Rapid gravitational wave parameter estimation with a single spin: Systematic uncertainties in parameter estimation with the SpinTaylorF2 approximation

B. Miller, R. O'Shaughnessy, T. B. Littenberg, and B. Farr

Phys. Rev. D **92**, 044056 — Published 31 August 2015

DOI: 10.1103/PhysRevD.92.044056

Rapid gravitational wave parameter estimation with a single spin: Systematic uncertainties in parameter estimation with the SpinTaylorF2 approximation

B. Miller and R. O'Shaughnessy
*Center for Computational Relativity and Gravitation, Rochester Institute of Technology, Rochester, NY 14623, USA**

T.B. Littenberg
Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA)& Dept. of Physics and Astronomy, 2145 Sheridan Rd, Evanston, IL 60208, USA

B. Farr
Enrico Fermi Institute, University of Chicago, Chicago, IL 60637, USA

Reliable low-latency gravitational wave parameter estimation is essential to target limited electromagnetic followup facilities toward astrophysically interesting and electromagnetically relevant sources of gravitational waves. In this study, we examine the tradeoff between speed and accuracy. Specifically, we estimate the astrophysical relevance of systematic errors in the posterior parameter distributions derived using a fast-but-approximate waveform model, `SpinTaylorF2` (`STF2`), in parameter estimation with `lalinference_mcmc`. Though efficient, the `STF2` approximation to compact binary inspiral employs approximate kinematics (e.g., a single spin) and an approximate waveform (e.g., frequency domain versus time domain). More broadly, using a large astrophysically-motivated population of generic compact binary merger signals, we report on the effectualness and limitations of this single-spin approximation as a method to infer parameters of generic compact binary sources. For most low-mass compact binary sources, we find that the `STF2` approximation estimates compact binary parameters with biases comparable to systematic uncertainties in the waveform. We illustrate by example the effect these systematic errors have on posterior probabilities most relevant to low-latency electromagnetic followup: whether the secondary has a mass consistent with a neutron star; whether the masses, spins, and orbit are consistent with that neutron star's tidal disruption; and whether the binary's angular momentum axis is oriented along the line of sight.

I. INTRODUCTION

Ground based gravitational wave detector networks (notably advanced LIGO [1] and Virgo [2]) are sensitive to the relatively well understood signal from the lowest-mass compact binaries $M = m_1 + m_2 \leq 16 M_\odot$ [3–14]. Strong signals permit high-precision constraints on binary parameters, particularly when the binary precesses. Precession arises only from spin-orbit misalignment; occurs on a distinctive timescale between the inspiral and orbit; and produces distinctive polarization and phase modulations [15–17]. As a result, the complicated gravitational wave signal from precessing binaries is unusually rich, allowing high-precision constraints on multiple parameters, notably the (misaligned) spin [18, 19]. Measurements of the spin orientations alone could provide insight into processes that affect spin alignment, such as supernova kicks [20, 21], tides and post-Newtonian resonances [22]. More broadly, gravitational waves constrain the pre-merger orbital plane and total angular momentum direction, both of which may correlate with the presence, beaming, and light curve [23–25] of any post-merger ultrarelativistic blastwave (e.g, short GRB) [26]. Moreover, spin-orbit coupling strongly influences orbital decay and hence the overall gravitational wave phase: the accuracy with which most other parameters can be determined is limited by knowledge of BH spins [19, 27–29]. Precession is known to break this degeneracy [18, 19, 30–33]. In sum, the rich gravitational waves emitted from a precessing binary allow higher-precision measurements of individual neutron star masses, black hole masses, and black hole spins, enabling constraints on their distribution across multiple events. In conjunction with electromagnetic measurements, the complexity of a fully precessing gravitational wave signal may enable correlated electromagnetic and gravitational wave measurements to much more tightly constrain the central engine of short gamma ray bursts.

Accurately simulating richness and complexity comes at a price: essentially,[1] the additional computational weight of numerically evolving several ODEs for the spin and orbit dynamics. This cost places a substantial bur-

*Electronic address: `oshaughn@mail.rit.edu`

[1] Recently, Kesden et al provided a new approach to evolving double-spin compact binaries, potentially enabling more rapid time- and frequency-domain solutions to the spin precesion and orbit equations. Though possible to use this approach to generate waveforms in principle, no implementation is available in `lalsimulation` for `lalinference` at the time of writing.

den on attempts to reconstruct the sources of an observed wave, since data must be systematically compared against all possible candidate signals [18, 34–41]. Owing both to the relatively large number of parameters needed to specify a precessing binary's orbit and to the seemingly-complicated evolution, Bayesian parameter estimation methods have only recently [41, 42] become efficient enough to draw inferences about gravitational waves from a statistically significant sample of generic precessing sources [43]. At the time of writing, successful calculations of this type require run times on the order of days with existing computational resources.

Except for a relatively small corner of parameter space, however, one black hole spin dominates any precession; the smaller mass usually has a strongly suppressed impact on angular momentum evolution because black hole spin ($\mathbf{S}$) scales as the black hole mass m squared ($\mathbf{S} = m^2\boldsymbol{\chi}$). The dynamics of a single spin are well-understood [15] and easily approximated in the time and frequency domain. In particular, the `SpinTaylorF2` (`STF2`) [17] approximation provides a fast and accurate model for binary inspiral, valid over a broad range of mass ratios and spins. At the time of writing, parameter estimation using `STF2` as a template requires roughly an order of magnitude less time on comparable resources, often of order a few tens of minutes.

Low-latency parameter estimation on these timescales enables transformative followup electromagnetic observations [44]. The most tantalizing proposed electromagnetic counterparts to compact binary merger are expected to be brief, potentially disappearing within days if not much sooner [45–50]. Given limited resources, reliable low-latency parameter estimation of gravitational wave signals will significantly enhance the science output of multimessenger, time-domain astronomy. Gravitational wave and electromagnetic observatories and observers have demonstrated a commitment to realizing the potential of low-latency followup [38, 51–57]. The use of rapid single-spin templates like `STF2` may be a critical ingredient in enabling these followup observations, if parameter estimation with this approximation provides suffciently robust predictions.

In this paper, we systematically assess the accuracy of parameter estimation with `STF2`. The `STF2` approximation is computationally efficient because (a) it uses only a single spin, eliminating three subdominant degrees of freedom and enabling fast, analytic solutions to the precession equations; and particularly (b) because it constructs a highly efficient stationary-phase approximation to the single-spin kinematics and gravitational wave emission. Though these approximations introduce systematic errors by neglecting higher-order post-Newtonian terms (e.g., associated with two-spin effects), on theoretical grounds one expects these uncertainties to be comparable to the neglected post-Newtonian terms. In other words, on theoretical grounds one anticipates comparable differences between (i) the predictions constructed using the standard adiabatic quasicircular inspiral models `SpinTaylorT4` (`STT4`) and `SpinTaylorT2` (`STT2`) where both black holes are allowed to have generic spins [3, 6, 14]; (ii) between `STT2` with two spins and `STT2` with only one nonzero spin, restricted to the more massive onbject; and (iii) between `STT2` with two or one spin and `STF2`. As part of a larger study of parameter estimation on an astrophysically-selected sample of sources [43], in this work we systematically evaluate these hypotheses and the astrophysical impact that systematic biases introduce.

This paper is organized as follows. In Section II we introduce our parameter estimation study, describing the population of events used; the specific models adopted to infer compact binary parameters; and the specific techniques we used to reconstruct each posterior parameter distribution. In Section III we introduce and employ standard statistical tools (i.e., Student's t disribution) to identify and quantify systematic differences between the posterior parameter distributions arrived at by using different waveform models. In Section IV we demonstrate these systematic errors, though statistically significant, rarely significantly impact our astrophysical conclusions, particularly because these errors are comparable to other systematic uncertainties. In Section V we summarize our conclusions. An appendix A briefly reviews statistical tools used in our analysis.

Context and related work

Our study is the first large-scale investigation of parameter estimation accuracy with approximate precessing templates, using production-scale code and an astrophysically-motivated sample. Several groups are pursuing complementary methods to accelerate parameter estimation for precessing binaries, including both fast and accurate waveform models [58–61] [62] [63–66], and alternative architectures for the likelihood function [67].

As an example of a fast but accurate waveform model, Klein and collaborators have developed methods to construct a highly faithful SPA-like fourier transform of the line-of-sight waveform $h(t, \hat{n})$ [58–61]. This accurate but technically sophisticated approach differs substantially from the simpler and more approximate `STF2`, defined as a term-by-term stationary-phase-approximated fourier transform of $h(t, \hat{n}) = \sum_{lm} h_{lm}(t) Y^{(-2)}_{lm}(\hat{n})$, using a corotating-frame expansion of $h_{lm}(t)$ to avoid precession-induced phase catastrophes. By design extremely faithful, the latest Klein et al. approximation [61] has been shown to enable faster parameter estimation, allowing the authors to assess hypotheses about precessing *double*-spin binaries [60]. As another example, reduced-order-modeling and SVD methods in principle offer a robust and rapid procedure to efficiently approximate *any* waveform, reconstructing the signal from a sparse set of basis signals and (interpolated) functions of parameters [63–66, 68, 69]. Given the challenge of high-dimensional interpolation, reduced order models have to

date been incrementally applied to high-cost and high-precision but low-dimensional models like nonprecessing EOBNR [66, 70, 71], an inspiral-merger-ringdown waveform; and IMRPhenomP [62], a cousin to `STF2` which includes approximate merger and ringdown (Raymond et al, private communication). At the time of writing, `STF2` is the fastest available waveform model including precession.

After the cost of waveform generation, the computational cost of parameter estimation is dominated by the cost per likelihood, which historically has been dominated by the cost of operating on and fourier transforming long arrays. Recently, several methods have been proposed to perform this comparison more efficiently [63–65, 67, 68], by interpolating some combination of the the waveform or likelihood; by adopting a sparse representation to reduce the computational cost of data handling; or by organizing the calculation to maximize the reuse of intermediate results obtained from these long and costly array operations.

II. PARAMETER ESTIMATION ON AN ASTROPHYSICAL SAMPLE

Gravitational wave parameter estimation involves systematic comparison of candidate waveforms to data [18, 34–41]. We use the `lalinference_mcmc` [41] parameter estimation code to infer posterior parameter distributions, a method and tool validated by extensive prior studies including [18, 29, 40, 43, 51, 72–75]. This code can construct synthetic data from and perform inference with any of a wide range of waveform modes present in the `lalsimulation` software library. In this study, we always generated our candidate precessing signals using the `STT2` approximation described in [76, 77]. Using this fixed data set, we recovered parameters using `STT4`, described in [3, 14], with one or two spins; `STT2`, with one or two spins; and `STF2`. In all cases we used a leading-order (Newtonian) amplitude; 3.5 PN order in orbital phase; up to 3.5 PN in spin-orbit and 2PN spin-spin terms; and treated all objects as point particles with black-hole like couplings in, for example, the quadrupole-monopole terms. This binary evolution is terminated prior to merger, either when it reaches the minimum energy circular orbit; when the orbital frequency ceases to increase monotonically; or when the post-Newtonian v/c is greater than unity. While our simulated binaries start their evolution when twice their orbital frequency is 20 Hz, as in previous work we specify binary spin parameters at a reference frequency $f_{\rm ref} = 100\,{\rm Hz}$ [19, 29, 42, 72]. In these analyses, we adopt a fiducial 3-detector network: advanced LIGO [1] and Virgo [2], assumed to have analytic design-sensitivity gaussian noise power spectra provided by `lalsimulation`; for example, we adopted the advanced LIGO zero-detuned high-power configuration [1, 78]. To simplify our analysis, following previous studies [29, 43, 72] we also adopt a unique preferred noise realization for all sources: exactly zero. Data was always sampled at a rate significantly in excess of the Nyquist frequency, at a sampling rate that depended on the source mass.

We apply these tools to a fixed set of 998 events, whose parameters has been selected as part of a large astrophysical study first reported by Littenberg et al [43]. The specific list of sources is available on request. The set's members were randomly selected, with masses m_1, m_2 uniform in a triangle with $m_{1,2} \geq 1M_\odot$ and $m_1 + m_2 \leq 30M_\odot$; spin magnitudes $\chi_{1,2}$ uniformly distributed between $[0, 1]$; and spin directions independently and uniformly distributed on a sphere. The set's members also have random positions in the universe, subject to the restriction that no source has network amplitude $\rho \leq 5$ in two or more detectors. For each member of the above set, we carried out parameter estimation with two spins using `STT2` and with one spin via `STT4`. To further validate our results, each double-spin `STT2` analysis was performed twice. In a followup investigation, we also carried out parameter estimation on a random subset of 250 events using the one spin `STT2` and `STT4` waveform models. Finally, since each Markov-Chain Monte Carlo instance returns between 900 and 3000 independent posterior samples, we standardize our statistical treatment by randomly selecting 900 such samples from each run.

In this work, we will investigate these results by comparing the estimated posterior distribution for each of the systems' parameters at 100 Hz [42], focusing particularly on parameters in common to all models: the component masses $m_{1,2}$; the magnitude and orientation of the most significant spin; and the angle θ_{JN} between the total angular momentum and the line of sight. Models with only one significant spin lack parameters for the subdominant spin.

A. Selection bias, spin priors, and their implications for parameter estimation

We select an astrophysically-motivated population of injections, drawn from an initially uniform distribution in volume and orientation and further constrained by a signal amplitude cut. As expected with real searches for astrophysical sources, this cut strongly favors distant, nearly-face-on sources ($\theta_{JN} \simeq 0 - 0.8$). For sources with small θ_{JN}, the angle between $\vec{L}$ and the line of sight is therefore nearly constant as the binary precesses. As a result, as first described in Littenberg et al. [79], the *detection-weighted* astrophysical population therefore strongly favors sources which are barely if at all modulated by precession of $\vec{L}$. As previous studies have shown [43, 80], the spins and masses of sources in these nearly-umodulated configurations are much more difficult to constrain than sources which exhibit strong precession along the line of sight [18, 72, 80].

By construction, our astrophysical sample also includes relatively few binaries with two dynamically significant

spins. Using $\mathcal{R} \equiv |\mathbf{S}_1 - \mathbf{S}_2|/|\mathbf{S}_1 + \mathbf{S}_2|$ to measure the relative magnitude of the spin difference, of the 2000 binaries in our sample, only 109 have $1 - \mathcal{R} < 0.5$ and only 387 have $1 - \mathcal{R} < 0.8$. Therefore, by design the dynamics of each binary in our synthetic population should be well-described by the dynamics of a single significant spin – an essential assumption of the `STF2` model. Under different but still plausible astrophysical assumptions, such as significant black hole birth spin and preferentially comparable binary black hole masses, both spins will play a more significant role in typical binaries' dynamics and gravitational wave signal.

Previous studies suggested higher harmonics had little impact on parameter estimation [19, 72], once precession broke degeneracies in the signal. For this reason, both our candidate signals and parameter estimation strategies use the leading-order gravitational wave amplitude. Due to the astrophysical selection process favoring small misalignment angles between the line of sight and $\vec{J}$, higher harmonics can play a relatively more important role in breaking degeneracies. Not least because `STF2` presently lacks higher harmonics, we defer a detailed investigation of the impact of higher harmonics to a subsequent study.

III. QUANTIFYING AND UNDERSTANDING SYSTEMATIC ERRORS

A. Global indicators of systematic error

The simplest way to demonstrate and quantify the existance of the systematic error is with the best-fit amplitude, measured by the "network amplitude" $\rho \equiv \sqrt{2 \ln L}$, where L is the likelihood (see, e.g., [29, 41, 67]). In gaussian noise, this quantity is nearly normally distributed with unit mean in the presence of a signal, with mean ampitude proportional to the overlap between the signal h in the data and the best-fitting member of the signal manifold. If two models A and B are applied to the same data, known to contain a signal from B, and $\langle \cdot \rangle$ denotes an expectation value, then the ratio $\langle \rho_A \rangle \,/\, \langle \rho_B \rangle$ should be less than 1, reflecting "mismatch" between A and B associated with the inability of the best-fitting members of A to reproduce the signal in B. For Markov-Chain Monte Carlo, this expectation can be efficiently implemented as a direct sample average: if the MCMC has N sample points with network amplitudes ρ_k, then $\bar{\rho} \equiv \frac{1}{N} \sum_k \rho_k$ is nearly equal to the expectation value.

Figure 1 shows the results of this analysis, expressed as a cumulative distribution of $r \equiv \bar{\rho}_A / \bar{\rho}_B$, where B is double-spin `STT2` and A is `STF2`. This figure indicates significant differences between `STF2` and both single- and double-spin `STF2`. In other words, for a significant fraction of sources, even the best fitting members of `STF2` do not completely reproduce our zero-noise data. Our results agree with the original investigations of `STF2` [16].

Though these mismatches allow systematic bias, not all systematic bias is astrophysically significant. For example, Cho and collaborators demonstrated that in many cases higher harmonics introduce a significant systematic bias in BH-NS parameter estimation, almost exclusively isolated to astrophysically irrelevant parameters [19, 29, 72].

B. Parameter biases

As anticipated from the discussion above, the posterior distributions derived from `STF2` differ slightly from the predictions produced using other approximations. Figure 2 shows a randomly-selected example. In this figure and in general, the posterior distributions are qualitatively similar, except for small systematic offsets typically comparable to but smaller than the widths of each posterior distribution.

To quantify the difference bewteen the two distributions' means, we use a tool from classical frequentist statistics: Student's t. As reviewed in the Appendix, for each simulation k and each parameter x, we evaluate the sample mean $\bar{x}_k = N^{-1} \sum_{\alpha=1}^{N} x_{\alpha,k}$ and sample standard deviation $s_{x,k}$, defined as $s^2_{x,k} = (N-1)^{-1} \sum_\alpha (x_{\alpha,k} - \bar{x}_k)^2$, where N is the number of samples in each simulation and $\alpha = 1 \ldots N$ indexes the posterior samples. (We require an equal number of samples from each posterior to simplify our interpretation; for a more general approach, see the Appendix). For each pair of simulations of the same data k with different approximations A and B, we then evaluate

$$\mathcal{T}_{x,k}(A,B) \;\equiv\; \frac{\bar{x}_{k,A} - \bar{x}_{k,B}}{\sqrt{(s^2_{x,k,A} + s^2_{x,k,B})/N}} \qquad (1)$$

Qualitatively speaking, the value of $\mathcal{T}_{x,k}(A,B)$ measures the difference in means between the two distributions, scaled to the standard deviation divided by $\sqrt{\mathbf{1000}} \simeq 30$; values less than or comparable to 30 are therefore small compared to typical statistical errors. The distribution of $\mathcal{T}_x(A,B)$ should be nearly t-distributed with $2(N-1)$ degrees of freedom. The empirical cumulative distribution of $\mathcal{T}_x(A,B)$ can be evaluated by sorting the array of 998 $\mathcal{T}_{x,k}(A,B)$ values and compared, both to the theoretical t distribution and to the results when the models A and B changes.

Figure 3 shows the distribution of $\mathcal{T}_{\mathcal{M}_c}$ for model A being `STF2`,`STT2` double spin, `STT2` single spin with model B being fixed to `STT2` double spin. First and foremost, this model shows that when $A = B =$`STT2`, the distribution of $\mathcal{T}_{\mathcal{M}_c}$ closely follows the expected t distribution. Second, the close agreement between a single and double-spin `STT2` model strongly suggests that a single-spin model accurately reproduces most sources. Third and critically, this figure suggests that `STF2` differs substantially from both double- and even single-spin `STT2`. These differences are smaller but still significant for low-

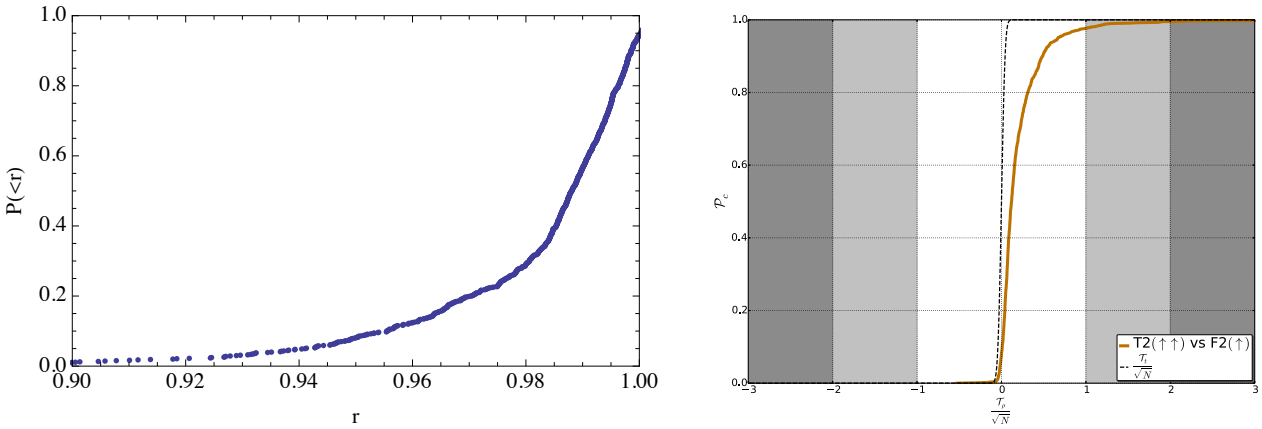

FIG. 1: **SNR ratio distribution**: *Left panel*: Ratio $r = (\bar{\rho}_B)/(\bar{\rho}_A)$ of the largest SNR recovered for `STF2` to the SNR recovered for STT2, described in Sec. III A. This distribution is significantly different from unity, with a magnitude consistent with the mismatch found in previous studies [17]. *Right panel*: Cumulative distribution of $\mathcal{T}_{\bar{\rho}}$ [Eq. (1)] divided by the square root of N; in these units, the horizontal axis measures the difference in mean in units of the standard deviation. The dashed curve is the theoretical cumulative t distribution. This plot demonstrate that there is a statistically significant difference between the average recovered SNR (ρ) by the two different models `STF2` and `STT2`.

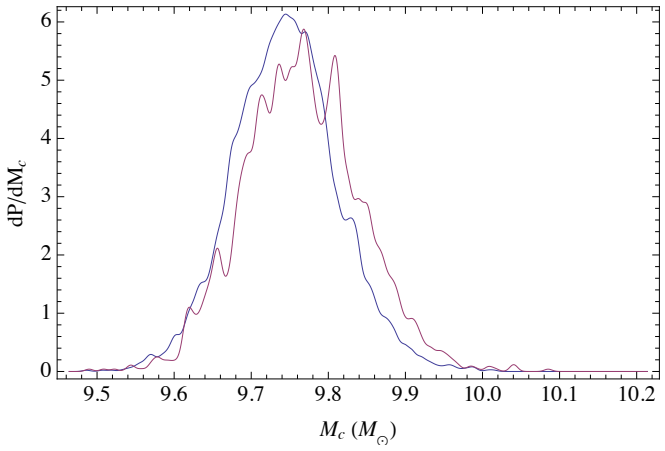

FIG. 2: **Example of posterior distributions**: Using one random example, this figure shows the one-dimensional distribution of chirp mass inferred from the `STF2` (red) and `STT2` (blue) distributions. This figure shows that though the two methods give very similar results, the two distributions are slightly offset. Our study demonstrates these small offsets occur ubiquitously, at a statistically significant level.

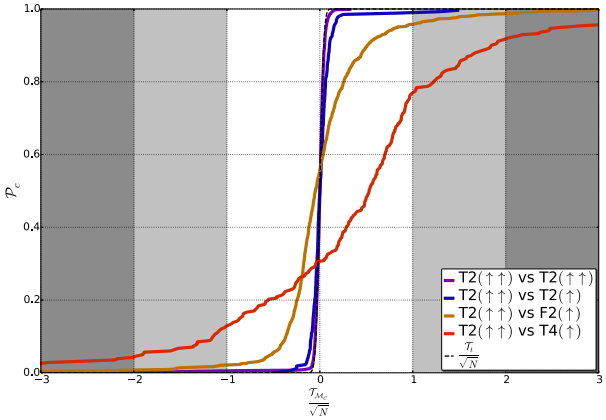

FIG. 3: **Scaled difference in mean chirp mass**: For the same data, we performed parameter estimation twice, reporting the cumulative distribution t-test scores in Eq. (1) divided by $\sqrt{N}$. In these units, the horizontal axis measures the difference in mean in units of the standard deviation. The dashed curve shows the expected result, a t distribution; the purple and blue curve shows the results when both models were `STT2` with the same or unequal numbers of nonzero spins; the orange curve shows the results when comparing (double spin) `STT2` to `STF2`; and the red curve shows a comparison between single-spin `STT4` and double-spin `STT2`.

mass and high-mass ratio signals, where `STF2` by design should be most accurate.

Though small but statistically significant differences exist between `STF2` and time-domain approximations, these differences are much smaller than the corresponding effect from systematic uncertainty in the (orbital phase of the) post-Newtonian approximation to precessing binaries. To illustrate the impact of systematic error in the post-Newtonian approximation, we perform the most conservative change possible: we adopt identical physics and identical termination conditions, but construct our gravitational wave spin and orbit evolution using the `STT4` scheme rather than the `STT2` scheme. By design, these two methods must agree up to unknown higher-order post-Newtonian (spin) terms in Taylor series for the factors of the (orbit-averaged) dv/dt or dt/dv. Figure 4 shows our results, expressed using the same strategy as described above to characterize differences between pos-

terior means. For this and all other parameters, the mean of the posterior derived from STT4 is often farther from the mean of STT2 than is the mean derived from STF2. Figure 4 provides a corresponding comparison using the intrinsic parameters η, χ_1. As with the chirp mass, these distributions usually show significant systematic differences between STT2 and STF2. Also like the chirp mass, these systematic differenes are comparable to the systematic uncertainty seen between STT2 and STT4. This analysis suggests that the systematic errors introduced by restricting to STF2 are relatively small, compared to the large systematic error currently inherent in a post-newtonian approximation to typical merging binaries.

Despite often statistically significant differences between approaches, parameter estimation employing these distinct approximations does produce consistent answers to questions less sensitive to the precise orbital phase, including extrinsic parameters. For example, Figure 4 also shows the models agree on θ_{JN}, the relative orientation between $\vec{J}$ and the line of sight. The close agreement between these results can be understood on two grounds. First, as has been described in the literature [17, 19, 81], the angle between $\vec{J}$ and the line of sight is closely related to the magnitude of precession-induced amplitude and phase modulations induced by the precession of $\vec{L}$ around $\vec{J}$. These modulations enter into the waveform as a rotation, multiplicatively, and hence approximately decouple from the orbital phase due to seperation of scales. Second and more broadly, precession is a robust effect at leading order, treated identically in both schemes. As discussed later, the ability to make robust statements about binary geometry just prior to merger may be critical in identifying or ruling out candiate short GRBs for extensive EM followup, since the directions of strongest emission should correlate with $\vec{J}$.

C. Extent of the confidence intervals

As illustrated by example in Figure 2, different waveform approximations generate posteriors that differ in mean but agree in shape, particularly width. To demonstrate this agreement quantitatively, we use another tool from classical frequentist statistics: the F statistic. As reviewed in the Appendix, the F statistic is a ratio of the sample standard deviations from two independent experiments, to assess whether the two distributions have the same width. For each pair of simulations of the same data but different waveform models A and B, and each parameter x, we evaluate

$$\mathcal{F}_{x,k}(A,B) = \frac{s^2_{x,k,A}}{s^2_{x,k,B}} \tag{2}$$

If both posterior distributions of x are gaussian with the same mean, then $\mathcal{F}_x(A,B)$ should be F_{n_1,n_2} distributed with $n_1 = n_2 = (N-1)$ degrees of freedom in the numerator and denominator.

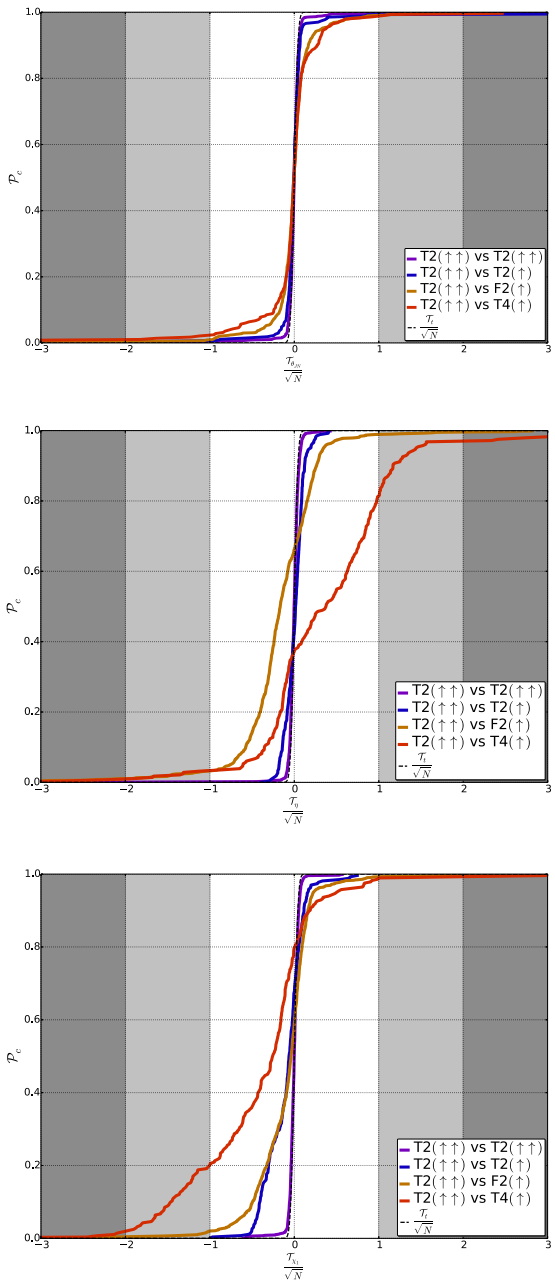

FIG. 4: **Results for additional parameters and PN approximants**: Like Figure 3, a cumulative plot of $\mathcal{T}_x$ for $x = \chi_1, \eta, \theta_{JN}$. The orange curve corresponds to comparisons between STT4 (single spin) and STT2 (double spin). These orange curve demonstrates that parameter estimation with post-Newtonian approximation schemes that adopt otherwise identical physics yield considerably different posterior distributions for *intrinsic* parameters. All approximations agree on geometric parameters like θ_{JN}.

Figure 5 shows the empirical distribution of $\mathcal{F}_{\mathcal{M}_c}(A,B)$ for $B =$ STT2 and A being STF2,STT2 double spin, STT2 single spin with model B being fixed to STT2 double spin. These results suggest that adopting a different ap-

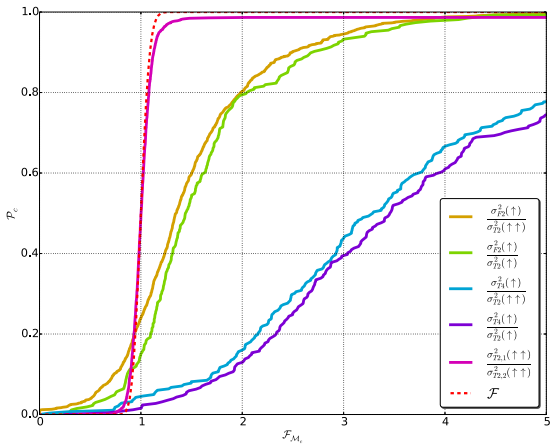

FIG. 5: **Simulation widths**: This plot shows that different approximations will predict posterior distributions of often significantly different widths. The width of the distribution is strongly impacted by nongaussian tails and is a less-robust measure of distribution similarity. These results suggest that the posteriors produced by F2 are more similar to the "true" distributions than the results from T4. [To emphasize the differences in distribution widths, we intentionally selected the parameter with the most gaussian distribution: the chirp mass. This test is less well suited to posterior distributions with significant nongaussian tails and cutoffs, notably η and $\chi_{1,2}$.]

proximation often yields a posterior distribution with a slightly different width. These differences can be ascribed in part to the different best-fitting point: the Fisher matrix and posterior distribution varies significantly across the parameter space. These differences can also be ascribed to the strong nongaussianity inherent in these distributions: the second moment used in the $\mathcal{F}$ distribution is sensitive to rare outliers. For these reasons, a detailed study of the change in extent and shape of the posterior distribution is substantially beyond the scope of this work. For the purposes of this study, we highlight the value of STF2 posterior distributions by two facts. First, that the posterior distribution obtained with either a single spin or STF2 typically differs in width by a few tens of percent from the full posterior distribution, comparable to the typical statistical error associated with the signal amplitude and only slightly greater than the sampling error ($\sqrt{1000}$) associated with our finite MCMC sample. Second, posterior distributions derived when using STT4 as a template are significantly broader than posterior distributions derived by any other means. In short, while the posterior distribution with STF2 is wider, we again assert that differences associated with using STF2 are larger than other systematic errors in our problem.

D. Unique evidence for two spins?

Lacking all degrees of freedom, a single-spin model like STF2 cannot reproduce precession-induced modulations induced by the subdominant spin. Our sample of events strongly favors face-on sources ($|\cos\theta_{JN}| \simeq 1$), with minimal modulation from the *secular* precession of $\vec{L}$ around $\vec{J}$. That said, our sample includes a significant fraction of comparable-mass BH-BH binaries with large and misaligned spins. The relative precession of the three angular momenta introduces additional modulations to $\vec{L}$ [15, 82] and hence to $h(t)$, potentially communicating more observationally-accessible information about the spins, including the subdominant spin.

The close agreement between single- and double-spin STT2 discussed above [Figure 1] strongly suggests that the subdominant spin rarely communicates observationally accessible information. To quantitatively and directly assess whether the subdominant spin's relative orientation can be measured, we examine the distribution of ϕ_{12}, the angle between $\vec{S}_1$ and $\vec{S}_2$ in the plane perpendicular to $\vec{L}$ [42]. As expected, in most cases the posterior distribution of ϕ_{12} is nearly uniform; however, a significant fraction of events ($\simeq 10\%$) have a nonuniform posterior distribution. These concentrated posterior distributions are associated with larger-than-average signal amplitude. To measure this effect quantitatively without adopting a preferred range of the periodic variable ϕ_{12}, we calculate the following quantity:

$$z \equiv \sum_{\alpha} e^{i\phi_{12,\alpha}} \tag{3a}$$

$$\sigma^2_{\phi_{12}} \equiv \sqrt{-2\ln|z|} \tag{3b}$$

In the limit that ϕ_{12} is narrowly distributed near some preferred value, ϕ_{12} is the standard deviation of the posterior. Figure 6 shows the distribution of $\sigma_{\phi_{12}}$. Based on human followup, the handful of cases with $\sigma_{\phi_{12}} \lesssim 1.8$ have posteriors that slightly or significantly favor some range of relative angles: the impact of the subdominant spin is measurable.

Not all large signal amplitudes are associated with binaries with two dynamically significant spins, nor with lines of sight that facilitate the measurement of both spins. Nonetheless, as one would expect, binaries with narrow posterior distributions of ϕ_{12} are associated with unusually large signal amplitudes.

IV. ASTROPHYSICAL IMPLICATIONS OF SYSTEMATIC ERRORS IN LOW-LATENCY FOLLOWUP WITH STF2

Having demonstrated the rapid but approximate parameter estimation enabled by STF2 introduces small but measurable systematic errors, we assess the practical astrophysical impact these errors introduce. Low-latency parameter estimation for precessing binaries fa-

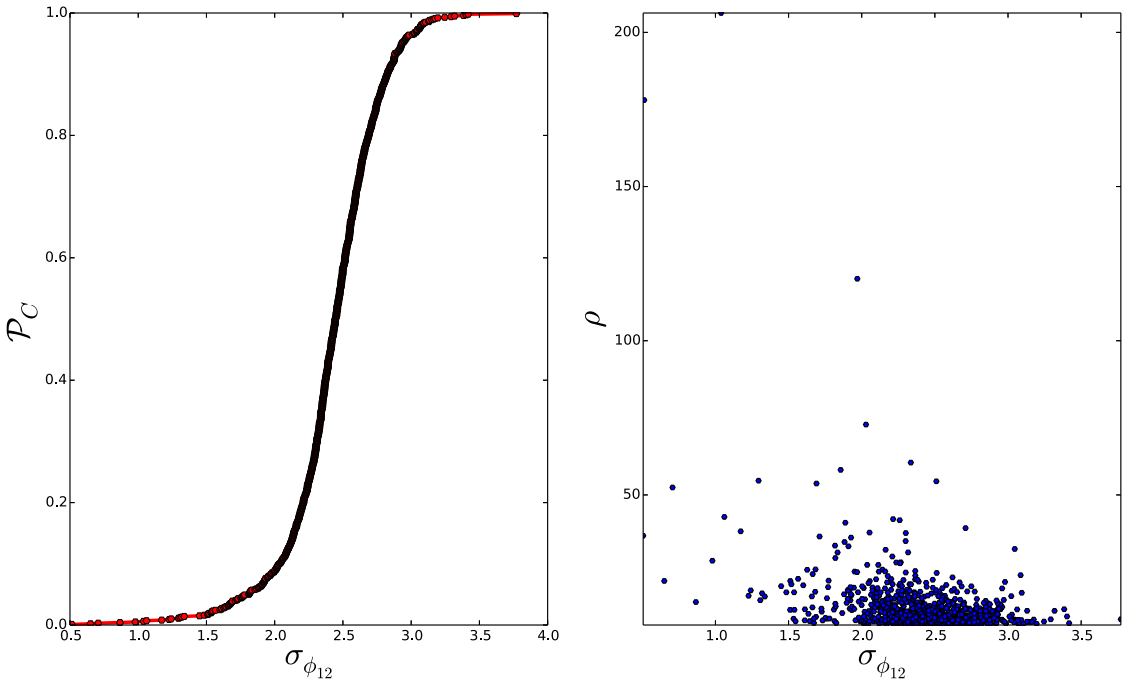

FIG. 6: **Distribution of ϕ_{12}**: Report on parmeter measurement accuracy of ϕ_{12} using two-spin `STT2`. [The parameter ϕ_{12} does not exist in an single-spin model, including `STF2`.] *Left panel*: Cumulative distribution of $\sigma_{\phi_{12}}$ [Eq. (3)]. For most cases, ϕ_{12} is approximately uniformly distributed. *Right panel*: Scatter plot of $\sigma_{\phi_{12}}$ versus ρ. Sources with high amplitude provide more information about all parameters, including ϕ_{12}. In some exceptional cases – with high amplitude, large spin, and fortuitious orientation – ϕ_{12} can thereby be tightly constrained.

cilitates electromagnetic followup by answering three critical questions about BH-NS binaries. First, is the smaller object a neutron star [43, 83–85]? Second, is the total angular momentum pointing towards the observer? Third and finally, are the BH and NS masses and spins consistent with tidal disruption prior to merger? We have used our large sample to evaluate the impact of systematic error on these three critical questions for low-latency parameter estimation.

A. Neutron star present

For the purposes of discussion, we will call a smaller object a "neutron star candidate" if its mass is less than 3 solar masses. Using the estimated posteriors derived from both models A and B (STT2 and STF2, respectively) and this condition, we evaluate posterior probabilities $P_A(NS)$ and $P_B(NS)$ that this condition is satisfied, for all simulations. As seen in Figure 7, the two probabilities largely agree, particularly when the NS mass is below $2M_\odot$. Given these results and systematic uncertainties in post-Newtonian waveforms with spin, electromagnetic followup will likely occur for all sources with $P(NS) > 0.1$. In this scenario, our calculations suggest that roughly 65-70% of all followed-up sources will actually be BH-NS binaries; the same result holds for either model.

In the transitional regime $m_2 \in [2M_\odot, 3M_\odot]$, the probabilities calculated using STF2 (red) systematically overpredict or underpredict $P(NS)$, with a probability $P(NS)$ often nearly equal to 1 below $2.5M_\odot$ while $P(NS)$ is often near 0 below $2.5M_\odot$. This property reflects the random but significant bias in the posterior distribution $p(m_2)$ predicted using STF2 compared to STT2: for each simulation, $p(m_2)_{STF2}$ and $p(m_2)_{STT2}$ have similar shape, but are offset. Depending on the offset, $P(NS) = \int^3 p(m_2)_{STF2} dm_2$ will be different, driven to be closer to 1 when the distribution should be centered well below $3M_\odot$ and driven to be closer to zero when the distribution should be centered near $3M_\odot$.

B. Total angular momentum direction

Both based on electromagnetic (jet break) and event rate arguments, short gamma ray bursts are assumed to be tightly beamed into a solid angle $\theta \lesssim 20^\circ$ [86]. For a precessing binary, we will conservatively assume the radiated energy is beamed along the total angular mo-

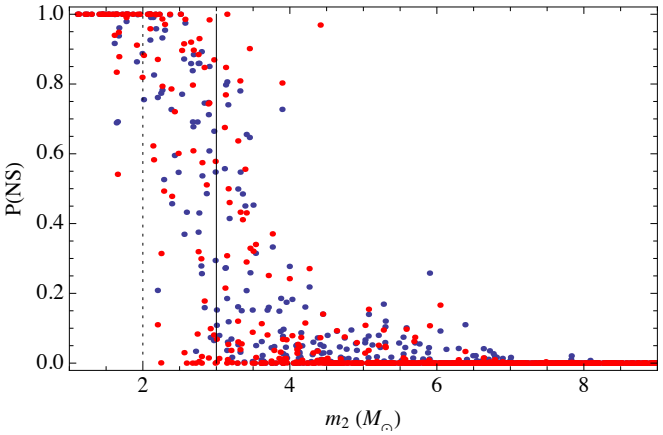

FIG. 7: **Identifying a candiate NS companion**: Plot of the posterior probability that the secondary is a neutron star versus the smaller object's mass, as estimated using STT2 (P_A; blue) and STF2(P_B; red).

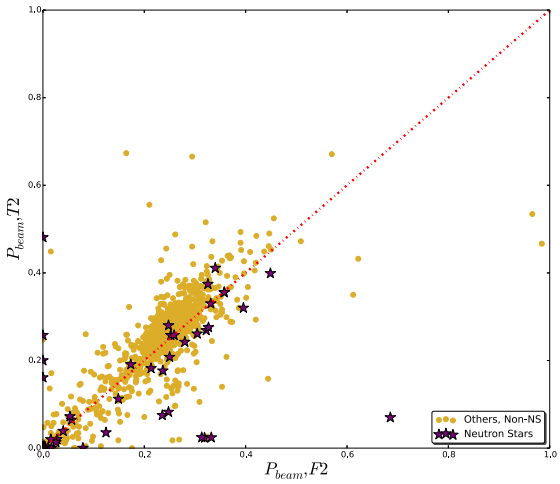

FIG. 8: **Aligned with the line of sight**: Distribution $P_A(beam)$ (red) and $P_B(beam)$ for the probability of alignment between the line of sight and the total angular momentum direction, shown both for all elements of the astrophysical sample (yellow) and for the sources with neutron star companions (purple).

mentum direction – for example, because it is powered by accretion of matter onto the final black hole, whose total angular momentum direction is nearly identical to the nearly-conserved total angular momentum direction of the binary from which it formed. Low-latency parameter estimation of gravitational waves can estimate the degree of misalignment θ_{JN} between the line of sight and the total angular momentum direction of the progenitor binary.

For the purposes of discussion, we will call a binary "aligned with the line of sight" if $\theta_{JN} < 20°$ or $\theta_{JN} > 180° - 20°$. This choice is highly arbitrary: neither theory nor observations motivate any hard cutoff, though all disfavor extreme misalignment. Using the estimated posteriors derived from both models, we define a probability P_A(beamed) and P_B(beamed) for each event that the binary is pointed towards us. Figure 8 shows the distribution of these probabilities for all events and for binaries containing a neutron star. Particularly given astrophysical systematic uncertainty in the choice of cutoff angle, these distributions strongly suggest the beaming probabilities derived from our two A and B models are nearly equivalent. The handful of cases with neutron star companions with inconsistent beaming probabilities ($P_{\rm beam}(F2) \simeq 0$ but $P_{\rm beam}(T2) > 0.1$ or vice-versa) were associated with either highly asymmetric ($m_1 \gtrsim 20M_\odot$) or nearly edge-on binaries.

C. Tidal disrupt prior to merger

To provide an unambiguous albeit approximate quantity to identify candidate tidal disruption events in binary mergers, we employ the following procecure to identify posterior samples consistent with tidal disruption of a neutron star. First, the smaller mass must lie within the range of masses allowed by our fiducial NS equation of state, here between $0.5M_\odot$ and $2.5M_\odot$. Second, the disruption process must leave behind a remnant disk with nonzero mass, as estimated using Foucart's expression [87] (his Eqs. (6,12-13)):

$$\frac{M_{disk}}{M_{NS}} = 0.288(3m_{bh}/m_{ns})^{1/3}[1-2\mathcal{C}] - 0.148(m_{bh}/m_{ns})\mathcal{C}R_{isco}(a)/m_{bh} \quad (4)$$

where $q = m_{bh}/m_{NS}$ is the binary mass ratio; $a = S_{bh}/m_{bh}^2$ is a dimensionless measure of the black hole spin; and where $\mathcal{C}(m_{ns}) = m_{ns}/R_{ns}$ is a mass- and equation-of-state dependent measure of the neutron star compactness. In Foucart's expression, shown above, $R_{isco}(a)$ is the radius of the innermost stable equitorial circular orbit of a test particle about a Kerr black hole [88]. When the black hole spin is not aligned with the orbit, we use the black hole spin magnitude a in this expression.[2] Using the fraction of all samples which satisfy this condition, we arrive at an (equation-of-state-dependent) probability that the candidate event produces a tidal disruption. Our Bayesian approach generalizes previously-reported Fisher-matrix-based [89] or search-template-based [90] estimates, incorporating

[2] We adopted an orientation-independent expression for tidal disruption probability for simplicity and to maximize the number of binaries in our sample which satisfy this condition. While physically more appropriate choice for the black hole spin would $a = \hat{L} \cdot S_{bh}/m_{bh}^2$ in this expression, extremely few

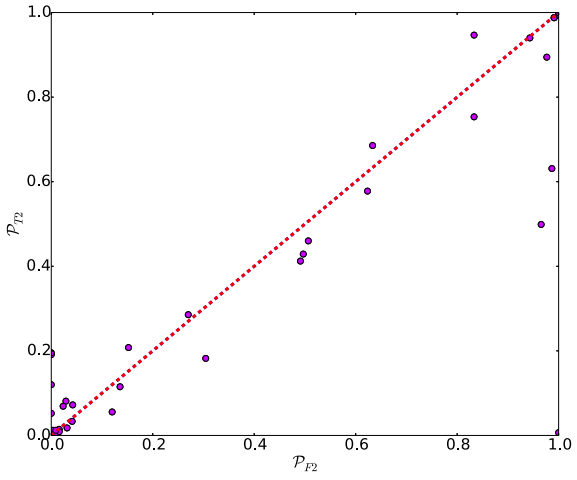

FIG. 9: **Tidal disruption probability**: Plot of the posterior tidal disruption probability evaluated via Eq. (4) using parameters drawn from `SpinTaylorF2` [P_{F2}] and using parameters drawn from `SpinTaylorT2` [P_{T2}].

state-of-the-art posterior estimates of each compact binary's parameters. Using `STF2`, these estimates can be evaluated in one to a few hours using current sampling algorithms in `lalinference_mcmc`, with further performance improvements expected before the first few detections.

Figure 9 shows a scatterplot of the tidal disruption probabilities evaluated using the posterior parameter distribution derived using `STF2` and `STT2`. As expected from Figure 7, the two approaches do not report precisely the same posterior probabilities for each event. Given systematic uncertainty in the nuclear equation of state and in post-Newtonian models for precessing binaries, however, the two models report adequately similar probabilities for targeted electromagnetic followup.

V. CONCLUSIONS

Motivated by the need for low-latency parameter estimation, in this work we estimate the systematic error introduced into posterior parameter distributions and astrophysical predictions by employing a rapid but approximate `STF2` waveform model in favor of time-domain models like `STT2` which include more physics. Though statistically significant differences exist, we demonstrate by repeated examples that these differences are small compared to the systematic post-Newtonian modeling uncertainty inherent in our present, still-approximate understanding of compact binary inspiral. Because the `STF2` model includes only one dynamically significant spin, our results also consistent with suggestions that only properties associated with a single effective spin will be observationally accessible with the first few gravitational wave detections. Finally, we show that while different approximants disagree on *intrinsic* parameters, the approximants lagely agree on extrinsic, geometric parameters like θ_{JN}. The stability of extrinsic parameter estimates will be important for targeting limited electromagnetic followup resources.

Our compelling results are in no way in conflict with prior and concurrent studies which demonstrated that, in specific moderate-amplitude cases, both black hole spins can be independently constrained in mangitude and direction. For the detection-weighted astrophysical sample adopted in these studies, most sources have only one significant spin (e.g., due the mass and spin prior) and have total angular momenta nearly along the line of sight. As described elsewhere [79], these circumstances minimize the ability of precession-induced-modulations to break degeneracies and enable both spins to be measured. Further investigations would be needed to determine if a different prior, favoring comparable-mass high-spin black hole binaries, will enable high-precision measurement of both black hole parameters.

Our analysis employed inspiral-only waveforms which lack the coalescence and ringdown signals present in real binary black hole merger signals. Their unphysical termination conditions are known to introduce convention-dependent artifacts into parameter estimation, with increasing impact as the total binary mass increases [91, 92]. A detailed discussion of waveform termination conditions is beyond the scope of this paper. That said, we anticipate that waveform termination conditions do not dominate the differences we observe. We found similar cumulative distributions when examining only low-mass sources, for which the impact of termination conditions is reduced. Particularly for the low-mass sources most likely to produce electromagnetic counterparts, we are optimistic that parameter estimation conducted using more complete models will produce similar results as our study. We understand that a similar study has begun using the `IMRPhenomP` model [93–95]. Like STF2, this frequency-domain model accounts for a single (effective) spin, both during inspiral and merger.

Though not emphasized here, our investigations suggest that STF2 can perform somewhat more reliably in some regions of the parameter space. Conversely biases between STF2 and STT2 are particularly large for some corners of parameter space; for example, detailed followup of Figure 7 suggests that STF2 may "miss" NS companions with mass $2.5M_\odot < m_2 < 3M_\odot$ when the binary is spin-dominated (at $2f_{orb} = 100\,\mathrm{Hz}$) and either edge-on or a high-amplitude source. Our sample size precludes robust investigation of how parameter biases change as a function of all source parameters (masses, spins, orientations, ...). That said, we encourage further study of biases between posterior distributions and even maximum-likelihood estimates recovered using different waveform models, as a critical ingredient to identify and diagnose the impact of systematic errors on gravitational wave astronomy.

Acknowledgments: We thank the anonymous referee for helpful feedback. TBL acknowledges NSF award PHY-1307020. BF was supported by the Enrico Fermi Institute at the University of Chicago as a McCormick Fellow. Computational resources were provided by the Northwestern University Grail cluster through NSF MRI award PHY-1126812.

Appendix A: Review of statistics

The Student-t distribution arises from the distribution of the ratio of $t = x/\sqrt{z/n}$ where x is normally distributed with zero mean and unit variance and where z is an independent χ^2-distributed variable with n degrees of freedom. After some algebra, the probability distribution function (PDF) of t is

$$p(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n}\Gamma(n/2)} \frac{1}{(1+\frac{t^2}{n})^{(n+1)/2}} \tag{A1}$$

The Student-t distribution is widely used in statistics to compare whether two populations have the same mean, given a priori equal variance. In the simplest case where two sets of measurements $x_1 \dots x_{N_x}$ and $y_1 \dots y_{N_y}$ have the same sample size, and $s_x, \bar{x}$ are defined as

$$\bar{x} \equiv \frac{1}{N_x} \sum_k x_k \tag{A2}$$

$$s_x^2 \equiv \frac{1}{N_x} \sum_k (x_k - \bar{x})^2 \tag{A3}$$

and similarly for $s_y, \bar{y}$, then the following quantity is t distributed with $N_1 + N_2 - 2$ degrees of freedom:

$$\mathcal{T} = \frac{\bar{x} - \bar{y}}{\sqrt{((N_x-1)s_x^2 + (N_y-1)s_y^2)\frac{N_x^{-1}+N_y^{-1}}{N_x+N_y-2}}} \tag{A4}$$

Note that a specific value of $\mathcal{T}$ corresponds to a difference in means between x, y by of order $\sigma/\sqrt{N}$. In our case, with $N \simeq 1000$

The F_{n_1,n_2} distribution arises from the distribution of the ratio $F = \frac{y_1/n_1}{y_2/n_2}$ of two independent χ^2-distributed random variables y_1, y_2 with n_1 and n_2 degrees of freedom, respectively. The probability distribution of F is

$$p(F) = \frac{\Gamma(\frac{n_1+n_2}{2})(n_1/n_2)^{n_1/2} F^{\frac{n_1}{2}-1}}{\Gamma(n_1/2)\Gamma(n_2/2)[1+(n_1F/n_2)]^{(n_1+n_2)/2}} \tag{A5}$$

The F distribution is most widely used in frequentist hypothesis testing (e.g., comparing the residuals after a fit to an independent estimate of the sample variance). For the purposes of this study, however, we point out that the unbiased estimate of the standard deviation [s_x^2; Eq. (A3)] is proportional to a χ^2-distributed random variable. As a result, the suitably-weighted ratio

$$F = s_x^2/s_y^2 \tag{A6}$$

can be used in a standard two-sample F test with $n-1, n-1$ degrees of freedom to assess whether the distributions of x and y have the same variance. In practice, because the second moment and therefore the F-test is very sensitive to non-normality, we construct an *empirical* distribution of F, based on two samples known to be drawn from the same distribution: independent repetitions of the `STT2` analysis.

[1] Aasi and The LIGO Scientific Collaboration, Classical and Quantum Gravity **32**, 074001 (2015), 1411.4547.

[2] F. Acernese et al. (VIRGO), Class.Quant.Grav. **32**, 024001 (2015), 1408.3978.

[3] A. Buonanno, Y. Chen, and M. Vallisneri, Phys. Rev. D **67**, 104025 (2003), URL `http://lanl.arxiv.org/pdf/gr-qc/0211087`.

[4] A. Buonanno, Y. Chen, Y. Pan, and M. Vallisneri, Phys. Rev. D **70**, 104003 (2004).

[5] T. Damour, A. Gopakumar, and B. R. Iyer, Phys. Rev. D **70**, 064028 (2004).

[6] Y. Pan, A. Buonanno, Y. Chen, and M. Vallisneri, Phys. Rev. D **69**, 104017 (2004), URL `http://xxx.lanl.gov/abs/gr-qc/0310034`.

[7] C. Königsdörffer and A. Gopakumar, Phys. Rev. D **71**, 024039 (2005).

[8] A. Buonanno, Y. Chen, Y. Pan, H. Tagoshi, and M. Vallisneri, Phys. Rev. D **72**, 084027 (2005).

[9] C. Königsdörffer and A. Gopakumar, Phys. Rev. D **73**, 124012 (2006).

[10] M. Tessmer and A. Gopakumar, MNRAS **374**, 721 (2007).

[11] M. Hannam, S. Husa, B. Brügmann, and A. Gopakumar, Phys. Rev. D **78**, 104007 (2008).

[12] I. Hinder, F. Herrmann, P. Laguna, and D. Shoemaker, ArXiv e-prints (2008).

[13] K. G. Arun, A. Buonanno, G. Faye, and E. Ochsner, Phys. Rev. D **79**, 104023 (2009).

[14] A. Buonanno, B. R. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, Phys. Rev. D **80**, 084043 (2009).

[15] T. A. Apostolatos, C. Cutler, G. J. Sussman, and K. S. Thorne, Phys. Rev. D **49**, 6274 (1994).

[16] D. A. Brown, A. Lundgren, and R. O'Shaughnessy, Phys. Rev. D **86**, 064020 (2012), 1203.6060, URL `http://arxiv.org/abs/1203.6060`.

[17] A. Lundgren and R. O'Shaughnessy, PRD in press (arXiv:1304.3332) (2013), URL `http://xxx.lanl.gov/abs/arXiv:1304.3332`.

[18] J. Aasi et al (The LIGO Scientific Collaboration and the Virgo Collaboration), Phys. Rev. D **88**, 062001 (2013), URL http://xxx.lanl.gov/abs/arXiv:1304.1775.

[19] H. Cho, E. Ochsner, R. O'Shaughnessy, C. Kim, and C. Lee, Submitted to PRD (arXiv:1209.4494) (2012), URL http://xxx.lanl.gov/abs/arXiv:1209.4494.

[20] H.-T. Janka, MNRAS **434**, 1355 (2013), 1306.0007.

[21] J. Nordhaus, T. D. Brandt, A. Burrows, and A. Almgren, MNRAS **423**, 1805 (2012), 1112.3342.

[22] D. Gerosa, M. Kesden, E. Berti, R. O'Shaughnessy, and U. Sperhake, Phys. Rev. D **87**, 104028 (2013), 1302.4442.

[23] H. van Eerten, W. Zhang, and A. MacFadyen, Astrophys. J. **722**, 235 (2010), 1006.5125.

[24] H. J. van Eerten and A. I. MacFadyen, ApJL **733**, L37 (2011), 1102.4571.

[25] H. van Eerten and A. MacFadyen, Astrophys. J. **767**, 141 (2013), 1209.1985.

[26] N. Gehrels, E. Ramirez-Ruiz, and D. B. Fox, ARAA **47**, 567 (2009), 0909.1531.

[27] E. Poisson and C. M. Will, Phys. Rev. D **52**, 848 (1995), arXiv:gr-qc/9502040.

[28] E. Baird, S. Fairhurst, M. Hannam, and P. Murphy, Phys. Rev. D **87**, 024035 (2013), 1211.0546.

[29] R. O'Shaughnessy, B. Farr, E. Ochsner, C. Cho, H.S. Kim, and C. Lee, Submitted to PRD (arXiv:1308.4704) (2013), URL http://arxiv.org/abs/1308.4704.

[30] R. N. Lang and S. A. Hughes, Phys. Rev. D **74**, 122001 (2006).

[31] A. Klein, P. Jetzer, and M. Sereno, Phys. Rev. D **80**, 064027 (2009).

[32] R. N. Lang, S. A. Hughes, and N. J. Cornish, Phys. Rev. D **84**, 022002 (2011).

[33] C. Van Den Broeck and A. S. Sengupta, Classical and Quantum Gravity **24**, 155 (2007).

[34] W. Del Pozzo, J. Veitch, and A. Vecchio, Phys. Rev. D **83**, 082002 (2011), 1101.1391.

[35] N. Cornish, L. Sampson, N. Yunes, and F. Pretorius, Phys. Rev. D **84**, 062003 (2011), 1105.2088.

[36] 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), 1110.0530.

[37] J. Veitch, I. Mandel, B. Aylott, B. Farr, V. Raymond, C. Rodriguez, M. van der Sluys, V. Kalogera, and A. Vecchio, Phys. Rev. D **85**, 104045 (2012), 1201.1195.

[38] S. Nissanke, J. Sievers, N. Dalal, and D. Holz, Astrophys. J. **739**, 99 (2011), 1105.3184.

[39] J. Veitch, I. Mandel, B. Aylott, B. Farr, V. Raymond, C. Rodriguez, M. van der Sluys, V. Kalogera, and A. Vecchio, Phys. Rev. D **85**, 104045 (2012), 1201.1195.

[40] V. Raymond, Ph.D. thesis, Northwestern University (2012), URL https://gwic.ligo.org/thesisprize/2012/raymond-thesis.pdf.

[41] J. Veitch, V. Raymond, B. Farr, W. M. Farr, P. Graff, S. Vitale, B. Aylott, K. Blackburn, N. Christensen, M. Coughlin, et al., Phys. Rev. D **91**, 042003 (2015), URL http://link.aps.org/doi/10.1103/PhysRevD.91.042003.

[42] B. Farr, E. Ochsner, W. M. Farr, and R. O'Shaughnessy, Phys. Rev. D **90**, 024018 (2014), 1404.7070, URL http://xxx.lanl.gov/abs/arXiv:1404.7070.

[43] T. B. Littenberg, B. Farr, S. Coughlin, V. Kalogera, and D. E. Holz, ArXiv e-prints (2015), 1503.03179.

[44] LIGO Scientific Collaboration, Virgo Collaboration, J. Aasi, J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. Abernathy, T. Accadia, F. Acernese, et al., (arXiv:1304.0670, available at http://arxiv.org/1304.0670) (2013), 1304.0670.

[45] B. D. Metzger and E. Berger, Astrophys. J. **746**, 48 (2012), 1108.6056.

[46] D. Grossman, O. Korobkin, S. Rosswog, and T. Piran, MNRAS **439**, 757 (2014), 1307.2943.

[47] B. D. Metzger and G. C. Bower, MNRAS **437**, 1821 (2014), 1310.4506.

[48] M. Tanaka, K. Hotokezaka, K. Kyutoku, S. Wanajo, K. Kiuchi, Y. Sekiguchi, and M. Shibata, Astrophys. J. **780**, 31 (2014), 1310.2774.

[49] J. Barnes and D. Kasen, Astrophys. J. **775**, 18 (2013), 1303.5787.

[50] K. Kyutoku, K. Ioka, and M. Shibata, MNRAS **437**, L6 (2014), 1209.5747.

[51] L. P. Singer, L. R. Price, B. Farr, A. L. Urban, C. Pankow, S. Vitale, J. Veitch, W. M. Farr, C. Hanna, K. Cannon, et al., Astrophys. J. **795**, 105 (2014), 1404.5623.

[52] D. Buskulic, Virgo Collaboration, and LIGO Scientific Collaboration, Classical and Quantum Gravity **27**, 194013 (2010).

[53] J. Abadie et al. (The LIGO Scientific Collaboration and the Virgo Collaboration), A&A **541** (2012), URL http://xxx.lanl.gov/abs/arXiv:1112.6005.

[54] J. Kanner, T. L. Huard, S. Márka, D. C. Murphy, J. Piscionere, M. Reed, and P. Shawhan, Classical and Quantum Gravity **25**, 184034 (2008), 0803.0312.

[55] LIGO Scientific Collaboration, Virgo Collaboration, J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. Abernathy, T. Accadia, F. Acernese, C. Adams, et al., A&A **539**, A124 (2012), 1109.3498.

[56] T. Sidery, B. Aylott, N. Christensen, B. Farr, W. Farr, F. Feroz, J. Gair, K. Grover, P. Graff, C. Hanna, et al., Phys. Rev. D **89**, 084060 (2014), 1312.6013.

[57] M. M. Kasliwal and S. Nissanke, ApJL **789**, L5 (2014), 1309.1554.

[58] K. Chatziioannou, N. Cornish, A. Klein, and N. Yunes, Phys. Rev. D **89**, 104023 (2014), 1404.3180, URL http://xxx.lanl.gov/abs/arXiv:1307.4418.

[59] A. Klein, N. Cornish, and N. Yunes, ArXiv e-prints (2013), 1305.1932.

[60] K. Chatziioannou, N. Cornish, A. Klein, and N. Yunes, ApJL **798**, L17 (2015), 1402.3581.

[61] A. Klein, N. Cornish, and N. Yunes, Phys. Rev. D **90**, 124029 (2014), 1408.5158.

[62] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, (Available at http://arxiv.org/1308.3271) (2013), 1308.3271.

[63] R. J. E. Smith, K. Cannon, C. Hanna, D. Keppel, and I. Mandel, Phys. Rev. D **87**, 122002 (2013), 1211.1254.

[64] K. Cannon, J. D. Emberson, C. Hanna, D. Keppel, and H. P. Pfeiffer, Phys. Rev. D **87**, 044008 (2013), 1211.7095.

[65] P. Canizares, S. E. Field, J. R. Gair, and M. Tiglio, Phys. Rev. D **87**, 124005 (2013), 1304.0462, URL http://xxx.lanl.gov/abs/arXiv:1304.0462.

[66] M. Pürrer, Classical and Quantum Gravity **31**, 195010 (2014), 1402.4146.

[67] C. Pankow, P. Brady, E. Ochsner, and R. O'Shaughnessy, Submitted to PRD

(arXiv:1502.04370) (2015), URL http://xxx.lanl.gov/abs/arXiv:1502.04370.

[68] P. Canizares, S. E. Field, J. Gair, V. Raymond, R. Smith, and M. Tiglio, Physical Review Letters **114**, 071104 (2015), 1404.6284.

[69] S. E. Field, C. R. Galley, J. S. Hesthaven, J. Kaye, and M. Tiglio, Phys. Rev. X **4**, 031006 (2014), URL http://link.aps.org/doi/10.1103/PhysRevX.4.031006.

[70] Y. Pan, A. Buonanno, L. T. Buchman, T. Chu, L. E. Kidder, H. P. Pfeiffer, and M. A. Scheel, Phys. Rev. D **81**, 084041 (2010), 0912.3466.

[71] A. Taracchini, Y. Pan, A. Buonanno, E. Barausse, M. Boyle, T. Chu, G. Lovelace, H. P. Pfeiffer, and M. A. Scheel, Phys. Rev. D **86**, 024011 (2012), 1202.0790.

[72] R. O'Shaughnessy, B. Farr, E. Ochsner, H.-S. Cho, V. Raymond, C. Kim, and C.-H. Lee, Phys. Rev. D **89**, 102005 (2014), URL http://link.aps.org/doi/10.1103/PhysRevD.89.102005.

[73] L. Wade, J. D. E. Creighton, E. Ochsner, B. D. Lackey, B. F. Farr, T. B. Littenberg, and V. Raymond, Phys. Rev. D **89**, 103012 (2014), 1402.5156.

[74] B. D. Lackey and L. Wade, Phys. Rev. D **91**, 043002 (2015), 1410.8866.

[75] C. P. L. Berry, I. Mandel, H. Middleton, L. P. Singer, A. L. Urban, A. Vecchio, S. Vitale, K. Cannon, B. Farr, W. M. Farr, et al., Astrophys. J. **804**, 114 (2015), 1411.6934.

[76] I. W. Harry, A. H. Nitz, D. A. Brown, A. P. Lundgren, E. Ochsner, and D. Keppel, Phys. Rev. D **89**, 024010 (2014), 1307.3562.

[77] A. H. Nitz, A. Lundgren, D. A. Brown, E. Ochsner, D. Keppel, and I. W. Harry, Phys. Rev. D **88**, 124039 (2013), 1307.1757.

[78] LIGO Scientific Collaboration (2009), URL https://dcc.ligo.org/cgi-bin/DocDB/ShowDocument?docid=m060056.

[79] T. B. Littenberg, B. Farr, S. Coughlin, V. Kalogera, and D. E. Holz, (in preparation) (????).

[80] D. Trifiro, R. O'Shaughnessy, G. Gerosa, M. Kesden, B. Berti, T. Littenberg, and U. Sperhake, (in preparation) (????).

[81] D. A. Brown, A. Lundgren, and R. O'Shaughnessy, Phys. Rev. D **86**, 064020 (2012), 1203.6060.

[82] M. Kesden, D. Gerosa, R. O'Shaughnessy, E. Berti, and U. Sperhake, Physical Review Letters **114**, 081103 (2015), 1411.0674.

[83] M. Hannam, D. A. Brown, S. Fairhurst, C. L. Fryer, and I. W. Harry, ApJL **766**, L14 (2013), 1301.5616.

[84] I. Mandel, C.-J. Haster, M. Dominik, and K. Belczynski, MNRAS **450**, L85 (2015), 1503.03172.

[85] R. O'Shaughnessy, B. Farr, E. Ochsner, H.-S. Cho, V. Raymond, C. Kim, and C.-H. Lee, Phys. Rev. D **89**, 102005 (2014), 1403.0544.

[86] E. Berger, ARAA **52**, 43 (2014), 1311.2603.

[87] F. Foucart, Phys. Rev. D **86**, 124007 (2012), 1207.6304.

[88] J. M. Bardeen, W. H. Press, and S. A. Teukolsky, Astrophys. J. **178**, 347 (1972).

[89] A. Maselli and V. Ferrari, Phys. Rev. D **89**, 064056 (2014), 1312.5133.

[90] F. Pannarale and F. Ohme, ApJL **791**, L7 (2014), 1406.6057.

[91] H.-S. Cho and C.-H. Lee, Classical and Quantum Gravity **31**, 235009 (2014), 1403.4681.

[92] I. Mandel, C. P. L. Berry, F. Ohme, S. Fairhurst, and W. M. Farr, Classical and Quantum Gravity **31**, 155005 (2014), 1404.2382.

[93] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, Physical Review Letters **113**, 151101 (2014), 1308.3271.

[94] M. Pürrer, M. Hannam, P. Ajith, and S. Husa, Phys. Rev. D **88**, 064007 (2013), 1306.2320.

[95] P. Schmidt, F. Ohme, and M. Hannam, Phys. Rev. D **91**, 024043 (2015), 1408.1810.