Strong, weak and pretty strong: Converses for quantum channel capacities

Andreas Winter (ICREA & UAB Barcelona)

1st (A)QIP: Aarhus 1998

1st (A)QIP: Aarhus 1998

64 participants: whom do you know?

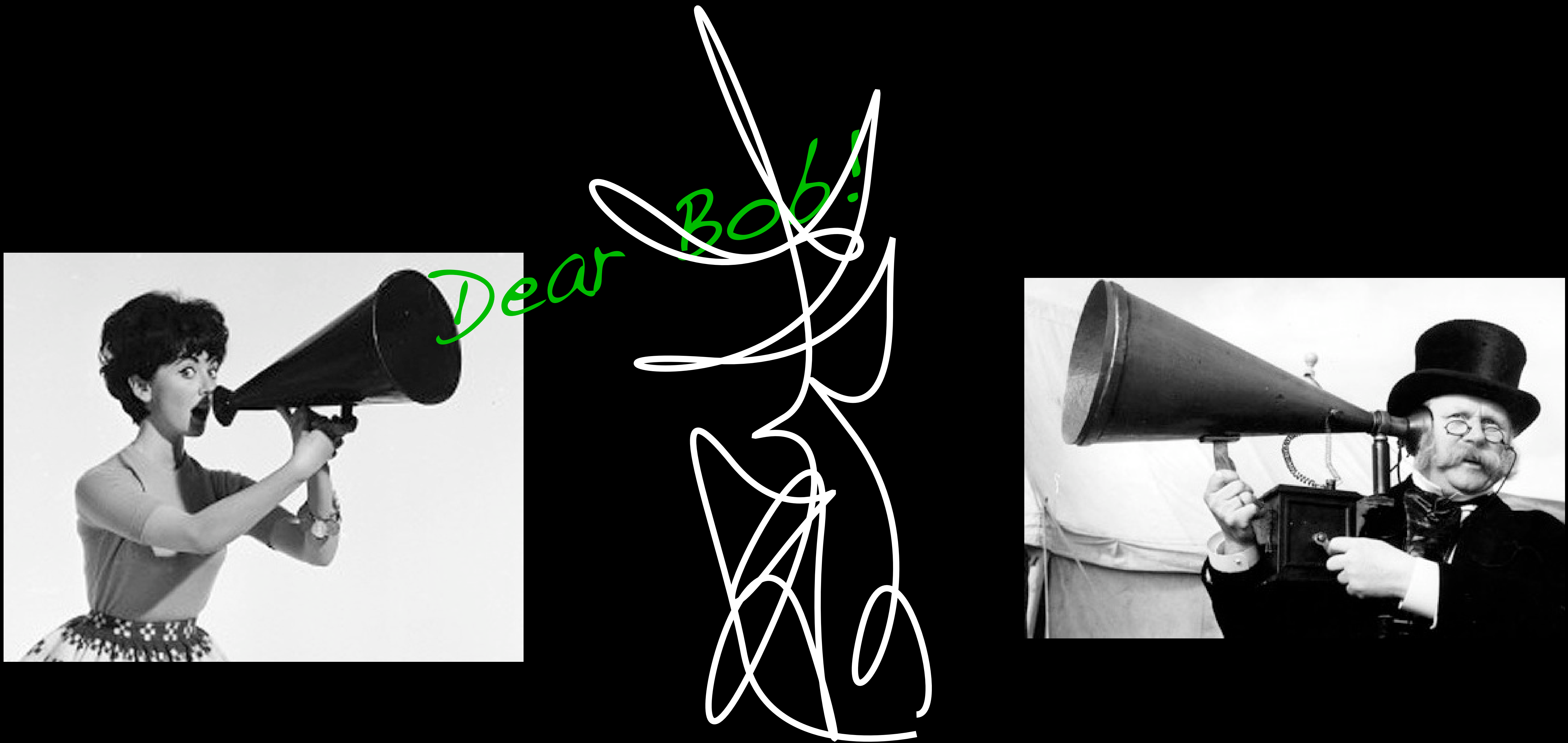
(circled: present @ QIP2014)

1. Communication

Dear Bob!

Shannon (1948): Fundamental problem is that of reproducing at one point a message selected at another point.

1. Communication

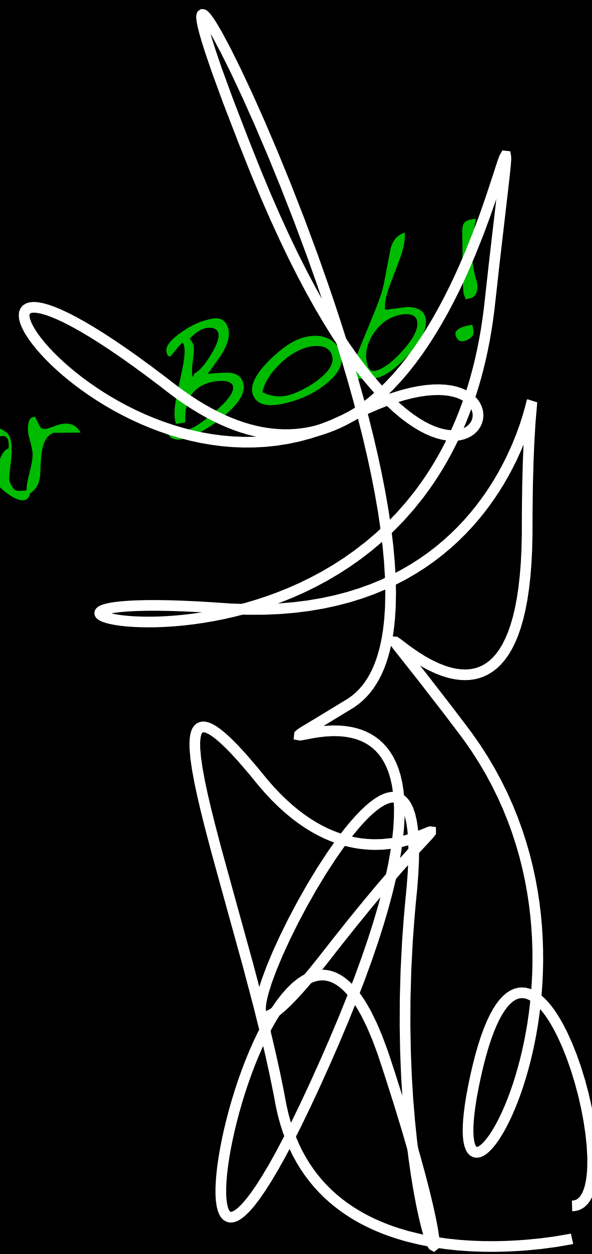


(noise)

1. Communication

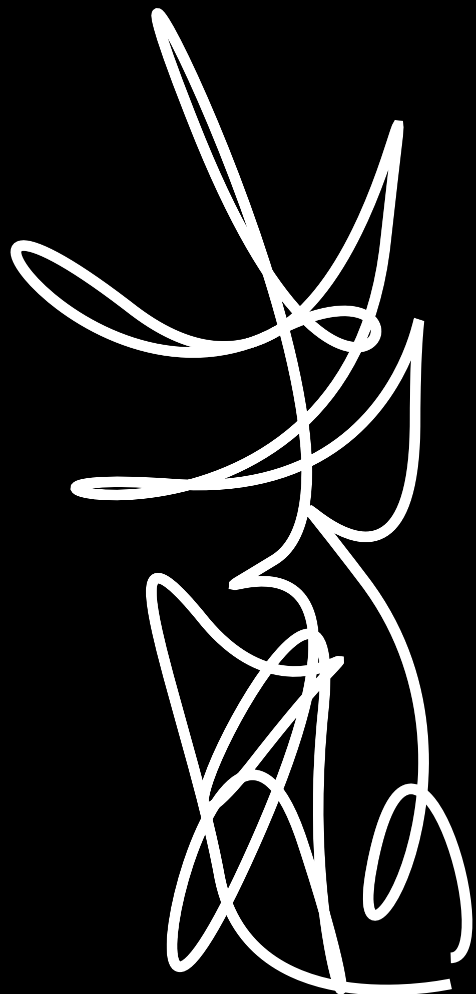
Deep throat?

Dear Bob!



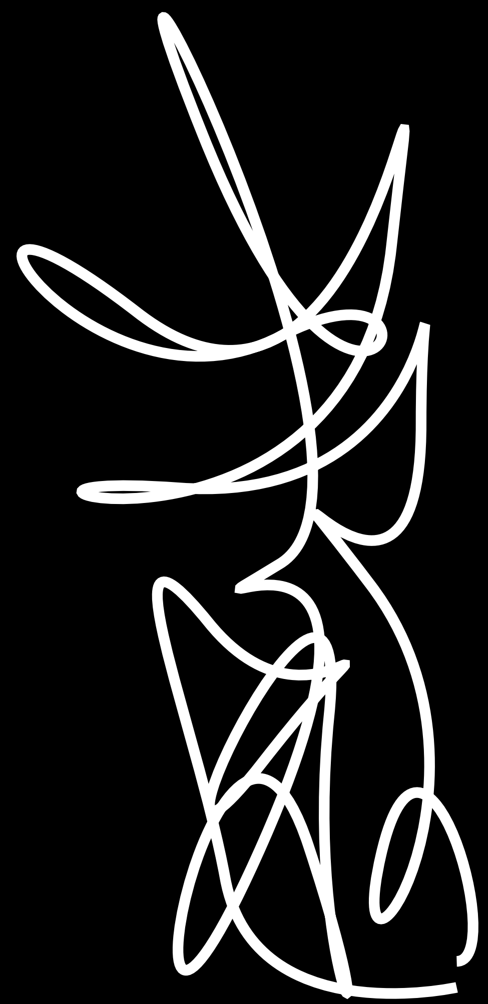
(noise)

1. Channels & capacity

Noise  modelled as "channel":

$$\xrightarrow{\;A\;} \boxed{\mathcal{N}} \xrightarrow{\;B\;}$$

1. Channels & capacity

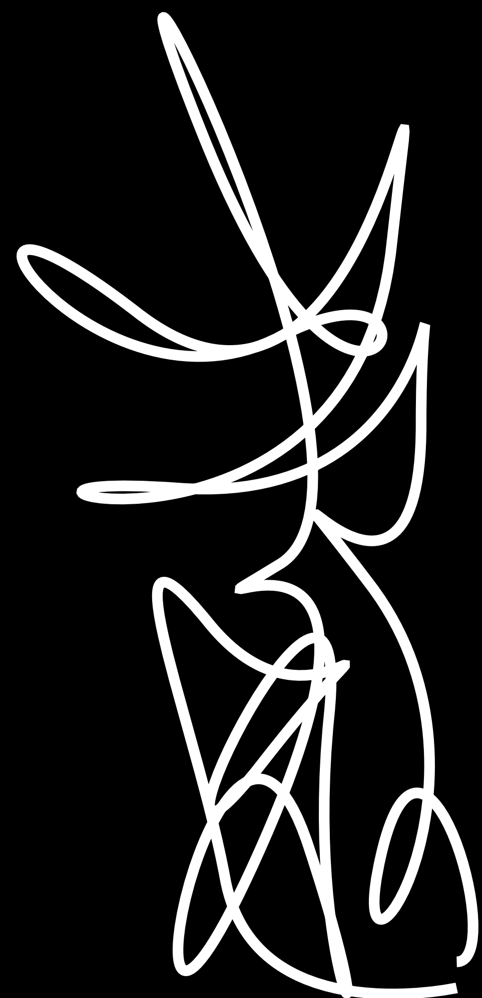
Noise  modelled as "channel":

$$A \longrightarrow \boxed{\mathcal{N}} \longrightarrow B$$

Channel = cptp map $\mathcal{N} : L(A) \longrightarrow L(B)$,

where A, B are finite-dim. Hilbert spaces.

1. Channels & capacity

Noise  modelled as "channel":

$$\xrightarrow{A} \boxed{\mathcal{N}} \xrightarrow{B}$$

Channel = cptp map $\mathcal{N} : L(A) \longrightarrow L(B)$,

where A, B are finite-dim. Hilbert spaces.

Kraus representation, you know...

1. Channels & capacity

Channel = cptp map $N: L(A) \longrightarrow L(B)$,

where A, B are finite-dim. Hilbert spaces.

Kraus representation, you know...

Stinespring: $N(\rho) = Tr_E\, V \rho V^\dagger$,

with an isometry $V: A \hookrightarrow B \otimes E$.

1. Channels & capacity

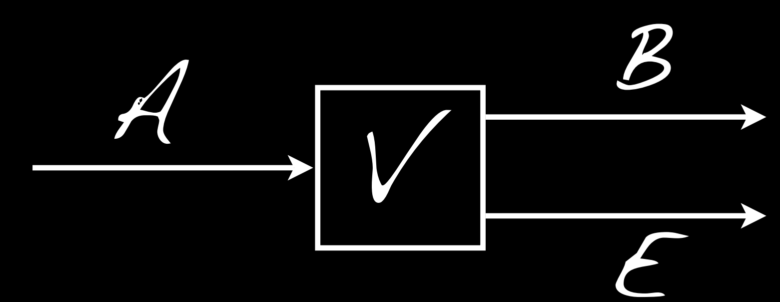
Channel = cptp map $N : L(A) \longrightarrow L(B)$,

where A, B are finite-dim. Hilbert spaces.

Kraus representation, you know...

Stinespring: $N(\rho) = Tr_E \, V \rho V^\dagger$,

with an isometry $V : A \hookrightarrow B \otimes E$.



Complementary channel:

$$\widehat{N}(\rho) = Tr_B \, V \rho V^\dagger.$$

1. Channels & capacity

Ex: 1) Noiseless channel = identity id_A.

2) Constant channel $K(\rho) = \omega_0$.

3) Depolarizing channels

4) Amplitude damping channels

5) Phase damping channels

6) Erasure channel $\mathcal{E}_q(\rho) = (1-q)\rho \oplus q|*\rangle\langle *|$

1. Channels & capacity

Ex: 1) Noiseless channel = identity id_A.

2) Constant channel $K(\rho) = \omega_0$.

3) Depolarizing channels

4) Amplitude damping channels

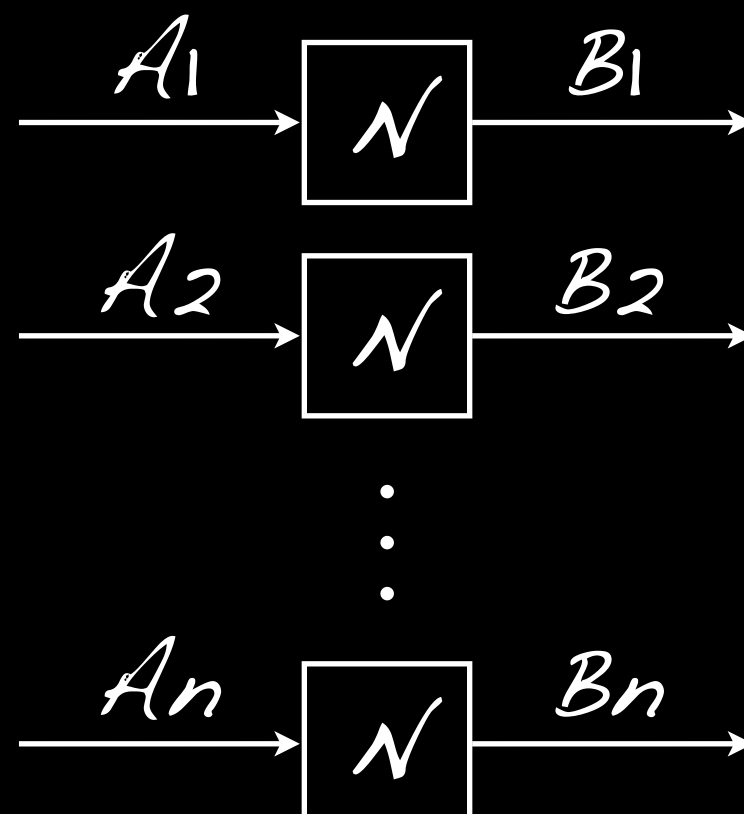
5) Phase damping channels

6) Erasure channel $\mathcal{E}_q(\rho) = (1-q)\rho \oplus q|*\rangle\langle *|$

(Later in this talk, we'll look at some special classes: degradable, Hadamard, entanglement-breaking, ...)

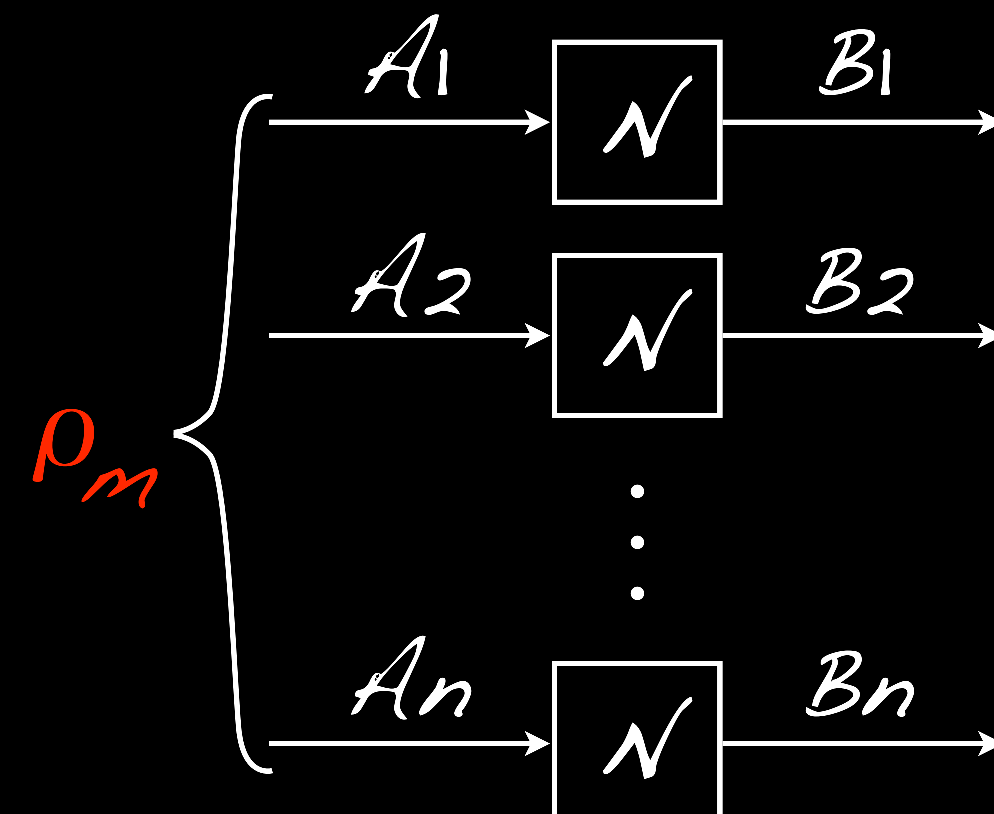
Classical capacity $C(\mathcal{N})$:= maximum cbit rate $\frac{k}{n}$ for asymptotically error-free transmission over $\mathcal{N}^{\otimes n}$.

Classical capacity $C(\mathcal{N}) :=$ maximum cbit rate $\frac{k}{n}$ for asymptotically error-free transmission over $\mathcal{N}^{\otimes n}$.



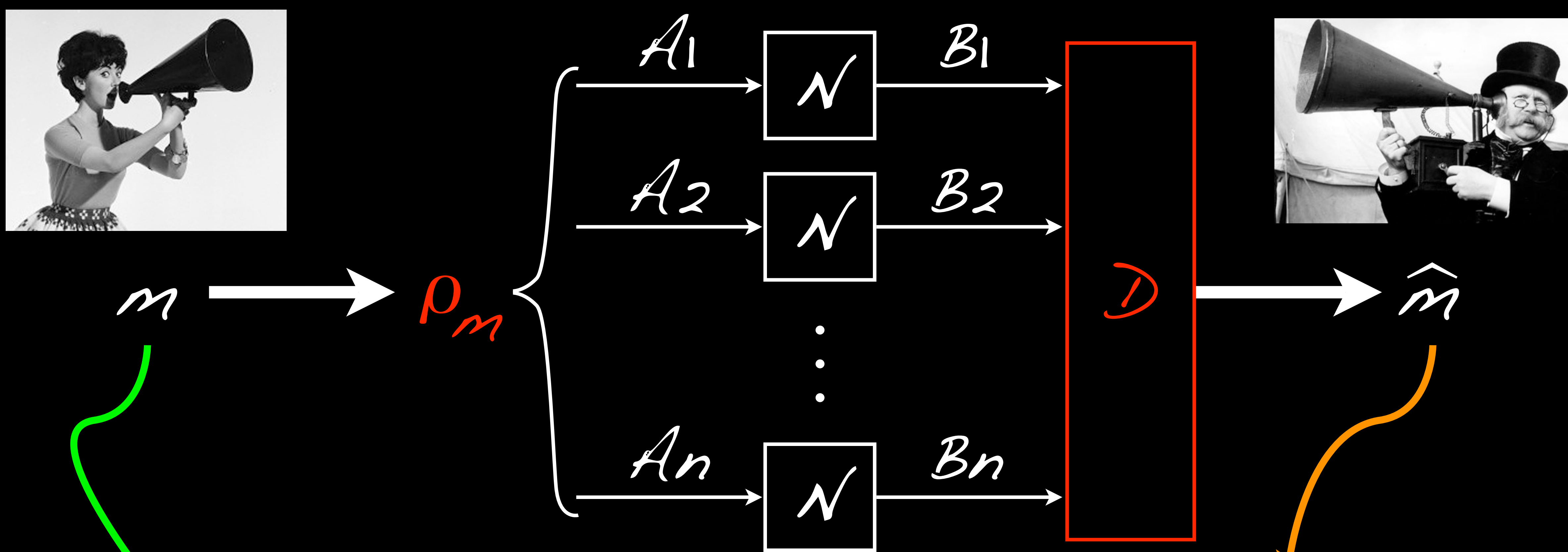
Classical capacity $C(\mathcal{N})$:= maximum cbit rate $\frac{k}{n}$ for asymptotically error-free transmission over $\mathcal{N}^{\otimes n}$:

m



message of
$k=k(n,\varepsilon)$ bits

Classical capacity $C(\mathcal{N})$:= maximum cbit rate $\frac{k}{n}$ for asymptotically error-free transmission over $\mathcal{N}^{\otimes n}$.



message of $k = k(n, \varepsilon)$ bits

$= m$ with prob $\geq 1 - \varepsilon$

...$C(N)$ is not the only capacity:

…$C(\mathcal{N})$ is not the only capacity:

Private capacity $P(\mathcal{N}) :=$ maximum cbit rate as before, in addition asymptotically secret: environment almost independent.

…$C(\mathcal{N})$ is not the only capacity:

Private capacity $P(\mathcal{N}) :=$ maximum cbit rate as before, in addition asymptotically secret: environment almost independent.

Quantum capacity $Q(\mathcal{N}) :=$ maximum qubit rate for asymptotically faithful transmission.

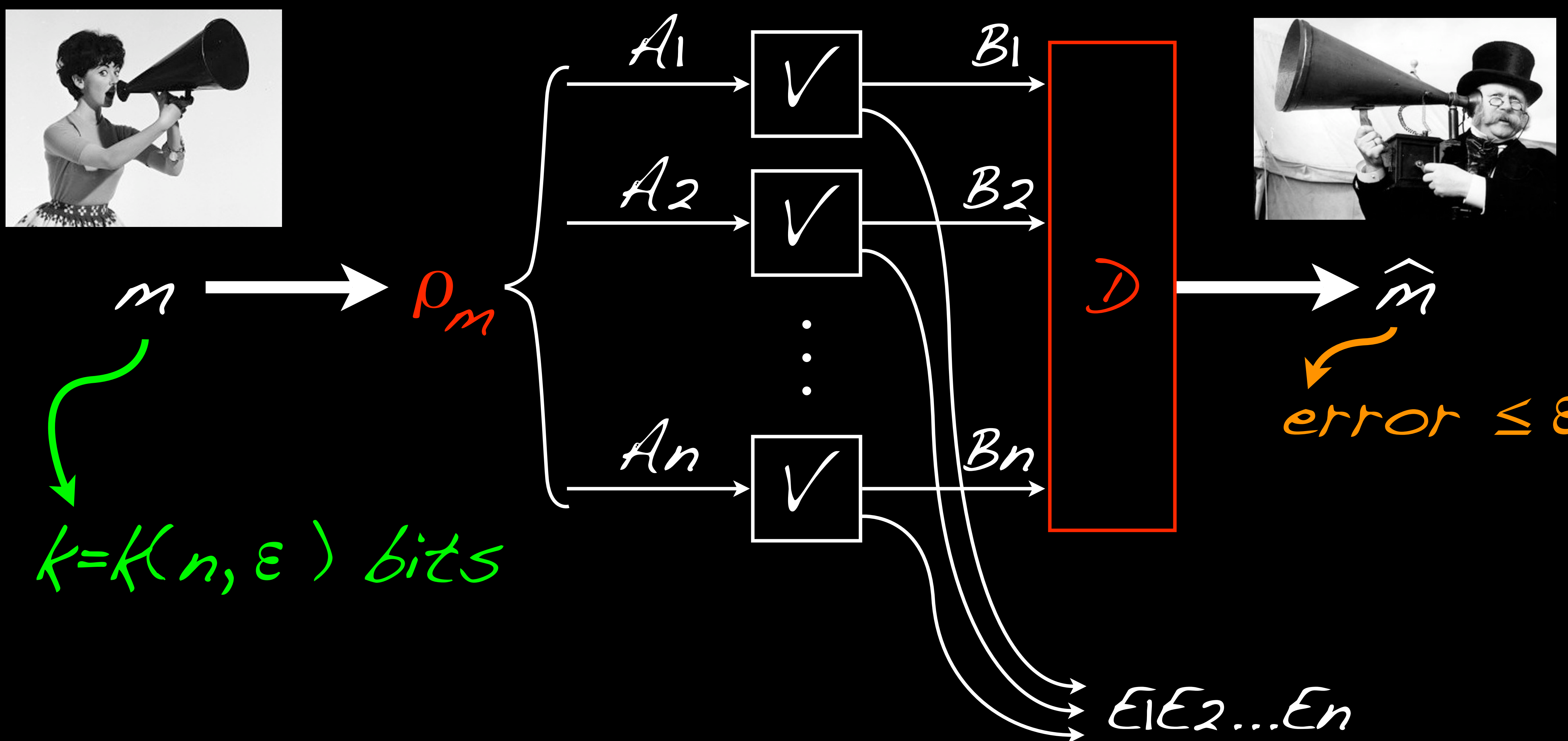
…$C(\mathcal{N})$ is not the only capacity:

Private capacity $P(\mathcal{N}) :=$ maximum cbit rate as before, in addition asymptotically secret: environment almost independent.

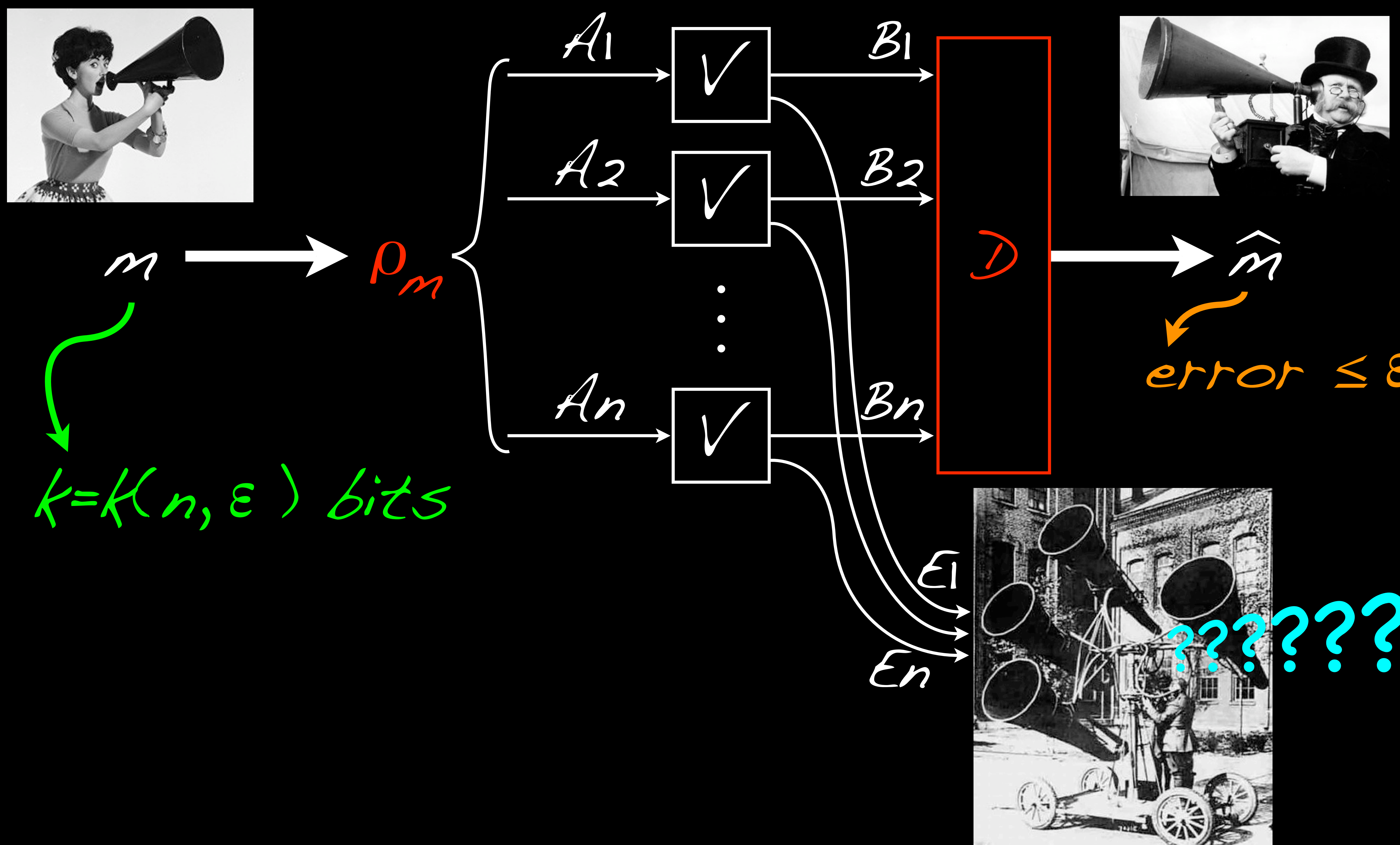
Quantum capacity $Q(\mathcal{N}) :=$ maximum qubit rate for asymptotically faithful transmission.

…and a veritable "zoo" when allowing other free resources: $\mathcal{E}, \leftarrow, \rightarrow, \leftrightarrow, \ldots$

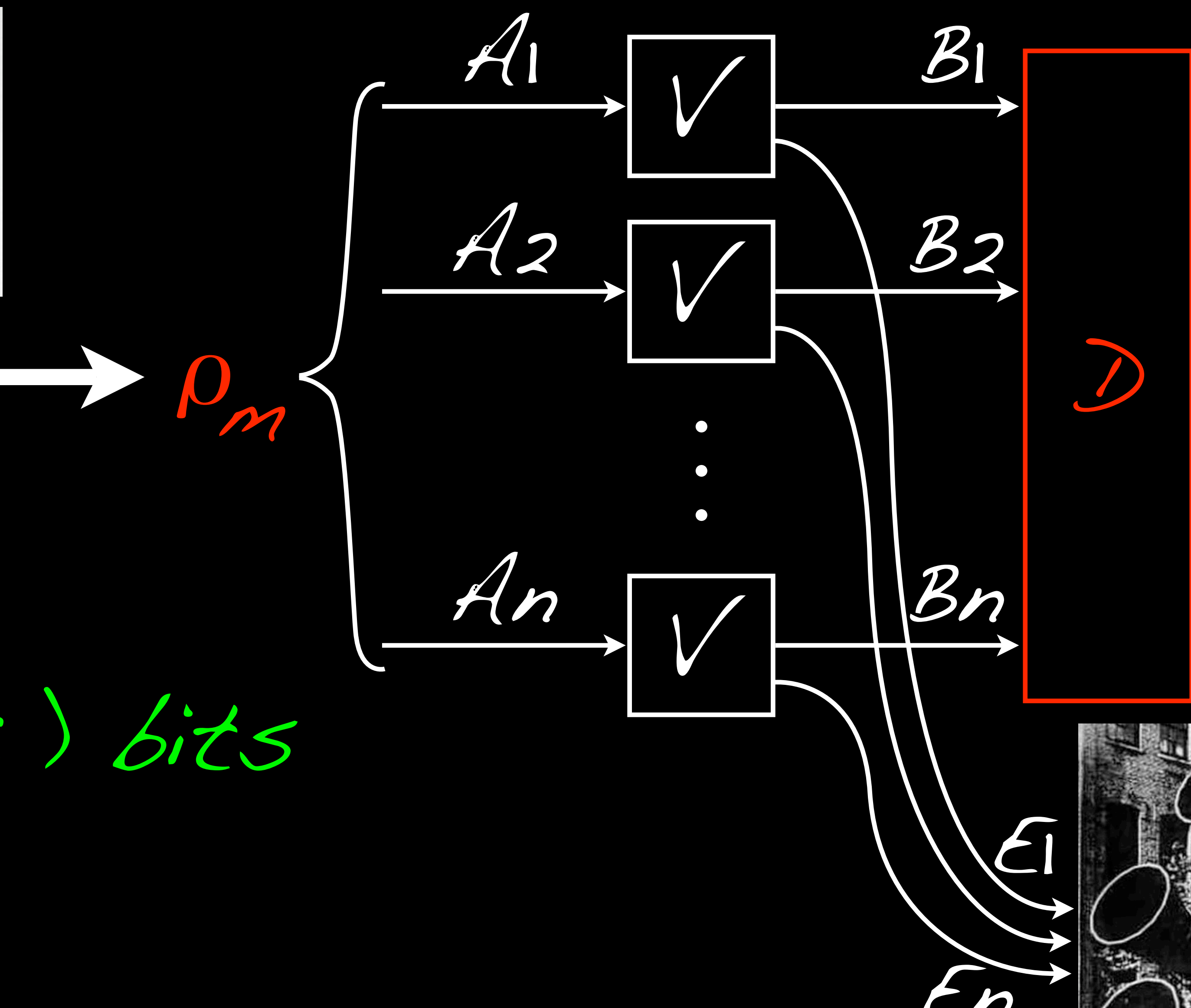
Private capacity $P(N)$:= maximum cbit rate $\frac{k}{n}$ for asymptotically error-free and secret transmission over $N^{\otimes n}$.



Private capacity $P(\mathcal{N})$:= maximum cbit rate $\frac{k}{n}$ for asymptotically error-free and secret transmission over $\mathcal{N}^{\otimes n}$.

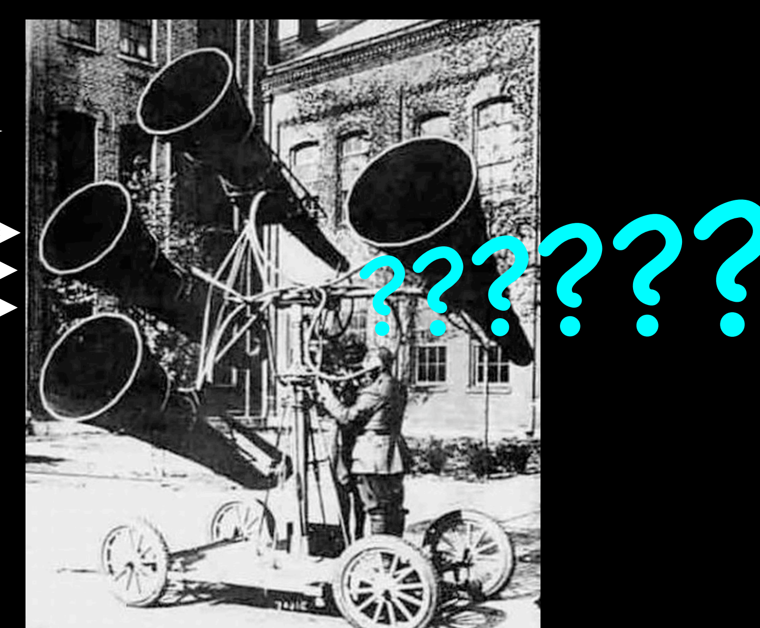


Private capacity $P(\mathcal{N})$:= maximum cbit rate $\frac{k}{n}$ for asymptotically error-free and secret transmission over $\mathcal{N}^{\otimes n}$.



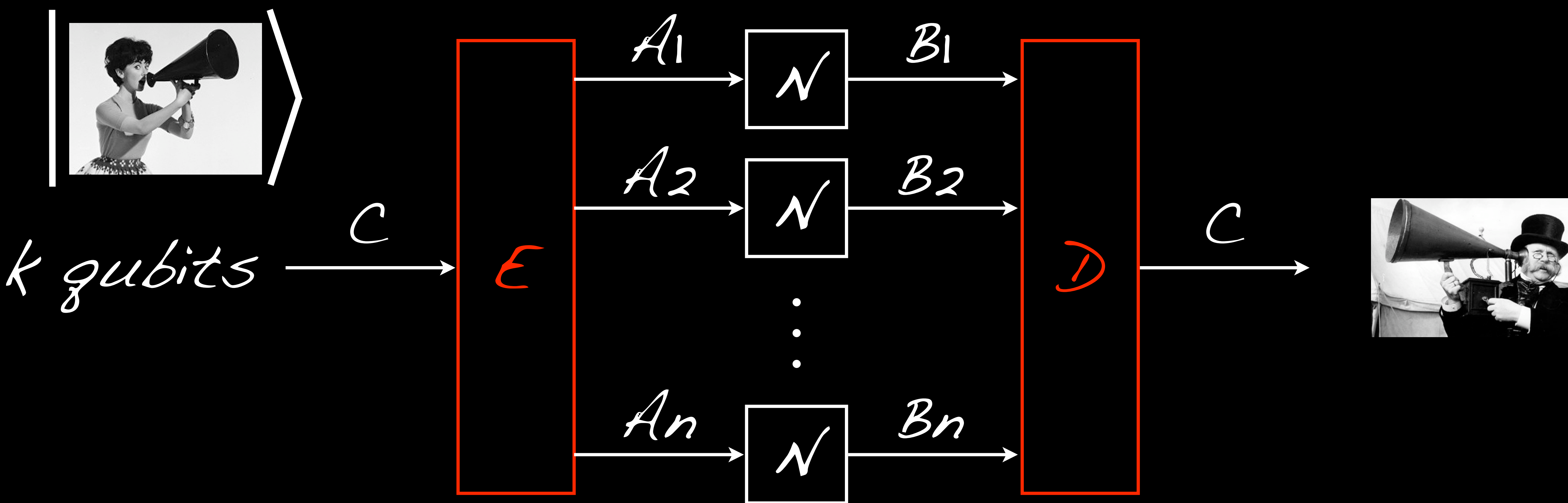
$k = k(n, \varepsilon)$ bits

error $\leq \varepsilon$

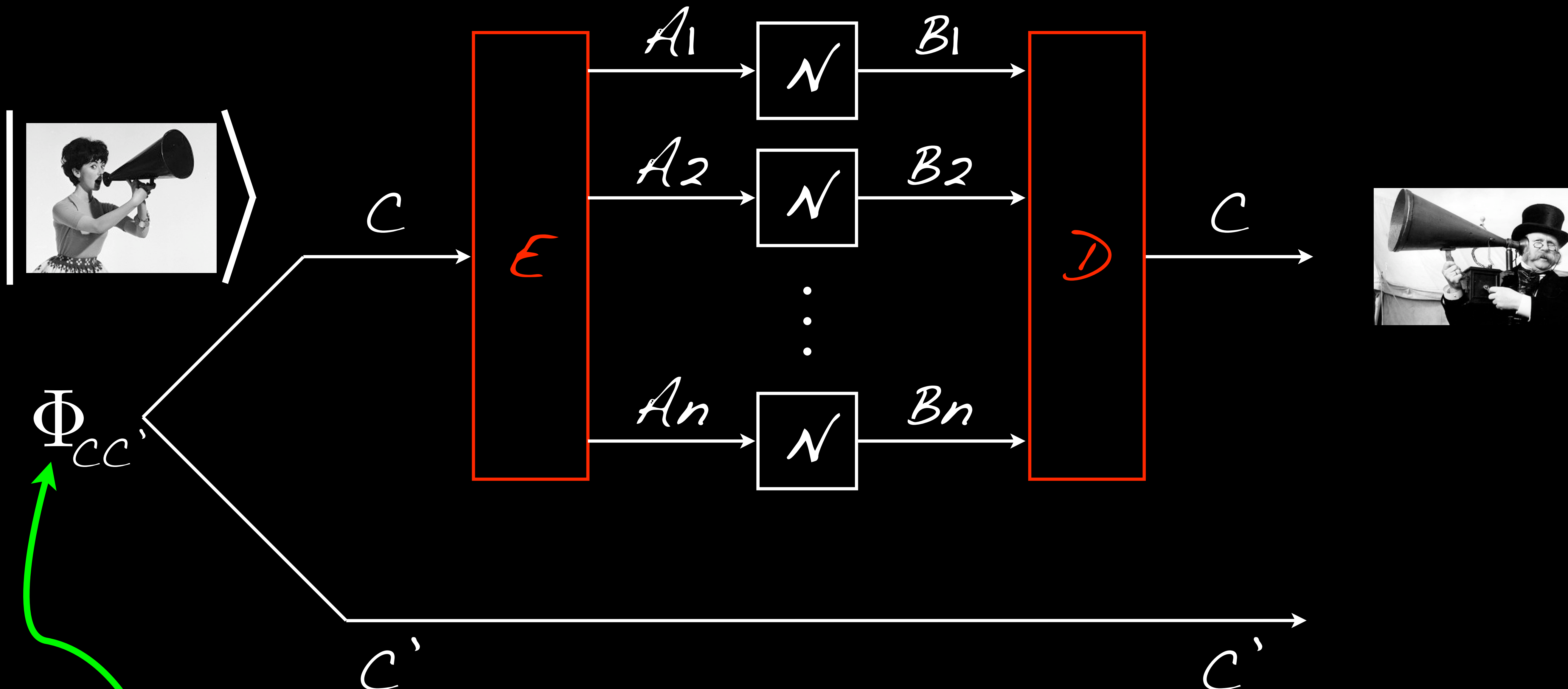


Secret: $\| \hat{\mathcal{N}}^{\otimes n}(\rho_m) - \omega_0 \|_1 \leq \delta$

Quantum capacity $Q(\mathcal{N})$ requires en- and decoding by cptp maps $\mathcal{E}$, $\mathcal{D}$:



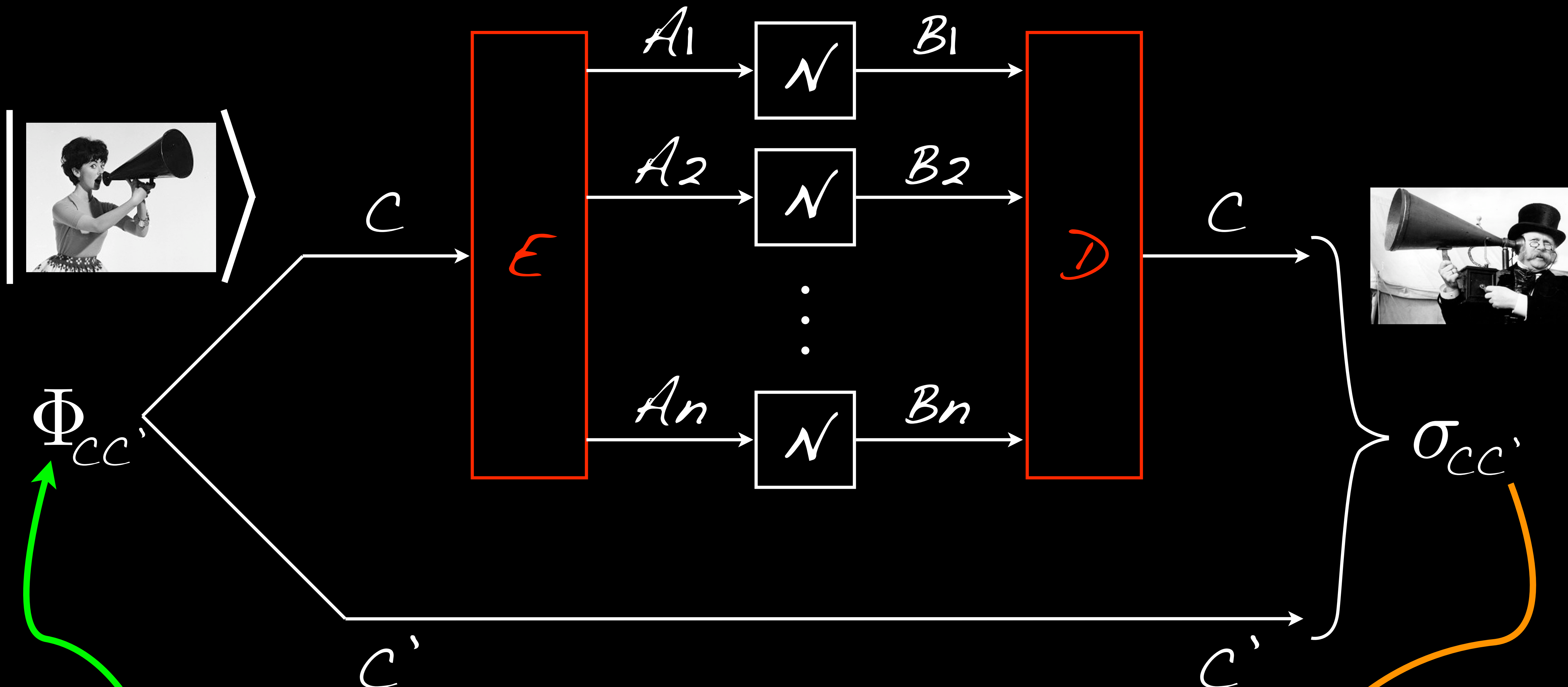
Quantum capacity $Q(\mathcal{N})$ requires en- and decoding by cptp maps $\mathcal{E}$, $\mathcal{D}$:



$k = k(n, \varepsilon)$ EPR pairs

(rate still $\frac{k}{n}$)

Quantum capacity $Q(\mathcal{N})$ requires en- and decoding by cptp maps $\mathcal{E}$, $\mathcal{D}$:



$k = k(n, \varepsilon)$ EPR pairs
(rate still $\frac{k}{n}$)

Approximates input:
$P(\Phi, \sigma) \leq \varepsilon$.

Digression on fidelity:

$$F(\rho,\sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$$

$$= \max |\langle\psi|\varphi\rangle| \text{ s.t.}$$

$|\psi\rangle$ purifies ρ, $|\varphi\rangle$ purifies σ.

$P(\rho,\sigma) := \sqrt{1-F(\rho,\sigma)^2}$ is a metric on states;

...and so is $A(\rho,\sigma) := \arcsin P(\rho,\sigma)$.

Note: Both are equivalent to the trace distance $\|\rho-\sigma\|_1$.

[cf. M. Tomamichel, PhD thesis, arXiv:1203.2142]

Outline

1. Quantum channels and their capacities ✓
2. Entropic capacity formulas; weak converse
3. What is a strong converse?
4. Ideal channel (warm-up); simulation argument
5. Rényi divergence paradigm: classical capacity
6. Min-entropies: "pretty strong" converse
7. End credits

2. Capacity formulas and weak converse

Thm (Holevo and Schumacher/Westmoreland, 1973 and 1996/7):

$$C(N) = \lim_{n\to\infty} \frac{1}{n} \chi(N^{\otimes n}), \text{ with}$$

$\chi(N) = \max I(X:B)$ wrt. $\{p_x, \rho_x\}$ and

$$\rho_{XB} = \sum_x p_x |x\rangle\langle x| \otimes N(\rho_x).$$

2. Capacity formulas and weak converse

Thm (Holevo and Schumacher/Westmoreland, 1973 and 1996/7):

$C(\mathcal{N}) = \lim_{n \to \infty} \frac{1}{n} \chi(\mathcal{N}^{\otimes n})$, with

$\chi(\mathcal{N}) = \max I(X:B)$ wrt. $\{p_x, \rho_x\}$ and

$$\rho_{XB} = \sum_x p_x |x\rangle\langle x| \otimes \mathcal{N}(\rho_x).$$

Holevo information $S(\rho_B) - \sum_x p_x S(\mathcal{N}(\rho_x))$

Von Neumann entropy: $S(\rho) = -\mathrm{Tr}\, \rho \log \rho$

Unfortunately,

$\chi(\mathcal{N}) = \max I(X:B)$ wrt. $\{p_x, \rho_x\}$ and

$$\rho_{XB} = \sum_x p_x |x\rangle\langle x| \otimes \mathcal{N}(\rho_x)$$

is not additive in general [Hastings, Nat. Phys 2009], hence $C(\mathcal{N}) > \chi(\mathcal{N})$ possible.

Unfortunately,

$\chi(\mathcal{N}) = \max I(X:B)$ wrt. $\{p_x, \rho_x\}$ and

$$\rho_{XB} = \sum_x p_x |x\rangle\langle x| \otimes \mathcal{N}(\rho_x)$$

is not additive in general [Hastings, Nat. Phys 2009], hence $C(\mathcal{N}) > \chi(\mathcal{N})$ possible.

However, for some classes of channels it is, and we know the classical capacity $C(\mathcal{N})$ as $\chi(\mathcal{N})$.

Interestingly, the upper bound ("converse") was proved first.

Interestingly, the upper bound ("converse") was proved first. In fact, Holevo [Probl. Inf. Transm. (1973), and other work in 1970's] showed that transmitting k bits over n uses of N with error ε,

$$k(1-\varepsilon) \leq 1 + \chi(N^{\otimes n}) \leq 1 + n\,C(N).$$

Interestingly, the upper bound ("converse") was proved first. In fact, Holevo [Probl. Inf. Transm. (1973), and other work in 1970's] showed that transmitting k bits over n uses of N with error ε,

$$k(1-\varepsilon) \leq 1 + \chi(N^{\otimes n}) \leq 1 + n\,C(N).$$

Uses only strong subadditivity (SSA) and continuity of von Neumann entropy (Fannes inequality): $S(\rho) = -\mathrm{Tr}\,\rho \log \rho$.

Interestingly, the upper bound ("converse") was proved first. In fact, Holevo [Probl. Inf. Transm. (1973), and other work in 1970's] showed that transmitting k bits over n uses of N with error ε,

$$k(1-\varepsilon) \leq 1 + \chi(N^{\otimes n}) \leq 1 + n\, C(N).$$

$$\boxed{\frac{k}{n} \lesssim (1+\varepsilon)\, C(N)}$$

Interestingly, the upper bound ("converse") was proved first. In fact, Holevo [Probl. Inf. Transm. (1973), and other work in 1970's] showed that transmitting k bits over n uses of N with error ε,

$$k(1-\varepsilon) \leq 1 + \chi(N^{\otimes n}) \leq 1 + n\, C(N).$$

$$\boxed{\frac{k}{n} \lesssim (1+\varepsilon)\, C(N)}$$

"weak converse"

Interestingly, the upper bound ("converse") was proved first. In fact, Holevo [Probl. Inf. Transm. (1973), and other work in 1970's] showed that transmitting k bits over n uses of N with error ε,

$$k(1-\varepsilon) \leq 1 + \chi(N^{\otimes n}) \leq 1 + n\, C(N).$$

$$\boxed{\frac{k}{n} \lesssim (1+\varepsilon)\, C(N)}$$

"weak converse"

...is the implied tradeoff real?

Analogous formulas for $P(N)$ and $Q(N)$:

Thm (Devetak and Cai/Yeung/AW, 2003):

$$P(N) = \lim_{n\to\infty} \frac{1}{n} P^{(1)}(N^{\otimes n}), \text{ with}$$

$$P^{(1)}(N) = \max I(X:B) - I(X:E) \text{ wrt. } \{p_x, \rho_x\}$$

Analogous formulas for $P(N)$ and $Q(N)$:

Thm (Devetak and Cai/Yeung/AW, 2003):

$$P(N) = \lim_{n\to\infty} \frac{1}{n} P^{(1)}(N^{\otimes n}), \text{ with}$$

$$P^{(1)}(N) = \max I(X:B) - I(X:E) \text{ wrt. } \{p_x, \rho_x\}$$

Thm (Schumacher and Lloyd-Shor-Devetak, 1996-2003):

$$Q(N) = \lim_{n\to\infty} \frac{1}{n} Q^{(1)}(N^{\otimes n}), \text{ with}$$

coherent information

$$Q^{(1)}(N) = \max I(A\rangle B)$$

$$= \max S(N(\rho)) - S(\hat{N}(\rho)) \text{ wrt. } \rho$$

Thm (Devetak and Cai/Yeung/AW, 2003):

$$P(N) = \lim_{n\to\infty} \frac{1}{n} P^{(1)}(N^{\otimes n})$$

Thm (Schumacher and Lloyd-Shor-Devetak, 1996-2003):

$$Q(N) = \lim_{n\to\infty} \frac{1}{n} Q^{(1)}(N^{\otimes n})$$

Thm (Devetak and Cai/Yeung/AW, 2003):

$$P(N) = \lim_{n \to \infty} \frac{1}{n} P^{(1)}(N^{\otimes n})$$

Thm (Schumacher and Lloyd-Shor-Devetak, 1996-2003):

$$Q(N) = \lim_{n \to \infty} \frac{1}{n} Q^{(1)}(N^{\otimes n})$$

Have analogous weak converses for $P(N)$ and $Q(N)$, and for much every other capacity we know how to characterize.

Thm (Devetak and Cai/Yeung/AW, 2003):

$$P(N) = \lim_{n \to \infty} \frac{1}{n} P^{(1)}(N^{\otimes n})$$

Thm (Schumacher and Lloyd-Shor-Devetak, 1996-2003):

$$Q(N) = \lim_{n \to \infty} \frac{1}{n} Q^{(1)}(N^{\otimes n})$$

Have analogous weak converses for $P(N)$ and $Q(N)$, and for much every other capacity we know how to characterize.

(Btw: also additivity issue with these!)

Thm (Devetak and Cai/Yeung/AW, 2003):

$$P(N) = \lim_{n \to \infty} \frac{1}{n} P^{(1)}(N^{\otimes n})$$

Thm (Schumacher and Lloyd-Shor-Devetak, 1996-2003):

$$Q(N) = \lim_{n \to \infty} \frac{1}{n} Q^{(1)}(N^{\otimes n})$$

Important to know: For all these capacities, at rates below, the error goes to zero exponentially, always!

Thm (Devetak and Cai/Yeung/AW, 2003):

$$P(N) = \lim_{n \to \infty} \frac{1}{n} P^{(1)}(N^{\otimes n})$$

Thm (Schumacher and Lloyd-Shor-Devetak, 1996-2003):

$$Q(N) = \lim_{n \to \infty} \frac{1}{n} Q^{(1)}(N^{\otimes n})$$

Important to know: For all these capacities, at rates below, the error goes to zero exponentially, always!

...so what about rates above capacity?

3. Strong converse?

The strong converse – in the sense of Wolfowitz [Ill. J. Math. 1:591 (1957)] –, is the statement that there is no rate-error trade-off. Viz., for rates R above the capacity, the error converges to 1.

3. Strong converse?

The strong converse - in the sense of Wolfowitz [Ill. J. Math. 1:591 (1957)] -, is the statement that there is no rate-error trade-off. Viz., for rates R above the capacity, the error converges to 1.

By contrapositive: If error < 1, then asymptotically the rate $\frac{k}{n}$ is bounded by the capacity.

Strong converse: If error < 1, then asymptotically the rate $\frac{k}{n}$ is bounded by the capacity.

Strong converse: If error < 1, then asymptotically the rate $\frac{k}{n}$ is bounded by the capacity.

Progress over the years:

- Classical channels [Shannon-Wolfowitz]
-
-

Strong converse: If error < 1, then asymptotically the rate $\frac{k}{n}$ is bounded by the capacity.

Progress over the years:

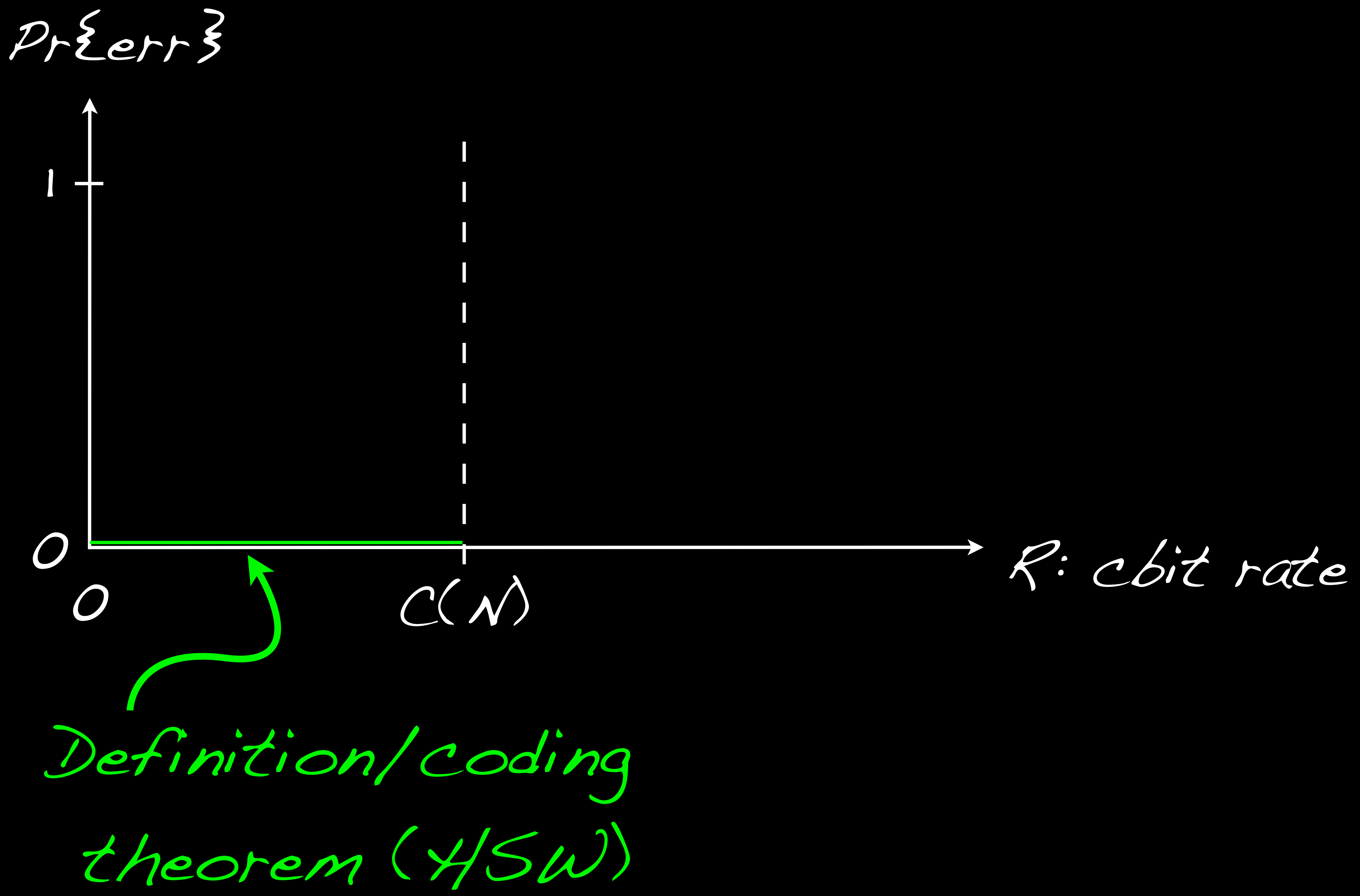
- Classical channels [Shannon-Wolfowitz]
- Classical capacity with product state inputs [Ogawa/Nagaoka; AW, IEEE-IT 45(7), 1999] - i.e., cq-channels
-

Strong converse: If error < 1, then asymptotically the rate $\frac{k}{n}$ is bounded by the capacity.

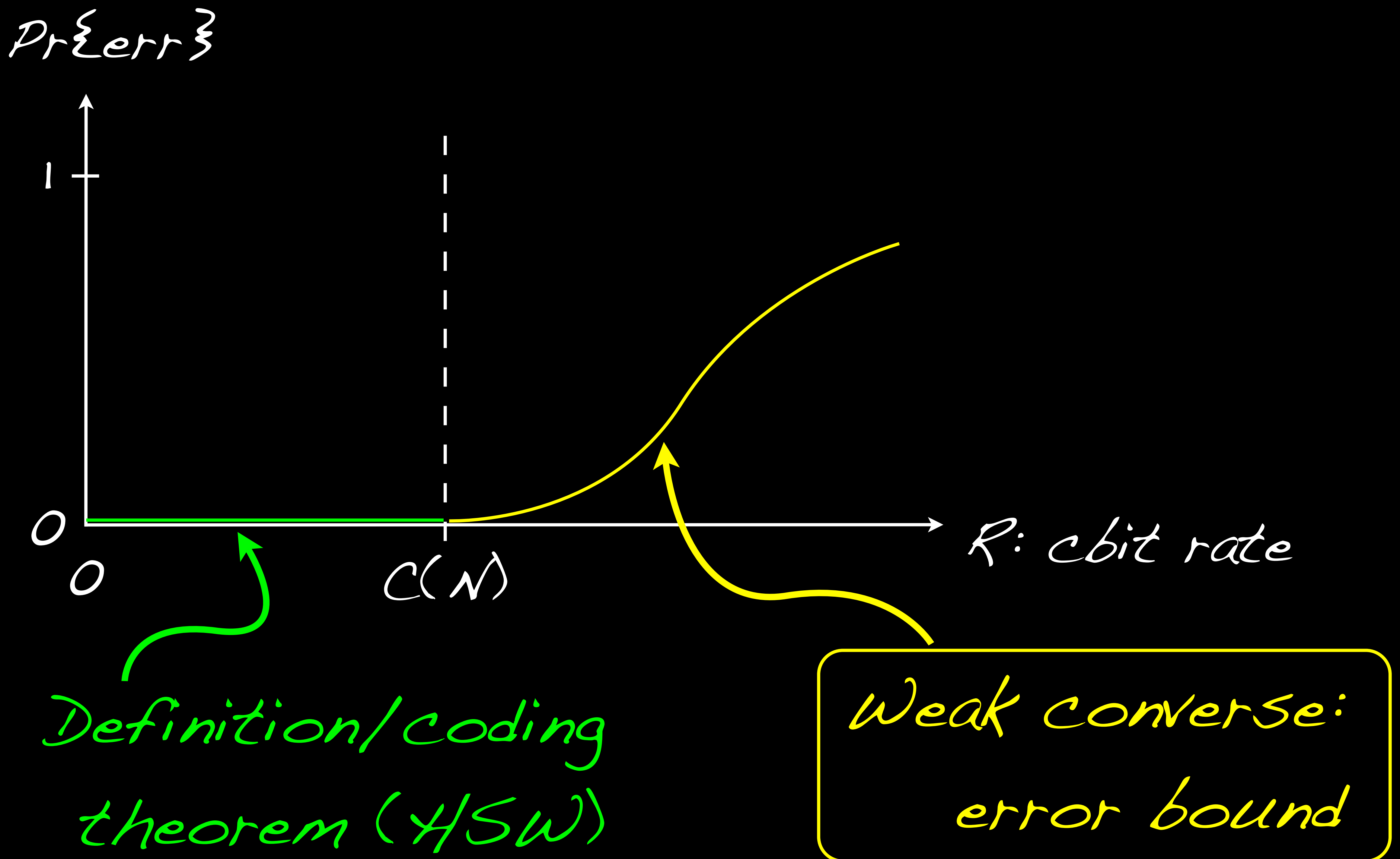
Progress over the years:

- Classical channels [Shannon-Wolfowitz]
- Classical capacity with product state inputs [Ogawa/Nagaoka; AW, IEEE-IT 45(7), 1999] - i.e., cq-channels
- Classical capacity of covariant channels [Koenig/Wehner, PRL 103:070504 (2009)]

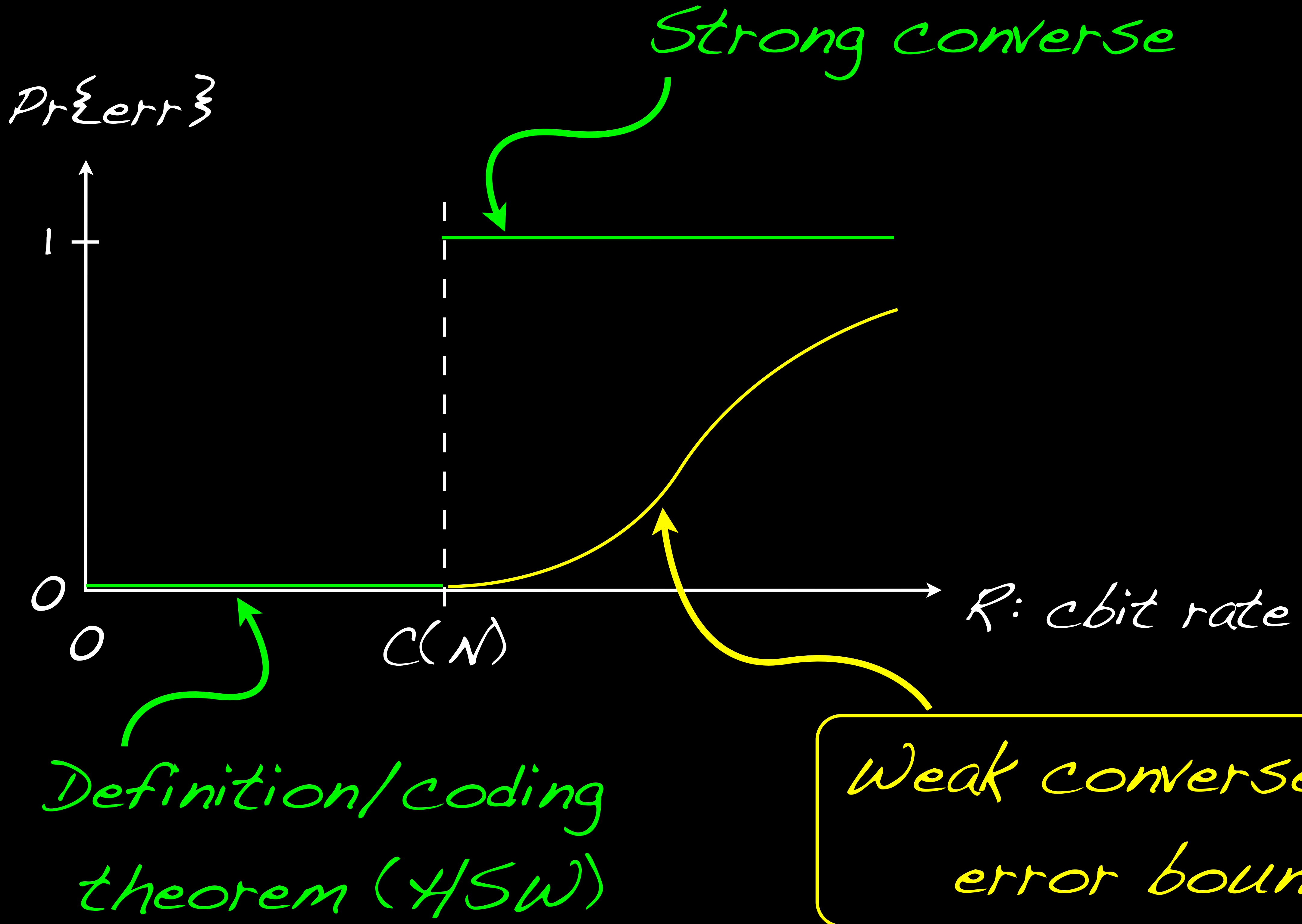
Rate vs asymptotic error:



Rate vs asymptotic error:



Rate vs asymptotic error:



4. Ideal channel

As a warm-up, prove strong converse for the noiseless qubit channel id_2. Note:

quantum code $\Rightarrow$ private code $\Rightarrow$ classical code

4. Ideal channel

As a warm-up, prove strong converse for the noiseless qubit channel id_2. Note:

quantum code $\Rightarrow$ private code $\Rightarrow$ classical code

Hence $Q(N) \leq P(N) \leq C(N)$ in general. Since $Q(\mathrm{id}_2) = P(\mathrm{id}_2) = C(\mathrm{id}_2) = 1$, enough to show it for the classical capacity.

Warm-up: strong converse for the noiseless qubit channel id_2.

Encode M message into id_L via states ρ_m and POVM elements D_m to decode:

[Nayak, Proc. 40th FOCS, pp. 369-376 (1999)]

Warm-up: strong converse for the noiseless qubit channel id_2.

Encode M message into id_L via states ρ_m and POVM elements D_m to decode:

$$1 - \varepsilon \leq \frac{1}{M} \sum_{m=1}^{M} Tr(\rho_m D_m)$$

[Nayak, Proc. 40th FOCS, pp. 369-376 (1999)]

Warm-up: strong converse for the noiseless qubit channel id_2.

Encode M message into id_L via states ρ_m and POVM elements D_m to decode:

$$1-\varepsilon \leq \frac{1}{M}\sum_{m=1}^{M} Tr(\rho_m D_m) \leq \frac{1}{M}\sum_{m=1}^{M} Tr\, D_m$$

[Nayak, Proc. 40th FOCS, pp. 369-376 (1999)]

Warm-up: strong converse for the noiseless qubit channel id_2.

Encode M message into id_L via states ρ_m and POVM elements D_m to decode:

$$1-\varepsilon \leq \frac{1}{M}\sum_{m=1}^{M} \mathrm{Tr}(\rho_m D_m) \leq \frac{1}{M}\sum_{m=1}^{M} \mathrm{Tr}\, D_m = \frac{L}{M}.$$

[Nayak, Proc. 40th FOCS, pp. 369-376 (1999)]

Warm-up: strong converse for the noiseless qubit channel id_2.

Encode M message into id_L via states ρ_m and POVM elements D_m to decode:

$$1-\varepsilon \leq \frac{1}{M}\sum_{m=1}^{M} \mathrm{Tr}(\rho_m D_m) \leq \frac{1}{M}\sum_{m=1}^{M} \mathrm{Tr}\, D_m = \frac{L}{M}.$$

For n uses of the channel and rate $R>1$: $L=2^n$ and $M=2^{nR}$, so $\varepsilon \geq 1-2^{-n(R-1)}$. QED

[Nayak, Proc. 40th FOCS, pp. 369-376 (1999)]

The simulation argument: If you can simulate a channel N by id_2 at rate K, then $C(N) \leq K$ and for rates $R > K$, the error $\varepsilon \geq 1 - 2^{-n(R-K)}$.

In particular: If $K = C(N)$, strong converse holds.

The simulation argument: If you can simulate a channel N by id_2 at rate K, then $C(N) \leq K$ and for rates $R > K$, the error $\varepsilon \geq 1 - 2^{-n(R-K)}$.

In particular: If $K = C(N)$, strong converse holds. Almost only trivial cases, except:

Thm (Wilde/AW, 1308.6732): For pure loss optical channel w/ transmissivity η and maximum mean photon number P, $C = g(\eta P)$, and the strong converse holds.

The simulation argument: If you can simulate a channel N by id_2 at rate K, then $C(N) \leq K$ and for rates $R > K$, the error $\varepsilon \geq 1 - 2^{-n(R-K)}$.

More interesting with free resources, eg.

$C_E(N)$ = ent.-assisted classical capacity

= minimal simulation cost assisted by ent. ("Qu. Reverse Shannon Thm.")

Ie. strong converse holds for C_E.

[Bennett et al., IEEE-IT 48:2637 (2002); Bennett et al. 0912.5537]

[Cf. Berta et al., IEEE-IT 59:6770 (2013) - $P(N)$ bound]

5. Rényi divergences for C

What can we do for $C(N)$? Nothing general it seems...

5. Rényi divergences for C

What can we do for $C(N)$? Nothing general it seems... However, unifying and extending the earlier results of Ogawa/Nagaoka, AW and König/Wehner:

Thm (Wilde/AW/Yang, 1306.1586): If N is entanglement-breaking (EB) or Hadamard (H), then for any code w rate $R > C(N)$, Pr{err} converges to 1, exponentially fast in the number n of channel uses.

5. Rényi divergences for C

Thm (Wilde/AW/Yang, 1306.1586): If N is EB or $\mathcal{H}$, then for any code w rate $R >$ $C(N)$, the error probability converges to 1, exponentially fast in the number n of channel uses:

There exists $t \geq \Omega\left((R-C(N))^2\right)$ s.t.

$$1-P\{err\} \leq \exp(-tn).$$

5. Rényi divergences for C

Thm (Wilde/AW/Yang, 1306.1586): If N is EB or $\mathcal{H}$, then for any code w rate $R >$ $C(N)$, the error probability converges to 1, exponentially fast in the number n of channel uses:

There exists $t \geq \Omega\left((R-C(N))^2\right)$ s.t.

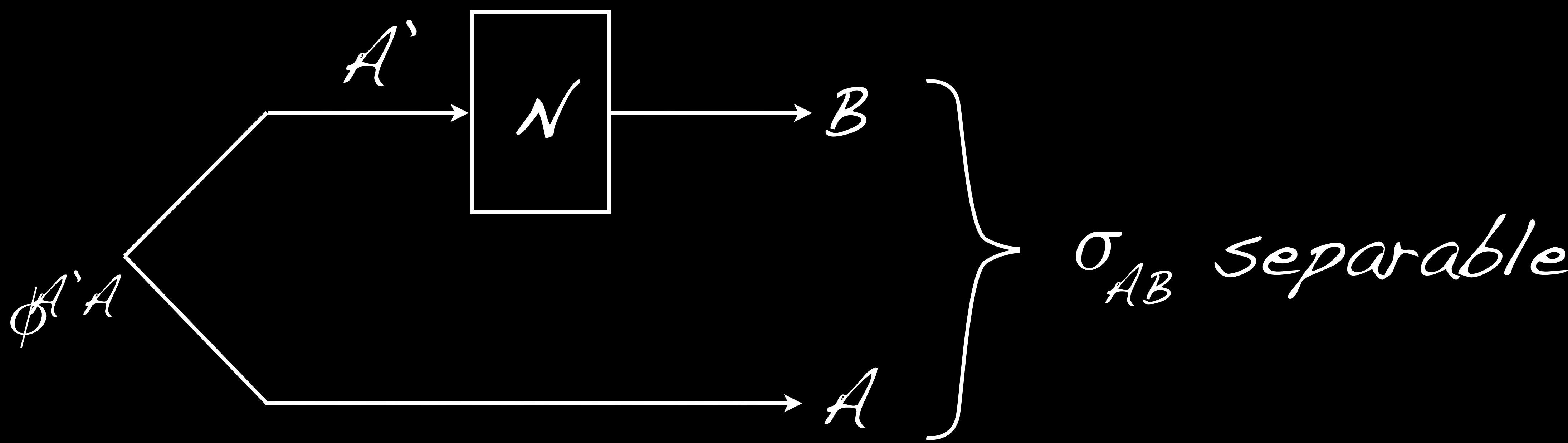
$$1-P\{err\} \leq \exp(-tn).$$

In other words, these channels satisfy the strong converse.

Hold on! I haven't even told you what these "EB" and "H" things are...

Hold on! I haven't even told you what these "EB" and "H" things are...

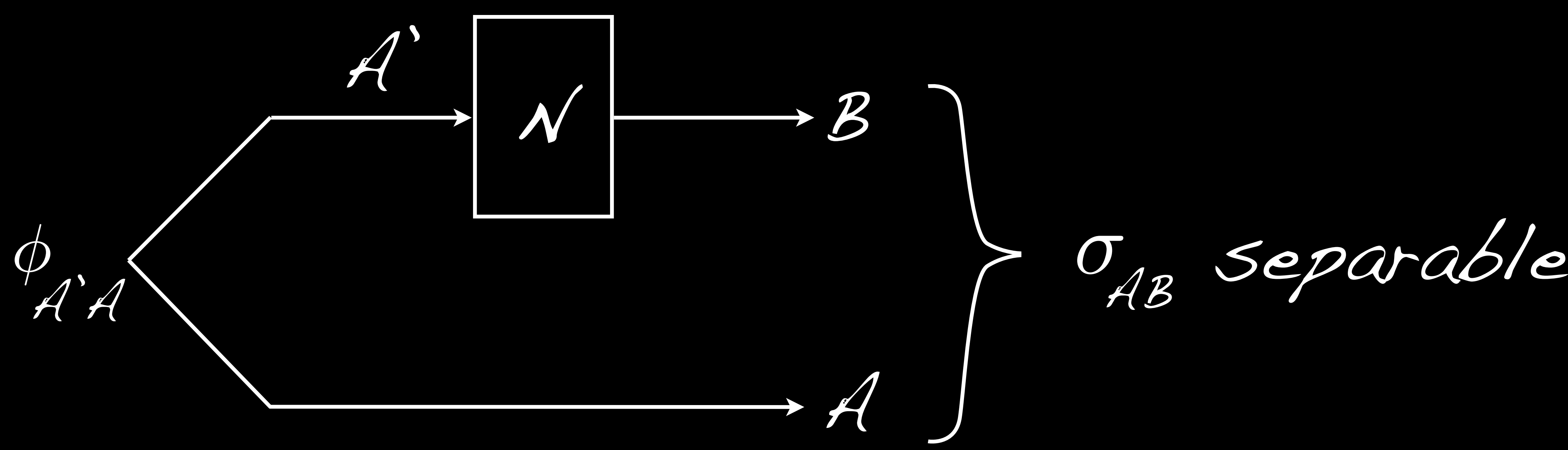
Entanglement-breaking (EB) channels:



Complementary to these:

Hadamard channels ($\mathcal{H}$)

Entanglement-breaking (EB) channels:

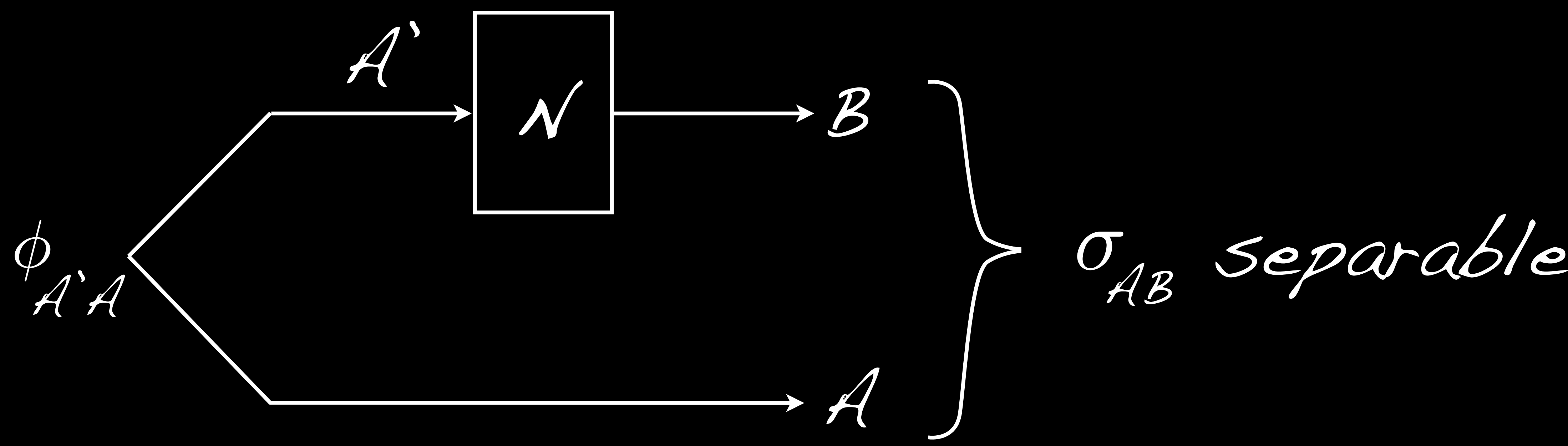


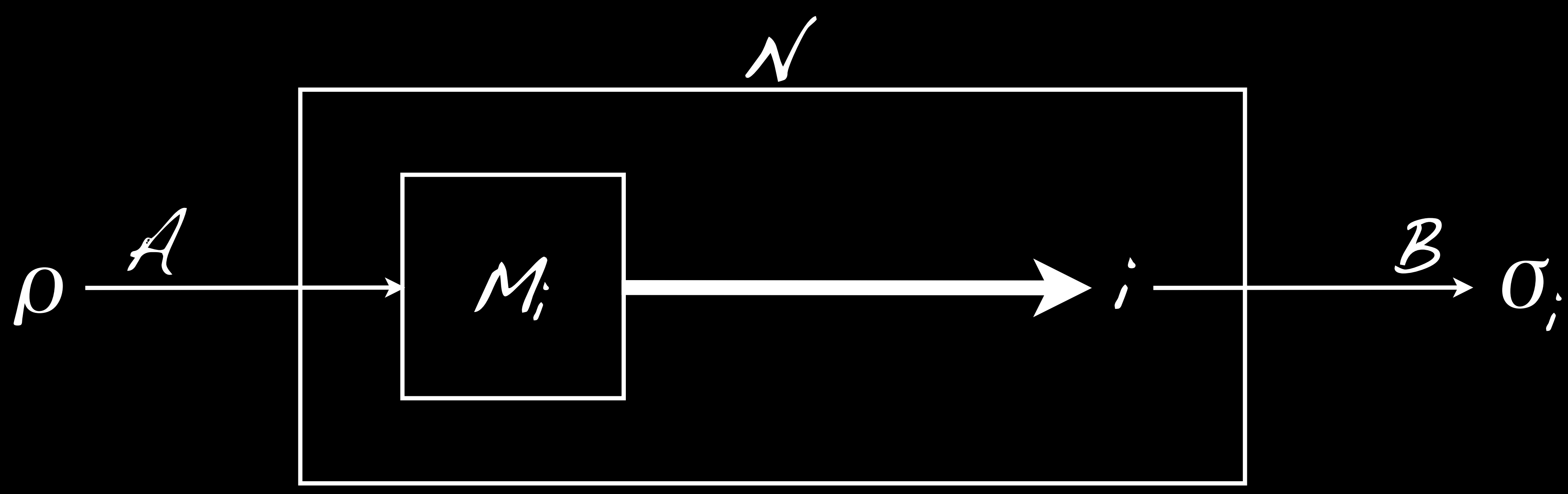
Fact: $\mathcal{N}$ entanglement-breaking iff

$$\mathcal{N}(\rho) = \sum_i Tr(\rho M_i)\sigma_i \quad s.t. \quad \sum_i M_i = \mathbb{1}$$

$$= \sum_j |\beta_j\rangle\langle\alpha_j|\rho|\alpha_j\rangle\langle\beta_j|$$

Entanglement-breaking (EB) channels:





$$\mathcal{N}(\rho) = \sum_i Tr(\rho M_i)\sigma_i \quad s.t. \quad \sum_i M_i = \mathbb{1}$$

$$= \sum_j |\beta_j\rangle\langle\alpha_j|\rho|\alpha_j\rangle\langle\beta_j|$$

Stinespring: $V: |\phi\rangle_A \to \sum_j \langle\alpha_j|\phi\rangle |\beta_j\rangle_B |j\rangle_E$

$$\mathcal{N}(\rho) = \sum_i Tr(\rho M_i)\sigma_i \quad s.t. \quad \sum_i M_i = \mathbb{1}$$

$$= \sum_j |\beta_j\rangle\langle\alpha_j|\rho|\alpha_j\rangle\langle\beta_j|$$

Stinespring: $V: |\phi\rangle_A \rightarrow \sum_j \langle\alpha_j|\phi\rangle |\beta_j\rangle_B |j\rangle_E$

$$\hat{\mathcal{N}}(\rho) = \sum_{jk} |j\rangle\langle k| \langle\alpha_j|\rho|\alpha_k\rangle \langle\beta_k|\beta_j\rangle$$

$$\mathcal{N}(\rho) = \sum_i Tr(\rho M_i)\sigma_i \quad s.t. \quad \sum_i M_i = \mathbb{1}$$

$$= \sum_j |\beta_j\rangle\langle\alpha_j|\rho|\alpha_j\rangle\langle\beta_j|$$

Stinespring: $V: |\phi\rangle_A \rightarrow \sum_j \langle\alpha_j|\phi\rangle\, |\beta_j\rangle_B\, |j\rangle_E$

$$\hat{\mathcal{N}}(\rho) = \sum_{jk} |j\rangle\langle k|\, \langle\alpha_j|\rho|\alpha_k\rangle\, \langle\beta_k|\beta_j\rangle$$

$$= \qquad\qquad\qquad \mathcal{U}\rho\,\mathcal{U}^\dagger \circ S$$

$$\mathcal{N}(\rho) = \sum_i Tr(\rho M_i)\sigma_i \quad s.t. \quad \sum_i M_i = \mathbb{1}$$

$$= \sum_j |\beta_j\rangle\langle\alpha_j|\rho|\alpha_j\rangle\langle\beta_j|$$

Stinespring: $V: |\phi\rangle_A \to \sum_j \langle\alpha_j|\phi\rangle\, |\beta_j\rangle_B |j\rangle_E$

$$\hat{\mathcal{N}}(\rho) = \sum_{jk} |j\rangle\langle k|\, \langle\alpha_j|\rho|\alpha_k\rangle\, \langle\beta_k|\beta_j\rangle$$

$$= \qquad \mathcal{U}\rho\,\mathcal{U}^\dagger \circ S$$

isometry $\mathcal{U} = \sum |j\rangle\langle\alpha_j| : A \to E$

$$\mathcal{N}(\rho) = \sum_i \mathrm{Tr}(\rho M_i)\sigma_i \quad s.t. \quad \sum_i M_i = \mathbb{1}$$

$$= \sum_j |\beta_j\rangle\langle\alpha_j|\rho|\alpha_j\rangle\langle\beta_j|$$

Stinespring: $V : |\phi\rangle_A \to \sum_j \langle\alpha_j|\phi\rangle\, |\beta_j\rangle_B |j\rangle_E$

$$\hat{\mathcal{N}}(\rho) = \sum_{jk} |j\rangle\langle k|\; \langle\alpha_j|\rho|\alpha_k\rangle\, \langle\beta_k|\beta_j\rangle$$

$$= \qquad U\rho U^\dagger \circ S$$

Schur product

isometry $U = \sum |j\rangle\langle\alpha_j| \; : \; A \to E$

$$\mathcal{N}(\rho) = \sum_i \mathrm{Tr}(\rho M_i)\,\sigma_i \quad s.t. \quad \sum_i M_i = \mathbb{1}$$

$$= \sum_j |\beta_j\rangle\langle\alpha_j|\rho|\alpha_j\rangle\langle\beta_j|$$

Stinespring: $V: |\phi\rangle_A \to \sum_j \langle\alpha_j|\phi\rangle\, |\beta_j\rangle_B\, |j\rangle_E$

$$\hat{\mathcal{N}}(\rho) = \sum_{jk} |j\rangle\langle k|\, \langle\alpha_j|\rho|\alpha_k\rangle\, \langle\beta_k|\beta_j\rangle$$

$$= U \rho U^\dagger \circ S$$

Schur product

isometry $U = \sum |j\rangle\langle\alpha_j| : A \to E$

Channels of this form: Hadamard channels

Examples – Entanglement-breaking channels:

1) cq-channels, i.e. classical input determines state preparation at output

Examples – Entanglement-breaking channels:

1) cq-channels, i.e. classical input determines state preparation at output

2) qc-channels, i.e. measurement with classical output

Examples - Entanglement-breaking channels:

1) cq-channels, i.e. classical input determines state preparation at output
2) qc-channels, i.e. measurement with classical output

Hadamard channels:

3) Phase damping channels, more generally Schur multipliers

Examples - Entanglement-breaking channels:

1) cq-channels, i.e. classical input determines state preparation at output
2) qc-channels, i.e. measurement with classical output

Hadamard channels:

3) Phase damping channels, more generally Schur multipliers
4) Cloning channels [cf. Brádler, IEEE-IT 2011]

The proof is beautiful but a bit long...

The proof is beautiful but a bit long...

Departure point minimax characterization of $\chi(\mathcal{N})$: [Schumacher/Westmoreland, PRA 2000]

$$\chi(\mathcal{N}) = \min_{\sigma} \max_{\rho} D(\mathcal{N}(\rho) \| \sigma)$$

The proof is beautiful but a bit long...

Departure point minimax characterization of $\chi(\mathcal{N})$: [Schumacher/Westmoreland, PRA 2000]

$$\chi(\mathcal{N}) = \min_{\sigma} \max_{\rho} D(\mathcal{N}(\rho) \| \sigma)$$

Relative entropy:

$$D(\rho \| \sigma) = \mathrm{Tr}\, \rho(\log\rho - \log\sigma)$$

The proof is beautiful but a bit long...

Departure point minimax characterization of $\chi(N)$: [Schumacher/Westmoreland, PRA 2000]

$$\chi(N) = \min_{\sigma} \max_{\rho} D(N(\rho) \| \sigma)$$

Relative entropy:

$$D(\rho \| \sigma) = \mathrm{Tr}\, \rho(\log \rho - \log \sigma)$$

Note: For EB and $\mathcal{H}$ channels N this is additive, and so $C(N) = \chi(N)$.

[Shor, JMP 2002 (EB); King et al., quant-ph/0509126 ($\mathcal{H}$)]

Relative entropy

$$D(\rho \| \sigma) = \mathrm{Tr}\, \rho(\log\rho - \log\sigma)$$

is a special case of a whole family of "generalized divergences"...

[Cf. Petz, 0909.3647; Müller-Lennert et al., 1306.3142]

(➢ talk by Marco Tomamichel, Fri 10:20)

Relative entropy

$$D(\rho \| \sigma) = \mathrm{Tr}\, \rho(\log \rho - \log \sigma)$$

is a special case of a whole family of "generalized divergences"...

[Cf. Petz, 0909.3647; Müller-Lennert et al., 1306.3142]

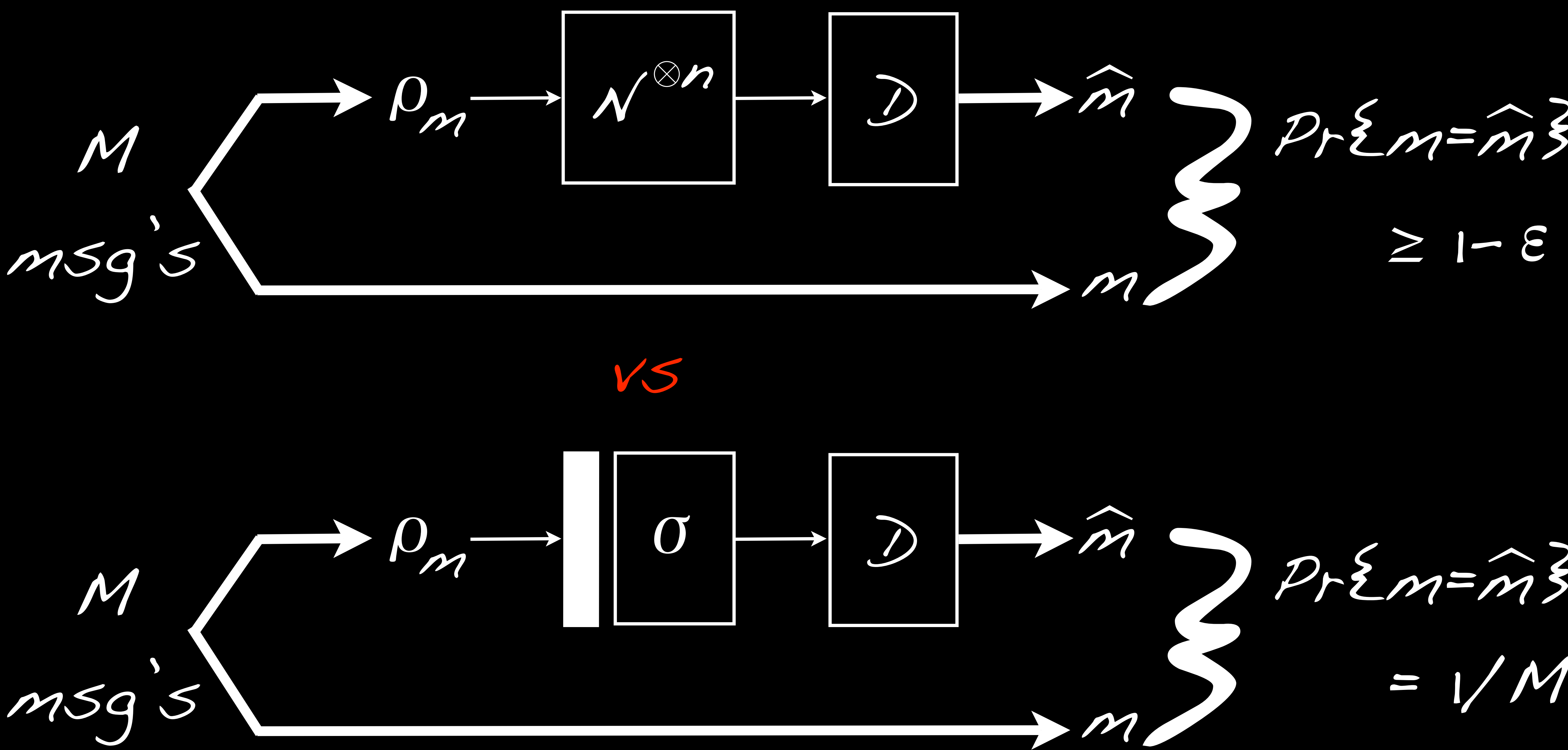
(➢ talk by Marco Tomamichel, Fri 10:20)

Fundamental property is monotonicity:

For any cptp map N,

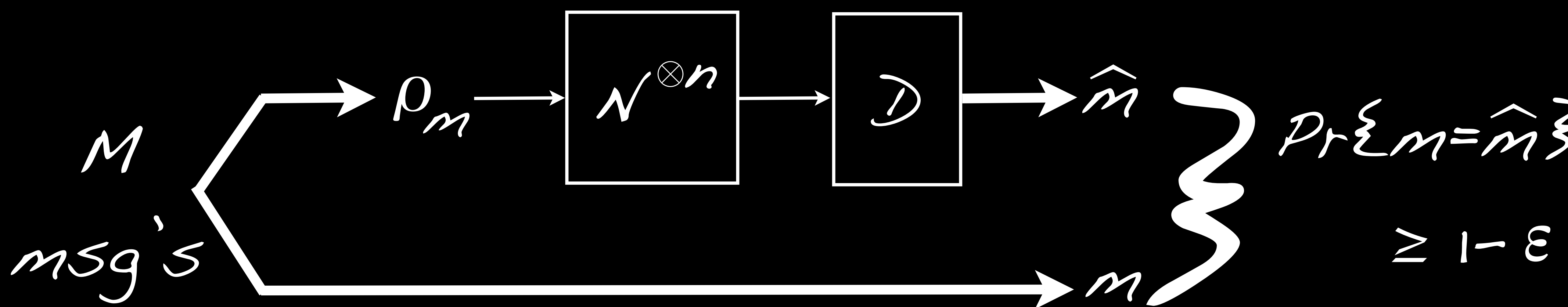
$$\tilde{D}(\rho \| \sigma) \geq \tilde{D}(N(\rho) \| N(\sigma)) \geq 0.$$

Compare code for $\mathcal{N}^{\otimes n}$ with trivial channel:

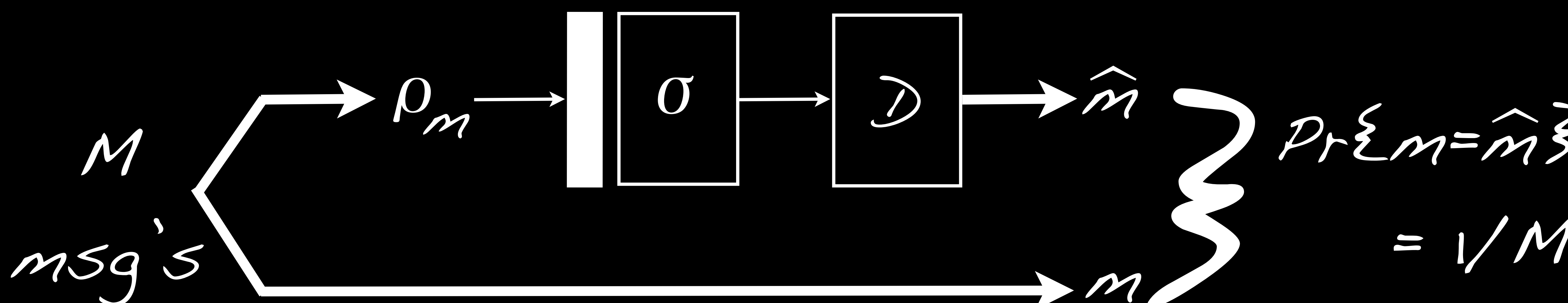


[Cf. Polyanskyi/Verdú, Proc. 48th Allerton CCC, 2010]

Compare code for $\mathcal{N}^{\otimes n}$ with trivial channel:



vs



$$\tilde{D}(1-\varepsilon \,\|\, 1/M) \leq \min_{\sigma} \max_{\rho} \tilde{D}(\mathcal{N}^{\otimes n}(\rho) \,\|\, \sigma)$$

[Cf. Polyanskyi/Verdú, Proc. 48th Allerton CCC, 2010]

Compare code for $N^{\otimes n}$ with trivial channel:

$$\tilde{D}(1-\varepsilon \,\|\, 1/M) \leq \min_{\sigma} \max_{\rho} \tilde{D}(N^{\otimes n}(\rho) \,\|\, \sigma)$$
$$=: \chi_{\tilde{D}}(N^{\otimes n})$$

Compare code for $\mathcal{N}^{\otimes n}$ with trivial channel:

$$\tilde{D}(1-\varepsilon \,\|\, 1/M) \leq \min_{\sigma} \max_{\rho} \tilde{D}(\mathcal{N}^{\otimes n}(\rho) \,\|\, \sigma)$$

$$=: \chi_{\tilde{D}}(\mathcal{N}^{\otimes n})$$

(For usual relative entropy, we recover the previous weak converse.)

Compare code for $N^{\otimes n}$ with trivial channel:

$$\tilde{D}(1-\varepsilon \,\|\, 1/M) \le \min_{\sigma} \max_{\rho} \tilde{D}(N^{\otimes n}(\rho) \,\|\, \sigma)$$
$$=: \chi_{\tilde{D}}(N^{\otimes n})$$

Sandwiched α-Rényi relative entropy ($\alpha > 1$):

$$\tilde{D}_\alpha(\rho \,\|\, \sigma) := \frac{1}{\alpha - 1} \log \mathrm{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \, \sigma^{\frac{1-\alpha}{2\alpha}} \right)^{\alpha}$$

[Cf. Müller-Lennert et al., 1306.3142; Beigi 1306.5920; Frank/Lieb 1306.5358]

Compare code for $\mathcal{N}^{\otimes n}$ with trivial channel:

$$\tilde{D}(1-\varepsilon \,\|\, 1/M) \leq \min_{\sigma} \max_{\rho} \tilde{D}(\mathcal{N}^{\otimes n}(\rho) \,\|\, \sigma)$$

$$=: \chi_{\tilde{D}}(\mathcal{N}^{\otimes n})$$

Sandwiched α-Rényi relative entropy ($\alpha > 1$):

$$\tilde{D}_{\alpha}(\rho \,\|\, \sigma) := \frac{1}{\alpha - 1} \log \mathrm{Tr} \left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \, \sigma^{\frac{1-\alpha}{2\alpha}} \right)^{\alpha}$$

[Cf. Müller-Lennert et al., 1306.3142; Beigi 1306.5920; Frank/Lieb 1306.5358]

Crucially additive: $\tilde{\chi}_{\alpha}(\mathcal{N}^{\otimes n}) = n \, \tilde{\chi}_{\alpha}(\mathcal{N})$.

($\triangleright$ talk by Marco Tomamichel, Fri 10:20)

$$\tilde{D}(1-\varepsilon \| 1/M) \le \min_{\sigma} \max_{\rho} \tilde{D}(\mathcal{N}^{\otimes n}(\rho) \| \sigma)$$

$$=: \chi_{\tilde{D}}(\mathcal{N}^{\otimes n})$$

Sandwiched α-Rényi relative entropy $(\alpha > 1)$:

$$\tilde{D}_{\alpha}(\rho \| \sigma) := \frac{1}{\alpha - 1} \log Tr \left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \, \sigma^{\frac{1-\alpha}{2\alpha}} \right)^{\alpha}$$

[Cf. Müller-Lennert et al., 1306.3142; Beigi 1306.5920; Frank/Lieb 1306.5358]

Crucially additive: $\tilde{\chi}_{\alpha}(\mathcal{N}^{\otimes n}) = n \, \tilde{\chi}_{\alpha}(\mathcal{N})$.

(x EB & $\mathcal{H}$ channels!)

$$\tilde{D}(1-\varepsilon \,\|\, 1/M) \le \min_{\sigma} \max_{\rho} \tilde{D}(\mathcal{N}^{\otimes n}(\rho) \,\|\, \sigma)$$

$$=: \chi_{\tilde{D}}(\mathcal{N}^{\otimes n})$$

Sandwiched α-Rényi relative entropy ($\alpha > 1$):

$$\tilde{D}_\alpha(\rho \,\|\, \sigma) := \frac{1}{\alpha - 1} \log \operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \, \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha$$

[Cf. Müller-Lennert et al., 1306.3142; Beigi 1306.5920; Frank/Lieb 1306.5358]

Crucially additive: $\tilde{\chi}_\alpha(\mathcal{N}^{\otimes n}) = n \, \tilde{\chi}_\alpha(\mathcal{N})$.

...and converges to $\chi(\mathcal{N})$ as $\alpha \to 1$!

(x EB & $\mathcal{H}$ channels!)

$$\tilde{D}(1-\varepsilon \,\|\, 1/M) \le \min_{\sigma} \max_{\rho} \tilde{D}(\mathcal{N}^{\otimes n}(\rho) \,\|\, \sigma)$$

$$=: \chi_{\tilde{D}}(\mathcal{N}^{\otimes n})$$

Sandwiched α-Rényi relative entropy ($\alpha > 1$):

$$\tilde{D}_\alpha(\rho \,\|\, \sigma) := \frac{1}{\alpha - 1} \log \operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \, \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha$$

[Cf. Müller-Lennert et al., 1306.3142; Beigi 1306.5920; Frank/Lieb 1306.5358]

Crucially additive: $\tilde{\chi}_\alpha(\mathcal{N}^{\otimes n}) = n \, \tilde{\chi}_\alpha(\mathcal{N})$.

...and converges to $\chi(\mathcal{N})$ as $\alpha \to 1$!

(x EB & 𝓗 channels!)

L.h.s.: $\frac{\alpha}{\alpha - 1} \log(1-\varepsilon) + \log M.$

For $\log M = nR$:

$$1-\varepsilon \leq \exp\left\{-n\,\frac{\alpha-1}{\alpha}\left(R-\tilde{\chi}_\alpha(N)\right)\right\},$$

which is exponentially small for $R > \chi(N)$ and $\alpha > 1$ small enough.

QED

For $\log M = nR$:

$$1-\varepsilon \leq \exp\left\{-n\,\frac{\alpha-1}{\alpha}\left(R-\tilde{\chi}_\alpha(\mathcal{N})\right)\right\},$$

which is exponentially small for $R > \chi(\mathcal{N})$ and $\alpha > 1$ small enough. QED

Combining with simulation argument and recent additivity of minimum output (Rényi) entropies [Giovannetti/García-Patrón/Cerf/Holevo, 1312.6225]: Strong converse for covariant Gaussian channels. ✓

[Bardhan/García-Patrón/Wilde/AW, 1401.4161]

6. Min-entropies: "pretty strong" converse for Q

Stinespring: $\mathcal{N}(\rho) = \mathrm{Tr}_E\, V\rho V^\dagger,$

with an isometry $V: A \hookrightarrow B \otimes E$.

Complementary channel:

$$\widehat{\mathcal{N}}(\rho) = \mathrm{Tr}_B\, V\rho V^\dagger.$$

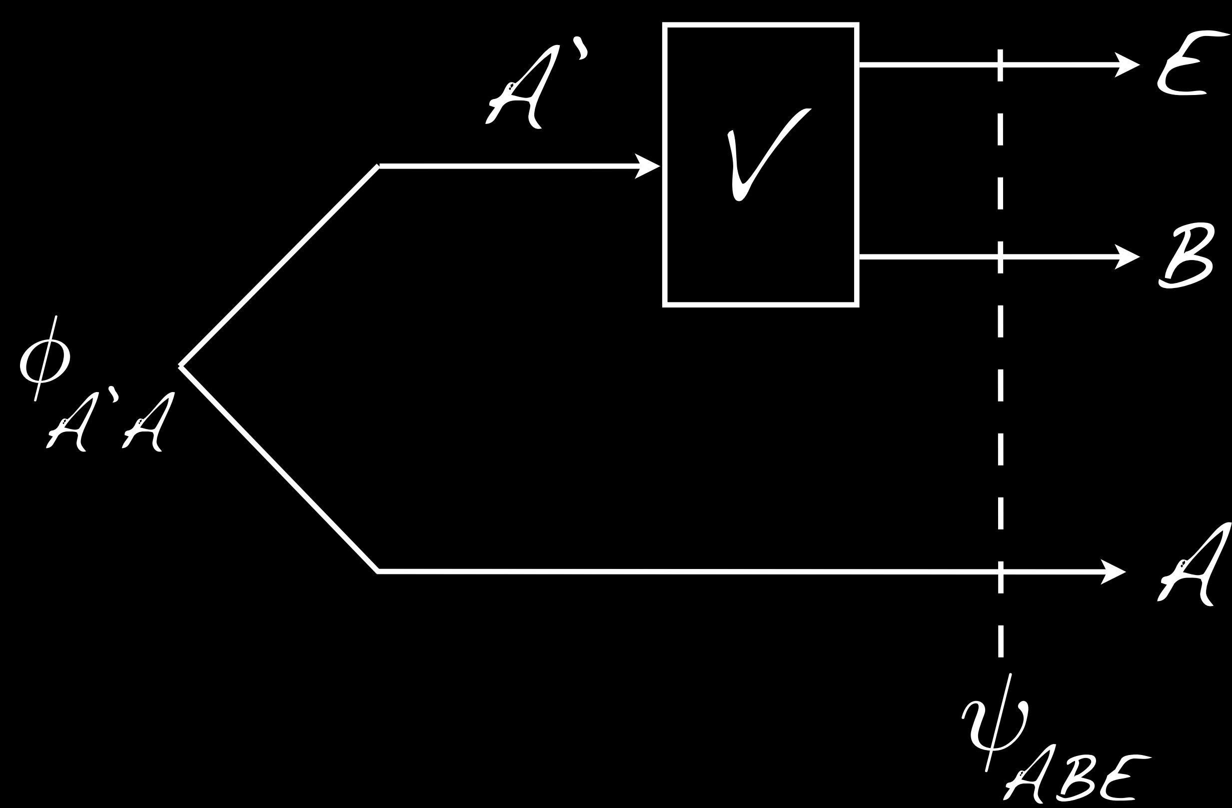
6. Min-entropies: "pretty strong" converse for Q

Stinespring: $\mathcal{N}(\rho) = \mathrm{Tr}_E\, V\rho V^\dagger,$
with an isometry $V: A \hookrightarrow B \otimes E.$

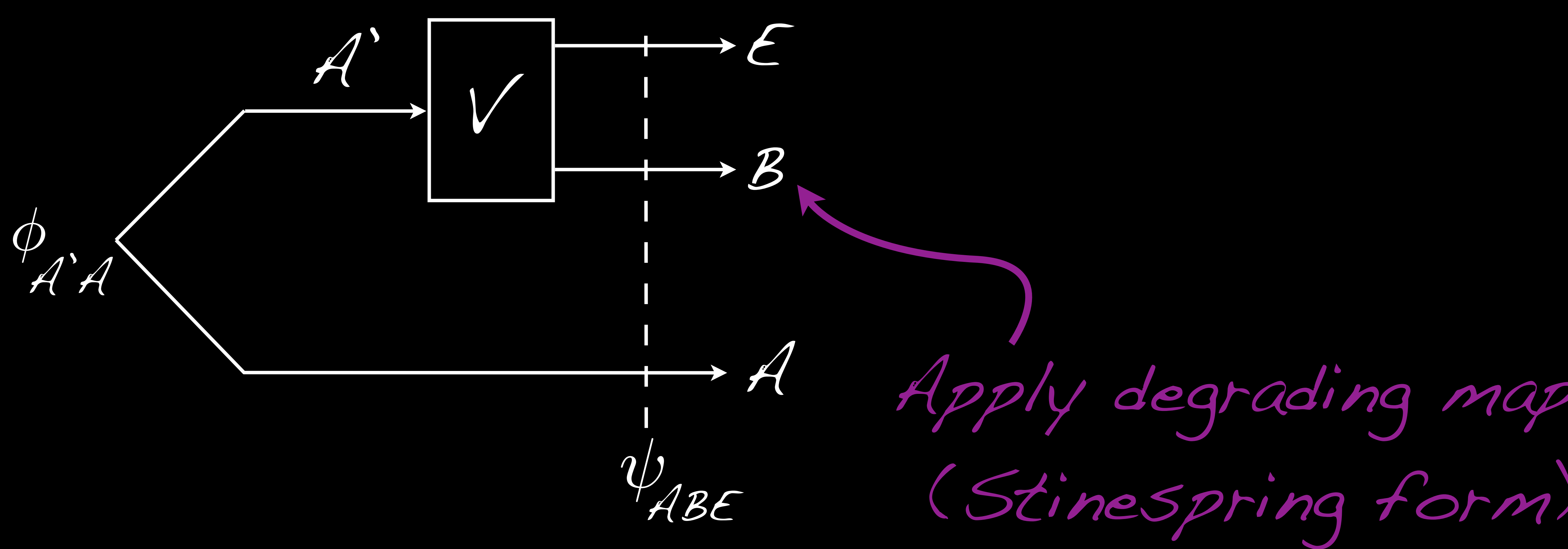
Complementary channel:
$\widehat{\mathcal{N}}(\rho) = \mathrm{Tr}_B\, V\rho V^\dagger.$

$\mathcal{N}$ is degradable if there exists a cptp map $\mathcal{D}$ s.t. $\widehat{\mathcal{N}} = \mathcal{D} \circ \mathcal{N}$. Vice-versa: anti-degradable.

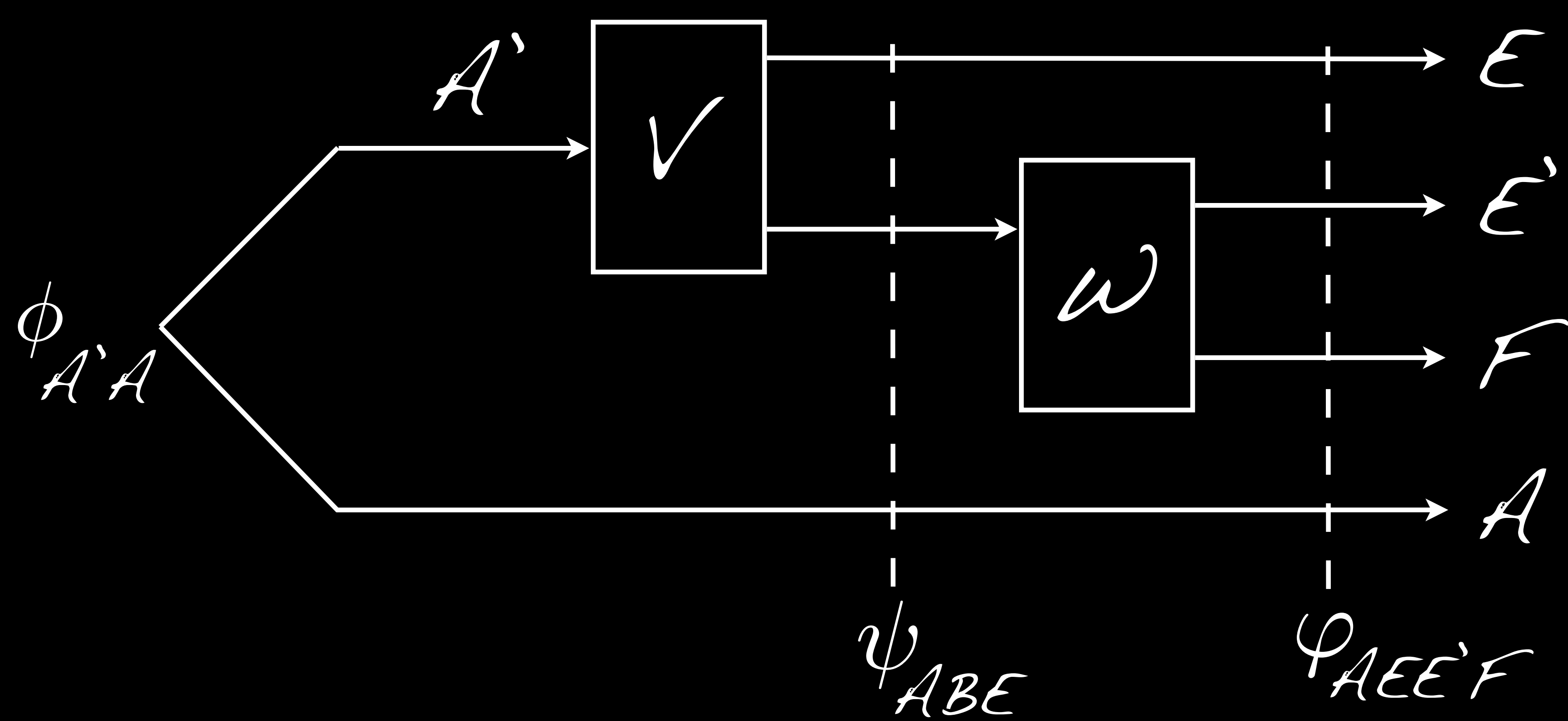
Degradability in the Church of the Larger Hilbert Space:



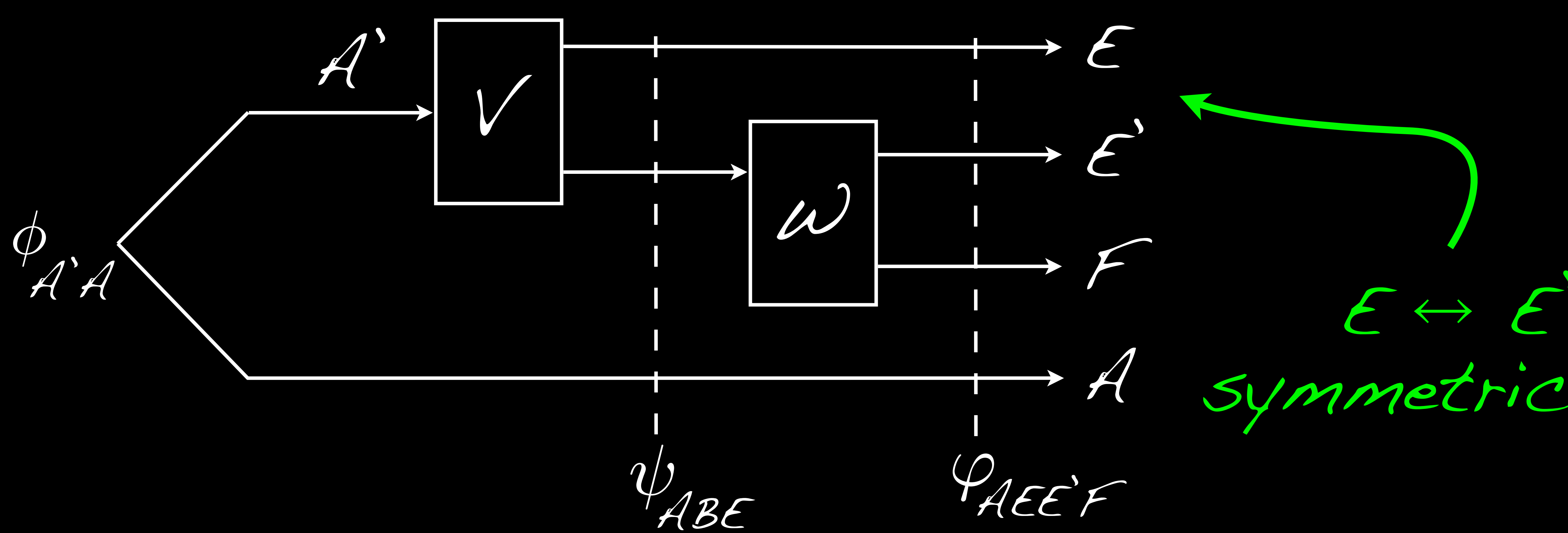
Degradability in the Church of the Larger Hilbert Space:

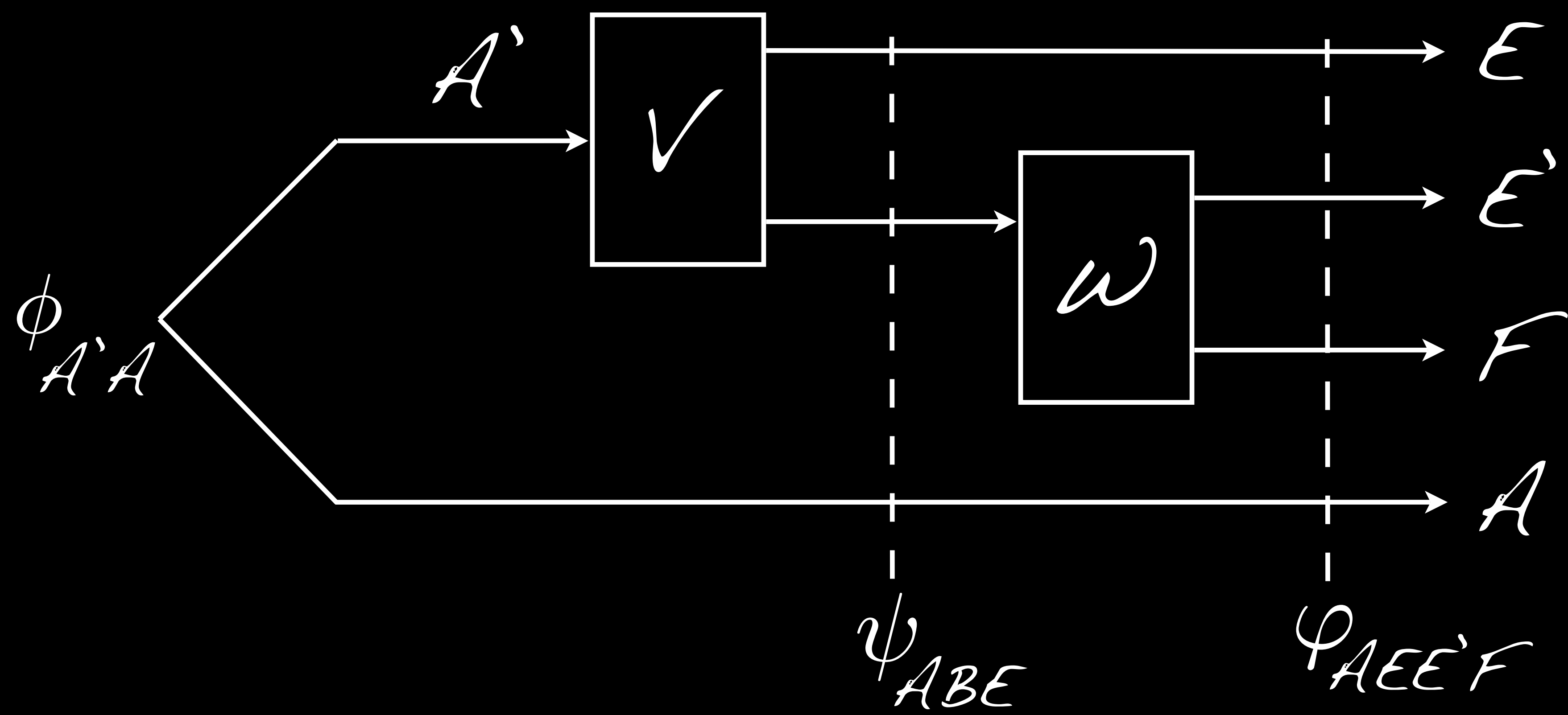


Degradability in the Church of the Larger Hilbert Space:



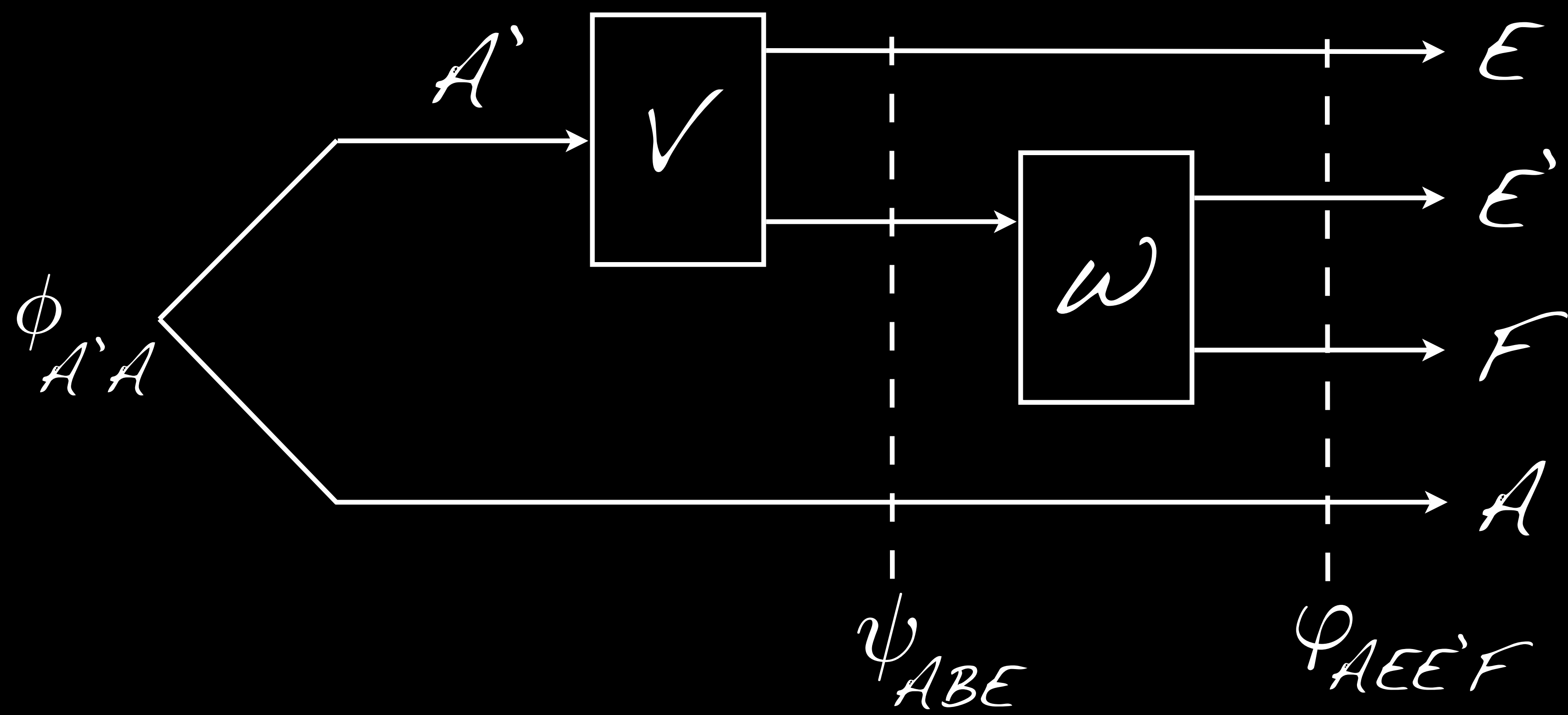
Degradability in the Church of the Larger Hilbert Space:





Examples:

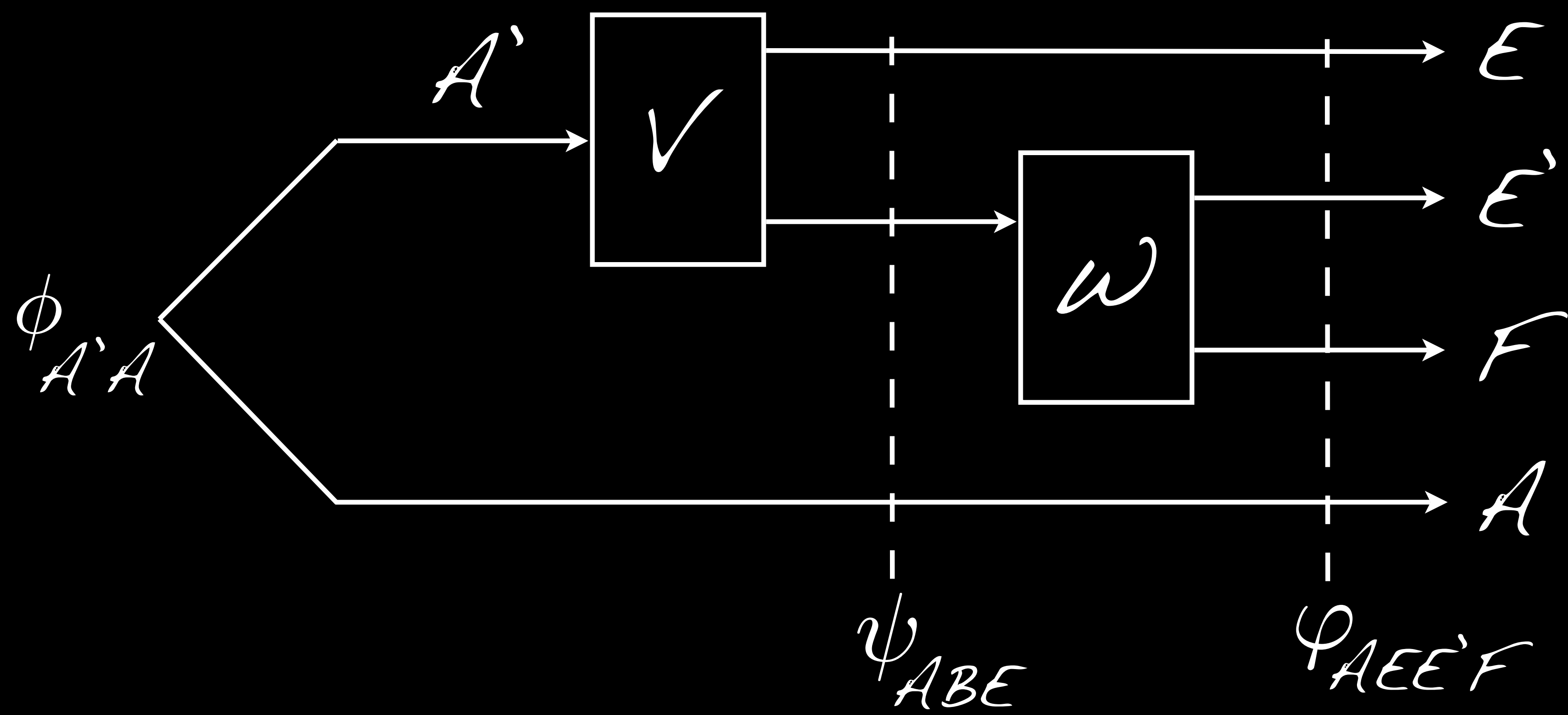
1) Phase damping channel, more generally Schur multipliers and Hadamard channels



Examples:

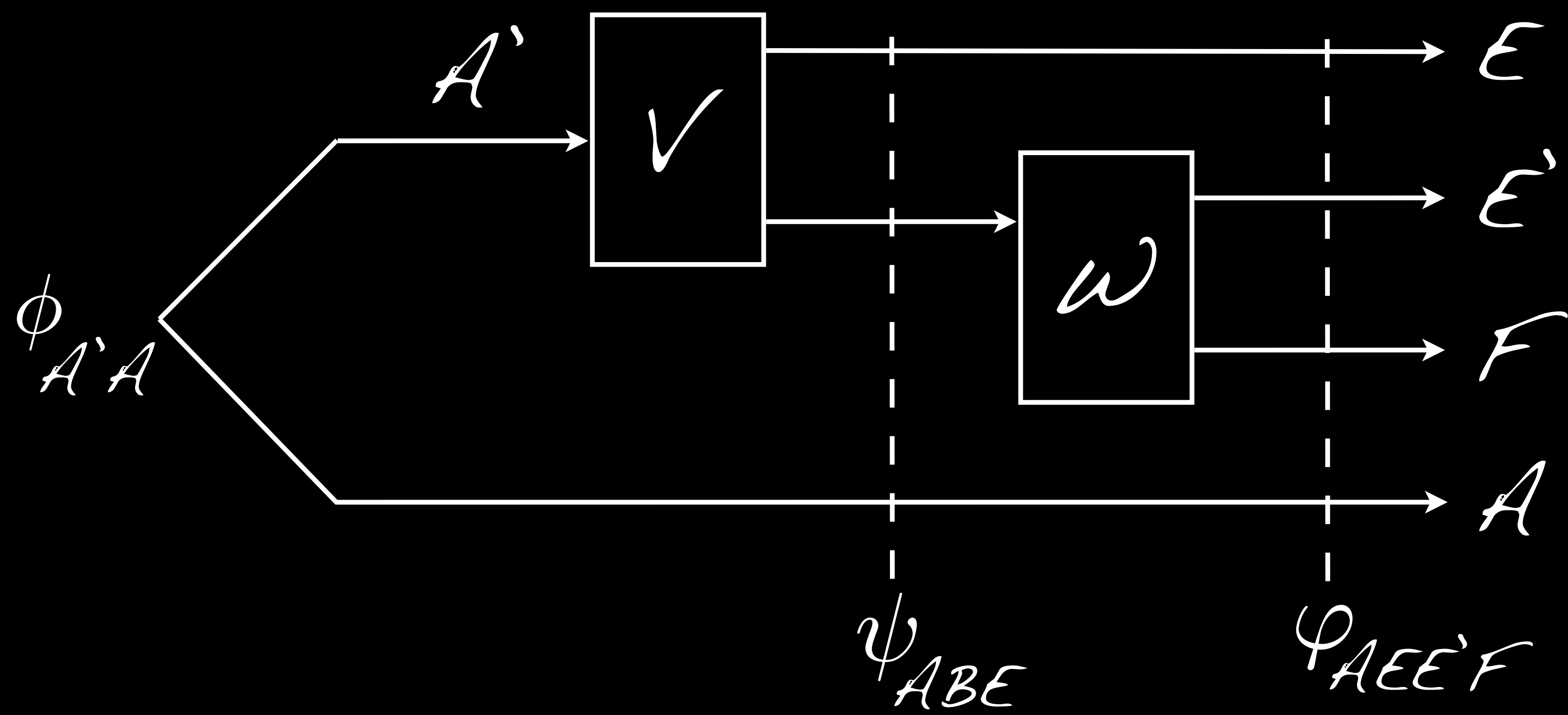
1) Phase damping channel, more generally Schur multipliers and Hadamard channels

2) Amplitude damping channel



Examples:

1) Phase damping channel, more generally Schur multipliers and Hadamard channels
2) Amplitude damping channel
3) Symmetric channels, i.e. trivial F, for instance 50% erasure channel

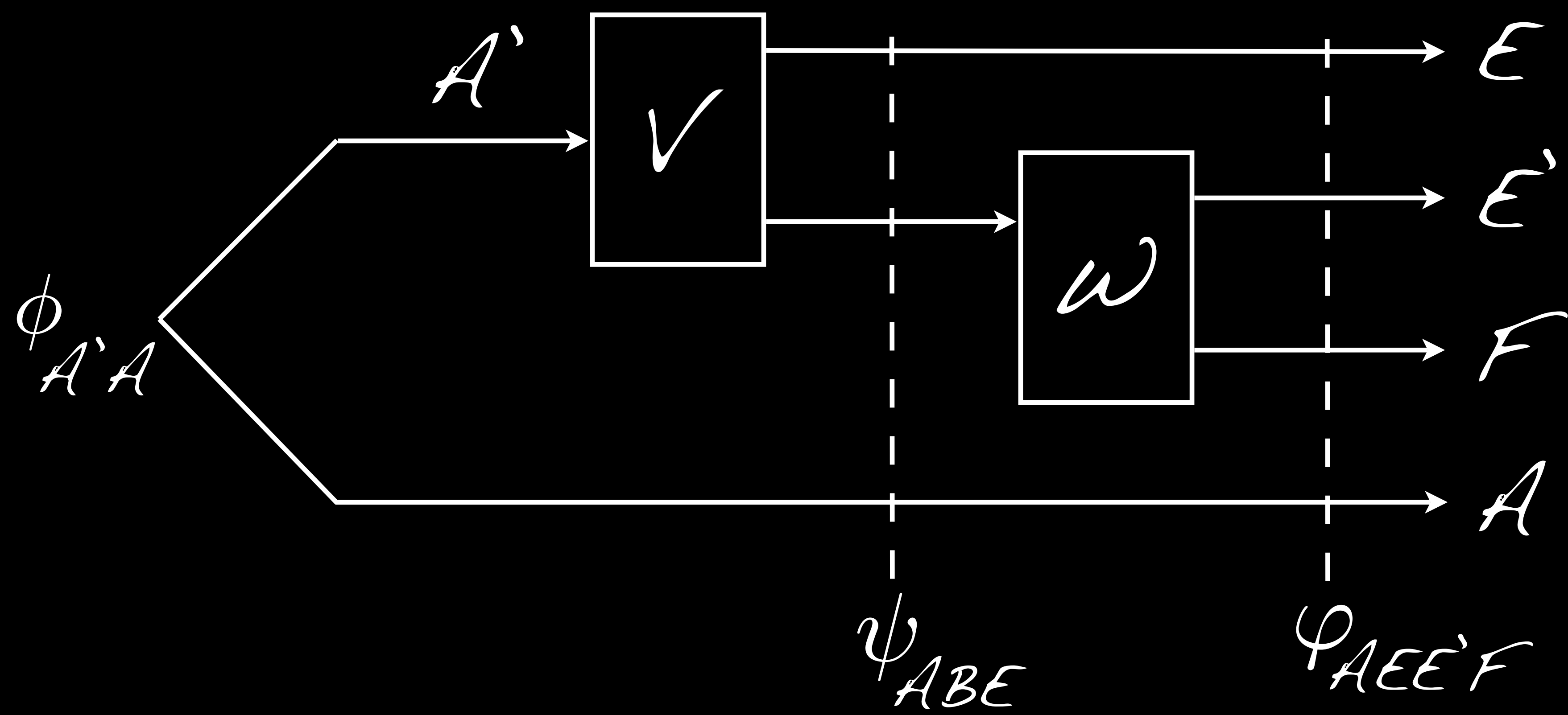


Why we are interested in degr. channels:

$Q^{(1)}(N)$ is additive and so $Q(N) = Q^{(1)}(N)$,

and the latter is a convex optimisation.

[Devetak/Shor, CMP 256:287 (2005)]



Why we are interested in degr. channels:

$Q^{(1)}(N)$ is additive and so $Q(N) = Q^{(1)}(N)$,

and the latter is a convex optimisation.

[Devetak/Shor, CMP 256:287 (2005)]

(For anti-degradable N: $Q(N) = 0$.)

A previous result [via E. Rains, IEEE-IT 47(7):2921-2933 (2001)]: If N is PPT entanglement-binding, then of course $Q(N)=0$, and strong converse holds (with error converging exponentially to 1).

A previous result [via E. Rains, IEEE-IT 47(7):2921-2933 (2001)]: If N is PPT entanglement-binding, then of course $Q(N)=0$, and strong converse holds (with error converging exponentially to 1).

Note: Already for symmetric (degradable & anti-degradable) channels - for which also $Q(N)=0$ - not clear at all.

Exercise: Strong converse for noiseless qubit id_2, even assisted by classical communication.

Exercise: Strong converse for noiseless qubit id_2, even assisted by classical communication.

Implies: Error goes to one for rates above $E_C(N)$, the entanglement cost of simulating the channel with free classical communication [Berta et al., IEEE-IT 59(10):6779-6795 (2013)].

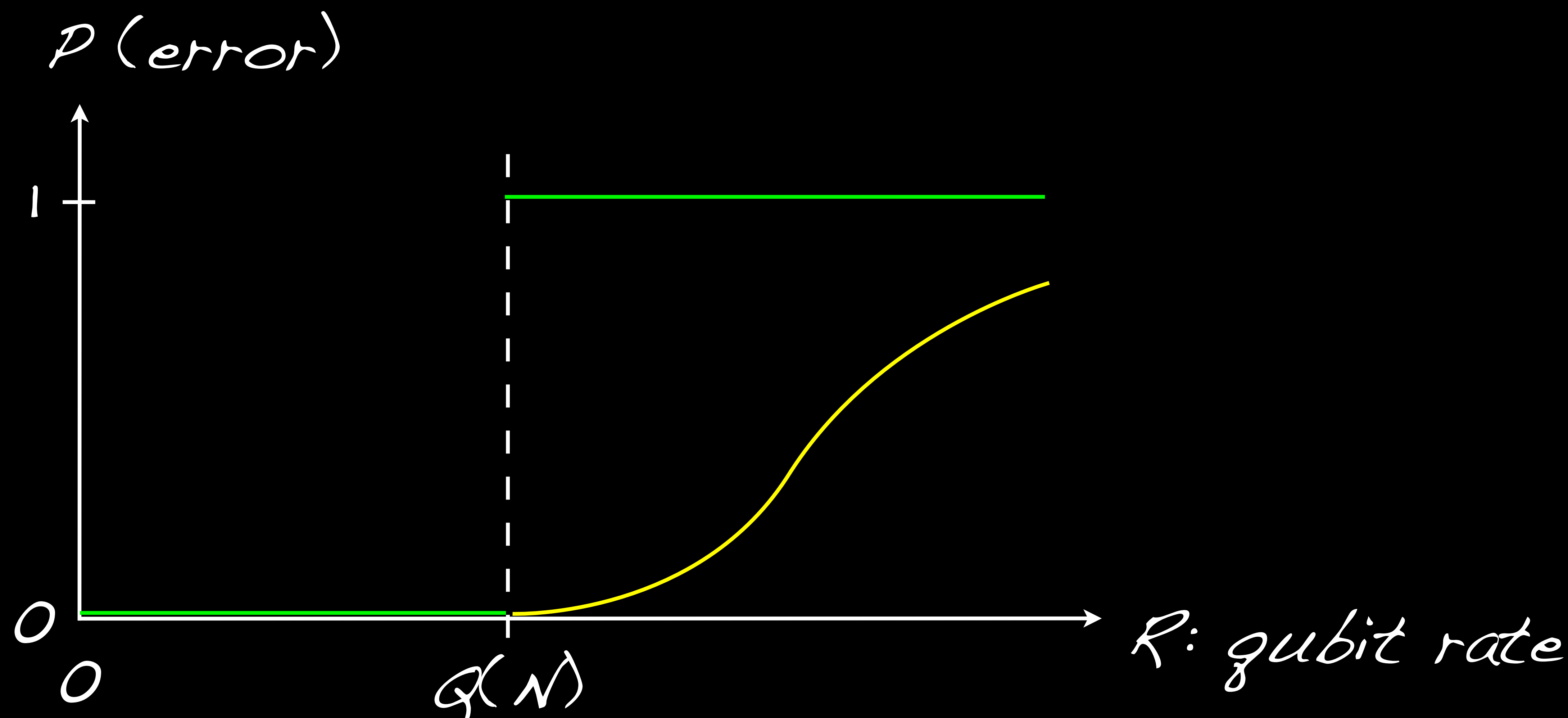
Exercise: Strong converse for noiseless qubit id_2, even assisted by classical communication.

Implies: Error goes to one for rates above $E_C(\mathcal{N})$, the entanglement cost of simulating the channel with free classical communication [Berta et al., IEEE-IT 59(10):6779-6795 (2013)].

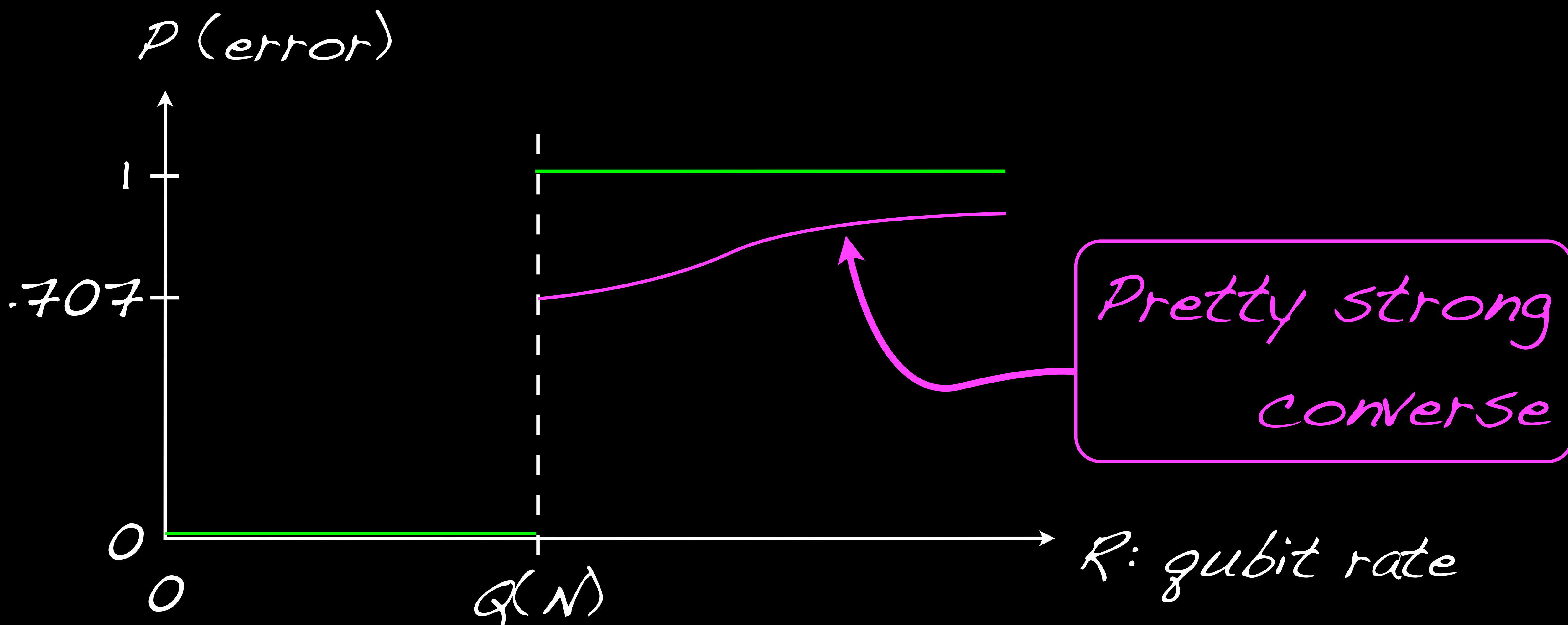
Still doesn't take care of 50% erasure channel, dephasing channels, etc!

Thm (Morgan/AW, 1301.4927): For any degradable channel N, all codes with rate $R > Q(N)$ have error at least 0.707, asymptotically.

Thm (Morgan/AW, 1301.4927): For any degradable channel N, all codes with rate $R > Q(N)$ have error at least 0.707, asymptotically. I.e., at $Q(N)$, the error has a finite "jump":



Thm (Morgan/AW, 1301.4927): For any degradable channel N, all codes with rate $R > Q(N)$ have error at least 0.707, asymptotically. I.e., at $Q(N)$, the error has a finite "jump":



Thm: For any degradable channel $\mathcal{N}$, codes with rate $R > Q(\mathcal{N})$ have error at least 0.707, asymptotically.

Error/fidelity achieved by a single 50% erasure channel - without encoding.

Thm: For any degradable channel N, codes with rate $R > Q(N)$ have error at least 0.707, asymptotically.

Error/fidelity achieved by a single 50% erasure channel - without encoding.

On the other hand: For larger error, any i.i.d. symmetric channel allows coding of $k = c\sqrt{n}$ qubits, by random codes. More?

Thm: For any degradable channel $\mathcal{N}$, codes with rate $R > Q(\mathcal{N})$ have error at least 0.707, asymptotically.

Similar result for private capacity:

Thm (1301.4927): For degradable channel $\mathcal{N}$, if decoding error and distance from perfect privacy are both below some universal threshold, then the rate is asymptotically bounded by $P(\mathcal{N}) = Q(\mathcal{N})$.

Thm: For any degradable channel N, codes with rate $R > Q(N)$ have error at least 0.707, asymptotically.

Significance of symmetric channels:

Thm (1301.4927): If symmetric channels (whose quantum capacity is 0) obey a strong converse, then so do all degradable channels N: for error below 1, rate would be asymptotically bounded by $Q(N)$.

Proof uses tight finite block length characterization of P and Q via (smooth) min-entropies & some tricks: symmetrization, de Finetti theorem, asymptotic equipartition property...

[Cf. R. Renner, PhD thesis, quant-ph/0512258 & M. Tomamichel, PhD thesis, arXiv:1203.2142]

Proof uses tight finite block length characterization of P and Q via (smooth) min-entropies & some tricks: symmetrization, de Finetti theorem, asymptotic equipartition property...

[Cf. R. Renner, PhD thesis, quant-ph/0512258 & M. Tomamichel, PhD thesis, arXiv:1203.2142]

Can be viewed as a complicated version of the proof of additivity: $P(N)=Q(N)=Q^{(1)}(N)$ for degradable N... :-/

[Devetak/Shor, CMP 256:287 (2005)]

7. Conclusion (sort of...)

- Lesson: To get more precise under-standing of code performance have to abandon von Neumann entropy and embrace non-standard entropies (Rényi entropy, min-entropy, ...)
- Price to pay: Each channel and each capacity requires its own approach. Many open - e.g. multi-user channels...

7. Conclusion (sort of...)

- The trick with the sandwiched channel reduces the additivity of $\chi(N)$ to that of the minimum output Rényi entropy of an associated family of cp (trace non-preserving) maps. Can it be applied to other channels? Other divergences?
  [Wilde/AW/Yang, 1306.1586]
- Can we also get "2nd order" behaviour?
  [Cf. Tomamichel/Tan, 1308.6503 for cq-channels]

7. Conclusion (sort of...)

- Big open problem: "pretty strong" is pretty ugly - how to get full strong converse for Q of degradable channels!? Bottleneck are the symmetric channels, e.g. 50% erasure channel...
- How to prove strong converses without additivity? Note that neither P, Q nor $P^{(1)}$, $Q^{(1)}$, χ are generally additive! (Not known for C.)

A. Proof ideas for C

A goody first: minimax characterisation of $\chi(\mathcal{N})$: [Schumacher/Westmoreland, PRA 2000]

$$\chi(\mathcal{N}) = \min_{\sigma} \max_{\rho} D(\mathcal{N}(\rho) \| \sigma)$$

Note: For EB and $\mathcal{H}$ channels $\mathcal{N}$ this is additive, and so $C(\mathcal{N}) = \chi(\mathcal{N})$.

[Shor, JMP 2002 (EB); King et al., quant-ph/0509126 ($\mathcal{H}$)]

A. Proof ideas for C

A goody first: minimax characterisation of $\chi(N)$: [Schumacher/Westmoreland, PRA 2000]

$$\chi(N) = \min_{\sigma} \max_{\rho} D(N(\rho) \| \sigma)$$

Relative entropy:

$$D(\rho \| \sigma) = \mathrm{Tr}\, \rho(\log \rho - \log \sigma)$$

Relative entropy

$$D(\rho \| \sigma) = \mathrm{Tr}\, \rho(\log \rho - \log \sigma)$$

is a special case of a whole family of "generalised divergences".

[Cf. Petz, 0909.3647; Müller-Lennert et al., 1306.3142]

Relative entropy

$$D(\rho \| \sigma) = \mathrm{Tr}\, \rho(\log \rho - \log \sigma)$$

is a special case of a whole family of "generalised divergences".

[Cf. Petz, 0909.3647; Müller-Lennert et al., 1306.3142]

Fundamental property is monotonicity: for any cptp map $\mathcal{N}$,

$$\tilde{D}(\rho \| \sigma) \geq \tilde{D}(\mathcal{N}(\rho) \| \mathcal{N}(\sigma)) \geq 0. \qquad (*)$$

Relative entropy

$$D(\rho\|\sigma) = \mathrm{Tr}\,\rho(\log\rho - \log\sigma)$$

is a special case of a whole family of "generalised divergences".

[Cf. Petz, 0909.3647; Müller-Lennert et al., 1306.3142]

Fundamental property is monotonicity: for any cptp map N,

$$\tilde{D}(\rho\|\sigma) \geq \tilde{D}(N(\rho)\|N(\sigma)) \geq 0. \qquad (*)$$

Notation: for binary distributions $P=(p,1-p)$ and $Q=(q,1-q)$, write $\tilde{D}(P\|Q) = \tilde{D}(p\|q)$.

Assume furthermore that

$$\tilde{D}\Big(\bigoplus_x p_x \rho_x \,\Big\|\, \bigoplus_x p_x \sigma_x\Big) = \sum_x p_x \tilde{D}(\rho_x \| \sigma_x). \quad (+)$$

Assume furthermore that

$$\tilde{D}\Big(\bigoplus_x p_x \rho_x \,\|\, \bigoplus_x p_x \sigma_x\Big) = \sum_x p_x \tilde{D}(\rho_x \| \sigma_x). \quad (+)$$

Then, for a code with M msg's, error $\leq \varepsilon$,

and $\rho_{XB} = \frac{1}{M} \sum_m |m\rangle\langle m| \otimes \mathcal{N}(\rho_m)$:

Assume furthermore that

$$\tilde{D}\Big(\bigoplus_x p_x \rho_x \,\Big\|\, \bigoplus_x p_x \sigma_x\Big) = \sum_x p_x \tilde{D}(\rho_x \| \sigma_x). \quad (+)$$

Then, for a code with M msg's, error $\leq \varepsilon$, and $\rho_{XB} = \frac{1}{M} \sum_m |m\rangle\langle m| \otimes \mathcal{N}(\rho_m)$:

$$\begin{aligned}
\tilde{D}(1-\varepsilon \,\|\, 1/M) &\overset{(*)}{\leq} \tilde{D}(\rho_{XB} \| \rho_X \otimes \sigma) \\
&\overset{(+)}{\leq} \frac{1}{M} \sum_m \tilde{D}(\mathcal{N}(\rho_m) \| \sigma) \\
&\leq \max_\rho \tilde{D}(\mathcal{N}(\rho) \| \sigma) =: \chi_{\tilde{D},\sigma}(\mathcal{N})
\end{aligned}$$

Assume furthermore that

$$\tilde{D}\Big(\bigoplus_x p_x \rho_x \,\Big\|\, \bigoplus_x p_x \sigma_x\Big) = \sum_x p_x \tilde{D}(\rho_x \| \sigma_x). \quad (+)$$

Then, for a code with M msg's, error $\leq \varepsilon$,

and $\rho_{XB} = \frac{1}{M} \sum_m |m\rangle\langle m| \otimes \mathcal{N}(\rho_m)$:

$$\begin{aligned}
\tilde{D}(1-\varepsilon \| 1/M) &\overset{(*)}{\leq} \tilde{D}(\rho_{XB} \| \rho_X \otimes \sigma) \\
&\overset{(+)}{\leq} \frac{1}{M} \sum_m \tilde{D}(\mathcal{N}(\rho_m) \| \sigma) \\
&\leq \max_\rho \tilde{D}(\mathcal{N}(\rho) \| \sigma) =: \chi_{\tilde{D},\sigma}(\mathcal{N})
\end{aligned}$$

[Cf. Nagaoka (≈2000); Polyanskiy/Verdú (2010); Sharma/Warsi, 1205.1712.]

$$\tilde{D}(1-\varepsilon \| 1/M) \le \max_{\rho} \tilde{D}(\mathcal{N}(\rho) \| \sigma) =: \chi_{\tilde{D},\sigma}(\mathcal{N})$$

$$\tilde{D}(1-\varepsilon \,\|\, 1/M) \leq \max_{\rho} \tilde{D}(\mathcal{N}(\rho) \,\|\, \sigma) =: \chi_{\tilde{D},\sigma}(\mathcal{N})$$

Everything depends on right choice of $\tilde{D}$:

$$\tilde{D}(1-\varepsilon \,\|\, 1/M) \leq \max_{\rho} \tilde{D}(\mathcal{N}(\rho) \,\|\, \sigma) =: \chi_{\tilde{D},\sigma}(\mathcal{N})$$

Everything depends on right choice of $\tilde{D}$:

Sandwiched α-Rényi relative entropy $(\alpha > 1)$

$$\tilde{D}_{\alpha}(\rho \,\|\, \sigma) := \frac{1}{\alpha - 1} \log \operatorname{Tr} \left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \, \sigma^{\frac{1-\alpha}{2\alpha}} \right)^{\alpha}$$

$$\tilde{D}(1-\varepsilon \,\|\, 1/M) \leq \max_{\rho} \tilde{D}(\mathcal{N}(\rho) \,\|\, \sigma) =: \chi_{\tilde{D},\sigma}(\mathcal{N})$$

Everything depends on right choice of $\tilde{D}$:

Sandwiched α-Rényi relative entropy ($\alpha > 1$)

$$\tilde{D}_{\alpha}(\rho \,\|\, \sigma) := \frac{1}{\alpha - 1} \log \mathrm{Tr} \left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \, \sigma^{\frac{1-\alpha}{2\alpha}} \right)^{\alpha}$$

[Cf. Müller-Lennert et al., 1306.3142; Beigi 1306.5920; Frank/Lieb 1306.5358]

It's monotonic, has property (+) and is $\leq D_{\alpha}(\rho \,\|\, \sigma) := \frac{1}{\alpha - 1} \log \mathrm{Tr}\, \rho^{\alpha} \sigma^{1-\alpha}$, with which it coincides when states commute.

$$\boxed{\tilde{D}_\alpha(\rho \| \sigma) := \frac{1}{\alpha - 1} \log \operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho\, \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha}$$

$$\tilde{D}_\alpha(1-\varepsilon \| 1/M) \le \max_\rho \tilde{D}_\alpha(\mathcal{N}(\rho) \| \sigma) =: \chi_{\alpha,\sigma}(\mathcal{N})$$

$$\boxed{\tilde{D}_\alpha(\rho \| \sigma) := \frac{1}{\alpha - 1} \log \mathrm{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho\, \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha}$$

$$\tilde{D}_\alpha(1-\varepsilon \| 1/M) \leq \max_\rho \tilde{D}_\alpha(\mathcal{N}(\rho) \| \sigma) =: \chi_{\alpha,\sigma}(\mathcal{N})$$

Lhs: $\tilde{D}_\alpha(1-\varepsilon \| 1/M) \geq \log M + \frac{\alpha}{\alpha - 1} \log(1-\varepsilon)$

$$\boxed{\tilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log \mathrm{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\,\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha}$$

$$\tilde{D}_\alpha(1-\varepsilon\|1/M) \le \max_\rho \tilde{D}_\alpha(\mathcal{N}(\rho)\|\sigma) =: \chi_{\alpha,\sigma}(\mathcal{N})$$

Lhs: $\tilde{D}_\alpha(1-\varepsilon\|1/M) \ge \log M + \frac{\alpha}{\alpha-1}\log(1-\varepsilon)$

Crucial: $-\chi_{\alpha,\sigma}(\mathcal{N})$ is the minimum α-Rényi output entropy of a perturbed cp map $\mathcal{N}'$,

$$\mathcal{N}'(\rho) = \sigma^{\frac{1-\alpha}{2\alpha}}\mathcal{N}(\rho)\sigma^{\frac{1-\alpha}{2\alpha}}.$$

Have:

$$\log(1-\varepsilon) \leq \left(1-\frac{1}{\alpha}\right)\left(\chi_{\alpha,\sigma}(N) - \log M\right)$$

Have:

$$\log(1-\varepsilon) \leq \left(1-\frac{1}{\alpha}\right)\left(\chi_{\alpha,\sigma}(\mathcal{N}) - \log M\right)$$

Now apply this to $\mathcal{N}^{\otimes n}$, $\sigma^{\otimes n}$, and $M=2^{nR}$.

Have:

$$\log(1-\varepsilon) \leq \left(1-\frac{1}{\alpha}\right)\left(\chi_{\alpha,\sigma}(\mathcal{N}) - \log M\right)$$

Now apply this to $\mathcal{N}^{\otimes n}$, $\sigma^{\otimes n}$, and $M = 2^{nR}$.

Key observation: Sandwiched channel is $(\mathcal{N}')^{\otimes n}$, and $\mathcal{N}'$ is EB if $\mathcal{N}$ is.

Have:

$$\log(1-\varepsilon) \le \left(1-\frac{1}{\alpha}\right)\left(\chi_{\alpha,\sigma}(N) - \log M\right)$$

Now apply this to $N^{\otimes n}$, $\sigma^{\otimes n}$, and $M = 2^{nR}$.

Key observation: Sandwiched channel is $(N')^{\otimes n}$, and N' is EB if N is.

$\Rightarrow$ Additivity, $\chi_{\alpha,\sigma}(N^{\otimes n}) = n\,\chi_{\alpha,\sigma}(N)$.

(Because of identity with min output entropy of N')

[King, QIC 2003; Holevo, Russ. Math. Surveys 2006]

Get, for n uses of $\mathcal{N}$ at rate R:

$$\log(1-\varepsilon) \leq n\left(1-\frac{1}{\alpha}\right)\left(\chi_{\alpha,\sigma}(\mathcal{N}) - R\right). \qquad (\&)$$

Get, for n uses of N at rate R:

$$\log(1-\varepsilon) \le n\left(1-\frac{1}{\alpha}\right)\left(\chi_{\alpha,\sigma}(N) - R\right). \quad (\&)$$

To complete the proof, need only to observe convergence of $\chi_{\alpha,\sigma}(N)$ to $\chi(N)$;

Get, for n uses of $\mathcal{N}$ at rate R:

$$\log(1-\varepsilon) \le n\left(1-\frac{1}{\alpha}\right)\left(\chi_{\alpha,\sigma}(\mathcal{N}) - R\right). \qquad (\&)$$

To complete the proof, need only to observe convergence of $\chi_{\alpha,\sigma}(\mathcal{N})$ to $\chi(\mathcal{N})$; hence can make r.h.s. of $(\&) \le -nt$, $t > 0$, by choosing $\alpha > 1$ close enough to 1.

Get, for n uses of $\mathcal{N}$ at rate R:

$$\log(1-\varepsilon) \leq n\left(1-\frac{1}{\alpha}\right)\left(\chi_{\alpha,\sigma}(\mathcal{N}) - R\right). \qquad (\&)$$

To complete the proof, need only to observe convergence of $\chi_{\alpha,\sigma}(\mathcal{N})$ to $\chi(\mathcal{N})$; hence can make r.h.s. of $(\&) \leq -nt$, $t > 0$, by choosing $\alpha > 1$ close enough to 1.

Takes care of EB channels; $\mathcal{H}$ similar but requires another small trick (...)

Get, for n uses of N at rate R:

$$\log(1-\varepsilon) \leq n\left(1-\frac{1}{\alpha}\right)\left(\chi_{\alpha,\sigma}(N) - R\right). \qquad (\&)$$

To complete the proof, need only to observe convergence of $\chi_{\alpha,\sigma}(N)$ to $\chi(N)$; hence can make r.h.s. of $(\&) \leq -nt$, $t > 0$, by choosing $\alpha > 1$ close enough to 1.

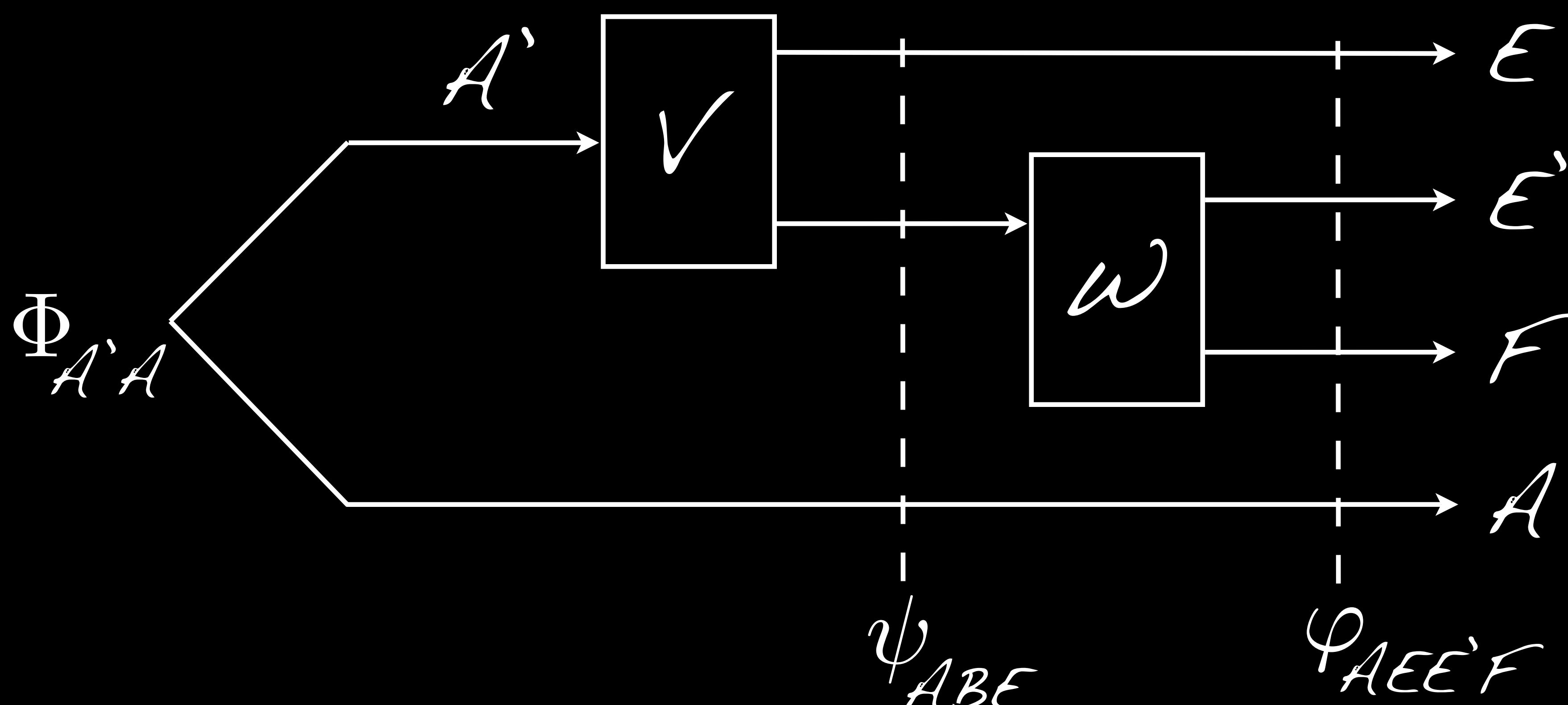
Takes care of EB channels; $\forall$ similar but requires another small trick (...) QED

B. Proof ideas for Q & P

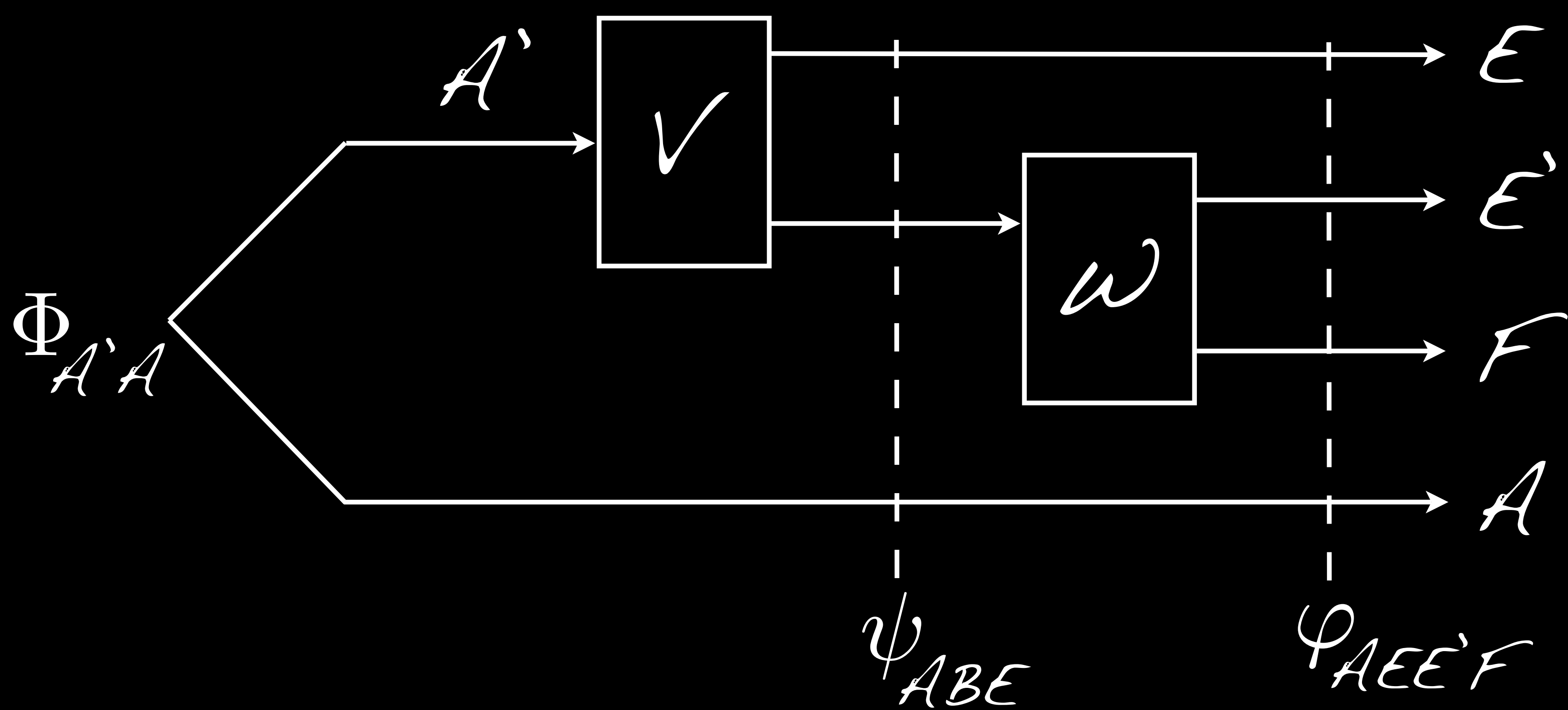
(smooth) min-entropies,
symmetrisation, de Finetti theorem,
AEP

Ideas: (smooth) min-entropies,
symmetrisation, de Finetti theorem,
AEP

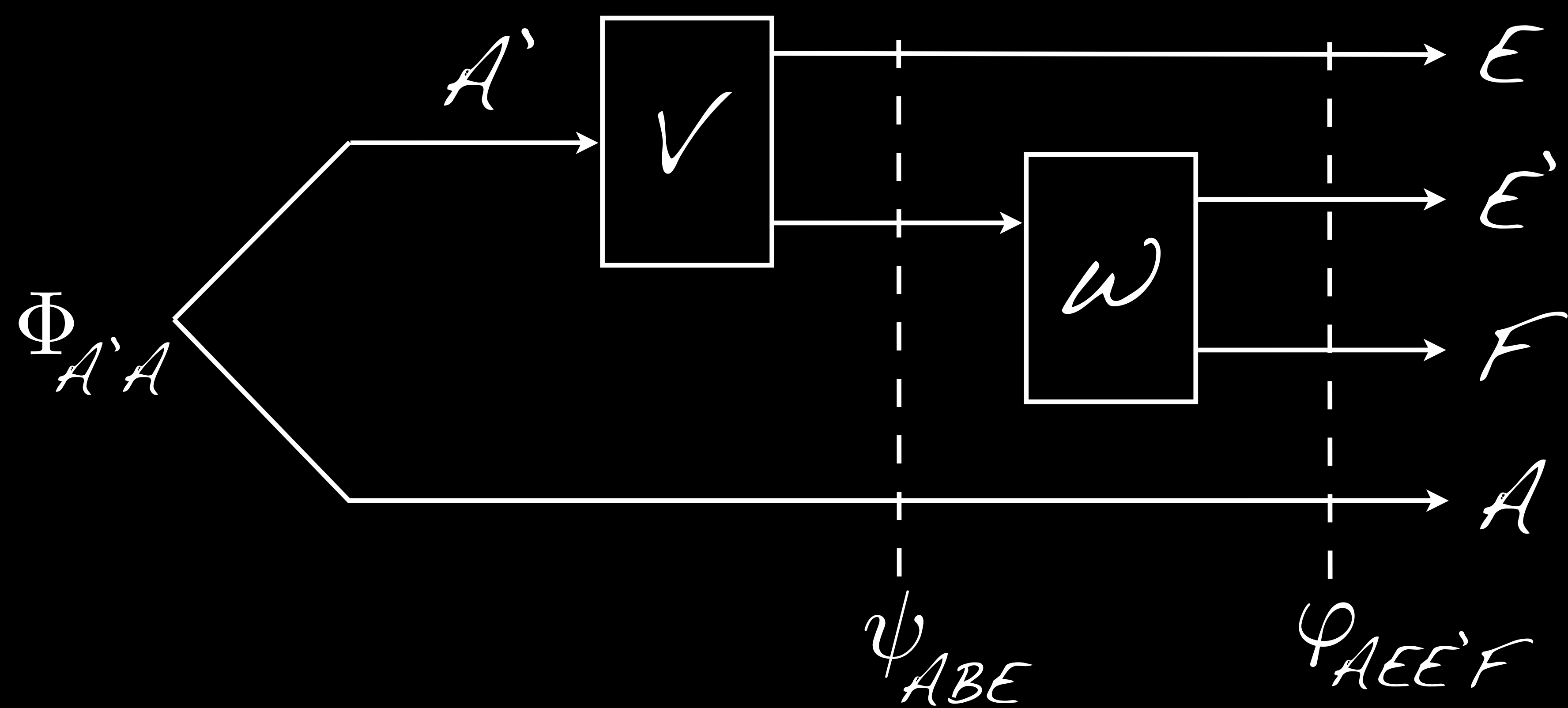
1) Use code - for simplicity subspace - with maximally entangled state Φ of k qubits:



Maximally entangled state Φ of k qubits:

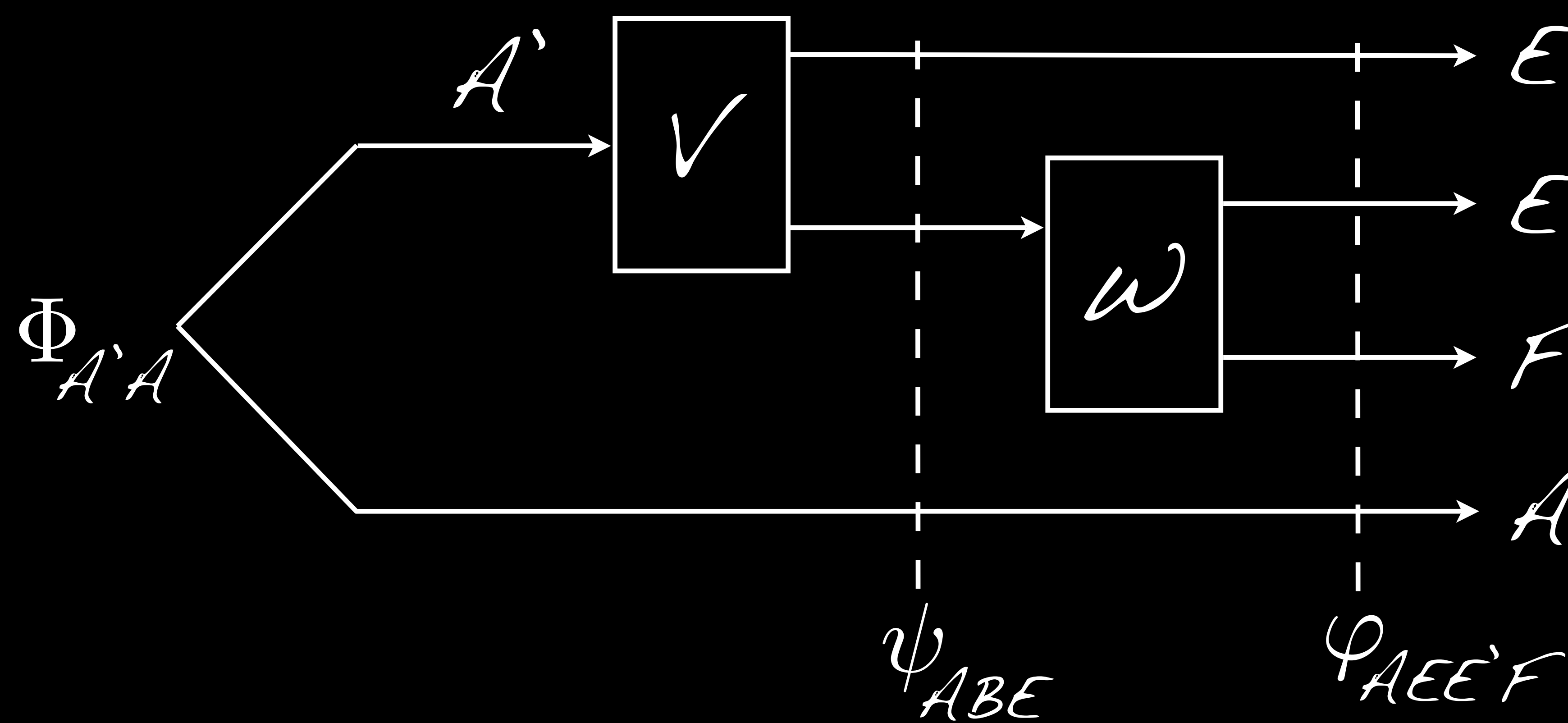


Maximally entangled state Φ of k qubits:



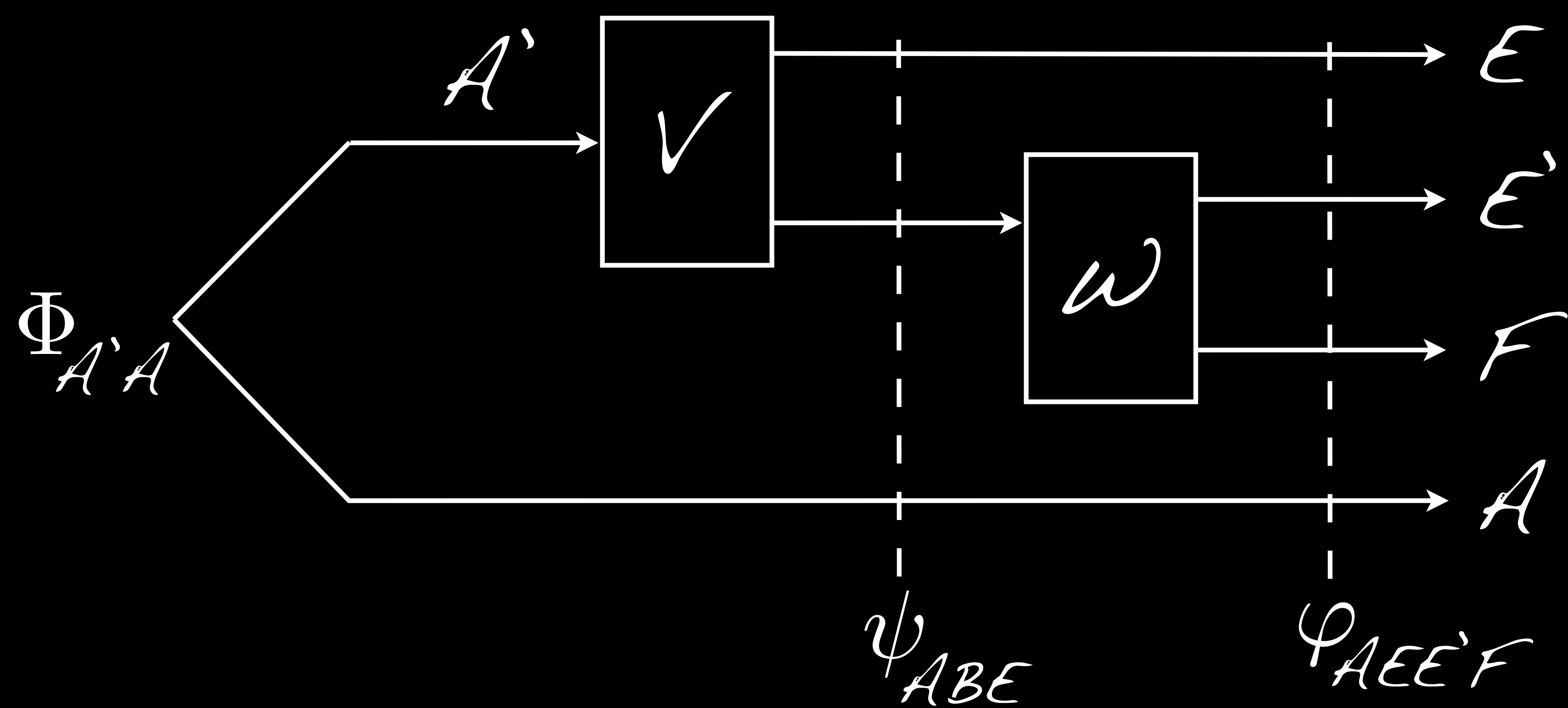
$$k \leq H_{min}^{\epsilon}(A|E)$$

Maximally entangled state Φ of k qubits:



$$k \leq H_{min}^{\epsilon}(A|E) = -H_{max}^{\epsilon}(A|E'F)$$

Maximally entangled state Φ of k qubits:



$$k \leq H^{\epsilon}_{min}(A|E) = -H^{\epsilon}_{max}(A|E'F)$$

[For min-entropy calculus, consult
R. Renner, PhD thesis, quant-ph/0512258
& M. Tomamichel, PhD thesis, arXiv:1203.2142]

$$
\begin{aligned}
k &\leq H_{min}^{\epsilon}(A|E) \\
&= -H_{max}^{\epsilon}(A|E'F)
\end{aligned}
$$

$$k \leq H^{\epsilon}_{min}(A|E)$$
$$= -H^{\epsilon}_{max}(A|E'F)$$

[Cf. also Buscemi/Datta, IEEE-IT 56(3), 2010; Datta/Hsieh, 1103.1135]

$$k \le H_{min}^{\epsilon}(A|E)$$
$$= -H_{max}^{\epsilon}(A|E'F)$$

Note: If we knew that for n channel uses, the maximum min-entropy is attained on a tensor product input, we'd be done by AEP (= asymptotic equipartition property)...

$$
\begin{aligned}
k &\le H_{min}^{\epsilon}(A|E) \\
&= -H_{max}^{\epsilon}(A|E'F) \\
&\le H_{max}^{\lambda}(F|E') - H_{max}^{\delta}(AF|E') + O(1)
\end{aligned}
$$

$$
\begin{aligned}
k &\leq H_{min}^{\epsilon}(A|E) \\
&= -H_{max}^{\epsilon}(A|E'F) \\
&\leq H_{max}^{\lambda}(F|E') - H_{max}^{\delta}(AF|E') + O(1)
\end{aligned}
$$

Chain rule, $\delta = \epsilon + 3\lambda$.

$$k \leq H_{min}^{\epsilon}(A|E)$$
$$= -H_{max}^{\epsilon}(A|E'F)$$
$$\leq H_{max}^{\lambda}(F|E') - H_{max}^{\delta}(AF|E') + O(1)$$

Chain rule, $\delta = \epsilon + 3\lambda$.

$$\leq H_{max}^{\lambda}(F|E') + O(1)$$

$$k \leq H_{min}^{\epsilon}(A|E)$$

$$= -H_{max}^{\epsilon}(A|E'F)$$

$$\leq H_{max}^{\lambda}(F|E') - H_{max}^{\delta}(AF|E') + O(1)$$

Chain rule, $\delta = \epsilon + 3\lambda$.

$$\leq H_{max}^{\lambda}(F|E') + O(1)$$

...if $\delta < 0.707$, by inequality H_{min} vs. H_{max}, and using symmetry between E and E'...

2) For n channel uses, have restricted concavity of H_{max}^{λ}:

$$k \leq H_{max}^{\lambda}(F^n|E'^n) + O(1)$$

2) For n channel uses, have restricted concavity of H_{max}^{λ}:

$$k \leq H_{max}^{\lambda}(F^{n}|E'^{n}) + O(1)$$

$$\leq H_{max}^{\lambda'}(F^{n}|E'^{n})_{\rho_A^{(n)}} + O(1)$$

2) For n channel uses, have restricted concavity of H_{max}^{λ}:

$$k \leq H_{max}^{\lambda}(F^n | E'^n) + O(1)$$

$$\leq H_{max}^{\lambda'}(F^n | E'^n)_{\rho_A^{(n)}} + O(1)$$

W.r.t. a permutation symmetric input state and $\lambda' = \lambda / \sqrt{2}$

2) For n channel uses, have restricted concavity of H_{max}^{λ}:

$$k \le H_{max}^{\lambda}(F^n|E'^n) + O(1)$$

$$\le H_{max}^{\lambda'}(F^n|E'^n)_{\rho_A^{(n)}} + O(1)$$

3) By de Finetti theorem [R. Renner, PhD thesis, quant-ph/0512258]:

$$k \le \max_{\rho_A} H_{max}^{\lambda''}(F^n|E'^n)_{\rho^{\otimes n}} + o(n)$$

4) By AEP (asymptotic equipartition property) [M. Tomamichel, arXiv:1203.2142]:

$$k \leq \max_{\rho_A} H_{max}^{\lambda''}(F^n | E'^n)_{\rho^{\otimes n}} + o(n)$$

4) By AEP (asymptotic equipartition property) [M. Tomamichel, arXiv:1203.2142]:

$$k \leq \max_{\rho_A} H_{max}^{\lambda''}(F^n|E'^n)_{\rho^{\otimes n}} + o(n)$$

$$= \max_{\rho_A} n\, S(F|E')_\rho + o(n)$$

4) By AEP (asymptotic equipartition property) [M. Tomamichel, arXiv:1203.2142]:

$$k \leq \max_{\rho_A} H_{max}^{\lambda''}(F^n|E'^n)_{\rho^{\otimes n}} + o(n)$$

$$= \max_{\rho_A} n\, S(F|E')_\rho + o(n)$$

$$= n\, Q^{(1)}(N) + o(n)$$

4) By AEP (asymptotic equipartition property) [M. Tomamichel, arXiv:1203.2142]:

$$k \leq \max_{\rho_A} H_{max}^{\lambda''}(F^n|E'^n)_{\rho^{\otimes n}} + o(n)$$

$$= \max_{\rho_A} n\, S(F|E')_\rho + o(n)$$

$$= n\, Q^{(1)}(N) + o(n)$$

(by the degradability argument)

4) By AEP (asymptotic equipartition property) [M. Tomamichel, arXiv:1203.2142]:

$$k \le \max_{\rho_A} H_{max}^{\lambda''}(F^n | E'^n)_{\rho^{\otimes n}} + o(n)$$

$$= \max_{\rho_A} n\, S(F|E')_\rho + o(n)$$

$$= n\, Q^{(1)}(N) + o(n)$$

(by the degradability argument)

QED