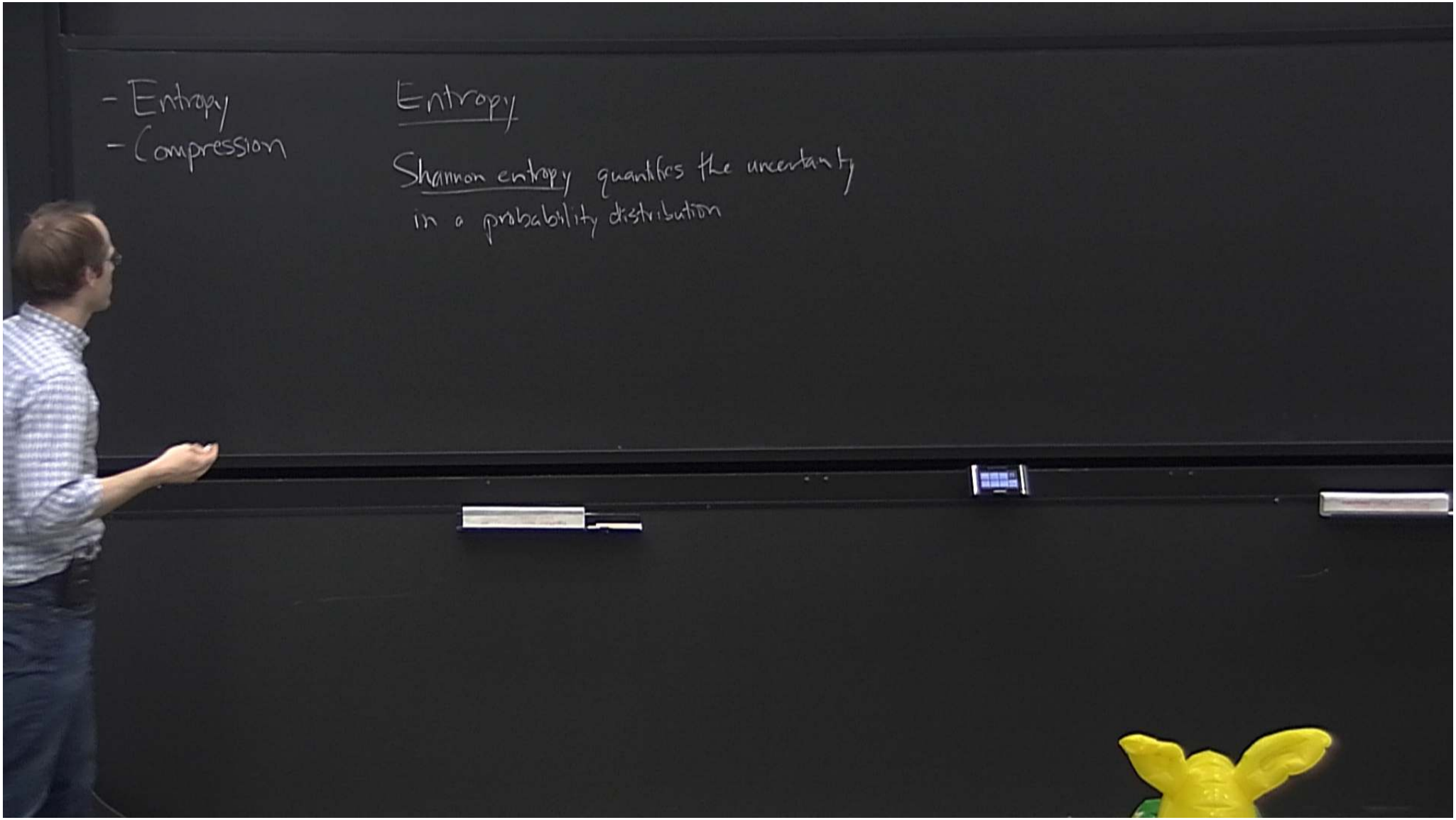
Title: 12/13 PSI - Quantum Information Review Lecture 11

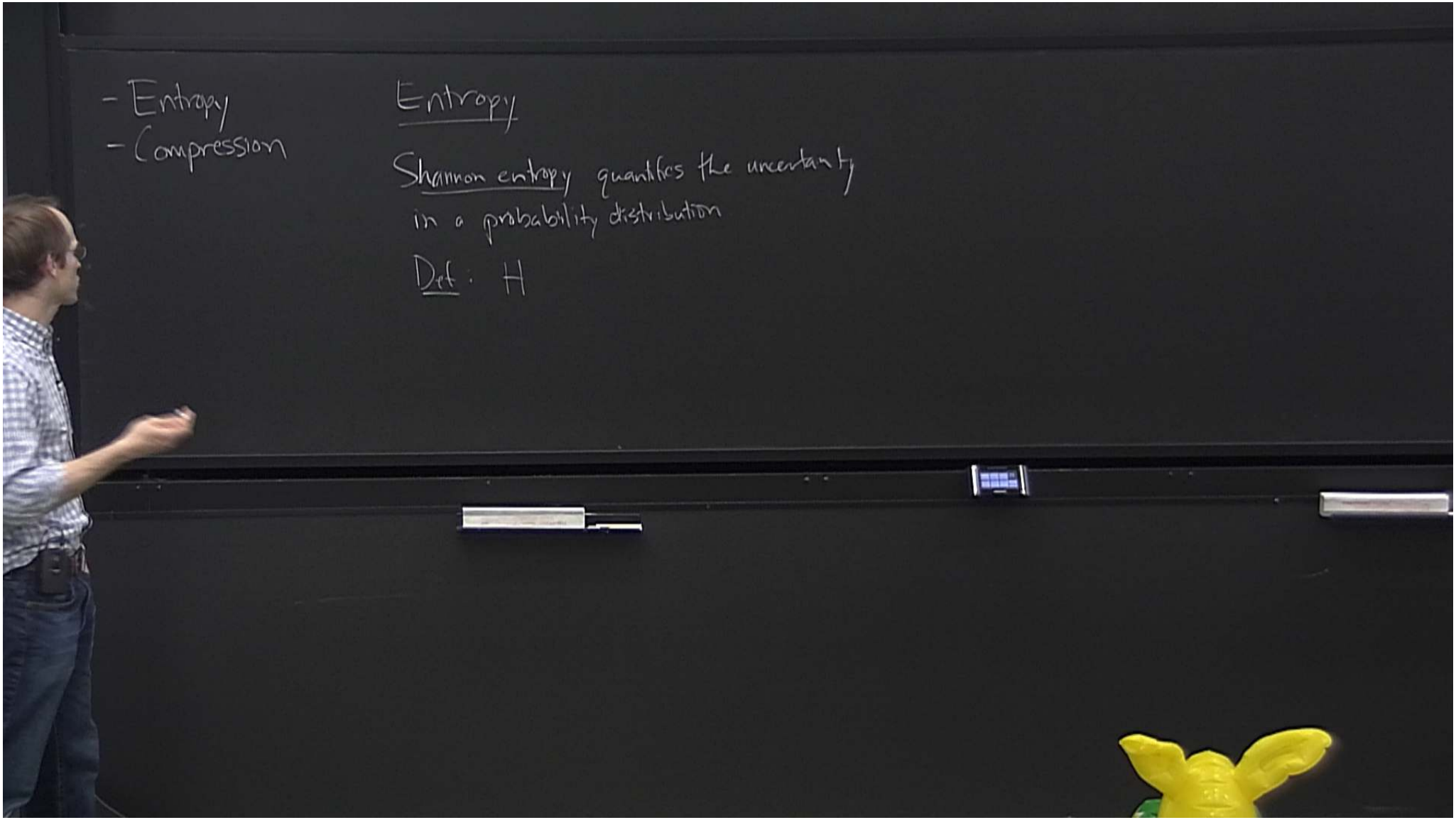
Date: Mar 05, 2013 10:15 AM

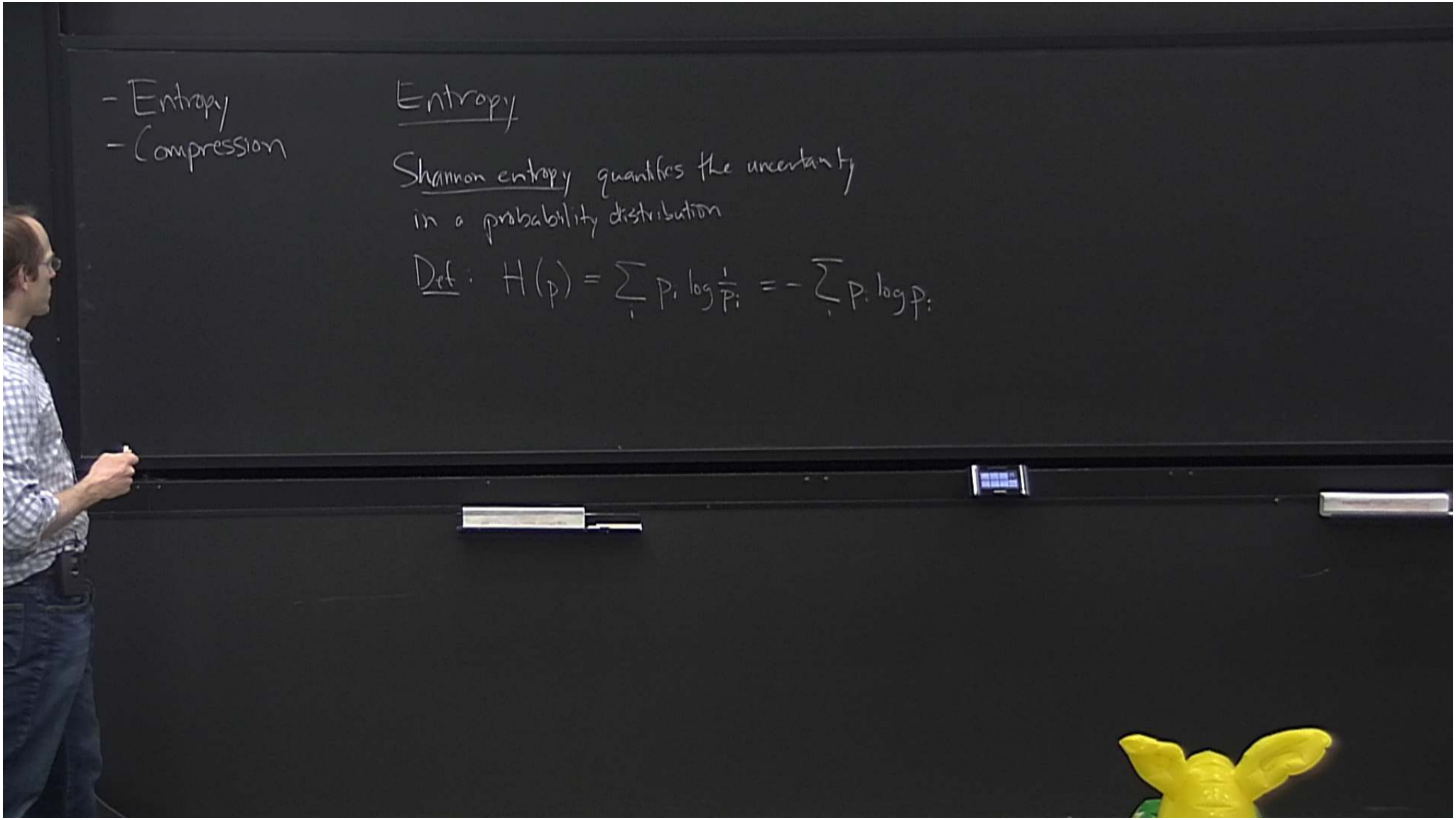
URL: http://www.pirsa.org/13030050

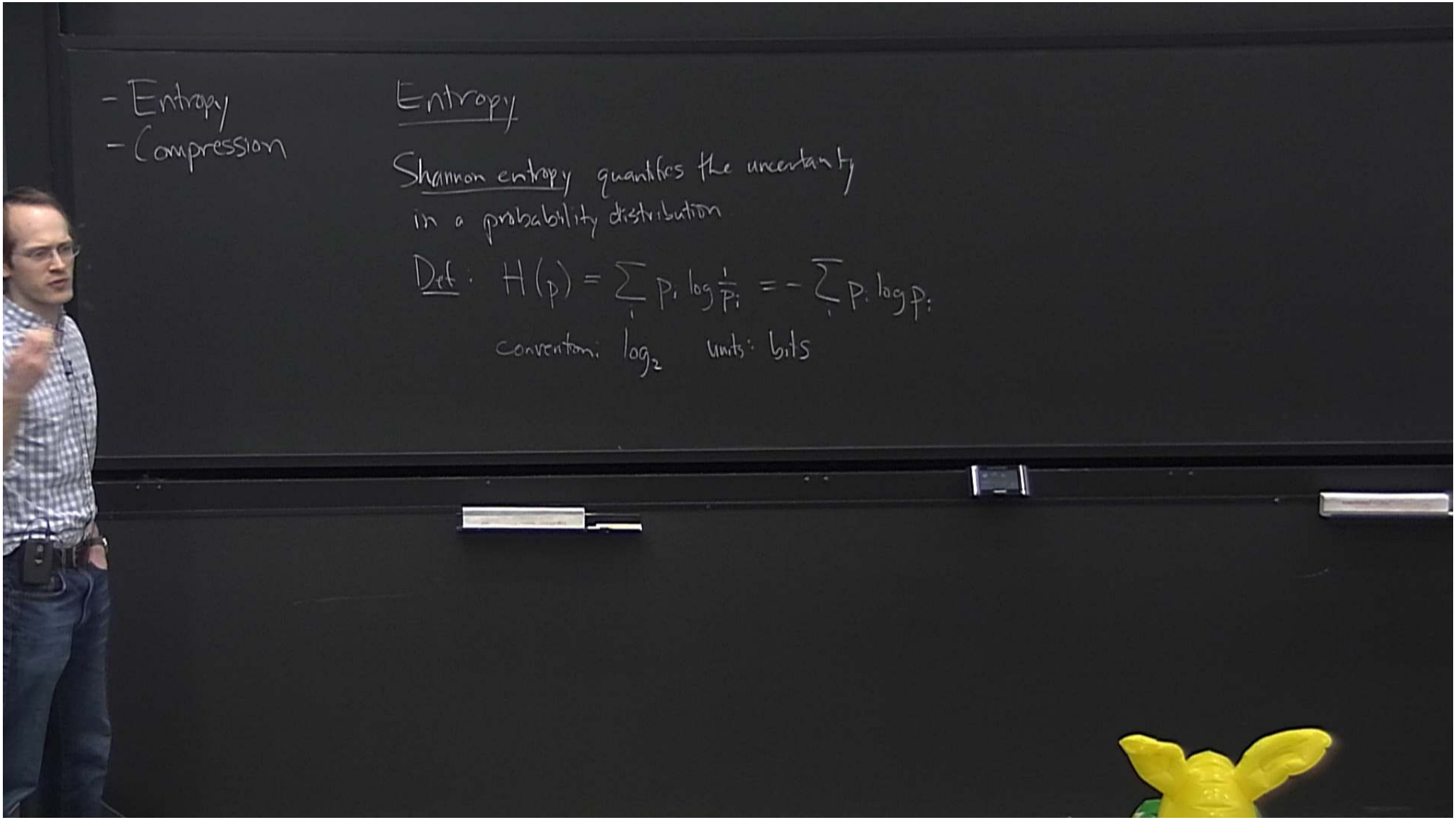
Abstract:

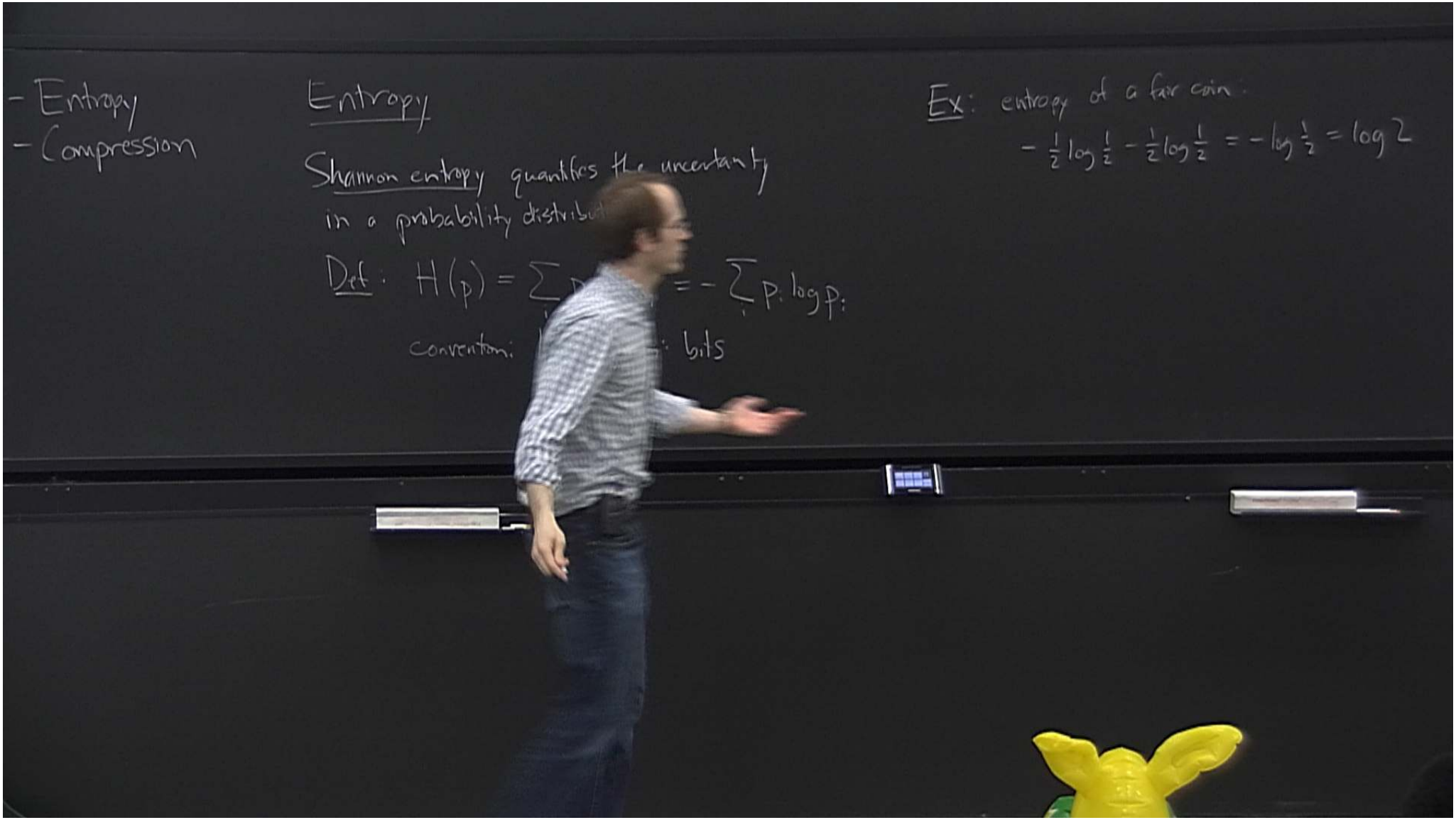
- Entropy
- Compression
Entropy

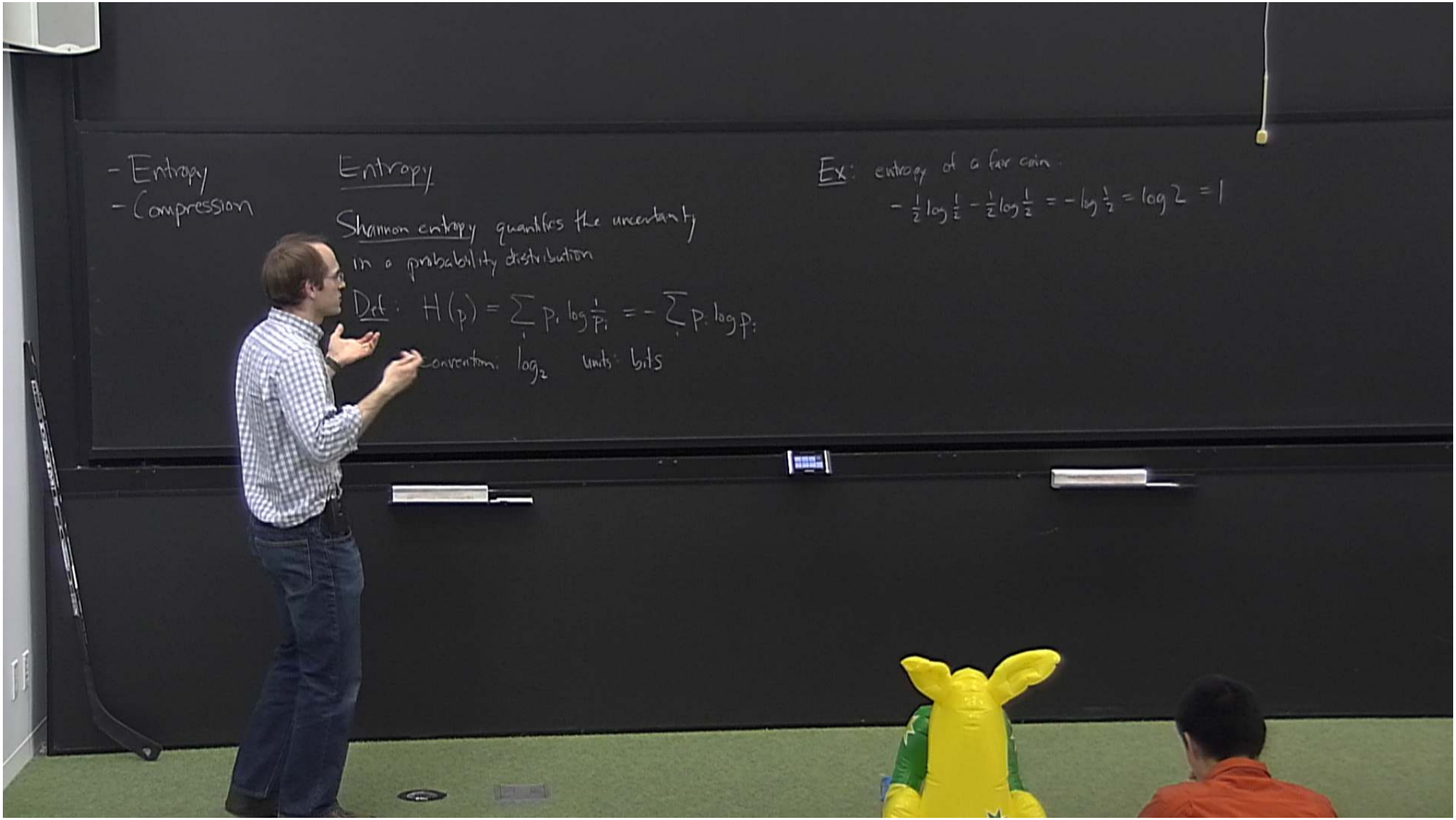

- Entropy
- Compression
Entropy
Shannon entropy quantifies the uncertainty
in a probability distribution

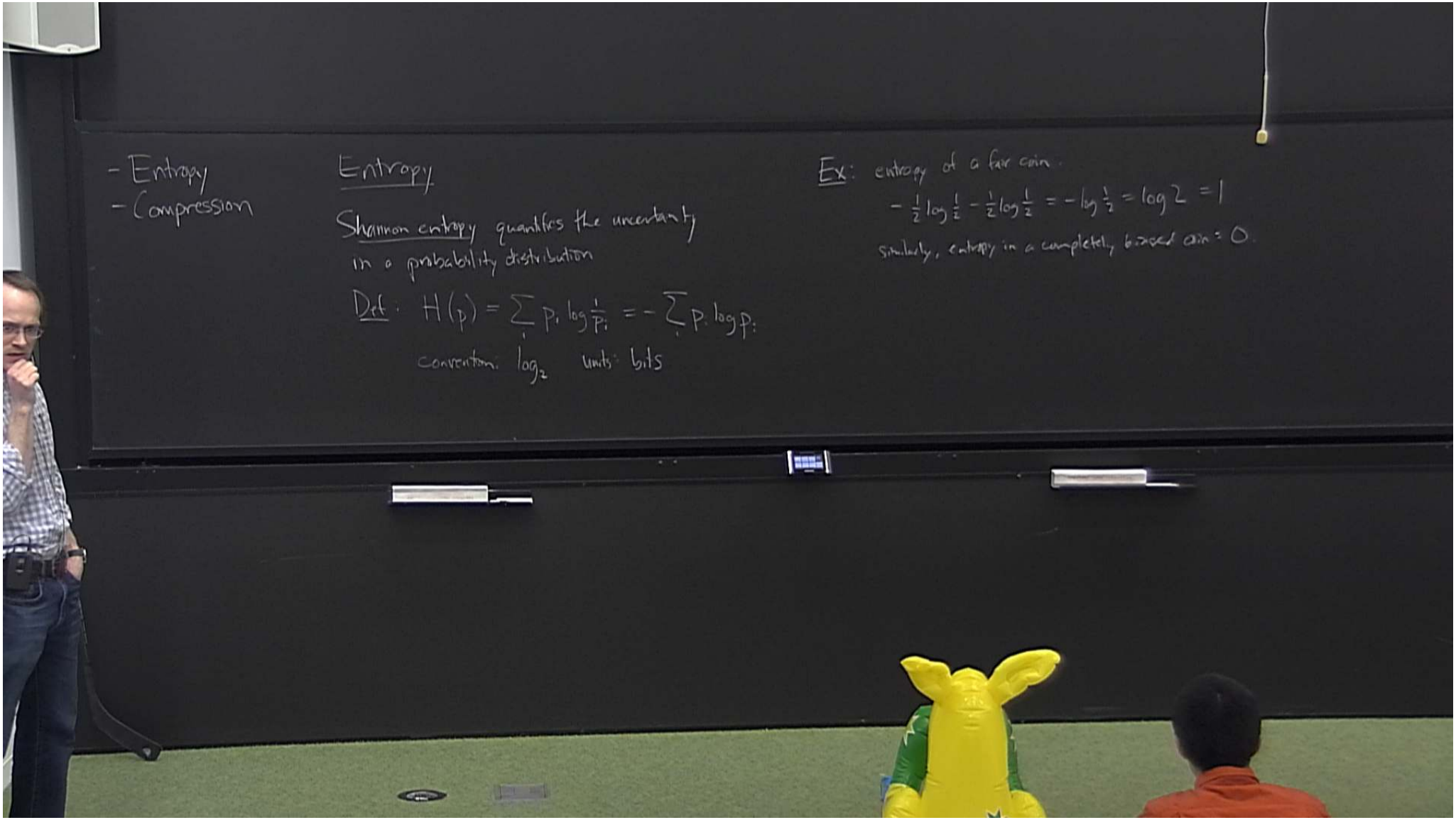

- Entropy
- Compression
Entropy
Shannon entropy quantifies the uncertainty
in a probability distribution
Def: H

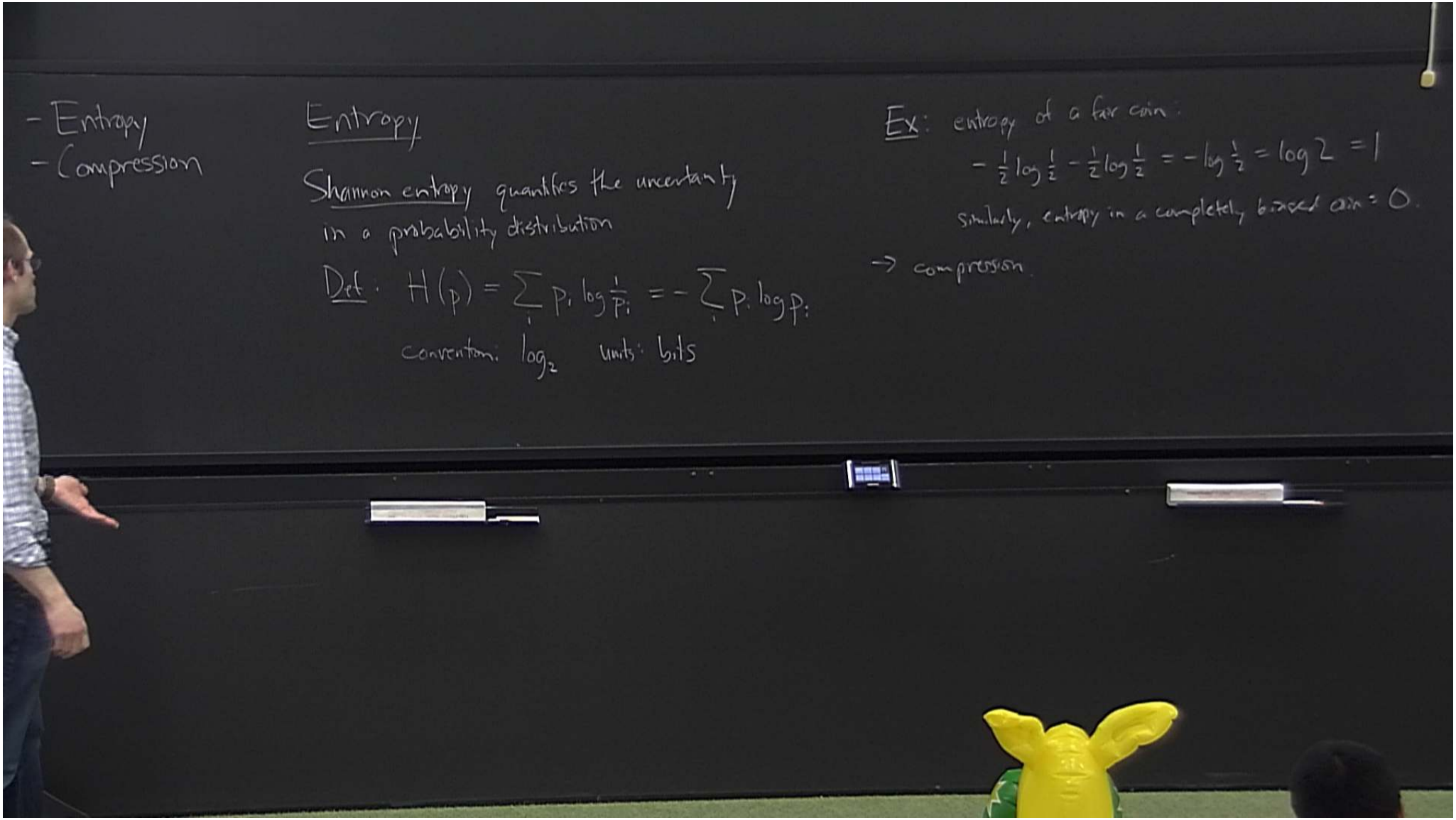

- Entropy
- Compression
Entropy
Shannon entropy quantifies the uncertainty
in a probability distribution
Def: $H(p) = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i$

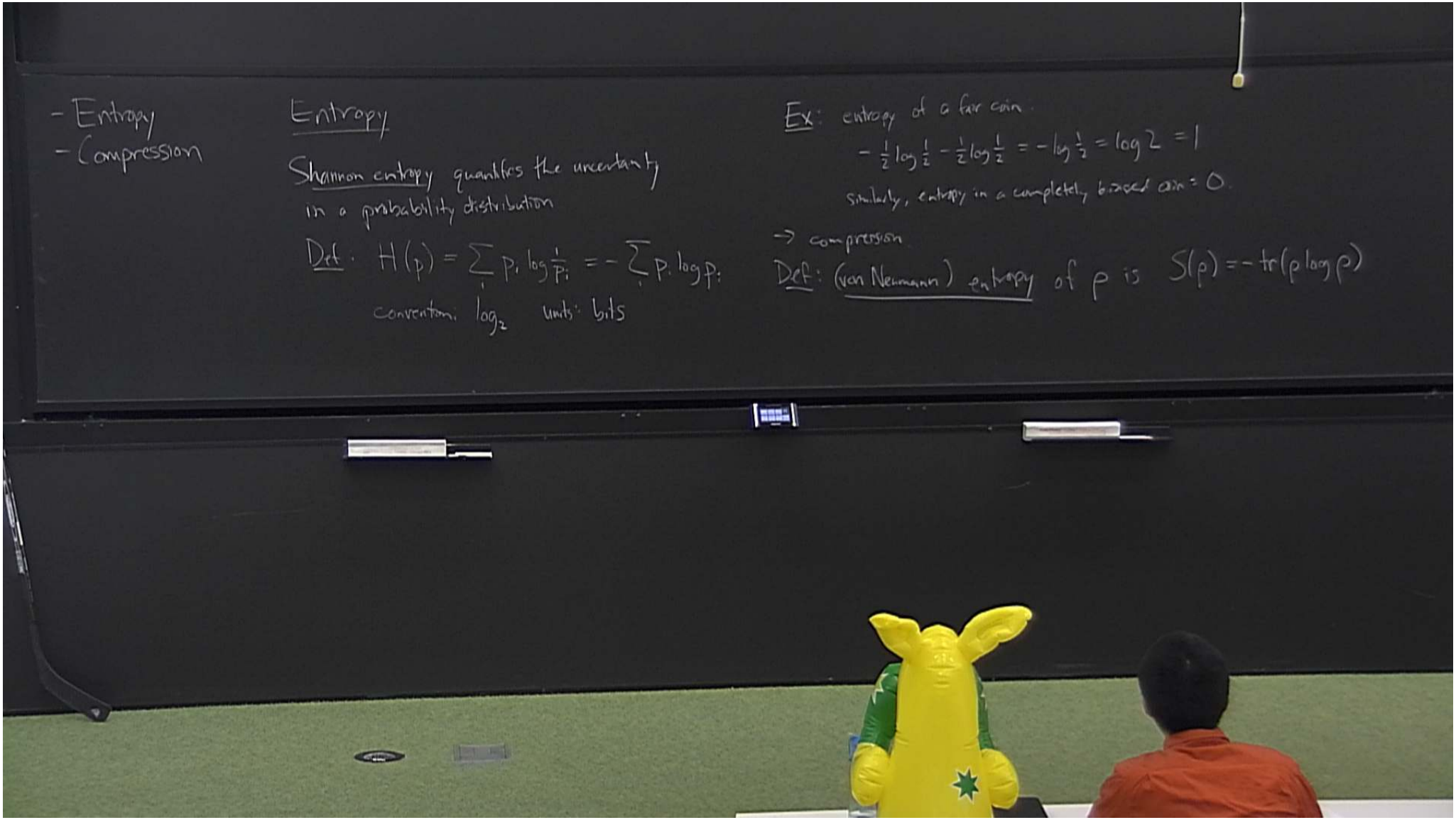

- Entropy
- Compression
Entropy
Shannon entropy quantifies the uncertainty
in a probability distribution
Def: $H(p) = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i$
convention: $\log_2$ units: bits

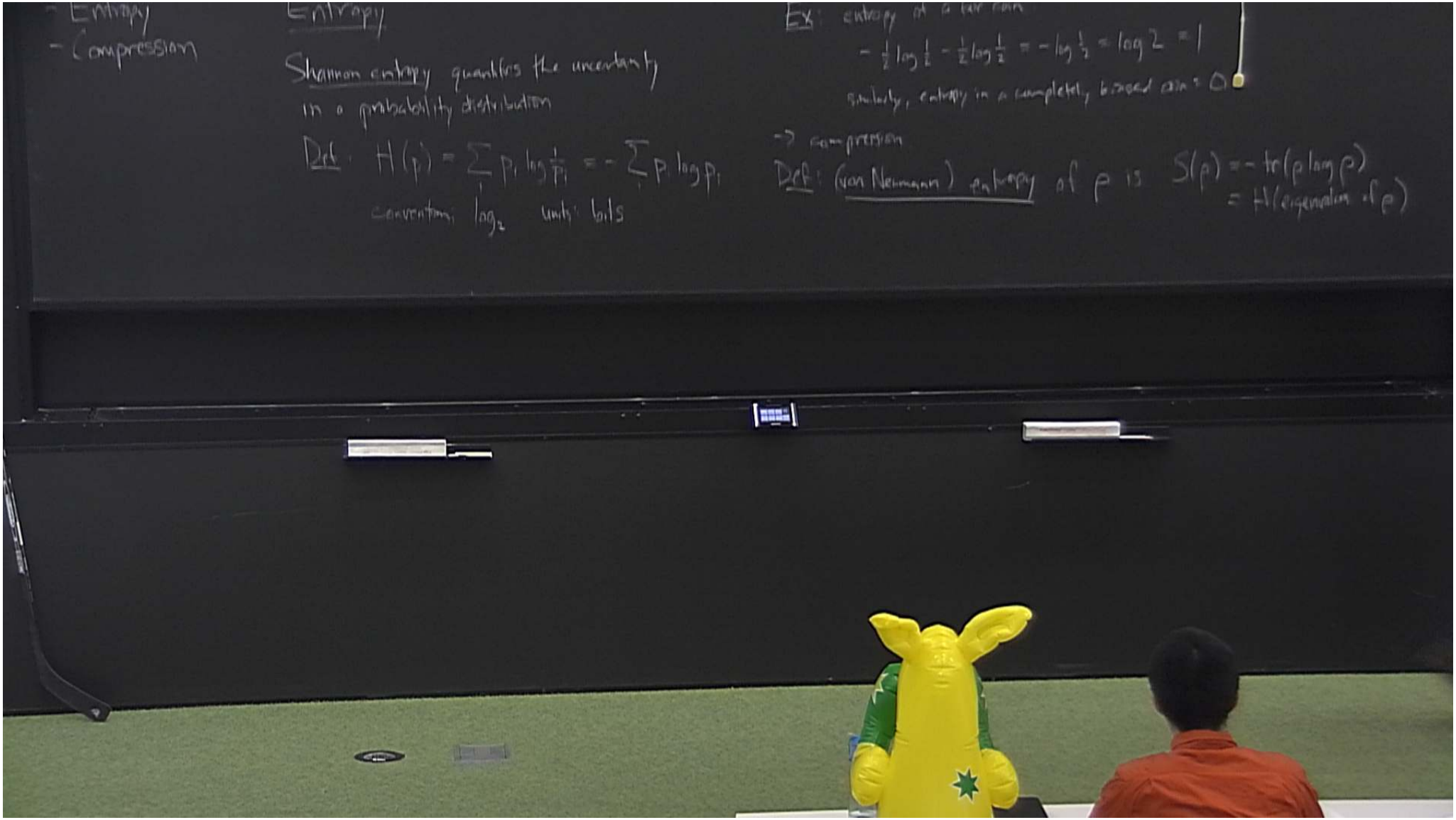

- Entropy
- Compression
Entropy
Shannon entropy quantifies the uncertainty
in a probability distribu
Def: H(p) = ∑ = - ∑ p_i log p_i
Convention: bits
Ex: entropy of a fair coin:
- 1/2 log 1/2 - 1/2 log 1/2 = - log 1/2 = log 2

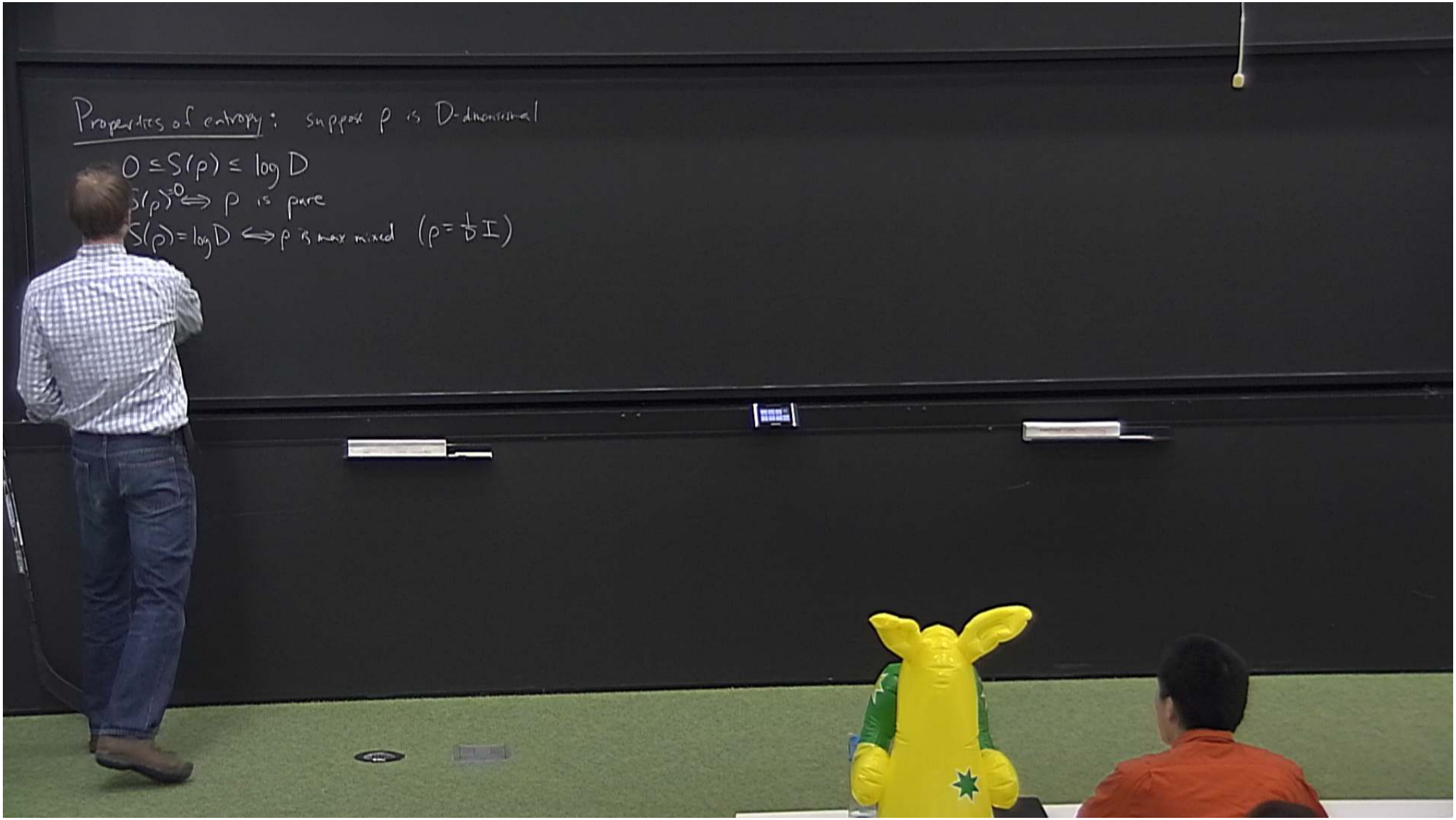

- Entropy
- Compression
Entropy
Shannon entropy quantifies the uncertainty in a probability distribution
Def: $H(p) = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i$
Conventions: $\log_2$ Units: bits
Ex: entropy of a fair coin
$-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = -\log\frac{1}{2} = \log 2 = 1$

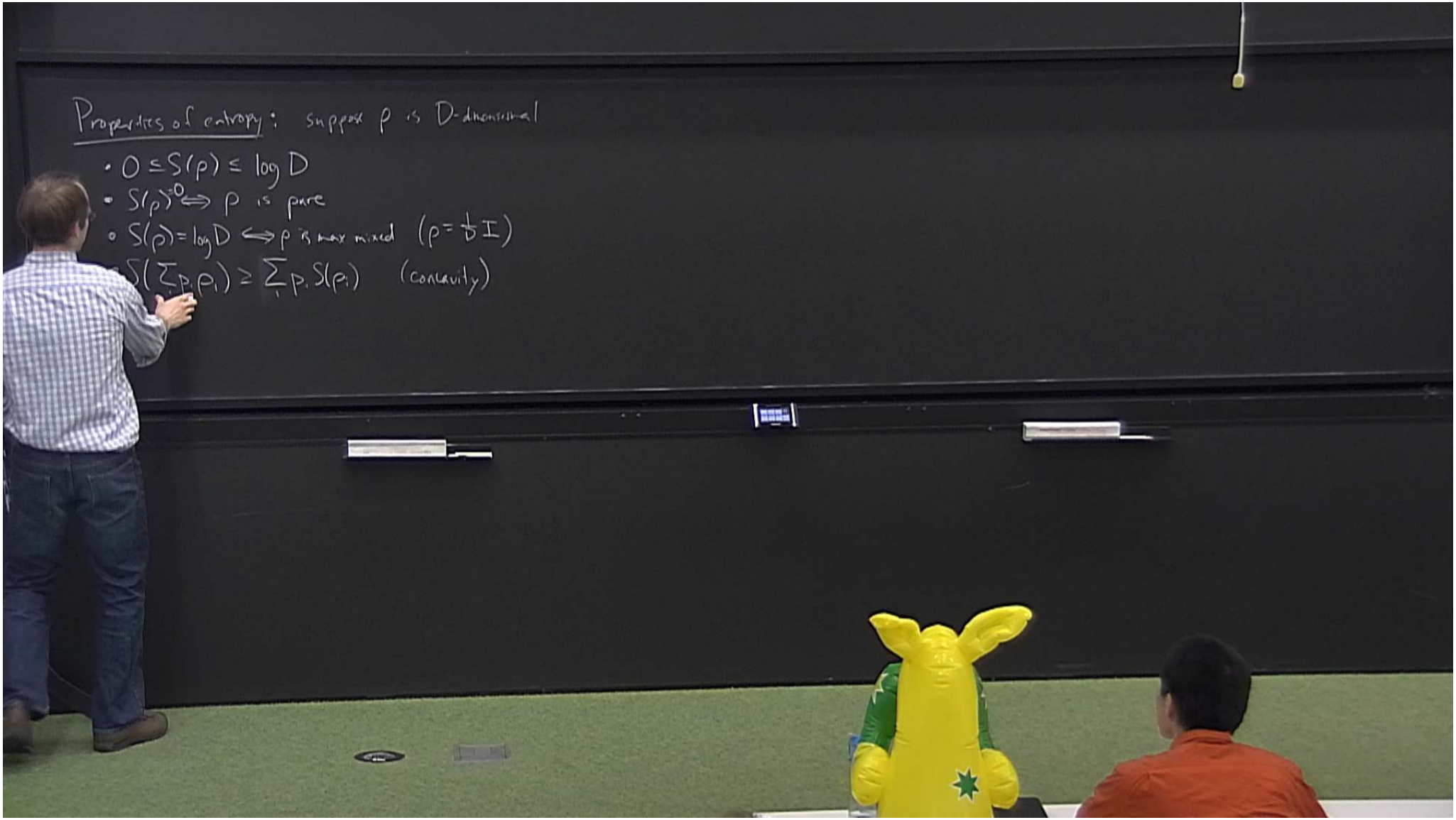

- Entropy
- Compression
Entropy
Shannon entropy quantifies the uncertainty in a probability distribution
Def: $H(p) = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i$
convention: $\log_2$ units: bits
Ex: entropy of a fair coin.
$-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = -\log\frac{1}{2} = \log 2 = 1$
Similarly, entropy in a completely biased coin = 0.

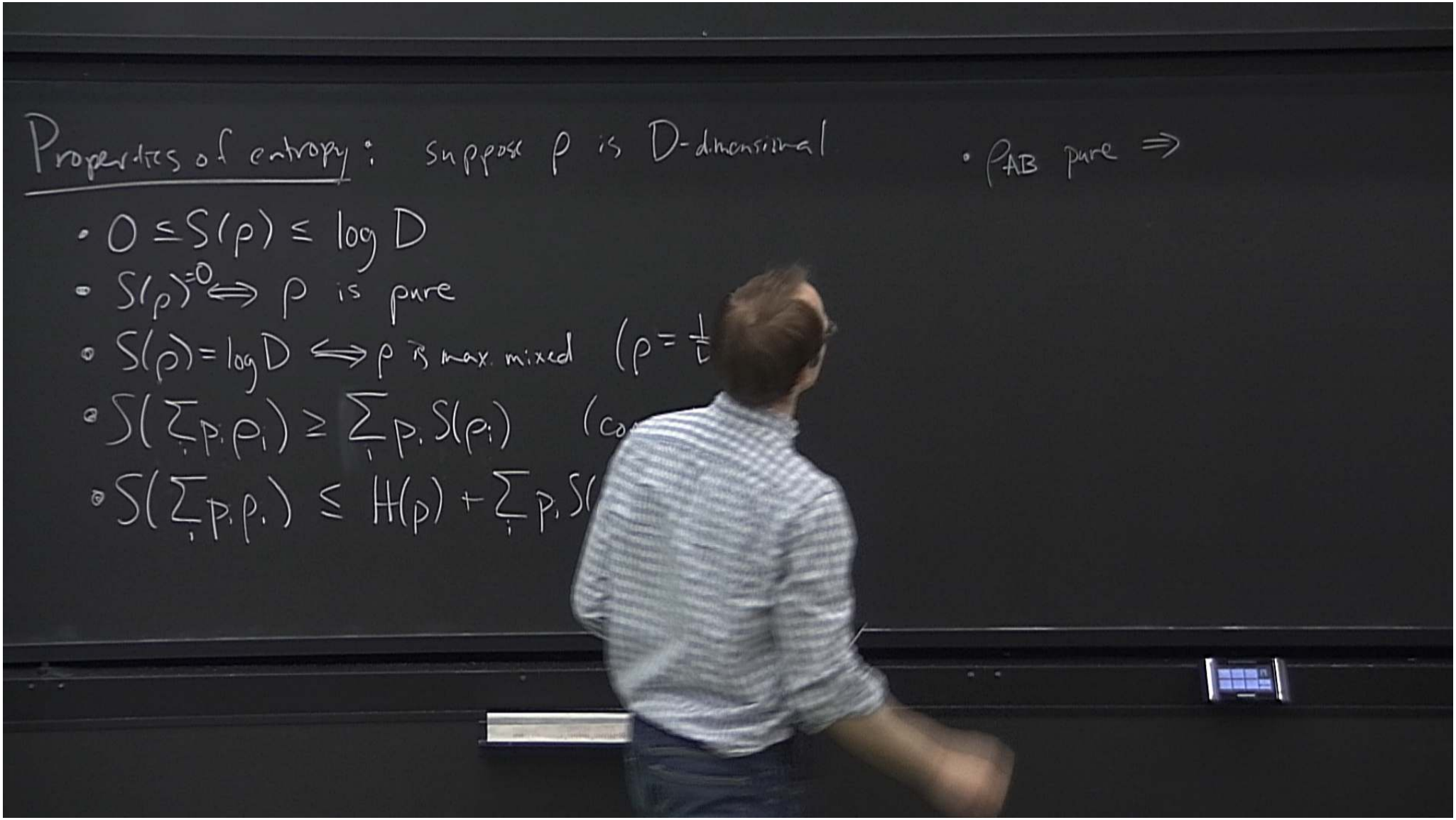

- Entropy
- Compression
Entropy
Shannon entropy quantifies the uncertainty in a probability distribution
Def: $H(p) = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i$
convention: $\log_2$ units: bits
Ex: entropy of a fair coin:
$-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = -\log\frac{1}{2} = \log 2 = 1$
similarly, entropy in a completely biased coin = 0.
→ compression

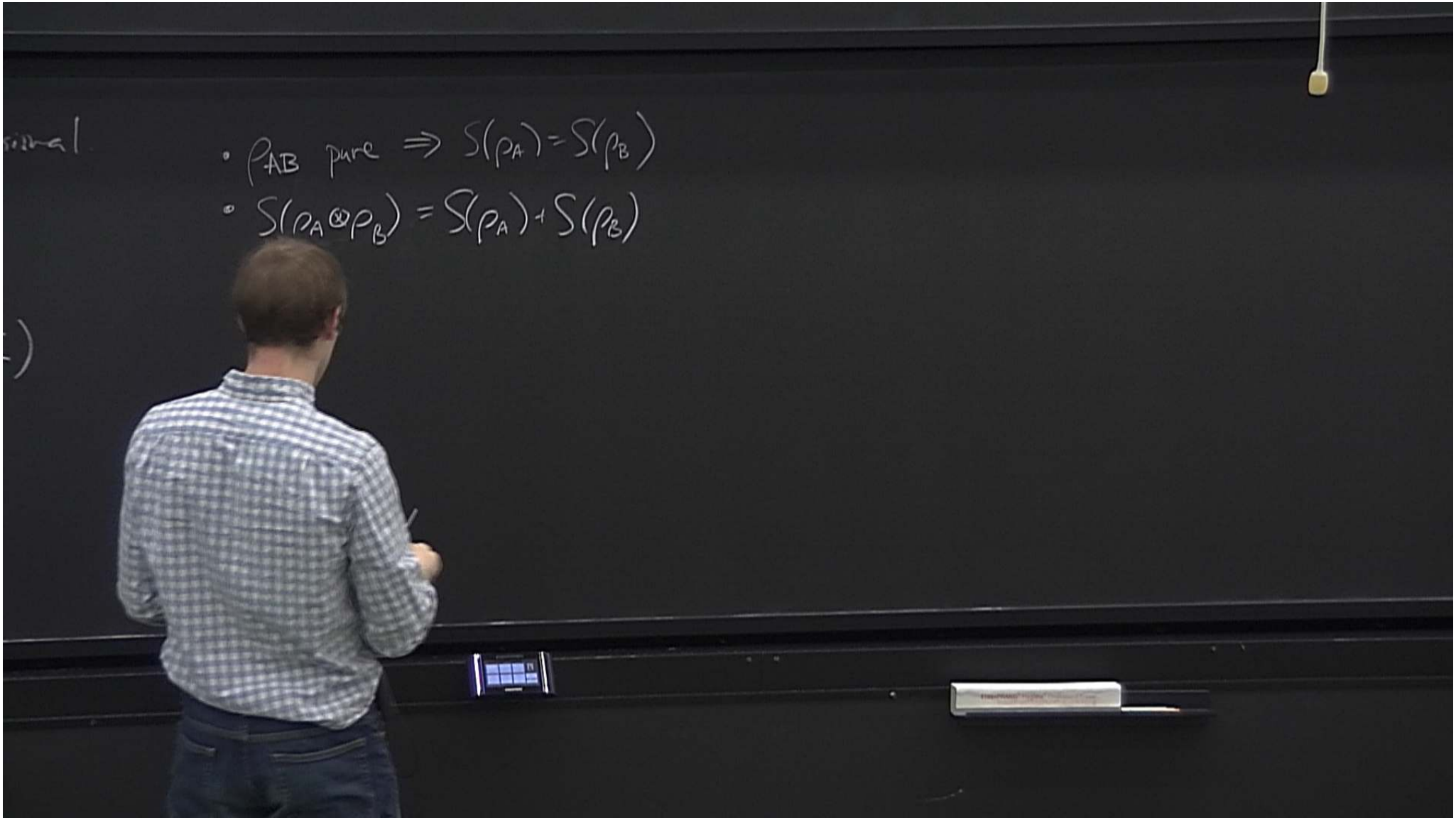

- Entropy
- Compression
Entropy
Shannon entropy quantifies the uncertainty in a probability distribution
Def: $H(p) = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i$
convention: $\log_2$ units: bits
Ex: entropy of a fair coin.
$-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = -\log\frac{1}{2} = \log 2 = 1$
Similarly, entropy in a completely biased coin = 0.
→ compression
Def: (von Neumann) entropy of ρ is $S(\rho) = -\mathrm{tr}(\rho \log \rho)$

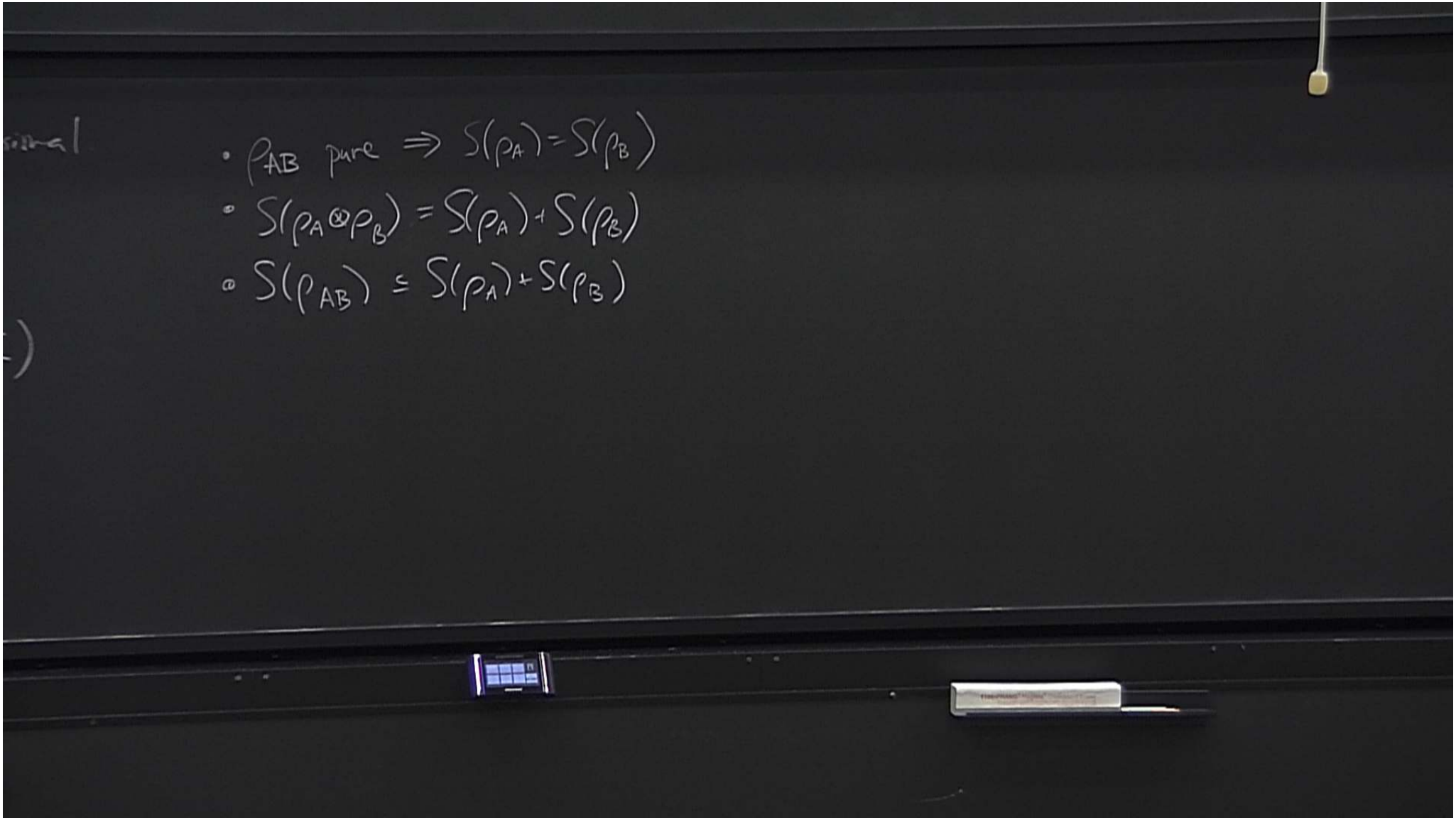

- Entropy
- Compression
Entropy
Shannon entropy quantifies the uncertainty in a probability distribution
Def: $H(p) = \sum_i p_i \log \frac{1}{p_i} = -\sum_i p_i \log p_i$
convention: $\log_2$ units: bits
Ex: entropy of a fair coin
$-\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = -\log\frac{1}{2} = \log 2 = 1$
similarly, entropy in a completely biased coin = 0
-> compression
Def: (von Neumann) entropy of ρ is $S(\rho) = -\mathrm{tr}(\rho \log \rho)$
$= H(\text{eigenvalues of } \rho)$

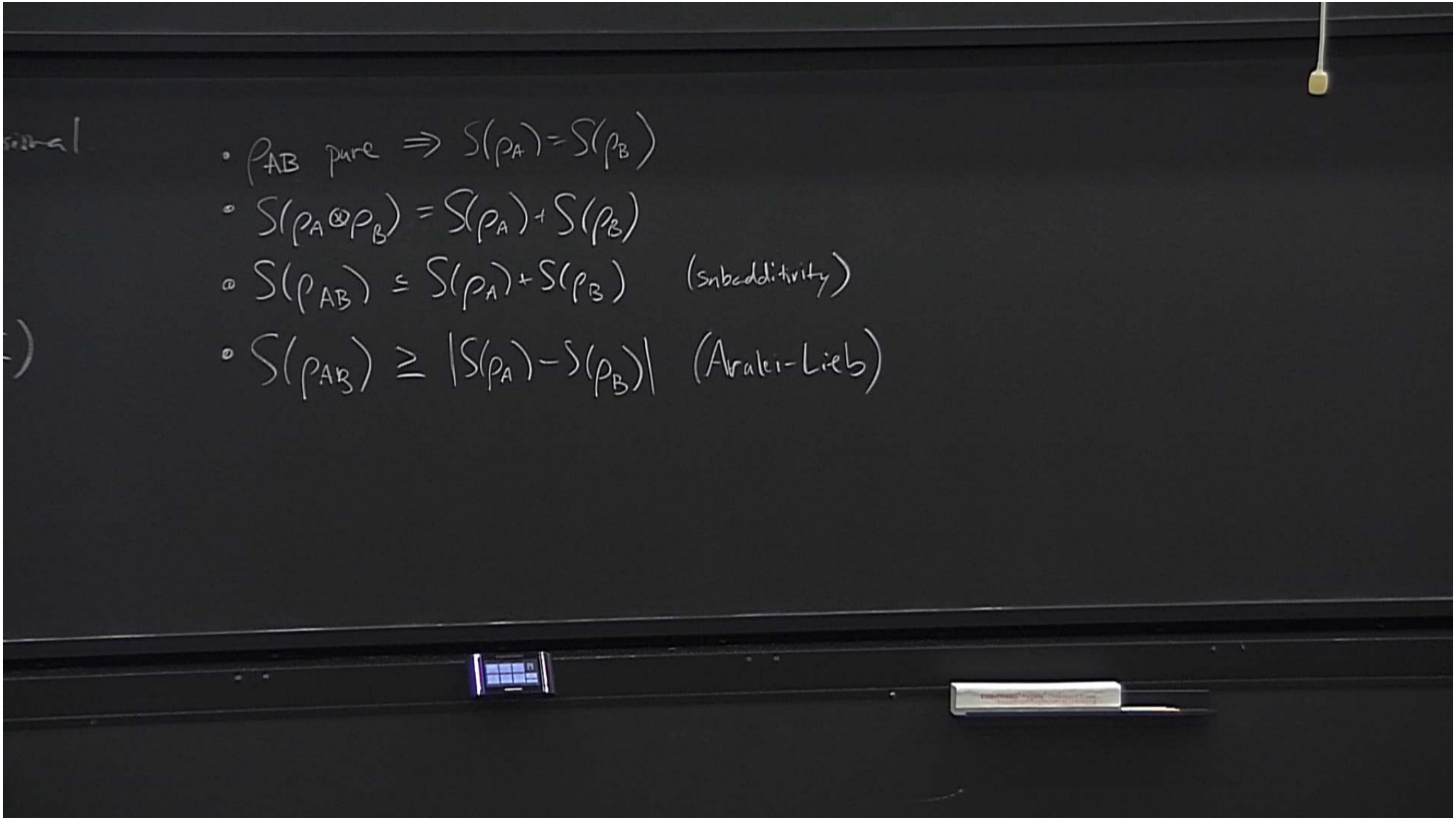

Properties of entropy: suppose ρ is D-dimensional
0 ≤ S(ρ) ≤ log D
S(ρ) = 0 ⇔ ρ is pure
S(ρ) = log D ⇔ ρ is max mixed (ρ = 1/D I)

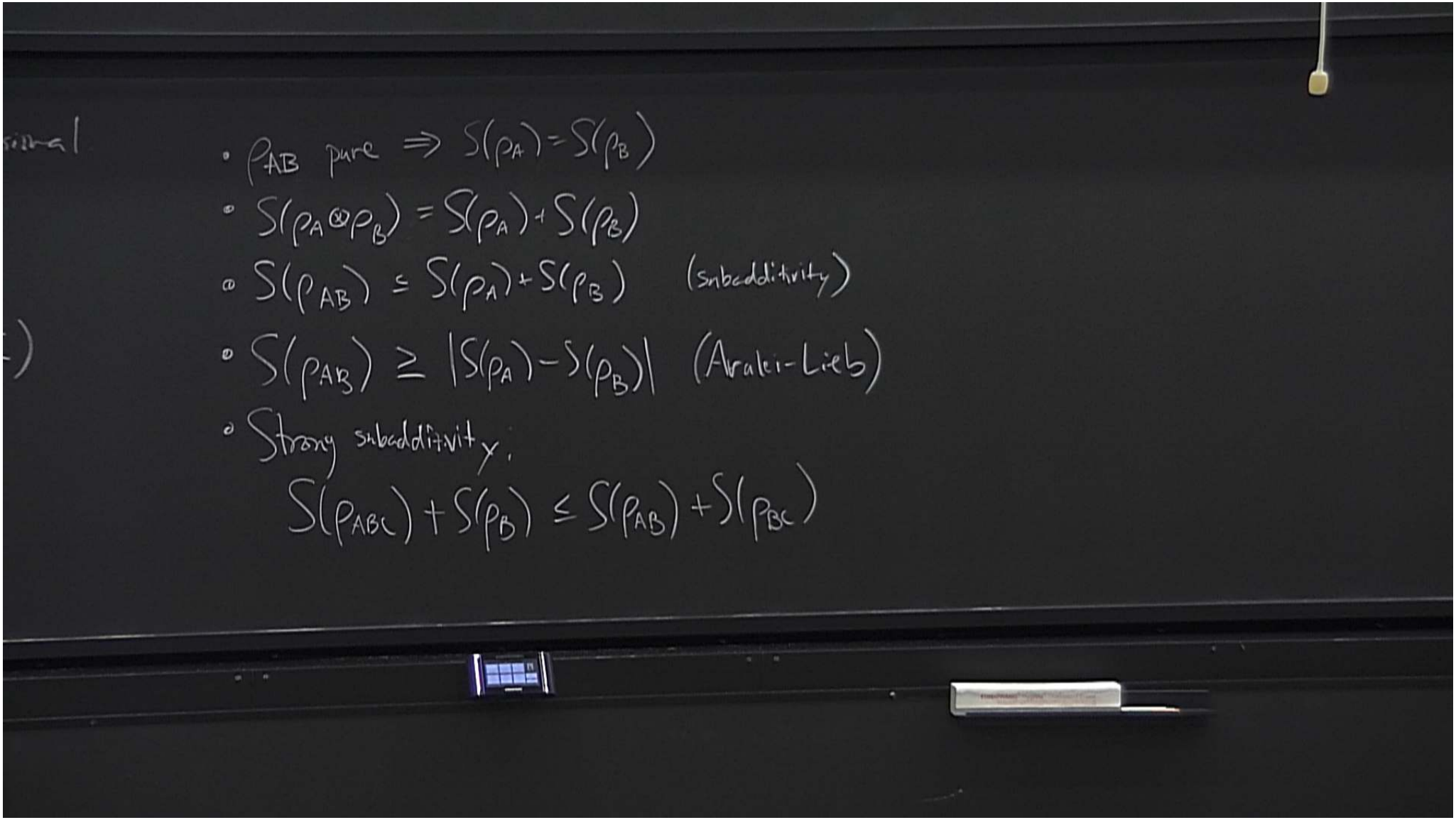

Properties of entropy: suppose ρ is D-dimensional
• 0 ≤ S(ρ) ≤ log D
• S(ρ) = 0 ⇔ ρ is pure
• S(ρ) = log D ⇔ ρ is max mixed (ρ = 1/D I)
S(Σ p_i ρ_i) ≥ Σ p_i S(ρ_i) (concavity)

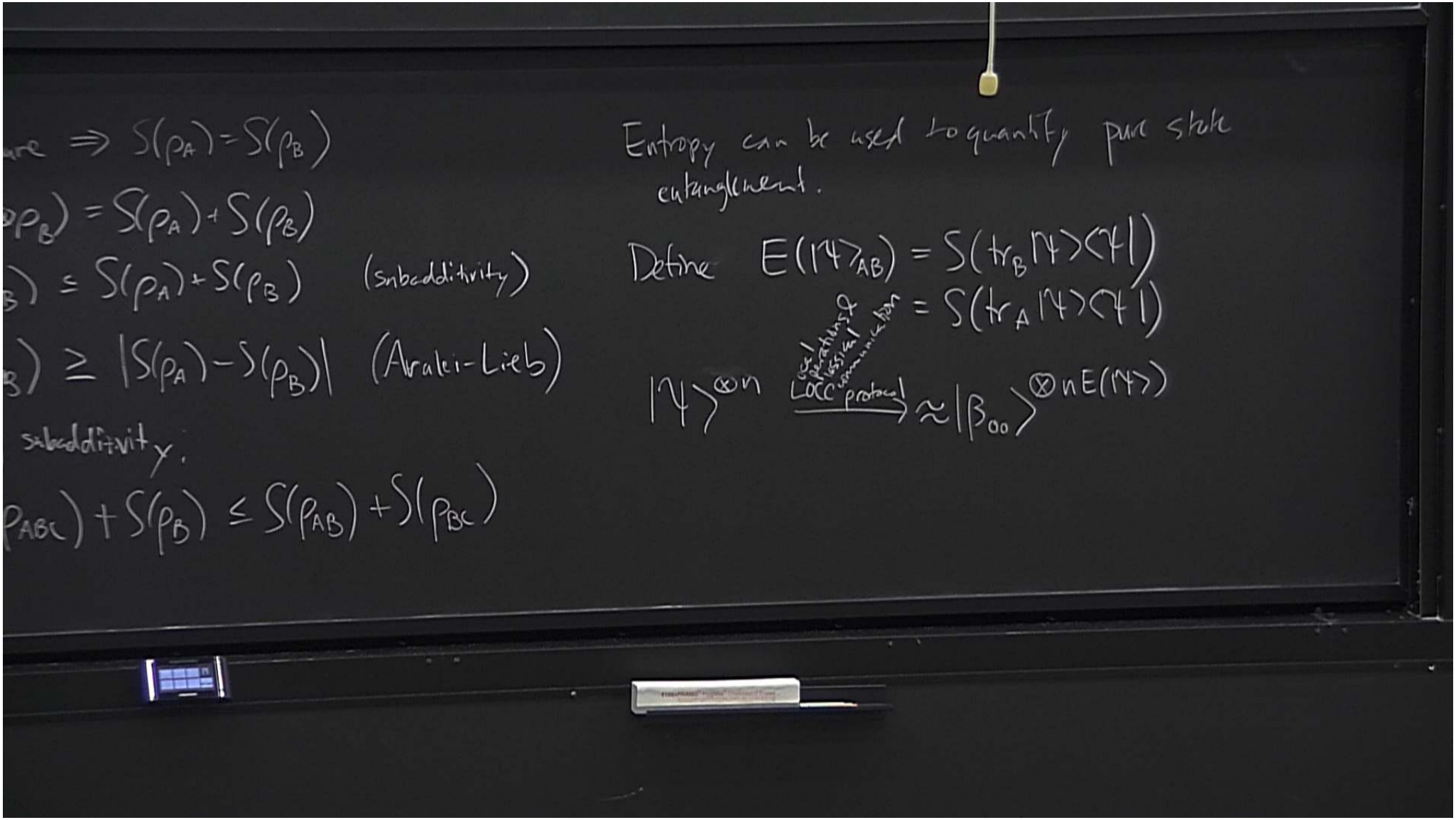

Properties of entropy: suppose ρ is D-dimensional
• ρAB pure ⇒
• 0 ≤ S(ρ) ≤ log D
• S(ρ) = 0 ⇔ ρ is pure
• S(ρ) = log D ⇔ ρ is max. mixed (ρ = 1/D
• S(Σi pi ρi) ≥ Σi pi S(ρi) (co
• S(Σi pi ρi) ≤ H(p) + Σi pi S(

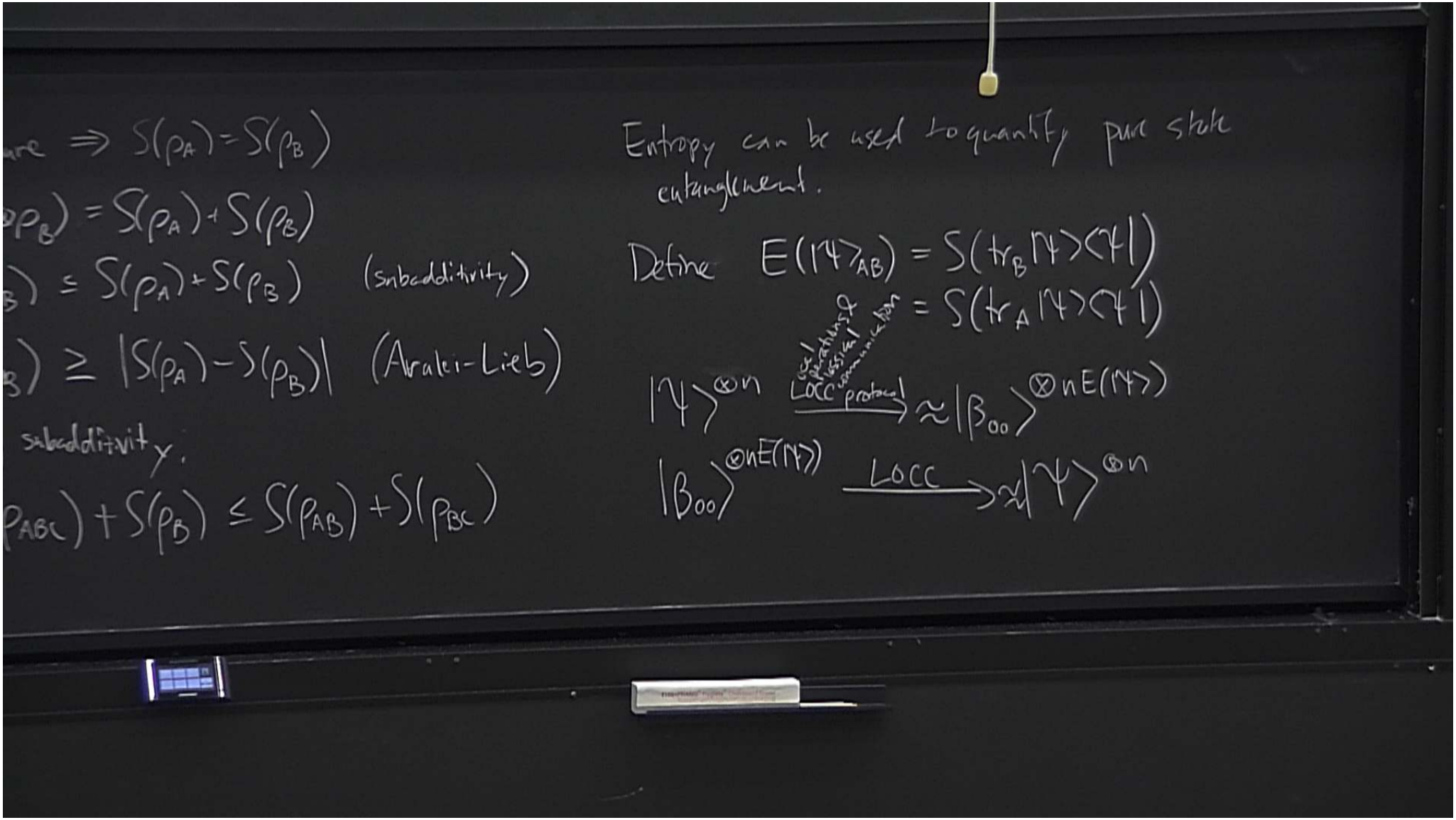

• ρ_{AB} pure $\Rightarrow S(\rho_A) = S(\rho_B)$
• $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$

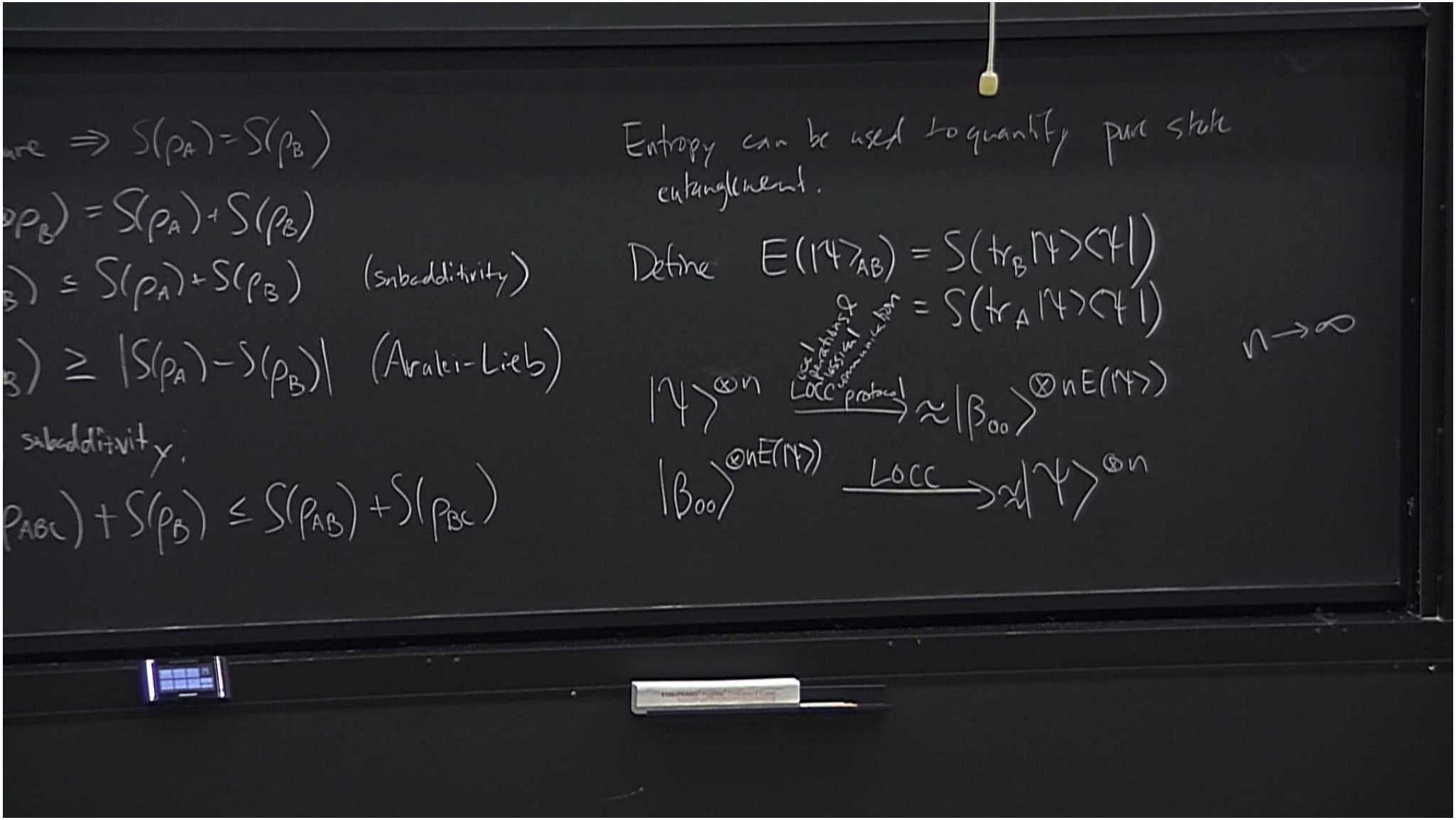

• ρ_{AB} pure $\Rightarrow S(\rho_A) = S(\rho_B)$
• $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$
• $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$

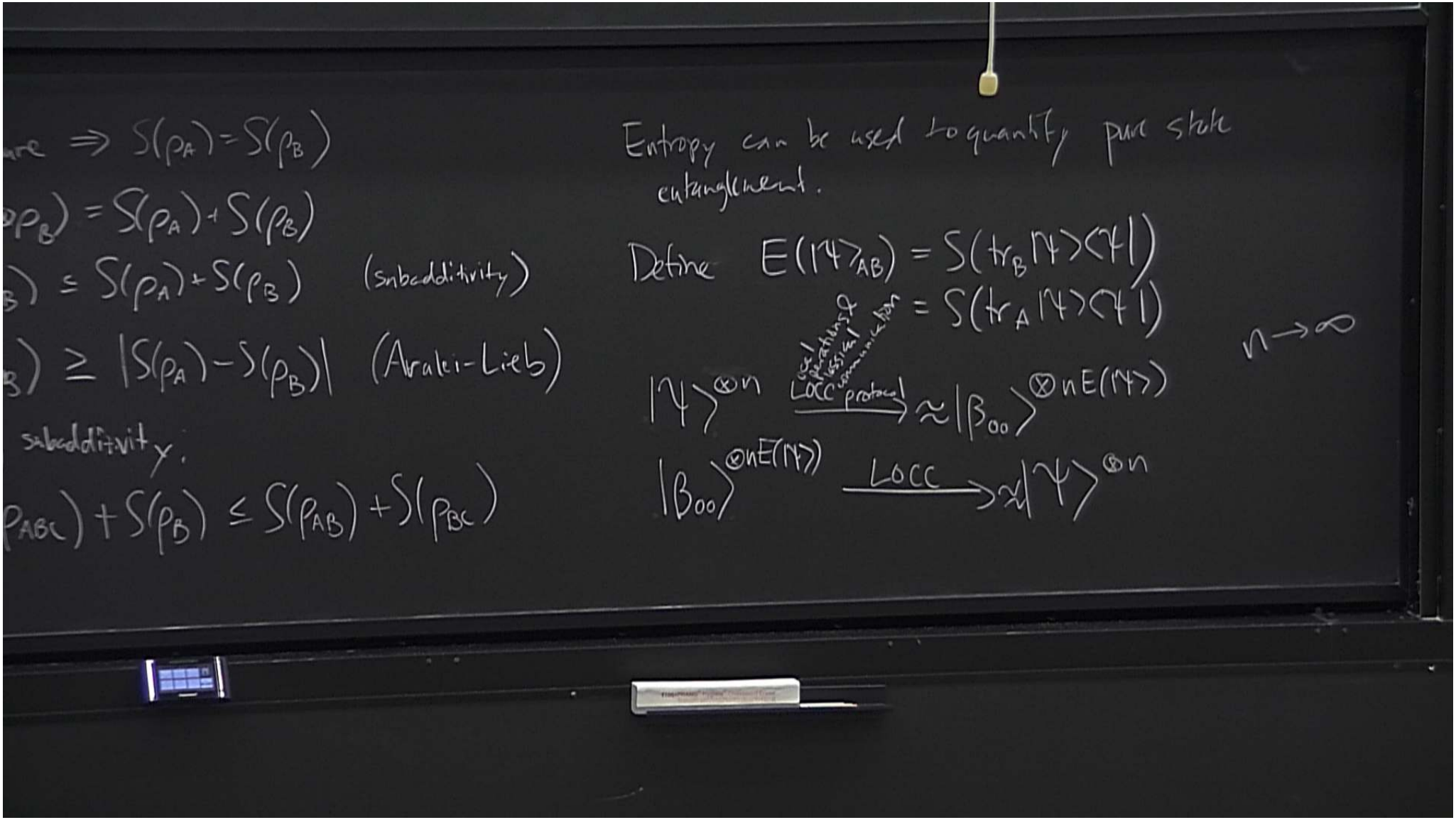

• ρ_{AB} pure $\Rightarrow S(\rho_A) = S(\rho_B)$
• $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$
• $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ (subadditivity)
• $S(\rho_{AB}) \geq |S(\rho_A) - S(\rho_B)|$ (Araki-Lieb)

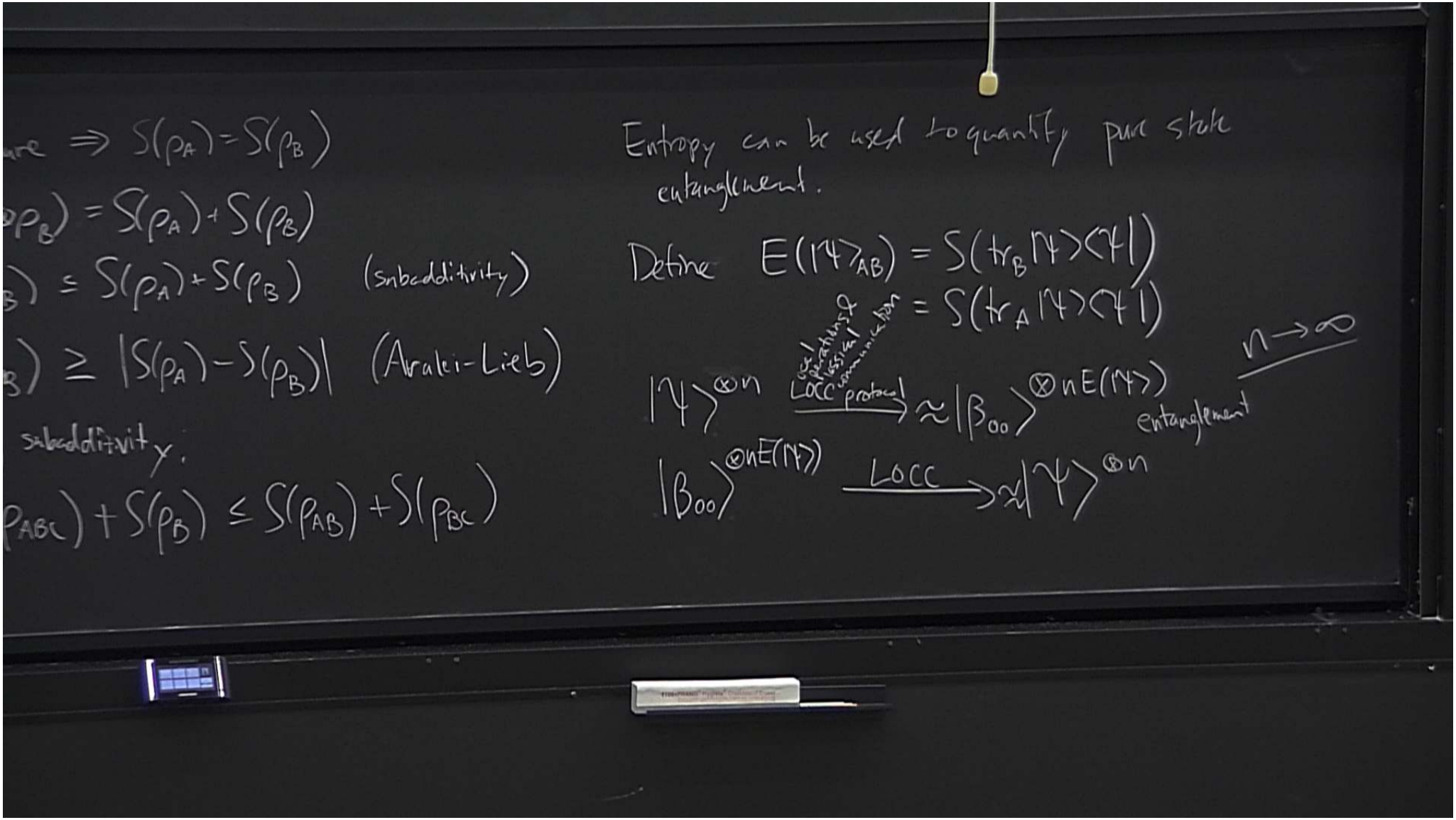

• ρ_{AB} pure $\Rightarrow S(\rho_A) = S(\rho_B)$
• $S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$
• $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ (subadditivity)
• $S(\rho_{AB}) \geq |S(\rho_A) - S(\rho_B)|$ (Araki-Lieb)
• Strong subadditivity:
$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$


re ⇒ S(ρA) = S(ρB)
ρB) = S(ρA) + S(ρB)
) ≤ S(ρA) + S(ρB) (subadditivity)
) ≥ |S(ρA) − S(ρB)| (Araki-Lieb)
subadditivity.
ρABC) + S(ρB) ≤ S(ρAB) + S(ρBC)
Entropy can be used to quantify pure state entanglement.
Define E(|ψ⟩AB) = S(trB|ψ⟩⟨ψ|)
= S(trA|ψ⟩⟨ψ|)
|ψ⟩⊗n
local operations & classical communication
LOCC protocol
≈ |β00⟩⊗nE(|ψ⟩)


$\Rightarrow S(\rho_A) = S(\rho_B)$
$S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$
$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ (subadditivity)
$S(\rho_{AB}) \geq |S(\rho_A) - S(\rho_B)|$ (Araki-Lieb)
subadditivity.
$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$
Entropy can be used to quantify pure state entanglement.
Define $E(|\psi\rangle_{AB}) = S(\mathrm{tr}_B |\psi\rangle\langle\psi|)$
$= S(\mathrm{tr}_A |\psi\rangle\langle\psi|)$
$|\psi\rangle^{\otimes n}$
local quantum & classical communication
LOCC protocol
$\approx |\beta_{00}\rangle^{\otimes nE(|\psi\rangle)}$
$|\beta_{00}\rangle^{\otimes nE(|\psi\rangle)}$
LOCC
$\approx |\psi\rangle^{\otimes n}$


are ⇒ S(ρA) = S(ρB)
ρB) = S(ρA) + S(ρB)
) ≤ S(ρA) + S(ρB) (subadditivity)
) ≥ |S(ρA) − S(ρB)| (Araki-Lieb)
subadditivity.
ρABC) + S(ρB) ≤ S(ρAB) + S(ρBC)
Entropy can be used to quantify pure state entanglement.
Define E(|ψ⟩AB) = S(trB|ψ⟩⟨ψ|)
= S(trA|ψ⟩⟨ψ|)
|ψ⟩⊗n
local operations & classical communication
LOCC protocol
≈ |β00⟩⊗nE(|ψ⟩)
n → ∞
|β00⟩⊗nE(|ψ⟩)
LOCC
≈ |ψ⟩⊗n


are ⇒ S(ρA) = S(ρB)
ρB) = S(ρA) + S(ρB)
) ≤ S(ρA) + S(ρB) (subadditivity)
) ≥ |S(ρA) − S(ρB)| (Araki-Lieb)
subadditivity.
ρABC) + S(ρB) ≤ S(ρAB) + S(ρBC)
Entropy can be used to quantify pure state entanglement.
Define E(|ψ⟩AB) = S(trB|ψ⟩⟨ψ|)
= S(trA|ψ⟩⟨ψ|)
|ψ⟩⊗n
local operations & classical communication
LOCC protocol
≈ |β00⟩⊗nE(|ψ⟩)
n → ∞
|β00⟩⊗nE(|ψ⟩)
LOCC
≈ |ψ⟩⊗n


$\Rightarrow S(\rho_A) = S(\rho_B)$
$S(\rho_A) + S(\rho_B)$
$\leq S(\rho_A) + S(\rho_B)$ (subadditivity)
$\geq |S(\rho_A) - S(\rho_B)|$ (Araki-Lieb)
subadditivity.
$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$
Entropy can be used to quantify pure state entanglement.
Define $E(|\psi\rangle_{AB}) = S(\mathrm{tr}_B |\psi\rangle\langle\psi|)$
$= S(\mathrm{tr}_A |\psi\rangle\langle\psi|)$
local operations & classical communication
$|\psi\rangle^{\otimes n}$ LOCC protocol $\approx |\beta_{00}\rangle^{\otimes nE(|\psi\rangle)}$
$n \to \infty$
entanglement
$|\beta_{00}\rangle^{\otimes nE(|\psi\rangle)}$ LOCC $\to \approx |\psi\rangle^{\otimes n}$

$\text{are} \Rightarrow S(\rho_A) = S(\rho_B)$

$\rho\rho_B) = S(\rho_A) + S(\rho_B)$

$_B) \leq S(\rho_A) + S(\rho_B)$ (subadditivity)

$_B) \geq |S(\rho_A) - S(\rho_B)|$ (Araki-Lieb)

subadditivity.

$\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$

Entropy can be used to quantify pure state entanglement.

Define $E(|\psi\rangle_{AB}) = S(\mathrm{tr}_B |\psi\rangle\langle\psi|)$
$= S(\mathrm{tr}_A |\psi\rangle\langle\psi|)$

$|\psi\rangle^{\otimes n}$ —(local operations & classical communication, LOCC protocol)→ $\approx |\beta_{00}\rangle^{\otimes nE(|\psi\rangle)}$ ($n \to \infty$: entanglement concentration)

$|\beta_{00}\rangle^{\otimes nE(|\psi\rangle)}$ —LOCC→ $\approx |\psi\rangle^{\otimes n}$ (entanglement dilution)

$\text{are} \Rightarrow S(\rho_A) = S(\rho_B)$

$\rho_B) = S(\rho_A) + S(\rho_B)$

$) \leq S(\rho_A) + S(\rho_B)$ (subadditivity)

$) \geq |S(\rho_A) - S(\rho_B)|$ (Araki-Lieb)

subadditivity.

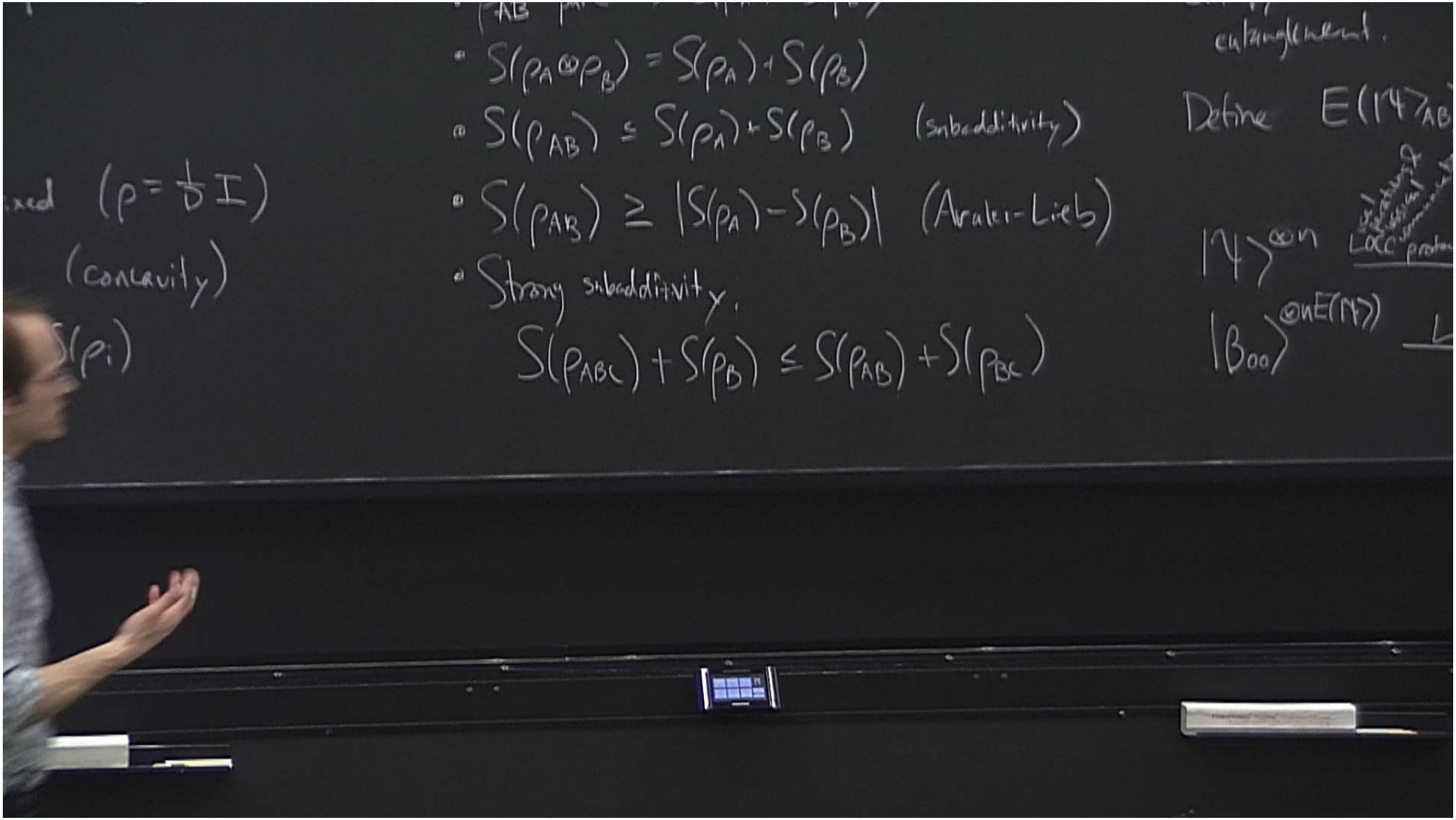
$\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$

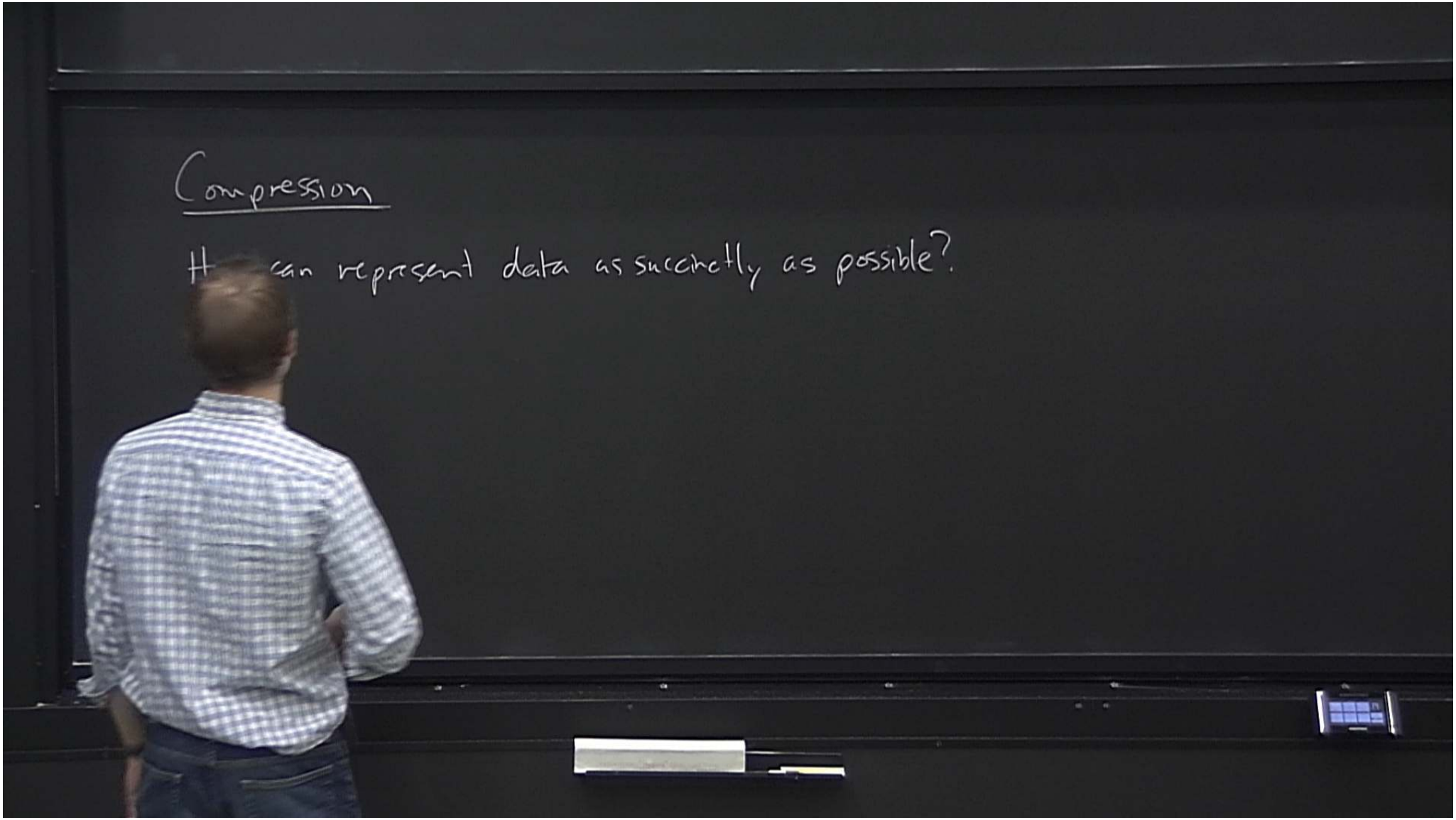
Entropy can be used to quantify pure state entanglement.

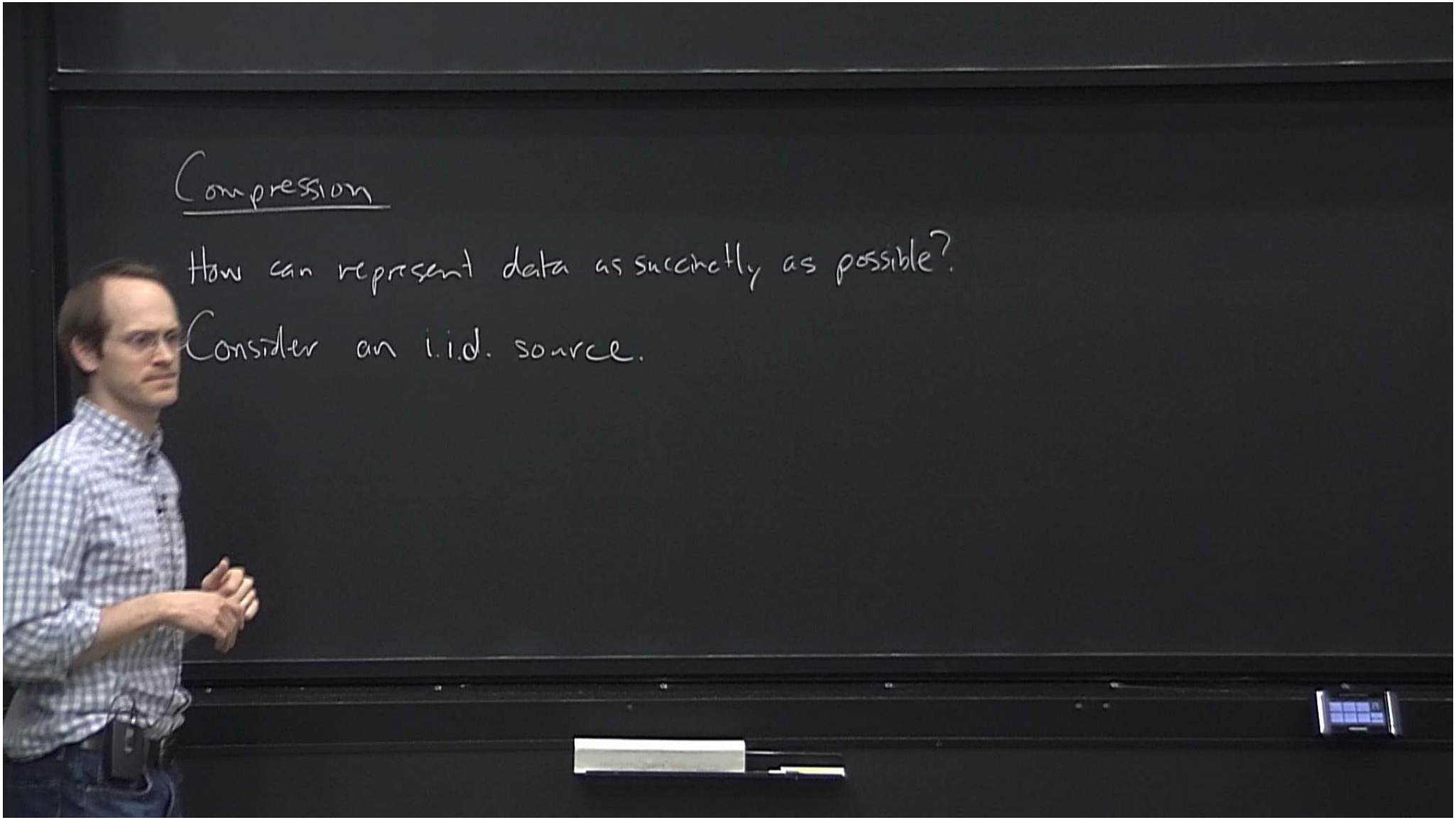
Define $E(|\psi\rangle_{AB}) = S(\mathrm{tr}_B |\psi\rangle\langle\psi|)$
$= S(\mathrm{tr}_A |\psi\rangle\langle\psi|)$

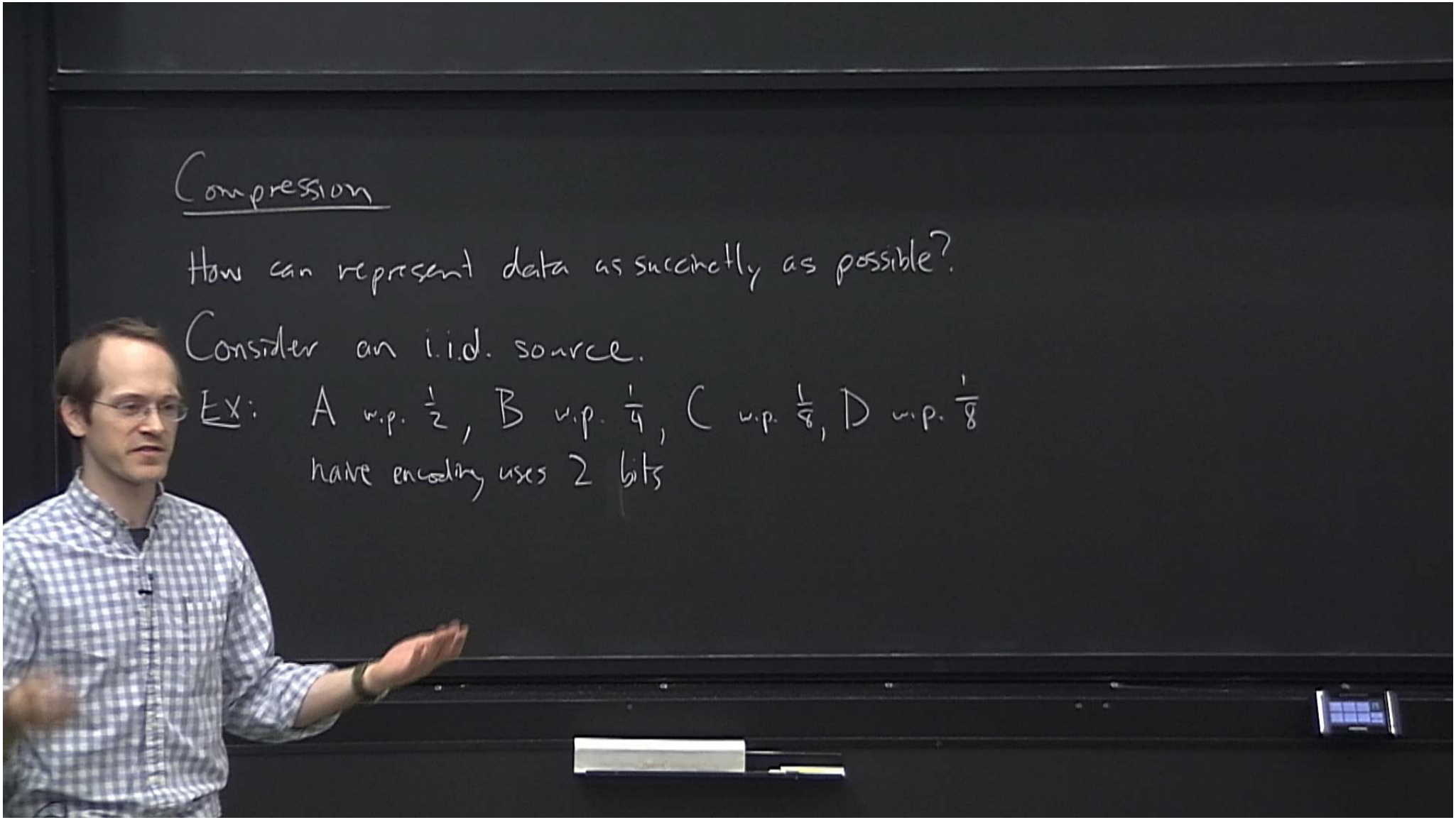
$|\psi\rangle^{\otimes n}$ —(local operations & classical communication, LOCC protocol)→ $\approx |\beta_{00}\rangle^{\otimes nE(|\psi\rangle)}$ ($n \to \infty$: entanglement concentration)

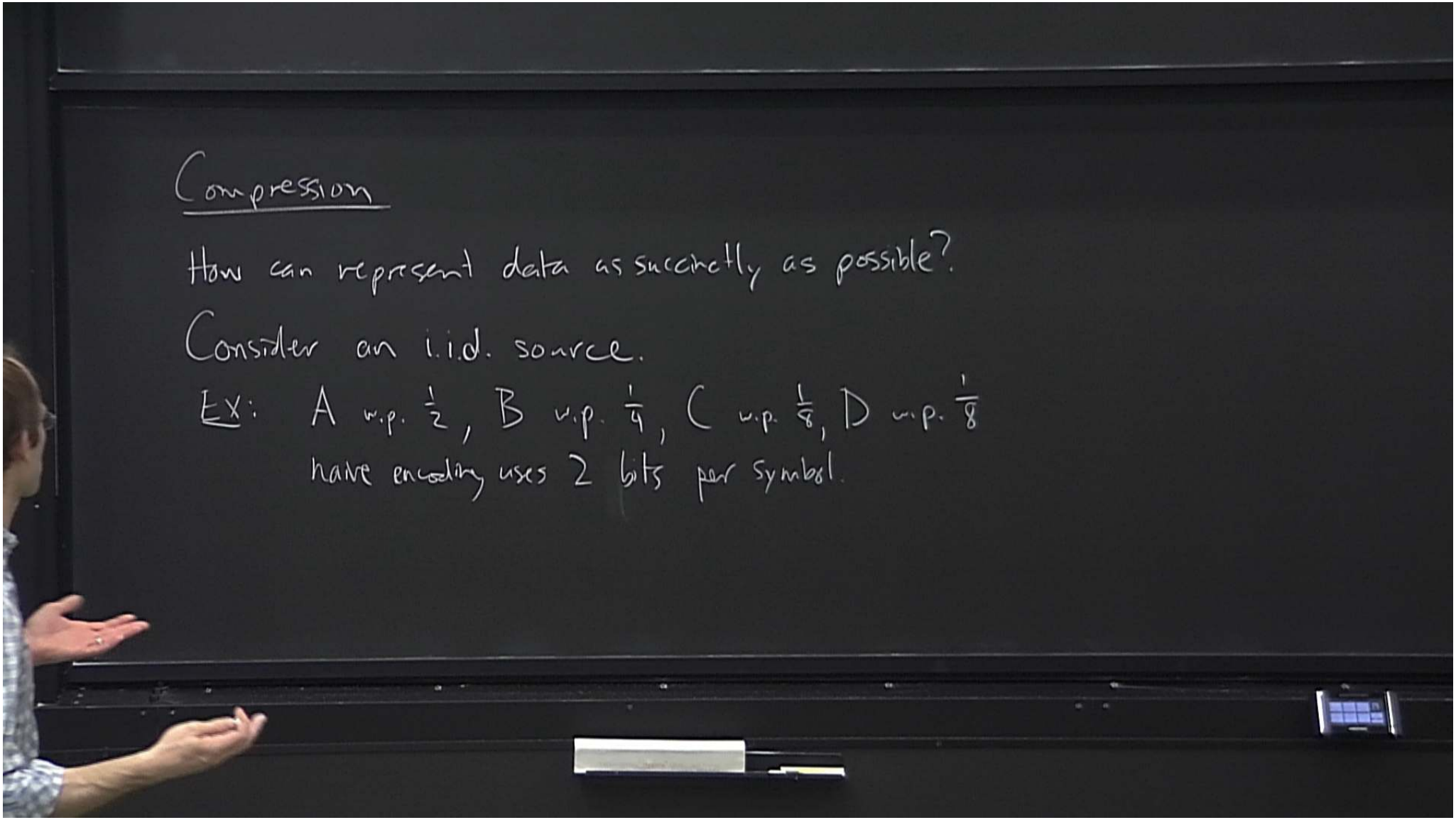
$|\beta_{00}\rangle^{\otimes nE(|\psi\rangle)}$ —LOCC→ $\approx |\psi\rangle^{\otimes n}$ (entanglement dilution)

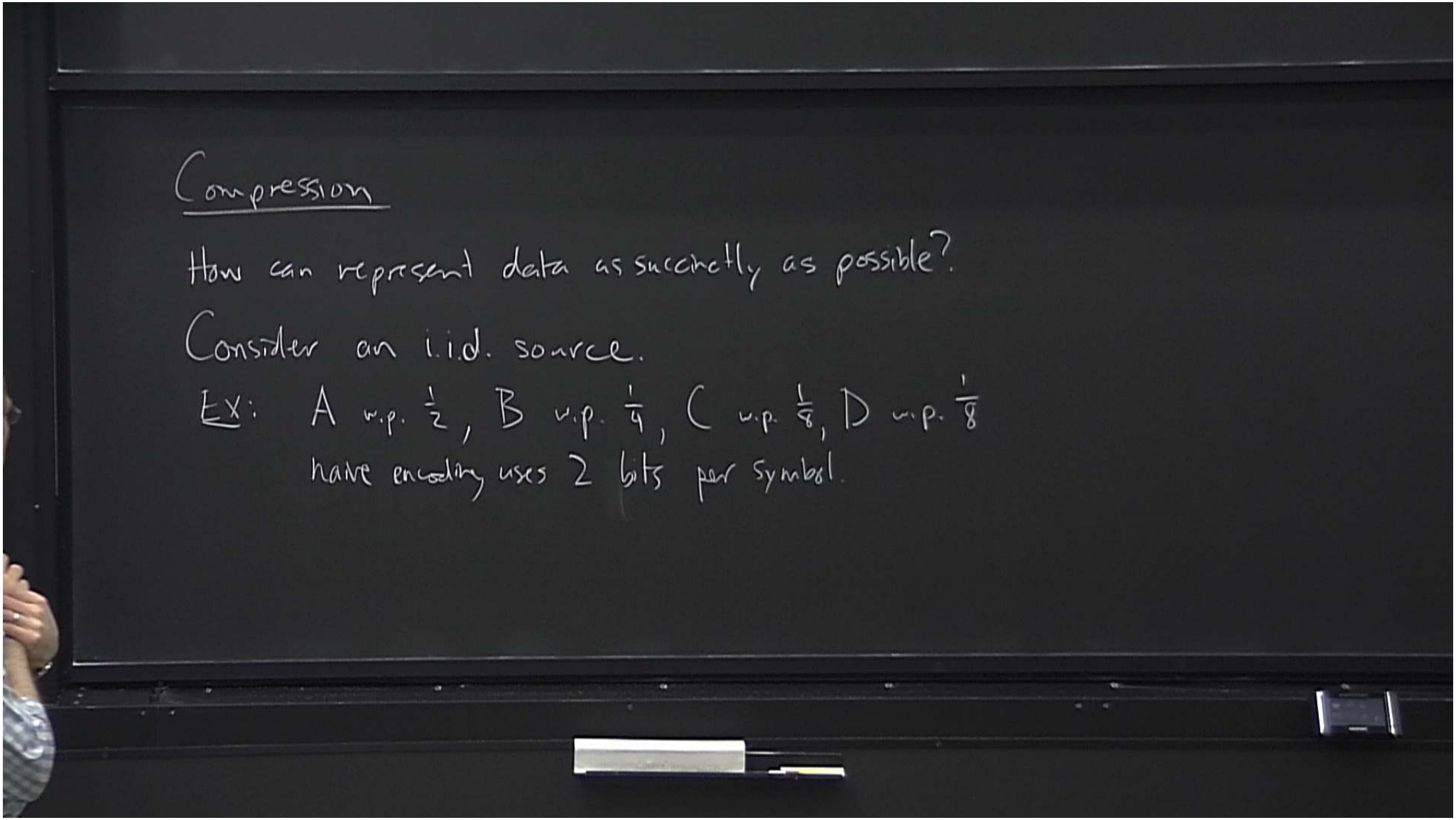

$S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)$
$S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$ (subadditivity)
$S(\rho_{AB}) \geq |S(\rho_A) - S(\rho_B)|$ (Araki-Lieb)
Strong subadditivity.
$S(\rho_{ABC}) + S(\rho_B) \leq S(\rho_{AB}) + S(\rho_{BC})$
entanglement.
Define $E(|\psi\rangle_{AB}$
$|\psi\rangle^{\otimes n}$
$|\beta_{00}\rangle^{\otimes nE(\psi)}$
$(\rho = \frac{1}{D} I)$
(concavity)
$S(\rho_i)$

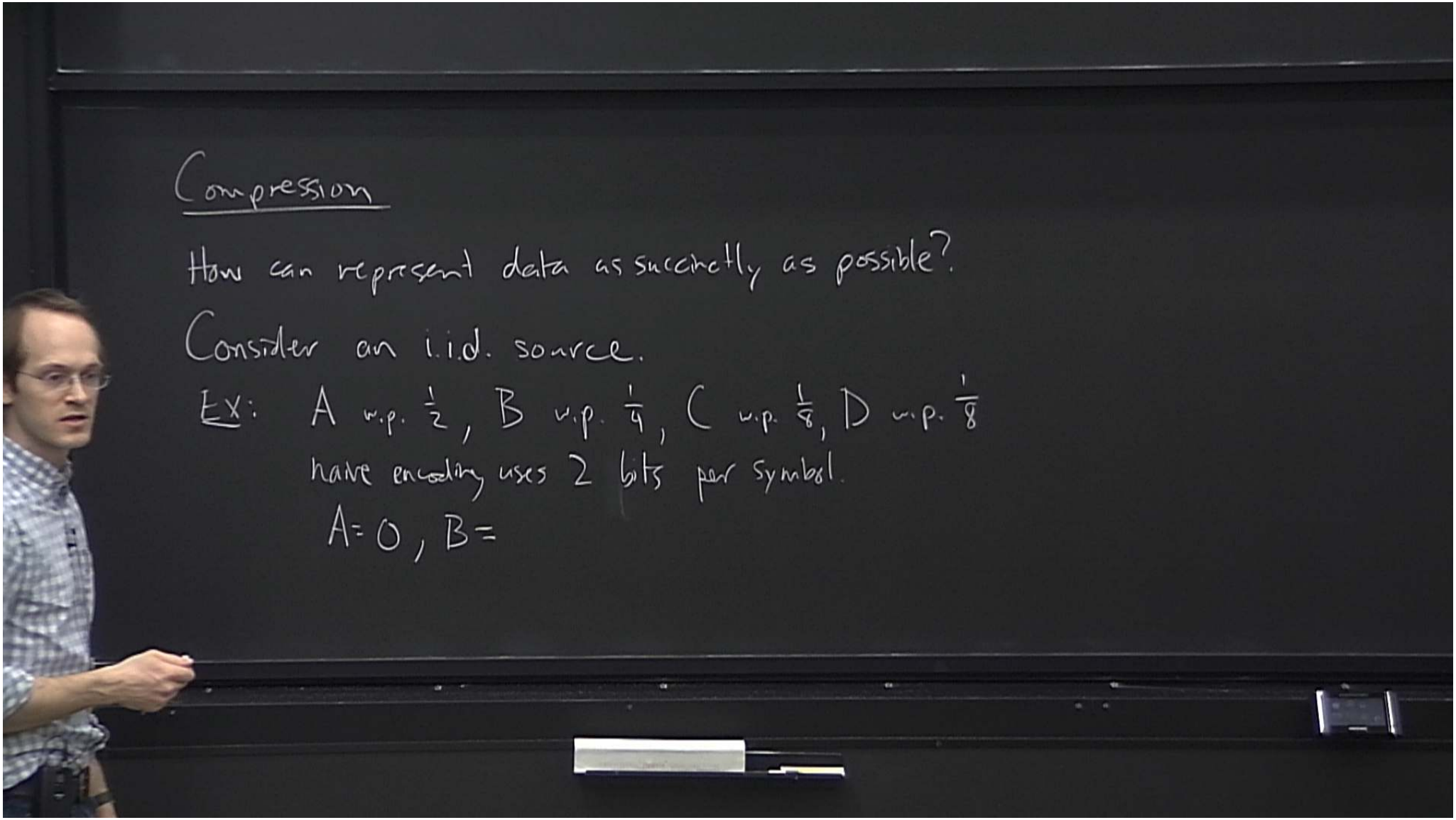

Compression
How can represent data as succinctly as possible?

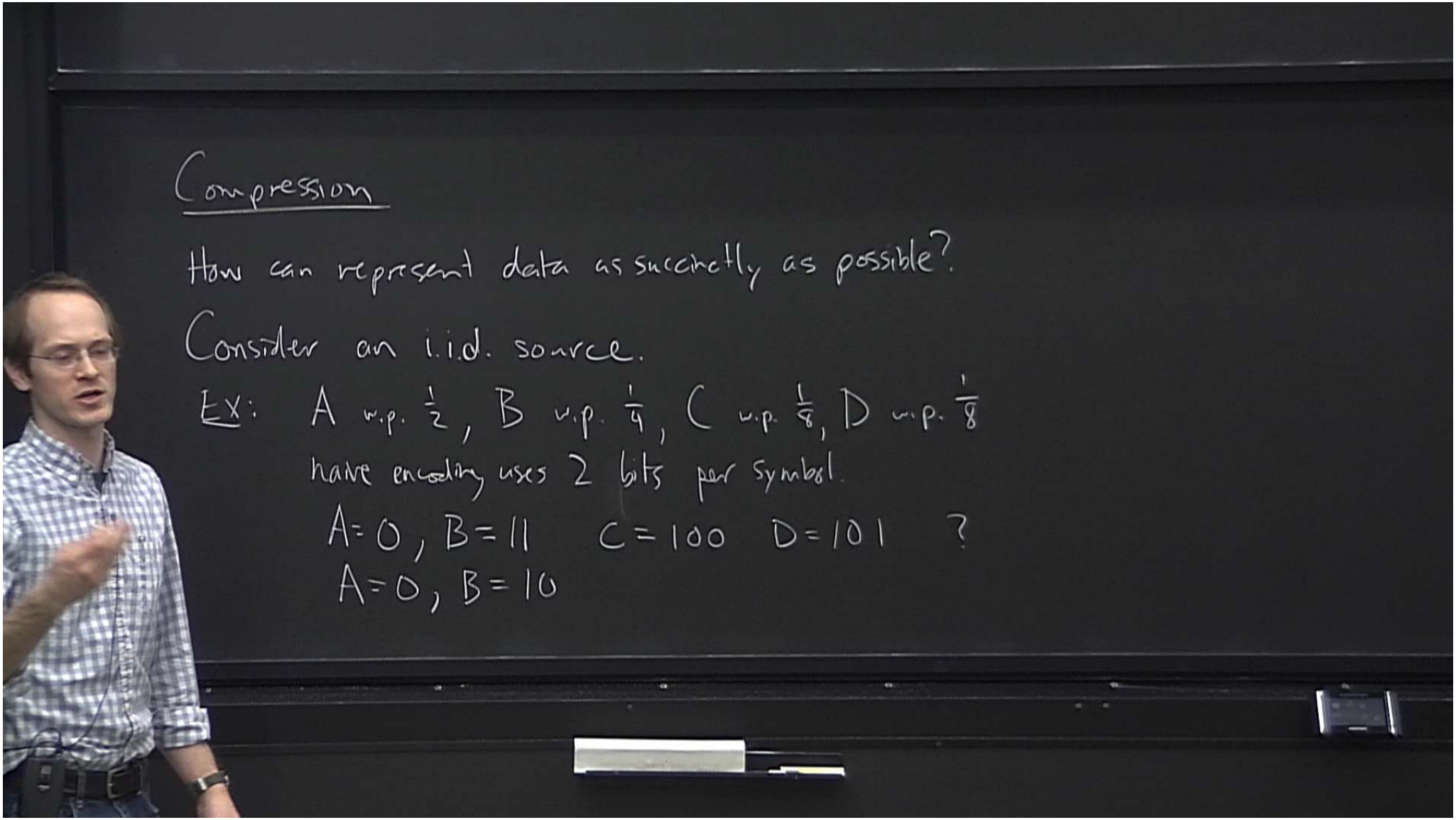

Compression
How can represent data as succinctly as possible?
Consider an i.i.d. source.

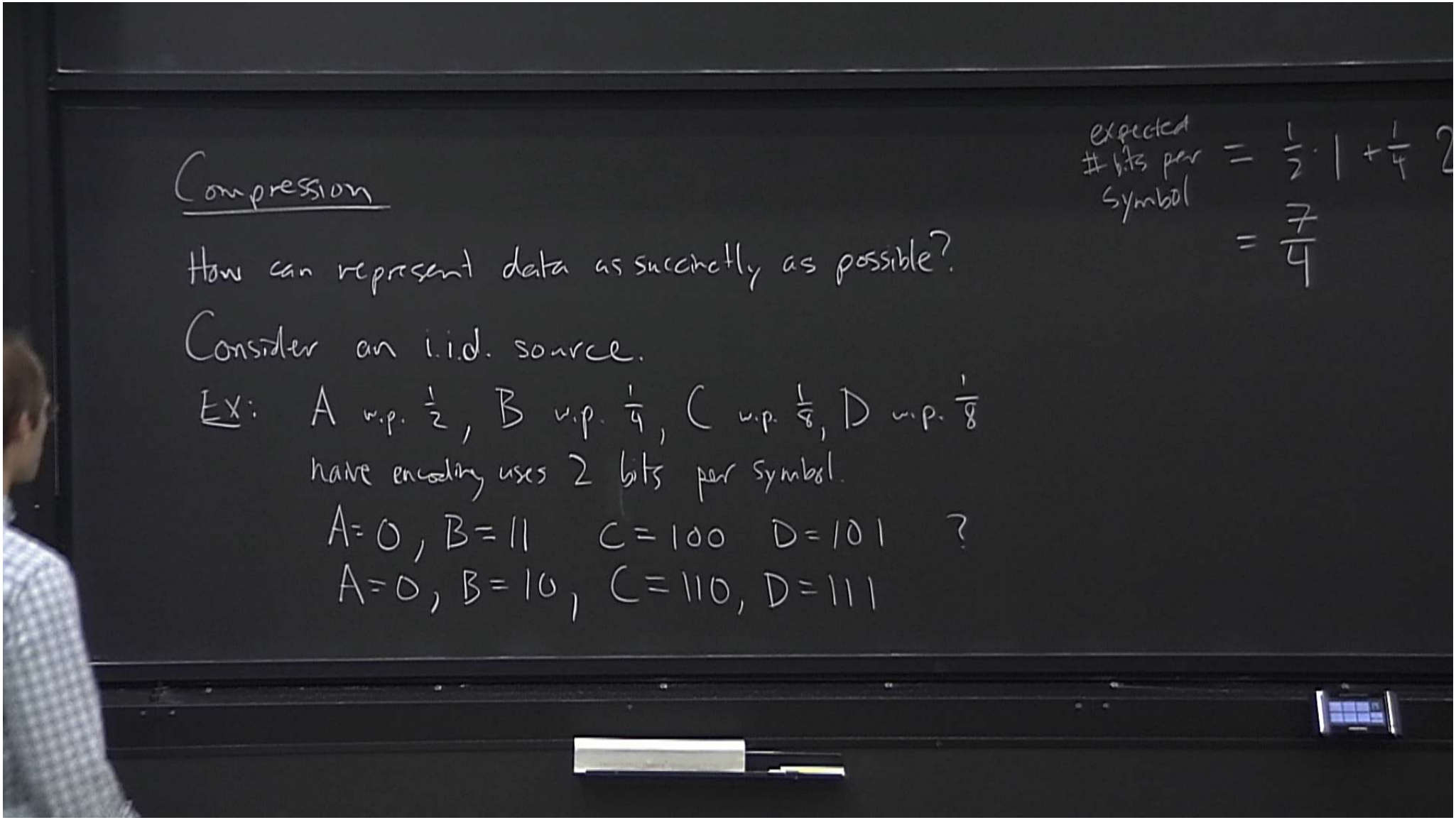

Compression
How can represent data as succinctly as possible?
Consider an i.i.d. source.
Ex: A w.p. 1/2, B w.p. 1/4, C w.p. 1/8, D w.p. 1/8
have encoding uses 2 bits


Compression
How can represent data as succinctly as possible?
Consider an i.i.d. source.
Ex: A w.p. 1/2, B w.p. 1/4, C w.p. 1/8, D w.p. 1/8
have encoding uses 2 bits per symbol.


Compression
How can represent data as succinctly as possible?
Consider an i.i.d. source.
Ex: A w.p. 1/2, B w.p. 1/4, C w.p. 1/8, D w.p. 1/8
have encoding uses 2 bits per symbol.


Compression
How can represent data as succinctly as possible?
Consider an i.i.d. source.
Ex: A w.p. 1/2, B w.p. 1/4, C w.p. 1/8, D w.p. 1/8
have encoding uses 2 bits per symbol.
A=0, B=


Compression
How can represent data as succinctly as possible?
Consider an i.i.d. source.
Ex: A w.p. 1/2, B w.p. 1/4, C w.p. 1/8, D w.p. 1/8
have encoding uses 2 bits per symbol.
A=0, B=11 C=100 D=101 ?
A=0, B=10


Compression
How can represent data as succinctly as possible?
Consider an i.i.d. source.
Ex: A w.p. 1/2, B w.p. 1/4, C w.p. 1/8, D w.p. 1/8
have encoding uses 2 bits per symbol.
A=0, B=11 C=100 D=101 ?
A=0, B=10, C=110, D=111
expected # bits per symbol = 1/2 · 1 + 1/4 2
= 7/4

Compression

How can represent data as succinctly as possible?

Consider an i.i.d. source.

Ex: A w.p. $\frac{1}{2}$, B w.p. $\frac{1}{4}$, C w.p. $\frac{1}{8}$, D w.p. $\frac{1}{8}$

have encoding uses 2 bits per symbol.

A=0, B=11 C=100 D=101 ?

A=0, B=10, C=110, D=111

expected # bits per symbol $= \frac{1}{2} \cdot 1 + \frac{1}{4} 2$

$= \frac{7}{4}$

$H(\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}) = \frac{7}{4}$

as possible?

$\frac{1}{8}$, D w.p. $\frac{1}{8}$

ymbol.

D = 101 ?

D = 111

expected # bits per symbol $= \frac{1}{2} \cdot 1 + \frac{1}{4} 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} 3$

$= \frac{7}{4}$

$H(\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}) = \frac{7}{4}$

In general, we may use block encodings: encode many characters at a time

ent data as succinctly as possible?

i.i.d. source.

$\frac{1}{2}$, B w.p. $\frac{1}{4}$, C w.p. $\frac{1}{8}$, D w.p. $\frac{1}{8}$

oding uses 2 bits per symbol.

B=11 C=100 D=101 ?

, B=10, C=110, D=111

expected # bits per symbol $= \frac{1}{2} \cdot 1 + \frac{1}{4} 2 + \frac{1}{8} 3 + \frac{1}{8} 3$

$= \frac{7}{4}$

$H(\{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}) = \frac{7}{4}$

In general, we may use block encodings: encode many characters at a time

$= \frac{1}{2} \cdot 1 + \frac{1}{4} 2 + \frac{1}{8} 3 + \frac{1}{8} 3$

$= \frac{7}{4}$

$\frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}) = \frac{7}{4}$

al, we may use block
s: encode many characters
ime

More generally, consider a source of bits

0 w.p p

1 w.p. 1-p

A string from this source has

$= \frac{1}{2} \cdot 1 + \frac{1}{4} 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} 3$

$= \frac{7}{4}$

$\frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}) = \frac{7}{4}$

al, we may use block

s: encode many characters

ine

More generally, consider a source of bits

0 w.p. p

1 w.p. $1-p$

A n-bit string from this source has αn 0s with prob

$$p^{\alpha n} (1-p)^{(1-\alpha)n}$$

$= \frac{1}{2} \cdot 1 + \frac{1}{4} 2 + \frac{1}{8} 3 + \frac{1}{8} 3$

$= \frac{7}{4}$

$\frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}) = \frac{7}{4}$

al, we may use blocks

s: encode many characters

ime

More generally, consider a source of bits

0 w.p. p

1 w.p. $1-p$

A particular n-bit string with αn 0s has prob.

$p^{\alpha n} (1-p)^{(1-\alpha)n}$

$= \frac{1}{2}\cdot 1 + \frac{1}{4}\,2 + \frac{1}{8}\cdot 3 + \frac{1}{8}\,3$

$= \frac{7}{4}$

$\frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}) = \frac{7}{4}$

al, we may use blocks

s: encode many characters

ine

More generally, consider a source of bits

0 w.p p

1 w.p. $1-p$

A particular n-bit string with αn 0s has prob.

$$p^{\alpha n}(1-p)^{(1-\alpha)n}$$

We expect $\approx pn$ 0s. The prob. of such a string is

$$p^{pn}(1-p)^{(1-p)n}$$

$$\log\left(p^{pn}(1-p)^{(1-p)n}\right) = pn\log p + (1-p)n\log(1-p)$$

$$= -nH_2(p)$$

where $H_2(p) = -p\log p - (1-p)\log(1-p)$

("binary entropy")

$$\log\left(p^{pn}(1-p)^{(1-p)n}\right) = pn\log p + (1-p)n\log(1-p)$$

$$= -nH_2(p)$$

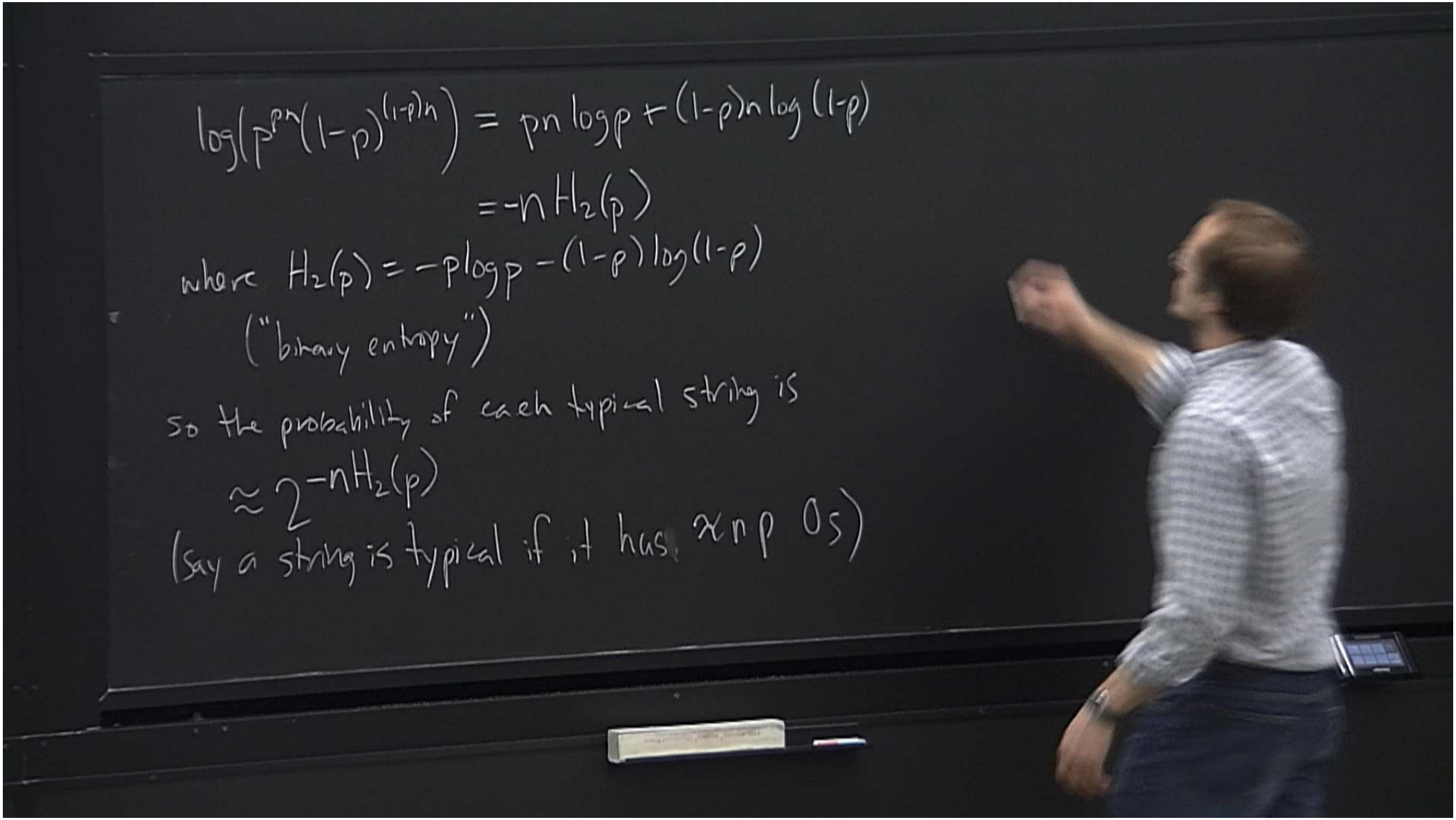
where $H_2(p) = -p\log p - (1-p)\log(1-p)$

("binary entropy")

So the probability of each typical string is

$$\approx 2^{-nH_2(p)}$$

(Say a string is typical if it has $\approx np$ 0s)



$$\log\left(p^{pn}(1-p)^{(1-p)n}\right) = pn\log p + (1-p)n\log(1-p)$$
$$= -nH_2(p)$$

where $H_2(p) = -p\log p - (1-p)\log(1-p)$

("binary entropy")

So the probability of each typical string is

$$\approx 2^{-nH_2(p)}$$

(Say a string is typical if it has $\approx np$ 0s)

There can be at most $\approx 2^{nH}$

$= pn\log p + (1-p)n\log(1-p)$

$= -nH_2(p)$

$= -p\log p - (1-p)\log(1-p)$

"entropy")

of each typical string is

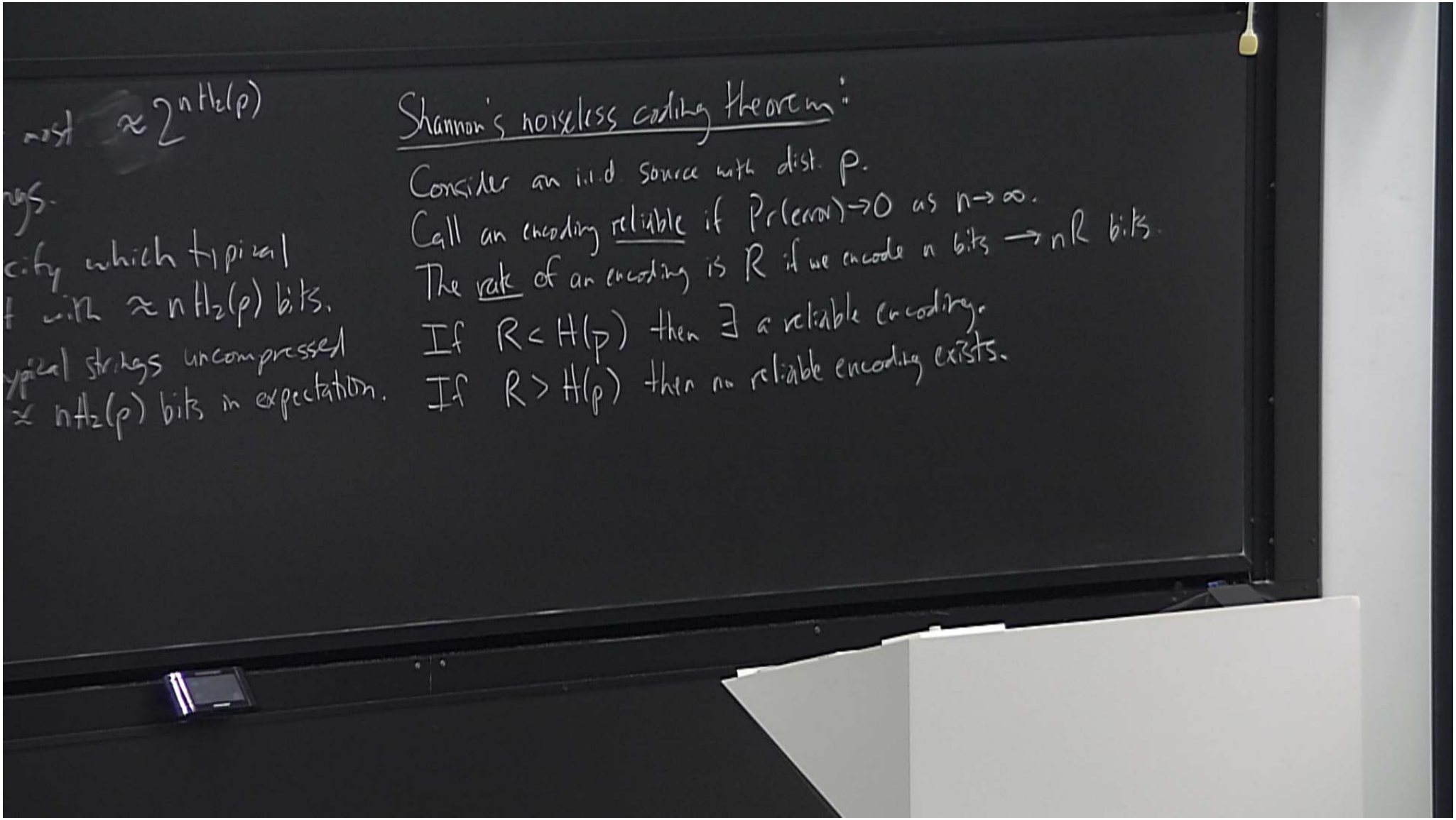
(p)

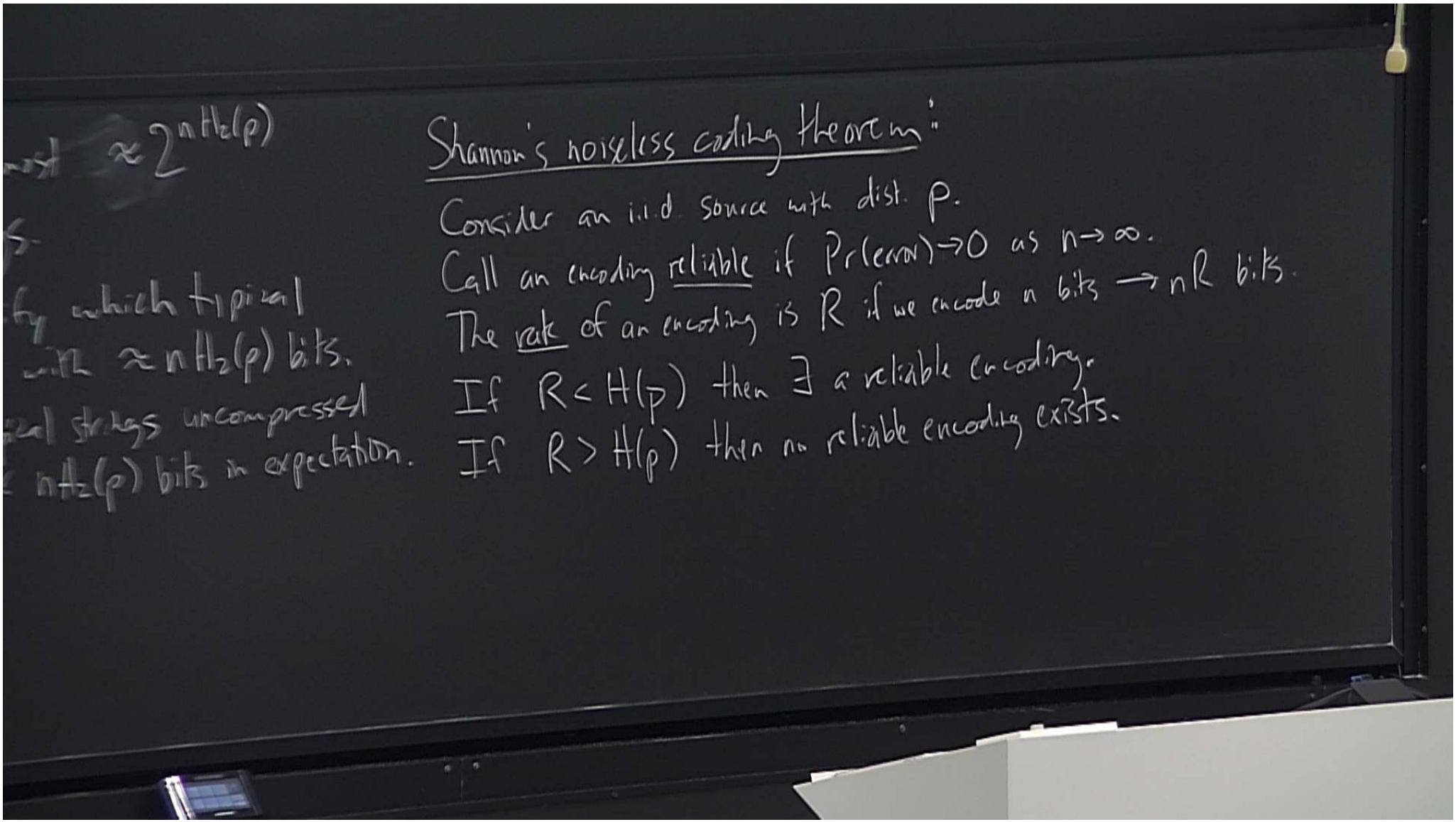
typical if it has $\approx np$ 0s)

There can be at most $\approx 2^{nH_2(p)}$ typical strings.

So we can specify which typical string we got with $\approx nH_2(p)$ bits.

We can encode atypical strings uncompressed and still use $\approx nH_2(p)$ bits in expectation.


most $\approx 2^{nH_2(p)}$
ngs.
city which typical
with $\approx nH_2(p)$ bits.
ypical strings uncompressed
$\approx nH_2(p)$ bits in expectation.
Shannon's noiseless coding theorem:
Consider an i.i.d source with dist. p.
Call an encoding reliable if $Pr(\text{error}) \to 0$ as $n \to \infty$.
The rate of an encoding is R if we encode n bits $\to nR$ bits.
If $R < H(p)$ then $\exists$ a reliable encoding.
If $R > H(p)$ then no reliable encoding exists.


$\approx 2^{nH_2(p)}$
which typical
with $\approx nH_2(p)$ bits.
strings uncompressed
$nH_2(p)$ bits in expectation.
Shannon's noiseless coding theorem:
Consider an i.i.d. source with dist. p.
Call an encoding reliable if $Pr(\text{error}) \to 0$ as $n \to \infty$.
The rate of an encoding is R if we encode n bits $\to nR$ bits.
If $R < H(p)$ then $\exists$ a reliable encoding.
If $R > H(p)$ then no reliable encoding exists.