Efficient decoding for the Hayden-Preskill protocol

(Joint work with Beni Yoshida)

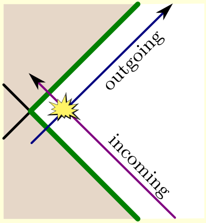
arXiv:1710.03363

Alexei Kitaev (Caltech)

Motivation: quantum black holes

- Quantum fields in classical space (Hawking): $T = \frac{\varkappa}{2\pi}, \quad S = \frac{A}{4}$

- Information scrambling: $\tau_{\rm scr} \approx (2\pi T)^{-1} \ln S$
  - Gravitational interaction between incoming and outgoing radiation
  - Dray-t'Hooft shock waves
  - OTOCs: $\langle W(t)\, Y(0)\, Z(t)\, X(0) \rangle$



- Evolution over Page's time, when half of the black hole evaporates
  - Full quantum gravity

?

Assumptions

- Thermal state is replaced with the maximally mixed state on a "typical subspace" $\mathcal{L}$:

$$\rho = Z^{-1} e^{-H/T} \quad \longrightarrow \quad \rho = \frac{I_{\mathcal{L}}}{d}, \quad d = \dim \mathcal{L} = \cancel{e^S}\, 2^S$$

- Late-time OTOCs (= almost perfect scrambling):

$$\big\langle W(t)\, Y(0)\, Z(t)\, X(0) \big\rangle \approx \langle WZ\rangle\langle Y\rangle\langle X\rangle + \langle W\rangle\langle Z\rangle\langle YX\rangle - \langle W\rangle\langle Z\rangle\langle Y\rangle\langle X\rangle$$

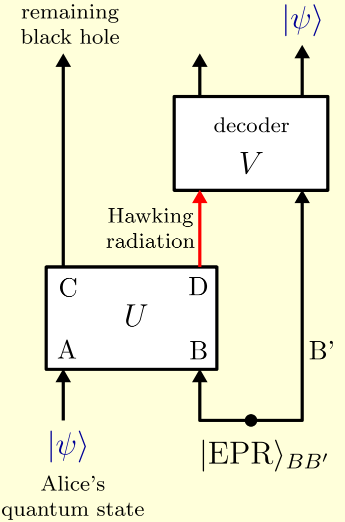
where $W(t) = U^\dagger W U, \quad Z(t) = U^\dagger Z U, \quad Y(0) = Y, \quad X(0) = X$

  – Holds for a Haar-random unitary U

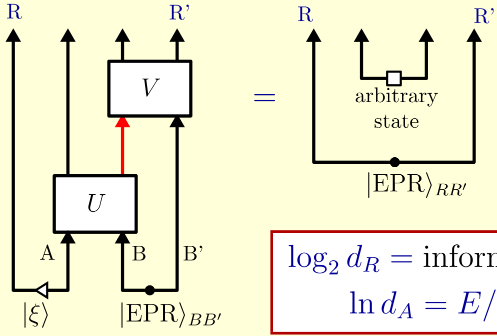
  – Broadly applicable if X, Y, Z, W act on small subsystems

The Hayden-Preskill problem

Basic version

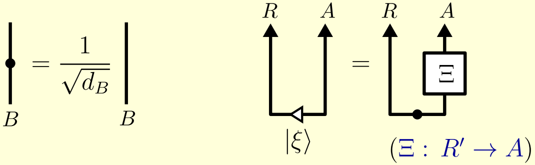


A variant using reference system R



$\log_2 d_R =$ information

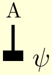
$\ln d_A = E/T$

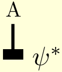


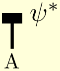
Tensor diagrams

- Nodes are tensors; lines are index contractions.

- Time goes up.
  - Vertical sections of lines are associated with Hilbert spaces. If a line goes up and then down, the Hilbert space changes to the dual space.
  - Let $\psi \in A$ be a vector with elements c_j and $\psi^* \in A'$ a vector with elements c_j^*. Then

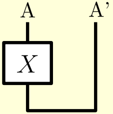
$|\psi\rangle =$ 

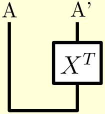
$|\psi^*\rangle =$ 

$\langle\psi| =$ 

$\langle\psi^*| =$ 

  - X^T is X upside-down:

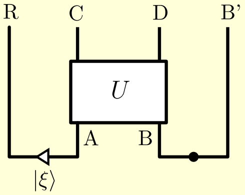


$=$


$= \sum_{j,k} X_{jk}|j,k\rangle$

When is the decoding possible?

State of the world: $|\widetilde{\Psi}\rangle =$



$\tilde{\rho} = |\widetilde{\Psi}\rangle\langle\widetilde{\Psi}|$

- Black hole has "forgotten" Alice's secret $\Leftrightarrow$ $\tilde{\rho}_{RC} \approx \tilde{\rho}_R \otimes \tilde{\rho}_C$
- Quantitative condition: Let

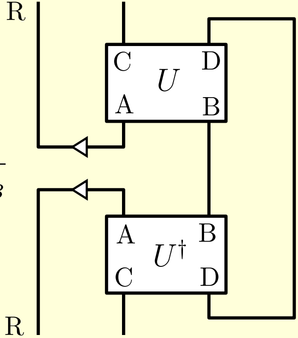
$$\delta = d_R d_C \operatorname{Tr} \tilde{\rho}_{RC}^2 - 1 \qquad (\delta \geqslant 0).$$

If $\delta \ll 1$, then Alice's secret can, in principle, be recovered from the Hawking radiation D and the purifying subsystem B'. In our algorithms (and in the original Hayden-Preskill work), δ determines the decoding fidelity.

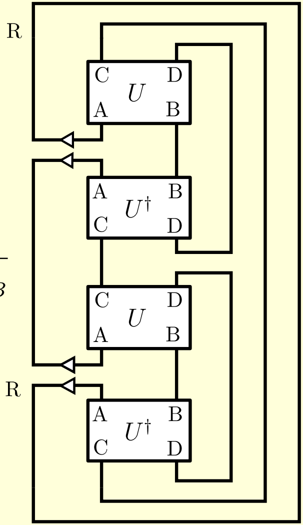
Calculation of $1+\delta = d_R d_C \,\mathrm{Tr}\,\tilde{\rho}_{RC}^2$

$\rho_{RC} = \mathrm{Tr}_{DB'} |\tilde{\Psi}\rangle\langle\tilde{\Psi}|$

$= \dfrac{1}{d_B}$

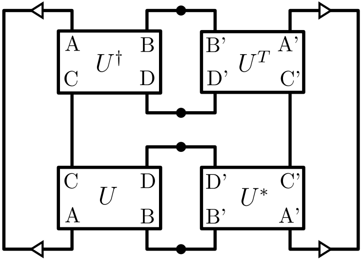


$\mathrm{Tr}\,\tilde{\rho}_{RC}^2 = \dfrac{1}{d_B^2}$



Expression for the fidelity parameter δ

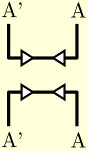
$$\delta = d_A d_R \Delta - 1\,, \qquad \Delta =$$



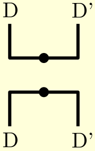
$$= \mathrm{OTOC}(L, M)$$

(Generalizes the result of P. Hosur, X.-L. Qi, D. Roberts, B. Yoshida, arXiv:1511.04021)

$$L = d_A$$



$$= \sum_j Y_j^T \otimes X_j\,, \qquad M =$$



$$= \sum_k W_k \otimes Z_k^T$$

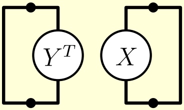
$$\mathrm{OTOC}(L, M) = \sum_{j,k} \frac{1}{d} \mathrm{Tr}\Big((U^\dagger W_k U)\, Y_j\, (U^\dagger Z_k U)\, X_j \Big)$$

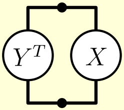
The late-time case

Assumption:

$$\langle \underbrace{W(t)}_{UWU^\dagger}\, Y(0)\, \underbrace{Z(t)}_{UZU^\dagger}\, X(0) \rangle \approx \langle WZ\rangle\langle Y\rangle\langle X\rangle + \langle W\rangle\langle Z\rangle\langle YX\rangle - \langle W\rangle\langle Z\rangle\langle Y\rangle\langle X\rangle$$

Used in calculations:

$\langle Y\rangle\langle X\rangle =$ 

$\langle YX\rangle =$ 

Result:

$$\Delta \approx \frac{1}{d_A d_R} + \frac{1}{d_D^2} - \frac{1}{d_A d_R d_D^2} \quad \Rightarrow \quad \delta = d_A d_R \Delta - 1 \leqslant \frac{d_A d_R}{d_D^2}$$

How hard is the decoding?

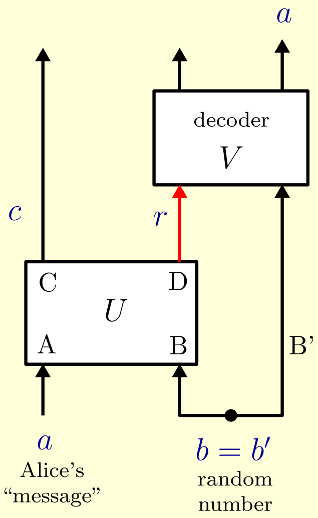
- The complexity is at least linear in d_R (i.e. exponential in the message size). Indeed, let $d_A = d_R$ and consider the classical case:

$$(c, r) = u(a, b)$$

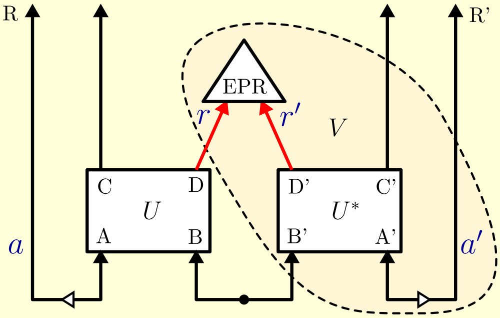
discarded (→ c) known to Bob (→ b)

Thus, $r = f(a)$, where f is random. The only general way to reconstruct a from r is exhaustive search.

- We show that the complexity is $O(d_A d_R \mathcal{C})$, where $\mathcal{C}$ is the size of the circuit for U.

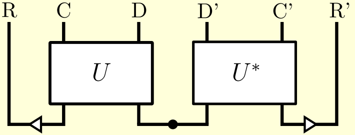


Probabilistic decoder



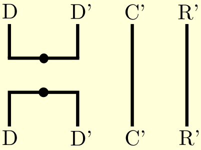
Classically, it's just random guessing. The Hawking radiation is $r = f(a)$; Bob picks a random a', computes $r' = f(a')$, and compares it with r.

$|\Psi_{\rm in}\rangle =$



Projector onto $|{\rm EPR}_{DD'}\rangle$,

$P_D =$



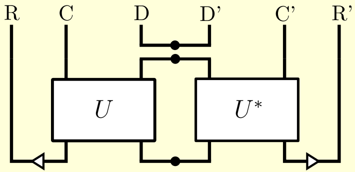
succeeds with probability

$$\langle \Psi_{\rm in}(I_{RC} \otimes P_D)\Psi_{\rm in}\rangle = \Delta \geqslant (d_A d_R)^{-1}$$

Fidelity of probabilistic decoding

Projected state:

$$|\Psi_{\rm out}\rangle = \frac{1}{\sqrt{\Delta}}(I_{RC}\otimes P_D)|\Psi_{\rm in}\rangle = \frac{1}{\sqrt{\Delta}}$$

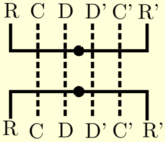


Fidelity:

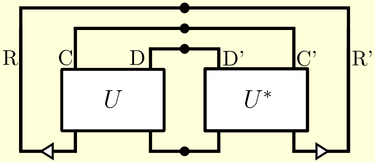
$$F = \langle\Psi_{\rm out}|P_R|\Psi_{\rm out}\rangle = \Delta^{-1}\langle\Psi_{\rm in}|\underbrace{P_R(I_{RC}\otimes P_D)}_{\geqslant|\mathrm{EPR}\rangle\langle\mathrm{EPR}|_{CD}}|\Psi_{\rm in}\rangle \geqslant \frac{1}{d_A d_R \Delta} = \boxed{\frac{1}{1+\delta}}$$

where

$$P_R =$$



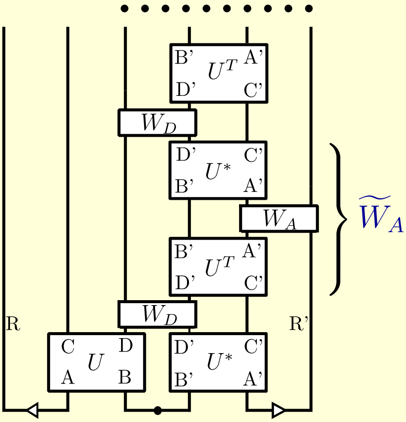
$$\langle\mathrm{EPR}|\Psi_{\rm in}\rangle =$$



$$= \frac{1}{\sqrt{d_A d_R}}$$

Deterministic decoder

- Uses Grover search to turn $|\Psi_{\rm in}\rangle$ to $|\Psi_{\rm out}\rangle$ without projection

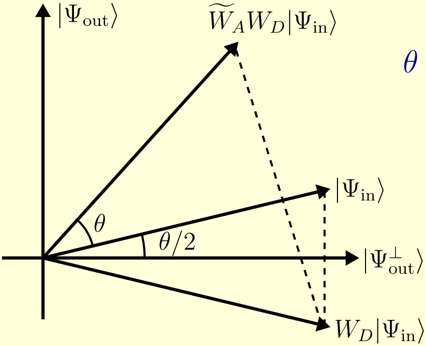


$\sim \sqrt{d_A d_R}$ iterations

Very roughly,

$$\widetilde{W}_A = 2|\Psi_{\rm in}\rangle\langle\Psi_{\rm in}| - 1$$

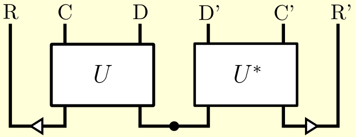
$$W_D = 1 - 2|\Psi_{\rm out}\rangle\langle\Psi_{\rm out}|$$



$$\theta \approx \frac{2}{\sqrt{d_A d_R}}$$

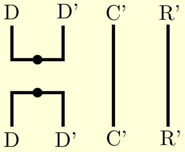
More accurate description of the algorithm

1) Apply U^* to produce $|\Psi_{\text{in}}\rangle =$

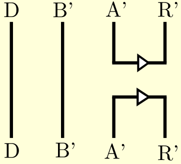


2) Let

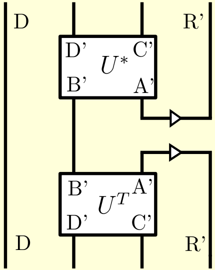
$P_D =$



$P_A =$



$\widetilde{P}_A =$



$$W_D = 1 - 2P_D, \qquad \widetilde{W}_A = 2\widetilde{P}_A - 1$$

3) Apply $\widetilde{W}_A W_D$ repeatedly $\dfrac{\pi}{2\theta_*}$ times, where $\theta_* = 2\arcsin\big((d_A d_R)^{-1/2}\big)$.

Analysis of the algorithm

- Let $\quad \widetilde{P}_A P_D \widetilde{P}_A = \Pi = \underbrace{\sum_{j=1}^{r} \alpha_j |\psi_j\rangle\langle\psi_j|}_{\text{eigenvalue decomposition, } \alpha_j > 0}, \qquad |\Psi_{\text{in}}\rangle = \sum_{j=1}^{r} \sqrt{p_j}\, \underbrace{|\eta_j\rangle_{RC} \otimes |\psi_j\rangle}_{|\Psi_j\rangle}$

Then each vector $|\Psi_j\rangle$ evolves under $I_{RC} \otimes (\widetilde{W}_A W_D)^m$ in a two-dimensional subspace with basis vectors $|\Phi_j\rangle$, $|\Phi_j^\perp\rangle$.

$$|\Psi(m)\rangle = \sum_{j=1}^{r} \sqrt{p_j} \Big(\sin\big((m + \tfrac{1}{2})\theta_j\big)\, |\Phi_j\rangle + \cos\big((m + \tfrac{1}{2})\theta_j\big)\, |\Phi_j^\perp\rangle \Big)$$

where $\quad \theta_j = 2 \arcsin \sqrt{\alpha_j}$

- We show that $\qquad r \leqslant d_R d_C, \quad \sum_{j=1}^{r} \alpha_j = \dfrac{d_C}{d_A}, \quad \sum_{j=1}^{r} \alpha_j^2 = \dfrac{d_C}{d_A}\, \Delta.$

  If $\delta = d_A d_R \Delta - 1 = 0$ (ideal case), then $\alpha_j = (d_A d_R)^{-1/2}$ for all j.

Analysis of the algorithm (cont.)

- Let $\delta = d_A d_R \Delta - 1, \quad m_* = \pi/(2\theta_*),$ where $\theta_* = 2\arcsin\big((d_A d_R)^{-1/2}\big)$.

  Then $\big(m_* + \frac{1}{2}\big)\theta_j \approx \frac{\pi}{2}$,

$$|\Psi(m_*)\rangle \approx \sum_{j=1}^{r} \sqrt{p_j}|\Phi_j\rangle \approx |\Psi_{\text{out}}\rangle\,; \quad \text{the Euclidean distance is } O\big(\sqrt{\delta}\big).$$

- Conclusions:

  – The algorithm involves $O\big(\sqrt{d_A d_R}\big)$ applications of U^* and U^T.

  – The fidelity of the reconstructed state $|\Psi(m_*)\rangle$ is $O(\delta)$. Recall that in the case of almost perfect scrambling,

$$\delta \leqslant \frac{d_R d_A}{d_D^2}$$

Open questions

1. How to generalize the algorithm to thermal density matrices? We can do it under these unrealistic assumptions:

$$\rho_{AB} = \rho_A \otimes \rho_B, \qquad \rho_{CD} = \rho_C \otimes \rho_D, \qquad \rho_{CD} = U \rho_{AB} U^\dagger.$$

2. In the traversible wormhole story, (Gao, Jafferis, Wall 2016; Maldacena, Stanford, Yang 2017), the decoding happens in one go. What are the necessary/sufficient conditions in terms of OTOCs?

3. The Grover iterations bear some similarity with multiple shocks (Shenker, Stanford 2014). What is the exact relation?