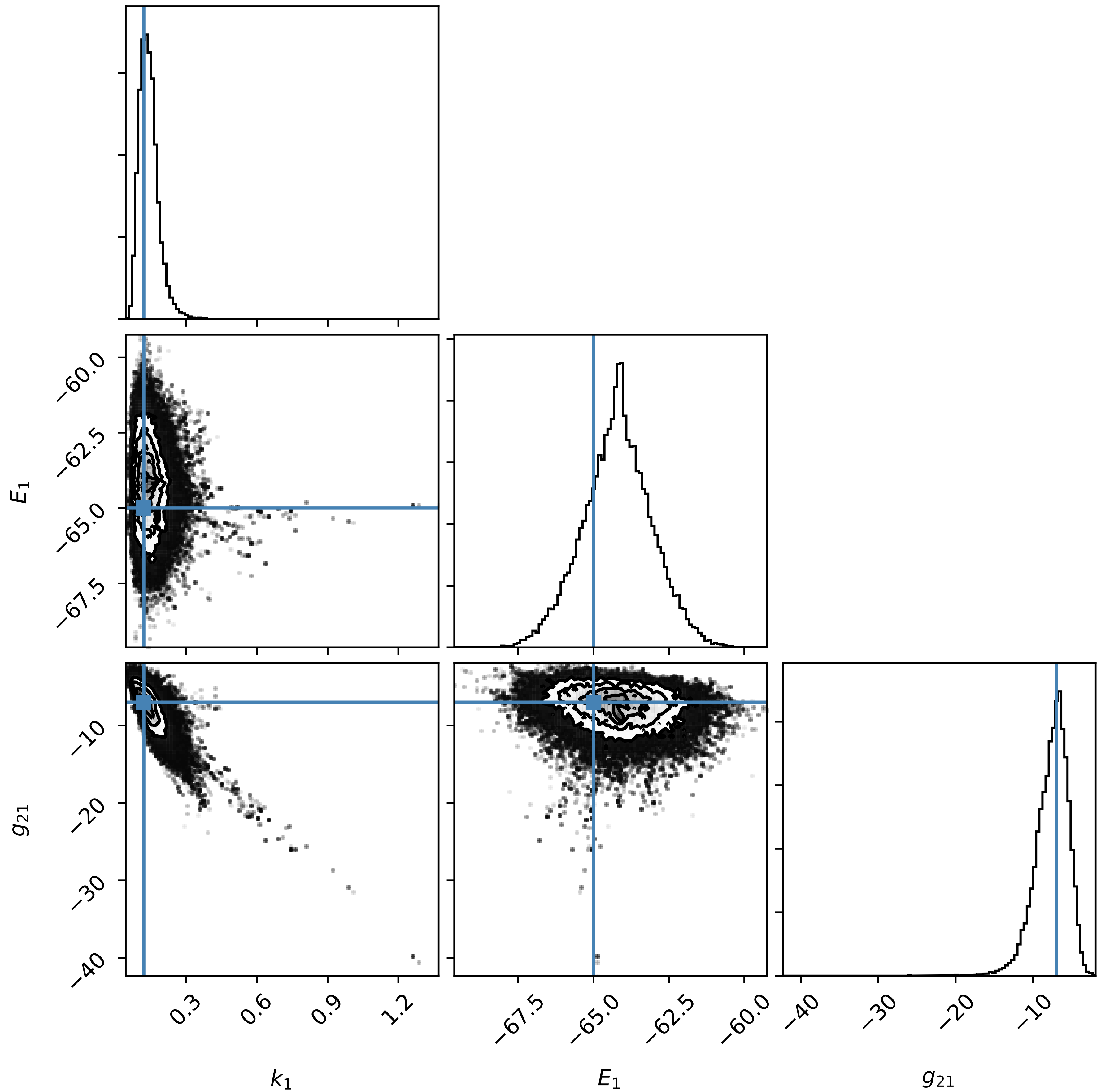
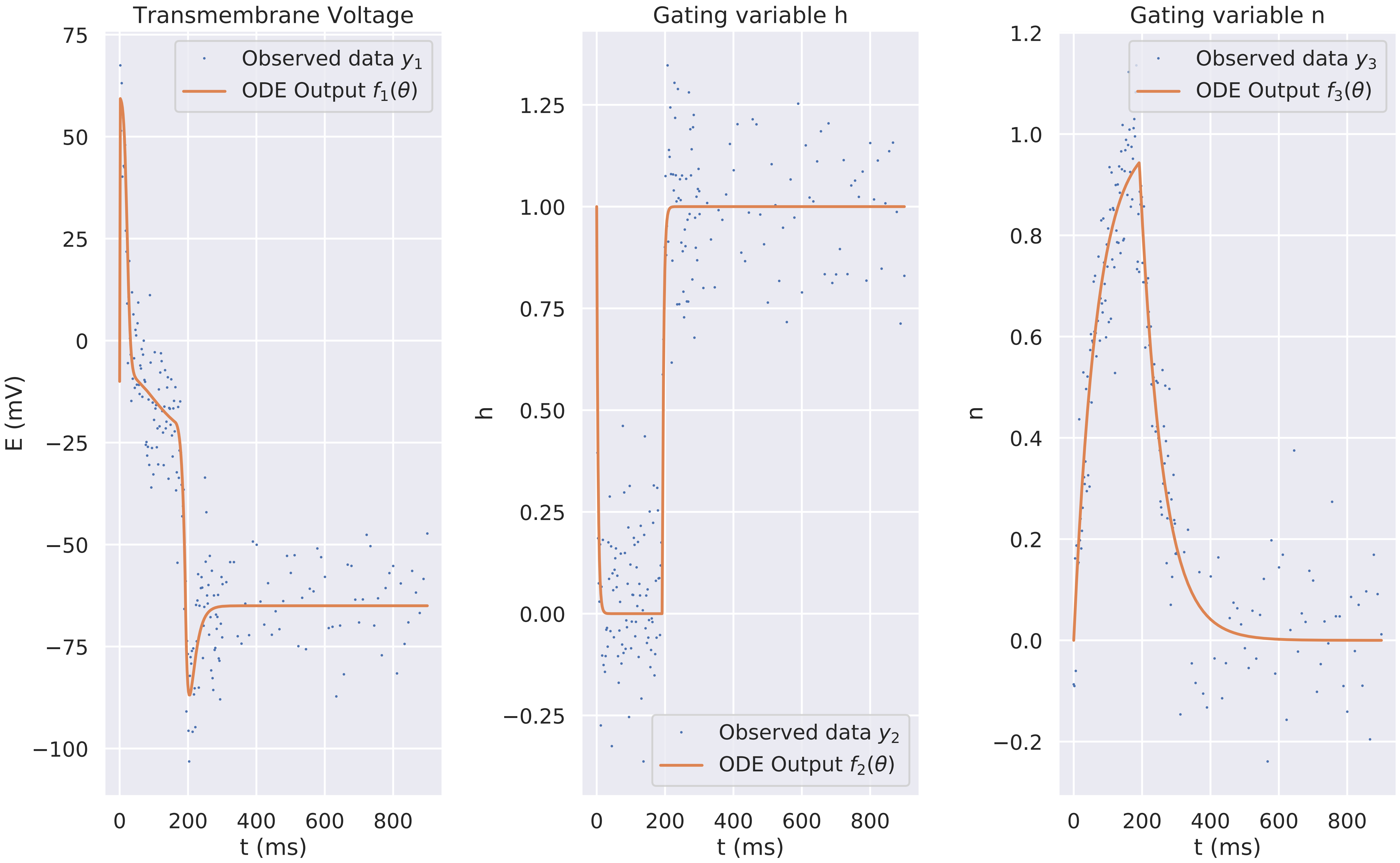
Obtaining first place estimates for Model 1

James Ritchie

(and Iain Murray)



−60.0
−62.5
−65.0
−67.5
E_1
−10
−20
−30
−40
g_{21}
0.3
0.6
0.9
1.2
−67.5
−65.0
−62.5
−60.0
−40
−30
−20
−10
k_1
E_1
g_{21}



Transmembrane Voltage
Observed data y_1
ODE Output $f_1(\theta)$
E (mV)
t (ms)
Gating variable h
Observed data y_2
ODE Output $f_2(\theta)$
h
t (ms)
Gating variable n
Observed data y_3
ODE Output $f_3(\theta)$
n
t (ms)

Bayesian Inference

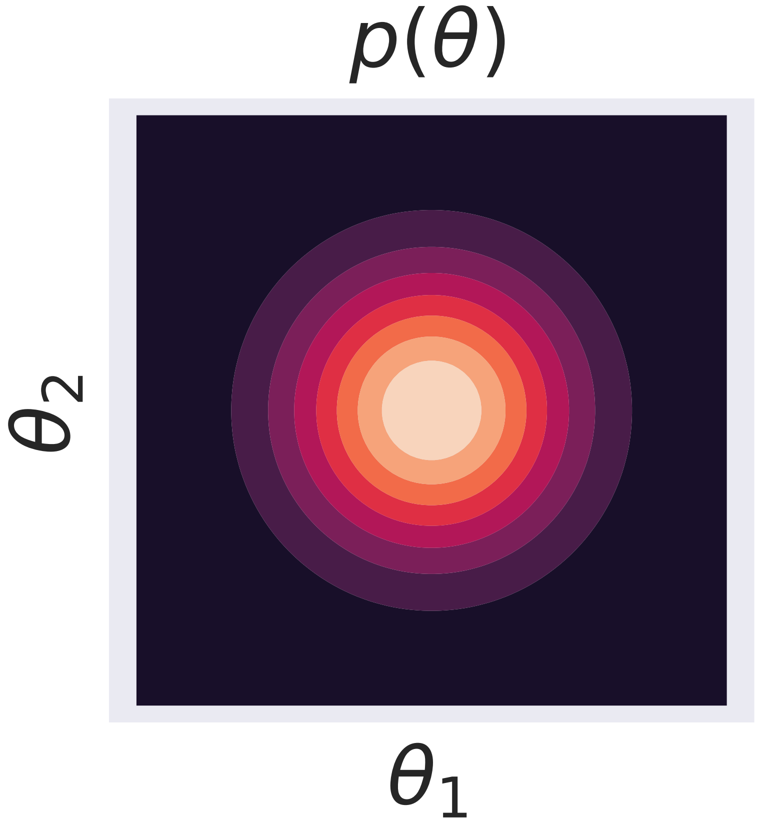
Bayes' rule:

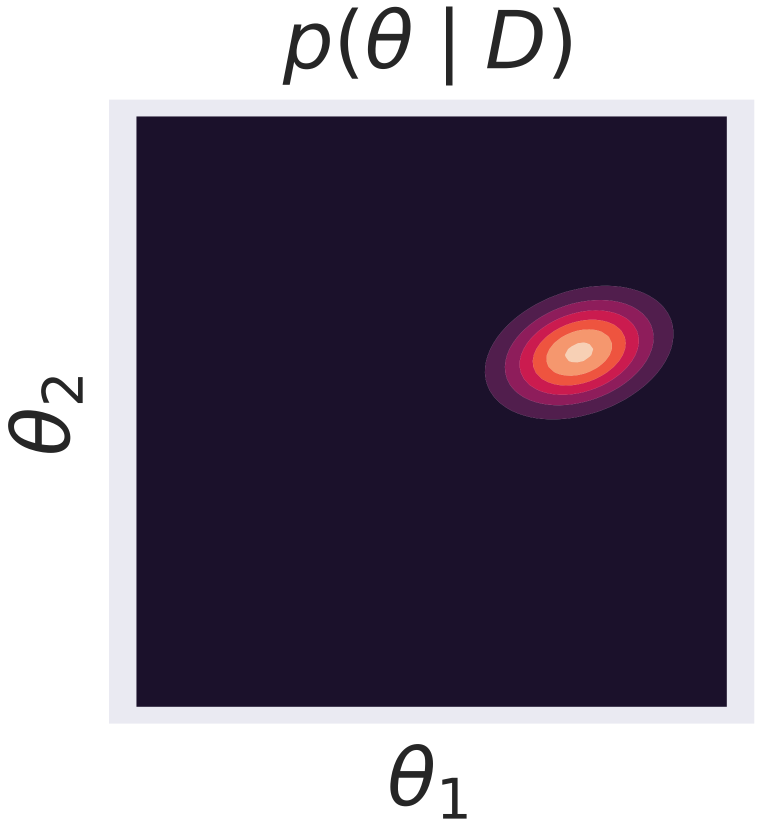
$$p(\theta \mid D) = \frac{p(D \mid \theta)p(\theta)}{p(D)}$$

Prior $p(\theta)$

Likelihood $p(D \mid \theta)$

Posterior $p(\theta \mid D)$



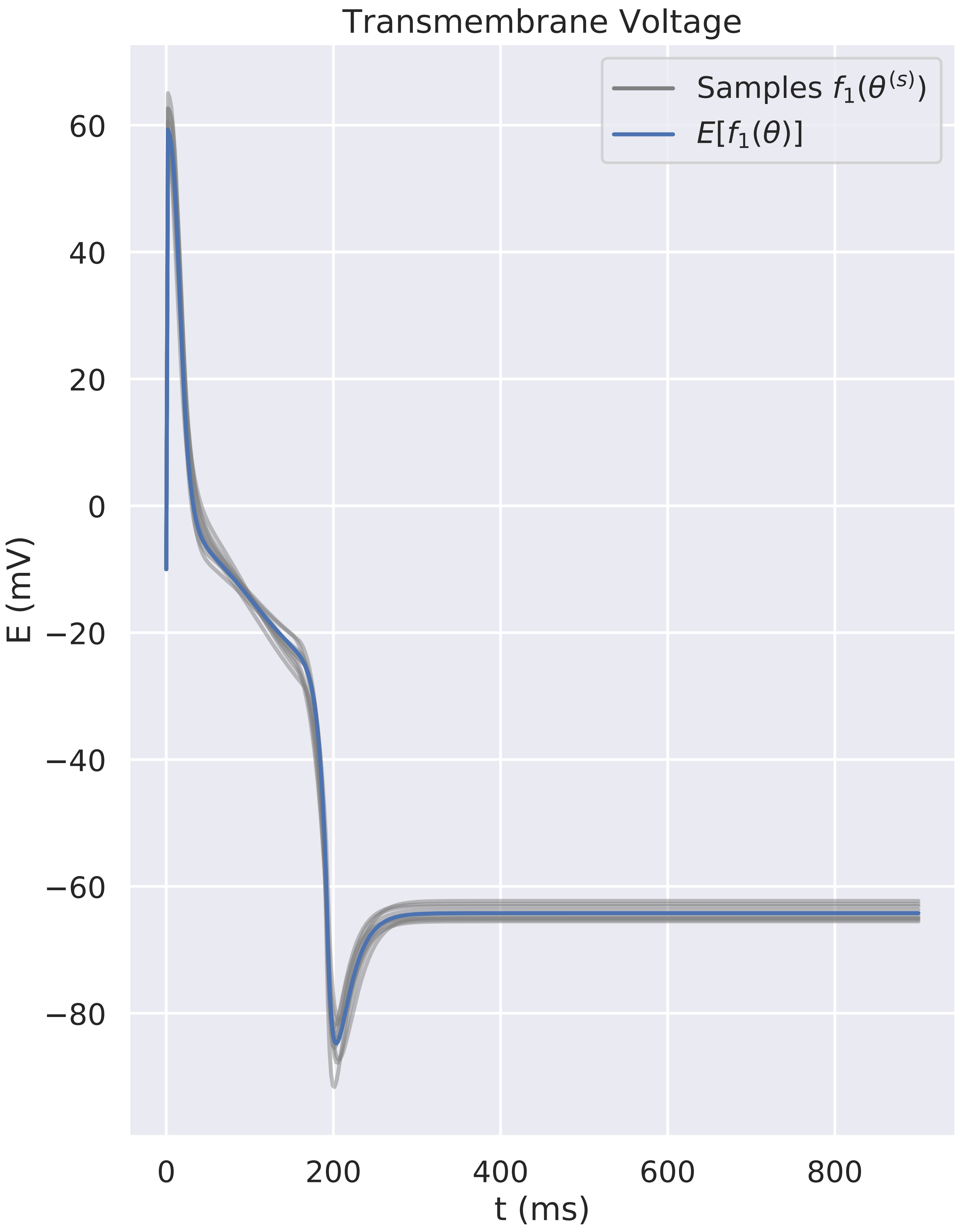


Bayesian Inference

$$E[f(\theta)] = \int f(\theta)p(\theta \mid D)d\theta$$

Approximate with S samples $\theta^{(s)}$ drawn from $p(\theta \mid D)$

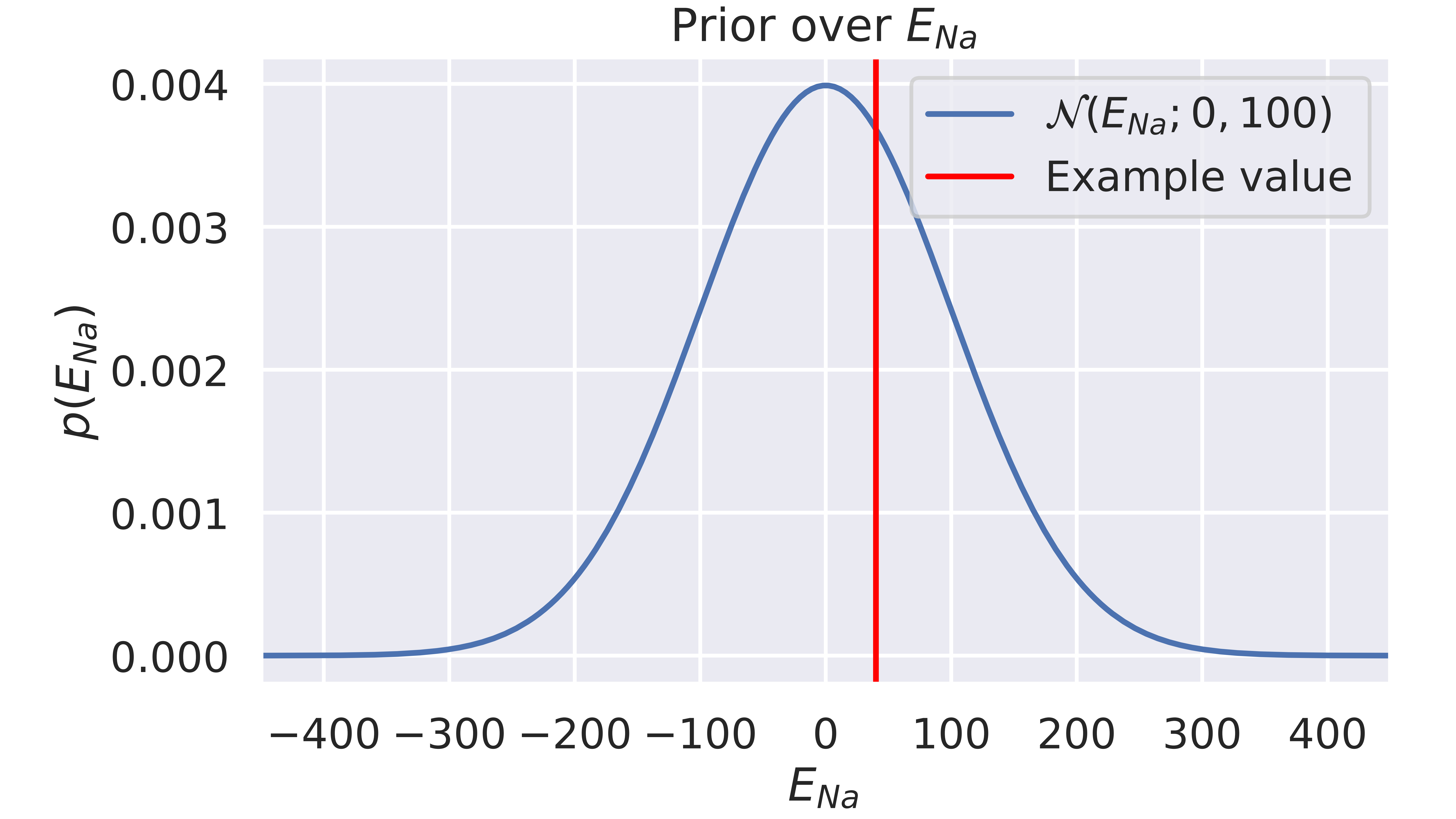
$$E[f(\theta)] \approx \frac{1}{S}\sum_{s=1}^{S} f(\theta^{(\mathbf{s})})$$



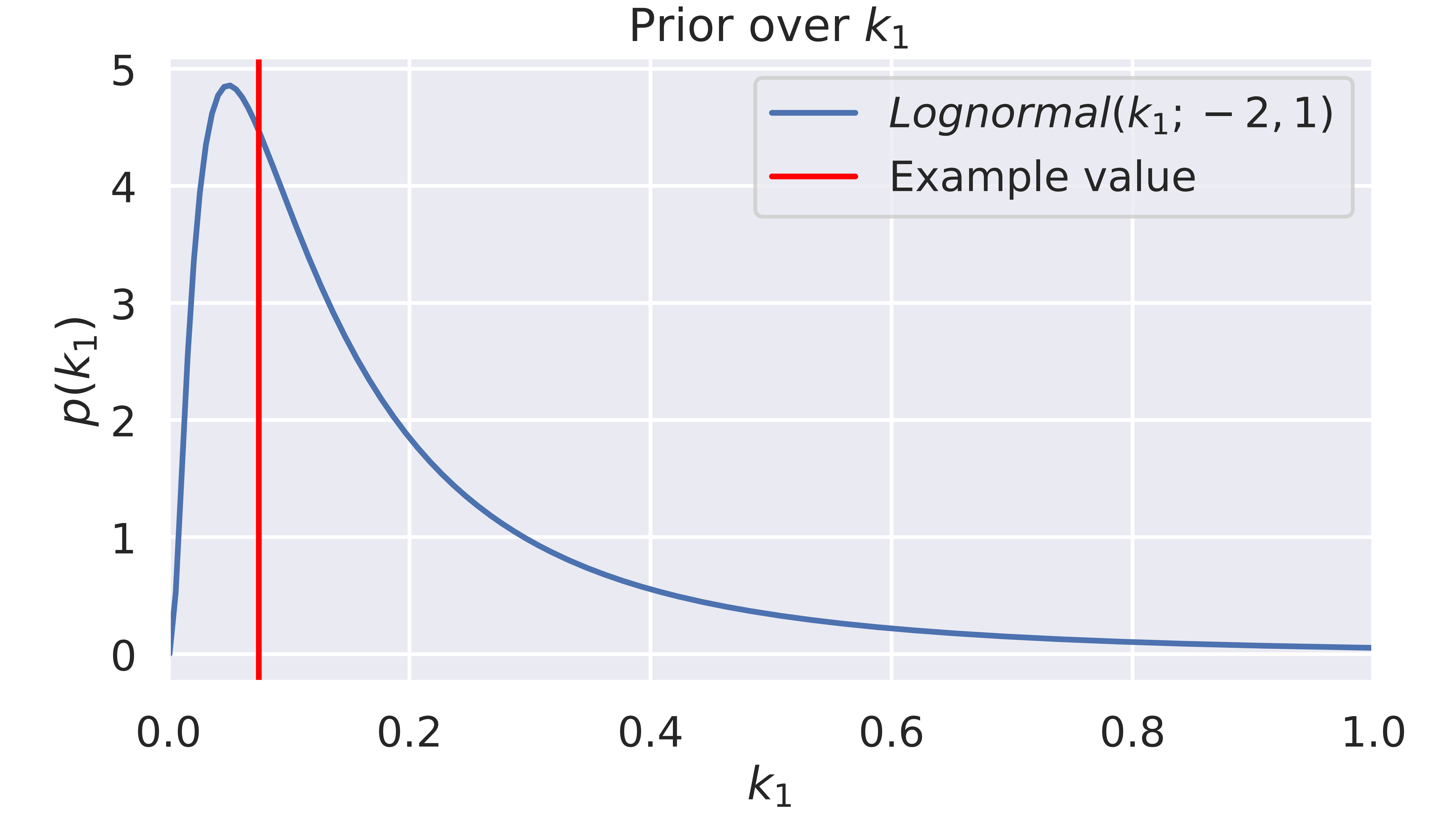
Priors

- Need to choose $p(\theta)$
- Read the paper![1]
- Types
  - Unconstrained parameters, e.g. E_{Na}
  - Constrained positive/negative: e.g. $k_1 > 0$
  - Ordered parameters, e.g. $E_1 < E_\dagger < E_\star$

[1] Simitev, R.D. and Biktashev, V.N., 2011. Asymptotics of Conduction Velocity Restitution in Models of Electrical Excitation in the Heart. Bulletin of Mathematical Biology, 73(1), pp.72-115.



Prior over E_{Na}
$\mathcal{N}(E_{Na}; 0, 100)$
Example value
0.004
0.003
0.002
0.001
0.000
$p(E_{Na})$
−400
−300
−200
−100
0
100
200
300
400
E_{Na}



Prior over k_1
$Lognormal(k_1; -2, 1)$
Example value
$p(k_1)$
k_1
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0

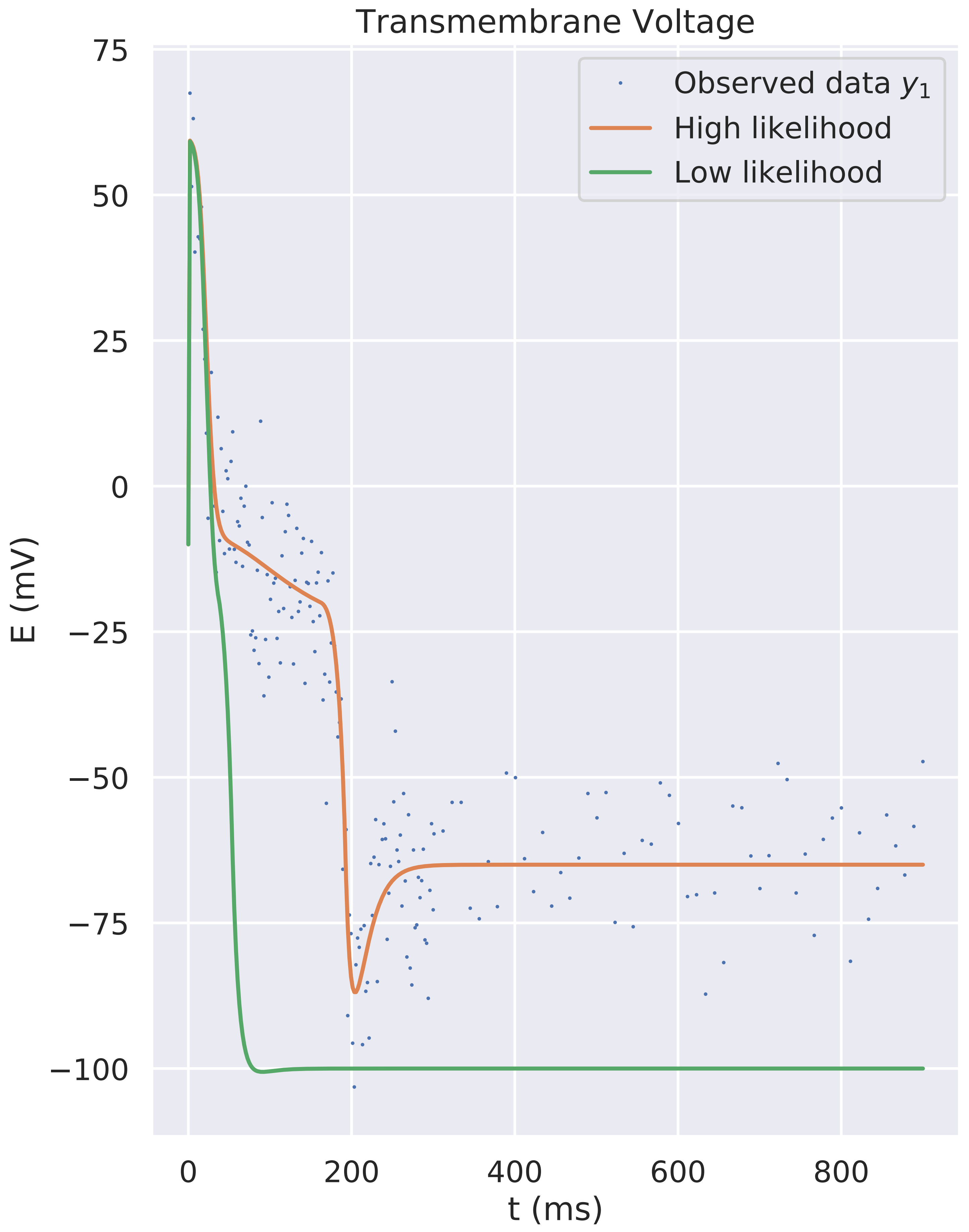
Priors on ordered parameters

- E.g. $E_1 < E_\dagger < E_\star$
- Parameterise as $E_1, \log(E_\dagger - E_1), \log(E_\star - E_\dagger)$
- Normal prior on E_1 as before
- Log-normal prior on $E_\dagger - E_1, E_\star - E_\dagger$ as before

Likelihood

- Run the provided solver for given θ
- Get outputs $f_j(\theta)^{(t)}$ for each signal j over the timeseries
- Gaussian log-likelihood

$$\log p(D \mid \theta, \sigma) = \sum_{j=1}^{3} \sum_{t=1}^{T} \log \mathcal{N}(y_j^{(t)}; f_j(\theta)^{(t)}, \sigma_j)$$



Error handling

```
Warning: Failure at t=3.000000e+01.  Unable
to meet integration tolerances without
reducing the step size below the smallest
value allowed (1.136868e-13) at time t.
```

Just return a likelihood of 0!

Posterior

$$\log p(\theta \mid D) = \log p(D \mid \theta) + \log p(\theta) + \text{constant}$$

— Could pass this to an MCMC tool

— First we need to find a starting point...

Maximum A Posteriori (MAP) Solution

$$\theta_{\mathrm{MAP}} = \arg\max_{\theta} \log p(\theta \mid D)$$

— Not differentiable[2].

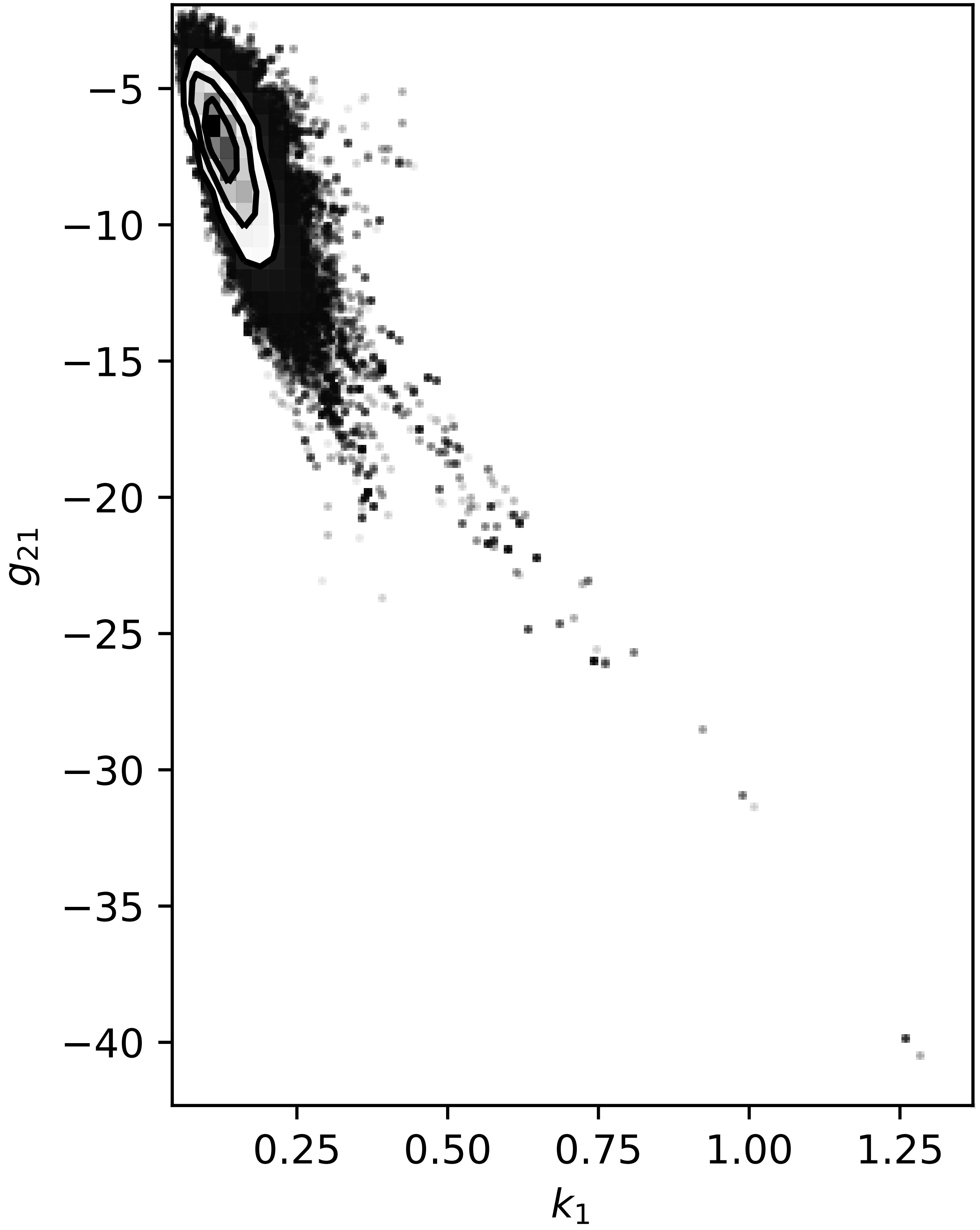
— Use Powell's method with multiple restarts

— Not Bayesian

[2] At least, not in provided implementation.

Markov Chain Monte Carlo (MCMC) Methods

- Generate S samples from posterior
- Use MCMC methods
- Strong correlations in posterior
- Use `emcee`[3]



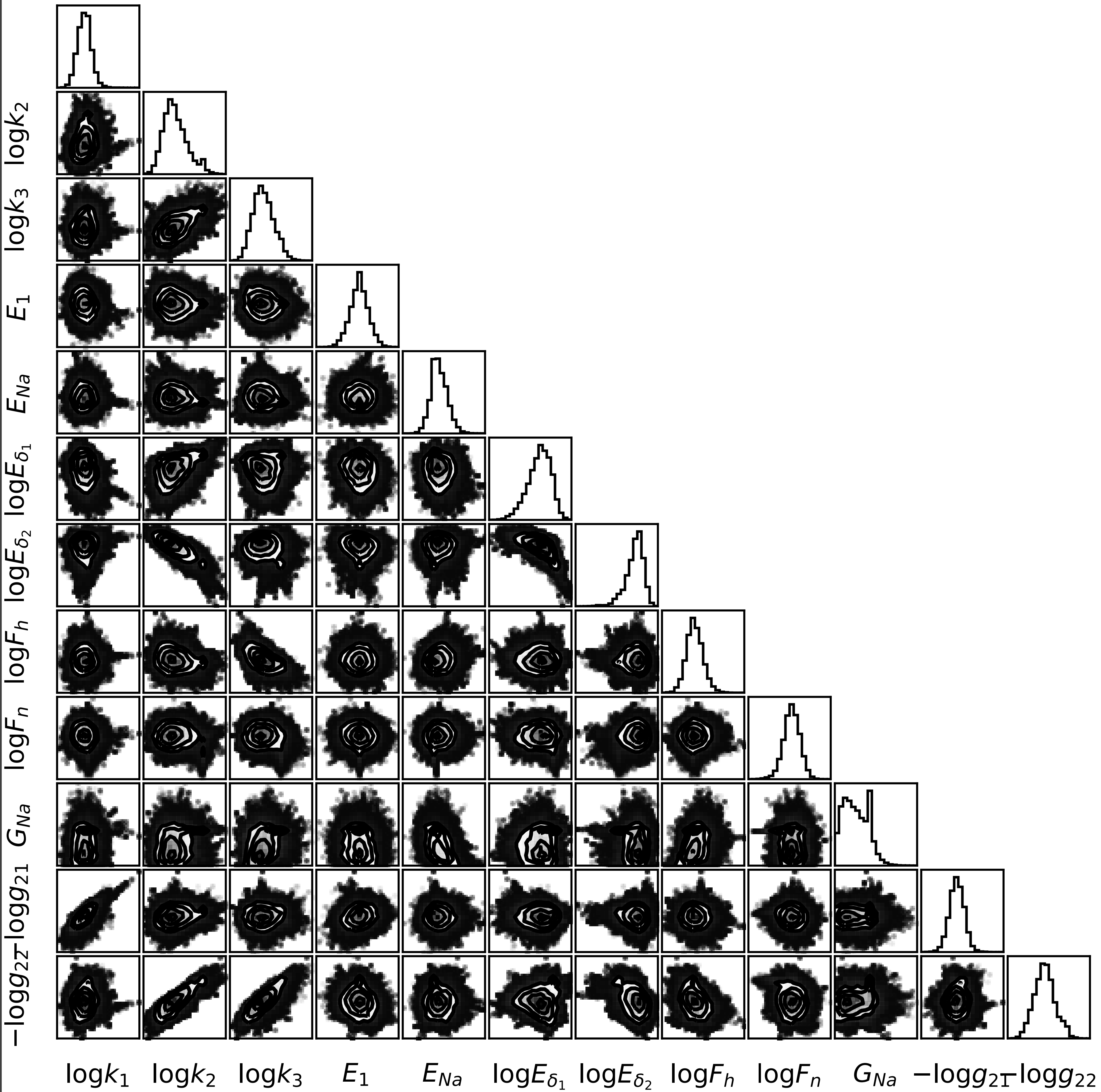
[3] dfm.io/emcee

emcee

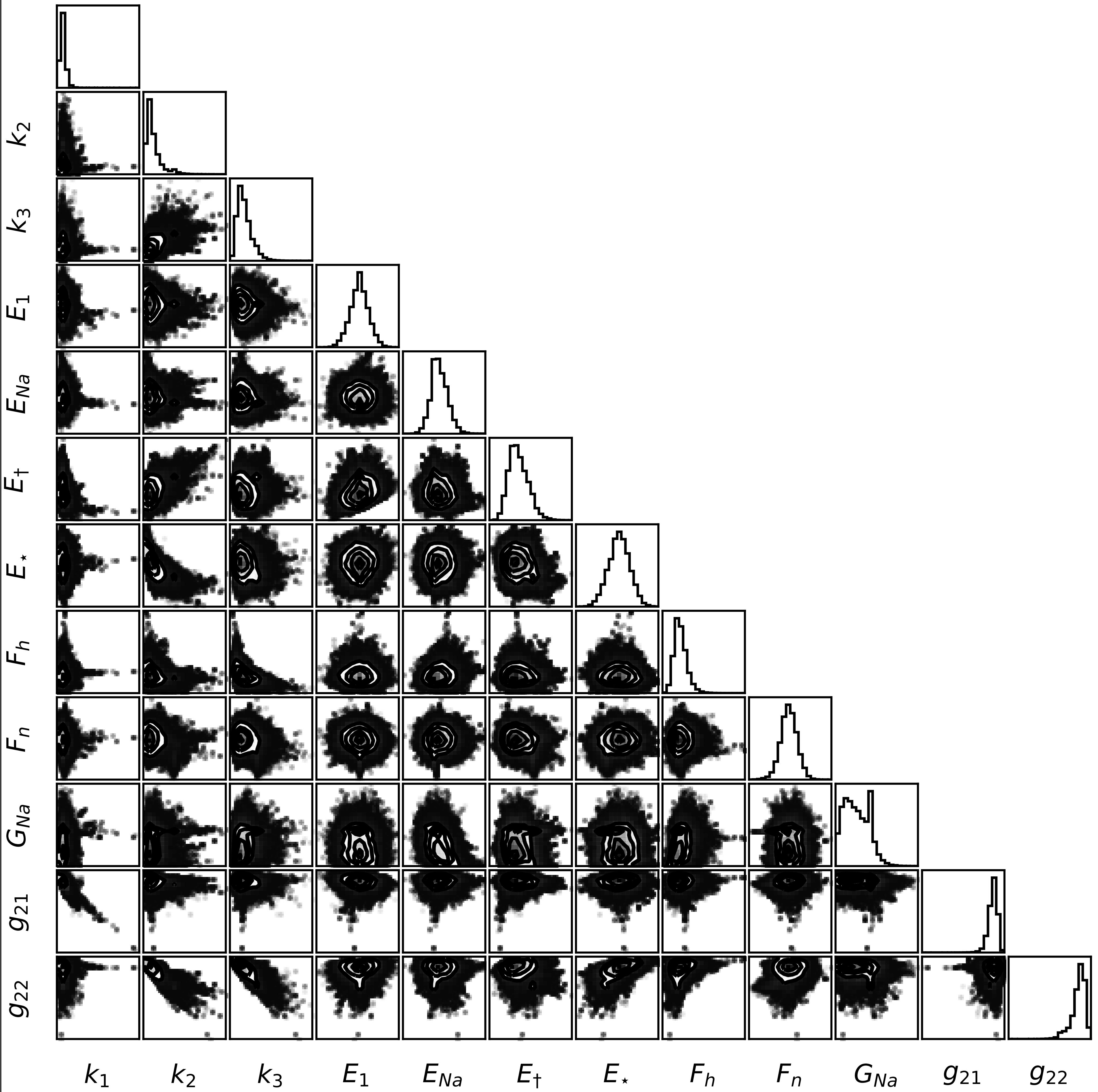
- Open source Python package
- Implements affine-invariant sampling[4]
  - Run many MCMC chains in parallel
  - Propose new samples based on other chains
- Ran for 10,000 steps with 100 chains each
- Discard first half of chain

---

[4] Goodman, J. and Weare, J., 2010. Ensemble Samplers With Affine Invariance. Communications in Applied Mathematics and Computational Science, 5(1), pp.65-80.



$\log k_2$
$\log k_3$
E_1
E_{Na}
$\log E_{\delta_1}$
$\log E_{\delta_2}$
$\log F_h$
$\log F_n$
G_{Na}
$-\log g_{21}$
$-\log g_{22}$
$\log k_1$
$\log k_2$
$\log k_3$
E_1
E_{Na}
$\log E_{\delta_1}$
$\log E_{\delta_2}$
$\log F_h$
$\log F_n$
G_{Na}
$-\log g_{21}$
$-\log g_{22}$

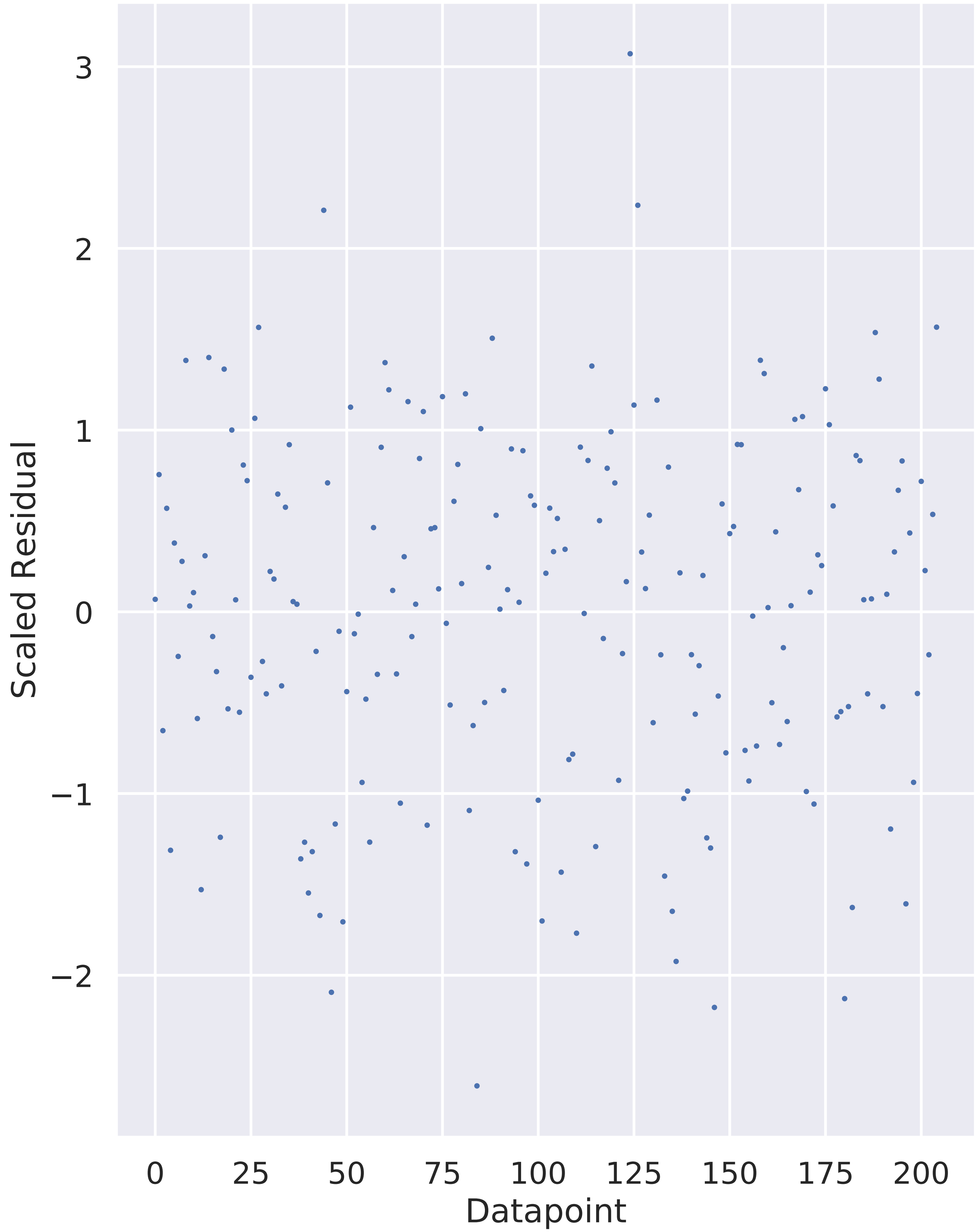


k_2
k_3
E_1
E_{Na}
$E_\dagger$
$E_\star$
F_h
F_n
G_{Na}
g_{21}
g_{22}
k_1
k_2
k_3
E_1
E_{Na}
$E_\dagger$
$E_\star$
F_h
F_n
G_{Na}
g_{21}
g_{22}

Procedure checking

1. Check convergence[5]
2. Check we can recover example parameters
3. Check other parameter settings
4. Check scaled residuals,

$$\frac{y_j^{(t)} - E[f_j(\theta)^{(t)}]}{\sigma_j}$$



[5] Gelman, A., Stern, H.S., Carlin, J.B., Dunson, D.B., Vehtari, A. and Rubin, D.B., 2013. Bayesian Data Analysis. Chapman and Hall/CRC.

Competition Entry

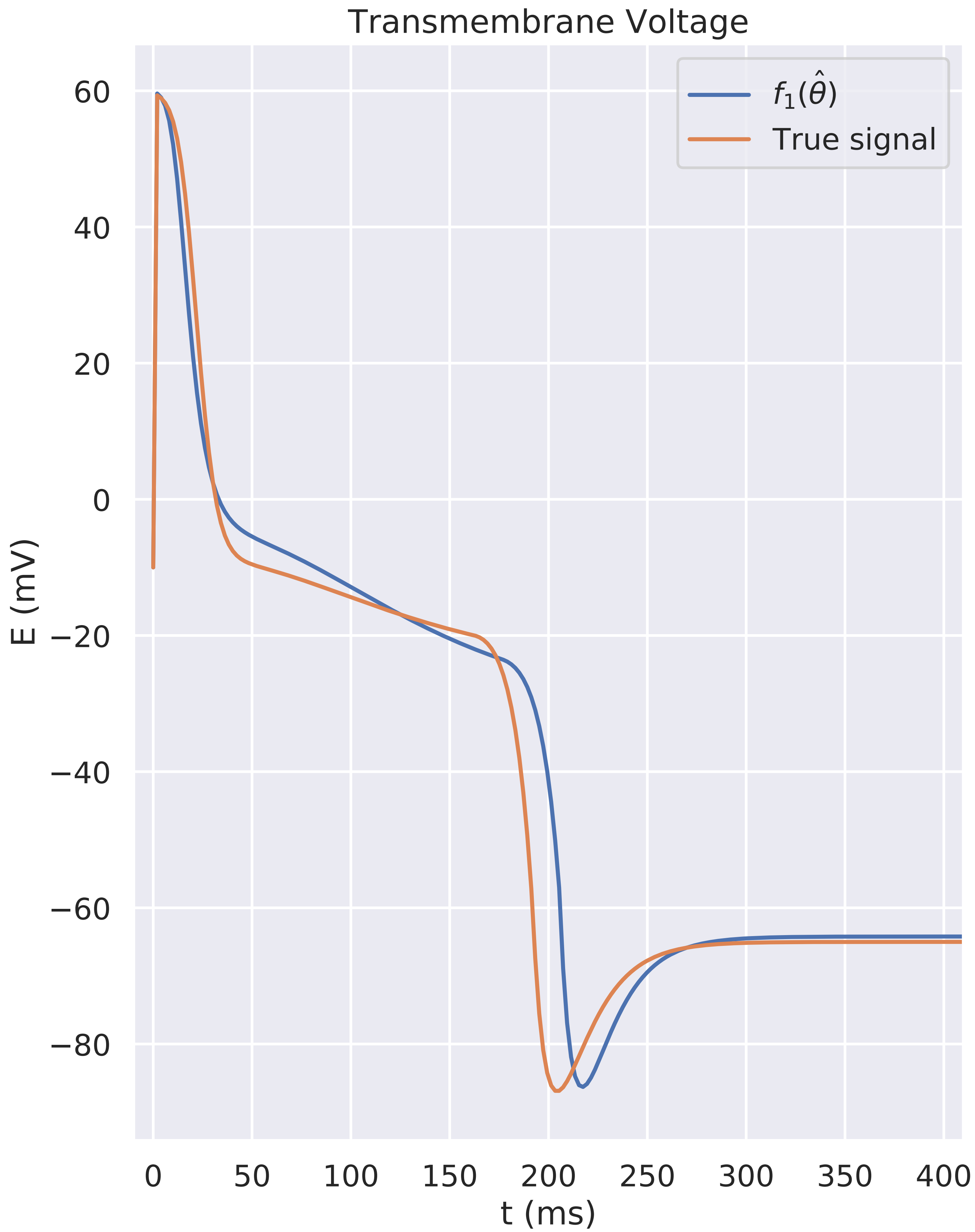
Tempting to submit MAP estimate

Instead submit sample mean

$$\hat{\theta} = \frac{1}{S}\sum_{s=1}^{S}\theta^{(s)}$$

Evaluate sample covariance similarly

Submit the mean output of the ODE, $E[f(\theta)]$, not $f(\hat{\theta})$!



Problems

1. Unstable Covariance
2. Slow
3. Can't handle multi-modal posteriors
4. Won't scale to high dimensional θ

Potential Improvements

1. Rescale the parameters
2. Don't evaluate the likelihood directly
3. Use Parallel-Tempered Ensemble Sampling
4. Use more scalable MCMC algorithms

Recommendations

1. Choose appropriate parameterisation
2. Find a good initialisation
3. Use tuning-free algorithms
4. Start with generic methods
5. Think about expectations you need