Multiparameter tests of general relativity using multiband gravitational-wave observations

Anuradha Gupta,[1,2,*] Sayantani Datta,[3] Shilpa Kastha,[4] Ssohrab Borhanian,[2] K. G. Arun,[3,2] and B. S. Sathyaprakash[2,5,6]

[1]Department of Physics and Astronomy, The University of Mississippi, Oxford MS 38677, USA
[2]Institute for Gravitation and the Cosmos, Department of Physics, Penn State University, University Park PA 16802, USA
[3]Chennai Mathematical Institute, Siruseri, 603103, India
[4]Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Callinstraße 38, D–30167 Hannover, Germany
Leibniz Universität Hannover, D–30167 Hannover, Germany
[5]Department of Astronomy and Astrophysics, Penn State University, University Park PA 16802, USA
[6]School of Physics and Astronomy, Cardiff University, Cardiff, CF24 3AA, United Kingdom

(Dated: May 20, 2020)

In this Letter we show that multiband observations of stellar-mass binary black holes by the next generation of ground-based observatories (3G) and the space-based Laser Interferometer Space Antenna (LISA) would facilitate a comprehensive test of general relativity by simultaneously measuring all the post-Newtonian (PN) coefficients. Multiband observations would measure most of the known PN phasing coefficients to an accuracy below a few percent—two orders-of-magnitude better than the best bounds achievable from even 'golden' binaries in the 3G or LISA bands. Such multiparameter bounds would play a pivotal role in constraining the parameter space of modified theories of gravity beyond general relativity.

Introduction: Gravitational wave (GW) observations have provided a first glimpse of the strong-field dynamics of binary black holes (BBHs) [1, 2]. They have also allowed us to place the first ever constraints on the possible departures from general relativity (GR) [3, 4] in this regime. Parametrized tests of the post-Newtonian (PN) approximation to GR [5–8] are among the most important theory-agnostic, null-tests of GR that are performed using GW observations. These tests make use of the analytical prediction of the structure of the phase evolution using the PN approximation to GR [9]. In the PN approximation the dynamics of the binary is treated as an adiabatic process and Einstein's field equations are solved under the assumption of *slow motion* and *weak gravitational fields*. This is an excellent approximation for the "inspiral" phase of the compact binary dynamics where the two stars spiral-in under the influence of radiation back reaction, but the time-scale of radiation reaction is large compared to the orbital time-scale.

Gravitational waveform from a compact binary coalescence, in the frequency domain, have the well-known form [10]

$$\tilde{h}(f) = \mathcal{A} f^{-7/6} e^{i\Phi(f)}\,, \tag{1}$$

where $\Phi(f)$ is the frequency domain phase of the emitted signal and $\mathcal{A}$ is the signal's amplitude. For inspiralling binaries in quasi-circular orbits, the waveform depends on the binary's masses, spins, distance, sky position, and the orientation of its orbit. More explicitly, the phase takes the form

$$\Phi(f) = 2\pi f\, t_c - \phi_c + \frac{3}{128\,\eta\, v^5}\left[\sum_{k=0}^{K} \phi_k\, v^k + \sum_{kl=0}^{K} \phi_{kl}\, v^{kl} \ln v\right], \tag{2}$$

where $v = (\pi m f)^{1/3}$ denotes the PN expansion parameter, m denotes the binary's total mass, and ϕ_{kl} and ϕ_k denote the logarithmic and non-logarithmic phasing coefficients, respectively. The PN coefficients are currently known up to 3.5 order in the PN expansion [11–14], which corresponds to $K = 7$ in the above equation. The parameters t_c and ϕ_c are the epoch when the signal's amplitude at the detector is the greatest and the phase of the signal at that epoch, respectively. For BBHs on quasi-circular orbits, the PN coefficients ϕ_k and ϕ_{kl} are functions of the component masses and spins. The assumption of a quasi-circular orbit is an excellent approximation for majority of the stellar-mass BBHs [15].

The parametrized tests rely on the unique prediction for the PN coefficients ϕ_k and ϕ_{kl} in GR and use GW BBH merger events to constrain possible departures of the coefficients from their GR prediction. A parametrized waveform replacing the GR phasing coefficients ϕ_a with $\phi_a(1 + \delta\hat{\phi}_a)$ $(a = k, kl)$ is employed for the test [8]. By construction, the *deformation* parameters $\delta\hat{\phi}_a$ are identically equal to zero in GR, while in a modified theory of gravity one or more of these parameters can deviate from zero. Thus, GW data allow the direct measurement of the PN coefficients and if their deformations are found to be consistent with zero, the uncertainty associated with the measurement provides an upper limit on the deviation of these parameters from their GR values.

Status of parametrized tests of post-Newtonian theory: Combining data for the ten BBH merger events found during the first and second observing runs of LIGO/Virgo, the current bound on the eight PN deformation parameters are given in Fig. 4 of Ref. [4]. Moreover, the bounds from this theory-agnostic test have been mapped onto specific modified theories of gravity in Ref. [16]. However, there is an important caveat while using these bounds to constrain a modified theory of gravity: The bound on the deviation from a particular PN coefficient reported in Ref. [4] is derived assuming that *all* the deformation parameters *except* the one that is being tested follow the predictions of GR with $\delta\hat{\phi}_a = 0$. This assumption is necessary because the most general test wherein all the PN coefficients are simultaneously measured yields very poor or no bounds due to the strong degree of covariance among the deformation parameters and the intrinsic parameters of the binary [17]. Hence one is compelled to replace this most general test with a series of tests wherein only one deformation parameter is varied at a time together with, of course, the intrinsic

* agupta1@olemiss.edu

parameters of the binary. This restricted suite of tests can still be expected to detect a deviation away from GR by finding statistically significant offsets away from zero in one or more of the PN deformation parameters [6, 8].

It is, however, incorrect to use the results from the single-parameter tests to constrain a specific modified theory of gravity, for two reasons. Firstly, any deviation from GR inferred for a particular PN coefficient cannot be attributed *uniquely* to a breakdown of GR at that PN order. This is because the waveform is largely degenerate in the PN coefficients. Consequently, deviation at a particular PN order can be captured by deformation of the coefficient at a different PN order. Hence a deviation in one or more of the PN coefficients in a set of tests does not necessarily give any fundamental insight into the true nature of the underlying theory of gravity. Secondly, if the single-parameter tests are all consistent with GR, the widths of the posterior distributions of the PN coefficients cannot be used to constrain the parameter space of modified theories of gravity. This is due to the expectation from effective field theoretic arguments, that deviations from GR, in a specific modified theory of gravity, show up starting *from* a certain PN order (see, for instance, [18, 19]). Therefore, to map the PN deformation parameters to the free parameters of a specific modified theory of gravity it is necessary to perform the most general multiparameter test. In other words, single-parameter tests would lead to an underestimation of the errors and hence yield bounds that are more stringent than what one might infer with multiparameter tests.

In this *Letter* we will show that combining data from the next generation (3G) of ground-based detectors, such as the Cosmic Explorer (CE) [20] and Einstein Telescope (ET) [21], with the space-based LISA observatory [22] is likely the only viable route to carry out this very challenging, but very general test of GR. Such tests are crucial to set reliable constraints on the parameter space of modified theories of gravity. Specifically, we demonstrate that multiband observations of a subclass of stellar-mass BBHs by LISA and CE would provide a unique opportunity to carry out the multiparameter test of PN theory. Combining the low-frequency sensitivity of LISA with the high-frequency sensitivity of CE helps in lifting the large degeneracies that prevent the use of multiparameter tests in either of these observatories. To demonstrate the advantage of multiparameter tests using multibanding we simulate a stellar-mass population of BBHs that obey the mass-, rate- and redshift-distribution inferred from the first and second observing runs of LIGO and Virgo. In a companion paper [23], we will discuss intermediate-mass BBHs as another important class of sources for multiband, multiparameter test of GR, although the bounds from stellar-mass BBHs are far better than their intermediate-mass counterparts [23].

Multiband visibility of stellar-mass binaries: The planned LISA observatory is sensitive to GWs in the frequency range ~ 0.1–100 mHz and the proposed 3G observatories (e.g. CE, which we have used in this paper as a representative of 3G detectors), will be sensitive in the frequency range ~ 1 Hz – 5 kHz. Though LISA is more sensitive to mergers of supermassive BBHs of millions of solar masses, it has been argued that the detection of stellar-mass BBHs using LISA would be possible despite the small signal-to-noise ratios (SNRs) [24–26] and would be of immense importance to astronomy and fundamental physics, as the mergers of these binaries would be detectable by the ground-based detectors operating at the same time. Observation of sources at earlier stages of their evolution in LISA, and later, more nonlinear, stages in 3G detectors is referred to as *multiband* observation.

Several authors have investigated the value added by multiband observations of GW sources. For example, Ref. [27–29] examined the projected constraints on the bounds on dipolar GW radiation, Ref. [30] investigated the bounds on single-parameter tests of GR and Ref. [31] studied the constraints on the parameter space of modified theories of gravity using multiband observations. These authors have used prototypical BBH systems, such as GW150914 [32], which will have good multiband visibility, and have studied the corresponding bounds for single-parameter tests of GR.

Here, we consider 5×10^5 BBHs corresponding to one year of CE observation [33], distributed uniformly in comoving volume up to redshift $z = 10$. The primary black hole masses are assumed to follow a power-law distribution with the power-law index $\alpha = 1.6$ (i.e., $p(m_1) \propto m_1^{-\alpha}$) in the mass range $[5, 100]M_\odot$ while secondary masses are uniform in the same mass range [34]. We assume the binary components to possess spins which are aligned or anti-aligned with respect to the orbital angular momentum vector. This assumption is consistent with the fact that none of the BBHs detected during the first and second observing runs of LIGO/Virgo showed evidence for spins misaligned with the binary's orbital angular momentum. The Kerr parameter of the companion black holes are drawn from two different distributions: (1) a uniform distribution in the range $[0, 1]$ and (2) a Gaussian with mean 0 and standard deviation 0.1.

3G detectors will be able to observe stellar-mass BBH mergers up to the epoch of the formation of first stars. The question, is what fraction of events detected by CE will have LISA counterparts. This joint population will be limited by the SNR in the LISA band. In Fig. 1 we plot the SNR distribution in LISA for this population. As expected, only a small subset of the population will have an SNR greater than 4. Such events will have an SNR of at least 2000 in CE, facilitating a very accurate measurement of the binary parameters, which in turn helps in digging the signals out of the LISA background noise. (See the Supplement for a discussion on the detectability of stellar-mass BBH signals in LISA with an SNR threshold as low as 4.) We find that among the hundreds of thousands of stellar-mass BBH merger that would be observable by CE in one year, $\sim$200 would cross this threshold and permit multiband, multiparameter tests of GR. These $\sim$200 BBHs would spend roughly 4.5 days to 7 weeks outside the LISA band before entering the CE band and eventually merge.

Multiparameter tests of GR via multiband GW observations: We now describe the efficacy of the multiparameter tests of GR using the population of $\sim$200 BBH merger events detectable by both CE and LISA. Our method here is based on the well-known Fisher information matrix which enables the computation of the projected 1σ errors on the various parameters describing a signal model for a given sensitivity of the detector configuration [35–37]. We use the sensitivity curves of CE and LISA given in [20] and [22], respectively. For simplicity, we do not consider the orbital motion of LISA as it is likely to have negligible impact on the parameter estimation of the

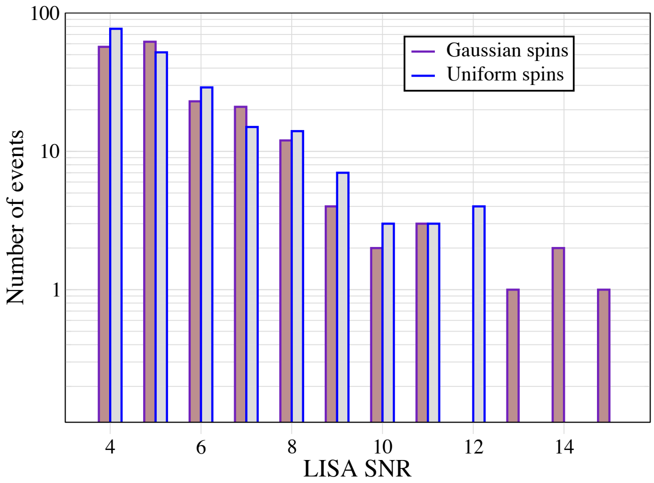

FIG. 1. Distribution of the SNR of stellar-mass BBHs in the LISA band using the mass and redshift distribution as inferred from the first and second observing runs, and spins following (i) a uniform distribution and (ii) a Gaussian distribution with mean 0 and standard deviation 0.1. Only SNR ≥ 4 events are shown. A small fraction of about ~200 of all sources (some 500,000) observed by CE in a year have SNR ≥ 4 in the LISA band. The plot also shows that the spin distribution of black holes does not have a significant effect on the visibility of stellar-mass BBHs in LISA.

intrinsic parameters of the binary, which is of interest here. Stellar-mass BBHs that merge in the CE band are assumed to have been observed for five years in LISA and the starting frequency for the signal in LISA is chosen accordingly following the prescription given in Eq. (2.15) of Ref. [38].

We employ the IMRPhenomD [39, 40] waveform model, a frequency-domain phenomenological model describing the complete inspiral-merger-ringdown phases of BBH systems. The waveform amplitude in this model is truncated at the quadrupolar order and we have introduced additional deformation parameters $\delta\hat{\phi}_a$ in the phase at different PN orders in the inspiral part of the waveform. We have set the four angles corresponding to the sky position of the binary and the orientation of its orbit with respect to the line-of-sight to zero. This amounts to assuming that the binaries are optimally located and oriented with respect to the detectors. Note, however, that the LISA sensitivity curve that we use is averaged over the sky and the polarization angle and we have included a factor of $\sqrt{4/5}$ in the calculation of the SNR and the Fisher matrix to account for the averaging over the inclination angle [41].

The Fisher information matrix for a single detector (CE or LISA) is defined as

$$\Gamma^{(0)}_{\alpha\beta} = \langle \tilde{h}_\alpha, \tilde{h}_\beta \rangle, \tag{3}$$

where $\tilde{h}(f;\vec{\theta})$ is the GW signal defined by a set of parameters $\vec{\theta}$, $\tilde{h}_\alpha = \partial\tilde{h}(f;\vec{\theta})/\partial\theta_\alpha$, and the angular bracket $\langle\, ,\, \rangle$ denotes the noise-weighted inner product defined by

$$\langle a, b\rangle = 2\int_{f_{\rm low}}^{f_{\rm high}} \frac{a(f)\, b^*(f) + a^*(f)\, b(f)}{S_h(f)}\, df\,, \tag{4}$$

where $S_h(f)$ is the one-sided noise power spectral density (PSD) of the detector and $f_{\rm low}, f_{\rm high}$ are the lower and upper limits of integration. For CE the lower limit of integration is taken to be 5 Hz and the upper frequency cut-off is chosen such that the characteristic amplitude $(2\sqrt{f}|\tilde{h}(f)|)$ of the GW signal is lower than that of the CE noise by 10% at maximum.

In order to combine the information from LISA and CE, we construct a multiband Fisher matrix by simply adding the Fisher matrices for the individual detectors, with the corresponding variance-covariance matrix $C^{\alpha\beta}$ defined by the inverse of the multiband Fisher matrix:

$$\Gamma_{\alpha\beta} = \Gamma^{\rm CE}_{\alpha\beta} + \Gamma^{\rm LISA}_{\alpha\beta}, \quad C^{\alpha\beta} = (\Gamma^{-1})^{\alpha\beta}, \tag{5}$$

The diagonal components, $C^{\alpha\alpha}$, are the variances of θ^α and the 1σ errors on θ^α are $\sigma^\alpha = \sqrt{C^{\alpha\alpha}}$.

The errors σ_a, where $a = 1, 2, \cdots 8$ denote the deformation parameters that are tested simultaneously, are obtained for each event in the population for different choices of the number of test parameters $\delta\hat{\phi}_a$, $a = 1, 2, \cdots, 8$. The bounds on the individual events are combined to obtain a net constraint by using the standard formula

$$\sigma_a^{-2} = \sum_{n=1}^{N} \left(\sigma_a^{(n)}\right)^{-2}, \tag{6}$$

where $n = 1, \ldots, N$ denotes the events in the BBH population and N is their total number.

Following Refs. [37, 42] we also add a prior matrix Γ^p to the Fisher information matrix, $\Gamma^{(0)}$, in order to account for certain properties of the signals that we assume. Specifically, we assume that the priors on the spin magnitudes and the phase of coalescence as $\Gamma^p_{\chi_1\chi_1} = \Gamma^p_{\chi_2\chi_2} = (0.5)^{-2}$ and $\Gamma^p_{\phi_c\phi_c} = (\pi)^{-2}$, respectively and all other elements of the prior matrix are set to zero. The Gaussian prior on spin magnitudes is a good approximation to the low-component spins of the BBHs reported in Ref. [2]. The prior on ϕ_c is somewhat adhoc, but helps the Fisher matrix to be better conditioned. We have verified that this choice of prior does not alter our conclusions reported in this paper. We now invert the resulting Fisher matrix given by $\Gamma_{\alpha\beta} = \Gamma^0_{\alpha\beta} + \Gamma^p_{\alpha\beta}$ to deduce the error bars.

Results and Discussions: Our main results combining LISA and CE observations of stellar-mass BBHs are summarized in Figure 2. As we increase the number of PN coefficients that are simultaneously tested, starting from the Newtonian order, the 1σ upper bounds on them are presented in the figure. For instance, the filled circles are the bounds where only one PN deformation parameter is estimated at a time, whereas the octagons denote the bounds when all the eight parameters are simultaneously estimated. In the eight parameter case, all the parameters are measured with an accuracy ~20%, of which the first three may be measured with an accuracy better than 1%, whereas the first two PN coefficients may yield bounds ~0.1%. Hence multiband observations of stellar-mass BBHs would permit us to test modified theories of gravity, which predict deviations at orders below 3PN to a precision less than ~ 1%.

One may notice interesting trends in the bounds as we add more and more parameters. The bounds on 0PN and 1PN deformation coefficient from 2-parameter estimation case are ~0.01%. The inclusion of the 1.5PN deformation coefficient results in a sudden worsening of the bounds by an order of magnitude. This may be understood by noting that 1.5PN is

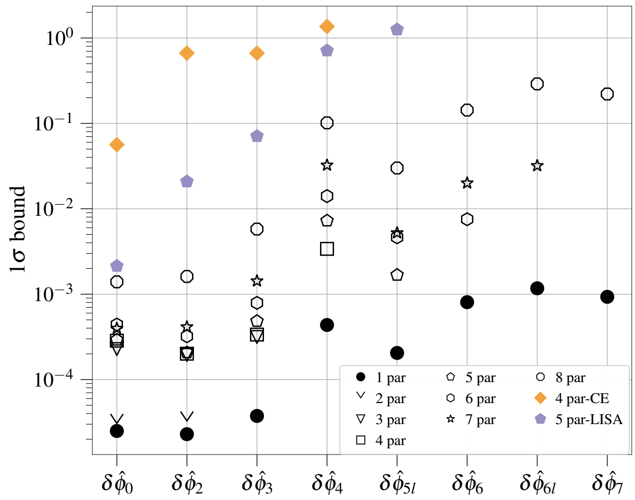

FIG. 2. Multiparameter tests using multiband observations with LISA and CE. Shown are combined 1σ bounds on various PN coefficients starting from 0PN through 3.5PN in the inspiral phase of the signal while measuring many of them together at a time. Different types of markers symbolize how many PN coefficients were constrained simultaneously. For example, '●' represents 'one PN coefficient at a time', '∨' represents 'two PN coefficients at a time', and so on. The figure represents results for the BBH population having Gaussian spin distribution, we get similar estimates for a uniform spin distribution. The filled diamonds and pentagons are bounds obtained with CE and LISA, respectively, on the first four and five PN coefficients from their respective golden binaries, respectively. The total masses of the CE and LISA golden binaries are 200 $M_\odot$ and 6.6×10^5 $M_\odot$, both binaries are 1 Gpc away and have component spins $\chi_1 = 0.6, \chi_2 = 0.5$.

the order at which spins first appear in the phasing formula. Adding a deformation parameter at that order, that is completely degenerate with spins, adversely affects the overall parameter estimation, which gets reflected in the bounds on the first two PN coefficients. The gradual worsening of the bounds as we go to even higher multiparameter tests is simply due to the increasing degeneracy brought in by each of the additional PN deformation parameters.

It can be seen that even in the era of 3G detectors we cannot obtain meaningful constraints with multiparameter tests. As is evident from the figure, for golden binaries in CE—binaries that have the smallest error for the multiparameter tests— the errors on $\delta\phi_a$, are $\sim$100% only for $a = 1, \ldots, 4$; if we vary more than four parameters at a time then the errors on PN coefficients with $a \neq 0$ are larger than 100%. In a year's time CE will observe a handful of such golden binaries and the joint error that one can obtain by combining golden binaries will still not be significantly smaller. Consequently, ground-based detectors alone cannot break the degeneracy among different PN coefficients. The same is true with LISA observations of supermassive BBHs. Even with a golden supermassive BBH we can perform the multiparameter test with only five parameters and LISA is not likely to observe more than a handful of such binaries over a 5 year period.

Conclusions: To conclude, we have shown the importance of multiband observations of GWs to carry out the multiparameter tests of GR. From our systematic study of a representative set of systems, we have also found that even for the best case scenario, observations of supermassive BBHs in the LISA band or stellar- or intermediate-mass BBHs in the CE band would not be able to place constraints as good as the one reported here. Hence multibanding would, perhaps, be the only way to carry out this test which in turn is necessary to make meaningful constraints on the parameter space of modified theories of gravity. As LIGO and Virgo detect several more BBHs in the future observing runs, the merger rate and the mass distribution would be more tightly constrained which is likely to further tighten the bounds derived here making this test an excellent science case for multiband observations.

Acknowledgments: We thank Chris Van Den Broeck, Bala Iyer, Arnab Dhani and M. Saleem for several useful discussions. B.S.S. is supported in part by NSF Grant No. PHY-1836779, AST-1716394 and AST-1708146. S.B. is supported by NSF Grant No. PHY-1836779. K.G.A. and S. D. are partially supported by a grant from the Infosys Foundation. They also acknowledge the Swarnajayanti grant DST/SJF/PSA-01/2017-18 DST-India. K.G.A acknowledges Core Research Grant EMR/2016/005594 of SERB. We thank all frontline workers combating the CoVID-19 pandemic without whose support this work would not have been possible.

Appendix: Supplemental Materials

In this Supplement we provide a discussion of the detectability of gravitational waves from stellar-mass binary black holes (BBHs) by the Laser Interferometer Space Antenna (LISA), an alternative to the multiparameter test presented in the paper and the accuracy of Fisher matrix inversion.

Archival searches for stellar-mass BBHs in the LISA data: Ground-based observatories such as the Cosmic Explorer (CE) [20] and Einstein Telescope (ET) [21], have the best sensitivity to stellar-mass BBHs of 10-100 $M_\odot$ and can detect them up to redshifts of $z \sim 10$ and beyond. A small fraction of such BBHs that are close enough will also be observable by LISA [22]. The observability of a source depends on the false alarm rate at a given signal-to-noise-ratio (SNR) and the number of trials needed to dig out the signal.

As an example, there is only a chance of 0.13% (i.e., a p-value of 0.0013) that a single draw from a Gaussian distribution with zero mean and unit variance yields a number larger than 3. On the other hand, multiple draws from the same distribution increases the p-value for getting a number larger than 3. Likewise, if it is necessary to carry out a blind search for GW signals without any knowledge of their parameters then one ought to employ a large number of templates which makes it computationally expensive to dig out weaker signals from a noisy background. Indeed, Ref. [43] argued that as many as 10^{40} templates would be needed to dig out a stellar-mass BBH signal from LISA data. This would require a matched filter SNR ≥ 14 for a p-value of 10^{-3}.

A third-generation (3G) network of CE and ET can aid in searching for stellar-mass BBH signals in LISA data as the former will detect them with extremely high fidelity and can therefore provide a very tight prior on the source's extrinsic (sky position, polarization, orientation of the orbit, and luminosity distance) as well as intrinsic (masses and spins of the companion black holes) parameters. Therefore, an archival search for signals can greatly reduce the number of templates/trials required and hence enhance their detectability in the LISA data.

The analyses in Ref. [43] assumed for the LISA archival search that the trigger time, the epoch when the signal's amplitude reaches its maximum value in the detector, is known precisely, while the errors in the intrinsic and other extrinsic parameters will be smaller in 3G detectors than those measured by LIGO [1] by a factor of 10. This contracted volume of the search space decreases the number of templates by a factor of 10^{29} compared to a blind search [43]. Yet the result is that LISA would need $\sim 10^{11}$ templates to identify a stellar-mass BBH signal, or an SNR threshold of ~9 for a p-value of 10^{-3}. Other studies [44] have shown that using an alternate method one can detect GW150914-like source even with an SNR~7 for the same p-value. We have computed the number of templates for the events considered in this work and find that LISA archival searches require far fewer templates owing to the greatly contracted search volume thanks to the high-fidelity of 3G observations.

In one year, the ground-based 3G network of CE and ET would find ~200 events that would have an SNR of 4 or more in LISA. For each such event we estimated the number of templates necessary for an archival search in LISA data. To this end, we computed the Fisher information matrix and its inverse—the variance-covariance matrix—to obtain the expected error in the measurement of various extrinsic and intrinsic parameters using a 3G network of two CE detectors, one each in the USA and Australia, and one ET in Europe. Fig. 3 plots a subset of our results: the distribution of the SNR, the error in the measurement of the chirpmass, and the uncertainty in the sky position (the left three panels). In making these plots we have rejected the worst 5% of the outliers in our simulation.

Due to the large SNR in the 3G network the parameters of the events are constrained very tightly. We assume that the events are contained within the 2-sigma region of the parameter uncertainties. This means that 90% of the sources will be resolved by the 3G network to within $\sim 60\,\mathrm{arcmin}^2$, which is better than the angular resolution of LISA for these sources; at an SNR of $\rho = 10$, $\Delta\Omega_{\rm LISA} \sim [1.22\lambda/(D\rho)]^2 \sim 1800\,\mathrm{arcmin}^2$, where $D = 2\,\mathrm{AU}$ is LISA's baseline and $\lambda \sim 3\times 10^7$ m is the wavelength of GWs corresponding to a frequency of 10 mHz. Moreover, the angles describing the orientation of the detector are also precisely determined by the 3G network—errors in the inclination and polarization angles are measured to within 1.0° and 2.3°, respectively, for 90% of the events. Consequently, it is not necessary to include the angular parameters in the search nor the trigger time, which will be known to better than 15 μs for 90% of the events. For the Gaussian distribution of spins considered in this paper it is not necessary to include spins either, leaving just the two masses over which a search should be carried out.

The number of templates required for an archival search for stellar-mass BBHs in the *LISA band* for the ~200 events in our population is shown in the right most panel of Fig. 3. These numbers are computed using a minimal match of 0.95 (or allowing for a loss of less than 5% in the SNR) [45] and assuming that the true event lies in the 2-sigma region of the uncertainty in the masses determined by the 3G network. A vast majority (90%) of the events require $< 15,000$ templates. One can employ singular value decomposition to find the number of *independent* templates [46], which is typically 1 to 2 orders-of-magnitude smaller than the number of templates found at this minimal match, or ~150-1500 for most of the events. Thus, an SNR-4 event in LISA will have a p-value 10^{-2} or smaller.

Multiparameter tests from the higher PN side: One may consider an interesting variant of the multiparameter tests of GR where more than one PN parameter is treated as independent, starting from the highest order that is currently known, which is 3.5PN. This may be thought of as tests of alternatives to GR where up to a particular PN order, the predictions of both GR and its alternative match but beyond that they deviate. This may be naturally motivated from an effective field theoretic perspective where the deviations may appear when the binary dynamics proceeds beyond a certain scale of velocity or field strength [18, 19].

Figure 4 shows the results for multiparameter tests starting from 3.5PN through 0PN order. For instance a three parameter test, would correspond to the case where only the last three PN deformation coefficients (3.5PN, 3PN *log* and 3PN) are simultaneously estimated and so on. The most significant result here is for the 7-parameter test for which it is found that the last seven PN parameters can be bounded with $\lesssim 10\%$ using the population we simulated earlier. This means, if we assume that the leading Newtonian coefficient is not modified, the simultaneous constraints on the remaining seven are of the order of a few percent. Since a modification to the Newtonian phasing would be at odds with the extremely stringent bounds on them from binary pulsar observations [47], this is the most general test we wish to carry out from this perspective. Fig. 4 also shows the best bounds that CE and LISA alone could yield from their golden binaries. This demonstrates that multibanding of GW signals is the only way to put meaningful constraints on multiple PN parameters with higher accuracy.

Inversion Accuracy of the Fisher matrix: Due to large degeneracies among the PN deformation parameters, there are high chances that the Fisher matrices (obtained from different events in the simulated population) corresponding to different multiparameter tests will be rendered ill-conditioned. Such Fisher matrices when inverted might lead to unreliable bounds. We therefore impose an inversion accuracy criterion on the Fisher matrices for all multiparameter tests. This criterion is defined to be $|\Gamma\cdot\Sigma - \mathrm{I}| \leq O(10^{-3})$, where Γ, Σ, and I are the multiband multiparameter Fisher matrix, the corresponding variance-covariance matrix and the identity matrix, respectively. Any Fisher matrix obtained from an event corresponding a particular multi-parameter test, that does not meet this criterion is dropped from our analysis and is not used to obtain the combined multiband multiparameter bounds.

[1] B. P. Abbott et al. (Virgo, LIGO Scientific), Phys. Rev. Lett. **116**, 061102 (2016), 1602.03837.

[2] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. **X9**, 031040 (2019), 1811.12907.

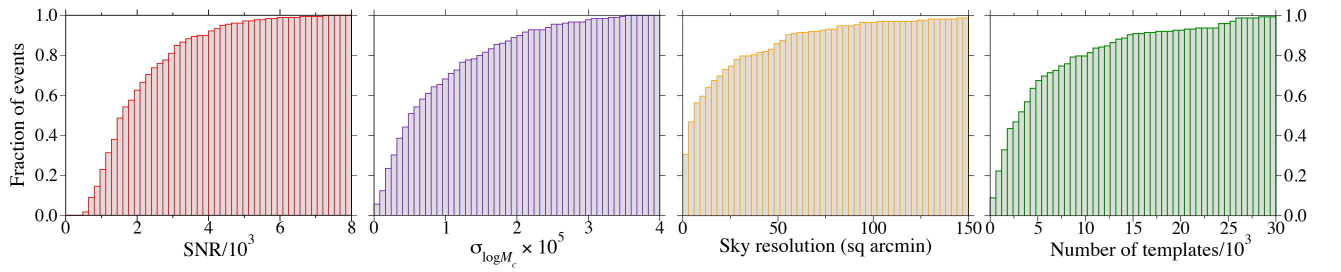

FIG. 3. The figure shows histograms of several parameters for the best 95% of the events that are visible both in 3G ground-based detectors and LISA (with an SNR ≥ 4): (i) the SNR, (ii) the error in the chirpmass, (iii) the error in the sky position, all determined by the 3G network, and (iv) the number of templates required to search for a stellar-mass BBH in the LISA data.

[3] B. P. Abbott et al. (Virgo, LIGO Scientific), Phys. Rev. Lett. **116**, 221101 (2016), 1602.03841.
[4] B. P. Abbott et al. (LIGO Scientific, Virgo), Phys. Rev. **D100**, 104036 (2019), 1903.04467.
[5] K. G. Arun, B. R. Iyer, M. S. S. Qusailah, and B. S. Sathyaprakash, Phys. Rev. D **74**, 024006 (2006), gr-qc/0604067.
[6] C. K. Mishra, K. G. Arun, B. R. Iyer, and B. S. Sathyaprakash, Phys. Rev. **D 82**, 064010 (2010), 1005.0304.
[7] N. Yunes and F. Pretorius, Phys. Rev. D **80**, 122003 (2009), 0909.3328.
[8] T. G. F. Li et al., Phys.Rev. D **85**, 082003 (2012), 1111.5274.
[9] L. Blanchet, Living Rev. Rel. **9**, 4 (2006), arXiv:1310.1528.

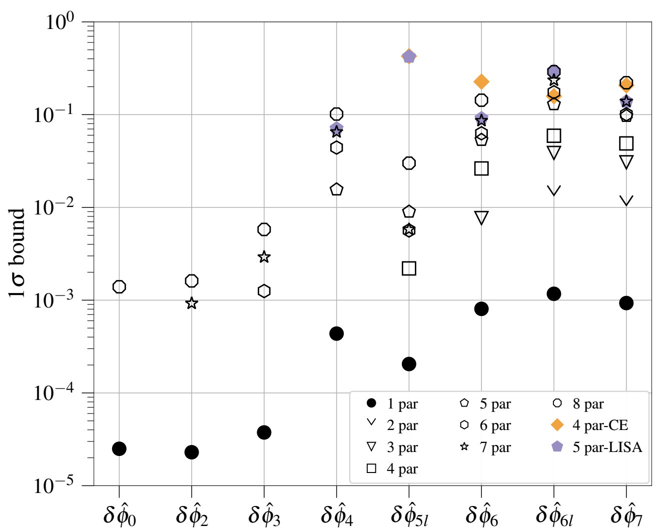

FIG. 4. Multiparameter tests using Multiband observations with LISA and CE. Shown are combined 1σ bounds on various PN coefficients starting from 3.5PN through 0PN in the inspiral phase of the signal while measuring many of them together at a time. Different types of markers symbolize how many PN coefficients were constrained simultaneously. For example, '●' represents 'one PN coefficient at a time', '∨' represents 'two PN coefficients at a time', and so on. The filled diamonds and pentagons are bounds obtained with CE and LISA, respectively, on the last four and five PN coefficients from their respective golden binaries. The total masses of the CE and LISA golden binaries are $200\,M_\odot$ and $6.6 \times 10^5\,M_\odot$, both binaries are 1 Gpc away and have component spins $\chi_1 = 0.6, \chi_2 = 0.5$. The figure represents results for the BBH population having Gaussian spin distribution (we get similar estimates for a uniform spin distribution).

[10] B. Sathyaprakash and S. Dhurandhar, Phys. Rev. D **44**, 3819 (1991).
[11] L. Blanchet, T. Damour, B. R. Iyer, C. M. Will, and A. G. Wiseman, Phys. Rev. Lett. **74**, 3515 (1995), gr-qc/9501027.
[12] L. Blanchet, T. Damour, G. Esposito-Farèse, and B. R. Iyer, Phys. Rev. Lett. **93**, 091101 (2004), gr-qc/0406012.
[13] K. G. Arun, A. Buonanno, G. Faye, and E. Ochsner, Phys. Rev. D **79**, 104023 (2009), 0810.5336.
[14] C. K. Mishra, A. Kela, K. G. Arun, and G. Faye, Phys. Rev. **D93**, 084054 (2016), 1601.05588.
[15] A. Buonanno, B. Iyer, E. Ochsner, Y. Pan, and B. Sathyaprakash, Phys. Rev. D **80**, 084043 (2009), 0907.0700.
[16] N. Yunes, K. Yagi, and F. Pretorius, Phys. Rev. **D94**, 084002 (2016), 1603.08955.
[17] K. G. Arun, B. R. Iyer, M. S. S. Qusailah, and B. S. Sathyaprakash, Class. Quantum Grav. **23**, L37 (2006), gr-qc/0604018.
[18] S. Endlich, V. Gorbenko, J. Huang, and L. Senatore, JHEP **09**, 122 (2017), 1704.01590.
[19] N. Sennett, R. Brito, A. Buonanno, V. Gorbenko, and L. Senatore (2019), 1912.09917.
[20] D. Reitze et al. (2019), 1907.04833.
[21] M. Punturo et al., Class. Quant. Grav. **27**, 194002 (2010).
[22] P. Amaro-Seoane, H. Audley, S. Babak, J. Baker, E. Barausse, P. Bender, E. Berti, P. Binetruy, M. Born, D. Bortoluzzi, et al., arXiv e-prints arXiv:1702.00786 (2017), 1702.00786.
[23] S. Datta et al. (in prep.).
[24] R. Nair, S. Jhingan, and T. Tanaka, PTEP **2016**, 053E01 (2016), 1504.04108.
[25] S. Vitale, Phys. Rev. Lett. **117**, 051102 (2016), 1605.01037.
[26] A. Sesana, Phys. Rev. Lett. **116**, 231102 (2016), 1602.06951.
[27] E. Barausse, N. Yunes, and K. Chamberlain, Phys. Rev. Lett. **116**, 241104 (2016), 1603.04075.
[28] A. Toubiana, S. Marsat, S. Babak, E. Barausse, and J. Baker (2020), 2004.03626.
[29] C. Liu, L. Shao, J. Zhao, and Y. Gao (2020), 2004.12096.
[30] Z. Carson and K. Yagi (2019), 1905.13155.
[31] G. Gnocchi, A. Maselli, T. Abdelsalhin, N. Giacobbo, and M. Mapelli, Phys. Rev. **D100**, 064024 (2019), 1905.13460.
[32] B. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al., Physical Review Letters **116** (2016), ISSN 1079-7114, URL http://dx.doi.org/10.1103/PhysRevLett.116.061102.
[33] V. Baibhav, E. Berti, D. Gerosa, M. Mapelli, N. Giacobbo, Y. Bouffanais, and U. N. Di Carlo, Phys. Rev. **D100**, 064060 (2019), 1906.04197.
[34] B. P. Abbott et al. (LIGO Scientific, Virgo), Astrophys. J. **882**, L24 (2019), 1811.12940.

[35] C. Rao, Bullet. Calcutta Math. Soc **37**, 81 (1945).
[36] H. Cramer, *Mathematical methods in statistics* (Pergamon Press, Princeton University Press, NJ, U.S.A., 1946).
[37] C. Cutler and E. Flanagan, Phys. Rev. D **49**, 2658 (1994).
[38] E. Berti, A. Buonanno, and C. M. Will, Phys. Rev. D **71**, 084025 (2005), gr-qc/0411129.
[39] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, Physical Review D **93** (2016), ISSN 2470-0029, URL http://dx.doi.org/10.1103/PhysRevD.93.044006.
[40] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Physical Review D **93** (2016), ISSN 2470-0029, URL http://dx.doi.org/10.1103/PhysRevD.93.044007.
[41] T. Robson, N. J. Cornish, and C. Liu, Class. Quant. Grav. **36**, 105011 (2019), 1803.01944.
[42] E. Poisson and C. Will, Phys. Rev. D **52**, 848 (1995).
[43] C. J. Moore, D. Gerosa, and A. Klein, Mon. Not. Roy. Astron. Soc. **488**, L94 (2019), 1905.11998.
[44] K. W. K. Wong, E. D. Kovetz, C. Cutler, and E. Berti, Phys. Rev. Lett. **121**, 251102 (2018), 1808.08247.
[45] B. J. Owen and B. Sathyaprakash, Phys. Rev. D **60**, 022002 (1999), gr-qc/9808076.
[46] K. Cannon et al., Astrophys. J. **748**, 136 (2012), 1107.2665.
[47] N. Yunes and S. A. Hughes, Phys. Rev. D **82**, 082002 (2010), 1007.1995.