Dataset for Low-Latency EM Follow-Up during O4 and O5 Gravitational Wave Observing Runs

Abigail Gray (UMN), Niharika Sravan (Drexel), Michael Coughlin (UMN)

Introduction

The current and upcoming Gravitational Wave (GW) observing runs by the LIGO/Virgo/KAGRA detectors will result in significantly more detections than previous runs (reference Table I). Preparation to follow up associated electromagnetic signals promptly and accurately from binary neutron star (BNS) and neutron star black hole (NSBH) detections now depends heavily on real-time ML implementation at the detectors. We develop a comprehensive low-latency data set for ML algorithms focused on classification and strategizing follow-up observations. It will incorporate all available data provided by IGWN prompt alerts. This dataset is built on GW localizations from observing scenarios simulations and draws extended data products from associated injections for BNS and NSBH sources.

Run	Distribution	BNS	NSBH	BBH
Annual number of detections				
O4	LRR	17^{+35}_{-13}	10^{+18}_{-8}	46^{+23}_{-17}
	GWTC-3	36^{+49}_{-22}	6^{+11}_{-5}	260^{+330}_{-150}
O5	LRR	86^{+171}_{-59}	48^{+71}_{-30}	190^{+80}_{-58}
	GWTC-3	180^{+220}_{-100}	31^{+42}_{-20}	870^{+1100}_{-480}

Table I: Annual Detection Rates of CBC GW signals expected during O4 and O5 observing runs. Rates are given in year^{-1}[1].

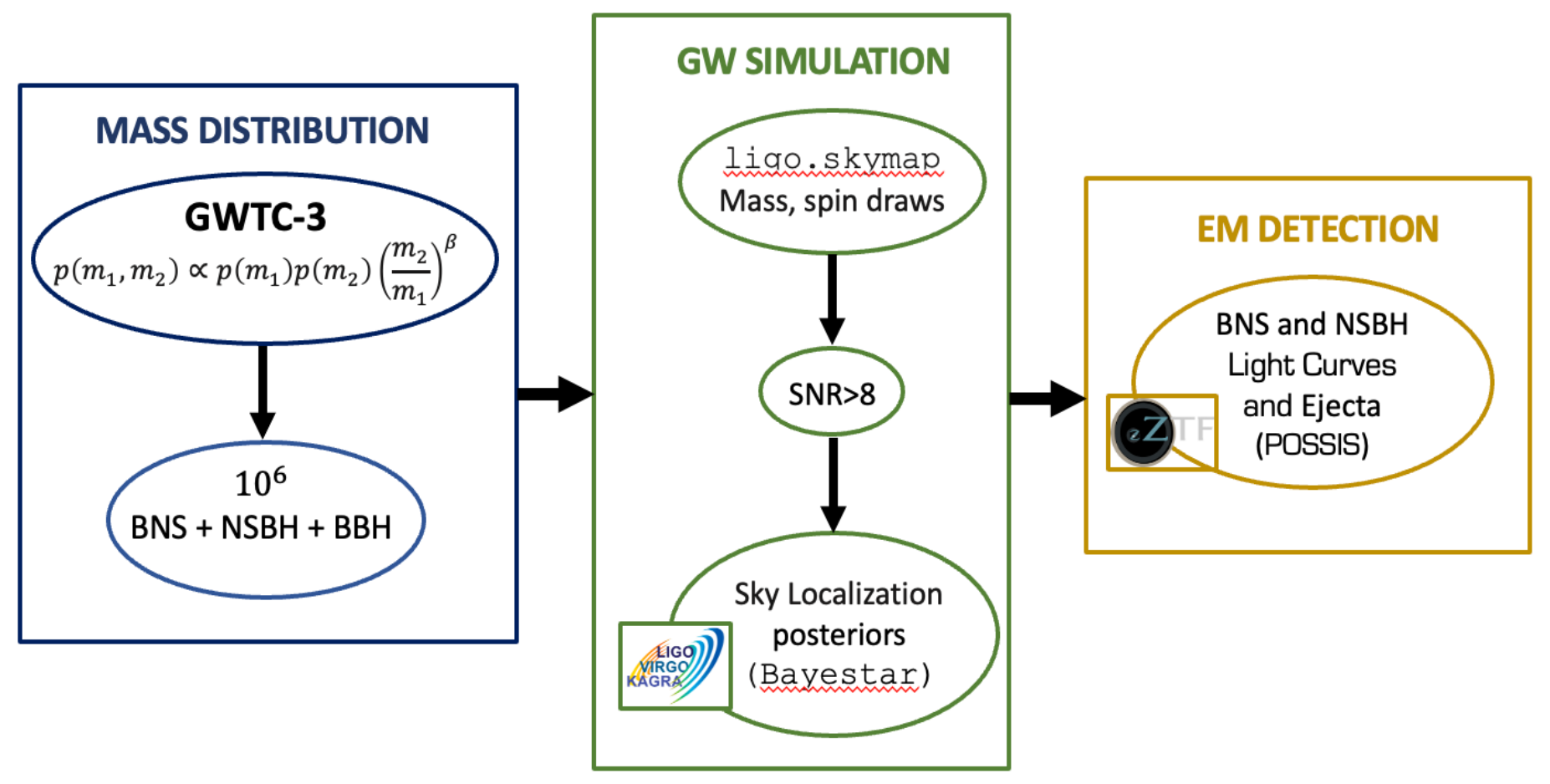
Methodology

We use the GW population injections and sky localizations generated by Observing Scenarios simulations (described below) to produce our extended data products from. We ignore the Binary Black Hole (BBH) populations associated with these populations as we are only interested only in the possibility of electromagnetic radiation from BNS and NSBH populations. For our EM analysis, use Zwicky Transient Facility (ZTF) for observations and scheduling.

Observing Scenarios

Observing Scenarios is an end-to-end GW-EM simulation that generates predictions and statistics for upcoming GW observing runs. A large CBC population is generated based on expected masses and spins of GW sources. Masses and spins are described by *Power Law+Dip+Break* [2]. These GW populations are uniformly and isotropically sampled in a comoving volume, creating our GWTC-3 distribution. An SNR cut of 8 is then applied. Based on LIGO/Virgo/KAGRA detector network, the sky position posteriors are generated for each event. To realize a realistic population for follow-up, we employ ZTF for observation/EM simulations.

Figure 1 (below): Flowchart of simulation process



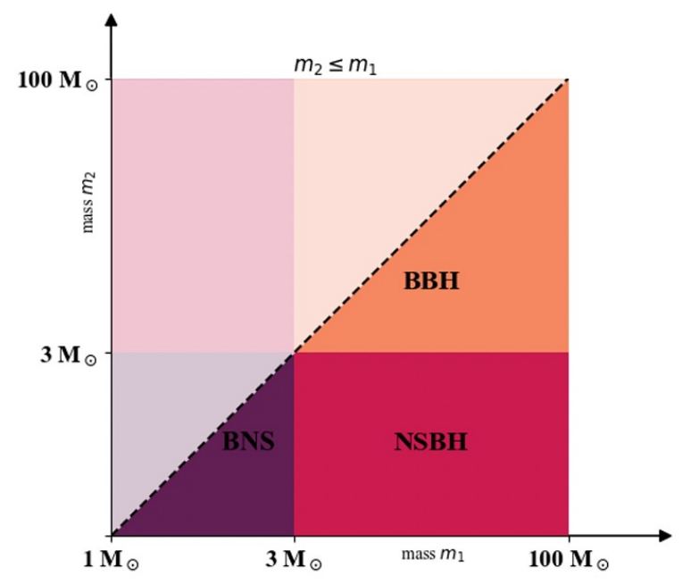


Figure 2 (above): Mass cuttoffs for all CBC population samples. In this work, we ignore BBH samples.

Light Curves and IGWN Products

Light curve posteriors and extended IGWN data products are generated from the realistic set of injections, i.e. the population with associated sky localizations. We simulate ejecta quantities and translate to light curves using a POSSIS-based grid of Bulla2019 Kilanovae Model and ZTF observing capabilities. We sample based on realistic ZTF ToO pointings during the first one or two days, randomly chosen, for exposure times of 180 and 300s and in g,r, and i filters. In conjunction with light curves, we generate realistic ZTF detection by employing observing schedules using *gwemopt* framework. Low-latency data products include event properties and classifications. Nuclear Multimessenger Astronomy (NMMA) code is used as the fundamental framework for all GW+EM simulations, and EMBright code is used to simulate event properties.

Results

A population of 10^6 fCBC's (BNS, NSBH, and BBH) pass the SNR threshold cut.. Eliminating BBH population from our set, we find we have 1004, 2003 BNS and 184, 356 NSBH for O4, O5 runs, respectively.

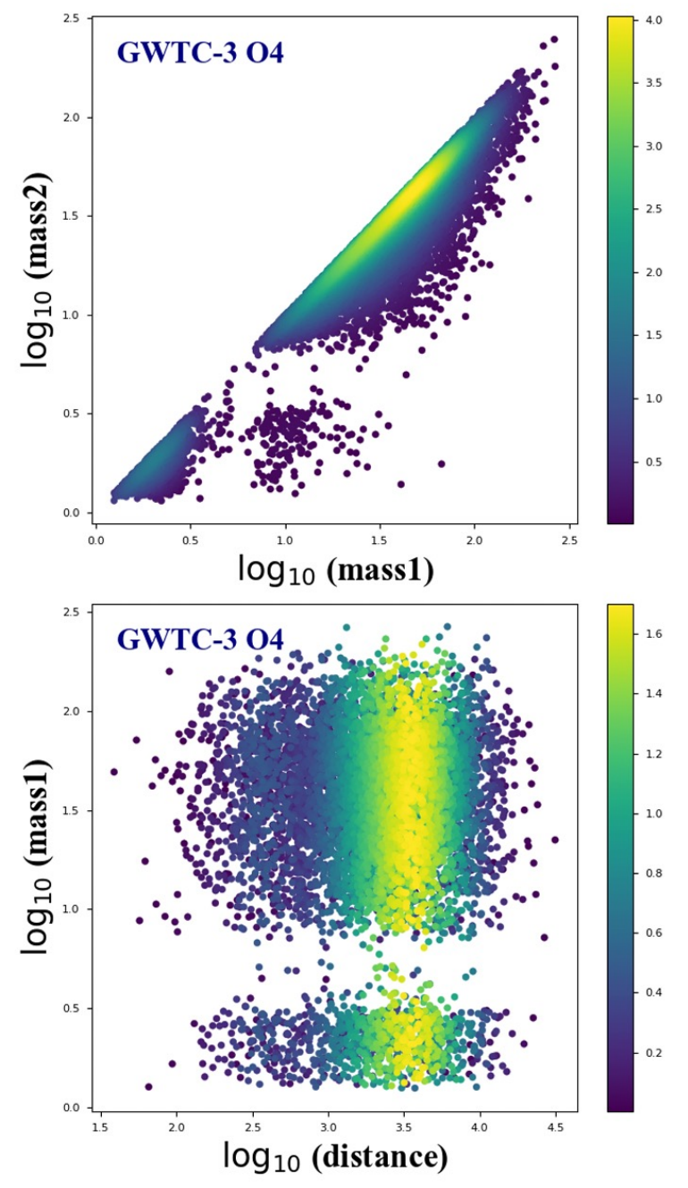


Figure 3 (right): Two-dimensional simulated mass distributions for O4 BNS, NSBH, and BBH. The upper panel shows the distribution of the primary and secondary CBC mass. The lower panel shows the primary mass and distance distribution. Note that this plot includes BBH population.

Based on LIGO/Virgo/KAGRA detector network, the sky localizations are given by probability contour regions, shown by shaded red. These regions are defined as the smallest area enclosing 50% and 90% of the total posterior probability and are realized as the sky area covered to have 50% and 90% chance of including the real kilanovae. A healpix projection fits file is generated for each injection (1188 and 2359 for BNS and NSBH combine for O4, O5, respectively). Distances and the probability contour of the true event are pulled for each injection as fundamental prompt data products. High SNR and nearby events dominate.

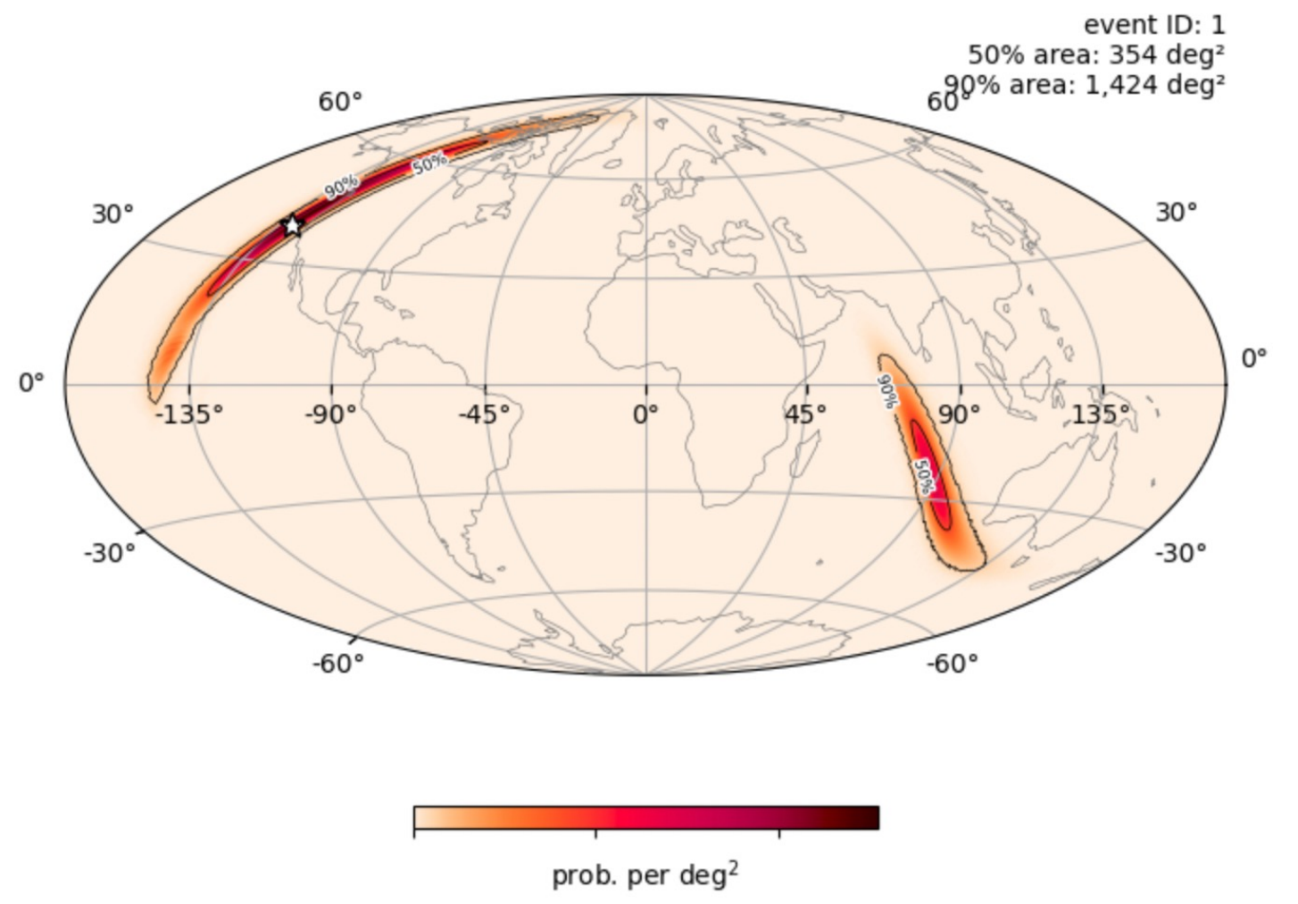


Figure 4 (right): Sample skymap for simulated O4 BNS event. Higher probability or credible regions are shown as darker red. The true location of the injected event is given by the star.

For the whole GW population, or the true Kilanovae in each skymap, we simulate light curves and realistic ZTF detection. Of the thousands of realistic GW injections, we find that only ~10% are detectable by ZTF. This is due to its limiting magnitudes of 21.7, 21.4, and 20.9 in g, r, and i bands respectively, despite its large FOV (47 deg^2). This leaves us with a significantly small set of detectable sources: 236 BNS and 4 NSBH (O4), and 118 BNS, and 0-1 NSBH (O5).

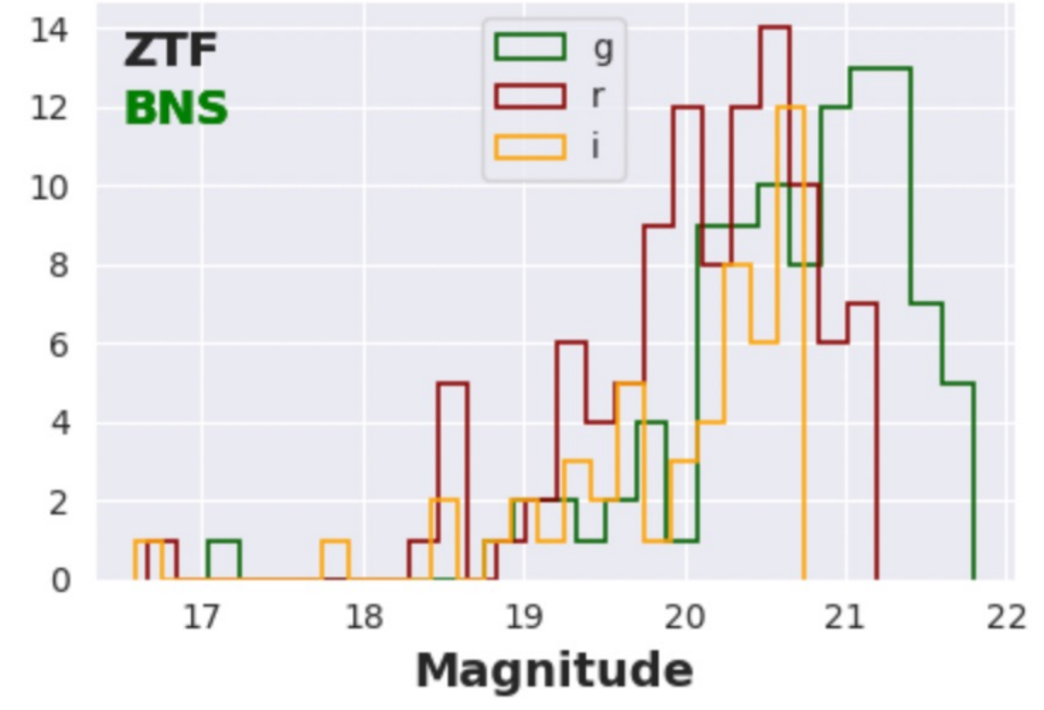
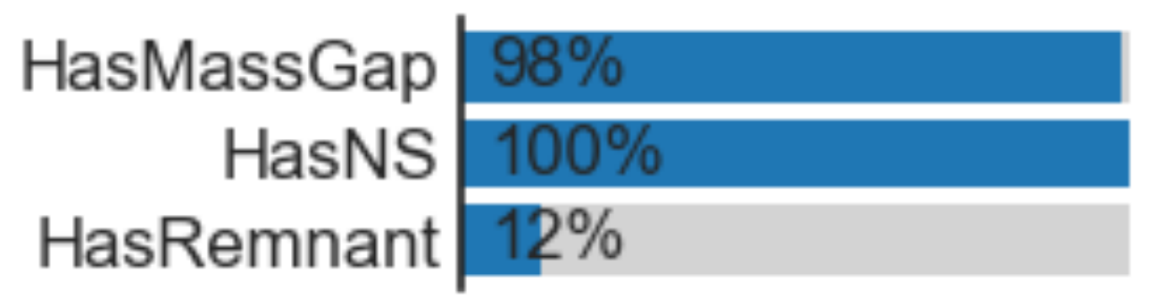


Figure 5(above): Histogram showing the mean magnitude of 236 light curves from O4 BNS detected sample.

We generate the event properties for the same GW population set. This is a fundamental prompt data product is given as the probability that at least one of the compact objects was a neutron star (HasNS), the system ejected a non-zero amount of matter(HasRemnant), and that one of the compact objects has a mass in the range 3-5 solar masses (HasMassGap). We use mass, spin, and SNR to generate this with EMBright codebase.

Table 2(right): Sample event property IGWN product.



We generate event classification probabilities or P-Astro values (in progress). This quantity defines the probability that the GW source is a BNS, NSBH, BBH, or Terrestrial (noise). This calculation depends on FAR, expected GW detection rate, and the number of detection pipelines.

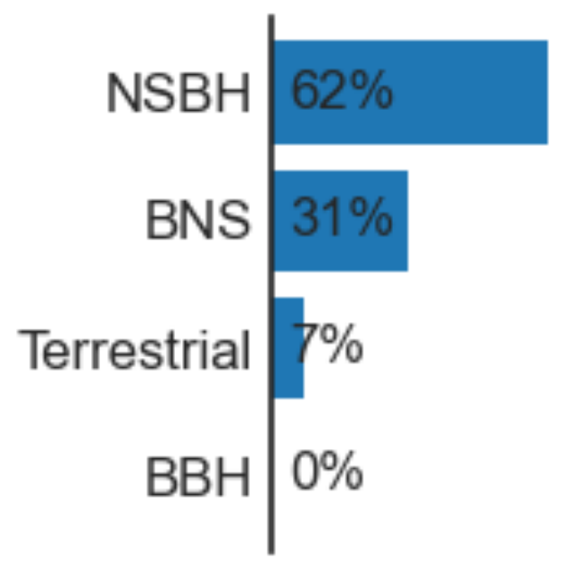


Table 3(left): Sample event classification IGWN product.

Discussion and Future Work

With upcoming observing runs, the data provided in low latency by IGWN is crucial for follow-up decision-making. We have aimed to provide the first full set of these products drawn from realistic end-to-end GW-EM simulation to employ on real-time classifiers. With cuts made in both GW and EM analysis, we look to scale this analysis to provide robust numbers for training. This would require changing selection effects in our population analysis. As we work toward this, we also look to add GRB data, spectra and host galaxy information, especially as we battle the imbalance of increasingly sensitive detectors but little improvement in localization. The current set of light curves, properties and classification is currently being implemented on classifiers. Evaluating the resulting performances will also be useful in evaluating the quality, quantity, and weight of each product. The full developed injection pipeline and codes will be made publicly available

References

1. Kiendrébéogo et al. "Updated observing scenarios and multi-messenger implications for the International Gravitational-wave Network's O4 and O5". In: *ArXiv* (Jun 2023) arXiv:2306.09234 [astro-ph.HE]
2. Farah et al. (2022). Bridging the Gap: Categorizing Gravitational-wave Events at the Transition between Neutron Stars and Black Holes *The Astrophysical Jouranl* 931(2) 108 https://iopscience.iop.org/article/10.3847/1538-4357/ac5f03

School of Physics and Astronomy
University of Minnesota

Accelerated AI Algorithms for Data-Driven Discovery

NSF Award 2117997