf{Sys}(\Theta)$ are labeling different system types: you can think of them as coins, electrons, Hilbert or Euclidian spaces, etc.

Events $\mathscr{T}, \mathscr{A}, \mathscr{B} \in \mathsf{Event}(\Theta)$ connect system types A and B, and may have outcomes $x, a, b \in \mathsf{Out}(\Theta)$ associated to it. You can think of them as generalisations of transformations and measurements. A test $\mathsf{T}_X^{A\to B}$ is essentially a set of events from system-type A to system-type B, with outcomes in the set X, i.e., $\mathsf{T}_X^{A\to B} = \{\mathscr{T}_x^{A\to B}\}_{x\in X}$. Diagrammatically,

[1]Notice that in this framework, the existence of a physical system is assumed to be true. Anti-realist interpretations are thus beyond the scope of the framework.

the operational language is represented by boxes and wires, such that

$$\mathscr{T}_x^{A\to B} \quad\mapsto\quad \overset{A}{-}\boxed{\mathscr{T}_x}\overset{B}{-} \;; \qquad \mathsf{T}_X^{A\to B} \quad\mapsto\quad \overset{A}{-}\boxed{\mathsf{T}_X}\overset{B}{-} \tag{2.1}$$

Notice that the diagrammatic representation is convenient because it allows us to detach the subscripts $A \to B$ from the labels of tests and events.

We will talk about the outcomes $\mathsf{Out}(\Theta)$ in a moment, but they will demand special treatment. Right now, let us see how we can combine $\mathsf{Sys}(\Theta)$, $\mathsf{Event}(\Theta)$, and $\mathsf{Test}(\Theta)$ to describe complex scenarios.

The first thing we might want to do is to perform tests in sequence over a system. This is a natural request when you think of an optical table, for instance, in which you can put several lenses in sequence on the same mode. Let us then define a *sequential composition* $\circ$:

$$\circ : \mathsf{Test}(B \to C) \times \mathsf{Test}(A \to B) \to \mathsf{Test}(A \to C) \tag{2.2}$$

$$\mathsf{T}'^{B\to C}_Y \times \mathsf{T}^{A\to B}_X \mapsto (\mathsf{T}' \circ \mathsf{T})^{A\to C}_{X\times Y}. \tag{2.3}$$

Diagrammatically, this is telling that

$$\overset{A}{-}\boxed{\mathsf{T}_X}\overset{B}{-}\boxed{\mathsf{T}'_Y}\overset{C}{-} \quad\equiv\quad \overset{A}{-}\boxed{(\mathsf{T}' \circ \mathsf{T})_{X\times Y}}\overset{C}{-} \;. \tag{2.4}$$

We will also require some features from this product. For instance, it must be associative, which means that

$$\overset{A}{-}\boxed{(\mathsf{T}' \circ \mathsf{T})_{X\times Y}}\overset{C}{-}\boxed{\mathsf{T}''_Z}\overset{D}{-} \quad=\quad \overset{A}{-}\boxed{\mathsf{T}_X}\overset{B}{-}\boxed{\mathsf{T}'_Y}\overset{C}{-}\boxed{\mathsf{T}''_Z}\overset{D}{-} \quad=\quad \overset{A}{-}\boxed{\mathsf{T}_X}\overset{B}{-}\boxed{(\mathsf{T}'' \circ \mathsf{T}')_{Y\times Z}}\overset{D}{-} \;. \tag{2.5}$$

Also, for every system-type A, there should exist a *trivial test* $\mathsf{I}^{A\to A}$, such that

$$\overset{A}{-}\boxed{\mathsf{T}_X}\overset{B}{-}\boxed{\mathsf{I}}\overset{B}{-} \quad=\quad \overset{A}{-}\boxed{\mathsf{I}}\overset{A}{-}\boxed{\mathsf{T}_X}\overset{B}{-} \quad=\quad \overset{A}{-}\boxed{\mathsf{T}_X}\overset{B}{-} \;. \tag{2.6}$$

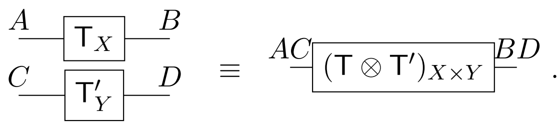
We are ignoring the outcome set of $\mathsf{I}^{A\to A}$ because it is possible to prove that it is the singleton set $\star$, and moreover I is unique for each system-type in $\mathsf{Sys}(\Theta)$. This invites us to represent $\mathsf{I}^{A\to A}$ as simply the wire associated with system-type A.

Another way of composing elements of the operational language is to have things co-exist. For instance, one could investigate a coin and a photon in the same experiment, or Alice and Bob performing simultaneous measurements in their respective labs. Take into consideration that pairs of systems, (A, B), must belong to $\mathsf{Sys}(\Theta)$ as well as their individual counterparts, so we say that $AB \in \mathsf{Sys}(\Theta)$. We thus define a *parallel composition* $\otimes$:

$$\otimes : \mathsf{Test}(A \to B) \times \mathsf{Test}(C \to D) \to \mathsf{Test}(AC \to BD) \tag{2.7}$$

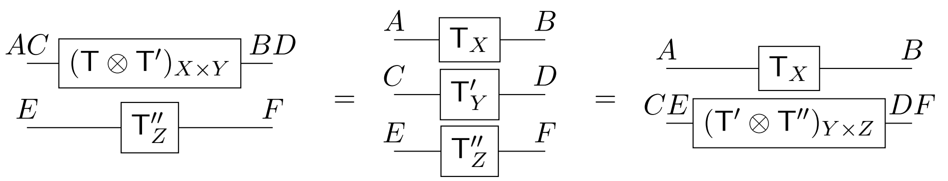
$$\mathsf{T}^{A\to B}_X \times \mathsf{T}'^{C\to D}_Y \mapsto (\mathsf{T} \otimes \mathsf{T}')^{AC\to BD}_{X\times Y}. \tag{2.8}$$

In diagrams, we have



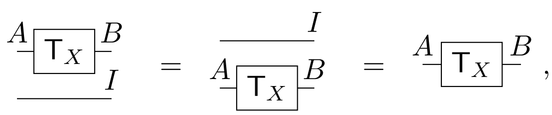
(2.9)

We will also demand associativity from this product,



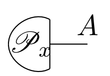
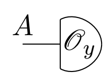
(2.10)

as well as the existence of a *trivial system-type* I, which represents not having a system and can thus be added to or ignored in any diagram:



(2.11)

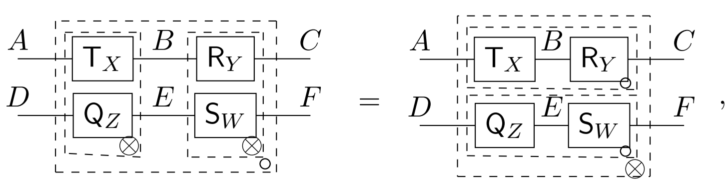
In particular, events of the form $\mathscr{P}_x^{I\to A}$ receive a special interpretation of *preparation events*. You can understand it as a procedure describing picking a coin out of a bag and tossing it, or a crystal emitting a photon in a particular state. Similarly, there are procedures going from non-trivial system types to the trivial I, $\mathscr{O}_y^{A\to I}$, which we are calling *observation events*. We will represent such events as

$\mathscr{P}_x^{I\to A} \quad \mapsto$  ; $\quad \mathscr{O}_y^{A\to I} \quad \mapsto$  . (2.12)

We finally require that for any tests $T_X^{A\to B}$, $R_Y^{B\to C}$, $Q_Z^{D\to E}$, $S_W^{E\to F}$, we have

$$(R_Y^{B\to C}\otimes S_W^{E\to F})\circ(T_X^{A\to B}\otimes Q_Z^{D\to E}) = (R_Y^{B\to C}\circ T_X^{A\to B})\otimes(S_W^{E\to F}\circ Q_Z^{D\to E}), \tag{2.13}$$

which looks very intricate algebraically, but when put into diagrams becomes very natural,



(2.14)

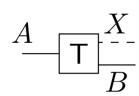
since if you just ignore all the dashed boxes it is immediate to see that the equality holds.

2.2 Outcomes & probabilistic models

Outcomes $\mathsf{Out}(\Theta)$ in the operational language represent the labels of the events in a particular test. As abstract as it might sound, you can think of a test "checking which face of a coin is upwards". The possible events correspond to the coin being with heads up, or tails up. Outcomes, therefore, are the labels H and T associated with each of these events.

Like systems and tests, there is a trivial outcome set $\star \in \mathsf{Out}(\Theta)$, representing the singleton set $\star := \{*\}$. Any test containing $\star$ as the set of outcomes is therefore related to a single event and is called a *deterministic test*. When this is the case, we will represent them without the subscript, $\mathsf{T}_\star^{A\to B} \equiv \mathsf{T}^{A\to B}$. The trivial test $\mathsf{I}^{A\to A}$, for instance, is an example of a deterministic test.

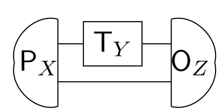
Another possible way of accommodating outcomes is by representing them diagrammatically as special system-types. In this convention, $T_X^{A\to B}$ is represented by

$\mathsf{T}_X^{A\to B} \quad \mapsto$  . (2.15)

Notice that this convention is perfectly compatible with the properties demanded from sequential and parallel compositions. We will mostly use the first notation introduced of outcomes as subscripts, but the notation with outcomes as wires will be convenient later when demonstrating some examples.

We also defined preparations and measurements as tests going from or to the trivial system I. A test going from *and* to the trivial system, i.e., $\mathsf{P}_X^{I\to I} \in \mathsf{Test}(\Theta)$ is a special element of the operational language and receives the name of *scalar*.

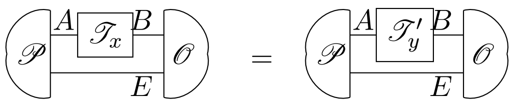
A *probabilistic model* consists of an operational language Π to which every scalar $P_X^{I\to I}$ is associated to a probability distribution $\{p(x)\}_{x\in X}$, i.e., every scalar is a function over the outcomes satisfying the Kolmogorov axioms introduced in the previous chapter. Most generally, they will have the form

$\{p(x,y,z)\}_{x\in X, y\in Y, z\in Z} =$  (2.16)

In fact, all these distributions are conditional to the tests P_X, T_Y, and O_Z being implemented as per the provided diagram. This association of scalars to probability distributions equips the operational language with all the features introduced in Chapter 1 and therefore results such as Bayes' rule can be rederived.

2.3 Quotient theories

Consider now two events $\mathscr{T}_x^{A\to B}$ and $\mathscr{T}_y'^{A\to B} \in \mathsf{Event}(\Pi)$, which are not necessarily equal or associated to the same outcome. However, it might be the case that



$$, \quad \forall \mathscr{P} \in \mathsf{Event}(I \to AE),\ \mathscr{O} \in \mathsf{Event}(BE \to I). \tag{2.17}$$

Whenever this happens, we say that $\mathsf{T}_x^{A\to B}$ *is probabilistically equivalent to* $\mathsf{T}_y'^{A\to B}$, and represent it as $\mathsf{T}_x^{A\to B} \sim \mathsf{T}_y'^{A\to B}$. It is possible to prove that $\sim$ indeed constitutes an equivalence relation.

We can then look not only to the set $\mathsf{Event}(\Pi)$ but to its partition $\mathsf{Event}(\Pi)/\sim$ in which every element is a subset of events probabilistically equivalent to each other. Such mapping is called *quotienting*. Notice that quotienting the events implies changes on the tests as well since they are just sets of events. In particular, some of these quotiented partitions receive special labels:

- The set $\mathsf{Test}(A \to B)/\sim$ of tests from non-trivial to non-trivial system-types is labeled as $\mathsf{Instr}(A \to B)$, the set of *instruments*;

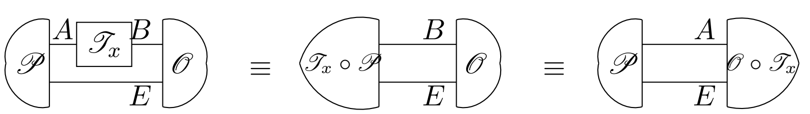
- The set $\mathsf{Event}(A \to B)/\sim$ from non-trivial to non-trivial system-types is labeled $\mathsf{Transf}(A \to B)$, the set of *transformations*;

- The set $\mathsf{Event}(I \to A)/\sim$ from the trivial to a non-trivial system-type is labeled $\mathsf{St}(A)$, the set of *states*;

- The set $\mathsf{Event}(A \to I)/\sim$ from a non-trivial to the trivial system-type is labeled $\mathsf{Eff}(A)$, the set of *effects*.

A *quotiented probabilistic model* $(\mathsf{Sys}(\Pi), \mathsf{Instr}(\Pi), \mathsf{Transf}(\Pi), \mathsf{St}(\Pi), \mathsf{Eff}(\Pi), \mathsf{Out}(\Pi))$ in which all scalars are combined through multiplication of real numbers constitutes an *operational probabilistic theory (OPT)*.

A very relevant feature of an OPT is that, since it assigns physical meaning to the scalars only (in the sense that they are the only directly observable element in the theory), every scalar boils down to a prepare-and-measure scenario, no matter how complex the underlying diagram might originally be. It means that, for example,



(2.18)

Conversely, prepare-and-measure scenarios are the only ones for which the OPT must assign a definite probability distribution. It doesn't mean that claims about probabilities of tests or events occurring cannot exist in an OPT, however, they do not represent an objective probability distribution, but merely a degree of belief of an agent about the presence of a particular element in the closed diagram.

2.4 Linear structure

We finally have tools to establish a linear structure to the OPT. Although this can be derived straightforwardly from the definition of OPTs and probability theory, we will just provide the theorem and explore its implications.

Theorem 2.4.1 (Linear structure for OPTs) *Let Θ be an OPT. Then* $\mathsf{Transf}(A \to B)$ *can be embedded into a real vector space* $\mathsf{Transf}_{\mathbb{R}}(A \to B)$*, such that the two operations* $+$ *(sum) and* $\cdot$ *(scalar multiplication) are well-defined and*

- $+$ *is distributive over parallel and sequential composition;*
- $\cdot$ *is compatible with parallel and sequential composition.*

What the theorem means is that, given a transformation $\mathscr{T}_i^{A\to B}$, $\mathscr{T}'^{B\to C}$, one has

$$\left(\sum_i q_i \overset{A}{\text{—}}\boxed{\mathscr{T}_i}\overset{B}{\text{—}}\right) \circ \overset{B}{\text{—}}\boxed{\mathscr{T}'}\overset{C}{\text{—}} = \sum_i q_i \overset{A}{\text{—}}\boxed{\mathscr{T}_i}\overset{B}{\text{—}}\boxed{\mathscr{T}'}\overset{C}{\text{—}} \; ; \tag{2.19}$$

$$\begin{matrix}\left(\sum_i q_i \overset{A}{\text{—}}\boxed{\mathscr{T}_i}\overset{B}{\text{—}}\right) \\ \overset{B}{\text{—}}\boxed{\mathscr{T}'}\overset{C}{\text{—}}\end{matrix} = \sum_i q_i \begin{matrix}\overset{A}{\text{—}}\boxed{\mathscr{T}_i}\overset{B}{\text{—}} \\ \overset{B}{\text{—}}\boxed{\mathscr{T}'}\overset{C}{\text{—}}\end{matrix} \; . \tag{2.20}$$

In particular, scalars and summations can "float" in the diagram, i.e., they can be moved in the diagram as one sees fit. They both have a clear interpretation with respect to the underlying operational language as well: summing is a *coarse-graining* of the underlying events so that the outcomes related to the summed-up events are being overlooked or ignored by the agent reasoning about the particular experiment. Scalar multiplication on its turn is a *randomisation*, in which the agent is just attributing to the particular diagram a chance of occurrence equal to the scalar it is being multiplied by.

Since we are embedding $\mathsf{Transf}(A \to B)$ into a real vector space, and ultimately any transformation can be absorbed by a state and an effect, it is very natural to read states and effects as vectors in this real vector space. In particular, states $\mathscr{P}_x \in \mathsf{St}(A)$ are mapped to vectors $|\rho_x) \in \mathbb{R}^m$, and effects $\mathscr{A}_y \in \mathsf{Eff}(A)$ are mapped to vectors $(a_y| \in \mathbb{R}^{m*}$ in the dual of $\mathbb{R}^m$. One can thus think of scalars as the inner product between states and effects[2],

[2] This comes from a representation theorem attesting that one can map vectors in the dual space of a vector space into the space itself, so the inner product is in fact taken between the states and the representations of the effects in the same vector space.

$p(x, y|\rho_x, a_y) = (a_y|\rho_x)$. Notice that coarse-graining over all possible outcomes implies

$$\sum_{y\in Y}(a_y|\rho_x) = \sum_{y\in Y} p(x, y|\rho_x, a_y) = p(x|\rho_x, a_y). \tag{2.21}$$

But as argued previously, the OPT cannot say anything about an event in its underlying diagrammatic explanation if there is no outcome associated with it. It is then natural to demand that the function above *does not depend on the choice of* a_y. This is called *causality*: demanding that the outcome statistics of a scalar test, when ignoring the outcomes of the observation, does not depend on the choice of measurement. It is very important to emphasize that not all OPTs will satisfy this property, and no OPT needs to satisfy this to admit a linear structure.

Nonetheless, admitting such an assumption does yield convenient properties to the OPT. For instance, it implies the existence of a unique unit effect $(u| \in \mathsf{Eff}_{\mathbb{R}}(A)$ for every system-type A, such that

$$(u|\rho) \leq 1, \quad \forall|\rho) \in \mathsf{St}_{\mathbb{R}}(\Pi). \tag{2.22}$$

It can be shown that the converse also holds: the existence of a unique unit effect for every system-type in an OPT implies that this OPT satisfies the causality assumption. This unit effect can be understood as taking the trace in standard quantum theory.

2.5 Examples

Classical bit

The classical bit consists of a subnormalised probability distribution $\{p(0), p(1)\}$. Let us assume this distribution is normalised, and call $p(0) = q$. States can straightforwardly be embedded into a 2-dimensional real vector space, such that any valid preparation has the form

$$|\rho) = \begin{pmatrix} q \\ 1-q \end{pmatrix}. \tag{2.23}$$

Effects in this case are all vectors $(a|$ such that $0 \leq (a|\rho) \leq 1$. We can represent them as per Figure 2.1. The unit effect is therefore simply $(u| = (1\,, 1)$, and the transformations are substochastic maps, taking states $|\rho) \in \mathsf{St}(\mathbb{R}^2)$ to $\mathsf{St}(\mathbb{R}^2)$.

Qubit

In quantum theory, the qubit is represented by a Hilbert space of dimension 2, with states of the form

$$\rho = \frac{1}{2}(Tr\{\rho\}\mathbb{1} + \langle X\rangle\,\sigma_X + \langle Y\rangle\,\sigma_Y + \langle Z\rangle\,\sigma_Z), \tag{2.24}$$

where $\mathbb{1}$ is the identity matrix, σ_X, σ_Y and σ_Z are the Pauli operators and $\langle X\rangle$, $\langle Y\rangle$ and $\langle Z\rangle$ are the expectation values of these operators.

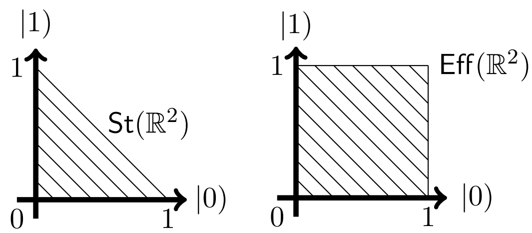


Figure 2.1: Real vector space representation for the sets of normalised states and effects associated with a classical bit distribution.

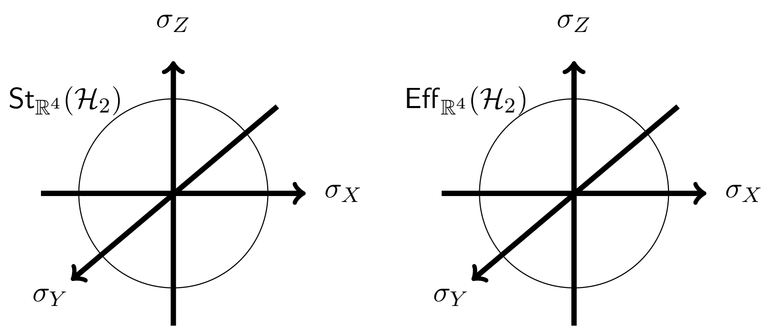


Figure 2.2: Real vector space representation for the sets of normalised states and effects associated with a quantum bit distribution.

Each density operator of a qubit already represents an equivalence class for preparations, since for instance the maximally mixed state will not distinguish whether the preparation employed was a convex mixture of pure Z states or X states, despite these being two different preparation procedures. All of these are represented by the same state, $\rho = \frac{1}{2}\mathbb{1}$.

We can then easily represent $|\rho)$ as vectors in $\mathbb{R}^4$:

$$|\rho) = \begin{pmatrix} Tr\{\rho\} \\ \langle X \rangle \\ \langle Y \rangle \\ \langle Z \rangle \end{pmatrix}. \tag{2.25}$$

The maximally mixed state, for instance, has the form

$$|\rho) = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{2.26}$$

The effects in the qubit are precisely the Bloch sphere again, so the dual of $\mathsf{St}_{\mathbb{R}^4}(\mathcal{H}_2)$

is itself. Effects are thus simply given by

$$(a| = (\ Tr\{a\} \quad \langle X\rangle' \quad \langle Y\rangle' \quad \langle Z\rangle'). \tag{2.27}$$

The unit effect is given by $(u| = (1\,,0\,,0\,,0)$. The standard graphical representation of the Bloch sphere is a 3-dimensional projection of the 4-dimensional hypersphere formed by the vectors $|\rho)$ satisfying $0 \leq (u|\rho) \leq 1$, and can be found in Figure 2.2. Transformations in the qubit are all the maps $T : \mathsf{St}_{\mathbb{R}^4}(\mathcal{H}_2) \rightarrow \mathsf{St}_{\mathbb{R}^4}(\mathcal{H}_2)$ that take valid states into valid states again, even when T is applied to one part of a bipartite system.

References

1. D. Frauchiger, *A Non-Probabilistic Framework for Scientific Theories.* PhD Thesis. ETH Zürich, Zürich (2016).

2. G. M. D'Ariano, G. Chiribella, P. Perinotti, *Quantum Theory from First Principles.* Cambridge University Press (2017).

3. G. Chiribella, G. M. D'Ariano, P. Perinotti, *Quantum from principles. In:*Quantum Theory: Informational Foundations and Foils", G. Chiribella and R. Spekkens eds., Springer (2016).

4. G. Chiribella, G. M. D'Ariano, P. Perinotti, *Probabilistic Theories with Purification.* Physical Review A 81, 062348 (2010).

5. G. M. D'Ariano, P. Perinotti, M. Erba, *Categorical Physics I – Review of a constructive framework for Operational Probabilistic Theories.* In preparation.

Chapter 3

Spekkens' Toy Theory

(...) Where can you find a place that will agree better with you and me? No schools, no teachers, no books! In that blessed place there is no such thing as study. Here, it is only on Saturdays that we have no school. In the Land of Toys, every day, except Sunday, is a Saturday. Vacation begins on the first of January and ends on the last day of December.(...)

— C. Collodi [1]

3.1 Ontic vs. epistemic – Spekkens' toy theory

In classical theory, a system is usually described by a point in the phase space in which each degree of freedom is associated with a canonical coordinate (for instance, the position and momentum (x, p) of a single particle). The possible states for this system to be in are given by the trajectory that satisfies the equations of motion, once the initial conditions are specified. If one is not certain about the initial conditions but rather assigns probabilities to each pair (x_0, p_0) of possible initial conditions in a region of the phase space, then the possible states for the system are all the trajectories associated with each of these initial conditions, weighted by the probability assigned to them.

This probability distribution of possible states in which your system might be in is what is called *epistemic state*. The name comes from the ancient Greek *epistḗmē*, knowledge, so epistemic states represent states of knowledge about a system's state. Naturally, if the probability distribution in question is a Dirac delta[1], then the agent is completely sure about which is the initial state of the system and can conclude which trajectory of the phase space the system will assume when it evolves with time. This state in which your system truly is is called *ontic state*, from the Greek *ón*, being, existing. Ontic states are therefore the fundamental states your system is occupying.

[1] A Dirac delta is a probability distribution that assigns nonzero probability to a single point and null probability to all others.

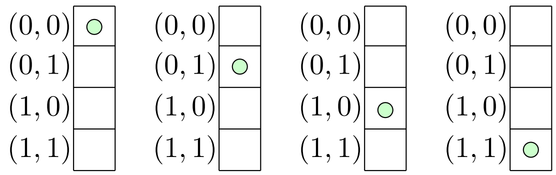


Figure 3.1: Ontic states in Spekkens' toy theory.

When we talk about quantum theory, the most widespread belief is that pure states are ontic, and mixed states are epistemic. This belief comes from the fact that as well as in classical theory, the evolution of a pure state is completely specified by initial conditions. Establishing a parallel with the previous chapter, an isolated qubit under the action of a constant magnetic field will completely specify a trajectory on the surface of the Bloch sphere, once the initial state $|\psi_0\rangle$ is given. Spekkens' toy theory is an attempt to defend the idea that all quantum states, both pure and mixed, are epistemic. In short, many of the apparently weird features of quantum theory, such as interference and entanglement, are totally mundane when quantum states are viewed as epistemic rather than ontic.

Consider a *toy bit*: a system with two degrees of freedom, (Q, P), such that to each of these physical quantities it can be assigned value 0 or 1. In this phase space, there will be 4 possible states for a system to be in: $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$. In Figure 3.1, you can see a representation of this phase space as a column. We also introduce the following rule in our theory:

> *An agent can only be certain about the value of one degree of freedom per system, being completely ignorant about the values of any other degree of freedom.*

This principle is known as *knowledge balance principle*, *epistemic restriction* or *principle of classical complementarity*, the last one being justified by its similarity with the quantum complementarity principle that acts over canonical pairs. What this principle is imposing is that an agent cannot know states of the form of Figure 3.1. Instead, when certain that $Q = 0$, an agent should not be able to tell whether the system is in state $(0, 0)$ or $(0, 1)$, and similarly for certainty about other values of Q or P. It means that the valid states are the ones given in Figure 3.2. There is a natural mapping from these valid epistemic states and *pure* quantum states of the qubit: the first two states represent the $|0\rangle$ and $|1\rangle$ of the Bloch sphere; the third and fourth states represent $|+\rangle$ and $|-\rangle$; the fifth and sixth represent $|+i\rangle$ and $|-i\rangle$. The last epistemic state is associated with a mixed state, in particular the maximally mixed one, $\mathbb{1}/2$.

Transformations in the toy theory must map valid epistemic states to valid epistemic states again, therefore no transformation can provide an agent with complete certainty about the ontic state of the system. But more than that, we demand that transformations

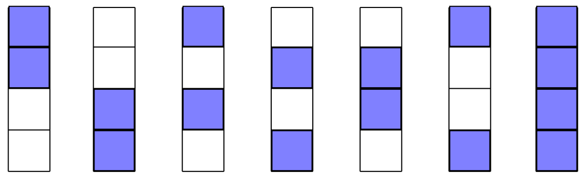
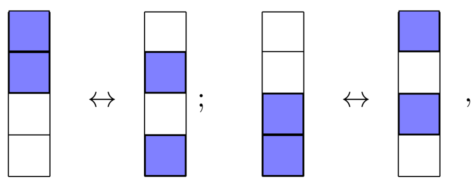


Figure 3.2: Valid epistemic states in Spekkens' toy theory

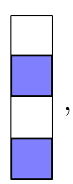
operate on the ontic level, i.e., they change epistemic states by acting directly on the ontic state. For instance, a transformation that swaps $(0, 0)$ and $(1, 1)$, while maintaining $(0, 1)$ and $(1, 0)$ invariant, will map epistemic states to valid epistemic states, since



(3.1)

leaving the other epistemic states unchanged. This definition will be particularly relevant when discussing the bipartite case later.

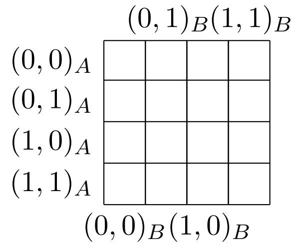
Valid measurements provide certainty about the value of a single degree of freedom, and "mix up" any information about the other one. For instance, measuring the value $Q = 0$ will just refresh any state of knowledge about the ontic state to the first epistemic state of Fig. 3.2. Measurements, however, are usually not reversible. Notice also that if for instance one measures $Q = 0$ on the state



(3.2)

one might conclude that the ontic state of the system *prior to the measurement* must have been $(0, 1)$. After the measurement however this might be no longer the case, due to the complementarity principle imposed over the epistemic states.

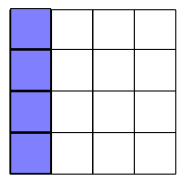
One can compose systems in the toy theory in the following manner: two systems will be in ontic states $(Q, P)_A$ and $(Q, P)_B$, respectively. The valid ontic states for the composite systems will be one out of the 16 entries of the following table



. (3.3)

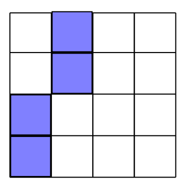
The complementarity principle tells us that one cannot have complete knowledge about a single system, which means that one cannot be certain about both Q_A and P_A. However, being certain about Q_A and P_B for instance does not posit any violation of the principle, yielding valid epistemic states. Some examples of valid epistemic states for two toy bits are given in Figure 3.3.

It is easier to understand what states are *not* valid for the case of two toy bits. For instance, the state



(3.4)

is not valid, since an agent would have maximal knowledge about the ontic state of system B. Another more intricate example is given by



. (3.5)

The problem with this state is that if one agent performs a measurement and learns that $Q_A = 1$ and $Q_B = 0$, which is allowed by the complementarity principle, the outcome of

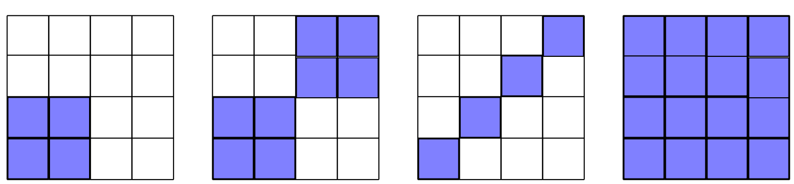
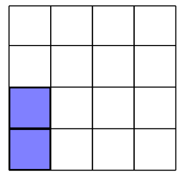


Figure 3.3: Some examples of valid epistemic states for two toy bits in the toy theory. Any permutations of rows and columns for these states are also valid in the theory.

the measurement could be given by



, (3.6)

which is not allowed since again the agent would know both Q_B and P_B. In fact, after some leveraging of possible states by considering this aspect of the principle, one can conclude that the only valid epistemic states are the ones in Figure 3.3 and permutations thereof.

3.2 "Quantum" features in the toy theory

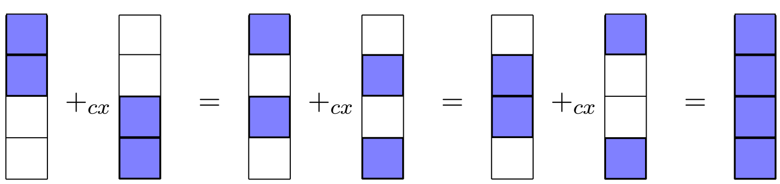
There are plenty of features displayed in the toy theory that have an immediate analogy with quantum behaviours. We briefly comment on some of them here.

Purity

Pure epistemic states are states of maximum knowledge. By the complementarity principle, this means that they are the states in which an agent is certain about exactly one outcome of the canonical pair for each system. By this definition, all but the last epistemic state in Figures 3.2 and 3.3 are pure states, and they are the only pure states in the theory.

Convex combination

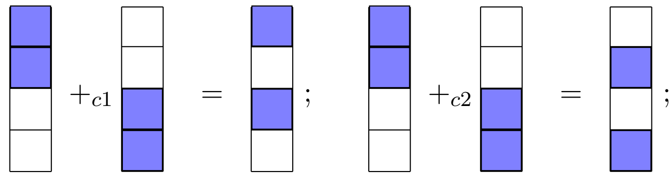
Like in quantum theory, it is possible to describe convex combinations of epistemic states in a very straightforward manner: the convex combination of two epistemic states consists of all possible ontic states inferred by both. For instance,



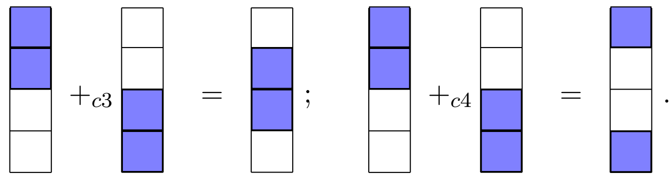
. (3.7)

Coherence

Coherence is the generalisation of the notion of superposition in quantum theory, i.e., the idea that two pure states can be combined into another pure state, as opposed to the convex combination that will always result in a mixed state (a state of non-maximum knowledge). There are in fact four of such operations in the toy theory. For the states corresponding to $Q = 0$ and $Q = 1$, for instance, they yield



(3.8)



(3.9)

The same can be defined for other combinations of pure states. The parallel with quantum theory comes by interpreting the above equalities as

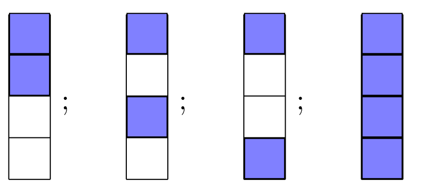
$$\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \;=\; |+\rangle\,; \qquad \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \;=\; |-\rangle\,; \tag{3.10}$$

$$\frac{1}{\sqrt{2}}(|0\rangle + i\,|1\rangle) \;=\; |+i\rangle\,; \qquad \frac{1}{\sqrt{2}}(|0\rangle - i\,|1\rangle) \;=\; |-i\rangle\,, \tag{3.11}$$

and so on for other pure states.

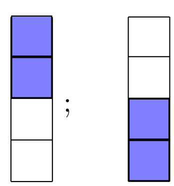
Non-orthogonality

We can say that two epistemic states are nonorthogonal if they share at least one ontic state among the possible ones in the state of knowledge they represent. For instance, the states



(3.12)

are all non-orthogonal, since all infer the state $(0, 0)$ as a possible ontic state. On the other hand, the states

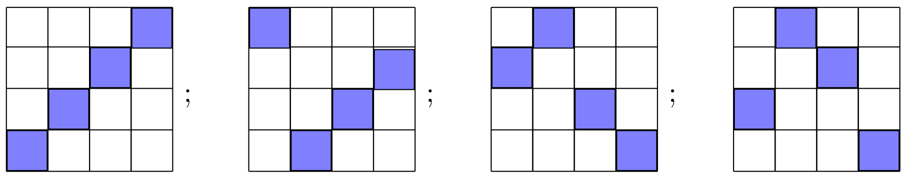


(3.13)

are said to be *orthogonal*, since the set of possible ontic states for both is disjoint. In classical theory, all pure states are orthogonal to each other, which does not hold in the toy theory.

Entanglement

The states



(3.14)

and their permutations of rows and columns are all valid and pure epistemic states in the theory. However, nothing is particularly known about the individual outcomes Q or P of each toy bit. Instead, the maximum knowledge is on how these quantities are related between the bits, such that once you find out the value of Q_A, for instance, you should be able to infer either Q_B or P_B with certainty. The first state, for example, tells that the ontic state of the bit A is always the same as the ontic state of the bit B, but it does not tell which ontic state is it. The others follow a similar reasoning.

Pure states for composite systems that tell nothing about the individual systems are called entangled states in quantum theory. In classical theories, having some maximal knowledge about a composite system always implies maximal knowledge about the individual systems as well, but the complementarity principle imposes a trade-off between local and global knowledge.

Geometrical structure

As introduced in the previous chapter, the toy theory can also be accommodated in the framework of OPTs and therefore embedded into a real vector space. Unlike both classical and quantum theory, its state of spaces $\mathsf{St}_{\mathbb{R}}(\texttt{ToyBit})$ is given by the one in Figure 3.4.

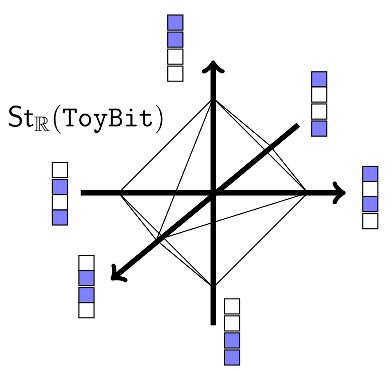
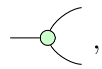


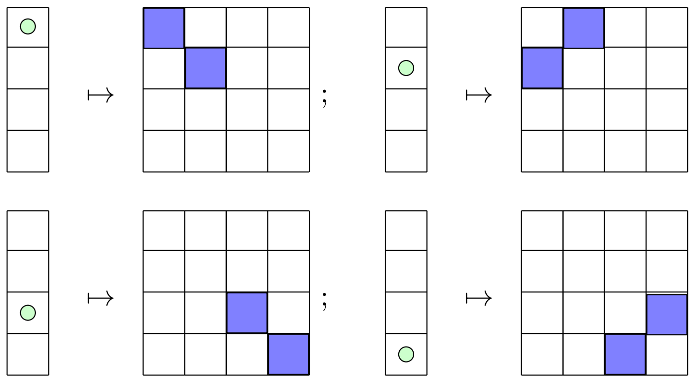
Figure 3.4: Set of states of the real vector space representation of the toy bit OPT.

3.3 The toy theory in ZX language

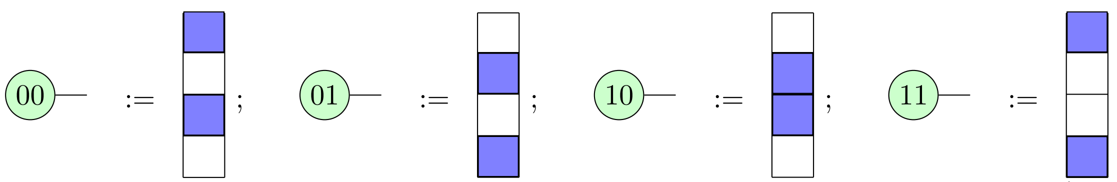
A very interesting feature of the toy theory is that it can easily be imported into the framework of ZX calculus. For this purpose, we merely need to associate some of its components with the graphical components of ZX. The first of such components will be the cloning spider:

 (3.15)

such that it maps ontic states to epistemic states in the following manner:

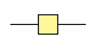
 (3.16)

Furthermore, we will establish the following convention: spiders with no input and one output will be always labeled with a phase 00, 01, 10, or 11, and will represent the following epistemic states:

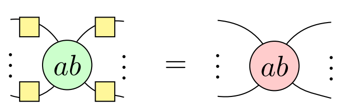


(3.17)

Finally, we introduce a Hadamard transformation

 (3.18)

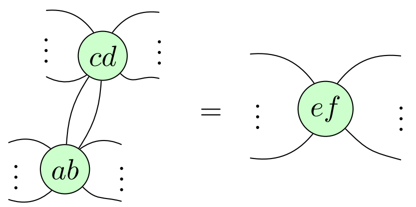
that will map ontic states to ontic states of single toy bits such that $(0,1) \leftrightarrow (1,0)$, while the other ontic states remain unaltered. We also establish that any spider with Hadamards applied to all its legs is written as a red spider, i.e.,



(3.19)

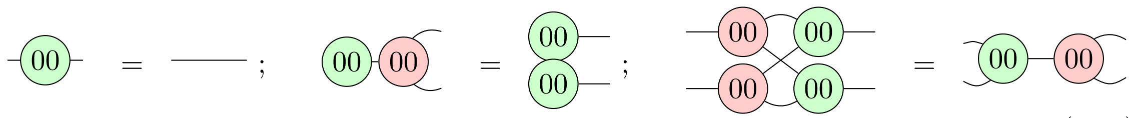
where a spider with multiple legs is just a concatenation of nested copying spiders and their respective adjoint spiders.

With only these tools, we can see how all properties of ZX calculus follow naturally. For instance, the spider rule will be such that



$$ef = (a \oplus c)(b \oplus d), \tag{3.20}$$

where $\oplus$ is the sum module 2. Finally, all bialgebra properties of ZX calculus hold for phase 00:



(3.21)

Finally, it is possible to demonstrate that the toy theory calculus is complete by employing the same completeness proofs as for standard ZX. This is a much more convenient result: now we can leverage any knowledge about ZX calculus to demonstrate things with the toy theory!

3.4 Example: the Peres-Mermin proof of non-classicality

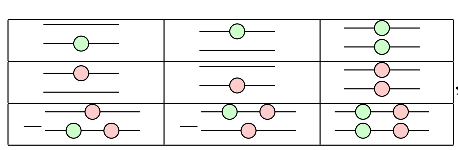
The Peres-Mermin square is a classic proof of non-classicality of quantum theory. It consists of a table of measurements over a two-qubit state, with the form

$\mathbb{1} \otimes \sigma_Z$	$\sigma_Z \otimes \mathbb{1}$	$\sigma_Z \otimes \sigma_Z$
$\sigma_X \otimes \mathbb{1}$	$\mathbb{1} \otimes \sigma_X$	$\sigma_X \otimes \sigma_X$
$-\sigma_X \otimes \sigma_Z$	$-\sigma_Z \otimes \sigma_X$	$\sigma_Y \otimes \sigma_Y$

. (3.22)

Each row of this matrix multiplies to $\mathbb{1} \otimes \mathbb{1}$, except for the last row that results in $-\mathbb{1} \otimes \mathbb{1}$. Then, we assume that each of these observables has a definite outcome ± 1 assigned to it, without giving attention to how these measurements will be implemented. This assumption is incompatible with quantum theory since it is simply impossible to replace the entries of the square by ± 1 in such a way that the quantum constraints are satisfied!

The Peres-Mermin square is considered a proof of *contextuality*, a signature of non-classicality that will be explored in the next chapter. Interestingly enough, Spekkens' toy theory cannot prove non-classicality in this case. That is because, by replacing the quantum measurements with the corresponding toy measurements, i.e.,



, (3.23)

where each node has phase 01. It is possible to check that every row and column of this square when concatenated sequentially, will yield merely two identities paired up, the last row inclusive. This means that all entries of the square can be assigned values ± 1 with no contradiction.

References

1. C. Collodi, *The Adventures of Pinocchio*. Translated by Carol Della Chiesa, 1892.

2. R. W. Spekkens, *Evidence for the epistemic view of quantum states: A toy theory*. Physical Review A 75, 032110 (2007).

3. M. Beckens & A. N. Duman, *A Complete Graphical Calculus for Spekkens' Toy Bit Theory*. Foundations of Physics 46, 70–103 (2016).

4. M. Beckens, *Completeness and the ZX-calculus*. PhD thesis, University of Oxford, Oxford (2016).

5. L. Hausmann, N. Nurgalieva, L. del Rio, *A consolidating review of Spekkens' toy theory*. arXiv:2105.03277 [quant-ph] (2021).

6. M. F. Pusey, *Stabilizer Notation for Spekkens' Toy Theory*. Foundations of Physics 42, 688–708 (2012).

7. M. Leiffer, *Epistemic Theories*. YouTube video. Lectures for the Solstice of Foundations 2022 — ETH Zürich Summer School.

Chapter 4

Signatures of Non-Classicality

> *To suppose two things indiscernible, is to suppose the same thing under two names. And therefore to suppose that the universe could have had at first another position of time and place, than that which it actually had; and yet that all the parts of the universe should have had the same situation among themselves, as that which they actually had; such a supposition, I say, is an impossible fiction.*

— G. W. F. von Leibniz [1]

4.1 Bell theorem & quantum violations

Last chapter, we learned that many quantum features are displayed by Spekkens' toy theory, a classical theory with an epistemic restriction. The last section introduced the Peres-Mermin square, an example of something that quantum theory predicts but the toy theory does not. This raises the question of what is truly non-classical in quantum theory, or put in other words, what does quantum theory can explain, but the toy theory cannot?

Bell nonlocality consists of a particular feature of operational theories that can accommodate a space-time structure. Consider the following Bell experiment: two agents, the ubiquitous Alice and Bob, are placed in separate laboratories. Each of them receives a system produced by the same source, chooses a measurement to perform over their share of the system, and registers the outcome observed. It is common to label $\mathbb{X}, \mathbb{Y}$ the sets of labels for the measurement choices of Alice and Bob, respectively, and $\mathbb{A}, \mathbb{B}$ the labels for the measurement outcomes. Without loss of generality, we will consider that every measurement choice has the same set of outcome labels, i.e., every measurement in $\mathbb{X}$ has the same set of outcomes $\mathbb{A}$, and the same for Bob. The setup of a Bell experiment is represented in Figure 4.1-a.

Finally, we assume that the laboratories are space-like separated. What it means is that their light cones have no intersections during the whole Bell experiment, from the moment

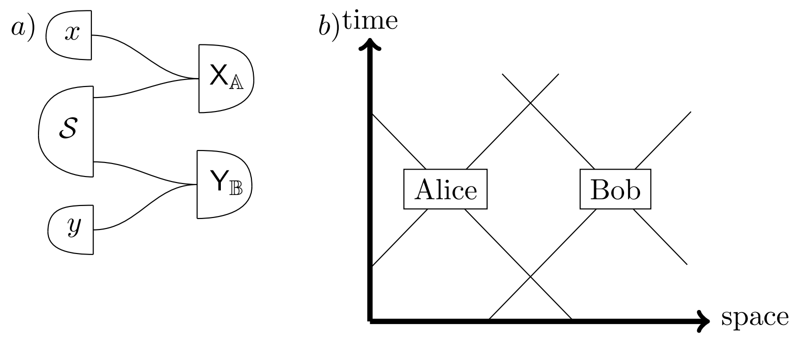


Figure 4.1: (a) OPT representation of a Bell experiment; (b) space-time diagram representing Alice and Bob's laboratories when the experiment happens. Notice that the diagonal lines, representing the trajectory of light, intersect only before or after the experiment is concluded.

they choose what measurements to perform and receive their halves of the system until they have performed the measurements and registered the outcomes. This is illustrated in Figure 4.1-b.

This operational setup produces a conditional probability distribution

$$\mathbb{P}_{\mathbb{AB}|\mathbb{XY}} := \{p(ab|xy)\}_{a\in\mathbb{A},b\in\mathbb{B},x\in\mathbb{X},y\in\mathbb{Y}}. \tag{4.1}$$

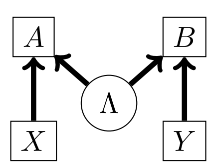
Furthermore, the constraint that they are space-like separated (which we will call *no-signaling condition*) implies that coarse-graining over the outcomes of a party implies ignoring also the measurement choice this party has made, i.e.,

$$\sum_{a\in\mathbb{A}} p(ab|xy) = p(b|y), \quad \forall b \in \mathbb{B}, x \in \mathbb{X}, y \in \mathbb{Y}; \tag{4.2}$$

$$\sum_{b\in\mathbb{B}} p(ab|xy) = p(a|x), \quad \forall a \in \mathbb{A}, x \in \mathbb{X}, y \in \mathbb{Y}. \tag{4.3}$$

This is similar to the *causality* condition defined for an OPT.

Now consider a causal structure given by the following graph



, (4.4)

where the square nodes represent observed random variables, and the circular node represents an unobserved or ignored variable. If we demand that the correlations in a Bell experiment are explained by this graph, in which all nodes are classical random variables, then they must necessarily have the form

$$p(ab|xy) = \sum_{\lambda\in\Lambda} p(\lambda)p(a|x\lambda)p(b|y\lambda), \quad \forall a \in \mathbb{A}, b \in \mathbb{B}, x \in \mathbb{X}, y \in \mathbb{Y}. \tag{4.5}$$

Bell's theorem demonstrates that there are operational theories, in particular quantum theory, satisfying the no-signaling condition and that yet cannot be explained by the above classical, causal structure (which is often called *locally causal*). It is easier to understand this result by looking at a particular case, where $\mathbb{X} = \mathbb{Y} = \mathbb{A} = \mathbb{B} = \{0, 1\}$. Mathematically, each experiment with fixed measurements $x, y \in \{0, 1\}$ is exactly the same as the example of tossing two coins simultaneously, provided in Chapter 1, where we assign values 0 to H and 1 to T. We can then compute the correlators

$$E_{xy} = p(00|xy) - p(01|xy) - p(10|xy) + p(11|xy), \quad \forall x, y \in \{0, 1\}. \tag{4.6}$$

If the probabilities $p(ab|xy)$ satisfy local causality, then there is Λ such that they have the form of Equation 4.5 and the correlators can be rewritten as

$$\begin{aligned} E_{xy} &= \sum_{a,b=0}^{1} (-1)^{a+b} p(ab|xy) &(4.7)\\ &= \sum_{a,b=0}^{1} (-1)^{a+b} \sum_{\lambda\in\Lambda} p(\lambda)p(a|x\lambda)p(b|y\lambda) &(4.8)\\ &= \sum_{\lambda\in\Lambda} p(\lambda) \left(\sum_{a=0}^{1}(-1)^a p(a|x\lambda)\right) \left(\sum_{b=0}^{1}(-1)^b p(b|y\lambda)\right) &(4.9)\\ &= \sum_{\lambda\in\Lambda} p(\lambda)E_x(\lambda)E_y(\lambda), &(4.10) \end{aligned}$$

where $E_x(\lambda)$ and $E_y(\lambda)$ are simply the expectation values of the individual coin tosses, parametrised by the choice of λ and x, y.

Consider now the Clauser-Horne-Shimony-Holt (CHSH) functional for this particular scenario

$$I_{CHSH}(\mathbb{P}_{\mathbb{AB}|\mathbb{XY}}) = E_{00} + E_{01} + E_{10} - E_{11}. \tag{4.11}$$

If each correlator factorises, it is possible to demonstrate that

$$|E_{00} + E_{01} + E_{10} - E_{11}| \leq 2, \tag{4.12}$$

which is the famous *CHSH inequality*.

Consider now the experiment in which Alice and Bob, each in their separate lab, receive a qubit. This pair of qubits was prepared in the state

$$|\Phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A |1\rangle_B + |1\rangle_A |0\rangle_B). \tag{4.13}$$

If Alice chooses $x = 0$, it means that she is going to perform measurement Z on her qubit, registering $a = 0$ if she gets outcome 0 and $a = 1$ otherwise. If she chooses $x = 1$, then she performs measurement X over the qubit registering outcomes in a similar way. Bob measures $\frac{1}{\sqrt{2}}(X + Z)$ when he chooses $y = 0$, and $\frac{1}{\sqrt{2}}(X - Z)$ when $y = 1$, registering outcomes like Alice.

By computing the correlators $E_{xy} = \langle \Phi | A_x \otimes B_y | \Phi \rangle$, one can get to the value

$$|E_{00} + E_{01} + E_{10} - E_{11}| = 2\sqrt{2} > 2. \tag{4.14}$$

Because assuming that the correlations satisfy local causality implies inequality 4.12, violating the inequality implies that the correlations do not admit of a locally causal explanation. Quantum theory is therefore deemed as *non-local* in the sense of Bell.

4.2 Non-signalling correlations and postquantum violations

The violation obtained in the previous example is in fact the best violation one can get with quantum theory for the CHSH inequality. However, if one considers all possible correlations $\mathbb{P}_{\mathbb{AB}|\mathbb{XY}}$ satisfying the no-signaling condition for the Bell scenario we investigated, one can conclude that

$$|E_{00} + E_{01} + E_{10} - E_{11}| \leq \sum_{x,y=0}^{1} |E_{xy}| \tag{4.15}$$

$$\leq 4. \tag{4.16}$$

This means that the optimal quantum violation, which we will refer to as the *Tsirelson bound*, is not tight: other probabilistic models out there are compatible with relativistic assumptions, and yet cannot be explained by quantum theory. This is illustrated in Figure 4.2.

Exploring these *postquantum theories* is more than a creative exercise. It has proven to be an invaluable tool for exploring what features of quantum theory are inherently quantum, in the sense that they cannot be explained by classical theory while at the same time ruling out other possible posquantum theories.

4.3 Generalised contextuality

A more general notion of non-classicality that is not present in the toy model is *generalised contextuality*, or simply contextuality from now on. Differently from Bell nonlocality, assessments of contextuality do not demand a space-time structure and thus can be present in many more operational theories. The object of study of contextuality is often a *prepare-and-measure scenario*, something that was briefly discussed in Chapter 2. Because any

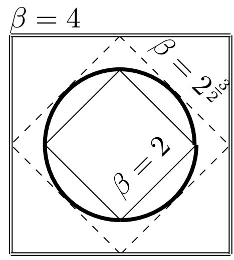


Figure 4.2: Representation of classical (normal line), Tsirelson (dashed line), and non-signaling (doubled line) bounds for the 2-measurements-2-outcomes Bell scenario. Quantum correlations (thick circle) never violate the Tsirelson bound, but other non-signaling correlations can.

operational scenario yielding probabilities can be reduced to a prepare-and-measure scenario, contextuality can also be assessed in a vast range of experiments.

To study prepare-and-measure scenarios, one must only know the preparations, observations, outcomes, and statistics in the operational language. We will call $\mathcal{P}$ the set of preparations, $\mathcal{M}$ the set of measurements, K the set of outcomes, and p the shortcut for $\{p(k|M,P)\}_{k\in K, M\in\mathcal{M}, P\in\mathcal{P}}$ the set of correlations derived from this scenario. As before, we can then represent the operational scenario by the tuple $(\mathcal{P}, \mathcal{M}, K, p)$.

As introduced in chapter 2, we want to quotient this operational language, so that any information that cannot be captured by preparing and measuring is ignored. We will keep using the symbol $\sim$ to tell that two preparations or measurement outcomes are operationally equivalent.

As in the Bell scenario, we want to supplement this operational scenario with some information about the causal structure of the experiment, a guess of what is happening to the system in the process of preparing and measuring it. We will call this extra information an *ontological model*. In the spirit of the toy theory, an ontological model will be composed of a space Λ containing all possible *ontic states* λ for the system. Each preparation will induce a state of knowledge about the system, i.e., an *epistemic state* $\{\mu(\lambda|P)\}_{P\in\mathcal{P},\lambda\in\Lambda}$, and each measurement is associated to a *response function* mapping ontic states to probabilities associated to each outcome, $\{\xi(k|M,\lambda)\}_{k\in K, M\in\mathcal{M},\lambda\in\Lambda}$. Finally, we want that this ontological model explains all statistics in the operational scenario,

$$p(k|M,P) = \int_{\lambda\in\Lambda} \xi(k|M,\lambda)\mu(\lambda|P)d\lambda, \quad \forall k\in K, M\in\mathcal{M}, P\in\mathcal{P}. \tag{4.17}$$

However, if there is information about preparations and measurement outcomes that cannot be captured by our experiment, there is no reason to include them in the ontological model. This is motivated by Leibniz's quotation at the beginning of this chapter: if two elements of a theory are indistinguishable, an explanation for that theory that needs to distinguish them is not the best explanation possible. What this means for our ontological model is that whenever two preparation procedures P, P' are equivalent, the epistemic states induced by them should be equal, and similar for measurement outcomes. This assumption is what we call *noncontextuality*:

$$P \sim P' \Rightarrow \mu(\lambda|P) = \mu(\lambda|P'), \quad \forall \lambda \in \Lambda,\ P, P' \in \mathcal{P}; \tag{4.18}$$

$$[k|M] \sim [k'|M'] \Rightarrow \xi(k|M, \lambda) = \xi(k'|M', \lambda), \quad \forall \lambda \in \Lambda\, k, k' \in K,\ M, M' \in \mathcal{M}. \tag{4.19}$$

It is the incompatibility with such an assumption that constitutes proof of contextuality. A theory is, therefore, *contextual* whenever it contains a prepare-and-measure scenario incompatible with the assumption of noncontextuality.

The reason why the toy model does not exhibit contextuality is that it is a noncontextual ontological model itself. It satisfies the principle of noncontextuality by construction, for instance, when the convex combination is defined by mapping three different combinations of epistemic states to the same maximally mixed state.

In fact, the geometric representation of states and effects in theories such as the toy model or the classical bit always ends up in *simplices*, i.e., generalisations of a triangle in multiple dimensions. Therefore, simplicial theories are often called *strictly classical.* In this sense, it has been shown that assessing contextuality for an operational theory is equivalent to assessing whether its geometric representation in the real vector space admits of a *simplex embedding*, i.e., a linear mapping from its states and effects into states and effects of a simplicial theory.

Formalising this concept, let $\Theta = (\mathsf{St}_{\mathbb{R}^m}, \mathsf{Eff}_{\mathbb{R}^m}, \mathsf{Sys})$ be the geometric representation of a quotiented operational theory in a real vector space. We say that Θ admits of a simplex embedding if there exists Δ_d a simplex in a (not necessarily the same) real vector space $\mathbb{R}^n$, and linear maps $\iota, \kappa : \mathbb{R}^m \to \mathbb{R}^n$ such that

- $\iota(\mathsf{St}_{\mathbb{R}^m}) \subseteq \Delta_d$;
- $\kappa(\mathsf{Eff}_{\mathbb{R}^m}) \subseteq \Delta_d^*$;
- The inner products are preserved by ι and κ.

This means that the quotiented operational theory Θ is a *subtheory* of the simplicial one, and therefore its statistics can be simulated by a strictly classical theory. This shows the strength of contextuality as a notion of non-classicality: every scenario that admits of a noncontextual ontological model is *classically explainable* in the sense that the whole setup can be simulated by a strictly classical quotiented theory. Quantum theory is not such a theory: the sets of states and effects of a qubit, and even small subsets of it, do not admit of a simplex embedding.

To prove it, consider the following six preparations:

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad \sigma_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}; \quad \sigma_2 = \frac{1}{4}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{pmatrix}; \tag{4.20}$$

$$\sigma_3 = \frac{1}{4}\begin{pmatrix} 3 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}; \quad \sigma_4 = \frac{1}{4}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{pmatrix}; \quad \sigma_5 = \frac{1}{4}\begin{pmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}, \tag{4.21}$$

with measurement outcomes being represented by the same matrices. Among the statistics for this experiment, we have that

$$p(1|0) = p(0|1) = Tr\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\} = 0; \tag{4.22}$$

$$p(3|2) = p(2|3) = Tr\left\{\frac{1}{4}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{pmatrix}\frac{1}{4}\begin{pmatrix} 3 & -\sqrt{3} \\ -\sqrt{3} & 1 \end{pmatrix}\right\} = 0; \tag{4.23}$$

$$p(5|4) = p(4|5) = Tr\left\{\frac{1}{4}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{pmatrix}\frac{1}{4}\begin{pmatrix} 3 & \sqrt{3} \\ \sqrt{3} & 1 \end{pmatrix}\right\} = 0, \tag{4.24}$$

and also that

$$\begin{aligned} \frac{1}{2}\mathbb{1} &= \frac{1}{2}(\sigma_0 + \sigma_1) && (4.25) \\ &= \frac{1}{2}(\sigma_2 + \sigma_3) && (4.26) \\ &= \frac{1}{2}(\sigma_4 + \sigma_5) && (4.27) \\ &= \frac{1}{3}(\sigma_0 + \sigma_2 + \sigma 4) && (4.28) \\ &= \frac{1}{3}(\sigma_1 + \sigma_3 + \sigma_5), && (4.29) \end{aligned}$$

and the same for the respective measurement outcomes.

Equations 4.25 will impose constraints on how the $\mu(\lambda|P)$ from the ontological model relate to each other, as well as the $\xi(k|\lambda)$. The statistics will impose constraints on how the epistemic states and response functions relate to each other, resulting in a long but simple system of equations that will only admit the trivial solution: all $\mu(\lambda|P)$ and $\xi(k|\lambda)$ must be null for all values of λ, which means that there is no ontological model compatible with this scenario.

References

1. G. W. F. von Leibniz & S. Clarke, *The Leibniz-Clarke Correspondence: Together Wiith Extracts from Newton's Principia and Opticks.* Manchester University Press, Manchester (1998).

2. V. Scarani, *Bell Nonlocality.* Oxford University Press, Oxford (2019).

3. A. B. Sainz, *Signatures of Non-Classicality.* Lecture notes for the Quantum Information Technology Masters program of the University of Gdańsk, Gdańsk (2022).

4. J. S. Bell, *On the Einstein-Podolsky-Rosen Paradox.* Physics 1(3), 195-290 (1964).

5. J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, *Proposed Experiment to Test Local Hidden-Variable Theories.* Physical Review Letters 23, 880 (1969).

6. R. W. Spekkens, *Contextuality for preparations, transformations, and unsharp measurements.* Physical Review A 71, 052108 (2005).

7. R. W. Spekkens, *The ontological identity of empirical indiscernibles: Leibniz's methodological principle and its significance in the work of Einstein.* arXiv:1909.04628 [physics.hist-ph] (2019).

8. D. Schmid. *Noncontextuality.* YouTube video. Lectures for the Solstice of Foundations 2022 — ETH Zürich Summer School.

Chapter 5

Resource Theories

> *The guiding philosophy in the pragmatic tradition is that understanding a phenomenon means being able to make use of it. Physical phenomena are studied in order to better leverage certain resources.*

—B. Coecke, T. Fritz, R. W. Spekkens [1]

5.1 Mathematical framework for resource theories

The concept of resource is borrowed from the economical concept of *scarcity*. A certain state of things is more or less valuable according to the easiness with which it can be extracted, obtained, or implemented. It is convenient thus to consider a theory that can keep track of how valuable a given state or experimental protocol is under certain physical restrictions. Structurally, a resource theory is a *commutative ordered monoid*. This can be defined as follows:

Definition 5.1.1 *(Commutative ordered monoids) Let $\mathcal{A}$ be a set of resources equipped with a binary operation $+$ (representing the situation in which one holds two resources simultaneously), a distinguish element e (a free resource), called identity, and an ordering relation $\geq$ (telling how valuable a resource is compared to another). Then $\mathcal{A}$ is said to be a commutative ordered monoid if, for any $a, a', a'' \in \mathcal{A}$, we have*

- *if $a \geq a'$ and $a' \geq a''$, then $a \geq a''$;*
- *if $a \geq a'$ and $a' \geq a$, then $a = a'$;*
- *$a + (a' + a'') = (a + a') + a''$ and $a + a' = a' + a$;*
- *$a + e = a$;*
- *if $a \geq a'$, then $a + a'' \geq a' + a''$.*

In physical terms, one should read the symbol $\geq$ as "*is convertible to*". Whenever $\rho \geq \sigma$, it means that, in the set of operations and descriptions allowed by the restriction imposed over the experimental scenario, there will be ways of transforming ρ into σ. If, otherwise, $\rho \not\geq \sigma$, it must be read as "*ρ is not convertible to σ*", meaning that there is no way of performing this transformation with the knowledge one has access to in the experimental scenario. Notice that not all resource theories need to satisfy the second property, i.e., there can be resource theories in which $(a \geq b) \wedge (b \geq a) \;\not\Longrightarrow\; a = b$. We call such structures *preordered monoids*, and they represent situations in which resources can be converted into one another albeit being different.

It is thus convenient to define a resource theory in terms of free operations and free states:

Definition 5.1.2 *(Resource theory) Let* $(\mathsf{Sys}(\Theta), \mathsf{St}(\Theta), \mathsf{Transf}(\Theta), \mathsf{Eff}(\Theta), \mathsf{Out}(\Theta))$ *be a quotiented operational theory. Let* $\mathcal{F} \subseteq \mathsf{St}(\Theta)$ *be a subset of states that* are not resourceful*, and* $\mathcal{O}(A \to B) \subseteq \mathsf{Transf}(A \to B)$*, for all* $A, B \in \mathsf{Sys}(\Theta)$ *be a subset of transformations that* cannot create resources*. Then, the tuple* $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ *is a resource theory if*

- $\mathsf{I} \in \mathcal{O}(A)$*,* $\forall A \in \mathsf{Sys}(\Theta)$*, where* I *is the trivial transformation;*
- $\mathsf{T} \in \mathcal{O}(A \to B)$ *and* $\mathsf{T}' \in \mathcal{O}(B \to C) \implies \mathsf{T}' \circ \mathsf{T} \in \mathcal{O}(A \to C)$*,* $\forall A, B, C \in \mathsf{Sys}(\Theta)$*.*

The set $\mathcal{F}$ is called the set of *free states*, while the set $\mathcal{O}$ is called the set of *free operations*. The demands for constructing a resource theory are perfectly reasonable: to do nothing is always a free operation, and a sequence of free operations must be free as well. A corollary of this definition is often highlighted due to its interpretational convenience:

Definition 5.1.3 *(Golden rule of resource theories) Let* $\mathcal{R} = (\mathcal{F}, \mathcal{O})$ *be a resource theory for the operational theory* Θ*. If* $\mathsf{T} \in \mathcal{O}(A \to B)$ *and* $\rho \in \mathcal{F}(A)$*, then* $\mathsf{T} \circ \rho \in \mathcal{F}(B)$*.*

The interpretation is as simple as it seems: performing a free operation over a free state will necessarily lead to a free state. In other words, free operations cannot convert free states into resource states.

Resource theories are always constructed operationally, taking into account what are the central phenomena to be studied to define the set of free operations or of free states (usually, one starts by defining just one of the sets, and the other is obtained from the Golden Rule). The quantum resource theory of entanglement, for example, starts from the assumption that for bipartite scenarios, operations performed locally over each of the parties and classical communication between them are always allowed. The set of free states, as a consequence, is restricted to separable states.

One of the most useful features of a resource theory is the possibility of identifying *monotones* associated with the resource property. Monotones are functions capable of witnessing or quantifying the convertibility between states and are mathematically described as homomorphisms between the resource theory $\mathcal{A}$ and $\mathbb{R}_{\geq 0}$, where $\mathbb{R}_{\geq 0}$ is the set of non-negative real numbers.

Definition 5.1.4 *(Homomorphism) Let $\mathcal{A}$ and $\mathcal{A}'$ be commutative ordered monoids. An ordered map $f : \mathcal{A} \to \mathcal{A}'$ is an homomorphism if, for every $a, a' \in \mathcal{A}$, we have*

- $a \geq a' \Rightarrow f(a) \geq f(a')$;
- $f(a + a') = f(a) + f(a')$;
- $f(0) = 0$.

Since $\mathbb{R}_{\geq 0}$ is also a commutative ordered monoid with respect to addition, we want to search for a functional f capable of quantifying the convertibility between elements of $\mathcal{A}$ with real numbers.

Catalytic convertibility is a common concept in many resource theories and is borrowed from the concept of catalysis in Chemistry. Mathematically, a resource theory equipped with catalytic convertibility can be defined by a *non-cancelative* commutative ordered monoid.

Definition 5.1.5 *(Non-cancelative commutative ordered monoid) Let $x, y, z \in \mathcal{A}$ be elements of a commutative ordered monoid. $\mathcal{A}$ is said to be non-cancelative if*

$$x + z \geq y + z \;\not\Longrightarrow\; x \geq y. \tag{5.1}$$

In resource-theoretic terms, it means that x is not convertible in y by itself, but in the presence of z, this process is allowed. The following example, due to Fritz, illustrates the idea: the conversion of *wood+nails* to *table* is not allowed, but the conversion *wood+nails+hammer* to *table+hammer* is possible. The state z is called the *catalyst* of this conversion. A resource theory that is non-cancellative can be turned into a cancellative one by redefining its ordering relation, such that for any $x, y, z \in \mathcal{A}$,

$$x + z \geq y + z \implies x \succeq y. \tag{5.2}$$

This relation can be read as "*x is catalytic convertible into y*", and a resource theory which is cancellative becomes an abelian ordered group[1]. Resource theories of this type are the ones that allow for borrowing resources, since the concept of a debt resource $-x$ is included in an abelian ordered group.

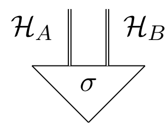
5.2 Example: Local Operations and Shared Randomness (LOSR)

We are interested in constructing a resource theory to quantify the non-classicality of common-cause scenarios. Common-cause scenarios consist of two parties, Alice and Bob,

[1] i.e., a commutative ordered monoid that has, for every $a \in \mathcal{A}$, an element $a^{-1} \in \mathcal{A}$ such that $aa^{-1} = a^{-1}a = e$.

sitting in separate labs, who have no direct way to influence one another. They might, however, have some systems in their lab which could have been interacting with one another at some point in the past when Alice and Bob met up with one another. It is these systems that are the common cause that can lead, for example, to correlations between what they observe in their labs. With this in mind, given that we are trying to understand non-classicality, it is natural to divide the things that Alice and Bob can do into free and nonfree by saying that the transformations they can do freely are those that rely only on a classical common cause — that is, some shared randomness — and the things that they can do non-freely are those that rely on a quantum common cause — that is, some shared entangled state.

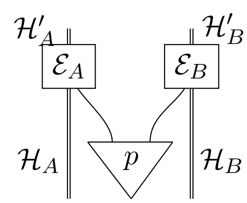
Consider the specific example of a resource state: a bipartite quantum state rho shared by Alice and Bob



. (5.3)

In this chapter, we will read diagrams from bottom to top for a change. Single wires will always represent classical systems, i.e., sets of random variables $\mathbb{A}, \mathbb{X}$, etc, while double wires will represent quantum systems (Hilbert spaces).

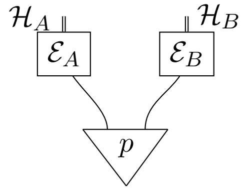
In this resource theory, we demand that they can freely convert σ into any other bipartite state ρ by performing local operations, i.e., complete-positive trace-preserving (CPTP) maps $\mathcal{E}^A$, $\mathcal{E}^B$ on their shares of the system, and by sharing some source of classical randomness $\{p(i)\}_{i\in I}$. These processes take the form



. (5.4)

It is easy to show that these operations will satisfy the transitivity and reflexibility of the order relation $\geq$, i.e., $(\rho \geq \sigma) \wedge (\sigma \geq \chi) \implies \rho \geq \xi$ and $\rho \geq \rho$. For the first one, it suffices to show that sequentially composing two of the above processes is again a free operation. The fact that the trivial process $\mathbb{1} \otimes \mathbb{1}$ belongs to the set of free operations proves the second property.

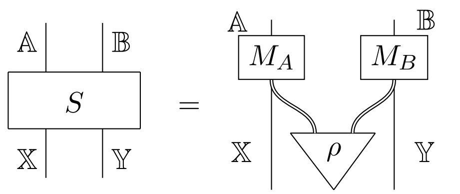
In particular, the type of states that can be created freely are separable states. That is because if you take maps $\mathcal{E}^A : \star \to \mathcal{H}_A$ and $\mathcal{E}^B : \star \to \mathcal{H}_B$, these are simply preparation procedures of quantum states, and



(5.5)

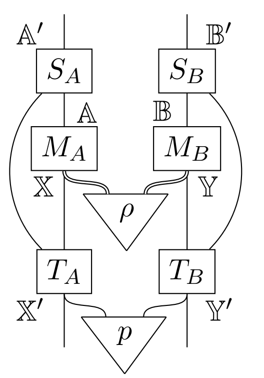
has the form of a free operation. Also, discarding is always a free operation. These two facts put together mean that *any resource can be converted to a separable state*, since one can always freely discard whatever resource is available and then freely create a separable state.

Notice however that this is one particular example of a resource. We can think of more general cases, such as bipartite stochastic maps. In our resource theory, we might want to say that they are free when they admit of a *quantum common-cause explanation*, in the spirit of a quantum Bell scenario, i.e.,



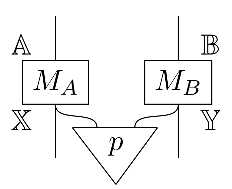
(5.6)

We can then ask ourselves what it means to perform local operations and have shared randomness in this scenario. These will be local stochastic maps acting on the inputs $\mathbb{X}$ and $\mathbb{Y}$ and outputs $\mathbb{A}$ and $\mathbb{B}$, such that



(5.7)

It turns out that for this set of free operations, *free resources are the classical-common cause Bell scenarios*, i.e.,

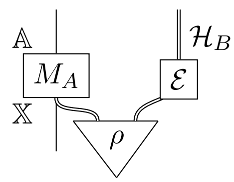


(5.8)

So by changing the notion of what is a free resource, we can use LOSR operations to quantify when something does not admit of a Bell classical common-cause explanation! In fact, this framework can be applied to virtually any process with classical, quantum, or even more general inputs/outputs. Take for instance the case with quantum Einstein-Podolsky-Rosen scenarios. They consist of a common cause for both Alice and Bob, with the difference that now Bob always receives a quantum system and never performs any measurement over it. The relevant objects in this scenario are called *assemblages*: sets of subnormalised quantum states labeled by Alices inputs and outputs,

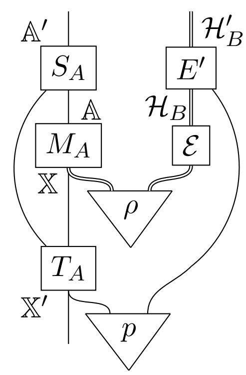
$$\Sigma_{\mathbb{A}|\mathbb{X}} := \{\sigma_{a|x}\}_{a\in\mathbb{A},x\in\mathbb{X}}. \tag{5.9}$$

If we consider the resources to be quantumly realisable assemblages, i.e., each element of the assemblage has the form



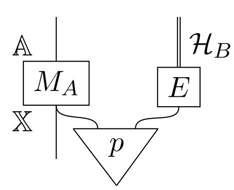
, (5.10)

then the LOSR operations will have the form



, (5.11)

i.e., some local operations on the inputs and outputs of Alice and quantum channels on Bob's state, all conditioned to a shared probability distribution. It is possible to verify that the free resources, in this case, are the *classical common-cause assemblages*, i.e.,



. (5.12)

We see therefore how this type-independent resource theory of LOSR allows us to quantify a myriad of interesting non-classicality scenarios without having to build a whole resource theory from scratch. One only needs to specify what processes in the theory are the resourceful ones, have a well-established notion of local operations and shared randomness for them, and identify when are the resources freely achievable. This framework can explore other common-cause scenarios beyond quantum theory — one just has to add postquantum system types and specify how LOSR processes will look for the sorts of scenarios under investigation.

References

1. B. Coecke, T. Fritz, R. W. Spekkens, *A mathematical theory of resources.* Information and Computation 250, 59-86 (2016).

2. T. Fritz, *Resource convertibility and ordered commutative monoids.* Mathematical Structures in Computer Science 27(6), 850–938 (2017).

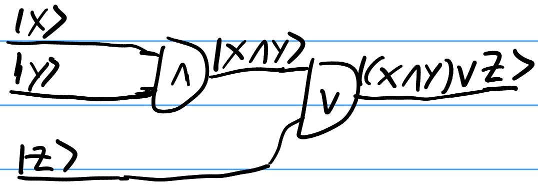
3. G. Gour & R. W. Spekkens, *The resource theory of quantum reference frames: manipulations and monotones.* New Journal of Physics 10, 033023 (2008).

4. D. Schmid, D. Rosset, F. Bruscemi, *The type-independent resource theory of local operations and shared randomness.* Quantum 4, 262 (2020)

5. J. H. Selby, *Signatures of Non-Classicality.* Tutorials for the Quantum Information Technology Masters program of the University of Gdańsk, Gdańsk (2022).

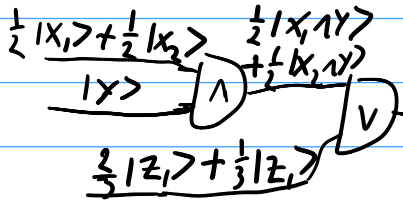
From Probabilities to Quantum Theory

Classical Logic



$x,y,z \in \{0,1\}$

Probabilistic Logic



$$\frac{2}{6}|(x_1\wedge y)\vee z_1\rangle + \frac{1}{6}|(x_1\wedge y)\vee z_2\rangle + \frac{2}{6}|(x_2\wedge y)\vee z_1\rangle + \frac{1}{6}|(x_2\wedge y)\vee z_2\rangle$$

Quantum Theory

Replace Prob. in $[0,1]$ by "Amplitudes" in $\mathbb{C}$

Instead of Bits $\{0,1\}$ or Prob. Bits $\{p|0\rangle+(1-p)|1\rangle \mid p\in[0,1]\}$

Have Quantum bits $\{a|0\rangle+b|1\rangle \mid a,b\in\mathbb{C}\}$ qubits

Example: $|0\rangle+|1\rangle$, $|0\rangle-|1\rangle$, $\frac{1}{3}i|0\rangle+e^{i\frac{2\pi}{3}}|1\rangle$

Note: $|0\rangle := \begin{pmatrix}1\\0\end{pmatrix}$
$|1\rangle := \begin{pmatrix}0\\1\end{pmatrix}$

Other way to think of it:

Go from $\{0,1\}$ to $\mathbb{C}^{\{0,1\}} := \{f:\{0,1\}\to\mathbb{C}\}$

For multiple bits:

Go from $\{0,1\}^n$ to $\mathbb{C}^{\{0,1\}^n} \cong \mathbb{C}^{2^n}$

Specifiying a classical state in $\{0,1\}^n$ we need n bits: 011 $(n=3)$

To specify a Q. State in $\mathbb{C}^{2^n}$ need 2^n numbers

$$a_0|000\rangle + a_1|001\rangle + a_2|010\rangle + a_3|011\rangle + a_4|100\rangle + \dots + a_8|111\rangle \qquad (n=3,\ 2^n=8)$$

Normalisation

Just like how Prob. Distributions are normalised:

For $\sum_\lambda p_\lambda |\lambda\rangle \Rightarrow \sum_\lambda p_\lambda = 1$

So Q. states have normalisation:

For $\vec{v}, \vec{w} \in \mathbb{C}^k$ $\quad \vec{v} = (v_1, v_2, \dots, v_k)$

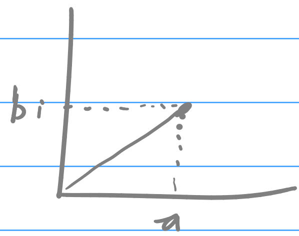
$\vec{w} = (w_1, w_2, \dots, w_k)$

$\langle \vec{v}, \vec{w} \rangle = \sum_j \bar{v}_j w_j$ where $\bar{v}_j$ is complex conjugate

Inner Product $\qquad \overline{(a+bi)} = a - bi$

$$\overline{(a+bi)}(a+bi) = a^2 + abi - bai - b^2 i^2 = a^2 + b^2$$

Norm: $|z| := \sqrt{\bar{z} z}$ $\qquad |a+bi| = \sqrt{a^2+b^2}$



$$\langle \vec{v}, \vec{v} \rangle = \sum_j \bar{v}_j v_j = \sum_j |v_j|^2$$

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle} = \sqrt{\sum_j |v_j|^2}$$

$\vec{v} \in \mathbb{C}^k$ normalised when $\|\vec{v}\| = 1$

Examples

$|0\rangle \in \mathbb{C}^2$ | $\sqrt{1^2+0^2}=1$

$\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle =: |+\rangle$ | $\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 = \frac{1}{2}+\frac{1}{2} = 1$

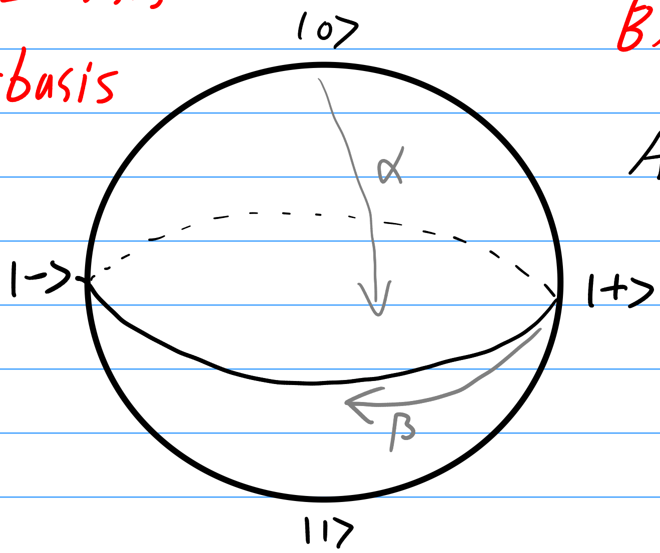
$\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle =: |-\rangle$ | $\left(\frac{1}{\sqrt{2}}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2 = 1$

$\cos\frac{\alpha}{2}|0\rangle + e^{i\beta}\sin\frac{\alpha}{2}|1\rangle$ | $\left(\cos\frac{\alpha}{2}\right)^2 + \left(\sin\frac{\alpha}{2}\right)^2 = 1$

$\{|0\rangle, |1\rangle\}$ Z-basis

$\{|+\rangle, |-\rangle\}$ X-basis

Bloch Sphere



All single-qubit states

transformations

Classical:

$0 \mapsto 0$	$0 \mapsto 1$	$0 \mapsto 0$
$1 \mapsto 1$	$1 \mapsto 0$	$1 \mapsto 0$
ID	NOT	

Not Allowed in QT

Quantum: 2 Rules

1. Linear $f: \mathbb{C}^{2^n} \to \mathbb{C}^{2^n}$

$$f(\vec{v}+\vec{w}) = f(\vec{v}) + f(\vec{w})$$
$$f(\lambda\vec{v}) = \lambda f(\vec{v})$$

Hence: Matrices M of size $2^n \times 2^n$

2. Preserve Normalisation:

$$\|\vec{v}\|=1 \Rightarrow \|f(\vec{v})\|=1$$

$\Longrightarrow$ Matrices are Unitary

$$\langle M\vec{v}, M\vec{w}\rangle = \langle \vec{v}, \vec{w}\rangle$$

OR $\quad MM^{\dagger} = ID \qquad (M^{\dagger})_{ij} = \overline{M_{ji}}$

$$MM^{\dagger} = ID$$

Examples

$$ID = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad NOT = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \quad \left(\text{Because } S^{\dagger} = \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} \text{ and } SS^{\dagger} = \begin{pmatrix} 1 & 0 \\ 0 & i \cdot -i \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right)$$

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \qquad H^2 = ID \quad \text{Hadamard}$$

$$CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

NON-Examples

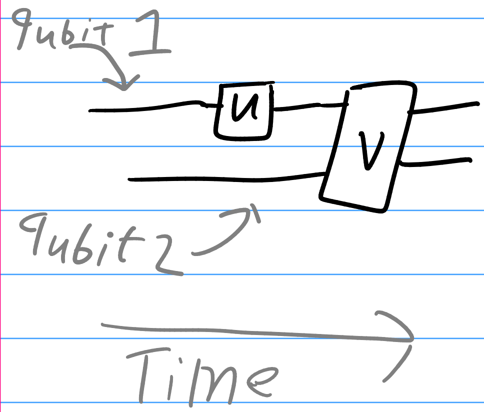
$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{Because} \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}^2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The operation $\begin{matrix} |0\rangle \mapsto |0\rangle \\ |1\rangle \mapsto |0\rangle \end{matrix}$ is $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$

Note $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}|-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

So $\||-\rangle\| = 1$, But $\left\|\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}|-\rangle\right\| = 0$

Q Circuits



Tensor Product

Let V and W be v. spaces w/ Bases $\{|v_i\rangle\}_{i=1}^{n}$ and $\{|w_j\rangle\}_{j=1}^{m}$

Then $V\otimes W$ is v. space w/ Basis $\{|v_i\rangle\otimes|w_j\rangle\}_{i,j=1}^{n,m}$

$\dim V = n \quad \dim W = m$

$\Rightarrow \dim(V\otimes W) = n\cdot m$

$V = W = \mathbb{C}^2$

$|v_0\rangle = |w_0\rangle = |0\rangle$

$|v_1\rangle = |w_1\rangle = |1\rangle$

Then Basis of $\mathbb{C}^2\otimes\mathbb{C}^2$ is $\{|0\rangle\otimes|0\rangle, |0\rangle\otimes|1\rangle, |1\rangle\otimes|0\rangle, |1\rangle\otimes|1\rangle\}$

Special Notation: $|00\rangle := |0\rangle\otimes|0\rangle$

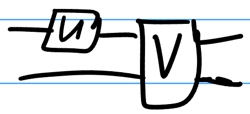
If $A: \mathbb{C}^k\to\mathbb{C}^k$, $B: \mathbb{C}^l\to\mathbb{C}^l$ are matrices, Then $A\otimes B: \mathbb{C}^k\otimes\mathbb{C}^l\to\mathbb{C}^k\otimes\mathbb{C}^l$ via

$(A\otimes B)|v\rangle\otimes|w\rangle = (A|v\rangle)\otimes(B|w\rangle)$

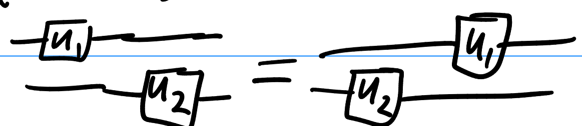
If $A = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$. Then $A\otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1k}B \\ \vdots & \ddots & \vdots \\ a_{k1}B & \cdots & a_{kk}B \end{pmatrix}$

Ex: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\otimes\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

Lemma: If U_1, U_2 Unitary, Then $U_1\otimes U_2$ also Unitary

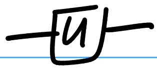
So $W =$  means $U: \mathbb{C}^2\to\mathbb{C}^2$, $V: \mathbb{C}^2\otimes\mathbb{C}^2\to\mathbb{C}^2\otimes\mathbb{C}^2$ and

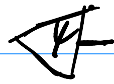
$W = V\circ(U\otimes ID)$

NOTE: 

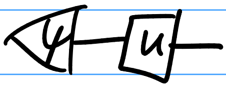
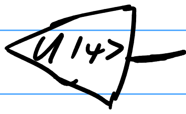
$(ID\otimes U_2)\circ(U_1\otimes ID) = (U_1\otimes ID)\circ(ID\otimes U_2)$

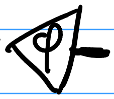
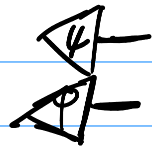
States & Effects

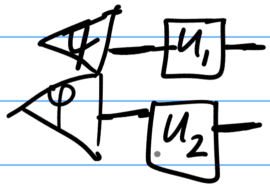
 has 1 input & 1 output

States have 0 inputs  $\in \mathbb{C}^2$

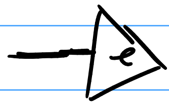
(They just "are")

 = New state 

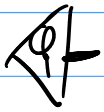
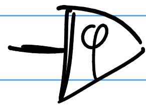
Tensor Prod.:  $\otimes$  = 

So  $= (U_1 \otimes U_2)(|\psi\rangle \otimes |\varphi\rangle)$

$= (U_1|\psi\rangle) \otimes (U_2|\varphi\rangle)$

Effects have 0 outputs 

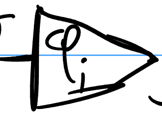
$e: \mathbb{C}^2 \to \mathbb{C}$ It takes a state and produces a number $e|\psi\rangle$.

Ex: for state  we have effect 

given by  $= \langle\varphi|\psi\rangle$ ← Inner Product

A Basis $|\varphi_1\rangle, \dots, |\varphi_k\rangle$ Forms a measurement

given state $|\psi\rangle$

$\{$  $\}_i$ w/ Probabilities $P(i \mid |\psi\rangle) = |\langle\varphi_i|\psi\rangle|^2$

observing outcome i

Check $\sum_i P(i \mid |\psi\rangle) = 1$

For instance, suppose

$$|\psi_i\rangle = |i\rangle = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Then $\sum_i P(i \mid |\psi\rangle) = \sum_i |\langle i|\psi\rangle|^2$

$$= \sum_i |\psi_i|^2 = \|\psi\|^2 = 1$$

After observing outcome i,

state $|\psi\rangle$ has collapsed to $|\varphi_i\rangle$

$$|\psi\rangle \rightsquigarrow |\varphi_i\rangle$$

measuring state again will then always give i

So: measuring Destroys Q. information

But: measuring is the only way to get information out of the system

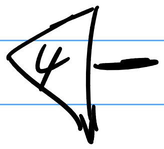
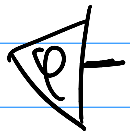
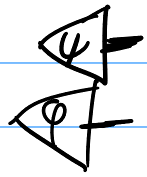
Quantum in a nutshell : scum

States are normalized vectors in complex V. space

AKA "Hilbert space"

$|\psi\rangle \in \mathbb{C}^n$ $\quad \| |\psi\rangle \|^2 = \sum |\psi_i|^2 = 1$

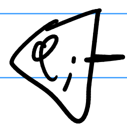
Compound systems are made by Tensor Product

 ,  $\in \mathbb{C}^2$ $\Rightarrow$  $\in \mathbb{C}^2 \otimes \mathbb{C}^2 \cong \mathbb{C}^4$

Unitaries are the allowed transformations

$U : \mathbb{C}^n \to \mathbb{C}^n \quad UU^\dagger = UU^\dagger = ID$

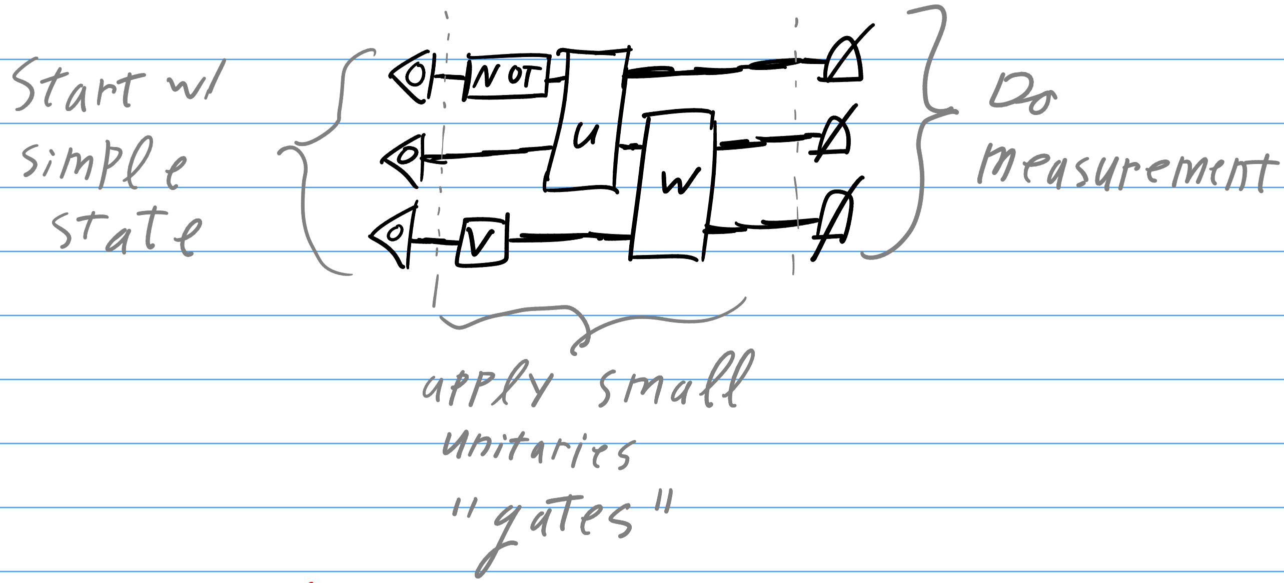
Measurements are how you get information out of a Q. System

get possible outcome  on input 

w/ Prob. $|\langle \psi | \varphi_i \rangle|^2$

Also known as The "Born rule"

Quantum Computation



A Quantum circuit

Hence, outcome is a bitstring following some prob. dist.

Goal: Find problems we can "solve" w/ high prob. using a "small" Q. circuit

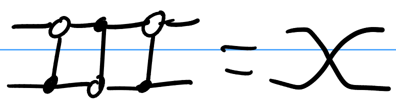
"solve" := we get some bitstring we can postprocess into an answer

Problem: n-qubit computations are $2^n \times 2^n$ matrices

$\Rightarrow$ Hard to work with

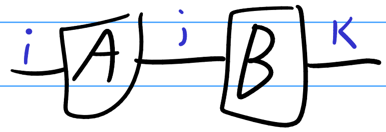
Solution: Diagrams

Instead of $\begin{pmatrix} \begin{smallmatrix}1 & \\ & 1\end{smallmatrix} & 0 \\ 0 & \begin{smallmatrix}0 & 1 \\ 1 & 0\end{smallmatrix} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \begin{smallmatrix}1 & \\ & 1\end{smallmatrix} & 0 \\ 0 & \begin{smallmatrix}0 & 1 \\ 1 & 0\end{smallmatrix} \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & 0 & 1 & \\ & 1 & 0 & \\ & & & 1 \end{pmatrix}$

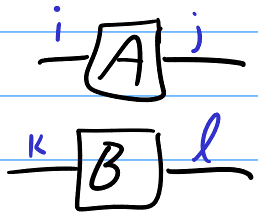
Have 

9

Tensor Networks

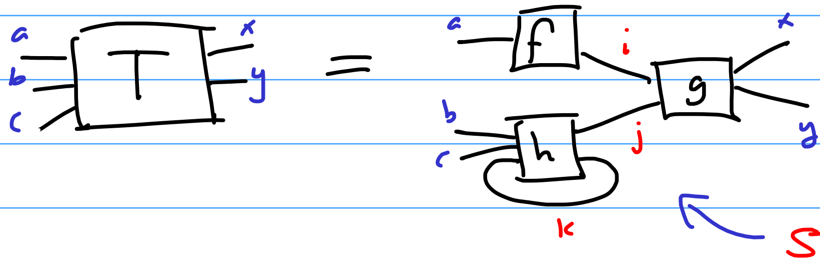


$$(BA)_i^k = \sum_j A_i^j B_j^k \qquad (A|\psi\rangle)^j = \sum_i A_i^j \psi^i$$



$$(A\otimes B)_{ik}^{jl} = A_i^j B_k^l$$

More generally:



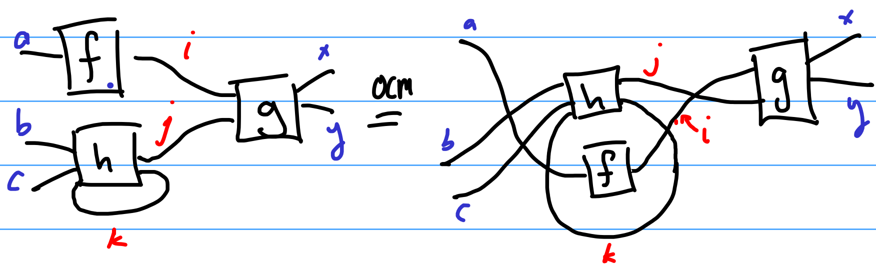
string diagram

has matrix elems:

$$T_{a,b,c}^{x,y} = \sum_{ijk} f_a^i \, g_{ij}^{xy} \, h_{bck}^j$$ ← tensor network

String Diagram MOTTO:

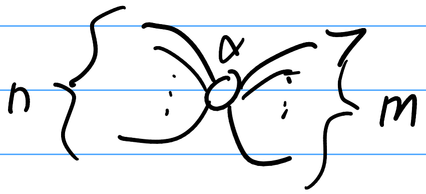
<u>Only connectivity matters</u>



ZX-Diagrams

a specific type of tensor network

Z-Spiders



$$:= \underbrace{|00\cdots0\rangle}_{m}\underbrace{\langle 00\cdots0|}_{n} + e^{i\alpha}|11\cdots1\rangle\langle 11\cdots1|$$

$$\longleftrightarrow \quad 2^n \left\{ \underbrace{\begin{pmatrix} 1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \\ & & & & e^{i\alpha} \end{pmatrix}}_{2^m} \right.$$

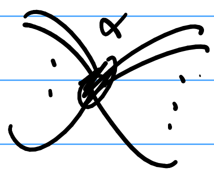
i.e $\begin{cases} |0\cdots0\rangle \mapsto |0\cdots0\rangle \\ |1\cdots1\rangle \mapsto e^{i\alpha}|1\cdots1\rangle \\ \text{other basis states} \mapsto 0 \end{cases}$

Note: in general not unitary

i.e a tensor $Z[\alpha]^{j_1\cdots j_m}_{i_1\cdots i_n} := \begin{cases} 1 & \text{if } i_1=\cdots=i_n=j_1=\cdots=j_m=0 \\ e^{i\alpha} & \text{if } i_1=\cdots=i_n=j_1=\cdots=j_m=1 \\ 0 & \text{else} \end{cases}$

X-Spiders

Recall $|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ $|-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$



$$:= |+\cdots+\rangle\langle +\cdots+| + e^{i\alpha}|-\cdots-\rangle\langle -\cdots-|$$

i.e same as Z-spider, but w/ $\{|+\rangle, |-\rangle\}$ basis instead of $\{|0\rangle, |-\rangle\}$ basis

ZX Examples

No inputs

 $= |0\rangle + |1\rangle = \begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\1\end{pmatrix} = \sqrt{2}\,|+\rangle$

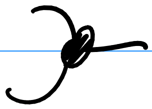
 $= |0\rangle + e^{i\pi}|1\rangle = \begin{pmatrix}1\\0\end{pmatrix} - \begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\-1\end{pmatrix} = \sqrt{2}\,|-\rangle$

 $= |+\rangle + |-\rangle = \frac{1}{\sqrt{2}}\left[\begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}1\\-1\end{pmatrix}\right] = \frac{1}{\sqrt{2}}\begin{pmatrix}2\\0\end{pmatrix} = \sqrt{2}\,|0\rangle$

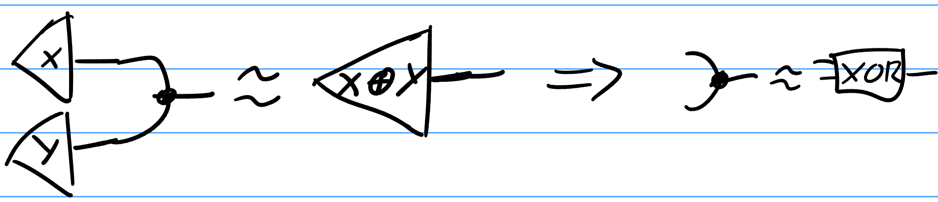
 $= |+\rangle - |-\rangle = \dots = \sqrt{2}\,|1\rangle$

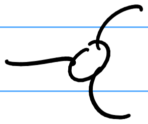
 $= |0\rangle\langle 0| + e^{i\alpha}|1\rangle\langle 1| = \begin{pmatrix}1&0\\0&0\end{pmatrix} + e^{i\alpha}\begin{pmatrix}0&0\\0&1\end{pmatrix} = \begin{pmatrix}1&0\\0&e^{i\alpha}\end{pmatrix}$

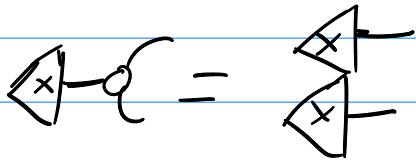
is unitary!

 $\approx \begin{pmatrix}1&0&0&1\\0&1&1&0\end{pmatrix} \leftrightsquigarrow \begin{cases}|00\rangle \mapsto |0\rangle\\ |01\rangle \mapsto |1\rangle\\ |10\rangle \mapsto |1\rangle\\ |11\rangle \mapsto |0\rangle\end{cases}$

means "proportional to"

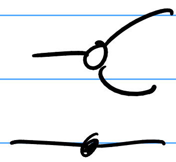


 acts as a "copy" for Z-basis states $\{|0\rangle, |1\rangle\}$

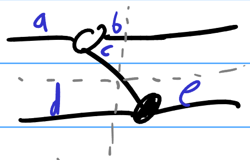


Composing ZX-Diagrams

Vertical = Tensor product



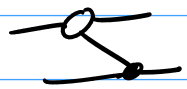
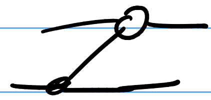
Horizontal = matrix mult



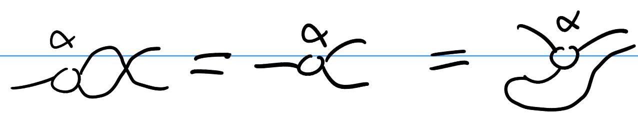
$$= (ID \otimes XOR) \circ (COPY \otimes ID)$$

$$= \sum_c Z[0]^{bc}_{a} \; X[0]^{e}_{cd}$$

is unitary

Note  = 

This is because spiders are symmetric tensors

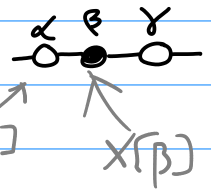


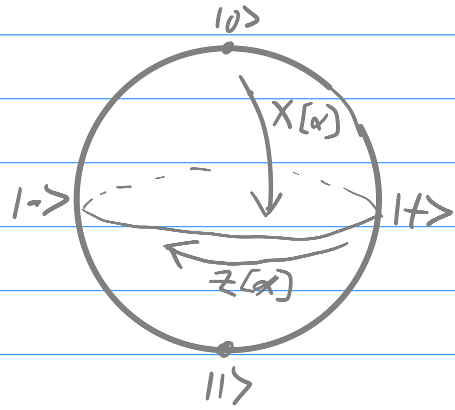
In general: Can treat ZX-Diagrams as undirected graphs

Universality

__Thm__ (Euler decomposition) For any single-qubit unitary U,

$\exists$ angles $\alpha, \beta, \gamma, \theta$ s.t.:

$U = e^{i\theta} \cdot$ 

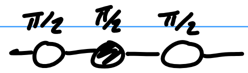
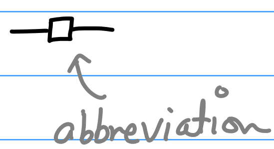


__Ex__ The Hadamard $H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

$H|0\rangle = |+\rangle \quad H|1\rangle = |-\rangle$

$H|+\rangle = |0\rangle \quad H|-\rangle = |1\rangle$

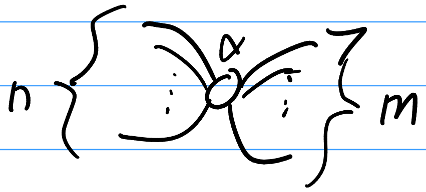
basis transform between Z-basis $\{|0\rangle, |1\rangle\}$ and X-basis $\{|+\rangle, |-\rangle\}$

 $= e^{-i\pi/4}$  $=:$ 

Thm: Any Linear map from $\mathbb{C}^{2^n}$ to $\mathbb{C}^{2^m}$ can be written as a ZX-Diagram

ZX-Calculus

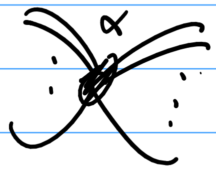
Z-Spiders



$$:= \underbrace{|00\ldots0\rangle}_{m}\underbrace{\langle 00\ldots0|}_{n} + e^{i\alpha}|11\ldots1\rangle\langle 11\ldots1|$$

X-Spiders

Recall $|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$ $|-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$



$$:= |+\ldots+\rangle\langle+\ldots+| + e^{i\alpha}|-\ldots-\rangle\langle-\ldots-|$$

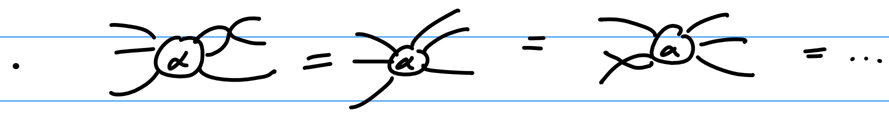
Thm: Any Linear map from $\mathbb{C}^{2^n}$ to $\mathbb{C}^{2^m}$ can be written as a ZX-Diagram

But: Can also reason diagrammatically

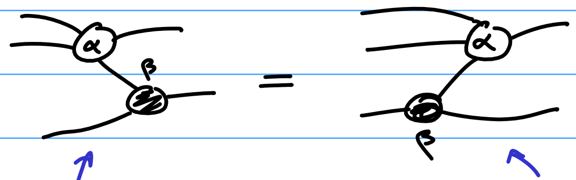
ZX diagrams have "extreme" OCM.

They are invariant under:

- SWAPPING SPIDER-LEGS:



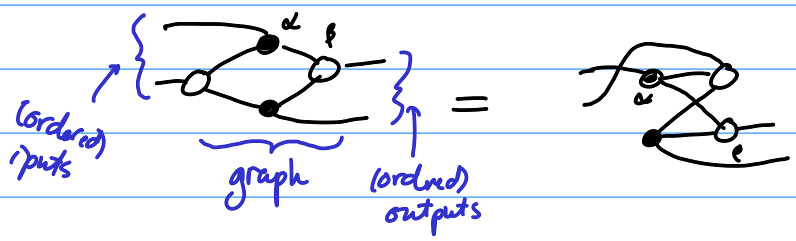
- CHANGING DIRECTION



$$(I\otimes X[\beta]^1_2)(Z[\alpha]^2_2\otimes I) = (Z[\alpha]^1_3\otimes I)(I\otimes I\otimes X[\beta]^2_1)$$

$\Rightarrow$ they can be treated as <u>undirected graphs</u> (w lists of inputs & outputs)

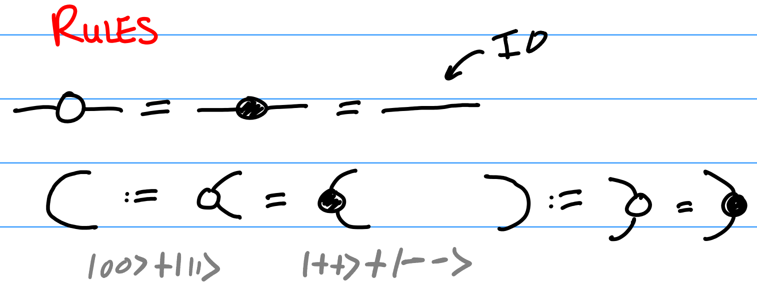
eg.



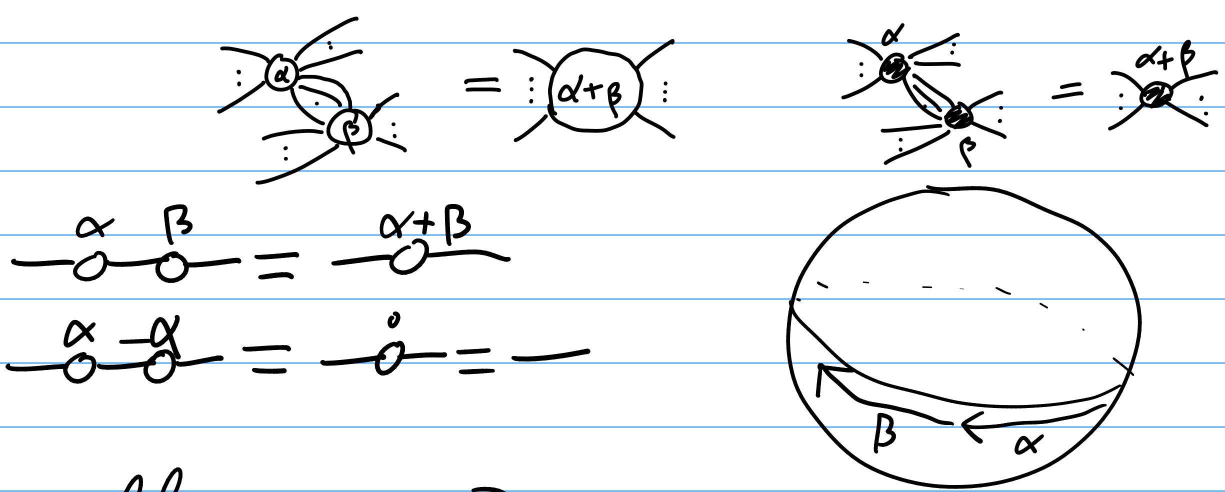
In addition, we have rewrite rules

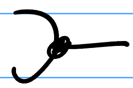
We call this the ZX-Calculus

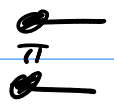
(0) "WIRE" RULES



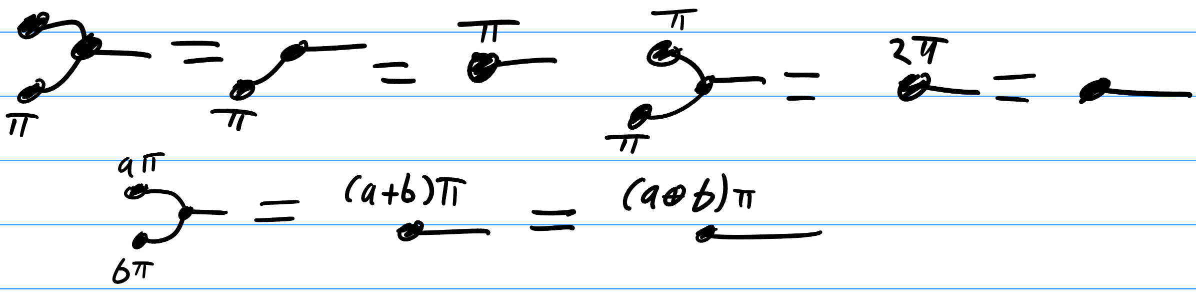
(1) SPIDER-FUSION

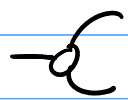


Recall: XOR ≈ 

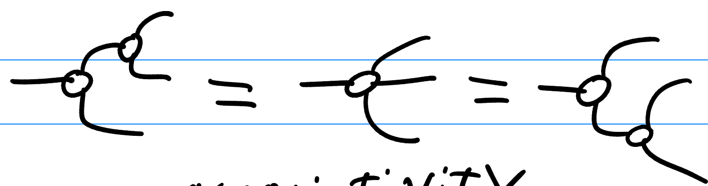
$|0\rangle \approx$ / $|1\rangle \approx$ 

$\Big\}$ $a\pi$ —— $\approx |a\rangle \quad a \in \{0,1\}$



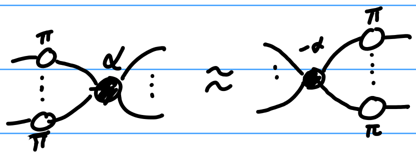
COPY = 

$|x\rangle \mapsto |xx\rangle$



associativity

(2) π - rule:

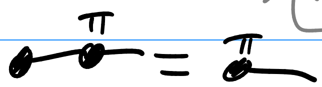


 $= |0\rangle\langle 0| - |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. We call this the Pauli Z

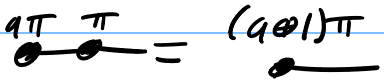
$Z := Z[\pi]$

 $= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ Pauli X

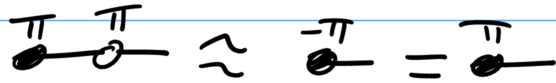
NOT gate



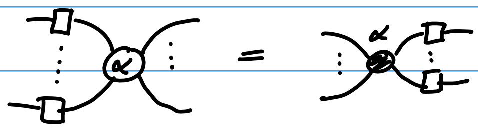
NOT$|0\rangle = |1\rangle$

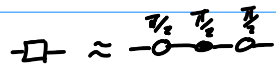


Ex: 



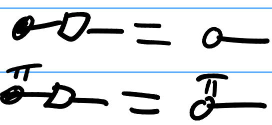
(3) Colour Change:

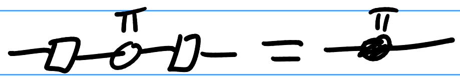


where  $= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

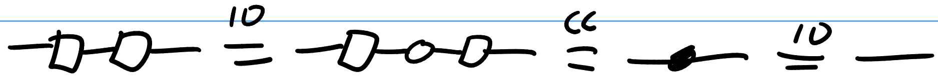
Hadamard

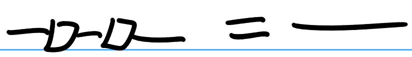
works because $H|0\rangle = |+\rangle$
$H|1\rangle = |-\rangle$



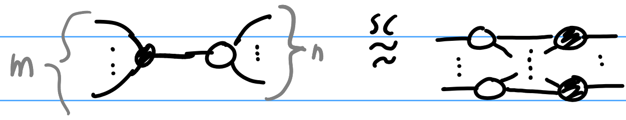
Ex: 

$HZH = HXH$

Note: 

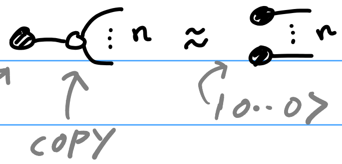
So: 

(4) Strong complementarity

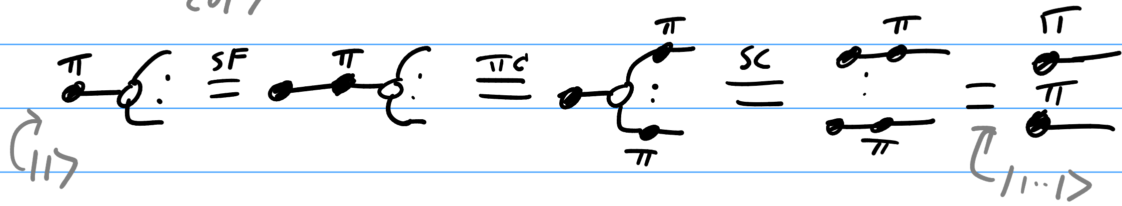


Special cases:

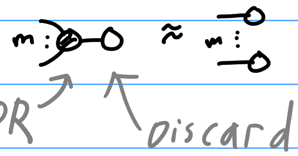
COPY: m=0 ⇒



Also:

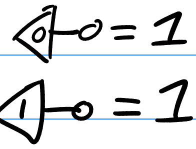


Discarding: n=0 ⇒

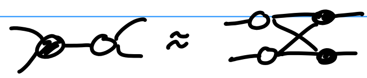


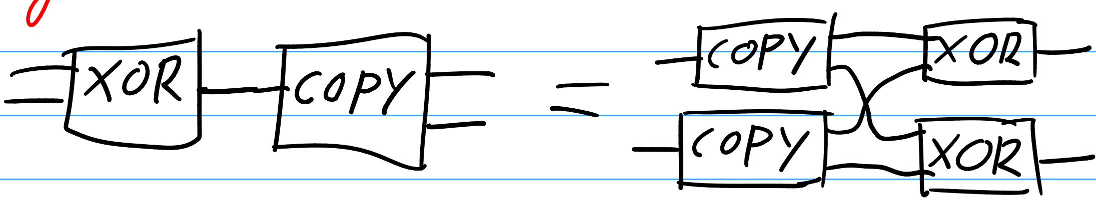
"Applying a function, Then throwing away the output, is the same as throwing away the inputs"

$-\circ = \langle 0| + \langle 1|$, So



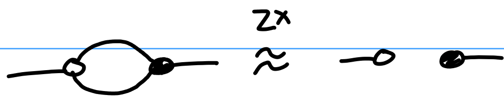
bialgebra m=2, n=2 ⇒



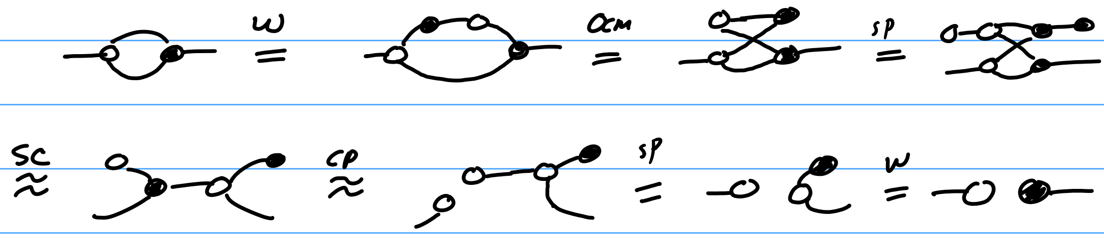


Rewriting examples

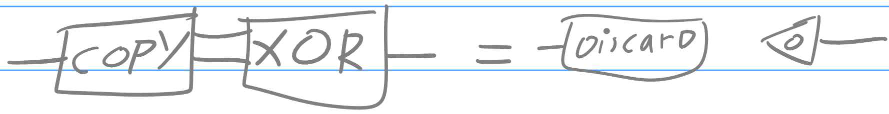
Thm (COMPLEMENTARITY)



Pf

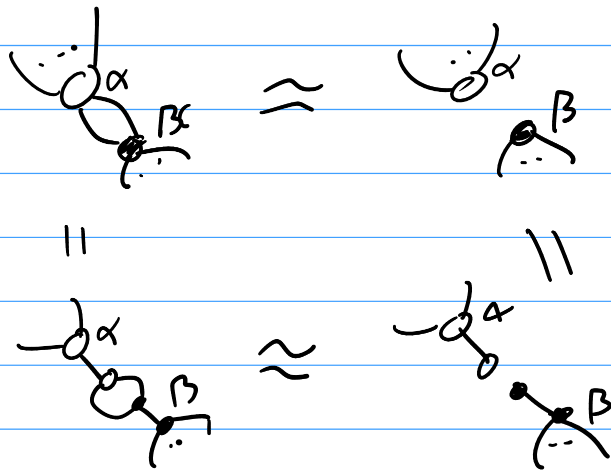


∎

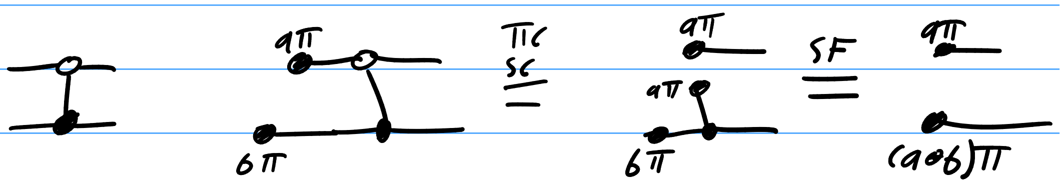


$x \oplus x = 0$

Corollary:

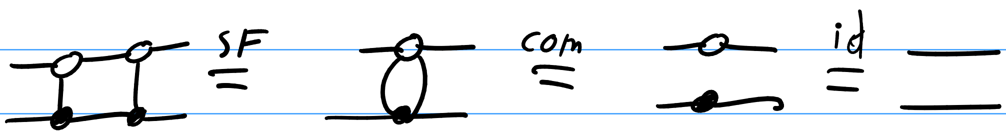


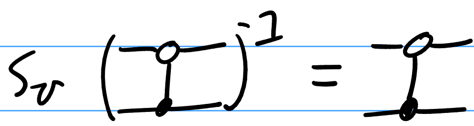
CNOT :=



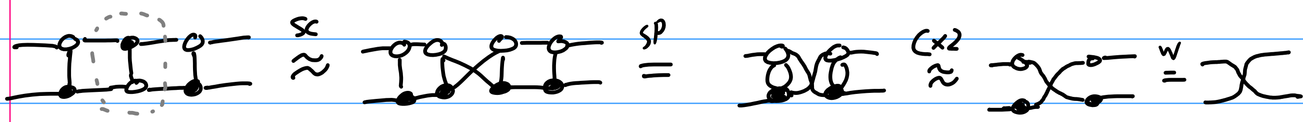
"controlled"

$CNOT|a,b\rangle = |a, a\oplus b\rangle$





Ex $3\,CNOT = SWAP$

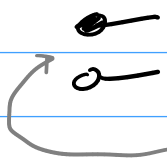


Note: CNOTs are important:

Thm: any n-qubit unitary can be written as a circuit of CNOT, $Z[\alpha]$ and $X[\alpha]$ gates

So: CNOT is the "only" 2-qubit interaction we need

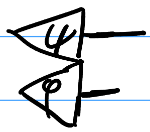
Bell state

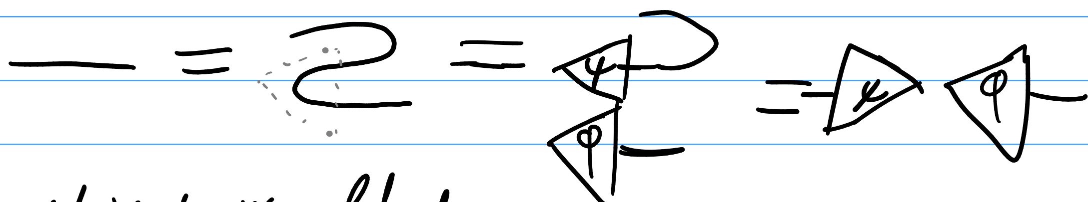
 is an example of a product state

No connection

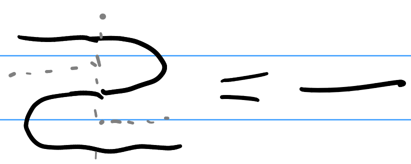
We also have entangled states

Bell state: $\subset \; = |00\rangle + |11\rangle$ The "cup"

Suppose $\subset =$  for some $|\psi\rangle, |\varphi\rangle$

Then 

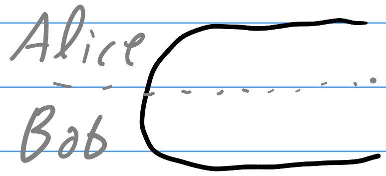
So everything would disconnect

Note  is known as the Yanking equation

$(CAP \otimes ID) \circ (ID \otimes CUP) = ID$

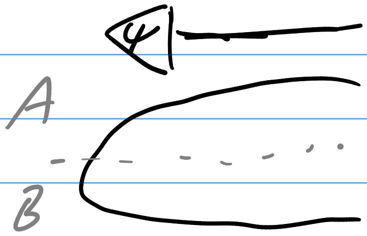
Teleportation

1. Alice & Bob start with a Shared Bell state

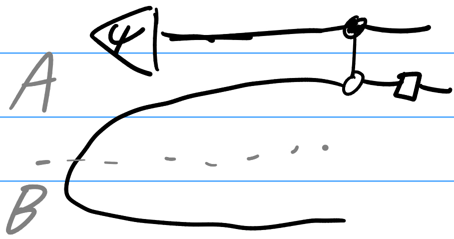


2. Then they may move far apart

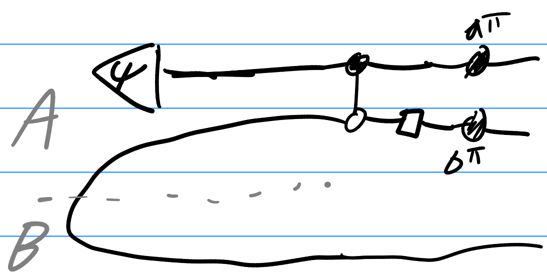
3. Alice Picks a Q. state $|\psi\rangle$ She wants to send to Bob

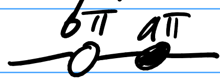


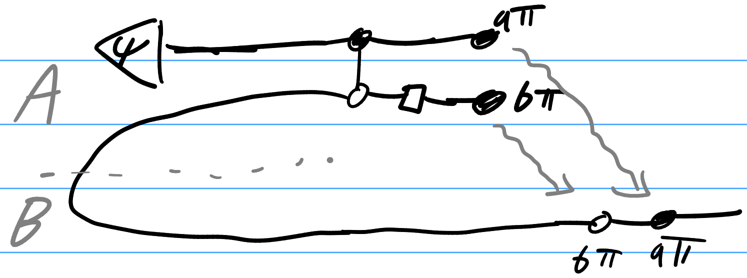
4. Alice performs CNOT & Had on her states



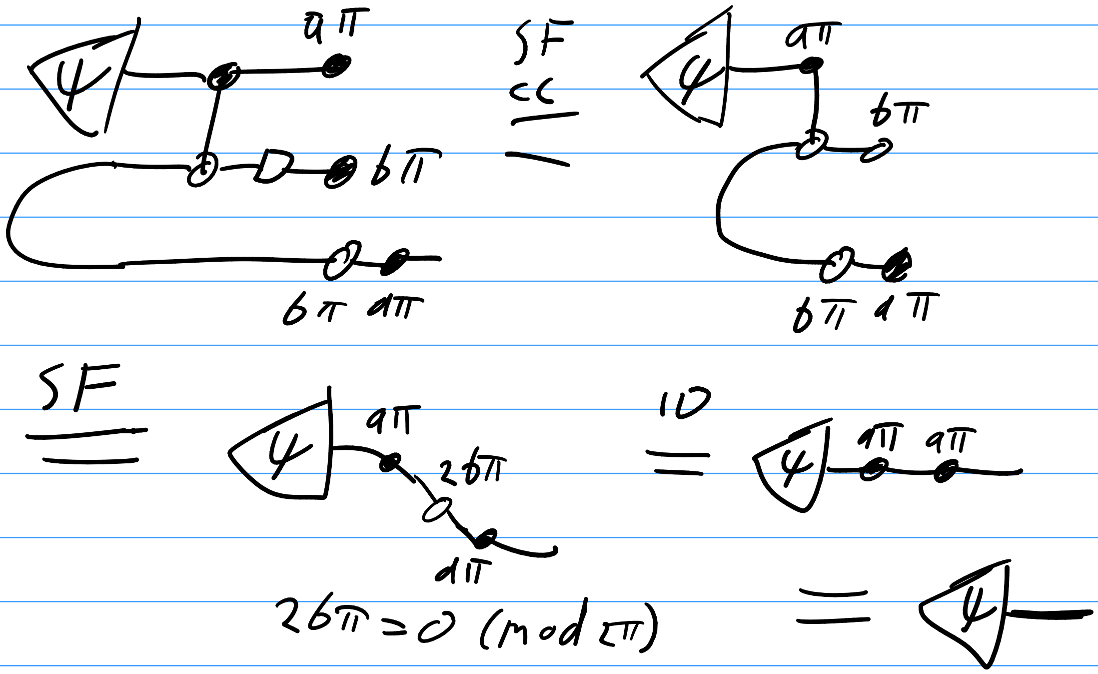
5. Then she measures both her qubits in $\{|0\rangle,|1\rangle\}$ basis, getting outcomes $a,b\in\{0,1\}$



6. She communicates a, b to Bob,
who performs a  correction



7. Now Bob's state is Alice's former state:

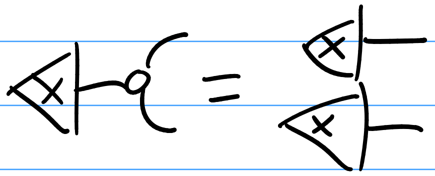


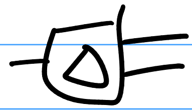
Conclusion: using Entanglement
& classical communication,
Alice can send Q. info to Bob

Needed! No FTL comm.

No cloning theorem

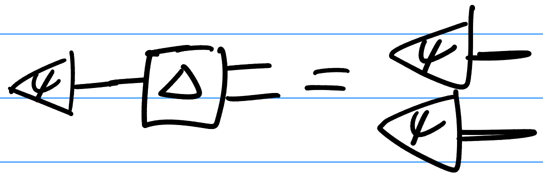
We can "clone", i.e. copy, classical states

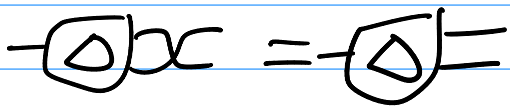


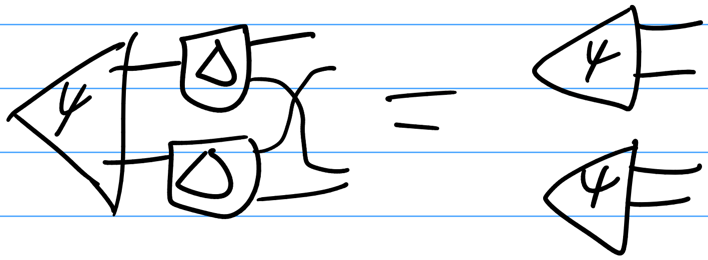
Is there some process  that

clones arbitrary Q. states?

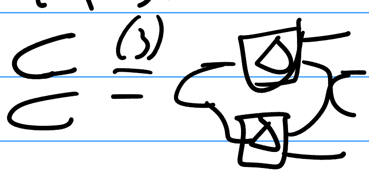
Def: A map Δ is a *cloning process* when

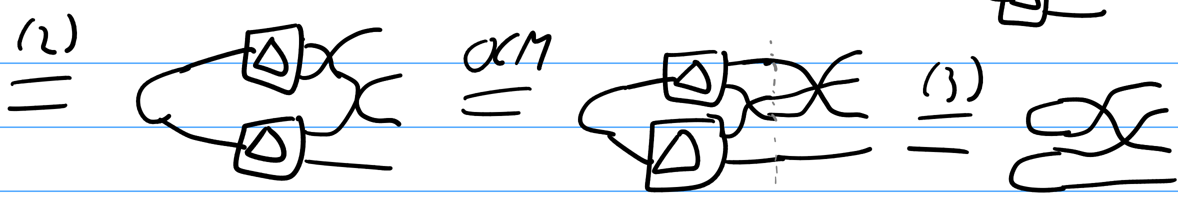
(1)  ∀ normalised 

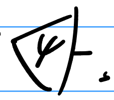
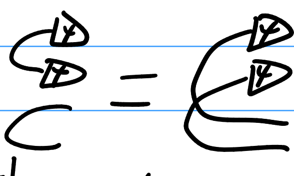
(2) 

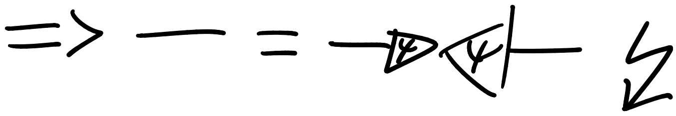
(3) 

Thm: No Cloning Process exists

Pf: Suppose it did. Then 

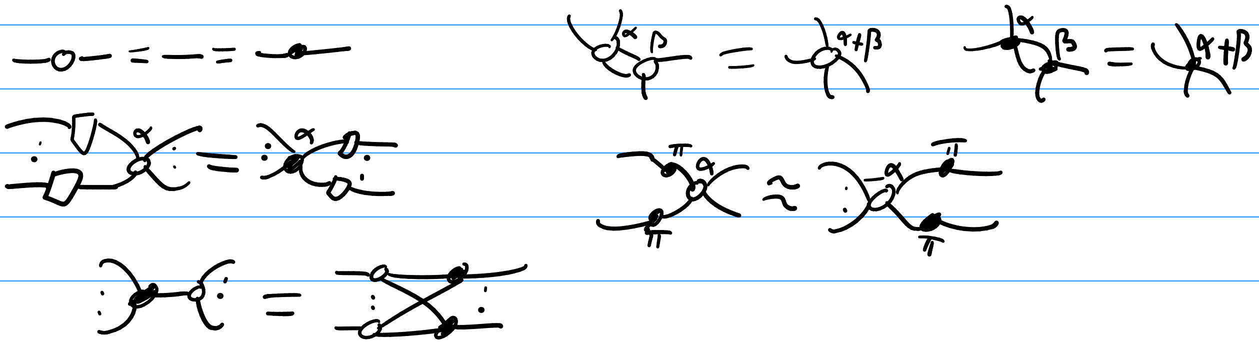


 Then pick a state : 



Completeness

We have 5 types of rewrites:

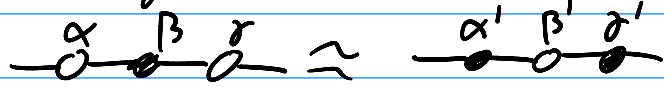


Q: How much can we prove w/ this?

Def: We say a set of rules is complete when two diagrams representing the same linear map can always be rewritten into each other by the rules

Thm: The rules above are complete for diagrams with all phases in set $\{0, \pi, \frac{\pi}{2}, -\frac{\pi}{2}\}$ The Clifford fragment

Thm: it is not complete over all phases, but adding one additional rule makes it complete



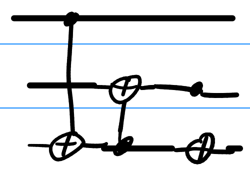
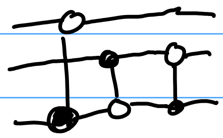
where
$$\left.\begin{aligned} \alpha' &= f_1(\alpha,\beta,\gamma) \\ \beta' &= f_2(\alpha,\beta,\gamma) \\ \gamma' &= f_3(\alpha,\beta,\gamma) \end{aligned}\right\} \text{complicated functions}$$

CNOT Circuits & Phase Free ZX Diagrams

Circuits made just out of 

ZX-Diags made out of  and 

Prop Any CNOT circuit is equal to a phase free ZX-diagram.

 $\rightsquigarrow$ 

Q: What about the converse?

Thm: (Unitary) phase-free ZX-diags $\rightsquigarrow$ CNOT circuits.

Parities

$$\text{CNOT}\,|x,y\rangle \mapsto |x, x\oplus y\rangle$$

$$\text{CNOT}\,|x,y\rangle \mapsto |f_1(x,y), f_2(x,y)\rangle \quad \text{where} \begin{cases} f_1(x,y) = x \\ f_2(x,y) = x\oplus y. \end{cases}$$

Def A function of the form $f(x_1,\ldots,x_n) = x_{i_1} \oplus \ldots \oplus x_{i_k}$ is called a parity map.

Def The field $\mathbb{F}_2$ has elements $\{0,1\}$ where:

$$x\cdot y := x \wedge y \qquad x+y = x\oplus y \qquad (\text{ie. } x+y \bmod 2)$$

Sometimes we call some $v \in \mathbb{F}_2$ a parity.

$$\text{par}(\vec{b}) = \sum_i b_i \quad \leftarrow \text{in } \mathbb{F}_2$$

$\text{par}(\vec{b}) = 0$ means $\vec{b}$ has an even # of 1's

$\text{par}(\vec{b}) = 1$ means odd #.

Parities for subsets of bits:

$$\begin{pmatrix} 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = b_1 \oplus b_3 \oplus b_4$$

Multiple parities at once:

$$\underbrace{\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}}_{\text{parity matrix.}} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} = \begin{pmatrix} b_1 \oplus b_3 \oplus b_4 \\ b_2 \oplus b_3 \\ b_1 \oplus b_4 \\ b_4 \end{pmatrix}$$

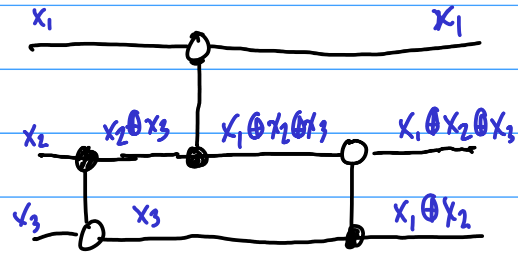
THM The action of a CNOT circuit on basis elements is defined by an invertible parity matrix:

$$C|b_1, \dots, b_n\rangle = |c_1, \dots, c_n\rangle$$

where $P\begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$. $\qquad PP^{-1} = ID$

Note: P is $n \times n$, so not exponentially large

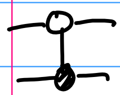
Ex:

$C =$ 

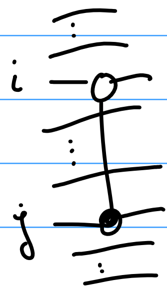
$$C|x_1, x_2, x_3\rangle = |x_1,\ x_1 \oplus x_2 \oplus x_3,\ x_1 \oplus x_2\rangle$$

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}}_{P} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_1 \oplus x_2 \oplus x_3 \\ x_1 \oplus x_2 \end{pmatrix}$$

Special case: Single CNOT.

 $|x, y\rangle \mapsto |x, x \oplus y\rangle$ $\qquad \underbrace{\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}}_{P} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ x \oplus y \end{pmatrix}$

More generally:



$\longleftrightarrow$

$$\begin{array}{cc} j \to & \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & \boxed{\ } & & 1 & \\ & & & & 1 \end{pmatrix} \\ & \quad\ \ \uparrow \\ & \quad\ \ i \end{array} = \underbrace{E^{ij}}_{\text{elementary matrix}}$$

$$E^{ij} A = \underset{\substack{\uparrow \\ \text{row } j = \text{row } j + \text{row } i}}{A'}$$

$$A E^{ji} = \underset{\substack{\uparrow \\ \text{col } j := \text{col } i + \text{col } j}}{A'}$$

Suppose $PE^{i_1j_1}\cdots E^{i_kj_k} = I$,

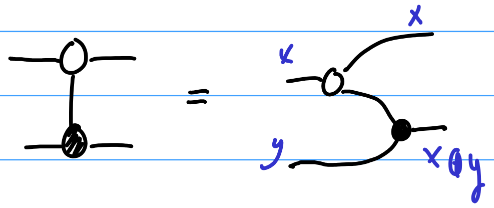
then $P = E^{i_kj_k}\cdots E^{i_1j_1}$

parity matrix CNOT gates!

Algorithm: CNOT-SYNTH:

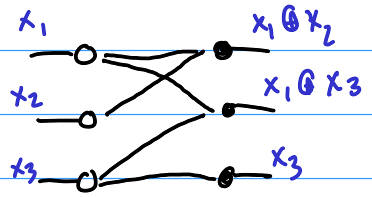
* Start w/ Parity matrix P, empty circ. C.
* Do Gauss-Jordan reduction of columns of P.
  - whenever an elem. col operation E^{ji} is applied, append CNOT^{ji} to C.
* C now implements P.

Parity maps in ZX



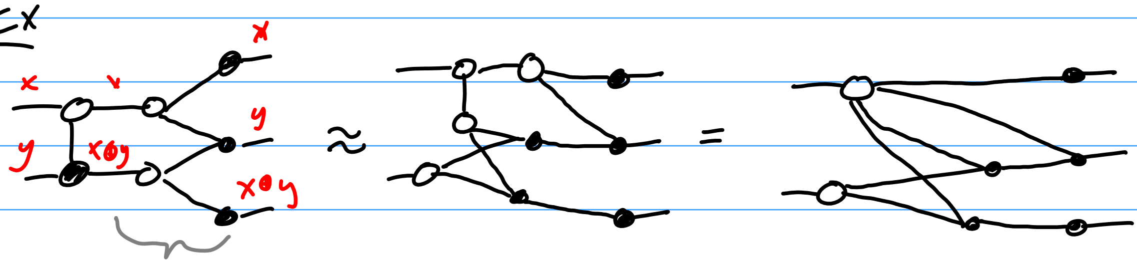
More general parity maps:

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \oplus x_2 \\ x_1 \oplus x_3 \\ x_3 \end{pmatrix}$$

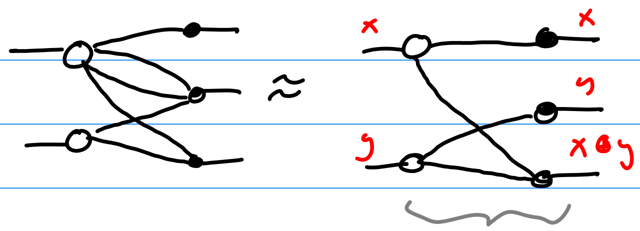


← implements P!

Ex

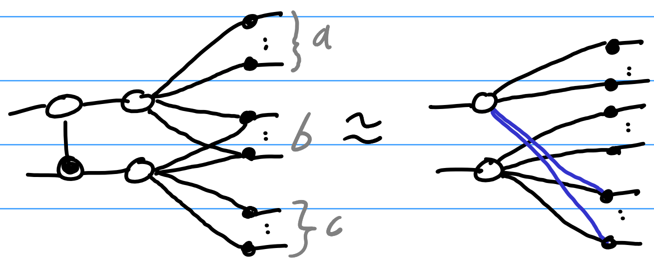


$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}$$



$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

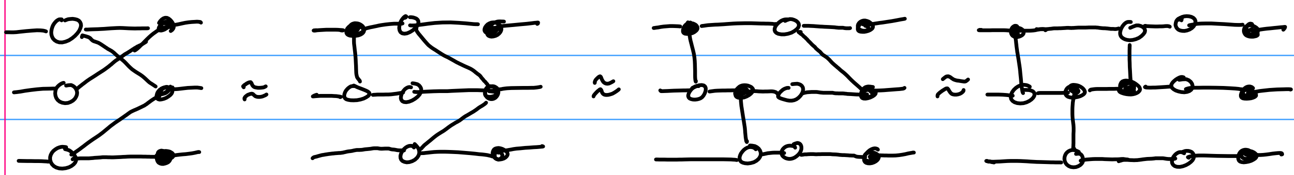
Lem 4.2.3



$$\begin{array}{c} a\{ \\ \\ b\{ \\ \\ c\{ \end{array} \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{pmatrix} \underset{\xleftarrow{c_1 = c_1 + c_2}}{\xrightarrow{c_1 = c_1 + c_2}} \begin{pmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & 1 \end{pmatrix}$$

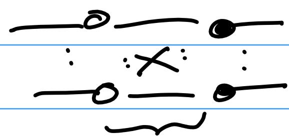
Ex

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_2 = C_2 + C_1} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_3 = C_3 + C_2} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{C_1 = C_1 + C_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$



Def A spider is called
- * an input spider if it is conn. to an input
- * an output spider . . . output
- * an interior spider otherwise.

Def A phase-free ZX-diagram is in parity normal form
- every Z spider is conn. to exactly 1 input
- every X . . . . . . output
- no wires between spiders of the same type
- no parallel wires



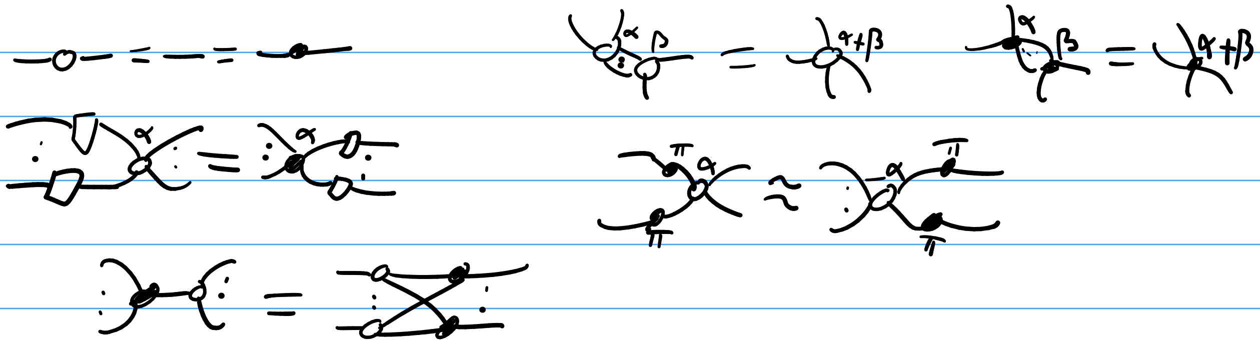
P parity matrix.

Thm: A ZX-diagram in Parity NF can be rewritten into a CNOT circuit

Missing piece: From arbitrary phase-free ZX-diagram to Parity NF

Completeness

We have 5 types of rewrites:

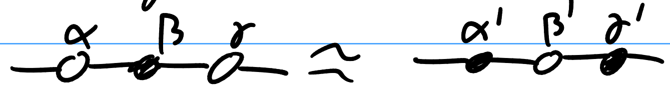


Q: How much can we prove w/ this?

Def: We say a set of rules is complete when two diagrams representing the same linear map can always be rewritten into each other by the rules

Thm: The rules above are complete for diagrams with all phases in set $\{0, \pi, \frac{\pi}{2}, -\frac{\pi}{2}\}$ The Clifford fragment

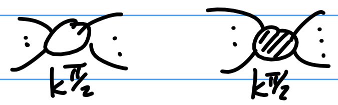
Thm: it is not complete over all phases, but adding one additional rule makes it complete



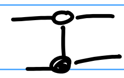
where $\alpha' = f_1(\alpha, \beta, \gamma)$
$\beta' = f_2(\alpha, \beta, \gamma)$
$\gamma' = f_3(\alpha, \beta, \gamma)$
complicated functions

Clifford diagrams and circuits

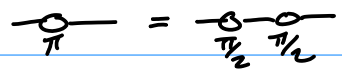
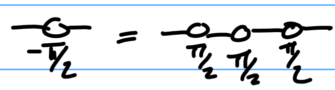
Def A ZX-diagram is **Clifford** when it is made of **Clifford spiders**

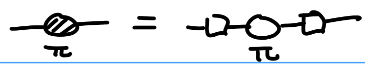
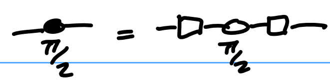
 $k \in \mathbb{Z}$.

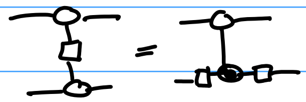
Def **Clifford circuits** are circuits made from:

$S =$  $H =$  $\text{CNOT} =$ 

Ex Some common Clifford gates:

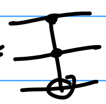
$Z =$  $S^{\dagger} =$ 

$X =$  $\sqrt{X} =$ 

$CZ =$ 

Ex Some non-Clifford gates:

$T =$  $(S = T^2)$

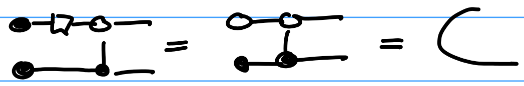
$\text{ToF} =$  $= \ldots$

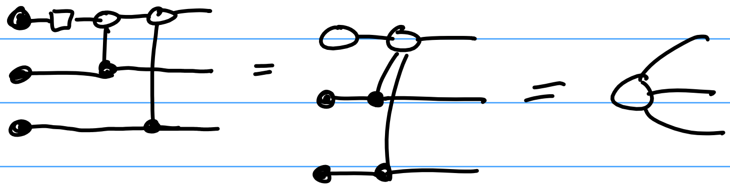
$|x, y, z\rangle \mapsto |x, y, (x \wedge y) \oplus z\rangle$

Def A Clifford state is a state $|\varphi\rangle = C|0\ldots0\rangle$ for a Clifford circuit C.

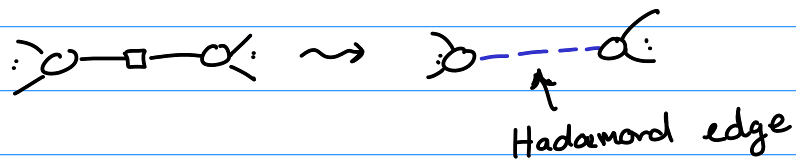
Q: Why care about Cliffords?

- Contains useful states, e.g.

Bell $\approx$ 

GHZ $\approx$ 

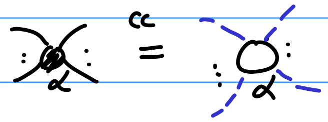
- Quantum error correction
- Eff. classical simulation (soon)
- Rich rewrite theory (now!)
- Cliffords + any other single qubit gate = approximately universal gateset

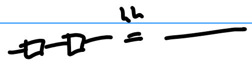
Notation: 

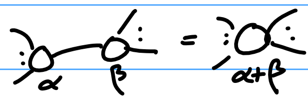
Def A ZX diagram is graph like if:

1. all spiders are Z spiders
2. all edges b/w spiders are Hadamard edges
3. no parallel edges or self-loops
4. every input/output is connected to a spider.

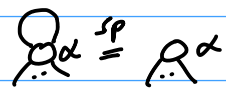
Prop: Every ZX-diagram is equal to a graph-like one.

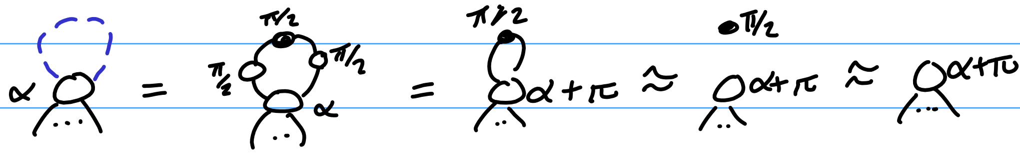
Pf 1. Use  to elim X spiders.

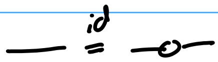
& use  to cancel extra H's.

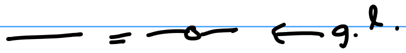
2. Use (sf) to elim non-H edges: 

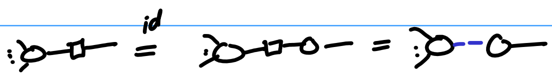
3. For parallel H-edges:

For self-loops: 

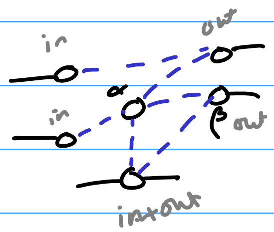


4. Use  if necessary.

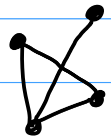
e.g. 

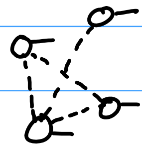


$\boxed{Z}$

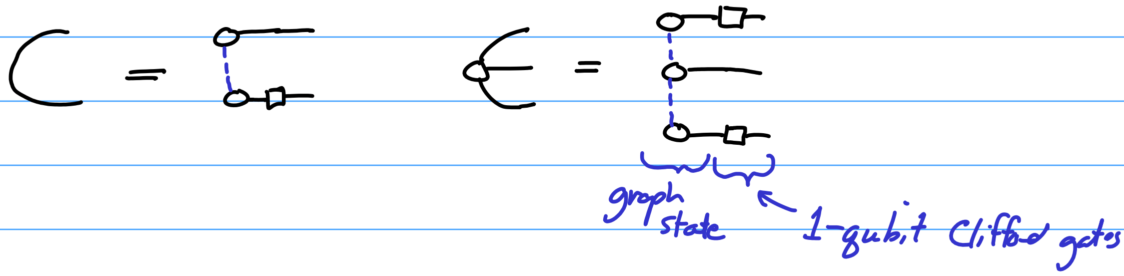
Ex 

Def A graph-like diagram is called a graph state
if: { • no inputs
• no interior spiders   interior = only connected to other spiders
• no phases

Ex: $G =$ 
graph

$|G\rangle =$ 
graph state

Some states are almost graph states, e.g.

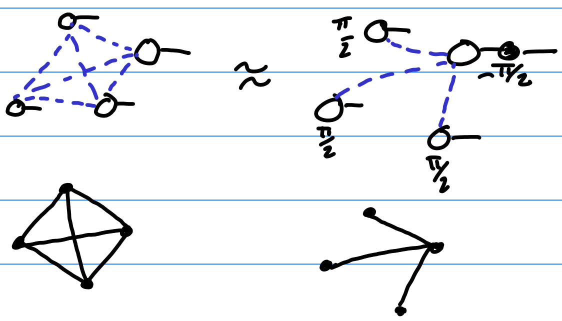


Def A graph state with local Cliffords (GSLC) is a state of the form $(U_1 \otimes \cdots \otimes U_n)|G\rangle$ for some graph state $|G\rangle$ and 1-qubit Clifford gates U_i.

Thm Any Clifford state is equal to a GSLC.

We'll need some new tools to prove this!

First, note that for GSLCs, the graph can be deceiving!



Local complementation

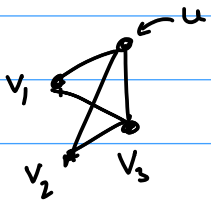
Def Let $G = (V, E)$ be a graph and $u \in V$.

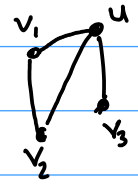
The Local complementation of G about u is a new graph $G * u = (V, E')$ where

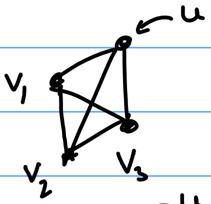
$$\forall v, w \in N_G(u). \quad (v,w) \in E' \Leftrightarrow (v,w) \notin E.$$

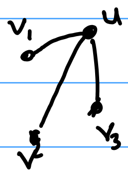
neighbourhood

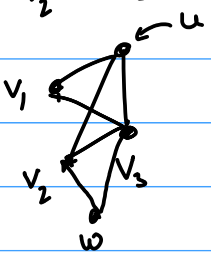
Ex

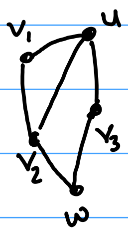
1. $G =$ 

$G * u =$ 

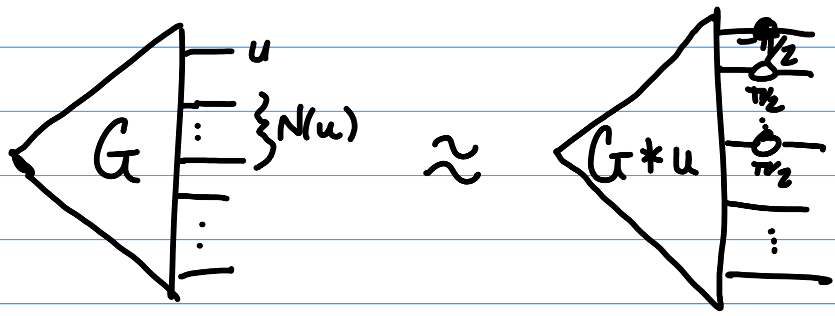
2. $G =$ 

$G * u =$ 

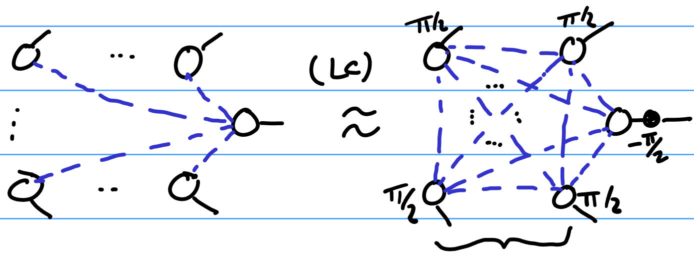
3. $G =$ 

$G * u =$ 

Prop



Graphically:

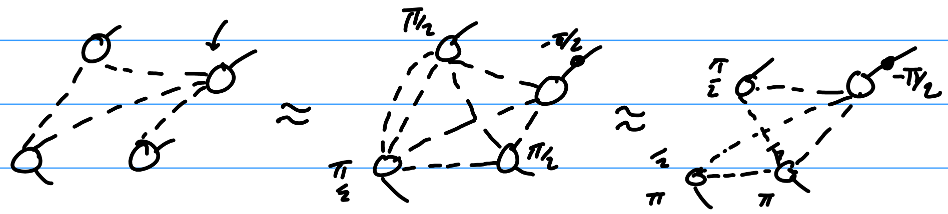


fully connected graph.

δ

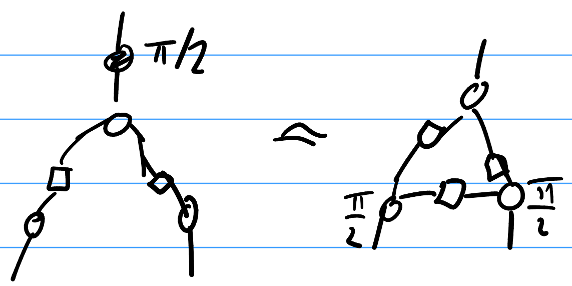
Q: Why is this the same as local comp?

A: Because 



Proving LC:

in Exercises saw base case:

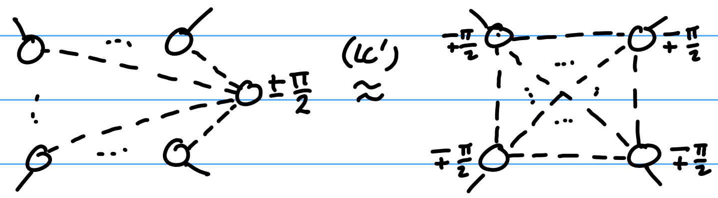


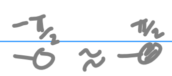
Prove rest by induction

So why do we care about LC?

it helps us remove spiders!

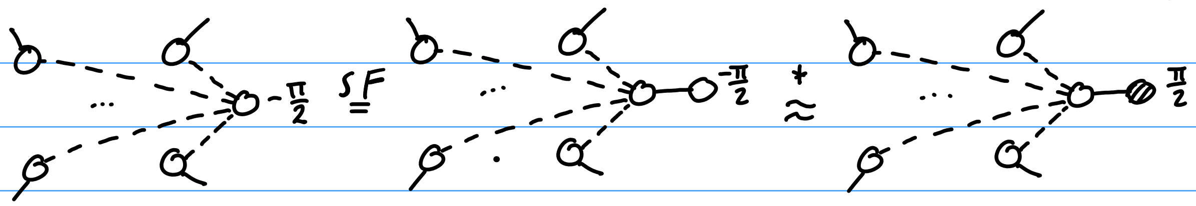
Prop

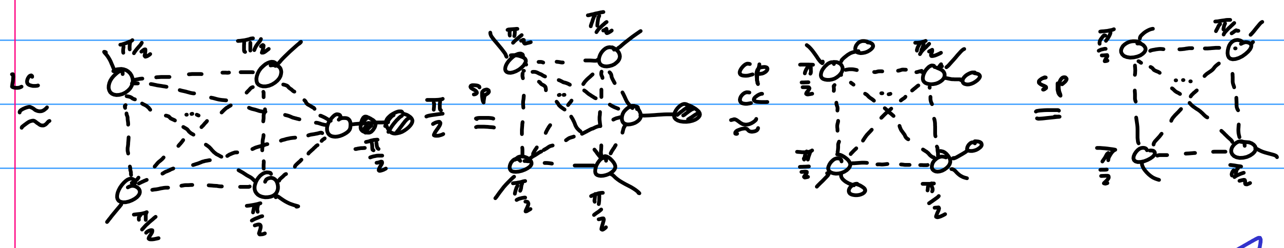


(*) 

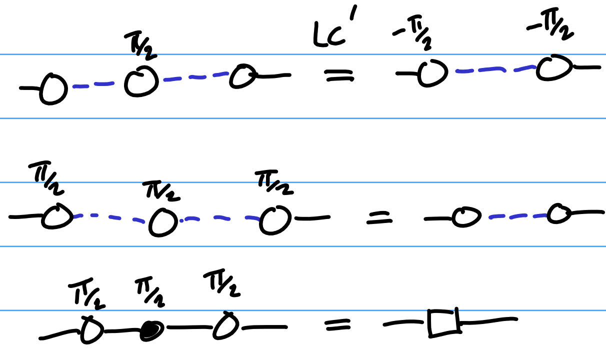
(follows from Euler + cc)

Pf



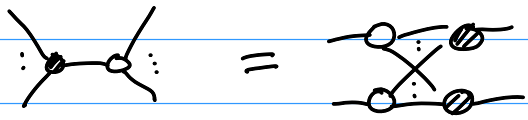


$LC' \Rightarrow EU \Rightarrow$



Pivoting.

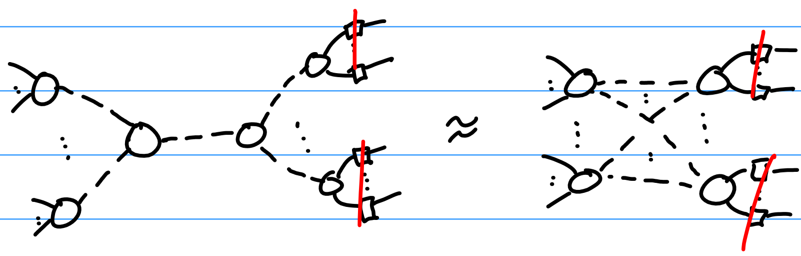
Consider the (sc) rule:

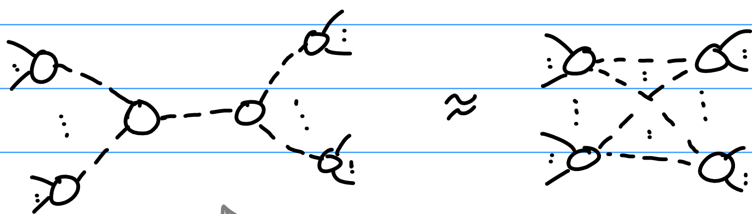


add some context:

always deletes spiders!

Now, (cc) both sides to elim X spiders:

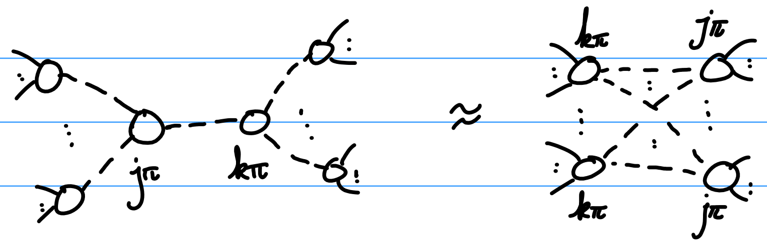




deletes 2 adj. phase-free spiders.

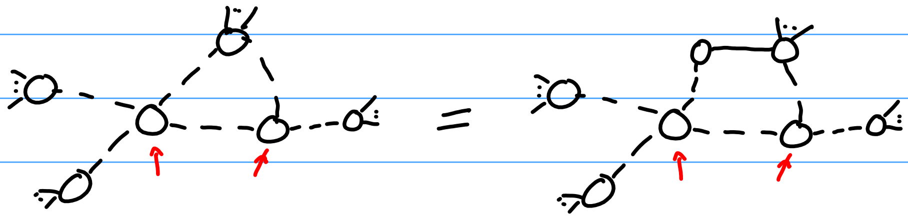
Generalisation:

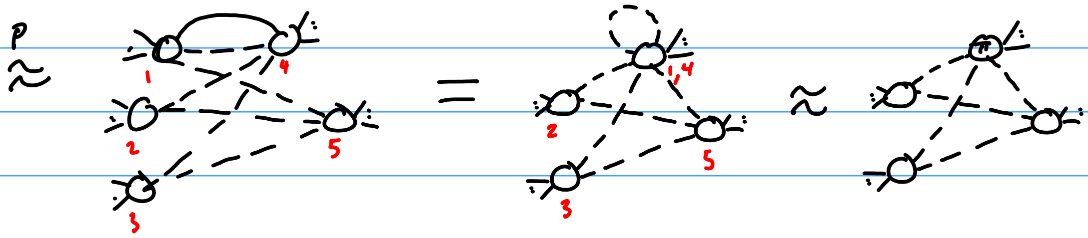
Pivot rule:



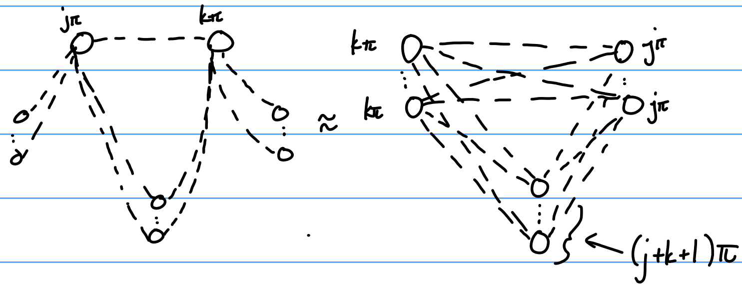
deletes 2 adjacent Pauli spiders.

Q: What if they share neighbours?

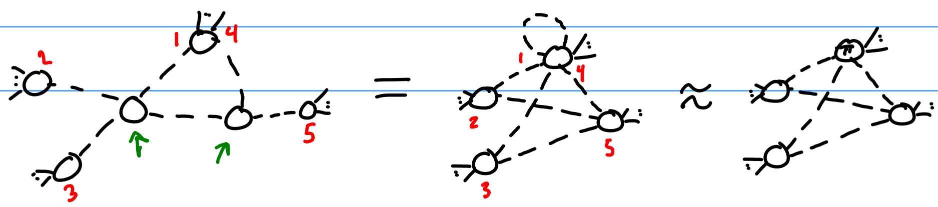




General pivot:



Or, as I prefer to think about it, use the simpler rule, but allow boundary sp's to match twice:



Rewrite strategy (Clifford-simp.)

1. convert to a graph-like diagram
2. apply LC' & P' as long a possible.
3. remove isolated $\{0,\pi\}$-spiders.

Prop1 Clifford-simp terminates for any ZX-diag and removes all interior:

- $\pm\frac{\pi}{2}$ spiders
- pairs of connected $\{0,\pi\}$-spiders

Recall: $m \vdots \boxed{} \vdots n$ is a $2^n \times 2^m$ matrix.

$m = n = 0 \Rightarrow 2^0 \times 2^0 = 1 \times 1$ matrix (a scalar)

Def A scalar ZX-diagram is a ZX-diag w/ no inputs and no outputs.

Cor (to Prop1) There exists a terminating rewrite strategy that removes all spiders from a scalar Clifford diagram.

Pf First apply Clifford-Simp. Then the only spiders left are $\bigcirc$ and $\overset{\pi}{\bigcirc}$. For these:

$$\bigcirc \to 2 \cdot \boxed{} \qquad \overset{\pi}{\bigcirc} \to \emptyset$$

∎

Q: What's left?

A: the scalar factor

$$\boxed{D_0} \to \lambda_1 \cdot \boxed{D_1} \to \lambda_2 \cdot \boxed{D_2} \to \dots \to \lambda_n \cdot \boxed{} = \lambda_n \in \mathbb{C}$$

Application 1 (Efficient) strong simulation of Clifford circuits.

Problem For a circuit C, compute:

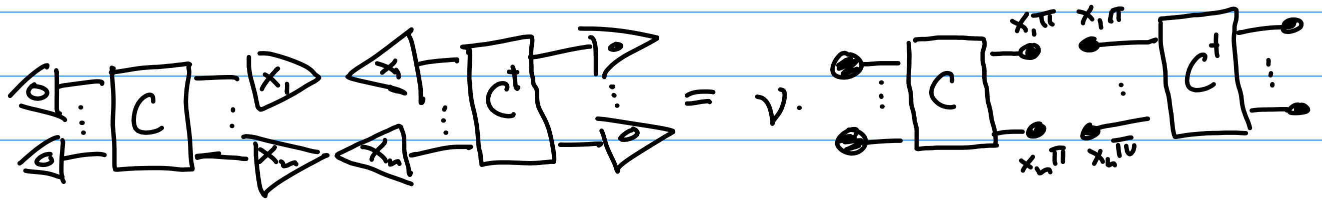
$$(*) \quad \mathrm{Prob}(x_1 \dots x_n \mid |\psi\rangle) \quad \text{where } |\psi\rangle = C|0\dots0\rangle.$$

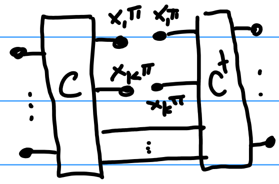
... or more generally, for $k \leq n$, compute the marginal probability:

$$(**) \quad \mathrm{Prob}(x_1 .. x_k \mid |\psi\rangle) = \sum_{x_{k+1}, .. x_n} \mathrm{Prob}(x_1 .. x_n \mid |\psi\rangle)$$

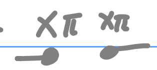
Born rule

$$\mathrm{Prob}(x_1 .. x_n \mid |\psi\rangle) := |\langle x_1 .. x_n | C | 0..0\rangle|^2$$

$=$ 

$\Rightarrow \mathrm{Prob}(x_1 .. x_k \mid |\psi\rangle) = \sum_{x_{k+1} .. x_n} \mathrm{Prob}(x_1 .. x_n \mid |\psi\rangle) = \nu' \cdot$ 

(ZX-diagram for (**))

$\Big(\sum_x$  $= 2 \cdot \sum_x |x\rangle\langle x| = 2I \Big)$

<u>Algorithm</u>: For a circuit C:

1. Let D be the ZX-diagram of $\mathrm{Prob}(x_1 .. x_k | C|0...0\rangle)$.
2. Apply Clifford-simp to get a number.

<u>Prop 1</u> Algorithm terminates in polynomial time (in the # of qubits & gates of C).

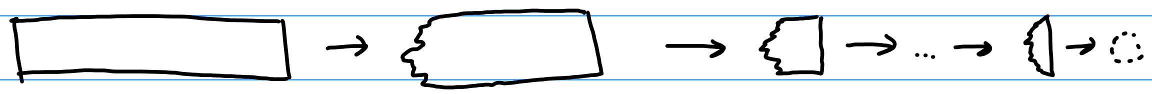
<u>Pf</u> Assume basic diagram operations (add/remove spider/wire) take constant time. If C has n qubits & k gates, D has at most $S := 2\cdot(2n+2k) = 4(n+k)$ spiders. Then:

- Each rewrite removes 1 or 2 spiders, so there are at most $4(n+k)$ steps.
- Each step adds/removes at most $(4(n+k))^2$ edges, so Algorithm 1 performs $O((n+k)^3)$ basic graph operations. $\boxtimes$

<u>Rem</u> this is <u>not</u> optimal. A good choice of LC' and P' steps actually takes $O(n^2 k)$ time.
$\Rightarrow$ if $k \gg n$, this makes a <u>big</u> difference!

<u>IDEA</u>:

1. Avoid big spiders: 
2. Apply LC' & P' from left-to-right:

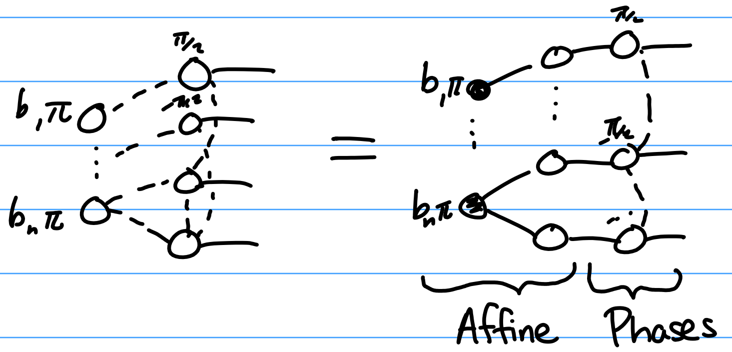


$\Rightarrow$ each step involves at most $O(n)$ spiders (hence $O(n^2)$ wires)

"Affine w/ Phases"

Def A graph-like ZX-diagram is in AP-form if all interior spiders:
- have phase $\in 0, \pi$
- are only connected to boundary spiders.

Application: completeness (see Exercises)

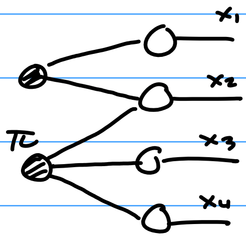


$$= \sum_{\vec{x},\, A\vec{x}=\vec{b}} e^{i\frac{\pi}{2}\cdot\phi} |\vec{x}\rangle$$

$A = \{\vec{x} \mid A\vec{x}=\vec{b}\}$ is an affine subspace of $\mathbb{F}_2^n$.

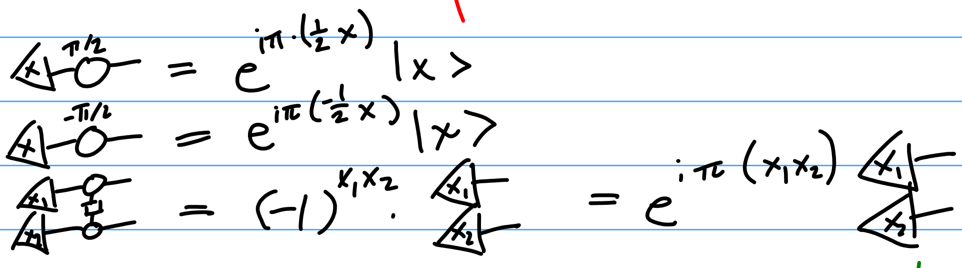
:= a solution to a set of linear eqns, e.g:

$$A = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \middle| \begin{matrix} x_1 \oplus x_2 = 0 \\ x_2 \oplus x_3 \oplus x_4 = 1 \end{matrix} \right\} \iff$$

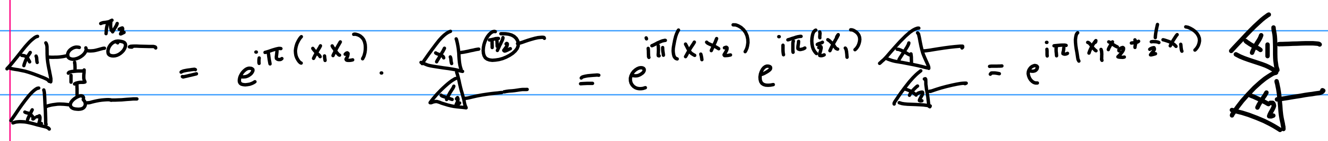


$$\approx \sum_{\vec{x},\, A\vec{x}=\vec{b}} |\vec{x}\rangle$$

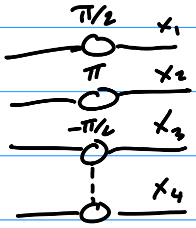
ϕ is a phase polynomial



phase polynomial



$U|\vec{x}\rangle = e^{i\pi\phi} \; |\vec{x}\rangle$ where $\phi = \frac{1}{2}x_1 - \frac{1}{2}x_3 + x_2 + x_3x_4$

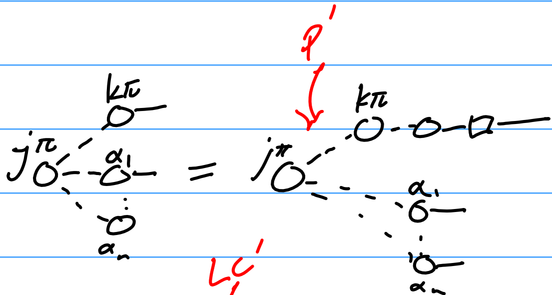
$U =$ 

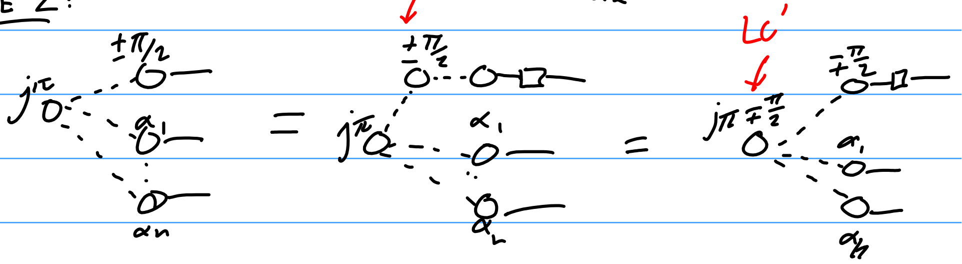
Def A ZX-diagram is in graph-state w/ local Clifford (GSLC) form if it has

* all Z spiders, fused as much as possible
* all spiders are connected to exactly 1 input (possibly via a 1-qubit Clifford unitary)

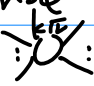
AP → GSLC :

2 cases :

CASE 1 : 

CASE 2 : 

Clifford diagram → AP-form → GSLC

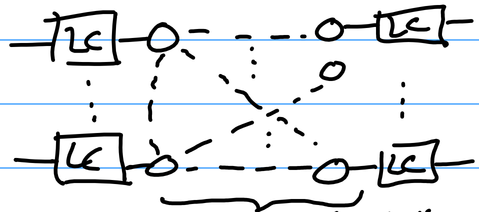
* only internal spiders are 
* no internal spiders.

Application: Efficient synthesis of Clifford circuits.

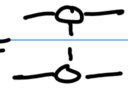
Algorithm 2 (Clifford re-synthesis)

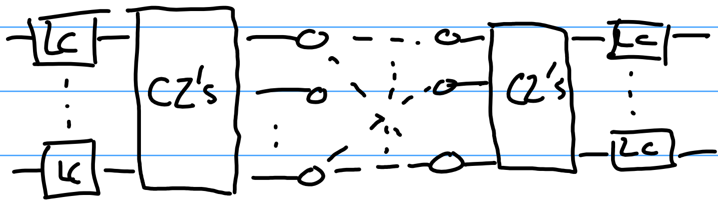
1. For an n-qubit Clifford circuit C, translate to ZX-diagram D.
2. Compute GSLC form.
3. Write as:

$O(n)$

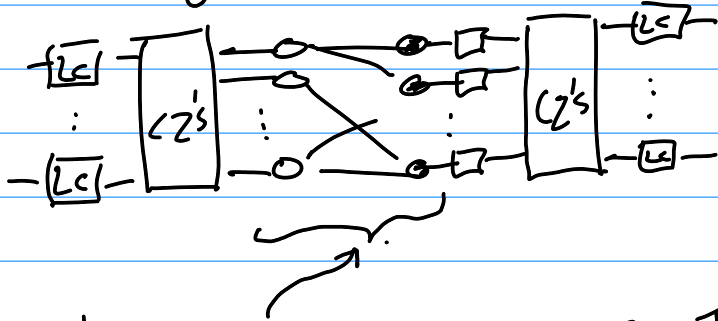


graph "state"

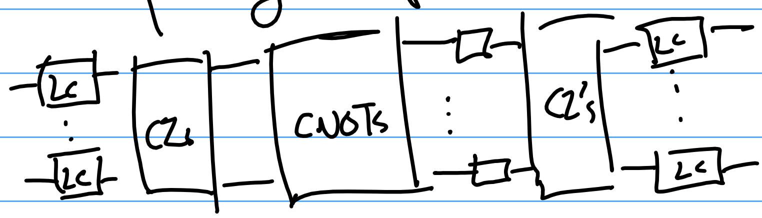
4. unfuse CZ gates =  $O(n^2)$



5. colour-change:



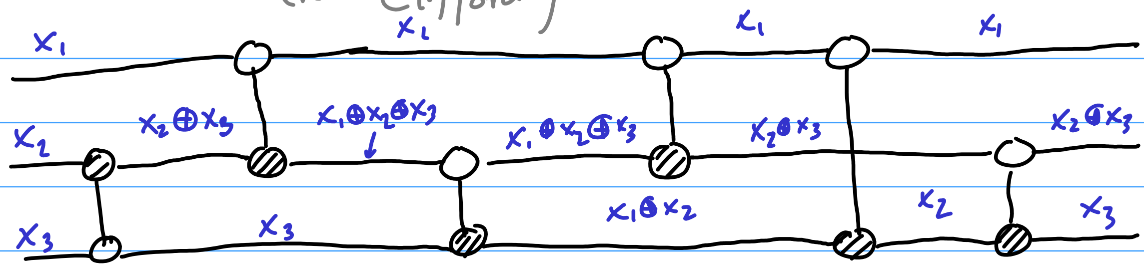
4. extract parity map as CNOTs $O(n^2)$



PROP Any Clifford circuit can be written w/ at most $O(n^2)$ gates!

CNOT + phase circuits

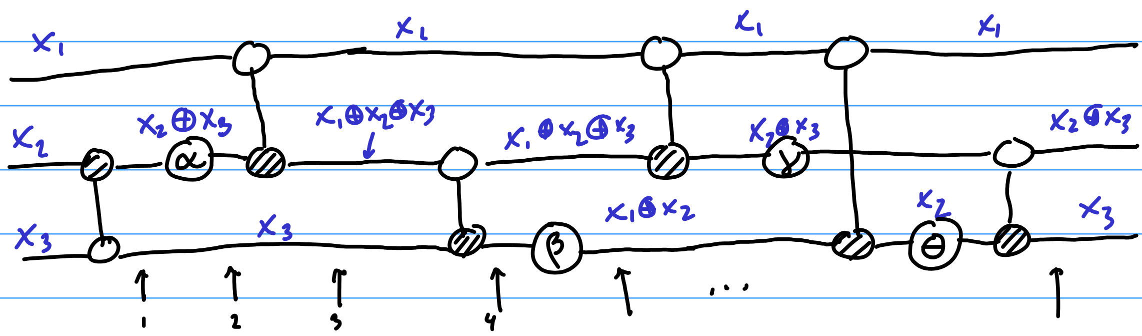
(non-Clifford)

$U =$ 

$$U|x_1 x_2 x_3\rangle = |x_1, x_2 \oplus x_3, x_3\rangle$$

Q: What happens when we add phase gates?

$$Z[\alpha] :: |x\rangle \mapsto e^{i\alpha \cdot x}|x\rangle$$

$V =$ 

$$|x_1 x_2 x_3\rangle \overset{1}{\mapsto} |x_1, x_2 \oplus x_3, x_3\rangle$$

$$\overset{2}{\mapsto} e^{i\alpha \cdot (x_2 \oplus x_3)} |x_1, x_2 \oplus x_3, x_3\rangle$$

$$\overset{3}{\mapsto} e^{i\alpha (x_2 \oplus x_3)} |x_1, x_1 \oplus x_2 \oplus x_3, x_3\rangle$$

$$\overset{4}{\mapsto} e^{i\alpha \cdot (x_2 \oplus x_3)} |x_1, x_1 \oplus x_2 \oplus x_3, x_1 \oplus x_2\rangle$$

$$\mapsto e^{i[\alpha \cdot (x_2 \oplus x_3) + \beta \cdot (x_1 \oplus x_2)]} |x_1, x_1 \oplus x_2 \oplus x_3, x_1 \oplus x_2\rangle$$

$$\mapsto \ldots$$

$$\mapsto e^{i[\alpha \cdot (x_2 \oplus x_3) + \beta \cdot (x_1 \oplus x_2) + \gamma \cdot (x_2 \oplus x_3) + \theta \cdot x_2]} |x_1, x_2 \oplus x_3, x_3\rangle$$

Prop Any CNOT+phase circuit describes a unitary of the form:

$$U :: |\vec{x}\rangle \mapsto e^{i\phi(\vec{x})} |L\vec{x}\rangle$$

phase polynomial   parity matrix.

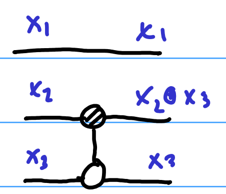
From the example above: $L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ and

$$\phi(x_1, x_2, x_3) = (\alpha + \gamma)\cdot(x_2 \oplus x_3) + \beta \cdot (x_1 \oplus x_2) + \theta \cdot x_2$$

phase-folding

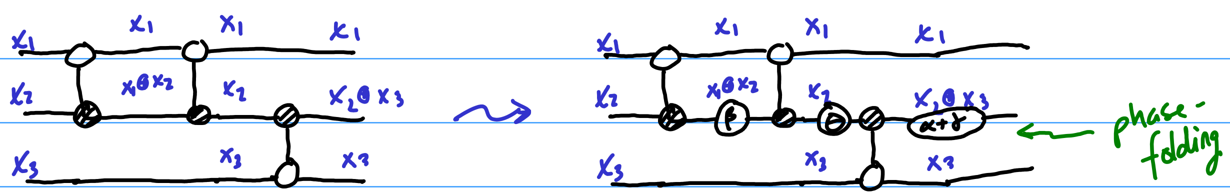
Q: can we re-synthesise a circuit for (L, ϕ)?

For L, we have:



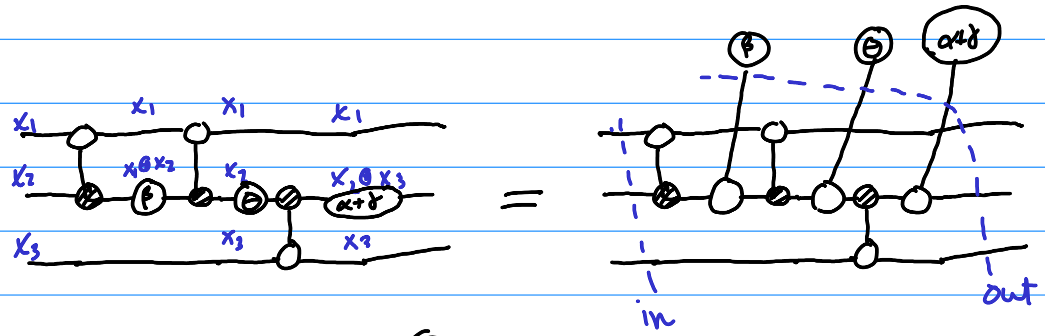
To get ϕ, we need to place Z-phases on wires labelled:
$x_2 \oplus x_3$, $x_1 \oplus x_2$, and x_2

Only $x_1 \oplus x_2$ is missing, so lets (temporarily) create it:

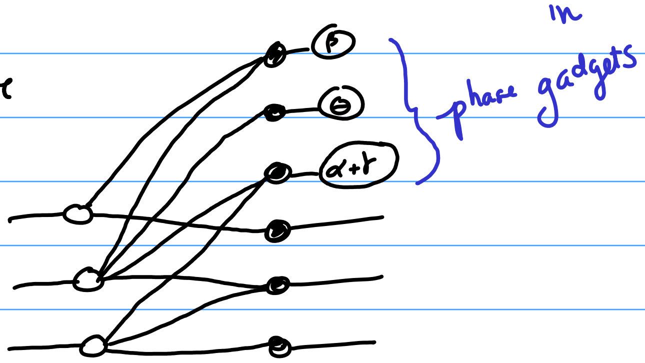


Phase polynomials, graphically (aka. phase gadgets)

Ex



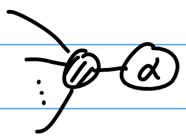
Phase-free simp



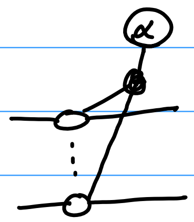
1-legged :

$$:: \ |x\rangle \longmapsto \begin{cases} 1 & \text{if } x=0 \\ e^{i\alpha} & \text{if } x=1 \end{cases} = e^{i\alpha \cdot x}$$

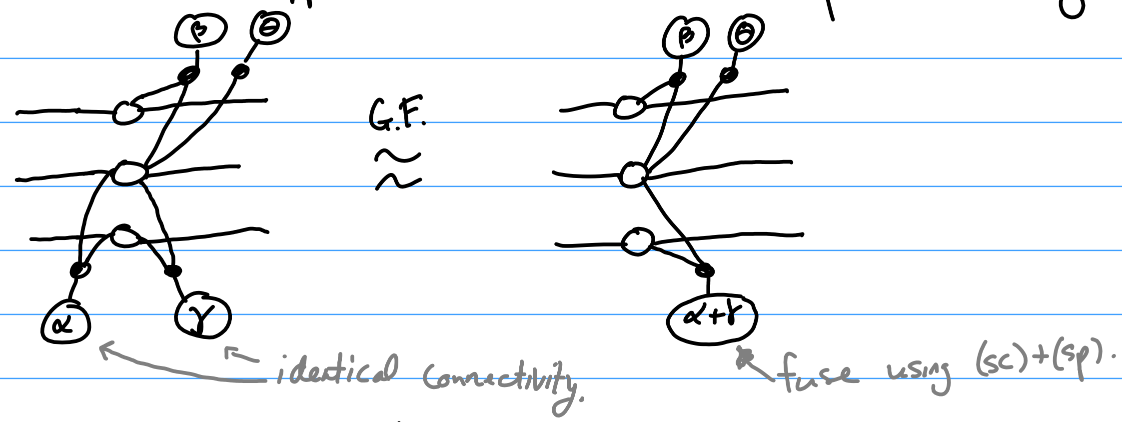
k-legged phase gadget:

$\sqrt{2}^{(k-1)}$  $:: \ |x_1 .. x_k\rangle \longmapsto e^{i\alpha \cdot (x_1 \oplus .. \oplus x_k)}$

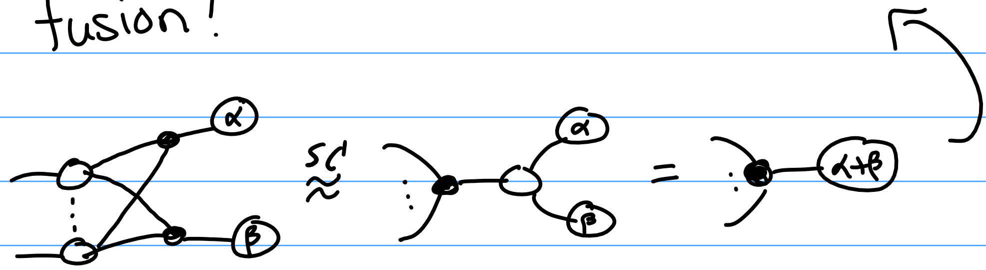
In a diagonal unitary :

$\sqrt{2}^{(k-1)}$  $:: \ |x_1 .. x_k\rangle \longmapsto e^{i\alpha \cdot (x_1 .. x_k)} |x_1 .. x_k\rangle$

Q: What happens when there is phase folding?



A: Gadget fusion!



Algorithm: CNOT + phase optimisation. (PNF = Parity NF)

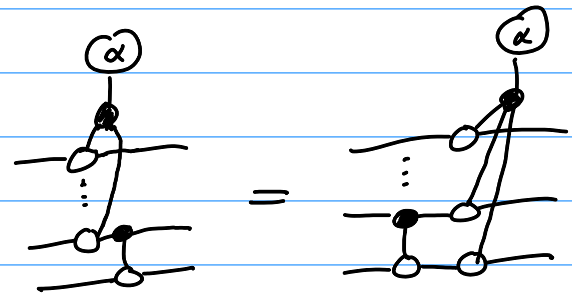
1. unfuse phases and treat as outputs.
2. Compute PNF of phase-free part.
3. perform gadget fusion (+ and other phase-poly reductions!)

?? → 4. extract a CNOT + phase circuit.

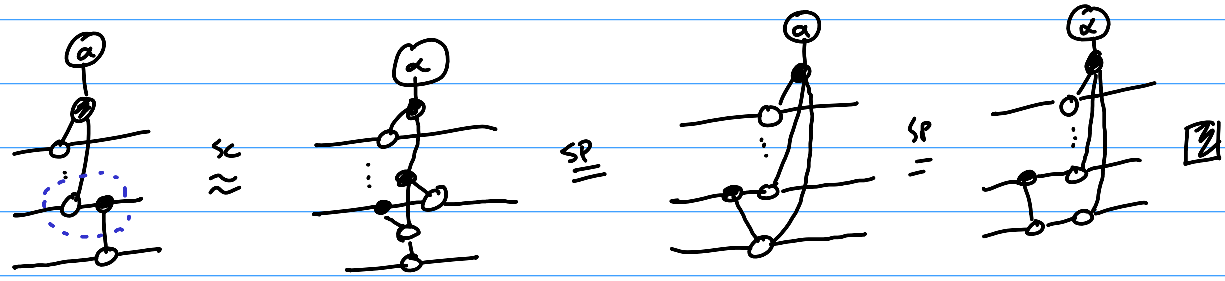
There are choices for step 4.

Naïve approach: "CNOT ladders"

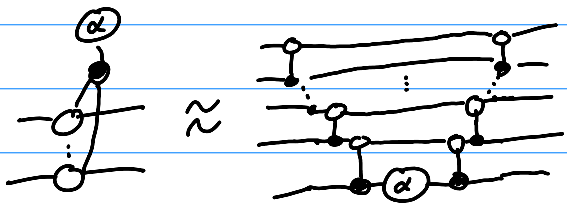
Prop



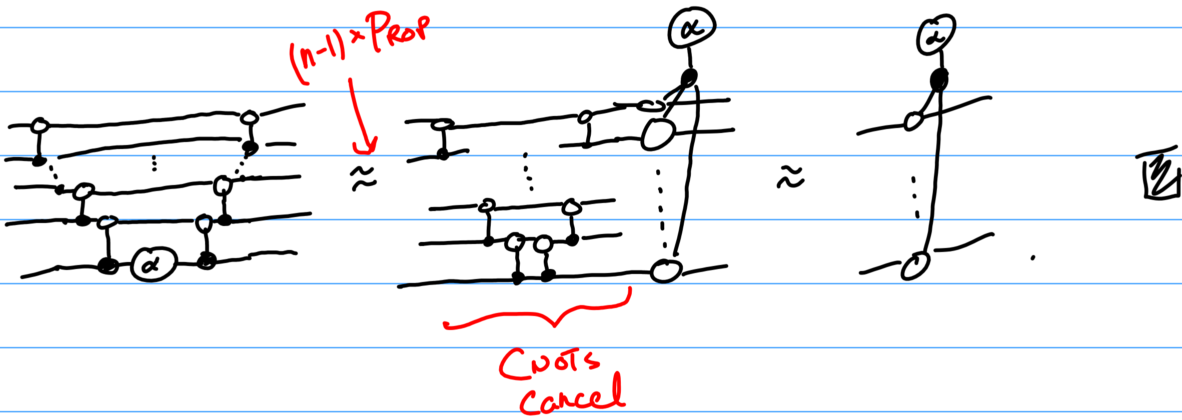
Pf



CoR

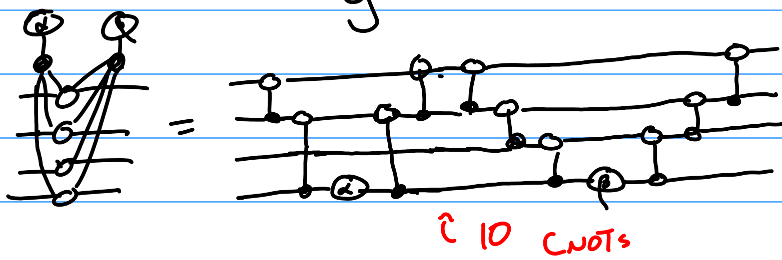


Pf

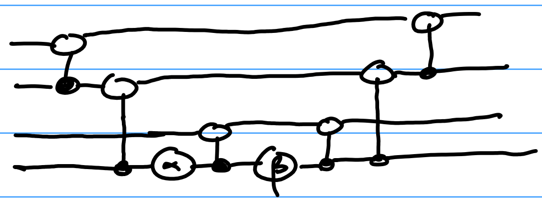


Naïve extraction:
1. unfuse a phase gadget & replace using CoR 1.
2. repeat until no phase gadgets
3. synthesise CNOT circuit from phase-free diag.

* Lots of wasted CNOT gates! e.g.



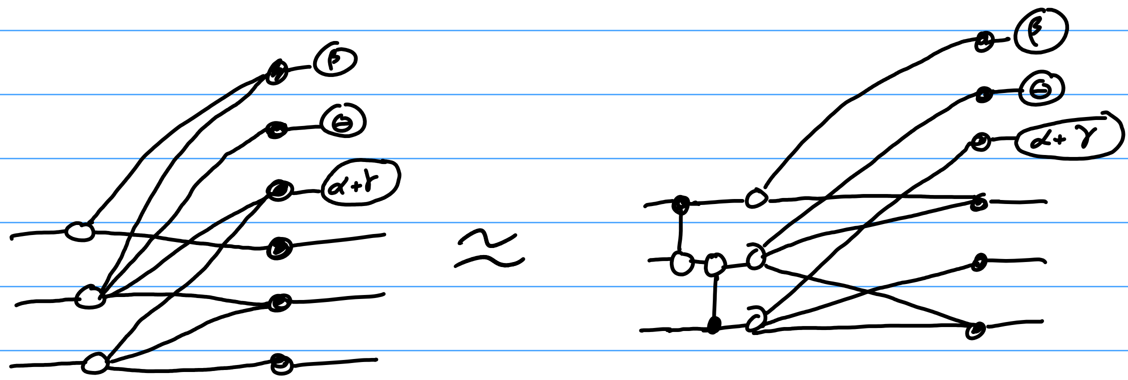
Vs.



6 CNOTs

Better extraction

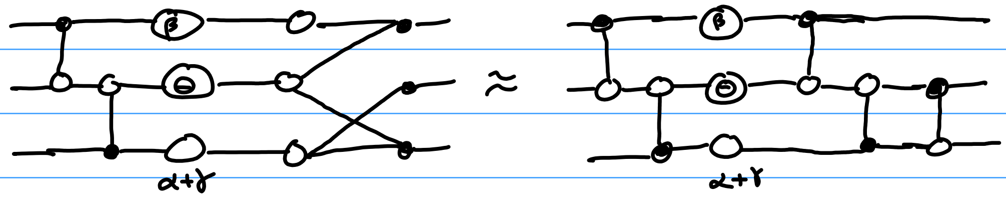
1. write an "extended biadjacency matrix"
2. identify a set of k linearly independent rows
3. reduce each row to a unit vector with column ops.
4. "extract phases" and repeat.



$\underset{10}{\overset{SF}{=}}$ $*$

$$\begin{array}{r}\text{gadgets}\left\{\begin{array}{c}\\ \\ \\ \end{array}\right.\\ \text{outputs}\left\{\begin{array}{c}\\ \\ \\ \end{array}\right.\end{array}\begin{pmatrix}1&1&0\\0&1&0\\0&1&1\\ \hline 1&0&0\\0&1&1\\0&0&1\end{pmatrix}\xrightarrow{C_2=C_2+C_1}\begin{pmatrix}1&0&0\\0&1&0\\0&1&1\\1&1&0\\0&1&1\\0&0&1\end{pmatrix}\xrightarrow{C_2=C_2+C_3}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\\ \hline 1&1&0\\0&0&1\\0&1&1\end{pmatrix}$$

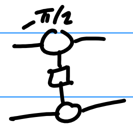
$* =$

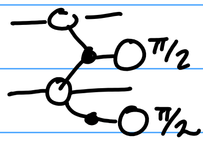


$$\begin{pmatrix}1&1&0\\0&0&1\\0&1&1\end{pmatrix}\xrightarrow{C_2=C_2+C_1}\begin{pmatrix}1&0&0\\0&0&1\\0&1&1\end{pmatrix}\xrightarrow{C_2=C_2+C_3}\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix}\xrightarrow{C_3=C_3+C_2}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$$

High-level gates

We've seen 2 kinds of phase polynomials:

"Multilinear" form, e.g.  :: $|x,y\rangle \mapsto e^{i\pi\cdot(-\frac{1}{2}x + x\cdot y)}|x,y\rangle$

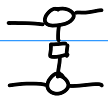
"XOR"/ Phase-gadget form, e.g.  :: $|x,y\rangle \mapsto e^{i\pi\cdot(x\oplus y + y)}$

These two forms are related:

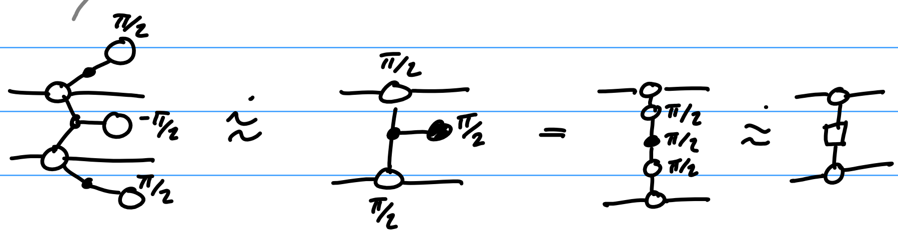
$$\underset{\text{XOR}}{x \oplus y} = x \underset{\text{plus}}{+} y \underbrace{- 2xy}_{\text{"correction"}} \qquad (x,y \in \{0,1\})$$

$$-2xy = x\oplus y - x - y$$

$$\Rightarrow xy = \tfrac{1}{2}(x + y - x\oplus y)$$

 :: $|xy\rangle \mapsto e^{i\pi\cdot(xy)}|xy\rangle$

$$= e^{i\pi(\frac{1}{2}x + \frac{1}{2}y - \frac{1}{2}x\oplus y)}|xy\rangle$$

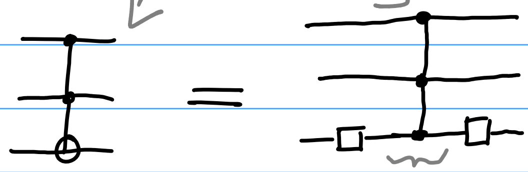


Some gates are easy to write in multilinear form.

Consider :

Q. Circuit Notation

Toffoli:



$|x,y,z\rangle$

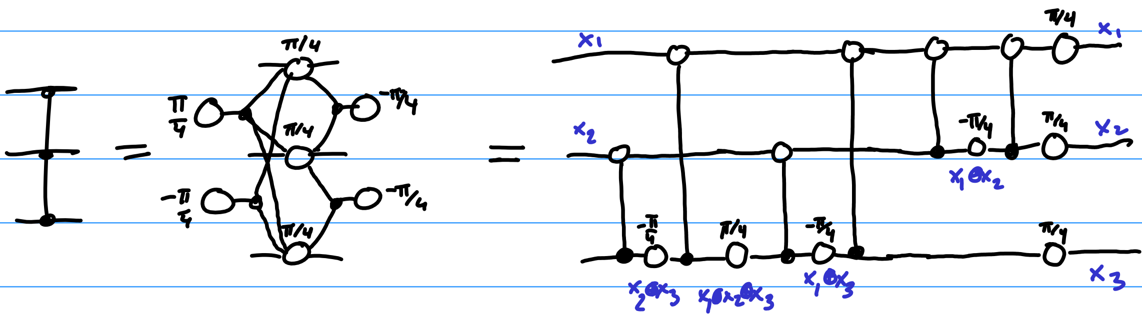
$\mapsto |x,y,(x\cdot y)\oplus z\rangle$

CCZ diagonal

$$CCZ|x_1x_2x_3\rangle = \begin{cases} |x_1x_2x_3\rangle & \text{if } x_1\cdot x_2 = 0 \\ |x_1x_2\rangle \otimes \underbrace{Z|x_3\rangle}_{(-1)^{x_3}|x_3\rangle} & \text{if } x_1\cdot x_2 = 1 \end{cases}$$

$$= (-1)^{x_1x_2x_3}|x_1x_2x_3\rangle = e^{i\pi\cdot x_1x_2x_3}|x_1x_2x_3\rangle$$

$$\begin{aligned} x_1(x_2x_3) &= \tfrac{1}{2}x_1\cdot(x_2+x_3-x_2\oplus x_3) \\ &= \tfrac{1}{2}(x_1x_2+x_1x_3-x_1(x_2\oplus x_3)) \\ &= \tfrac{1}{4}(x_1+x_2-x_1\oplus x_2+\cancel{x_1}+x_3-x_1\oplus x_3-\cancel{x_1}-x_2\oplus x_3+x_1\oplus x_2\oplus x_3) \\ &= \tfrac{1}{4}(x_1+x_2+x_3-x_1\oplus x_2-x_1\oplus x_3-x_2\oplus x_3+x_1\oplus x_2\oplus x_3) \end{aligned}$$



Translation of CCZ into XOR form is a special case of discrete Fourier transform.

dot product, i.e. parities.

Prop For any function $\phi : \mathbb{F}_2^n \to \mathbb{R}$,

$$\phi(\vec{x}) = \frac{-1}{2^{n-1}} \sum_{\vec{y}} \tilde{\alpha}_{\vec{y}} \, (\vec{x} \cdot \vec{y})$$

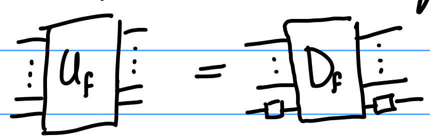
where $\tilde{\alpha}_{\vec{y}} = \frac{-1}{2^{n-1}} \sum_{\vec{z}} (-1)^{\vec{y}\cdot\vec{z}} \cdot \phi(\vec{z})$ are the Fourier coefficients.

In the CCZ case, taking the Fourier xform of $\phi(\vec{x}) = \begin{cases} 1 & \text{if } x_1 x_2 x_3 = 1 \\ \emptyset & \text{o.w.} \end{cases}$

gives: $\begin{cases} \tilde{\alpha}_{100} = \tilde{\alpha}_{010} = \tilde{\alpha}_{001} = \frac{1}{4} \\ \tilde{\alpha}_{110} = \tilde{\alpha}_{101} = \tilde{\alpha}_{011} = -\frac{1}{4} \\ \tilde{\alpha}_{111} = \frac{1}{4}. \end{cases}$

This gives a general strategy for synthesising classical oracles $f : \{0,1\}^n \to \{0,1\}$

1. Let $U_f|\vec{x}, y\rangle := |\vec{x}, f(\vec{x}) \oplus y\rangle$



$$D_f|\vec{x}, y\rangle := e^{i\pi \cdot \phi} |\vec{x}, y\rangle \quad \text{where } \phi(\vec{x}, y) = f(\vec{x}) \cdot y$$

2. Compute Fourier coeffs of ϕ.
3. Synthesise D_f as CNOT+Phase circuit.

Path sums

We know how to deal w/ Cliffords

& w/ CNOT + Phase

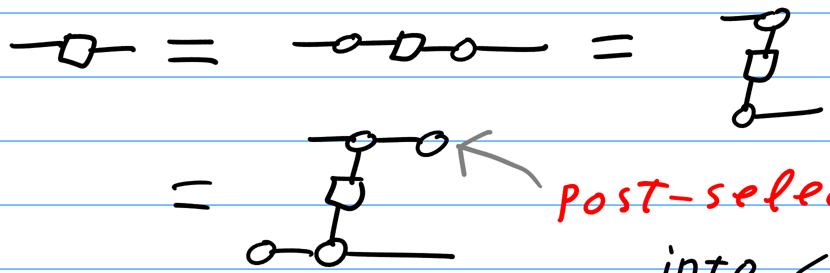
Combination: CNOT + Phase + Hadamard

$Z[\alpha]$ has $S := Z[\frac{\pi}{2}]$
as special case

$\Rightarrow$ This is a universal gateset

$\Rightarrow$ Do not expect efficient rewriting

One approach: Path sums



post-selection
into $\langle +|$

ancilla in $|+\rangle$ state

$|+\rangle = |0\rangle + |1\rangle = \sum_y |+\rangle$ $\qquad$ $\langle +| :: |x\rangle \mapsto 1$

deleting variable

$CNOT :: |x_1, x_2\rangle \mapsto |x_1, x_1 \oplus x_2\rangle$ $\quad$ updating variable

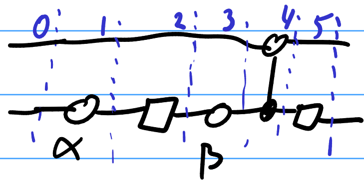
$Z[\alpha] :: |x\rangle \mapsto e^{i\alpha}|x\rangle$ $\quad$ updating phase polynomial

$-\square- :: |x\rangle \mapsto \sum_y (-1)^{x \cdot y}|y\rangle$

Path variable

Ex:

Step 0: $|x_1, x_2\rangle$



$$\overset{1}{\Longrightarrow} e^{i\alpha x_1}|x_1, x_2\rangle$$

$$\overset{2}{\Longrightarrow} \sum_{y} e^{i\alpha x_1}(-1)^{y\cdot x_1}|x_1, y\rangle$$

$$\overset{3}{\Longrightarrow} \sum_{y} e^{i\alpha x_1 + \beta y}(-1)^{y\cdot x_1}|x_1, y\rangle$$

$$\overset{4}{\Longrightarrow} \sum_{y} e^{i\alpha x_1 + \beta y}(-1)^{y\cdot x_1}|x_1, x_1\oplus y\rangle$$

$$\overset{5}{\Longrightarrow} \sum_{y,z} e^{i\alpha x_1 + \beta y}(-1)^{y\cdot x_1 + (x_1\oplus y)\cdot z}|x_1, z\rangle$$

Hadamards create new paths, "branches", these interfere via phases

Classical simulation method: Just sum all the branches

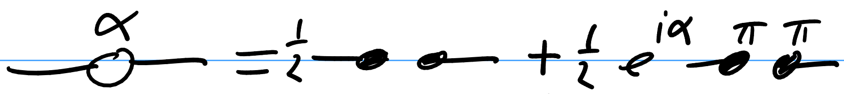
cost: $\mathcal{O}(n \cdot k \cdot 2^h)$ # Hadamards

qubits # gates

Power of Q. computation = Hadamards?

=> get bunch of 1-qubit circuits

cost: $\mathcal{O}(n \cdot k \cdot 2^c)$ # CNOTs



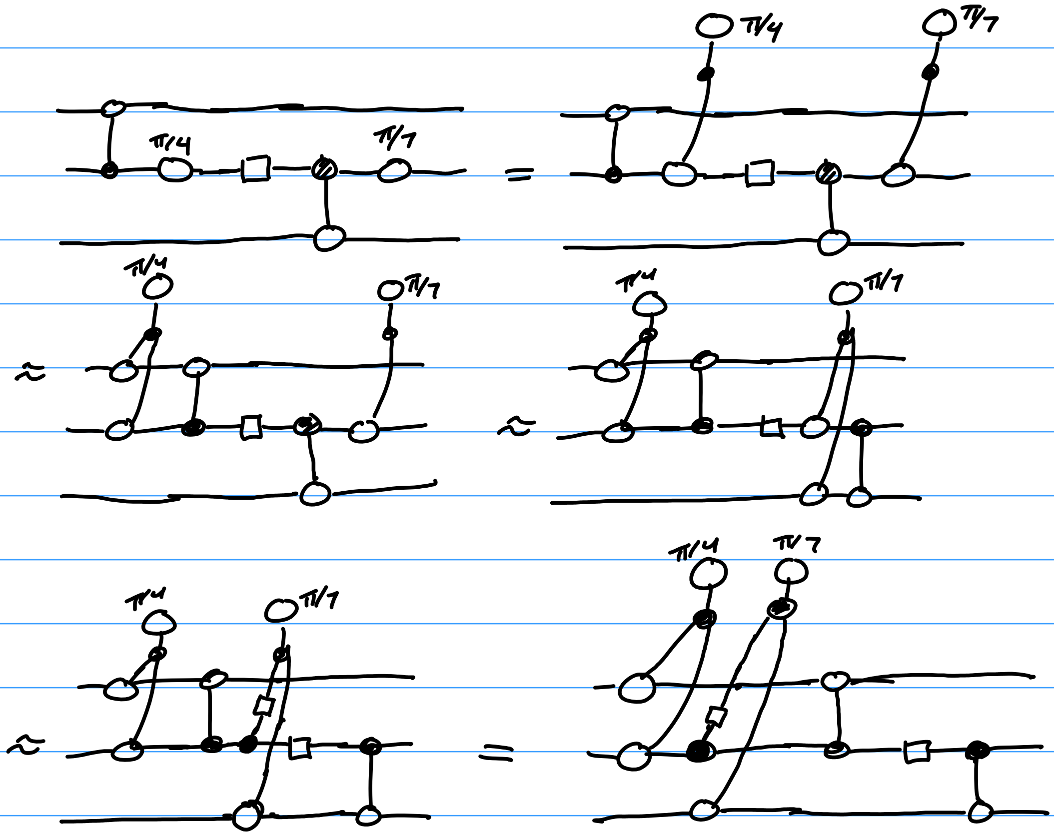
=> get Clifford circuits

cost: $\mathcal{O}(n^2 \cdot k \cdot 2^t)$ # non-Clifford phases

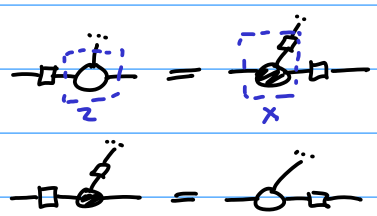
Pauli Gadgets

Clifford + Phase is a universal family.

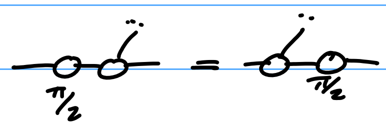
Q: Can we move all the non-Clifford phases out?

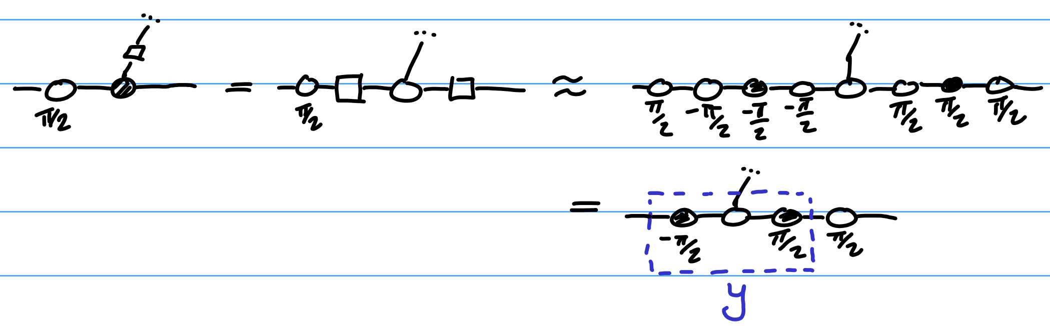


H gates:

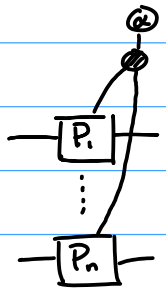


S gates:

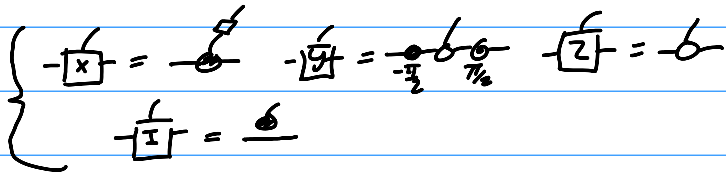




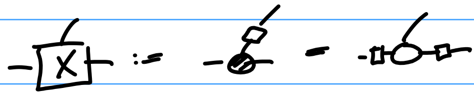
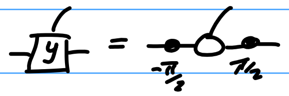
Prop For $\vec{P} = P_1 \otimes .. \otimes P_n$ with $P_i \in \{I, X, Y, Z\}$ the map:

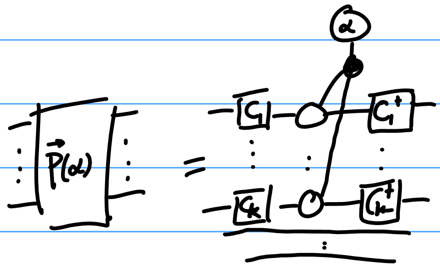


where:



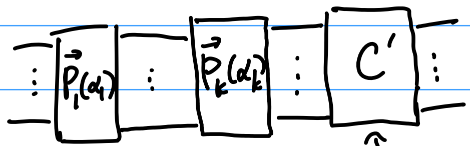
is unitary. It is called the <u>Pauli gadget</u> $\vec{P}(\alpha)$.

<u>Pf</u> Note  and . So



for Cliff. unitaries C_i. Since phase gadgets are unitary, so is $\vec{P}(\alpha)$. ☒

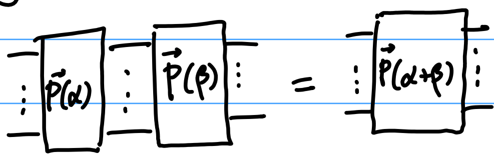
<u>Thm</u> Any Clifford+phase circuit can be written as:

$C =$ 

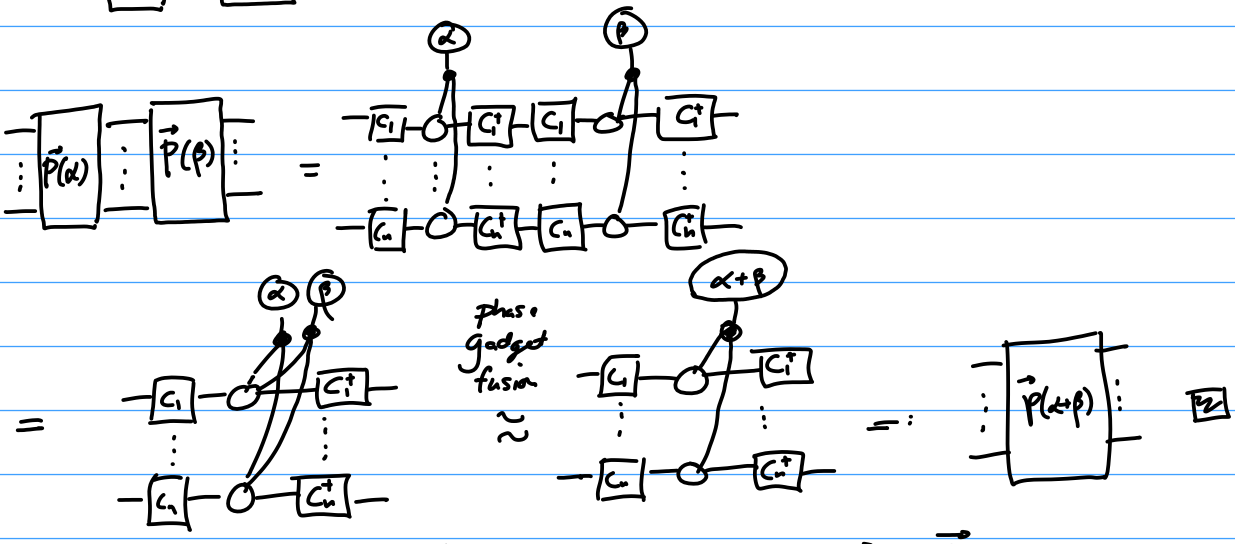
Clifford

<u>Pf</u> (Idea) • Show Pauli gadgets commute past all Clifford gates.
• Move phases out of C, one at a time. ☒

Prop (Pauli gadget fusion.)



Pf



Prop For Paulis $\vec{P}, \vec{Q}$ if $\vec{P}\vec{Q} = \vec{Q}\vec{P}$, then $\vec{P}(\alpha)\vec{Q}(\beta) = \vec{Q}(\beta)\vec{P}(\alpha)$.

Pf Exercise/book (Hint: it's complementarity!)

Algorithm Pauli "phase folding".

1. Compute Pauli gadget form of a circuit.
2. Commute PG's and combine phases where possible.
3. Merge PG's with Clifford phases into the Clifford part.
4. Repeat until no more reductions.
5. Extract circuit.*

* like with CNOT+phase, there are many options.

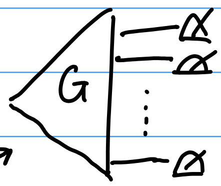
Measurement-based quantum computing (MBQC)

:= QC where measurements make up most of the computation.

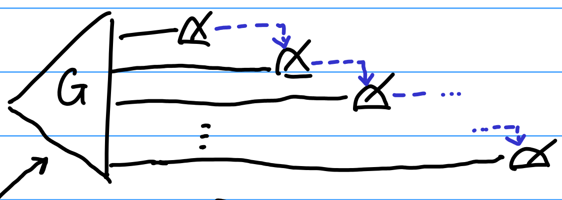
The "code" of MBQC is a measurement pattern:
:= measurement choices + classical control (feed-forward)

Several models:
- (gate teleportation)
- one-way model *
- hypergraph MBQC
- fault tolerant QC
  - lattice surgery (*)
  - topological FTQC
- ...

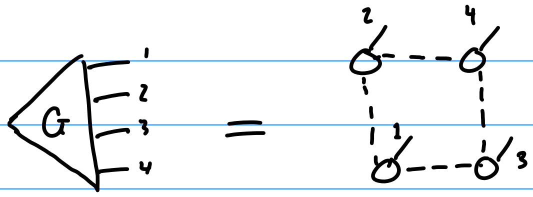
One-way model of MBQC (Raussendorf/Briegel 2001)



graph state   single-qubit measurements

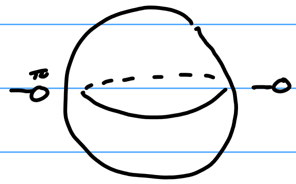


graph state   single-qubit measurements with feed-forward

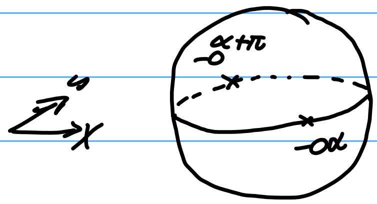


Single-Qubit Measurements:

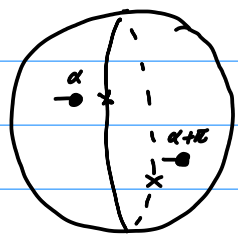
X-measurement: $\left\{ -\overset{k\pi}{\circ} \right\}_{k=0,1}$



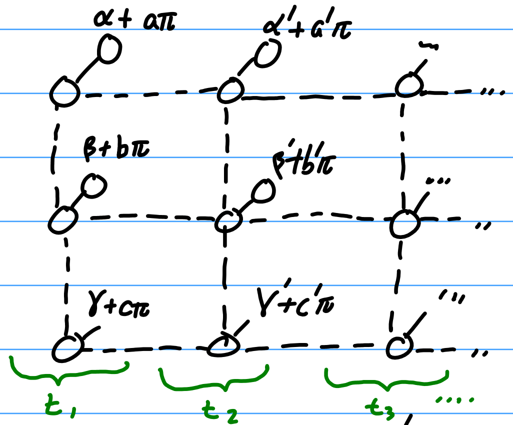
More generally: XY-plane measurements: $\left\{ -\overset{\alpha+k\pi}{\circ} \right\}_{k=0,1}$



Similarly, YZ-plane measurements: $\left\{ -\overset{\alpha+k\pi}{\bullet} \right\}_{k=0,1}$



(Z-measurements $\Rightarrow \alpha = 0$)



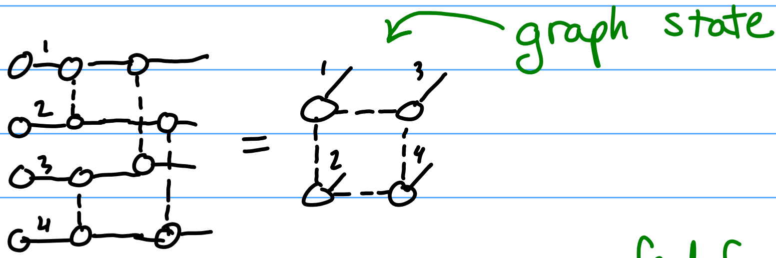
Feed-forward : $\alpha' = \alpha'(a,b,c)$ ← fn of (earlier) measurement outcomes. (a.k.a. signals)
$\beta' = \beta'(a,b,c)$
"""" ""

Def A measurement pattern for the one-way model consists of a sequence of instructions:

* $N_j :=$ ○ ^{j} — prepare a new qubit in $|+\rangle$
* $E_{jk} :=$ — entangle qubits j & k
* $M_j^\alpha := \{ \overset{\alpha + s_j\pi}{—○} \}_{s_j \in 0,1}$ — measure qubit j in XY plane
  * store result in signal $s_j \in \{0,1\}$
  * $\alpha = \alpha(s_{k_1}, s_{k_2}, \ldots)$
* $M_j^{YZ,\alpha} := \{ \overset{\alpha + s_j\pi}{—●} \}_{s_j \in 0,1}$ — " " " " YZ plane "
* $M_j^{XZ,\alpha} := \{ \overset{-\pi/2}{—●}\overset{s_j\pi}{—○} \}_{s_j \in 0,1}$ — " " " " XZ plane "
* $Z_j^b := \overset{b\pi}{—○—}$, $X_j^b := \overset{b\pi}{—●—}$ — perform Pauli corrections, where
  * $b = b(s_{k_1}, s_{k_2}, \ldots)$

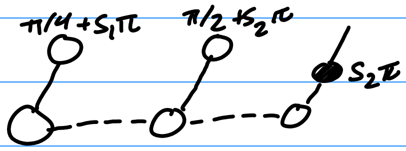
nodes edges

$$P := N_1; N_2; N_3; N_4; E_{12}; E_{34}; E_{13}; E_{24}$$



feed forward

$$Q := N_1; N_2; N_3; E_{12}; E_{23}; M_1^{\pi/4}; M_2^{\pi/2}; X_3^{s_2}$$



<u>Def</u> A measurement pattern is:

* <u>runnable</u> if all angles / corrections are fns of <u>past</u> measurement outcomes.

  OK

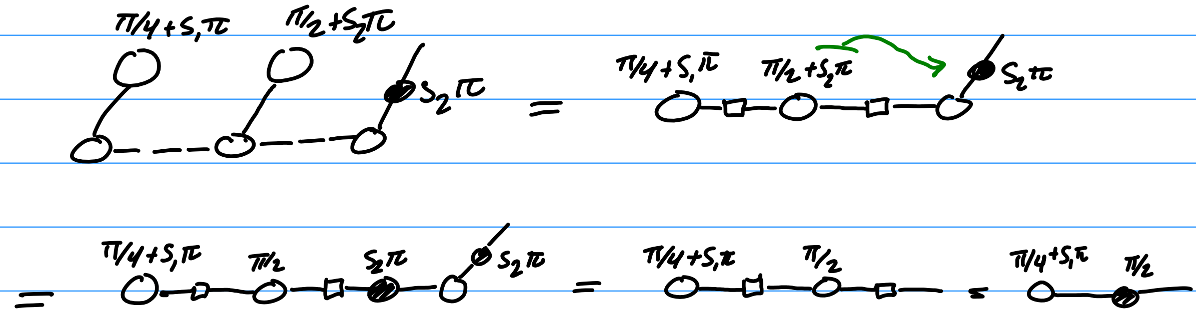
  $$M_j^{\alpha}; \ldots; Z_k^{s_j}$$

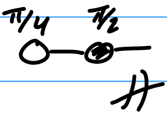
  BAD

  $$Z_k^{s_j}; \ldots; M_j^{\alpha}$$

* <u>deterministic</u> if all choices of measurement outcomes give the same map (up to scalars)

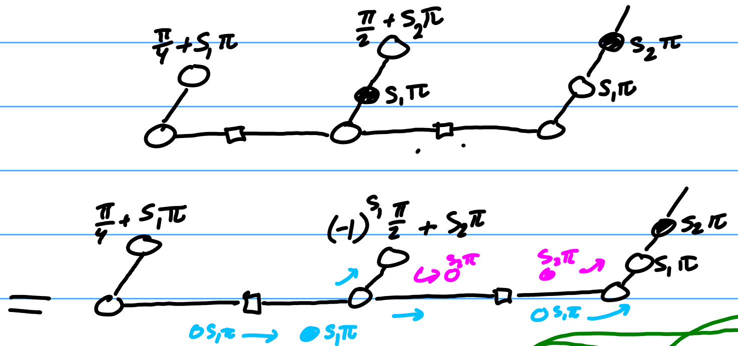
Q: runnable? ✓ deterministic ✗



$s_1 = 0 \Rightarrow * =$ 

$s_1 = 1 \Rightarrow * =$ 

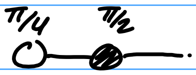
$Q' := N_1; N_2; N_3; E_{12}; E_{23}; M_1^{\frac{\pi}{4}}; X_2^{s_1}; Z_3^{s_1}; X_3^{s_2}$



$= **$

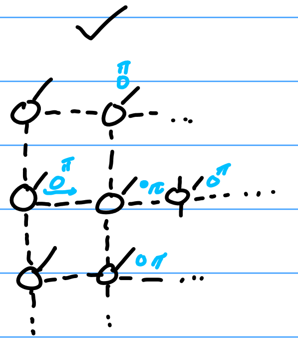
$Q' = Q'' := N_1; N_2; N_3; E_{12}; E_{23}; M_1^{\frac{\pi}{4}}; M_2^{(-1)^{s_1}\frac{\pi}{2}}; Z_3^{s_1}; X_3^{s_2}$

Q'' : runnable? ✓ deterministic? ✓

$s_1, s_2 \in 0,1 \Rightarrow ** =$ 

Question Can I alway "push" errors forward in time?

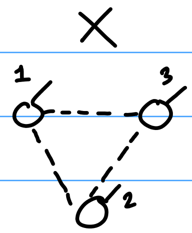
Answer: It depends on the graph.



$t_1 < t_2 < t_3 \ldots$

CLUSTER STATE

(:= graph state shaped like a square lattice)

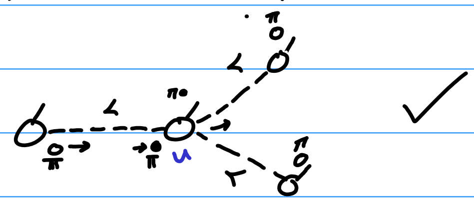


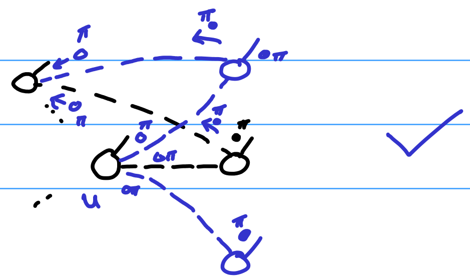
there is no time ordering for qubits $\{1,2,3\}$ that works.

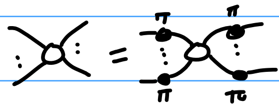
Q: how can we classify which graph states "work"?

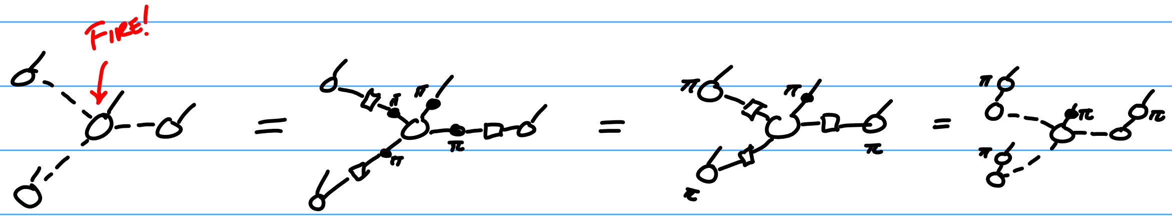
IDEA: 1. Fix a time-ordering $<$: $\begin{cases} \text{past}(u) := \{v \mid v < u\} \\ \text{future}(u) := \{v \mid u < v\} \end{cases}$

2. push errors from u into future(u) (without messing up past(u))





Equivalently, think about "firing" a spider with 

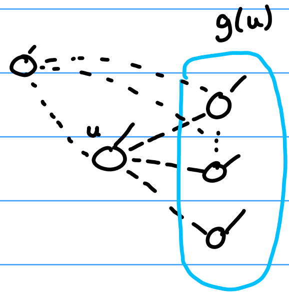


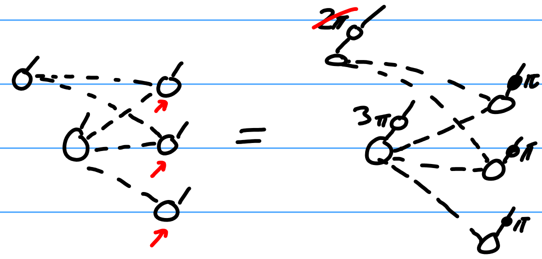
The game: for each u, find a set $g(u)$ that is:

(i) in the future of u

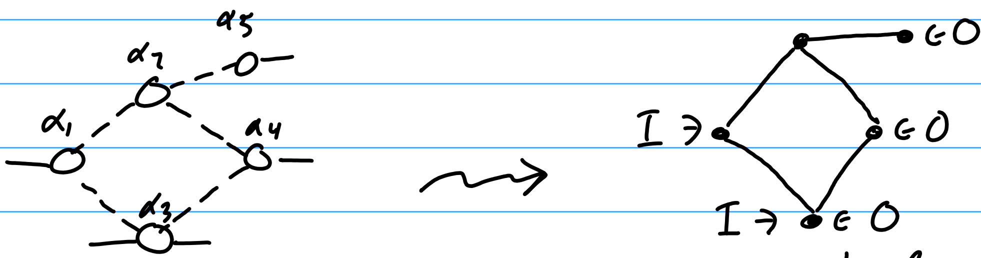
(ii) connected to u an <u>odd</u> number of times

(iii) connected to the past of u an <u>even</u> number of times





Def An open graph is a graph G with a set of inputs $I_G \subseteq V_G$ and outputs $O_G \subseteq V_G$.



graph-like ZX-diag

open graph G

Def An open graph has generalised flow (gflow) if there exists a partial order $\preceq$ on V_G and a function $g: V_G \setminus O_G \to \mathcal{P}(V_G \setminus I_G)$ such that $\forall u$:

(i) $g(u) \subseteq \text{future}(u)$

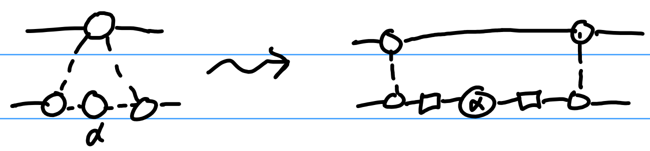
(ii) $g(u)$ connects to u an odd # of times

(iii) $\forall v \in V_G \setminus O_G$. if $v \neq u$, $v \notin \text{future}(u)$ then $g(u)$ connects to v an even # of times.

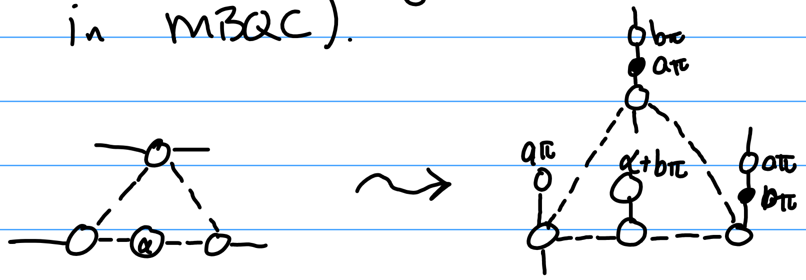
THM (Determinism) For any graph-like ZX-diagram D with gflow, there exists a runnable, deterministic pattern P that implements it.

$\Rightarrow$ There are at least 2 ways that a ZX-diagram can be "run" on a quantum computer:

1. If it can be transformed into a circuit.



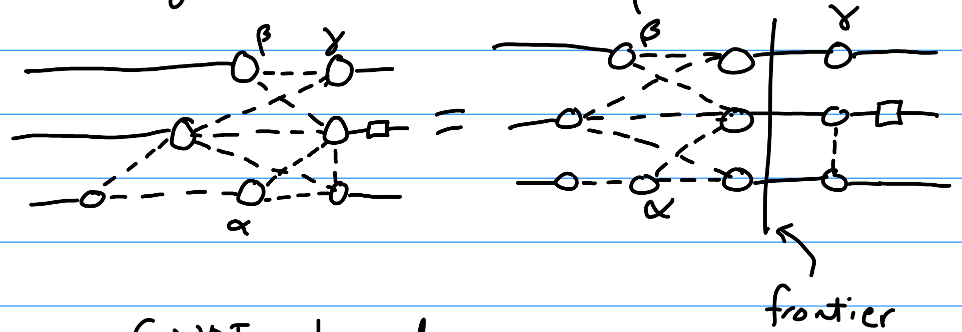
2. If it has gflow (hence can be implemented in MBQC).



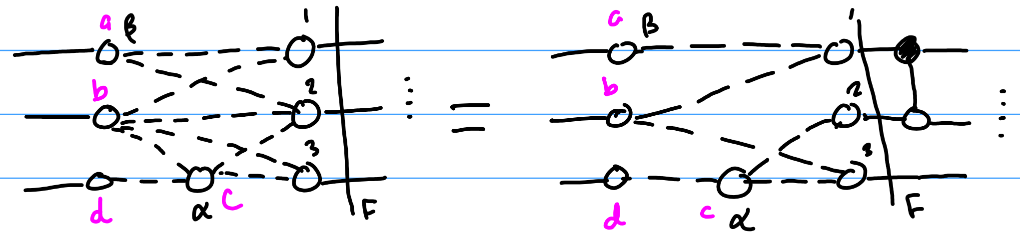
Now: 2. $\Rightarrow$ 1. (circuit extraction)

Algorithm (Circuit Extraction)

1. unfuse gates as much as possible:



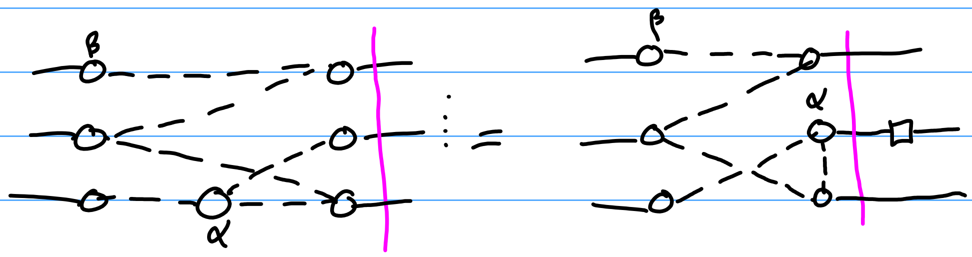
2. use CNOTs to do row operations until we get an "extractible" spider (= unit-vector row)

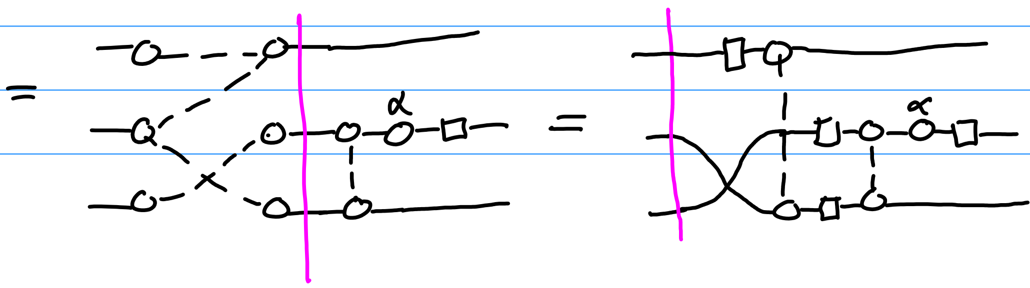


$$\begin{array}{c c} & \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} \end{array} \xrightarrow{R_2 = R_2 + R_1} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} \end{array}$$

extractible

3. Repeat 1 & 2 until nothing is left of the frontier.

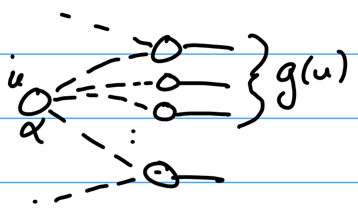




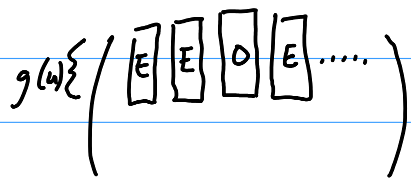
THM If a ZX-diagram has gflow, CIRCUIT EXTRACTION terminates with a quantum circuit.

Pf Step 1 never adds spiders to the left of the frontier, so s.t.s. Step 2 always removes a spider.

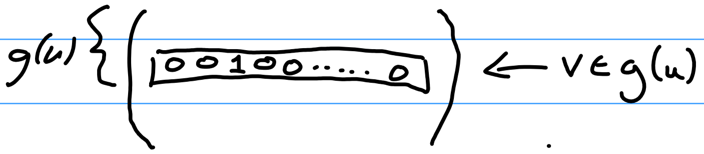
Take a maximal non-output u, w.r.t $\prec$. Then $g(u) \subseteq \text{future}(u)$ must be all outputs:



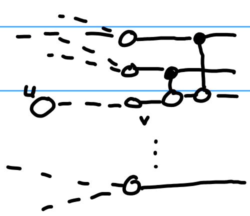
By gflow, the only node connected an odd # of times to $g(u)$ is u.



If we add all the rows to a single row, then we get:



So, doing CNOTs ctrl'ed on a single $v \in g(u)$ to all other $v' \in g(u)$ gives:

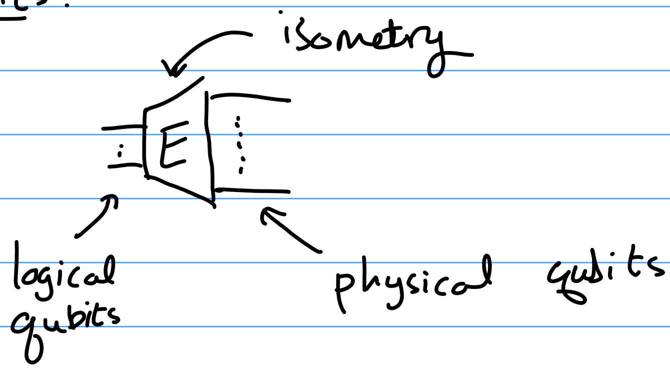


← u becomes extractible

10

Extract & make u an output. The result still <u>has</u> <u>gflow</u> and there is one fewer spider left of the frontier. $\boxtimes$

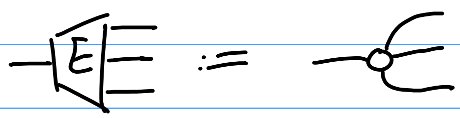
<u>Quantum error correction</u> works by encoding some <u>logical</u> <u>qubits</u> into a space of (more) <u>physical</u> <u>qubits</u>.



Q: Why?

A: Because some errors can be <u>detected</u> and/or <u>corrected</u> using quantum measurements without destroying the logical state.

<u>Ex.</u> The GHZ code:



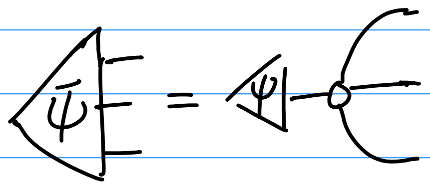
$$\mathbb{C}^2 = \text{span}\{|0\rangle, |1\rangle\} \xmapsto{\;E\;} \text{span}\{|000\rangle, |111\rangle\} \subseteq (\mathbb{C}^2)^{\otimes 3}$$

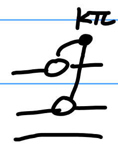
2D   2D   8D

$$|\bar{0}\rangle := |000\rangle, \quad |\bar{1}\rangle := |111\rangle$$

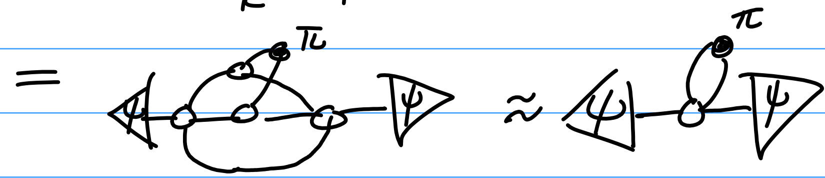
MORE GENERALLY: $|\bar{\psi}\rangle := E|\psi\rangle$.

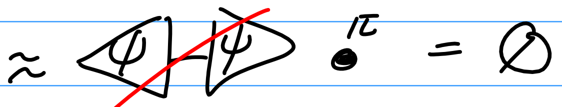
Suppose I measure ZZI:



$M_{ZZI} := \{ P_k :=$  $\}_{k=0,1}$

$$\text{Prob}(1 / |\bar{\psi}\rangle) = \langle \bar{\psi} | P_k | \bar{\psi} \rangle$$

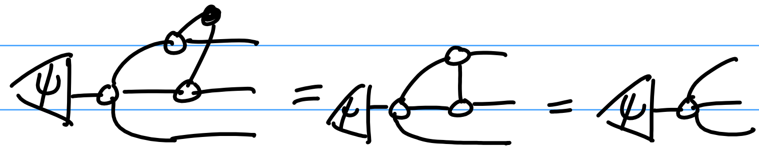




1

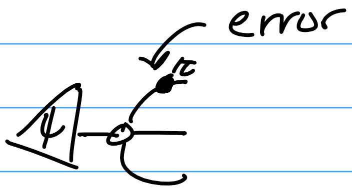
$\Rightarrow \text{Prob}(0 \mid |\bar{\psi}\rangle) = 1.$

Also:

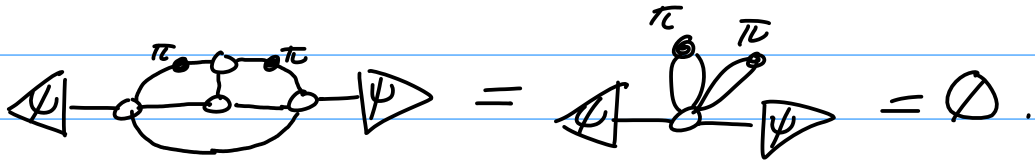
$P_0|\bar{\Psi}\rangle =$  $= |\bar{\Psi}\rangle$

$\Rightarrow$ measuring ZZI does not disturb $|\bar{\Psi}\rangle$.

But



$\text{Prob}(0 \,|\, (X \otimes I \otimes I)|\bar{\Psi}\rangle) =$



$\Rightarrow \text{Prob}(1 \,|\, (X \otimes I \otimes I)|\bar{\Psi}\rangle) = 1$.

So a ZZI measurement can detect the error $X \otimes I \otimes I$.

Thm The GHZ code can detect (and correct) any error in the set $\{XII, IXI, IIX\}$.

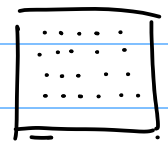
bit-flip errors

Better codes correct more errors (e.g. "phase flips" like $ZII\cdots$, multi-qubit errors, etc.)

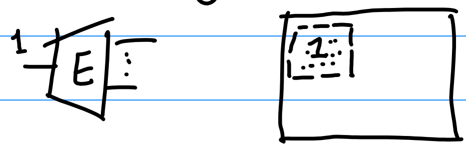
Q: How can I <u>compute</u> with encoded states.

A: FTQC!

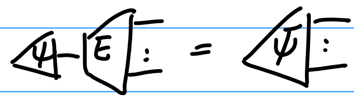
<u>Scheme</u>: Lattice Surgery.

<u>Idea</u>: • Use a grid of qubits 

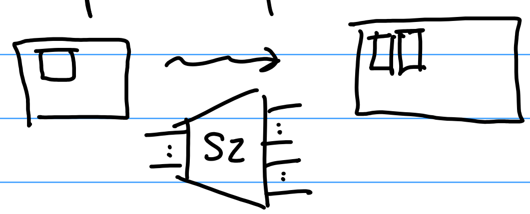
• encode 1 qubit as a "patch"



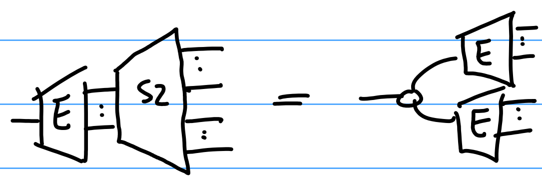
• implement operations to:

— prepare 1Q states: 

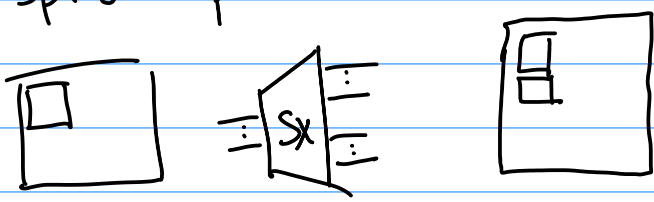
— "Z-split" patches:

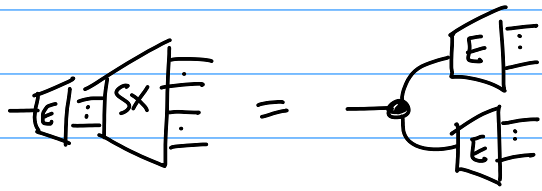


where:

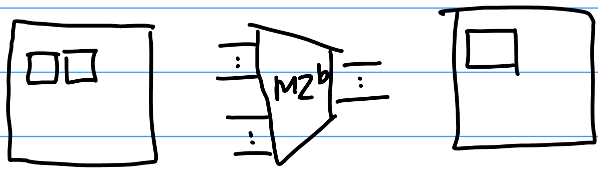


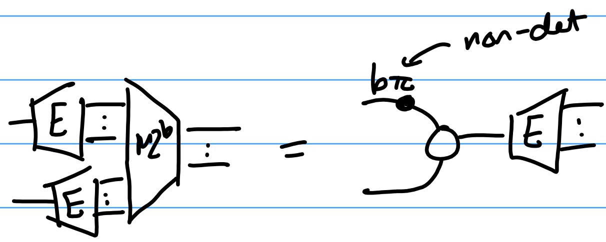
— "X-split" patches



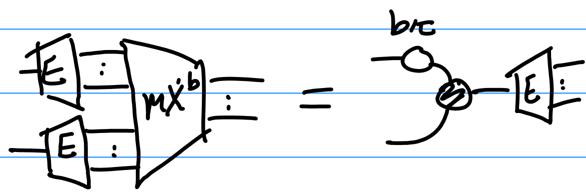


— "Z-merge"

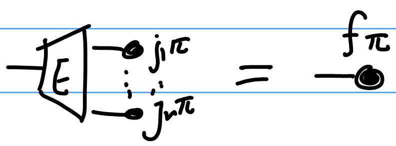


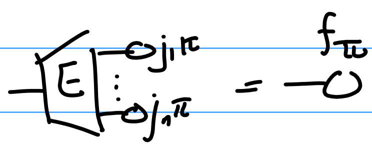


— "X-merge"



— Pauli measurements





Encoded computation: