Assessing the impact of gaps in LISA data

Sylvain Marsat (L2IT, Toulouse)

Ollie Burke (L2IT, Toulouse)

Jonathan Gair (AEI, Potsdam, DE)

Outline

- Introducing gaps for the LISA data
- Theory: stochastic noise process, covariance and mismodeling error
- Analysis settings: LISA TDI data, short MBHB signals
- Results: mismodelling errors
- Conclusion

Gaps in LISA data

- Gaps are missing data periods; can be **planned** and **unplanned**
- The stochastic noise process could be **coherent** (masked data) or **incoherent** (instrument reset)
- Masking out data is also a crude way of addressing glitches

Gap type	Frequency	Duration	Total loss (hr/yr)
Antenna repointing	every 2 weeks	3.3h	1%
PAAM angle adjust	3 per day	100s	0.3%
TM stray pot. est.	2/yr	1 day	0.56%
TTL coupling est.	4/yr	2 days	2.22%
Unplanned: platform	3/yr	2.5 days	2%
Unplanned: payload	4/yr	2.75 days	3%
Unplanned: micro-meteorites	30/yr	1 day	8%

[A. Petiteau, FMT]

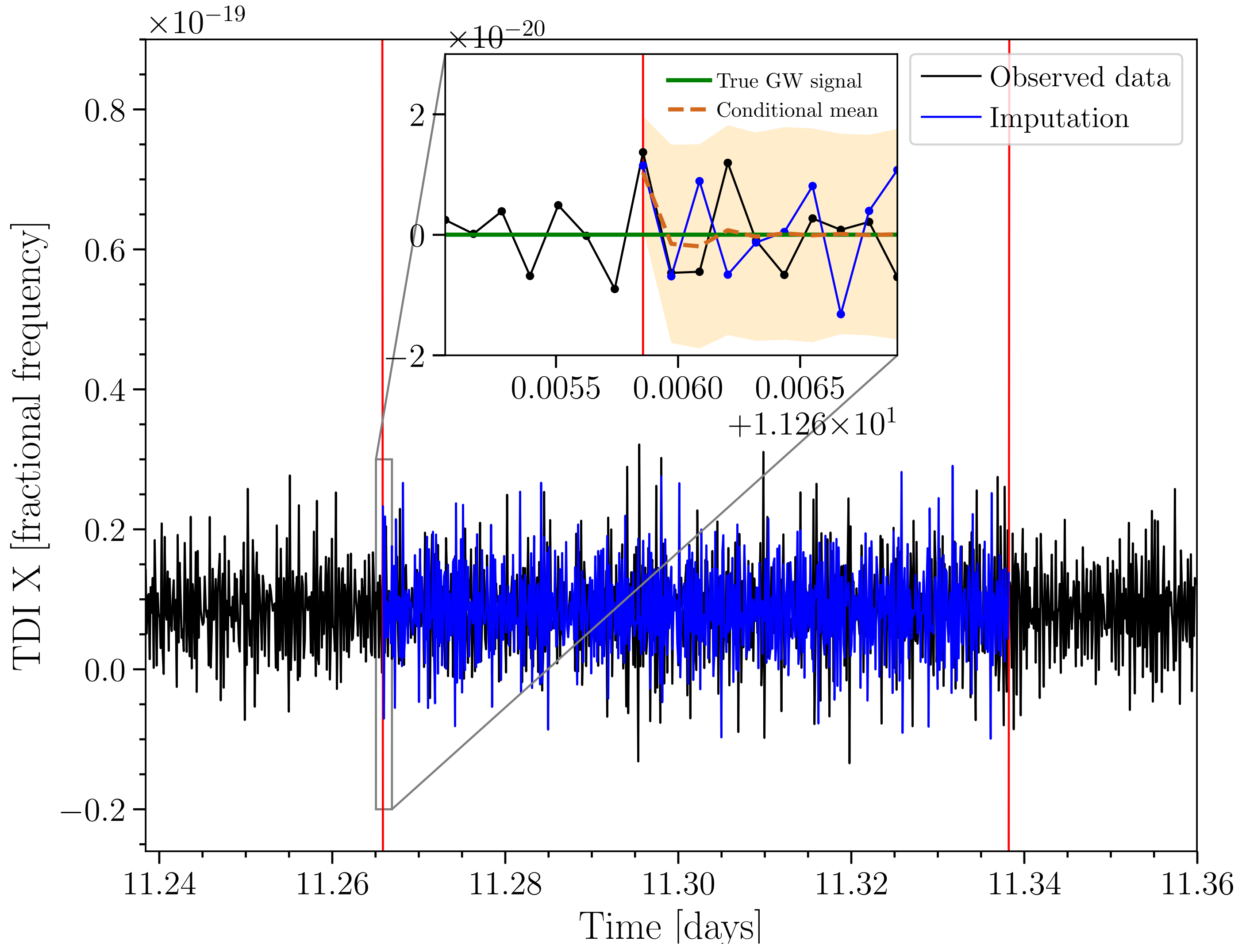
Gaps in LISA data: different approaches

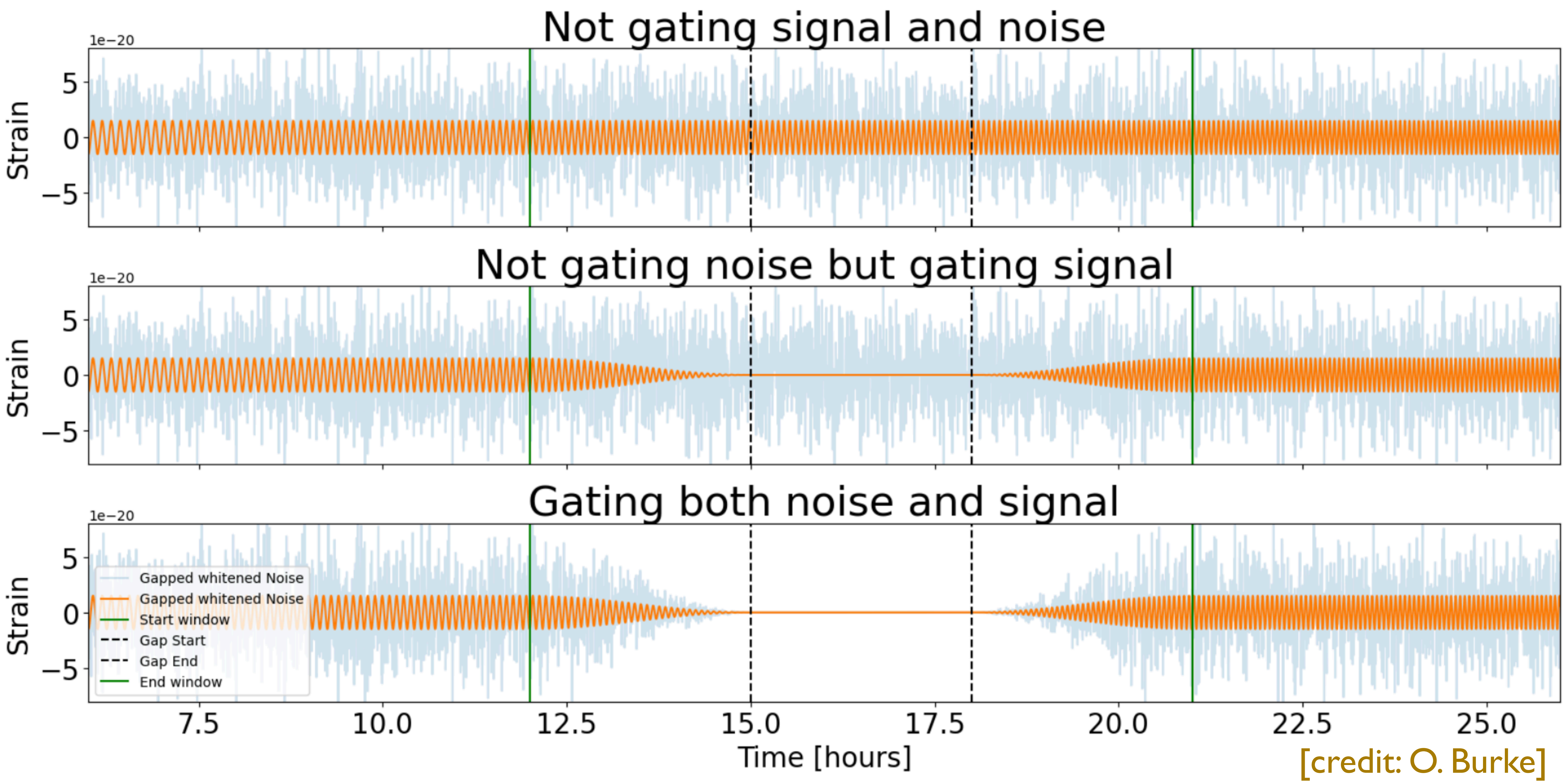
- **Windowing** the data around gaps, having well-behaved Fourier-domain signals

[Dey&al arXiv:2104.12646]

- **Data augmentation** (or imputation): the data inside the gaps is inferred in a Gibbs sampling between GW params and missing data

[Baghi&al arXiv:1907.04747]





[credit: O. Burke]

- **Modelling of the full covariance:** direct marginalization over missing data, useful as a test bed for our assumptions — limited by the size of the $N \times N$ covariance matrix — **This work**

The basics: Stationary Gaussian process

- Assumption: noise as **Gaussian process** described by its covariance

$$C(t,t') = \langle n(t) n(t') \rangle$$

- Assumption: underlying noise (before introducing gaps) **Stationary**, with autocorrelation depending on lag only

$$C(t,t') = C(0, t'-t) \equiv C(t'-t)$$

- In the Fourier domain, this leads to independence:

$$\langle \tilde{n}(f) \tilde{n}^*(f') \rangle = \frac{1}{2} S_n(f) \delta(f - f')$$

- Usually represented with 1-sided PSD $S_n(f)$

- Noise-weighted inner product over positive frequencies:

$$(a|b) = 4 \mathrm{Re} \int \frac{df}{S_n(f)} \tilde{a}^*(f) \tilde{b}(f)$$

- Likelihood:

$$\ln \mathcal{L}(\theta) = -\frac{1}{2} (h(\theta) - d | h(\theta) - d)$$

Whittle likelihood

The Fourier-domain covariance matrix is diagonal: from $N \times N$ to N !

TD/FD data vectors and covariance

Time domain covariance matrix

$$\langle N N^T \rangle = \Sigma$$

Sationarity imposes a Toeplitz structure

$$C(t,t') \equiv C(t-t')$$

$$A = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & \cdots & \cdots & a_{-(n-1)} \\ a_1 & a_0 & a_{-1} & \ddots & & \vdots \\ a_2 & a_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & a_{-1} & a_{-2} \\ \vdots & & \ddots & a_1 & a_0 & a_{-1} \\ a_{n-1} & \cdots & \cdots & a_2 & a_1 & a_0 \end{bmatrix}$$

Diagonality after DFT requires in fact **Circulant** structure (periodicity)

$$C = \begin{bmatrix} c_0 & c_{n-1} & \cdots & c_2 & c_1 \\ c_1 & c_0 & c_{n-1} & & c_2 \\ \vdots & c_1 & c_0 & \ddots & \vdots \\ c_{n-2} & & \ddots & \ddots & c_{n-1} \\ c_{n-1} & c_{n-2} & \cdots & c_1 & c_0 \end{bmatrix}$$

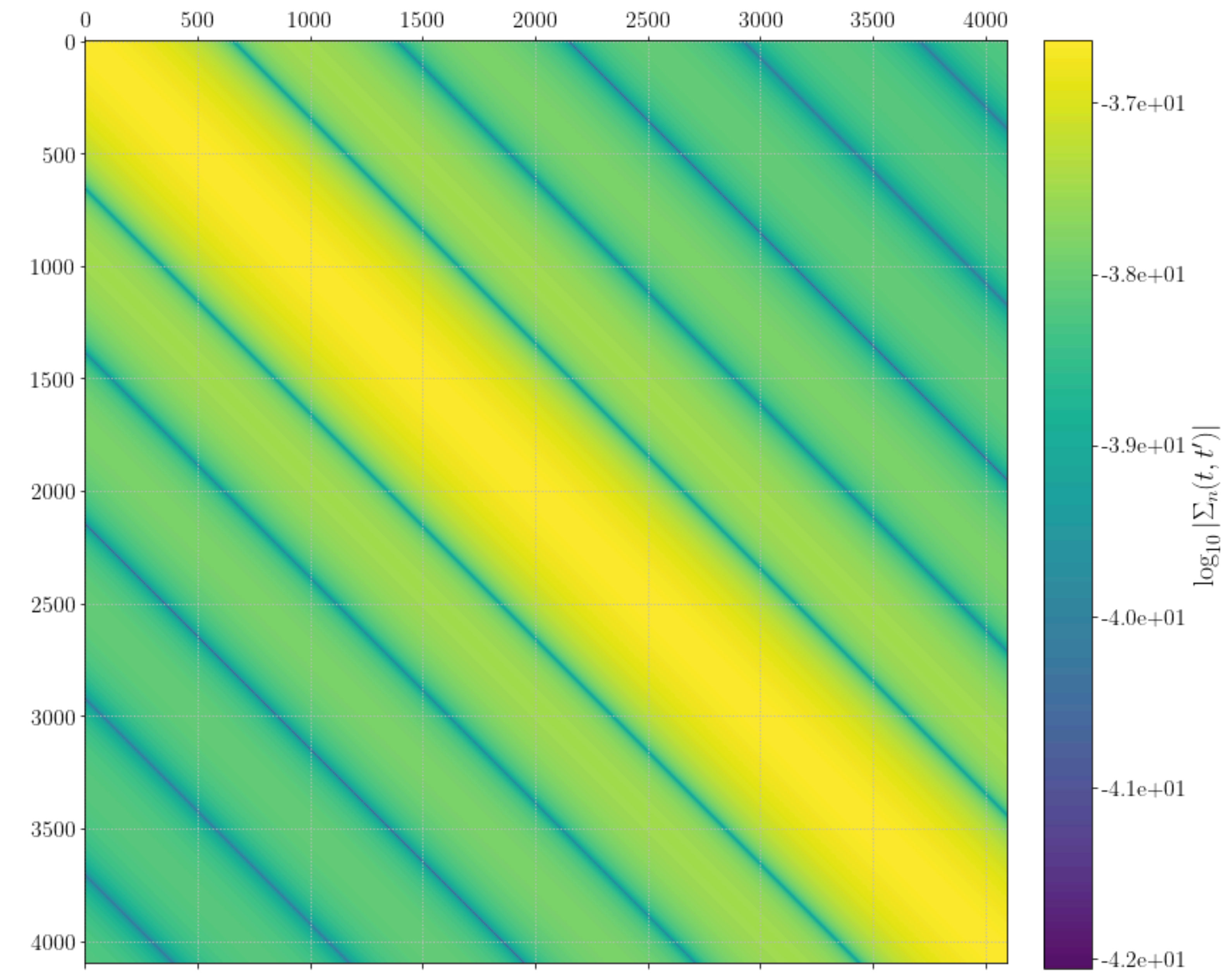
Discrete Fourier transform

$$\tilde{F}(f) = \int dt\, e^{-2i\pi f t} F(t)$$

$$\tilde{F}(f_j) = \Delta t \sum_{i=0}^{N-1} \omega^{-ij} F(t_i) \qquad \omega = e^{\frac{2i\pi}{N}}$$

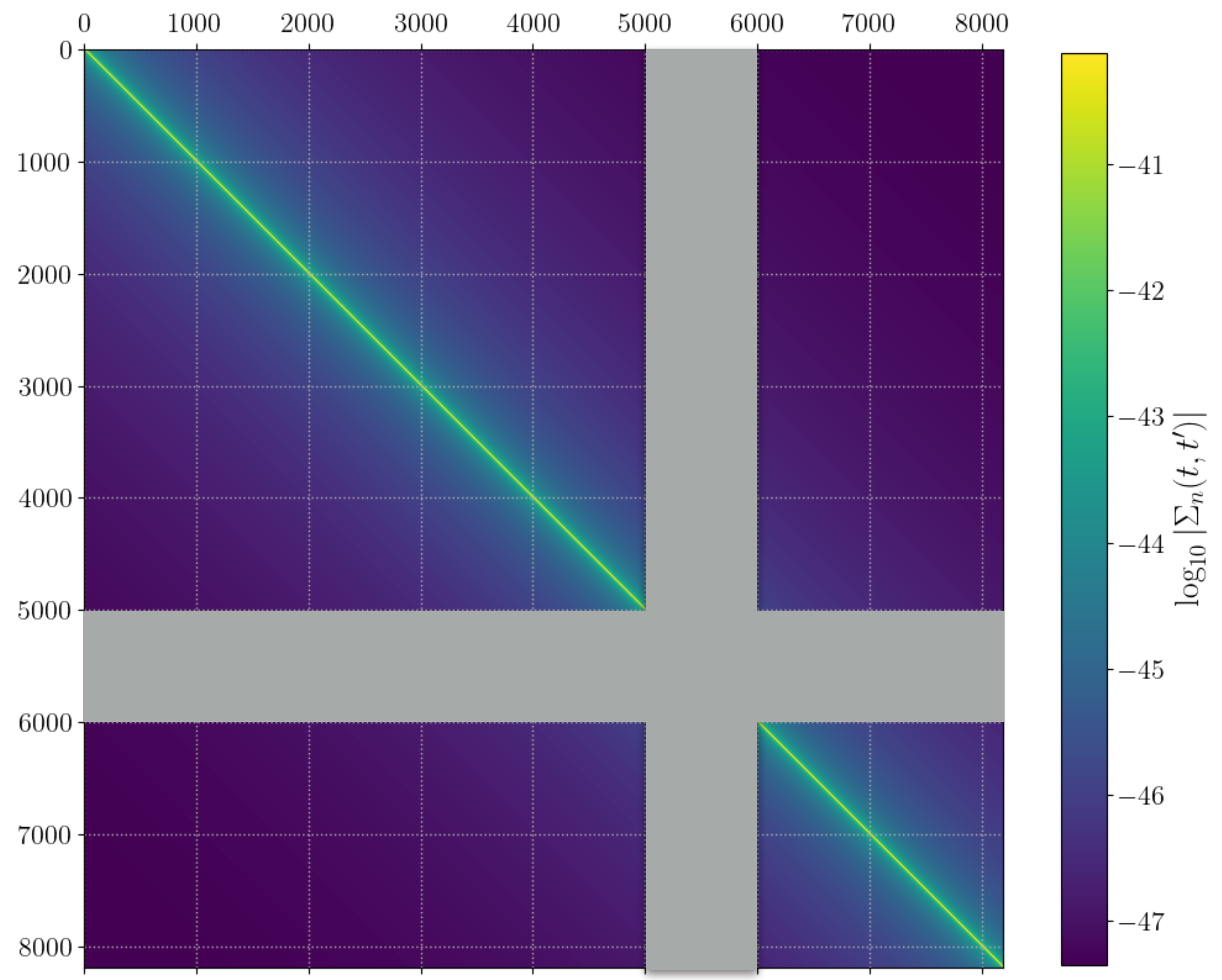
$$\langle \tilde{N} \tilde{N}^\dagger \rangle = \tilde{\Sigma}$$ with $\tilde{\Sigma}$ diagonal

Example TD covariance



Marginalizing out the missing data: Time Domain

Direct marginalization:



Truncation of the Gaussian covariance:

$$X = H(\theta) - D = -N \qquad X = (X_1, X_2)$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{pmatrix}$$

$$\mathcal{L}_{\mathrm{gap}} = p(N_1) = \int dN_2\, p(N_1, N_2)$$

$$\ln \mathcal{L}_{\mathrm{gap}} = -\frac{1}{2} X_1^T \Sigma_{11}^{-1} X_1$$

Gap seen as a gated process:

For an invertible modulation W:

$$X \to WX$$

$$X^T \Sigma^{-1} X = (WX)^T (W\Sigma W)^{-1} (WX)$$

For an gating function W, non-invertible: $$W = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}$$

Moore-Penrose **pseudo-inverse** (uniquely defined):

$$A^+ A A^+ = A^+$$
$$A A^+ A = A$$
$$(AA^+)^\dagger = AA^+$$
$$(A^+ A)^\dagger = A^+ A$$

$$(W\Sigma W)^+ = \begin{pmatrix} \Sigma_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

Equivalence with the marginalized likelihood:

$$X^T (W\Sigma W)^+ X = X_1^T \Sigma_{11}^{-1} X_1$$

Note that windowing the data becomes facultative !

$$(WX)^T (W\Sigma W)^+ (WX) = X^T (W\Sigma W)^+ X$$

Marginalizing out the missing data: Fourier Domain

DFT in linear algebra:

$$\tilde{F}(f) = \int dt\, e^{-2i\pi f t} F(t)$$

$$\tilde{F}(f_j) = \Delta t \sum_{i=0}^{N-1} \omega^{-ij} F(t_i) \quad \omega = e^{\frac{2i\pi}{N}}$$

Using the DFT matrix P:

$$\tilde{F} = \Delta t \sqrt{N} P F$$

$$P_{ij} = \frac{1}{\sqrt{N}} \omega^{-ij}$$

$$P P^\dagger = \mathbb{1}$$

FD covariance and FD window:

$$\tilde{\Sigma} = N \Delta t^2 P \Sigma P^\dagger$$

$$\tilde{W} = P W P^\dagger$$

Translation from time to Fourier domain:

The pseudo-inverse is transparent to the unitary matrix P:

$$P\,(W\Sigma W)^+\, P^\dagger = \left(P W \Sigma W P^\dagger\right)^+$$

This allows for a direct translation:

$$\ln \mathcal{L}_{\text{gap}} = -\frac{1}{2}(WX)^T\,(W\Sigma W)^+\,(WX) = -\frac{1}{2}\left(\widetilde{WX}\right)^\dagger \left(\tilde{W}\tilde{\Sigma}\tilde{W}\right)^+ \left(\widetilde{WX}\right)$$

Windowing the data is facultative, in FD also:

$$\ln \mathcal{L}_{\text{gap}} = -\frac{1}{2}\left(\widetilde{WD} - \tilde{H}\right)^\dagger \left(\tilde{W}\tilde{\Sigma}\tilde{W}\right)^+ \left(\widetilde{WD} - \tilde{H}\right)$$

In practice, computing the FD gated covariance matrix is a convolution:

$$\left(\tilde{W}\tilde{\Sigma}\tilde{W}\right)_{ij} = \frac{\Delta f}{2} \sum_k S_n^k \tilde{w}_{i-k} \tilde{w}^*_{j-k}$$

cost scales as $\mathcal{O}(N^2 \log N)$

$$\tilde{\Sigma}_w(f, f') = \frac{1}{2}\int_{-\infty}^{+\infty} dv\, S_n(v)\tilde{w}(f - v)\tilde{w}^*(f' - v)$$

Numerics: pseudo-inverse of the FD covariance

- The Moore-Penrose pseudo-inverse is uniquely defined

- Most straightfoward algorithm: use the **singular value decomposition** (SVD)

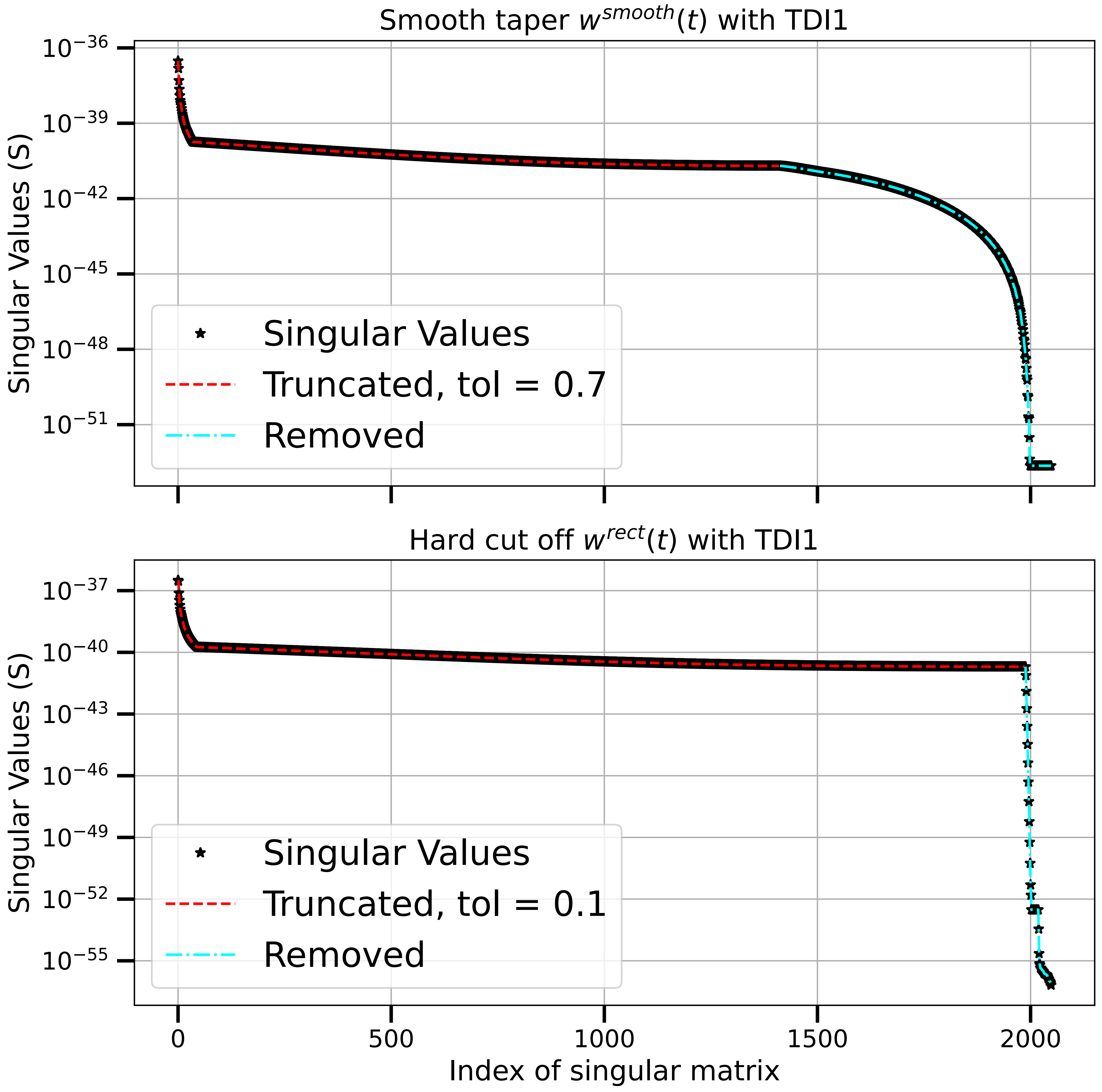
$$M = U\Sigma V^\dagger$$

$$\Sigma = (\sigma_1^2, \ldots, \sigma_p^2, 0, \ldots, 0)$$

And use the pseudo-inverse of singular values:

$$\Sigma^+ = (1/\sigma_1^2, \ldots, 1/\sigma_p^2, 0, \ldots, 0)$$

- The procedure also allows for regularizing the covariance (for instance when dealing with smooth windows)



Effect of noise on posterior and Cutler-Vallisneri biases

Linearized signal approximation:

$$H(\theta) \simeq H(\theta_0) + \Delta\theta^i \partial_i H$$

Likelihood: $(A|B)_\Sigma = A^T \Sigma^{-1} B$

$$\ln \mathcal{L} = -\frac{1}{2}(H - D|H - D)_\Sigma$$

$$\ln \mathcal{L} \simeq -\frac{1}{2}\Gamma_{ij}\Delta\theta^i\Delta\theta^j$$

Fisher information matrix:

$$\Gamma_{ij} = (\partial_i H|\partial_j H)_\Sigma$$

Fisher covariance matrix between parameters:

$$\Gamma^{-1}_{ij}$$

Cutler-Vallisneri bias due to noise realization:

$$D = H(\theta_0) + N$$

$$\ln \mathcal{L} \simeq -\frac{1}{2}\Delta\theta^i\Delta\theta^j(\partial_i H|\partial_j H)_\Sigma + \Delta\theta^i(\partial_i H|N)_\Sigma$$

CV bias and variance:

$$\Delta\theta^i_{\mathrm{bf}} = \Gamma^{-1}_{ij}(\partial_j H|N)_\Sigma \qquad \langle\Delta\theta^i_{\mathrm{bf}}\rangle = 0$$

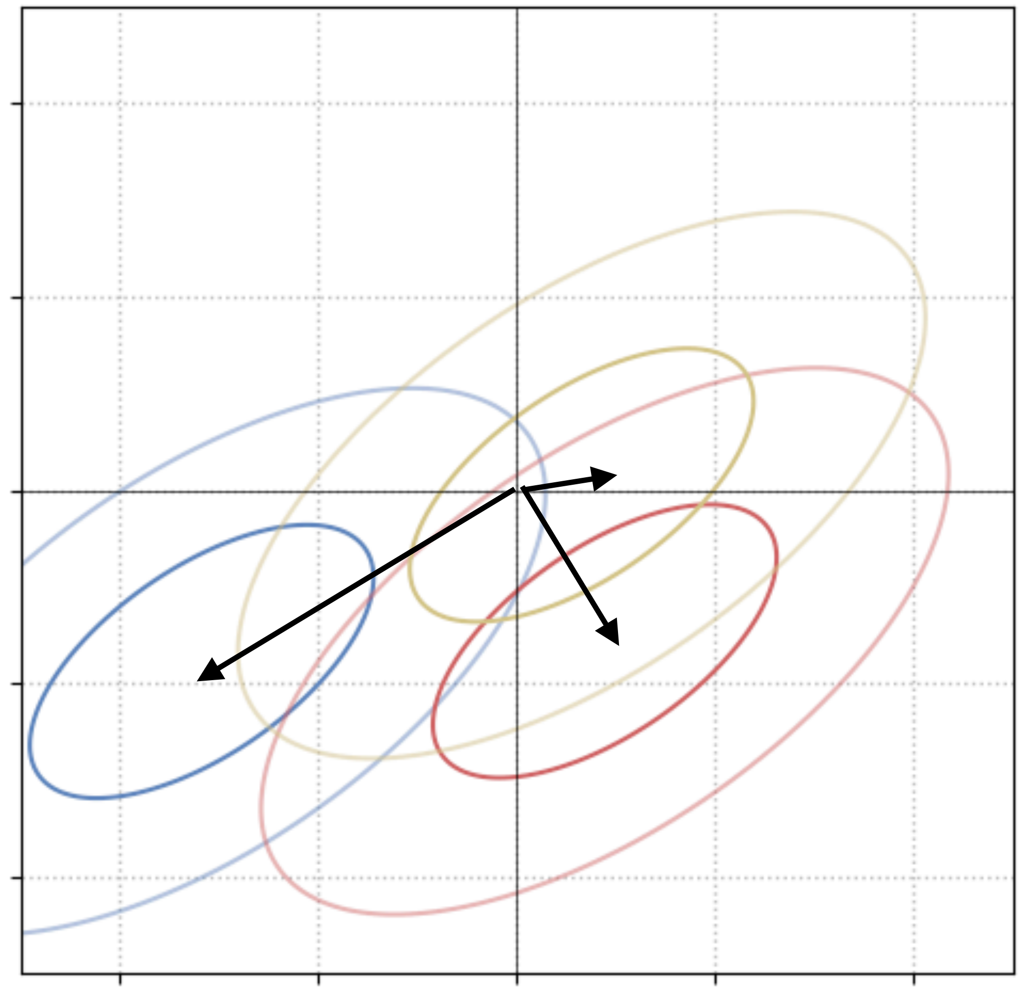
CV bias variance:

$$\langle\Delta\theta^i_{\mathrm{bf}}\Delta\theta^j_{\mathrm{bf}}\rangle = \Gamma^{-1}_{ik}\Gamma^{-1}_{jl}\partial_k H^T \Sigma^{-1}\langle N N^T\rangle\Sigma^{-1}\partial_l H$$

$$\langle N N^T\rangle = \Sigma$$

$$\langle\Delta\theta^i_{\mathrm{bf}}\Delta\theta^j_{\mathrm{bf}}\rangle = \Gamma^{-1}_{ij}$$

The CV bias variance matches the Fisher covariance



The noise ‘scatters’ the posteriors:

Mismodelling of the gated noise process

Introduce mismodelling in the covariance itself:

- Wrong covariance Σ' assuming stationarity, gating the signals
- Could also be used to introduce a misestimation of the PSD

$$(A|B)_{\Sigma'} = A^T \Sigma'^{-1} B$$

We use the wrong likelihood:

$$\ln \mathcal{L} = -\frac{1}{2}(W(H-D)|W(H-D))_{\Sigma'}$$

$$\ln \mathcal{L} \simeq -\frac{1}{2}\Delta\theta^i \Delta\theta^j (W\partial_i H|W\partial_j H)_{\Sigma'} + \Delta\theta^i (W\partial_i H|WN)_{\Sigma'}$$

$$\Delta\theta^i_{\rm bf'} = {\Gamma'}^{-1}_{ij}(W\partial_j H|WN)_{\Sigma'} \qquad \langle \Delta\theta^i_{\rm bf'} \rangle = 0$$

bias still zero-mean !

$${\Gamma'}_{ij} = (W\partial_i H|W\partial_j H)_{\Sigma'}$$

true covariance

$$\langle \Delta\theta^i_{\rm bf'} \Delta\theta^j_{\rm bf'} \rangle = {\Gamma'}^{-1}_{ik}{\Gamma'}^{-1}_{jl}\partial_k H^T W \Sigma'^{-1}\langle N N^T \rangle W \Sigma'^{-1} W \partial_l H$$

$$\langle \Delta\theta^i_{\rm bf'} \Delta\theta^j_{\rm bf'} \rangle = {\Gamma'}^{-1}_{ik}{\Gamma'}^{-1}_{jl}\partial_k H^T W \Sigma'^{-1} W \Sigma W \Sigma'^{-1} W \partial_l H$$

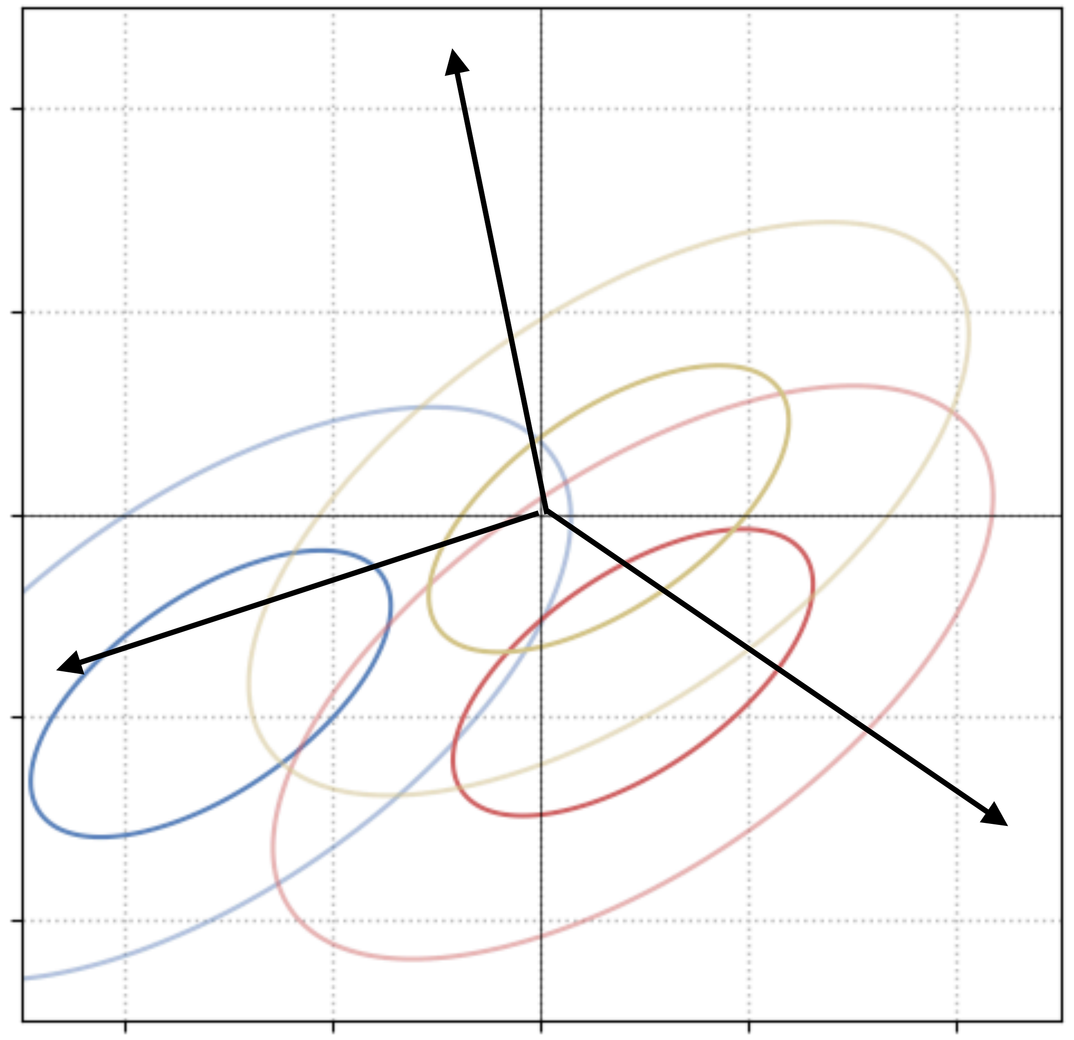
[Edy&al arXiv:2101.07743]

A measure of inconsistency:

$$\Upsilon_i = \frac{\sqrt{\langle \Delta\theta^i_{\rm bf'} \Delta\theta^j_{\rm bf'} \rangle}}{\sqrt{\Gamma^{-1}_{ii}}}$$

- Would be 1 for the correct covariance
- Measures the inconsistency of the posterior 'scatter'
- The bias is still 0-mean here
- SNR-independent

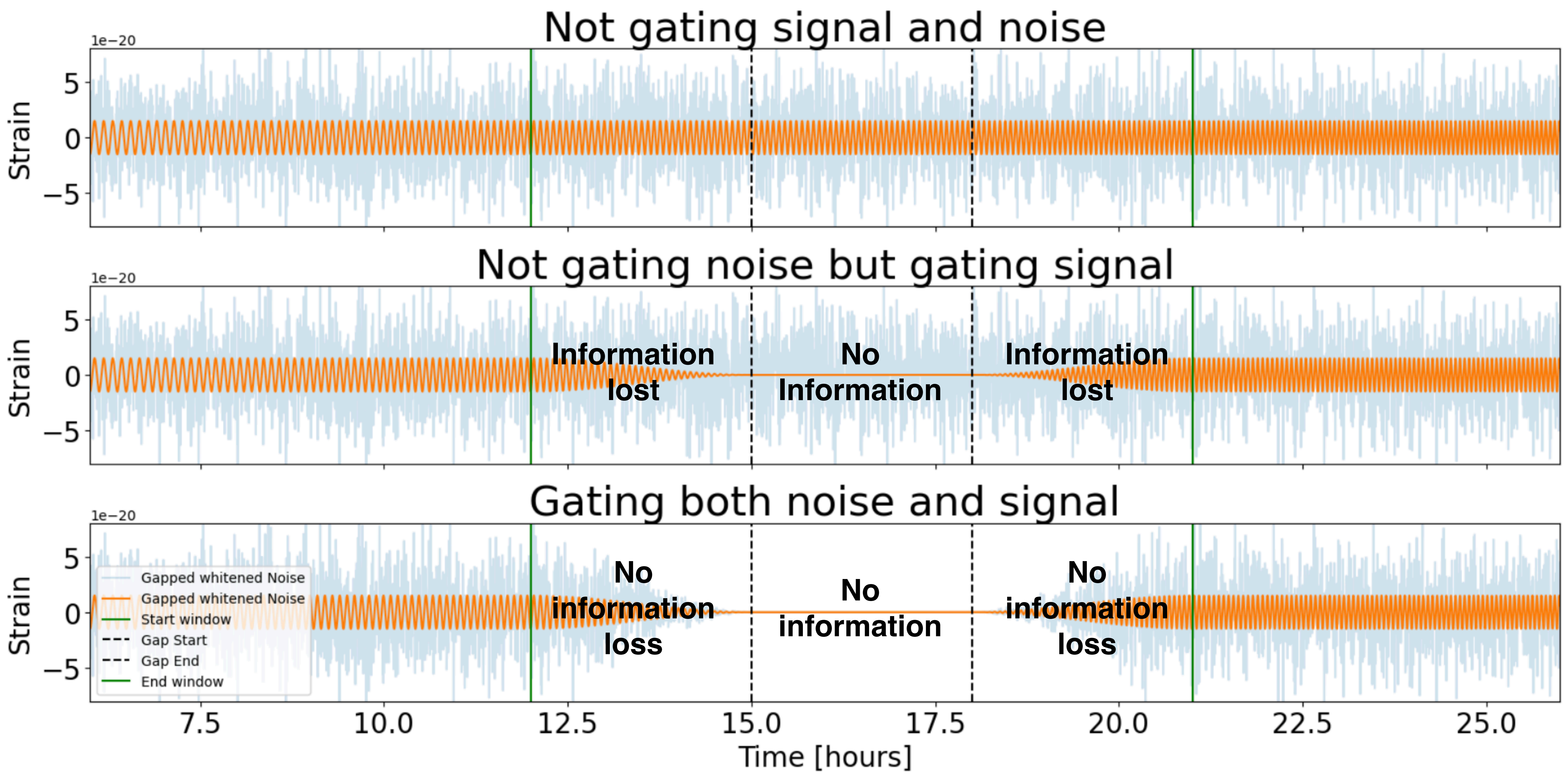
Inconsistent 'scatter'



Windowing for data gaps

Different logics:

- Windowing only the data/signal: loss of information
- Windowing the data/signal and the covariance: no loss of information
- Windowing the noise: keep diagonal-dominated structure of the FD covariance

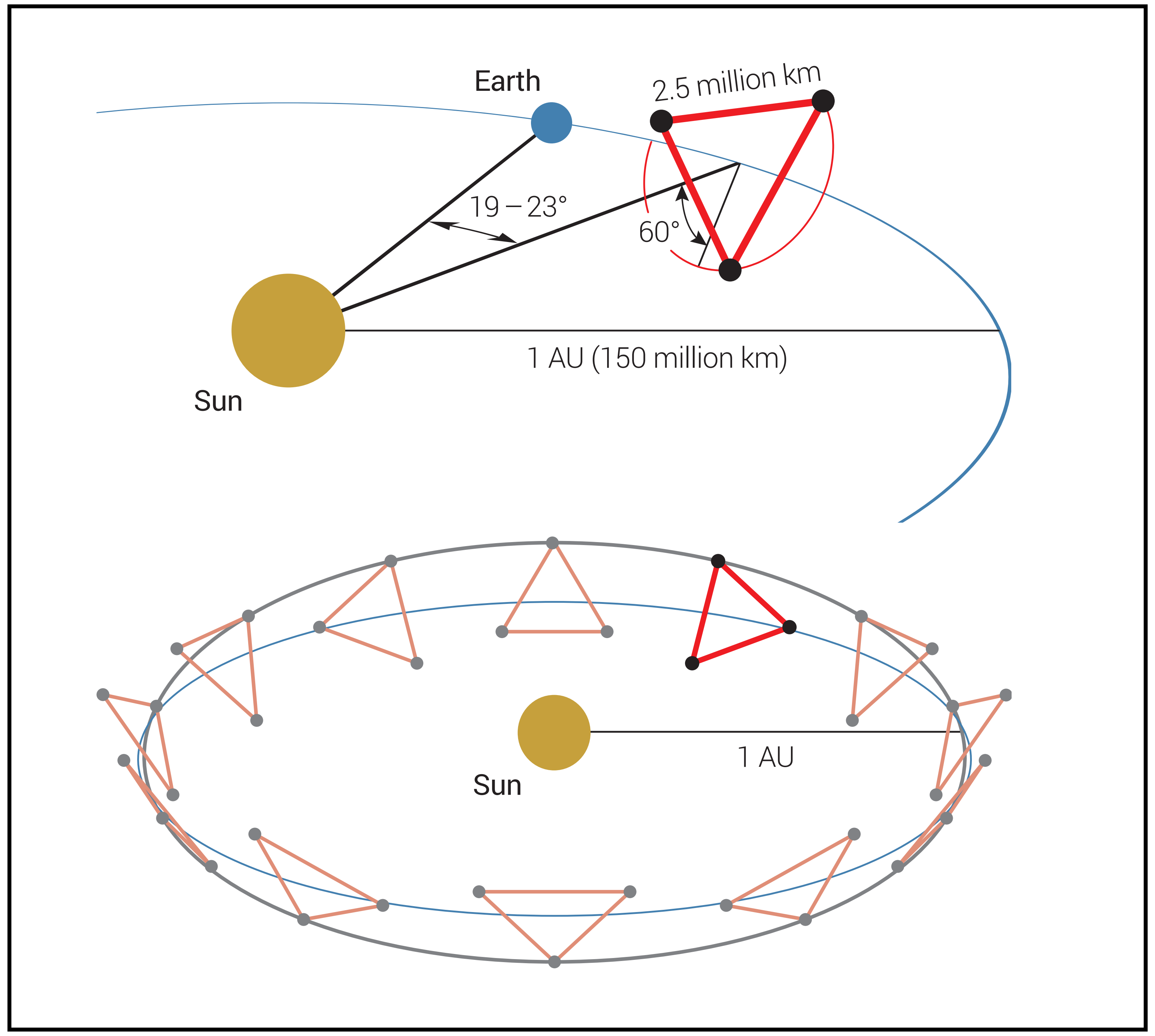


Effects of windowing:

- Windowing: the gap edge: alleviate the statistical offense by avoiding a jump to 0
- Windowing the signal: avoids Gibbs oscillation in the Fourier domain
- Windowing the noise: keep diagonal-dominated structure of the FD covariance, but delicate (pseudo)inversion of the matrix

Work in progress

LISA data and TDI generations



Doppler delay from orbit, change in orientation

Analogous to 2 LIGO in motion at low frequencies only

One-link observables: from spacecraft s to spacecraft r through link s:

$$y = \Delta\nu/\nu$$

$$y_{slr} = \frac{1}{2}\frac{1}{1-\hat{k}\cdot n_l} n_l \cdot (h(t_s) - h(t_r)) \cdot n_l$$

Response time and frequency-dependent:

$$\mathcal{T}_{slr} = \frac{i\pi f L}{2} \mathrm{sinc}\left[\pi f L\left(1 - k\cdot n_l\right)\right] \exp\left[i\pi f\left(L + k\cdot(p_r + p_s)\right)\right] n_l \cdot P \cdot n_l(t_f)$$

+ Time-delay interferometry (TDI), 1st and 2nd generations:

$$X_2^{GW} = \underbrace{\left[(y_{31}^{\mathrm{GW}} + y_{13,2}^{\mathrm{GW}}) + (y_{21}^{\mathrm{GW}} + y_{12,3}^{\mathrm{GW}})_{,22} - (y_{21}^{\mathrm{GW}} + y_{12,3}^{\mathrm{GW}}) - (y_{31}^{\mathrm{GW}} + y_{13,2}^{\mathrm{GW}})_{,33}\right]}_{X^{\mathrm{GW}}(t)}$$
$$- \underbrace{\left[(y_{31}^{\mathrm{GW}} + y_{13,2}^{\mathrm{GW}}) + (y_{21}^{\mathrm{GW}} + y_{12,3}^{\mathrm{GW}})_{,22} - (y_{21}^{\mathrm{GW}} + y_{12,3}^{\mathrm{GW}}) - (y_{31}^{\mathrm{GW}} + y_{13,2}^{\mathrm{GW}})_{,33}\right]_{,2233}}_{X^{\mathrm{GW}}(t-2L_2-2L_3)\simeq X^{\mathrm{GW}}(t-4L)} .$$

TDI variables amount to taking (discretized) time derivatives of the GW:

$$y_{slr} \sim h(t) - h(t-L)$$
$$X \sim y_{slr}(t) - y_{slr}(t-2L)$$
$$X_2 \sim X(t) - X(t-4L)$$

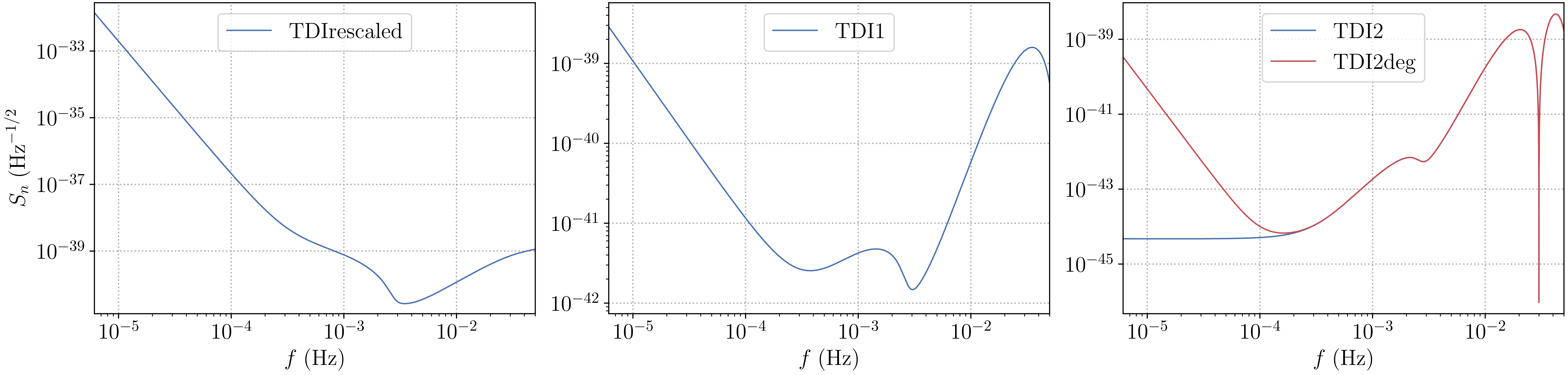
$$\tilde{y}_{slr} \sim \sin(\pi f L(1-k\cdot n))\tilde{h}$$
$$\tilde{X} \sim \sin(\pi f L(1-k\cdot n))\sin(2\pi f L)\tilde{h}$$
$$\tilde{X}_2 \sim \sin(\pi f L(1-k\cdot n))\sin(2\pi f L)\sin(4\pi f L)\tilde{h}$$

Noise PSDs for LISA

$$\texttt{TDIrescaled:} \quad \tilde{y}_{slr} \sim \sin(\pi f L(1 - k \cdot n))\tilde{h}$$

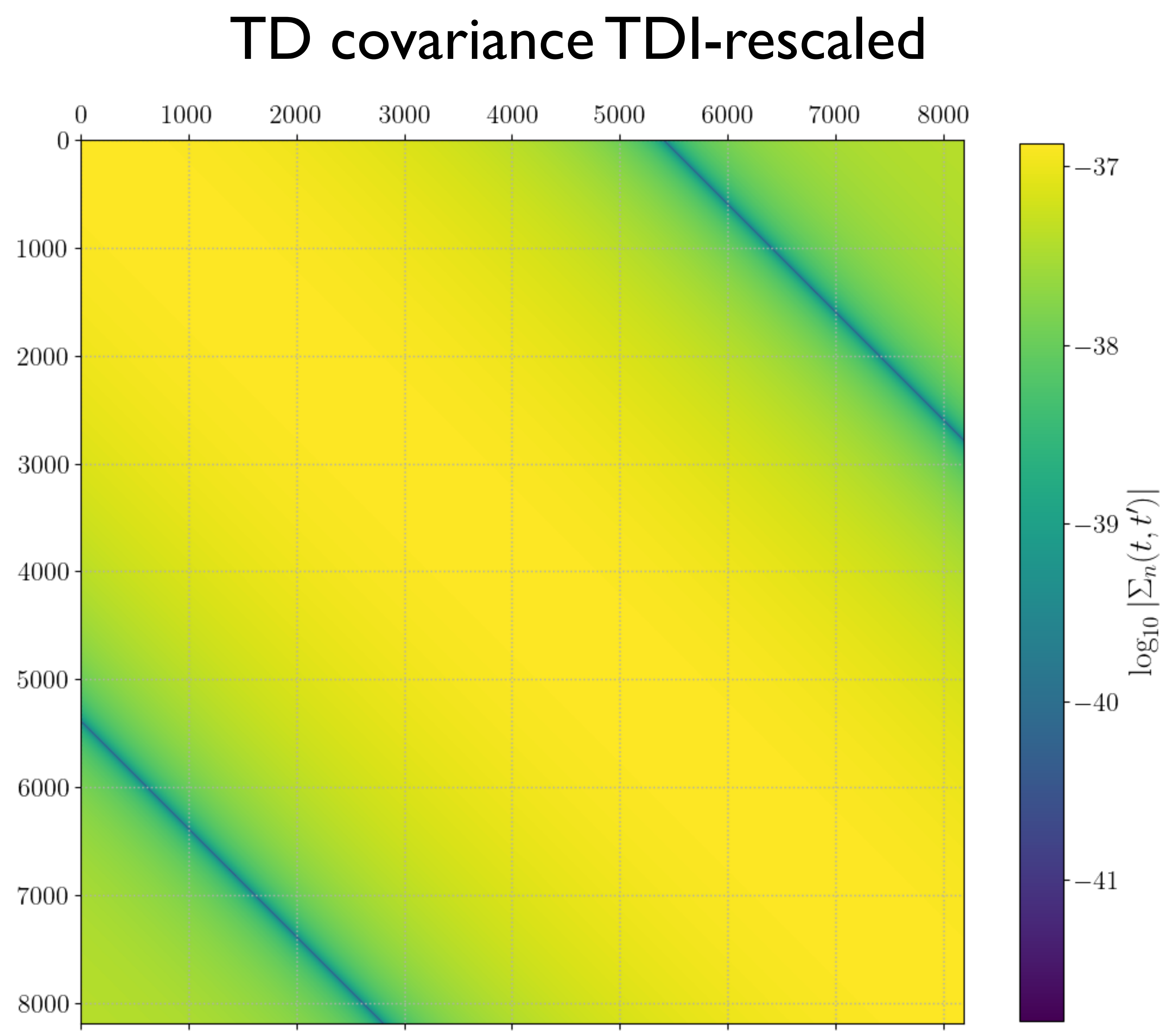
$$\texttt{TDI1:} \quad \tilde{X} \sim \sin(\pi f L(1 - k \cdot n)) \sin(2\pi f L)\tilde{h}$$

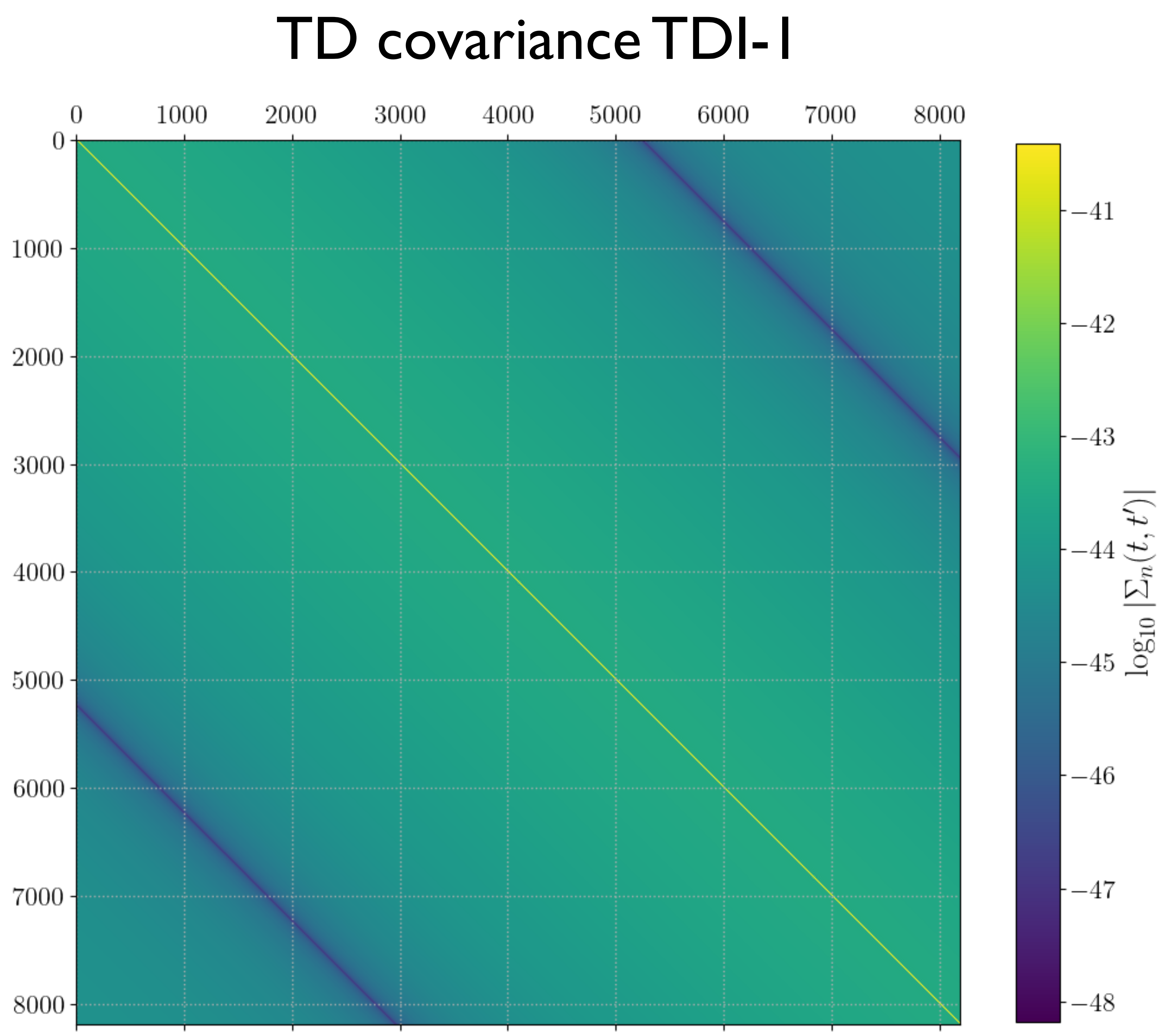
$$\texttt{TDI2:} \quad \tilde{X}_2 \sim \sin(\pi f L(1 - k \cdot n)) \sin(2\pi f L) \sin(4\pi f L)\tilde{h}$$

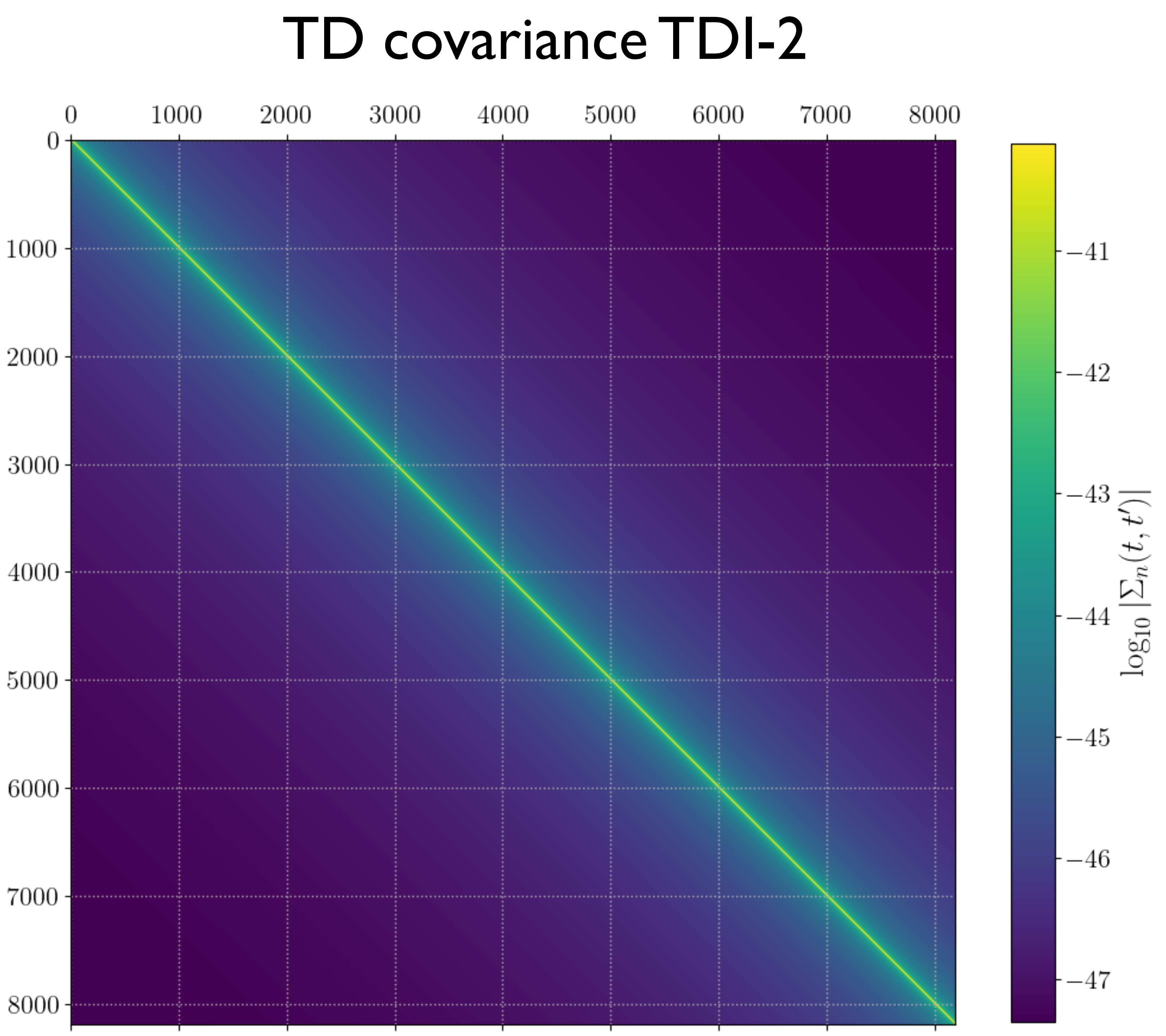


- Dependency on the sampling rate: a smaller Δt gives access to higher frequencies
- Dependency on the total duration: a smaller Δf gives access to lower frequencies

Covariance matrices in Time Domain

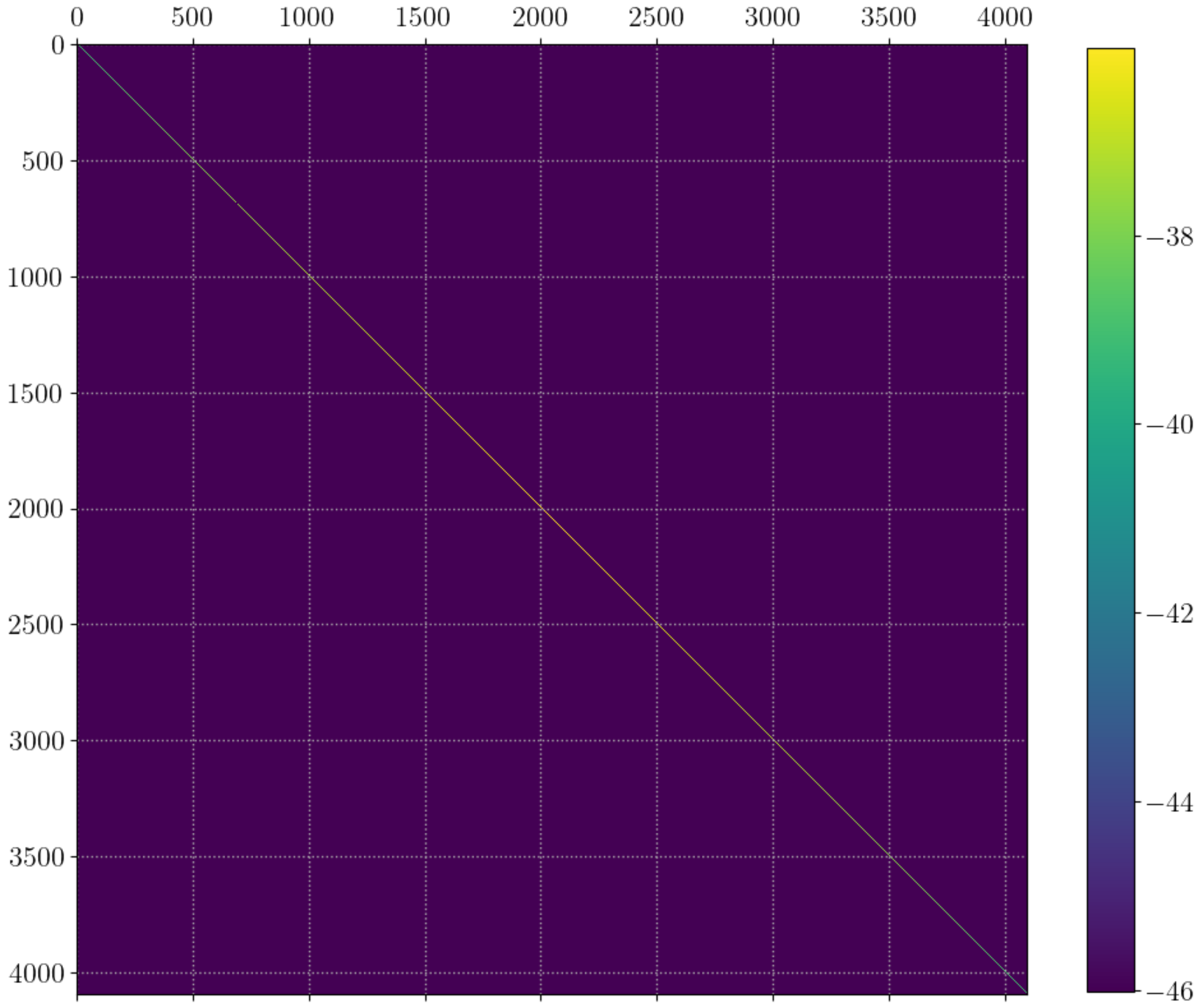




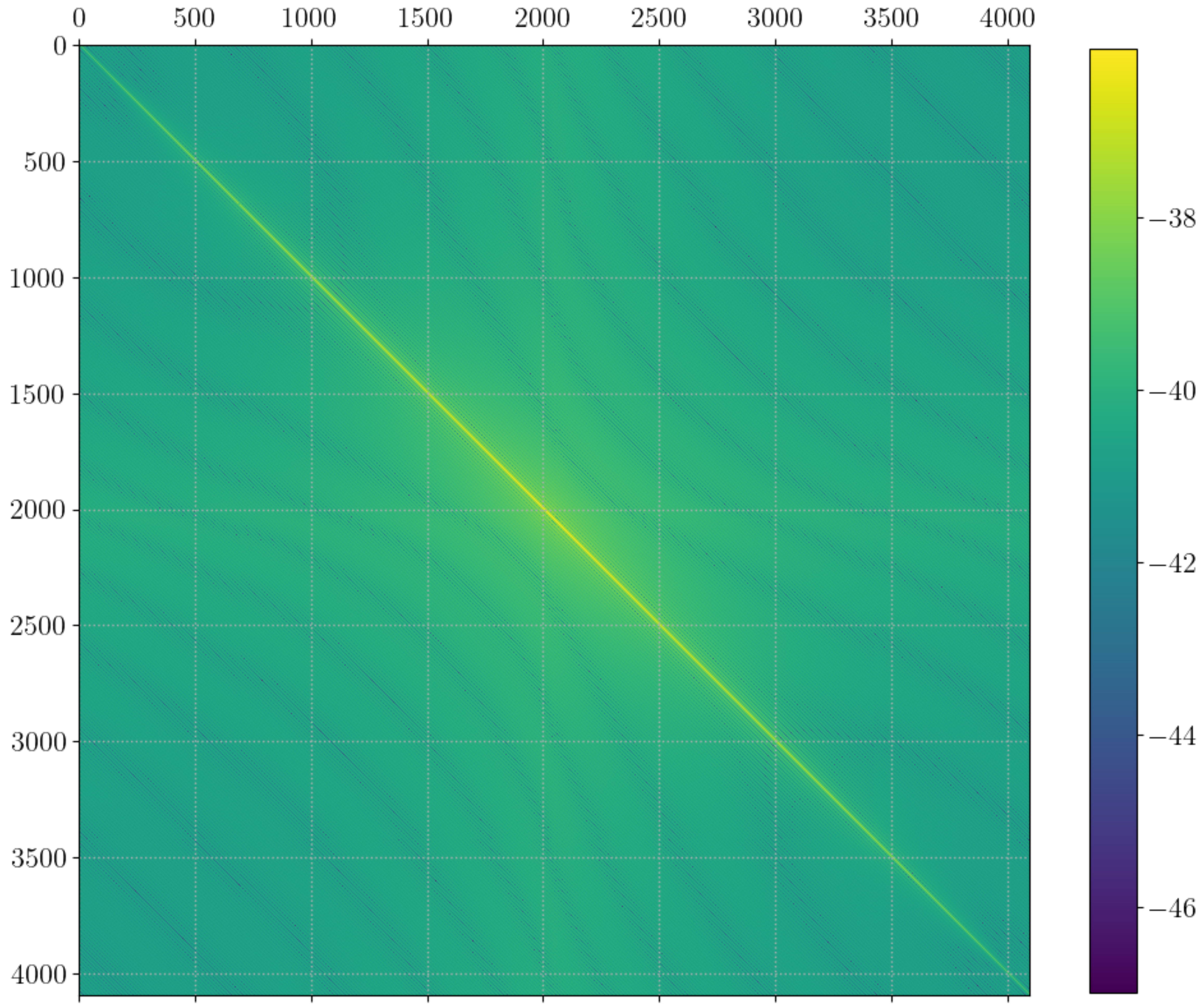


Covariance matrices in Fourier Domain

FD covariance TDI-2 without gap (stationary)

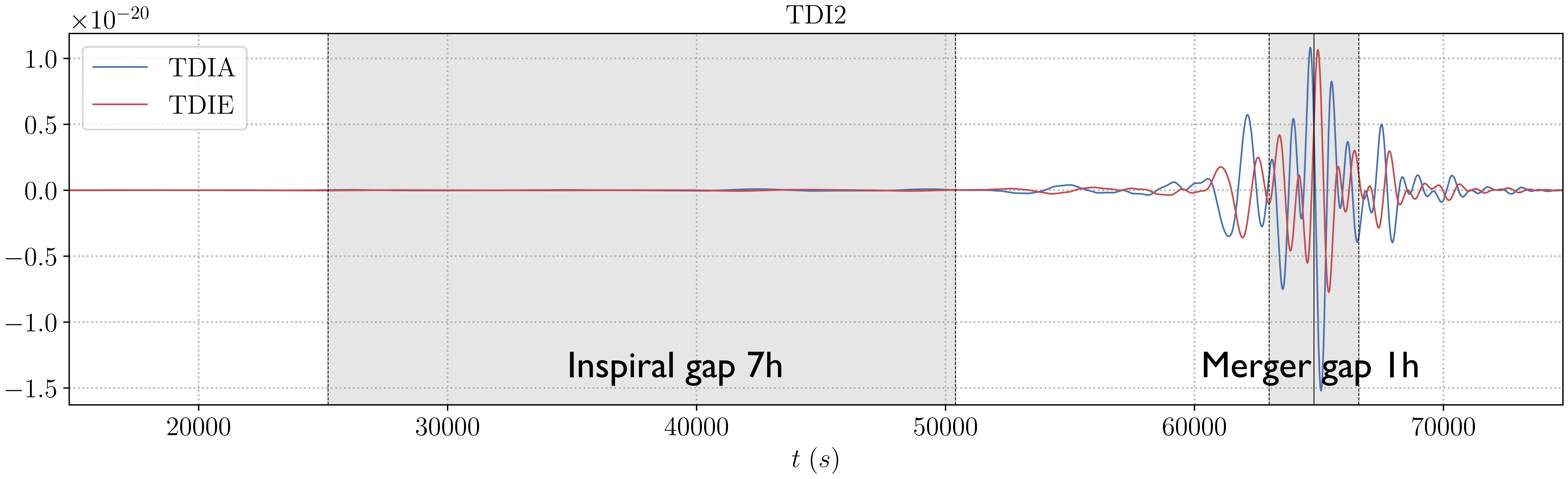


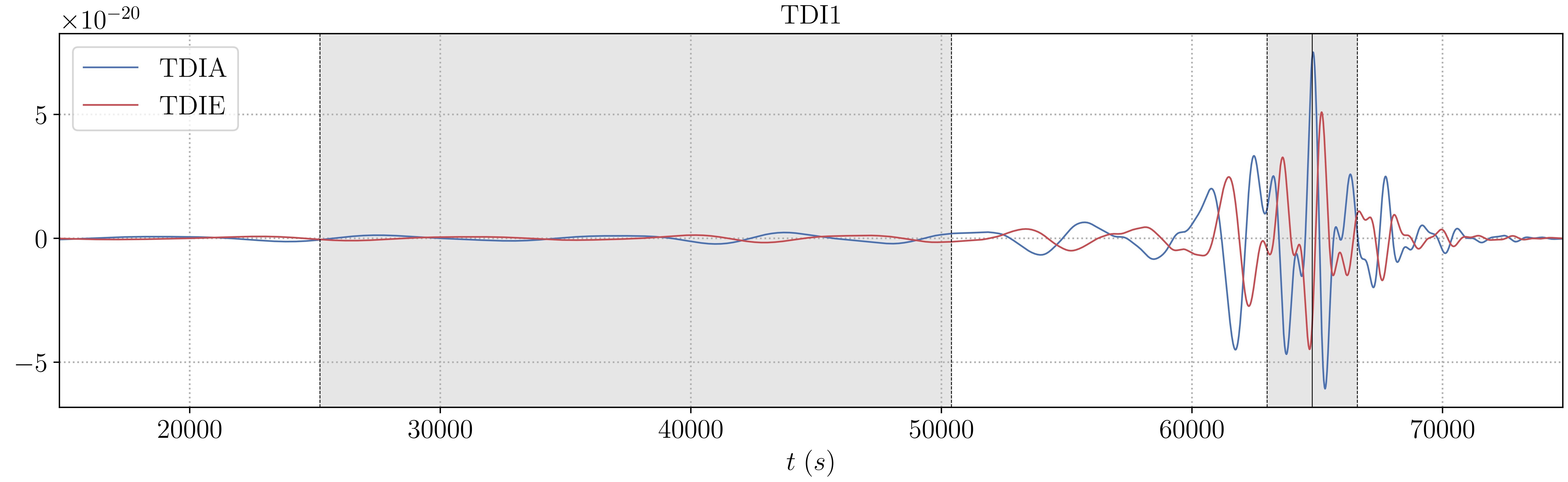
FD covariance TDI-2 with gap



Note: need to consider both positive and negative frequencies

MBHBs: signals and gap configurations



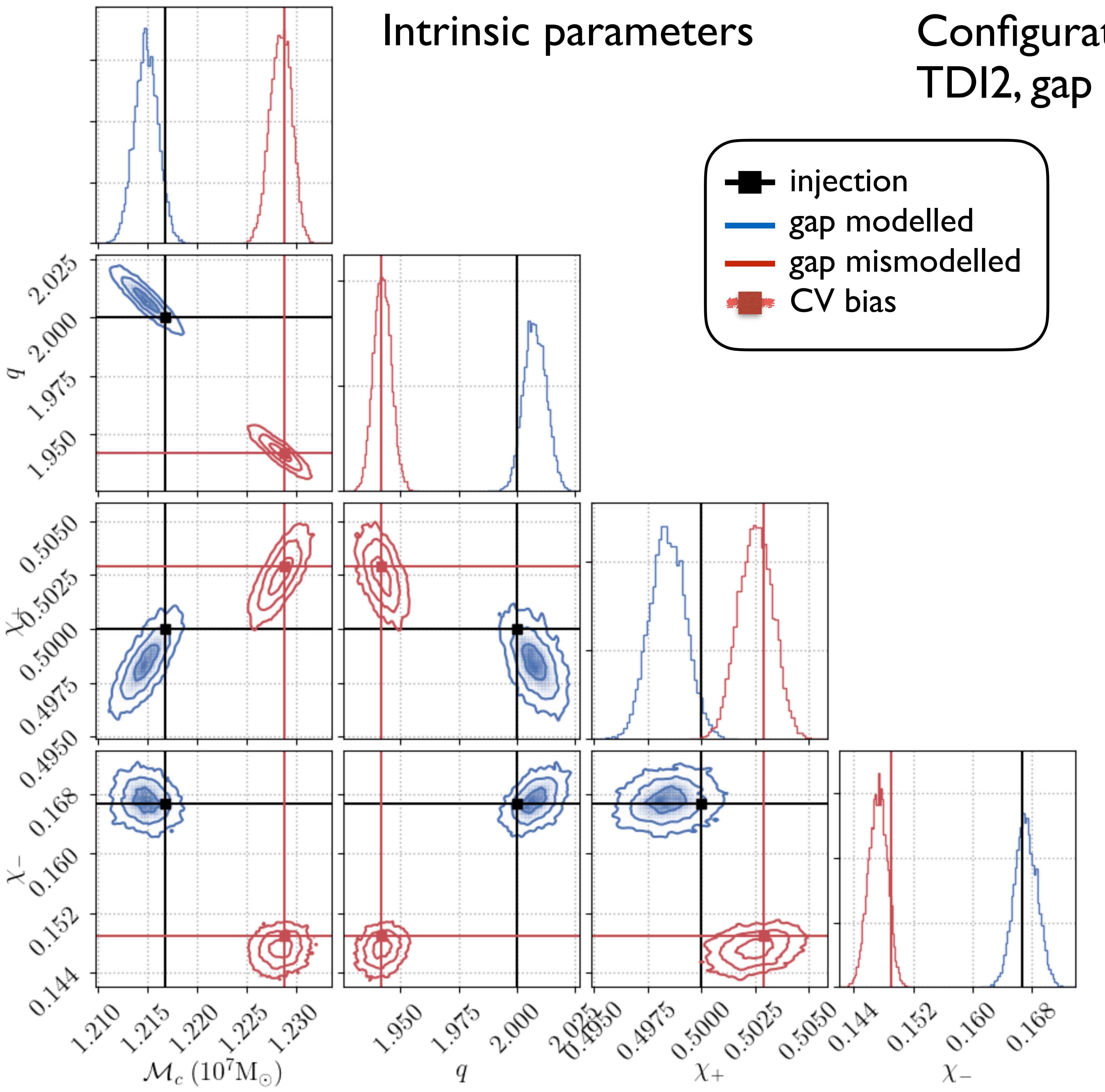


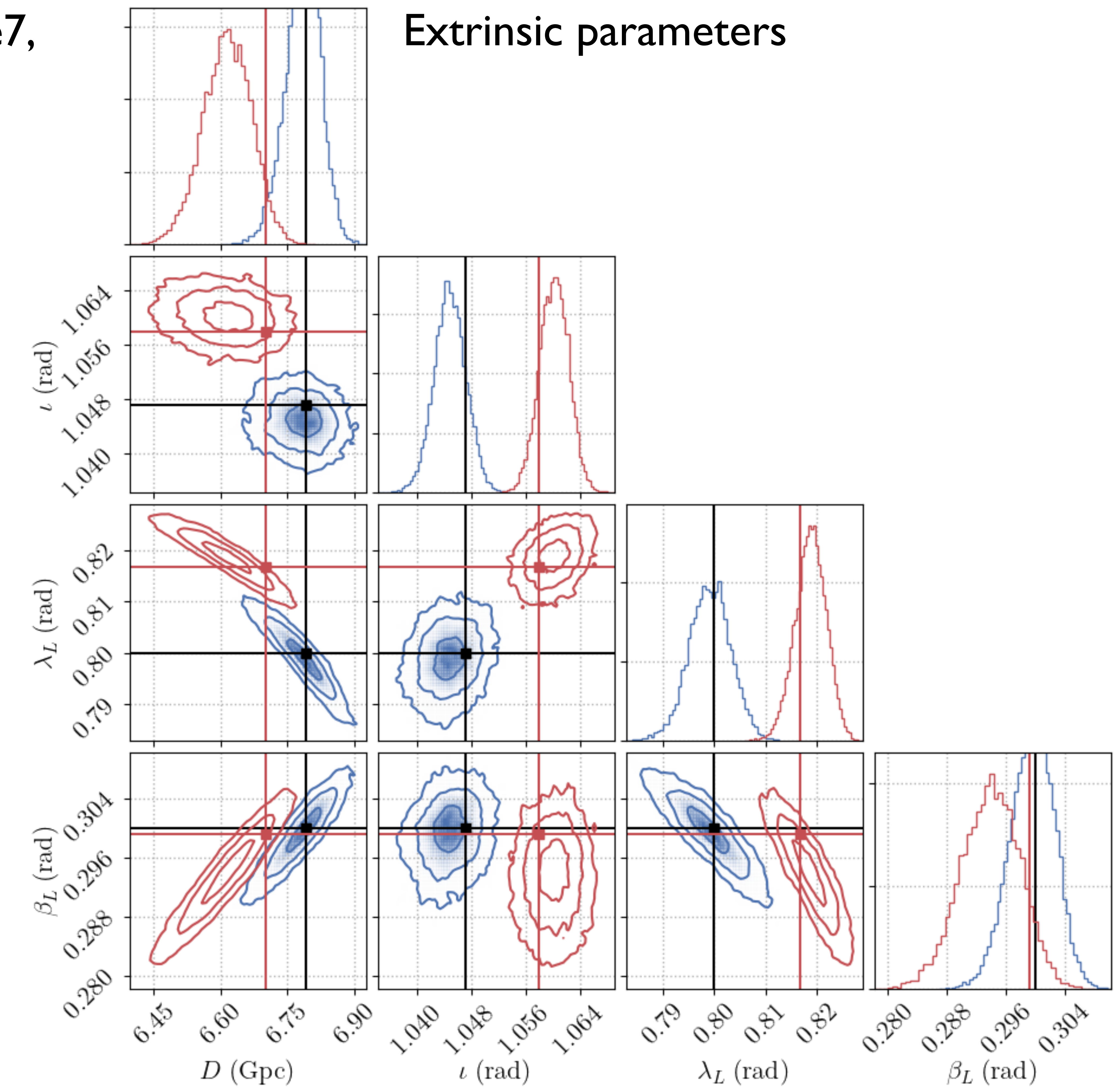
Results: example PE, comparing with CV bias

Intrinsic parameters

Configuration: M=3e7, TDI2, gap at merger

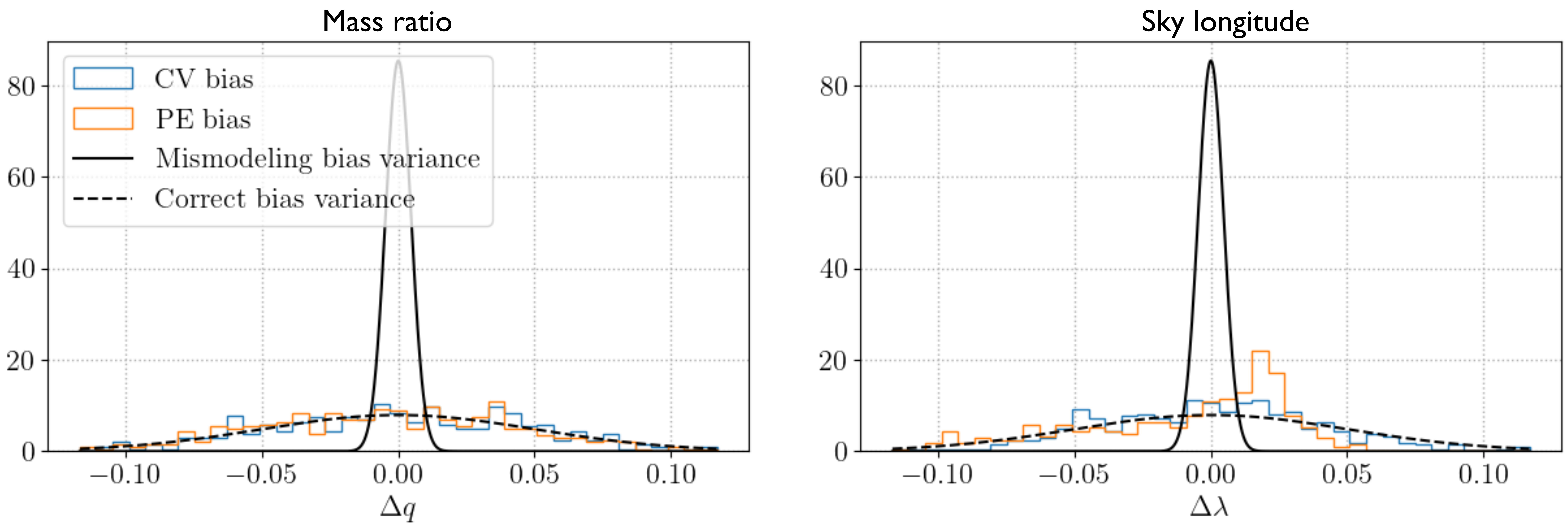
Extrinsic parameters





- Significant biases from mismodelling
- Bias well captured by CV

Results: comparing CV mismodeling prediction to PE

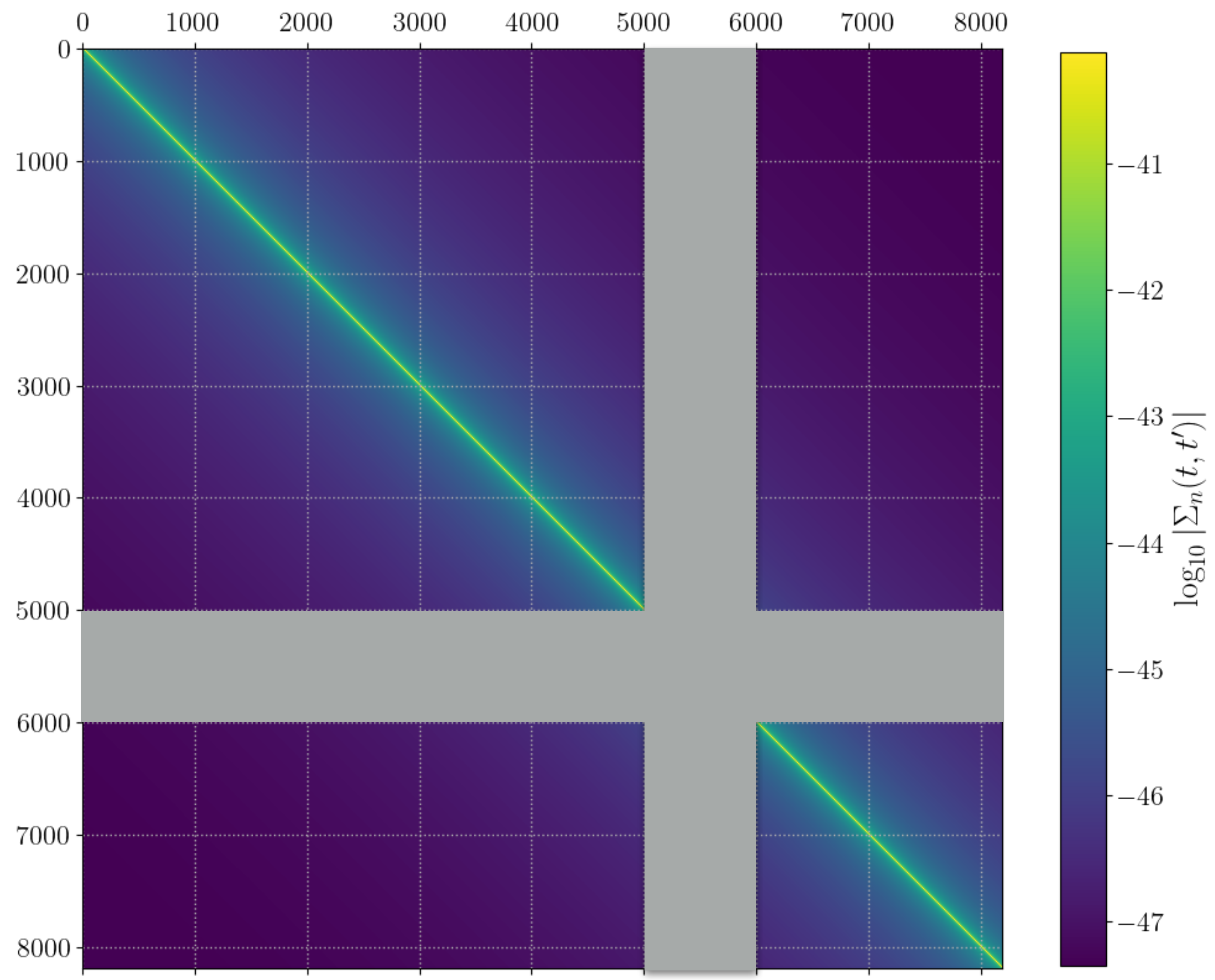


- Cutler-Vallisneri bias: calculated with the wrong covariance
- PE bias: from full PE runs, calculated with the wrong covariance
- Mismodelling bias variance: prediction $\langle \Delta\theta^i_{\mathrm{bf}'} \Delta\theta^j_{\mathrm{bf}'} \rangle$
- Correct bias variance: Fisher prediction with the correct covariance

Validation of the CV-inspired Υ to assess the impact of mismodelling noise

Results: biases from using the Whittle likelihood on gated data

Covariance with gap



- Correct modelling: pseudo-inverse of gated covariance

$$\ln \mathcal{L}_{\text{right}} = -\frac{1}{2} X^T (W \Sigma W)^{+} X$$

- Incorrect modelling: Whittle for gated data and signal

$$\ln \mathcal{L}_{\text{wrong}} = -\frac{1}{2} X^T W \Sigma^{-1} W X$$

Question: how wrong is it to mismodel the covariance using Whittle ?

Gap at merger

TDI2 M=3e7, inj.: gap, model: nogap

θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	7.8	10.8	9.5	10.9	11.4	9.3	11.4	11.0	9.5	10.6	9.6

TDI1 M=3e7, inj.: gap, model: nogap

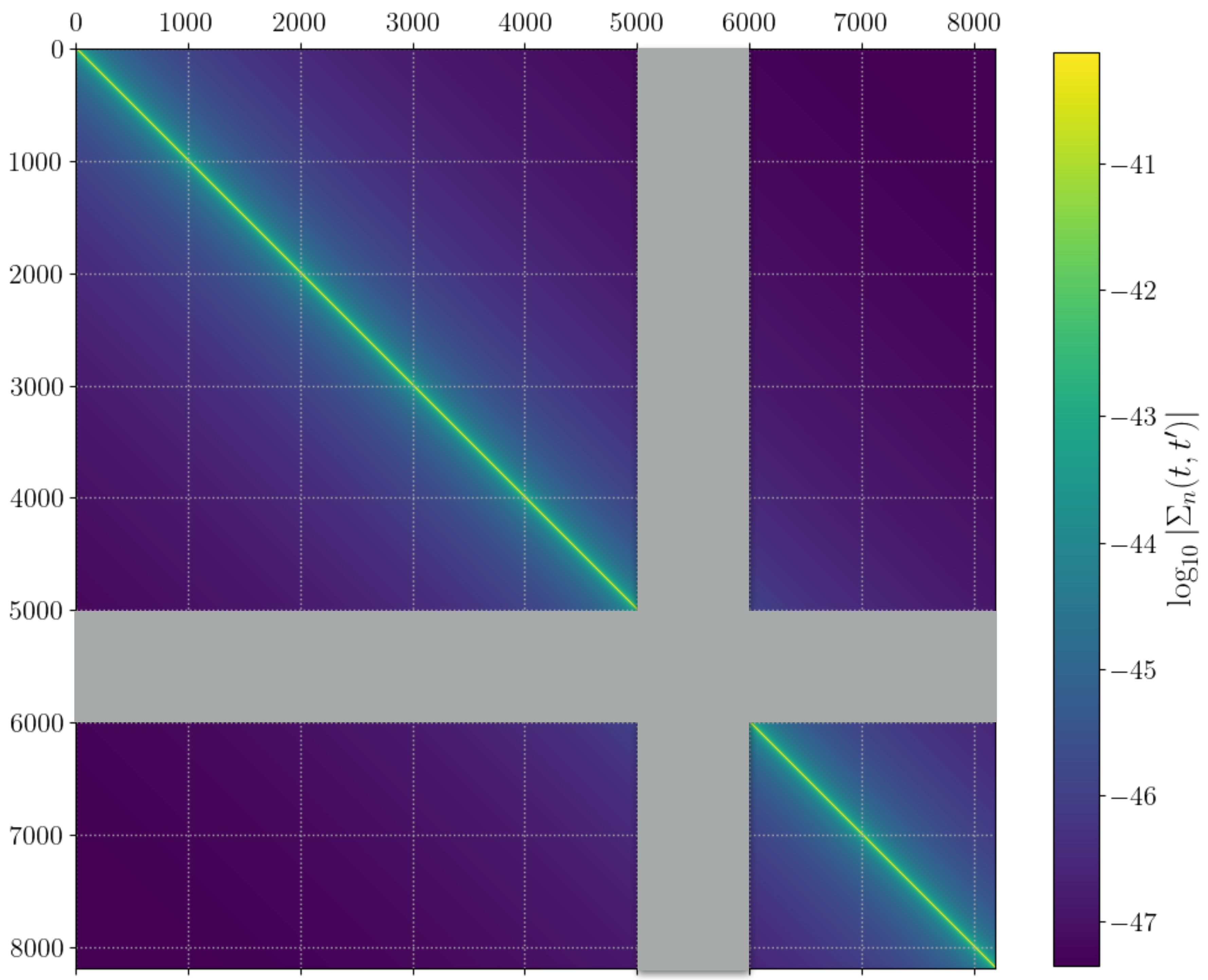
θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	1.0	1.1	1.1	1.2	1.1	1.1	1.1	1.1	1.1	1.1	1.1

TDIrescaled M=3e7, inj.: gap, model: nogap

θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	33.4	27.6	35.2	35.9	45.4	38.3	28.3	39.2	32.4	40.2	39.4

Results: biases from using the Whittle likelihood on gated data

Covariance with gap



- Correct modelling: pseudo-inverse of gated covariance

$$\ln \mathcal{L}_{\rm right} = -\frac{1}{2} X^T (W \Sigma W)^{+} X$$

- Incorrect modelling: Whittle for gated data and signal

$$\ln \mathcal{L}_{\rm wrong} = -\frac{1}{2} X^T W \Sigma^{-1} W X$$

Gap in the inspiral

TDI2 M=3e7, inj.: gap, model: nogap

θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	1.2	1.8	2.0	1.2	1.3	1.3	1.6	1.1	1.1	1.2	1.1

TDI1 M=3e7, inj.: gap, model: nogap

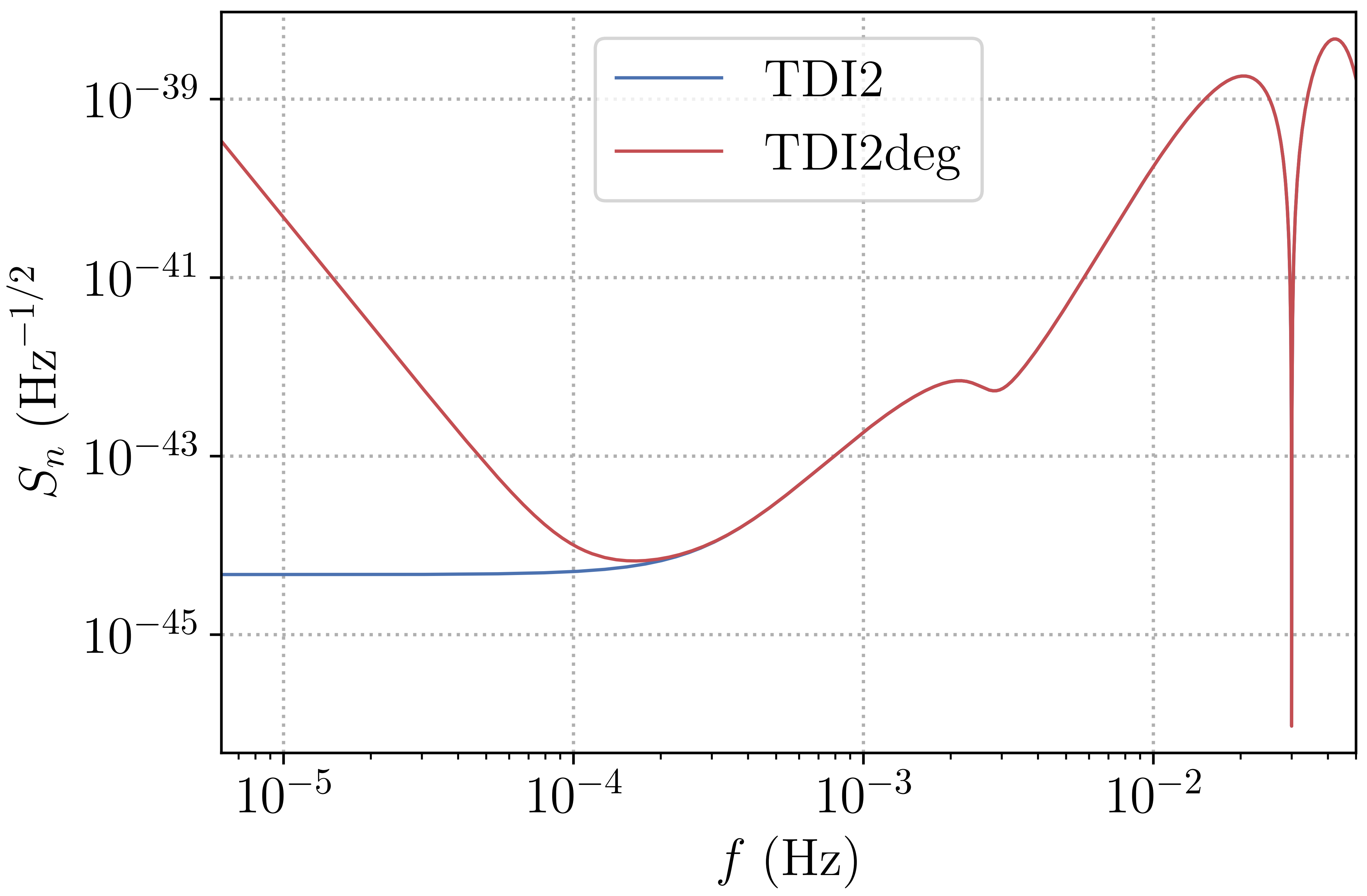
θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

TDIrescaled M=3e7, inj.: gap, model: nogap

θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	36.4	26.6	53.5	12.7	30.3	19.9	26.4	11.6	4.4	20.6	7.4

Question: how wrong is it to mismodel the covariance using Whittle ?

Results: mismodeling of low-frequency noise



Question: how wrong is it to mismodel the PSD at low frequencies ?

For a stationary process (no gap), the effect should be weak by frequency independence.

For a non-stationary process (gap), can the mismodeling error at low frequencies affect different frequencies where the signal is ?

TDI2 M=3e7, inj.: nogap, TDI2deg model: nogap, TDI2

θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	1.0	1.1	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

TDI2 M=3e7, inj.: gap, TDI2deg model: gap, TDI2

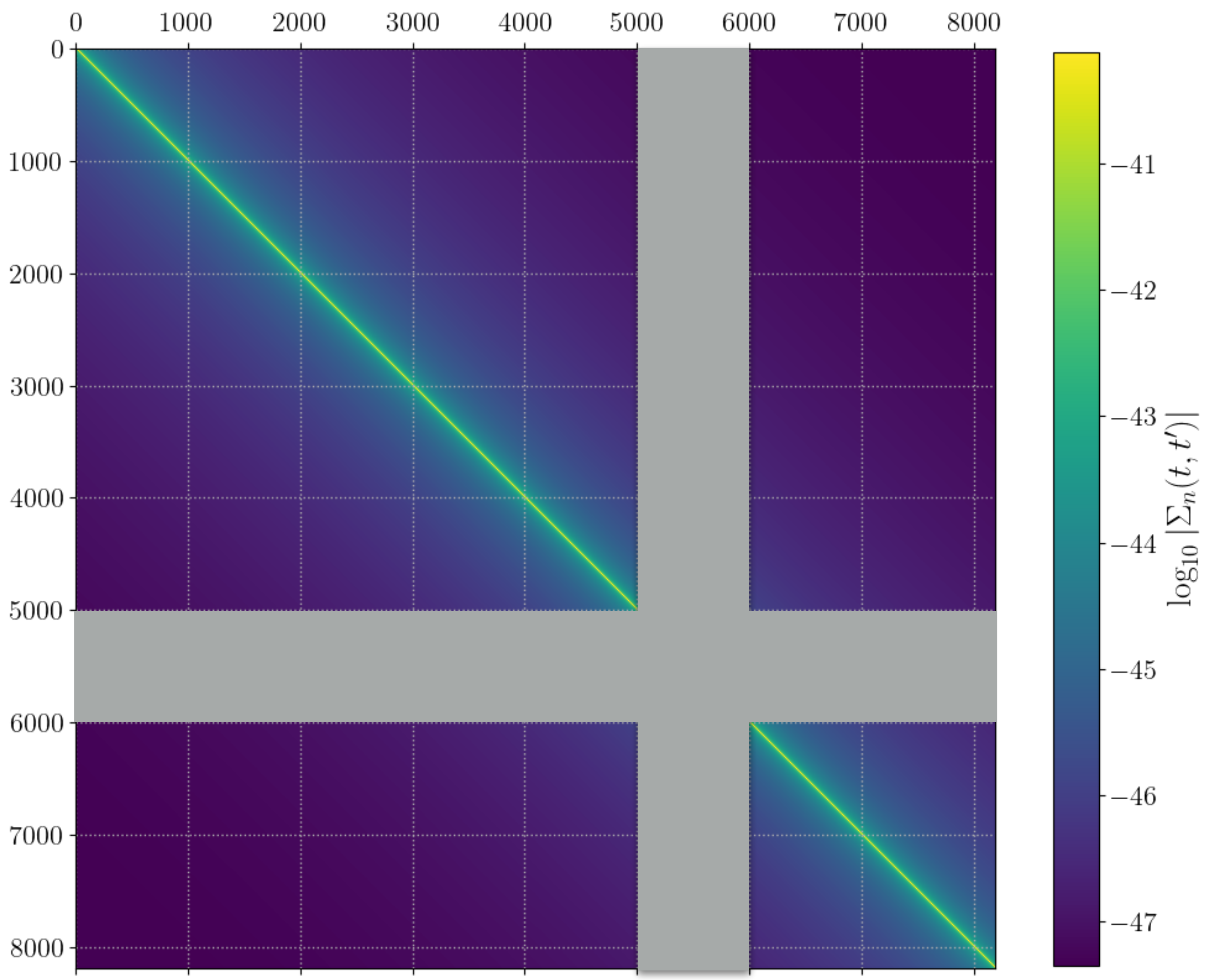
θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	2.4	2.7	1.7	2.1	2.0	2.3	1.5	2.3	1.9	2.7	2.3

TDI2 M=3e7, inj.: gap, TDI2deg model: nogap, TDI2

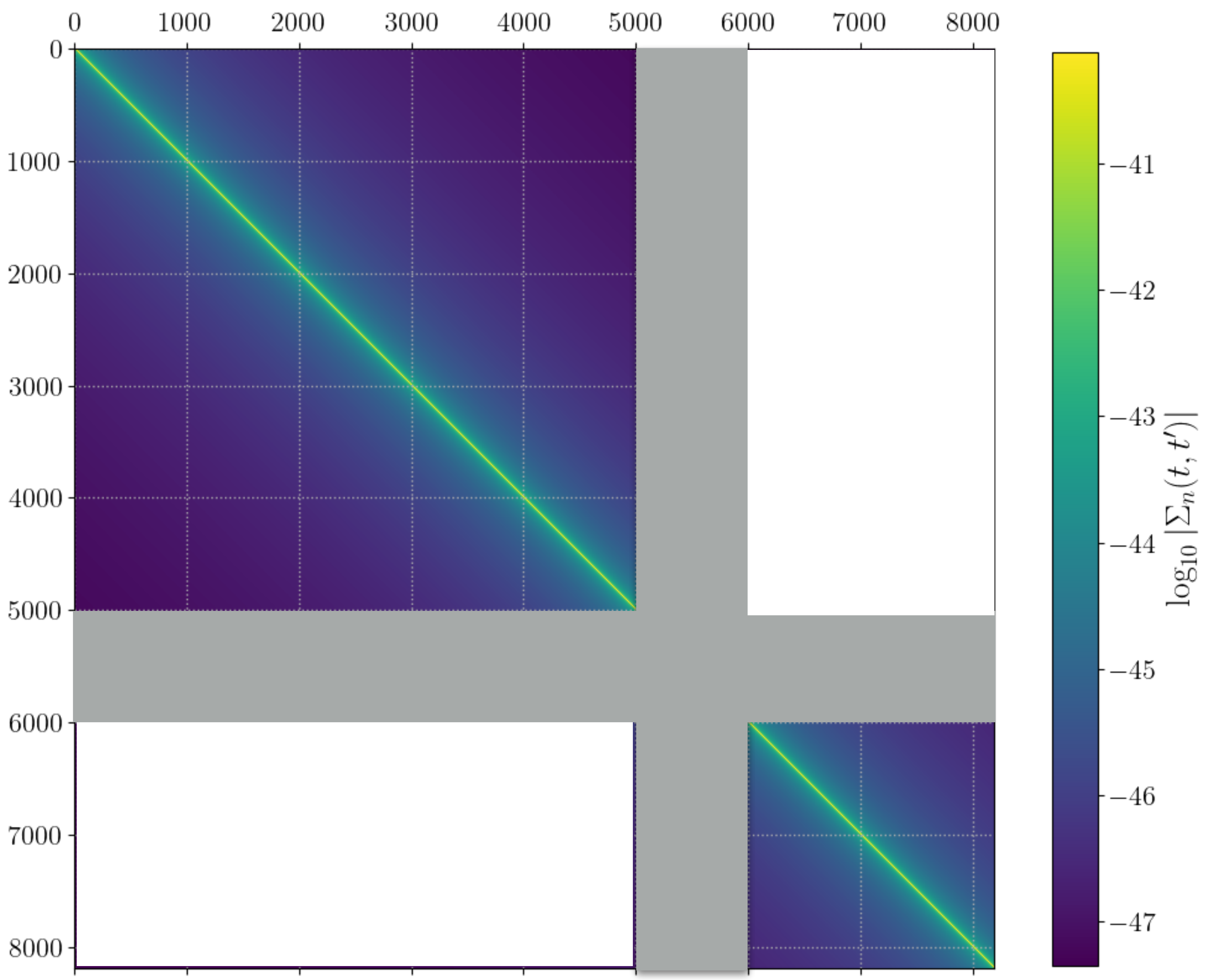
θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	8.5	11.5	9.9	11.9	11.7	10.3	11.4	11.8	10.2	11.8	10.6

Results: assumptions about segment independence

Covariance with gap, correlated segments



Covariance with gap, independent segments



Question: how wrong is it to mismodel data segments as correlated/independent ?

Only assuming correlations when the segments are independent gives $\Upsilon \neq 1$

TDI2 M=3e7, inj.: independent model: correlated

θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	1.5	2.0	1.0	1.2	1.1	1.1	1.4	1.1	1.0	1.2	1.1

TDIrescaled M=3e7, inj.: independent model: correlated

θ	M	q	χ_1	χ_2	Δt	D	ι	ϕ	λ	β	ψ
Υ	47.5	66.6	18.0	25.8	13.1	21.5	37.2	18.7	14.2	26.2	16.2

Results and outlook

Results

[Preliminary]

- Derived a framework handling missing data in time or frequency domain — **caveat**: for short segments only
- Derived a measure of inconsistency for the scatter of best-fit parameters
- Application to the LISA case: exploration of different mismodelling settings
- This framework is a test-bed: allows to test assumptions, possibility to compare to e.g. imputation

Outlook

- Longer signals ? Low-mass MBHBs, EMRIs ?
- Extension to other sources of non-stationarity
- More exploration to be done…