

# Gaussian Processes for Mitigating Systematics

Kavli Science Focus Meeting – Science with the 21-cm line

Christian Kirkham

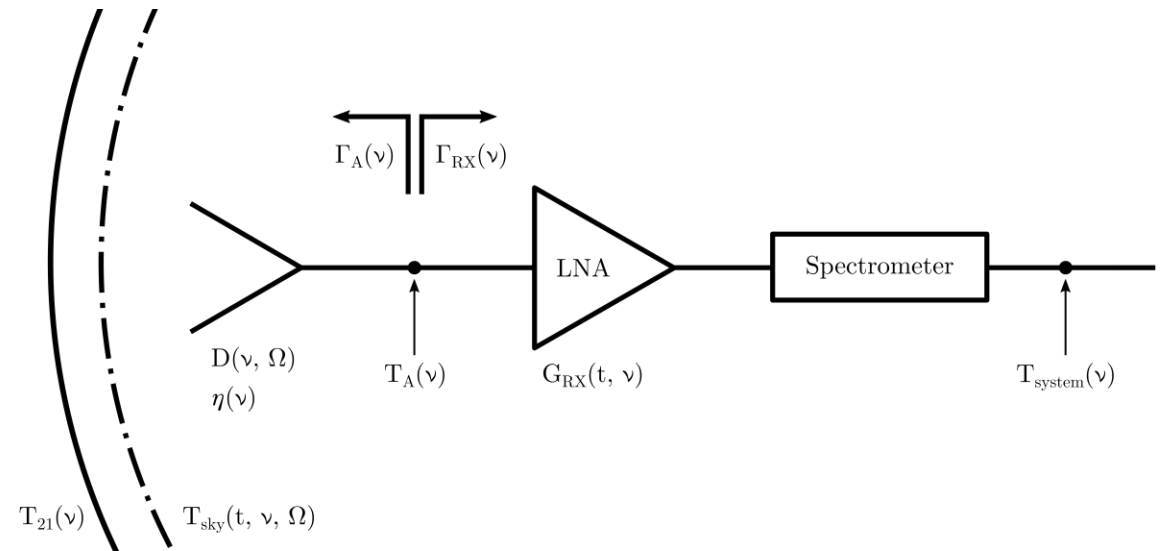
# Introduction

- Origins of time-varying systematics
- Systematic model
- Incorporating Gaussian Processes into the REACH pipeline
- Results
- Further work
  - GP Regression
  - Automatic Kernel Selection



# Time Varying Systematics in the REACH System

- Many sources of systematic error in the REACH system
- Some are expected to be static, others vary with time
  - Reflections from the soil vary with rainfall (Bevins et al. 2021, 2022b)
  - Impedences in the system are temperature dependent
  - Improperly modelled beams can cause systematics which vary with the galactic foreground



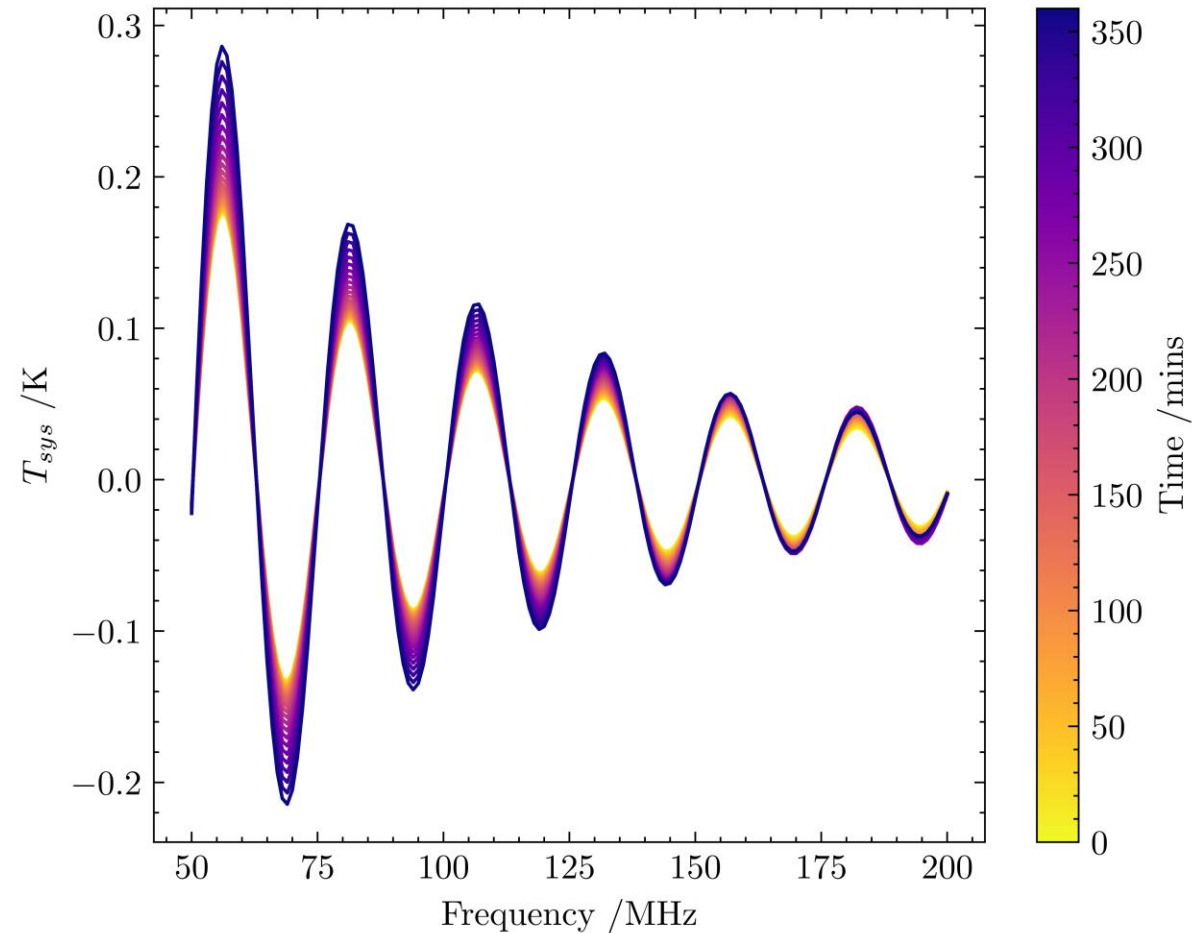
# Modelling Time Varying Systematics

- Model systematic as a sinusoid whose amplitude is modulated by the power of the galactic foreground

$$T_{\text{sys}}(\nu) = A_{\text{sys}} \left( \frac{\nu}{\nu_{0,\text{sys}}} \right)^{-\alpha_{\text{sys}}} \sin \left( \frac{2\pi\nu}{P_{\text{sys}}} + \phi_{\text{sys}} \right)$$

- Systematic amplitude is varied like

$$A_{\text{sys}}(t_j) = A_{\text{sys}}(t_0) \cdot \frac{T_{\text{fg}}(\nu = \nu_0, t = t_j)}{T_{\text{fg}}(\nu = \nu_0, t = t_0)}$$



# Introducing Gaussian Processes

- Define a GP as a collection of random variables which have consistent joint Gaussian distributions (Rasmussen 2004)
- Distribution is defined by the covariance matrix, or covariance function
$$\mathbf{C}_{ij} = K(t_i, t_j)$$
- $K$  is called the “kernel” – form can be chosen freely

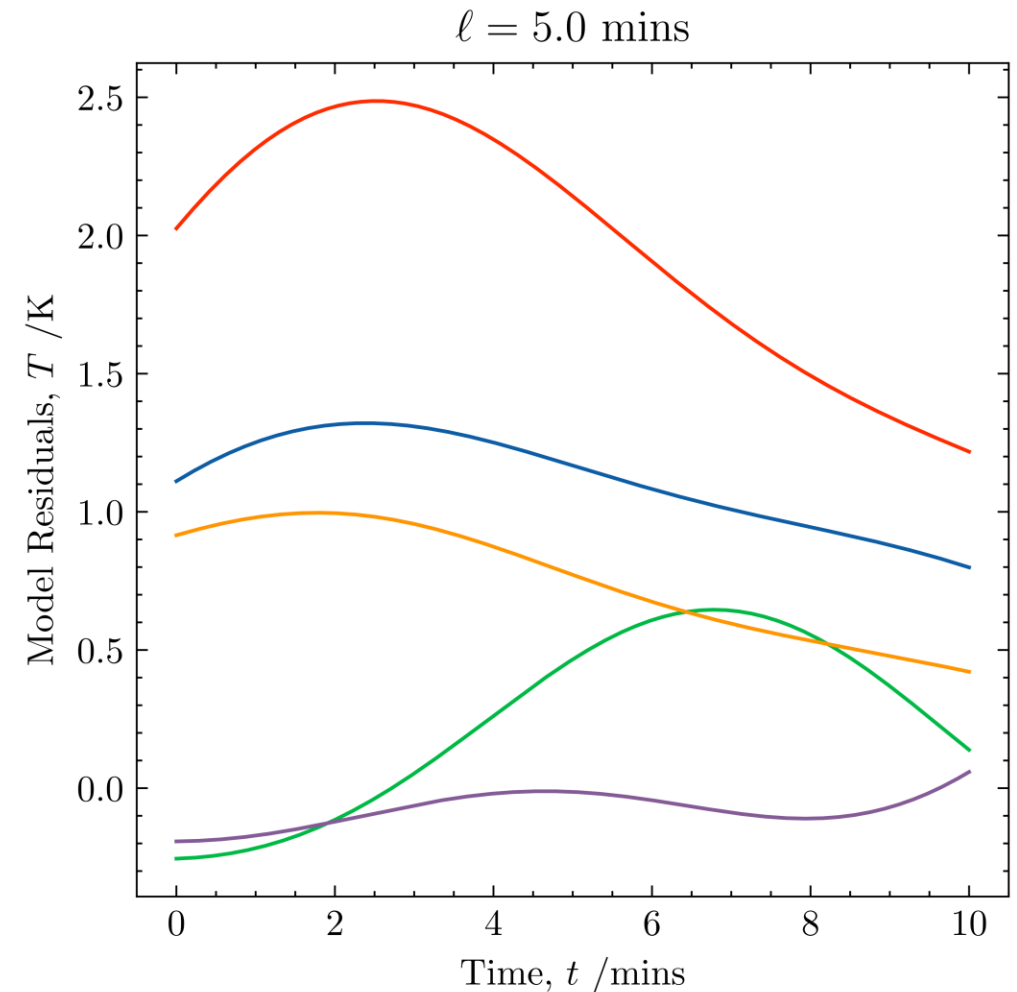


# Squared Exponential Kernel

- In this work I use a squared exponential kernel with a white noise term

$$\mathbf{C}_{ij} = \sigma_{0,\text{GP}}^2 + \sigma_{\text{SE}}^2 \exp\left(-\frac{|t_i - t_j|^2}{2\ell^2}\right)$$

- We have uncorrelated white noise term, scale factor of the SE kernel and the covariance length – “hyperparameters”
- Define smooth functions

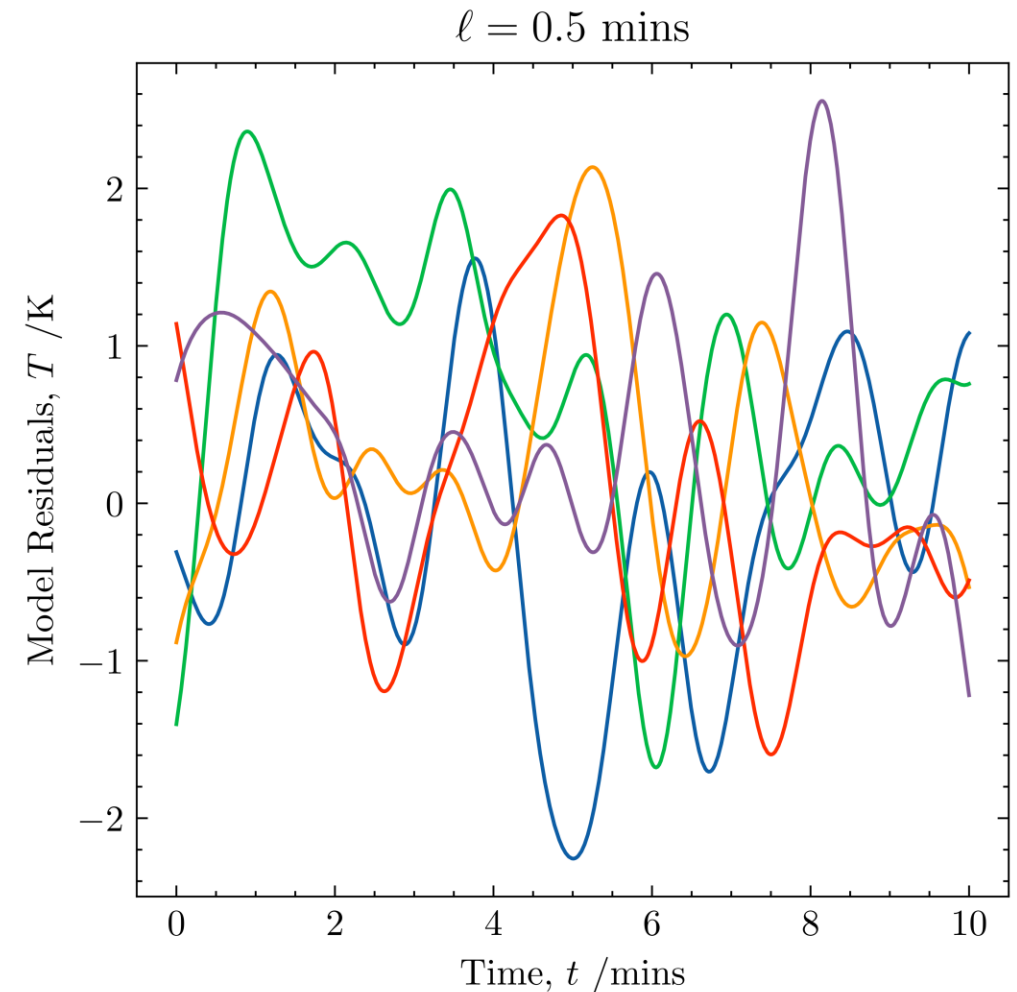


# Squared Exponential Kernel

- In this work I use a squared exponential kernel with a white noise term

$$\mathbf{C}_{ij} = \sigma_{0,\text{GP}}^2 + \sigma_{\text{SE}}^2 \exp\left(-\frac{|t_i - t_j|^2}{2\ell^2}\right)$$

- We have uncorrelated white noise term, scale factor of the SE kernel and the covariance length – “hyperparameters”
- Define smooth functions



# Gaussian Processes in the REACH Pipeline

- Hereafter called “Gaussian Process pipeline”

$$\mathcal{L}(\theta) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{C}|}} \exp\left(-\frac{1}{2}(\mathbf{D} - \mathbf{M}(\theta))^T \mathbf{C}^{-1}(\mathbf{D} - \mathbf{M}(\theta))\right)$$

- $\mathbf{D}$  is the data,  $\mathbf{M}$  is the model (foreground, beam model, Gaussian signal model)
- Gaussian process covariance matrix,  $\mathbf{C}$ , fits covariance of model residuals between time bins
- “Standard pipeline” (Anstey et al. 2021, 2022) – uncorrelated time bins



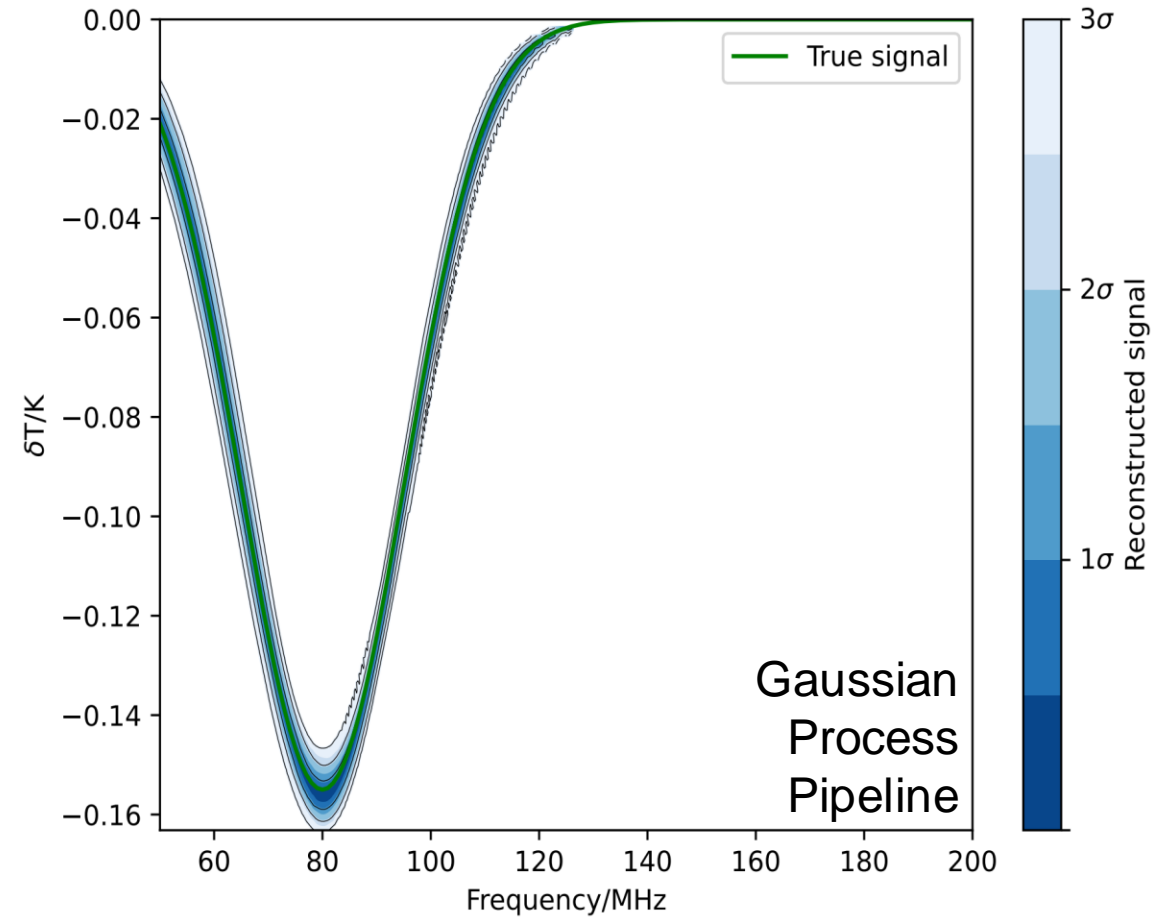
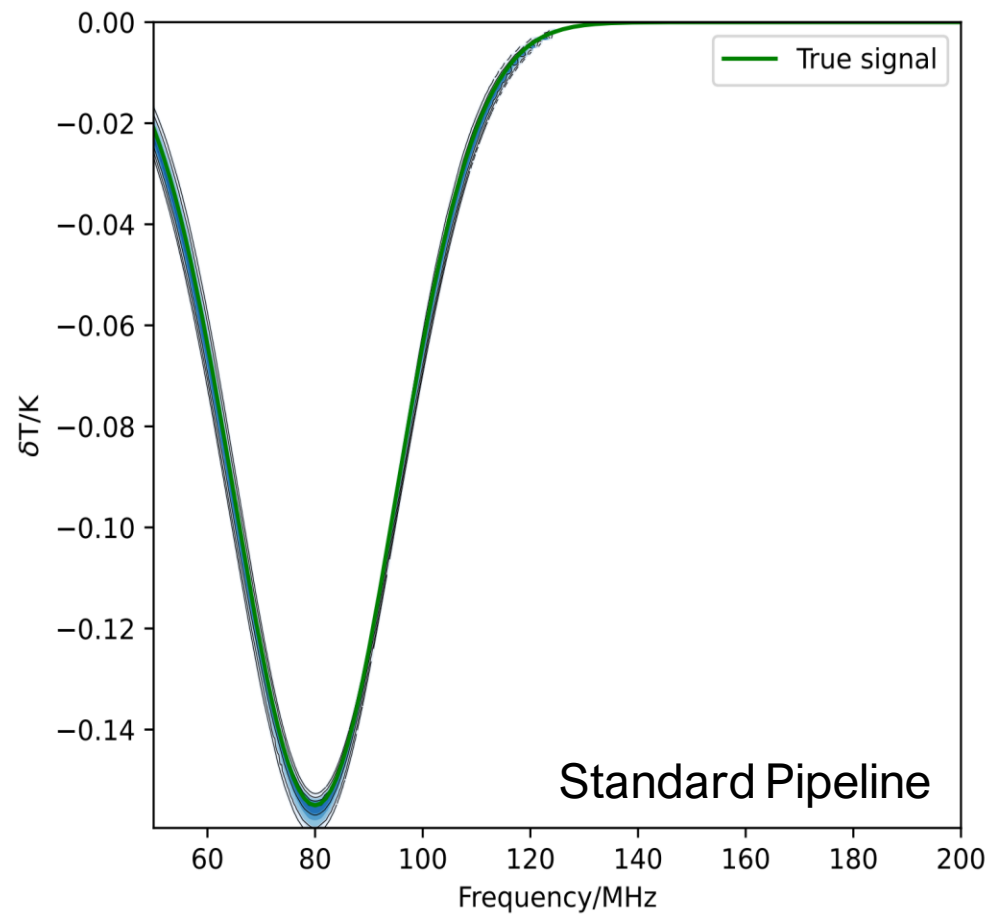


# Results

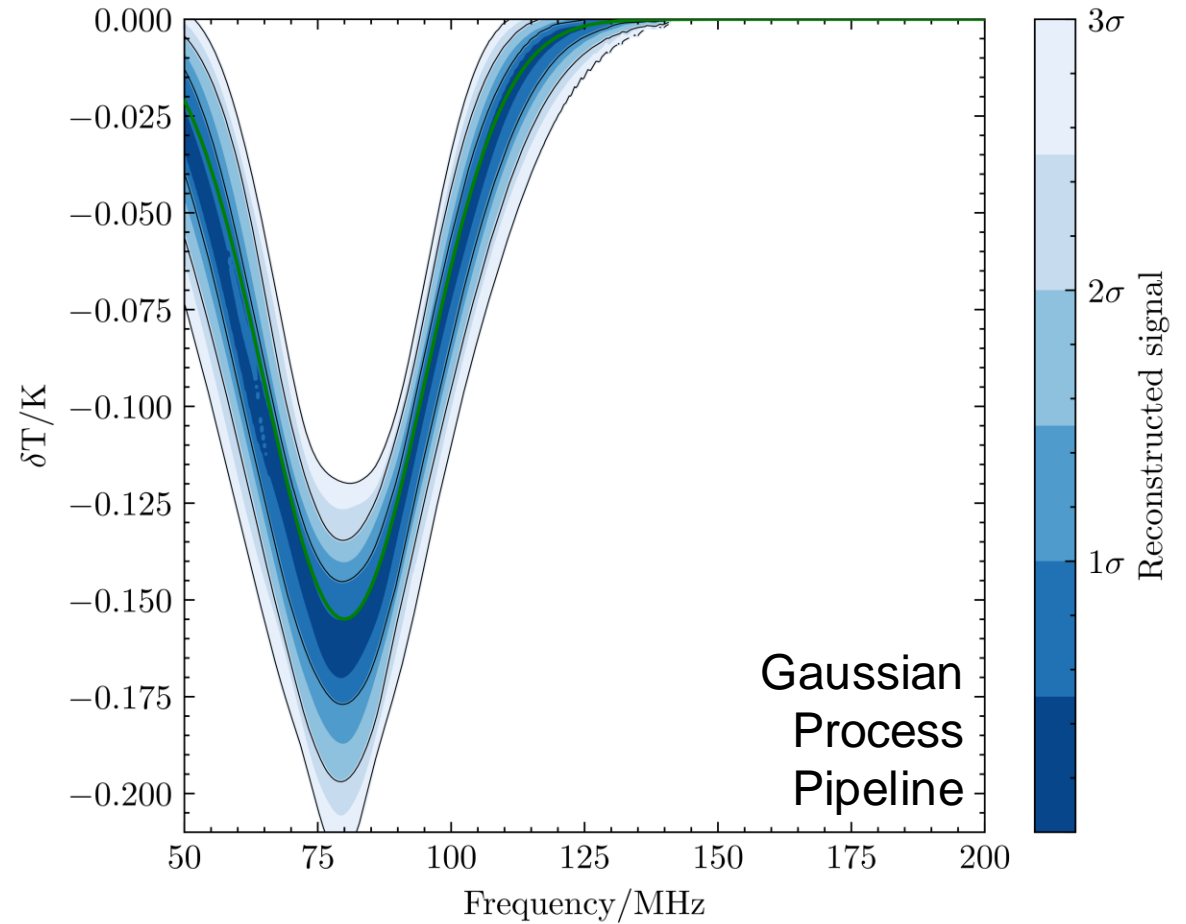
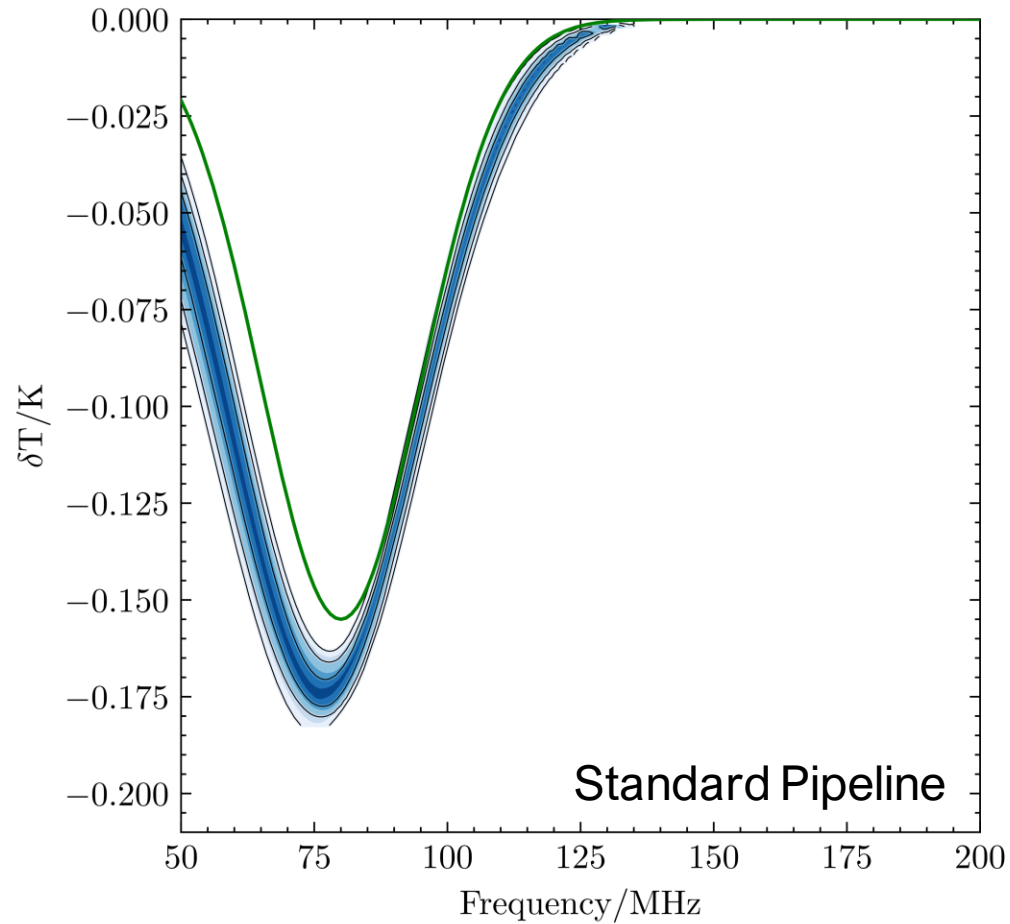
- In all cases added a Gaussian signal model into the data and fit for Gaussian
  - $A_{21} = 0.155 \text{ K}$
  - $\sigma_{21} = 15 \text{ MHz}$
  - $\nu_{0,21} = 85 \text{ MHz}$
- Parameters of systematic varied relative to signal parameters
- Data generated for 24 time bins of length 15 minutes – 6 hours of observation in total



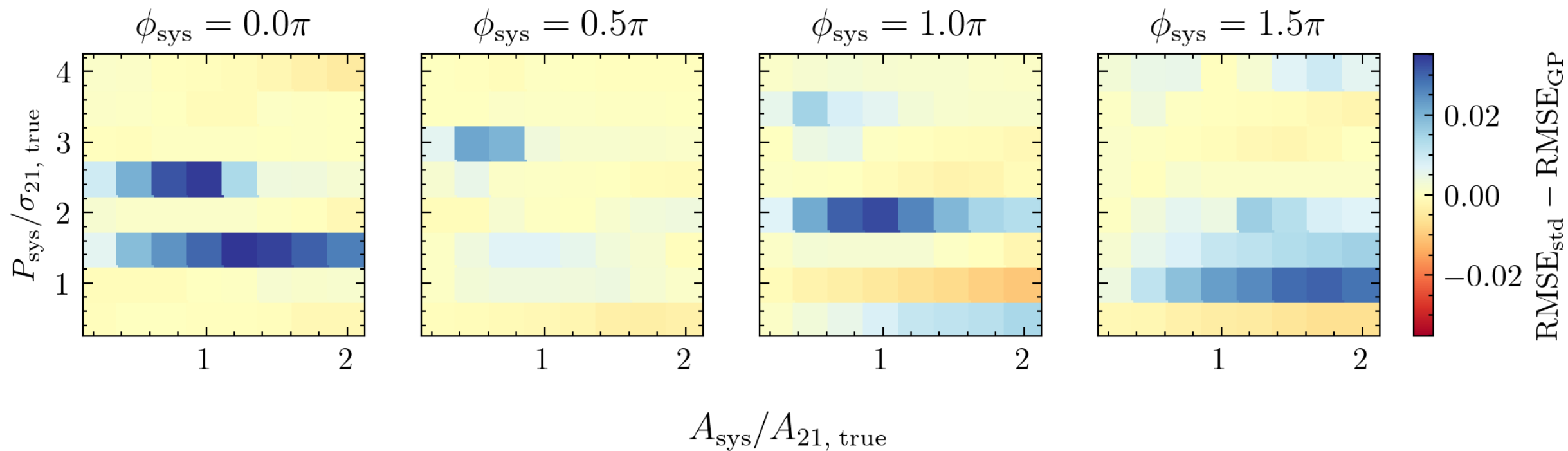
# Results – Example Comparison of Signal Posteriors – No Systematic



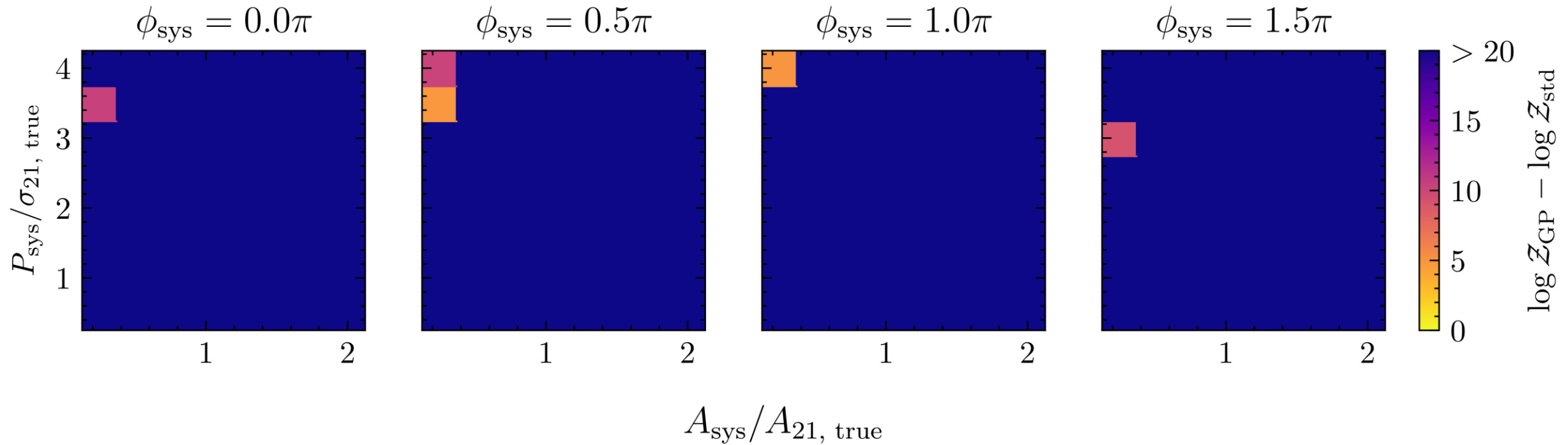
# Results – Example Comparison of Signal Posteriors – Systematic



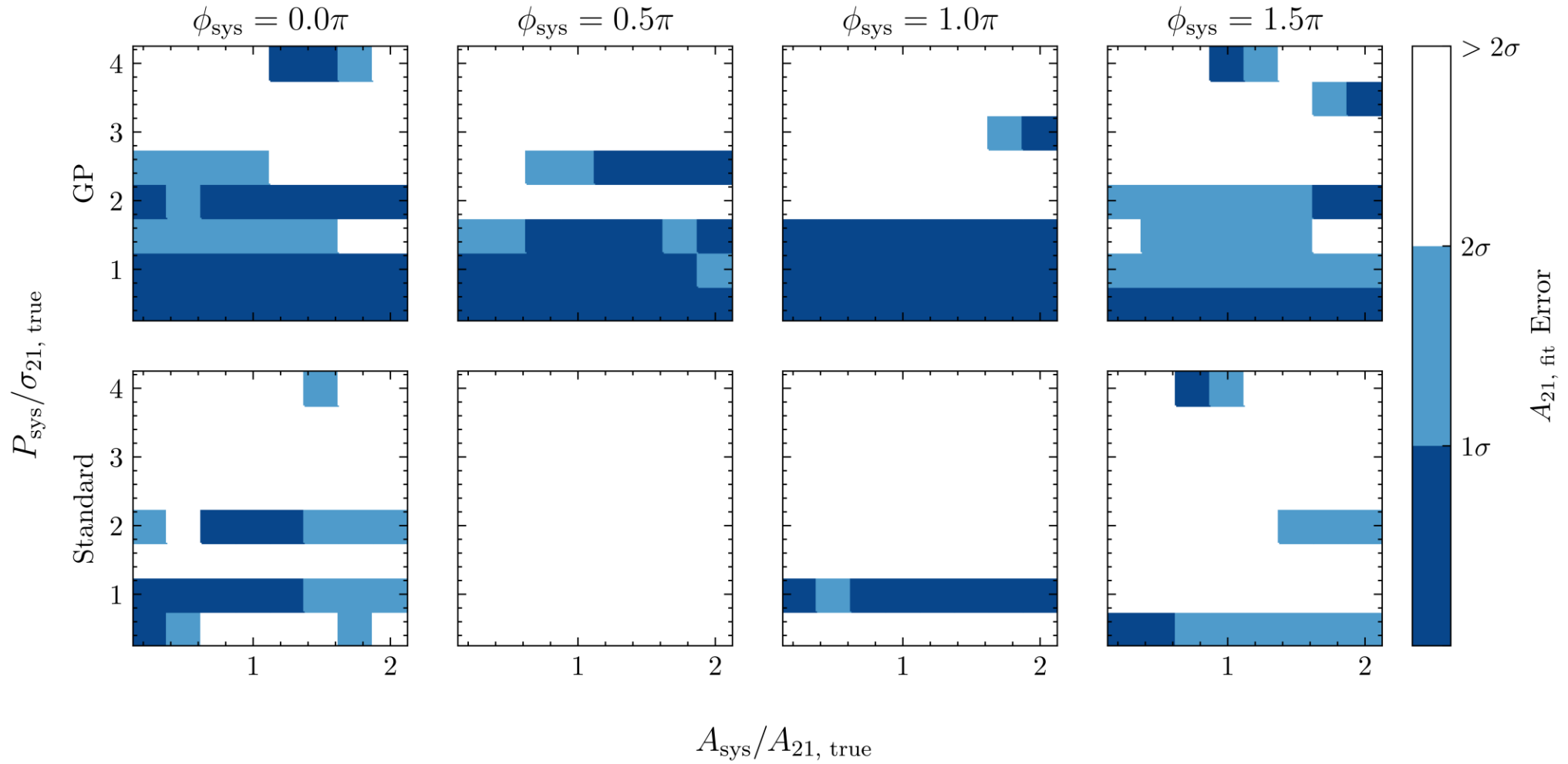
# Results – Weighted RMSE



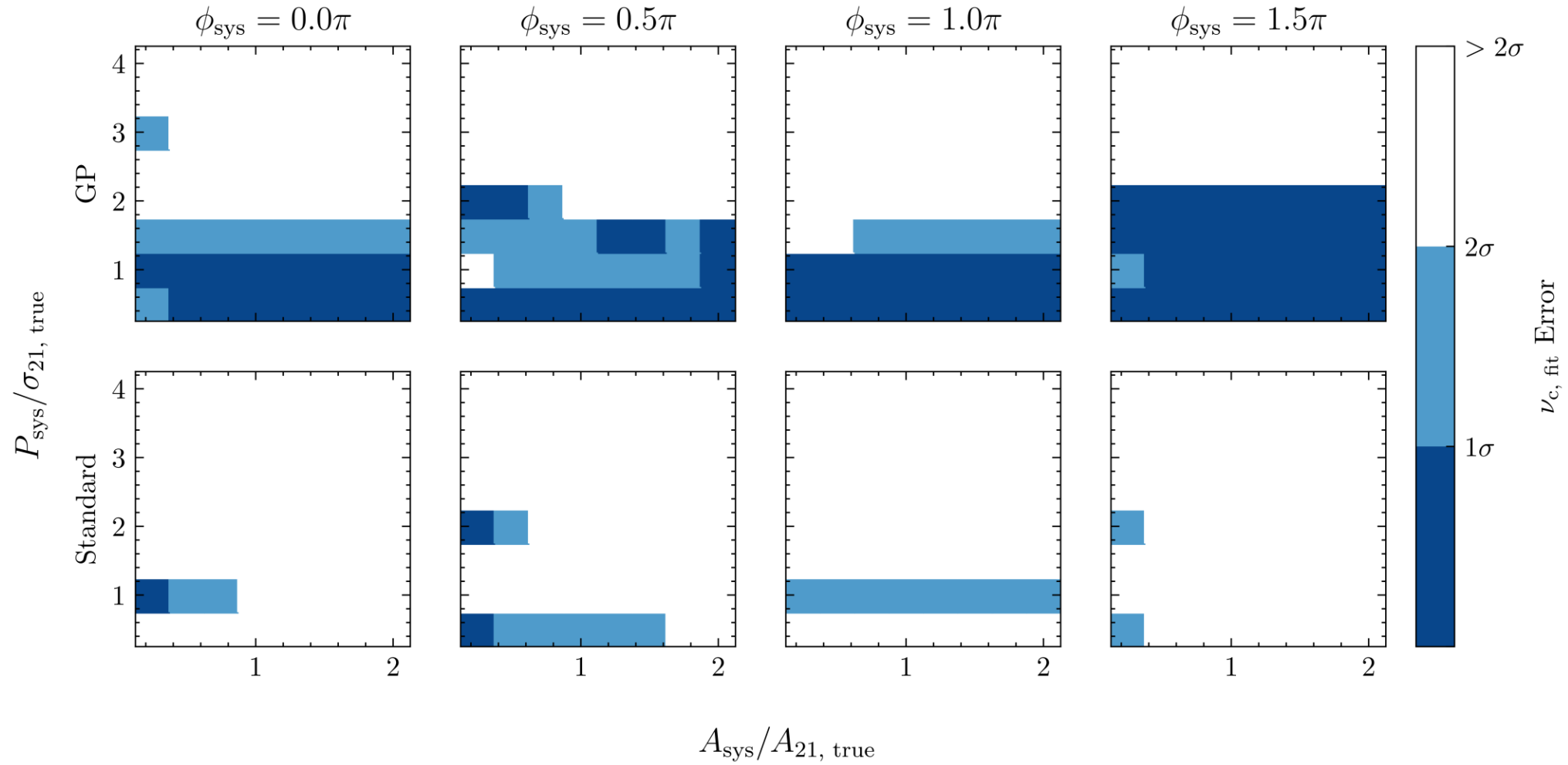
# Results – Bayes Factor



# Results – Fitted Signal Amplitude



# Results – Fitted Centre Frequency



# Future – Gaussian Process Regression

- Can also use Gaussian Processes for regression
- Mean function of the Gaussian Process posterior given by

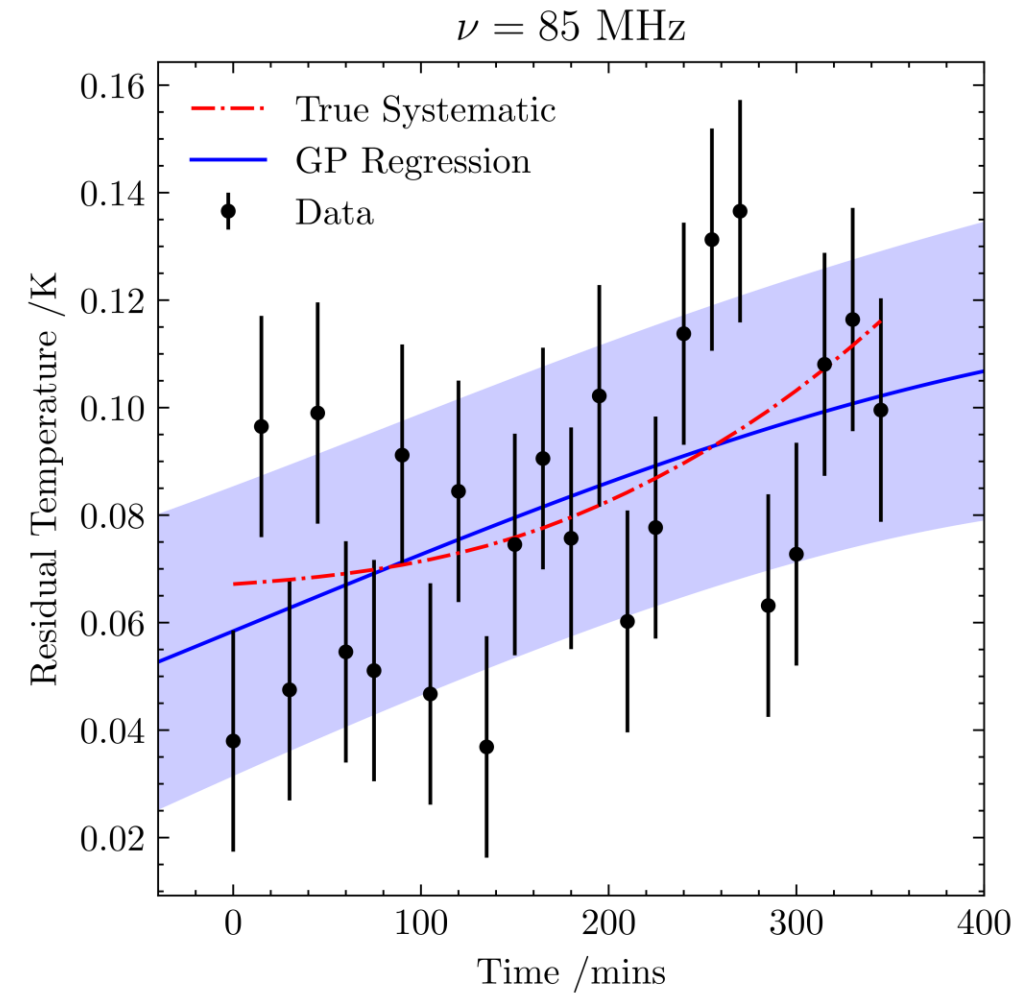
$$\mu(\mathbf{t}_{\text{pred}}) = K(\mathbf{t}_{\text{pred}}, \mathbf{t}_{\text{data}})K(\mathbf{t}_{\text{data}}, \mathbf{t}_{\text{data}})^{-1}\mathbf{T}_{\text{data}}$$

- Can be used to predict temperature,  $\mathbf{T}_{\text{pred}}$ , of model residuals at some time  $\mathbf{t}_{\text{pred}}$



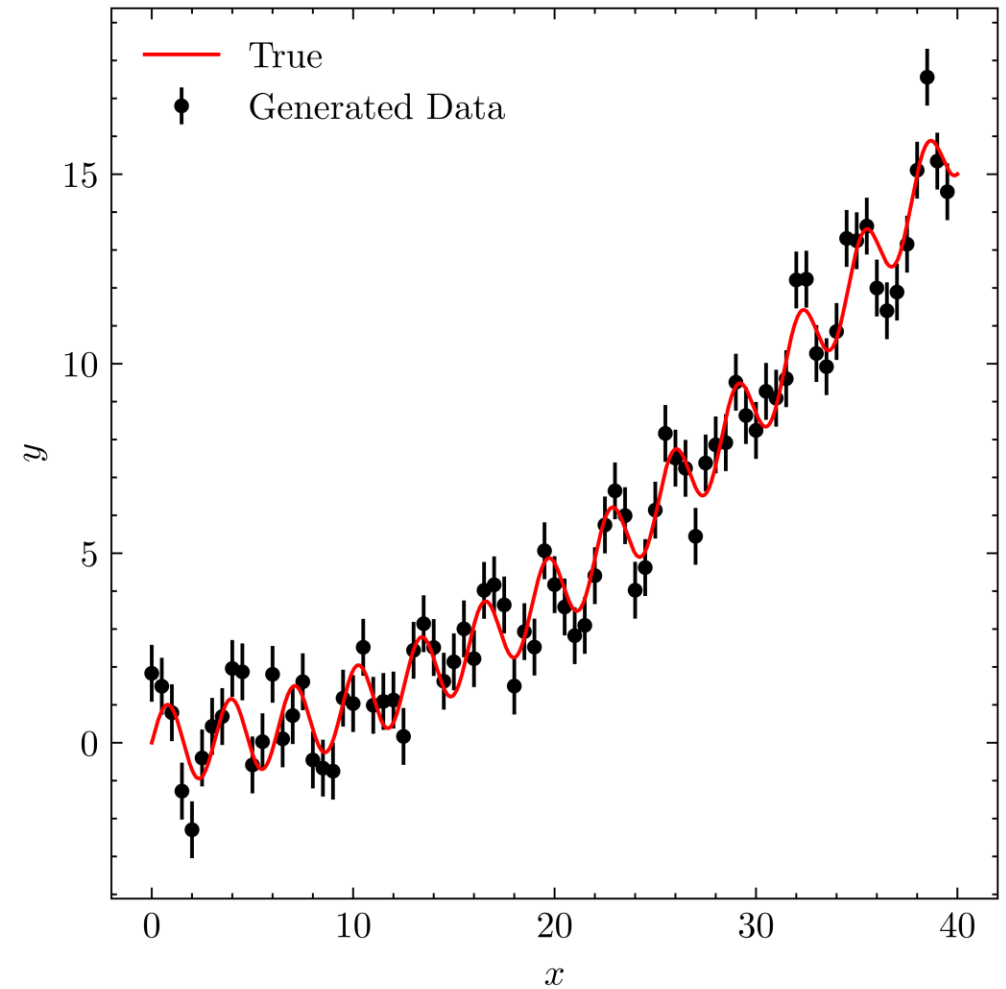
# Future – Gaussian Process Regression

- Use weighted mean Gaussian Process hyperparameters to get a smooth fit to model residuals
- Could possibly be used to inform future time-varying systematic models



# Future – Automatic Kernel Selection

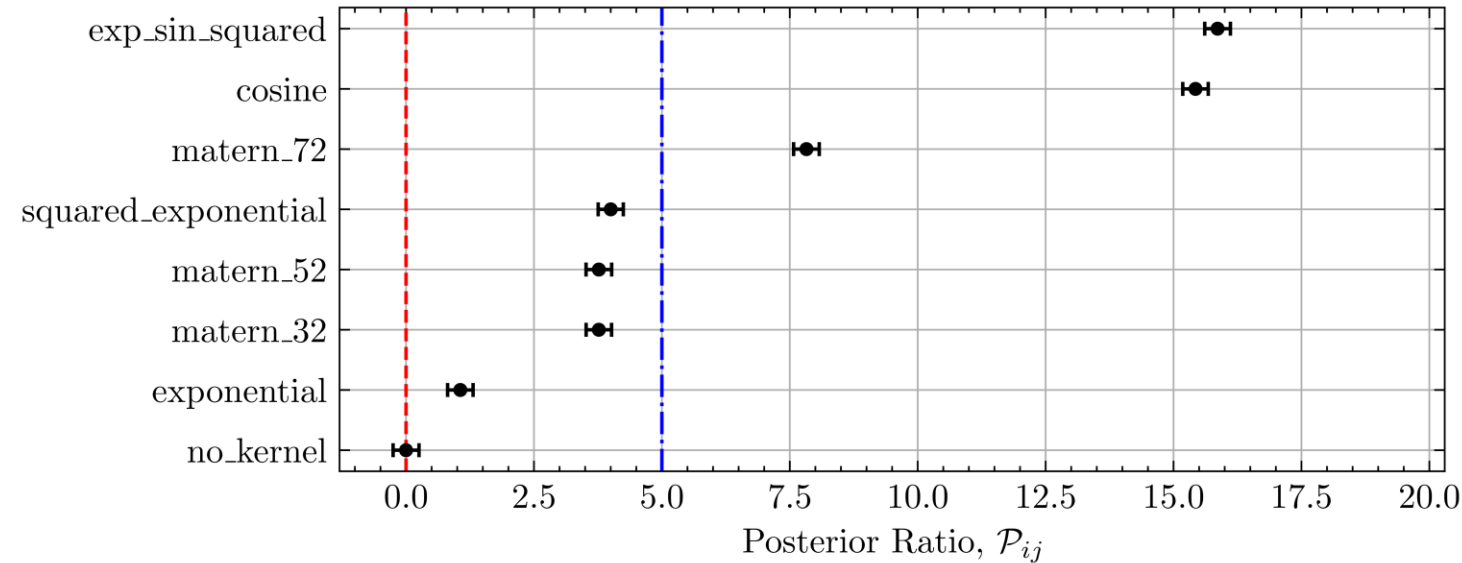
- Kernel choice is arbitrary so can use Bayes factor to inform us (Hee et al. 2015, Kroupa et al. in prep)
- Basic test with a quadratic curve with a sinusoidal residual
- Fit for quadratic but not sinusoid



# Future – Automatic Kernel Selection

- Uses PolyChord to sample over all kernels using a choice parameter,  $c$
- Uses the posterior ratio as a proxy for Bayes Factor

$$\mathcal{P}_{ij} = \log \frac{\Pr(c = j | \mathbf{D}, \mathbf{M})}{\Pr(c = i | \mathbf{D}, \mathbf{M})}$$



# Conclusions

- Using Gaussian Processes to account for time correlated residuals improves fitting
- General method – no systematic model required
- Regression can inform future models of systematics
- Automatic Kernel Selection can help select most appropriate kernel

