Space-time Generalization of Mutual Information

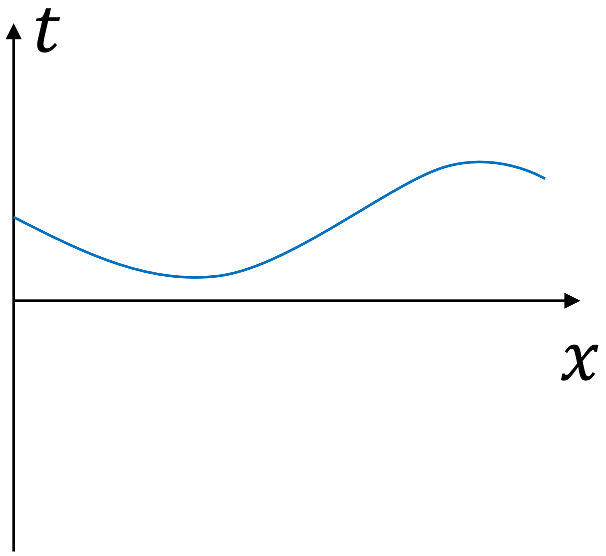
Xiao-Liang Qi

Stanford University

QIEP2022, YITP, 3/24/2022

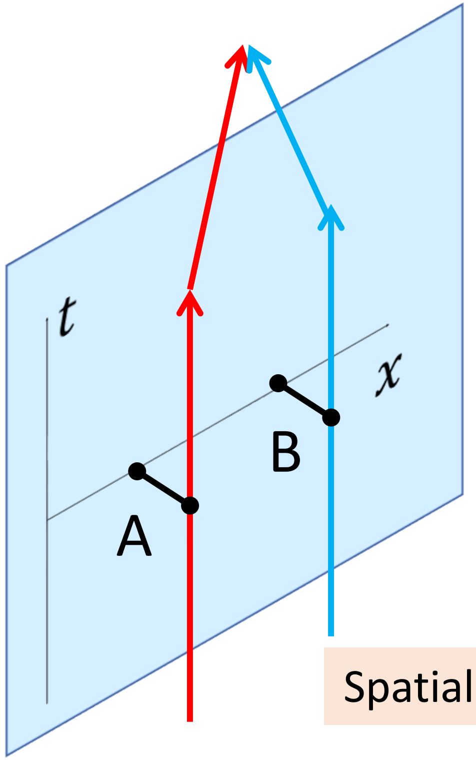
Motivation

- An essential insight of Einstein's theory of *relativity*: Don't talk about absolute "reality" but talk about what an **observer** will see.
- Spacetime is not a fixed given background but a description of correlation.
- A lesson from holographic duality (also known as AdS/CFT): spacetime may be emergent from quantum information.
- Quantum mechanics still depend on an "absolute" time
- $i\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle$
- Quantum information mainly characterizes quantum states.

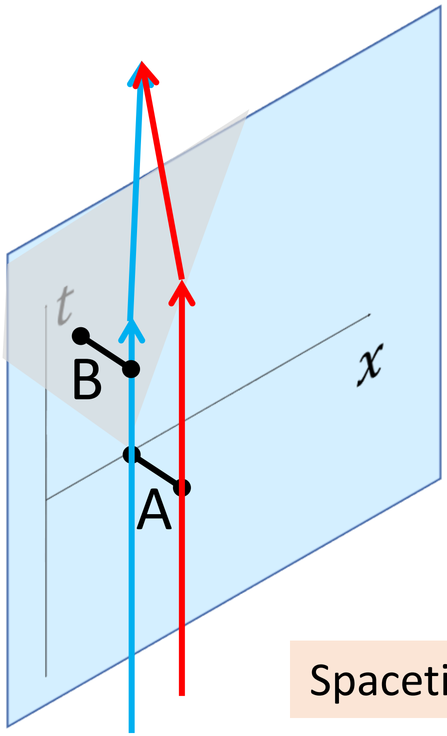


Motivation

- Inspired by relativity, we should focus on observables.
- A (quantum) experimentalist can measure correlation between different spatial points in a similar way as between different time.
- Spacetime correlation should be characterized by quantum information measures similar to those for spatial correlation.



Spatial correlation



Spacetime correlation

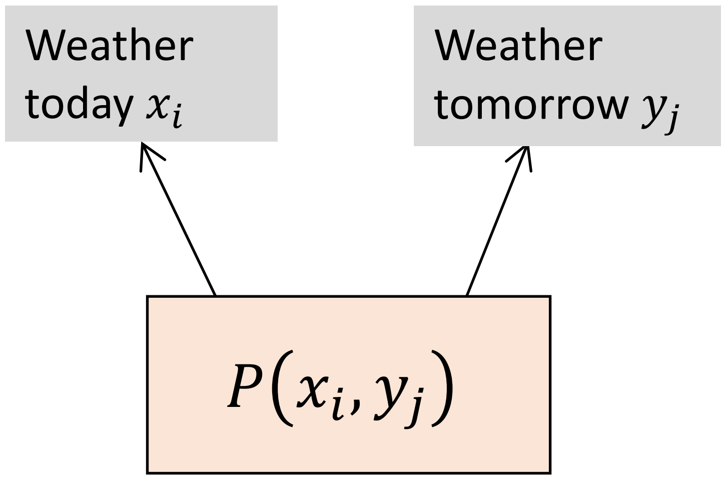
Outline

- Review of mutual information and spatial correlation
- Relative entropy and hypothesis testing
- Setup for the spacetime generalization
- Properties
- Examples

- Ref: Ongoing work with Paolo Glorioso and Zhenbin Yang
- Related previous works:
[1] Cotler, J., Jian, C. M., Qi, X. L., & Wilczek, F. (2018). Superdensity operators for spacetime quantum mechanics. Journal of High Energy Physics, 2018(9), 1-57.
- [2] Cotler, J., Han, X., Qi, X. L., & Yang, Z. (2019). Quantum causal influence. Journal of High Energy Physics, 2019(7), 1-67.
- [3] Aharonov, D., Cotler, J., & Qi, X. L. (2021). Quantum algorithmic measurement. arXiv preprint arXiv:2101.04634.

Classical mutual information

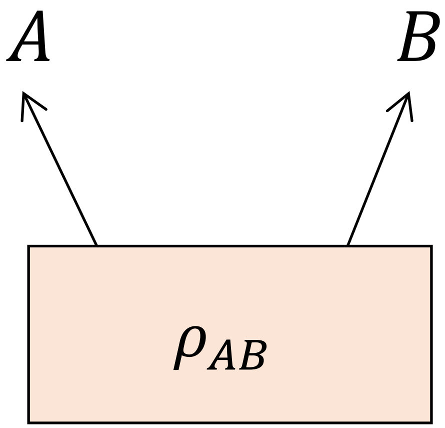
- Two random variables X and Y with values x_i, y_j and joint probability distribution $P(x_i, y_j)$
- Shannon entropy $S(X) = -\sum_i P(x_i) \log P(x_i)$
- If I measure Y, how much do I know about X?
- $I(X:Y) = S(X) + S(Y) - S(XY) = S(X) - S(X|Y)$
- Perfect correlation: $x_i = y_i, P(x_i, y_j) = p(x_i)\delta_{ij}$
- $I(X:Y) = S(X)$
- Knowing Y is the same as knowing X



		
	p_{00}	p_{10}
	p_{01}	p_{11}

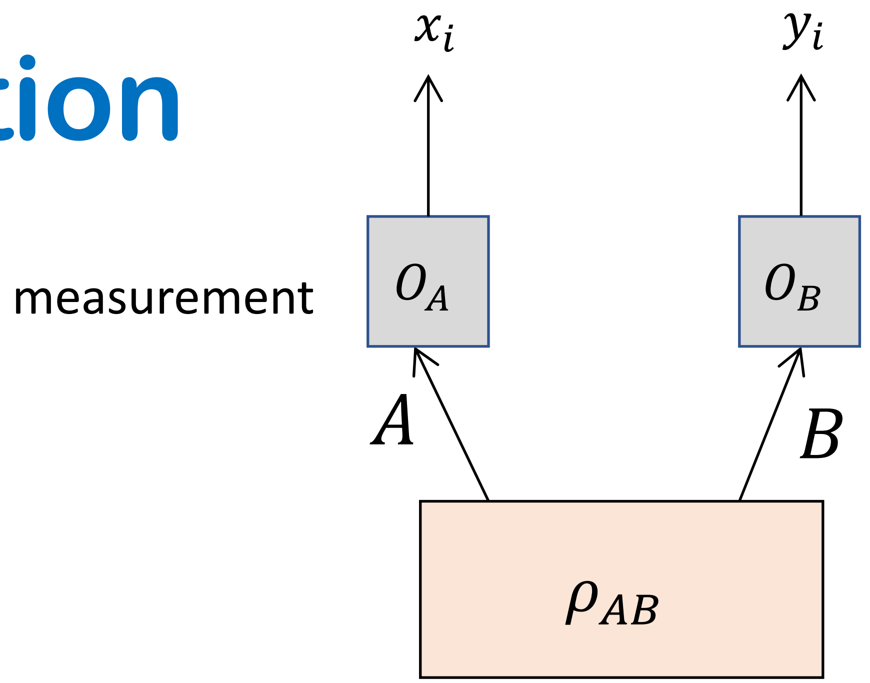
Quantum mutual information

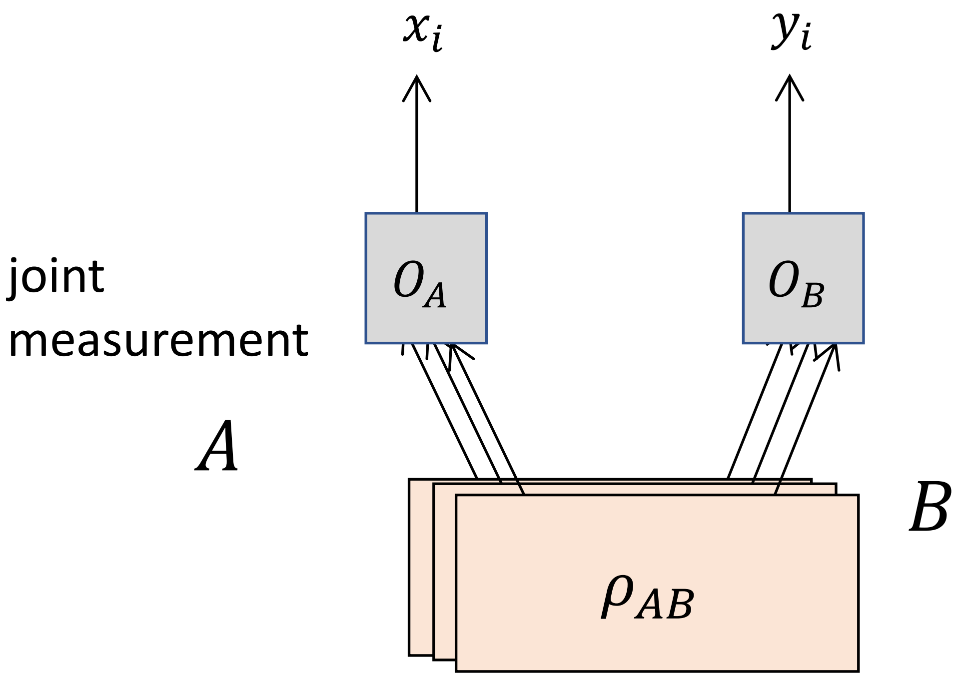
- A quantum system with subsystems A, B
- $\rho_{AB} = tr_{\overline{AB}}(|\Psi\rangle\langle\Psi|)$
- ρ_{AB} determines all correlation functions $\langle O_A O_B \rangle = tr(\rho_{AB} O_A O_B)$
- A quantum information measure independent from the choice of operators is the mutual information
- $I(A{:}B) = S_A + S_B - S_{AB}$
- $S(A) = -tr(\rho_A \log \rho_A)$ is the von Neumann entropy



Quantum mutual information

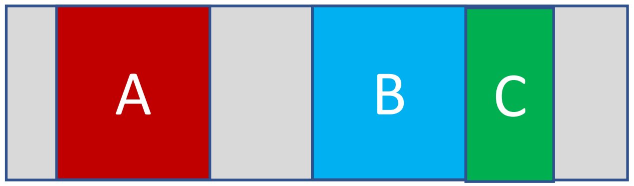
- In general $S_A + S_B \geq S_{AB}$
- $S_A + S_B = S_{AB}$ if and only if $\rho_{AB} = \rho_A \otimes \rho_B$
- Mutual information measures the difference between ρ_{AB} and $\rho_A \otimes \rho_B$
- Mutual information gives an upper-bound to correlation (Wolf Verstraete Hastings Cirac '08)
- Connected correlation function $\langle O_A O_B \rangle_{\mathbf{c}} \equiv \langle O_A O_B \rangle - \langle O_A \rangle \langle O_B \rangle$
- $\dfrac{\langle O_A O_B \rangle_{\mathrm{c}}^2}{2\|O_A\|_\infty \|O_B\|_\infty} \leq I(A{:}B)$
- The bound also works for multiple copies
- $\dfrac{\langle O_A O_B \rangle_{\mathrm{c}}^2}{2\|O_A\|_\infty \|O_B\|_\infty} \leq NI(A{:}B)$

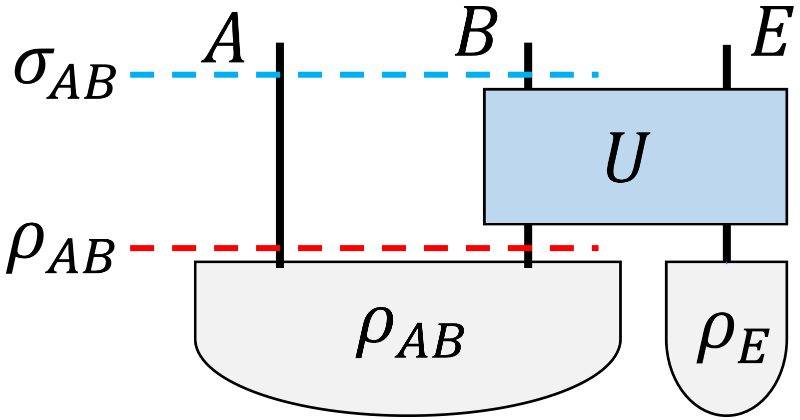




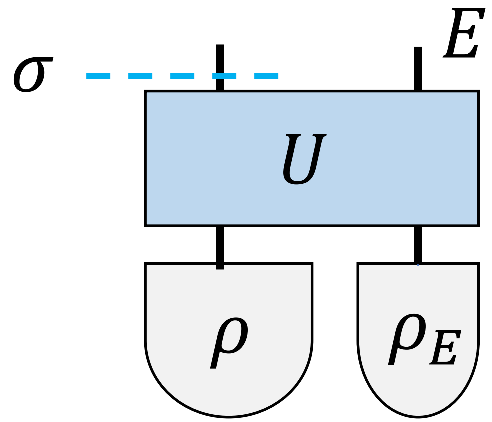
Quantum mutual information

- $I(A\!:\!BC) \geq I(A\!:\!B)$
- Equivalent to strong subadditivity
  $S_{BC} + S_{AB} \geq S_B + S_{ABC}.$
- Mutual information does not increase
  when a quantum channel is applied to A or B
- $I(A\!:\!B)[\sigma_{AB}] \leq I(A\!:\!B)[\rho_{AB}]$

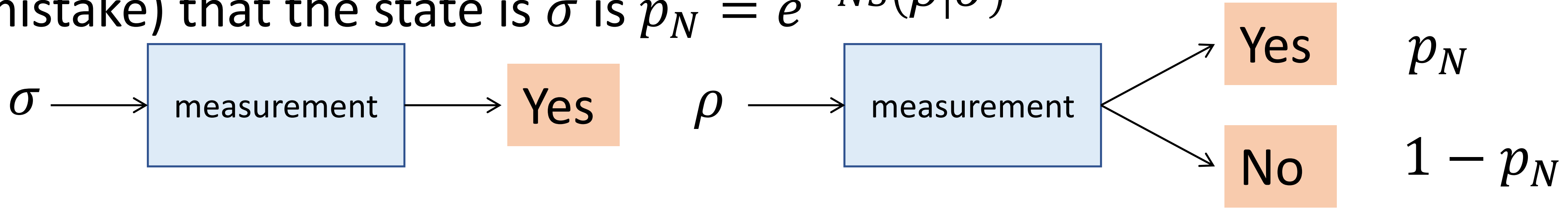




Mutual information and relative entropy



- Relative entropy $S(\rho|\sigma) = tr(\rho \log \rho - \rho \log \sigma)$
- $S(\rho_{AB}|\rho_A \otimes \rho_B) = I(A{:}B)$
- Relative entropy is monotonous under quantum channel $S(\mathcal{N}(\rho)|\mathcal{N}(\sigma)) \leq S(\rho|\sigma)$
- Relative entropy is an asymmetric measure of how different are ρ and σ
- **Hypothesis testing:** An observer is given N copies of the same state. She has a hypothesis that the state is σ, and does an experiment to test if this is true. If the actual state is ρ, the chance that the observer will conclude (by mistake) that the state is σ is $p_N = e^{-NS(\rho|\sigma)}$



Mutual information and relative entropy

- A classical example of hypothesis testing

	0	1
Hypothetical state σ	p	$1-p$
Actual state σ	q	$1-q$

Throw the coin N times
What's the chance that my hypothesis is true?

$$p_N = \binom{N}{M} p^M (1-p)^{N-M} \simeq e^{-NS(q|p)}$$

$$S(q|p) = q\log\frac{q}{p} + (1-q)\log\frac{1-q}{1-p}, q = \frac{M}{N}$$

When the actual state is ρ, after the measurement, we get an output with $M \simeq qN$. The prior probability (that the hypothesis is correct) is p_N

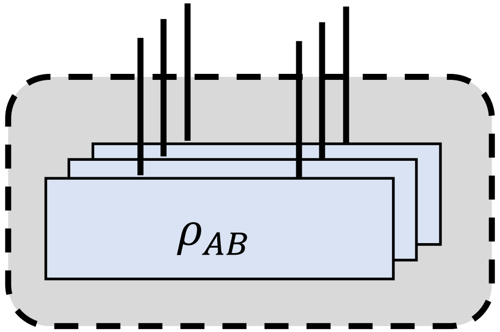
Mutual information and relative entropy

- Bound of connected correlation is also a property of relative entropy
- $S(\rho|\sigma) \geq \frac{1}{2} \parallel \rho - \sigma \parallel_1^2 \geq \frac{|tr(\rho O) - tr(\sigma O)|^2}{2\|O\|^2}$
- Take $O = O_A O_B, \rho = \rho_{AB}, \sigma = \rho_A \otimes \rho_B$
- $tr(O\rho - O\sigma) = \langle O_A O_B \rangle - \langle O_A \rangle \langle O_B \rangle \equiv \langle O_A O_B \rangle_{\mathbf{c}}$
- Thus

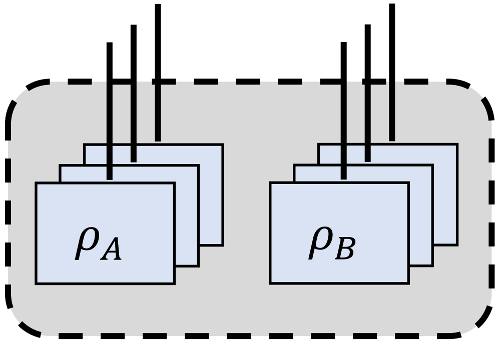
$$\frac{|\langle O_A O_B \rangle_{\mathbf{c}}|^2}{2 \parallel O_A \parallel^2 \parallel O_B \parallel^2} \leq I(A:B)$$

Space-time generalization

- Hypothesis testing in spatial case



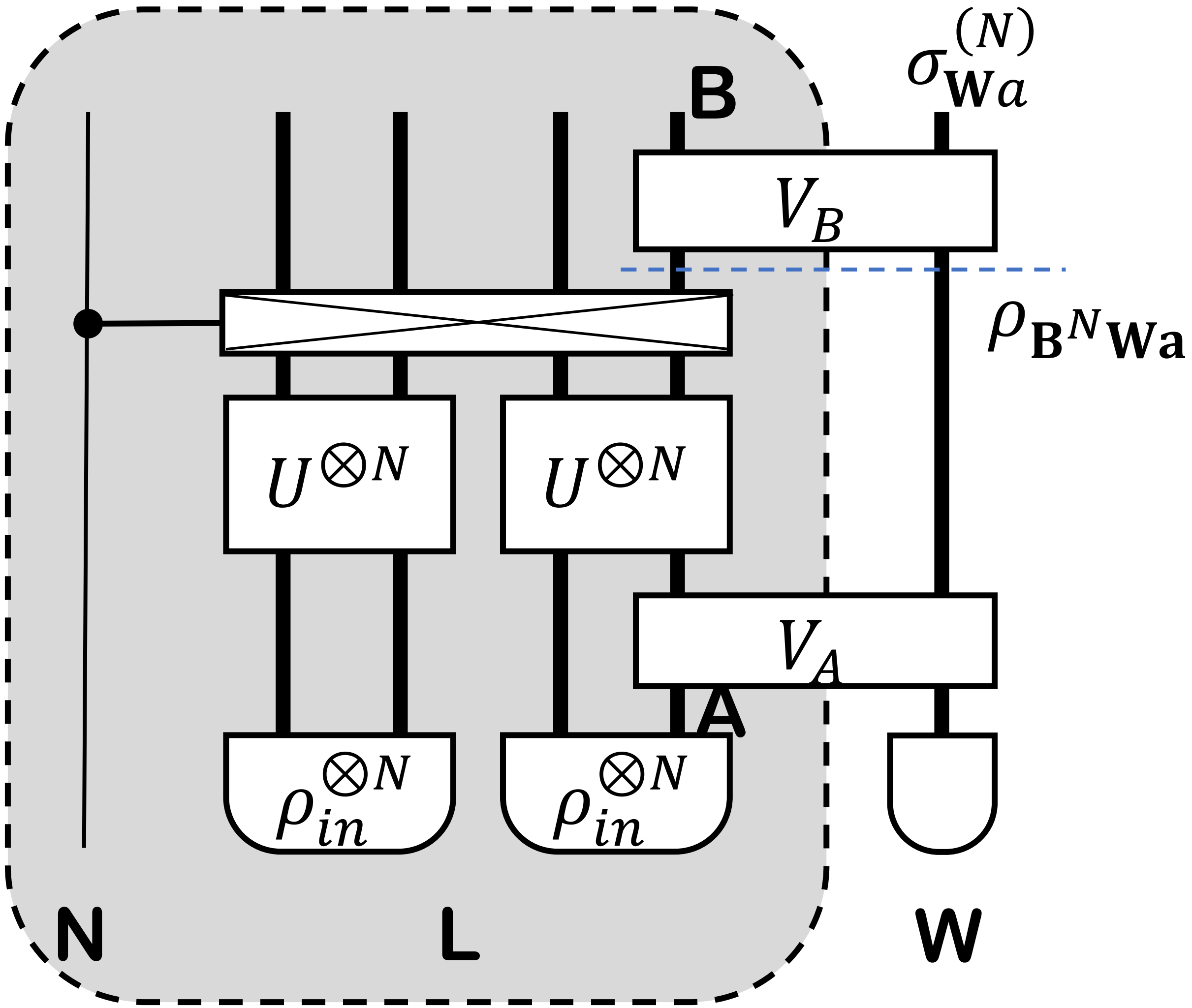
or

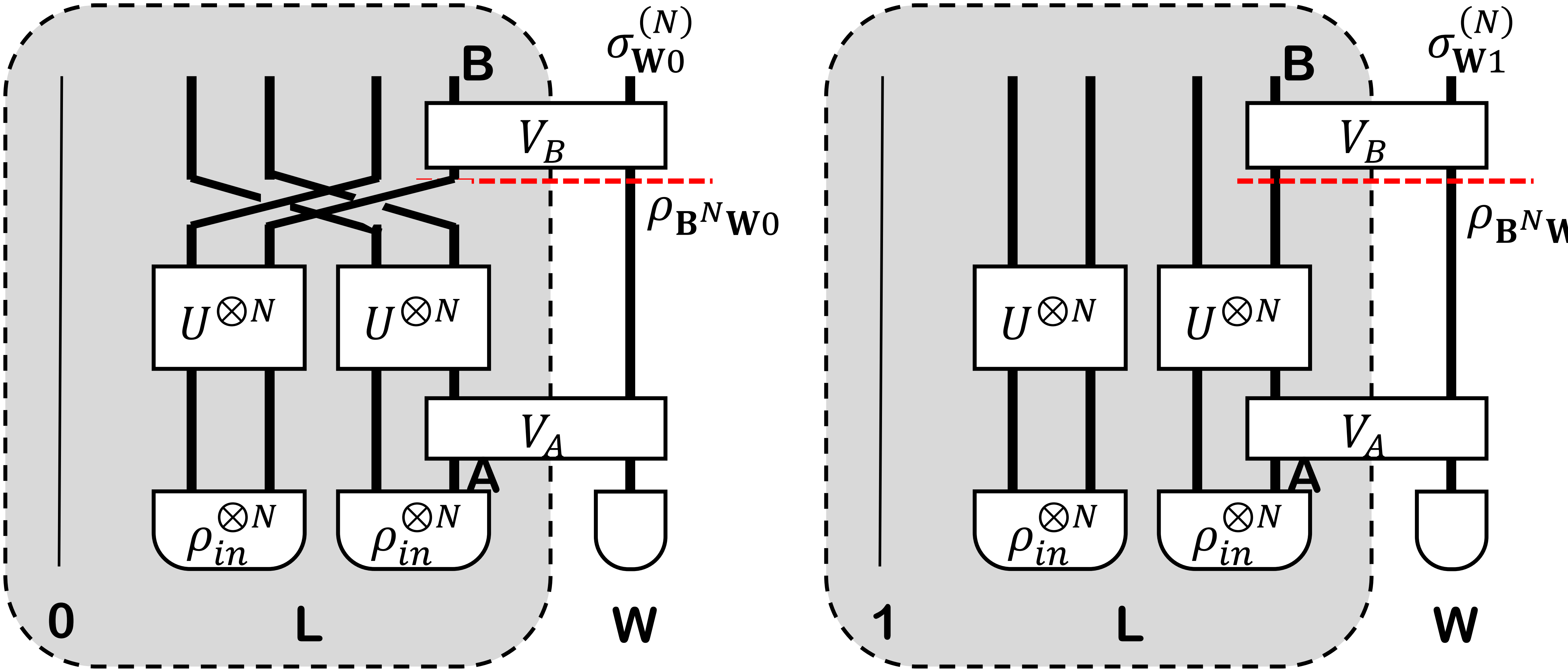


- When the conjectured state is $\rho_A \otimes \rho_B$, and the actual state is ρ_{AB}, the chance that the conjecture is true (after measuring N copies of systems) is $p_N = e^{-NI(A:B)}$
- **Generalize the hypothesis testing to spacetime**: a black box which can be accessed only at A and B. <u>Define a generalized mutual information by the error probability of hypothesis testing</u>

Space-time mutual information

- $2N$ copies of the physical system (**L**) coupled with lab equipment **W** by V_A, V_B (related to Cotler Jian XLQ Wilczek '18)
- A classical bit **N**=0,1 controls two situations
- Case 0: SWAP before applying V_B
- Case 1: Identity before applying V_B
- Task: measure W to distinguish the two cases.
- This is an example of quantum algorithmic measurement (QUALM, Aharonov-Cotler-XLQ '21)

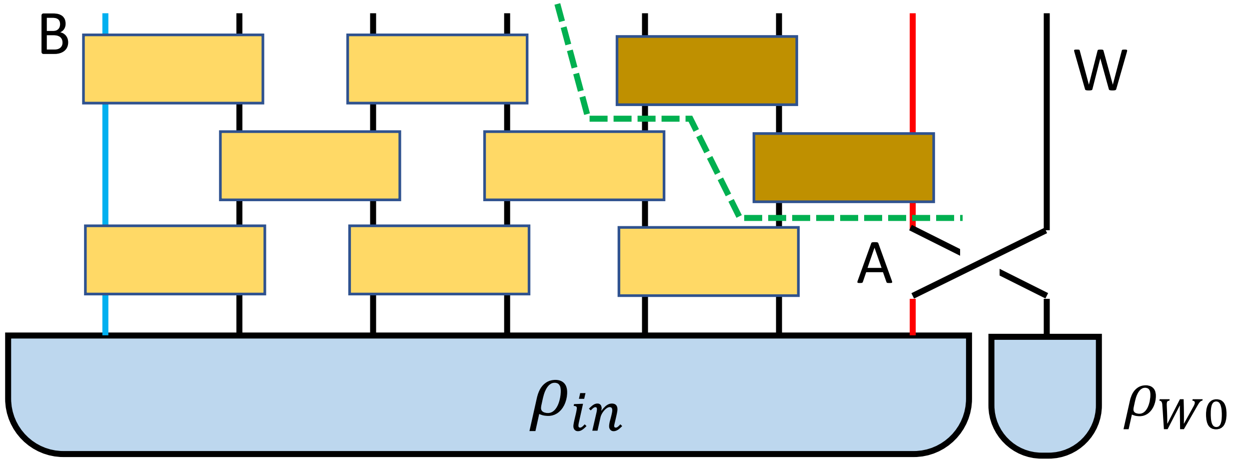




B
$\sigma_{\mathbf{W0}}^{(N)}$
V_B
$\rho_{\mathbf{B}^N\mathbf{W0}}$
$U^{\otimes N}$
$U^{\otimes N}$
V_A
A
$\rho_{in}^{\otimes N}$
$\rho_{in}^{\otimes N}$
0
L
W
B
$\sigma_{\mathbf{W1}}^{(N)}$
V_B
$U^{\otimes N}$
$U^{\otimes N}$
V_A
A
$\rho_{in}^{\otimes N}$
$\rho_{in}^{\otimes N}$
1
L
W

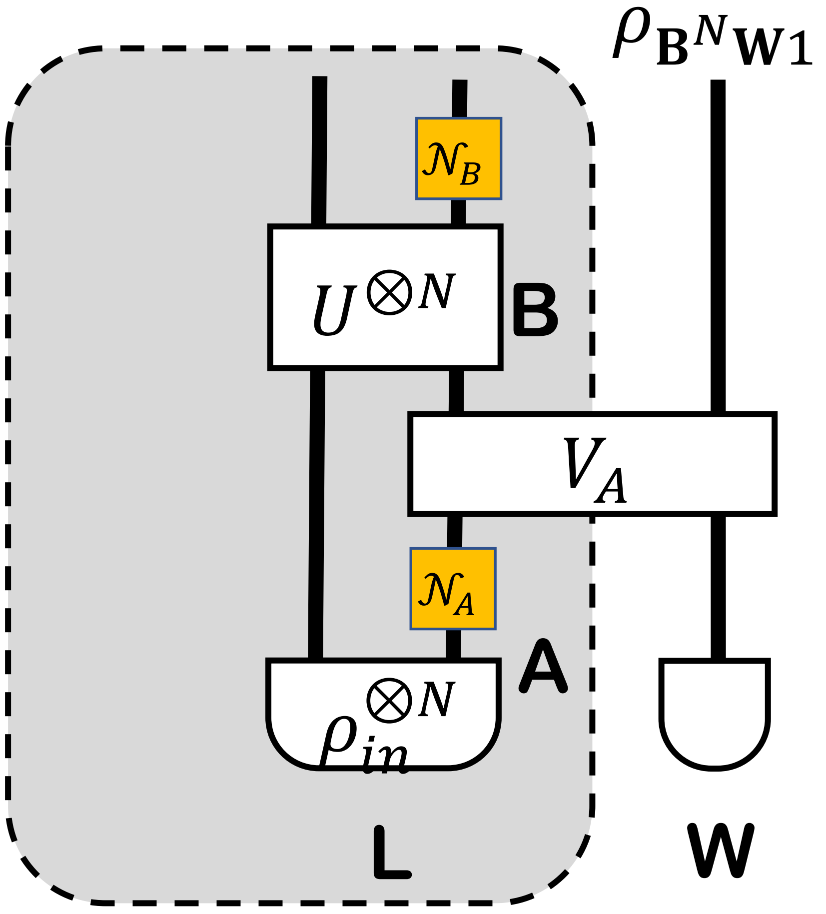
Space-time mutual information

- **W** couples with 2N copies of **L**
- **W** is arbitrarily large.
- Definition $J_N(A:B) = \frac{1}{N} \sup_{V_A,V_B} S(\sigma_{W1}|\sigma_{W0}) , J(A:B) = \lim_{N\to+\infty} J_N(A:B)$
- V_B optimization is trivial: $J_N(A:B) = \frac{1}{N} \sup_{V_A} S(\rho_{BW1}|\rho_{BW0})$
- If A, B are space-like separated, $J_N(A:B) = I(A:B)$ reduces to the mutual information. Supreme is achieved by $V_A = SWAP$



Properties of space-time mutual information

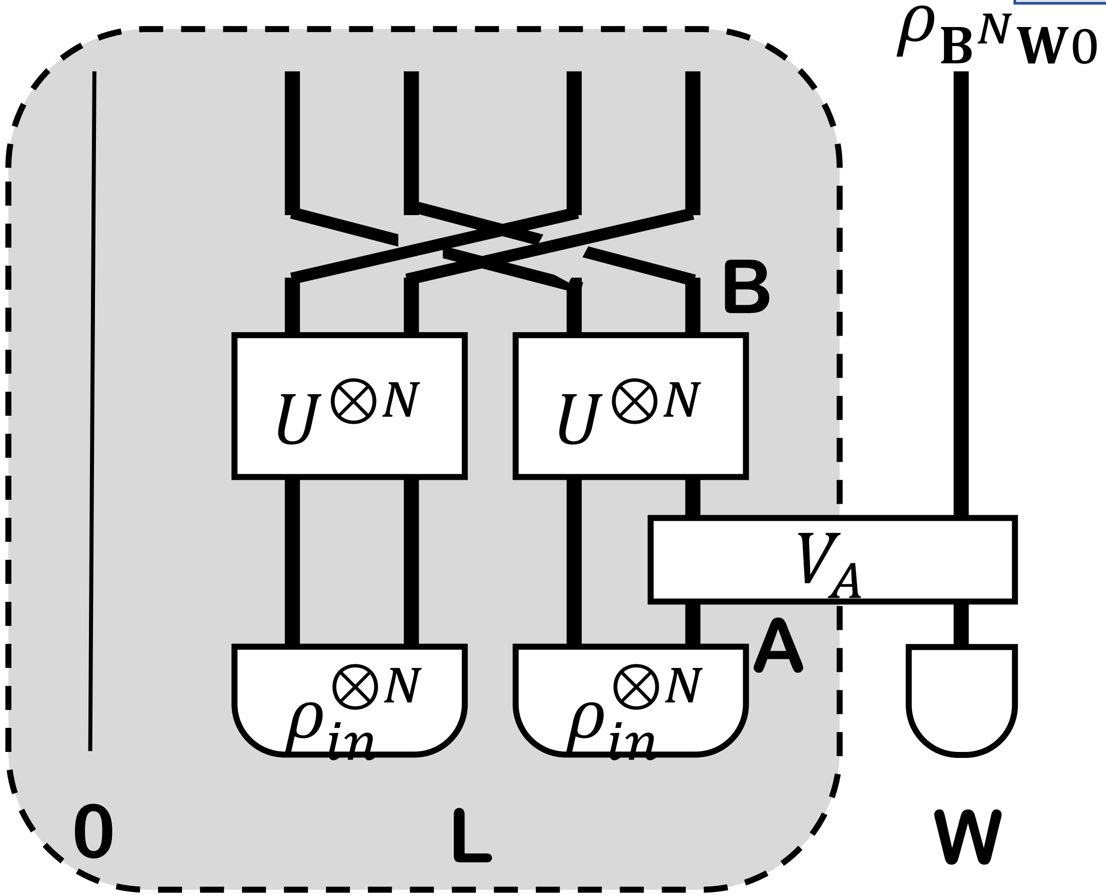
- Monotonicity under quantum channels. $J_N(A{:}B) \geq J_N\big(\mathcal{N}_A(A){:}\mathcal{N}_B(B)\big)$
- $J_N(A{:}B) = 0$ implies that there is no connected correlation function between A and B.
- We can always choose **W** initial state to be $|0\rangle$. This allows us to view V_A as an isometry from **A** to **AW.**

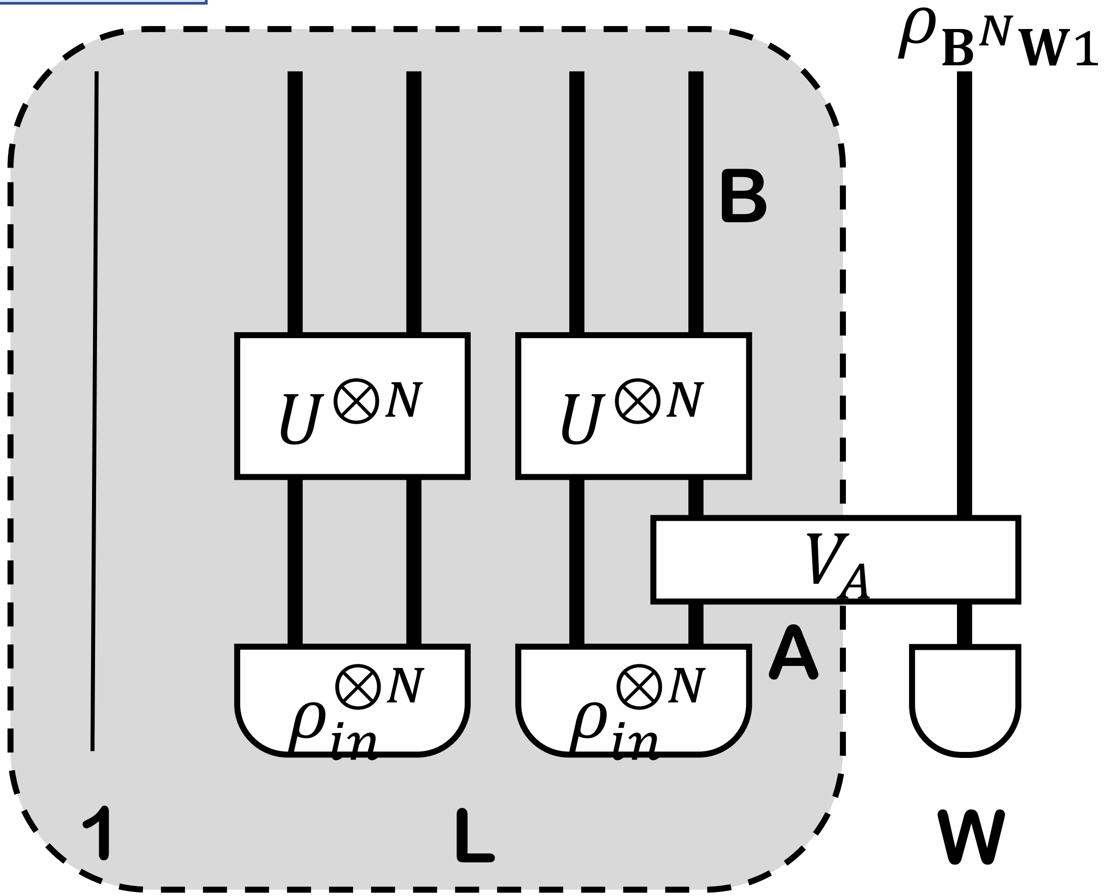


- $S(\rho_{BW1}|\rho_{BW0}) = -S(\rho_{BW1}) - tr(\rho_{BW1}\log(\rho_{B0}\otimes\rho_W))$
  $= I(B{:}W)[\rho_{BW1}] + S(\rho_B|\rho_{B0})$
- $\rho_B = tr_W(\rho_{BW1})$

Mutual information between B and W

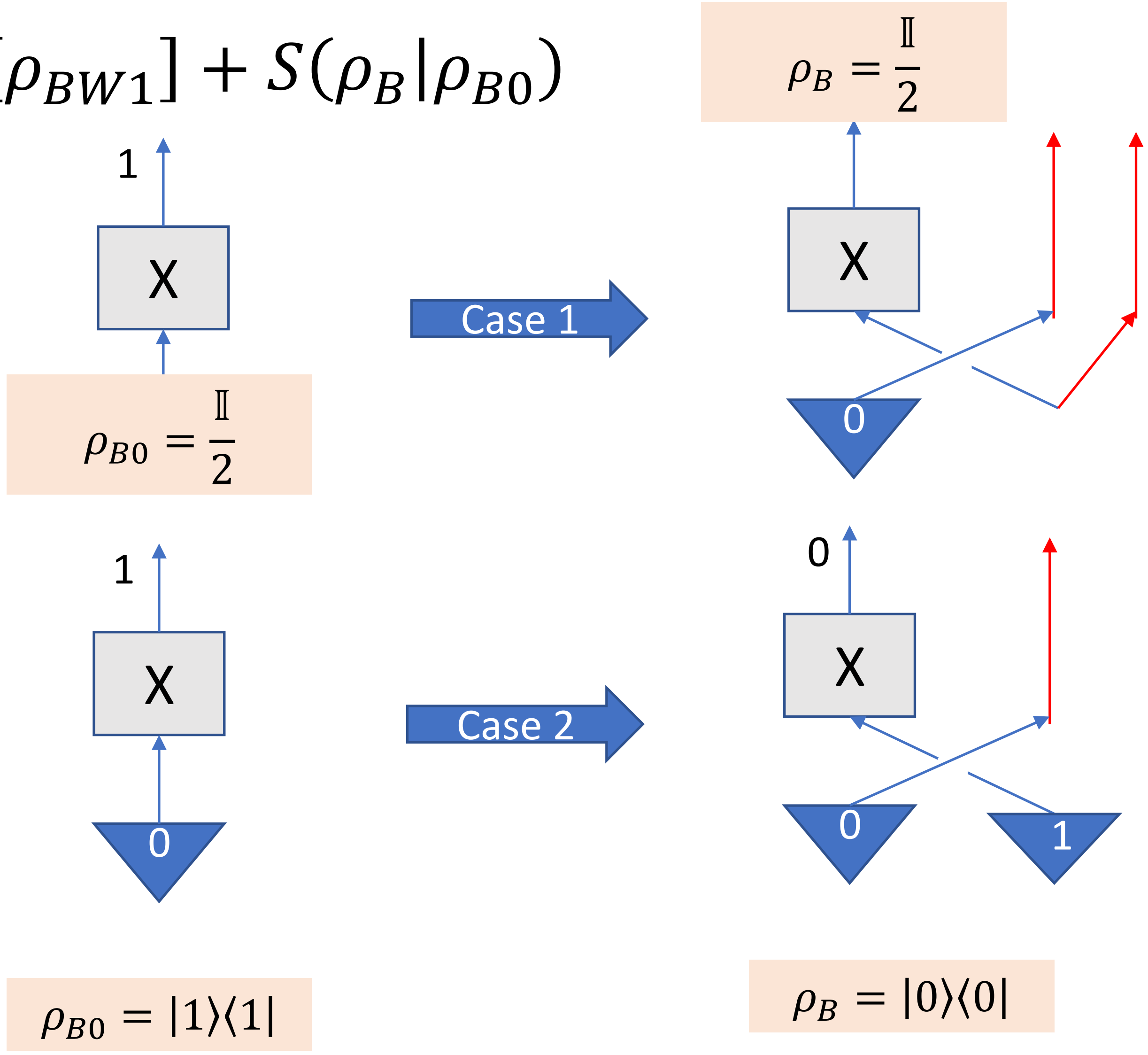
Relative entropy of "affected state" ρ_B and original state ρ_{B0}. New term in spacetime case. This term may diverge.





Interpretation of the decomposition

- $S(\rho_{BW1}|\rho_{BW0}) = I(B:W)[\rho_{BW1}] + S(\rho_B|\rho_{B0})$
- Example
- Case 1:
- $S(\rho_B|\rho_{B0}) = 0$
- $I(B:W) = \log 2$
- Case 2:
- $S(\rho_B|\rho_{B0}) = +\infty$
- $I(B:W) = 0$

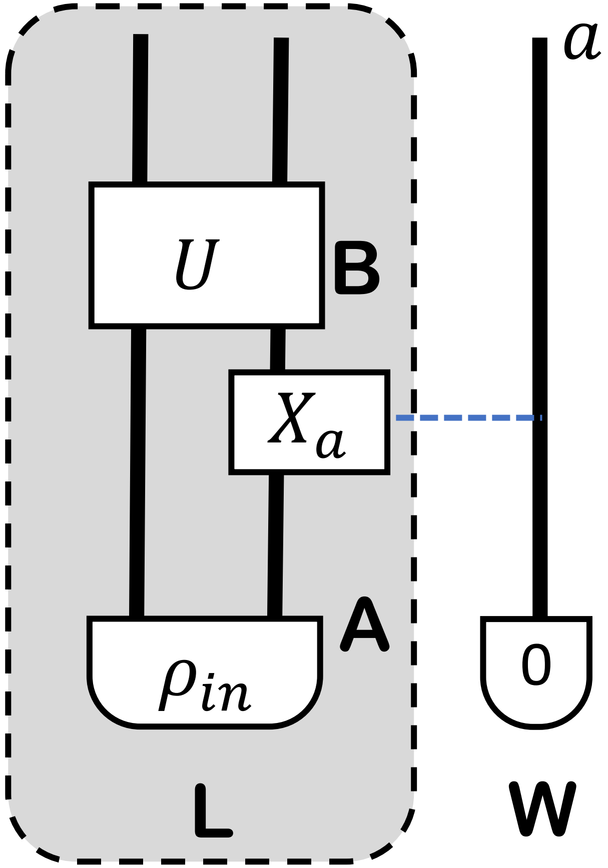


Bound of correlation function

- Theorem: For arbitrary Hermitian operators O_A, O_B acting on A, B, denote $\hat{O}_A, \hat{O}_B$ as the corresponding Heisenberg operators. We have
- $\frac{1}{8}\left(\frac{tr(\rho_{in}(-i)[\hat{O}_B,\hat{O}_A])}{\|O_B\|\|O_A\|}\right)^2 \leq J_1(A:B)$
- $\frac{1}{8}\left(\frac{tr(\rho_{in}\{\hat{O}_B,\hat{O}_A\})_c}{\|O_B\|\|O_A\|}\right)^2 \leq J_1(A:B)$
- Proof: Define three operators acting on A:

$$X_0 = \sqrt{1-p}\mathbb{I}, X_1 = \sqrt{p}\frac{O_A}{\|O_A\|}, X_2 = \sqrt{p}\sqrt{\mathbb{I} - \frac{O_A^2}{\|O_A\|^2}}.$$

- $X_0^+ X_0 + X_1^+ X_1 + X_2^+ X_2 = \mathbb{I}$



Bound of correlation function

- Now define an operator applying to W:

$$Y_W = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

- One can prove

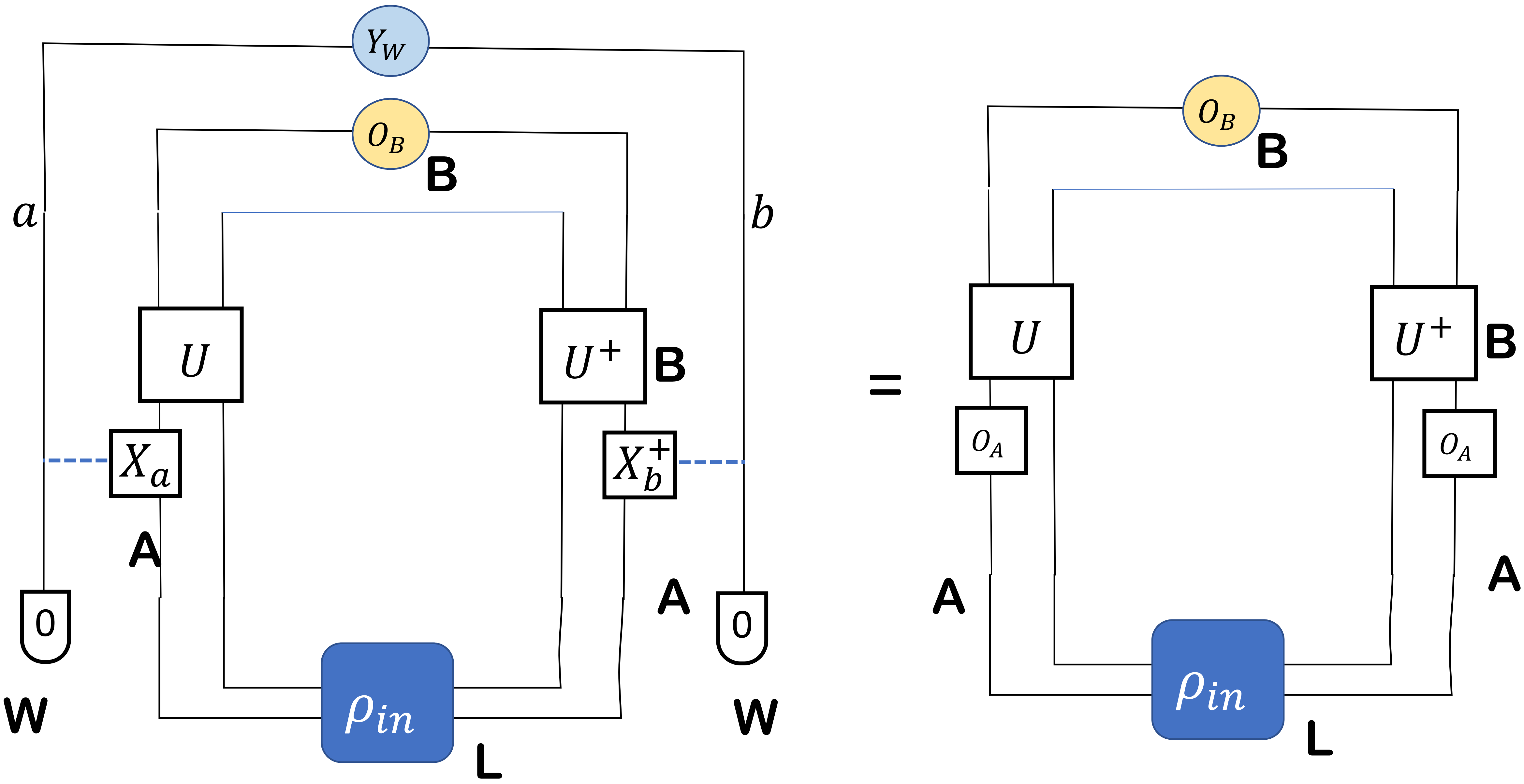
$$tr(\rho_{BW1} Y_W O_B) = -i \frac{\sqrt{p(1-p)}}{\|O_A\|} tr([\hat{O}_A, \hat{O}_B]\rho_{in})$$
$$tr(\rho_{BW0} Y_W O_B) = 0$$

- On the other hand,

$$\left(\frac{|tr((\rho_{BW1} - \rho_{BW0}) Y_W O_B)|}{\|Y_W\| \|O_B\|}\right)^2 \le \|\rho_{BW1} - \rho_{BW0}\|_1^2 \le 2S(\rho_{BW1}|\rho_{BW0})$$

- ➔ $\frac{1}{8}\left(\frac{-itr([\hat{O}_A, \hat{O}_B]\rho_{in})}{\|O_A\|\|O_B\|}\right)^2 \le J_1(A:B)$ (we have taken $p = \frac{1}{2}$)

For related ideas on measuring correlation with ancilla, see e.g. R. Somma, et al , Phys. Rev. A 65, 042323



Y_W
O_B
B
a
b
U
U^+
B
X_a
X_b^+
A
A
0
0
W
W
ρ_in
L
=
O_B
B
U
U^+
B
O_A
O_A
A
A
ρ_in
L

Bound of correlation function

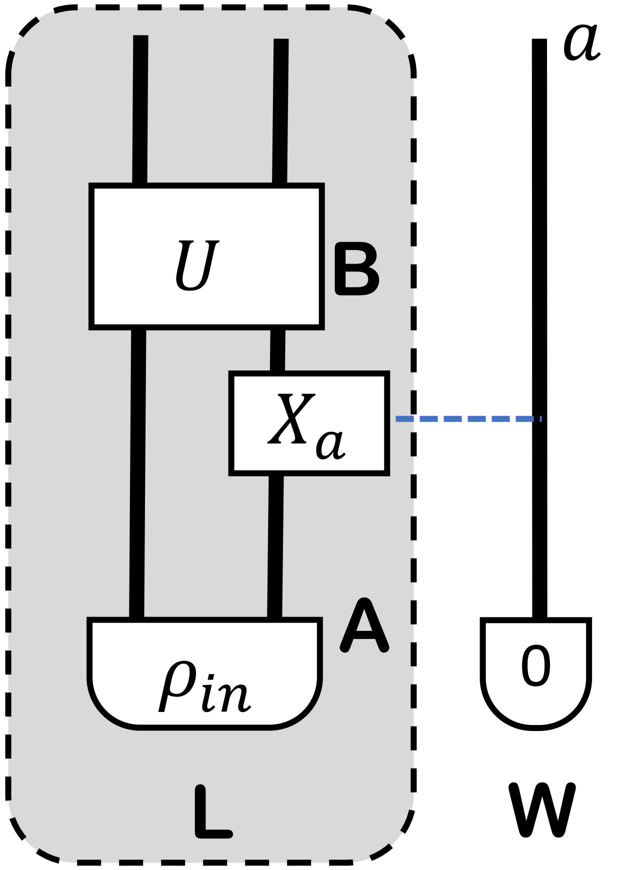
- The other formula $\frac{1}{8}\left(\frac{tr(\rho_{in}\{\hat{O}_B,\hat{O}_A\})_c}{\|O_B\|\|O_A\|}\right)^2 \leq J_1(A:B)$ can be proven similarly by replacing Y_W with

$$X_W = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

- Comments
- 1. For spacelike separated regions, this bound reduces to the Wolf et al result.
- 2. Physically, a bound can be found because all correlators can be measured by an appropriate W.
- 3. In the case of response function (commutator), $J_1(A:B)$ can be replaced by $\sup_{V_A} I(B:W)$, which is a tighter bound. (This is because the disconnected term vanishes.)

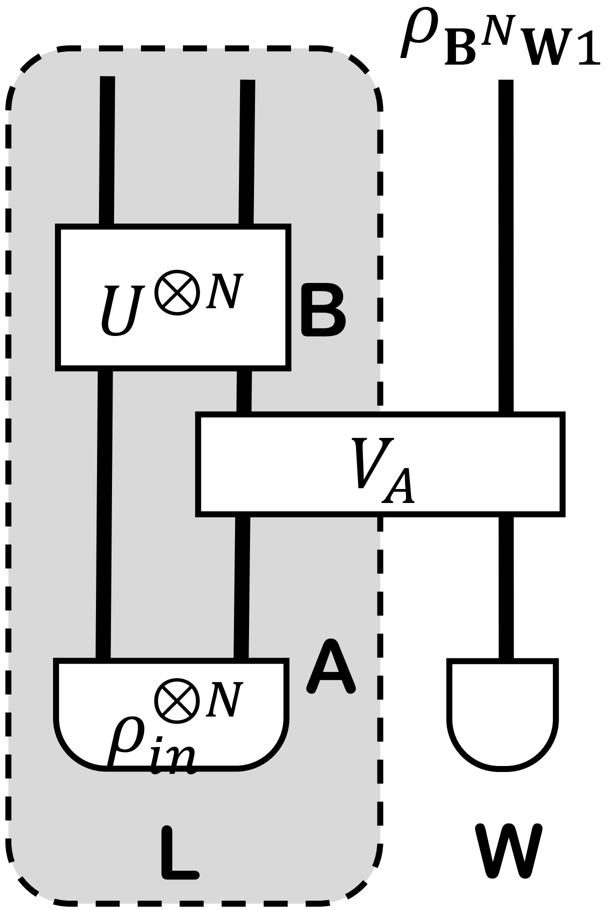
Bound of causal influence

- Consider a similar setup but with W a qubit.
- $X_0 = \frac{1}{\sqrt{2}}\mathbb{I}, X_1 = \frac{1}{\sqrt{2}}R_A$, R_A is unitary
- Define $Z_W = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
- For Hermitian operator O_B, $\frac{tr((\rho_{BW1}-\rho_{BW0})Z_W O_B)}{\|O_B\|} = \frac{\langle O_B\rangle_{R_A}-\langle O_B\rangle_{\mathbb{I}}}{2\|O_B\|}$.
- This gives a bound of causal influence (Cotler-Han-XLQ-Yang). If we define $CI(A:B) = \sup_{R_A}\frac{\left|\langle O_B\rangle_{R_A}-\langle O_B\rangle_{\mathbb{I}}\right|}{2\|O_B\|}$, then $CI(A:B)^2 \leq 2J_1(A:B)$



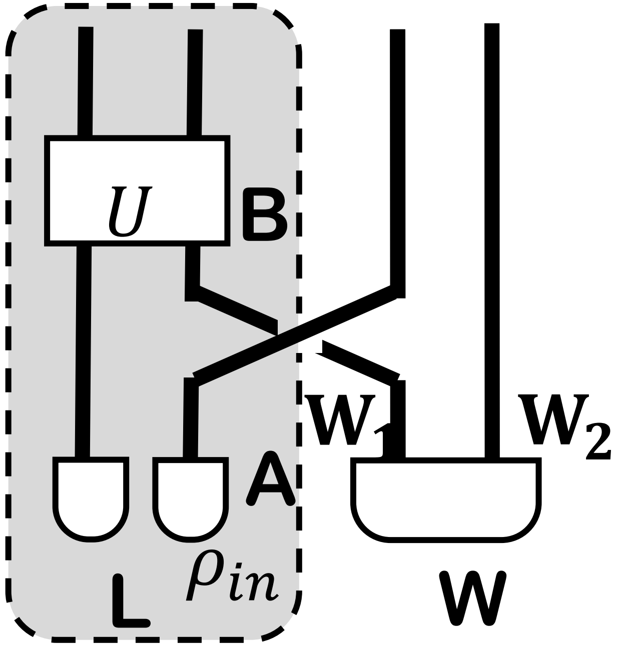
An upper bound of spacetime MI

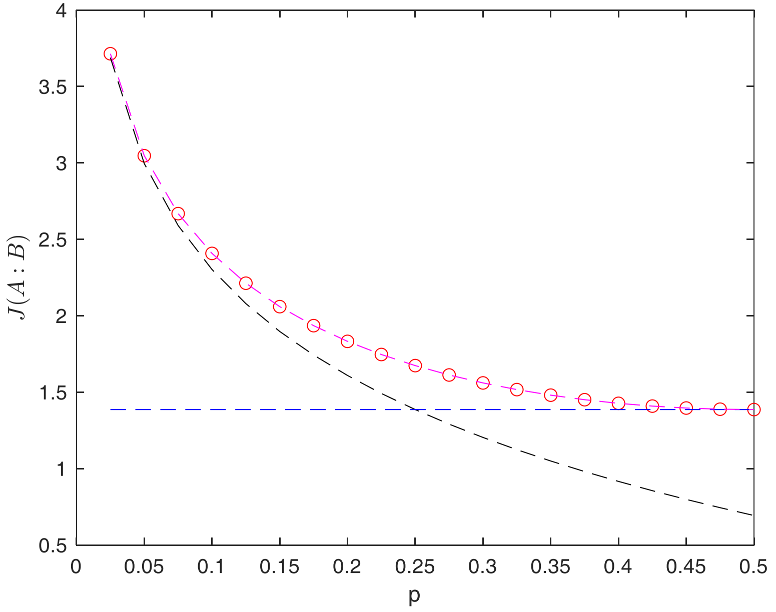
- How big can the space-time mutual information be?
- Spatial mutual information $I(A\colon B) \leq \min\{2S_A, 2S_B\}$
- Spacetime mutual information:
- $I(B\colon W) \leq \min\{2S_B, 2S_W\} \leq \min\{2S_B, 2S_A + 2|A|\}$
- $S(\rho_B|\rho_{B0}) \leq -\log p_{min}(\rho_{B0})$
- p_{min} is the minimal eigenvalue of ρ_{B0}.
- For given size of A, B, $\mathrm{I(B\colon W)} \leq \min\{2|B|, 4|A|\}$ but the relative entropy term does not have a universal upper bound.



Analytic results

- If the initial state of A is not entangled
- $\rho_{B0} = \mathcal{N}(\rho_{in})$, we prove that $V_A = SWAP$ is always a saddle point
- We only need to variationally determine ρ_{W1}
- In addition if $\mathcal{N}: A \to B$ is unitary,
  $\rho_{W1} = C\rho_{in}^{-1}$,
  $J_1(A:B) = J(A:B) = 2S(\rho_{W1}) + S(\rho_{W1}|\rho_{in})$
  $= \log(\sum_i p_i^{-1})$.
- p_i are eigenvalues of ρ_{in}.
- For example, $\rho_{in} = \begin{pmatrix} p & 0 \\ 0 & 1-p \end{pmatrix}$
- $J(A:B) = -\log(\mathrm{p}) - \log(1-\mathrm{p})$



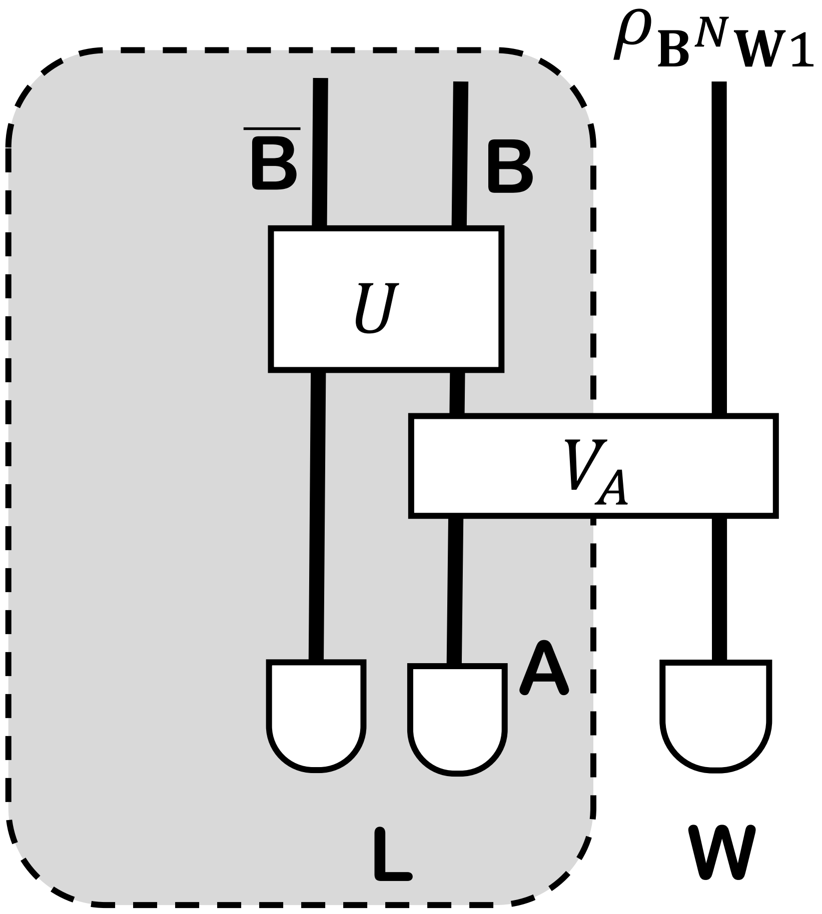


Relation with quantum channel discrimination

- If the initial state is pure $\rho_{in} = |\psi_{in}\rangle\langle\psi_{in}|$, then $V_A|\psi_{in}\rangle^{\otimes N} \otimes |0\rangle_W$ can be an arbitrary state, denoted as ρ_{AW}. $\rho_{B^N W 1}$ and $\rho_{B^N W 0}$ are output of two different channels for the same input state.
- This is the quantum channel discrimination problem
- $D_N(\mathcal{N}_1|\mathcal{N}_2) = \frac{1}{N}\sup_{\rho_{AW}} S\left(\mathcal{N}_1^{\otimes N}(\rho_{AW})\middle|\mathcal{N}_2^{\otimes N}(\rho_{AW})\right)$
- $J_N(A:B) = D_N(\mathcal{N}|\mathcal{R})$
- $\mathcal{R}$ is a replacer channel. $\mathcal{R}(\rho_{AW}) = \rho_{A0}^{\otimes N} \otimes tr_A(\rho_{AW})$. For replacer channel, it is known that $D_N(\mathcal{N}|\mathcal{R})$ is independent from N (Cooney,Mosonyi,Wilde '16)
- This additivity does not hold for general channels (Fang,Fawzi,Renner,Sutter '20).

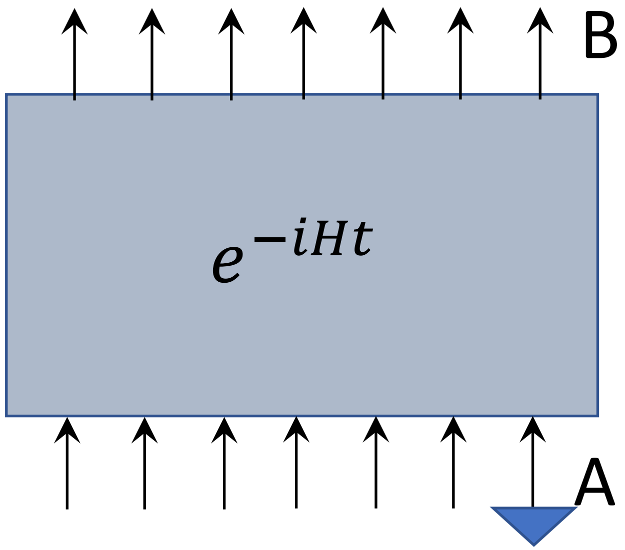
Example: thermodynamics

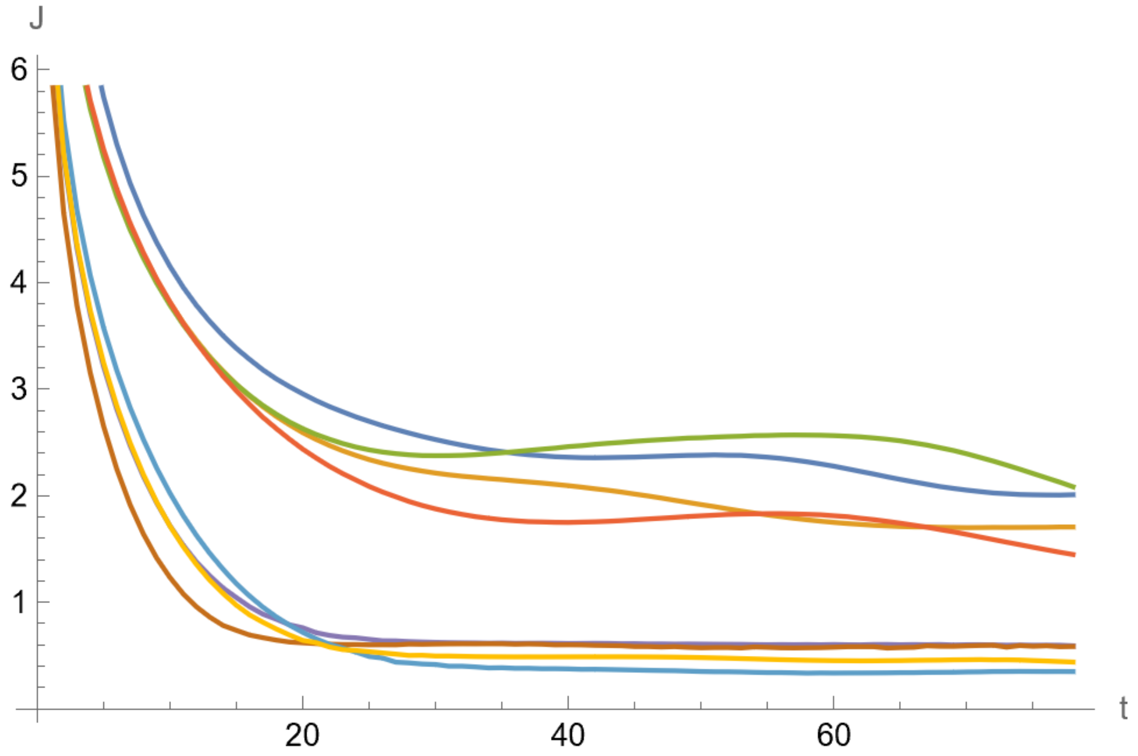
- Consider a U that is a Haar random unitary, and the initial state is a product of pure states.
- If B is smaller than half system, ρ_B is almost maximally mixed independent from V_A, so that $S(\rho_B|\rho_{B0}) = 0$.
- If $U = e^{-iHt}$, in long time, small region B thermalizes to $\rho = \frac{1}{Z}e^{-\beta H_B}$. In this case, the correlation is nonzero. For a small B, the mutual information $I(B{:}W)$ comes from classical correlation of β with the energy input from A. The leading contribution to $J_1(A{:}B)$ comes from $S(\rho_B|\rho_{B0}) = \beta_0\big(\mathrm{F}(\beta) - F(\beta_0)\big)$



Example: many-body localization

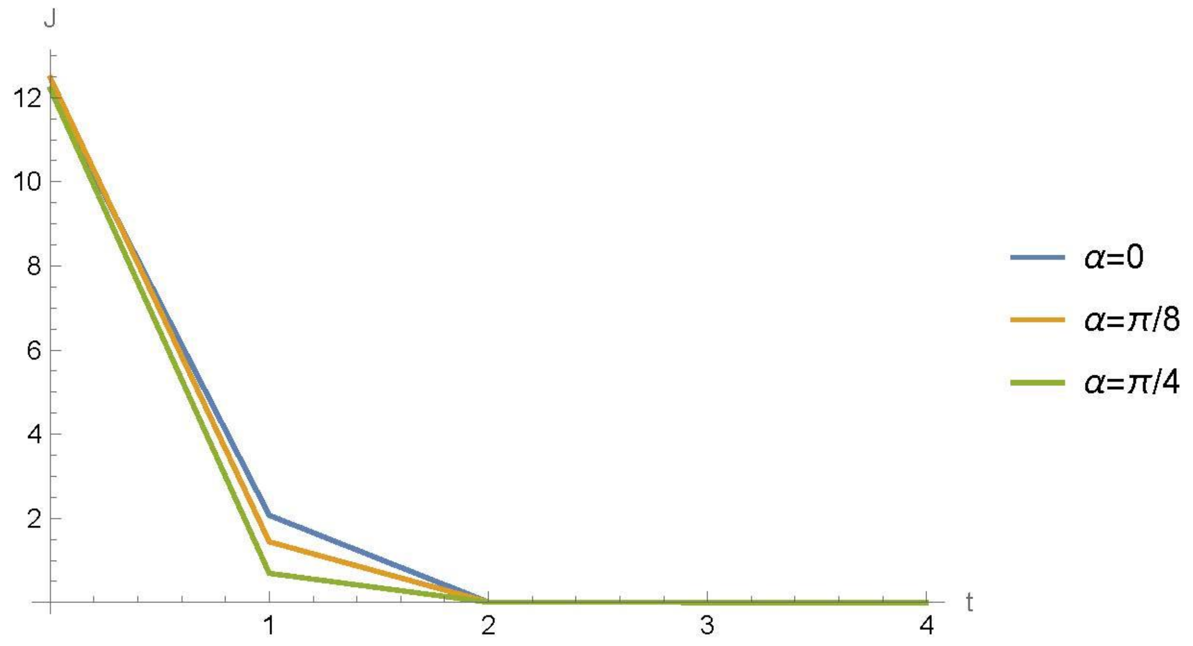
- MBL in random-field spin chain $H = \sum_{j=1}^{L-1} \vec{\sigma}_j \cdot \vec{\sigma}_{j+1} + \sum_{j=1}^{L} h_j \sigma_j^z$
- Initial state $|\chi\rangle\langle\chi| \otimes \frac{\mathbb{I}}{D}$, $|\chi\rangle = \cos\alpha\,|0\rangle + \sin\alpha\,|1\rangle$
- B is the same site as A at later time.
- Long time residual J_{AB} nonzero, $|\chi\rangle$ dependent.





Example: thermalizing Floquette system

- $U = e^{-i\frac{\tau}{2}H_x} e^{-i\tau H_z} e^{-i\frac{\tau}{2}H_x}$
- $H_x = g\sum_{j=1}^{L}\sigma_j^x,\ H_z = \sum_{j=1}^{L-1}\sigma_j^z\sigma_{j+1}^z + h\sum_{j=1}^{L}\sigma_j^z$
- $J_1(A\!:B)$ vanishes after some finite time.



Summary and open questions

- We proposed related quantities $J_N(A:B)$ and $J(A:B)$ which are space-time generalizations of mutual information.
- They have interpretation in hypothesis testing, and provide upper bounds to different kinds of correlation functions.
- For pure un-entangled initial state, $J_N = J_1$
- For mixed un-entangled initial state, probably $J_N = J_1$ but not proven.
- Open questions:
  1. Explicit computation of $J_N(A:B)$, especially $J_1(A:B)$ and $J(A:B)$. Does $J_N(A:B)$ actually grow with N?
  2. Relation to other quantum information measures, such as quantum channel capacity.
  3. Other examples.