# BayesWave++: Implementing a Bayesian Inference Package for Gravitational Waves in C++

Maven Holst
Katerina Chatziioannou, Sophie Hourihane

September 2024

## Abstract

Gravitational waves (GWs) are minute ripples in spacetime detectable only by a global network of interferometer observatories that sense length changes on the scale of a proton. Despite the careful engineering of these observatories, GW signals are notoriously difficult to extract and characterize, often mimicked or masked by the significant detector noise caused by terrestrial events, meteorological interference, and instrument glitches. BayesWave++ is an improved C++ implementation of BayesWave, a Bayesian inference software written to enable signal modeling in this noisy environment. Using a transdimensional Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm, BayesWave++ fits sine-exponential wavelets to burst signals of both GW events and instrument glitches, and specific waveform templates from compact binary coalescence (CBC) simulations to GW events. Here, pipeline support for automated analysis of real GW data is added to BayesWave++, and a thorough review is conducted of the GW/glitch wavelet and CBC models to ensure proper algorithm performance. From this review, new issues throughout the BayesWave++ software are identified and corrected, and bounds are determined on the subset of parameter space in which the algorithm works well.

## Introduction

Gravitational waves are released by some of the most violent events in the cosmos; since their first detection in 2015 (GW150914, a binary black hole merger event [2]), they have become an increasingly useful way to learn about the universe.

A global network of interferometer observatories detect the minute strain signals caused by gravitational waves as they pass through Earth, and coherence between these multiple detectors helps eliminate detector-specific instrumental and environmental noise [2].

Weak compared to the significant noise due to terrestrial events, meteorological interference, and instrument glitches, gravitational-wave signals are nonetheless notoriously difficult to extract and characterize. To address this, BayesWave was developed to enable signal modeling in this noisy environment, using a continuous basis of Morlet-Gabor wavelets to fit burst signals due to both gravitational-wave events and instrument glitches. A transdimensional Reversible Jump Markov Chain

Monte Carlo (RJMCMC) algorithm is used to determine the placement and features of these wavelets [1]. As compared to a typical Markov Chain Monte Carlo (MCMC) algorithm, the RJMCMC used in BayesWave enables comparison between posterior spaces of varying dimensionality: this allows for any number of wavelets to be fit to the signal, enabling a model to be fit to any arbitrary burst waveform, including glitches and gravitational waves from sources that are not yet well-modeled in general relativity [1]. As models for these unknown sources are proposed, they can then be overlaid with the wavelet models constructed by BayesWave, allowing for such models to be checked against observation and for information about the source signal to be extracted.

For those sources that are well-modeled in general relativity (as is the case for GW150914 [2]), BayesWave also allows for the direct fitting of simpler, fixed-dimensionality waveform templates as generated from numerical relativity simulations. Given the parameters of this model fit, interesting features of the signal source can be extracted: for combact binary coalescences (CBCs) like GW150914, this includes its location in the sky, the inclination of its orbit relative to Earth, and the masses of the constituent bodies.

BayesWave++ is a more recent effort to rewrite old BayesWave, originally written in C, using C++, making a number of improvements along the way in functionality, documentation, modularity, usability and post-processing. In this project, some new functionality to BayesWave++ was added, including the development of an automated pipeline for downloading real gravitational-wave strain data from LIGO servers and inputting it into BayesWave++ for analysis. Additionally, metrics of MCMC performance and convergence, including prior testing and P-P plots, were conducted, enabling a robust evaluation of BayesWave++ and allowing for any existing bugs to be identified and resolved.

# Methods

For a full description of the archetypal Markov Chain Monte Carlo method, see Numerical Recipes § 15.8 [3].

Two methods of testing are used to evaluate BayesWave++ performance in this review: prior testing and P-P plots. During prior testing, the likelihood evaluation step in the MCMC algorithm is skipped, meaning that the best-fit model parameters are drawn directly from their prior distributions, without any further modification (and without using any input signal information). The parameter posterior distributions as seen in corner plots generated from the MCMC results should then directly match the corresponding prior distributions if there is no bias or other problem in the sampling step. Thus, prior testing tests the health of the MCMC samplers in the code, and a pass is indicated if the posterior distributions match the prior distributions.

During P-P testing, signals are injected into the BayesWave++ model fitting step with parameters drawn directly from their prior distributions, and the code is then run on these injections as it would any other signal. Over many such prior-draw runs (we use 400 runs per P-P plot), plotting for each run the percentile of each injected parameter in the respective posterior distribution vs. the percentile of this percentile across all runs should result in a straight line of unit slope (within some error due to statistical uncertainty). Deviations from a straight line indicate that the best-fit posterior distributions overall do not match the prior distributions of the injections, and thus either that there is some mathematical error in the MCMC algorithm, or that the algorithm is failing to converge. All fit parameters remaining within statistical uncertainty of the $y = x$ line in a P-P plot constitutes a pass, indicating successful convergence for the given parameter range specified.

# Results and Analysis

Prior testing was conducted for all three major BayesWave++ model types: glitch wavelet, GW wavelet, and GW CBC. These tests indicate that the MCMC samplers properly recreate the prior distributions for each parameter when likelihood is ignored; this constitutes a pass across the board for the prior tests (see Appendix A: Example Prior Test Results).

Additionally, P-P plots for the three model types were generated and analyzed across a wide portion of the available parameter space. In the case of the GW/glitch wavelet models, this includes varying the shape and center of the wavelet amplitude prior distributions, spanning different signal-to-noise ratio (SNR) ranges, varying the dimension range, and fixing/unfixing the sky location of the injected signals. For the GW CBC model, this includes varying the CBC approximant model (precessing or non-precessing, etc.). For the most part, these indicate expected and proper MCMC function over all the parameter values tested.

The notable exception is a substantial deviation from the expected P-P plot $y = x$ line in unfixed multidimensional wavelet runs using an SNR-dependent prior for wavelet amplitude: the empirical distributions for model dimension were been observed to exhibit marked shifts from the expected prior distributions from which the signals injected into these runs were drawn. In particular, the extracted model dimension distribution routinely had a lower mean and variance than the injected model dimension (prior) distribution, which is expected to be uniform (see Appendix B: Example Glitch Model P-P Plots and Appendix C: Example GW Wavelet Model P-P Plots).

Upon further discussion and analysis of this issue, it was determined that certain approximations made in the proposal step of the BayesWave++ MCMC algorithm assume high SNR; for runs that include injected signals of SNR on the order of 1 or below, this approximation is clearly invalid. Further, since BayesWave++ attempts to use the most parsimonious fit to the signal [1], it would make sense for the algorithm to have a tendency to omit lower-SNR wavelets from the best-fit model if they are largely eclipsed by a higher-SNR signal. It was confirmed that runs that ignore SNR dependence in the wavelet amplitude prior, and those that only include injected wavelets with SNR above around 5 (and in some cases as high as 7; see Appendix C: Example GW Wavelet Model P-P Plots), tend to succeed in passing the P-P tests better than those in which the SNR range from 0.1 to 5 is included. Further P-P plot analysis of multidimensional runs with SNR exclusively bounded from 0.1 to 5 was conducted to confirm that the MCMC algorithm does indeed behave poorly in this regime, at least in terms of the dimensionality of the reconstructed signal. All other parameters, including the extrinsic GW parameters common to both the GW wavelet and CBC models, pass the P-P test even in this SNR range.

During the review process, a bug was discovered in the GW extrinsic sampler (used both in the GW wavelet and CBC models) that resulted in extracted parameter distributions for many extrinsic GW parameters with lower variance than the injected parameter (prior) distributions. P-P testing for the GW models confirmed this bug, with clear deviations in the P-P plots from the expected $y = x$ line. Following a fix made to the algorithm, P-P testing also confirmed successful implementation thereof, with the extracted extrinsic parameter distributions behaving as expected (see Figure 1).

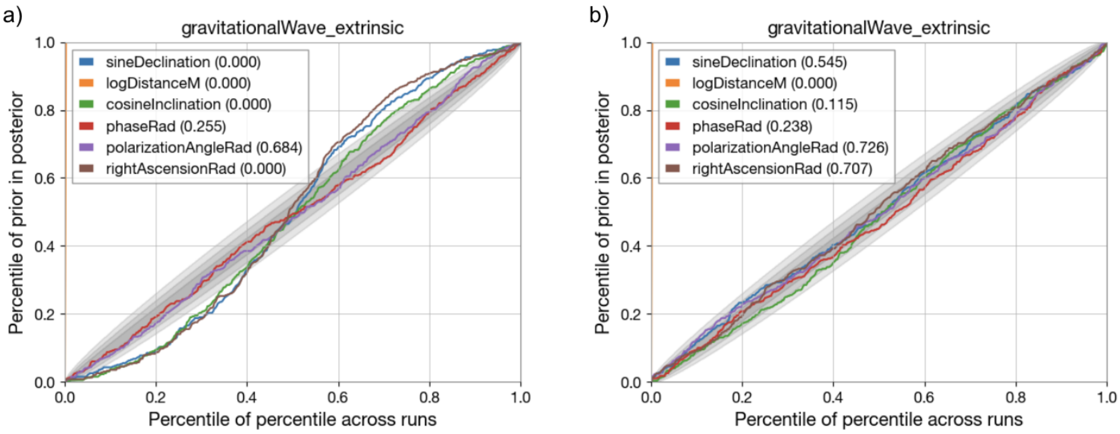



Figure 1: Pre- and Post-Fix GW Extrinsic P-P Plots. Panel (a): Pre-fix GW extrinsic parameter P-P plot for wavelet dimension 1-5, uniform amplitude prior, elliptical polarization. Clear deviations from a straight line are present for declination, inclination, and right ascension parameters indicating smaller variance than expected. Panel (b): Post-fix GW extrinsic parameter P-P plot for wavelet dimension 1-5, uniform amplitude prior, elliptical polarization. Problematic parameters now correctly follow the expected $y = x$ line, within uncertainty. Similar results are seen for generic polarization.

# Conclusions and Discussion

Given the SNR-dependence of the multidimensional wavelet model P-P test results discussed in the section above, a major conclusion of this review is that in the case of signals best represented by a multidimensional wavelet model, BayesWave++ works optimally in the mid-and-high-SNR regime. This is an acceptable working parameter subspace, since candidates for gravitational-wave events in currently-active LIGO event detection pipelines typically have a minimum SNR threshold [2], but still bears further investigation for possible avenues for improvement. Dimension-independent parameters, such as the extrinsic GW parameters (distance, inclination, etc.) seem to be unaffected, however; dimension is the primary suspect parameter due to the preference BayesWave++ gives to less complex best-fit models.

Other than this SNR bound, all other reasonable combinations of parameters tested with prior testing and P-P plots during this review successfully pass, indicating that BayesWave++ functions well across nearly the entire parameter space available in terms of metrics of convergence and sampling bias.

Additionally, given that the GW CBC model was still in active development during this review process, the ongoing review results proved very helpful in identifying bugs and confirming proper implementation of the in-progress CBC model: the results from Figure 1 are a prime example thereof. In this vein, we suggest that any further models or other developments added to the BayesWave++ codebase undergo a similar concurrent review process to catch and correct bugs during implementation.

# Future Work

Plots and conclusions for review results will be gathered on the BayesWave++ GitLab page (and eventually posted on a LIGO wiki page). Testing of the majority of the parameter space of interest has been conducted and indicates a pass across the board; as aforementioned, however, further investigation into the SNR lower bound requirement for multidimensional wavelet runs should be conducted both to determine the true source of the test failures, and to find any reasonable means by which to reduce the lower bound by as much as possible.

Additionally, as the CBC model continues to be more completely implemented in BayesWave++, using increasingly-complex numerical relativity waveform approximants, more parameters will become available; this will merit further prior and P-P testing to confirm proper code functionality.

# References

[1] Neil J. Cornish and Tyson B. Littenberg. BayesWave: Bayesian Inference for Gravitational Wave Bursts and Instrument Glitches. *Class. Quant. Grav.*, 32, 2015.

[2] B.P. Abbott et al. A guide to LIGO-Virgo detector noise and extraction of transient gravitational-wave signals. *Class. Quant. Grav.*, 37, 2020.

[3] William H. et al Press. Numerical Recipes: The Art of Scientific Computing (Third Edition). *Cambridge University Press*, 2007.

# Acknowledgements

Many thanks to Dr. Katerina Chatziioannou and Sophie Hourihane for providing indispensable mentorship and guidance throughout this project; without their help, this project would not have been possible.

Funding for this work was supported by the National Science Foundation Research Experiences for Undergraduates (NSF REU) and the Caltech Student-Faculty Program (Caltech SFP).

# Appendix A: Example Prior Test Results

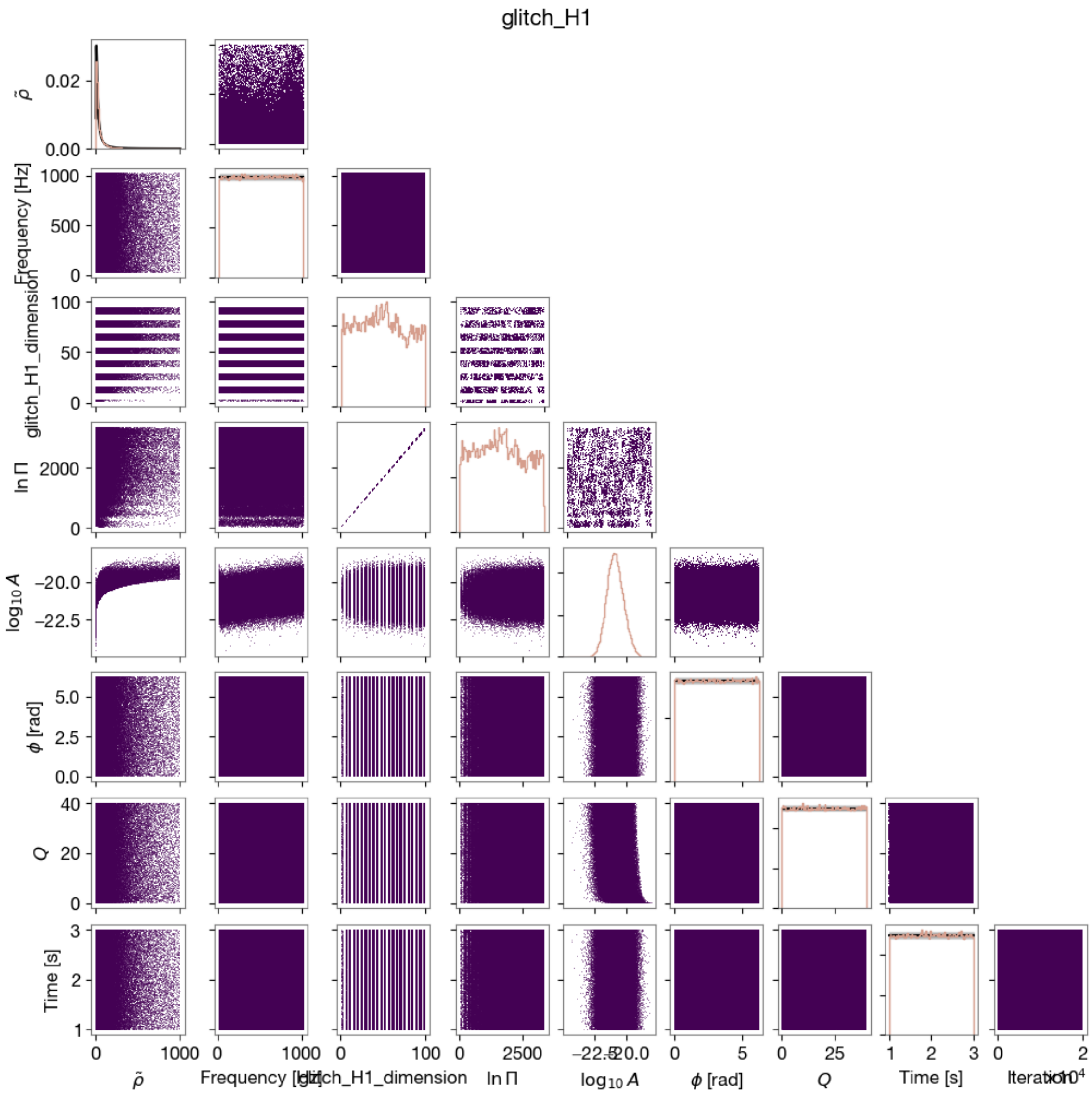



Figure 2: Example Prior Test Results for Glitch Model. All of the 1-D posterior distributions match their respective priors: uniform for most parameters, and singly-peaked for log wavelet amplitude. Likelihood $\overline{\rho}$ is a function of all parameters. Default parameters used, segment length 4.0 seconds, sampling rate 2048 Hz.

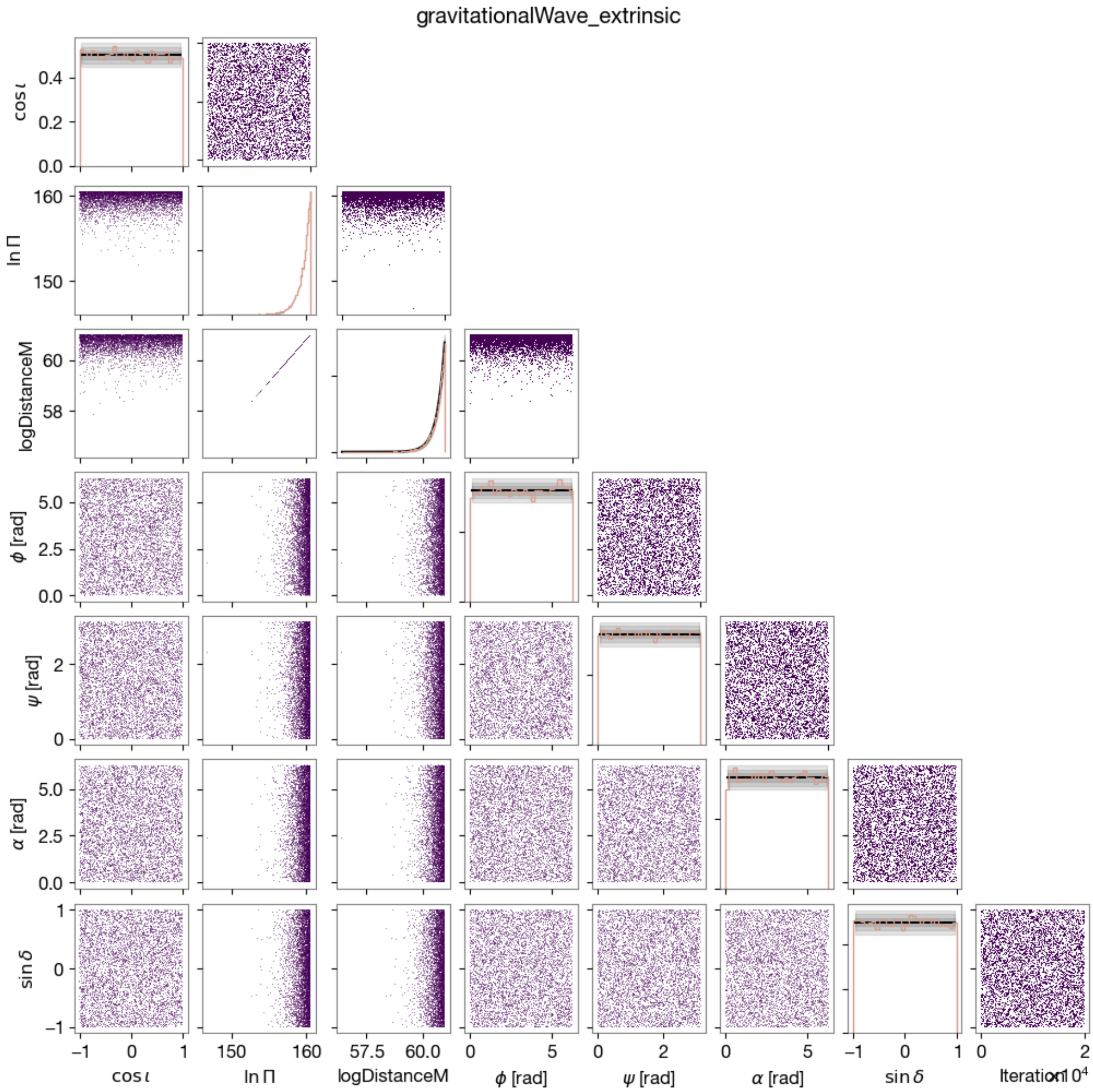



Figure 3: Example Prior Test Results for GW CBC Model, Extrinsic Parameters. All of the 1-D posterior distributions match their respective priors: uniform for most parameters, and exponential in log distance (transformed from uniform in volume). Prior $\Pi$ is a function of all parameters. Default parameters used, segment length 4.0 seconds, sampling rate 2048 Hz.

# Appendix B: Example Glitch Model P-P Plots

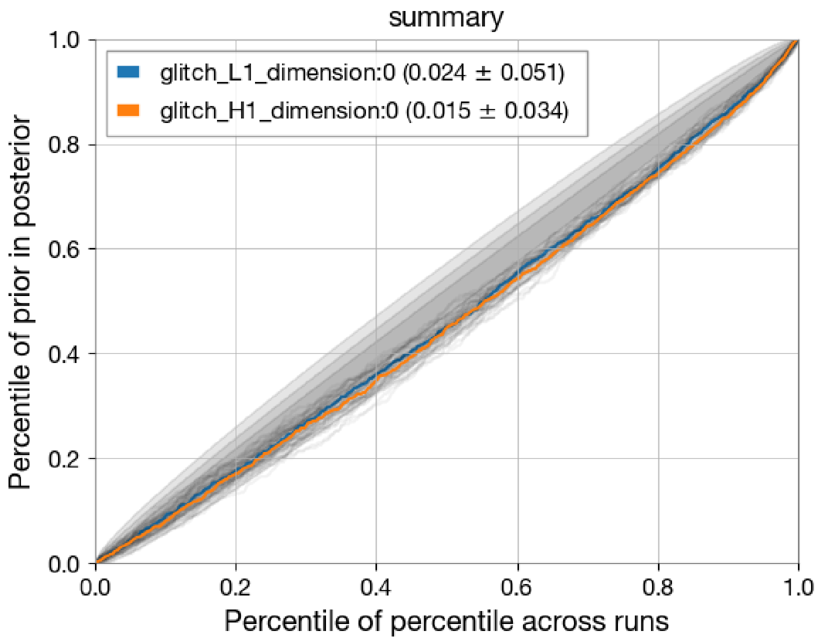



Figure 4: Glitch 5D P-P Plot, No SNR Bound. For SNR within the default range (0.1 to 1000), injected signal dimension uniform from 1 to 5, clear deviations from a straight line are present for best-fit model dimension, just bordering on the pass/fail line.

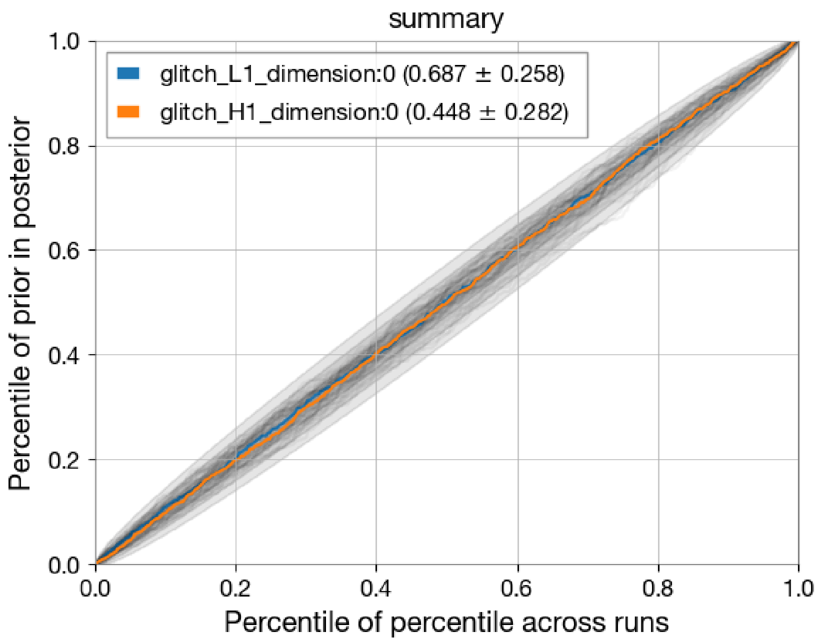



Figure 5: Glitch 5D P-P Plot, No SNR Bound, Uniform Amplitude Prior. For SNR within the default range (0.1 to 1000), injected signal dimension uniform from 1 to 5, using the uniform amplitude prior, the P-P test clearly passes. This is likely because the uniform amplitude prior is uniform arithmetically and not geometrically, so the majority of injected signals have very high amplitude, and thus SNR.

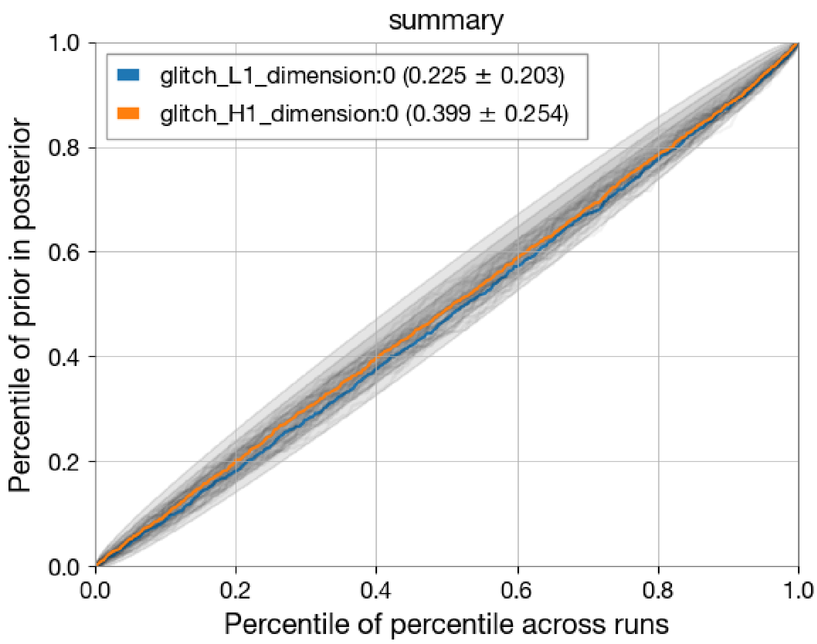



Figure 6: Glitch 5D P-P Plot, SNR Bound of 5. For SNR within the range 5 to 500, injected signal dimension uniform from 1 to 5, the P-P plot is much healthier, and lies fully within the passing regime.

# Appendix C: Example GW Wavelet Model P-P Plots

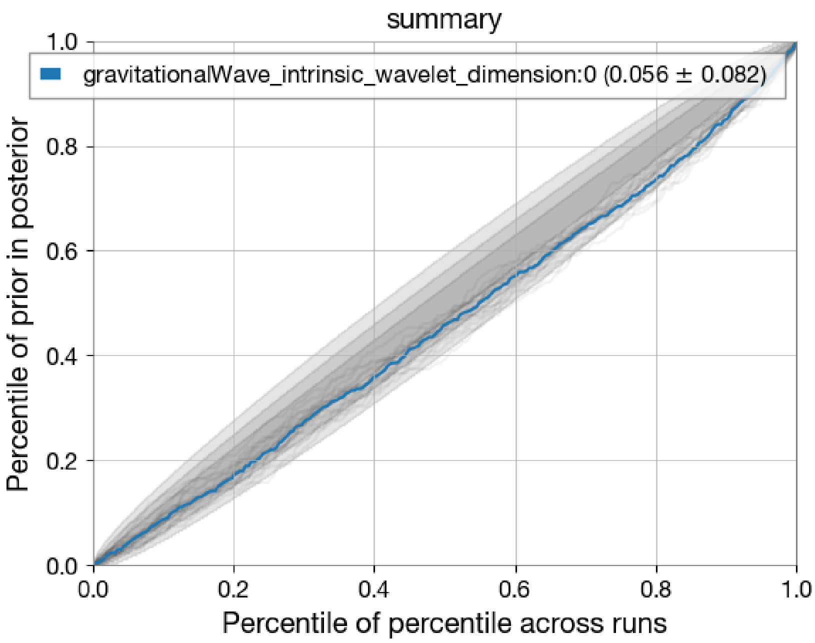



Figure 7: GW Wavelet 5D Dimensional P-P Plot, No SNR Bound. For SNR within the default range (0.1 to 1000), injected signal dimension uniform from 1 to 5, using elliptical polarization, clear deviations from a straight line are present for best-fit model dimension, at certain points bordering the pass/fail line. Similar results are seen for generic polarization.

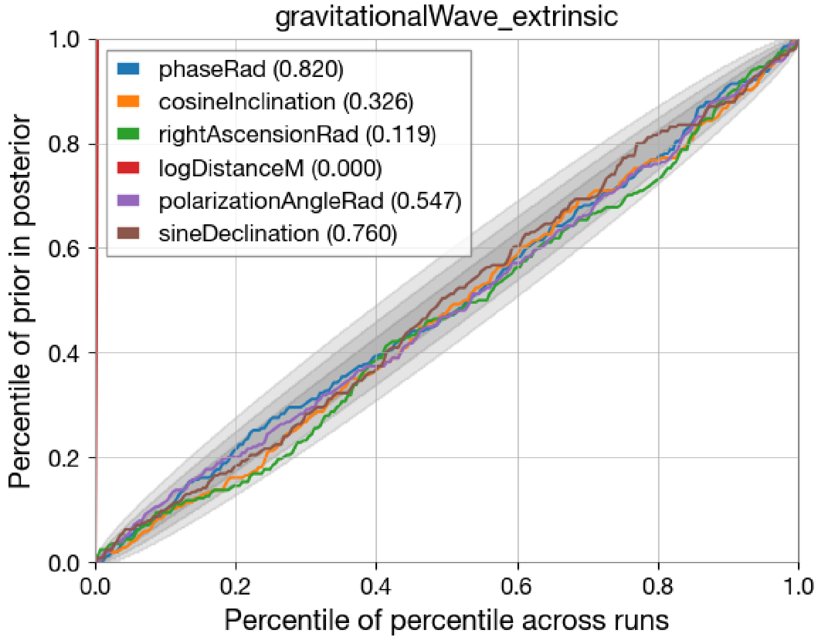



Figure 8: GW Wavelet 5D Extrinsic P-P Plot, No SNR Bound. For SNR within the default range (0.1 to 1000), injected signal dimension uniform from 1 to 5, using elliptical polarization, the P-P test for extrinsic parameters still passes, despite the above issues with the dimension parameter. Similar results are seen for generic polarization.

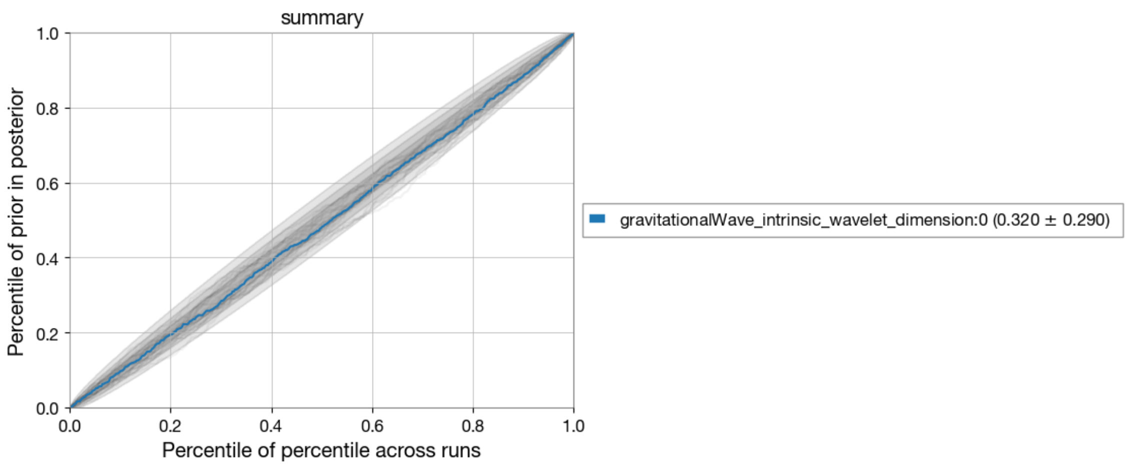
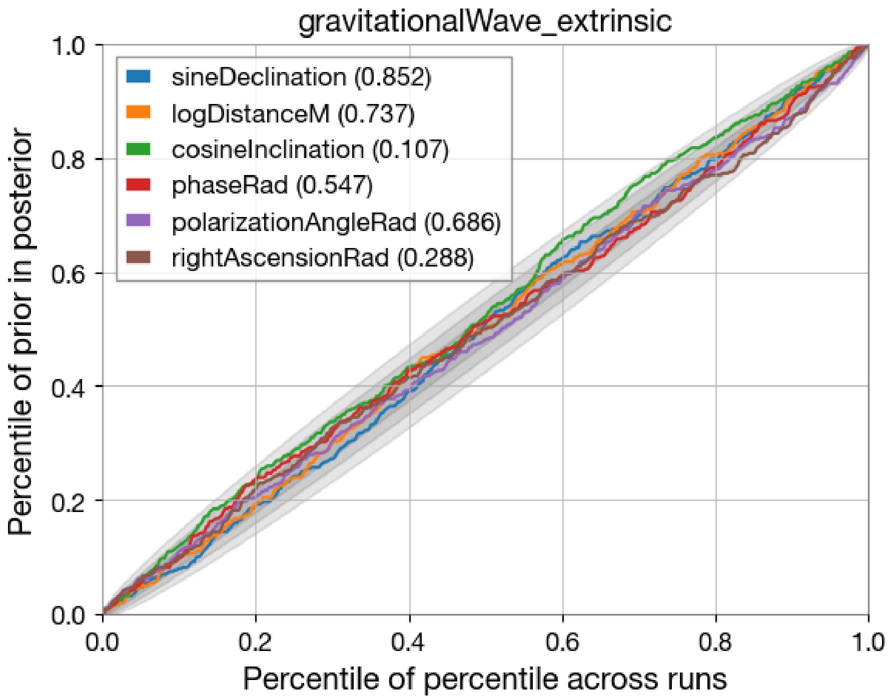



Figure 9: GW Wavelet 5D P-P Plot, SNR Bound of 7. For SNR within the range 7 to 1000, injected signal dimension uniform from 1 to 5, using elliptical polarization, the dimension P-P plot is much healthier, and lies fully within the passing regime. Extrinsic parameters (not shown) also pass. Similar results are seen for generic polarization.

# Appendix D: Example GW CBC Model P-P Plots



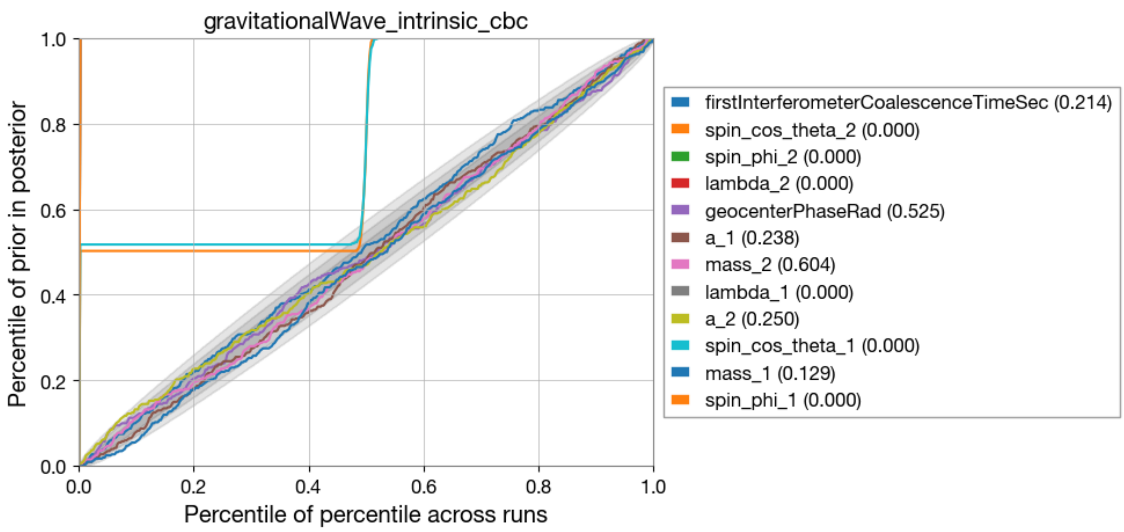




Figure 10: GW CBC P-P Plots, No Precession. Using CBC waveform approximant IMRPhenomD (the simplest CBC waveform implemented in BayesWave++), the P-P tests for both extrinsic parameters (top) and intrinsic parameters (bottom) pass. The heavily-kinked curves in the bottom intrinsic P-P plot correspond to the cosine precession angles of each body in the binary system, an unused parameter in the non-precessing waveform model.

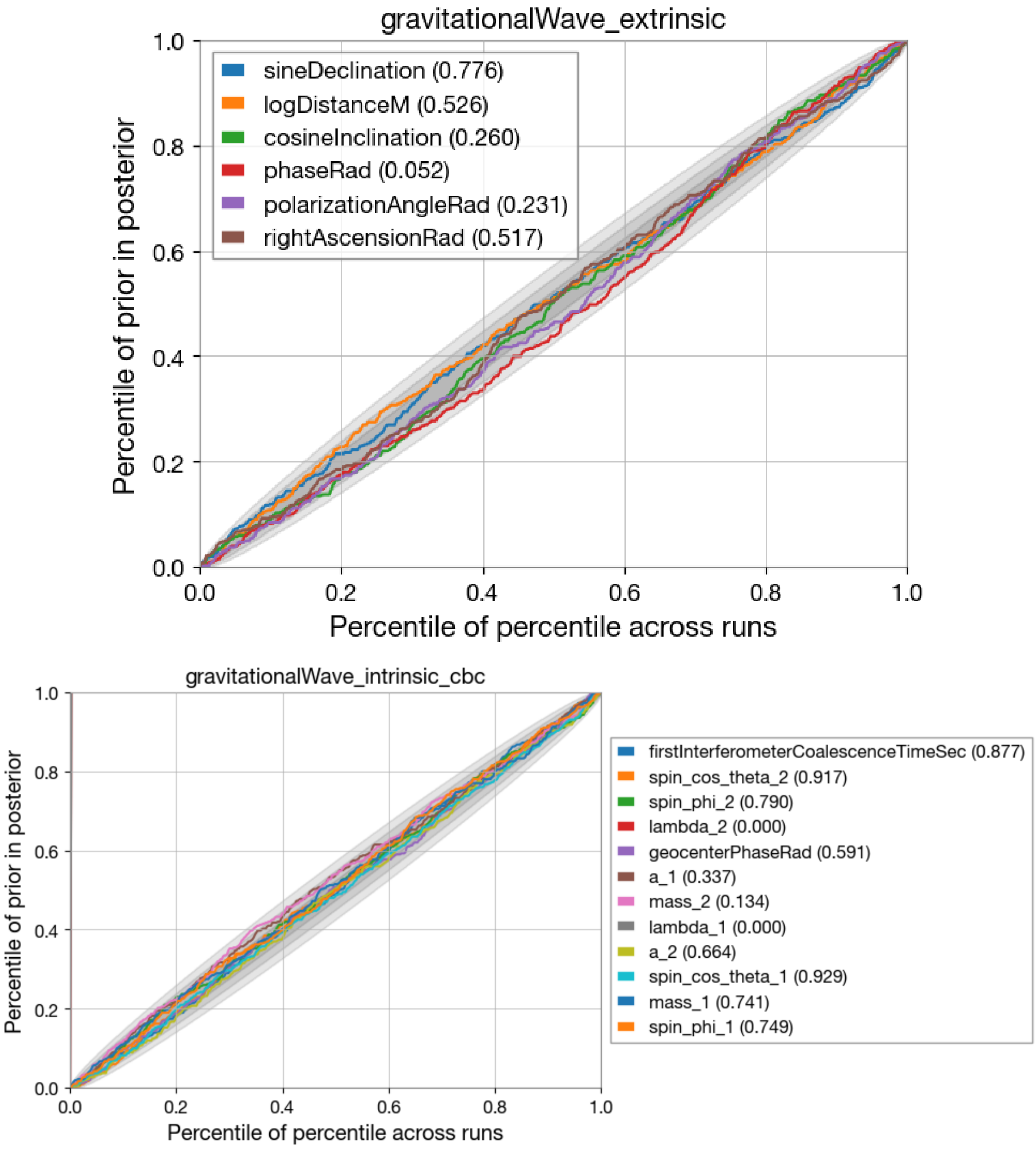



Figure 11: GW CBC P-P Plots, With Precession. Using CBC waveform approximant IMRPhenomXP (the simplest CBC waveform implemented in BayesWave++ that includes precession), the P-P tests for both extrinsic parameters (top) and intrinsic parameters (bottom) pass.