

Searches for stochastic gravitational-wave backgrounds

Lecture 2
Les Houches Summer School
20 July 2018

Joe Romano
University of Texas Rio Grande Valley
Texas Tech University



Plan for lectures

Yesterday: Overview / Basics

1. Motivation / context
2. Different types of stochastic backgrounds
3. Characterizing a stochastic GW background
4. Correlation methods
5. Some simple examples

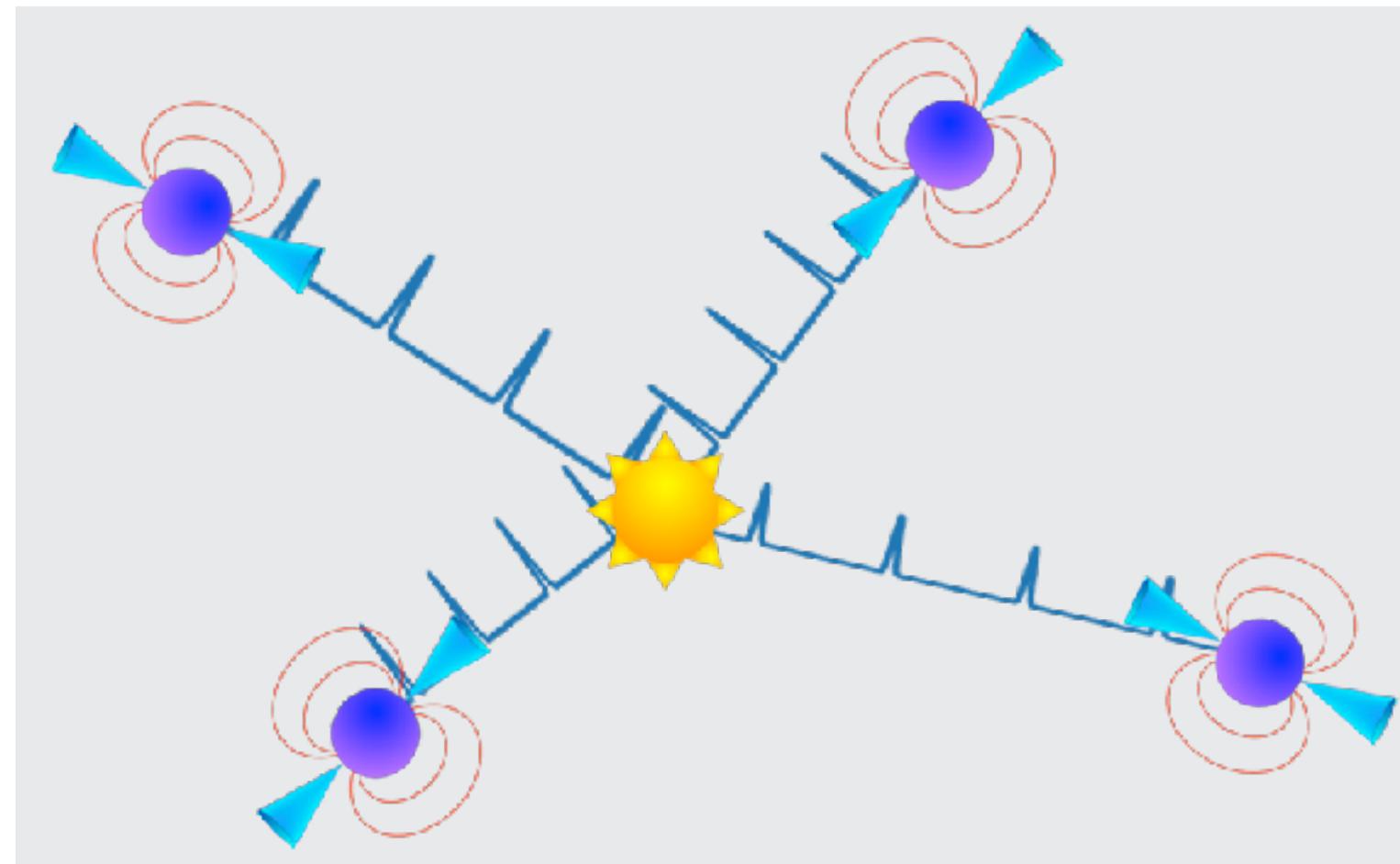
Today: Details / Example

1. Non-trivial detector response
2. Non-trivial correlated response
3. What to do in the absence of correlations (e.g., for LISA)?
4. Frequentist and Bayesian methods
5. Example: searching for the background from BBH mergers

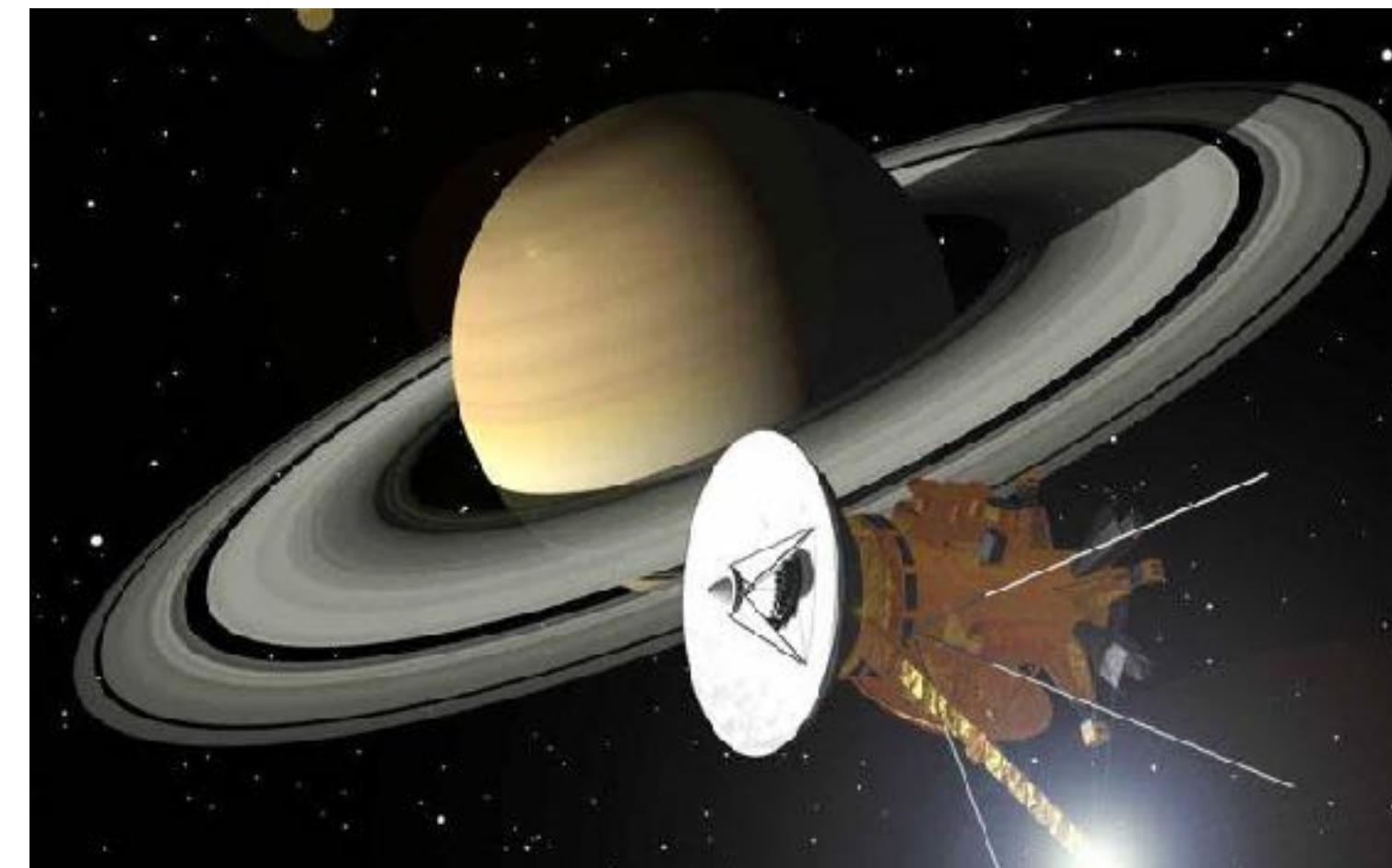
1 . Non-trivial detector response

Beam detectors

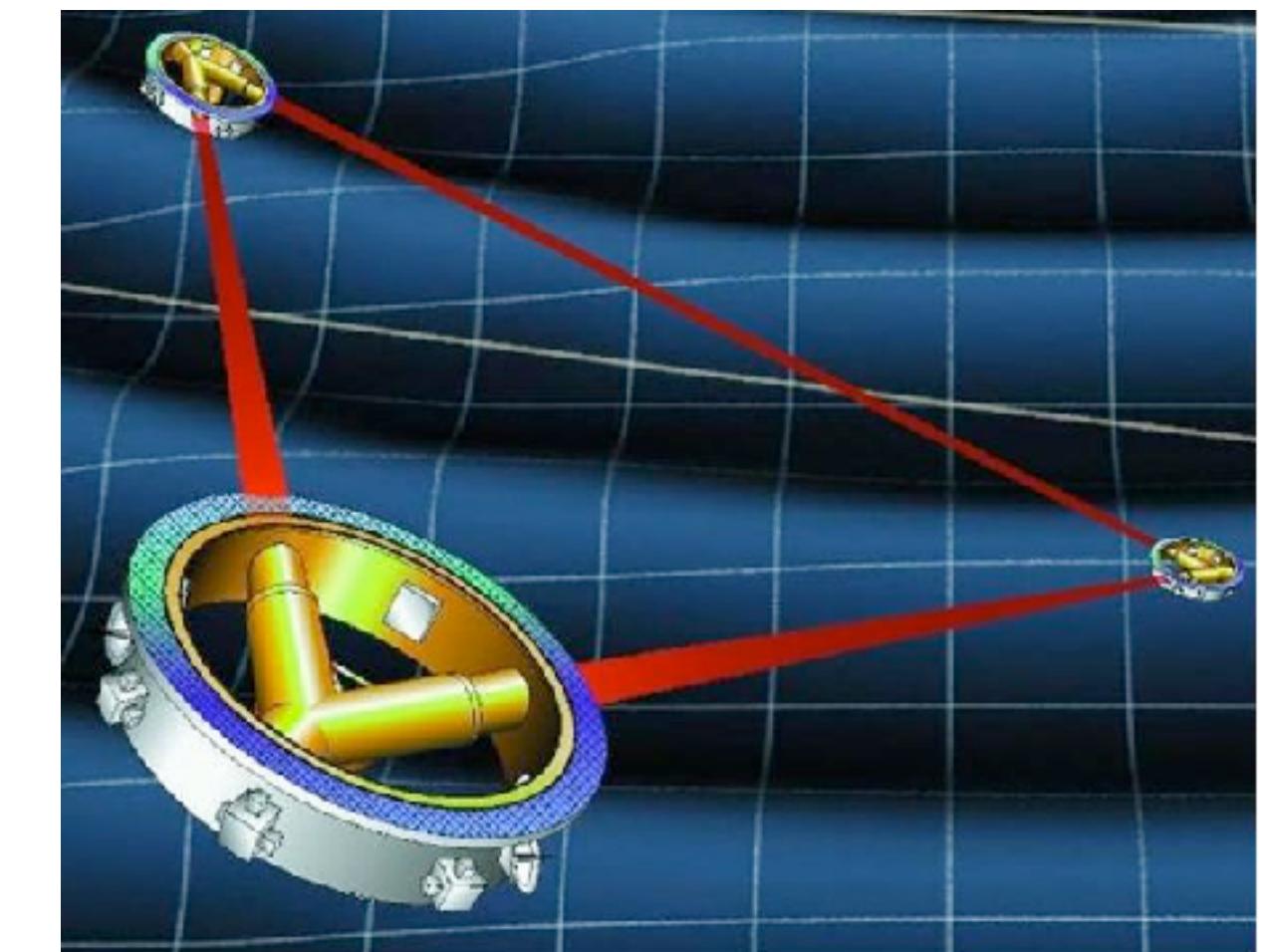
pulsar timing

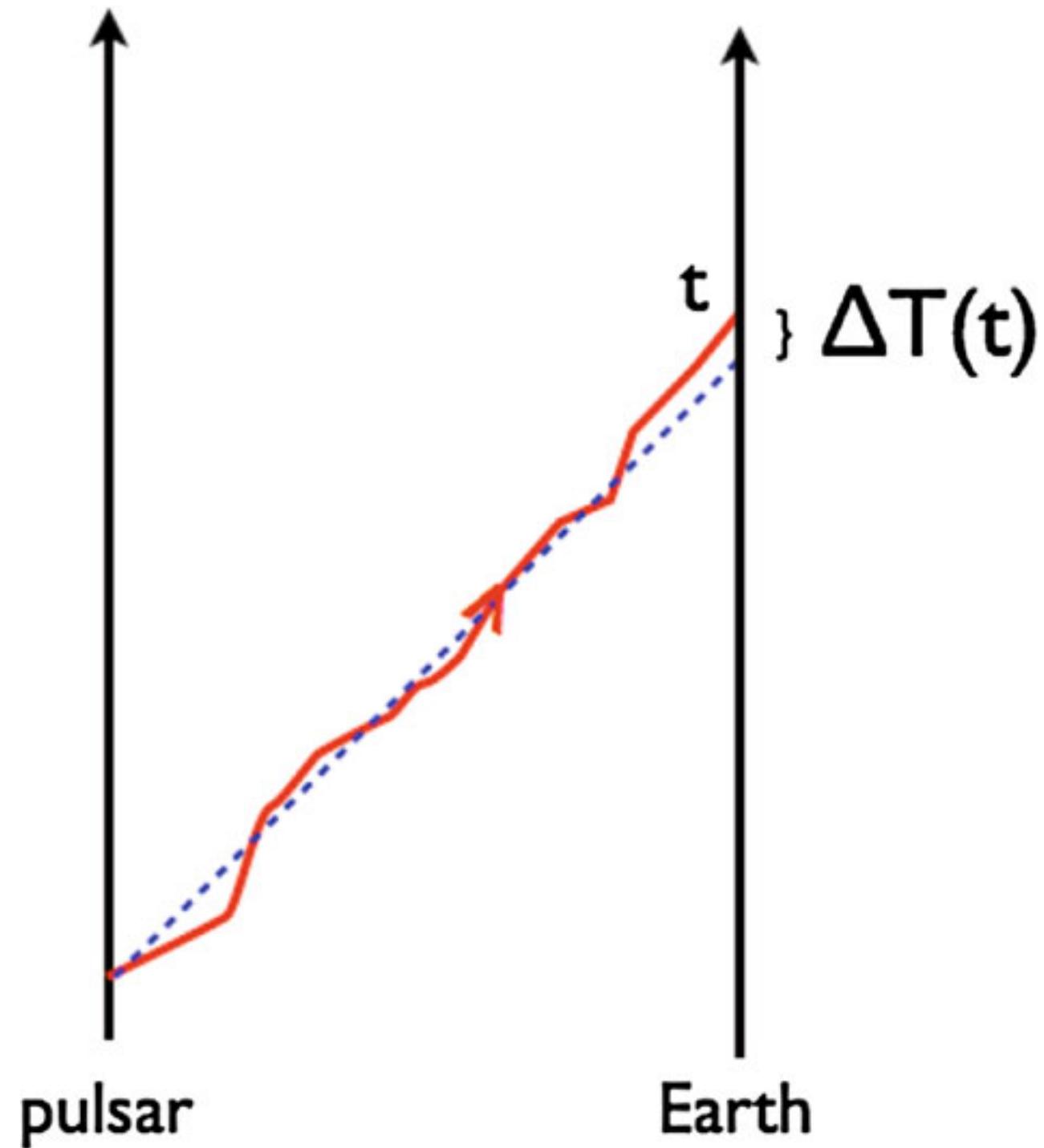
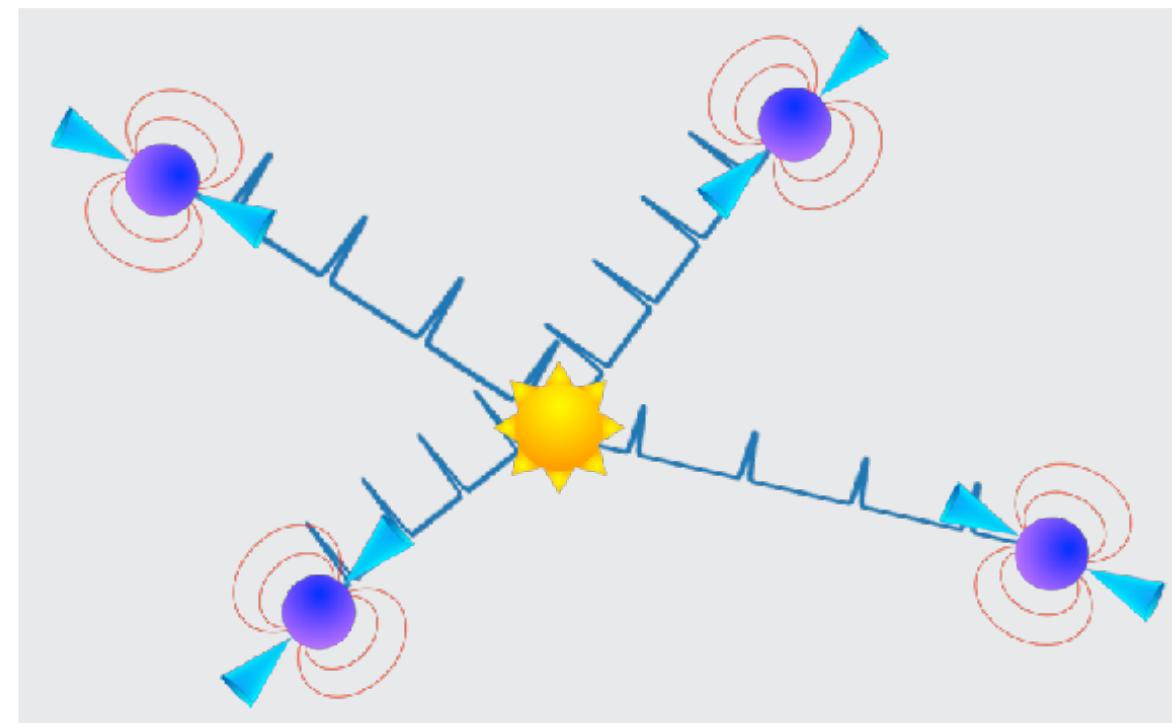


spacecraft Doppler tracking

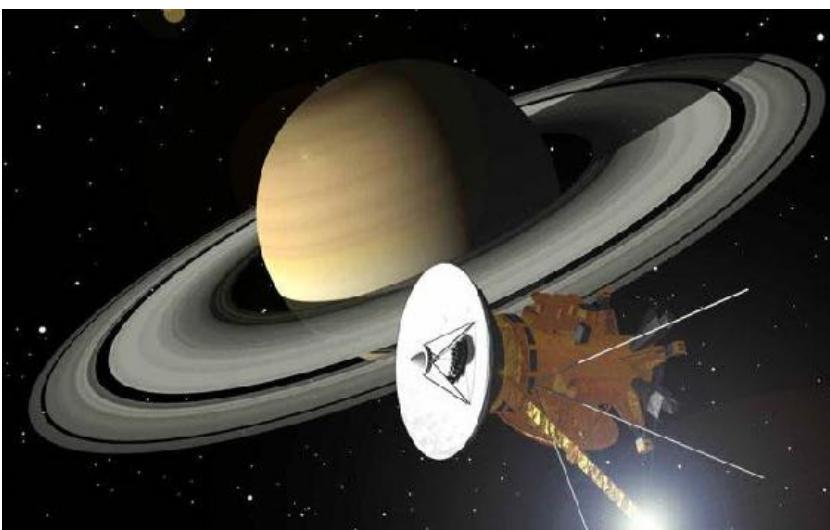


laser interferometers





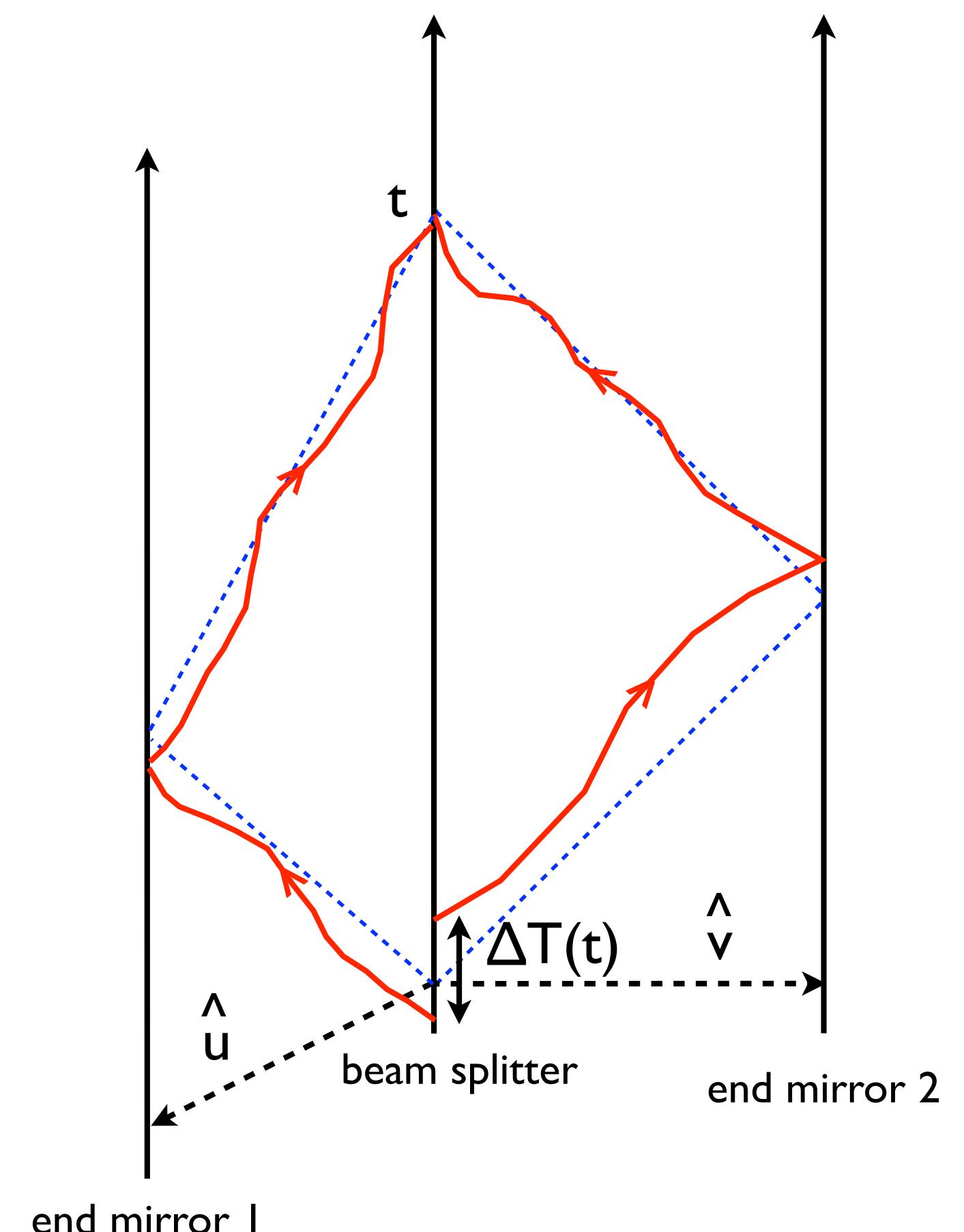
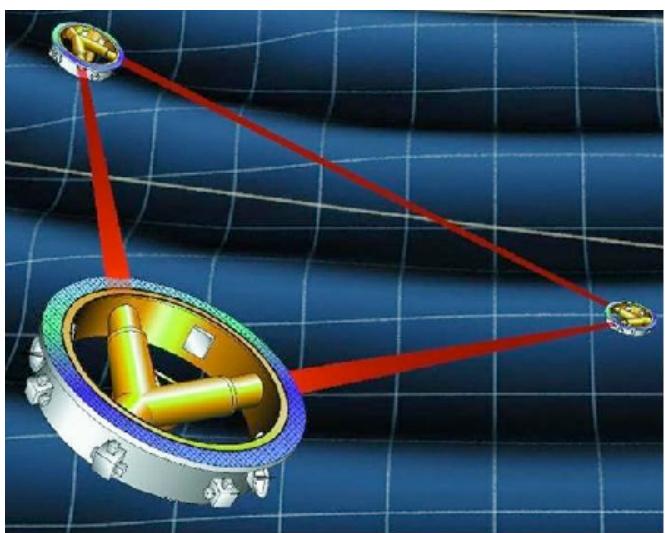
(1-arm, 1-way)



Earth

spacecraft

(1-arm, 2-way)



(2-arm, 2-way)

Different types of response

timing:

$$h(t) \equiv \Delta T(t) \quad (\text{pulsar timing})$$

Doppler frequency:

$$h(t) \equiv \frac{\Delta\nu(t)}{\nu_0} = \frac{d\Delta T(t)}{dt} \quad (\text{pulsar timing, spacecraft Doppler tracking})$$

strain:

$$h(t) \equiv \frac{\Delta L(t)}{L} = \frac{\Delta T(t)}{T} \quad (\text{LIGO, Virgo, ...})$$

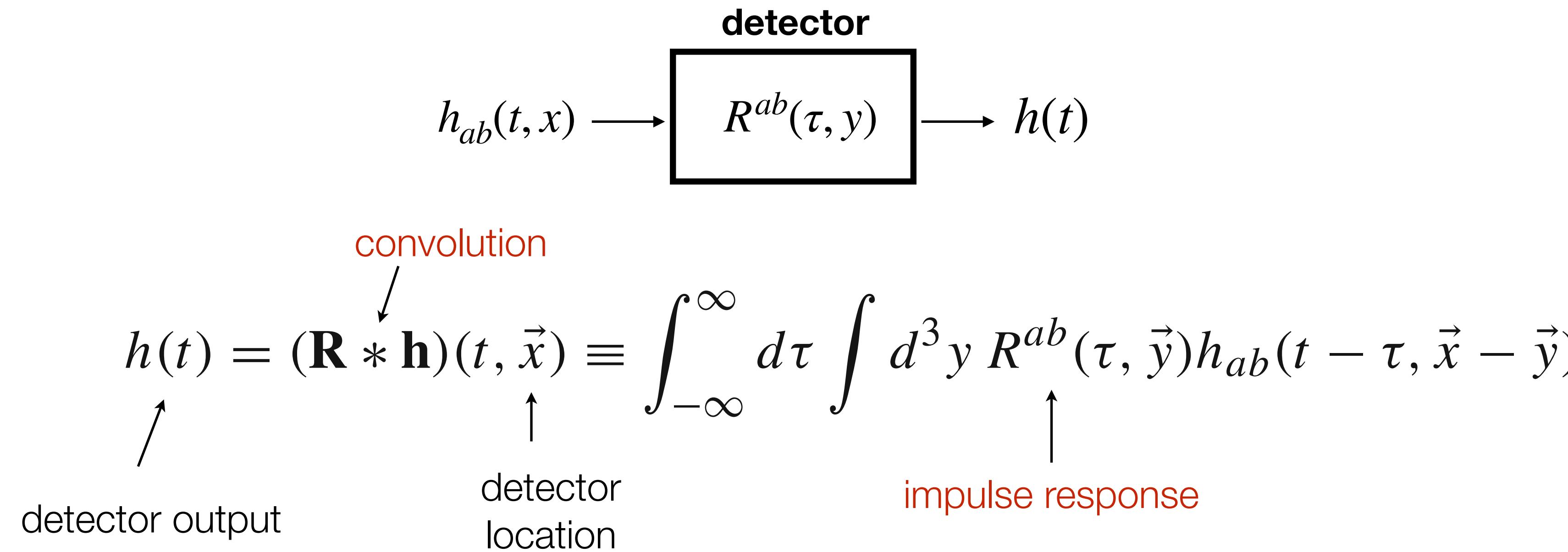
phase:

$$h(t) \equiv \Delta\Phi(t) = 2\pi\nu_0 \Delta T(t) \quad (\text{LISA})$$

All of these responses derivable from the change in light travel time $\Delta T(t)$

Detector response

Detector = **linear system** which converts GW metric perturbations to detector output



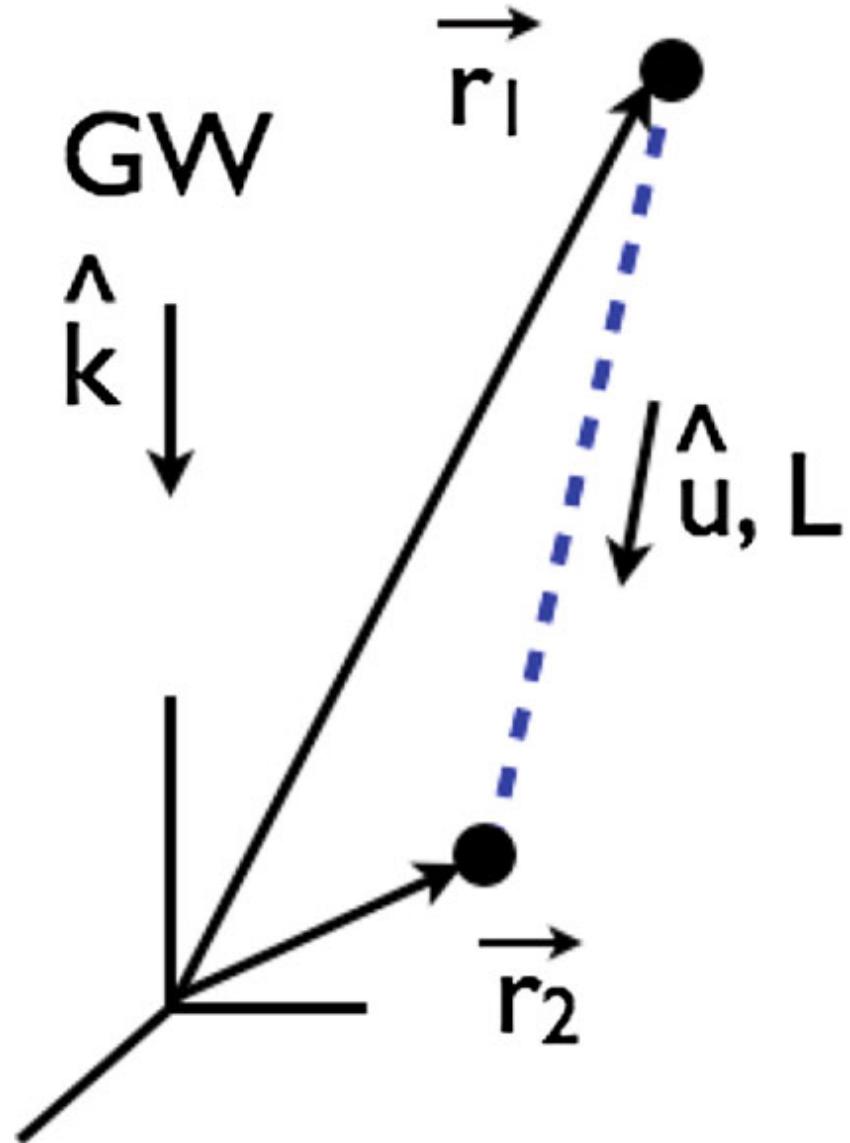
$$\implies \tilde{h}(f) = \int d^2\Omega_{\hat{n}} \sum_A \boxed{R^A(f, \hat{n})} h_A(f, \hat{n})$$

detector response for a plane-wave
with frequency f, direction k, polarization A

$$R^A(f, \hat{n}) \equiv R^{ab}(f, \hat{n}) e_{ab}^A(\hat{n})$$

$$R^{ab}(f, \hat{n}) \equiv e^{i2\pi f \hat{n} \cdot \vec{x}/c} \int_{-\infty}^{\infty} d\tau \int d^3y R^{ab}(\tau, \vec{y}) e^{-i2\pi f(\tau + \hat{n} \cdot \vec{y}/c)}$$

Example: Single-link response function (e.g., pulsar timing)



$$\Delta T(t) = \frac{1}{2c} u^a u^b \int_0^L ds h_{ab}(t(s), \vec{x}(s))$$

where

$$t(s) = (t - L/c) + s/c, \quad \vec{x}(s) = \vec{r}_1 + s\vec{v}$$

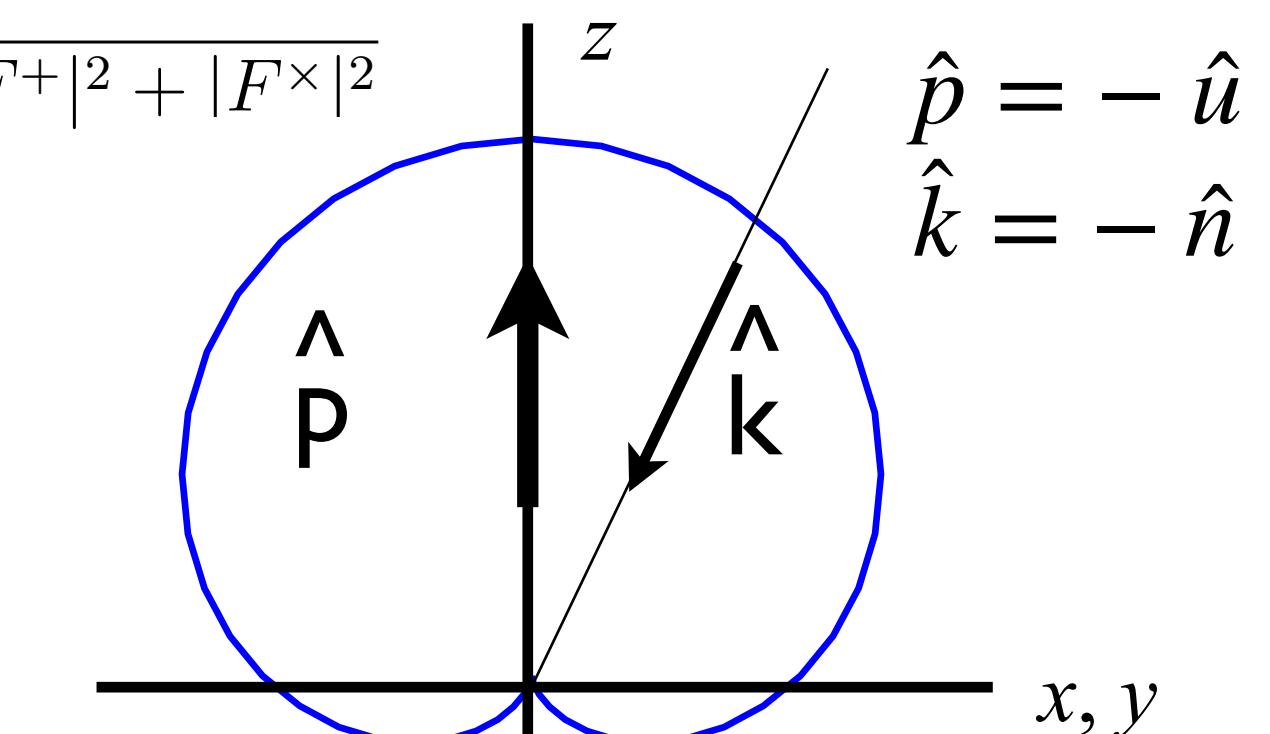
$$R^A(f, \hat{n}) = \frac{1}{i2\pi f} \left[\frac{1}{2} u^a u^b e_{ab}^A(\hat{n}) \frac{1}{1 + \hat{n} \cdot \hat{u}} \right] \left[1 - e^{-\frac{i2\pi f L}{c}(1 + \hat{n} \cdot \hat{u})} \right] e^{i2\pi f \hat{n} \cdot \vec{r}_2 / c}$$

= 1 for Doppler
freq measurement

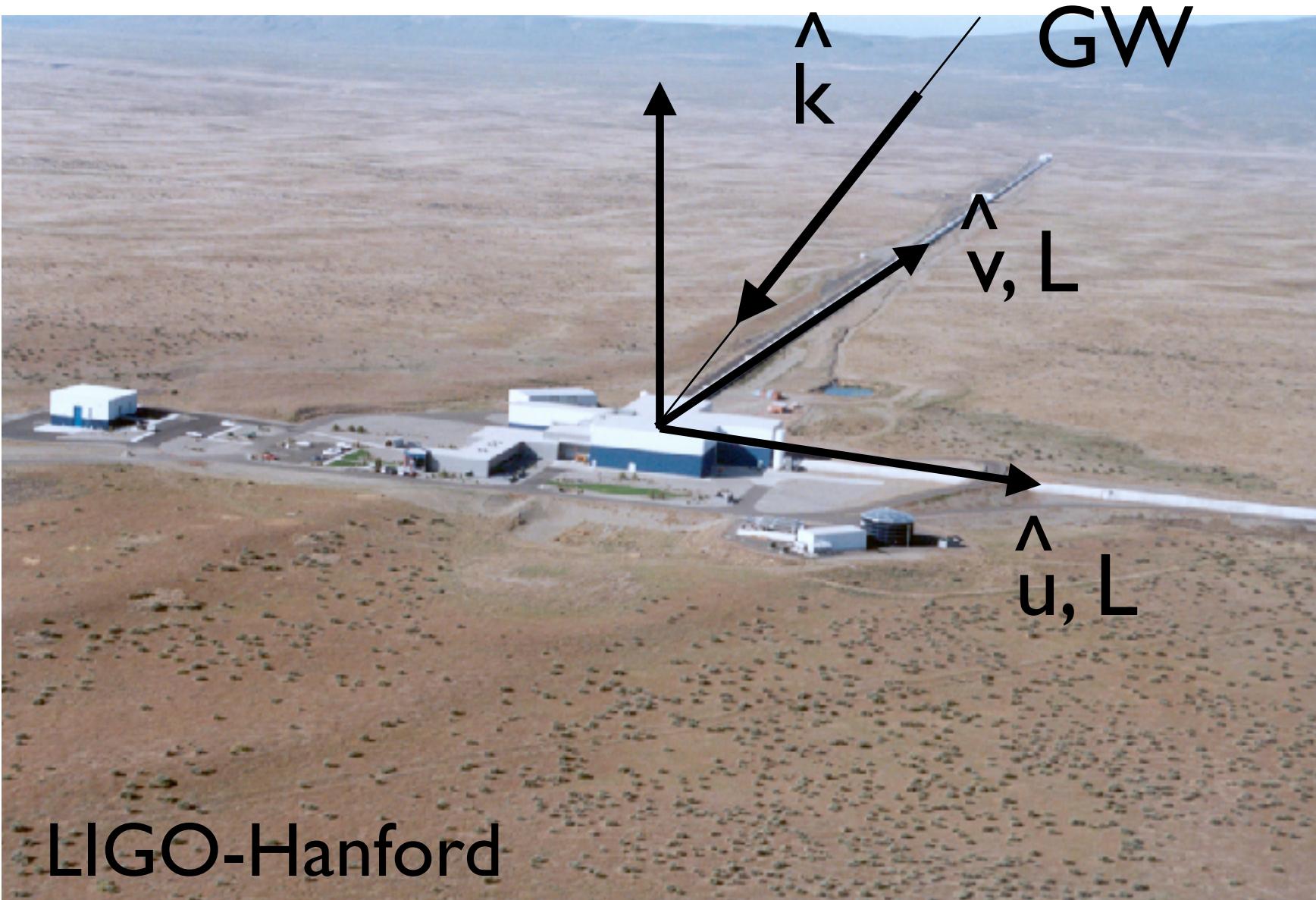
$F^A(\hat{n})$

(typically ignored
for pulsar timing)

Exercise 6: Derive this expression for the response function



Example: LIGO response (equal-arm, short-antenna limit)

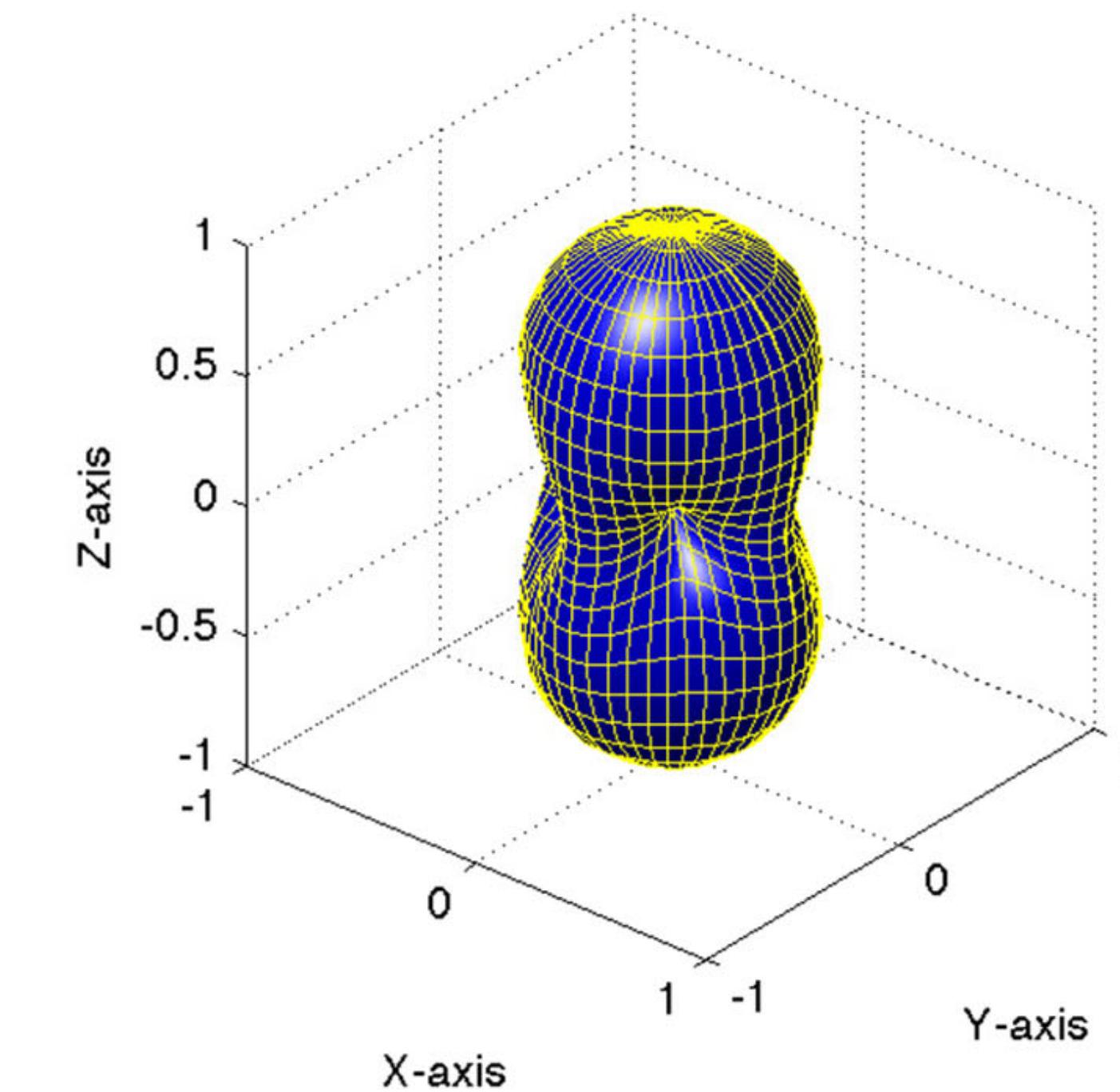


$$h(t) = \frac{1}{2} \left(\frac{\Delta T_{\vec{u}, \text{roundtrip}}(t)}{T} - \frac{\Delta T_{\vec{v}, \text{roundtrip}}(t)}{T} \right)$$

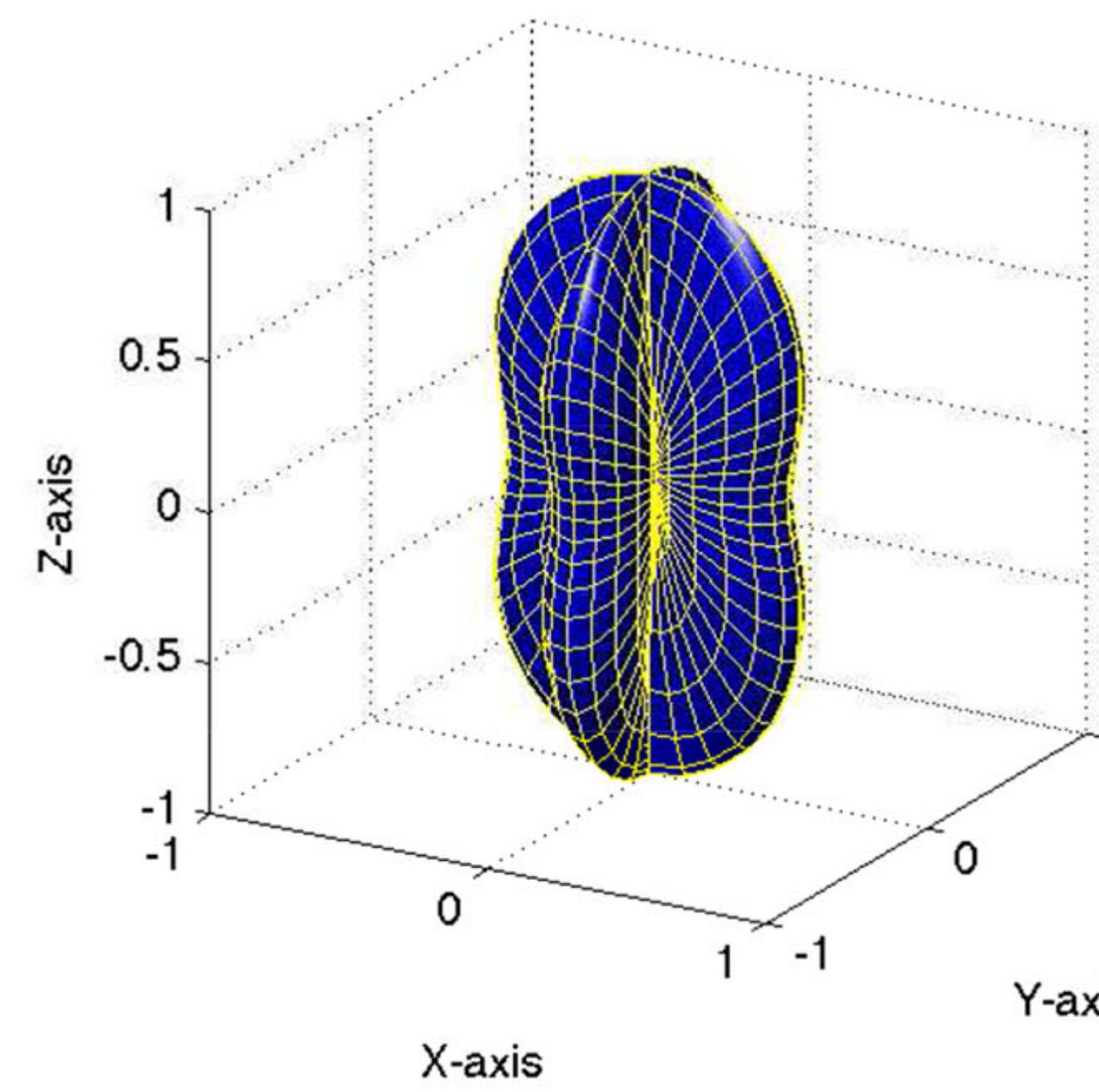
$$R^A(f, \hat{n}) \simeq \boxed{\frac{1}{2} (u^a u^b - v^a v^b)} e_{ab}^A(\hat{n})$$

detector tensor

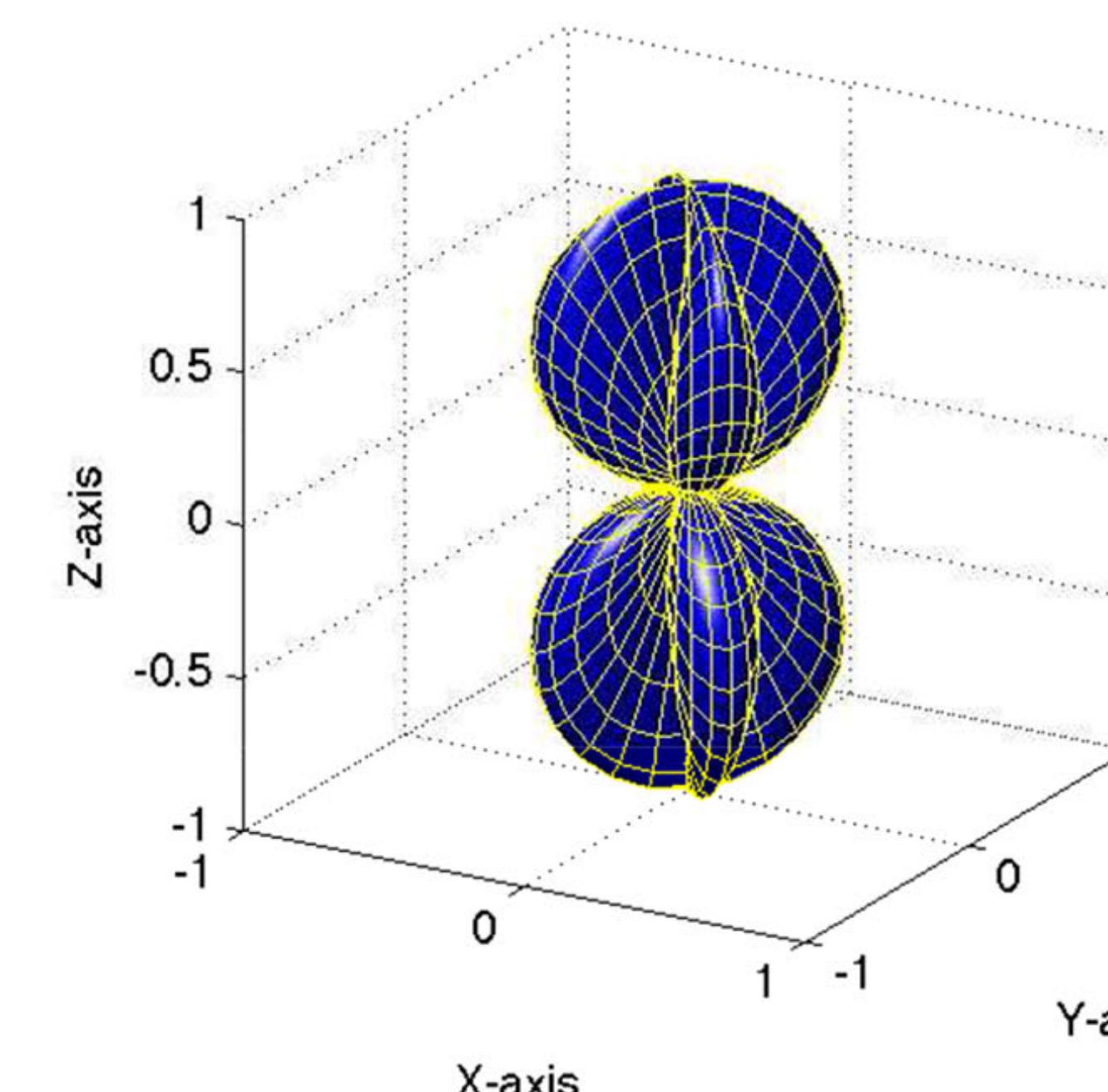
$$\sqrt{|R^+|^2 + |R^\times|^2}$$



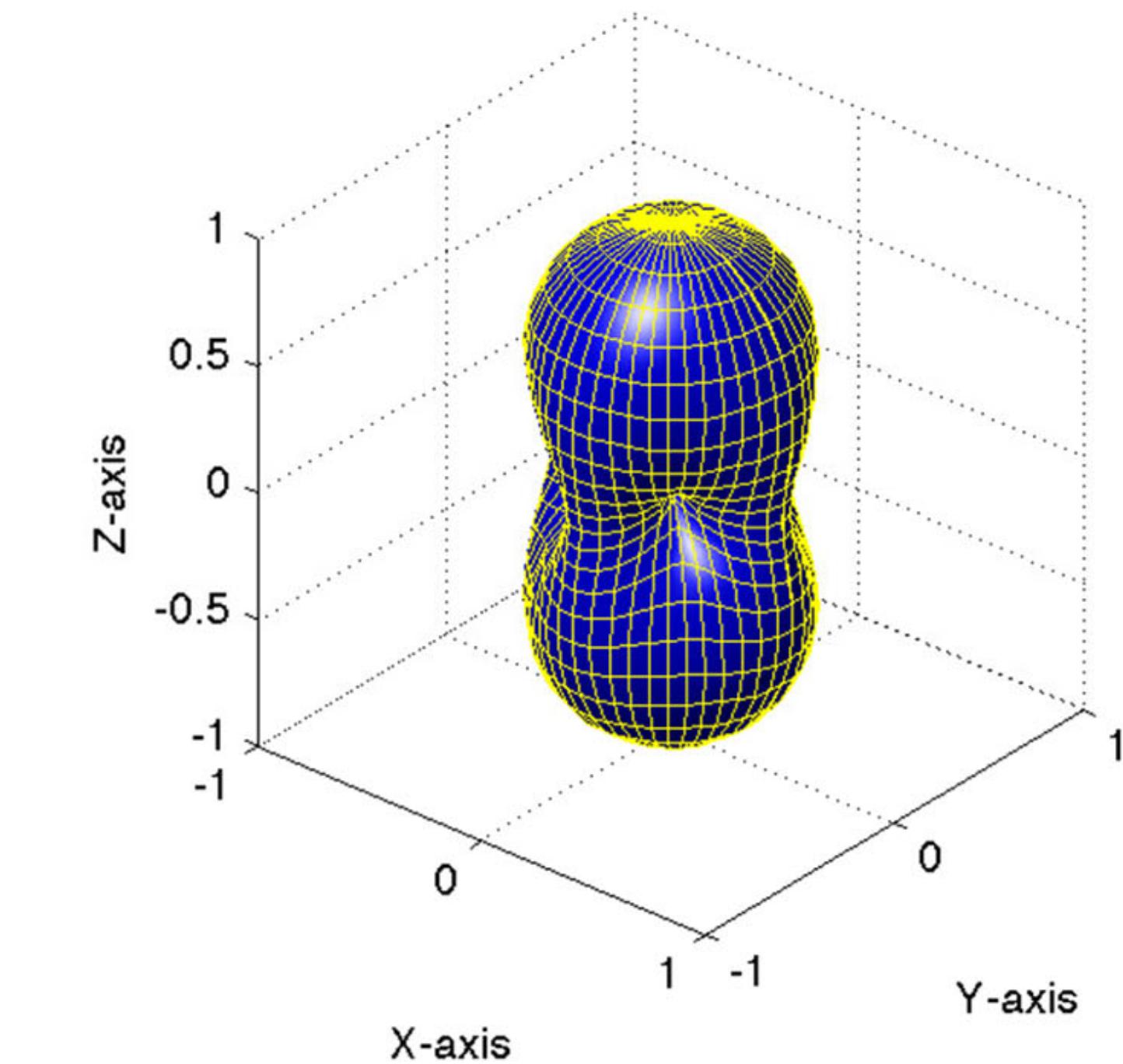
Beam pattern functions



$$|R^+(f, \hat{n})|$$



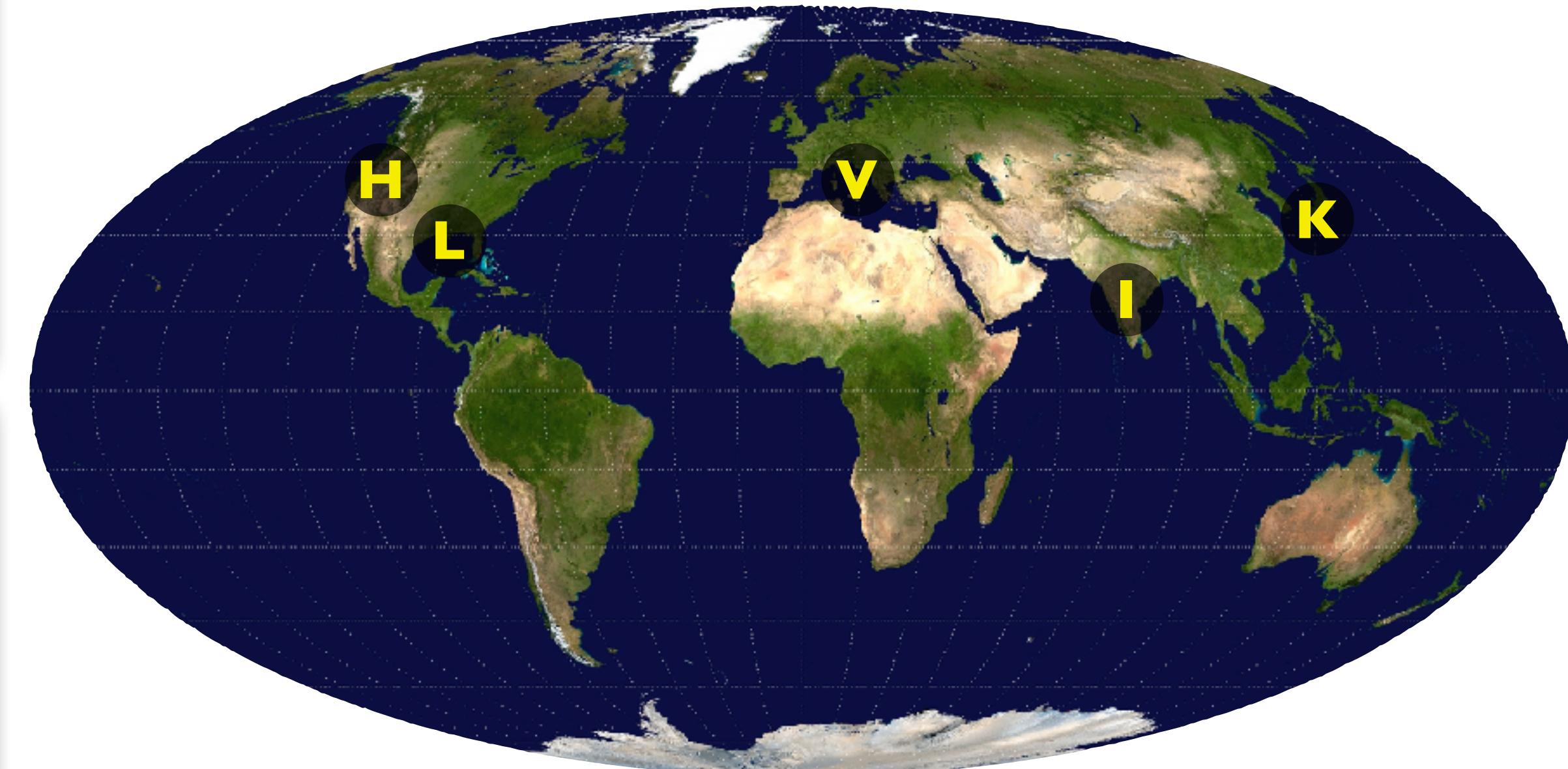
$$|R^\times(f, \hat{n})|$$



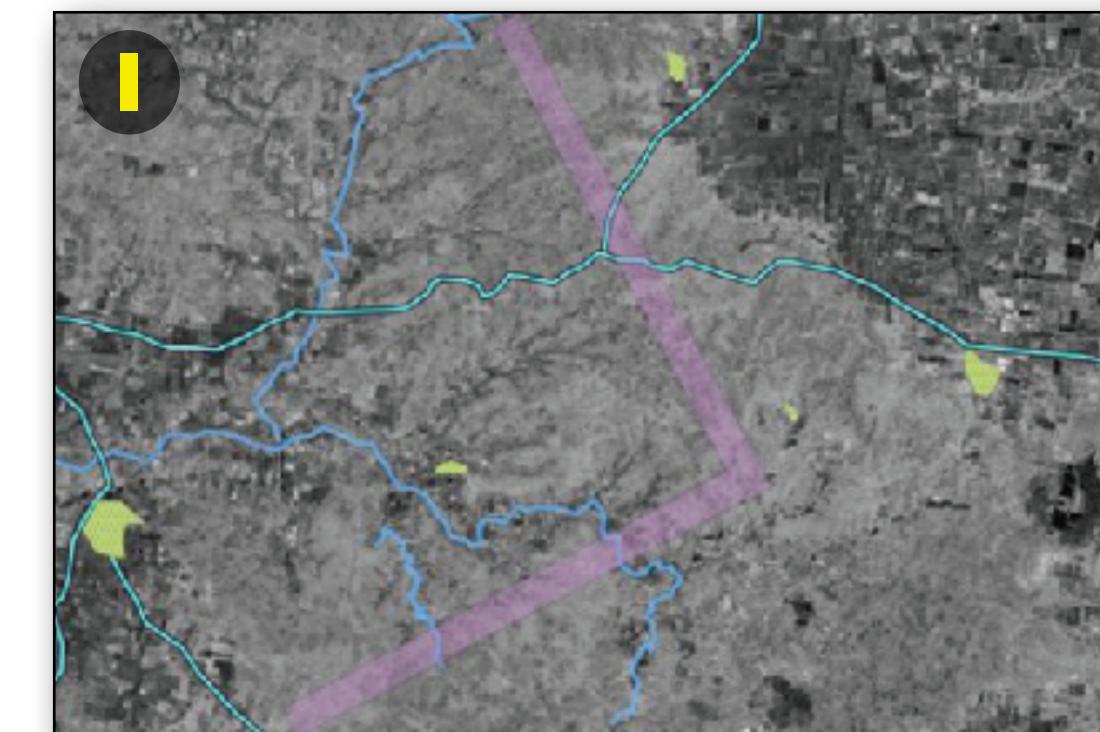
$$\left(|R^+(f, \hat{n})|^2 + |R^\times(f, \hat{n})|^2 \right)^{1/2}$$

$(f < \text{a few kHz})$

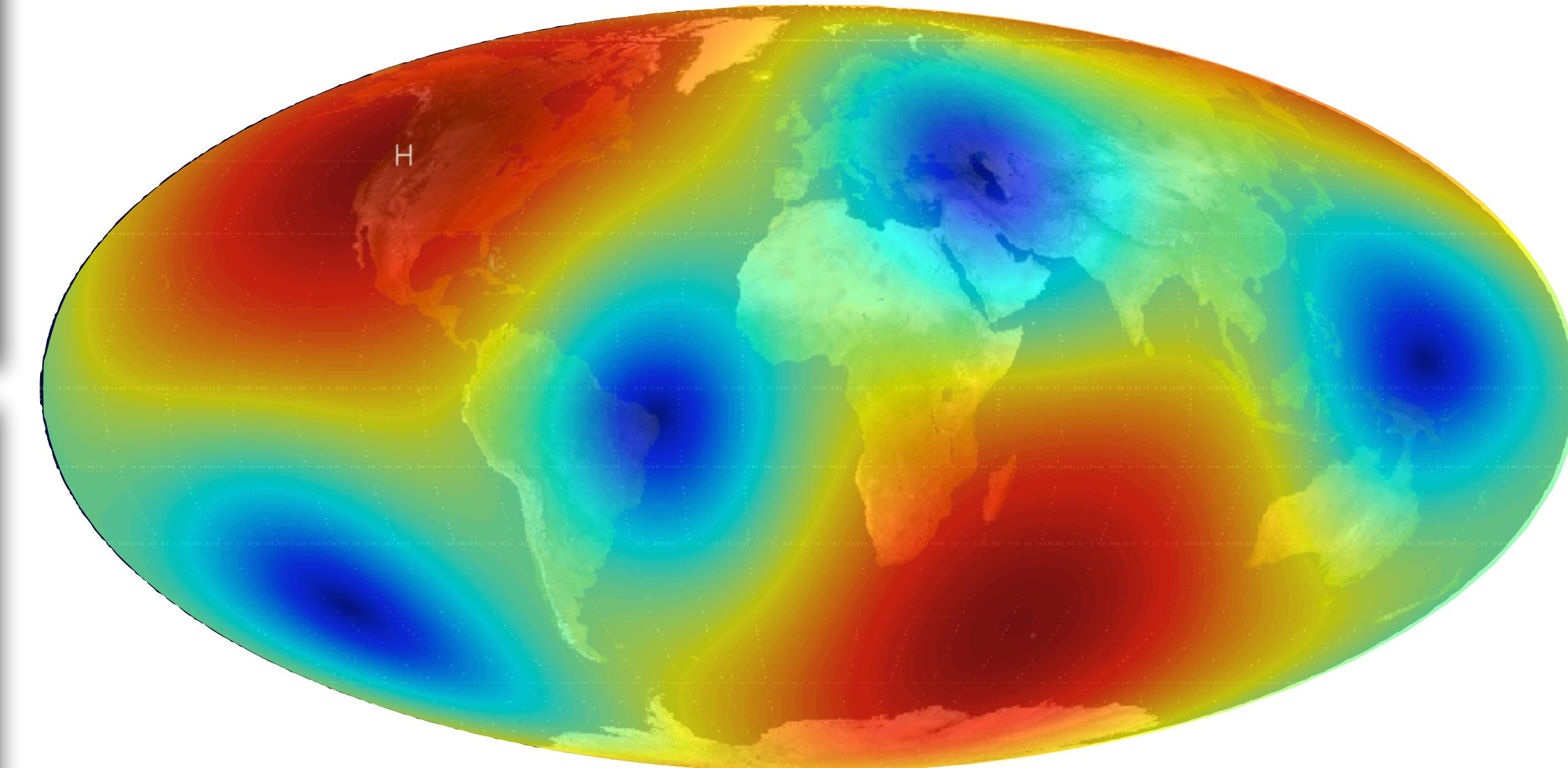
Terrestrial Network



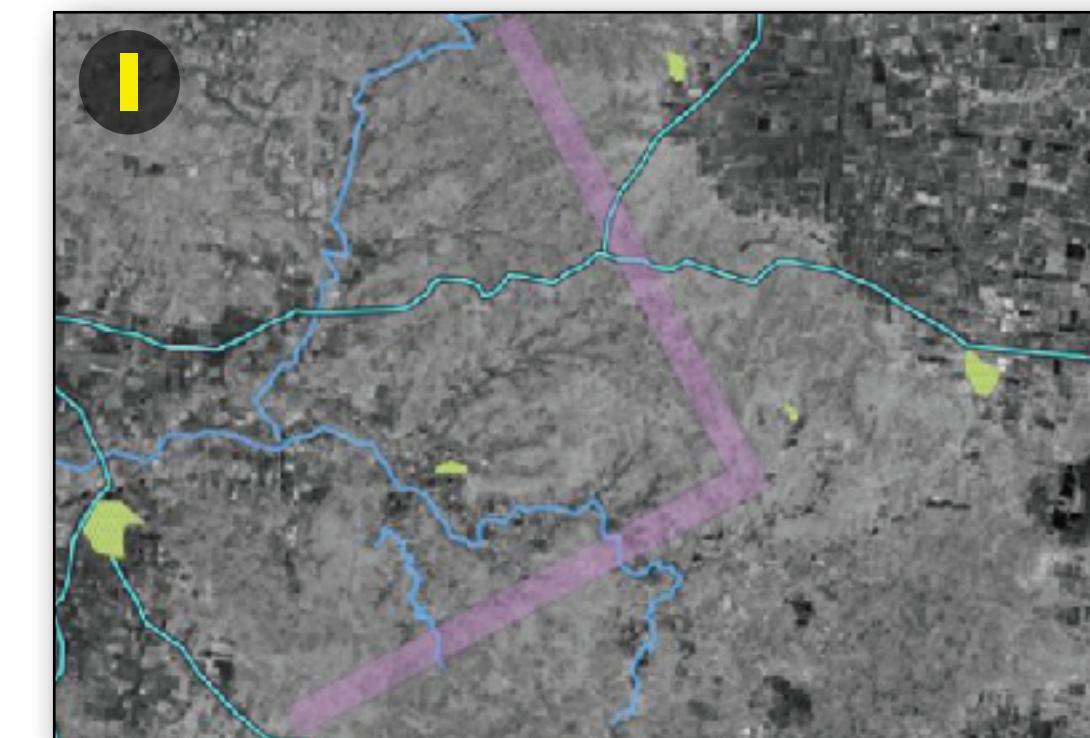
(credit: N. Cornish)



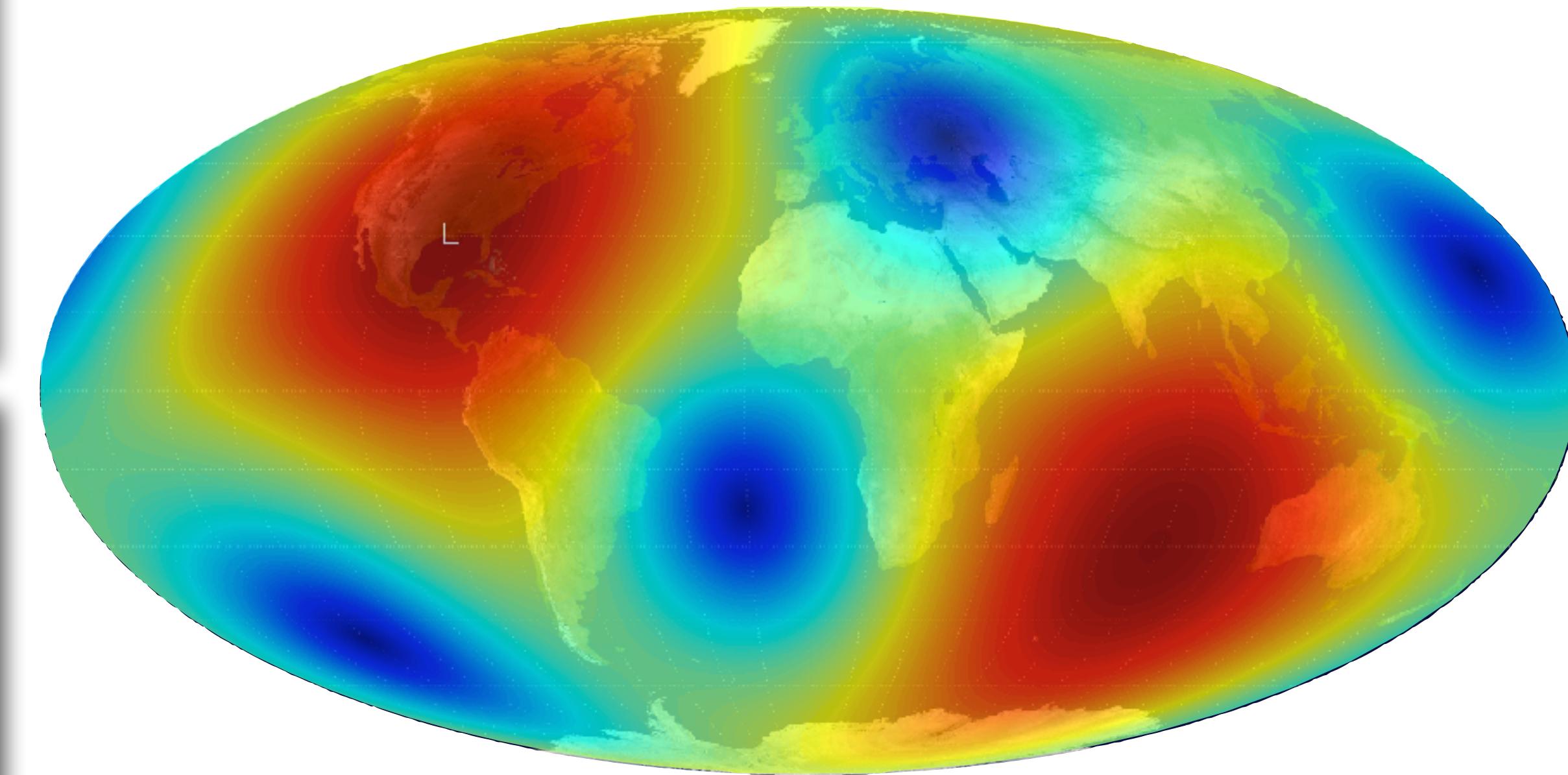
Terrestrial Network



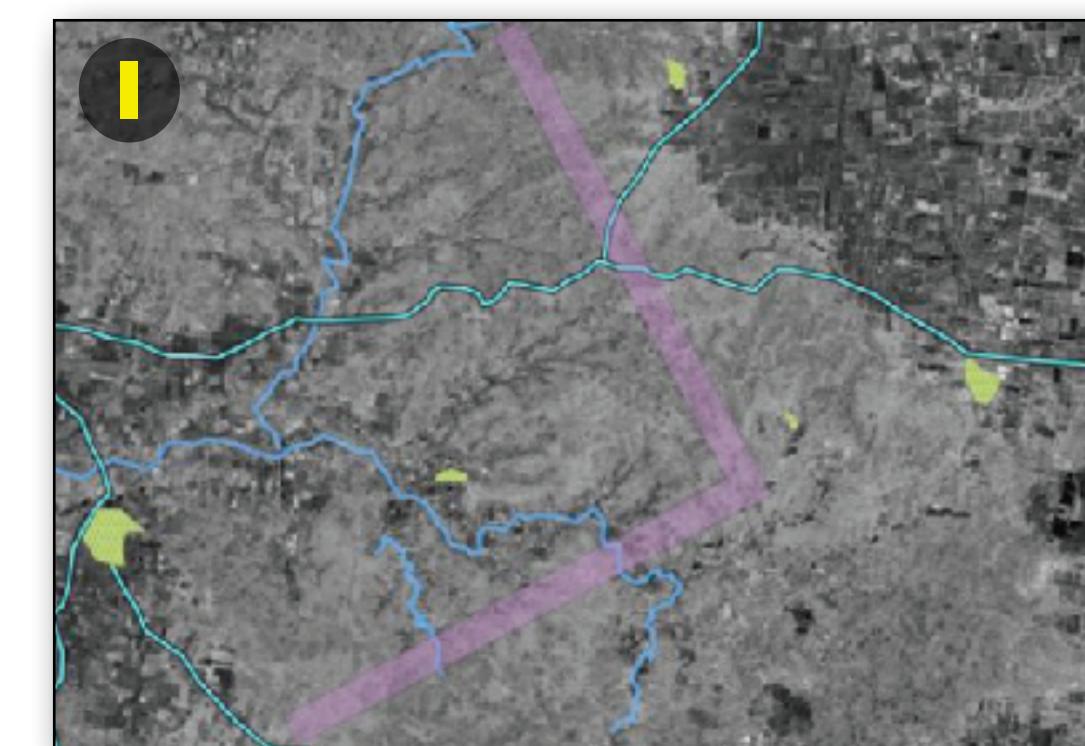
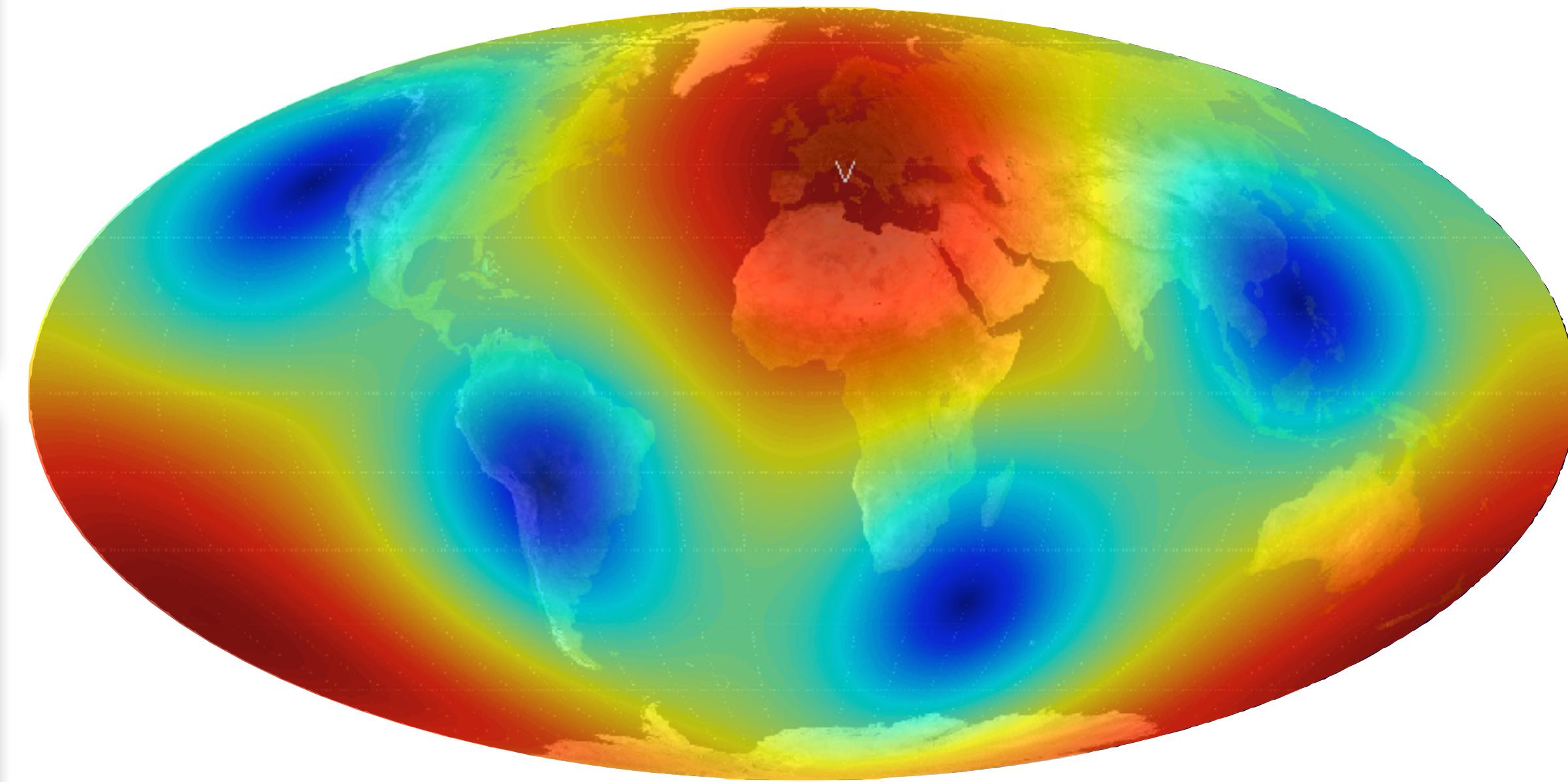
(credit: N. Cornish)



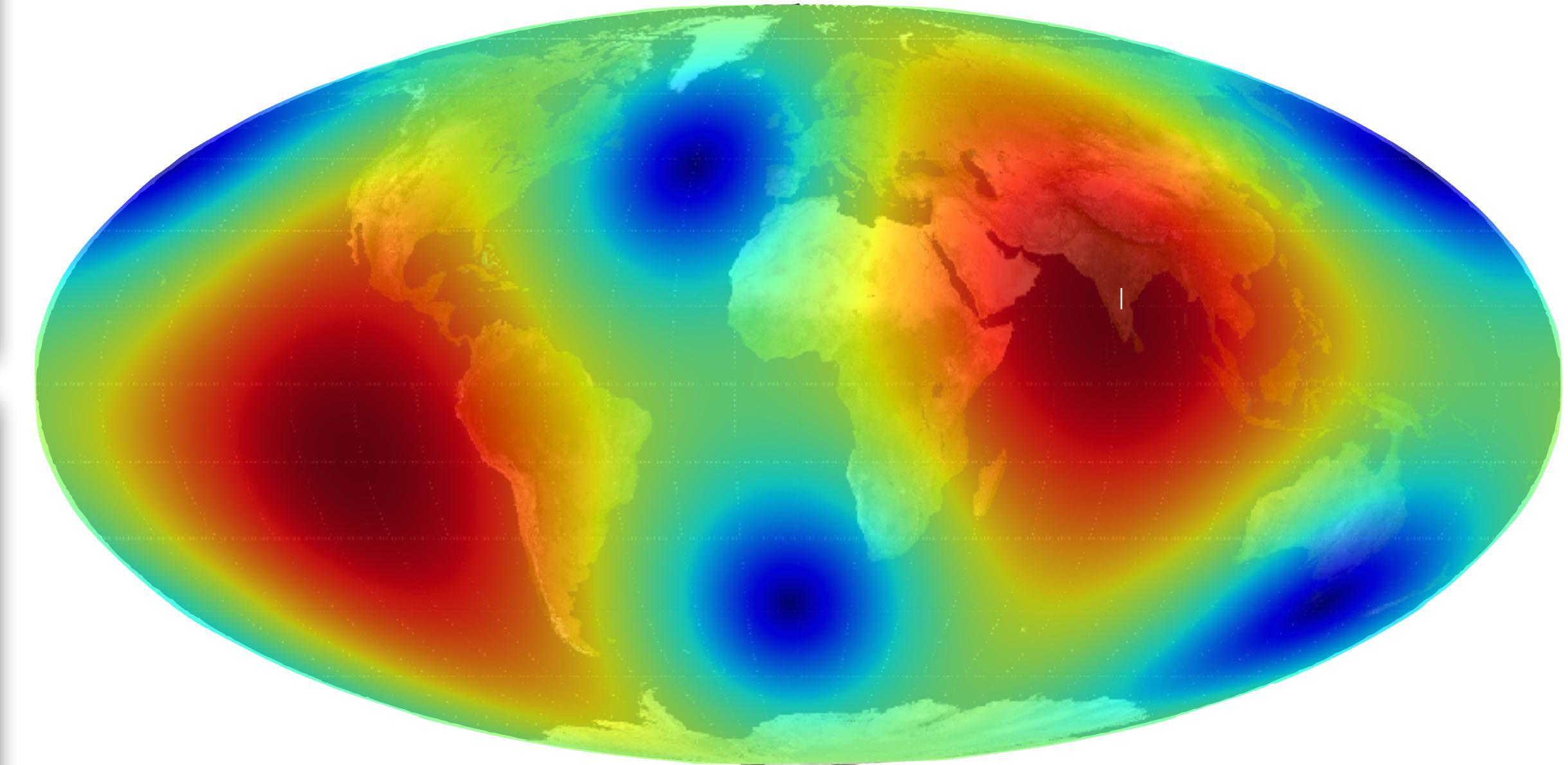
Terrestrial Network



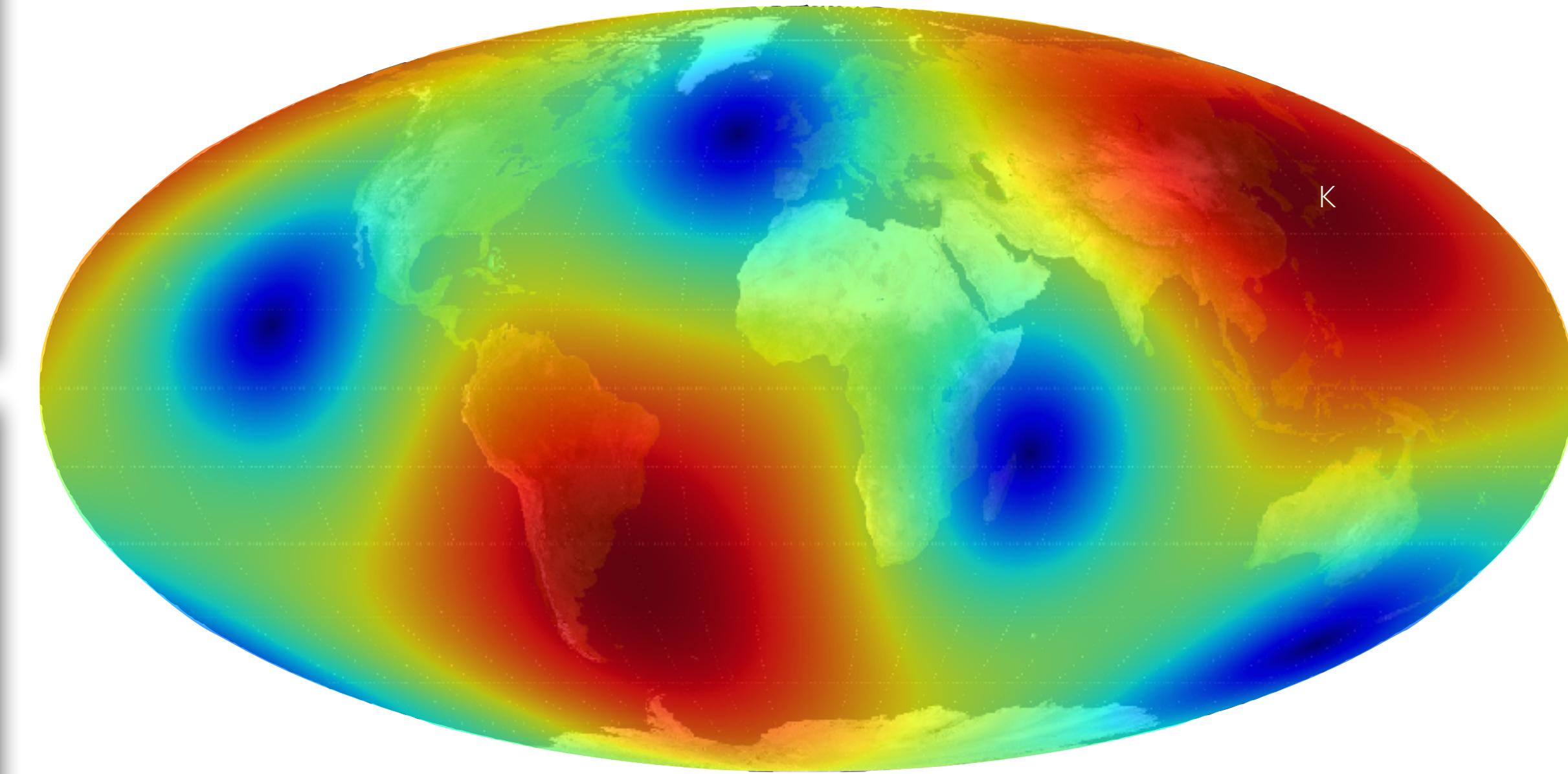
Terrestrial Network



Terrestrial Network



Terrestrial Network



(credit: N. Cornish)



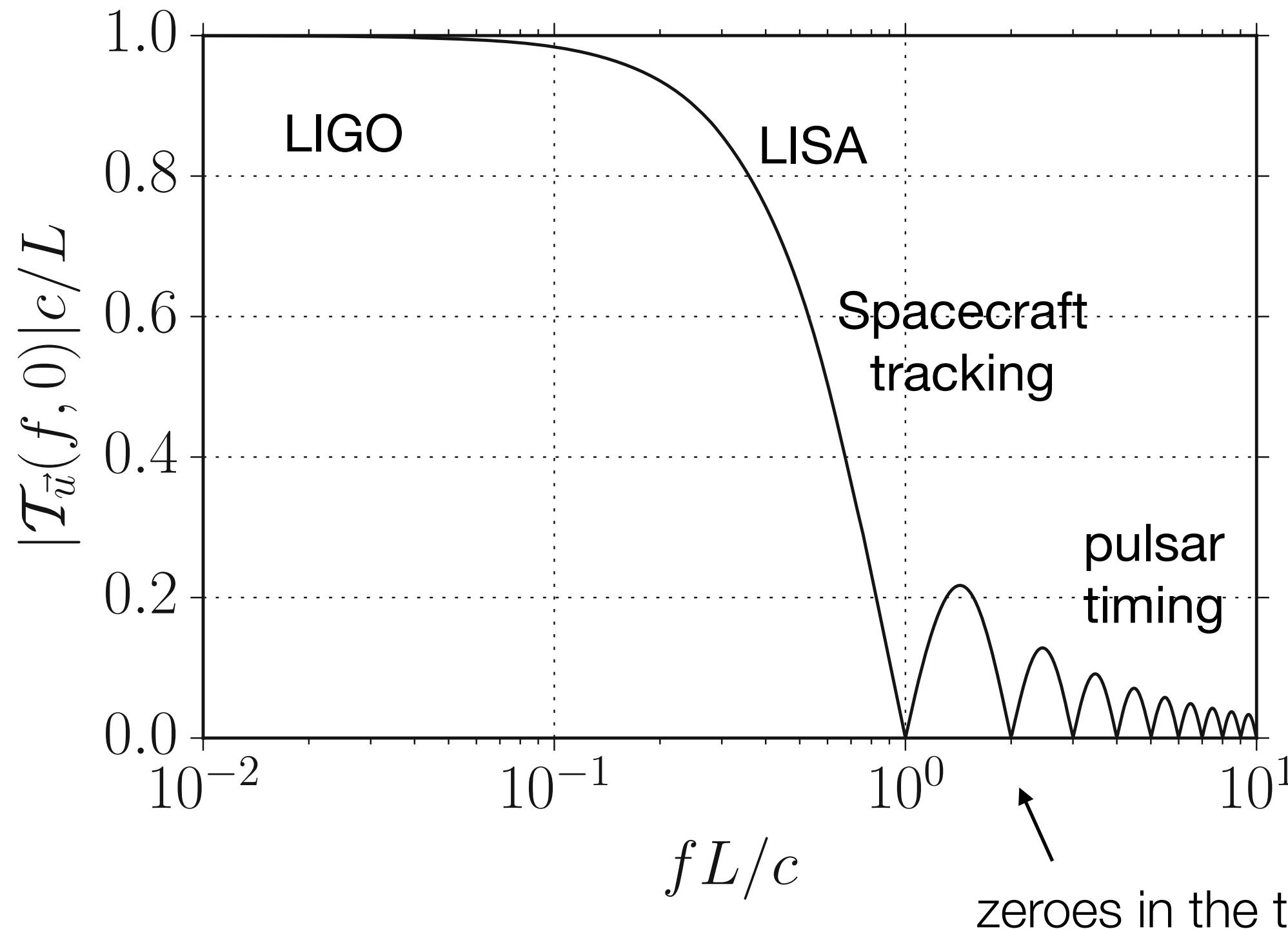
Beyond the short-antenna limit

LISA, spacecraft Doppler tracking and pulsar timing all operate outside of the short-antenna limit

Recall response function:
(one-arm, one-way)

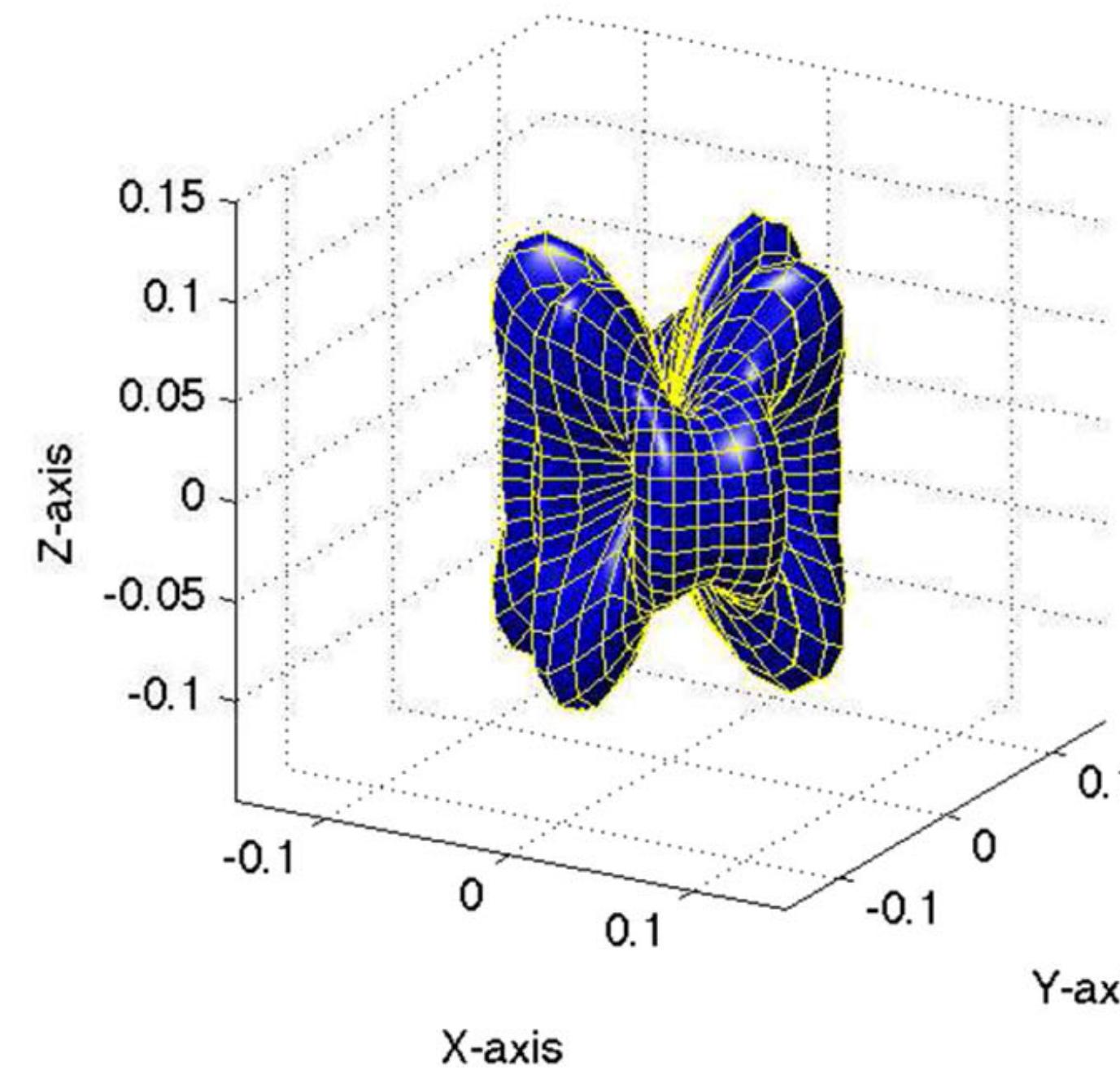
$$R^A(f, \hat{n}) = \frac{1}{2} u^a u^b e_{ab}^A(\hat{n}) \frac{1}{i2\pi f} \frac{1}{1 + \hat{n} \cdot \hat{u}} \left[1 - e^{-\frac{i2\pi f L}{c}(1 + \hat{n} \cdot \hat{u})} \right] e^{i2\pi f \hat{n} \cdot \vec{r}_2 / c}$$

$$\mathcal{T}_{\hat{u}}(f, \hat{n} \cdot \hat{u}) \equiv \frac{1}{i2\pi f} \frac{1}{1 + \hat{n} \cdot \hat{u}} \left[1 - e^{-\frac{i2\pi f L}{c}(1 + \hat{n} \cdot \hat{u})} \right] = \frac{L}{c} e^{-\frac{i\pi f L}{c}(1 + \hat{n} \cdot \hat{u})} \text{sinc}\left(\frac{\pi f L}{c}[1 + \hat{n} \cdot \hat{u}]\right)$$

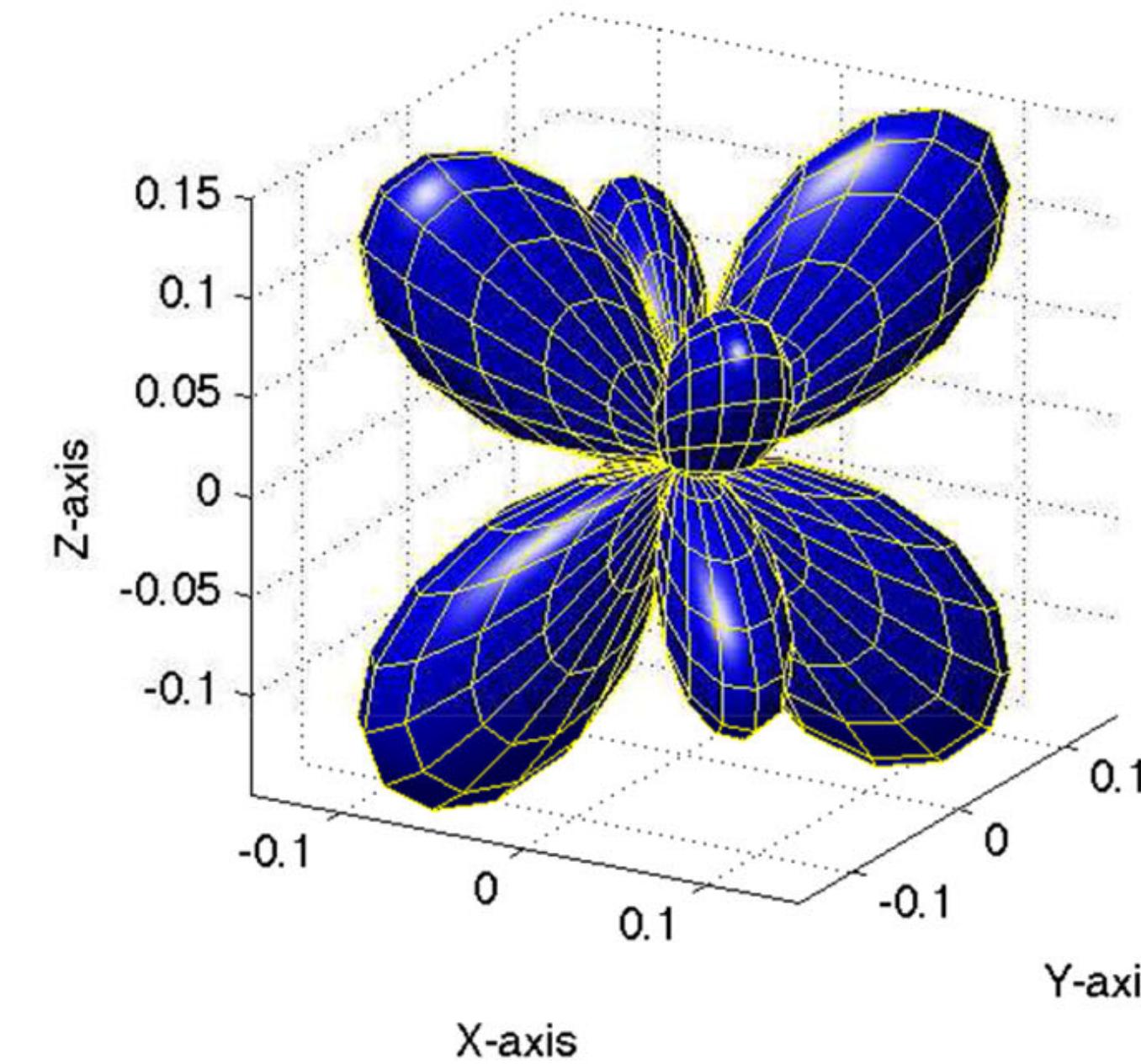


Beam detector	L (km)	f_* (Hz)	f (Hz)	f/f_*	Relation
Ground-based interferometer	~ 1	~ 10^5	$10 - 10^4$	$10^{-4} - 10^{-1}$	$f \ll f_*$
Space-based interferometer	~ 10^6	~ 10^{-1}	$10^{-4} - 10^{-1}$	$10^{-3} - 1$	$f \lesssim f_*$
Spacecraft Doppler tracking	~ 10^9	~ 10^{-4}	$10^{-6} - 10^{-3}$	$10^{-2} - 10$	$f \sim f_*$
Pulsar timing	~ 10^{17}	~ 10^{-12}	$10^{-9} - 10^{-7}$	$10^3 - 10^5$	$f \gg f_*$

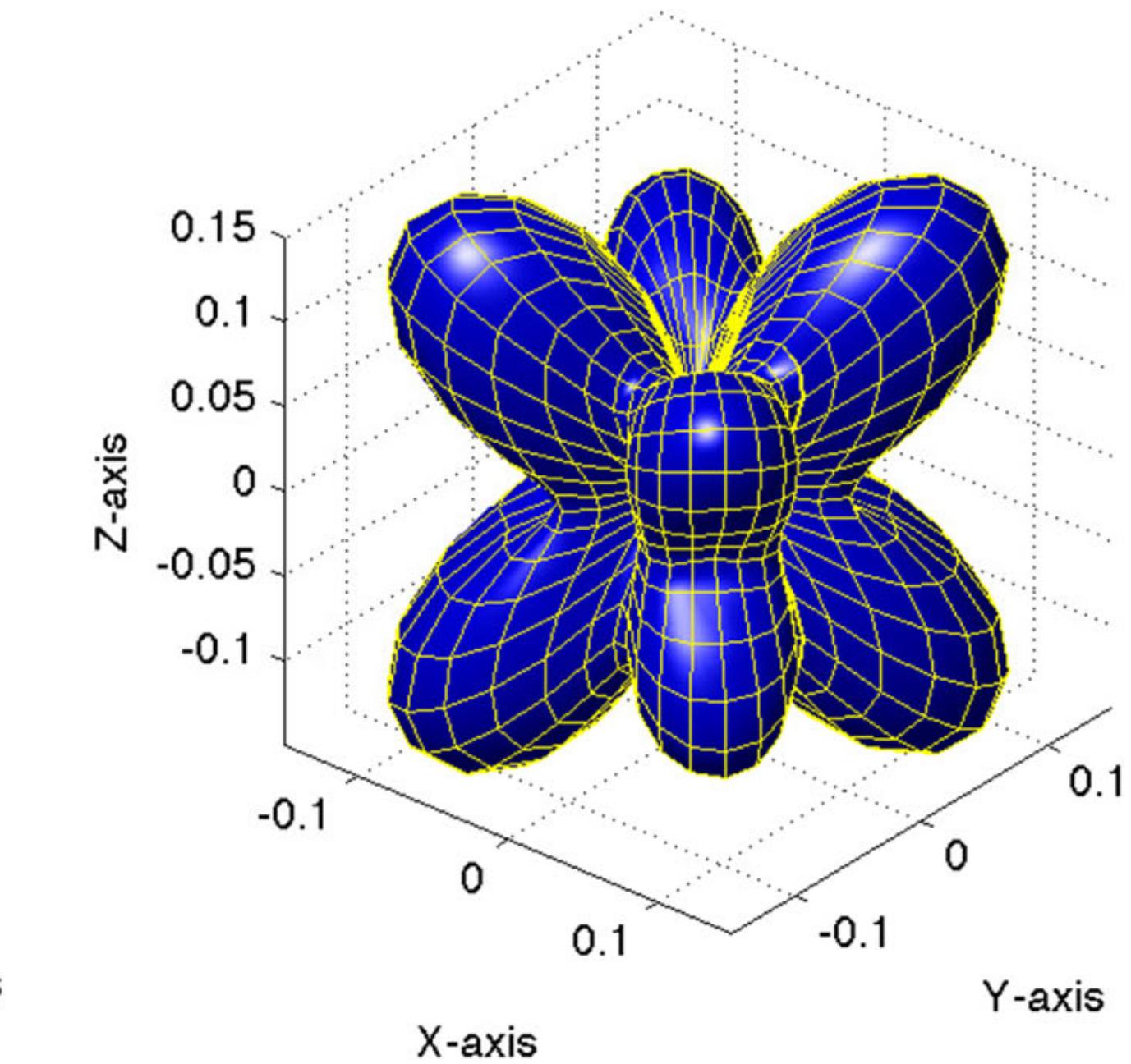
Beam pattern functions



$$|R^+(f, \hat{n})|$$



$$|R^\times(f, \hat{n})|$$



$$\left(|R^+(f, \hat{n})|^2 + |R^\times(f, \hat{n})|^2 \right)^{1/2}$$

$(f = c/(2L) = 37.5 \text{ kHz})$

2. Non-trivial correlated response

Overlap function (correlation coefficient)

- Detectors in **different locations** and with **different orientations** respond differently to a passing GW.
- Overlap function encodes reduction in sensitivity of a cross-correlation analysis due to **separation** and **misalignment** of the detectors

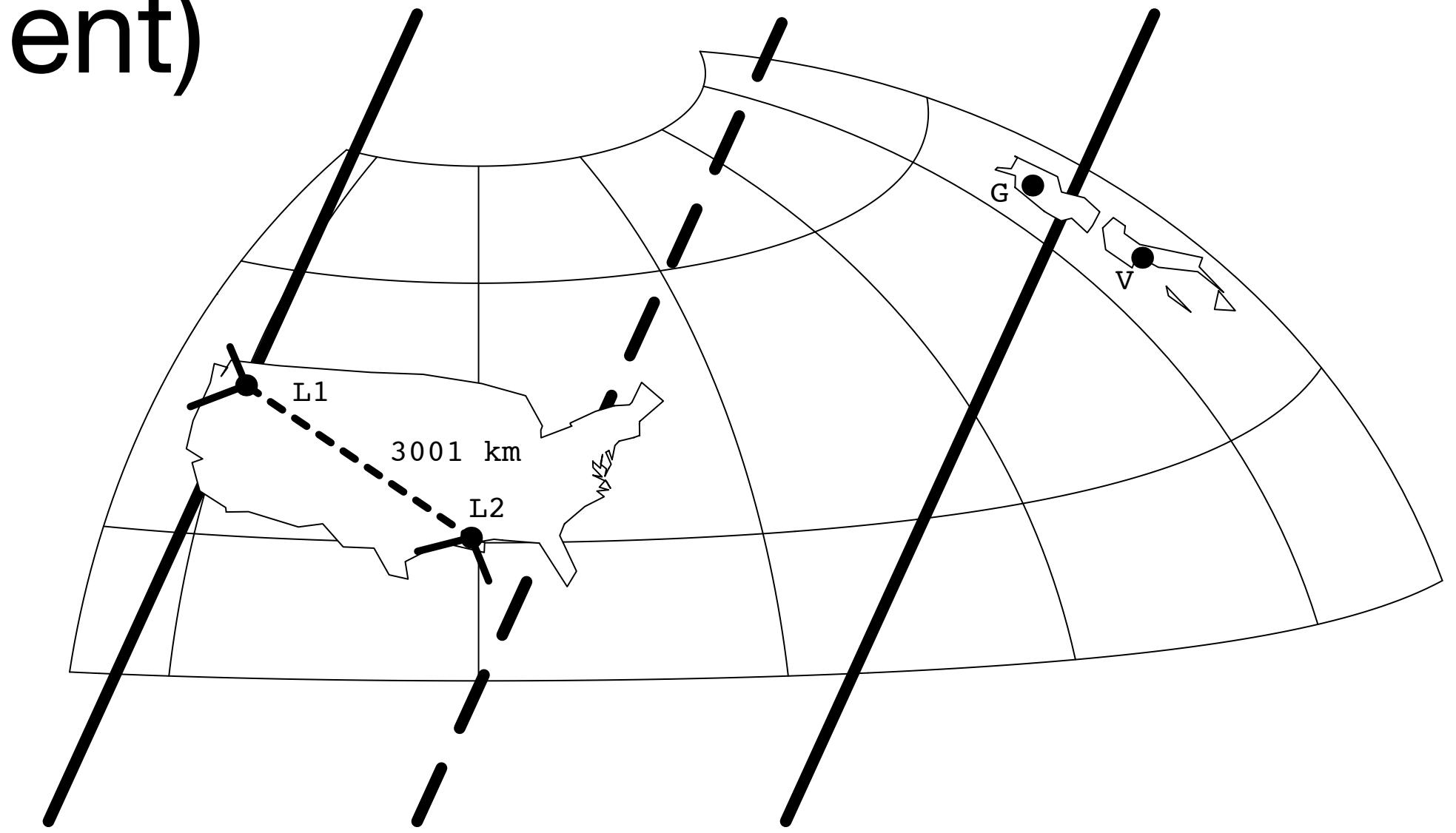
Expected correlation:

$$\langle h_I(t)h_J(t') \rangle = \frac{1}{2} \int_{-\infty}^{\infty} df e^{i2\pi f(t-t')} \Gamma_{IJ}(f) S_h(f) \quad \Leftrightarrow \quad \langle \tilde{h}_I(f)\tilde{h}_J^*(f') \rangle = \frac{1}{2} \delta(f-f') \Gamma_{IJ}(f) S_h(f)$$

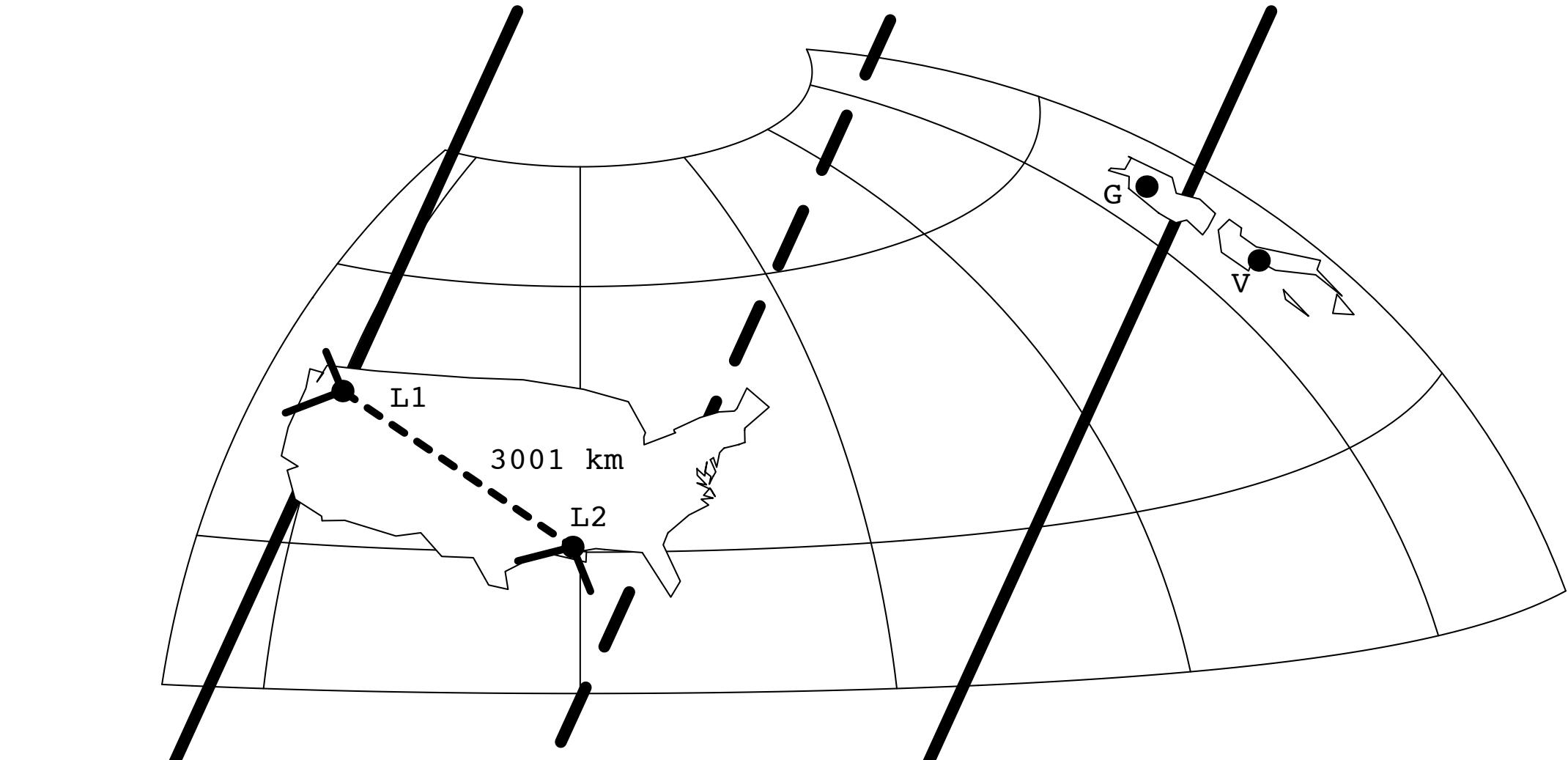
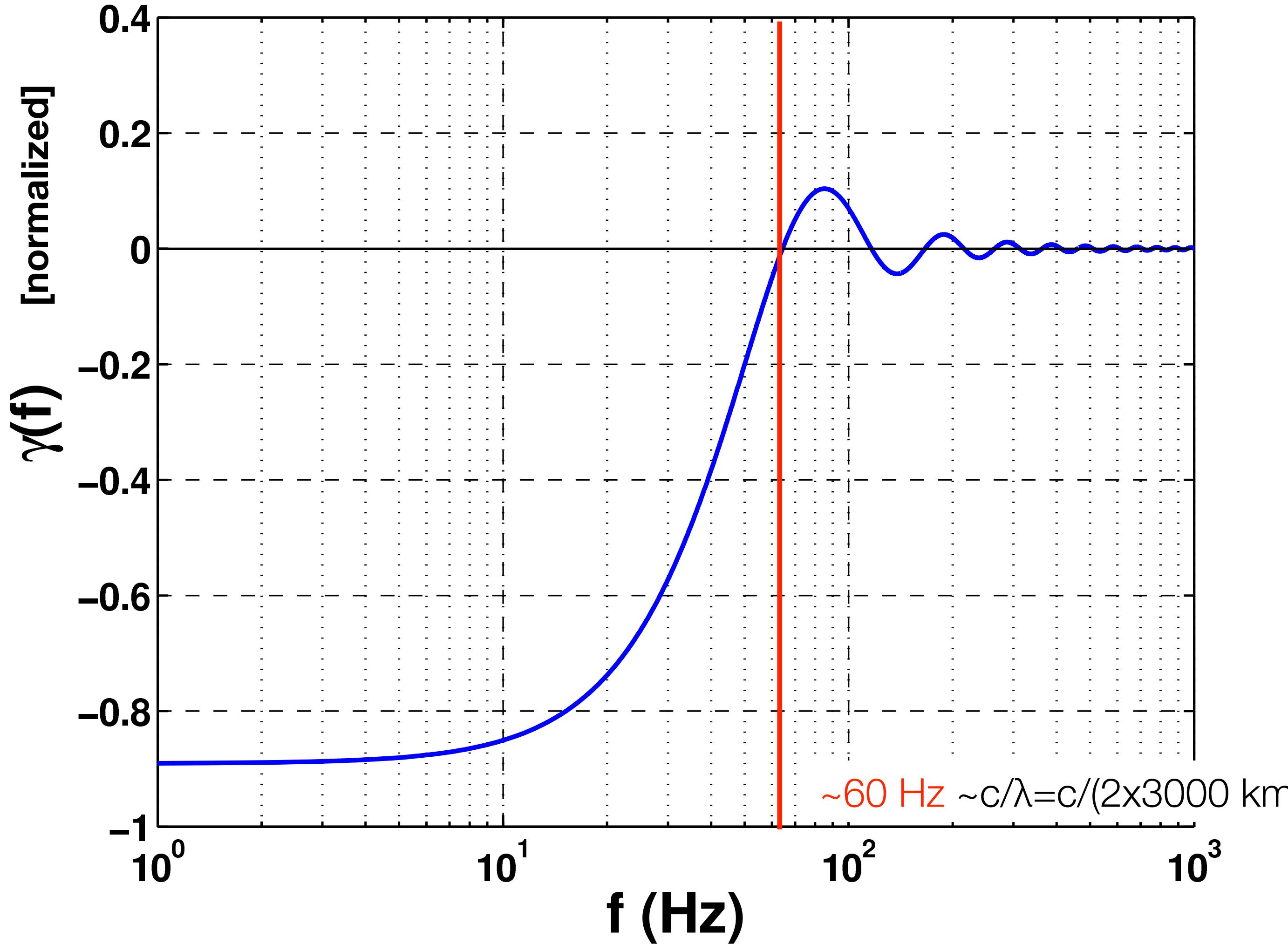
$$\Gamma_{IJ}(f) = \frac{1}{8\pi} \int d^2\Omega_{\hat{n}} \sum_A R_I^A(f, \hat{n}) R_J^{A*}(f, \hat{n})$$

(assumes unpolarized, isotropic background)

$\Gamma_{IJ}(f)$ is the **transfer function** between GW power and detector cross-power; **integrand** of $\Gamma_{IJ}(f)$ is important for **anisotropic** stochastic backgrounds

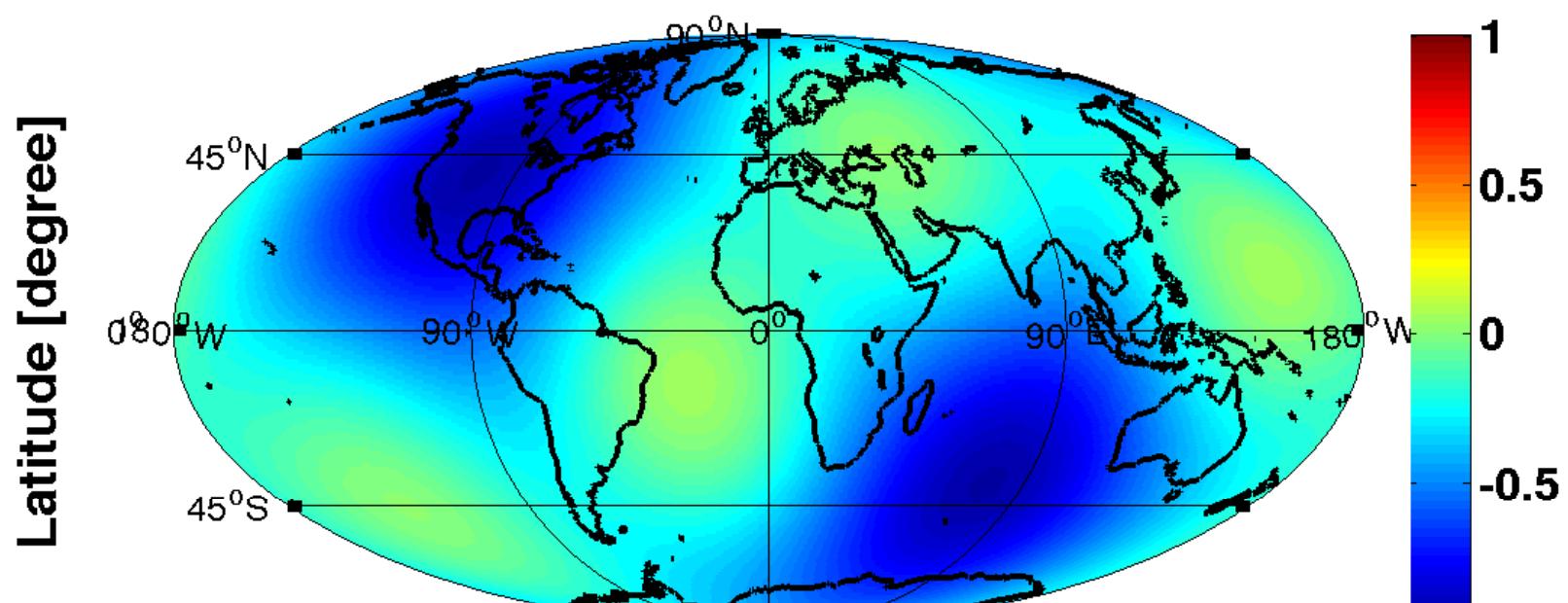


LIGO Hanford-LIGO Livingston overlap function (small-antenna approximation)

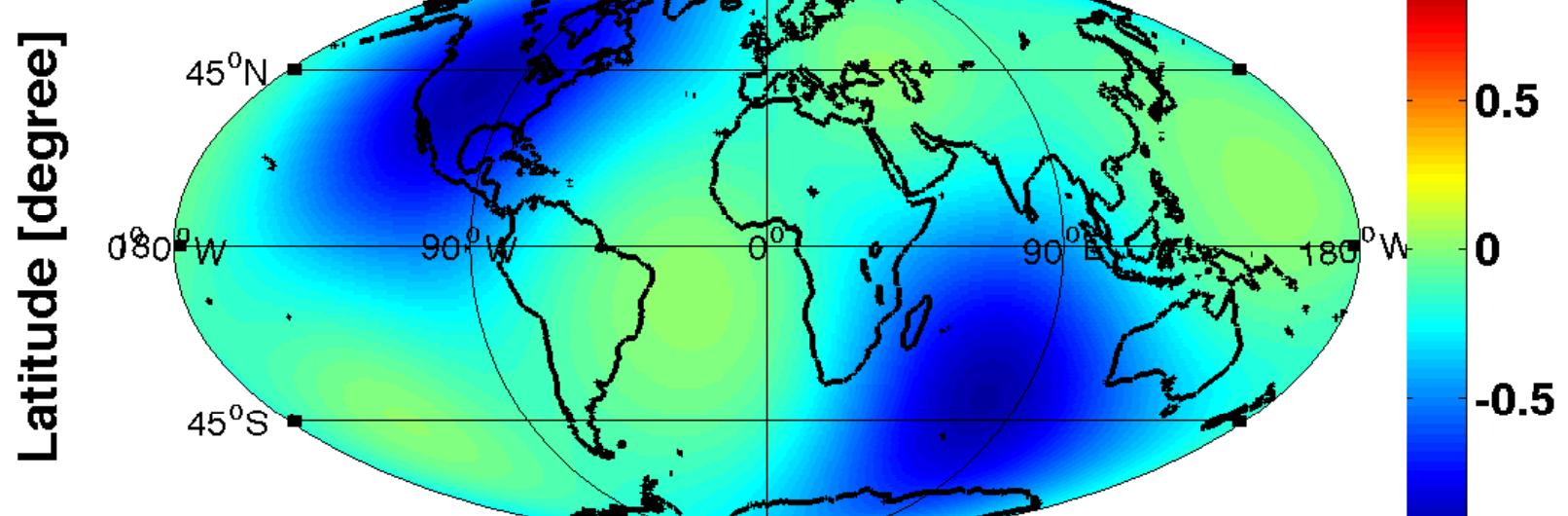


Overlap integrand (HL)

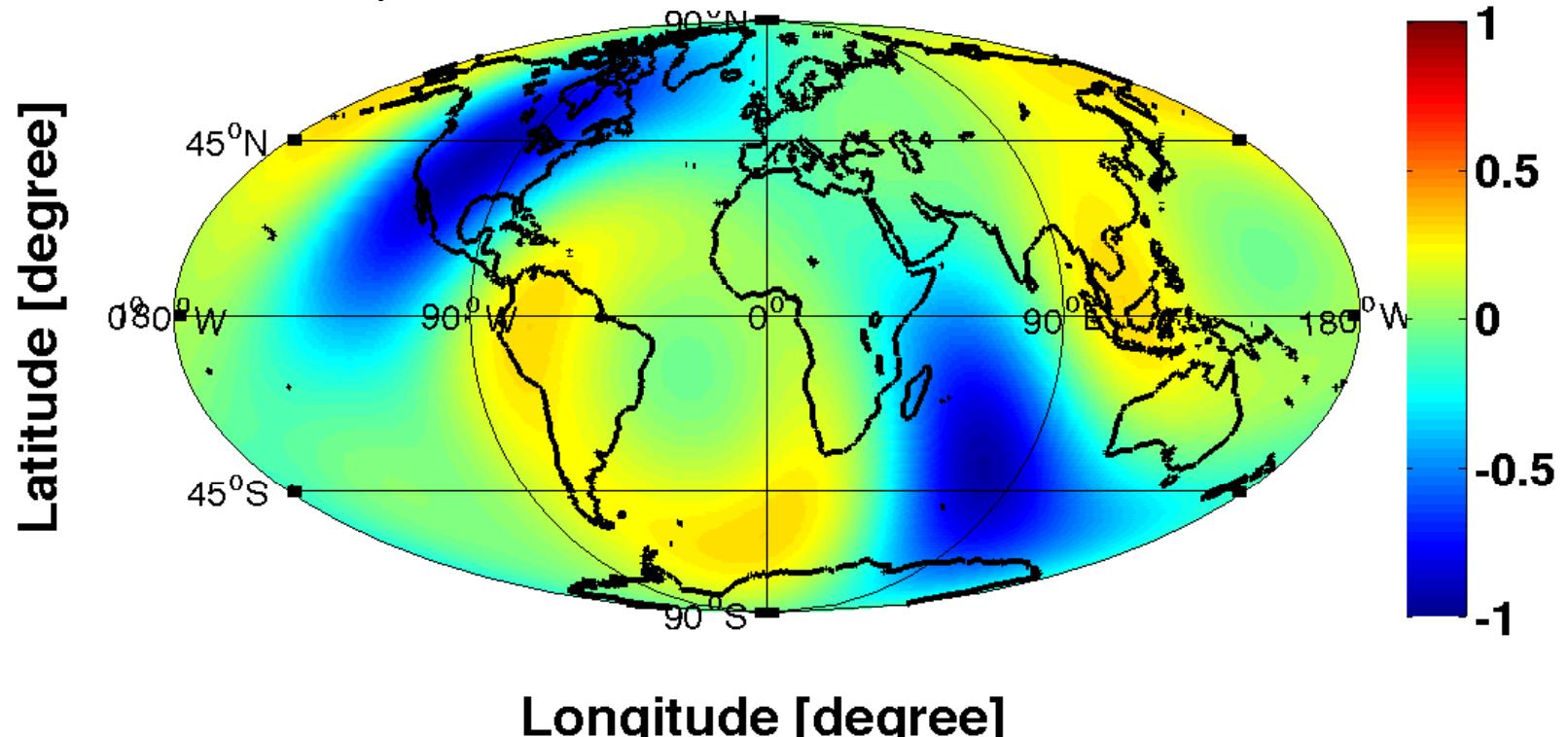
$$\gamma(0 \text{ Hz}) = -0.89$$



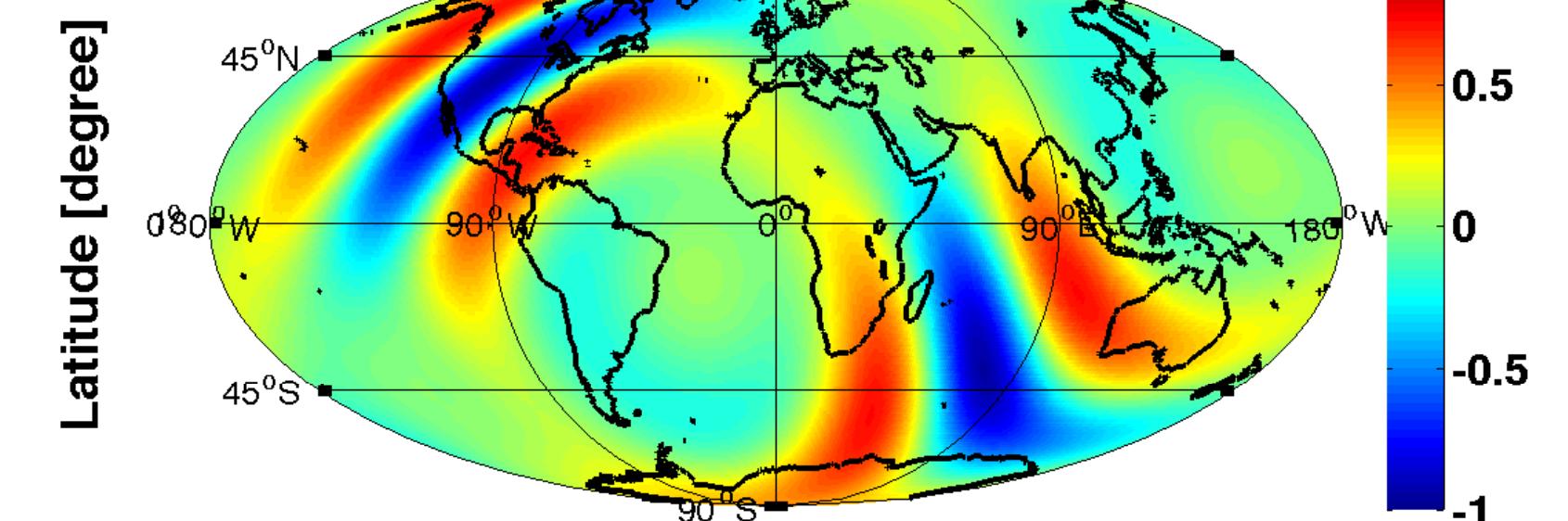
$$\gamma(20 \text{ Hz}) = -0.74$$



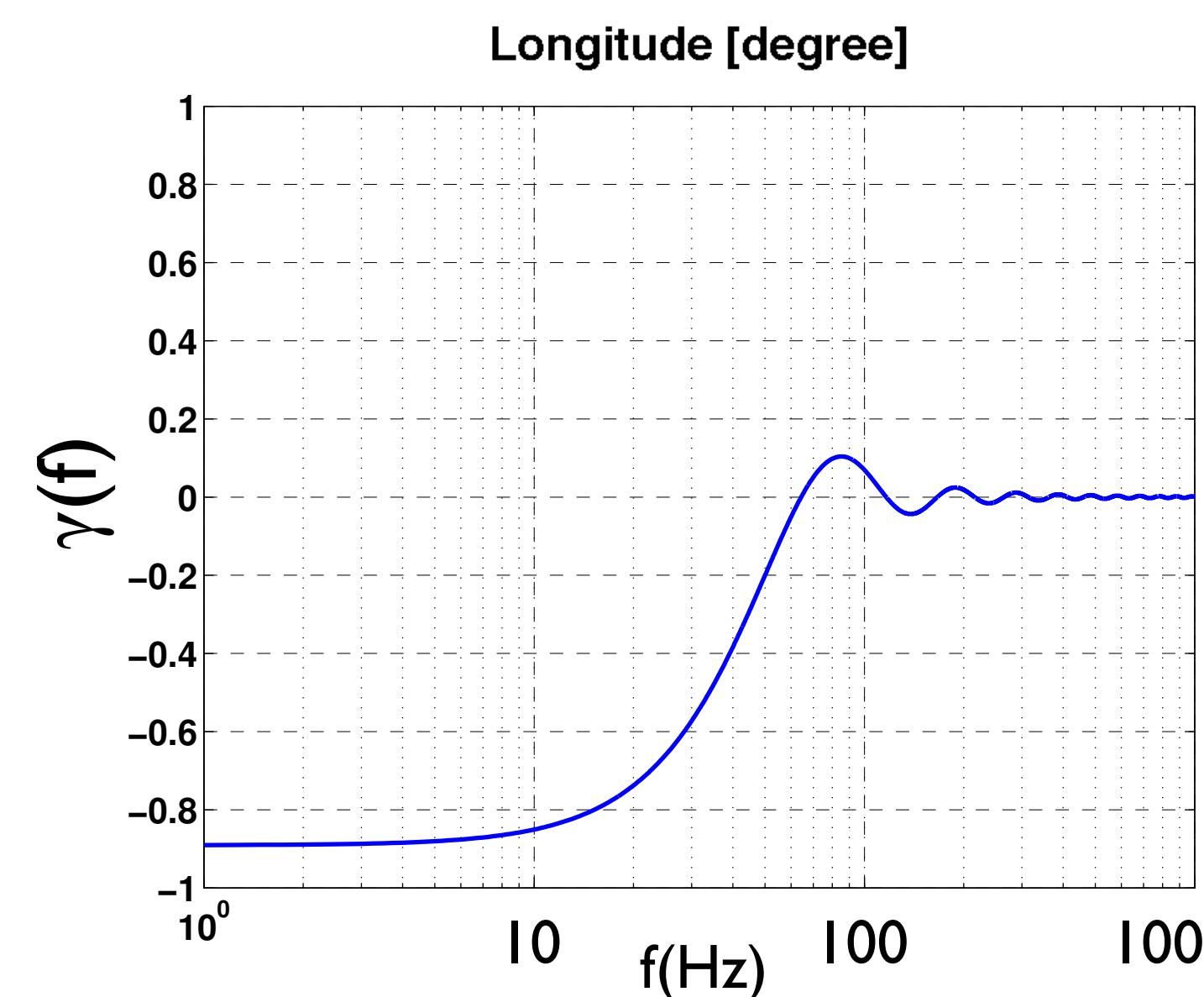
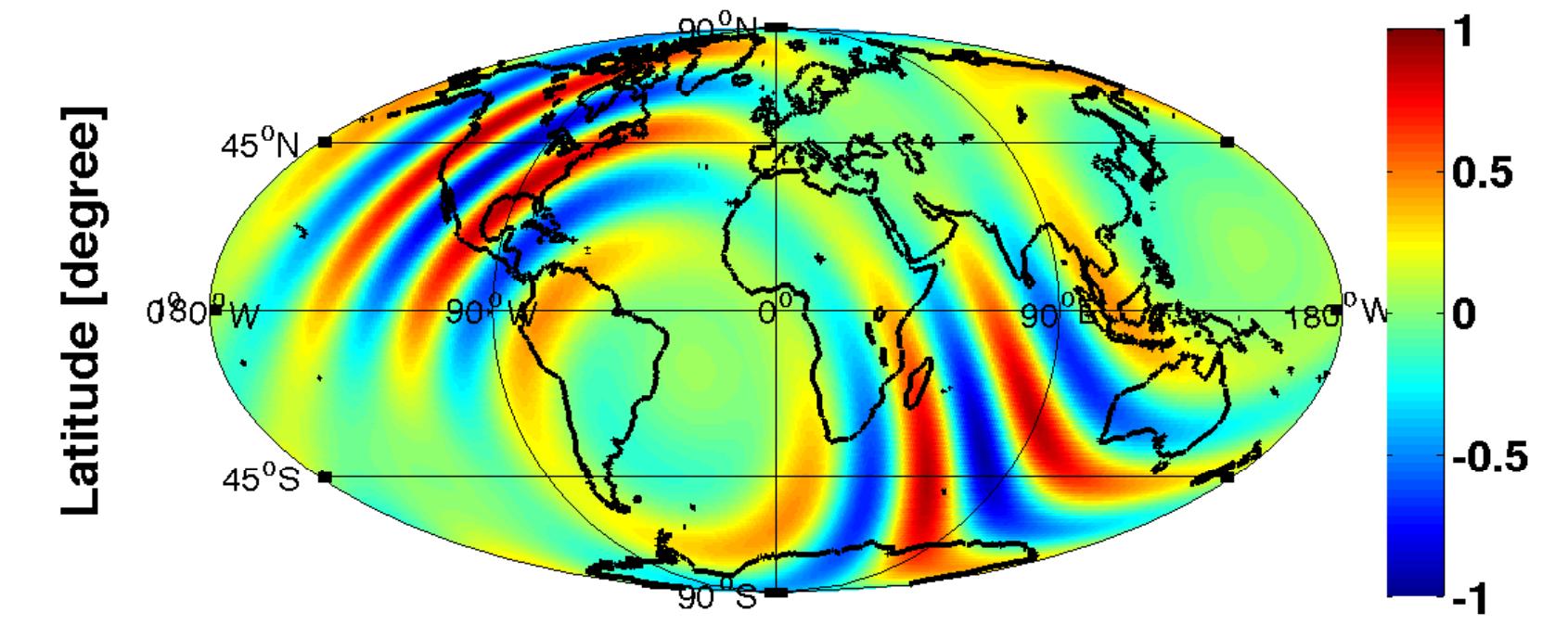
$$\gamma(50 \text{ Hz}) = -0.20$$



$$\gamma(100 \text{ Hz}) = 0.07$$

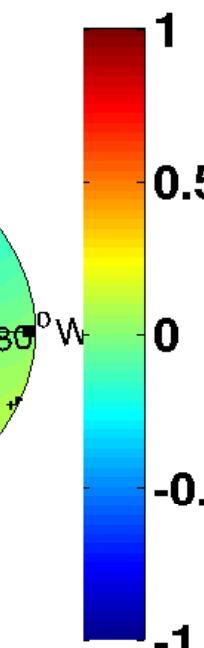
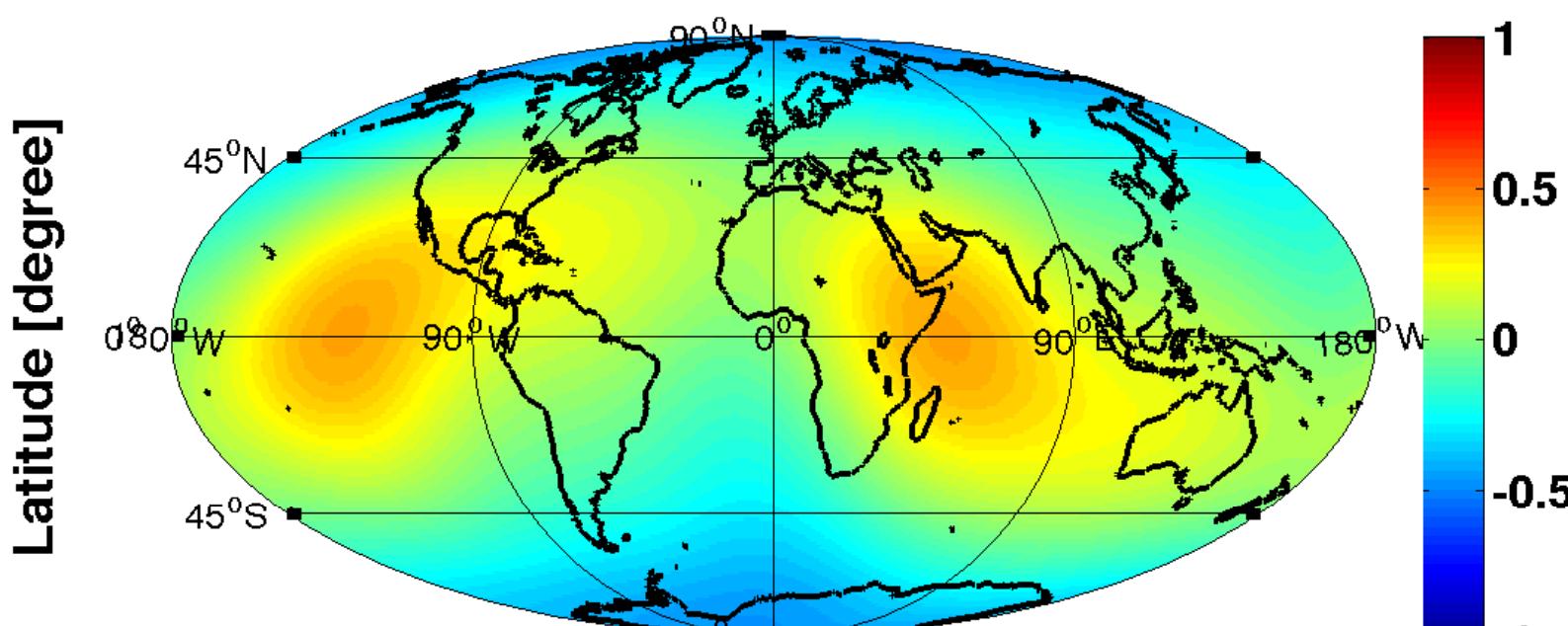


$$\gamma(200 \text{ Hz}) = 0.02$$

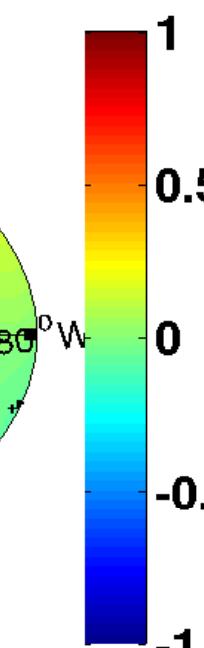
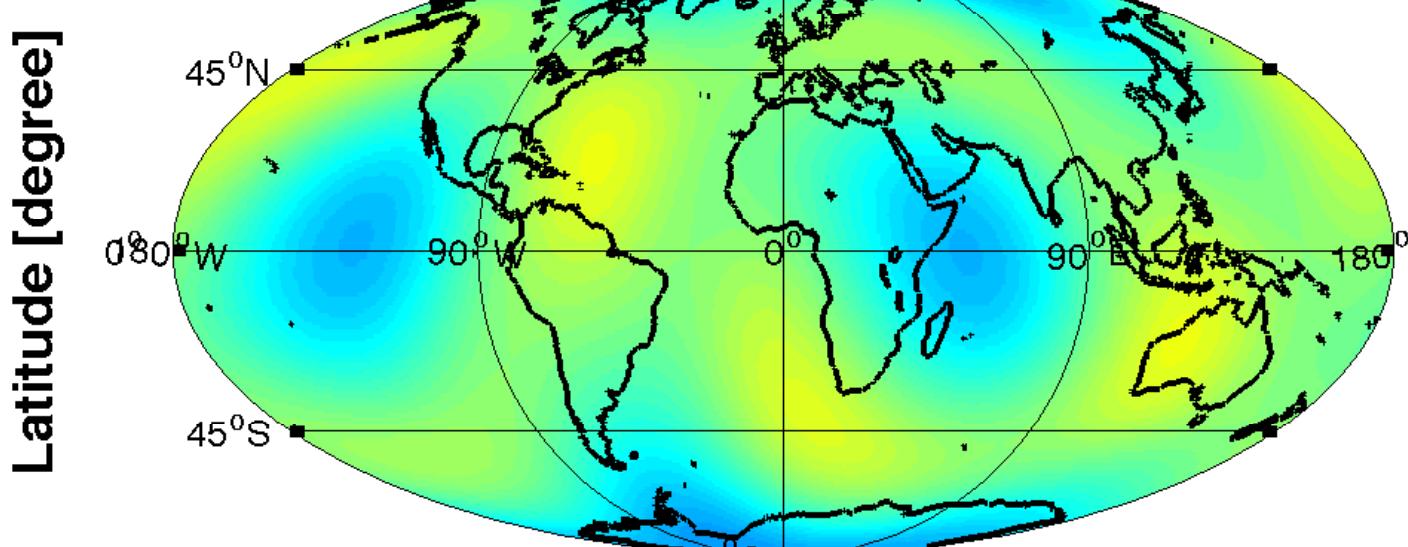


Overlap integrand (HV)

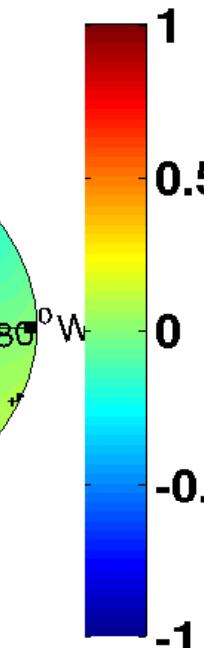
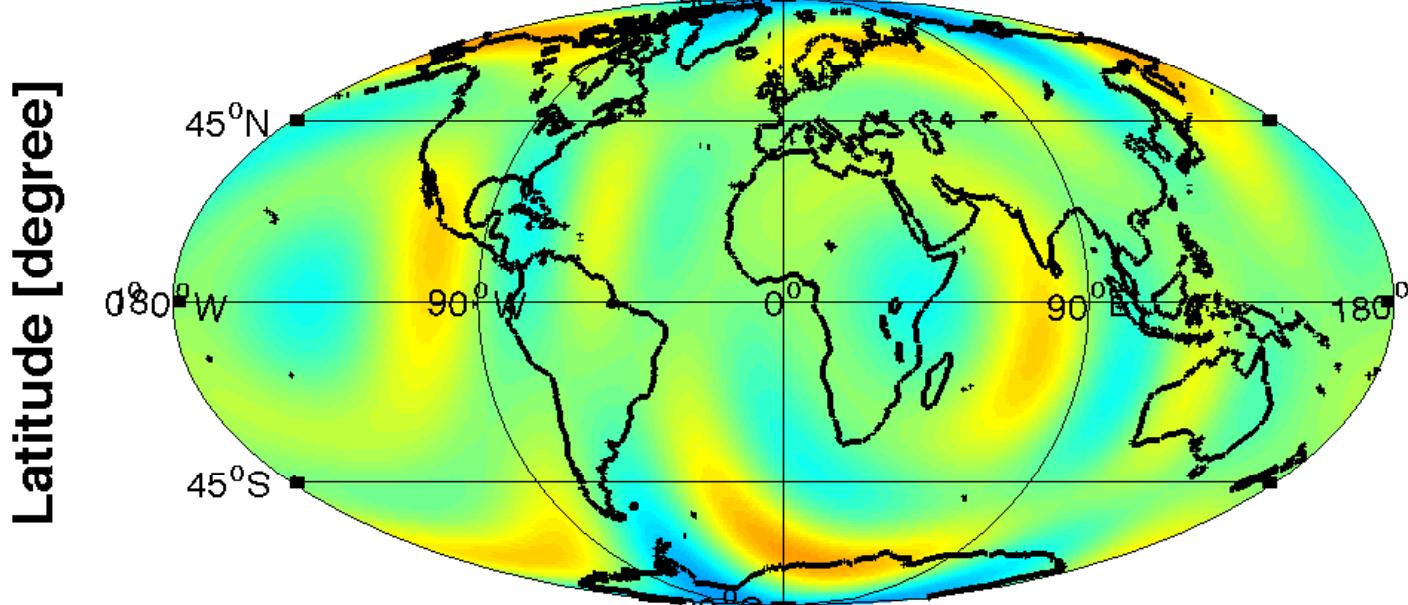
$$\gamma(0 \text{ Hz}) = -0.016$$



$$\gamma(20 \text{ Hz}) = -0.204$$

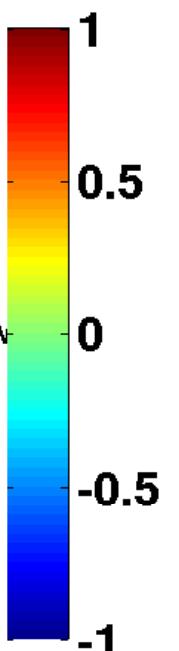
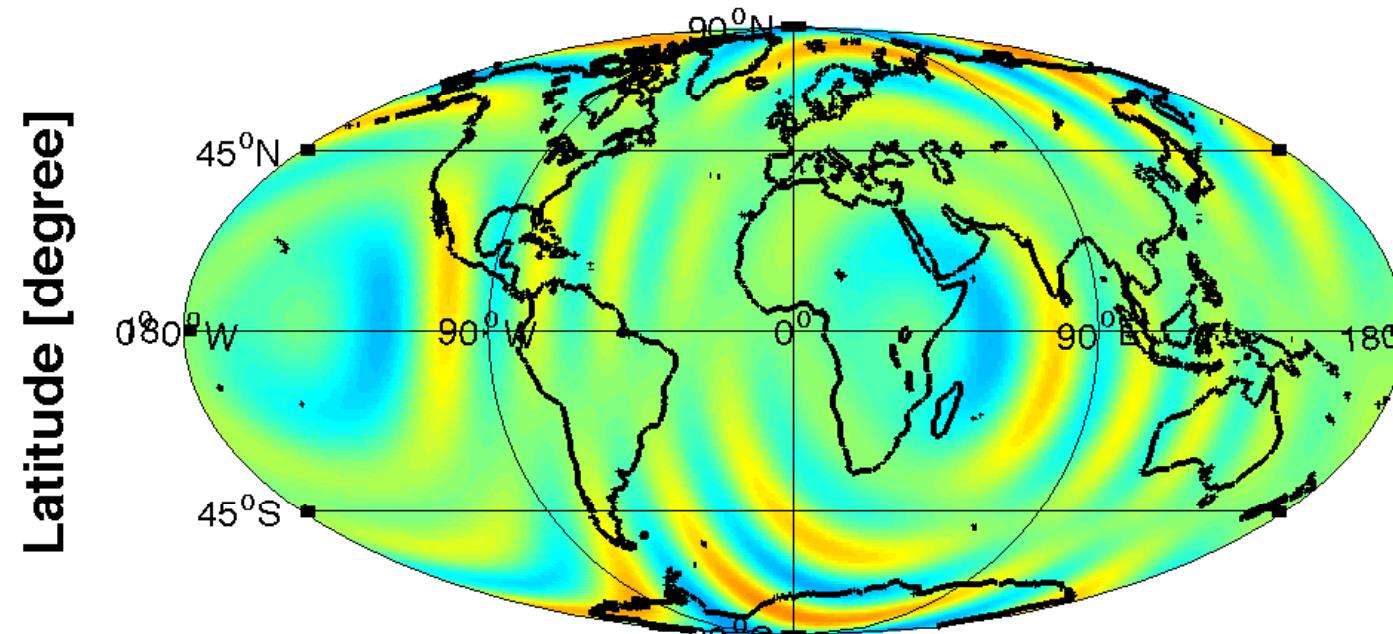


$$\gamma(50 \text{ Hz}) = 0.033$$

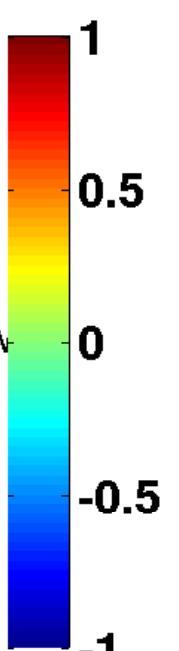
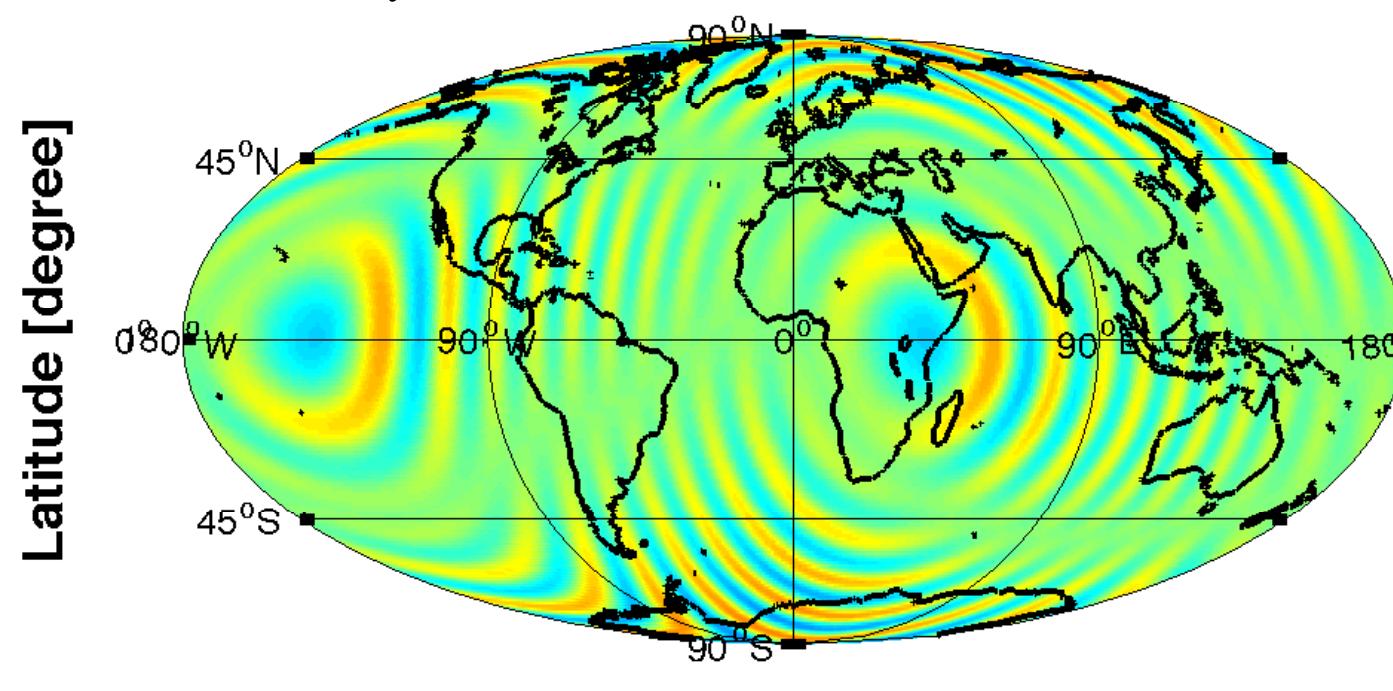


Longitude [degree]

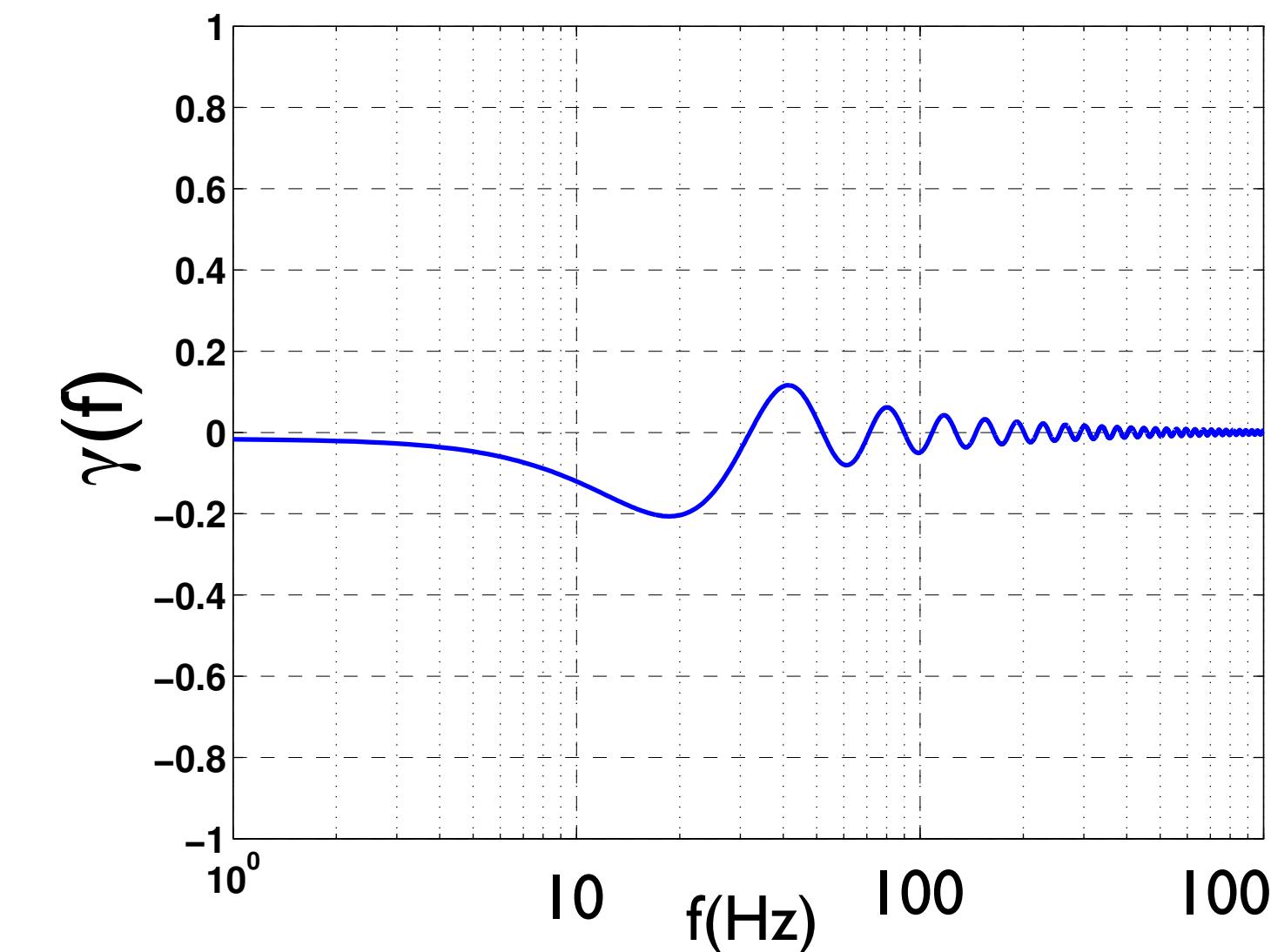
$$\gamma(100 \text{ Hz}) = -0.049$$



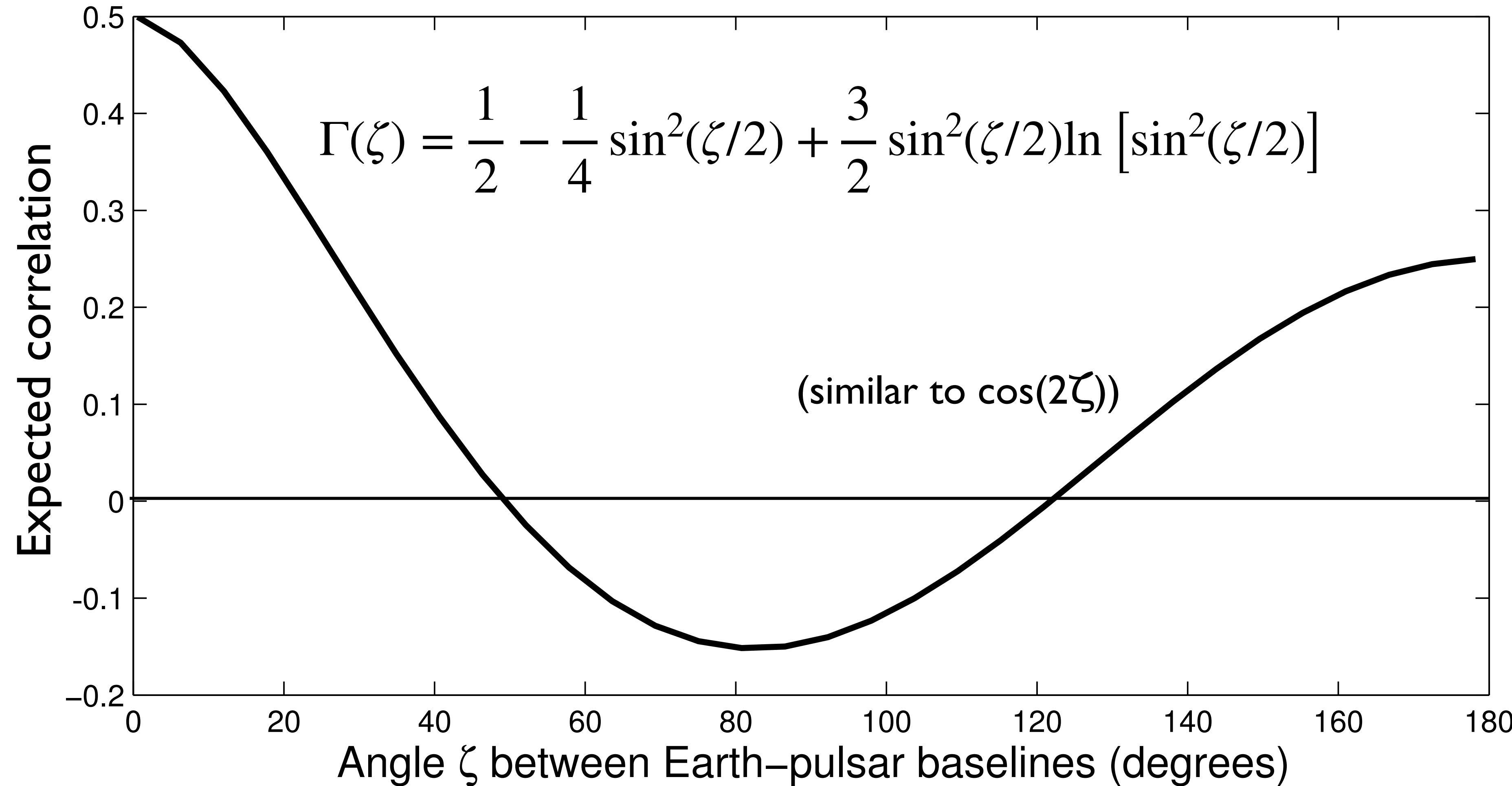
$$\gamma(200 \text{ Hz}) = 0.004$$



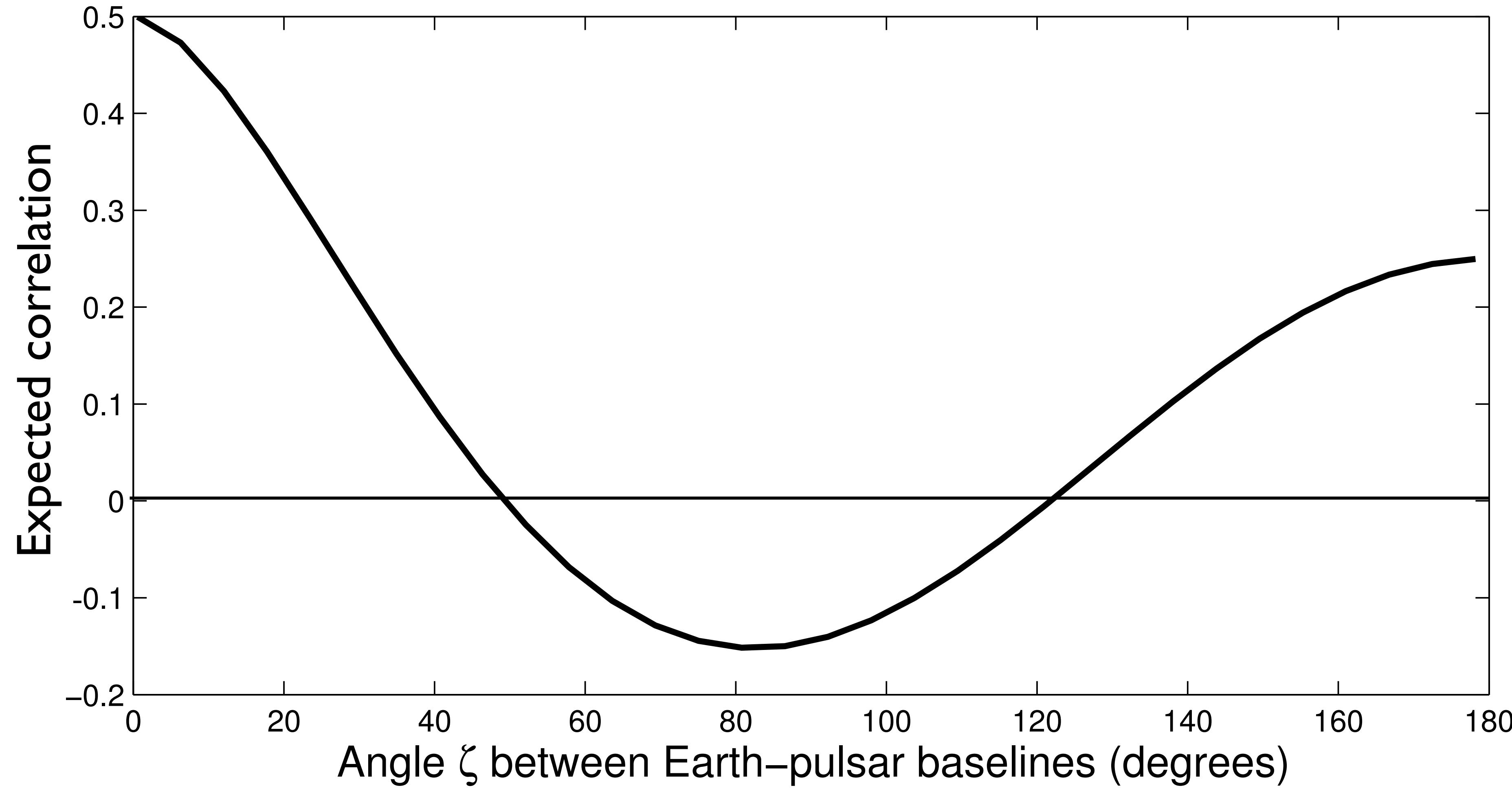
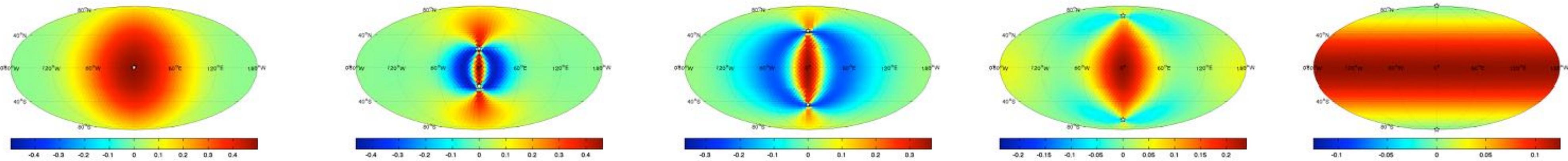
Longitude [degree]



Pulsar timing correlations (Hellings & Downs curve)
(correlation for an isotropic, unpolarized GW background in GR)



overlap integrand for different angular separations

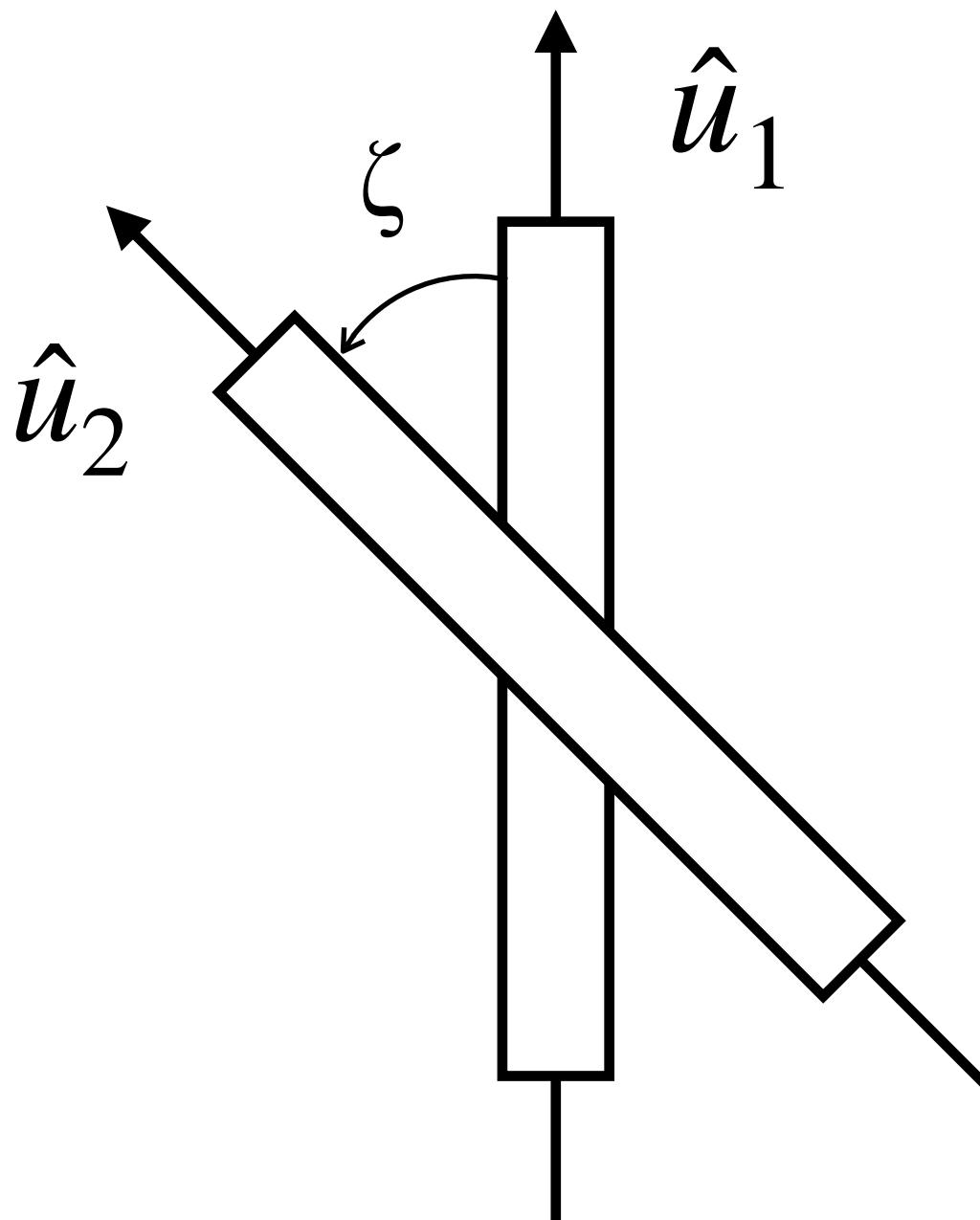


Exercise 7: Show that the overlap function for a pair of short, colocated electric dipole antennae pointing in direction \hat{u}_1 and \hat{u}_2 is given by

$$\Gamma_{12} \propto \hat{u}_1 \cdot \hat{u}_2 \equiv \cos \zeta$$

for an unpolarized, isotropic electromagnetic field.

Jenet and Romano, AJP 83 (7), 2015



Hint: $r_I(t) = \hat{u}_I \cdot \vec{E}(t, \vec{x}_0)$

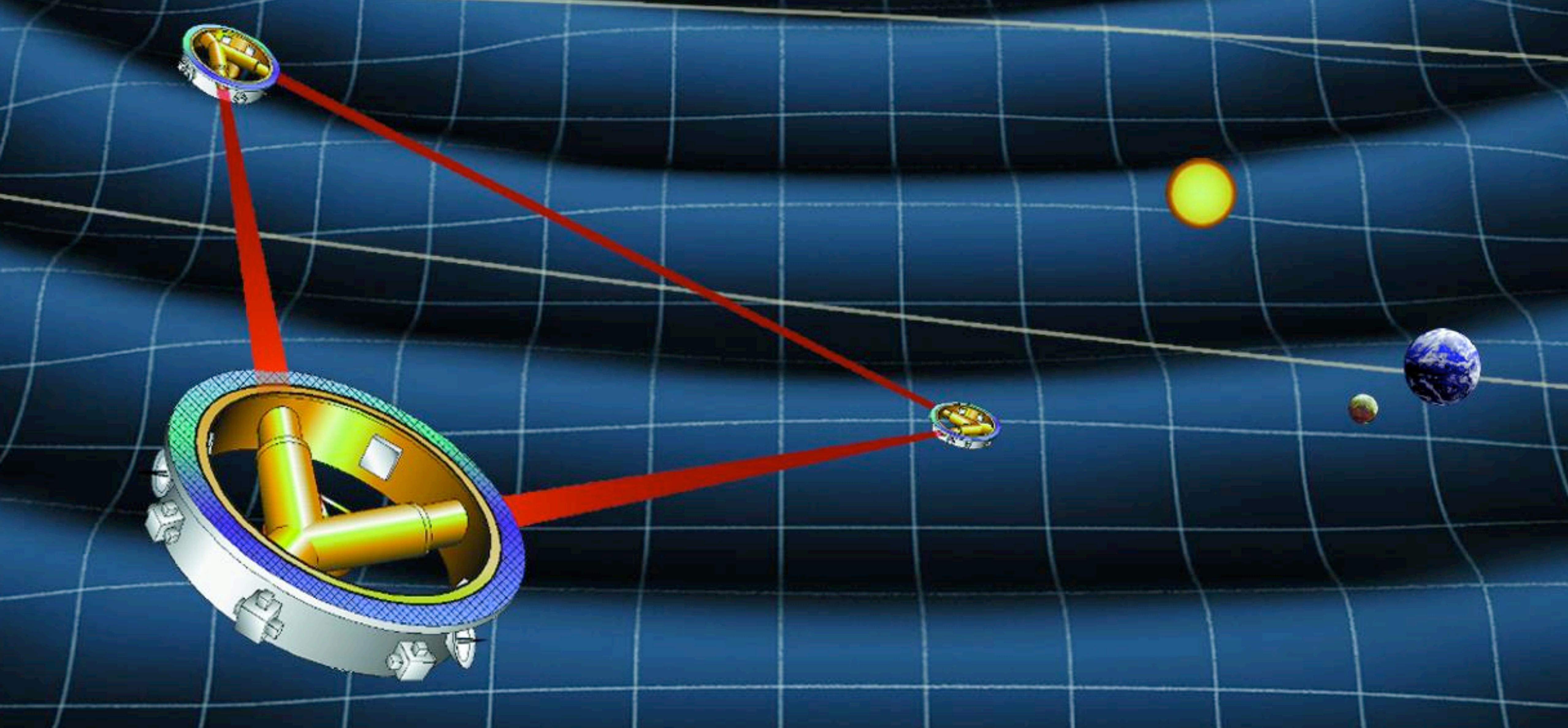
$$\vec{E}(t, \vec{x}) = \int_{-\infty}^{\infty} df \int d^2\Omega_{\hat{n}} \sum_{\alpha=1}^2 \tilde{E}_{\alpha}(f, \hat{n}) \hat{e}_{\alpha}(\hat{n}) e^{i2\pi f(t + \hat{n} \cdot \vec{x}/c)}$$

$$\hat{e}_1(\hat{n}) = \hat{\theta}, \quad \hat{e}_2(\hat{n}) = \hat{\phi}$$

etc. ...

3. What to do in the absence of correlations (e.g., for LISA)?

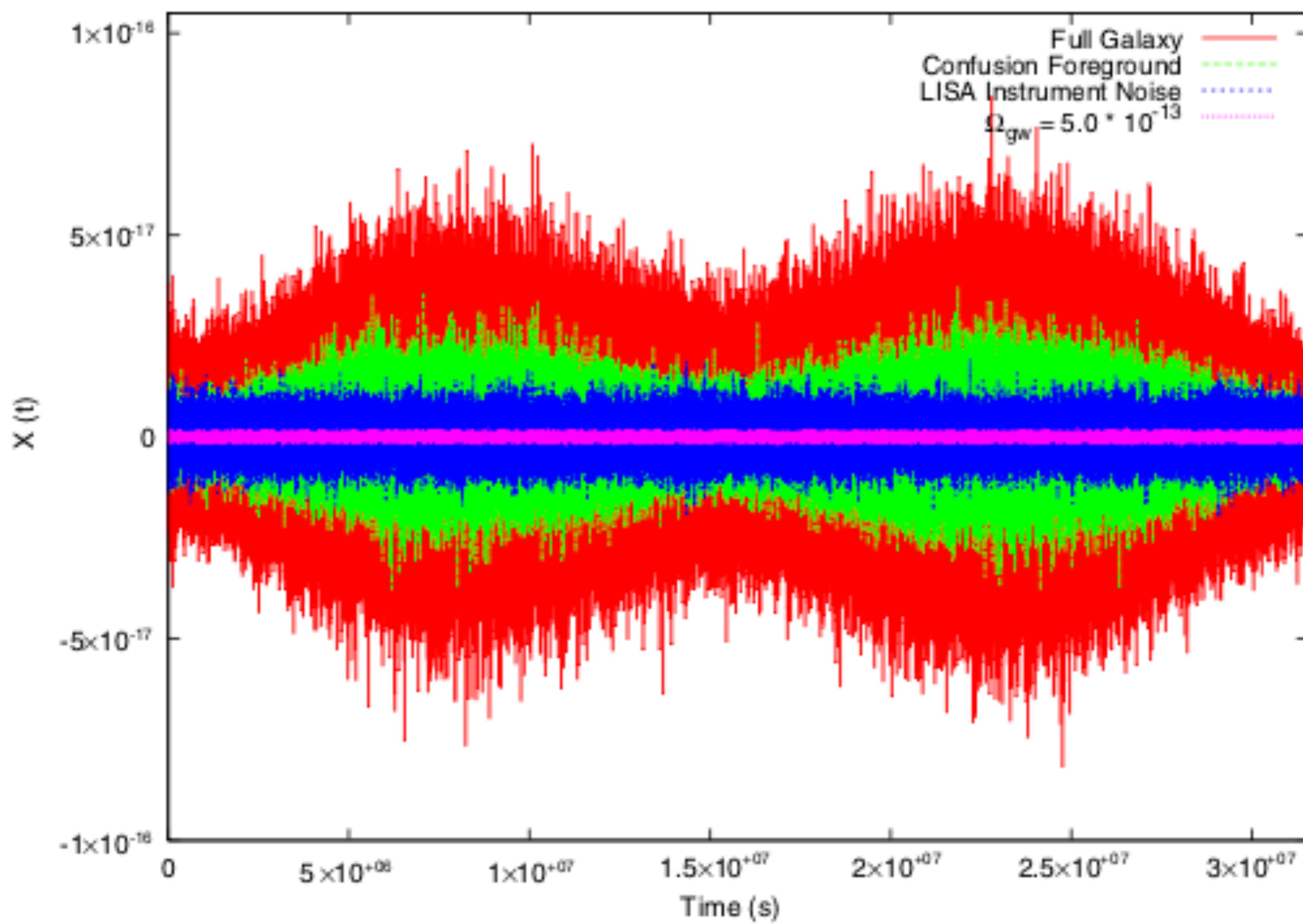
LISA (Laser Interferometer Space Antenna)

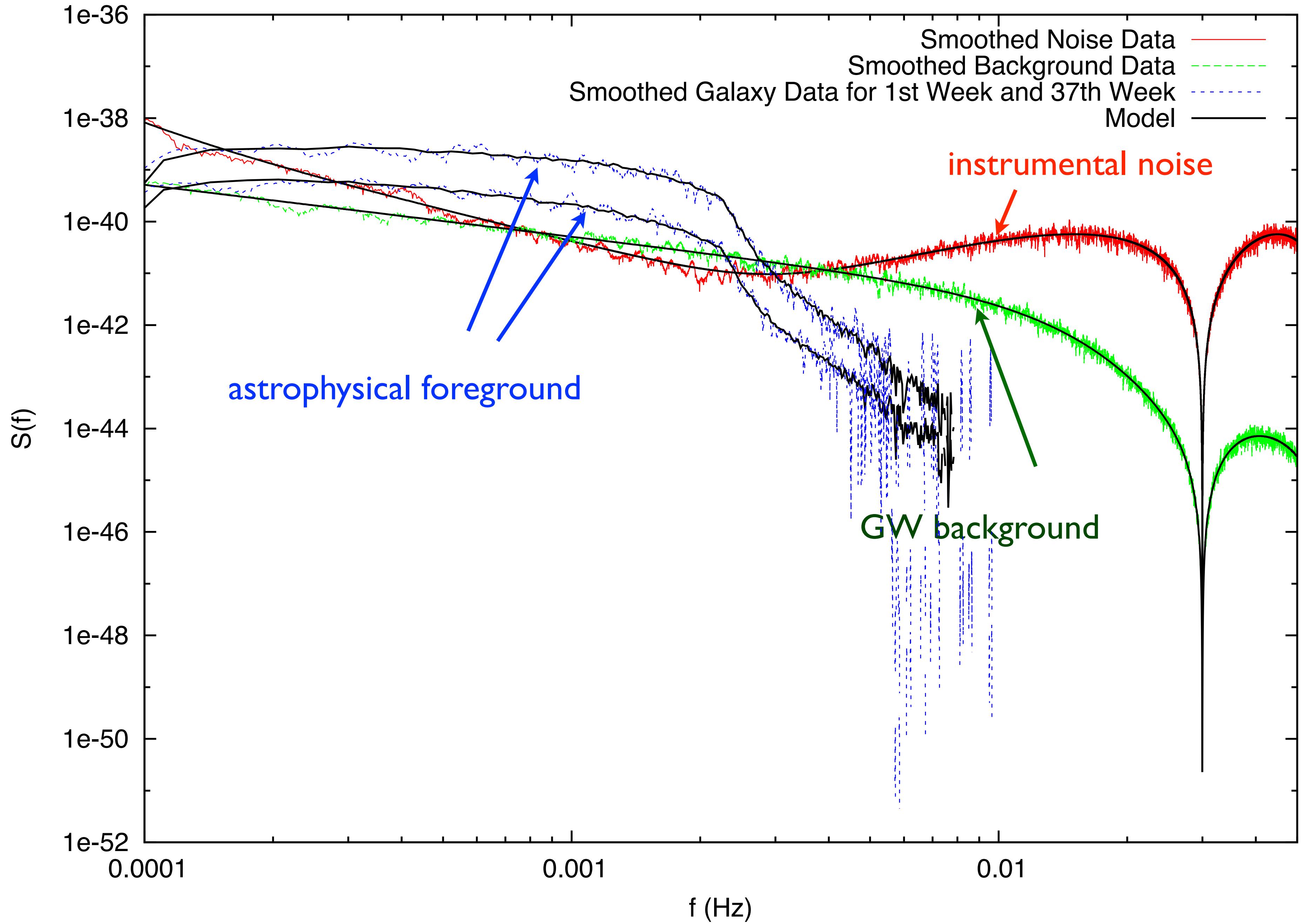


Single-detector methods for LISA

- Cross-correlation is **not an option** for LISA (at least at low frequencies)
 - although there are 3 vertices, there are only two noise-orthogonal Michelson combinations that are sensitive to GWs (so-called A, E), which turn out to be **signal orthogonal** (+, x polarizations)
 - The third noise-orthogonal channel (so-called T) is **insensitive** to GWs (**null** channel)
- Nonetheless, proper modeling of **instrumental noise**, astrophysical **foregrounds** (galactic WD binaries), and **GWB** allows you to discriminate all three components (Adams & Cornish, 2010, 2014)

Detailed questions? Ask Neil when he arrives!





4. Frequentist and Bayesian methods

Frequentist statistics	Bayesian inference
Probabilities are long-run relative occurrences of outcomes of repeatable expts —> can't be assigned to hypotheses	Probabilities are degree of belief —> can be assigned to hypotheses
Usually start with a likelihood function $p(d H)$	Same as frequentist
Construct statistics for parameter estimation / hypothesis testing	Specify priors for parameters and hypotheses
Calculate probability distribution of the statistics (e.g., using time slide)	Use Bayes' theorem to update degree of belief
Calculates confidence intervals and p-values	Construct postriors and odds ratios (Bayes factors)

Likelihood function

Starting point for both frequentist & Bayesian analyses:

$$\text{likelihood} = p(\text{data} \mid \text{parameters, model})$$

Gaussian detector noise and GWB:

$$p(d \mid C_n, \mathcal{M}_0) = \frac{1}{\sqrt{\det(2\pi C_n)}} \exp \left[-\frac{1}{2} d^T C_n^{-1} d \right] \quad (\text{noise-only model})$$

$$p(d \mid C_n, S_h, \mathcal{M}_1) = \frac{1}{\sqrt{\det(2\pi C)}} \exp \left[-\frac{1}{2} d^T C^{-1} d \right] \quad (\text{signal+noise model})$$

N samples of white noise, white GWB, in two colocated and coaligned detectors:

$$C_n = \begin{bmatrix} S_{n_1} \mathbf{1}_{N \times N} & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & S_{n_2} \mathbf{1}_{N \times N} \end{bmatrix} \quad \& \quad C = \begin{bmatrix} (S_{n_1} + S_h) \mathbf{1}_{N \times N} & S_h \mathbf{1}_{N \times N} \\ S_h \mathbf{1}_{N \times N} & (S_{n_2} + S_h) \mathbf{1}_{N \times N} \end{bmatrix}$$

Frequentist analysis

Use maximum-likelihood (ML) ratio for detection, and maximum-likelihood parameter values as estimators

Maximum-likelihood detection statistic:

$$\ln(\Lambda_{\text{ML}}(d)) \equiv \frac{\max_{S_{n_1}, S_{n_2}, S_h} p(d | S_{n_1}, S_{n_2}, S_h, \mathcal{M}_1)}{\max_{S_{n_1}, S_{n_2}} p(d | S_{n_1}, S_{n_2}, \mathcal{M}_0)}$$
$$\Lambda(d) \equiv 2 \ln(\Lambda_{\text{ML}}(d)) \simeq \frac{\hat{S}_h^2}{\hat{S}_{n_1} \hat{S}_{n_2} / N} \quad \leftarrow \text{SNR}^2$$

Maximum-likelihood estimators:

$$\hat{S}_h \equiv \frac{1}{N} \sum_{i=1}^N d_{1i} d_{2i} \quad \leftarrow \text{cross-correlation statistic}$$

$$\hat{S}_{n_1} \equiv \frac{1}{N} \sum_{i=1}^N d_{1i}^2 - \hat{S}_h$$

$$\hat{S}_{n_2} \equiv \frac{1}{N} \sum_{i=1}^N d_{2i}^2 - \hat{S}_h$$

Exercise 8: Verify the expressions for the ML estimators.

Exercise 9: Verify the expression for the detection statistic $2 \ln(\Lambda_{\text{ML}}(d))$

Bayesian analysis

Use Bayes' theorem to calculate posterior distributions for parameter estimation and odds ratios (Bayes factors) for model selection

Bayes' theorem:

$$p(H|d) = \frac{p(d|H)p(H)}{p(d)}$$

likelihood

prior

normalization factor

Posteriors:

$$p(S_{n_1}, S_{n_2}, S_h | d, \mathcal{M}_1) = \frac{p(d | S_{n_1}, S_{n_2}, S_h, \mathcal{M}_1)p(S_{n_1}, S_{n_2}, S_h | \mathcal{M}_1)}{p(d | \mathcal{M}_1)}$$

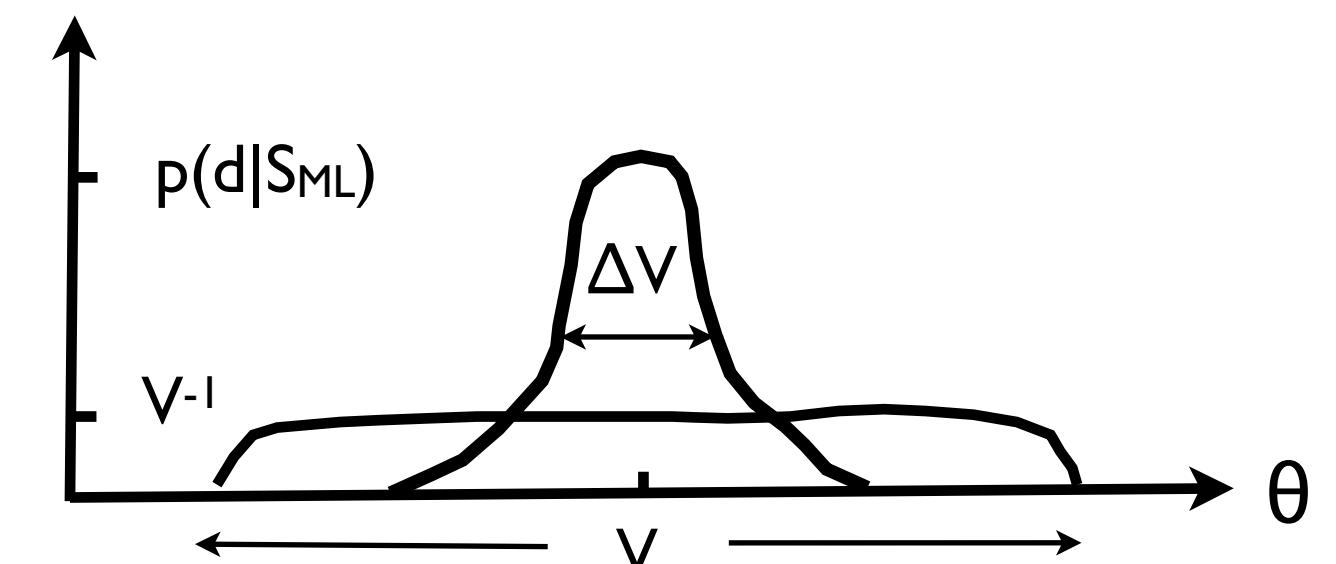
Model selection:

$$\frac{p(\mathcal{M}_1 | d)}{p(\mathcal{M}_0 | d)} = \frac{p(d | \mathcal{M}_1)p(\mathcal{M}_1)}{p(d | \mathcal{M}_0)p(\mathcal{M}_0)}$$

Bayes factor $\mathcal{B}_{10}(d)$

Relationship to frequentist approach:

$$\mathcal{B}_{10}(d) \equiv \frac{p(d | \mathcal{M}_1)}{p(d | \mathcal{M}_0)} = \frac{\int dS_{n_1} \int dS_{n_2} \int dS_h p(d | S_{n_1}, S_{n_2}, S_h, \mathcal{M}_1) p(S_{n_1}, S_{n_2}, S_h | \mathcal{M}_1)}{\int dS_{n_1} \int dS_{n_2} p(d | S_{n_1}, S_{n_2} | d, \mathcal{M}_0) p(S_{n_1}, S_{n_2} | \mathcal{M}_0)} \simeq \Lambda_{\text{ML}}(d) \frac{\Delta V_1 / V_1}{\Delta V_0 / V_0}$$



Derivation of standard stochastic likelihood using “stochastic templates”

Generic likelihood:

$$p(d | C_n, h) \equiv p_n(d - h | C_n) = \frac{1}{\sqrt{\det(2\pi C_n)}} \exp \left[-\frac{1}{2} (d - h)^T C_n^{-1} (d - h) \right]$$

↑
signal model

↑
covariance matrix for noise, e.g., $C_n = \begin{bmatrix} S_{n_1} & 0 \\ 0 & S_{n_2} \end{bmatrix}$

stochastic signal model:

$$p(h | S_h) = \frac{1}{\sqrt{2\pi S_h}} \exp \left[-\frac{1}{2} \frac{h^2}{S_h} \right]$$

Marginalized likelihood:

$$p(d | C_n, S_h) = \int dh p_n(d - h | C_n) p(h | S_h) = \frac{1}{\sqrt{\det(2\pi C)}} \exp \left[-\frac{1}{2} d^T C^{-1} d \right]$$

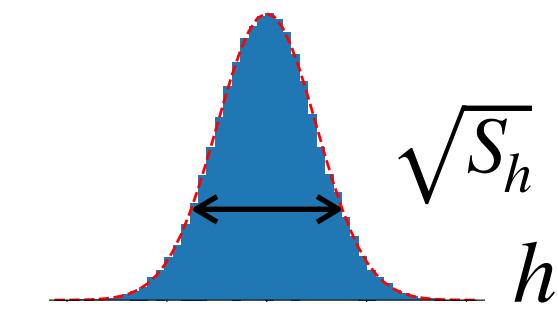
covariance matrix
for signal + noise $\longrightarrow C = \begin{bmatrix} S_{n_1} + S_h & S_h \\ S_h & S_{n_2} + S_h \end{bmatrix}$

Exercise 10: Do the marginalization over h to obtain this final result.

Signal priors define the signal model...

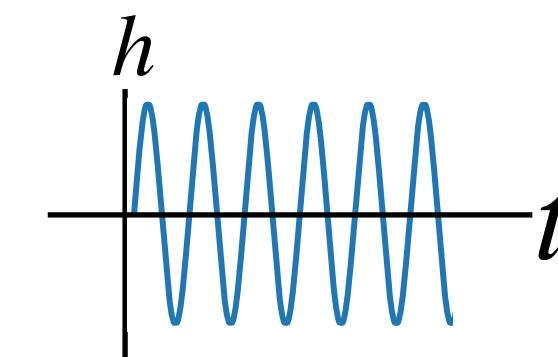
stochastic:

$$p(h | S_h) = \frac{1}{\sqrt{2\pi S_h}} \exp \left[-\frac{1}{2} \frac{h^2}{S_h} \right]$$



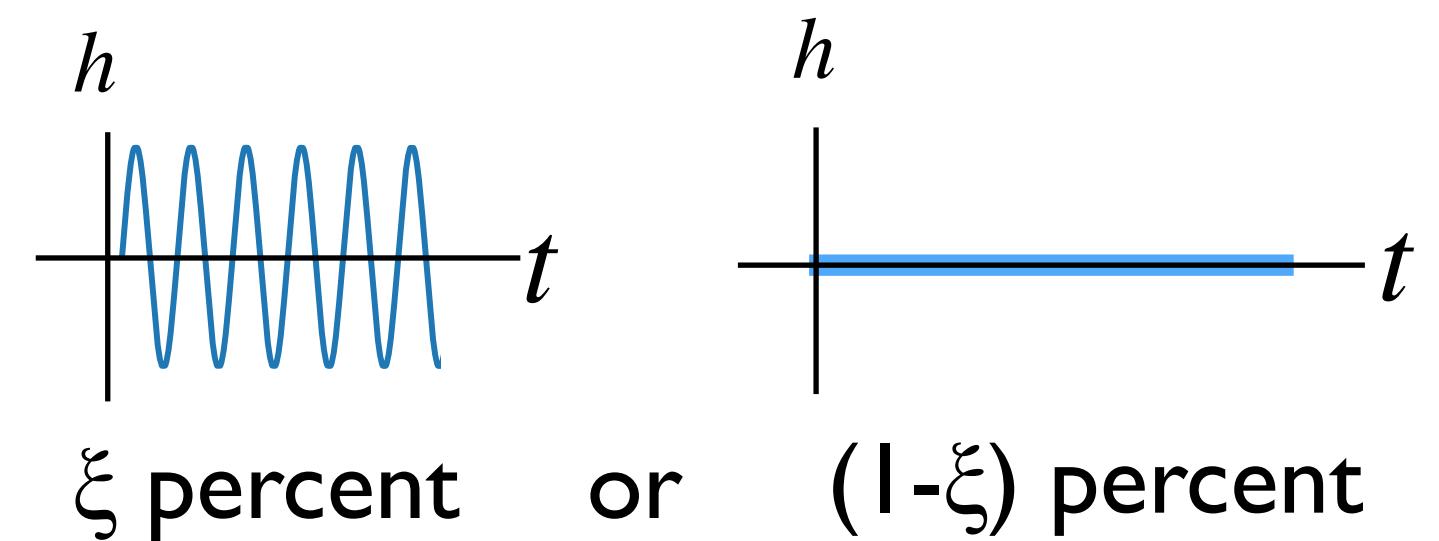
deterministic:

$$p(h | A, t_0, f_0) = \delta(h - A \sin[2\pi f_0(t - t_0)])$$



hybrid:

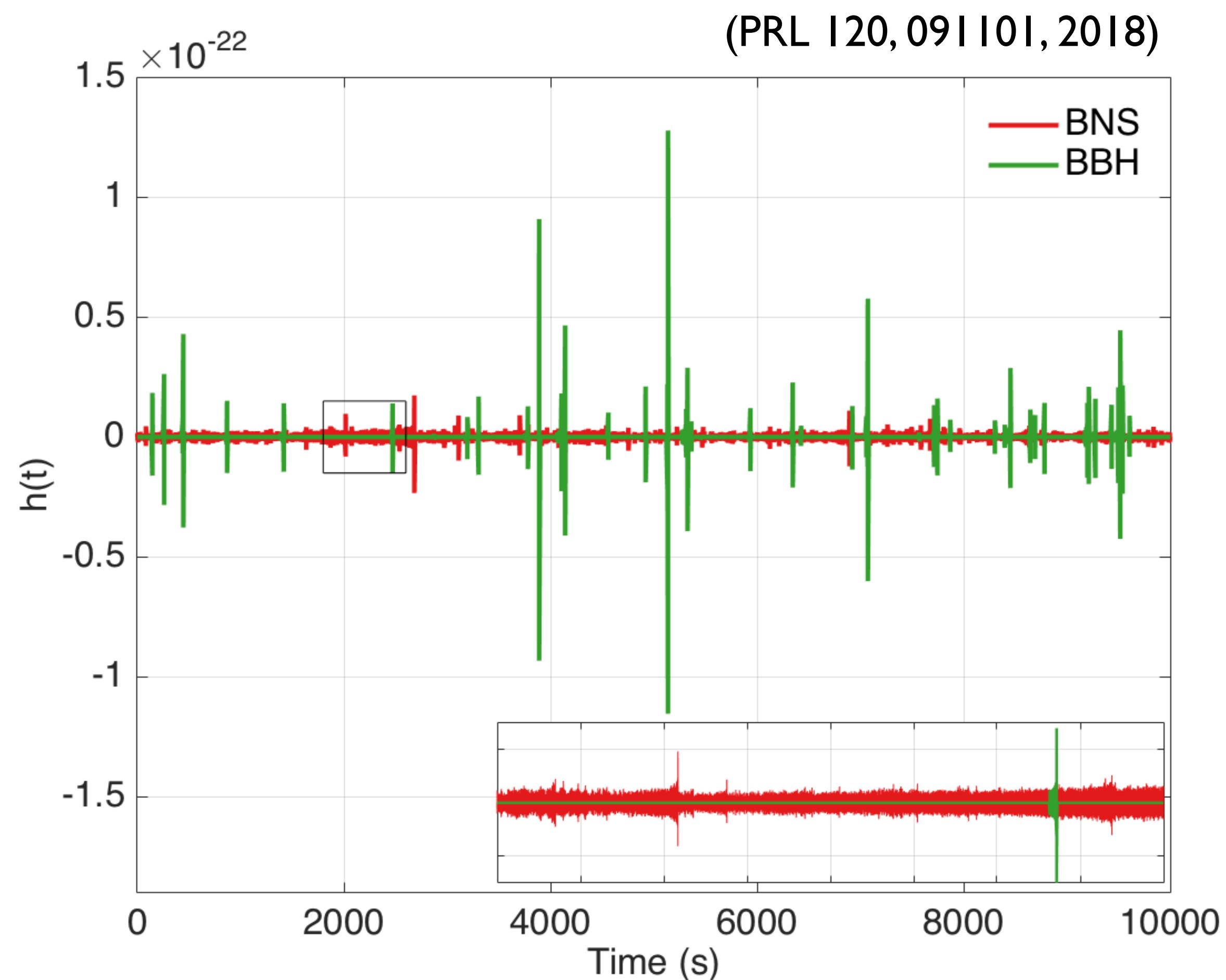
$$p(h | \xi, A, t_0, f_0) = \xi \delta(h - A \sin[2\pi f_0(t - t_0)]) + (1 - \xi) \delta(h)$$



5. Example: searching for the background from BBH mergers

RECALL: Non-stationary background from BBH mergers is a potential signal for advanced LIGO,Virgo

- Recent detections of BBH and BNS mergers by advanced LIGO, Virgo imply the existence of a stochastic background of weaker events
- Smith & Thrane (PRX 8, 021019,2018) have proposed an alternative method to search for the BBH component, optimally suited for the non-stationarity
- Describe BBH background with a hybrid signal model
- Average over chirp parameters to infer only rate of mergers
- Use two detectors to discriminate against glitches



Mathematical details

Likelihood:

$$p(d | C_n, h) \equiv p_n(d - h | C_n)$$

Hybrid signal model:

$$p(h | \xi, \vec{\lambda}) = \xi \delta\left(h - \text{chirp}(\vec{\lambda})\right) + (1 - \xi) \delta(h)$$

Marginalized likelihoods:

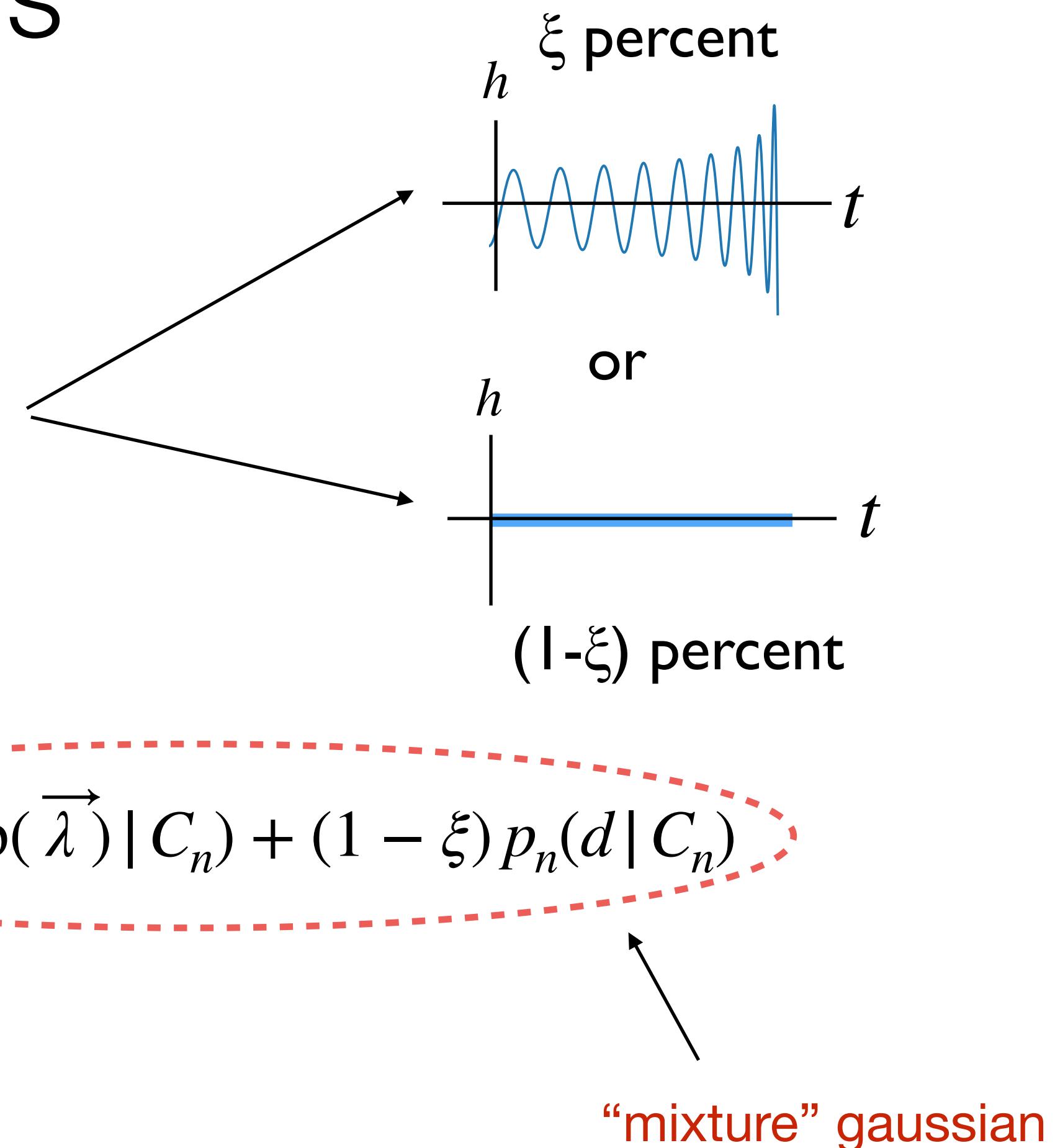
$$p(d | \xi, \vec{\lambda}) = \int dh p(d | C_n, h) p(h | \xi, \vec{\lambda}) = \xi p_n(d - \text{chirp}(\vec{\lambda}) | C_n) + (1 - \xi) p_n(d | C_n)$$

$$p(d | \xi) = \int d \vec{\lambda} p(d | \xi, \vec{\lambda}) p(\vec{\lambda}) = (S - N)\xi + N$$

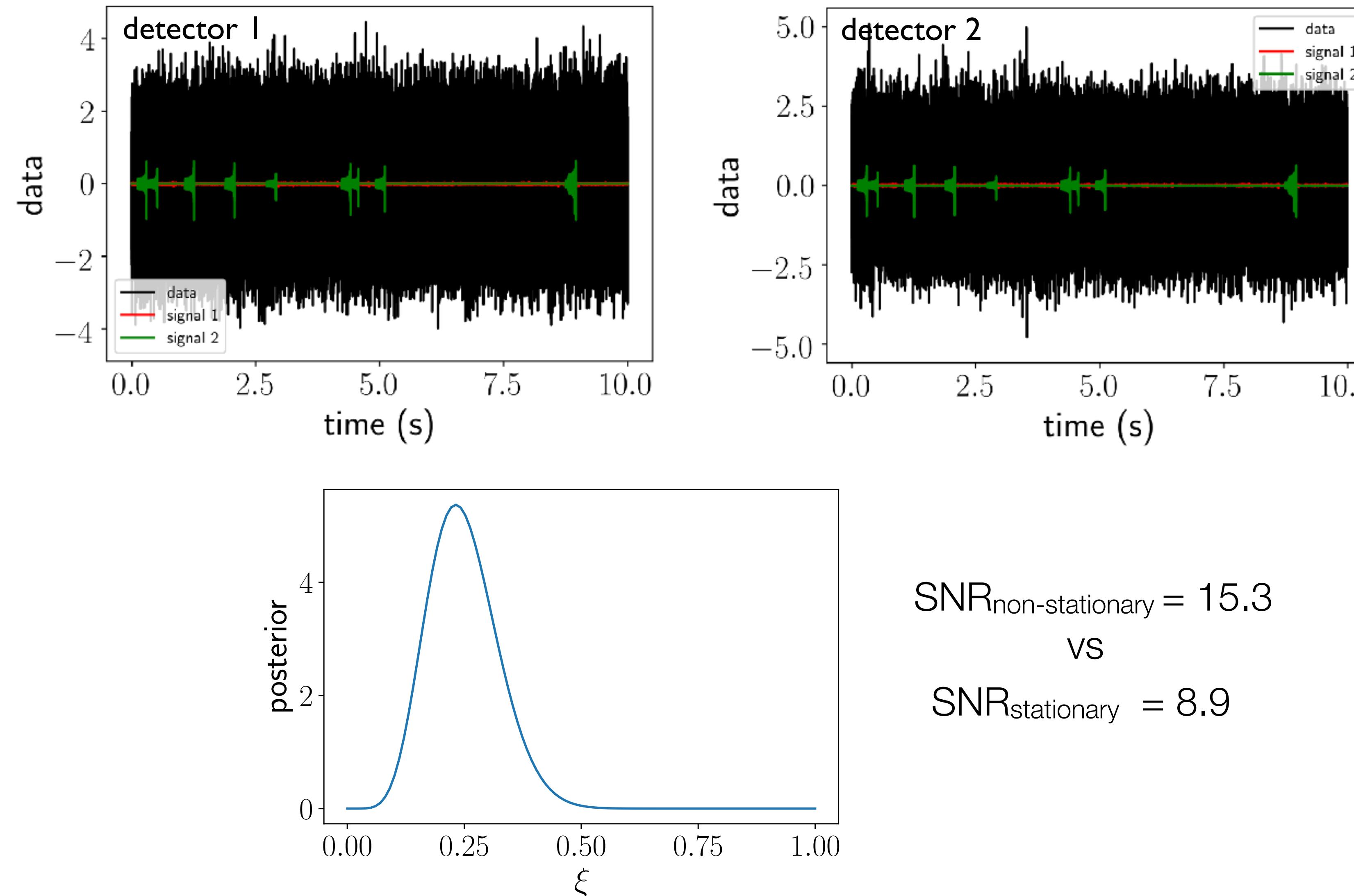
Posterior:

$$p(\xi | d) = \frac{p(d | \xi) p(\xi)}{p(d)}$$

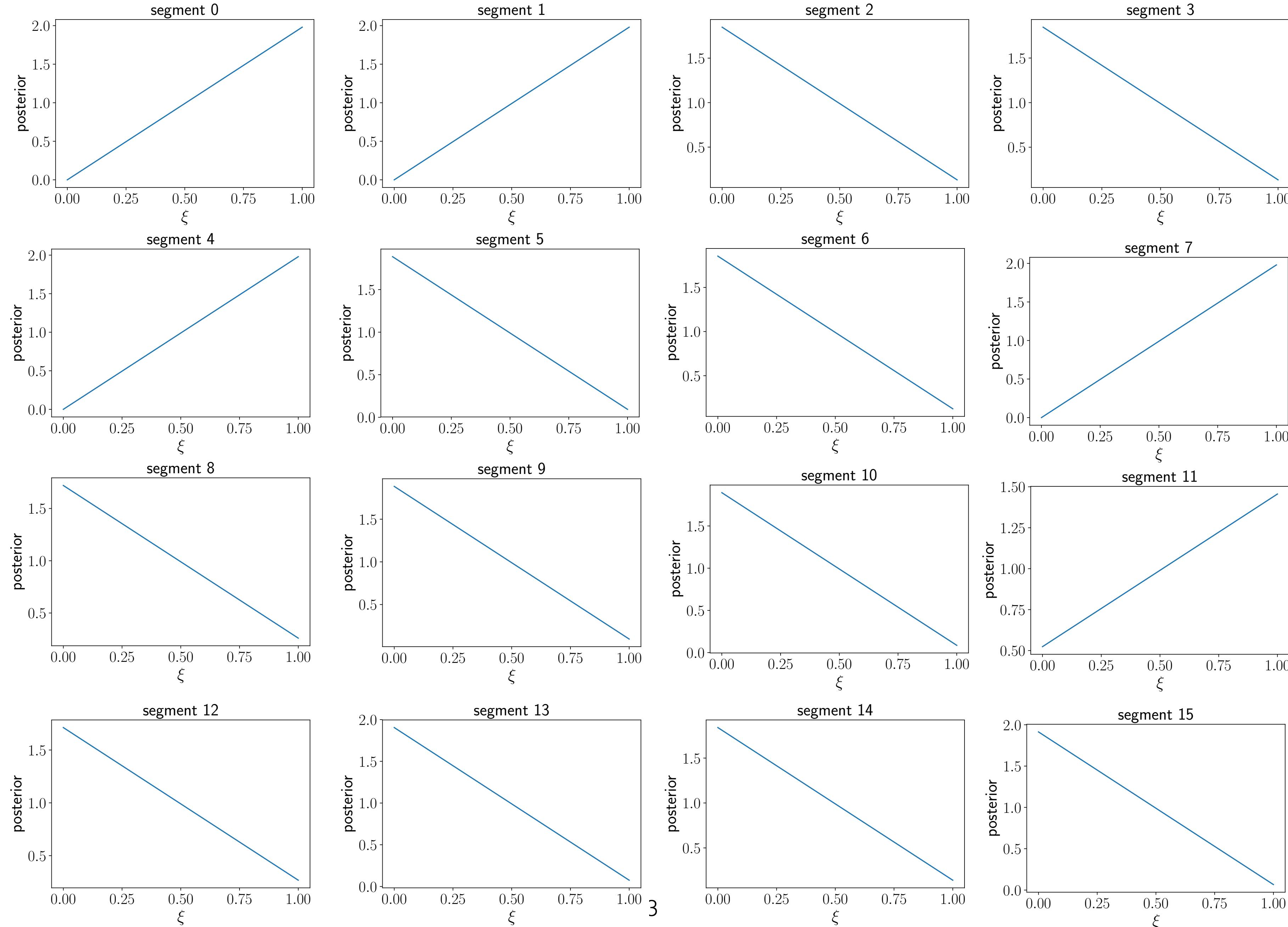
Combine segments by multiplying likelihoods, ...



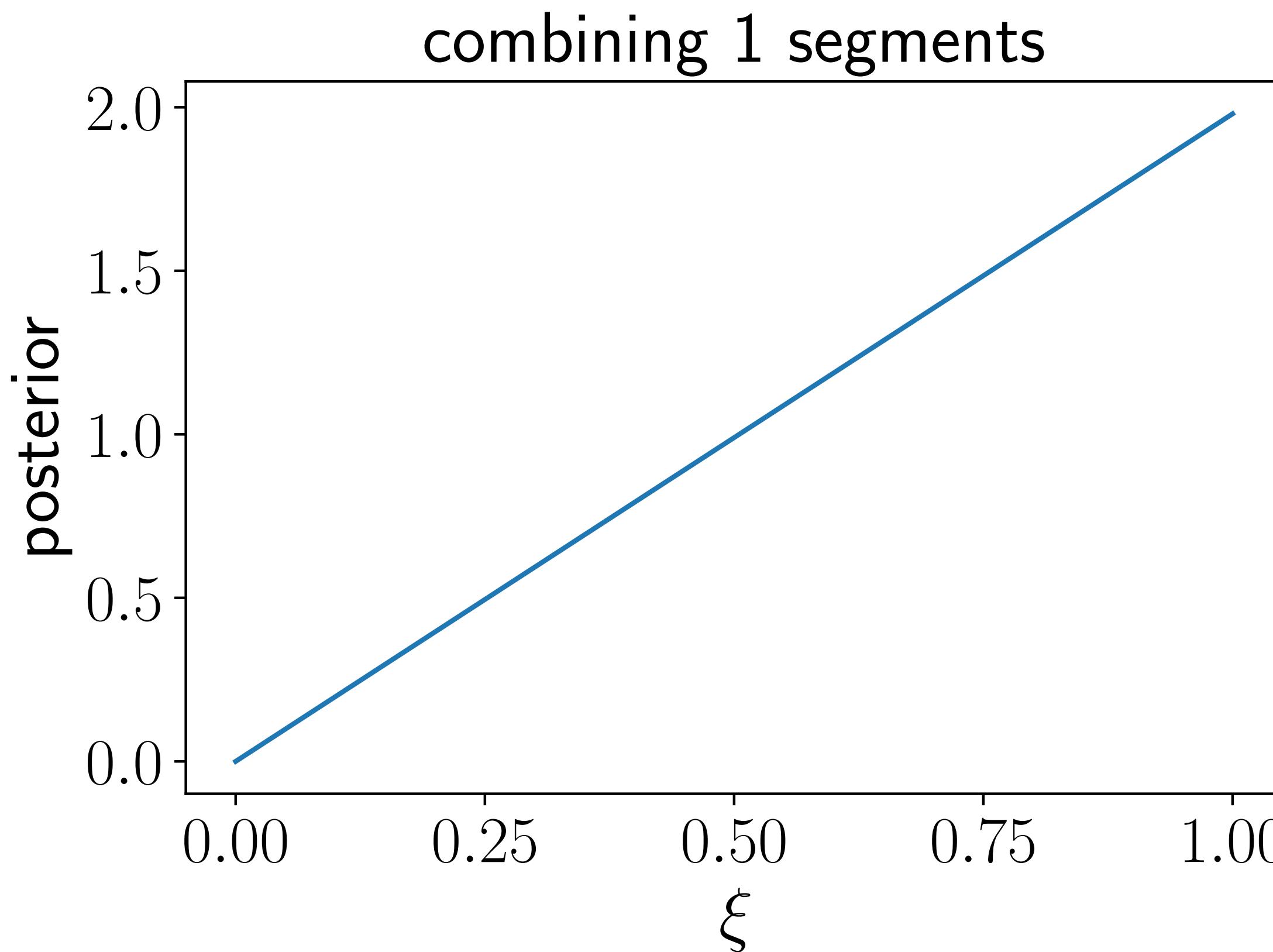
Example: Simulated BBH background in white detector noise and confusion-limited BNS background



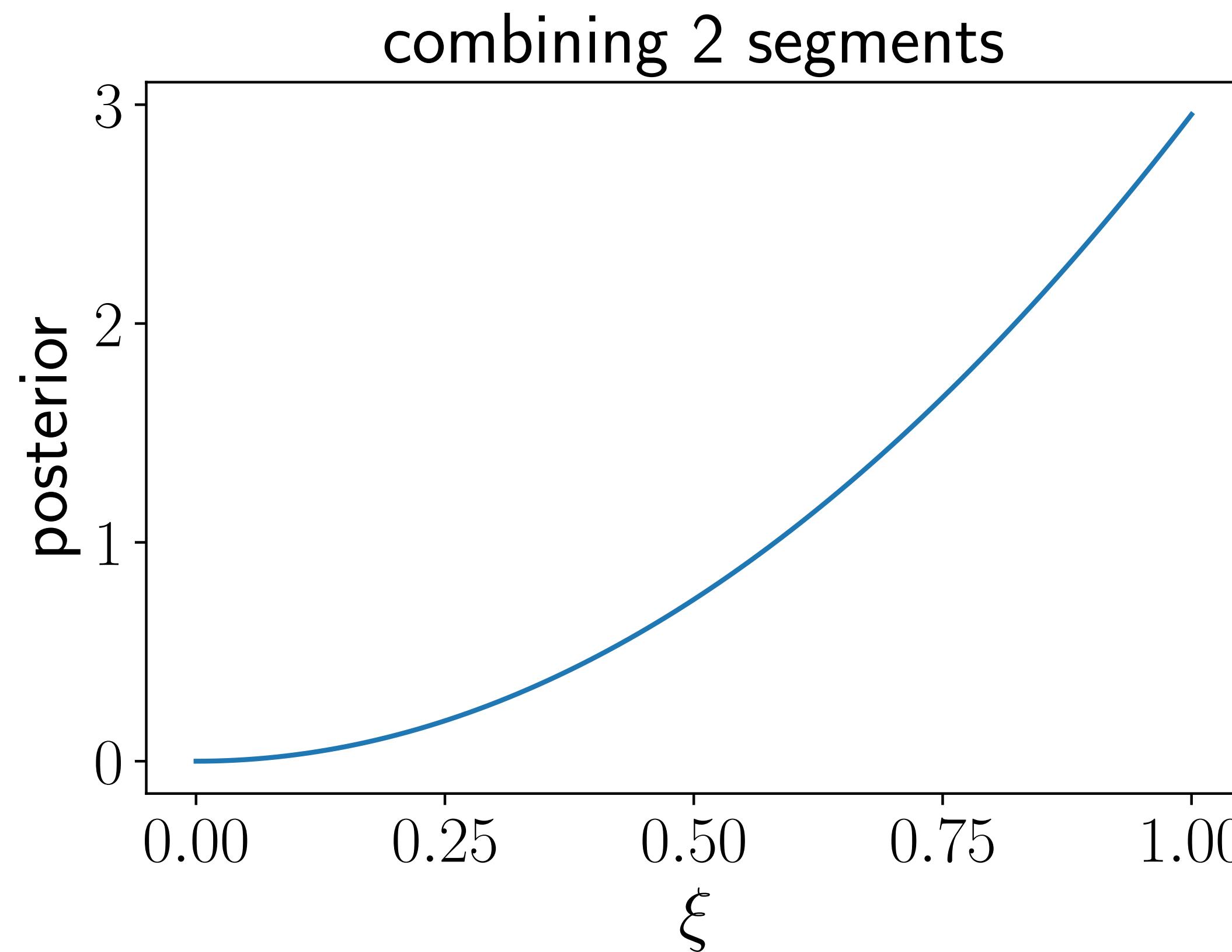
Posteriors $p(\xi|d)$ for individual segments



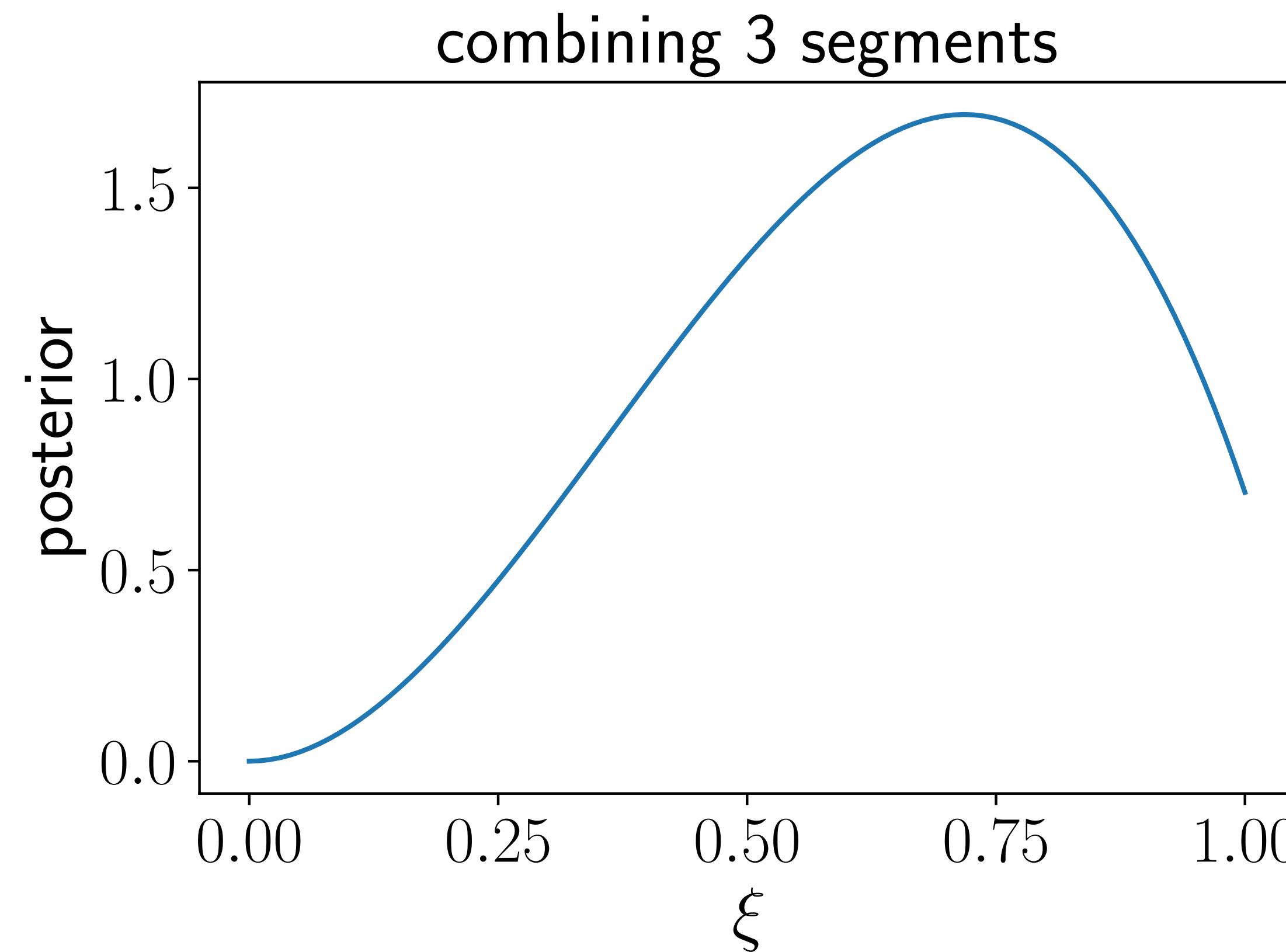
Combining segment posteriors



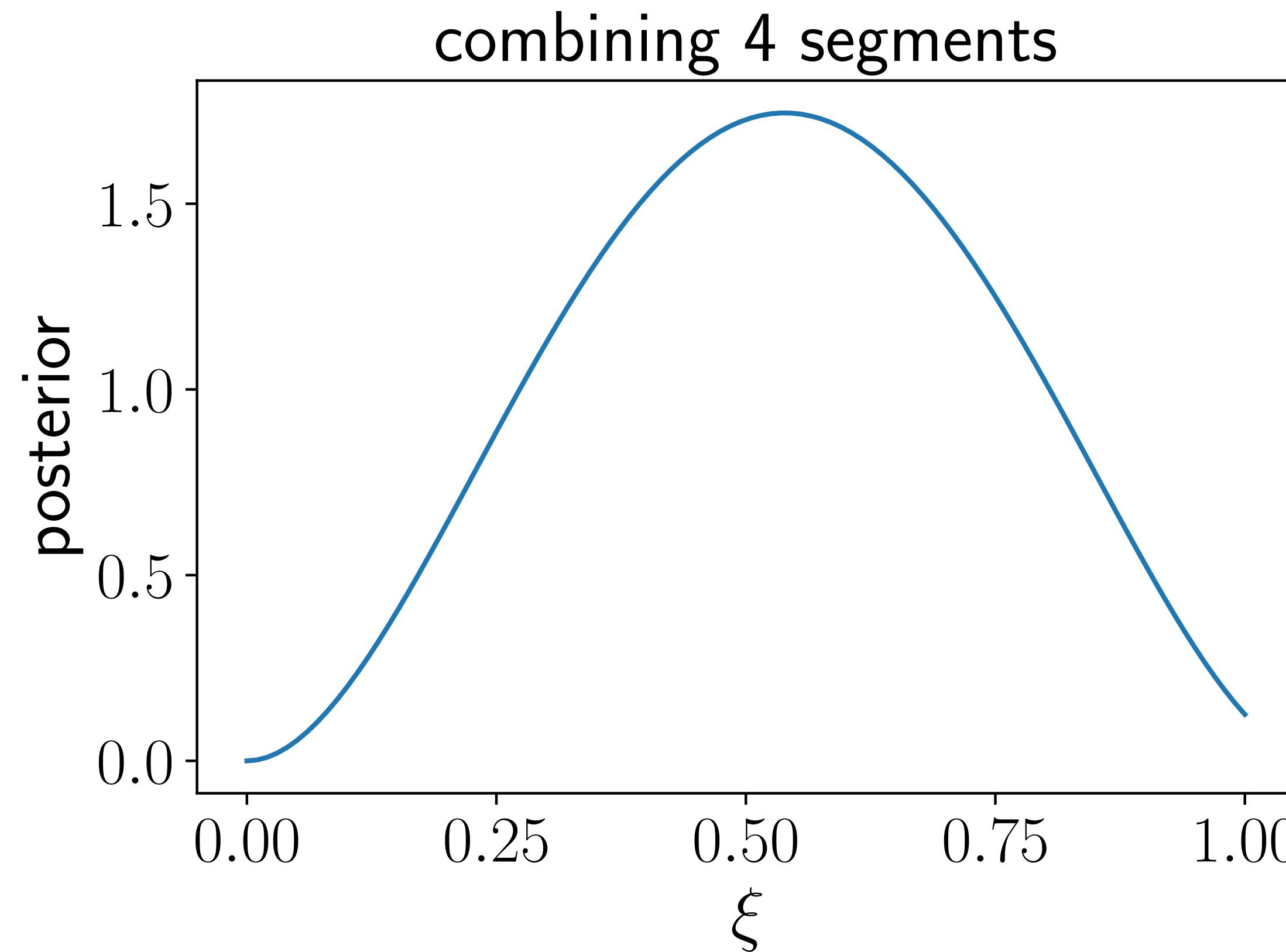
Combining segment posteriors



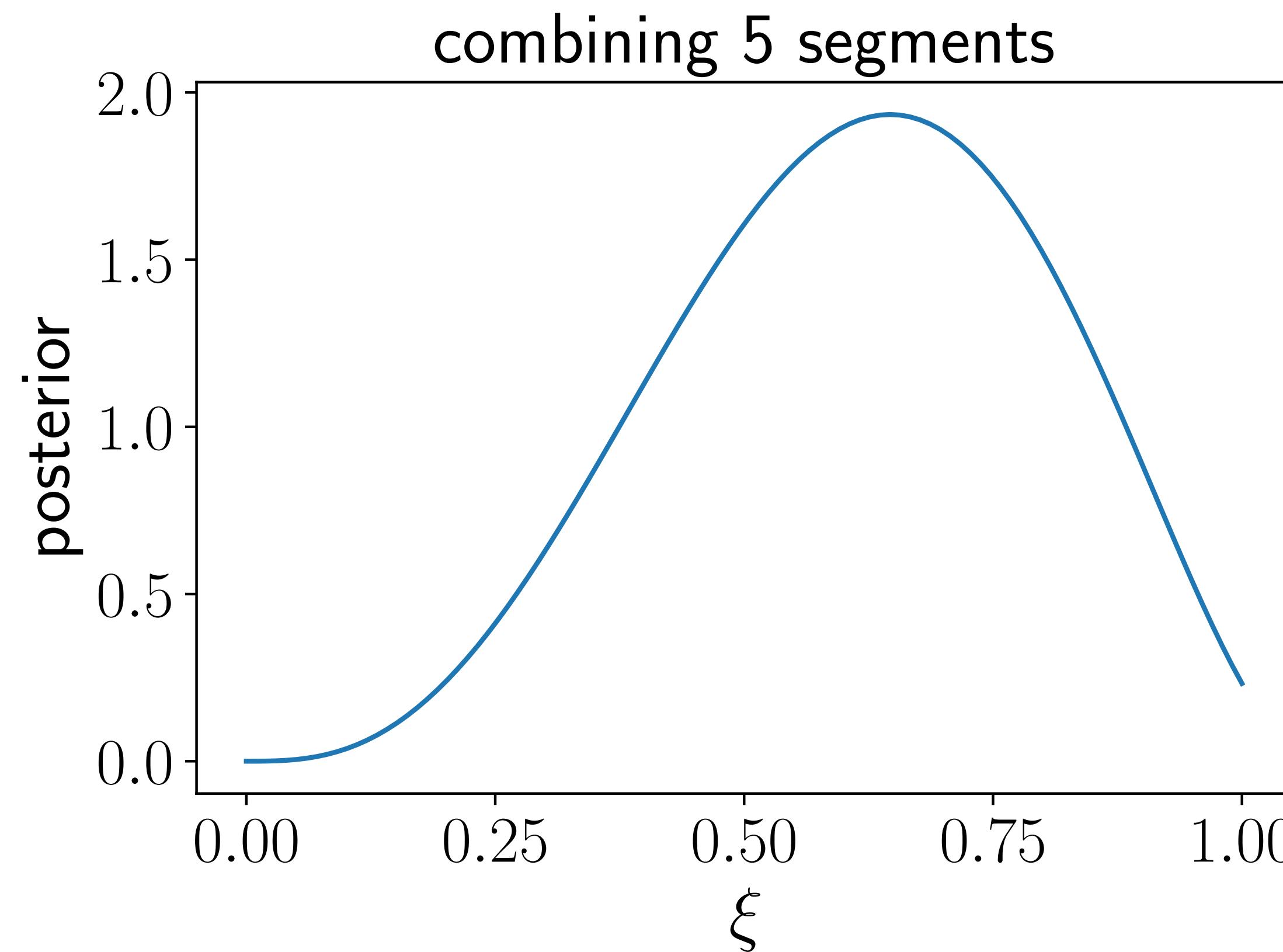
Combining segment posteriors



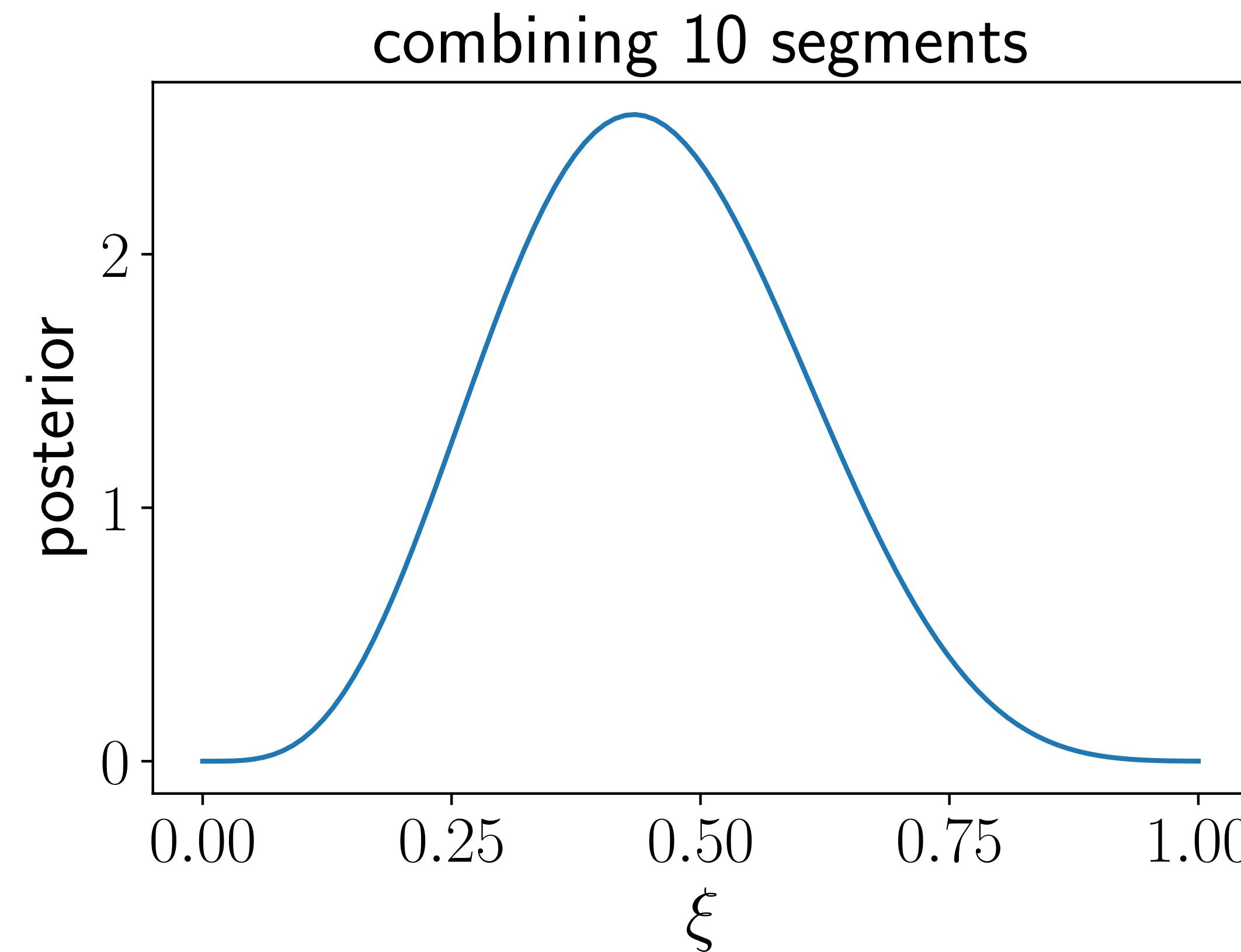
Combining segment posteriors



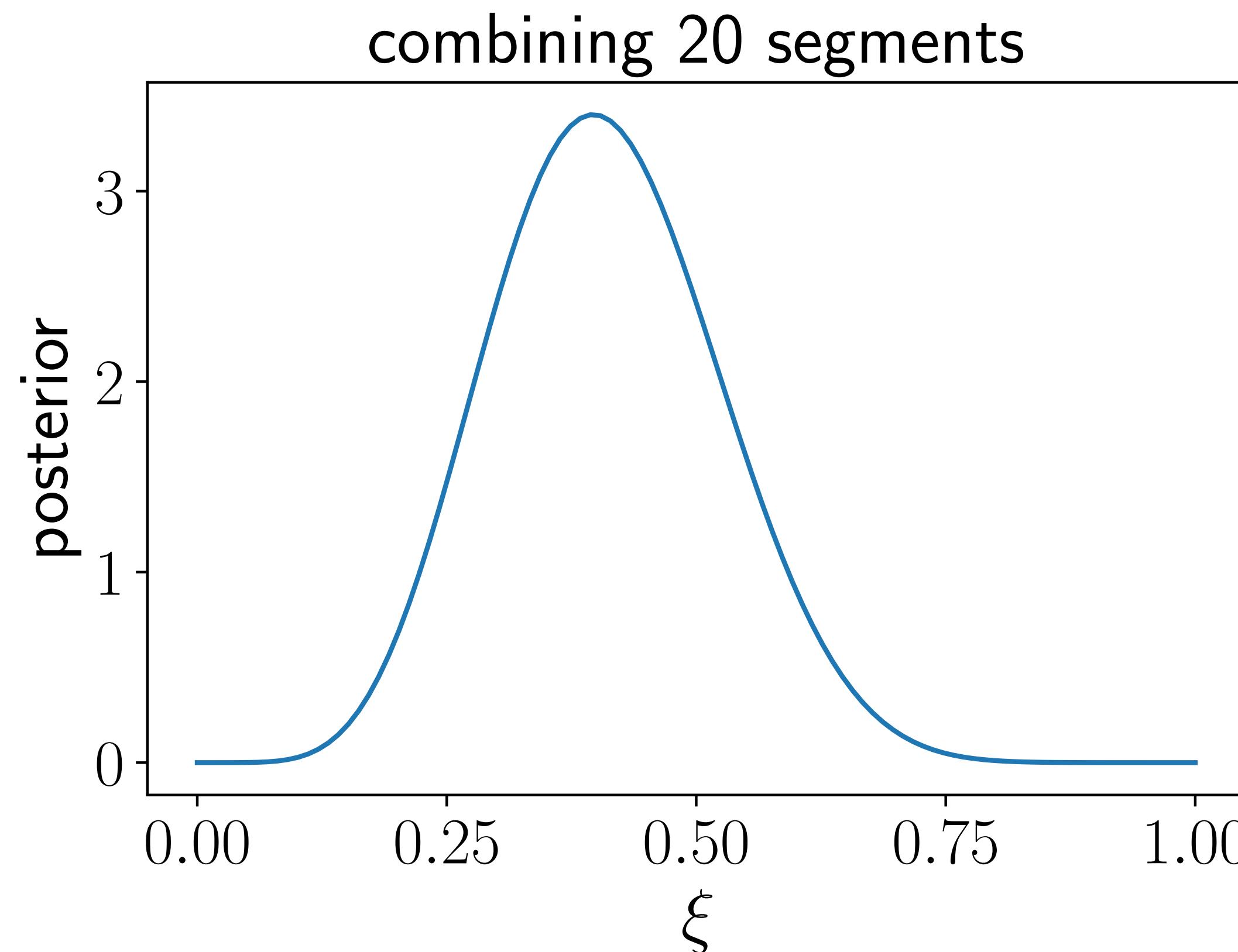
Combining segment posteriors



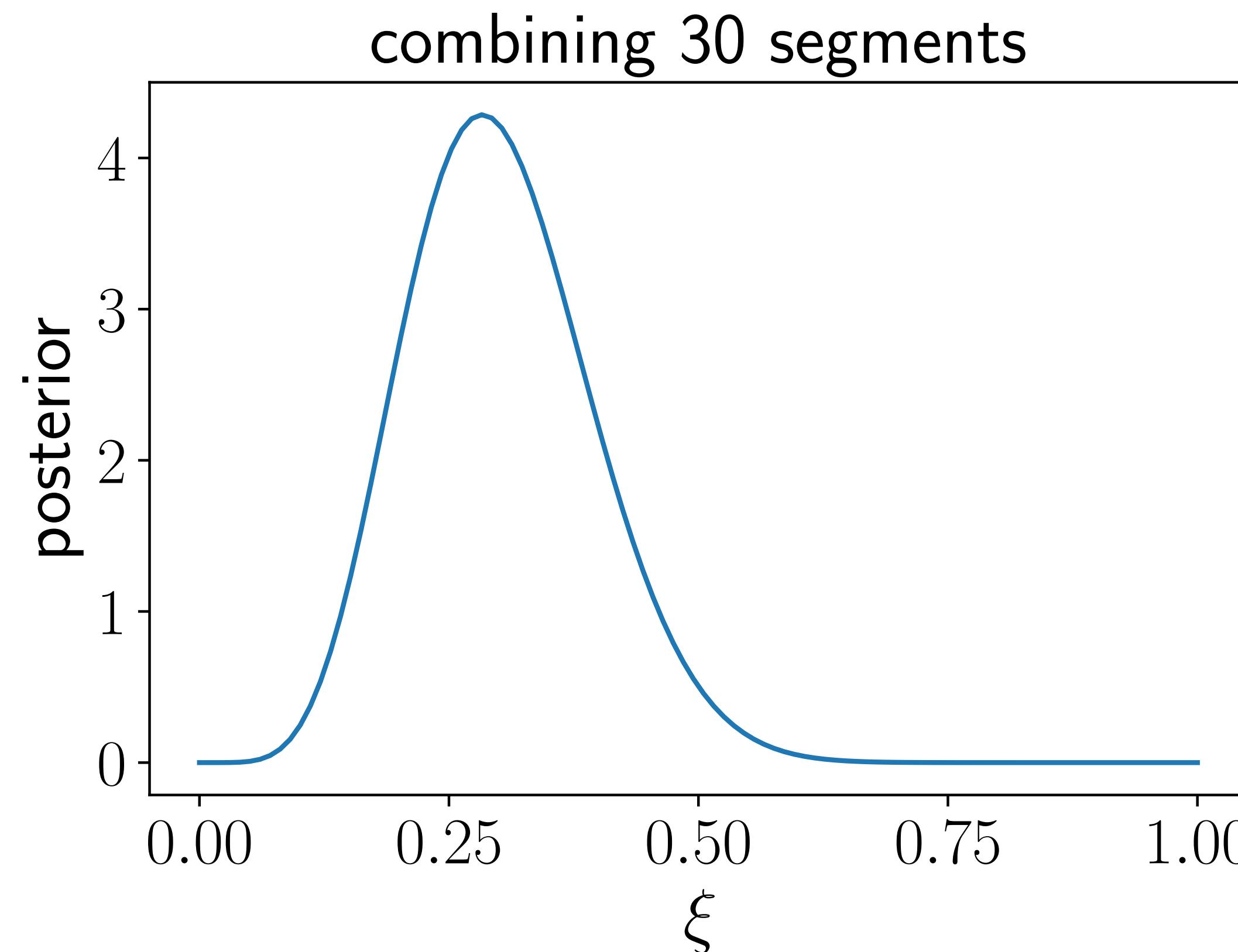
Combining segment posteriors



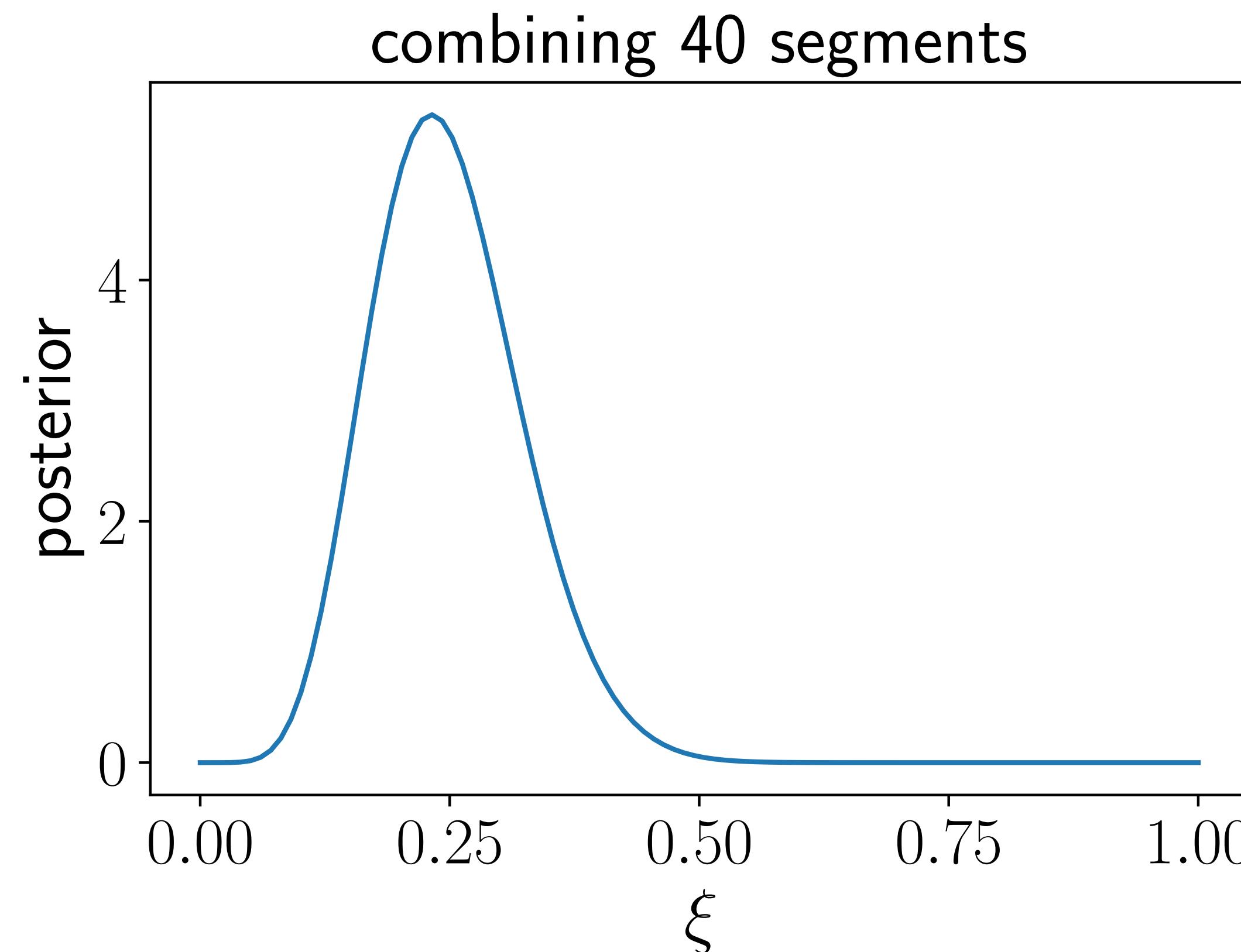
Combining segment posteriors



Combining segment posteriors



Combining segment posteriors



The optimal analysis reduces time to detection because...

- All segments contribute to estimating signal parameter ξ
- BBH chirp signal is deterministic and not stochastic

$$\frac{\text{SNR}_{\text{non-stationary}}}{\text{SNR}_{\text{stationary}}} \sim \sqrt{\frac{N_{\text{cycles}}}{\xi}}$$

haven't been able to rigorously prove the N_{cycles} part!!

~40 months of observation reduces to ~1 day!!

So stay tuned!!

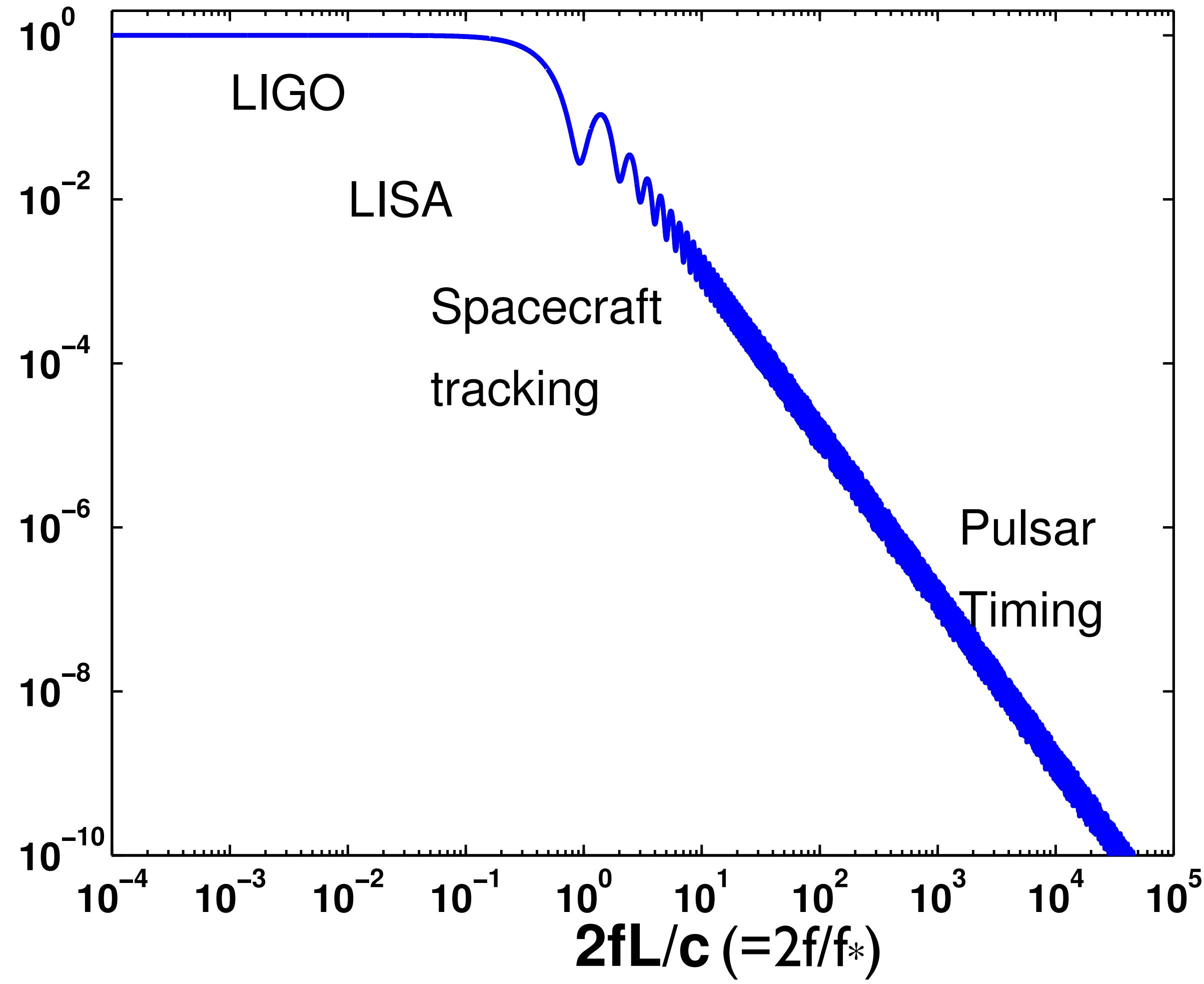
extra slides

Summary

- A stochastic background / foreground is the superposition of signals that are either **too weak** or **too numerous** to individually detect
- **Correlation methods** are typically used to search for such signals, but spectral and temporal differences between the detector noise and GWB allow separation even in one detector (e.g., LISA)
- The collection of sub-threshold signals from BBH mergers in the LIGO band gives rise to a **non-stationary** background
- A new approach targeting non-stationary signals has a **good chance of detecting** this background before the end of LIGO's third observing run

So stay tuned!!

Frequentist statistics	Bayesian inference
Probabilities are long-run relative occurrences of outcomes of repeatable expts; cannot be assigned to hypotheses, which have fixed but unknown values	Probabilities are degree of belief , so they can be assigned to hypotheses
Assumes measured data drawn from a probability distribution, which assumes the truth of a particular hypothesis / model (likelihood function)	Same
Constructs a statistic to estimate parameters or test a hypothesis / model	Needs to specify priors for parameters and hypotheses
Calculates probability distribution of this statistic (sampling distribution)	Uses Bayes' theorem to update prior degree of belief in light of new data
Calculates confidence intervals and p-values for parameter estimation and hypothesis testing	Constructs posteriors and odds ratios (Bayes factors) for parameter estimation and hypothesis testing



Beam detector	L (km)	f_* (Hz)	f (Hz)	f/f_*	Relation
Ground-based interferometer	~ 1	$\sim 10^5$	$10 - 10^4$	$10^{-4} - 10^{-1}$	$f \ll f_*$
Space-based interferometer	$\sim 10^6$	$\sim 10^{-1}$	$10^{-4} - 10^{-1}$	$10^{-3} - 1$	$f \lesssim f_*$
Spacecraft Doppler tracking	$\sim 10^9$	$\sim 10^{-4}$	$10^{-6} - 10^{-3}$	$10^{-2} - 10$	$f \sim f_*$
Pulsar timing	$\sim 10^{17}$	$\sim 10^{-12}$	$10^{-9} - 10^{-7}$	$10^3 - 10^5$	$f \gg f_*$

$\mathcal{B}_{\alpha\beta}(d)$	$2 \ln \mathcal{B}_{\alpha\beta}(d)$	Evidence for model \mathcal{M}_α relative to \mathcal{M}_β
<1	<0	Negative (supports model \mathcal{M}_β)
1–3	0–2	Not worth more than a bare mention
3–20	2–6	Positive
20–150	6–10	Strong
>150	>10	Very strong

Adapted from [Kass and Raftery \(1995\)](#)

Matched-filtering determination of measured TOAs

$$C(\Delta t) = \mathcal{N} \int dt y(t)p(t - \Delta t)$$

