

Bias from small-scale leakage in Pulsar Timing Array maps

Federico Semenzato,^{a,b} Nicola Bellomo,^{a,b,c} Alvise Raccanelli,^{a,b,c}
Chiara M. F. Mingarelli^d

^aDipartimento di Fisica e Astronomia “G. Galilei”, Università degli Studi di Padova, via Marzolo 8, I-35131 Padova, Italy

^bINFN, Sezione di Padova, Via Marzolo 8, I-35131, Padova, Italy

^cINAF - Osservatorio Astronomico di Padova, Vicolo dell’Osservatorio 5, I-35122 Padova, Italy

^dDepartment of Physics, Yale University, New Haven, CT, 06520, USA

E-mail: federico.semenzato.1@phd.unipd.it, nicola.bellomo@unipd.it,
alvise.raccanelli.1@unipd.it, chiara.mingarelli@yale.edu

Abstract. Pulsar Timing Array experiments are rapidly approaching the era of gravitational wave background anisotropy detection. The timing residuals of each pulsar are an integrated measure of the gravitational-wave power across all angular scales. However, due to the limited number of monitored pulsars, current analyses are only able to reconstruct the angular structure of the background at large scales. We show analytically that this mismatch between the integrated all-sky signal and the truncated reconstruction introduces a previously unaccounted source of systematic bias in anisotropic background map reconstruction. The source of this systematic error, that we call “small-scale leakage”, is the intrinsic presence of unaccounted gravitational wave power at scales smaller than the reconstructed scales. This unmodeled power leaks into large-scale modes, artificially increasing the recovered value of the inferred angular power spectrum by at least one order of magnitude in a wide range of scales. Importantly, this effect is fundamentally independent of the geometry of the pulsar configuration, the anisotropy reconstruction method, the use of different regularization schemes, and the presence of pulsar noise. As the quality of pulsar timing array experiments improves, a robust understanding of small-scale leakage will become paramount for reliable detection and characterization of the gravitational wave background. Thus, the theoretical formalism developed here will be essential to estimate the magnitude of this systematic uncertainty in anisotropy searches.

Contents

1	Introduction	1
2	Pulsar timing array response to a gravitational wave background	3
3	Emergence of small-scale leakage	4
3.1	Map-making process	4
3.2	Small-scale leakage - Geometrical intuition	7
3.3	Small-scale leakage - Statistical intuition	8
3.4	Regularization scheme and interplay with small-scale leakage	10
4	Impact of pulsar timing noise	12
5	Discussion and conclusions	13
A	Persistence of small-scale residuals	15
B	Complementary material on small-scale leakage derivation	15
B.1	Geometrical intuition	15
B.2	Statistical intuition	17
B.3	Covariance matrix	18
C	Regularization schemes	20
C.1	Likelihood regularization	20
C.2	Covariance matrix regularization	20
C.3	Combining regularization schemes	20

1 Introduction

Anisotropies of the nanohertz gravitational wave background (GWB) contain invaluable information regarding its origin, whether astrophysical [1–4], cosmological [5–7] or a mixture of both. At this time, numerous Pulsar Timing Array (PTA) collaborations already reported evidence for a GWB [8–12]. However, despite a robust understanding of the capabilities of detecting its anisotropies [13–16], we have just started to gather hints about their presence [17, 18]. Their detection opens up the possibility of locating individual supermassive black hole binaries (SMBHBs) in well-localized host galaxies [15, 19–23] and, for future experiments, to cross-correlate the anisotropic GWB with galaxy surveys to firmly establish its origin [24–26].

However, measuring and interpreting the statistics of the GWB anisotropies still represents a formidable challenge [27–30]. Any PTA is comprised of a finite and sparsely distributed set of millisecond pulsars, and each pulsar timing residual is generated by the integrated GW power distributed across the entire sky. Thus, because of the limited number of pulsar pairs, we can reconstruct the GWB angular distribution only at the largest angular scales [30–32]. On the other hand, the actual GW sky, potentially shaped by a large number of SMBHBs, contains a structure that extends to rather small angular scales, which appears to be discarded from traditional analyses [15, 20].

Although it is not possible to draw any perfect analogy with the case discussed in this work, we note that cosmology has already dealt with potential issues regarding the interpretation of information coming from small-scales, in particular in the realm of simulation and data analysis. In the case of Cosmic Microwave Background, two instances immediately come to mind: the need to remove small-scale power at scales larger than the desired resolution of the analysis to avoid aliasing effects [33, 34], and the need to remove small-scale power when downgrading an existing map to a lower resolution [35]. Failure to perform such removal results in both cases in overestimating the magnitude of anisotropies

or, equivalently, of their statistical n -point functions. Additionally, it is well understood how aliasing effects appear when analyzing dark matter or galaxy catalogs to extract their power spectrum due to the finite resolution of the Fourier mode grid [36–38]. As in the previous case, the consequence is a conspicuous overestimation of the two-point function. The saving grace in all these instances is the possibility to directly access the map of spatial anisotropies, and to perform some form of “spatial smoothing” that removes small-scale power before performing any statistical analysis.

In a PTA, GWB anisotropies contribute to each correlated timing residual as an integrated signal over the sky. Consequently, power on small angular scales (although unreconstructable) is still embedded in every residual measurement. When the GWB sky is modeled only up to a finite multipole ℓ_{max} , this unresolved small-scale power is projected onto the available large-scale modes. As a result, forcing all measured power into low- ℓ modes inevitably biases the reconstruction, leading to an overestimation of the anisotropy amplitude and a distorted angular power spectrum.

We call this phenomenon *small-scale leakage*: it is a systematic bias that has not yet been properly investigated in the PTA context. The relevance of this effect is expected to become a fatal source of error in any anisotropy measurement, as PTA collaborations rapidly improve the precision of timing residual measurements, the number of monitored pulsars, and the total observation time. Therefore, without a rigorous understanding of this systematic bias and the development of effective mitigation strategies, this could emerge as an important challenge for future claims about the detection of anisotropies and their physical interpretation.

This work presents the first complete analytical framework for understanding and quantifying small-scale leakage in PTA anisotropy measurements. First, we demonstrate the presence and magnitude of small-scale power in timing residuals measurements. Then, we derive analytical expressions describing how unmodeled power propagates through the reconstruction process, revealing the large-scale modes contamination sourced by small-scale power. Furthermore, we explicitly show which interplay exists between small-scale leakage and regularization strategies implemented in data analyses. These results provide the theoretical foundation for the development of statistical estimators that account for such a systematic effect, thus allowing for the reconstruction of unbiased GWB sky maps.

This paper is organized as follows. Section 2 provides a quick overview of the basic concepts used in this work. Section 3 demonstrates the existence of small-scale leakage bias and its interplay with regularization schemes in the idealized case of noiseless measurements. Section 4 shows how the small-scale leakage effect persists even when noisy measurements are considered. Finally, we conclude in section 5 with a discussion of the implications of our findings for upcoming anisotropy searches. The appendices A, B, and C contain additional material that supports the discussion in the main text. In this work, we adopt natural units $c = G = 1$. A summary of our notation is provided in table 1.

Symbol	Description	Defining equation
R_p	Frequency-space timing residuals for pulsar p	(2.4)
$\mathcal{R}_{pq}$	Correlated residual for pulsar pair (p, q)	(2.7)
$\mathcal{R}$	Vector of correlated residuals for all pulsar pairs	(3.3)
$\mathcal{R}_{\text{LS}}, \mathcal{R}_{\text{SS}}$	Large/Small-Scale correlated residual	(3.7)
$\Gamma, \Gamma_{\text{LS}}, \Gamma_{\text{SS}}$	Geometry matrix (full; LS/SS blocks)	(3.4), (3.7)
$\mathbf{a}_{\text{LS}}$	LS harmonic coefficients estimator	(3.9), (3.15)
$\mathbb{P}_\ell$	Projection matrix selecting the $(2\ell + 1)$ LS block	(3.10), (B.4)
$M_{\ell\ell'}, M'_{\ell\ell'}, M''_{\ell\ell'}$	Mode-mixing leakage kernel	(3.11), (3.16), (3.20)
$\text{Cov}, \text{Cov}_{\text{reg}}$	Residual covariance and SVD-regularized version	(3.14), (4.3)
$\lambda, \mathcal{P}$	Ridge parameter and penalty matrix	(3.17)
$K_{\ell\ell'}$	Mode-mixing regularization-bias kernel	(3.19)
$H, \mathbf{v}$	Likelihood Hessian and score	(B.8)
e_{thres}	SVD eigenvalue cutoff	(C.4)

Table 1: Summary of the notation used in this paper in order of appearance.

2 Pulsar timing array response to a gravitational wave background

The metric fluctuation h_{ij} describing the GWB is usually decomposed in individual plane waves as

$$h_{ij}(t, \mathbf{x}) = \sum_A \int_{-\infty}^{\infty} df \int d\hat{\mathbf{n}} h_A(f, \hat{\mathbf{n}}) e_{ij}^A(\hat{\mathbf{n}}) e^{i2\pi f(t - \hat{\mathbf{n}} \cdot \mathbf{x})}, \quad (2.1)$$

where $(t, \mathbf{x})$ are the cosmic time and comoving coordinates, respectively, $(f, \hat{\mathbf{n}})$ are the GW frequency and propagation direction, respectively, $A = \{+, \times\}$ are the two polarization degrees of freedom, $e_{ij}^A(\hat{\mathbf{n}})$ is the polarization tensor, and $h_A(f, \hat{\mathbf{n}})$ is the mode amplitude. In the case of a stationary, Gaussian, and unpolarized background, we have

$$\langle h_A(f, \hat{\mathbf{n}}) h_{A'}^*(f', \hat{\mathbf{n}}') \rangle_s = \delta_{AA'}^K \delta^D(f - f') \delta^D(\hat{\mathbf{n}} - \hat{\mathbf{n}}') \frac{H(f, \hat{\mathbf{n}})}{8\pi}, \quad (2.2)$$

where $\langle \cdot \rangle_s$ indicates the ensemble average over the properties of the population of sources (masses, spins, eccentricities, etc.), $*$ indicates the complex conjugate, δ^K and δ^D are Kronecker and Dirac deltas, respectively, and $H(f, \hat{\mathbf{n}})$ is the (one-sided) GWB power spectrum. In this work, we consider a GWB produced by a large number of SMBHBs stochastically distributed in host galaxies. Therefore, since their spatial distribution is intrinsically anisotropic, the GWB power spectrum retains a dependence on the direction. Although in this work we are interested in the “weak field” regime, where the anisotropies in the emitted power in GWs are expected to be small since they are not generated by individual “loud sources”, the effect discussed in this paper exists for every form of the GWB power spectrum. Under these assumptions, we can conveniently factorize the power spectrum as

$$H(f, \hat{\mathbf{n}}) = H(f) [1 + P(f, \hat{\mathbf{n}})], \quad (2.3)$$

where $H(f)$ is the isotropic value, $P(f, \hat{\mathbf{n}})$ characterizes the anisotropic angular distribution of GWB, and it has zero mean across the entire sky, i.e., $\int d\hat{\mathbf{n}} P(f, \hat{\mathbf{n}}) = 0$. In this notation, the physically meaningful values of the anisotropic power are $P(f, \hat{\mathbf{n}}) \geq -1$.

The presence of a GWB causes a delay in the arrival time of pulsar electromagnetic signals. The induced, frequency-dependent timing residual in the arrival time from a pulsar p at distance D_p in the direction $\hat{\mathbf{p}}$ reads as

$$R_p(f) = \frac{1}{2\pi i f} \int d\hat{\mathbf{n}} \sum_A h_A(f, \hat{\mathbf{n}}) F_p^A(\hat{\mathbf{n}}) \left[1 - e^{-2\pi i f D_p (1 + \hat{\mathbf{n}} \cdot \hat{\mathbf{p}})} \right], \quad (2.4)$$

where the pulsar antenna pattern function is given by

$$F_p^A(\hat{\mathbf{n}}) = \frac{\hat{p}^i \hat{p}^j}{2(1 + \hat{\mathbf{n}} \cdot \hat{\mathbf{p}})} e_{ij}^A(\hat{\mathbf{n}}). \quad (2.5)$$

Therefore, the expectation value of the timing residual coming for a pair of different pulsars (p, q) reads as

$$\langle R_p(f) R_q^*(f') \rangle_s = \frac{\delta^D(f - f')}{2} [\mathcal{R}_{pq}^{\text{iso}} + \mathcal{R}_{pq}^{\text{anis}}], \quad (2.6)$$

where we define as “isotropic” and “anisotropic” correlated residuals those contributions sourced by the isotropic and anisotropic components of the GWB, respectively. They are formally defined as

$$\begin{aligned} \mathcal{R}_{pq}^{\text{iso}} &= \frac{H(f)}{(2\pi f)^2} \int \frac{d\hat{\mathbf{n}}}{4\pi} \gamma_{pq}(\hat{\mathbf{n}}) = \frac{H(f)}{(2\pi f)^2} \text{HD}(\hat{\mathbf{p}} \cdot \hat{\mathbf{q}}), \\ \mathcal{R}_{pq}^{\text{anis}} &= \frac{H(f)}{(2\pi f)^2} \int \frac{d\hat{\mathbf{n}}}{4\pi} P(f, \hat{\mathbf{n}}) \gamma_{pq}(\hat{\mathbf{n}}), \end{aligned} \quad (2.7)$$

where the overlap reduction functions read as

$$\gamma_{pq}(\hat{\mathbf{n}}) = \sum_A F_p^A(\hat{\mathbf{n}}) F_q^A(\hat{\mathbf{n}}) \left[1 - e^{-2\pi i f D_p (1 + \hat{\mathbf{n}} \cdot \hat{\mathbf{p}})} \right] \left[1 - e^{2\pi i f D_q (1 + \hat{\mathbf{n}} \cdot \hat{\mathbf{q}})} \right] \approx \sum_A F_p^A(\hat{\mathbf{n}}) F_q^A(\hat{\mathbf{n}}), \quad (2.8)$$

since the Earth–pulsar phase factors in square brackets give a rapidly oscillatory contribution that can be neglected in the long-arm limit [39], and

$$\text{HD}(\hat{\mathbf{p}} \cdot \hat{\mathbf{q}}) = \int \frac{d\hat{\mathbf{n}}}{4\pi} \gamma_{pq}(\hat{\mathbf{n}}) = \frac{1}{3} - \frac{1 - \hat{\mathbf{p}} \cdot \hat{\mathbf{q}}}{12} + \frac{1 - \hat{\mathbf{p}} \cdot \hat{\mathbf{q}}}{2} \log \frac{1 - \hat{\mathbf{p}} \cdot \hat{\mathbf{q}}}{2} \quad (2.9)$$

is the Hellings-Downs curve [40].¹

3 Emergence of small-scale leakage

For the sake of clarity, we demonstrate the existence of a purely physical effect called “small-scale leakage” in the case of noiseless timing residual measurements. Pulsar noise is introduced in section 4, alongside with a discussion of its impact on small-scale leakage.

3.1 Map-making process

Different clustering properties of GW-emitting sources are expected to have a different impact on the magnitude of anisotropic correlated timing residuals. Since these properties are typically scale dependent, it is convenient to decompose the GWB power anisotropic distribution appearing in equation (2.7) in a basis that makes explicit how different scales contribute to the correlated residual. The considerations raised in this article are independent from the particular choice of basis; however, in order to present a fully worked-out scenario, here we adopt a real spherical harmonic decomposition. In this basis, the GWB power angular distribution reads as

$$P(f, \hat{\mathbf{n}}) = \sum_{\ell m} a_{\ell m}^f \mathcal{Y}_{\ell m}(\hat{\mathbf{n}}), \quad (3.1)$$

where $\mathcal{Y}_{\ell m}$ are the real spherical harmonics, and $a_{\ell m}^f$ are the frequency-dependent harmonic coefficients.² These coefficients represent a true random variable associated with the stochasticity of the position of the sources. In other words, even considering a fixed set of GW sources, the scenario presents a residual degree of stochasticity given by how GW sources populate different host galaxies. The mean and variance of the harmonic coefficients read as

$$\langle a_{\ell m}^f \rangle = 0, \quad \langle a_{\ell m}^f a_{\ell' m'}^{f'} \rangle = \delta_{\ell \ell'} \delta_{m m'} C_{\ell}^{ff'}, \quad (3.2)$$

where the angular power spectrum $C_{\ell}^{ff'} \neq 0$ even when $f \neq f'$ when the background is sourced by eccentric SMBHBs [41, 42], and, more generally, because the spatial distribution of the hosts of GW sources emitting in different frequency bands is correlated. The angle brackets $\langle \cdot \rangle$ appearing in equation (3.2) indicate averaging over the $a_{\ell m}$ ensemble, i.e., over all the possible spatial distributions of the same SMBHB population across different galaxies. Finally, from now on, we consider only multipoles $\ell \geq 2$ because $P(f, \hat{\mathbf{n}})$ has no isotropic component by construction and, in the weak field regime, the value of the harmonic coefficients with $\ell = 1$ is completely dominated by the kinetic dipole.

Recovering the angular power spectrum that characterizes the anisotropic distribution of GW sources is tightly connected to our ability to reconstruct the $P(f, \hat{\mathbf{n}})$ map. Suppose to monitor a set of N_{psr} pulsars, and to have obtained a set of correlated timing residual $\mathcal{R}_{pq}^{\text{anis}}$ for each one of the $N_{\text{pairs}} = N_{\text{psr}}(N_{\text{psr}} - 1)/2$ pulsar pairs of the sample. Consider also a single frequency bin for

¹Multiple definitions of Hellings-Downs are currently used in the literature. Our definition is related to the one employed by the NANOGrav Collaboration by a rescaling factor as $\text{HD}_{\text{NANOGrav}}(\hat{\mathbf{p}} \cdot \hat{\mathbf{q}}) = (3/2)\text{HD}(\hat{\mathbf{p}} \cdot \hat{\mathbf{q}})$.

²Real spherical harmonics are defined in terms of spherical harmonics $Y_{\ell m}$ as $\mathcal{Y}_{\ell m} = \sqrt{2}\text{Re}(Y_{\ell m})$ ($m > 0$), $\mathcal{Y}_{\ell 0} = Y_{\ell 0}$ ($m = 0$), and $\mathcal{Y}_{\ell m} = \sqrt{2}\text{Im}(Y_{\ell |m|})$ ($m < 0$).

simplicity. Data can then be organized in a vectorial structure as [31, 32, 43]

$$\mathcal{R} = \begin{pmatrix} \mathcal{R}_{p_1 p_2}^{\text{anis}} \\ \mathcal{R}_{p_1 p_3}^{\text{anis}} \\ \vdots \\ \mathcal{R}_{p_{N-1} p_N}^{\text{anis}} \end{pmatrix} = \begin{pmatrix} \Gamma_{2-2}^{p_1 p_2} & \Gamma_{2-1}^{p_1 p_2} & \Gamma_{20}^{p_1 p_2} & \dots & \Gamma_{\ell m}^{p_1 p_2} & \dots \\ \vdots & & \vdots & & \vdots & \\ \vdots & & \vdots & & \vdots & \\ \Gamma_{2-2}^{p_{N-1} p_N} & \Gamma_{2-1}^{p_{N-1} p_N} & \Gamma_{20}^{p_{N-1} p_N} & \dots & \Gamma_{\ell m}^{p_{N-1} p_N} & \dots \end{pmatrix} \begin{pmatrix} a_{2-2} \\ a_{2-1} \\ a_{20} \\ \vdots \\ a_{\ell m} \\ \vdots \end{pmatrix} = \Gamma \mathbf{a}, \quad (3.3)$$

where $\mathcal{R}$ is a $(N_{\text{pairs}} \times 1)$ column vector containing the correlated residuals, the elements of the Γ matrix, using equations (2.7) and (3.1), read as

$$\Gamma_{\ell m}^{p_i p_j} = \frac{H(f)}{(2\pi f)^2} \int \frac{d\hat{\mathbf{n}}}{4\pi} \gamma_{p_i p_j}(\hat{\mathbf{n}}) \mathcal{Y}_{\ell m}(\hat{\mathbf{n}}), \quad (3.4)$$

and $\mathbf{a}$ is a column vector that contains the harmonic coefficients. Since it is not known a priori how the anisotropic power is distributed on different scales, the dimension of the $\mathbf{a}$ vector, and thus the number of columns of the Γ matrix, has the potential of being extremely large. On the other hand, we probe a finite amount of pulsar pairs, therefore we have the ability of constraining at most N_{pairs} real spherical harmonic coefficients or, equivalently, multipoles up to a maximum reconstructed scale $\ell_{\text{max}}^{\text{rec}} = \left\lfloor \sqrt{\frac{N_{\text{psr}}(N_{\text{psr}}-1)}{2}} + 4 - 1 \right\rfloor$ [30, 31], contrariwise to what reported in ref. [44].

Regarding the structure of the $\mathbf{a}$ vector, suppose that we have a GWB with anisotropic power up to some maximum scale $\ell_{\text{max}}^{\text{GWB}} > \ell_{\text{max}}^{\text{rec}}$, i.e., that $a_{\ell m} = 0$ for every multipole $\ell > \ell_{\text{max}}^{\text{GWB}}$. In this scenario, we can divide the set of harmonic coefficients into a set of “large-scale” (LS) coefficients with multipoles $2 \leq \ell \leq \ell_{\text{max}}^{\text{rec}}$, and a set of “small-scale” (SS) coefficients with multipoles $\ell_{\text{max}}^{\text{rec}} + 1 \leq \ell \leq \ell_{\text{max}}^{\text{GWB}}$, that we then arrange in two vectors as

$$\mathbf{a}_{\text{LS}} = \begin{pmatrix} a_{2-2} \\ a_{2-1} \\ \vdots \\ a_{\ell_{\text{max}}^{\text{rec}} \ell_{\text{max}}^{\text{rec}} - 1} \\ a_{\ell_{\text{max}}^{\text{rec}} \ell_{\text{max}}^{\text{rec}}} \end{pmatrix}, \quad \mathbf{a}_{\text{SS}} = \begin{pmatrix} a_{\ell_{\text{max}}^{\text{rec}} + 1 - (\ell_{\text{max}}^{\text{rec}} + 1)} \\ a_{\ell_{\text{max}}^{\text{rec}} + 1 - (\ell_{\text{max}}^{\text{rec}} + 1) + 1} \\ \vdots \\ a_{\ell_{\text{max}}^{\text{GWB}} \ell_{\text{max}}^{\text{GWB}} - 1} \\ a_{\ell_{\text{max}}^{\text{GWB}} \ell_{\text{max}}^{\text{GWB}}} \end{pmatrix}, \quad (3.5)$$

where the dimension of each vector is

$$N_{\text{LS}} = \dim(\mathbf{a}_{\text{LS}}) = (\ell_{\text{max}}^{\text{rec}} + 1)^2 - 4, \quad N_{\text{SS}} = \dim(\mathbf{a}_{\text{SS}}) = (\ell_{\text{max}}^{\text{GWB}} + 1)^2 - (\ell_{\text{max}}^{\text{rec}} + 1)^2, \quad (3.6)$$

respectively. Under these assumptions, equation (3.3) nicely factorizes as

$$\boxed{\mathcal{R} = \mathcal{R}_{\text{LS}} + \mathcal{R}_{\text{SS}} = \Gamma_{\text{LS}} \mathbf{a}_{\text{LS}} + \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}} \quad (3.7)$$

where Γ_{LS} contains the first $\ell_{\text{max}}^{\text{rec}} - 1$ columns of Γ , while Γ_{SS} contains the second $\ell_{\text{max}}^{\text{GWB}} - \ell_{\text{max}}^{\text{rec}}$ ones. For later convenience, we define the correlated residuals coming from large- and small-scale modes only as $\mathcal{R}_{\text{LS}}$ and $\mathcal{R}_{\text{SS}}$, respectively.

The core idea of this work is contained in equation (3.7). Correlated timing residuals generated by the anisotropic distribution of sources are integrated across all scales, even if limited samples of pulsars allow for the reconstruction of only a limited subset of harmonic coefficients. Therefore, the extra small-scale power is unavoidably incorporated into the large-scale harmonic coefficients when we interpret the total correlated residuals as a large-scale only contribution by reconstructing the map up to $\ell_{\text{max}}^{\text{rec}} < \ell_{\text{max}}^{\text{GWB}}$. In other words, small-scale power leaks into large-scale modes when we attempt to perform a partial reconstruction of the anisotropic GWB power map, artificially increasing large-scale mode variance, i.e., their C_{ℓ} s.

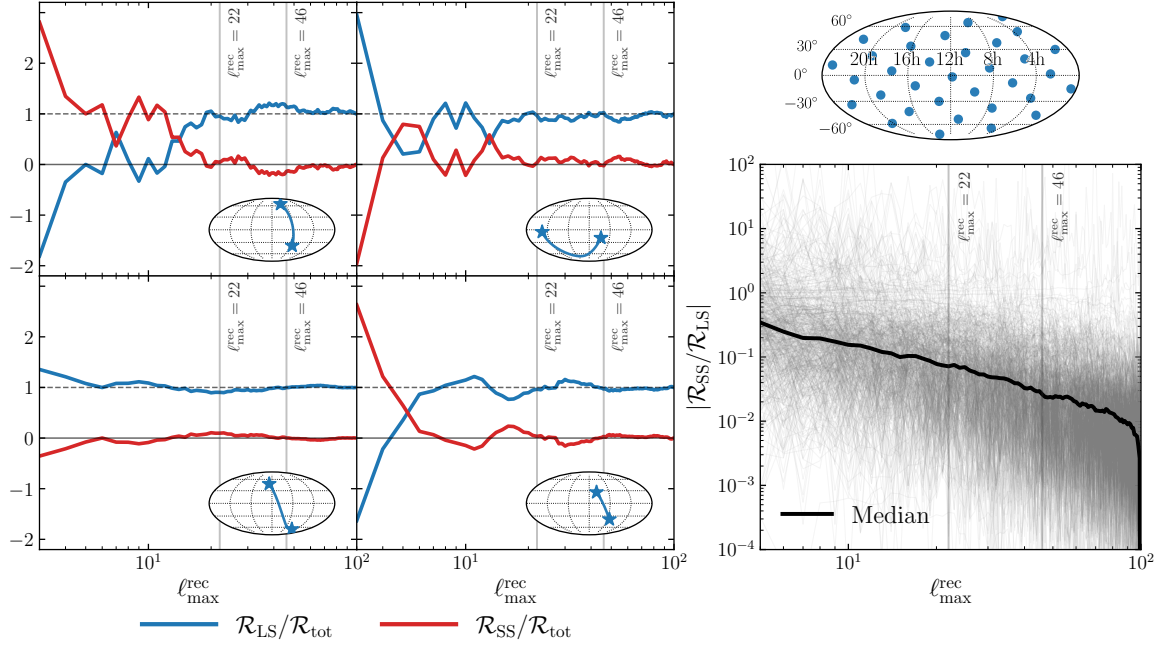


Figure 1: *Left panels:* fractional contribution of large-scale (blue lines) and small-scale (red lines) correlated residuals for four randomly displaced pulsars as a function of the reconstruction multipole $\ell_{\text{max}}^{\text{rec}}$. For $\ell_{\text{max}}^{\text{rec}} = 0$ (monopole only), the entire signal originates from small scales. As $\ell_{\text{max}}^{\text{rec}}$ increases, progressively more power shifts into the large-scale component. The sum of both components, i.e., the total correlated residual, remains constant when varying $\ell_{\text{max}}^{\text{rec}}$. *Right panel:* absolute value of the small- to large-scale correlated residuals ratio for all pulsar pairs in the I34 configuration. The black line represents the median of the absolute value of the ratio.

Although the systematic bias effect presented in this work holds for every injected GWB, our proof-of-principle analysis considers exclusively the scenario where the GWB is sourced by a population of SMBHBs spatially distributed in galaxies, i.e., of a GWB with anisotropies resembling those of galaxy distributions. Anisotropic maps of the GWB are generated using the 10^3 realizations and the pipeline created in ref. [25], and are compatible with the scenario in which the GWB is generated by a population of SMBHBs emitting at 4 nHz, with no loud continuous-wave source present in the signal. This set of realizations is compatible with the weak-field regime assumption, and the inferred $C_\ell \simeq 3 \times 10^{-4}$ follows the expected Wishart distribution [25]. Despite the fact that GWB maps created with this approach presents power even at scales $\ell \gtrsim 10^3$, for the sake of reducing computational cost, we consider maps where we truncate it at $\ell_{\text{max}}^{\text{GWB}} = 250$. In other words, all the results should to be interpreted as conservative, since there is considerably more power at small-scales than the one used to perform this analysis. Regarding the pulsar distribution, we consider three sets of geometries with a different number of pulsars to showcase the robustness of the physical effect discussed in this article with respect to observational limitations. These configurations are labeled as

- **I34, I68:** two sets of $N_{\text{psr}} = 34$ and $N_{\text{psr}} = 68$ pulsars *isotropically* distributed on the sky, respectively;
- **U34, U68:** two sets of $N_{\text{psr}} = 34$ and $N_{\text{psr}} = 68$ pulsars *uniformly* distributed on the sky, respectively;
- **NG34, NG68:** a subset of $N_{\text{psr}} = 34$ 15-year dataset NANOGrav pulsars, and the full sample.

Given these numbers of pulsars, the maximum reconstruction multipole is $\ell_{\text{max}}^{\text{rec}} = 22$ ($\ell_{\text{max}}^{\text{rec}} = 46$ for the {I34, U34, NG34} ({I68, U68, NG68}) configurations).

In the left panel of figure 1 we show the relative importance of the small- and large-scale coefficients for four random pairs of pulsars as we change the maximum reconstruction multipole. Although in the limit of large $\ell_{\max}^{\text{rec}}$ the small-scale residuals carry a negligible amount of power, the same cannot be said for $\ell_{\max}^{\text{rec}} \lesssim 10$, where the large pile-up of small-scale power is clearly exacerbated. These examples are not peculiar, as we show in the right panel of the same figure, where we report the ratio of small- to large-scale correlated residuals for all the pulsar pairs of the I34 configuration. In particular, we observe that small- and large-scale residuals easily have the same order of magnitude for plausible values of the maximum reconstruction multipole. Moreover, this effect is fundamentally independent of the pulsar geometry considered, as explicitly shown in appendix A for the other pulsar configurations.

3.2 Small-scale leakage - Geometrical intuition

There are multiple fashions in which it is possible to establish the consequences of assigning small-scale power to large-scale modes. An intuitive approach is to view the issue as a purely geometrical map-making process, where we attempt to directly reconstruct the GWB anisotropic map by implementing a “minimum least squares” estimator. Suppose to look for a set of large-scale only correlated residuals that match the “true” ones $\mathcal{R}$ via the chi-squared statistics

$$\chi_{\text{LMS}}^2 = \min_{\mathbf{a}_{\text{LS}}} (\mathcal{R} - \Gamma_{\text{LS}} \mathbf{a}_{\text{LS}})^2 = \min_{\mathbf{a}_{\text{LS}}} (\Gamma_{\text{LS}} \mathbf{a}_{\text{LS}}^{\text{true}} + \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}} - \Gamma_{\text{LS}} \mathbf{a}_{\text{LS}})^2. \quad (3.8)$$

It is commonly known that the solution to this minimization problem is given by

$$\mathbf{a}_{\text{LS}} = \mathbf{a}_{\text{LS}}^{\text{true}} + \Gamma_{\text{LS}}^+ \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}}, \quad (3.9)$$

where $\Gamma_{\text{LS}}^+ = (\Gamma_{\text{LS}}^T \Gamma_{\text{LS}})^{-1} \Gamma_{\text{LS}}^T$ is the Moore-Penrose inverse, and T indicates the transpose (see also appendix B for a derivation of this result). The set of harmonic coefficients that minimize the χ_{LMS}^2 statistics is not the set of “true” large-scale ones due to a contamination of small-scale modes. This is what we refer to as “small-scale leakage”.

The implications of this fact can be immediately understood by constructing an estimator of the angular power spectrum. Consider a vector $\mathbf{a}_{\ell}^T = (a_{\ell - \ell}, \dots, a_{\ell m}, \dots, a_{\ell \ell}) = \mathbb{P}_{\ell} \mathbf{a}_{\text{LS}}$ containing all the harmonic coefficients corresponding to a given multipole ℓ , where the explicit form of the projection matrix $\mathbb{P}_{\ell}$ is reported in appendix B. For each multipole, the estimator of the angular power spectrum reads as

$$\hat{C}_{\ell} = \frac{\mathbf{a}_{\ell}^T \mathbf{a}_{\ell}}{2\ell + 1} = \frac{|\mathbb{P}_{\ell} \mathbf{a}_{\text{LS}}^{\text{true}} + \mathbb{P}_{\ell} \Gamma_{\text{LS}}^+ \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}}|^2}{2\ell + 1}. \quad (3.10)$$

Therefore, as we show in appendix B, the expectation value of this estimator is given by

$$C_{\ell}^{\text{rec}} = \langle \hat{C}_{\ell} \rangle = C_{\ell}^{\text{true}} + C_{\ell}^{\text{leakage}}, \quad (3.11)$$

where C_{ℓ}^{rec} and C_{ℓ}^{true} are the reconstructed and true large-scale angular power spectrum, respectively, and the leakage contribution reads as

$$C_{\ell}^{\text{leakage}} = \frac{1}{2\ell + 1} \langle |\mathbb{P}_{\ell} \Gamma_{\text{LS}}^+ \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}}|^2 \rangle = \sum_{\ell' = \ell_{\max}^{\text{rec}} + 1}^{\ell_{\max}^{\text{GWB}}} M_{\ell \ell'} C_{\ell'}^{\text{true}}, \quad (3.12)$$

and $M_{\ell \ell'}$ is a *mode-mixing* function defined in appendix B that depends exclusively on the geometry of the pulsar distribution. This additional variance is sourced exclusively by the power at small scales, as the upper and lower bounds of the summation indicate. Any term containing mixed large- and small-scale harmonic coefficients has a zero expectation value since the different multipoles are not correlated.

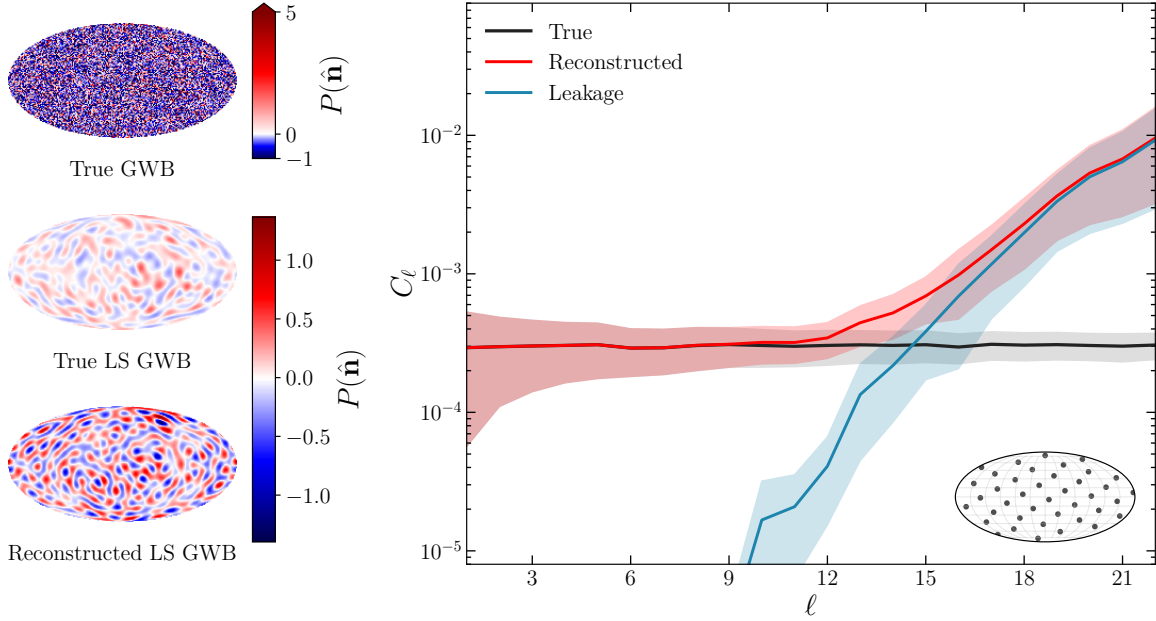


Figure 2: *Left panels:* individual realization of a true GWB with power up to $\ell_{\text{max}}^{\text{GWB}} = 250$ (*top panel*), its large-scale component with power up to $\ell_{\text{max}}^{\text{rec}} = 22$ (*central panel*), and reconstructed large-scale background, including small-scale leakage, up to $\ell_{\text{max}}^{\text{rec}} = 22$ (*bottom panel*). *Right panel:* envelopes of true (*black*), reconstructed (*red*), and leakage (*blue*) correlated residual angular power spectra for 10^3 realizations of the GWB and the I34 pulsar geometry. Solid lines indicate the median value of angular power spectra.

In figure 2 we compare the reconstructed (biased) angular power spectra with the true one for 10^3 realizations, and showcase that the additional variance is precisely described by the small-scale leakage formula derived in equation (3.12). The relevance of this effect is displayed also on the left panels of the figure, where we show how the reconstructed map obtained with the harmonic coefficients of equation (3.9) is significantly different from the original one, and it even reaches non physical values of $P(\hat{\mathbf{n}}) < -1$ because of the small-scale contamination. The importance of this effect cannot be understated: due to the unmodeled small-scale power, the reconstructed large-scale angular power spectrum is overestimated by at least one order of magnitude with respect to its true value, far exceeding the level of uncertainty associated with cosmic variance $\sigma_{\text{CV}}/C_\ell = \sqrt{2/(2\ell+1)}$, which is of order 20% and 15% for $\ell = 22$ and $\ell = 46$, respectively.

The origin of this contamination can be traced back to the mode-mixing function $M_{\ell\ell'}$ showed in figure 3 for the I34 pulsar geometry. The structure of this function is such that large-scale multipoles far from $\ell_{\text{max}}^{\text{rec}}$ are minimally affected by small-scale leakage. However, the magnitude of the mode-mixing function increases rapidly the closer we get to $\ell_{\text{max}}^{\text{rec}}$, reaching and exceeding values of order unity, driven mainly by small-scale multipoles close to $\ell_{\text{max}}^{\text{rec}}$. This structure is consistent across different pulsar geometries, as we show in appendix B.

3.3 Small-scale leakage - Statistical intuition

The geometric picture presented above has, in reality, a deeper implication in terms of statistical inference of the harmonic coefficients, as in a likelihood-based analysis. Harmonic coefficients in the weak field regime have a Gaussian distribution; therefore, we can introduce a Gaussian likelihood statistics as [31]

$$-2 \log \mathcal{L}(\mathbf{a}_{\text{LS}}|\mathbf{R}) = (\mathbf{R} - \mathbf{R}^{\text{th}})^T \text{Cov}^{-1} (\mathbf{R} - \mathbf{R}^{\text{th}}), \quad (3.13)$$

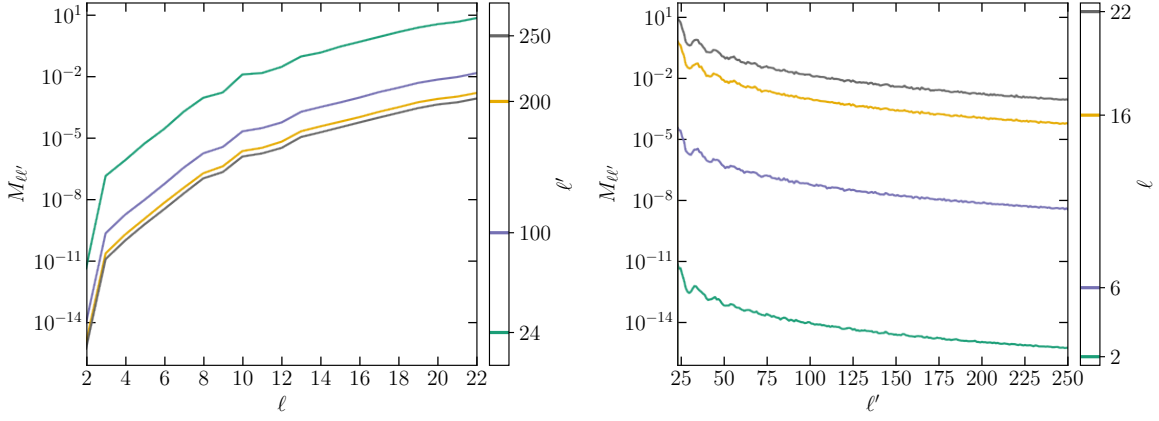


Figure 3: Value of the mode-mixing function $M_{\ell\ell'}$ from equation (3.12) in terms of large-scale ℓ and small-scale ℓ' multipoles. This matrix quantifies how the unmodeled scales ℓ' bias the reconstructed C_ℓ (see equation (3.11)). In both panels we show how low large-scale multipoles are almost unaffected by small-scale leakage, however the situation radically changes for multipoles close to $\ell_{\text{max}}^{\text{rec}}$, where the order of magnitude of the mode-mixing function is around order unity.

where $\mathcal{R}^{\text{th}}$ is the vector of theoretical correlated residuals, and $\text{Cov} = \left\langle \left(\mathcal{R} - \mathcal{R}^{\text{th}} \right) \left(\mathcal{R} - \mathcal{R}^{\text{th}} \right)^T \right\rangle$ is a covariance matrix.³ In the pure signal limit, as shown in appendix B, the elements of the covariance matrix read as

$$\text{Cov}_{i_{pq}j_{rs}} = \sum_{\ell m} \left(\Gamma_{\ell m}^{pr} \Gamma_{\ell m}^{qs} + \Gamma_{\ell m}^{ps} \Gamma_{\ell m}^{qr} \right) (C_\ell + N_\ell^{\text{shot}}), \quad (3.14)$$

where N_ℓ^{shot} is the Poissonian shot-noise term due to the discreteness of the SMBHBs, and with the composite indices i_{pq} we explicitly indicate which pulsar pair corresponds to each row/column of the covariance matrix.

As we demonstrate in appendix B, under the incorrect assumption that correlated residuals are described only by large-scale modes, i.e., that $\mathcal{R}^{\text{th}} = \Gamma_{\text{LS}} \mathbf{a}_{\text{LS}}$, we observe a shift in the maximum likelihood position given by

$$\mathbf{a}_{\text{LS}} = \mathbf{a}_{\text{LS}}^{\text{true}} + \left(\Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{LS}} \right)^{-1} \Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}}. \quad (3.15)$$

Therefore, also in this formulation of the issue, the expectation value of the reconstructed harmonic coefficients contains a leakage term that reads as

$$C_\ell^{\text{leakage}} = \frac{\left\langle \left| \mathbb{P}_\ell \left(\Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{LS}} \right)^{-1} \Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}} \right|^2 \right\rangle}{2\ell + 1} = \sum_{\ell'=\ell_{\text{max}}^{\text{rec}}+1}^{\ell_{\text{max}}^{\text{GWB}}} M'_{\ell\ell'} C_{\ell'}^{\text{true}}. \quad (3.16)$$

Since the form and structure of the mode-mixing function $M'_{\ell\ell'}$ is similar to that presented in figure 3, we choose to report them both in appendix B. Moreover, since the case presented in section 3.2 can be exactly recovered by assuming that $\text{Cov} = I_{N_{\text{pairs}}}$, in the following we focus only on this interpretation, drawing a connection to the simple map-making process only when needed.

³The likelihood in equation (3.13) is used exclusively to determine the harmonic coefficients that best match the realization of the Universe in which we live in. It does not serve the purpose of estimating the posterior of the theoretical model parameters used to generate our set of realizations. Therefore, the extra term “log detCov” that should appear in the case of Gaussian likelihoods has been removed from equation (3.13) since it is just a fixed additive constant for the purpose at hand.

3.4 Regularization scheme and interplay with small-scale leakage

Although the analysis presented above is statistically well posed, in reality its numerical structure is ill-conditioned. In particular, we can foresee two potential sources of numerical instability in the likelihood analysis which might affect the inference of the harmonic coefficients.

The first source of instability comes from the likelihood itself. Its origin can be traced back to the presence of quasi-singular pulsar configurations that make the Γ_{LS} matrix lines almost linearly dependent. In this instance, the $\mathbf{a}_{\text{LS}}$ space presents an almost-degenerate direction. In this analysis, we introduce a *ridge regularization* scheme to address this potentially critical issue. We add a penalization term to equation (3.13) given by $\lambda \mathbf{a}_{\text{LS}}^T \mathcal{P} \mathbf{a}_{\text{LS}}$, where $\lambda > 0$ is the regularization parameter, and $\mathcal{P}$ is a symmetric, positive semi-definite penalty matrix. The new maximum likelihood position is now given by

$$\mathbf{a}_{\text{LS}} = \left[I - (\Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{LS}} + \lambda \mathcal{P})^{-1} \lambda \mathcal{P} \right] \mathbf{a}_{\text{LS}}^{\text{true}} + (\Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{LS}} + \lambda \mathcal{P})^{-1} \Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}}, \quad (3.17)$$

as we show in appendix C. First of all, we note that regularization schemes come with an intrinsic trade-off: even though they remove almost-degenerate directions from the likelihood analysis, they are also responsible for the introduction of an unavoidable bias in the determination of the true large-scale coefficients, as we read from the first term of equation (3.17). Second, we observe that the leakage term still appears in the second term of equation (3.17); however, we can already foresee that this term is suppressed for large values of the regularization parameter. Although this choice of regularization scheme is arbitrary, the point we are raising is not: removing numerical instabilities from the analysis in an uncontrolled fashion has the potential of biasing the inference of an unknown physical signal.

The second potential source of instability is due to an ill-conditioned covariance matrix. Also in this case, in the presence of pulsar geometries that present near-perfect degeneracies, the inverse of the covariance matrix could be numerically unstable. A typical solution involves the introduction of a regularized version of the covariance matrix, $\text{Cov}_{\text{reg}}^{-1}$, see, e.g., appendix C for the practical implementation used in this work. In terms of the inferred bias in the value of large-scale coefficients, this regularization does not solve the small-scale leakage, since this only exchanges Cov^{-1} for $\text{Cov}_{\text{reg}}^{-1}$ in equation (3.17). The large-scale coefficients are still biased by a small-scale leakage, even though by a different amount. We note that it is not possible to disentangle the small-scale leakage and the additional bias due to covariance regularization.

Once both of these regularization schemes are taken into account, we find that the reconstructed angular power spectrum reads as

$$C_{\ell}^{\text{rec}} = C_{\ell}^{\text{true}} - C_{\ell}^{\text{reg}} + C_{\ell}^{\text{leakage}}, \quad (3.18)$$

where the *regularization bias* term C_{ℓ}^{reg} is given by

$$C_{\ell}^{\text{reg}} = \sum_{\ell'=2}^{\ell_{\text{max}}^{\text{rec}}} K_{\ell\ell'} C_{\ell'}^{\text{true}}, \quad (3.19)$$

and $K_{\ell\ell'}$ is a second, regularization scheme dependent, mode-mixing matrix, this time coupling different large-scale modes. Its explicit form is reported in appendix C, alongside with the form of the new mode-mixing matrix $M_{\ell\ell'}''$ appearing in the leakage term

$$C_{\ell}^{\text{leakage}} = \sum_{\ell'=\ell_{\text{max}}^{\text{rec}}+1}^{\ell_{\text{max}}^{\text{GWB}}} M_{\ell\ell'}'' C_{\ell'}^{\text{true}}. \quad (3.20)$$

The functional form of $M_{\ell\ell'}''$ and $K_{\ell\ell'}$ is also shown in appendix C.

We show in figure 4 how different choices of the ridge parameter affect the reconstruction of the angular power spectra for NG34 and NG68 pulsar configurations in the top and bottom panels, respectively. First, we observe how for small values of λ the amount of small-scale leakage extends to very low multipoles for both configurations, especially compared to the isotropic case of section 3.2.

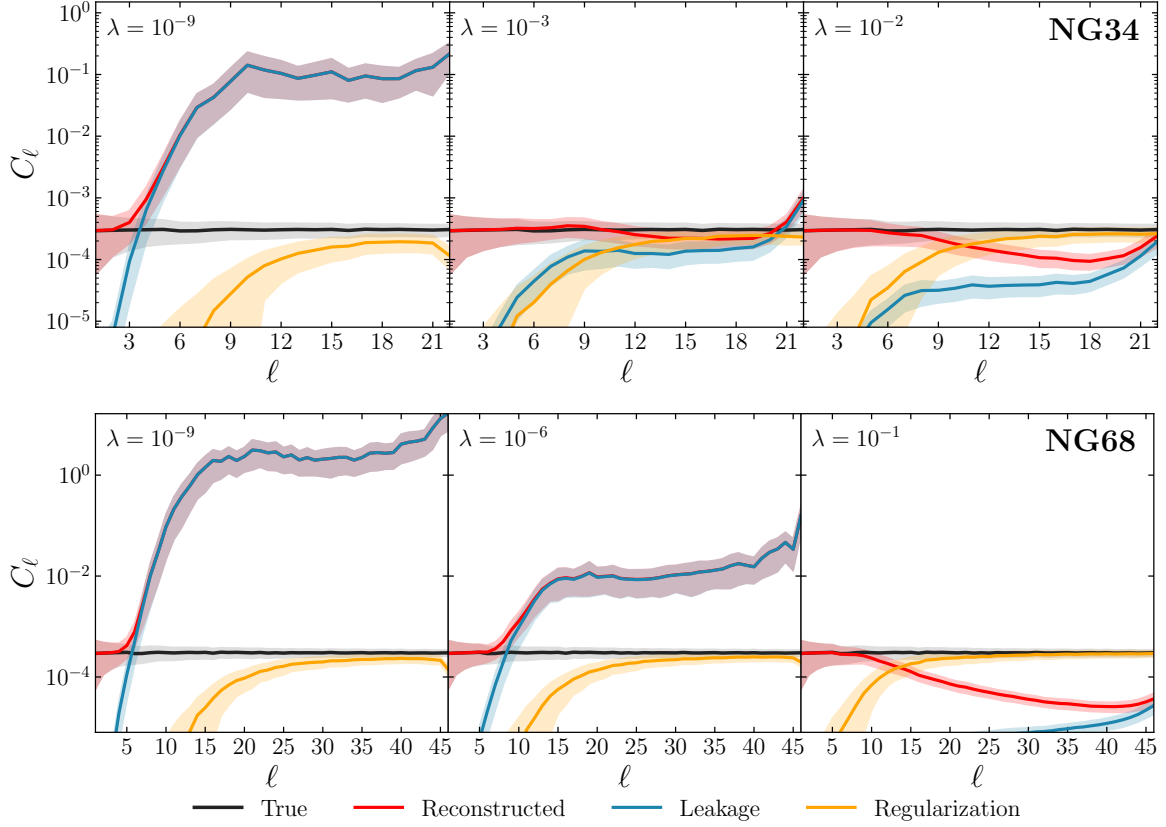


Figure 4: Envelopes of true (*black*), reconstructed (*red*), leakage (*blue*) and regularization (*orange*) C_ℓ for 10^3 realizations of the GWB and the NG34 (*top panels*) and NG68 (*bottom panels*) pulsar geometry. Different panels show the effect of different values of the ridge regularization parameter, indicated on the top left corner. Solid lines indicate the median value of angular power spectrum. Although a mild regularization appears to successfully suppress the small-scale leakage effect, more aggressive choices suppress the reconstructed signal below its true value. Since the true value of GWB angular power spectrum is not known a priori, it is not possible to fine-tune the regularization process to filter out the spurious effect.

Increasing the magnitude of λ effectively reduces the leakage contribution; however, it also increases the regularization bias, making it of the same order of magnitude of the true angular power spectrum, thus effectively suppressing the reconstructed signal. This phenomenology is supported by the overall form of the $M''_{\ell\ell'}$ and $K_{\ell\ell'}$ functions, as we show in appendix C. In particular, we notice $M''_{\ell\ell'}$ presents the same shape as in the purely geometric case, but its amplitude decreases as the ridge parameter increases, i.e., $M''_{\ell\ell'}(\lambda \rightarrow \infty) \rightarrow 0$, effectively canceling the leakage term. In contrast, we observe for the regularization matrix that $K_{\ell\ell'}(\lambda \rightarrow \infty) \rightarrow \delta_{\ell\ell'}$, so that $C_\ell^{\text{reg}} \rightarrow C_\ell^{\text{true}}$ and a null signal is reconstructed. The phenomenology in the isotropic and uniformly distributed pulsar is qualitatively identical to the NANOGrav pulsar geometry, therefore it is not displayed.

Finally, it is natural to wonder whether small-scale leakage can be eliminated by making an appropriate choice of the ridge parameter. Unfortunately, this is not the case, since a priori we lack any knowledge regarding the amount of power at small scales. In other words, since we are not aware from first principles of the magnitude of the small-scale leakage, we cannot use this regularization scheme to recover the true power spectrum in the multipole range that has already been contaminated by small-scale power.

4 Impact of pulsar timing noise

Although so far we considered a purely deterministic analysis in which pulsar noise plays no role, this is not the case in a real-life scenario. Here, we want to demonstrate that the small-scale leakage exists even in a scenario where noise is included, even though some simplifying assumptions are still taken to maintain the formalism completely analytical. In general, the measured pulsar timing residual R_p^{meas} receives two contributions, i.e.

$$R_p^{\text{meas}} = R_p + n_p, \quad (4.1)$$

where R_p is the residual defined in equation (2.4), and the noise n_p is assumed to have zero mean $\langle n_p \rangle = 0$ and variance [31, 43]

$$\langle n_p n_q \rangle_n = \frac{\delta^D(f - f')}{2} \delta_{pq}^K \sigma_p^2(f), \quad (4.2)$$

where $\langle \cdot \rangle_n$ indicates ensemble averages over noise realizations, $\sigma_p^2(f)$ is the (one-sided) noise power spectrum, and we implicitly assume that the pulsar noise is uncorrelated between different pulsars and is also uncorrelated with the GWB signal, i.e., $\langle R_p n_q \rangle = 0$. In this case, the inclusion of the noise transforms the form of the full covariance matrix introduced in the likelihood, which now receives an additional contribution and reads as

$$\text{Cov}_{i_{pq}j_{rs}} = [\delta_{pr}^K \delta_{qs}^K + \delta_{ps}^K \delta_{qr}^K] \sigma_p^2 \sigma_q^2 + \sum_{\ell m} (\Gamma_{\ell m}^{pr} \Gamma_{\ell m}^{qs} + \Gamma_{\ell m}^{ps} \Gamma_{\ell m}^{qr}) (C_\ell + N_\ell^{\text{shot}}). \quad (4.3)$$

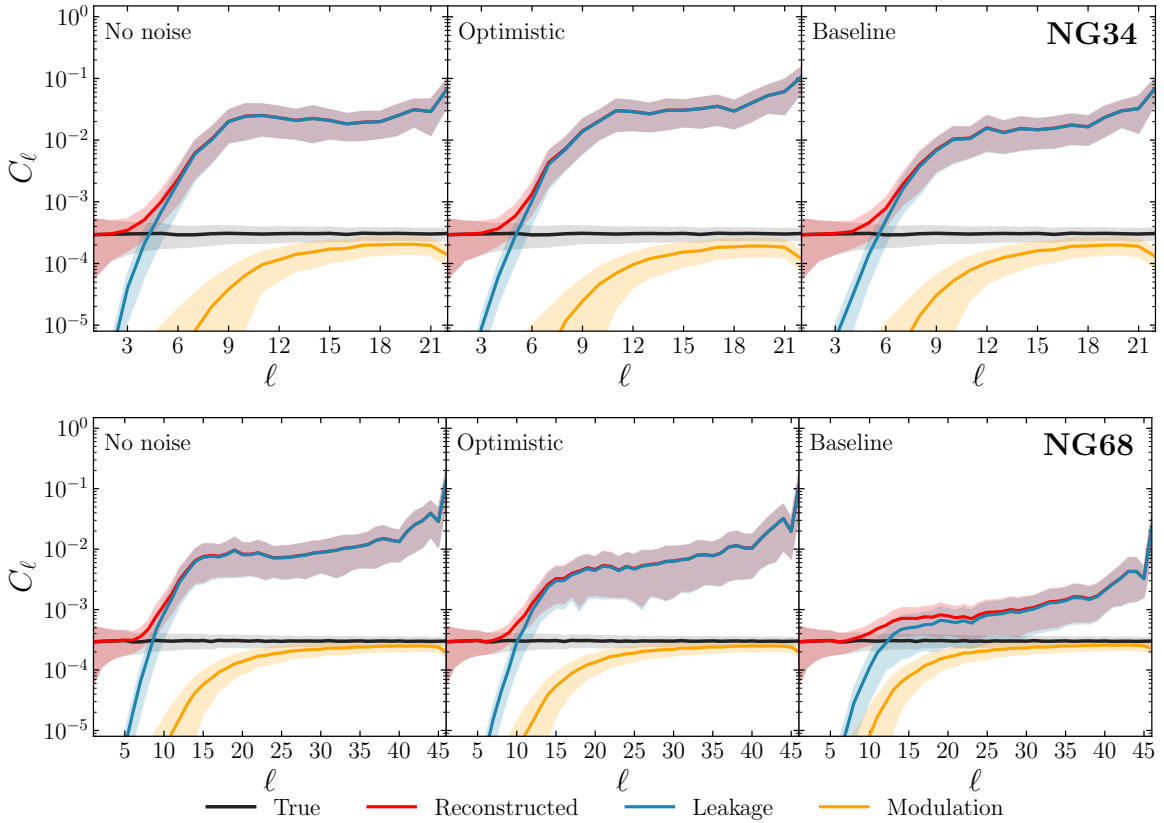


Figure 5: Same of figure 4, but for different amounts of injected pulsar red noise. Although in both cases the injected noise does not significantly affect the reconstruction of the signal, we observe as it increases the variance of the bands.

We created a realistic set of pulsar noise values by rescaling the current noise budget of pulsars in the NANOGrav 15-year dataset in the $f = 4$ nHz frequency bin [45]. The magnitude of this rescaling makes the error budget on average compatible with the detection of GWB anisotropies. In particular, we define a *baseline* case, where the noise and the signal terms appearing in equation (4.3) have approximately the same order of magnitude (i.e., $\sim 50\%$ of the pulsar pairs are signal dominated); and an *optimistic* case, where the noise is on average subdominant with respect to the signal (i.e., $\sim 90\%$ of the pulsar pairs are signal dominated). We show in figure 5 the impact of considering noise in the measurements of the correlated residuals. As we observe by contrasting figures 4 and 5, the inclusion of this additional term in the covariance does not substantially alter the phenomenology already described in the previous section. This is not unexpected, since also in this case the position of the maximum likelihood has still the form reported in equation (3.17). Moreover, for any practical purpose, one could even think of the noise contribution as some form of natural regularization of the covariance matrix, which in this case necessitates a milder regularization cut-off. However, the effect of small-scale leakage is still present in all its relevance, regardless of the noise configuration, the pulsar geometry, or the regularization scheme (even if, for conciseness, we show only one case for the latter). Also in this scenario, the phenomenology in the isotropic and uniformly distributed pulsar is qualitatively identical to the NANOGrav pulsar geometry, therefore it is not displayed.

5 Discussion and conclusions

Since each pulsar timing residual measurement is sensitive to the GW power emitted across the entire sky, unresolved small-scale power unavoidably leaks into the large-scale one when uncontrolled partial reconstruction of the GWB is attempted. This effect is particularly significant for the case in which the origin of the background is astrophysical, since the GW power generated by SMBHBs tracing the cosmic web is spread across an extremely wide range of angular scales. The unmodeled power beyond the current resolution limits of $\ell_{\text{max}}^{\text{rec}} \sim 10$ does not simply vanish from the correlated timing residuals; therefore, it has the potential to be misinterpreted as large-scale power.

The magnitude of this systematic effect strongly depends on the amount of power at small scales. In this work, we show that, for realistic realizations of the astrophysical background, this contamination exceeds the cosmic variance level by at least one order of magnitude across a broad range of scales, fundamentally altering how we interpret anisotropy measurements from PTA experiments. These numbers should be interpreted as conservative, as we artificially removed any power existing at scales $\ell > \ell_{\text{max}}^{\text{GWB}} = 250$ to reduce computational costs. Furthermore, as demonstrated in this article, the presence of this systematic error is independent of the specific details of the pulsar spatial configuration, the regularization scheme introduced in the analysis, and the presence of pulsar noise. In particular, we show how regularization represents both a necessity and a compromise, since it introduces its own regularization bias, and therefore it cannot be thought of as a possible solution to the small-scale leakage issue.

Perfect reconstruction of the large-scale angular structure of the background appears challenging without any prior knowledge of the amount of small-scale power, since it would require the fulfillment of at least one of these conditions: (i) complete absence of small-scale power, i.e., $\mathbf{a}_{\text{SS}} = \mathbf{0}$; or (ii) an astonishing large resolution, so that $\mathcal{R}_{\text{SS}} \ll \mathcal{R}_{\text{LS}}$ for all pairs. Since none of these conditions can be achieved for realistic PTA experiments, the bias introduced by the small-scale leakage effect is unavoidable, and some mitigation strategy should be implemented in the statistical analysis. A possible way forward in solving this issue would be to create an estimator that has some intrinsic form of “orthogonality” between large- and small-scale correlated residuals.

Additionally, the presence of this systematic effect should also be addressed on the side of numerical simulations. If the simulated signal does not incorporate a realistic realization of the power present at small scales in realistic, physically motivated scenarios, the subsequent analysis will be automatically blind to the presence of this effect. Moreover, since such simulations are also used to validate optimal estimators, the statistical pipeline developed to analyze current and future datasets might not reach the desired level of accuracy.

On the physical interpretation side, the consequences of small-scale leakage are potentially catastrophic. The distorted maps and angular power spectra are fundamentally incompatible with any large-scale structure pattern, and thus with any astrophysical interpretation. Therefore, any sort of “cross-correlation” analysis with galaxy catalogs, either to locate sources in host galaxies or to study the clustering properties of SMBHBs, will return spurious correlations sourced by this systematic effect.

Although this work considers only the weak-field regime scenario for GWB anisotropies, we do not expect the situation to improve once individual sources are introduced into the analysis. In contrast, loud continuous-wave sources create localized hot spots and are associated with scale-invariant angular power spectrum with power spread over a wide range of scales. Moreover, their presence also breaks the assumption of Gaussianity of the inferred harmonic coefficients. Therefore, in addition to the guaranteed presence of a small-scale leakage, we expect further complications to arise in terms of dealing with the statistics of the correlated residuals. In this sense, the nature of the leakage effect is fully general in different scenarios.

Statistical uncertainties will continue to decrease as PTAs start monitoring a larger number of pulsars and extend their observation baselines. However, systematic effects will not disappear and will eventually dominate the total error budget, unless addressed explicitly. In particular, small-scale leakage represents a fundamental limitation of current analysis methods when reconstructing the GW sky from PTA data; thus, it can be mitigated but not fully eliminated by increasing the number of pulsars. In this sense, understanding and mitigating these systematic effects will mark a crucial step toward reliable characterization of the nHz GWB angular structure.

Acknowledgements

The authors thank Deepali Agarwal, Mesut Çalışkan, Neha Anil Kumar, Michele Liguori, Giorgio Mentasti, Nihan Pol, Luca Reali, Joseph Romano, Stephen Taylor, and Benjamin Wandelt for useful discussions. The authors also thank Bjorn Larsen for providing the red noise data for the NANOGrav set of pulsars.

This work is partly supported by ICSC - Centro Nazionale di Ricerca in High Performance Computing, Big Data and Quantum Computing, funded by European Union - NextGenerationEU. NB acknowledges support from the European Union’s Horizon Europe research and innovation program under the Marie Skłodowska-Curie grant agreement no. 101207487 (GWSKY - Mapping the Universe with Gravitational Waves), and by PRD/ARPE 2022 “Cosmology with Gravitational waves and Large Scale Structure - CosmoGraLSS”. CMFM acknowledges support from the National Science Foundation from Grants NSF PHY-2020265 and AST-2414468. This work was also supported by the Flatiron Institute, part of the Simons Foundation. AR acknowledges funding from the Italian Ministry of University and Research (MIUR) through the “Dipartimenti di eccellenza” project “Science of the Universe”.

A Persistence of small-scale residuals

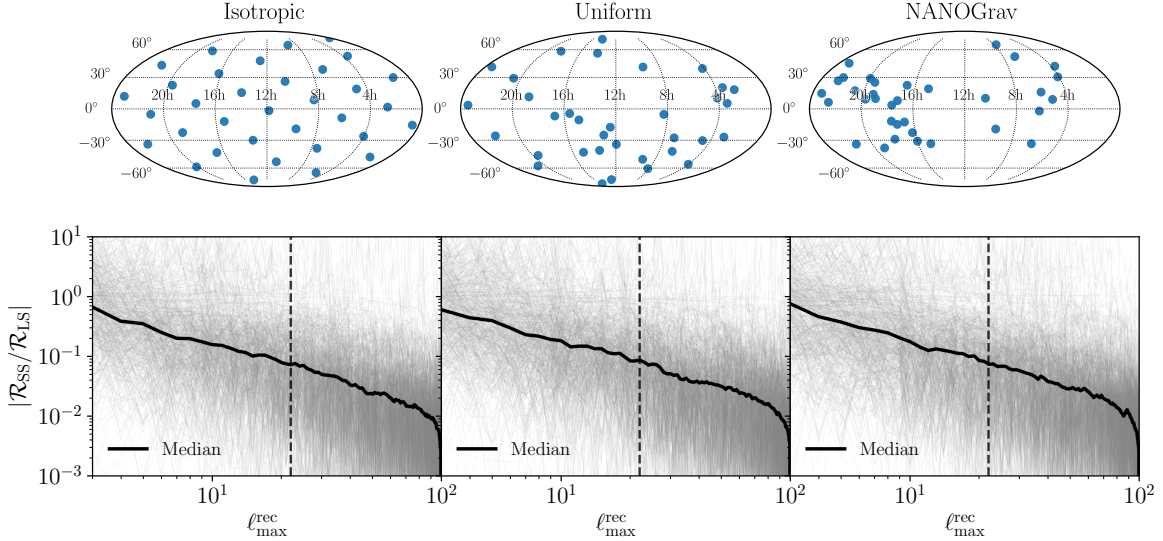


Figure 6: *Top panels:* isotropic (I34), uniform (U34) and NANOGrav (NG34) samples of pulsars considered in this work. *Bottom panels:* absolute value of the ratio of small-scale to large-scale correlated residuals for all the pulsar pairs in the sample. The *black solid line* represents the median value of $|\mathcal{R}_{\text{SS}}/\mathcal{R}_{\text{LS}}|$.

The existence of small-scale correlated residuals is a physical effect completely independent of the geometrical distribution of pulsars. We show in figure 6 the absolute value of the small-scale to large-scale correlated residuals ratio for all possible pairs of the I34, U34 and NG34 geometrical configurations of pulsars. As it immediately appears, especially for very small maximum reconstruction multipoles, this ratio is order unity. Therefore, by attempting any partial reconstruction of a GWB anisotropic map, we are effectively misplacing almost the entire amount of small-scale power into large-scale modes. This behavior is consistent across different geometries, as it becomes even more evident by comparing the probability distribution functions of the small- to large-scale correlated residual ratio in figure 7 for the pulsar geometries considered in this work at their respective maximum reconstruction multipole. We also report in the tables in the top left corners of the panels some statistical quantities of interest that describe such probability distribution functions. We observe the existence of a small, but not zero, number of “outlier” pulsar pairs N_{out} , for which this ratio takes values $|\mathcal{R}_{\text{SS}}/\mathcal{R}_{\text{LS}}| > 2$. Moreover, by comparing the proportion between the 68% and 95% CL regions, we note that these distributions have strong non-Gaussian tails, allowing for a large number of pairs where the error on the assignment of power is almost order unity. The dispersion of the probability distribution function at multipole $\ell < \ell_{\text{max}}^{\text{rec}}$ is larger than in the case shown in the figure.

B Complementary material on small-scale leakage derivation

B.1 Geometrical intuition

The minimum of the χ_{LMS}^2 is found by solving the *normal equations*

$$\nabla_{\mathbf{a}_{\text{LS}}} \chi_{\text{LMS}}^2 \simeq \Gamma_{\text{LS}}^T (\Gamma_{\text{LS}} \mathbf{a}_{\text{LS}}^{\text{true}} + \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}} - \Gamma_{\text{LS}} \mathbf{a}_{\text{LS}}) = 0, \quad (\text{B.1})$$

where $\nabla_{\mathbf{a}_{\text{LS}}}$ is an operator whose elements are derivatives with respect to each individual large-scale $a_{\ell m}$. The use of the Moore-Penrose inverse immediately returns equation (3.9) since by construction $\Gamma_{\text{LS}}^+ \Gamma_{\text{LS}} = I_{N_{\text{LS}}}$, with I_n being the identity matrix of dimension n .

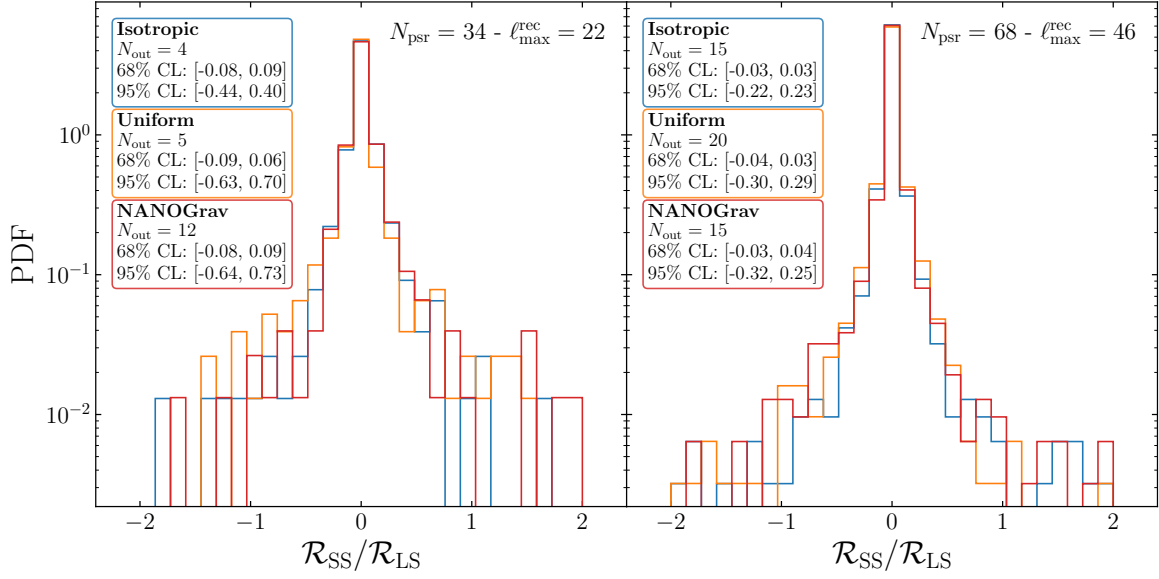


Figure 7: Probability distribution function (PDF) of the $\mathcal{R}_{\text{SS}}/\mathcal{R}_{\text{LS}}$ correlated residual ratio for the $N_{\text{psr}} = 34$ (left panel) and $N_{\text{psr}} = 68$ (right panel) pulsar configurations. Each PDF is computed at the maximum reconstruction multipole indicate in the top right corner. Blue, orange, and red lines indicate the isotropic, uniform, and NANOGrav configurations, respectively. The tables on the top left corners report for each distribution the number of outliers N_{out} such that $|\mathcal{R}_{\text{SS}}/\mathcal{R}_{\text{LS}}| > 2$, and the 68%, 95% confidence limits (CL). As the confidence limits suggest, even though the PDFs seem Gaussian, they have generally larger tails.

For each value of ℓ , the vector $\mathbf{a}_\ell$ is constructed by introducing a projection matrix $\mathbb{P}_\ell$ to isolate the $(2\ell + 1)$ components of $\mathbf{a}_{\text{LS}}$ corresponding to $a_{\ell m}$. In practice, $\mathbb{P}_\ell$ is a rectangular matrix with dimensions $(2\ell + 1) \times N_{\text{LS}}$ defined as

$$\mathbb{P}_\ell = \begin{pmatrix} 0_{(2\ell+1) \times (\ell^2-4)} & I_{(2\ell+1)} & 0_{(2\ell+1) \times (N_{\text{LS}}-\ell^2-2\ell+3)} \end{pmatrix}, \quad (\text{B.2})$$

where $0_{i \times j}$ is a rectangular matrix where all elements are zeros. For instance, we have that

$$\mathbb{P}_2 = \begin{pmatrix} I_5 & 0_{5 \times N_{\text{LS}}-5} \end{pmatrix}, \quad \mathbb{P}_3 = \begin{pmatrix} 0_{7 \times 5} & I_7 & 0_{7 \times N_{\text{LS}}-12} \end{pmatrix}, \quad \mathbb{P}_4 = \begin{pmatrix} 0_{9 \times 12} & I_9 & 0_{9 \times N_{\text{LS}}-21} \end{pmatrix}, \quad \text{etc.}, \quad (\text{B.3})$$

and, for the purpose of developing future intuition, we note that the square matrix $\mathbb{P}_\ell^T \mathbb{P}_\ell$ is given by

$$\mathbb{P}_\ell^T \mathbb{P}_\ell = \begin{pmatrix} 0_{(\ell^2-4) \times (\ell^2-4)} & 0_{(\ell^2-4) \times (2\ell+1)} & 0_{(\ell^2-4) \times (N_{\text{LS}}-\ell^2-2\ell+3)} \\ 0_{(2\ell+1) \times (\ell^2-4)} & I_{(2\ell+1)} & 0_{(2\ell+1) \times (N_{\text{LS}}-\ell^2-2\ell+3)} \\ 0_{(N_{\text{LS}}-\ell^2-2\ell+3) \times (\ell^2-4)} & 0_{(N_{\text{LS}}-\ell^2-2\ell+3) \times (2\ell+1)} & 0_{(N_{\text{LS}}-\ell^2-2\ell+3) \times (N_{\text{LS}}-\ell^2-2\ell+3)} \end{pmatrix}. \quad (\text{B.4})$$

The convoluted matrix product of equation (3.12) can be simplified by defining a “leakage” matrix

$$\mathcal{M}_\ell = \Gamma_{\text{SS}}^T (\Gamma_{\text{LS}}^+)^T \mathbb{P}_\ell^T \mathbb{P}_\ell \Gamma_{\text{LS}}^+ \Gamma_{\text{SS}}, \quad (\text{B.5})$$

in such a way that

$$\begin{aligned}
C_\ell^{\text{leakage}} &= \frac{1}{2\ell+1} \langle (\mathbf{a}_{\text{SS}}^{\text{true}})^T \mathcal{M}_\ell \mathbf{a}_{\text{SS}}^{\text{true}} \rangle = \frac{1}{2\ell+1} \sum_{ij} \langle \mathbf{a}_{\text{SS},i}^{\text{true}} \mathcal{M}_{\ell,ij} \mathbf{a}_{\text{SS},j}^{\text{true}} \rangle \\
&= \frac{1}{2\ell+1} \sum_{\ell'm'} \sum_{\ell''m''} \langle a_{\ell'm'}^{\text{true}} \mathcal{M}_{\ell,i_{\ell'm'} j_{\ell''m''}} a_{\ell''m''}^{\text{true}} \rangle = \frac{1}{2\ell+1} \sum_{\ell'm'} \sum_{\ell''m''} \mathcal{M}_{\ell,i_{\ell'm'} j_{\ell''m''}} \delta_{\ell'\ell''}^K \delta_{m'm''}^K C_{\ell'}^{\text{true}} \\
&= \sum_{\ell'} \left(\frac{1}{2\ell+1} \sum_{m'} \mathcal{M}_{\ell,i_{\ell'm'} i_{\ell'm'}} \right) C_{\ell'}^{\text{true}} = \sum_{\ell'} M_{\ell\ell'} C_{\ell'}^{\text{true}},
\end{aligned} \tag{B.6}$$

where we indicate with $i_{\ell m}$ the matrix row/column index corresponding to the (ℓ, m) pair of indices.

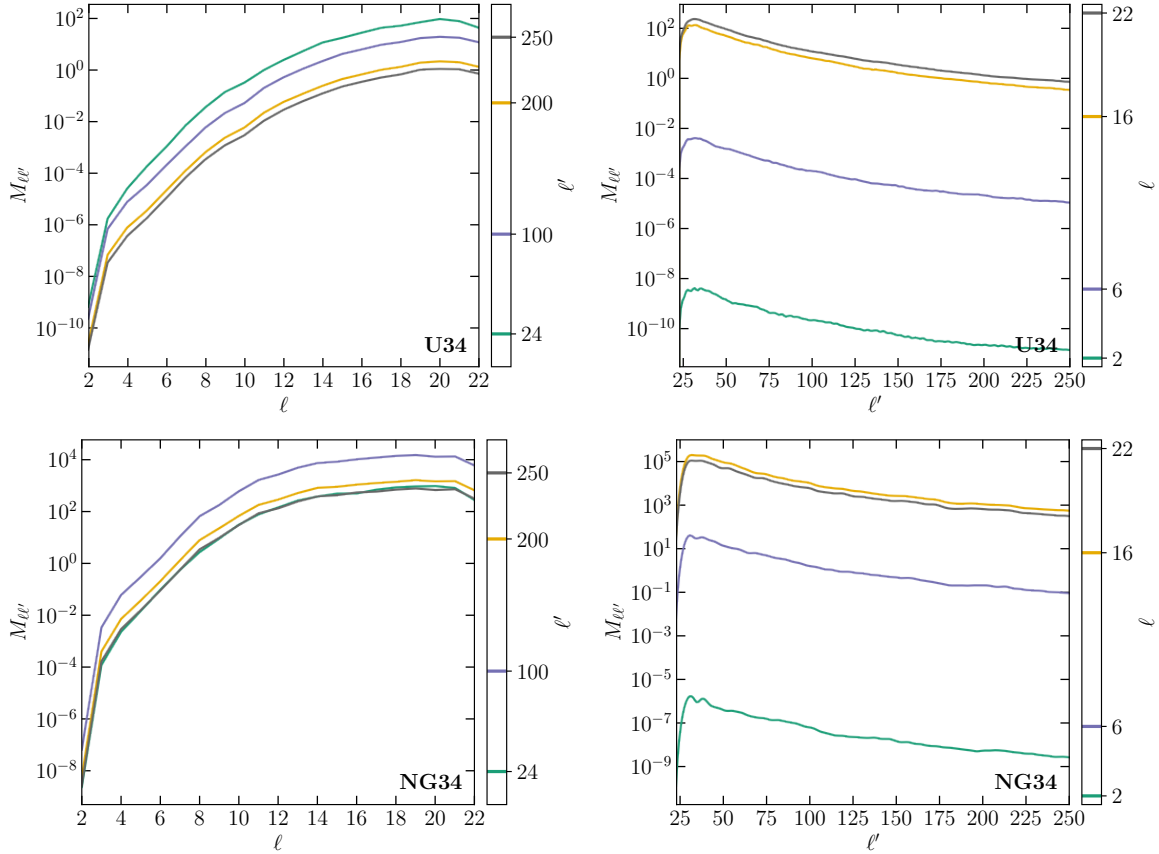


Figure 8: Value of the mode-mixing function $M_{\ell\ell'}$ in term of large-scale ℓ and small-scale ℓ' multipoles for the U34 (*top panels*) and NG34 (*bottom panels*) pulsar configurations.

We show in figure 8 the structure of the mode-mixing function for the pulsar configurations U34 and NG34 in the top and bottom panels, respectively. By comparing this figure with the I34 pulsar case shown in figure 3, the overall shape of this function does not vary significantly between different pulsar geometries. However, there is a significant difference in terms of amplitude of $M_{\ell\ell'}$, likely introduced by the presence of nearly singular pulsar configurations. Therefore, we expect that the small-scale leakage in these cases is considerably larger than that obtained for the I34 configuration.

B.2 Statistical intuition

Similarly to the purely geometric case, the position of the maximum of the likelihood is given by

$$\nabla_{\mathbf{a}_{\text{LS}}} \log \mathcal{L} \simeq \Gamma_{\text{LS}}^T \text{Cov}^{-1} (\mathcal{R} - \Gamma_{\text{LS}} \mathbf{a}_{\text{LS}}) = 0. \tag{B.7}$$

In this case, the normal equations is commonly recast as

$$H\mathbf{a}_{\text{LS}} = \mathbf{v}, \quad (\text{B.8})$$

where the $H = \Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{LS}} \simeq \nabla_{\mathbf{a}_{\text{LS}}}^2 \log \mathcal{L}$ is proportional to the likelihood Hessian matrix, and $\mathbf{v} = \Gamma_{\text{LS}}^T \text{Cov}^{-1} \mathbf{R}$ is the “likelihood score”. As before, it is straightforward to recover equation (3.15) from the set of normal equations. In this case, the bias introduced depends not only on the pulsar geometry but also on how the covariance matrix is modeled. This result is totally analogous of what in Cosmology is typically referred to as “bestfit shift” due to an incorrect modeling of the physical observable, see, e.g., ref. [46] and refs. therein.

Finally, we observe that the algebra underlying equation (3.16) is completely equivalent to that leading to equation (B.6), except for using as leakage matrix

$$\mathcal{M}'_\ell = [\Gamma_{\text{LS}}^{-1} \text{Cov}^{-1} \Gamma_{\text{SS}}]^T \left[(\Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{LS}})^{-1} \right]^T \mathbb{P}_\ell^T \mathbb{P}_\ell (\Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{LS}})^{-1} \Gamma_{\text{LS}}^{-1} \text{Cov}^{-1} \Gamma_{\text{SS}}, \quad (\text{B.9})$$

leading to an analogous definition of a matrix $M'_{\ell\ell'}$.

B.3 Covariance matrix

In this section we compute for the first time the form of the covariance matrix for anisotropies searches without discarding the intrinsic cosmological signal, along the lines of what described in ref. [31]. For the sake of simplicity we consider a single, narrow frequency bin; however, the derivation presented here can easily being generalized to the case of multiple, wide frequency bins. Suppose to define for each pulsar an estimator of the measured timing residual in equation (4.1) as

$$\hat{R}_p^{\text{meas}} = \hat{R}_p + \hat{n}_p, \quad (\text{B.10})$$

so that the correlated residual estimator simply reads as

$$\hat{\mathcal{R}}_{pq}^{\text{meas}} = \hat{R}_p^{\text{meas}} \hat{R}_q^{\text{meas}} = (\hat{R}_p + \hat{n}_p) (\hat{R}_q + \hat{n}_q). \quad (\text{B.11})$$

This estimator is built in such a way that its expectation value over the source and noise properties is given by

$$\langle \hat{\mathcal{R}}_{pq}^{\text{meas}} \rangle_{s,n} = \mathcal{R}_{pq}^{\text{iso}} + \mathcal{R}_{pq}^{\text{anis}} + \delta_{pq}^K \sigma_p^2, \quad (\text{B.12})$$

where, as already commented, the anisotropic component still has a residual degree of stochasticity such that

$$\langle \mathcal{R}_{pq}^{\text{anis}} \rangle_p = \sum_{\ell m} \Gamma_{\ell m}^{pq} \langle a_{\ell m} \rangle_p = 0, \quad (\text{B.13})$$

and

$$\langle \mathcal{R}_{pq}^{\text{anis}} \mathcal{R}_{rs}^{\text{anis}} \rangle_p = \sum_{\ell m} \sum_{\ell' m'} \Gamma_{\ell m}^{pq} \Gamma_{\ell' m'}^{rs} \langle a_{\ell m} a_{\ell' m'} \rangle_p = \sum_{\ell m} \sum_{\ell' m'} \Gamma_{\ell m}^{pq} \Gamma_{\ell' m'}^{rs} C_{\ell}. \quad (\text{B.14})$$

In this section we use the angle bracket subscripts (s, n, p) to indicate ensemble averages over source, noise and position properties, respectively, and should not be confused with the pulsar indices. In particular, the angle bracket subscript p has been omitted in the main text for simplicity; however, in this section it is useful to keep track about what sort of averages we are implementing in different stages of the derivation.

Therefore, it is immediate to construct an estimator of the measured anisotropic correlated residual as

$$\hat{\mathcal{R}}_{pq}^{\text{anis, meas}} = \hat{\mathcal{R}}_{pq}^{\text{meas}} - \mathcal{R}_{pq}^{\text{iso}} - \delta_{pq}^K \sigma_p^2 = \hat{R}_p^{\text{meas}} \hat{R}_q^{\text{meas}} - \mathcal{R}_{pq}^{\text{iso}} - \delta_{pq}^K \sigma_p^2, \quad (\text{B.15})$$

where, as expected,

$$\langle \hat{\mathcal{R}}_{pq}^{\text{anis, meas}} \rangle_{s,n} = \mathcal{R}_{pq}^{\text{anis, th}}, \quad \langle \hat{\mathcal{R}}_{pq}^{\text{anis, meas}} \rangle_{s,n,p} = 0, \quad (\text{B.16})$$

and in this case the “theoretical” correlated residual $\mathcal{R}_{pq}^{\text{anis,th}}$ has the meaning of the correlated residual of the realization of the sky we observe. At this point, the (i_{pq}, j_{rs}) element of the covariance matrix read as

$$\begin{aligned} \text{Cov}_{i_{pq}j_{rs}} &= \left\langle \left[\hat{\mathcal{R}}_{pq}^{\text{anis,meas}} - \mathcal{R}_{pq}^{\text{anis,th}} \right] \left[\hat{\mathcal{R}}_{rs}^{\text{anis,meas}} - \mathcal{R}_{rs}^{\text{anis,th}} \right] \right\rangle_{s,n,p} \\ &= \left\langle \left[\hat{R}_p^{\text{meas}} \hat{R}_q^{\text{meas}} - \mathcal{R}_{pq}^{\text{iso}} - \delta_{pq}^K \sigma_p^2 - \mathcal{R}_{pq}^{\text{anis,th}} \right] \left[\hat{R}_r^{\text{meas}} \hat{R}_s^{\text{meas}} - \mathcal{R}_{rs}^{\text{iso}} - \delta_{rs}^K \sigma_r^2 - \mathcal{R}_{rs}^{\text{anis,th}} \right] \right\rangle_{s,n,p}. \end{aligned} \quad (\text{B.17})$$

Given the statistical properties described before, it is straightforward to evaluate all the individual terms appearing in this equation. In particular, we have

$$\begin{aligned} \left\langle \hat{R}_p^{\text{meas}} \hat{R}_q^{\text{meas}} \hat{R}_r^{\text{meas}} \hat{R}_s^{\text{meas}} \right\rangle_{s,n,p} &= \left\langle \left(\hat{R}_p + \hat{n}_p \right) \left(\hat{R}_q + \hat{n}_q \right) \left(\hat{R}_r + \hat{n}_r \right) \left(\hat{R}_s + \hat{n}_s \right) \right\rangle_{s,n,p} \\ &= \left\langle \hat{R}_p \hat{R}_q \hat{R}_r \hat{R}_s + \left(\hat{R}_p \hat{R}_q \hat{R}_r \hat{n}_s + 3 \text{ perm.} \right) \right. \\ &\quad \left. + \left(\hat{R}_p \hat{R}_q \hat{n}_r \hat{n}_s + 5 \text{ perm.} \right) + \left(\hat{R}_p \hat{n}_q \hat{n}_r \hat{n}_s + 3 \text{ perm.} \right) \right. \\ &\quad \left. + \hat{n}_p \hat{n}_q \hat{n}_r \hat{n}_s \right\rangle_{s,n,p} \\ &= \left(\mathcal{R}_{pq}^{\text{iso}} + \delta_{pq}^K \sigma_p^2 \right) \left(\mathcal{R}_{rs}^{\text{iso}} + \delta_{rs}^K \right) \\ &\quad + \left(\mathcal{R}_{pr}^{\text{iso}} + \delta_{pr}^K \right) \left(\mathcal{R}_{qs}^{\text{iso}} + \delta_{qs}^K \right) \\ &\quad + \left(\mathcal{R}_{ps}^{\text{iso}} + \delta_{ps}^K \right) \left(\mathcal{R}_{qr}^{\text{iso}} + \delta_{qr}^K \right) \\ &\quad + \sum_{\ell m} \left[\Gamma_{\ell m}^{pq} \Gamma_{\ell m}^{rs} + \Gamma_{\ell m}^{pr} \Gamma_{\ell m}^{qs} + \Gamma_{\ell m}^{ps} \Gamma_{\ell m}^{qr} \right] C_\ell, \end{aligned} \quad (\text{B.18})$$

where we used the fact that the only non-vanishing terms are

$$\begin{aligned} \left\langle \hat{R}_p \hat{R}_q \hat{R}_r \hat{R}_s \right\rangle_{s,n,p} &= \left\langle \left\langle \hat{R}_p \hat{R}_q \right\rangle_s \left\langle \hat{R}_r \hat{R}_s \right\rangle_s \right\rangle + \left\langle \left\langle \hat{R}_p \hat{R}_r \right\rangle_s \left\langle \hat{R}_q \hat{R}_s \right\rangle_s \right\rangle + \left\langle \left\langle \hat{R}_p \hat{R}_s \right\rangle_s \left\langle \hat{R}_q \hat{R}_r \right\rangle_s \right\rangle_p \\ &= \left\langle \left(\mathcal{R}_{pq}^{\text{iso}} + \hat{\mathcal{R}}_{pq}^{\text{anis}} \right) \left(\mathcal{R}_{rs}^{\text{iso}} + \hat{\mathcal{R}}_{rs}^{\text{anis}} \right) + \left(\mathcal{R}_{pr}^{\text{iso}} + \hat{\mathcal{R}}_{pr}^{\text{anis}} \right) \left(\mathcal{R}_{qs}^{\text{iso}} + \hat{\mathcal{R}}_{qs}^{\text{anis}} \right) \right. \\ &\quad \left. + \left(\mathcal{R}_{ps}^{\text{iso}} + \hat{\mathcal{R}}_{ps}^{\text{anis}} \right) \left(\mathcal{R}_{qr}^{\text{iso}} + \hat{\mathcal{R}}_{qr}^{\text{anis}} \right) \right\rangle_p \\ &= \mathcal{R}_{pq}^{\text{iso}} \mathcal{R}_{rs}^{\text{iso}} + \mathcal{R}_{pr}^{\text{iso}} \mathcal{R}_{qs}^{\text{iso}} + \mathcal{R}_{ps}^{\text{iso}} \mathcal{R}_{qr}^{\text{iso}} \\ &\quad + \sum_{\ell m} \left[\Gamma_{\ell m}^{pq} \Gamma_{\ell m}^{rs} + \Gamma_{\ell m}^{pr} \Gamma_{\ell m}^{qs} + \Gamma_{\ell m}^{ps} \Gamma_{\ell m}^{qr} \right] C_\ell, \end{aligned} \quad (\text{B.19})$$

$$\begin{aligned} \left\langle \hat{n}_p \hat{n}_q \hat{n}_r \hat{n}_s \right\rangle_{s,n,p} &= \left\langle \left\langle \hat{n}_p \hat{n}_q \right\rangle_n \left\langle \hat{n}_r \hat{n}_s \right\rangle_n \right\rangle + \left\langle \left\langle \hat{n}_p \hat{n}_r \right\rangle_n \left\langle \hat{n}_q \hat{n}_s \right\rangle_n \right\rangle + \left\langle \left\langle \hat{n}_p \hat{n}_s \right\rangle_n \left\langle \hat{n}_q \hat{n}_r \right\rangle_n \right\rangle_{s,p} \\ &= \left\langle \delta_{pq}^K \delta_{rs}^K \sigma_p^2 \sigma_r^2 + \delta_{pr}^K \delta_{qs}^K \sigma_p^2 \sigma_q^2 + \delta_{ps}^K \delta_{qr}^K \sigma_p^2 \sigma_q^2 \right\rangle_{s,p} \\ &= \delta_{pq}^K \delta_{rs}^K \sigma_p^2 \sigma_r^2 + \delta_{pr}^K \delta_{qs}^K \sigma_p^2 \sigma_q^2 + \delta_{ps}^K \delta_{qr}^K \sigma_p^2 \sigma_q^2, \\ \left\langle \hat{R}_p \hat{R}_q \hat{n}_r \hat{n}_s \right\rangle_{s,n,p} &= \left\langle \left\langle \hat{R}_p \hat{R}_q \right\rangle_s \left\langle \hat{n}_r \hat{n}_s \right\rangle_n \right\rangle_p = \mathcal{R}_{pq}^{\text{iso}} \delta_{rs}^K \sigma_r^2, \end{aligned}$$

and permutations of the latter. Similarly, we have that

$$\begin{aligned} \left\langle \hat{R}_p^{\text{meas}} \hat{R}_q^{\text{meas}} \left(\mathcal{R}_{rs}^{\text{iso}} + \delta_{rs}^K \sigma_r^2 + \mathcal{R}_{rs}^{\text{anis,th}} \right) \right\rangle_{s,n,p} &= \left\langle \left(\mathcal{R}_{pq}^{\text{iso}} + \delta_{pq}^K \sigma_p^2 + \mathcal{R}_{pq}^{\text{anis,th}} \right) \hat{R}_r^{\text{meas}} \hat{R}_s^{\text{meas}} \right\rangle_{s,n,p} \\ &= \left\langle \left(\mathcal{R}_{pq}^{\text{iso}} + \delta_{pq}^K \sigma_p^2 + \mathcal{R}_{pq}^{\text{anis,th}} \right) \left(\mathcal{R}_{rs}^{\text{iso}} + \delta_{rs}^K \sigma_r^2 + \mathcal{R}_{rs}^{\text{anis,th}} \right) \right\rangle_{s,n,p} \\ &= \left(\mathcal{R}_{pq}^{\text{iso}} + \delta_{pq}^K \sigma_p^2 \right) \left(\mathcal{R}_{rs}^{\text{iso}} + \delta_{rs}^K \sigma_r^2 \right) + \sum_{\ell m} \Gamma_{\ell m}^{pq} \Gamma_{\ell m}^{rs} C_\ell. \end{aligned} \quad (\text{B.20})$$

Therefore, it is immediate to recover that the elements of the covariance matrix read as in equation (3.15), when the noise is absent, or as in equation (4.3), when the noise is present.

C Regularization schemes

C.1 Likelihood regularization

After the introduction of the ridge regularization factor, the normal equations read as

$$[H + \lambda\mathcal{P}] \mathbf{a}_{\text{LS}} = \mathbf{v} = H\mathbf{a}_{\text{LS}}^{\text{true}} + \Gamma_{\text{LS}}^T \text{Cov}^{-1} \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}}, \quad (\text{C.1})$$

therefore, the explicit form of equation (3.17) is immediately recovered by substituting $H = (H + \lambda\mathcal{P}) - \lambda\mathcal{P}$ in the RHS of this equation. For the purpose of this work, we choose as penalty matrix $\mathcal{P} = \text{med}_H \times I_{N_{\text{LS}}}$, where med_H is median value of the Hessian matrix diagonal. For $\lambda = 1$, this coefficient introduces a regularization with an order of magnitude comparable to that of the ill-conditioned Hessian matrix.

As noted previously, it is possible to infer the phenomenology due to the introduction of a regularization scheme on the χ_{LMS}^2 statistics, i.e., in the map-making scenario, by substituting $\text{Cov} = I_{N_{\text{LS}}}$ in equation (3.17). In that case, we have that

$$\mathbf{a}_{\text{LS}} = \left[I - (\Gamma_{\text{LS}}^T \Gamma_{\text{LS}} + \lambda\mathcal{P})^{-1} \lambda\mathcal{P} \right] \mathbf{a}_{\text{LS}}^{\text{true}} + (\Gamma_{\text{LS}}^T \Gamma_{\text{LS}} + \lambda\mathcal{P})^{-1} \Gamma_{\text{LS}}^T \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}}. \quad (\text{C.2})$$

Therefore, also in this case, the biased angular power spectra have an identical structure as those obtained via the likelihood approach discussed in section 3.4. Finally, even if we present only the case of ridge regularization, other schemes as, for instance, SVD regularization, can be recast as a special ridge regularization case, thus the generality of this formulation and of our conclusions.

C.2 Covariance matrix regularization

In the case of regularizing the covariance matrix, we implement a SVD (Single Value Decomposition) regularization scheme. First, we perform an eigendecomposition of the covariance matrix as $\text{Cov} = U\Sigma U^{-1}$, where U is the orthogonal matrix of eigenvectors, and $\Sigma = \text{diag}(e_j)$ is a diagonal matrix containing the eigenvalues $\{e_j\}$ of the covariance. The regularized version of the inverse of the covariance is then explicitly computed as

$$\text{Cov}_{\text{reg}}^{-1} = U\Sigma_{\text{reg}}^{-1}U^{-1}, \quad (\text{C.3})$$

where the elements of the diagonal matrix Σ_{reg}^{-1} read as

$$(\Sigma_{\text{reg}}^{-1})_{jj} = \begin{cases} e_j^{-1}, & e_j \geq e_{\text{thres}}, \\ 0, & e_j < e_{\text{thres}}, \end{cases} \quad (\text{C.4})$$

and e_{thres} is an arbitrary threshold value. When only this regularization scheme is implemented, the normal equations read as

$$H_{\text{reg}} \mathbf{a}_{\text{LS}} = \mathbf{v}_{\text{reg}}, \quad (\text{C.5})$$

where the regularized Hessian and likelihood score are given by $H_{\text{reg}} = \Gamma_{\text{LS}}^T \text{Cov}_{\text{reg}}^{-1} \Gamma_{\text{LS}}$ and $\mathbf{v}_{\text{reg}} = \Gamma_{\text{LS}}^T \text{Cov}_{\text{reg}}^{-1} \mathbf{R}$, respectively.

C.3 Combining regularization schemes

When these two regulation schemes are simultaneously implemented, the estimate of the large-scale coefficients that maximize the likelihood is given by

$$\mathbf{a}_{\text{LS}} = \left[I - (H_{\text{reg}} + \lambda\mathcal{P})^{-1} \lambda\mathcal{P} \right] \mathbf{a}_{\text{LS}}^{\text{true}} + (H_{\text{reg}} + \lambda\mathcal{P})^{-1} \Gamma_{\text{LS}}^T \text{Cov}_{\text{reg}}^{-1} \Gamma_{\text{SS}} \mathbf{a}_{\text{SS}}^{\text{true}}, \quad (\text{C.6})$$

which now explicitly depends on the arbitrary choice of both λ and e_{thres} . In this case, we have to introduce a regularization bias matrix defined as

$$\begin{aligned} \mathcal{K}_\ell = \lambda \left[\left[\mathbb{P}_\ell^T \mathbb{P}_\ell (H_{\text{reg}} + \lambda\mathcal{P})^{-1} \mathcal{P} \right]^T + \mathbb{P}_\ell^T \mathbb{P}_\ell (H_{\text{reg}} + \lambda\mathcal{P})^{-1} \mathcal{P} \right] \\ - \lambda^2 \left[(H_{\text{reg}} + \lambda\mathcal{P})^{-1} \mathcal{P} \right]^T \mathbb{P}_\ell^T \mathbb{P}_\ell (H_{\text{reg}} + \lambda\mathcal{P})^{-1} \mathcal{P}, \end{aligned} \quad (\text{C.7})$$

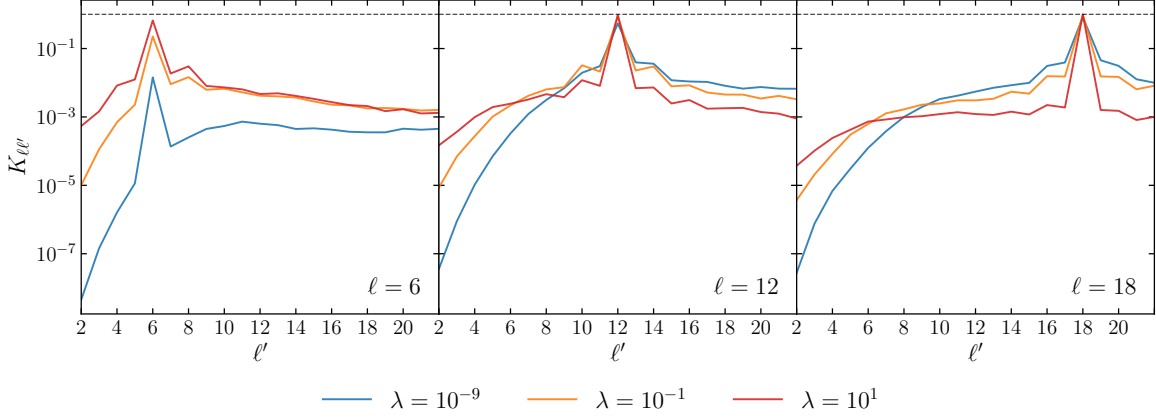


Figure 9: Values of the regularization bias matrix for different $\ell = 6, 12, 18$ large-scale multipoles (*left, center, and right panel*, respectively) for growing values of the regularization parameter λ in the NG34 pulsar configuration.

therefore, the bias introduced by ridge regularization is given by

$$C_\ell^{\text{reg}} = \frac{1}{2\ell+1} \left\langle (\mathbf{a}_{\text{LS}}^{\text{true}})^T \mathcal{K}_\ell \mathbf{a}_{\text{LS}}^{\text{true}} \right\rangle = \sum_{\ell'} \left(\frac{1}{2\ell+1} \sum_{m'} \mathcal{K}_{\ell, i_{\ell' m'} i_{\ell' m'}} \right) C_{\ell'}^{\text{true}} = \sum_{\ell'=2}^{\ell_{\text{max}}^{\text{rec}}} K_{\ell\ell'} C_{\ell'}^{\text{true}}, \quad (\text{C.8})$$

where in this instance the sum runs over all the large-scale modes.

We show in figure 9 the value of the regularization bias matrix for three different multipoles and ridge regularization parameters for the NG34 pulsar configuration. These functions sharply peak around $\ell' = \ell$; therefore in this scenario there is no strong large-scale mode coupling. However, different geometries can potentially exhibit a stronger coupling, especially between nearby multipoles. The same considerations apply also to the NG68 configuration, which is not shown for brevity.

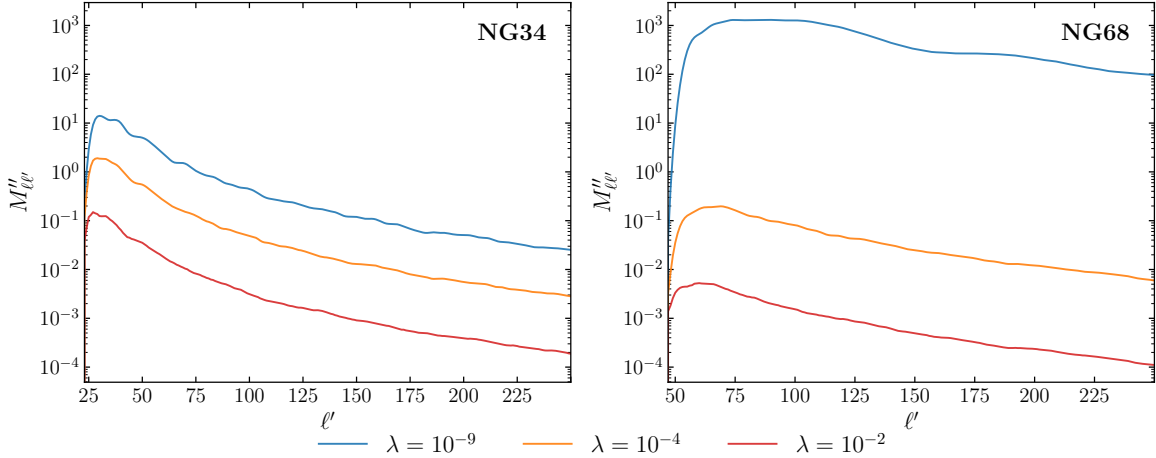


Figure 10: Values of the mode-mixing function $M''_{\ell\ell'}$ for the large-scale multipole $\ell = 18$ and growing values of the regularization parameter λ , both for the NG34 (*left panel*) and NG68 (*right panel*) pulsar configurations. The magnitude of the function is strongly dependent on the presence of near perfect degeneracies which maximizes the leakage, especially in the NG68 case.

On the other hand, in this scenario, the leakage matrix reads as

$$\mathcal{M}_\ell'' = \left[(H_{\text{reg}} + \lambda \mathcal{P})^{-1} \Gamma_{\text{LS}}^T \text{Cov}_{\text{reg}}^{-1} \Gamma_{\text{SS}} \right]^T (H_{\text{reg}} + \lambda \mathcal{P})^{-1} \Gamma_{\text{LS}}^T \text{Cov}_{\text{reg}}^{-1} \Gamma_{\text{SS}}, \quad (\text{C.9})$$

and it is responsible of an equivalent $M_{\ell\ell'}''$, as described before. We show in figure 10 how the mode-mixing matrix varies for growing values of the ridge regularization parameter, both for the NG34 and NG68 pulsar configurations. In particular, we observe how the coupling between large- and small-scale modes is effectively diminished when the strength of the regularization increases. We also observe that in the presence of nearly-degenerate configurations, as for instance in NG68, coupling between different modes is significantly boosted. The trend of these functions is consistent across the range of large-scale multipoles and thus is not shown here for conciseness.

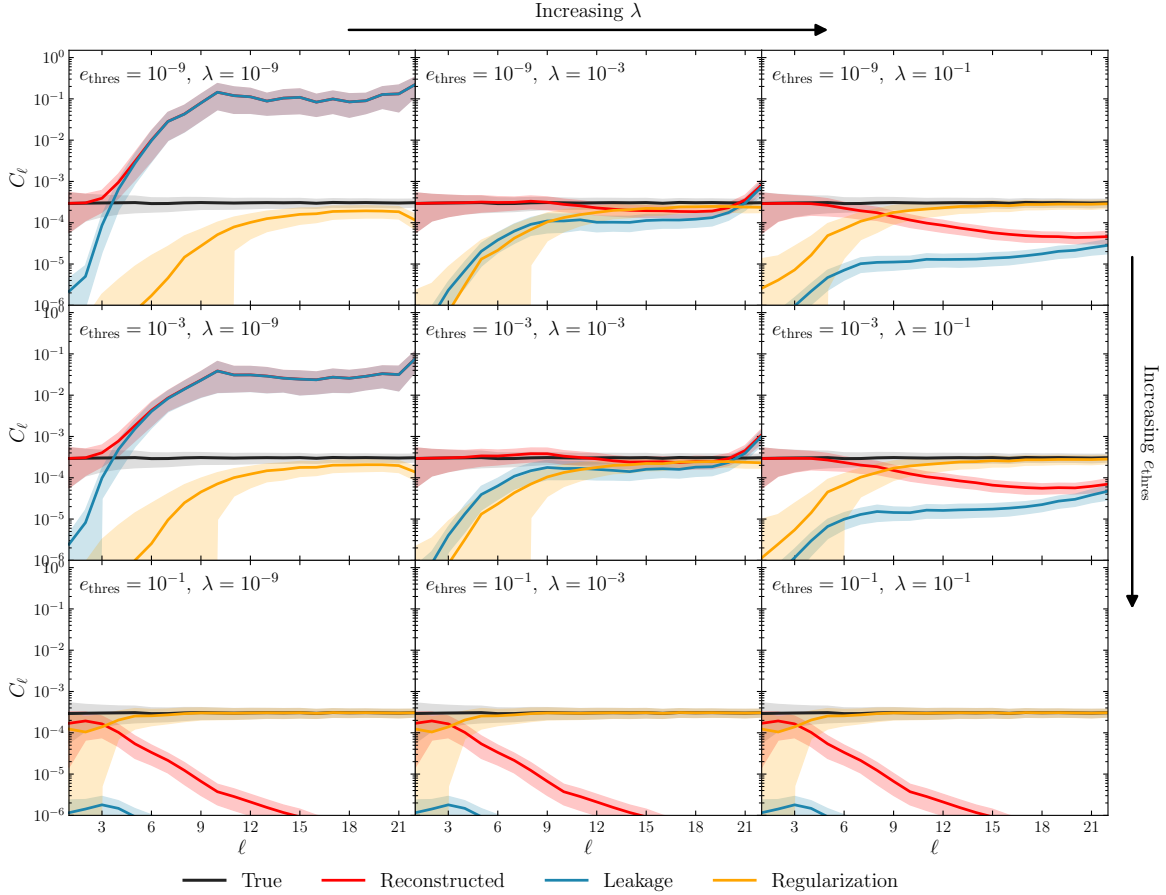


Figure 11: Reconstructed angular power spectrum for the NG34 pulsar configuration for different values of the two regularization parameters, λ (increasing from left to right) and e_{thres} (increasing from top to bottom).

Finally, we show in figure 11 how the inferred angular power spectra change when both regularization parameters λ and e_{thres} are varied independently of each other. In all these cases, we note that there exists a subtle interplay between different choices for the regularization scheme that should be carefully evaluated in creating any anisotropy reconstruction algorithm. Otherwise, the risk is to effectively remove power at small scales. This behavior is replicated identically in the NG68 configuration.

References

- [1] A. Sesana, A. Vecchio, and C. N. Colacino, “The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with Pulsar Timing Arrays”, *Mon. Not. Roy. Astron. Soc.* **390** (2008) 192, [arXiv:0804.4476 \[astro-ph\]](#).
- [2] The **NANOGrav Collaboration**, G. Agazie *et al.*, “The NANOGrav 15 yr Data Set: Constraints on Supermassive Black Hole Binaries from the Gravitational-wave Background”, *Astrophys. J. Lett.* **952** no. 2, (2023) L37, [arXiv:2306.16220 \[astro-ph.HE\]](#).
- [3] G. Sato-Polito, M. Zaldarriaga, and E. Quataert, “Where are the supermassive black holes measured by PTAs?”, *Phys. Rev. D* **110** no. 6, (2024) 063020, [arXiv:2312.06756 \[astro-ph.CO\]](#).
- [4] M. R. Sah, S. Mukherjee, V. Saeedzadeh, A. Babul, M. Tremmel, and T. R. Quinn, “Imprints of supermassive black hole evolution on the spectral and spatial anisotropy of nano-hertz stochastic gravitational-wave background”, *Mon. Not. Roy. Astron. Soc.* **533** no. 2, (2024) 1568–1582, [arXiv:2404.14508 \[astro-ph.CO\]](#).
- [5] A. A. Starobinsky, “Spectrum of relict gravitational radiation and the early state of the universe”, *JETP Lett.* **30** (1979) 682–685.
- [6] S. Matarrese, O. Pantano, and D. Saez, “General-relativistic approach to the nonlinear evolution of collisionless matter”, *Phys. Rev. D* **47** (Feb, 1993) 1311–1323.
- [7] The **NANOGrav Collaboration**, A. Afzal *et al.*, “The NANOGrav 15 yr Data Set: Search for Signals from New Physics”, *Astrophys. J. Lett.* **951** no. 1, (2023) L11, [arXiv:2306.16219 \[astro-ph.HE\]](#).
- [8] The **NANOGrav Collaboration**, G. Agazie *et al.*, “The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background”, *Astrophys. J. Lett.* **951** no. 1, (2023) L8, [arXiv:2306.16213 \[astro-ph.HE\]](#).
- [9] The **EPTA and InPTA Collaborations**, J. Antoniadis *et al.*, “The second data release from the European Pulsar Timing Array - III. Search for gravitational wave signals”, *Astron. Astrophys.* **678** (2023) A50, [arXiv:2306.16214 \[astro-ph.HE\]](#).
- [10] D. J. Reardon *et al.*, “Search for an Isotropic Gravitational-wave Background with the Parkes Pulsar Timing Array”, *Astrophys. J. Lett.* **951** no. 1, (2023) L6, [arXiv:2306.16215 \[astro-ph.HE\]](#).
- [11] H. Xu *et al.*, “Searching for the Nano-Hertz Stochastic Gravitational Wave Background with the Chinese Pulsar Timing Array Data Release I”, *Res. Astron. Astrophys.* **23** no. 7, (2023) 075024, [arXiv:2306.16216 \[astro-ph.HE\]](#).
- [12] M. T. Miles *et al.*, “The MeerKAT Pulsar Timing Array: the first search for gravitational waves with the MeerKAT radio telescope”, *Monthly Notices of the Royal Astronomical Society* **536** no. 2, (Dec., 2024) 1489–1500, [arXiv:2412.00153](#).
- [13] C. M. F. Mingarelli, T. Sidery, I. Mandel, and A. Vecchio, “Characterizing gravitational wave stochastic background anisotropy with pulsar timing arrays”, *Phys. Rev. D* **88** no. 6, (2013) 062005, [arXiv:1306.5394 \[astro-ph.HE\]](#).
- [14] S. R. Taylor and J. R. Gair, “Searching for anisotropic gravitational-wave backgrounds using pulsar timing arrays”, *Physical Review D* **88** no. 8, (Oct, 2013), [arXiv:1306.5395](#).
- [15] C. M. F. Mingarelli, T. J. W. Lazio, A. Sesana, J. E. Greene, J. A. Ellis, C.-P. Ma, S. Croft, S. Burke-Spolaor, and S. R. Taylor, “The Local Nanohertz Gravitational-Wave Landscape From Supermassive Black Hole Binaries”, *Nature Astron.* **1** no. 12, (2017) 886–892, [arXiv:1708.03491 \[astro-ph.GA\]](#).
- [16] E. C. Gardiner, L. Z. Kelley, A.-M. Lemke, and A. Mitridate, “Beyond the Background: Gravitational-wave Anisotropy and Continuous Waves from Supermassive Black Hole Binaries”, *ApJ Lett.* **965** no. 2, (Apr., 2024) 164, [arXiv:2309.07227 \[astro-ph.HE\]](#).
- [17] The **NANOGrav Collaboration**, G. Agazie *et al.*, “The NANOGrav 15 yr Data Set: Search for Anisotropy in the Gravitational-wave Background”, *Astrophys. J. Lett.* **956** no. 1, (2023) L3, [arXiv:2306.16221 \[astro-ph.HE\]](#).

- [18] K. Grunthal *et al.*, “The MeerKAT Pulsar Timing Array: Maps of the gravitational wave sky with the 4.5-yr data release”, *Monthly Notices of the Royal Astronomical Society* **536** no. 2, (Dec., 2024) 1501–1517.
- [19] J. A. Casey-Clyde, C. M. F. Mingarelli, J. E. Greene, K. Pardo, M. Nañez, and A. D. Goulding, “A Quasar-based Supermassive Black Hole Binary Population Model: Implications for the Gravitational Wave Background”, *Astrophys. J.* **924** no. 2, (2022) 93, [arXiv:2107.11390 \[astro-ph.HE\]](#).
- [20] G. Agazie *et al.*, “The NANOGrav 15 yr Data Set: Looking for Signs of Discreteness in the Gravitational-wave Background”, (4, 2024) , [arXiv:2404.07020 \[astro-ph.HE\]](#).
- [21] L.-W. Tian, Y.-C. Bi, Y.-M. Wu, and Q.-G. Huang, “Targeted search for an individual SMBHB in NANOGrav 15-year and EPTA DR2 data sets”, [arXiv:2508.14742 \[astro-ph.GA\]](#).
- [22] The **NANOGrav Collaboration**, N. Agarwal *et al.*, “The NANOGrav 15 yr Data Set: Targeted Searches for Supermassive Black Hole Binaries”, [arXiv:2508.16534 \[astro-ph.HE\]](#).
- [23] J. A. Casey-Clyde, C. M. F. Mingarelli, J. E. Greene, A. D. Goulding, S. Chen, and J. R. Trump, “Quasars Can Signpost Supermassive Black Hole Binaries”, *ApJ Lett.* **987** no. 2, (July, 2025) 106, [arXiv:2405.19406 \[astro-ph.HE\]](#).
- [24] N. Bellomo, D. Bertacca, A. C. Jenkins, S. Matarrese, A. Raccanelli, T. Regimbau, A. Ricciardone, and M. Sakellariadou, “CLASS_GWB: robust modeling of the astrophysical gravitational wave background anisotropies”, *JCAP* **06** no. 06, (2022) 030, [arXiv:2110.15059 \[gr-qc\]](#).
- [25] F. Semenzato, J. A. Casey-Clyde, C. M. F. Mingarelli, A. Raccanelli, N. Bellomo, N. Bartolo, and D. Bertacca, “Cross-Correlating the Universe: The Gravitational Wave Background and Large-Scale Structure”, [arXiv:2411.00532 \[astro-ph.CO\]](#).
- [26] M. R. Sah and S. Mukherjee, “Discovering the Cosmic Evolution of Supermassive Black Holes using Nano-Hertz Gravitational Waves and Galaxy Surveys”, (7, 2024) , [arXiv:2407.11669 \[astro-ph.CO\]](#).
- [27] N. Pol, S. R. Taylor, and J. D. Romano, “Forecasting Pulsar Timing Array Sensitivity to Anisotropy in the Stochastic Gravitational Wave Background”, *ApJ Lett.* **940** no. 2, (Dec., 2022) 173, [arXiv:2206.09936 \[astro-ph.HE\]](#).
- [28] T. Konstandin, A.-M. Lemke, A. Mitridate, and E. Perboni, “The impact of cosmic variance on PTAs anisotropy searches”, [arXiv:2408.07741 \[astro-ph.CO\]](#).
- [29] T. Konstandin, A.-M. Lemke, A. Mitridate, and E. Perboni, “Prospects and Limitations of PTAs Anisotropy Searches – The Frequentist Case”, 2025. <https://arxiv.org/abs/2509.07074>.
- [30] V. Domcke, G. Franciolini, and M. Pieroni, “Cosmic Variance in Anisotropy Searches at Pulsar Timing Arrays”, [arXiv:2508.21131 \[astro-ph.CO\]](#).
- [31] Y. Ali-Haïmoud, T. L. Smith, and C. M. F. Mingarelli, “Fisher formalism for anisotropic gravitational-wave background searches with pulsar timing arrays”, *Phys. Rev. D* **102** no. 12, (2020) 122005, [arXiv:2006.14570 \[gr-qc\]](#).
- [32] S. R. Taylor, R. van Haasteren, and A. Sesana, “From Bright Binaries To Bumpy Backgrounds: Mapping Realistic Gravitational Wave Skies With Pulsar-Timing Arrays”, *Phys. Rev. D* **102** no. 8, (2020) 084039, [arXiv:2006.04810 \[astro-ph.IM\]](#).
- [33] W. Hu, M. M. Hedman, and M. Zaldarriaga, “Benchmark parameters for CMB polarization experiments”, *Physical Review D* **67** no. 4, (Feb., 2003) , [arXiv:astro-ph/0210096](#).
- [34] D. Alonso, J. Sanchez, and A. Slosar, “A unified pseudo- $\mathcal{C}\ell$ framework”, *Monthly Notices of the Royal Astronomical Society* **484** no. 3, (Jan., 2019) 4127–4151, [arXiv:1809.09603](#).
- [35] R. M. Sullivan, L. T. Hergt, and D. Scott, “Methods for CMB map analysis”, [arXiv:2410.12951 \[astro-ph.CO\]](#).
- [36] R. W. Hockney and J. W. Eastwood, *Computer Simulation Using Particles*. 1981.
- [37] E. Sefusatti, M. Crocce, R. Scoccimarro, and H. M. P. Couchman, “Accurate estimators of correlation functions in Fourier space”, *Monthly Notices of the Royal Astronomical Society* **460** no. 4, (May, 2016) 3624–3636, [arXiv:1512.07295](#).

- [38] N. Hand, Y. Feng, F. Beutler, Y. Li, C. Modi, U. Seljak, and Z. Slepian, “nbodykit: An Open-source, Massively Parallel Toolkit for Large-scale Structure”, *The Astronomical Journal* **156** no. 4, (Sept., 2018) 160, [arXiv:1712.05834](#).
- [39] C. M. F. Mingarelli and A. B. Mingarelli, “Proving the short-wavelength approximation in Pulsar Timing Array gravitational-wave background searches”, *Journal of Physics Communications* **2** no. 10, (Oct., 2018) 105002, [arXiv:1806.06979](#) [[astro-ph.IM](#)].
- [40] R. W. Hellings and G. S. Downs, “Upper limits on the isotropic gravitational radiation background from pulsar timing analysis”, *Astrophys. J. Lett.* **265** (1983) L39–L42.
- [41] B. E. Moreschi, S. Valtolina, A. Sesana, G. Shaifullah, M. Falxa, L. Speri, D. Izquierdo-Villalba, and A. Chalumeau, “Dissecting the nanoHz gravitational wave sky: frequency-correlated anisotropy induced by eccentric supermassive black hole binaries”, [arXiv:2506.14882](#) [[astro-ph.GA](#)].
- [42] M. R. Sah, A. Maurya, S. Mukherjee, P. Kumar, V. Saeedzadeh, A. Babul, C. K. Mishra, K. Paul, T. R. Quinn, and M. Tremmel, “An Accurate Modeling of Nano-Hertz Gravitational Wave Signal from Eccentric Supermassive Binary Black Holes: Essential Step Towards a Robust Discovery”, [arXiv:2505.22745](#) [[astro-ph.CO](#)].
- [43] Y. Ali-Haïmoud, T. L. Smith, and C. M. F. Mingarelli, “Insights into searches for anisotropies in the nanohertz gravitational-wave background”, *Phys. Rev. D* **103** no. 4, (2021) 042009, [arXiv:2010.13958](#) [[gr-qc](#)].
- [44] J. D. Romano and N. J. Cornish, “Detection methods for stochastic gravitational-wave backgrounds: a unified treatment”, *Living Reviews in Relativity* **20** no. 1, (Apr., 2017) , [arXiv:1608.06889](#).
- [45] The **NANOGrav Collaboration**, G. Agazie *et al.*, “The NANOGrav 15 yr Data Set: Detector Characterization and Noise Budget”, *Astrophys. J. Lett.* **951** no. 1, (2023) L10, [arXiv:2306.16218](#) [[astro-ph.HE](#)].
- [46] J. L. Bernal, N. Bellomo, A. Raccanelli, and L. Verde, “Beware of commonly used approximations. Part II. Estimating systematic biases in the best-fit parameters”, *Journal of Cosmology and Astroparticle Physics* **2020** no. 10, (Oct., 2020) 017–017, [arXiv:2005.09666](#).