

Mass-spin reparametrization for a rapid parameter estimation of inspiral gravitational-wave signals

Eunsub Lee¹, Soichiro Morisaki², and Hideyuki Tagoshi¹

¹*Institute for Cosmic Ray Research, The University of Tokyo,
5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8582, Japan*

²*Department of Physics, University of Wisconsin-Milwaukee,
Milwaukee, Wisconsin 53201, USA*



(Received 12 March 2022; accepted 26 May 2022; published 23 June 2022)

Estimating the source parameters of gravitational waves from compact binary coalescence (CBC) is a key analysis task in gravitational-wave astronomy. To deal with the increasing detection rate of CBC signals, optimizing the parameter estimation analysis is crucial. The analysis typically employs a stochastic sampling technique such as Markov Chain Monte Carlo (MCMC), where the source parameter space is explored and regions of high-Bayesian posterior probability density are found. One of the bottlenecks slowing down the analysis is the nontrivial correlation between masses and spins of colliding objects, which makes the exploration of mass-spin space extremely inefficient. We introduce a new set of mass-spin sampling parameters which makes the posterior distribution simpler in the new parameter space, regardless of the true values of the parameters. The new parameter combinations are obtained as the principal components of the Fisher matrix for the restricted 1.5 post-Newtonian waveform. Our reparametrization improves the efficiency of MCMC by a factor of ~ 10 for a binary neutron star with a narrow-spin prior ($|\vec{\chi}| < 0.05$) and ~ 100 for a binary neutron star with a broad-spin prior ($|\vec{\chi}| < 0.99$), under the assumption that the binary has spins aligned with its orbital angular momentum.

DOI: [10.1103/PhysRevD.105.124057](https://doi.org/10.1103/PhysRevD.105.124057)

I. INTRODUCTION

In September 2015 the first direct detection of gravitational waves (GWs), which was radiated from the coalescence of a binary black hole (BBH), took place [1]. Since the first detection, 90 gravitational-wave signals from compact binary coalescence (CBC) have been reported by the LIGO-Virgo-KAGRA Collaboration [2–5], including the coalescence of a binary neutron star (BNS) with electromagnetic (EM) follow-up observations [6].

Estimating source parameters from a CBC signal is an important task in gravitational-wave astronomy. The estimated source location is crucial for the EM follow-up observations, and the masses and spins of colliding objects are important for studying the formation history of compact binaries [7,8].

This parameter estimation analysis typically employs Bayesian inference using stochastic sampling techniques, such as Markov Chain Monte Carlo (MCMC) [9,10] and nested sampling [11]. While the stochastic sampling is known to be efficient for estimating high-dimensional parameters, it is still computationally costly, taking more than weeks for a BNS event without any approximate methods. Speeding it up is necessary to deal with the increasing detection rate of CBC signals. It is also crucial

for the follow-up observations of EM counterparts rapidly fading after the merger.

Efficient exploration of the parameter space is essential for optimizing the stochastic sampling. One-dimensional jumps can efficiently explore the parameter space if parameters are not strongly correlated. However, they are extremely inefficient if parameters are strongly correlated, which is the case for a CBC signal. One solution for this is to use parametrizations minimizing the correlations between parameters. Based on this idea, current analysis pipelines, such as LALInference [12], PyCBC Inference [13], and BILBY [14], use chirp mass $\mathcal{M} \equiv (m_1 m_2)^{3/5} (m_1 + m_2)^{-1/5}$ and mass ratio $q \equiv m_2/m_1$, to reduce the correlation between component masses, m_1 and m_2 , where $m_1 \geq m_2$. However, the masses are correlated with spins [15], and this choice of parameters does not fully minimize their correlations.

In this paper, we introduce an alternative set of mass-spin parameters which has significantly reduced correlations between parameters. The new parameters, referred to as μ_1 and μ_2 , have been studied in literature for efficient template placement [16–18] and to find the best measurable combinations of physical parameters [19–21]. In contrast to them, we use these parameters in the MCMC sampling process.

We show that using the new parametrizations significantly speeds up the parameter estimation, without any loss of accuracy of the estimation.

This paper is organized as follows. In Sec. II we introduce the basics of parameter estimation and the correlation between mass and spin parameters. In Sec. III we introduce a new set of mass-spin parameters, which has significantly reduced correlations between parameters, and explain a sampling method using them. In Sec. IV we describe injection tests for studying the performance of our new sampling strategy and introduce their results. Finally we present our conclusion in Sec. V.

II. BACKGROUND

In this section we explain the basics of Bayesian parameter estimation using MCMC. Then, we introduce the correlation between mass and spin parameters, which makes the parameter estimation analysis of a CBC signal significantly inefficient.

A. Parameter estimation using MCMC

In the Bayesian inference, the inference result is the posterior distribution $p(\theta|d)$ which is the probability distribution of the parameters θ when the observation data d is given. The posterior distribution is calculated by the Bayes' theorem,

$$p(\theta|d) \propto p(d|\theta)p(\theta), \quad (1)$$

where $p(d|\theta)$ is the likelihood at θ and $p(\theta)$ is the prior distribution. We assume that the data of a detector is modeled as the sum of Gaussian, stationary random noise n and a gravitational-wave signal h ,

$$d(t) = n(t) + h(t; \theta). \quad (2)$$

Then the likelihood for a single detector becomes

$$p(d|\theta) \propto \exp \left[-\frac{1}{2} (d - h(\theta), d - h(\theta)) \right]. \quad (3)$$

Here (x, y) indicates the noise-weighted inner product,

$$(x, y) \equiv 4\Re \left[\int_{f_{\min}}^{f_{\max}} df \frac{\tilde{x}^*(f)\tilde{y}(f)}{S_n(f)} \right], \quad (4)$$

where f_{\min} and f_{\max} are the low- and high-frequency cutoffs of the analysis respectively, and $S_n(f)$ is the one-sided power spectral density (PSD) of the detector. For multiple detectors, the likelihood of the combined data is the product of that of each detector, assuming the noise at each detector is statistically independent.

Even though the posterior distribution has a simple form of Eq. (1), it is not easy to get the properties of this

distribution. A practical method is to generate samples that follow the distribution. MCMC is an efficient method for the sampling from the posterior distribution. A new sample is drawn stochastically based on the current sample. The most basic MCMC is the Metropolis-Hastings (MH) algorithm [10]. In the MH algorithm, drawing a sample is divided into two steps; proposal and acceptance-rejection. Selecting appropriate proposal distribution is critical to the performance of MH algorithm. However, if the posterior distribution has a complicated structure owing to the correlation of parameters, it is a hard task to find efficient proposal distribution in advance.

B. Correlation between masses and spins

The phase evolution of gravitational waves is predominantly characterized by the masses and spins of colliding objects, especially, the leading term in the post-Newtonian expansion solely depends on chirp mass \mathcal{M} . Thus, the chirp mass is precisely determined from gravitational waves, and the contour of the mass distribution approximately follows the fixed line of \mathcal{M} . This leads to the strong correlation between m_1 and m_2 , which makes the stochastic sampling in the m_1 - m_2 coordinate system inefficient. As explained in Sec. I, the `LALInference` software uses \mathcal{M} and mass ratio q as independent sampling parameters to solve this issue.

The dominant spin contribution appears at the 1.5 post-Newtonian order through the following combination,

$$\beta = \frac{1}{12} \sum_{i=1}^2 \left[113 \left(\frac{m_i}{m_1 + m_2} \right)^2 + 75\eta \right] \chi_i, \quad (5)$$

where η is symmetric mass ratio,

$$\eta = \frac{m_1 m_2}{(m_1 + m_2)^2}, \quad (6)$$

and χ_i is the component of the spin angular momentum \vec{S}_i along the orbital angular momentum \vec{L}_i normalized by m_i^2 ,

$$\chi_i \equiv \frac{\mathbf{L}}{|\mathbf{L}|} \cdot \frac{c\mathbf{S}_i}{Gm_i^2}. \quad (7)$$

Since it also affects the frequency evolution, it is correlated with \mathcal{M} and q . Figure 1 shows the posterior samples for a binary neutron star signal in the \mathcal{M} - q - β space. It shows that the mass and spin parameters are correlated non-trivially. Even \mathcal{M} and q are strongly correlated with each other again in this case. The samples are along a characteristic curve; thus, the mass-spin space needs to be explored along the curve, which makes the sampling difficult and inefficient.

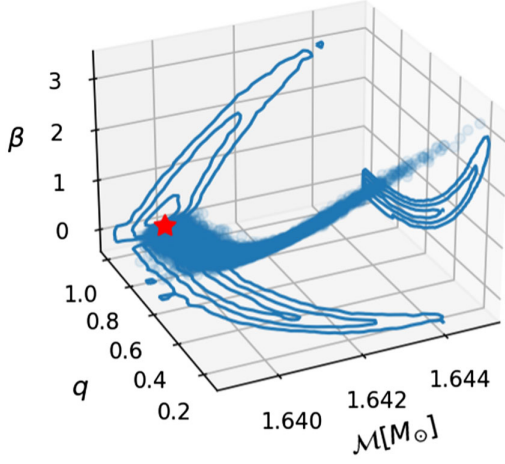


FIG. 1. Posterior samples from a simulated signal, which is presented as #S2 in Sec. IV. Two-dimensional plots are marginal distributions and contours represent the 1σ , 2σ , and 3σ regions. The red star marks the true values of the parameters.

III. METHODOLOGY

In this section we construct an alternative set of mass-spin sampling parameters, which has a significantly reduced correlation between the parameters. Following [19,21], we construct them as the principal components of the Fisher matrix for the restricted post-Newtonian waveform. We also discuss the practical choice of sampling parameters, sampling method and its extension to the case of multiple detectors.

A. Restricted post-Newtonian waveform

To construct the efficient sampling parameters, we make use of a restricted post-Newtonian waveform model. Here we take into account terms up to the 1.5 post-Newtonian order to incorporate the dominant spin contribution. The waveform is given by

$$\tilde{h}(f) = \mathcal{A} \left(\frac{f}{f_{\text{ref}}} \right)^{-\frac{7}{6}} e^{-i\Psi(f)}, \quad (8)$$

with the phase function

$$\begin{aligned} \Psi(f) = & \psi_1 \left(\frac{f}{f_{\text{ref}}} \right)^{-\frac{5}{3}} + \psi_2 \left(\frac{f}{f_{\text{ref}}} \right)^{-1} + \psi_3 \left(\frac{f}{f_{\text{ref}}} \right)^{-\frac{2}{3}} \\ & + \psi_4 + \psi_5 \left(\frac{f}{f_{\text{ref}}} \right). \end{aligned} \quad (9)$$

The phase expansion coefficients are related to the physical parameters as

$$\psi_1 = \frac{3}{128} (\pi G \mathcal{M} f_{\text{ref}} / c^3)^{-\frac{5}{3}}, \quad (10a)$$

$$\psi_2 = \frac{55}{384} \left(\eta + \frac{743}{924} \right) \eta^{-\frac{2}{3}} (\pi G \mathcal{M} f_{\text{ref}} / c^3)^{-1}, \quad (10b)$$

$$\psi_3 = \frac{3}{32} (\beta - 4\pi) \eta^{-\frac{2}{3}} (\pi G \mathcal{M} f_{\text{ref}} / c^3)^{-\frac{2}{3}}, \quad (10c)$$

$$\psi_4 = -2\phi_c - \frac{\pi}{4}, \quad (10d)$$

$$\psi_5 = 2\pi f_{\text{ref}} t_c, \quad (10e)$$

where ϕ_c is merger phase, t_c is merger time, and β is the combination of spins as defined in (5). f_{ref} is a reference frequency introduced to make phase expansion coefficients dimensionless, and we use $f_{\text{ref}} = 200$ Hz following [19,21]. The amplitude \mathcal{A} in Eq. (8) is a function of the chirp mass and the extrinsic parameters, and the signal-to-noise ratio (SNR) $\rho \equiv (h, h)^{-1/2}$ is proportional to \mathcal{A} .

B. Principal component analysis

We use the restricted post-Newtonian waveform to study the approximate structure of the posterior distribution in the mass-spin space. By substituting the restricted 1.5P post-Newtonian waveform into the likelihood for a single detector (3), we obtain

$$\begin{aligned} p(d|\psi) \propto \exp \left[-\frac{1}{2} (h(\hat{\psi}) - h(\psi), h(\hat{\psi}) - h(\psi)) \right. \\ \left. - (n, h(\hat{\psi}) - h(\psi)) \right], \end{aligned} \quad (11)$$

where we use the phase coefficients $\psi = \{\psi_i\}_{i=1}^5$ instead of physical parameters to parameterize the waveform, and $\hat{\psi}_i$ is their true values. In the limit of a high signal-to-noise ratio, the posterior distribution has a sharp peak around $\hat{\psi}$. We expand $h(\psi)$ as a Taylor series around $h(\hat{\psi})$ and approximate it to leading order. Then Eq. (11) becomes Gaussian,

$$p(d|\psi) \propto \exp \left[-\frac{1}{2} \sum_{i,j} \Gamma_{ij} \Delta\psi_i \Delta\psi_j \right], \quad (12)$$

where

$$\Gamma_{ij} \equiv \left(\frac{\partial h}{\partial \psi_i}, \frac{\partial h}{\partial \psi_j} \right) \quad (13)$$

is the Fisher information matrix (FIM) for ψ , and

$$\Delta\psi_i \equiv \psi_i - \hat{\psi}_i - \sum_k (\Gamma^{-1})^{ik} \left(n, \frac{\partial h}{\partial \psi_k} \right). \quad (14)$$

We construct combinations which depend on $(\mathcal{M}, q, \chi_1, \chi_2)$ but not t_c and ϕ_c , in consideration of the easy availability. We thus consider the posterior marginalized over ϕ_c and t_c , or equivalently ψ_4 and ψ_5 . The marginal posterior is then

$$p(\{\psi_1, \psi_2, \psi_3\}|d) \propto p(\{\psi_1, \psi_2, \psi_3\}) \times \exp \left[-\frac{1}{2} \sum_{i,j} \tilde{\Gamma}_{ij} \Delta\psi_i \Delta\psi_j \right], \quad (15)$$

where $p(\{\psi_1, \psi_2, \psi_3\})$ is the prior distribution of (ψ_1, ψ_2, ψ_3) , $\tilde{\Gamma}$ is the 3×3 FIM,

$$\tilde{\Gamma}_{ij} = \Gamma_{ij} - \sum_{k,l} \Gamma_{ik} \gamma^{kl} \Gamma_{lj}, \quad (16)$$

$i, j = 1, 2, 3, k, l = 4, 5$ and

$$\gamma = \begin{bmatrix} \Gamma_{44} & \Gamma_{45} \\ \Gamma_{54} & \Gamma_{55} \end{bmatrix}^{-1}. \quad (17)$$

We can diagonalize $\tilde{\Gamma}$ using an orthogonal matrix U as

$$\tilde{\Gamma}_{ij} = \sum_{m,n} U_{im}^T \lambda_m \delta_{mn} U_{nj}, \quad (18)$$

where $\{\lambda_m\}_{m=1}^3$ are the eigenvalues in descending order $\lambda_1 > \lambda_2 > \lambda_3$. Using this, Eq. (15) takes a simple form,

$$p(\mu|d) \propto p(\mu) \prod_n e^{-\frac{1}{2} \lambda_n \Delta\mu_n^2}, \quad (19)$$

with a new set of parameters

$$\mu_n \equiv \sum_i U_{ni} \psi_i, \quad (n = 1, 2, 3). \quad (20)$$

Equation (19) can be represented as product of each parameter's function if the prior distribution $p(\mu)$ is separable. This implies that if the parameters are not strongly correlated in the prior, the posterior distribution in the μ space becomes very simple. Thus, we use μ_n as an alternative sampling parameter.

Before discussing the sampling method using μ_n , we discuss the dependence of μ_n on masses and spins. The original FIM is given as

$$\Gamma = 4|\mathcal{A}|^2 \begin{bmatrix} I_{-17/3} & I_{-5} & I_{-14/3} & I_{-4} & I_{-3} \\ I_{-5} & I_{-13/3} & I_{-4} & I_{-10/3} & I_{-7/3} \\ I_{-14/3} & I_{-4} & I_{-11/3} & I_{-3} & I_{-2} \\ I_{-4} & I_{-10/3} & I_{-3} & I_{-7/3} & I_{-4/3} \\ I_{-3} & I_{-7/3} & I_{-2} & I_{-4/3} & I_{-1/3} \end{bmatrix}, \quad (21)$$

with

$$I_a \equiv \int_{f_{\min}}^{f_{\max}} \left(\frac{f}{f_{\text{ref}}} \right)^a \frac{df}{S_n(f)}. \quad (22)$$

f_{\max} is usually set to the innermost stable circular orbit frequency,

$$f_{\text{isco}} = \frac{1}{6^{3/2} \pi G (m_1 + m_2) / c^3}, \quad (23)$$

for inspiral-only waveform models. Thus, Γ , and hence μ_n , generally depends on masses. However, in this work, we fix f_{\max} to 2048 Hz instead. Then U does not depend on the physical parameters, and depends only on f_{\min} and the power spectral density of the detector. For typical BNS events, $f_{\text{isco}} \simeq 4400/(M/M_\odot)$ Hz is higher than 1000 Hz for BNS events. Since $S_n(f)$ gets larger in $f > 1000$ Hz for the current ground-based detectors, and $(f/f_{\text{ref}})^\alpha$ simply decreases for $\alpha < 0$, the contributions to I_a for 2048 Hz $< f < f_{\text{isco}}$ or $f_{\text{isco}} < f < 2048$ Hz are suppressed. Thus, our choice of f_{\max} is a reasonable approximation for BNS events. Actually, even if we change f_{\max} to 1000 Hz, the coefficients of μ_1 and μ_2 are changed up to 5.6% compared to when $f_{\max} = 2048$. In Sec. IV we show that μ_n obtained under this choice of f_{\max} is useful even for BBH events.

C. Practical choice of sampling parameters and sampling strategy

Next, we discuss our choice of sampling parameters including μ_n . In this paper, we restrict ourself to a system whose spins are aligned with the orbital angular momentum and consider only four mass-spin parameters $(m_1, m_2, \chi_1, \chi_2)$. We will discuss a potential extension of our method to the case of full-spin components in Sec. V.

As a surrogate set of mass-spin parameters, we choose $(\mu_1, \mu_2, q, \chi_2)$. The prior distribution in $(\mu_1, \mu_2, q, \chi_2)$ space can be represented as

$$p(\mu_1, \mu_2, q, \chi_2) = \frac{p(m_1, m_2, \chi_1, \chi_2)}{|J|}. \quad (24)$$

$|J|$ is the Jacobian determinant, which can be calculated as

$$|J| = \left| (U_{11}U_{23} - U_{21}U_{13}) \frac{\partial\psi_1}{\partial\mathcal{M}} \frac{\partial\psi_3}{\partial\chi_1} + (U_{12}U_{23} - U_{22}U_{13}) \frac{\partial\psi_2}{\partial\mathcal{M}} \frac{\partial\psi_3}{\partial\chi_1} \right| \frac{|\mathcal{M}|}{m_1^2}, \quad (25)$$

where

$$\frac{\partial\psi_1}{\partial\mathcal{M}} = -\frac{5}{128} \frac{\pi G f_{\text{ref}}}{c^3} (\pi G M f_{\text{ref}}/c^3)^{-\frac{8}{3}}, \quad (26)$$

$$\frac{\partial\psi_2}{\partial\mathcal{M}} = -\frac{55}{384} \left(\eta + \frac{743}{924} \right) \eta^{-\frac{2}{3}} \frac{\pi G f_{\text{ref}}}{c^3} (\pi G M f_{\text{ref}}/c^3)^{-2}, \quad (27)$$

$$\frac{\partial\psi_3}{\partial\chi_1} = \frac{113 + 75q}{12(1+q)^2} \frac{3}{32} \eta^{-\frac{3}{3}} (\pi G M f_{\text{ref}}/c^3)^{-\frac{2}{3}}. \quad (28)$$

To include q is necessary to avoid the singularity of the Jacobian at $m_1 = m_2$, which is an obstacle for the use of η as a sampling parameter [12]. A natural candidate for the sampling parameters including q might be (μ_1, μ_2, μ_3, q) . However, in that case the Jacobian determinant becomes 0, since (μ_1, μ_2, μ_3) depend on χ_1 and χ_2 only via ψ_3 .

We can generate posterior samples efficiently in $(\mu_1, \mu_2, q, \chi_2)$. To obtain the posterior distribution for physical parameters, we need a transformation from $(\mu_1, \mu_2, q, \chi_2)$ to $(m_1, m_2, \chi_1, \chi_2)$. m_1 and m_2 are calculated as follows: First, $x \equiv \mu_1 - (U_{13}/U_{23})\mu_2$, is calculated. Then \mathcal{M} is calculated as a solution of the following equation,

$$\left(U_{11} - \frac{U_{13}}{U_{23}} U_{21} \right) \psi_1 + \left(U_{12} - \frac{U_{13}}{U_{23}} U_{22} \right) \psi_2 = x. \quad (29)$$

Since the left-hand side of (29) is a decreasing function of \mathcal{M} , we can use a simple bisectional search to find the solution. m_1 and m_2 are easily calculated from \mathcal{M} and q . Finally, given $(\mu_2, m_1, m_2, \chi_2)$, χ_1 is easily calculated.

TABLE I. Test cases.

| Spin prior range | Case | $\mathcal{M}[M_\odot]$ | q | χ_1 | χ_2 |
|-------------------------------|------|------------------------|------|----------|----------|
| <i>Single detector cases</i> | | | | | |
| Narrow | #S1 | 1.64 | 1.0 | 0.02 | 0.02 |
| $ \chi_1 , \chi_2 < 0.05$ | | | | | |
| Semibroad | #S2 | 1.64 | 1.0 | 0 | 0 |
| $ \chi_1 , \chi_2 < 0.4$ | #S3 | 1.64 | 0.7 | 0 | 0 |
| | #S4 | 1.64 | 1.0 | 0.3 | 0.3 |
| | #S5 | 1.64 | 0.7 | 0.3 | 0.3 |
| Broad | #S6 | 1.64 | 0.7 | 0.5 | 0.5 |
| $ \chi_1 , \chi_2 < 0.99$ | #S7 | 28.2 | 0.84 | 0.26 | 0.32 |
| <i>Multiple detector case</i> | | | | | |
| $ \chi_1 , \chi_2 < 0.4$ | #M | 1.64 | 0.9 | 0 | 0 |

TABLE II. Prior range of μ_1 and μ_2 parameters used in each case.

| Case | Minimum | Parameter | Maximum |
|---------|-------------|-----------|------------|
| #S1 | 148.615862 | μ_1 | 156.940475 |
| | -76.106636 | μ_2 | -41.460751 |
| #S2-#S5 | 145.212081 | μ_1 | 158.693980 |
| | -97.311606 | μ_2 | -30.738354 |
| #S6 | 139.474278 | μ_1 | 162.409836 |
| | -133.057127 | μ_2 | -12.663454 |
| #S7 | -1.520207 | μ_1 | 1.834994 |
| | -20.906248 | μ_2 | -0.742144 |
| #M | 145.231364 | μ_1 | 158.702923 |
| | -96.789911 | μ_2 | -30.549108 |

TABLE III. Injected extrinsic parameters. They are shared in all cases.

| Parameter | Injected value | Unit |
|-----------|---------------------|-------------|
| ϕ_c | Merger phase | 1.3 rad |
| t_c | Merger time | 0 s |
| d_L | Luminosity distance | 200 Mpc |
| ψ | Polarization angle | 2.659 rad |
| ι | Orbital inclination | 0.4 rad |
| α | Right ascension | 1.375 rad |
| δ | Declination | -1.2108 rad |

D. Multiple detector case

We can also construct μ parameters satisfying Eq. (19) when we have data from multiple detectors. For instance, suppose we have three detectors, LIGO-Livingston (L), LIGO-Hanford (H), Virgo (V), and the Fisher matrix for the detectors are Γ_L , Γ_H , and Γ_V respectively, then the likelihood becomes

$$p(d|\psi) \propto \exp \left[-\frac{1}{2} \Gamma_{\text{LHV}} \Delta\psi_i \Delta\psi_j \right], \quad (30)$$

where $\Gamma_{\text{LHV}} \equiv \Gamma_L + \Gamma_H + \Gamma_V$. The Fisher matrix for each detector is proportional to the square of the SNR of the signal at the detector, so we write

$$\Gamma_{\text{LHV}} = q_L^2 \hat{\Gamma}_L + q_H^2 \hat{\Gamma}_H + q_V^2 \hat{\Gamma}_V, \quad (31)$$

where

$$\hat{\Gamma}_L = \begin{bmatrix} I_{-17/3} & I_{-5} & I_{-14/3} & I_{-4} & I_{-3} \\ I_{-5} & I_{-13/3} & I_{-4} & I_{-10/3} & I_{-7/3} \\ I_{-14/3} & I_{-4} & I_{-11/3} & I_{-3} & I_{-2} \\ I_{-4} & I_{-10/3} & I_{-3} & I_{-7/3} & I_{-4/3} \\ I_{-3} & I_{-7/3} & I_{-2} & I_{-4/3} & I_{-1/3} \end{bmatrix}_L (I_{-7/3})_L^{-1}, \quad (32)$$

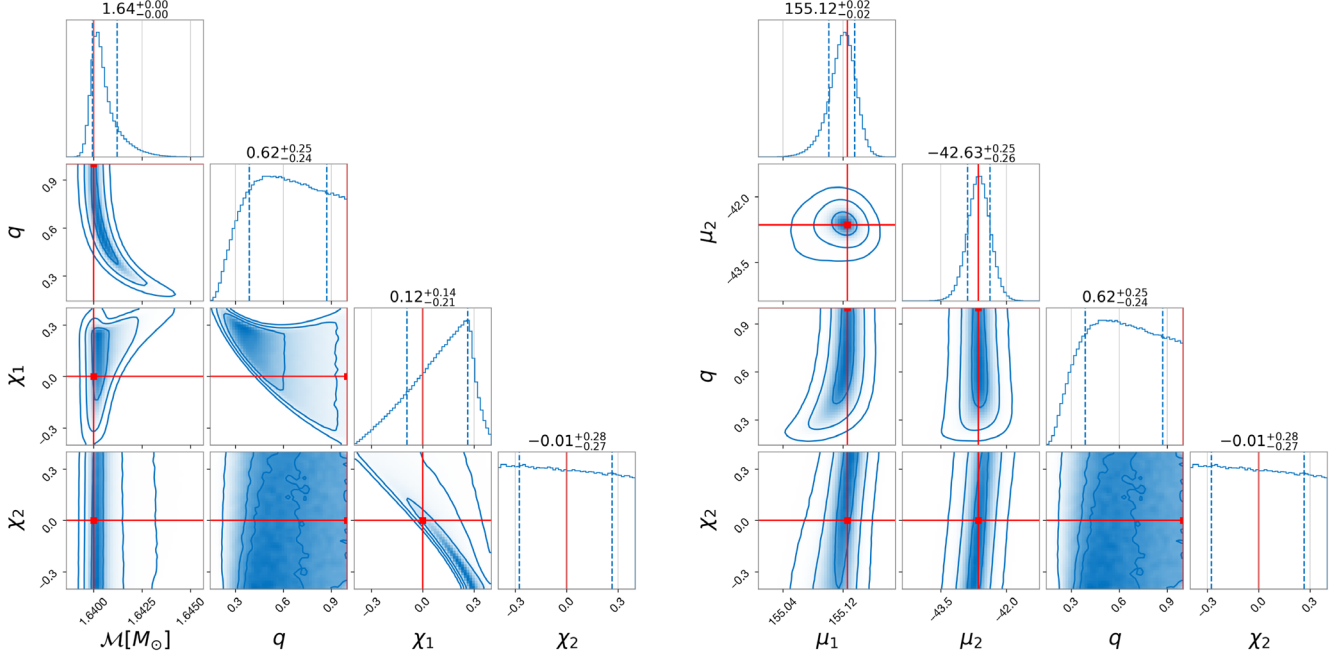


FIG. 2. Mass-spin part of the estimated posterior distribution in the injection test case #S2. The same posterior samples are plotted in the two different sets of mass and spin parameters. Dashed lines on each one-dimensional marginal distribution represent the 1σ region. Three contours on each two-dimensional marginal distribution represent the 1σ , 2σ , and 3σ regions, respectively. The red lines represent the injected value.

is parameter independent and therefore can be calculated in advance with the PSD of the detector L, and similarly for H and V.

Using the point estimate of SNRs given by the detection pipeline, we can calculate Γ_{LHV} , $\tilde{\Gamma}_{\text{LHV}}$, and U_{LHV} quickly. Note that if the PSDs for detectors are very similar we can just use U_{ij} calculated from any single detector regardless of the SNRs.

IV. INJECTION TEST

To confirm the effectiveness of the new sampling parameters, we conduct a series of injection tests, where CBC signals are artificially injected into simulated Gaussian noise and their source parameter values are recovered by parameter estimation analyses. The analyses are performed under the assumption that the sources have spins aligned with their angular momenta, and only the two spin components (χ_1, χ_2) are sampled. They are performed with the conventional set of sampling parameters $(\mathcal{M}, q, \chi_1, \chi_2)$ and our new set of sampling parameters $(\mu_1, \mu_2, q, \chi_2)$, and their efficiency and estimation results are compared.

The signal injection and parameter estimation process are managed through BILBY [14]. The injected CBC signals

are generated by the IMRPhenomD waveform model [22,23], and the same waveform model is used for recovering the source parameter values. The IMRPhenomD describes not only inspiral phase, but also merger and ringdown phases in contrast with the PN waveform we used in constructing the parameters. The waveform calculations are done via LALSimulation [24]. The integration range of the likelihood is from $f_{\min} = 20$ Hz to $f_{\max} = 2048$ Hz.

For parameter estimation, we use PTMCMCSampler [25]. Parallel tempering (PT) [26,27] is the main feature of this sampler and it makes the sampling efficient especially when the probability distribution is multimodal. Though PT can reduce the autocorrelation of samples, it is not used in our tests since we are more interested in the convergence than searching modes. For the same reason, we fix the starting point of the sampling to the injected parameter values.

PTMCMCSampler provides several built-in jump proposals, and custom proposals can be added to the sampling process. In here, two built-in proposals, single component adaptive metropolis (SCAM), and adaptive metropolis (AM) are used with the same weights. We use default options, with a minor modification; if there is no accepted proposal until the adaptation stage, the scale of jump proposals is changed by $1/2$.

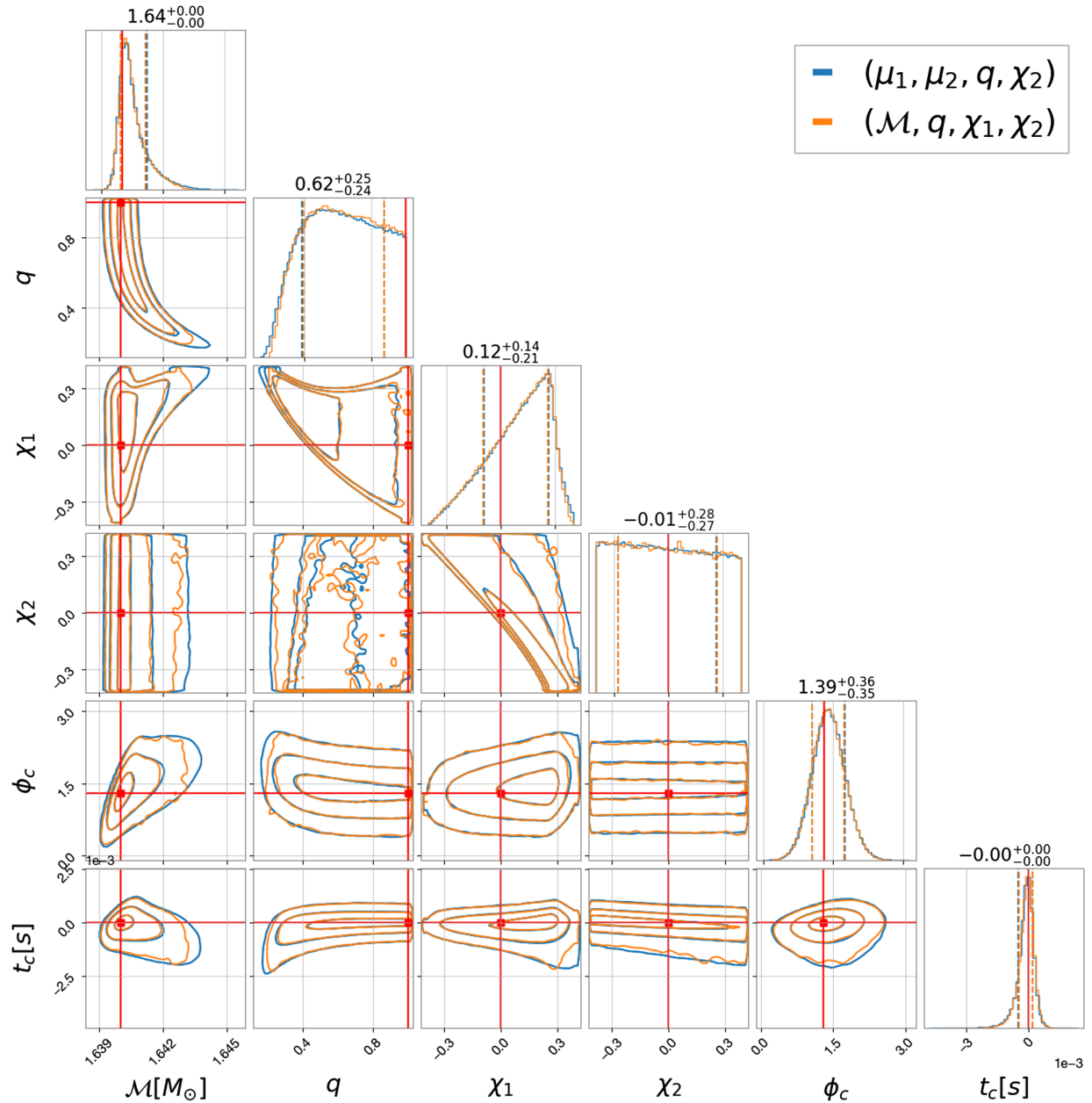


FIG. 3. Comparison of the estimated posterior distribution with (blue) and without (orange) reparametrization at the 10^8 th iteration in the injection test case #S2. Dashed lines on each one-dimensional marginal distribution represent the 1σ region. Three contours on each two-dimensional marginal distribution represent the 1σ , 2σ , and 3σ regions, respectively. The red lines represent the injected value.

A. Test cases

Table I lists the CBC signals we simulated and the spin prior range for each. Most of our injected signals are in the BNS mass range, where the parameter estimation analysis is computationally costly and its speedup is necessary. The massive case #S7 is to test the effectiveness of our sampling parameters in the BBH mass region, where our choice of $f_{\max} = 2048$ Hz for calculating μ_n is not valid. The lengths of data used for the analyses are 4 s for #S7 and 128 s for the other cases. For 128 s-data cases, we use the focused,

reduced-order quadrature (FROQ) technique in the calculation of likelihoods to speed up the estimation [21].

Our new set of sampling parameters deals with the correlation of masses and spins, which are measured by the frequency evolution of the signal. Thus, it is expected to be enough to test them with data from a single detector. In most of the tests (#S1–#S7), the analysis takes into account only the single detector, LIGO-Livingston. The last case #M is the exception, where a signal is injected into data of LIGO-Livingston, LIGO-Hanford, and Virgo, and all the

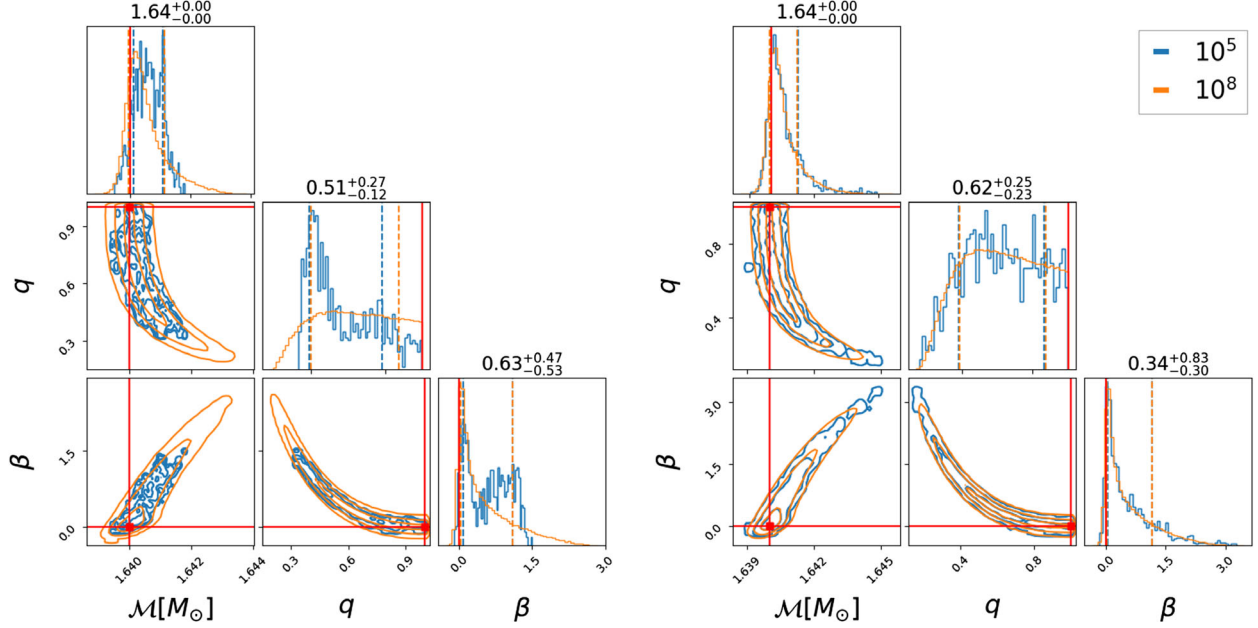


FIG. 4. Snapshots of the posterior samples at the 10^5 th (blue) and the 10^8 th (orange) iteration in the injection test case #S2. The left and the right panels show samples generated without and with our reparametrization, respectively. Dashed lines on each one-dimensional marginal distribution represent 1σ region. Three contours on each two-dimensional marginal distribution represent the 1σ , 2σ , and 3σ regions, respectively. The red lines represent the injected value.

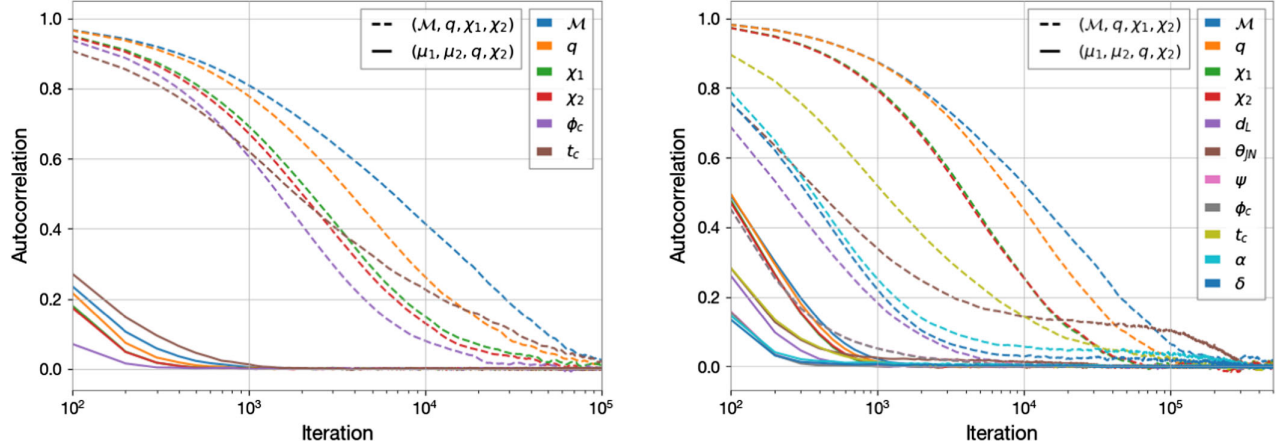


FIG. 5. Normalized autocorrelations of the generated samples in the two injection test cases, #S2 (left) and #M (right). The dashed and solid lines represent the sampling using conventional sampling parameters and new parameters, respectively. The symbols of parameters other than masses and spins are introduced in Table III.

data are used for parameter estimation. Their design sensitivities are used for generating Gaussian noise, and also as the PSDs used for calculating likelihood.

Since our method deals with the inefficiency coming from mass-spin degeneracy in the waveform, it is expected to have greater effect when spin parameters have broader prior range. To confirm this, we test cases with different spin prior range. The narrow spin prior $|\chi_{1,2}| < 0.05$ covers expected spins at merger of known BNS [6,28]. The semi-broad spin prior $|\chi_{1,2}| < 0.4$ covers all known neutron stars [6,29]. For this prior range, we consider

four injected signals (#S2–#5) to test the effectiveness of our sampling parameters for various mass and spin values. The broad spin prior $|\chi_{1,2}| < 0.99$ is used as a prior agnostic about the astrophysical nature of compact binaries.

In all the cases, the prior is uniform in component masses m_1, m_2 , and spins χ_1, χ_2 . For the sampling to be efficient, the explored range of \mathcal{M} is restricted to be $1.63 M_\odot \leq \mathcal{M} \leq 1.65 M_\odot$ for the cases except for #S7, and $25 M_\odot \leq \mathcal{M} \leq 31 M_\odot$ for #S7. The range of q is restricted to be $q > 1/8$.

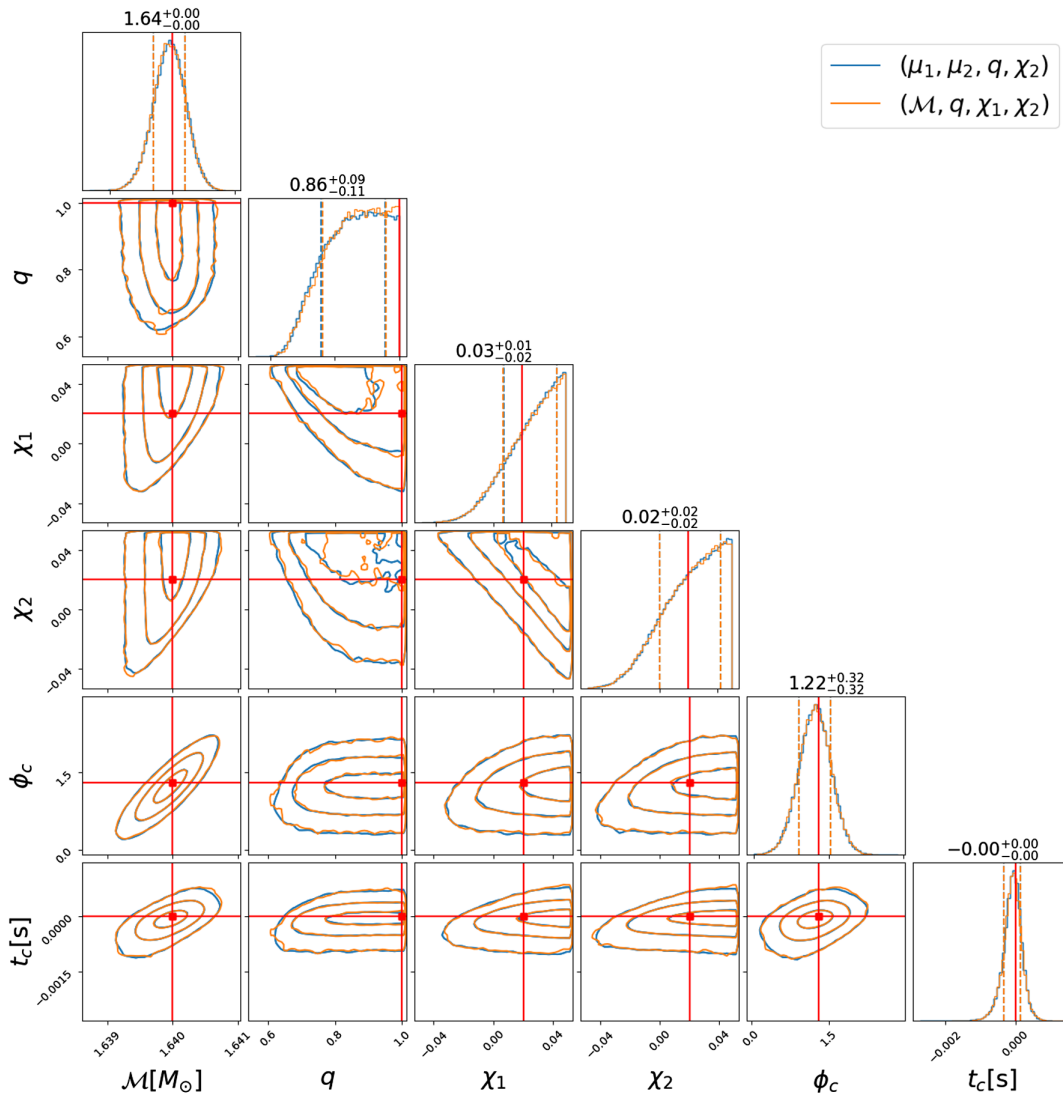


FIG. 6. Estimated posterior distribution with reparametrization in the injection test case #S1. Dashed lines on each one-dimensional marginal distribution represent the 1σ region. Three contours on each two-dimensional marginal distribution represent the 1σ , 2σ , and 3σ regions, respectively. The red lines represent the injected value.

For the estimation with reparametrization, we also set constraints on the range of μ_1 and μ_2 directly, to suppress invalid jump proposals. We just set the range by calculating the minimum and maximum values in the prior ranges of $(\mathcal{M}, q, \chi_1, \chi_2)$. Table II shows the values for each case. Since the range of \mathcal{M} is narrow, when the Jacobian (25) is evaluated during the sampling, it is evaluated with the injected value of \mathcal{M} rather than its current value. This approximation makes the prior $p(\mu_1, \mu_2, q, \chi_2)$ dependent only on q , and thus easy to be implemented. This approximation is optional and not necessary for our method. In the real data analyses, we should not use it since the chirp mass estimated by the detection pipeline could be biased significantly, especially for massive BBH events.

Other than masses and spins, there are seven parameters characterizing a CBC signal; merger phase, merger time,

luminosity distance to the source, polarization angle, the inclination angle between the line of sight and the orbital angular momentum, and right ascension and declination of the source. Their injected values are common in all the test cases and listed in the Table III. Their prior is the standard one used in the analysis by the LIGO-Virgo-KAGRA (See Appendix B of [2]). The range of merger time is $-0.1 \text{ s} < t_c < 0.1 \text{ s}$.

In the multiple-detector case #M, we infer all the 11 source parameters. For the single-detector cases, geometrical parameters such as right ascension and declination are not measurable. Thus, we infer only six of them; masses, spins, merger phase, and time, with the other parameter values being fixed to their injected values.

The PSD difference between #M and the other cases results in the difference in the parameter conversion. For the

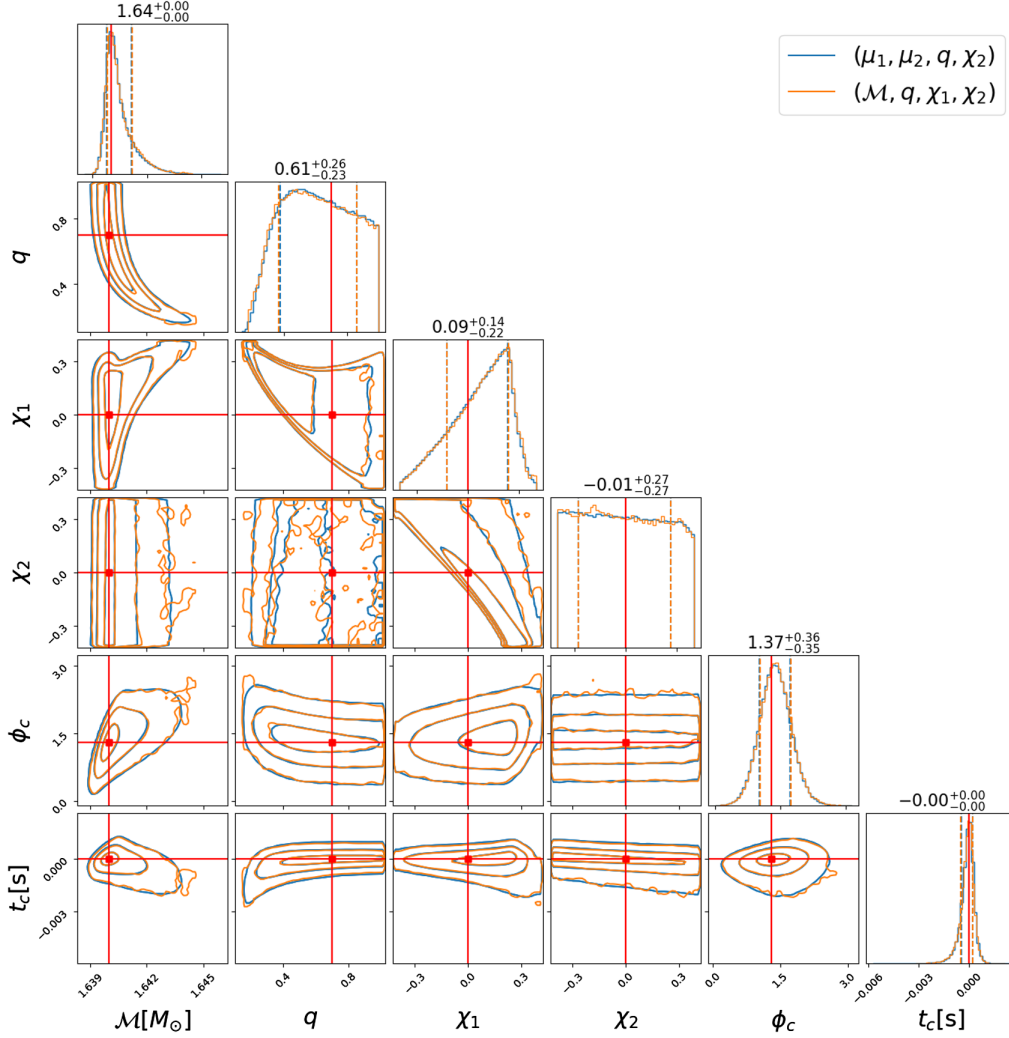


FIG. 7. Estimated posterior distribution with re-parameterization in the injection test case #S3. Dashed lines on each one-dimensional marginal distribution represent the 1σ region. Three contours on each two-dimensional marginal distribution represent the 1σ , 2σ , and 3σ regions, respectively. The red lines represent the injected value.

single-detector cases, setting $f_{\min} = 20$ Hz and $S_n(f)$ to the design sensitivity of LIGO-Livingston, μ_1 and μ_2 become

$$\mu_1 = 0.97320942\psi_1 + 0.21269341\psi_2 + 0.08732089\psi_3, \quad (33a)$$

$$\mu_2 = -0.22571628\psi_1 + 0.81153242\psi_2 + 0.53895018\psi_3. \quad (33b)$$

In the multiple detector case, assuming $q_L^2 \simeq q_H^2 \simeq q_V^2$ (the optimal SNRs of the injected signal are 10.93, 14.41, and 8.81 for L, H, and V respectively), μ_1 and μ_2 become

$$\mu_1 = 0.973164\psi_1 + 0.21292605\psi_2 + 0.08726009\psi_3, \quad (34a)$$

$$\mu_2 = -0.22599821\psi_1 + 0.81299212\psi_2 + 0.53662708\psi_3. \quad (34b)$$

B. Results

In this subsection, we visualize the results, from case #S2 as an example, and discuss what we can find from them. After that we list the improvement in the estimation efficiency of all cases.

First, we check whether the posterior distribution becomes simple in the new parameter space. In Fig. 2, we visualize the generated samples as one-dimensional and two-dimensional marginal distributions using CORNER.PY [30]. The left group is the distributions in $(\mathcal{M}, q, \chi_1, \chi_2)$ subspace, which shows strong correlations of parameters in the two-dimensional plots. On the other hand, in the right group, $(\mu_1, \mu_2, \chi_1, \chi_2)$ subspace, the posterior distribution shows weak correlation between any two parameters.

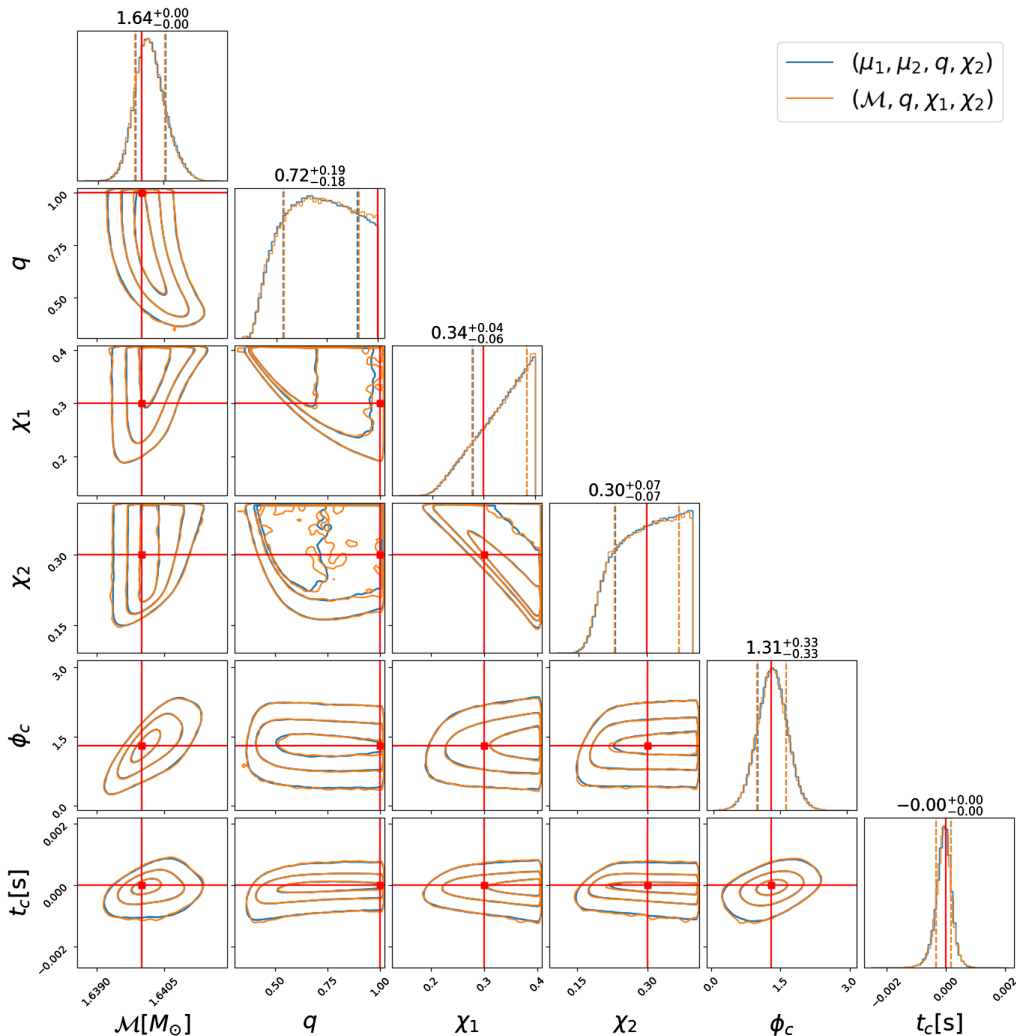


FIG. 8. Estimated posterior distribution with reparametrization in the injection test case #S4. Dashed lines on each one-dimensional marginal distribution represent the 1σ region. Three contours on each two-dimensional marginal distribution represent the 1σ , 2σ , and 3σ regions, respectively. The red lines represent the injected value.

Especially the μ_1 – μ_2 plot shows a hardly correlated distribution, which makes the exploration efficient greatly.

Next, we check that the estimation results are the same regardless to the sampling parameters. Figure 3 shows the estimation results with conventional and our new sets of sampling parameters. For better comparison, the density is not plotted in the two-dimensional plots, and only 1σ , 2σ , and 3σ contours are plotted. We can see the contours are very consistent between samplings with two different sampling parameters. Also, the one-dimensional distributions clearly agree.

While the reparametrization does not change the result of estimation, it can reduce the estimation time. Figure 4 is a comparison of the distribution of posterior samples generated by the 10^5 th iteration and the 10^8 th iteration. Only \mathcal{M} , q , β of samples are plotted here for simplicity. In the left plot, which is from the estimation without reparametrization, the distribution at the 10^5 th iteration is quite

different from the converged distribution. In the two-dimensional plots, we can see an unexplored region. It takes many more iterations to explore the region and converge. On the contrary, in the right plot, which is from the estimation with the mass-spin reparametrization, the samples are already distributed in the entire converged distribution region at the 10^5 th iteration; the one-dimensional marginal distribution for \mathcal{M} and β are quite stable at the 10^5 th iteration.

The autocorrelation function of samples can be used to quantify the enhancement of the convergence speed of MCMC algorithms [31–34]. In Fig. 5, each curve represents a normalized autocorrelation function of samples for each parameter. The left panel is the result from #S2, and the right panel is from #M. The dashed lines are from the estimation without reparametrization, and the solid lines are from the estimation with reparametrization. Compared to dashed lines, we can see that the solid lines fall to zero

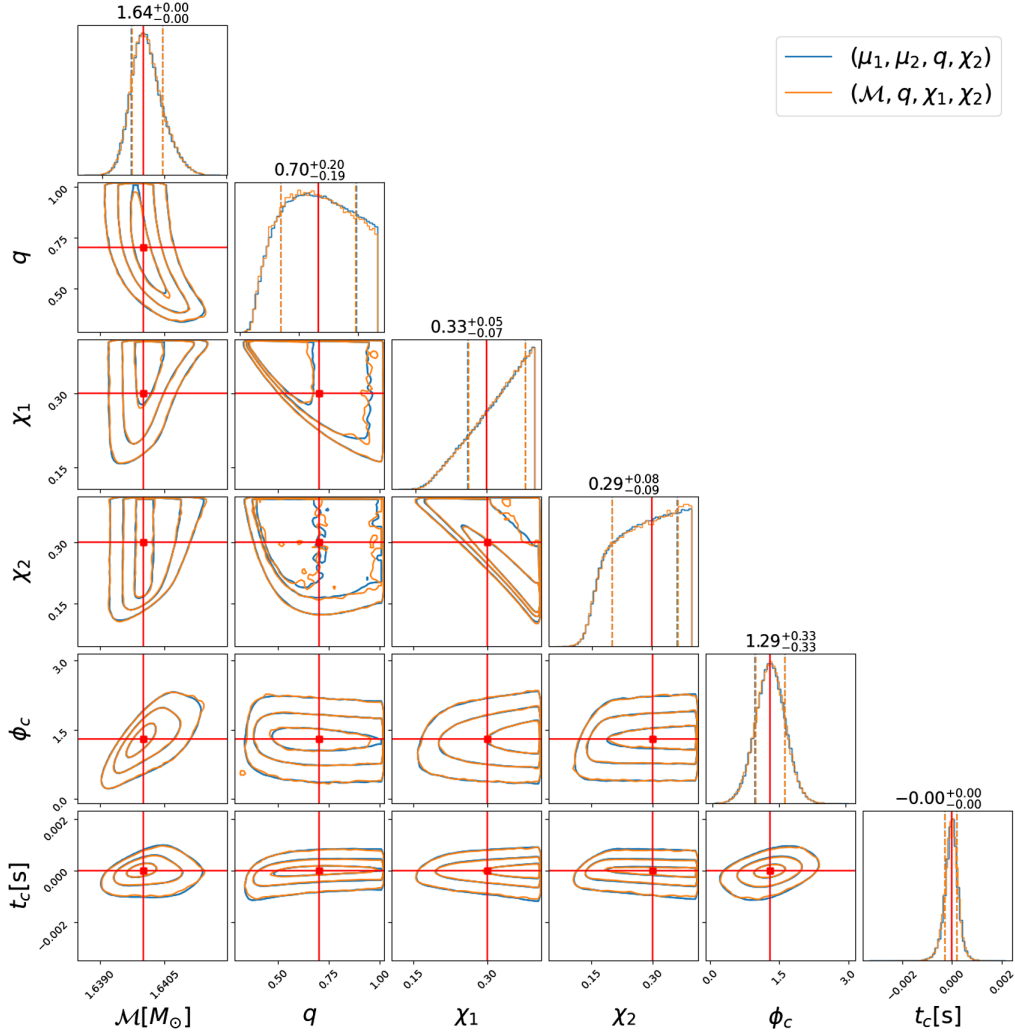


FIG. 9. Estimated posterior distribution with reparametrization in the injection test case #S5. Dashed lines on each one-dimensional marginal distribution represent the 1σ region. Three contours on each two-dimensional marginal distribution represent the 1σ , 2σ , and 3σ regions, respectively. The red lines represent the injected value.

faster, which means that the samples are less correlated with near ones, and thus, statistically independent samples are generated more frequently. Note that, the autocorrelation functions for extrinsic parameters also fall to zero faster with reparametrization.

The number of iterations to obtain a statistically independent sample is called integrated autocorrelation time (IAT), and can be calculated as

$$\text{IAT} \approx 1 + 2 \sum_{\tau} \hat{c}(\tau), \quad (35)$$

where τ indicates each of iterations and $\hat{c}(\tau)$ is the normalized autocorrelation function. Therefore we can approximate the speed-up gain by the ratio of IAT, between the estimation with and without the reparametrization.

The IAT from all cases are listed in Table IV. A set of samples has different autocorrelation functions and IATs for different parameters. Here, we compare IAT maximized over physical parameters. We can see that our reparametrization reduces the IAT by a factor of ~ 10 for BNS with a narrow-spin prior ($|\vec{\chi}| < 0.05$) and ~ 100 for a broad-spin prior ($|\vec{\chi}| < 0.99$). Even for the massive BBH case, the IAT becomes about $1/3$ using our reparametrization.

Although additional time is taken for parameter conversion in each iteration with our method, its effect to the total estimation time is small. To check this, we compare the estimation time directly. In the case #M, it takes 17 minutes to get 1000 statistically independent samples with reparametrization on ICRR common computer system [35] [CPU; Intel Xeon Gold 6230 (2.1 GHz)], while 41 hours are needed without reparametrization. In all cases

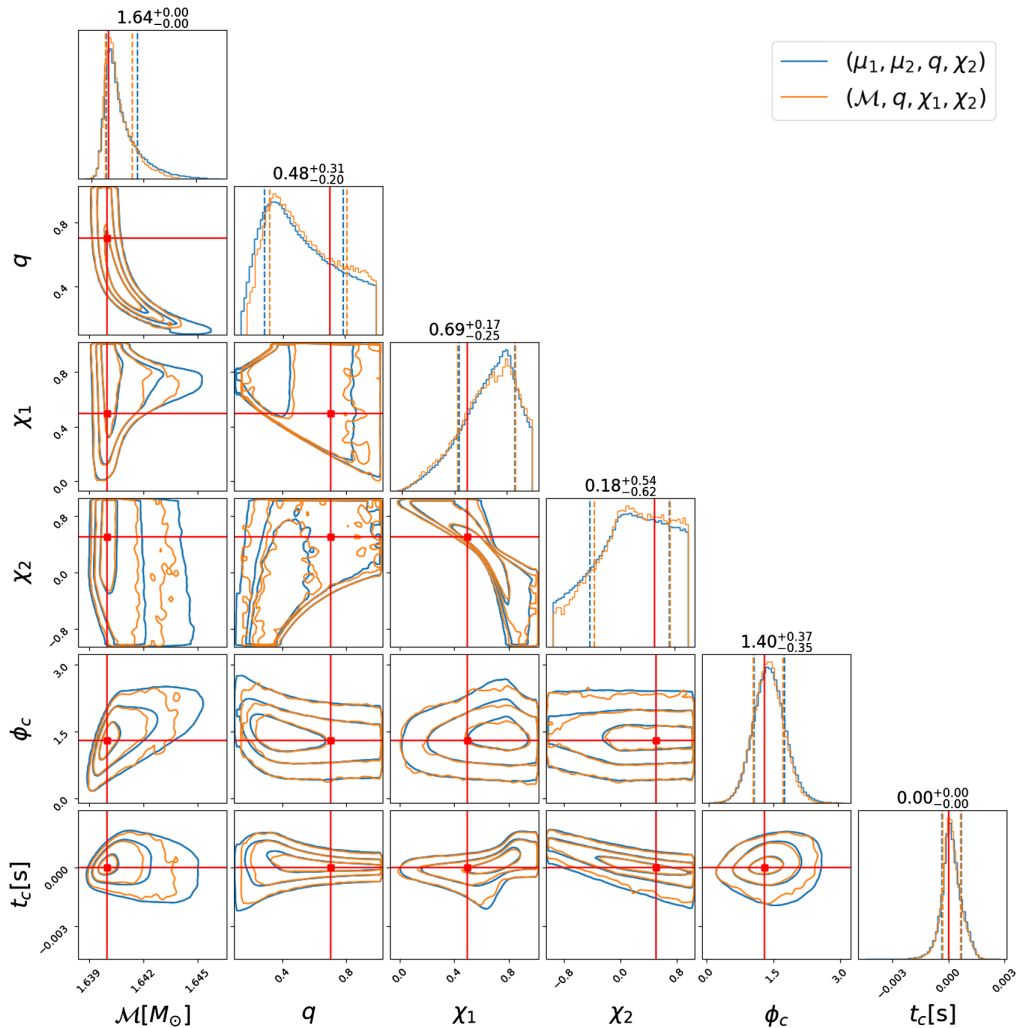


FIG. 10. Estimated posterior distribution with reparametrization in the injection test case #S6. Dashed lines on each one-dimensional marginal distribution represent the 1σ region. Three contours on each two-dimensional marginal distribution represent the 1σ , 2σ , and 3σ regions, respectively. The red lines represent the injected value.

the total estimation time is reduced. For reference, we plot the estimation results comparison from the test cases other than #S2 too, in Figs. 6–12.

V. CONCLUSIONS

In this paper we have introduced a new set of mass-spin parameters for aligned-spin compact binary inspiral waveform, which makes the posterior distribution simple and therefore the estimation efficient. To inspect its effect on the sampling efficiency, we performed parameter estimation runs on simulated signals using the new set of mass-spin sampling parameters. In all test cases, the new set of parameters improves the efficiency of the sampling process. The improvement is especially remarkable for the analysis

of binary neutron star signals with a broad prior range of spins, where the effects of mass-spin correlations are significant. Quantitatively, the speed-up gain in the analysis of binary neutron star signals is ~ 10 for narrow-spin priors ($|\vec{\chi}| < 0.05$), ~ 10 – 100 for semibroad-spin priors ($|\vec{\chi}| < 0.4$), and ~ 100 for broad-spin priors ($|\vec{\chi}| < 0.99$).

The results are case dependent, thus they have to be understood carefully. In the tests, we adopted single component adaptive Metropolis- and adaptive Metropolis-jump proposals. The choice is natural, but there could be other jump proposals that make the sampling with complicated posteriors more efficient. Those proposals may reduce the improvement of our method since it settles the inefficiency in a different way. The comparison of the sampling with our method, the sampling with other

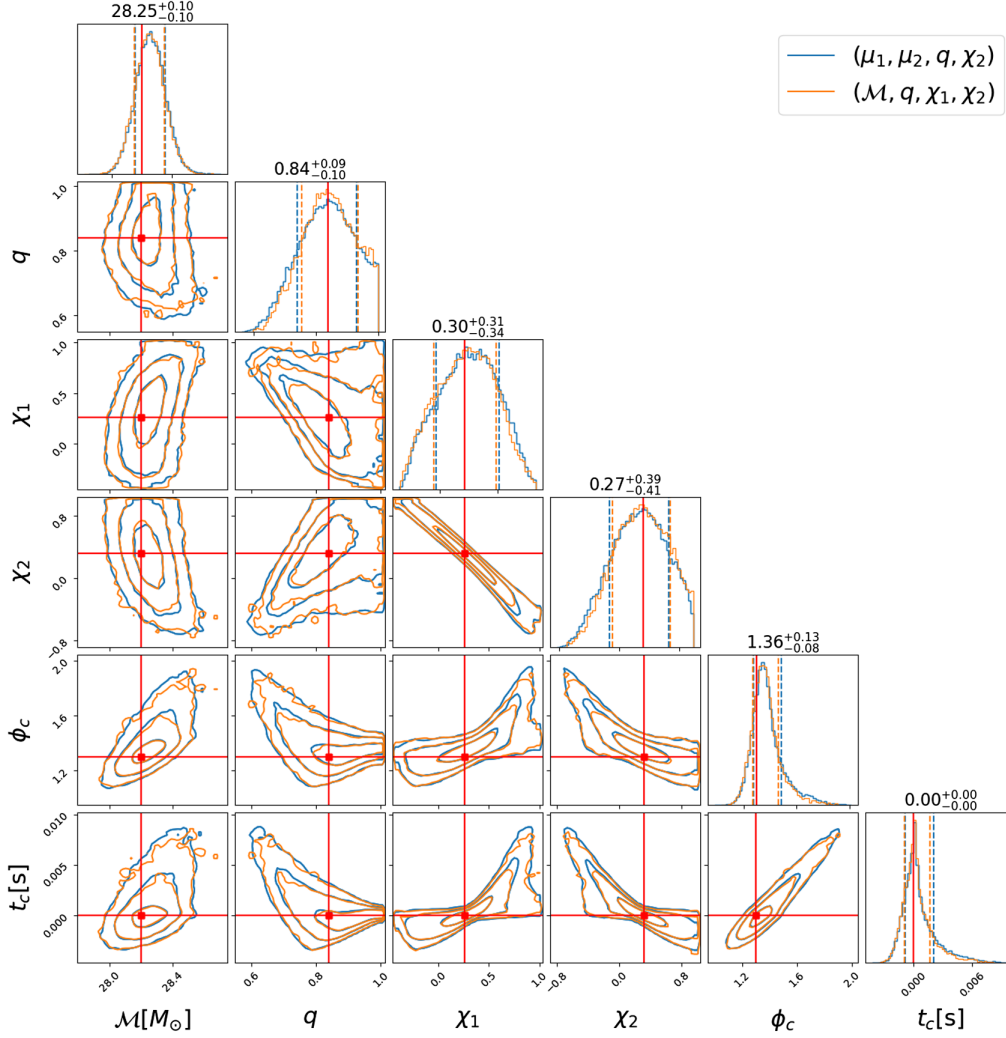


FIG. 11. Estimated posterior distribution with re-parameterization in the injection test case #S7. Dashed lines on each one-dimensional marginal distribution represent the 1σ region. Three contours on each two-dimensional marginal distribution represent the 1σ , 2σ , and 3σ regions, respectively. The red lines represent the injected value.

proposals, and the sampling with both could be a future work. Nevertheless, our method has the advantage that it can solve the problem of degeneracy in any parameter estimation pipeline with any sampler.

We can combine other parameter estimation techniques with our reparameterization. Since our method reduces the number of likelihood evaluations, it could be a nice duo with methods that cut down the single-likelihood evaluation time, such as the focused reduced-order quadrature technique. Our method also can be used with parallel tempering, which can also reduce the estimation time.

In the narrow-spin prior case (#S1) and the massive case (#S7), the posterior distribution already has a relatively simple form in the usual mass-spin parameter space. Even in these cases, our reparameterization improves the sampling process, and at least does not worsen the sampling. This fact, along with the improvement in the multiple-detectors

case, shows our method could be applied to the actual observation comprehensively.

In this paper we only consider a binary system whose spins are aligned with its orbital angular momentum. If the spins are misaligned with the orbital angular momentum, the precession of the orbital plane is induced, and the amplitude and phase of the signal are modulated [36]. Since it can break the degeneracy between the distance and orbital-inclination angle, and improve the accuracy of source localization [37,38], rapid parameter estimation (taking into account the precession effects) can be helpful for multimessenger observations. Even in this case, the frequency evolution of the signal is predominantly determined by masses and spin components along the orbital angular momentum, and our reparameterization may still make the posterior distribution simple and the analysis more efficient. We leave the extension of our method to precessing binary systems for a future work.

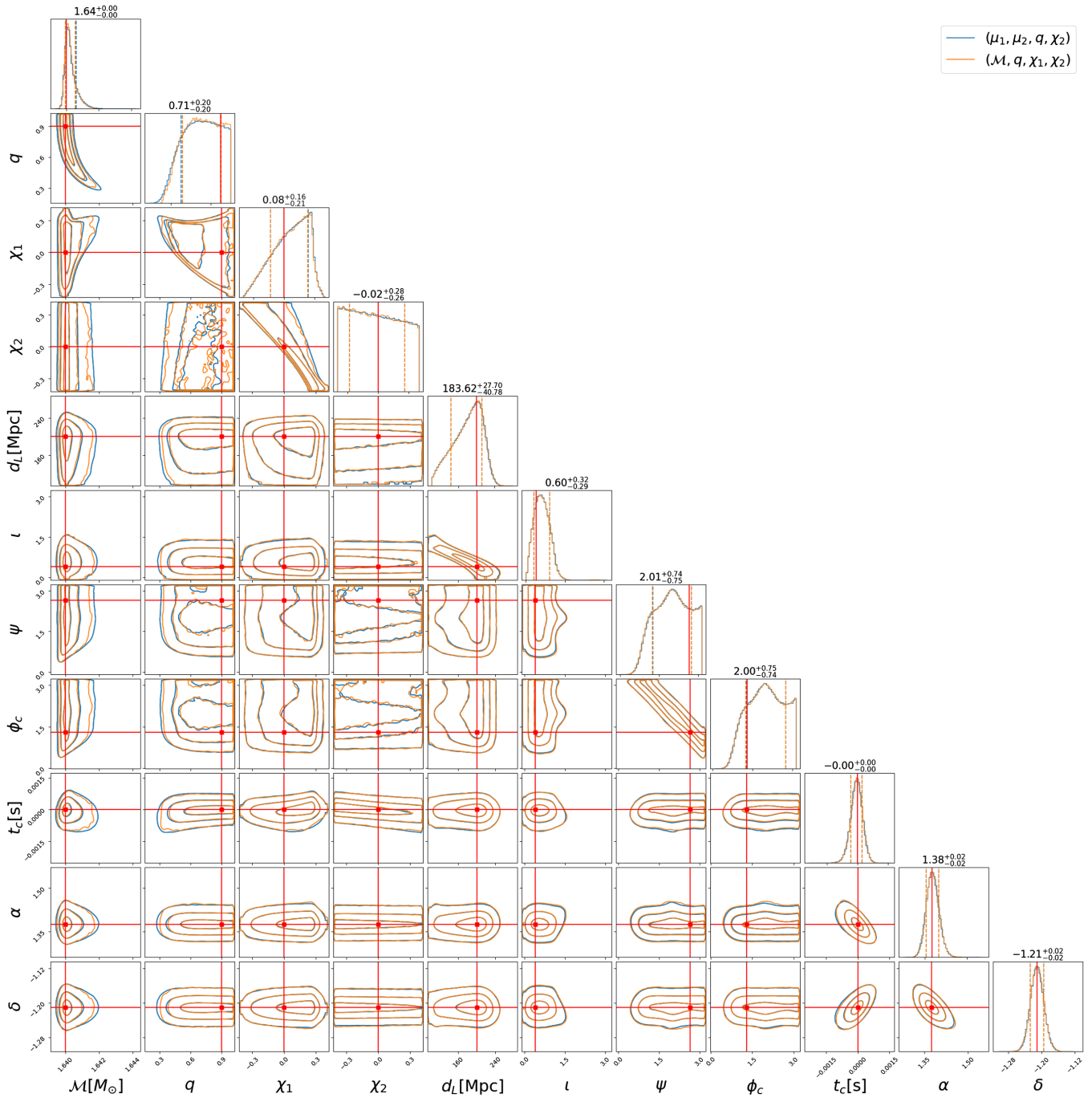


FIG. 12. Estimated posterior distribution with reparametrization in the injection test case #M. Dashed lines on each one-dimensional marginal distribution represent the 1σ region. Three contours on each two-dimensional marginal distribution represent the 1σ , 2σ , and 3σ regions, respectively. The red lines represent the injected value.

TABLE IV. The comparison of maximal (over physical parameters) IAT values estimated from the samples.

| Case | $(\mathcal{M}, q, \chi_1, \chi_2)$ | $(\mu_1, \mu_2, q, \chi_2)$ | Ratio |
|------|------------------------------------|-----------------------------|-------|
| #S1 | 427 | 47.8 | 8.93 |
| #S2 | 3.44×10^4 | 250 | 138 |
| #S3 | 7.68×10^4 | 276 | 278 |
| #S4 | 1.46×10^3 | 50.4 | 29.0 |
| #S5 | 2.00×10^3 | 53.9 | 37.1 |
| #S6 | 1.10×10^5 | 717 | 153 |
| #S7 | 1.45×10^4 | 5.04×10^3 | 2.88 |
| #M | 6.17×10^4 | 401 | 154 |

ACKNOWLEDGMENTS

We thank Kyohei Kawaguchi, Tatsuya Narikawa, Nami Uchikata, Bin-Hua Hsieh, and Takashi Kato for useful discussions and comments. This work was supported by NSF Grant No. PHY-1912649 (S.M.), MEXT, JSPS

Leading-edge Research Infrastructure Program, JSPS Grant-in-Aid for Specially Promoted Research 26000005, JSPS Grant-in-Aid for Scientific Research on Innovative Areas 2905 Grants No. JP17H06358, No. JP17H06361, No. JP16H02183, and No. JP17H06364, JSPS Core-to-Core Program A. Advanced Research Networks, JSPS Grant-in-Aid for Scientific Research (S) Grants No. 17H06133 and No. 15H00787, the joint research program of the Institute for Cosmic Ray Research, the cooperative research program of the Institute of Statistical Mathematics, National Research Foundation (NRF) and Computing Infrastructure Project of KISTI-GSDC in Korea, Academia Sinica, AS Grid Center and the Ministry of Science and Technology (MoST) in Taiwan under grants including Grant No. AS-CDA-105-M06, Advanced Technology Center of NAOJ, Mechanical Engineering Center of KEK, the LIGO project, and the Virgo project.

-
- [1] B. P. e. Abbott (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.* **116**, 061102 (2016).
- [2] B. Abbott, R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt *et al.*, *Phys. Rev. X* **9**, 031040 (2019).
- [3] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Phys. Rev. X* **11**, 021053 (2021).
- [4] R. Abbott *et al.* (LIGO Scientific, VIRGO Collaborations), [arXiv:2108.01045](https://arxiv.org/abs/2108.01045).
- [5] R. Abbott *et al.* (LIGO Scientific, VIRGO, KAGRA Collaborations), [arXiv:2111.03606](https://arxiv.org/abs/2111.03606).
- [6] B. Abbott, R. Abbott, T. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, V. Adya *et al.*, *Phys. Rev. Lett.* **119**, 161101 (2017).
- [7] R. Abbott *et al.* (LIGO Scientific, Virgo Collaborations), *Astrophys. J. Lett.* **913**, L7 (2021).
- [8] R. Abbott *et al.* (The LIGO Scientific Collaboration, The Virgo Collaboration, and The KAGRA Collaboration), [arXiv:2111.03634](https://arxiv.org/abs/2111.03634).
- [9] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**, 1087 (1953).
- [10] W. K. Hastings, *Biometrika* **57**, 97 (1970).
- [11] J. Skilling, *Bayesian Anal.* **1**, 833 (2006).
- [12] J. Veitch *et al.*, *Phys. Rev. D* **91**, 042003 (2015).
- [13] C. M. Biwer, C. D. Capano, S. De, M. Cabero, D. A. Brown, A. H. Nitz, and V. Raymond, *Publ. Astron. Soc. Pac.* **131**, 024503 (2019).
- [14] G. Ashton *et al.*, *Astrophys. J. Suppl. Ser.* **241**, 27 (2019).
- [15] C. Cutler and E. E. Flanagan, *Phys. Rev. D* **49**, 2658 (1994).
- [16] T. Tanaka and H. Tagoshi, *Phys. Rev. D* **62**, 082001 (2000).
- [17] B. S. Sathyaprakash and B. F. Schutz, *Classical Quantum Gravity* **20**, S209 (2003).
- [18] D. A. Brown, I. Harry, A. Lundgren, and A. H. Nitz, *Phys. Rev. D* **86**, 084017 (2012).
- [19] F. Ohme, A. B. Nielsen, D. Keppel, and A. Lundgren, *Phys. Rev. D* **88**, 042002 (2013).
- [20] A. Pai and K. G. Arun, *Classical Quantum Gravity* **30**, 025011 (2013).
- [21] S. Morisaki and V. Raymond, *Phys. Rev. D* **102**, 104020 (2020).
- [22] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, *Phys. Rev. D* **93**, 044006 (2016).
- [23] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, *Phys. Rev. D* **93**, 044007 (2016).
- [24] LIGO Scientific Collaboration, LIGO Algorithm Library—LALSuite, free software (GPL), [10.7935/GT1W-FZ16](https://doi.org/10.7935/GT1W-FZ16) (2018).
- [25] J. Ellis and R. van Haasteren, jellis18/ptmcmcsampler: Official release, [10.5281/zenodo.1037579](https://doi.org/10.5281/zenodo.1037579) (2017).
- [26] R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.* **57**, 2607 (1986).
- [27] C. J. Geyer, Computing science and statistics, in *Proceedings of the 23rd Symposium on the Interface* (American Statistical Association, New York, 1991), p. 156.
- [28] M. Burgay, N. D’Amico, A. Possenti, R. N. Manchester, A. G. Lyne, B. C. Joshi, M. A. McLaughlin, M. Kramer, J. M. Sarkissian, F. Camilo, V. Kalogera, C. Kim, and D. R. Lorimer, *Nature (London)* **426**, 531 (2003).
- [29] J. W. T. Hessels, S. M. Ransom, I. H. Stairs, P. C. C. Freire, V. M. Kaspi, F. Camilo, *Science* **311**, 1901 (2006).
- [30] D. Foreman-Mackey, *J. Open Source Software* **1**, 24 (2016).

- [31] A. Sokal, Bosonic algorithms, in *Quantum Fields on the Computer*, edited by M. Creutz (World Scientific, Singapore, 1992), [10.1142/1634](https://doi.org/10.1142/1634).
- [32] N. Kawashima, J.E. Gubernatis, and H.G. Evertz, *Phys. Rev. B* **50**, 136 (1994).
- [33] K. Binder and E. Luijten, *Phys. Rep.* **344**, 179 (2001).
- [34] H. Müller-Krumbhaar and K. Binder, *J. Stat. Phys.* **8**, 1 (1973).
- [35] Icr computer system, <https://www.icrr.u-tokyo.ac.jp/cal2/>.
- [36] T.A. Apostolatos, C. Cutler, G.J. Sussman, and K.S. Thorne, *Phys. Rev. D* **49**, 6274 (1994).
- [37] S. Vitale and H.-Y. Chen, *Phys. Rev. Lett.* **121**, 021303 (2018).
- [38] T. Tsutsui, A. Nishizawa, and S. Morisaki, *Phys. Rev. D* **104**, 064013 (2021).