Classical vs Quantum Information

Jeffrey Bub

Department of Philosophy
and
IPST
University of Maryland

RIT on Quantum Information and Computation, 2010

Outline

Web Resources

- RIT website: http://www.math.umd.edu/ dio/RIT/QI-Spring10
- Sam Lomonaco: *A Rosetta Stone for Quantum Mechanics with an Introduction to Quantum Computation*: http://arxiv.org/pdf/quant-ph/0007045
- Todd Brun: *Lecture Notes on Quantum Information Processing*: http://almaak.usc.edu/ tbrun/Course/index.html
- Valerio Scarani: *Quantum Information: Primitive Notions and Quantum Correlations*: http://arxiv.org/pdf/0910.4222
- John Preskill: *Lecture Notes on Quantum Computation*: http://www.theory.caltech.edu/people/preskill/ph229/

Print Resources

- Sam Lomonaco: *A Rosetta Stone for Quantum Mechanics with an Introduction to Quantum Computation*, in *AMS Short Course Lecture Notes: Quantum Computation* (Providence: AMS, 2000).
- Michael A Nielsen and Isaac L. Chuang: *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press, 2000).
- Chris J. Isham: *Lectures on Quantum Theory: Mathematical and Structural Foundations* (London: Imperial College Press, 1995).

Print Resources

- Hoi-Kwong Lo, Sandu Popescu, Tom Spiller (eds.): *Introduction to Quantum Computation and Information* (World Scientific: 1998).
- L. Diosi: *A Short Course in Quantum Information Theory* (Springer, 2007).
- Michel Le Bellac: *A Short Introduction to Quantum Information and Quantum Computation* (Cambridge University Press, 2005).

Correlations

Quantum probabilities are puzzling because quantum correlations are puzzling, and quantum correlations are puzzling in the way they differ from classical correlations.

Classical correlations

- The space of classical probability distributions, considered as a convex set, has the structure of a simplex.
- An n-simplex is a particular sort of convex set: a convex polytope generated by $n+1$ vertices that are not confined to any $(n-1)$-dimensional subspace (e.g., a triangle or a tetrahedron as opposed to a square or a cube).

Classical correlations

- The simplest classical probability space is the 1-bit space (1-simplex), consisting of two extremal (or pure) probability distributions.
- These are deterministic states, $\mathbf{0} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\mathbf{1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, represented by the vertices of the simplex, with mixtures—convex combinations of extremal states—represented by the line segment between the two vertices: $\mathbf{p} = p\,\mathbf{0} + (1-p)\,\mathbf{1}$, for $0 \leq p \leq 1$.

'No signaling' correlations

- A simplex has the rather special property that any state (probability distribution) can be represented in one and only one way as a mixture of extremal states, the vertices of the simplex. No other state space has this feature: if the state space is not a simplex, the representation of mixed states as convex combinations of extremal states is not unique.
- The simplest quantum system is the qubit, whose state space as a convex set has the structure of a sphere (the Bloch sphere), which is not a simplex.
- The space of all 'no signaling'correlations is a convex polytope that is not a simplex.

'No signaling' correlations

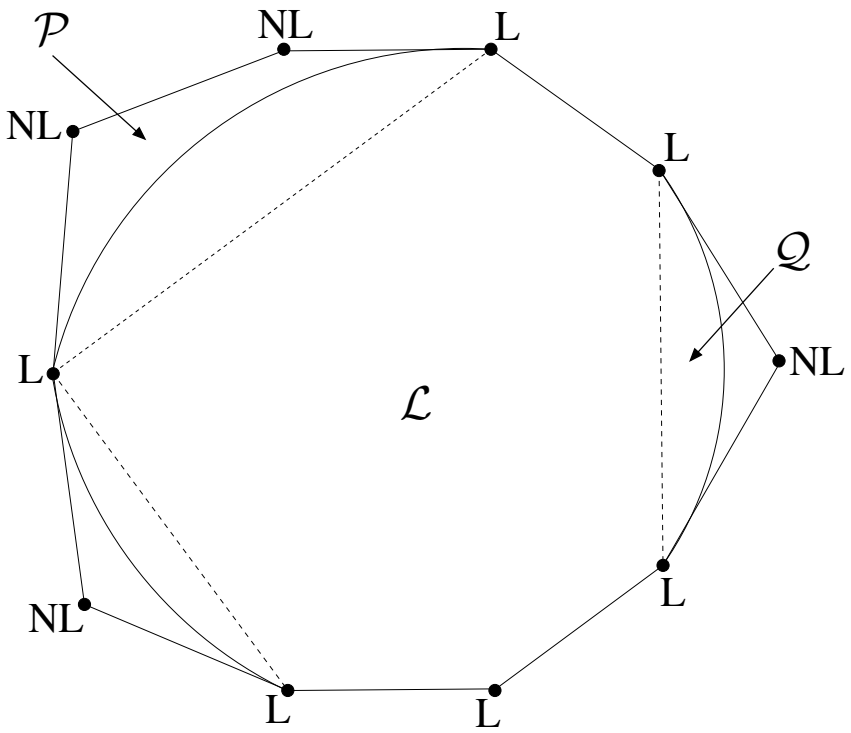


Figure: A schematic representation of the space of no-signaling correlations. The vertices are labelled L and NL for local and nonlocal. Bell inequalities characterize the facets represented by dashed lines. The set bounded by these is $\mathcal{L}$. The region accessible to quantum mechanics is $\mathcal{Q}$. Superquantum correlations lie in region $\mathcal{P}$.

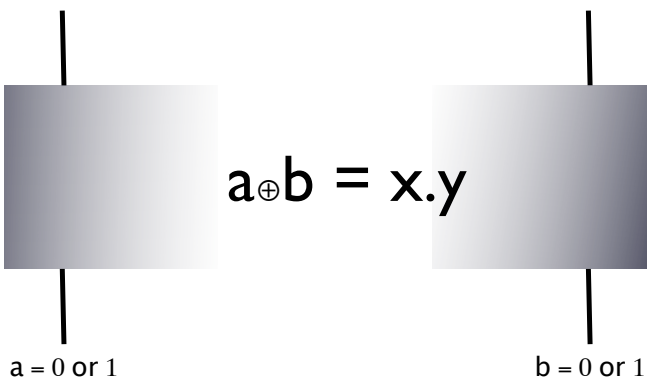
PR-box

- The vertices of the 'no signaling' polytope are deterministic states or non-deterministic Popescu-Rohrlich (PR) boxes (S. Popescu and D. Rohrlich, *Foundations of Physics* **24**, 379 (1994)).
- A PR-box is a hypothetical device or nonlocal information channel that is more nonlocal than quantum mechanics, in the sense that the correlations between outputs of the box for given inputs maximally violate the Tsirelson bound.

PR-box

- A PR-box is defined as follows: there are two inputs, $x \in \{0,1\}$ and $y \in \{0,1\}$, and two outputs, $a \in \{0,1\}$ and $b \in \{0,1\}$. The box is bipartite and nonlocal in the sense that the x-input and a-ouput can be separated from the y-input and b-output by any distance without altering the correlations.
- For convenience, we can think of the x-input as controlled by Alice, who monitors the a-ouput, and the y-input as controlled by Bob, who monitors the b-output.

PR-box



PR-box correlations

Alice's and Bob's inputs and outputs are required to be correlated according to:

$$a \oplus b = x \cdot y$$

where $\oplus$ is addition mod 2, i.e.,

- same outputs (i.e., 00 or 11) if the inputs are 00 or 01 or 10
- different outputs (i.e., 01 or 10) if the inputs are 11

'No signaling'

- The 'no signaling' condition is a requirement on the marginal probabilities: the marginal probability of Alice's outputs do not depend on Bob's input, i.e., Alice cannot tell what Bob's input was by looking at the statistics of her outputs, and conversely.
- Formally:

$$\sum_{b\in\{0,1\}} p(a,b|x,y) = p(a|x),\ a,x,y \in \{0,1\}$$

$$\sum_{a\in\{0,1\}} p(a,b|x,y) = p(b|y),\ b,x,y \in \{0,1\}$$

PR-box marginals

The correlations together with the 'no signaling' condition entail that the marginals are equal to $1/2$ for all inputs $x, y \in \{0, 1\}$ and all outputs $a, b \in \{0, 1\}$:

$$p(a = 0|x) = p(a = 1|x) = p(b = 0|y) = p(b = 1|y) = 1/2$$

PR-box joint probabilities

A PR-box can be defined equivalently in terms of the joint probabilities for all inputs and all outputs. For bipartite probability distributions, with two input values and two output values, the vertices of the 'no signaling' polytope are all PR-boxes (differing only with respect to permutations of the input values and/or output values) or deterministic boxes.

x / y	0		1	
0	$p(00\|00) = 1/2$	$p(10\|00) = 0$	$p(00\|10) = 1/2$	$p(10\|10) = 0$
	$p(01\|00) = 0$	$p(11\|00) = 1/2$	$p(01\|10) = 0$	$p(11\|10) = 1/2$
1	$p(00\|01) = 1/2$	$p(10\|01) = 0$	$p(00\|11) = 0$	$p(10\|11) = 1/2$
	$p(01\|01) = 0$	$p(11\|01) = 1/2$	$p(01\|11) = 1/2$	$p(11\|11) = 0$

Table: Joint probabilities for the PR-box

A Game: simulating a PR-box

- Consider the problem of simulating a PR-box: how close can Alice and Bob come to simulating the correlations of a PR-box for random inputs if they are limited to certain resources?
- In units where $a = \pm 1, b = \pm 1$,

$$\langle 00 \rangle = p(\text{same output}|00) - p(\text{different output}|00)$$

so:

$$\begin{aligned} p(\text{same output}|00) &= \frac{1 + \langle 00 \rangle}{2} \\ p(\text{different output}|00) &= \frac{1 - \langle 00 \rangle}{2} \end{aligned}$$

and similarly for input pairs 01, 10, 11.

CHSH correlation

It follows that the probability of a successful simulation is given by:

$$\begin{aligned}\text{prob(successful sim)} &= \frac{1}{4}(p(\text{same output}|00) + p(\text{same output}|01) \\ &\quad + p(\text{same output}|10) + p(\text{different output}|11)) \\ &= \frac{1}{2}(1 + \frac{CHSH}{4}) = \frac{1}{2}(1 + E)\end{aligned}$$

where

$$CHSH = \langle 00\rangle + \langle 01\rangle + \langle 10\rangle - \langle 11\rangle$$

is the Clauser-Horne-Shimony-Holt (CHSH) correlation.

Bell's locality argument

- Bell's locality argument shows that if Alice and Bob are limited to classical resources, i.e., if they are required to reproduce the correlations on the basis of shared randomness or common causes established before they separate (after which no communication is allowed), then $CHSH_C \leq 2$ (i.e., $E \leq \frac{1}{2}$), so the optimal probability of success is $\frac{1}{2}(1 + \frac{1}{2}) = \frac{3}{4}$.
- If Alice and Bob are allowed to base their strategy on shared entangled states prepared before they separate, then the Tsirelson inequality requires that $CHSH_Q \leq 2\sqrt{2}$ (i.e., $E \leq \frac{1}{\sqrt{2}}$), so the optimal probability of success limited by quantum resources is $\frac{1}{2}(1 + \frac{1}{\sqrt{2}}) \approx .85$.
- For the PR-box, $CHSH = 4$ (i.e., $E = 1$), so the probability of success is, of course, $\frac{1}{2}(1 + 1) = 1$.

Information causality

- M. Pawlowski, T. Patarek, D. Kaszlikowski, V. Scarani, A. Winter, M. Zukowski: 'A New Physical Principle: Information Causality,' *Nature* **461**, 1101 (2009), or quant-ph/0905.2292v1 (at http://www.arxiv.org).
- Information Causality states that the information gain for Bob about an unknown data set of Alice, using all his local resources and m classical bits communicated by Alice, is at most m bits.
- The no-signaling condition is just Information Causality for $m = 0$.

New game: oblivious transfer

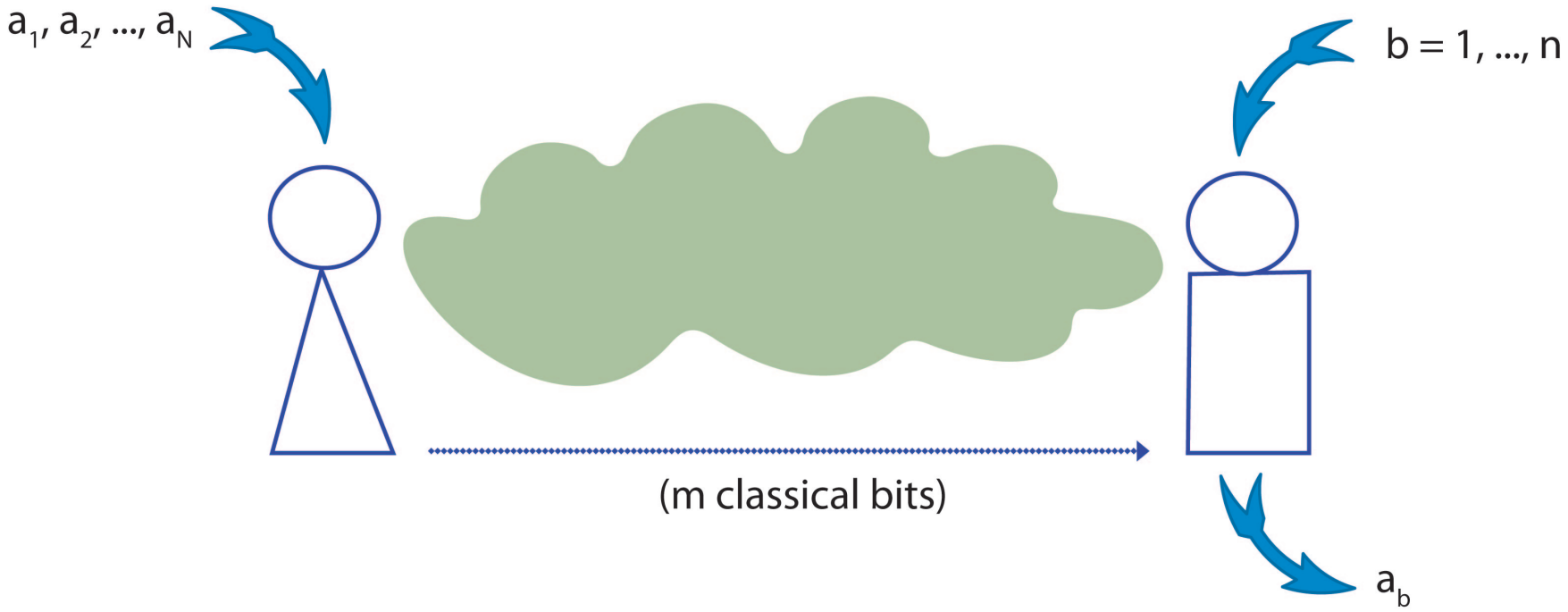


Figure: Alice receives N random and independent bits $\vec{a} = (a_1, a_2, ..., a_N)$. In a separate location, Bob receives a random variable $b \in \{1, 2, ..., n\}$. Alice can send m classical bits to Bob with the help of which Bob is asked to guess the value of the b-th bit in the Alice's list, a_b.

Classical strategy

- Bob can correctly give the value of at least m bits. If Alice sends him a message $\vec{x} = (a_1, ..., a_m)$ Bob will guess a_b perfectly whenever $b \in \{1, ..., m\}$.
- The price to pay is that he is bound to make a sheer random guess when $b \in \{m+1, ..., N\}$.

Information causality condition

- More formally, denote Bob's output by β. The efficiency of Alice's and Bob's strategy is quantified by

$$I \equiv \sum_{k=1}^{N} H(a_k : \beta | b = k)$$

where $H(a_i : \beta | b = k)$ is the Shannon mutual information between a_i and β, computed under the condition that Bob has received $b = k$, i.e.

$$I \equiv \sum_{k=1}^{N} H(a_k) + H(\beta) - H(a_k, \beta)$$

- By definition, Information Causality is fulfilled if

$$I \leq m$$

Shannon entropy and mutual information

- The Shannon entropy of a random variable X is defined as:

$$H(X) = -\sum_i p_i \log p_i$$

- The mutual information $H(X:Y)$—sometimes $I(X:Y)$— of two random variables is a measure of how much information they have in common: the sum of the information content of the two random variables, as measured by the Shannon entropy (in which joint information is counted twice), minus their joint information:

$$H(X:Y) = H(X) + H(Y) - H(X,Y)$$

No signaling' boxes

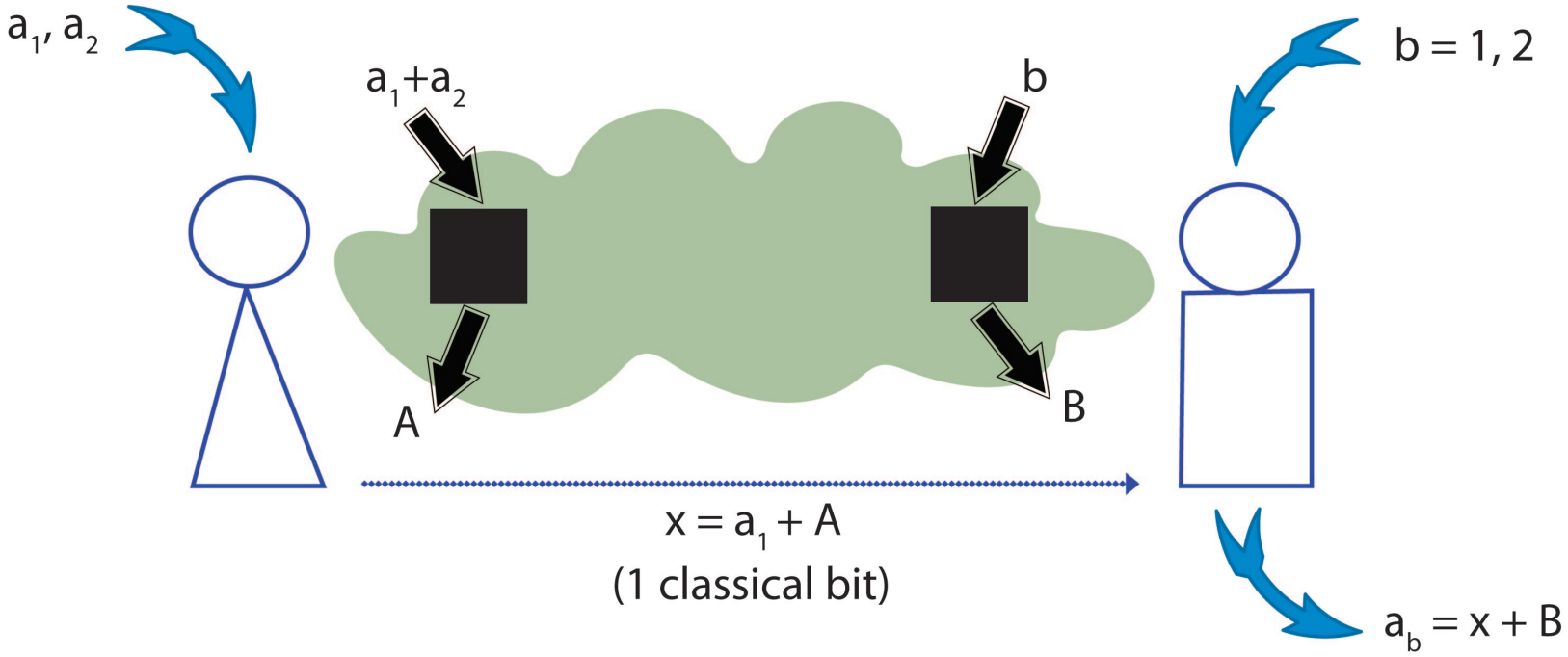


Figure: Simplest case ($m = 1$): Alice receives two bits (a_1, a_2) and is allowed to send only one bit $(m = 1)$ to Bob. A convenient way of thinking about no-signaling resources is to consider black boxes shared between Alice and Bob (NS-boxes). Note: the $+$ here should be $\oplus$.

Violating information causality

- The correlations between inputs a, b and outputs A, B of the boxes are described by probabilities $P(A \oplus B = ab|a, b)$. Assume random local outputs, so no-signaling is satisfied.
- With suitable NS-boxes Alice and Bob can violate Information Causality. Alice uses $a = a_1 \oplus a_2$ as an input to the shared NS-box and obtains the outcome A, which is used to compute her message bit $x = a_1 \oplus A$ for Bob. Bob, on his side, inputs $b = 1$ if he wants to learn a_1, and $b = 2$ if he wants to learn a_2; he gets the outcome B.
- Upon receiving x from Alice, Bob computes his guess $\beta = x \oplus B = a_1 \oplus A \oplus B$.

Probability of guessing the value of a bit

The probability that Bob correctly guesses the value of the bit a_1 is

$$P_{\mathrm{I}} = \frac{1}{2}\left[P(A \oplus B = 0|0,0) + P(A \oplus B = 0|1,0)\right],$$

and the analogous probability for the bit a_2 reads

$$P_{\mathrm{II}} = \frac{1}{2}\left[P(A \oplus B = 0|0,1) + P(A \oplus B = 1|1,1)\right].$$

Unbiased case

In the unbiased case, where Bob's bit can be 0 or 1 with equal probability, the probability that Bob guesses successfully (i.e., gives the correct value of Alice's k'th bit if his input bit is k, for $k = 1, 2$) is:

$$\begin{aligned} P(\text{success}) &= \frac{1}{2}(P_{\text{I}} + P_{\text{II}}) = \frac{1}{4}\sum_{a,b} P(A \oplus B = ab|a, b) \\ &= \frac{1}{4}((p(\text{same}|0, 0) + p(\text{same}|0, 1) \\ &\quad + p(\text{same}|1, 0) + p(\text{different}|1, 1))) \\ &= \frac{1}{2}(1 + E) \end{aligned}$$

Bounds

Recall:

- classical bound: $E = \frac{1}{2}$ $(CHSH = 2)$
- quantum (Tsirelson) bound: $E = \frac{1}{\sqrt{2}}$ $(CHSH = 2\sqrt{2})$
- PR-box: $E = 1$ $(CHSH = 4)$
- Note that for uncorrelated random bits: $E = 0$ $(CHSH = 0)$

Alice receives n-bits

- In the case where the outcomes of the boxes are uniformly random, the correlations are given by:

$$P(A \oplus B = ab|a, b) = \frac{1}{2}(1 + E)$$

with $0 \leq E \leq 1$.

- So $P_k = \frac{1}{2}(1 + E)$
- If Alice receives $N = 2^n$ bits and Bob receives n input bits b_n that describe the index of the bit he has to guess, and Alice is allowed to send Bob 1 bit, the probability that Bob guesses a_k correctly can be shown to be given by:

$$P_k = \frac{1}{2}(1 + E^n)$$

N-bit case

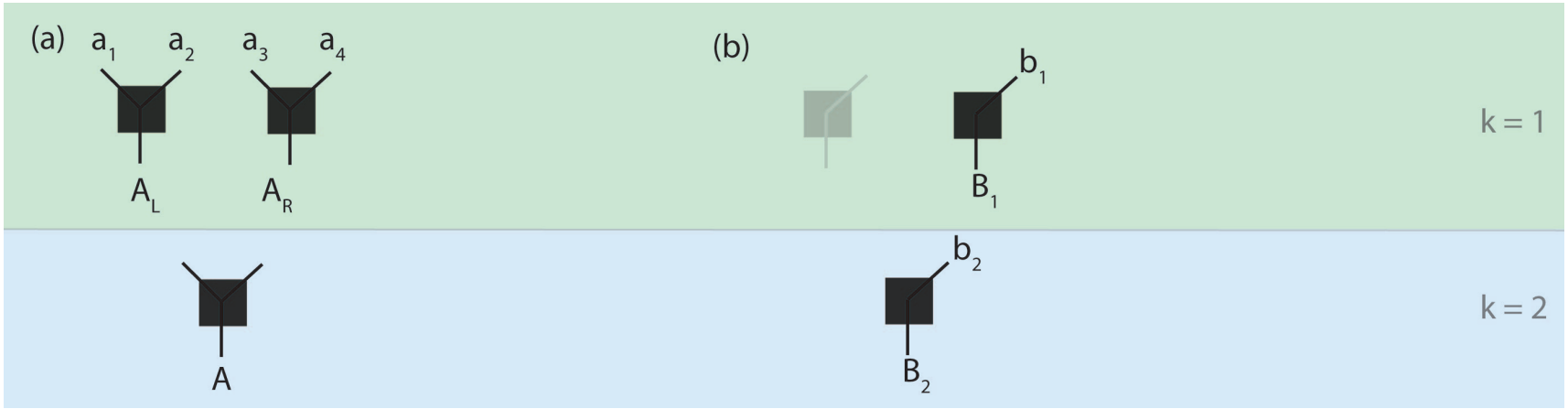


Figure: Alice receives $N = 2^n$ input bits and Bob receives n input bits b_n that label the index of the bit he has to guess: $b = 1 + \sum_{k=1}^{n} b_k 2^{k-1}$. Alice is allowed to send Bob a single bit, $m = 1$. (Note: The two inputs to the boxes at (a) should be one input $a_1 \otimes a_2$ to the Left box, one input $a_3 \otimes a_4$ to the Right box at level $k = 1$, and one input $(a_1 \otimes A_L) \otimes (a_3 \otimes A_R)$ to the box at the level $k = 2$.)

$n = 2$ protocol

- To encode information about her data, Alice uses a pyramid of NS-boxes as shown in the panel (a) for $n = 2$. Previously we saw how Bob can correctly guess the first or second bit of Alice using a single pair of the boxes.
- The probabilities of guessing correctly the first or the second bit are P_{I} and P_{II}, respectively. If Alice has more bits, Bob recursively uses this protocol.

$n = 2$ protocol

- For four input bits of Alice, two pairs of NS-boxes on the level $k = 1$ allow Bob to make the guess of a value of any one of Alice's bits as soon as he knows either $a_1 \oplus A_L$ or $a_3 \oplus A_R$, which are the one-bit messages of the 1-box protocol.
- These can be encoded using the third box, on the level $k = 2$, by inserting their sum to the Alice's box and sending $x = a_1 \oplus A_L \oplus A$ to Bob.
- Depending on the bit he is interested in, he now reads a suitable message using the box on the level $k = 2$ and uses one of the boxes on the level $k = 1$.
- The situation in which Bob aims at the value of a_3 or a_4 is depicted in the panel (b). Bob's final answer is $x \oplus B_2 \oplus B_1$.

$n = 2$ protocol example

- Generally, Alice and Bob use a pyramid of $N - 1$ pairs of boxes placed on n levels. Looking at the binary decomposition of b Bob aims $(n - k)$ times at the left bit and k times at the right, where $k = b_1 + ... + b_n$.
- His final guess is the sum of $\beta = x \oplus B_1 \oplus ... \oplus B_n$. Therefore, Bob's final guess is correct whenever he has made an even number of errors in the intermediate steps.
- This leads to $P_k = \frac{1}{2}(1 + E^n)$ for the probability of his correct final guess.

$n = 2$ protocol

- Consider the case $n = 2$. Bob receives 2 input bits b_1, b_2 that label the index of the bit he has to predict, as follows:

$$b = 1 + b_1 2^1 + b_2 2^2 = 1 + b_1 + b_2$$

- So:

$$\text{if } b_1 = 0, b_2 = 0 : b = 1$$
$$\text{if } b_1 = 1, b_2 = 0 : b = 2$$
$$\text{if } b_1 = 0, b_2 = 1 : b = 3$$
$$\text{if } b_1 = 1, b_2 = 1 : b = 4$$

$n = 2$ protocol

- So $b_2 = 0$ or 1 distinguishes between the pairs:
  - $b_2 = 0$: 1st bit, 2nd bit
  - $b_2 = 1$: 3rd bit, 4th bit
- $b_1 = 0$ or 1 then distinguishes between the two members of the pair chosen by b_2

How the $n = 2$ protocol works

The protocol works as follows:

- At the first level, there are two boxes. Call them L and R.
- Alice inputs $a_1 \oplus a_2$ into the L box, and $a_3 \oplus a_4$ into the R box.
- Bob inputs b_1 into both boxes (the input to one of these boxes is going to be irrelevant, depending on what bit Bob aims to guess).
- At the second level, they use one box. Alice inputs $(a_1 \oplus A_L) \oplus (a_3 \oplus A_R)$ into this box, where A_L is the Alice-output of the L box and aA_R is the Alice output of the R box.
- Bob inputs b_2 into this box.
- Alice then sends Bob 1 bit: $a_1 \oplus A_L \oplus A$, where A is the Alice-output of the second level box.

How the $n = 2$ protocol works

- Now Bob's output of this second level box is: B_2
- Bob can predict either $a_1 \oplus A_L$ or $a_3 \oplus A_R$ (using the protocol for the single box) as:

$$[(a_1 \oplus A_L) \oplus A] \oplus B_2$$

- Suppose Bob wants to predict a_3. This corresponds to $b_1 = 0, b_2 = 1$. He takes the Bob-output of the R box (which corresponds to the pair (a_3, a_4)) and adds this to the above bit:

$$[(a_1 \oplus A_L) \oplus A] \oplus B_1 \oplus B_2$$

This is his final response (his guess for the value of a_3).

How the $n = 2$ protocol works

- Now Bob's response will be correct if $[(a_1 \oplus A_L) \oplus A] \oplus B_2$ correctly predicts the required pair (in this case a_3, a_4), and if $[(a_1 \oplus A_L) \oplus A] \oplus B_1 \oplus B_2$ correctly predicts the correct member of the pair (in this case, a_3).
- But Bob's response will also be correct if he is incorrect in both cases (because the errors will cancel out, i.e., $B_1 \oplus B_2$ is the same if B_1, B_2 are both correct or both incorrect, i.e., if B_1 differs from the correct bit by 1 and B_2 differs from the correct bit by 1).

Probabiity that Bob is correct

- The probability of being correct at both levels is:

$$\frac{1}{2}(1+E)\cdot\frac{1}{2}(1+E)=\frac{1}{4}(1+E)^2$$

- So the probability of being correct at both levels or incorrect at both levels is:

$$\frac{1}{4}(1+E)^2+\frac{1}{4}(1-E)^2=\frac{1}{2}(1+E^2)$$

since $1-\frac{1}{2}(1+E)=\frac{1}{2}(1-E)$

- In the general case, iterating the procedure, you get: $\frac{1}{2}(1+E^n)$

Probabiity that Bob is correct

- So far we've shown that if Alice has $N = 2^n$ bits then the probability of Bob correctly guessing Alice's k'th bit is:

$$P_k = \frac{1}{2}(1 + E^n)$$

- We now show that if $E > \frac{1}{\sqrt{2}}$ then information causality is violated, i.e., $I > m$.

Simple case $n = 1$

- Consider the simple case $n = 1$, where Alice has $2^1 = 2$ bits, Bob receives $n = 1$ bit that indicates the index of Alice's bit that he has to guess, and Alice can send Bob 1 bit of information.
- Information causality is the condition that $I \leq 1$, where:

$$I = (H(a_1) + H(\beta) - H(a_1, \beta)) + (H(a_2) + H(\beta) - H(a_2, \beta))$$

- In the unbiased case $H(a_1) = H(\beta) = 1$ and $H(a_1, \beta) = H(a_2, \beta)$.
- For $I \leq 1$ to be *violated* (i.e, $I > 1$), we require $4 - 2H(a_k, \beta) > 1$, i.e.,

$$H(a_k, \beta) < 1 + \frac{1}{2}$$

Condition for a violation of information causality

- We can show that

$$H(a_k, \beta) = 1 + h(P_k)$$

where $h(P_k)$ is the binary entropy of P_k:

$$h(P_k) = -(\frac{1}{2}(1+E)\log(1+E) + \frac{1}{2}(1-E)\log(1-E))$$

- So the condition for a violation of information causality can be expressed as:

$$\begin{aligned} \text{For } n = 1\text{: } H(a_k, \beta) &< 1 + \frac{1}{2} \\ \text{i.e., } h(P_k) &\leq \frac{1}{2} \end{aligned}$$

Probability that Bob correctly guesses Alice's k'th bit

- In the unbiased case, the probability that Bob correctly guesses Alice's k'th bit is:

$$P_k = \frac{1}{2}(1+E) = p(a_k = 0, \beta = 0) + p(a_k = 1, \beta = 1)$$

- $p(a_k = 0, \beta = 0) = p(a_k = 1, \beta = 1) = \frac{1}{2}P_k = \frac{1}{4}(1+E)$
- $p(a_k = 0, \beta = 1) = p(a_k = 1, \beta = 0) = \frac{1}{2}(1-P_k) = \frac{1}{4}(1-E)$

$H(a_k, \beta) = h(P_k) + 1$

$$
\begin{aligned}
H(a_k,\beta) &= -(p(a_k=0,\beta=0)\log p(a_k=0,\beta=0) \\
&\quad +p(a_k=0,\beta=1)\log p(a_k=0,\beta=1) \\
&\quad +p(a_k=1,\beta=0)\log p(a_k=1,\beta=0) \\
&\quad +p(a_k=1,\beta=1)\log p(a_k=1,\beta=1)) \\
&= -(\frac{1}{4}(1+E)\log\frac{1}{4}(1+E)+\frac{1}{4}(1-E)\log\frac{1}{4}(1-E) \\
&\quad +\frac{1}{4}(1-E)\log\frac{1}{4}(1-E)+\frac{1}{4}(1+E)\log\frac{1}{4}(1+E)) \\
&= -(\frac{1}{2}(1+E)\log\frac{1}{2}(1+E)+\frac{1}{2}(1-E)\log\frac{1}{2}(1-E))+1 \\
&= h(P_k)+1
\end{aligned}
$$

where $h(P_k) = -(P_k \log P_k + (1-P_k)\log(1-P_k))$ is the binary entropy of P_k.

Case $n = 2$

- Now consider the case $n = 2$, where Alice has $2^2 = 4$ bits, Bob receives $n = 2$ bits that indicate the index of Alice's bit that he has to guess, and Alice can send Bob 1 bit of information.
- As we saw above, $P_k = \frac{1}{2}(1 + E^2)$
- In this case:

$$\begin{aligned} I &= [H(a_1) + H(\beta) - H(a_1, \beta)] \\ &\quad + [H(a_2) + H(\beta) - H(a_2, \beta)] \\ &\quad + [H(a_3) + H(\beta) - H(a_3, \beta)] \\ &\quad + [H(a_4) + H(\beta) - H(a_4, \beta)] \end{aligned}$$

Information causality violation condition

- So for $I > 1$ to be satisfied, we require:

$$8 - 4H(a_k, \beta) > 1$$

i.e., $H(a_k, \beta) < \frac{7}{4}$

- So the condition for a violation of information causality can be expressed as:

$$\begin{aligned} \text{For } n = 2\text{: } H(a_k, \beta) &< 1 + \frac{3}{4} \\ \text{i.e., } h(P_k) &\leq \frac{3}{4} \end{aligned}$$

General case

- For the general case: $P_k = \frac{1}{2}(1 + E^n)$
- The condition for a violation of information causality becomes:

$$\begin{aligned} \text{For } n\text{: } H(a_k, \beta) &< 1 + \frac{2^n - 1}{2^n} \\ i.e., h(P_k) &< \frac{2^n - 1}{2^n} \end{aligned}$$

General case

- In the general case, we have a violation of information causality when:

$$\begin{aligned} h(P_k) &< \frac{2^n - 1}{2^n} \\ \text{i.e., } h(\frac{1}{2}(1 + E^n)) &< \frac{2^n - 1}{2^n} \end{aligned}$$

- The following inequality can be proved:

$$h(\frac{1}{2}(1 + y)) \leq 1 - \frac{y^2}{2 \ln 2}$$

where $\ln 2 \approx .693$ is the natural log of 2 (base e).

General case

- So you get a violation of information causality when:

$$1 - \frac{E^{2n}}{2 \ln 2} < \frac{2^n - 1}{2^n}$$

- This becomes:

$$(2E^2)^n > 2 \ln 2 \approx 1.386$$

The Tsirelson bound

- The condition for a violation of information causality is: $(2E^2)^n > 2\ln 2 \approx 1.386$.
- If $2E^2 = 1$, i.e., if $E = \frac{1}{\sqrt{2}}$, the Tsirelson bound, then you *don't* get a violation.
- If $2E^2 = 1 + a$, for some a, no matter how small, then:

$$\begin{aligned} (1+a)^n &= 1 + na + \frac{n(n-1)}{2!} + \frac{n(n-1)(n-2)}{3!} + \cdots \\ \text{i.e., } (2E^2)^n &> 1 + na \end{aligned}$$

- But $1 + na > 2\ln 2 \approx 1.386$ for some n, i.e., $na > .386$, or,

$$n > \frac{.386}{a}$$

for some n, for any a, however small.

Some numbers

- Let's take some numbers for E and n to see how things work (remembering that $\log_2 x = \frac{\log_{10} x}{\log_{10} 2} \approx \frac{\log_{10} x}{.301}$).
- We get a violation of information causality when $h(P_k) < \frac{2^n - 1}{2^n}$.
- For the case $n = 1$, where Alice has $2^1 = 2$ bits, and $E = \frac{1}{\sqrt{2}}$ (the Tsirelson bound):

$$\begin{aligned} h(P_k) &= -(\frac{1}{2}(1 + \frac{1}{\sqrt{2}})\frac{\log_{10} \frac{1}{2}(1 + \frac{1}{\sqrt{2}})}{.301} \\ &\quad + \frac{1}{2}(1 - \frac{1}{\sqrt{2}})\frac{\log_{10} \frac{1}{2}(1 - \frac{1}{\sqrt{2}})}{.301}) \\ &\approx .600 \end{aligned}$$

There is no violation of information causality because $.600 > \frac{2-1}{2} = \frac{1}{2}$.

Case $n = 8$ for Tsirelson bound

- For the case $n = 8$, where Alice has $2^8 = 256$ bits, and $E = \frac{1}{\sqrt{2}}$ (the Tsirelson bound), we get a violation of information causality when:

$$h(P_k) < \frac{2^8 - 1}{2^8} = \frac{255}{256}$$

- In this case:

$$\begin{aligned} h(P_k) &= -(\frac{1}{2}(1 + \frac{1}{\sqrt{2}^8})\frac{\log_{10}\frac{1}{2}(1 + \frac{1}{\sqrt{2}^8})}{.301} \\ &\quad + \frac{1}{2}(1 - \frac{1}{\sqrt{2}^8})\frac{\log_{10}\frac{1}{2}(1 - \frac{1}{\sqrt{2}^8})}{.301}) \\ &\approx .997 \end{aligned}$$

There is still no violation of information causality because $.997 > \frac{255}{256} \approx .996$.

Case $n = 8$ and $E = .725$

- Take $n = 8$ and $E = .725$ (i.e., $E > \frac{1}{\sqrt{2}} \approx .707$, i.e., $E >$ Tsirelson bound).
- In this case:

$$\begin{aligned} h(P_k) &= -(\frac{1}{2}(1+.725^8)\frac{\log_{10}\frac{1}{2}(1+.725^8)}{.301} \\ &\quad +\frac{1}{2}(1-.725^8)\frac{\log_{10}\frac{1}{2}(1-.725^8)}{.301}) \\ &\approx .9958 \end{aligned} \tag{1}$$

Now there is a violation of information causality because $.9958 < \frac{255}{256} \approx .996$.

E close to Tsirelson bound

- If E is very close to the Tsirelson bound, then n must be very large for a violation of information causality.
- For example, $E_T \approx .707$. For $n = 10$, where Alice has $2^{10} = 1024$ bits, $\frac{1023}{1024} \approx .9990$, there is no violation:

$$\begin{aligned} P_k &= -(\frac{1}{2}(1+\frac{1}{\sqrt{2}^{10}})\frac{\log_{10}\frac{1}{2}(1+\frac{1}{\sqrt{2}^{10}})}{.301} \\ &\quad +\frac{1}{2}(1-\frac{1}{\sqrt{2}^{10}})\frac{\log_{10}\frac{1}{2}(1-\frac{1}{\sqrt{2}^{10}})}{.301}) \approx .99939 \end{aligned}$$

- If we take $E = .708$, then there is still no violation:

$$\begin{aligned} P_k &= -(\frac{1}{2}(1+.708^{10})\frac{\log_{10}\frac{1}{2}(1+.708^{10})}{.301} \\ &\quad +\frac{1}{2}(1-.708^{10})\frac{\log_{10}\frac{1}{2}(1-.708^{10})}{.301}) \approx .99937 \end{aligned}$$

Looking at it another way

- Another way to look at this: As $n \to \infty$, for $E = \frac{1}{\sqrt{2}}$, $P_k = \frac{1}{2}(1 + E^n) \to \frac{1}{2}$ (so $H(a_k, \beta) \to 1 + \frac{1}{2}$), and:

$$\begin{aligned} h(P_k) &\to \frac{2^n - 1}{2^n} \\ \text{i.e., } I &\to 1 \,(\text{from below}) \end{aligned}$$

- For a PR-box, $E = 1$, $h(P_k) = 0$, so $I = N = 2^n$.
- Note that in the unbiased case:

$$I = 2N - \sum_{k=1}^{N} H(a_k, \beta)$$

So: $0 \leq I \leq N$

Looking at it another way

- If Bob guesses randomly for all k, then:

$$h(P_k) = 1 \,(\text{so } H(a_k, \beta) = 1 + 1 = 2)$$

and

$$I = 0$$

- In fact, $0 \leq I \leq N$, with a violation of information causality when $I > 1$ (for $m = 1$, where m is the number of bits Alice is allowed to send to Bob).