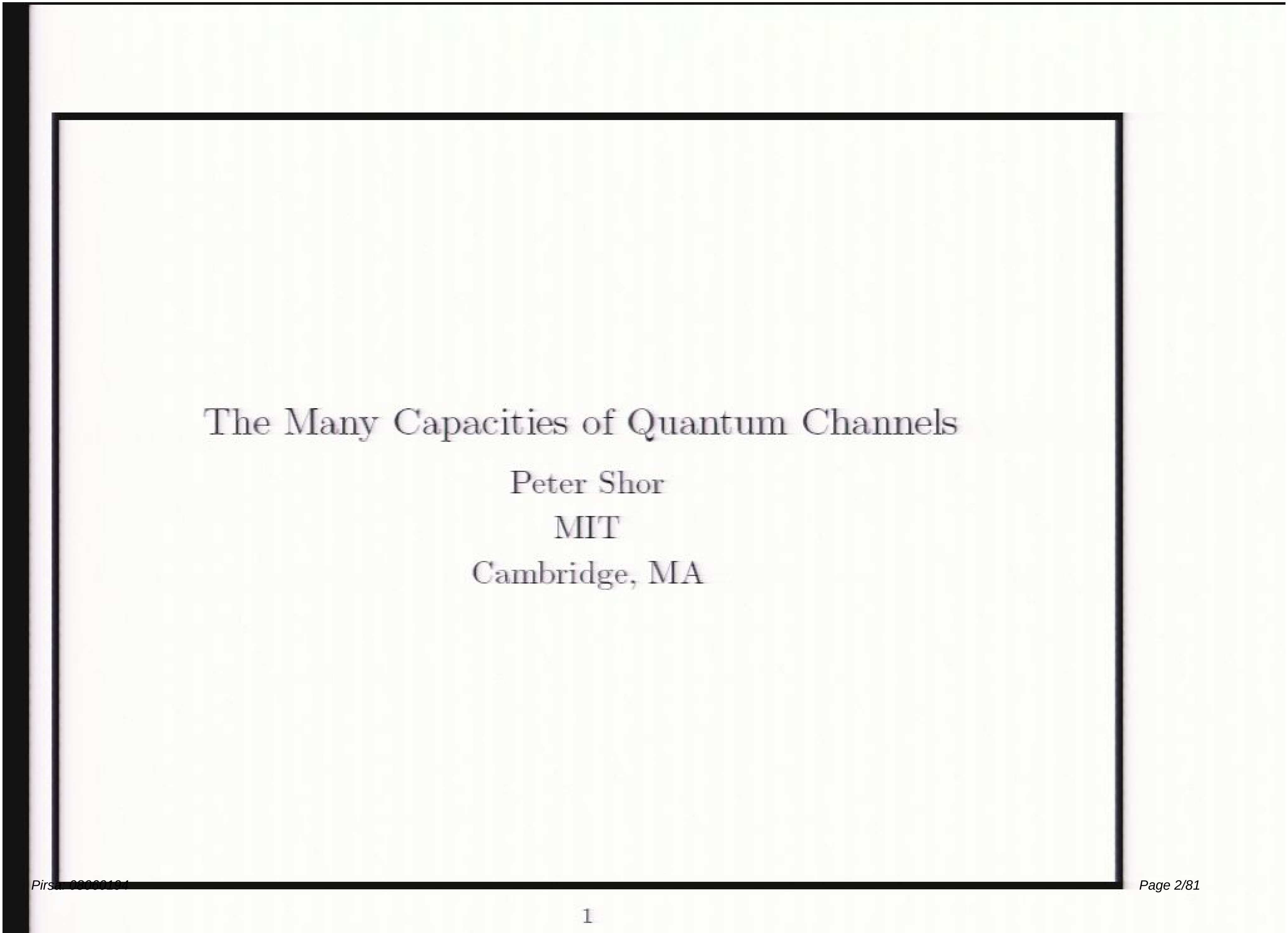
Title: The Many Capacities of a Quantum Channel

Date: Jun 25, 2008 09:00 AM

URL: http://pirsa.org/08060194

Abstract:



The Many Capacities of Quantum Channels

Peter Shor
MIT
Cambridge, MA

Classical channels have essentially only one capacity. For quantum channels, we can define many capacities.

We will mention:

1 Accessible informtation.

2,2$\frac{1}{2}$ Adaptively accessible information.

3 Classical capacity.

3$\frac{1}{2}$ Holevo capacity.

4 Entanglement-assisted capacity.

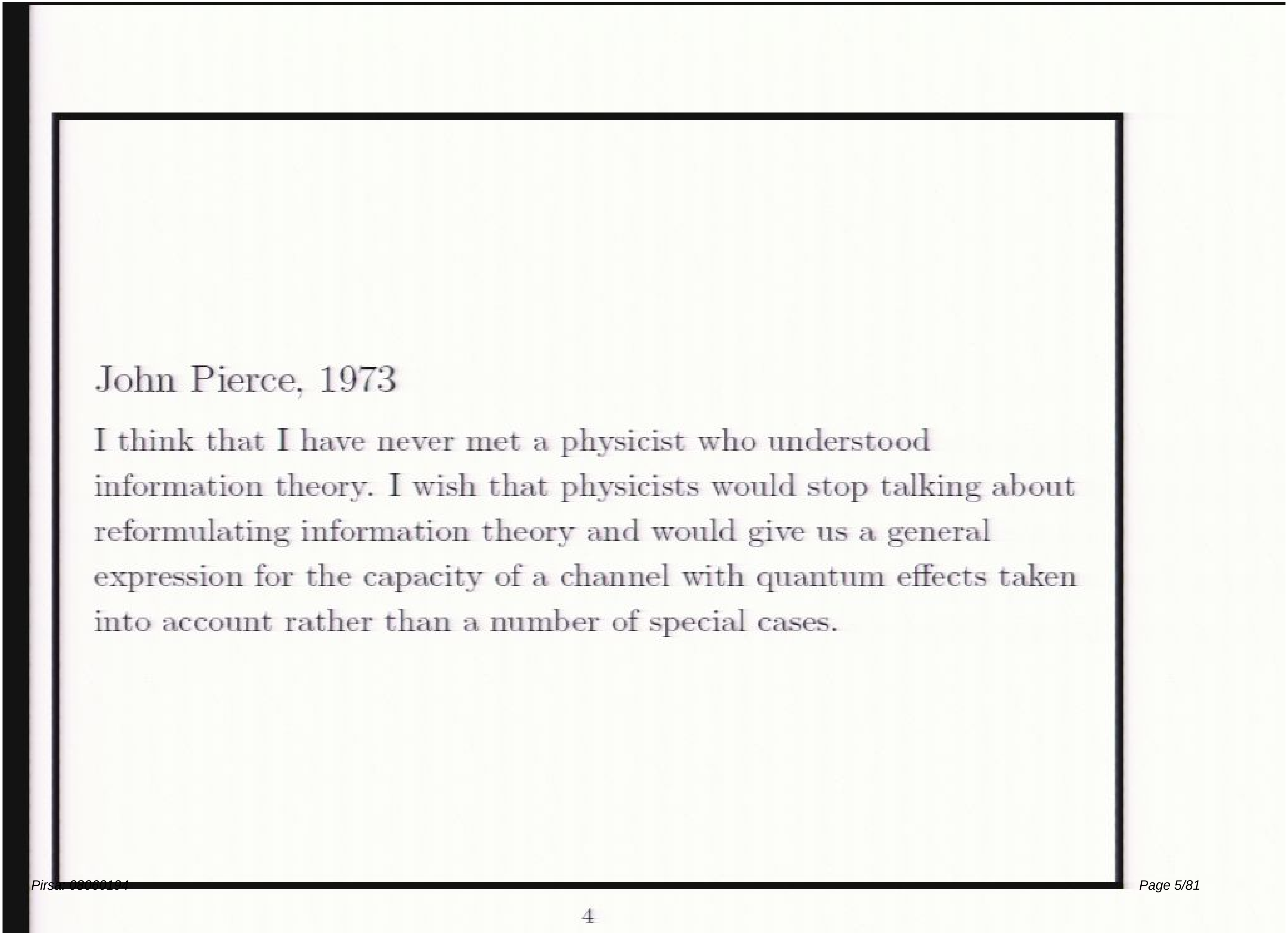
5 Quantum capacity.

6,7 Capacity (q and c) with feedback channel.

8,9 Capacity (q and c) with side channel.

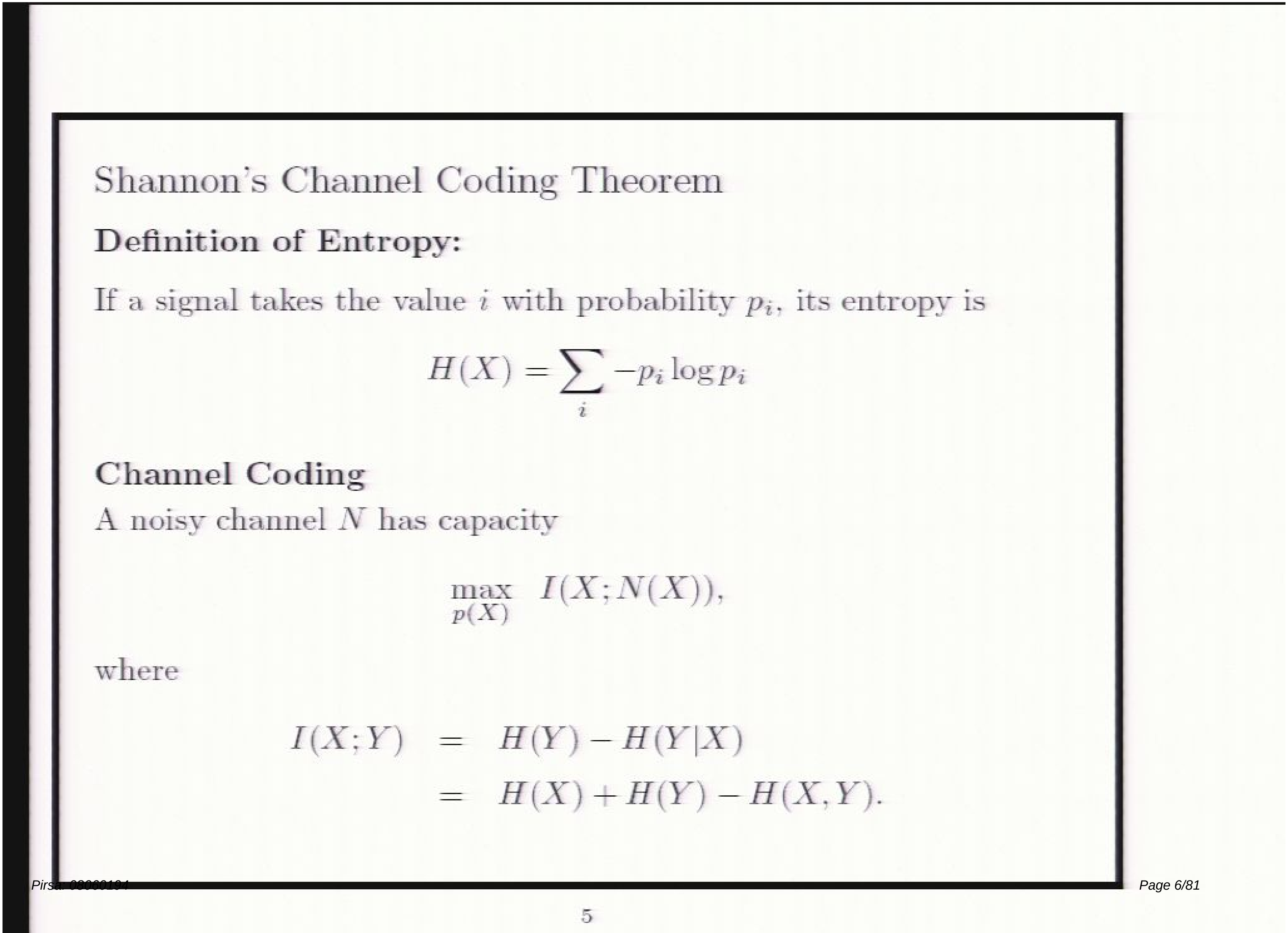
Claude Shannon, 1948

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point.



John Pierce, 1973

I think that I have never met a physicist who understood information theory. I wish that physicists would stop talking about reformulating information theory and would give us a general expression for the capacity of a channel with quantum effects taken into account rather than a number of special cases.



Shannon's Channel Coding Theorem

Definition of Entropy:

If a signal takes the value i with probability p_i, its entropy is

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

Channel Coding

A noisy channel N has capacity

$$\max_{p(X)} \; I(X;N(X)),$$

where

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X,Y). \end{aligned}$$

Entropy of a quantum state

Classical Case

Given n photons, each in state $|\updownarrow\rangle$ or $|\leftrightarrow\rangle$, with probability $\frac{1}{2}$. Any two of these states are completely distinguishable. The entropy is n bits.

Quantum Case

Given n photons, each in state $|\updownarrow\rangle$ or $|\nearrow\rangle$, with probability $\frac{1}{2}$. If the angle between the polarizations is small, any two of these states are barely distinguishable. Intuitively, the entropy should be much less than n bits.

By thermodynamic arguments, von Neumann deduced the entropy of a quantum system with density matrix ρ is

$$H(\rho) = -\mathrm{Tr}(\rho \log \rho)$$

Recall ρ was positive semidefinite, so $\rho \log \rho$ is defined.

If ρ is diagonal with eigenvalues λ_i, then $\rho \log \rho$ is diagonal with eigenvalues $\lambda_i \log \lambda_i$.
Thus, $H(\rho) = H_{\mathrm{Shan}}(\lambda_i)$ so the von Neumann entropy is the Shannon entropy of the eigenvalues.
(Recall $\mathrm{Tr}\rho = 1 = \sum_i \lambda_i$.)

You can ask: Does this definition give the right quantum channel capacity?

Accessible Information

Suppose that we have a source that outputs signal ρ_i with probability p_i. How much Shannon information can we extract about the sequence of i's?

Let X be the random variable telling which signal ρ_i was sent.

Answer (from classical information theory):
Optimize over all possible measurements M on the signals (with outcomes M_1, M_2, $\ldots$).

$$I_{\text{acc}} = \max_M I(X, M)$$

Example 1: Two states in ensemble

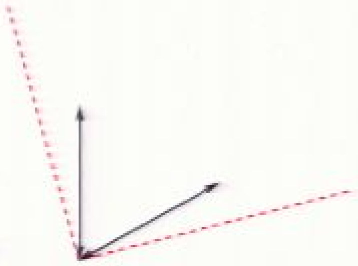
$v_1 =$ ↕ $\qquad v_2 =$ ⤢

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$$

Then

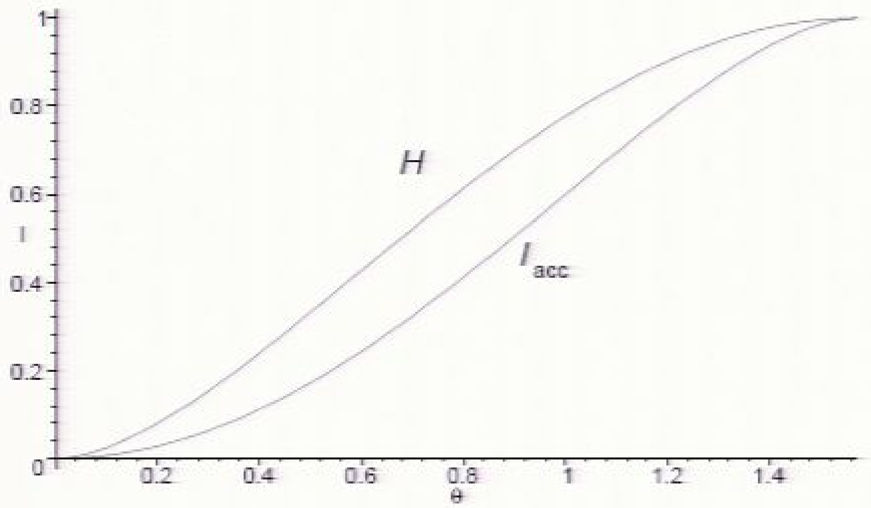
$$\rho = \frac{1}{2} \begin{pmatrix} 1 + \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & 1 - \cos^2\theta \end{pmatrix}$$

and $H = H(\frac{1}{2} + \frac{\cos\theta}{2})$.



The optimal measurement is and $I_{\text{acc}} = 1 - H(\frac{1}{2} + \frac{\sin\theta}{2})$.

We see that $I_{\text{acc}} < H(\rho)$.



A plot of H and I_{acc} for the ensemble of two pure quantum states with $p = 1/2$ that differ by an angle of θ, $0 \le \theta \le \pi/2$.

The top curve is the von Neumann entropy $H = H(\frac{1}{2} + \frac{\cos\theta}{2})$ and the bottom the accessible information $I_{\text{acc}} = 1 - H(\frac{1}{2} + \frac{\sin\theta}{2})$.

POVM Measurements

(Positive Operator Valued Measurements).

We are given a set of positive semidefinite matrices E_i satisfying $\sum_i E_i = I$.

The probability of the i'th outcome is

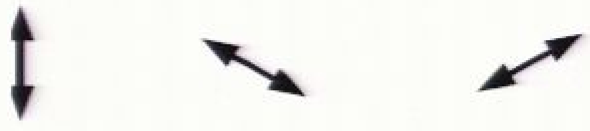
$$p_i = \mathrm{Tr}(E_i \rho)$$

For von Neumann measurements, $E_i = \Pi_{S_i}$

To obtain the maximum information, we can assume that E_i's are pure states. Then $E_i = v_i v_i^\dagger$ for some vector v_i.

Example 2:

Three signal states differing by 60°.

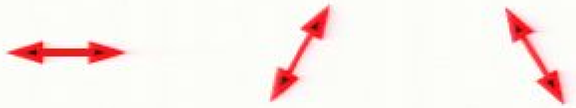
v_i:  (prob $\frac{1}{3}$)

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \end{pmatrix} \quad v_3 = \begin{pmatrix} -1/2 \\ -\sqrt{3}/2 \end{pmatrix}$$

Optimal Measurement:

POVM corresponding to vectors $w_i \perp v_i$.

$E_i = \frac{2}{3} w_i w_i^\dagger$

w_i:  (prob $\frac{1}{3}$)

Each outcome rules out one state, leaves other two equally likely

$I_{\mathrm{acc}} = \log 3 - 1 = .585$ bits; $H = 1$ bit. Again, $I_{\mathrm{acc}} \leq H$.

Holevo Bound χ

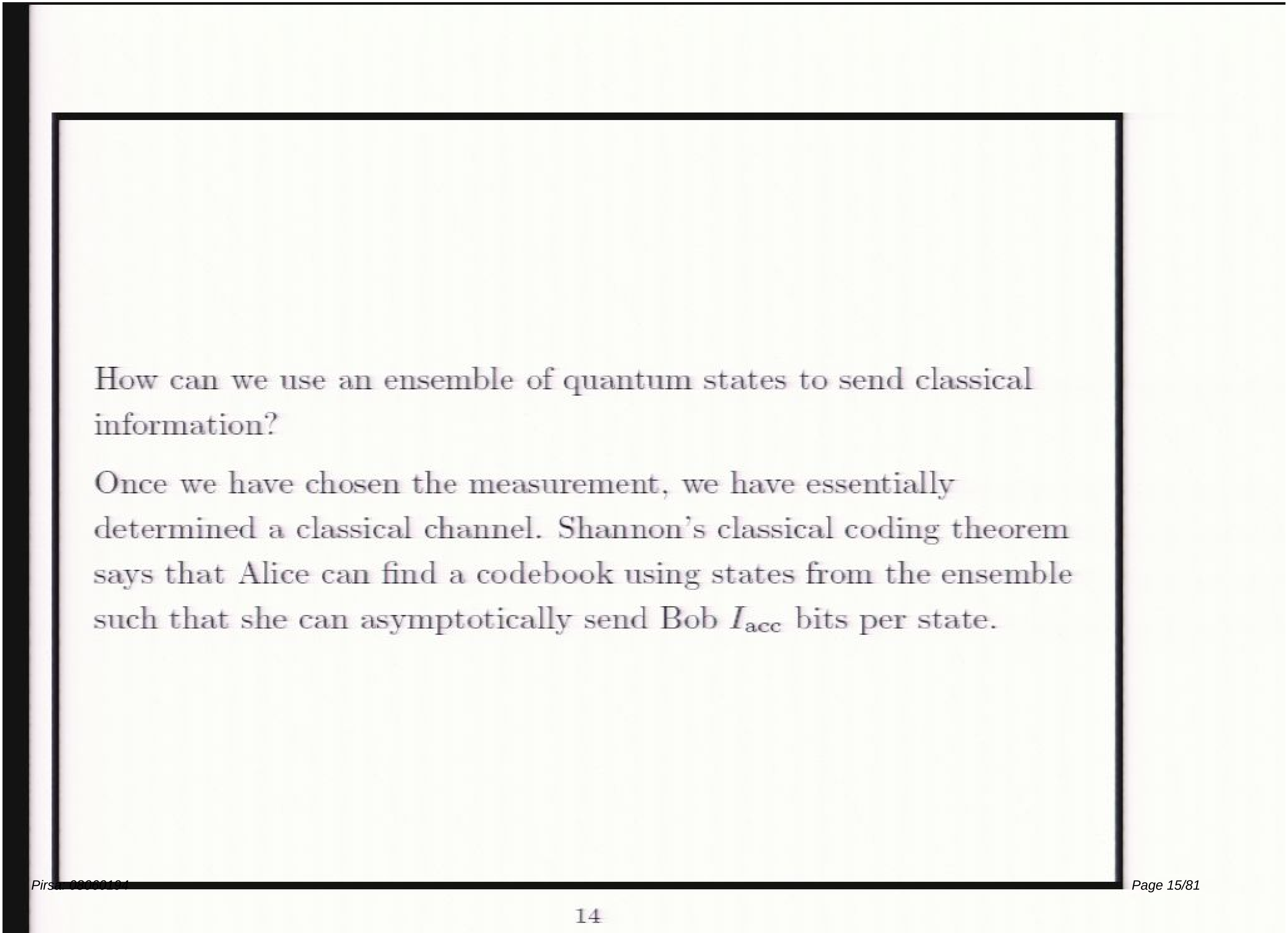
Suppose we have a source emitting ρ_i with probability p_i.

$$\chi = H(\sum_i p_i \rho_i) - \sum_i p_i H(\rho_i)$$

Theorem (Holevo, 1973)

$$I_{\text{acc}} \leq \chi$$

If all the ρ_i commute, the situation is essentially classical, and we get $I_{\text{acc}} = \chi$. Otherwise $I_{\text{acc}} < \chi$.



How can we use an ensemble of quantum states to send classical information?

Once we have chosen the measurement, we have essentially determined a classical channel. Shannon's classical coding theorem says that Alice can find a codebook using states from the ensemble such that she can asymptotically send Bob I_{acc} bits per state.

Example 2, Continued:

Suppose we use just two of the three signal states differing by 60°.

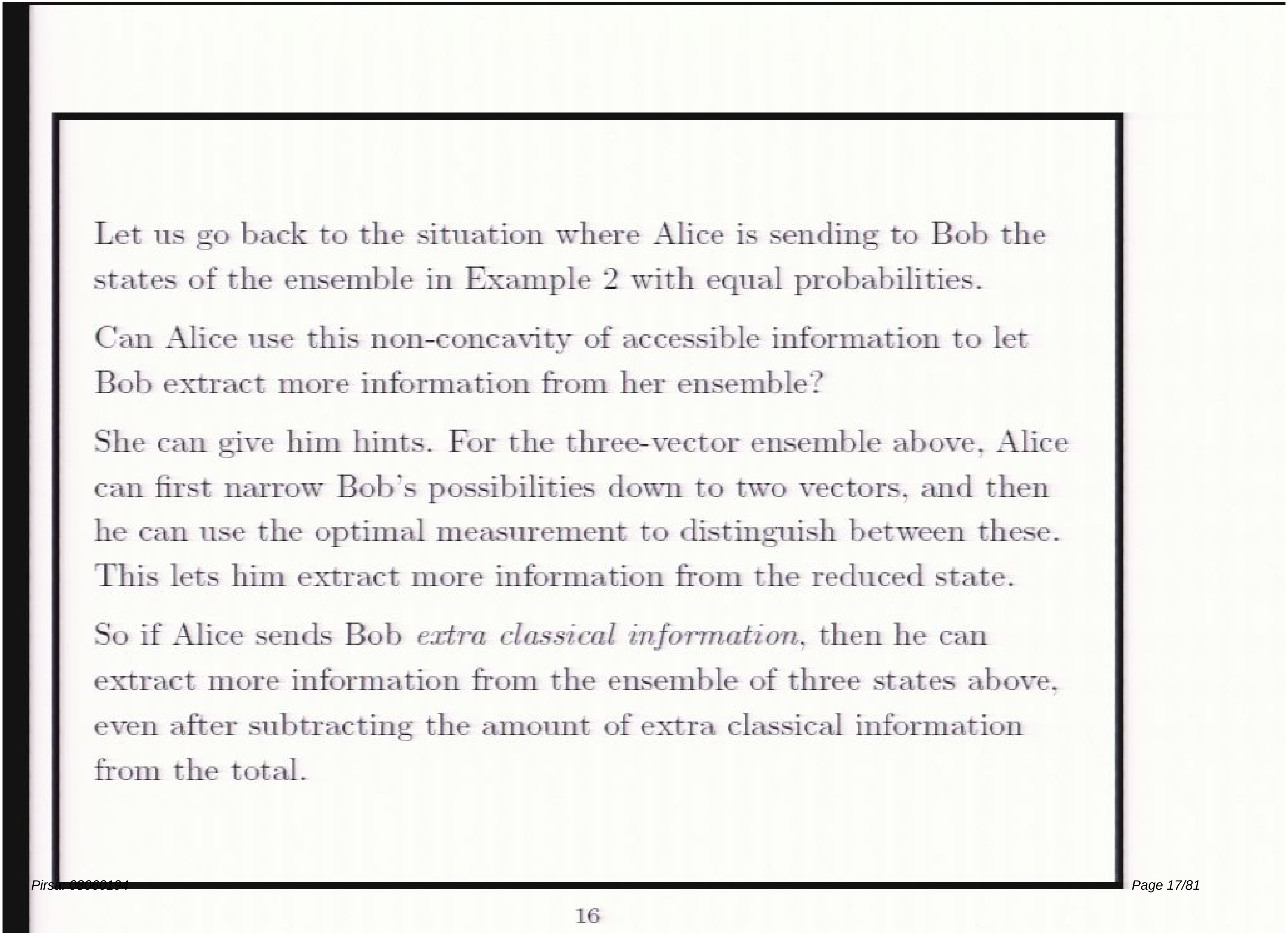
v_i: ↕ ⤡ (prob $\frac{1}{2}$)

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \end{pmatrix}$$

Optimal Measurement for two vectors:

$$I_{\text{acc}} = 1 - H(\tfrac{1}{2} + \tfrac{\sqrt{3}}{2}) = .6454 \text{ bits}$$

This is larger than the accessible information for the ensemble containing all three states with equal probability, showing that the accessible information is not concave.

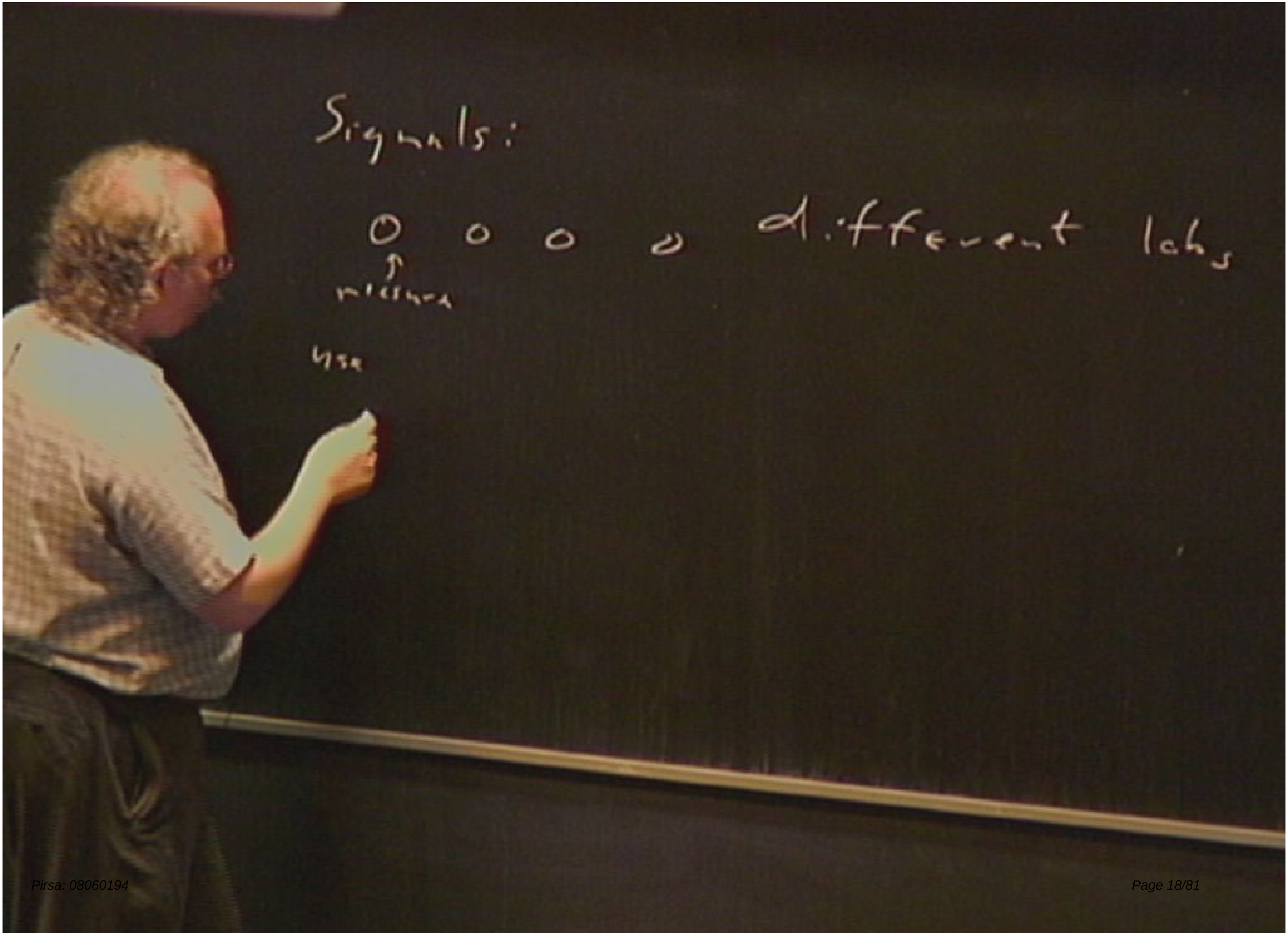


Let us go back to the situation where Alice is sending to Bob the states of the ensemble in Example 2 with equal probabilities.

Can Alice use this non-concavity of accessible information to let Bob extract more information from her ensemble?

She can give him hints. For the three-vector ensemble above, Alice can first narrow Bob's possibilities down to two vectors, and then he can use the optimal measurement to distinguish between these. This lets him extract more information from the reduced state.

So if Alice sends Bob *extra classical information*, then he can extract more information from the ensemble of three states above, even after subtracting the amount of extra classical information from the total.

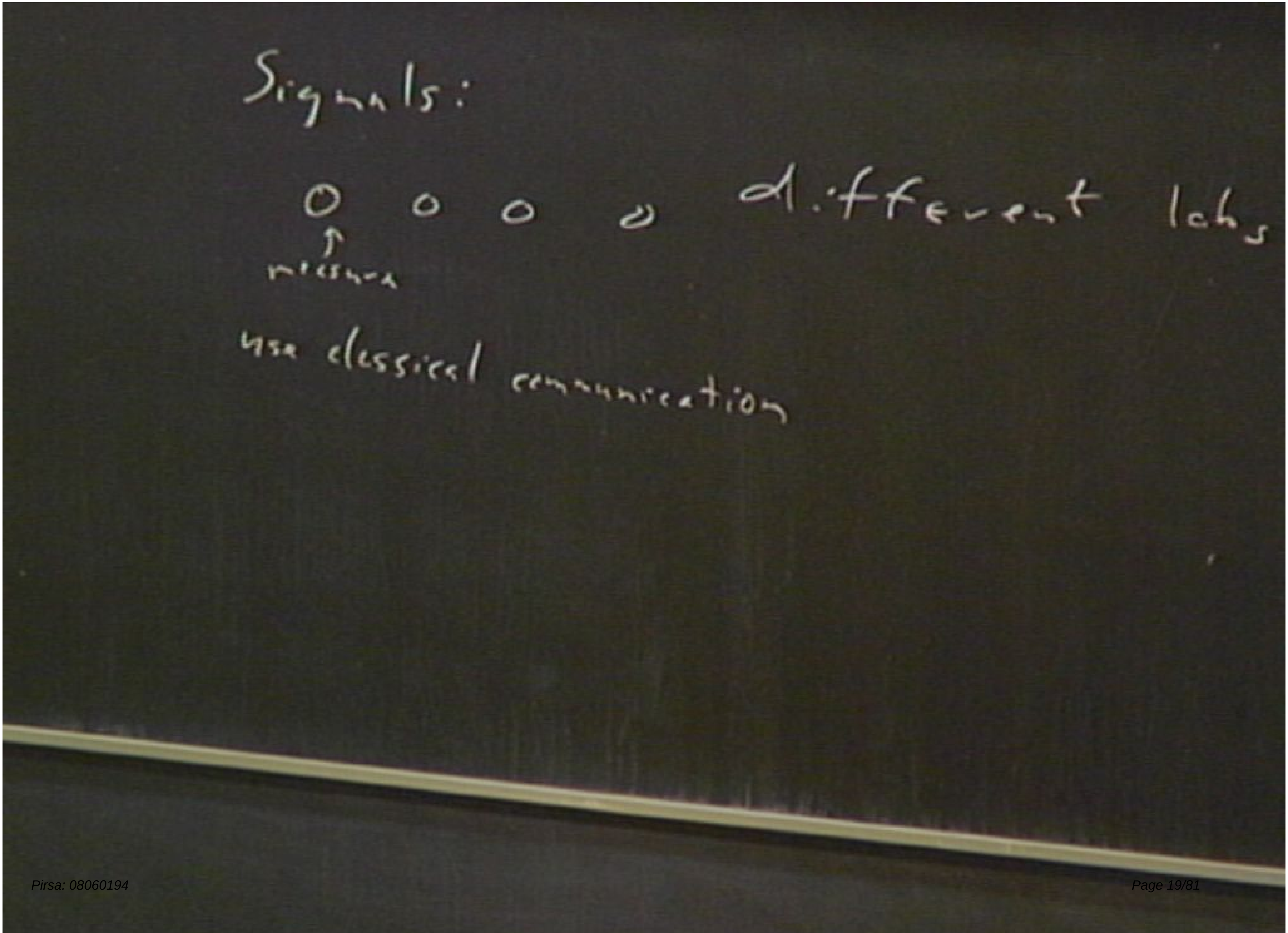


Signals:

O o o o different labs

↑ measured

use

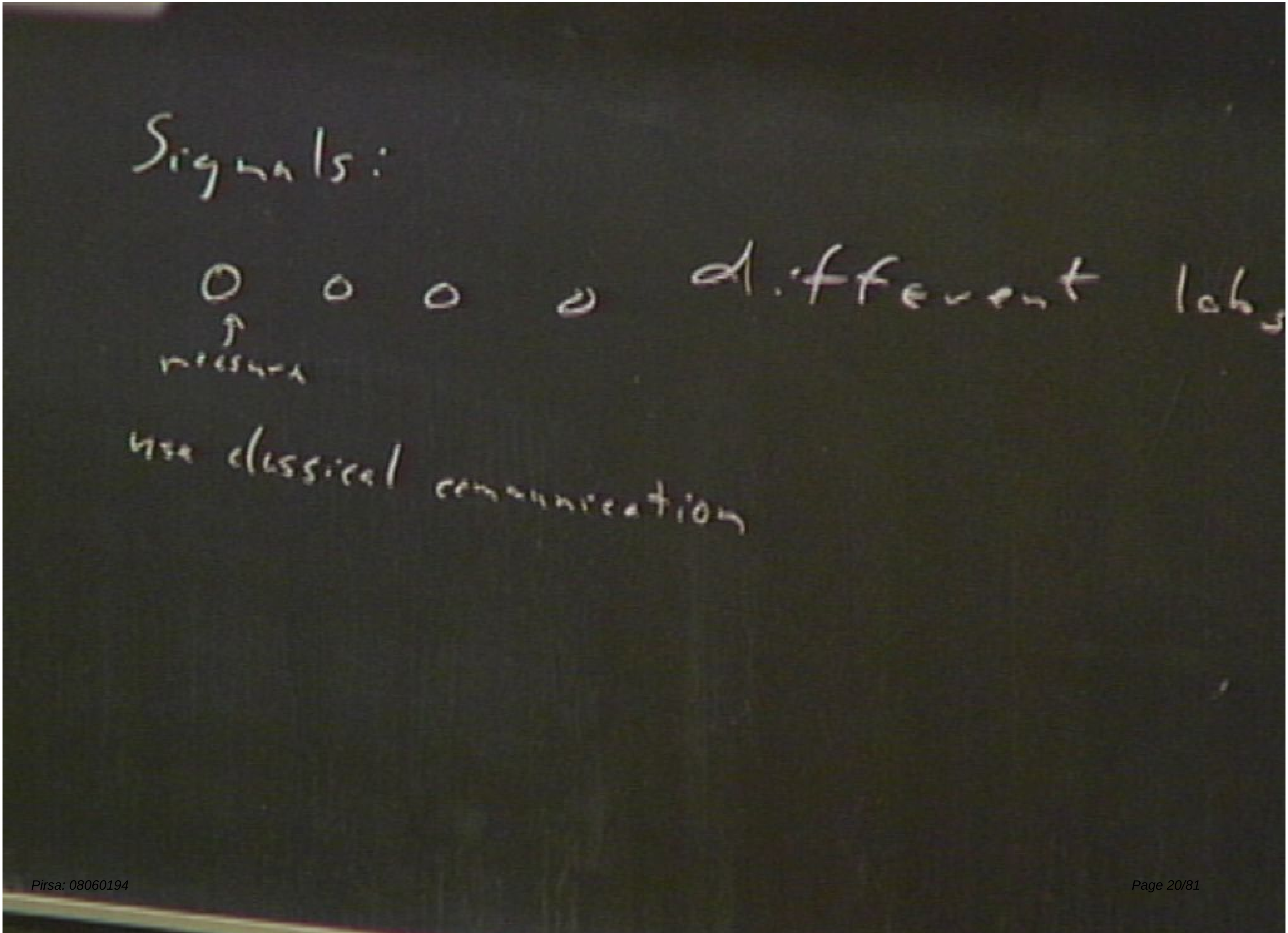


Signals:

○ ○ ○ ○ different labs

↑
measure

use classical communication



Signals:

○ ○ ○ ○ different labs

↑
measure

use classical communication

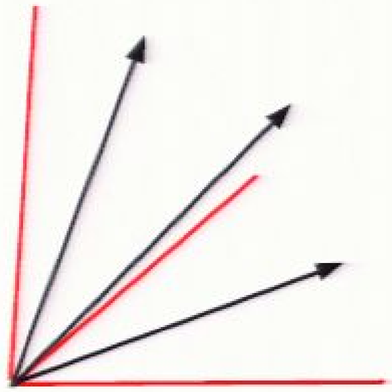
This situation does not occur for classical probability distributions. Suppose Alice has three labels, each corresponding to a probability distribution on Bob's classically correlated information. If Alice sends Bob more information about the label, he can now necessarily extract less information from the reduced state.

Alice's extra information tells Bob how to make a better measurement.

Classically, Bob has complete information, so extra information can't help.

How to do better: use three codewords $v_1 \otimes v_1$, $v_2 \otimes v_2$, $v_3 \otimes v_3$.

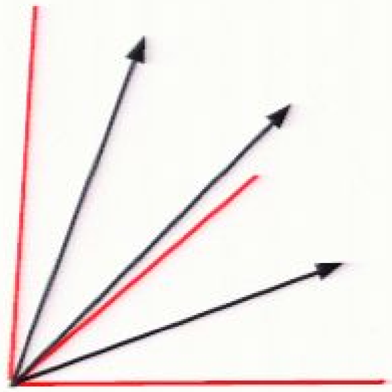
The optimal measurement for these three states gives 1.369 bits, which is larger than $2 \cdot 0.6545 = 1.309$ bits.



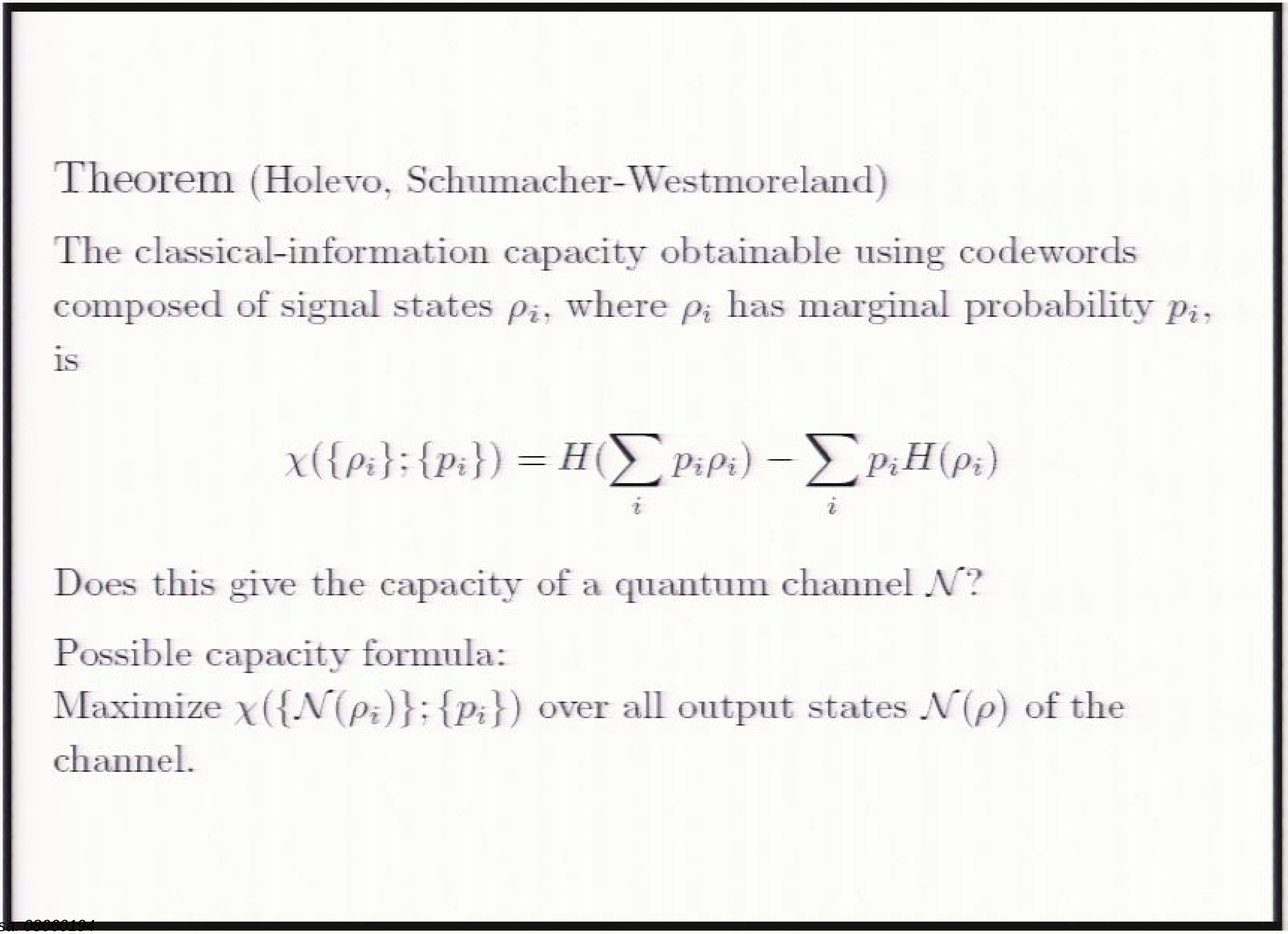
What about still longer codewords?

How to do better: use three codewords $v_1 \otimes v_1$, $v_2 \otimes v_2$, $v_3 \otimes v_3$.

The optimal measurement for these three states gives 1.369 bits, which is larger than $2 \cdot 0.6545 = 1.309$ bits.



What about still longer codewords?



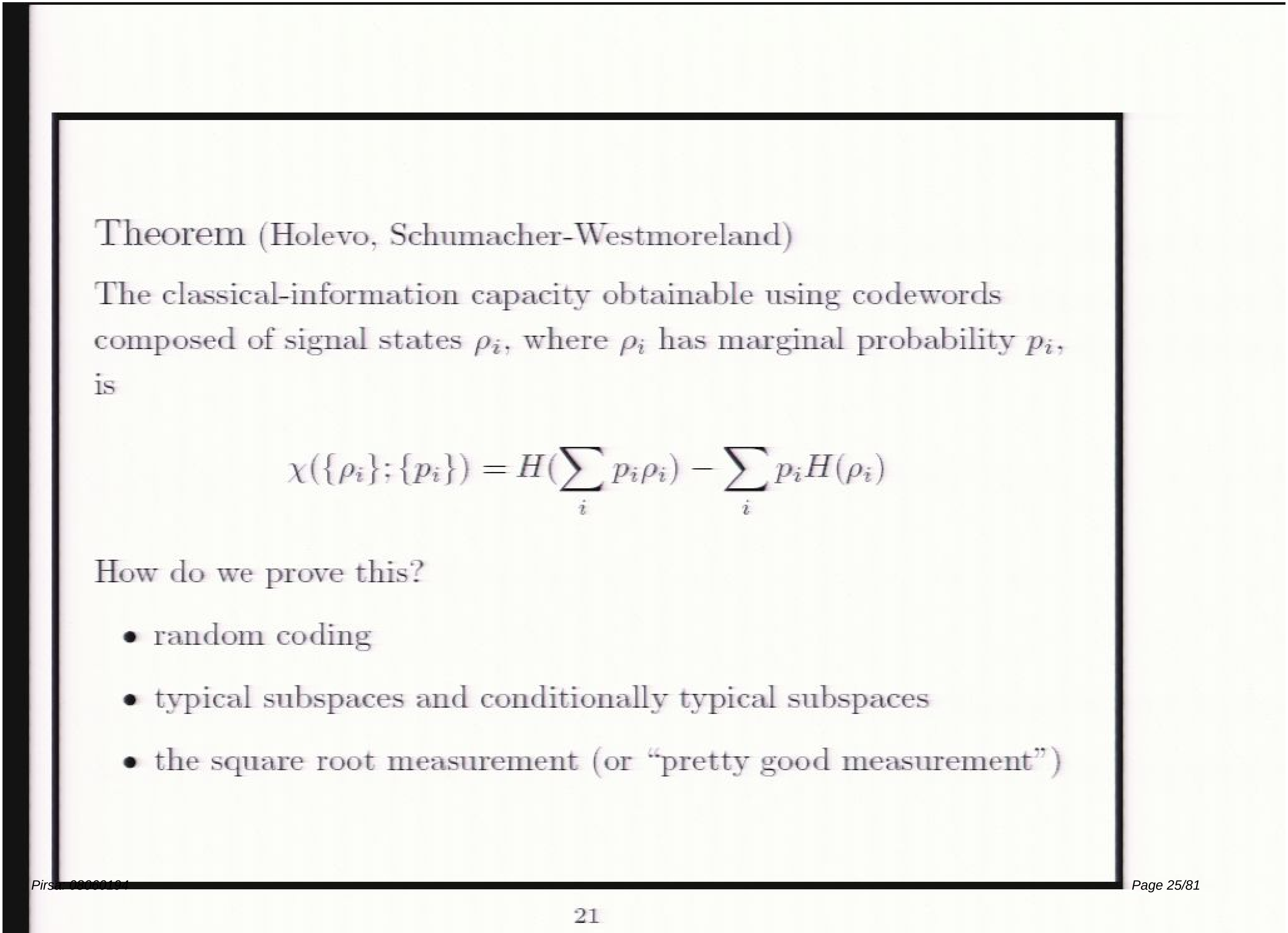
Theorem (Holevo, Schumacher-Westmoreland)

The classical-information capacity obtainable using codewords composed of signal states ρ_i, where ρ_i has marginal probability p_i, is

$$\chi(\{\rho_i\};\{p_i\}) = H(\sum_i p_i\rho_i) - \sum_i p_i H(\rho_i)$$

Does this give the capacity of a quantum channel $\mathcal{N}$?

Possible capacity formula:
Maximize $\chi(\{\mathcal{N}(\rho_i)\};\{p_i\})$ over all output states $\mathcal{N}(\rho)$ of the channel.



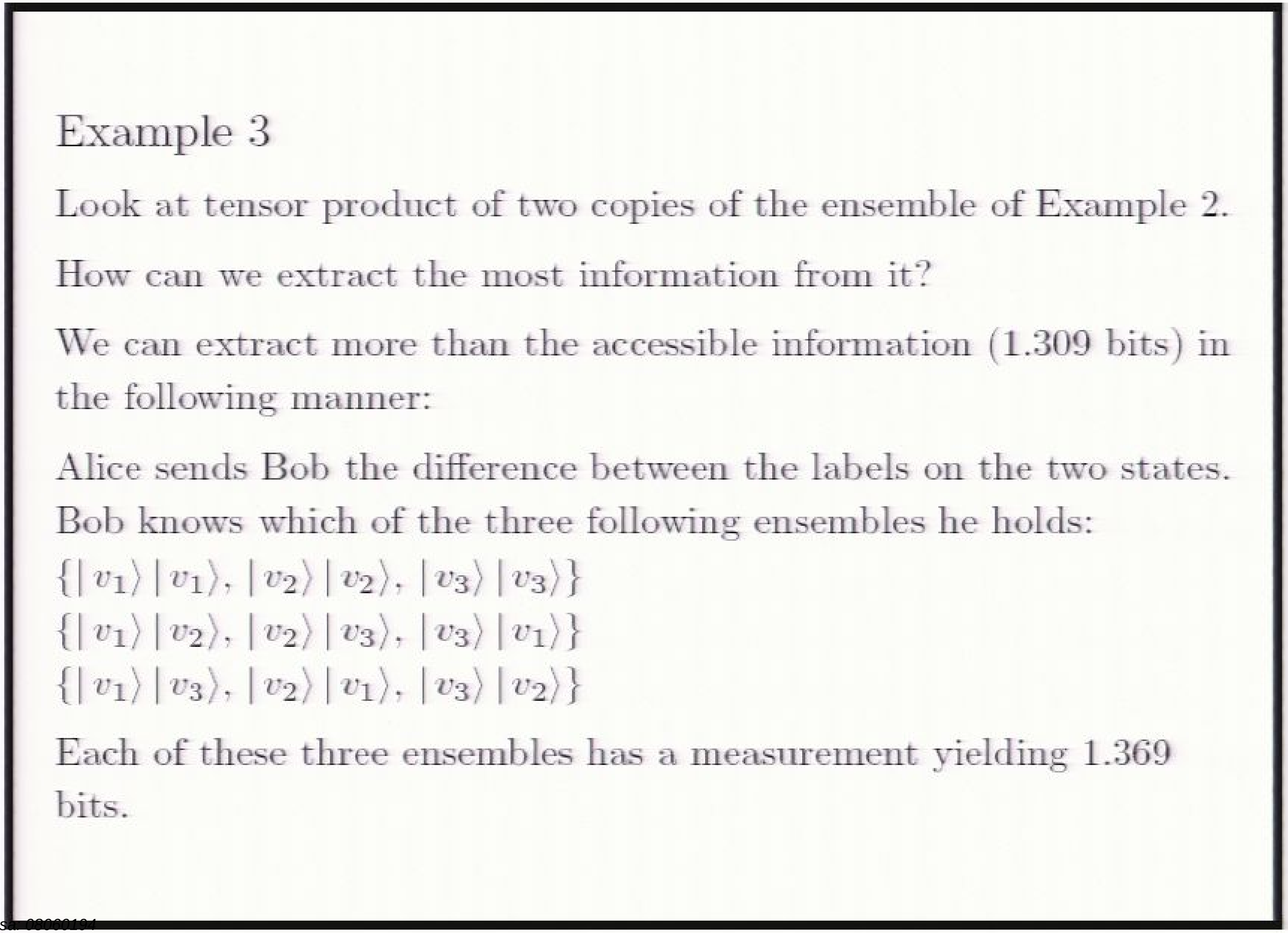
Theorem (Holevo, Schumacher-Westmoreland)

The classical-information capacity obtainable using codewords composed of signal states ρ_i, where ρ_i has marginal probability p_i, is

$$\chi(\{\rho_i\};\{p_i\}) = H(\sum_i p_i\rho_i) - \sum_i p_i H(\rho_i)$$

How do we prove this?

- random coding
- typical subspaces and conditionally typical subspaces
- the square root measurement (or "pretty good measurement")



Example 3

Look at tensor product of two copies of the ensemble of Example 2.

How can we extract the most information from it?

We can extract more than the accessible information (1.309 bits) in the following manner:

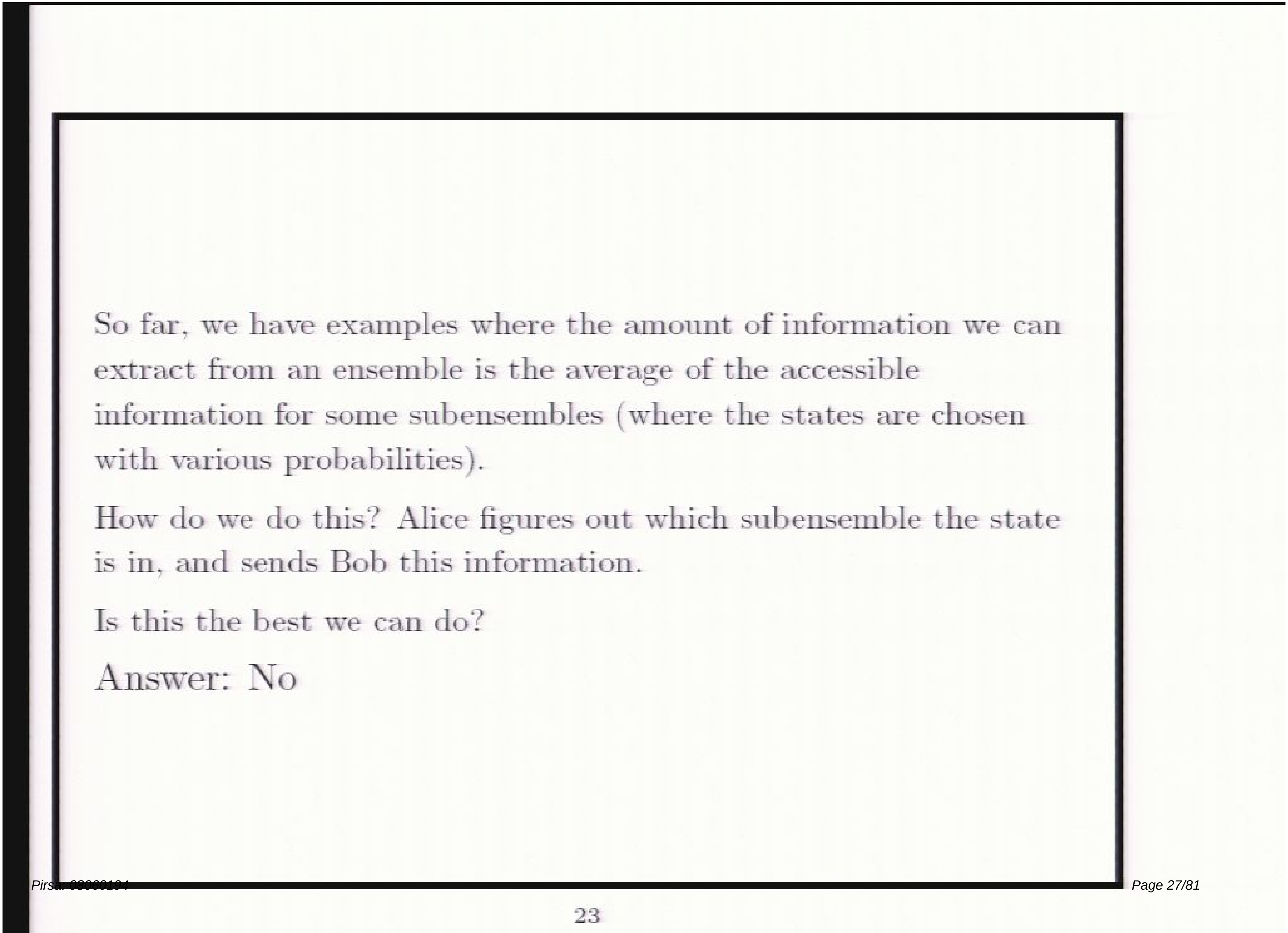
Alice sends Bob the difference between the labels on the two states. Bob knows which of the three following ensembles he holds:

$\{|\, v_1\rangle\,|\, v_1\rangle,\ |\, v_2\rangle\,|\, v_2\rangle,\ |\, v_3\rangle\,|\, v_3\rangle\}$

$\{|\, v_1\rangle\,|\, v_2\rangle,\ |\, v_2\rangle\,|\, v_3\rangle,\ |\, v_3\rangle\,|\, v_1\rangle\}$

$\{|\, v_1\rangle\,|\, v_3\rangle,\ |\, v_2\rangle\,|\, v_1\rangle,\ |\, v_3\rangle\,|\, v_2\rangle\}$

Each of these three ensembles has a measurement yielding 1.369 bits.

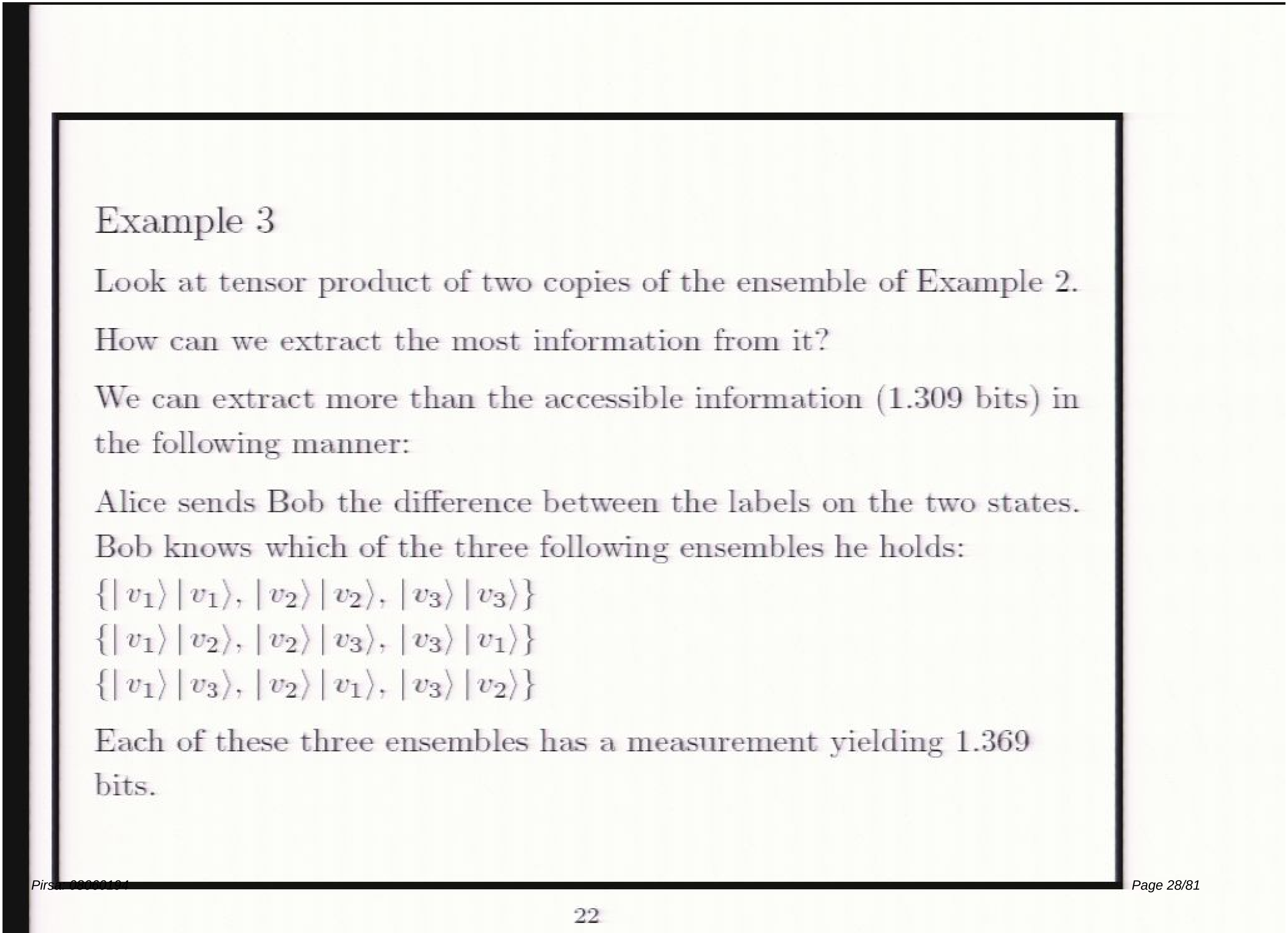


So far, we have examples where the amount of information we can extract from an ensemble is the average of the accessible information for some subensembles (where the states are chosen with various probabilities).

How do we do this? Alice figures out which subensemble the state is in, and sends Bob this information.

Is this the best we can do?

Answer: No



Example 3

Look at tensor product of two copies of the ensemble of Example 2.

How can we extract the most information from it?

We can extract more than the accessible information (1.309 bits) in the following manner:

Alice sends Bob the difference between the labels on the two states. Bob knows which of the three following ensembles he holds:
$\{|v_1\rangle\,|v_1\rangle,\ |v_2\rangle\,|v_2\rangle,\ |v_3\rangle\,|v_3\rangle\}$
$\{|v_1\rangle\,|v_2\rangle,\ |v_2\rangle\,|v_3\rangle,\ |v_3\rangle\,|v_1\rangle\}$
$\{|v_1\rangle\,|v_3\rangle,\ |v_2\rangle\,|v_1\rangle,\ |v_3\rangle\,|v_2\rangle\}$

Each of these three ensembles has a measurement yielding 1.369 bits.

So far, we have examples where the amount of information we can extract from an ensemble is the average of the accessible information for some subensembles (where the states are chosen with various probabilities).

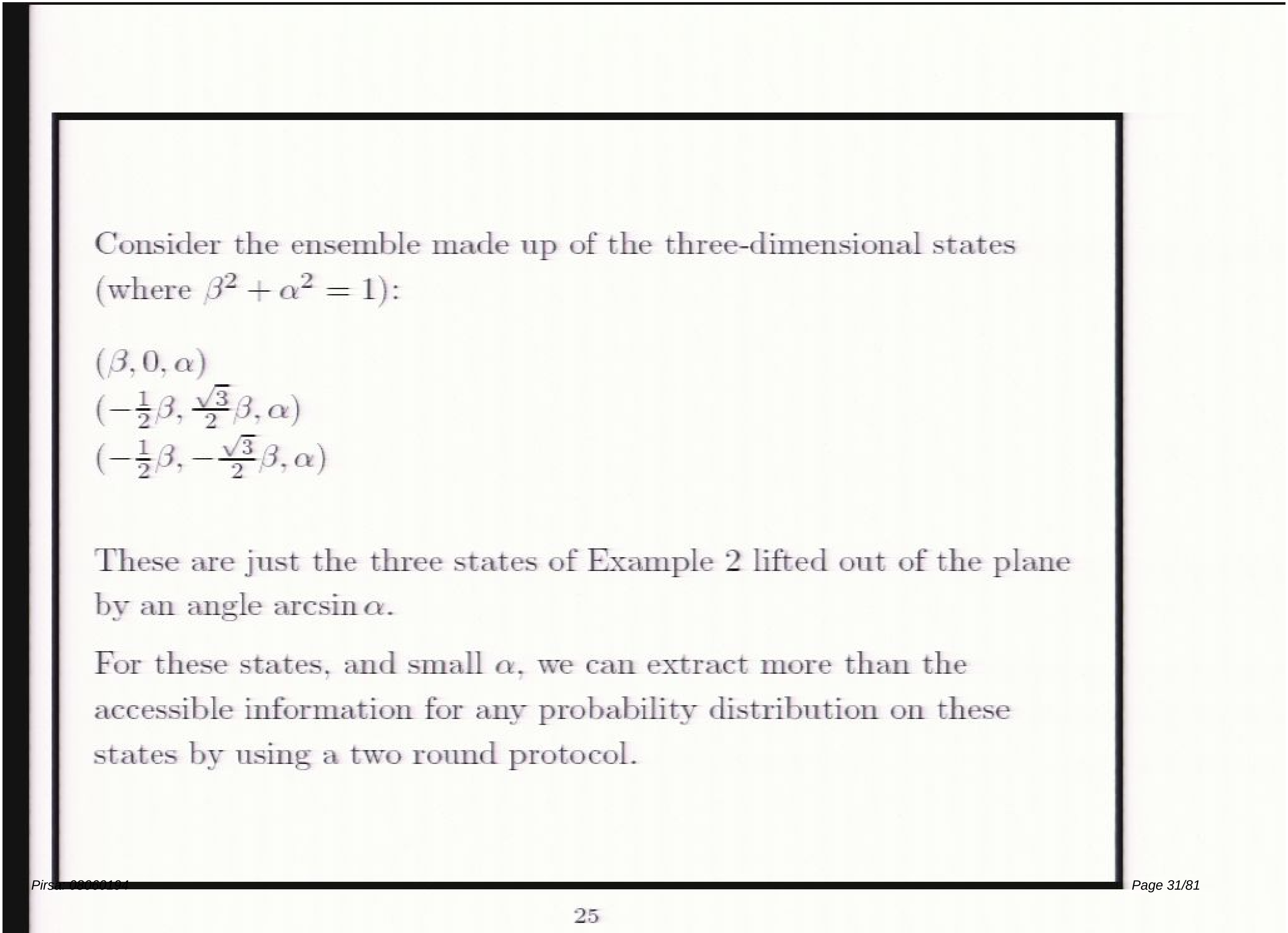
How do we do this? Alice figures out which subensemble the state is in, and sends Bob this information.

Is this the best we can do?

Answer: No

How do we do better?

We have an example where Bob can do better using two measurements sequentially.

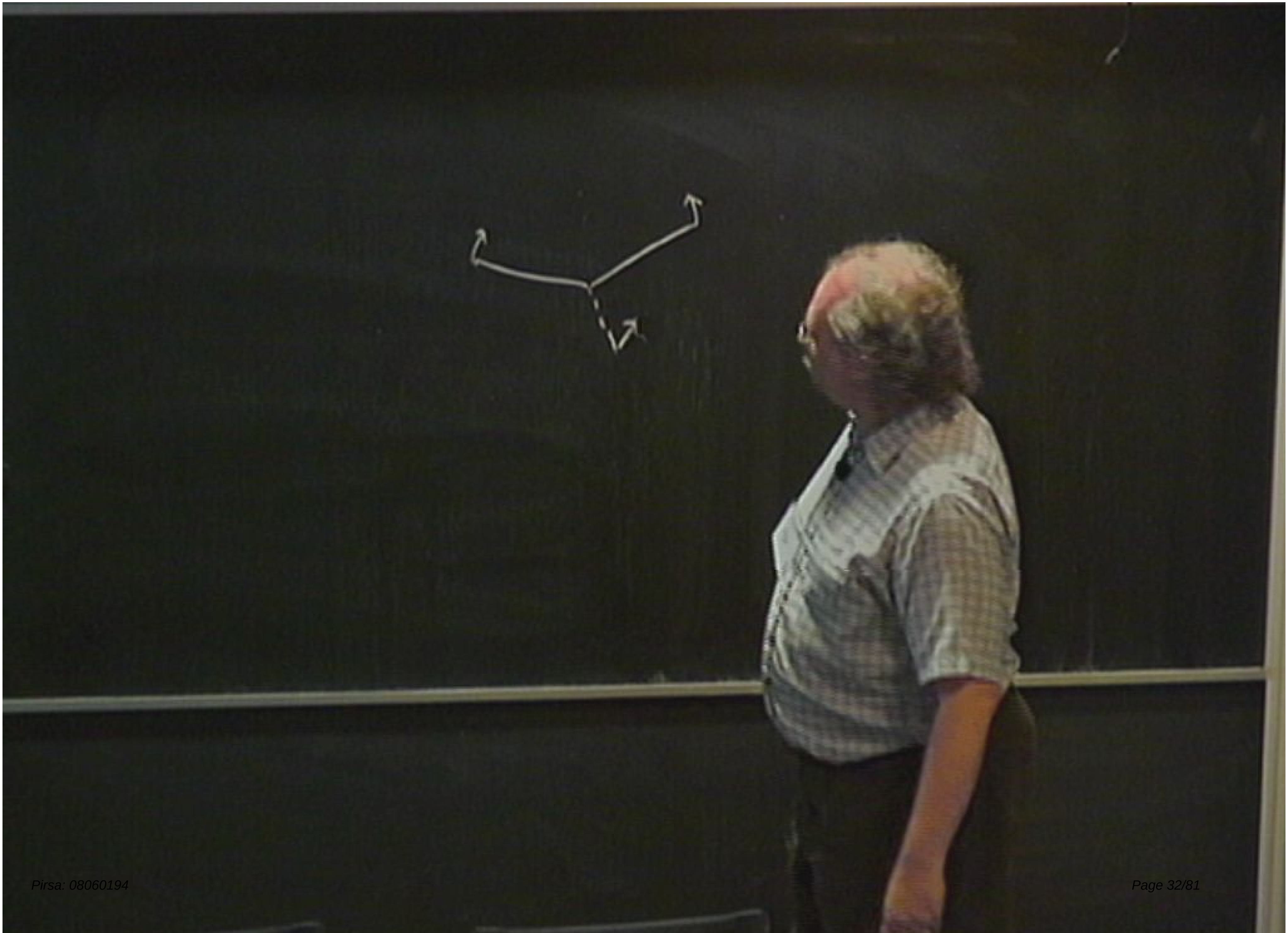


Consider the ensemble made up of the three-dimensional states (where $\beta^2 + \alpha^2 = 1$):

$(\beta, 0, \alpha)$
$(-\frac{1}{2}\beta, \frac{\sqrt{3}}{2}\beta, \alpha)$
$(-\frac{1}{2}\beta, -\frac{\sqrt{3}}{2}\beta, \alpha)$

These are just the three states of Example 2 lifted out of the plane by an angle $\arcsin \alpha$.

For these states, and small α, we can extract more than the accessible information for any probability distribution on these states by using a two round protocol.



Consider the ensemble made up of the three-dimensional states (where $\beta^2 + \alpha^2 = 1$):

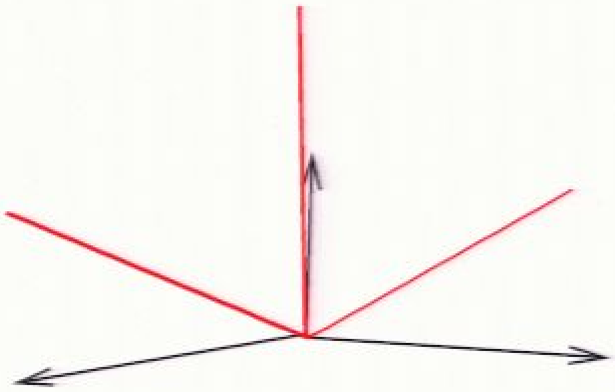
$(\beta, 0, \alpha)$
$(-\frac{1}{2}\beta, \frac{\sqrt{3}}{2}\beta, \alpha)$
$(-\frac{1}{2}\beta, -\frac{\sqrt{3}}{2}\beta, \alpha)$

These are just the three states of Example 2 lifted out of the plane by an angle $\arcsin \alpha$.

For these states, and small α, we can extract more than the accessible information for any probability distribution on these states by using a two round protocol.

First, Bob makes a measurement which either projects these three trine states down into the plane (Example 2) or lifts them further out of the plane (akin to example 3 but with different angles).

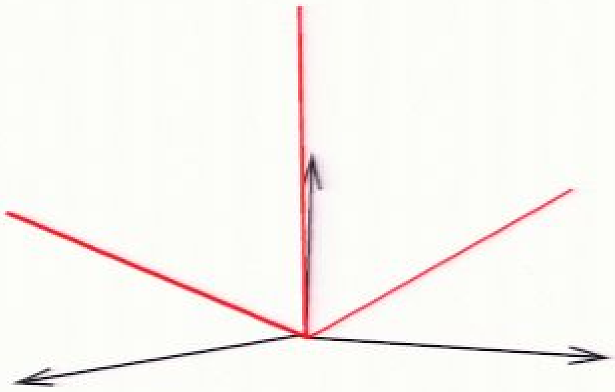
If they're lifted out of the plane, Bob makes the optimal measurement distinguishing all three of them.



If they're projected into the plane, Bob asks Alice to narrow the choice down to two possibilities, and then makes the optimal measurement distinguishing them (Example 1).

First, Bob makes a measurement which either projects these three trine states down into the plane (Example 2) or lifts them further out of the plane (akin to example 3 but with different angles).

If they're lifted out of the plane, Bob makes the optimal measurement distinguishing all three of them.



If they're projected into the plane, Bob asks Alice to narrow the choice down to two possibilities, and then makes the optimal measurement distinguishing them (Example 1).

Can we use this protocol for extracting information for communication of classical information using quantum states chosen from this ensemble?

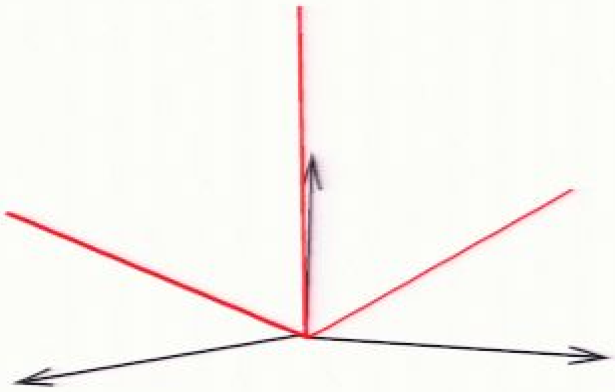
Answer: Yes

We need to show that Alice can use this protocol without a feedback channel from Bob to Alice.

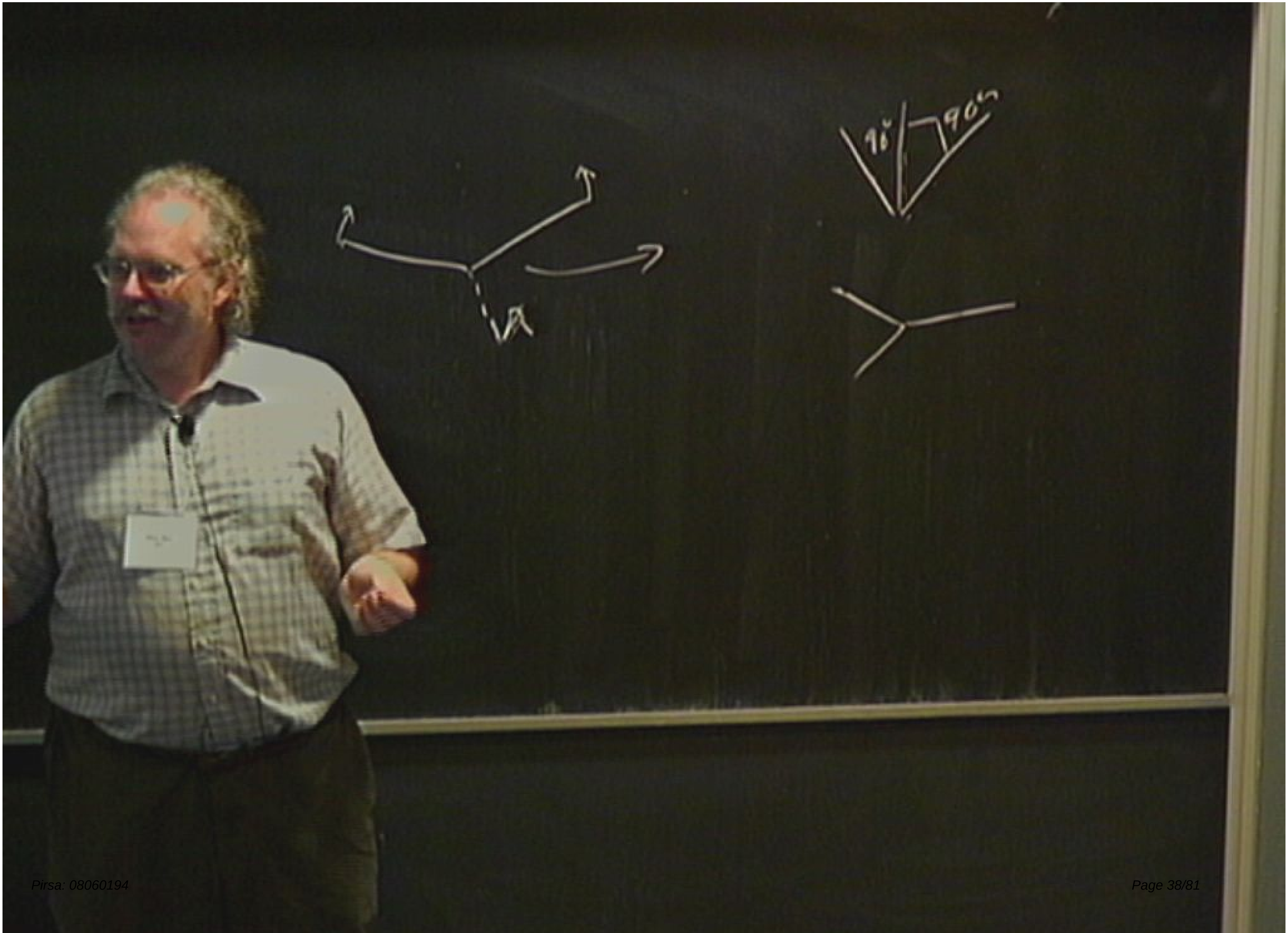
We use two classical codes, one corresponding to each of Bob's measurement steps, and combine them by bitwise addition (mod 3). By the fact that random classical codes achieve Shannon's channel capacity, this means that Bob can decode Alice's message.

First, Bob makes a measurement which either projects these three trine states down into the plane (Example 2) or lifts them further out of the plane (akin to example 3 but with different angles).

If they're lifted out of the plane, Bob makes the optimal measurement distinguishing all three of them.

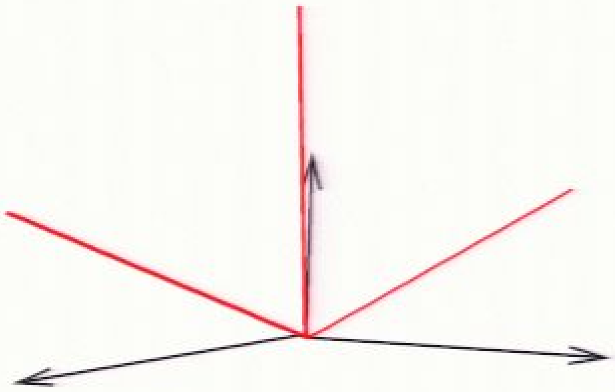


If they're projected into the plane, Bob asks Alice to narrow the choice down to two possibilities, and then makes the optimal measurement distinguishing them (Example 1).


√2
90°
90°

First, Bob makes a measurement which either projects these three trine states down into the plane (Example 2) or lifts them further out of the plane (akin to example 3 but with different angles).

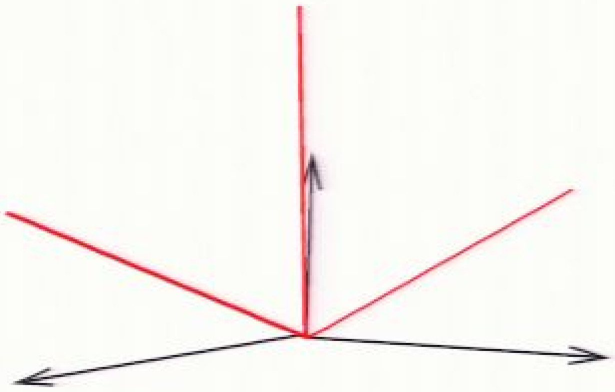
If they're lifted out of the plane, Bob makes the optimal measurement distinguishing all three of them.



If they're projected into the plane, Bob asks Alice to narrow the choice down to two possibilities, and then makes the optimal measurement distinguishing them (Example 1).

First, Bob makes a measurement which either projects these three trine states down into the plane (Example 2) or lifts them further out of the plane (akin to example 3 but with different angles).

If they're lifted out of the plane, Bob makes the optimal measurement distinguishing all three of them.



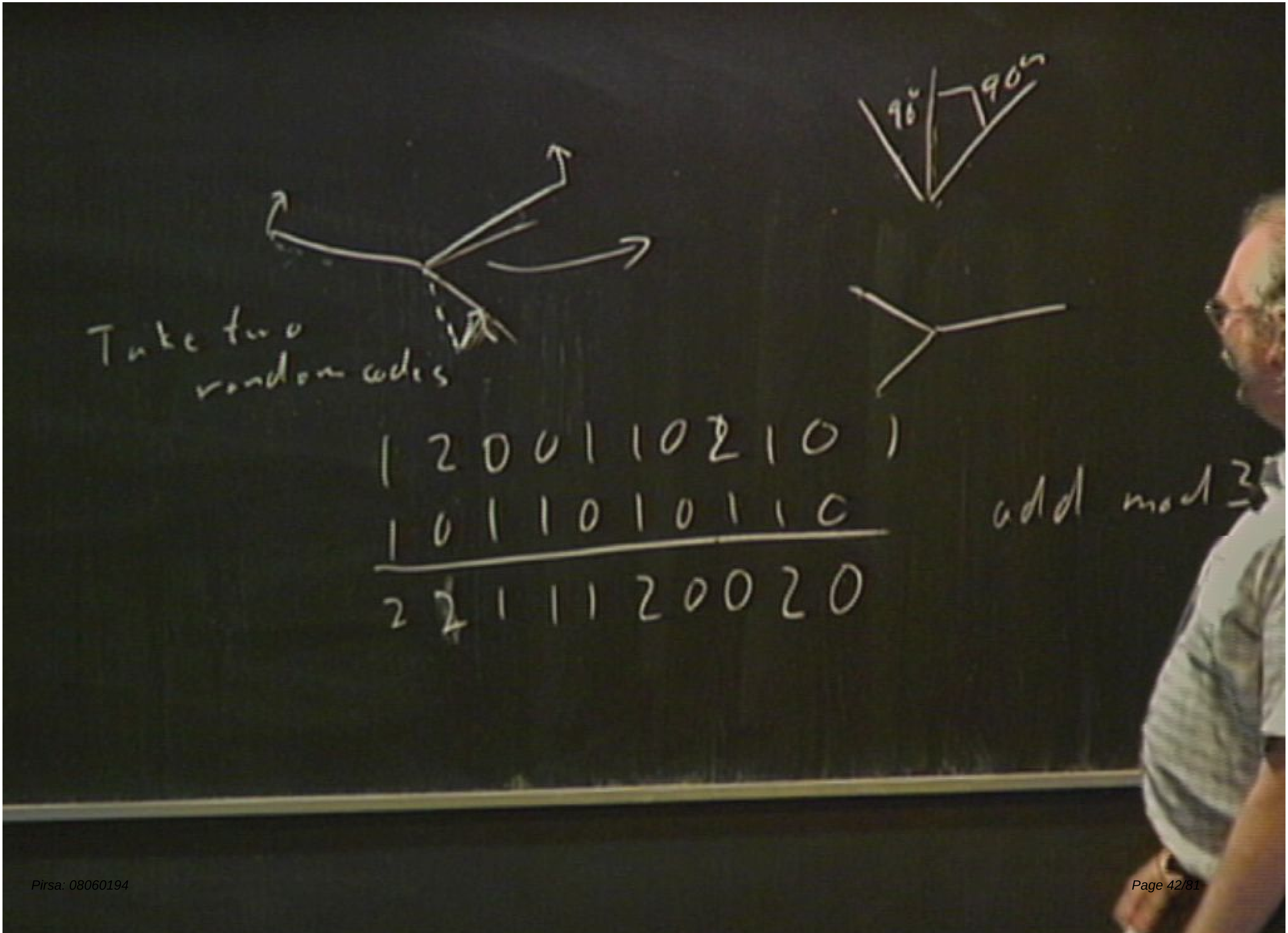
If they're projected into the plane, Bob asks Alice to narrow the choice down to two possibilities, and then makes the optimal measurement distinguishing them (Example 1).

Can we use this protocol for extracting information for communication of classical information using quantum states chosen from this ensemble?

Answer: Yes

We need to show that Alice can use this protocol without a feedback channel from Bob to Alice.

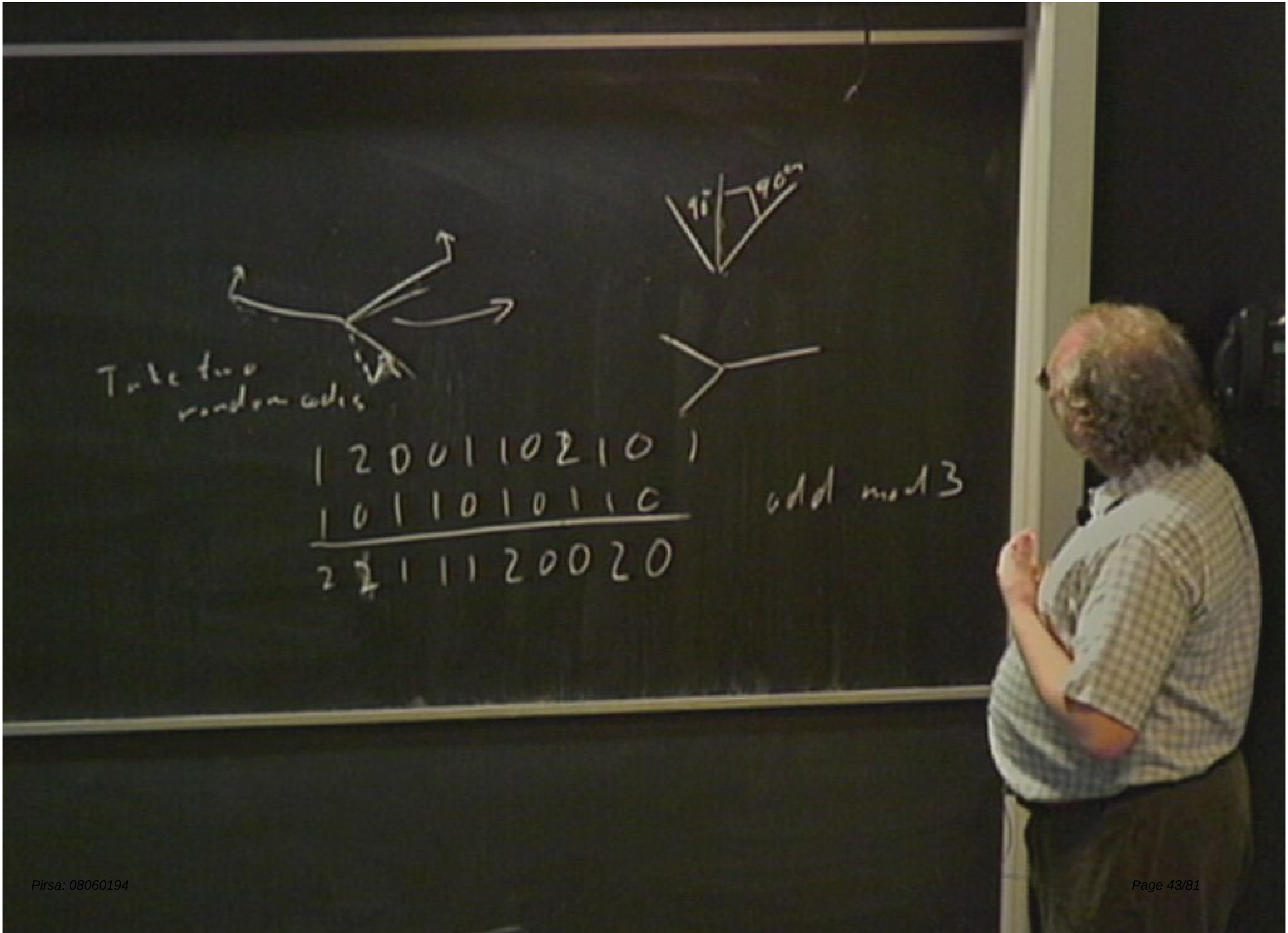
We use two classical codes, one corresponding to each of Bob's measurement steps, and combine them by bitwise addition (mod 3). By the fact that random classical codes achieve Shannon's channel capacity, this means that Bob can decode Alice's message.



Take two random codes

1 2 0 0 1 1 0 2 1 0 1
1 0 1 1 0 1 0 1 1 0
2 2 1 1 1 2 0 0 2 0

add mod 3

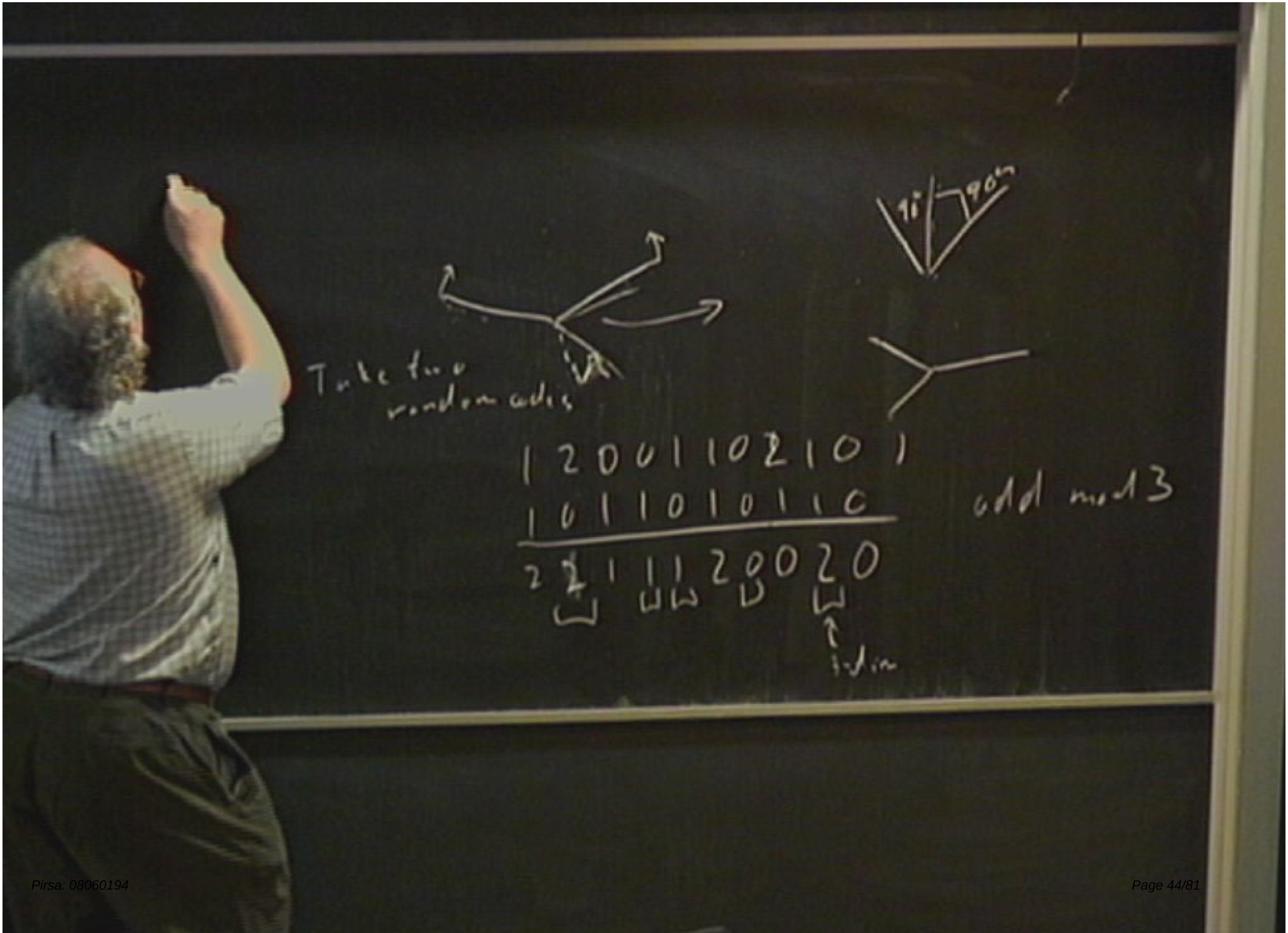


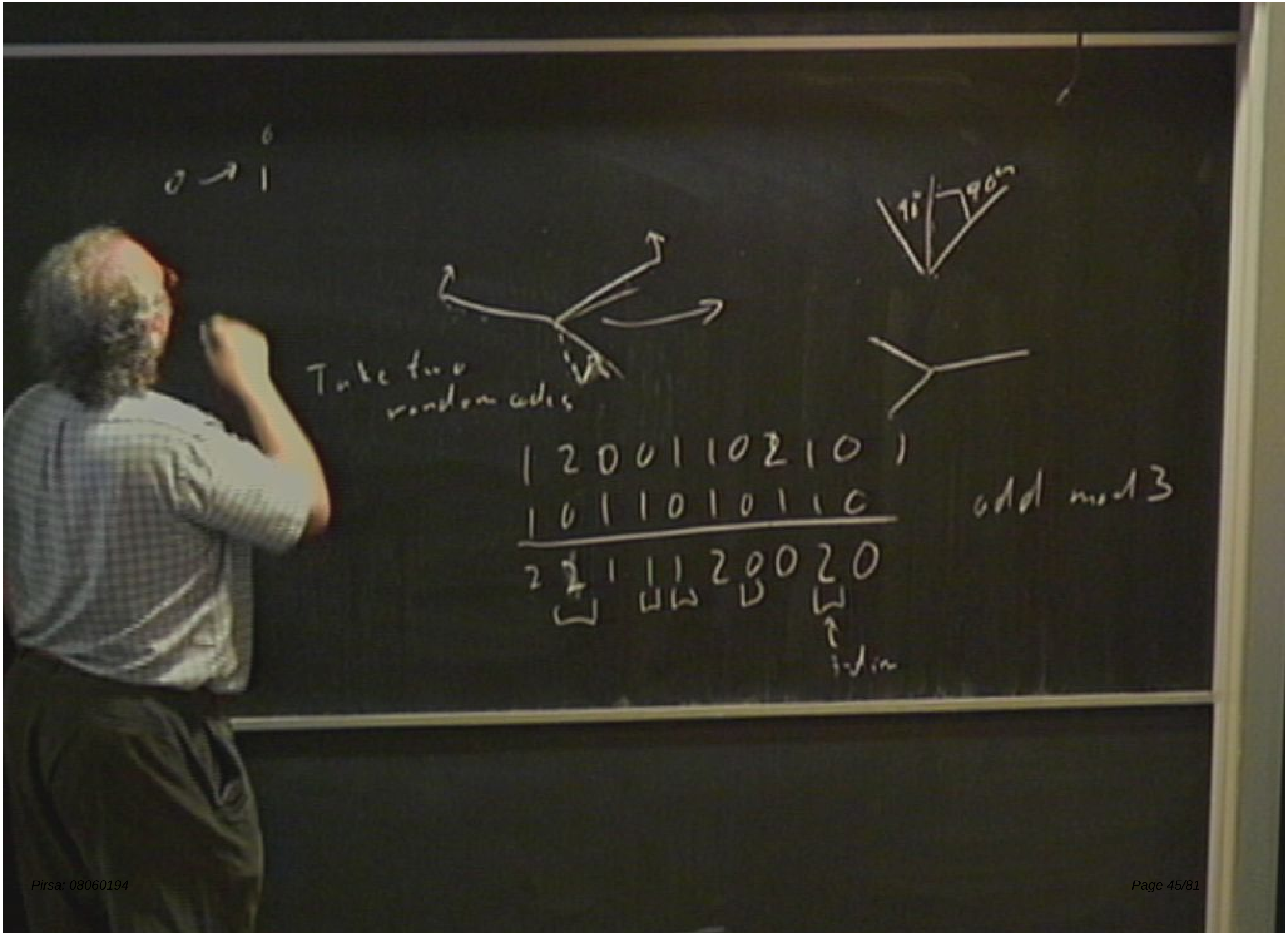
Take two random codes

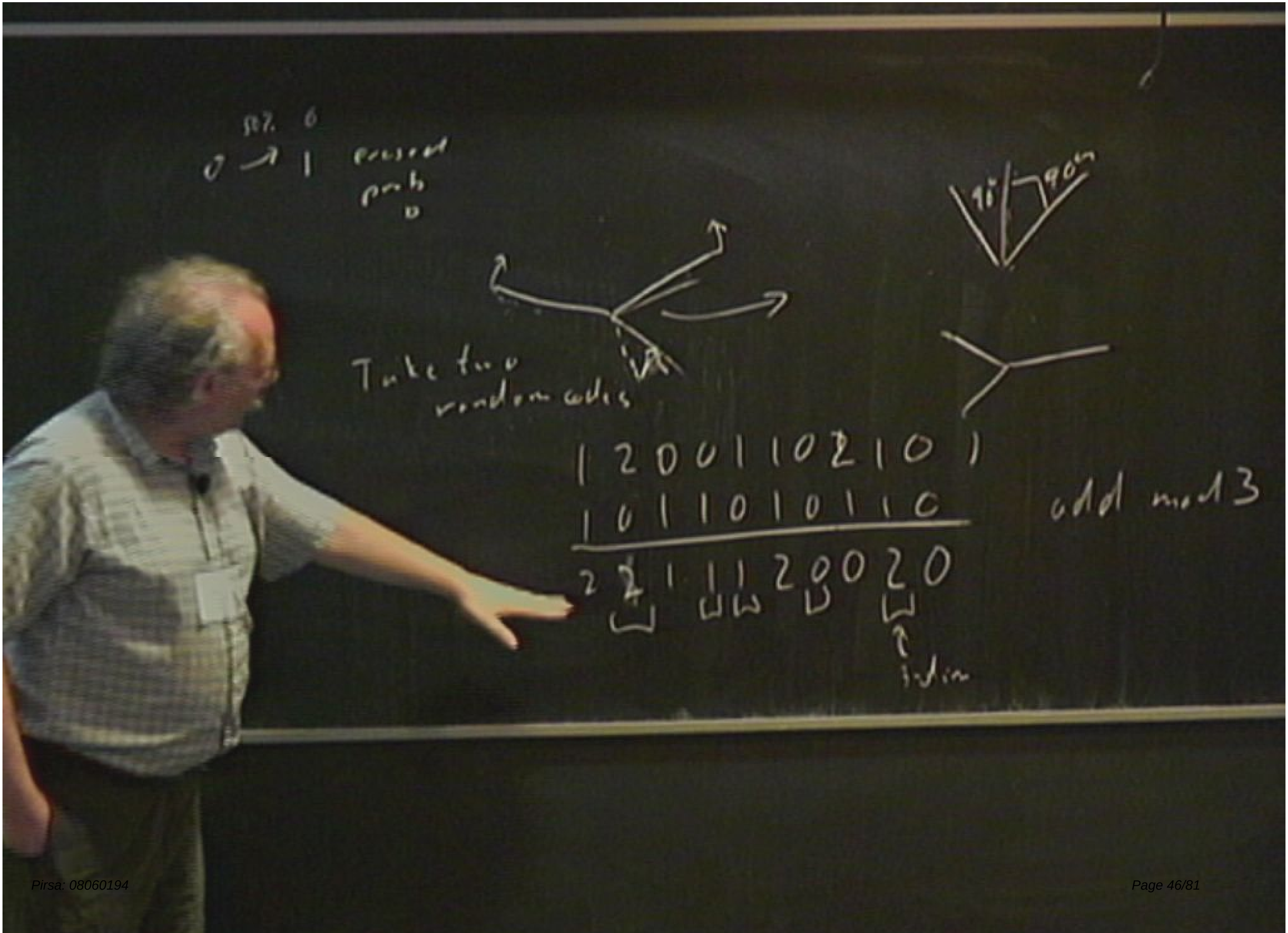
1 2 0 0 1 1 0 2 1 0 1
1 0 1 1 0 1 0 1 1 0

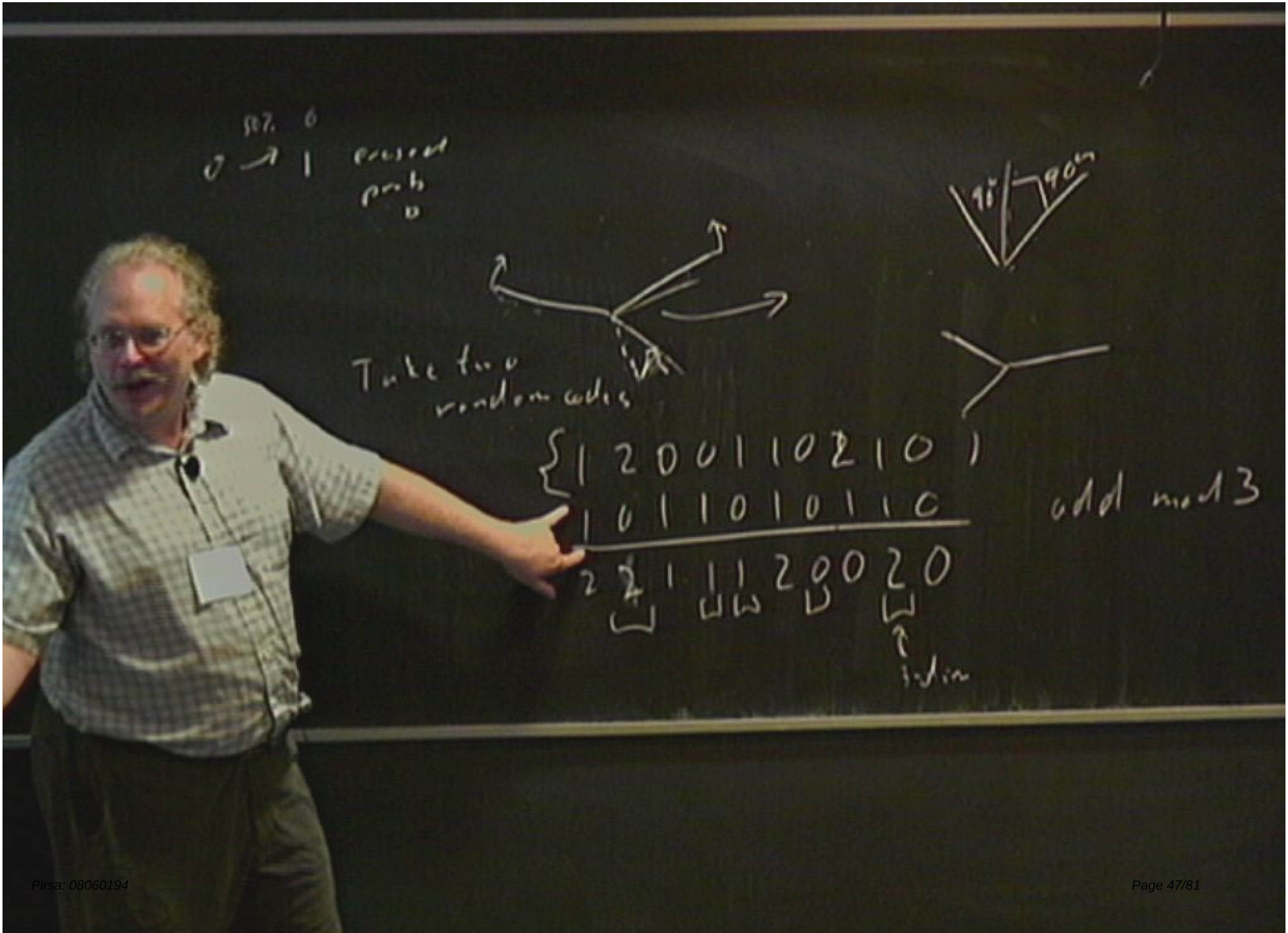
2 2 1 1 1 2 0 0 2 0

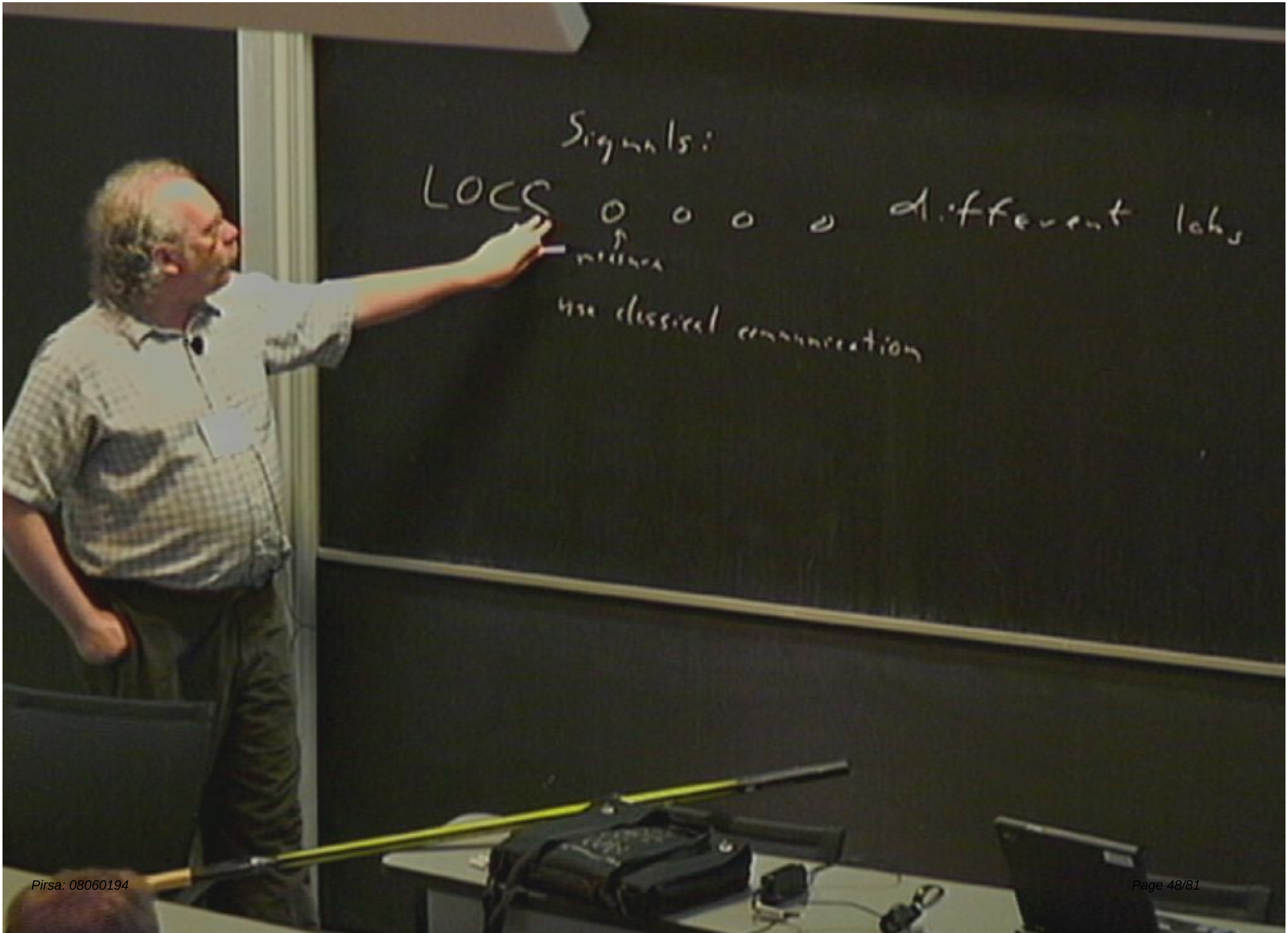
odd mod 3


Take two random codes
1 2 0 0 1 1 1 0 2 1 0 1
1 0 1 1 0 1 0 1 1 0
2 2 1 1 1 2 0 0 2 0
add mod 3
91°
90°


0 → 1
0
91°
90°
Take two
random codes
1 2 0 0 1 1 0 2 1 0 1
1 0 1 1 0 1 0 1 1 0
2 2 1 1 1 2 0 0 2 0
add mod 3


5%   0
0 → 1 erased
prob
0
Take two random codes
1 2 0 0 1 1 1 0 2 1 0 1
1 0 1 1 0 1 0 1 1 0
2 2 1 1 1 2 0 0 2 0
add mod 3
90°
90°


90°
90°
0 → 1 erased
path
Take two
random codes
1 2 0 0 1 1 1 0 2 1 0 1
1 0 1 1 0 1 0 1 1 1 0
2 2 1 1 1 2 0 0 0 2 0
add mod 3

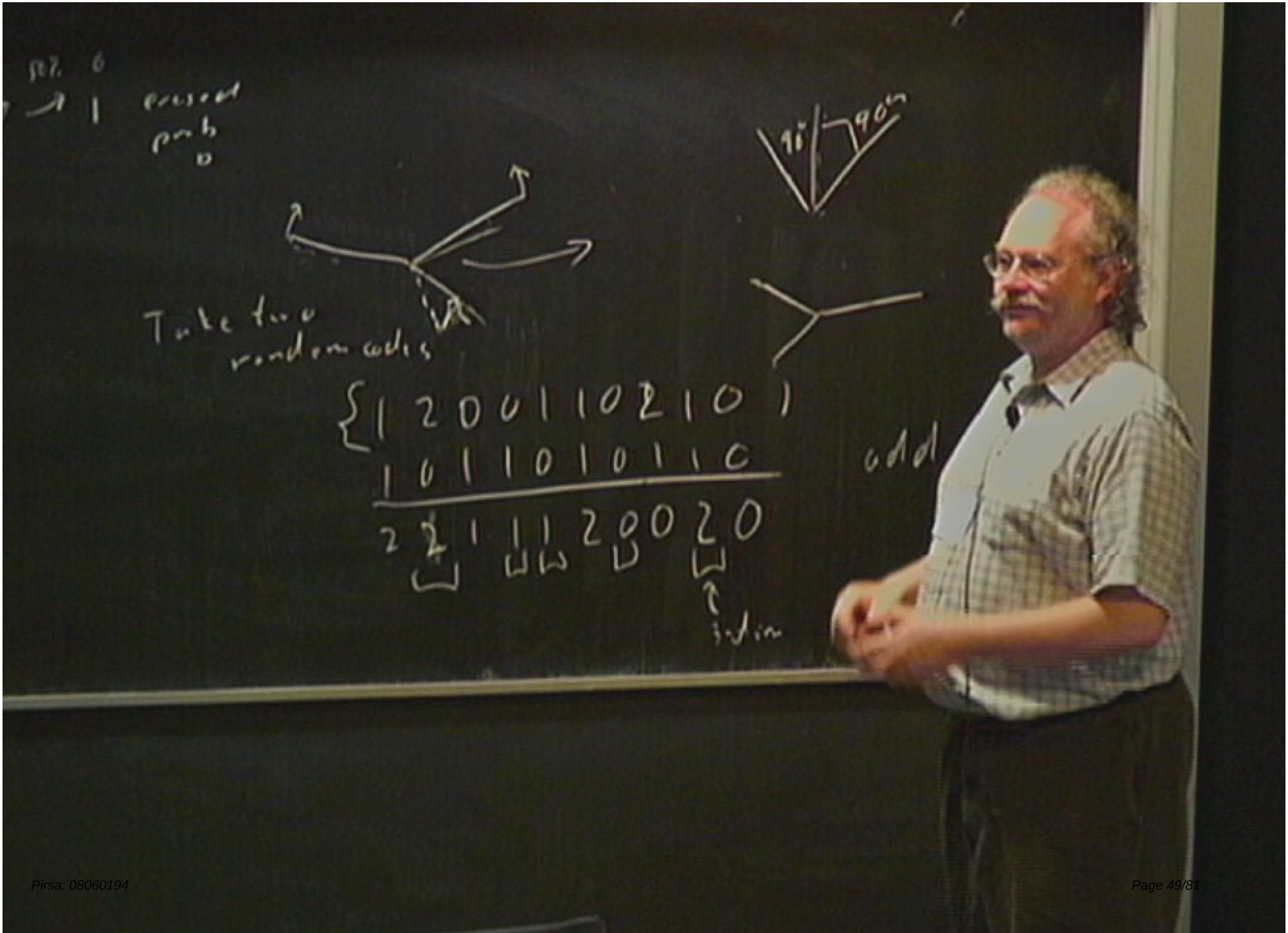


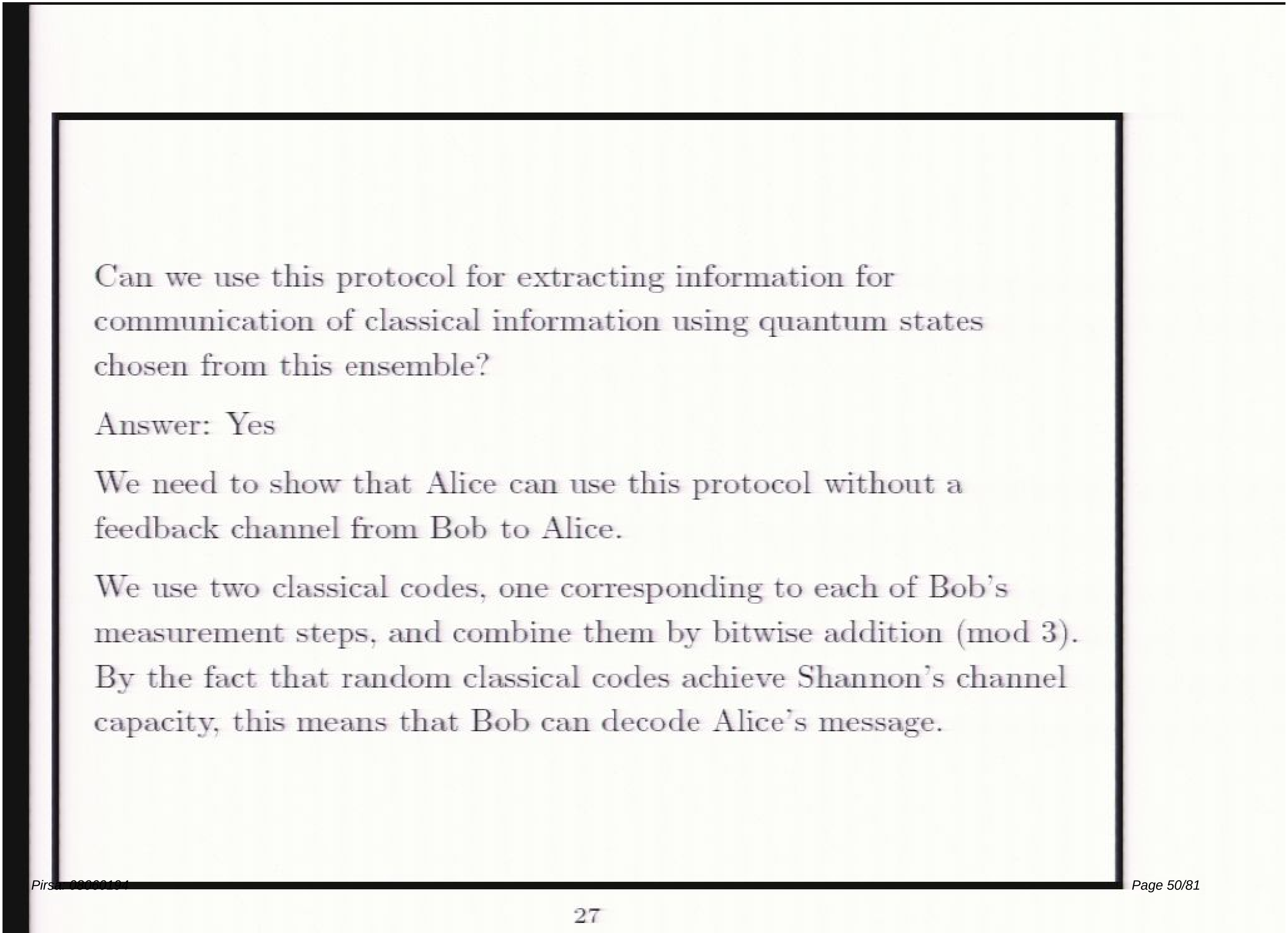
Signals:

LOCC ○ ○ ○ ○ different labs

nucleus

use classical communication


90% 0
1 present prob
Take two random codes
{1 2 0 0 1 1 0 2 1 0 1
1 0 1 1 0 1 0 1 1 0
2 2 1 1 1 2 0 0 2 0
odd
90° 90°



Can we use this protocol for extracting information for communication of classical information using quantum states chosen from this ensemble?

Answer: Yes

We need to show that Alice can use this protocol without a feedback channel from Bob to Alice.

We use two classical codes, one corresponding to each of Bob's measurement steps, and combine them by bitwise addition (mod 3). By the fact that random classical codes achieve Shannon's channel capacity, this means that Bob can decode Alice's message.

Can we use this protocol for extracting information for communication of classical information using quantum states chosen from this ensemble?

Answer: Yes

We need to show that Alice can use this protocol without a feedback channel from Bob to Alice.

We use two classical codes, one corresponding to each of Bob's measurement steps, and combine them by bitwise addition (mod 3). By the fact that random classical codes achieve Shannon's channel capacity, this means that Bob can decode Alice's message.

This gives us two additional capacities:

1. The capacity of a set of quantum states where Bob is allowed to make LOCC measurements.
2. The capacity of a set of quantum states where Bob has a back channel to Alice, and Alice has a side channel to Bob (but the amount of information sent over the side channel is subtracted from the total amount Bob receives).

Are they the same? I don't know.

They appear to be different if you look at the natural generalization to classical black boxes.

So far we have discussed communication using quantum states. We now discuss communication over quantum channels.

Formula for arbitrary memoryless quantum channel $\mathcal{N}$.
$\mathcal{N}$ must be trace-preserving completely positive operator.

$$\rho \longrightarrow \mathcal{N}(\rho) = \sum_i A_i \rho A_i^\dagger$$

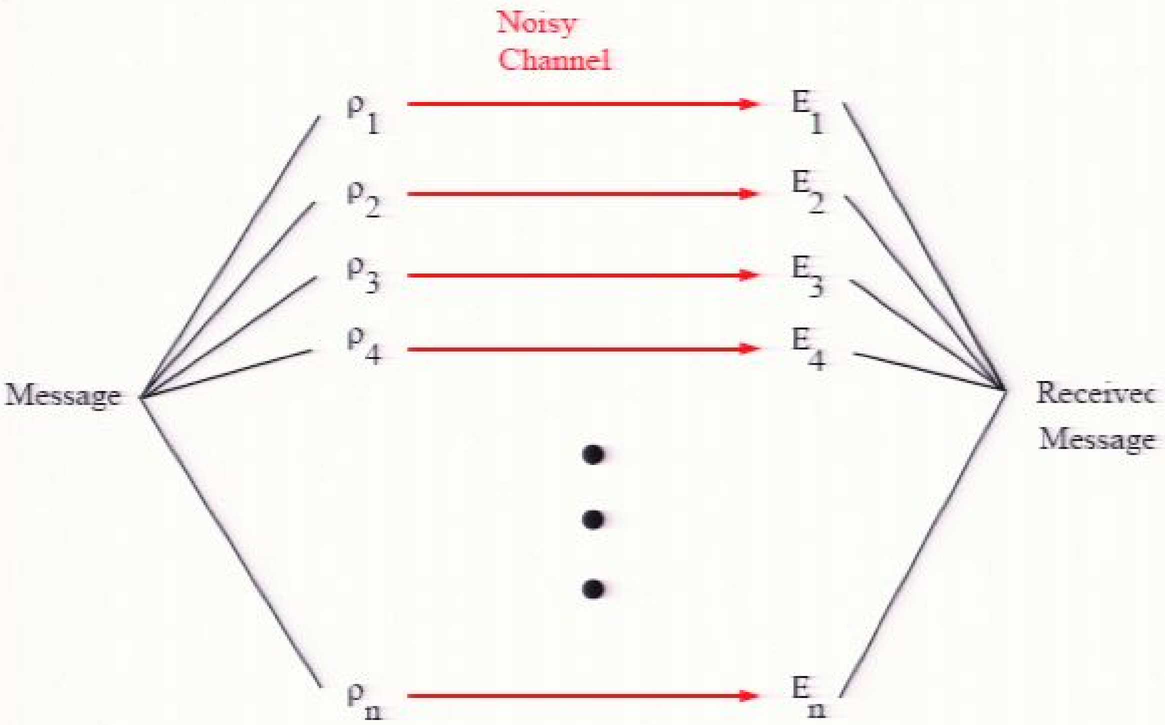
where

$$\sum_i A_i^\dagger A_i = I$$

Positive: takes positive semi-definite matrices to positive semi-definite matrices.

Completely positive: is positive even when tensored with the identity channel.

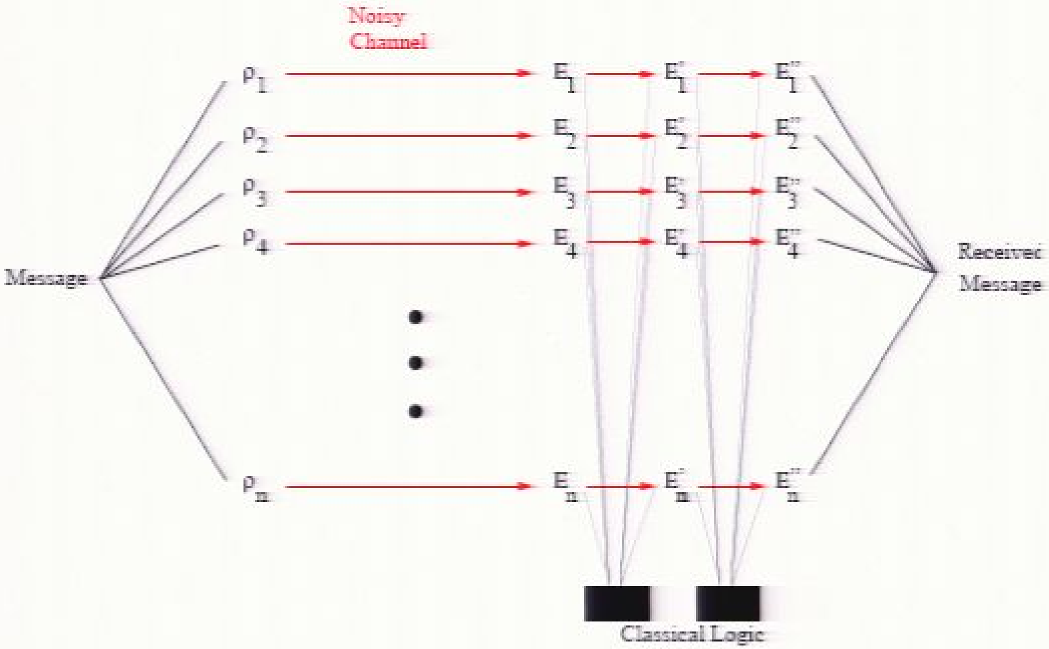
Unentangled Inputs, Separate Measurements



Maximize, over probability distributions p_i on inputs ρ_i to the channel,

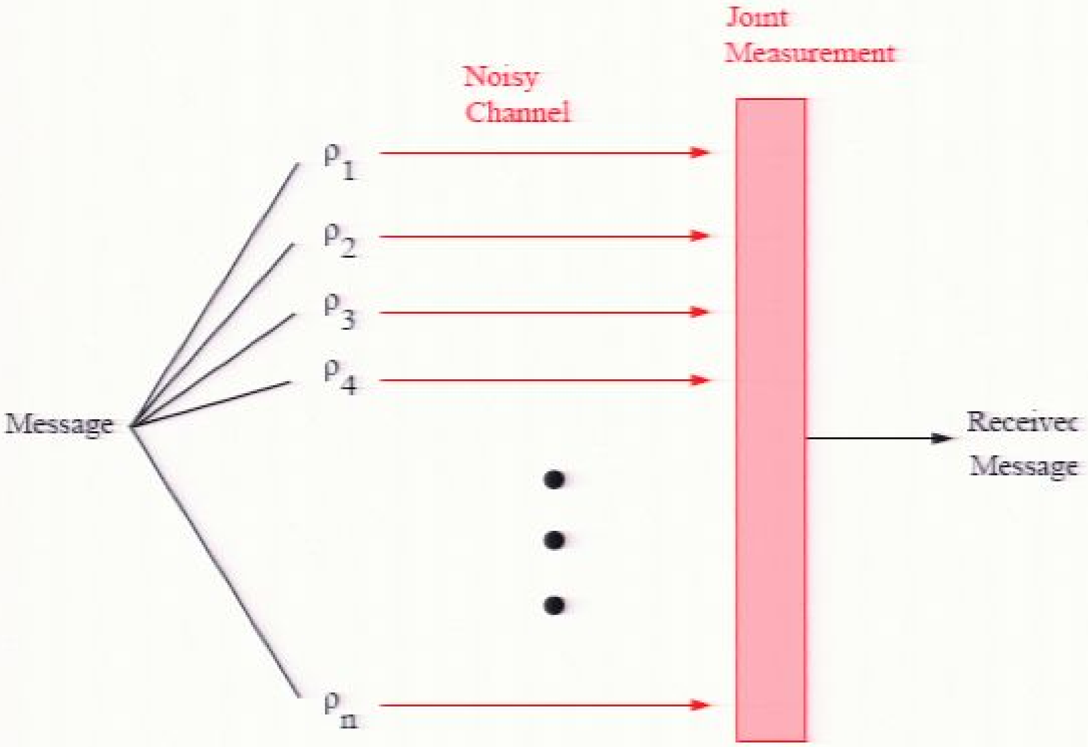
$$I_{\text{acc}}(\{\mathcal{N}(\rho_i)\};\{p_i\})$$

Unentangled Inputs, Adaptive Separate Measurements



Does this increase if you add a feedback channel from Bob to Alice? Allow a classical side channel from Alice to Bob that must be paid for?

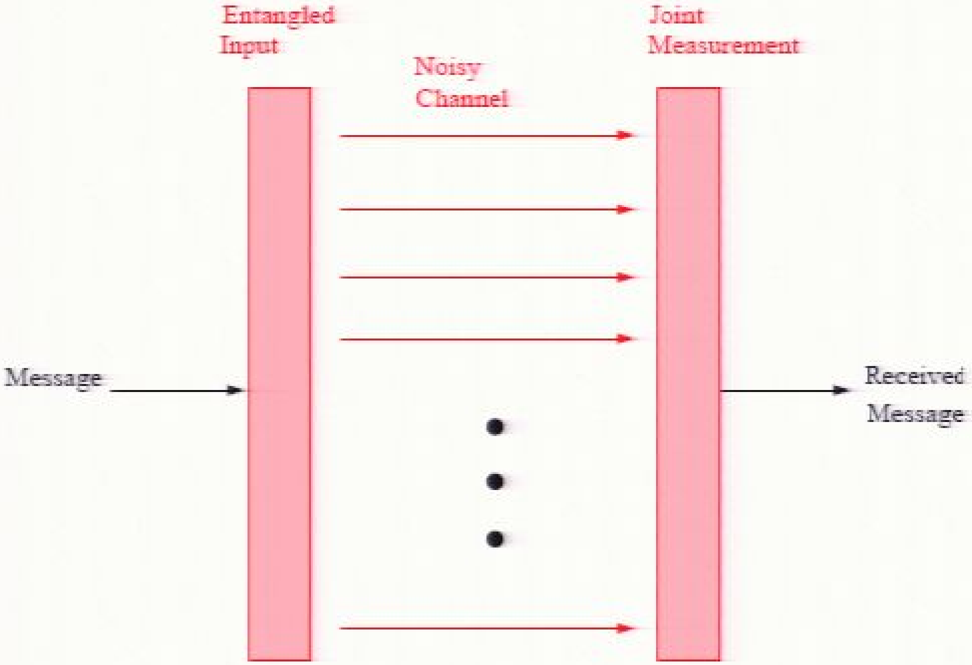
Unentangled Inputs, Joint Measurements



Maximize over probability distributions p_i on inputs ρ_i to the channel

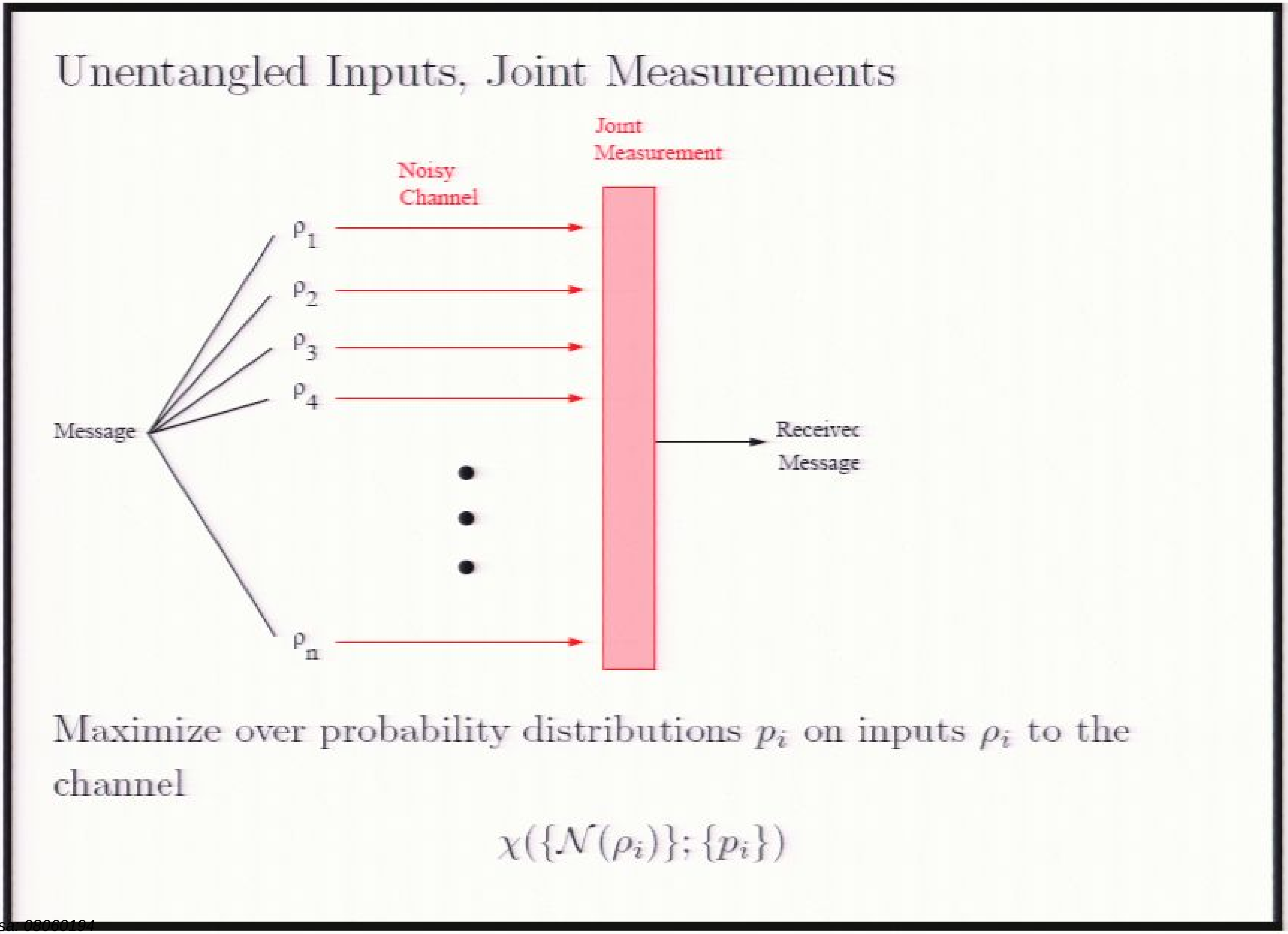
$$\chi(\{\mathcal{N}(\rho_i)\};\{p_i\})$$

Entangled Inputs, Joint Measurements



Maximize over probability distributions p_i on inputs ρ_i to the channel, where ρ_i is in the tensor product space of n inputs:

$$\lim_{n\to\infty} \chi(\{\mathcal{N}^{\otimes n}(\rho_i)\};\{p_i\})$$

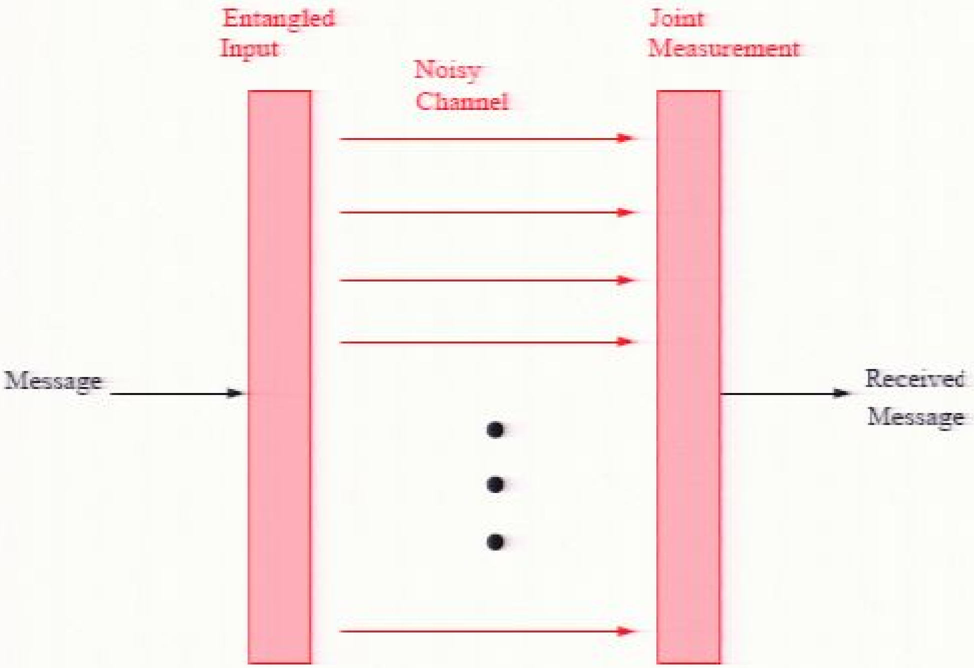


Unentangled Inputs, Joint Measurements

Maximize over probability distributions p_i on inputs ρ_i to the channel

$$\chi(\{\mathcal{N}(\rho_i)\};\{p_i\})$$

Entangled Inputs, Joint Measurements



Maximize over probability distributions p_i on inputs ρ_i to the channel, where ρ_i is in the tensor product space of n inputs:

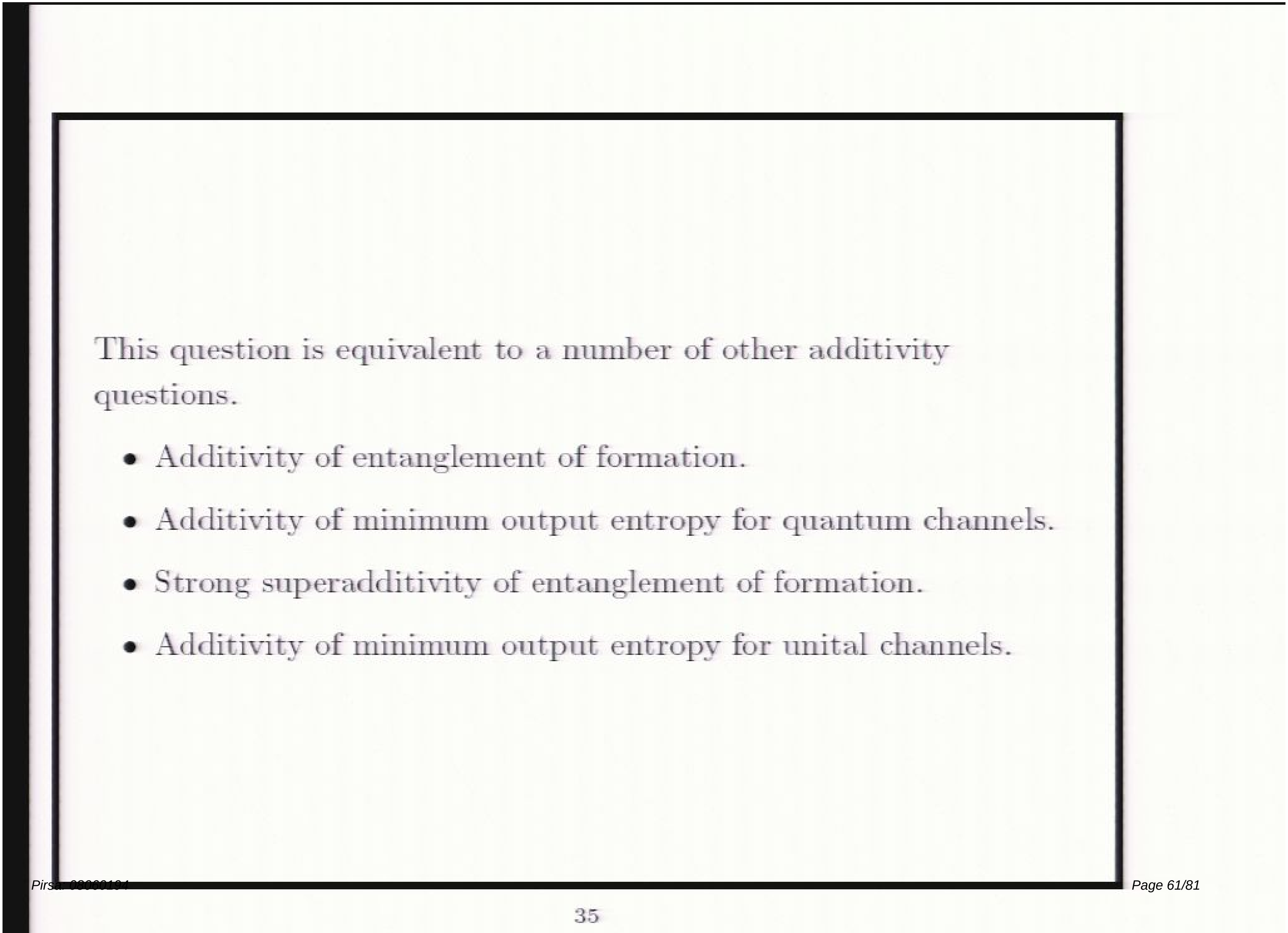
$$\lim_{n\to\infty} \chi(\{\mathcal{N}^{\otimes n}(\rho_i)\};\{p_i\})$$

Open Question

Is channel capacity additive?

Is $\max\ \chi(\mathcal{N}_1 \otimes \mathcal{N}_2) = \max\ \chi(\mathcal{N}_1) + \max\ \chi(\mathcal{N}_2)$?

If it is, then χ gives the classical-information capacity of a quantum channel.



This question is equivalent to a number of other additivity questions.

- Additivity of entanglement of formation.
- Additivity of minimum output entropy for quantum channels.
- Strong superadditivity of entanglement of formation.
- Additivity of minimum output entropy for unital channels.

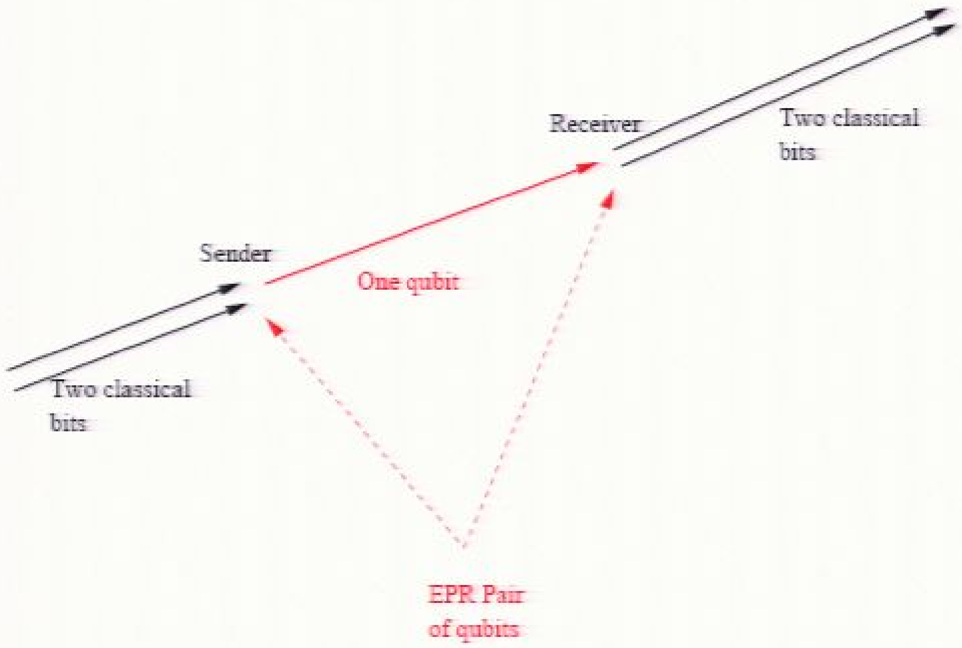
What things might increase the capacity of a quantum channel which don't affect the capacity of a classical channel?

a) Entanglement between different channel uses? Unknown: this is the additivity question.

b) A classical feedback channel from the receiver to the sender? Not without a). More later.

c) Prior entanglement shared between the sender and the receiver. This helps!

The phenomenon called superdense coding lets you send two bits per qubit over a noiseless quantum channel if the sender and receiver share entanglement.



Receiver

Two classical
bits

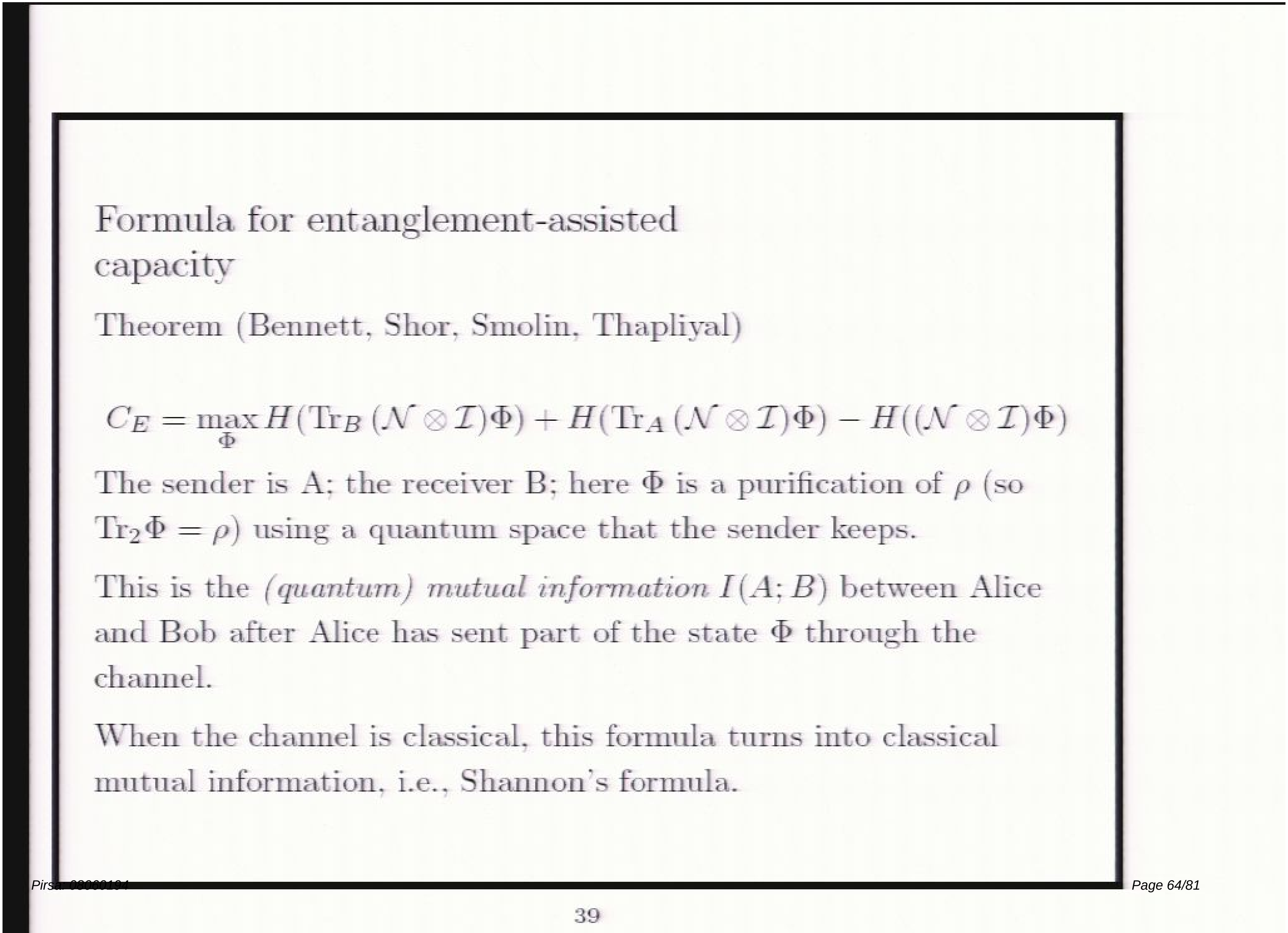
Sender

One qubit

Two classical
bits

EPR Pair
of qubits

By Holevo's theorem, the bound without prior shared entanglement is one bit per qubit.



Formula for entanglement-assisted capacity

Theorem (Bennett, Shor, Smolin, Thapliyal)

$$C_E = \max_{\Phi} H(\mathrm{Tr}_B\,(\mathcal{N} \otimes \mathcal{I})\Phi) + H(\mathrm{Tr}_A\,(\mathcal{N} \otimes \mathcal{I})\Phi) - H((\mathcal{N} \otimes \mathcal{I})\Phi)$$

The sender is A; the receiver B; here Φ is a purification of ρ (so $\mathrm{Tr}_2\Phi = \rho$) using a quantum space that the sender keeps.

This is the *(quantum) mutual information* $I(A;B)$ between Alice and Bob after Alice has sent part of the state Φ through the channel.

When the channel is classical, this formula turns into classical mutual information, i.e., Shannon's formula.

Suppose that we have a quantum channel $\mathcal{N}$. From superdense coding, if $\mathcal{N}$ is a noiseless quantum channel, the sender could communicate twice as much classical information to a receiver if they share EPR pairs than if they don't.

This new capacity generalizes to noisy channels. We call this quantity the entanglement-assisted capacity and denote it by C_E.

Formula for entanglement-assisted capacity

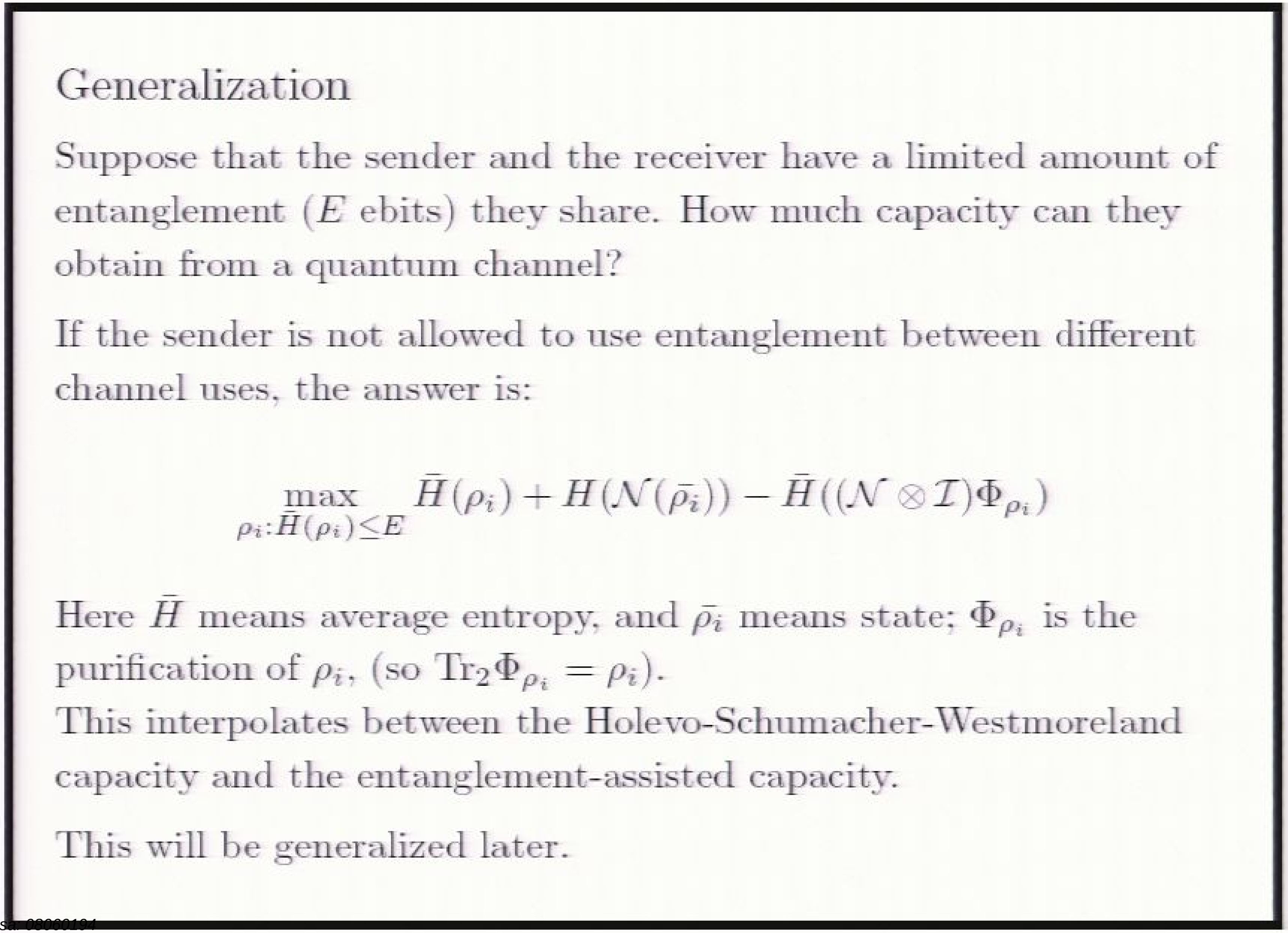
Theorem (Bennett, Shor, Smolin, Thapliyal)

$$C_E = \max_{\Phi} H(\mathrm{Tr}_B\,(\mathcal{N} \otimes \mathcal{I})\Phi) + H(\mathrm{Tr}_A\,(\mathcal{N} \otimes \mathcal{I})\Phi) - H((\mathcal{N} \otimes \mathcal{I})\Phi)$$

The sender is A; the receiver B; here Φ is a purification of ρ (so $\mathrm{Tr}_2\Phi = \rho$) using a quantum space that the sender keeps.

This is the *(quantum) mutual information* $I(A;B)$ between Alice and Bob after Alice has sent part of the state Φ through the channel.

When the channel is classical, this formula turns into classical mutual information, i.e., Shannon's formula.



Generalization

Suppose that the sender and the receiver have a limited amount of entanglement (E ebits) they share. How much capacity can they obtain from a quantum channel?

If the sender is not allowed to use entanglement between different channel uses, the answer is:

$$\max_{\rho_i : \bar{H}(\rho_i) \leq E} \bar{H}(\rho_i) + H(\mathcal{N}(\bar{\rho_i})) - \bar{H}((\mathcal{N} \otimes \mathcal{I})\Phi_{\rho_i})$$

Here $\bar{H}$ means average entropy, and $\bar{\rho_i}$ means state; Φ_{ρ_i} is the purification of ρ_i, (so $\mathrm{Tr}_2 \Phi_{\rho_i} = \rho_i$).
This interpolates between the Holevo-Schumacher-Westmoreland capacity and the entanglement-assisted capacity.

This will be generalized later.

Formula for entanglement-assisted capacity

Theorem (Bennett, Shor, Smolin, Thapliyal)

$$C_E = \max_{\Phi} H(\mathrm{Tr}_B\,(\mathcal{N}\otimes\mathcal{I})\Phi) + H(\mathrm{Tr}_A\,(\mathcal{N}\otimes\mathcal{I})\Phi) - H((\mathcal{N}\otimes\mathcal{I})\Phi)$$

The sender is A; the receiver B; here Φ is a purification of ρ (so $\mathrm{Tr}_2\Phi = \rho$) using a quantum space that the sender keeps.

This is the *(quantum) mutual information* $I(A; B)$ between Alice and Bob after Alice has sent part of the state Φ through the channel.

When the channel is classical, this formula turns into classical mutual information, i.e., Shannon's formula.

Generalization

Suppose that the sender and the receiver have a limited amount of entanglement (E ebits) they share. How much capacity can they obtain from a quantum channel?

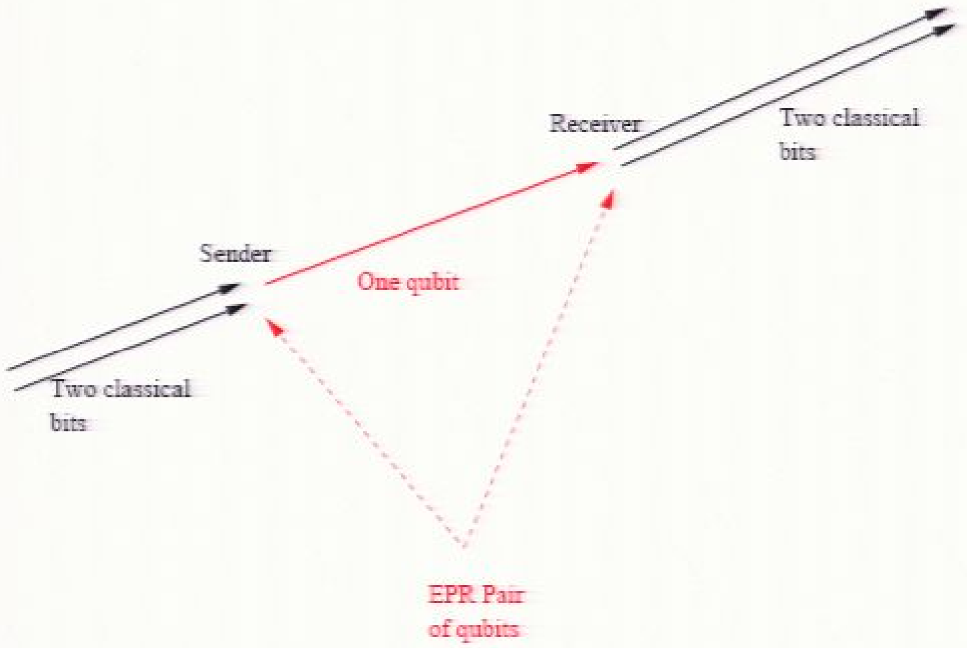
If the sender is not allowed to use entanglement between different channel uses, the answer is:

$$\max_{\rho_i : \bar{H}(\rho_i) \leq E} \bar{H}(\rho_i) + H(\mathcal{N}(\bar{\rho_i})) - \bar{H}((\mathcal{N} \otimes \mathcal{I})\Phi_{\rho_i})$$

Here $\bar{H}$ means average entropy, and $\bar{\rho_i}$ means state; Φ_{ρ_i} is the purification of ρ_i, (so $\mathrm{Tr}_2 \Phi_{\rho_i} = \rho_i$).
This interpolates between the Holevo-Schumacher-Westmoreland capacity and the entanglement-assisted capacity.

This will be generalized later.

The phenomenon called superdense coding lets you send two classical bits per qubit over a noiseless quantum channel if the sender and receiver share entanglement.



By Holevo's theorem, the bound without prior shared entanglement is one bit per qubit.

Quantum capacity of a quantum channel.

How many qubits can you send from the sender to the receiver per channel use?

Some channels (e.g. prob $\frac{1}{2}$ erasure channel) have no quantum capacity.

This channel can be viewed as a channel which sends the output to the receiver with probability $\frac{1}{2}$ and to an eavesdropper with probability $\frac{1}{2}$. If you could send qubits reliably through this channel, you could clone them.

Formula for quantum capacity.

$$\lim_{n\to\infty} \frac{1}{n} \max_{\rho} H(\mathcal{N}^{\otimes n}(\rho)) - H((\mathcal{N}^{\otimes n} \otimes I))(\Phi_\rho))$$

where Φ_ρ is a purification of ρ.

This formula is *not* additive, so we need the limit as the number of channel uses n goes to infinity.

Now you can ask questions such as: for a given channel $\mathcal{N}$, with E shared entanglement, how many qubits Q and how many classical bits C can you send.

For example, in the *father* protocol (Devetak, Harrow, Winter), $C = 0$ and:

$$\begin{aligned} E &= \frac{1}{2}\Big[H(\rho) + H(\mathcal{N} \otimes I(\Phi_\rho)) - H(\mathcal{N}(\rho))\Big] \\ Q &= \frac{1}{2}\Big[H(\rho) + H(\mathcal{N}(\rho)) - H(\mathcal{N} \otimes I(\Phi_\rho))\Big] \end{aligned}$$

This can be expressed more simply by

$$\begin{aligned} E &= \frac{1}{2} I(\text{Alice}; \text{Environment}) \\ Q &= \frac{1}{2} I(\text{Alice}; \text{Bob}) \end{aligned}$$

The father formula can be combined with superdense coding to give the entanglement-assisted capacity formula.

The father formula can be combined with the transmission of halves of EPR pairs to give the formula for quantum channel capacity. capacity formula.

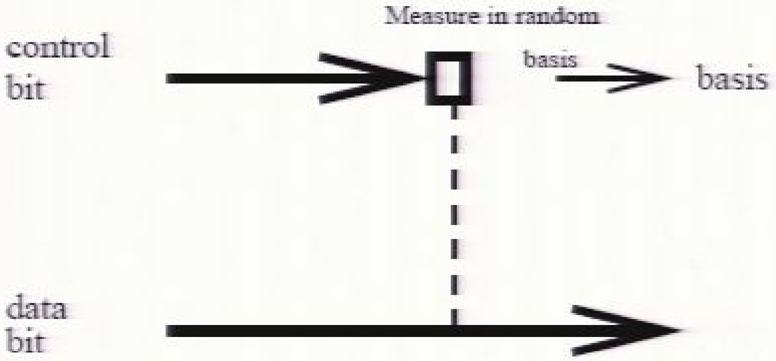
There is a similar *mother* protocol when sender and receiver share an entangled mixed state.

This is part of the very nice theory of *resource inequalities* in quantum information.

Now, let's revisit the question of what things might increase the capacity of a quantum channel which don't affect the capacity of a classical channel?

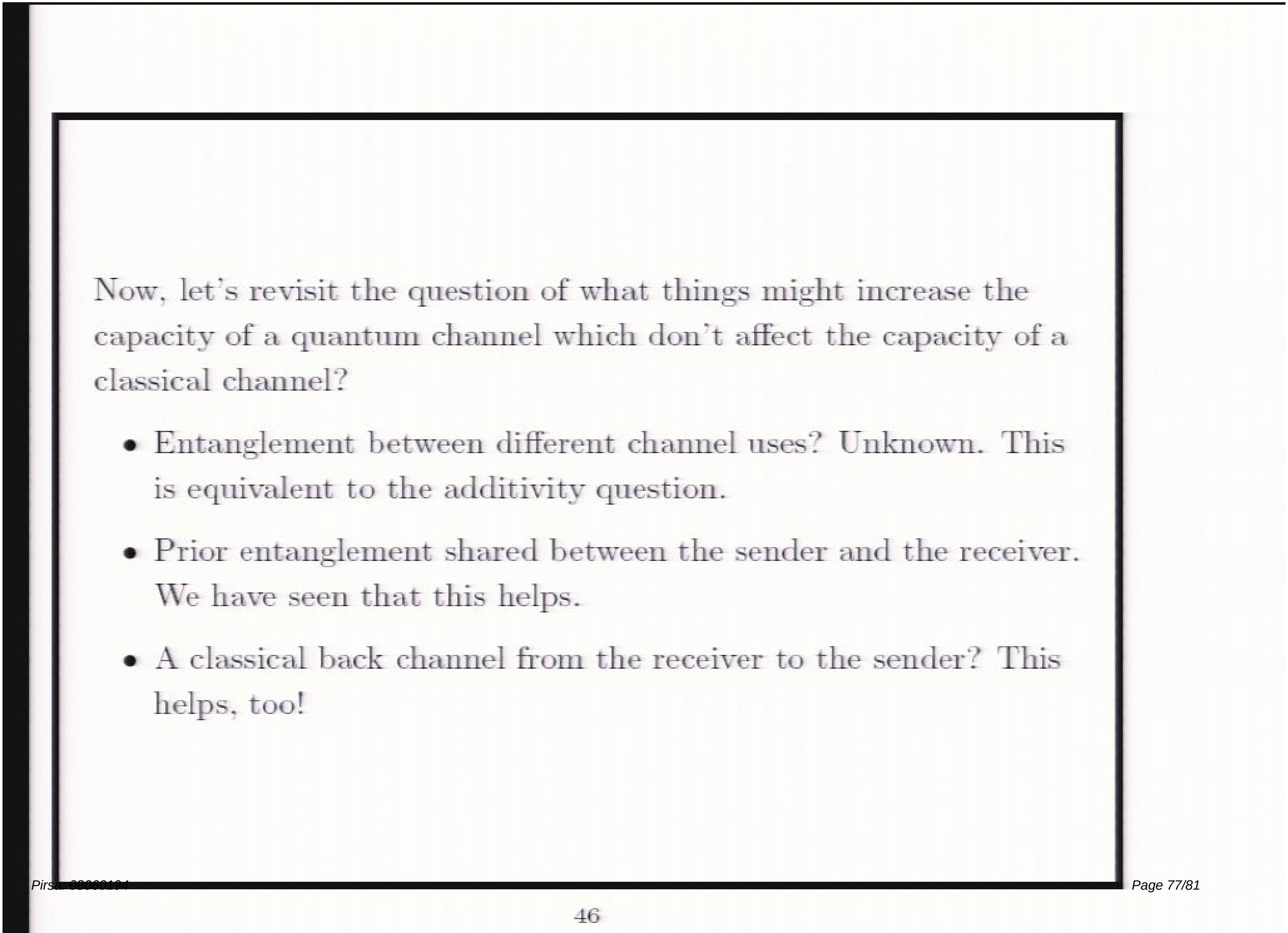
- Entanglement between different channel uses? Unknown. This is equivalent to the additivity question.
- Prior entanglement shared between the sender and the receiver. We have seen that this helps.
- A classical back channel from the receiver to the sender? This helps, too!

Why does a back channel help?



Consider a channel with two input registers. The control register is measured in a random basis, with outcome m, and then this basis and $U_m \,|\, \psi\rangle$ is given to Bob, where ψ is the input of the data register.

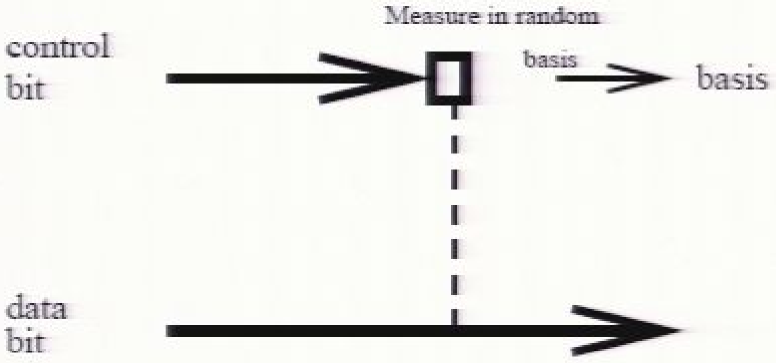
If Bob can send the measurement basis to Alice, she can input half of an EPR pair into the control bit, and measure the other half to learn the measurement outcome m. Then she can tell Bob m, and he can undo the unitary U_m.



Now, let's revisit the question of what things might increase the capacity of a quantum channel which don't affect the capacity of a classical channel?

- Entanglement between different channel uses? Unknown. This is equivalent to the additivity question.
- Prior entanglement shared between the sender and the receiver. We have seen that this helps.
- A classical back channel from the receiver to the sender? This helps, too!

Why does a back channel help?



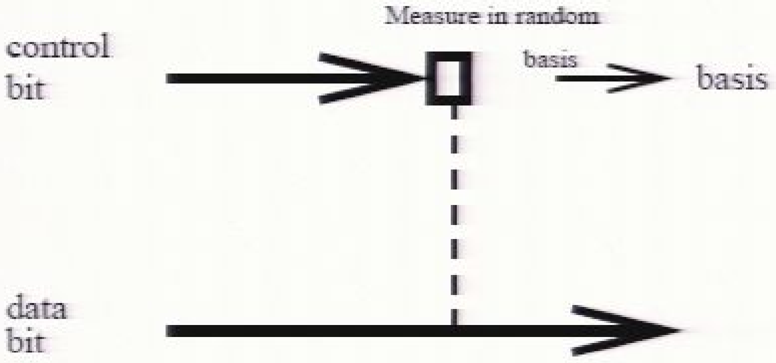
Consider a channel with two input registers. The control register is measured in a random basis, with outcome m, and then this basis and $U_m \,|\, \psi\rangle$ is given to Bob, where ψ is the input of the data register.

If Bob can send the measurement basis to Alice, she can input half of an EPR pair into the control bit, and measure the other half to learn the measurement outcome m. Then she can tell Bob m, and he can undo the unitary U_m.

Now, let's revisit the question of what things might increase the capacity of a quantum channel which don't affect the capacity of a classical channel?

- Entanglement between different channel uses? Unknown. This is equivalent to the additivity question.
- Prior entanglement shared between the sender and the receiver. We have seen that this helps.
- A classical back channel from the receiver to the sender? This helps, too!

Why does a back channel help?



Consider a channel with two input registers. The control register is measured in a random basis, with outcome m, and then this basis and $U_m \,|\, \psi\rangle$ is given to Bob, where ψ is the input of the data register.

If Bob can send the measurement basis to Alice, she can input half of an EPR pair into the control bit, and measure the other half to learn the measurement outcome m. Then she can tell Bob m, and he can undo the unitary U_m.

With the above type of construction, we can discover that the eight capacities

$$\begin{array}{ccccccc} C & \leq & C_B & \leq & C_2 & \leq & C_E \\ \text{\rotatebox{90}{$\leq$}} & & \text{\rotatebox{90}{$\leq$}} & & \text{\rotatebox{90}{$\leq$}} & & \text{\rotatebox{90}{$\leq$}} \\ Q & \leq & Q_B & \leq & Q_2 & \leq & Q_E \end{array}$$

appear to be all different, although $C_E = 2Q_E$. They can be proved to be different, assuming the additivity conjecture is true.

Here, C and Q are the classical and quantum capacities, respectively; C_B and Q_B are the capacities with a back-channel, C_E and Q_E are the capacities when the sender and receiver have shared entanglement to aid them in communication, and C_2 and Q_2 are the private and quantum capacities with a two-way side channel.