

# Semiquantum Algorithms for Characterization and Control

Christopher Granade

*[www.cgranade.com/research/talks/msr-2014](http://www.cgranade.com/research/talks/msr-2014)*

---

Joint work with:

Nathan Wiebe   Christopher Ferrie   D. G. Cory

Institute for Quantum Computing  
University of Waterloo, Ontario, Canada

June 30, 2013  
Microsoft Research

# Characterizing Quantum Systems

Characterizing quantum systems is an essential task in quantum information.

- Accurate knowledge required for high-fidelity control.
- Allows for comparing to proven and estimated thresholds.
- Characterization allows for *validating* control.

# State Tomography

Common task: characterize the *state*  $\rho$  of a quantum system.

Tomographic approach: measure  $p_i = \text{Tr}(E_i \rho)$  for a *positive operator-valued measure*  $\{E_i\}$ .

Given measurement record  $\{d_i\}$ , what should  $\hat{\rho}$  be?

- Need to ensure  $\rho \geq 0$ , is full-rank.
- Exponentially many parameters needed.
- How to parameterize uncertainty?

# Process Tomography

Can also consider learning about quantum processes,  
 $S : \rho_i \mapsto \rho_f$ .

- Even more parameters
- Negativity: difficult to separate sampling error from violation of assumptions (e.g. initially-correlated states)

(Altepeter et al. 2003 [10/dtdk4z](#); Boulant et al. 2003 [10/fgvbg9](#); Weinstein et al. 2004 [10/bn6sn2](#))

# Bayesian Approaches

Model data collection as a probability distribution, called a *likelihood function*

$$\Pr(d|\underline{x}; \underline{e}).$$

$d$ : data,    $\underline{x}$ : model,    $\underline{e}$ : experiment

# Bayesian Approaches

Model data collection as a probability distribution, called a *likelihood function*

$$\Pr(d|\underline{x}; \underline{e}).$$

$d$ : data,    $\underline{x}$ : model,    $\underline{e}$ : experiment

## Example

Single qubit, Larmor precession at an unknown frequency  $\omega$ , unknown dephasing time  $T_2$ :

$$H(\omega) = \frac{\omega}{2}\sigma_z, \quad |\psi_{\text{in}}\rangle = |+\rangle, \quad M = \{|+\rangle\langle+|, |-\rangle\langle-|\}$$

$$\Pr(d = 0|\underline{x} = (\omega, T_2); \underline{e} = (t)) = \frac{1}{2}(1 - e^{-t/T_2}) + e^{-t/T_2} \cos^2(\omega t/2)$$

# Updating Knowledge

Once we have a likelihood function for our model, we can reason about

$$\Pr(\underline{x}|\underline{d}, \underline{e}),$$

what we know about our model having seen some data.

# Updating Knowledge

Once we have a likelihood function for our model, we can reason about

$$\Pr(\underline{x}|\underline{d}, \underline{e}),$$

what we know about our model having seen some data.  
By Bayes' rule,

$$\Pr(\underline{x}|\underline{d}, \underline{e}) = \frac{\Pr(\underline{d}|\underline{x}; \underline{e})}{\Pr(\underline{d}|\underline{e})} \Pr(\underline{x}),$$

telling us that our knowledge is intimately connected to our ability to simulate.



# Updating Knowledge

Once we have a likelihood function for our model, we can reason about

$$\Pr(\underline{x}|\underline{d}, \underline{e}),$$

what we know about our model having seen some data.  
By Bayes' rule,

$$\Pr(\underline{x}|\underline{d}, \underline{e}) = \frac{\Pr(\underline{d}|\underline{x}; \underline{e})}{\Pr(\underline{d}|\underline{e})} \Pr(\underline{x}),$$

telling us that our knowledge is intimately connected to our ability to simulate.

Estimate  $\hat{\underline{x}}$  as the expectation over  $\underline{x}$ ,

$$\hat{\underline{x}} = \mathbb{E}[\underline{x}] = \int \underline{x} \Pr(\underline{x}) \, \mathrm{d}\underline{x}.$$

# Loss

*Figure of merit:* how well have we learned a model?

Assign to estimate  $\hat{x}$  of a “true” model  $x$  a *loss*, describing how bad  $\hat{x}$  does at estimating  $x$ .

# Loss

*Figure of merit:* how well have we learned a model?

Assign to estimate  $\hat{\underline{x}}$  of a “true” model  $\underline{x}$  a *loss*, describing how bad  $\hat{\underline{x}}$  does at estimating  $\underline{x}$ .

## Definition (Quadratic Loss)

$$L_{\underline{\underline{Q}}}(\hat{\underline{x}}, \underline{x}) = (\hat{\underline{x}} - \underline{x})^T \underline{\underline{Q}} (\hat{\underline{x}} - \underline{x}),$$

where  $\underline{\underline{Q}}$  is a positive semidefinite scale matrix.

# Loss

*Figure of merit:* how well have we learned a model?

Assign to estimate  $\hat{\underline{x}}$  of a “true” model  $\underline{x}$  a *loss*, describing how bad  $\hat{\underline{x}}$  does at estimating  $\underline{x}$ .

## Definition (Quadratic Loss)

$$L_{\underline{\underline{Q}}}(\hat{\underline{x}}, \underline{x}) = (\hat{\underline{x}} - \underline{x})^T \underline{\underline{Q}} (\hat{\underline{x}} - \underline{x}),$$

where  $\underline{\underline{Q}}$  is a positive semidefinite scale matrix.

The quadratic loss generalizes the MSE for multiple parameters.

# Risk and Bayes Risk

**Estimator:** function from data records  $D$  to estimates  $\hat{x}(D)$ .  
What is the expected loss?

# Risk and Bayes Risk

**Estimator:** function from data records  $D$  to estimates  $\hat{x}(D)$ .  
What is the expected loss?

## Definition (Risk)

$$R(\hat{x}, \underline{x}) = \mathbb{E}_D[L(\hat{x}(D), \underline{x})]$$

# Risk and Bayes Risk

**Estimator:** function from data records  $D$  to estimates  $\hat{x}(D)$ .  
What is the expected loss?

## Definition (Risk)

$$R(\hat{x}, \underline{x}) = \mathbb{E}_D[L(\hat{x}(D), \underline{x})]$$

Since we don't know the true model *a priori*, we average again to obtain the Bayes risk.

# Risk and Bayes Risk

**Estimator:** function from data records  $D$  to estimates  $\hat{x}(D)$ .  
What is the expected loss?

## Definition (Risk)

$$R(\hat{x}, \underline{x}) = \mathbb{E}_D[L(\hat{x}(D), \underline{x})]$$

Since we don't know the true model *a priori*, we average again to obtain the Bayes risk.

## Definition (Bayes Risk)

$$r(\hat{x}, \pi) = \mathbb{E}_{\underline{x} \sim \pi}[R(\hat{x}, \underline{x})]$$



# Cramér-Rao Bound

The Fisher information

$$\underline{\underline{I}}(\underline{x}) = \mathbb{E}_D[(\nabla_{\underline{x}} \log \Pr(D|\underline{x}))(\nabla_{\underline{x}} \log \Pr(D|\underline{x}))^T]$$

describes how much information about  $\underline{x}$  is obtained by sampling data.

# Cramér-Rao Bound

The Fisher information

$$\underline{\underline{I}}(\underline{x}) = \mathbb{E}_D[(\nabla_{\underline{x}} \log \Pr(D|\underline{x}))(\nabla_{\underline{x}} \log \Pr(D|\underline{x}))^T]$$

describes how much information about  $\underline{x}$  is obtained by sampling data.

The Cramér-Rao bound tells how well any unbiased estimator can do. If  $\underline{\underline{Q}} = \mathbb{1}$ , then

$$R(\hat{\underline{x}}, \underline{x}) = \text{Tr}(\text{Cov}(\hat{\underline{x}})) \geq \text{Tr}(\underline{\underline{I}}(\underline{x})^{-1}).$$

# Cramér-Rao Bound

The Fisher information

$$\underline{\underline{I}}(\underline{x}) = \mathbb{E}_D[(\nabla_{\underline{x}} \log \Pr(D|\underline{x}))(\nabla_{\underline{x}} \log \Pr(D|\underline{x}))^T]$$

describes how much information about  $\underline{x}$  is obtained by sampling data.

The Cramér-Rao bound tells how well any unbiased estimator can do. If  $\underline{\underline{Q}} = \mathbb{1}$ , then

$$R(\hat{\underline{x}}, \underline{x}) = \text{Tr}(\text{Cov}(\hat{\underline{x}})) \geq \text{Tr}(\underline{\underline{I}}(\underline{x})^{-1}).$$

Compare: quantum Cramér-Rao bound (Heisenberg limit).  
Not necessarily the limit of practical interest.

# Bayesian Cramér-Rao Bound

Integrating the Fisher information over the prior  $\pi$  results in a Bayesian analog, the Bayesian Cramér-Rao bound:

$$\underline{\underline{B}} := \mathbb{E}_{\underline{\underline{x}}}[\underline{\underline{I}}(\underline{\underline{x}})], \quad r(\pi) \geq \underline{\underline{B}}^{-1}.$$

If experiments are designed adaptively, then the current posterior is used instead of the prior.

The BCRB can be computed iteratively, making it useful for tracking optimality in an online fashion.

$$\underline{\underline{B}}_{k+1} = \underline{\underline{B}}_k + \begin{cases} \mathbb{E}_{\underline{\underline{x}} \sim \pi}[\underline{\underline{I}}(\underline{\underline{x}}; \underline{\underline{e}}_{k+1})] & \text{(non-adaptive)} \\ \mathbb{E}_{\underline{\underline{x}}|d_1, \dots, d_k}[\underline{\underline{I}}(\underline{\underline{x}}; \underline{\underline{e}}_{k+1})] & \text{(adaptive)} \end{cases}$$

# Sequential Monte Carlo

SMC is a numerical algorithm for generating samples from a distribution.

$$\text{prior} \xrightarrow{\text{Bayes' Rule}} \text{posterior}$$

Bayes' rule acts as a transition kernel from prior samples to posterior samples.

Posterior samples then give Monte Carlo approximations to integrals/expectations.

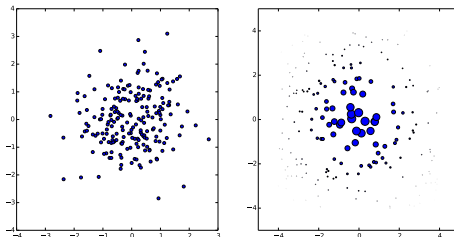
## SMC Approximation

$$\Pr(\underline{x}) \approx \sum_i^n w_i \delta(\underline{x} - \underline{x}_i)$$

(Doucet and Johansen 2011; Huszár and Houlby [10/s86](#); Granade et al. 2012 [10/s87](#))

# Ambiguity and Impovrishment

The SMC approximation can represent distributions by density of *particles* (left), or by weight (right).



Using weight is less numerically stable, results in smaller *effective* number of particles.

$$n_{\text{ess}} := 1 / \sum_i w_i^2$$

# Numerical Stability and Resampling

As data  $D$  is collected,  $\Pr(\underline{x}_i|D) \rightarrow 0$  for initial particles  $\{x_i\}$ .

- Results in  $n_{\text{ess}} \rightarrow 0$  as data is collected.

Can mitigate by *resampling*: moving information from the weights to the density of SMC particles.

Resampling when  $n_{\text{ess}}/n \leq 0.5$  helps preserve representative sample. Moreover, monitoring  $n_{\text{ess}}$  can herald some kinds of failures.

# Liu and West Algorithm

Draw new particles  $\underline{x}'$  from kernel density estimate

$$\begin{aligned}\Pr(\underline{x}') &\propto \sum_i w_i \exp \left( (\underline{x}' - \underline{\mu}_i)^T \underline{\underline{\Sigma}} (\underline{x}' - \underline{\mu}_i) \right) \\ \underline{\mu}_i &= a \underline{x}_i + (1 - a) \mathbb{E}[\underline{x}] \\ \underline{\underline{\Sigma}} &= (1 - a^2) \text{Cov}[\underline{x}]\end{aligned}$$

Set new weights to be uniform, so that  $n_{\text{ess}} = n$ .

- $a = 1, h = 0$ : Bootstrap filter, used in state-space applications like CONDENSATION.
- $a^2 + h^2 = 1$ : Ensures  $\mathbb{E}[\underline{x}'] = \mathbb{E}[\underline{x}]$  and  $\text{Cov}(\underline{x}') = \text{Cov}(\underline{x})$ , but assumes unimodality.
- $a = 1, h \geq 0$ : Allows for multimodality, emulating state-space with synthesized noise.

(West 1993; Liu and West 2001)

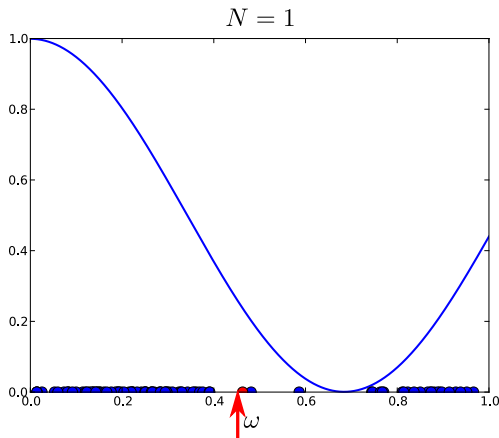


# Putting it All Together: The SMC Algorithm

- 1 Draw  $\{\underline{x}_i\} \sim \pi$ , set  $\{w_i\} = 1/n$ .
- 2 For each datum  $d_j \in D$ :
  - 1  $w_i \leftarrow w_i \times \Pr(d_j | \underline{x}_i; \underline{e}_j)$ .
  - 2 Renormalize  $\{w_i\}$ .
  - 3 If  $n_{\text{ess}}/n \leq 0.5$ , resample.
- 3 Report  $\hat{\underline{x}} := \mathbb{E}[\underline{x}] \approx \sum_i w_i \underline{x}_i$ .

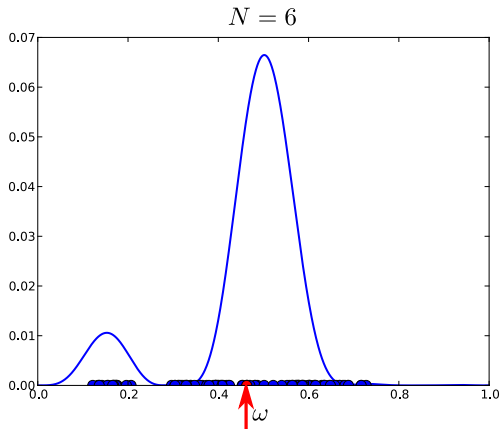
# Sequential Monte Carlo

With SMC and resampling, particles move towards the true model as data is collected.



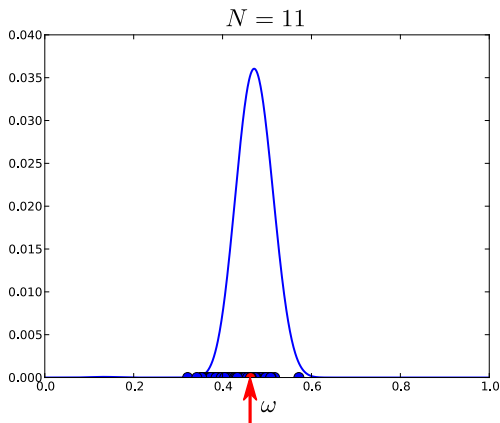
# Sequential Monte Carlo

With SMC and resampling, particles move towards the true model as data is collected.



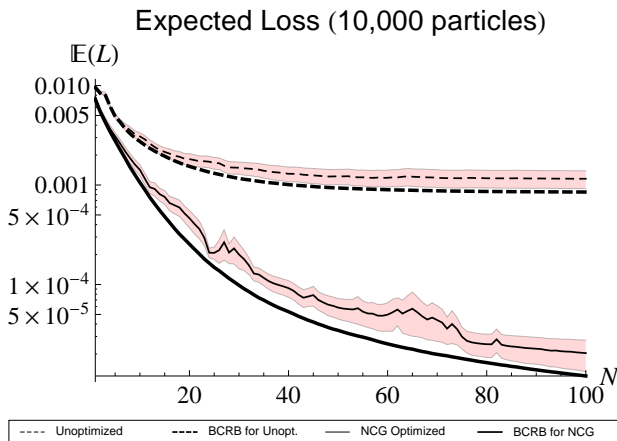
# Sequential Monte Carlo

With SMC and resampling, particles move towards the true model as data is collected.



# Near-Optimality for $\cos^2$

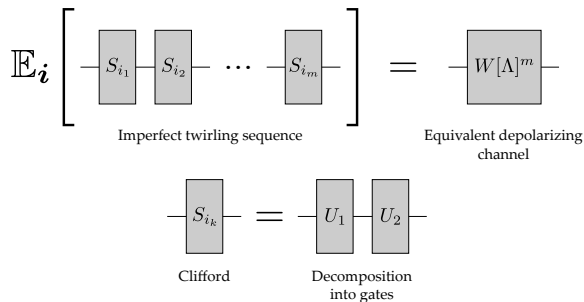
Using adaptive experiment design with Newton Conjugate-Gradient:



(Granade et al. 2012 [10/s87](#))

# Randomized Benchmarking Example

Applying sequences of random Clifford gates *twirls* errors in a gateset, such that they can be simulated using depolarizing channels.



(Knill et al. 2008 [10/cxz9vm](#); Magesan et al. 2012 [10/tfz](#); Magesan et al. 2012 [10/s8j](#))

# Randomized Benchmarking Example

SMC: interpret survival probability as likelihood. For interleaved case, the lowest-order model is:

$$\Pr(\text{survival}|A, B, \tilde{p}, p_{\text{ref}}; m, \text{mode}) = \begin{cases} Ap_{\text{ref}}^m + B & \text{reference} \\ A(\tilde{p}p_{\text{ref}})^m + B & \text{interleaved} \end{cases}$$

$A, B$ : state preparation and measurement

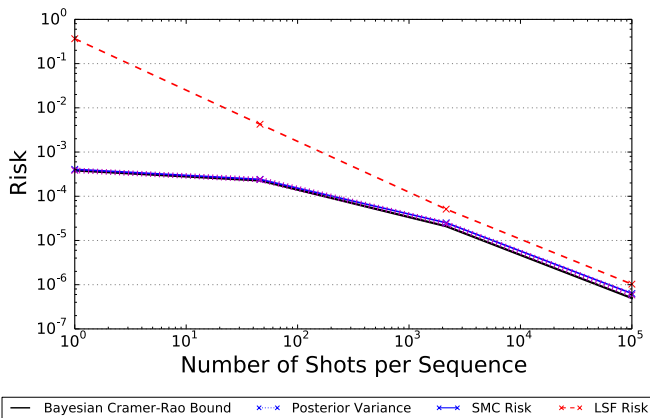
$m$ : sequence length

$p_{\text{ref}}$ : reference depolarizing parameter

$\tilde{p}$ : depolarizing parameter for gate of interest

# Randomized Benchmarking Example

Using SMC, useful conclusions can be reached with significantly less data than with least-squares fitting.



(Granade, Ferrie and Cory 2014 [1404.5275](#))



# Method of Hyperparameters

If “true” model  $\underline{x} \sim \Pr(\underline{x}|\underline{y})$ , for some *hyperparameters*  $\underline{y}$ , can est.  $\underline{y}$  directly:

$$\Pr(d|\underline{y}; \underline{e}) = \int \Pr(d|\underline{x}, \underline{y}; \underline{e}) \Pr(\underline{x}|\underline{y}; \underline{e}) \, d\underline{x}.$$

# Method of Hyperparameters

If “true” model  $\underline{x} \sim \Pr(\underline{x}|\underline{y})$ , for some *hyperparameters*  $\underline{y}$ , can est.  $\underline{y}$  directly:

$$\Pr(d|\underline{y}; \underline{e}) = \int \Pr(d|\underline{x}, \underline{y}; \underline{e}) \Pr(\underline{x}|\underline{y}; \underline{e}) \, d\underline{x}.$$

## Example

For Larmor precession with  $\omega \sim \text{Cauchy}(\omega_0, T_2^{-1})$ ,

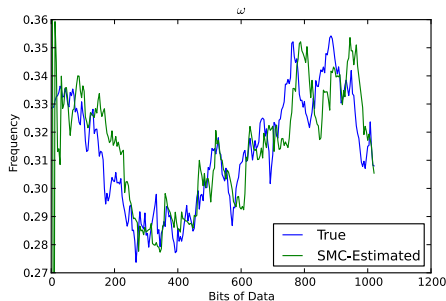
$$\Pr(d|(\omega_0, T_2^{-1}); t) = e^{-tT_2^{-1}} \cos^2(\omega_0 t/2) + (1 - e^{-tT_2^{-1}})/2.$$

Let  $\underline{y} = (\omega_0, T_2^{-1})$ .

# State-Space SMC

Alternatively, can move particles at each timestep  
 $\underline{x}(t_k) \sim \Pr(\underline{x}(t_k) | \underline{x}(t_{k-1}))$ .

This represents *tracking* of a stochastic process.



# Confidence and Credible Regions

Characterizing uncertainty of estimates is critical for many applications:

## Definition (Confidence Region)

$X_\alpha$  is an  $\alpha$ -confidence region if  $\Pr_D(\underline{x}_0 \in X_\alpha(D)) \geq \alpha$ .

## Definition (Credible Region)

$X_\alpha$  is an  $\alpha$ -credible region if  $\Pr_{\underline{x}}(\underline{x} \in X_\alpha | D) \geq \alpha$ .

Credible regions can be calculated from posterior  $\Pr(\underline{x}|D)$  by demanding

$$\int_{X_\alpha} d\Pr(\underline{x}|D) \geq \alpha.$$

# High Posterior Density

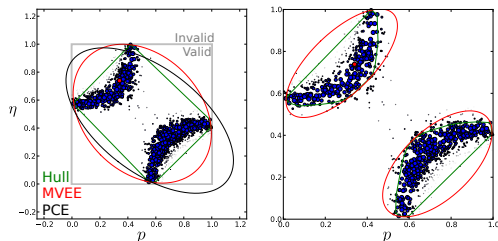
Want credible regions that are *small* (most powerful).

- Posterior covariance ellipses (PCE)— good for approximately normal posteriors
- Convex hull— very general, but verbose description
- Minimum volume enclosing ellipses (MVEE)— good approximation to hull

(Granade et al. 2012 [10/s87](#); Ferrie 2014 [10/tb4](#))

# Comparison of HPD Estimators

For multimodal distributions, clustering algorithms can be used to exclude regions of small support. For a noisy coin model (heads probability  $p$ , visibility  $\eta$ ):

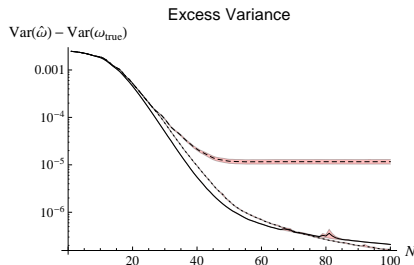


Left, no clustering. Right, DBSCAN.

Plot courtesy of Chris Ferrie. (Ferrie 2014 [10/tb4](#))

# Hyperparameters and Region Estimation

In some hyperparameter models, can also express as region estimator on underlying parameters.



**Figure :** Larmor precession model w/  $\omega \sim N(\mu, \sigma^2)$ , three exp. design strategies

Critically, the covariance region for  $\omega$  is not smaller than the true covariance given by the hyperparameter  $\sigma^2$ .

(Granade et al. 2012 [10/s87](#))

# Bayes Factors and Model Selection

In SMC update  $w_i \mapsto w_i \times \Pr(d|\underline{x}; \underline{e})/\mathcal{N}$ ,

$$\mathcal{N} \approx \Pr(d|\underline{e}).$$

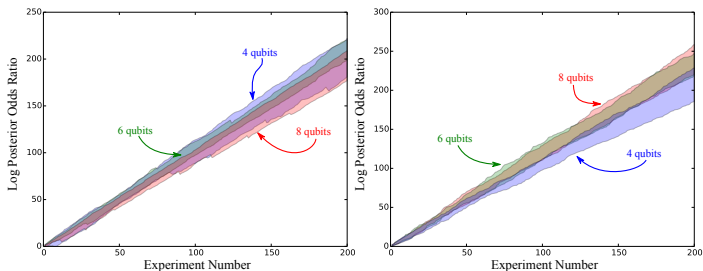
Running SMC updaters for distinct models  $A$  and  $B$ , collecting normalizations  $\mathcal{N}_A$  and  $\mathcal{N}_B$  at each step gives

$$\text{BF} = \frac{\mathcal{N}_A}{\mathcal{N}_B} \approx \frac{\Pr(d|A; \underline{e})}{\Pr(d|B; \underline{e})}$$

For full data record, can multiply normalization records to select  $A$  versus  $B$ .



For example, deciding between linear- (left) and complete-graph (right) Ising models:



(Wiebe, Granade, Ferrie and Cory 2014 [10/tdk](#))

Main cost to SMC: simulation calls.  $n$  each Bayes update.

Simulation and learning are intimately connected: if we can simulate, then we can learn.

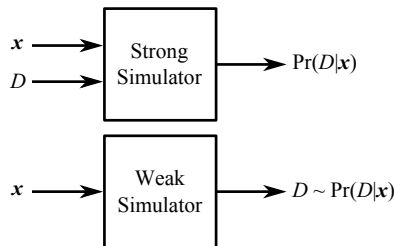
Main cost to SMC: simulation calls.  $n$  each Bayes update.

Simulation and learning are intimately connected: if we can simulate, then we can learn.

## Big Idea

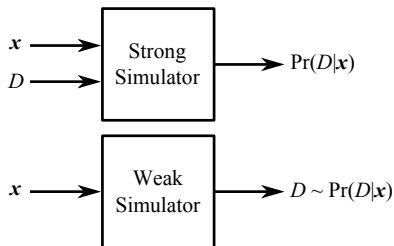
Use quantum simulation to learn about unknown quantum systems.

# Weak and Strong Simulation



(Ferrie and Granade 2014 [10/tdj](#))

# Weak and Strong Simulation



Quantum simulation produces data, not likelihoods. Must sample to estimate likelihood.

# Adaptive Likelihood Estimation

## Solution

Treat estimating the likelihood as a secondary estimation problem.

(Ferrie and Blume-Kohout 2012 [10/tf2](#), Ferrie and Granade 2014 [10/tdj](#))

# Adaptive Likelihood Estimation

## Solution

Treat estimating the likelihood as a secondary estimation problem.

2-outcome model: hedged binomial estimator finds the probability  $p_0$  of a “0” outcome by repeatedly sampling a weak simulator.

(Ferrie and Blume-Kohout 2012 [10/tf2](#), Ferrie and Granade 2014 [10/tdj](#))

# Adaptive Likelihood Estimation

## Solution

Treat estimating the likelihood as a secondary estimation problem.

2-outcome model: hedged binomial estimator finds the probability  $p_0$  of a “0” outcome by repeatedly sampling a weak simulator.

Variance well-known, so collect until a fixed *tolerance* is reached.

We will show that SMC is robust to likelihood estimation errors.

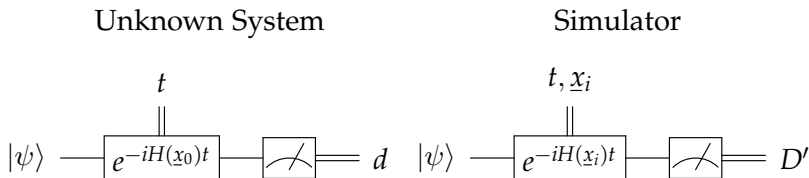


# Quantum Likelihood Evaluation

First approach: compare classical outcomes of unknown and trusted quantum systems.

Evolve state  $|\psi\rangle$  for time  $t$  then measure, getting  $d$ .

For each particle  $\underline{x}_i$ , repeatedly sample from quantum simulation of  $e^{-iH(\underline{x}_i)t}$ , getting  $D'$ .



Estimated likelihood  $\hat{\ell}_i := |\{d' \in D' | d' = d\}|$ . SMC update:

$$w_i \mapsto w_i \hat{\ell}_i / \sum_i w_i \hat{\ell}_i.$$

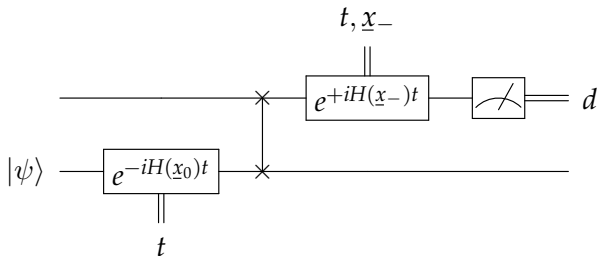
(Wiebe, Granade, Ferrie and Cory 2014 [10/tf3](#))

QLE can work, but as  $t \rightarrow \infty$ ,  $\Pr(d|\underline{x}; t)$  equilibrates. Thus,  $t \geq t_{\text{eq}}$  is uninformative.

By CRB, error then scales as  $O(1/Nt_{\text{eq}}^2)$ .

# Interactive QLE

Solution: couple unknown system to a quantum simulator, then invert evolution by hypothesis  $\underline{x}_-$ .



## Echo

If  $\underline{x}_- \approx \underline{x}_0$ , then  $|\langle \psi | e^{-it(H(\underline{x}_0) - H(\underline{x}_-))} | \psi \rangle|^2 \approx 1$ .

(Wiebe, Granade, Ferrie and Cory 2014 10/13)

# Alternate Interpretation

QHL finds  $\hat{\underline{x}}$  such that  $H(\hat{\underline{x}})$  most closely approximates “unknown” system  $H_0$ .

Gives an  $\alpha$ -credible bound on error introduced by replacing  $H_0 \rightarrow H(\hat{\underline{x}})$ .

# Posterior Guess Heuristic

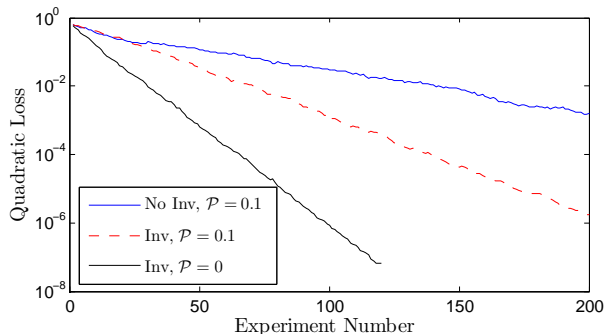
Inversion connects the model and experiment spaces. Use to come up with a heuristic for experiment designs.

- Choose  $\underline{x}_e, \underline{x}'_e \sim \Pr(\underline{x})$ , the most recent posterior.
- Choose  $t = 1/\|\underline{x}_e - \underline{x}'_e\|$ .
- Return  $\underline{e} = (\underline{x}_e, t)$ .

# Ising Model on Spin Chains

Hamiltonian: nearest-neighbor Ising models on a chain of nine qubits.

Interactivity allows for dramatic improvements over QLE.

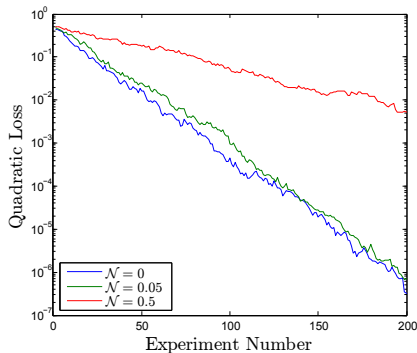


$\mathcal{P}$ : adaptive likelihood estimation tolerance.

(Wiebe, Granade, Ferrie and Cory 2014 10/f3)

# Ising Model on the Complete Graph

With IQLE, can also learn on complete interaction graphs. We show the performance as a function of the depolarization strength  $\mathcal{N}$ .

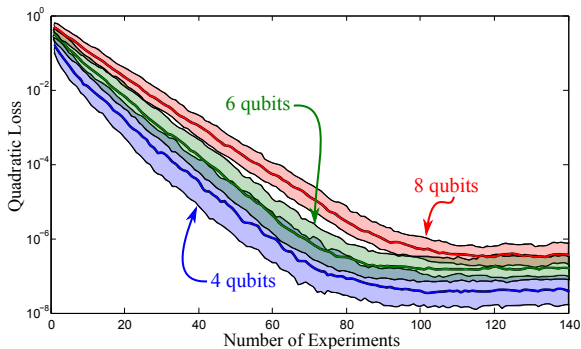


$\mathcal{N}$ : depolarizing noise following SWAP gate.

(Wiebe, Granade, Ferrie and Cory 2014 [10/tdk](#))

# Ising Model with the Wrong Graph

Simulate with spin chains, suppose “true” system is complete, with non-NN couplings  $O(10^{-4})$ .



(Wiebe, Granade, Ferrie and Cory 2014 [10/tdk](#))



# Scaling Parameter

$\dim \underline{x}$ , not  $\dim \mathcal{H}$ , determines scaling of IQLE.

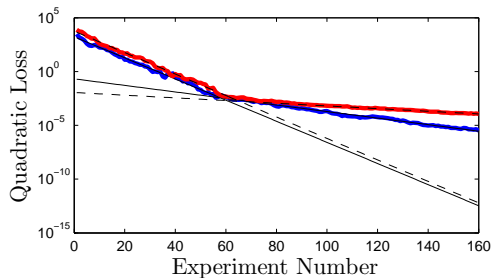


Figure : 4 qubit (red) and 6 qubit (blue) complete graph IQLE

(Wiebe, Granade, Ferrie and Cory 2014 [10/tf3](#))

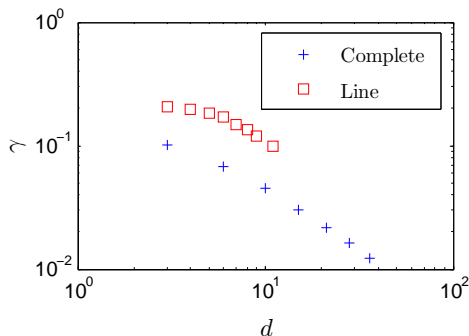
# Scaling and Dimensionality

In both the spin-chain and complete graph cases, the quadratic loss on average decays exponentially,  $L_Q \propto e^{-\gamma N}$ , for some rate constant  $\gamma$ .

(Wiebe, *Granade*, Ferrie and Cory 2014 [10/tf3](#))

# Scaling and Dimensionality

In both the spin-chain and complete graph cases, the quadratic loss on average decays exponentially,  $L_Q \propto e^{-\gamma N}$ , for some rate constant  $\gamma$ . Consider  $\gamma = \gamma(\dim \underline{x})$ :



This suggests that, with access to a quantum simulator, learning *may* scale efficiently.

(Wiebe, Granade, Ferrie and Cory 2014 [10/tf3](#))

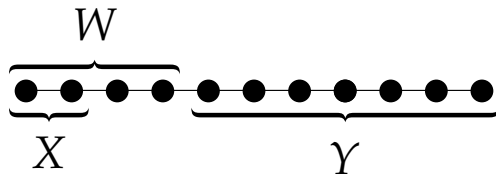
## SMC + IQLE:

- Possibly scalable with quantum resources.
- Robust to finite sampling.
- Robust to approximate models.

Still requires simulator be at least as large as system of interest.

# Information Locality

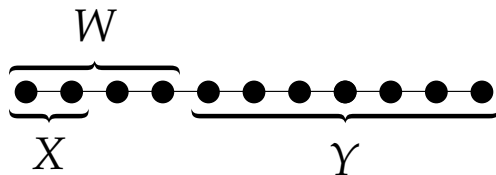
To go further, we want to *localize* our experiment, such that we can simulate on a smaller system.



Measure on  $X$ , simulate on  $W$ , and ignore all terms with support over  $Y$ .

# Information Locality

To go further, we want to *localize* our experiment, such that we can simulate on a smaller system.

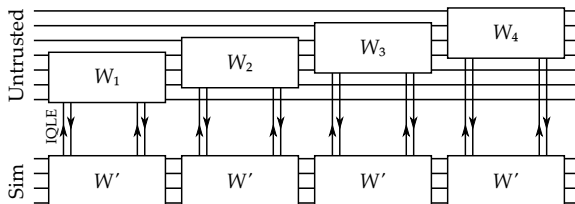


Measure on  $X$ , simulate on  $W$ , and ignore all terms with support over  $Y$ .

Gives *approximate* model that can be used to learn Hamiltonian restricted to  $X$ .

# Local and Global Particle Clouds

To reconstruct the entire system, we need to combine data from different partitions.



Separate out one partition  $W_k$  at a time, maintain a *global* cloud of particles.

# Local and Global Particle Clouds

Initialize  $\{\underline{x}_i\}$  over entire system. Then, for each simulated subregister  $W_k$ :

- 1 Make “local” particle cloud  $\{\underline{x}_i|_{W_k}\}$  by slicing  $\{\underline{x}_i\}$ .
- 2 Run SMC+IQLE with  $\{\underline{x}_i|_{W_k}\}$  as a prior.
- 3 Ensure that the final “local” cloud has been resampled (has equal weights).
- 4 Overwrite parameters in “global” cloud  $\{\underline{x}_i\}$  corresponding to post-resampling  $\{\underline{x}_i|_{W_k}\}$ .

In this way, all parameters are updated by an SMC run.



## Q50 Example

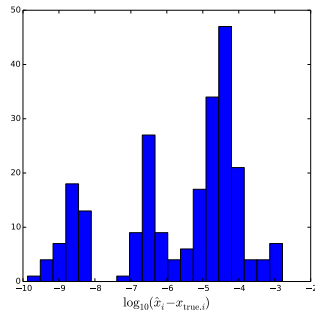
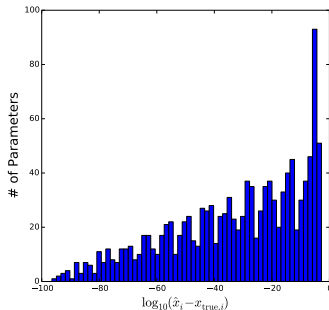
Goal: characterize a 50-qubit Ising model (complete graph) with unknown ZZ couplings.

All Hamiltonian terms commute, but initial state doesn't. Let  $A_X$  be observable,  $A_{X'}$  be sim. observable.

$$\begin{aligned}\|A_X(t) - A_{X'}(t)\| &\leq \|A_X(t)\| (e^{2\|H|_Y\|t} - 1) \\ \Rightarrow t &\leq \ln \left( \frac{\delta}{\|A_X(t)\|} + 1 \right) (2\|H|_Y\|)^{-1},\end{aligned}$$

where  $\delta$  is the tolerable likelihood error.

# Example Q50 Run



$|X_k| = 4$ ,  $|W_k| = 8$ ,  $n = 20,000$ ,  $N = 500$ , exp. decaying interactions.

NB: 1225 parameter model,  $L_2$  error of 0.3%.

# Lieb-Robinson Bounds

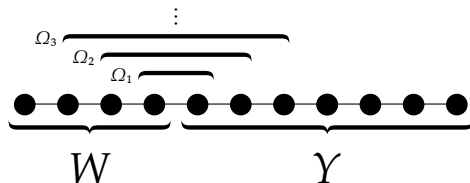
More generally, for  $[H|_W, H_Y] \neq 0$ , use *Lieb-Robinson bound*.  
 If interactions between  $X$  and  $Y$  decay sufficiently quickly, then there exists  $C, \mu$  and  $v$  s. t. for any observables  $A_X(t), B_Y$ :

$$\|[A_X(t), B_Y]\| \leq C\|A_X(t)\|\|B_Y\|\|X\|\|Y\|(e^{v|t|} - 1)e^{-\mu d(X,Y)}$$

This *guarantees* that error due to truncation is bounded if we choose small  $t$ .

# Lieb-Robinson Bounds

Can find bound in terms of Hamiltonian by considering  $H$  site-by-site.

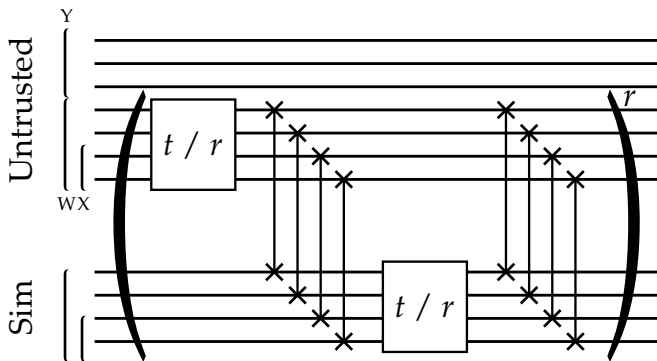


Let  $H_j$  be the Hamiltonian term containing distance- $j$  interactions between  $W$  and  $Y$ , acting on sites  $\Omega_j$ .

$$\|A(t) - e^{iH|w|t} A e^{-iH|w|t}\| \leq \sum_j C \|A\| \|H_j\| |X| |\Omega_j| e^{-\mu j} (e^{v|t|} - 1)$$

# Trotterization

Can improve the Lieb-Robinson bound by “shaking” between simulator and system. Using  $r \approx vt$  SWAP gates, error is  $O(t)$ .



- Bayesian inference: simulation as a characterization/validation resource.

- Bayesian inference: simulation as a characterization/validation resource.
- Sequential Monte Carlo: numerical algorithm for inference.

- Bayesian inference: simulation as a characterization/validation resource.
- Sequential Monte Carlo: numerical algorithm for inference.
- Robust to many practical concerns.



- Bayesian inference: simulation as a characterization/validation resource.
- Sequential Monte Carlo: numerical algorithm for inference.
- Robust to many practical concerns.
- Can use quantum simulation to offer potential scaling.

- Bayesian inference: simulation as a characterization/validation resource.
- Sequential Monte Carlo: numerical algorithm for inference.
- Robust to many practical concerns.
- Can use quantum simulation to offer potential scaling.
- Using robustness of SMC, can truncate simulation → bootstrapping.

## Further Information

Slides, a journal reference for this work, a full bibliography and a software implementation can be found at [\*http://www.cgranade.com/research/talks/msr-2014/\*](http://www.cgranade.com/research/talks/msr-2014/).



Thank you for your kind attention!