Title: A separation of out-of-time-ordered correlator and entanglement

Speakers: Linghang Kong

Series: Perimeter Institute Quantum Discussions

Date: August 08, 2019 - 4:00 PM

URL: http://pirsa.org/19080079

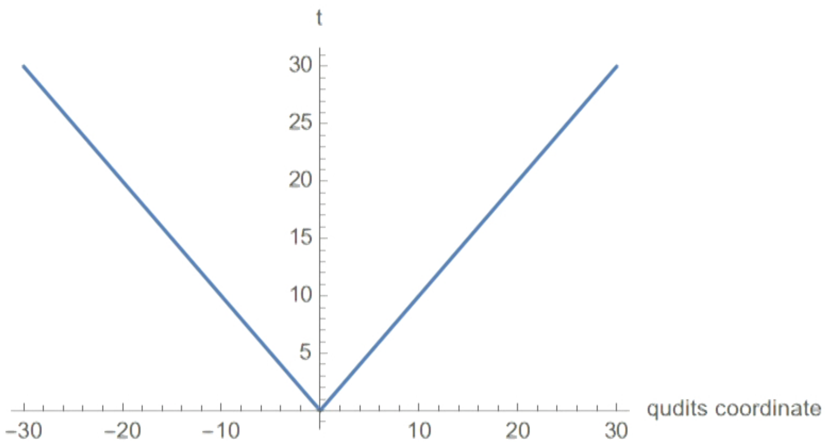
Abstract: The out-of-time-ordered correlator (OTOC) and entanglement are two physically motivated and widely used probes of the ``scrambling" of quantum information, which has drawn great interest recently in quantum gravity and many-body physics. By proving upper and lower bounds for OTOC saturation on graphs with bounded degree and a lower bound for entanglement on general graphs, we show that the time scales of scrambling as given by the growth of OTOC and entanglement entropy can be asymptotically separated in a random quantum circuit model defined on graphs with a tight bottleneck. Our result counters the intuition that a random quantum circuit mixes in time proportional to the diameter of the underlying graph of interactions. It also serves as a more rigorous justification for an argument of [Shor, 1807.04363], that black holes may be very slow scramblers in terms of entanglement generation. Such observations may be of fundamental importance in the understanding of the black hole information problem. The bound we obtained for OTOC is interesting in its own right in that it generalized previous studies of OTOC on lattices to the geometries on graphs and proved it rigorously.   Based on [Harrow-Kong-Liu-Mehraban-Shor, 1906.02219]

Scrambling

- Scrambling describes a quantum phenomenon that
  - Dynamically: initially localized information gets spread onto the whole quantum system so that it can no longer be retrieved from local measurements
  - Kinematically: the overall evolution and state of the system "looks" random

Probes of Scrambling

- Operator growth — how local perturbations acting on one part of the system spreads; quantified by OTOC



正在共享桌面 停止共享

Probes of Scrambling

- Operator growth — how local perturbations acting on one part of the system spreads; quantified by OTOC
- Entanglement growth — scrambling happens because the generation of global multipartite entanglement "hides" information from any local measurement; quantified by entanglement entropies (von Neumann, Rényi...)

OTOC

- Let $O_1(x, 0)$ be an operator acting on qudit x
- $O_2(y, t)$ be an Heisenberg operator at time t that only acts on qudit y at time 0

OTOC

- Let $O_1(x,0)$ be an operator acting on qudit x
- $O_2(y,t)$ be an Heisenberg operator at time t that only acts on qudit y at time 0
- The commutator form of out-of-time-ordered correlator (OTOC) is defined to be

$$C(t) = \frac{1}{2}\langle [O_1(x,0), O_2(y,t)]^\dagger [O_1(x,0), O_2(y,t)]\rangle_\beta$$

- The average is taken w.r.t. thermal state with inverse temperature β, which we take to be 0 in this work (infinite temperature)

正在共享桌面 停止共享

Scrambling Time Scale

- After an enough amount of time, the unitary produced by the circuit will be random enough

正在共享桌面 停止共享

Scrambling Time Scale

- After an enough amount of time, the unitary produced by the circuit will be random enough
- The OTOC between x and y will become close to the equilibrium value $\Theta(1)$
- The entanglement entropy between a set of vertices A and its complement $\bar{A}$ will be close to the equilibrium value $\Theta(\min\{|A|, |\bar{A}|\})$

正在共享桌面 停止共享

Scrambling Time Scale

- After an enough amount of time, the unitary produced by the circuit will be random enough
- The OTOC between x and y will become close to the equilibrium value $\Theta(1)$
- The entanglement entropy between a set of vertices A and its complement $\bar{A}$ will be close to the equilibrium value $\Theta(\min\{|A|,|\bar{A}|\})$
- We define $\tau_{\text{OTOC}}^{(x,y)}$ and τ_{ent}^{A} as the time needed for OTOC and entanglement to reach a constant times their equilibrium value

正在共享桌面 停止共享

Scrambling Time Scale

- After an enough amount of time, the unitary produced by the circuit will be random enough
- The OTOC between x and y will become close to the equilibrium value $\Theta(1)$
- The entanglement entropy between a set of vertices A and its complement $\bar{A}$ will be close to the equilibrium value $\Theta(\min\{|A|, |\bar{A}|\})$
- We define $\tau_{\text{OTOC}}^{(x,y)}$ and τ_{ent}^{A} as the time needed for OTOC and entanglement to reach a constant times their equilibrium value
- For a given graph, we characterize its properties with $\tau_{\text{OTOC}}^{(x,y)}$ and τ_{ent}^{A} with (x, y) and A that maximizes them respectively

正在共享桌面 停止共享

Random Circuit Model

- Consider a graph G with V vertices and E edges. d-dimensional qudits placed on each vertex

Random Circuit Model

- Consider a graph G with V vertices and E edges. d-dimensional qudits placed on each vertex
- On average, a Haar-random 2-qudit unitary is acted on the vertices connected by each edge in each time unit
  - A Haar-random unitary is uniformly random over the group $U(d^2)$

正在共享桌面 停止共享

Random Circuit Model

- Consider a graph G with V vertices and E edges. d-dimensional qudits placed on each vertex
- On average, a Haar-random 2-qudit unitary is acted on the vertices connected by each edge in each time unit
  - A Haar-random unitary is uniformly random over the group $U(d^2)$
  - More precisely, unitaries are applied according to independent Poisson distributions

正在共享桌面 停止共享

Random Circuit Model

- Consider a graph G with V vertices and E edges. d-dimensional qudits placed on each vertex
- On average, a Haar-random 2-qudit unitary is acted on the vertices connected by each edge in each time unit
  - A Haar-random unitary is uniformly random over the group $U(d^2)$
  - More precisely, unitaries are applied according to independent Poisson distributions

正在共享桌面 停止共享

Main Result

- Let G be a graph with V vertices and E edges. Consider any pair of vertices x and y with $D(x, y)$ being the distance between them. Suppose the degree for each vertex is at most d^2, we have shown that $\tau_{\text{OTOC}}^{(x,y)} = \Theta(D(x, y))$

正在共享桌面 停止共享

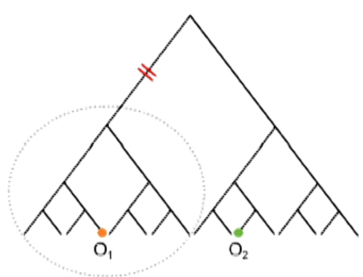
Main Result

- Let G be a graph with V vertices and E edges. Consider any pair of vertices x and y with $D(x, y)$ being the distance between them. Suppose the degree for each vertex is at most d^2, we have shown that $\tau_{\text{OTOC}}^{(x,y)} = \Theta(D(x, y))$
- Suppose the graph is partitioned into A and $\bar{A}$. We have also shown that $\tau_{\text{ent}}^{A} = \Omega\left(\frac{\min\{|A|, |\bar{A}|\}}{C(A, \bar{A})}\right)$, where $C(A, \bar{A})$ is the number of edges with one endpoint in A and one in $\bar{A}$

正在共享桌面 停止共享

Main Result

- Let G be a graph with V vertices and E edges. Consider any pair of vertices x and y with $D(x,y)$ being the distance between them. Suppose the degree for each vertex is at most d^2, we have shown that $\tau_{\text{OTOC}}^{(x,y)} = \Theta(D(x,y))$
- Suppose the graph is partitioned into A and $\bar{A}$. We have also shown that $\tau_{\text{ent}}^{A} = \Omega\left(\frac{\min\{|A|,|\bar{A}|\}}{C(A,\bar{A})}\right)$, where $C(A,\bar{A})$ is the number of edges with one endpoint in A and one in $\bar{A}$
- Now we apply the bounds above to a perfect binary tree



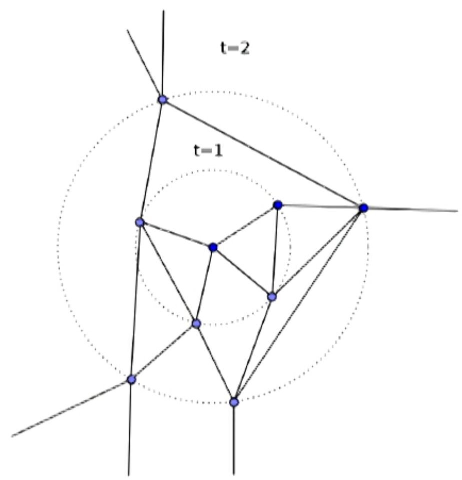
- $\tau_{\text{OTOC}} = \Theta(\ln V)$ and $\tau_{\text{ent}} = \Omega(V)$ where V is the number of vertices

Implications

1. Scrambling in non-Euclidean geometries
   - Existing work has studied scrambling mostly on Euclidean lattices. The general assumption is that after time t, a localized perturbation will affect everything within some ball of radius $v_{\text{butterfly}}t$
   - However, this has not been proved and previous works only gave heuristic arguments for it that included uncontrolled approximations
   - For the random circuit models defined on general graphs, we find that if the graph degree is small then indeed there is a linear butterfly velocity

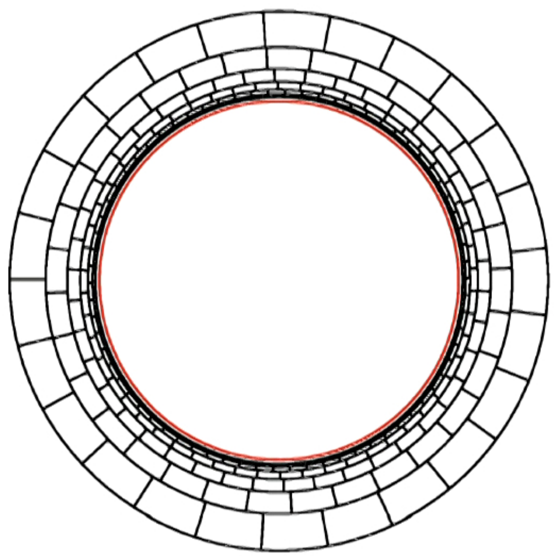
Implications

1. Scrambling in non-Euclidean geometries



Implications

2. Black hole information scrambling
   - Can be viewed as a more rigorous argument for [1807.04363]



Implications

2. Black hole information scrambling
   - Can be viewed as a more rigorous argument for [1807.04363]
   - OTOC saturation is needed for Yoshida-Kitaev protocol [1710.03363] of Hayden-Preskill decoding task

Implications

3. Counterexample of constructing 2-designs in time proportional to diameter

- A 2-design is a distribution that approximately agree with the Haar measure up to the first two moments (important in quantum information; admit efficient constructions)
- Having 2-designs implies saturation of OTOC, entanglement, etc.
- Therefore our results implies that 2-designs might not be achieved in time $\propto$ diameter

正在共享桌面 停止共享

Outline

Markov Chain

- Using the properties of Poisson distribution, $\Theta(Et)$ edges are picked in time t with high probability. Here E is the total number of edges

Markov Chain

- Using the properties of Poisson distribution, $\Theta(Et)$ edges are picked in time t with high probability. Here E is the total number of edges
- Each edge picked is uniformly random among all edges

Markov Chain

- Using the properties of Poisson distribution, $\Theta(Et)$ edges are picked in time t with high probability. Here E is the total number of edges
- Each edge picked is uniformly random among all edges
- Equivalent to picking a random edge every $1/E$ unit of time, up to a constant factor

正在共享桌面 停止共享

Markov Chain

- Using the properties of Poisson distribution, $\Theta(Et)$ edges are picked in time t with high probability. Here E is the total number of edges
- Each edge picked is uniformly random among all edges
- Equivalent to picking a random edge every $1/E$ unit of time, up to a constant factor
- From now on the process of picking an edge and apply a unitary is called a "step", and the time elapsed will be the number of steps divided by E

Markov Chain

- The Heisenberg operator can be expanded into the Pauli basis $\{\sigma_{\vec{p}} | \vec{p} \in \{0, 1, \ldots, d^2 - 1\}^n\}$

$$U^\dagger \sigma_{\vec{p}} U = \sum_{\vec{q}} \alpha_{\vec{q}} \sigma_{\vec{q}}, \quad \alpha_{\vec{q}} \equiv \frac{1}{d^n} \mathrm{Tr}[U^\dagger \sigma_{\vec{p}} U \sigma_{\vec{q}}^\dagger]$$

正在共享桌面 停止共享

Markov Chain

- The Heisenberg operator can be expanded into the Pauli basis $\{\sigma_{\vec{p}} | \vec{p} \in \{0, 1, \ldots, d^2 - 1\}^n\}$

$$U^\dagger \sigma_{\vec{p}} U = \sum_{\vec{q}} \alpha_{\vec{q}} \sigma_{\vec{q}}, \quad \alpha_{\vec{q}} \equiv \frac{1}{d^n} \mathrm{Tr}[U^\dagger \sigma_{\vec{p}} U \sigma_{\vec{q}}^\dagger]$$

- $\mathbb{E}_U \alpha_{\vec{q}} \alpha_{\vec{q}'}^*$ transforms linearly in each step. Due to the properties of the distribution in each step, this is zero for $\vec{q} \neq \vec{q}'$

正在共享桌面 停止共享

Markov Chain

- The Heisenberg operator can be expanded into the Pauli basis $\{\sigma_{\vec{p}} | \vec{p} \in \{0, 1, \ldots, d^2 - 1\}^n\}$

$$U^\dagger \sigma_{\vec{p}} U = \sum_{\vec{q}} \alpha_{\vec{q}} \sigma_{\vec{q}}, \quad \alpha_{\vec{q}} \equiv \frac{1}{d^n} \mathrm{Tr}[U^\dagger \sigma_{\vec{p}} U \sigma_{\vec{q}}^\dagger]$$

- $\mathbb{E}_U \alpha_{\vec{q}} \alpha_{\vec{q}'}^*$ transforms linearly in each step. Due to the properties of the distribution in each step, this is zero for $\vec{q} \neq \vec{q}'$
- The quantities $\mathbb{E}_U \alpha_{\vec{q}} \alpha_{\vec{q}}^*$ sum to 1, and can be viewed as a probability distribution. The process can be described by a Markov chain on all the Pauli operators

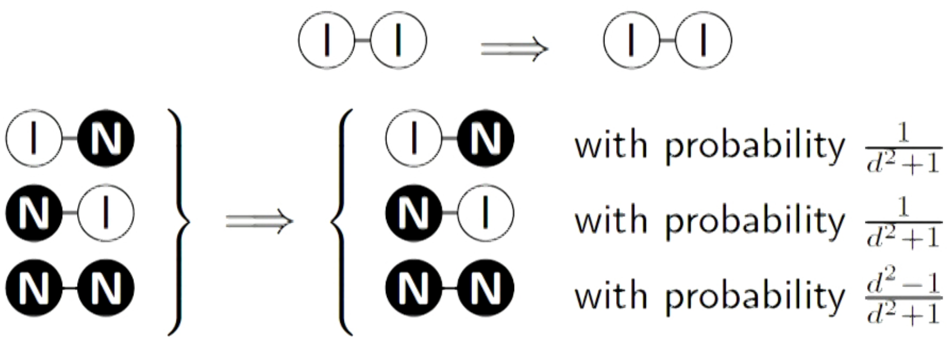
正在共享桌面 停止共享

Markov Chain

- On each vertex the non-identity Pauli operators have the same probability, so we only care if the operator on a vertex is identity (I) or non-identity (N)
  - i.e. the state space simplifies from $(d^2)^n$ Pauli operators to 2^n configurations in which each vertex is assigned a label "N" or "I"

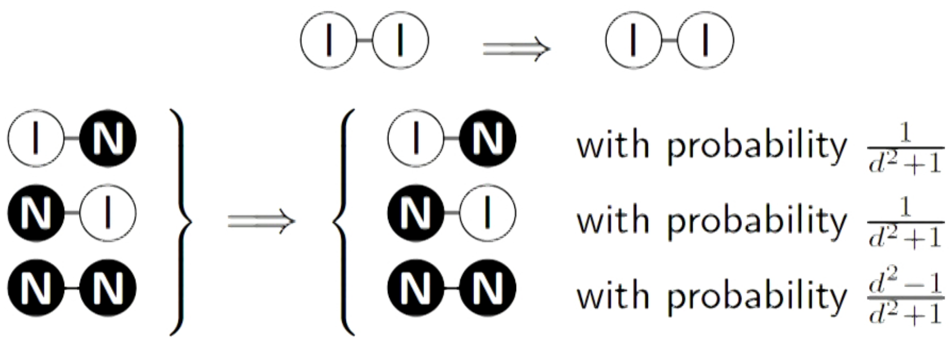
Markov Chain

- On each vertex the non-identity Pauli operators have the same probability, so we only care if the operator on a vertex is identity (I) or non-identity (N)
- The update rule is shown as follows



Markov Chain

- On each vertex the non-identity Pauli operators have the same probability, so we only care if the operator on a vertex is identity (I) or non-identity (N)
- The update rule is shown as follows



- For OTOC with $O_1(x,0)$ and $O_2(y,t)$, the Markov chain starts from a single "N" on y, and OTOC is proportional to the probability of having "N" on x

正在共享桌面 停止共享

Upper Bound

Let G be a graph with V vertices and E edges. The maximum degree of vertices is z. Consider any pair of vertices x and y with $D(x, y)$ being the distance between them.

Theorem (OTOC Upper Bound)

Suppose $z \leq d^2$*, where* d *is the Hilbert space dimension for each vertex. Then* $\tau_{OTOC}^{(x,y)} = O(D(x, y))$ *with probability at least* $1 - e^{-\Theta(D(x,y))}$.

正在共享桌面 停止共享

Upper Bound

- The first part of the proof is to show that a label "N" reaches vertex x in $O(E \cdot D(x,y))$ number of steps with high probability

正在共享桌面 停止共享

Upper Bound

- The first part of the proof is to show that a label "N" reaches vertex x in $O(E \cdot D(x, y))$ number of steps with high probability
- To show this we consider a modified Markov chain that keeps only one label "N" that is closest to vertex x, and other labels "N" are set to "I" after each step

正在共享桌面 停止共享

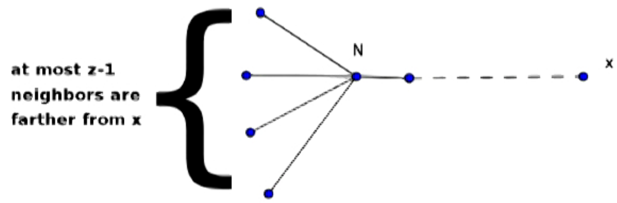
Upper Bound

- The first part of the proof is to show that a label "N" reaches vertex x in $O(E \cdot D(x, y))$ number of steps with high probability
- To show this we consider a modified Markov chain that keeps only one label "N" that is closest to vertex x, and other labels "N" are set to "I" after each step
- The number of steps needed in this modified chain gives an upper bound for that in the original chain

正在共享桌面 停止共享

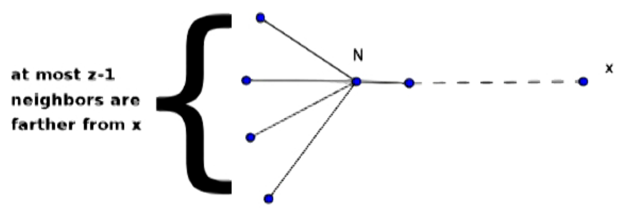
Upper Bound

- The distance between this "N" label and vertex x increases by 1 with probability at most $\frac{z-1}{E(d^2+1)}$, and decreases by 1 with probability at least $\frac{d^2}{E(d^2+1)}$



Upper Bound

- The distance between this "N" label and vertex x increases by 1 with probability at most $\frac{z-1}{E(d^2+1)}$, and decreases by 1 with probability at least $\frac{d^2}{E(d^2+1)}$



- As long as $z \le d^2$, this distance tend to decrease in each step and becomes 0 in $O(E \cdot D(x, y))$ number of steps with high probability

Upper Bound

- The next step is to show that the probability for having label "N" on vertex x is constant after it reaches x

正在共享桌面 停止共享

Upper Bound

- The next step is to show that the probability for having label "N" on vertex x is constant after it reaches x
- Again we study the Markov chain with a single "N"
- The distance between this label "N" and vertex x still tends to decrease in each step, so the probability for this distance to be 0 is constant in the equilibrium distribution as long as $z \leq d^2$
- It could be shown that the probability is monotonically non-increasing w.r.t. the number of steps, and therefore is lower bounded by a constant

正在共享桌面 停止共享

Lower Bound

Let G be a graph with V vertices and E edges. The maximum degree of vertices is z. Consider any pair of vertices x and y with $D(x, y)$ being the distance between them.

Theorem (OTOC Lower Bound)

Suppose $z = O(1)$. *Then* $\tau_{OTOC}^{(x,y)} = \Omega(D(x, y))$ *with probability at least* $1 - e^{-\Theta(D(x,y))}$.

This bound works as long as the degree is constant, but it becomes weaker (in terms of the constant hidden in $\Omega(\cdot)$) when the degree increases.

Lower Bound

- The lower bound is proved using a union bound
- Consider a path starting from vertex y with length D. For any $\lambda < 1$, the probability for labels "N" to propagate through this path within λDE steps is $e^{-\Theta(D)}$

正在共享桌面 停止共享

Lower Bound

- The lower bound is proved using a union bound
- Consider a path starting from vertex y with length D. For any $\lambda < 1$, the probability for labels "N" to propagate through this path within λDE steps is $e^{-\Theta(D)}$
- For any integer D, there are at most $(z-1)^D$ paths between vertex y and any vertex with distance D away from y

正在共享桌面 停止共享

Lower Bound

- The lower bound is proved using a union bound
- Consider a path starting from vertex y with length D. For any $\lambda < 1$, the probability for labels "N" to propagate through this path within λDE steps is $e^{-\Theta(D)}$
- For any integer D, there are at most $(z-1)^D$ paths between vertex y and any vertex with distance D away from y
- By union bound, the probability for any vertex x with distance D away from y to get a label "N" is at most $(z-1)^D \cdot e^{-\Theta(D)}$, and is exponentially small in D if we pick a small λ

Outline

Introduction

Our Results

Statement of the results

Implications

Proof Techniques for OTOC

Conversion into Markov Chain

Upper Bound

Lower Bound

Proof Techniques for Entanglement

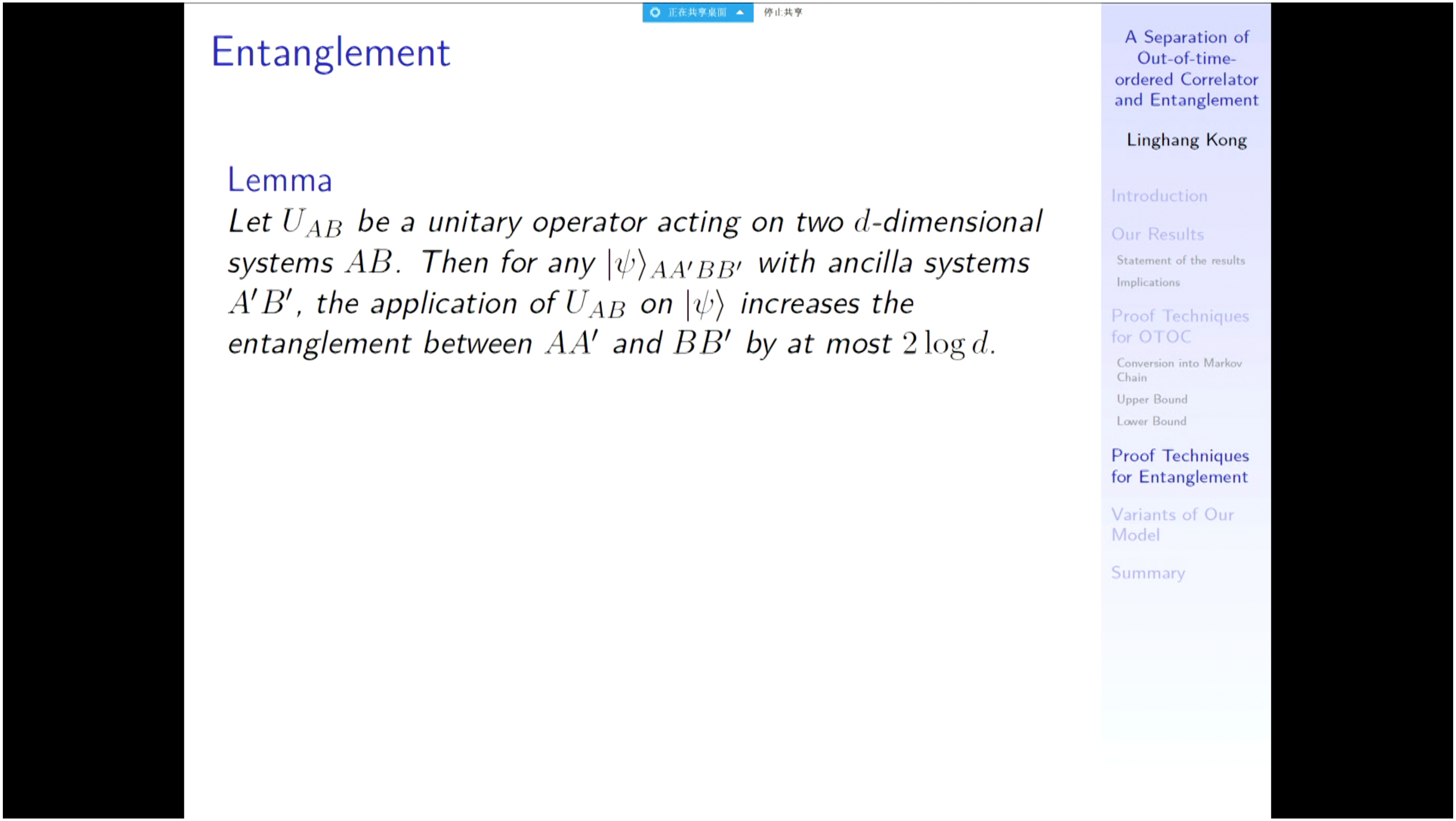
Variants of Our Model

正在共享桌面 停止共享

Statement of the Bound

Theorem (Entanglement Lower Bound)

For a general graph G with vertices partitioned into sets A and $\bar{A}$, the entanglement saturation time is $\Omega\left(\frac{\min\{|A|,|\bar{A}|\}}{C(A,\bar{A})}\right)$, where $C(A,\bar{A})$ is the number of edges with one endpoint in A and one in $\bar{A}$.


正在共享桌面
停止共享
Entanglement
Lemma
Let U_{AB} be a unitary operator acting on two d-dimensional systems AB. Then for any $|\psi\rangle_{AA'BB'}$ with ancilla systems $A'B'$, the application of U_{AB} on $|\psi\rangle$ increases the entanglement between AA' and BB' by at most $2\log d$.
A Separation of Out-of-time-ordered Correlator and Entanglement
Linghang Kong
Introduction
Our Results
Statement of the results
Implications
Proof Techniques for OTOC
Conversion into Markov Chain
Upper Bound
Lower Bound
Proof Techniques for Entanglement
Variants of Our Model
Summary

Entanglement

Lemma

Let U_{AB} be a unitary operator acting on two d-dimensional systems AB. Then for any $|\psi\rangle_{AA'BB'}$ with ancilla systems $A'B'$, the application of U_{AB} on $|\psi\rangle$ increases the entanglement between AA' and BB' by at most $2\log d$.

Proof.

Suppose Alice hold AA' and Bob holds BB' parts of $|\psi\rangle$. Besides that they also share 2 copies of EPR pairs. Consider the LOCC protocol in which Alice teleports register A to Bob, followed by Bob's local application of U_{AB} and teleporting A back to Alice. The entanglement is non-increasing under this LOCC protocol, which gives the bound. □

正在共享桌面 停止共享

Entanglement

- The theorem above shows that the entanglement between subsystem A and $\bar{A}$ increases by at most $2 \log d$ when a cut edge between A and $\bar{A}$ is selected
- The equilibrium value of entanglement is $\Theta(\min\{|A|, |\bar{A}|\} \log d)$
- If there are $C(A, \bar{A})$ cut edges in total, the time needed will be $\Omega(\frac{\min\{|A|,|\bar{A}|\}}{C(A,\bar{A})})$

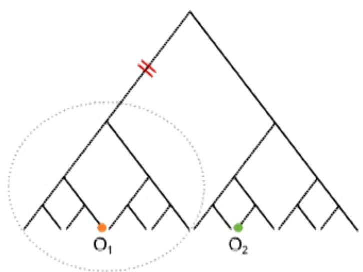
Different Order

- One might argue that the slow entanglement saturation might be due to the low chance of picking the cut edge in the binary tree

正在共享桌面 停止共享

Different Order

- One might argue that the slow entanglement saturation might be due to the low chance of picking the cut edge in the binary tree
- We considered an alternative model which picks one edge on one side of the cut edge, then picks the cut edge, and finally picks an edge from the other side



Different Order

- One might argue that the slow entanglement saturation might be due to the low chance of picking the cut edge in the binary tree
- We considered an alternative model which picks one edge on one side of the cut edge, then picks the cut edge, and finally picks an edge from the other side
- There is still a separation in this model

正在共享桌面 停止共享

Different Order

- One might argue that the slow entanglement saturation might be due to the low chance of picking the cut edge in the binary tree
- We considered an alternative model which picks one edge on one side of the cut edge, then picks the cut edge, and finally picks an edge from the other side
- There is still a separation in this model
- The reason is that the unitary acted on the cut edge can commute through other unitaries unless they share a vertex. In this way $\Theta(V)$ unitaries on the cut edge will collapse into one

正在共享桌面 停止共享

Different Interaction Geometries

- Instead of binary trees, we could also consider two complete graphs connected by a bottleneck edge

Different Interaction Geometries

- Instead of binary trees, we could also consider two complete graphs connected by a bottleneck edge
- Even though the graph has high degree and our OTOC bound does not hold, a separation of OTOC and entanglement could still be shown, both in the original Poisson clock order and the one shown on last page

正在共享桌面 停止共享

Different Interaction Geometries

- Instead of binary trees, we could also consider two complete graphs connected by a bottleneck edge
- Even though the graph has high degree and our OTOC bound does not hold, a separation of OTOC and entanglement could still be shown, both in the original Poisson clock order and the one shown on last page
- This geometry allows a third order of choosing the edges, which is to choose a random perfect match in the complete graphs followed by choosing the cut edge. Still a separation can be shown

正在共享桌面 停止共享

High Degree Graphs

- In graphs with high degrees, the OTOC time can significantly deviate from the linear bound we obtained

A Separation of Out-of-time-ordered Correlator and Entanglement

Linghang Kong

High Degree Graphs

- In graphs with high degrees, the OTOC time can significantly deviate from the linear bound we obtained
- One example is star graph, which has constant diameter but $\log n$ OTOC time [1903.01468]

High Degree Graphs

- In graphs with high degrees, the OTOC time can significantly deviate from the linear bound we obtained
- One example is star graph, which has constant diameter but $\log n$ OTOC time [1903.01468]
- Another example is to consider a perfect tree with high degree $z \gg d^2 \gg 1$. Starting from one leaf, the label "N" has to fill the entire subtree when climbing up to a new height, thus taking a significant amount of time to reach the root

正在共享桌面 停止共享

High Degree Graphs

- On the other hand, for a complete graph the OTOC time is $O(\log n/n)$ while the diameter is 1

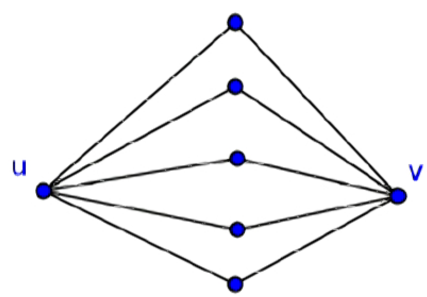
正在共享桌面 停止共享

High Degree Graphs

- On the other hand, for a complete graph the OTOC time is $O(\log n/n)$ while the diameter is 1
- By replacing each edge by k-fold parallel edges, the OTOC on any graph can be scaled down by a factor of $\Theta(1/k)$

High Degree Graphs

- On the other hand, for a complete graph the OTOC time is $O(\log n/n)$ while the diameter is 1
- By replacing each edge by k-fold parallel edges, the OTOC on any graph can be scaled down by a factor of $\Theta(1/k)$
- Alternatively one can replace each edge (u, v) into the following object



Summary

- We have shown linear upper and lower bounds for OTOC saturation for random circuits on general graphs with low degree

Summary

- We have shown linear upper and lower bounds for OTOC saturation for random circuits on general graphs with low degree
- Together with a lower bound for entanglement, we show a separation between the scrambling time scale defined using OTOC and entanglement

Thank you for your attention!