# A stochastic search for intermittent gravitational-wave backgrounds

Jessica Lawrence[1,*] Kevin Turbang[2,3,†] Andrew Matas,[4] Arianna I. Renzini[5,‡]
Nick van Remortel[3,§] and Joseph Romano[1,∥]

[1]*Department of Physics, Texas Tech University, Lubbock, Texas 79409, USA*
[2]*Theoretische Natuurkunde, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium*
[3]*Universiteit Antwerpen, Prinsstraat 13, B-2000 Antwerpen, Belgium*
[4]*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), D-14476 Potsdam, Germany*
[5]*Department of Physics, California Institute of Technology, Pasadena, California 91125, USA*

A likely source of a gravitational-wave background (GWB) in the frequency band of the Advanced LIGO, Virgo, and KAGRA detectors is the superposition of signals from the population of unresolvable stellar-mass binary-black-hole (BBH) mergers throughout the Universe. Since the duration of a BBH merger in band ($\sim 1$ s) is much shorter than the expected separation between neighboring mergers ($\sim 10^3$ s), the observed signal will be "popcornlike" or intermittent with duty cycles of order $10^{-3}$. However, the standard cross-correlation search for stochastic GWBs currently performed by the LIGO-Virgo-KAGRA Collaboration is based on a continuous-Gaussian signal model, which does not take into account the intermittent nature of the background. The latter is better described by a Gaussian mixture model, which includes a duty cycle parameter that quantifies the degree of intermittence. Building on an earlier paper by Drasco and Flanagan [Detection methods for non-gaussian gravitational wave stochastic backgrounds, Phys. Rev. D **67**, 082003 (2003).], we propose a stochastic-signal-based search for intermittent GWBs. For such signals, this search performs better than the standard continuous cross-correlation search. We present results of our stochastic-signal-based approach for intermittent GWBs applied to simulated data for some simple models, and we compare its performance to the other search methods, in terms of both detection and signal characterization. Additional testing on more realistic simulated datasets, e.g., consisting of astrophysically motivated BBH merger signals injected into colored detector noise containing noise transients, will be needed before this method can be applied with confidence on real gravitational-wave data.

## I. INTRODUCTION

The Advanced LIGO [1], Virgo [2], and KAGRA [3] (LVK) detectors have completed their third observing run (O3), increasing the number of confident detections of gravitational-wave (GW) signals to 90 overall [4]. The detected signals are primarily associated with stellar-mass binary-black-hole (BBH) mergers, although a handful of binary-neutron-star (BNS) and neutron star-black hole (NSBH) coalescences have also been observed [5,6]. All of these signals are relatively large signal-to-noise ratio (SNR) events, which stand out above the detector noise when searched for using matched-filtering techniques [7,8].

In addition to these loud, individually resolvable events, the LVK detectors are also being showered by GW signals produced by much weaker (e.g., more distant and/or less massive) sources, whose combined effect gives rise to a low-level background of gravitational radiation—a so-called gravitational-wave background (GWB) (see, e.g., [9,10] and references cited within). This background signal is expected to be stochastic (i.e., random) in the sense that there is no single deterministic waveform that we can use to perform a matched-filter search for this type of GW signal. Nonetheless, because this signal is present in all detectors, we can cross-correlate the data from multiple detectors to observe the GWB, despite its weakness relative to the noise [11,16]. Although to date there has not been a direct detection of a GWB using a stochastic pipeline, we know from Advanced LIGO's and Virgo's detections of individual resolvable sources that a background arising from compact binary mergers must exist. Assuming our detectors are upgraded as planned in the coming years [12], and given current projections for the signal [13], detecting

[*]jessica.lawrence@ttu.edu
[†]kevin.turbang@vub.be
[‡]arenzini@caltech.edu
[§]nick.vanremortel@uantwerpen.be
[∥]joseph.d.romano@ttu.edu

the GWB may just be a matter of time. On the other hand, we can improve our detection methods to measure this signal sooner. We assume the latter strategy in this paper.

## A. Motivation

A likely source of a GWB in the frequency band of the LVK detectors is the population of stellar-mass BBH mergers throughout the Universe. Rate estimates calculated from the BBH signals detected to date [13,14] predict a BBH merger in the observable universe every ∼5–10 minutes on average. Since the duration of a BBH merger in the LVK band is of order 1 s, the duty cycle $\xi$ of such events (defined as the time in band for one merger signal divided by the average time between successive mergers) is of order $10^{-3}$. Thus, the expected GWB signal is "popcornlike" or *intermittent*, with the signal being "on" a small fraction of the total observation time. A similar calculation for the population of BNS mergers predicts (on average) roughly one event every 15 s, while the duration of a BNS signal in band is approximately 100 s. Thus, BNS merger signals overlap in time leading to a continuous (and possibly confusion-limited) background.

The total expected BBH signal is potentially detectable with the Advanced LIGO and Virgo detectors when observing at design sensitivity [13,15]. Although the SNRs for the individual events are small, the combined SNR of the correlated data summed over all events grows like the square root of the observation time, reaching a detectable level of $3\sigma$ (corresponding to a false alarm probability of approximately $10^{-3}$) after ∼40 months of observation [15]. This estimate of the time to detection is based on the standard cross-correlation search [16], which looks for evidence of excess cross-correlated signal power, assuming that the amplitude of the GW signal component is drawn from a continuous-Gaussian distribution. This search assumes that the signal is on all the time, in conflict with the intermittent nature of the stellar-mass BBH background, which is expected to be the dominant signal. Thus, although the standard cross-correlation search is able to detect the time-averaged signal from an intermittent GWB [17], this search is suboptimal in the sense that the time to detection will be longer than that for a search that properly takes into account the intermittent nature of the background.

## B. Purpose and outline

The purpose of this paper is to introduce a new stochastic-signal-based search that specifically targets intermittent GWBs, and hence can potentially reduce the time to detection of the BBH background signal. This new search is built on the seminal work of Drasco and Flanagan [18], who proposed a Gaussian mixture-model (GMM) likelihood function for analyzing intermittent GWBs (Sec. II A). Our proposed search for intermittent

GWBs looks for excess cross-correlated power in short stretches of data. Conversely, a deterministic-signal-based search for the intermittent BBH background was proposed by Smith and Thrane [19], which involves marginalizing over the signal parameters for deterministic BBH chirp waveforms in short (∼4 s) stretches of data (Sec. II B). By construction, our proposed search is adaptable to a generic intermittent GWB since it looks only for excess cross-correlated power. We also expect our proposed search to be computationally more efficient in detecting a signal than the deterministic-signal-based approach of Smith and Thrane, since our search ignores the deterministic form of the GW signal waveforms and hence the need to marginalize over all the associated signal parameters.

A brief outline of this paper is as follows: first, we give an overview of the current searches for intermittent GWBs in Sec. II. We then proceed by introducing our proposed stochastic search for intermittent GWBs in Sec. III. To compare the performance of the various search methods mentioned above, we analyze a series of datasets that are tailored to highlight the merits and shortcomings of each style of search. We start in Sec. IV A by considering stationary-Gaussian white noise in two colocated and coaligned detectors, and we inject an intermittent GWB made up of white GW bursts with Gaussian signal amplitudes scaled by distances to the sources drawn from a uniform-in-volume distribution. We then consider a background made up of colored GW bursts[1] in Sec. IV B, which follow the expected spectral shape of BBH mergers. Finally, we analyze a set of deterministic BBH chirp waveforms in Sec. IV C, where the chirp parameters are fixed except for the distance to the source, which is also drawn from a uniform-in-volume distribution. We conclude in Sec. V by discussing possible extensions of our method and additional tests that are needed on more realistic simulated data before it can be run on real LVK data.

## II. PROPOSED SEARCHES FOR INTERMITTENT GWBs—OVERVIEW

The standard continuous cross-correlation search [16] aims to measure the fractional energy density of a GWB, defined as

$$\Omega_{\mathrm{gw}}(f) = \frac{1}{\rho_c} \frac{\mathrm{d}\rho_{\mathrm{gw}}}{\mathrm{d}\ln f}, \qquad (1)$$

where the critical energy density of the Universe is $\rho_c = 3H_0^2 c^2/(8\pi G)$, $H_0$ is the Hubble constant, $c$ is the speed of light, and $G$ is Newton's constant. Alternatively, a

---

[1]The term "burst" will be used throughout this paper as it is the most general, irrespective of the type of signal. In the context of compact binary mergers, these bursts of GWs are often referred to as "transients."

GWB can be characterized by its power spectral density (PSD) $P_{\mathrm{gw}}(f)$, which is related to $\Omega_{\mathrm{gw}}(f)$ by [16]

$$\Omega_{\mathrm{gw}}(f) = \frac{10\pi^2}{3H_0^2} f^3 P_{\mathrm{gw}}(f). \tag{2}$$

For the target signal of a BBH GWB, it is well known that the fractional energy density spectrum is $\Omega_{\mathrm{gw}}(f) \propto f^{2/3}$ to good approximation [20], in the frequency ranges probed by the LVK interferometers. This knowledge can be incorporated into the search, reducing it to the measurement of a single quantity $\Omega_{\mathrm{gw}}(f_{\mathrm{ref}})$, where $f_{\mathrm{ref}}$ is a reference frequency chosen where the sensitivity of the LVK detectors is best (typically 25 Hz) [21]. For the remainder of the paper, we will refer to $\Omega_{\mathrm{gw}}(f_{\mathrm{ref}})$ simply as $\Omega_{\mathrm{gw}}$ for brevity. For a set of data containing enough events to be statistically significant, $\Omega_{\mathrm{gw}}$ is the amplitude of the time and population-averaged energy density. We will refer to this stochastic search for continuous backgrounds described above as SSC.

Since this search assumes a continuous-in-time signal in the data, it does not properly model an important feature of the BBH GWB signal—the intermittency. To remedy this improper modeling, several searches targeting intermittent GWBs specifically have been proposed. We start by giving a high-level overview of these different analysis methods. We refrain from giving details about the actual form of the likelihoods and refer to Appendix for more information.

### A. Gaussian mixture-model likelihood function for intermittent GWBs

In 2003, Drasco and Flanagan [18] proposed a search for an intermittent GWB that makes use of a GMM likelihood function of the form

$$\mathcal{L}_{\mathrm{tot}} = \prod_{I}^{N_{\mathrm{seg}}} [\xi \mathcal{L}_{s,I} + (1-\xi)\mathcal{L}_{n,I}], \tag{3}$$

where $\xi$ is the probability that a particular segment contains a GW signal, and $\mathcal{L}_{s,I}$ and $\mathcal{L}_{n,I}$ are the likelihood functions for segment $I$ in the presence and absence of a GW signal, i.e., the signal and noise likelihoods. For the simple toy model considered in their paper (i.e., single-sample GW "bursts," occurring with probability $\xi$ drawn from a fixed Gaussian distribution with variance $\sigma_b^2$, and injected into uncorrelated white noise in two colocated and coaligned detectors), the signal and noise parameters that enter the likelihood functions $\mathcal{L}_{s,I}$ and $\mathcal{L}_{n,I}$ are the variances $(\sigma_b^2, \sigma_{n_1}^2, \sigma_{n_2}^2)$ and $(\sigma_{n_1}^2, \sigma_{n_2}^2)$, respectively. Single-sample bursts are bursts whose duration is less than the sample period $\Delta t$. By maximizing $\mathcal{L}_{\mathrm{tot}}$ with respect to all four parameters $(\xi, \sigma_b^2, \sigma_{n_1}^2, \sigma_{n_2}^2)$, Drasco and Flanagan obtained a detection statistic (the maximum-likelihood statistic), which they could use to search for intermittent GWBs. Note that in the case $\xi = 1$, i.e., assuming the signal is always present, one recovers the standard continuous-Gaussian search introduced above.

Although Drasco and Flanagan tested their proposed method with a test statistic within a frequentist framework, we have decided to work within a Bayesian framework in this paper. We define several concepts of importance within this framework before moving on to the discussion of the results of Drasco and Flanagan.

Given a likelihood function $\mathcal{L}_{\mathrm{tot}}$ and priors $\pi$, the joint posterior distribution for the duty cycle and the signal + noise parameters can be computed using Bayes' theorem:

$$p(\xi, \sigma_b^2, \sigma_{n_1}^2, \sigma_{n_2}^2 | d)$$
$$= \frac{\mathcal{L}_{\mathrm{tot}}(d|\xi, \sigma_b^2, \sigma_{n_1}^2, \sigma_{n_2}^2)\pi(\xi)\pi(\sigma_b^2)\pi(\sigma_{n_1}^2)\pi(\sigma_{n_2}^2)}{\mathcal{Z}(d)}, \tag{4}$$

where

$$\mathcal{Z}(d) \equiv \int \mathrm{d}\xi \int \mathrm{d}\sigma_b^2 \int \mathrm{d}\sigma_{n_1}^2 \int \mathrm{d}\sigma_{n_2}^2 \mathcal{L}_{\mathrm{tot}}(d|\xi, \sigma_b^2, \sigma_{n_1}^2, \sigma_{n_2}^2)\pi(\xi)\pi(\sigma_b^2)\pi(\sigma_{n_1}^2)\pi(\sigma_{n_2}^2) \tag{5}$$

is the model evidence. Marginalized posterior distributions (for each parameter separately) are obtained by integrating the joint posterior distribution over all the other parameters, e.g.,

$$p(\xi) = \int \mathrm{d}\sigma_b^2 \int \mathrm{d}\sigma_{n_1}^2 \int \mathrm{d}\sigma_{n_2}^2 p(\xi, \sigma_b^2, \sigma_{n_1}^2, \sigma_{n_2}^2). \tag{6}$$

Of course, likelihood functions, priors, etc., are all calculated in the context of a particular choice of analysis model

$\mathcal{M}_\alpha$ (e.g., a GMM likelihood search for intermittent GWBs or the standard continuous-Gaussian search), which we have not indicated in the above expressions. If we explicitly denote the dependence of the above distributions on the choice of analysis model, we can define the Bayes factor between models $\mathcal{M}_\alpha$ and $\mathcal{M}_\beta$ as

$$\mathcal{B}_{\alpha\beta}(d) \equiv \frac{\mathcal{Z}(d|\mathcal{M}_\alpha)}{\mathcal{Z}(d|\mathcal{M}_\beta)}. \tag{7}$$

Assuming equal prior odds for the two models, the Bayes factor tells us how much more the data favors model $\mathcal{M}_\alpha$ relative to $\mathcal{M}_\beta$. Throughout this paper, we will make plots of the natural logarithm of the Bayes factor as a function of the duty cycle to compare the various search methods.

With these concepts in mind, we now move to the discussion of the results of the proposed GMM likelihood. Drasco and Flanagan showed that their detection statistic for intermittent GWBs performs better than the standard cross-correlation statistic for continuous-Gaussian backgrounds when the duty cycle $\xi$ is sufficiently small. To illustrate this, we implement their proposed GMM likelihood in a Bayesian framework. Instead of using their proposed frequentist detection statistic, we use the Bayes factor as a measure of efficiency. To be able to study its behavior as a function of the duty cycle, we combine 100 data realizations for each $\xi$ value. Each data realization consists of 40,000 data points (i.e., 40,000 segments, each 1 sample long), where a fraction of them contains single-sample bursts drawn from a Gaussian distribution with variance $\sigma_b^2 = 1$.

We keep the total continuous-Gaussian signal-to-noise ratio fixed to 3, computed using (8) and (10), by adjusting the noise variances for each value of the duty cycle, rather than adjusting the signal parameters. So, as $\xi$ decreases, the segment signal-to-noise ratios must increase, which means that the noise variances must decrease. This is illustrated in Fig. 1, where both the continuous-in-time, i.e., $\xi = 1$ in (3), and the intermittent GMM likelihood analysis methods are used. Each plotted point corresponds to the mean of the ln Bayes factor over 100 realizations of data, while the error



FIG. 1. ln Bayes factors of the signal + noise model to the noise-only model as a function of the duty cycle $\xi$ for the intermittent search (blue) and the continuous search (orange) where the signal consists of single sample bursts drawn from a Gaussian distribution of variance $\sigma_b^2$.

bars correspond to the standard deviation of the ln Bayes factor.

While the continuous search performs equally well for all duty cycles (since it assumes $\xi = 1$), the Bayes factor for the GMM likelihood increases as $\xi$ decreases, exceeding the continuous stochastic likelihood Bayes factor, illustrating that the GMM likelihood performs better than the continuous likelihood for smaller values of $\xi$. Equivalently, the relative performance of the Bayes factors shown in Fig. 1 can be expressed in terms of

$$\rho_{\text{seg}} \equiv \frac{\sigma_b^2}{\sigma_{n_1}\sigma_{n_2}}, \tag{8}$$

which is the expected signal-to-noise ratio in an individual segment assuming the presence of a GW signal with burst variance $\sigma_b^2$. In terms of $\rho_{\text{seg}}$, the condition for the GMM likelihood to perform better than the continuous likelihood is

$$\rho_{\text{seg}} \sim 1. \tag{9}$$

In the limit where $\rho_{\text{seg}} \ll 1$, the GW signals in an individual segment are sufficiently weak that the GMM likelihood does not perform any better than the standard stochastic continuous likelihood. Conversely, when $\rho_{\text{seg}} \gg 1$, the GW signals in the individual segments are so strong that they are individually resolvable, with segment signal-to-noise ratios exceeding the threshold needed for detection with a single-detector burst statistic. In other words, a search for an intermittent GWB is the most sensitive search when the GW signals in the individual segments are marginally subthreshold ($\rho_{\text{seg}} \sim 1$).

Furthermore, we can determine an approximate value of $\rho_{\text{seg}}$ for the LVK detectors, for the population of stellar-mass BBH mergers throughout the Universe. As mentioned in Sec. I, it should take $\sim 40$ months of observation using the standard continuous-Gaussian cross-correlation statistic to observe the BBH background with a total signal-to-noise ratio $\rho_{\text{tot}} = 3$ [22]. Since the segment duration proposed by Smith and Thrane [19] for an intermittent search is of order $T_{\text{seg}} \sim 4$ s (see Sec. II B for more details), 40 months of observation corresponds to $N_{\text{seg}} \sim 2.5 \times 10^7$ segments. The final input that we need to do the calculation is the expected duty cycle of the signal, which for stellar-mass BBH mergers throughout the Universe is $\xi \sim 10^{-3}$. These values imply

$$\rho_{\text{seg}} = \frac{\rho_{\text{tot}}}{\xi\sqrt{N_{\text{seg}}}} \sim 0.6, \tag{10}$$

which is in the regime where a search for an intermittent GWB should start to perform better than the standard continuous-Gaussian cross-correlation search. The value of
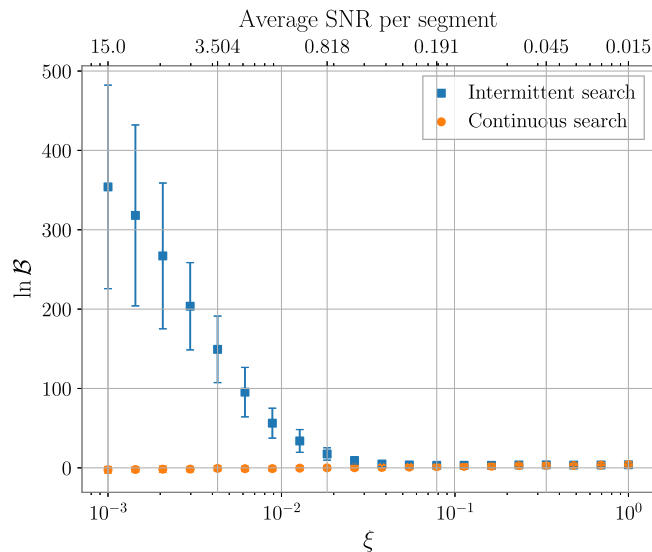
$\rho_{\text{seg}}$ at which the intermittent search begins to outperform the continuous search in Fig. 1 matches this result.

### B. Deterministic-signal-based search for intermittent GWBs

In 2018, Smith and Thrane [19] extended the work of Drasco and Flanagan [18] by proposing an optimal fully Bayesian deterministic-signal-based search for the intermittent GWB produced by the population of stellar-mass BBH mergers throughout the Universe. As in [18], Smith and Thrane [19] assume a mixture model for the intermittent GW signals. They chose a segment duration ∼4 s, which is long enough to include a typical BBH chirp signal, yet short enough that the probability of two such signals occurring in a single segment is negligibly small ($\sim 10^{-4}$). However, instead of considering single-sample GW bursts drawn from a fixed Gaussian distribution, they considered finite-duration deterministic BBH chirp waveforms $h = h_{\text{chirp}}(t; \theta)$, where $\theta$ are the chirp parameters (e.g., the component masses and spins of the two BHs and the inclination angle of the orbital plane relative to the line of sight). Smith and Thrane then marginalized (instead of maximized) over the signal parameters for each segment of data, assuming prior probability distributions for these parameters, while replacing the noise parameters by measured estimates of these quantities. If the signal priors are conditioned on segment-independent population parameters $\theta_{\text{pop}}$, which parametrize the distributions from which the individual masses, spins, etc., are drawn, then the final (marginalized) likelihood function $\mathcal{L}_{\text{tot}} \equiv \mathcal{L}_{\text{tot}}(d|\xi, \theta_{\text{pop}})$ depends only on the duty cycle $\xi$ and the population parameters $\theta_{\text{pop}}$. Finally, doing Bayesian inference calculations given $\mathcal{L}_{\text{tot}}$ and a prior for $\xi$ and $\theta_{\text{pop}}$, Smith and Thrane were able to construct joint posterior distributions for $\xi$ and $\theta_{\text{pop}}$ as well as Bayes factors comparing the evidence for this intermittent signal model and, e.g., that for the standard cross-correlation search for a continuous-Gaussian GWB.

The deterministic-signal-based search of Smith and Thrane is expected to decrease the time to detection of the intermittent GWB produced by stellar-mass BBH mergers by a factor of ∼1000 relative to the standard continuous-Gaussian search [19], by taking into account both the intermittent nature of the signal as well as the knowledge of the form of the individual waveforms, whose parameters are marginalized over. For this factor of ∼1000 determination, they did not consider any population parameters, so the only parameter that they needed to infer from the data was the duty cycle $\xi$. A posterior distribution for $\xi$ sufficiently bounded away from zero would be evidence of a confident detection of an intermittent GWB signal. The gain in time to detection comes at the computational cost of having to perform Bayesian marginalization over all the BBH chirp signal parameters for

every 4 s segment of data. This search is currently in the testing phase, in preparation for running on real LVK data in the near future.

Within this paper, for comparative purposes, we will implement a much simpler version of this deterministic-signal-based search. We will use the acronym DSI throughout this work to refer to the deterministic-signal-based search for intermittent GWBs.

### III. SSI: STOCHASTIC SEARCH FOR INTERMITTENT GWBs

Building off the work of Drasco and Flanagan [18], we propose a new search based on a stochastic-signal model consisting of intermittent bursts of correlated stochastic GWs with unknown duty cycle $\xi$, in otherwise uncorrelated noise in two detectors. We call this search SSI, for stochastic search for intermittent GWBs, referencing both the signal model the analysis assumes and the type of background for which it is designed. To make the connection with BBH mergers, we assume that these bursts of GWs last on the order of a few seconds so the data are split into short stretches as in Smith and Thrane, and that the power spectrum in the LVK detectors goes like $f^{-7/3}$, appropriate for binary inspiral. This corresponds to a fractional energy density spectrum $\Omega_{\text{gw}}(f) \propto f^{2/3}$, as introduced in (1).

Rather than marginalize over the parameters of deterministic BBH chirp waveforms as in the deterministic-signal-based approach, our search looks for excess cross-correlated power when the signal is assumed to be present, using a mixture-model likelihood function. Thus, we trade off optimality for computational efficiency and flexibility relative to the deterministic-signal-based approach, while still accounting for the intermittent nature of the BBH background, which is missing from the standard cross-correlation search for continuous-Gaussian GWBs.

We begin by dividing up the data into short segments such that the probability of a segment containing more than one signal is small. The total likelihood is given by a product over segments of the GMM likelihood function

$$
\begin{aligned}
&\mathcal{L}_{\text{tot}}(d|\xi, \theta_{s,\text{pop}}, \theta_n) \\
&= \prod_I [\xi \mathcal{L}_s(d_I|\theta_{s,\text{pop}}, \theta_n) + (1-\xi)\mathcal{L}_n(d_I|\theta_n)], \quad (11)
\end{aligned}
$$

where $\theta_n$ represents the noise parameters, $\theta_{s,\text{pop}}$ represents the signal population parameters, and $d_I$ represents the data in segment $I$.

For our stochastic-signal-based search, the segment-dependent signal likelihood takes the form

$$
\mathcal{L}_s(d_I|\theta_{s,\text{pop}}, \theta_n) \equiv \int d\theta_{s,I} \mathcal{L}_n(d_I|\theta_{s,I}, \theta_n) \pi(\theta_{s,I}|\theta_{s,\text{pop}}), \quad (12)
$$

where the segment-dependent signal parameters $\theta_{s,I}$ are marginalized over. Marginalizing over the correct segment

prior is an important and necessary step in order to recover correct and unbiased results.

We choose to write the likelihood for a specific set of parameters, $\theta_{s,\mathrm{pop}} = \langle \Omega_b \rangle$, $\theta_{s,I} = \Omega_{b,I}$, and $\theta_n = \{\sigma_{n_1}^2, \sigma_{n_2}^2\}$, where $\langle \Omega_b \rangle$ is the population-averaged energy density amplitudes of bursts of GW power and $\Omega_{b,I}$ is the energy density amplitude in data segment $I$. The population parameter $\langle \Omega_b \rangle$ is related to $\Omega_{\mathrm{gw}}$, introduced at the beginning of Sec. II, by

$$\Omega_{\mathrm{gw}} = \xi \langle \Omega_b \rangle. \tag{13}$$

Recall that $\Omega_{\mathrm{gw}}$ is what the standard cross-correlation search for a continuous-Gaussian GWB estimates. For the analyses included in this paper, we simulate stationary, white-Gaussian noise. This means that the power spectrum

of the noise is independent of frequency and has the value

$$P_{n_\mu} = \frac{\sigma_{n_\mu}^2}{f_{\mathrm{high}} - f_{\mathrm{low}}}, \tag{14}$$

where $\mu = 1, 2$ is the detector index and $f_{\mathrm{low}}$ and $f_{\mathrm{high}}$ are the low- and high-frequency cutoffs for our search. We will take $f_{\mathrm{high}}$ to equal the Nyquist critical frequency $f_{\mathrm{nyq}} \equiv 1/(2\Delta t)$, where $\Delta t$ is the sample period. Each segment of time-domain data of duration $T$ is Fourier transformed and coarse-grained to frequencies $f_k$ having frequency resolution $M/T$. We then take our noise parameters to be the variance of the noise in each detector. Under these assumptions, the segment-dependent signal likelihood (12) becomes

$$\mathcal{L}_s(d_I | \langle \Omega_b \rangle, \sigma_{n_1}^2, \sigma_{n_2}^2) = \int d\Omega_{b,I} \pi(\Omega_{b,I} | \langle \Omega_b \rangle) \prod_k \frac{1}{(\pi T/2)^{2M}(P_{1,I}(f_k)P_{2,I}(f_k) - P_{b,I}^2(f_k))^M}$$

$$\times \exp\left\{ -\frac{M}{(P_{1,I}(f_k)P_{2,I}(f_k) - P_{b,I}^2(f_k))}[\hat{P}_{1,Ik}P_{2,I}(f_k) + \hat{P}_{2,Ik}P_{1,I}(f_k) - 2\hat{P}_{b,Ik}P_{b,I}(f_k)] \right\}, \tag{15}$$

where

$$P_{1,I}(f) \equiv \frac{\sigma_{n_1}^2}{f_{\mathrm{high}} - f_{\mathrm{low}}} + P_{b,I}(f),$$

$$P_{2,I}(f) \equiv \frac{\sigma_{n_2}^2}{f_{\mathrm{high}} - f_{\mathrm{low}}} + P_{b,I}(f),$$

$$P_{b,I}(f) \equiv \Omega_{b,I}H(f) \tag{16}$$

are the total autocorrelated power spectra in each detector and the power spectrum for a GW burst in segment $I$, and $k$ runs over the coarse-grained frequencies $f_k$. The derivation of the product within the integrand of (15) is given in Sec. IV A. of [23]. The prior $\pi(\Omega_{b,I} | \langle \Omega_b \rangle)$ is not specified at this point as its functional form depends on the expected distribution of the energy density of the signals injected in Sec. IV. The spectral shape $H(f)$ is of the form

$$H(f) \equiv \frac{3H_0^2}{10\pi^2} \frac{1}{f_{\mathrm{ref}}^3} \left(\frac{f}{f_{\mathrm{ref}}}\right)^{-7/3}. \tag{17}$$

The Fourier transformed data enter the evidence via the following quadratic combinations:

$$\hat{P}_{1,Ik} \equiv \frac{2}{T} \frac{1}{M} \sum_{k'=k-M/2}^{k+M/2-1} |\tilde{d}_{1,Ik'}|^2,$$

$$\hat{P}_{2,Ik} \equiv \frac{2}{T} \frac{1}{M} \sum_{k'=k-M/2}^{k+M/2-1} |\tilde{d}_{2,Ik'}|^2,$$

$$\hat{P}_{b,Ik} \equiv \frac{2}{T} \frac{1}{M} \sum_{k'=k-M/2}^{k+M/2-1} \mathrm{Re}(\tilde{d}_{1,Ik'}^* \tilde{d}_{2,Ik'}), \tag{18}$$

which are coarse-grained estimators (i.e., averaged over fine-grained frequencies labeled by $k'$) of the total autocorrelated and cross-correlated power spectra in the two detectors.

The segment-dependent noise likelihood can similarly be written as

$$\mathcal{L}_n(d_I | \sigma_{n_1}^2, \sigma_{n_2}^2) = \prod_k \frac{1}{(\pi T/2)^{2M}(P_{n_1}(f_k)P_{n_2}(f_k))^M}$$

$$\times \exp\left\{ -M\left[\frac{\hat{P}_{1,Ik}}{P_{n_1}} + \frac{\hat{P}_{2,Ik}}{P_{n_2}}\right] \right\}. \tag{19}$$

In principle, the noise parameters $\theta_n = \{\sigma_{n_1}^2, \sigma_{n_2}^2\}$ in the likelihood functions above should be inferred together with the signal population parameters $\theta_{s,\mathrm{pop}} = \langle \Omega_b \rangle$, as part of the Bayesian inference procedure. Doing so defines the so-called *full* version of the analyses. However, as LVK noise is stationary to a good approximation, it is typically sufficient to use measured estimates of the noise parameters

[denoted by $\bar{\sigma}^2_{n_1}$ and $\bar{\sigma}^2_{n_2}$ and computed using (A16)] in the likelihood function, thereby avoiding having to infer them in this analysis. We refer to this approach as the *reduced* form of the analyses, which is computationally cheaper than the full form. The reduced version of the likelihood

requires that the cross-correlation estimators be approximately Gaussian, which holds only if the number of samples per segment $N$ is sufficiently large.

The reduced segment-dependent signal likelihood is given by [23]

$$\mathcal{L}_s(d_I|\langle\Omega_b\rangle, \bar{\sigma}^2_{n_1}, \bar{\sigma}^2_{n_2}) = \int d\Omega_{b,I}\pi(\Omega_{b,I}|\langle\Omega_b\rangle)\frac{1}{\sqrt{2\pi\mathrm{var}(\bar{\Omega}_{b,I})}}\exp\left[-\frac{(\hat{\Omega}_{b,I}-\Omega_{b,I})^2}{2\mathrm{var}(\bar{\Omega}_{b,I})}\right], \tag{20}$$

where

$$\hat{\Omega}_{b,I} \equiv \frac{\sum_k Q_I(f_k)\hat{P}_{b,Ik}}{\sum_{k'} Q_I(f_{k'})H(f_{k'})},$$

$$\mathrm{var}(\bar{\Omega}_{b,I}) \equiv \left(2M\sum_k Q_I(f_k)H(f_k)\right)^{-1} \tag{21}$$

are the optimally filtered cross-correlation estimators and corresponding variances, which are constructed from coarse-grained estimates of the cross-correlated power $\hat{P}_{b,Ik}$ [given by (18)] and the segment-dependent optimal filter function

$$Q_I(f) \equiv \frac{H(f)}{\bar{P}_{1,I}(f)\bar{P}_{2,I}(f)}, \tag{22}$$

where

$$\bar{P}_{1,I}(f) \equiv \frac{\bar{\sigma}^2_{n_1}}{f_{\mathrm{high}}-f_{\mathrm{low}}} + \Omega_{b,I}H(f),$$

$$\bar{P}_{2,I}(f) \equiv \frac{\bar{\sigma}^2_{n_2}}{f_{\mathrm{high}}-f_{\mathrm{low}}} + \Omega_{b,I}H(f). \tag{23}$$

Note that $Q_I(f)$ is a generalization of the standard optimal filter for an $f^{-7/3}$ power spectrum (see, e.g., [16,24]), extended to include the segment-dependent burst contribution, i.e., dependent on the likelihood parameter $\Omega_{b,I}$, to the total autocorrelated power estimates $\bar{P}_{1,I}(f)$, $\bar{P}_{2,I}(f)$.

The reduced segment-dependent noise likelihood $\mathcal{L}_n(d_I|\bar{\sigma}^2_{n_1}, \bar{\sigma}^2_{n_2})$ is given by

$$\mathcal{L}_n(d_I|\bar{\sigma}^2_{n_1}, \bar{\sigma}^2_{n_2}) = \frac{1}{\sqrt{2\pi\mathrm{var}(\bar{\Omega}_b)}}\exp\left[-\frac{(\hat{\Omega}_{b,I})^2}{2\mathrm{var}(\bar{\Omega}_b)}\right], \tag{24}$$

where $\hat{\Omega}_{b,I}$ and $\mathrm{var}(\bar{\Omega}_b)$ are the same as for the segment-dependent signal likelihood, but with a segment- *independent*, noise-only optimal filter function

$$Q(f) \equiv \frac{H(f)}{\bar{P}_{n_1}\bar{P}_{n_2}}. \tag{25}$$

## IV. ANALYSES

In this section, we describe in detail a set of analyses, which we use to illustrate various aspects of the search methods described above. The tests that these analyses allow us to perform should be thought of as providing a "proof-of-principle" demonstration of our proposed stochastic-signal-based search for intermittent GWBs. A more rigorous test of this search on actual LVK noise and realistic injected BBH chirp signals is a topic for future investigation (see Sec. V for more details).

For all the analyses we consider, we assume white, stationary-Gaussian noise in two colocated and coaligned detectors with variances $\sigma^2_{n_1}$ and $\sigma^2_{n_2}$, respectively. The assumption of colocated and coaligned detectors means that we can ignore the so-called overlap reduction function [25,26], which encodes the reduction in cross-correlated power that comes from correlating two physically separated and possibly misaligned detectors. To calculate the total SNR for each set of data, we use the average SNR per segment computed using formulas specified below for each dataset and rearrange (10) to solve for $\rho_{\mathrm{tot}}$. We note that this $\rho_{\mathrm{tot}}$ is the total SNR of the continuous-in-time cross-correlation search, which assumes the signal exists in every segment of data. For our intermittent analyses, we use this definition of total SNR to quantify the strength of the GW signal.

### A. Extension of previous work

In Sec. II A, the results of Drasco and Flanagan [18] are reproduced within a Bayesian framework (see Fig. 1). We remind the reader that the signals considered there are single-sample GW bursts drawn from a fixed Gaussian distribution with variance $\sigma^2_b$. We proceed with the generalization of the proposed GMM likelihood to allow for more realistic signals.

As a first step, we now allow multisample ($N \gg 1$) bursts of white stochastic GWs having duty cycle $\xi$, with signal samples drawn from a probability distribution that depends on the distance $r$ to an individual source. We require each burst to span the entire length of the segment it lies in so that the number of time-series samples of each burst is always equal to the number of samples per segment

*N*. For a source at arbitrary reference distance $r_{ref}$, we draw the signal samples from a Gaussian distribution with fixed variance $\sigma_{ref}^2$. For a source at a general distance $r$, we first draw the signal samples from a Gaussian distribution with variance $\sigma_{ref}^2$ as explained above, and then we rescale the samples by a factor of $r_{ref}/r$, since GW signal amplitudes fall off as $1/r$ [27]. Thus,

$$\sigma_b^2(r) \equiv \sigma_{ref}^2 \frac{r_{ref}^2}{r^2} \tag{26}$$

is the burst variance for a source at distance $r$.

For the population model, we will assume that the source distances are drawn from a *uniform-in-volume* probability distribution

$$p(r|r_{max}) \equiv \frac{3r^2}{r_{max}^3 - r_{min}^3}, \tag{27}$$

where $r_{max}$ is the maximum distance out to which the sources are formed (i.e., an unknown population parameter that will eventually be inferred from the data). The parameter $r_{min}$ is taken to be a fixed, known parameter, for simplicity. Note that choosing $r_{min} \neq 0$ in the simulation process limits the number of GW bursts that are so loud that they are individually detectable in a single segment of data. We also note that this choice of population model is a simplification as it does not take into account cosmology.

It follows from (26) and (27) that

$$p(\sigma_b^2(r)|r_{max}) = \frac{3r_{ref}^3}{2(r_{max}^3 - r_{min}^3)}(\sigma_{ref}^2)^{3/2}(\sigma_b^2(r))^{-5/2} \tag{28}$$

is the probability distribution for the signal variance $\sigma_b^2(r)$ associated with a source at distance $r$. We also define the population-averaged burst variance:

$$\langle\sigma_b^2\rangle \equiv \int_{r_{min}}^{r_{max}} dr\, p(r|r_{max})\sigma_b^2(r)$$
$$= 3\sigma_{ref}^2 \frac{r_{ref}^2(r_{max} - r_{min})}{r_{max}^3 - r_{min}^3}, \tag{29}$$

which is obtained by averaging $\sigma_b^2(r)$ over the uniform-in-volume-distributed source distances $r$. We define $\sigma_{gw}^2 \equiv \xi\langle\sigma_b^2\rangle$, which has the interpretation of being the time and population-averaged variance of the signals. This quantity is what the standard cross-correlation search for a continuous-Gaussian GWB (SSC) estimates.

Since the probability distribution for $\sigma_b^2(r)$ depends on just one free parameter, i.e., $r_{max}$ in (28), we can equally well use the population-averaged variance $\langle\sigma_b^2\rangle$ as the population parameter for the probability distribution. Solving (29) for $r_{max}$ in terms of $\langle\sigma_b^2\rangle$, we find

$$r_{max} = r_{min}\left(\sqrt{-\frac{3}{4} + 3\frac{\sigma_{b,max}^2}{\langle\sigma_b^2\rangle}} - \frac{1}{2}\right),$$
$$\sigma_{b,max}^2 \equiv \sigma_b^2(r_{min}) = \sigma_{ref}^2 \frac{r_{ref}^2}{r_{min}^2}, \tag{30}$$

leading to

TABLE I. Parameters used for the different analyses in Sec. IV. Parameters listed in "Extension of previous work" and "Stochastic bursts" were used in the production of Figs. 3 and 5, respectively. Note the units of the distance parameters in "Extension of previous work" are arbitrary since they have no effect on the variances of the bursts. The first nine columns in "Deterministic chirps" were used in the production of Fig. 7, while the last five columns specified the additional parameters used for Fig. 8.

| Extension of previous work | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $N_{seg}$ | $N$ | $\xi$ | $r_{min}$ | $r_{max}$ | $r_{ref}$ | $\sigma_{ref}^2$ | $\langle\sigma_b^2\rangle$ | $\sigma_n^2$ | $\langle\rho_{seg}\rangle$ | $\rho_{tot}$ |
| $4 \times 10^4$ | 2048 | $2.98 \times 10^{-3}$ | 2 | 5 | 1 | 1 | 0.0769 | 0.691 | 5.04 | 3 |

| Stochastic bursts | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $N_{seg}$ | $N$ | $T$ | $\xi$ | $r_{min}$ | $r_{max}$ | $r_{ref}$ | $\Omega_{ref}$ | $\langle\Omega_b\rangle$ | $f_{low}$ | $f_{high}$ | $\langle\rho_{seg,stoch}\rangle$ | $\rho_{tot,stoch}$ |
| $4 \times 10^4$ | 2048 | 4 s | $2.98 \times 10^{-3}$ | 2 Mpc | 5 Mpc | 2 Mpc | 2.61 | 0.803 | 20 Hz | 256 Hz | 5.04 | 3 |

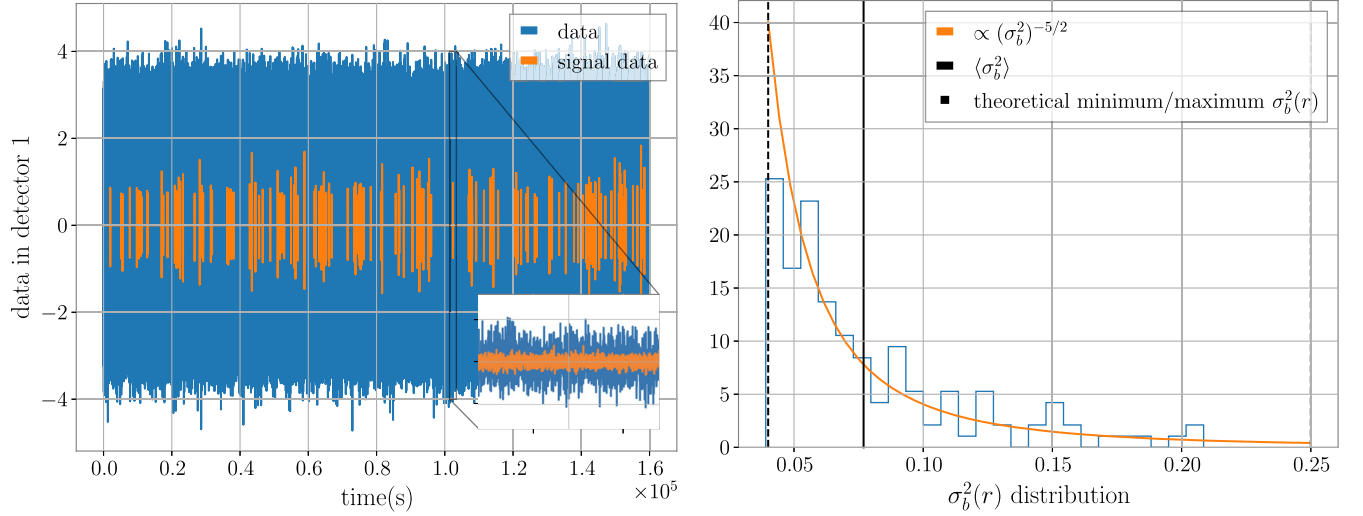| Deterministic chirps | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $N_{seg}$ | $N$ | $T$ | $r_{min}$ | $r_{max}$ | $f_{low}$ | $f_{high}$ | $m$ | $\langle\Omega_b\rangle$ | $\xi$ | $\langle\rho_{seg,stoch}\rangle$ | $\rho_{tot,stoch}$ | $\langle\rho_{seg,det}\rangle$ | $\rho_{tot,det}$ |
| $4 \times 10^4$ | 2048 | 4 s | 2 Mpc | 5 Mpc | 20 Hz | 256 Hz | $30M_\odot$ | 0.803 | $2.98 \times 10^{-3}$ | 5.04 | 3 | 13.2 | 7.86 |

FIG. 2.    Left: Example of simulated data with amplitudes drawn from a uniform-in-volume distribution along with a subplot zooming in on a single burst. The parameters used for this injection are given in the "Extension of previous work" section of Table I. Right: Distribution of the burst variances drawn from a uniform-in-volume distribution, with theoretical minimum and maximum burst variances evaluated at $r_{\max}$ and $r_{\min}$, respectively, and average burst variance $\langle \sigma_b^2 \rangle$ computed according to (29).

$$p(\sigma_b^2(r)|\langle\sigma_b^2\rangle) = \frac{\langle\sigma_b^2\rangle(\sigma_{b,\max}^2)^{1/2}}{\sqrt{-3 + 12\sigma_{b,\max}^2/\langle\sigma_b^2\rangle - 3}}(\sigma_b^2(r))^{-5/2}.$$

$$\tag{31}$$

The above expression is somewhat messy, but it will become the prior when we perform Bayesian inference on $\langle\sigma_b^2\rangle$ [i.e., $\pi(\sigma_{b,I}^2|\langle\sigma_b^2\rangle) = p(\sigma_b^2(r_I)|\langle\sigma_b^2\rangle)$]. Building on the above, we define the average segment SNR of the distribution in a similar manner as (29)

$$\langle\rho_{\text{seg}}\rangle = \int_{r_{\min}}^{r_{\max}} \mathrm{d}r\, p(r|r_{\max})\rho_{\text{seg}}(r), \tag{32}$$

where $\rho_{\text{seg}}(r)$ for these signals is given by (8) with $\sigma_b^2$ replaced by $\sigma_b^2(r)$.

We generate multisample ($N = 2048$) bursts of white stochastic GWs having duty cycle $\xi = 2.98 \times 10^{-3}$, with signal samples drawn from a probability distribution that depends on the distance $r$ to an individual source, as described above. With the chosen parameters (listed explicitly in Table I) the population-averaged variance is $\langle\sigma_b^2\rangle = 0.0769$ and the noise variances are $\sigma_{n_1}^2 = \sigma_{n_2}^2 = 0.691$. An example of the simulated data is shown in Fig. 2, together with the distribution of the burst variances $\sigma_b^2(r)$.

We analyze the data with SSC and SSI, using the full version of the likelihoods, i.e. inferring the noise parameters as well as the population parameters. We will not consider DSI for this particular data. The concrete expressions for the likelihoods can be found in the Appendix with the prior specified by (31). In Fig. 3, we display the recovery of our SSI search, illustrating that the

generalizations made in this section still allow for a successful recovery of the population and noise parameters.

We note that given the large number of samples per segment ($N = 2048$) used for this analysis, one could have resorted to the reduced version of the likelihoods, where the estimates of the noise parameters are used (as provided in the Appendix). We refrain from entering into a detailed comparison between full and reduced implementations of the likelihoods, as this was the topic of work by Matas and Romano [23]. Throughout the remainder of the paper, we will work with a large number of samples per segment and will employ the reduced version of the likelihoods.

### B. Stochastic bursts

We extend the analysis described in the previous section to include frequency dependence. We analyze data defined by multisample ($N \gg 1$) bursts of stochastic GWs having duty cycle $\xi$ and an $f^{-7/3}$ power spectrum for a uniform-in-volume distribution of source distances between $r_{\min}$ to $r_{\max}$, as in Sec. IV A. The choice of spectral index $-7/3$ is appropriate for compact binary inspiral. We first simulate data for a source at reference distance $r_{\text{ref}}$ so that it has the power spectrum[2]

$$P_{\text{ref}}(f) = A_{\text{ref}}\left(\frac{f}{f_{\text{ref}}}\right)^{-7/3}, \tag{33}$$

where $A_{\text{ref}}$ is some fixed amplitude and $f_{\text{ref}}$ is a reference frequency, usually taken to be 25 Hz in line with LVK

---

[2]In practice, we first simulate the data in the frequency domain with an amplitude spectral density $\sqrt{P_{\text{ref}}(f)}$ and random phases, and then inverse-Fourier transform the data back to the time domain.
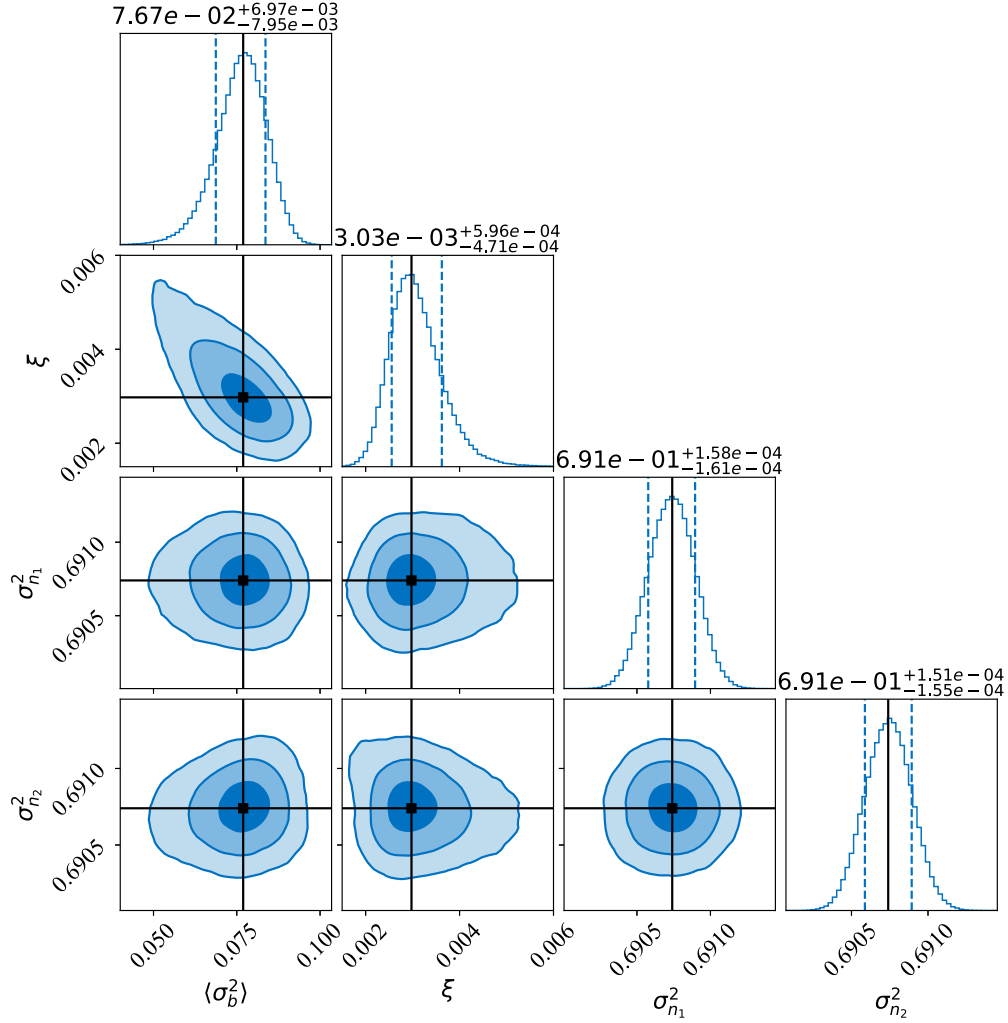
FIG. 3.   Corner plot for the full version of the SSI analysis, combining the posteriors of 100 realizations of the data. The black lines show the injected values of the parameters used for the simulated data, and the three shaded regions for the two-dimensional joint posteriors correspond to $1\sigma$, $2\sigma$, and $3\sigma$ uncertainty levels. All parameters are recovered within a $1\sigma$ credible interval.

searches. For a source at a general distance $r$, we do the same as above, and then rescale the amplitude of the simulated signal by a factor of $r_{\rm ref}/r$, which is equivalent to having

$$A_b(r) \equiv A_{\rm ref} \frac{r_{\rm ref}^2}{r^2} \qquad (34)$$

as the amplitude of the power spectral density for a GW burst at source distance $r$. The power spectrum of a burst is therefore

$$P_b(r; f) = A_{\rm ref} \frac{r_{\rm ref}^2}{r^2} \left( \frac{f}{f_{\rm ref}} \right)^{-7/3}. \qquad (35)$$

Note that by using (2), we can also write the above expression in terms of the fractional energy density spectrum $\Omega_b(r; f)$. Then by taking $f = f_{\rm ref}$, we can define the amplitude of the energy density at reference frequency $f_{\rm ref}$ of a burst at distance $r$,

$$\Omega_b(r) \equiv \frac{10\pi^2}{3H_0^2} f_{\rm ref}^3 P_b(r; f_{\rm ref}) = \Omega_{\rm ref} \frac{r_{\rm ref}^2}{r^2}, \qquad (36)$$

where

$$\Omega_{\rm ref} \equiv \frac{10\pi^2}{3H_0^2} f_{\rm ref}^3 A_{\rm ref}. \qquad (37)$$

By following the same derivation given in (29), the population-averaged energy density amplitude for sources distributed uniformly in volume between $r_{\rm min}$ and $r_{\rm max}$ is

$$\langle \Omega_b \rangle = 3\Omega_{\rm ref} \frac{r_{\rm ref}^2 (r_{\rm max} - r_{\rm min})}{r_{\rm max}^3 - r_{\rm min}^3}. \qquad (38)$$

The probability distribution of the amplitude of the energy density of the bursts $\Omega_b(r)$ has the same form as (31)

$$p(\Omega_b(r)|\langle\Omega_b\rangle) = \frac{\langle\Omega_b\rangle\Omega_{b,\max}^{1/2}}{\sqrt{-3 + 12\Omega_{b,\max}/\langle\Omega_b\rangle - 3}}\Omega_b^{-5/2}(r),$$

$$(39)$$

where $\Omega_{b,\max} \equiv \Omega_b(r_{\min})$.

Thus, the signal segment likelihood used for SSI is given by (15) (full) and (20) (reduced) with prior given by (39) [i.e., $\pi(\Omega_{b,I}|\langle\Omega_b\rangle) = p(\Omega_b(r_I)|\langle\Omega_b\rangle)$]. The integration bounds are then $\Omega_{b,\min}(\langle\Omega_b\rangle)$ and $\Omega_{b,\max}$ where $\Omega_{b,\min} = \Omega_b(r_{\max})$ and $r_{\max}$ is written in terms of the population parameter $\langle\Omega_b\rangle$, in the same manner as (30).

For reference, we note that the expected value of the stochastic (optimally filtered) signal-to-noise ratio for a segment that contains a GWB burst is

$$\rho_{\text{seg,stoch}} = \sqrt{2T}\left[\int_{f_{\text{low}}}^{f_{\text{high}}} df \frac{P_b^2(f)}{P_{n_1}P_{n_2}}\right]^{1/2}, \quad (40)$$

where $P_{n_1}$ and $P_{n_2}$ are the power spectra of the noise in each detector. Note that if the two detectors were not colocated and coaligned, we would need to include a factor of the overlap reduction squared in the numerator of the integrand in (40). The above expression for $\rho_{\text{seg,stoch}}$ is a *power* signal-to-noise ratio, defined as the expected value of the optimally filtered cross-correlation statistic divided by its standard deviation (see, e.g., [24]).

As mentioned before, our stochastic-signal-based search looks for a GWB consistent with a power spectrum of spectral index $-7/3$, as expected for BBH mergers. In contrast, the deterministic-signal-based search described in Sec. II B (which we call DSI) looks for deterministic BBH chirp waveforms, where the signal parameters of the individual chirps must be marginalized over. We inject

intermittent, stochastic bursts with an $f^{-7/3}$ power spectrum and duty cycle $\xi = 2.98 \times 10^{-3}$. The parameters used for the injection are displayed in Table I. We arbitrarily choose the reference distance $r_{\text{ref}} = r_{\min}$. The value of $\Omega_{\text{ref}}$ is chosen to be 2.61 (to be consistent with the parameters chosen in Sec. IV C). With these parameters, the population-averaged energy density amplitude of the bursts is $\langle\Omega_b\rangle = 0.803$. The noise is then set such that the average SNR per segment, as computed with (40), is 5.04 to give a total SNR of 3, as obtained by using (10). An example of one stochastic burst is shown in Fig. 4 (left panel), along with the corresponding average power spectral density of a population of stochastic bursts (right panel).

We analyze our data with the reduced forms which estimate the noise parameters of our stochastic-signal-based search and the deterministic-signal-based search. The exact form of the likelihood is given in Sec. III (with coarse-graining factor $M = 16$) and the Appendix, respectively. The population parameter recovered by SSI is $\langle\Omega_b\rangle$ while the population parameter recovered by DSI is $r_{\max}$. Note, these are related by (38). In Fig. 5, we demonstrate that DSI cannot recover the signal in the data, since no chirp waveform exists. While this result is in a sense obvious, it highlights the challenges that a deterministic-signal-based search faces. Incorrectly modeling the waveforms of the chirps could lead the search to overlook a signal which is present. Conversely, SSI recovers both stochastic bursts of GW power as well as deterministic waveforms, as we will see in the next section.

### C. Deterministic chirps

Finally, we consider multisample bursts of GWs produced by deterministic BBH chirp signals, for a uniform-in-volume distribution of sources (27). The corresponding
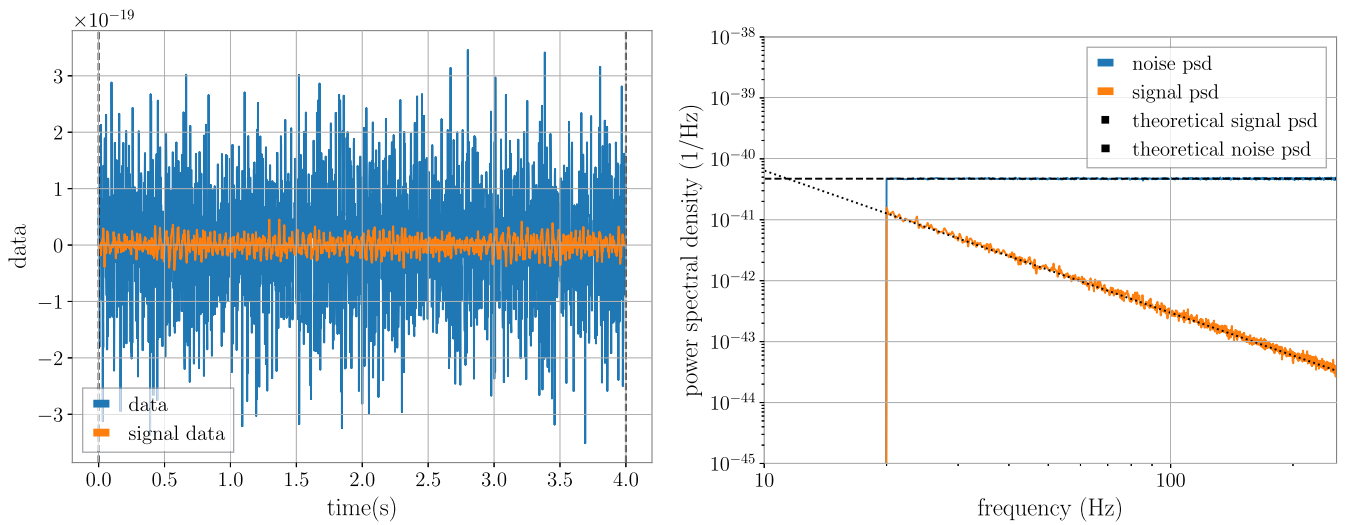


FIG. 4. Left: Example of one stochastic burst in the time-domain. Right: Averaged power spectral density of a population of stochastic burst signals as a function of frequency for the noise and signal separately, along with their theoretical predictions according to the injected values.
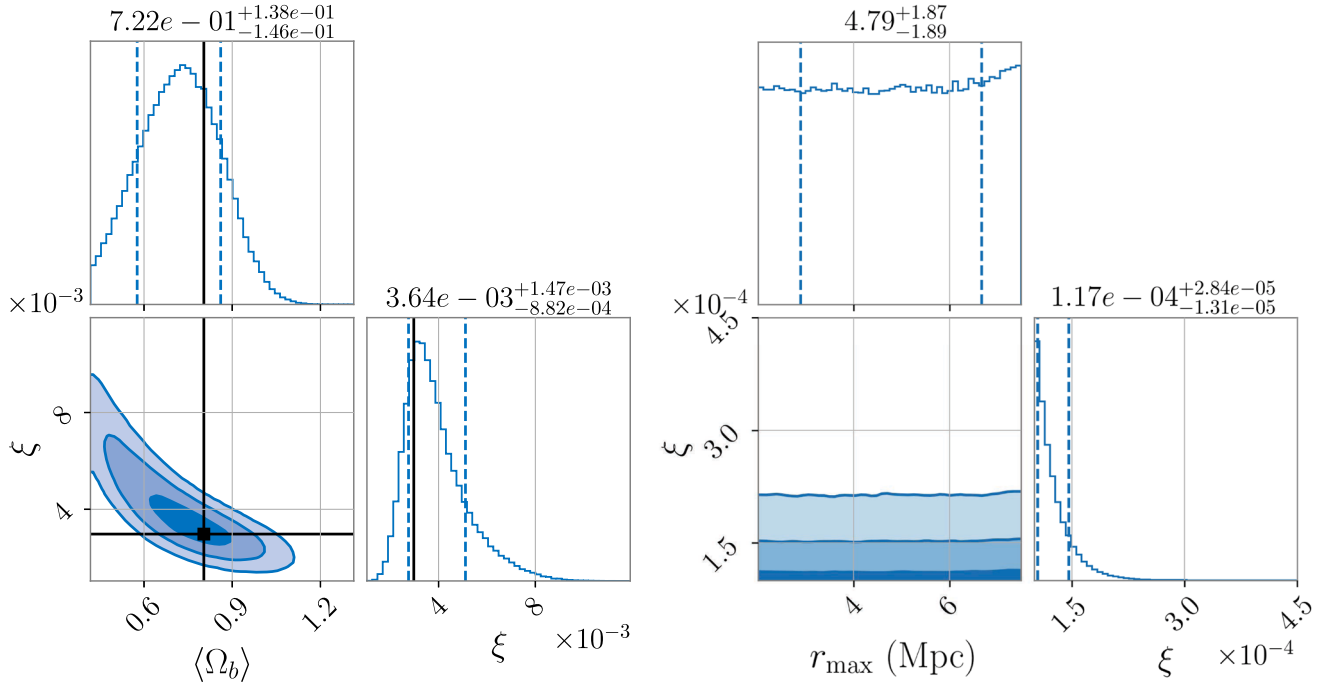
FIG. 5.   For intermittent, stochastic bursts with an $f^{-7/3}$ power spectrum, we demonstrate recovery of our search (left) and compare it to that of a deterministic-signal-based search (right). Our search recovers the injected signal parameters within a $1\sigma$ credible interval, while DSI recovers the uniform prior on $r_{\max}$ and the lower boundary of the prior imposed on the duty cycle ($\xi = 10^{-4}$). Thus, the DSI analysis finds no signal in the data.

power spectrum will necessarily have an approximate $f^{-7/3}$ frequency dependence. By using deterministic BBH chirp signals, this analysis is more in line with the assumptions made by the deterministic-signal-based search DSI.

We assume that all parameters defining the chirp waveforms except for the distances to the sources [e.g., the chirp mass $\mathcal{M}_c \equiv (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$, the inclination angle $\iota$, the coalescence time $t_{\mathrm{col}}$, and the phase of coalescence within a segment] have fixed values and are known *a priori* by the DSI search. For simplicity, we choose the two component masses to be equal (i.e., $m_1 = m_2 \equiv m$); the inclination angle $\iota = \pi/2$ so that the source is linearly polarized [i.e., $h(t) = h_+(t)$, $h_\times(t) = 0$]; the phase at coalescence $\Phi_0$ to be zero; and the coalescence time $t_{\mathrm{col}}$ to occur at the end of a segment, so $t_{\mathrm{col}} = T$, the segment duration. For a source drawn from the population with distance $r$, the explicit form for the simulated deterministic chirp signal is given in the time domain by [27]

$$h_{\mathrm{chirp}}(t;r) = \frac{1}{2r}\left(\frac{G\mathcal{M}_c}{c^2}\right)^{5/4}\left(\frac{5}{c\tau}\right)^{1/4}\cos\left[\Phi(\tau)\right], \quad (41)$$

where $\tau \equiv t_{\mathrm{col}} - t$ and

$$\Phi(\tau) \equiv -2\left(\frac{5G\mathcal{M}_c}{c^3}\right)^{-5/8}\tau^{5/8} + \Phi_0 \quad (42)$$

encodes the frequency evolution of the chirp,

$$f(t) \equiv -\frac{1}{2\pi}\frac{\mathrm{d}}{\mathrm{d}\tau}\Phi(\tau) = \frac{1}{\pi}\left(\frac{G\mathcal{M}_c}{c^3}\right)^{-5/8}\left(\frac{5}{256}\frac{1}{\tau}\right)^{3/8}. \quad (43)$$

The corresponding BBH chirp power spectrum is

$$P_{\mathrm{chirp}}(r;f) = \frac{2}{T}|\tilde{h}_{\mathrm{chirp}}(r;f)|^2 \equiv A_{\mathrm{chirp}}(r)\left(\frac{f}{f_{\mathrm{ref}}}\right)^{-7/3}, \quad (44)$$

where $\tilde{h}_{\mathrm{chirp}}$ is the Fourier transform of the chirp waveform and

$$A_{\mathrm{chirp}}(r) = A_{\mathrm{ref}}\frac{r_{\mathrm{ref}}^2}{r^2}, \quad (45)$$

where

$$A_{\mathrm{ref}} \equiv \frac{2}{T}\frac{c^2}{4r_{\mathrm{ref}}^2}\left(\frac{5\pi}{24}\right)\left(\frac{G\mathcal{M}_c}{c^3}\right)^{5/3}(\pi f_{\mathrm{ref}})^{-7/3}. \quad (46)$$

Note one can express the chirp PSD, $P_{\mathrm{chirp}}$, in terms of the fractional energy density of the chirps by using (2). For reference, we note that the expected value of the deterministic (matched-filter) signal-to-noise ratio for a segment which contains a BBH chirp signal is [24]

$$\rho_{\text{seg,det}} = \left[ 4 \sum_{\mu=1}^{2} \int_{f_{\text{low}}}^{f_{\text{high}}} df \frac{|\tilde{h}_{\text{chirp}}(f)|^2}{P_{n_\mu}} \right]^{1/2}$$

$$= \sqrt{2T} \left[ \sum_{\mu=1}^{2} \int_{f_{\text{low}}}^{f_{\text{high}}} df \frac{P_{\text{chirp}}(f)}{P_{n_\mu}} \right]^{1/2}, \quad (47)$$

where $P_{n_\mu}$ is the noise power spectral density in detector $\mu = 1, 2$ [see (14)]. The above expression for $\rho_{\text{seg,det}}$ is an *amplitude* signal-to-noise ratio, defined as the expected value of the matched-filter statistic divided by its standard deviation. The quadrature sum takes into account the contribution from using both detectors to do the analysis.

Figure 6 shows a plot of a representative BBH chirp signal in the time -domain (left panel) and an average over an ensemble of BBH chirp signals in the frequency domain (right panel).

As mentioned in Sec. II A, the detection statistic in our Bayesian framework is the Bayes factor where the models in (7) are the signal + noise model and the noise-only model for a particular search. While SSC and SSI contain the same noise model, the noise model in DSI does not take the same form. Hence, the Bayes factors for the different searches are not computed with respect to the same noise model, and one cannot compare these methods with one another in terms of the Bayes factor. Instead, we evaluate how the intermittent nature of the signal impacts each search method's effectiveness in recovering the signal by plotting the ln Bayes factor as a function of the duty cycle. In other words, we wish to answer two questions: (i) How well does SSI do in recovering the signal at different duty cycles for a constant total stochastic signal-to-noise ratio? and (ii) How well does DSI do in recovering the signal at different duty cycles for a constant total deterministic

signal-to-noise ratio? The answers to the questions are independent of one another and cannot be used as a way to assess if one search is "better" than the other. However, since SSC and SSI contain the same noise model, these searches can be compared to one another using the Bayes factor.

To assess the efficiency of the methods with respect to their respective noise-only models, we simulate 40,000 segments of data with each segment being 4 s long. We choose values of $r_{\min} = 2$ Mpc, $r_{\max} = 5$ Mpc, and the black hole component masses to each be $30M_\odot$. These parameters give a value of $\langle \Omega_b \rangle = 0.803$. The parameters used for this analysis are tabulated in the first nine columns of the "Deterministic chirps" section in Table I. Thus, the signal has the same strength as in Sec. IV B, but it is now composed of deterministic chirps. The same coarse-graining factor and low- and high-frequency cutoffs that were used in Sec. IV B are used for this case as well when analyzing the data.

Figure 7 shows the ln Bayes factors for the stochastic-signal-based searches (left panel) and for the deterministic-signal-based search (right panel) as a function of the duty cycle. Analogous to what was done in Sec. II A, the total SNR is kept constant by adjusting the noise levels. For the stochastic searches, we keep the total power SNR, computed using (40), constant, while for the deterministic search we keep the total amplitude SNR constant, obtained using (47). We see that both intermittent searches (SSI and DSI) perform well at low duty cycles, with values of the ln Bayes factors reaching over 1000 for some of the smallest values of the duty cycle considered.

To directly compare SSI with DSI, we run both analyses on the same dataset. The data are generated such that the duty cycle is $2.98 \times 10^{-3}$, the signal is the same as
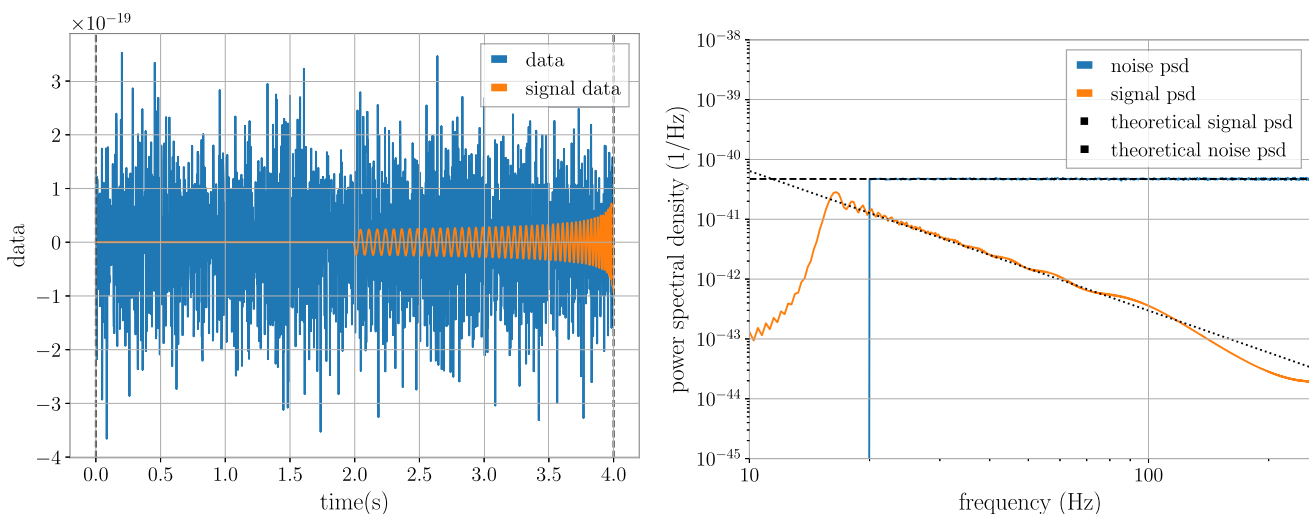


FIG. 6.   Left: Example BBH chirp signal in the time domain as given by (41). Right: Averaged power spectral density of an ensemble of BBH chirp signals as a function of frequency for the noise and signal separately, together with their theoretical predictions according to the injected values.
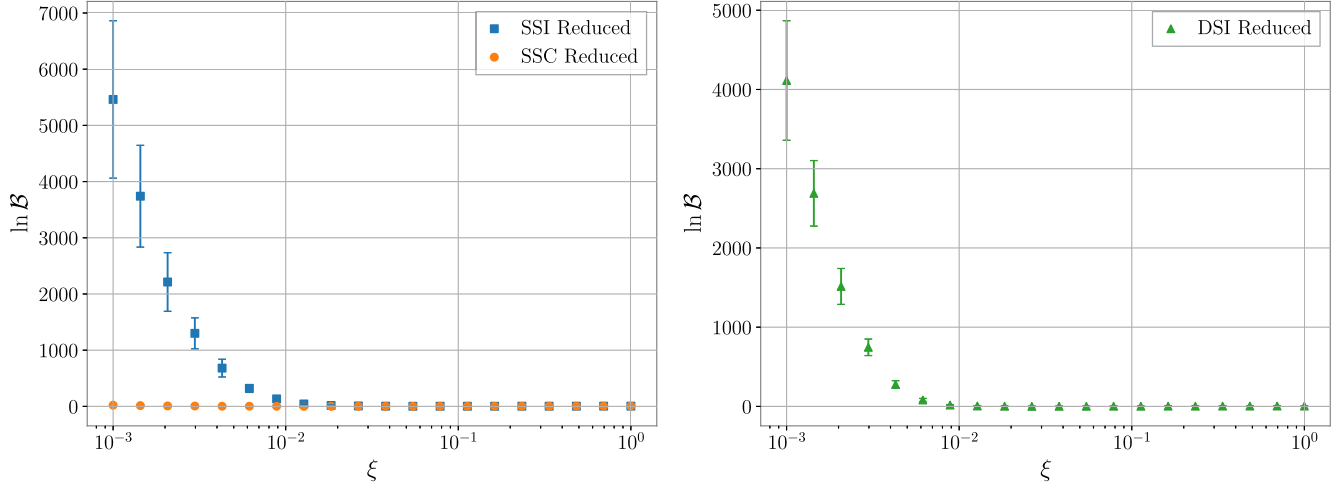
FIG. 7. Plots of the ln Bayes factor averaged over 100 data realizations for SSC and SSI (left) and DSI (right) for deterministic chirp signals occurring with various values of the duty cycle $\xi$. Both intermittent searches are well-suited for detecting signals with a low duty cycle.

described above, and the noise variance is chosen such that the average stochastic SNR per segment, computed using (40), is equal to 5.04 and the total stochastic SNR is equal to 3.0. Note for these values, the average deterministic SNR per segment, computed using (47), is 13.20 with the total deterministic SNR being 7.86, which is considerably larger than the total stochastic SNR. Note these parameters are displayed in the remaining columns of the "Deterministic chirps" section of Table I. A comparison of the recovered corner plots is shown in Fig. 8 (left panel). We see that for

these data, both searches recover the signal within a $1\sigma$ credible interval, with the error bars for DSI much smaller than SSI, due to the deterministic approach appropriately modeling the chirp waveform of the signal. We also show a comparison of one-dimensional (1D) posterior plots of $\Omega_{gw}$ in Fig. 8 (right panel). Similar to the corner plot, the posterior width is smaller for DSI than SSI, although SSI still performs better than SSC.

One notes a small bias in the recovery of $\Omega_{gw}$ for SSI in the right panel of Fig. 8. In Fig. 9 we show the relative
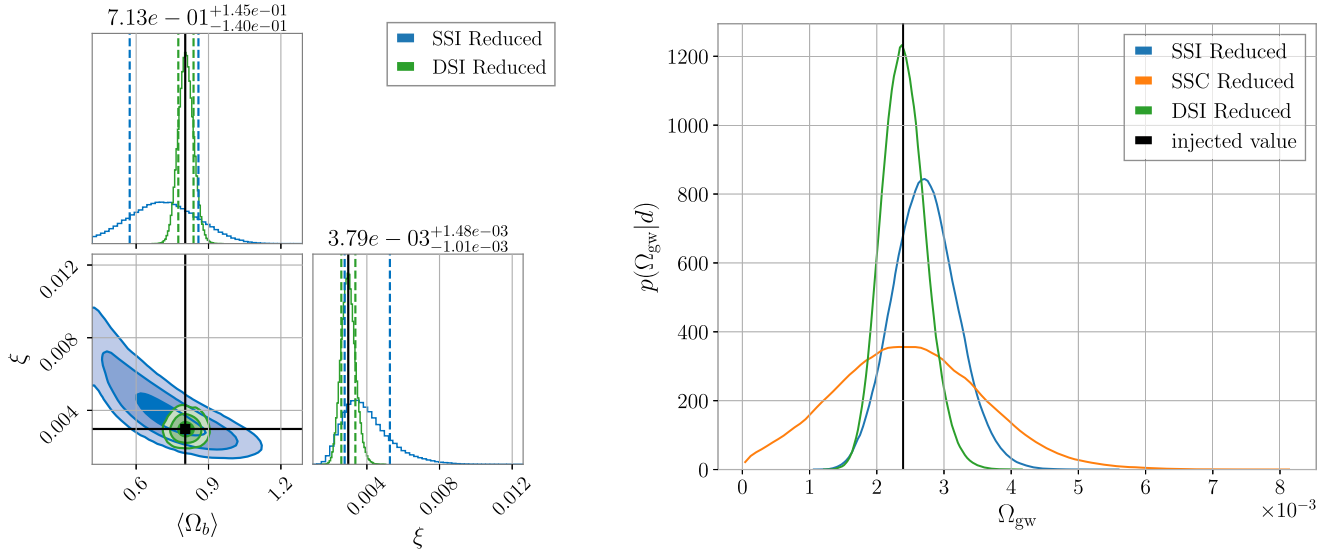


FIG. 8. Left: Posterior corner plot combined over 100 data realizations analyzed with SSI reduced (blue) and DSI reduced (green). Both searches recover the injected signal parameters ($\xi = 2.98 \times 10^{-3}$ and $\langle \Omega_b \rangle = 0.803$) within a $1\sigma$ confidence interval. The recovered values and error bars are those recovered by the SSI reduced search. Right: The 1D posterior plot of $\Omega_{gw}$ samples from SSI reduced (blue), SSC reduced (orange), and DSI reduced (green) constructed by combining posterior samples for $\xi$ and $\langle \Omega_b \rangle$ using (13). Note, the inference done with the DSI likelihood gives posterior samples for the parameters $\xi$ and $r_{max}$, and the values of $r_{max}$ are then converted to samples in $\langle \Omega_b \rangle$ by (38), since the other variables in (38) are fixed and known.
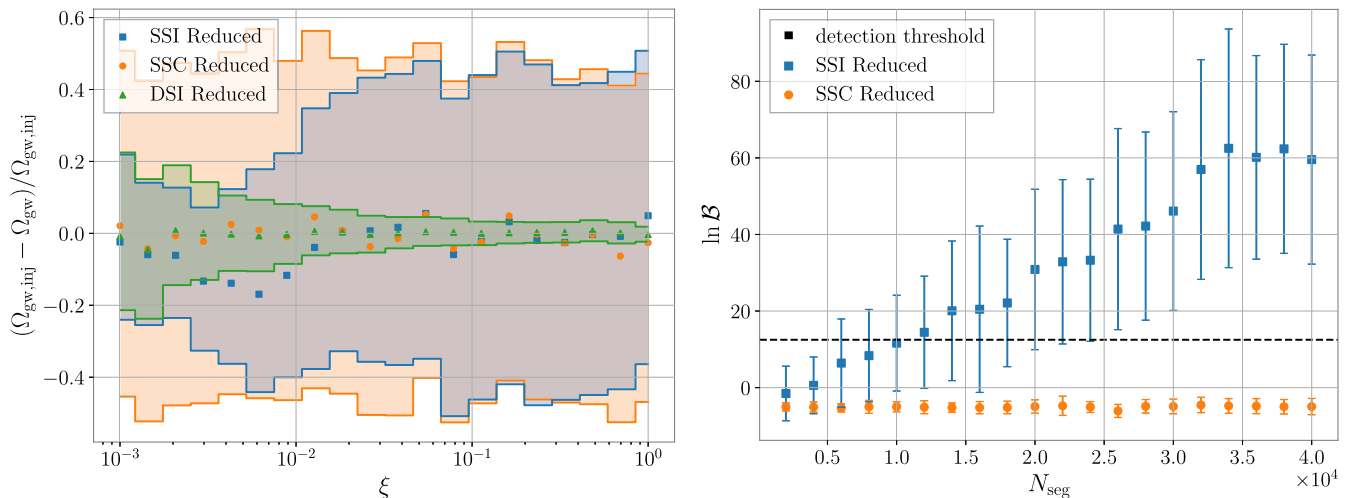
FIG. 9. Left: Comparison of recovered values to the injected value of $\Omega_{\mathrm{gw}}$ for SSI reduced (blue), SSC reduced (orange), and DSI reduced (green) for different values of the duty cycle. All injected parameters are equivalent to the parameters used in the left panel of Fig. 7, and the recovered values are those after combining 100 realizations of data. The shaded regions represent the $1\sigma$ credible interval of the combined 100 realizations of data. Right: ln Bayes factor vs $N_{\mathrm{seg}}$ for data with $\langle \rho_{\mathrm{seg,stoch}} \rangle = 2$ and $\xi = 2.98 \times 10^{-3}$. We define a detection threshold of $\ln \mathcal{B} = 12.5$. SSI crosses this threshold after $\sim 12{,}000$ segments of data, while SSC crosses this threshold after $\sim 650{,}000$ segments of data, corresponding to a factor of improvement in detecting the signal of roughly 54 for SSI relative to SSC.

difference of the injected value and the recovered value of $\Omega_{\mathrm{gw}}$ as a function of $\xi$ for the three searches, together with the $1\sigma$ uncertainty band, after combining the posterior over 100 realizations of data. We note that the biased recovery is not always toward higher values of $\Omega_{\mathrm{gw}}$. We also note that the width of the uncertainty for the DSI analysis improves as $\xi$ increases because the total deterministic SNR is not held constant and increases.

To conclude, we give an estimate of the improvement in time to detection of a GWB with our search. Note that this estimate is computed under the assumptions adopted in this paper and will therefore most likely differ for a realistic detector configuration, with realistic detector noise. We also note that the strength of the signal may affect these values. Nevertheless, to obtain such an estimate, we simulate a GWB consisting of deterministic chirps with parameters $\langle \rho_{\mathrm{seg,stoch}} \rangle = 2$ (corresponding to $\langle \rho_{\mathrm{seg,det}} \rangle = 8.3$) and $\xi = 2.98 \times 10^{-3}$. We then vary the number of data segments and assess how many 4 s segments are needed to reach a threshold value of the ln Bayes factor which is large enough to claim a detection. We define this threshold to be of value 12.5, corresponding to a detection of SNR equal to 5. This is shown in the right panel of Fig. 9 for SSI and SSC. Due to the large difference in deterministic and stochastic SNR, the ln Bayes factor for DSI already reaches $\sim 160$ at the first value of $N_{\mathrm{seg}}$ considered. We therefore do not include DSI on this plot to avoid scaling issues. We estimate that the SSC search would cross this threshold after $650{,}000$ segments of data. This corresponds to a factor of $\sim 54$ improvement in detection of SSI versus SSC for these parameters and assumptions.

## V. DISCUSSION

Developing data-analysis techniques to reduce the time to detection of an astrophysical GWB with the LVK detectors is one of the current challenges that the GW community faces. Searches that include the intermittency of the BBH background to improve detection statistics have been proposed in the past [18,19,28,29]. In this work, we propose a new, stochastic search for intermittent GWBs and compare its efficiency with other searches. Our stochastic-based search looks for excess cross-correlated power in short stretches of data, ignoring the deterministic form of the GW signal waveforms and, hence the need to marginalize over all the associated signal parameters, as is done in the deterministic-signal-based approach of Smith and Thrane [19]. Not only is it beneficial to develop multiple searches in order to cross-check a potential detection, but there is an added benefit to running a search which does not look for a specific waveform in the data. The stochastic signal model allows our search to be flexible with respect to the type of signal it can detect. By changing the spectral index $\alpha$ in the search (or by allowing $\alpha$ to be inferred as a population parameter) we could detect other intermittent signals which might exist in the data.

For a series of analyses on data of increasing complexity, we show that for data with low duty cycles our search performs better than the standard continuous cross-correlation search, which does not take the intermittent nature of the BBH background into account. Furthermore, we show that a stochastic search for intermittent GWBs is more flexible to the source of the intermittent GWB than our implementation of the Smith and Thrane approach [19] and should be

more computationally efficient in detecting a signal. The detection of an intermittent background will allow us to test existing theoretical models, as described in [30,31].

Before being able to apply this search method on real GW data, further generalizations need to be made. We give several examples of such generalizations, which will be addressed in future work.

For all of our data in this paper, we only simulate signals which lie completely within the segment boundaries. A crucial next step is investigating how a signal which extends past a segment boundary will impact our results.

In addition, the distribution of the injected signals was used in the Bayesian inference as the segment-dependent prior. Since a poorly chosen prior could bias results, more investigations will need to be done in order to determine a reasonable prior to use for inference on real data.

Further, the most realistic data we consider consists of individual BBH chirps injected in white, Gaussian noise. However, various assumptions were made about the source distribution that generates these chirps. For example, the two component masses were chosen to be equal, and the resulting chirp mass chosen to be identical for all the chirps (with only the distance to the source varying from one data segment to another). In reality, the black hole masses will most likely follow a power-law + peak distribution as shown by the latest LVK results [4]. Generalizing our method to allow for such mass distributions, as well as the performance of our search in that case, is left for future work.

Several simplifications regarding the detectors were made as well. First, we worked under the assumption that the detectors are colocated and coaligned. This needs to be generalized by taking into account the effect of the overlap reduction function. Second, it was assumed that the noise in the detector is white and Gaussian. However, realistic detector noise follows a colored, i.e., frequency-dependent, power spectral density. An additional complication related to noise estimation arises from the presence of a continuous GWB of BNS mergers. At any time, several BNS mergers are expected to be emitting GWs in the LVK frequency band. Not only does this violate the assumption that a segment contains either one signal or noise only, but it will

also affect the noise PSD estimation. Challenges related to the correct noise estimation will be addressed in future work. Furthermore, the Gaussian noise assumption will likely be violated as well, due to the presence of noise transients, so-called glitches. During the third observing run of the LVK Collaboration, these glitches were omnipresent in the data [21,32]. Therefore, before analyzing real detector data, the sensitivity of our search to the presence of such glitches will have to be investigated. Analyzing real detector data will introduce many challenges, which we plan to address incrementally, considering more and more realistic detectors and signals.

## APPENDIX: LIKELIHOODS

Throughout this work, various searches for GWBs are compared. In this appendix, we provide the likelihoods corresponding to those searches. We start by giving an overview of the likelihoods used in Sec. IV A, i.e., applicable to white signals, and conclude with the likelihoods for colored signals used in Secs. IV B and IV C. We also remind the reader that all likelihoods considered in this work are for stationary, white-Gaussian noise [see (14)].

### 1. Likelihoods for white signals

#### a. SSC-full

For white signals, we define the likelihood functions for a continuous stochastic search (SSC-full) as [23]

$$\mathcal{L}(d|\sigma_{\text{gw}}^2,\sigma_{n_1}^2,\sigma_{n_2}^2) = \prod_{I=1}^{N_{\text{seg}}} \frac{1}{(2\pi)^N(\sigma_1^2\sigma_2^2 - (\sigma_{\text{gw}}^2)^2)^{N/2}} \exp\left\{-\frac{1}{2}\frac{N}{(\sigma_1^2\sigma_2^2 - (\sigma_{\text{gw}}^2)^2)}[\hat{\sigma}_{1,I}^2\sigma_2^2 + \hat{\sigma}_{2,I}^2\sigma_1^2 - 2\hat{\sigma}_{\text{gw},I}^2\sigma_{\text{gw}}^2]\right\}, \quad (A1)$$

where

$$\sigma_1^2 \equiv \sigma_{n_1}^2 + \sigma_{\text{gw}}^2, \qquad \sigma_2^2 \equiv \sigma_{n_2}^2 + \sigma_{\text{gw}}^2, \quad (A2)$$

are parameters describing the total autocorrelated power in detectors 1 and 2, and

$$\hat{\sigma}_{1,I}^2 \equiv \frac{1}{N}\sum_i d_{1,Ii}^2, \quad (A3a)$$

$$\hat{\sigma}_{2,I}^2 \equiv \frac{1}{N}\sum_i d_{2,Ii}^2, \quad (A3b)$$

$$\hat{\sigma}_{\text{gw},I}^2 \equiv \frac{1}{N}\sum_i d_{1,Ii}d_{2,Ii} \quad (A3c)$$

are the quadratic combinations of the data from segment $I$ that enter the likelihood function. (Here, $i$ labels the time sample in data segment $I$.) The noise variances in each detector are $\sigma_{n_1}^2$ and $\sigma_{n_2}^2$. It turns out that $\hat{\sigma}_{1,I}^2, \hat{\sigma}_{2,I}^2, \hat{\sigma}_{\text{gw},I}^2$ are the maximum-likelihood estimates of $\sigma_1^2, \sigma_2^2, \sigma_{\text{gw}}^2$ for segment $I$.

### b. SSC-reduced

For a large number of samples per segment ($N \gg 1$), one can define a reduced version of the likelihood function, which is given by [23]

$$
\mathcal{L}(d|\sigma_{\text{gw}}^2, \bar{\sigma}_{n_1}^2, \bar{\sigma}_{n_2}^2)
$$

$$
= \prod_{I=1}^{N_{\text{seg}}} \frac{1}{\sqrt{2\pi \text{var}(\bar{\sigma}_{\text{gw}}^2)}} \exp\left[-\frac{(\hat{\sigma}_{\text{gw},I}^2 - \sigma_{\text{gw}}^2)^2}{2\text{var}(\bar{\sigma}_{\text{gw}}^2)}\right], \quad \text{(A4)}
$$

where

$$
\text{var}(\bar{\sigma}_{\text{gw}}^2) \equiv \frac{1}{N} \bar{\sigma}_1^2 \bar{\sigma}_2^2, \quad \text{(A5)}
$$

with

$$
\bar{\sigma}_1^2 \equiv \frac{1}{N_{\text{tot}}} \sum_{I,i} d_{1,Ii}^2, \qquad \bar{\sigma}_2^2 \equiv \frac{1}{N_{\text{tot}}} \sum_{I,i} d_{2,Ii}^2 \quad \text{(A6)}
$$

being estimates of the total autocorrelated power in the two detectors constructed from all the data. We expect SSC-reduced and SSC-full to perform equally well, assuming $N \gg 1$, which is needed for the cross-correlation data to be approximately Gaussian.

### c. SSI-full

For our proposed stochastic search for intermittent GWBs, we build upon the framework of Drasco and Flanagan [18] and extend their proposed formalism to a larger number of samples per segment ($N \gg 1$) and allow for the amplitudes to be drawn from a uniform-in-volume distribution. The likelihood takes the same form as (3), where the segment-dependent signal and noise likelihoods are now, respectively, given by

$$
\mathcal{L}_s(d_I|\langle\sigma_b^2\rangle, \sigma_{n_1}^2, \sigma_{n_2}^2) = \int_{\sigma_{b,\min}^2(\langle\sigma_b^2\rangle)}^{\sigma_{b,\max}^2} d\sigma_{b,I}^2 \pi(\sigma_{b,I}^2|\langle\sigma_b^2\rangle) \frac{1}{(2\pi)^N(\sigma_{1,I}^2\sigma_{2,I}^2 - (\sigma_{b,I}^2)^2)^{N/2}}
$$

$$
\times \exp\left\{-\frac{1}{2}\frac{N}{(\sigma_{1,I}^2\sigma_{2,I}^2 - (\sigma_{b,I}^2)^2)}[\hat{\sigma}_{1,I}^2\sigma_{2,I}^2 + \hat{\sigma}_{2,I}^2\sigma_{1,I}^2 - 2\hat{\sigma}_{b,I}^2\sigma_{b,I}^2]\right\}, \quad \text{(A7)}
$$

$$
\mathcal{L}_n(d_I|\sigma_{n_1}^2, \sigma_{n_2}^2) = \frac{1}{(2\pi)^N(\sigma_{n_1}^2\sigma_{n_2}^2)^{N/2}} \exp\left\{-\frac{N}{2}\left[\frac{\hat{\sigma}_{1,I}^2}{\sigma_{n_1}^2} + \frac{\hat{\sigma}_{2,I}^2}{\sigma_{n_2}^2}\right]\right\}, \quad \text{(A8)}
$$

where

$$
\hat{\sigma}_{b,I}^2 \equiv \frac{1}{N} \sum_i d_{1,Ii} d_{2,Ii}, \quad \text{(A9a)}
$$

$$
\hat{\sigma}_{1,I}^2 \equiv \frac{1}{N} \sum_i d_{1,Ii}^2, \quad \text{(A9b)}
$$

$$
\hat{\sigma}_{2,I}^2 \equiv \frac{1}{N} \sum_i d_{2,Ii}^2. \quad \text{(A9c)}
$$

In the above expression for the signal likelihood, we used

$$
\sigma_{1,I}^2 \equiv \sigma_{n_1}^2 + \sigma_{b,I}^2, \qquad \sigma_{2,I}^2 \equiv \sigma_{n_2}^2 + \sigma_{b,I}^2, \quad \text{(A10)}
$$

which are parameters describing the *segment-dependent* total autocorrelated power, with the segment dependence coming from the burst variance $\sigma_{b,I}^2$.

Note that the segment-dependent signal likelihood requires a marginalization over the segment-dependent burst variances $\sigma_{b,I}^2$, which is taken into account by the appropriate use of prior distribution, as introduced in (31),

$$
\pi(\sigma_{b,I}^2|\langle\sigma_b^2\rangle) = \frac{\langle\sigma_b^2\rangle(\sigma_{b,\max}^2)^{1/2}}{\sqrt{-3 + 12\sigma_{b,\max}^2/\langle\sigma_b^2\rangle - 3}}(\sigma_{b,I}^2)^{-5/2}, \quad \text{(A11)}
$$

where

$$
\sigma_{b,\min}^2(\langle\sigma_b^2\rangle) = \frac{2\sigma_{b,\max}^2}{6\sigma_{b,\max}^2/\langle\sigma_b^2\rangle - 1 - \sqrt{-3 + 12\sigma_{b,\max}^2/\langle\sigma_b^2\rangle}}, \quad \text{(A12)}
$$

and

$$
\sigma_{b,\max}^2 = \sigma_{\text{ref}}^2 \frac{r_{\text{ref}}^2}{r_{\text{min}}^2} \quad \text{(A13)}
$$

are the limits of integration, which depend on the fixed (known) parameter $r_{\min}$ and the (unknown) population-averaged variance $\langle\sigma_b^2\rangle$.

#### d. SSI-reduced

Similar to the case of SSC, one can define a reduced version of the SSI likelihood, provided the number of samples per segment $N$ is large. The segment-dependent signal likelihood still requires a marginalization over the segment-dependent burst variances $\sigma_{b,I}^2$:

$$\mathcal{L}_s(d_I|\langle\sigma_b^2\rangle,\bar{\sigma}_{n_1}^2,\bar{\sigma}_{n_2}^2) = \int_{\sigma_{b,\min}^2(\langle\sigma_b^2\rangle)}^{\sigma_{b,\max}^2} d\sigma_{b,I}^2\,\pi(\sigma_{b,I}^2|\langle\sigma_b^2\rangle)$$
$$\times \frac{1}{\sqrt{2\pi\mathrm{var}(\bar{\sigma}_{b,I}^2)}}\exp\left[-\frac{(\hat{\sigma}_{b,I}^2-\sigma_{b,I}^2)^2}{2\mathrm{var}(\bar{\sigma}_{b,I}^2)}\right],$$
(A14)

where the prior and limits of integration are the same as those used for SSI-full. In addition,

$$\mathrm{var}(\bar{\sigma}_{b,I}^2) \equiv \frac{1}{N}\bar{\sigma}_{1,I}^2\bar{\sigma}_{2,I}^2 \qquad (A15)$$

with

$$\bar{\sigma}_{1,I}^2 \equiv \bar{\sigma}_{n_1}^2 + \sigma_{b,I}^2, \qquad \bar{\sigma}_{1,I}^2 \equiv \bar{\sigma}_{n_2}^2 + \sigma_{b,I}^2, \quad (A16)$$

where we estimate the white noise variances from the autocorrelated and cross-correlated power in the two detector outputs using the full set of data:

$$\bar{\sigma}_{\mathrm{gw}}^2 \equiv \hat{\sigma}_{\mathrm{gw}}^2\theta(\hat{\sigma}_{\mathrm{gw}}^2), \qquad (A17a)$$

$$\bar{\sigma}_{n_1}^2 \equiv (\hat{\sigma}_1^2 - \bar{\sigma}_{\mathrm{gw}}^2)\theta(\hat{\sigma}_1^2 - \bar{\sigma}_{\mathrm{gw}}^2), \qquad (A17b)$$

$$\bar{\sigma}_{n_2}^2 \equiv (\hat{\sigma}_2^2 - \bar{\sigma}_{\mathrm{gw}}^2)\theta(\hat{\sigma}_2^2 - \bar{\sigma}_{\mathrm{gw}}^2), \qquad (A17c)$$

where

$$\hat{\sigma}_{\mathrm{gw}}^2 \equiv \frac{1}{N_{\mathrm{tot}}}\sum_{I,i}d_{1,Ii}d_{2,Ii}, \qquad (A18a)$$

$$\hat{\sigma}_1^2 \equiv \frac{1}{N_{\mathrm{tot}}}\sum_{I,i}d_{1,Ii}^2, \qquad (A18b)$$

$$\hat{\sigma}_2^2 \equiv \frac{1}{N_{\mathrm{tot}}}\sum_{I,i}d_{2,Ii}^2. \qquad (A18c)$$

In the above expressions, $\theta(x)$ is the usual Heaviside step function, which is defined as $\theta(x) = 0$ or 1 depending on whether $x < 0$ or $x > 0$, and the hatted quantities $\hat{\sigma}_{\mathrm{gw}}^2$, $\hat{\sigma}_1^2$, $\hat{\sigma}_2^2$ are the quadratic combinations of the data in the two detectors. This simplification is possible since the simulated noise is stationary.

The segment-dependent noise likelihood $\mathcal{L}_n(d_I|\bar{\sigma}_{n_1}^2,\bar{\sigma}_{n_2}^2)$ is given as before by

$$\mathcal{L}_n(d_I|\bar{\sigma}_{n_1}^2,\bar{\sigma}_{n_2}^2) = \sqrt{\frac{N}{2\pi\bar{\sigma}_{n_1}^2\,\bar{\sigma}_{n_2}^2}}\exp\left[-\frac{N}{2}\frac{(\hat{\sigma}_{b,I}^2)^2}{\bar{\sigma}_{n_1}^2\,\bar{\sigma}_{n_2}^2}\right]. \quad (A19)$$

### 2. Likelihoods for colored signals

The signal and noise dependent likelihoods for SSI are specified in Sec. III for both the full (infer noise parameters) and reduced (use estimated noise parameters) analyses. When analyzing stochastic bursts (Sec. IV B) and deterministic chirps (Sec. IV C), the segment prior and integration bounds are specified in (39) and the subsequent paragraph.

#### a. SSC-full

For the continuous search, SSC, the full likelihood is specified by

$$\mathcal{L}(d|\Omega_{\mathrm{gw}},\sigma_{n_1}^2,\sigma_{n_2}^2) = \prod_{I=1}^{N_{\mathrm{seg}}}\prod_k \frac{1}{(\pi T/2)^{2M}(P_1(f_k)P_2(f_k) - P_{\mathrm{gw}}^2(f_k))^M}$$
$$\times \exp\left\{-\frac{M}{(P_1(f_k)P_2(f_k) - P_{\mathrm{gw}}^2(f_k))}[\hat{P}_{1,Ik}P_2(f_k) + \hat{P}_{2,Ik}P_1(f_k) - 2\hat{P}_{\mathrm{gw},Ik}P_{\mathrm{gw}}(f_k)]\right\}, \quad (A20)$$

where

$$P_1(f) \equiv P_{n_1}(f) + P_{\mathrm{gw}}(f),$$
$$P_2(f) \equiv P_{n_2}(f) + P_{\mathrm{gw}}(f), \qquad (A21)$$

with

$$P_{n_1}(f) \equiv \frac{\sigma_{n_1}^2}{(f_{\mathrm{high}} - f_{\mathrm{low}})}, \qquad (A22a)$$

$$P_{n_2}(f) \equiv \frac{\sigma_{n_2}^2}{(f_{\mathrm{high}} - f_{\mathrm{low}})}, \qquad (A22b)$$

$$P_{\mathrm{gw}}(f) \equiv \Omega_{\mathrm{gw}}H(f), \qquad (A22c)$$

and $H(f)$ is given by (17). Note that the population parameter for SSC is $\Omega_{gw}$, the time and population-averaged energy density amplitude. In addition, the data enter the signal evidence via the same quadratic combinations as for SSI-full [see (18)], but with the cross-correlation combination now defining $\hat{P}_{gw,Ik}$ as opposed to $\hat{P}_{b,Ik}$.

### b. SSC-reduced

For SSC-reduced, we have [23]

$$\mathcal{L}(d|\Omega_{gw}, \bar{\sigma}_{n_1}^2, \bar{\sigma}_{n_2}^2) = \prod_{I=1}^{N_{seg}} \frac{1}{\sqrt{2\pi \mathrm{var}(\bar{\Omega}_{gw})}}$$
$$\times \exp\left[-\frac{(\hat{\Omega}_{gw,I} - \Omega_{gw})^2}{2\mathrm{var}(\bar{\Omega}_{gw})}\right], \qquad (A23)$$

where

$$\hat{\Omega}_{gw,I} \equiv \frac{\sum_k Q(f_k)\hat{P}_{gw,Ik}}{\sum_{k'} Q(f_{k'})H(f_{k'})},$$
$$\mathrm{var}(\bar{\Omega}_{gw}) \equiv \left(2M\sum_k Q(f_k)H(f_k)\right)^{-1} \qquad (A24)$$

are the optimally filtered cross-correlation estimators and corresponding variances, which are constructed from coarse-grained estimates of the cross-correlated power $\hat{P}_{gw,Ik}$, and the optimal filter function

$$Q(f) \equiv \frac{H(f)}{\bar{P}_1(f)\bar{P}_2(f)}. \qquad (A25)$$

In the above expression,

$$\bar{P}_1(f) \equiv \frac{\bar{\sigma}_{n_1}^2}{(f_{high} - f_{low})} + \bar{\Omega}_{gw}H(f),$$
$$\bar{P}_2(f) \equiv \frac{\bar{\sigma}_{n_2}^2}{(f_{high} - f_{low})} + \bar{\Omega}_{gw}H(f), \qquad (A26)$$

where $\bar{\sigma}_{n_1}^2$ and $\bar{\sigma}_{n_2}^2$ are measured estimates of the detector noise power as defined in (A16) and $\bar{\Omega}_{gw}$ is related to $\bar{\sigma}_{gw}^2$ [also defined in (A16)] via

$$\bar{\Omega}_{gw} = \frac{4}{3}\frac{\bar{\sigma}_{gw}^2}{f_{ref}}\left(\frac{3H_0^2}{10\pi^2}\frac{1}{f_{ref}^3}\right)^{-1}\left[\left(\frac{f_{ref}}{f_{low}}\right)^{4/3} - \left(\frac{f_{ref}}{f_{high}}\right)^{4/3}\right]^{-1}. \qquad (A27)$$

This last equation follows from the general relation between variance and power spectrum,

$$\sigma_{gw}^2 \equiv \int_{f_{low}}^{f_{high}} df\, P_{gw}(f) = \Omega_{gw} \int_{f_{low}}^{f_{high}} df\, H(f)$$
$$= \Omega_{gw}\left(\frac{3H_0^2}{10\pi^2}\frac{1}{f_{ref}^3}\right)\int_{f_{low}}^{f_{high}} df\left(\frac{f}{f_{ref}}\right)^{-7/3}. \qquad (A28)$$

### c. DSI-full

We also analyze the colored data with DSI, our much simpler implementation of the deterministic-signal-based search. Following [19] for two detectors, we define the DSI segment-dependent signal likelihood to be

$$\mathcal{L}_s(d_I|r_{max}, \sigma_{n_1}^2, \sigma_{n_2}^2) \propto \int_{r_{min}}^{r_{max}} dr_I\, \pi(r_I|r_{max})\exp\left\{-\frac{1}{2}(4\Delta f)\sum_k\sum_{\mu=1,2}\frac{|(\tilde{d}_{\mu,Ik} - \tilde{h}_{chirp}(r_I;f_k))|^2}{P_{n_\mu}}\right\}, \qquad (A29)$$

where $\tilde{d}_{\mu,Ik}$ and $\tilde{h}_{chirp}(r_I;f_k)$ are the Fourier transforms of the data and chirp waveform, respectively, with all of the other chirp parameters assumed to be known *a priori*. In the above signal evidence, we are marginalizing over the segment-dependent source distance $r_I$, which is drawn from a uniform-in-volume distribution $\pi(r_I|r_{max})$ as given by (27).

By taking $\tilde{h}_{chirp}(r_I;f_k) = 0$ (corresponding to no signal in the data) the corresponding segment-dependent noise likelihood is

$$\mathcal{L}_n(d_I|\sigma_{n_1}^2, \sigma_{n_2}^2) \propto \exp\left\{-\frac{1}{2}(4\Delta f)\sum_k\sum_{\mu=1,2}\frac{|\tilde{d}_{\mu,Ik}|^2}{P_{n_\mu}}\right\}. \qquad (A30)$$

#### d. DSI-reduced

For the reduced implementation, we substitute the noise parameters with the autocorrelated power estimates which give

$$\mathcal{L}_s(d_I | r_{\max}, \bar{\sigma}_{n_1}^2, \bar{\sigma}_{n_2}^2) \propto \int_{r_{\min}}^{r_{\max}} dr_I \pi(r_I | r_{\max}) \exp\left\{ -\frac{1}{2}(4\Delta f) \sum_k \sum_{\mu=1,2} \frac{|(\tilde{d}_{\mu,Ik} - \tilde{h}_{\text{chirp}}(r_I; f_k))|^2}{\bar{P}_{n_\mu}} \right\} \tag{A31}$$

and

$$\mathcal{L}_n(d_I | \bar{\sigma}_{n_1}^2, \bar{\sigma}_{n_2}^2) \propto \exp\left\{ -\frac{1}{2}(4\Delta f) \sum_k \sum_{\mu=1,2} \frac{|\tilde{d}_{\mu,Ik}|^2}{\bar{P}_{n_\mu}} \right\} \tag{A32}$$

for the segment-dependent signal and noise likelihoods, respectively.

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

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

[3] T. Akutsu Knkai, Overview of KAGRA: Detector design and construction history, Prog. Theor. Exp. Phys. **2021**, 05A101 (2020).

[4] R. Abbott *et al.* (The LIGO Scientific, Virgo, and KAGRA Collaborations), GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run, arXiv:2111.03606.

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

[6] R. Abbott *et al.*, Observation of gravitational waves from two neutron star-black hole coalescences, Astrophys. J. **915**, L5 (2021).

[7] Carl W. Helstrom, *Statistical Theory of Signal Detection, 2nd edition* (Pergamon Press, Oxford, London, Edinburgh, New York, Toronto, Sydney, Paris, Braunschweig, 1968).

[8] L. A. Wainstein and V. D. Zubakov, *Extractions of Signals from Noise* (Dover Publications Inc., Mineola, 1971).

[9] Nelson Christensen, Stochastic gravitational wave backgrounds, Rep. Prog. Phys. **82**, 016903 (2018).

[10] Nick van Remortel, Kamiel Janssens, and Kevin Turbang, Stochastic gravitational wave background: Methods and implications, Prog. Part. Nucl. Phys. **128**, 104003 (2023).

[11] P. F. Michelson, On detecting stochastic background gravitational radiation with terrestrial detectors, Mon. Not. R. Astron. Soc. **227**, 933 (1987).

[12] B. P. Abbott *et al.* (The LIGO Scientific, Virgo, and KAGRA Collaborations), Prospects for observing and localizing gravitational-wave transients with Advanced LIGO, Advanced Virgo and KAGRA, Living Rev. Relativity **23**, 3 (2020).

[13] R. Abbott *et al.* (LIGO Scienic, VIRGO, KAGRA), The population of merging compact binaries inferred using gravitational waves through GWTC-3, arXiv:2111.03634.

[14] B. P. Abbott *et al.*, Binary black hole population properties inferred from the first and second observing runs of Advanced LIGO and Advanced Virgo, Astrophys. J. Lett. **882**, L24 (2019).

[15] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and V. B. Adya, GW170817: Implications for the Stochastic Gravitational-Wave Background from Compact Binary Coalescences, Phys. Rev. Lett. **120**, 091101 (2018).

[16] Bruce Allen and Joseph D. Romano, Detecting a stochastic background of gravitational radiation: Signal processing strategies and sensitivities, Phys. Rev. D **59**, 102001 (1999).

[17] Duncan Meacher, Michael Coughlin, Sean Morris, Tania Regimbau, Nelson Christensen, Shivaraj Kandhasamy, Vuk Mandic, Joseph D. Romano, and Eric Thrane, Mock data and science challenge for detecting an astrophysical stochastic gravitational-wave background with Advanced LIGO and Advanced Virgo, Phys. Rev. D **92**, 063002 (2015).

[18] Steve Drasco and Éanna É. Flanagan, Detection methods for non-gaussian gravitational wave stochastic backgrounds, Phys. Rev. D **67**, 082003 (2003).

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

[20] Tania Regimbau, The astrophysical gravitational wave stochastic background, Res. Astron. Astrophys. **11**, 369 (2011).

[21] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), Upper limits on the isotropic gravitational-wave background from Advanced LIGO and Advanced Virgo's third observing run, Phys. Rev. D **104**, 022004 (2021).

[22] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW170817: Implications for the Stochastic Gravitational-Wave Background from Compact Binary Coalescences, Phys. Rev. Lett. **120**, 091101 (2018).

[23] Andrew Matas and Joseph D. Romano, Frequentist versus bayesian analyses: Cross-correlation as an approximate sufficient statistic for LIGO-Virgo stochastic background searches, Phys. Rev. D **103**, 062003 (2021).

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

[25] N. Christensen, Measuring the stochastic gravitational-radiation background with laser-interferometric antennas, Phys. Rev. D **46**, 5250 (1992).

[26] Éanna É. Flanagan, Sensitivity of the laser interferometer gravitational wave observatory to a stochastic background, and its dependence on the detector orientations, Phys. Rev. D **48**, 2389 (1993).

[27] M. Maggiore, *Gravitational Waves, Vol. 1, Theory and Experiments* (Oxford University Press, New York, 2008).

[28] Takahiro S. Yamamoto, Sachiko Kuroyanagi, and Guo-Chin Liu, Deep learning for intermittent gravitational wave signals, Phys. Rev. D **107**, 044032 (2023).

[29] Riccardo Buscicchio, Anirban Ain, Matteo Ballelli, Giancarlo Cella, and Barbara Patricelli, Detecting non-gaussian gravitational wave backgrounds: A unified framework, Phys. Rev. D **107**, 063027 (2023).

[30] Suvodip Mukherjee and Joseph Silk, Time-dependence of the astrophysical stochastic gravitational wave background, Mon. Not. R. Astron. Soc. **491**, 4690 (2020).

[31] Suvodip Mukherjee and Joseph Silk, Fundamental physics using the temporal gravitational wave background, Phys. Rev. D **104**, 063518 (2021).

[32] D. Davis *et al.*, LIGO detector characterization in the second and third observing runs, Classical Quantum Gravity **38**, 135014 (2021).

[33] Gregory Ashton *et al.*, BILBY: A user-friendly bayesian inference library for gravitational-wave astronomy, Astrophys. J. Suppl. Ser. **241**, 27 (2019).

[34] Joshua S. Speagle, DYNESTY: A dynamic nested sampling package for estimating bayesian posteriors and evidences, Mon. Not. R. Astron. Soc. **493**, 3132 (2020).