

Earl Campbell

SHORTER GATE SEQUENCES FOR QUANTUM COMPUTING BY MIXING UNITARIES

PHYSICAL REVIEW A **95**, 042306 (2017)

ARXIV:1612.02689

1

Gate-synthesis / compiling

2

Measuring noise and coherence

3

The RESULT

4

The PROOF



**Gate-synthesis
/compiling**

Gate set \mathcal{G}

A collection of unitaries used to build circuits

e.g. from $\mathcal{G} = \{A, B, C\}$ we can build $ABC, ABBC, CBC,$ etc

Universality - Informal statement

\mathcal{G} is **universal** if it can implement any unitary (upto finite precision)

Universality - Formal statement

\mathcal{G} is **universal** if for any target unitary V and $\epsilon > 0$ there exists a finite circuit $U \in \langle \mathcal{G} \rangle$ such that $d(U, V) \leq \epsilon$

Example

Clifford+T or Clifford+Toffoli

Cost model $\mathfrak{C} : \mathcal{G} \rightarrow \mathbb{R}_+$

Each elementary gate given a positive valued “cost”

This induces a circuit cost

$$\text{if } U = \prod_i G_i \text{ then } \mathfrak{C}(U) = \sum_i \mathfrak{C}(G_i)$$

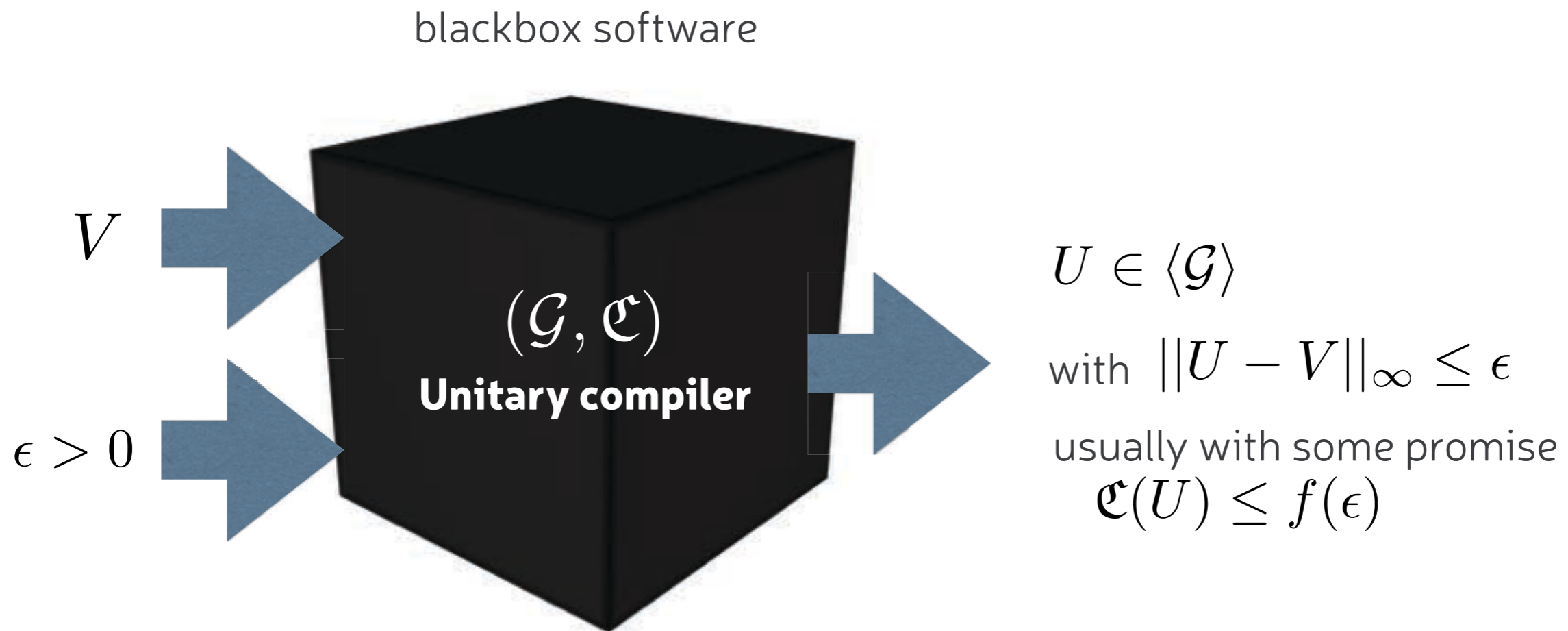
Example

Uniform cost model:

$$\mathfrak{C}(G) = 1 \text{ for all } G \in \mathcal{G}$$

Magic state cost model / T-count:

$$\mathfrak{C}(T) = 1 \text{ and } \mathfrak{C}(C) = 0 \text{ for all } C \text{ in the Clifford group.}$$



For **efficient** compilers The promise function $f(\epsilon)$ is often polylog $f(\epsilon) \leq A \log(1/\epsilon)^\gamma$

An **optimal** compiler will have the lowest possible $\mathfrak{C}(U)$ and $f(\epsilon)$

Solovay-Kitaev

Consider any universal gate set \mathcal{G} (generating a group) with uniform cost.

We can efficiently solve the compiling problem using $\mathfrak{c}(U) \leq O(\log(1/\epsilon)^\gamma)$ where γ is a constant dependent on \mathcal{G}

Comments:

For Clifford+T, the Solovay-Kitaev method gives $\gamma \leq 3.97$

See: Dawson & Nielsen *QIC*, 6 81 (2006)

Solovay-Kitaev is efficient but not optimal.



Random compiling problem

Given a V and $\epsilon > 0$ output a probability distribution of circuits, realising

$$\mathcal{E}(\rho) = \sum_i p_i U_i \rho U_i^\dagger$$

such that $d(\mathcal{E}, \mathcal{V}) \leq \epsilon$ and minimise $\mathfrak{C}(\mathcal{E})$

Def:

$$\mathcal{U}(\cdot) := U \cdot U^\dagger$$

Cost models for random circuits

Our results for worst-case

$$\mathfrak{C}(\mathcal{E}) := \max(\mathfrak{C}(U_i))$$

But average cost also interesting

$$\mathfrak{C}(\mathcal{E}) := \sum_i p_i \mathfrak{C}(U_i)$$



Measuring noise & Coherence

Desiderata for $d(\mathcal{E}, \mathcal{V})$: must compose nicely

$$d\left(\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \boxed{V} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array}, \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \boxed{\mathcal{E}} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \right) = d\left(\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \boxed{V} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array}, \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \boxed{\mathcal{E}} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right)$$

$$d\left(\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \begin{array}{c} \boxed{V^{(3)}} \\ \boxed{V^{(4)}} \end{array} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \begin{array}{c} \boxed{V^{(2)}} \\ \boxed{V^{(1)}} \end{array} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array}, \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \begin{array}{c} \boxed{\mathcal{E}^{(3)}} \\ \boxed{\mathcal{E}^{(4)}} \end{array} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \begin{array}{c} \boxed{\mathcal{E}^{(2)}} \\ \boxed{\mathcal{E}^{(1)}} \end{array} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \text{---} \end{array} \right)$$

$$\leq \sum_j \left(\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \boxed{V^{(j)}} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array}, \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \boxed{\mathcal{E}^{(j)}} \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \right) \quad \text{to leading order}$$

What distance we use matters!

Average infidelity

$$d_F(\mathcal{E}, U) := 1 - \int_{\psi} F(\mathcal{E}(\psi), U|\psi\rangle) = 1 - \int_{\psi} \langle \psi | U^\dagger \mathcal{E}(|\psi\rangle\langle\psi|) U | \psi \rangle$$

- looks simple,
- often used to report experimental results.
- **fails** our desiderata!

What distance we use matters!

Diamond norm distance

$$d_{\diamond}(\mathcal{E}, \mathcal{U}) := \frac{1}{2} \|\mathcal{E} - \mathcal{U}\|_{\diamond} = \frac{1}{2} \max_{\rho} \frac{\|(\mathcal{E} \otimes \mathbb{I})(\rho) - (\mathcal{U} \otimes \mathbb{I})(\rho)\|_1}{\|\rho\|_1}$$

where $\|X\|_1 := \text{Tr}[\sqrt{X^\dagger X}]$ is the Schatten 1-norm

- composed nicely (**yes to desiderata**);
- due to nice properties, used in proofs, e.g. of threshold theorem.
- So <1% threshold statements refer to diamond distance!

Algorithm level “coherent noise”:

Consider gate set $\mathcal{G} = \{A, B, C\}$



each individual gate is perfect

When we compile a target gate V

we perform $U \sim ABACABABABA = V e^{i\delta}$

The approximation error (due to finite sequence length) is a coherent error $e^{i\delta}$.

For coherent noise

$$d_{\diamond}(\mathcal{U}, \mathcal{V}) \sim \sqrt{d_F(\mathcal{U}, \mathcal{V})}$$

so choice of distance measures is important!

Contrast physical level coherent noise

Consider a scenario where we

want to implement sequence from gate set $\mathcal{G} = \{A, B, C\}$

but our experiment systematically performs $\tilde{\mathcal{G}} = \{\tilde{A}, \tilde{B}, \tilde{C}\}$

we have $\tilde{A} = Ae^{i\delta_A}$, $\tilde{B} = Be^{i\delta_B}$...

where δ are small noise terms.

Randomised compiling can help

say \mathcal{U} is target unitary

$\tilde{\mathcal{U}}$ is unitary with coherent noise

so that $d_{\diamond}(\tilde{\mathcal{U}}, \mathcal{U}) = \epsilon$ and $d_F(\tilde{\mathcal{U}}, \mathcal{U}) \sim \epsilon^2$

then there exist twirling schemes $\mathcal{E}(\rho) = \sum_i p_i U_i \rho U_i^\dagger$ where $U_i = P_i \tilde{\mathcal{U}} P_i'$

with $d_{\diamond}(\mathcal{E}, \mathcal{U}) \sim \epsilon^2$

E. Knill arXiv:quant-ph/0404104

Wallman & Emerson,

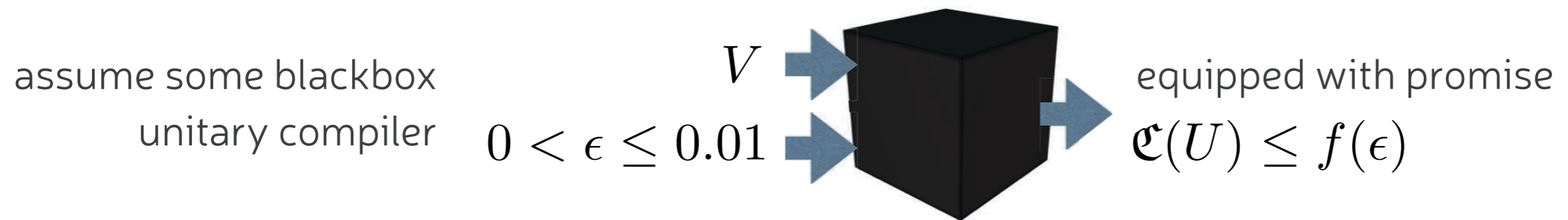
Phys Rev A **94**, 052325 (2016)



The RESULT

Preamble

Let \mathcal{G} be a universal gate set with cost measure \mathfrak{C}



Theorem

...then there exists a random sequence $\mathcal{E}(\rho) = \sum p_j U_j \rho U_j^\dagger$

with $d_\diamond(\mathcal{V}, \mathcal{E}) \leq 10\epsilon^2$ and cost $\mathfrak{C}(\mathcal{E}) \leq f(\epsilon)$

2nd Theorem (paraphrased)

For single qubit axial rotations: the assumptions can be relaxed and inequalities tightened slightly.

“same cost gets you better error suppression”

Theorem

...then there exists a random sequence $\mathcal{E}(\rho) = \sum_j p_j U_j \rho U_j^\dagger$
 with $d_\diamond(\mathcal{V}, \mathcal{E}) \leq 10\epsilon^2$ and cost $\mathfrak{C}(\mathcal{E}) \leq f(\epsilon)$

“same error suppression for lower cost?”

Corollary

if the unitary cost is polylog $\mathfrak{C}(U) \leq f(\epsilon) = A \log_2(1/\epsilon)^\gamma$

...then there exists a random circuit $\mathcal{E}(\rho) = \sum_j p_j U_j \rho U_j^\dagger$

with $d_\diamond(\mathcal{V}, \mathcal{E}) \leq \epsilon$ and cost $\mathfrak{C}(\mathcal{E}) \leq C^\gamma f(\epsilon) \sim (1/2)^\gamma$

$$C = \left(\frac{1}{2}\right) \left(1 + \frac{\log(A)}{\log(1/\epsilon)}\right) \xrightarrow{\epsilon \rightarrow 0} \left(\frac{1}{2}\right)$$

WE SAVE $\sim 2^\gamma$

Known “gamma” values:

For single qubit Clifford+T gate set
and T-count cost metric
 $\mathcal{C}(U) \leq f(\epsilon) = 9 \log(1/\epsilon)$
so $\gamma = 1$
and roughly 2x saving

	optimal unitary	random compiling
$\epsilon = 10^{-20}$	600	314 or less
$\epsilon = 10^{-10}$	300	165 or less
$\epsilon = 10^{-5}$	150	90 or less

$$\gamma \sim 3.97 \qquad 2^{3.97} \sim 15 - 16$$

Using Solovay-Kitaev for single qubits, so save

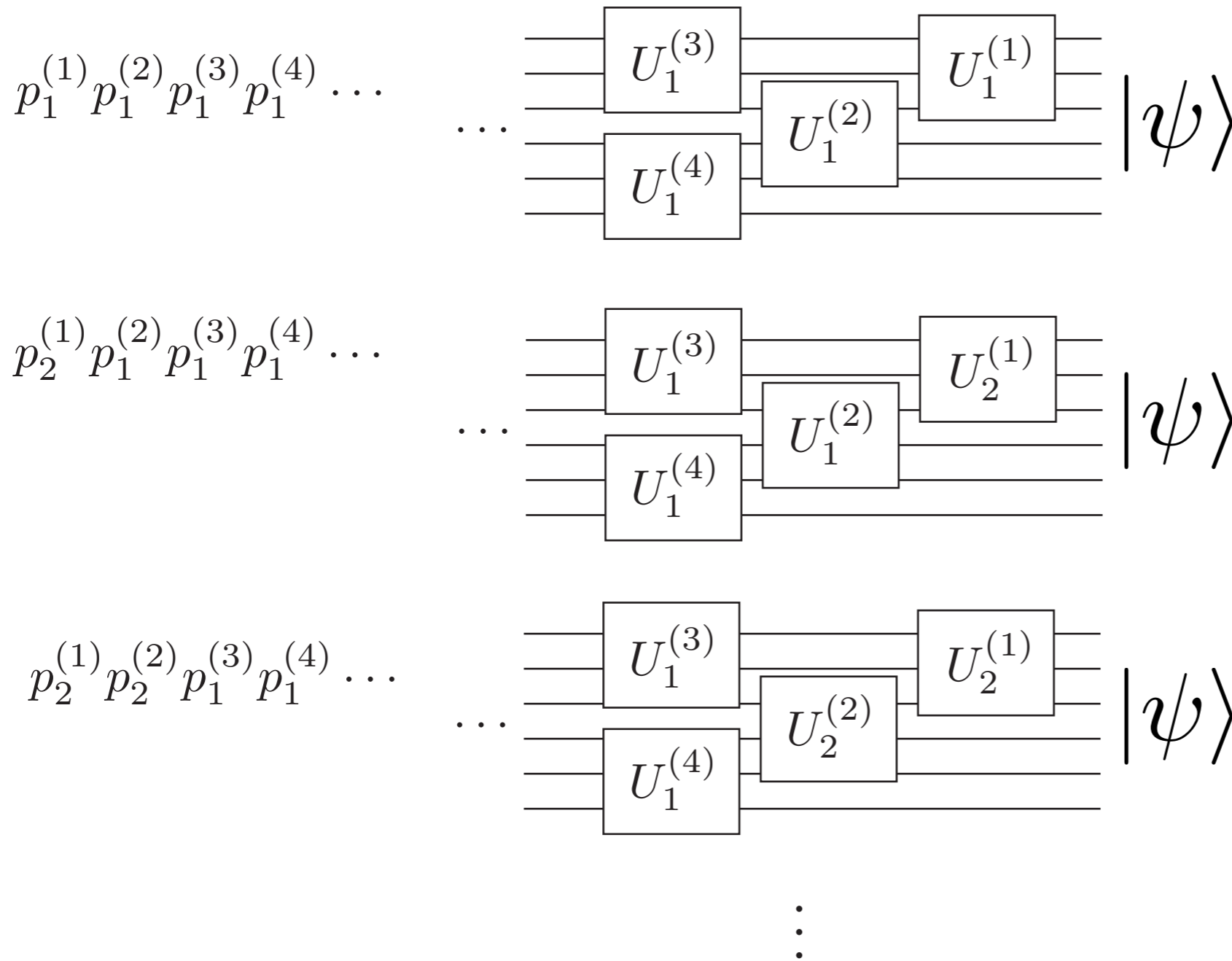
$$\gamma = ?$$

$$1 \ll \gamma$$

For optimal multiqubit synthesis

unknown, conjecture

Is derandomisation possible?



TRUE THAT:

For fixed input there is one “best” choice of unitaries that maximise fidelity

BUT:

- best unitary depends on choice of input state;
- can't be calculated without simulating the whole computation!



The PROOF

STEP 1. Show from “suitable” unitaries/Hamiltonians we can find random circuit with quadratically reduced error suppression.

STEP 2. Give concrete algorithm finding suitable Hamiltonians.

STEP 1. Show from “suitable” unitaries/Hamiltonians we can find random circuit with quadratically reduced error suppression.*

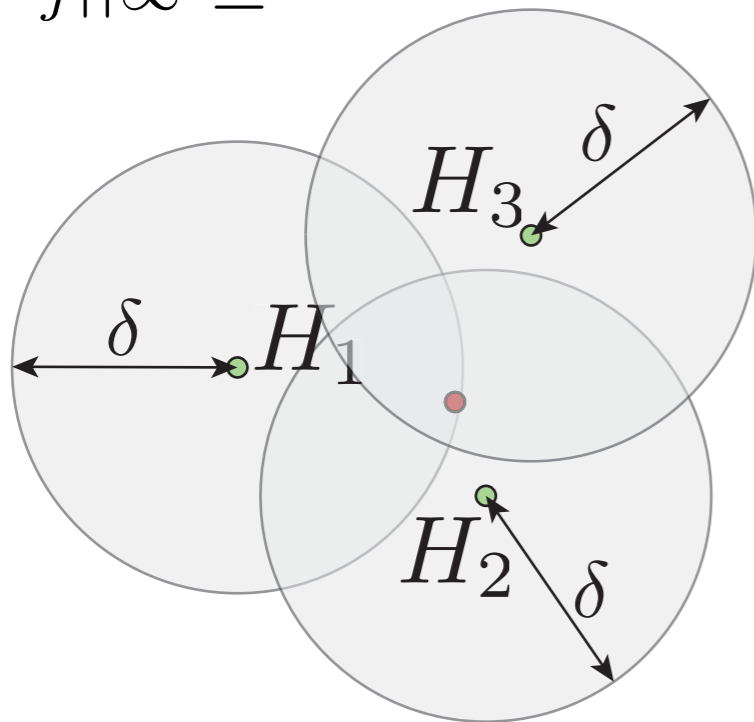
STEP 2. Give concrete algorithm finding suitable Hamiltonians.

* Similar “step 1” result found simultaneously by Hastings
arXiv:1612.01011 QIC **17** 0488 (2017)

STEP 1. Show from “suitable” Hamiltonians we can find random circuit

Lemma. Given target unitary V and a set of unitaries $U_j = V e^{iH_j}$ such that

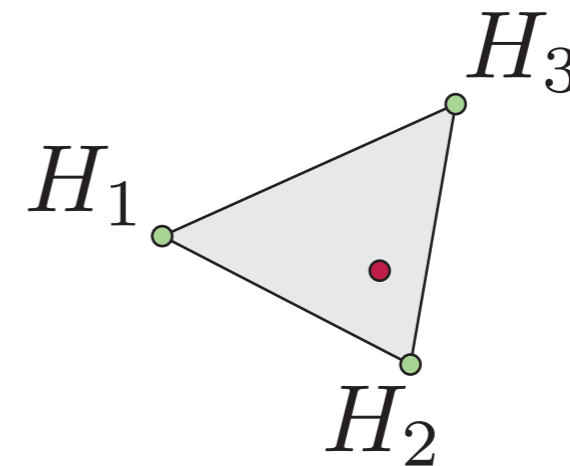
(1) All Hamiltonians within δ of origin
 $\|H_j\|_\infty \leq \delta$



(2) Hamiltonians enclose the origin

$$\exists p_j \quad 0 \leq p_j \quad \sum p_j = 1$$

$$\sum_j p_j H_j = 0$$



then $\mathcal{E}(\rho) = \sum_j p_j U_j \rho U_j^\dagger$ **satisfies** $d_\diamond(\mathcal{E}, \mathcal{V}) \leq 1.1 * \delta^2$

Sketch

start with

$$d_{\diamond}(\mathcal{E}, \mathcal{V}) = d_{\diamond}(\mathcal{V}^{\dagger} \circ \mathcal{E}, \text{identity})$$

and
$$\mathcal{V}^{\dagger} \circ \mathcal{E}(\rho) = \sum_j p_j e^{-iH_j} \rho e^{iH_j}$$

Expand exponential and find all first order terms cancel due to $\sum_j p_j H_j = 0$

$$\mathcal{V}^{\dagger} \circ \mathcal{E}(\rho) = \rho + O(H_j^2, \rho)$$

Subtracting identity gives

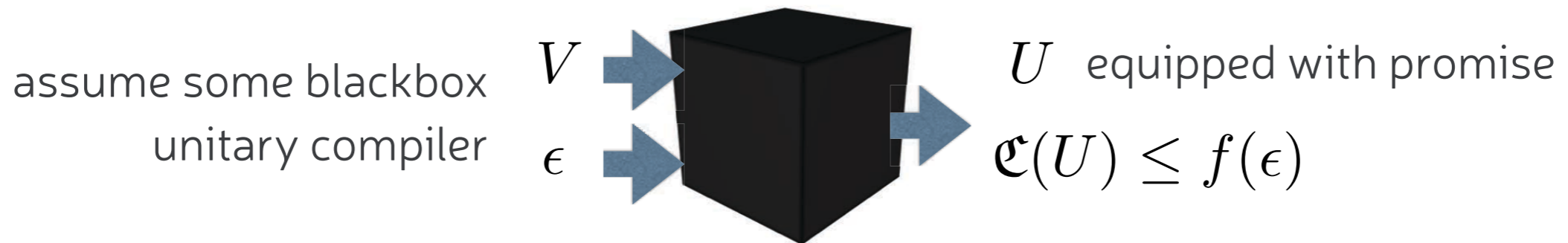
$$\mathcal{V}^{\dagger} \circ \mathcal{E}(\rho) - \rho = O(H_j^2, \rho)$$

Leaving H_j^2 and higher order terms

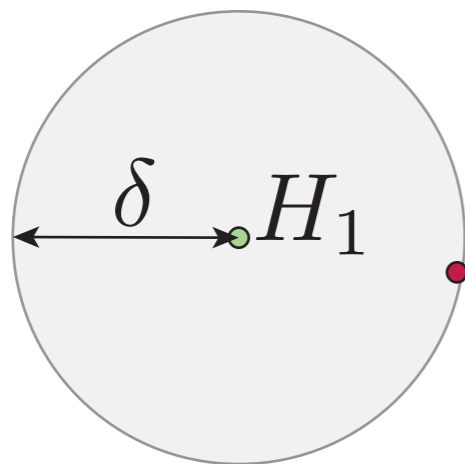
Take 1-norm, use triangle inequality, Hölder's inequality, carefully bound higher order terms...

$$\|\mathcal{V}^{\dagger} \circ \mathcal{E}(\rho) - \rho\|_1 = O(\|H_j\|_{\infty}^2) = O(\delta^2)$$

STEP 2. Give concrete algorithm finding suitable Hamiltonians.

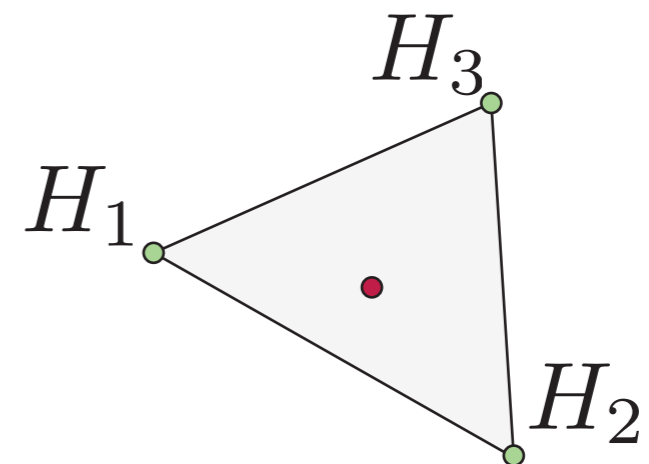


It follows $U_1 = V e^{iH_1}$ where $\|H_j\|_\infty \leq \delta = 3\epsilon + 7\epsilon^2$

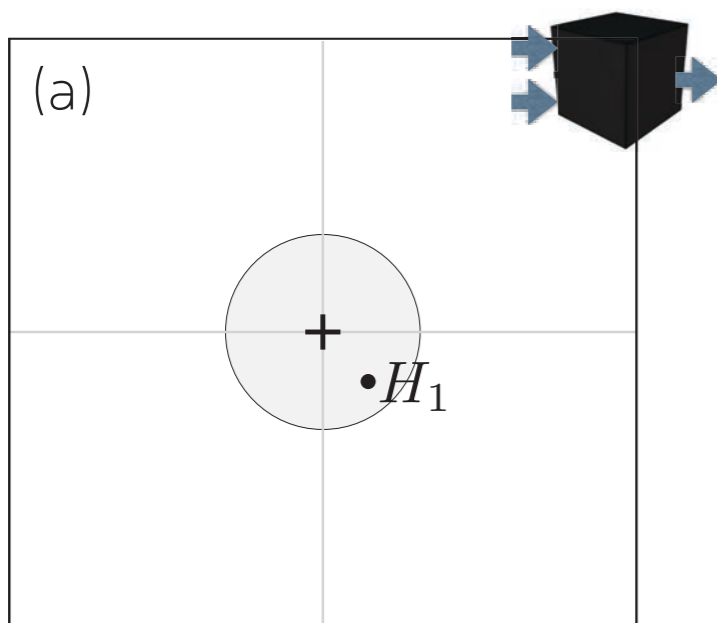


This gives 1 nearby Hamiltonian,

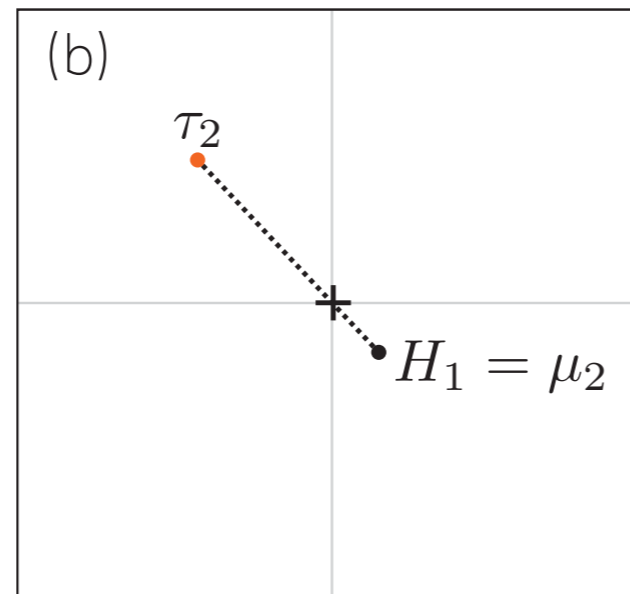
but in a D variable problem we typically need $D+1$ points to enclose the origin! e.g. For 2 variable problems we need 3 points/Hamiltonians



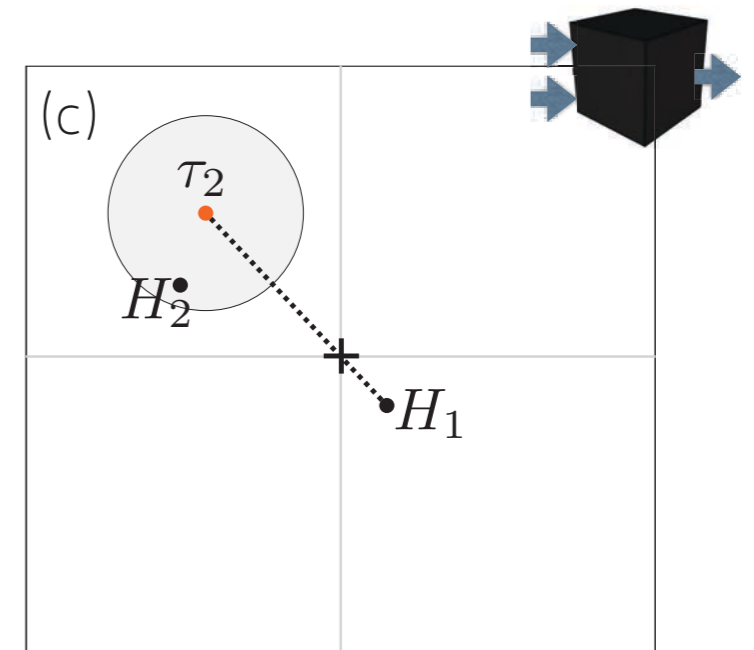
STEP 2. Give concrete algorithm finding suitable Hamiltonians.



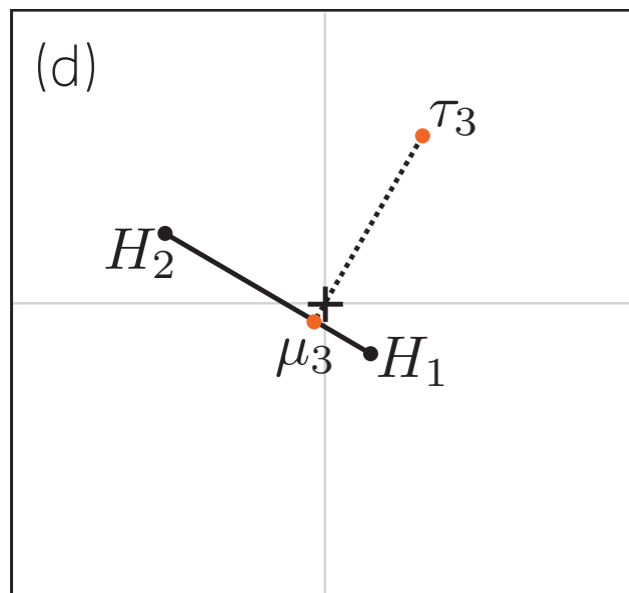
Find H_1 near origin



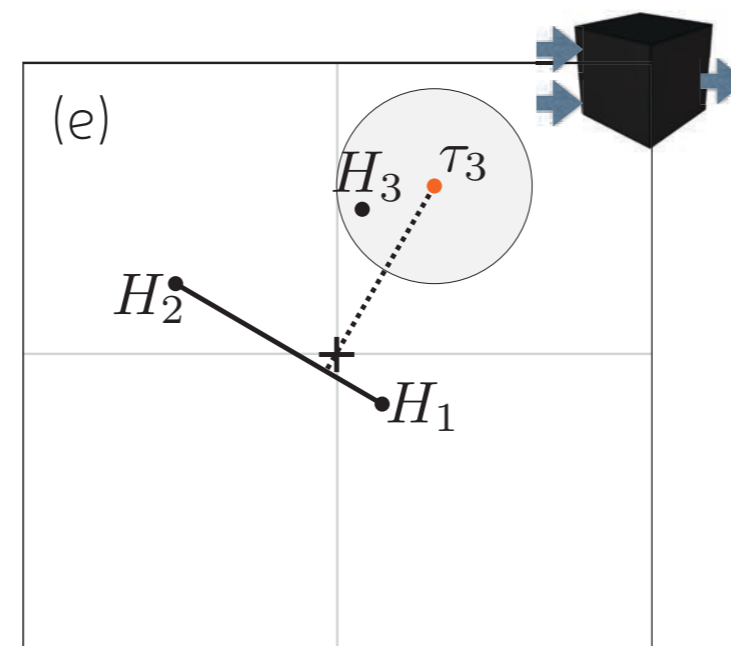
Select new target
 $\tau_2 \propto -H_1$



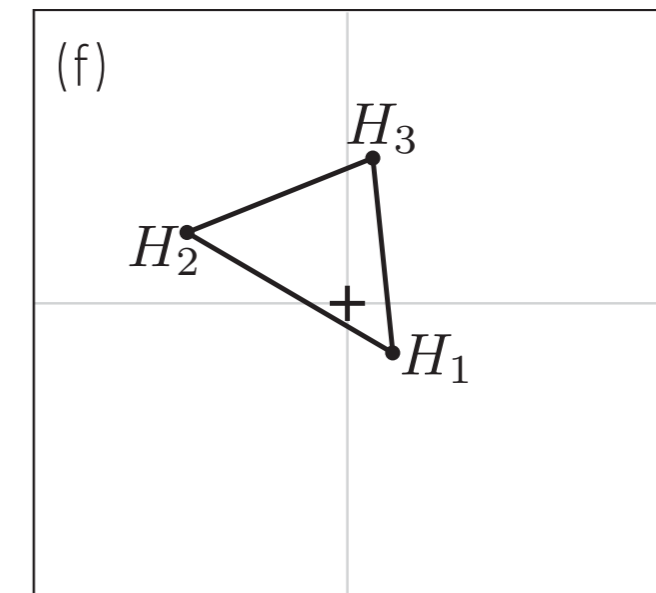
Find H_2 near τ_2



Find $\mu_3 = p_1 H_1 + p_2 H_2$
near origin and $\tau_3 \propto -\mu_3$



Find H_3 near τ_3



Find μ_4 near origin
EXIT when origin enclosed
or continue looping

STEP 2. Give concrete algorithm finding suitable Hamiltonians.

We prove that:

- Each iteration give a new H_j within $3\epsilon + 7\epsilon^2$ of origin.
- Each oracle call gives a new mixture

$$\mu_n = \sum_{j=1}^n p_j H_j$$

- When $\|\mu_n\|_\infty = 0$ the program EXITS, but can also EXIT when $\|\mu_n\|_\infty \ll \epsilon^2$
- and convergence is exponentially fast $\|\mu_n\|_\infty \leq 6\epsilon e^{-0.62n}$

Furthermore (unproven),

- Geometric intuition hints that algorithm terminates in constant number of steps

STEP 1. Show from “suitable” Hamiltonians we can find random circuit

showed $\mathcal{E}(\rho) = \sum_j p_j U_j \rho U_j^\dagger$ *satisfies* $d_\diamond(\mathcal{E}, \mathcal{V}) \leq 1.1 * \delta^2$

STEP 2. Give concrete algorithm finding suitable Hamiltonians.

found unitaries with $\|H_j\|_\infty \leq \delta = 3\epsilon + 7\epsilon^2$

COMBINED RESULT $d_\diamond(\mathcal{E}, \mathcal{V}) \leq 10\epsilon^2$



The END bit

We reviewed how

- Unitary compilers have coherent noise on the algorithm level.
- Coherent noise at the physical level can be quadratically reduced using random circuits.

This work showed

- Our random circuits give free quadratic error suppression;
- Free quadratic error suppression can be swapped for shorter gate sequences;
- For single qubit gates we save a factor 2x;
- Savings could be larger for multiqubit unitaries.

Future work

- Implement and numerically test.
- Classical runtime convergence in constant time?
- Better algorithms?

HIRING 4 POSTDOCS

QCDA - Quantum Code Design & Architecture

www.qcda.eu

Sheffield - Earl Campbell

London - Dan Browne

Paris - Anthony Leverrier & Jean Pierre Tillich

Delft - Barbara Terhal, Koen Bertels & Carmina Garcia Almudever

Munich - Robert Koenig

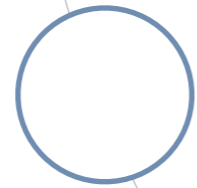
Funded through



QUANTERA



The
University
Of
Sheffield.



THANK YOU!

EPSRC

Engineering and Physical Sciences
Research Council