SAND2021-1099C

Non-Gaussian noise characterization with factorial hidden Markov models

Quantum Computing Program Review

February 2, 2021

Dylan Albrecht

PI: Toby Jacobson

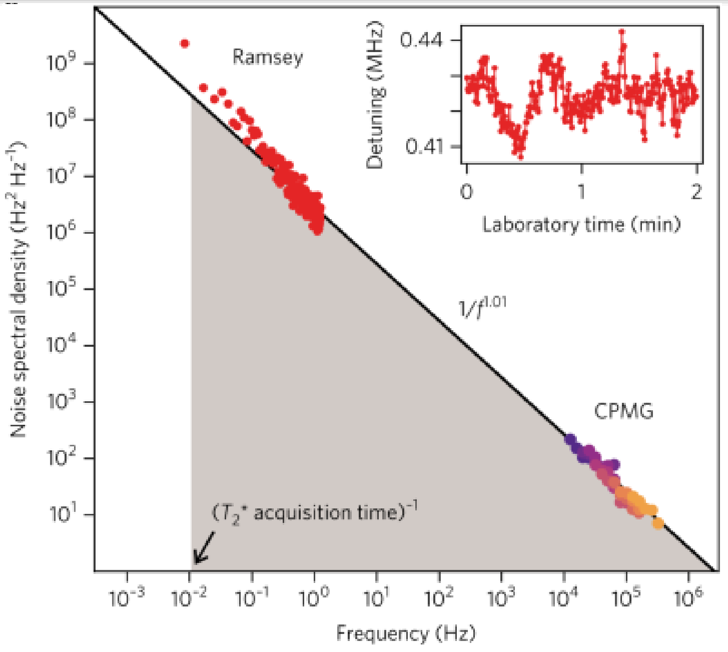
Center for Computing Research

Sandia National Laboratories

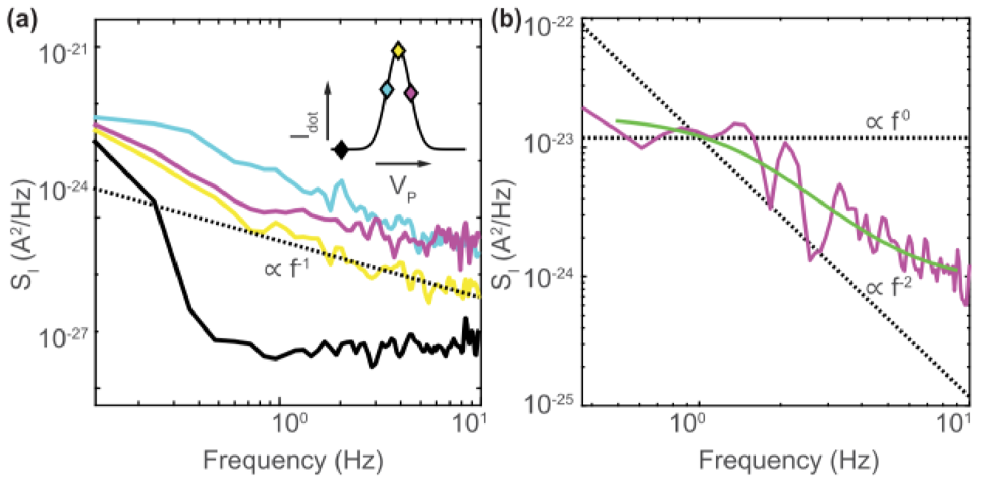
Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. SAND2021-nnnn X

Motivation and problem statement

- Low-frequency (LF) noise is a nuisance for qubit operation
    - Driver of qubit drift and non-Markovianity
- The power spectral density (PSD) is a standard noise metric, but incompletely characterizes non-Gaussian noise
- Given a measured noise timeseries, what else can we infer about it?
    - Interested in a more “big data”-inspired approach

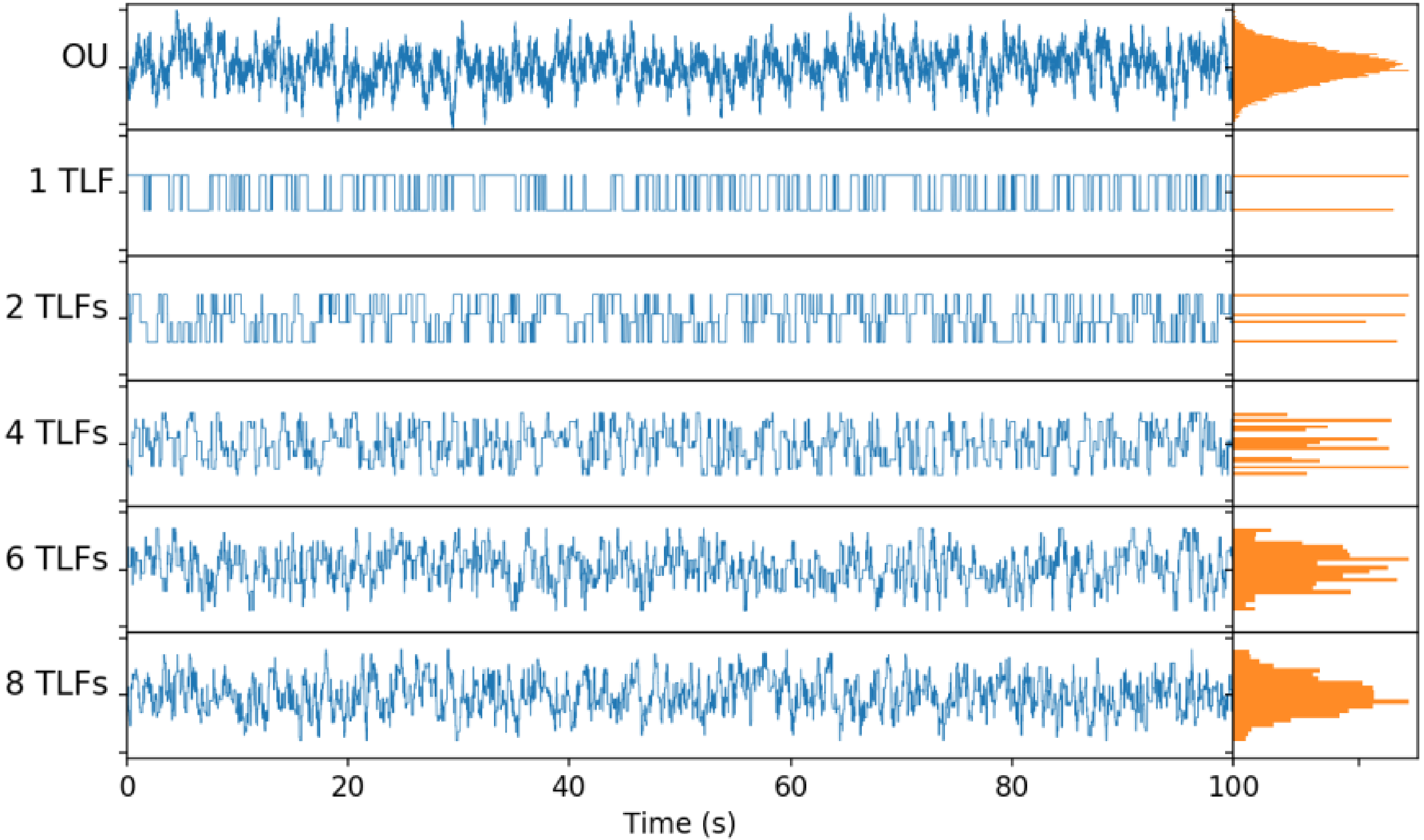


[Yoneda, et al. Nat. Nano. (2018)]



[Connors, et al. PRB (2019)]

The PSD only partially characterizes non-Gaussian noise

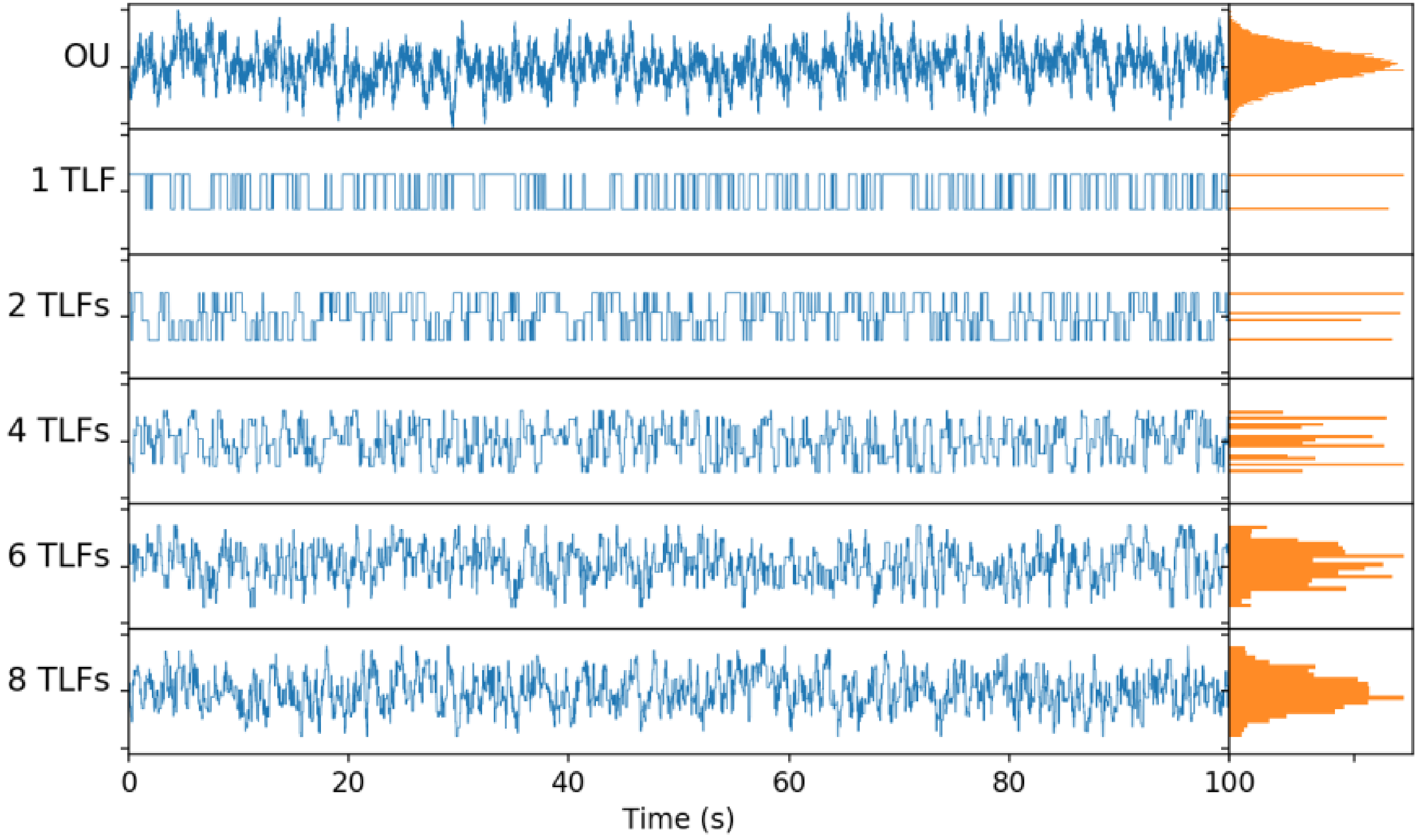


OU = Ornstein-Uhlenbeck process

TLF = Two-level fluctuator

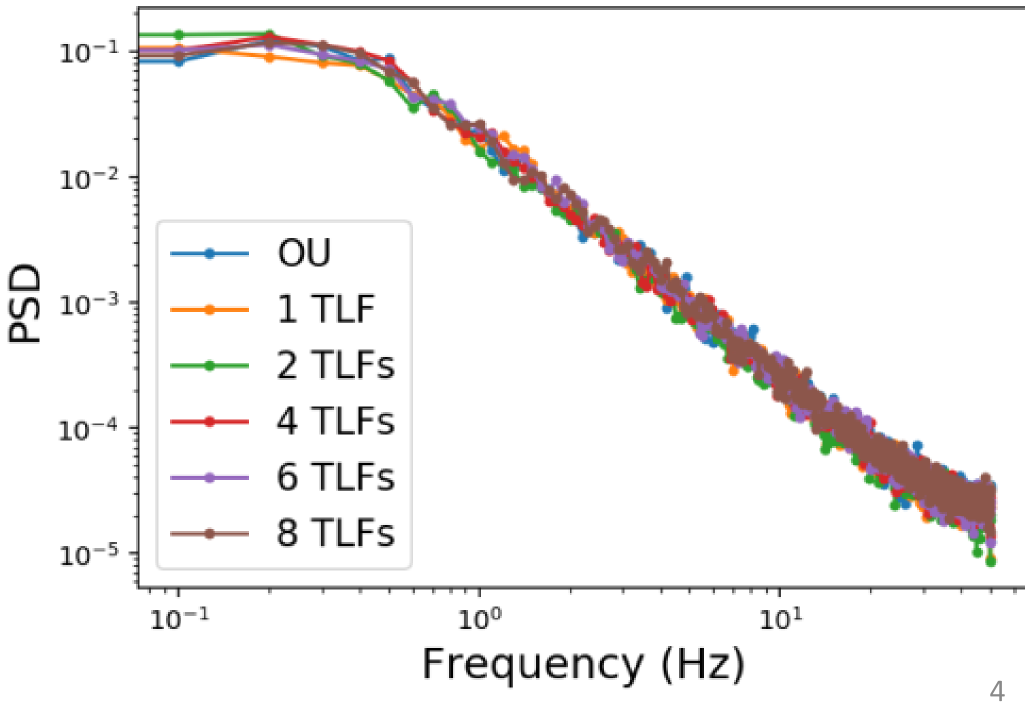
Sandia National Laboratories

The PSD only partially characterizes non-Gaussian noise



OU = Ornstein-Uhlenbeck process
TLF = Two-level fluctuator

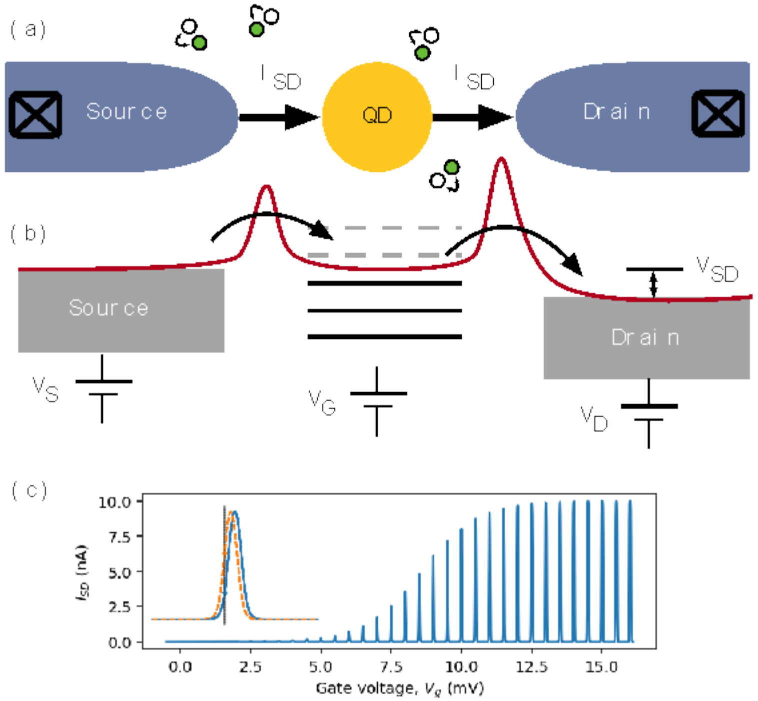
Same (Lorentzian) power spectra



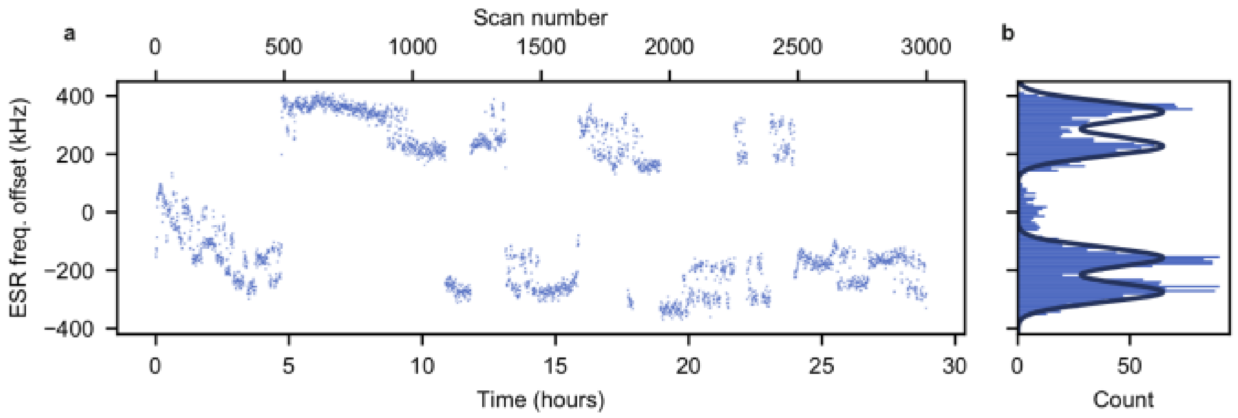
- Few TLFs are needed before histogram looks Gaussian
- Even if certain the noise were a sum of telegraph noise processes, infeasible to distinguish relative contributions based on PSD alone

Candidate noise types

Charge noise



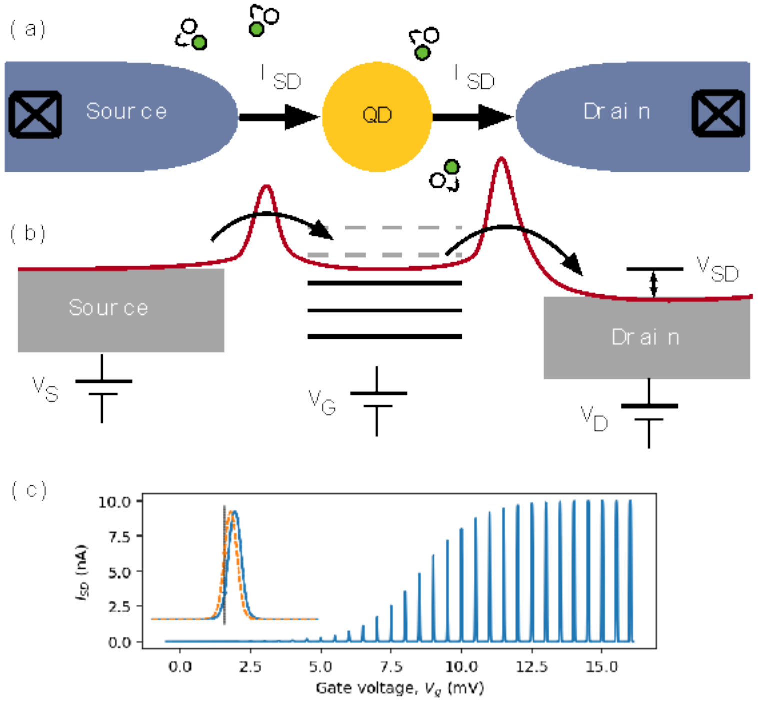
Nuclear contact HF coupling



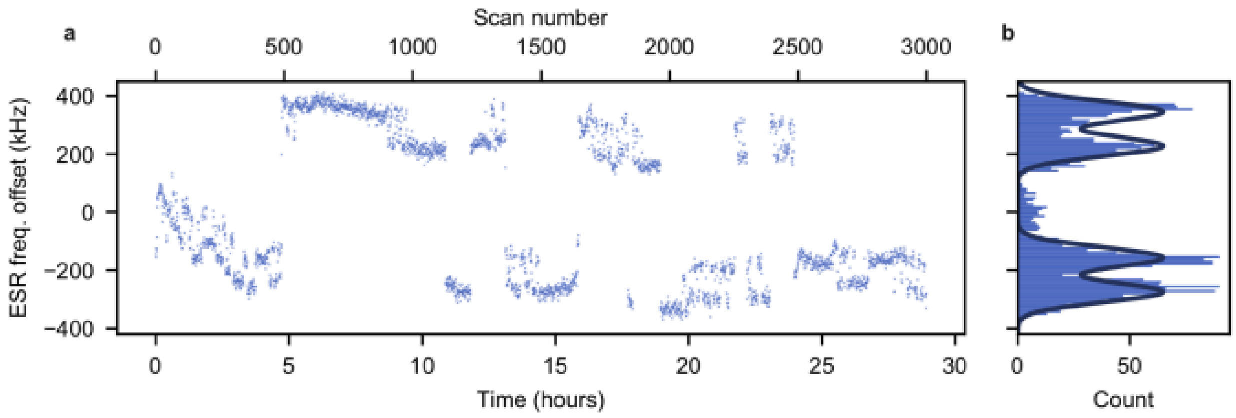
[Hensen, et al. Nat. Nano. (2020)]

Candidate noise types

Charge noise



Nuclear contact HF coupling



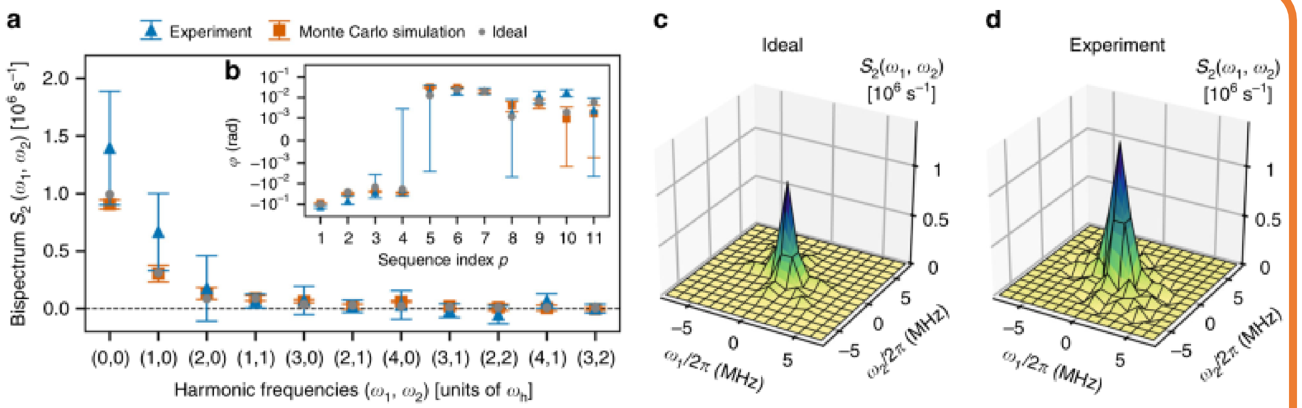
[Hensen, et al. Nat. Nano. (2020)]

Noise due to few-level fluctuators:

- Can we identify properties of the fluctuators themselves?
- Model parameters: coupling strengths, transition rates, biases
- If feasible, may provide more complete characterization of underlying noise mechanisms, suggest mitigation steps, and enable iterative perturbation-response measurements

Background

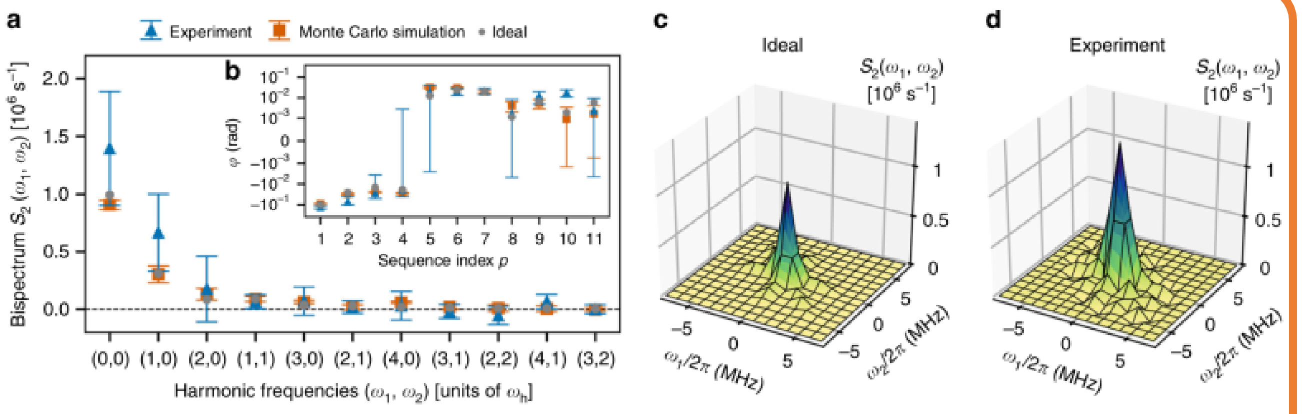
- For non-Gaussian noise, other measures based on higher-order correlations can be informative (bispectrum, second spectrum, etc.)



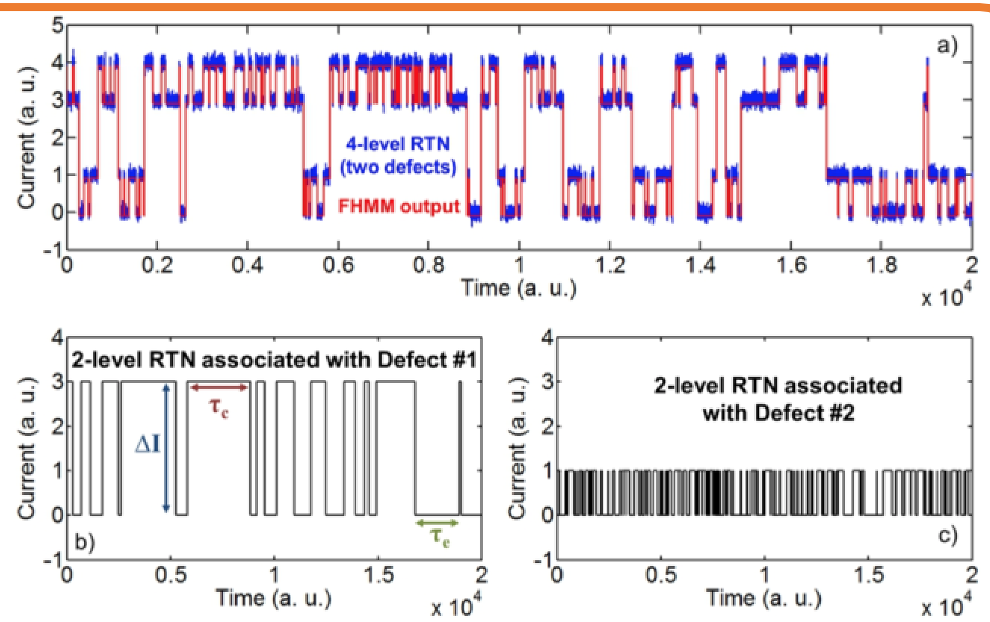
[Sung, et al. Nat. Comm. (2019)]

Background

- For non-Gaussian noise, other measures based on higher-order correlations can be informative (bispectrum, second spectrum, etc.)
- The approach we're taking here is to model the noise as a ***factorial*** hidden Markov model (FHMM)
  - An explicit model for underlying discrete noise degrees of freedom
  - Factorial → independent degrees of freedom and structured (factorized) reduced-dimensional parameter space
  - Previously applied to study telegraph noise in transistors [Puglisi, et al.]



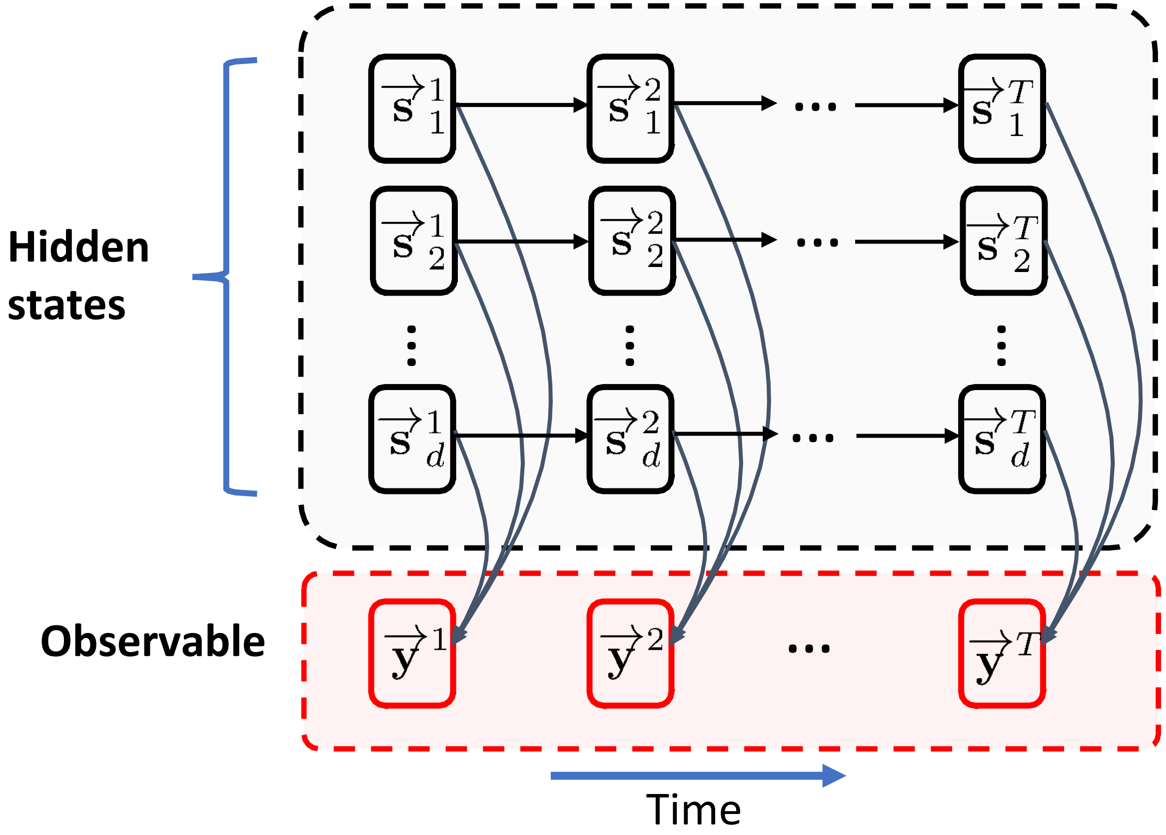
[Sung, et al. Nat. Comm. (2019)]



[Puglisi, et al. IEEE IPFA (2017)]

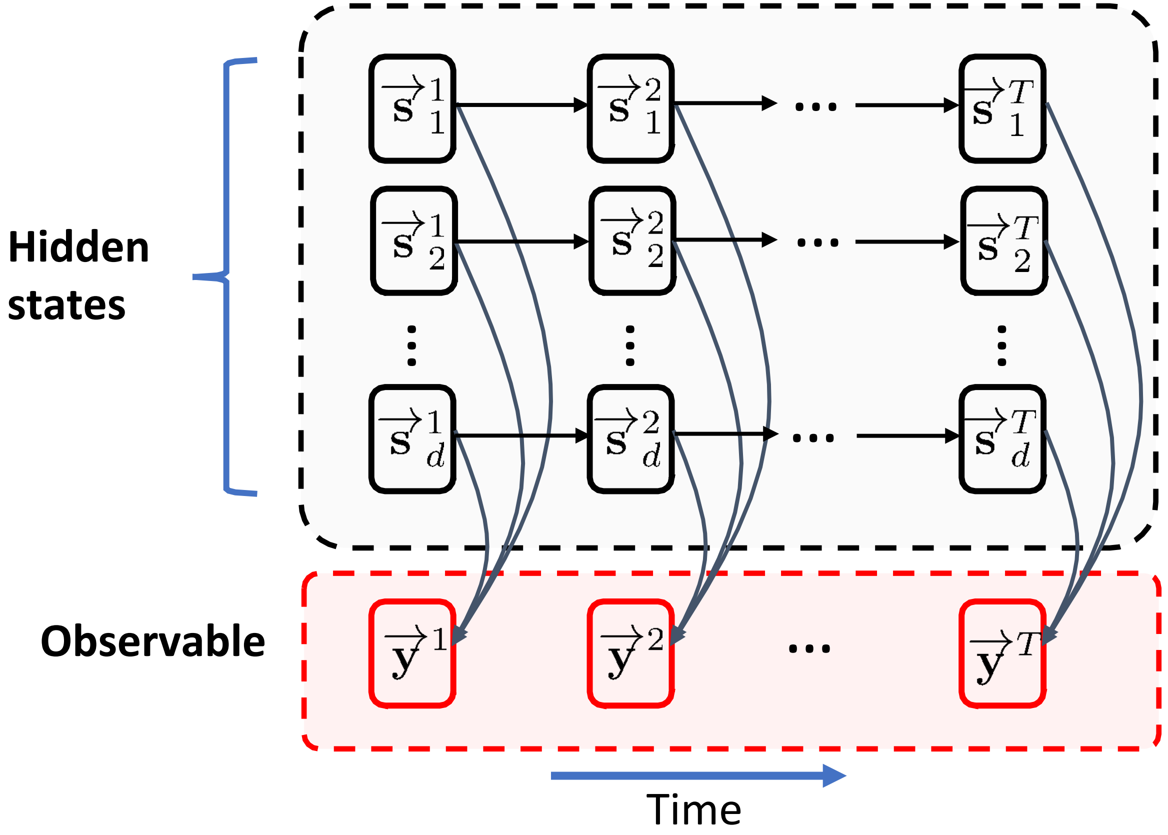
Sandia National Laboratories

Factorial hidden Markov models (FHMMs)



- **d** independent **k**-level fluctuators
- Additive contributions to the observable **y**
- $\mathbf{k}^{\mathbf{d}}$–dimensional state space for **d** fluctuators
- White noise background of strength **σ**

Factorial hidden Markov models (FHMMs)



- **d** independent **k**-level fluctuators
- Additive contributions to the observable **y**
- $\mathbf{k^d}$–dimensional state space for **d** fluctuators
- White noise background of strength **σ**

- Result of this project: **NoMoPy** ("Noise Modeling in Python") includes implementation of exact expectation maximization algorithm for FHMMs [Ghahramani & Jordan, Machine Learning 29, 245 (1997)] and uncertainty estimation machinery
- Maximum likelihood estimation (MLE) to infer the underlying model parameters and hidden state trajectory (Viterbi algorithm) for given noise timeseries.

Uncertainty quantification for FHMMs

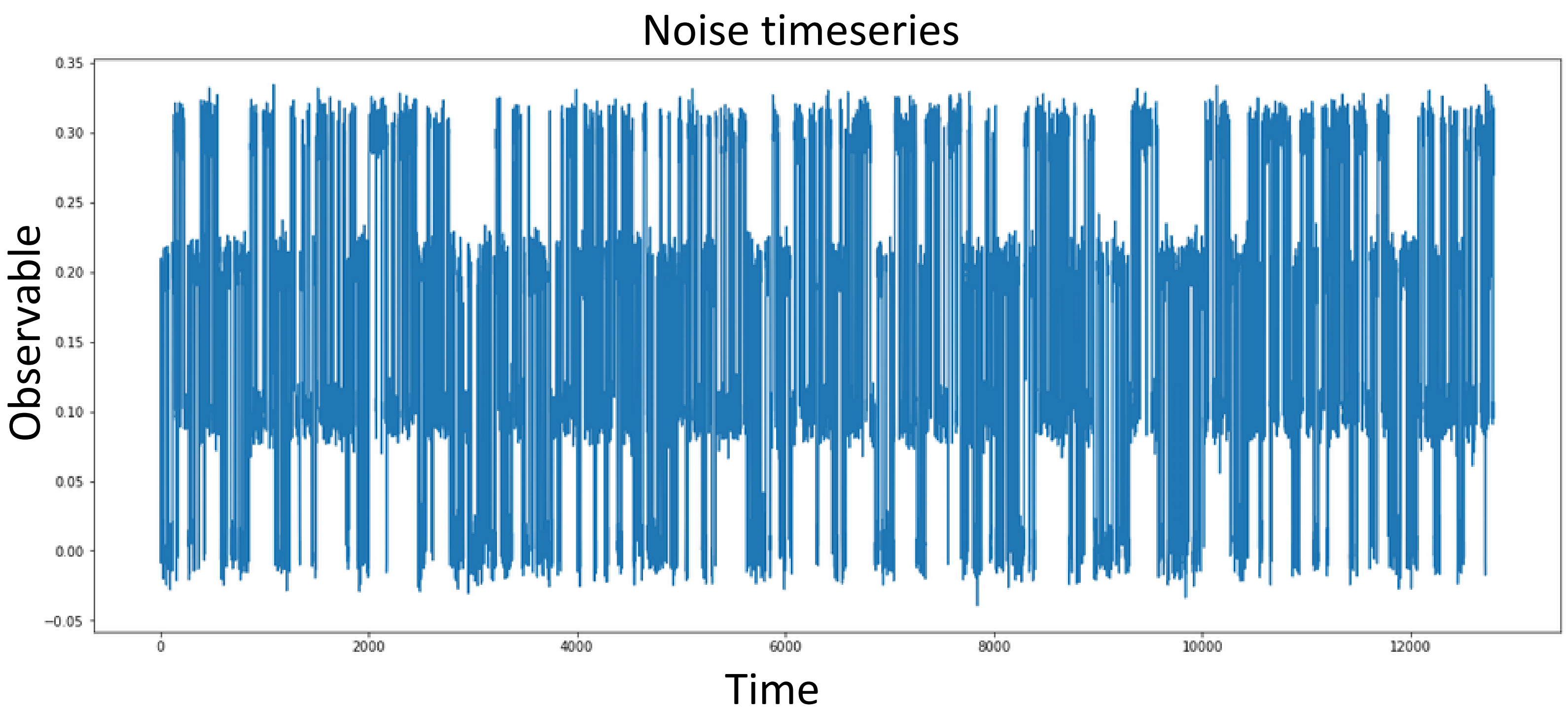
- An estimate for model parameters is much more useful with **confidence intervals**

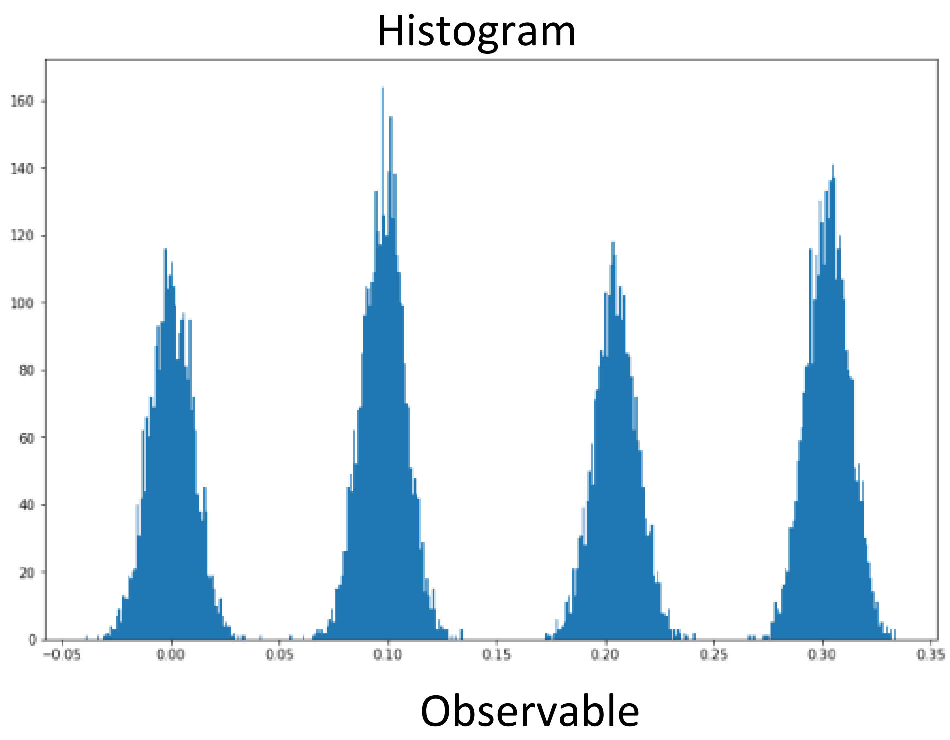
- We compute confidence intervals for underlying parameters in two ways:
  - Based on **Hessian** of loglikelihood at the optimal parameter values
  - **Bootstrapping** (more pessimistic error bars than Hessian-based estimate, in practice)

Uncertainty quantification for FHMMs

- An estimate for model parameters is much more useful with **confidence intervals**
- We compute confidence intervals for underlying parameters in two ways:
  - Based on **Hessian** of loglikelihood at the optimal parameter values
  - **Bootstrapping** (more pessimistic error bars than Hessian-based estimate, in practice)
- **Model selection problem:**
  - How many fluctuators are needed to satisfactorily describe the measurement timeseries?
  - At what point does adding more model degrees of freedom not improve predictive power?
  - Confidence intervals help inform when to stop adding more model degrees of freedom (identifying superfluous degrees of freedom, poorly constrained weights relative to noise background)

Example: 2 TLFs, **weak** white noise



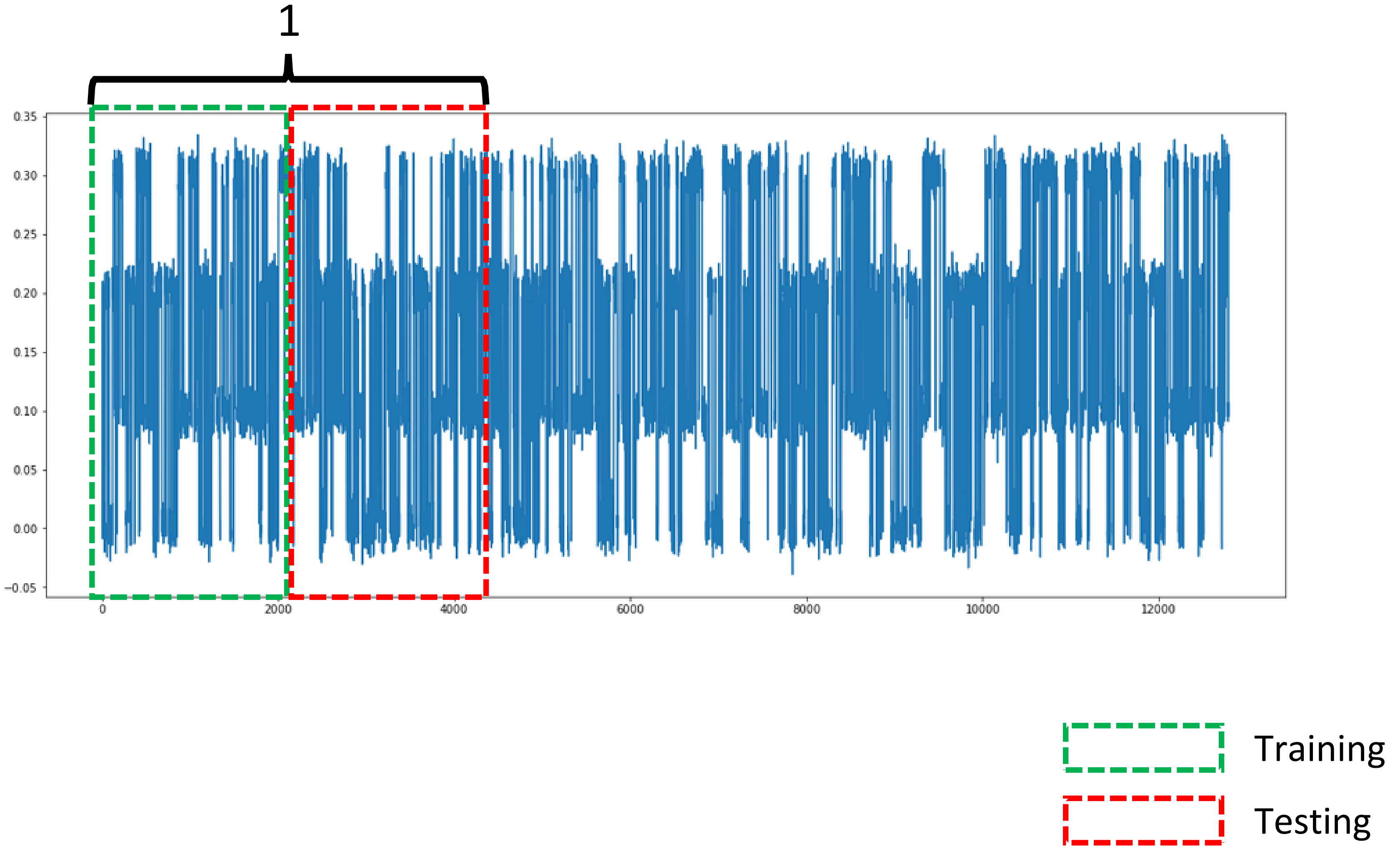


- Exact model has d=2 (four total hidden states)
- Weak white noise → levels easily distinguished

Example: 2 TLFs, **weak** white noise

General idea: A "good" model should perform well on data that haven't been used to train it

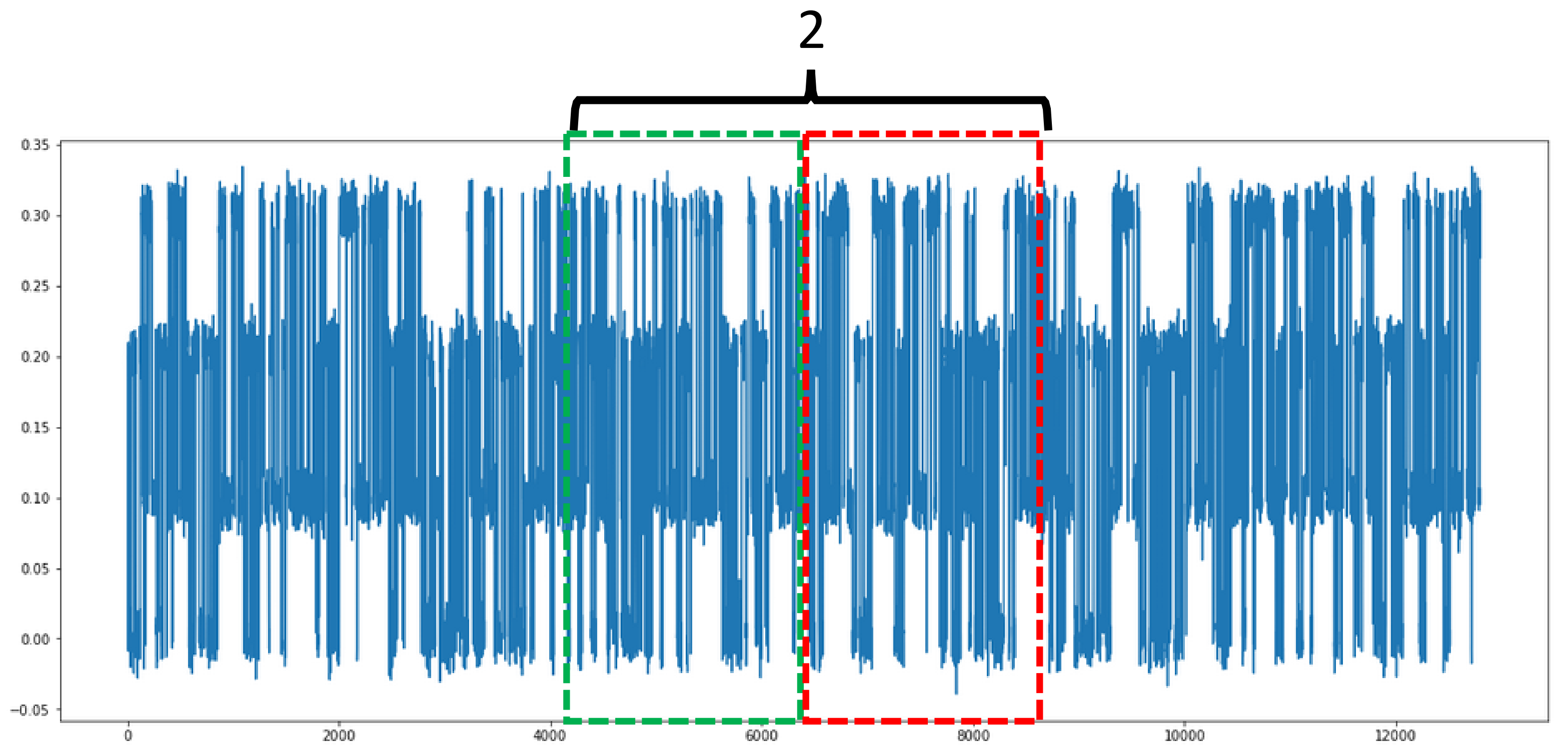
Three-fold cross-validation (CV)



Example: 2 TLFs, **weak** white noise

General idea: A "good" model should perform well on data that haven't been used to train it

Three-fold cross-validation (CV)



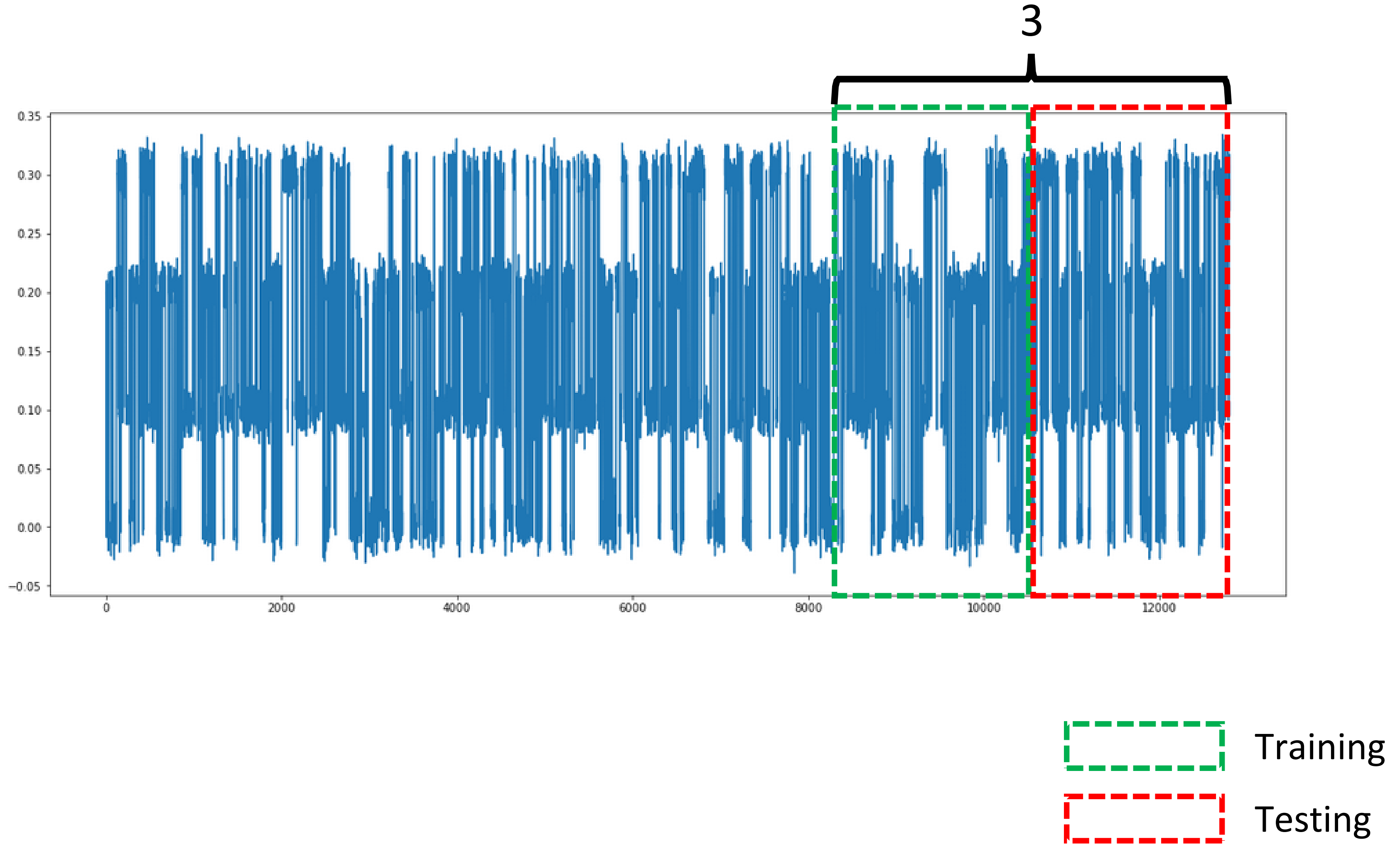
Training

Testing

Example: 2 TLFs, **weak** white noise

General idea: A "good" model should perform well on data that haven't been used to train it

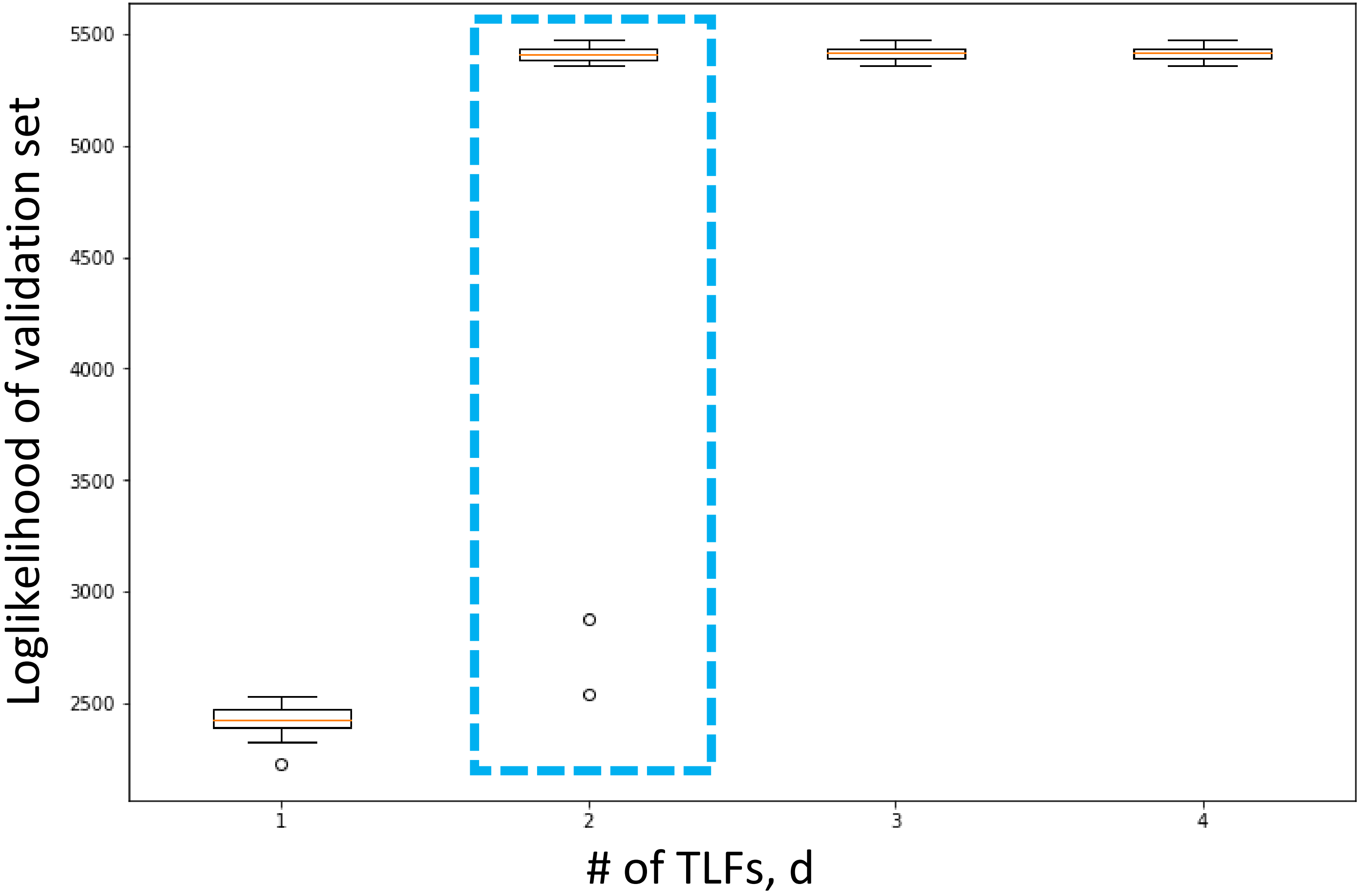
Three-fold cross-validation (CV)



Example: 2 TLFs, **weak** white noise

All confidence intervals reported are at the 95% level

20-fold cross-validation to pick d (# of TLFs)

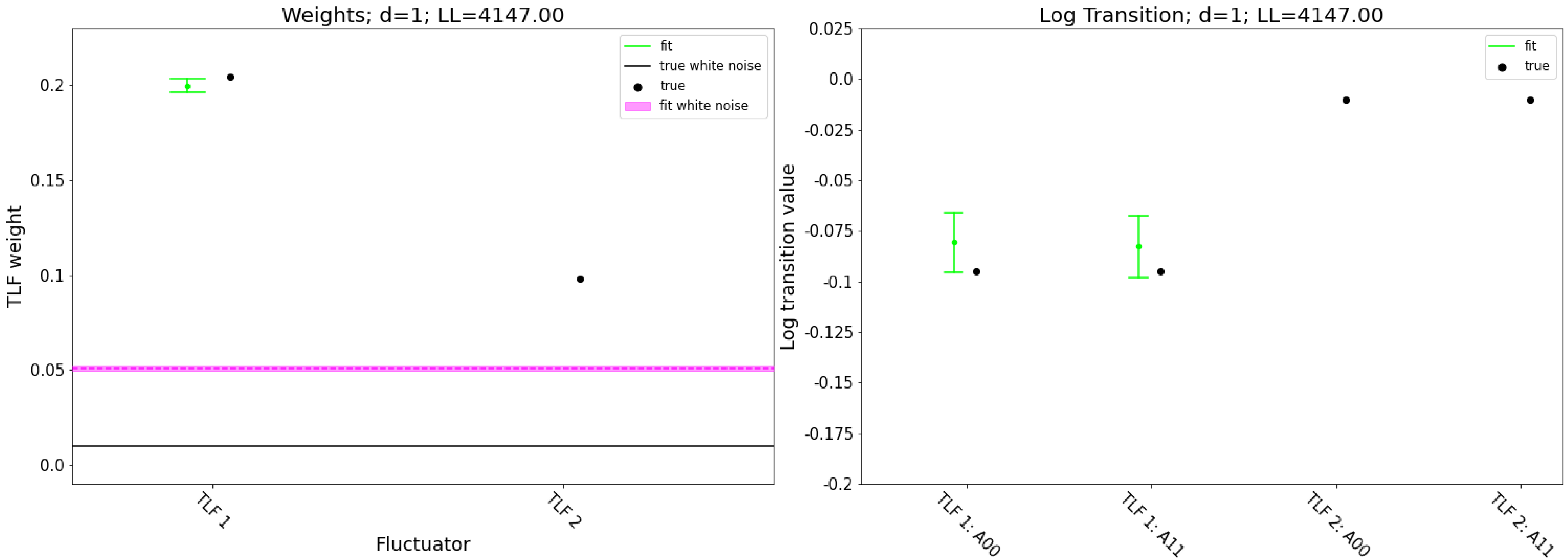


Sandia National Laboratories

Example: 2 TLFs, **weak** white noise

Transition matrix: A_{ij} = probability to transition $j \to i$

Ansatz: **d=1**

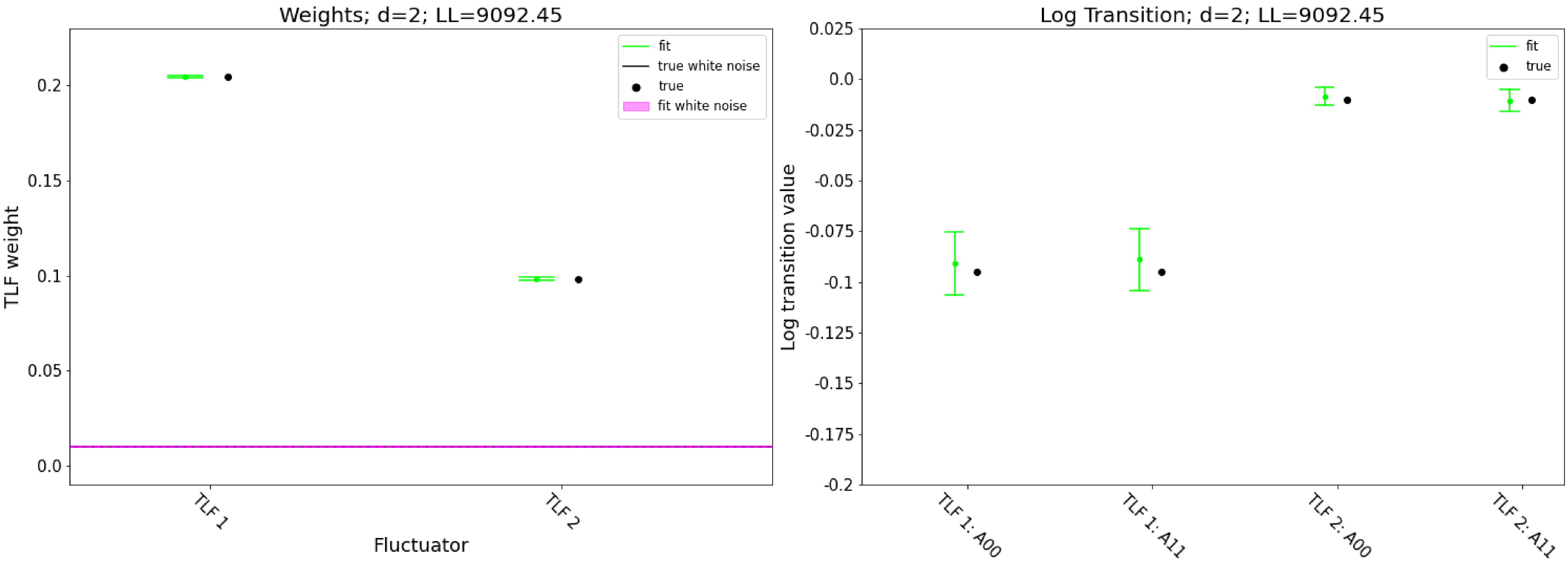


Sandia National Laboratories

Example: 2 TLFs, **weak** white noise

Transition matrix: A_{ij} = probability to transition $j \rightarrow i$

Ansatz: **d=2**



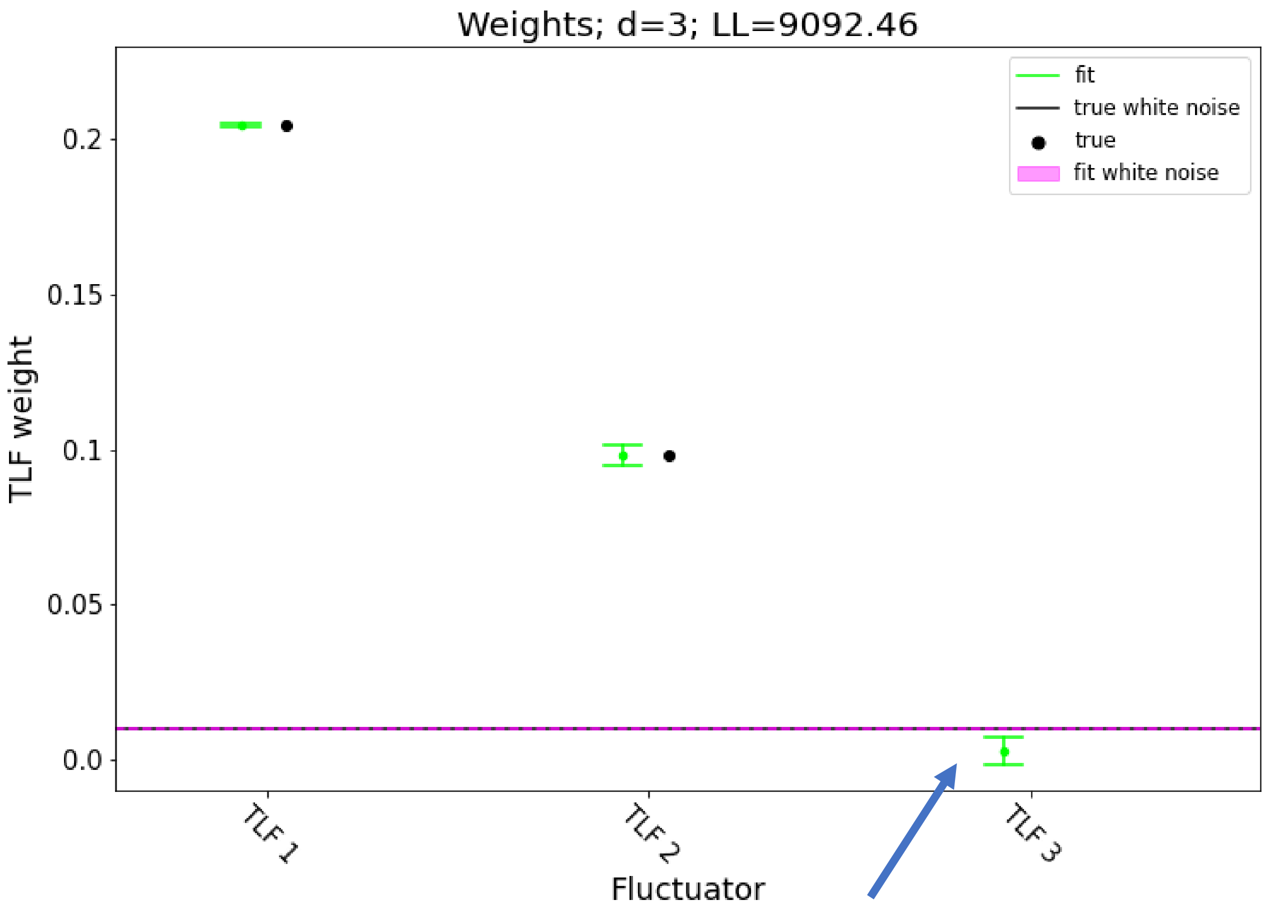
- True dipoles and transition probabilities sit within confidence intervals

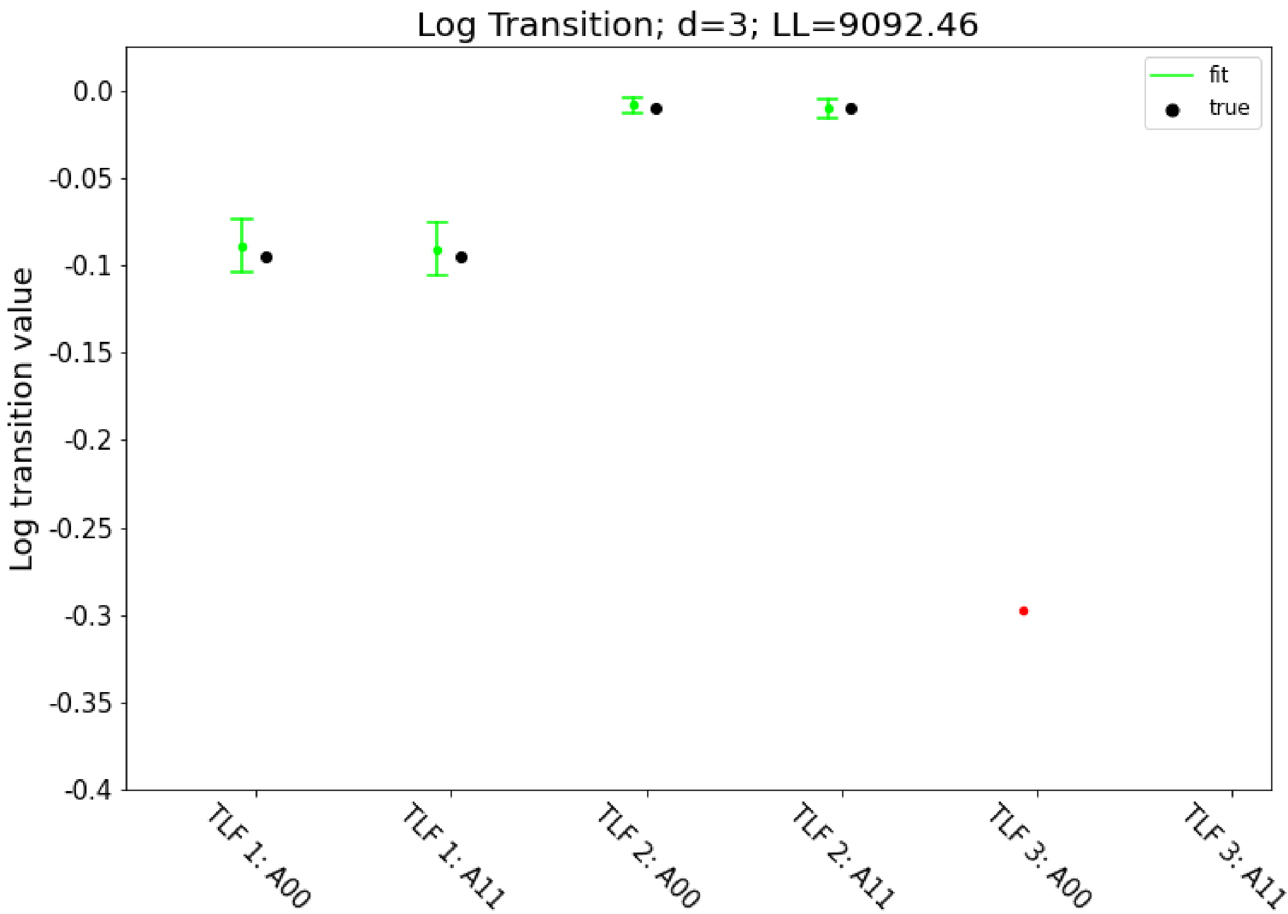
Sandia National Laboratories

Example: 2 TLFs, **weak** white noise

Transition matrix: A_{ij} = probability to transition $j \to i$

Ansatz: **d=3**



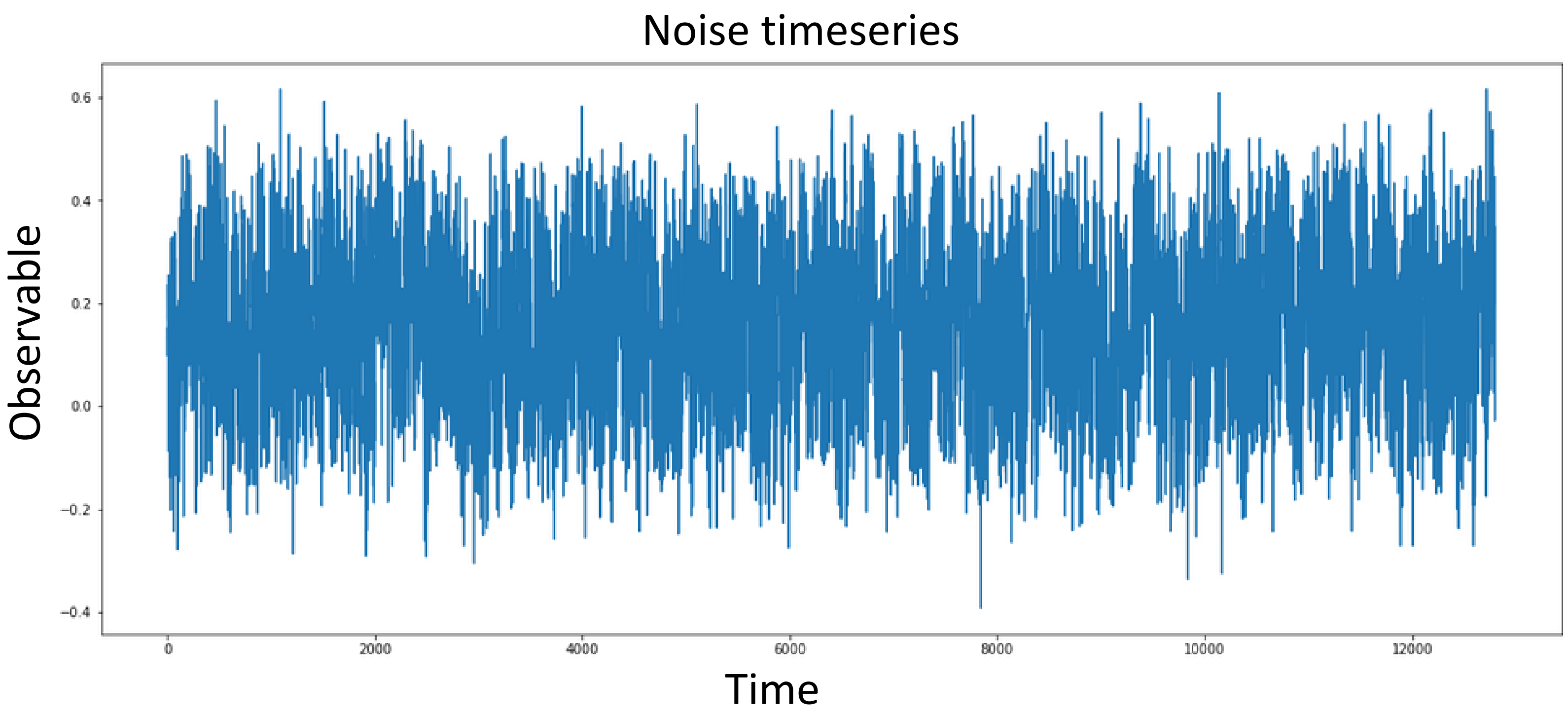


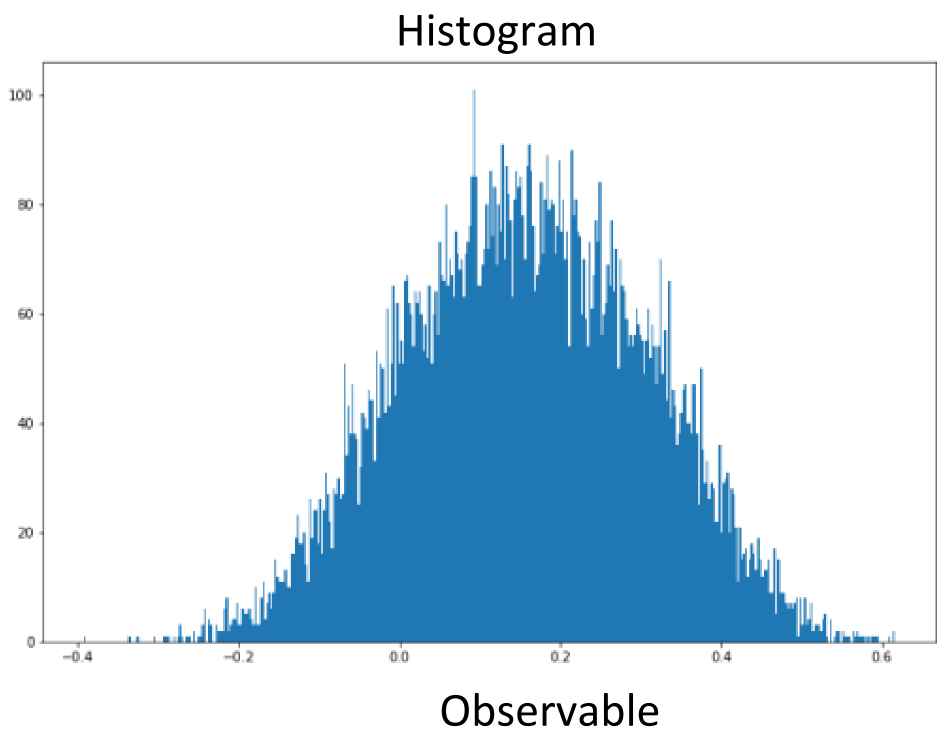
Third fluctuator estimated to have effectively zero dipole (superfluous degree of freedom)

Red points are NaN confidence intervals (due to excessive insensitivity)

Sandia National Laboratories

Example: 2 TLFs, **strong** white noise

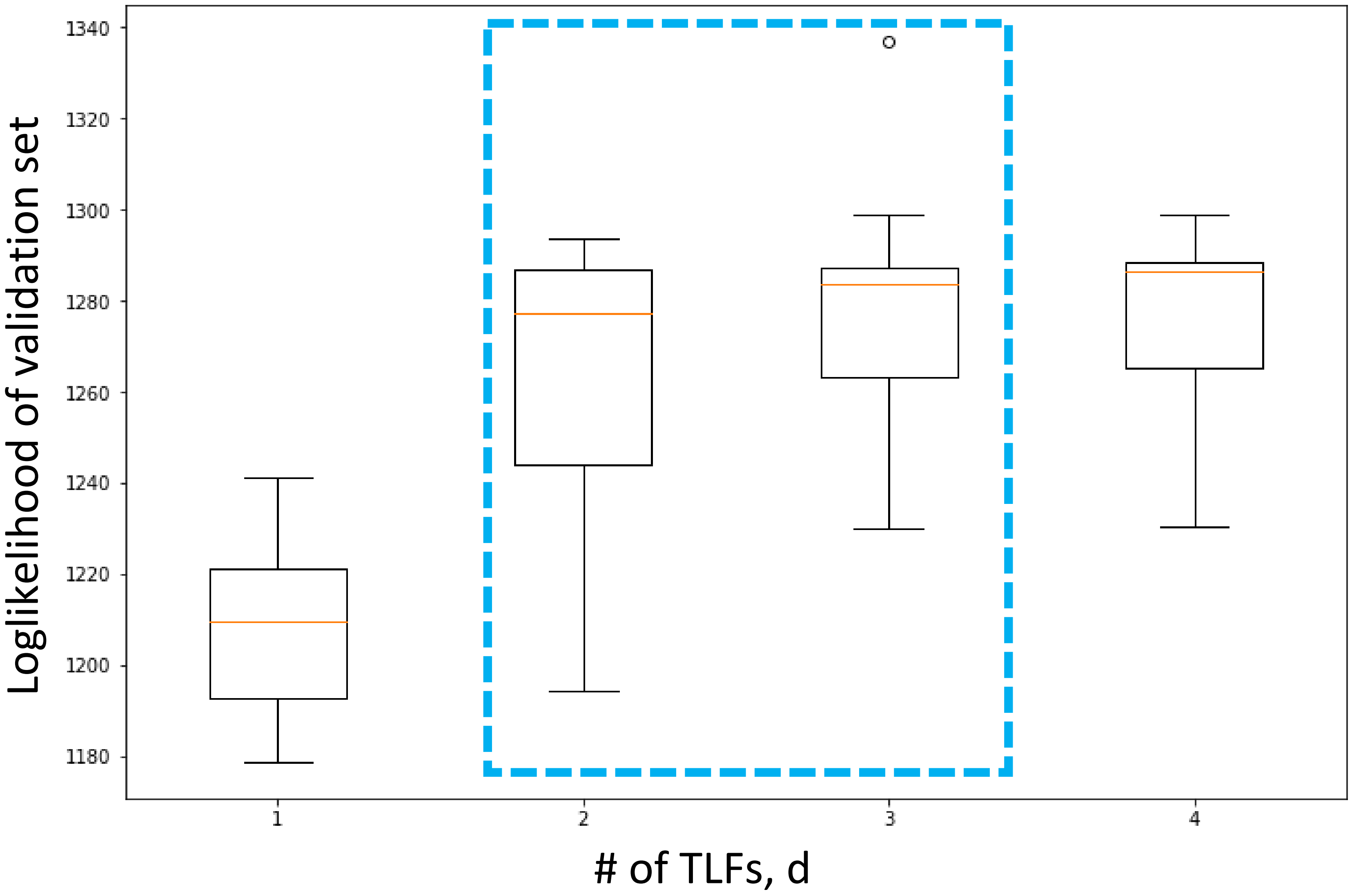




- Exact model has d=2 (four total hidden states)
- Strong white noise → no discrete levels visible

Example: 2 TLFs, **strong** white noise

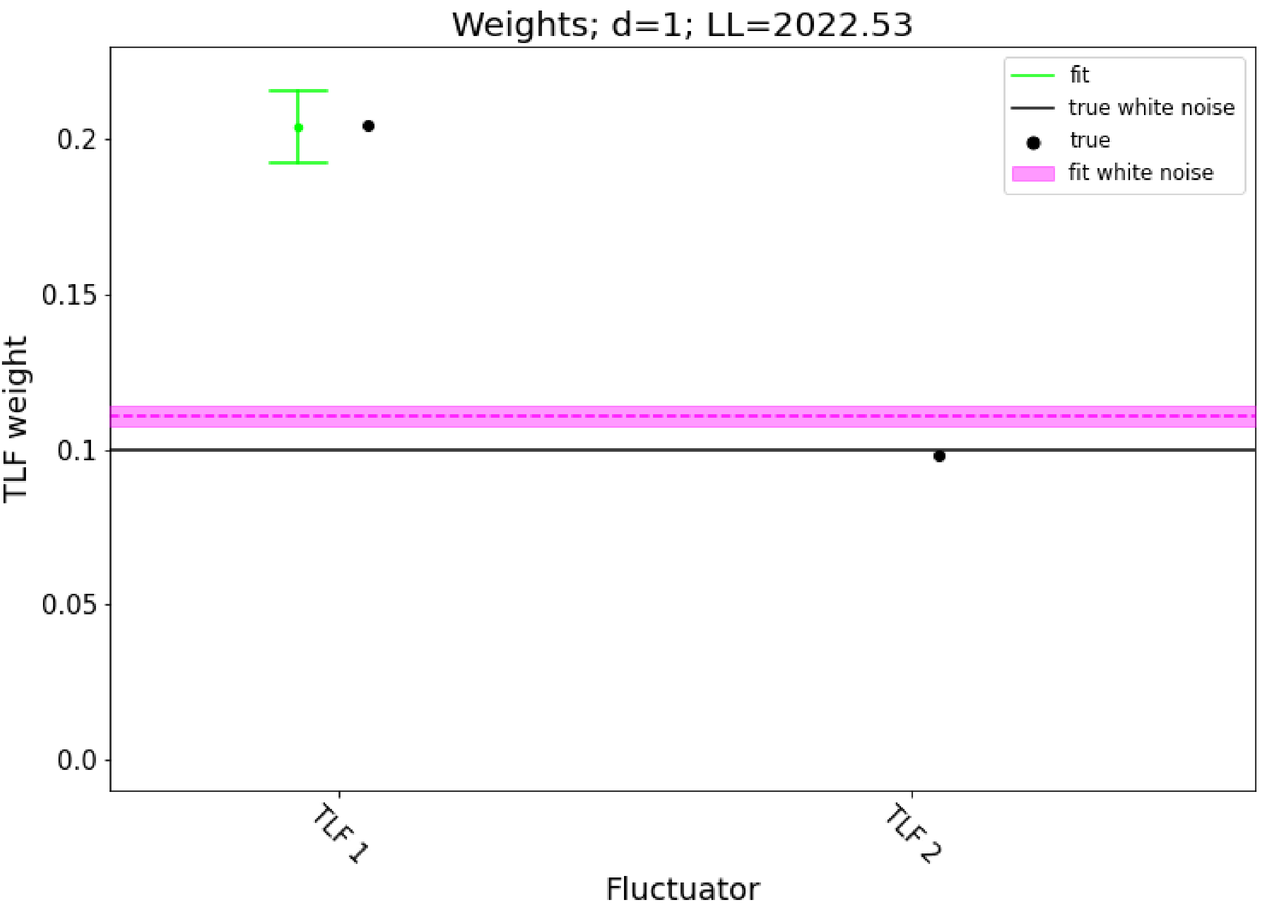
20-fold cross-validation to pick d (# of TLFs)

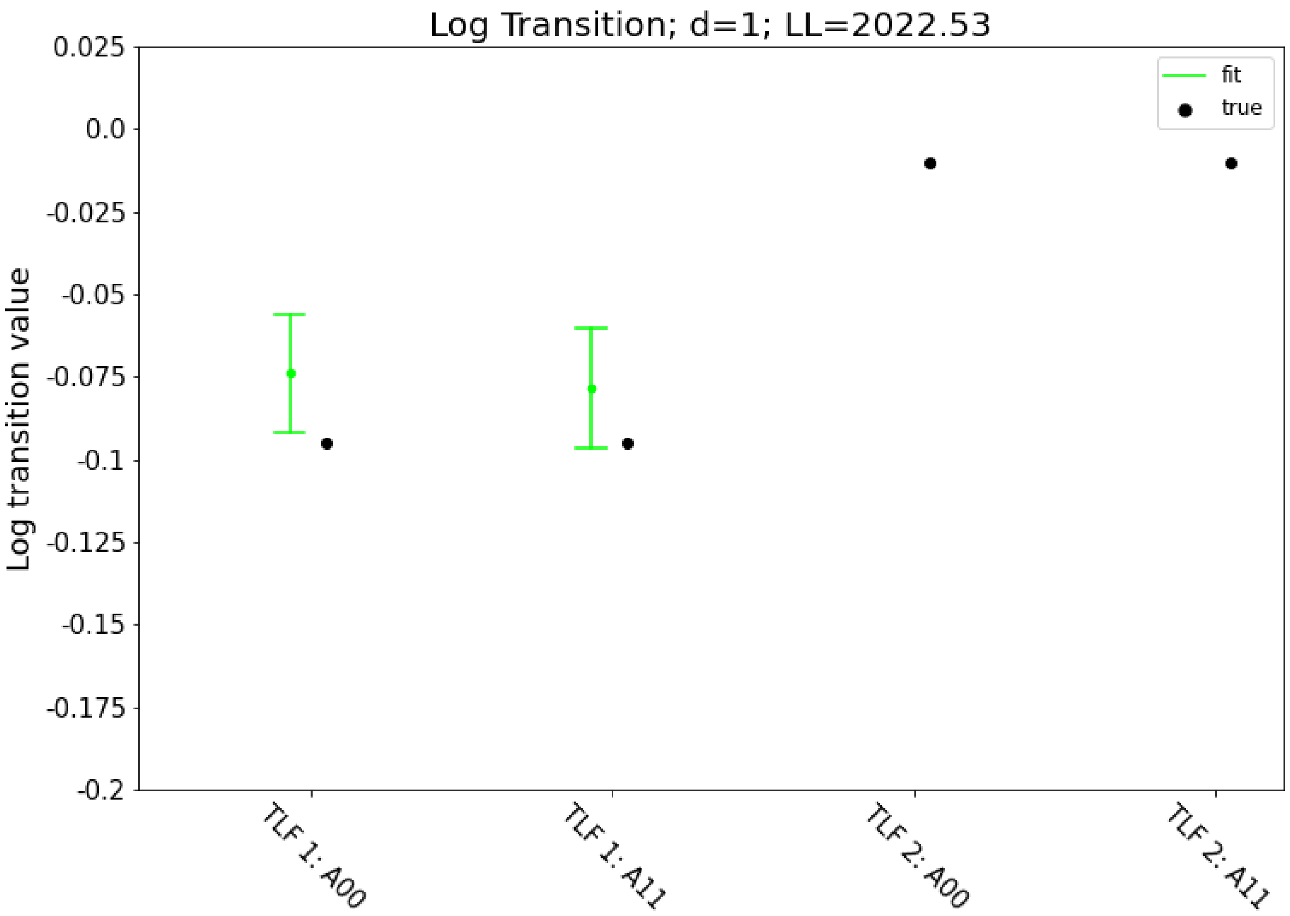


Example: 2 TLFs, **strong** white noise

Transition matrix: A_{ij} = probability to transition $j \rightarrow i$

Ansatz: **d=1**



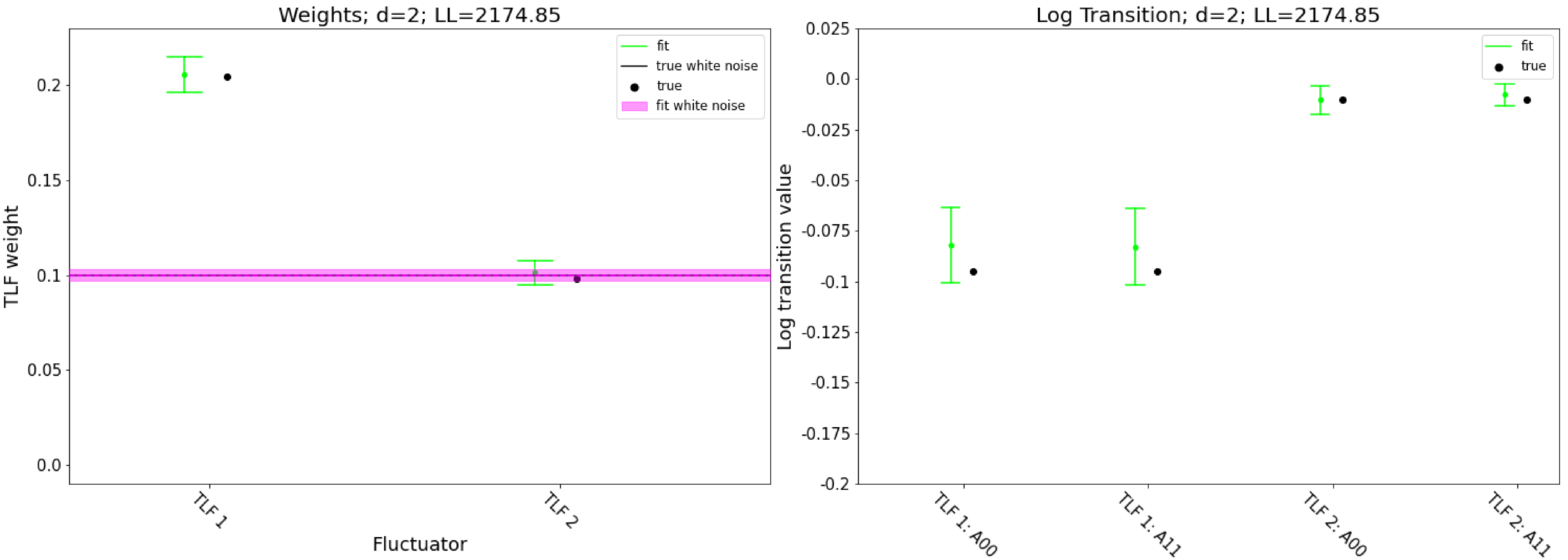


Sandia National Laboratories

Example: 2 TLFs, **strong** white noise

Transition matrix: A_{ij} = probability to transition $j \rightarrow i$

Ansatz: **d=2**

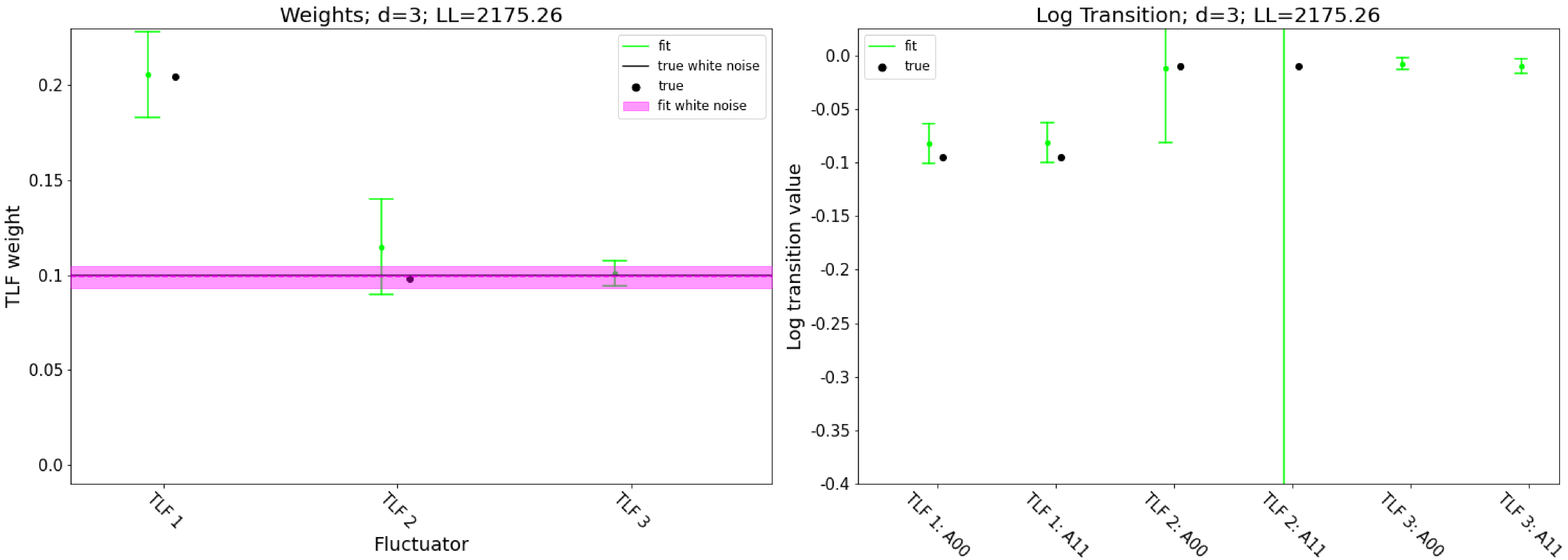


- Again, true dipoles and transition probabilities sit within confidence intervals

Example: 2 TLFs, **strong** white noise

Transition matrix: A_{ij} = probability to transition $j \rightarrow i$

Ansatz: **d=3**



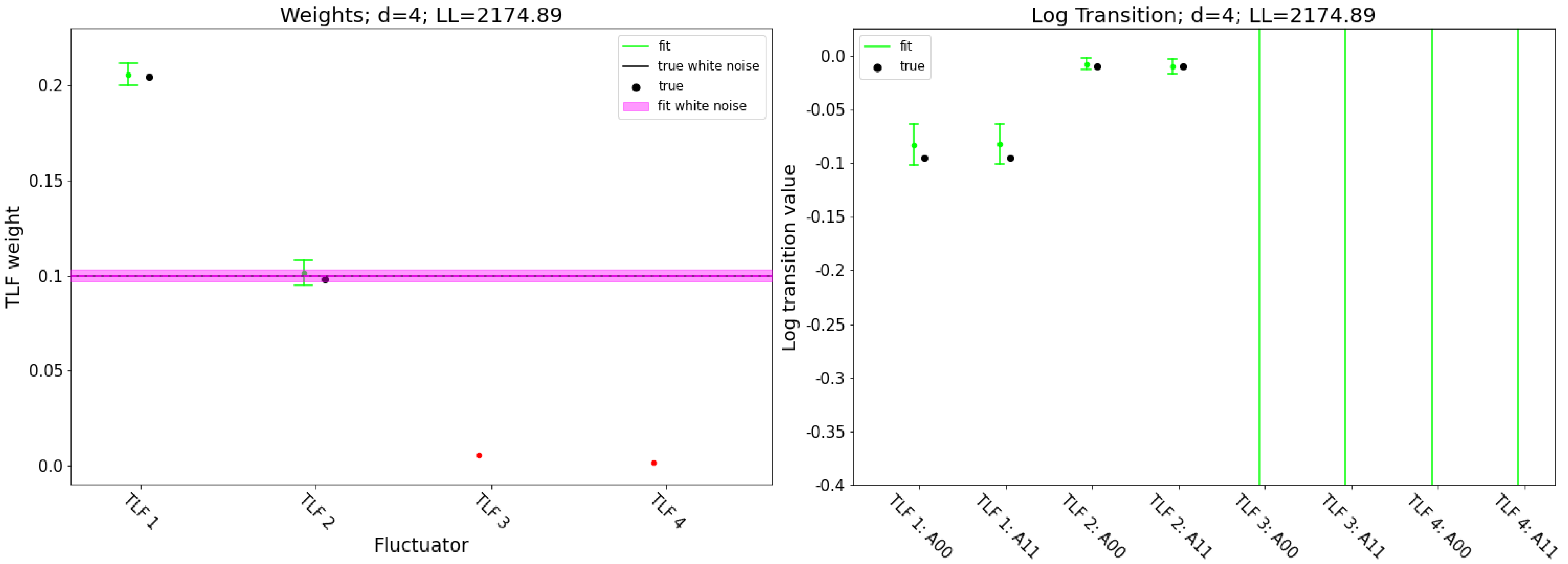
- Two TLFs tightly constrained around true values, additional TLF poorly constrained around white noise level

Sandia National Laboratories

Example: 2 TLFs, **strong** white noise

Transition matrix: A_{ij} = probability to transition $j \to i$

Ansatz: **d=4**



- Two TLFs tightly constrained around true values, additional TLFs fit to have negligible weight

Sandia National Laboratories

Example: 2 TLFs

Observations:

- Cross validation jumps/saturates near true **d** (# TLFs)
- At larger **d**, the model fitting starts to break down or the uncertainties are excessive

- Fitting more difficult when background white noise level is high
- For the strong noise case, the white noise level was chosen to be near that of the smallest TLF dipole

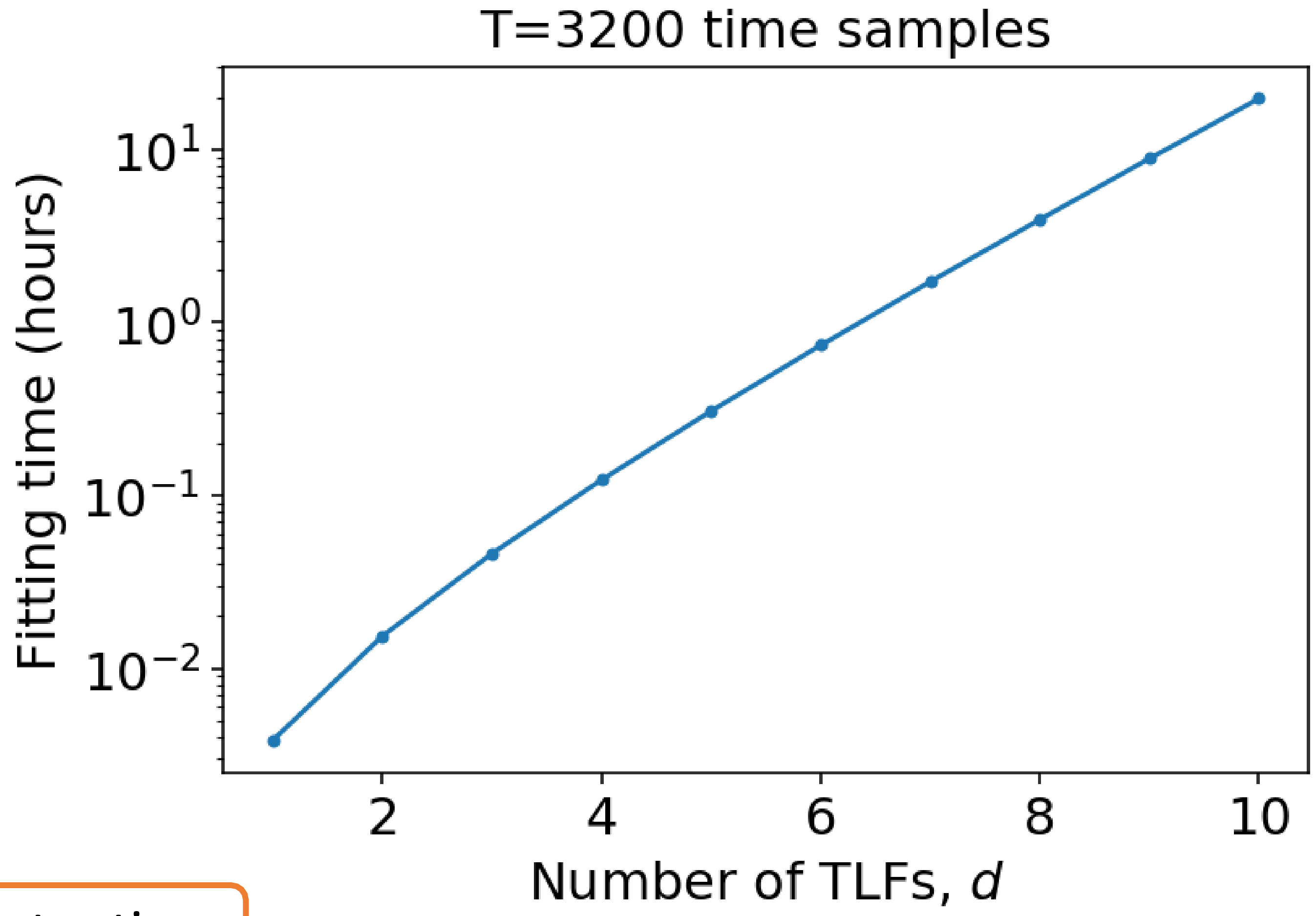
- **d** = 2 found to be the best fit in both cases

Scaling of computational complexity: How large can you go?

Log likelihood computation: $O(Tdk^{d+1})$

For **d=6** (64 hidden states):
- log(L) evaluation in ~3 seconds
- Fitting in ~45 minutes
- Hessian evaluation in ~18 minutes

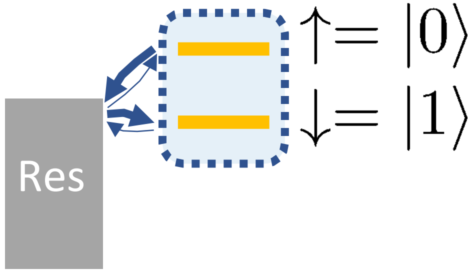
For **d=10** (1024 hidden states):
- log(L) evaluation in ~80 seconds
- Fitting in ~18 hours
- Hessian evaluation in ~20 hours



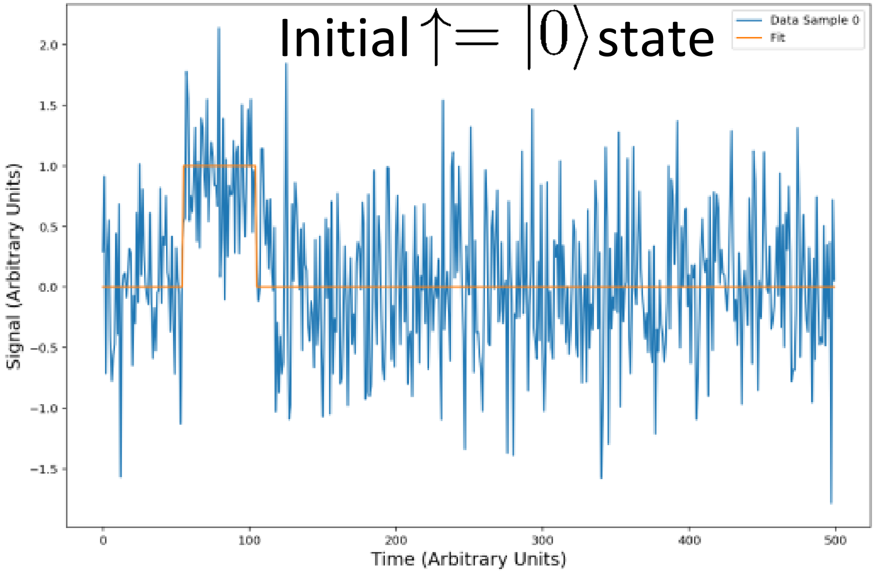
More degrees of freedom → more computer time

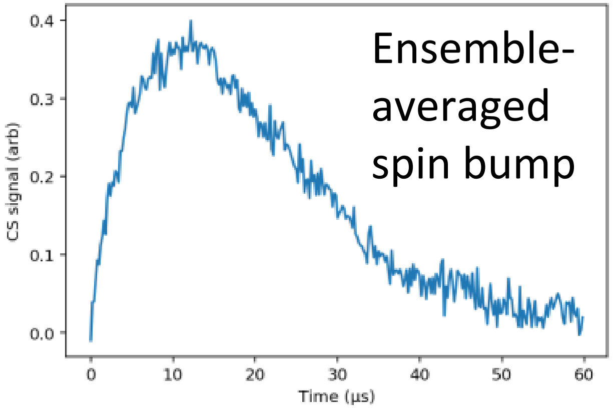
Other applications suitable for *NoMoPy*

- Optimal spin-to-charge readout
  - [D'Anjou & Coish PRA 89, 012313 (2014)]
  - [Gambetta, et al. PRA 76, 012325 (2007)]



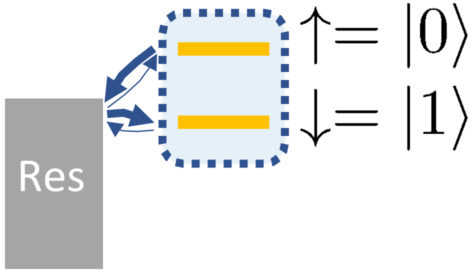
Example (3-level system):

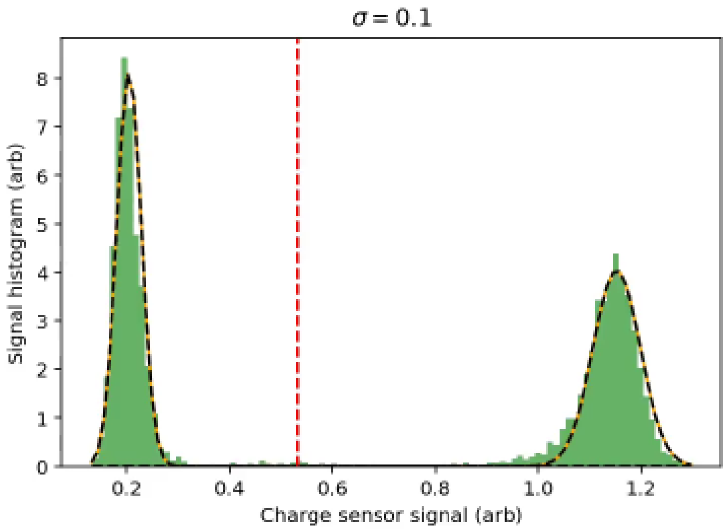




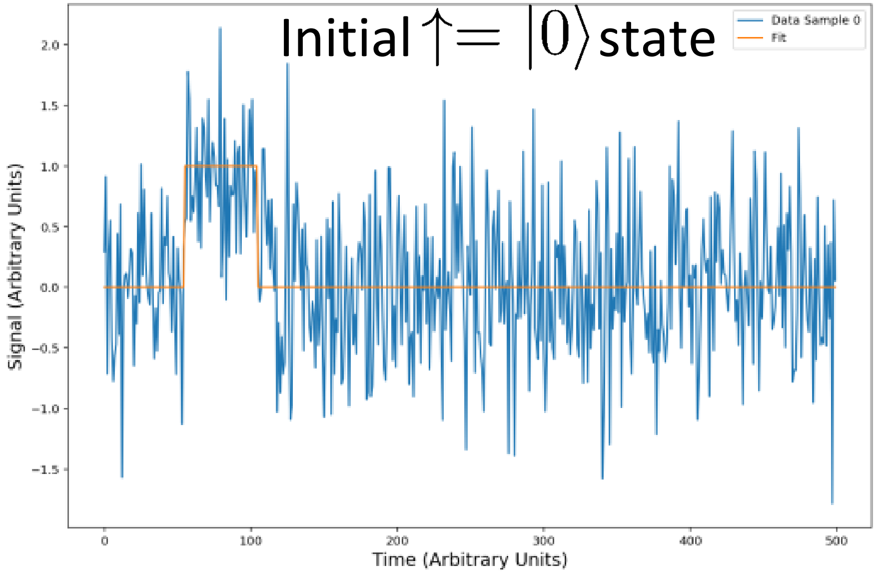
Other applications suitable for *NoMoPy*

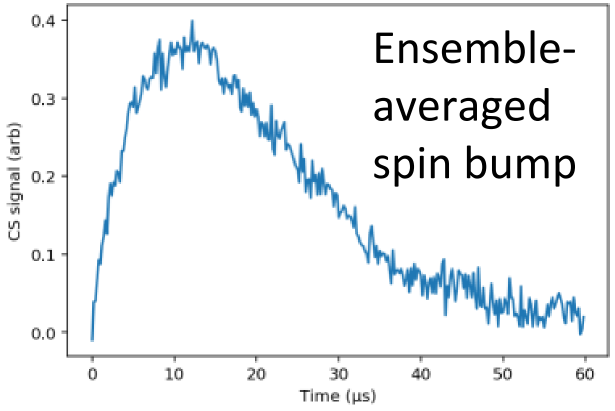
- Optimal spin-to-charge readout
  - [D'Anjou & Coish PRA 89, 012313 (2014)]
  - [Gambetta, et al. PRA 76, 012325 (2007)]



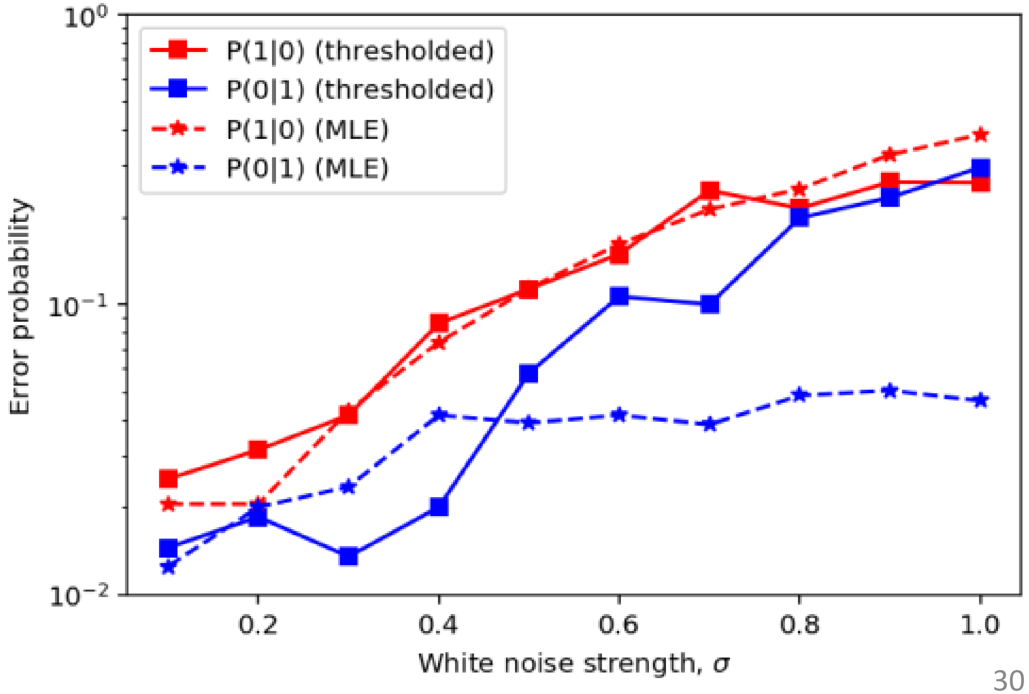


Example (3-level system):





2000 shots w/ known initial state



- Given a model and measurement timeseries, identify the most probable initial state
- Performance relative to simple thresholding depends on measurement details

Challenges

- High quality data needed (low systematic background, high precision) to capture many fluctuators
  - Any systematic background (e.g. 60 Hz line noise) needs to be well-filtered

- Inevitably, at some point the central limit theorem kicks in. Regime of interest here is between obvious telegraph noise and nearly Gaussian noise

- In practice, should be feasible to fit up to d~10 fluctuators (1k hidden states)
  - Cost of exact algorithm grows exponentially with d, soon requiring HPC resources

Summary

- We’ve developed a code, ***NoMoPy***, that implements factorial hidden Markov model fitting and parameter uncertainty quantification for noise generated by independent discrete fluctuations
  - Protocols for model selection and investigation of robustness to noise “signal-to-noise” (i.e. strength of fluctuators of interest relative to white noise background)

- Capability may expand inferences feasible from measurements of qubit-relevant non-Gaussian noise, e.g. charge and nuclear hyperfine noise
  - A “big data”-inspired approach to qubit noise analysis

- If you have noise measurements that might benefit from such an approach, we’d be interested to talk with you!