



Low-latency gravitational wave alert infrastructure and new data products during LIGO-Virgo-KAGRA fourth observing run

SUSHANT SHARMA CHAUDHARY* ¹, ANDREW TOIVONEN ^{1,2} AND
HASSSM TEAM, CGMI TEAM, LLPIC BUILDERS, LL TEAM AND PIPELINE FOLKS³

¹*School of Physics and Astronomy, University of Minnesota, Minneapolis, Minnesota 55455, USA*

²*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

³*ABC*

ABSTRACT

Multi-messenger investigations of compact object mergers, such as binary neutron star (BNS) and neutron star-black hole (NSBH) systems, remain among the most dynamic frontiers of modern astronomy. During the fourth observing run (O4) of Advanced LIGO, Advanced Virgo, and Kagra, the search for joint electromagnetic and neutrino counterparts to gravitational wave (GW) signals continued using public, semi-automated data products distributed in near real-time. The sky localizations and preliminary source characterizations provided by these alerts helped guide multiple rapid follow-up observations to candidate GW events. During this run, the low-latency search parameter space was expanded to include the sub-solar mass regime and associated public data products were developed. Further, chirp mass estimates of the binary merger system were introduced as a new public data product and distributed to enable follow-up searches to concentrate their resources on the most promising transients. In order to streamline the implementation of new searches and data products, a Mock Data Challenge (MDC), known as the Low Latency Production Integration Challenge (LLPIC), was used as a real-time replay of a simulated population of GW sources to optimize and profile the software infrastructure and scientific deliverables. In this paper we describe these new public data products, their performance and the result of end-to-end testing of the low-latency infrastructure with the LLPIC.

Keywords: Gravitational waves, Multi-Messenger Astronomy, Compact Binary Mergers

1. INTRODUCTION

Following the landmark detection of the first binary black hole (BBH) merger during Advanced LIGO (aLIGO)’s first observing run (O1) (Abbott et al. 2016), the subsequent discovery of the first BNS merger in aLIGO’s and Advanced Virgo (AdVirgo)’s second observing run (O2) (Abbott et al. 2017a) and its associated electromagnetic counterparts AT2017gfo (Coulter et al. 2017; Smartt et al. 2017; Abbott et al. 2017b) and GRB170817A (Goldstein et al. 2017; Savchenko et al. 2017; Abbott et al. 2017c), and the detection of an NSBH (Abbott et al. 2021) in aLIGO’s and AdVirgo’s third observing run (O3), the global network of advanced ground-based interferometers entered

O4 with Laser Interferometer GW Observatory (LIGO), Virgo, and KAGRA GRAVitational-wave observatory (KAGRA) operating in coordination. The first half of O4, which spanned close to 8 months from May 24th, 2023, to January 16th, 2024, yielded nearly 100 candidate compact binary mergers (Abac et al. 2025a). As O4 progresses toward completion, the LIGO-Virgo-KAGRA (LVK) collaboration continues to release a high volume of real-time gravitational-wave alerts, enabling rapid electromagnetic and neutrino follow-up crucial for advancing Multi-Messenger Astronomy (MMA). These efforts support a broad range of science goals, including constraining the equation of state (EoS) of a neutron star (NS) (Bauswein et al. 2017; Margalit & Metzger 2017; Coughlin et al. 2019a, 2018, 2019b; Annala et al. 2018; Most et al. 2018; Radice et al. 2018; Lai et al. 2019; Dietrich et al. 2020; Huth et al. 2022), measuring the Hubble constant (Coughlin et al. 2020a,b; Abbott et al.

*ssharmac@umn.edu

2017; Hotokezaka et al. 2018; Dietrich et al. 2020), and probing r -process nucleosynthesis (Chornock et al. 2017; Coulter et al. 2017; Cowperthwaite et al. 2017; Pian et al. 2017; Rosswog et al. 2017; Smartt et al. 2017; Watson et al. 2019; Kasliwal et al. 2019).

The LVK’s Low-Latency Alert Infrastructure (LLAI) is a composite system comprised of multiple subsystems that enable the real-time analysis of GW candidates and the generation of alerts. The overarching infrastructure collects GW triggers from various low-latency search pipelines (All  n   et al. 2025; Ewing et al. 2024; Dal Canton et al. 2021; Chu et al. 2021; Klimentenko et al. 2016a; Marx et al. 2025; Skliris et al. 2024), produces annotated data products—including classifications, source properties, and sky localizations—and issues automated public alerts. This system plays a crucial role in supporting MMA and target of opportunity (ToO) searches. A detailed description of the components of the LLAI and the LVK’s public data products is provided in Chaudhary et al. (2024). Prior to O4, the O3 Replay MDC exercise (Chaudhary et al. 2024) demonstrated the performance of these data products, characterized the various latencies involved in alert processing, and served as a reference dataset for continuous development and integration testing by both LLAI developers and search pipeline teams.

During the course of O4, the LLAI has been enhanced with new capabilities and data products designed to broaden the scientific reach of the alert system. A key development was enabling public alerts for Sub-Solar Mass (SSM) candidates, which expands the parameter space for low-latency searches (All  n   et al. 2025; Hanna et al. 2024). To support this effort and help observers prioritize these novel triggers, we introduced HasSSM, a new public data product based on a machine learning classifier to quantify the presence of a SSM object in the binary merger. In addition to expanding the search scope, O4 alerts are enriched with more detailed source information. During the run, the LVK introduced a new chirp mass data product in the form of binned source-frame chirp mass estimates for significant GW candidate alerts. This product provides a more direct look at the intrinsic properties of the source in low-latency. To ensure new searches and data products were robust and fully integrated, a new, week-long MDC named Low-Latency Production Integration Challenge (LLPIC), was generated using simulated O4 background noise. In addition to compact binary coalescence (CBC) signals from BNS, NSBH and BBH, this MDC also included SSM and transient GWs which do not originate from a CBC source, known as “bursts.” Many astrophysical objects can produce bursts of GWs: core-

collapse supernovae Janka (2012); Abdikamalov et al. (2020), magnetars Nazari & Roshan (2020); Contopoulos et al. (2023); Kouvaris (2025), starquakes of neutron stars Watts & Strohmayer (2007); Giliberti & Cambiotti (2022), non-linear memory Ebersold & Tiwari (2020), hyperbolic encounters of compact objects Bini et al. (2024), and cosmic strings Damour & Vilenkin (2000, 2001, 2005). Unlike the O3 replay MDC (which was primarily used to optimize and stress-test the infrastructure), these new signals provided a crucial testbed for continuous development of the LLAI.

This paper details these significant enhancements to the LVK low-latency infrastructure for O4, focusing on the new data products and the validation framework that supports them. In Sec. 2, we briefly describe various components of LLAI and introduce the Mock Event Generator (MEG) and its significance in continuous integration-testing of LLAI. In Sec. 3, we describe the extension of public alerts to include sub-solar mass candidates, detailing the training and performance of the HasSSM machine learning classifier designed to identify them. Following this, in Sec. 4, we introduce chirp mass estimates, outlining the methodology performance of this new low-latency data product. In Sec. 5, we present the results from the LLPIC which provided a comprehensive end-to-end validation of these new features using a realistic O4 environment. We conclude in Sec. 6 with a summary of these changes and a future outlook.

2. LLAI

The LVK’s LLAI is a complex, distributed system that enables the real-time detection, analysis, and dissemination of GW alerts. It integrates several critical components: low-latency data calibration and transfer, online modeled and unmodeled search pipelines, and the suite of infrastructure elements for event management and alert distribution. These include GRAvitational-wave Candidate Event DataBase (GraceDB)¹, which serves as the central event database; `igwn-alert`,² an internal Kafka-based alert system for communicating event-state updates; and `GWCelery`,³ a distributed task queue that clusters, annotates, and orchestrates events, as well as publishes public alerts for the broader astronomical community.

The process of detection of GW and subsequent alerts begins at the detectors. The real-time GW strain data is calibrated at the detector sites and sent to California

¹ <https://gracedb.ligo.org/>

² <https://igwn-alert.readthedocs.io>

³ <https://git.ligo.org/emfollow/gwcelery>

Institute of Technology computing clusters. The low-latency search pipelines operate continuously on this LVK data stream to identify potential GW signals in near real-time. Modeled searches (CBC)—including GStreamer LIGO Scientific Collaboration Algorithm Library (GstLAL) (Messick et al. 2017; Tsukada et al. 2023; Ewing et al. 2024), PyCBC (Nitz et al. 2018; Dal Canton et al. 2021), Multi-Band Template Analysis (MBTA) (All  n   et al. 2025; Aubin et al. 2021), Summed Parallel Infinite Impulse Response (SPIIR) (Chu et al. 2021; Kovalam et al. 2022), and Aframe (Marx et al. 2025)—use matched filtering or Machine Learning (ML) based techniques to detect GW signals of a CBC. Unmodeled searches (Burst)—such as Coherent WaveBurst (cWB) (Klimenko et al. 2008, 2016b; Drago et al. 2021) and MLy (Skliris et al. 2024)—rely on excess power or convolutional neural network (CNN) algorithms to identify generic transient signals. When a search pipeline identifies a candidate (also called a GW trigger), it assigns a false alarm rate (FAR) and uploads the trigger to GraceDB. Multiple triggers may arrive nearly simultaneously. Through `igwn-alert`, `GWCelery` continuously monitors GraceDB for updates, performs further annotation (such as rapid sky localization and source classification), and clusters temporally coincident triggers within ± 1 s into a *superevent*. A preferred event is then selected—based on the highest network signal-to-noise ratio (SNR) for CBC pipelines or the lowest FAR for Burst pipelines. The preferred event’s data products form the basis for public alert generation once a predefined FAR threshold is crossed (see Chaudhary et al. 2024 for details). Up-to-date information about current FAR thresholds can be found in the LVK user guide⁴

Over long observing runs, continuous development and integration of new features, search pipelines, and data products are essential. However, the complexity of the system makes end-to-end testing challenging. The O3 Replay MDC helped address this by providing a cyclic 40-day data stream processed by production-equivalent pipeline instances. Triggers were uploaded to a dedicated internal GraceDB instance named “playground”, with a parallel `GWCelery` deployment managing event processing and annotations. This configuration enabled realistic end-to-end validation of the low-latency system. However, maintaining such parallel pipeline instances over extended runs is computationally expensive. The subsequent development of the MEG framework and its associated features, which we detail in the following section, provided a more scalable and

efficient solution for continued testing and improvement of the LLAI.

2.1. Mock Event Generator

MEG is designed to *replay GW triggers* in real-time by injecting simulated events from the MDC into the full LVK low-latency pipeline stack. MEG reproduces the behavior of real detections by producing GraceDB entries, sky maps, and associated JSON data products (*pastro*, EM-Bright, etc.) in order to enable end-to-end testing of the alert system under realistic operating conditions prior to and during observing runs.

MEG ensures that these signals appear as *live alerts* to downstream infrastructure. This design allows for systematic validation of every stage of the workflow – from the search pipelines, through GraceDB ingestion, to the orchestration layer in `GWCelery`. A key advantage is that MEG provides reproducibility: identical events can be replayed across multiple cycles, allowing direct comparison of data product consistency.

As part of its expanded functionality, MEG introduced a structured review process to ensure the reliability and scientific integrity of replayed events. Each replay undergoes a series of automated and manual validation steps designed to mirror production-level quality control. These include verifying event metadata, such as ensuring that JSON fields like *pastro*, `HasNS`, and binned chirp mass probabilities (see Sec. 4) are correctly populated—along with cross-checking sky maps generated by BAYESian TriAngulation and Rapid localization (BAYESTAR) against injected parameters to confirm statistical consistency. The framework also performs timing and latency assessments to ensure that alert generation meets the O4 design goal of publishing initial alerts. Additionally, cross-pipeline consistency checks are carried out by comparing detections from multiple search pipelines (e.g., GstLAL, PyCBC, MBTA, etc) to verify agreement in classification and parameter recovery. Together, these measures provide a robust mechanism for validating LLAI performance and ensuring that any updates or new features integrated through MEG maintain the expected reliability and accuracy standards.

The inclusion of MEG in the latest MDC (see Sec. 5) framework provides confidence that the low-latency alerts distributed to the astronomical community are scientifically robust. The replay-and-review framework has enabled verification of new data products — such as sub-solar mass classifications and binned chirp mass estimates – and a new ML based search pipeline Aframe Marx et al. (2025) prior to their release on GraceDB. More broadly, the MEG review lowers the risk of false or

⁴ <https://emfollow.docs.ligo.org/userguide/>

malformed alerts reaching electromagnetic and neutrino follow-up facilities, which rely on reliable, low-latency triggers for efficient allocation of telescope resources.

As the LVK moves toward O5, we envision expanding MEG to cover new categories of data products, including early-warning alerts (Ryan Magee & Zweizig 2021), low-latency Bayesian classification (Romero-Shaw et al. 2020a), and potentially machine-learning-based event vetting. Another priority is to further automate the MEG Review with standardized validation scripts, enabling daily replay checks without manual inspection. The long-term goal is a continuous integration-style framework, where every new low-latency feature is tested through MEG prior to deployment.

3. SUB-SOLAR MASS SEARCH AND DATA PRODUCTS

Previously, all CBC low-latency searches had limited their search on compact objects above one solar mass, however new SSM searches specifically intended for objects below one solar mass have been introduced in O4. Largely, searches for sub-solar mass compact objects are motivated by the possibility to explore new physics, and potentially detect dark matter since low-mass objects more compact than white-dwarfs cannot be produced by standard astrophysical channels.

Most invoked of these are primordial black holes (PBH) which are hypothesized to have formed from large density fluctuations in the early universe, potentially during cosmic inflation (Carr & Hawking 1974; Carr et al. 2016). While the mass spectrum of PBHs may overlap with that of stellar-origin black holes, the existence of SSM PBHs remains an open question. Notably, even the first BBH event, GW150914, was hypothesized to be of primordial origin, adding to the long-standing debate on whether PBHs could contribute significantly to dark matter (Bird et al. 2016; Clesse & García-Bellido 2018).

Another proposed scenario involves the cooling and gravitational collapse of dissipative dark matter halos, which can produce BHs with sub-solar masses and lead to merger rates that may be within the sensitivity reach of advanced GW detectors (Shandera et al. 2018; Kouvaris & Nielsen 2015).

Additionally, alternative exotic compact objects such as boson stars, composed of scalar field condensates, and novel formation channels like imploding neutron stars mediated by interactions with dark matter, could contribute to the mass spectrum near a solar mass (Cardoso & Pani 2019; Schunck & Mielke 2003). These scenarios fall within the scope of ongoing and future SSM GW searches.

Beyond the dark matter channels, sub-solar mass neutron stars have been proposed to form through the fragmentation of collapsar disks or core fission inside a core-collapse supernova producing a “superkilonova” (Piro & Pfahl 2007; Metzger et al. 2024). Therefore, triggers from searches for coalescing SSM binaries form interesting candidates for multi-messenger follow-up to also probe neutron star formation.

Although no confirmed GW detections from SSM binary mergers have been reported, recent LVK searches for sub-solar mass compact binaries in data from the second half of the third observing run have placed upper limits on their merger rate density (Abbott et al. 2022). These limits continue to refine our understanding of exotic compact objects and their potential role in astrophysics and cosmology.

3.1. Online SSM Searches

During the fourth observing run (O4), the low latency SSM search was carried out by the GstLAL (Cannon et al. 2012; Messick et al. 2017; Sachdev et al. 2019; Hanna et al. 2020; Cannon et al. 2020; Sakon et al. 2022; Tsukada et al. 2023; Hanna et al. 2024) and MBTA (Abadie et al. 2012; Adams et al. 2016; Aubin et al. 2021; Alléné et al. 2025) search pipelines. Both searches target systems with at least one SSM compact object.

In the template bank used in the low-latency search, GstLAL restricts the component masses to $m_{1,2} \geq 0.5M_{\odot}$ and m_1 (m_2) $\leq 10M_{\odot}$ ($1M_{\odot}$), unlike in the offline search, where component masses extend down to $0.2M_{\odot}$. The bank contains only spin-aligned templates with $|\chi_{1,2}| < 0.3$. The template bank is constructed with a minimal match of 0.97, using the IMR-PHENOMD (Husa et al. 2016; Khan et al. 2016) waveform model with a cut-off on the waveform duration set to 128 s (Hanna et al. 2024). Since the GW signals from these low-mass systems can last for minutes in the LIGO-Virgo sensitivity band, the filtering is restricted to 45-512 Hz to alleviate memory usage.

The MBTA low-latency SSM search covers the same parameter space as the offline search with the component masses $m_{1,2} \geq 0.2M_{\odot}$ and m_1 (m_2) $\leq 10M_{\odot}$ ($1M_{\odot}$). MBTA bank also contains only spin-aligned templates with the dimensionless spin magnitude restricted to $-0.1 \leq \chi_{1,2} \leq 0.1$ if $m_{1,2} \leq 0.5M_{\odot}$ while $-0.9 \leq \chi_{1,2} \leq 0.9$ for all other component masses. The bank was constructed with a minimal match of 0.97 using the TAYLORF2 waveform approximant (Buonanno et al. 2009). MBTA uses a low frequency cut-off of 45 Hz, and a high frequency cut-off of 1024 Hz.

For more details on the construction and performance of the template bank, and configuration of low-latency SSM searches, see [Hanna et al. \(2024\)](#) and [All  n   et al. \(2025\)](#).

3.2. SSM Injection Study

In the $m_1 - m_2$ space, chirp mass forms an ambiguity contour leading to a degeneracy where binaries with different component masses can have the same chirp mass. Due to this, the chirp mass of templates in the SSM bank coincides with the typical chirp mass of BNS and NSBH signals causing SSM searches to detect BNS and NSBH signals. It is therefore important to provide public data products indicating probabilities for presence of an SSM, NS and even mass-gap compact objects. To create such data products, we adopt an ML based approach similar to [Berbel et al. \(2024\)](#); [Chatterjee et al. \(2020\)](#).

In order to train the ML classifier, an injection set was constructed spanning the parameter space of the online SSM searches. Approximately 1.3×10^5 simulated CBC waveforms were used for injection with intrinsic properties (masses and spins) given in Table 1. The optimal network SNR is greater than 5 for all injections and is set to prevent ‘‘hopeless’’ injections, which are unlikely to be detected in reality. IMRPhenomXAS ([Pratten et al. 2020](#)) was used as waveform approximant for the injections. The injection density was set to be as low as 10 s and the injection study was carried out in O4a data starting from Sep 06, 2023 (GPS time:1378079744) lasting two weeks. The events were distributed uniformly in co-moving volume assuming flat Λ CDM cosmology with $H_0 = 67.3 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $\Omega_m = 0.3$ based on Planck 2018 results mentioned in Table 1 of [Planck Collaboration \(2020\)](#). Both MBTA and GstLAL pipelines analyzed the injection with their online template bank and produced triggers in response. Additional cuts $q > 0.1$, $M_c^{det} \leq 2.56$, redshift < 0.25 were also applied.

Fig. 1 shows the injection recovery from GstLAL and MBTA pipelines respectively after a FAR cut of $FAR < 3.8 \times 10^{-7} \text{ Hz}$. MBTA recovered $\sim 1.3 \times 10^4$ injections whereas GstLAL recovered $\sim 1.5 \times 10^4$ injections at the aforementioned FAR threshold. For both pipelines, chirp mass is extremely well-recovered. However, mass ratio q is poorly recovered, typically estimating lower values than what was injected. This causes non SSM triggers being recovered in SSM parameter space as seen in Fig. 2. This injection recovery data was used for training the classifier for individual pipelines.

3.3. SSM Search Data Products

As witnessed in Fig. 2 for GstLAL and Fig. 16 for MBTA, SSM searches can recover signals beyond the

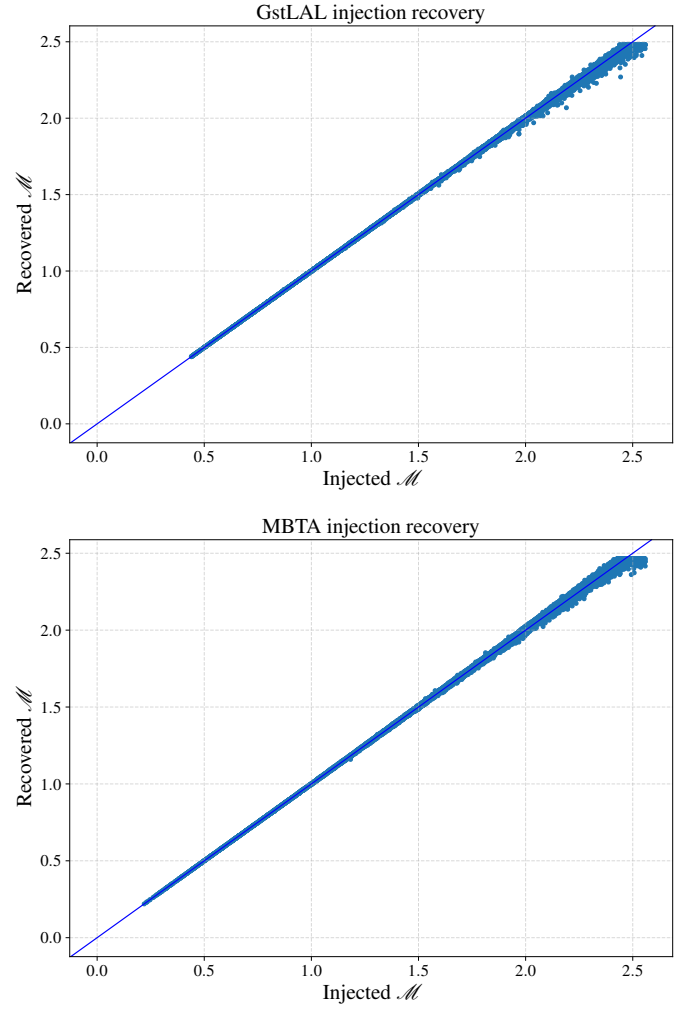


Figure 1. chirp mass injection-recovery plots for GstLAL (upper) and MBTA (lower) pipeline. The diagonal line in both plots is the $x=y$ line indicating perfect recovery. Both pipelines accurately recover chirp mass of the injections. (use chirp mass instead of \mathcal{M})

SSM parameter regime with significant FAR. This implies that a possible SSM search trigger could contain compact objects that are not SSM in nature i.e. a low mass BNS merger could easily be recovered as SSM object. In order to assess the source properties of SSM triggers in low latency, three data products were constructed:

- **HasSSM** : Probability that at least one SSM compact object is present in the binary merger.
- **HasNS**: Probability that at least one NS is present in the binary merger. For SSM searches, we have defined the lower NS boundary is set to $1 M_\odot$ whereas upper NS boundary is set to $3 M_\odot$.

Table 1. Distribution of intrinsic properties (component masses m and spins a) of binary systems in SSM injections dataset.

SSM Injection Properties				
Object	m/M_\odot (min/max)	m distribution	Max a	a distribution
Primary	0.19 - 11	m^{-1}	0.5	uniform & isotropic
Secondary	0.19 - 5	m^{-1}	0.5	uniform & isotropic

- **HasMassGap**: Probability that at least one mass gap compact object ($3\text{--}5 M_\odot$) is present in the binary merger.

HasNS data product is also defined for traditional non-SSM CBC searches. Unlike SSM search, **HasNS** for other searches consider only the upper boundary of NS based on various EoS and the final value is marginalized over multiple EoS. Another important quantity, **HasRemnant**, is not computed for SSM searches since we make no distinction whether the SSM trigger has a NS or black hole (BH), only that it has a sub-solar mass component.

To construct these data products, we trained on the injection dataset using a multi-labeling scheme. Each sub-section of the mass parameter (SSM-SSM, SSM-BNS, BNS-MassGap, etc.) were assigned separate labels. Although the injections were carried with max m_1 and max m_2 of $11 M_\odot$ and $5 M_\odot$ respectively, the injections recovered by the pipelines were limited to $\sim 11 M_\odot$ for m_1 and $\sim 3 M_\odot$ for m_2 . Hence, mass gap only injections were absent from the dataset. The dataset was further divided into 70-30% split for training and testing purpose. Following Chatterjee et al. (2020); Berbel et al. (2024), both K-Nearest Neighbors (KNN) and Random Forest (RF) algorithms were used in the training process. The input data for training consisted of reconstructed parameters $\{m_1^{rec}, m_2^{rec}, \chi_{1,z}^{rec}, \chi_{2,z}^{rec}, \mathcal{M}^{rec}, \text{SNR}\}$ as input features. After a series of cross-validation, the KNN algorithm with distance as weights and 25 neighbors was chosen for the training purpose since it slightly out performed RF.

The final data products **HasSSM**, **HasNS**, **HasMassGap** were computed by combining probabilities of labels contained within the definition of each data product. Using the testing set, Receiver Operating Characteristic (ROC) curves were generated showing the performance for each data product in Fig. 3. The ROC curves show that **HasSSM** performs better compared to other data products. The SSM searches primarily target sources in the region $m_2 < 1 M_\odot$ and the classifier shows ~ 0.85 True Positive Rate (TPR) at 0.2 False Positive Rate (FPR). MBTA shows better performance than GstLAL, which can be linked to the greater coverage of the MBTA bank with the SSM injections used to train the classifier, in both masses and

spins. SSM searches are not designed to pick signals in the BNS and NSBH regimes. Despite that, they can recover BNS and NSBH with significant FAR as seen in the injection-recovery plots. Signals in this region are typically identified with lower q values, and it becomes challenging for the classifier to learn this behavior consistently. This can be reflected in the poorer performance of **HasNS** and **HasMassGap** (see Fig. 3).

3.4. Low-Latency Implementation

The low-latency SSM pipeline data products have been publicly available since early 2025. In the current implementation, a significant GW trigger from SSM searches is released publicly only when no significant counterpart is identified by the standard (non-SSM) CBC searches. In such cases, the public release includes the sky localization map along with the SSM data products—**HasSSM**, **HasNS**, and **HasMassGap**—contained within the **em-bright** packet. As of O4, p_{astro} values are not available for SSM searches. The significance threshold for SSM triggers is set to a FAR of one per year (3.17×10^{-8} Hz) before accounting for trials factors.

Following completion of the low-latency Parameter Estimation (PE), the data products are updated based on the posterior samples. Since the introduction of SSM searches and associated data products, the **em-bright** PE update now includes four quantities: **HasNS**, **HasRemnant**, **HasMassGap**, and **HasSSM**. The **HasSSM** quantity represents samples with component masses $< 1 M_\odot$, while **HasNS** includes samples $\geq 1 M_\odot$ constrained by the neutron star EoS. The **HasNS** value is marginalized over 10,000 EoS draws (Legred et al. 2021, 2022). The **HasRemnant** quantity is evaluated using samples from the BNS and NSBH regions following Foucart et al. (2018) and marginalized over a large ensemble of EoS realizations. Finally, **HasMassGap** corresponds to component masses between 3 and $5 M_\odot$. The same set of updated data products is provided for all classes of CBC triggers.

4. PUBLIC CHIRP MASS ESTIMATES

The LVK is now providing low-latency, public mass estimates for significant candidate compact merger events. Low-latency GW searches have traditionally provided estimates of FAR, sky localization, and source classification, but no estimate of mass. Mass information can be

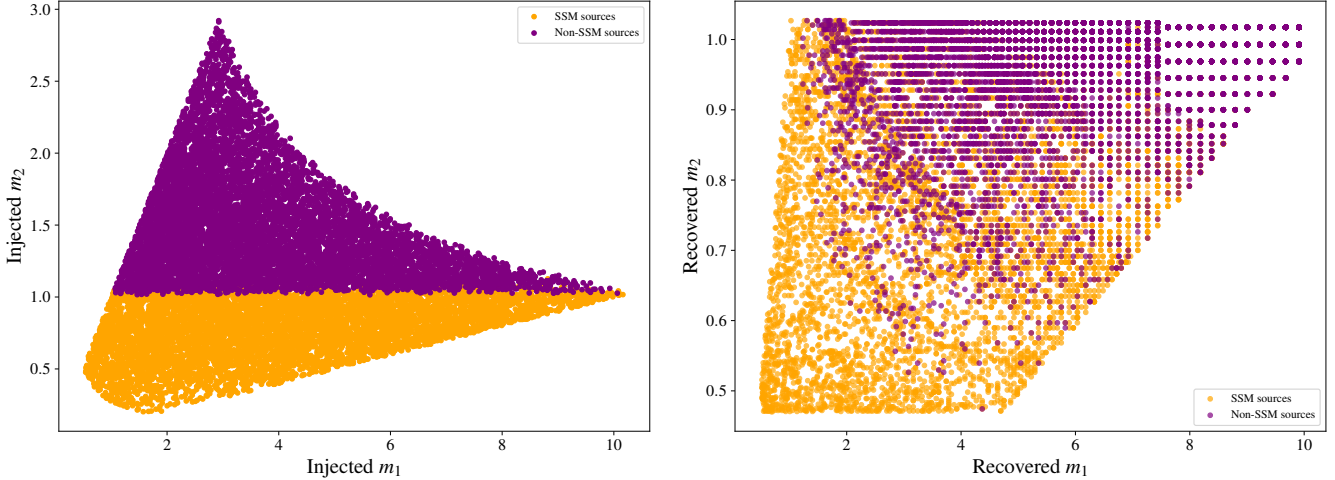


Figure 2. Left: m_1 and m_2 scatter plot of the injection dataset successfully recovered by GstLAL pipeline. Right: m_1 and m_2 scatter plot of the injection dataset successfully recovered by GstLAL pipeline. The colors show whether the injection was in SSM mass space or otherwise. We see significant number of non-SSM injections being recovered by the SSM search with template parameters in the SSM region.

used to determine properties of associated counterparts (Margalit & Metzger 2019), so it is in high demand. These new mass estimates are released as probabilities for a predetermined set of chirp mass bins, and can be later updated by parameter estimation. The chirp mass is estimated through a combination of search pipeline information, sky maps, and parameter estimation. This data product enables better interpretation of GW candidate events and aids multi-messenger ToO searches by providing mass information alongside source classification to better characterize the progenitor system. We introduce this new public GW data product below, and intend for it to be ingested by consumers of our alerts and enable downstream scientific efforts.

Chirp mass is well-measured from a GW signal due to its relationship with the frequency evolution of the signal, making GW searches are well-positioned to provide low-latency chirp mass estimates. Chirp mass, \mathcal{M} , is defined as:

$$\mathcal{M} = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}}, \quad (1)$$

and is degenerate with the individual component masses, meaning a single chirp mass value can correspond to multiple component mass pairs.

We provide a new public data product for chirp mass, currently being released for significant O4 candidate events, in the form of binned estimates of chirp mass for significant events uploaded by CBC searches, including cWB BBH (Klimenko et al. 2008, 2016b; Drago et al. 2021). In solar masses, the chirp mass bins are:

$$[0.1, 0.87, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.7, 1.9, 2.1, 2.3, 3, 5.5, 11, 22, 44, 88, 1000]. \quad (2)$$

These bins are in the source frame of the merger, *except* in the case of the low-latency cWB BBH estimate, where there is no distance or redshift information available. In this case, we report the “detector frame”, or redshifted estimate while still using the same bins. To shift to the source frame, we can relate the redshifted detector frame value \mathcal{M}_z to \mathcal{M} using the redshift z :

$$\mathcal{M}_z = (1 + z)\mathcal{M}. \quad (3)$$

The mass estimates are made publicly available on **GraceDB** alongside the Second Preliminary Alert, and may later be the subject of an Update Alert. On **GraceDB**, they can be found in the form of a `.json` file of the probabilities for each bin and a `.png` plot of the same probabilities, as seen in example Fig. 4. The highest probability bin will be reported in the circular, but this information is *not currently* included in the alert packet `.xml`. This was done to ease integration for O4; inclusion is a possible goal for O5.

Fig. 5 shows the low-latency and parameter estimation-based estimates run on MDC Chaudhary et al. (2024) events compared to their injected chirp mass values. We see that the distribution of chirp mass estimates closely match the injected values, particularly at low mass where there is large interest in multi-messenger follow-up. This demonstrates the chirp mass estimates are extremely accurate at recovering the true chirp mass in low-mass regime.

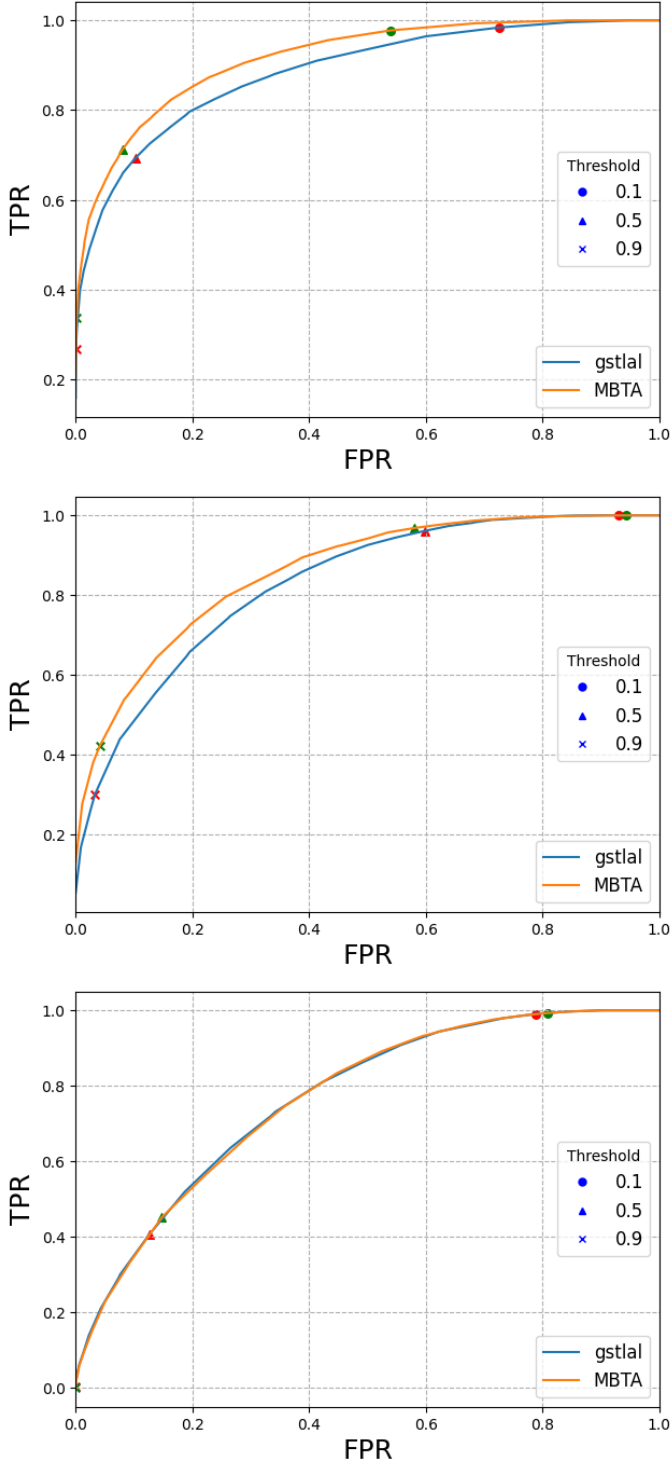


Figure 3. ROC curves on the top panel represents *HasSSM*, the middle represents *HasNS* and the bottom panel represents *HasMassGap*. The two different lines represents classifiers trained for individual pipelines. The markers represent the threshold used within the ROC curve.

The binned chirp mass estimates are intended to help inform time-sensitive follow-up activities. This includes

multi-messenger follow-up of events with at least one neutron star (Paterson et al. 2021; Agayeva et al. 2022; Paek et al. 2024; Ahumada et al. 2024; Takahashi et al. 2025; Coughlin et al. 2019c), as well as follow-up of high mass BBH events (Abdalla et al. 2021; Vajpeyi et al. 2022) in the search for electromagnetic counterparts that could result from the flaring and disruption of AGN accretion disks during merger.

4.1. Bin Rationale

The bins were designed to be finely spaced at lower masses where chirp mass is recovered very accurately, and as these low masses are of particular interest due to their potential of containing a NS. While specific bins do not exclusively relate to a given source class, they were designed with the different regimes or source classes in mind, and can be loosely split into the following three regimes: the SSM regime, the *HasNS* regime, and the BBH regime. As we cover the different regimes, keep in mind there are multiple component mass pairs for a given chirp mass, and that the regimes themselves are simply meant to be illustrative of the different source classes we can expect to observe.

The SSM regime corresponds to the bin $[0.1, 0.87] M_{\odot}$, likely to contain an event with a component mass $< 1 M_{\odot}$. The upper edge of this bin, a chirp mass of $0.87 M_{\odot}$, may correspond to a $1 M_{\odot}, 1 M_{\odot}$ merger, for example, demonstrating events with $\mathcal{M} < 0.87 M_{\odot}$ will contain a component mass $< 1 M_{\odot}$. The *HasNS* regime roughly corresponds to chirp mass values for mergers with a strong possibility of containing a NS: either BNS or NSBH. The bins in the chirp mass range $[0.87, 3] M_{\odot}$ roughly fall within this regime, again with a reminder these regimes are not exclusive and are for illustrative purposes. For a *HasNS* example, a chirp mass of $3 M_{\odot}$ may correspond to a $1.4 M_{\odot}, 10 M_{\odot}$ NSBH merger. Lastly, the coarse, high mass bins are likely to correspond to BBH events. Chirp mass bins above $3 M_{\odot}$, which go all the way to $1000 M_{\odot}$, have a strong chance of being a BBH.

4.2. Implementation in alert infrastructure

Binned chirp mass estimates come in two forms: (i) estimates based on information available in low latency (pipeline uploads, sky maps), and (ii) estimates based on Bilby parameter estimation (Ashton et al. 2019; Romero-Shaw et al. 2020b). The former is immediately available in low latency, while the latter accompanies update alerts \sim hour(s) later when Bilby parameter estimation is complete. Chirp mass estimates are only be provided for significant events. Similar to the procedure with other update information (EM-Bright, sky maps,

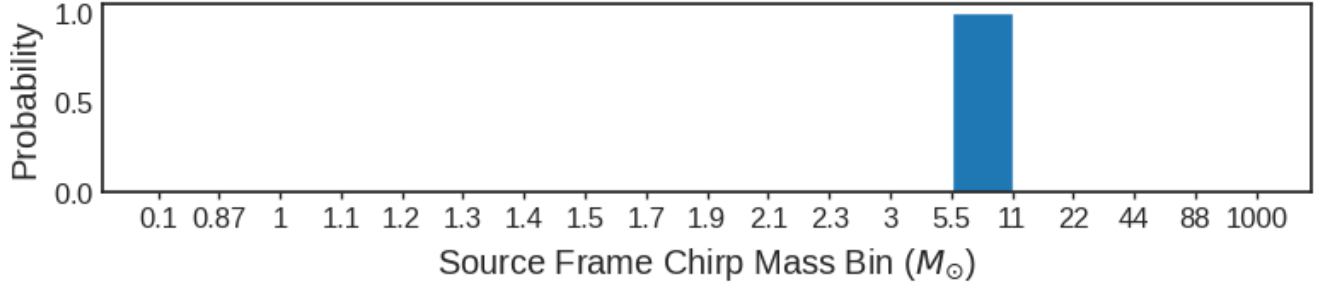


Figure 4. Example histogram of chirp mass estimate that will be available on **GraceDB** for significant events.

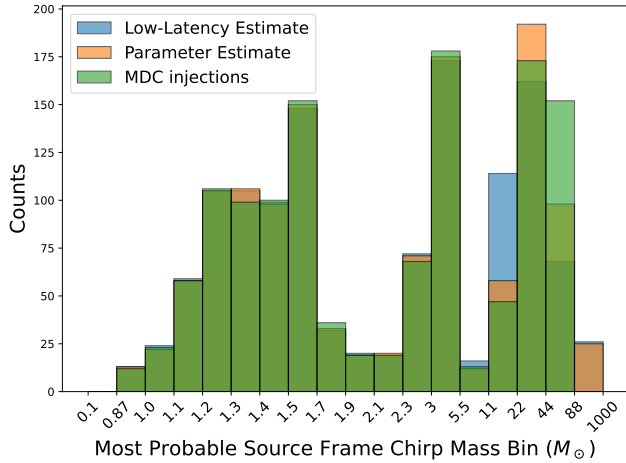


Figure 5. The most likely bin from the low-latency and parameter estimation-based estimates run on MDC Chaudhary et al. (2024) events compared to their injected chirp mass.

etc), the updated chirp mass based on parameter estimation will be uploaded and identified as resulting from parameter estimation, without deleting or overwriting the original low-latency estimate. Both these estimates assume the candidate is astrophysical in origin.

The low-latency chirp mass estimates use the pipeline point estimate of the detector chirp mass, combined with distance information provided by the **BAYESTAR** sky map (Singer & Price 2016), in order to correct the detector chirp mass value to the source frame. The result is binned and converted to an estimate of the probability corresponding to each bin. In the case of cWB BBH, there is no distance information, and therefore we assume a redshift of zero when reporting the source chirp mass. The estimate is uploaded in the form of a `.json` file named `mchirp_source.json`.

Once the low-latency PE is completed, the chirp mass estimates are updated using the source frame posterior samples directly. The result is binned and converted to an estimate of the probability corresponding to each

bin, and that estimate will be uploaded in the form of a `.json` file named `mchirp_source_PE.json`

4.3. Multi-Messenger Applications

In this following, we highlight two direct multi-messenger applications of the chirp mass estimates: i) inform follow-up of potential NS mergers by helping determine whether the merger likely contained a NS, and how likely that event is to produce a kilonova, and ii) targeted follow-up of high mass BBH events in the hope of detecting an electromagnetic (EM) counterpart in the form of a flare from an active galactic nucleus (AGN).

4.3.1. Neutron Star Merger Follow-up

GW candidate events with at least one NS are often subjects of ToO searches for EM counterparts such as Gamma-Ray Bursts (GRBs) or kilonovae. ToO searches rely on low-latency data products and timely decision making when determining which GW candidate events are worth devoting telescope resources to. When searching for kilonovae, one main criteria astronomers consider is whether the candidate event likely contains a NS. Alongside quantities such as `HasNS` and `p_astro`, we expect the chirp mass estimates to help answer this question. In addition, not only do these quantities offer insight on the possible presence of a NS, they also provide an actual mass estimate. This is useful as we expect different masses to produce varying amounts of mass ejecta and consequently have varying likelihoods of producing a kilonova. The chirp mass could also be used to predict the resulting remnant of the merger Margalit & Metzger (2019).

4.4. BBH Follow-up

It is expected that high-mass BBH mergers that take place in the accretion disk of an AGN may produce an observable EM flare Graham et al. (2023). As an example, **BBHBot** is a automatic triggering pipeline designed to monitor GW triggers in real time, and schedule ToO observations with ZTF Bellm (2014) to observe promising BBH candidates. **BBHBot** ingests public low-latency

GW data products such as the new chirp mass estimates as of recently, in addition to the sky map and FAR in order to determine which events meet the triggering criteria. If candidate events pass cuts on mass and FAR as well as meet our observing criteria, including cuts on the 90% sky map area and air mass, an observing plan is automatically drafted and submitted to trigger ZTF.

Specifically, we are interested in events with $> 50 M_{\odot}$ of total mass, and even more so those with $> 100 M_{\odot}$ of total mass, so we filter for events most likely to fall within chirp mass bins $> 22 M_{\odot}$. Keeping in mind the definition of chirp mass in Eq. 1, the merger of a pair of $25 M_{\odot}$ BHs corresponds to a $\mathcal{M} \sim 22 M_{\odot}$, while the merger of a pair of $50 M_{\odot}$ BHs corresponds to a $\mathcal{M} \sim 44 M_{\odot}$. Conveniently, there are chirp mass bin edges at both 22 and 44 M_{\odot} . Following the addition of chirp mass to the public GW data products in June 2025, BBHBot has used it in the real-time evaluation of more than a dozen candidate BBH mergers as of August 2025. Of these, nine exceeded the selection threshold of a chirp mass $> 22 M_{\odot}$.

4.5. A timely example: S250818k

S250818k was a recent low-significance GW candidate of interest, given a possible external coincidence with an EM source⁵, AT2025ulz. Extensive follow-up of the counterpart AT2025ulz was carried out, with the possibility the source was associated with the GW event and compact object merger (Franz et al. 2025; Gillanders et al. 2025; Kasliwal et al. 2025). Due to this possible coincidence, chirp mass estimates were provided for this GW event⁶, even though it was not significant. Alongside the possible counterpart, the chirp mass estimates identified S250818k as a standout event given its low mass and high chance of containing a NS and SSM object. There has never been a confirmed SSM detection, so this possibility is exciting and demonstrates the use of the chirp mass data product, given this information would not have been publicly released in the past. The chirp mass estimates are intended to give astronomers the additional information needed to make these informed follow-up decisions and identify standout events such as S250818k. It is important to remember, however, that while both the low-latency and parameter estimation-based chirp mass estimates firmly place this event in the SSM regime, those estimates assume the event is astrophysical in origin, and that is not necessarily a safe assumption in the case of this low-significance

event. In Fig. 6, we show the chirp mass estimate for this event, falling within the $[0.1, 0.87] M_{\odot}$ bin.

5. LLPIC

(Needs an intro paragraph?)

The LLPIC was commissioned to streamline the implementation new searches and their downstream data products in the form of a new MDC. This simulation campaign consists of a real-time replay of injected waveforms constituting a population of gravitational-wave sources, and was used to validate the performance of new searches and data products, as well as ease integration into the larger LLAI.

)

To create a background for the LLPIC, we consider a week long stretch of data taken between April 27, 2024 – May 4, 2024 (GPS times ranging between: 1398297600 - 1398902400) by the LVK instruments during O4b. Simulated CBC and Burst waveforms were injected into the data and then streamed in similar fashion to the low-latency data distribution. The data was analyzed by downstream low-latency CBC and Burst pipelines and the GW triggers produced by them were submitted to an internal-only instance of GraceDB named “playground”. Further clustering, annotation, and alerts for these events was done by GWCelery. The flow of data from analysis to alerts is kept identical to the O4 low-latency production system. The one week data containing injection is replayed every week (starting April 29th, 2025) leading to multiple LLPIC cycles.

5.1. CBC Injections

A total of 2,364 simulated CBC waveforms with mass and spin distributions detailed in Table 2 were injected into LLPIC data stretch with an interval of ~ 4 minute between injections. The SNR cutoff was set to 6 to allow recovery of more injections relative to the O3 MDC. The IMRPhenomXPHM waveform approximant is used to generate waveforms for injections. For injection purposes, the boundary between NSs and BHs was considered to be $3M_{\odot}$. The binary sources in Table 2 were mixed to ensure an equal amount of injections belonging to each source class. While the injection rate density and distribution used in the study is artificially high and not representative of the true population, these injections are suitable for testing the low-latency infrastructure and data products other than p_{astro} , which relies on a background distribution expected for normal detector operations. The injection was carried out in all three detectors—Hanford, Livingston, and Virgo (H, L and V) (Aasi et al 2015; Acernese et al 2015)—and a fraction of them fall during detector down time. The CBC injec-

⁵ <https://gcn.nasa.gov/circulars/41414>

⁶ <https://gcn.nasa.gov/circulars/41440>

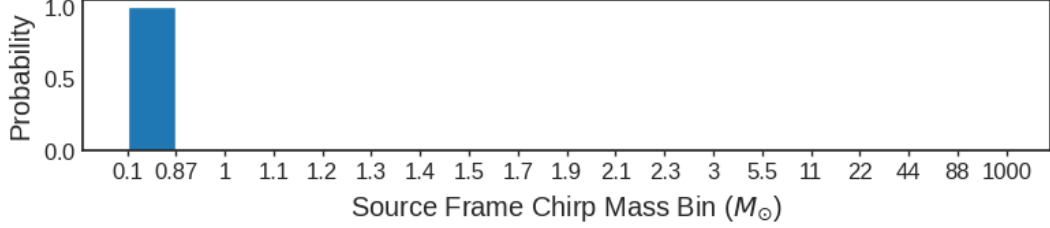


Figure 6. Histogram of chirp mass estimate for S250818k publicly available on *GraceDB* and placing the candidate event in the SSM regime.

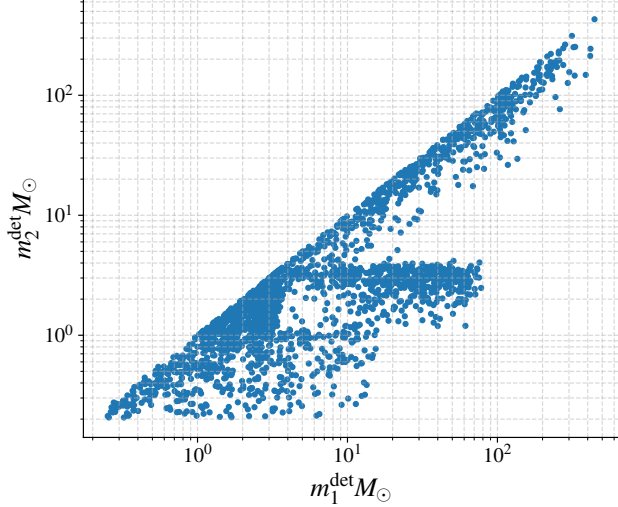


Figure 7. Distribution of LLPIC CBC injections in detector frame masses.

tion set consists of 600 SSM, 626 BNS, 553 NSBH, and 585 BBH injections.

5.2. Burst Injections

The goal of injecting Burst signals in the LLPIC data was different unlike the regular injection studies commonly conducted by the LVK search groups to characterize the sensitivity of Burst search pipelines [Abac et al. \(2025b\)](#). In LLPIC, the number of Burst injections were limited to only 82. The time of injections were chosen to test different detector network configurations (2 or 3 detectors observing simultaneously), to be spaced by more than one hour from each others and not to overlap a CBC injection. This ensured that no single injection fell during interferometer down time. Several waveform models were considered. First, generic waveforms were selected to describe the unmodeled nature of Burst sources. They include sine functions in a Gaussian envelope, band- and time-limited stochastic white noise bursts, and Gaussian pulses. Intrinsic parameters

for these waveforms such as duration, frequency or bandwidth were varied to scan the search parameter space. In addition, astrophysically-motivated signals were also injected: they include core-collapse supernova waveforms derived from numerical simulations [Mezzacappa et al. \(2023\)](#); [Pan et al. \(2021\)](#); [Powell & Müller \(2019\)](#) and phenomenological gravitational waveforms from cosmic strings [Damour & Vilenkin \(2000, 2001, 2005\)](#). The injection amplitudes were randomly drawn in a range tuned to offer a good balance between significant detections and signals invisible to searches. Finally, the incoming sky direction of GWs is randomly drawn on the sphere.

5.3. Results

5.3.1. Pipeline Performance

The replay of LLPIC started from 29 April 2025. A majority of the low-latency search pipelines (GstLAL, PyCBC, MBTA, cWB, MLy) participated in analyzing the LLPIC data on its 4th cycle ranging from 20 – 27 May 2025. **(Previous cycles were used for validation and integration testing.)** In the following, we study the GW triggers and their corresponding data products uploaded to *GraceDB* during this time frame. Triggers and injections are cross matched using the merger time; all triggers within 1s of an injection merger time are matched to that injection and included in the analysis. This process removes noise triggers between injections and ensures trigger times correspond to an injection time window. This process, however, does not completely filter noise triggers and they can still be coincident with the injection time window. This cross matching process allows us to evaluate data products by comparing the results with the injected quantities in a similar manner to the O3 MDC [Chaudhary et al. \(2024\)](#).

During the 4th cycle of LLPIC, 332 significant CBC signals were cross-matched with the injections among which 45 were SSM injections, 86 were BNS, 63 were NSBH, and 138 were BBH injections. Despite an SNR cutoff of 6, $\sim 75\%$ of CBC signals are not recovered. This is because majority of CBC signals were lost since

		CBC Injection Properties			
Binary Type	Object	m/M_\odot (min/max)	m distribution	a distribution	max redshift
SSM	m_1	0.2 - 15	$m^{-2.5}$	75% Uniform(0 - 0.4) + 25% Uniform(0-1.0)	0.25
	m_2	0.2 - 1.0	m^{-1}	75% Uniform(0 - 0.4) + 25% Uniform(0-1.0)	
BNS	m_1	1.0 - 3.0	Uniform	75% Uniform(0 - 0.4) + 25% Uniform(0-1.0)	0.25
	m_2	1.0 - 3.0	Uniform	75% Uniform(0 - 0.4) + 25% Uniform(0-1.0)	
NSBH	m_1	3.0 - 60.0	m^{-1}	Uniform(0 - 1)	0.75
	m_2	1.0 - 3.0	m^1	75% Uniform(0 - 0.4) + 25% Uniform(0-1.0)	
BBH	m_1	3.0 - 1000	$m^{-2.35}/m^{-3}/m^{-1}$	Uniform(0 - 1)	2.5
	m_2	3.0 - 1000	m^1	Uniform(0 - 1)	

Table 2. Distribution of intrinsic properties (component masses m and spins a) of binary systems in the CBC injection sample in LLPIC. The primary mass (m_1) for the BBH population use broken power law with breaks at $100 M_\odot$ and $200 M_\odot$. The spin distributions are isotropic in orientation.

they were injected in time period where one or more detectors were either not operational or not in science mode. Other factors like large distance and possible glitch overlap also affect successful recovery of the signals.

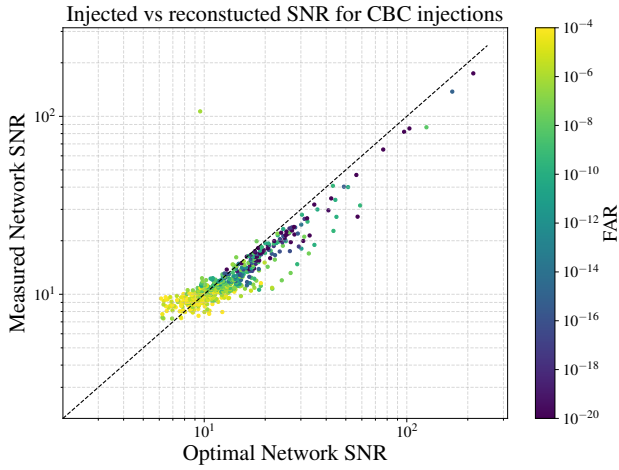


Figure 8. The measured network SNR recovered during the LLPIC cycle 4 compared to the optimal, or injected, network SNR with the points colored by FAR.

We plot the simulated vs. recovered network SNR in Figure 8. Overall, signals with moderate SNR are recovered well, with some bias at low SNR due to the FAR threshold imposed for upload. Extremely loud signals are recovered with slightly less SNR than the injected signal. This could be due to the fact that simulated optimal SNRs were calculated using fixed detector sensitivities, whereas actual detector data has significant fluctuations in sensitivity over time. The optimal SNRs for injections are calculated using a global average PSD, instead of using a local estimate of the PSD, which may cause some of the off-diagonal outliers.

In case of Burst injections, 45 of them were detected as significant triggers by a Burst pipeline, cWB or MLy. All of them were detected by cWB while MLy offered a redundant detection for 7 injections. Interestingly, Burst injections can also be detected by CBC pipelines. Fig. 9 presents the number of Burst injections detected by the different low-latency pipelines. The version of cWB configured to target BBH signals, labeled “cwb_bbh”, detected 13 Burst injections. This result is expected as the pipeline is only mildly modeled for BBHs and remains sensitive to non-CBC signals. The CBC match-filter pipelines, MBTA and GstLAL, detected 4 and 7 Burst injections respectively showing that CBC templates can trigger on Burst waveforms, including sine-Gaussian, cosmic string and supernova waveforms. These injections were all recovered with $P(\text{BBH}) > 96\%$. Indeed, Burst signals can be mis-interpreted as heavy BBHs with very few cycles in the bandwidth of the detectors (Abbott et al. (2020)). The majority of these signals were detected with a higher significance by Burst pipelines. However one injection, a cosmic string signal, was best detected by GstLAL: the SNR reported by GstLAL is 19.7 while the cWB SNR is 18.2, and the false-alarm rate for this event is 3.0×10^{-9} Hz and 1.8×10^{-8} Hz for GstLAL and cWB respectively. This example shows that very short duration GW candidates can be mis-identified as CBC and that careful follow-up studies must be performed to state on the nature of a candidate. Fig. 9 also shows that the PyCBC pipeline did not detect any Burst injections, making it the most robust pipeline to detect CBC signals.

5.3.2. Latency Measures

To enable prompt multi-messenger follow-up, a central goal of the LLA is the rapid dissemination of GW alerts along with data products. In this study, we systematically examine the alert latency arising from

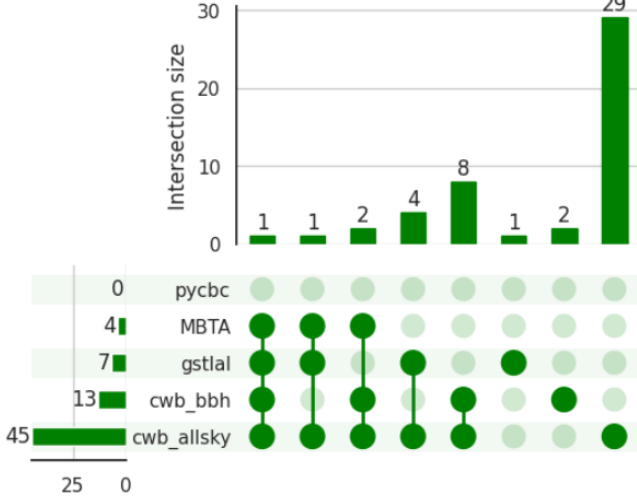


Figure 9. Overlap between Burst and CBC searches for Burst injections. The number of detected Burst injections is indicated for each search pipeline in the left-side of the diagram. The histogram counts the number of detected injections for each combination of pipelines. The unmodeled Burst search using cWB is labeled “cwb_allsky” while the BBH-model version of cWB is labeled “cwb_bbh”.

three primary components of the pipeline: (i) the search pipelines, (ii) the event orchestrator **GWcelery**, and (iii) **GraceDB**. Latency contributions from data calibration, frame construction, and inter-site transfer – typically taking about 5–10 s – are not included, nor are delays associated with alert ingestion and redistribution through GCN⁷ or SCIMMA⁸. The O3 MDC (Chaudhary et al. 2024) has shown that the LLAI system is capable of sending alerts at a median latency of ~ 30 s. We note that occasional technical issues during the LLPIC can lead to unusually high-latency outliers; therefore, the latencies reported here represent conservative estimates, excluding the time required for strain data preparation and transfer.

Event latencies are computed by comparing the time an event is created and appears on **GraceDB** to that of the known merger time. For GW event candidates during an observing run, the merger time is defined as the time the signal peak reaches Earth’s center.

For CBC pipelines, we find a median (90%) latency of 8.89 s (62.89 s); for Burst unmodeled pipelines, we find a median (90%) latency of 90.68 s (166.0 s). During O4 various search pipelines implemented a SNR-Optimized search (cite here) which is a refinement over the first

Latency Measure	Description	50% (s)	90% (s)
Superevents	$t_{\text{superevent}} - t_0$	8.37	49.77
CBC Events	$t_{\text{event}} - t_0$	8.89	62.89
CBC Events (SNR Optimized)	$t_{\text{event}} - t_0$	129.42	231.61
Burst Events (cWB-BBH)	$t_{\text{event}} - t_0$	70.88	123.60
Burst Events (unmodeled)	$t_{\text{event}} - t_0$	90.68	166.0
Early Warning Events	$t_{\text{event}} - t_0$	-3.87	5.98
GW Advocate Request	$t_{\text{ADV_REQ}} - t_0$	13.09	60.64
GCN Preliminary Sent	$t_{\text{GCN_PRELIM}} - t_0$	26.18	77.64

Table 3. Measured latencies for a number of steps in the pipeline. t_0 corresponds to the event merger time reported by the pipeline, while $t_{\text{superevent}}$ and t_{event} correspond to the time of superevent or event creation. For the case of superevent latencies, t_0 is determined by the preferred event.

CBC trigger. The median latency for the SNR-Optimized triggers stands at 129.42 s which makes them a candidate for second preliminary alert. For Early Warning alerts, shown on the right of Figure 10; we find a median (90%) latency of -3.1 s (5.98 s).

Within **GWcelery** the median (90%) time for creation of a superevent latency of 8.37 s (49.77 s). The median superevent latency is lower than the event latency due to the fact that the superevent may be created upon the first trigger, and that a superevent often consists of multiple events. Once a significant superevent has been created, there is a request for human vetting of the alert and the start of rapid response procedure, called the Advocate Request; we find a median (90%) latency of 13.09 s (60.64 s) to notify the advocate. To measure the latency of event communication to the community, we also measure the latency for sending of the GCN preliminary alert, which occurs for superevents that the pass significance threshold; we find a median (90%) latency of 26.18 s (77.64 s) for full bandwidth events. We show this statistic for GCN preliminary alerts on the left of Figure 10. This latency reported specifically measures that time until the GCN preliminary label is applied. During the entire one week cycle, only 23 significant early warning events were found and we found the median latency for public early warning alerts to be 6.95 s. Some early warning events were sent with negative latency with the earliest being -30.84 s.

We also compare this to the measured median GCN latency during O4 which, including data calibration, construction, and transfer time, is 30.4 s. The difference between GCN latency during the LLPIC and O4 is due to the fact that reak O4 data includes data calibration, construction, and transfer time latency, while the LLPIC measurement does not. This is smaller than previously

⁷ <https://gcn.nasa.gov/>

⁸ <https://scimma.org/>

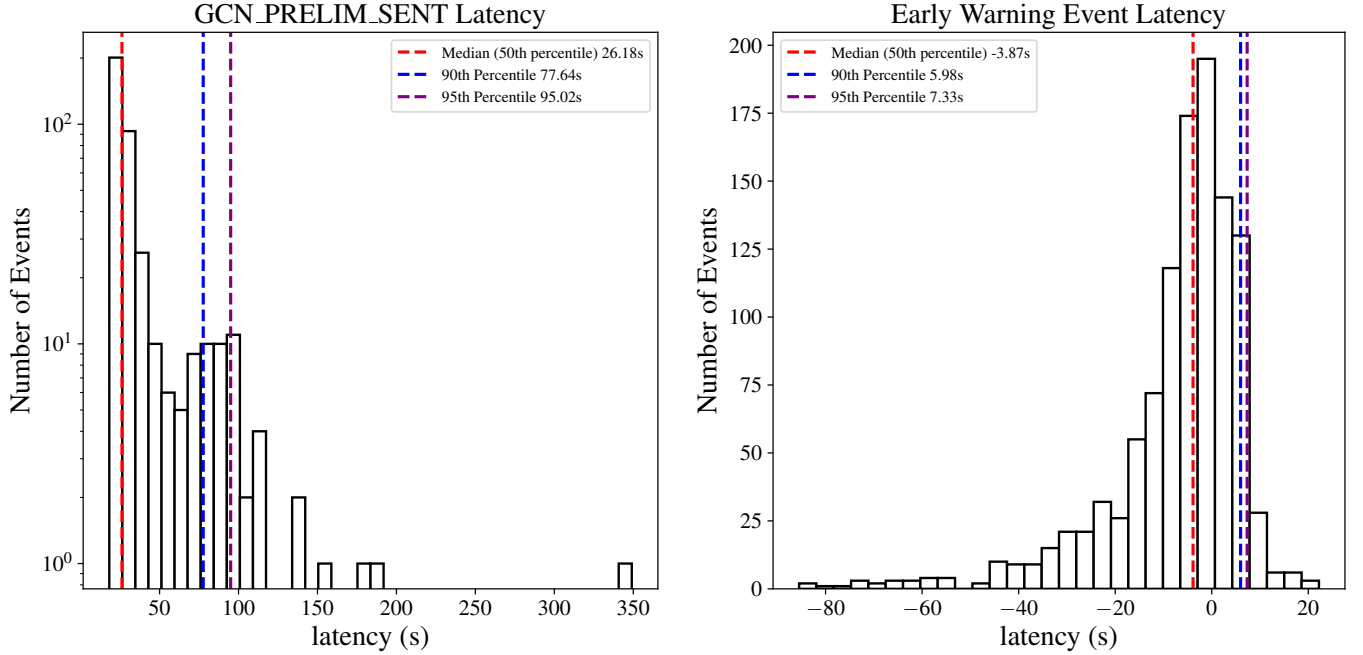


Figure 10. *Left:* Histogram of latencies for the sending of the GCN preliminary alert. We compare the LLPIC latencies to the median latency measured during O4. The O4 measurement includes data calibration, construction, and transfer time, while the LLPIC latencies do not. *Right* Histogram of latencies for Early Warning alerts. (add the O4 latency line on GCN Prelim sent)

reported number of 29.5s due to various changes made in `GraceDB` and `igwn-alert`. Table 3 shows a compilation of these latency statistics for comparison.

5.3.3. Probability of astrophysical origin

p_{astro} is defined as the probability that a particular GW trigger is caused by an astrophysical signal as opposed to the probability of being caused by a noise fluctuation or terrestrial contamination, $p_{\text{terrestrial}}$. p_{astro} is further divided into probability that the binary is a BNS, or NSBH or BBH given by p_{BNS} , p_{NSBH} , and p_{BBH} , respectively, to provide more information regarding the GW source class. Each GW superevent inherits p_{astro} estimate from its preferred CBC event. Burst unmodeled searches do not provide an estimate to p_{astro} . The p_{astro} mass parameter space has not yet been expanded onto SSM regions and thus SSM CBC searches report an empty p_{astro} .

By matching LLPIC’s injected parameters to the recovered p_{astro} values from the superevent, we can test the accuracy of p_{astro} . Figure 11 shows the recovered $P(\text{source})$ for true sources (e.g., $P(\text{BNS})$ for true injected BNS events), for superevents which pass the public alert FAR threshold. In matching these injections, we place the cut between NS and BH at $3 M_{\odot}$. “True” terrestrial events correspond to superevents which were not temporally matched to injections, indicating that they arose from detector noise. Injected BBHs are typically

recovered confidently, with the vast majority resulting in $P(\text{BBH}) > 0.5$. Over 90% of BBHs are recovered with $P(\text{BBH}) > 0.9$. The $P(\text{BNS})$ and $P(\text{NSBH})$ distributions are less confident than the $P(\text{BBH})$ distribution; about 20% of true BNSs and NSBHs are recovered with $P(\text{BNS})$ or $P(\text{NSBH}) < 0.1$. As a check against contamination across $P(\text{BNS})$ and $P(\text{NSBH})$ due to errors in recovered masses or variation in the definitions of the mass border between NS and BH (i.e., the Tolman–Oppenheimer–Volkoff mass), we can also look at the distribution of the sum of $P(\text{BNS})$ and $P(\text{NSBH})$ for injected events where $m_2 \leq 3M_{\odot}$. This is shown in the bottom panel of Figure 11.

Compared to $P(\text{BNS})$ or $P(\text{NSBH})$ alone, $P(\text{BNS}) + P(\text{NSBH})$ performs better, with $\sim 40\%$ of true BNS or NSBHs receiving $P(\text{BNS}) + P(\text{NSBH}) > 0.9$, compared to 10% and 20% for NSBH and BNS respectively. If instead we use a threshold $P(\text{source})$ of .5, we find a TPR of $\sim 98\%$ for BBH, $\sim 40\%$ for BNS, and $\sim 30\%$ for NSBH injections. (may be put this in appendix?)

(There is a significant drops in p-astro accuracy compared to O3 MDC. Need to explore this)

5.3.4. Sky Localization

One of the key data products to enable multi-messenger follow-up is the rapid inference of the sky localization from GW observations. This sky localization consists of the posterior probability distribution of the

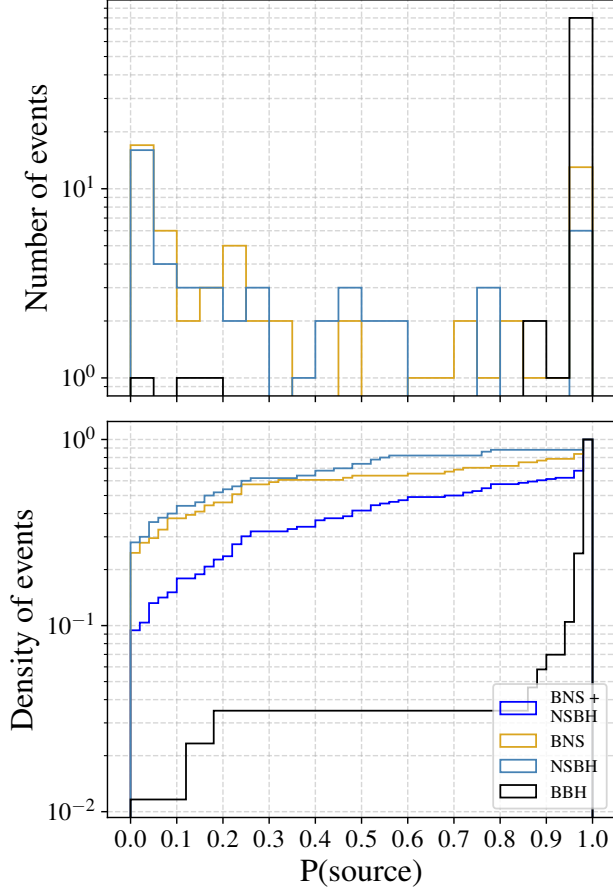


Figure 11. *Top:* Histogram of “preferred” recovered $P(\text{source})$ for true sources, for superevents which pass the significant public alert threshold. The possible source classes are BNS, NSBH, and BBH. This excludes early warning events, which were not fully functional during the time of this analysis. *Bottom:* Cumulative density of the same data. We also include the distribution of correctly recovered BNS or NSBH, which checks for contamination between the two classes due to misrecovered secondary masses or effects from varying the NS/BH mass boundary. As this is a cumulative histogram, the fraction of events above a certain $P(\text{source})$ corresponds to the TPR. We see that the majority of events with a $P(\text{source})$ greater than 0.5 correctly recover the injection source type.

source location in the sky. The sky localization, mapped either over a 2D map of right ascension and declination, or a 3D volume which also includes a distance estimate, is known as a “sky map.” Sky localization (and parameter estimation more generally) is conducted in multiple stages once a candidate is identified. For CBC sources, BAYESTAR, a rapid sky localization algorithm (Singer & Price 2016), is used to generate sky maps, and may be updated by Bilby, a python-based parameter estimation pipeline that uses stochastic sampling methods

(Ashton et al. 2019; Romero-Shaw et al. 2020b). Sky maps from BAYESTAR are released with Preliminary Notices and sky maps from Bilby are released in Update Notices. Burst pipelines, on the other hand, produce their own sky maps and upload them on to GraceDB.

In order to evaluate localization performance of the sky maps, the localization area, and searched area are computed. Localization area is the area (measured in deg^2) which encloses a given probability contour in the sky map; in this work, we use 90% as the total cumulative probability threshold. Searched area is the smallest 2D area, starting with the regions of highest probability, that contains the true location of the source; it represents a measurement of the sky area that a telescope with a small field of view relative to the sky map size would need to cover before imaging the true location. We refer the reader to Singer & Price (2016) for more details on these parameters, and use the `ligo.skymap`⁹ package to compute all metrics.

When evaluating skymap performance, we consider the preferred events for superevents that fall under the significant FAR threshold before trials factor and were detected by more than one interferometer. We exclude single detector triggers as most resulted from injections that occurred during a portion of LLPIC data where one or more detectors was not in science mode, and they may have sky localizations on the order of the entire sky.

Further, in the top left panel of Fig. 12, we show the accuracy of BAYESTAR sky maps through a P - P plot. Sky maps corresponding to the preferred event of a superevent that passing the current CBC FAR threshold of 1/6 month were considered. P - P plots of this format show the fraction of injections found within a given credible interval across all levels of credible intervals. The three gray regions around the diagonal shows the three different levels of confidence (1-3 σ) for the combined BAYESTAR map sample. We find that the BAYESTAR sky maps fall well within the credible intervals which implies a good accuracy. In the bottom left panel of Fig. 12, we show the cumulative trend of the searched area statistics from the combined sample. We find a median searched area of $100 - 200 \text{ deg}^2$. We then compare the two interferometer events with the three interferometer events, to show that the latter produced smaller searched areas. The numbers in the legend indicate the total number of events used in the plot.

Further, we probed the accuracy of the BAYESTAR sky maps compared with Bilby sky maps. Due to the dense rate of injections and limited resources, not all PE runs

⁹ <https://git.ligo.org/leo-singer/ligo.skymap>

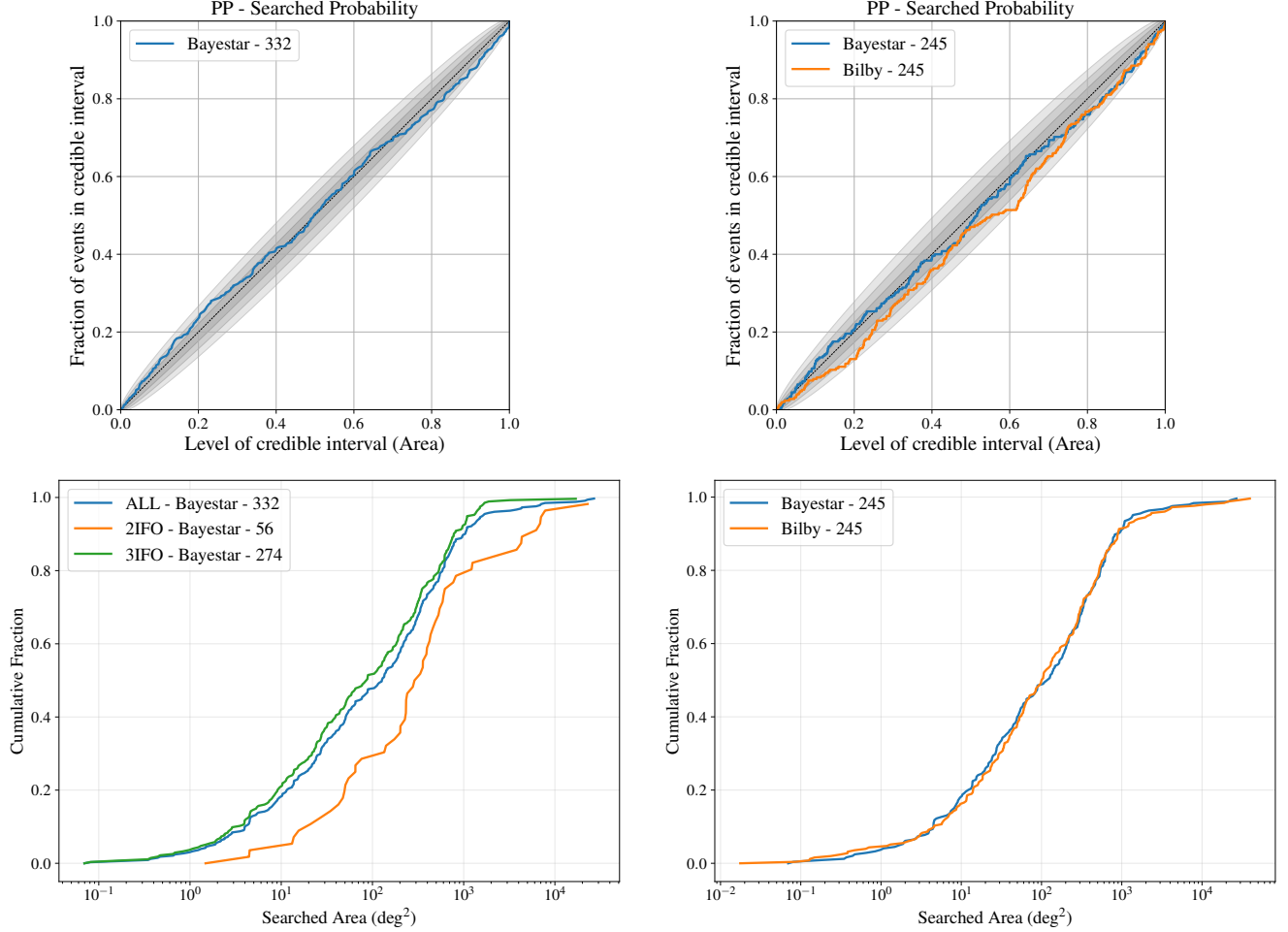


Figure 12. *Top left:* A P - P plot showing the BAYESTAR sky map statistics for the preferred event. The credible intervals shown in gray are based on the total number of events. *Bottom left:* Cumulative histograms of BAYESTAR searched area for all events (blue), compared to two-interferometer (Orange), three-interferometer (Green). *Top right:* A P - P plot showing the performance of BAYESTAR (blue) and Bilby (orange) generated sky maps for BNS events. The credible intervals shown in gray based on the total number of such preferred events where both BAYESTAR and Bilby sky maps are available. *Bottom right:* Cumulative histograms showing searched area statistics for BAYESTAR and Bilby sky maps.

were completed for significant superevents in LLPIC. The total number of Bilby runs crossmatched with injection during the fourth LLPIC cycle was found to be 245. We compare the P - P plot and searched area histogram in right side of Fig. 12 for the preferred event of all superevents for which both sky maps were produced. From the top right panel of Figure 12, we see that both BAYESTAR and Bilby sky maps include $\sim 100\%$ of events within their searched area. In some places the Bayesian Inference Library for CBC GW signal (Bilby) sky maps tend to sag below BAYESTAR's implying that their precision was slightly overstated as compared to Bilby. The cumulative searched area plot in the bottom right panel of Fig. 12 shows that both BAYESTAR and Bilby sky maps have comparable searched area. In the O3 Replay MDC, we saw that Bilby sky maps tend to have

lower searched area compared to BAYESTAR (Chaudhary et al. 2024). The comparable cumulative searched area is likely due to smaller statistics and generally larger SNR events compared to O3 MDC. The median SNR of events for which Bilby runs were completed is ~ 14 .

5.3.5. EM-Bright

EM-bright is a pipeline designed to assess binary source properties and whether a GW candidate is capable of producing an electromagnetic counterpart (Chatterjee et al. 2020). During O3, EM-Bright data products included HasNS and HasRemnant, where HasNS is the probability of the binary having a NS component, while HasRemnant is the probability of the merger leaving remnant matter post-merger in the form of dynamical or tidal ejecta. At the start of O4, another data

product **HasMassGap** was added which is the probability of the binary having at least one object in the lower mass gap regime between $3 - 5M_{\odot}$ and recently **HasSSM** data product (ref. Section 3) was also included for SSM searches. EM-Bright quantities are reported as a part of the automated and update discovery notices.

In order to access the performance of the EM-Bright data products in LLPIC, their Receiver Operating Characteristic (ROC) curves. For this, the NSBH boundary is chosen according to SLy, EoS but the probabilities are EoS marginalized, meaning we include some uncertainty in the NS EoS in our classifications. The markers indicate three different representative thresholds (a score above which events are considered to be positively classified as the source type in question) along the ROC curves for each pipeline.

In Fig. 13, we see that the **HasRemnant** quantity for all participating CBC pipelines has greater than a $\sim 90\%$ TPR for a $\sim 10\%$ FPR. In case of **HasNS**, we see it perform at $\sim 90\%$ TPR but at $\sim 30\%$ FPR. The FPR is significantly higher than what was observed in the O3 Replay MDC. Similarly, for **HasMassGap**, we see a drop in performance with $\sim 80\%$ TPR at $\sim 40\%$ FPR.

Fig. 14 is a scatter plot of source m_1 and m_2 depicting misclassified triggers. A trigger is considered misclassified if its **HasNS** score incorrectly crosses the decision threshold of 0.5 – that is, when the score is below 0.5 for a true NS injection (ground truth = 1) or above 0.5 for a non-NS injection (ground truth = 0). We can see that majority of the triggers that are misclassified are between $2-3 M_{\odot}$. Since SLy EoS is considered as truth for classifying a NS, all the injections above $2.05 M_{\odot}$ are not considered as NS. This causes an increase in FPR which is clearly shown by the ROC curve. If we choose a more stiff EoS with maximum NS masses around 2.8 , the FPR reduces by $\sim 20\%$ indicating that the classifier’s performance near the NS boundary depends on selection of NS EoS. In the O3 replay MDC, the NS injection boundary was set to $2.05 M_{\odot}$ which resulted in many fewer triggers in the region between $2-5 M_{\odot}$ consequently the change in **HasNS** performance was not observed (reword, what does this mean?). However, since the classifier were trained on an O2 era dataset, this drop in performance suggests retraining the classifier with latest dataset might help in improving the classification performance.

5.3.6. Chirp mass

Fig. 15 shows a comparison of the highest probability chirp mass bin distribution based on low-latency estimates with the injected chirp mass for the significant CBC events found in LLPIC cycle 4. Ideally, the high-

est probability chirp mass bin distribution should match injection chirp mass distribution. We see that the mismatch between the two is limited to a small number of events in each bins. This shows that the low-latency binned chirp mass estimates are performing fairly accurately for the LLPIC dataset.

6. CONCLUSION

This work summarizes key advances made in the LLAI during O4 of the LVK. We introduce the deployment of two new low-latency public data products – notably **HasSSM**, a machine learning classifier quantifying the probability of a SSM compact object, and binned chirp mass estimates. Both enhancements significantly expand the astrophysical and multi-messenger scope of LVK’s real-time alert infrastructure.

We also describe the LLPIC – a comprehensive, end-to-end validation framework designed to test and strengthen the LVK’s LLAI ahead of the next observing runs. LLPIC integrates realistic simulated CBC and Burst GW signals into archived O4 data and replays them through the entire alert generation pipeline. This approach enables systematic testing of new data products, latency performance, and cross-pipeline consistency under production-like conditions. (Do we mention Rubin MDC?)

Public, low-latency chirp mass estimates are a significant step forward for enabling multi-messenger searches, as the mass information they provide can be used to characterize the source and possible counterparts. We intend for these mass estimates, which are provided in the form binned probabilities and plotted and uploaded to **GraceDB**, to be ingested alongside source classification by all consumers of our alerts to gain a clearer picture of the candidate events we detect and enable both ToO searches and downstream scientific efforts. S250818k has already demonstrated the potential usefulness of this data product by highlighting the standout nature of this candidate event, and more broadly, by estimating that it is likely a low-mass **HasNS**-type event, if it is in fact astrophysical. In the future, we intend for these mass estimates to identify many more candidates of interest, whether they be low mass BNS or high mass BBH.

Performance analyses on LLPIC shows that that low-latency data products (**HasNS** and p_{astro}) shows reduced performance compared to the previous MDC. However, this can be attributed to the peculiar injection distribution of LLPIC and a very small dataset. Tuning these data products more in the NS-BH boundaries would help increase the performance in the upcoming observing runs. Other data products like sky-localization (**BAYESTAR** and **Bilby**) and the new chirp

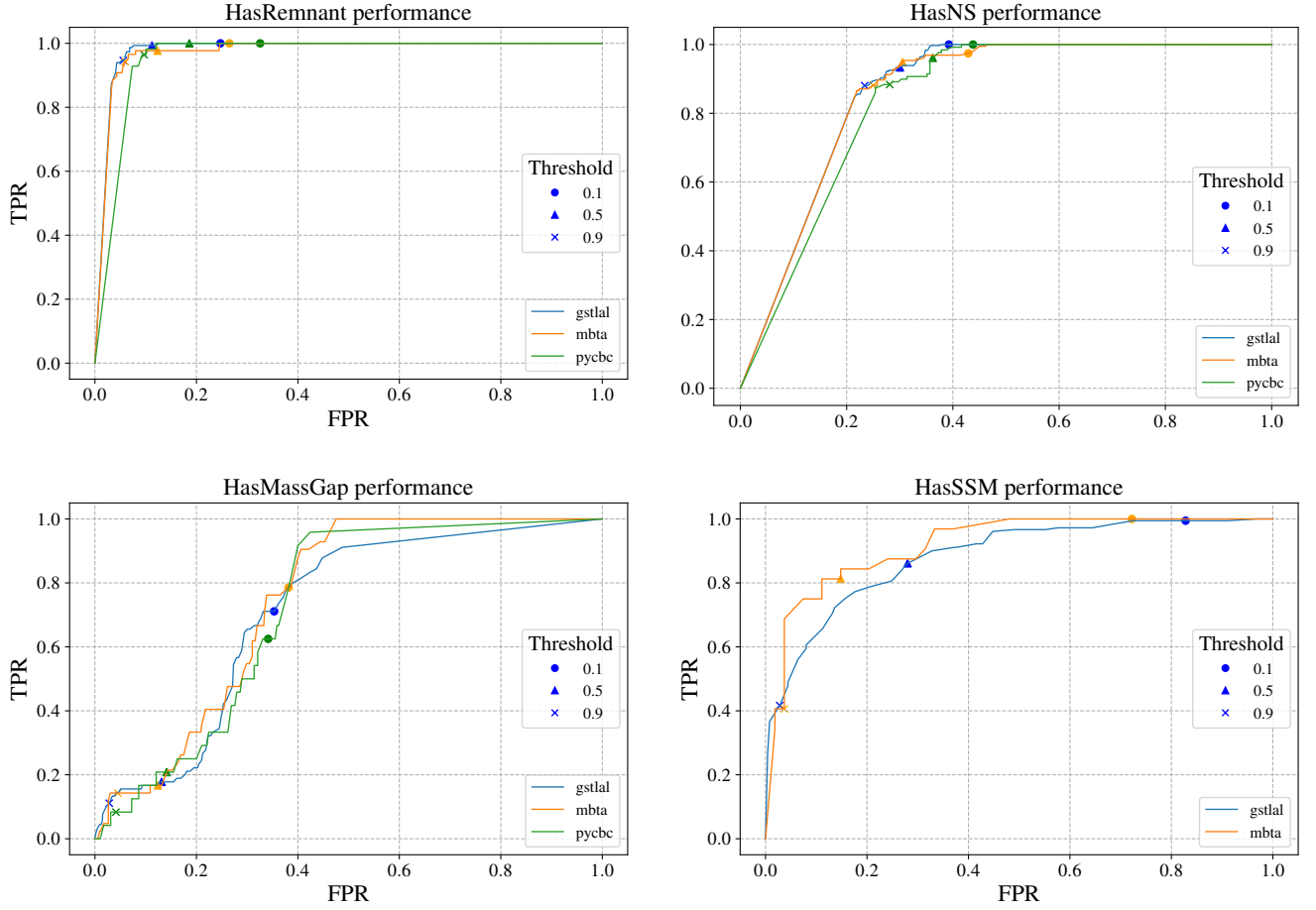


Figure 13. The ROC curves for the different EM-Bright classifiers are shown here for LLPIC events. The *HasRemnant*, *HasNS*, and *HasMassGap* performances are evaluated on regular CBC alerts whereas *HasSSM* is computed for CBC SSM searches only. The markers denote different representative thresholds along the curve. (Missing fpr tpr in hasns plots)

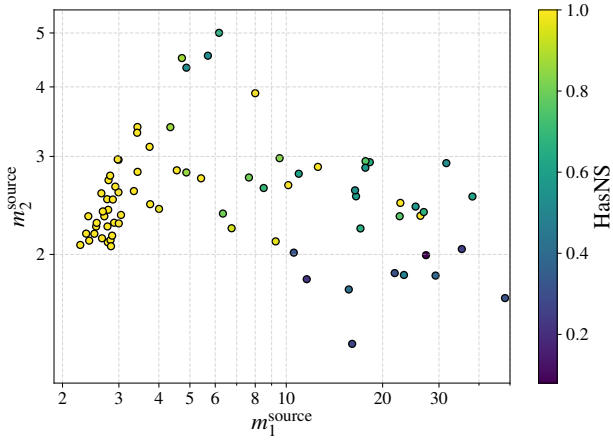


Figure 14. Primary and secondary mass of significant LLPIC triggers with misclassified *HasNS*.

mass estimates are performing as expected for O4 sensitivity. Compared to O3 Replay MDC, we see a slight decrease in the median latency for preliminary alerts

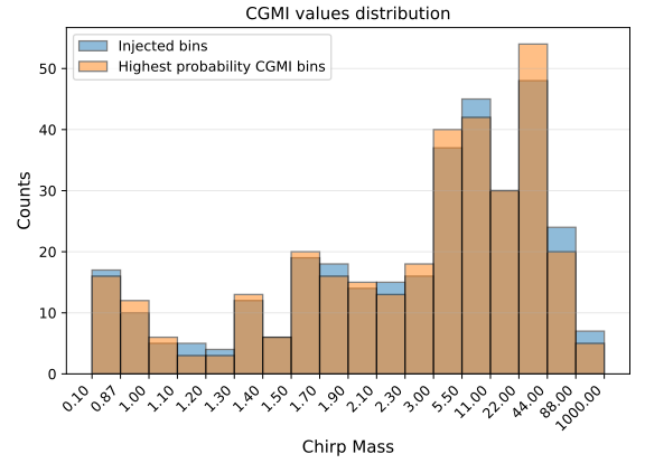


Figure 15. Comparison of the highest probable chirp mass bin distribution from low-latency estimates with the injections chirp mass. (remove CGMI)

attributed to continuous developments in the *GraceDB*

and `igwn-alert` frameworks. Burst–CBC overlap tests further underscore the need for careful classification of short-duration transients, which can be misidentified across search pipelines.

We also developed and integrated the MEG, which simulates live replay of events through `GraceDB` and `GWCelery`. MEG enables robust verification of data integrity, latency, and consistency across all alert components. The combination of LLPIC and MEG form the foundation for a continuous integration–style testing framework, ensuring that every new alert feature—such as early-warning alerts, low-latency classification, and machine-learning-based vetting—is validated automat-

ically prior to deployment. These developments mark a crucial step toward a more automated, reliable, and scientifically rich GW alert infrastructure as the LVK prepares for O5 and future observing runs.

- 1 We thank (fill in) for review of this paper.
- 2 A.T and M.W.C. acknowledge support from the Na-
- 3 tional Science Foundation with grant numbers PHY-
- 4 2117997, PHY-2308862 and PHY-2409481. AT acknowl-
- 5 edges support from the Director’s Postdoc Fellowship at
- 6 Los Alamos National Laboratory. This material is based
- 7 upon work supported by NSF’s LIGO Laboratory which
- 8 is a major facility fully funded by the National Science
- 9 Foundation.

APPENDIX

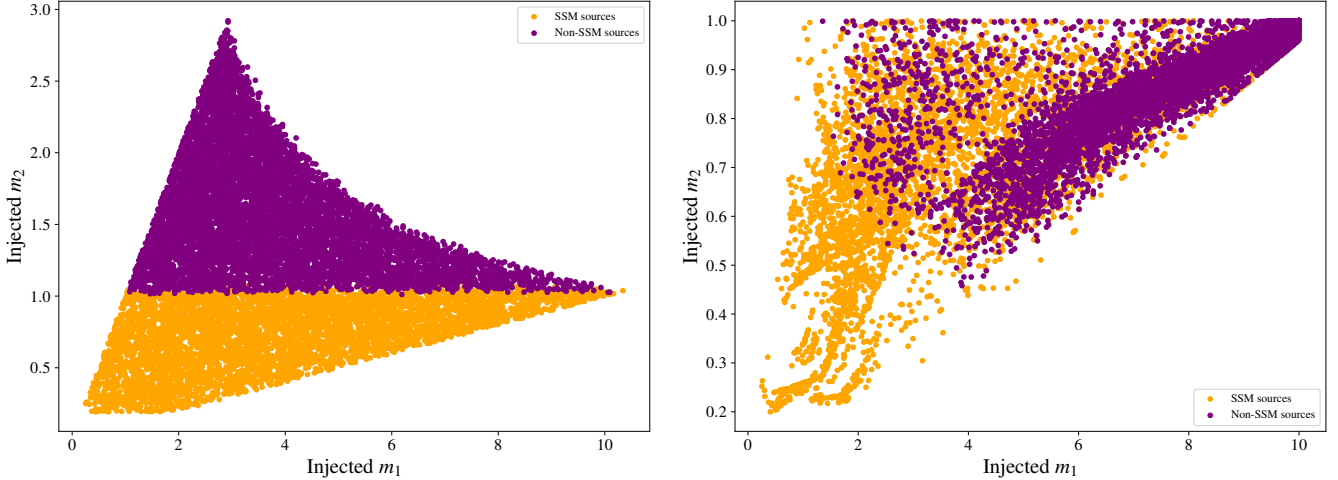


Figure 16. Left: m_1 and m_2 scatter plot of the injection dataset successfully recovered by MBTA pipeline. Right: m_1 and m_2 scatter plot of the injection dataset successfully recovered by MBTA pipeline. The colors show whether the injection was in SSM mass space or otherwise. We see significant number of non-SSM injections being recovered by the SSM search with template parameters in the SSM region.

REFERENCES

- Aasi et al. 2015, *Classical and Quantum Gravity*, 32, 074001
- Abac, A. G., et al. 2025a. <https://arxiv.org/abs/2508.18081>
- . 2025b. <https://arxiv.org/abs/2507.12374>
- Abadie, J., et al. 2012, *Astrophys. J.*, 760, 12, doi: [10.1088/0004-637X/760/1/12](https://doi.org/10.1088/0004-637X/760/1/12)
- Abbott, B. P., et al. 2017a, *Phys. Rev. Lett.*, 119, 161101, doi: [10.1103/PhysRevLett.119.161101](https://doi.org/10.1103/PhysRevLett.119.161101)
- . 2017b, *The Astrophysical Journal Letters*, 848, L12, doi: [10.3847/2041-8213/aa91c9](https://doi.org/10.3847/2041-8213/aa91c9)
- . 2017c, *The Astrophysical Journal Letters*, 848, L13. <http://stacks.iop.org/2041-8205/848/i=2/a=L13>
- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, *Nature*, 551, 85, doi: [10.1038/nature24471](https://doi.org/10.1038/nature24471)
- Abbott, R., Abbott, T. D., Acernese, F., et al. 2022, *Phys. Rev. Lett.*, 129, 061104, doi: [10.1103/PhysRevLett.129.061104](https://doi.org/10.1103/PhysRevLett.129.061104)
- Abbott, R., et al. 2020, *Astrophys. J. Lett.*, 900, L13, doi: [10.3847/2041-8213/aba493](https://doi.org/10.3847/2041-8213/aba493)
- Abbott, R., Abbott, T. D., Abraham, S., et al. 2021, *The Astrophysical Journal Letters*, 915, L5, doi: [10.3847/2041-8213/ac082e](https://doi.org/10.3847/2041-8213/ac082e)
- Abbott et al. 2016, *Astrophys. J.*, 826, L13, doi: [10.3847/2041-8205/826/1/L13](https://doi.org/10.3847/2041-8205/826/1/L13)
- Abdalla, H., et al. 2021, *Astrophys. J.*, 923, 109, doi: [10.3847/1538-4357/ac2e04](https://doi.org/10.3847/1538-4357/ac2e04)
- Abdikamalov, E., Pagliaroli, G., & Radice, D. 2020, doi: [10.1007/978-981-15-4702-7_21-1](https://doi.org/10.1007/978-981-15-4702-7_21-1)
- Acernese et al. 2015, *Classical and Quantum Gravity*, 32, 024001
- Adams, T., Buskulic, D., Germain, V., et al. 2016, *Class. Quant. Grav.*, 33, 175012, doi: [10.1088/0264-9381/33/17/175012](https://doi.org/10.1088/0264-9381/33/17/175012)
- Agayeva, S., et al. 2022, *Proc. SPIE Int. Soc. Opt. Eng.*, 12186, 121861H, doi: [10.1117/12.2630240](https://doi.org/10.1117/12.2630240)
- Ahumada, T., et al. 2024, *Publ. Astron. Soc. Pac.*, 136, 114201, doi: [10.1088/1538-3873/ad8265](https://doi.org/10.1088/1538-3873/ad8265)
- All  n  , C., et al. 2025, *Class. Quant. Grav.*, 42, 105009, doi: [10.1088/1361-6382/add234](https://doi.org/10.1088/1361-6382/add234)
- All  n  , C., Aubin, F., Bentara, I., et al. 2025, *The MBTA Pipeline for Detecting Compact Binary Coalescences in the Fourth LIGO-Virgo-KAGRA Observing Run*. <https://arxiv.org/abs/2501.04598>
- Annala, E., Gorda, T., Kurkela, A., & Vuorinen, A. 2018, *Phys. Rev. Lett.*, 120, 172703, doi: [10.1103/PhysRevLett.120.172703](https://doi.org/10.1103/PhysRevLett.120.172703)
- Ashton, G., H  bner, M., Lasky, P. D., et al. 2019, *ApJS*, 241, 27, doi: [10.3847/1538-4365/ab06fc](https://doi.org/10.3847/1538-4365/ab06fc)
- Aubin, F., et al. 2021, *Class. Quant. Grav.*, 38, 095004, doi: [10.1088/1361-6382/abe913](https://doi.org/10.1088/1361-6382/abe913)
- Bauswein, A., Just, O., Janka, H.-T., & Stergioulas, N. 2017, *ApJL*, 850, L34, doi: [10.3847/2041-8213/aa9994](https://doi.org/10.3847/2041-8213/aa9994)
- Bellm, E. 2014, in *The Third Hot-wiring the Transient Universe Workshop*, ed. P. R. Wozniak, M. J. Graham, A. A. Mahabal, & R. Seaman, 27–33. <https://arxiv.org/abs/1410.8185>
- Berbel, M., Miravet-Ten  s, M., Sharma Chaudhary, S., et al. 2024, *Classical and Quantum Gravity*, 41, 085012, doi: [10.1088/1361-6382/ad3279](https://doi.org/10.1088/1361-6382/ad3279)
- Bini, S., Tiwari, S., Xu, Y., et al. 2024, *Phys. Rev. D*, 109, 042009, doi: [10.1103/PhysRevD.109.042009](https://doi.org/10.1103/PhysRevD.109.042009)
- Bird, S., Cholis, I., Mu  oz, J. B., et al. 2016, *Phys. Rev. Lett.*, 116, 201301, doi: [10.1103/PhysRevLett.116.201301](https://doi.org/10.1103/PhysRevLett.116.201301)
- Buonanno, A., Iyer, B. R., Ochsner, E., Pan, Y., & Sathyaprakash, B. S. 2009, *Phys. Rev. D*, 80, 084043, doi: [10.1103/PhysRevD.80.084043](https://doi.org/10.1103/PhysRevD.80.084043)
- Cannon, K., Hanna, C., & Keppel, D. 2012, *Phys. Rev. D*, 85, 081504, doi: [10.1103/PhysRevD.85.081504](https://doi.org/10.1103/PhysRevD.85.081504)
- Cannon, K., et al. 2020. <https://arxiv.org/abs/2010.05082>
- Cardoso, V., & Pani, P. 2019, *Living Reviews in Relativity*, 22, 4, doi: [10.1007/s41114-019-0020-4](https://doi.org/10.1007/s41114-019-0020-4)
- Carr, B., K  hnel, F., & Sandstad, M. 2016, *Phys. Rev. D*, 94, 083504, doi: [10.1103/PhysRevD.94.083504](https://doi.org/10.1103/PhysRevD.94.083504)
- Carr, B. J., & Hawking, S. W. 1974, *Monthly Notices of the Royal Astronomical Society*, 168, 399, doi: [10.1093/mnras/168.2.399](https://doi.org/10.1093/mnras/168.2.399)
- Chatterjee, D., Ghosh, S., Brady, P. R., et al. 2020, *The Astrophysical Journal*, 896, 54, doi: [10.3847/1538-4357/ab8dbe](https://doi.org/10.3847/1538-4357/ab8dbe)
- Chaudhary, S. S., Toivonen, A., et al. 2024, *Proc. Nat. Acad. Sci.*, 121, e2316474121, doi: [10.1073/pnas.2316474121](https://doi.org/10.1073/pnas.2316474121)
- Chornock et al. 2017, *The Astrophysical Journal Letters*, 848, L19. <http://stacks.iop.org/2041-8205/848/i=2/a=L19>
- Chu, Q., Kovalam, M., Wen, L., et al. 2021, *The SPIIR online coherent pipeline to search for gravitational waves from compact binary coalescences*. <https://arxiv.org/abs/2011.06787>
- Clesse, S., & Garc  a-Bellido, J. 2018, *Seven Hints for Primordial Black Hole Dark Matter*. <https://arxiv.org/abs/1711.10458>
- Contopoulos, I., Kazanas, D., & Papadopoulos, D. B. 2023, *Mon. Not. Roy. Astron. Soc.*, 527, 11198, doi: [10.1093/mnras/stad3913](https://doi.org/10.1093/mnras/stad3913)

- Coughlin, M. W., Dietrich, T., Margalit, B., & Metzger, B. D. 2019a, *Monthly Notices of the Royal Astronomical Society: Letters*, 489, L91, doi: [10.1093/mnrasl/slz133](https://doi.org/10.1093/mnrasl/slz133)
- Coughlin, M. W., Dietrich, T., Doctor, Z., et al. 2018, *Monthly Notices of the Royal Astronomical Society*, 480, 3871, doi: [10.1093/mnras/sty2174](https://doi.org/10.1093/mnras/sty2174)
- Coughlin, M. W., Dietrich, T., Antier, S., et al. 2019b, *Monthly Notices of the Royal Astronomical Society*, 492, 863, doi: [10.1093/mnras/stz3457](https://doi.org/10.1093/mnras/stz3457)
- Coughlin, M. W., et al. 2019c, *Astrophys. J. Lett.*, 885, L19, doi: [10.3847/2041-8213/ab4ad8](https://doi.org/10.3847/2041-8213/ab4ad8)
- Coughlin, M. W., Dietrich, T., Heinzel, J., et al. 2020a, *Phys. Rev. Research*, 2, 022006, doi: [10.1103/PhysRevResearch.2.022006](https://doi.org/10.1103/PhysRevResearch.2.022006)
- Coughlin, M. W., Antier, S., Dietrich, T., et al. 2020b, *Nature Communications*, 11, 4129, doi: [10.1038/s41467-020-17998-5](https://doi.org/10.1038/s41467-020-17998-5)
- Coulter, D. A., Foley, R. J., Kilpatrick, C. D., et al. 2017, *Science*, 358, 1556, doi: [10.1126/science.aap9811](https://doi.org/10.1126/science.aap9811)
- Coulter et al. 2017, *Science*, 358, 1556, doi: [10.1126/science.aap9811](https://doi.org/10.1126/science.aap9811)
- Cowperthwaite, P. S., Berger, E., Villar, V. A., et al. 2017, *ApJL*, 848, L17, doi: [10.3847/2041-8213/aa8fc7](https://doi.org/10.3847/2041-8213/aa8fc7)
- Dal Canton, T., Nitz, A. H., Gadre, B., et al. 2021, *Astrophys. J.*, 923, 254, doi: [10.3847/1538-4357/ac2f9a](https://doi.org/10.3847/1538-4357/ac2f9a)
- Damour, T., & Vilenkin, A. 2000, *Phys. Rev. Lett.*, 85, 3761, doi: [10.1103/PhysRevLett.85.3761](https://doi.org/10.1103/PhysRevLett.85.3761)
- . 2001, *Phys. Rev. D*, 64, 064008, doi: [10.1103/PhysRevD.64.064008](https://doi.org/10.1103/PhysRevD.64.064008)
- . 2005, *Phys. Rev. D*, 71, 063510, doi: [10.1103/PhysRevD.71.063510](https://doi.org/10.1103/PhysRevD.71.063510)
- Dietrich, T., Coughlin, M. W., Pang, P. T. H., et al. 2020, *Science*, 370, 1450, doi: [10.1126/science.abb4317](https://doi.org/10.1126/science.abb4317)
- Drago, M., Klimentko, S., Lazzaro, C., et al. 2021, *SoftwareX*, 14, 100678
- Ebersold, M., & Tiwari, S. 2020, *Phys. Rev. D*, 101, 104041, doi: [10.1103/PhysRevD.101.104041](https://doi.org/10.1103/PhysRevD.101.104041)
- Ewing, B., Huxford, R., Singh, D., et al. 2024, *Phys. Rev. D*, 109, 042008, doi: [10.1103/PhysRevD.109.042008](https://doi.org/10.1103/PhysRevD.109.042008)
- Foucart, F., Hinderer, T., & Nissanke, S. 2018, *Phys. Rev. D*, 98, 081501, doi: [10.1103/PhysRevD.98.081501](https://doi.org/10.1103/PhysRevD.98.081501)
- Franz, N., et al. 2025. <https://arxiv.org/abs/2510.17104>
- Giliberti, E., & Cambiotti, G. 2022, *Mon. Not. Roy. Astron. Soc.*, 511, 3365, doi: [10.1093/mnras/stac245](https://doi.org/10.1093/mnras/stac245)
- Gillanders, J. H., et al. 2025. <https://arxiv.org/abs/2510.01142>
- Goldstein, A., Veres, P., Burns, E., et al. 2017, *The Astrophysical Journal*, 848, L14, doi: [10.3847/2041-8213/aa8f41](https://doi.org/10.3847/2041-8213/aa8f41)
- Graham, M. J., et al. 2023, *Astrophys. J.*, 942, 99, doi: [10.3847/1538-4357/aca480](https://doi.org/10.3847/1538-4357/aca480)
- Hanna, C., Kennington, J., Niu, W., et al. 2024, *Template bank for sub solar mass compact binary mergers in the fourth observing run of Advanced LIGO, Advanced Virgo, and KAGRA.* <https://arxiv.org/abs/2412.10951>
- Hanna, C., et al. 2020, *Phys. Rev. D*, 101, 022003, doi: [10.1103/PhysRevD.101.022003](https://doi.org/10.1103/PhysRevD.101.022003)
- Hotokezaka, K., Nakar, E., Gottlieb, O., et al. 2018. <https://arxiv.org/abs/1806.10596>
- Husa, S., Khan, S., Hannam, M., et al. 2016, *Phys. Rev. D*, 93, 044006, doi: [10.1103/PhysRevD.93.044006](https://doi.org/10.1103/PhysRevD.93.044006)
- Huth, S., et al. 2022, *Nature*, 606, 276, doi: [10.1038/s41586-022-04750-w](https://doi.org/10.1038/s41586-022-04750-w)
- Janka, H.-T. 2012, *Ann. Rev. Nucl. Part. Sci.*, 62, 407, doi: [10.1146/annurev-nucl-102711-094901](https://doi.org/10.1146/annurev-nucl-102711-094901)
- Kasliwal, M. M., Kasen, D., Lau, R. M., et al. 2019, *Monthly Notices of the Royal Astronomical Society: Letters*, doi: [10.1093/mnrasl/slz007](https://doi.org/10.1093/mnrasl/slz007)
- Kasliwal, M. M., et al. 2025. <https://arxiv.org/abs/2510.23732>
- Khan, S., Husa, S., Hannam, M., et al. 2016, *Phys. Rev. D*, 93, 044007, doi: [10.1103/PhysRevD.93.044007](https://doi.org/10.1103/PhysRevD.93.044007)
- Klimentko, S., Yakushin, I., Mercer, A., & Mitselmakher, G. 2008, *Classical and Quantum Gravity*, 25, 114029
- Klimentko, S., Vedovato, G., Drago, M., et al. 2016a, *Phys. Rev. D*, 93, 042004, doi: [10.1103/PhysRevD.93.042004](https://doi.org/10.1103/PhysRevD.93.042004)
- Klimentko, S., et al. 2016b, *Phys. Rev. D*, 93, 042004, doi: [10.1103/PhysRevD.93.042004](https://doi.org/10.1103/PhysRevD.93.042004)
- Kouvaris, C. 2025, *Int. J. Mod. Phys. D*, 34, 2550037, doi: [10.1142/S0218271825500373](https://doi.org/10.1142/S0218271825500373)
- Kouvaris, C., & Nielsen, N. G. 2015, *Phys. Rev. D*, 92, 063526, doi: [10.1103/PhysRevD.92.063526](https://doi.org/10.1103/PhysRevD.92.063526)
- Kovalam, M., Patwary, M. A. K., Sreekumar, A. K., et al. 2022, *The Astrophysical Journal Letters*, 927, L9, doi: [10.3847/2041-8213/ac5687](https://doi.org/10.3847/2041-8213/ac5687)
- Lai, X., Zhou, E., & Xu, R. 2019, *The European Physical Journal A*, 55, 60, doi: [10.1140/epja/i2019-12720-8](https://doi.org/10.1140/epja/i2019-12720-8)
- Legred, I., Chatziioannou, K., Essick, R., Han, S., & Landry, P. 2021, *Phys. Rev. D*, 104, 063003, doi: [10.1103/PhysRevD.104.063003](https://doi.org/10.1103/PhysRevD.104.063003)
- . 2022, *Impact of the PSR J0740+6620 radius constraint on the properties of high-density matter: Neutron star equation of state posterior samples*, Zenodo, doi: [10.5281/zenodo.6502467](https://doi.org/10.5281/zenodo.6502467)
- Margalit, B., & Metzger, B. 2017, *The Astrophysical Journal Letters*, 850, doi: [10.3847/2041-8213/aa991c](https://doi.org/10.3847/2041-8213/aa991c)
- Margalit, B., & Metzger, B. D. 2019, *Astrophys. J. Lett.*, 880, L15, doi: [10.3847/2041-8213/ab2ae2](https://doi.org/10.3847/2041-8213/ab2ae2)

- Marx, E., et al. 2025, *Phys. Rev. D*, 111, 042010, doi: [10.1103/PhysRevD.111.042010](https://doi.org/10.1103/PhysRevD.111.042010)
- Messick, C., et al. 2017, *Phys. Rev. D*, 95, 042001, doi: [10.1103/PhysRevD.95.042001](https://doi.org/10.1103/PhysRevD.95.042001)
- Metzger, B. D., Hui, L., & Cantiello, M. 2024, *Astrophys. J. Lett.*, 971, L34, doi: [10.3847/2041-8213/ad6990](https://doi.org/10.3847/2041-8213/ad6990)
- Mezzacappa, A., et al. 2023, *Phys. Rev. D*, 107, 043008, doi: [10.1103/PhysRevD.107.043008](https://doi.org/10.1103/PhysRevD.107.043008)
- Most, E. R., Weih, L. R., Rezzolla, L., & Schaffner-Bielich, J. 2018, *Phys. Rev. Lett.*, 120, 261103, doi: [10.1103/PhysRevLett.120.261103](https://doi.org/10.1103/PhysRevLett.120.261103)
- Nazari, E., & Roshan, M. 2020, *Mon. Not. Roy. Astron. Soc.*, 498, 110, doi: [10.1093/mnras/staa2322](https://doi.org/10.1093/mnras/staa2322)
- Nitz, A. H., Dal Canton, T., Davis, D., & Reyes, S. 2018, *Phys. Rev. D*, 98, 024050, doi: [10.1103/PhysRevD.98.024050](https://doi.org/10.1103/PhysRevD.98.024050)
- Paek, G. S. H., et al. 2024, *Astrophys. J.*, 960, 113, doi: [10.3847/1538-4357/ad0238](https://doi.org/10.3847/1538-4357/ad0238)
- Pan, K.-C., Liebendörfer, M., Couch, S. M., & Thielemann, F.-K. 2021, *Astrophys. J.*, 914, 140, doi: [10.3847/1538-4357/abfb05](https://doi.org/10.3847/1538-4357/abfb05)
- Paterson, K., et al. 2021, *Astrophys. J.*, 912, 128, doi: [10.3847/1538-4357/abeb71](https://doi.org/10.3847/1538-4357/abeb71)
- Pian et al. 2017, *Nature*, 551, 67 EP . <http://dx.doi.org/10.1038/nature24298>
- Piro, A. L., & Pfahl, E. 2007, *Astrophys. J.*, 658, 1173, doi: [10.1086/511672](https://doi.org/10.1086/511672)
- Planck Collaboration. 2020, *Astronomy and Astrophysics*, 641, A6, doi: [10.1051/0004-6361/201833910](https://doi.org/10.1051/0004-6361/201833910)
- Powell, J., & Müller, B. 2019, *Mon. Not. Roy. Astron. Soc.*, 487, 1178, doi: [10.1093/mnras/stz1304](https://doi.org/10.1093/mnras/stz1304)
- Pratten, G., Husa, S., García-Quirós, C., et al. 2020, *Phys. Rev. D*, 102, 064001, doi: [10.1103/PhysRevD.102.064001](https://doi.org/10.1103/PhysRevD.102.064001)
- Radice, D., Perego, A., Zappa, F., & Bernuzzi, S. 2018, *The Astrophysical Journal Letters*, 852, L29. <http://stacks.iop.org/2041-8205/852/i=2/a=L29>
- Romero-Shaw, I. M., et al. 2020a, *Monthly Notices of the Royal Astronomical Society*, 499, 3295–3319, doi: [10.1093/mnras/staa2850](https://doi.org/10.1093/mnras/staa2850)
- . 2020b, *Mon. Not. Roy. Astron. Soc.*, 499, 3295, doi: [10.1093/mnras/staa2850](https://doi.org/10.1093/mnras/staa2850)
- Rosswog, S., Feindt, U., Korobkin, O., et al. 2017, *Class. Quant. Grav.*, 34, 104001, doi: [10.1088/1361-6382/aa68a9](https://doi.org/10.1088/1361-6382/aa68a9)
- Ryan Magee, Deep Chatterjee, L. P. S. S. M. K. G. M. S. A. P. B. P. B. K. C. T. D. C. Q. C. P. C. A. C. M. D. P. G. S. G. G. G. C. H. S. J. K. E. K. V. O. A. E. P. F. P. A. P. R. D. P. B. P. T. P. L. R. A. K. S. M. J. S. V. V. A. V. M. W. L. W., & Zweizig, J. 2021, *The Astrophysical Journal*, 910, L21, doi: [10.3847/2041-8213/abed54](https://doi.org/10.3847/2041-8213/abed54)
- Sachdev, S., et al. 2019. <https://arxiv.org/abs/1901.08580>
- Sakon, S., et al. 2022, Template bank for compact binary mergers in the fourth observing run of Advanced LIGO, Advanced Virgo, and KAGRA. <https://arxiv.org/abs/2211.16674>
- Savchenko, V., Ferrigno, C., Kuulkers, E., et al. 2017, *The Astrophysical Journal*, 848, L15, doi: [10.3847/2041-8213/aa8f94](https://doi.org/10.3847/2041-8213/aa8f94)
- Schunck, F. E., & Mielke, E. W. 2003, *Classical and Quantum Gravity*, 20, R301, doi: [10.1088/0264-9381/20/20/201](https://doi.org/10.1088/0264-9381/20/20/201)
- Shandera, S., Jeong, D., & Gebhardt, H. S. G. 2018, *Phys. Rev. Lett.*, 120, 241102, doi: [10.1103/PhysRevLett.120.241102](https://doi.org/10.1103/PhysRevLett.120.241102)
- Singer, L. P., & Price, L. R. 2016, *Phys. Rev. D*, 93, 024013, doi: [10.1103/PhysRevD.93.024013](https://doi.org/10.1103/PhysRevD.93.024013)
- Singer, L. P., & Price, L. R. 2016, *PhRvD*, 93, 024013, doi: [10.1103/PhysRevD.93.024013](https://doi.org/10.1103/PhysRevD.93.024013)
- Skiridis, V., Norman, M. R. K., & Sutton, P. J. 2024, Real-Time Detection of Unmodelled Gravitational-Wave Transients Using Convolutional Neural Networks. <https://arxiv.org/abs/2009.14611>
- Smartt et al. 2017, *Nature*, 551, 75 EP . <http://dx.doi.org/10.1038/nature24303>
- Takahashi, I., et al. 2025. <https://arxiv.org/abs/2510.15534>
- Tsukada, L., et al. 2023. <https://arxiv.org/abs/2305.06286>
- Vajpeyi, A., Smith, R., Thrane, E., et al. 2022, *Mon. Not. Roy. Astron. Soc.*, 516, 5309, doi: [10.1093/mnras/stac2332](https://doi.org/10.1093/mnras/stac2332)
- Watson, D., Hansen, C. J., Selsing, J., et al. 2019, *Nature*, 574, 497, doi: [10.1038/s41586-019-1676-3](https://doi.org/10.1038/s41586-019-1676-3)
- Watts, A. L., & Strohmayer, T. E. 2007, *Adv. Space Res.*, 40, 1446, doi: [10.1016/j.asr.2006.12.021](https://doi.org/10.1016/j.asr.2006.12.021)