

Title: Constant-Overhead Magic State Distillation

Speakers: Hayata Yamasaki

Collection/Series: Quantum Information

Subject: Quantum Information

Date: November 20, 2024 - 11:00 AM

URL: <https://pirsa.org/24110068>

Abstract:

Magic state distillation is a crucial yet resource-intensive process in fault-tolerant quantum computation. The protocol's overhead, defined as the number of input magic states required per output magic state with an error rate below ϵ , typically grows as $O(\log^\gamma(1/\epsilon))$ as $\epsilon \rightarrow 0$. Achieving smaller overheads, i.e., smaller exponents γ , is highly desirable; however, all existing protocols require polylogarithmically growing overheads with some $\gamma > 0$, and identifying the smallest achievable exponent γ for distilling magic states of qubits has remained challenging. To address this issue, we develop magic state distillation protocols for qubits with efficient, polynomial-time decoding that achieve an $O(1)$ overhead, meaning the optimal exponent $\gamma = 0$; this improves over the previous best of $\gamma \approx 0.678$ due to Hastings and Haah. In our construction, we employ algebraic geometry codes to explicitly present asymptotically good quantum codes for 2^{10} -dimensional qudits that support transversally implementable logical gates in the third level of the Clifford hierarchy. These codes can be realized by representing each 2^{10} -dimensional qudit as a set of 10 qubits, using stabilizer operations on qubits. We prove that the use of asymptotically good codes with non-vanishing rate and relative distance in magic state distillation leads to the constant overhead. The 10-qubit magic states distilled with these codes can be converted to and from conventional magic states for the controlled-controlled-Z (CCZ) and T gates on qubits with only a constant overhead loss, making it possible to achieve constant-overhead distillation of such standard magic states for qubits. These results resolve the fundamental open problem in quantum information theory concerning the construction of magic state distillation protocols with the optimal exponent.

The talk is based on the following paper.

<https://arxiv.org/abs/2408.07764>

Constant-Overhead Magic State Distillation

Hayata Yamasaki

The University of Tokyo, NanoQT Inc.

hayata.yamasaki@gmail.com [@hayatayamasaki](https://twitter.com/hayatayamasaki)

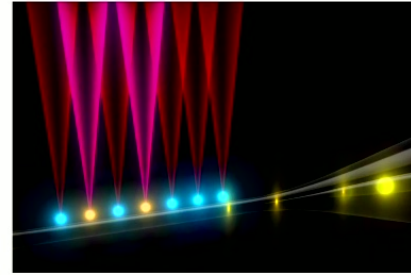
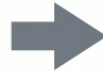
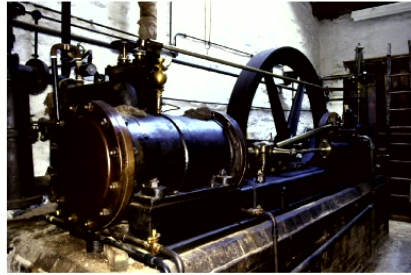
20th November, 2024

Reference:

Adam Wills, Min-Hsiu Hsieh, Hayata Yamasaki, [arXiv:2408.07764](https://arxiv.org/abs/2408.07764)

Future of Quantum Technologies

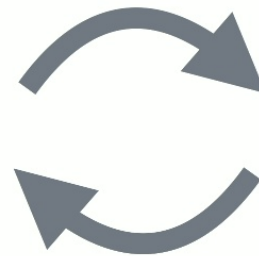
19th century
Steam engine
Thermodynamics



<https://www.nano-qt.com/>

21st century
Quantum devices
Quantum information

Quantum technologies
New technological advances
→ quantum computers



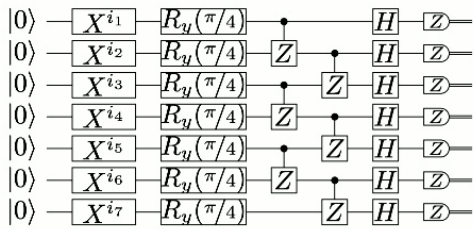
Quantum information theory
Theory of physics to understand what
we can do with quantum mechanics

Goal: What can we achieve in the quantum era, and how?

Adam Wills, Min-Hsiu Hsieh, Hayata Yamasaki, [arXiv:2408.07764](https://arxiv.org/abs/2408.07764)

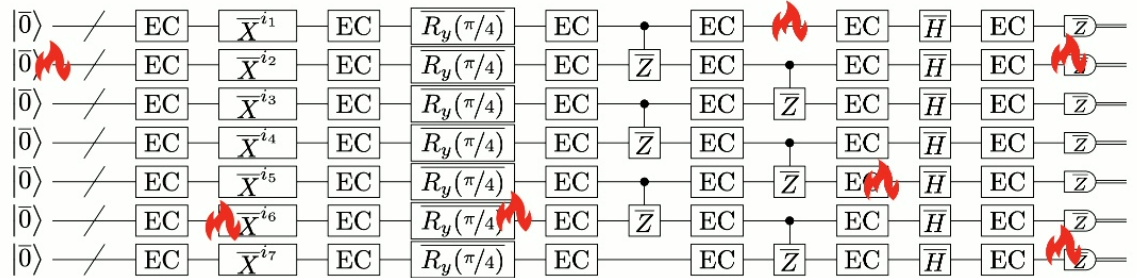
Fault-Tolerant Quantum Computation

Original circuit



Overall error within δ

Noisy circuit with physical error rate $p > 0$



Challenge:

To overcome the effect of noise

Without quantum error correction, error per gate has to be too small

$$p \leq \frac{\delta}{\#gates} \rightarrow 0 \quad 1/500=0.2\% \text{ required already for 500 qubits}$$

NISQ algorithms are not expected to attain any large quantum advantage because of this

Solution:

Fault-tolerant quantum computation (FTQC)

Use quantum error correction to suppress logical error rate ϵ arbitrarily

$$p \leq p_{th} \approx 0.1 \sim 1\% \quad \text{Below threshold}$$

$$\Rightarrow \epsilon \lesssim \frac{\delta}{\#gates} \quad \text{Suppress logical error rate as required}$$

Clifford Gates and Non-Clifford Gates

Clifford gate H, S, CNOT, \dots

- Transversal in many protocols
- Easy to implement



Non-Clifford gate T, CCZ, \dots

- We need error-suppressed **magic states**
- Implemented by gate teleportation

Typical protocols for FTQC

Color code

Transversal

$$\bar{U} = \bigotimes_{j=1}^n U_j$$

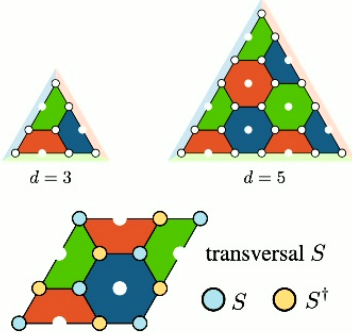


Figure from arXiv:1704.01589

Surface code

Fold-transversal

Moussa, arXiv:1603.02286

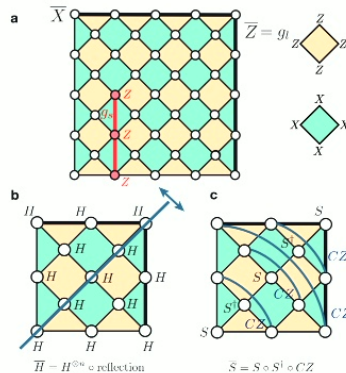


Figure from arXiv:2406.17653

Clifford hierarchy

$$\mathcal{C}^{(1)} = \mathcal{P} : \text{Pauli}$$

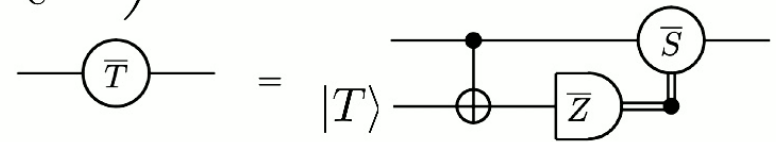
$$\mathcal{C}^{(k)} = \{U : UPU^\dagger \in \mathcal{C}^{(k-1)}, P \in \mathcal{P}\}$$

$$\mathcal{C}^{(2)} : \text{Clifford} \quad \mathcal{C}^{(3)}, \dots : \text{Non-Clifford}$$

Magic state for T gate: $|T\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\pi/4}|1\rangle)$

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \in \mathcal{C}^{(3)}$$

Gate teleportation



Magic State Distillation

Magic state distillation: Resource-intensive process to prepare magic states for FTQC



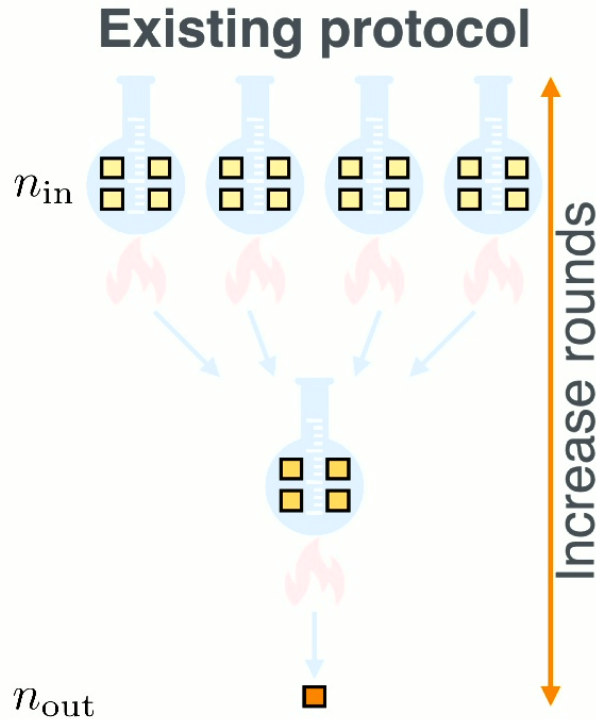
- State injection: Physical magic state \rightarrow Logical magic state \approx **physical error rate**
- **Stabilizer operations:** Preparation of $|0\rangle$, Clifford gates, Z measurement \approx **noiseless**

Goal: Convert n_{in} noisy magic states into n_{out} below error rate ϵ by stabilizer operations

\rightarrow Overhead: n_{in}/n_{out}

Question: What is the required overhead, and how to reduce it as much as possible?

Problem: Overheads



Error suppression by $[[15,1,3]]$ triorthogonal code

Bravyi, Kitaev arXiv:quant-ph/0403025

$$|+\rangle \xrightarrow[\text{Transversal } T]{\text{noisy } T^{\otimes 15}} \mathcal{N}^{\otimes 15}(|T\rangle) \xrightarrow{\text{post-selection}} |\overline{T}\rangle \xrightarrow{\text{Clifford}} |T\rangle$$

Error rate p $O(p^3)$

→ Repeat this L times so that the error rate should be below ϵ

Formula on overhead for $[[n, k, d]]$ triorthogonal codes

Error-suppression bound: $\epsilon = O(p^{d^L})$ n #physical qubits

Overhead: $\frac{n_{\text{in}}}{n_{\text{out}}} = \left(\frac{n}{k}\right)^L = O\left(\log^\gamma\left(\frac{1}{\epsilon}\right)\right)$ k #logical qubits

Exponent: $\gamma = \frac{\log \frac{n}{k}}{\log d} > 0$ d distance

Bravyi, Haah, arXiv:1209.2426

Open question: What is the optimal exponent γ ?

Results: Constant-Overhead Magic State Distillation

Year	Authors	Overhead Exponent	Note
2004	Bravyi, Kitaev	$\gamma \approx 2.46$	Discovery of Magic State Distillation
2012	Meier, Eastin, Knill	$\gamma \approx 2.32$	
2012	Bravyi, Haah	$\gamma \approx 1.58$	Conjectured that $\gamma < 1$ would be impossible
2012	Jones	$\gamma \rightarrow 1$	
2017	Hastings, Haah	$\gamma \approx 0.678$	Falsified the above conjecture
2024	Wills, Hsieh, Yamasaki	$\gamma = 0$ (optimal)	This work

Krishna, Tillich, arXiv:1811.08461 achieved arbitrarily small $\gamma > 0$ for prime-dimensional qudits, while this is not applicable to magic state distillation for qubits

Improvement of $[[n,k,d]]$ codes

→ Polylog overhead with small γ

$$O\left(\log^\gamma\left(\frac{1}{\epsilon}\right)\right) \quad \gamma = \frac{\log \frac{n}{k}}{\log d} > 0$$



Our work: Asymptotically good code + New protocol

→ We prove our protocol achieves **constant overhead**

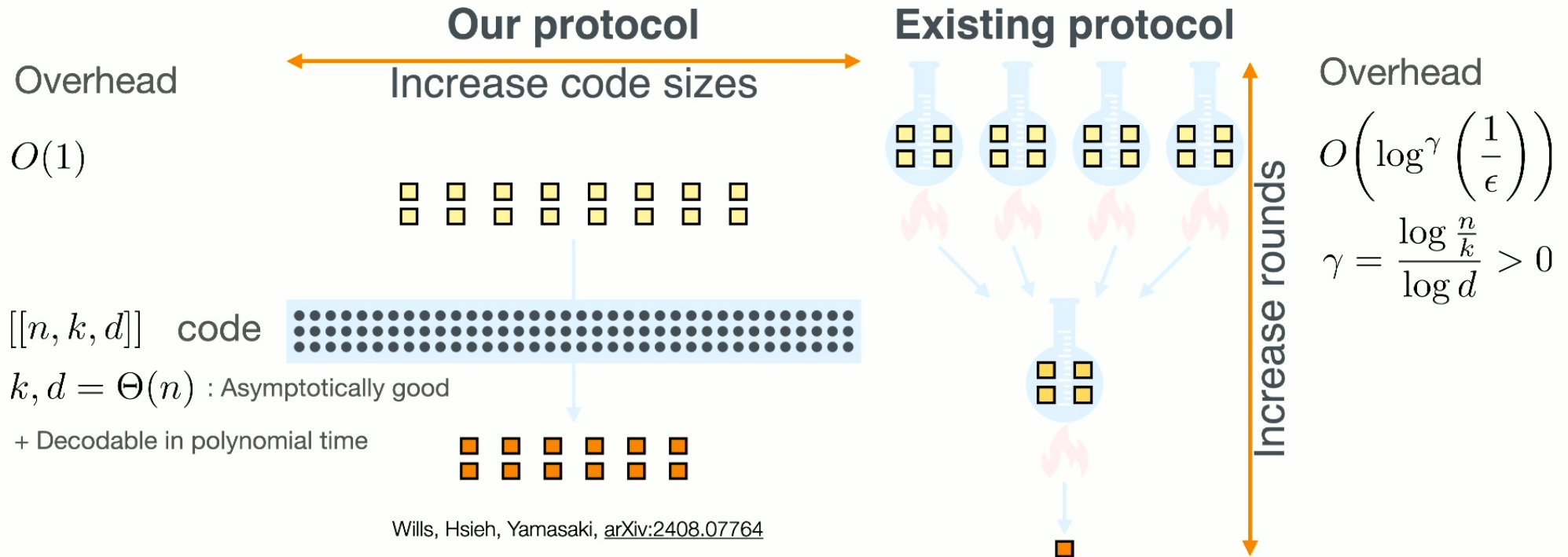
$$O(1) \text{ i.e., optimal overhead exponent } \gamma = 0$$

After posting our work, other works on code construction appeared, but without our protocol, existing protocols with such codes only achieve polylog overheads with small $\gamma > 0$
 Golowich, Guruswami, arXiv:2408.09254; Nguyen, arXiv:2408.10140; Scruby, Pesah, Webster, arXiv:2408.13130; Lin, arXiv:2410.14631; Golowich, Lin, arXiv:2410.14662

Adam Wills, Min-Hsiu Hsieh, Hayata Yamasaki, arXiv:2408.07764

7

Idea: Asymptotically Good Triorthogonal Code + New Protocol



Our results

- **Asymptotically good** triorthogonal code
- + **New single-round protocol**



Achieving constant overhead
 $O(1) \Rightarrow \gamma = 0$: optimal

Triorthogonal Codes

Triorthogonal matrix: $m \times n$ matrix $G \in \mathbb{F}_2^{m \times n}$ with rows $(g^a)_{a=1}^m$ satisfying **algebraic relations**

$$\sum_{i=1}^n g_i^a g_i^b = 0 \quad (a < b) \quad : \text{For commutativity of X and Z stabilizer generators}$$

$$\sum_{i=1}^n g_i^a g_i^b g_i^c = 0 \quad (a < b < c) : \text{For transversal T gates}$$

Arithmetics over finite field $\mathbb{F}_2 = \{0, 1\}$

$$G = \begin{pmatrix} G_1 \\ G_0 \end{pmatrix} \begin{matrix} \text{n columns} \\ \text{k rows: odd weight} \\ \text{(m-k) rows: even weight} \end{matrix}$$

Triorthogonal codes: $\text{CSS}(X, \mathcal{G}_0; Z, \mathcal{G}^\perp)$ with stabilizer generators $X^x (x \in \mathcal{G}_0); Z^z (z \in \mathcal{G}^\perp)$

$\mathcal{G}_0, \mathcal{G}_1, \mathcal{G} \subset \mathbb{F}_2^n$: Linear subspaces spanned by rows of G_0, G_1, G

$\mathcal{G}^\perp := \{u \in \mathbb{F}_2^n : u \cdot v = 0, v \in \mathcal{G}\}$: Dual code

→ $[[n, k, d]]$ codes with distance $d = d_Z \geq \min_{v \in \mathcal{G}_0^\perp \setminus \mathcal{G}^\perp} |v|$ **determined by dual codes**

General framework for constructing quantum codes with transversal T

Bravyi, Haah, arXiv:1209.2426 for qubits; Krishna, Tillich, arXiv:1811.08461 for prime-dimensional qudits

Technique: Algebraic Geometry Codes in 2^s Dimensions

Previous: Codes with **polynomials**

- Reed-Muller code
Hastings, Haah, arXiv:1709.03543 for qubits
- Reed-Solomon code
Krishna, Tillich, arXiv:1811.08461 for prime-dimensional qudits



This work: Algebraic geometry code

- Using **rational functions** instead of polynomials
- Leading to **non-vanishing rate & linear distance**
Goppa, Geometry and Codes (1988)

Algebraic geometry code over q -dimensional finite fields

We have an infinite family of fields of rational functions **parameterized by n_i** (number of rational places) and **g_i** (genus) with

$$\limsup_{i \rightarrow \infty} \frac{n_i}{g_i} = \sqrt{q} - 1 : \text{Better constant factors for large } q$$

$$g_i = \Theta(n_i)$$

Given a field of rational functions parameterized by n_i and g_i , for **any parameter a_i** satisfying $2g_i - 1 \leq a_i < n_i$

$$a_i = \Theta(n_i)$$

- We have a linear code $\mathcal{C} \subset \mathbb{F}_q^{n_i}$ with dimension $k_i = a_i + 1 - g_i = \Theta(n_i)$
- The dual code has distance $d_i \geq a_i - (2g_i - 2) = \Theta(n_i)$
- The dual code is efficiently decodable up to radius $t_i = (a_i - 3g_i + 1)/2 = \Theta(n_i)$

See also Stichtenoth, Algebraic function fields and codes (2009); Preliminaries of our paper

Adam Wills, Min-Hsiu Hsieh, Hayata Yamasaki, [arXiv:2408.07764](https://arxiv.org/abs/2408.07764)

10

Technique: Algebraic Geometry Codes in 2^s Dimensions

Challenge: Algebraic geometry codes require large dimension q , but we want qubit codes

→ **Solution:** Work on 2^s -dimensional qudits ($q=2^s$), **isomorphic to sets of s qubits**

$$\mathbb{F}_2 = \{0, 1\}$$

with arithmetics mod 2



$$\mathbb{F}_{2^{10}} = \left\{ \eta = \sum_{i=0}^9 a_i \alpha^i : a_i \in \mathbb{F}_2 \right\} \quad (s = 10)$$

Set of polynomials

with arithmetics mod irreducible polynomial $\alpha^{10} + \alpha^3 + 1 = 0$

Trace map: $\text{tr}(\gamma) := \sum_{i=0}^9 \gamma^{2^i} : \mathbb{F}_{2^{10}} \rightarrow \mathbb{F}_2$

Self-dual basis: $\{\alpha_i \in \mathbb{F}_{2^{10}}\}_{i=0,\dots,9}$ such that $\text{tr}(\alpha_i \alpha_j) = \delta_{i,j} \Rightarrow \eta = \sum_{i=0}^9 b_i \alpha_i \in \mathbb{F}_{2^{10}}$
Sum of polynomials in the basis

$$2^{10}\text{-dim qudit } \mathbb{C}^{2^{10}} = \text{span} \left\{ \left| \eta = \sum_{i=0}^9 b_i \alpha_i \right\rangle \right\}$$

$$X^\beta |\eta\rangle = |\eta + \beta\rangle \quad Z^\gamma |\eta\rangle = (-1)^{\text{tr}(\gamma\eta)} |\eta\rangle$$



$$\text{set of 10 qubits } (\mathbb{C}^2)^{\otimes 10} = \text{span} \left\{ \bigotimes_{i=0}^9 |b_i\rangle \right\}$$

$$X^{\mathbf{b}} = \bigotimes_{i=0}^9 X^{b_i} \quad Z^{\mathbf{b}} = \bigotimes_{i=0}^9 Z^{b_i}$$

Clifford hierarchy: In the same way $\mathcal{C}^{(1)} = \mathcal{P} : \text{Pauli}; \quad \mathcal{C}^{(k)} = \{U : UPU^\dagger \in \mathcal{C}^{(k-1)}, P \in \mathcal{P}\}$

For each fixed self-dual basis, we have **one-to-one correspondence**

Formulation of Triorthogonality for 2^s -dimensional Qudits

Triorthogonal Matrix: $m \times n$ matrix $G \in \mathbb{F}_{2^s}^{m \times n}$ with rows $(g^a)_{a=1}^m$ satisfying **algebraic relations**

$$\sum_{i=1}^n \sigma_i g_i^a g_i^b = \begin{cases} \tau_a & \text{if } 1 \leq a = b \leq k \\ 0 & \text{otherwise} \end{cases} \quad \text{For commutativity of X and Z stabilizer generators}$$

It is OK to allow coefficients $\neq 1$

$$\sum_{i=1}^n (g_i^a)^4 (g_i^b)^2 (g_i^c) = \begin{cases} 1 & \text{if } 1 \leq a = b = c \leq k \\ 0 & \text{otherwise} \end{cases} \quad \text{For transversal non-Clifford gates}$$

Arithmetics over finite field \mathbb{F}_{2^s}

$$G = \begin{pmatrix} G_1 \\ G_0 \end{pmatrix} \begin{matrix} \text{n columns} \\ \text{k rows} \\ \text{(m-k) rows} \end{matrix}$$

New family of single-qudit non-Clifford gates

$$U^{(n)} := \sum_{\gamma \in \mathbb{F}_{2^s}} \exp[i\pi \text{tr}(\gamma^n)] |\gamma\rangle \langle \gamma|$$

Instead of $i\pi/4$, use nonlinearity of γ^n for non-Cliffordness

$U^{(1)}, U^{(2)}, U^{(4)}$: Pauli $U^{(3)}, U^{(5)}, U^{(6)}$: Clifford

$U^{(7)}$: non-Clifford gate in the 3rd level of Clifford hierarchy for any $s \geq 5$

Triorthogonal codes for 2^s -dimensional qudits: In the same way $\text{CSS}(X, \mathcal{G}_0; Z, \mathcal{G}^\perp)$

- Transversal non-Clifford gate $(U^{(7)})^{\otimes n} = \overline{(U^{(7)})^{\otimes k}}$
- Distance $d = d_Z \geq \min_{v \in \mathcal{G}_0^\perp \setminus \mathcal{G}^\perp} |v|$ determined by dual codes

Construction of Asymptotically Good Triorthogonal Codes

Main theorem: Consider 2^s -dimensional qudits for any fixed $s \geq 10$ with $s \neq 0 \pmod{3}$.

Using an infinite family of fields of rational functions **parameterized by n'_i (number of rational places) and g_i (genus)** with

$$n'_i - 4 + g_i \geq 7(3g_i + 2)$$

we can construct triorthogonal matrices that give rise to $[[n_i, k_i, d_i]]$ triorthogonal codes with

- **Linear** number of logical qubits $k_i = \Theta(n_i)$
- **Linear** distance $d_i = \Theta(n_i)$ **Asymptotically good**
- **Linear** radius for efficient decoding of Z errors $t_i = \Theta(n_i)$: Relevant for magic state distillation

+ We **explicitly construct these codes** while not optimizing constant factors

Examples of code parameters: $s = 10, n_i \approx 29 \times 32^i, k_i \approx 1.3 \times 32^i, d_i \approx 1.3 \times 32^i, t_i \approx 0.13 \times 32^i$ ($i \in \{3, 4, \dots\}$)

- The inequality condition: For ensuring **triorthogonality**
- Algebraic geometry codes + Puncturing \rightarrow **Asymptotically good** code parameters
- **Linear t_i for efficient decoding is must for constant-overhead magic state distillation**

Implementation of 2^s -dimensional Qudits with Qubits

Problem: How to distill conventional magic states, e.g., CCZ and T, with constant overheads

→ **Key finding:** Qubit codes with **transversal CCZ and T gates are unnecessary**

Requirement for $O(1)$ overhead: **Exact conversion**

→ Resource theory of magic



Mere use of Solovay–Kitaev theorem

→ Polylog overhead thus insufficient

$$|M\rangle := U^{(7)} \left| +_{(2^{10})} \right\rangle \stackrel{\text{stabilizer}}{\leftrightarrow} |CCZ\rangle := CCZ |+\rangle^{\otimes 3} \quad U^{(n)} := \sum_{\gamma \in \mathbb{F}_2^s} \exp[i\pi \text{tr}(\gamma^n)] |\gamma\rangle \langle \gamma|$$

Magic state on a set of 10 qubits

- **CCZ → U⁽⁷⁾:** For each self-dual basis, any **diagonal** non-Clifford gate in 3rd level of Clifford hierarchy is **decomposed into a finite sequence of Z, CZ, CCZ** $|CCZ\rangle^{\otimes 70} \rightarrow |M\rangle$
Houshmand, Zamani, Sedighi, Arabzadeh, arXiv:1405.6741
- **U⁽⁷⁾ → CCZ:** Convert **hypergraph states** by Z measurements to delete edges $|M\rangle \rightarrow |CCZ\rangle$
- It is also possible to further **convert from/to T** $|T\rangle^{\otimes 4} \rightarrow |CCZ\rangle \quad |CCZ\rangle \otimes |T\rangle \rightarrow |T\rangle^{\otimes 3}$
Jones, arXiv:1212.5069; Selinger, arXiv:1210.0974; Gidney, Fowler, arXiv:1812.01238; Beverland, Campbell, Howard, Kliuchnikov arXiv:1904.01124

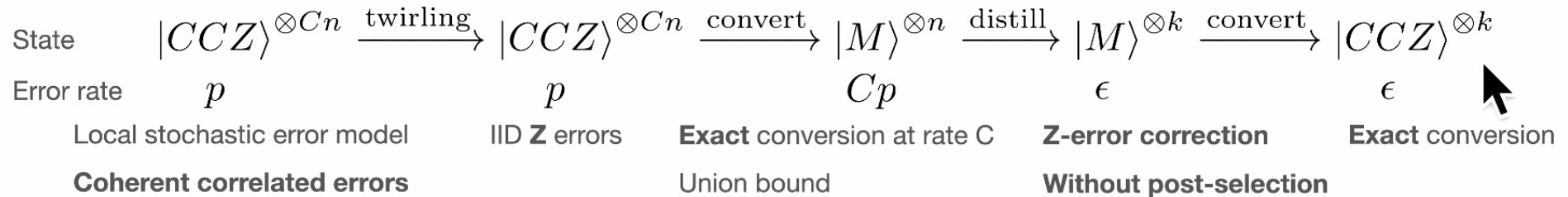
Apart from our approach, Golowich, Guruswami, arXiv:2408.09254; Nguyen, arXiv:2408.10140 showed another approach to use algebraic geometry codes for arguing existence of asymptotically good codes with transversal CCZ gates, but it is currently unknown which of these two approaches leads to better constant factors

Adam Wills, Min-Hsiu Hsieh, Hayata Yamasaki, arXiv:2408.07764

14

Single-Round Protocol by Efficient Quantum Error Correction

New single-shot protocol



Our proof of threshold theorem and constant overhead

- **We analyze local stochastic error model**, more general than IID errors in existing work
- **Linear decoding radius t**
 → **Nonzero threshold** $p_{\text{th}} > 0$ for error suppression in single round $\epsilon \leq \left(\frac{p}{p_{\text{th}}}\right)^{t+1}$
- **Non-vanishing rate + Efficient decoding** without post-selection
 → **Constant overhead** $\frac{Cn}{k} = \Theta(1)$

Development of new protocol and analysis is essential for achieving constant overhead

Connection to Generalized Quantum Stein's Lemma

Constant-overhead magic state distillation (this week): Bottom-up approach

- **Stabilizer operations \mathcal{O} :** Preparation of $|0\rangle$, Clifford gates, Z measurement
- **We prove nonzero asymptotic conversion rate** $r_{\mathcal{O}}(\rho \rightarrow |CCZ\rangle) > 0$



Implications of generalized quantum Stein's lemma (last week): Top-down approach

- **Asymptotically resource-non-generating operations $\tilde{\mathcal{O}}$**
- **Exact characterization of asymptotic conversion rate** $r_{\tilde{\mathcal{O}}}(\rho \rightarrow |CCZ\rangle) = \frac{R_{\mathbb{R}}^{\infty}(\rho)}{R_{\mathbb{R}}^{\infty}(|CCZ\rangle)}$

Masahito Hayashi, Hayata Yamasaki, [arXiv:2408.02722](https://arxiv.org/abs/2408.02722)

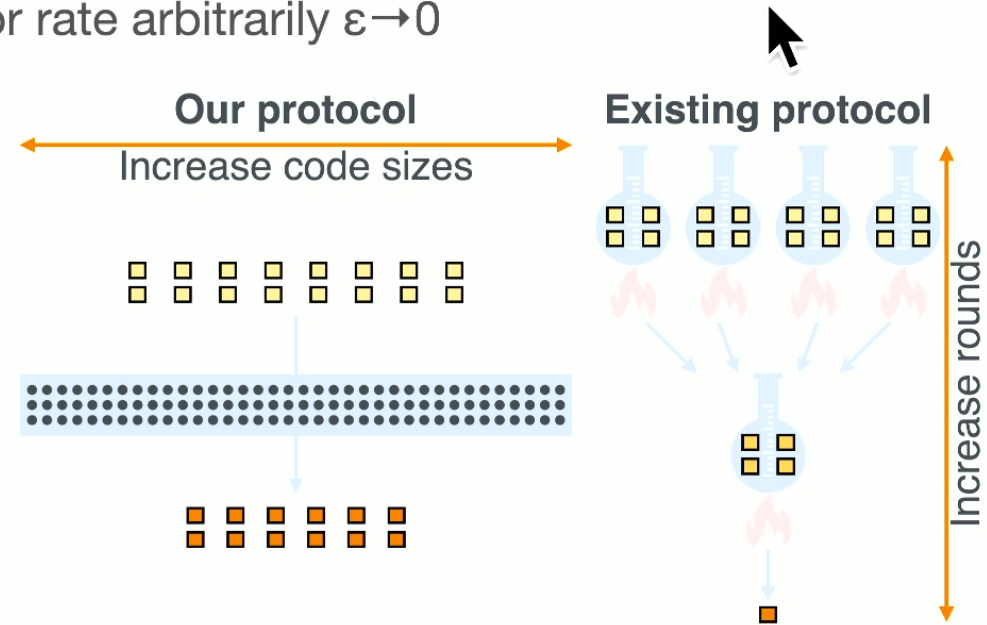
Open: What is the optimal rate of magic state distillation under stabilizer operations?

Implications for Future Protocol Design

New scale-up strategy for magic state distillation

For large-scale FTQC, we need to suppress error rate arbitrarily $\epsilon \rightarrow 0$

- A **single round** is all you need
 - **Do not post-select** but correct errors on large scales
 - **Start from small** triorthogonal codes
[[10,1,2]], [[15,1,3]],...
- Bravyi, Kitaev arXiv:quant-ph/0403025; Vasmer, Kubica, arXiv:2112.01446
- **Design a sequence of codes** with linearly growing distances to scale up



An exciting time has come to optimize magic state distillation **with a new design principle**

My Research

Social
implementation

Advance of IT society
by quantum technology

Useful quantum algorithm

Quantum machine learning
with high speed/applicability

**Theoretical
foundation
= my works**

Implementation of QC

Low-overhead/scalable
fault-tolerant QC (FTQC)

Efficient Q operations

Quantitative analysis of use
of quantum resources

Experimental
foundation

Advance of
quantum technology



<https://www.hayatayamasaki.com/>

18

Summary

- We develop **asymptotically good triorthogonal codes** and a **new protocol** to achieve constant-overhead magic state distillation
- **We prove the threshold theorem and constant overhead** of our protocol
- Techniques from algebraic geometry codes illuminate **what we can do optimally**, up to constant factors that are to be optimized further from this point forward

References:

Adam Wills, Min-Hsiu Hsieh, Hayata Yamasaki, [arXiv:2408.07764](https://arxiv.org/abs/2408.07764)

Reach me out for further discussion

Hayata Yamasaki hayata.yamasaki@gmail.com

Thank you for your attention.

<https://www.hayatayamasaki.com/>

19