Identifying the Information Gain of a Quantum Measurement

Mario Berta, Joseph M. Renes, Mark M. Wilde - full version IEEE Transactions on Information Theory, vol. 60, no. 12, pages 7987-8006, 2014

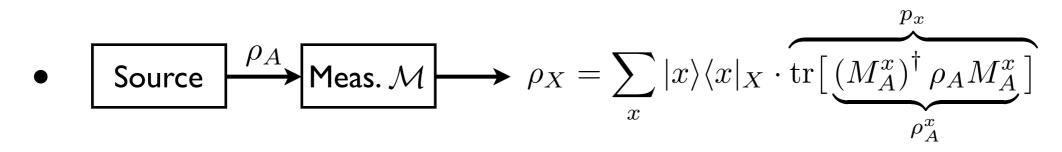
- Setup making the question precise (information-theoretically)
- Main result answer
- Proof ideas
- Conclusions

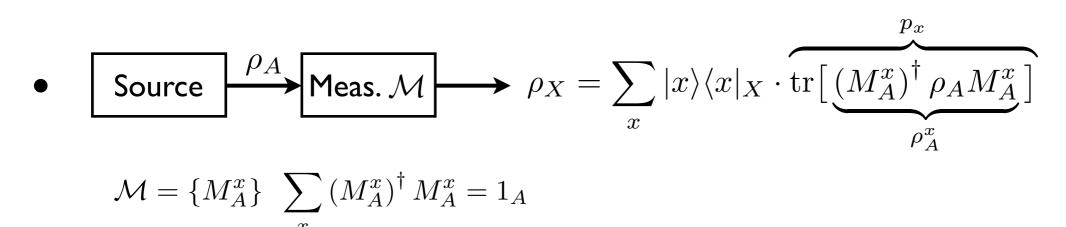
- Setup making the question precise (information-theoretically)
- Main result answer
- Proof ideas
- Conclusions

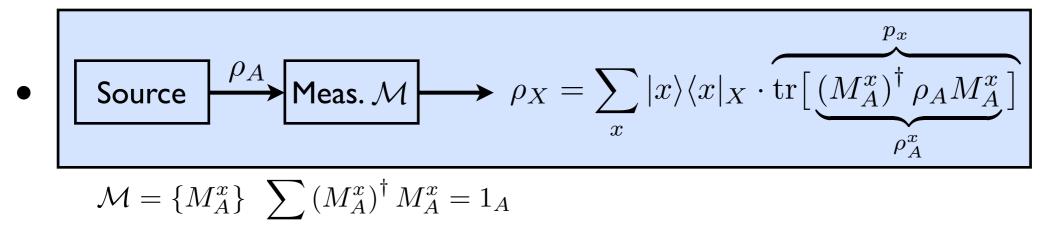
- Setup making the question precise (information-theoretically)
- Main result answer
- Proof ideas
- Conclusions

- Setup making the question precise (information-theoretically)
- Main result answer
- Proof ideas
- Conclusions

- Setup making the question precise (information-theoretically)
- Main result answer
- Proof ideas
- Conclusions







• Source
$$\rho_A$$
 Meas. \mathcal{M} $\rightarrow \rho_X = \sum_x |x\rangle \langle x|_X \cdot \operatorname{tr}\left[\underbrace{(M_A^x)^{\dagger} \rho_A M_A^x}_{\rho_A^x}\right]$
 $\mathcal{M} = \{M_A^x\} \sum_x (M_A^x)^{\dagger} M_A^x = 1_A$

• First guess: amount of data acquired = information gain? $I(\mathcal{M}) \stackrel{?}{=} \log(\#x)$

• Source
$$\rho_A$$
 Meas. \mathcal{M} $\rightarrow \rho_X = \sum_x |x\rangle \langle x|_X \cdot \operatorname{tr}\left[\underbrace{(M_A^x)^{\dagger} \rho_A M_A^x}{\rho_A^x}\right]$
 $\mathcal{M} = \{M_A^x\} \sum_x (M_A^x)^{\dagger} M_A^x = 1_A$

• First guess: amount of data acquired = information gain? $I(\mathcal{M}) \stackrel{?}{=} \log(\#x)$

Or for asymptotic independent and identically distributed (iid) setup: $I(\mathcal{M}) \stackrel{?}{=} H(X)_{\rho} \left(= -\sum_{x} p_{x} \log p_{x} \right)$

• Source
$$\rho_A$$
 Meas. \mathcal{M} $\rightarrow \rho_X = \sum_x |x\rangle \langle x|_X \cdot \operatorname{tr}\left[\underbrace{(M_A^x)^{\dagger} \rho_A M_A^x}_{\rho_A^x}\right]$
 $\mathcal{M} = \{M_A^x\} \sum (M_A^x)^{\dagger} M_A^x = 1_A$

• First guess: amount of data acquired = information gain? $I(\mathcal{M}) \stackrel{?}{=} \log(\#x)$

Or for asymptotic independent and identically distributed (iid) setup: $I(\mathcal{M}) \stackrel{?}{=} H(X)_{\rho} \left(= -\sum_{x} p_{x} \log p_{x} \right) \qquad \underbrace{\text{But:}}_{X} M_{A}^{x} \neq |x\rangle \qquad \qquad \underbrace{\text{But:}}_{X} M_{A}^{x} \neq |x\rangle$

• Source
$$\rho_A$$
 Meas. $\mathcal{M} \longrightarrow \rho_X = \sum_x |x\rangle \langle x|_X \cdot \operatorname{tr}\left[\underbrace{(M_A^x)^{\dagger} \rho_A M_A^x}_{\rho_A^x}\right]$
 $\mathcal{M} = \{M_A^x\} \sum_x (M_A^x)^{\dagger} M_A^x = 1_A$

• First guess: amount of data acquired = information gain?
$$I(\mathcal{M}) \stackrel{?}{=} \log(\#x)$$

Or for asymptotic independent and identically distributed (iid) setup: $I(\mathcal{M}) \stackrel{?}{=} H(X)_{\rho} \left(= -\sum_{x} p_{x} \log p_{x} \right) \qquad \underbrace{\text{But:}}_{x} M_{A}^{x} \neq |x\rangle \qquad \qquad \underbrace{\text{But:}}_{x} M_{A}^{x} \neq |x\rangle$

• Intuitive reasoning [1]: reduction of entropy? $I(\mathcal{M}) \stackrel{?}{=} H(A)_{\rho} - \sum_{x} p_{x} \cdot H(A)_{\rho^{x}}$

 $H(A)_{\rho} = -\mathrm{tr}[\rho_A \log \rho_A]$

• Source
$$\rho_A$$
 Meas. \mathcal{M} $\rightarrow \rho_X = \sum_x |x\rangle \langle x|_X \cdot \operatorname{tr}\left[\underbrace{(M_A^x)^{\dagger} \rho_A M_A^x}{\rho_A^x}\right]$

$$\mathcal{M} = \{M_A^x\} \quad \sum_x \left(M_A^x\right)^\dagger M_A^x = \mathbf{1}_A$$

First guess: amount of data acquired = information gain? $I(\mathcal{M}) \stackrel{?}{=} \log(\#x)$ Or for asymptotic independent and identically distributed (iid) setup:

$$I(\mathcal{M}) \stackrel{?}{=} H(X)_{\rho} \left(= -\sum_{x} p_{x} \log p_{x} \right) \qquad \underbrace{\text{But:}}_{X} M_{A}^{x} \neq |x\rangle \qquad \underbrace{\text{But:}}_{\rho} M^{\otimes n}(\rho_{A}^{\otimes n})$$

- •
- Intuitive reasoning [1]: reduction of entropy? $I(\mathcal{M}) \stackrel{?}{=} H(A)_{\rho} \sum_{x} p_{x} \cdot H(A)_{\rho^{x}}$ Non-negativity [2]: $H(A)_{\rho} \sum p_{x} \cdot H(A)_{\rho^{x}} \ge 0$ $H(A)_{\rho} = -\operatorname{tr}[\rho_{A} \log \rho_{A}]$ **Non-negativity [2]:** $H(A)_{\rho} - \sum_{x} p_{x} \cdot H(A)_{\rho^{x}} \ge 0$

• Source
$$\rho_A$$
 Meas. $\mathcal{M} \longrightarrow \rho_X = \sum_x |x\rangle \langle x|_X \cdot \operatorname{tr}\left[(M_A^x)^{\dagger} \rho_A M_A^x\right]_{\rho_A^x}$
 $\mathcal{M} = \{M_A^x\} \sum_x (M_A^x)^{\dagger} M_A^x = 1_A$

• First guess: amount of data acquired = information gain?
$$I(\mathcal{M}) \stackrel{?}{=} \log(\#x)$$

- Intuitive reasoning [1]: reduction of entropy? $I(\mathcal{M}) \stackrel{?}{=} H(A)_{\rho} \sum_{x} p_{x} \cdot H(A)_{\rho^{x}}$ Non-negativity [2]: $H(A)_{\rho} \sum_{x} p_{x} \cdot H(A)_{\rho^{x}} \ge 0$ $H(A)_{\rho} = -\operatorname{tr}[\rho_{A} \log \rho_{A}]$ **Non-negativity [2]:** $H(A)_{\rho} - \sum p_x \cdot H(A)_{\rho^x} \ge 0$
- Information and disturbance in quantum measurements [3]:

$$I(\mathcal{M}) \stackrel{?}{=} I(X:R)_{\omega} (= H(X)_{\omega} + H(R)_{\omega} - H(XR)_{\omega}) \ge 0$$
$$\omega_{XR} = \sum_{x} |x\rangle \langle x|_{X} \otimes \operatorname{tr}_{A} \left[\left(M_{A}^{x} \otimes 1_{R} \right)^{\dagger} \underbrace{\rho_{AR}}_{\text{pure}} \left(M_{A}^{x} \otimes 1_{R} \right) \right]$$

[1] Groenewold, Int. J. of Th. Phys. (1971)

[2] Ozawa, J. of Math. Phys. (1986)

[3] Buscemi et al., Phys. Rev. Lett. (2008)

• Source
$$\rho_A$$
 Meas. \mathcal{M} $\rightarrow \rho_X = \sum_x |x\rangle \langle x|_X \cdot \operatorname{tr}\left[\underbrace{(M_A^x)^{\dagger} \rho_A M_A^x}_{\rho_A^x}\right]$
 $\mathcal{M} = \{M_A^x\} \sum_x (M_A^x)^{\dagger} M_A^x = 1_A$

• First guess: amount of data acquired = information gain?
$$I(\mathcal{M}) \stackrel{?}{=} \log(\#x)$$

Or for asymptotic independent and identically distributed (iid) setup: $I(\mathcal{M}) \stackrel{?}{=} H(X)_{\rho} \left(= -\sum_{x} p_{x} \log p_{x} \right) \qquad \underbrace{\text{But:}}_{X} M_{A}^{x} \neq |x\rangle$

- Intuitive reasoning [1]: reduction of entropy? $I(\mathcal{M}) \stackrel{?}{=} H(A)_{\rho} \sum_{x} p_{x} \cdot H(A)_{\rho^{x}}$
- Non-negativity [2]: $H(A)_{\rho} \sum_{x} p_x \cdot H(A)_{\rho^x} \ge 0$ $H(A)_{\rho} = -\operatorname{tr}[\rho_A \log \rho_A]$
- Information and disturbance in quantum measurements [3]:

$$I(\mathcal{M}) \stackrel{?}{=} I(X:R)_{\omega} (= H(X)_{\omega} + H(R)_{\omega} - H(XR)_{\omega}) \ge 0$$
$$\omega_{XR} = \sum_{x} |x\rangle \langle x|_{X} \otimes \operatorname{tr}_{A} \left[\left(M_{A}^{x} \otimes 1_{R} \right)^{\dagger} \underbrace{\rho_{AR}}_{\text{pure}} \left(M_{A}^{x} \otimes 1_{R} \right) \right]$$

• Precise information-theoretic meaning in [4]!

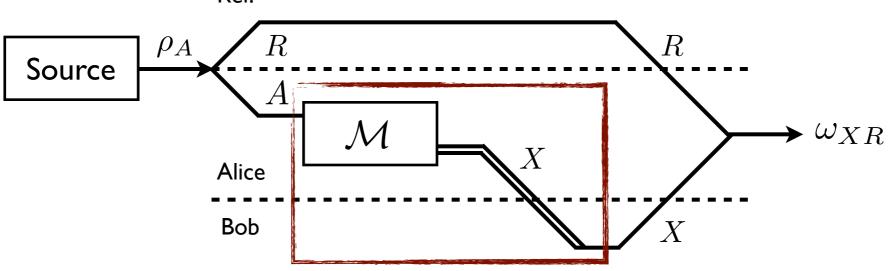
[1] Groenewold, Int. J. of Th. Phys. (1971)[4] Winter, Comm. in Math. Phys. (2004)

[2] Ozawa, J. of Math. Phys. (1986)

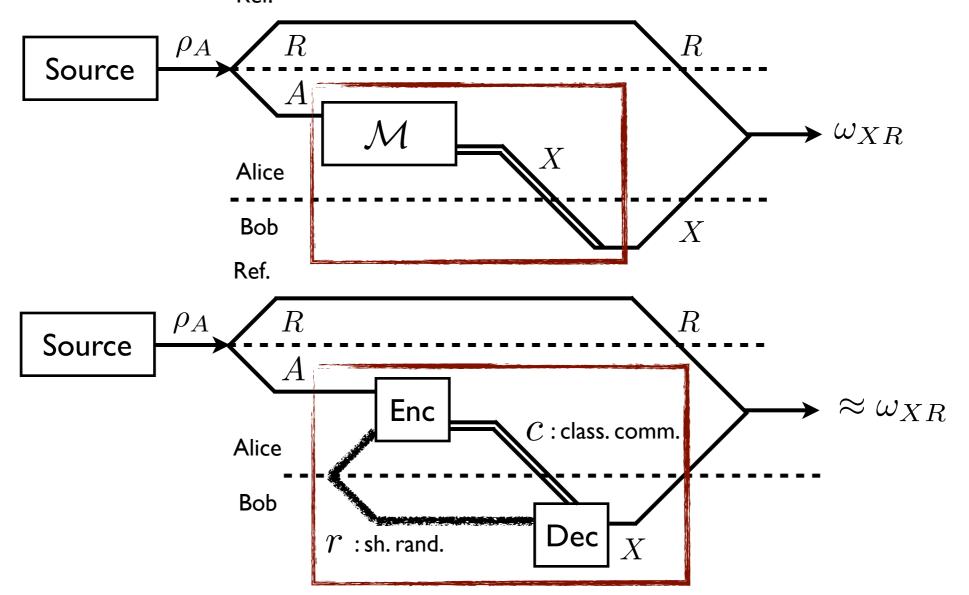
[3] Buscemi et al., Phys. Rev. Lett. (2008)

• Only output data that is correlated to the input should be counted as information gain! Information-theoretic idea of **simulating** the measurement [4]:

 Only output data that is correlated to the input should be counted as information gain! Information-theoretic idea of simulating the measurement [4]: Ref.

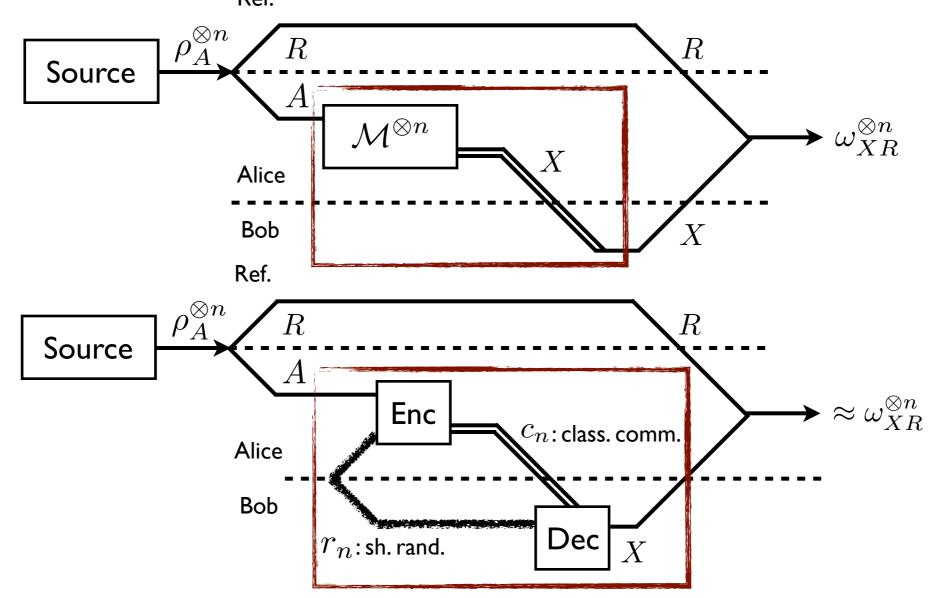


 Only output data that is correlated to the input should be counted as information gain! Information-theoretic idea of simulating the measurement [4]: Ref.

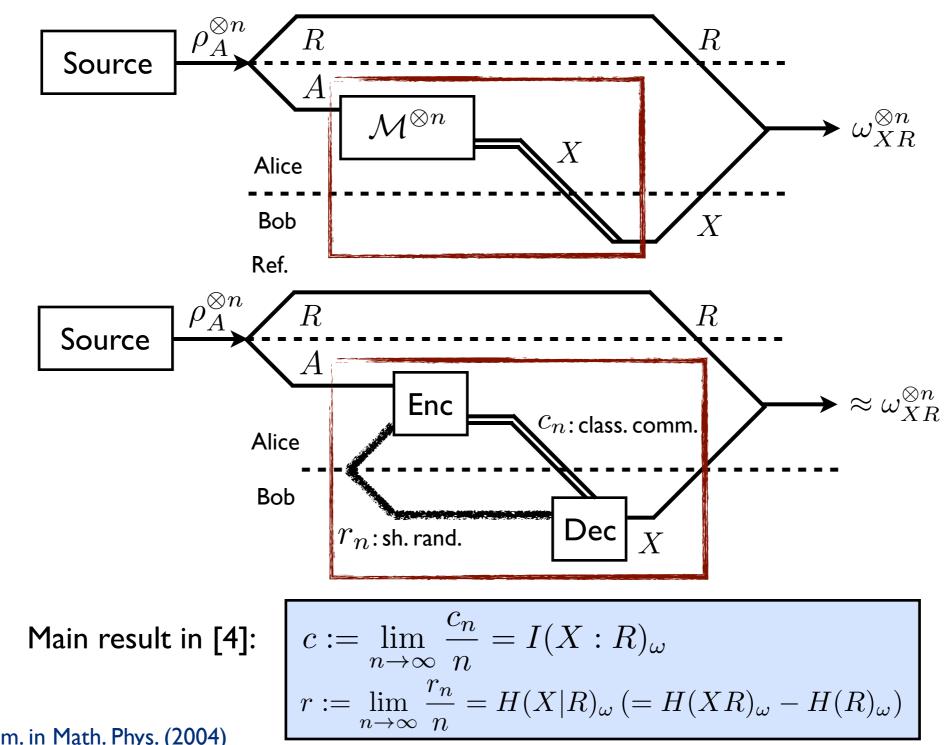


[4] Winter, Comm. in Math. Phys. (2004)

 Only output data that is correlated to the input should be counted as information gain! Information-theoretic idea of simulating the measurement [4]: Ref.



 Only output data that is correlated to the input should be counted as information gain! Information-theoretic idea of simulating the measurement [4]: Ref.



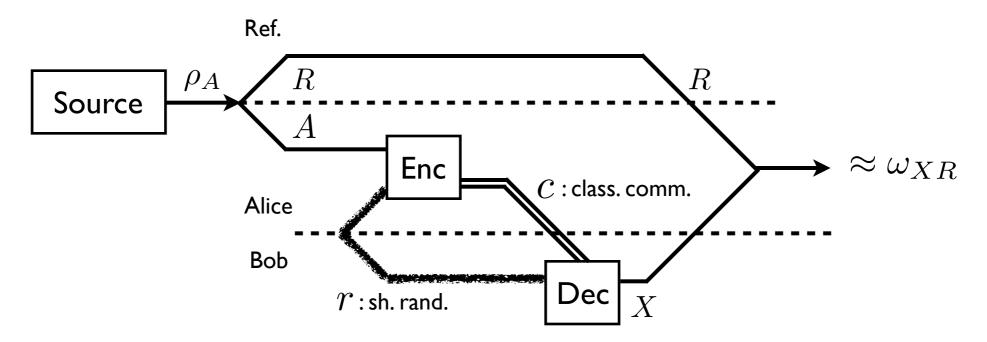
[4] Winter, Comm. in Math. Phys. (2004)

• An easy example for [4]:

$$\rho_A = \frac{1}{2} \mathbf{1}_A \qquad \mathcal{M} = \left\{ \frac{1}{2} |0\rangle \langle 0|, \frac{1}{2} |1\rangle \langle 1|, \frac{1}{2} |+\rangle \langle +|, \frac{1}{2} |-\rangle \langle -| \right\}$$

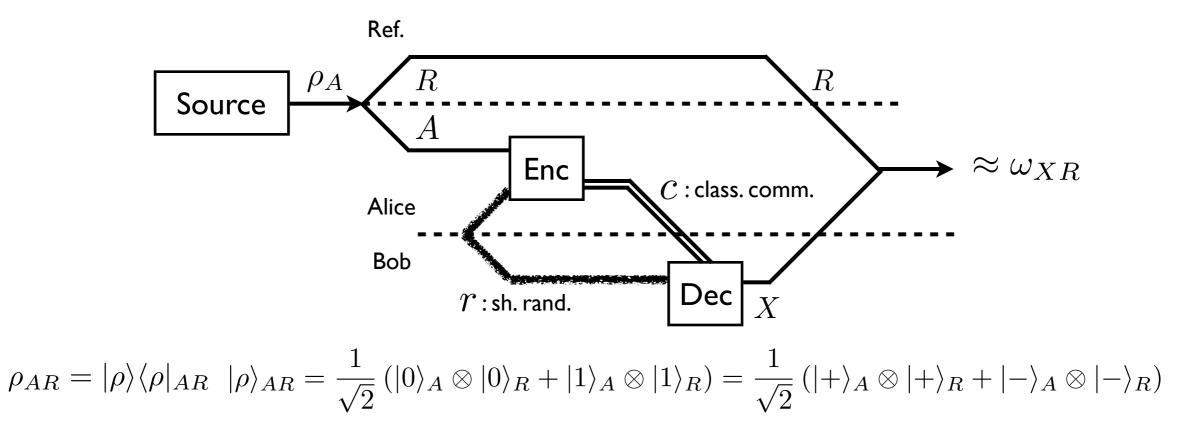
• An easy example for [4]:

$$\rho_A = \frac{1}{2} \mathbf{1}_A \qquad \mathcal{M} = \left\{ \frac{1}{2} |0\rangle \langle 0|, \frac{1}{2} |1\rangle \langle 1|, \frac{1}{2} |+\rangle \langle +|, \frac{1}{2} |-\rangle \langle -| \right\}$$



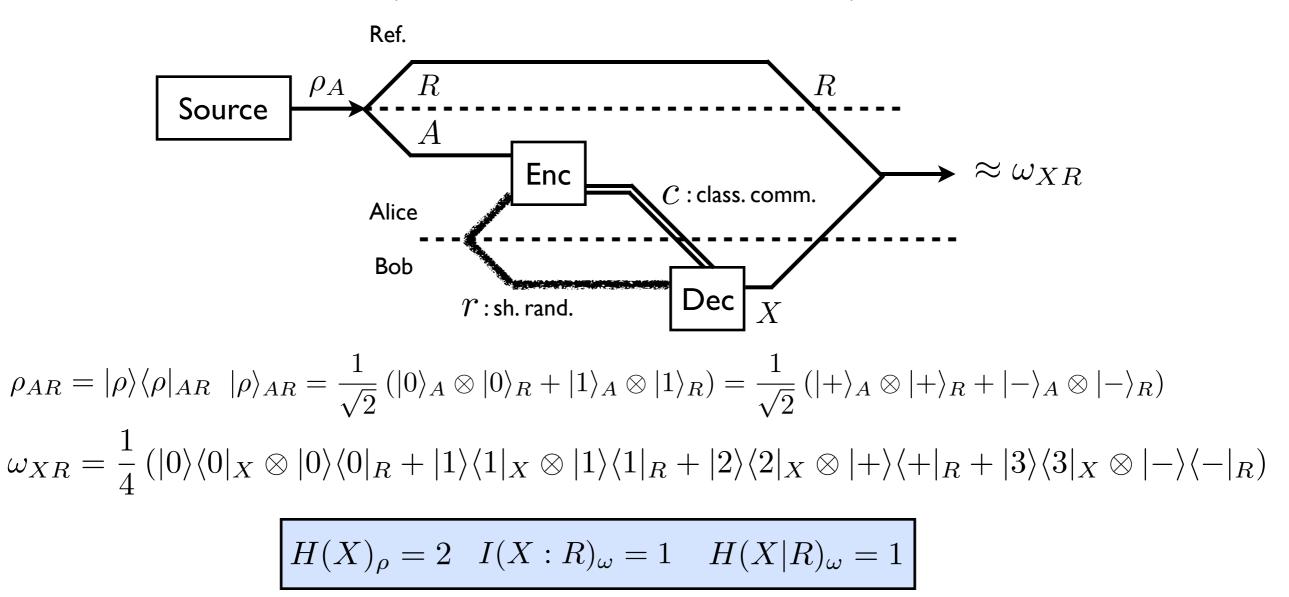
• An easy example for [4]:

$$\rho_A = \frac{1}{2} \mathbf{1}_A \qquad \mathcal{M} = \left\{ \frac{1}{2} |0\rangle \langle 0|, \frac{1}{2} |1\rangle \langle 1|, \frac{1}{2} |+\rangle \langle +|, \frac{1}{2} |-\rangle \langle -| \right\}$$



• An easy example for [4]:

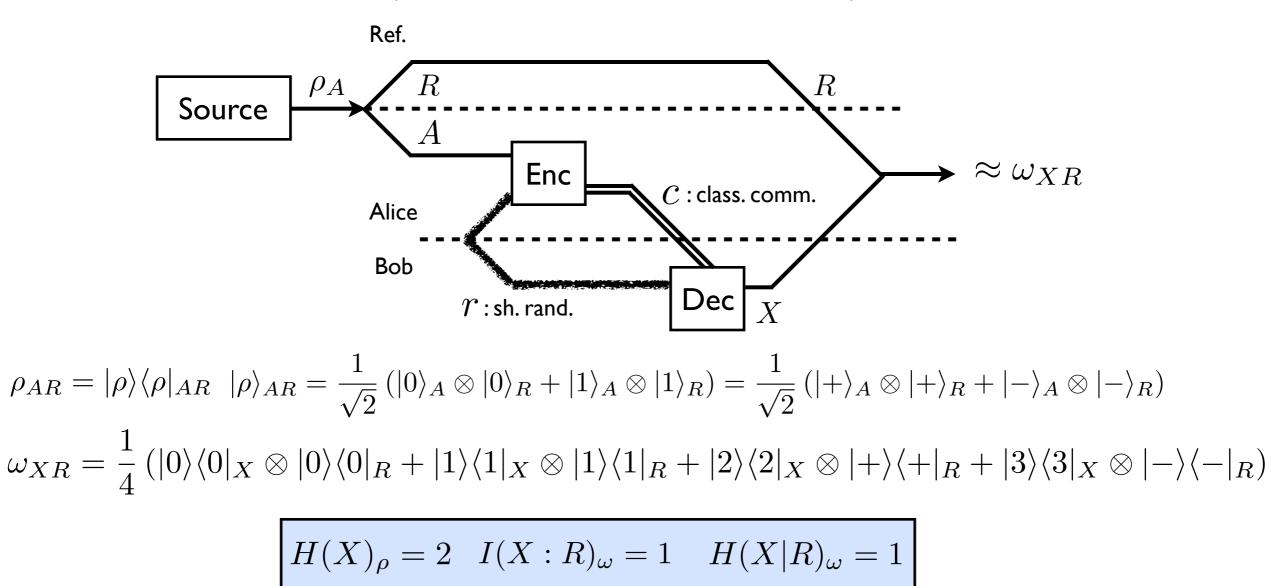
$$\rho_A = \frac{1}{2} \mathbf{1}_A \qquad \mathcal{M} = \left\{ \frac{1}{2} |0\rangle \langle 0|, \frac{1}{2} |1\rangle \langle 1|, \frac{1}{2} |+\rangle \langle +|, \frac{1}{2} |-\rangle \langle -| \right\}$$



[4] Winter, Comm. in Math. Phys. (2004)

• An easy example for [4]:

$$\rho_A = \frac{1}{2} \mathbf{1}_A \qquad \mathcal{M} = \left\{ \frac{1}{2} |0\rangle \langle 0|, \frac{1}{2} |1\rangle \langle 1|, \frac{1}{2} |+\rangle \langle +|, \frac{1}{2} |-\rangle \langle -| \right\}$$



• Apply according to shared randomness: $\sigma_X = \left\{\frac{1}{2}|0\rangle\langle 0|, \frac{1}{2}|1\rangle\langle 1|\right\} \sigma_Z = \left\{\frac{1}{2}|+\rangle\langle +|, \frac{1}{2}|-\rangle\langle -|\right\}$ [4] Winter, Comm. in Math. Phys. (2004)

- Setup making the question precise (information-theoretically)
- Main result answer
- Proof ideas
- Conclusions

• Weaknesses of [4]: (i) only asymptotic iid (ii) only for known fixed input (protocol depends on input).

- Weaknesses of [4]: (i) only asymptotic iid (ii) only for known fixed input (protocol depends on input).
- We generalize (ii) to **arbitrarily varying sources** and even **entangled sources** (measurement simulation on general inputs):

$$\rho_A^{\otimes n}$$
 vs. $\rho_A^1 \otimes \rho_A^2 \otimes \cdots \otimes \rho_A^n$ vs. $\rho_{A^n}^n$

- Weaknesses of [4]: (i) only asymptotic iid (ii) only for known fixed input (protocol depends on input).
- We generalize (ii) to **arbitrarily varying sources** and even **entangled sources** (measurement simulation on general inputs):

$$\rho_A^{\otimes n} \text{ vs. } \rho_A^1 \otimes \rho_A^2 \otimes \cdots \otimes \rho_A^n \text{ vs. } \rho_{A^n}^n$$

- Weaknesses of [4]: (i) only asymptotic iid (ii) only for known fixed input (protocol depends on input).
- We generalize (ii) to **arbitrarily varying sources** and even **entangled sources** (measurement simulation on general inputs):

- Weaknesses of [4]: (i) only asymptotic iid (ii) only for known fixed input (protocol depends on input).
- We generalize (ii) to **arbitrarily varying sources** and even **entangled sources** (measurement simulation on general inputs):

• Hence, we determined the **information gain of quantum measurements**:

$$I(\mathcal{M}) = \max_{\rho} I(X:R)_{\omega}$$

[4] Winter, Comm. in Math. Phys. (2004)

- Weaknesses of [4]: (i) only asymptotic iid (ii) only for known fixed input (protocol depends on input).
- We generalize (ii) to **arbitrarily varying sources** and even **entangled sources** (measurement simulation on general inputs):

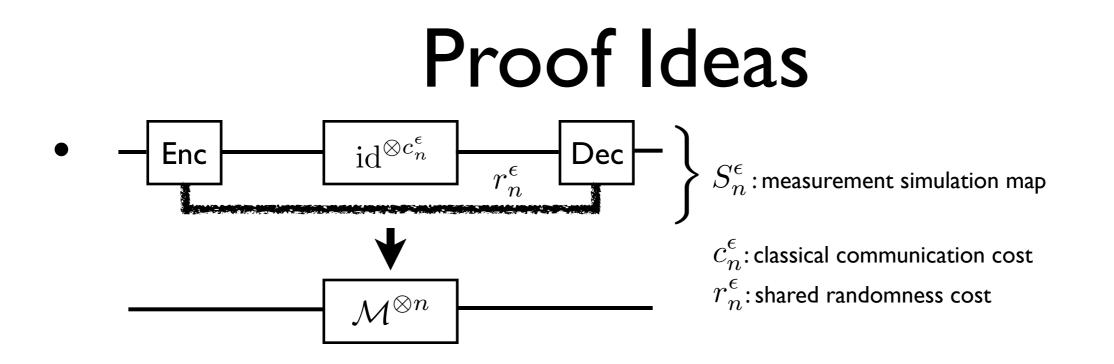
• Hence, we determined the **information gain of quantum measurements**:

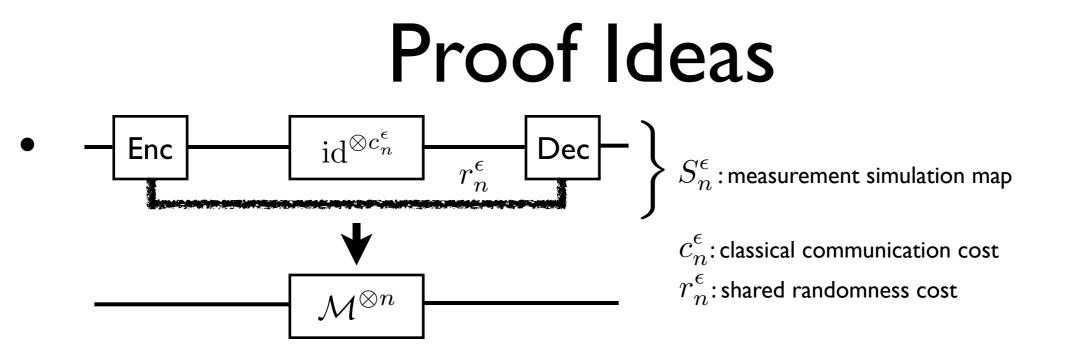
$$I(\mathcal{M}) = \max_{\rho} I(X:R)_{\omega}$$

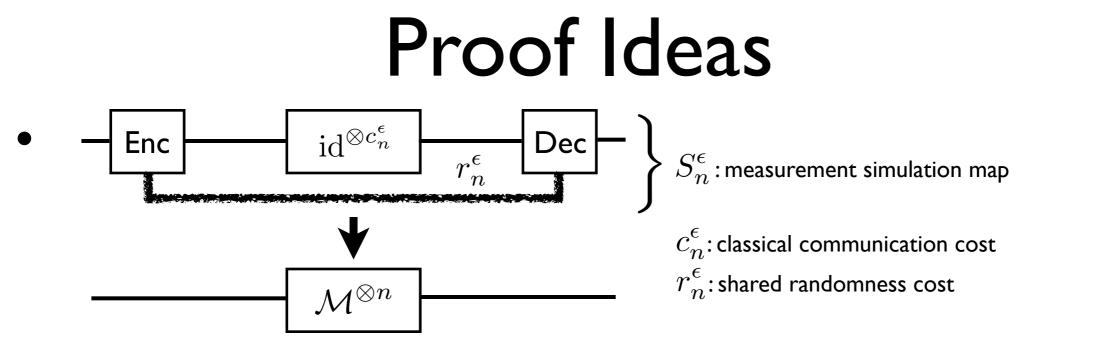
• By proving this, we also generalize (i) to the one-shot case: $c^{\epsilon} \approx I_{\max}^{\epsilon}(X:R)_{\omega}$

[4] Winter, Comm. in Math. Phys. (2004)

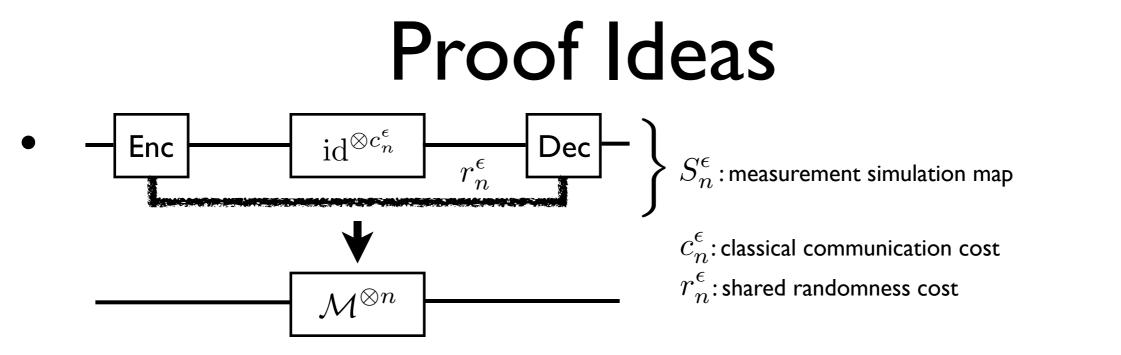
- Setup making the question precise (information-theoretically)
- Main result answer
- Proof ideas
- Conclusions







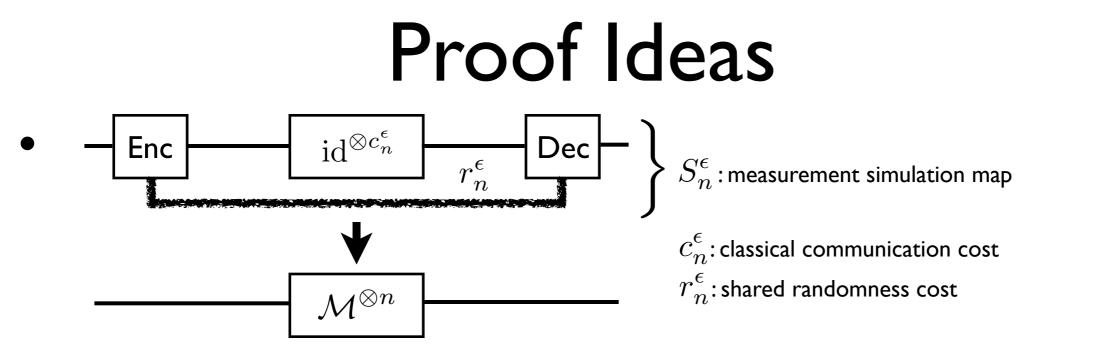
- In [4] for fixed iid input ρ_{AR} : $\left\| \left(\mathcal{M}_A^{\otimes n} \otimes \operatorname{id}_R^{\otimes n} \right) \left(\rho_{AR}^{\otimes n} \right) \left(S_n^{\epsilon} \otimes \operatorname{id}_R^{\otimes n} \right) \left(\rho_{AR}^{\otimes n} \right) \right\|_1 \leq \epsilon$
- Now: $\left\| \mathcal{M}_{A}^{\otimes n} S_{n}^{\epsilon} \right\|_{\diamond} \leq \epsilon$ $\left\| \mathcal{M} \right\|_{\diamond} = \sup_{k \in \mathbb{N}} \sup_{\|\sigma\|_{1} \leq 1} \| (\mathcal{M} \otimes \mathrm{id}_{k})(\sigma) \|_{1}$



• Now:
$$\left\| \mathcal{M}_{A}^{\otimes n} - S_{n}^{\epsilon} \right\|_{\diamond} \leq \epsilon$$
 $\left\| \mathcal{M} \right\|_{\diamond} = \sup_{k \in \mathbb{N}} \sup_{\|\sigma\|_{1} \leq 1} \left\| (\mathcal{M} \otimes \mathrm{id}_{k})(\sigma) \right\|_{1}$

• **Post-selection technique** for quantum channels [5]:

$$\left\|\mathcal{M}_{A}^{\otimes n} - S_{n}^{\epsilon}\right\|_{\diamond} \leq \operatorname{poly}(n) \cdot \left\|\left(\left(\mathcal{M}_{A}^{\otimes n} - S_{n}^{\epsilon}\right) \otimes \operatorname{id}_{R'}\right)\left(\zeta_{AR'}^{n}\right)\right\|_{1}$$

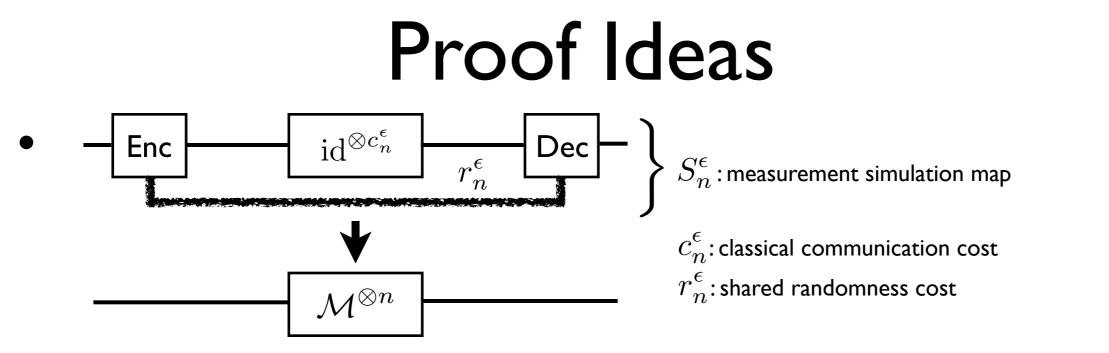


• Now:
$$\|\mathcal{M}_A^{\otimes n} - S_n^{\epsilon}\|_{\diamond} \leq \epsilon$$
 $\|\mathcal{M}\|_{\diamond} = \sup_{k \in \mathbb{N}} \sup_{\|\sigma\|_1 \leq 1} \|(\mathcal{M} \otimes \mathrm{id}_k)(\sigma)\|_1$

• **Post-selection technique** for quantum channels [5]:

$$\left\|\mathcal{M}_{A}^{\otimes n} - S_{n}^{\epsilon}\right\|_{\diamond} \leq \operatorname{poly}(n) \cdot \left\|\left(\left(\mathcal{M}_{A}^{\otimes n} - S_{n}^{\epsilon}\right) \otimes \operatorname{id}_{R'}\right)\left(\zeta_{AR'}^{n}\right)\right\|_{1}\right\|_{1}$$

<u>But</u>: $\zeta_{AR'}^n$ purification of a special de Finetti state (a state which consists of n iid copies of a state on a single subsystem), i.e., no iid structure and [4] not applicable!



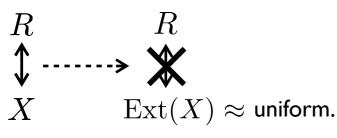
• Now:
$$\|\mathcal{M}_A^{\otimes n} - S_n^{\epsilon}\|_{\diamond} \leq \epsilon$$
 $\|\mathcal{M}\|_{\diamond} = \sup_{k \in \mathbb{N}} \sup_{\|\sigma\|_1 \leq 1} \|(\mathcal{M} \otimes \mathrm{id}_k)(\sigma)\|_1$

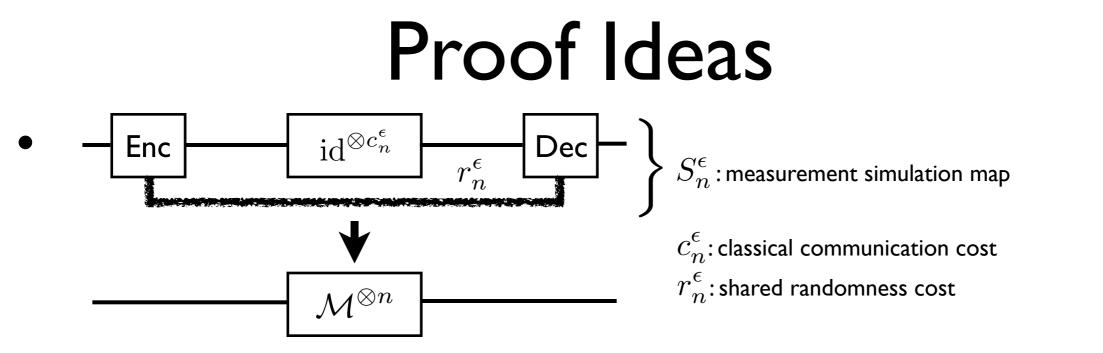
• **Post-selection technique** for quantum channels [5]:

$$\left\|\mathcal{M}_{A}^{\otimes n} - S_{n}^{\epsilon}\right\|_{\diamond} \leq \operatorname{poly}(n) \cdot \left\|\left(\left(\mathcal{M}_{A}^{\otimes n} - S_{n}^{\epsilon}\right) \otimes \operatorname{id}_{R'}\right)\left(\zeta_{AR'}^{n}\right)\right\|_{1}$$

<u>But</u>: $\zeta_{AR'}^n$ purification of a special de Finetti state (a state which consists of n iid copies of a state on a single subsystem), i.e., no iid structure and [4] not applicable!

• New one-shot protocol based on quantum-proof randomness extractors:





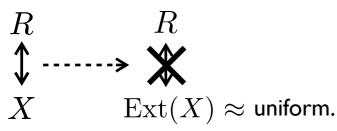
• Now:
$$\|\mathcal{M}_A^{\otimes n} - S_n^{\epsilon}\|_{\diamond} \leq \epsilon$$
 $\|\mathcal{M}\|_{\diamond} = \sup_{k \in \mathbb{N}} \sup_{\|\sigma\|_1 \leq 1} \|(\mathcal{M} \otimes \mathrm{id}_k)(\sigma)\|_1$

• **Post-selection technique** for quantum channels [5]:

$$\left\|\mathcal{M}_{A}^{\otimes n} - S_{n}^{\epsilon}\right\|_{\diamond} \leq \operatorname{poly}(n) \cdot \left\|\left(\left(\mathcal{M}_{A}^{\otimes n} - S_{n}^{\epsilon}\right) \otimes \operatorname{id}_{R'}\right)\left(\zeta_{AR'}^{n}\right)\right\|_{1}$$

<u>But</u>: $\zeta_{AR'}^n$ purification of a special de Finetti state (a state which consists of n iid copies of a state on a single subsystem), i.e., no iid structure and [4] not applicable!

• New one-shot protocol based on quantum-proof randomness extractors:



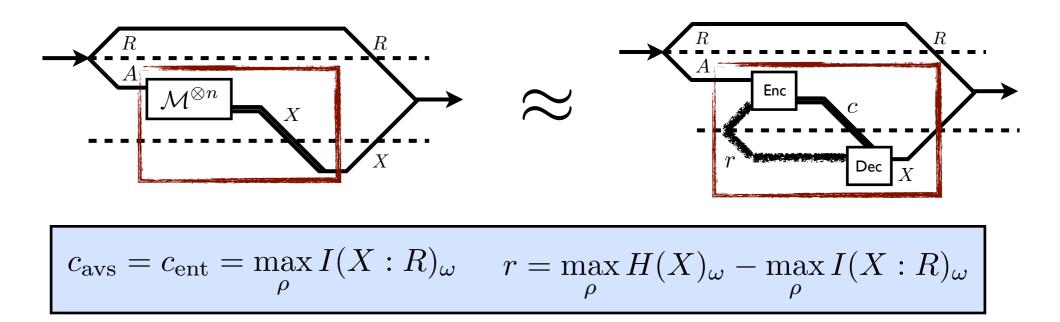
<u>Idea</u>: extract all randomness from measurement data and only send the rest from Alice to Bob to simulate the measurement

<u>Question</u>: How much information (about the input) is gained by performing a given quantum measurement?

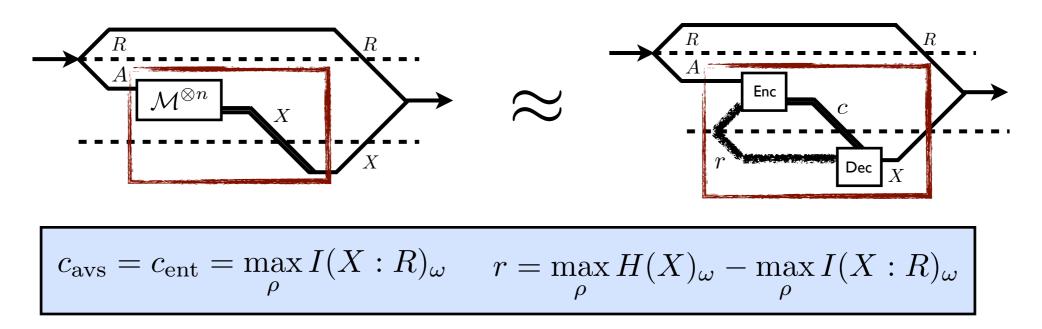
Outline

- Setup making the question precise (information-theoretically)
- Main result answer
- Proof ideas
- Conclusions

• Our main result is measurement simulation on general inputs:



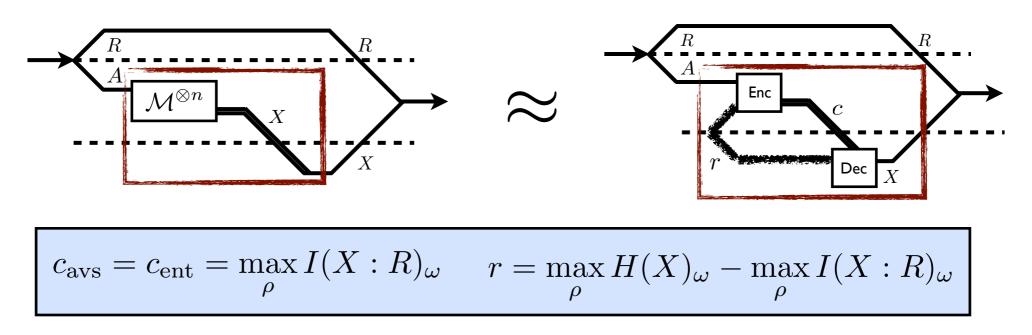
• Our main result is measurement simulation on general inputs:



• Hence, we determined the **information gain of quantum measurements**:

$$I(\mathcal{M}) = \max_{\rho} I(X:R)_{\omega}$$

• Our main result is measurement simulation on general inputs:

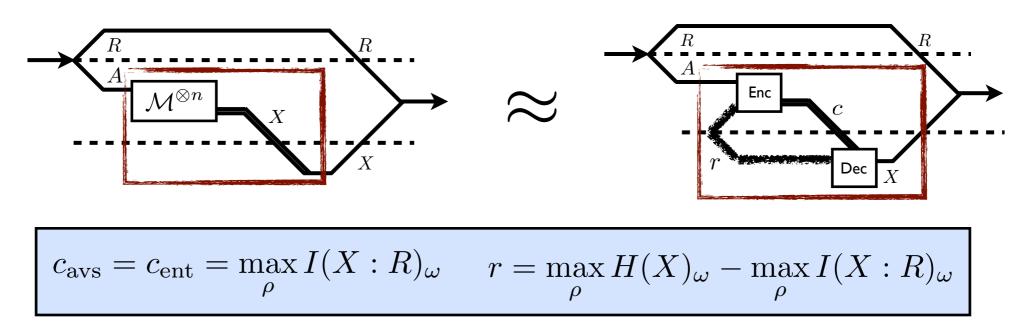


• Hence, we determined the information gain of quantum measurements:

$$I(\mathcal{M}) = \max_{\rho} I(X:R)_{\omega}$$

• Result between classical [6] and quantum [7,8] reverse Shannon theorem

• Our main result is measurement simulation on general inputs:

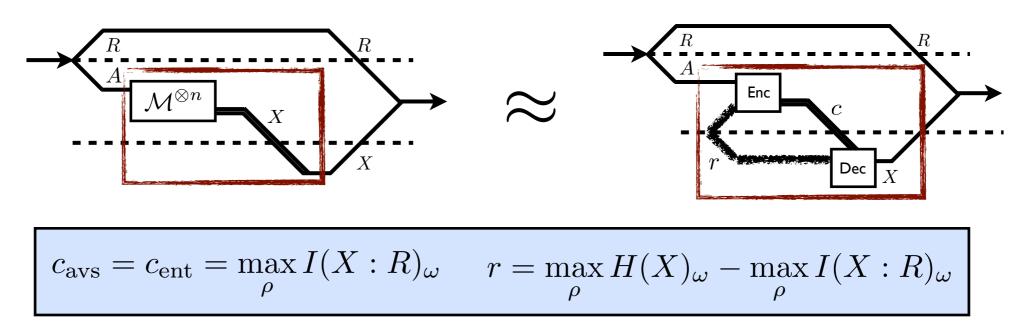


• Hence, we determined the information gain of quantum measurements:

$$I(\mathcal{M}) = \max_{\rho} I(X:R)_{\omega}$$

- Result between classical [6] and quantum [7,8] reverse Shannon theorem
- Extension: rate region for non-feedback vs. feedback measurement simulation $c(r) = \max\left\{\max_{\rho} I(X:R)_{\omega}, \max_{\rho} H(X)_{\omega} r\right\}$

• Our main result is measurement simulation on general inputs:



• Hence, we determined the information gain of quantum measurements:

$$I(\mathcal{M}) = \max_{\rho} I(X:R)_{\omega}$$

- Result between classical [6] and quantum [7,8] reverse Shannon theorem
- Extension: rate region for non-feedback vs. feedback measurement simulation $c(r) = \max\left\{\max_{\rho} I(X:R)_{\omega}, \max_{\rho} H(X)_{\omega} r\right\}$
- Extension: quantum instrument simulation
- Extension: explicit protocols

[6] Bennett et al., IEEE Trans. on Inf. Th. (2002) [7] Bennett et al., arXiv:0912.5537v5 [8] Berta et al., Comm. in Math. Phys. (2011)

Thanks!