SAND2016-1759C

Gate Set Tomography and Beyond

Robin Blume-Kohout

John K. Gamble, Erik Nielsen, Kenny Rudinger, Peter Maunz, Dan Lobser, Kevin Fortier

Center for Computing Research
Sandia National Labs

CCR Center for Computing Research

Marcus da Silva, Diego Riste, Matt Ware, Colm Ryan (BBN)

U.S. DEPARTMENT OF ENERGY

Sandia National Laboratories

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Why do tomography?

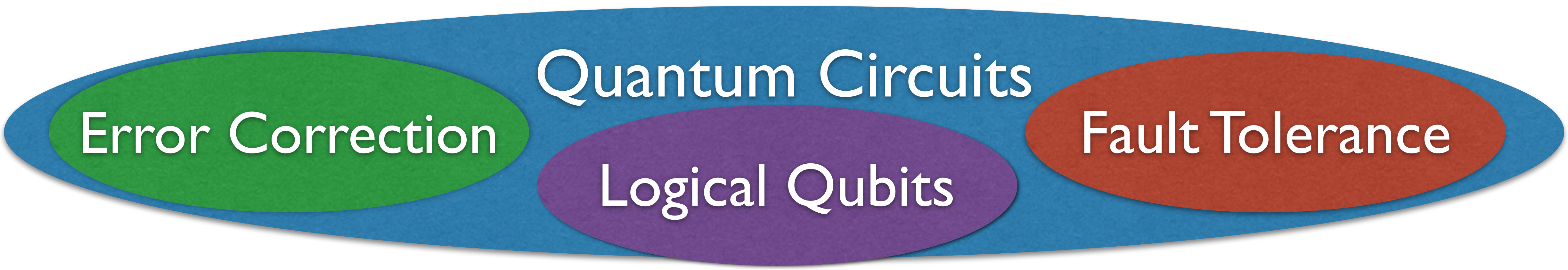
- To figure out what your quantum system is doing.

- To see how suitable it is for some task.

- To get information that helps you improve it (for that task).

QIP: an excuse for tomography

- Quantum information processing (QIP) is a very compelling task that justifies tomography and other QCVV methods.



- **Key ingredients**: initialize, logic gates, measure (+collapse).
  = "Gateset" describing the device (e.g. qubit).

- **Suitability for task**: low "error rates" (below FT threshold).

Tasks for tomographers

- Identify the critical parameters of gatesets (for QIP).

- Measure them accurately, efficiently, reliably.

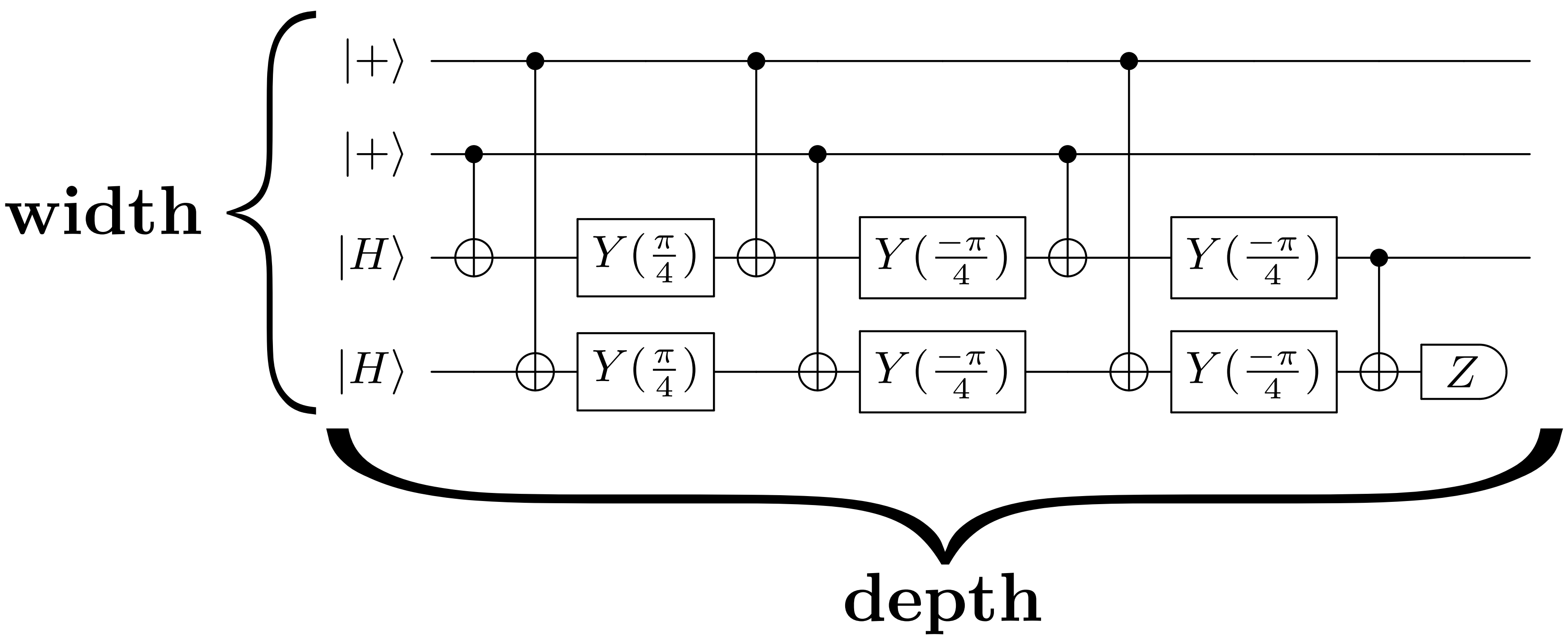
- Provide debugging information (to improve them).

Outline of this talk

1. “Error rates” for quantum gates.
2. Gate set tomography (GST)
3. “...and beyond”: open QCVV problems for logical qubits
4. Lessons learned

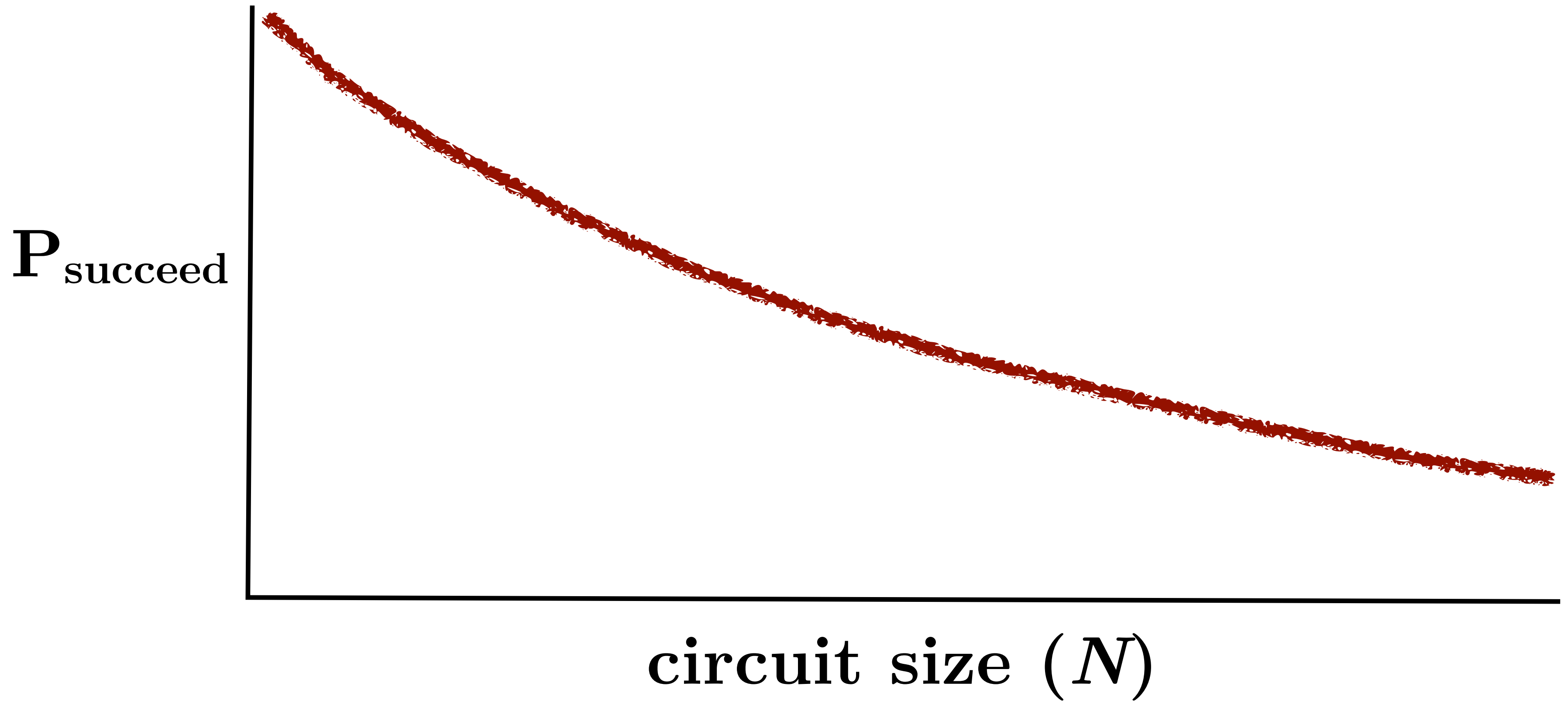
What does "error rate" mean?

- Goal of QIP: successfully perform quantum circuits.

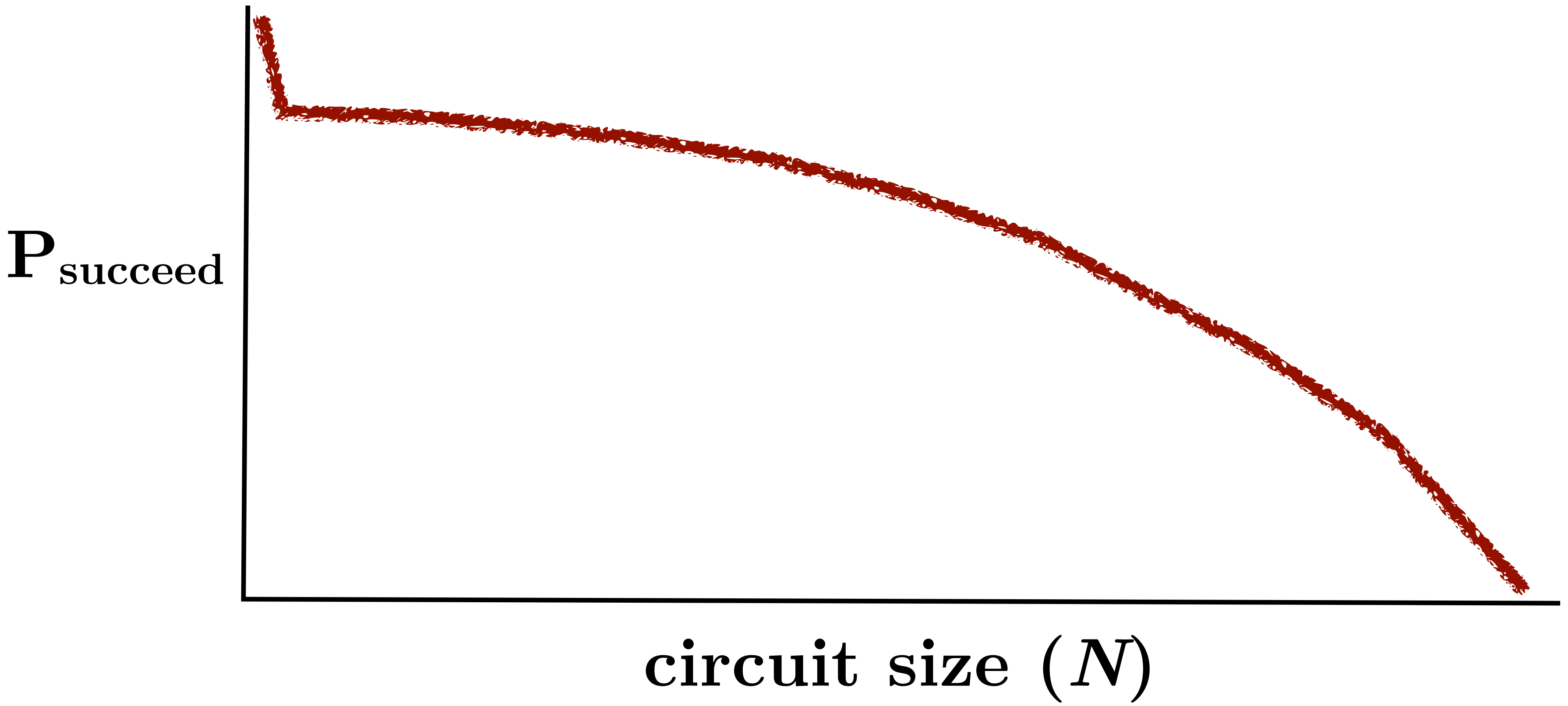


- Circuit <u>size</u> $\boldsymbol{N}$ = width x depth = # of gates that could fail.
- "Error rate p" ⇒ circuits with N ≪ 1/p probably won't fail.

If things were simple

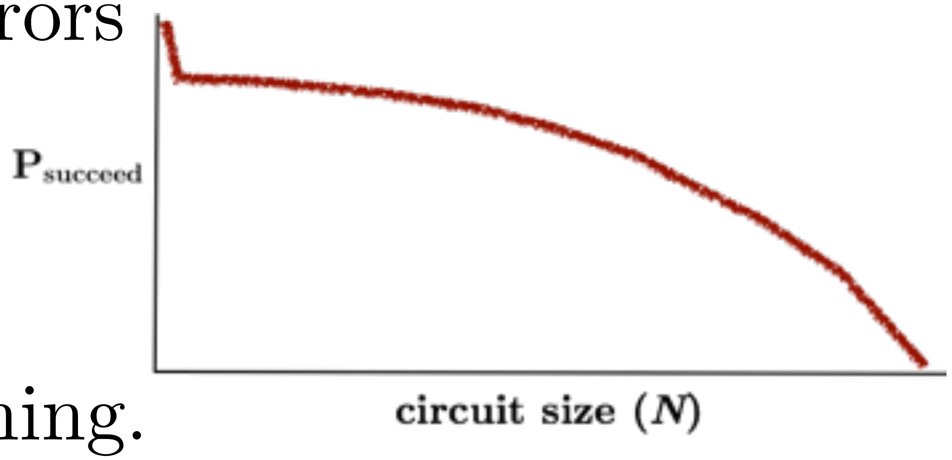


What can actually happen

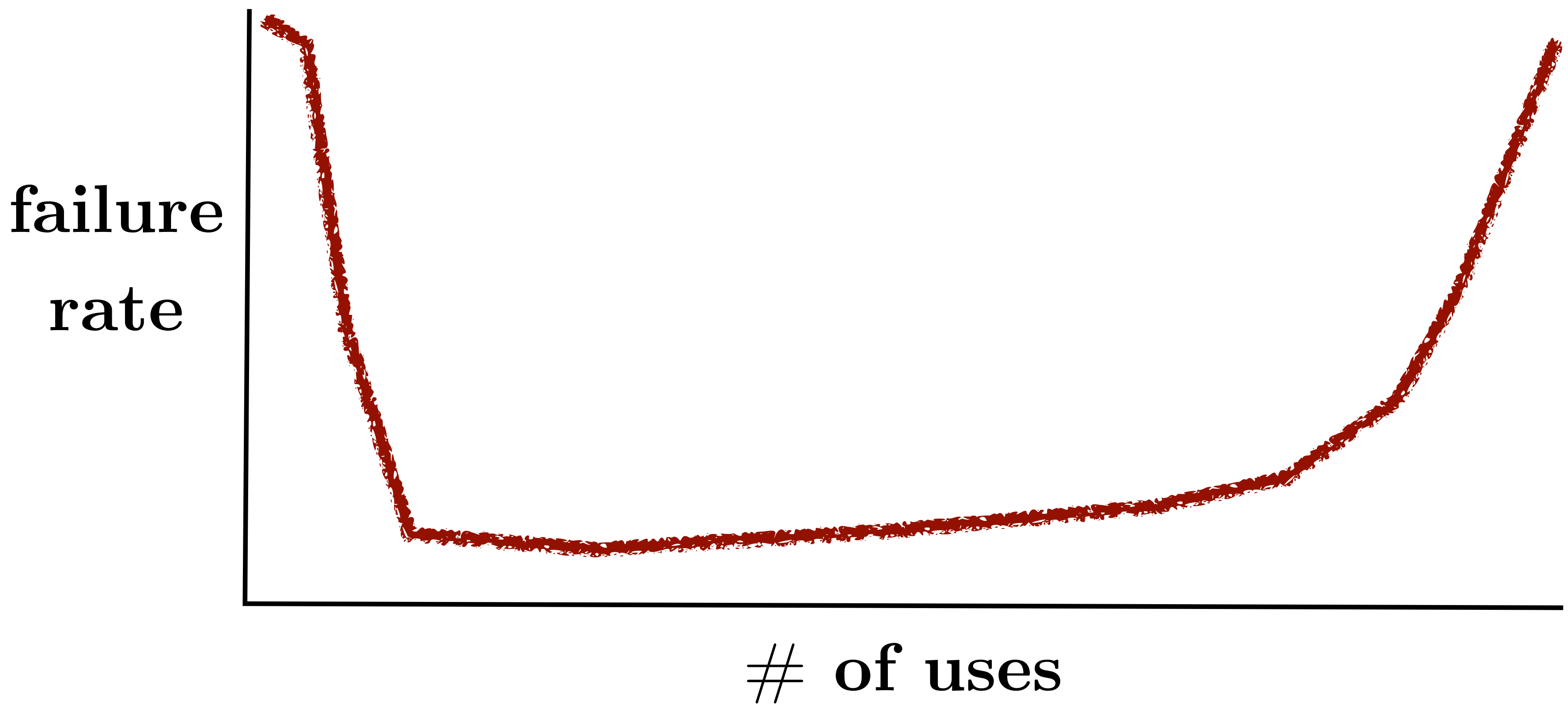


The bathtub curve

- In QIP, this is caused by coherent errors if they add up in a circuit.
  Or non-Markovian noise.

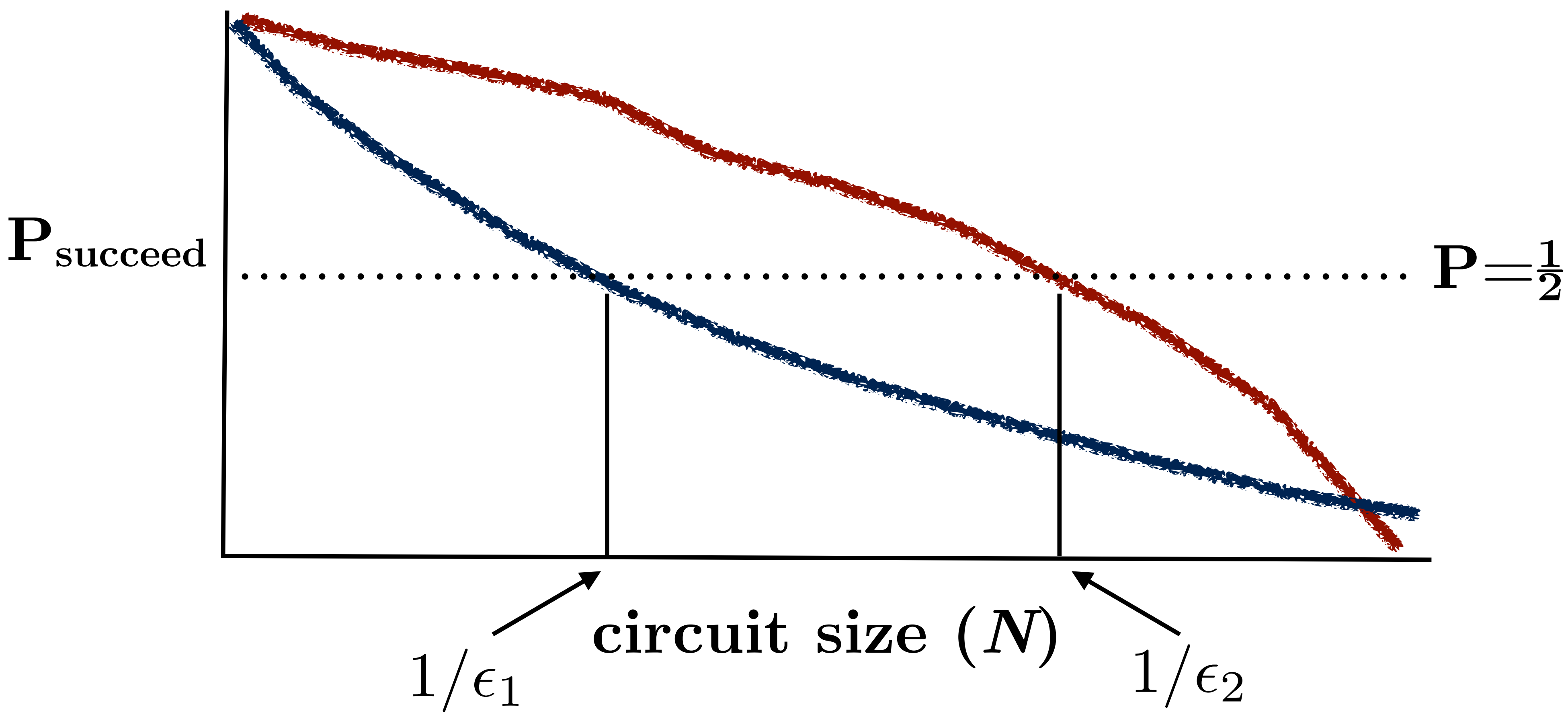


- Fortunately, this isn't <u>really</u> a new thing.
  e.g., consider metal fatigue...



Error rate ≈ 1/MTBF

- Define error rate in terms of smallest circuit with $P_{\mathrm{fail}} = O(1)$.



- Remaining question: **what circuits do we consider?**

Circuits that appear in FTQEC

- Many *different* circuits may be performed to achieve fault-tolerant quantum error correction.

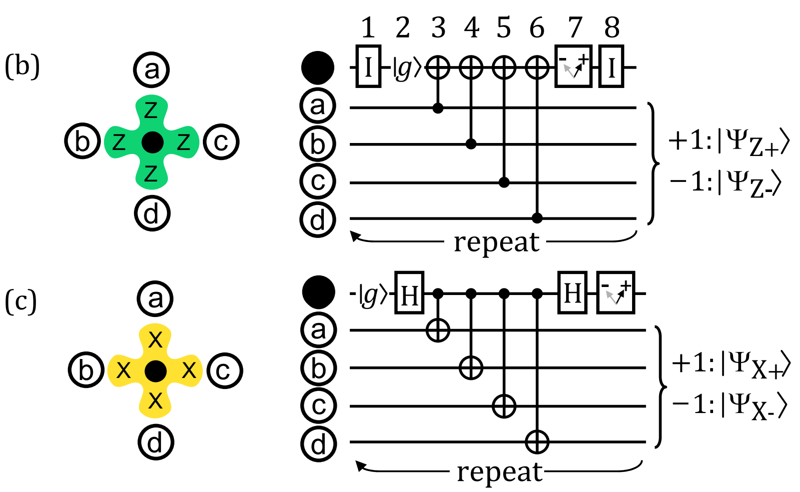
Circuits that appear in FTQEC

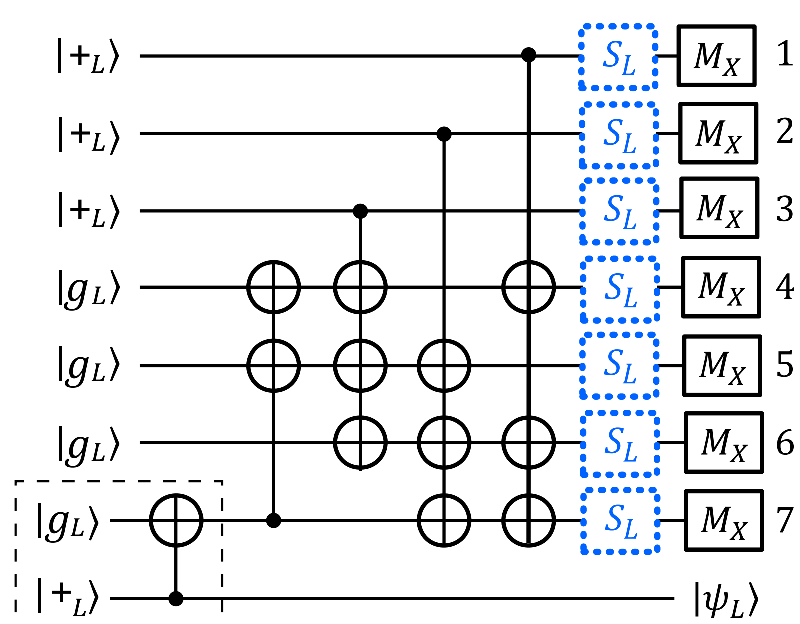
Surface codes: Towards practical large-scale quantum computation

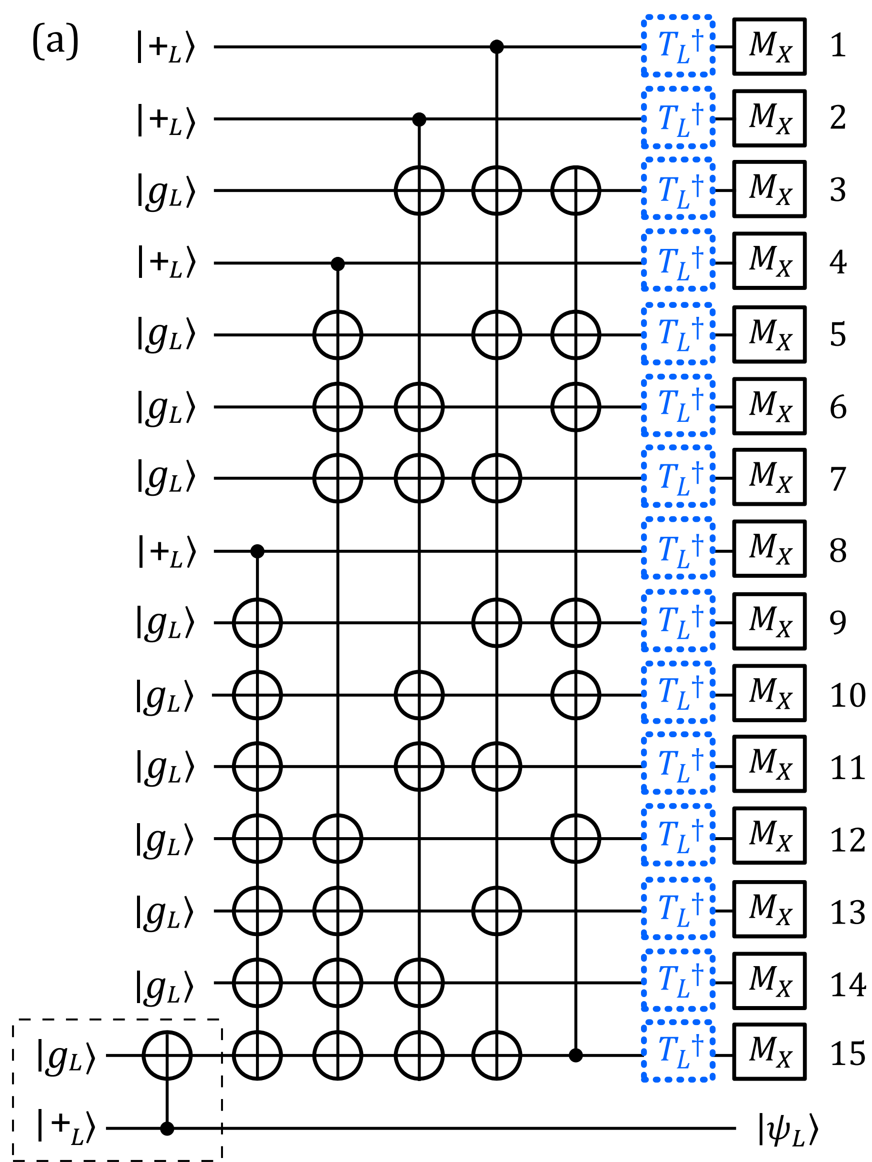
Austin G. Fowler
Centre for Quantum Computation and Communication Technology,
School of Physics, The University of Melbourne, Victoria 3010, Australia

Matteo Mariantoni
Department of Physics, University of California, Santa Barbara, CA 93106-9530, USA and
California Nanosystems Institute, University of California, Santa Barbara, CA 93106-9530, USA

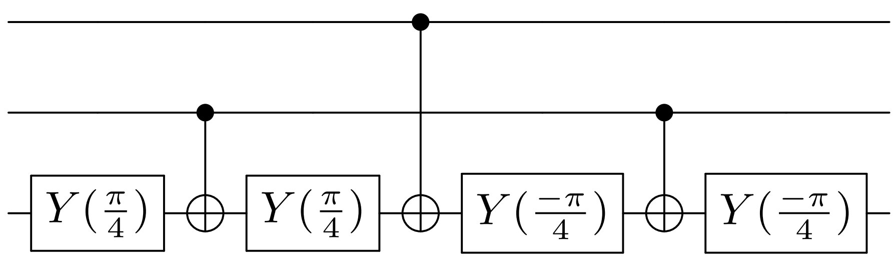
John M. Martinis and Andrew N. Cleland
California Nanosystems Institute, University of California, Santa Barbara, CA 93106-9530, USA
(Dated: October 26, 2012)





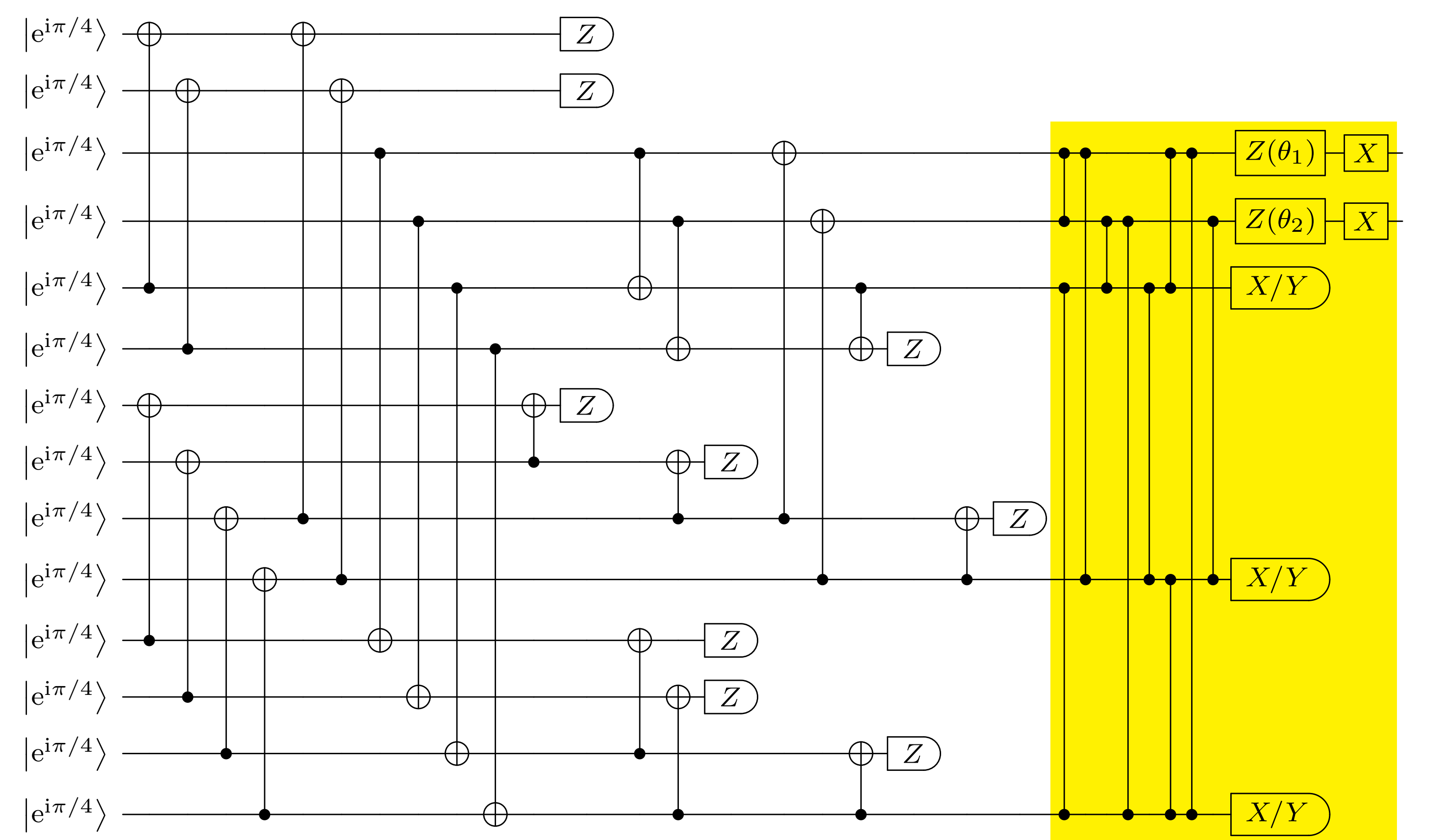


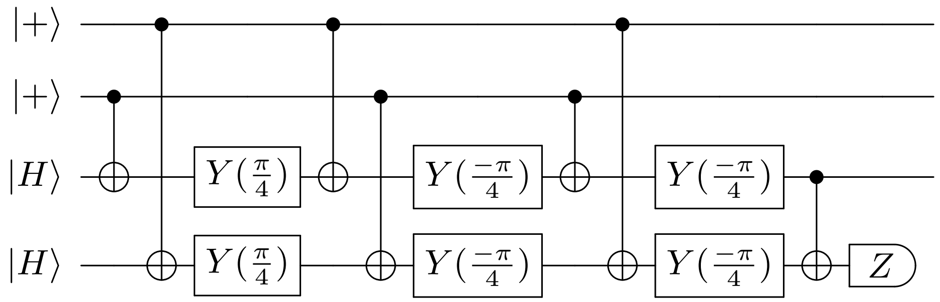
Circuits that appear in FTQEC

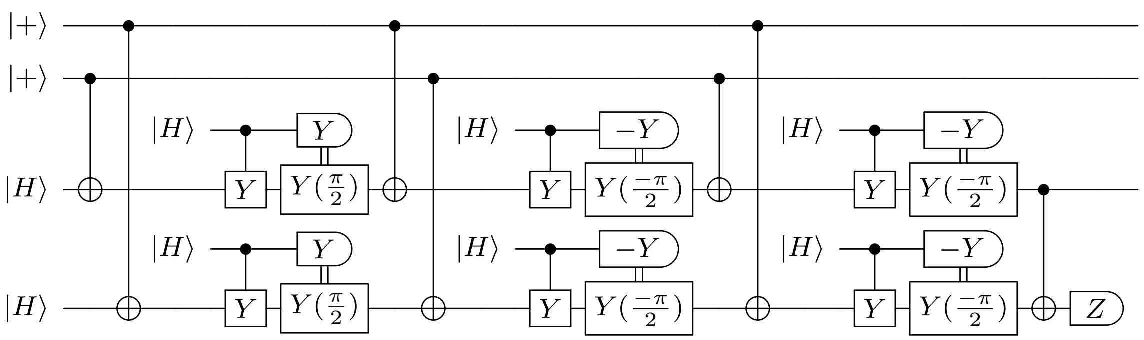


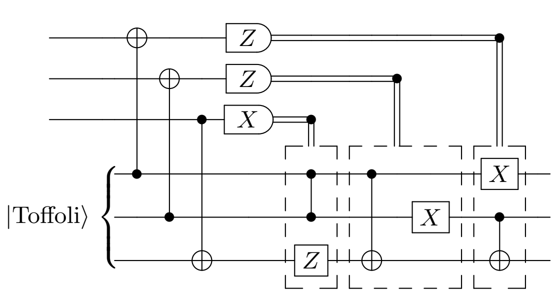
Distilling one-qubit magic states into Toffoli states

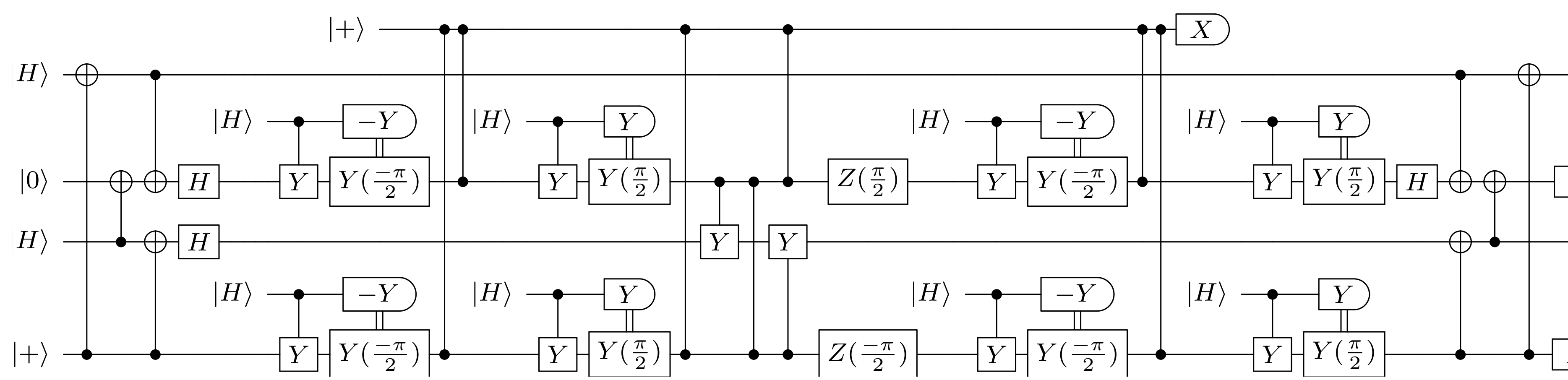
Bryan Eastin*
Northrop Grumman Corporation, Baltimore, MD







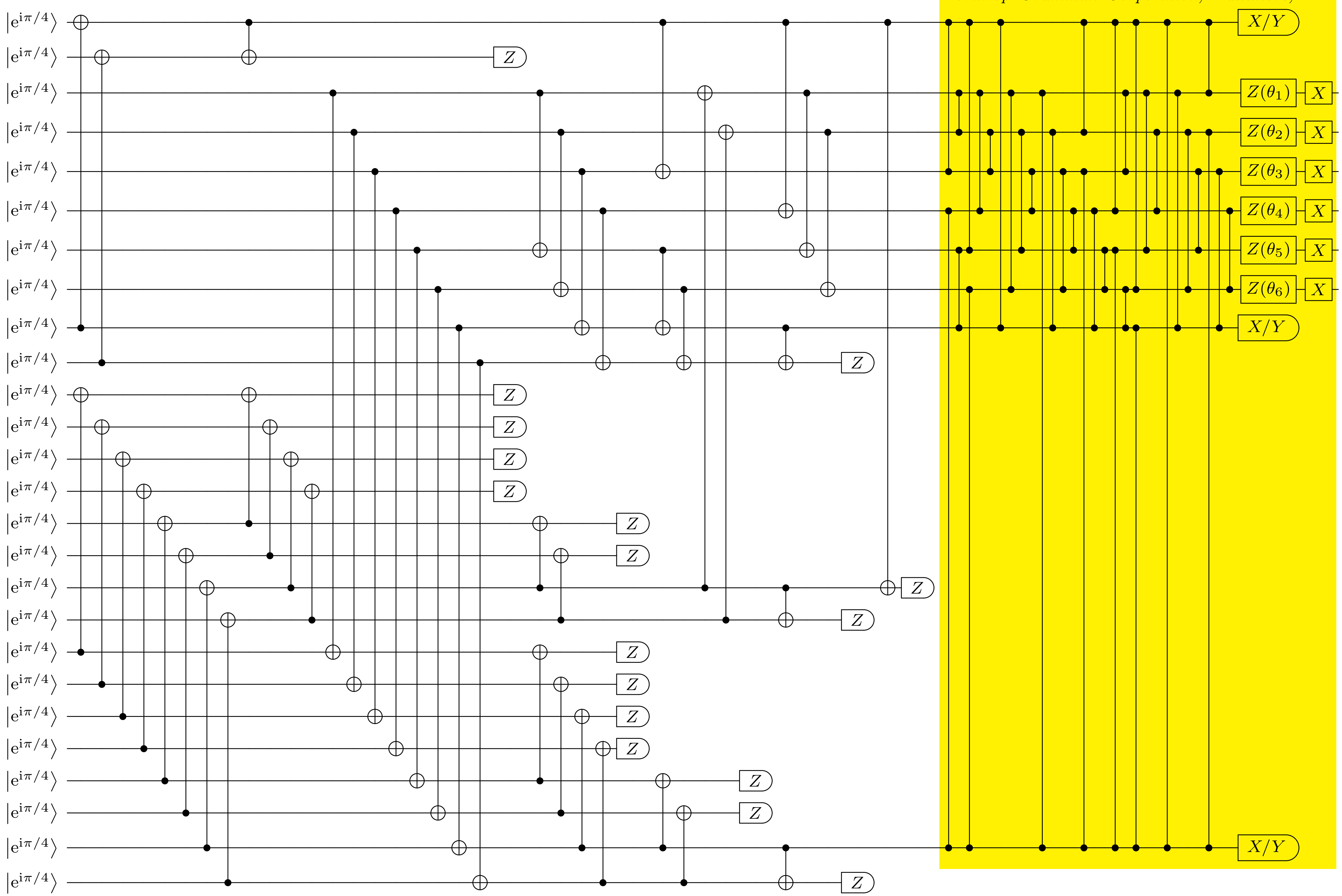




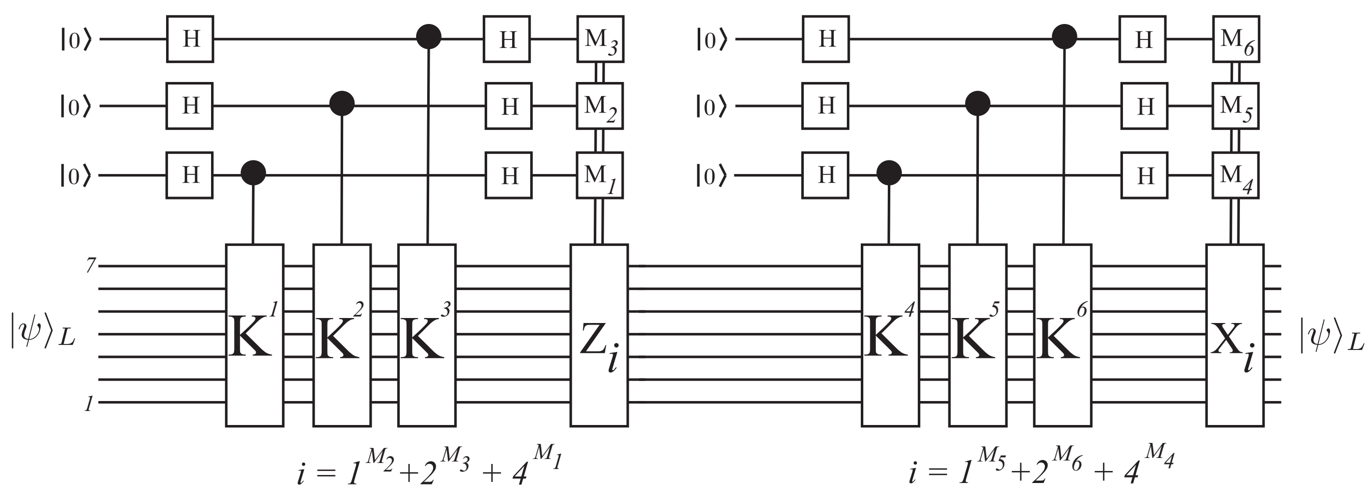
Circuits that appear in FTQEC

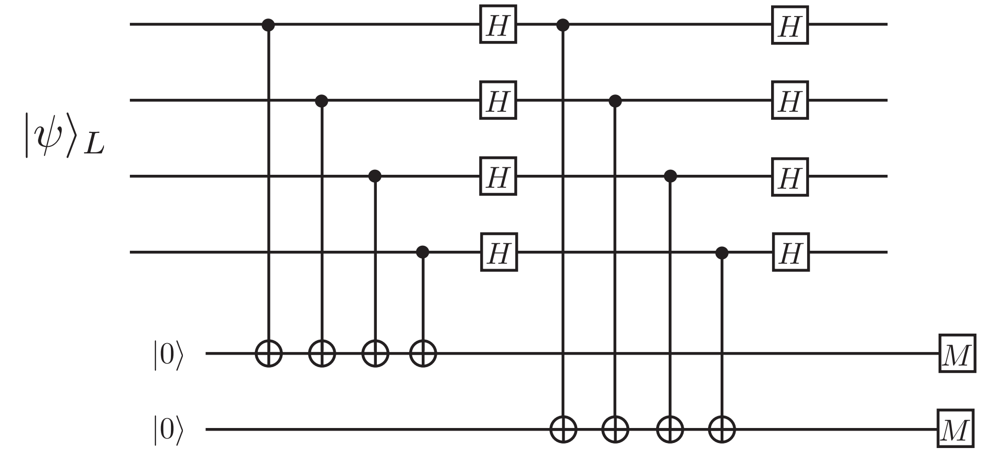
Distilling one-qubit magic states into Toffoli states

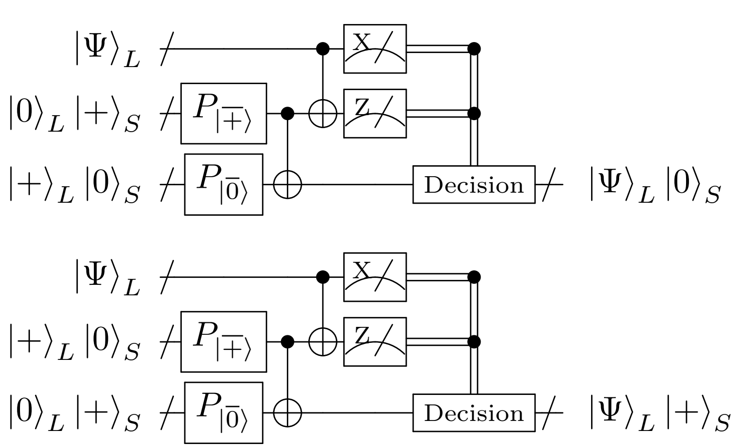
Bryan Eastin*
Northrop Grumman Corporation, Baltimore, MD

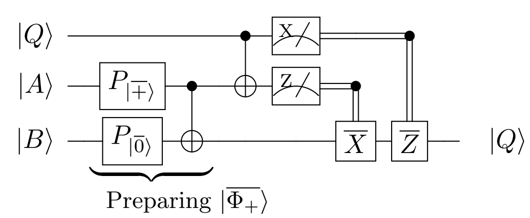


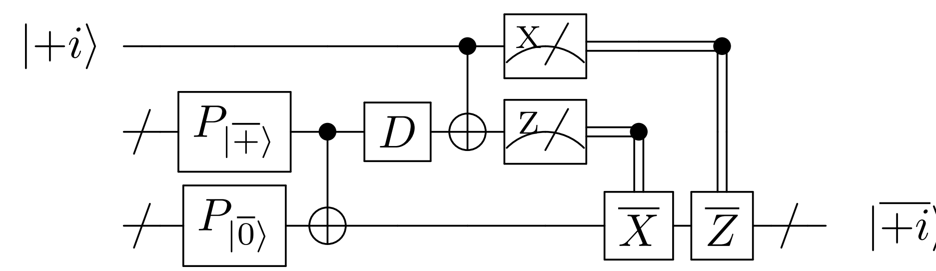
Circuits that appear in FTQEC

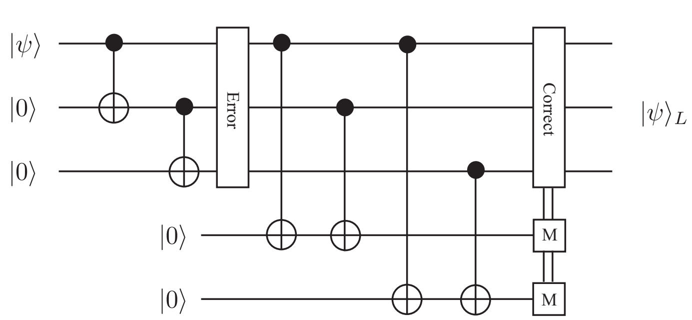


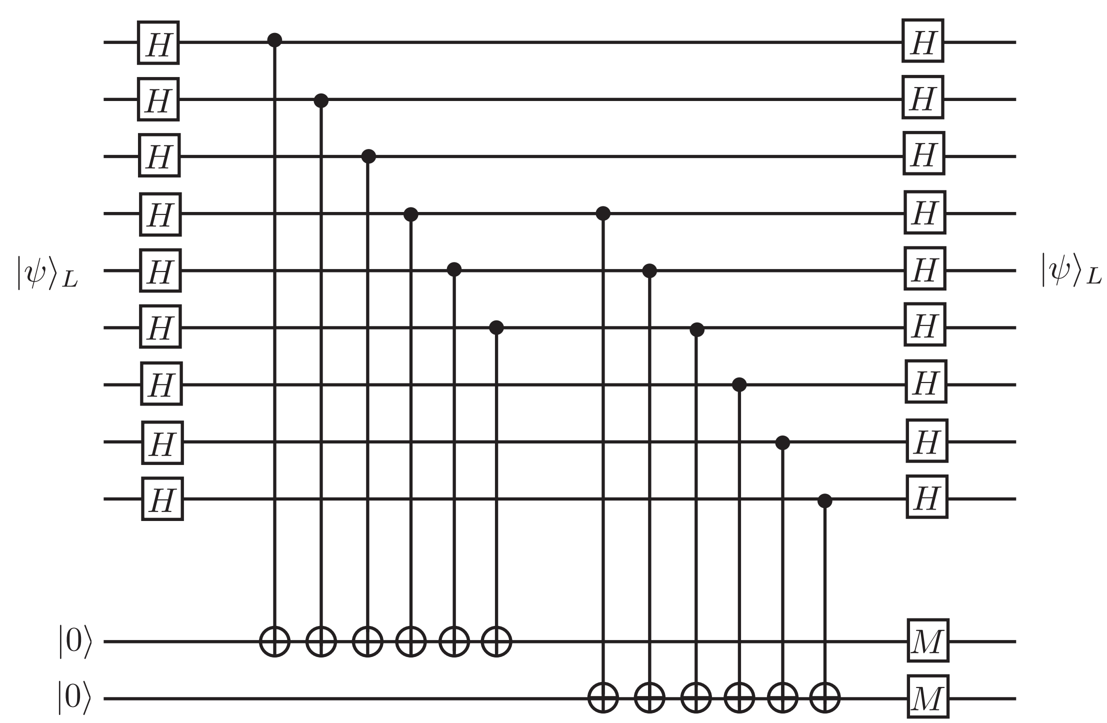


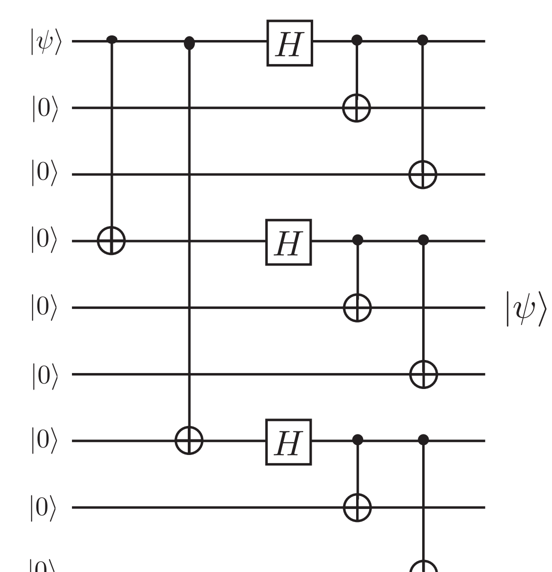












Circuits that appear in FTQEC

- Many *different* circuits may be performed to achieve fault-tolerant quantum error correction.
- We don't know which ones will be used.
- The "error rate" we should worry about is the <u>worst case with respect to all possible (plausible) quantum circuits</u>.

  (*not* just the average over <u>random</u> circuits)

- How do we quantify this without searching over 2^N circuits?

The diamond norm

- The diamond norm error metric (for gates) is a tight upper bound for the worst-case-over-circuits error rate.

$$\begin{aligned}\epsilon_\diamond &\equiv \|G - G_{\text{ideal}}\|_\diamond \\ &= \max_\rho \|(G \otimes \mathbb{1})[\rho] - (G_{\text{ideal}} \otimes \mathbb{1})[\rho]\|_{\text{tr}}\end{aligned}$$

- Why?
  1. It is sub-additive when gates are composed.
  2. When it gets to O(1), it implies $\boldsymbol{P}_{\text{fail}} = \text{O}(1)$.
  3. It can be saturated (by coherent errors).

- *BTW, the max over input states is a red herring...*

  ...and so is the extension to an ancilla...

  ...and so is the single shot (Helstrom) interpretation.

Summary

- One thing tomographers need to do is estimate “error rate”.
- Success probability doesn’t always decay exponentially.
  - c.f. coherent *and* non-Markovian errors.
- 1/MTBF is a pretty good way to define “error rate”.
- Worst case (over circuits) is important.
- Diamond norm captures this behavior pretty well.

- (MTBF can be applied to non-Markovian errors too!
  But will require a different mathematical definition)

Part 2: Gate set tomography

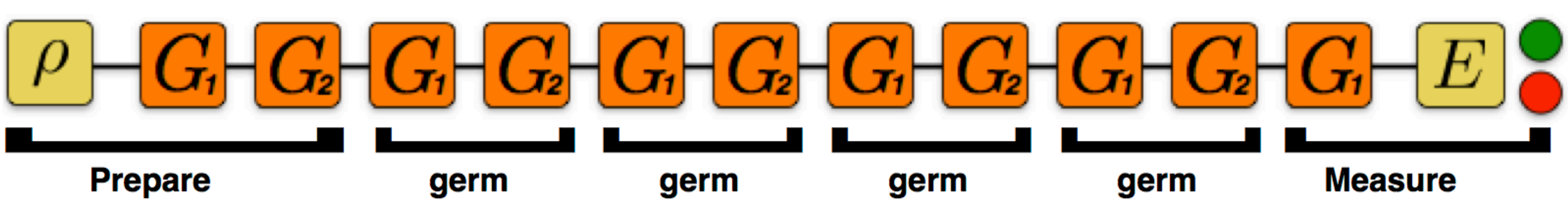
Gate set tomography

We developed GST to fill two needs: *debugging* and *certifying* as-built qubits.

Capabilities vs Protocols	No Calibration Required	Detailed Debug Info	Efficiently Certifies ◊-norm	Detects Non-Markovian noise
Randomized Benchmarking	✓	✗	✗	✓
Process Tomography	✗	✓	✗	✗
Gate Set Tomography	✓	✓	✓	✓

GST characterizes *all* the logic gates at once (like RB)...
...but reports full process matrices for them (like process tomography).

GST uses data from *gate sequences* (like RB), but these sequences are *structured* and *periodic*. Small errors get amplified and are measured precisely.



Gate set tomography

General GST Properties

- No reliance on pre-calibrated operations (gates, POVMs, etc).
- Unconditional reliability (except due to non-Markovian effects).
- Resource complexity (# settings, clicks, etc) is only slightly higher than that of process tomography on all gates in the gateset.

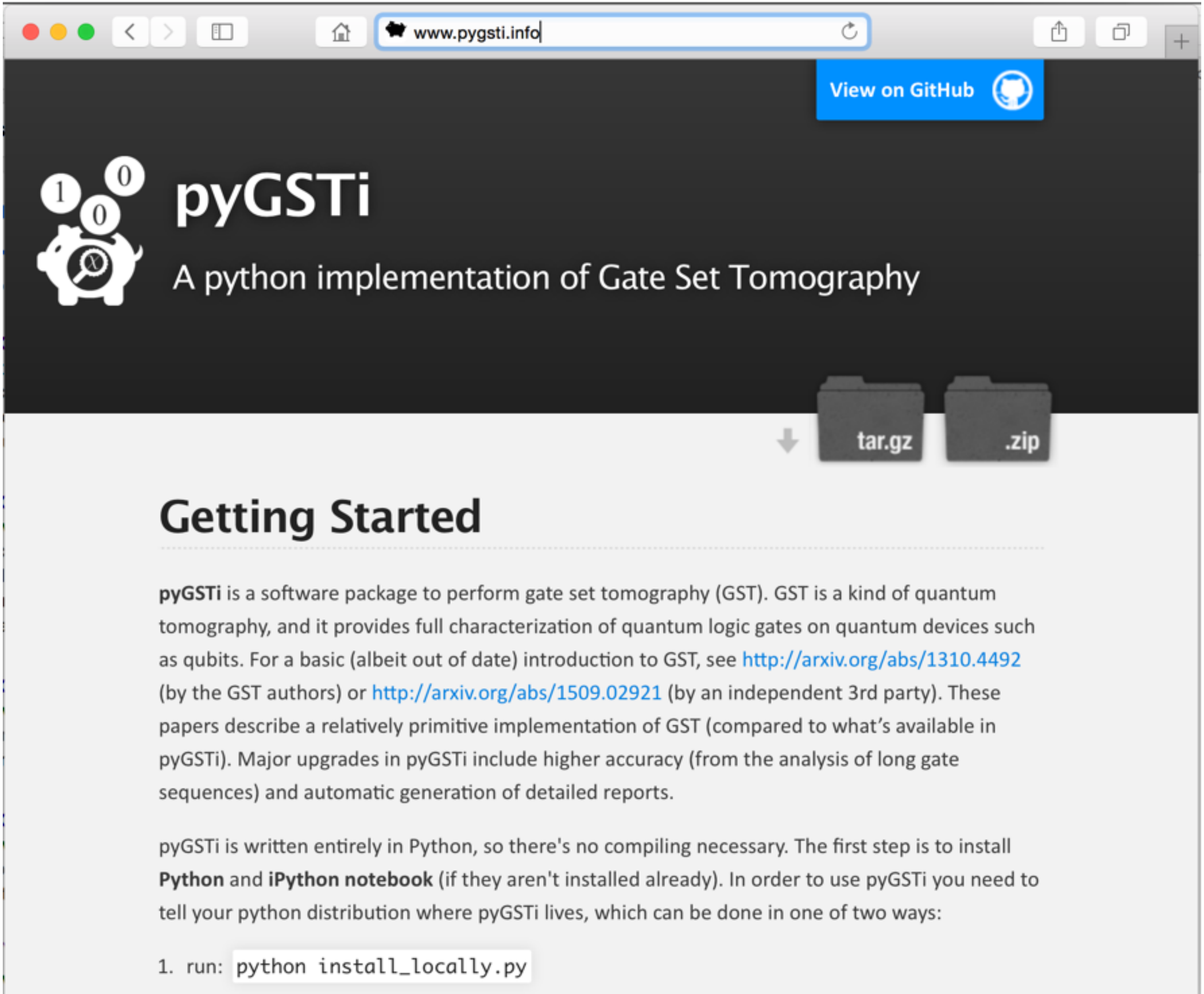
Sandia's "PyGSTi" code

- Estimates the RB number as accurately as RB itself.
- Estimates *all* gate parameters to high accuracy (including derived quantities, e.g. $\diamondsuit$-norm)
- Usually detects non-Markovian noise/errors/effects.
- Gateset estimate and derived quantities can be equipped with error bars.

Labs that have used GST

- Eriksson Si hybrid qubit (UWisc)
- Maunz Yb^+ ion qubit (Sandia)
- Carroll Si spin qubit (Sandia)
- Saffman neutral atom (UWisc)
- Morello Si donor qubit (UNSW)
- BBN transmon qubit
- Schoelkopf "cat" (Yale)
- Monroe ion qubit (UMaryland)
- Palmer superconducting (LPS)
- (...several others...)

...and now anyone!



Getting Started

pyGSTi is a software package to perform gate set tomography (GST). GST is a kind of quantum tomography, and it provides full characterization of quantum logic gates on quantum devices such as qubits. For a basic (albeit out of date) introduction to GST, see http://arxiv.org/abs/1310.4492 (by the GST authors) or http://arxiv.org/abs/1509.02921 (by an independent 3rd party). These papers describe a relatively primitive implementation of GST (compared to what's available in pyGSTi). Major upgrades in pyGSTi include higher accuracy (from the analysis of long gate sequences) and automatic generation of detailed reports.

pyGSTi is written entirely in Python, so there's no compiling necessary. The first step is to install **Python** and **iPython notebook** (if they aren't installed already). In order to use pyGSTi you need to tell your python distribution where pyGSTi lives, which can be done in one of two ways:

1. run: `python install_locally.py`

Principles of GST

- **Gates are *relational*.** Initial states are prepared using gates, and final measurements are performed using gates.

  Process tomography is not about "How does *this* process transform *these* input states?"

  It's about "How does *this* process relate to these *other* processes?"

- The existence of a *gauge* for gatesets is a direct consequence.

- The probabilities for sufficiently many *gate sequences* (circuits) determine a gateset, up to gauge. These are estimated from data.

- All ensuing discussion is about: (1) what sequences to measure; and/or (2) how to analyze the data.

“Overkill tomography”. IBM 2012-13

- What to measure: all gate sequences of length ≤ 3.

$$P_{ijk} = \langle\langle E|\, G_i G_j G_k \,|\rho\rangle\rangle$$

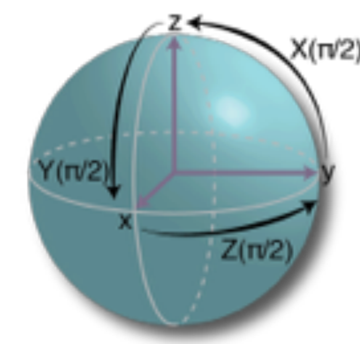
- Analysis method: maximum likelihood (MLE). Local optimization using target gates as a starting point.

- **The Good**: first implementation — original idea. Usually worked.

- **The Bad**: both aspects ad-hoc. Likelihood function known to be non concave (probabilities nonlinear in gates), so hard to optimize reliably.

Linear GST (LGST). Sandia, 2013

- What to measure: specific “fiducial sandwich” sequences of length L ≤ 7.



$$P_{ijk} = \langle\langle E|\,\{F_i\}G_j\{F_k\}\,|\rho\rangle\rangle$$

$$\{F_i\} = \{\emptyset, G_x, G_y, G_xG_x, G_xG_xG_x, G_yG_yG_y\}$$

- Analysis method: closed-form linear algebra.

- **The Good**: incredibly fast, 100% reliable.

- **The Bad**: Not very accurate.
  Unweighted linear inversion is statistically unsophisticated.
  Minimal ability to detect non-Markovian noise.

Long sequence GST. Sandia, 2013-16

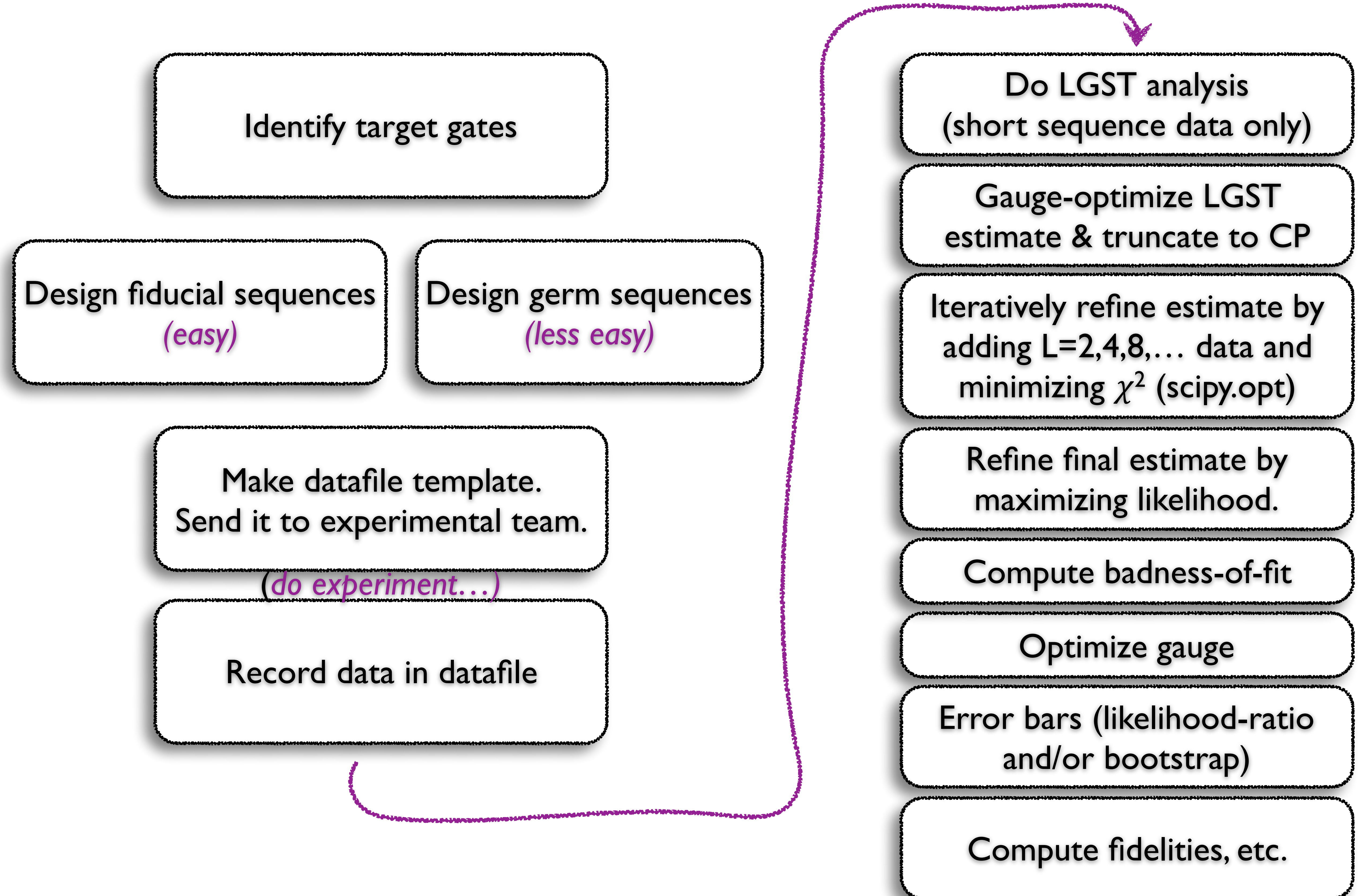
- What to measure: Specific "germ power" fiducial sandwich sequences of length $2^L + \mathrm{O}(1)$, up to $2^L = 8192$.

$$P_{ijkL} = \langle\langle E|\,\{F_i\}\{g_j\}^L\{F_k\}\,|\rho\rangle\rangle$$

$$\{g_j\} = \{G_x, G_y, G_i, G_xG_y, G_xG_xG_iG_y, \mathrm{etc}\}$$

- Analysis method: naive least squares, min-χ^2, or MLE.
- **The Good**: Heisenberg accuracy. Highly reliable. Extensive detection of non-Markovian noise.
- **The Bad**: computationally demanding.

PyGSTi work flow



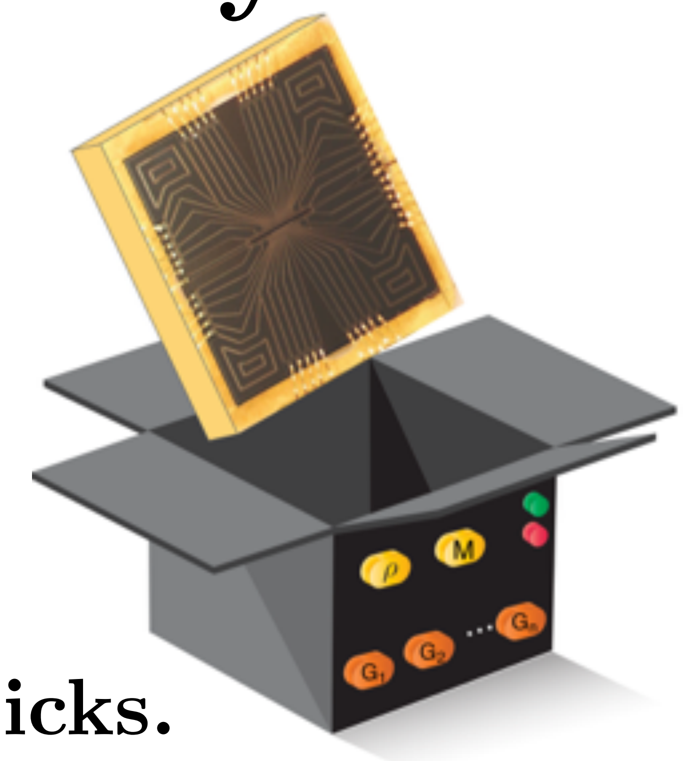
Report generation

- With PyGSTi, you can read in a dataset, analyze it, and generate a comprehensive human-readable report on the gates with a single command.
- Reports run 15-30 pages and contain:
  - Summary of experimental protocol & target gates
  - Estimated gates (including state prep and measurement)
  - Derived quantities (fidelities, $\diamondsuit$-norms, rotation angles, rotation axes, SPAM parameter, etc...)
  - Model violation info ("signature of non-Markovianity")

GST in action

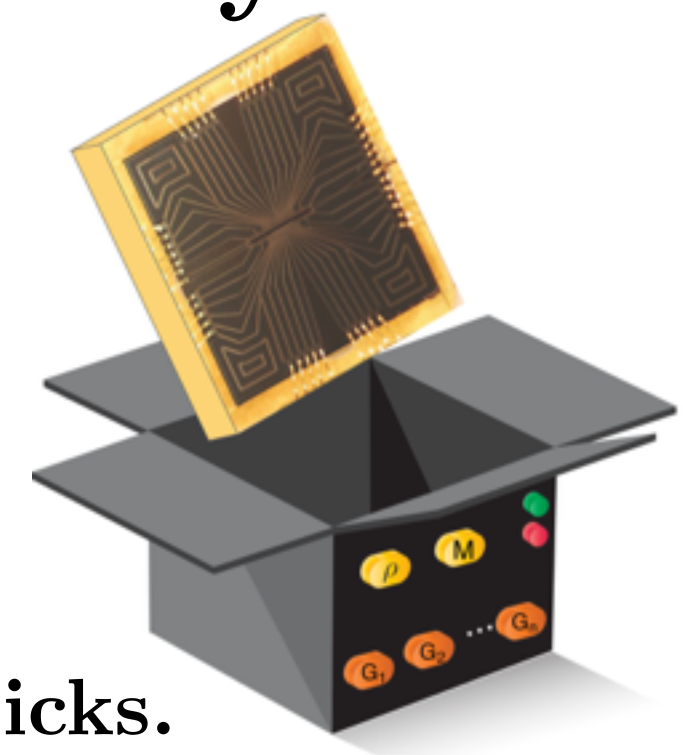
1. 1-qubit gates with $<10^{-4}$ $\Diamond$-norm error.
2. Error bars of $\pm 10^{-5}$.
3. 2-qubit GST

We tomographed a qubit really really really precisely.



- Trapped-ion (Yb^+) qubit in Peter Maunz's lab (March 2015)
- 3 gates ($\mathbf{X}_{\pi/2}$, $\mathbf{Y}_{\pi/2}$, Idle).
  6 fiducials. 11 germs. L = 1,2,4,...,8192.
  4657 sequences. 50 counts/each. **230,000 clicks.**

We tomographed a qubit really really really precisely.



- Trapped-ion (Yb^{+}) qubit in Peter Maunz's lab (March 2015)
- 3 gates ($\mathbf{X}_{\pi/2}$, $\mathbf{Y}_{\pi/2}$, Idle).
  6 fiducials. 11 germs. L = 1,2,4,...,8192.
  4657 sequences. 50 counts/each. **230,000 clicks.**

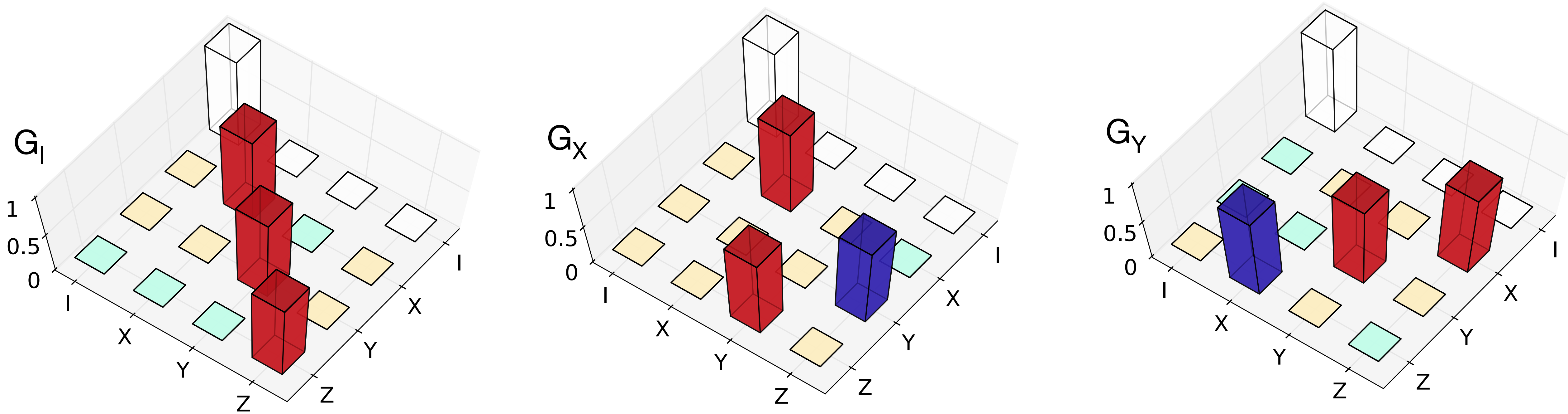
$$\widehat{E} = \begin{bmatrix} 0 & 0 \\ 0 & 0.9929 \end{bmatrix} \pm \begin{bmatrix} 0.6 & 2 \\ 2 & 0.6 \end{bmatrix} \times 10^{-3}$$

$$\widehat{\rho} = \begin{bmatrix} 0.9957 & (2+4i)\times 10^{-3} \\ (2-4i)\times 10^{-3} & 4.3\times 10^{-3} \end{bmatrix} \pm \begin{bmatrix} 0.4 & 2 \\ 2 & 0.4 \end{bmatrix} \times 10^{-3}$$

$$P_{\text{spam}} = (4.2 \pm 0.5) \times 10^{-3}$$

We tomographed a qubit really really really precisely.

- Trapped-ion (Yb^+) qubit in Peter Maunz's lab (March 2015)



$$\mathbb{1} \qquad X_{\pi/2} \qquad Y_{\pi/2}$$

We tomographed a qubit really really really precisely.

- Trapped-ion (Yb^+) qubit in Peter Maunz's lab (March 2015)

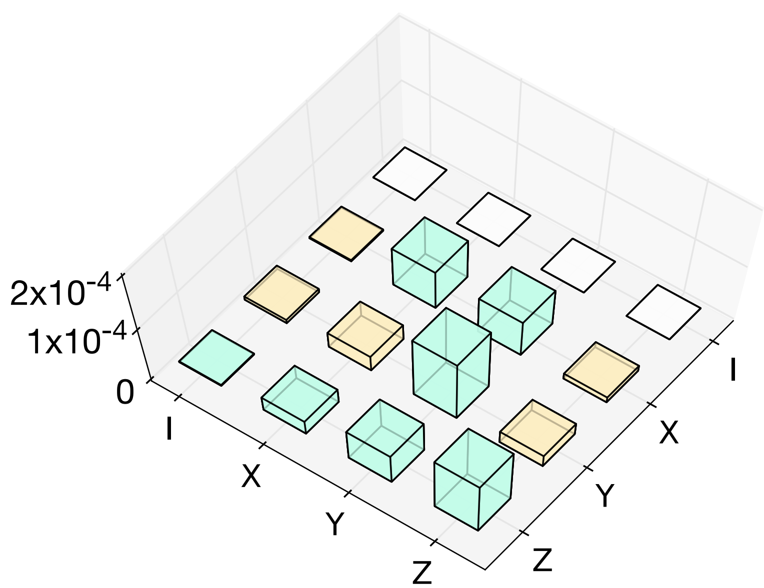
$$\widehat{G}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.999932 & -6\times10^{-5} & 1\times10^{-5} \\ 6\times10^{-6} & 3\times10^{-5} & 0.999891 & 2\times10^{-5} \\ 0 & -3\times10^{-5} & -6\times10^{-5} & 0.999900 \end{bmatrix} \pm \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.4 & 0.7 & 1.3 & 1.2 \\ 0.4 & 1.3 & 0.9 & 1.3 \\ 0.5 & 1.2 & 1.3 & 0.9 \end{bmatrix} \times 10^{-5}$$

$$\widehat{G}_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.999946 & 5\times10^{-5} & 0 \\ 0 & 2\times10^{-5} & 5\times10^{-5} & -0.999904 \\ 0 & 4\times10^{-5} & 0.999904 & 6\times10^{-5} \end{bmatrix} \pm \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.4 & 0.7 & 1.6 & 1.5 \\ 0.6 & 1.4 & 1.7 & 0.7 \\ 0.6 & 1.5 & 0.7 & 1.7 \end{bmatrix} \times 10^{-5}$$

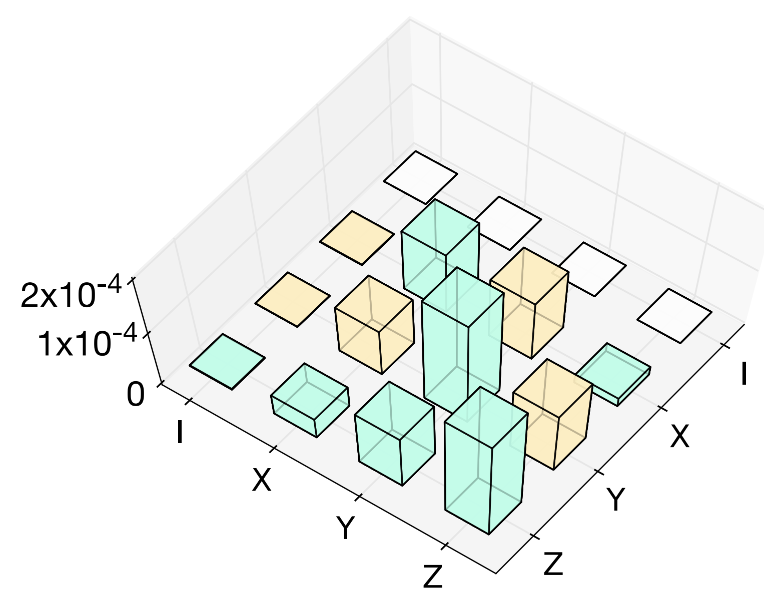
$$\widehat{G}_y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2\times10^{-5} & 2\times10^{-5} & 0.999876 \\ 0 & -2\times10^{-5} & -0.999962 & 0 \\ 0 & 0.999876 & 3\times10^{-5} & -5\times10^{-5} \end{bmatrix} \pm \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.6 & 1.6 & 1.5 & 0.7 \\ 0.4 & 1.5 & 0.6 & 1.5 \\ 0.5 & 0.7 & 1.6 & 1.7 \end{bmatrix} \times 10^{-5}$$

We tomographed a qubit really really really precisely.

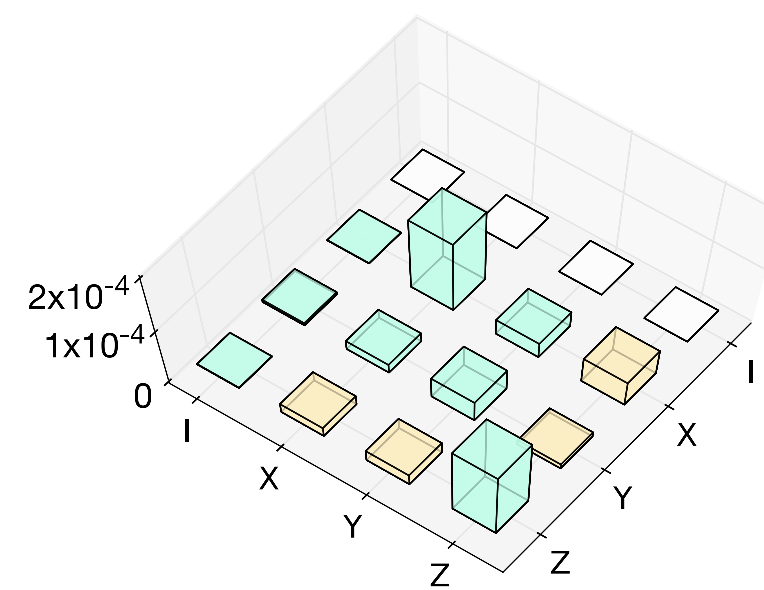
- Trapped-ion (Yb^+) qubit in Peter Maunz's lab (March 2015)
- Shown: *error generators*: $G = G_{\text{ideal}} \circ e^{-\mathcal{L}}$



$\mathbb{1}$



$X_{\pi/2}$



$Y_{\pi/2}$

We tomographed a qubit really really really precisely.

- Trapped-ion (Yb^+) qubit in Peter Maunz's lab (March 2015)

$$\widehat{G}_i = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.999932 & -6\times10^{-5} & 1\times10^{-5} \\ 6\times10^{-6} & 3\times10^{-5} & 0.999891 & 2\times10^{-5} \\ 0 & -3\times10^{-5} & -6\times10^{-5} & 0.999900 \end{bmatrix}$$

$$1 - F(\widehat{G}, G_{\text{ideal}}) = (6.9 \pm 0.3)\times 10^{-5}$$
$$\|\widehat{G} - G_{\text{ideal}}\|_\diamond = (7.9 \pm 0.9)\times 10^{-5}$$

$$\widehat{G}_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.999946 & 5\times10^{-5} & 0 \\ 0 & 2\times10^{-5} & 5\times10^{-5} & -0.999904 \\ 0 & 4\times10^{-5} & 0.999904 & 6\times10^{-5} \end{bmatrix}$$

$$1 - F(\widehat{G}, G_{\text{ideal}}) = (6.1 \pm 0.4)\times 10^{-5}$$
$$\|\widehat{G} - G_{\text{ideal}}\|_\diamond = (7.0 \pm 1.3)\times 10^{-5}$$

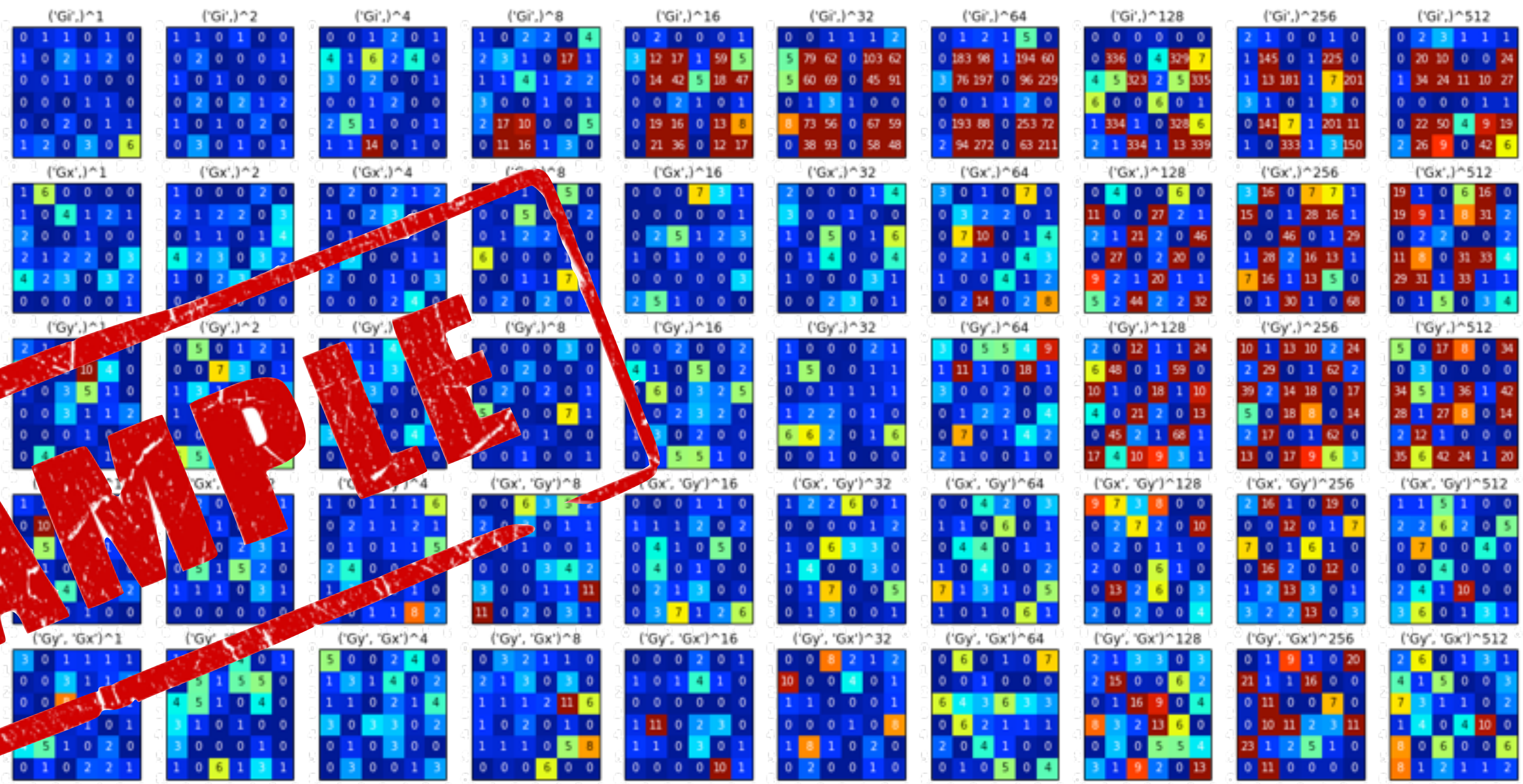
$$\widehat{G}_y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2\times10^{-5} & 2\times10^{-5} & 0.999876 \\ 0 & -2\times10^{-5} & -0.999962 & 0 \\ 0 & 0.999876 & 3\times10^{-5} & -5\times10^{-5} \end{bmatrix}$$

$$1 - F(\widehat{G}, G_{\text{ideal}}) = (7.1 \pm 0.4)\times 10^{-5}$$
$$\|\widehat{G} - G_{\text{ideal}}\|_\diamond = (8.1 \pm 1.3)\times 10^{-5}$$

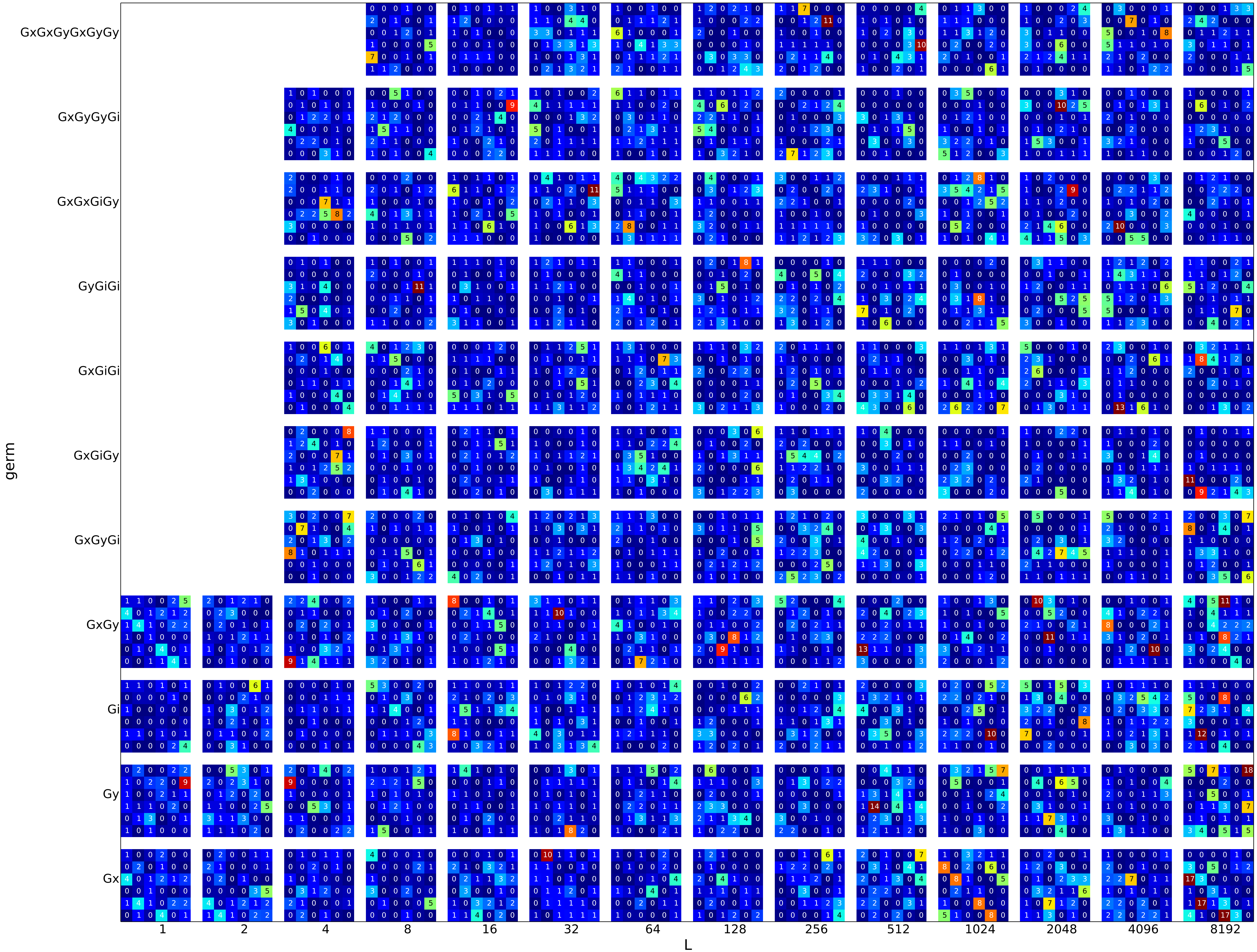
Detecting non-Markovian noise

- Fitting ~30 parameters to ~4500 experiments is *overcomplete.*
  => we have a lot of residual data for *model testing/selection.*

- Badly fit data points = model violation = non-Markovian.



Our gates: *really* Markovian!



GST on two qubits

- We have been developing 2-qubit GST for about a year.
- Challenges:
  - Many parameters in the gateset (~1200)
  - Need to design germs to amplify all those parameters.
  - Many sequences (up to 54,000).
  - Memory requirements for LSGST analysis (10+ GB)
- June 2015: Simulations (it worked).
- June-Dec. 2015: "Biqubit" experiments in Maunz lab
- Feb. 2016: Complete 2-qubit GST in BBN lab

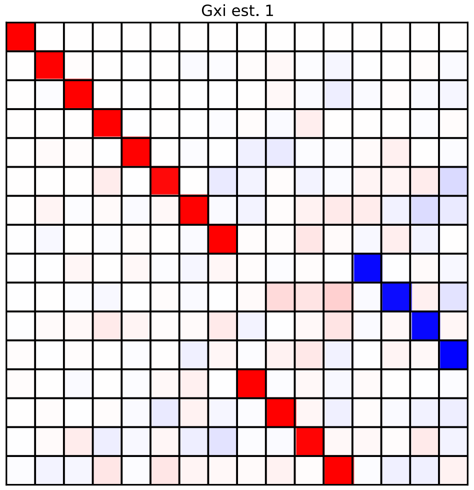
2-Transmon expt: parameters

- Two transmon qubits at BBN* (fabricated by IBM)
- $\boldsymbol{X}_{\pi/2}$ and $\boldsymbol{Y}_{\pi/2}$ gates on each qubit, $e^{-i\boldsymbol{Z}\otimes\boldsymbol{Y}}$ entangling gate.
- 16 state preps x 9 measurements = 144 "fiducial pairs".
- 71 germ sequences, repeated L=1,2,4,8,16 times each.
- 27,017 distinct gate sequences. 2500 shots/sequence
  => 68 million clicks.
- Analysis ran for 14 hours on a laptop.
- **Gates are not optimized!!!**

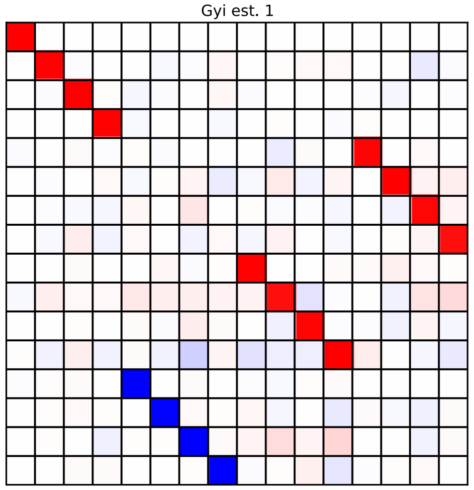
* Thanks to Marcus da Silva, Diego Riste, Matt Ware, Colm Ryan, and the ARO QCVV program (#W911NF-14-C-0048)

Results: Linear GST

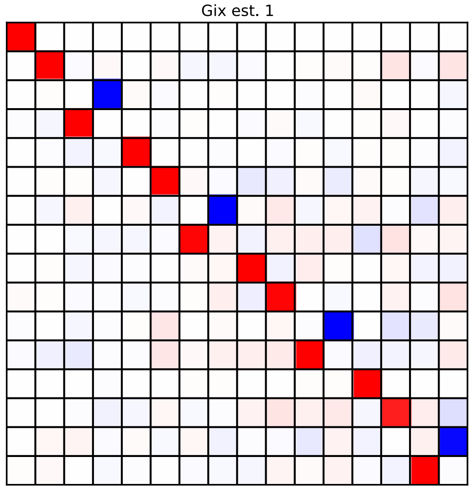
G_{XI} (LGST)



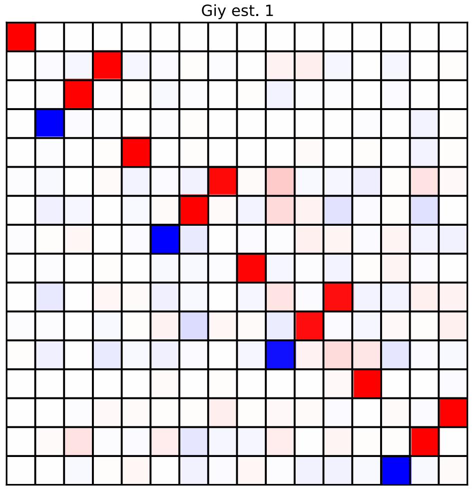
G_{YI} (LGST)



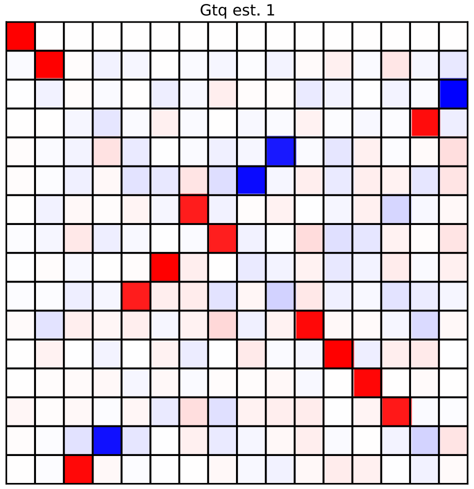
G_{IX} (LGST)



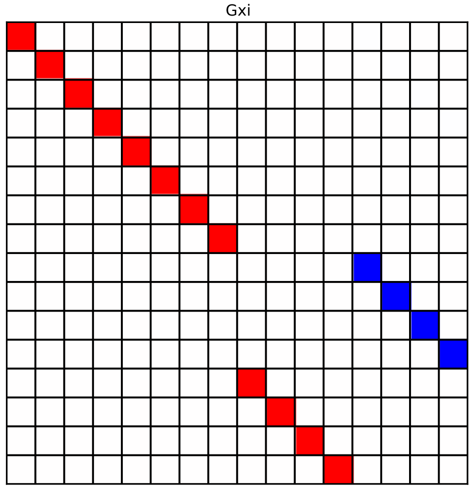
G_{IY} (LGST)



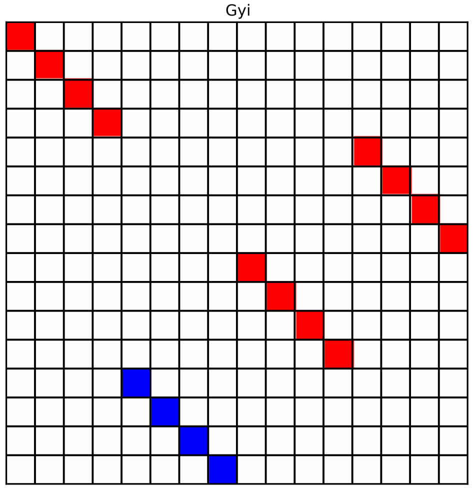
G_{2Q} (LGST)



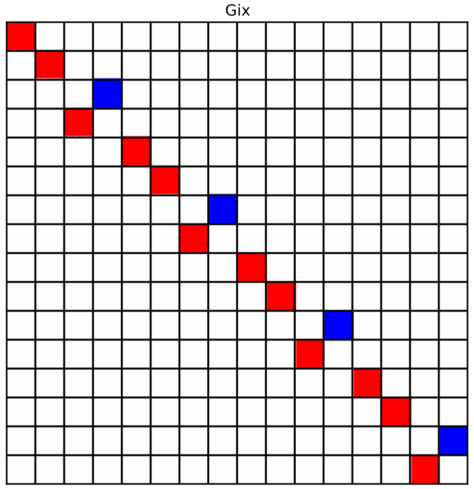
G_{XI} (target)



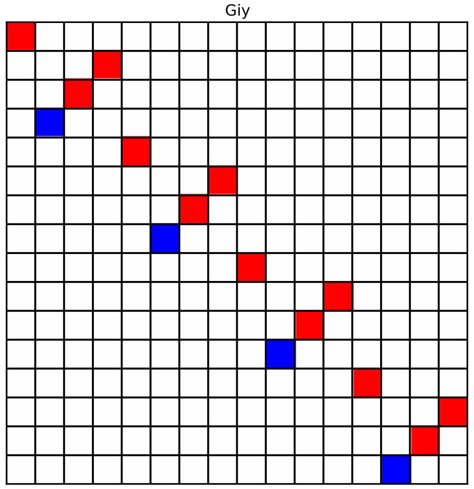
G_{YI} (target)



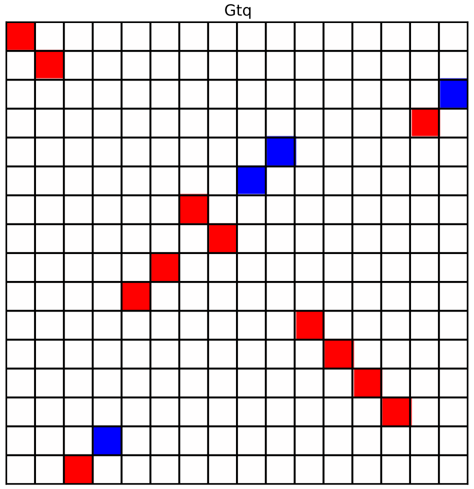
G_{IX} (target)



G_{IY} (target)



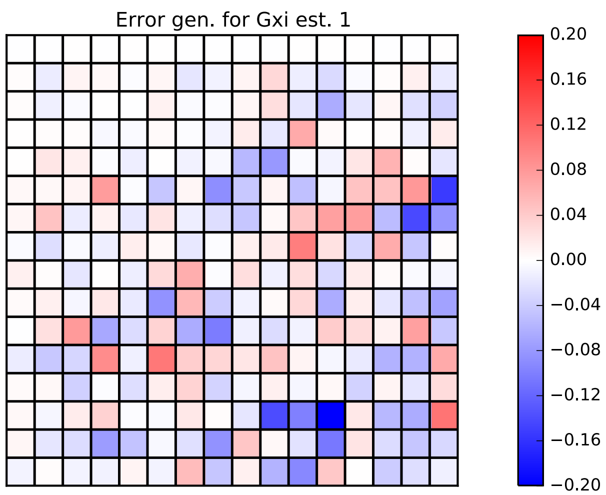
G_{2Q} (target)



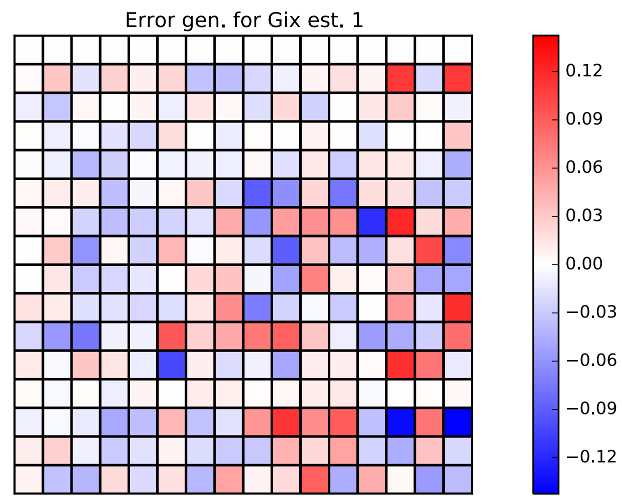
Error generators (Linear GST)

Scale on plots: ± 0.2

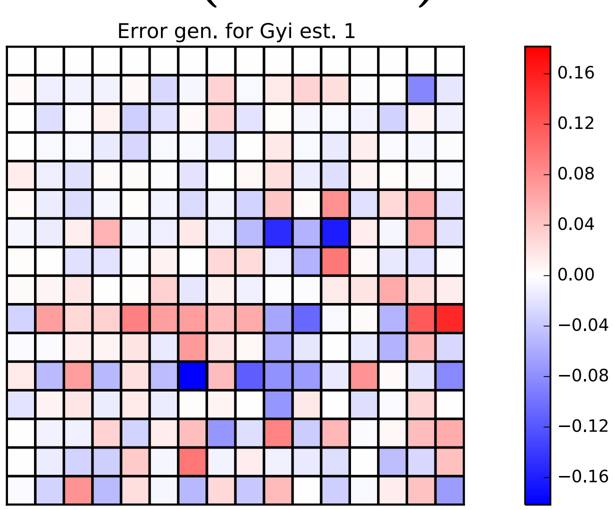
G_{XI} (LGST)



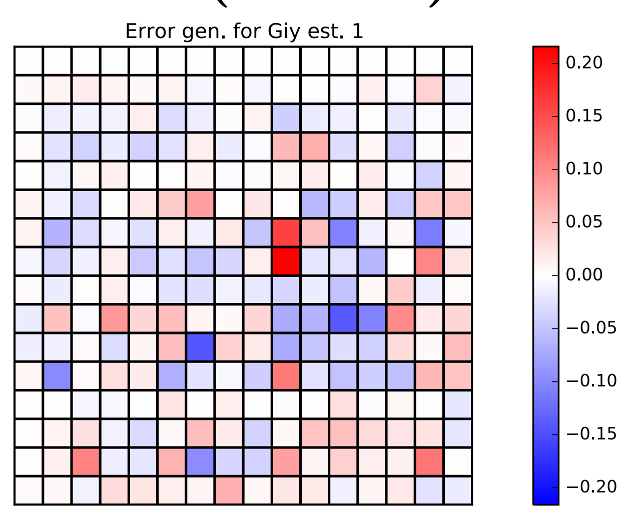
G_{IX} (LGST)



G_{YI} (LGST)

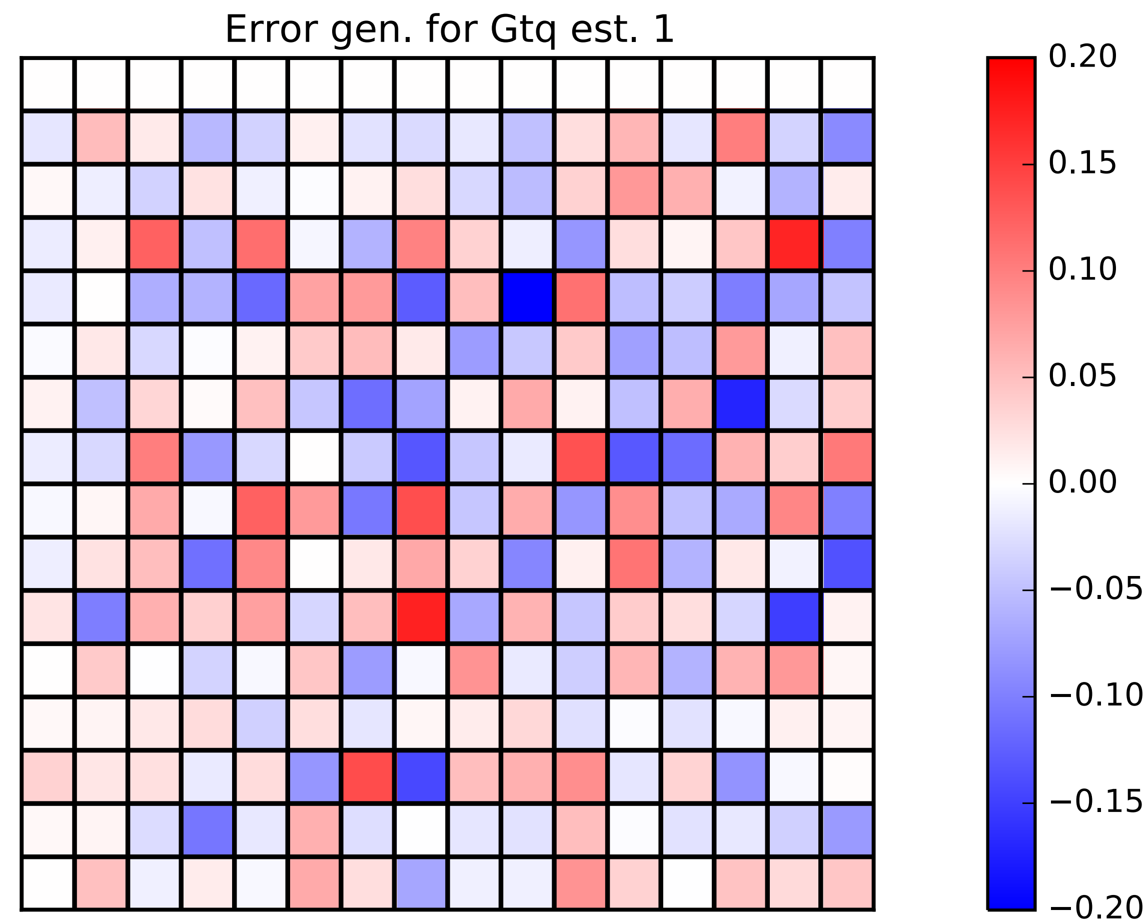


G_{IY} (LGST)



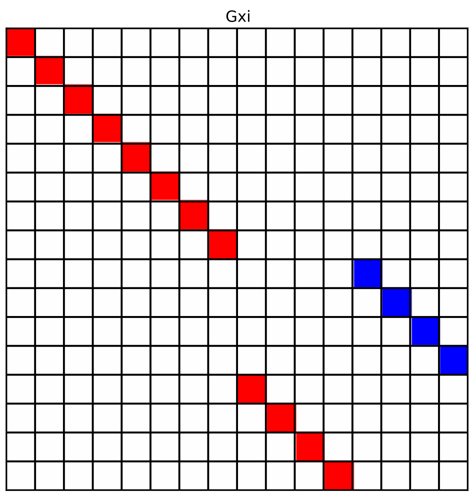
G_{2Q} (LGST)

Scale: ± 0.2

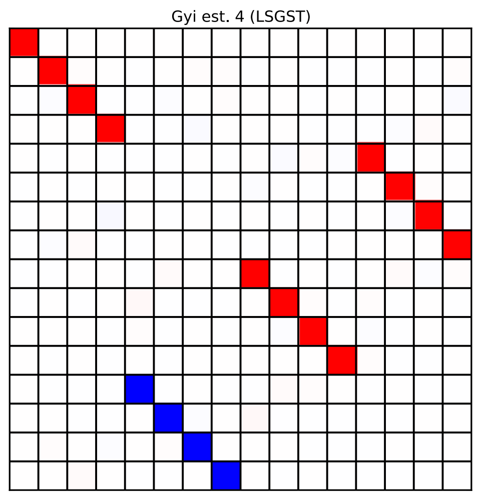


Long sequence (L=16) GST

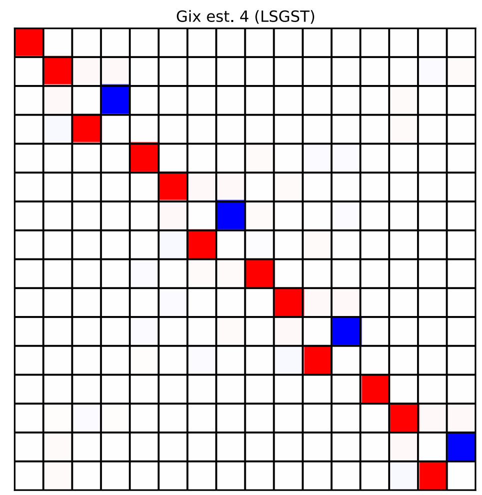
G_{XI} (LGST)



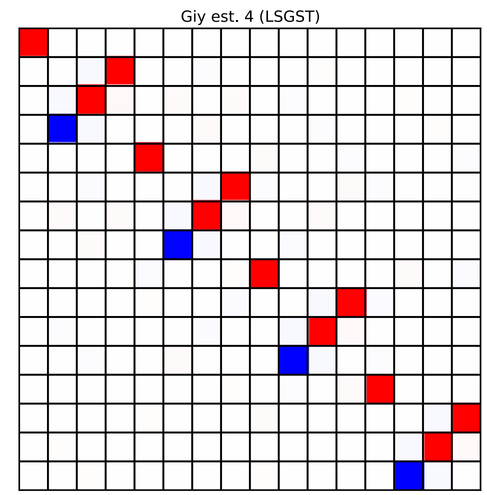
G_{YI} (LGST)



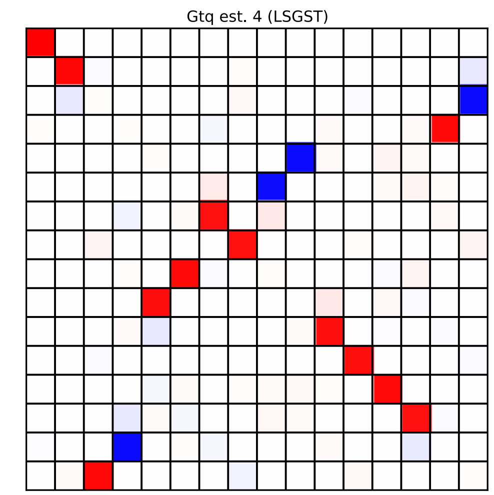
G_{IX} (LGST)



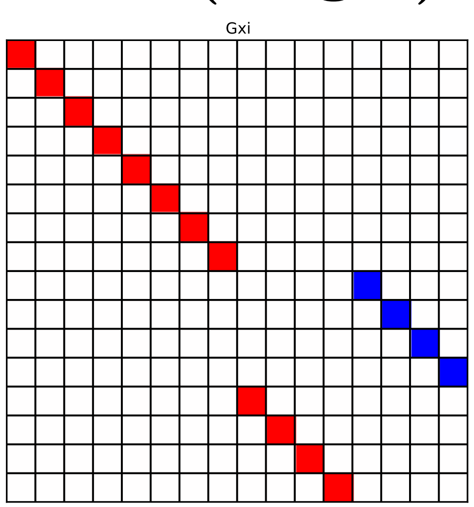
G_{IY} (LGST)



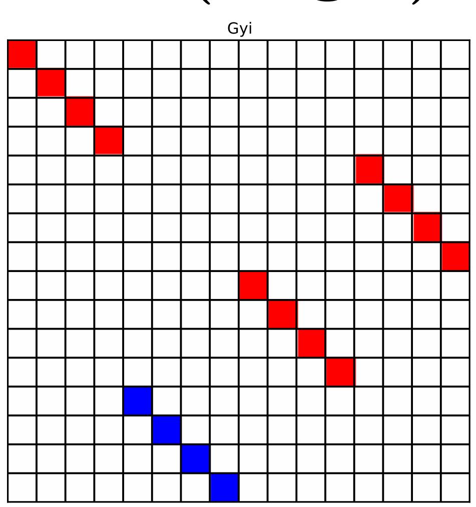
G_{2Q} (LGST)



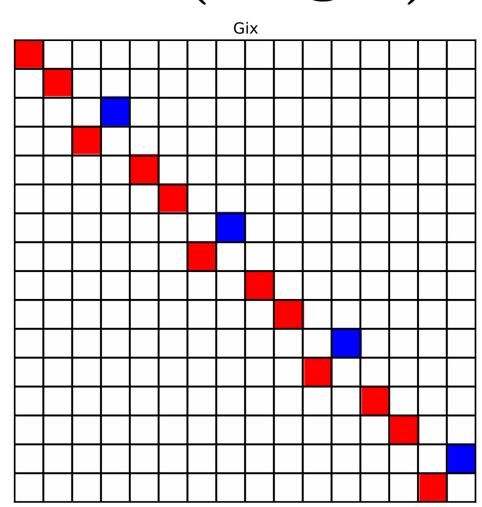
G_{XI} (target)



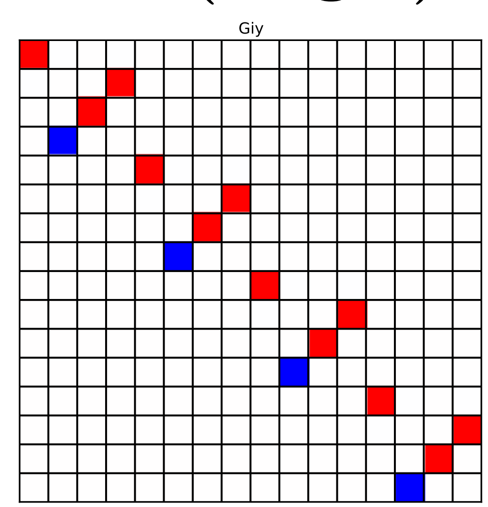
G_{YI} (target)



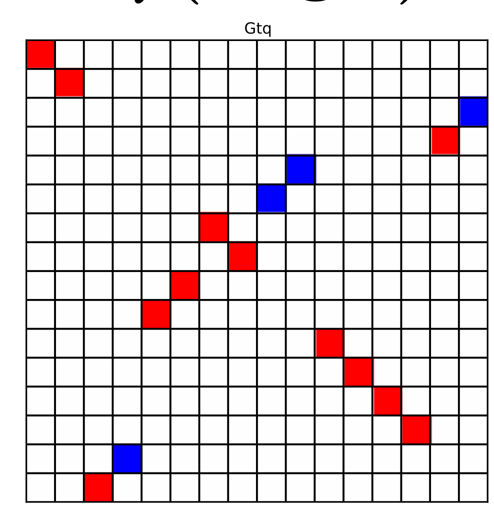
G_{IX} (target)



G_{IY} (target)



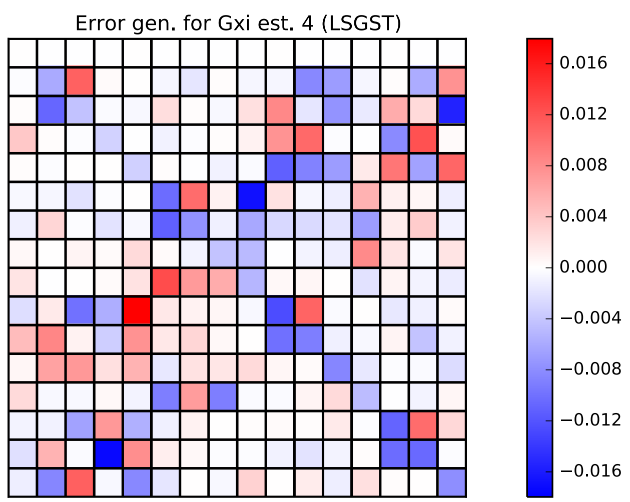
G_{2Q} (target)



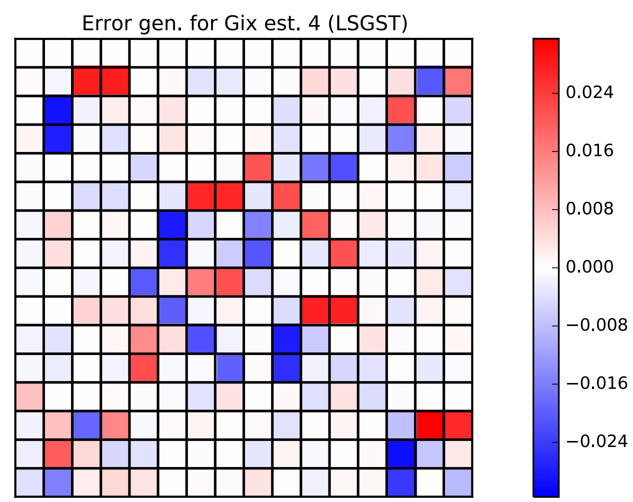
Error generators (long-seq GST)

Scale on plots: ±0.024

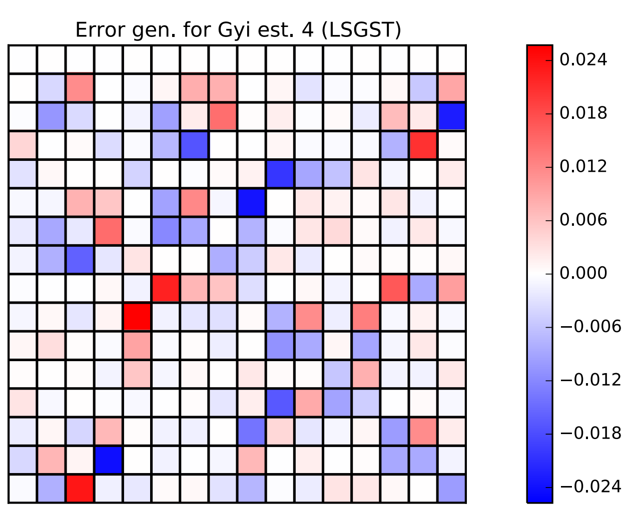
G_{XI}



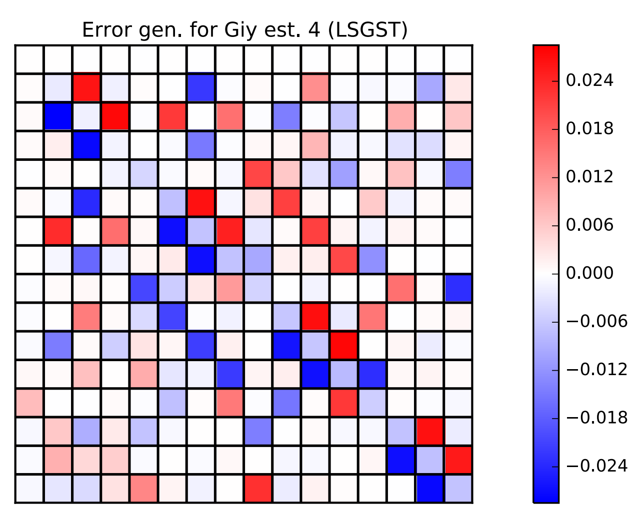
G_{IX}



G_{YI}

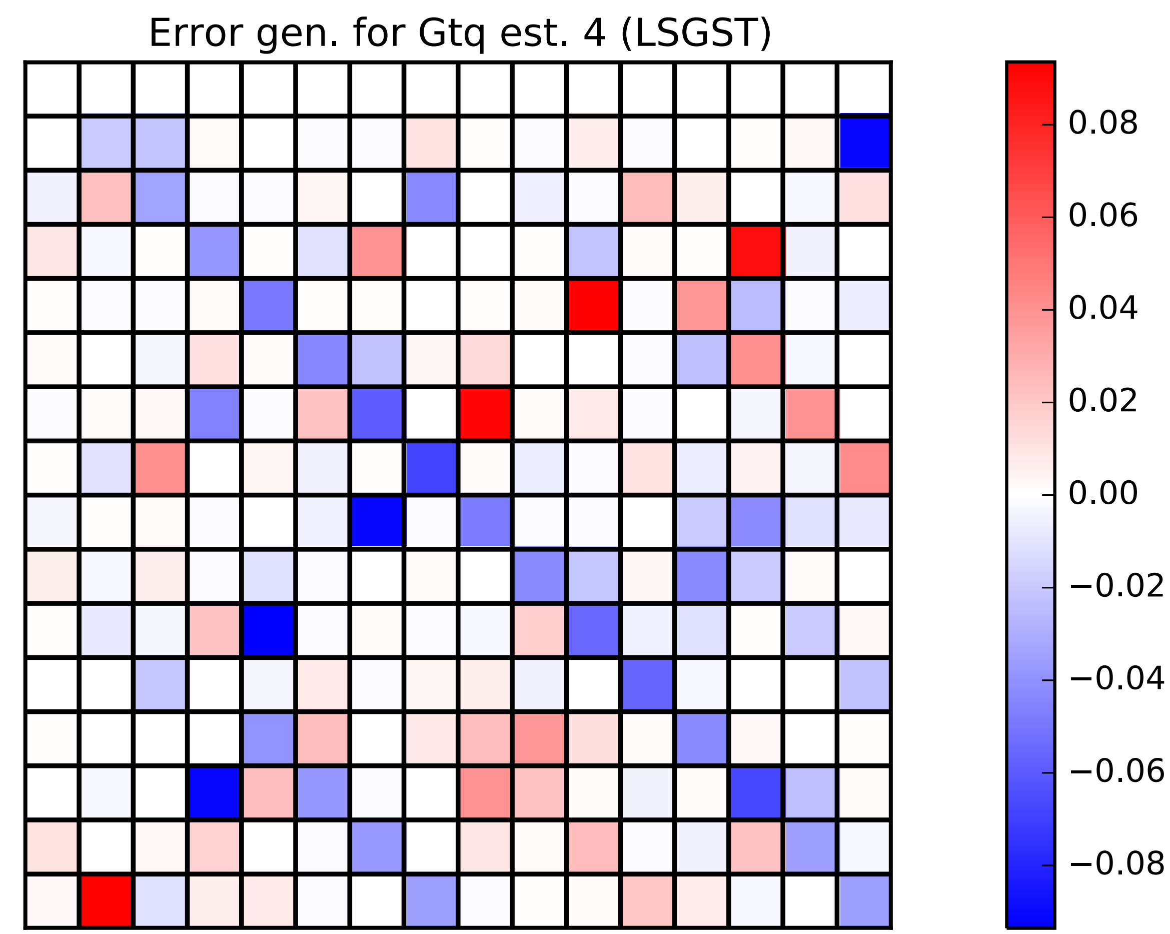


G_{IY}



G_{2Q}

Scale: ±0.08



Part 3: ...and beyond

The next 5 years

- IARPA's "LogiQ" project

- Logical qubit
- 17+ physical qubits (working together)
- Repeated error correction
- Breakeven (logical $\boldsymbol{P}_{\text{fail}} < 0.1\%$)

- **This is getting real!**

The role of tomography

- Certify uniformly low error rates.

- Generate predictive models for circuit simulation.

- Provide rapid debugging feedback for experimentalists
  => Bring $\Diamond$-norm down to infidelity by removing coherent errors.

- Motivate bespoke QEC codes for asymmetric errors.

But we need MORE!

- Great 1- and 2-qubit gates are just the beginning.

Fun problems that terrify me

- Crosstalk (do operations on A affect B)
- Independence (do parallel ops on A & B commute?)
- Correlated errors (*very* low probabilities relevant)
- Non-Markovian noise (important? how to characterize it?)
- Effects of non-logiqal operations (shuttling, cooling, etc).
- Measuring logical error rates
  - very low probabilities => long experiments!
  - can't do RB/GST without full (Clifford) gate set.

Lessons we have learned

- If it's important, you can probably measure it.
- Don't look under the lamppost. Measure important things.
- Calibration-free and gauge-invariant is good.
- Run circuits to predict circuits.
- Use ~~assumptions~~ models. Test/validate them!
- Expect, admit, and quantify violation of models.
- Ground "estimates" in raw data if possible. Verify!
- *Thank the audience for their attention.*