Quantum Rewinding Tutorial Part 2:

How to Run a Quantum Attacker Many Times (or: The Unreasonable Effectiveness of Alternating Projectors)

> Fermi Ma (Simons & Berkeley)

Based on:

• "Post-Quantum Succinct Arguments: Breaking the Quantum Rewinding Barrier" by Alessandro Chiesa, Fermi Ma, Nicholas Spooner, and Mark Zhandry (2021)

However, this technique has some major drawbacks:

However, this technique has some major drawbacks:

We get bad soundness guarantees (can't rule out a quantum prover that breaks Blum with probability 0.7)

However, this technique has some major drawbacks:

- We get bad soundness guarantees (can't rule out a quantum prover that breaks Blum with probability 0.7)
- More serious issue: this technique only applies to a *very limited class* of protocols (e.g., Blum but not [GMW86] 3-coloring)

However, this technique has some major drawbacks:

- We get bad soundness guarantees (can't rule out a quantum prover that breaks Blum with probability 0.7)
- More serious issue: this technique only applies to a *very limited class* of protocols (e.g., Blum but not [GMW86] 3-coloring)

Plan for this talk: we'll see a significantly more powerful rewinding technique due to [CMSZ21].

Motivating example: **Succinct Arguments for NP**

Motivating example: **Succinct Arguments for NP**

"Succinct" = communication $+$ verifier efficiency is $poly(\lambda, \log(|x| + |w|))$

"Succinct" = communication $+$ verifier efficiency is $poly(\lambda, \log(|x| + |w|))$

"Argument" = sound against *efficient* cheating

[Kilian92] constructs a 4-message succinct argument for NP from collision-resistant hash functions (CRHFs).

[Kilian92] constructs a 4-message succinct argument for NP from collision-resistant hash functions (CRHFs).

Many applications: universal arguments [BG01], zero knowledge [Barak01], SNARGs [Micali94, BCS16], …

Extra Motivation: studying quantum rewinding for succinct arguments will force us to develop general-purpose techniques.

Extra Motivation: studying quantum rewinding for succinct arguments will force us to develop general-purpose techniques.

• Typically prove soundness using several transcripts to specify a witness.

Extra Motivation: studying quantum rewinding for succinct arguments will force us to develop general-purpose techniques.

- Typically prove soundness using several transcripts to specify a witness.
- Succinct arguments inherently require many transcripts to specify a witness, so *lots* of rewinding is required.

Let's see how Kilian's protocol works

Compile a *probabilistically checkable proof** (PCP) into an interactive argument system using cryptography.

*[BFLS91,FGLSS91,AS92,ALMSS92]

Compile a *probabilistically checkable proof** (PCP) into an interactive argument system using cryptography.

*[BFLS91,FGLSS91,AS92,ALMSS92]

Encode w as PCP π

ζ P $\hat{\beta}$ sends short commitment to PCP π .

Encode w as PCP π

$\mathsf{P} \stackrel{\cdot\cdot}{\mathcal{S}}$ sends short commitment to PCP π .

Encode w as PCP π

$\mathsf{P} \stackrel{\cdot\cdot}{\mathcal{S}}$ sends short commitment to PCP π .

Kilian's protocol $x, w \longrightarrow x$ CRHF h \Box P h h h h PCP π

$\mathsf{P} \stackrel{\cdot}{\mathcal{S}}$ sends short commitment to PCP π .

P \hat{B} sends short commitment to PCP π.

$P \hat{B}$ sends short commitment to PCP π .

$P(\hat{B})$ sends short commitment to PCP π .

PCP π

Intuition: want to show that the CRHF forces $\left\{ \right. \right. \left\{ \right. \right\}$ 3 to respond consistently with some PCP string π .

Intuition: want to show that the CRHF forces $\binom{m}{n}$ to respond consistently with some PCP string π .

Formalize by *rewinding* last two messages many times.

Reduction's goal: record *many* accepting transcripts (r_i, z_i)

Reduction's goal: record *many* accepting transcripts (r_i, z_i) Eventually finds impossible π OR collision. Pr[PCP verifier accepts π] > PCP soundness error

Define success probability as $p \coloneqq \Pr_{r \in R} [\text{ss}^2]$ wins]

Problem: |S′⟩ might not be a successful adversary!

Define *success probability* as $p \coloneqq \Pr$ $r \leftarrow R$ $\left[\frac{\partial^2 y}{\partial S}\right]$ wins]

Problem: |S′⟩ might not be a successful adversary!

This work: we devise a "repair" procedure to restore the original success probability.

Problem: |S′⟩ might not be a successful adversary!

This work: we devise a "repair" procedure to restore the original success probability.

First, recall a key idea from the first talk: As long as the prover's response is "collapsing", measuring the prover's response amounts to measuring the bit indicating accept/reject.

Naïve Measurement:

Measure Σ |z) right away.

Naïve Measurement: Measure Σ |z) right away.

"Lazy" Measurement: (1) Compute + measure $V(r, z)$. (2) Measure *z* if $V(r, z) = 1$.

Naïve Measurement:

Measure Σ |z) right away.

"Lazy" Measurement:

(1) Compute + measure $V(r, z)$. (2) Measure *z* if $V(r, z) = 1$.

[U16]: As long as z is "collapsing", measurement in step (2) causes *is undetectable to the prover*!

Naïve Measurement:

Measure Σ |z) right away.

"Lazy" Measurement:

(1) Compute + measure $V(r, z)$. (2) Measure z if $V(r, z) = 1$.

[U16]: As long as z is "collapsing", measurement in step (2) causes *is undetectable to the prover*!

Kilian's protocol satisfies this property if the CRHF is a "collapsing hash function".

Naïve Measurement: Measure Σ |z) right away.

"Lazy" Measurement: (1) Compute + measure $V(r, z)$. (2) Measure z if $V(r, z) = 1$.

As in the first talk, collapsing allows us to treat the measurement of the prover's response on r as a binary-outcome measurement $(\Pi_r, \mathbb{I} - \Pi_r)$

Rest of this talk: "repair" the prover's state after a binary-outcome measurement.

State repair task: Given $|S'\rangle$, efficiently produce

State repair task: Given $|S'\rangle$, efficiently produce a p -good adversary state.

We'll use the [MW05] alternating projectors idea.

State repair task: Given $|S'\rangle$, efficiently produce a p -good adversary state.

We'll use the [MW05] alternating projectors idea.

But which projectors do we use? Recall what we did last time.

State repair task: Given $|S'\rangle$, efficiently produce a p -good adversary state.

Watrous rewinding task: given verifier state $|\psi\rangle$ and projector Π_G indicating "successful simulation", output the state $\Pi_G|\psi\rangle_V|0\rangle_{Aux}$

State repair task: Given $|S'\rangle$, efficiently produce a p -good adversary state.

Watrous rewinding task: given verifier state $|\psi\rangle$ and projector Π_G indicating "successful simulation", output the state $\Pi_G|\psi\rangle_V|0\rangle_{Aux}$

Algorithm: alternate Π_0 , Π_6 measurements until Π_6 accepts.

State repair task: Given $|S'\rangle$, efficiently produce a p -good adversary state.

Watrous rewinding task: given verifier state $|\psi\rangle$ and projector Π_G indicating "successful simulation", output the state $\Pi_G|\psi\rangle_V|0\rangle_{Aux}$

Algorithm: alternate Π_0 , Π_6 measurements until Π_6 accepts.

Why these projectors?

State repair task: Given $|S'\rangle$, efficiently produce a p -good adversary state.

Watrous rewinding task: given verifier state $|\psi\rangle$ and projector Π_G indicating "successful simulation", output the state $\Pi_G|\psi\rangle_V|0\rangle_{Aux}$

Algorithm: alternate Π_0 , Π_6 measurements until Π_6 accepts.

Why these projectors?

- \lim_{Ω} (Π_0) contains the *initial state* $|\psi\rangle|0\rangle$
- \lim_{G} (Π_G) contains the *target state* Π _{*G}* $|\psi\rangle$ _{*V*} $|0\rangle$ _{*Aux*}</sub>

State repair task: Given $|S'\rangle$, efficiently produce a p -good adversary state.

How do we apply the [MW05,W05] approach to our setting?

State repair task: Given $|S'\rangle$, efficiently produce a p -good adversary state.

How do we apply the [MW05,W05] approach to our setting?

Oversimplification: suppose we can efficiently implement $(\Pi_p, \mathbb{I} - \Pi_p)$ where image(Π_p) exactly corresponds to p-good adversary states.

State repair task: Given $|S'\rangle$, efficiently produce a p -good adversary state.

How do we apply the [MW05,W05] approach to our setting?

Oversimplification: suppose we can efficiently implement $(\Pi_p, \mathbb{I} - \Pi_p)$ where image(Π_n) exactly corresponds to p-good adversary states.

$$
\begin{array}{c}\n\begin{array}{c}\n\hline\n\text{S} \\
\text{S}\n\end{array}\n\end{array}
$$
\nApply Π_r

\n(S')

Oversimplification: suppose we can efficiently implement $(\Pi_p, \mathbb{I} - \Pi_p)$ where image(Π_n) exactly corresponds to p-good adversary states.

Oversimplification: suppose we can efficiently implement $(\Pi_p, \mathbb{I} - \Pi_p)$ where image(Π_n) exactly corresponds to p-good adversary states.

Oversimplification: suppose we can efficiently implement $(\Pi_p, \mathbb{I} - \Pi_p)$ where image(Π_n) exactly corresponds to p-good adversary states.

Oversimplification: suppose we can efficiently implement $(\Pi_p, \mathbb{I} - \Pi_p)$ where image(Π_n) exactly corresponds to p-good adversary states.

Oversimplification: suppose we can efficiently implement $(\Pi_p, \mathbb{I} - \Pi_p)$ where image(Π_n) exactly corresponds to p-good adversary states.

Oversimplification: suppose we can efficiently implement $(\Pi_p, \mathbb{I} - \Pi_p)$ where image(Π_n) exactly corresponds to p-good adversary states.

Oversimplification: suppose we can efficiently implement $(\Pi_p, \mathbb{I} - \Pi_p)$ where image(Π_n) exactly corresponds to p-good adversary states.

Oversimplification: suppose we can efficiently implement $(\Pi_p, \mathbb{I} - \Pi_p)$ where image(Π_n) exactly corresponds to p-good adversary states.

Missing Pieces

- 1) Why does this terminate?
- 2) How do we define/implement Π_p ?

In the last talk, we used special properties of the two projectors to bound the runtime, but it's not clear what we can say about Π_r , Π_p .

In the last talk, we used special properties of the two projectors to bound the runtime, but it's not clear what we can say about Π_r , Π_n .

Insight: analyze runtime starting from $|S\rangle$, not $|S'\rangle$.

In the last talk, we used special properties of the two projectors to bound the runtime, but it's not clear what we can say about Π_r , Π_n .

Insight: analyze runtime starting from $|S\rangle$, not $|S'\rangle$.

In the last talk, we used special properties of the two projectors to bound the runtime, but it's not clear what we can say about Π_r , Π_n .

Insight: analyze runtime starting from $|S\rangle$, not $|S'\rangle$. Why does this help?

Consider the 2-D case.

Consider the 2-D case.

Consider the 2-D case.

Simple calculation: time to return to $|S\rangle$ is *independent* of θ .

Consider the 2-D case.

Simple calculation: time to return to $|S\rangle$ is *independent* of θ .

This extends to the general case by Jordan's lemma.

Repairing the Prover After Measurement

Missing Pieces

- Why does this terminate?
- 2) How do we define/implement Π_p ?

Repairing the Prover After Measurement

Missing Pieces

- Why does this terminate?
- 2) How do we define/implement Π_p ?

How do we define/implement Π_p ?

Rephrased: how do we measure the prover's success probability?

How do we define/implement Π_p ?

Rephrased: how do we measure the prover's success probability?

Bad news: we can't do this efficiently.

How do we define/implement $\Pi_{\mathcal{D}}$?

Rephrased: how do we measure the prover's success probability?

Bad news: we can't do this efficiently.

Good news: we can *approximately measure* the success probability…

How do we define/implement $\Pi_{\mathcal{D}}$?

Rephrased: how do we measure the prover's success probability?

Bad news: we can't do this efficiently.

Good news: we can *approximately measure* the success probability…

How?

How do we define/implement Π_n ?

Rephrased: how do we measure the prover's success probability?

Bad news: we can't do this efficiently.

Good news: we can *approximately measure* the success probability…

How?

Alternating projectors *again*!

•
$$
|+\rangle_R := \frac{1}{\sqrt{R}} \sum_{r \in R} |r\rangle
$$
 (uniform superposition of challenges)

- $\left| + \right\rangle _R \coloneqq$ % $\frac{1}{R} \sum_{r \in R} |r\rangle$ (uniform superposition of challenges)
- $\Pi_{\text{Acc}} \coloneqq \sum_r |r\rangle\langle r|_R \otimes \Pi_r$.

$$
|+\rangle_R
$$

$$
|\text{S}\rangle
$$

$$
-\frac{\text{Measure}}{\text{I}_{\text{Acc}}}
$$

- $\left| + \right\rangle _R \coloneqq$ % $\frac{1}{R} \sum_{r \in R} |r\rangle$ (uniform superposition of challenges)
- $\Pi_{\text{Acc}} \coloneqq \sum_r |r\rangle\langle r|_R \otimes \Pi_r$.

$$
|+\rangle_R
$$

 $|S\rangle$
 $\frac{\Pi_{\text{Acc}}}{b}$
 $\text{Pr}[b = 1]$ is the success probability of $|S\rangle$

- $\left| + \right\rangle _R \coloneqq$ % $\frac{1}{R} \sum_{r \in R} |r\rangle$ (uniform superposition of challenges)
- $\Pi_{\text{Acc}} \coloneqq \sum_r |r\rangle\langle r|_R \otimes \Pi_r$.

[MW05, Z20]: learn success probability by alternating Π_{Acc} measurements with $\Pi_{\text{Unif}} = |+\rangle\langle +|_{R} \otimes \mathbb{I}$ measurements

$$
|+\rangle_R
$$

 $|S\rangle$
 $\frac{\Pi_{Acc}}{b}$
 $\frac{\Pi_{Acc}}{b}$
 $|B\rangle$
 $|S\rangle$
 $|B\rangle$
 $|B\rangle$
 $|B\rangle$
 $|B\rangle$
 $|B\rangle$
 $|B\rangle$
 $|B\rangle$

- $\left| + \right\rangle _R \coloneqq$ % $\frac{1}{R} \sum_{r \in R} |r\rangle$ (uniform superposition of challenges)
- $\Pi_{\text{Acc}} \coloneqq \sum_r |r\rangle\langle r|_R \otimes \Pi_r.$
- $\Pi_{\text{Init}} := |+\rangle \langle +|_R \otimes \mathbb{I}.$

1) Initialize $|+_R\rangle$ (S).

$$
\left| + \right\rangle_R
$$

$$
\left| S \right\rangle
$$

- $\left| + \right\rangle _R \coloneqq$ % $\frac{1}{R} \sum_{r \in R} |r\rangle$ (uniform superposition of challenges)
- $\Pi_{\text{Acc}} \coloneqq \sum_r |r\rangle\langle r|_R \otimes \Pi_r.$
- $\Pi_{\text{Init}} := |+\rangle \langle +|_R \otimes \mathbb{I}.$

1) Initialize $|+_R\rangle$ (S). 2) Alternate M_{Acc} , M_{Unif} measurements, obtaining $(b_1, b_2, ..., b_T)$

$$
|+\rangle_R
$$

$$
|S\rangle
$$

$$
\frac{\text{Measure}}{\text{I}_{\text{Acc}}}
$$

$$
\frac{\text{Measure}}{\text{I}_{\text{Unif}}}
$$

$$
\frac{\text{Measure}}{\text{I}_{\text{L}_{\text{loc}}}}
$$

$$
\frac{\text{Measure}}{\text{I}_{\text{Unif}}}
$$

$$
\frac{\text{Measure}}{\text{I}_{\text{Unif}}}
$$

$$
\frac{\text{Measure}}{\text{I}_{\text{Unif}}}
$$

$$
\frac{\text{Value}}{\text{I}_{\text{L}_{\text{lin}}}}
$$

- $\left| + \right\rangle _R \coloneqq$ % $\frac{1}{R} \sum_{r \in R} |r\rangle$ (uniform superposition of challenges)
- $\Pi_{\text{Acc}} \coloneqq \sum_r |r\rangle\langle r|_R \otimes \Pi_r.$
- $\Pi_{\text{Init}} := |+\rangle \langle +|_R \otimes \mathbb{I}.$

1) Initialize $|+_R\rangle$ (S). 2) Alternate M_{Acc} , M_{Unif} measurements, obtaining $(b_1, b_2, ..., b_T)$ 3) Output $p = #$ of times $b_i = b_{i+1}/(T-1)$

- $\left| + \right\rangle _R \coloneqq$ % $\frac{1}{R} \sum_{r \in R} |r\rangle$ (uniform superposition of challenges)
- $\Pi_{\text{Acc}} \coloneqq \sum_r |r\rangle\langle r|_R \otimes \Pi_r.$
- $\Pi_{\text{Init}} := |+\rangle \langle +|_R \otimes \mathbb{I}.$

Why does this work?

Eigenvalue $p_i = \cos^2(\theta_i) = ||\Pi_{\text{Acc}}| + \rangle_R |S\rangle||^2$

 $(p_i$ is an eigenvalue of $\Pi_{\text{Unif}}\Pi_{\text{Acc}}\Pi_{\text{Unif}})$

•
$$
p_j
$$
 = success prob of $|S\rangle$.

Eigenvalue $p_i = \cos^2(\theta_i) = ||\Pi_{\text{Acc}}| + \rangle_R |S\rangle||^2$ $(p_i$ is an eigenvalue of $\Pi_{\text{Unif}}\Pi_{\text{Acc}}\Pi_{\text{Unif}})$

Eigenvalue $p_i = \cos^2(\theta_i) = ||\Pi_{\text{Acc}}| + \rangle_R |S\rangle||^2$ $(p_i$ is an eigenvalue of $\Pi_{\text{Unif}}\Pi_{\text{Acc}}\Pi_{\text{Unif}})$

Suppose $|+\rangle_R$ $|S\rangle$ lies in a 2dim Jordan subspace S_i .

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

Unif

dim Jordan subspace S_i .

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ Acc Unif

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

 $b_0 = 1$ $b_1 = 1$ Acc Unif

dim Jordan subspace S_i .

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ Acc Unif Unif

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ Acc Unif Unif

dim Jordan subspace S_i .

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ $b_3 = 0$ Acc Unif Acc Unif

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ $b_3 = 0$ Acc Unif Acc Unif

dim Jordan subspace S_i .

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ $b_3 = 0$ $b_4 = 0$ Acc Unif Acc Unif Unif

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ $b_3 = 0$ $b_4 = 0$ Acc Unif Acc Unif Unif

dim Jordan subspace S_i .

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ $b_3 = 0$ $b_4 = 0$ $b_5 = 0$ Acc Unif Acc Unif Unif Acc

Suppose $|+\rangle_R$ $|S\rangle$ lies in a 2dim Jordan subspace S_i .

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ $b_3 = 0$ $b_4 = 0$ $b_5 = 0$ Acc Unif Acc Unif Unif Acc

dim Jordan subspace S_i .

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ $b_3 = 0$ $b_4 = 0$ $b_5 = 0$ Acc Unif Acc Unif Unif Acc $b_6 = 1$ Unif

Suppose $|+\rangle_R$ $|S\rangle$ lies in a 2dim Jordan subspace S_i .

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ $b_3 = 0$ $b_4 = 0$ $b_5 = 0$ Acc Unif Acc Unif Unif Acc $b_6 = 1$ Unif

dim Jordan subspace S_i .

- p_i = success prob of $|S\rangle$.
- alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ $b_3 = 0$ $b_4 = 0$ $b_5 = 0$ $b_6 = 1$ $b_7 = 1$ Acc Unif Acc Unif Unif Acc Unif Acc

dim Jordan subspace S_i .

• p_i = success prob of $|S\rangle$.

alternating Π_{Acc} , Π_{Unif} measurements gives p_i

outcomes

 $b_0 = 1$ $b_1 = 1$ $b_2 = 1$ $b_3 = 0$ $b_4 = 0$ $b_5 = 0$ $b_6 = 1$ $b_7 = 1$ $b_8 = 1$ Acc Unif Acc Unif Unif Acc Unif Acc Unif

In general, $|+\rangle_R \otimes |S\rangle$ can have components in more than one Jordan subspace S_i .

Suppose
$$
|+\rangle_R \otimes |S\rangle = \alpha_1 |u_1\rangle + \alpha_2 |u_2\rangle + \alpha_3 |u_3\rangle
$$
.
success prob p_1 success prob p_2 sprob p_3 prob p_3

Key fact: Alternating measurement outcomes distributed as though +)_R \otimes (S) were contained in S_j with prob α_j & .

Key fact: Alternating measurement outcomes distributed as though +)_R \otimes (S) were contained in S_j with prob α_j & .

Leftover state concentrated on S_i 's most consistent w/ outcomes.

[MW05] Estimation "approximately" projects onto $\{S_i\}$

• w/prob $\approx |\alpha_j|^2$ obtain estimate $\approx p_j$ and leftover state $\approx |u_j\rangle$

Key Properties 1) $\mathbb{E}[p] = p_0$ 2) If we apply MW twice, the two outcomes p , q are close with high probability.

 p Rey Properties 1) $\mathbb{E}[p] = p_0$ 2) If we apply MW twice, the two outcomes p , q are close with high probability. Formally, MW achieves $Pr[|p - q| \leq \varepsilon] \geq 1 - \delta$ with $poly($ 1 $\mathcal{E}_{\mathcal{E}}$, $\log\left(\frac{1}{5}\right)$ $\frac{1}{\delta}$) runtime.

As in [Zha20], we call this " (ε, δ) -almost-projective."

Key Properties 1) $\mathbb{E}[p] = p_0$ 2) If we apply MW twice, the two outcomes p , q are close with high probability. Formally, MW achieves $Pr[|p - q| \leq \varepsilon] \geq 1 - \delta$

For this talk, we'll need to know two

things about the MW estimator.

with $poly($

1

, $\log\left(\frac{1}{5}\right)$

 $\frac{1}{\delta}$) runtime.

 $\mathcal{E}_{\mathcal{E}}$

Let's see how [MW05] fits into our approach.

We don't know how to measure Π_n , but we can approximate it:

MW_n: run the MW estimator and accept if the output is $\geq p$.

We don't know how to measure Π_n , but we can approximate it:

MW_n: run the MW estimator and accept if the output is $\geq p$.

Idea: run Marriott-Watrous on Marriott-Watrous!

We don't know how to measure Π_n , but we can approximate it:

MW_n: run the MW estimator and accept if the output is $\geq p$.

Idea: run Marriott-Watrous on Marriott-Watrous!

Subtle point: Just restoring "success probability" is not enough!

Subtle point: Just restoring "success probability" is not enough! Definition: (S) is *strongly p-successful* if it is concentrated on $(\Pi_{\text{Acc}}, \Pi_{\text{Unif}})$ -Jordan subspaces with eigenvalue $\geq p$

Subtle point: Just restoring "success probability" is not enough! Definition: (S) is *strongly p-successful* if it is concentrated on $(\Pi_{\text{Acc}}, \Pi_{\text{Unif}})$ -Jordan subspaces with eigenvalue $\geq p$ We want: If $|S\rangle$ is strongly p-successful, then $|S_1\rangle$ is strongly psuccessful

This seems promising, but we have a problem: Our proof that this procedure terminates requires the measurements to be projective, but MW_p is not!

This seems promising, but we have a problem:

Our proof that this procedure terminates requires the measurements to be projective, but MW_p is not!

(running it twice may give different outcomes)

This seems promising, but we have a problem:

Our proof that this procedure terminates requires the measurements to be projective, but MW_p is not!

Easy(?) fix: Make MW_p projective by expanding the Hilbert space.

Measuring $|S'\rangle$ with MW_p can be implemented as a projective measurement of some Π_p^* on $\ket{S'}_A\ket{0}_W \in A \otimes W$.

adversary state register workspace/ancilla

Measuring $|S'\rangle$ with MW_p can be implemented as a projective measurement of some Π_p^* on $\ket{S'}_A\ket{0}_W \in A \otimes W$.

adversary state register workspace/ancilla

Measuring $|S'\rangle$ with MW_p can be implemented as a projective measurement of some Π_p^* on $|S'\rangle_A|0\rangle_W \in A \otimes W$.

But we need to be careful: Simply being in image(Π_p^*) doesn't tell us anything! If the ancilla is not $|0\rangle$, then measuring Π^*_p does not correspond to MW_{p} .

Our solution is re-define Π_r to $\Pi_r^* := \Pi_r \otimes |0\rangle\langle 0|_W$, so that each measurement of Π^*_r attempts to "reset" the W to $|0\rangle_W$.

This is essentially the full repair procedure!

Our solution is re-define Π_r to $\Pi_r^* := \Pi_r \otimes |0\rangle\langle 0|_W$, so that each measurement of Π^*_r attempts to "reset" the W to $|0\rangle_W$.

This is essentially the full repair procedure!

Our solution is re-define Π_r to $\Pi_r^* := \Pi_r \otimes |0\rangle\langle 0|_W$, so that each measurement of Π^*_r attempts to "reset" the W to $|0\rangle_W$.

Not obvious: why does this choice of Π^*_r make repair work?

In any 2-D Jordan subspace: if we start at S')|0) we end up at $|\psi\rangle$ after Π_p^* accepts.

In any 2-D Jordan subspace: if we start at S')|0) we end up at $|\psi\rangle$ after Π_p^* accepts.

Claim: ψ_A corresponds to a strongly ($p \varepsilon$)-successful adversary.

In any 2-D Jordan subspace: if we start at S')|0) we end up at $|\psi\rangle$ after Π_p^* accepts.

Claim: ψ_A corresponds to a strongly ($p \varepsilon$)-successful adversary.

Proof Sketch

1) If we run MW twice, two estimates are ε -close with prob $1 - \delta$.

In any 2-D Jordan subspace: if we start at S')|0) we end up at $|\psi\rangle$ after Π_p^* accepts.

Claim: ψ_A corresponds to a strongly ($p \varepsilon$)-successful adversary.

Proof Sketch

1) If we run MW twice, two estimates are ε -close with prob $1 - \delta$. 2) $|\psi\rangle = \Pi_p^*|S'\rangle|0\rangle$ corresponds to running MW($|S'\rangle$) $\rightarrow q$ and *conditioning* on $q \geq p$.

In any 2-D Jordan subspace: if we start at S')|0) we end up at $|\psi\rangle$ after Π_p^* accepts.

Claim: ψ_A corresponds to a strongly ($p \varepsilon$)-successful adversary.

Proof Sketch

1) If we run MW twice, two estimates are ε -close with prob $1 - \delta$. 2) $|\psi\rangle = \Pi_p^*|S'\rangle|0\rangle$ corresponds to running MW($|S'\rangle$) $\rightarrow q$ and *conditioning* on $q \geq p$. 3) Markov: if we run MW on ψ_A , get $\geq p - \varepsilon$ with prob $1 - \delta / \gamma$.

In any 2-D Jordan subspace: if we start at S')|0) we end up at $|\psi\rangle$ after Π_p^* accepts.

Claim: ψ_A corresponds to a strongly ($p \varepsilon$)-successful adversary.

Proof Sketch

For the general case, need to show that most of the state is on subspaces where γ_i is not too small.

• We showed how to rewind and obtain *arbitrarily many* accepting protocol transcripts.

- We showed how to rewind and obtain *arbitrarily many* accepting protocol transcripts.
- This gives post-quantum soundness of Kilian's protocol as well as *optimal* soundness error for many other protocols (e.g., Blum).

- We showed how to rewind and obtain *arbitrarily many* accepting protocol transcripts.
- This gives post-quantum soundness of Kilian's protocol as well as *optimal* soundness error for many other protocols (e.g., Blum).
- This technique has found many other applications:
	- [Bitansky-Brakerski-Kalai22]: "advice preserving" non-interactive quantum reductions
	- [Lai-Malavolta-Spooner22]: quantum rewinding for many-round protocols
	- [Gunn-Ju-Ma-Zhandry22]: quantum-communication succinct arguments

- We showed how to rewind and obtain *arbitrarily many* accepting protocol transcripts.
- This gives post-quantum soundness of Kilian's protocol as well as *optimal* soundness error for many other protocols (e.g., Blum).
- This technique has found many other applications:
	- [Bitansky-Brakerski-Kalai22]: "advice preserving" non-interactive quantum reductions
	- [Lai-Malavolta-Spooner22]: quantum rewinding for many-round protocols
	- [Gunn-Ju-Ma-Zhandry22]: quantum-communication succinct arguments

So have we resolved quantum rewinding?

The [CMSZ21] technique is *still* not as powerful as classical rewinding:

The [CMSZ21] technique is *still* not as powerful as classical rewinding:

• Needs advice: an explicit lower bound on the prover's initial success probability is needed to guarantee extraction. In general, this lower bound may not be physically accessible.

The [CMSZ21] technique is *still* not as powerful as classical rewinding:

- Needs advice: an explicit lower bound on the prover's initial success probability is needed to guarantee extraction. In general, this lower bound may not be physically accessible.
- Much slower than classical rewinding: if prover is ε -successful, it takes $1/\varepsilon^5$ steps to extract!

The [CMSZ21] technique is *still* not as powerful as classical rewinding:

- Needs advice: an explicit lower bound on the prover's initial success probability is needed to guarantee extraction. In general, this lower bound may not be physically accessible.
- Much slower than classical rewinding: if prover is ε -successful, it takes $1/\varepsilon^5$ steps to extract!
- Doesn't preserve the prover's state: prover state after extraction may be completely different than the adversary's real (post-execution) state. This is *by design*, since repair only restores success probability.

This concludes: the unreasonable effectiveness of alternating projectors in quantum rewinding.

Thank You!

Questions?