Spatio-Temporal Inference Strategies In The Quest For Gravitational Wave Detection With Pulsar Timing Arrays

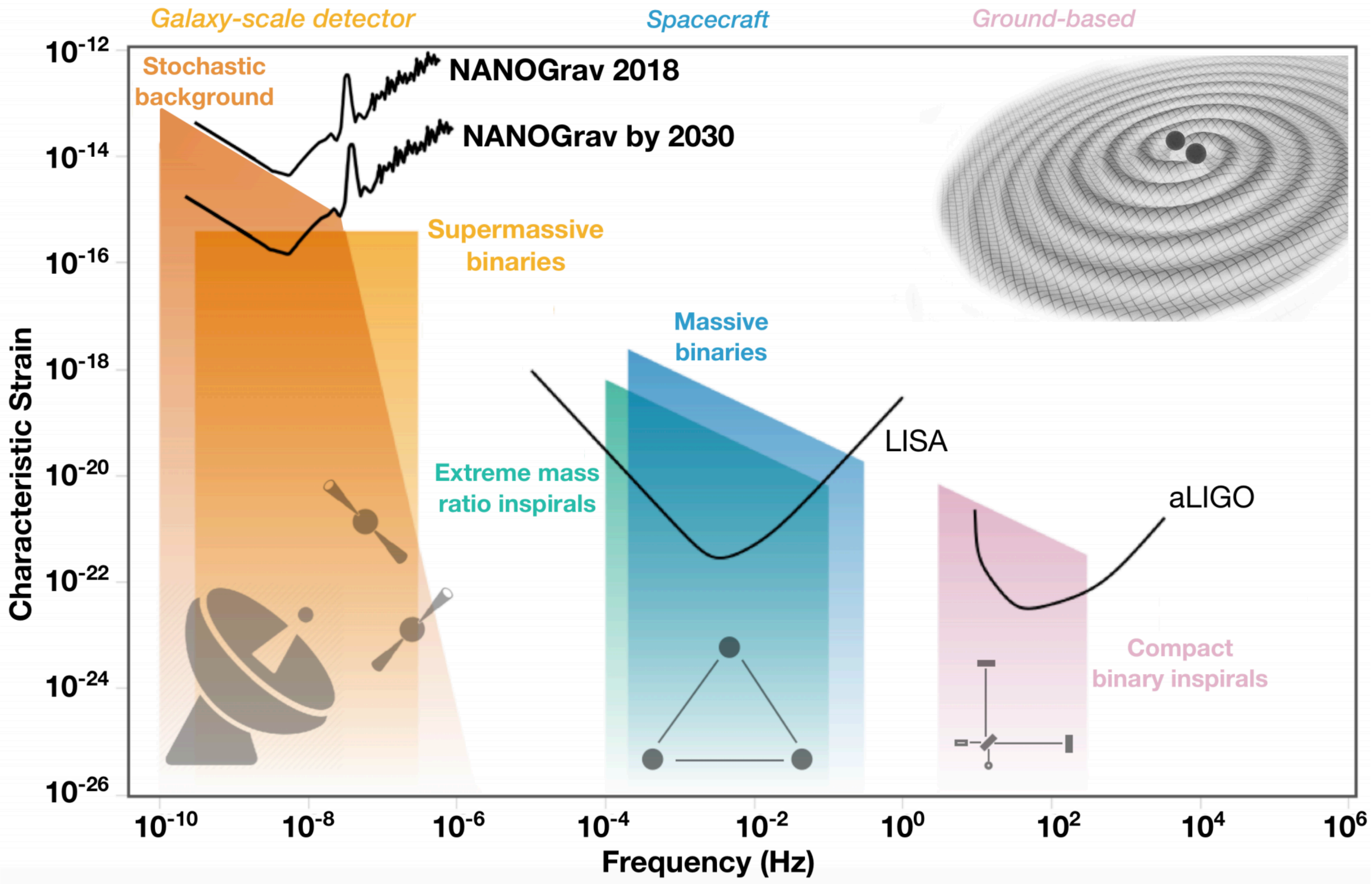
Stephen Taylor

Vanderbilt University

ICERM, Brown University, November 19th 2020

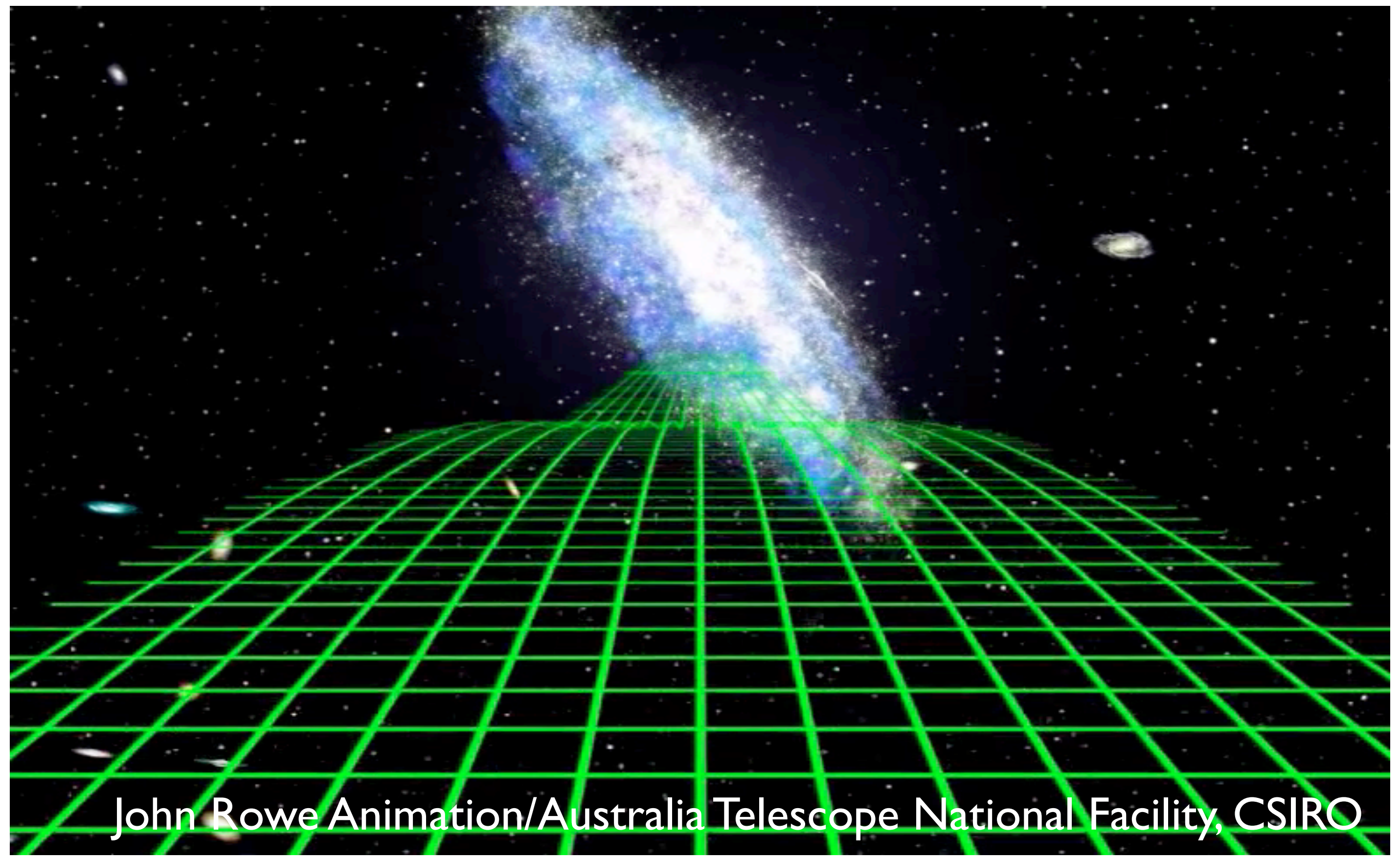
Image courtesy of *Science*, credit: Nicolle Rager Fuller *[modified]*

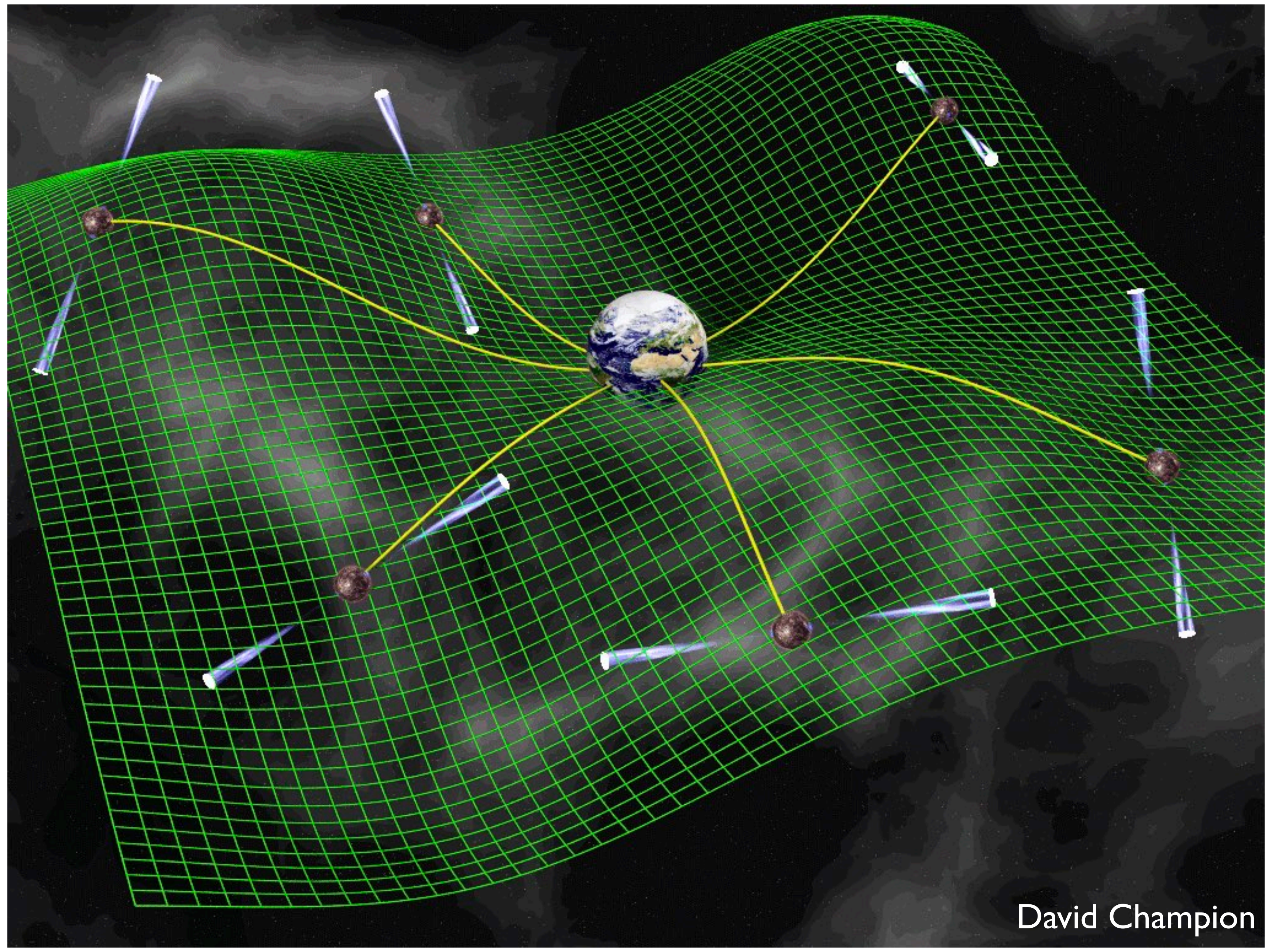
PTAs — The Elevator Pitch

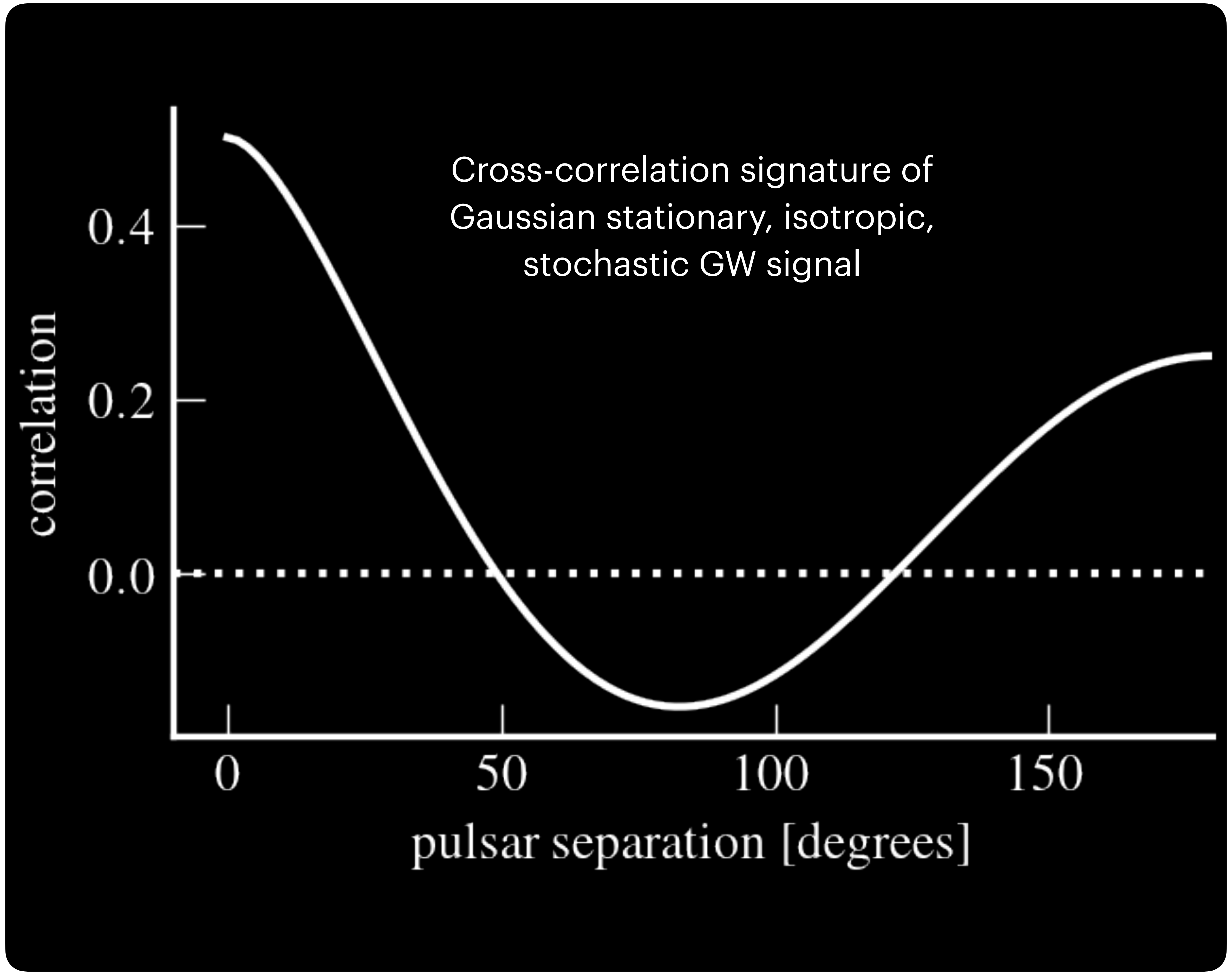


S. Taylor & C. Mingarelli, adapted from gwplotter.org (Moore, Cole, Berry 2014) and based on a figure in Mingarelli & Mingarelli (2018). Illustration of merging black holes adapted from R. Hurt/Caltech-JPL/EPA

PTAs — The Elevator Pitch



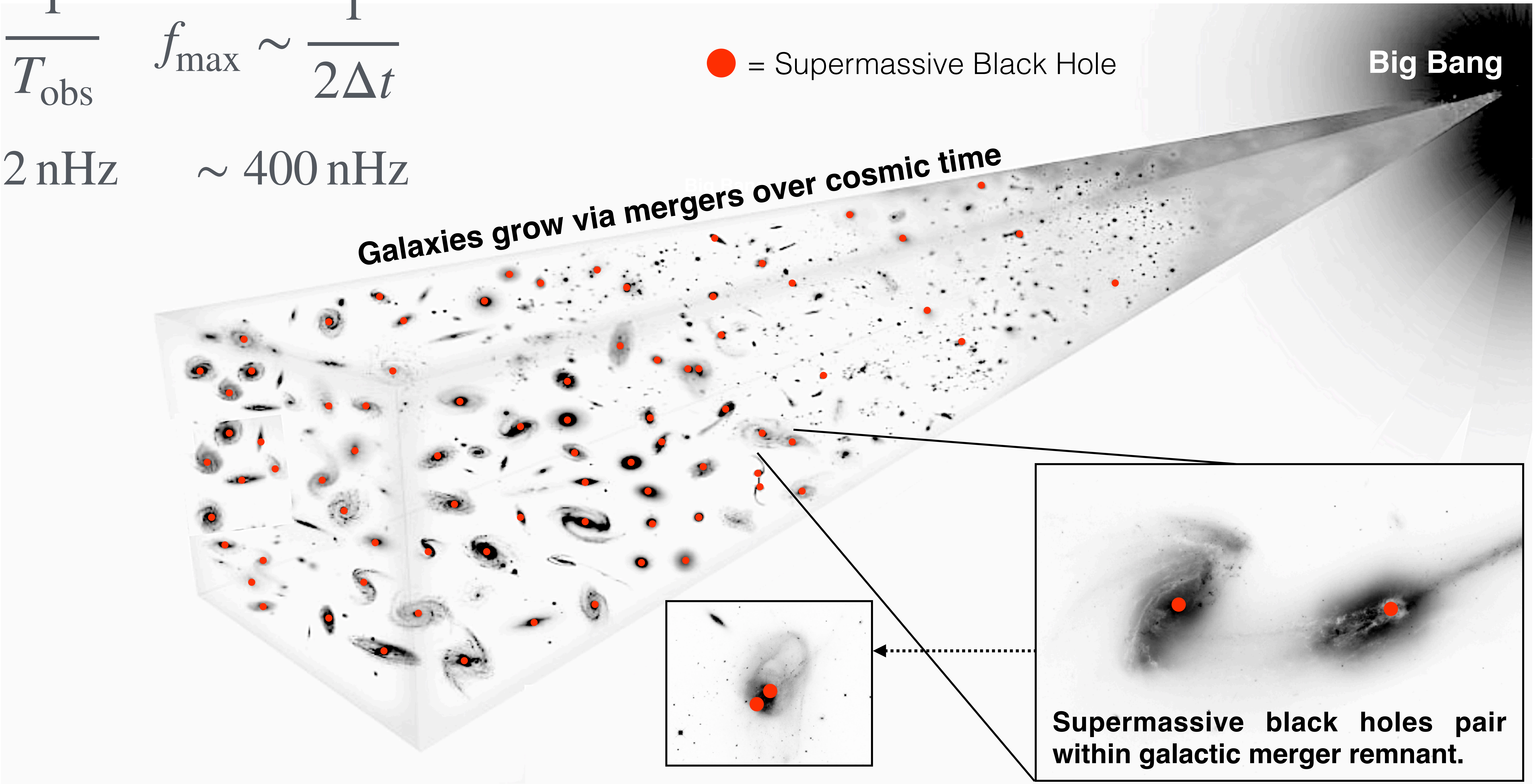




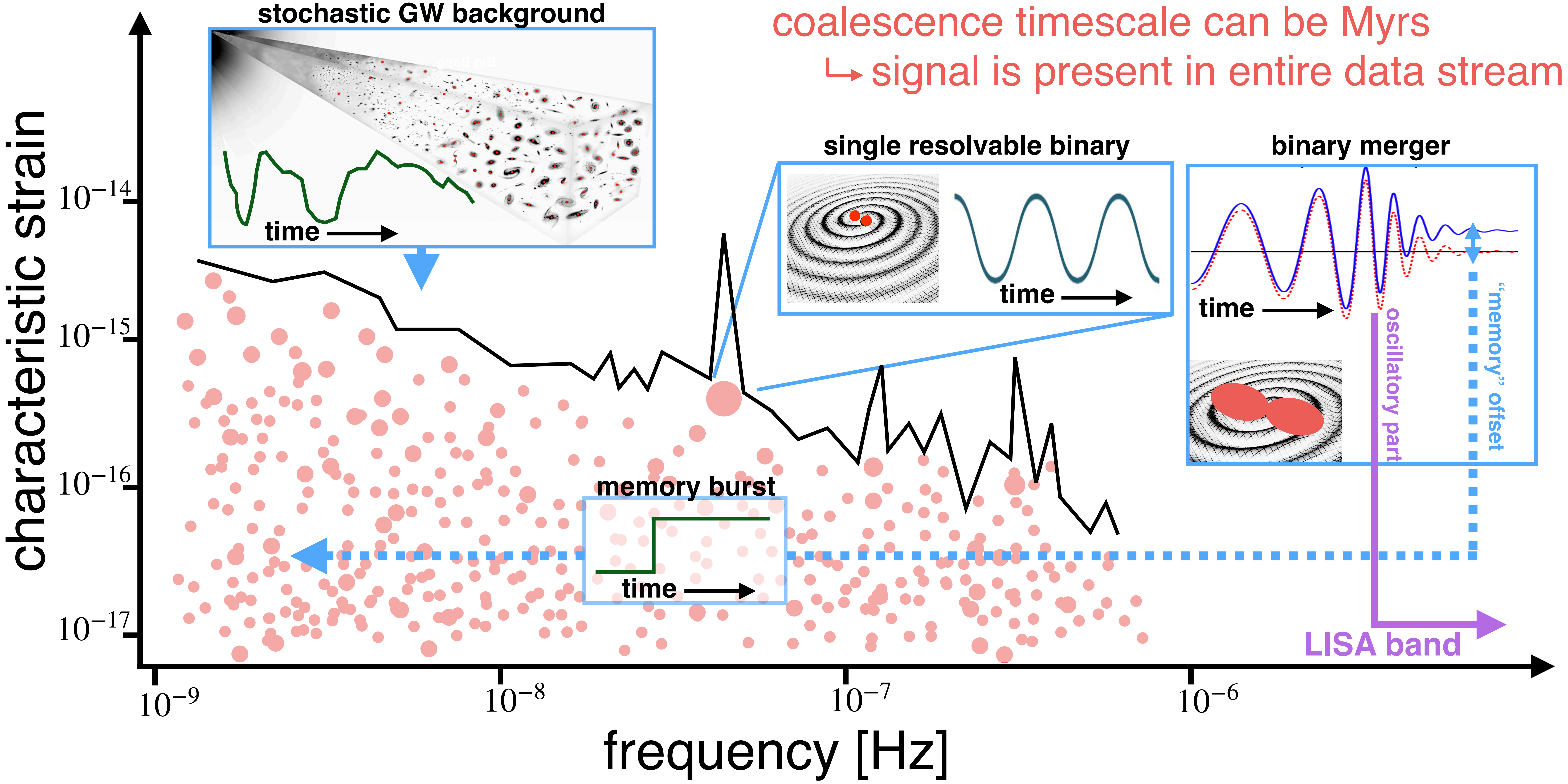
PTAs — The Elevator Pitch

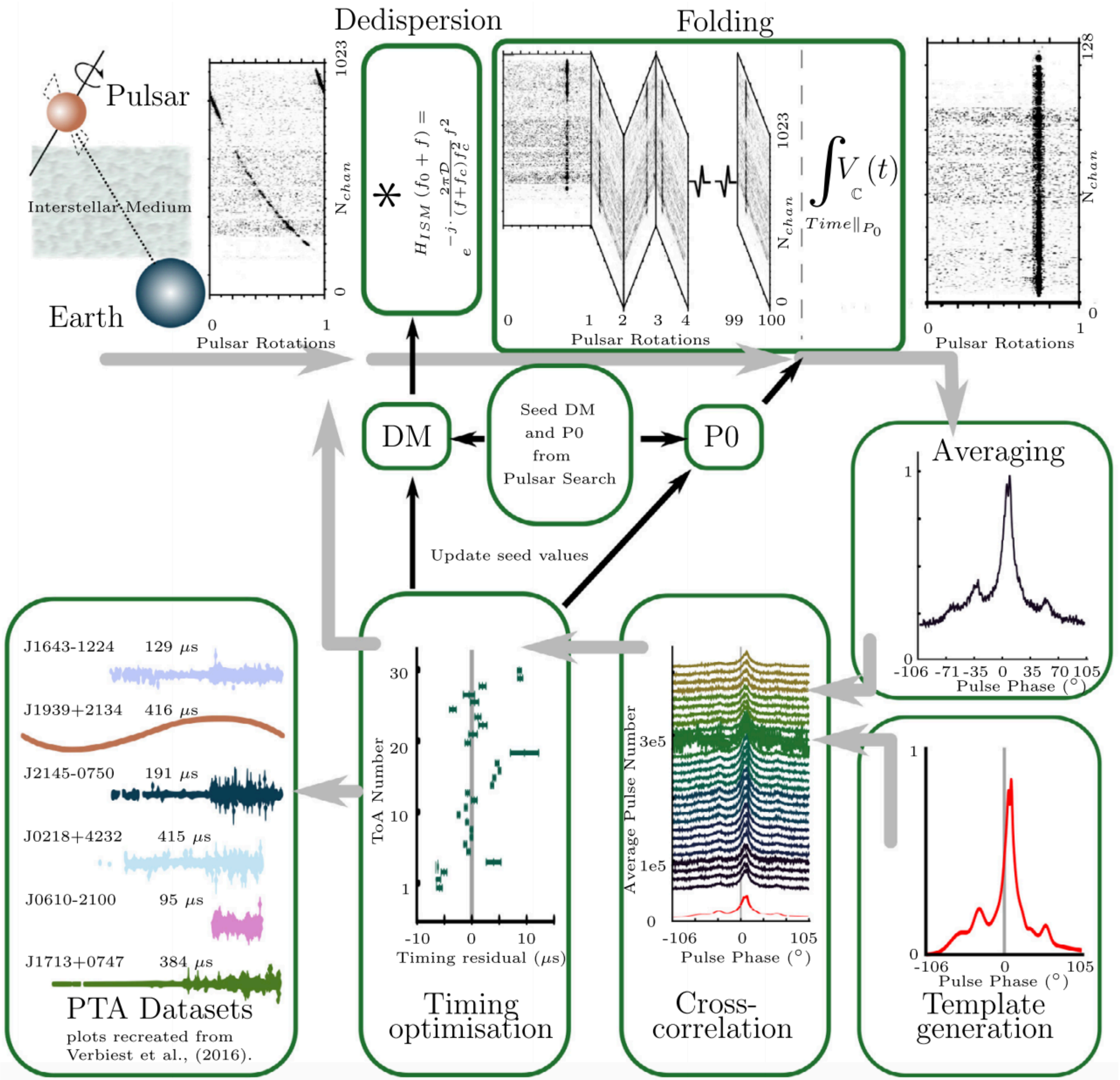
$$f_{\min} = \frac{1}{T_{\text{obs}}} \qquad f_{\max} \sim \frac{1}{2\Delta t}$$

$$\sim 2\,\text{nHz} \qquad \sim 400\,\text{nHz}$$



PTAs — The Elevator Pitch



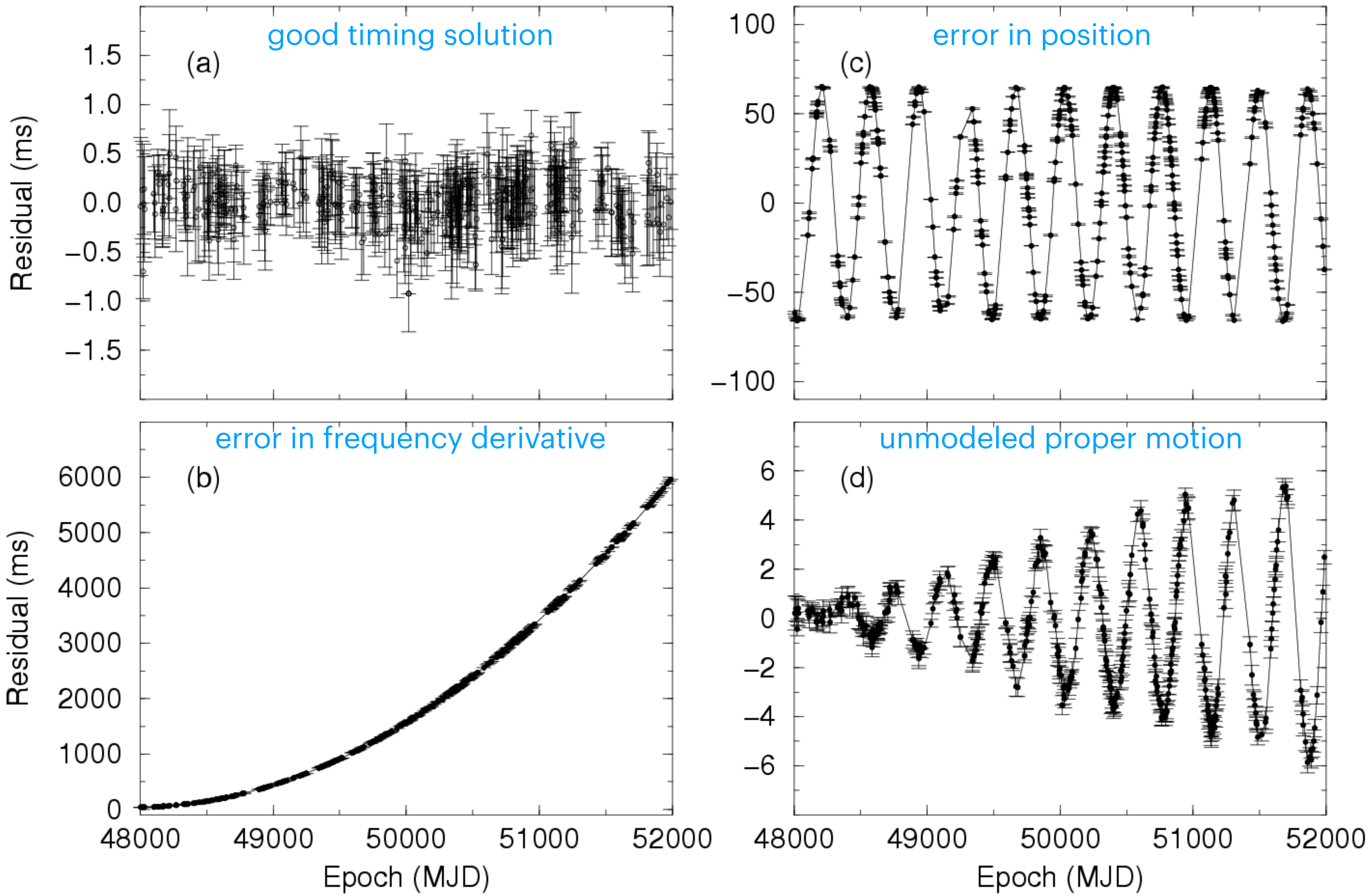


Verbiest & Shaifullah (2018)

From pulses to TOAs

**TOA = times of arrival*

Creating a timing ephemeris



Lorimer & Kramer (2005)

Pulsar-timing Data Model

random Gaussian processes

$$\vec{t}_{\mathrm{TOA}} = \vec{t}_{\mathrm{det}} + \vec{t}_{\mathrm{stoch}}$$

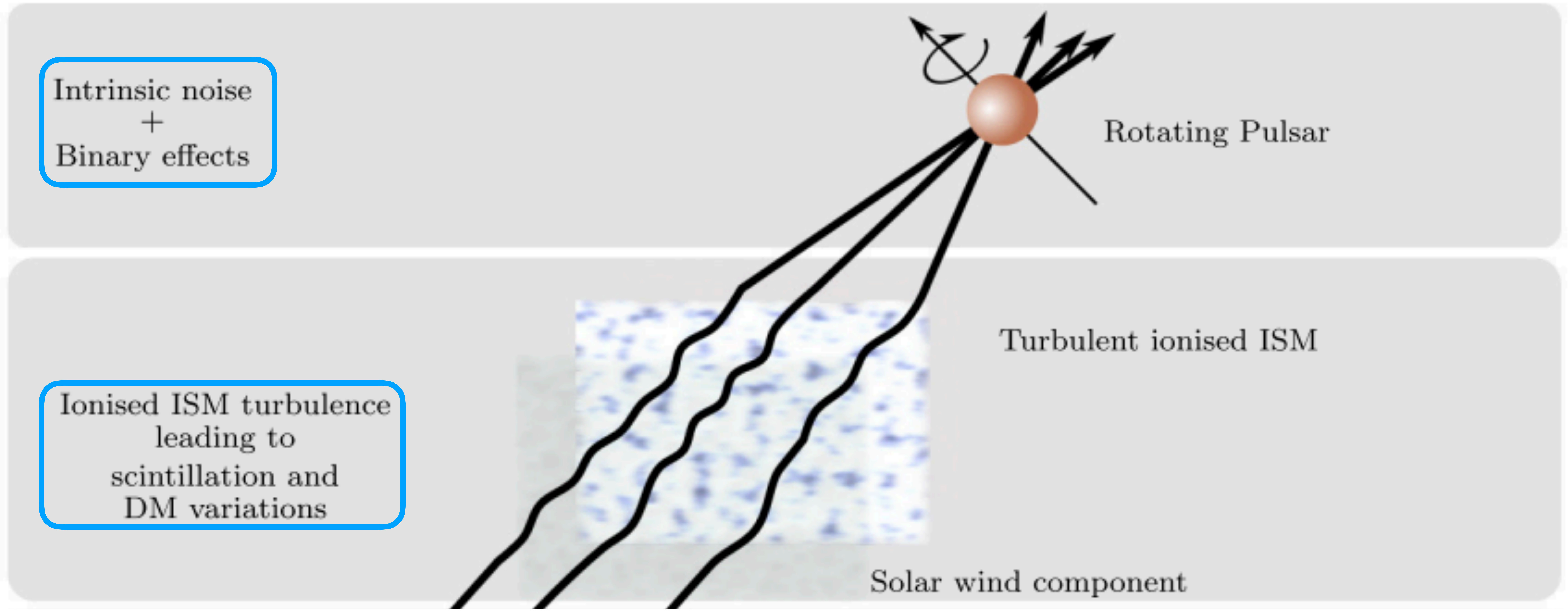
$$\overrightarrow{\delta t} \equiv \vec{t}_{\mathrm{TOA}} - \vec{t}_{\mathrm{det}}(\vec{\beta}_0)$$

Timing residuals

Deterministic	Stochastic
timing ephemeris	per-pulsar achromatic red noise
	per-pulsar white noise
transient noise features	per-pulsar chromatic red noise
single resolvable GW signals	interpulsar-correlated achromatic processes

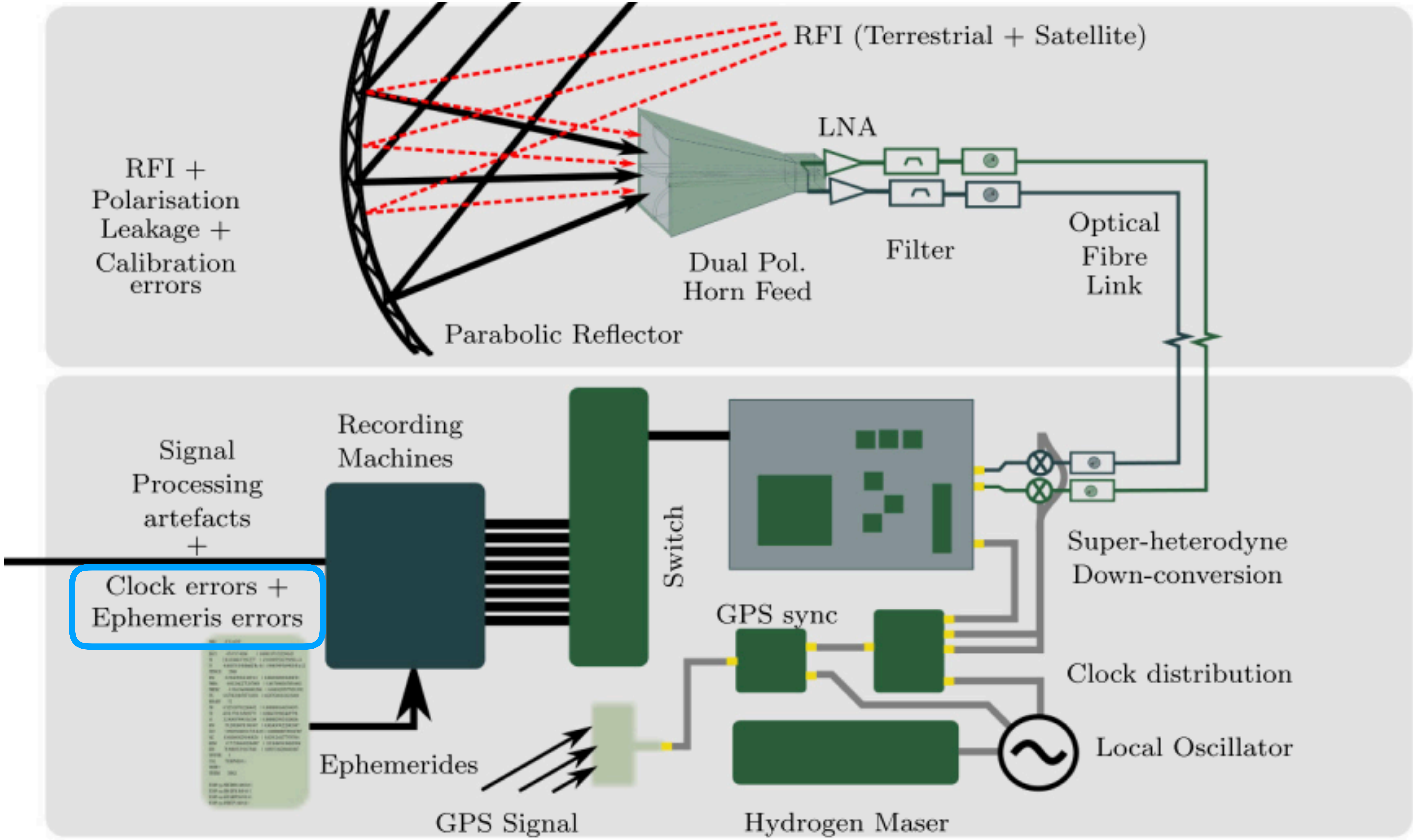
GWB

Sources of noise



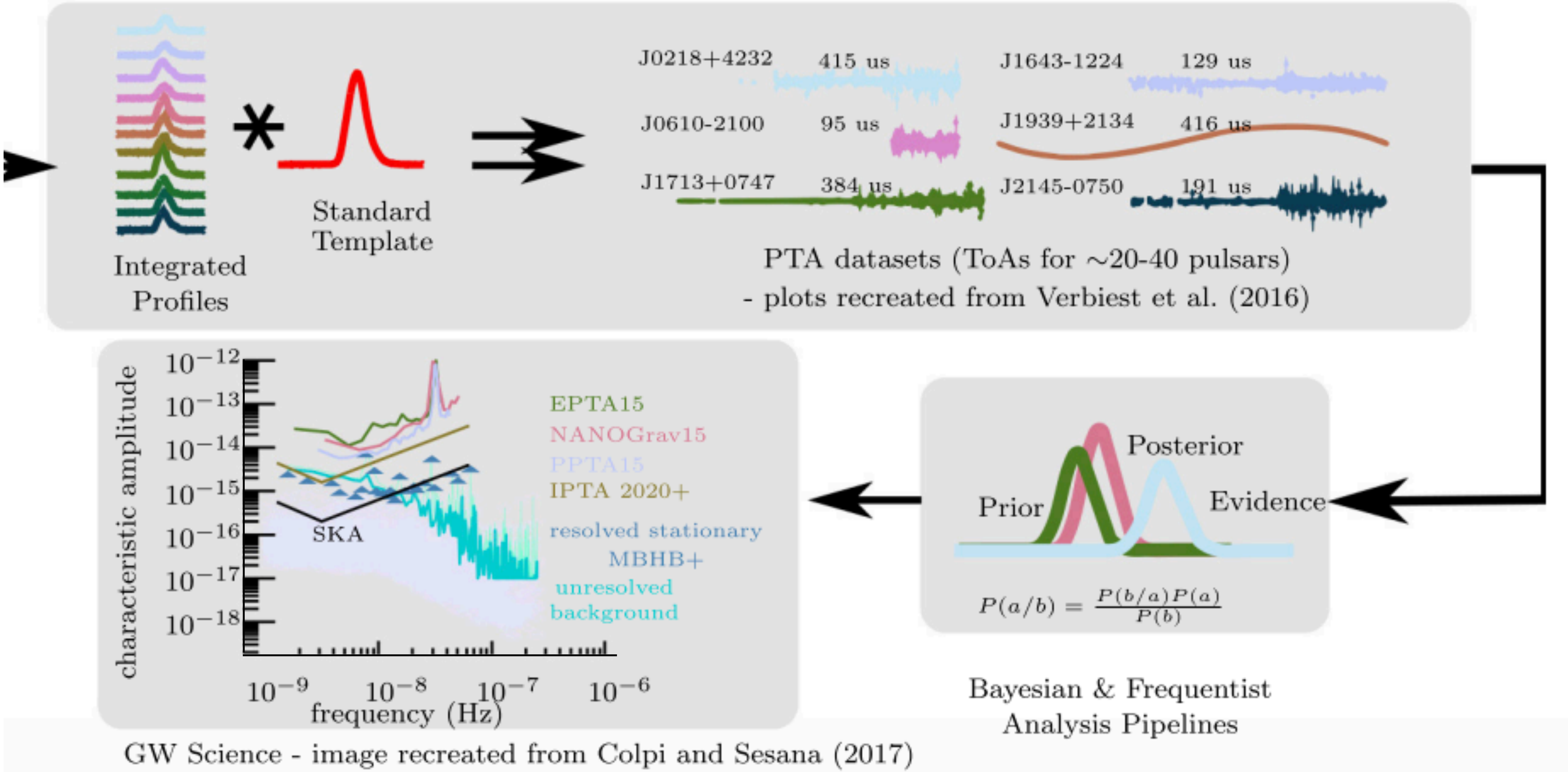
Verbiest & Shaifullah (2018)

Sources of noise



Verbiest & Shaifullah (2018)

Sources of noise



Verbiest & Shaifullah (2018)

Pulsar-timing Data Model

$$\delta t = \delta t_{\mathrm{tm}} + \delta t_{\mathrm{white}} + \delta t_{\mathrm{red}}$$

- **Deviations around best-fit of timing ephemeris**
- **White noise**
  - TOA measurement uncertainties
  - Extra unaccounted white-noise from receivers
  - Pulse phase "jitter"
- **Intrinsic low-frequency processes**
  - Rotational instabilities lead to random walk in phase, period, period-derivative
  - Radio-frequency dependent dispersion-measure variations
- **Spatially-correlated low-frequency processes**
  - Stochastic variations in time standards
  - Solar-system ephemeris errors
  - **Gravitational-wave background**

Timing Ephemeris

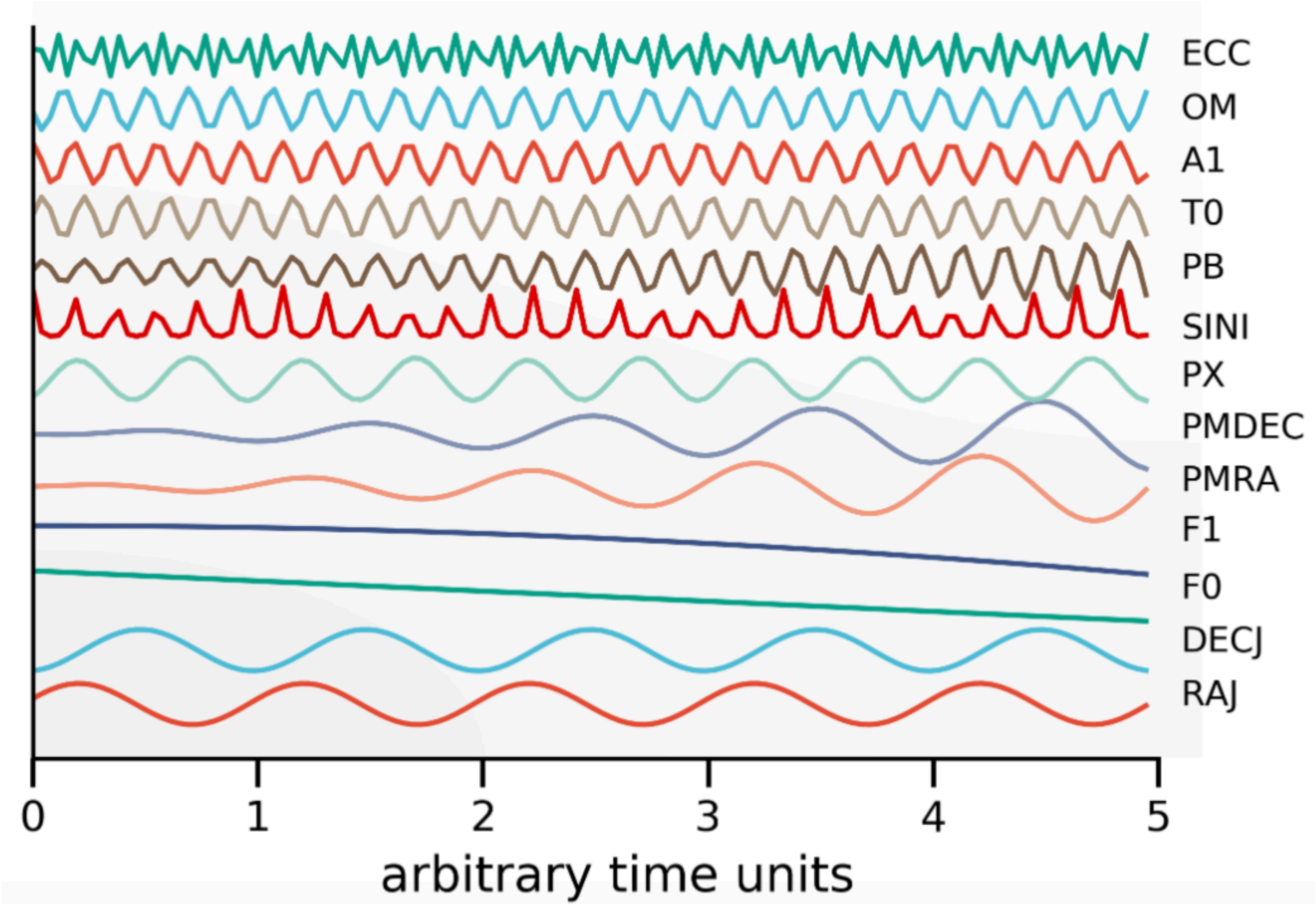
$$t_{\mathrm{det},i}(\vec{\beta}) = t_{\mathrm{det},i}(\vec{\beta}_0) + \left[\sum_j \left.\frac{\partial t_{\mathrm{det},i}}{\partial \beta_j}\right|_{\vec{\beta}_0} \times (\beta_j - \beta_{0,j}) \right]$$

$$\vec{t}_{\mathrm{det}}(\vec{\beta}) = \vec{t}_{\mathrm{det}}(\vec{\beta}_0) + \mathbf{M}\vec{\epsilon}$$

Timing ephemeris design
matrix for linear offsets

Timing Ephemeris

Temporal behavior of timing ephemeris basis



White Noise (1/2)

- Flat power-spectral density across all sampling frequencies
- No inter-pulsar correlations

$$\langle n_{i,\mu} n_{j,\nu} \rangle = F_\mu^2 \sigma_i^2 \delta_{ij} \delta_{\mu\nu} + Q_\mu^2 \delta_{ij} \delta_{\mu\nu}$$

EFAC = Extra FACtor
to correct uncertainties

"Radiometer noise"—
pulse template fitting
uncertainties

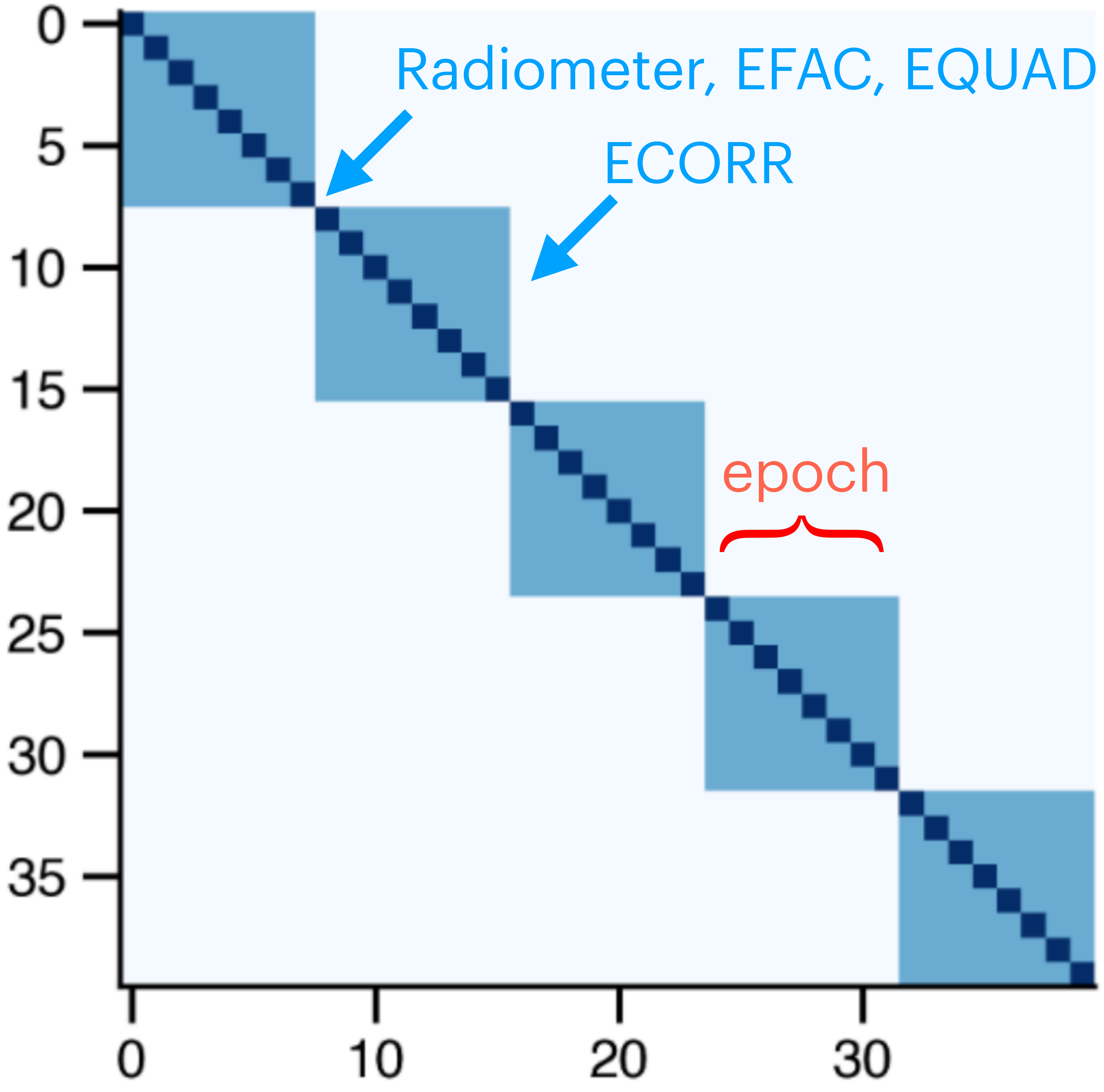
EQUAD = Extra QUADrature

White Noise (2/2)

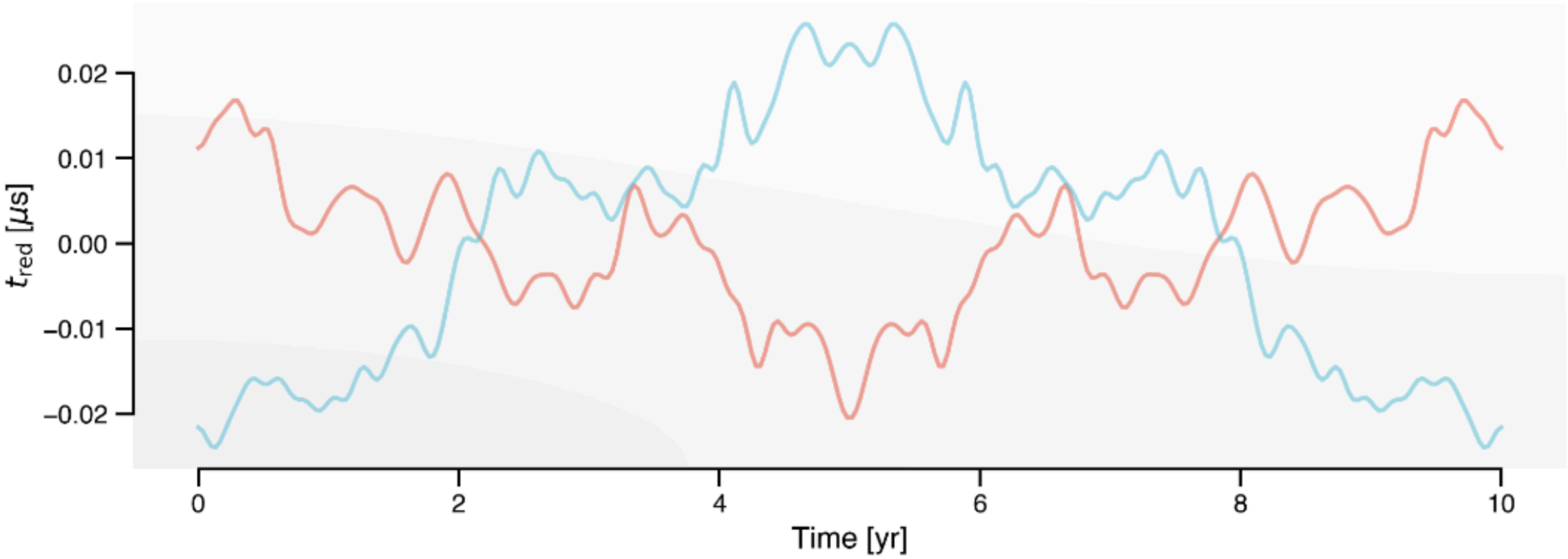
- Fitting a template to a finite-pulse folded observation can give "jitter" errors
- Simultaneous observations across many radio sub-bands in an epoch will have correlated jitter errors

$$\langle n^J_{i,\mu} n^J_{j,\nu} \rangle = J^2_\mu \delta_{e(i)e(j)} \delta_{\mu\nu}$$

ECORR = Extra CORRelated white noise



Red Processes (1/5)



Red Processes (2/5)

- Time-domain covariance matrix is large and dense

$$\langle \delta t_i \delta t_j \rangle = C(|t_i - t_j|)$$

- But we only care abut the lowest frequencies
- Use a rank-reduced formalism for covariance

$$\overrightarrow{\delta t}_{\mathrm{red}} = \mathbf{F}\vec{a}$$

$$\langle \overrightarrow{\delta t}_{\mathrm{red}} \overrightarrow{\delta t}^{\mathrm{T}}_{\mathrm{red}} \rangle = \mathbf{F} \langle \vec{a}\,\vec{a}^{\mathrm{T}} \rangle \mathbf{F}^{\mathrm{T}}$$

$$C = \mathbf{F} \phi \mathbf{F}^{\mathrm{T}}$$

Red Processes (3/5)

$$\overrightarrow{\delta t}_{\rm red} = \mathbf{F}\,\vec{a}$$

Fourier design matrix over small number of modes

$$\mathbf{F} = \begin{pmatrix} \sin(2\pi t_1/T) & \cos(2\pi t_1/T) & \cdots & \sin(2\pi N_f t_1/T) & \cos(2\pi N_f t_1/T) \\ \sin(2\pi t_2/T) & \cos(2\pi t_2/T) & \cdots & \sin(2\pi N_f t_2/T) & \cos(2\pi N_f t_2/T) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sin(2\pi t_N/T) & \cos(2\pi t_N/T) & \cdots & \sin(2\pi N_f t_N/T) & \cos(2\pi N_f t_N/T) \end{pmatrix}$$

Red Processes (4/5)

$$\overrightarrow{\delta t}_{\rm red} = \mathbf{F}\boxed{\vec{a}}$$

Fourier coefficients

$$p(\vec{a}\,|\,\vec{\eta}) = \frac{\exp\left(-\frac{1}{2}\vec{a}^{\rm T}\phi(\vec{\eta})^{-1}\vec{a}\right)}{\sqrt{\det(2\pi\phi(\vec{\eta}))}}$$

$$[\phi]_{(ak)(bj)} = \Gamma_{ab}\rho_k\delta_{kj} + \kappa_{ak}\delta_{kj}\delta_{ab}$$

Overlap Reduction Function

GWB PSD

Intrinsic red-noise PSD

Red Processes (5/5)

- power laws
- per frequency
- GP emulators

GWB PSD

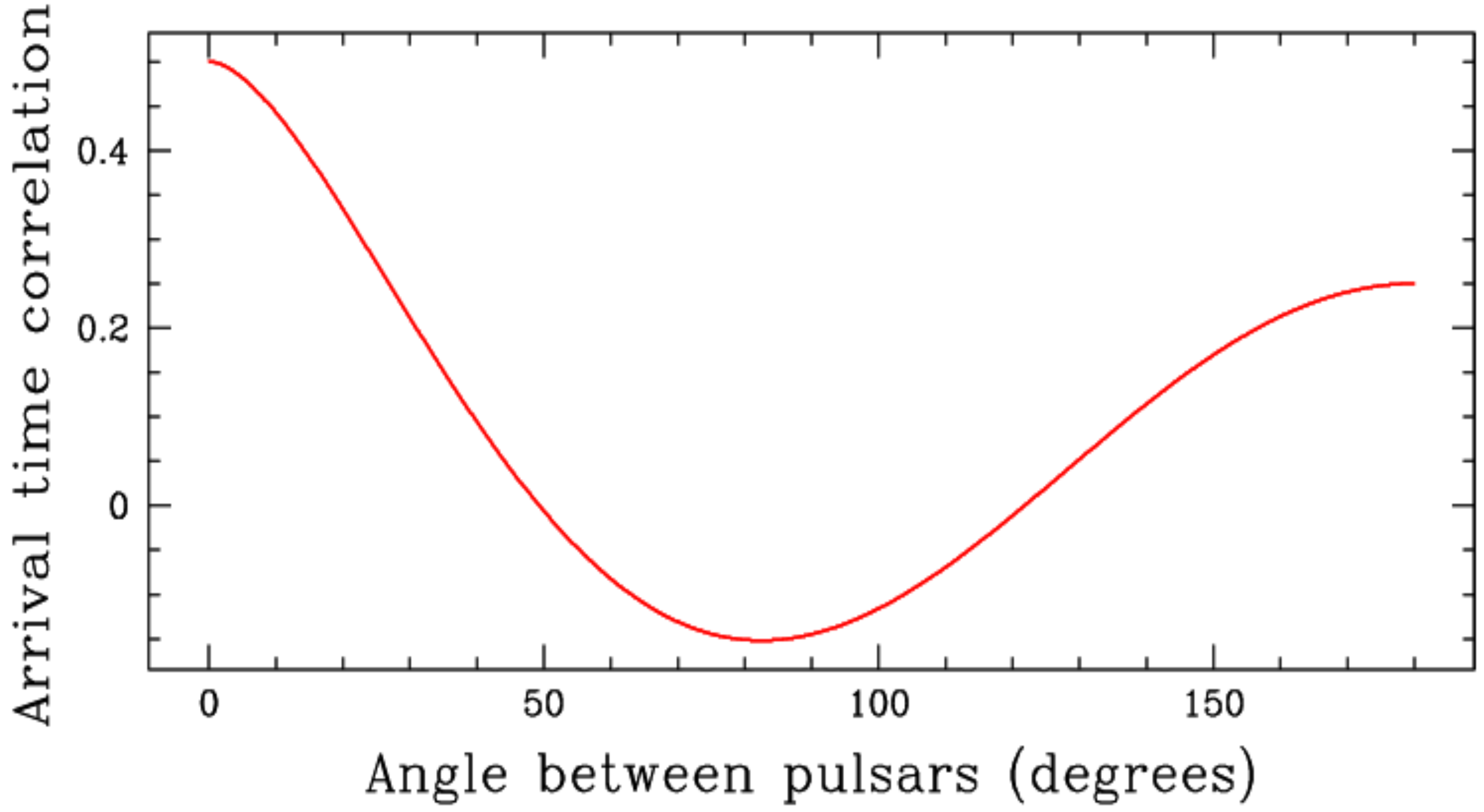
$$\rho(f) = S(f)\Delta f = \frac{h_c(f)^2}{12\pi^2 f^3}\frac{1}{T}$$

GWB ORF

$$\Gamma_{ab} \propto (1+\delta_{ab})\int_{S^2} d^2\hat{\Omega}\, P(\hat{\Omega})\left[F_a^{+}(\hat{\Omega})F_b^{+}(\hat{\Omega}) + F_a^{\times}(\hat{\Omega})F_b^{\times}(\hat{\Omega})\right]$$

PTA overlap reduction function for Gaussian stationary, isotropic stochastic GWB

"Hellings & Downs Curve" (1983)



The PTA Likelihood

$$\vec{\delta t} = M\vec{\epsilon} + F\vec{a} + U\vec{j} + \vec{n}$$

small linear perturbations around best-fit timing solution

low-frequency processes in Fourier basis

jitter

white noise

$$[M] = N_{\mathrm{TOA}} \times N_{\mathrm{tm}}$$
$$[\vec{\epsilon}] = N_{\mathrm{tm}}$$

"M" is matrix of TOA derivatives wrt timing-model parameters

~ few tens

$$[F] = N_{\mathrm{TOA}} \times 2N_{\mathrm{freqs}}$$
$$[\vec{a}] = 2N_{\mathrm{freqs}}$$

"F" has columns of sines and cosines for each frequency

~ few tens

$$[U] = N_{\mathrm{TOA}} \times N_{\mathrm{epochs}}$$
$$[\vec{j}] = N_{\mathrm{epochs}}$$

"U" has block diagonal structure, with ones filling each block

~ couple of hundred

Lentati et al. (inc Taylor) (2013)
van Haasteren & Vallisneri (2014a,b)

The PTA Likelihood

Start with Gaussian white noise likelihood

$$p(\vec{n}) = \frac{\exp\left(-\frac{1}{2}\vec{n}^{\mathrm{T}} N^{-1} \vec{n}\right)}{\sqrt{\det(2\pi N)}}$$

$$p(\vec{\delta t}\,|\,\vec{\epsilon}, \vec{a}, \vec{j}) = \frac{\exp\left[-\frac{1}{2}\left(\vec{\delta t} - M\vec{\epsilon} - F\vec{a} - U\vec{j}\right)^{\mathrm{T}} N^{-1}\left(\vec{\delta t} - M\vec{\epsilon} - F\vec{a} - U\vec{j}\right)\right]}{\sqrt{\det(2\pi N)}}$$

$$p(\vec{\delta t}\,|\,\vec{b}) = \frac{\exp\left[-\frac{1}{2}\left(\vec{\delta t} - T\vec{b}\right)^{\mathrm{T}} N^{-1}\left(\vec{\delta t} - T\vec{b}\right)\right]}{\sqrt{\det(2\pi N)}}$$

$$T\vec{b} = M\vec{\epsilon} + F\vec{a} + U\vec{j}$$

$$\mathbf{b} = \begin{bmatrix} \boldsymbol{\epsilon} \\ \mathbf{a} \\ \mathbf{j} \end{bmatrix} \qquad T = [M \quad F \quad U]$$

The PTA Likelihood

But we're describing all stochastic terms as random Gaussian processes...

$$p(\vec{b}\,|\,\vec{\eta}) = \frac{\exp\left(-\frac{1}{2}\vec{b}^{\mathrm{T}}\mathbf{B}^{-1}\vec{b}\right)}{\sqrt{\det(2\pi\mathbf{B})}}$$

$$\boldsymbol{B} = \begin{pmatrix} \infty & 0 \\ 0 & \phi \end{pmatrix}$$

$$p(\vec{\eta}, \vec{b}\,|\,\vec{\delta t}) \propto p(\vec{\delta t}\,|\,\vec{b})p(\vec{b}\,|\,\vec{\eta})p(\vec{\eta})$$

hierarchical modelling

$$p(\vec{\eta}\,|\,\vec{\delta t}) = \int p(\vec{\eta}\,|\,\vec{\delta t})\,\mathrm{d}\vec{b}$$

(analytically!) marginalize over coefficients

$$p(\vec{\eta}\,|\,\vec{\delta t}) \propto \frac{\exp\left(-\frac{1}{2}\vec{\delta t}^{\mathrm{T}}\mathbf{C}^{-1}\vec{\delta t}\right)}{\sqrt{\det(2\pi\mathbf{C})}}p(\vec{\eta})$$

$$C = N + TBT^{\mathrm{T}}$$

The PTA Likelihood

$$C = N + TBT^{\mathrm{T}}$$

what are we actually doing here?

$$[TBT^{\mathrm{T}}]_{(ab),\tau} = \sum_{k}^{N_f} [\phi]_{ab} \cos(2\pi k\tau/T)$$

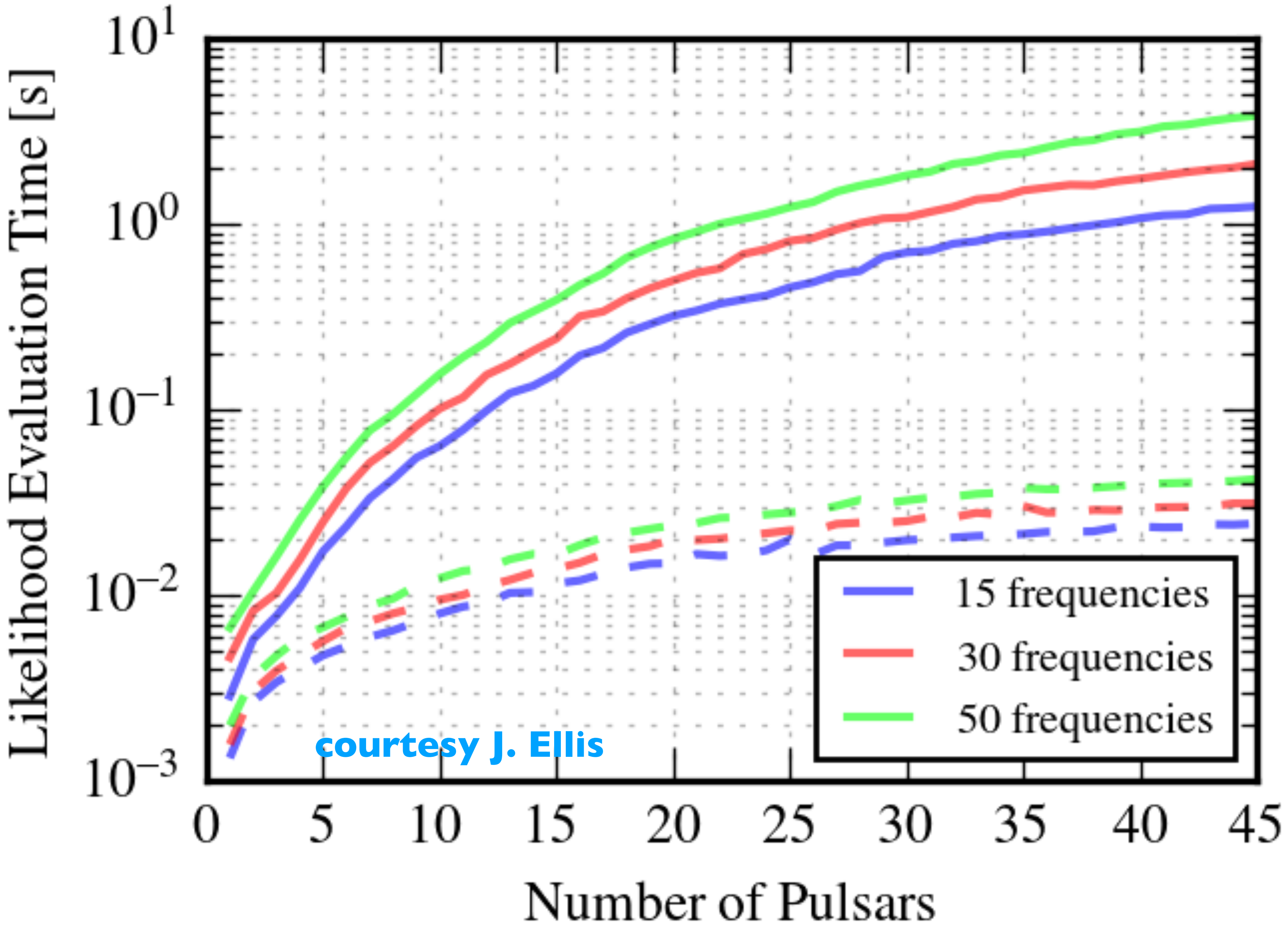
this is just the Wiener-Khinchin theorem!

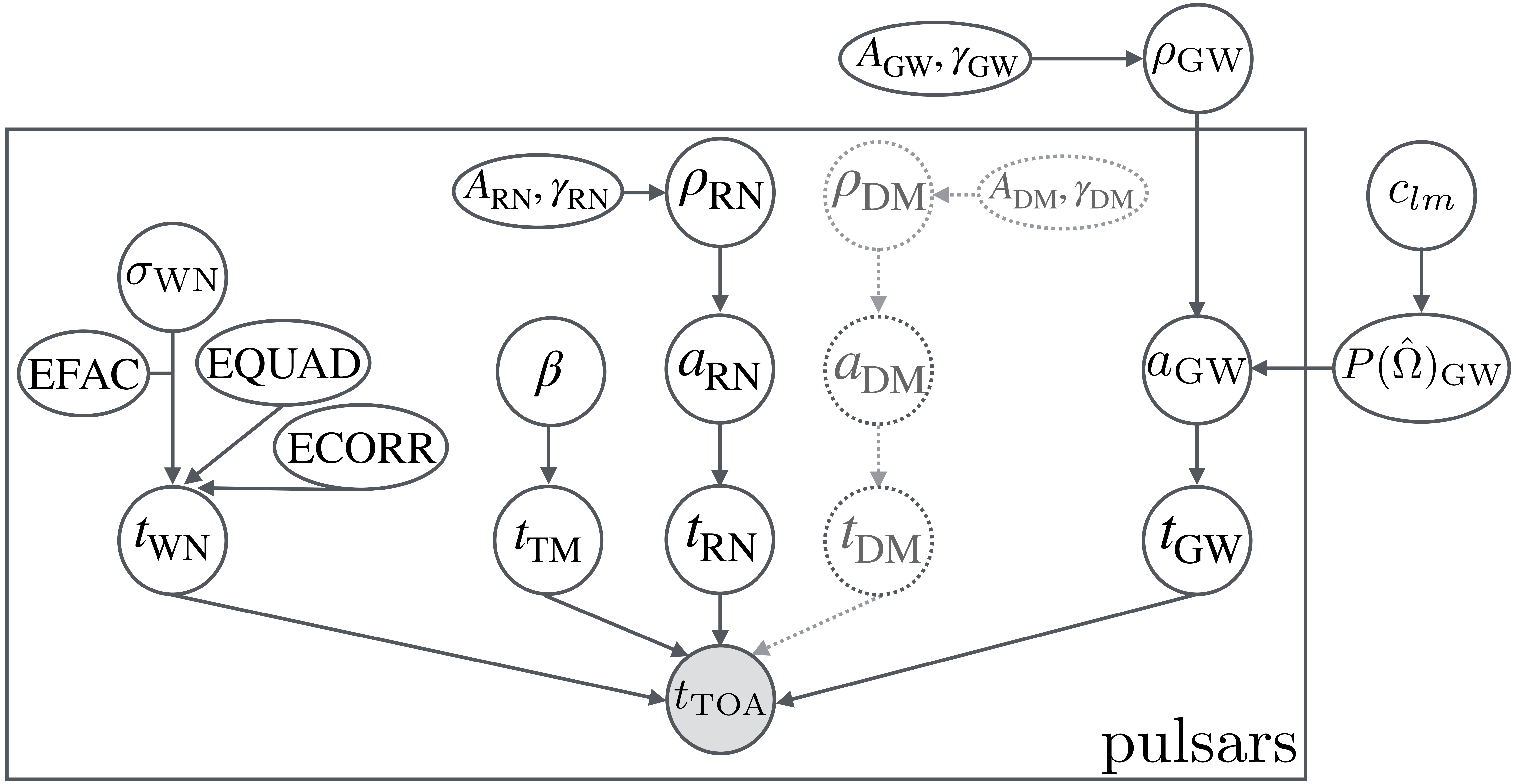
Woodbury lemma

$$\begin{aligned} C^{-1} &= (N^{-1} + TBT^{\mathrm{T}})^{-1} \\ &= N^{-1} - N^{-1}T\underbrace{(B^{-1} + T^{\mathrm{T}}N^{-1}T)^{-1}}T^{\mathrm{T}}N^{-1} \end{aligned}$$

Much easier and faster than $N_{\mathrm{TOA}} \times N_{\mathrm{TOA}}$ inversion

The PTA Likelihood





Without inter-pulsar correlations
[~ tens of ms]

With inter-pulsar correlations
[~few seconds]

The PTA Bayesian Network

The NANOGrav 12.5-year Data Set: Search For An Isotropic Stochastic Gravitational-Wave Background

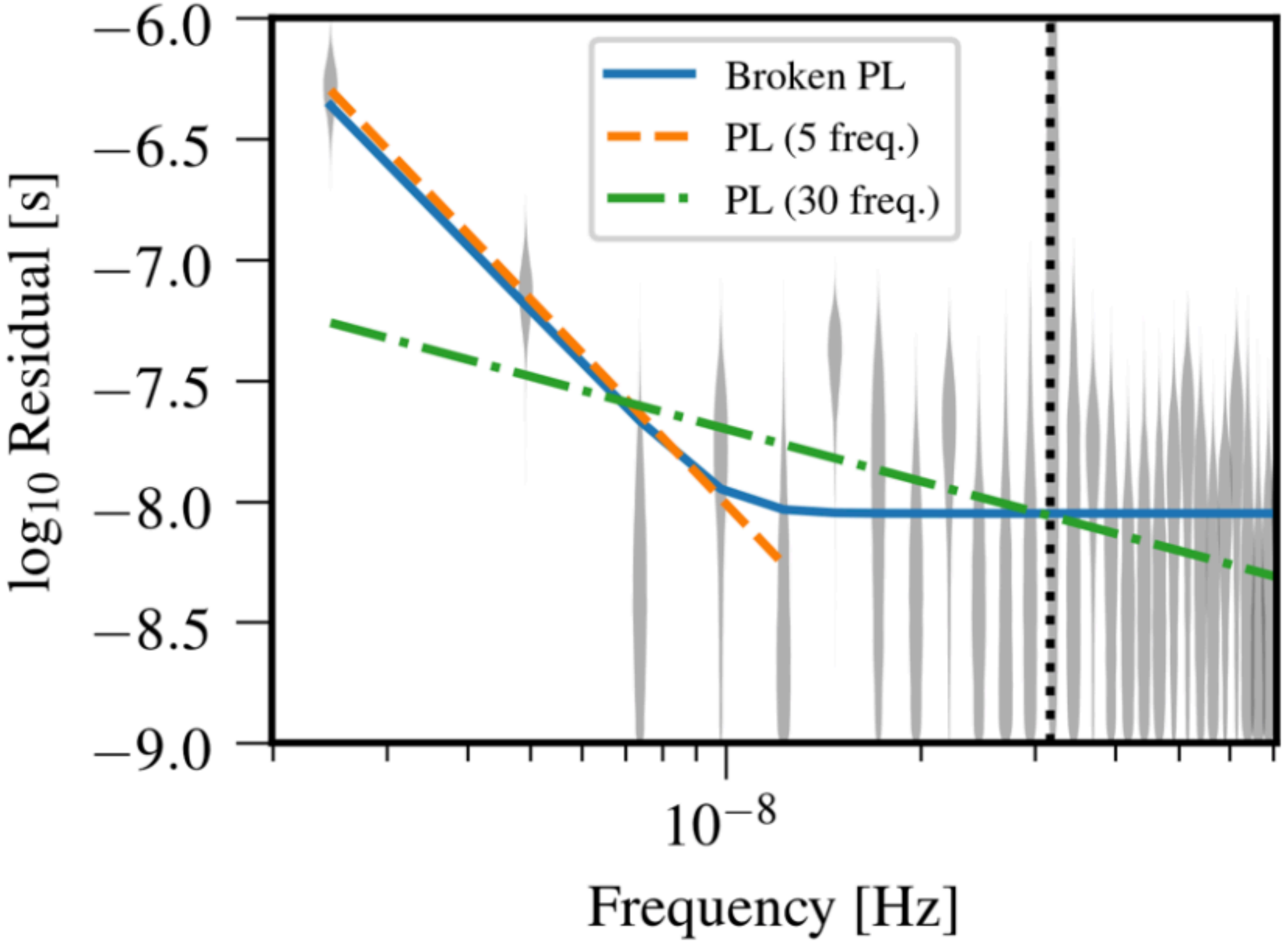
Zaven Arzoumanian,[1] Paul T. Baker,[2] Harsha Blumer,[3,4] Bence Bécsy,[5] Adam Brazier,[6] Paul R. Brook,[3,4] Sarah Burke-Spolaor,[3,4,7] Shami Chatterjee,[6] Siyuan Chen,[8,9,10] James M. Cordes,[6] Neil J. Cornish,[5] Fronefield Crawford,[11] H. Thankful Cromartie,[12] Megan E. DeCesar,[13,14,*] Paul B. Demorest,[15] Timothy Dolch,[16] Justin A. Ellis,[17] Elizabeth C. Ferrara,[18] William Fiore,[3,4] Emmanuel Fonseca,[19] Nathan Garver-Daniels,[3,4] Peter A. Gentile,[3,4] Deborah C. Good,[20] Jeffrey S. Hazboun,[21,*] A. Miguel Holgado,[22] Kristina Islo,[23] Ross J. Jennings,[6] Megan L. Jones,[23] Andrew R. Kaiser,[3,4] David L. Kaplan,[23] Luke Zoltan Kelley,[24] Joey Shapiro Key,[21] Nima Laal,[25] Michael T. Lam,[26,27] T. Joseph W. Lazio,[28] Duncan R. Lorimer,[3,4] Jing Luo,[29] Ryan S. Lynch,[30] Dustin R. Madison,[3,4,*] Maura A. McLaughlin,[3,4] Chiara M. F. Mingarelli,[31,32] Cherry Ng,[33] David J. Nice,[13] Timothy T. Pennucci,[34,35,*] Nihan S. Pol,[3,4] Scott M. Ransom,[34] Paul S. Ray,[36] Brent J. Shapiro-Albert,[3,4] Xavier Siemens,[25,23] Joseph Simon,[28,37] Renée Spiewak,[38] Ingrid H. Stairs,[20] Daniel R. Stinebring,[39] Kevin Stovall,[15] Jerry P. Sun,[25] Joseph K. Swiggum,[13,*] Stephen R. Taylor,[40] Jacob E. Turner,[3,4] Michele Vallisneri,[28] Sarah J. Vigeland,[23] Caitlin A. Witt,[3,4]

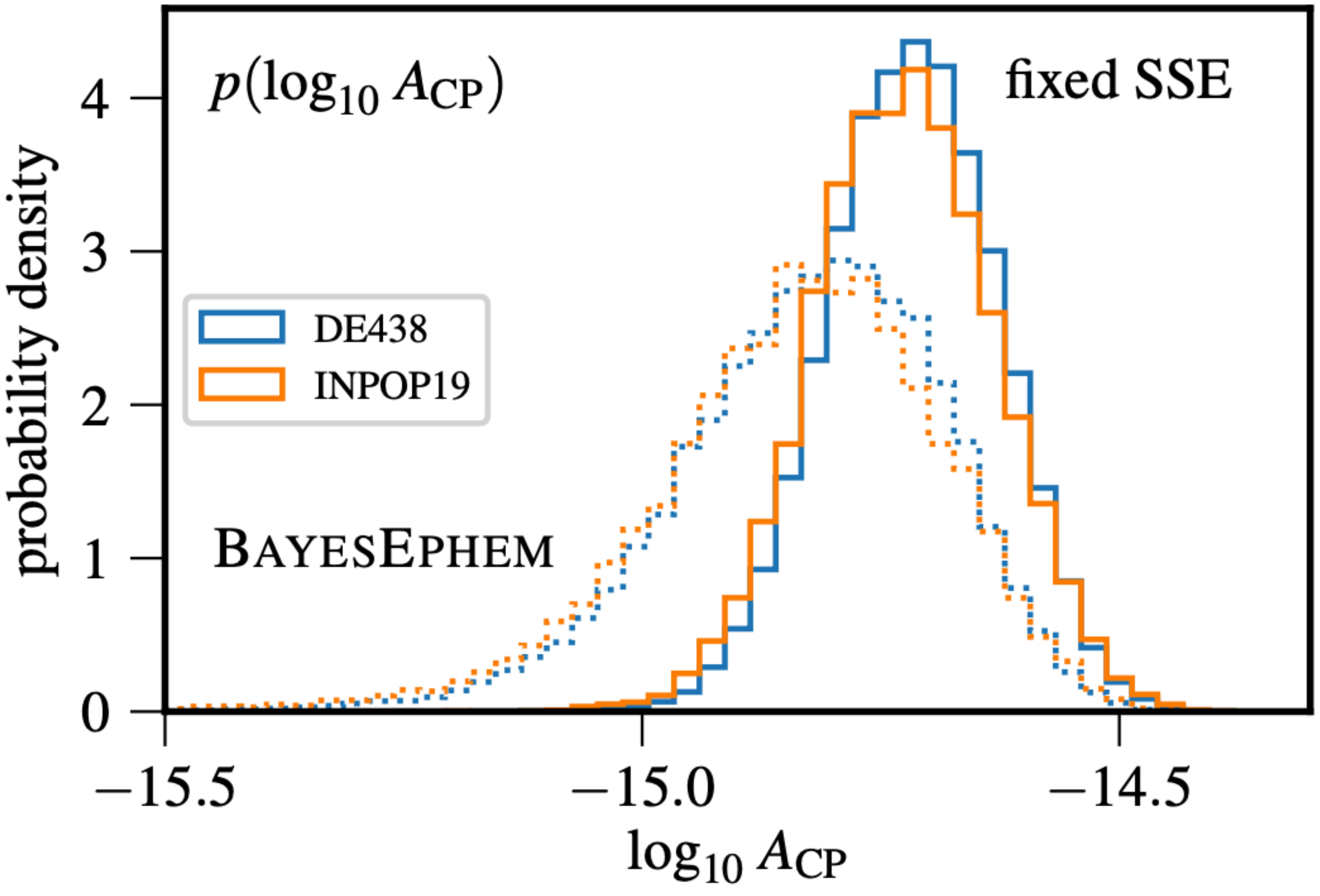
The NANOGrav Collaboration

NANOGrav 12.5yr Dataset Search (arXiv:2009.04496),
corresponding author: Joe Simon (JPL / CU-Boulder)

A Common-spectrum Process

NANOGrav 12.5yr Dataset Search (arXiv:2009.04496),
corresponding author: Joe Simon (JPL / CU-Boulder)

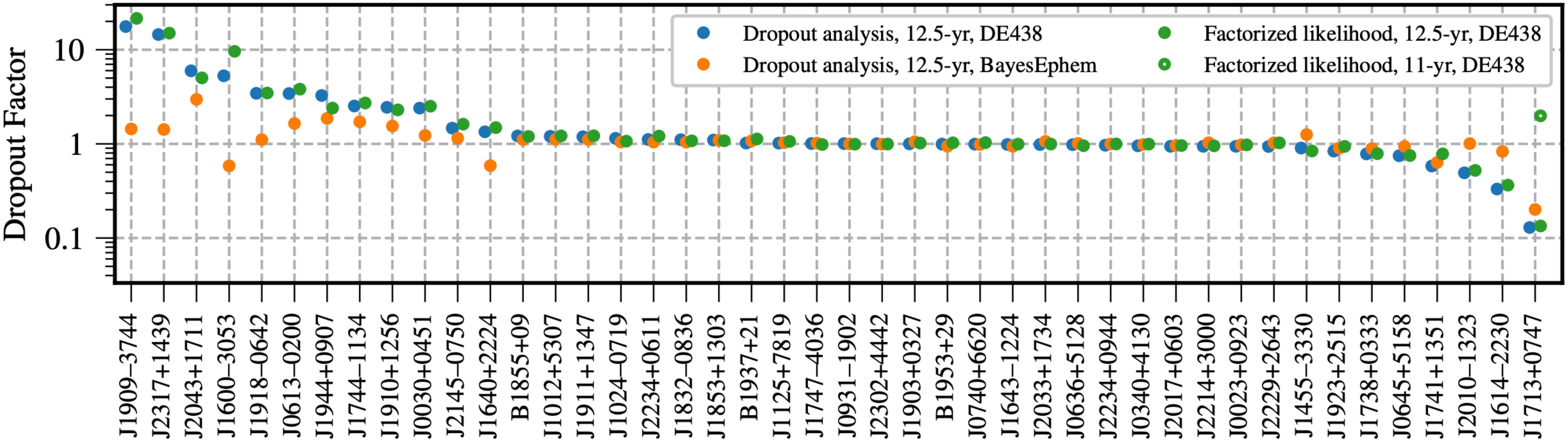




A steep-spectrum process in common across NANOGrav's 45-pulsar array with max baseline of 12.9 years

A Common-spectrum Process

NANOGrav 12.5yr Dataset Search (arXiv:2009.04496),
corresponding author: Joe Simon (JPL / CU-Boulder)



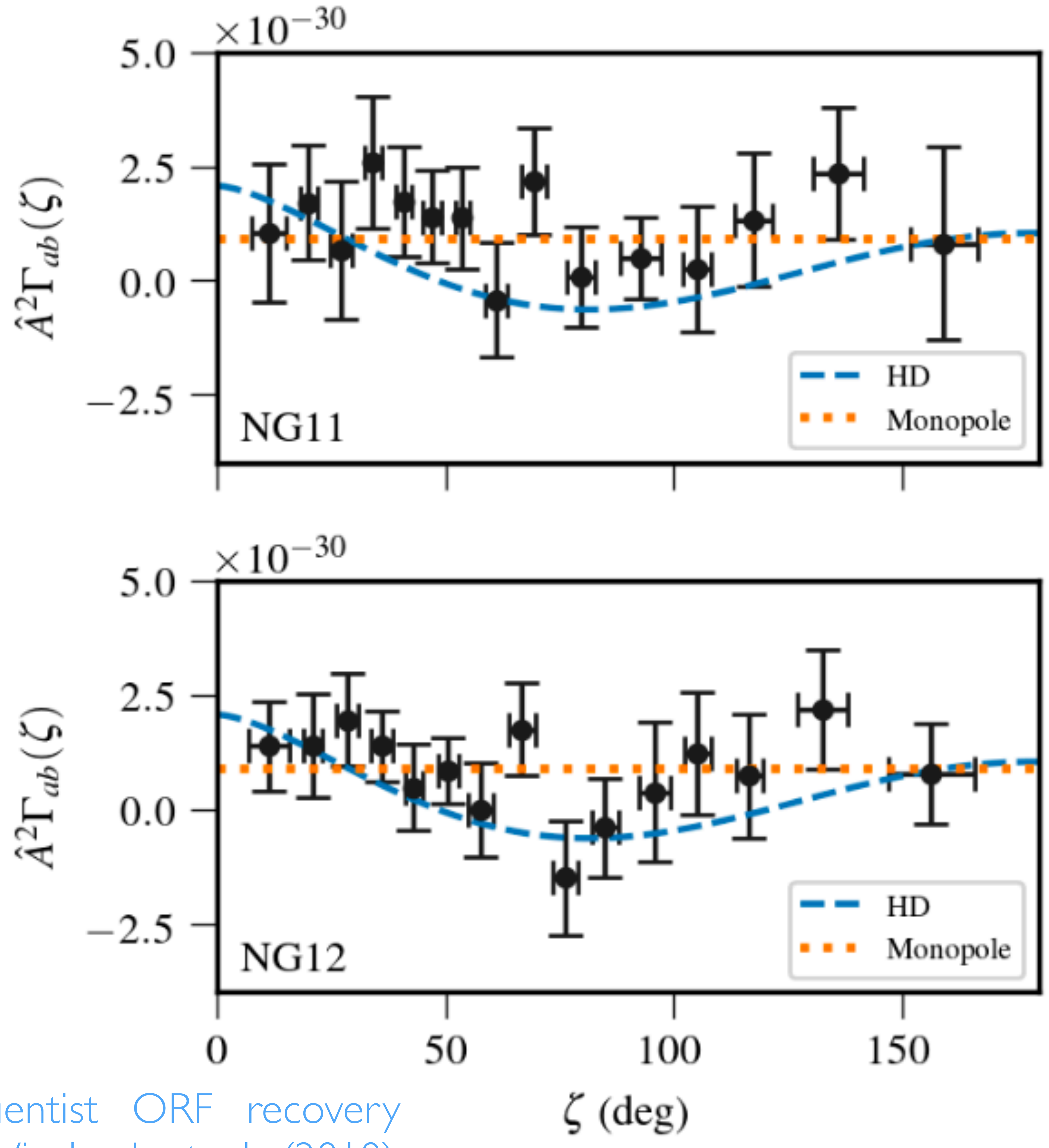
Dropout factor = cross-validation probability

i.e. how much does each pulsar support what is found by all other pulsars?

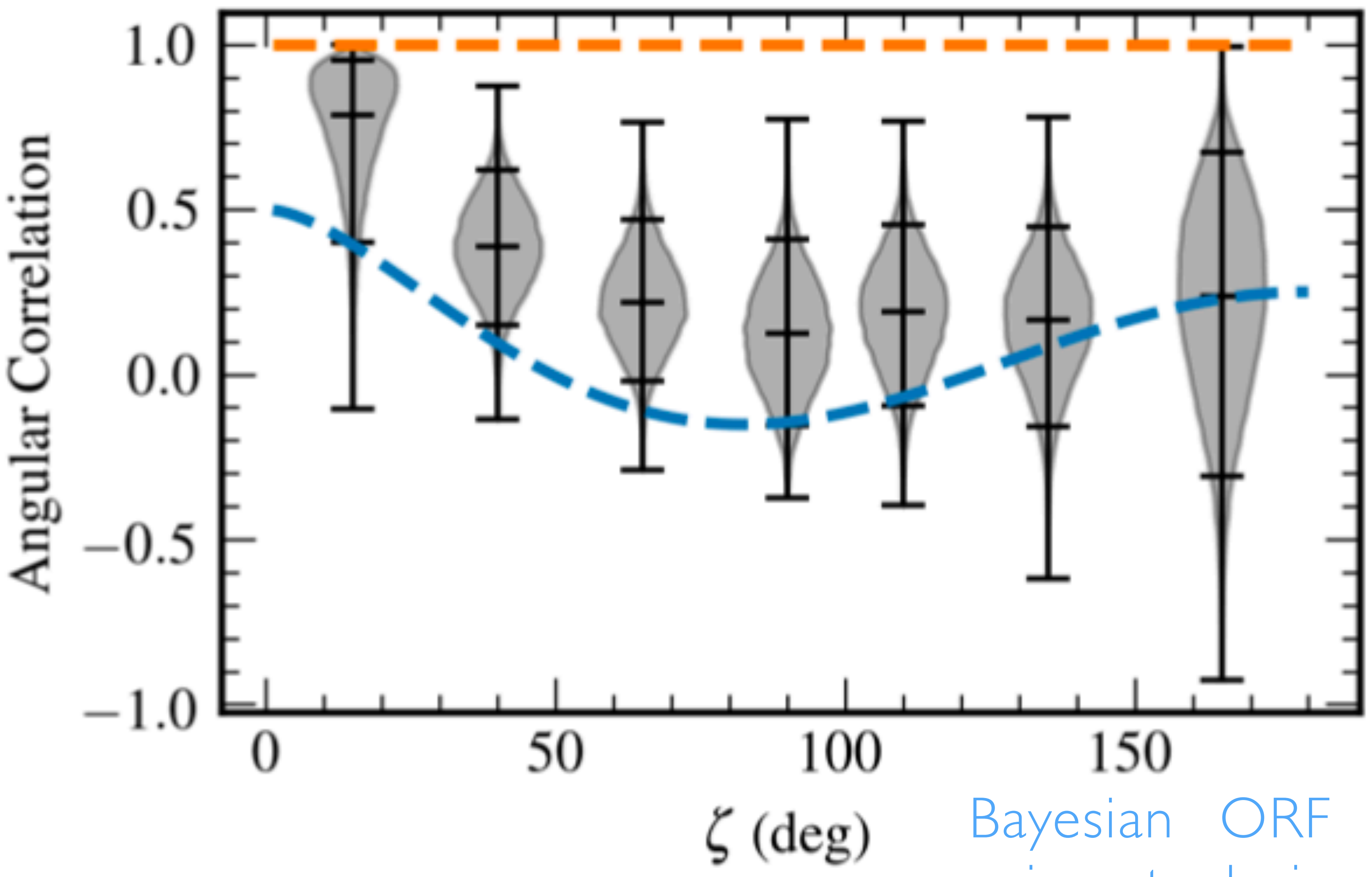
S. Vigeland, S. Taylor, M. Vallisneri

A Common-spectrum Process

NANOGrav 12.5yr Dataset Search (arXiv:2009.04496), *corresponding author: Joe Simon (JPL / CU-Boulder)*



Frequentist ORF recovery —> Vigeland et al. (2018), Chamberlin et al. (2015), etc.

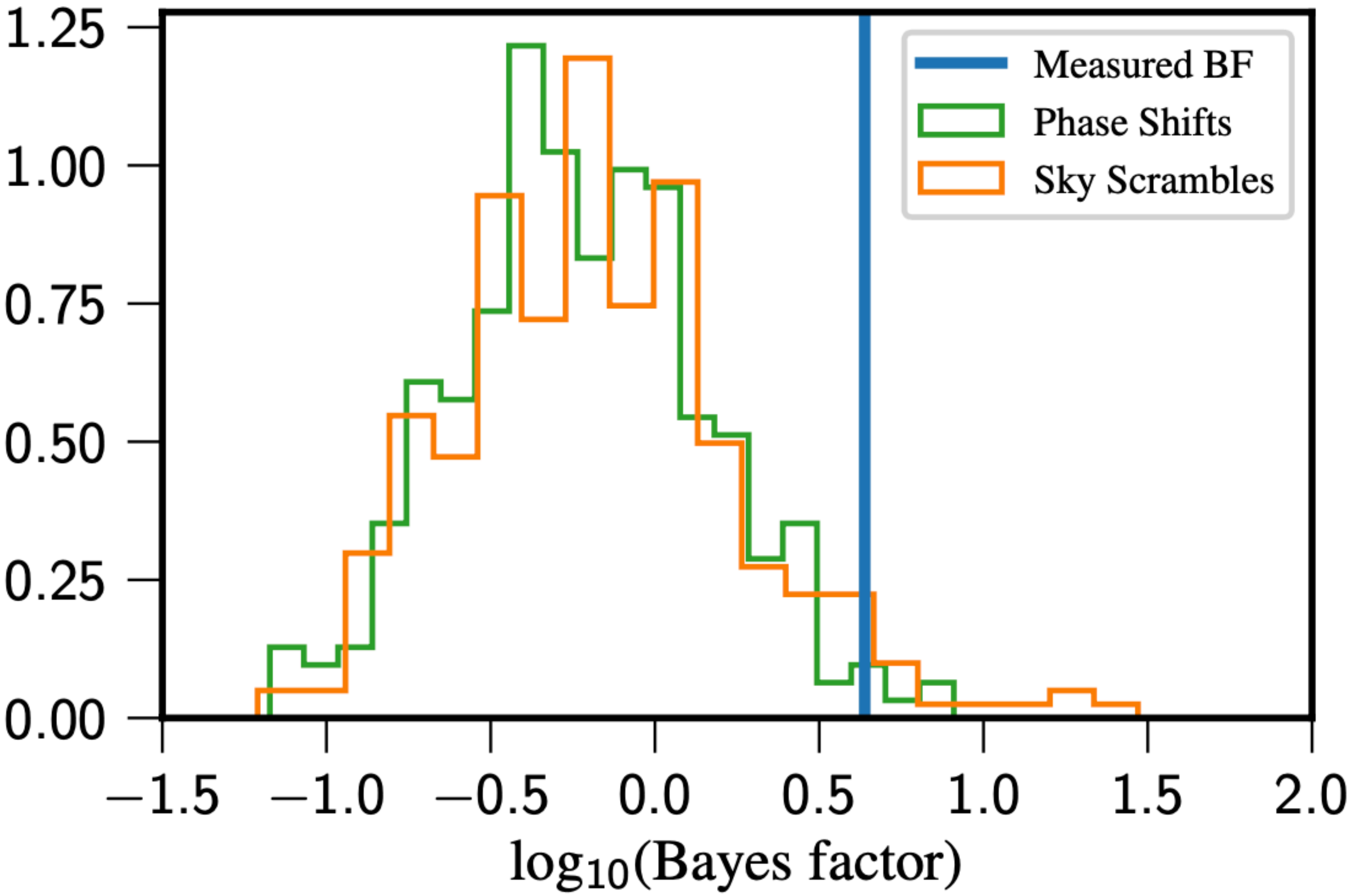


Bayesian ORF recovery using techniques from Taylor, Gair, Lentati (2013)

- Inter-pulsar correlations remain insignificant.
- Odds ratios for Hellings & Downs correlations **~2–4** depending on ephemeris modeling.

A Common-spectrum Process

NANOGrav 12.5yr Dataset Search (arXiv:2009.04496),
corresponding author: Joe Simon (JPL / CU-Boulder)



- Assess the significance of spatial correlations by constructing null distribution.
- LIGO-Virgo-KAGRA use time slides... we use **phase shifts (Taylor et al. 2017) and sky scrambles (Cornish & Sampson 2016; Taylor et al. 2017)**.
- **p ~ 5 - 10%**

$$\mathbf{C}_{\rm gwb} = \mathbf{F}\varphi_{\rm gwb}\mathbf{F}^{\mathbf{T}}$$

Phase Shifting

Sky Scrambles

The Road To & Beyond Detection

…Or "what to expect when you're expecting to detect a signal".

Dr. Nihan Pol

Simulate up to 20 years of PTA data, forecasting from the 45 pulsars in the NG 12.5yr data

$\hat{\rho} =$ total S/N *(from full log-likelihood ratio)*

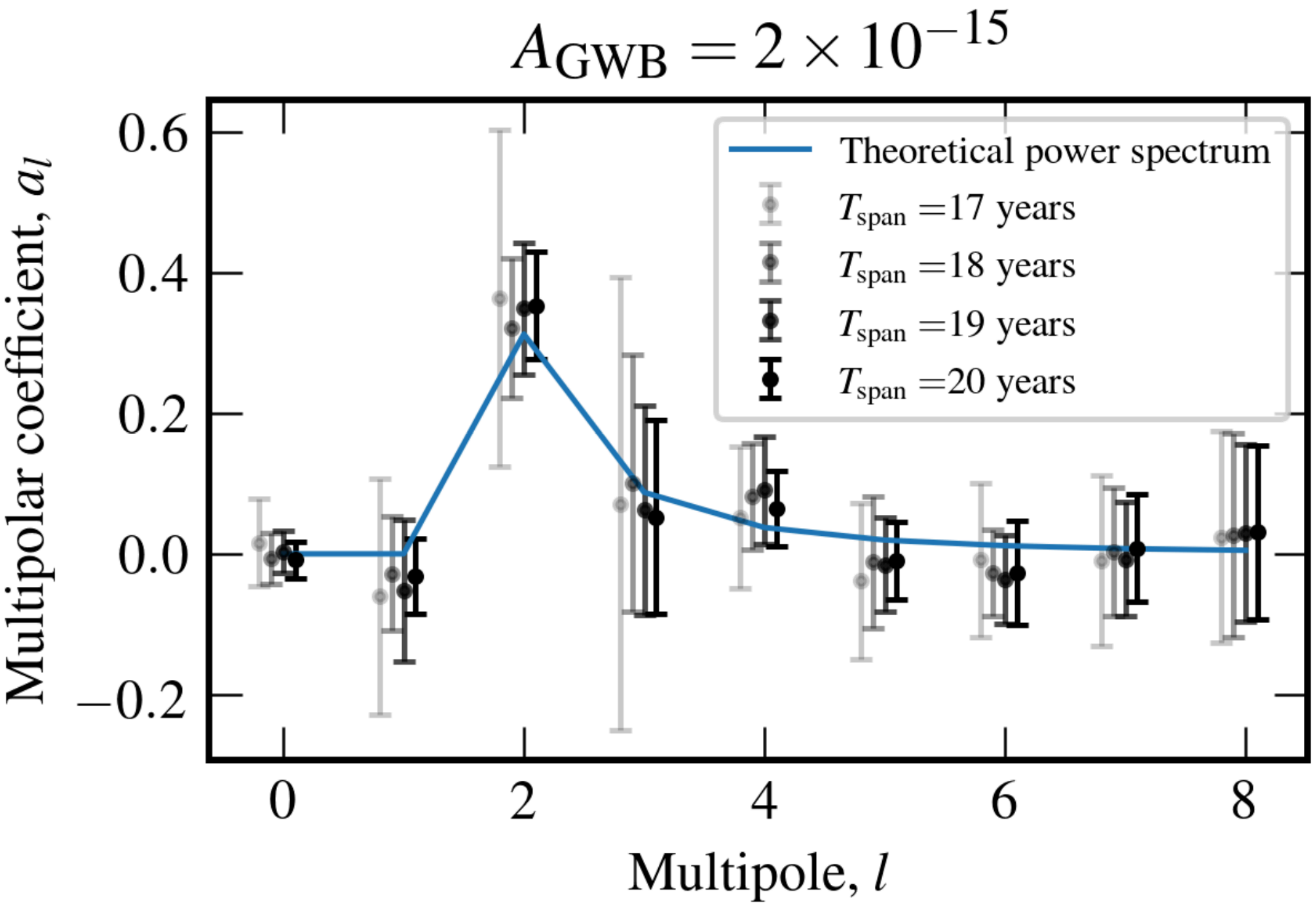
$\rho_{\mathrm{HD}} =$ cross-correlation S/N

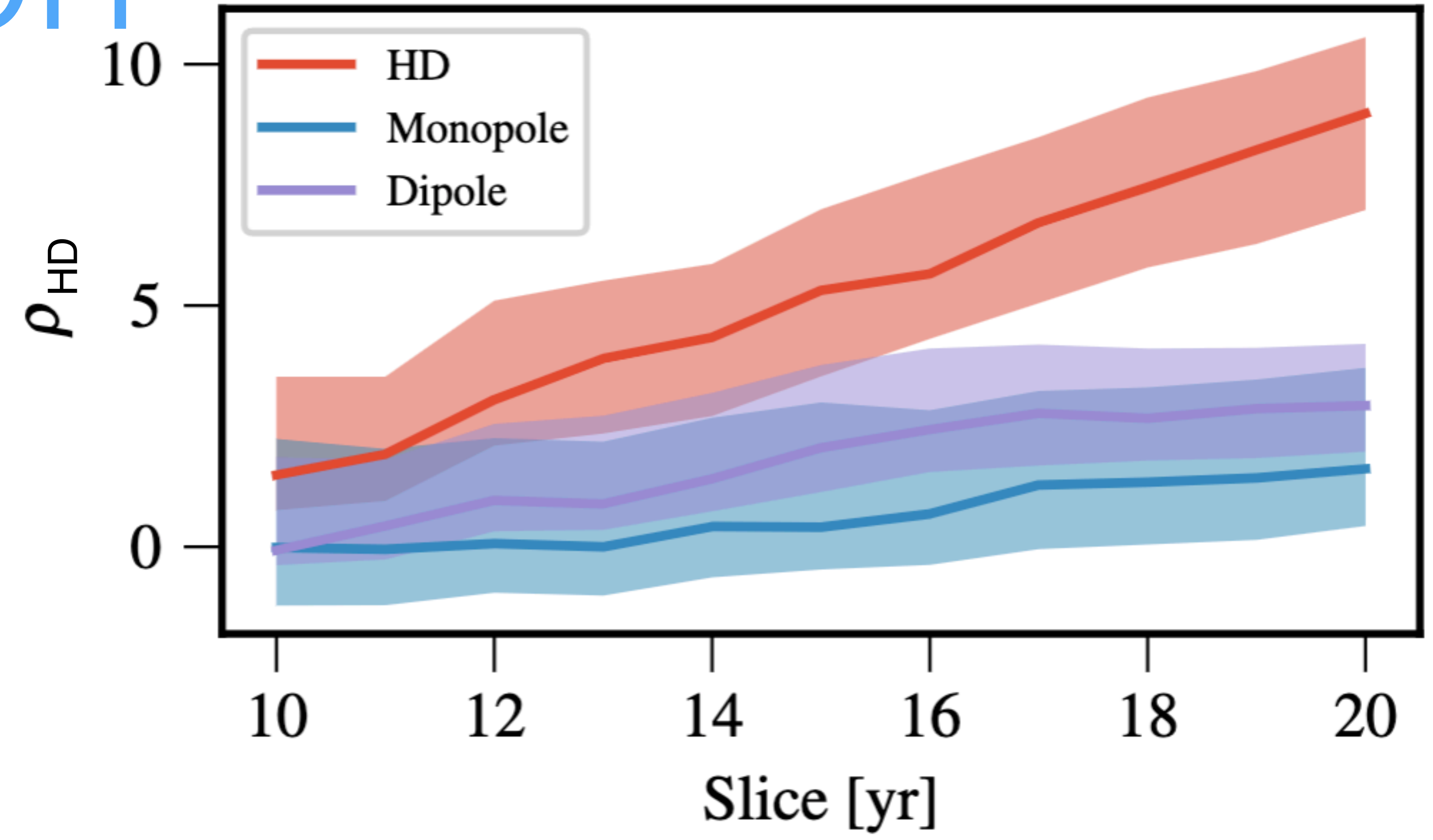
Full team: Nihan Pol, Stephen Taylor, Luke Kelley, Joe Simon, Sarah Vigeland, Siyuan Chen

The Road To & Beyond Detection

…Or "*what to expect when you're expecting to detect a signal*".

Probe the multipolar structure of the inter-pulsar correlations





$$\Gamma_{ab} = \sum_{l=0}^{\infty} a_l P_l(\cos\theta_{ab})$$

Isotropic GWB:
$$\begin{cases} a_l = \dfrac{3}{4} N_l^2 (2l+1) \\ N_l = \sqrt{\dfrac{2(l-2)!}{(l+2)!}} \end{cases}$$

Gair, Romano, Taylor, Mingarelli (2014)

The Road To & Beyond Detection

"*Astrophysics Milestones For Pulsar Timing Array Gravitational Wave Detection*", Pol, Taylor et al., arXiv:2010.11950

…Or "*what to expect when you're expecting to detect a signal*".

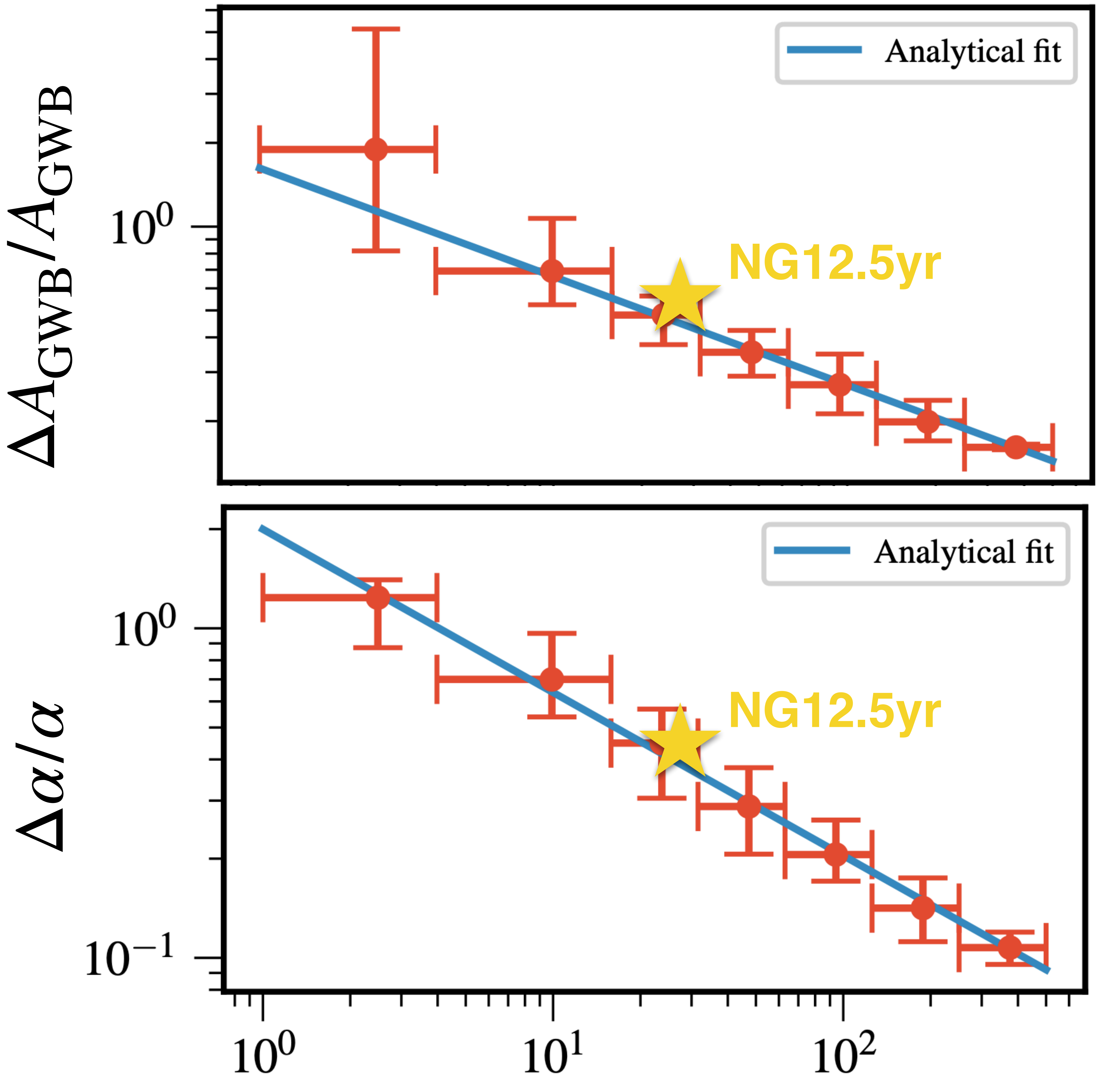
$$h_c(f) = A_{\mathrm{GWB}} \left(\frac{f}{1\ \mathrm{yr}^{-1}} \right)^{\alpha}$$

parameter uncertainty scaling laws

$$\Delta A_{\mathrm{GWB}}/A_{\mathrm{GWB}} = 44 \times \left(\frac{\hat{\rho}}{25} \right)^{-2/5} \%$$

$$\Delta \alpha/\alpha = 40 \times \left(\frac{\hat{\rho}}{25} \right)^{-1/2} \%$$

Can relate $\hat{\rho}$ to ρ_{HD} and factors like T, σ_{RMS}, N_{pulsar}, etc.



Summary

- **Pulsar Timing Arrays** are sensitive to nanohertz gravitational waves.
- We use rank-reduced time-domain modeling of stochastic processes across dozens of pulsars and over decades of observations.
- If the NANOGrav result hints at a GWB, then **detection and characterization could be within a few years** (expedited by fusing datasets together in the IPTA).
- The road beyond detection will inform demographics and final-parsec binary dynamical interactions of supermassive binary black holes.