Model Calibration and Circuit Optimization for Quantum Hardware

Dissertation Defense by Stefan Krastanov

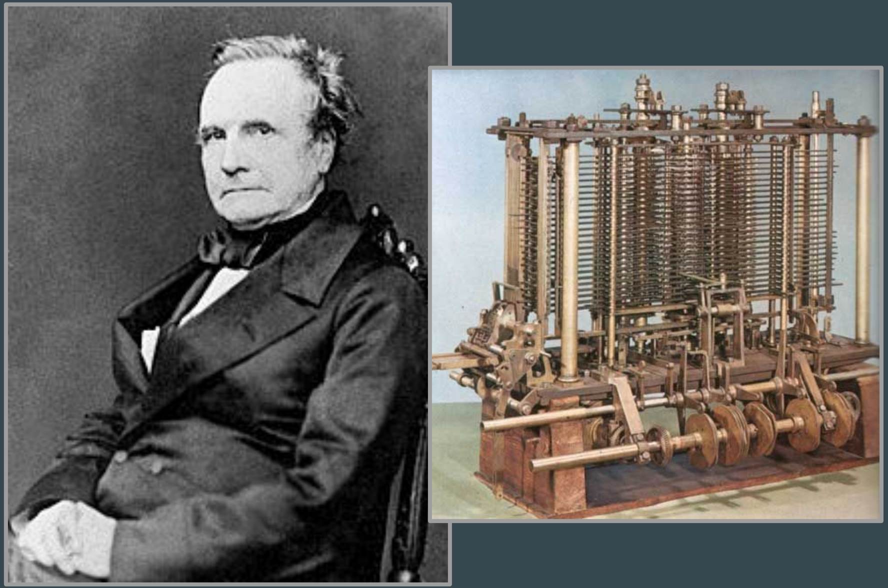
Doctoral Advisor Prof. Liang Jiang

Yale Quantum Institute

The work done during my doctoral studies would not have been possible without valuable input from many colleagues, including Chang-Ling Zou, Chao Shen, Kyungjoo Noh, Michel Devoret, Philip Reinhold, Reinier Heeres, Robert Schoelkopf, Sisi Zhou, Steven Flammia, Steven Girvin, and Victor Albert.

Brief (Opinionated) History of Computing Machinery

A Roman Abacus and an Inca Quipo;
abacus in use since 2700 BCE

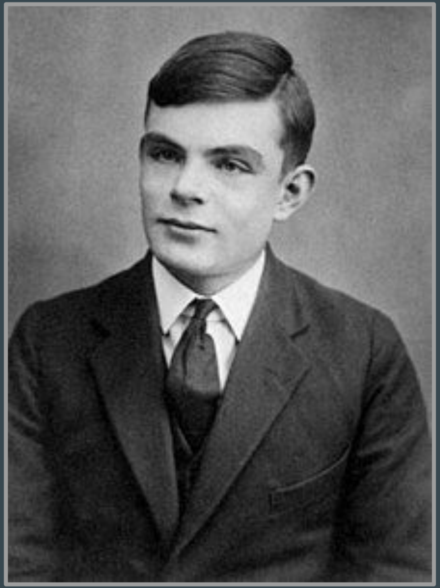


Babbage’s Analytical Engine, late 1800s

The Notion of Universality*

"[...] the Analytical Engine weaves algebraical patterns just as the Jacquard loom weaves flowers and leaves."

Ada Lovelace
late 1800s



“The idea behind digital computers [... is] to carry out any operations which could be done by a human computer.”

Alan Turing
mid 1900s

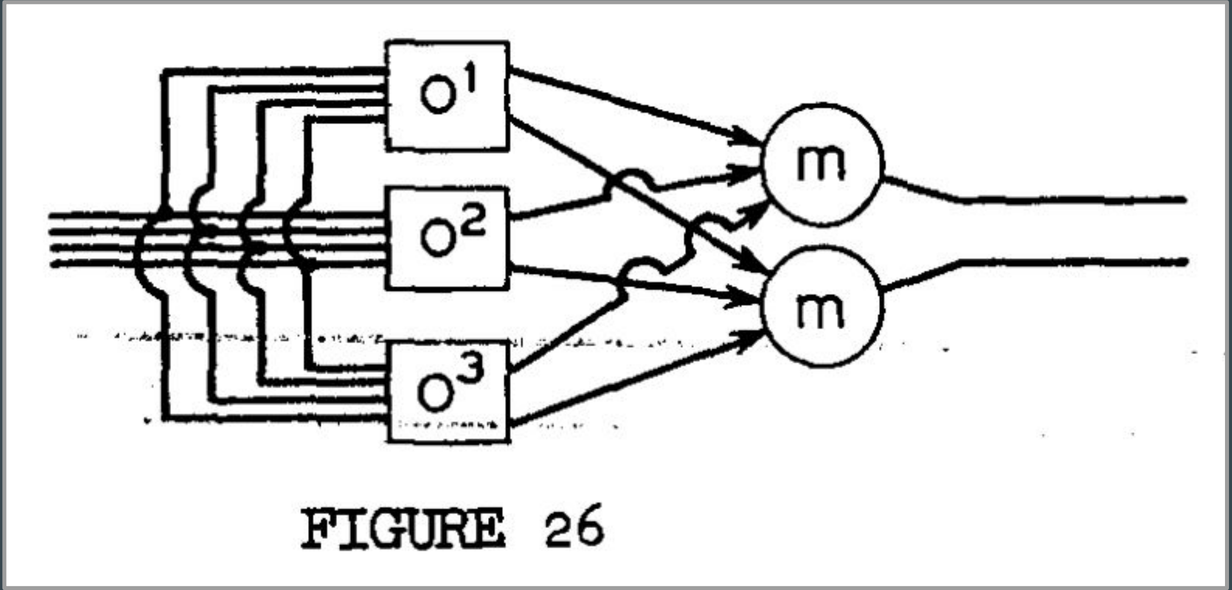
*Universality: the notion that a computer is a general purpose problem solver that can simulate any other computer / problem solver.

Ideal Machines Can Be Difficult to Build in The Real World

Babbage never completed the Analytical Engine.

Many doubted that cogs or relays or vacuum tubes could ever be sufficiently noiseless.

Thankfully, von Neumann proved a “threshold theorem” providing a noise reducing “gadget”.

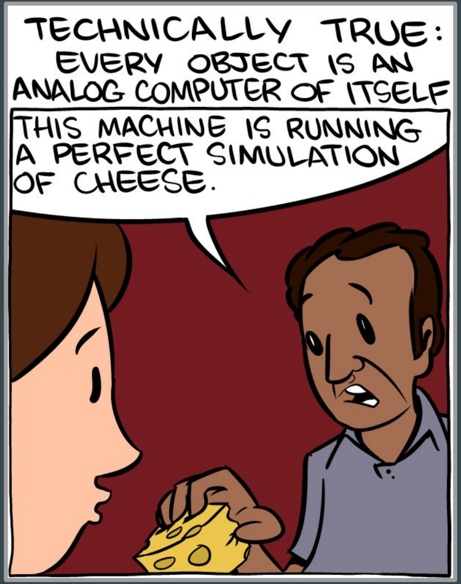


A triply-redundant implementation of an otherwise unreliable operation.
“[…] The Synthesis of Reliable Organisms […]”, von Neumann, 1952

How Much Computational Power Do the Laws of the Physical* Universe Ultimately Provide?

Some problems are difficult, but the universe still solves them.

Quantum mechanics, in particular, is difficult to simulate.



smbc-comics.com, 2013

*Skipping scalable analog computers and hypercomputation as unphysical.

Quantum Computers can be universal! And they outperform* classical ones!

Computable and Noncomputable, Yuri Manin, 1980
Simulating Physics with Computers, Richard Feynman, 1982
Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer, David Deutsch, 1985
Universal Quantum Simulators, Seth Lloyd, 1996

*Probably...

However, Imperfect Quantum Hardware is Hard to Control
even more so than Babbage’s cogs and cams

In that spirit, an Outline

- Control and Calibration
- Purifying Resources with Imperfect Circuits

Control and Calibration

arxiv:1812.05120 / QST 2019

STEADY

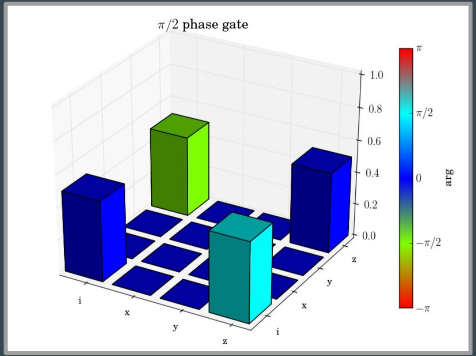
STochastic Estimation Algorithm for DYnamical variables

How Control and Calibration Work
naive interpretation of a theorist

Measure the model parameters governing the hardware.

Devise some control protocol, informed by the model.

e.g. SNAP based control
or the OCT work by Reinier and Philip
or Philip and Wenlong’s error resistant control



Perform measurements validating the quality of the control protocol.

e.g. Process Tomography or Randomized Benchmarking

Survey of Methods: Parameter Estimation/Evaluation

let us start with a simplified version of Process Tomography

A Quantum Process

$$\mathcal{E}(\rho_{\text{input}}) = \rho_{\text{output}}$$

is a linear map (super-operator)

$$\mathcal{E}(\rho_i) = \sum_j \mathcal{E}_{ij}\rho_j$$

i.e. it can be represented as a matrix $\mathcal{E}_{ij}$ where each row/column corresponds to a basis vector for the output/input.

We can use Process Tomography to measure all coefficients $\mathcal{E}_{ij}$:

1. Prepare ρ_i

2. Apply the quantum process to get $\mathcal{E}(\rho_i)$

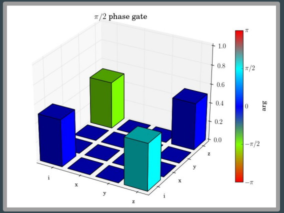
3. Measure the overlap* with ρ_j, which gives you $\mathcal{E}_{ij}$

Susceptible to State Preparation and Measurement Errors

* There are some questions of orthogonality to address.

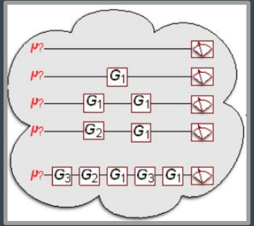
Survey of Methods: Parameter Estimation/Control/Evaluation
(biased and incomplete)

Process Tomography



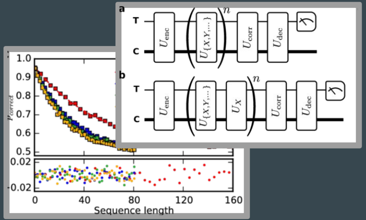
1. Inefficient dense representation
2. Expensive
3. SPAM*-sensitive

Gate Set Tomography



1. Inefficient dense representation
2. Expensive
3. Immune to most SPAM*

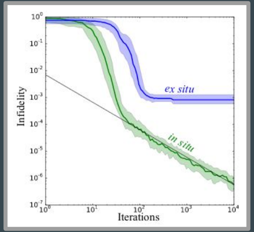
Randomized Benchmarking



1. Only overall “scalar” fidelity
2. Immune to SPAM*

Only verification, not informative on how to improve control.

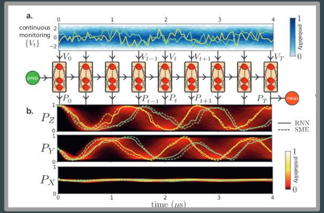
Stochastic in situ Methods



1. Very inefficient
2. Model independent
3. Immune to SPAM*
4. Very high quality

Phys. Rev. A 91, 052306

Recurrent Neural Nets



1. Inscrutable black-box model
2. Can be sparse
3. Very general

arXiv:1811.12420

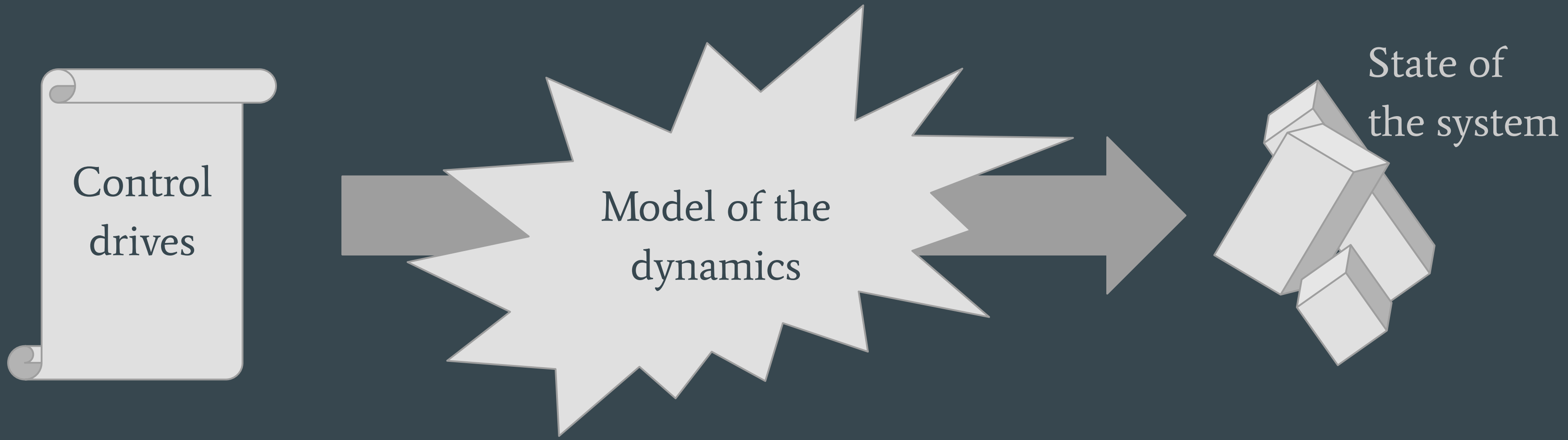
*SPAM: State Preparation and Measurement errors

We want it all!

Immune to SPAM*
Sparse Efficient Representation
Continuous (for use in optimal control)
Saturating Information Theory Bounds
<u>General</u> while still <u>Interpretable</u>

*State Preparation and Measurement errors

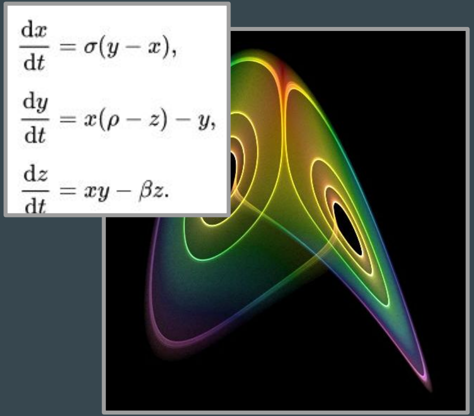
What is a Model of Controllable Hardware?



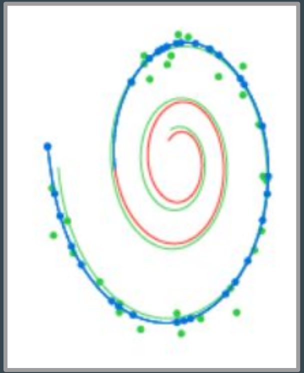
$$\boldsymbol{d}(t) \qquad \boldsymbol{d} \to H\left(\boldsymbol{d}\right) \qquad |\dot{\psi}(t)\rangle = -iH\left(\boldsymbol{d}(t)\right)|\psi(t)\rangle$$

Spectrum of Models

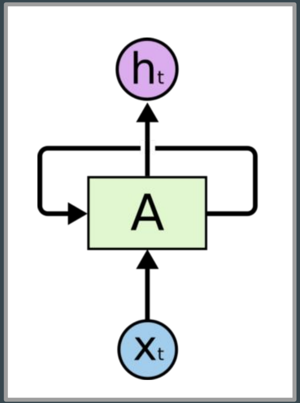
Differential Equations



Neural ODEs



Recurrent Neural Networks



Neural Networks

Sparse, Understandable
Susceptible to model errors

Dense, Inscrutable
What “model error”?
(Just get me more training data)

Model Hamiltonian

$$\boxed{\tilde{H}}(\boxed{\omega}; \boxed{d})$$

Model Hamiltonian Parameters Control Drives

$$\boxed{H}(d) = \tilde{H}(\boxed{\omega_0}; d)$$

"True" Hamiltonian "True" Parameters

Model Hamiltonian
A Particular Parameterization

$$\tilde{H}(\boldsymbol{\omega}; \boldsymbol{d})$$

Model Hamiltonian (for Q qubits)

Parameters

Control Drives (D-dimensional vector)

$$\tilde{H}(\boldsymbol{\alpha}, \boldsymbol{\beta}; \boldsymbol{d}) = \sum_{k=1}^{M} a_k A_k,$$

Predetermined permitted operators (M different operators, not spanning the entire space of operators)

$$\text{where } a_k = \sum_{l=1}^{D} \alpha_{kl} d_l + \beta_k.$$

$M \times (D+1)$ real parameters to be estimated

Model Hamiltonian
General Linear Control Drives

$$\tilde{H}(\boldsymbol{\omega};\boldsymbol{d})$$

Model Hamiltonian Parameters Control Drives

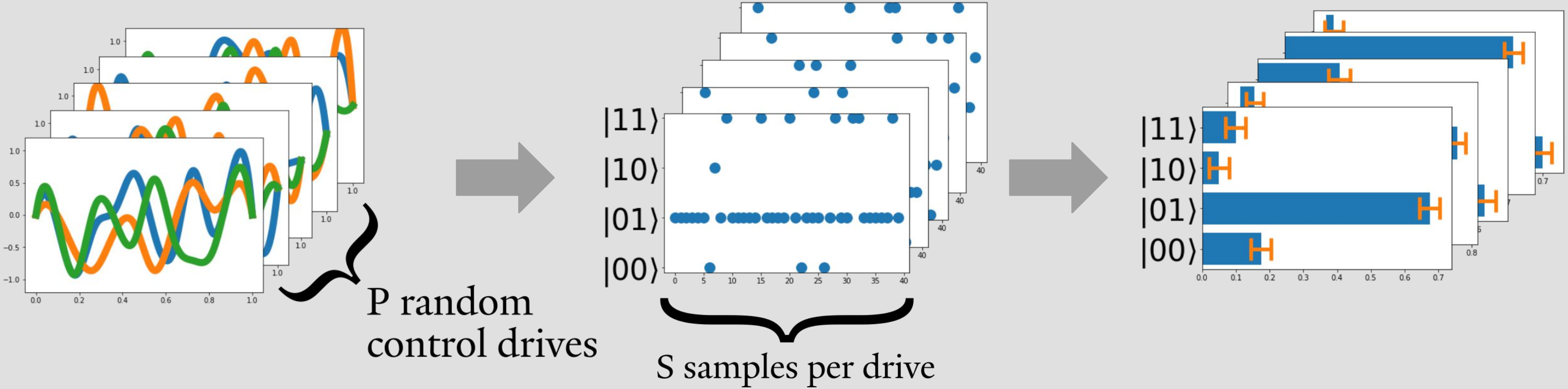
$$\tilde{H}_{ij}(\boldsymbol{\sigma},\boldsymbol{h};\boldsymbol{d}) = h_{ij} + \sum_{k=1}^{D} \sigma_{ijk} d_k$$

$2^{2Q} \times (D+1)$ real parameters to be estimated

Calibration

Run a bunch of random control drives.
Compare the result to the model prediction.
Fudge the model parameters until they match.

Run a Bunch of Control Drives and Repeatedly Sample



1. Prepare random pulses

2. Sample the result of each pulse repeatedly.

3. Estimate populations in each basis state

Compare the Result to Prediction

Fundamentally, you can not *exactly* measure a quantum state.
You can *estimate* its population in a given basis
with *repeated* projective measurements.

(or a bit more generally—and more expensively—do tomography on the state)

$|\psi_0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

run control drive

$e^{-iHt}|\psi_0\rangle = \begin{pmatrix} i\sqrt{2/3} \\ -\sqrt{1/3} \end{pmatrix}$

phases are lost in projective measurement

$\boldsymbol{p_{Born}} = \{|\langle k|e^{-iHt}|\psi_0\rangle|^2, |k\rangle \text{ spans basis}\} = \begin{pmatrix} 2/3 \\ 1/3 \end{pmatrix}$

sample from this distribution

0 1 0 0

the estimate after S samples $\boldsymbol{p^S} = \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix} \pm \sqrt{\frac{1}{S}}$

Finally, compare the estimate of the truth to the prediction:

$$\boldsymbol{\tilde{p}_{Born}} = \{|\langle k|e^{-i\tilde{H}t}|\psi_0\rangle|^2\}_k$$

Compare Model and Measurement

Current estimate of the hardware parameters

$$C(\omega) = \frac{1}{P}\sum_{i=1}^{P} \mathrm{dist}\left(p_i^S, \tilde{p}_{iBorn}(\omega)\right)$$

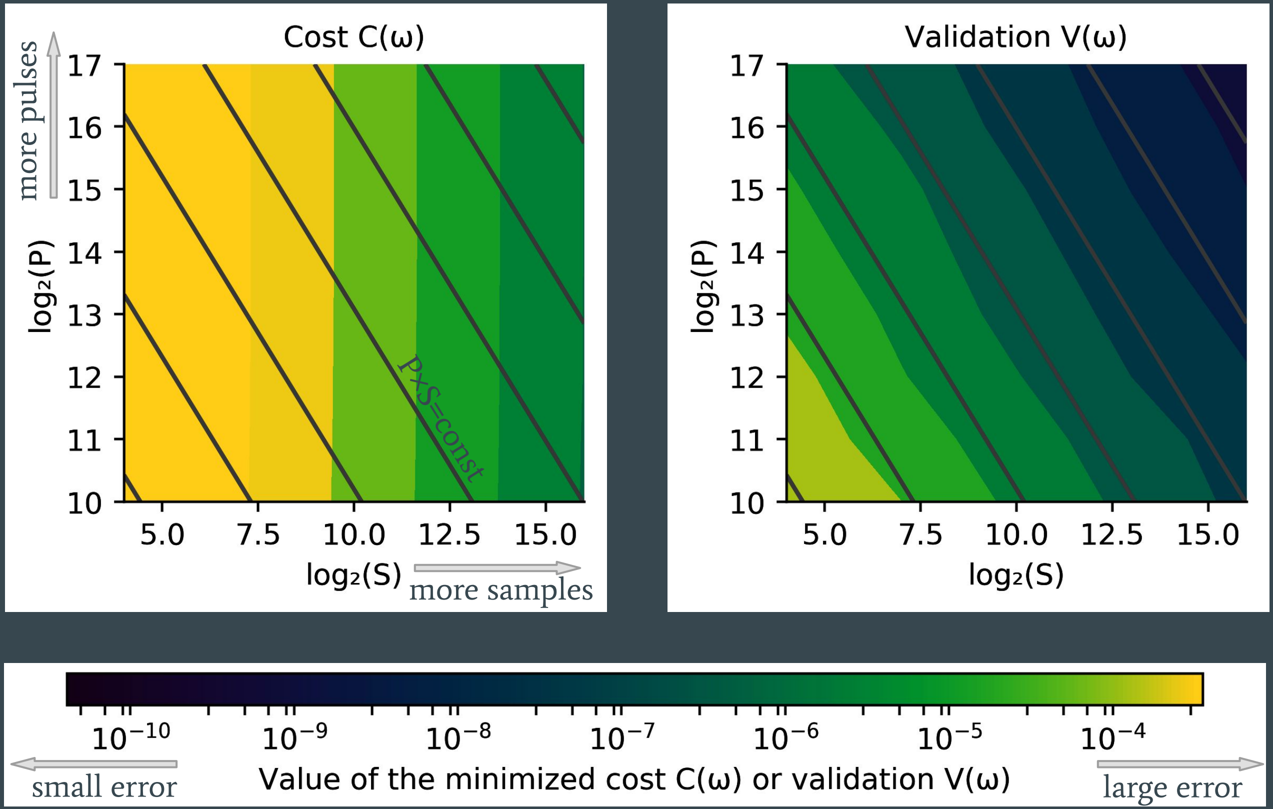
Sum over a bunch of random control pulses

Measured estimate of the actual final state obtained from S repeated samples per pulse

Prediction given the parameters

Fundamentally, you can not *exactly* measure a quantum state.
You can *estimate* its population in a given basis
with *repeated* projective measurements.
(or a bit more generally—and more expensively—do tomography on the state)

P - number of different random pulses
S - number of samples per pulse



$$C(\omega) = \frac{1}{P}\sum_{i=1}^{P} \text{dist}\left(\boldsymbol{p}_i^S, \tilde{\boldsymbol{p}}_{i\,Born}(\omega)\right)$$

$$V(\omega) = \frac{1}{P}\sum_{i=1}^{P} \text{dist}\left(\boldsymbol{p}_{i\,Born}, \tilde{\boldsymbol{p}}_{i\,Born}(\omega)\right)$$

The predicted final state given the parameters

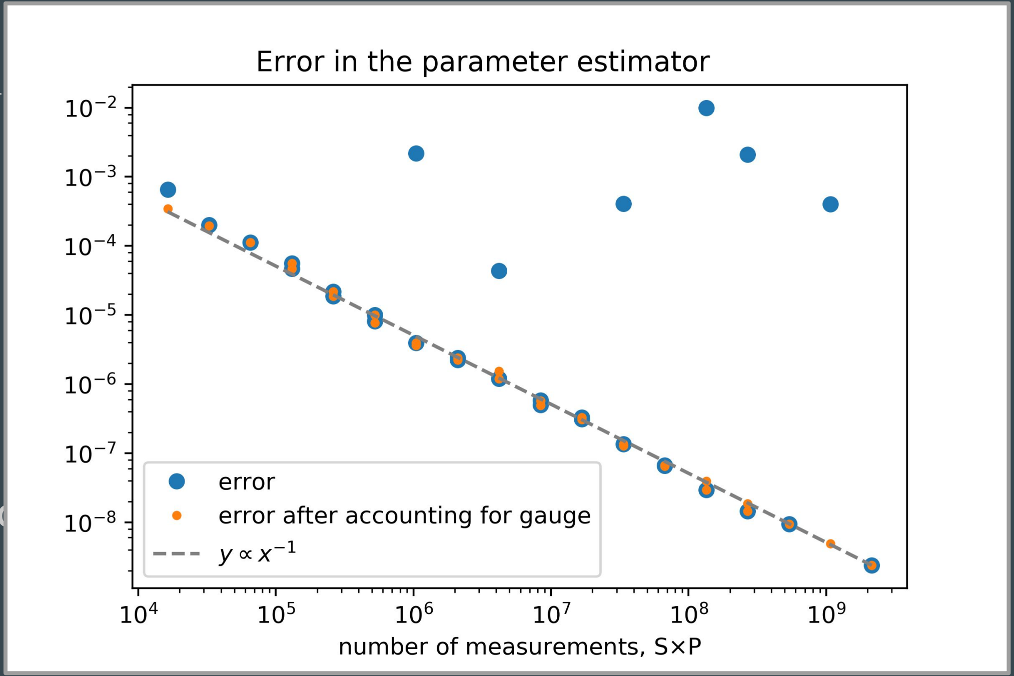
Measured estimate of the actual final state

The actual final state

Optimized Cost (P unique pulses, each ran and sampled S times)

Theoretical Limit of Calibration Fidelity

Fisher Information

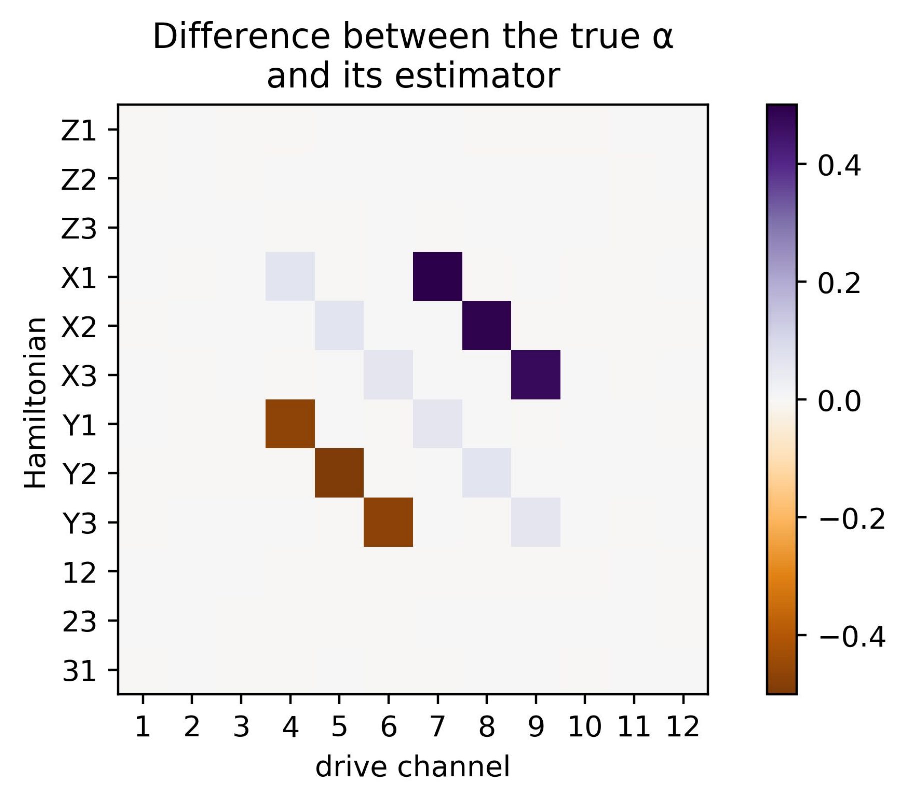


... is additive

Cramér–Rao bound

$$\mathrm{var}(\omega_i) \propto \frac{1}{P \times S}$$

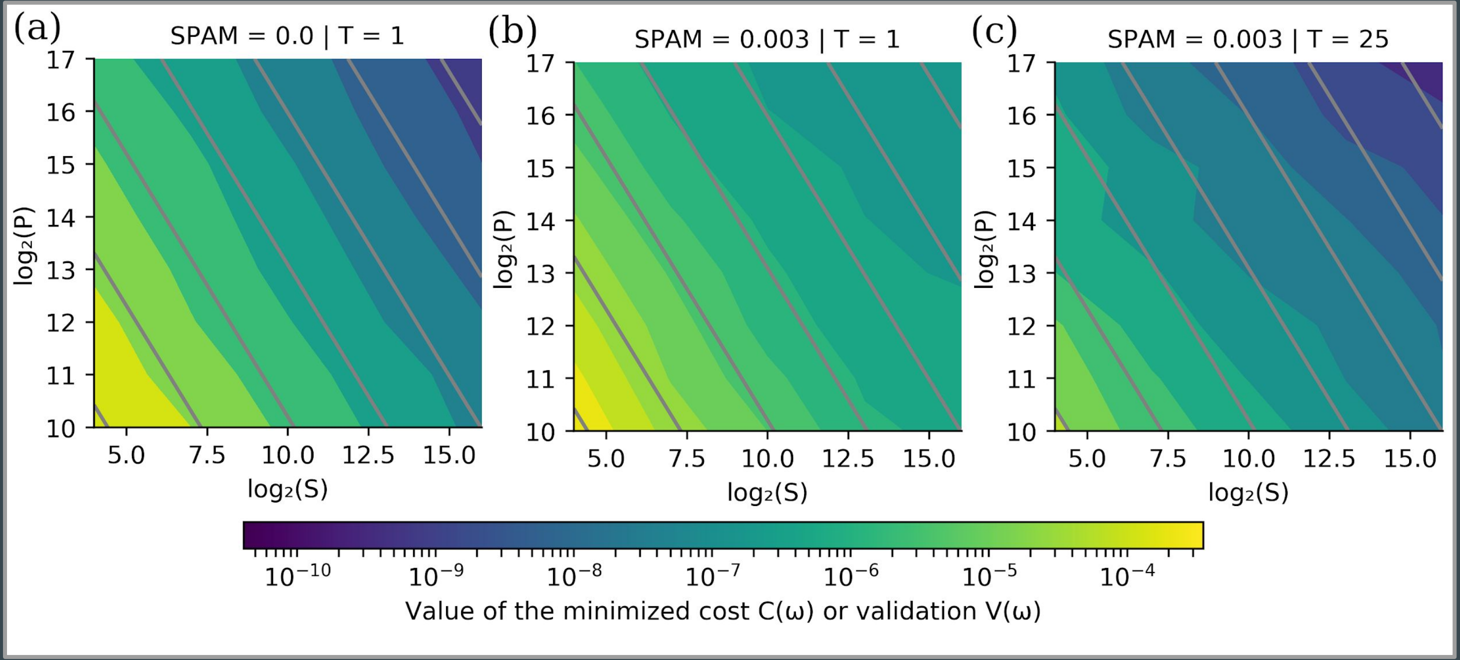
“Gauge” Degrees of Freedom



Permitted Operators (i = 1, 2 or 3):

$$\{\sigma_i^x,\ \sigma_i^y,\ \sigma_i^z,\ \sigma_i^+\sigma_{i+1}^- + \text{h.c.}\}$$

Intrinsic State Preparation and Measurement Errors

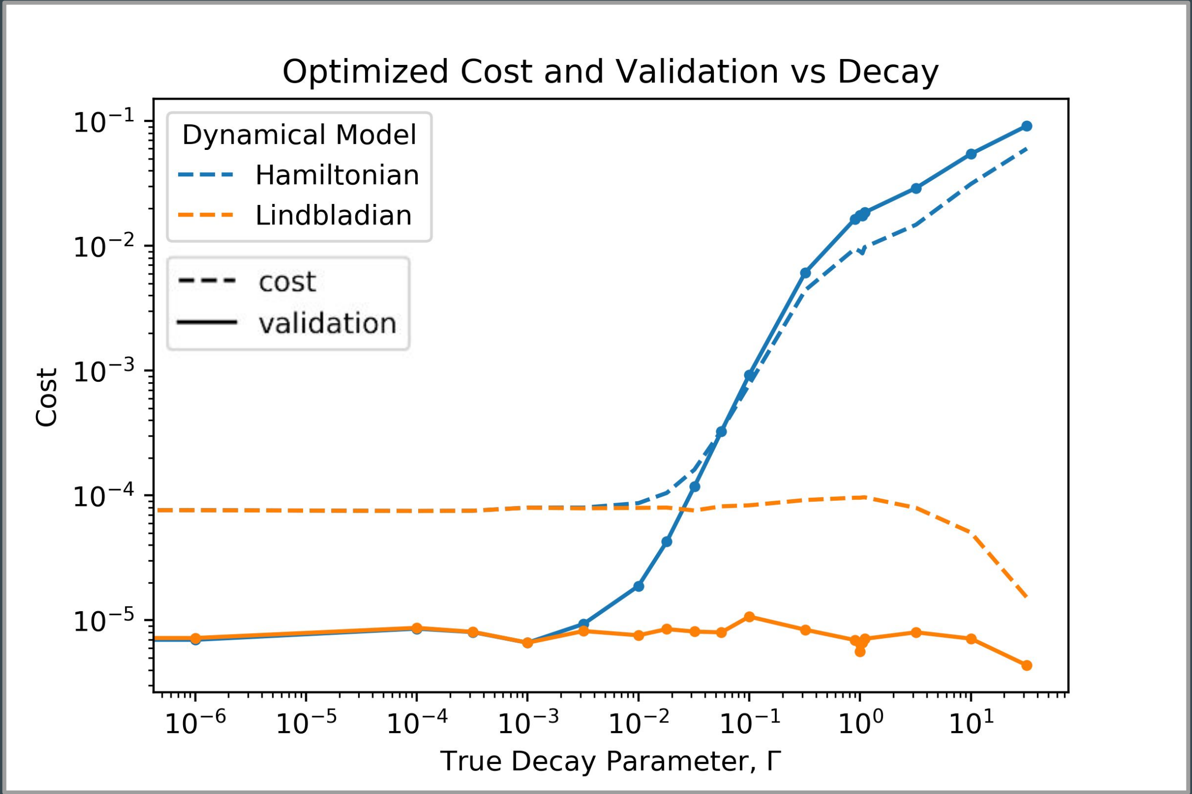


Intrinsic SPAM causes a bias in our estimator!

Non-unitary Evolution

A Hamiltonian model is incapable of capturing decoherence.

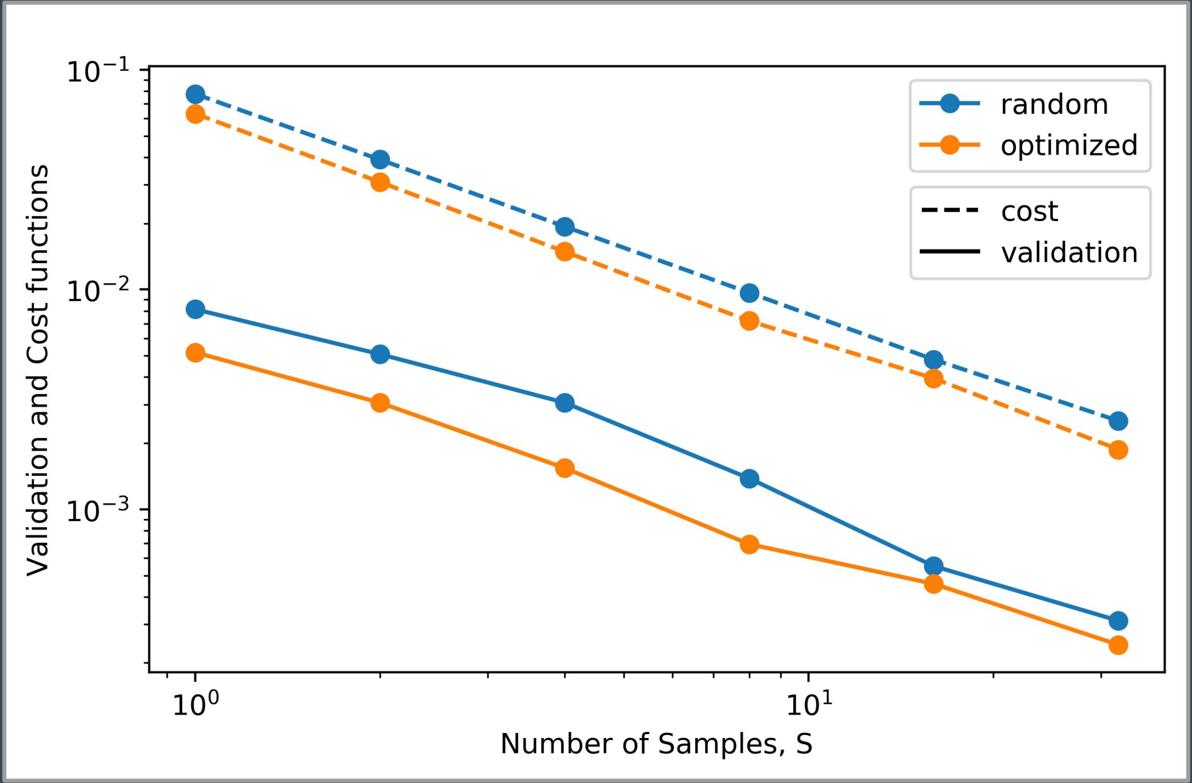
A Lindbladian model can cover most of the dynamics.



Improving Sensitivity Through Optimal Control

Naturally, the gradient descent procedure can be run in reverse to devise a control pulse for desired final state or another metric.

E.g. Fisher Information can be maximized.



What about Stochastic Master Equations

Alternative blackbox RNN approach in arxiv:1811.12420

$$d\rho = \boxed{-i\,[H,\rho]\,dt} + \boxed{\mathcal{D}\,(c,\rho)\,dt} + \boxed{\sqrt{\eta}\,\mathcal{H}\,(c,\rho)\,dW}$$

Unitary (Hamiltonian) Dissipative (Lindbladian) Weak Measurement (Backaction)

Parameterization

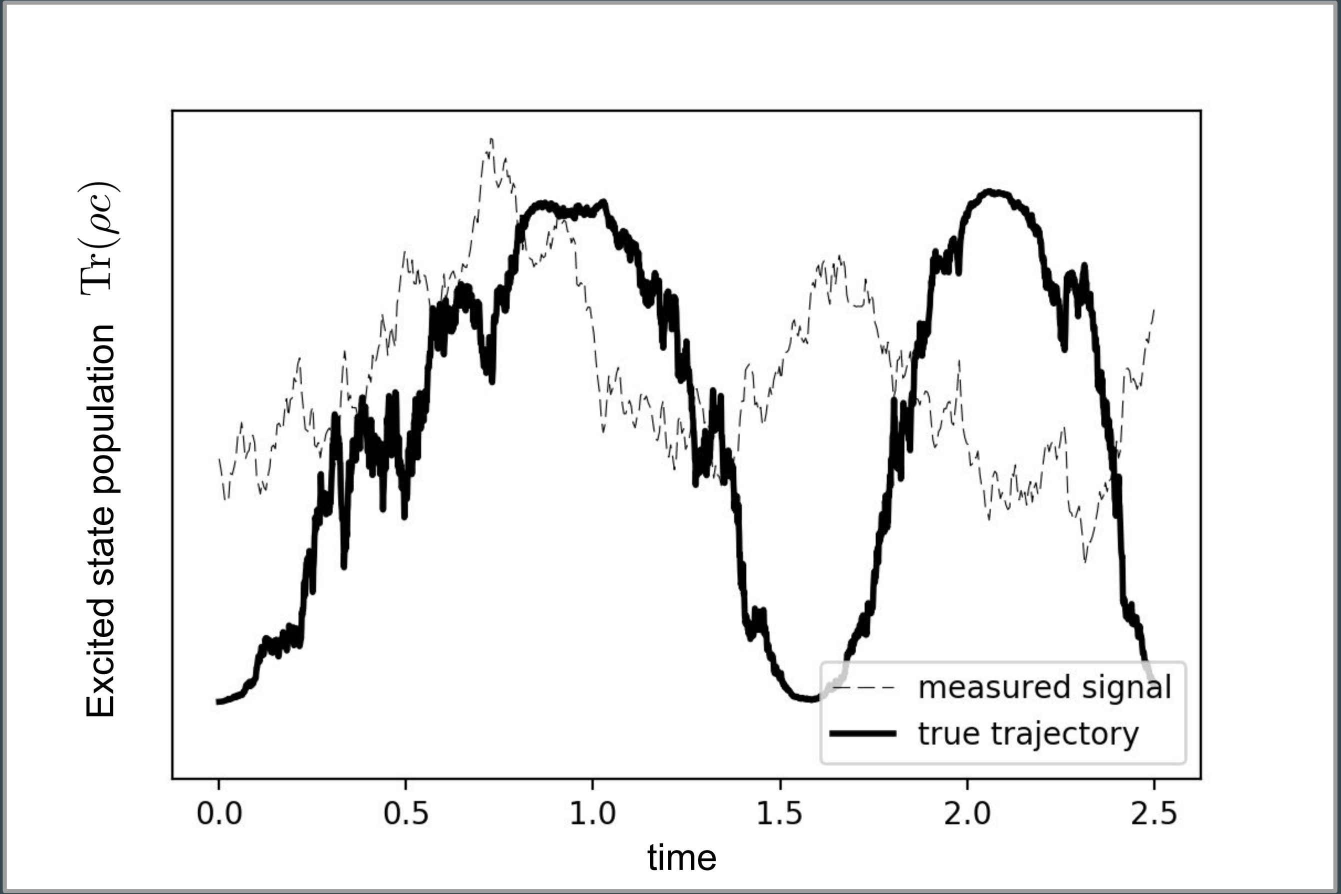
$H = \boxed{\Omega}\sigma_x$ Oscillation frequency

$c = \sqrt{\boxed{\gamma}}\sigma_z$ Dissipation strength

$\boxed{\eta}$ Measurement efficiency

$$dW = \left(\boxed{V} - 2\sqrt{\eta}\,\mathrm{Tr}\,(\rho c)\right) dt$$

Measurement Record



Example Trajectory: Weakly measuring the excited state population

$$\mathrm{d}W = \left(V - 2\sqrt{\eta}\mathrm{Tr}\left(\rho c\right)\right)\mathrm{d}t$$

The Model to which we fit

$$\mathrm{d}\rho = -i\left[H,\rho\right]\mathrm{d}t$$

Unitary
(Hamiltonian)

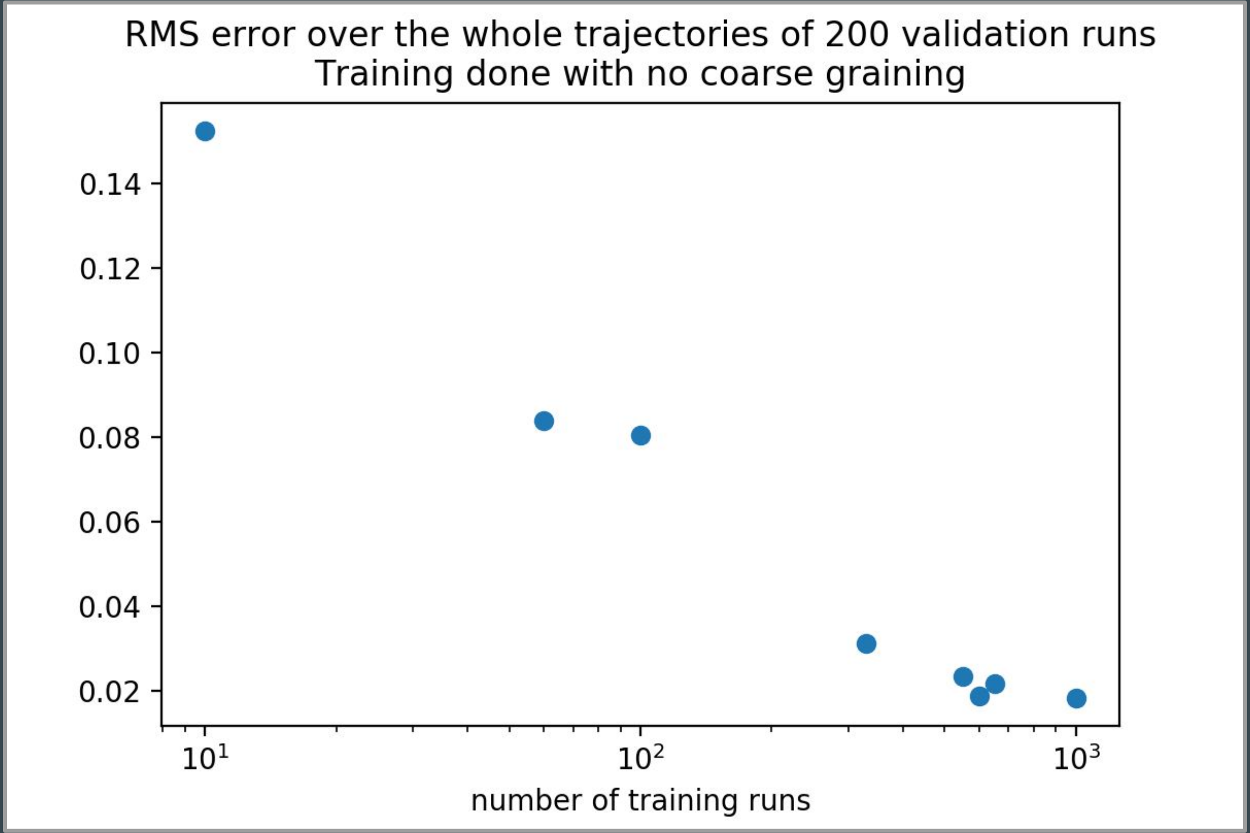
$$+\,\mathcal{D}\left(c,\rho\right)\mathrm{d}t$$

Dissipative
(Lindbladian)

$$-\,2\eta\mathcal{H}(c,\rho)\mathrm{Tr}\left(\rho c\right)\mathrm{d}t$$

$$+\,\sqrt{\eta}\mathcal{H}(c,\rho)dV$$

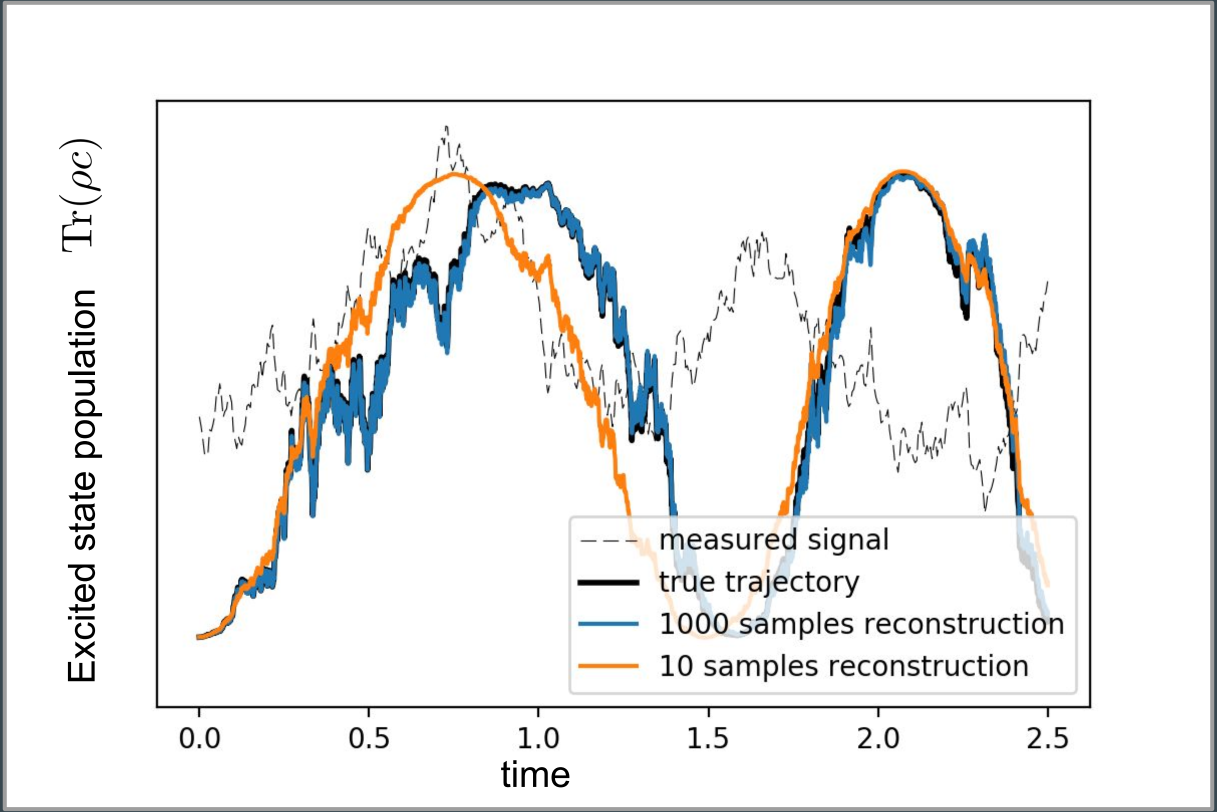
Weak
Measurement
(Backaction)



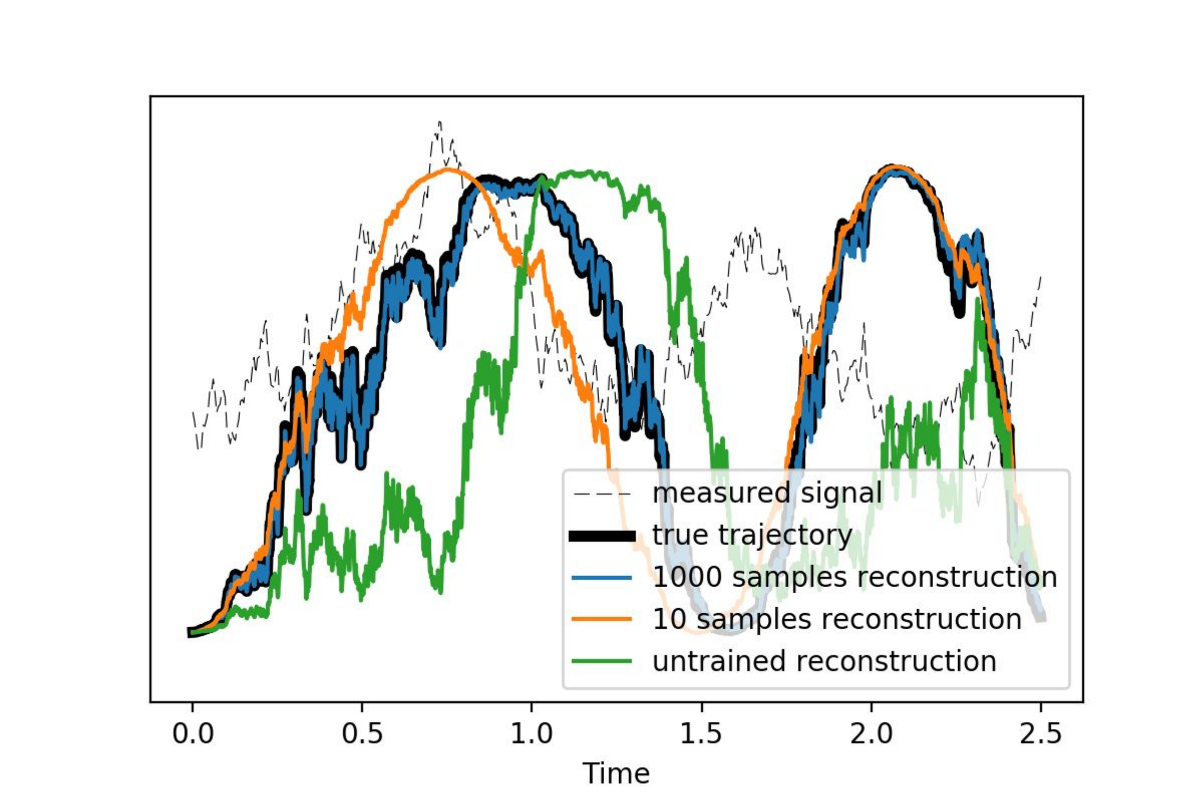
Estimator performance vs Amount of measurement data

Example Reconstructions

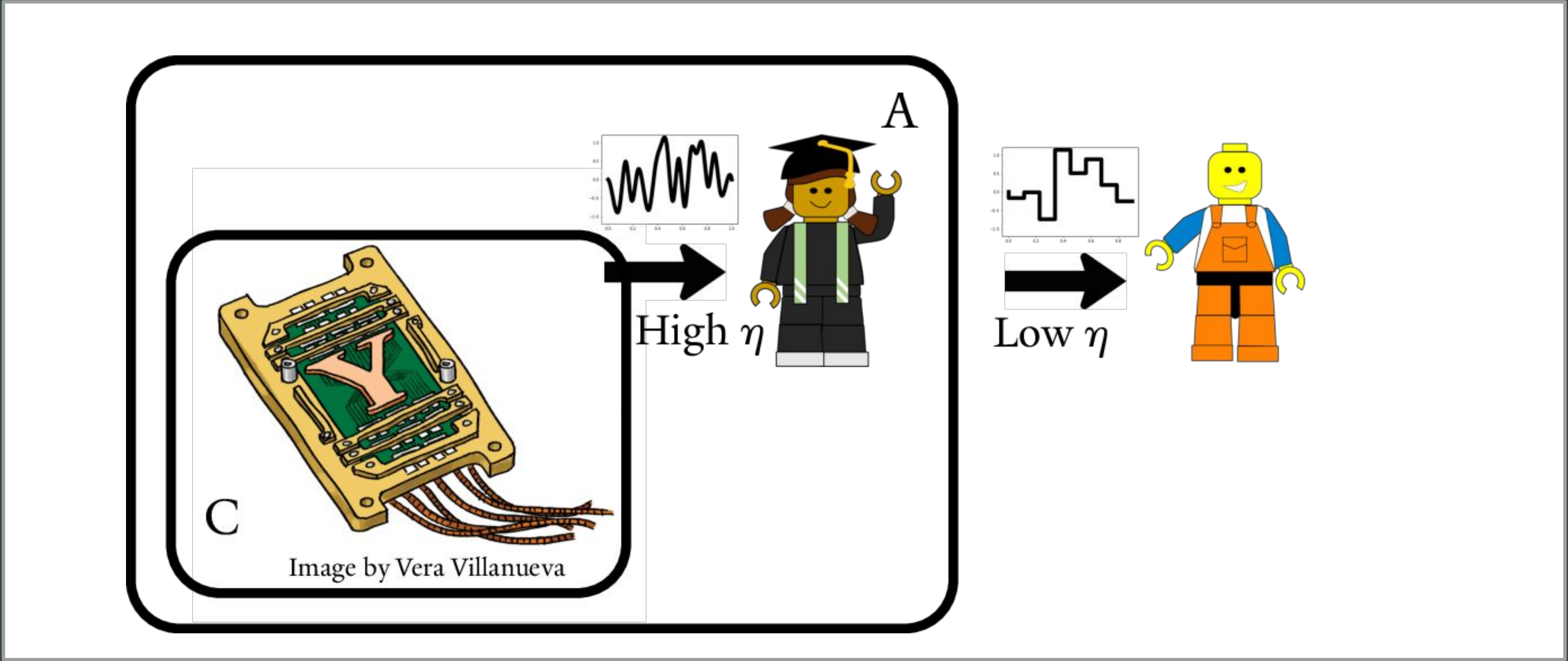
1000× less measurement data required compared to RNN approaches.



Example Reconstruction (including untrained run)

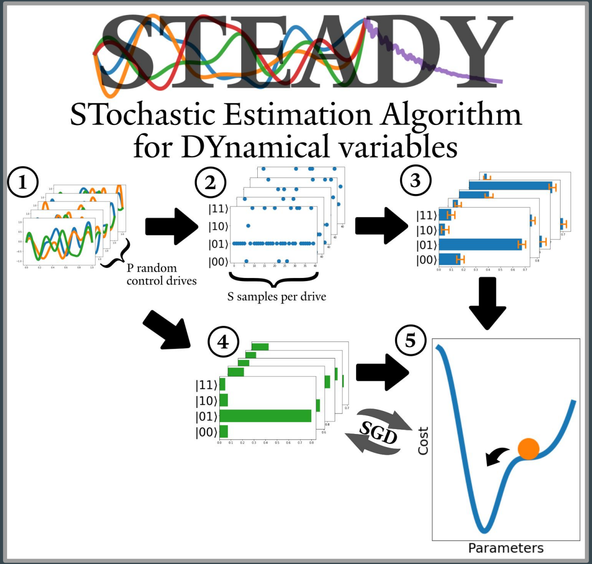


Weak Measurement Sampling Rate vs Numerical Timestep



Stochastic Estimation Overview

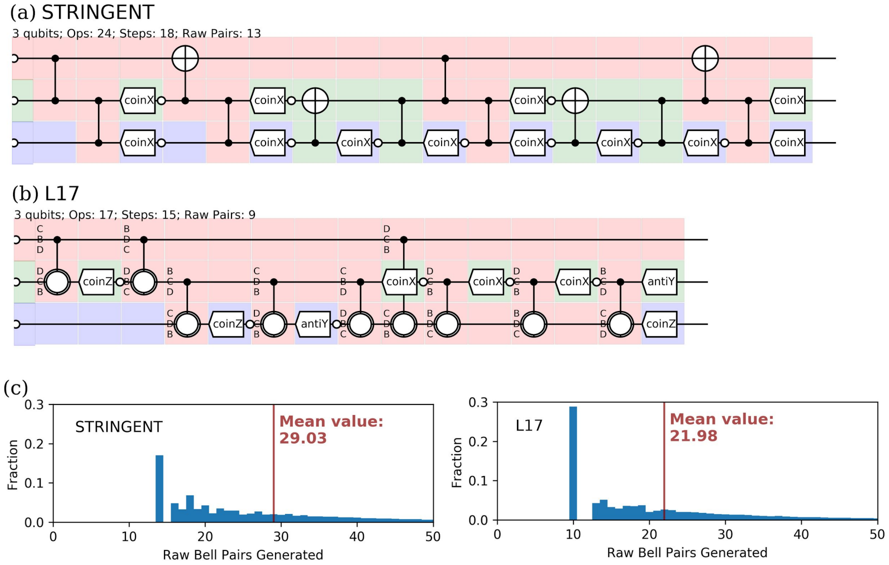
- Reaching for the information theory limit of performance
- Learning the whole Hamiltonian/Lindbladian, not just a set of gates
  - Sparse Efficient Representation
  - Continuous (for use in optimal control)
  - “Experimental design” is easy to plug in
- Deals with SPAM and non-unitary errors
- General while still Interpretable
- Embarrassingly simple implementation



Outline

- Control and Calibration
- Purifying Resources with Imperfect Circuits

Purifying Resources through Imperfect Circuits



arxiv:1712.09762 / Quantum 2019

Real Bell Pairs are Imperfect

$$A = |\phi_+\rangle = |00\rangle + |11\rangle$$

Y error

$$B = |\psi_-\rangle = |01\rangle - |10\rangle$$

bit flip error (X)

$$C = |\psi_+\rangle = |01\rangle + |10\rangle$$

phase error (Z)

$$D = |\phi_-\rangle = |00\rangle - |11\rangle$$

Purification of Imperfect Resources

$$A = |\phi_+\rangle = |00\rangle + |11\rangle$$
$$B = |\psi_-\rangle = |01\rangle - |10\rangle$$
$$C = |\psi_+\rangle = |01\rangle + |10\rangle$$
$$D = |\phi_-\rangle = |00\rangle - |11\rangle$$

Four possible Bell states A, B, C, and D.

Typically an entanglement generator would not provide perfect pairs, rather (for instance):

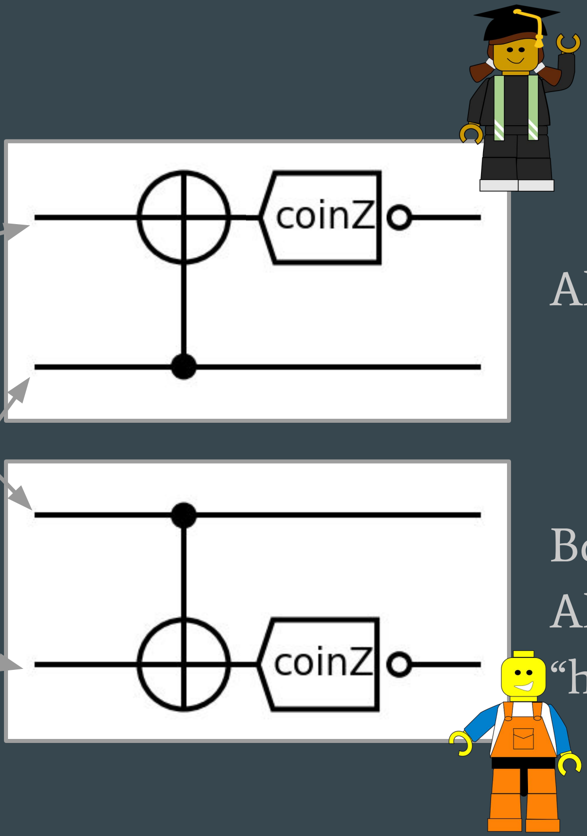
A at 90%
B, C, and D at 3.3%

We use two such imperfect Bell pairs to distil one higher quality pair.

Purification

Sacrificial Bell pair.

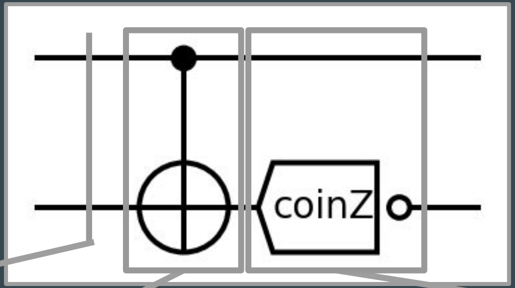
Bell pair to be purified.



Alice’s side of the circuit.

Bob’s side of the circuit.
Alice does the same to her “half pairs”.

Purification



Bob's side of the circuit.
Alice does the same to her "half pairs".

AA	81%
AB	3%
AC	3%
AD	3%
others	...

AA remains AA
DD becomes AD
etc

Trace over the second pair, selecting only A or D.

Result:
93% A in the first pair.

Tracing over the second pair returns the same quality pair as the one we started with.

$$A = |\phi_+\rangle = |00\rangle + |11\rangle$$
$$B = |\psi_-\rangle = |01\rangle - |10\rangle$$
$$C = |\psi_+\rangle = |01\rangle + |10\rangle$$
$$D = |\phi_-\rangle = |00\rangle - |11\rangle$$

The Initial State

Single pair:

A - F
B - q
C - q
D - q

e.g. F=90% and q=3.3%

Two pairs:

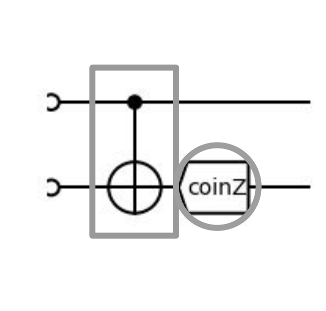
AA - FF
AB - Fq
AC - Fq
AD - Fq
BA - Fq
BB - qq
BC - qq
BD - qq
CA - Fq
CB - qq
CC - qq
CD - qq
DA - Fq
DB - qq
DC - qq
DD - qq

Restating in the form of permutations and selections

Case study:

CNOT as
a permutation
on the Bell basis

Coincidence Z
measurement
as selecting half
of the set



AA	:	FF	⇨	FF
AB	:	Fq	⇨	~~qq~~
AC	:	Fq	⇨	~~Fq~~
AD	:	Fq	⇨	qq
BA	:	Fq	⇨	qq
BB	:	qq	⇨	~~qq~~
BC	:	qq	⇨	~~Fq~~
BD	:	qq	⇨	qq
CA	:	Fq	⇨	qq
CB	:	qq	⇨	~~qq~~
CC	:	qq	⇨	~~Fq~~
CD	:	qq	⇨	qq
DA	:	Fq	⇨	Fq
DB	:	qq	⇨	~~Fq~~
DC	:	qq	⇨	~~qq~~
DD	:	qq	⇨	Fq

$$F_{final} = F^2 + q^2$$

CNOT:
AA⇨AA
AB⇨DB
AC⇨AC
AD⇨DD
BA⇨BC
BB⇨CD
BC⇨BA
BD⇨CB
CA⇨CC
CB⇨BD
CC⇨CA
CD⇨BB
DA⇨DA
DB⇨AB
DC⇨DC
DD⇨AD

```
Coincidence Z:
take A and D
```

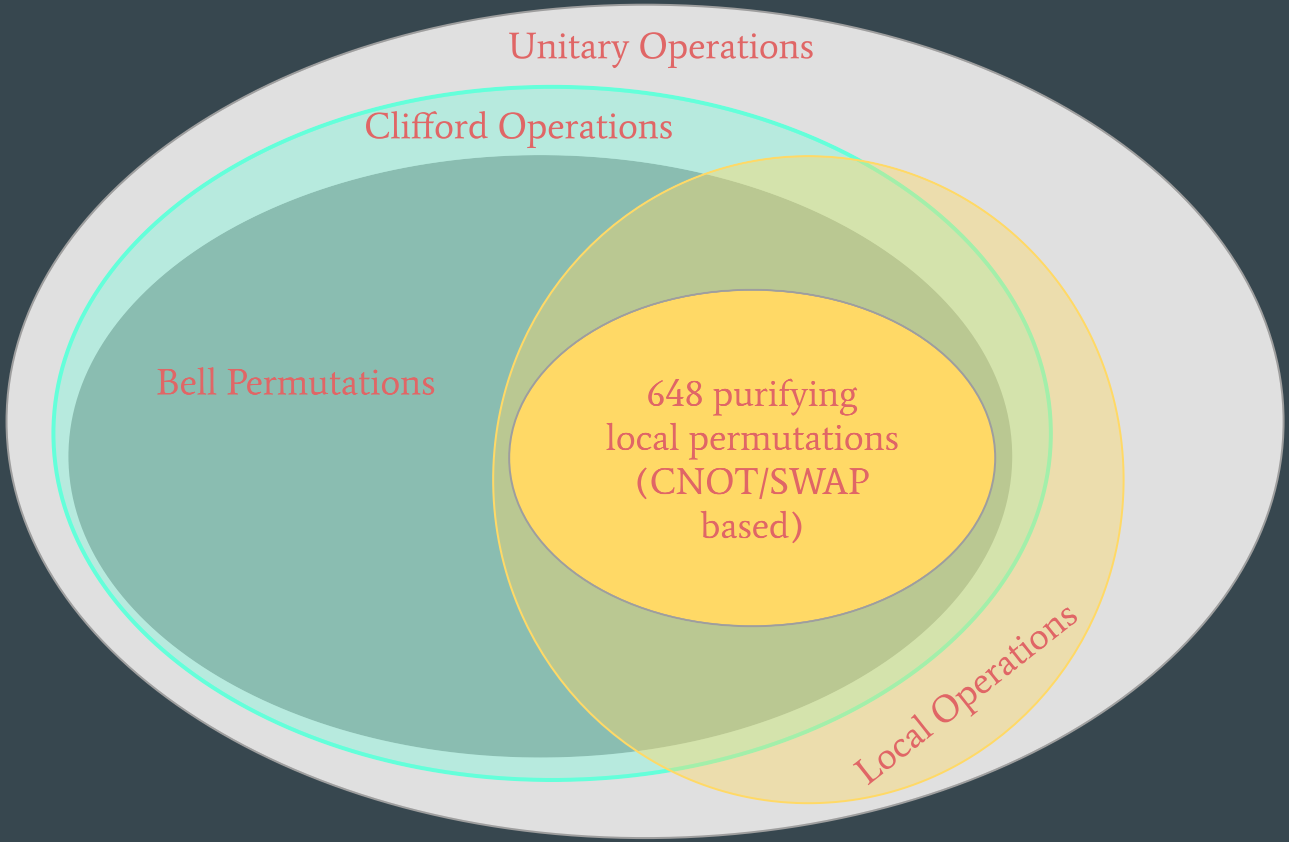
$$A = |\phi_+\rangle = |00\rangle + |11\rangle$$

$$B = |\psi_-\rangle = |01\rangle - |10\rangle$$

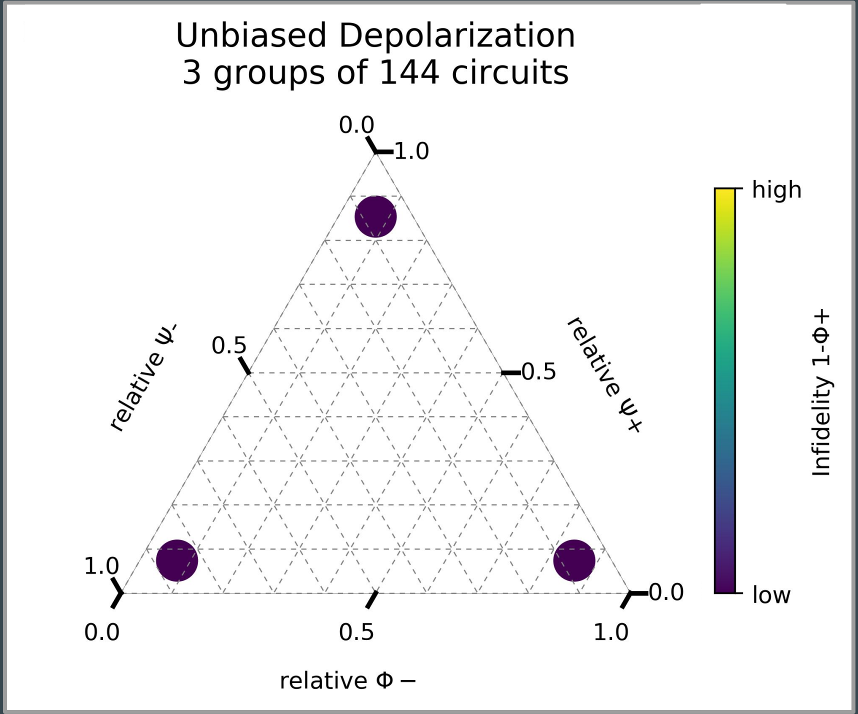
$$C = |\psi_+\rangle = |01\rangle + |10\rangle$$

Why don’t we pick a better permutation?

Purification-enabling Gates



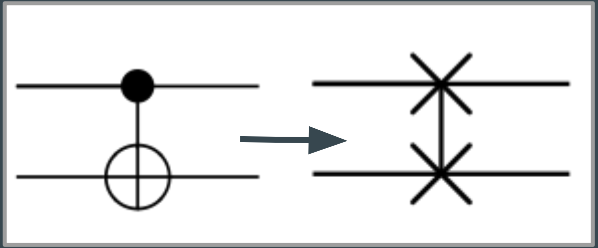
Useful Permutation Subgroups and Cosets



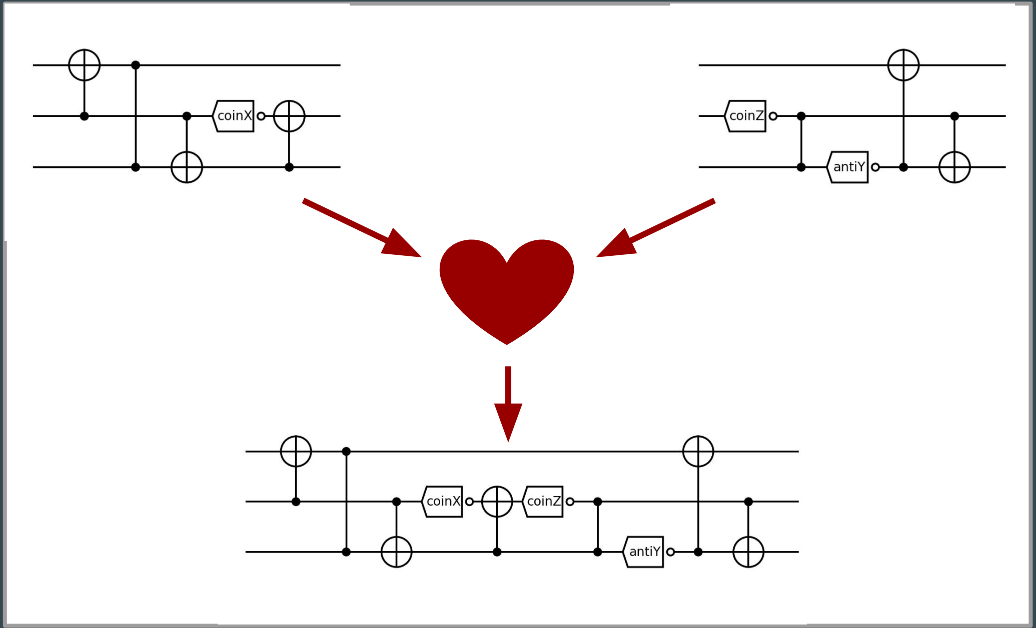
Making Bigger and Better Circuits

Optimized Purification (Genetic Algorithms)

Mutations



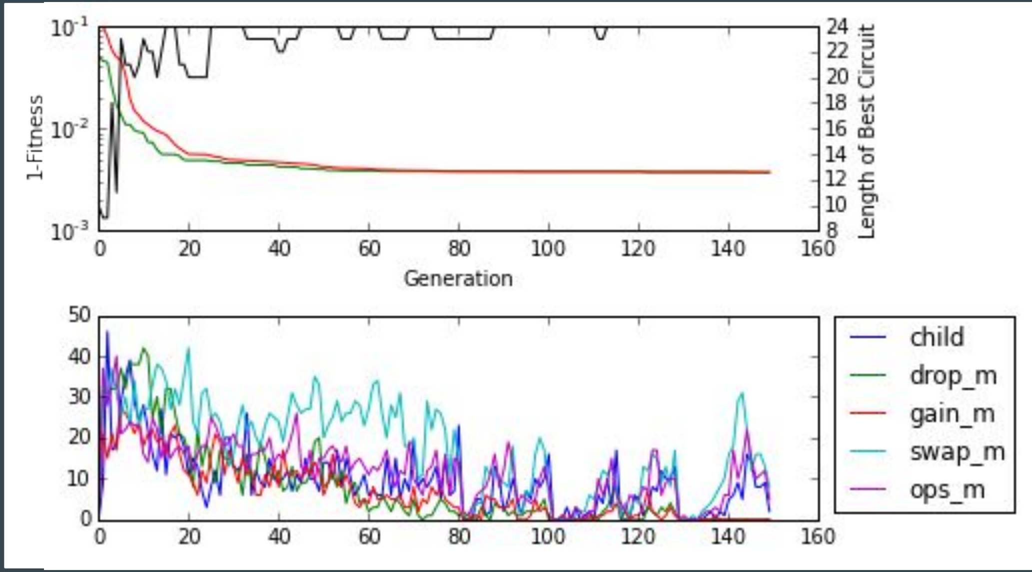
Creating Offspring Circuits



A Hundred Generations Later (5min laptop time)

For a given error model and specified error parameters quickly find near-optimal purification circuits.

Various optimizations related to the structure of the problem were implemented.



Real-time trace of the optimization process.

Better Than Prior Art

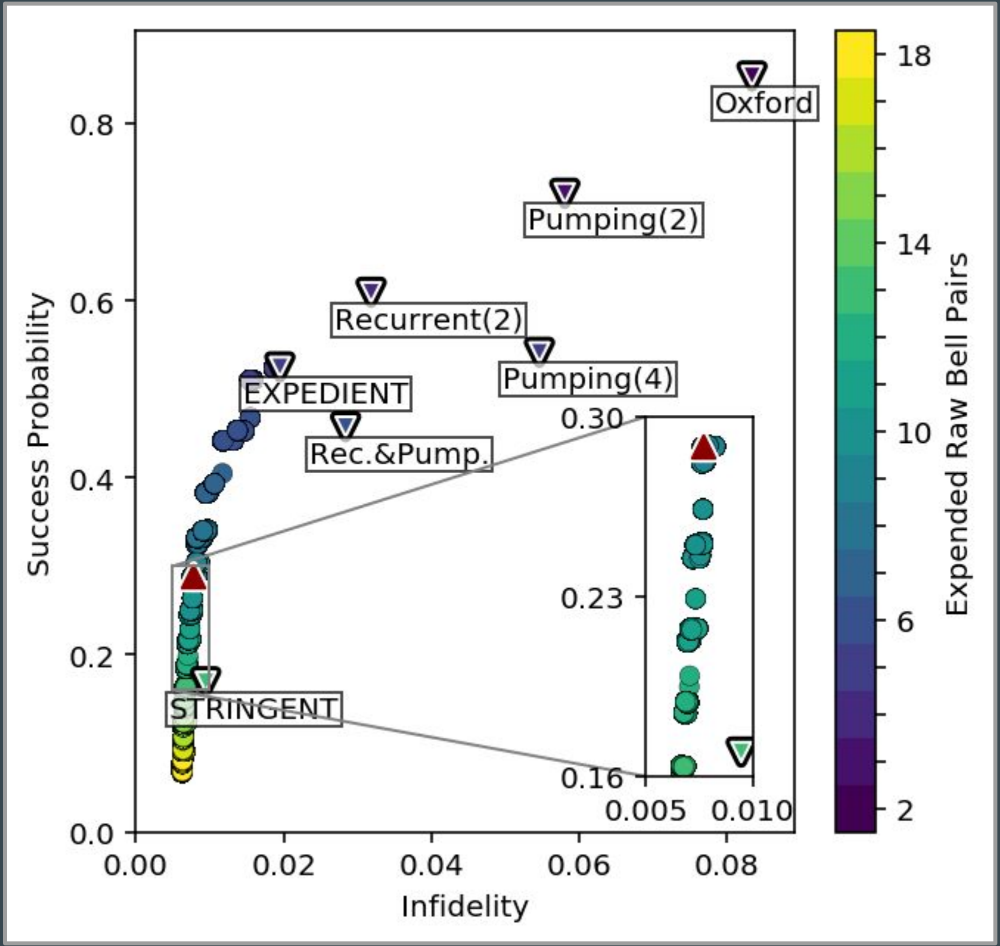
Breadth 3, simulated for

- Initial Bell Pair Fidelity 0.9
- Local Operations Fidelity 0.99
- Depolarization Error Model

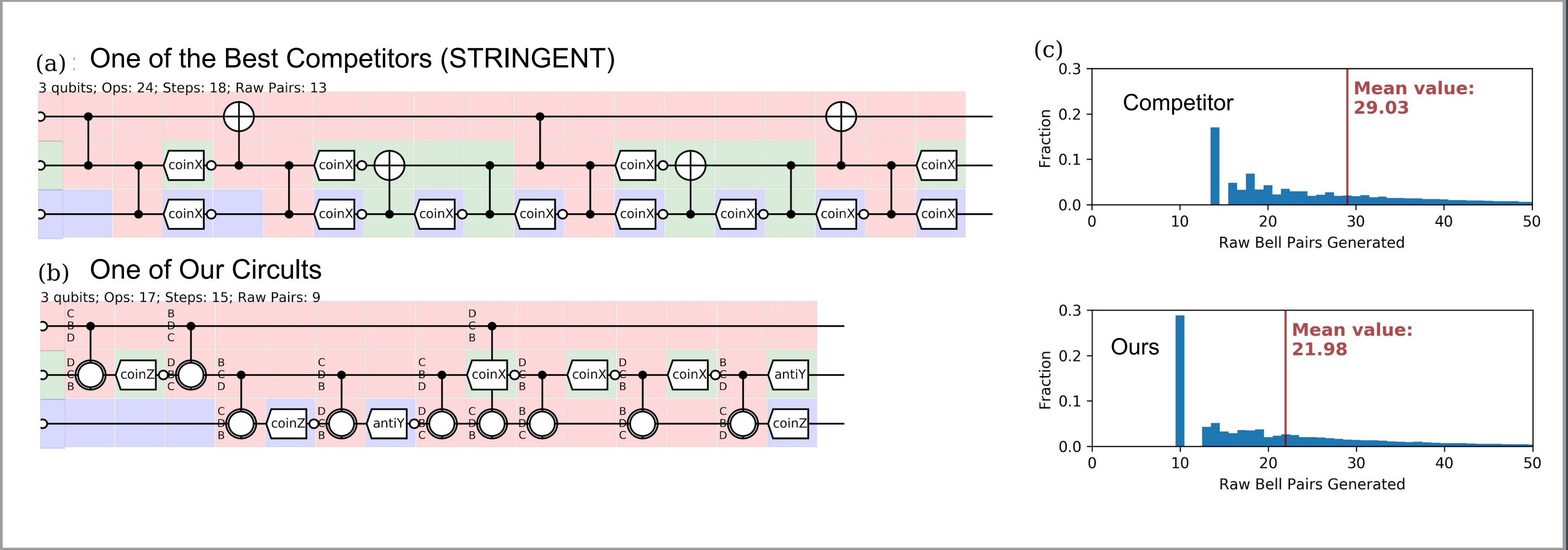
Available at

qevo.krastanov.org

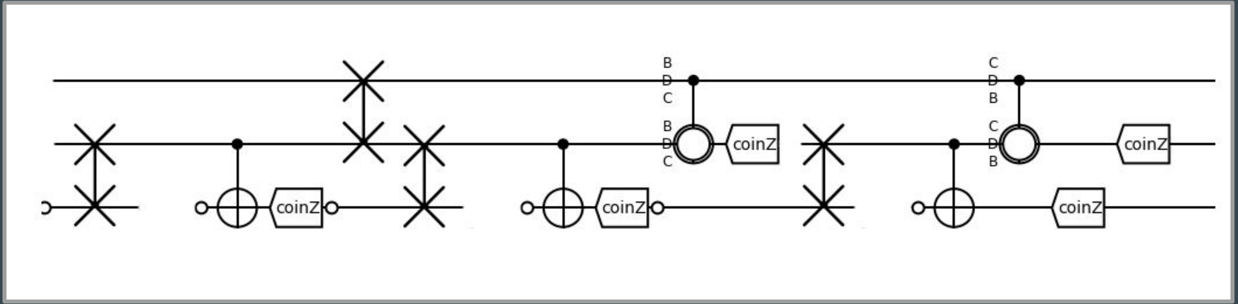
Circuits From:
- Nickerson, Nature Comm. 4, 1756 (2014)
- Deutsch, PRL 77, 2818 (1996)
- Dür, PRA, 59, 169 (1999)



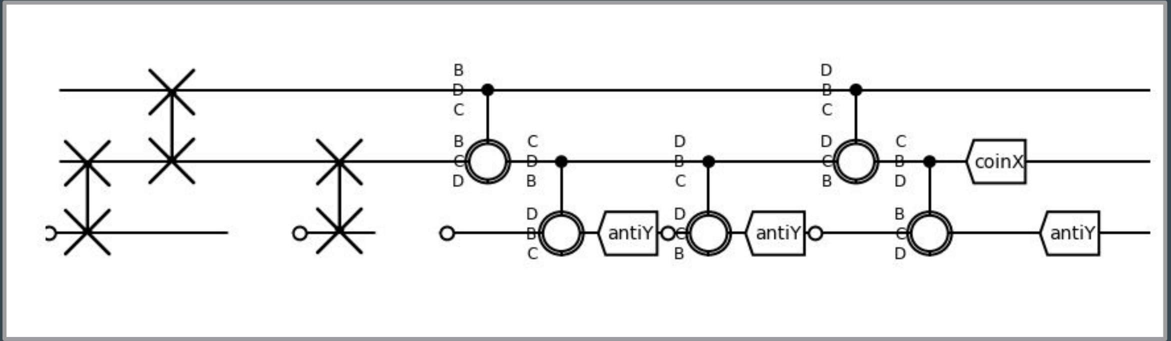
A Particular Example



Example with a Communication Qubit

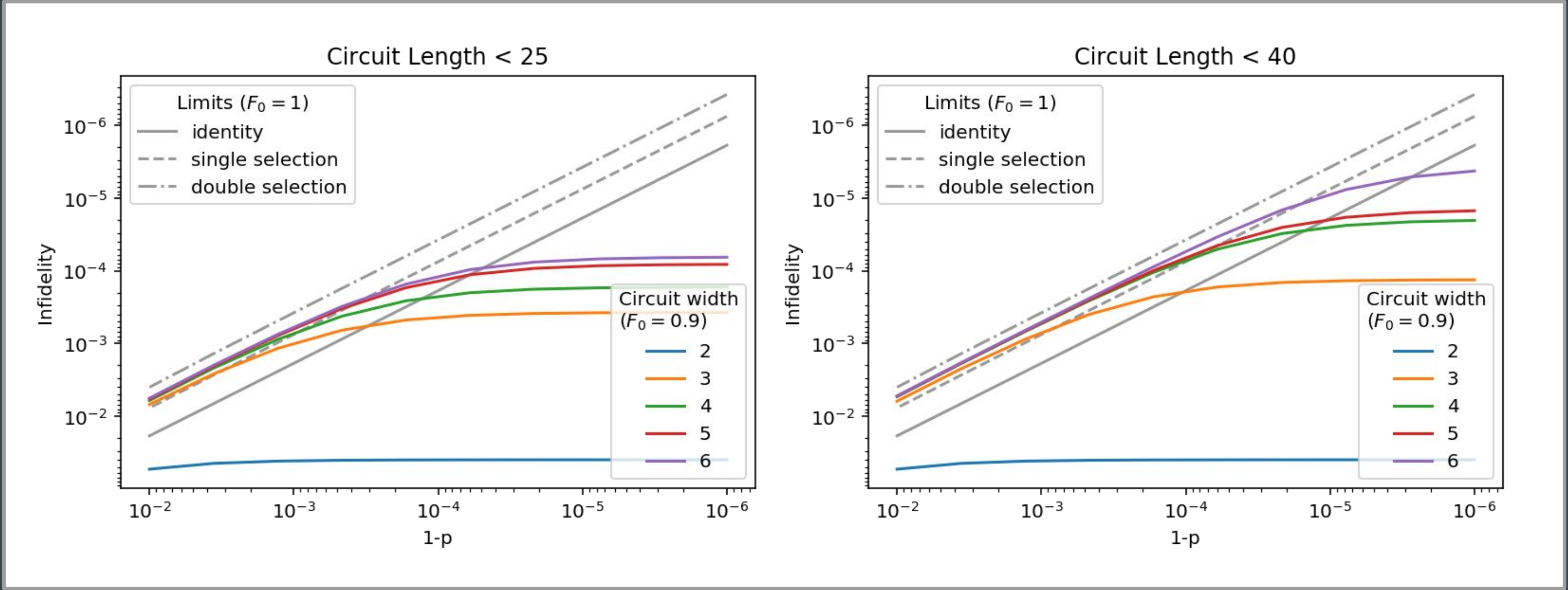


Nigmatulin et al:
- 6 raw Bell pairs
- 2.46% infidelity

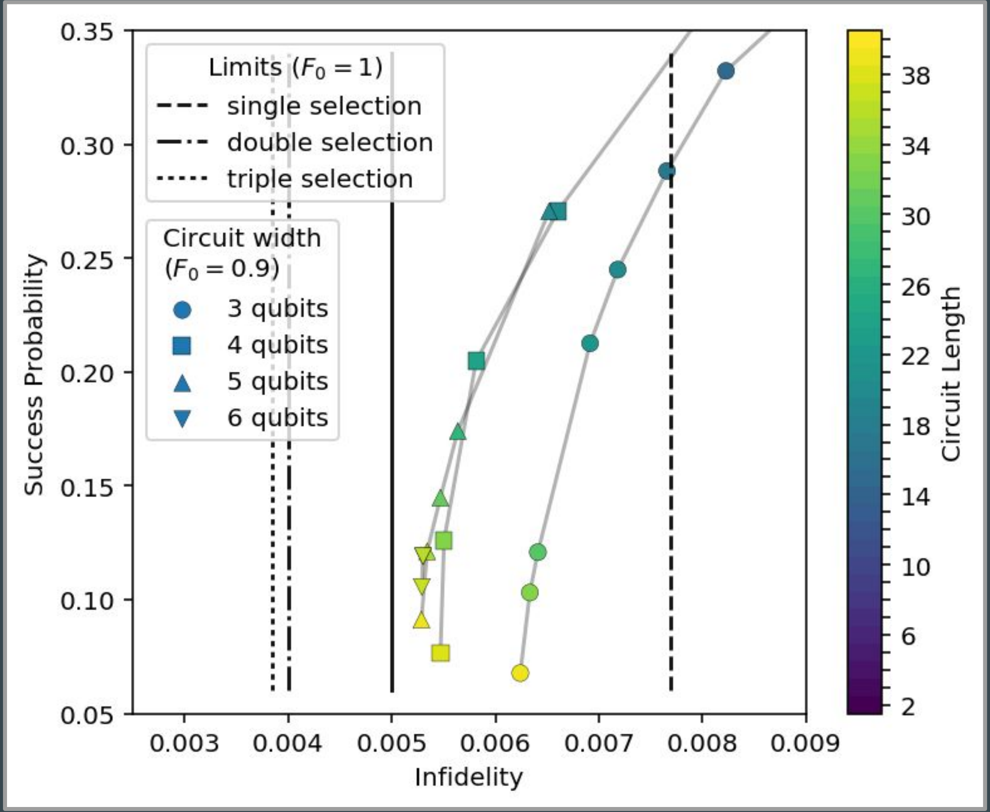


Our optimized circuit:
- 5 raw Bell pairs
- 1.77% infidelity

Initialization vs Operations' Infidelity



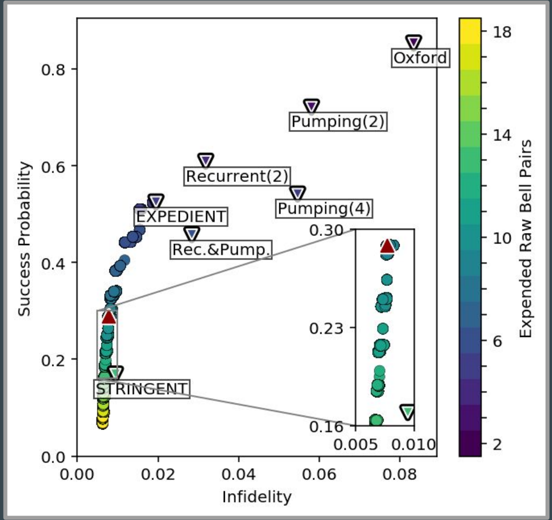
Diminishing Returns



Optimized Circuits Overview

The optimizer is giving the best known circuits

- They are customized to your hardware's error model
- The software also provides:
  - Monte Carlo resource estimates
  - Exact symbolic evaluation for the infidelities
  - code and examples at qevo.krastanov.org



Speculative Outlook

Bringing the two toolkits together:

- The circuit optimization required working efficiently with noisy Clifford circuits
- The parameter estimation is not yet scalable, as it requires quantum simulations

However, with some finite effort we can:

- Design Clifford gates through optimal control over the estimated model
- Run a random circuit of such imperfect Clifford gates
- Perform infidelity estimation across this entire pipeline
- Thus the parameter estimation would run efficiently even for larger systems

Thanks to Collaborators, Classmates, and Friends!

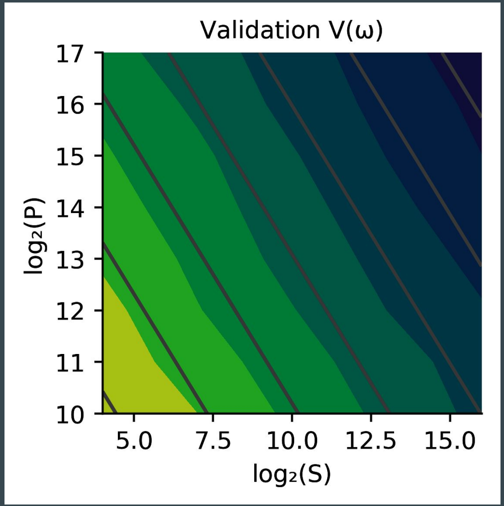
To colleagues and collaborators for working together, expanding our knowledge of the world.

And to friends and classmates, for supporting each other, and making life at Yale so much better.

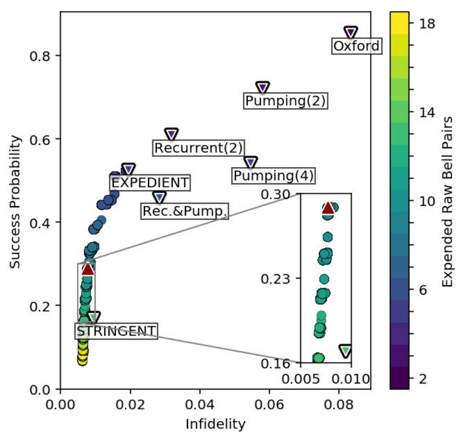
And thank you, Liang, for mentoring me and for enabling my teaching and outreach interests!

High schoolers building motion-sensing LED bracelets (Eng. Day 2017) and learning the math and CS behind drawing fractals (Math Art 2018).

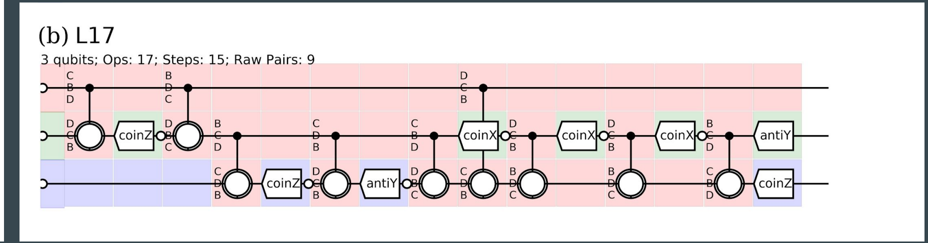
Questions?



Validation V(ω)
log₂(P)
log₂(S)
17
16
15
14
13
12
11
10
5.0
7.5
10.0
12.5
15.0



Oxford
Pumping(2)
Recurrent(2)
Pumping(4)
EXPEDIENT
Rec.&Pump.
STRINGENT
Success Probability
Infidelity
Expended Raw Bell Pairs
0.0
0.2
0.4
0.6
0.8
0.00
0.02
0.04
0.06
0.08
0.30
0.23
0.16
0.005
0.010
2
6
10
14
18



(b) L17
3 qubits; Ops: 17; Steps: 15; Raw Pairs: 9
coinZ
coinX
antiY