12-15-2020

Fisher Formalism For Anisotropic Gravitational-Wave Background Searches With Pulsar Timing Arrays

Y. Ali-Haïmoud

Tristan L. Smith
Swarthmore College, tsmith2@swarthmore.edu

C. M. F. Mingarelli

Recommended Citation

Y. Ali-Haïmoud, Tristan L. Smith, and C. M. F. Mingarelli. (2020). "Fisher Formalism For Anisotropic Gravitational-Wave Background Searches With Pulsar Timing Arrays". *Physical Review D*. Volume 102, Issue 12. DOI: 10.1103/PhysRevD.102.122005
https://works.swarthmore.edu/fac-physics/428

Fisher formalism for anisotropic gravitational-wave background searches with pulsar timing arrays

Yacine Ali-Haïmoud,[1] Tristan L. Smith,[2] and Chiara M. F. Mingarelli[3,4]

[1]*Center for Cosmology and Particle Physics, Department of Physics, New York University, New York, New York 10003, USA*

[2]*Department of Physics and Astronomy, Swarthmore College, 500 College Avenue, Swarthmore, Pennsylvania 19081, USA*

[3]*Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, New York 10010, USA*

[4]*Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs, Connecticut 06269-3046, USA*

(Received 3 July 2020; accepted 17 November 2020; published 22 December 2020)

Pulsar timing arrays (PTAs) are currently the only experiments directly sensitive to gravitational waves with decade-long periods. Within the next five to ten years, PTAs are expected to detect the stochastic gravitational-wave background (SGWB) collectively sourced by inspiraling supermassive black hole binaries. It is expected that this background is mostly isotropic, and current searches focus on the monopole part of the SGWB. Looking ahead, anisotropies in the SGWB may provide a trove of additional information on both known and unknown astrophysical and cosmological sources. In this paper, we build a simple yet realistic Fisher formalism for anisotropic SGWB searches with PTAs. Our formalism is able to accommodate realistic properties of PTAs and allows simple and accurate forecasts. We illustrate our approach with an idealized PTA consisting of identical, isotropically distributed pulsars. In a companion paper, we apply our formalism to current PTAs and show that it can be a powerful tool to guide and optimize real data analysis.

DOI: 10.1103/PhysRevD.102.122005

I. INTRODUCTION

The promise of timing pulsars to detect nanohertz (nHz) gravitational waves (GWs) was pointed out more than four decades ago [1,2], and the application to a *stochastic* gravitational-wave background (SGWB) was studied shortly after [3]. In that seminal paper, Hellings and Downs derived the response of pulsar timing residual correlations to an isotropic SGWB, and were the first to combine several pulsars to extract an upper limit on the SGWB amplitude. Since then, several collaborations [4–7] have been timing arrays of pulsars, some for over two decades, and getting an increasingly stringent upper limit on the SGWB amplitude [8–10]. If our understanding of galaxy formation and merger history is correct, pulsar timing arrays (PTAs) should detect the SGWB generated by inspiraling supermassive black hole binaries (SMBHBs) within the next decade [11–13]. In addition to this astrophysical background, other, more exotic processes could also contribute to the nHz SGWB; see, e.g., Refs. [12,14,15].

While most current searches assume perfect isotropy, there is likely some level of anisotropy in the SGWB.[1] For one, the finite number of SMBHBs should inevitably imply some level of anisotropy [16–18]. On large scales, one also expects the distribution of SMBHBs to trace cosmic structure [12]. Independent of their physical origin, it is important to understand what kind and what level of SGWB anisotropy PTAs might in principle be able to detect. Indeed, such anisotropy might eventually prove a powerful discriminant between different models of the SGWB.

The standard approach to study the detectability of SGWB anisotropies has been to harness the full power of Bayesian analysis pipelines used for real data [19,20]. While this approach provides the most accurate results, it is computationally demanding and does not allow for making quick estimates or building intuition. Recently, the authors of Ref. [21] developed a simplified approach, holding for

[1]Given that the SGWB is the power spectrum of the gravitational-wave strain, we are technically referring to *statistical* anisotropy. However, since the gravitational-wave strain itself is *necessarily* anisotropic (the only isotropic rank-2 and trace-free tensor is the null tensor), we will drop the qualifier "statistical" when referring to anisotropies of the (scalar) SGWB intensity, as there is no risk of confusion.

an idealized PTA constituted of a large number of identical, isotropically distributed pulsars.[2] In this paper, we fill the gap between these two methods, by deriving a simple, yet realistic, Fisher formalism for SGWB anisotropies. We moreover break away from the spherical-harmonic decomposition of the SGWB that most past works have relied on thus far [17,20,22,23], as we argue it is poorly adapted to real PTAs (see Ref. [24] for a different mapping approach, in the case of continuous waves). Our formalism allows for real pulsar distributions and noise properties, and yet it permits us to make quick detectability estimates without running time-consuming Monte Carlo Markov chains. While our formalism is not a substitute for a full-on data analysis, it provides useful tools to make *forecasts*, as well as *guide* and *optimize* SGWB searches with PTAs. Our philosophy is inspired by what has long been the norm in the field of cosmology, where Fisher analyses (e.g., [25,26]) are routinely used and have proven extremely useful, not only to produce forecasts, but also to make the field accessible to a broader community.

This work is the first of a series of two articles. In the present paper, we expound the theoretical formalism, culminating in the derivation of the PTA Fisher matrix for an anisotropic SGWB, Eq. III C. Along the way, we rederive some known results with a fresh approach, using only frame-independent, geometric expressions. In the companion paper (hereafter, Paper II), we shall apply this tool to several practical examples, and illustrate how to make forecasts or optimize searches for SGWB anisotropies.

This paper is organized as follows. In Sec. II, we start by describing the statistical properties of the SGWB in a new frame-independent, geometric fashion, and then derive the response of a timing array to the SGWB. In Sec. III, we derive the Fisher matrix for the SGWB intensity, which is our main result. In Sec. IV, we apply our results to the idealized case of a dense array of identical pulsars isotropically distributed on the sky. We present a new calculation of the Hellings and Downs curve in Appendix A, derive the covariance matrix of time residual band powers in Appendix B, and compute the Fisher matrix in the limit of a large number of identical, isotropically distributed pulsars in Appendix C. Throughout we use units in which the speed of light is unity. A summary of our notation can be found in Table I.

II. GEOMETRIC DESCRIPTION OF THE RESPONSE OF PULSAR PAIRS TO A SGWB

A. Geometric decomposition of the SGWB power spectrum

In this section we present a new, geometric, and frame-independent decomposition of the SGWB power spectrum. In Sec. II B, we relate the new expressions, Eqs. (6), (9), and (11), to those commonly found in the PTA literature. Our frame-independent expressions will prove very powerful later on as they allow us to express all relevant observables through explicit functions of scalar products between unit vectors.

We decompose the GW strain $h_{ab}(t,\vec{x})$ in the Fourier domain as follows:

$$h_{ab}(t,\vec{x}) = \int_{-\infty}^{\infty} df \int d^2\,\hat{\Omega}\, h_{ab}(f,\hat{\Omega}) e^{2i\pi f(t-\hat{\Omega}\cdot\vec{x})}, \quad (1)$$

where $h_{ab}(f,\hat{\Omega})$ is symmetric, trace-free, and transverse to the direction of propagation $\hat{\Omega}$, i.e., $\hat{\Omega}^a h_{ab}(f,\hat{\Omega}) = 0$.

If we assume that the SGWB is a stationary Gaussian random field (as would be the case if it is generated by a large number of uncorrelated sources), it is entirely determined by its power spectrum $\mathcal{P}_{abcd}$, which we normalize as follows:

$$\langle h_{ab}(f,\hat{\Omega}) h^*_{cd}(f',\hat{\Omega}')\rangle = \frac{\delta_{\rm D}(\hat{\Omega}',\hat{\Omega})}{4\pi}\frac{\delta_{\rm D}(f'-f)}{2} \times \mathcal{P}_{abcd}(f,\hat{\Omega}), \quad (2)$$

where $\delta_{\rm D}$ is the Dirac function. The Dirac function in frequency stems from time-translation invariance (i.e., stationarity) of the correlation function $\langle h_{ab}(t) h_{cd}(t+\Delta t)\rangle$, and the angular Dirac function $\delta_{\rm D}(\hat{\Omega}',\hat{\Omega})$ stems from spatial-translation invariance (i.e., statistical isotropy).

The definition (2) implies the following Hermiticity property:

$$\mathcal{P}_{cdab}(f,\hat{\Omega}) = \mathcal{P}^*_{abcd}(f,\hat{\Omega}). \quad (3)$$

In addition, the reality of $h_{ab}(t,\vec{x})$ implies that $h_{ab}(-f,\hat{\Omega}) = h^*_{ab}(f,\hat{\Omega})$, in turn implying

$$\mathcal{P}_{abcd}(-f,\hat{\Omega}) = \mathcal{P}^*_{abcd}(f,\hat{\Omega}). \quad (4)$$

The GW power spectrum $\mathcal{P}_{abcd}$ is a rank-4 tensor, which is symmetric and trace-free for the first and last pairs of indices, and transverse to $\hat{\Omega}$ in each index. Hence it has four independent components, which are *a priori* complex. The Hermiticity property (3) reduces the number of independent components to four *real* quantities. This is the same number of components as the (rank-2) electromagnetic intensity tensor, i.e., the power spectrum of the electromagnetic field. Just as the latter, we may decompose $\mathcal{P}_{abcd}$ into a component proportional to its trace (the total intensity)

$$\mathcal{I}(f,\hat{\Omega}) \equiv \frac{1}{4}\mathcal{P}_{abab}(f,\hat{\Omega}), \quad (5)$$

[2]This underlying assumption is not explicitly spelled out in Ref. [21], but is required for the harmonic transform of the timing residuals to be uncorrelated and have ℓ-independent noise, as assumed there.

TABLE I. Summary of the notation used in this paper, in alphabetic order, with the defining equations.

Symbol	Description	Dimensions	Defining equation
$\mathbf{1}(\hat{\Omega})$	Isotropic map equal to unity for all directions $\hat{\Omega}$	Dimensionless	$\mathbf{1}(\hat{\Omega}) = 1 \ \forall \ \hat{\Omega}$
$\mathcal{C}_{IJ}$	Covariance of estimators of time-residual band power	Time^4	$\mathcal{C}_{IJ} \equiv \text{cov}(\hat{\mathcal{R}}_I, \hat{\mathcal{R}}_J)$
$\delta_{ab}^{\perp\hat{\Omega}}$	Identity tensor in the plane orthogonal to $\hat{\Omega}$	Dimensionless	(10)
Δt_p	Observation cadence of pulsar p	Time	
Δf	Frequency bandwidth for band powers	Frequency	
f	Gravitational-wave frequency	Frequency	(1)
$\mathcal{F}(\hat{\Omega}, \hat{\Omega}')$	Fisher matrix of band-integrated GW intensity	Dimensionless	(55)
$F(\hat{\Omega}, \hat{\Omega}')$	Reduced Fisher matrix for identical pulsars	Dimensionless	(69)
F_{IJ}	$N_{\text{pair}} \times N_{\text{pair}}$ discretized reduced Fisher matrix	Dimensionless	(84)
$\boldsymbol{\gamma}_{\hat{p}\,\hat{q}}(\hat{\Omega}) = \boldsymbol{\gamma}_I(\hat{\Omega})$	Pairwise timing response function at pulsar pair $I = (p, q)$	Dimensionless	II D
$\boldsymbol{\gamma}_I^*(\hat{\Omega})$	Dual map of $\boldsymbol{\gamma}_I(\hat{\Omega})$	Dimensionless	$\boldsymbol{\gamma}_I^* \cdot \boldsymbol{\gamma}_J = \delta_{IJ}$
$h_{ab}(t, \hat{\Omega})$	Gravitational-wave strain	Dimensionless	(27)
$h_{ab}(f, \hat{\Omega})$	Fourier transform of $h_{ab}(t)$	1/Frequency	(1)
$h_c(f)$	Characteristic gravitational-wave strain	Dimensionless	(13)
$\mathcal{H}(\mu)$	Hellings and Downs function (response to an isotropic SGWB)	Dimensionless	(45)
I, J	Labels of unique pulsar pairs	Indices	$I = (p, q)$, $J = (p', q')$
$\mathcal{I}(f, \hat{\Omega})$	Total intensity of the SGWB	1/Frequency	(5)
$\mathcal{I}_f(\hat{\Omega})$ or $\mathcal{I}(\hat{\Omega})$	Band-integrated SGWB intensity	Dimensionless	(49)
$\mathfrak{T}_{abcd}(\hat{\Omega})$	Geometric dependence of the SGWB total intensity piece	Dimensionless	(9)
$\mathcal{L}_{abcd}(f, \hat{\Omega})$	Linear polarization tensor of the SGWB	1/Frequency	(6)
$N_p(t)$	Intrinsic time-residual noise of pulsar p	Time	(47)
$N_p(f)$	Fourier transform of $N_p(t)$	Time/Frequency	
$\hat{\Omega}$	Gravitational-wave direction of propagation	Unit vector	(1)
p, q	Labels of individual pulsars	Indices	
$\hat{p}$, $\hat{q}$	Unit vectors pointing in the direction of individual pulsars	Unit vectors	
$\mathcal{P}_{abcd}(f, \hat{\Omega})$	Rank-4 power spectrum of the SGWB	1/Frequency	(2)
$R_p^{\text{GW}}(t)$	Gravitational-wave-induced time residual of pulsar p	Time	(30)
$R_p^{\text{GW}}(f)$	Fourier transform of $R_p^{\text{GW}}(t)$	Time/Frequency	(32)
$R_p(t)$	Total time residual of pulsar p	Time	(47)
$R_p(f)$	Fourier transform of $R_p(t)$	Time/Frequency	
$\mathcal{R}_I(f) = \mathcal{R}_{pq}(f)$	Cross-power spectrum of time residuals of pulsar pair $I = (p, q)$	Time^2/Frequency	(33)
$\mathcal{R}_{I,f}$ or $\mathcal{R}_I$	Band-integrated time-residual power spectrum	Time^2	(51)
$\hat{\mathcal{R}}_I$	Estimator for $\mathcal{R}_I$	Time^2	
$\sigma_p^2(f)$	Pulsar timing noise power spectrum	$\text{Time}^2/\text{Frequency}$	(48)
$\sigma_{p,f}^2$ or σ_p^2	Band-integrated pulsar timing noise	Time^2	(50)
$\mathcal{T}_p(f)$	Timing-model-fitting transmission function of pulsar p	Dimensionless	III C
T_p	Total observation time of pulsar p	Time	
T_{pq}	Effective total observation time of pulsar pair p, q	Time	$T_{pq} = \min(T_p, T_q)$
$\mathcal{V}(f, \hat{\Omega})$	Circular polarization amplitude of the SGWB	1/Frequency	(12)
$\mathfrak{V}_{abcd}(\hat{\Omega})$	Geometric dependence of the SGWB circular polarization piece	Dimensionless	(11)
$\mathcal{Y}_{\ell m}(\hat{\Omega})$	Real spherical harmonics	Dimensionless	

a component proportional to a real pseudoscalar $\mathcal{V}(f, \hat{\Omega})$ (the circular polarization), and a real, fully trace-free linear-polarization tensor $\mathcal{L}_{abcd}(f, \hat{\Omega})$, carrying the two remaining independent components. More specifically, we want to decompose the power spectrum as follows:

$$\mathcal{P}_{abcd}(f, \hat{\Omega}) = \mathcal{I}(f, \hat{\Omega})\mathfrak{T}_{abcd}(\hat{\Omega}) + i\mathcal{V}(f, \hat{\Omega})\mathfrak{V}_{abcd}(\hat{\Omega}) + \mathcal{L}_{abcd}(f, \hat{\Omega}), \tag{6}$$

where the (real) tensors $\mathfrak{T}_{abcd}(\hat{\Omega})$ and $\mathfrak{V}_{abcd}(\hat{\Omega})$ are purely geometric, frequency-independent objects. The traces of the tensors appearing in Eq. (6) are

$$\mathfrak{T}_{abab} = 4, \qquad \mathfrak{V}_{abab} = 0, \qquad \mathcal{L}_{abad} = 0. \tag{7}$$

The reality condition (4) implies that $\mathcal{I}(f, \hat{\Omega})$ and $\mathcal{L}_{abcd}(f, \hat{\Omega})$ are even functions of f, while $\mathcal{V}(f, \hat{\Omega})$ is an odd function of f.

The geometric objects $\mathfrak{T}_{abcd}(\hat{\Omega})$ and $\mathfrak{V}_{abcd}(\hat{\Omega})$ must be built exclusively out of isotropic tensors, i.e., the Kroneker delta and the Levi-Civita tensor ϵ_{abc}, and of $\hat{\Omega}$, which is the only preferred direction. For $\mathcal{I}$ to be a real scalar, the tensor $\mathfrak{T}_{abcd}$ must only contain Kronecker deltas and $\hat{\Omega}$, i.e., be of the form

$$\mathfrak{T}_{abcd}(\hat{\Omega}) \propto \delta_{ab}\delta_{cd}, \ldots, \delta_{ab}\hat{\Omega}_c\hat{\Omega}_d, \ldots, \hat{\Omega}_a\hat{\Omega}_b\hat{\Omega}_c\hat{\Omega}_d, \quad (8)$$

where the ellipses include all permutations of indices. By imposing that $\mathfrak{T}_{abcd}$ has the symmetries of $\mathcal{P}_{abcd}$, and, from Eq. (3), is symmetric under exchange of the first and last pairs of indices, one finds that the *only* rank-4 tensor satisfying these properties, with the appropriate normalization $\mathfrak{T}_{abab} = 4$ is

$$\mathfrak{T}_{abcd}(\hat{\Omega}) = \delta^{\perp\hat{\Omega}}_{ac}\delta^{\perp\hat{\Omega}}_{bd} + \delta^{\perp\hat{\Omega}}_{ad}\delta^{\perp\hat{\Omega}}_{bc} - \delta^{\perp\hat{\Omega}}_{ab}\delta^{\perp\hat{\Omega}}_{cd}, \quad (9)$$

where $\delta^{\perp\hat{\Omega}}_{ab}$ is the identity tensor projected on the plane orthogonal to $\hat{\Omega}$,

$$\delta^{\perp\hat{\Omega}}_{ab} \equiv \delta_{ab} - \hat{\Omega}_a\hat{\Omega}_b. \quad (10)$$

For $\mathcal{V}$ to be a pseudoscalar, the geometric object $\mathfrak{V}_{abcd}(\hat{\Omega})$ must be proportional to $\hat{\Omega}^a\epsilon_{abc}$, and otherwise be built out of Kronecker deltas and $\hat{\Omega}$. It must have the same symmetry properties as $\mathcal{P}_{abcd}$ and satisfy $\mathfrak{V}_{cdab} = -\mathfrak{V}_{abcd}$. Up to a proportionality constant (which we chose in order to match existing derivations, as we will see shortly), the only possible tensor with the appropriate symmetry properties is

$$\mathfrak{V}_{abcd}(\hat{\Omega}) = \hat{\Omega}^e(\epsilon_{ea(c}\delta^{\perp\hat{\Omega}}_{d)b} + \epsilon_{eb(c}\delta^{\perp\hat{\Omega}}_{d)a}), \quad (11)$$

where $X_{(ab)} \equiv (X_{ab} + X_{ba})/2$ represents symmetrization. With this convention, we have $\mathfrak{V}_{abad} = 2\hat{\Omega}^e\epsilon_{ebd}$; thus the amplitude of circular polarization $\mathcal{V}$ can be obtained from

$$\mathcal{V} = \frac{1}{4i}\mathcal{P}_{abad}\epsilon_{bde}\hat{\Omega}^e. \quad (12)$$

Finally, $\mathcal{L}_{abad}$ contains information about the *linear* polarization of the SGWB.

To conclude, Eqs. (6), (9), and (11) form a geometric, frame-independent decomposition of the GW power spectrum, with the most general frequency and angular dependence. In the remainder of this paper, we will focus on the total-intensity part of the SGWB, i.e., assume that the circular and linear polarization components are subdominant (it is conceptually straightforward to generalize our formalism to a polarized SGWB). In the majority of works on the SGWB, the intensity is assumed to be isotropic, $\mathcal{I}(f,\hat{\Omega}) = \mathcal{I}(f)$. In this case, the SGWB intensity is just half of the one-sided GW strain spectral density $S_h(f) = h_c^2(f)/f$, where $h_c(f)$ is the characteristic strain [27]. More generally, these quantities are related to the angle average of $\mathcal{I}(f,\hat{\Omega})$ through

$$S_h(f) = \frac{h_c^2(f)}{f} = 2\int\frac{d\hat{\Omega}}{4\pi}\mathcal{I}(f,\hat{\Omega}), \quad (13)$$

as can be seen from taking the trace of Eq. (2) and integrating it over angles, and comparing to, e.g., Ref. [28]. The authors of Ref. [17] consider the possibility of an anisotropic GW intensity, with a factorized frequency and angular dependence. Their convention corresponds to $\mathcal{I}(f,\hat{\Omega}) = 8\pi H(f)P(\hat{\Omega})$. A similar convention is adopted (up to a factor of 2) in Ref. [22].

B. Connection with standard frame-dependent notation

We now relate the geometric, frame-independent description given above to the more standard frame-dependent expressions found in the literature. For a given direction of GW propagation $\hat{\Omega}$, one may pick two arbitrary vectors $\hat{m}$ and $\hat{n} = \hat{\Omega}\times\hat{m}$ orthogonal to $\hat{\Omega}$, and define the two polarization basis tensors

$$e^+_{ab} \equiv \hat{m}_a\hat{m}_b - \hat{n}_a\hat{n}_b, \qquad e^\times_{ab} \equiv \hat{m}_a\hat{n}_b + \hat{n}_a\hat{m}_b. \quad (14)$$

Since the triad $\hat{m}$, $\hat{n}$, $\hat{\Omega}$ forms an orthonormal basis, we have

$$\hat{m}_a\hat{m}_b + \hat{n}_a\hat{n}_b = \delta_{ab} - \hat{\Omega}_a\hat{\Omega}_b = \delta^{\perp\hat{\Omega}}_{ab}, \quad (15)$$

$$\hat{m}_b\hat{n}_d - \hat{n}_b\hat{m}_d = \hat{\Omega}^e\epsilon_{ebd}, \quad (16)$$

independent of the choice of $\hat{m}$, $\hat{n}$. From these expressions, one can show that the tensors $\mathfrak{T}_{abcd}$ and $\mathfrak{V}_{abcd}$ defined in Eqs. (9) and (11) are given by

$$\mathfrak{T}_{abcd} = e^+_{ab}e^+_{cd} + e^\times_{ab}e^\times_{cd}, \quad (17)$$

$$\mathfrak{V}_{abcd} = e^+_{ab}e^\times_{cd} - e^\times_{ab}e^+_{cd}. \quad (18)$$

Let us now project the strain onto the basis e^+_{ab}, $e^\times_{ab}$:

$$h_{ab}(f,\hat{\Omega}) = h_+(f,\hat{\Omega})e^+_{ab} + h_\times(f,\hat{\Omega})e^\times_{ab}. \quad (19)$$

For the sake of compactness, for any two stochastic variables X, Y, we define the quantity $\langle XY^*\rangle'$ such that

$$\langle X(f,\hat{\Omega})Y^*(f',\hat{\Omega}')\rangle = \frac{\delta_{\rm D}(\hat{\Omega}',\hat{\Omega})}{4\pi}\frac{\delta_{\rm D}(f'-f)}{2}\langle XY^*\rangle'. \quad (20)$$

In words, $\langle XY^*\rangle'$ is the cross-power spectrum of X and Y. With this convention, the GW power spectrum is

$$\mathcal{P}_{abcd} = \langle h_{ab} h^*_{cd} \rangle' = \langle |h_+|^2 \rangle' e^+_{ab} e^+_{cd} + \langle h_+ h^*_\times \rangle' e^+_{ab} e^\times_{cd} + \langle h_\times h^*_+ \rangle' e^\times_{ab} e^+_{cd} + \langle |h_\times|^2 \rangle' e^\times_{ab} e^\times_{cd}. \quad (21)$$

Now using Eqs. (17) and (18), we see that we can write the GW power spectrum in the form (6), with

$$\mathcal{I} = \frac{1}{2} \langle |h_+|^2 + |h_\times|^2 \rangle', \quad (22)$$

$$\mathcal{V} = \frac{1}{2i} \langle h_+ h^*_\times - h_\times h^*_+ \rangle' = \mathrm{Im} \langle h_+ h^*_\times \rangle', \quad (23)$$

$$\mathcal{L}_{abcd} = \mathcal{Q}(e^+_{ab} e^+_{cd} - e^\times_{ab} e^\times_{cd}) + \mathcal{U}(e^+_{ab} e^\times_{cd} + e^\times_{ab} e^+_{cd}), \quad (24)$$

$$\mathcal{Q} \equiv \frac{1}{2} \langle |h_+|^2 - |h_\times|^2 \rangle', \quad (25)$$

$$\mathcal{U} \equiv \frac{1}{2} \langle h_+ h^*_\times + h_\times h^*_+ \rangle' = \mathrm{Re} \langle h_+ h^*_\times \rangle'. \quad (26)$$

These relations clearly show the analogy with the standard Stokes parameters of electromagnetic waves (see, e.g., [29]). Our normalization matches precisely that of Ref. [30]—see Ref. [31] for similar expressions, with a different normalization.

C. Concise derivation of the timing residuals from GWs

A common derivation of the time-residual induced by GWs consists of deriving expressions for a null geodesic in the presence of gravitational plane waves using Killing vectors [2,32,33]. Here we provide a new and concise derivation in the spirit of the first calculation by Ref. [1]. Our derivation has the advantage of not being limited to a plane wave, but directly applies to a generic superposition of waves, with no special symmetries and hence no Killing vector fields. Consider null geodesics in the metric

$$ds^2 = -dt^2 + (\delta_{ab} + h_{ab}) dx^a dx^b, \quad (27)$$

where the GW strain $h_{ab}(t, \vec{x})$ is symmetric, trace-free, and transverse ($\partial^a h_{ab} = 0$). Specifically, consider light rays originating at a pulsar p (event P) and received on Earth (event E). We define $d\ell^2 = \delta_{ab} dx^a dx^b$. The null geodesic condition implies that

$$dt = (d\ell^2 + h_{ab} dx^a dx^b)^{1/2} = d\ell \left(1 + h_{ab} \frac{dx^a}{d\ell} \frac{dx^b}{d\ell}\right)^{1/2} = d\ell \left(1 + \frac{1}{2} h_{ab} \frac{dx^a}{d\ell} \frac{dx^b}{d\ell}\right) + \mathcal{O}(h^2), \quad (28)$$

where we have expanded to linear order in h_{ab}. At this order, we only need to compute $dx^a/d\ell$ along unperturbed geodesics. For unperturbed geodesics traveling along the direction $-\hat{p}$ (so that the unit vector $\hat{p}$ points from Earth to the pulsar), we have $dx^a/d\ell = -\hat{p}^a$. Integrating Eq. (28), we therefore get

$$t_E - t_P = \ell_E - \ell_P + \frac{1}{2} \hat{p}^a \hat{p}^b \int_{t_P}^{t_E} dt h_{ab}(t, \vec{x}(t)), \quad (29)$$

where we have substituted $d\ell$ by dt in the integral, as they are equal to zeroth order in h_{ab}, and $\vec{x}(t)$ is the spatial position along the geodesic. Now, in this gauge the pulsar and Earth (seen as test particles) stay at the same spatial coordinates [34]. This implies (i) $\ell_E - \ell_P$ takes the same value with and without GWs, and (ii) the proper time measured at Earth is also the coordinate time t. Therefore the last term in Eq. (29) is precisely the sought-after GW-induced timing residual R^{GW}_p. Assuming the Earth is at the origin of spatial coordinates, we have

$$R^{\mathrm{GW}}_p(t) = \frac{1}{2} \hat{p}^a \hat{p}^b \int_{t-D_p}^{t} dt' h_{ab}(t', (t - t')\hat{p}), \quad (30)$$

where D_p is the distance (or time) between Earth and the pulsar.

It is useful to recast this result in terms of the Fourier transform of the strain. Inserting Eq. (1) into the time-residual (30), we obtain

$$R^{\mathrm{GW}}_p(t) = \frac{1}{2} \hat{p}^a \hat{p}^b \int df \int d^2\hat{\Omega}\, h_{ab}(f, \hat{\Omega}) \times \int_{t-D_p}^{t} dt' e^{2\pi i f (t' - \hat{p}\cdot\hat{\Omega}(t-t'))} \equiv \int df e^{2\pi i f t} R^{\mathrm{GW}}_p(f). \quad (31)$$

Upon performing the time integral, we find the Fourier transform of the GW-induced time residual $R^{\mathrm{GW}}_p(f)$:

$$R^{\mathrm{GW}}_p(f) = \frac{\hat{p}^a \hat{p}^b}{4\pi i f} \int d^2\hat{\Omega} \frac{h_{ab}(f, \hat{\Omega})}{(1 + \hat{\Omega}\cdot\hat{p})} \times (1 - e^{-2\pi i f D_p (1 + \hat{p}\cdot\hat{\Omega})}). \quad (32)$$

The first term in the parentheses corresponds to the "Earth term" and the second term to the "pulsar term."

D. Time-residual correlations

We define the (one-sided) cross-power spectrum $\mathcal{R}^{\mathrm{GW}}_{pq}(f)$ of the GW-induced time residuals at different pulsars p, q as follows:

$$\langle R^{\mathrm{GW}}_p(f) R^{*\mathrm{GW}}_q(f') \rangle = \frac{\delta_{\mathrm{D}}(f' - f)}{2} \mathcal{R}^{\mathrm{GW}}_{pq}(f). \quad (33)$$

Using Eq. (2), we find

$$\mathcal{R}^{\rm GW}_{pq}(f) = \frac{1}{(4\pi f)^2}\int \frac{d^2\hat{\Omega}}{4\pi}\frac{\hat{p}^a\hat{p}^b\hat{q}^c\hat{q}^d\mathcal{P}_{abcd}(f,\hat{\Omega})}{(1+\hat{p}\cdot\hat{\Omega})(1+\hat{q}\cdot\hat{\Omega})} \times (1-e^{-2\pi i f D_p(1+\hat{p}\cdot\hat{\Omega})})(1-e^{2\pi i f D_q(1+\hat{q}\cdot\hat{\Omega})}). \quad (34)$$

We can think of the pulsar-term contributions as taking the harmonic transform of the integrand at multipole $\ell \sim 2\pi f D$ (note that the numerator vanishes as $\hat{\Omega}\to-\hat{p}$ and $\hat{\Omega}\to-\hat{q}$ so the integrand is well behaved there). In practice, we have $D\sim$ kpc $\sim 3\times10^3$ light-years and $f\sim 1/\mathrm{yr}$; thus

$$2\pi f D \approx 2\times10^4 \frac{D}{\rm kpc}\frac{f}{\rm yr^{-1}}. \quad (35)$$

Therefore, as long as angular fluctuations of the SGWB on a scale $\ell \gtrsim 10^4$ are negligible, we may safely approximate the terms in parentheses by

$$(1-e^{-2\pi i f D_p(1+\hat{p}\cdot\hat{\Omega})})(1-e^{2\pi i f D_q(1+\hat{q}\cdot\hat{\Omega})}) \to (1+\delta_{pq}),$$

where the Kronecker delta accounts for the factor of 2 if the two pulsars are *identical*, i.e., have the same location on the sky *and* are at the same distance. See Ref. [35] for an explicit proof of the validity of this approximation for an isotropic SGWB and [36] for an anisotropic one.

It will be useful in what follows to introduce some compact notation to denote integrals over the sky. For any two functions of sky location (which from here on we refer to as "maps" and represent by bolded symbols throughout) $\boldsymbol{\mathcal{M}}_1(\hat{\Omega})$ and $\boldsymbol{\mathcal{M}}_2(\hat{\Omega})$, we denote for short the scalar product

$$\boldsymbol{\mathcal{M}}_1\cdot\boldsymbol{\mathcal{M}}_2 \equiv \int\frac{d^2\hat{\Omega}}{4\pi}\boldsymbol{\mathcal{M}}_1(\hat{\Omega})\boldsymbol{\mathcal{M}}_2(\hat{\Omega}). \quad (36)$$

Specializing to the total-intensity piece of the SGWB power spectrum, Eq. (34) then becomes

$$\mathcal{R}^{\rm GW}_{pq}(f) = \frac{1+\delta_{pq}}{(4\pi f)^2}\int\frac{d^2\hat{\Omega}}{4\pi}\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega})\mathcal{I}(f,\hat{\Omega}) = \frac{1+\delta_{pq}}{(4\pi f)^2}\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega})\cdot\boldsymbol{\mathcal{I}}(f), \quad (37)$$

where we have defined the geometric quantity

$$\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega}) \equiv \frac{\hat{p}^a\hat{p}^b\hat{q}^c\hat{q}^d\mathfrak{T}_{abcd}(\hat{\Omega})}{(1+\hat{p}\cdot\hat{\Omega})(1+\hat{q}\cdot\hat{\Omega})}, \quad (38)$$

which can be written explicitly in terms of dot products as follows:

$$\boxed{\gamma_{\hat{p}\hat{q}}(\hat{\Omega}) = 2\frac{(\hat{p}\cdot\hat{q}-(\hat{p}\cdot\hat{\Omega})(\hat{q}\cdot\hat{\Omega}))^2}{(1+\hat{p}\cdot\hat{\Omega})(1+\hat{q}\cdot\hat{\Omega})} - (1-\hat{p}\cdot\hat{\Omega})(1-\hat{q}\cdot\hat{\Omega}).} \quad (39)$$

In what follows, we shall refer to $\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega})$ as the *pairwise timing response function*. It can be expressed in the standard, frame-dependent, notation in terms of the so-called antenna beam patterns $F^+_{\hat{p}}$, $F^\times_{\hat{p}}$:

$$\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega}) = 4\sum_{A=+,\times}F^A_{\hat{p}}(\hat{\Omega})F^A_{\hat{q}}(\hat{\Omega}), \quad (40)$$

$$F^A_{\hat{p}}(\hat{\Omega}) \equiv \frac{1}{2}\frac{\hat{p}^a\hat{p}^b}{1+\hat{p}\cdot\hat{\Omega}}e^A_{ab}(\hat{\Omega}), \quad (41)$$

as can be seen from Eqs. (17) and (38). Our geometric approach allowed us to obtain the explicit and clearly frame-independent expression for this function, Eq. II D. The so-called overlap reduction function is then obtained by integrating the angular dependence of the SGWB intensity multiplied by the pairwise timing response function.

The kernel $\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega})$ is symmetric in $\hat{p}$, $\hat{q}$, and so $\mathcal{R}_{pq}(f)$ is symmetric in p, q. We show plots of $\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega})$ in Fig. 1 as a function of the separation between pulsars $\hat{p}$, $\hat{q}$ (see also Ref. [37] for similar plots).

As an aside, it is interesting to consider the response function for a single pulsar ($\hat{p}=\hat{q}$):

$$\boldsymbol{\gamma}_{\hat{p}\hat{p}}(\hat{\Omega}) = (1-\hat{p}\cdot\hat{\Omega})^2 = \frac{4}{3}-2\hat{p}\cdot\hat{\Omega}+\hat{p}^a\hat{p}^b\left(\hat{\Omega}_a\hat{\Omega}_b-\frac{1}{3}\delta_{ab}\right). \quad (42)$$

Therefore a single pulsar is sensitive to a specific linear combination of the SGWB monopole, dipole projected onto $\hat{p}$, and quadrupole twice projected onto $\hat{p}$. Single-pulsar upper limits typically assume an isotropic background. These limits would be weakened if accounting for anisotropies (see, e.g., models 1, 2A–2D considered in Ref. [10]).

E. Response to an isotropic SGWB

Let us now compute the response to an isotropic SGWB. We define the unit-norm monopole map

$$\mathbf{1}(\hat{\Omega}) \equiv 1, \quad \forall\ \hat{\Omega}. \quad (43)$$

If the SGWB intensity is isotropic, $\boldsymbol{\mathcal{I}}(f,\hat{\Omega}) \propto \mathbf{1}(\hat{\Omega})$. In that case the time-residual power spectrum $\mathcal{R}_{pq}$ only depends on the scalar product $\mu\equiv\hat{p}\cdot\hat{q}$, with the well-known Hellings and Downs functional dependence [3]. We define

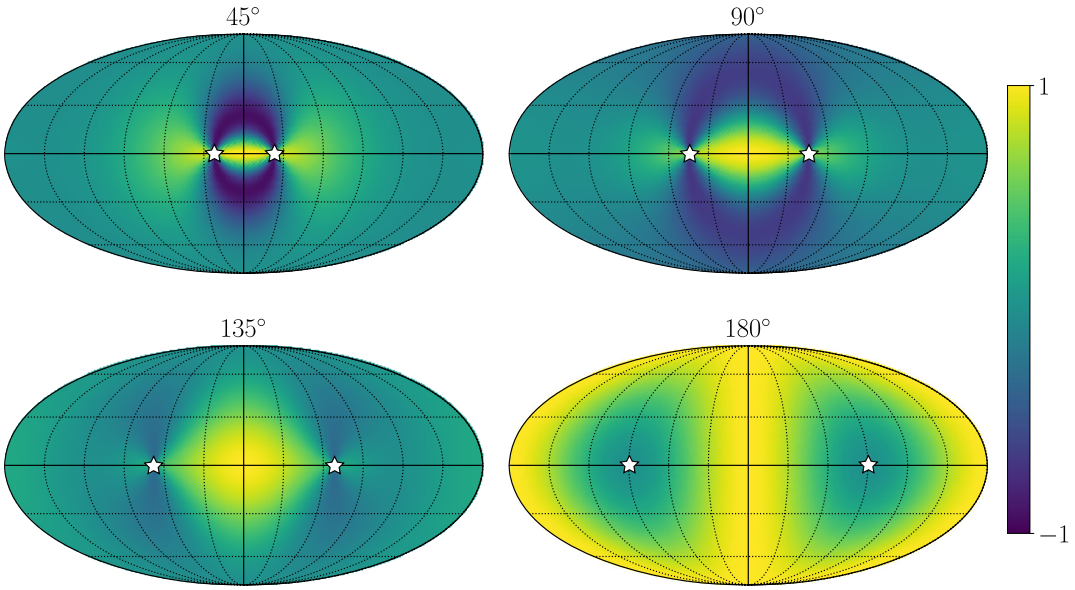

FIG. 1. Pairwise timing response function $\gamma_{\hat{p}\hat{q}}(\hat{\Omega})$, as a function of sky location $\hat{\Omega}$, for four different pulsar separations. The pulsars locations $\hat{p}$ and $\hat{q}$ are shown as stars, and the angle between them is indicated on the top of each figure.

$$\mathcal{H}_{\hat{p}\hat{q}} \equiv \mathbf{1}\cdot\gamma_{\hat{p}\hat{q}} = \int \frac{d^2\hat{\Omega}}{4\pi}\gamma_{\hat{p}\hat{q}}(\hat{\Omega}) = \mathcal{H}(\hat{p}\cdot\hat{q}). \tag{44}$$

We provide a new and concise derivation of this function in Appendix A, making use of our geometric, frame-invariant formalism. With our normalization convention, it is given by

$$\mathcal{H}(\mu) = \frac{3+\mu}{3} + 2(1-\mu)\ln\left(\frac{1-\mu}{2}\right). \tag{45}$$

III. PTA FISHER MATRIX FOR THE GWB INTENSITY

A. Motivations and general considerations

The analyses of real PTA data are typically built on a Bayesian framework and deal directly with the times of arrival (TOAs) of pulsar pulses [8,10]. The final product of such analyses is to estimate how likely a GWB signal $\mathcal{I}(f,\hat{\Omega})$ is given the data. If the data sample is sufficiently large, the likelihood $\mathcal{L}$ of the GWB intensity $\mathcal{I}(f,\hat{\Omega})$ ought to be approximately Gaussian (see, e.g., Fig. 2 of [38]), i.e., *formally* of the form

$$-2\ln\mathcal{L}(\mathcal{I}) = \text{const} + \iint df\,df'[\mathcal{I}(f)-\mathcal{I}^{\rm ml}(f)]\cdot\mathcal{G}_{ff'}\cdot[\mathcal{I}(f')-\mathcal{I}^{\rm ml}(f')], \tag{46}$$

where $\mathcal{G}_{ff'}(\hat{\Omega},\hat{\Omega}')$ is a generalized inverse-covariance "matrix" and $\mathcal{I}^{\rm ml}(f,\hat{\Omega})$ is the maximum-likelihood SGWB intensity. In full generality, $\mathcal{G}_{ff'}$ itself ought to depend on the SGWB intensity (so that the likelihood is not actually Gaussian); nevertheless, we expect that this dependence should only be important once the GWB is *detected* to sufficient significance, as we will quantify shortly. Until then, a *weak-signal* Fisher matrix is sufficiently accurate.

Our goal here is to provide an *approximate* Fisher matrix that can be used as a guide to data analysis. This bears similarities with the study of cosmic microwave background (CMB) anisotropies (see, e.g., Ref. [39]): while the full analysis of CMB data uses a Bayesian framework and deals with the temperature and polarization maps directly, the simple Fisher matrix of CMB *power spectra* allows one to make quick and rather accurate detectability forecasts, which serve to inform full data analyses.

B. Approximate Fisher matrix of band-integrated GWB intensity

In addition to the stochastic timing residual caused by a SGWB, arrival times are noisy, due to intrinsic pulsar noise and instrumental noise:

$$R_p = R_p^{\rm GW} + N_p, \tag{47}$$

where N_p is the (non-GWB-sourced) timing noise, which we assume to be uncorrelated between pulsars,[3] and whose power spectrum is $\sigma_p^2(f)$:

[3]A more realistic analysis includes several additional sources of *common* noise, correlated among pulsars, such as global clock errors or ephemeris errors [10]. These additional noise sources do not appear to significantly affect current upper limits on the amplitude of the SGWB [10], and we do not include them here. We leave to future work a more detailed treatment including these common noise sources within our Fisher framework.

$$\langle N_p(f)N_q^*(f')\rangle = \delta_{pq}\frac{\delta_{\rm D}(f'-f)}{2}\sigma_p^2(f). \tag{48}$$

The standard pulsar analysis fits for several different pulsar-specific sources of noise (e.g., [38,40]).

In the remainder of this paper, we will work with *band-integrated* quantities: given a frequency bandwidth Δf, we define the dimensionless band-integrated SGWB intensity

$$\mathcal{I}_f(\hat{\Omega}) \equiv \int_{f-\Delta f/2}^{f+\Delta f/2} df' \mathcal{I}(f',\hat{\Omega}), \tag{49}$$

and the band-integrated noise (with dimensions of time squared)

$$\sigma_{p,f}^2 \equiv \int_{f-\Delta f/2}^{f+\Delta f/2} df' \sigma_p^2(f'). \tag{50}$$

We denote by $\mathcal{R}_{pq}(f) = \mathcal{R}_{pq}^{\rm GW}(f) + \delta_{pq}\sigma_p^2(f)$ the total timing–residual cross power spectrum, and by $\mathcal{R}_{pq,f}$ the timing–residual cross-band-powers (with dimensions of time squared), given by

$$\begin{aligned}\mathcal{R}_{pq,f} &= \mathcal{R}_{pq,f}^{\rm GW} + \delta_{pq}\sigma_{p,f}^2 \\ &= \frac{1+\delta_{pq}}{(4\pi f)^2}\boldsymbol{\gamma}_{pq}\cdot\boldsymbol{\mathcal{I}}_f + \delta_{pq}\sigma_{p,f}^2.\end{aligned} \tag{51}$$

In what follows, and unless explicitly specified, we always work with band-integrated quantities centered at frequency f. To keep the notation manageable, we drop the subscripts f on all band powers.

We label *unique* pairs of *distinct* pulsars by capital indices I, J, K. For instance, $I = (p,q) = (q,p)$ represents a unique pair of distinct pulsars $p \neq q$. For $N_{\rm psr}$ pulsars, there are $N_{\rm pair} = N_{\rm psr}(N_{\rm psr}-1)/2$ such distinct pairs. For a pair of distinct pulsars I, assuming the SGWB is the only source of correlated noise between distinct pulsars, Eq. (51) simplifies to $\mathcal{R}_I = \boldsymbol{\gamma}_I\cdot\boldsymbol{\mathcal{I}}/(4\pi f)^2$.

Let us denote by $\hat{\mathcal{R}}_I$ unbiased estimators of the band powers. Let us assume that these estimators are constructed from a large number of effectively uncorrelated samples, implying that they are approximately Gaussian distributed. Their statistics are thus entirely determined by their $N_{\rm pair}\times N_{\rm pair}$ covariance matrix $\mathcal{C}$, with elements $\mathcal{C}_{IJ}$ [with dimensions of $(\text{time})^4$]. Note that this matrix depends on frequency f. Under the Gaussian approximation, the joint probability distribution $\mathcal{L}$ of the estimators $\hat{\mathcal{R}}_I$ is therefore

$$\begin{aligned}-2\ln\mathcal{L} &= \text{const} \\ &+ \sum_{I,J}\left(\hat{\mathcal{R}}_I - \frac{\boldsymbol{\gamma}_I\cdot\boldsymbol{\mathcal{I}}}{(4\pi f)^2}\right)(\mathcal{C}^{-1})_{IJ}\left(\hat{\mathcal{R}}_J - \frac{\boldsymbol{\gamma}_J\cdot\boldsymbol{\mathcal{I}}}{(4\pi f)^2}\right).\end{aligned} \tag{52}$$

As is standard in Bayesian data analysis, we view this probability distribution as the likelihood of the signal—the GWB background bandpower $\boldsymbol{\mathcal{I}}(\hat{\Omega})$—given the data. To be precise, this statement assumes a uniform prior on the amplitude of $\boldsymbol{\mathcal{I}}(\hat{\Omega})$.

In order to write an estimator for the SGWB intensity $\hat{\boldsymbol{\mathcal{I}}}$, we define the *dual maps* $\boldsymbol{\gamma}_I^*(\hat{\Omega})$ (not to be mistaken with complex conjugates), which are the unique linear combinations of the $\boldsymbol{\gamma}_I(\hat{\Omega})$ satisfying

$$\boldsymbol{\gamma}_I^*\cdot\boldsymbol{\gamma}_J = \delta_{IJ}. \tag{53}$$

We then define

$$\hat{\boldsymbol{\mathcal{I}}}(\hat{\Omega}) \equiv (4\pi f)^2\sum_K \hat{\mathcal{R}}_K\boldsymbol{\gamma}_K^*(\hat{\Omega}), \tag{54}$$

which satisfies $(\boldsymbol{\gamma}_I\cdot\hat{\boldsymbol{\mathcal{I}}})/(4\pi f)^2 = \hat{\mathcal{R}}_I$. We are now finally in the position of defining the Fisher matrix for the band powers,

$$\boxed{\boldsymbol{\mathcal{F}}_f(\hat{\Omega},\hat{\Omega}') \equiv \frac{1}{(4\pi f)^4}\sum_{I,J}\boldsymbol{\gamma}_I(\hat{\Omega})(\mathcal{C}^{-1})_{IJ}\boldsymbol{\gamma}_J(\hat{\Omega}')}. \tag{55}$$

With these definitions, we see that the likelihood for timing residual band powers can be rewritten as

$$\mathcal{L} \propto \exp\left\{-\frac{1}{2}\sum_{\text{band}(f)}[\boldsymbol{\mathcal{I}}_f - \hat{\boldsymbol{\mathcal{I}}}_f]\cdot\boldsymbol{\mathcal{F}}_f\cdot[\boldsymbol{\mathcal{I}}_f - \hat{\boldsymbol{\mathcal{I}}}_f]\right\}. \tag{56}$$

It might appear at first sight that Eq. (56) is a probability distribution on the inifinite-dimensional space of maps $\boldsymbol{\mathcal{I}}(\hat{\Omega})$. However, the $N_{\rm pair}$ pairwise-time-residual correlations $\mathcal{R}_I$ can only possibly measure $N_{\rm pair}$ projections of the SGWB map. To see what these are precisely, decompose $\boldsymbol{\mathcal{I}}(\hat{\Omega})$ onto a piece which is a linear combination of the functions $\boldsymbol{\gamma}_I(\hat{\Omega})$—hence of the $\boldsymbol{\gamma}_I^*(\hat{\Omega})$—and a piece which is orthogonal to all of them:

$$\begin{aligned}\boldsymbol{\mathcal{I}}(\hat{\Omega}) &= \boldsymbol{\mathcal{I}}_{||}(\hat{\Omega}) + \boldsymbol{\mathcal{I}}_{\perp}(\hat{\Omega}), \\ \boldsymbol{\mathcal{I}}_{||}(\hat{\Omega}) &\equiv (4\pi f)^2\sum_K \mathcal{R}_K\boldsymbol{\gamma}_K^*(\hat{\Omega}), \\ \boldsymbol{\mathcal{I}}_{\perp}\cdot\boldsymbol{\gamma}_I &= 0, \quad \forall\ I.\end{aligned} \tag{57}$$

We purposefully denoted by $\mathcal{R}_K$ the coefficient of $\boldsymbol{\gamma}_K^*$, as it is indeed the time-residual correlation measured by the pulsar pair K [see Eq. (51)]. From the expression of $\boldsymbol{\mathcal{F}}$, we see that $\boldsymbol{\mathcal{F}}\cdot\boldsymbol{\mathcal{I}}_{\perp} = 0$. Thus the likelihood function only depends on $\boldsymbol{\mathcal{I}}_{||}$, which spans a $N_{\rm pair}$-dimensional space, and contains no information about $\boldsymbol{\mathcal{I}}_{\perp}$. Put differently, the components of $\boldsymbol{\mathcal{F}}$ orthogonal to the space spanned by the $\boldsymbol{\gamma}_I$'s have *infinite noise*. A consequence of this property is that one cannot hope to simultaneously constrain more than

$N_{\rm pair}$ statistically independent components of the SGWB map—be they harmonic coefficients, independent pixels, or any other linear projections.

C. Weak-signal limit for the Fisher matrix

In Appendix B, we derive the following approximation of the covariance matrix of the estimators for the pairwise-time-residual band powers $\hat{\mathcal{R}}_{pq}$: for two pairs $I=(p,q)$ and $J=(p',q')$, we have

$$\mathcal{C}_{IJ}=\mathcal{C}_{pq,p'q'}\equiv\mathrm{cov}(\hat{\mathcal{R}}_{pq},\hat{\mathcal{R}}_{p'q'}) \approx\frac{1}{2T_{IJ}\Delta f}(\mathcal{R}_{pp'}\mathcal{R}_{qq'}+\mathcal{R}_{pq'}\mathcal{R}_{qp'}), \tag{58}$$

where T_{IJ} is the *effective* total time of observation of the four pulsars p, q, p', q', which we found to be approximately

$$T_{IJ}\approx\max\left[\min(T_p,T_q),\min(T_{p'},T_{q'})\right], \tag{59}$$

if each pulsar p is observed for a total time T_p. This equation is a generalization of the radiometer equation for electromagnetic intensity [41] and holds provided the bandwidth Δf satisfies

$$1/T\ll\Delta f\ll f, \tag{60}$$

where T is the minimum of all observation times. In particular, it only applies for $f\gg 1/T$.

We now specialize to the *weak-signal limit*, i.e., assume that, for *every pulsar* p (and in particular, for the least noisy pulsar),

$$\mathcal{I}\ll(4\pi f)^2\sigma_p^2. \tag{61}$$

In other words, we assume that the SGWB-induced signal is subdominant to the intrinsic pulsar noise in *each individual pulsar*. In this limit, we may approximate $\mathcal{R}_{pq}\approx\delta_{pq}\sigma_p^2$ on the right-hand side of Eq. (58). As a result, the weak-signal correlation matrix $\mathcal{C}_{IJ}$ is diagonal, and so is its inverse:

$$\mathcal{C}_{IJ}\approx\frac{\sigma_p^2\sigma_q^2}{2T_{pq}\Delta f}\delta_{IJ},\qquad(\mathcal{C}^{-1})_{IJ}\approx\frac{2T_{pq}\Delta f}{\sigma_p^2\sigma_q^2}\delta_{IJ},$$
$$I=(p,q),\qquad T_{pq}\equiv\min(T_p,T_q). \tag{62}$$

In addition to stochastic contributions discussed thus far, the timing residual R_p contains a *deterministic* piece, resulting from the pulsar's intrinsic motion, spin down, etc. To account for these deterministic contributions, a timing model is fitted to pulsars' times of arrival. This process results in a loss of information, quantified by a "transmission function" $\mathcal{T}_p(f)$ [40]. For our purposes, let us note that for all pulsars $\mathcal{T}_p(f)\simeq 1$ for $f\gtrsim 1/T$ and $\mathcal{T}_p(f)\simeq(fT_p)^6$ for $fT_p\ll 1$ for most pulsars[4] [40]. In addition (and more relevant for us since we only consider the regime $fT_p\gtrsim 1$), the transmission function filters out harmonics of 1/year due to degeneracies of timing-model parameters with the motion of the Earth around the Sun.

Combining Eq. (55) with Eq. (62) and multiplying the contribution of each pair $I=(p,q)$ by $\mathcal{T}_p\mathcal{T}_q$, our final expression for the Fisher matrix for the band-integrated SGWB is therefore

$$\boxed{\boldsymbol{\mathcal{F}}_f(\hat{\Omega},\hat{\Omega}')=\frac{1}{(4\pi f)^4}\sum_{p\neq q}\mathcal{T}_p(f)\mathcal{T}_q(f)\times\frac{2T_{pq}\Delta f}{\sigma_{p,f}^2\sigma_{q,f}^2}\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega})\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega}').} \tag{63}$$

This weak-signal Fisher matrix is the main result of this paper.[5] It allows us to estimate the signal-to-noise ratio (SNR) of the GWB band-integrated intensity $\boldsymbol{\mathcal{I}}_f(\hat{\Omega})$ with an *arbitrary angular dependence*:

$$\mathrm{SNR}^2[\boldsymbol{\mathcal{I}}_f]=\boldsymbol{\mathcal{I}}_f\cdot\boldsymbol{\mathcal{F}}_f\cdot\boldsymbol{\mathcal{I}}_f =\sum_{p\neq q}\mathcal{T}_p(f)\mathcal{T}_q(f)2T_{pq}\Delta f\left[\frac{\boldsymbol{\gamma}_{\hat{p}\hat{q}}\cdot\boldsymbol{\mathcal{I}}_f}{(4\pi f)^2\sigma_{p,f}\sigma_{q,f}}\right]^2. \tag{64}$$

Provided the bandwidth is much wider than the inverse of the observation time for each pulsar, $\Delta f\gg 1/T_p$, different bands are uncorrelated, so that the total SNR^2 is obtained from summing that of each band:

$$\mathrm{SNR}^2[\mathrm{total}] \approx\sum_{\mathrm{band}(f)}\mathrm{SNR}^2[\boldsymbol{\mathcal{I}}_f] \approx 2\sum_{p\neq q}T_{pq}\int_{1/T_{pq}}^{f_{\max}}df\mathcal{T}_p(f)\mathcal{T}_q(f)\left[\frac{\boldsymbol{\gamma}_{\hat{p}\hat{q}}\cdot\boldsymbol{\mathcal{I}}(f)}{(4\pi f)^2\sigma_p(f)\sigma_q(f)}\right]^2, \tag{65}$$

where we replaced the sum over bands by an integral under the assumption that $\Delta f\ll f$. The lower frequency bound is such that $f_{\min}=\max(1/T_p,1/T_q)=1/T_{pq}$, and depends on the pulsar pair. The upper frequency bound is the Nyquist frequency $f_{\max}=\min(1/\Delta t_p,1/\Delta t_q)/2$, inversely proportional to the observation cadence. Given the factor f^{-4} in the integrand, unless the SGWB is significantly blue the total SNR is typically dominated by the lowest frequencies, and the upper cutoff has little impact.

[4]This scaling applies to pulsars with a quadradic spin-down.

[5]Note that the "point-spread function" defined in Ref. [37] is proportional to our Fisher matrix, in the case where all pulsars have identical noise.

Equation (65) generalizes Eq. (17) of Ref. [42] in several ways. First, it accounts for different observation times for each pulsar. Second, it accounts for the loss of information in the timing-model-fitting process, through the transmission functions $\mathcal{T}_p(f)$. Last but not least, it accounts for an arbitrary angular dependence of the SGWB, rather than assuming a monopole.

Before moving on to applications, let us quantify when the weak-signal limit applies. Suppose all pulsars have a typical observation time T and noise σ_f^2. Consider moreover frequencies for which $\mathcal{T}_p(f) \simeq 1$ (note that for our simple covariance matrix to hold, we require $f \gg 1/T$, and thus $\mathcal{T}(f) \simeq 1$ except at harmonics of 1/year). Equation (64) then gives

$$\mathrm{SNR}^2[\boldsymbol{\mathcal{I}}_f(\hat{\Omega})] \simeq N_{\rm pair} 2T\Delta f \left(\frac{\mathcal{I}_f}{(4\pi f)^2\sigma_f^2}\right)^2. \tag{66}$$

The weak-signal approximation (61) requires the last term in parentheses to be less than unity. It is thus self-consistent as long as the band-integrated SGWB is detected with a signal-to-noise ratio $\mathrm{SNR} \lesssim N_{\rm psr}\sqrt{T\Delta f}$ *in each band*. Unless the SGWB is significantly blue, the total SNR is dominated by the lowest frequencies, so that for the weak-signal limit to be appropriate we must have a total (frequency-integrated) $\mathrm{SNR} \lesssim N_{\rm psr}$.

IV. IDEALIZED CASE: ISOTROPICALLY DISTRIBUTED IDENTICAL PULSARS

In this section we apply our results to an idealized PTA consisting of $N_{\rm psr} \gg 1$ identical pulsars approximately isotropically distributed on the sky. This limiting case is amenable to analytic approximations, and it will serve to cross-check our numerical algorithms when we apply our formalism to real PTAs. We moreover compare our results with those of Ref. [21], which apply in this limit.

A. Analytic expression for $N_{\rm psr} \to \infty$

Suppose all the pulsars have the same noise $\sigma_p = \sigma$, are observed for the same time T, and have the same transmission function $\mathcal{T}(f)$. In that limit the Fisher matrix $\boldsymbol{\mathcal{F}}$ is given by

$$\boldsymbol{\mathcal{F}}(\hat{\Omega},\hat{\Omega}') = C\boldsymbol{F}(\hat{\Omega},\hat{\Omega}'), \tag{67}$$

$$C \equiv \frac{\mathcal{T}(f)^2}{(4\pi f)^4}\frac{2T\Delta f}{\sigma^4}N_{\rm pair}, \tag{68}$$

$$\boldsymbol{F}(\hat{\Omega},\hat{\Omega}') \equiv \frac{1}{N_{\rm pair}}\sum_I \boldsymbol{\gamma}_I(\hat{\Omega})\boldsymbol{\gamma}_I(\hat{\Omega}'). \tag{69}$$

In the limit that $N_{\rm psr} \to \infty$, assuming the pulsars are isotropically distributed, we find

$$\boldsymbol{F}(\hat{\Omega},\hat{\Omega}') \underset{N_{\rm psr}\to\infty}{\longrightarrow} \mathcal{F}_\infty(\hat{\Omega}\cdot\hat{\Omega}') \equiv \int \frac{d^2\hat{p}}{4\pi}\frac{d^2\hat{q}}{4\pi}\boldsymbol{\gamma}_{\hat{p}\,\hat{q}}(\hat{\Omega})\boldsymbol{\gamma}_{\hat{p}\,\hat{q}}(\hat{\Omega}'). \tag{70}$$

By symmetry, this is a function of $\chi \equiv \hat{\Omega}\cdot\hat{\Omega}'$ only, which we compute explicitly in Appendix C. We derive the following analytic expression:

$$\mathcal{F}_\infty(\chi) = \frac{16}{9(1+\chi)^2} \times \left[\left(\frac{1-\chi^2}{4} + 2 - \chi + 3\frac{1-\chi}{1+\chi}\log\frac{1-\chi}{2}\right)^2 + \left(2-\chi+3\frac{1-\chi}{1+\chi}\log\frac{1-\chi}{2}\right)^2\right]. \tag{71}$$

We show $\mathcal{F}_\infty(\chi)$ as a solid line in Fig. 2. For comparison, we also show the reduced Fisher matrix $\boldsymbol{F}(\hat{\Omega},\hat{\Omega}')$ for a finite number of identical, quasi-isotropically distributed pulsars,[6] for 1000 randomly selected pairs of sky directions $(\hat{\Omega},\hat{\Omega}')$. Only in the limit $N_{\rm psr} \to \infty$ is the Fisher matrix a function of the angle $\angle(\hat{\Omega},\hat{\Omega}')$ only; otherwise, it depends on both $\hat{\Omega}$ and $\hat{\Omega}'$, which translates into a scatter of the values of $F(\hat{\Omega},\hat{\Omega}')$ when plotted as a function of $\angle(\hat{\Omega},\hat{\Omega}')$. We see that $F(\hat{\Omega},\hat{\Omega}')$ indeed converges to the function $\mathcal{F}_\infty$ as $N_{\rm psr}$ increases, with a difference (both in running mean and scatter) scaling as $\sim 1/N_{\rm psr}$.

The dense-PTA Fisher matrix can be decomposed into Legendre polynomials:

$$\mathcal{F}_\infty(\hat{\Omega}\cdot\hat{\Omega}') = \sum_\ell (2\ell+1)\mathcal{F}_\ell P_\ell(\hat{\Omega}\cdot\hat{\Omega}') = 4\pi\sum_{\ell,m}\mathcal{F}_\ell \mathcal{Y}_{\ell m}(\hat{\Omega})\mathcal{Y}_{\ell m}(\hat{\Omega}'), \tag{72}$$

where the $\mathcal{Y}_{\ell m}$ are the real spherical harmonics.

We show the Legendre coefficients $\mathcal{F}_\ell$ in Fig. 3. Interestingly, the amplitude of Legendre coefficients decreases monotonically with ℓ, *except* for $\mathcal{F}_1 \approx \mathcal{F}_0/7$, which is significantly lower than $\mathcal{F}_2$ and comparable to $\mathcal{F}_3$.

B. Minimum detectable dipolar anisotropy

Suppose the GWB takes the form

$$\boldsymbol{\mathcal{I}} = \mathcal{I}_0\left(\mathbf{1} + \sum_{\ell\geq 1,m} g_{\ell m}\boldsymbol{\mathcal{Y}}_{\ell m}\right). \tag{73}$$

From Eq. (72), we see that in the limit of a dense array of identical pulsars, the spherical harmonic coefficients of the

[6]To place pulsars quasi-isotropically we arrange them in equal intervals in the azimuthal angle and with the polar angle $\theta = \cos^{-1}(\mathcal{U})$, where $\mathcal{U}$ are a set of uniformly spaced numbers in $[-1, 1]$.

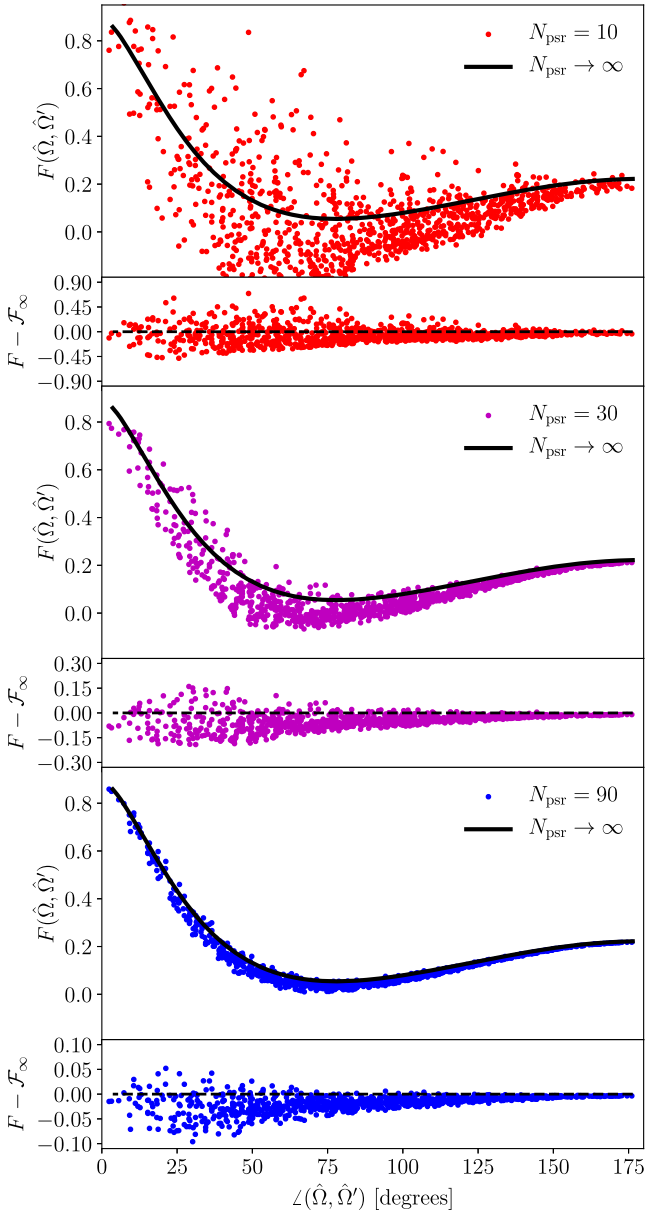

FIG. 2. Values of the rescaled Fisher matrix for a finite number of quasi-isotropically distributed identical pulsars (10, 30, and 90, respectively), for 1000 randomly selected pairs of SGWB directions in the sky $(\hat{\Omega}, \hat{\Omega}')$, as a function of the angle between them. The solid black line shows our analytic result, holding for an infinite number of isotropically distributed identical pulsars. We also show the difference between F and its infinite-pulsar limit, $\mathcal{F}_\infty$. We see that the difference decreases as $|F - \mathcal{F}_\infty| \sim 1/N_{\rm psr}$ (note the different y-axis scales in the difference plots).

SGWB are uncorrelated, with noise proportional to $1/\mathcal{F}_\ell$. Explicitly, the signal-to-noise ratio of the coefficients $g_{\ell m}$ is such that

$$\frac{\mathrm{SNR}^2[\mathcal{I}_0 g_{\ell m} \boldsymbol{\mathcal{Y}}_{\ell m}]}{\mathrm{SNR}^2[\mathcal{I}_0 \mathbf{1}]} = \frac{g_{\ell m}^2}{4\pi} \frac{\mathcal{F}_\ell}{\mathcal{F}_0}, \tag{74}$$

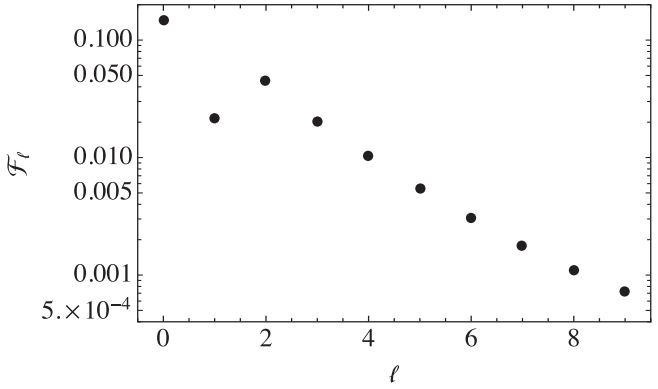

FIG. 3. First few Legendre coefficients of the dense-PTA Fisher matrix.

where we used the fact that $\mathbf{1} = \sqrt{4\pi}\boldsymbol{\mathcal{Y}}_{00}$. In particular, in order to detect the ℓmth harmonic with SNR ≥ 3 requires an amplitude

$$g_{\ell m} \geq g_{\ell m,\min} \equiv \sqrt{\frac{\mathcal{F}_0}{\mathcal{F}_\ell}} \frac{6\sqrt{\pi}}{\mathrm{SNR}_0}, \tag{75}$$

where $\mathrm{SNR}_0 \equiv \mathrm{SNR}[\mathcal{I}_0 \mathbf{1}]$ is the signal-to-noise ratio of the monopole. For the dipole, we find

$$g_{1m,\min} \approx \frac{28}{\mathrm{SNR}_0}, \tag{76}$$

which is identical to the result of Ref. [21] in the weak-signal limit.

C. Hot spot in the SGWB

Now consider a SGWB with a hot spot[7] in a *known* direction $\hat{\Omega}_0$. Such a hot spot could be generated, for instance, by a concentration of supermassive black hole binaries, sufficiently numerous that the GW background can still be approximated as *stochastic*. Specifically, we assume

$$\boldsymbol{\mathcal{I}}(\hat{\Omega}) = \mathcal{I}_0 + \mathcal{I}_0 g(4\pi\delta_{\rm D}(\hat{\Omega}; \hat{\Omega}_0) - 1), \tag{77}$$

where we chose the normalization such that the fraction of GW energy density (proportional to the SGWB intensity) in the hot spot is g. With this convention, a physical SGWB ought to have $g \leq 1$.

The joint probability distribution of the monopole and hot spot amplitudes can be obtained from Eq. (56), and is a two-dimensional uncorrelated Gaussian distribution:

[7] A GW "beam" in the nomenclature of Ref. [21].

$$\mathcal{L}(\mathcal{I}_0, \mathcal{I}_0 g) \propto \exp\left[-2\pi C(\mathcal{I}_0^2 \mathcal{F}_0 + (\mathcal{I}_0 g)^2(\mathcal{F}_\infty(1) - \mathcal{F}_0))\right] = \exp\left[-\frac{8\pi}{27} C(\mathcal{I}_0^2 + 5(\mathcal{I}_0 g)^2)\right], \tag{78}$$

where the coefficient C is given in Eq. (68), and in the second line we used $\mathcal{F}_0 = 4/27$ and $\mathcal{F}_\infty(1) = 8/9$. The variances of the monopole and hot spot amplitudes are thus given by

$$\mathrm{var}[\mathcal{I}_0] = 5\mathrm{var}[\mathcal{I}_0 g] = \frac{27}{16\pi} C^{-1}. \tag{79}$$

Hence, for the hot spot to be detectable at the $3-\sigma$ level, its amplitude needs to be

$$g \geq g_{\min} = 3\frac{\sqrt{\mathrm{var}[\mathcal{I}_0 g]}}{\mathcal{I}_0} = \frac{3/\sqrt{5}}{\mathrm{SNR}_0}, \tag{80}$$

where $\mathrm{SNR}_0 = \mathcal{I}_0/\sqrt{\mathrm{var}[\mathcal{I}_0]}$ is again the signal-to-noise ratio of the monopole amplitude. This estimate is in agreement with the numerical result of Ref. [21] in the weak-signal limit. We thus conclude that, provided with the knowledge of the direction of the hot spot, an idealized PTA would be able to detect a hot spot with amplitude $g \simeq 1$ shortly after the monopole is detected. Without any prior information on the hot spot's direction, of course, this conclusion does not hold.

D. Eigenmaps

From Eq. (72), we can see that the eigenmaps of the dense-PTA Fisher matrix are the real spherical harmonics. As one can expect, and as we shall see in greater detail in Paper II, the real spherical harmonics are no longer the eigenmaps of realistic PTAs, and therefore do not provide a particularly well adapted basis for searches for anisotropies. To illustrate this, we diagonalize the reduced Fisher matrix $\boldsymbol{F}(\hat{\Omega}, \hat{\Omega}')$ of an idealized array of a *finite number* of identical, quasi-isotropically distributed pulsars. Specifically, we seek unit-norm maps $\boldsymbol{\mathcal{M}}_n(\hat{\Omega})$ such that

$$\boldsymbol{F} \cdot \boldsymbol{\mathcal{M}}_n = \frac{1}{\Sigma_n^2} \boldsymbol{\mathcal{M}}_n. \tag{81}$$

This continuous eigenvalue problem can be transformed into a regular, discrete, eigenvalue problem by seeking $\boldsymbol{\mathcal{M}}_n$ as a linear combination of the $\boldsymbol{\gamma}_I$:

$$\boldsymbol{\mathcal{M}}_n(\hat{\Omega}) = \sum_I \mathcal{M}_n^I \boldsymbol{\gamma}_I(\hat{\Omega}). \tag{82}$$

The eigenvalue problem (81) is then equivalent to the discrete $N_{\mathrm{pair}} \times N_{\mathrm{pair}}$ eigenvalue problem

$$\sum_J F_{IJ} \mathcal{M}_n^J = \frac{1}{\Sigma_n^2} \mathcal{M}_n^I, \tag{83}$$

$$F_{IJ} \equiv \frac{\boldsymbol{\gamma}_I \cdot \boldsymbol{\gamma}_J}{N_{\mathrm{pair}}}. \tag{84}$$

We thus see that there are exactly N_{pair} principal maps. They do not form a complete set of all possible maps. However, they are a complete set of *observable* maps for a given PTA. Note that the eigenmaps that we derive here are *scalar* maps, corresponding to the intensity of a *stochastic* GW background; this is to be contrasted with the strain eigenmaps derived in Ref. [24], which apply to *continuous* (i.e., deterministic) GW searches. There does not appear to be a straightforward connection between our N_{pair} SGWB intensity eigenmaps and the $2N_{\mathrm{psr}}$ strain eigenmaps of Ref. [24].

We show the first 50 eigenvalues Fig. 4 for $N_{\mathrm{psr}} = 10$, 30, 90, where we compare them against the dense-pulsar limit $N_{\mathrm{psr}} \to \infty$. We see that, as N_{psr} increases, the eigenvalues do converge toward the dense pulsar limit. For $N_{\mathrm{psr}} = 90$, one recognizes the sequences of quasidegenerate eigenvalues, corresponding to the degenerate harmonics for $N_{\mathrm{psr}} \to \infty$. For lower N_{psr}, as the Fisher matrix departs further from its $N_{\mathrm{psr}} \to \infty$ limit, eigenmaps "mix" and are no longer grouped in subsets with similar eigenvalues. This is very similar to the breaking of degeneracy in atomic levels in the presence of a perturbed Hamiltonian. We show the first five eigenmaps in Fig. 5, as a function of N_{psr}. We see that as N_{psr} becomes large, the first eigenmap approaches the monopole, and the next two become quadrupolar. For $N_{\mathrm{psr}} = 10$, however, the eigenmaps do not at all resemble spherical harmonics. More importantly, as we shall see in Paper II, for realistic pulsar distributions, there exist anisotropies to

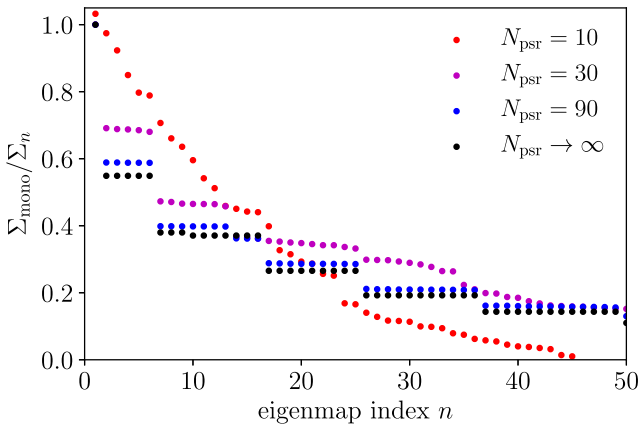

FIG. 4. First 50 eigenvalues of quasi-isotropically distributed identical pulsars compared against the dense-pulsar limit $N_{\mathrm{psr}} \to \infty$. The sequences of equal-noise black dots correspond to multipoles $\ell = 0$, 2, 1, 3, 4, 5, in that order. Having a finite number of pulsars perturbs the eigenmaps away from spherical harmonics and breaks the degeneracies in their eigenvalues.

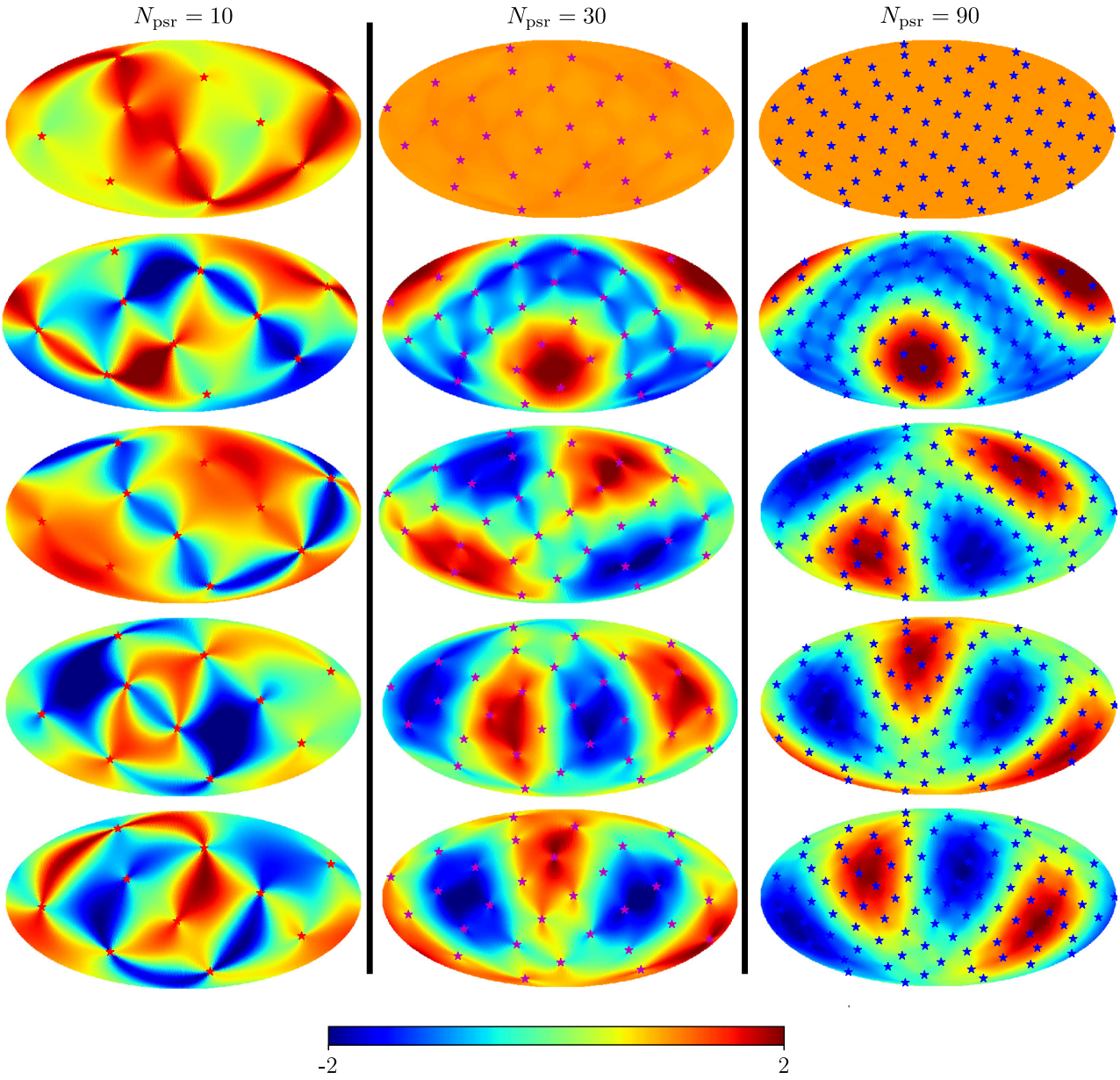

FIG. 5. First five eigenmaps of the Fisher matrix for $N_{\rm psr} = 10$ (left column), 30 (middle column), and 90 (right column) identical, quasi-isotropically distributed pulsars. As $N_{\rm psr}$ is increased, the first eigenmap approaches the monopole, and the next few eigenmaps become more and more quadrupolar. For $N_{\rm psr} = 10$, the eigenmaps do not resemble spherical harmonics at all. The stars indicate the location of the identical pulsars.

which a PTA is much more sensitive than the lowest-order spherical harmonics.

V. CONCLUSIONS

We have derived a band-integrated Fisher matrix for the intensity of a weak, anisotropic SGWB measured by a PTA, Eq. (III C). This Fisher matrix provides a versatile tool with which we can better study the detectability of anisotropies in the SGWB by PTAs. We derived a simple expression of the SNR of an anisotropic SGWB, Eq. (65), generalizing previous results. We moreover derived an exact analytic expression for the Fisher matrix of an idealized PTA consisting of a dense and isotropic distribution of pulsars on the sky. With this matrix, we could recover the results of Ref. [21] for the detectability of dipolar and hot-spot anisotropies. We illustrated how our formalism is better adapted to realistic PTAs by quantifying the convergence of the Fisher matrix of a finite number of pulsars to that of the dense-pulsar limit. In particular, we showed that, for a finite number of pulsar pairs, the eigenmaps of the Fisher matrix are *not* spherical harmonics, commonly used to

study SGWB anisotropies. These $N_{\rm pair}$ eigenmaps best characterize the information content of the Fisher matrix. In a follow-up paper, we will further explore the information content of real PTAs, with unevenly distributed pulsars of unequal noise properties.

In order to arrive at our new Fisher formalism, we rederived existing results with a fresh look and presented them in a geometric, coordinate-free form. Let us highlight, in particular, the SGBW power spectrum (a rank-4 tensor) given in Eq. (6), and the pairwise timing response function, Eq. II D, which characterizes the correlated response of a pair of pulsars to a generic SGWB intensity map. While in this paper we focused on the total intensity of the SGWB, we have provided all the ingredients needed to extend our results to a circularly or linearly polarized SGWB. Our work could also be generalized to non-Einsteinian polarizations [43]. Last, our Fisher formalism can easily be made more realistic: it can accommodate other sources of correlated pulsar timing residuals, such as global clock errors, and can be generalized to a nonweak SGWB, by using the full expression for the Fisher matrix, Eq. (55). Some elements of our Fisher formalism may moreover carry over to other gravitational-wave detection techniques (such as space and ground-based laser interferometers).

The strength of the approach outlined in this paper lies in its ability to clearly and concisely describe the information content of GW measurements. A similar approach for measurements of the CMB [25] has allowed accurate, rigorous, and intuitive estimates of the CMB's sensitivity to a variety of effects. At the dawn of GW astronomy, the development of such a tool is both timely and necessary in order to learn as much as we can from the first GW signals that have been and will be measured.

ACKNOWLEDGMENTS

We thank Marc Kamionkowski, Selim Hotinli, Jeffrey Hazboun, David Nice, and Joseph Romano for useful discussions. Y. A. H. acknowledges support from the National Science Foundation through Grant No. 1820861. T. L. S. acknowledges support from NASA (though Grant No. 80NSSC18K0728) and the Research Corporation. The Flatiron Institute is supported by the Simons Foundation.

APPENDIX A: NEW DERIVATION OF THE HELLINGS AND DOWNS FUNCTION

We want to compute the following function of $\mu \equiv \hat{p}\cdot\hat{q}$, with $\hat{p} \neq \hat{q}$:

$$\mathcal{H}(\mu) \equiv \int \frac{d^2\hat{\Omega}}{4\pi} \gamma_{\hat{p}\,\hat{q}}(\hat{\Omega}), \tag{A1}$$

where $\gamma_{\hat{p}\,\hat{q}}(\hat{\Omega})$ is given in Eq. II D. Let us define the new variables $x \equiv \hat{p}\cdot\hat{\Omega}$, $y \equiv \hat{q}\cdot\hat{\Omega}$, so that

$$\gamma_{\hat{p}\,\hat{q}}(\hat{\Omega}) = \frac{2(\mu - xy)^2 - (1-x^2)(1-y^2)}{(1+x)(1+y)}. \tag{A2}$$

The numerator can be rewritten as

$$\begin{aligned} &2(\mu-xy)^2-(1-x^2)(1-y^2) \\ &\quad = 2(x^2+y^2-2\mu xy-1+\mu^2)+(1-x^2)(1-y^2). \end{aligned} \tag{A3}$$

The second part simplifies with the denominator and the integral can readily be computed, so we get

$$\mathcal{H}(\mu) = \mathcal{J}(\mu) + (1+\mu/3), \tag{A4}$$

$$\mathcal{J}(\mu) \equiv 2\int \frac{d^2\hat{\Omega}}{4\pi} \frac{x^2+y^2-2\mu xy-(1-\mu^2)}{(1+x)(1+y)}. \tag{A5}$$

One can show that the coordinates x, y are restricted to the region

$$\mathcal{E}(\mu) = \{(x,y); x^2+y^2-2\mu xy < 1-\mu^2\}. \tag{A6}$$

The boundary of $\mathcal{E}(\mu)$ is an ellipse whose principal axes are at 45 degree angles from the (x,y) coordinate axes, and with semimajor and semiminor axes $\sqrt{1\pm\mu}$. Moreover, we can show that the area element is

$$d^2\hat{\Omega} = \frac{2dxdy}{\sqrt{1-\mu^2-x^2-y^2+2\mu xy}}. \tag{A7}$$

With these new variables, the integral $\mathcal{J}(\mu)$ simplifies to

$$\mathcal{J}(\mu) = -\int_{\mathcal{E}(\mu)} \frac{dxdy}{\pi} \frac{\sqrt{1-\mu^2-x^2-y^2+2\mu xy}}{(1+x)(1+y)}. \tag{A8}$$

For a given $x \in [-1,1]$, $y \in [y_-, y_+]$, where the boundaries are given by

$$y_\pm \equiv \mu x \pm \sqrt{(1-\mu^2)(1-x^2)}. \tag{A9}$$

We therefore rewrite the integral as

$$\mathcal{J}(\mu) = \int_{-1}^{1} dx \frac{\mathcal{K}(x,\mu)}{1+x}, \tag{A10}$$

where the inner integral is

$$\begin{aligned} \mathcal{K}(x,\mu) &\equiv -\frac{1}{\pi}\int_{y_-}^{y_+} dy \frac{\sqrt{(y_+-y)(y-y_-)}}{1+y} \\ &= |x+\mu| - (1+\mu x). \end{aligned} \tag{A11}$$

After performing the simple outer integral, we arrive at

$$\mathcal{J}(\mu) = 2(1-\mu)\ln\left(\frac{1-\mu}{2}\right). \tag{A12}$$

Inserting this result into Eq. (A4), we finally arrive at the Hellings and Downs function, given in Eq. (45).

APPENDIX B: PROBABILITY DISTRIBUTION OF TIMING POWER SPECTRA

In this Appendix we derive a simple estimate of the covariance matrix of the pairwise-time-residual cross-power spectra. This simple estimate is not meant to follow nor replace a realistic data analysis. Yet, it should provide accurate qualitative scalings, and be quantitatively accurate at the factor-of-few level.

1. Continuous sampling case

Let us suppose that we sample the time residuals $R_p(t)$ of each pulsar p *continuously* over some finite time interval $t \in [-T_p/2, T_p/2]$. Given a frequency f, we define

$$\begin{aligned}\tilde{R}_p(f) &\equiv \int_{-T_p/2}^{T_p/2} dt e^{-2\pi i f t} R_p(t) \\ &= T_p \int df_1 R_p(f_1)\mathrm{sinc}(\pi T_p(f_1 - f)).\end{aligned} \tag{B1}$$

The covariance of these quantities is such that

$$\begin{aligned}&\langle \tilde{R}_p(f)\tilde{R}_q^*(f')\rangle \\ &= \frac{1}{2}T_pT_q\int df_1 \mathcal{R}_{pq}(f_1)\mathrm{sinc}(\pi T_p(f_1 - f)) \\ &\quad \times \mathrm{sinc}(\pi T_q(f_1 - f')),\end{aligned} \tag{B2}$$

where $\mathcal{R}_{pq}$ is the total timing residual cross-power spectrum, defined as in Eq. (33). Now, assume $\mathcal{R}_{pq}$ varies on a scale $\delta f \sim f$, and that $T_p f$, $T_q f \gg 1$. Suppose moreover, for definiteness, that $T_p > T_q$. The sinc function with T_p is narrower and can be approximated as

$$T_p\mathrm{sinc}(\pi T_p(f' - f)) \approx \delta_{\mathrm{D}}(f' - f). \tag{B3}$$

We define $T_{pq} \equiv \min(T_p, T_q)$. We then get

$$\langle \tilde{R}_p(f)\tilde{R}_q^*(f')\rangle \approx \frac{T_{pq}}{2}\mathcal{R}_{pq}(f)\mathrm{sinc}(\pi T_{pq}(f' - f)). \tag{B4}$$

Let us now define for $f > 0$

$$\hat{\mathcal{R}}_{pq}(f) \equiv \frac{1}{T_{pq}}(\tilde{R}_p(f)\tilde{R}_q^*(f) + \tilde{R}_q(f)\tilde{R}_p^*(f)). \tag{B5}$$

From the previous result, $\langle \hat{\mathcal{R}}_{pq}(f)\rangle = \mathcal{R}_{pq}(f)$, which means that $\hat{\mathcal{R}}_{pq}$ is an *unbiased estimator* of $\mathcal{R}_{pq}(f)$. Let us now compute its covariance,

$$\begin{aligned}C_{pq,p'q'}(f,f') &\equiv \langle(\hat{\mathcal{R}}_{pq}(f) - \mathcal{R}_{pq}(f))(\hat{\mathcal{R}}_{p'q'}(f') - \mathcal{R}_{p'q'}(f'))\rangle = \langle \hat{\mathcal{R}}_{pq}(f)\hat{\mathcal{R}}_{p'q'}(f') - \mathcal{R}_{pq}(f)\mathcal{R}_{p'q'}(f')\rangle \\ &= \frac{1}{2T_{pq}T_{p'q'}}\{T_{pp'}\mathrm{sinc}(\pi T_{pp'}(f' - f))T_{qq'}\mathrm{sinc}(\pi T_{qq'}(f' - f))\mathcal{R}_{pp'}(f)\mathcal{R}_{qq'}(f) \\ &\quad + T_{pq'}\mathrm{sinc}(\pi T_{pq'}(f' - f))T_{qp'}\mathrm{sinc}(\pi T_{qp'}(f' - f))\mathcal{R}_{pq'}(f)\mathcal{R}_{qp'}(f)\}.\end{aligned} \tag{B6}$$

We now define

$$T_{\min} \equiv \min(T_{pp'}, T_{qq'}) = \min(T_{pq'}, T_{qp'}) = \min(T_p, T_q, T_{p'}, T_{q'}), \quad T_1 \equiv \max(T_{pp'}, T_{qq'}), \quad T_2 \equiv \max(T_{pq'}, T_{qp'}). \tag{B7}$$

The broader sinc functions can be evaluated at $f' = f$, and the expression above simplifies to

$$C_{pq,p'q'}(f,f') \approx \frac{T_{\min}}{2T_{pq}T_{p'q'}}\{T_1\mathrm{sinc}(\pi T_1(f' - f))\mathcal{R}_{pp'}(f)\mathcal{R}_{qq'}(f) + T_2\mathrm{sinc}(\pi T_2(f' - f))\mathcal{R}_{pq'}(f)\mathcal{R}_{qp'}(f)\}. \tag{B8}$$

This result shows that the estimators are correlated for frequencies separated by less than $\sim 1/T$, and that their correlation drops for wider frequency separations.

Let us consider the *band powers*, centered at frequencies $f_n = n\Delta f$, where Δf is some fixed bandwidth:

$$\mathcal{R}_{pq,f_n} \equiv \int_{f_n - \Delta f/2}^{f_n + \Delta f/2} df' \mathcal{R}_{pq}(f'). \tag{B9}$$

The unbiased estimator $\hat{\mathcal{R}}_{pq,f_n}$ is obtained by integrating $\hat{\mathcal{R}}_{pq}(f)$ over a frequency band. Provided $\Delta f/f_n \ll 1$, we have $\mathcal{R}_{pq,f_n} \approx \Delta f \mathcal{R}_{pq}(f_n)$. The covariance of the band power estimators is obtained by integrating Eq. (B8) over the bandwidth Δf for both frequencies f, f'. Provided $T_1\Delta f$, $T_2\Delta f \gg 1$, the sinc functions integrate out, and we are left with

$$\begin{aligned}&\mathrm{cov}(\hat{\mathcal{R}}_{pq,f_n},\hat{\mathcal{R}}_{p'q',f_{n'}})\\&\quad\approx\frac{\delta_{nn'}\Delta f}{2T_{IJ}}\{\mathcal{R}_{pp'}(f_n)\mathcal{R}_{qq'}(f_n)+\mathcal{R}_{pq'}(f_n)\mathcal{R}_{qp'}(f_n)\}\\&\quad\equiv\delta_{nn'}\Delta f\mathcal{C}_{IJ}(f_n),\end{aligned}\tag{B10}$$

where the indices $I\equiv(p,q), J\equiv(p',q')$ label pairs of pulsars, and

$$T_{IJ}\equiv\max\left[\min(T_p,T_q),\min(T_{p'},T_{q'})\right].\tag{B11}$$

2. Discrete sampling

Let us now consider the more realistic case where each pulsar p is timed at $(N_p+1)\gg 1$ *discrete* times $t_k=k\Delta t_p,\ k=-N_p/2,\ldots,N_p/2,$ where $T_p=N_p\Delta t_p$. Typically, $\Delta t_p\sim 2$–4 weeks. We now define

$$\begin{aligned}\tilde{R}_p(f)&\equiv\Delta t_p\sum_{k=-N_p/2}^{N_p/2}\mathrm{e}^{-2\pi i f t_k}R_p(t_k)\\&=T_p\int df_1 R_p(f_1)\frac{\mathrm{sinc}(\pi T_p(f_1-f))}{\mathrm{sinc}(\pi\Delta t_p(f_1-f))}.\end{aligned}\tag{B12}$$

The derivation follows the same route as in the continuous case, except for the issue of *aliasing*, translated mathematically by

$$T\frac{\mathrm{sinc}(\pi T(f'-f))}{\mathrm{sinc}(\pi\Delta t(f'-f))}\approx\sum_{n=-\infty}^{\infty}(-1)^n\delta_{\mathrm{D}}(f'-f-n/\Delta t).\tag{B13}$$

If the timing cross-power spectrum $\mathcal{R}_{pq}(f)$ scales as $f^{-\alpha}$, with $\alpha>1$, then aliasing does not affect any of the results above, as the contribution from higher-order multiples of $1/\Delta t_p$ is negligible relative to the fundamental mode $n=0$. This is expected to be the case for $p\neq q$. However, the single-pulsar timing residual power spectrum $\mathcal{R}_{pp}(f)$ has a constant white noise piece $P_p(f)=\sigma^2_{p,\mathrm{wn}}t_{\mathrm{obs}}$ at sufficiently high frequencies, up to a maximum frequency $|f_{\max}|=1/t_{\mathrm{obs}}$. Here t_{obs} is the duration of an *individual observation* (typically, $t_{\mathrm{obs}}\sim 30$ min), from which a single, averaged TOA is obtained, and $\sigma^2_{p,\mathrm{wn}}$ is the variance of the timing residual (after fitting a timing model) between individual observations. Thus, we find

$$\begin{aligned}&\langle\tilde{R}_p(f)\tilde{R}_p^*(f')\rangle\\&\quad\approx\frac{T_p}{2}\sum_{n=-\Delta t_p/t_{\mathrm{obs}}}^{\Delta t_p/t_{\mathrm{obs}}}\mathcal{R}_{pp}(f-n/\Delta t_p)\mathrm{sinc}(\pi T_p(f'-f))\\&\quad=\frac{T_p}{2}\mathrm{sinc}(\pi T_p(f'-f))(\mathcal{R}_{pp}(f)+2\sigma^2_{p,\mathrm{wn}}\Delta t_p).\end{aligned}\tag{B14}$$

Hence, the results obtained for the continuum-sampling case carry over to the discrete-sampling case, provided one includes the white noise contribution $2\sigma^2_{p,\mathrm{wn}}\Delta t_p$ in pulsar autocorrelations. We emphasize that this term accounts for aliasing, i.e., from the white noise power at all harmonics of $1/\Delta t_p$, up to the maximum frequency $1/t_{\mathrm{obs}}$.

APPENDIX C: DENSE AND ISOTROPIC PULSAR DISTRIBUTION LIMIT

In the limit where pulsars are densely and isotropically distributed across the sky, the Fisher matrix becomes proportional to

$$\mathcal{F}_\infty(\chi)\equiv\int\frac{d^2\hat{p}}{4\pi}\frac{d^2\hat{q}}{4\pi}\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega})\boldsymbol{\gamma}_{\hat{p}\hat{q}}(\hat{\Omega}'),\quad\chi\equiv\hat{\Omega}\cdot\hat{\Omega}'.\tag{C1}$$

Now remember that the pairwise timing response function is given by

$$\gamma_{\hat{p}\hat{q}}(\hat{\Omega})=\frac{\hat{p}^a\hat{p}^b\hat{q}^c\hat{q}^d\mathfrak{T}_{abcd}(\hat{\Omega})}{(1+\hat{p}\cdot\hat{\Omega})(1+\hat{q}\cdot\hat{\Omega})}.\tag{C2}$$

The double angular integral over pulsar directions can thus be factorized:

$$\begin{aligned}\mathcal{F}_\infty(\hat{\Omega}\cdot\hat{\Omega}')&=\tilde{\mathcal{K}}_{aba'b'}(\hat{\Omega},\hat{\Omega}')\mathfrak{T}_{a'b'c'd'}(\hat{\Omega}')\\&\quad\times\tilde{\mathcal{K}}_{c'd'cd}(\hat{\Omega}',\hat{\Omega})\mathfrak{T}_{cdab}(\hat{\Omega}),\end{aligned}\tag{C3}$$

$$\tilde{\mathcal{K}}_{aba'b'}(\hat{\Omega},\hat{\Omega}')\equiv\int\frac{d^2\hat{p}}{4\pi}\frac{\hat{p}_a\hat{p}_b\hat{p}_{a'}\hat{p}_{b'}}{(1+\hat{p}\cdot\hat{\Omega})(1+\hat{p}\cdot\hat{\Omega}')}.\tag{C4}$$

Since $\mathfrak{T}_{a'b'c'd'}(\hat{\Omega}')$ is orthogonal to $\hat{\Omega}'$ in all indices, and trace-free in the first and last pairs of indices, one may replace $\tilde{\mathcal{K}}_{aba'b'}(\hat{\Omega},\hat{\Omega}')$ by its projection orthogonal to $\hat{\Omega}'$ and trace-free on the right two indices. The same holds for the left two indices. Upon projecting on $\mathcal{I}$, we find

$$\mathcal{F}_\infty(\hat{\Omega}\cdot\hat{\Omega}')=4\mathcal{K}_{aba'b'}(\hat{\Omega},\hat{\Omega}')\mathcal{K}_{aba'b'}(\hat{\Omega},\hat{\Omega}'),\tag{C5}$$

$$\mathcal{K}_{aba'b'}(\hat{\Omega},\hat{\Omega}')\equiv\int\frac{d^2\hat{p}}{4\pi}\frac{(\hat{p}^\perp_a\hat{p}^\perp_b-\frac{1}{2}(\hat{p}^\perp)^2\delta^\perp_{ab})(\hat{p}^{\perp'}_{a'}\hat{p}^{\perp'}_{b'}-\frac{1}{2}(\hat{p}^{\perp'})^2\delta^{\perp'}_{a'b'})}{(1+\hat{p}\cdot\hat{\Omega})(1+\hat{p}\cdot\hat{\Omega}')},\tag{C6}$$

where $\hat{p}^\perp\equiv\hat{p}-(\hat{p}\cdot\hat{\Omega})\hat{\Omega}$ and $\hat{p}^{\perp'}\equiv\hat{p}-(\hat{p}\cdot\hat{\Omega}')\hat{\Omega}'$.

The tensor $\mathcal{K}_{aba'b'}(\hat{\Omega},\hat{\Omega}')$ is symmetric, trace-free, and orthogonal to $\hat{\Omega}$ in its first two indices, and symmetric, trace-free, and orthogonal to $\hat{\Omega}'$ in its last two indices. It therefore has four independent components.

Given the preferred directions $\hat{\Omega}$, $\hat{\Omega}'$, one may construct two rank-2 tensors that are symmetric, trace-free, and orthogonal to $\hat{\Omega}$ on both indices. Defining $\boldsymbol{V}=\hat{\Omega}\times\hat{\Omega}'$, those two tensors are

$$\begin{aligned}\boldsymbol{\mathcal{A}}(\hat{\Omega},\hat{\Omega}') &\equiv (\hat{\Omega}'-\chi\hat{\Omega})\otimes(\hat{\Omega}'-\chi\hat{\Omega})-\boldsymbol{V}\otimes\boldsymbol{V},\\ \boldsymbol{\mathcal{B}}(\hat{\Omega},\hat{\Omega}') &\equiv (\hat{\Omega}'-\chi\hat{\Omega})\otimes\boldsymbol{V}+\boldsymbol{V}\otimes(\hat{\Omega}'-\chi\hat{\Omega}).\end{aligned}\tag{C7}$$

Note that both $\hat{\Omega}'-\chi\hat{\Omega}$ and $\boldsymbol{V}$ have norm $\sqrt{1-\chi^2}$, where $\chi\equiv\hat{\Omega}\cdot\hat{\Omega}'$, which is why $\boldsymbol{\mathcal{A}}$ is indeed trace-free.

Therefore the rank-4 tensor $\boldsymbol{\mathcal{K}}(\hat{\Omega},\hat{\Omega}')$ ought to take the form

$$\begin{aligned}\boldsymbol{\mathcal{K}}(\hat{\Omega},\hat{\Omega}') &= A\boldsymbol{\mathcal{A}}(\hat{\Omega},\hat{\Omega}')\otimes\boldsymbol{\mathcal{A}}(\hat{\Omega}',\hat{\Omega})\\ &\quad+B\boldsymbol{\mathcal{B}}(\hat{\Omega},\hat{\Omega}')\otimes\boldsymbol{\mathcal{B}}(\hat{\Omega}',\hat{\Omega})\\ &\quad+C\boldsymbol{\mathcal{A}}(\hat{\Omega},\hat{\Omega}')\otimes\boldsymbol{\mathcal{B}}(\hat{\Omega}',\hat{\Omega})\\ &\quad+D\boldsymbol{\mathcal{B}}(\hat{\Omega},\hat{\Omega}')\otimes\boldsymbol{\mathcal{A}}(\hat{\Omega}',\hat{\Omega}),\end{aligned}\tag{C8}$$

where A, B, C, D only depend on χ. Now, $\boldsymbol{\mathcal{K}}$ is symmetric under the exchange of the first two indices and the last two indices, simultaneously with the exchange of $\hat{\Omega}$, $\hat{\Omega}'$. Since $\boldsymbol{\mathcal{B}}(\hat{\Omega}',\hat{\Omega})=-\boldsymbol{\mathcal{B}}(\hat{\Omega},\hat{\Omega}')$ (if we do not change the definition of $\boldsymbol{V}=\hat{\Omega}\times\hat{\Omega}'$), then we must have $D=-C$. Last, $\boldsymbol{\mathcal{K}}(-\hat{\Omega},-\hat{\Omega}')=\boldsymbol{\mathcal{K}}(\hat{\Omega},\hat{\Omega}')$, which implies $C=D=0$. We have thus found that

$$\begin{aligned}\boldsymbol{\mathcal{K}}(\hat{\Omega},\hat{\Omega}') &= A(\chi)\boldsymbol{\mathcal{A}}(\hat{\Omega},\hat{\Omega}')\otimes\boldsymbol{\mathcal{A}}(\hat{\Omega}',\hat{\Omega})\\ &\quad+B(\chi)\boldsymbol{\mathcal{B}}(\hat{\Omega},\hat{\Omega}')\otimes\boldsymbol{\mathcal{B}}(\hat{\Omega}',\hat{\Omega}).\end{aligned}\tag{C9}$$

The desired function is the contraction of $\boldsymbol{\mathcal{K}}$ with itself in its first two indices and in its last two indices. Using the fact that $(\boldsymbol{\mathcal{A}}:\boldsymbol{\mathcal{B}})\equiv\mathcal{A}_{ab}\mathcal{B}_{ab}=0$, we get

$$\mathcal{F}_\infty(\chi)=4[A^2(\boldsymbol{\mathcal{A}}:\boldsymbol{\mathcal{A}})^2+B^2(\boldsymbol{\mathcal{B}}:\boldsymbol{\mathcal{B}})^2].\tag{C10}$$

Last, we have

$$\boldsymbol{\mathcal{A}}:\boldsymbol{\mathcal{A}}=2(1-\chi^2)^2=\boldsymbol{\mathcal{B}}:\boldsymbol{\mathcal{B}}.\tag{C11}$$

Hence we have found

$$\mathcal{F}_\infty(\chi)=16(1-\chi^2)^4[A^2+B^2].\tag{C12}$$

The next step is now to determine $A(\chi)$ and $B(\chi)$. We do so by computing the following contractions:

$$(\hat{\Omega}'\otimes\hat{\Omega}'):\boldsymbol{\mathcal{A}}(\hat{\Omega},\hat{\Omega}')=(1-\chi^2)^2=(\hat{\Omega}'\otimes\boldsymbol{V}):\boldsymbol{\mathcal{B}}(\hat{\Omega},\hat{\Omega}'),\tag{C13}$$

$$(\hat{\Omega}'\otimes\boldsymbol{V}):\boldsymbol{\mathcal{A}}(\hat{\Omega},\hat{\Omega}')=0=(\hat{\Omega}'\otimes\hat{\Omega}'):\boldsymbol{\mathcal{B}}(\hat{\Omega},\hat{\Omega}').\tag{C14}$$

We therefore arrive at

$$(1-\chi^2)^4A=(\hat{\Omega}'\otimes\hat{\Omega}'):\boldsymbol{\mathcal{K}}(\hat{\Omega},\hat{\Omega}'):(\hat{\Omega}\otimes\hat{\Omega}),\tag{C15}$$

$$(1-\chi^2)^4B=(\hat{\Omega}'\otimes\boldsymbol{V}):\boldsymbol{\mathcal{K}}(\hat{\Omega},\hat{\Omega}'):(\boldsymbol{V}\otimes\hat{\Omega}).\tag{C16}$$

It is now "only a matter of" computing these contractions, which are scalar integrals. To do so, let us introduce some notation:

$$\begin{aligned}&x\equiv\hat{p}\cdot\hat{\Omega},\quad y\equiv\hat{p}\cdot\hat{\Omega}',\\ &P(x,y,\chi)\equiv1-\chi^2-x^2-y^2+2\chi xy=(\boldsymbol{V}\cdot\hat{p})^2\geq0.\end{aligned}\tag{C17}$$

We then get

$$\begin{aligned}&\hat{\Omega}'_a\hat{\Omega}'_b\left(\hat{p}^\perp_a\hat{p}^\perp_b-\frac{1}{2}(\hat{p}^\perp)^2\delta^\perp_{ab}\right)\\ &\quad=(y-\chi x)^2-\frac{1}{2}(1-\chi^2)(1-x^2)\\ &\quad=\frac{1}{2}(1-\chi^2)(1-x^2)-P(x,y,\chi),\end{aligned}\tag{C18}$$

$$\begin{aligned}&\hat{\Omega}'_aV_b\left(\hat{p}^\perp_a\hat{p}^\perp_b-\frac{1}{2}(\hat{p}^\perp)^2\delta^\perp_{ab}\right)\\ &\quad=(y-\chi x)\boldsymbol{V}\cdot\hat{p}=\pm(y-\chi x)\sqrt{P(x,y,\chi)}.\end{aligned}\tag{C19}$$

So we find

$$\begin{aligned}(1-\chi^2)^4A(\chi)&=\int\frac{d^2\hat{p}}{4\pi}\frac{[\frac{1}{2}(1-\chi^2)(1-x^2)-P(x,y,\chi)][\frac{1}{2}(1-\chi^2)(1-y^2)-P(x,y,\chi)]}{(1+x)(1+y)}\\ (1-\chi^2)^4B(\chi)&=\int\frac{d^2\hat{p}}{4\pi}\frac{(y-\chi x)(x-\chi y)P(x,y,\chi)}{(1+x)(1+y)}.\end{aligned}\tag{C20}$$

Now recall, from Appendix A, that

$$d^2\hat{p}=\frac{2dxdy}{\sqrt{P(x,y,\chi)}}.\tag{C21}$$

Evaluating the integrals, and simplifying, we find

$$(1-\chi^2)^4 A(\chi) = \frac{(1-\chi)^2}{3}\left(\frac{1+\chi}{4}(9-\chi(4+\chi)) + 3(1-\chi)\log\frac{1-\chi}{2}\right), \tag{C22}$$

$$(1-\chi^2)^4 B(\chi) = \frac{(1-\chi)^2}{3}\left((\chi+1)(\chi-2) - 3(1-\chi)\log\frac{1-\chi}{2}\right). \tag{C23}$$

After simplifying, we thus arrive at our final expression, Eq. (71).

[1] M. V. Sazhin, Sov. Astron. **22**, 36 (1978), https://ui.adsabs.harvard.edu/abs/1978SvA....22...36S/abstract.
[2] S. Detweiler, Astrophys. J. **234**, 1100 (1979).
[3] R. W. Hellings and G. S. Downs, Astrophys. J. **265**, L39 (1983).
[4] European Pulsar Timing Array (EPTA), http://www.epta.eu.org.
[5] North American Nanohertz Observatory for Gravitational Waves (NANOGrav), http://nanograv.org.
[6] Parkes Pulsar Timing Array (PPTA), https://www.atnf.csiro.au/research/pulsar/ppta.
[7] International Pulsar Timing Array (IPTA), http://ipta4gw.org.
[8] L. Lentati *et al.* (EPTA Collaboration), Mon. Not. R. Astron. Soc. **453**, 2576 (2015).
[9] R. M. Shannon *et al.* (PPTA Collaboration), Science **349**, 1522 (2015).
[10] Z. Arzoumanian *et al.* (NANOGrav Collaboration), Astrophys. J. **859**, 47 (2018).
[11] S. R. Taylor, M. Vallisneri, J. A. Ellis, C. M. F. Mingarelli, T. J. W. Lazio, and R. van Haasteren, Astrophys. J. Lett. **819**, L6 (2016).
[12] S. Burke-Spolaor *et al.*, Astron. Astrophys. Rev. **27**, 5 (2019).
[13] C. M. F. Mingarelli, Nat. Astron. **3**, 8 (2019).
[14] X. Siemens, J. S. Hazboun, P. T. Baker, S. Burke-Spolaor, D. Madison, C. Mingarelli, J. Simon, and T. Smith, arXiv:1907.04960.
[15] P. D. Lasky *et al.*, Phys. Rev. X **6**, 011035 (2016).
[16] N. J. Cornish and A. Sesana, Classical Quantum Gravity **30**, 224005 (2013).
[17] C. M. Mingarelli, T. Sidery, I. Mandel, and A. Vecchio, Phys. Rev. D **88**, 062005 (2013).
[18] 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, Nat. Astron. **1**, 886 (2017).
[19] S. R. Taylor and J. R. Gair, Phys. Rev. D **88**, 084001 (2013).
[20] S. R. Taylor *et al.*, Phys. Rev. Lett. **115**, 041101 (2015).
[21] S. C. Hotinli, M. Kamionkowski, and A. H. Jaffe, Open J. Astrophys. **2**, 8 (2019).
[22] J. Gair, J. D. Romano, S. Taylor, and C. M. F. Mingarelli, Phys. Rev. D **90**, 082001 (2014).
[23] S. R. Taylor, R. van Haasteren, and A. Sesana, Phys. Rev. D **102**, 084039 (2020).
[24] N. J. Cornish and R. van Haasteren, arXiv:1406.4511.
[25] L. Knox, Phys. Rev. D **52**, 4307 (1995).
[26] M. Tegmark, D. J. Eisenstein, W. Hu, and A. de Oliveira-Costa, Astrophys. J. **530**, 133 (2000).
[27] E. S. Phinney, arXiv:astro-ph/0108028.
[28] A. H. Jaffe and D. C. Backer, Astrophys. J. **583**, 616 (2003).
[29] G. B. Rybicki and A. P. Lightman, *Radiative Processes in Astrophysics* (Wiley-VCH, 1986).
[30] T. L. Smith and R. Caldwell, Phys. Rev. D **95**, 044036 (2017).
[31] R. Kato and J. Soda, Phys. Rev. D **93**, 062003 (2016).
[32] W. L. Burke, Astrophys. J. **196**, 329 (1975).
[33] F. B. Estabrook and H. D. Wahlquist, Gen. Relativ. Gravit. **6**, 439 (1975).
[34] N. J. Cornish, Phys. Rev. D **80**, 087101 (2009).
[35] C. M. F. Mingarelli and A. B. Mingarelli, J. Phys. Commun. **2**, 105002 (2018).
[36] C. M. F. Mingarelli and T. Sidery, Phys. Rev. D **90**, 062011 (2014).
[37] M. Anholm, S. Ballmer, J. D. E. Creighton, L. R. Price, and X. Siemens, Phys. Rev. D **79**, 084030 (2009).
[38] S. J. Chamberlin, J. D. E. Creighton, X. Siemens, P. Demorest, J. Ellis, L. R. Price, and J. D. Romano, Phys. Rev. D **91**, 044048 (2015).
[39] M. Tegmark, A. Taylor, and A. Heavens, Astrophys. J. **480**, 22 (1997).
[40] J. S. Hazboun, J. D. Romano, and T. L. Smith, Phys. Rev. D **100**, 104028 (2019).
[41] A. R. Thompson, J. M. Moran, J. Swenson, and W. George, *Interferometry and Synthesis in Radio Astronomy*, 3rd ed. (Springer International Publishing, 2017).
[42] X. Siemens, J. Ellis, F. Jenet, and J. D. Romano, Classical Quantum Gravity **30**, 224015 (2013).
[43] D. Eardley, D. Lee, A. Lightman, R. Wagoner, and C. Will, Phys. Rev. Lett. **30**, 884 (1973).