SAND2022-2313C

Sandia National Laboratories

Out-of-Model Effects and Overdispersion in Gate Set Tomography

COREY OSTROVE*, ERIK NIELSEN*, KEVIN YOUNG*, ROBIN BLUME-KOHOUT*

*SANDIA NATIONAL LABORATORIES

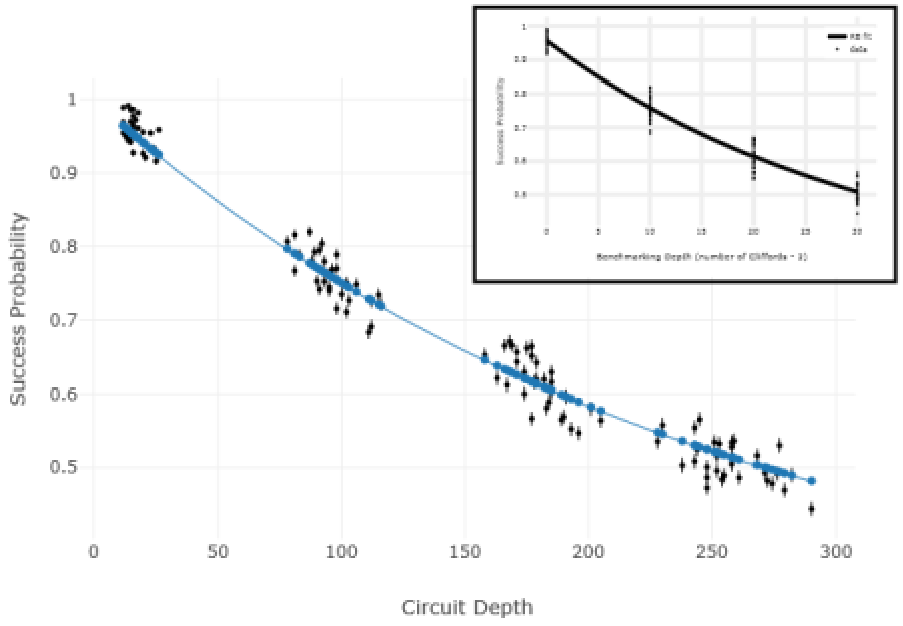
MARCH MEETING 2022

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

A Problem With QCVV

- In real QCVV experiments it is the norm to find that our models fail to fully predict the data they are fit from.
- Sometimes this is expected. As when simplifying assumptions intentionally neglect some effects which may turn out to have been important
- In practice this happens even when fitting the largest class of Markovian models supported by our methods, indicating non-Markovianity.
- This problem is especially relevant to gate set tomography (GST), but also arises in essentially all model-based QCVV protocols.



Fitting a depolarizing channel to RB data simulated using a gate set with a mix of stochastic and coherent errors. The model is clearly incomplete and as such there is substantial dispersion about the predictions.

E. Nielsen *et al. New J. Phys.* **23** 093020 (2021)

Model Violation

- How can we know a model is violated? We can retroactively inspect the quality of the fit by asking: Does this model fit the data as well as we'd expect given solely finite-sample fluctuations?
- Quantitatively, we use log-likelihood ratio statistics and Wilks' Theorem:

Theorem. *Let $\mathcal{L}$ be the likelihood an estimated model produced a data set and $\mathcal{L}_{max}$ the likelihood of a saturated model which fits the data exactly. If the estimated model is valid the log-likelihood ratio statistic will be χ_k^2 distributed,*

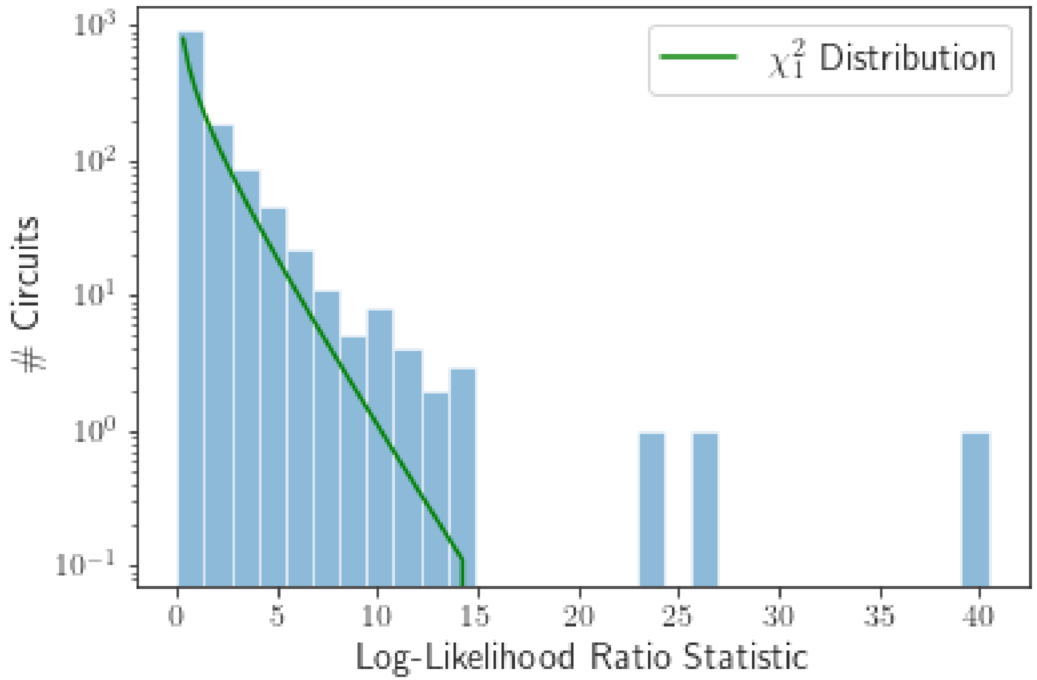
$$LLRS = 2(\log(\mathcal{L}_{max}) - \log(\mathcal{L})) \sim \chi_k^2$$

where k is the difference in the number of paramaters between the models.

- For a 2-outcome measurement this implies that the per-circuit contributions to the log-likelihood ratio should be approximately χ_1^2 distributed.

Model Violation

- How can we know a model is violated? We can retroactively inspect the quality of the fit by asking: Does this model fit the data as well as we'd expect given solely finite-sample fluctuations?

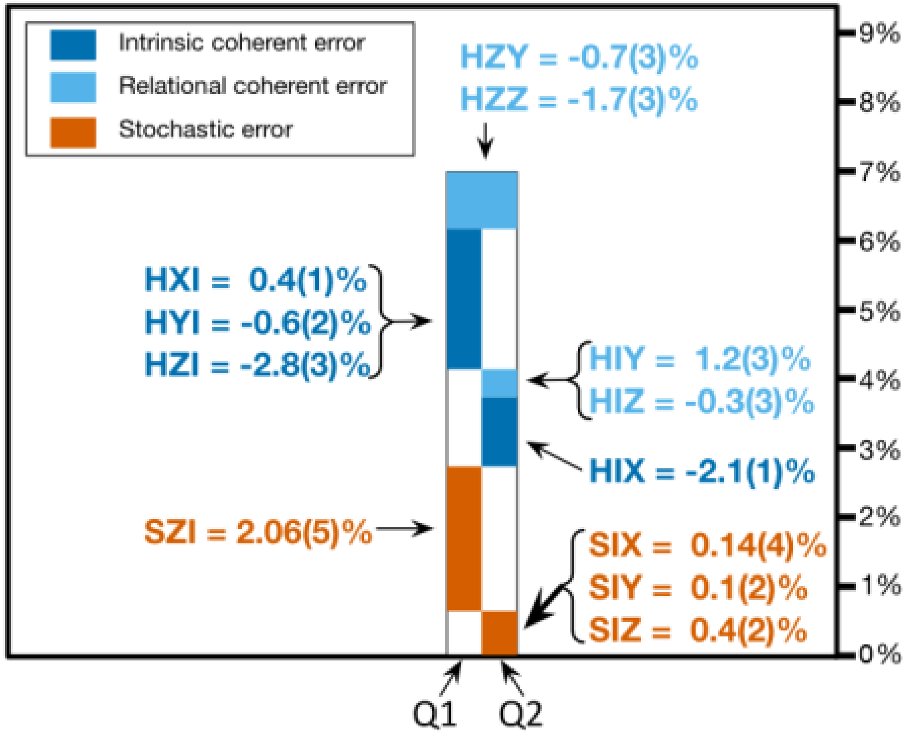


Distribution of the per-circuit log-likelihood ratio contributions for a model fit using the largest class of Markovian models supported by GST on experimental data from the SNL trapped ion group.

Another way of phrasing this problem: The data are further from the predictions than we expect. i.e. they are *overdispersed.*

Why Does This Matter?

- There is an inherent tension between the high levels of precision reported by our model parameter estimates and the fact that we know these models can't be fully correct.
- One may argue: we never claimed the models were literally *correct*, taking the models too seriously is user error. If you take these methods off-label you can't be surprised that there are side-effects…
- In the real world end-users *do* take the results of modeling seriously and use insights derived models to evaluate device performance and imperfections.



Estimated error rates for an $R_x(\frac{\pi}{2})$ gate on a silicon donor system obtained using GST show unambiguous signatures of anomalous 2-qubit entangling ZZ interactions. This led to the discovery of unexpected physics in the device.

Mądzik, M.T., *et al.* Precision tomography of a three-qubit donor quantum processor in silicon. *Nature* **601,** 348–353 (2022)

What Should We Do About Model Violation?

- We propose the following:
  - Let's promote the overdispersion in our data from simply being an indication of the failure of our models to a quantity which we can statistically model in and of itself.
- What does modeling the overdispersion directly achieve?
  - We will create, by construction, a model consistent with the data, even with data inconsistent with *any* non-overdispersed Markovian model.
  - Even a single overdispersion parameter will be sufficient in practice to restore consistency.
  - It naturally inflates the confidence intervals for model parameter estimates, more honestly accounting for uncertainty.

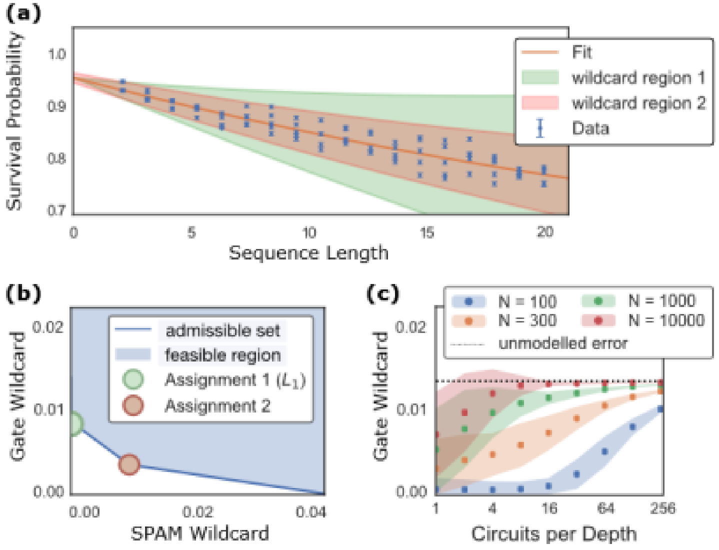
How Do We Model Overdispersion?

<u>Prior Art:</u>

•<u>Wildcard Error:</u> arXiv:2012.12231, R. Blume-Kohout, K. Rudinger, E. Nielsen, T. Proctor, and K. Young

•Ad-hoc and inconsistent with standard tools for statistical model analysis.

•Still useful as a heuristic for unmodelled effects.

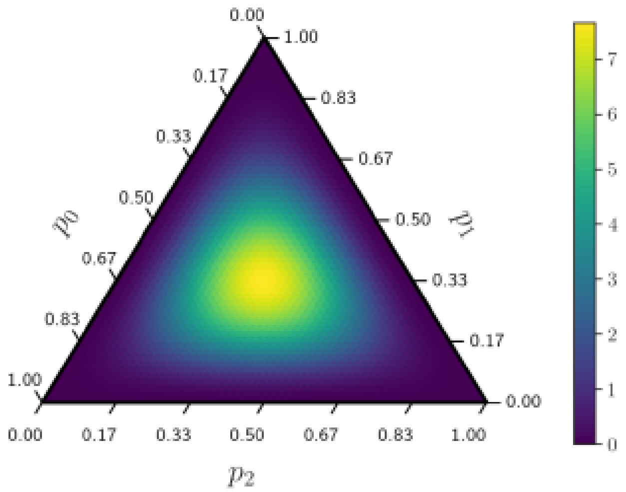


- <u>Current Proposal:</u> Systematically weaken the predictions of our model by, instead of reporting a probability distribution over circuit outcomes, reporting a distribution *over* probability distributions for each circuit.
- The width of the distribution over distributions is parameterized by an *overdispersion parameter* fit from the data.
- A set of overdispersion parameters with a set of rules for associating circuits with them is an *overdispersion model.*

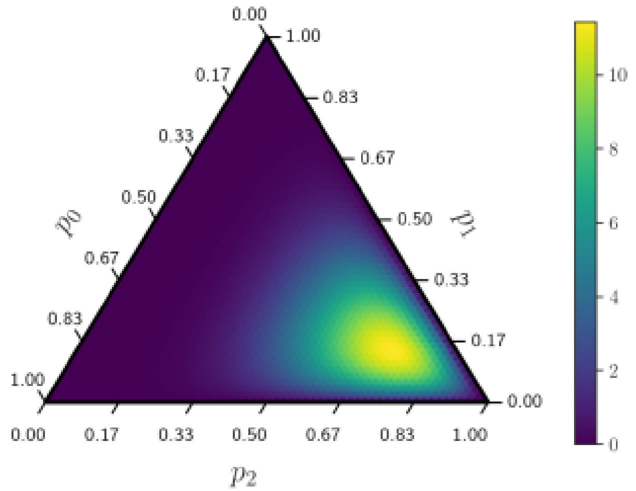
How Do We Model Overdispersion?

- The Dirichlet-multinomial ansatz is adopted for the generation of overdispersed count data.
- Operationally this is broken into two steps:

1. Sample a probability vector from a Dirichlet distribution centered on the base model's prediction.
2. Sample counts from a multinomial distribution according to this probability vector.

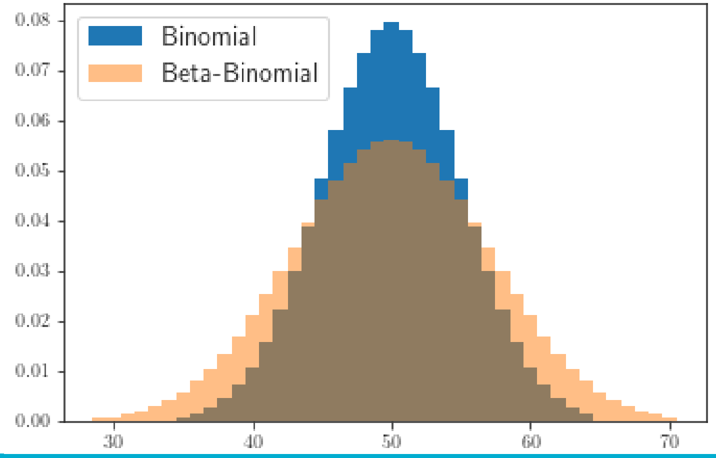


Dirichlet distribution with $E[p_0] = [p_1] = E[p_2] = \frac{1}{3}$ $\alpha_0 = 10$



Dirichlet distribution with $E[p_0] = .6, E[p_1] = E[p_2] = .2$ $\alpha_0 = 10$

Comparision of binomial distribution with $n = 100, p = .5$ and beta-binomial distribution with $n = 100, E[p] = .5, \alpha_0 = 100$



The Dirichlet-Multinomial Ansatz

- The likelihood a set of counts $\{n_0, \cdots, n_{K-1}\}$ for a circuit c is generated by an overdispersed model is:

$$\mathcal{L}_c = \frac{\Gamma(\frac{1}{\phi_{0,c}})\Gamma(n+1)}{\Gamma(n+\frac{1}{\phi_{0,c}})} \prod_{k=0}^{K-1} \frac{\Gamma(n_k + \frac{1}{\phi_{0,c}} E[p_k])}{\Gamma(\frac{1}{\phi_{0,c}} E[p_k])\Gamma(n_k+1))}$$

$$\phi_{0,c} = \text{Overdispersion Parameter}$$

- In practice it is more convenient to work with the log-likelihood:

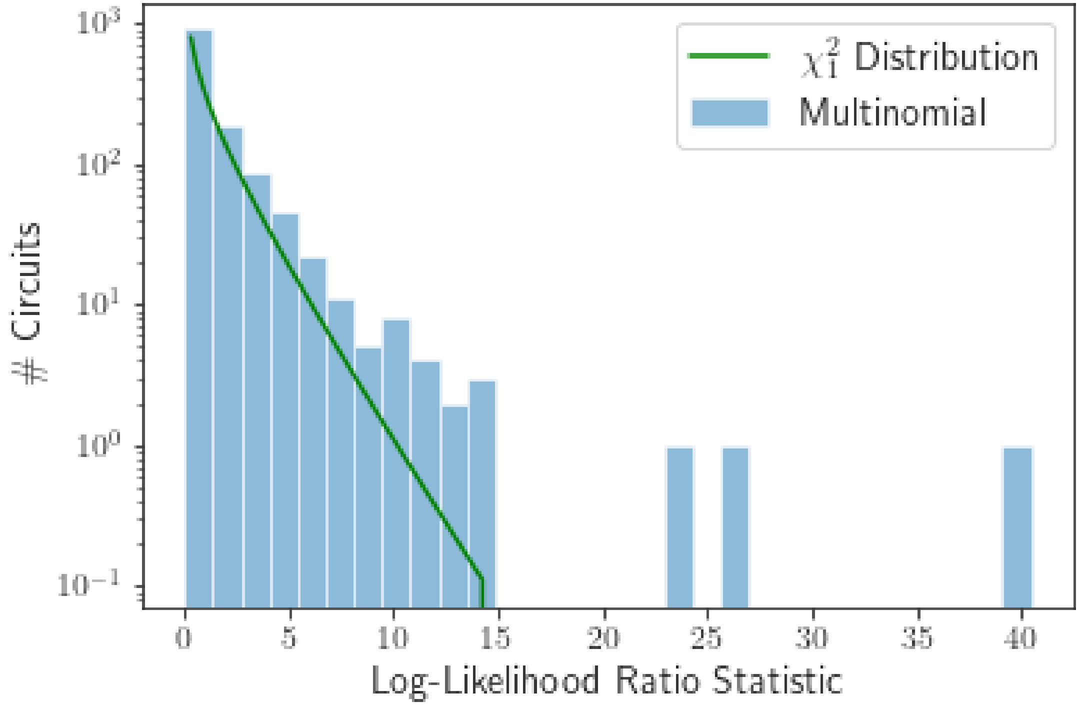
$$l_c = \log\left(\Gamma\left(\frac{1}{\phi_{0,c}}\right)\right) + \log\left(\Gamma(n+1)\right) - \log\left(\Gamma\left(n + \frac{1}{\phi_{0,c}}\right)\right) + \sum_{k=0}^{K-1} \log\left(\Gamma\left(n_k + \frac{1}{\phi_{0,c}} E[p_k]\right)\right) - \log\left(\Gamma\left(\frac{1}{\phi_{0,c}} E[p_k]\right)\right) - \log\left(\Gamma(n_k+1)\right)$$

- The overdispersion parameters are fit using MLE over the total log-likelihood for a set of circuits $\mathcal{C}$

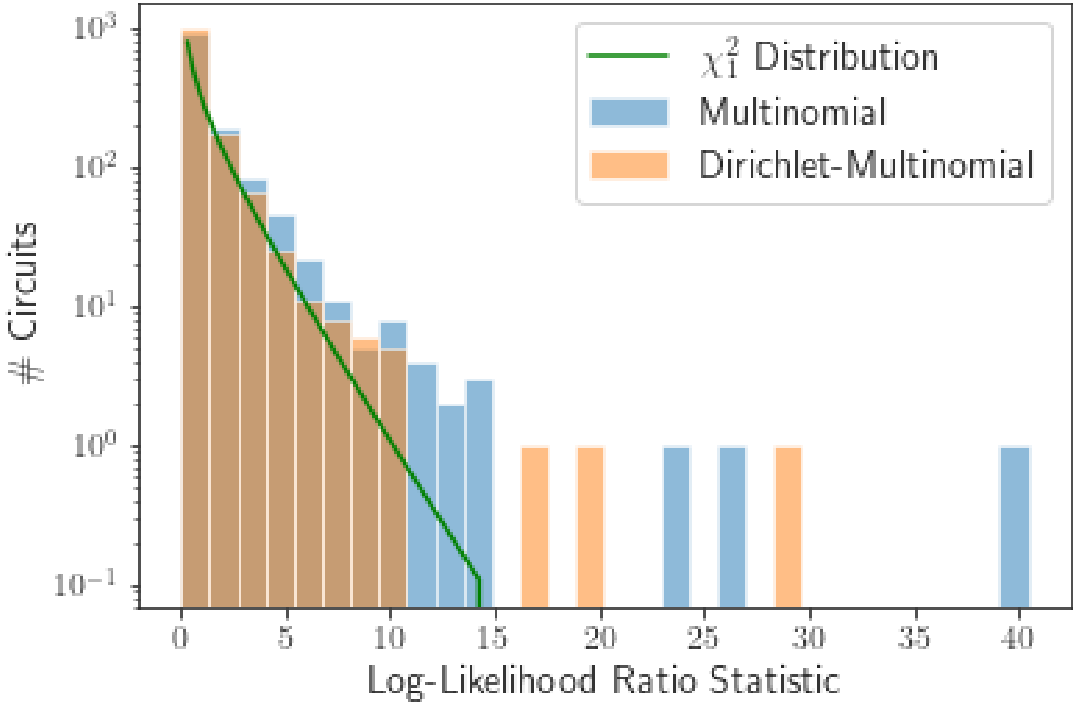
$$l_{\mathcal{C}} = \sum_{c\in\mathcal{C}} \left[\log\left(\Gamma\left(\frac{1}{\phi_{0,c}}\right)\right) + \log\left(\Gamma(n+1)\right) - \log\left(\Gamma\left(n + \frac{1}{\phi_{0,c}}\right)\right) + \sum_{k=0}^{K-1} \log\left(\Gamma\left(n_k + \frac{1}{\phi_{0,c}} E[p_k]\right)\right) - \log\left(\Gamma\left(\frac{1}{\phi_{0,c}} E[p_k]\right)\right) - \log\left(\Gamma(n_k+1)\right) \right]$$

Single Parameter Overdispersion Model

The simplest model of overdispersion is the single parameter model where the overdispersion parameters, $\phi_{0,c}$, are equal for every circuit.



Per-circuit log-likelihood ratio distribution under the standard multinomial ansatz.



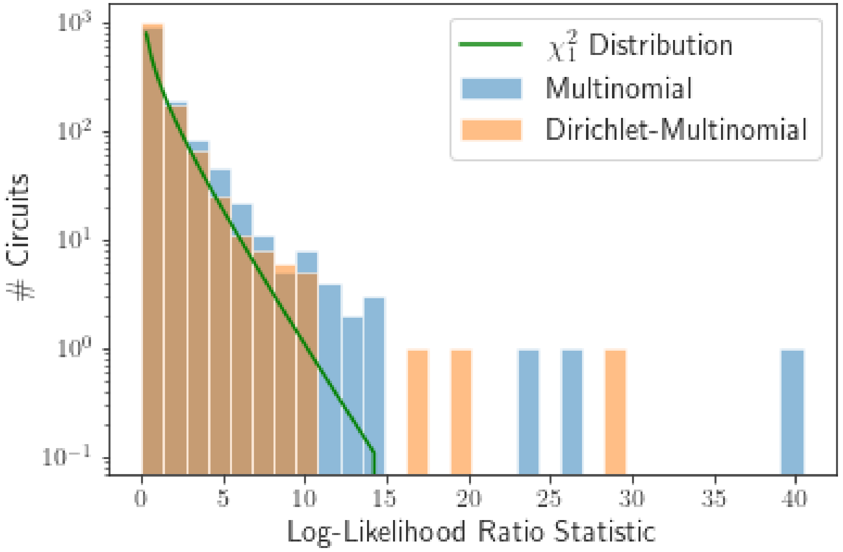
Distribution with a single overdispersion parameter.

Total log-likelihood: $l_{\mathcal{C}} = -8111$ $\Delta l_{\mathcal{C}} = 60$ Total log-likelihood: $l_{\mathcal{C}} = -8051$

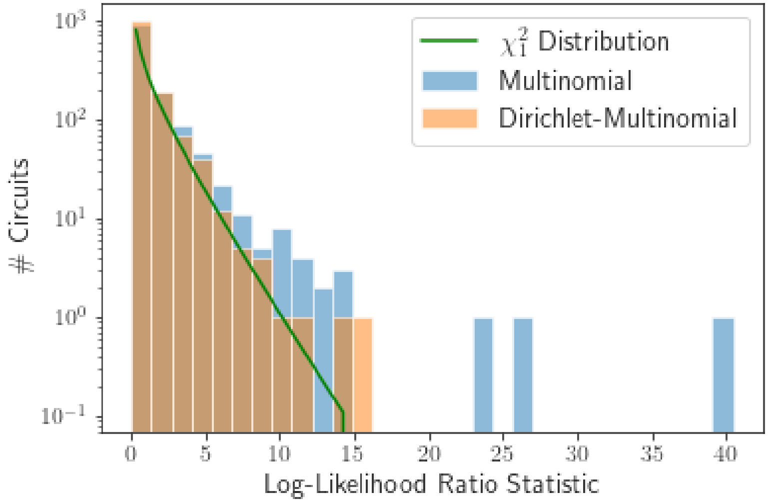
Total log-likelihood ratio is consistent with Wilks' Theorem.

More Sophisticated Overdispersion Models

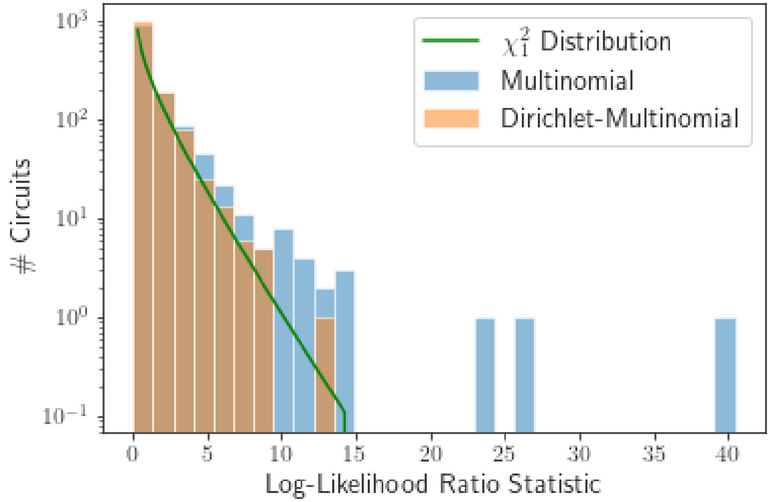
- Given the success of the single overdispersion parameter model, what is the value in more sophisticated multi-parameter models?



The single parameter model resolves consistency issues with the total log-likelihood, but there are still problems with the per-circuit distribution.



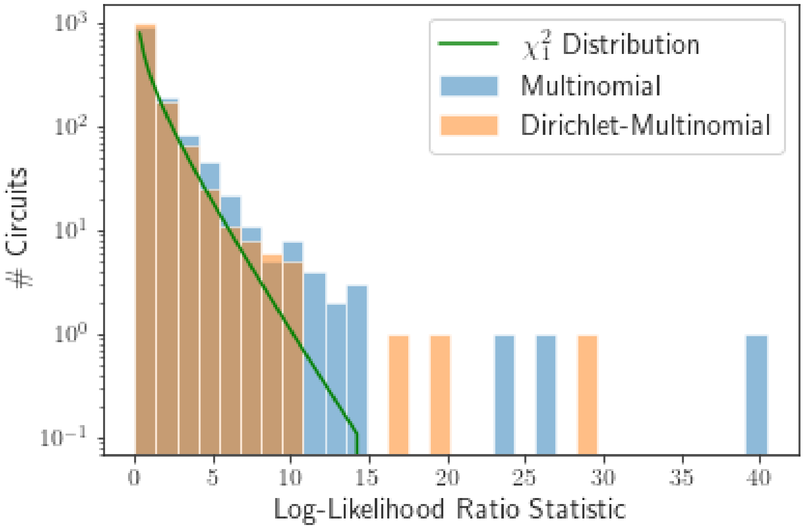
The “per-iteration” model course-grains the circuit depth and assigns a different overdispersion parameter to each.



The “per-depth” model overdispersion parameter to each depth (sans course-graining).

More Sophisticated Overdispersion Models

Single Parameter



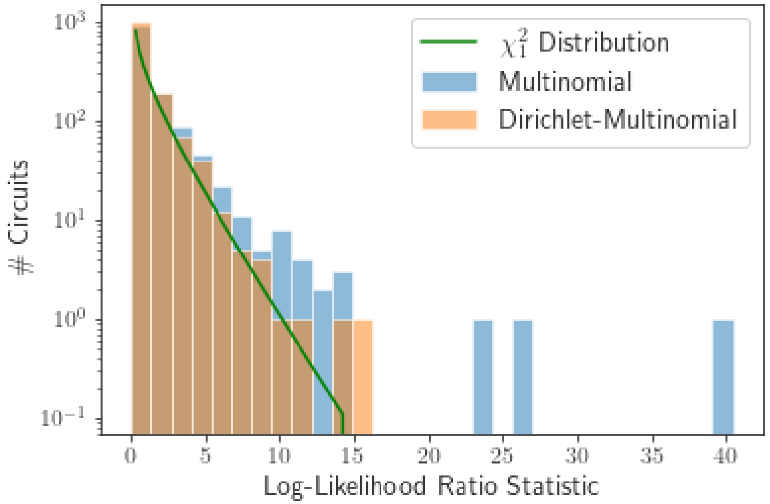
Total log-likelihood: $l_{\mathcal{C}} = -8051$

$$\Delta l_{\mathcal{C}} = 60$$

Additional Parameters: 1

$$\gamma = 60$$

Per-Iteration



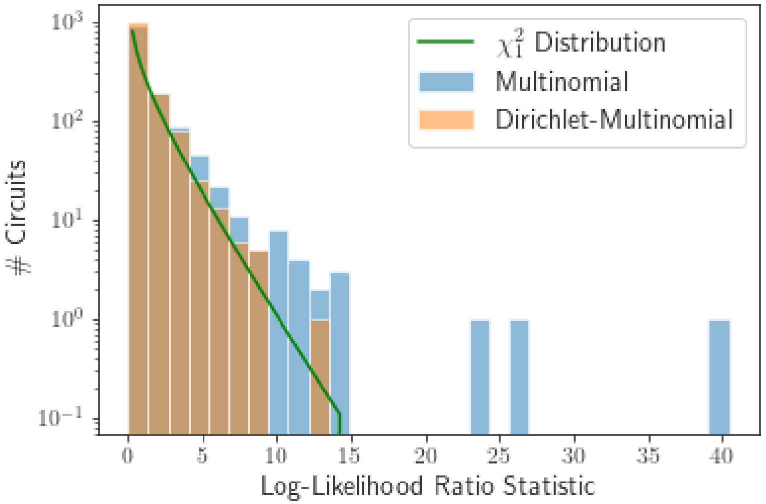
Total log-likelihood: $l_{\mathcal{C}} = -7995$

$$\Delta l_{\mathcal{C}} = 116$$

Additional Parameters: 9

$$\gamma = 12.9$$

Per-Depth



Total log-likelihood: $l_{\mathcal{C}} = -7960$

$$\Delta l_{\mathcal{C}} = 151$$

Additional Parameters: 57

$$\gamma = 2.6$$

Evidence Ratio: $\gamma = \frac{2\Delta l_{\mathcal{C}}}{\Delta k}$, where Δk is the change in the number of parameters.

Summary

- In real QCVV experiments it is the norm to find that our models fail to predict the data they are fit from as well as we'd expect.
- There is an inherent tension between the high levels of precision we can achieve in our model parameter estimates and the fact that we know these models can't be fully correct.
- Overdispersion models allow us to systematically relax the predictions of our models in order to restore consistency with the experimental data.
- This can be achieved at the cost of very few additional parameters in practice.
- Overdispersion models are compatible with the standard tools of statistical model analysis and model selection, and provide a new tool for the quantification of unmodelled effects in QCVV analysis.