Spectral separation of the stochastic gravitational-wave background for LISA: Observing both cosmological and astrophysical backgrounds

Guillaume Boileau[*] and Nelson Christensen[†]

*Artemis, Observatoire de la Côte d'Azur, Université Côte d'Azur,
CNRS, CS 34229, F-06304 Nice Cedex 4, France*

Renate Meyer[‡]

Department of Statistics, University of Auckland, Auckland, New Zealand

Neil J. Cornish[§]

*eXtreme Gravity Institute, Department of Physics, Montana State University,
Bozeman, Montana 59717, USA*

With the goal of observing a stochastic gravitational-wave background (SGWB) with LISA, the spectral separability of the cosmological and astrophysical backgrounds is important to estimate. We attempt to determine the level with which a cosmological background can be observed given the predicted astrophysical background level. We predict detectable limits for the future LISA measurement of the SGWB. Adaptive Markov chain Monte Carlo methods are used to produce estimates with the simulated data from the LISA Data Challenge. We also calculate the Cramer-Rao lower bound on the variance of the SGWB parameter estimates based on the inverse Fisher information using the Whittle likelihood. The estimation of the parameters is done with the three LISA channels A, E, and T. We simultaneously estimate the noise using a LISA noise model. Assuming the expected astrophysical background around $\Omega_{\mathrm{GW,astro}}(25\ \mathrm{Hz}) = 0.355 \rightarrow 35.5 \times 10^{-9}$, a cosmological SGWB normalized energy density of around $\Omega_{\mathrm{GW,Cosmo}} \approx 1 \times 10^{-12}$ to 1×10^{-13} can be detected by LISA after 4 years of observation.

I. INTRODUCTION

Since the accomplishment of the first detection of gravitational waves from the merger of two stellar mass black holes [1] by Advanced LIGO [2,3] and thereafter with Advanced Virgo [4,5], gravitational-wave observatories have become a new means to observe astronomical phenomena. So far LIGO and Virgo have announced the observation of 50 signals produced from compact binary coalescence [6,7], including two from binary neutron star mergers [8,9]. Gravitational wave detections are expanding our understanding of astrophysics and of the Universe.

The Laser Interferometer Space Antenna (LISA) [10] is a future ESA mission, also supported by NASA, with the aim to observe gravitational waves in the low frequency band $[10^{-5}, 1]$ Hz. The mission lifetime will nominally be 4 years, but could be extendable to 6 or 10 years of scientific observations. LISA is a triangular constellation of three spacecraft, separated from one another at a distance of

$L = 2.5 \times 10^9$ m. The low-frequency band is rich with gravitational-wave signals. The foreground of LISA will be dominated by sources from our galaxy, the Milky Way. White dwarf binaries [11–13] are numerous ($\sim$35 million binaries), and relatively near the LISA constellation. For example, recently the Zwicky Transient Facility measured a double white dwarf with an orbital period estimated at 7 minutes [14], which corresponds to a gravitational-wave emission of $\simeq$30 mHz. LISA can be expected to observe many resolved binaries, many of which are already known from photometry studies and constitute the so-called verification binaries [15,16]. Well-studied systems like this can be used to verify the LISA performance, acting as a way to confirm the sensitivity of LISA. We can expect to have one in a thousand binaries which are resolvable. The large majority of the galactic binaries are unresolved and form a stochastic signal. The stochastic gravitational-wave background from white dwarf binaries or galactic foreground will be anisotropic and the signal will not be a pure power law. A stochastic gravitational-wave background (SGWB) [17,18] will have a significant contribution from unresolved binaries, such as binary black holes and binary neutron stars. This background is essentially isotropic, and

[*]guillaume.boileau@oca.eu
[†]nelson.christensen@oca.eu
[‡]renate.meyer@auckland.ac.nz
[§]ncornish@montana.edu

its level can be predicted from the signals observed by LIGO and Virgo [19,20]. Another important SGWB would be from cosmological sources [17]. The origin of this background comes from the early Universe [21,22], with the possibility to measure the inflation scenario parameters [23]. Cosmic strings could be another observable source [24]. A cosmologically produced background can be modeled as a flat spectral energy density $\propto f^0$ [25].

In this paper, we present a strategy to separate the two SGWBs (astrophysical and cosmological), as well as the LISA noise, using a Bayesian strategy [12,26] based on an adaptive Markov chain Monte Carlo (A-MCMC) algorithm. We then show LISA's ability to measure a cosmological SGWB for different magnitudes for the astrophysical background. The SGWB from astrophysical sources today represents an important goal, especially considering the current observations by LIGO and Virgo [27,28].

Numerous studies have recently been presented which address how to possibly detect a cosmologically produced SGWB in the presence of an astrophysically produced SGWB. For example a recent study displayed the use of principal component analysis to model and observe a SGWB in the presence of a foreground from binary black holes and binary neutron stars in the LISA observation band [29]. A component separation method was proposed in Ref. [30], where they showed that it is possible to detect an isotropic SGWB. The method uses maximum likelihood parameter estimation with Fisher information matrices. This is proposed to replace an MCMC approach, and applied to the LIGO-Virgo observational band.

The proposal in Ref. [31] is to use a number of broken power-law filters to separate different backgrounds with gravitational-wave detectors on the Earth. In the study of Ref. [32] the proposal is to divide the data into individual short time segments. The method used the procedures described in Ref. [33] to search the segments for the presence of a binary black hole signal, either through direct detection or subthreshold by generating a Bayesian evidence. A cosmological SGWB would be present in all segments, whereas a probability would exist for the presence of a binary black hole merger for the segments. The method is general, and could be applied to LIGO-Virgo or LISA. The study presented in Ref. [34] noted that the sensitivity of third-generation gravitational-wave detectors, such as Einstein Telescope [35] or Cosmic Explorer [36], will be so good that almost every binary black hole merger in the observable Universe can be directly detected, and then removed from the search for a cosmological SGWB. The study of Ref. [37] then explored how to do such a subtraction of binary black hole merger signals, and the consequences of the effect of residuals from such subtractions. Another study used Bayesian methods to address spectral separation for LIGO-Virgo observations, but tried to address how to separate a SGWB from a correlated magnetic noise background produced by the Schumann

resonances [38–40]; the study is, however, general and can be applied to spectral separation for different types of backgrounds [41]. This study was then expanded to address the simultaneous estimation of astrophysical and cosmological SGWBs, and displayed that this will be especially important for third-generation ground-based detectors [42]. Another study, specifically dedicated to LISA observations [43] proposed to divide the data into bins, and then within in each bin, a fit is made to a power law or a constant amplitude; a variation on this approach is presented here [44]. The claim is that this method is more dynamic and able to fit arbitrarily shaped SGWBs. The study of Ref. [45] showed how to assign Bayes factors and probabilities to differentiate a SGWB signal from instrumental noise.

All the SGWB studies referenced above are summarized in Tables II, III, IV, respectively for LIGO/Virgo, LISA, and third-generation detectors. We compare the goals, methods, the performance, the limitations and the application; see the Appendix. The study we present in this paper, using Bayesian parameter estimation methods, has the advantage of fitting two backgrounds and the LISA noise simultaneously. We note the possibility to expand the work presented here to estimate more complex LISA noise, and adding new models for the SGWB; for example, more complex SGWBs could include broken power laws, peaks in the frequency domain, or an anisotropic SGWB from our galaxy.

The organization of the paper is as follows. In Sec. II we introduce the SGWB spectral separation problem for LISA, and then describe the inverse of the Fisher information matrix of the SGWB parameters, and how this provides the Cramer-Rao lower bound on the variance of the parameter estimates. In Sec. III we describe the A-MCMC. The simulated LISA mock data is presented in Sec. IV. Presented in Sec. V are the parameter estimation procedures and results using the LISA A and T channels; Sec. IV presents similar results using the LISA A, E and T channels. Conclusions are given in Sec. VI.

II. SPECTRAL SEPARATION

An isotropic SGWB observed today $\Omega_{\mathrm{GW}}(f)$ can be modeled with the frequency variation of the energy density of the gravitational waves, ρ_{GW}, where $\mathrm{d}\rho_{\mathrm{GW}}$ is the gravitational-wave energy density contained in the frequency band $[f, f + df]$) [46]. The distribution of the energy density over the frequency domain can be expressed as,

$$\Omega_{\mathrm{GW}}(f) = \frac{f}{\rho_c} \frac{\mathrm{d}\rho_{\mathrm{GW}}}{\mathrm{d}\ln(f)} = \sum_k \Omega_{\mathrm{GW}}^{(k)}(f) \qquad (1)$$

where the critical density of the Universe is $\rho_c = \frac{3H_0^2 c^2}{8\pi G}$. In this paper we approximate the spectral energy density as a collection of power-law contributions (this is a simplified model), $\Omega_{\mathrm{GW}}(f) \simeq \sum_k A_k (\frac{f}{f_{\mathrm{ref}}})^{\alpha_k}$ where the energy spectral

density amplitude of the component k (representing the different SGWBs) is A_k, with the respective slope α_k and $f_{\rm ref}$ is some characteristic frequency. The SGWB is predicted to have a slope component $\alpha \approx 0$ for the cosmological background. This is true for scale-invariant processes, and this is approximately true for the standard inflation and certainly false for cosmic strings and turbulence. However for our study here we will model the cosmologically produced SGWB with $\alpha = 0$. In addition, we will use $\alpha = \frac{2}{3}$ for a compact-binary-produced astrophysical background. According to Farmer and Phinney the slope is $\alpha = \frac{2}{3}$ for quasicircular binaries evolving purely under gravitational-wave emission [47]. The eccentricity and environmental effects can modify the slope. We also note the limitations of our power-law model as phase transitions in the early Universe can produce two-part power laws, with a traction between the rising and falling power-law component at some peak frequency. But we start in this study with two power-law backgrounds. As the two backgrounds are superimposed, the task is to simultaneously extract both the astrophysical and cosmological properties, i.e., to simultaneously estimate the astrophysical and the cosmological contributions to the energy spectral density.

To avoid identification issues, we use a Bayesian approach by putting informative priors on the individual slope and amplitude parameters. Our work here builds on that of Adams and Cornish [48] where they demonstrated that it is possible to separate a SGWB from the instrumental noise in a Bayesian context. Similarly Adams and Cornish then showed that one could detect a cosmological SGWB in the presence of a background produced by white dwarf binaries in our galaxy [11]. Since the production of those studies LIGO and Virgo have observed gravitational waves from binary black hole and binary neutron star coalescence. We now know that there will definitely be an astrophysically produced background across the LISA observation band produced by compact binary coalescences over the history of the Universe [20], and if LISA is to observe a cosmologically produced background it will be necessary to separate the two.

The literature displays large differences in the estimation of the magnitude of the astrophysically produced SGWB. A recent simulation of the SGWB from merging compact binary sources with the StarTrack code [49] predicts an amplitude around $\Omega_{\rm GW} \simeq 4.97 \times 10^{-9}$ to 2.58×10^{-8} at 25 Hz. However another study considered the binary black hole and binary neutron star observations by LIGO/Virgo, and produced predictions going from the LISA observational band to the LIGO/Virgo band. They estimated an amplitude for the astrophysical SGWB of $\Omega_{\rm GW} \simeq 1.8 \times 10^{-9}$ to 2.5×10^{-9} at 25 Hz [20]. These amplitudes can be propagated to the LISA band by recalling Eq. (1) and using $f_{\rm ref} = 25$ Hz and $\alpha = 2/3$. In the context of an effort to observe a cosmological SGWB we have large variations due to the predictions of the astrophysical component.

In our study here we predict the accuracy of a measurement of $\Omega_{\rm GW}^{(0)}$ with astrophysical inputs of differing magnitudes using $f_{\rm ref} = 25$ Hz, $\Omega_{\rm GW}^{(\frac{2}{3})} = [3.55 \times 10^{-10}, 1.8 \times 10^{-9}, 3.55 \times 10^{-9}, 3.55 \times 10^{-8}]$ after 4 years of observation. We use the orthogonal LISA A, E, and T channels, which are created from the time-delay interferometry (TDI) variables X, Y, and Z [50]. Our method fits the parameters of two stochastic backgrounds, and simultaneously the LISA noise with the help of the channel T. We assume uncorrelated noise TDIs between the "science" channels (A, E) and the noise channel (T). The T channel is "signal insensitive" for gravitational-wave wavelengths larger than the arm lengths. The noise channel T is obtained from a linear combination [50] of the TDIs channel (X, Y, Z). We demonstrate a good ability to estimate the noise present in the two science data channels A and E. We can then set a limit on the ability to detect the cosmological SGWB. The predictions from the Bayesian study are confirmed via a study of the frequentist estimation of the error. Namely, we use a Fisher information analysis, performed for the spectral separation independently of the Bayesian A-MCMC approach. The inverse of the Fisher information matrix of the SGWB parameters, presented in Sec. II, provides the Cramer-Rao lower bound on the variance of the SGWB parameter estimates.

A useful toy model to consider is the problem of separating two independent stationary mean-zero Gaussian noise processes that have different power spectra $S_{n_1}(f) = A_1 f^{\alpha_1}$ and $S_{n_2}(f) = A_2 f^{\alpha_2}$. Suppose we have data that is formed from the sum of these two independent noise processes

$$d(t) = n_1(t) + n_2(t), \quad t = 1, \ldots, T. \quad (2)$$

After a Fourier transform to $\tilde{d}(f_k) = \frac{1}{\sqrt{T}} \sum_{i=1}^{T} d(t) e^{-it f_k}$ at Fourier frequencies $f_k = 2\pi k/T$, $k = 0, \ldots, N = \frac{T}{2} - 1$ (for T even), we can write

$$\tilde{d}(f_k) = \tilde{n}_1(f_k) + \tilde{n}_2(f_k), \quad k = 0, \ldots, N. \quad (3)$$

Then the vector $\tilde{d}$ has an asymptotic complex multivariate Gaussian distribution with a diagonal covariance matrix. The diagonal elements are given by the values of the spectral density $S(f_k) = A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2}$. Our assumption of independence implies that one can simply sum the individual spectral densities of the two noise processes.

The Whittle likelihood approximation in the frequency domain can then be written as

$$p(d|A_1, \alpha_1, A_2, \alpha_2) = \prod_{k=1}^{N} \frac{1}{\pi S(f_k)} e^{-\frac{\tilde{d}(f_k)^\star \tilde{d}(f_k)}{S(f_k)}} \quad (4)$$

where $S(f_k) = A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2}$. The product $I_n(f_k) = \tilde{d}(f_k)^\star \tilde{d}(f_k)$ is the *periodogram*, the squared magnitude

of the Fourier coefficients at the frequency f_k. The log likelihood (up to an additive constant) is thus

$$\ln p(d|A_1, \alpha_1, A_2, \alpha_2) = -\sum_{k=1}^{N}\left(\frac{I_n(f_k)}{S(f_k)} + \ln S(f_k)\right). \quad (5)$$

A. The Fisher information

The Fisher information matrix Γ for a parameter vector $\theta = (\theta_1, ..., \theta_p)$ is given by the expected value of the negative Hessian of the log likelihood. The element in row i and column j of the Fisher information is given by

$$\Gamma_{ij} = E\left[-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln p(d|\theta)\right]. \quad (6)$$

The Fisher information can be easily obtained for the parameter vector $(A_1, \alpha_1, A_2, \alpha_2)$ by using that (asymptotically) $E[I_n(f_k)] = S(f_k)$ and $\Gamma_{ij} = \Gamma_{ji}$,

$$\Gamma_{11} = \sum_{k=1}^{N}\frac{f_k^{2\alpha_1}}{(A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2})^2}, \quad (7)$$

$$\Gamma_{22} = \sum_{k=1}^{N}\frac{(A_1 f_k^{\alpha_1}\ln f_k)^2}{(A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2})^2}, \quad (8)$$

$$\Gamma_{33} = \sum_{k=1}^{N}\frac{f_k^{2\alpha_2}}{(A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2})^2}, \quad (9)$$

$$\Gamma_{44} = \sum_{k=1}^{N}\frac{(A_2 f_k^{\alpha_2}\ln f_k)^2}{(A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2})^2}, \quad (10)$$

$$\Gamma_{12} = \Gamma_{21} = \sum_{k=1}^{N}\frac{A_1 f_k^{2\alpha_1}\ln f_k}{(A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2})^2}, \quad (11)$$

$$\Gamma_{13} = \Gamma_{31} = \sum_{k=1}^{N}\frac{f_k^{\alpha_1+\alpha_2}}{(A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2})^2}, \quad (12)$$

$$\Gamma_{14} = \Gamma_{41} = \sum_{k=1}^{N}\frac{A_2 f_k^{\alpha_1+\alpha_2}\ln f_k}{(A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2})^2}, \quad (13)$$

$$\Gamma_{23} = \Gamma_{32} = \sum_{k=1}^{N}\frac{A_1 f_k^{\alpha_1+\alpha_2}\ln f_k}{(A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2})^2}, \quad (14)$$

$$\Gamma_{24} = \Gamma_{42} = \sum_{k=1}^{N}A_1 A_2\frac{A_1 A_2 f_k^{\alpha_1+\alpha_2}\ln^2 f_k}{(A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2})^2}, \quad (15)$$

$$\Gamma_{34} = \Gamma_{43} = \sum_{k=1}^{N}\frac{A_2 f_k^{2\alpha_2}\ln f_k}{(A_1 f_k^{\alpha_1} + A_2 f_k^{\alpha_2})^2}. \quad (16)$$

B. The Cramer-Rao bound

The Fisher information can be used to give a lower bound for the variance of any unbiased estimator, the so-called Cramer-Rao bound. For any unbiased estimator $\hat{\theta}_i$ of the unknown parameter θ_i, its standard error $\Delta\hat{\theta}_i$ satisfies

$$(\Delta\hat{\theta}_i)^2 \geq \Gamma_{ii}(\theta)^{-1} = \frac{1}{E[-\frac{\partial}{\partial\theta_i}\frac{\partial}{\partial\theta_i}\ln p(d|\theta)]}. \quad (17)$$

Under certain regularity conditions, the posterior distribution of a parameter θ is asymptotically Gaussian, centered at the posterior mode and covariance matrix equal to the inverse of the negative Hessian of the posterior distribution evaluated at the posterior mode. For flat priors, the posterior density is proportional to the likelihood, the posterior mode is the maximum likelihood estimate and the standard error $\Delta\hat{\theta}_i$ of the Bayesian estimator $\hat{\theta}_i$ of the parameter θ_i can be approximated by evaluating the Fisher information at $\hat{\theta}_i$, i.e.,

$$\Delta\hat{\theta}_i \approx \Gamma_{ii}(\hat{\theta}_i)^{-1/2}. \quad (18)$$

Defining the *uncertainty* of an estimate $\hat{\theta}_i$ by

$$\frac{\Delta\hat{\theta}_i}{\hat{\theta}_i} \quad (19)$$

we say that we can estimate the parameter θ_i with on error of 10% based on the Fisher analysis if the uncertainty of a parameter estimate is equal to 0.1. The purpose of this study is to derive a threshold on the separability by an A-MCMC routine with the likelihood of Eq. (4). In the following we will thus have a limiting value for the separability of the cosmological SGWB parameters and the astrophysical SGWB.

We use a toy problem to display the separability of two stochastic backgrounds according to their slope difference. For this we fix one background $\Omega_1(f) = A_1(\frac{f}{f_{\text{ref}}})^{\alpha_1} = \Omega_{2/3}(\frac{f}{f_{\text{ref}}})^{\alpha_{2/3}} = 3.55 \times 10^{-9}(\frac{f}{25\text{ Hz}})^{2/3}$, and we leave free the slope of the second background $\Omega_2(f) = A_2(\frac{f}{f_{\text{ref}}})^{\alpha_2} = \Omega_0(\frac{f}{f_{\text{ref}}})^{\alpha_0} = 1 \times 10^{-12}(\frac{f}{25\text{ Hz}})^{\alpha_0}$. We show the uncertainties $(\frac{\Delta\hat{\theta}_i}{\hat{\theta}_i}$ for $\theta_i \in [\Omega_{2/3}, \alpha_{2/3}, \Omega_0, \alpha_0]$, with $\Delta\hat{\theta}_i$ being the error from the Fisher information; see Sec. II B) for the amplitudes and spectral slopes as a function of the difference between the spectral slopes ($\delta\alpha = \alpha_0 - \alpha_{2/3}$). This quantity is also called the coefficient of variation or the relative standard deviation, and this is the absolute value of the standard deviation divided by the mean of the parameter. We use this quantity to appreciate the dispersion of values around the mean. It is preferable to use this quantity because it is unitless. Thus it is easier to compare parameters of different units and ranges of values. Figure 1
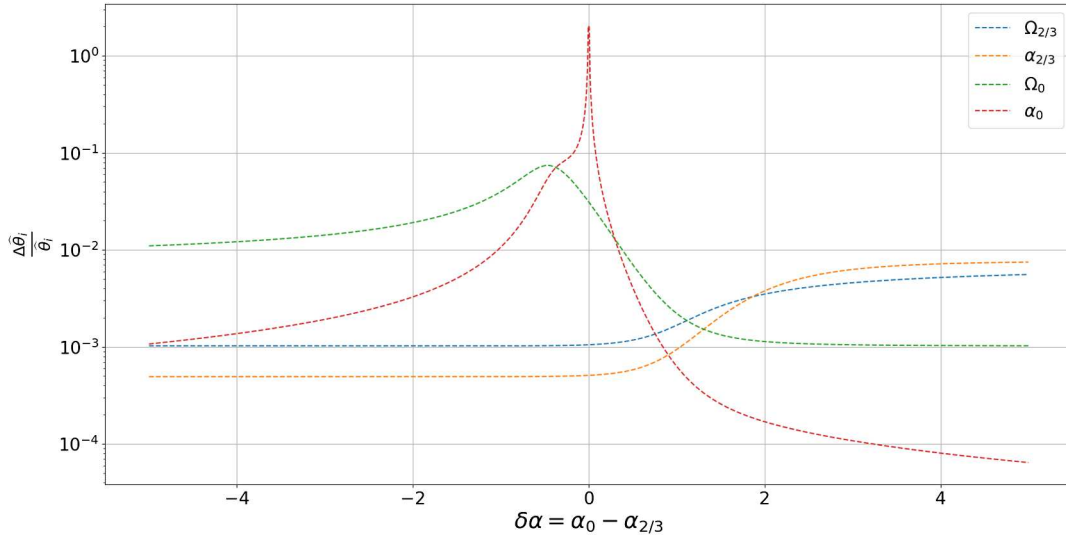
FIG. 1.   Uncertainties ($\frac{\Delta\hat{\theta}_i}{\hat{\theta}_i}$) of the amplitudes and spectral slopes as a function of the difference in the differential spectral slopes ($\delta\alpha = \alpha_0 - \alpha_{2/3}$).

displays the uncertainties ($\frac{\Delta\hat{\theta}_i}{\hat{\theta}_i}$) as a function of $\delta\alpha$ between -5 and 5.

The uncertainty of the parameter α_0 becomes larger when the slope difference $\delta\alpha$ is near zero. Here it is more difficult to separate the two backgrounds when their slopes are similar. The uncertainties are also not symmetric about $\delta\alpha = 0$ because when the slope changes the amplitude is also changing by a factor $f_{\rm ref}^{-\alpha}$. The uncertainty of the amplitude parameter Ω_0 is maximal when the two amplitude parameters are identical. The position of the maximum changes for different inputs of Ω_0; if Ω_0 increases the position of the maximum converges to $\delta\alpha = 0$.

III. ADAPTIVE MARKOV CHAIN MONTE CARLO

A. Markov chain Monte Carlo

Bayesian inference quantifies the estimation and uncertainties of unknown parameters based on the observation of events that depend on these parameters. The quantification uses the *posterior* probability distribution. It is obtained using Bayes' theorem [see Eq. (20)] by updating the *prior* distribution of the parameters with the *likelihood* $p(d|\theta)$, the conditional distribution of the observations given the parameters:

$$p(\theta|d) = \frac{p(d|\theta)p(\theta)}{p(d)} \qquad (20)$$

where $p(\theta)$ is the prior distribution, $p(\theta|d)$ is the posterior distribution, and $p(d) = \int p(d|\theta)p(\theta)d\theta$ is the *evidence*.

MCMC methods [51] provide a numerical strategy to compute the joint posterior distribution and its marginal distributions. It is a sampling-based approach that simulates a Markov chain constructed in such a way that its invariant distribution is the joint posterior.

B. Metropolis-Hasting sampler

As it is generally difficult to sample independently from a multivariate distribution, MCMC methods draw dependent samples from Markov chains. The predominant MCMC algorithm is the Metropolis-Hastings (MH) algorithm. It is based on the rejection or acceptance of a candidate parameter θ' where the acceptance probability is given by the likelihood ratio between the candidate and the previously sampled parameter value. Thus, any move in the direction of higher likelihood (towards the maximum likelihood estimation) will always be accepted, but because downhill moves still have a chance to be accepted, the MH algorithm avoids getting stuck in local maxima.

Metropolis-Hastings algorithm
(1) Randomly select an initial point $\theta^{(0)}$
(2) At the nth iteration:
   (a) Generation of candidate θ' with the proposal distribution $g(\theta'|\theta^{(n)})$
   (b) Calculation of acceptance probability $\alpha = \min\left[1, \frac{p(d|\theta')}{p(d|\theta^{(n)})}\frac{p(\theta^{(n)})}{p(\theta')}\right]$
   (c) Accept/Reject
      (i) Generation of a uniform random number u on [0, 1]
      (ii) if $u \leq \alpha$, accept the candidate: $\theta^{(n+1)} = \theta'$
      (iii) if $u > \alpha$, reject the candidate: $\theta^{(n+1)} = \theta^{(n)}$

Note that the proposal distribution g is often chosen to be Gaussian centered around the current parameter value. While executing the algorithm, we can monitor the acceptance rate, the proportion of candidates that were accepted. On the one

hand, if this number is too close to 0 then the algorithm makes large moves into the tails of the posterior distribution which have low acceptance probability causing the chain to stay at one value for a long time. On the other hand, a high acceptance rate indicates that the chain makes only small moves causing slow mixing. To control the mixing of the Markov chain we can introduce an adaptive step-size parameter that controls the size of the moves; this is the standard deviation in the case of a univariate Gaussian proposal or the covariance matrix of a multivariate Gaussian proposal. As the iterations of the algorithm proceed, it is possible to dynamically modify the step size to improve the convergence of the chain. Intuitively, an optimal proposal would be as close to the posterior distribution as possible. Using a Gaussian proposal, its covariance matrix should thus be as close to the covariance matrix of the posterior distribution. Since the previous MCMC samples can be used to provide a consistent estimate of the covariance matrix, this estimate can be used to adapt the proposal on the fly, as detailed in Sec. III C.

C. Adaptive Markov chain Monte Carlo

We use the version of the adaptive Metropolis MCMC from Robert and Rosenthal [52]. For a p-dimensional MCMC we can perform the Metropolis-Hasting algorithm with a proposal density $g_n(.|\theta^{(n)})$ in iteration n defined by a mixture of Gaussian proposals:

$$g_n(.|\theta^{(n)}) = (1-\beta)N\left(\theta^{(n)}, \frac{(2.28)^2}{p}\Sigma_n\right)$$
$$+ \beta N\left(\theta^{(n)}, \frac{(0.1)^2}{p}I_p\right) \quad (21)$$

where Σ_n is the current empirical estimate of the covariance matrix, $\beta = 0.25$ is a constant, p is the dimensionality of the parameter space, N is the multinormal distribution and I_p is the $p \times p$ identity matrix. We compute an estimate Σ_n of the covariance matrix using the last hundred samples of the chain. The chain generated from an adaptive algorithm is not Markovian but the diminishing adaptation condition ensures ergodicity and thus the convergence to the stationary distribution.

IV. DATA FROM THE MOCK LISA DATA CHALLENGE

A. Noise and SGWB energy spectral density of the MLDC

The Mock LISA Data Challenge (MLDC) provides simulations of the signal and noise of LISA in the approximation of one arm. We use the (X, Y, Z) time series of the *LDC1-6* data set from the MLDC webpage [53]. These are simulations of a binary-produced SGWB of the form $\Omega_{GW}(f) = \Omega_{2/3}(\frac{f}{f_{ref}})^\alpha$ for $f_{ref} = 25$ Hz with a slope $\alpha = \frac{2}{3}$ and an amplitude of $\Omega_{2/3} = 3.55 \times 10^{-9}$ (at 25 Hz). Figures 2 and 3 display the gravitational-wave periodograms for the (X, Y, Z) and (A, E, T) channels.

We can transform the X, Y, Z time series to the A, E, T channels according to

$$\begin{cases} A = \frac{1}{\sqrt{2}}(Z - X), \\ E = \frac{1}{\sqrt{6}}(X - 2Y + Z), \\ T = \frac{1}{\sqrt{3}}(X + Y + Z). \end{cases} \quad (22)$$

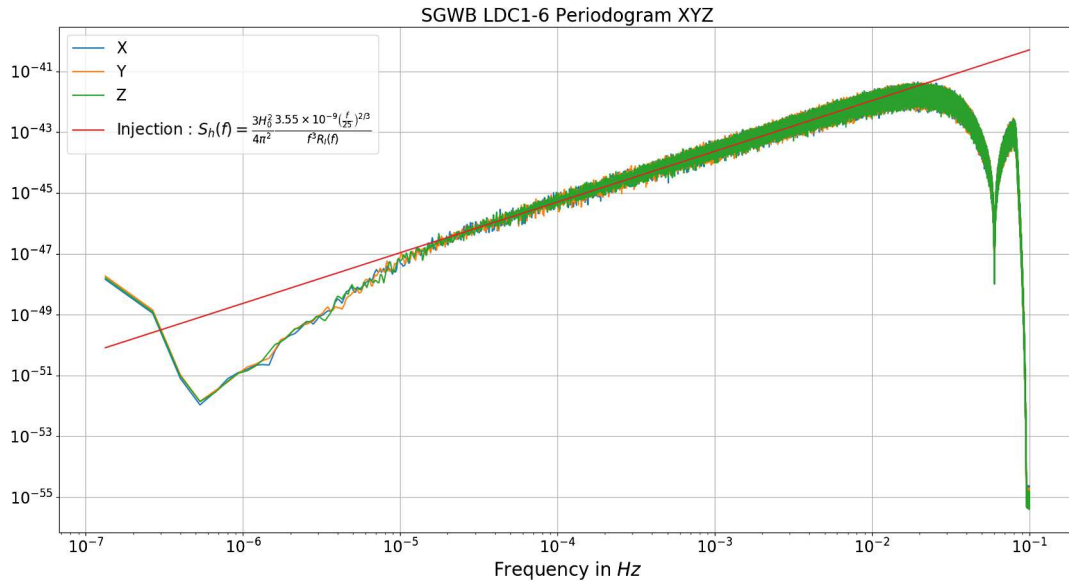

FIG. 2. Periodogram of the channels (X, Y, Z) of the SGWB from MLDC (*LDC1-6* noiseless) with a single background $[\Omega_{GW}(f) = 3.55 \times 10^{-9}(\frac{f}{25\,Hz})^{2/3}]$

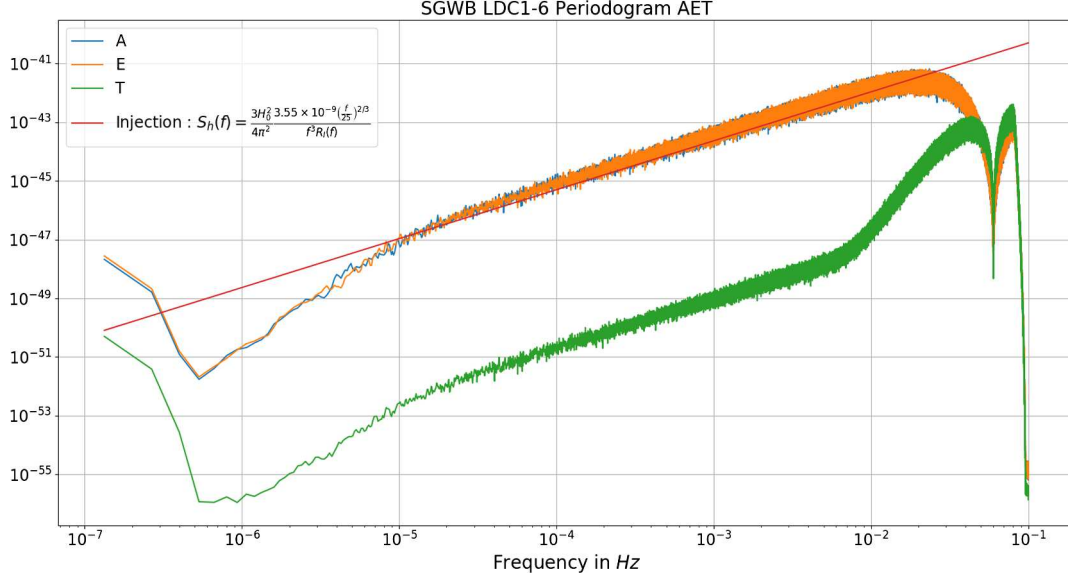FIG. 3. Periodogram of the channels (A, E, T) of the SGWB from MLDC (*LDC1-6* noiseless) with an single background $(\Omega_{\mathrm{GW}}(f) = 3.55 \times 10^{-9}(\frac{f}{25\,\mathrm{Hz}})^{2/3})$.

This linear combination of the original channels used to define T has been shown to be insensitive to the gravitational-wave signal. While this is not exactly true, we will maintain that assumption for this analysis. As such, T can be regarded as a null channel which contains mainly only noise, while channels A and E are the science channels, containing the gravitational-wave signal in the presence of noise [18]. In the following we focus on the science channels, A and E.

In this study we use a simplified model where we assume equal noise levels on each spacecraft. According to Adams and Cornish [11] one can use a more complicated model that allowed for different noise levels. Future work will address this, plus the situation where the slope parameters for the noise can also vary. These parameters could then also be estimated by Bayesian parameter estimation methods.

For the following studies we restrict the frequency band to correspond to the LISA band $[10^{-5}, 1]$ Hz. The power spectral density (PSD) of the channel T, S_T, can be described as (according to Ref. [53])

$$S_T(x) = 16 S_{\mathrm{Op}}(x)(1 - \cos(x))\sin^2(x)$$
$$+ 128 S_{\mathrm{pm}}(x)\sin^2(x)\sin^4\left(\frac{x}{2}\right) \qquad (23)$$

with $x = \frac{2\pi L}{c}f$, where S_{Op} is the optical metrology system noise and S_{pm} is the acceleration and displacement noise. The LISA noise budget is

$$\begin{cases} S_{\mathrm{Op}}(f) = N_{\mathrm{Opt}}L^2\left(1 + \left(\frac{8\ \mathrm{mHz}}{f}\right)^4\right), \\ S_{\mathrm{Pm}}(f) = N_{\mathrm{Acc}}L^2 S_{\mathrm{Acc}}(f)S_{\mathrm{Dis}}(f) \end{cases} \qquad (24)$$

with

$$\begin{cases} S_{\mathrm{Acc}}(f) = \left(1 + \left(\frac{0.4\ \mathrm{mHz}}{f}\right)^2\right)\left(1 + \frac{f}{8\ \mathrm{mHz}}\right)^4, \\ S_{\mathrm{Dis}}(f) = (2\pi f)^{-4}\left(\frac{2\pi f}{c}\right)^2. \end{cases} \qquad (25)$$

The two free parameters, N_{Opt} and N_{Acc}, are the respective levels of the two principal sources of noise in the LISA noise budget. In the LISA Science Requirements Document [54], the level of the LISA noise acceleration is $N_{\mathrm{Acc}} = 1.44 \times 10^{-48}\ \mathrm{s}^{-4}\ \mathrm{Hz}^{-1}$ and the upper limit on the level of the optical metrology system noise is $N_{\mathrm{Opt}} = 3.6 \times 10^{-47}\ \mathrm{Hz}^{-1}$. From the modeling of the strain requirements of the mission performance requirements, this is a maximization of the noise level. The LISA noise budget corresponds to all sources of contamination that contribute to the power spectral density of the LISA detection system. The two noise sources correspond to estimates of different physical effects. We clearly do not yet have the true values for these physical effects; we presently only have estimates from experiments. The LISA requirements fixed the limit of the two magnitude levels so as to respect LISA's detection performance. In Fig. 4, the green curve is the analytic noise model of the PSD of the channel T with the parameters from the proposal [54]. The blue curve is the periodogram for the channel T of the MLDC data (*LDC1-6*
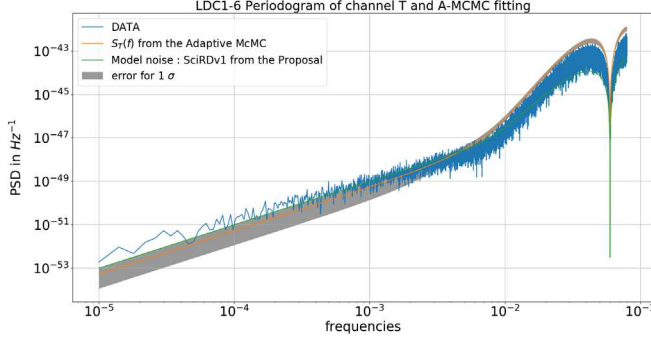
FIG. 4. Power spectral density of the channel T from the MLDC (in blue) [53]. The green line represents the analytic noise model of the power spectral density of the channel T with the parameters from the proposal [54]. The orange line is the model from Eq. (23) with the values fit with the MCMC. In grey is the 1σ error. This is the uncertainty calculated from Eq. (37), where we take $d\text{PSD}_T$ with $dN_{\text{pos}} = \sigma_{N_{\text{pos}}}$ and $dN_{\text{acc}} = \sigma_{N_{\text{acc}}}$; σ is the standard deviation of the posterior estimation. See Fig. 5 and Eq. (26).

SGWB signal); this is the magnitude squared of the Fourier coefficients for the data [see Eq. (22)]. Assuming the functional form of the noise PSD in channel T is given by Eq. (23), we can use the A-MCMC (see Sec. III) to fit the LISA noise parameters N_{Opt} and N_{Acc}. The priors for the two components are flat log-uniform distributions and we specify $\beta = 0.01$ and $N = 200000$ in the A-MCMC algorithm. The orange curve in Fig. 4 is the estimated PSD based on Eq. (23) with N_{Opt} and N_{Acc} replaced by the posterior means of samples obtained via the A-MCMC, given in Eq. (26). The 1σ error bands are overlaid in grey. Figure 5 shows the corner plot for the posterior samples of the two parameters, and the empirical posterior distributions seem to be well approximated by Gaussian distributions. It shows that this model yields a reasonable fit to the simulated channel T data. We acknowledge that this is a rigid noise model for the purpose of this study, and future work will include more realistic scenarios: allowing for different noise levels on each spacecraft [11], allowing for small modifications of the transfer functions, and allowing for small modifications in the spectral slopes of the noise components. The posterior means of the two noise parameters are

$$
\begin{cases}
\hat{N}_{\text{acc}} = 7.08 \times 10^{-51} \pm 4 \times 10^{-53} \ \text{s}^{-4} \, \text{Hz}^{-1}, \\
\hat{N}_{\text{Opt}} = 1.91 \times 10^{-47} \pm 4 \times 10^{-49} \ \text{Hz}^{-1}.
\end{cases}
\tag{26}
$$

The gravitational-wave energy spectral density Ω_{GW} can be defined as

$$
\Omega_{\text{GW},I}(f) = \frac{2\pi^2}{3H_0^2} f^3 \frac{\text{PSD}_I(f)}{R_I(f)}
\tag{27}
$$

for $I = A, E$, where H_0 is the Hubble-Lemaître constant ($H_0 \simeq 2.175 \times 10^{-18}$ Hz), PSD_I is the power spectral
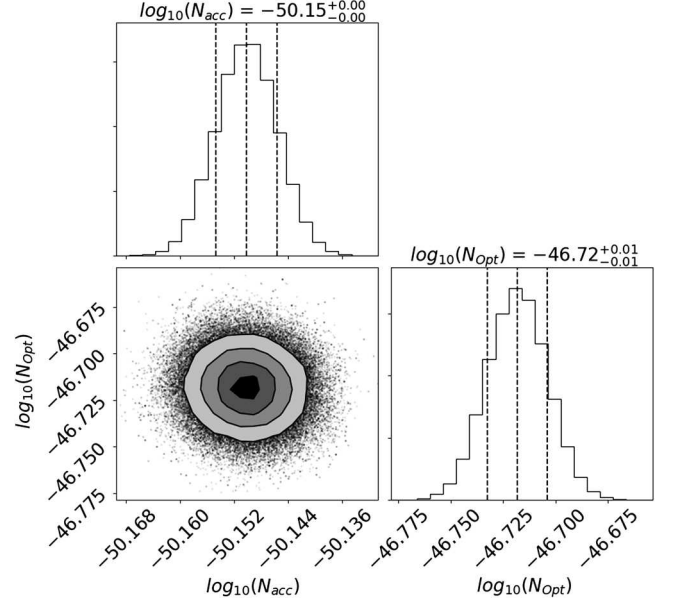

FIG. 5. Corner plot for the A-MCMC generated posterior distributions for the power spectral density of the channel T of the MLDC data set, estimating the two magnitudes of the LISA noise model from the proposal [54]. The vertical dashed lines on the posterior distributions represent, from left to right, the quantiles [16%, 50%, 84%].

density of the channel I and R_I is the response function. An asymptotically unbiased estimate of PSD_I is given by the periodogram $I_n(f) = \sum_{k=1}^{N} |\tilde{d}(f_k)|^2 = \tilde{d}_I^*(f_k) \tilde{d}_I(f_k)$.

We use two different response functions for the MLDC data: one system of equations for the noiseless data Eq. (28), and one for the noisy data Eq. (30)

$$
\begin{cases}
R_A(f) = R_{AA}(f) \frac{16}{9} \frac{2}{\pi} \left(\frac{f}{f_*} \right)^4 \sin^{-2}(f/f_*), \\
R_E(f) = R_{EE}(f) \frac{16}{7} \frac{2}{\pi} \left(\frac{f}{f_*} \right)^4 \sin^{-2}(f/f_*)
\end{cases}
\tag{28}
$$

with R_{II} given in Ref. [48], $f_* = \frac{c}{2\pi L}$, and

$$
\begin{aligned}
R_{AA}(f) &= R_{EE}(f) \\
&= 4\sin^2\left(\frac{f}{f_*} \right) \left[\frac{3}{10} + \frac{169}{1680} \left(\frac{f}{f_*} \right)^2 \right. \\
&\quad + \frac{85}{6048} \left(\frac{f}{f_*} \right)^4 - \frac{178273}{15667200} \left(\frac{f}{f_*} \right)^6 \\
&\quad \left. + \frac{19121}{2476656000} \left(\frac{f}{f_*} \right)^8 \right],
\end{aligned}
\tag{29}
$$

$$
R_{\text{I}}(f) = \frac{S_{\text{II}}(f)L}{3cS_p} \left[\frac{36}{10} \frac{f}{f_*} \sin^{-2}(f/f_*) \right]^2
\tag{30}
$$

where $S_{\text{II}}(f) = 8\sin^2(\frac{f}{f_*})[4S_a(1 + \cos(\frac{f}{f_*}) + \cos^2(\frac{f}{f_*})) + S_p(2 + \cos(\frac{f}{f_*}))]$ was defined in Ref. [18] with $S_a = \frac{9 \times 10^{-50}}{(2\pi f)^4}(1 + (\frac{10^{-4}}{f})^2)$, $S_p = 4.10^{-42}$ Hz^{-1} and $f_* = \frac{c}{2\pi L}$.

The energy spectral density of the astrophysical background from the MLDC is a power law according to the documentation of the LISA Data Challenge Manual [53] given by $\Omega_{GW}(f) = 3.55 \times 10^{-9}(\frac{f}{25\,Hz})^{2/3}$. Figures 6 and 7 show the energy periodogram $\hat{\Omega}_{GW,I}(f) = \frac{2\pi^2}{3H_0^2}f^3\frac{I_n(f)}{R_I(f)}$ for channel A in blue and for channel E in orange. The green curve is the power-law model with the parameters $(\Omega_\alpha, f_{ref}, \alpha)$ with $\Omega_{GW} = \Omega_\alpha(\frac{f}{f_{ref}})^\alpha$ from the MLDC documentation. The data at high frequency cannot be used because the transformations of Eqs. (28) and (30) are valid for low frequency. We use the frequency band $[2.15 \times 10^{-5}, 9.98 \times 10^{-3}]$ Hz.

B. Uncertainty of the cosmological component Ω_0 from the A-MCMC

According to Sec. II B, one can calculate the uncertainty of the estimation of the parameter Ω_0 (the cosmological amplitude of the spectral energy density), namely $\frac{\Delta\Omega_0}{\Omega_0}$. To estimate this quantity from the Fisher information, we use the formulas given in Sec. II and the inverse matrix of the Fisher information (blue line in Fig. 11).

Not surprisingly we can predict a better separability (uncertainty is less) for high values of the cosmological background. The uncertainty can be calculated independently with the A-MCMC calculation:

$$\frac{\Delta\Omega_0}{\Omega_0} = \frac{\sigma_{\Omega_0}}{\Omega_0}. \tag{31}$$

This ratio is calculated and represented as the scatter points in Fig. 11. We can also estimate the error of the uncertainty estimation [see Eq. (32)] from the estimation of the full width at half maximum of the posterior distributions. The uncertainties (from the A-MCMC) are given by

$$\begin{cases} \text{Error}_{+,I} = \frac{\sigma_{\Omega_0}}{|\Omega_0 - \sigma_{\Omega_0}|}, \\ \text{Error}_{-,I} = \frac{\sigma_{\Omega_0}}{|\Omega_0 + \sigma_{\Omega_0}|}. \end{cases} \tag{32}$$
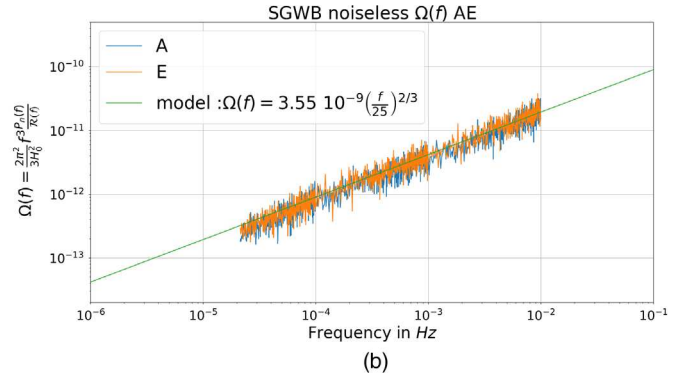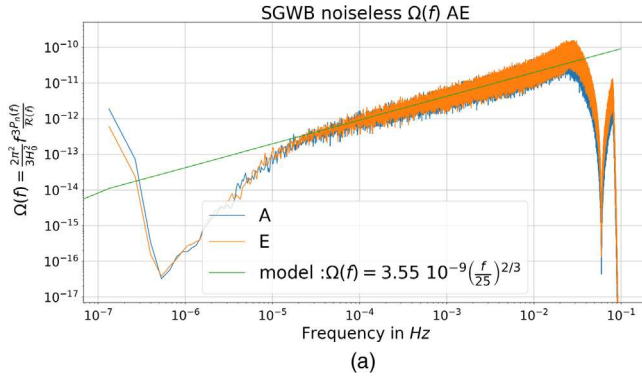

FIG. 6. Observations in channels $[A, E]$ of the spectral energy density of the SGWB from astrophysical background $\Omega_{GW}(f)$ of the MLDC for the noiseless channel, Eq. (28). (a) Total frequency band of Channels A and E. (b) Reduced frequency band 2.15×10^{-5} to 9.98×10^{-3} Hz of Channels A and E.
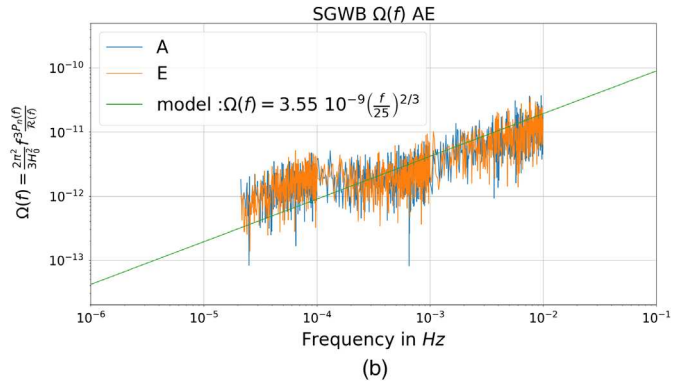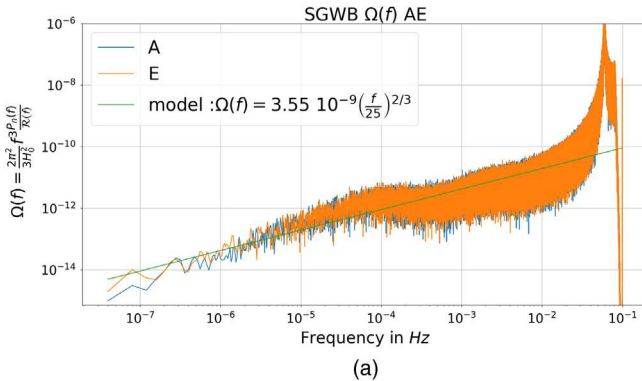

FIG. 7. Observations in channels $[A, E]$ of the spectral energy density of the SGWB from astrophysical background $\Omega_{GW}(f)$ of the MLDC for the noisy channel, Eq. (30). (a) Total frequency band of Channels A and E. (b) Reduced frequency band 2.15×10^{-5} to 9.98×10^{-3} Hz of Channels A and E.

V. STOCHASTIC GRAVITATIONAL-WAVE BACKGROUND FITTING WITH ADAPTIVE MARKOV CHAIN MONTE CARLO USING THE CHANNEL T AND THE TWO SCIENCE CHANNELS A AND E

In this section we consider the null channel T and the science channels A and E. We assume that the observation of the noise in channel T informs us of the noise in channels A and E. We follow the formalism of Smith and Caldwell [55].

We can simulate the noise and SGWB in the frequency domain:

$$
\begin{cases}
\text{PSD}_A = S_A + N_A, \\
\text{PSD}_E = S_E + N_E, \\
\text{PSD}_T = N_T.
\end{cases} \tag{33}
$$

With $S_A(f) = S_E(f) = \frac{3H_0^2}{4\pi^2}\frac{\Omega_{GW,\alpha}(\frac{f}{f_{\text{ref}}})^\alpha}{f^3}$, $f_{\text{ref}} = 25$ Hz, the noise components $N_A(f) = N_E(f)$ and $N_T(f)$ can be written as

$$
\begin{cases}
N_A = N_1 - N_2, \\
N_T = N_1 + 2N_2,
\end{cases} \tag{34}
$$

with

$$
\begin{cases}
N_1(f) = (4S_s(f) + 8(1 + \cos^2(\frac{f}{f_*}))S_a(f))|W(f)|^2, \\
N_2(f) = -(2S_s(f) + 8S_a(f))\cos(\frac{f}{f_*})|W(f)|^2,
\end{cases} \tag{35}
$$

where $W(f) = 1 - e^{-\frac{2if}{f_*}}$ and

$$
\begin{cases}
S_s(f) = N_{\text{Pos}}, \\
S_a(f) = \frac{N_{\text{acc}}}{(2\pi f)^4}\left(1 + \left(\frac{0.4 \text{ mHz}}{f}\right)^2\right).
\end{cases} \tag{36}
$$

The LISA noise budget is given from the LISA Science Requirements Document [54]. To create the data for our example, we use an acceleration noise of $N_{\text{acc}} = 1.44 \times 10^{-48}$ s^{-4} Hz^{-1} and the optical path-length fluctuation $N_{\text{Pos}} = 3.6 \times 10^{-41}$ Hz^{-1}. We can estimate the magnitude of the noise from the channel T. One should note the importance of using the channel T to estimate the noise in the channels A and E, as it is then possible to parametrize an A-MCMC of six parameters, $\theta = (N_{\text{acc}}, N_{\text{Pos}}, \Omega_{2/3}, \alpha_{2/3}, \Omega_0, \alpha_0)$. We can also calculate the propagation of uncertainties for the power spectral densities with the partial derivative method. As such, we can estimate the error on the measurement realized by a fit of the parameters θ, $d\text{PSD}_{\text{I}} = \sqrt{\sum_\theta (\frac{\partial \text{PSD}_{\text{I}}}{\partial\theta})^2 d\theta^2}$. We then obtain for two SGWBs $\Omega_{\text{astro}}(f) = \Omega_{2/3}(\frac{f}{f_{\text{ref}}})^{2/3}$, $\Omega_{\text{cosmo}}(f) = \Omega_0(\frac{f}{f_{\text{ref}}})^0$,

$$
\begin{cases}
d\text{PSD}_{\text{I}} = \left[N_I(0, dN_{\text{acc}}, f)^2 + N_I(dN_{\text{pos}}, 0, f)^2 + S_I(\Omega_{2/3}, \alpha_{2/3}, \Omega_0, \alpha_0, f)^2(d\Omega_0^2 + d\Omega_{2/3}^2 + \ln\left(\frac{f}{f_{\text{ref}}}\right)^2(\Omega_{2/3}^2 d\alpha_{2/3}^2 + \Omega_0^2 d\alpha_0^2))\right]^{1/2}, \\
d\text{PSD}_T = [N_T(0, dN_{\text{acc}}, f)^2 + N_T(dN_{\text{pos}}, 0, f)^2]^{1/2}
\end{cases} \tag{37}
$$

with $\{dN_{\text{acc}}, dN_{\text{pos}}, d\Omega_{\text{astro}}, d\alpha_{\text{astro}}, d\Omega_{\text{cosmo}}, d\alpha_{\text{cosmo}}\}$ being the positive error estimations of the parameters; $I = A, E$. We take 1σ for the posterior distributions. We can also estimate the error of the power spectral density fit using the MCMC chains to produce the error. With the MCMC chains we can calculate a histogram of $\text{PSD}_{\text{I}}(f)$ at each frequency. For each histogram we compute the 68% credible band. This method is similar to that of BayesWave; see Fig. 7 of Ref. [56]. The two methods produce the same error bands, but we need to assume that the posterior distributions are Gaussian. The quadratic sum of the partial errors calculation yields a good estimation of error from MCMC chains if the posterior distributions of the chains are Gaussian.

We can calculate the covariance matrix:

$$
\langle \text{PSD}_{\text{I}}(f), \text{PSD}_J(f) \rangle = \mathcal{C}_{I,J}(\theta, f) \tag{38}
$$

with $I, J = [A, E, T]$. As such, it is possible to parametrize an A-MCMC with six parameters: $\theta = (N_{\text{acc}}, N_{\text{Pos}}, \Omega_{\text{GW}\alpha}, \alpha)$. We can calculate the covariance matrix of $(\tilde{d}_A(f), \tilde{d}_E(f), \tilde{d}_T(f))$

$$
\mathcal{C}(\theta, f) = \begin{pmatrix} S_A + N_A & 0 & 0 \\ 0 & S_E + N_E & 0 \\ 0 & 0 & N_T \end{pmatrix}, \tag{39}
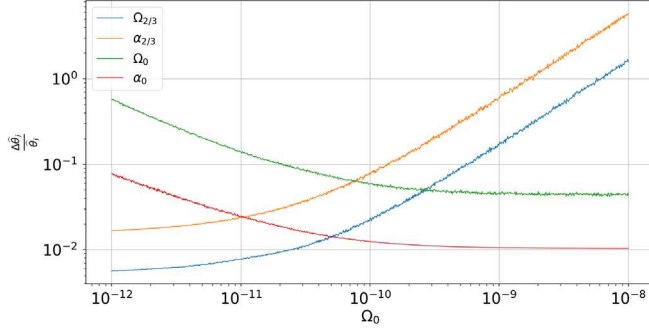$$

FIG. 8. Evolution of the relative uncertainties for the estimation of the parameters $[\Omega_0, \alpha_0, \Omega_{2/3}, \alpha_{2/3}]$ versus the cosmological background amplitude Ω_0. The precision for estimating the parameters is affected by the value of the cosmological amplitude Ω_0. We use $\Omega_{2/3} = 3.55 \times 10^{-9}$, $\alpha_{2/3} = \frac{2}{3}$ and $\alpha_0 = 0$.

$$
C^{-1}(\theta, f) = K \begin{pmatrix} (S_A + N_A)^{-1} & 0 & 0 \\ 0 & (S_E + N_E)^{-1} & 0 \\ 0 & 0 & N_T^{-1} \end{pmatrix} \quad (40)
$$

and $K(f_k) = \det(\mathcal{C}) = \frac{1}{(S_A+N_A)(S_E+N_E)N_T}$. We use the definition of the Whittle likelihood from Ref. [18], and the log likelihood is

$$
\mathcal{L}(\mathbf{d}|\theta) = -\frac{1}{2} \sum_{k=0}^{N} \left[\sum_{I,J=[A,E,T]} (\sqrt{d_I(f)}(\mathcal{C}^{-1})_{IJ}\sqrt{d_J(f)}) \right.
$$
$$
\left. + \ln(2\pi K(f_k)) \right]
$$
$$
= -\frac{1}{2} \sum_{k=0}^{N} \left[\frac{d_A^2}{S_A + N_A} + \frac{d_E^2}{S_E + N_E} + \frac{d_T^2}{N_T} \right.
$$
$$
\left. + \ln(8\pi^3 (S_A + NA)(S_E + N_E)N_T) \right], \quad (41)
$$

$$
F_{ab} = \frac{1}{2} \mathrm{Tr} \left(\mathcal{C}^{-1} \frac{\partial \mathcal{C}}{\partial \theta_a} \mathcal{C}^{-1} \frac{\partial \mathcal{C}}{\partial \theta_b} \right)
$$
$$
= \sum_{k=0}^{N} \left[\frac{\frac{\partial(S_A+N_A)}{\partial \theta_a} \frac{\partial(S_A+N_A)}{\partial \theta_b}}{2(S_A + N_A)^2} \right.
$$
$$
\left. + \frac{\frac{\partial(S_E+N_E)}{\partial \theta_a} \frac{\partial(S_E+N_E)}{\partial \theta_b}}{2(S_E + N_E)^2} + \frac{\frac{\partial N_T}{\partial \theta_a} \frac{\partial N_T}{\partial \theta_b}}{2N_T^2} \right]. \quad (42)
$$

If we have the channel T as zero and we consider the two science channels A and E as independent, we obtain

$$
F_{ab} = \frac{1}{2} \sum_{I=A,E} \sum_{k=0}^{N} \frac{\frac{\partial S_I(f)+N_I(f)}{\partial \theta_a} \frac{\partial S_I(f)+N_I(f)}{\partial \theta_b}}{(S_I(f) + N_I(f))^2}. \quad (43)
$$

TABLE I. Results of the A-MCMC runs with six parameters (two for the LISA noise, two for the astrophysical background and two for the cosmological background). We use the data from the A, E and T channels. The four columns of values correspond to the output of 13 A-MCMC runs. The study is conducted using four values for the amplitude of the astrophysical background after 4 years of observation: 3.55×10^{-8}, 3.55×10^{-9}, 1.8×10^{-9} and 3.55×10^{-10}, and respectively, the same for the error columns. The error estimations come from the posterior distributions.

Input	Values of the A-MCMC				Errors (σ)			
				Ω_{Astro}				
Ω_0	3.55×10^{-8}	3.55×10^{-9}	1.8×10^{-9}	3.55×10^{-10}	3.55×10^{-8}	3.55×10^{-9}	1.8×10^{-9}	3.55×10^{-10}
$1. \times 10^{-8}$	1.011×10^{-8}	9.982×10^{-9}	9.987×10^{-9}	9.992×10^{-9}	3.395×10^{-10}	3.057×10^{-10}	3.106×10^{-10}	2.588×10^{-10}
$5. \times 10^{-9}$	5.014×10^{-9}	4.971×10^{-9}	5.007×10^{-9}	4.960×10^{-9}	1.754×10^{-10}	1.464×10^{-10}	1.506×10^{-10}	1.462×10^{-10}
$2. \times 10^{-9}$	2.005×10^{-9}	1.984×10^{-9}	2.007×10^{-9}	2.083×10^{-9}	7.481×10^{-11}	5.600×10^{-11}	6.588×10^{-11}	5.492×10^{-11}
$1. \times 10^{-9}$	9.972×10^{-10}	1.008×10^{-9}	1.046×10^{-9}	1.046×10^{-9}	4.480×10^{-11}	2.828×10^{-11}	3.196×10^{-11}	3.196×10^{-11}
$5. \times 10^{-10}$	4.965×10^{-10}	4.975×10^{-10}	5.076×10^{-10}	4.956×10^{-10}	2.529×10^{-11}	1.497×10^{-11}	1.703×10^{-11}	1.385×10^{-11}
$2. \times 10^{-10}$	2.002×10^{-10}	1.984×10^{-10}	1.976×10^{-10}	1.976×10^{-10}	1.394×10^{-11}	6.647×10^{-11}	8.251×10^{-12}	5.157×10^{-11}
$1. \times 10^{-10}$	9.981×10^{-11}	1.065×10^{-10}	9.941×10^{-11}	1.003×10^{-10}	9.228×10^{-12}	5.322×10^{-12}	4.050×10^{-12}	3.048×10^{-12}
$5. \times 10^{-11}$	5.013×10^{-11}	5.057×10^{-11}	5.058×10^{-11}	5.163×10^{-11}	7.078×10^{-11}	5.171×10^{-12}	2.879×10^{-12}	1.706×10^{-12}
$2. \times 10^{-11}$	2.006×10^{-11}	2.014×10^{-11}	1.989×10^{-11}	2.016×10^{-11}	5.389×10^{-12}	2.558×10^{-12}	1.130×10^{-12}	8.457×10^{-13}
$1. \times 10^{-11}$	1.001×10^{-11}	1.008×10^{-11}	1.002×10^{-11}	1.026×10^{-11}	4.269×10^{-12}	1.406×10^{-12}	5.902×10^{-13}	4.472×10^{-13}
$5. \times 10^{-12}$	5.011×10^{-12}	4.959×10^{-12}	5.001×10^{-12}	5.024×10^{-12}	3.583×10^{-12}	9.843×10^{-13}	4.526×10^{-13}	2.556×10^{-13}
$2. \times 10^{-12}$	2.196×10^{-12}	1.952×10^{-12}	1.948×10^{-12}	1.985×10^{-12}	3.001×10^{-12}	7.460×10^{-13}	3.190×10^{-13}	1.433×10^{-13}
$1. \times 10^{-12}$	1.019×10^{-12}	1.064×10^{-12}	9.936×10^{-13}	1.013×10^{-12}	2.155×10^{-12}	5.119×10^{-13}	2.233×10^{-13}	1.040×10^{-13}
$1. \times 10^{-13}$		9.891×10^{-14}	1.040×10^{-13}	9.936×10^{-14}		2.002×10^{-13}	1.036×10^{-13}	4.054×10^{-14}

We have a comparable result to that given in Ref. [55], and the inverse of the Fisher information matrix on the diagonal gives the uncertainties of the estimation of the parameters. We see the importance to estimate the "noise" channel T for the estimation of the SGWB.

In Fig. 8 we display the influence of the precision on the fitted parameter versus the value of the cosmological background Ω_0. Obviously, we understand that if the astrophysical background is large it will be harder to measure the cosmological background with high precision.

We have also conducted an A-MCMC study with six parameters: two for the noise channel T, two for the astrophysical background, and two for the cosmological background. We use the data from the two science channels, A and E, along with channel T. Given the magnitude level of the LISA noise budget from the LISA Science Requirements Document [54], we use the acceleration noise $N_{\text{acc}} = 1.44 \times 10^{-48}$ s^{-4} Hz^{-1} and the optical path-length fluctuation $N_{\text{Pos}} = 3.6 \times 10^{-41}$ Hz^{-1}. We make the assumption that the data in channels A and T are independent. The noises in both channels depend on the two parameters N_{pos} and N_{acc}. We aim to estimate the SGWB and noise parameters simultaneously using data from channels A, E and T via our A-MCMC algorithm. Using the additional data from channel T will yield a more efficient estimation procedure and a gain in precision of parameter estimates than using the data from channels A, E only. For four different magnitudes of the astrophysical SGWB, we conduct A-MCMC runs with different values for the amplitude of the cosmological background; see Table I. The A-MCMC is characterized by $\beta = 0.01$, $N = 4000000$ (see Sec. III C) and we use 2000 samples to estimate the covariance matrix. We use log-uniform priors with ten magnitude intervals for the two noise channel parameters $[N_{\text{Opt}}, N_{\text{Acc}}]$ and for the two background amplitudes $[\Omega_{\text{cosmo}}, \Omega_{\text{astro}}]$, a uniform prior for the slope between -0.4 and 0.4 for the cosmological slope α_{cosmo}, and a uniform prior between 0.27 and 1.07 for the astrophysical slope α_{astro}.

We note for comparison purposes the results given in Ref. [55] where the diagonal elements of the inverse of the Fisher information F_{ab} provide the uncertainties of the respective parameter estimates. The Fisher information matrix is a block matrix. Indeed, we have a 6×6 matrix, assuming the parameters are independent. We can thus distinguish two independent types. The first comes from derivatives related to the noise of LISA which generates a 2×2 matrix, $N_{2 \times 2}$. The second type corresponds to a 4×4 matrix giving the derivatives linked to the SGWB, $S_{4 \times 4}$. This second matrix is the same as the one calculated in Sec. II A. So we have
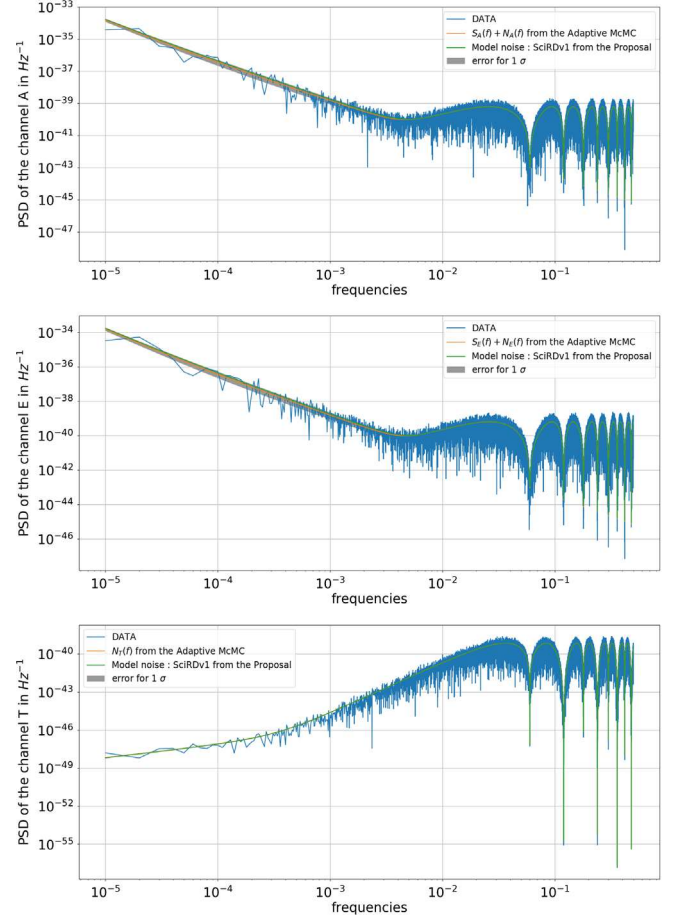

FIG. 9.   Power spectral density of the channels A, E and T from the LISA noise model [55] and an astrophysical SGWB ($\Omega_{2/3} = 3.55 \times 10^{-9}$ at 25 Hz). The figures show the power spectral densities: channel A (top), E (middle), and T (bottom). The parameters are from the proposal [54]. The orange line is the LISA noise model from Ref. [55], the green line is the values from the A-MCMC, and the 1σ error is in grey.

$$F_{ab} = \left[\begin{array}{c|c} N_{2 \times 2} & 0 \\ \hline 0 & S_{4 \times 4} \end{array} \right]. \qquad (44)$$

In Fig. 9, the blue line is the data for $\theta = (N_{\text{acc}}, N_{\text{Pos}}, \Omega_{\text{GW}\alpha}, \alpha) = (1.44 \times 10^{-48}$ s^{-4} Hz^{-1}, 3.6×10^{-41} Hz^{-1}, 3.55×10^{-9}, $\frac{2}{3})$. The data are simulated with the LISA noise model of Eq. (33) with a SGWB of binary origin. The green line is the LISA noise model from Ref. [55]. The A-MCMC is characterized by $\beta = 0.01$, $N = 1000000$ (see Sec. III C) and we use 2000 samples to estimate the covariance matrix. We use log-uniform priors with ten magnitude intervals for the three first parameters and a uniform prior for the slope between $-\frac{4}{3}$ and $\frac{8}{3}$. The orange line in Fig. 9 displays the result of the A-MCMC,
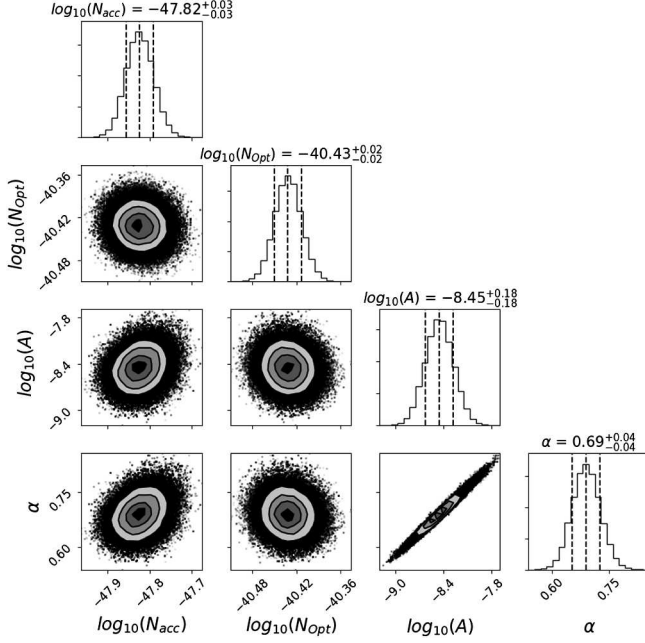
FIG. 10. Corner plot for the A-MCMC using the channels A, E and T. The results are for the two magnitudes for the LISA noise model from the proposal [54], and a single SGWB (amplitude and spectral slope). The vertical dashed lines on the posterior distribution represent from left to right the quantiles [16%, 50%, 84%]. The *true* values for the parameters are $\theta = (N_{\mathrm{acc}}, N_{\mathrm{Pos}}, \Omega_{\mathrm{GW}\alpha}, \alpha) = (1.44 \times 10^{-48}\ \mathrm{s}^{-4}\ \mathrm{Hz}^{-1}, 3.6 \times 10^{-41}\ \mathrm{Hz}^{-1}, 3.55 \times 10^{-9}, \frac{2}{3})$.

and the 1σ error is shown in grey. Figure 10 displays the corner plot from the A-MCMC; the posterior distributions are well approximated by Gaussian distributions. We have evidence of good fits. The estimation of the noise level magnitudes from the parametric estimation yields a positive result because we have the possibility to fit the background with the noise level throughout the frequency domain; it is also possible to have a very efficient estimation of the different noise components thanks to the signal T being devoid of a science signal source.

The advantage of two science channels, A and E, as opposed to one, A or E, is a factor of $\sqrt{2}$ for the error estimation, and hence the overall sensitivity. Indeed, the error of the cosmological amplitude is given by the coefficient (Ω_0, Ω_0) of the square root of the inverse of the Fisher information matrix. We have for one channel (A or E), $\Delta\Omega_{0(A\ \mathrm{or}\ E)} = \sqrt{F_{\Omega_0,\Omega_{0(A\ \mathrm{or}\ E)}}^{-1}}$. For a combination of A and E, we have $\Delta\Omega_{0(A\ \mathrm{and}\ E)} = \dfrac{\Delta\Omega_{0(A\ \mathrm{or}\ E)}}{\sqrt{2}}$, because the two channels respond identically. If the models for the spectrum of A and E were the same, then $F_{a,b_{(A\ \mathrm{and}\ E)}} = 2F_{a,b_{(A\ \mathrm{or}\ E)}}$.

Note that in the LISA observing band we have a ratio of $\frac{\Omega_{\mathrm{astro}}}{\Omega_{\mathrm{Cosmo}}} = 5.29$ at 1 mHz and 1.15 at 0.1 mHz. The importance in being able to distinguish between two backgrounds is not the absolute amplitude of the background, but the ratio between the two backgrounds' magnitudes $\frac{\Omega_{\mathrm{astro}}}{\Omega_{\mathrm{Cosmo}}}$. For a smaller ratio we can fit the cosmological background with
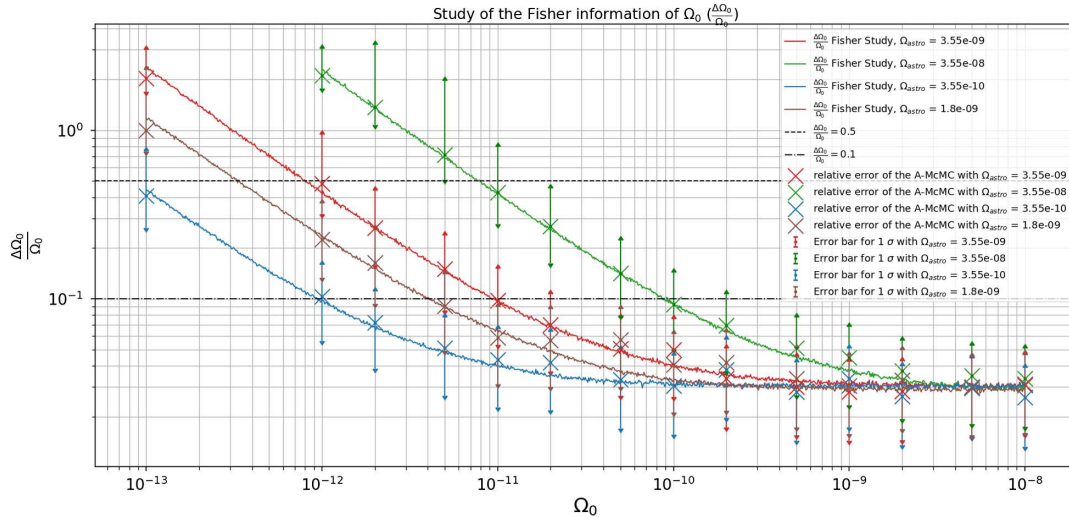

FIG. 11. Uncertainty of the estimation of the parameter Ω_0 (the spectral energy density of the cosmological SGWB) from the Fisher information study (displayed as lines), and the parametric estimation from the A-MCMC (displayed as scatter points) for the channel A and E, with the noise channel T. We conduct the study with different values for the astrophysical magnitude Ω_{astro}. There are error bars for the four sets of A-MCMC runs; see Eq. (32). The horizontal dashed line represents the error level of 50%. This is the limit where it is possible to observe the cosmological SGWB. The dot-dashed line represents the 10% error.
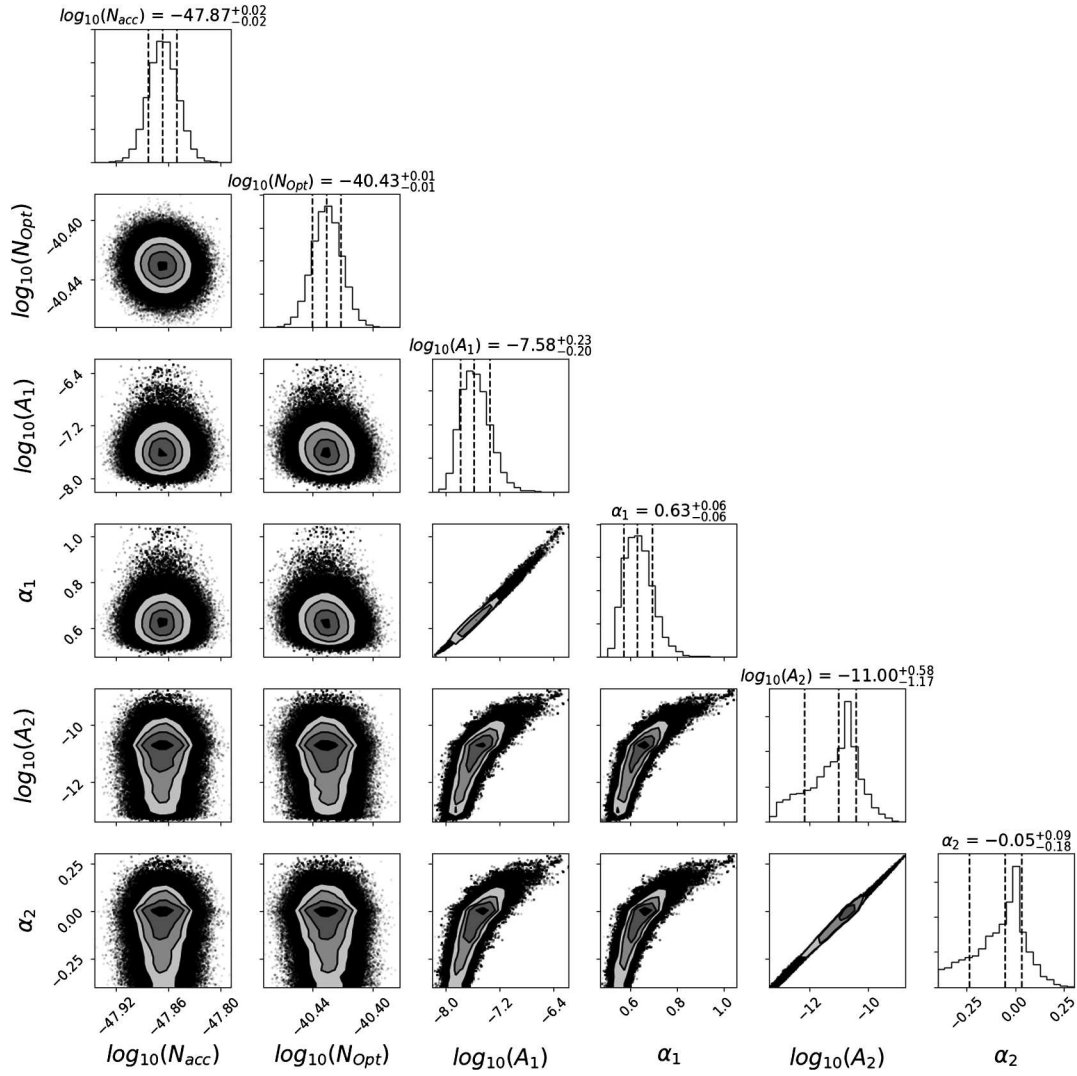
FIG. 12. Corner plot giving the A-MCMC-generated posterior distributions for a run with six parameters with $\Omega_{\mathrm{GW,Astro}} = 3.55 \times 10^{-8}$ and $\Omega_{\mathrm{GW,Cosmo}} = 1 \times 10^{-11}$. The vertical dashed lines on the posterior distributions represent from left to right the quantiles [16%, 50%, 84%]. This is from a run of using the data from channels A, E and T. These results are presented in Table I and also in Fig. 11.

less uncertainty. From Fig. 11, we can separate the cosmological background from the astrophysical background with a magnitude ratio of 4610 with $\Omega_{\mathrm{astro}} = 3.55 \times 10^{-9}$ and a reference frequency of 25 Hz. Here we have a fitting uncertainty of 50%, which is the limit for making a measurement. In fact, we can consider making a measurement of the cosmological background if the uncertainty is less than 50%; note the dashed line in Fig. 11. This example corresponds to a cosmological background of $\Omega_{\mathrm{Cosmo}} = 7.7 \times 10^{-13}$ In Fig. 11 the same study is presented with four values for the astrophysical background:

$\Omega_{\mathrm{astro}} = 3.55 \times 10^{-8}$, 3.55×10^{-9}, 1.8×10^{-9} and 3.55×10^{-10}. The same ratio produces similar results for different inputs of astrophysical amplitude. We obtain respectively the limits to constrain the cosmological background: $\Omega_{\mathrm{Cosmo}} = 7.8 \times 10^{-12}$, 7.8×10^{-13}, 3.6×10^{-13} and 7.6×10^{-14}. The values of these A-MCMC results are given in the Table I. Figures 12 and 13 present respective examples of corner plots and posterior distributions for a run of a six-parameter A-MCMC with $\Omega_{\mathrm{GW,Astro}} = 3.55 \times 10^{-8}$ and $\Omega_{\mathrm{GW,Cosmo}} = 1 \times 10^{-10}$, $\Omega_{\mathrm{GW,Astro}} = 3.55 \times 10^{-9}$ and $\Omega_{\mathrm{GW,Cosmo}} = 5 \times 10^{-12}$.
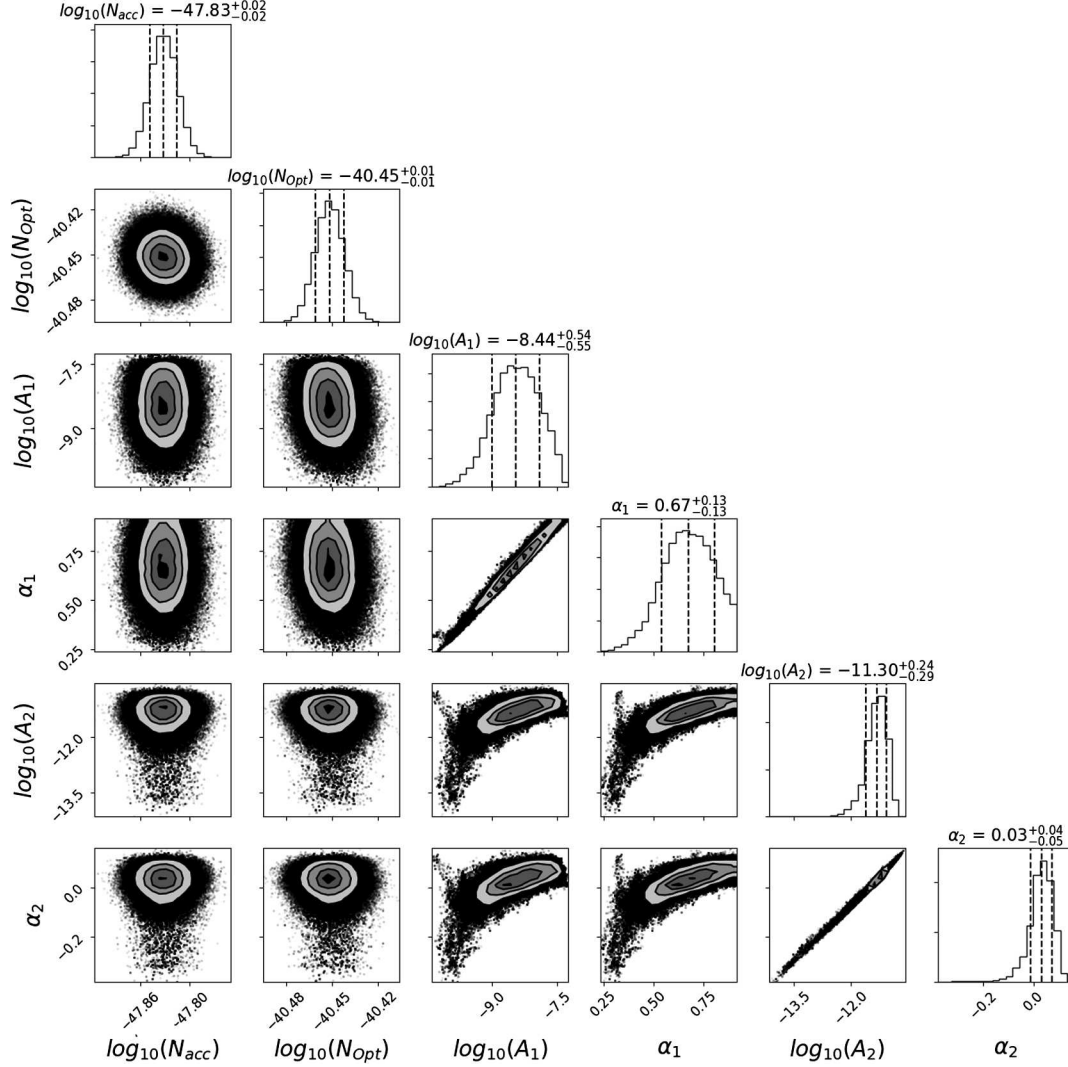
FIG. 13.   Corner plot giving the A-MCMC-generated posterior distributions for a run of six parameters, with $\Omega_{\rm GW,Astro} = 3.55 \times 10^{-9}$ and $\Omega_{\rm GW,Cosmo} = 5 \times 10^{-12}$. The vertical dashed lines on the posterior distribution represent from left to right the quantiles [16%, 50%, 84%]. This is from a run using the data from channels A, E and T. These results are presented in Table I, and also in Fig. 11.

VI. CONCLUSION

In this paper we presented the potential for separating the spectral components of the two SGWBs with an adaptive MCMC method. We also implemented a Fisher information study, predicting the measurement uncertainty from the A-MCMC analysis. The two independent studies produced consistent results. We obtained an uncertainty around 1 for the low level ($\Omega_0 = 1 \times 10^{-12}$) and around 0.03 for the high level ($\Omega_0 = 1 \times 10^{-8}$). For example, with an astrophysical background of $\Omega_{\rm GW,Astro} = 3.55 \times 10^{-9}(\frac{f}{25 \text{ Hz}})^{2/3}$ a cosmological background at $\Omega_{\rm GW,Cosmo} = 7.6 \times 10^{-13}$ can be detected. This corresponds to an uncertainty $\frac{\Delta\Omega_0}{\Omega_0}$ of 0.5 (dashed line in Fig. 11). The study presented in Sec. IV B displays the possibility to fit the parametric components of the SGWB.

In Sec. V we discussed and demonstrated the possibility to analyze the "noise" channel (the T channel) to fit the noise parameters of the LISA noise budget. The advantage of this method is that it increases the efficiency of the parameter estimates and utilize the total frequency domain $[1 \times 10^{-5} \text{ Hz}, 1 \text{ Hz}]$. We also applied a Fisher information study to the LISA noise. According to Fig. 11 we showed the possibility to separate the two SGWBs with a spectral separation with a factor of 4610 (for $f_{\rm ref} = 25$ Hz). Using a realistic range for the predicted magnitude of the astrophysically produced SGWB the methods demonstrated in this paper show that it is possible for LISA to also observe a cosmologically produced SGWB in the range of $\Omega_{\rm GW,Cosmo} \approx 1 \times 10^{-12}$ to 1×10^{-13}.

We note some limitations in this study and give some expectations for future work. In this paper we assumed no

difference in the noise levels on each spacecraft. According to Ref. [11] it is possible to include such a noise variation for each spacecraft. We could also include small modifications of the transfer functions R_I, and allow for some modification of the spectral slopes of the noise components. We can have a varying slope but with a narrow Gaussian prior centered on the theoretical value. It will be important to address more detailed models of both the LISA noise and the astrophysical and cosmological contributions to the stochastic background.

ACKNOWLEDGMENTS

APPENDIX: SIGNAL SEPARATION LITERATURE SUMMARY

In this Appendix we present in tabular form a list of various studies that have been conducted in order to separate different SGWBs and detector noise sources. Much has been published on this subject.

1. SGWB studies for LIGO/Virgo

Table II presents a summary of the literature addressing SGWB signal separation for LIGO and Virgo.

TABLE II.   Methods to measure and to separate SGWBs for LIGO/Virgo.

Reference	Goal	Method	Performance	Limitations and applications
Chen *et al.* [20]	Astrophysical SGWB from binary black holes and binary neutrons stars	Estimation of the SGWB from LIGO/Virgo observations; local merger rate $\mathcal{R}$	$\Omega_{2/3} = 4.4^{+6.3}_{-3.0} \times 10^{-12}$, $f_{\text{ref}} = 3$ mHz)	The error on the local merger rate is important
Abbott *et al.* [19]	Astrophysical SGWB from binary black holes and binary neutrons stars	Estimation of the SGWB from LIGO/Virgo observations with the local merger rate $\mathcal{R}$ estimation from GW150914	$\Omega_{2/3} = 1.1^{+2.7}_{-0.9} \times 10^{-12}$ $f_{\text{ref}} = 25$ Hz	The error on the local merge rate is important
Abbott *et al.* [27,28]	Three backgrounds considered, power laws $\alpha = 0, \frac{2}{3}, 3$	Results from cross-correlation analysis with Advanced LIGO O3 combined O1 and O1 results	$\Omega_0 < 5.8 \times 10^{-9}$; $\Omega_{2/3} < 3.4 \times 10^{-9}$, $f_{\text{ref}} = 25$ Hz	No correlated noise due to the magnetic Schumann resonances
Parida *et al.* [30]	Separate different isotropic SGWBs for LIGO	Component separation of power laws avoiding use of MCMC methods	Simulation demonstration for Advanced LIGO target sensitivity: $\Omega_0 = (1 \pm 0.676) \times 10^{-8}$; $\Omega_{2/3} = (1 \pm 1.719) \times 10^{-8}$; $\Omega_3 = (1 \pm 3.284) \times 10^{-8}$; $f_{\text{ref}} = 100$ Hz.	Requires a negligible amount of computation and would be simple to apply to real data
Ungarelli and Vecchio [31]	Fit broken power-law SGWB with data from Earth-based detectors	Filters based on broken power-law spectra	Achieved fitting factor greater than 97%	Small number of filters needed to measure SGWB in the first-generation laser interferometers
Smith and Thrane [33]	To detect astrophysical SGWB with LIGO/Virgo	Bayesian parameter estimation to detect unresolved binary black hole background	Less data needed to observe background, as opposed to traditional correlation based search	Gives a unified method for a search for resolvable signals and a SGWB of unresolvable signals

(Table continued)

TABLE II. *(Continued)*

Reference	Goal	Method	Performance	Limitations and applications
S. Biscoveanu *et al.* [32]	To detect a primordial SGWB in the presence of unresoved binary black holes in LIGO/ Virgo band	Use method of [33]; individual short time segments analyzed	Measurement of a simulated power law: $\log \Omega_\alpha = -5.96^{+0.08}_{-0.16}$, $\alpha = 0.49^{+1.14}_{-0.49}$	Limitations from the precision of the compact binary signal waveforms, and non-Gaussian noise
E. Thrane *et al.* [38]	SGWB measurement in the context of correlated magnetic noise in LIGO/Virgo band.	Correlated noise between detectors creates a systematic error in cross correlation study	Measurement of the correlated noise from the Schumann resonances	Possibility to use Wiener Filter to subtract the correlation.
P. M. Meyers *et al.* [41]	LIGO/Virgo SGWB measurement in the context of correlated magnetic noise	Parameter estimation of the correlated magnetic noise and SGWB	Demonstration with $\Omega_{2/3} \simeq 3 \times 10^{-9}$, $f_{\rm ref} = 25$ Hz and realistic magnetic coupling in LIGO/Virgo	An alternative to Wiener filtering

2. SGWB studies in the LISA band

Table III presents a summary of the literature addressing SGWB signal separation for LISA.

TABLE III.   Methods to measure and to separate SGWBs for LISA.

Reference	Goal	Method	Perfomance	Limitations and applications
Cornish and Larson [25]	Observe cosmic SGWB with astrophysical foregrounds	Strategies for individual, or two LISA interferometers using cross-correlation	LISA could detect a cosmic SGWB at the level of $\Omega_{\rm GW}(f)h_0^2 > 7 \times 10^{-12}$	The LISA sensibility is derived for LISA arm length of $L = 5 \times 10^9$
Pieroni and Barauss [29]	Extraction of the cosmological SGWB and astrophysical foreground with LISA noise	Principal component analysis to model and extract SGWBs	LISA can measure a cosmological SGWB of $\Omega_0 = 6 \times 10^{-13}$ with SNR = 31	A robust technique that can be extended to different detectors
Caprini *et al.* [43]	Observe SGWBs with LISA	Reconstruction of SGWB as a function of frequency for simple and broken power-laws	Detects a power law of $\Omega_{2/3} = 5.4 \times 10^{-12}$, $f_{\rm ref} = 0.001$ Hz, with SNR = 601.	Signal and noise are assumed to be stationary for all times.
Flauger *et al.* [44]	Observe SGWBs with LISA, building on the work of [43]	Reconstruction of the spectral shape of a SGWB with the LISA A, E, T channels	Improvement of $\sqrt{2}$ over the method of [43]	Will be expanded to account for unequal arm lengths for LISA constellation.
Karnesis *et al.* [45]	Fast methodology to assess LISA detectability of a stationary, Gaussian, and isotropic SGWB	Testing the *Radler* simulated data set from the LISA Data Challenge	Successful demonstration for $\Omega_{2/3}(f) = 3.6 \times 10^{-9}(\frac{f}{25\,{\rm Hz}})^{2/3}$	Analysis done with simple LISA noise model

3. SGWB studies for the future third-generation detectors

Table IV presents a summary of the literature addressing SGWB signal separation for third-generation gravitational-wave detectors.

TABLE IV. Methods to measure and to separate SGWBs for the third-generation detectors.

Reference	Goal	Method	Perfomance	Limitation and application
Regimbau *et al.* [34]	Observing a primordial SGWB below the compact binary produced background	The data will be cleaned of the direct observations of binaries by the third-generation detectors	Possible limit of $\Omega_{GW} \simeq 10^{-13}$ after 5 years of observation with third-generation detectors [35,36]	Potential limitation to sensitivity comes from other astrophysical gravitational-wave emission.
Sharma and Harms [37]	Cosmological SGWB with third-generation detectors in the presence of an astrophysical foreground	Matched filtering and residual study for the astrophysical foreground and cross-correlation for cosmological SGWB	Cosmological SGWB (flat) $\Omega_{GW} = 2 \times 10^{-12}$ observed with SNR ≈ 5.2 after 1.3 years	Limitation for cosmological SGWB is instrumental noise and unremoved astrophysical sources
Martinovic *et al.* [42]	Astrophysical (compact binary coalescence) and cosmological SGWB (cosmic strings and first order phase transitions)	Bayesian parameter estimation for simultaneous estimation of astrophysical and cosmological SGWB with third-generation detectors	Possible limit at 25 Hz of $\Omega_{GW} = 2.2 \times 10^{-13}$ (broken power-law model for primordial SGWB) and $\Omega_{GW} = 4.5. \times 10^{-13}$ for cosmic strings	Methods will be applicable for LISA

[1] B. Abbott *et al.*, Observation of Gravitational Waves from a Binary Black Hole Merger, Phys. Rev. Lett. **116**, 061102 (2016).

[2] J. Aasi *et al.*, Advanced LIGO, Classical Quantum Gravity **32**, 074001 (2015).

[3] G. M. Harry, Advanced LIGO: The next generation of gravitational wave detectors, Classical Quantum Gravity **27**, 084006 (2010).

[4] B. P. Abbott *et al.*, GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence, Phys. Rev. Lett. **119**, 141101 (2017).

[5] F. Acernese *et al.*, Advanced Virgo: A second-generation interferometric gravitational wave detector, Classical Quantum Gravity **32**, 024001 (2015).

[6] B. P. Abbott *et al.*, GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs, Phys. Rev. X **9**, 031040 (2019).

[7] R. Abbott *et al.*, GWTC-2: Compact binary coalescences observed by LIGO and Virgo during the first half of the third observing run, arXiv:2010.14527 [Phys. Rev. X (to be published)].

[8] B. P. Abbott *et al.*, GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral, Phys. Rev. Lett. **119**, 161101 (2017).

[9] B. P. Abbott *et al.*, GW190425: Observation of a compact binary coalescence with total mass $\sim 3.4 \, M_\odot$, Astrophys. J. Lett. **892**, L3 (2020).

[10] P. Amaro-Seoane *et al.*, Laser interferometer space antenna, arXiv:1702.00786.

[11] M. R. Adams and N. J. Cornish, Detecting a stochastic gravitational wave background in the presence of a galactic foreground and instrument noise, Phys. Rev. D **89**, 022001 (2014).

[12] N. J. Cornish and T. B. Littenberg, Tests of bayesian model selection techniques for gravitational wave astronomy, Phys. Rev. D **76**, 083006 (2007).

[13] A. Lamberts, S. Blunt, T. B. Littenberg, S. Garrison-Kimmel, T. Kupfer, and R. E. Sanderson, Predicting the LISA white dwarf binary population in the Milky Way with cosmological simulations, Mon. Not. R. Astron. Soc. **490**, 5888 (2019).

[14] K. B. Burdge *et al.*, General relativistic orbital decay in a seven-minute-orbital-period eclipsing binary system, Nature (London) **571**, 528 (2019).

[15] K. B. Burdge *et al.*, A systematic search of Zwicky transient facility data for ultracompact binary LISA-detectable gravitational-wave sources, Astrophys. J. **905**, 32 (2020).

[16] V. Korol, S. Toonen, A. Klein, V. Belokurov, F. Vincenzo, R. Buscicchio, D. Gerosa, C. J. Moore, E. Roebber,

E. M. Rossi *et al.*, Populations of double white dwarfs in Milky Way satellites and their detectability with LISA, Astron. Astrophys. **638,** A153 (2020).

[17] N. Christensen, Stochastic gravitational wave backgrounds, Rep. Prog. Phys. **82,** 016903 (2019).

[18] J. D. Romano and N. J. Cornish, Detection methods for stochastic gravitational-wave backgrounds: A unified treatment, Living Rev. Relativity **20,** 2 (2017).

[19] B. Abbott *et al.*, Gw150914: Implications for the Stochastic Gravitational-Wave Background from Binary Black Holes, Phys. Rev. Lett. **116,** 131102 (2016).

[20] Z.-C. Chen, F. Huang, and Q.-G. Huang, Stochastic gravitational-wave background from binary black holes and binary neutron stars and implications for LISA, Astrophys. J. **871,** 97 (2019).

[21] J. Garcia-Bellido and D. G. Figueroa, A stochastic Background of Gravitational Waves from Hybrid Preheating, Phys. Rev. Lett. **98,** 061302 (2007).

[22] L. E. Mendes, A. B. Henriques, and R. G. Moorhouse, Exact calculation of the energy density of cosmological gravitational waves, Phys. Rev. D **52,** 2083 (1995).

[23] P. Campeti, E. Komatsu, D. Poletti, and C. Baccigalupi, Measuring the spectrum of primordial gravitational waves with CMB, PTA and laser interferometers, J. Cosmol. Astropart. Phys. 01 (2021) 012.

[24] C.-F. Chang and Y. Cui, Stochastic gravitational wave background from global cosmic strings, Phys. Dark Universe **29,** 100604 (2020).

[25] N. J. Cornish and S. L. Larson, Space missions to detect the cosmic gravitational-wave background, Classical Quantum Gravity **18,** 3473 (2001).

[26] N. Christensen and R. Meyer, Markov chain Monte Carlo methods for Bayesian gravitational radiation data analysis, Phys. Rev. D **58,** 082001 (1998).

[27] B. Abbott *et al.*, Search for the isotropic stochastic background using data from Advanced LIGO's second observing run, Phys. Rev. D **100,** 061101 (2019).

[28] R. Abbott *et al.*, Upper limits on the isotropic gravitational-wave background from Advanced LIGO's and Advanced Virgo's third observing run, arXiv:2101.12130 [Phys. Rev. D (to be published)].

[29] M. Pieroni and E. Barausse, Foreground cleaning and template-free stochastic background extraction for LISA, J. Cosmol. Astropart. Phys. 07 (2020) 021.

[30] A. Parida, S. Mitra, and S. Jhingan, Component separation of a isotropic gravitational wave background, J. Cosmol. Astropart. Phys. 04 (2016) 024.

[31] C. Ungarelli and A. Vecchio, A family of filters to search for frequency-dependent gravitational wave stochastic backgrounds, Classical Quantum Gravity **21,** S857 (2004).

[32] S. Biscoveanu, C. Talbot, E. Thrane, and R. Smith, Measuring the Primordial Gravitational-Wave Background in the Presence of Astrophysical Foregrounds, Phys. Rev. Lett. **125,** 241101 (2020).

[33] R. Smith and E. Thrane, Optimal Search for an Astrophysical Gravitational-Wave Background, Phys. Rev. X **8,** 021019 (2018).

[34] T. Regimbau, M. Evans, N. Christensen, E. Katsavounidis, B. Sathyaprakash, and S. Vitale, Digging Deeper: Observing Primordial Gravitational Waves below the Binary-Black-Hole-Produced Stochastic Background, Phys. Rev. Lett. **118,** 151105 (2017).

[35] M. Punturo *et al.*, The Einstein telescope: A third-generation gravitational wave observatory, Classical Quantum Gravity **27,** 194002 (2010).

[36] D. Reitze *et al.*, Cosmic explorer: The U.S. contribution to gravitational-wave astronomy beyond LIGO, Bull. Am. Astron. Soc. **51,** 035 (2019), arXiv:1907.04833.

[37] A. Sharma and J. Harms, Searching for cosmological gravitational-wave backgrounds with third-generation detectors in the presence of an astrophysical foreground, Phys. Rev. D **102,** 063009 (2020).

[38] E. Thrane, N. Christensen, and R. Schofield, Correlated magnetic noise in global networks of gravitational-wave interferometers: Observations and implications, Phys. Rev. D **87,** 123009 (2013).

[39] D. D. Sentman, *Handbook of Atmospheric Electrodynamics* (CRC Press, Boca Raton, 1995), Vol. 1.

[40] M. Füllekrug, Schumann resonances in magnetic field components, J. Atmos. Terr. Phys. **57,** 479 (1995).

[41] P. M. Meyers, K. Martinovic, N. Christensen, and M. Sakellariadou, Detecting a stochastic gravitational-wave background in the presence of correlated magnetic noise, Phys. Rev. D **102,** 102005 (2020).

[42] K. Martinovic, P. M. Meyers, M. Sakellariadou, and N. Christensen, Simultaneous estimation of astrophysical and cosmological stochastic gravitational-wave backgrounds with terrestrial detectors, Phys. Rev. D **103,** 043023 (2021).

[43] C. Caprini, D. G. Figueroa, R. Flauger, G. Nardini, M. Peloso, M. Pieroni, A. Ricciardone, and G. Tasinato, Reconstructing the spectral shape of a stochastic gravitational wave background with LISA, J. Cosmol. Astropart. Phys. 11 (2019) 017.

[44] R. Flauger, N. Karnesis, G. Nardini, M. Pieroni, A. Ricciardone, and J. Torrado, Improved reconstruction of a stochastic gravitational wave background with LISA, J. Cosmol. Astropart. Phys. 01 (**2021**) 059.

[45] N. Karnesis, M. Lilley, and A. Petiteau, Assessing the detectability of a stochastic gravitational wave background with LISA, using an excess of power approach, Classical Quantum Gravity **37,** 215017 (2020).

[46] J. B. Camp and N. J. Cornish, Gravitational wave astronomy, Annu. Rev. Nucl. Part. Sci. **54,** 525 (2004).

[47] A. J. Farmer and E. Phinney, The gravitational wave background from cosmological compact binaries, Mon. Not. R. Astron. Soc. **346,** 1197 (2003).

[48] M. R. Adams and N. J. Cornish, Discriminating between a stochastic gravitational wave background and instrument noise, Phys. Rev. D **82,** (2010).

[49] C. Périgois, C. Belczynski, T. Bulik, and T. Regimbau, StarTrack predictions of the stochastic gravitational-wave background from compact binary mergers, Phys. Rev. D **103,** 043002 (2021).

[50] T. A. Prince, M. Tinto, S. L. Larson, and J. W. Armstrong, LISA optimal sensitivity, Phys. Rev. D **66,** 122002 (2002).

[51] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, 3rd ed. (Chapman & Hall, New York, NY, USA, 2014).

[52] G. O. Roberts and J. S. Rosenthal, Examples of adaptive MCMC, J. Comput. Graph. Stat. **18**, 349 (2009).

[53] S. Babak and A. Petiteau, LISA data challenge manual LISA-LCST-SGD-MAN-001, LISA Consortium, 2018.

[54] ESA, LISA science study team LISA science requirement document ESA-L3-EST-SCI-RS-001_LISA_SciRD, LISA Consortium, 2018.

[55] T. L. Smith and R. R. Caldwell, LISA for cosmologists: Calculating the signal-to-noise ratio for stochastic and deterministic sources, Phys. Rev. D **100**, 104055 (2019).

[56] LIGO Scientific and the Virgo Collaborations, A guide to LIGO-Virgo detector noise and extraction of transient gravitational-wave signals, Classical Quantum Gravity **37**, 055002 (2020).