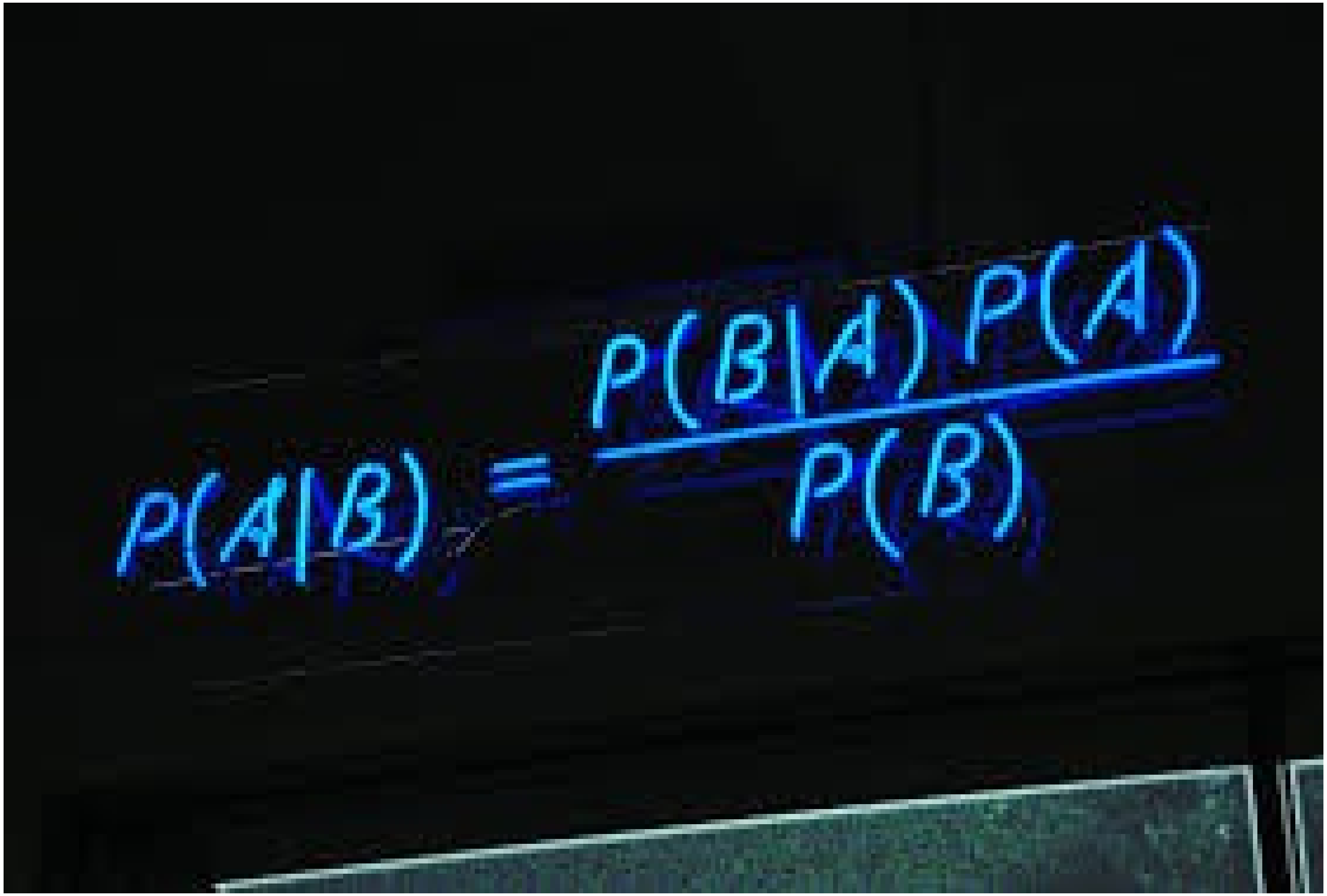
Dimensions and Gravitational Waves

Rutger van Haasteren
(Jet Propulsion Lab)

Road to the next dimension

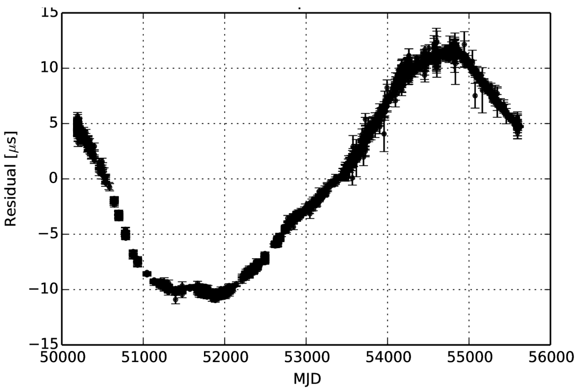
1. Gravitational waves and pulsar timing
2. The curse of dimensionality
3. Solution 1: reduced-rank representations
4. Solution 2: Gibbs sampling
5. Outlook: beating the clock

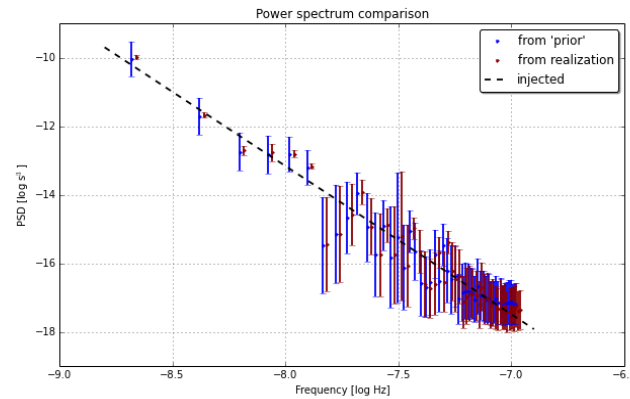
Bayesian analysis

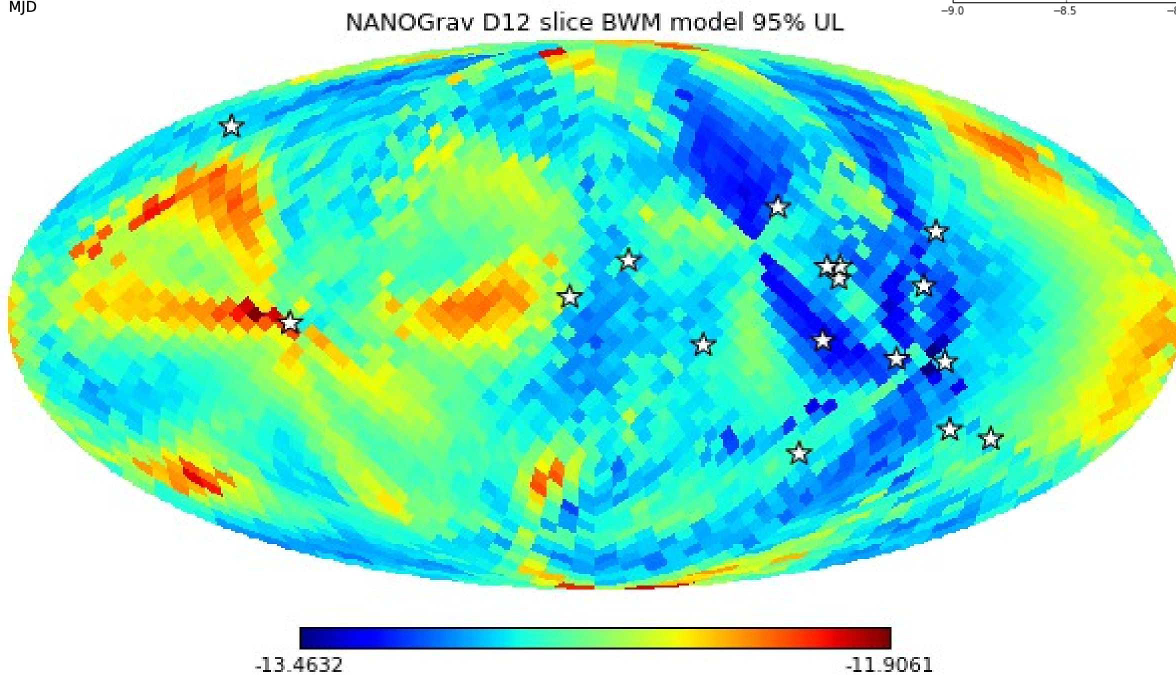


Posterior ~ Likelihood x Prior

Degrees of freedom in model



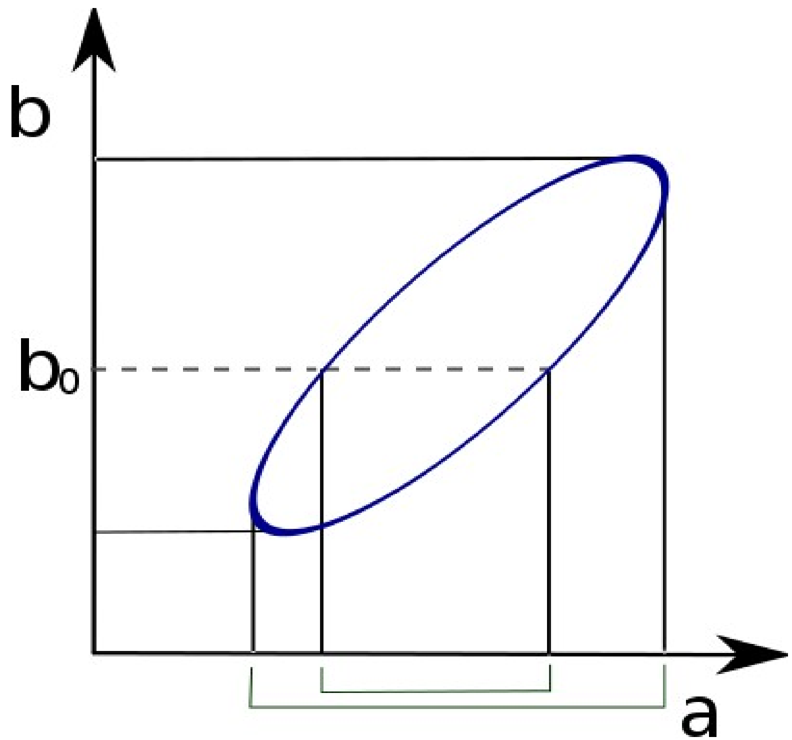




Bayesians: “Marginalize over the nuisance parameters”

BIG BAD: Parameter covariance

- Covariant parameters: parameters that depend on each other
- Fitting for the parameters one-by-one gives biased estimators.
- Solution: treat all parameters **simultaneously**
- Simple example: linear least-squares fit



We do noise analysis

- Timing residuals: $$TOA = f_0(t) + \underbrace{\delta t_{TM} + \delta t_{RN} + \delta t_{WN} + \delta t_{DM} + \delta t_J}_{\delta t_{all}}$$

  - TM: timing model correction
  - RN: red spin noise
  - WN: white noise
  - DM: dispersion measure variations
  - J: Pulse jitter/correlated white noise

- Likelihood is trivial to write down:

$$P(TOA) = \delta\left(TOA - f_0(t) - \delta t_{TM} - \delta t_{RN} - \delta t_{WN} - \delta t_{DM} - \delta t_J\right)$$

Likelihood & white noise

- Timing residuals: $TOA = f_0(t) + \delta t_{Noise}$

Data

Timing model

Noise model

Model

- Data likelihood: $P(TOA) = \delta\left(TOA - f_0(t) - \delta t_{TM} - \delta t_{RN} - \delta t_{WN} - \delta t_{DM} - \delta t_J\right)$

- Prior white noise: $P(\delta t_{WN}) = \prod_i \exp\left(\frac{-\delta t_{WN,i}^2}{2\sigma_i^2}\right)$

- The big trick: marginalize (integrate) over all δt_{WN}

Least-squares fitting

- Hierarchical likelihood is now:

$$P(TOA,\delta t_{WN}) = \delta\left(TOA - f_0(t) - \delta t_{TM} - \delta t_{RN} - \delta t_{WN} - \delta t_{DM} - \delta t_J\right) \prod_i \exp\left(\frac{-\delta t_{WN,i}^2}{2\sigma_i^2}\right)$$

- Prior $P(\delta t_{WN})$ is modeled with 'hyper parameters'
- Result of marginalization: $P(TOA) \propto \exp\left(-\frac{1}{2}\sum_i \frac{\left(TOA - f_0(t) - \sum_a \delta t_{a,i}\right)^2}{\sigma_i^2}\right)$
- Equivalent to weighted least-squares fit

Hierarchical modeling

- Say we want to do both $P(\delta t_{WN})$ and $P(\delta t_{RN})$:

$$P(\delta t | \delta t_{TM}, \delta t_{RN}, \delta t_{DM}, \delta t_{J}, \{\sigma_i\}) \, P(\delta t_{RN} | A_{RN}, \gamma_{RN}) =$$

$$P(TOA) \propto \exp\left(-\frac{1}{2} \sum_i \frac{\left(TOA - f_0(t) - \sum_a \delta t_{a,i}\right)^2}{\sigma_i^2} \right) \exp \frac{\left(-\frac{1}{2} \delta t_{RN} C_{RN}^{-1} \delta t_{RN} \right)}{\sqrt{\det C_{RN}}}$$

- Can do this for ANY component/distribution

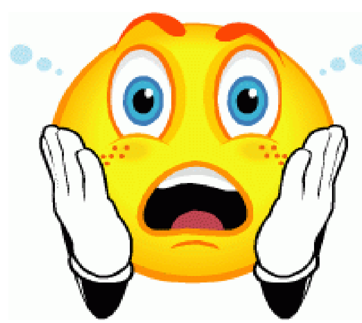
$$TOA = f_0(t) + \delta t_{TM} + \delta t_{RN} + \delta t_{WN} + \delta t_{DM} + \delta t_J + \delta t_{GW}$$

The curse of dimensionality

- We can model anything that way. Two problems remain
  - What do we choose as our model?
  - Can we run our analysis fast enough?

- Fully including all parameters (FAST likelihood)
  - $\sim 10^4$ residuals (times 5)
  - $\sim 10^3$ hyper parameters
  - ~ 50 pulsars

- For a full array, we have $\sim 10^6$ parameters



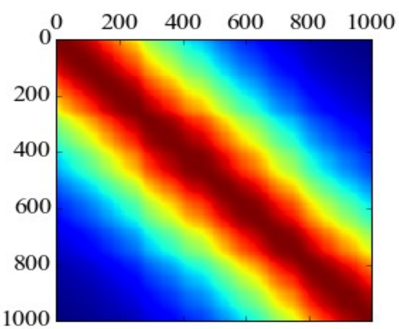
Solution 1: rank-reduction

$$\langle t_i t_j \rangle = K(t_i - t_j)$$

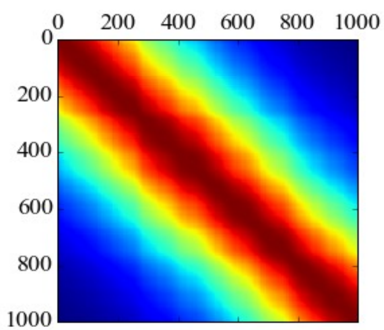
$$K_{ij} = N\delta_{ij} + C_{ij}$$

$$C(\tau) = \int df\, S(f) \cos(2\pi f \tau)$$

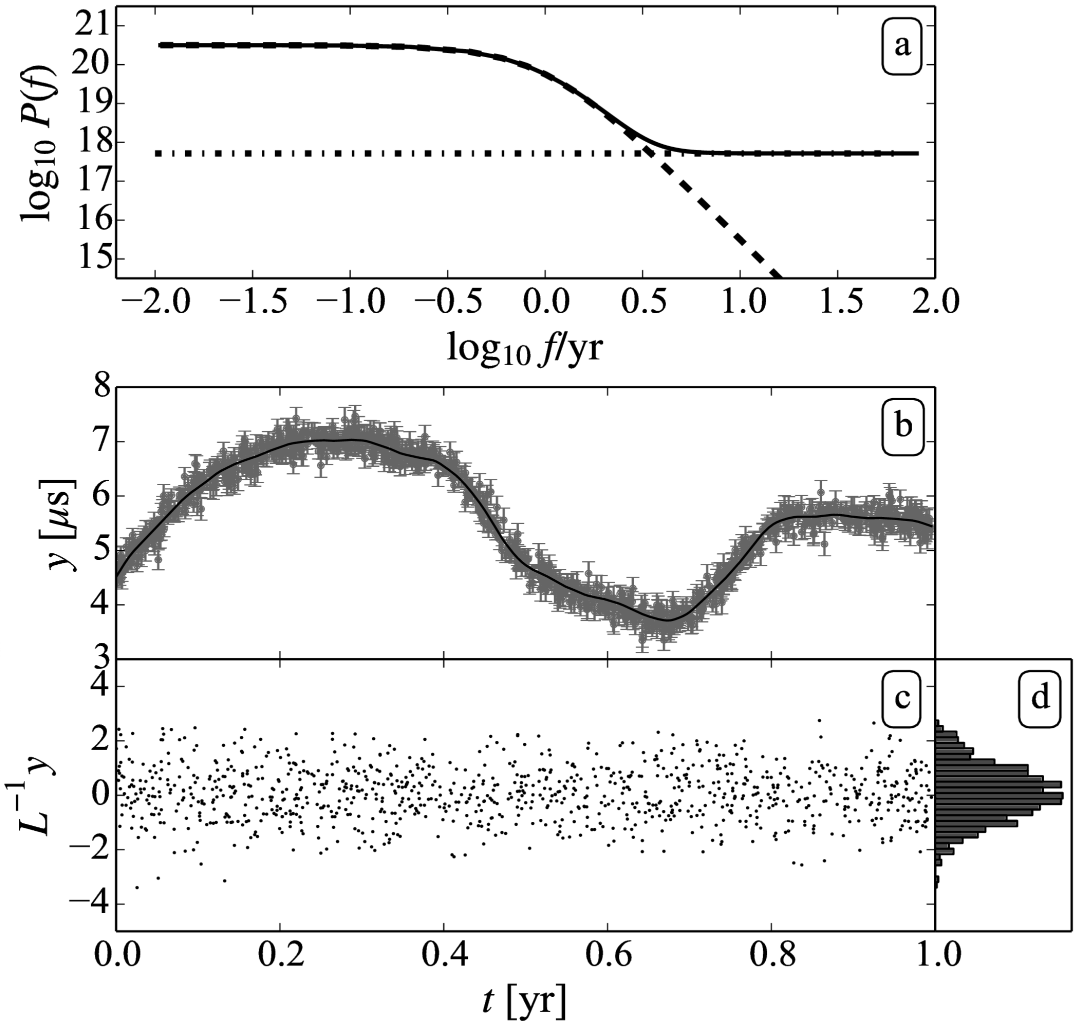
Reduce rank! $C = F \varphi F^T$



True

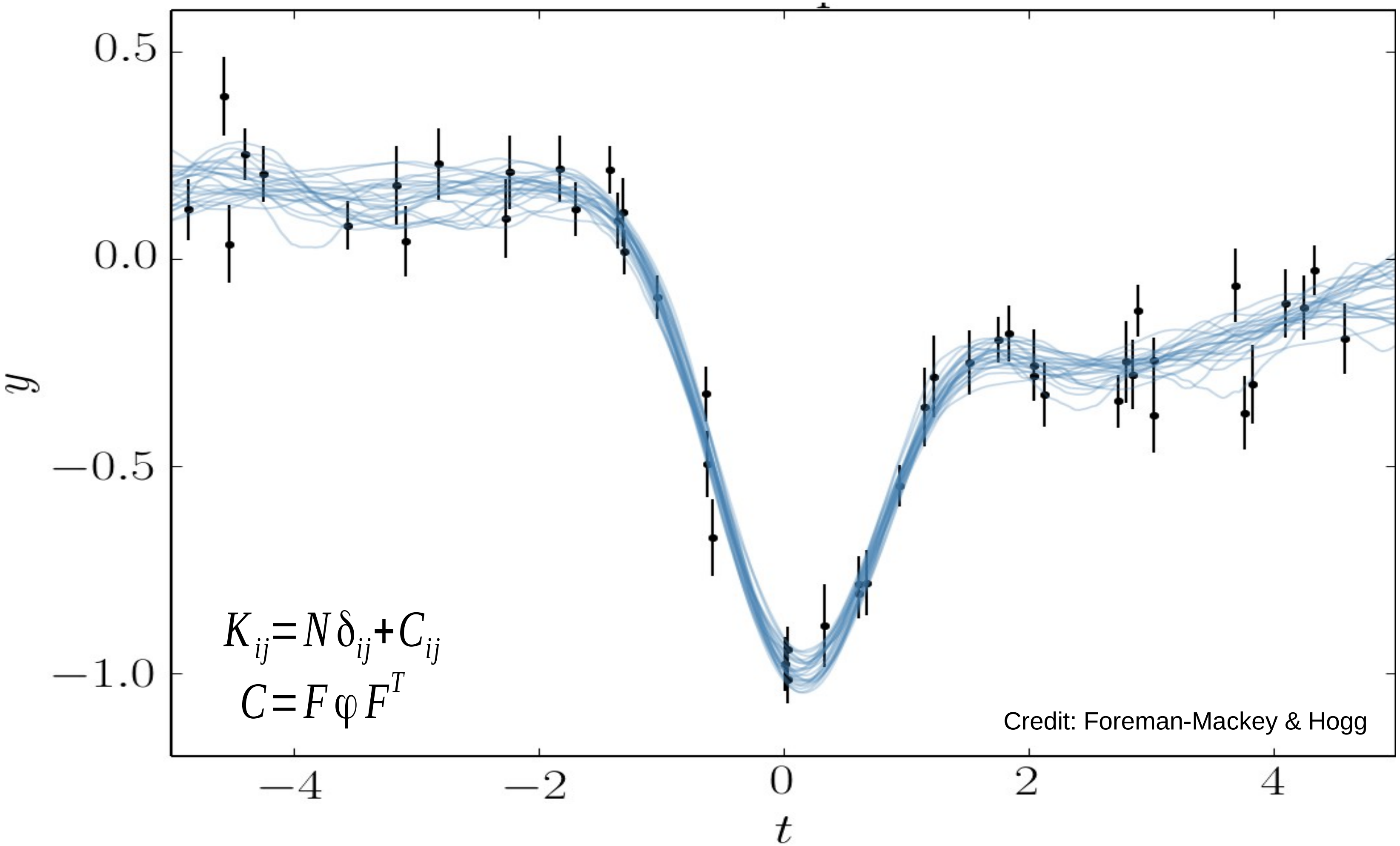


Reduced



van Haasteren & Vallisneri, MNRAS, 2014

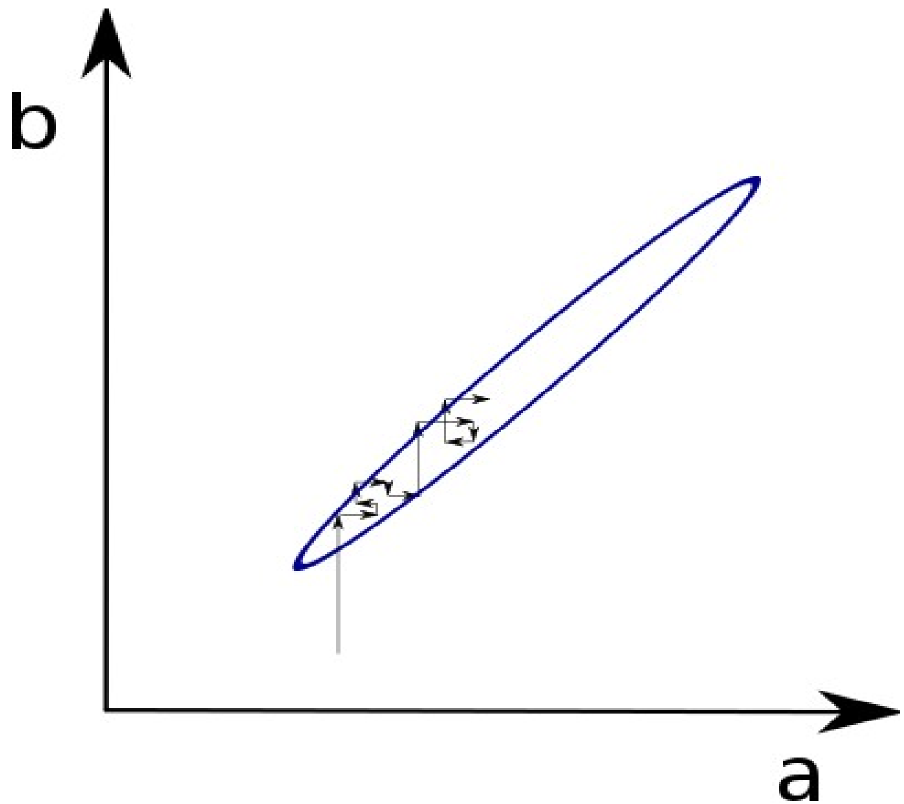
Pulsar timing & Kepler



Dimensionality reduction: ~100 x (speedup O(n^3))

Solution 2: reduce autocorrelation

- MCMC methods have limited efficiency. Characterized by autocorrelation length of the chain



See animation for Metropolis

Gibbs sampling

- Taken from the Planck data analysis: Gibbs sampling (van Haasteren & Vallisneri, PRD, 2014)

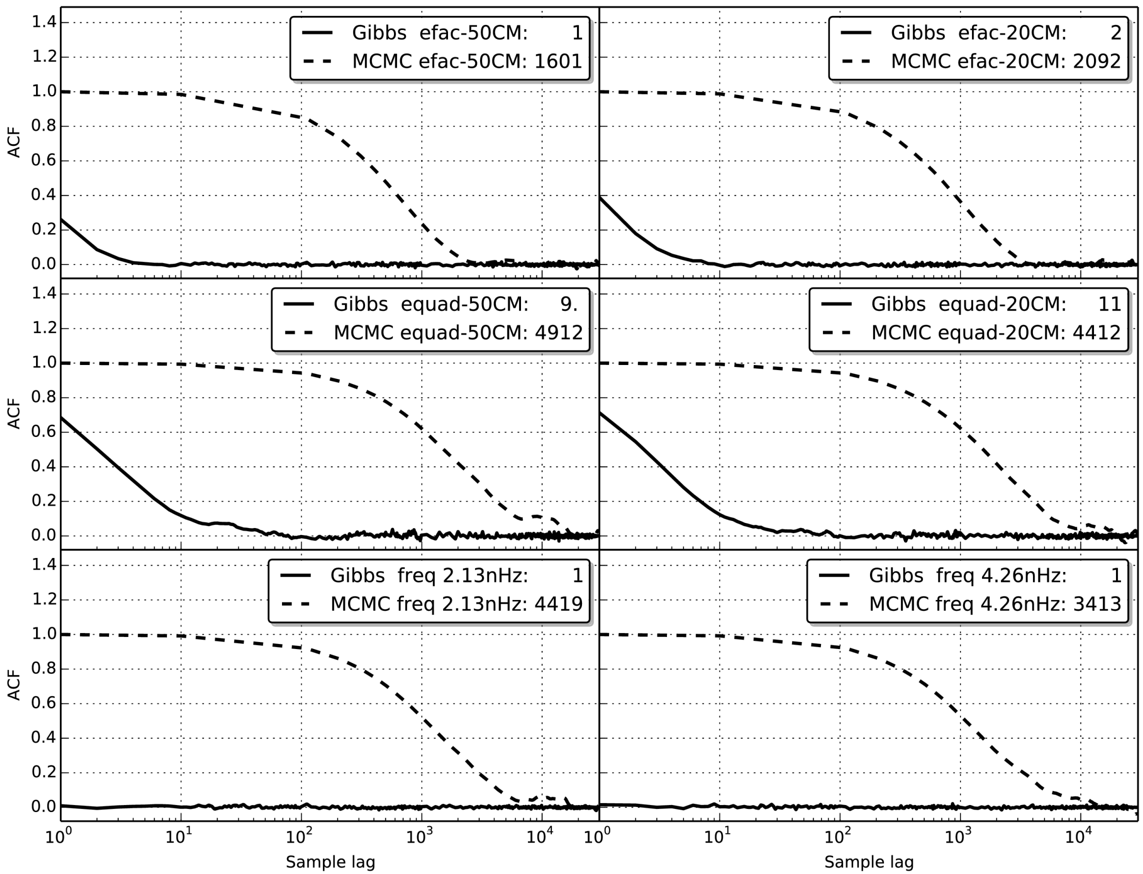
$$P\left(\delta t \,|\, \delta t_{TM}, \delta t_{RN}, \delta t_{DM}, \delta t_{J}, \{\sigma_i\}\right) P\left(\delta t_{RN} \,|\, A_{RN}, \gamma_{RN}\right) =$$

$$P(TOA) \propto \exp\left(-\frac{1}{2}\sum_i \frac{\left(TOA - f_0(t) - \sum_a \delta t_{a,i}\right)^2}{\sigma_i^2}\right) \exp \frac{\left(-\frac{1}{2}\delta t_{RN} C_{RN}^{-1} \delta t_{RN}\right)}{\sqrt{\det C_{RN}}}$$

- Group the parameters in non-covariant blocks! Sample the blocks analytically.

Gibbs sampling

- Gibbs sampling reduces the autocorrelation!



van Haasteren & Vallisneri, PRD, 2014

See animation

Summary

- Hierarchical modeling is awesome. We can do <u>anything</u>!
- Curse of dimensionality a pain. But we know tricks
- Reduce the number of dimensions where we can
- Exploit knowledge of non-covariant remaining parameters
- Revving up to search for GWs with these methods...
  …. which I'll tell you about next year