Calibrating Tension Statistics with Neural Ratio Estimators

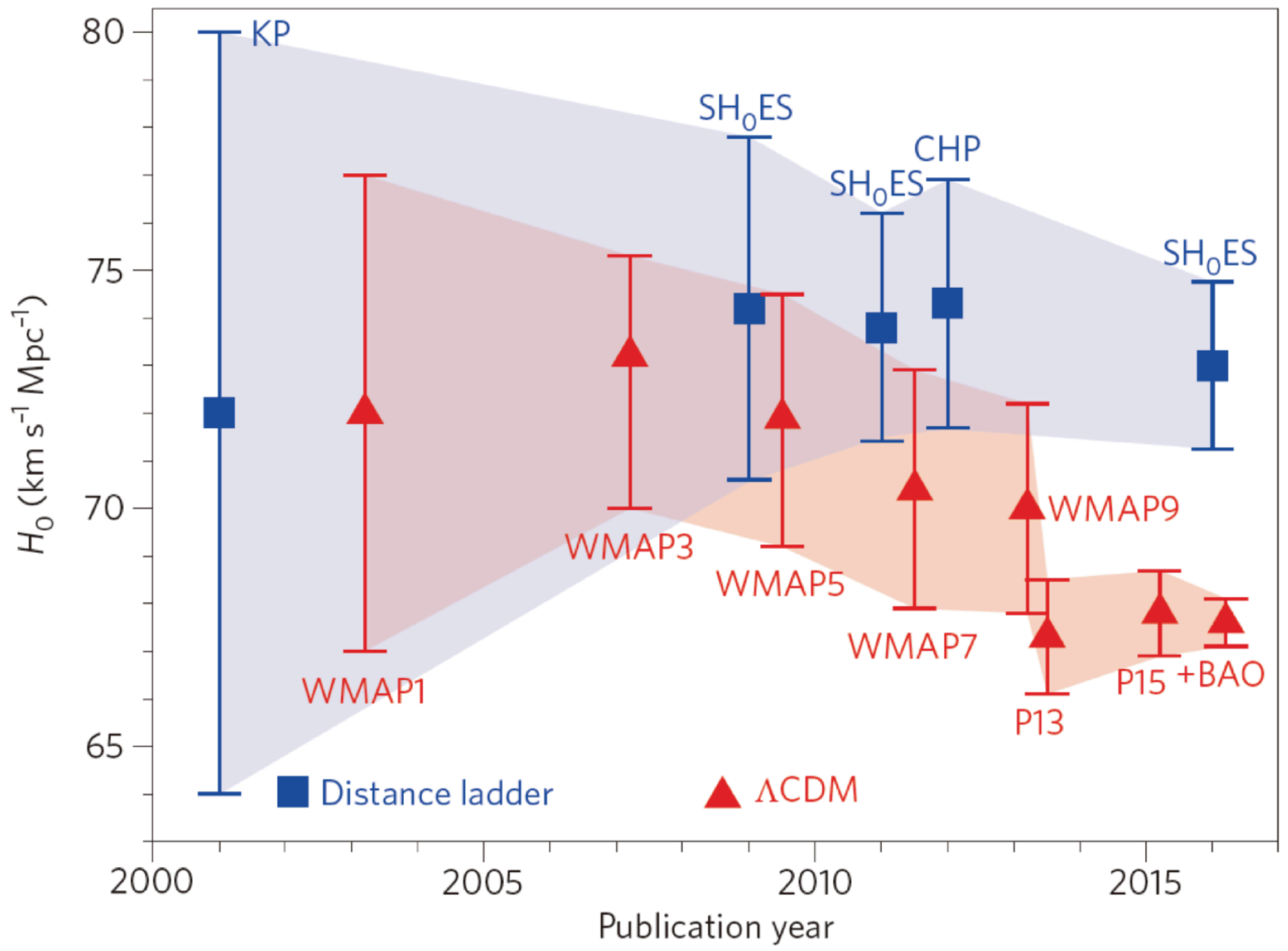
Harry Bevins
with Thomas Gessey-Jones and Will Handley
University of Cambridge

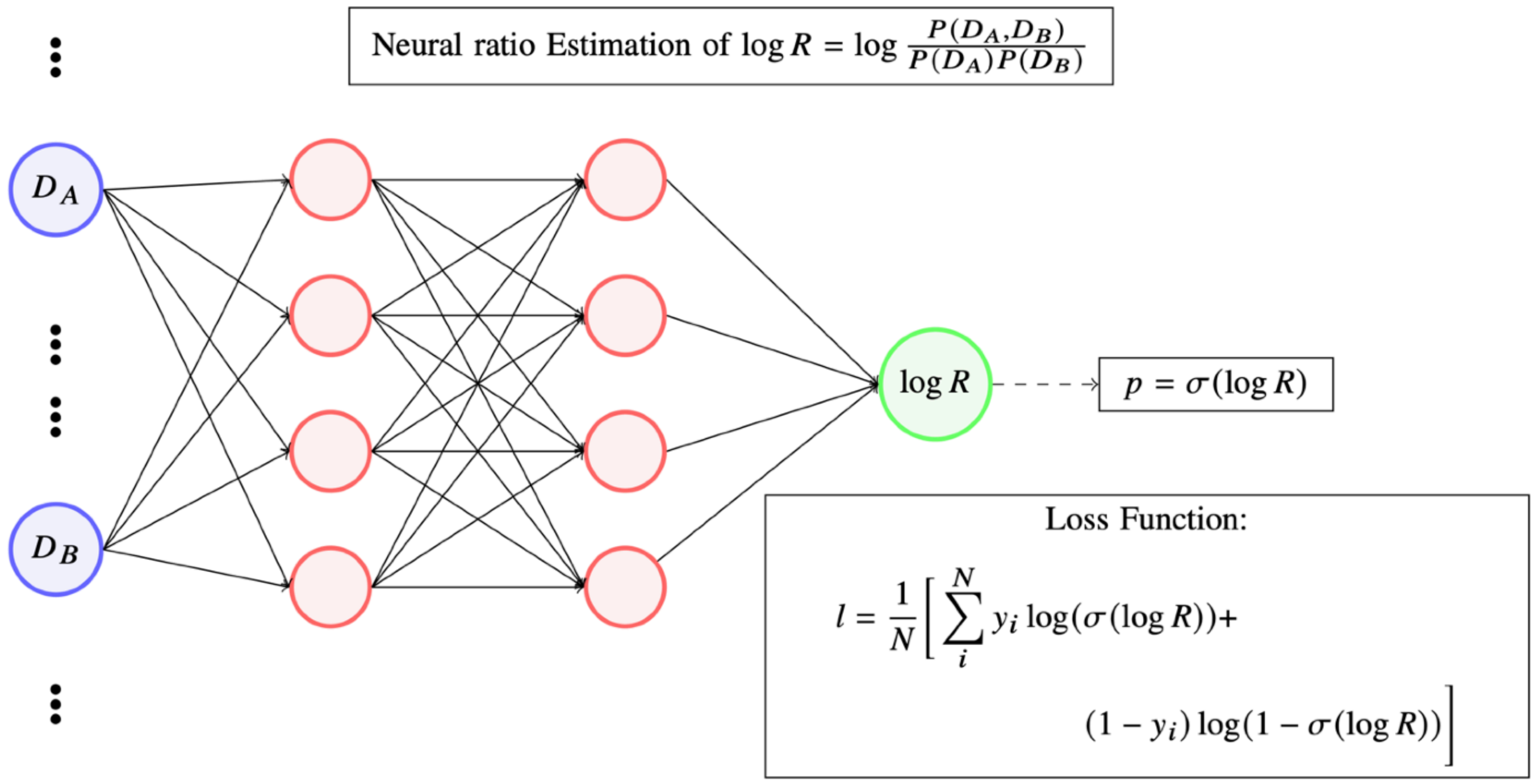
Calibrating Tensions

1. Why are we interested and how do we measure tension?

2. Calibrating with Neural Ratio Estimation

3. Demonstrations

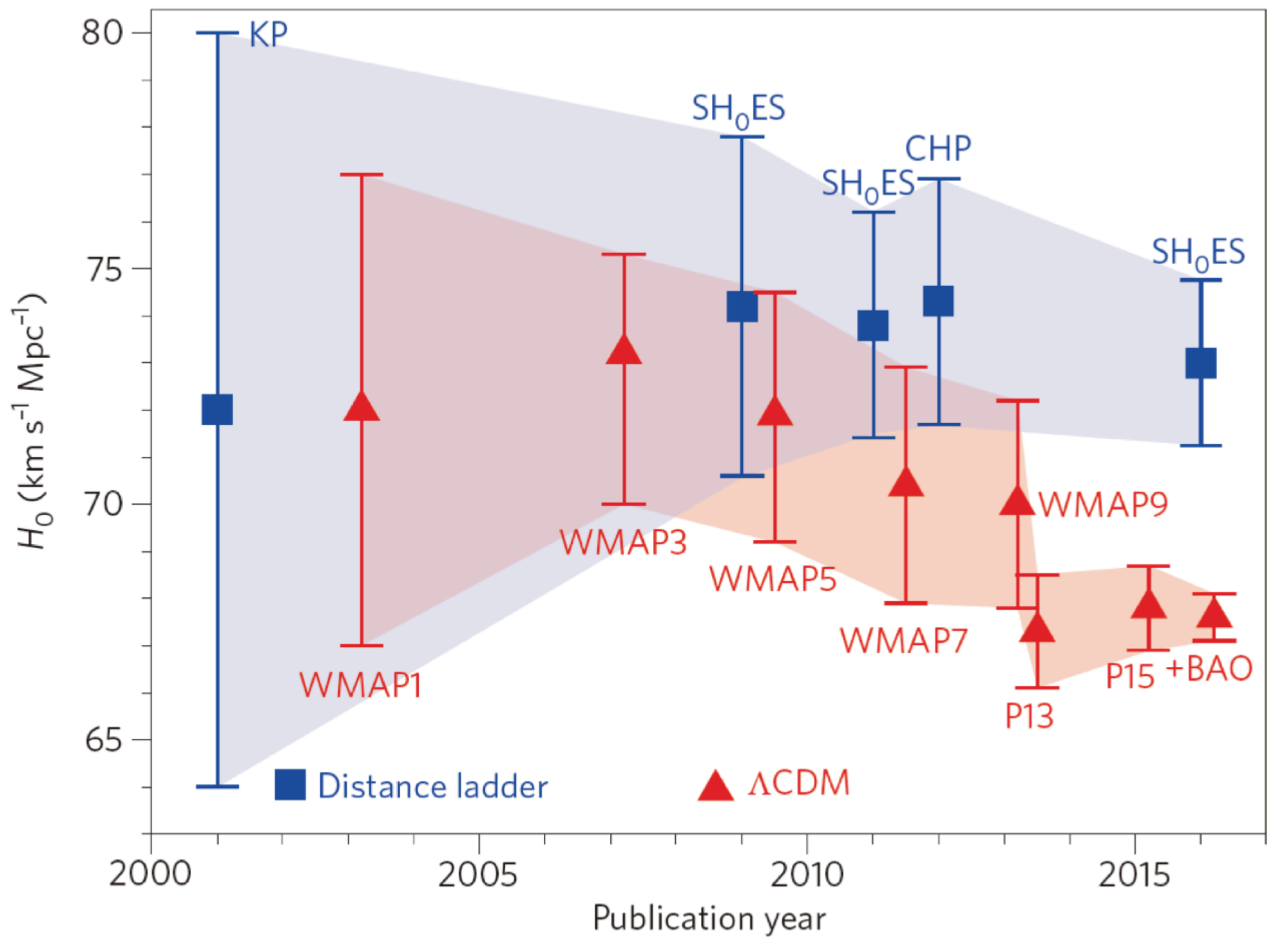


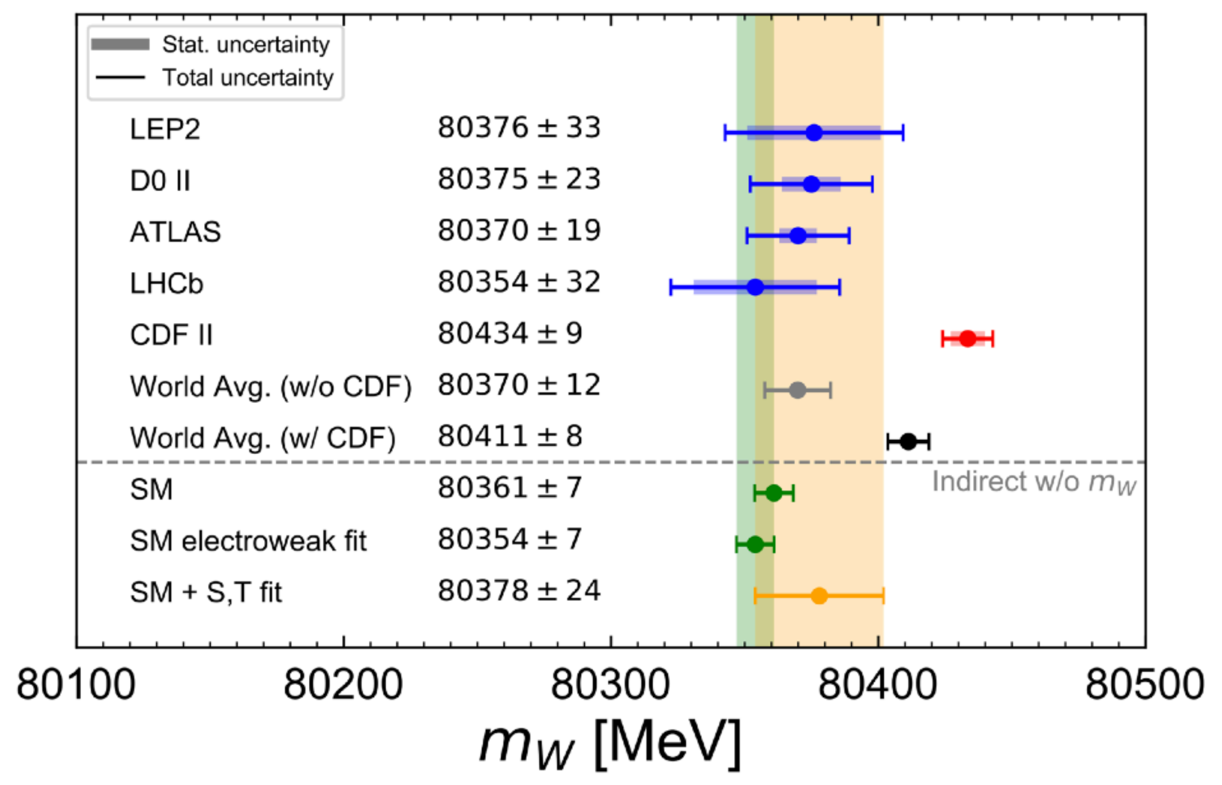


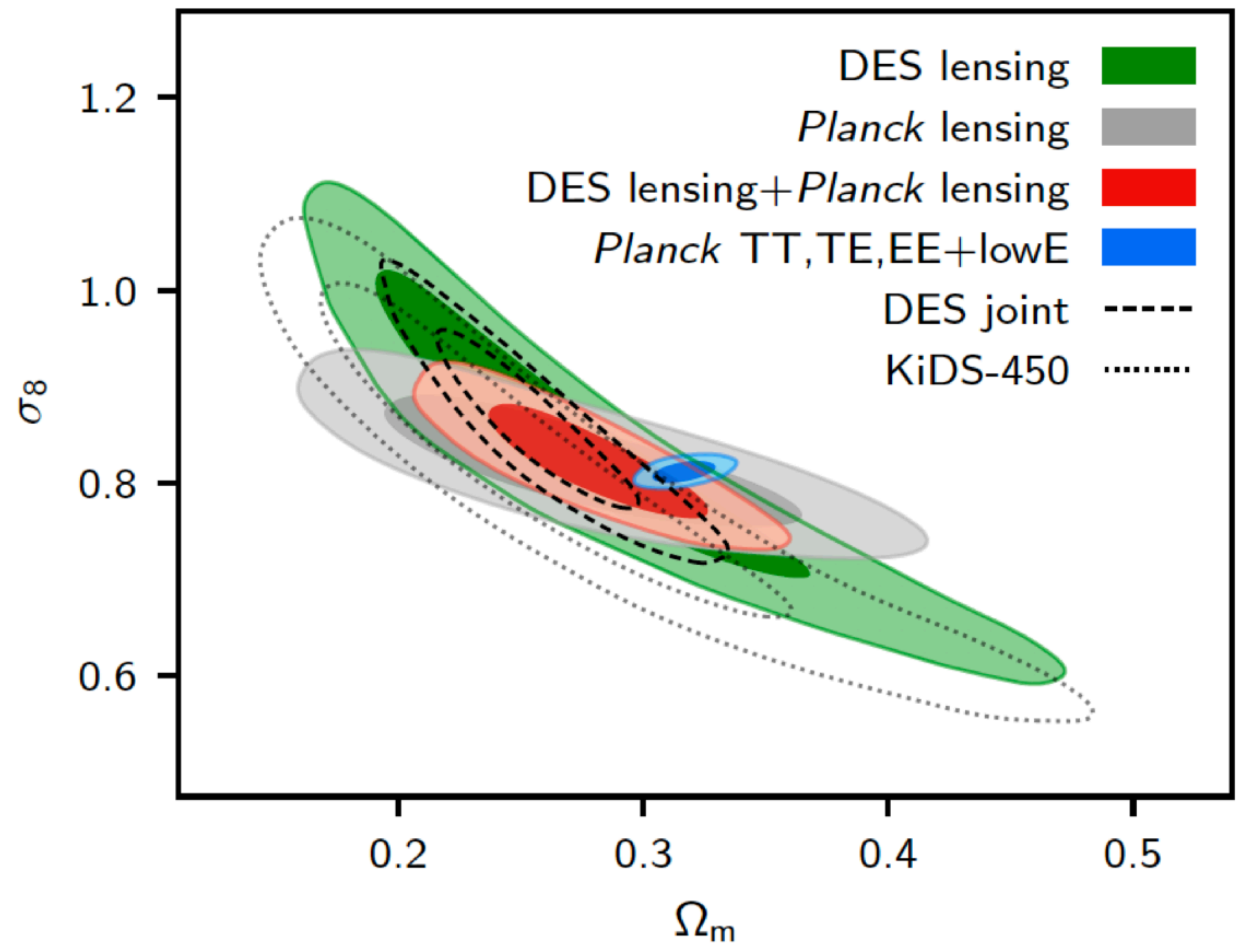
Why are we interested?

Why are we interested in tension?

- Important to be able to independently observe and confirm experimental results
- When two experiments give different results we call this a tension
- Understanding where tension comes from can lead us to new physics and a better understanding of our instruments







[2204.0526]
[1905.11122]
[1807.06209]

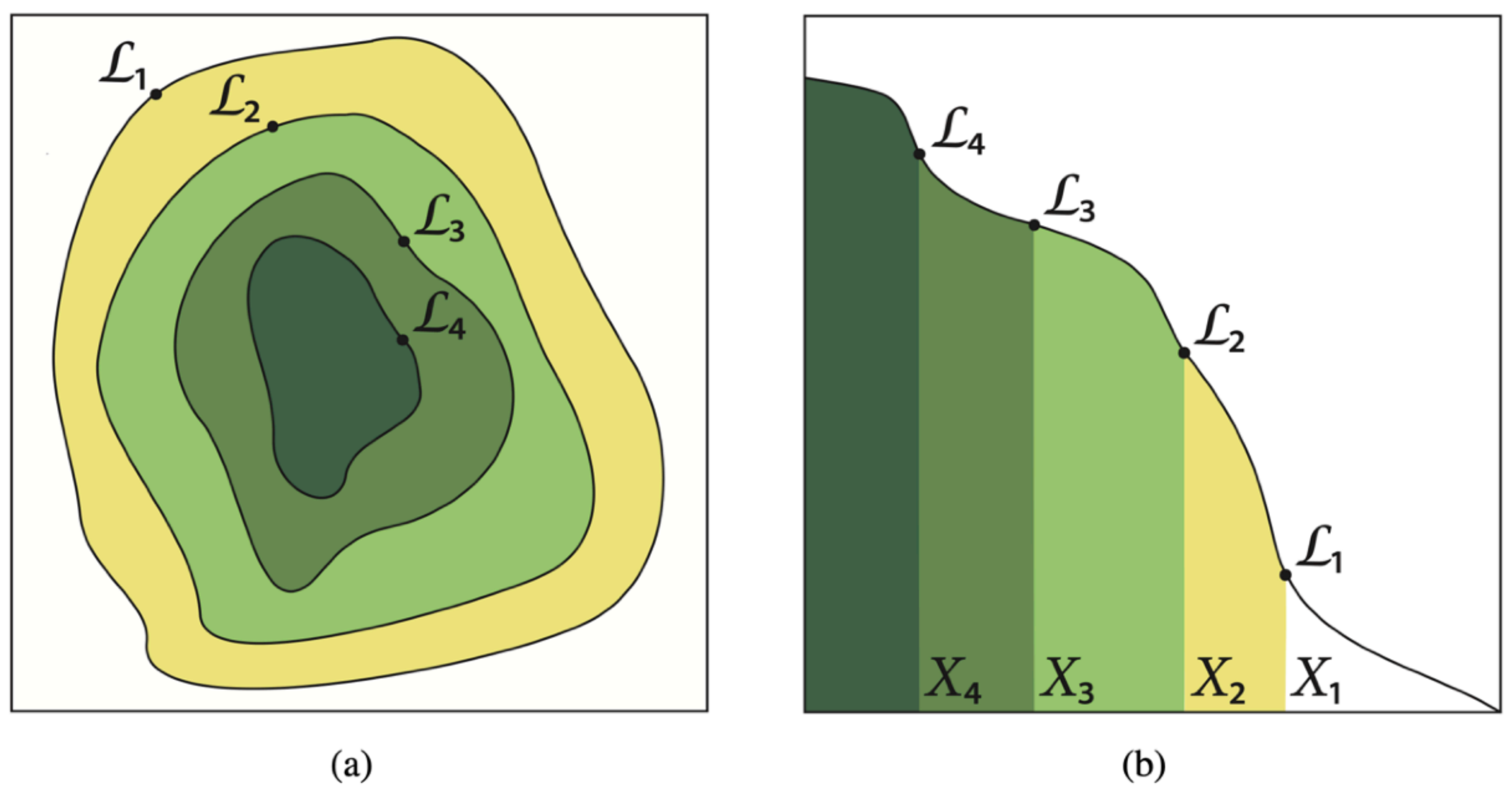
Measuring tension

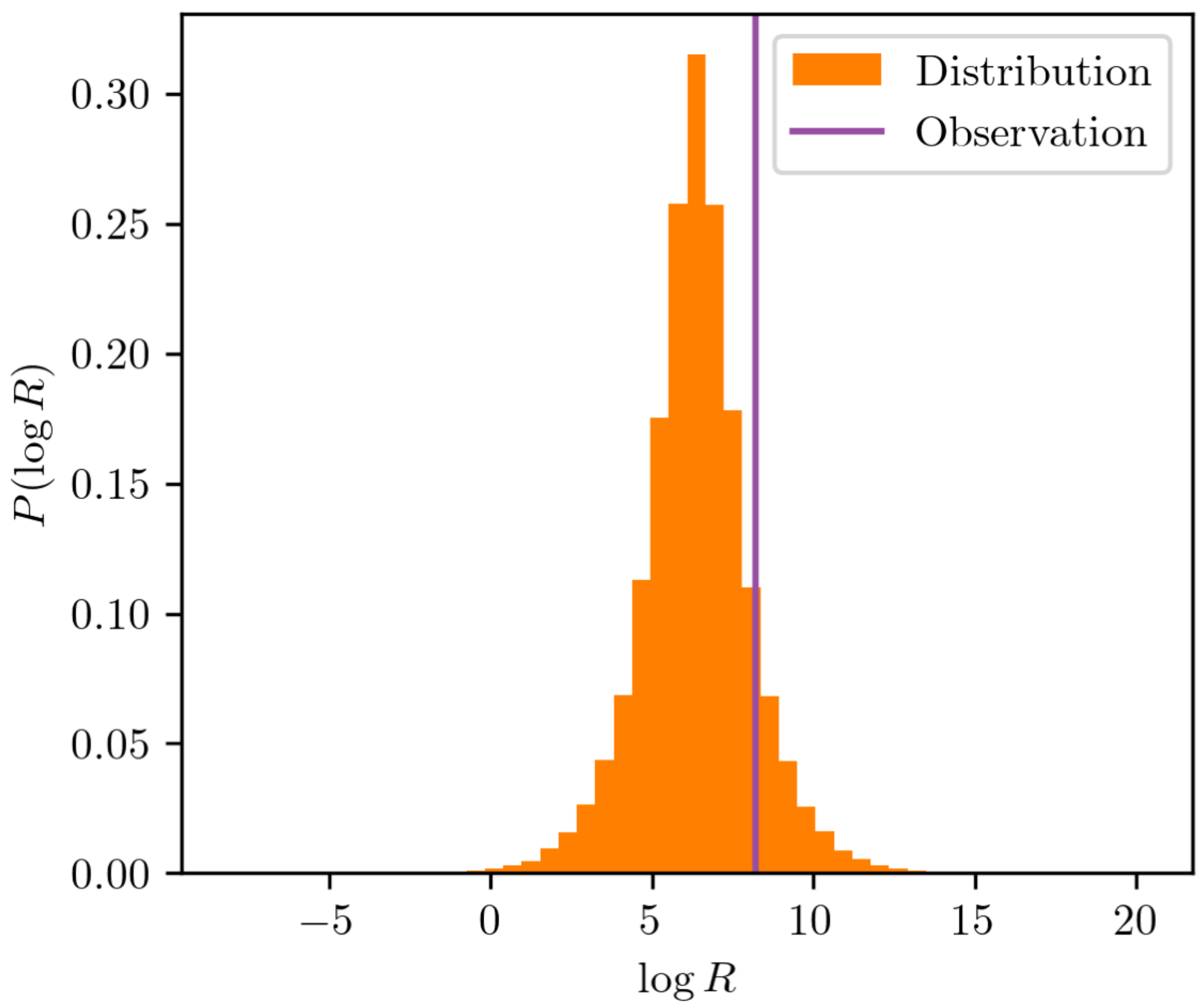
- Parameter differences, goodness of fit degradation, suspiciousness (see 2012.09554 for a review)

- Here, interested in evidence ratio

$$R = \frac{P(D_A, D_B)}{P(D_A)P(D_B)} = \frac{Z_{AB}}{Z_A Z_B}$$

- For any pair of experiments, model and prior there is a distribution of in concordance R values





Measuring tension

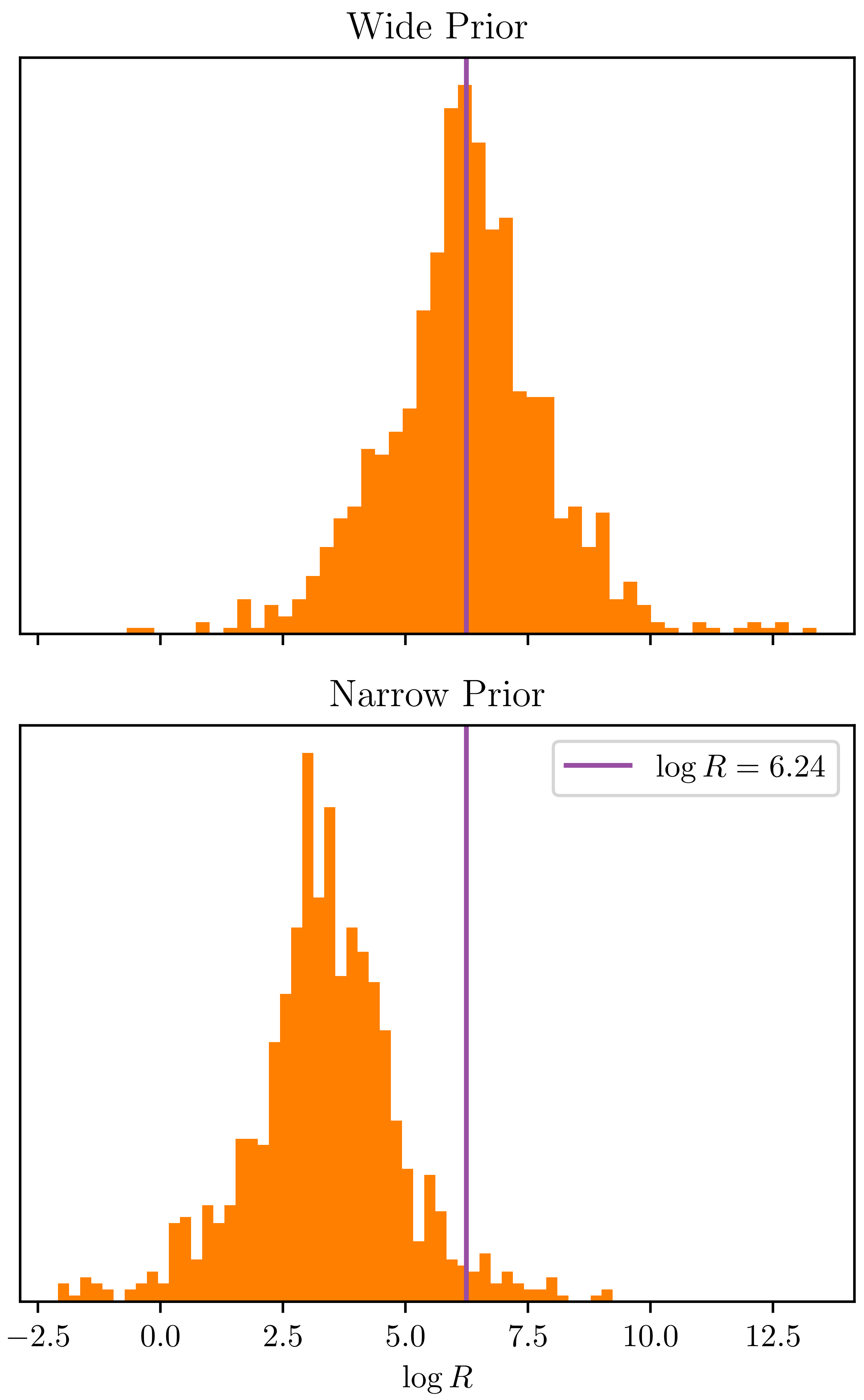
- The fractional increase in our confidence in one experiment given data from another

- Dimensionally consistent and parameterisation invariant

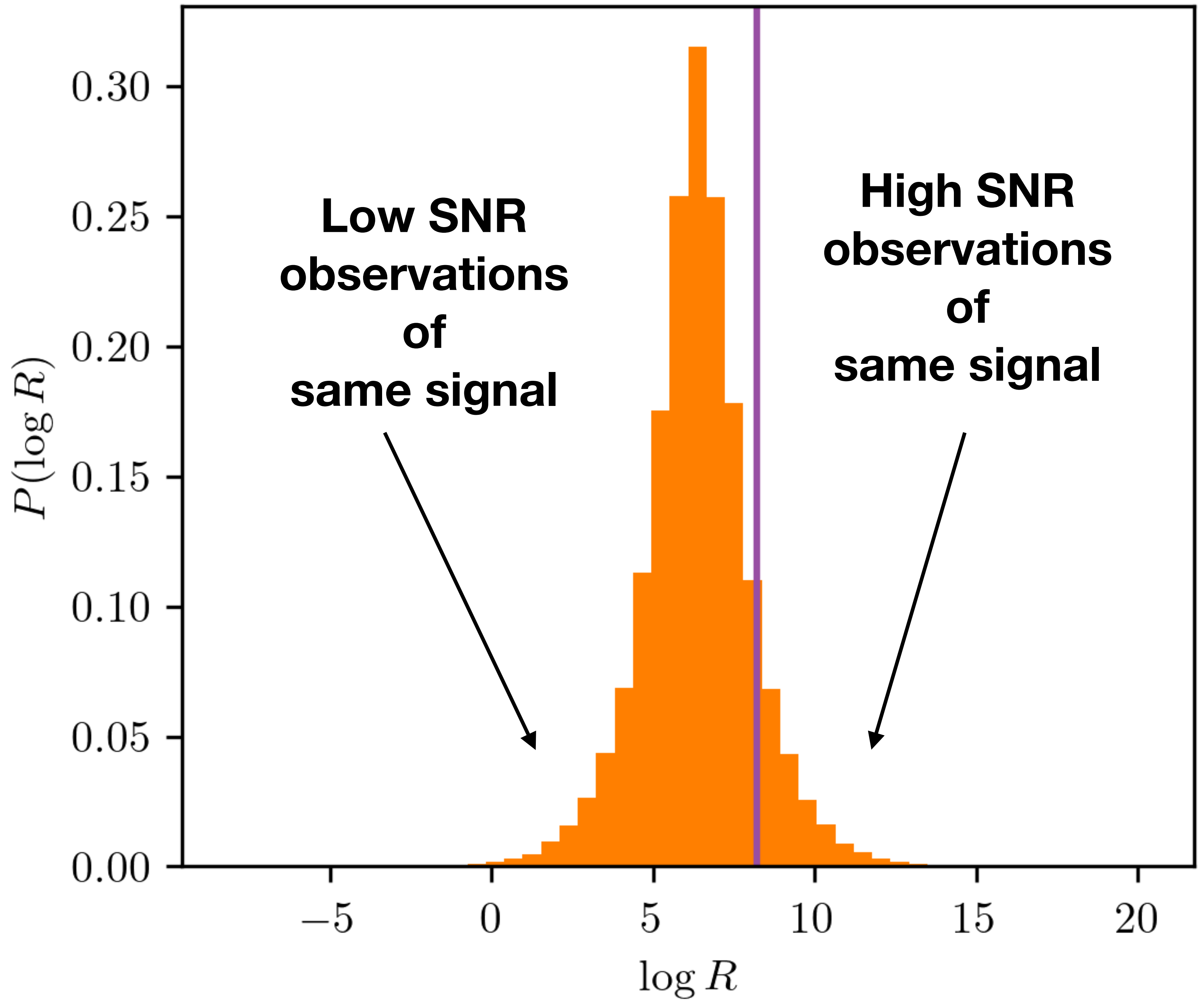
- But prior dependent and hard to interpret

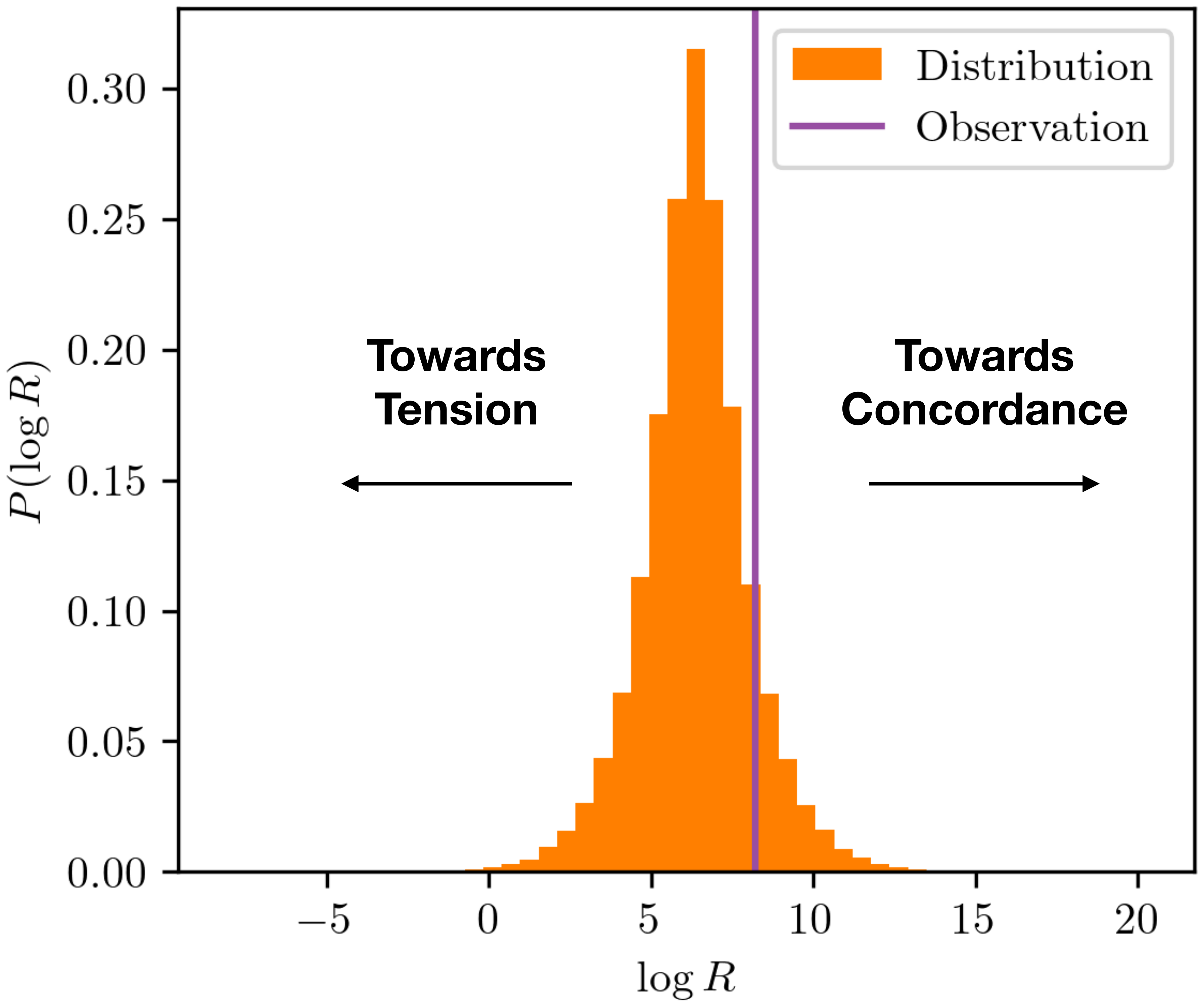
  - $R \gg 1 \rightarrow$ in concordance

  - $R \ll 1 \rightarrow$ in tension



Measuring tension





Calibrating with Neural Ratio Estimation

Neural Ratio Estimation

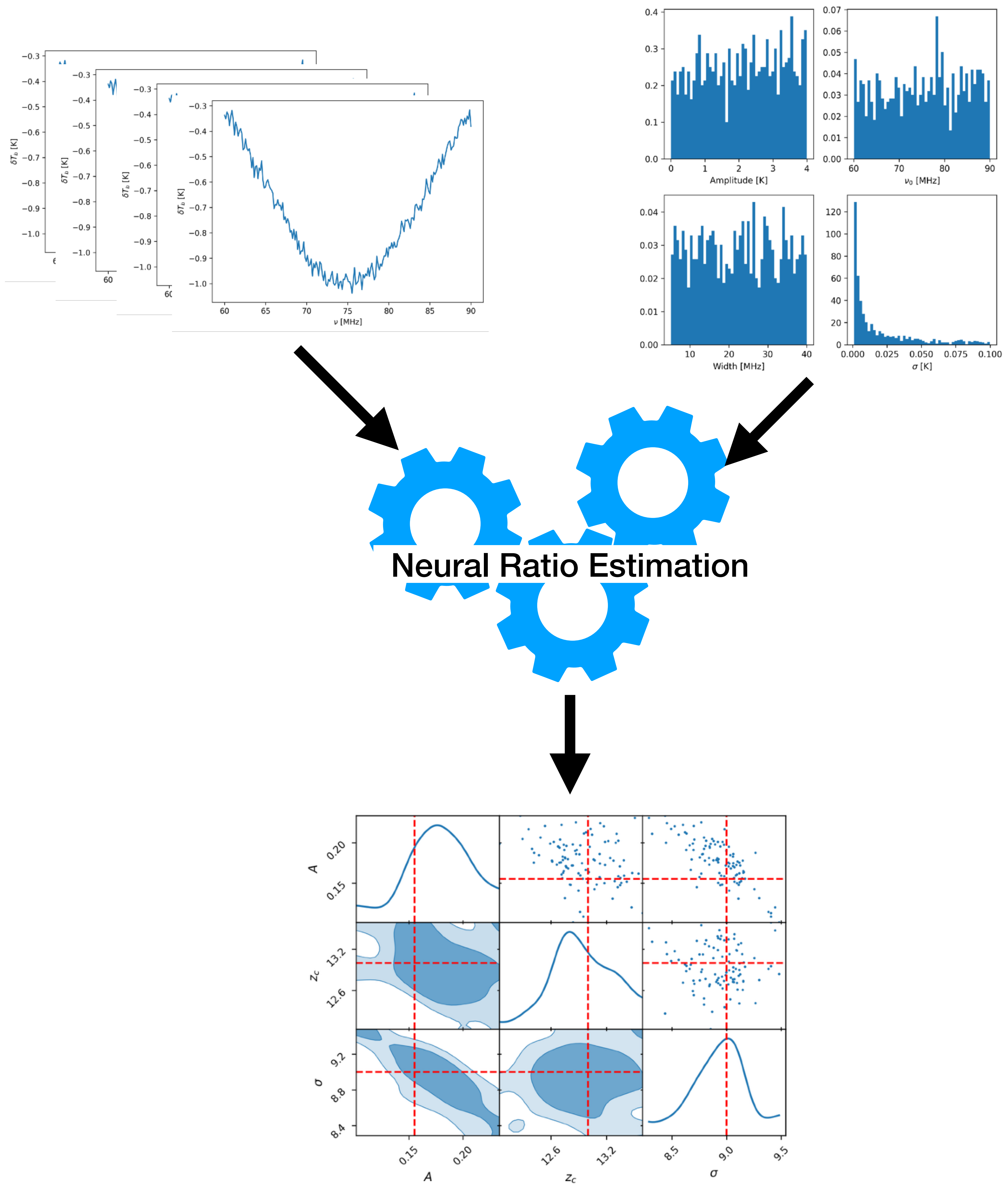
- Essentially just classifiers

- Take in two inputs A and B and estimate the probability that they are drawn from joint distribution vs independent

$$r = \frac{P(A, B)}{P(A)P(B)}$$

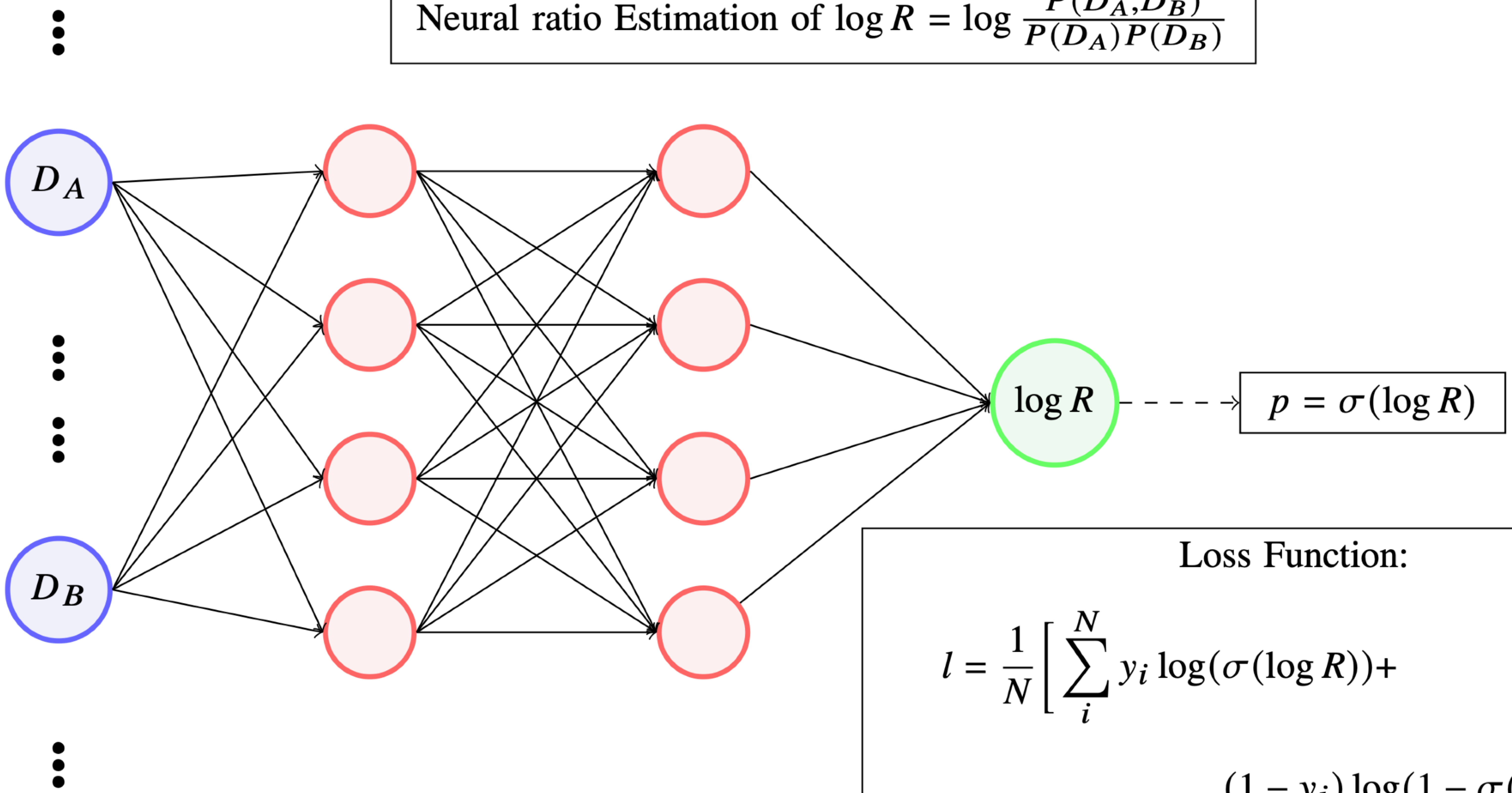
- Used for parameter inference

$$r = \frac{P(D, \theta)}{P(D)P(\theta)} = \frac{P(D|\theta)}{P(D)} = \frac{L(\theta)}{Z}$$



R with NREs

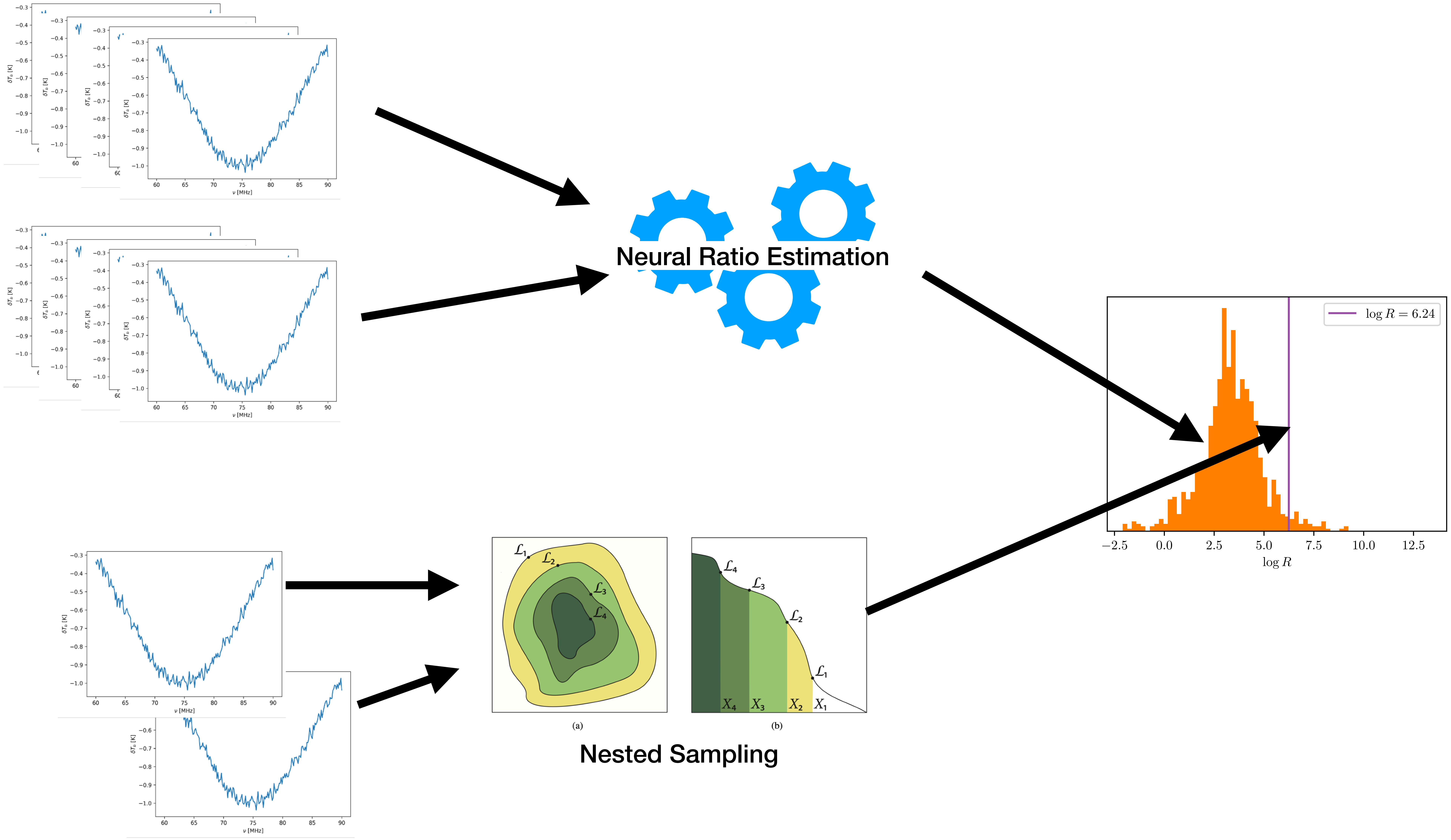
Neural ratio Estimation of $\log R = \log \frac{P(D_A, D_B)}{P(D_A)P(D_B)}$



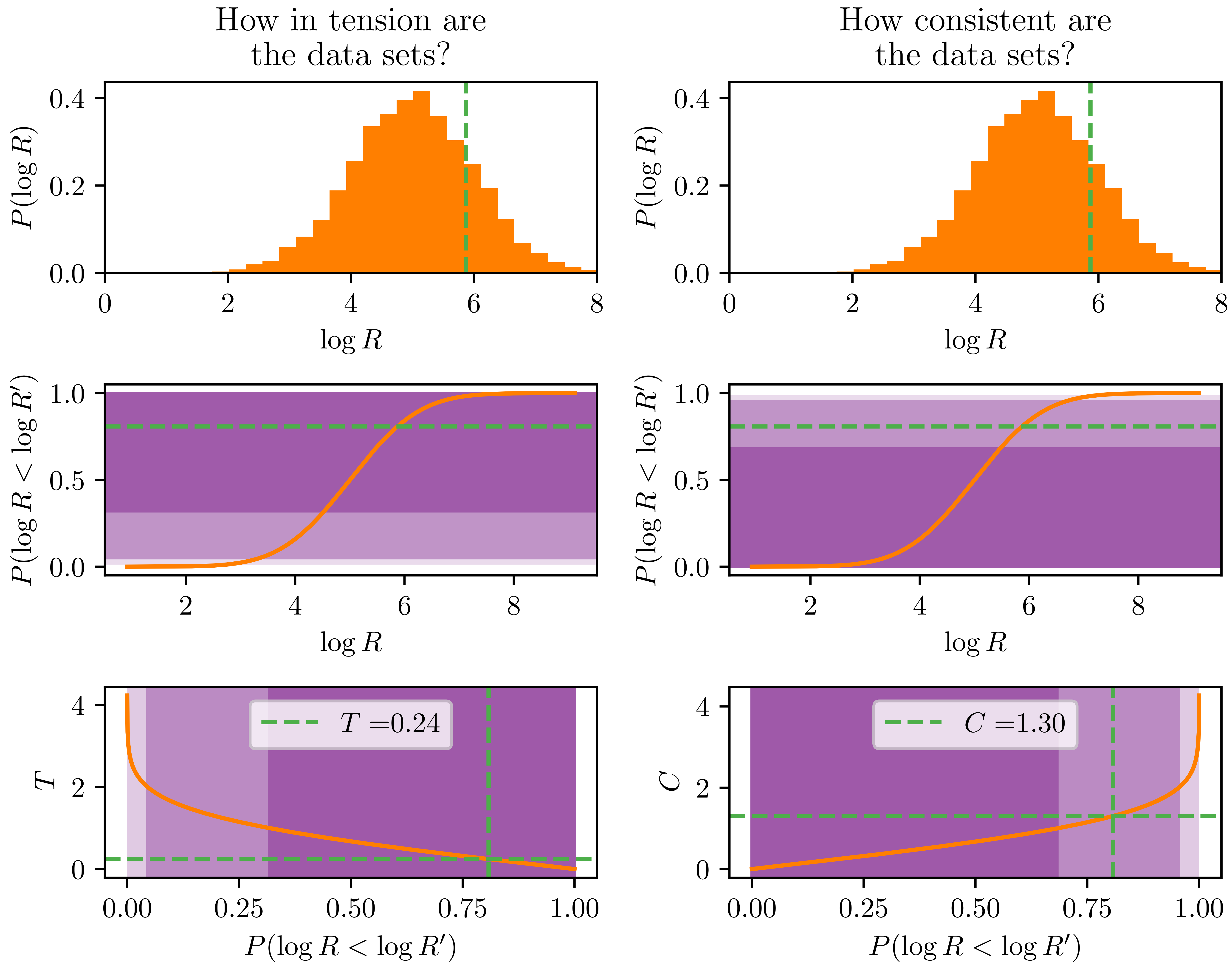
Loss Function:

$$l = \frac{1}{N}\left[\sum_i^N y_i \log(\sigma(\log R)) + (1 - y_i)\log(1 - \sigma(\log R))\right]$$

Direct predictions or calibration?



Calibration of R



Examples

Analytic Example: Set Up

- Define a linear model

$$D_A = M_A\theta + m_A \pm \sqrt{C_A}$$

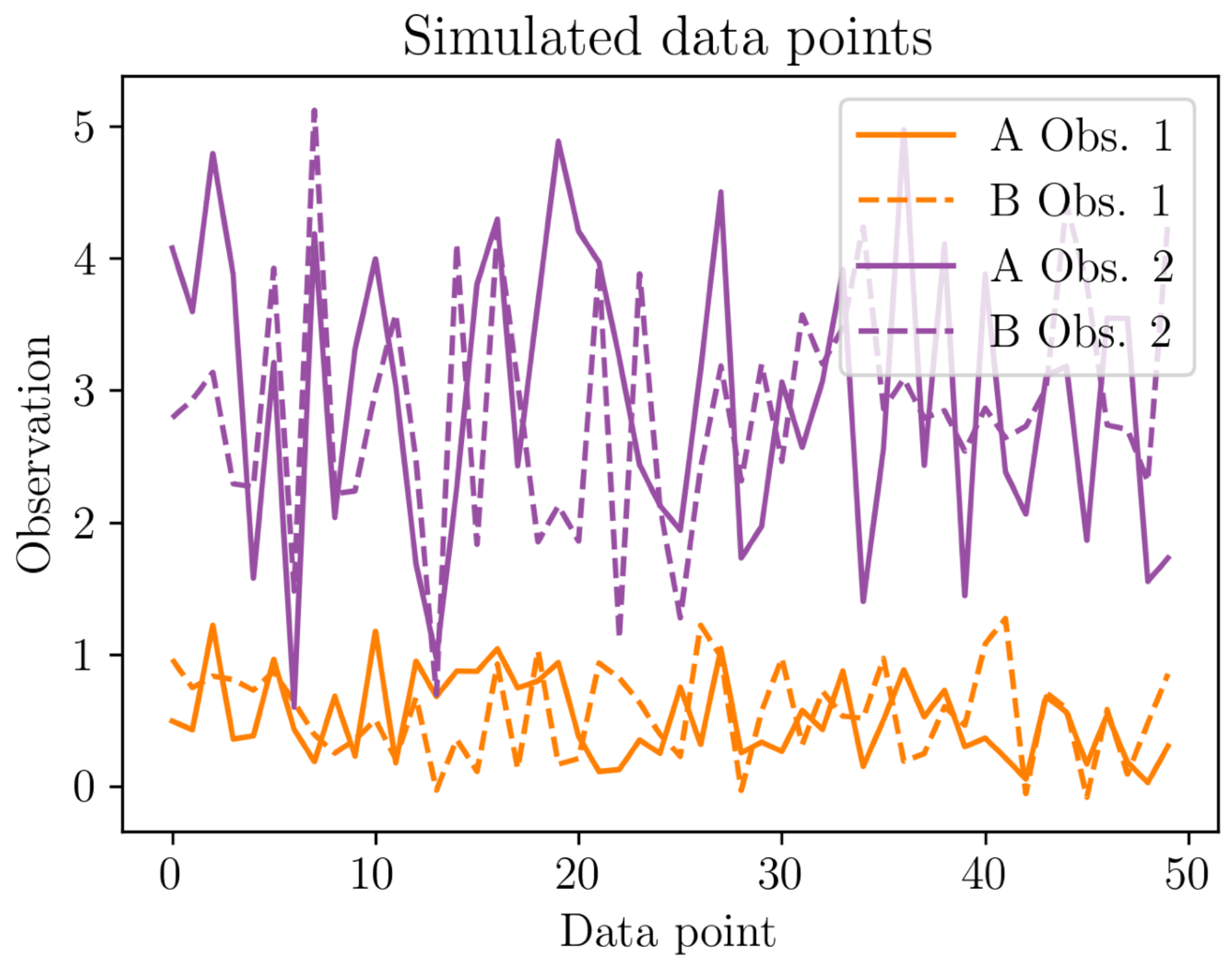
$$D_B = M_B\theta + m_B \pm \sqrt{C_B}$$

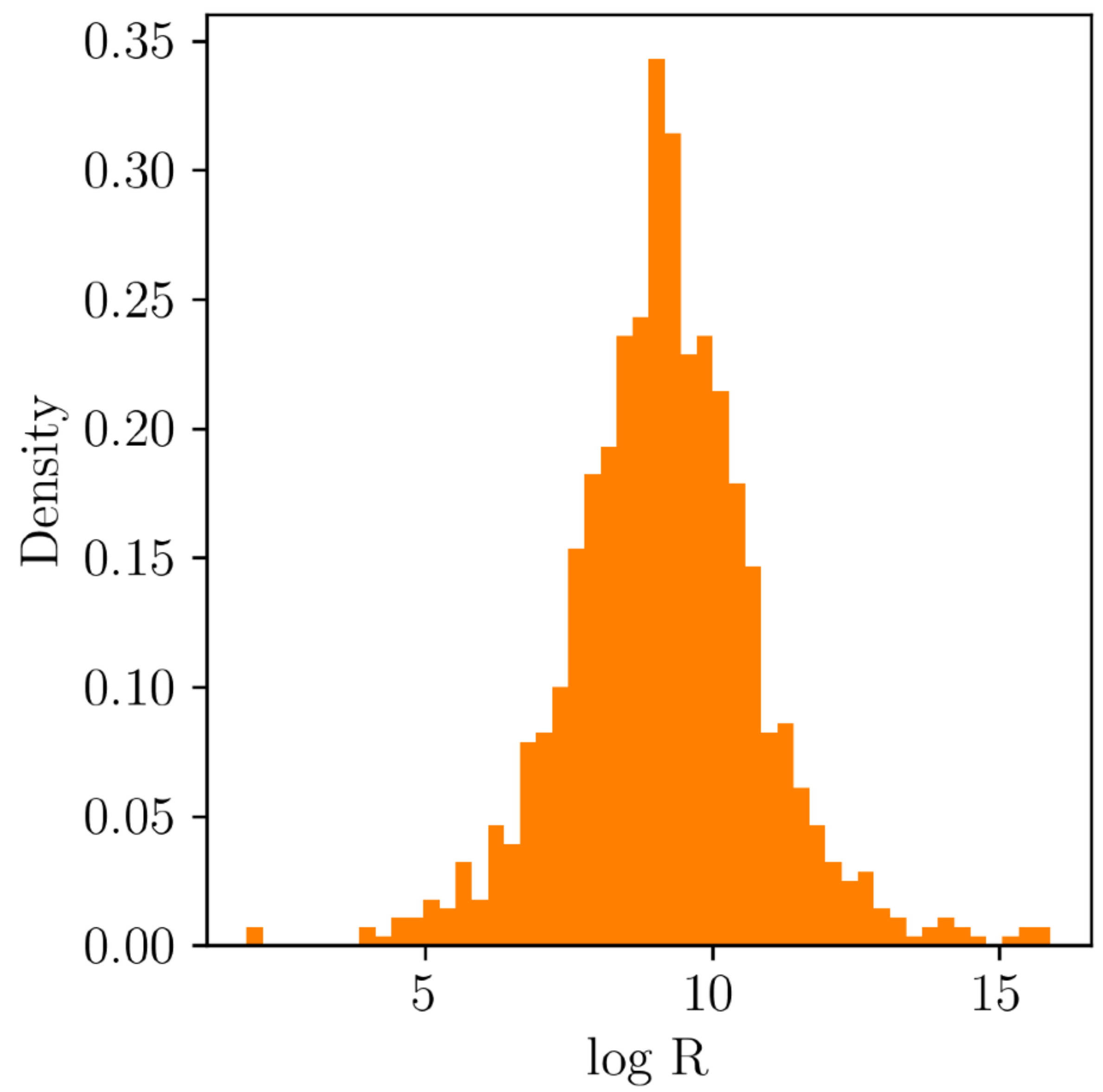
- $n_{dims} = 3, n_{data} = 50$

- Gaussian prior and likelihood

- Can analytically calculate $Z_A = P(D_A)$, Z_B and Z_{AB} and therefore get $\log R$

- Using lsbi package (https://github.com/handley-lab/lsbi)



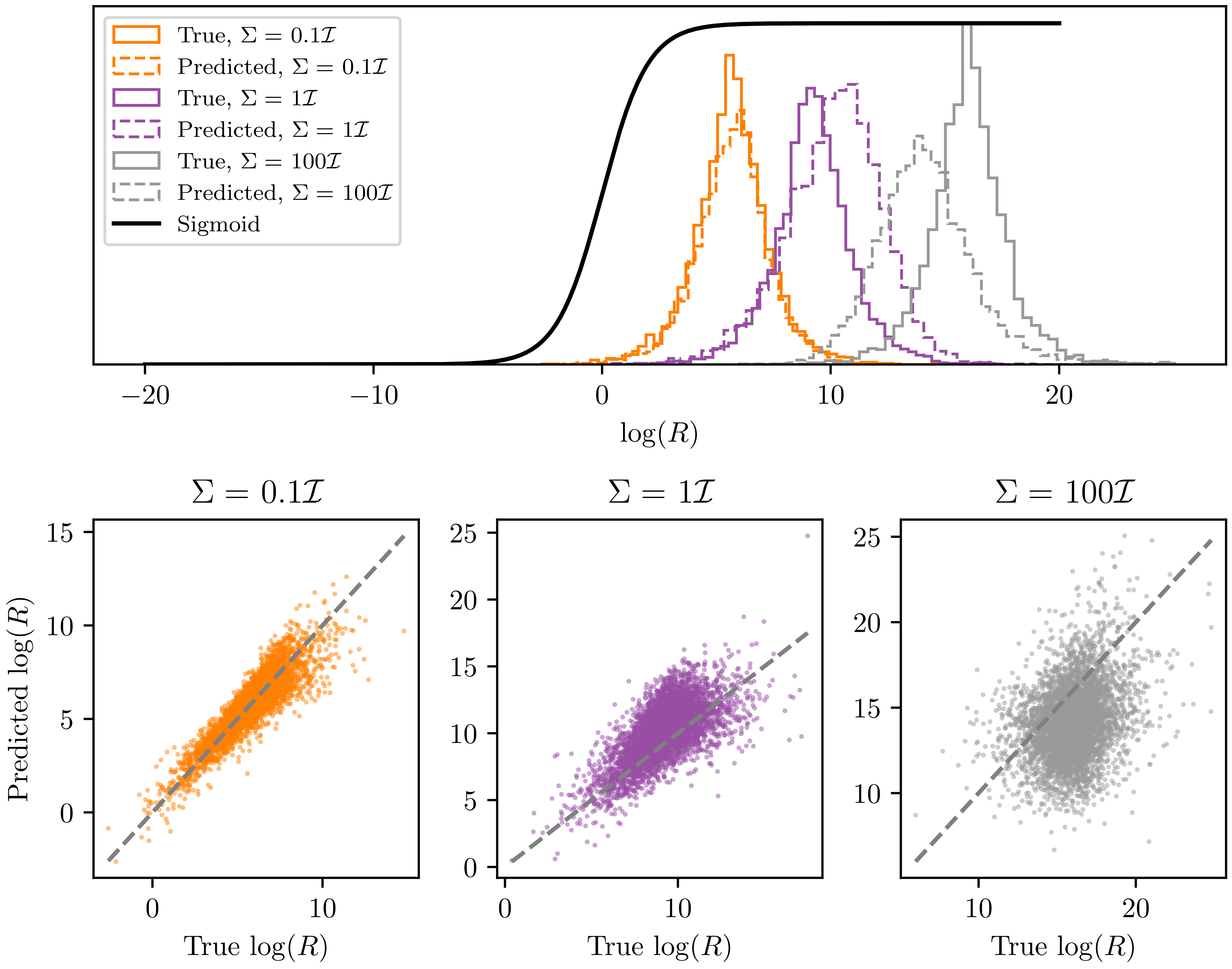


Analytic Example: Results

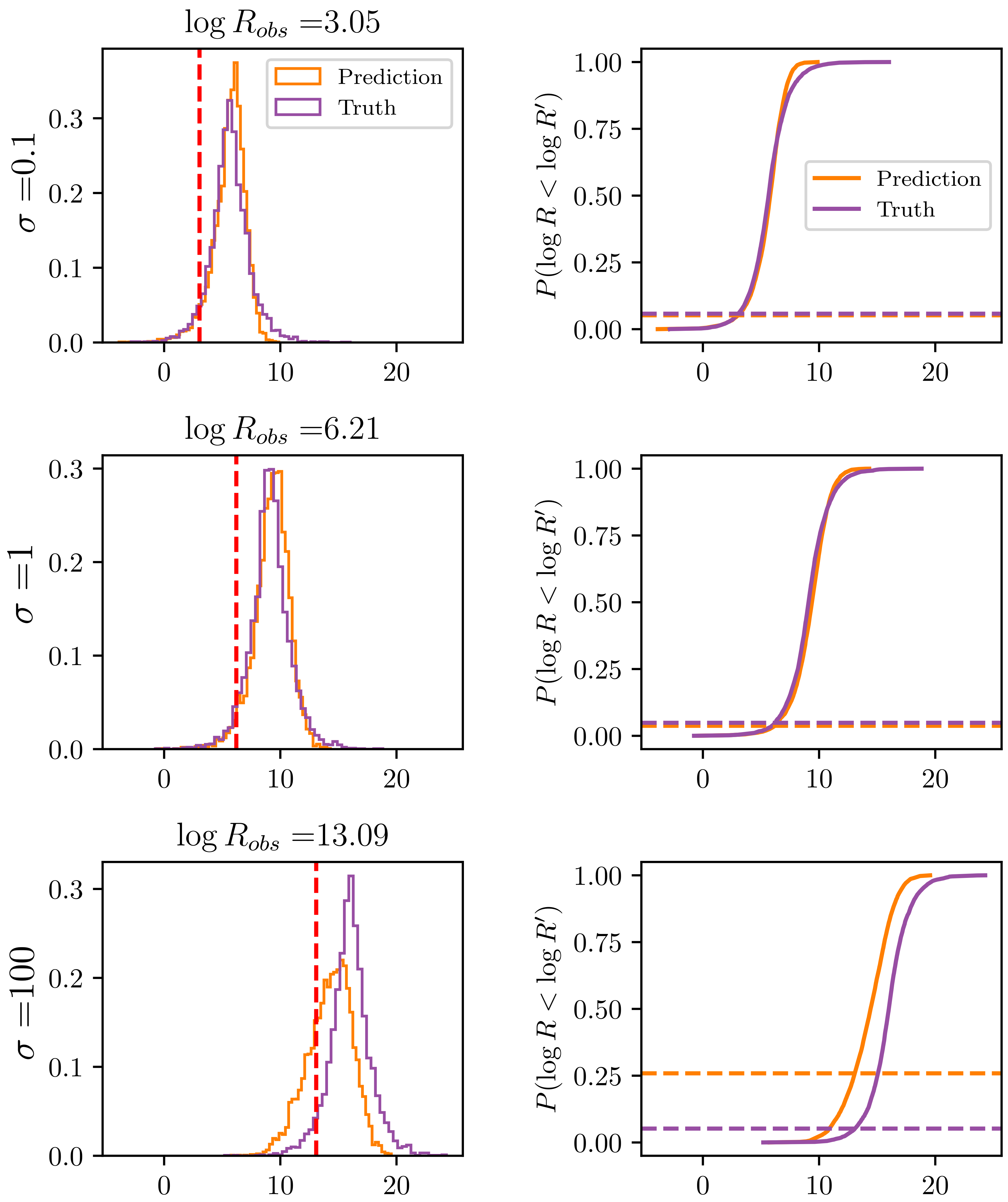
- Assess accuracy with changing prior width

- Performance is good for narrow priors

- Can push the performance for higher $\log R$ by tuning hyperparameters



Analytic Example: Prior Dependence



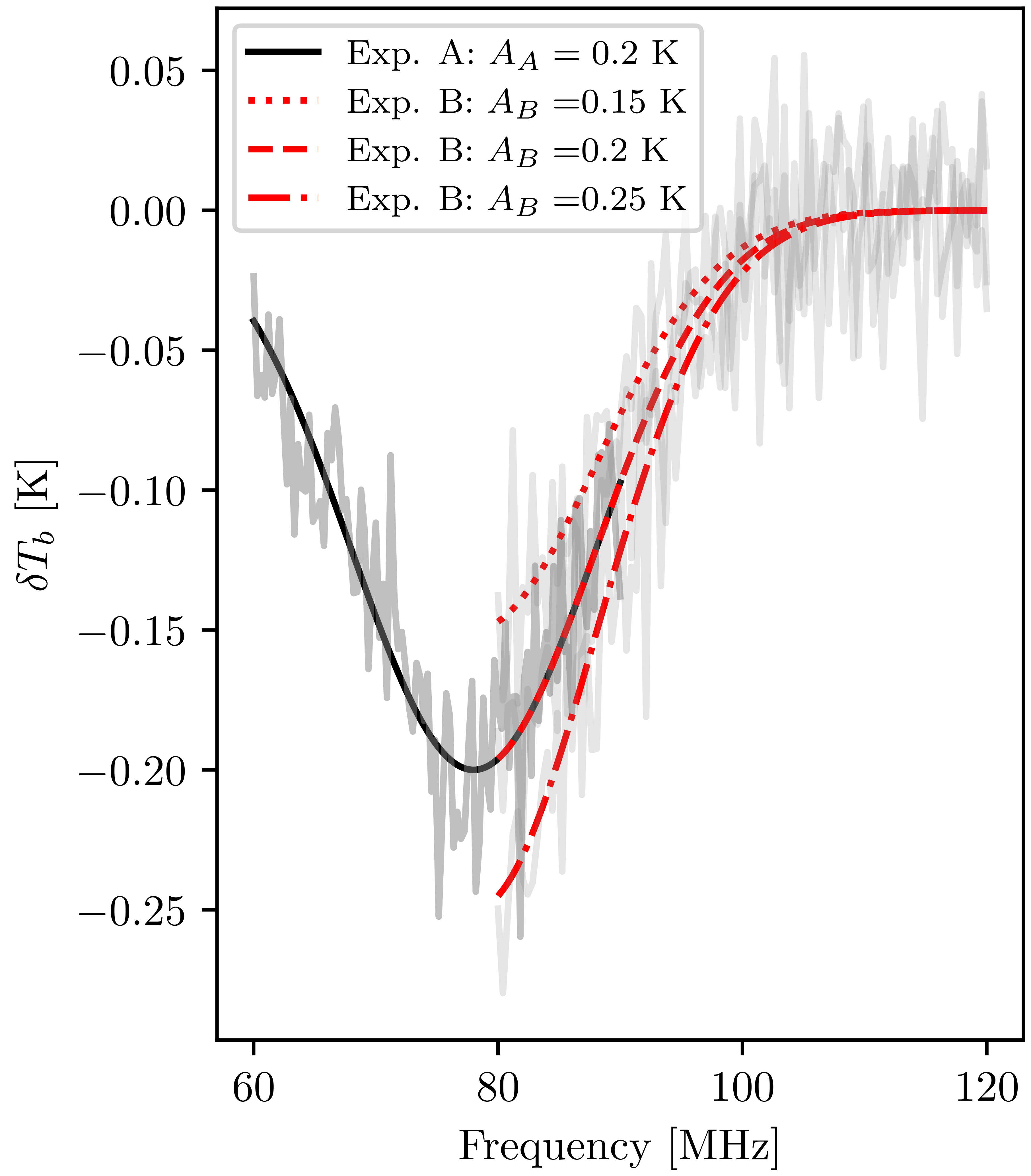
	σ_D	σ_A
Truth	1.896	0.073
Prediction	1.943	0.065

	σ_D	σ_A
Truth	1.970	0.061
Prediction	2.079	0.047

	σ_D	σ_A
Truth	1.943	0.065
Prediction	1.129	0.330

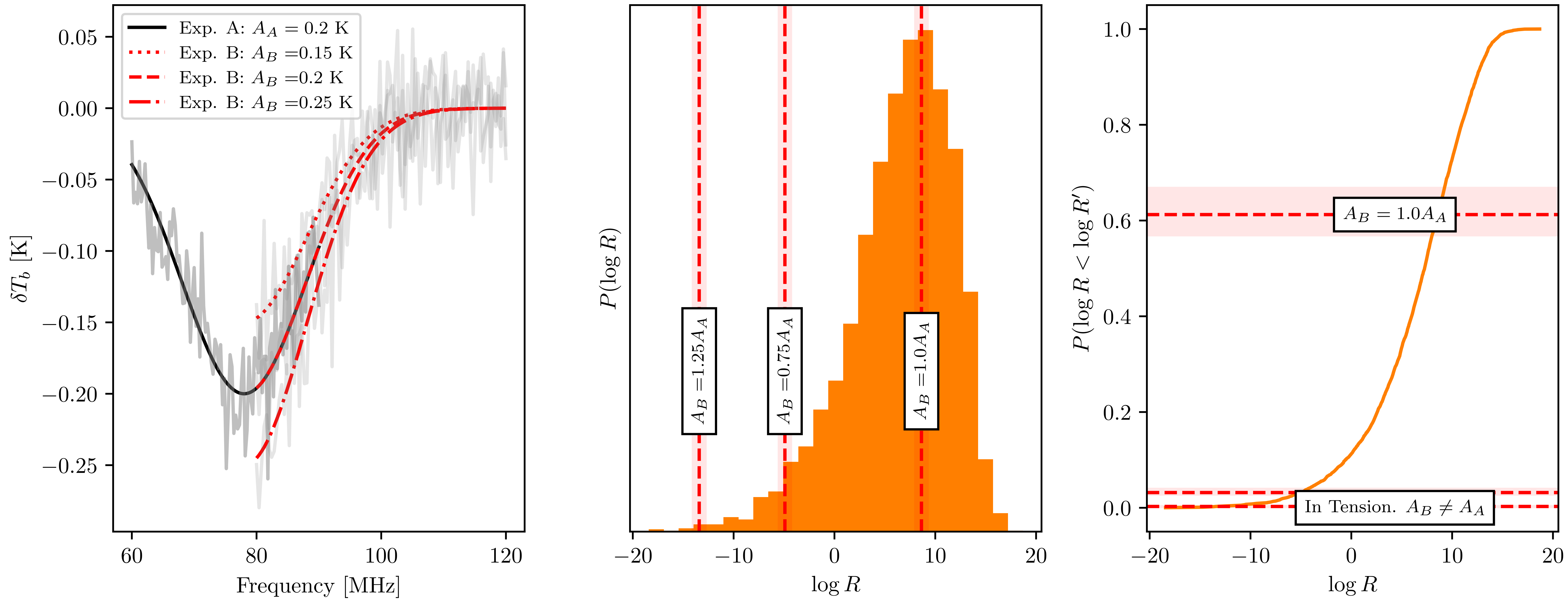
21cm Example: Set Up

- Looking for an absorption feature in the CMB a low frequencies
- Information rich signal that tells us about the properties of the first stars
- Observationally challenging and tension estimation is becoming more important [2112.06778]
- Two experiments observing in different bands
- Three scenarios



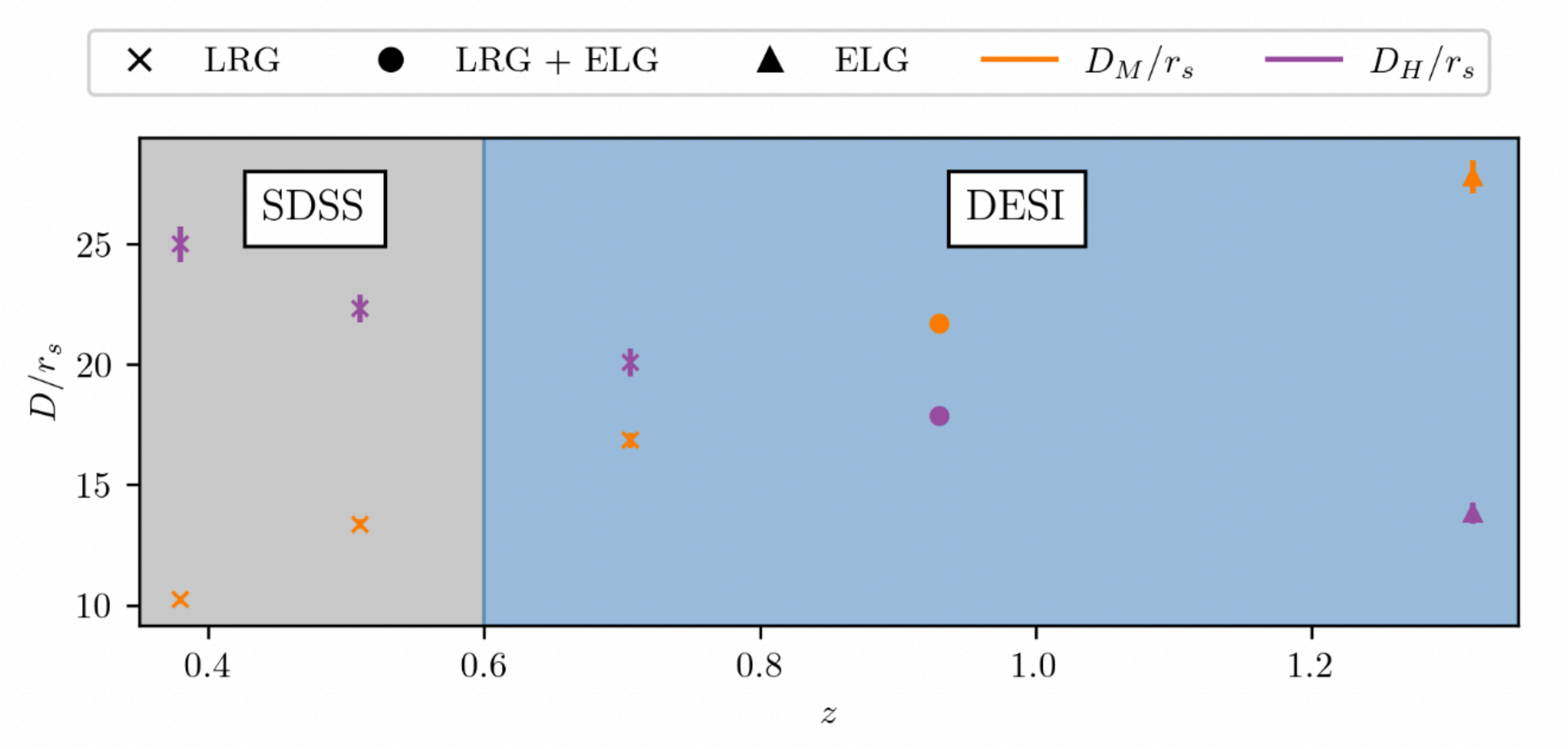
21cm Example: Results

- For the two in tension datasets $T = 2.989^{+0.167}_{-0.060}$ and $T = 2.147^{+0.056}_{-0.089}$
- For the inconcordance dataset $T = 0.507^{+0.063}_{-0.078}$ and $C = 0.864^{+0.107}_{-0.076}$

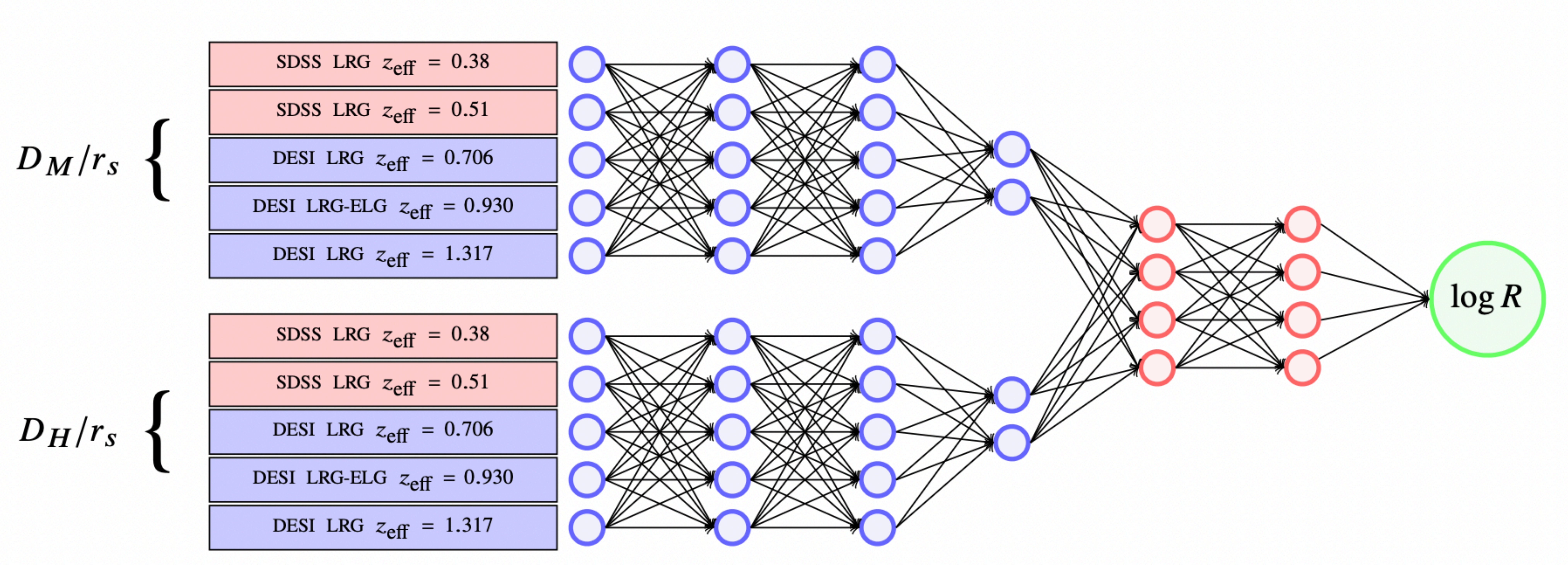


DESI + SDSS: Joint Data Set

- No existing correlated likelihood to evaluate a true R_{obs} with Nested Sampling
- Select different measurements from each survey to maximise the effective volume [e.g.]
- Focusing on LRG and ELG
- Add Quasars and Lyα in the future

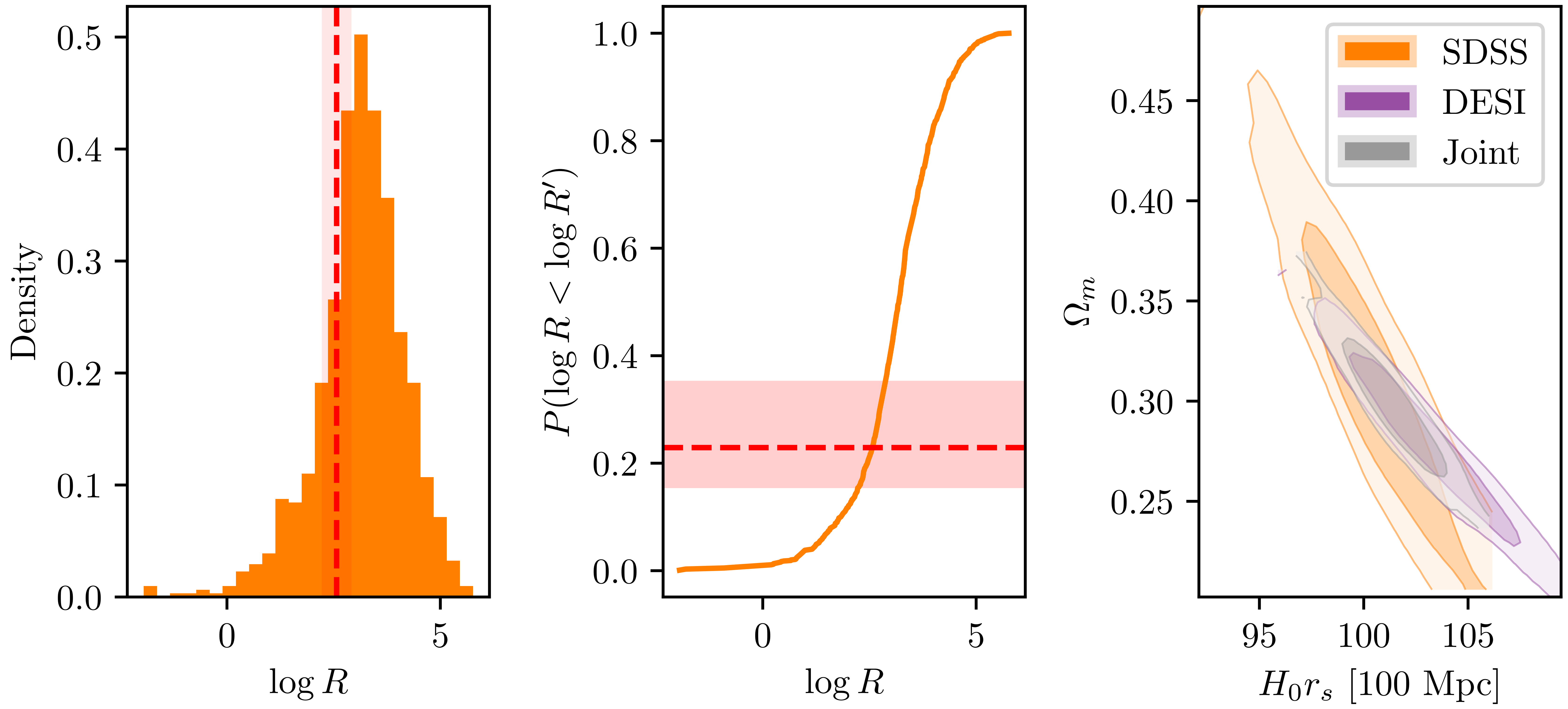


DESI + SDSS: NRE Set Up



DESI + SDSS: Results

- We find $T = 1.22 \pm 0.20$



Conclusions

- Understanding tensions can help us identify new physics or instrumental systematics
- R statistic is an appropriately Bayesian choice
- We can use Neural Ratio Estimation to help us interpret the tension between different experiments
- Paper coming soon!
- Email: htjb2@cam.ac.uk
- Website: harrybevins.co.uk