

Confronting general relativity with principal component analysis: Simulations and results from GWTC-3 events

Parthapratim Mahapatra^{1,2,*}, Sayantani Datta^{2,3}, Ish Gupta⁴, Poulami Dutta Roy^{2,5}, Muhammed Saleem⁶,
Purnima Narayan⁷, Soumen Roy^{8,9}, Jan Steinhoff¹⁰, Deirdre Shoemaker⁶, Alan J. Weinstein¹¹,
Anuradha Gupta⁷, B. S. Sathyaprakash^{4,12,1} and K. G. Arun^{2,4}

¹*School of Physics and Astronomy, Cardiff University, Cardiff, CF24 3AA, United Kingdom*

²*Chennai Mathematical Institute, Siruseri, 603103, India*

³*Department of Physics, University of Virginia,*

P.O. Box 400714, Charlottesville, Virginia 22904-4714, USA

⁴*Institute for Gravitation and the Cosmos, Department of Physics,
Penn State University, University Park, Pennsylvania 16802, USA*

⁵*Department of Physics, University of Florida, PO Box 118440, Gainesville, Florida 32611-8440, USA*

⁶*Center for Gravitational Physics, The University of Texas at Austin, Austin, Texas 78712, USA*

⁷*Department of Physics and Astronomy, The University of Mississippi, University, Mississippi 38677, USA*

⁸*Centre for Cosmology, Particle Physics and Phenomenology—CP3,*

Université Catholique de Louvain, Louvain-La-Neuve, B-1348, Belgium

⁹*Royal Observatory of Belgium, Avenue Circulaire, 3, 1180 Uccle, Belgium*

¹⁰*Max-Planck-Institute for Gravitational Physics (Albert-Einstein-Institute),*

Am Mühlenberg 1, 14476 Potsdam-Golm, Germany

¹¹*LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, USA*

¹²*Department of Astronomy and Astrophysics, Penn State University,
University Park, Pennsylvania 16802, USA*



(Received 9 August 2025; accepted 8 October 2025; published 6 November 2025)

We present a comprehensive assessment of multiparameter tests of general relativity (GR) in the inspiral regime of compact binary coalescences using principal component analysis (PCA). Our analysis is based on an extensive set of simulated gravitational-wave (GW) signals, including both general relativistic and non-GR sources, injected into zero-noise data colored by the noise power spectral densities of the LIGO and Virgo GW detectors at their designed sensitivities. We evaluate the performance of PCA-based methods in the context of two established frameworks: TIGER and FTI. For GR-consistent signals, we find that PCA enables stringent constraints on potential deviations from GR, even in the presence of multiple free parameters. Applying the method to simulated signals that explicitly violate GR, we demonstrate that PCA is effective at identifying such deviations. We further test the method using numerical relativity waveforms of eccentric binary black hole systems and show that missing physical effects—such as orbital eccentricity—can lead to apparent violations of GR if not properly included in the waveform models used for analysis. Finally, we apply our PCA-based test to selected real gravitational-wave events from GWTC-3, including GW190814 and GW190412. We present joint constraints from selected binary black hole events in GWTC-3, finding that the 90% credible bound on the most informative PCA parameter is $0.03^{+0.08}_{-0.08}$ in the TIGER framework and $-0.01^{+0.05}_{-0.04}$ in the FTI framework, both of which are consistent with GR. These results highlight the sensitivity and robustness of the PCA-based approach and demonstrate its readiness for application to future observational data from the fourth observing runs of LIGO, Virgo, and KAGRA.

DOI: [10.1103/c1sj-jc4v](https://doi.org/10.1103/c1sj-jc4v)

*Contact author: MahapatraP@cardiff.ac.uk

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

I. INTRODUCTION

Parametric models play a central role in searches for new physics beyond the established paradigms, such as potential deviations from general relativity (GR) [1–6]. Since these models aim to capture hitherto unknown physics, they often involve a high-dimensional parameter space. When

tested against data, this high dimensionality can hinder inference by diluting constraints or inflating uncertainties. Principal component analysis (PCA), a linear algebra technique, offers an effective solution to this challenge. PCA reduces the effective dimensionality of the parameter space by identifying the best-measured linear combinations of the original parameter set without compromising on the information content in the data. It has been used recently in the context of gravitational-wave-based (GW-based) tests of GR [7–13].

Our focus here is on parametrized tests of compact binary inspirals [14–23]. The adiabatic inspiral phase of a compact binary is well described by post-Newtonian (PN) theory (see Ref. [24] for a review), in which the amplitude and phase of the gravitational-wave signal are expanded in powers of the characteristic orbital velocity v of the binary [25]. Since the phase evolution carries more critical information for parameter inference than the amplitude, a widely used approach for parametric models of the GW phase in a non-GR theory is to introduce additive deviations to the PN phasing coefficients at various orders—eight in total till 3.5PN, corresponding to $\mathcal{O}(v/c)^7$ [19,21,23]. Recently, Ref. [26] reported the 4.5PN phasing expression for quasicircular, nonspinning compact binaries; however, in this work, we restrict ourselves to the 3.5PN order since the newly computed terms are yet to be integrated into the standard tests of GR.

The PN phasing coefficients elegantly capture the richness of nonlinear effects in the compact binary dynamics (such as “tail” effects [25,27,28], tails of tails [29], tail square [30], memory effects [31–34], spin-orbit effects [35,36], spin-spin effects [35,37], etc.) at different PN orders in GR. Since these effects in a non-GR theory can differ from GR (see, for instance, Ref. [38] for scalar-tensor gravity theories), the additive parameters in the phasing are sensitive to such deviations. The parametrization is termed *null* because all deviations vanish identically in GR; thus, any statistically significant departure from the null hypothesis could hint at physics beyond GR. This formulation of the inspiral phase enables several distinct strategies for testing GR, which we now discuss.

A. Single versus multiparameter tests

There are different ways in which these tests can be performed.¹ Along with the usual GR parameters, one can choose to (a) estimate only one parameter at a time

¹Currently, the inspiral tests performed in the LVK papers use ten PN deformation parameters. These include eight fractional PN deformation parameters corresponding to the eight non-vanishing PN coefficients in GR up to 3.5PN order, along with absolute PN deformation parameters at the -1 PN dipole term and the 0.5PN term. We do not include the absolute deformation terms in this analysis and consider deformations only to those which are nonzero in GR.

(keeping all other testing GR parameters fixed at zero), (b) estimate all the parameters simultaneously, or (c) estimate a subset of them together while keeping the others fixed at zero. The single-parameter tests in (a) are the simplest among the tests that can be performed and are the ones implemented by the LIGO-Virgo-KAGRA (LVK) Collaboration [1,2,4,5], but they are subject to a fundamental limitation related to the theoretical framework of gravity. Should the true theory of gravity differ from GR, several phasing coefficients may deviate from their GR values. Therefore, a test of GR where all eight (or a sufficiently large number of) PN deformation coefficients are varied simultaneously would be ideal. Such a test involves estimating several additional parameters besides the standard GR ones from the data. This is a daunting task due to the strong degeneracies between the parameters that lead to uninformative posteriors. For example, Ref. [1] performed such a multiparameter test on GW150914 [39] and found uninformative posteriors on the PN deformation parameters due to the aforementioned correlations among parameters (see Fig. 7 of Ref. [1]).

The curse of dimensionality could be mitigated by multiband observations, where data from both space-based detectors, such as the Laser Interferometer Space Antenna, and ground-based observatories operating at that time are jointly analyzed for the same source [40,41]. Unfortunately, current GW detectors lack the sensitivity required to break the correlations between various parameters. The question then becomes whether there exist interesting parameters in terms of which a multiparameter test can be cast without having to deal with the parameter degeneracies. In this context, Ref. [42] proposed that a waveform model parametrizing the GW amplitude and phase in terms of the radiative multipole moments of the binary system [43–46] could serve as an effective approach to enable high-precision, multiparameter tests of GR.

In the current implementation of parametrized PN tests within the LVK tests of GR, PCA can be of help since it can identify the best-measured linear combinations of the original eight PN parameters, bringing down the dimensionality of the non-GR parameter space. Despite having the advantage of dealing with lower-dimensional parameter space, the posteriors of the PCA parameters carry the information of the original multiparameter test. In other words, instead of choosing to test one of the eight PN parameters at a time, PCA allows one to optimize the multiparameter test of GR by identifying the new set of parameters that account for most of the information and have the least correlation between them.

B. Review of PCA-based tests in the literature

Reference [7] proposed the idea of PCA to address the problem of degeneracies caused when one looks for simultaneous departures from GR predictions in several of the PN coefficients [16]. Using 3.5PN accurate

nonspinning phasing in the frequency domain [25,47–49], the authors argued that the linear combination of the original PN coefficients can be measured with significantly better precision [$\mathcal{O}(10^{-6})$ – $\mathcal{O}(10^{-4})$] than the original PN coefficients using second-generation or third-generation ground-based detector configurations. Recently, Ref. [9] carried out such a test in the case of GW170817 [50], for which the inspiral-based PN waveform model `TaylorF2` [51] was accurate enough, and derived the bounds on the PCA parameters.

Reference [10] revisited the problem in terms of PN fractional deformation parameters (as opposed to directly using the PN phasing coefficients as observables [7,9]) that are regularly employed for testing GR using the Test Infrastructure for General Relativity (TIGER) [21,22] and Flexible Theory-Independent (FTI) [23] frameworks. As mentioned earlier, these are fractional additive deformations introduced at every PN order which, by definition, are zero in GR. First, they performed a multiparameter test of GR varying six out of the eight PN deformation parameters along with all the standard GR parameters in a binary black hole (BBH) signal. Then, they obtained the covariance matrix corresponding to the six-dimensional posteriors of the deformation parameters by marginalizing over the GR parameters, in the TIGER framework. The covariance matrix was then diagonalized to obtain the new set of parameters that are linear combinations of the original deformation parameters. These are referred to as PCA parameters. The precision with which the PCA parameters are inferred was studied for a variety of simulated signals. As expected, the PCA-based method yields significantly tighter constraints on deviations from GR than the original multiparameter approach.

When fiducial non-GR signals are injected, the PCA-based method is capable of excluding GR with high statistical confidence, provided the signal-to-noise ratio (SNR) is sufficiently large: For a non-GR signal with an SNR of 54, GR was ruled out at a confidence level of $\sim 14\sigma$, while an SNR of 40 yielded a 11σ exclusion. In contrast, a signal with an SNR of 27 led to a GR exclusion at only $\sim 4\sigma$. This demonstrates the method’s sensitivity to both the SNR and the injected deviation.

Additionally, Ref. [10] also carried out this test on a selected subset of GW events in the first GW transient catalog [52] and obtained the corresponding bounds on the PCA parameters.

C. This work

In this paper, we extend the work of Ref. [10] by performing a comprehensive study that includes both the TIGER and FTI frameworks, a wider set of injections covering precession and nonquadrupole modes, and a broader range of GR and non-GR signals. This is an important step before deploying the PCA-based test of GR (or “PCA-TGR”) on the real signals as they will exhibit

several of these physical effects. Precisely quantifying the test’s response to a diverse set of injections is crucial for identifying potential systematic biases and for accurately interpreting results from real events—the primary objective of the present study.

More specifically, in this paper, we address the following questions:

- (1) How does PCA-TGR respond to different types of GR injections, which cover different physical effects present in the waveform, and how confidently does it recover GR?
- (2) How well can PCA-TGR detect a GR violation, and how does this ability depend on SNR, with analyses carried out independently using both the TIGER and FTI frameworks and considering a broader range of deviation scenarios, beyond the case where all PN coefficients are deformed by the same magnitude?
- (3) How does PCA-TGR respond to unmodeled physics in the waveform, such as eccentricity? Which of the PCA parameters capture this “false” GR violation?
- (4) What do the PCA-TGR results tell us about possible GR violation when applied to selected GWTC-3 [4,53] events?

This paper is organized as follows. Section II details the analysis framework. Sections III and IV report the results from GR and non-GR injections, respectively. Section V discusses how PCA-TGR responds to eccentric numerical relativity (NR) injections, and Sec. VI discusses the results from selected events from the third observing run of LIGO. Section VII summarizes the conclusions and caveats of the method.

II. ANALYSIS FRAMEWORK

In GR, the early inspiral phase of a gravitational-wave signal from a compact binary coalescence is well approximated by the PN formalism [24,25]. In this formalism, the amplitude and phase of gravitational waves are expressed as a series in powers of the characteristic orbital velocity v of the system. The frequency domain GW phase of the inspiral part from a quasicircular compact binary, computed using the stationary phase approximation [54,55], for the leading quadrupolar harmonic [56–58], takes the following form in GR

$$\Phi_{\text{insp}}(f) = 2\pi f t_c - \phi_c - \frac{\pi}{4} + \frac{3}{128\eta v^5} \times \left[\sum_{k=0}^{k=7} (\phi_k + \phi_{kl} \ln v) v^k + \mathcal{O}(v^8) \right], \quad (1)$$

where t_c and ϕ_c are the time and phase of coalescence, and η is the symmetric mass ratio defined as the ratio of the reduced mass to the total mass of the binary. The characteristic orbital velocity parameter, in geometrical units, is given by $v = (\pi M f)^{1/3}$, where f is the GW frequency and

M is the (redshifted) total mass of the binary.² Terms with a k th power of v are identified as the $k/2$ PN order corrections to the GW phase. The PN coefficients, denoted by ϕ_k (nonlogarithmic) and $\phi_{k\ell}$ (logarithmic), are unique functions of the compact binary's component masses and spins, capturing several physical effects and nonlinearities in GR. In GR, up to 3.5PN order, the logarithmic coefficients ϕ_{kl} are nonzero only for $k = 5, 6$.

A. Multiparameter test of general relativity

In alternative theories of gravity, one or more of the PN coefficients may show deviation from GR [6,38,59,60]. They may have different dependencies on the component masses and spins than GR, and/or they may depend on additional parameters that characterize the alternative theory (see, for example, Ref. [38] for scalar-tensor gravity theories).

The parametrized inspiral tests search for possible deviations from the unique structure of the PN coefficients predicted by GR. This is done by introducing phenomenological, dimensionless fractional deviation parameters into the PN coefficients [1,19–21]:

$$\phi_b \rightarrow \phi_b^{\text{GR}}(1 + \delta\hat{\phi}_b), \quad (2)$$

where the subscript “ b ” collectively represents both logarithmic and nonlogarithmic coefficients of Eq. (1). This is a theory-agnostic approach wherein the parametrized fractional deviations $\delta\hat{\phi}_b$ are identically zero in GR, whereas in alternative theories of gravity, one or more of these parameters could differ from zero. Therefore, a direct measurement of $\delta\hat{\phi}_b$ using GW data allows us to constrain any deviation from GR. Such tests are referred to as “null” tests of GR, with the null hypothesis being GR.

Currently, among LVK analyses, there are two different frameworks that can perform this type of GR null test. The first one is TIGER [21,61], and the second is FTI [23]. The two frameworks differ in the choice of baseline GR waveform models and in how the inspiral test is implemented. TIGER [62] uses the IMRPhenomX waveform family [63–65], which accounts for precession, as the baseline GR waveform model. The FTI [66] relies on the SEOBNRv4_ROM waveform family [67–69] and is currently limited to aligned-spin systems. In the TIGER framework, the fractional PN deviation parameters are added directly to the PN coefficients in the frequency-domain inspiral phase of the IMRPhenomX waveform family, which ends at the minimum energy circular orbit (MECO) frequency. In the FTI framework, the fractional PN deviation parameters are added to the full

frequency-domain inspiral-merger-ringdown phase of the SEOBNRv4_ROM waveform family and are smoothly tapered off at the frequency corresponding to the peak of the (2, 2) mode. Note that parametric fractional deviations are only applied to the nonspinning portion of the inspiral phase in both frameworks.³

In the current implementation of this null test by the LVK Collaboration, only one fractional deviation parameter, $\delta\hat{\phi}_b$, is varied at a time while keeping the others fixed at their GR values to obtain the bounds on all the fractional deviation parameters. We refer to this approach as *single-parameter tests*, which leads to a set of null tests of GR, each corresponding to a PN coefficient, although the tests are not necessarily all independent. While single-parameter tests can *detect* deviations from GR [70,71], they *cannot be uniquely attributed* to a real deviation at the corresponding PN order due to degeneracies in the waveform. In other words, it is possible that a deviation at particular PN order in a non-GR theory can be suboptimally captured by a test which looks for departures at a different PN order. Additionally, single-parameter tests ignore potential modifications of multiple PN coefficients, which can lead to an underestimation of errors and biases [72]. As a consequence, despite their ability to detect GR deviations, the single-parameter tests provide a somewhat limited insight into the nature of alternative theories of gravity.

In light of the limitations of single-parameter tests, we advocate for a more robust multiparameter approach to test GR, wherein multiple fractional deviation parameters are varied simultaneously [7,16,61,70]. More specifically, we consider simultaneous variation of all the fractional deformation parameters: $\{\vec{\delta\hat{\phi}}_k, \vec{\delta\hat{\phi}}_{k\ell}\}$. Taking into account the inspiral phase in GR up to 3.5PN order with nonzero PN coefficients, we have a set of eight fractional deformation parameters⁴:

$$\vec{\delta\hat{\phi}}_k = \{\delta\hat{\phi}_0, \delta\hat{\phi}_2, \delta\hat{\phi}_3, \delta\hat{\phi}_4, \delta\hat{\phi}_6, \delta\hat{\phi}_7\}, \quad (3)$$

$$\vec{\delta\hat{\phi}}_{k\ell} = \{\delta\hat{\phi}_{5\ell}, \delta\hat{\phi}_{6\ell}\}. \quad (4)$$

The major obstacle of performing this class of tests is the uninformative posteriors due to high correlations between the parameters. We employ principal component analysis to remedy this shortcoming and identify the best measured

²We note that in the LVK papers, $(1+z)M$ is used to denote the redshifted total mass, where z is the cosmological redshift and M is the source-frame total mass. Here, however, we denote M as the redshifted (detector-frame) total mass for brevity.

³This choice is motivated by the convenience of mapping the results to nonlinear effects in general relativity, such as tails, tails of tails, and so on. Since we are performing null tests, whether the deformation is applied to the total phasing or only to the nonspinning part will have a negligible impact on the results. However, this distinction should be kept in mind when interpreting the outcomes of the tests.

⁴We do not consider fractional deviations to the nonlogarithmic term at 2.5PN order, $\delta\hat{\phi}_5$, as it is independent of frequency and can therefore be absorbed by redefining the phase at coalescence, ϕ_c .

linear combinations of the deformation parameters [7,9,41,73]. The details of the method are described below.

B. Bayesian formulation

The baseline GR waveform model for a BBH is a function of several binary parameters (denoted as $\vec{\theta}_{\text{GR}}$), including component masses, spins, luminosity distance, and angular parameters describing the orientation of the binary orbit. For nonprecessing and precessing BBHs on quasicircular orbits, $\vec{\theta}_{\text{GR}}$ contains 11 and 15 parameters,⁵ respectively. In addition to $\vec{\theta}_{\text{GR}}$, the parametrized waveform models under consideration are also functions of eight fractional PN deviation parameters: $\vec{\theta}_{\text{D}} = \{\vec{\delta\phi}_k, \vec{\delta\phi}_{k\ell}\}$. In general, the parametrized waveform models depend on m GR parameters and n fractional deviation parameters.

In a multiparameter test of GR, this implies running a Bayesian parameter estimation algorithm to infer m GR parameters $\vec{\theta}_{\text{GR}}$ and n non-GR parameters $\vec{\theta}_{\text{D}}$. Given a hypothesis $\mathcal{H}$ that the data d_j of event j are a sum of Gaussian noise and a GW signal described by a waveform model that is a function of $\vec{\theta}_{\text{GR}}$ and $\vec{\theta}_{\text{D}}$ parameters, the posterior probability distribution of parameters $\vec{\theta}_{\text{GR}}$ and $\vec{\theta}_{\text{D}}$ can be written using Bayes' theorem as

$$p(\{\vec{\theta}_{\text{GR}}, \vec{\theta}_{\text{D}}\} | d_j, \mathcal{H}) = \frac{\mathcal{L}(d_j | \{\vec{\theta}_{\text{GR}}, \vec{\theta}_{\text{D}}\}, \mathcal{H}) \pi(\{\vec{\theta}_{\text{GR}}, \vec{\theta}_{\text{D}}\} | \mathcal{H})}{\mathcal{Z}(d_j | \mathcal{H})}. \quad (5)$$

Here, $\pi(\{\vec{\theta}_{\text{GR}}, \vec{\theta}_{\text{D}}\} | \mathcal{H})$ is the *prior probability* distribution of $\vec{\theta}_{\text{GR}}$ and $\vec{\theta}_{\text{D}}$, $\mathcal{L}(d_j | \{\vec{\theta}_{\text{GR}}, \vec{\theta}_{\text{D}}\}, \mathcal{H})$ is the *likelihood function*, and $\mathcal{Z}(d_j | \mathcal{H})$ is the *Bayesian evidence* in favor of the hypothesis $\mathcal{H}$. To generate discrete samples for the probability distribution function $p(\{\vec{\theta}_{\text{GR}}, \vec{\theta}_{\text{D}}\} | d_j, \mathcal{H})$, we used the Bayesian parameter inference package BILBY TGR [74] based on BILBY [75,76], with the DYNESTY [77] sampler (which uses the nested sampling algorithm [78]). We adopt the prior probability distribution $\pi(\{\vec{\theta}_{\text{GR}}, \vec{\theta}_{\text{D}}\} | \mathcal{H})$ following the standard setup used in the

⁵The observed GW signal emitted by quasicircular, precessing BBHs within GR depends on 15 parameters: detector-frame masses m_1 (primary black hole) and m_2 (secondary black hole); spin magnitudes χ_1 and χ_2 ; spin-orbit tilt angles θ_1 and θ_2 (angles between the spin vectors and the orbital angular momentum); spin azimuthal angle difference ϕ_{12} (the difference between the azimuthal angles of the two spin vectors); ϕ_{JL} (the angle between the orbital and total angular momentum); inclination angle θ_{JN} (between the line of sight and total angular momentum); luminosity distance D_L ; sky location (right ascension α and declination δ); polarization angle ψ ; time of coalescence, t_c ; and phase at coalescence, ϕ_c . For aligned-spin binaries, only two spin parameters (χ_1, χ_2) are relevant, reducing the total number of parameters to 11.

publicly available LVK data [4,53,79]. This means for the GR sector the detector-frame component masses and spin magnitudes of individual black holes have uniform priors, while spin orientations are distributed isotropically over a sphere. The luminosity distance D_L is assigned a prior so that sources are uniformly distributed in the comoving volume and source-frame time. The priors for all other angular parameters are chosen to be isotropic over a sphere. In general, the chosen priors of the standard GR parameters are uninformative and sufficiently broad to encompass the regions of the parameter space where the posterior distributions are supported while also balancing computational efficiency and remaining within the calibration range of the waveform models. For the non-GR parameters, we assume a uniform symmetric prior between -20 and 20 , centered at zero.

As our focus is on the fractional deformation parameters, we marginalize Eq. (5) over the m GR parameters to obtain n -dimensional joint posterior distribution of $\vec{\theta}_{\text{D}}$,

$$p(\vec{\theta}_{\text{D}} | d_j, \mathcal{H}) = \int p(\{\vec{\theta}_{\text{GR}}, \vec{\theta}_{\text{D}}\} | d_j, \mathcal{H}) d\vec{\theta}_{\text{GR}}. \quad (6)$$

Moreover, we only vary six deviation parameters from 1.5PN to 3.5PN, fixing the 0PN and 1PN deviation parameters to their GR value of zero, leading to $\vec{\theta}_{\text{D}} = \{\delta\hat{\phi}_3, \delta\hat{\phi}_4, \delta\hat{\phi}_{5\ell}, \delta\hat{\phi}_6, \delta\hat{\phi}_{6\ell}, \delta\hat{\phi}_7\}$. The reason for this is explained at the end of Sec. II C. We do not consider -1 PN and 0.5 PN deviations, which we set to zero.

C. Principal component analysis

As noted earlier, the fractional deformation parameters $\vec{\theta}_{\text{D}}$ are generally correlated. Consequently, a multiparameter test using the fractional deformation parameters described in the previous section can lead to degenerate posteriors, resulting in a nondiagonal covariance matrix in their joint distribution. To address this, we aim to construct a new set of orthogonal parameters by rotating the axes of the original fractional deformation parameters such that the resulting covariance matrix becomes diagonal. This is accomplished by using the principal component analysis, following the method outlined in Ref. [10], with the key steps summarized below.

- (i) We compute the variance-covariance matrix for fractional deviation parameters from the $\vec{\theta}_{\text{D}}$ posteriors obtained in Eq. (6),

$$C_{ab} = \langle (\delta\hat{\phi}_a - \langle \delta\hat{\phi}_a \rangle) (\delta\hat{\phi}_b - \langle \delta\hat{\phi}_b \rangle) \rangle, \quad (7)$$

where the angular brackets denote averaging over the marginalized posterior distribution.

- (ii) We diagonalize the covariance matrix C as

$$C = USU^T, \quad (8)$$

where S is the diagonal matrix containing the eigenvalues, and U is the unitary matrix whose columns are the corresponding eigenvectors representing the transformation of the basis vectors. Here, we use the `numpy.linalg.svd` function in Python [80] to diagonalize the covariance matrix—or more specifically, to obtain the U and S matrices.

- (iii) In this context, S represents a diagonal variance-covariance matrix corresponding to the new deformation parameters, which are obtained by applying the transformation matrix U (with elements α_{ib}) on the original deformation parameters,

$$\delta\hat{\phi}_{\text{PCA}}^{(i)} = \sum_b \alpha_{ib} \delta\hat{\phi}_b, \quad (9)$$

which we refer to as the PCA parameters.

- (iv) PCA has an inherent sign ambiguity because the direction of eigenvectors is not uniquely defined; flipping the sign of an eigenvector does not change the variance it explains. In other words, both $\vec{u}_i$ and $-\vec{u}_i$ are valid eigenvectors of the covariance matrix C , yielding the same diagonalized form S . This ambiguity carries over to the principal components $\delta\hat{\phi}_{\text{PCA}}^{(i)}$, potentially causing inconsistencies in interpretation across analyses of different GW events. To resolve this, we adopt a sign convention in which each eigenvector, and the corresponding principal component, is adjusted based on the sign of the eigenvector's largest (in magnitude) element—ensuring it is always positive. This choice does not affect the statistical properties of the PCA parameters but guarantees consistency in the orientation of the new basis. As a result, the PCA parameters remain well defined and on a comparable footing across different GW events in our analysis.
- (v) We perform numerical consistency checks to validate the diagonalization procedure. First, we assess the *orthonormality* of the transformation matrix U by evaluating the maximum absolute value of the off-diagonal elements of $UU^T - \mathbb{I}$:

$$\max_{a \neq b} |(UU^T - \mathbb{I})_{ab}|, \quad (10)$$

where $(\mathbb{I})_{ab}$ denotes the element in the a th row and b th column of the identity matrix. This should vanish for a perfectly orthonormal matrix. Second, we reconstruct the covariance matrix using $C = USU^T$ and compute the Frobenius norm of the difference with the original covariance matrix:

$$\|USU^T - C\|_F, \quad (11)$$

to quantify the total reconstruction error introduced by the numerical diagonalization.

In all cases applied here, we find that the maximum off-diagonal element of $|UU^T|$ is $\mathcal{O}(10^{-16})$, and the Frobenius norm of $USU^T - C$ is $\mathcal{O}(10^{-13})$, confirming the numerical stability and internal consistency of the PCA-based diagonalization.

- (vi) The primary advantage of the PCA formalism is dimensionality reduction. In our case, this is guided by examining the eigenvalues, i.e., the diagonal elements of the matrix S . Each eigenvalue corresponds to the variance (squared uncertainty) of the posterior along the direction defined by the associated eigenvector in U . A hierarchy naturally emerges: eigenvectors associated with smaller eigenvalues correspond to directions with tighter constraints and, hence, carry more information. Conversely, directions with large eigenvalues (i.e., large error bars) are poorly constrained and largely uninformative—often yielding posteriors that resemble the priors. Therefore, such directions can be safely excluded from further analysis, enabling a truncated, lower-dimensional representation of the parameter space depending on the acceptable tolerance in reproducing the data with the lower dimensional PCA parameters.
- (vii) The information content encoded in the posterior distribution can be assessed using the inverse of the covariance matrix, also known as the information matrix. The overall information can then be quantified by computing the determinant of the inverse of S . Previous studies [10,12] have shown that retaining only the two most tightly constrained principal components often captures more than 99% of the total information. In this work, rather than relying solely on eigenvalue hierarchy, we also use the Jensen-Shannon (JS) divergence [81] to determine the number of PCA parameters to retain. Specifically, we retain only those PCA components for which the JS divergence between the posterior and prior distributions exceeds 0.1 bits, ensuring that the selected components carry meaningful information gained from the data.

Equation (9) defines a new set of dimensionless deformation parameters, each of which is a linear combination of the original PN fractional deviation parameters. These transformed parameters effectively encapsulate the joint information content of a multiparameter test. The approach outlined above is data-driven, meaning that the new parameters are derived directly from the posterior distribution, making the resulting parametrization inherently unique to each event. Because the posterior is unique to each event, this motivates the use of a data-driven parametrization, as it allows the method to adapt to the information content of each event while avoiding assumptions about the underlying parameter values. An alternative strategy involves estimating the covariance matrix using the Fisher information

matrix formalism [82–85]. However, this requires prior knowledge of the true parameter values—information that is generally inaccessible, especially in scenarios where deviations from GR might be present. Furthermore, preliminary investigations have shown inconsistencies between Fisher matrix predictions and the outcomes of the data-driven approach, likely due to the approximations and limitations inherent in the Fisher formalism for the SNRs of typical LVK events, such as bimodality and other non-Gaussianities in the posteriors.

In the current formulation of the PCA-based multiparameter test, we exclude 0PN and 1PN fractional deformation parameters and instead vary coefficients from 1.5PN to 3.5PN. This is due to the limited sensitivity of current GW detectors, which cannot effectively disentangle these lower-order fractional deformations from intrinsic binary parameters like the chirp mass and symmetric mass ratio. Furthermore, the lower PN deviation parameters are already tightly constrained by binary pulsar observations [86,87]. For this reason, the current analysis is limited to a six-dimensional multiparameter test, with deformation parameters introduced at orders ranging from 1.5PN to 3.5PN. However, with future improvements in detector sensitivity or the observation of high-SNR events with longer inspirals, the PCA framework can be extended to perform a full eight-dimensional test, enabling a more complete exploration of PN deformations.

D. Combining information from multiple events

The PCA parameters, obtained by the procedure outlined above, will be different for different events. Therefore, joint constraints on the PCA parameters have to be clearly defined. There are two commonly used approaches for combining results from single-parameter tests of GR across multiple GW events. The first approach, which we adopt in this work, involves multiplying the individual likelihoods of the deformation parameters obtained for each event [20,88]. This method assumes that the value of each deformation parameter is identical across all events, an admittedly strong and potentially unrealistic assumption given the variation in source properties, but one that is valid if GR holds true. An alternative approach involves hierarchical inference [89,90], which models event-to-event variation through a population distribution, typically modeled as Gaussian.

In the context of PCA-based multiparameter tests, combining results across events presents additional challenges. Unlike single-parameter analyses, PCA parameters are event-specific due to differences in the transformation matrix that defines them. As a result, the same PCA index corresponds to different linear combinations of the original fractional PN deviation parameters across events, making the post-PCA combination technically inconsistent. We therefore combine events at the level of the original multiparameter posteriors—using the shared-parameter likelihood multiplication method adopted in this work—and

then perform PCA on the combined posterior. While a hierarchical approach could, in principle, be extended to this setting, modeling the joint distribution of six correlated parameters would require estimating 27 hyperparameters under the Gaussian model assumption, a task not yet explored in high-dimensional GW analyses. We leave such developments to future work.

To combine data from a set of, say, N events, we begin by computing the six-dimensional marginalized likelihood for the deformation parameters, $\vec{\theta}_D$, for each event, following Eqs. (5) and (6). The combined likelihood is then constructed as the product of the N multidimensional likelihoods:

$$\mathcal{L}^{\text{combined}}(\{d_j\}|\vec{\theta}_D, \mathcal{H}) = \prod_{j=1}^N \mathcal{L}(d_j|\vec{\theta}_D, \mathcal{H}), \quad (12)$$

where each $\mathcal{L}(d_j|\vec{\theta}_D, \mathcal{H})$ is computed by dividing the marginalized posterior distribution in Eq. (6) by the prior on $\vec{\theta}_D$, $\pi(\vec{\theta}_D|\mathcal{H})$. Given that uniform priors were used for all six parameters in the original analyses, the likelihoods are directly proportional to the posteriors, and because the posteriors of each event are provided as discrete samples, the likelihoods are constructed by sampling from the six-dimensional kernel density estimator fits using the `sklearn.mixture.GaussianMixture` function [91]. The combined posterior on $\vec{\theta}_D$ is then obtained by multiplying the combined likelihood, as given in Eq. (12), with the same uniform prior on $\vec{\theta}_D$ used in the original single-event analysis:

$$p(\vec{\theta}_D|\{d_j\}, \mathcal{H}) \propto \pi(\vec{\theta}_D|\mathcal{H}) \mathcal{L}^{\text{combined}}(\{d_j\}|\vec{\theta}_D, \mathcal{H}). \quad (13)$$

Furthermore, to generate discrete samples for the probability distribution function $p(\vec{\theta}_D|\{d_j\}, \mathcal{H})$, we again used the Bayesian parameter inference library `BILBY` [75] with the `dynesty` [77] sampler. Finally, the PCA formalism, described in Sec. II C, is applied to the discrete samples of the combined six-dimensional posterior probability distribution $p(\vec{\theta}_D|\{d_j\}, \mathcal{H})$ to obtain the joint constraints on the PCA parameters.

After describing the analysis framework of the PCA-TGR, we validate its efficacy through various types of injection studies, as discussed below.

III. RESPONSE OF PCA-TGR TO GR INJECTIONS

First, we briefly summarize the key characteristics common to all injections discussed in this section and the following two sections (Secs. IV and V). Unless otherwise specified, all reported injections assume the “zero-noise” approximation in a three-detector LIGO-Virgo network [92–94] using the predicted noise power spectral densities

at design sensitivity: aLIGOZeroDetHighPower for the LIGO detectors [95] and AdvVirgo for the Virgo detector [96]. This approximation is used to isolate potential systematic biases and to avoid statistical fluctuations arising from specific noise realizations. A lower cutoff frequency of 20 Hz is used in the computation of the likelihood, while the upper cutoff frequency varies across systems, primarily based on their masses. These values are chosen to ensure that no signal content is lost, at the same time avoiding unnecessarily high computational costs arising from arbitrarily large cutoffs. All other parameter estimation settings—such as signal duration, sampling frequency, and sampler configurations—generally follow the standard setup used in previous LVK publications [4,53,79]. It is worth noting that signal duration and sampling frequency are system dependent, chosen based on the physical properties of the injections. In contrast, the sampler settings remain largely consistent across all analyses.

To quantify if the posterior distributions of deformation parameters include or exclude the GR value, for each injection we compute the GR quantiles $Q_{\text{GR}} = P(\delta\hat{\phi}_{\text{PCA}}^{(i)} < 0)$ of the PCA posterior distributions. A value of Q_{GR} significantly different from 0.5 indicates that the GR hypothesis falls in the tail of the distribution. We also compute the GR z score (z_{GR}), defined as the number of standard deviations by which the GR value (zero) deviates from the statistical median of the posterior.

A. TIGER framework

1. Aligned-spin systems

Our first goal is to assess the efficacy of the PCA-TGR in response to GR-consistent signals across different regions of the BBH parameter space (i.e., varying total masses, mass ratios, and spin magnitudes, and network SNRs). Because of the computational cost of multiparameter analyses, we begin with a simplified setup using aligned-spin quadrupolar waveform models to reduce computational overhead. In later sections, however, we incorporate more complex models that include spin precession and higher-order modes.

Results for the GR-consistent, aligned-spin injections are shown in Table I and Fig. 1. As shown in Table I, we consider binaries with fixed intrinsic parameters evaluated at two network SNRs: ~ 20 (low) and ~ 40 (high), achieved by adjusting the luminosity distance. The extrinsic parameters—such as sky position, polarization angles, and phase at coalescence—are chosen randomly. To span a broad parameter space, we use total detector-frame masses of $25M_{\odot}$ and $60M_{\odot}$, with mass ratios $q = 1.5$ and 4 . For spin magnitudes, we choose $(\chi_1, \chi_2) = (0.2, 0.1)$ and $(0.8, 0.7)$ as representative low- and high-spin configurations, respectively. These combinations yield eight distinct binary configurations, each analyzed at both low and high network SNRs, as listed in Table I. We use the IMRPhenomXAS

TABLE I. Summary of results from the simulated aligned-spin BBH GR injections study within the TIGER framework, demonstrating the efficacy of the PCA-TGR method as discussed in Sec. III A 1. We simulate GW signals using the IMRPhenomXAS waveform model and perform parameter estimation with the parametrized IMRPhenomXAS model. The specifications of the binary systems and the network SNRs (rounded to the nearest integer) of the injected signals are listed. For each injection, the table shows the median value, as well as the upper and lower limits of the 90% credibility interval, Q_{GR} , and z_{GR} for the posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$.

No.	Properties of GR injections					Properties of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$		Properties of $\delta\hat{\phi}_{\text{PCA}}^{(2)}$	
	$M (M_{\odot})$	q	(χ_1, χ_2)	D_L (Mpc)	SNR	Median and 90% errors	Q_{GR} (z_{GR})	Median and 90% errors	Q_{GR} (z_{GR})
1.	25	1.5	(0.2, 0.1)	1200	20	$-0.02^{+0.17}_{-0.21}$	0.57 (0.16)	$-0.09^{+0.8}_{-1.0}$	0.57 (0.15)
				650	40	$-0.0^{+0.1}_{-0.1}$	0.52 (0.05)	$0.0^{+0.39}_{-0.43}$	0.5 (0.01)
2.	25	1.5	(0.8, 0.7)	1400	20	$-0.07^{+0.11}_{-0.26}$	0.76 (0.66)	$-0.18^{+0.81}_{-1.37}$	0.63 (0.26)
				700	40	$-0.04^{+0.06}_{-0.13}$	0.77 (0.73)	$-0.09^{+0.44}_{-0.64}$	0.63 (0.28)
3.	25	4	(0.2, 0.1)	900	20	$-0.01^{+0.21}_{-0.2}$	0.53 (0.06)	$0.12^{+1.3}_{-1.14}$	0.42 (0.16)
				500	40	$-0.0^{+0.12}_{-0.12}$	0.5 (0.0)	$0.04^{+1.06}_{-0.76}$	0.46 (0.07)
4.	25	4	(0.8, 0.7)	1000	20	$-0.04^{+0.13}_{-0.22}$	0.66 (0.34)	$-0.21^{+0.94}_{-1.42}$	0.62 (0.24)
				550	40	$-0.03^{+0.04}_{-0.1}$	0.75 (0.62)	$-0.02^{+0.51}_{-0.65}$	0.53 (0.06)
5.	60	1.5	(0.2, 0.1)	2400	20	$-0.04^{+0.22}_{-0.29}$	0.6 (0.22)	$-0.18^{+1.57}_{-2.25}$	0.56 (0.15)
				1200	40	$-0.0^{+0.16}_{-0.15}$	0.5 (0.01)	$-0.04^{+1.18}_{-1.04}$	0.52 (0.06)
6.	60	1.5	(0.8, 0.7)	2800	20	$0.0^{+0.3}_{-0.28}$	0.49 (0.02)	$0.59^{+1.98}_{-0.73}$	0.23 (0.71)
				1400	40	$0.02^{+0.23}_{-0.19}$	0.45 (0.13)	$0.21^{+0.87}_{-0.47}$	0.29 (0.52)
7.	60	4	(0.2, 0.1)	1800	20	$-0.02^{+0.36}_{-0.34}$	0.54 (0.09)	$0.32^{+2.38}_{-2.21}$	0.41 (0.22)
				950	40	$0.0^{+0.21}_{-0.17}$	0.5 (0.0)	$0.21^{+1.68}_{-1.34}$	0.41 (0.23)
8.	60	4	(0.8, 0.7)	2100	20	$0.01^{+0.29}_{-0.26}$	0.48 (0.03)	$-0.11^{+1.17}_{-1.43}$	0.56 (0.13)
				1100	40	$-0.01^{+0.16}_{-0.17}$	0.55 (0.11)	$0.03^{+0.83}_{-0.7}$	0.47 (0.06)

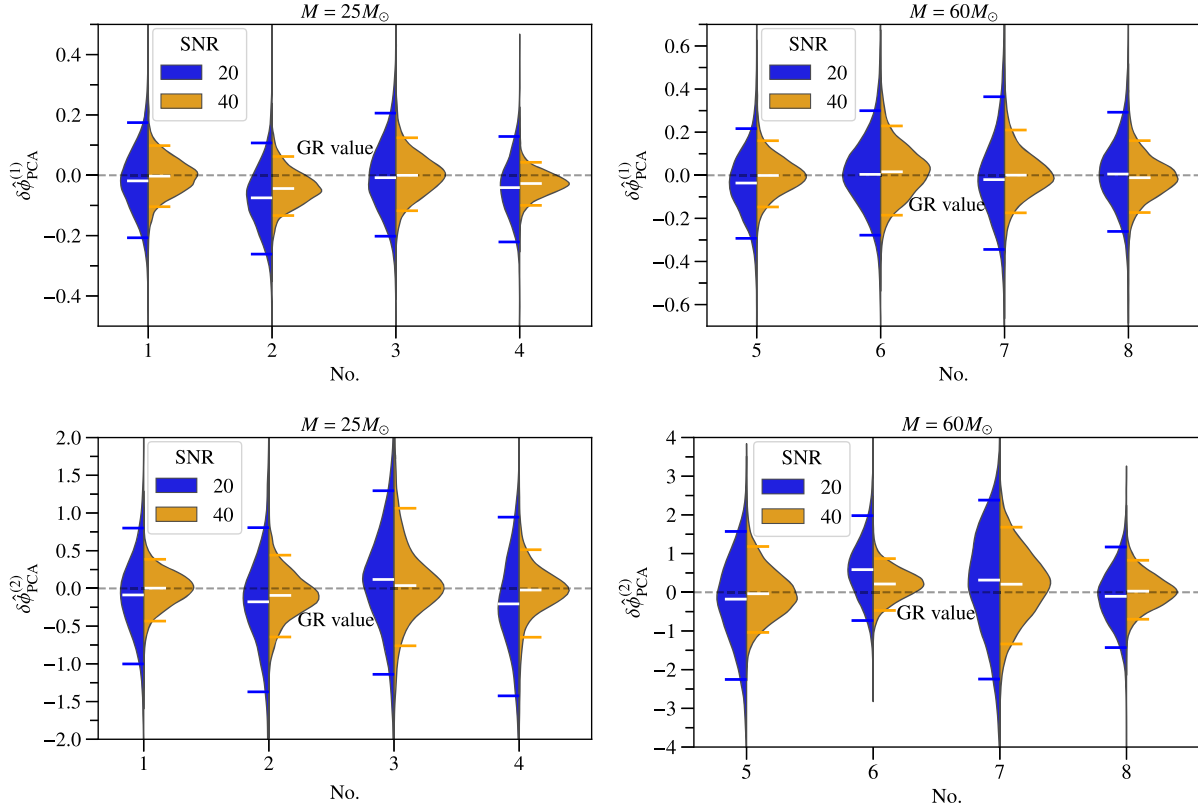


FIG. 1. Posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ (upper panel) and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ (lower panel) for each simulated injection—listed in Table I and discussed in Sec. III A 1—with the left and right panels displaying results from BBHs with detector-frame total masses of $25M_{\odot}$ and $60M_{\odot}$, respectively. In each violin plot, the colored horizontal bars and the horizontal white solid line denote the 90% credible intervals and the posterior median, respectively. We mark the GR value of zero with dashed gray lines.

waveform model [97] to generate injected GW signals consistent with GR and employ the parametrized IMRPhenomXAS model [62] for parameter estimation, producing discrete samples of the posterior probability distribution given in Eq. (5) within the TIGER framework. We then perform PCA on the marginalized six-dimensional posteriors of the fractional PN deviation parameters to construct the dominant linear combinations that are best measured, as described in Sec. II C.

We find that the JS divergence between the posterior and prior distributions exceeds 0.1 bits⁶ for the first two leading PCA parameters in all injections, while it is marginally greater than 0.1 bits for the third leading PCA parameter for some cases. As expected, due to the nature of PCA, the widths of the first two dominant PCA parameters, $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$, are significantly smaller than those of the original fractional PN deformation parameters.

In Fig. 1, we show the posterior distributions for the leading two PCA parameters presented as violin plots for each injection analysis. Each binary configuration is

labeled by a number on the x axis, corresponding to the list of binaries in Table I. In Table I, we also report the median and 90% credible intervals, along with Q_{GR} and z_{GR} for the posteriors of the two leading PCA parameters. First, we observe that the GR value is well recovered in all cases, as zero lies within the 90% credible interval. This confirms that PCA-TGR can successfully recover GR across different binary configurations. Second, as expected, the constraints on the PCA parameters improve with increasing SNR. We also find that the constraints are slightly tighter for $q = 4$ compared to $q = 1.5$, which can be attributed to the larger number of signal cycles within the detector’s sensitivity band.

2. Precessing systems

We now present results for a set of GR-consistent injection runs involving precessing BBHs. The configurations span a range of effective precession parameters χ_p [98,99] and network SNRs, allowing us to investigate the performance of the PCA-TGR in recovering the injected GR signal under varying levels of precessional dynamics and signal strength.

All injections correspond to binary systems with total mass $30M_{\odot}$ and mass ratio $q = 2$. To probe both

⁶When the JS divergence is 0 bit, the two distributions are identical, and when it is 1 bit, the two distributions are maximally different.

TABLE II. Summary of results from the simulated precessing-spin BBH GR injection study conducted within the TIGER framework, as described in Sec. III A 2. GW signals are simulated using the IMRPhenomXP waveform model, and parameter estimation is performed using the parameterized IMRPhenomXP model. The statistics of the two leading PCA parameters are reported, and their corresponding posterior distributions are shown in Fig. 2.

No.	Properties of GR injections							Properties of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$		Properties of $\delta\hat{\phi}_{\text{PCA}}^{(2)}$	
	M ($M_{\odot}$)	q	(χ_1, χ_2)	(θ_1, θ_2)	χ_p	D_L (Mpc)	SNR	Median and 90% errors	$\mathcal{Q}_{\text{GR}}$ (z_{GR})	Median and 90% errors	$\mathcal{Q}_{\text{GR}}$ (z_{GR})
1.	30	2	(0.3,0.2)	$(\pi/9, \pi/4)$	0.10	1200	20	$0.0^{+0.18}_{-0.18}$	0.49 (0.02)	$-0.08^{+0.95}_{-0.97}$	0.56 (0.14)
							650	$0.02^{+0.11}_{-0.13}$	0.41 (0.21)	$-0.05^{+0.59}_{-0.6}$	0.56 (0.15)
2.	30	2	(0.8,0.7)	$(\pi/3, \pi/4)$	0.69	1200	20	$-0.04^{+0.2}_{-0.19}$	0.63 (0.29)	$-0.19^{+0.97}_{-0.93}$	0.64 (0.32)
							650	$-0.02^{+0.12}_{-0.11}$	0.61 (0.27)	$-0.09^{+0.58}_{-0.56}$	0.62 (0.27)

moderately and strongly precessing regimes, we consider two configurations differing in spin magnitudes and orientations. For a high-precession configuration, we consider $(\chi_1, \chi_2) = (0.8, 0.7)$ with tilt angles $(\theta_1, \theta_2) = (\pi/3, \pi/4)$, corresponding to $\chi_p = 0.69$. For a low-precession configuration, we use $(\chi_1, \chi_2) = (0.3, 0.2)$ with smaller tilt angles $(\theta_1, \theta_2) = (\pi/9, \pi/4)$, yielding $\chi_p = 0.102$. In each case, we choose $\theta_{JN} = 2.7$ radians, $\phi_{JL} = 1.7$ radians, and $\phi_{12} = 3$ radians. Each system is analyzed at two network SNRs: ~ 20 and ~ 40 , corresponding to low and high signal strengths, respectively.

For GR-consistent injection, we employ the IMRPhenomXP waveform model [63], a frequency-domain phenomenological waveform approximant that extends the IMRPhenomXAS [97] by incorporating spin-induced precession effects in the quadrupolar mode. For recovery in the parameter estimation process, we use the parametrized IMRPhenomXP [62] waveform model approximant to produce discrete samples of the posterior probability distribution given in Eq. (5) within the TIGER framework and subsequently perform the PCA-TGR analysis.

The binary configurations used in this analysis are summarized in Table II, which also reports the median and 90% credible intervals of the posteriors for the leading

two PCA parameters. Here, we find that the JS divergence between the posterior and prior distributions exceeds 0.1 bits only for the first two leading PCA parameters. The values of $\mathcal{Q}_{\text{GR}}$ and z_{GR} for the PCA posteriors are also provided. Figure 2 displays the posterior distributions of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ as violin plots for all injection configurations. We find that, in all cases, the GR value lies well within the 90% credible regions of the PCA parameters. This is reflected in the corresponding $\mathcal{Q}_{\text{GR}}$ and z_{GR} values reported in Table II. Moreover, as expected, the constraints on the PCA parameters tighten with increasing SNR, illustrating the dependence of the PCA-TGR method's constraining power on the signal strength.

3. Systems with precession and higher-order harmonics

We now extend our analysis to investigate the impact of higher-order harmonics and spin-induced precession by performing injection studies using binary parameters of known detected BBH systems. Specifically, we consider configurations inspired by three well-studied BBH merger events: GW151226 [100], GW190412 [101], and GW190814 [102]. GW190412 and GW190814 exhibit pronounced mass asymmetry, which enhances the amplitude of higher-order harmonics in the GW signal. In

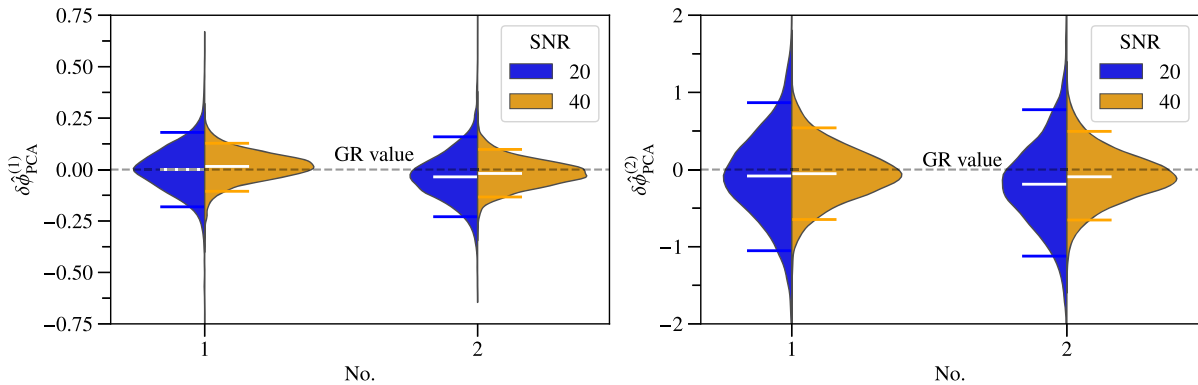


FIG. 2. Posterior distributions of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ (left panel) and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ (right panel) for each simulated precessing-spin GR injection—listed in Table II and discussed in Sec. III A 2—presented as violin plots. The colored and white horizontal bars have the same meaning as in Fig. 1. For all the binary configurations, both the PCA parameters are consistent with the GR value (dashed gray line) of 0.

TABLE III. Summary of the PCA-TGR results from injection studies using simulated GR-consistent signals from GW151226-like, GW190412-like, and GW190814-like BBH systems, conducted within the TIGER framework, as discussed in Sec. III A 3. Each row lists the waveform model used for injection (with recovery performed using its corresponding parametrized version, which is not separately listed), along with the binary parameters and the statistical properties of the leading two PCA parameters. Posterior distributions for the dominant two PCA components are shown in Fig. 3.

No.	Event	Waveform	Properties of GR injections					Properties of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$		Properties of $\delta\hat{\phi}_{\text{PCA}}^{(2)}$	
			$M (M_{\odot})$	q	(χ_1, χ_2)	(θ_1, θ_2)	SNR	Median and 90% errors	$\mathcal{Q}_{\text{GR}} (z_{\text{GR}})$	Median and 90% errors	$\mathcal{Q}_{\text{GR}} (z_{\text{GR}})$
1.	GW151226-like	IMRPhenomXPHM	23.72	1.91	(0.61, 0.455)	(1.02, 1.385)	20	$-0.08^{+0.18}_{-0.18}$	0.77 (0.7)	$0.06^{+0.83}_{-0.82}$	0.44 (0.13)
							40	$-0.02^{+0.08}_{-0.09}$	0.69 (0.46)	$-0.02^{+0.41}_{-0.4}$	0.53 (0.06)
2.	GW190412-like	IMRPhenomXHM	46.66	4.24	(0.46, 0.56)	(0,0)	35	$-0.03^{+0.14}_{-0.13}$	0.65 (0.32)	$0.11^{+0.75}_{-0.67}$	0.39 (0.26)
		IMRPhenomXPHM				(0.68, 1.66)	35	$0.01^{+0.15}_{-0.12}$	0.43 (0.15)	$-0.03^{+0.86}_{-0.79}$	0.53 (0.06)
3.	GW190814-like	IMRPhenomXHM	27.2	9	(0.03, 0.47)	(0,0)	25	$-0.02^{+0.1}_{-0.09}$	0.64 (0.32)	$0.13^{+0.6}_{-0.6}$	0.35 (0.34)
		IMRPhenomXPHM				(1.57, 1.46)	25	$0.0^{+0.08}_{-0.07}$	0.48 (0.04)	$0.08^{+0.56}_{-0.54}$	0.4 (0.23)

contrast, GW151226 is more symmetric in masses and features moderate spin-induced precession. The injection parameters—component masses, spin magnitudes, and tilt angles—are selected to approximate the inferred properties of these observed systems and are summarized in Table III.

To systematically disentangle the effects of precession and higher modes, we perform injections and recoveries using two types of waveform models. The first, IMRPhenomXHM [64] (for injection) and its parametrized version (for recovery), includes higher-order modes but assumes spins aligned with the orbital angular momentum. The second, IMRPhenomXPHM [63] (for injection) and its parametrized version (for recovery), extends the former by incorporating spin-precession effects.

For GW151226-like injections (mass ratio $q \sim 2$, $\chi_p \sim 0.52$), we use IMRPhenomXPHM for injection and parametrized IMRPhenomXPHM for recovery, with network SNRs of 20 and 40. For the mass-asymmetric GW190412-like ($q \sim 4$, $\chi_p \sim 0.29$) and GW190814-like ($q \sim 9$, $\chi_p \sim 0.07$) systems, we conduct separate injection studies: one set using IMRPhenomXHM with zero spin tilts (aligned-spin case) and its parametrized version for recovery, and another set using IMRPhenomXPHM along with its parametrized version to include spin precession effects. This strategy allows for a direct comparison of aligned and precessing versions of the same system, thereby isolating the contributions from spin precession and higher-order modes. Such comparisons are essential, as spin-induced precession can alter the mixing of angular modes, thereby modulating the detectability and influence of higher-order harmonics on the waveform. Verifying that PCA-based tests remain robust in these scenarios is crucial for assessing PCA-TGR’s applicability to future complex signals.

For GW190814-like injections, we use the noise PSDs estimated with BAYESWAVE [103,104] from the actual event data for each detector [102]. For the other two events,

we use the same detector-specific noise PSDs as in the previous sections. The results of this injection analysis are summarized in Table III and Fig. 3. Here, we also find that the JS divergence between the posterior and prior distributions exceeds 0.1 bits for the first two leading PCA parameters in all cases while it is marginally greater than 0.1 bits for the third leading PCA parameter for some cases. The statistics of the two leading PCA posteriors are listed in Table III. Figure 3 shows their posterior distributions. The top row corresponds to $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ and the bottom row to $\delta\hat{\phi}_{\text{PCA}}^{(2)}$. In each row, the left panel shows the posteriors of GW151226-like injections at low and high SNR. The right panel shows the posteriors for GW190412-like and GW190814-like injections, considering both aligned-spin and precessing cases. In summary, the PCA posteriors are consistent with GR in all cases, with the injected value of zero lying within the 90% credible intervals. The bounds on the PCA parameters for the GW190814-like injection are relatively tighter compared to the others, due to the greater strength of higher-order modes in the injected signal, which helps break degeneracies between different model parameters.

B. FTI framework

In the previous subsection, we studied the efficacy of PCA-TGR within the TIGER framework using a variety of simulated GR-consistent injected signals. We now shift out focus to the efficacy of PCA-TGR within the FTI framework. One important point to note is that, in its current implementation, the FTI framework uses aligned-spin waveform models.

We consider GW150914-like BBH systems, with injected binary parameters set to the median values obtained from the parameter estimation of the GW150914 event [105]. The luminosity distance is varied to achieve the desired network SNR values of 20 (low) and 40 (high). To simulate GR-consistent GW signals, we first

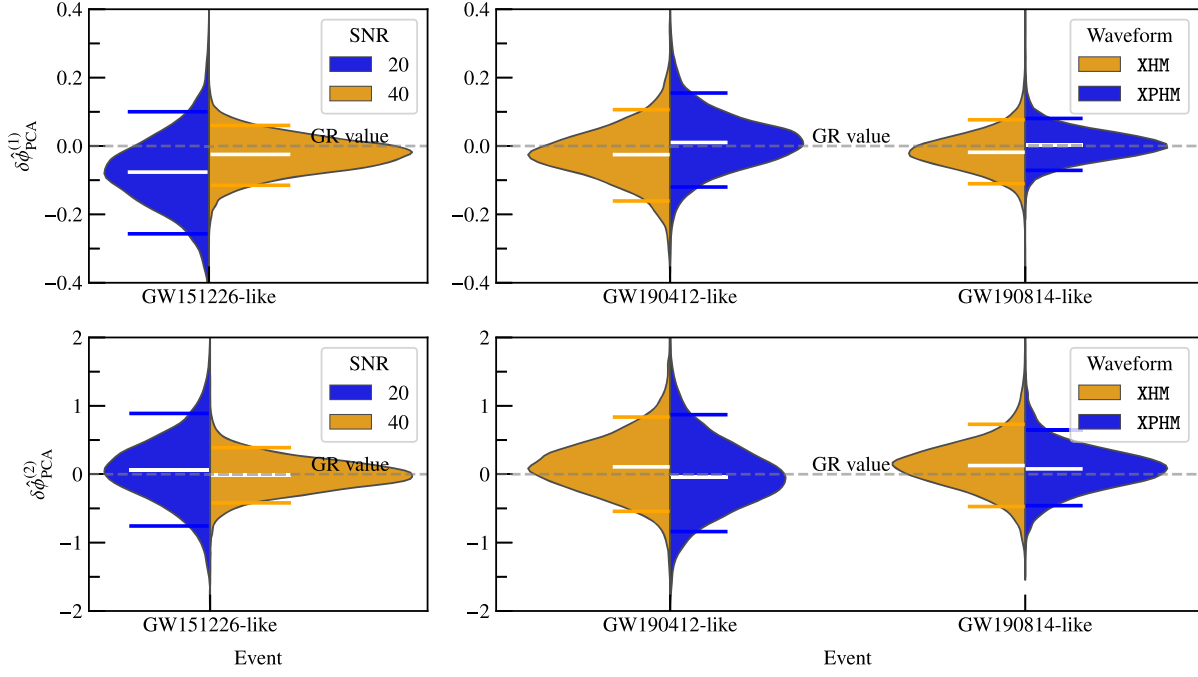


FIG. 3. Posterior distributions of the dominant two PCA parameters ($\delta\hat{\phi}_{\text{PCA}}^{(1)}$, top row; $\delta\hat{\phi}_{\text{PCA}}^{(2)}$, bottom row) obtained from GR-consistent injections based on GW151226-, GW190412-, and GW190814-like BBH systems, as described in Sec. III A 3. The left panel compares PCA posteriors for GW151226-like injections at network SNRs of 20 and 40, using IMRPhenomXPHM for injection and its parametrized version for recovery. The right panel compares aligned-spin (IMRPhenomXHM for injection and its parametrized version for recovery) and precessing (IMRPhenomXPHM for injection and its parametrized version for recovery) injections for GW190412-like and GW190814-like systems. The colored and white horizontal bars have the same meaning as in Fig. 1. In all cases, the PCA constraints are consistent with the GR value of 0, marked by the dashed gray line.

use the SEOBNRv4_ROM waveform approximant [67]. For parameter estimation, we employ the parametrized SEOBNRv4_ROM [23,66] model for signal recovery, producing discrete samples from the posterior probability distribution given in Eq. (5), within the FTI framework. We repeat the same exercise but now using the SEOBNRv4HM_ROM waveform model [68,69], which incorporates higher-order modes beyond the quadrupolar ones.

The results from this analysis are summarized in Table IV (first two entries) and Fig. 4. The JS divergence between the posterior and prior distributions exceeds 0.1 bits for the first three leading PCA parameters in all cases; however, we present results only for the first two. Similarly to the previous subsection, the statistical properties of the two leading PCA parameters are listed in Table IV, while their posterior distributions are shown as violin plots in Fig. 4. In every case, the PCA posteriors are found to be consistent with GR, with the injected value of zero lying within the 90% credible intervals. The PCA-TGR analysis with similar GR injections has also been repeated using the SEOBNRv5_ROM waveform family [106], yielding results similar to those from SEOBNRv4_ROM, although these results are not included here for brevity.

Within the FTI framework, we also explore the effect of spin precession on the PCA-TGR analysis since the

waveform models used in the FTI analysis do not account for it. We consider the BBH system with a total mass of $30M_{\odot}$, a mass ratio of $q = 2$, and a network SNR of approximately 20. We study this system using two different values of the effective spin-precession parameter, $\chi_p = 0.102$ (weak precession) and $\chi_p = 0.69$ (strong precession), obtained by varying the spin magnitudes and tilt angles. For each case, GR-consistent injected signals are simulated using the IMRPhenomXPHM waveform model, which incorporates both higher-order modes and precession effects. Parameter estimation is performed using the parametrized SEOBNRv4HM_ROM model, which accounts for higher-order modes but not for precession.

Similar to the previous injection analysis, we again find that the JS divergence between the posterior and prior distributions exceeds 0.1 bits for the three leading PCA parameters in both injections. The statistical properties of the first two PCA parameters for both cases are shown in entries 3 and 4 of Table IV, respectively. Since this is a GR injection, any apparent GR deviation observed in the posteriors of the PCA parameters can be attributed to the effects of unaccounted spin precession in the waveform model used for parameter estimation during recovery. As shown in Fig. 5, the posterior medians of the leading two PCA parameters exhibit significant offsets from zero in both cases. In particular, the GR value is excluded at

TABLE IV. Summary of the PCA-TGR results from the simulated BBH GR injection study conducted within the FTTI framework, as described in Sec. III B. Each row lists the binary parameters, the waveform models used for injection and recovery, and the statistical properties of the leading two PCA parameters. Posterior distributions for the leading two PCA components are shown in Figs. 4 and 5 corresponding to the first two and the last two entries in the table, respectively.

No.			Properties of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$		Properties of $\delta\hat{\phi}_{\text{PCA}}^{(2)}$		
			Median and		Median and		
Event	Waveform injection Waveform recovery	SNR	90% errors	$\mathcal{Q}_{\text{GR}} (z_{\text{GR}})$	90% errors	$\mathcal{Q}_{\text{GR}} (z_{\text{GR}})$	
1.	GW150914-like	SEOBNRv4_ROM	20	$0.03^{+0.22}_{-0.13}$	0.39 (0.23)	$-0.04^{+0.55}_{-0.65}$	0.55 (0.12)
		parametrized SEOBNRv4_ROM	40	$0.02^{+0.15}_{-0.08}$	0.35 (0.31)	$0.01^{+0.35}_{-0.32}$	0.49 (0.04)
2.	GW150914-like	SEOBNRv4HM_ROM	20	$0.02^{+0.15}_{-0.08}$	0.38 (0.26)	$-0.03^{+0.53}_{-0.61}$	0.53 (0.08)
		parametrized SEOBNRv4HM_ROM	40	$0.01^{+0.1}_{-0.06}$	0.38 (0.26)	$0.02^{+0.32}_{-0.28}$	0.46 (0.09)
Effect of precession on PCA							
3.	$M = 30M_{\odot}, q = 2, \chi_p = 0.69$	IMRPhenomXPHM	20	$-0.07^{+0.02}_{-0.15}$	0.91 (1.31)	$-0.62^{+0.26}_{-1.04}$	1.0 (2.59)
		parametrized SEOBNRv4HM_ROM					
4.	$M = 30M_{\odot}, q = 2, \chi_p = 0.102$	IMRPhenomXPHM	20	$-0.06^{+0.04}_{-0.15}$	0.86 (0.93)	$-0.23^{+0.15}_{-0.68}$	0.83 (0.9)
		parametrized SEOBNRv4HM_ROM					

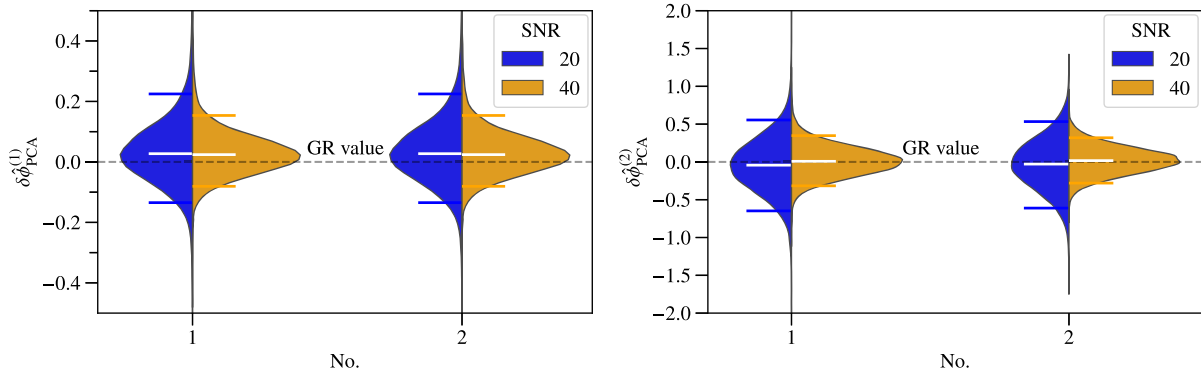


FIG. 4. Violin plots showing the posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ from the GR-consistent aligned-spin injection analysis within the FTTI framework, using the SEOBNRv4_ROM (x-label:1) and SEOBNRv4HM_ROM (x-label:2) waveform models for GW150914-like BBH systems, as listed in entries 1 and 2 of Table IV and discussed in Sec. III B. The horizontal bars have the same meaning as in Fig. 1. In every case, the PCA posteriors remain consistent with the GR prediction of zero, indicated by the dashed gray line.

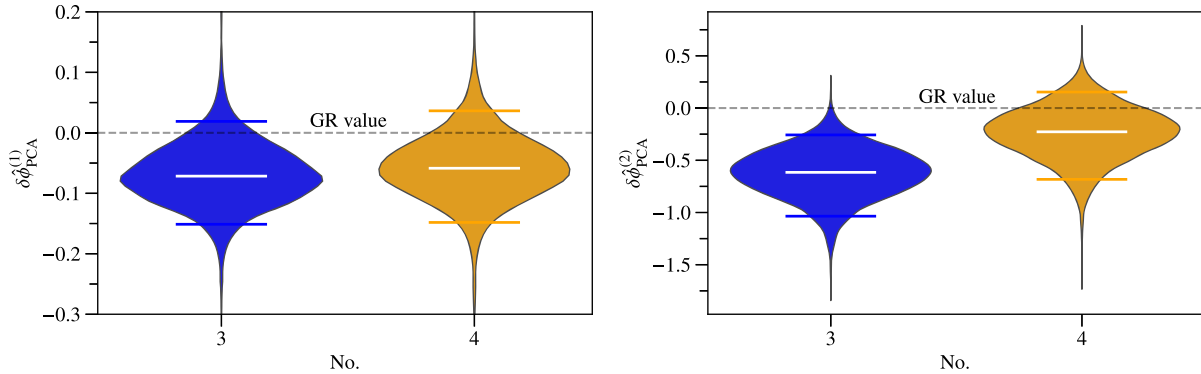


FIG. 5. Violin plots showing the posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ from the GR-consistent precessing-spin injection analysis within the FTTI framework, as presented in entries 3 and 4 of Table IV and discussed in Sec. III B. GR-consistent GW signals are simulated using the IMRPhenomXPHM waveform model, and parameter estimation is performed using the parametrized SEOBNRv4HM_ROM model. The horizontal bars have the same meaning as in Fig. 1. The median values of the PCA parameter posteriors show notable deviations from zero in all cases. Specifically, for the strongly precessing binary configuration, the GR value is excluded from the 90% credible interval of the $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ posterior.

90% confidence in the posterior of $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ for the strongly precessing binary configuration. The posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(3)}$ are almost an order of magnitude less constrained than those of $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ and are not shown here. Although their posteriors are consistent with GR, the posterior medians show a slight offset from zero. The conclusion from this injection study is that using an aligned-spin waveform model in the FTI framework could lead to a false violation of GR in the PCA-TGR analysis if significant spin precession is present in the signal.

IV. RESPONSE OF PCA-TGR TO NON-GR INJECTIONS

After demonstrating the ability of PCA-TGR to recover GR in an unbiased manner when the signal in the data is that of a BBH merger in GR—with the exception of its response in the FTI framework when an aligned-spin waveform is used to analyze strongly precessing signals—we now turn

our attention to non-GR injections and PCA-TGR’s ability to detect GR deviations.

A. TIGER framework

In order to simulate GR violating signals, we consider BBH waveforms whose inspiral phasing coefficients are deformed at different PN order using the TIGER framework. We consider different types of non-GR injections which are summarized in Table V. Broadly speaking, the injections cover the following three scenarios:

- (1) All the PN coefficients above 1PN are deformed by the same magnitude. We consider small, moderate, and large deviations which correspond to fractional deviations of 0.1, 0.5, and 0.9, respectively.
- (2) The fractional deviations in the PN coefficients above 1PN increase (decrease) with PN order, ranging from 0.3 to 0.8 in steps of 0.1 (or vice versa).
- (3) All the PN coefficients above 1PN are deformed by the same magnitude but by -0.5 to study the

TABLE V. Summary of results from the simulated non-GR injection study within the TIGER framework that demonstrates the capability of the PCA-TGR method to detect GR violations in the PN phasing coefficients. The values of the fractional PN deviation parameter, binary systems, and the network SNRs (rounded to the nearest integer) of the injected signals are listed. For each non-GR injection, the table shows the median value, as well as the upper and lower limits of the 90% credibility interval, $\mathcal{Q}_{\text{GR}}$, and z_{GR} for the posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$. The posteriors on $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ exclude GR at $>3\sigma$ in most of the cases, while the posteriors on $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ encompass the GR value within 90% credibility in all cases. The PCA-TGR method can rule out GR at $>3\sigma$ for a fractional deviation of magnitude >0.5 in the PN phasing coefficients.

Properties of non-GR injections			Properties of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$				Properties of $\delta\hat{\phi}_{\text{PCA}}^{(2)}$			
No.	$\{\delta\hat{\phi}_k\}$	SNR	Median and 90% errors		$\mathcal{Q}_{\text{GR}}$ (z_{GR})		Median and 90% errors		$\mathcal{Q}_{\text{GR}}$ (z_{GR})	
			GW1509	GW1512	GW1509	GW1512	GW1509	GW1512	GW1509	GW1512
			14-like	26-like	14-like	26-like	14-like	26-like	14-like	26-like
1.	$\delta\hat{\phi}_a = 0(a = -2, 0, 1, 2)$	20	$0.10^{+0.25}_{-0.26}$	$0.00^{+0.20}_{-0.23}$	0.26 (0.62)	0.50 (0.01)	$-0.07^{+1.71}_{-2.04}$	$0.07^{+0.84}_{-1.00}$	0.53 (0.06)	0.45 (0.11)
	$\delta\hat{\phi}_b = 0.1(b = 3, 4, 5l, 6, 6l, 7)$	40	$0.11^{+0.17}_{-0.14}$	$0.06^{+0.11}_{-0.11}$	0.09 (1.17)	0.16 (0.94)	$-0.07^{+0.98}_{-1.15}$	$0.03^{+0.41}_{-0.44}$	0.55 (0.11)	0.45 (0.10)
2.	$\delta\hat{\phi}_a = 0(a = -2, 0, 1, 2)$	20	$0.51^{+0.19}_{-0.19}$	$0.41^{+0.24}_{-0.24}$	0.00 (4.09)	0.01 (2.74)	$0.03^{+2.16}_{-2.05}$	$0.54^{+0.96}_{-0.84}$	0.49 (0.02)	0.15 (0.98)
	$\delta\hat{\phi}_b = 0.5(b = 3, 4, 5l, 6, 6l, 7)$	40	$0.52^{+0.13}_{-0.12}$	$0.48^{+0.14}_{-0.14}$	0.00 (6.50)	0.00 (5.66)	$-0.17^{+1.16}_{-1.23}$	$0.15^{+0.53}_{-0.46}$	0.59 (0.23)	0.30 (0.49)
3.	$\delta\hat{\phi}_a = 0(a = -2, 0, 1, 2)$	20	$0.92^{+0.15}_{-0.16}$	$0.90^{+0.20}_{-0.16}$	0.00 (9.16)	0.00 (7.90)	$-0.03^{+2.89}_{-2.67}$	$0.48^{+1.11}_{-0.95}$	0.51 (0.02)	0.21 (0.75)
	$\delta\hat{\phi}_b = 0.9(b = 3, 4, 5l, 6, 6l, 7)$	40	$0.93^{+0.12}_{-0.10}$	$0.93^{+0.12}_{-0.10}$	0.00 (12.83)	0.00 (13.17)	$-0.31^{+1.49}_{-1.64}$	$0.12^{+0.57}_{-0.53}$	0.63 (0.33)	0.35 (0.37)
4.	$\delta\hat{\phi}_a = 0(a = -2, 0, 1, 2)$	20	$0.35^{+0.21}_{-0.23}$	$0.18^{+0.25}_{-0.26}$	0.01 (2.48)	0.12 (1.16)	$0.09^{+1.89}_{-1.97}$	$0.48^{+0.93}_{-0.77}$	0.47 (0.08)	0.15 (0.89)
	$\delta\hat{\phi}_3 = 0.3, \delta\hat{\phi}_4 = 0.4$	40	$0.37^{+0.14}_{-0.14}$	$0.30^{+0.13}_{-0.14}$	0.00 (4.13)	0.00 (3.58)	$-0.04^{+1.07}_{-1.15}$	$0.28^{+0.45}_{-0.39}$	0.52 (0.05)	0.12 (1.06)
	$\delta\hat{\phi}_{5l} = 0.5, \delta\hat{\phi}_6 = 0.6$									
	$\delta\hat{\phi}_{6l} = 0.7, \delta\hat{\phi}_7 = 0.8$									
5.	$\delta\hat{\phi}_a = 0(a = -2, 0, 1, 2)$	20	$0.77^{+0.17}_{-0.16}$	$0.73^{+0.21}_{-0.17}$	0.00 (7.17)	0.00 (6.17)	$-0.05^{+2.64}_{-2.33}$	$0.43^{+1.17}_{-0.95}$	0.51 (0.03)	0.23 (0.67)
	$\delta\hat{\phi}_3 = 0.8, \delta\hat{\phi}_4 = 0.7$	40	$0.78^{+0.11}_{-0.10}$	$0.78^{+0.14}_{-0.11}$	0.00 (11.03)	0.00 (10.43)	$-0.37^{+1.35}_{-1.41}$	$0.07^{+0.60}_{-0.52}$	0.67 (0.44)	0.42 (0.21)
	$\delta\hat{\phi}_{5l} = 0.6, \delta\hat{\phi}_6 = 0.5$									
	$\delta\hat{\phi}_{6l} = 0.4, \delta\hat{\phi}_7 = 0.3$									
6.	$\delta\hat{\phi}_a = 0(a = -2, 0, 1, 2)$	20	$-0.50^{+0.30}_{-0.27}$	$-0.56^{+0.21}_{-0.20}$	0.99 (2.83)	1.00 (4.33)	$0.00^{+1.86}_{-2.42}$	$0.00^{+0.90}_{-1.03}$	0.50 (0.00)	0.50 (0.00)
	$\delta\hat{\phi}_b = -0.5(b = 3, 4, 5l, 6, 6l, 7)$	40	$-0.49^{+0.21}_{-0.17}$	$-0.53^{+0.13}_{-0.11}$	1.00 (4.11)	1.00 (7.59)	$-0.03^{+1.15}_{-1.56}$	$-0.11^{+0.51}_{-0.52}$	0.52 (0.04)	0.65 (0.34)

difference between a positive fractional deviation and a negative one.

We consider GW150914- and GW151226-like BBH systems, using the same binary parameters as in the previous section. We take two different values of the luminosity distance for each of the chosen BBH systems to produce two different network SNRs of ~ 20 and ~ 40 . We use the parametrized IMRPhenomXP waveform model to generate the injected GW signal with GR violations and to perform parameter estimation while analyzing the injected signal in the TIGER framework. Indeed, the injected GR violations are *ad hoc* and do not follow the predictions of any known modified theories of gravity. Our aim here is to assess the capability of the PCA-TGR to detect a violation of GR when the GW signal does not follow the GR prediction.

Our results are summarized in Table V and Fig. 6. We find that, in all injections, the JS divergence between the posterior and prior distributions exceeds 0.1 bits only for the first two leading PCA parameters. Figure 6 shows the violin plots for the posterior probability distributions of the leading two PCA parameters from our injection analysis using the TIGER framework. From Table V and Fig. 6, we find that the posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ exclude GR at $> 3\sigma$ level in most of the cases when we consider a fractional deviation of magnitude > 0.1 in the PN phasing coefficients and for all

cases with SNR of 40. However, for a fractional deviation of magnitude 0.1, the PCA-TGR is unable to detect GR violation with a high confidence level even for signals with a network SNR of 40. Our simulations currently are not exhaustive enough to find the SNR at which a fractional deviation of 0.1 at every PN order will show up as a GR violation in the PCA parameters. However, a simple scaling suggests that a minimum SNR of ~ 103 (126) is required to detect this level of GR violation at more than the 3σ level for a GW150914-like (GW151226-like) system. In all the simulations considered here, the subleading PCA parameters show consistency with GR within 90% credibility. For deviations that are fractional in nature, such as those considered here, $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ is most efficient in detecting a GR violation. However, we cannot comment whether this is a generic feature for all possible types of GR violations with our limited set of injections. As these deviations are not motivated by any specific theory of gravity but are representative examples for demonstration of the method, we do not consider more combinations of the fractional deformation values. We will discuss other forms of GR deviations and the response of the PCA-TGR in the next section.

In most of the non-GR injections, the error bars on the first PCA parameter $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ are comparable for GW150914-like and GW151226-like systems, with GW150914-like injections providing slightly better constraints in most

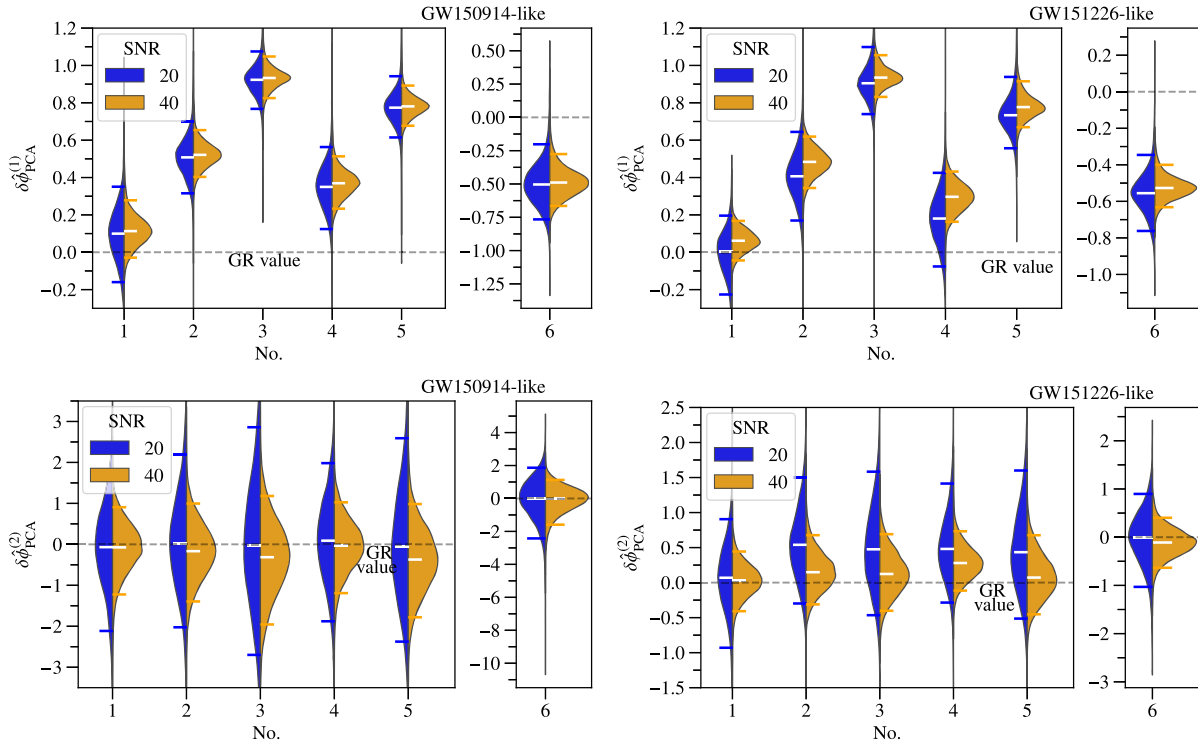


FIG. 6. Posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ (upper panel) and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ (lower panel) for each simulated non-GR injection (the left and right panels show the results obtained from GW150914- and GW151226-like BBH injections, respectively) listed in Table V presented as violin plots. The results from the injections with network SNR ~ 20 (~ 40) are shown in blue (orange). The horizontal bars have the same meaning as in Fig. 1. We mark the GR value of zero with dashed gray lines.

TABLE VI. Summary of results from the simulated non-GR injection study within the FTI framework that demonstrates the capability of the PCA-TGR to detect GR violations in the PN phasing coefficients. The injected values of the fractional PN deviation parameter, injected binary systems, and the network SNRs (rounded to the nearest integer) of the injected signals are listed. For each non-GR injection, the table shows the median value, as well as the upper and lower limits of the 90% credibility interval, $\mathcal{Q}_{\text{GR}}$, and z_{GR} for the posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$. The posteriors on $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ exclude GR at $>3\sigma$ in the second case but not in the first case, while the posteriors on $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ encompass the GR value within 90% credibility in both cases. The PCA-TGR can rule out GR at $>3\sigma$ for a fractional deviation of magnitude >0.5 in the PN phasing coefficients.

No.	Properties of Non-GR injections $\{\delta\hat{\phi}_k\}$	SNR	Properties of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$				Properties of $\delta\hat{\phi}_{\text{PCA}}^{(2)}$			
			Median and 90% errors		$\mathcal{Q}_{\text{GR}}$ (z_{GR})		Median and 90% errors		$\mathcal{Q}_{\text{GR}}$ (z_{GR})	
			GW1509 14-like	GW1512 26-like	GW1509 14-like	GW1512 26-like	GW1509 14-like	GW1512 26-like	GW1509 14-like	GW1512 26-like
1.	$\delta\hat{\phi}_a = 0 (a = -2, 0, 2),$ $\delta\hat{\phi}_b = 0.1 (b = 3, 4, 5l, 6, 6l, 7)$	20	$0.11^{+0.11}_{-0.09}$	$0.04^{+0.14}_{-0.12}$	0.02 (1.76)	0.27 (0.47)	$-0.05^{+0.55}_{-0.60}$	$-0.24^{+0.44}_{-0.59}$	0.56 (0.14)	0.81 (0.76)
2.	$\delta\hat{\phi}_a = 0 (a = -2, 0, 2),$ $\delta\hat{\phi}_b = 0.5 (b = 3, 4, 5l, 6, 6l, 7)$	20	$0.51^{+0.11}_{-0.10}$	$0.46^{+0.13}_{-0.14}$	0.00 (7.51)	0.00 (5.28)	$0.11^{+0.62}_{-0.64}$	$-0.03^{+0.42}_{-0.55}$	0.39 (0.29)	0.54 (0.09)

cases. In contrast, GW151226-like systems consistently provide better constraints on the second PCA parameter $\delta\hat{\phi}_{\text{PCA}}^{(2)}$. For $\delta\hat{\phi}_{\text{PCA}}^{(1)}$, GW150914-like injections generally yield higher z_{GR} values compared to GW151226-like injections, except for injection No. 6 and the SNR 40 case in injection No. 3. This is likely due to the trade-off between the relative weights of the different $\delta\hat{\phi}_b$ in the PCA parameters, the magnitudes of these deformations, and the measurement uncertainties associated with these deviations. For detector configurations with better low-frequency sensitivity, GW151226-like systems may still produce high z_{GR} values. For the second PCA parameter $\delta\hat{\phi}_{\text{PCA}}^{(2)}$, GW151226-like injections generally give higher z_{GR} values than GW150914-like injections, except in the SNR 40 case of injection Nos. 1 and 5. However, for $\delta\hat{\phi}_{\text{PCA}}^{(2)}$, we find the deviations from GR are much less significant across the simulations, as mentioned earlier. This makes the comparison across different systems less meaningful, as we have not quantified the statistical error bars on the z scores, and the observed differences may, in many cases, lie within these unquantified uncertainties. Broadly speaking, this trend indicates that the second PCA parameter is unable to detect GR violations for the specific types of fractional deviations used in this analysis, particularly for signals with $\text{SNR} \leq 40$.

B. FTI framework

We now examine a couple of non-GR injections previously considered in the TIGER framework but now analyzed within the FTI framework. Specifically, we consider the first two cases in Table V, which correspond to fractional deviations of 0.1 (where the PCA-TGR method

could not rule out GR at $>3\sigma$ in the TIGER framework) and 0.5 (where the PCA-TGR method could rule out GR at $>3\sigma$ in the TIGER framework). The injection setups are identical to those in the previous subsection, except that here we use the parametrized SEOBNRv4HM_ROM waveform model both to generate the injected GW signals with GR violations (which therefore do not include the effects of precession) and to perform parameter estimation within the FTI framework. We also restrict our analysis to scenarios with a network SNR of approximately 20.

The results from this analysis are presented in Fig. 7 and Table VI. Figure 7 shows violin plots of the posterior probability distributions for the leading two PCA parameters from our injection analysis using the FTI framework.

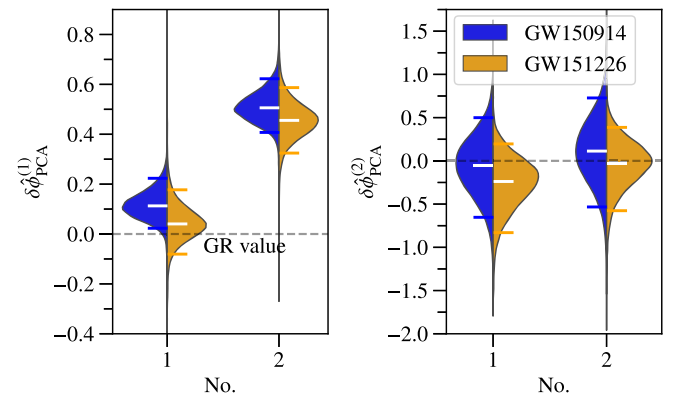


FIG. 7. Posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ (left panel) and $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ (right panel) for each simulated injection listed in Table VI presented as violin plots. The results from GW150914-like (GW151226-like) BBH injections are shown in blue (orange). The horizontal bars have the same meaning as in Fig. 1. We mark the GR value of zero with dashed gray lines.

The posterior medians, 90% credible intervals, $\mathcal{Q}_{\text{GR}}$, and z_{GR} values for the two leading PCA parameters for each injection are summarized in Table VI. The posteriors on $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ exclude GR at $>3\sigma$ in the second case (corresponding to a fractional deviation of 0.5) but not in the first case (with a fractional deviation of 0.1). In contrast, the posteriors on $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ encompass the GR value within the 90% credible interval in both cases. These results demonstrate that the PCA-TGR, within the FTI framework, can robustly identify GR violations at the $>3\sigma$ level when the deviation in the PN phasing coefficients exceeds 0.5 in magnitude—consistent with our earlier findings using the TIGER framework. However, in the case with a fractional deviation of 0.5, the posteriors on $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ in the FTI framework show a slightly stronger deviation from GR compared to the TIGER framework (see the z_{GR} values in Tables V and VI). This enhancement arises because the PCA parameters are slightly better constrained in the FTI framework than in the TIGER framework, due to differences in how the end of the inspiral regime is defined in the two frameworks (see the first paragraph of Sec. VI).

In summary, we find that the PCA-TGR is successfully able to detect GR violations of magnitude 0.5 with a high confidence level. The ability to detect a GR violation depends on the magnitude of violation and the SNR of the signal.

V. RESPONSE OF PCA-TGR TO INJECTIONS WITH ECCENTRICITY

Previous studies have found that the unmodeled eccentricity in the parametrized waveform models can lead to false violation of GR in TIGER/FTI tests if the BBH has

significant residual eccentricity by the time it enters the detector frequency band [107,108]. In this section, we study the response of PCA-TGR to eccentric BBH signals when we employ a parametrized quasicircular inspiral-merger-ringdown waveform model for inference. The question is how the different PN deformation parameters of the quasicircular inspiral-merger-ringdown waveform respond to the neglect of eccentricity in the signal and how the response translates into the PCA parameters.

We adopt the simulated eccentric BBH signals used in Ref. [108]. These eccentric signals are nonspinning BBH NR simulations from the SXS catalog [109] (see also Ref. [110]) with mass ratios $q = 1, 2, 3$ and a redshifted total mass of $80M_{\odot}$ at a luminosity distance of 400 Mpc. Other extrinsic parameters take the values given in [108]. For each mass ratio, we consider simulated signals with eccentricities, defined at 17 Hz, $e_{17} \sim 0.05, \sim 0.1$ (representing binaries with small eccentricities) and $<10^{-4}$ (representative of BBH in quasicircular orbits). For each set of BBH parameters, we consider injections which have (2, 2) and (3, 2) modes of the NR simulation and analyze the simulated signal using the parametrized IMRPhenomXPHM waveform model. The injection SNRs for the (2, 2) + (3, 2) mode for mass ratios $q = 1, 2, 3$ are 67, 230, and 137, respectively. We briefly summarize the properties of the simulations adopted in Table VII. The reader may refer to Sec. III of [108] for a more detailed discussion of the simulations used here.

The JS divergence between the posterior and prior distributions exceeds 0.1 bits for the first three leading PCA parameters in all injections, except for those with ($q = 1, e_{17} = 0.053, 0.097$) and ($q = 3, e_{17} = 0.093, <10^{-4}$), where the JS divergence for the third leading

TABLE VII. Summary of results from the simulated eccentricity injection analysis performed in Sec. V that studies the response of the current implementation of the PCA-TGR framework to the eccentric BBH signals. The properties of the simulated eccentricity injections are listed (see the text for more details). For each injection, the table shows the median value, as well as the upper and lower limits of the 90% credibility interval, $\mathcal{Q}_{\text{GR}}$, and z_{GR} for the posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$, $\delta\hat{\phi}_{\text{PCA}}^{(2)}$, and $\delta\hat{\phi}_{\text{PCA}}^{(3)}$. The posteriors on PCA parameters in all the cases with $e_{17} < 10^{-4}$ show consistency with GR at 90% credibility. For the injections with $e_{17} \sim 0.05$ and $e_{17} \sim 0.1$, the posteriors on at least one of the PCA parameters exclude GR at $>90\%$ credibility.

ID	Properties of injections			Properties of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$		Properties of $\delta\hat{\phi}_{\text{PCA}}^{(2)}$		Properties of $\delta\hat{\phi}_{\text{PCA}}^{(3)}$	
	Mass ratio	e_{17}	Mode	Median and 90% errors	$\mathcal{Q}_{\text{GR}}$ (z_{GR})	Median and 90% errors	$\mathcal{Q}_{\text{GR}}$ (z_{GR})	Median and 90% errors	$\mathcal{Q}_{\text{GR}}$ (z_{GR})
SXS:BBH:1155	1	$<10^{-4}$	(2, 2) + (3, 2)	$0.01^{+0.01}_{-0.02}$	0.27 (0.57)	$0.00^{+0.17}_{-0.19}$	0.48 (0.04)	$1.26^{+9.02}_{-10.33}$	0.41 (0.22)
SXS:BBH:1355	1	0.053	(2, 2) + (3, 2)	$-0.04^{+0.02}_{-0.03}$	1.00 (2.44)	$-0.54^{+0.86}_{-0.47}$	0.85 (1.33)	$1.59^{+12.27}_{-12.08}$	0.42 (0.22)
SXS:BBH:1357	1	0.097	(2, 2) + (3, 2)	$0.25^{+0.05}_{-0.06}$	0.00 (7.83)	$-1.30^{+0.63}_{-0.86}$	1.00 (2.91)	$9.26^{+11.76}_{-22.08}$	0.25 (0.88)
SXS:BBH:1222	2	$<10^{-4}$	(2, 2) + (3, 2)	$0.01^{+0.03}_{-0.04}$	0.37 (0.26)	$-0.01^{+0.28}_{-0.27}$	0.54 (0.08)	$-0.68^{+10.59}_{-10.94}$	0.54 (0.10)
SXS:BBH:1364	2	0.044	(2, 2) + (3, 2)	$0.07^{+0.04}_{-0.04}$	0.00 (2.75)	$-0.08^{+0.21}_{-0.21}$	0.73 (0.59)	$-14.25^{+8.67}_{-4.41}$	0.99 (3.49)
SXS:BBH:1368	2	0.097	(2, 2) + (3, 2)	$0.06^{+0.03}_{-0.02}$	0.00 (4.77)	$-0.64^{+0.25}_{-0.25}$	1.00 (4.20)	$-11.22^{+9.18}_{-9.16}$	0.98 (2.03)
SXS:BBH:2265	3	$<10^{-4}$	(2, 2) + (3, 2)	$-0.08^{+0.22}_{-0.25}$	0.76 (0.52)	$0.80^{+1.70}_{-1.58}$	0.20 (0.82)	$-0.52^{+19.46}_{-18.93}$	0.52 (0.04)
SXS:BBH:1371	3	0.055	(2, 2) + (3, 2)	$-0.03^{+0.02}_{-0.03}$	0.99 (2.49)	$-0.18^{+0.16}_{-0.15}$	0.97 (1.87)	$-10.10^{+7.38}_{-6.07}$	0.98 (2.43)
SXS:BBH:1373	3	0.093	(2, 2) + (3, 2)	$0.28^{+0.04}_{-0.03}$	0.00 (13.22)	$-4.91^{+0.32}_{-0.39}$	1.00 (22.28)	$-3.53^{+18.34}_{-15.67}$	0.63 (0.34)

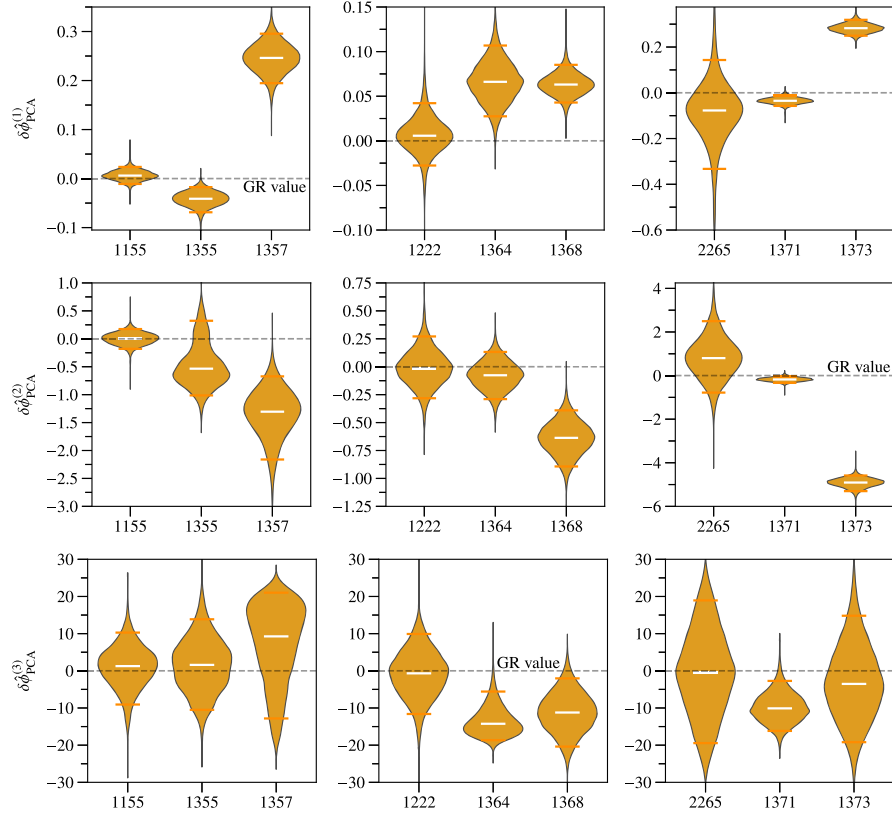


FIG. 8. Posteriors of $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ (upper panel), $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ (middle panel), and $\delta\hat{\phi}_{\text{PCA}}^{(3)}$ (lower panel) for each simulated eccentric injection listed in Table VII presented as violin plots. The left, middle, and right panels show the results obtained from eccentric BBH injections with mass ratios 1, 2, and 3, respectively. The horizontal bars have the same meaning as in Fig. 1. We mark the GR value of zero with dashed gray lines.

PCA parameter falls below 0.1 bits. Nevertheless, we include a discussion of the third PCA parameter $\delta\hat{\phi}_{\text{PCA}}^{(3)}$, as it can carry a significant imprint of apparent GR violation in certain cases, as discussed below. We show the violin plots for the posterior probability distributions of the leading three PCA parameters for all the eccentric injections in Fig. 8. We also present the posterior median and 90% credible intervals, $\mathcal{Q}_{\text{GR}}$, and z_{GR} of the leading three PCA parameters for each injection in Table VII. As expected, the posteriors on PCA parameters in all the quasicircular ($e_{17} < 10^{-4}$) injection analyses show consistency with GR at 90% credibility. In all other eccentricity ($e_{17} \sim 0.05$ and $e_{17} \sim 0.1$) injection analyses, the posteriors on at least one of the PCA parameters (more than one PCA parameter in most of the cases) exclude GR at $> 90\%$ credibility. It can be seen that $\delta\hat{\phi}_{\text{PCA}}^{(3)}$ is the least informative with respect to detecting a GR violation for most of the cases. For all the higher-eccentricity ($e_{17} \sim 0.1$) injection cases, the posterior distributions on at least one of the PCA parameters show strong GR violation with GR excluded at the $> 3\sigma$ level. More precisely, as expected, the amount of GR violation detected increases as eccentricity increases for all mass ratios. For $q = 1, 2$, $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ picks

larger GR violation than $\delta\hat{\phi}_{\text{PCA}}^{(2)}$, while for $q = 3$ this trend is reversed and the second PCA parameter shows larger GR violation. In the previous section, we saw that only the posteriors on $\delta\hat{\phi}_{\text{PCA}}^{(1)}$ show strong GR deviation when one considers fractional deviations in PN coefficients from GR. However, in the eccentric injection analyses, in many cases the posteriors on $\delta\hat{\phi}_{\text{PCA}}^{(2)}$ (and occasionally even the $\delta\hat{\phi}_{\text{PCA}}^{(3)}$) also exclude GR at $> 90\%$ credibility.

Unlike the “linear” fractional deviations we considered in the previous section, the eccentric BBH signal in GR induces “nonlinear” modifications to the PN phasing of the waveform of BBH in a quasicircular orbit in GR. By nonlinear we mean deviations that cannot be simply absorbed into deformations of the existing PN phasing coefficients. Hence, when we perform a parameter estimation with parametrized IMRPhenomXPHM on these injections, the eccentricity-induced corrections will show up as deviations in all of the PN coefficients with different degrees of departures which propagate into the posteriors on the PCA parameters, as we see. Therefore, it is reasonable to expect that for generic departures from GR (or unknown waveform systematics)—not limited to the specific types of fractional deviations considered in the previous section—

more than one PCA parameter can show statistically significant deviations from GR.

To summarize, we find that the unmodeled eccentricity in the parametrized waveform models might show up as an apparent GR violation in the posteriors of PCA parameters if we apply the PCA-TGR method to the eccentric BBH signals. Furthermore, we find that the nonlinear modifications to the PN phasing could make the posteriors of the higher-order PCA parameters deviate significantly from GR.

VI. ANALYSIS OF GWTC-3 EVENTS

We now present the results obtained by applying the PCA-TGR within the TIGER and FTI frameworks to selected GW events detected during the first three observing runs (O1, O2, and O3) of the Advanced LIGO and Advanced Virgo detectors [52,53,111,112]. The multiparameter runs involve sampling from 21- and 17-dimensional parameter spaces for precessing and aligned-spin BBHs, respectively, which is computationally expensive. Moreover, the combined bounds on the PCA parameters are largely dominated by the loudest and longest-duration inspiral signals in the catalog, which contribute the majority of the overall information content. To manage computational cost, ensure a longer inspiral phase, and minimize the impact of noise artifacts, we impose selection criteria that reduce the number of BBH events analyzed while retaining the most informative signals. Specifically, we focus on events that are sufficiently loud—thereby reducing potential biases due to noise—by requiring a false-alarm rate of $\leq 10^{-3} \text{ yr}^{-1}$, a network SNR ≥ 15 in the inspiral regime, and a redshifted chirp mass $\leq 20M_{\odot}$.

The end of the inspiral regime is defined by the MECO frequency in the TIGER framework and by the peak frequency of the (2, 2) mode in the FTI framework. The optimal SNR in the inspiral regime is computed using the maximum likelihood parameters from standard BBH parameter estimation under the GR hypothesis: The maximum likelihood parameters are extracted from parameter estimation samples with the tag C01:IMRPhenomXPHM, provided by the LVK Collaboration and accessible on Zenodo [113]. To compute the inspiral SNR, we use the IMRPhenomXPHM waveform model for the TIGER analysis and SEOBNRv4HM_ROM waveform model for FTI analysis. For the chirp mass, we use the median of its posterior distribution from the same GR parameter estimation analysis. Both the optimal SNR in the inspiral regime and the chirp mass are rounded to the nearest integer. We find that only four events in GWTC-3 satisfy the selection criteria: GW190412, GW190814, GW191204_171526, and GW191216_213338. The optimal SNRs in the inspiral regime, along with the upper frequency cutoffs that define the end of the inspiral regime for these four selected events, are listed in Table VIII.

We perform the PCA-TGR analysis on the above-chosen events within the TIGER and FTI frameworks. The FTI multiparameter runs are performed with the parametrized SEOBNRv5HM_ROM waveform [106], while the TIGER multiparameter runs are performed with the parametrized IMRPhenomXPHM-SpinTaylor waveform [65].⁷ Figure 9 shows the posterior distributions of the first three dominant PCA parameters that capture the most information (JS divergence between the prior and posterior distributions greater than 0.1 bits). Although there are few cases, especially for the TIGER framework, where only the first two leading PCA parameters have a JS divergence greater than 0.1 bits, we present the posterior distributions of the three most dominant PCA parameters for both FTI and TIGER for broad comparison. The reason why the third-dominant PCA parameter for FTI is more informative than TIGER is because the SNR and number of cycles in the inspiral regime for FTI are larger than TIGER. As mentioned before, the inspiral cutoff frequency for FTI is the peak frequency of the (2, 2) mode, which is larger than the MECO frequency used for TIGER as the inspiral cutoff frequency. The longer inspiral for the FTI analysis compared to TIGER enables the measurement of more PCA parameters and provides overall tighter constraints, if not comparable, as can be seen in Fig. 9. At the same time, the PN phase in Eq. (1) becomes increasingly inaccurate at higher frequencies. Hence, the differences between the FTI and TIGER results can be considered as an estimate of the typical systematics of inspiral tests, with TIGER providing more conservative bounds.

The bounds on the leading PCA parameters for all four events are $|\delta\hat{\phi}_{\text{PCA}}^{(1)}| \leq 0.5$ and are consistent with GR for both TIGER and FTI. Of the four events, GW190814 provides the best constraints, which has the highest inspiral SNR and is the most asymmetric system, leading to the largest number of GW cycles in the inspiral regime. The second-best constraint is provided by GW191204_171526 and GW190412, which have comparable inspiral SNRs, as shown in Table VIII.

The second leading PCA parameters for GW190412, GW191204_171526, and GW191216_213338 are consistent with GR for TIGER and FTI within the 90% credible interval. However, for GW190814, the median of the posterior distribution for TIGER shows a significant offset

⁷Note that in all the injection analyses presented in the previous section, we used the parametrized IMRPhenomX or parametrized IMRPhenomX-MSA-version and the parametrized SEOBNRv4_ROM-version within the TIGER and FTI frameworks, respectively. However, in this section, we have used the newer parametrized IMRPhenomX-SpinTaylor-version and the parametrized SEOBNRv5_ROM-version in the TIGER and FTI frameworks, respectively. These newer versions are slightly improved compared to the previous versions, so we do not expect any changes in the results or conclusions presented in the previous sections.

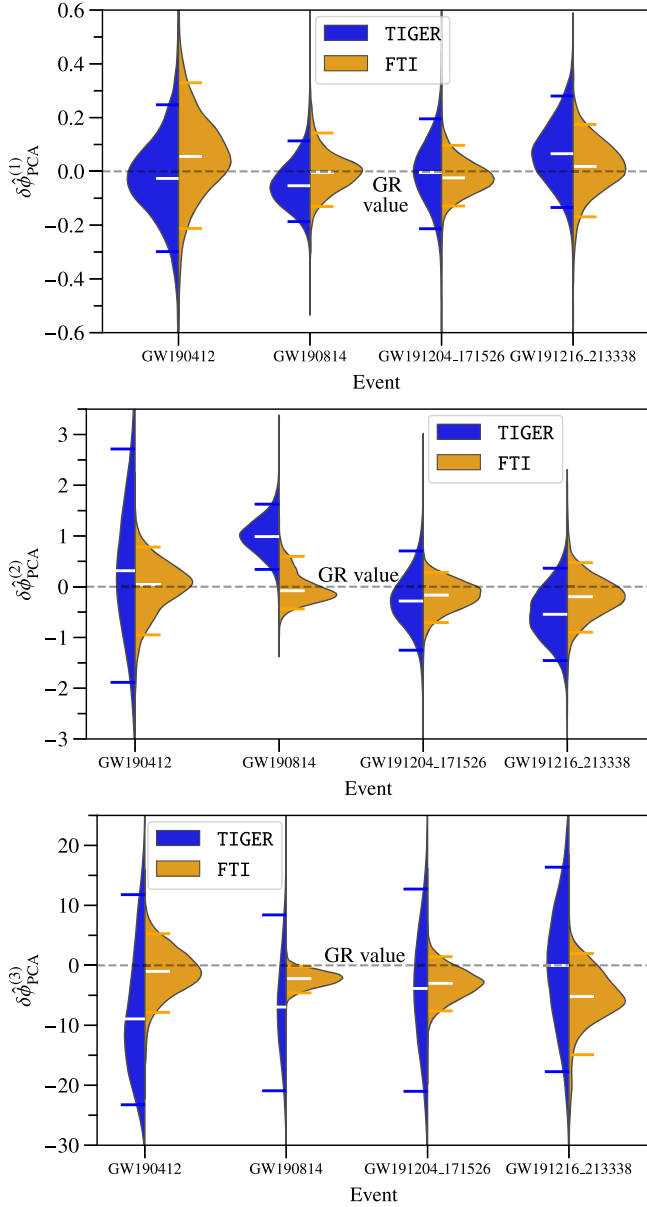


FIG. 9. Posterior probability distribution of the first three leading PCA parameters from the GW events from GWTC-3 passing the selection criteria described in the first paragraph of Sec. VI presented as violin plots. In each violin plot, the colored horizontal bars and the white solid line denote the 90% credible intervals and the posterior median, respectively. We mark the GR value of zero with dashed gray lines.

from the GR value. For TIGER, the GR value lies at 0.01 quantile, indicating a statistically significant violation from the GR value. The posterior distribution of the second leading PCA parameter for FTI is statistically consistent with GR. The composition of the second leading PCA eigenvector shows significant contributions from higher-order PN deformation parameters at 2.5PN, 3PN, and 3.5PN, indicating that these terms play a relatively larger role in shaping the posterior distribution peak away from

TABLE VIII. Rounded-off optimal SNRs in the inspiral regime and the upper cutoff frequencies ($f_{\max}$), which define the end of the inspiral regime, of the four GWTC-3 events selected for the PCA-TGR analysis using the TIGER and FTI frameworks.

Events	TIGER		FTI	
	SNR	$f_{\max}$	SNR	$f_{\max}$
GW190412	17	127	20	257
GW190814	23	159	24	359
GW191204_171526	17	260	18	522
GW191216_213338	16	126	19	322

the GR value. This is consistent with the results obtained by the LVK Collaboration [3], which showed that the GR value falls near the tails of the posterior distribution of the PN deformation parameters.

The maximum likelihood values of all mass parameters obtained from the multiparameter TIGER/FTI runs differ from those obtained through standard parameter estimation using BBH waveforms in GR. In Sec. III A 3, we performed a GW190814-like GR injection under the zero-noise approximation in the TIGER framework, using the IMRPhenomXPHM waveform for injection and the parametrized IMRPhenomXPHM waveform for recovery. The posterior distributions of the first two dominant PCA parameters from this analysis are shown in blue in the rightmost plots of Fig. 3. We see that the GR value is recovered well, which demonstrates that PCA-TGR works robustly in the parameter space spanned by GW190814-like events. This injection study confirms that, at least, correlations between different parameters are not responsible for the offset observed in the PCA posteriors. For example, Ref. [66] discusses a case where a degeneracy between the OPN deviation parameter and the chirp mass led to an apparent GR violation in the posterior of the OPN fractional PN deviation parameter for GW230529 [114]. Therefore, the offsets in the posterior distribution of the second leading PCA parameter away from zero could be due to waveform systematics or noise artifacts in the data.

The posterior distributions of the third PCA parameter for TIGER are not particularly informative for all events except GW190814, which also shows a significant offset from the GR value. For FTI, the third PCA parameter also shows deviation from GR, especially for GW190814, for which the GR value lies at 0.96 quantile. The corresponding eigenvector of the third dominant PCA parameter has a significant contribution from the higher order PN deformation parameters, with the 3.5PN deformation parameter contributing $\sim 50\%$. To understand the underlying reason for the deviations from GR shown by the most dominant PCA parameters, one first needs to examine various sources of possible false indications of GR violation [115], such as the impact of unmodeled physics in the current waveform model or noise artifacts in the data. A detailed investigation

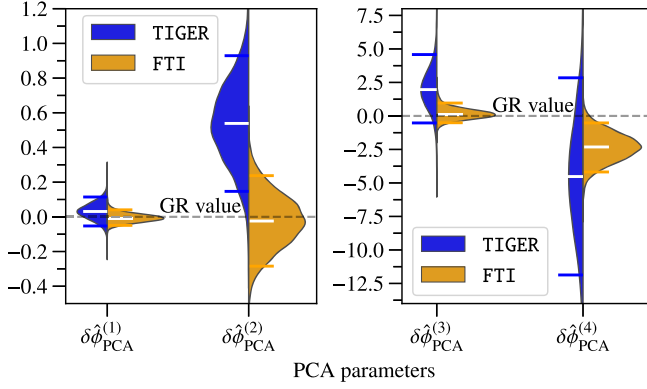


FIG. 10. Joint posterior probability distributions on the first four leading PCA parameters from all the selected GWTC-3 events. The joint bounds are obtained through the marginalized likelihood multiplication technique. We mark the GR value of zero with dashed gray lines.

of such waveform systematics or noise artifacts is beyond the scope of this paper.

A. Joint bounds on the PCA parameters from selected GWTC-3 events

We follow the method described in Sec. IID to combine information from the selected GWTC-3 events, GW190412, GW190814, GW191204_171526, and GW191216_213338 to obtain joint bounds on the PCA parameters. We find that the joint posterior distribution of the first four most dominant PCA parameters, plotted in Fig. 10, have JS divergence greater than 0.1 bits, indicating that they carry significant information. The median and the 90% credible interval for joint posteriors along with the corresponding GR quantiles are shown in Table IX. For the joint posterior of the leading PCA parameter, the GR value lies at the 0.27 quantile in the TIGER framework and the 0.61 quantile in the FTI framework, indicating good consistency with the GR. However, the medians of the joint posterior distributions of the subdominant PCA parameters show a large offset from zero. In the TIGER framework, the posterior

TABLE IX. Combined constraints on the leading four PCA parameters from all selected GWTC-3 events. The table shows the median, 90% credible intervals, $\mathcal{Q}_{\text{GR}}$, and z_{GR} of the combined PCA posteriors.

$\delta\hat{\phi}_{\text{PCA}}^{(i)}$	Median and 90% errors		$\mathcal{Q}_{\text{GR}}$ (z_{GR})	
	TIGER	FTI	TIGER	FTI
$\delta\hat{\phi}_{\text{PCA}}^{(1)}$	$0.03^{+0.08}_{-0.08}$	$-0.01^{+0.05}_{-0.04}$	0.27 (0.60)	0.61 (0.26)
$\delta\hat{\phi}_{\text{PCA}}^{(2)}$	$0.54^{+0.39}_{-0.39}$	$-0.02^{+0.26}_{-0.26}$	0.01 (2.26)	0.56 (0.15)
$\delta\hat{\phi}_{\text{PCA}}^{(3)}$	$1.97^{+2.61}_{-2.49}$	$0.14^{+0.82}_{-0.65}$	0.09 (1.24)	0.36 (0.29)
$\delta\hat{\phi}_{\text{PCA}}^{(4)}$	$-4.52^{+7.36}_{-7.35}$	$-2.32^{+1.80}_{-1.86}$	0.84 (1.01)	0.98 (2.06)

distribution of the second PCA parameter exhibits the highest deviation from the GR value (GR at the 0.01 quantile), followed by the third (GR at the 0.09 quantile) and fourth (GR at 0.84 quantile). The trend is the opposite for FTI, where the highest deviation from GR is shown by the least dominant fourth PCA parameter (GR at the 0.98 quantile), followed by the third (GR at the 0.36 quantile) and second (GR at the 0.56 quantile) PCA parameters. We find the above interesting characteristics because of GW190814,⁸ for which we get a statistically significant deviation from GR for the second PCA parameter in the case of TIGER and a statistically significant deviation from GR for the third PCA parameter for FTI as shown in Fig. 9. The same is reflected in the joint posterior distributions of the subdominant PCA parameters but to a varied degree, as the joint posteriors are obtained after diagonalizing the joint posterior distribution of the original PN deformation parameters and not simply by multiplying the posteriors of a PCA parameter across events.

VII. CONCLUSIONS

It is likely that a deviation from GR will manifest as changes in multiple PN coefficients in the phase evolution of a compact binary. Therefore, a test that measures possible simultaneous departures from GR in these PN coefficients can act as a powerful test of GR. However, the high dimensionality of the parameter space—which includes both the standard GR parameters and fractional deviations in various PN coefficients, along with strong parameter degeneracies—poses a significant challenge to this test. Diagonalization of the covariance matrix that corresponds to the space of deformation parameters helps in identifying the best-measured linear combination of these parameters, and the relative weights of the new set of parameters helps to reduce the dimensionality of the parameter space. This variant of the multiparameter test is referred to as the principal component assisted test of GR or PCA-TGR.

We have done an extensive study of the efficiency of PCA-TGR, using the TIGER and FTI frameworks, that uses Phenom and EOB families of waveforms, respectively, to perform parametrized tests. Besides an extensive zero-noise injection campaign that covers aspects such as spin-precession and higher modes for GR as well as non-GR injections, we also studied how nonzero eccentricity would manifest as a (false) GR violation and how PCA-TGR would respond to this for different values of eccentricities. We then applied this framework to some of the selected events in O3 that have long inspirals and found no statistically significant deviation from GR. We reported

⁸We have verified that the combined posteriors on the PCA parameters, using only three events, excluding GW190814, do not show any significant offset away from GR.

joint bounds by combining the constraints from these events.

Listed below are some of the caveats and future directions.

- (1) It is likely that any true deviation from GR in the inspiral phase will manifest as nonzero (and correlated) values of multiple PN deviation parameters $\delta\hat{\phi}_b$, which is the primary motivation for this PCA analysis. However, it is not possible to predict the pattern of these deviations for any generic beyond-GR model, or even (at present) for a specific beyond-GR theory. Additionally, it is not unlikely that any true deviation from GR in the inspiral phase will manifest as a different pattern of PN deviation parameters for different events (different masses, mass ratio, and spins). Thus, this analysis must be thought of as a test of the null hypothesis (GR is correct). Results that are inconsistent with zero deviation in the best-measured PCA combination (s) (both for the single-event analyses and especially for the multievent “multiplication of likelihoods” analysis) will require careful follow-up with other methods to understand the origin of the deviation. However, in the scenario that a true deviation from GR in the inspiral phase manifests as a fractional (small compared with 1, so approximately linear) deviation of a single PN parameter from zero (which may be the same or different parameters for different events), the PCA analysis has been demonstrated to be essentially as sensitive as the single parameter tests with TIGER and FTI.
- (2) Our results are based on 3.5PN phasing. However, when recently computed terms at 4 and 4.5PN are integrated into the TIGER/FTI frameworks, this method can be straightforwardly extended to include them.
- (3) Currently, the multiparameter test on which PCA is performed simultaneously varies PN fractional deformation coefficients starting from 1.5PN onwards and does not include 0PN, 0.5PN, and 1PN orders. This is because the sensitivity of the current detectors is not good enough to break the degeneracies these PN order deformations have with chirp mass (0PN) and symmetric mass ratio (1PN). However, the algorithm allows this, and hence, if there is an event with high enough SNR and large enough inspiral cycles, such a test can be performed.
- (4) In its current implementation, PCA-TGR deals with only fractional deformation parameters in the inspiral phase. Frameworks such as TIGER allow for simultaneous deformations in the postinspiral phase of the compact binary waveform. This method can be straightforwardly extended to include such deformations, though a fresh injection campaign similar to

the present one will be required to understand the results.

- (5) Our mock non-GR waveforms are based on parametric waveforms and not from any specific theory of modified gravity. It would be interesting to use one of the inspiral-merger-ringdown waveforms from numerical simulations of modified gravity theories (such as Ref. [116]) and test how PCA-TGR responds to such waveforms.

Finally, we note that PCA-TGR is not limited to parametrized PN tests—it can be applied to any parametrized null test of GR to derive constraints from multiparameter analyses. The framework is designed to be general and flexible, making it applicable to a wide range of theory-agnostic, beyond-GR null tests, such as BBH mimicker tests [117,118], parametrized ringdown tests [119–121], and more.

ACKNOWLEDGMENTS

We thank Anuradha Samajdar for critically reading the manuscript and providing useful comments. We are grateful to Nathan K. Johnson-McDaniel for his assistance with the NR eccentricity injection studies. P.M. thanks Alex B. Nielsen and N. V. Krishnendu for their insightful comments on the manuscript. P.M. thanks Duncan MacLeod for his suggestions on various computing issues. P.M. acknowledges the Science and Technology Facilities Council (STFC) for support through Grant No. ST/V005618/1. S.D. acknowledges support from the UVA Arts and Sciences Rising Scholars Fellowship. M.S. acknowledges support from the Weinberg Institute for Theoretical Physics at the University of Texas at Austin. K. G. A. acknowledges support from the Core Research Grant No. CRG/2021/004565 of the Science and Engineering Research Board of India and a grant from the Infosys Foundation. K. G. A. also acknowledges support from the Department of Science and Technology and the Science and Engineering Research Board of India via the Swarnajayanti Fellowship Grant No. DST/SJF/PSA-01/2017-18. P. D. R. acknowledges support from the National Science Foundation (NSF) via NSF Award No. PHY-2409372. S. R. is supported by the Fonds de la Recherche Scientifique—FNRS (Belgium). We also acknowledge NSF support via NSF Awards No. PHY-2308887 to P.N., No. AST-2205920 and No. PHY-2308887 to A.G., No. PHY-2207780 to DS, No. PHY-2309200 and No. PHY-2207758 to A.J.W., and No. AST-2307147, No. PHY-2207638, No. PHY-2308886, and No. PHY-2309064 to B.S.S. The authors are grateful for computational resources provided by the Cardiff University and support by STFC Grants No. ST/I006285/1 and No. ST/V005618/1. The authors are also grateful for computational resources provided by the LIGO Laboratory and supported by National Science Foundation Grants No. PHY-0757058 and No. PHY-0823459. This research has made use of data obtained from the

Gravitational Wave Open Science Center [122], a service of the LIGO Laboratory, the LIGO Scientific Collaboration, and the Virgo Collaboration. LIGO Laboratory and Advanced LIGO are funded by the U.S. NSF as well as the STFC of the United Kingdom, the Max-Planck Society, and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory, by the French Centre National de Recherche Scientifique, the Italian Istituto Nazionale di Fisica Nucleare, and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary,

Ireland, Japan, Monaco, Poland, Portugal, and Spain. The construction and operation of KAGRA are funded by the Ministry of Education, Culture, Sports, Science and Technology, and Japan Society for the Promotion of Science, the National Research Foundation and Ministry of Science and ICT in Korea, and Academia Sinica and the Ministry of Science and Technology in Taiwan.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. They are available from the authors upon reasonable request.

-
- [1] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 221101 (2016); **121**, 129902 (E) (2018).
 - [2] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **100**, 104036 (2019).
 - [3] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **103**, 122002 (2021).
 - [4] R. Abbott *et al.* (LIGO Scientific, VIRGO, and KAGRA Collaborations), [arXiv:2112.06861](https://arxiv.org/abs/2112.06861).
 - [5] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **123**, 011102 (2019).
 - [6] N. Yunes, K. Yagi, and F. Pretorius, *Phys. Rev. D* **94**, 084002 (2016).
 - [7] A. Pai and K. Arun, *Classical Quantum Gravity* **30**, 025011 (2013).
 - [8] K. G. Arun and A. Pai, *Int. J. Mod. Phys. D* **22**, 1341012 (2013).
 - [9] A. A. Shoom, P. K. Gupta, B. Krishnan, A. B. Nielsen, and C. D. Capano, *Gen. Relativ. Gravit.* **55**, 55 (2023).
 - [10] M. Saleem, S. Datta, K. G. Arun, and B. S. Sathyaprakash, *Phys. Rev. D* **105**, 084062 (2022).
 - [11] S. H. Völkel, N. Franchini, and E. Barausse, *Phys. Rev. D* **105**, 084046 (2022).
 - [12] S. Datta, M. Saleem, K. G. Arun, and B. S. Sathyaprakash, *Phys. Rev. D* **109**, 044036 (2024).
 - [13] S. Datta, [arXiv:2303.04399](https://arxiv.org/abs/2303.04399).
 - [14] L. Blanchet and B. S. Sathyaprakash, *Classical Quantum Gravity* **11**, 2807 (1994).
 - [15] L. Blanchet and B. S. Sathyaprakash, *Phys. Rev. Lett.* **74**, 1067 (1995).
 - [16] K. G. Arun, B. R. Iyer, M. S. S. Qusailah, and B. S. Sathyaprakash, *Classical Quantum Gravity* **23**, L37 (2006).
 - [17] K. G. Arun, B. R. Iyer, M. S. S. Qusailah, and B. S. Sathyaprakash, *Phys. Rev. D* **74**, 024006 (2006).
 - [18] N. Yunes and F. Pretorius, *Phys. Rev. D* **80**, 122003 (2009).
 - [19] C. K. Mishra, K. G. Arun, B. R. Iyer, and B. S. Sathyaprakash, *Phys. Rev. D* **82**, 064010 (2010).
 - [20] T. G. F. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos, J. Veitch, K. Grover, T. Sidery, R. Sturani, and A. Vecchio, *Phys. Rev. D* **85**, 082003 (2012).
 - [21] M. Agathos, W. Del Pozzo, T. G. F. Li, C. V. D. Broeck, J. Veitch *et al.*, *Phys. Rev. D* **89**, 082001 (2014).
 - [22] J. Meidam *et al.*, *Phys. Rev. D* **97**, 044033 (2018).
 - [23] A. K. Mehta, A. Buonanno, R. Cotesta, A. Ghosh, N. Sennett, and J. Steinhoff, *Phys. Rev. D* **107**, 044020 (2023).
 - [24] L. Blanchet, *Living Rev. Relativity* **9**, 4 (2006).
 - [25] L. Blanchet, T. Damour, and B. R. Iyer, *Phys. Rev. D* **51**, 5360 (1995).
 - [26] L. Blanchet, G. Faye, Q. Henry, F. Larrouturou, and D. Trestini, *Phys. Rev. Lett.* **131**, 121402 (2023).
 - [27] L. Blanchet and T. Damour, *Phys. Rev. D* **37**, 1410 (1988).
 - [28] L. Blanchet and G. Schäfer, *Classical Quantum Gravity* **10**, 2699 (1993).
 - [29] L. Blanchet, *Classical Quantum Gravity* **15**, 113 (1998).
 - [30] L. Blanchet, *Classical Quantum Gravity* **15**, 89 (1998).
 - [31] D. Christodoulou, *Phys. Rev. Lett.* **67**, 1486 (1991).
 - [32] K. Thorne, *Phys. Rev. D* **45**, 520 (1992).
 - [33] K. G. Arun, L. Blanchet, B. R. Iyer, and M. S. S. Qusailah, *Classical Quantum Gravity* **21**, 3771 (2004); **22**, 3115(E) (2005).
 - [34] M. Favata, *Phys. Rev. D* **80**, 024002 (2009).
 - [35] L. Kidder, C. Will, and A. Wiseman, *Phys. Rev. D* **47**, R4183 (1993).
 - [36] L. Blanchet, A. Buonanno, and G. Faye, *Phys. Rev. D* **74**, 104034 (2006); **75**, 049903(E) (2007); **81**, 089901(E) (2010).
 - [37] A. Buonanno, G. Faye, and T. Hinderer, *Phys. Rev. D* **87**, 044009 (2013).
 - [38] L. Bernard, L. Blanchet, and D. Trestini, *J. Cosmol. Astropart. Phys.* **08** (2022) 008.
 - [39] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 061102 (2016).

- [40] A. Gupta, S. Datta, S. Kastha, S. Borhanian, K. G. Arun, and B. S. Sathyaprakash, *Phys. Rev. Lett.* **125**, 201101 (2020).
- [41] S. Datta, A. Gupta, S. Kastha, K. G. Arun, and B. S. Sathyaprakash, *Phys. Rev. D* **103**, 024036 (2021).
- [42] P. Mahapatra, S. Kastha, A. Gupta, B. S. Sathyaprakash, and K. G. Arun, *Phys. Rev. D* **109**, 064036 (2024).
- [43] S. Kastha, A. Gupta, K. G. Arun, B. S. Sathyaprakash, and C. Van Den Broeck, *Phys. Rev. D* **98**, 124033 (2018).
- [44] S. Kastha, A. Gupta, K. G. Arun, B. S. Sathyaprakash, and C. Van Den Broeck, *Phys. Rev. D* **100**, 044007 (2019).
- [45] P. Mahapatra, *Phys. Rev. D* **109**, 024050 (2024).
- [46] P. Mahapatra and S. Kastha, *Phys. Rev. D* **109**, 084069 (2024).
- [47] L. Blanchet, T. Damour, B. R. Iyer, C. M. Will, and A. G. Wiseman, *Phys. Rev. Lett.* **74**, 3515 (1995).
- [48] L. Blanchet, T. Damour, G. Esposito-Farèse, and B. R. Iyer, *Phys. Rev. Lett.* **93**, 091101 (2004).
- [49] K. G. Arun, B. R. Iyer, B. S. Sathyaprakash, and P. A. Sundararajan, *Phys. Rev. D* **71**, 084008 (2005); **72**, 069903 (E) (2005).
- [50] B. Abbott *et al.*, *Phys. Rev. Lett.* **119**, 161101 (2017).
- [51] A. Buonanno, B. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, *Phys. Rev. D* **80**, 084043 (2009).
- [52] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **9**, 031040 (2019).
- [53] R. Abbott *et al.* (KAGRA, VIRGO, and LIGO Scientific Collaborations), *Phys. Rev. X* **13**, 041039 (2023).
- [54] B. S. Sathyaprakash and S. V. Dhurandhar, *Phys. Rev. D* **44**, 3819 (1991).
- [55] T. Damour, B. R. Iyer, and B. S. Sathyaprakash, *Phys. Rev. D* **62**, 084036 (2000).
- [56] A. Buonanno, B. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, *Phys. Rev. D* **80**, 084043 (2009).
- [57] K. G. Arun, A. Buonanno, G. Faye, and E. Ochsner, *Phys. Rev. D* **79**, 104023 (2009).
- [58] C. K. Mishra, A. Kela, K. G. Arun, and G. Faye, *Phys. Rev. D* **93**, 084054 (2016).
- [59] S. Endlich, V. Gorbenko, J. Huang, and L. Senatore, *J. High Energy Phys.* **09** (2017) 122.
- [60] N. Sennett, R. Brito, A. Buonanno, V. Gorbenko, and L. Senatore, *Phys. Rev. D* **102**, 044056 (2020).
- [61] M. Agathos, W. Del Pozzo, T. G. Li, C. Van Den Broeck, J. Veitch, and S. Vitale, *Phys. Rev. D* **89**, 082001 (2014).
- [62] S. Roy, M. Haney, G. Pratten, P. T. H. Pang, and C. Van Den Broeck, *arXiv:2504.21147*.
- [63] G. Pratten *et al.*, *Phys. Rev. D* **103**, 104056 (2021).
- [64] C. García-Quirós, M. Colleoni, S. Husa, H. Estellés, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena, and R. Jaume, *Phys. Rev. D* **102**, 064002 (2020).
- [65] M. Colleoni, F. A. R. Vidal, C. García-Quirós, S. Akçay, and S. Bera, *Phys. Rev. D* **111**, 104019 (2025).
- [66] E. M. Slinger *et al.*, *arXiv:2406.03568*.
- [67] A. Bohé *et al.*, *Phys. Rev. D* **95**, 044028 (2017).
- [68] R. Cotesta, S. Marsat, and M. Pürrer, *Phys. Rev. D* **101**, 124040 (2020).
- [69] R. Cotesta, A. Buonanno, A. Bohé, A. Taracchini, I. Hinder, and S. Ossokine, *Phys. Rev. D* **98**, 084028 (2018).
- [70] T. Li, W. Del Pozzo, S. Vitale, C. Van Den Broeck, M. Agathos *et al.*, *Phys. Rev. D* **85**, 082003 (2012).
- [71] L. Sampson, N. Cornish, and N. Yunes, *Phys. Rev. D* **87**, 102001 (2013).
- [72] Z. Lyu, N. Jiang, and K. Yagi, *Phys. Rev. D* **105**, 064001 (2022); **106**, 069901(E) (2022).
- [73] M. Saleem *et al.*, *Classical Quantum Gravity* **39**, 025004 (2022).
- [74] G. Ashton *et al.*, Bilby tgr (2024), 10.5281/zenodo.10940210.
- [75] G. Ashton *et al.*, *Astrophys. J. Suppl. Ser.* **241**, 27 (2019).
- [76] I. M. Romero-Shaw *et al.*, *Mon. Not. R. Astron. Soc.* **499**, 3295 (2020).
- [77] J. S. Speagle, *Mon. Not. R. Astron. Soc.* **493**, 3132 (2020).
- [78] J. Skilling, *Bayesian Anal.* **1**, 833 (2006).
- [79] LIGO Scientific Collaboration and Virgo Collaboration, GWTC-3 data release (2021), 10.7935/b024-1886.
- [80] C. R. Harris *et al.*, *Nature (London)* **585**, 357 (2020).
- [81] J. Lin, *IEEE Trans. Inf. Theory* **37**, 145 (1991).
- [82] C. Rao, *Bull. Calcutta Math. Soc.* **37**, 81 (1945).
- [83] H. Cramer, *Mathematical Methods in Statistics* (Pergamon Press, Princeton University Press, NJ, USA, 1946).
- [84] C. Cutler and E. Flanagan, *Phys. Rev. D* **49**, 2658 (1994).
- [85] E. Poisson and C. Will, *Phys. Rev. D* **52**, 848 (1995).
- [86] N. Yunes and S. A. Hughes, *Phys. Rev. D* **82**, 082002 (2010).
- [87] M. Kramer *et al.*, *Phys. Rev. X* **11**, 041050 (2021).
- [88] W. Del Pozzo, J. Veitch, and A. Vecchio, *Phys. Rev. D* **83**, 082002 (2011).
- [89] A. Zimmerman, C.-J. Haster, and K. Chatziioannou, *Phys. Rev. D* **99**, 124044 (2019).
- [90] M. Isi, K. Chatziioannou, and W. M. Farr, *Phys. Rev. Lett.* **123**, 121101 (2019).
- [91] F. Pedregosa *et al.*, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [92] F. Acernese, M. Agathos *et al.*, *Classical Quantum Gravity* **32**, 024001 (2014).
- [93] J. Aasi *et al.*, *Classical Quantum Gravity* **32**, 074001 (2015).
- [94] B. P. Abbott *et al.* (KAGRA, LIGO Scientific, and Virgo Collaborations), *Living Rev. Relativity* **19**, 1 (2016).
- [95] LIGO Scientific, VIRGO, and KAGRA Collaborations, <https://dcc.ligo.org/LIGO-T2000012/public> (2022).
- [96] LIGO Scientific, VIRGO, and KAGRA Collaborations, <https://dcc.ligo.org/LIGO-P1200087-v42/public> (2022).
- [97] G. Pratten, S. Husa, C. Garcia-Quiros, M. Colleoni, A. Ramos-Buades, H. Estelles, and R. Jaume, *Phys. Rev. D* **102**, 064001 (2020).
- [98] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, *Phys. Rev. Lett.* **113**, 151101 (2014).
- [99] P. Schmidt, F. Ohme, and M. Hannam, *Phys. Rev. D* **91**, 024043 (2015).
- [100] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **116**, 241103 (2016).
- [101] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **102**, 043015 (2020).
- [102] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Astrophys. J. Lett.* **896**, L44 (2020).
- [103] C. Pankow *et al.*, *Phys. Rev. D* **98**, 084016 (2018).
- [104] N. J. Cornish, T. B. Littenberg, B. Bécsy, K. Chatziioannou, J. A. Clark, S. Ghonge, and M. Millhouse, *Phys. Rev. D* **103**, 044006 (2021).

- [105] B. P. Abbott *et al.* (Virgo and LIGO Scientific Collaborations), *Phys. Rev. X* **6**, 041014 (2016).
- [106] L. Pompili *et al.*, *Phys. Rev. D* **108**, 124035 (2023).
- [107] P. Saini, M. Favata, and K. G. Arun, *Phys. Rev. D* **106**, 084031 (2022).
- [108] P. Narayan, N. K. Johnson-McDaniel, and A. Gupta, *Phys. Rev. D* **108**, 064003 (2023).
- [109] M. Boyle *et al.*, *Classical Quantum Gravity* **36**, 195006 (2019).
- [110] I. Hinder, L. E. Kidder, and H. P. Pfeiffer, *Phys. Rev. D* **98**, 044015 (2018).
- [111] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **6**, 041015 (2016); **8**, 039903(E) (2018).
- [112] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. X* **11**, 021053 (2021).
- [113] L. S. Collaboration, V. Collaboration, and K. Collaboration, GWTC-3: Compact binary coalescences observed by LIGO and Virgo during the second part of the third observing run—parameter estimation data release (2023), [10.5281/zenodo.8177023](https://zenodo.org/record/8177023).
- [114] A. G. Abac *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), *Astrophys. J. Lett.* **970**, L34 (2024).
- [115] A. Gupta *et al.*, [10.21468/SciPostPhysCommRep.5](https://arxiv.org/abs/10.21468/SciPostPhysCommRep.5) (2024).
- [116] M. Okounkova, L. C. Stein, J. Moxon, M. A. Scheel, and S. A. Teukolsky, *Phys. Rev. D* **101**, 104016 (2020).
- [117] N. V. Krishnendu, K. G. Arun, and C. K. Mishra, *Phys. Rev. Lett.* **119**, 091101 (2017).
- [118] N. K. Johnson-McDaniel, A. Mukherjee, R. Kashyap, P. Ajith, W. Del Pozzo, and S. Vitale, *Phys. Rev. D* **102**, 123010 (2020).
- [119] G. Carullo, W. Del Pozzo, and J. Veitch, *Phys. Rev. D* **99**, 123029 (2019); **100**, 089903(E) (2019).
- [120] E. Maggio, H. O. Silva, A. Buonanno, and A. Ghosh, *Phys. Rev. D* **108**, 024043 (2023).
- [121] L. Pompili, E. Maggio, H. O. Silva, and A. Buonanno, *Phys. Rev. D* **111**, 124040 (2025).
- [122] <https://www.gwosc.org/>.