The Astrophysical Journal, 927:50 (16pp), 2022 March 1 https://doi.org/10.3847/1538-4357/ac4d34

A Systematic Exploration of Kilonova Candidates from Neutron Star Mergers during the Third Gravitational-wave Observing Run

J. C. Rastinejad[1], K. Paterson[1], W. Fong[1], D. J. Sand[2], M. J. Lundquist[3], G. Hosseinzadeh[2], E. Christensen[4], P. N. Daly[2], A. R. Gibbs[4], S. Hall[1,5], F. Shelly[4], and S. Yang[6]
[1] Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA; jillianrastinejad2024@u.northwestern.edu
[2] Steward Observatory, The University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721-0065, USA
[3] W. M. Keck Observatory, 65-1120 Mamalahoa Highway, Kamuela, HI 96743, USA
[4] Lunar and Planetary Lab, Department of Planetary Sciences, University of Arizona, Tucson, AZ 85721, USA
[5] Department of Physics and Astronomy, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, USA
[6] Department of Astronomy, The Oskar Klein Center, Stockholm University, AlbaNova, SE-10691 Stockholm, Sweden
Received 2021 December 17; revised 2022 January 13; accepted 2022 January 18; published 2022 March 4

Abstract

We present a comprehensive analysis of 653 optical candidate counterparts reported during the third gravitational-wave (GW) observing run. Our sample concentrates on candidates from the 15 events (published in GWTC-2, GWTC-3, or not retracted on GraceDB) that had a >1% chance of including a neutron star in order to assess their viability as true kilonovae. In particular, we leverage tools available in real time, including pre-merger detections and cross-matching with catalogs (i.e., point-source, variable-star, quasar and host-galaxy redshift data sets), to eliminate 65% of candidates in our sample. We further employ spectroscopic classifications, late-time detections, and light-curve behavior analyses and conclude that 66 candidates remain viable kilonovae. These candidates lack sufficient information to determine their classifications, and the majority would require luminosities greater than that of AT 2017gfo. Pre-merger detections in public photometric survey data and comparison of cataloged host-galaxy redshifts with the GW event distances are critical to incorporate into vetting procedures, as these tools eliminated >20% and >30% of candidates, respectively. We expect that such tools that leverage archival information will significantly reduce the strain on spectroscopic and photometric follow-up resources in future observing runs. Finally, we discuss the critical role prompt updates from GW astronomers to the EM community play in reducing the number of candidates requiring vetting.

Unified Astronomy Thesaurus concepts: Gravitational wave astronomy (675); Neutron stars (1108)

Supporting material: machine-readable table

1. Introduction

Since the first detection of a compact object merger by gravitational waves (GWs) in 2015 (Abbott et al. 2016a), the large number of detected mergers of black holes (BHs) and/or neutron stars (NSs) has contributed to the rapidly growing field of multimessenger astronomy. Each subsequent GW observing run has brought increased detector sensitivity and a larger survey volume to detect the mergers of binary black holes (BBHs), binary neutron stars (BNSs) and neutron star−black hole (NSBH) mergers (Abbott et al. 2019; The LIGO Scientific Collaboration et al. 2021; Abbott et al. 2021a, 2021b). BNSs and some NSBH mergers are expected to produce kilonovae, optical−near-IR thermal transients powered by the radioactive decay of heavy *r*-process elements (Li & Paczyński 1998; Metzger et al. 2010; Barnes & Kasen 2013; Tanaka & Hotokezaka 2013; Kawaguchi et al. 2016). Given their relatively low peak luminosities ($\sim 10^{41}$–10^{42} erg s^{-1}) and fast-fading nature (observable on ∼week timescales), discerning kilonovae from the wide array of optical transients is a long-standing challenge in this field.

The discovery of the first GW-detected BNS merger, GW170817 (Abbott et al. 2017a), and its kilonova AT 2017gfo (Arcavi et al. 2017; Coulter et al. 2017; Lipunov et al. 2017; Soares-Santos et al. 2017; Tanvir et al. 2017; Valenti et al. 2017) was the first proof of concept for multimessenger astronomy between gravitational and electromagnetic (EM) waves. Positively identifying new kilonova counterparts to GW events will help to constrain their intrinsic and extrinsic diversity (Metzger & Fernandez 2014; Gompertz et al. 2018; Ascenzi et al. 2019; Shibata & Hotokezaka 2019; Rossi et al. 2020; Kawaguchi et al. 2020a; Rastinejad et al. 2021). Further, by matching kilonova observations to models, one may infer their ejecta masses and compositions, therein elucidating the contribution of NS mergers' *r*-process enrichment in the universe. Given the high angular resolution of typical optical instruments, the discoveries of kilonova counterparts to GW events provide subarcsecond localizations and thus crucial identifications to host galaxies and stellar populations. This in turn can enable constraints on the Hubble constant (through identification of the host galaxy; Abbott et al. 2017b) and lend insight into the environments that give rise to BNS/NSBH mergers. Finally, one can indirectly constrain the maximum mass of NSs (as, with a greater sample of mergers with optical counterparts, we can probe the upper end of component masses that produce kilonovae; Fryer et al. 2015; Nicholl et al. 2021).

The third and most recent LIGO-Virgo Collaboration (LVC) observing run (O3) took place from 2019 April to 2020 March. The improved sensitivity of detectors produced a higher rate of detected compact object mergers at greater distances. This

resulted in 125 published O3 events between the second and third Gravitational-Wave Transient Catalogs (GWTC-2 and GWTC-3), augmenting the previous collection of GW-detected mergers by a factor of ∼10 (The LIGO Scientific Collaboration et al. 2021; Abbott et al. 2021a). The O3 literature includes five mergers for which the mass distribution of at least one component falls within the upper limit of an NS, accounting for uncertainties ($\lesssim 3\ M_{\odot}$). In addition, tens of merger events involving an NS were announced via the Gamma-ray Coordinates Network circulars (GCNs) in O3 that did not pass the traditional thresholds for inclusion in the published samples. Despite these numerous opportunities and subsequent efforts by the community, no credible EM counterpart to a GW event has been identified since AT 2017gfo (Andreoni et al. 2019a; Coughlin et al. 2019; Dobie et al. 2019; Goldstein et al. 2019; Gomez et al. 2019; Hosseinzadeh et al. 2019; Lundquist et al. 2019; Andreoni et al. 2020a; Antier et al. 2020a, 2020b; Ackley et al. 2020; Garcia et al. 2020; Gompertz et al. 2020; Kasliwal et al. 2020; Morgan et al. 2020; Pozanenko et al. 2020; Thakur et al. 2020; Vieira et al. 2020; Watson et al. 2020; Alexander et al. 2021; Anand et al. 2021; Becerra et al. 2021; Bhakta et al. 2021; Chang et al. 2021; de Wet et al. 2021; Dichiara et al. 2021; Dobie et al. 2022; Kilpatrick et al. 2021; Oates et al. 2021; Ohgami et al. 2021; Paterson et al. 2021; Tucker et al. 2021; deJaeger et al. 2022). A significant challenge for EM follow-up is the need to search large localization areas, which spanned ∼10–10,000 deg^2 for events in O3.

Previously in Paterson et al. (2021), the Searches After Gravitational-waves Using ARizona Observatories (SAGUARO) collaboration presented an analysis of optical candidate counterparts to 17 O3 events. Similar to other surveys, SAGUARO's observations of the large localizations returned thousands of candidate counterparts, roughly tens of which remained viable candidates after initial vetting (Lundquist et al. 2019; Paterson et al. 2021). In Paterson et al. (2021), we also examined optical follow-up by the community, finding that only 65% of reported candidates were ever reobserved. Among this follow-up, we found a high potential for redundant spectroscopic or photometric observations of candidates. In addition, we eliminated 12 previously "open" candidates as kilonovae by examining their photometric light curves and host-galaxy redshifts.

With the LIGO-Virgo-KAGRA (LVK) fourth observing run (O4) on the horizon, we are still confronted with the challenge of the correct identification of optical counterparts amid large localization areas. This is evidenced by the many remaining viable O3 kilonova candidates and the limited nature of spectroscopic and photometric follow-up resources. Thus, we are motivated to leverage the full arsenal of tools available at the time of follow-up (e.g., contextual catalog matching, existing survey observations) to conduct a uniform analysis of all kilonova candidates of any O3 merger involving an NS. We aim to examine what fraction of kilonova candidates could have been eliminated without targeted follow-up and what fraction remain viable after thorough vetting.

In this work, we analyze 653 O3 kilonova candidates across 15 GW events gathered from the GCNs and the Transient Name Server (TNS). We aim to (i) identify the most promising methods to eliminate candidates as kilonovae in real time (i.e., shortly after candidates have been identified and before they are reported in GCNs) and (ii) determine whether, after exploiting all tools at our disposal, any of the candidates are still physically viable as kilonovae. In Section 2 we describe our selection of 15 GW events and the corresponding 653 candidate counterparts. In Section 3 we apply tools to eliminate candidates that will be available in real time for O4. In Section 4 we utilize all tools, regardless of whether they are available in real time, to eliminate any remaining candidates. We examine the results of our analysis in Section 5 and make recommendations for future GW observing runs. Finally, we present our conclusions in Section 6. All magnitudes are reported in the AB system and are corrected for Milky Way dust extinction based on Schlafly & Finkbeiner (2011). Throughout, we assume a standard cosmology of $H_0 = 69.6\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$, $\Omega_M = 0.286$, $\Omega_{\rm vac} = 0.714$ (Bennett et al. 2014).

2. Collection of Kilonova Candidates

2.1. Event Selection

Toward our aim of identifying any remaining, plausible kilonovae, we examine the candidate counterparts to GW events involving at least one NS. Though potentially only a small fraction of NSBH mergers produce EM emission (Foucart et al. 2013; Shibata & Hotokezaka 2019; Broekgaarden et al. 2021), we search for candidates from all NSBH events, regardless of their mass ratio, spin, or other properties. Our sample of GW events includes (i) events published in the LVC literature whose final mass distribution includes at least one component with a >5% probability of being $<3\ M_{\odot}$ (following the conservative upper limit on an NS mass assumed by Abbott et al. 2021a based on Rhoades & Ruffini 1974; Kalogera & Baym 1996) and (ii) nonretracted events announced in the GCNs whose most recent classification likelihood of being a BNS or NSBH merger is >1%.

We first gather events from the GWTC-2 and GWTC-3 catalogs or other published LVC works (Abbott et al. 2020a, 2020b, 2021a, 2021b; The LIGO Scientific Collaboration et al. 2021). Five GW events that were initially announced in the GCNs meet the first criterion above (GW190425, GW190426_152155, GW190814, GW200105_162426, and GW200115_042309; Abbott et al. 2020a, 2020b, 2021a, 2021b). These events have false-alarm rates (FARs) of $<2.0\ \mathrm{yr^{-1}}$, which corresponds to an expected contamination fraction of <10% (Abbott et al. 2021a, 2021b). We also note that the GWTC-3 catalog (The LIGO Scientific Collaboration et al. 2021) published three events that meet the first criterion, GW191113_071753, GW191219_163120, and GW200210_092254, but they were not announced in real time via GCNs and all have estimated median distances beyond which typical nightly surveys depths would be sensitive to an AT 2017gfo-like kilonova ($D_{\rm L} \gtrsim 550$ Mpc; The LIGO Scientific Collaboration et al. 2021); thus, we do not include them in our sample.

Following the second criterion above, we augment our sample with 10 additional nonretracted events (S190510g, S190718y, S190901ap, S190910d, S190910h, S190923y, S190930t, S191205ah, S191213g, and S200213t) that were reported in the GCNs but did not meet the threshold for inclusion in the GWTC-2 or GWTC-3 catalogs owing to low FAR values found by offline analyses (The LIGO Scientific Collaboration et al. 2021; Abbott et al. 2021a). As most of these events were the subject of targeted optical searches reported to the GCNs and the aim of this work is to investigate how to improve future follow-up, we add them to our sample.

Table 1
Total Sample of Gravitational-wave Events and Electromagnetic Counterpart Candidates

Event	Classification[a]	Distance (Mpc)	GCN (Number of Candidates)	TNS	Combined Unique Total
GW190425	BNS, NSBH	157^{+70}_{-70}	16	33	34
GW190426[b]	NSBH	377^{+180}_{-160}	22	20	28
S190510g	BNS[c]	227^{+92}_{-92}	15	19	21
S190718y	BNS[c]	227^{+165}_{-165}	3	15	15
GW190814	Mass Gap, NSBH	241^{+40}_{-50}	31	95	96
S190901ap	BNS[c]	241^{+79}_{-79}	12	94	95
S190910d	NSBH[c]	632^{+186}_{-186}	0	7	7
S190910h	BNS[c]	230^{+88}_{-88}	16	59	59
S190923y	NSBH[c]	483^{+133}_{-133}	1	1	2
S190930t	NSBH[c]	108^{+38}_{-38}	10	181	183
S191205ah	NSBH[c]	385^{+164}_{-164}	8	27	27
S191213g	BNS[c]	201^{+81}_{-81}	12	23	23
GW200105[b]	NSBH	280^{+110}_{-110}	7	38	38
GW200115[b]	NSBH	300^{+150}_{-100}	10	13	13
S200213t	BNS[c]	201^{+80}_{-80}	5	12	12
Total			168	638	653

Notes. Bolded event names indicate that the event met the significance threshold for inclusion in GWTC-2 or GWTC-3 (The LIGO Scientific Collaboration et al. 2021; Abbott et al. 2021a). Unbolded events were announced in the GCNs but not included in the GWTC-2 or GWTC-3 catalogs, though they have not been retracted.
[a] Most probable nonterrestrial classifications.
[b] We abbreviate these event titles from their full names: GW190426_152155, GW200105_162426, and GW200115_042309.
[c] Most recent classification, according to GraceDB.

We use the most recent localizations and properties reported on GraceDB[7] for these unretracted, low-significance events.

Though an optical counterpart to a BBH merger has been claimed (Graham et al. 2020; but see also Ashton et al. 2021), we do not include BBH mergers in our analysis, as their probability of producing detectable EM counterparts is exceedingly low compared to mergers involving an NS (Perna et al. 2018). We note that mergers in which one object falls in the "Mass Gap" ($3 < M/M_{\odot} < 5$; Abbott et al. 2016b) also have the potential to create EM emission and thus received some follow-up during O3. Three events were initially announced in the GCNs as having a >95% chance of being a Mass Gap event (GW190924_021846, GW190930_133541, and GW200316_215756). However, final analyses of these events (The LIGO Scientific Collaboration et al. 2021; Abbott et al. 2021a) find that all three are most likely BBHs, and thus they are not included in our sample. Our final sample of 15 GW events and their properties are listed in Table 1.

2.2. GCN and TNS Candidates

We gather candidate counterparts to each GW event from the GCNs and TNS.[8] GCNs are real-time notices to the community of EM follow-up to gamma-ray bursts (GRBs) or GW events. GCNs primarily contain candidates reported immediately following the event, while TNS is a more comprehensive database of newly discovered transients that is independent of GW events.

Our first step in optical candidate selection is to define initial criteria based on time, location, and luminosity to use in our GCN and TNS searches. Our initial criteria for inclusion are as follows: candidates (i) with $0 < \delta t < 5$ days (where δt is the time between the event merger time and the discovery time of the candidate) and (ii) that are within the 90% contour on the final localization map (for events published in the O3 catalogs or other LVK papers; Abbott et al. 2020a, 2020b, 2021a, 2021b) or the most recent LALInference map available in GraceDB. We apply a third criterion, which filters out candidates that would be more luminous than the brightest kilonova model predictions (e.g., Barbieri et al. 2021; Fong et al. 2021) and short GRB kilonova candidates (see analyses of Fong et al. 2021; Rastinejad et al. 2021) at the GW-inferred event distance. We calculate the luminosity, νL_{ν}, at the 1σ lower bound on the GW-inferred event distance (providing a conservative estimate) using the discovery filter pivot wavelength (or the r-band pivot wavelength if no filter is reported) and include only candidates fitting the criteria $\nu L_{\nu} < 10^{43}$ erg s^{-1}, ≈10 times the luminosity of AT 2017gfo and ≈5 times the peak luminosities of the brightest kilonova models (Barbieri et al. 2021; Fong et al. 2021).

We apply these initial criteria to our GCN searches for the 15 events, collecting the name, R.A., decl., discovery magnitude, and time of discovery of each reported candidate. Since our goal is to investigate real-time tools that might eliminate the need for follow-up, we include all GCN candidates that pass the initial criteria, regardless of whether they were subsequently eliminated by further GCNs or the literature. However, we do keep track of which candidates were later eliminated. In total, we find 168 candidates across 15 events reported to the GCNs, 96 (57.1%) of which were subsequently eliminated as reported in the GCNs. One event had no candidates reported (S190910d), while GW190814 had the most candidates reported that meet our initial criteria (31).

We next apply the initial criteria to all transients reported to TNS in 2019 and 2020, regardless of their TNS classification.

[7] https://gracedb.ligo.org
[8] https://www.wis-tns.org

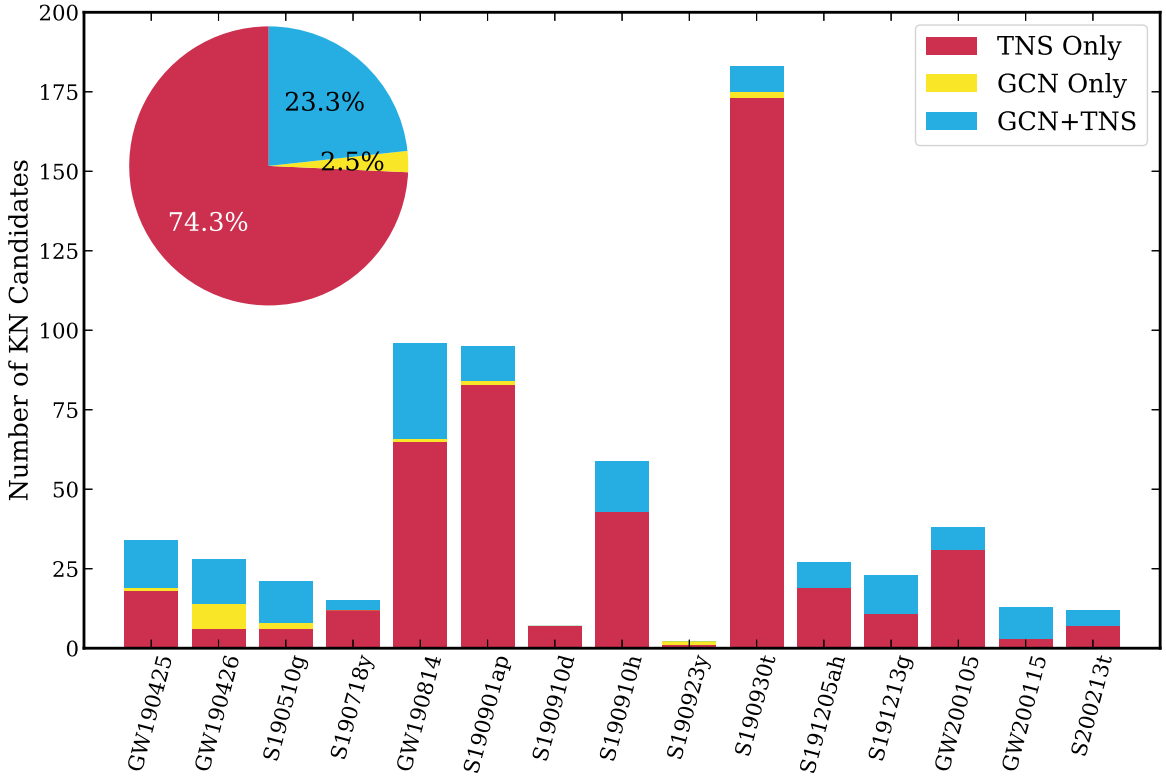

Figure 1. The number of plausible kilonova candidates sorted by GW event after our initial criteria but before we eliminate candidates using "Tools Available in Real Time" (Section 3). Each bar is color-coded by the source of the candidate (GCN, TNS, or both; Section 2). The most likely GW event classification is labeled at the top of each bar. We list two classifications for GW190425 and GW190814, as these events are well studied and the literature is divided on their origin. Inset: a pie chart shows the distribution of the sources of all candidates meeting our initial criteria. Approximately one-quarter of candidates that we consider in our sample are reported in the GCNs, and almost all are reported to the TNS.

We gather 638 candidates from TNS across the 15 events, the majority ($\gtrsim$90%) of which are not classified as a particular transient type as of 2021 December. The event with the second-largest 90% localization region, S190930t, had the greatest number of candidates from TNS (181).

To remove any duplicate candidates, we cross-match between the GCN and TNS samples using the unique TNS name (if reported in GCNs) and by eliminating matches within 2″. Together, our sample includes 652 unique candidates. However, one transient is a candidate to both S190910d and S190910h, and thus we "double-count" it by independently considering it for each GW event. Thus, after our preliminary cuts, we consider a total of 653 candidates corresponding to the 15 GW events.

2.3. Total Sample

We present the number of candidates collected in the GCNs, in TNS, and in our total sample in Table 1. In Figure 1 we show a histogram of the number of candidates per event. With the exception of GW190814, for which deep, follow-up observations were conducted (Gomez et al. 2019; Ackley et al. 2020; Morgan et al. 2020; Thakur et al. 2020; Vieira et al. 2020; Andreoni et al. 2020a; Kilpatrick et al. 2021; Tucker et al. 2021), the four events with the highest number of candidates have localizations of >7500 deg^2. Despite the initial classification of >99% chance Mass Gap event (LIGO Scientific Collaboration & Virgo Collaboration 2019a), the relatively small localization and subsequent >99% NSBH event classification (LIGO Scientific Collaboration & Virgo Collaboration 2019b; reported ∼11 hr later) resulted in extensive community follow-up of GW190814. The substantial targeted follow-up is also evidenced by the high fraction of GCN-reported candidates in comparison with other events.

3. Vetting Candidates with Tools Available in Real Time

We begin by cross-matching candidates with catalogs that provide critical contextual information and searching public surveys for pre-merger detections to vet candidates with tools that would be available in real time. Motivated to find the most efficient means of ruling out kilonova imposters, we apply the first four conditions in parallel (point-source/stellar catalogs, moving-object catalog, quasar catalogs, and pre-explosion detections; Sections 3.1–3.4) to every candidate in our sample, regardless of whether it was classified in the GCNs or TNS. We apply the steps described in Section 3.5, host-galaxy matching, to only the remaining candidates. In Figure 2 (left) we show the fraction of candidates eliminated by each tool in this section.

3.1. Point-source and Variable-star Catalogs

Variable stars are sources of contamination in transient surveys and may be eliminated by cross-matching transients with the locations of known point sources. We query the Gaia early Data Release 3 (eDR3; Gaia Collaboration et al. 2016, 2021) catalog, the Pan-STARRS (PS1) point-source catalog (Tachibana & Miller 2018), and the ASAS-SN variable-star catalog (Jayasinghe et al. 2019) for objects within 2″ of each candidate. Though the astrometrical uncertainty varies by catalog, we elect to use a uniform cross-matching radius that reflects the seeing and pixel size of the surveys. For Gaia, a

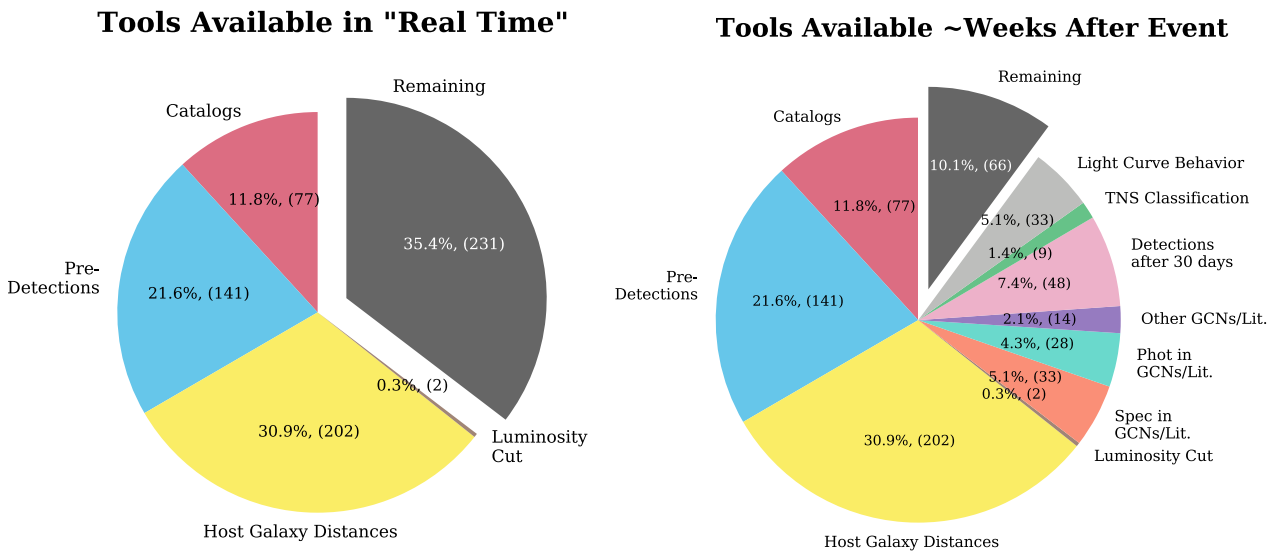

Figure 2. Left: pie chart demonstrating the fractions of candidates ruled out by cross-matching with stellar, quasar, and moving-object catalogs (Sections 3.1–3.2; red); photometric detections before the associated GW event (Section 3.4; blue); host galaxies inconsistent with the GW event distance (Section 3.5; yellow); and a luminosity cut using the host-galaxy distance (Section 3.6; brown). Each of these tools can be applied in real time during future observing runs. Together, host-galaxy distances and pre-explosion detections rule out over half of the candidates in our sample, demonstrating their utility. Right: same as the left panel, but with additional eliminations made using tools available roughly days to weeks after the event. Eliminations made in the GCNs and literature are divided into those made with spectroscopy (orange), photometry resources (turquoise), or other follow-up (mostly pre-detections, purple; Section 4.1). In addition, candidates ruled out as kilonovae owing to detections after 30 days (Section 4.2; pink), TNS classifications (Section 4.3; green), and light-curve behavior inconsistent with a kilonova (Section 4.4; gray) are shown. At the conclusion of the analysis, 66 candidates (10.1% of the 653 candidates in the original sample) remain viable kilonova candidates.

candidate is considered stellar if (i) its match has an absolute proper motion value >3 times the error in total proper motion, (ii) it is flagged as variable, or (iii) it has a parallax significance >8 (following Tachibana & Miller 2018). The PS1 point-source catalog assigns nearly all PS1 sources a score indicating their likelihood of being stellar based on a random forest machine-learning algorithm. We consider sources to be stellar if they have a single counterpart in the catalog whose point-source score is greater than 0.83 (reflecting a true-positive rate of ∼0.995, and a false-positive rate of 0.005; Tachibana & Miller 2018). Finally, we inspect the light curves of the four ASAS-SN variable-star matches, finding that three are good matches to variable stars. The remaining source, AT 2019rup or ASASSN-V J094204.78 +234107.0, is classified as a young stellar object (YSO) and is coincident with the nucleus of a galaxy. However, as its ASAS-SN light curve shows pre-merger variability, we rule this out as a viable candidate. In total, we conclude that 51 sources are stellar after cross-matching with three catalogs. Notably, four were reported in the GCNs as initially viable counterparts. The PS1 point-source catalog eliminates the greatest number of candidates as kilonovae (38).

3.2. Moving-object Catalog

Near-Earth moving objects are another potential source of contaminants. We use a radius of 20″ to cross-match candidates with the IAU Minor Planet Center Orbit Database[9] (MPCORB). As many surveys already include moving-object cross-matching in their pipeline, we expect few matches within our sample. Accordingly, we identify 10 candidates with a match in MPCORB.

We do not find any detections at the same coordinates prior to or following the initial discoveries of the 10 candidates in the public ZTF database, the SAGUARO database, or the TNS, or using the ATLAS forced photometry tool (the use of these tools is further described in Section 3.4). However, we note that for three candidates (AT 2019nri, AT 2019nsn, AT 2019nsl) detections were reported ∼2–8 minutes apart in the *i*- and *z*-band filters, though these are consistent with the radial velocities of known moving objects. We eliminate all 10 candidates with matches in MPCORB as viable kilonovae.

3.3. Quasar Catalogs

Active galactic nuclei (AGNs) and quasars are known variable sources that may masquerade as real transients. We next query the Million Quasar Catalog (MILLIQUAS; Flesch 2015, 2021) for objects within 1″ of candidates. We employ a smaller match radius than was used for point-source matching to avoid confusion between an offset candidate and an active host-galaxy center. We accept candidates as quasars if their probability of being a quasar (calculated based on the association of photometric data with radio or X-ray detections) is >97%, following Flesch (2015), which found this threshold to yield good agreement with confirmed quasars in the Sloan Digital Sky Survey (SDSS) DR16 Quasar Catalog (Lyke et al. 2020). We also cross-match candidates to the SDSS quasar catalog using the same matching radius. We consider objects marked by the catalog as questionable quasars to still be viable candidates. We also inspect the offsets of candidates to the host-galaxy nucleus in the Legacy Survey viewer[10] to ensure that quasars are not falsely attributed to real transients (for future, larger samples of candidates, calculating the offsets using cataloged coordinates may be prudent). Finally, we examine the candidates' light curves (further described in Section 3.4) and find that 20 (none of which were reported in

[9] http://www.minorplanetcenter.net/iau/MPCORB.html

[10] http://legacysurvey.org/viewer

the GCNs) had credible pre-merger detections. Twenty-six candidates are marked as quasars, thus eliminating them from being viable kilonova candidates. In total, cross-matching to point-source/stellar, moving-object, and quasar catalogs (Sections 3.1–3.3) results in the elimination of 77 candidates, or 11.8% of our sample (Figure 2, left).

3.4. Pre-merger Detections

Compact object mergers are not typically expected to produce optical or IR emission prior to a GW event, allowing us to eliminate any candidate with a pre-merger detection. We search for and combine all available photometry of the 653 unique candidates in our sample from TNS, the public Zwicky Transient Facility (ZTF; Bellm et al. 2019) database, the Asteroid Terrestrial-impact Last Alert System (ATLAS; Tonry et al. 2018; Smith et al. 2020) forced photometry tool, and our own observations from the SAGUARO database. As SAGUARO utilizes the Steward Observatory 1.5 m Mt. Lemmon telescope with its 5 $\deg^2$ imager (operated by the Catalina Sky Survey; Christensen et al. 2018) as its discovery engine, $\sim$3 yr of observations at an average depth of 21.1 mag are available to search for pre-explosion detections (see Lundquist et al. 2019; Paterson et al. 2021, for more details). In Section 5.3 we discuss additional surveys to search for pre-merger detections.

For ZTF, we gather image-subtracted photometry when available (Masci et al. 2019). For ATLAS, we perform forced point-spread function (PSF) photometry at the positions of candidates covered by the survey for 200 days preceding the GW events using the publicly available service (simulating what would be computationally reasonable in real time; Tonry et al. 2018; Smith et al. 2020). We also stack multiple ATLAS epochs of photometry in a given filter on a single night using the publicly available script provided by the service.[11] Finally, we query the SAGUARO database for candidates using a matching radius of 1″ and obtain difference-imaged photometry or 5σ limits (Lundquist et al. 2019; Paterson et al. 2021). We convert all magnitudes to AB units and correct all detections for Milky Way extinction in the direction of the candidate (Schlafly & Finkbeiner 2011).

As not all difference images are publicly available, it is possible that poor galaxy subtractions can masquerade as pre-merger detections at or near candidate positions. Further, asteroid detections may contaminate the forced photometry tool. Thus, we take a conservative approach by requiring that any of the following criteria are met to eliminate a candidate as viable: (i) there is a pre-merger detection $<$12 days prior to the GW event that, upon manual inspection, follows a smooth trajectory consistent with the subsequent light curve; (ii) there are multiple pre-merger detections within 10 days of each other; (iii) there are multiple pre-merger detections at any time by different surveys; or (iv) there is a single pre-merger detection that does not meet any of the criteria above but lacks a contaminating source within $\sim$5″ of the candidate position in archival PS1 imaging.

Ultimately, we rule out 186 candidates ($>$20% of our initial sample; Figure 2, left) based on pre-merger detections. Thirty-three of these eliminated candidates were reported in the GCNs. Combined with eliminations made in Sections 3.1–3.3, we eliminate 218 candidates, leaving 435 candidates to be carried to the following step.

3.5. Host-galaxy Associations and Distances

We next attempt to associate each candidate with its most likely host galaxy and, if available, use the cataloged photometric or spectroscopic redshift to eliminate candidates whose hosts are outside the 95% credible interval of the GW-inferred event distance. We search three catalogs for potential host galaxies: SDSS Data Release 12 (SDSS DR12; Alam et al. 2015), Pan-STARRS Source Types and Redshifts with Machine Learning (PS1-STRM; Chambers et al. 2016; Beck et al. 2021), and Legacy Survey Data Release 9 (LS DR9; Dey et al. 2019). SDSS DR12 covers over 14,000 $\deg^2$ to an average depth of $r > 22.7$ AB mag and supplies spectroscopic redshifts of more than 1.4 million galaxies (Alam et al. 2015). The catalog presents $>$200 million photometric redshifts and includes star–galaxy probabilistic classifications of objects (Beck et al. 2016). The PS1-STRM catalog analyzes over 2.9 billion objects from PS1 DR1 (DR2 is unavailable), presenting star–galaxy classifications and a large catalog of photometric redshifts (Beck et al. 2021). PS1 DR1 covers $\sim$30,000 $\deg^2$, reaching similar depths to SDSS (Metcalfe et al. 2013; Chambers et al. 2016). Finally, LS DR9 combines observations from the fourth BASS (Zou et al. 2017) and MzLS (Silva et al. 2016) data release and the seventh DECaLS (Dey et al. 2019) release. The combined survey covers $\approx$14,000 $\deg^2$ to depths of $r > 23.4$ AB mag (Dey et al. 2019). The LS DR9 photometric redshift catalog includes over 2.7 million objects that are identified as galaxies using color and magnitude cuts (Zhou et al. 2021). All three photometric redshift catalogs are trained on a sample of spectroscopically classified galaxies spanning the range $0 \lesssim z \lesssim 0.8$. Together, these three catalogs provide nearly complete coverage of the candidates in our sample, with the positions of only 14 candidates, or 3.1% of those considered in this step, not covered by the combined footprints of the surveys.

For host associations, we begin by searching for galaxies near each candidate's position in SDSS DR12 and PS1-STRM through Vizier[12] and in LS DR9 through NOIRLab's Data Lab.[13] Following Zhou et al. (2021), we do not use LS DR9 photometric redshifts of sources $z < 21$ mag. Based on the observed offsets of short GRBs from the centers of their host galaxies (and thus the maximum observed offsets of NS mergers; Fong & Berger 2013), we determine that a 100 kpc offset between the transient and host center is a conservative search radius. At the nearest GW event distance in our sample (108 Mpc; Table 1), this corresponds to a search radius of 3.′138. We cross-match between the three catalogs using a radius of 2″ and remove sources identified as stellar by either SDSS or STRM. If a spectroscopic redshift is available, we consider this value as the galaxy redshift. Otherwise, we record any photometric redshifts reported in STRM, SDSS, or LS DR9.

Next, we compute the probability of chance coincidence ($P_{\rm cc}$; Bloom et al. 2002) for each source. $P_{\rm cc}$ calculates the probability of chance alignment between the transient and potential host galaxies in the field using the galaxy magnitudes and angular offsets from the candidate's position. A low $P_{\rm cc}$

[11] https://gist.github.com/thespacedoctor/86777fa5a9567b7939e8d84fd8cf6a76

[12] https://vizier.cds.unistra.fr

[13] https://datalab.noirlab.edu.

Table 2
Candidates Whose Hosts Are within the GW Event Distance Range

Category	Description	No. of Candidates
Platinum	Highly Confident Host Association and Spectroscopic Redshift within GW Distance Uncertainty	19
Gold	Moderately Confident Host Association and Spectroscopic Redshift within GW Distance Error	1
Silver	Highly Confident Host Association and Photometric Redshift within GW Distance Error	94
Bronze	Moderately Confident Host Association and Photometric Redshift within GW Distance Error	12
Inconclusive	Either candidate region uncatalogued, no redshift of best galaxy, or no confident association	107
Eliminated	Highly or Moderately Confident Host Association and Redshift outside GW Distance Error	202

Note. Candidates remaining after cross-matching to catalogs and searching for pre-merger detections (Sections 3.1–3.4) separated into categories of host-galaxy association confidence. Confidence that each candidate is associated with a host galaxy in the GW-inferred distance range descends from Platinum to Bronze.

value, especially in comparison to the values of other galaxies in the field, indicates a more likely host. Using the hosts' $P_{\rm cc}$ values, we sort the candidates into four categories based on our confidence in their host-galaxy associations:

1. *Highly Confident*: Across all galaxies considered, the minimum $P_{\rm cc}$ value ($P_{\rm cc,min}$) < 0.01, while the second smallest $P_{cc} \geqslant 3 \times P_{\rm cc,min}$.
2. *Moderately Confident*: $P_{\rm cc,min} < 0.15$ and the second smallest $P_{cc} \geqslant 3 \times P_{\rm cc,min}$.
3. *Not Confident*: Sources were found within the vicinity of the candidate, but neither the "Highly Confident" nor "Moderately Confident" criterion was met.
4. *Uncatalogued*: Candidate not covered by SDSS, STRM, and LS DR9 footprint.

We remove duplicate entries of a single host in multiple catalogs on visual inspection and, when applicable, keep track of both photometric redshifts.

In Table 2 we define categories that combine our confidence in the host association (see above) and the relative robustness between a spectroscopic and photometric redshift in determining whether a host is within the GW distance uncertainty. Candidates for which we are unable to make a confident association or are not covered by the footprint of the catalogs we query are marked inconclusive. In total, 364 candidates have "Highly Confident" associations. Of these, 37 have spectroscopic redshifts available ("Platinum" associations) and 266 have photometric redshifts ("Silver"). Thirty-one candidates have "Moderately Confident" associations. One "Moderately Confident" host has a spectroscopic redshift ("Gold"), and the remaining have photometric redshifts ("Bronze"). As we cannot draw any firm conclusions for candidates with "Not Confident" associations, we mark the 107 "Not Confident" and uncatalogued candidates as "inconclusive."

Finally, we use the spectroscopic and photometric redshifts of the "Highly Confident" and "Moderately Confident" associations to determine whether the candidates' host-galaxy redshifts are consistent with the 95% distance credible interval inferred from GWs (Table 1), which is inherently position dependent (Singer et al. 2016). We thus use the calculated GW distance uncertainties from the 3D localization maps at the position of each candidate for comparison. For hosts with photometric redshifts, we determine whether the redshifts' 1σ error range falls within the 95% GW credible interval. For hosts with spectroscopic redshifts, we simply use the central value, as spectroscopic redshift errors are negligible. Between the "Highly Confident" and "Moderately Confident" samples, we find hosts of 202 candidates (30.9% of our initial sample; Figure 2, left) that do not fall within the event distance range; thus, these sources are ruled out as GW counterpart candidates. We state the number of remaining candidates in each of the Platinum, Gold, Silver, and Bronze categories, as well as the number of inconclusive and eliminated candidates, in Table 2.

3.6. Luminosity Cut Using Host Distance

Finally, we attempt to eliminate candidates with luminosities inconsistent with kilonovae, calculated using the host-galaxy distances determined in Section 3.5 (which generally have lower errors than the GW event distances). We calculate the luminosity (νL_ν) of each candidate with Platinum, Gold, Silver, or Bronze associations using its host-galaxy redshift derived in Section 3.5. For νL_ν we use the candidate's discovery magnitude, filter pivot wavelength, and the 1σ lower bound on the host-galaxy distance, as this provides a lower bound on luminosity. We rule out any candidate that exceeds $\nu L_\nu > 10^{43}$ erg s^{-1} (following the reasoning of Section 2.2). To rule out a candidate whose host has multiple photometric redshift measurements, the luminosities calculated from all inferred redshifts for a given host must exceed the cut. We rule out two candidates as viable kilonovae based on this criterion.

After applying the steps in Sections 3.1–3.6 in which only information available in real time is used, we have eliminated 422 of the 653 candidates (64.6%) in our total initial sample. Of the candidates we rule out using tools available in real time, 88 were reported in the GCNs, or 52.3% of GCN-reported candidates in our sample. This demonstrates the power of using available contextual information for winnowing down kilonova candidates. For clarity, we break down exactly how these candidates were eliminated, using the methods in this section, in Figure 2 (left).

4. Employing Information Available Roughly Days to Weeks after the Event

After eliminating candidates based on cross-matching to catalogs, pre-merger detections, host associations, and luminosities in Section 3, we examine the remaining viable candidates (214) using tools or information available roughly days to weeks after the event. Through this process of further eliminating any of the 214 remaining candidates, we determine which (if any) candidate counterparts to the events in our sample remain viable. Exploring the sample of viable candidates in aggregate will inform what future tools would be most useful to eliminate them. In addition, tracking classifications from the GCNs and literature allows us to quantify what fraction of targeted follow-up would be

Table 3
Remaining Viable Kilonova Candidates and Their Properties

Name	R.A. (deg)	Decl. (deg)	Source	Event	Event Dist. (Mpc)	δt[a] (days)	Mag.[a] (AB mag)	Filt.[a]	Host[b]	Host z	νL_ν (erg s^{-1})	Notes
AT 2019eig	9.92875	−31.99236	Both	GW190425	157 ± 70	4.29	18.5	G	N		2.1×10^{42}	[d]
AT 2019efr	246.73323	10.93689	TNS	GW190425	157 ± 70	2.10	20.6	w	S	$0.035^{+0.124}_{-0.027}$	4.6×10^{41}	[d]
AT 2019frd	180.68045	−15.08255	TNS	GW190426	377 ± 160	4.69	20.7	w	S	$0.086^{+0.006}_{-0.006}$	4.5×10^{41}	[d]
AT 2019fxg	165.75082	−7.19200	TNS	GW190426	377 ± 160	4.62	22.1	w	S	$0.102^{+0.009}_{-0.006}$	1.2×10^{41}	
AT 2019ioy	172.08144	−11.04771	TNS	GW190426	377 ± 160	4.68	21.8	w	S	$0.107^{+0.036}_{-0.036}$	1.7×10^{41}	[d]

Notes. Table 3 is published in its entirety in the electronic edition of the Astrophysical Journal. A portion is shown here for guidance regarding its form and content.
[a] Time, magnitude, and filter of discovery.
[b] Host-galaxy association class, where P, G, S, and B are abbreviations for the Platinum, Gold, Silver, and Bronze classes (as defined in Table 2), respectively. T indicates a redshift reported to TNS, and N indicates that no host galaxy can be confidently associated with the candidate.
[c] Kilpatrick et al. (2021) eliminate the candidate based on an inconsistent host distance. As we employ different methods of eliminating candidates with host-galaxy redshifts (see Section 3.5), this candidate remains viable in our analysis.
[d] The candidate's light curve (based on photometry gathered from TNS, ATLAS, ZTF, and SAGUARO, further described in Section 3.4) consists of a single detection.
[e] Time of discovery is not included in reporting GCN, and object is not included in reporting group's published work summarizing candidates from this event (Goldstein et al. 2019; Andreoni et al. 2019b). We approximate its discovery time with those of candidates reported in the literature by the same discovery group.
(This table is available in its entirety in machine-readable form.)

considered redundant if all tools we apply in Section 3 had been available in O3 (further discussed in Section 5.1).

4.1. GCN and Literature Follow-up

We consider optical follow-up reported in the GCNs and literature (submitted or published papers) by numerous EM counterpart follow-up groups (Coughlin et al. 2019; Goldstein et al. 2019; Gomez et al. 2019; Hosseinzadeh et al. 2019; Andreoni et al. 2019a; Ackley et al. 2020; Garcia et al. 2020; Gompertz et al. 2020; Kasliwal et al. 2020; Morgan et al. 2020; Thakur et al. 2020; Vieira et al. 2020; Watson et al. 2020; Andreoni et al. 2020a; Antier et al. 2020a, 2020b; Anand et al. 2021; Becerra et al. 2021; Chang et al. 2021; Dichiara et al. 2021; Kilpatrick et al. 2021; Oates et al. 2021; Ohgami et al. 2021; Tucker et al. 2021). We note that the majority of these works were focused on follow-up of GW190814, the most precisely localized event potentially involving an NS throughout O3.

In total, we gather follow-up information of 199 of the 653 candidates in our initial sample from the GCNs and literature. We find that 77 candidates (48 and 29 reported to the GCNs and the literature, respectively) still considered viable after employing real-time tools (e.g., after Section 3) are eliminated. A total of 28 and 11 of the GCN eliminations are due to spectroscopic classifications and photometric follow-up, respectively. Of the 26 candidates eliminated in the literature, 2 are reported as image artifacts, 2 as moving objects, 1 as an AGN based on PS1 imaging, and 5 as stellar based on the Gaia, PS1, or ASAS-SN catalogs. An additional 14 transients are eliminated using serendipitous or targeted photometry of each candidate. Five candidates are eliminated by a spectroscopic classification.

Kilpatrick et al. (2021) determine that nine candidates cannot be kilonova counterparts to GW190814 based on host-galaxy cross-matching. As we employ a different method of associating candidates with host galaxies and, in cases where multiple photometric redshifts are available, require all to be inconsistent with the GW event distance (Section 3.5), we retain the nine candidates eliminated by Kilpatrick et al. (2021) (and note them in Table 3). Thus, ignoring the nine eliminations based on host distance, we rule out 29 candidates that have not already been rejected based on reasoning from the literature. In Figure 2 (right) we show the 71 candidates ruled out as viable kilonovae in the GCNs or literature, classified by the method of elimination. Later, we explore what fraction of the targeted follow-up could be considered redundant in light of eliminations made in Section 3 in Section 5.2.

4.2. Detections after $\delta t = 30$ days

The optical and/or near-IR emission emitted by NS mergers is not predicted to be detectable at the GW distances considered in this work (see Table 1) beyond $\delta t \sim 2$ weeks. Thus, we examine late-time light curves of the 144 remaining candidates, built using the same databases and process as described in Section 3.4. We rule out transients as viable kilonovae if their light curves contain optical and/or near-IR detections beyond a conservative timescale of $\delta t > 30$ days using criteria similar to those employed in Section 3.4. In this case, to be eliminated, the transient light curve must show (i) at least one detection past $\delta t = 30$ days that follows the shape of the preceding light curve, (ii) multiple detections past $\delta t = 30$ days within 10 days of each other, or (iii) multiple detections past $\delta t = 30$ days by different telescopes. Forty-eight candidates meet one of these criteria and are eliminated at this stage.

4.3. TNS Classifications

Next, we consider candidate classifications reported to TNS. Generally, these classifications are based on spectroscopic observations. Unlike in the GCNs, the follow-up and classifications reported to TNS are not necessarily performed with targeted kilonova or EM counterpart searches in mind. Rather, much of the TNS follow-up is focused on reporting newly discovered transients from targeted or untargeted surveys often unconnected to GW follow-up. Nine candidates not eliminated in previous steps were classified in TNS as Type Ia, IIn, or Ic supernovae (SNe; Brennan et al. 2019; Fremling et al. 2019; Gromadzki 2019; Dahiwale & Fremling 2019; Strader 2019; Zimmerman et al. 2020), leaving 99 remaining candidates at this stage.

4.4. Analysis of Photometric Light Curves

Finally, we analyze the light curves of the remaining candidates and rule out those whose behaviors are inconsistent with predictions for kilonovae, even accounting for the most luminous kilonova models. We again employ photometric data gathered from TNS, the public ZTF stream, ATLAS forced photometry, and the SAGUARO database (see Section 3.4).

Other O3 EM follow-up works analyzed light curves on a case-by-case basis, eliminating candidates whose light curves are flat (Ackley et al. 2020) or decline more slowly than 0.3 mag day^{-1} (Kasliwal et al. 2020). Cowperthwaite & Berger (2015) established cuts of $(i-z) > 0.4$ mag and $\delta t_{\rm rise} < 4$ days (where $\delta t_{\rm rise}$ is defined as the time it takes for a transient to rise from 1 mag pre-peak to peak in z band) to discriminate between kilonovae, Type Ia SNe, other fast-evolving transients (e.g., Drout et al. 2014; Margutti et al. 2019; Coppejans et al. 2020), and other contaminants.

We elect a more conservative approach grounded in the predicted luminosities and time evolution of several diverse kilonova models. The observed colors and luminosities of kilonovae are predicted to vary with both extrinsic (e.g., viewing angle) and intrinsic (primarily the ejecta mass and composition) properties. Kilonovae viewed at an edge-on (pole-on) angle are expected to be redder (bluer) owing to the lanthanide-rich tidal tails produced during the merger (Kasen et al. 2015; Chase et al. 2021; Korobkin et al. 2021). The ejecta mass and composition are directly affected by the progenitor (BNS or NSBH) and remnant (e.g., BH, short-lived rotationally supported NS, or magnetar) types. In particular, a bluer, slow-decaying kilonova is a predicted product of a long-lived NS or magnetar remnant owing to the effects of neutrino irradiation unbinding the disk material surrounding the remnant (Metzger & Fernandez 2014; Kasen et al. 2015; Lippuner et al. 2017; Fong et al. 2021). Taking into account predicted diversity, we consider an AT 2017gfo-like model (Kasen et al. 2017), a fiducial NSBH kilonova model (Kawaguchi et al. 2020b), a stable NS remnant kilonova model (Kasen et al. 2015), a model of the kilonova from massive progenitors (Barbieri et al. 2021), and a kilonova model for a magnetar remnant (Fong et al. 2021). We find that none rise in luminosity beyond $\delta t \approx 4$ days in a single band, and thus we employ this as our first criterion to eliminate candidates as kilonovae. We explain our use of detections in multiple filters below. In light of bluer kilonova models (Kasen et al. 2015), the early, blue peak of AT 2017gfo, and the lack of color measurements for most candidates, we do not employ any color criteria in eliminating candidates.

Eleven candidates show rising behavior in a single filter after $\delta t = 4$ days and thus are not viable kilonovae. An additional three candidate light curves show nearly constant magnitudes in a single filter over periods of ∼3–9 days, inconsistent with any kilonova models. When enough data points were available, we find that 13 candidate light curves decline by $\lesssim$1.5 mag day^{-1} over periods of 5−25 days in a single filter, again inconsistent with AT 2017gfo or any kilonova model. We further eliminate five candidates whose light curves decline by <1.5 mag over 13-to-18-day periods in "adjacent" (e.g., g- and o-band) optical filters. Similarly, unlike any expected kilonova, we eliminate one candidate that brightens over 6 days in adjacent filters. The majority of the remaining light curves contain only one detection, preventing us from making any conclusions about their behavior. In all, we determine that 33 candidates are not kilonovae based on their light curves.

After applying all tools at our disposal, including those available in real time and those available roughly weeks after the merger, 66 candidates remain viable kilonovae. Our initial sample of 653 is a fairly comprehensive compilation of candidates to O3 NS mergers, and this work eliminates 587 (90%) of these candidates as kilonovae.

5. Discussion

We now explore the results and implications of our analysis. We first determine whether there is sufficient evidence to claim any of the remaining viable candidates as a real kilonova counterpart. The remaining viable candidates and their properties are summarized in Table 3. Next, we examine sources of redundancy in candidate follow-up and make recommendations for improvements in future observing runs. Finally, we discuss the results of this analysis in the context of previous work and practices by the EM community.

5.1. Aggregate Analysis of the Remaining Viable Candidates

We now analyze the available information for the 66 remaining candidates still considered viable kilonovae after our vetting procedures, including their maximum observed luminosities and time evolution. We do not find convincing evidence that any of the remaining candidates are indeed kilonovae (though we cannot eliminate them using available information).

We show the initial and remaining number of candidates for each event, ordered by increasing localization area, in Figure 3. With the exception of GW190814, the three events with the largest localization areas have the highest numbers of remaining and initial candidates, as expected. Barring GW190814 and the three events with the largest localizations, the number of candidates (both initial and remaining) per event is roughly consistent. Notably, despite being the most precisely localized event in the sample, GW190814 has the second-greatest number of initial kilonova candidates (96). This indicates that well-localized, distant ($\gtrsim$200 Mpc) events with a high astrophysical probability discovered in O4 will likely receive wide, deep coverage by the EM community, resulting in a much higher number of candidates per square degree than the average event. This is further demonstrated by the higher fraction of GCN-reported candidates for GW190814 compared to most other events (Figure 1).

In an effort to determine whether our remaining candidates are either plausible kilonovae or clear contaminants, we next compare our sample in aggregate to the landscape of known transients. We first calculate the maximum observed luminosities ($\nu L_{\nu,\rm max}$) of remaining candidates based on the brightest observation in their photometric light curve and their host distance (determined based on TNS, redshift surveys, or the median GW event distance). Figure 4 shows a histogram of $\nu L_{\nu,\rm max}$ values for the remaining candidates in blue. The peak r- and K-band luminosities of AT 2017gfo, calculated based on best-fit models from Kasen et al. (2017), are shown as red vertical bars. The expected peak luminosity range of plausible kilonovae (Barnes & Kasen 2013; Tanaka & Hotokezaka 2013; Metzger & Fernandez 2014; Kasen et al. 2015; Metzger 2019; Fong et al. 2021) and short GRB afterglows (based on observations from Fong et al. 2015) are shown with red horizontal bars. The peak luminosity ranges of possible contaminating transients (including novae and various SN types) are overlaid in black bars at the top. Broadly, a majority

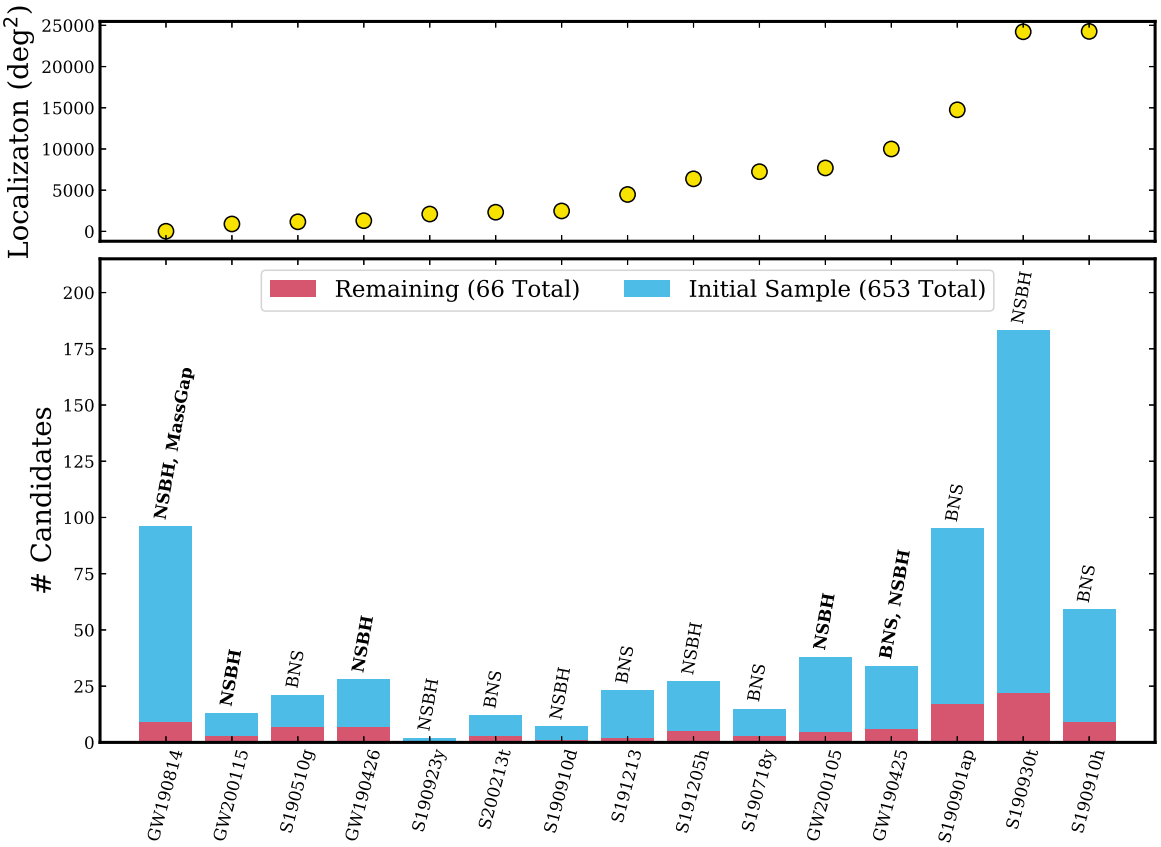

Figure 3. Bottom: a bar chart showing the initial number of candidates per event (blue) and the remaining number (red) per event after our critical analysis discussed in Sections 3 and 4. The events are sorted by the 90% confidence level localization size, shown in the top panel. The most likely event classifications are labeled at the top of each bar, with labels in bold showing candidates included in GWTC-2 or GWTC-3 (The LIGO Scientific Collaboration et al. 2021; Abbott et al. 2021a). With the exception of GW190814, the best-localized and most followed-up event of O3, the initial and remaining sample sizes generally scale with the localization size.

of the remaining candidates' $\nu L_{\nu,\max}$ are consistent with either the expected kilonova peak range or the on-axis short GRB afterglow range. However, we note that no coincident short GRBs were detected during O3, despite at least partial coverage of all events in our sample by either the Swift Observatory or Fermi Space Telescope. In addition, many of the maximum observed luminosities of remaining candidates are consistent with those of core-collapse, Type Ia, thermonuclear, or superluminous SNe (Figure 4).

In Figure 5, we plot the remaining candidates' light curves using all available information. We calculate luminosities and rest-frame times using the host redshift reported to TNS (points marked as squares), the spectroscopic (stars) or photometric (circles) redshift of a host found in Section 3.5 (when multiple photometric redshifts are available, we take a weighted average), or the GW event distance (open triangles). As the majority of photometric points are in the optical (filter pivot wavelengths are denoted by the marker colors), we overplot five diverse kilonova models in the r band.

Figure 5 shows that the majority of light curves are more luminous than the most optimistic kilonova model plotted, although we caution that the large errors of the GW event distance and photometric redshifts (including some that are consistent with $z = 0$) imply that the actual luminosity may vary widely. Additionally, 43 (71.7%) of the remaining candidates' light curves include only a single detection, indicating that we have no knowledge of their behavior over time and their true peak luminosity may be much higher than what we observe. The majority of candidate light curves observed beyond ∼5 days decline more slowly than all models except that of the stable NS remnant (Kasen et al. 2015). However, all of the slow-decaying light curves are ∼10 times more luminous than the model at the respective times. In general, the light curves of candidates that fall within the luminosity range of the plotted kilonova models consist of a single detection, preventing a comparison of their fading behavior or colors to the kilonova models. In addition, as demonstrated by the large fraction of triangle symbols in Figure 5, we are unable to associate the majority (53.3%) of the remaining candidates with a host galaxy cataloged in SDSS, LS DR9, or PS1-STRM. Finally, we note the lack of multiple bands of photometry (and thus color information) for nearly all candidates at early times, when color is a useful tool to distinguish kilonovae from other contaminants (Cowperthwaite & Berger 2015).

Ultimately, we do not find sufficient evidence in the photometric light curves to claim that any of the remaining viable candidates is a true kilonova. Future observations of kilonovae from blind searches (Doctor et al. 2017; Kasliwal et al. 2017; Smartt et al. 2017; Yang et al. 2017; Andreoni et al. 2020b, 2021), follow-up to short GRBs (see compilation studies of Gompertz et al. 2018; Ascenzi et al. 2019; Rossi et al. 2020; Rastinejad et al. 2021), or mergers detected by GWs during O4 will further our knowledge on the acceptable range of kilonova luminosities, colors, and decay timescales.

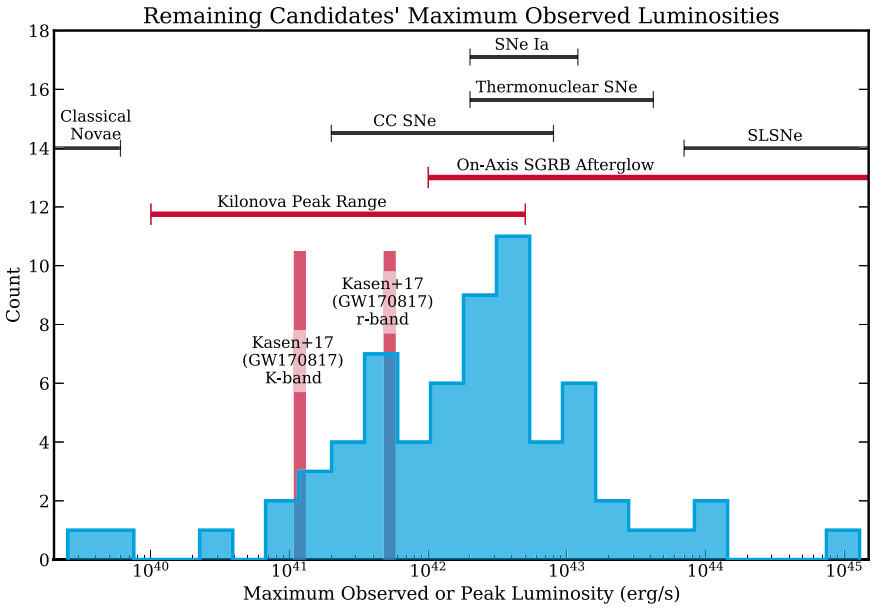

Figure 4. Histogram of the peak luminosities of 66 kilonova candidates still considered viable after our analysis. Luminosities are calculated using the host distance if known from TNS or our host analysis (Section 3.5), and the GW event distance otherwise. Horizontal bars at the top of the figure demonstrate the range of peak luminosities of other optical transients (from Bildsten et al. 2007; Darbha et al. 2010; Shen et al. 2010; Li et al. 2011; Kasliwal 2012; Cenko 2017). The average luminosity of remaining candidates is 2.4×10^{42} erg s^{-1}, on the upper end of the kilonova peak range, and consistent with the peak luminosities of short GRB afterglows and core-collapse, Type Ia, and thermonuclear SNe.

5.2. Examining Redundancy in Optical Candidate Follow-up

With the aim of making recommendations to improve candidate follow-up in future observing runs, we investigate two sources of redundancy. First, we contrast the use of targeted spectroscopic and photometric follow-up of O3 kilonova candidates with eliminations made using the real-time tools we consider in Section 3. Second, we examine how often multiple spectra of a single object were reported to the GCNs and discuss methods to reduce the number of redundant observations during O4.

Both spectroscopic follow-up and multiband imaging follow-up are time-intensive and expensive and are best reserved for promising candidates one cannot eliminate with other methods. Generally, the GCNs and literature are a reflection of real-time follow-up in connection to GW events. However, many relevant follow-up observations are only reported in the literature (which is often not published until months following the event), implying that relying on real-time tools and reports is key. Sixty-one candidates in our initial sample of 653 were followed up spectroscopically. Of these, 15 could have been eliminated based on pre-merger detections (although we note that the ATLAS forced photometry tool, on which the majority of these eliminations were based, was not publicly available during O3). Moreover, an additional 13 could have been eliminated based on inconsistencies between the GW and host-galaxy distances, all of which are "highly confident" host associations. Twelve eliminations were made based on photometric redshifts, while one was made based on a spectroscopic redshift. Our analysis utilizes the 95% GW distance credible interval to eliminate candidates based on their host-galaxy redshift. In cases where there are fewer candidate counterparts, follow-up groups may elect to use a wider interval. Though photometric redshifts are often inaccurate at low distances, they should still be used to prioritize follow-up spectra. We find that ≈46% of candidates that received spectroscopic follow-up observations could have instead benefited from real-time archival information, making follow-up observations somewhat redundant (Figure 6).

In a similar vein, 67 candidates had photometric follow-up reported. Of these, 1 was a match to a star in PS1, 13 have pre-merger detections, and 21 have associated host galaxies with photometric redshifts outside of the GW event distance range. Thus, ≈51% of candidates with reported photometric follow-up could have benefited from archival information. In Figure 6 we summarize the number of candidates with GCN or literature-reported spectra (left) or photometry (right) in pink. The black bars show the fraction of this follow-up that is redundant or could have been avoided by employing all the real-time tools described in Section 3.

We next examine another potential source of redundancy in follow-up resources: cases in which multiple follow-up spectra are taken of a single candidate. After examining the GCNs, we find seven initially promising candidates for which more than three spectra were reported, including AT 2019dzk (SN IIn; 4 spectra taken), AT 2019dzw (SN II; 5), AT 2019ebq (dust-reddened SN Ib/c; 8), AT 2019wqj (SN II; 3), AT 2019wxt (SN IIb, 7), AT 2020cja (blue, featureless continuum; 3), ZTFabvizsw (CV; 3). In general, all redundant spectra of a given candidate were obtained before any definitive classification results were reported. Notably, the majority of the candidates with multiple spectra were Type II SNe. For AT 2019ebq, the candidate with the most spectra taken, a near-IR spectrum was necessary to classify the dust-reddened SN (Carini et al. 2019; Dimitriadis et al. 2019; Jencson et al. 2019; McCully et al. 2019; Morokuma et al. 2019). Three out of seven of these candidates were potential counterparts to the NS merger GW190425. Generally, these well-followed candidates met some combination of the following criteria: (i) reported early in the search ($\delta t \lesssim 1$ day), (ii) showed red colors, and (iii) were associated with a host consistent with the GW distance. Overall, we find that obtaining multiple spectra of the same source was not a large sink on resources and was most notable

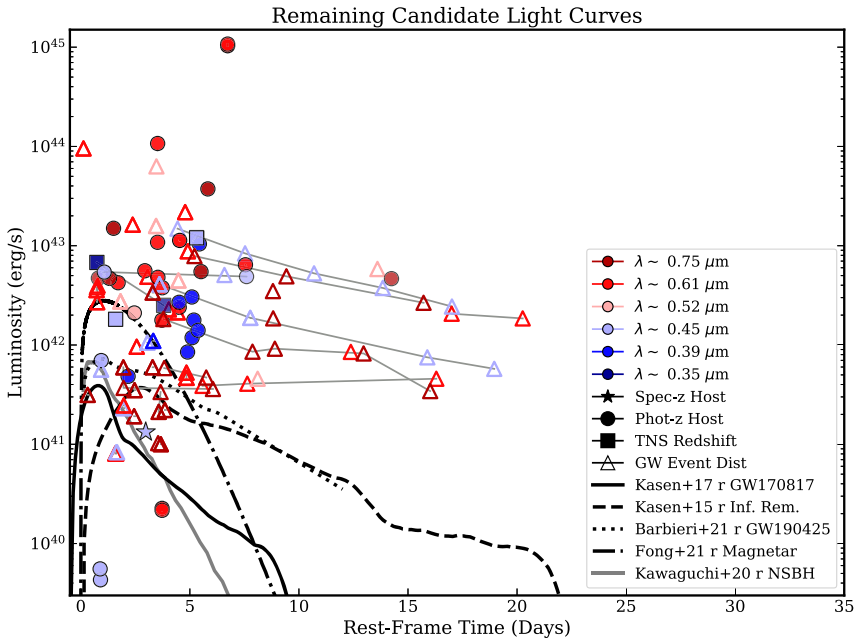

Figure 5. Light curves of the remaining viable kilonova candidates after our vetting analysis in Sections 3 and 4. We build light curves using photometry from the ATLAS forced photometry tool, public ZTF data, TNS, and the SAGUARO database. We correct all photometry for Milky Way extinction and convert the observations to luminosity and rest-frame time using the TNS-reported redshift, the redshift of the associated galaxy (we average the distances if multiple are available; Section 3.5) or, if none are available, the GW event distance. Marker symbols denote the source of the distance, and marker colors show the filter pivot wavelength. We also plot a diverse set of *r*-band ($\lambda \sim 0.61$ μm) kilonova models, including those of an AT 2017gfo-like kilonova (black solid line; Kasen et al. 2017), a fiducial kilonova from an NSBH merger (gray solid line; Kawaguchi et al. 2020b), an infinite-lifetime NS remnant kilonova (dashed line; Kasen et al. 2015), a kilonova modeled for the event GW190425 (dotted line; Barbieri et al. 2021), and a magnetar-boosted kilonova (dashed−dotted line; Fong et al. 2021).

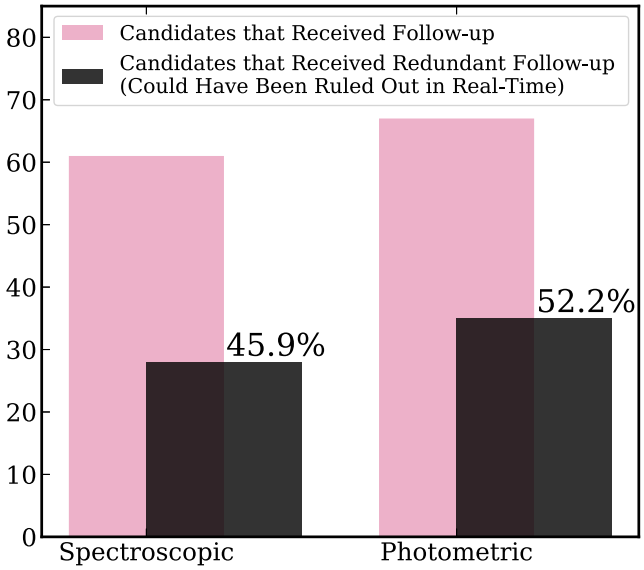

Figure 6. Bar chart representing the number of candidates with spectroscopic and photometric follow-up reported to the GCNs and literature (pink), and the number of these that we rule out with tools available in real time (black).

for events of high interest (GW190425 and GW190814). Indeed, if the candidate was the true counterpart, high-cadence early spectra would be essential for characterizing the kilonova (and would be critical for a comparison to the early, blue component of AT 2017gfo; Andreoni et al. 2017; Arcavi et al. 2017; Coulter et al. 2017; Cowperthwaite et al. 2017; Drout et al. 2017; Evans et al. 2017; Hu et al. 2017; Lipunov et al. 2017; Pian et al. 2017; Shappee et al. 2017; Smartt et al. 2017; Valenti et al. 2017; Villar et al. 2017) and its overall color evolution.

We note here several reasons for redundancy in follow-up and lessons learned. First, several of the tools that provide essential archival information were not available in O3, such as the ATLAS forced photometry tool and the LS DR9 photometric redshift catalog. The next era of EM follow-up to GW events will greatly benefit from the use of such tools. Additionally, we utilize the most updated GW localizations and distances in our analysis, which are generally closer to final values, thus allowing us to eliminate candidates that might have been viable kilonovae at the time of follow-up. For example, the photometric redshift of the host we associate with AT 2019dzk in Section 3.5 is inconsistent with the updated GW190425 distance but was consistent with the GW distance reported at the time follow-up spectra were taken. We note that some groups may choose to follow candidates that exhibit color and fading behavior similar to AT 2017gfo, even if the photometric redshift is inconsistent (especially given that they become less robust for $z \lesssim 0.1$). Further, we calculate the number of GCN and TNS candidates per event that meet our initial criteria (Section 2.2) using the localization maps first announced in the GCNs. We compare these numbers to those found using the final localizations. Predictably, for events with large final localizations (e.g., S190901ap, S190930t) the number of candidates does not change significantly using the preliminary maps. However, for half of the events in our sample, the number of candidates increases by 30%–140% when the preliminary localizations are utilized (GW190425, GW190426, GW190814, S190910d, GW200105, GW200115, S200213t), indicating the power of prompt localization updates in reducing the strain on follow-up resources. Together, these points highlight the importance of real-time updates from the

GW community to the EM community, specifically of localization maps and event distances.

Overall, an examination of follow-up redundancies highlights the need for implementation of real-time tools that leverage archival information, improved organization of the community's follow-up plans and results, and updated key GW parameters (such as distance, localization maps, component masses, and mass ratios) in future observing runs. Potential ways to assist community organization include reducing and reporting spectra as quickly as possible, reporting intended follow-up to the GCNs, or the creation of a database to report planned observations (similar to the Gravitational Wave Treasure Map, which aids the community by posting pointings; Wyatt et al. 2020).

5.3. Comparison to Other Kilonova Candidate Studies

Finally, for context we comment on the methods used in our analysis compared to those employed by other GW follow-up groups. Many works focused on the candidates reported in their own searches, with the exception of some works that analyzed all publicly reported candidates for a given event (e.g., Hosseinzadeh et al. 2019; Kilpatrick et al. 2021), and often the original samples of candidates were defined differently. For instance, Kilpatrick et al. (2021) took a more conservative approach in their study of GW190814 by including candidates reported to TNS within 14 days of the merger and within the 99% localization. Their sample included 214 candidates, over twice the size of our initial GW190814 candidate sample. As we are analyzing a larger sample of GW events, we utilize the 90% localization contour, a stricter cut for δt motivated by kilonova light-curve models, and make a preliminary luminosity cut (Section 2).

We find that cross-matching to stellar catalogs, such as PS1, Gaia, the Two Micron All Sky Survey, SDSS, USNO-B, and DES, is a fairly ubiquitous practice in GW optical candidate vetting. In our analysis, the PS1 point-source classification eliminated the greatest number of candidates (38) of the stellar catalogs that we considered. Cross-matching to AGNs was not always standard practice. Two examples of strategies include cross-matching to the MILLIQUAS catalog (Flesch 2015) and identifying AGN-like colors with Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010) observations in combination with a nuclear position (Ackley et al. 2020; Kasliwal et al. 2020). Our analysis found 26 quasars by cross-matching (Section 3.3), 20 of which also had pre-merger detections, indicating that the majority of quasars could be eliminated using the methods outlined in Section 3.4.

The largest variation in candidate vetting is observed in the methods employed to associate candidates with host galaxies and the use of photometric or spectroscopic redshifts. Databases or catalogs used include the NASA/IPAC Extragalactic Database, the Census of the Local Universe (Cook et al. 2019), Galaxy List for the Advanced Detector Era (Dalya et al. 2018), LS DR9, PS1-STRM, and DES Y3 (Hartley et al. 2022; Kasliwal et al. 2017; Ackley et al. 2020; Antier et al. 2020a; Morgan et al. 2020; Kilpatrick et al. 2021; Paterson et al. 2021). We find that SDSS DR12, PS1-STRM, and LS DR9 footprints combined cover >97% of kilonova candidates queried (Section 3.5). Precise association methods varied between searching web interfaces and determining hosts by eye (Ackley et al. 2020), calculating the galaxy with the minimum projected offset (Kilpatrick et al. 2021), and calculating probability based on angular offset and redshift (following the prescriptions of Singer et al. 2016; Morgan et al. 2020). In our analysis, the SDSS, PS1, and LS DR9 redshift catalogs are the most effective tools at eliminating candidates, although we note that at low distances photometric redshifts are not always robust. Future spectroscopic redshift surveys (e.g., DESI, Subaru; DESI Collaboration et al. 1611; Takada et al. 2014) will be valuable tools for eliminating candidates.

Photometric detections were also frequently used in the literature to eliminate candidates based on their light curves, although it was not clear whether this is standard practice in real time. Some works examined their own data streams (e.g., GRAWITA, ZTF, SAGUARO/CSS), and some utilized public data streams (e.g., the PS1 Detection catalog, TNS, ZTF, the VISTA archive; Ackley et al. 2020; Kasliwal et al. 2020; Kilpatrick et al. 2021; Paterson et al. 2021). Others obtained follow-up observations at later times to examine their candidates' late-time light curves (Morgan et al. 2020). This tool was the second most effective at eliminating candidates in our analysis. Looking forward, we recommend that surveys make these available, as they promise to be an invaluable tool in eliminating candidates in real time, especially as the number of detected NS mergers grows in subsequent years. Eventually, publicly available, deep, multiband observations from the Vera C. Rubin Observatory will transform the search for pre-explosion detections, at least in the southern hemisphere.

Overall, we find that several of the steps we apply in Section 3 are not standard practice among the O3 literature and can significantly improve community candidate vetting in future observing runs. The expansion of new surveys and resources will also certainly enhance the candidate vetting process in future observing runs.

6. Conclusion and Future Prospects

We analyze 653 optical candidate counterparts to 15 O3 GW mergers involving at least one NS using a combination of information available at the time of the GW event ("real time") and that available days to weeks afterward. The number of candidates in our initial sample per event roughly scales with the size of the 90% c.l. localization. The notable exception to this is GW190814, the best-localized multimessenger prospect in O3, indicating that similarly well-localized O4 events will also result in large numbers of candidates. At the conclusion of our analysis, we find that only 66 candidates remain viable, none of which have sufficient information to be claimed as a real kilonova. We also review the GCNs and literature and make recommendations for avoiding redundant observations to classify a candidate in O4. Our main conclusions are as follows:

1. Employment of the real-time tools (including pre-merger detections and cross-matching with catalogs) that use archival information eliminated 65% of the original candidate sample as viable kilonovae. In particular, pre-merger detections in public surveys account for >20% of eliminations alone, and 15 of these still received follow-up observations. Availability and incorporation of these tools into follow-up of future GW events will allow the community to focus limited follow-up resources and reduce redundancy.
2. The most effective real-time tool at eliminating candidates as viable kilonovae was association with host

galaxies in public photometric or spectroscopic redshift surveys outside the 95% GW event distance. The combination of PS1-STRM, SDSS, and LS DR9 covered the footprint of >97% of candidates queried. Future spectroscopic redshift surveys will increase the robustness of host-galaxy redshifts. Meanwhile, photometric redshifts are an important tool for prioritizing classification resources.

3. At the conclusion of our analysis, 66 candidates remain viable as kilonovae, although the majority have insufficient information to be considered otherwise (single data point, unidentified host). The remaining candidates with light curves and redshifts would be particularly luminous if they were kilonovae, although given the diversity of model luminosities, they cannot be confidently eliminated as such.
4. Increased collaboration and transparency between and within the GW and EM communities would facilitate the search for EM counterparts. For instance, tools that can reduce redundancy or increase transparency among EM follow-up groups should be more widely adopted. Moreover, the prompt release of updated localization maps and distance measurements in particular would reduce the number of kilonova candidates that pass initial vetting, and updated component masses would help to prioritize the use of limited follow-up resources.

Looking forward, the larger volumes probed by GW detectors make the issue of candidate contamination increasingly urgent. It is imperative to take advantage of any available tool that leverages the wealth of existing or follow-up data, as well as to build tools that facilitate community follow-up (Chang et al. 2019; Wyatt et al. 2020; Tak et al. 2021).

In tandem with observational strides, theoretical works predict a wide diversity in the timescales, colors, and peak luminosities of kilonovae (e.g., Li & Paczyński 1998; Metzger & Fernandez 2014; Lippuner et al. 2017; Shibata & Hotokezaka 2019; Kawaguchi et al. 2020a). With concurrent GW observations, it will be possible to connect each kilonova and its *r*-process abundance to the observed population of NSs and BHs. Each successive GW observing run has brought new and exciting discoveries; the methods presented in this work, along with many other developments in GW-EM astronomy, set the stage for novel, multimessenger revelations in forthcoming observing runs.

The authors gratefully acknowledge Charlie Kilpatrick, Chase Kimball, Zoheyr Doctor, Adam Miller, Stefano Valenti, Stephanie Juneau, and Dustin Lang for their helpful discussions in preparing this manuscript. SAGUARO is supported by the National Science Foundation (NSF) under award Nos. AST-1909358 and AST-1908972.

Time domain research by D.J.S. is supported by NSF grants AST-1821987, 1813466, and 2108032 and by the Heising-Simons Foundation under grant No. 2020-1864. The Fong Group at Northwestern acknowledges support by the National Science Foundation under grant No. AST-1814782 and CAREER grant No. AST-2047919. W.F. gratefully acknowledges support by the David and Lucile Packard Foundation. S.Y. has been supported by the research project grant Understanding the Dynamic Universe funded by the Knut and Alice Wallenberg Foundation under Dnr KAW 2018.0067, and the G.R.E.A.T research environment, funded by Vetenskapsrådet, the Swedish Research Council, project No. 2016-06012, and the Wenner-Gren Foundations.

The operation of the facilities of Steward Observatory is supported in part by the state of Arizona.

This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

This research has made use of data and/or services provided by the International Astronomical Union's Minor Planet Center.

This work has made use of data from the Asteroid Terrestrial-impact Last Alert System (ATLAS) project. The Asteroid Terrestrial-impact Last Alert System (ATLAS) project is primarily funded to search for near-Earth asteroids through NASA grants NN12AR55G, 80NSSC18K0284, and 80NSSC 18K1575; by-products of the NEO search include images and catalogs from the survey area. This work was partially funded by Kepler/K2 grant J1944/80NSSC19K0112 and HST GO-15889 and STFC grants ST/T000198/1 and ST/S006109/1. The ATLAS science products have been made possible through the contributions of the University of Hawaii Institute for Astronomy, the Queens University Belfast, the Space Telescope Science Institute, the South African Astronomical Observatory, and the Millennium Institute of Astrophysics (MAS), Chile.

The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID #2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing-Arizona Sky Survey (BASS; NOAO Prop. ID #2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; Prop. ID #2016A-0453; PI: Arjun Dey). DECaLS, BASS, and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo a Pesquisa do

Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientifico e Tecnologico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft, and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenossische Technische Hochschule (ETH) Zurich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciencies de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig Maximilians Universitat Munchen and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories of China, the Chinese Academy of Sciences (the Strategic Priority Research Program "The Emergence of Cosmological Structures" grant No. XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. BASS is also supported by the External Cooperation Program of Chinese Academy of Sciences (grant No. 114A11KYSB 20160057) and the Chinese National Natural Science Foundation (grant No. 11433005).

The Legacy Survey team makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a project of the Jet Propulsion Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration.

The Legacy Surveys imaging of the DESI footprint is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under contract No. DE-AC02-05CH1123, by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; and by the U.S. National Science Foundation, Division of Astronomical Sciences under contract No. AST-0950945 to NOAO.

Facility: SO:1.5m.

Software: astropy (Astropy Collaboration et al. 2013, 2018), Numpy (Harris et al. 2020), The IDL Astronomy User's Library (Landsman 1993), SCAMP (Bertin 2006, 2010a), SWarp (Bertin 2010b), IRAF (Tody 1986, 1993), SExtractor (Bertin & Arnouts 1996), ZOGY (https://github.com/pmvreeswijk/ZOGY).

ORCID iDs

J. C. Rastinejad https://orcid.org/0000-0002-9267-6213
K. Paterson https://orcid.org/0000-0001-8340-3486
W. Fong https://orcid.org/0000-0002-7374-935X
D. J. Sand https://orcid.org/0000-0003-4102-380X
M. J. Lundquist https://orcid.org/0000-0001-9589-3793
G. Hosseinzadeh https://orcid.org/0000-0002-0832-2974
A. R. Gibbs https://orcid.org/0000-0002-2575-2618
S. Hall https://orcid.org/0000-0002-3841-380X
S. Yang https://orcid.org/0000-0002-2898-6532

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016a, PhRvL, 116, 061102
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016b, LRR, 21, 3
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, PhRvL, 119, 161101
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, Natur, 551, 85
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2019, PhRvX, 9, 031040
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2020a, ApJL, 892, L3
Abbott, R., Abbott, T. D., Abraham, S., et al. 2020b, ApJL, 896, L44
Abbott, R., Abbott, T. D., Abraham, S., et al. 2021a, PhRvX, 11, 021053
Abbott, R., Abbott, T. D., Abraham, S., et al. 2021b, ApJL, 915, L5
Ackley, K., Amati, L., Barbieri, C., et al. 2020, A&A, 643, A113
Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, ApJS, 219, 12
Alexander, K. D., Schroeder, G., Paterson, K., et al. 2021, ApJ, 923, 66
Anand, S., Coughlin, M. W., Kasliwal, M. M., et al. 2021, NatAs, 5, 46
Andreoni, I., Ackley, K., Cooke, J., et al. 2017, PASA, 34, e069
Andreoni, I., Coughlin, M. W., Kool, E. C., et al. 2021, ApJ, 918, 63
Andreoni, I., Goldstein, D. A., Anand, S., et al. 2019a, ApJL, 881, L16
Andreoni, I., Goldstein, D. A., Coughlin, M., et al. 2019b, GCN, 24268, 1
Andreoni, I., Goldstein, D. A., Kasliwal, M. M., et al. 2020a, ApJ, 890, 131
Andreoni, I., Kool, E. C., Sagués Carracedo, A., et al. 2020b, ApJ, 904, 155
Antier, S., Agayeva, S., Aivazyan, V., et al. 2020a, MNRAS, 492, 3904
Antier, S., Agayeva, S., Almualla, M., et al. 2020b, MNRAS, 497, 5518
Arcavi, I., Hosseinzadeh, G., Howell, D. A., et al. 2017, Natur, 551, 64
Ascenzi, S., Coughlin, M. W., Dietrich, T., et al. 2019, MNRAS, 486, 672
Ashton, G., Ackley, K., Hernandez, I. M., & Piotrzkowski, B. 2021, CQGra, 38, 235004
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Barbieri, C., Salafia, O. S., Colpi, M., Ghirlanda, G., & Perego, A. 2021, A&A, 654, A12
Barnes, J., & Kasen, D. 2013, ApJ, 775, 18
Becerra, R. L., Dichiara, S., Watson, A. M., et al. 2021, MNRAS, 507, 1401
Beck, R., Dobos, L., Budavari, T., Szalay, A. S., & Csabai, I. 2016, MNRAS, 460, 1371
Beck, R., Szapudi, I., Flewelling, H., et al. 2021, MNRAS, 500, 1633
Bellm, E. C., Kulkarni, S. R., Graham, M. J., et al. 2019, PASP, 131, 018002
Bennett, C. L., Larson, D., Weiland, J. L., & Hinshaw, G. 2014, ApJ, 794, 135
Bertin, E. 2006, in ASP Conf. Ser. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel (San Francisco, CA: ASP), 112
Bertin, E. 2010a, SCAMP: Automatic Astrometric and Photometric Calibration, Astrophysics Source Code Library, ascl:1010.063
Bertin, E. 2010b, SWarp: Resampling and Co-adding FITS Images Together, Astrophysics Source Code Library, ascl:1010.068
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bhakta, D., Mooley, K. P., Corsi, A., et al. 2021, ApJ, 911, 77
Bildsten, L., Shen, K. J., Weinberg, N. N., & Nelemans, G. 2007, ApJL, 662, L95
Bloom, J. S., Kulkarni, S. R., & Djorgovski, S. G. 2002, AJ, 123, 1111
Brennan, S., Callis, E., Ihanec, N., Gromadzki, M., & Irani, I. 2019, TNSCR, 2019-1773, 1
Broekgaarden, F. S., Berger, E., Neijssel, C. J., et al. 2021, MNRAS, 508, 5028
Carini, R., Izzo, L., Palazzi, E., et al. 2019, GCN, 24252, 1
Cenko, S. B. 2017, NatAs, 1, 0008
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
Chang, P., Allen, G., Anderson, W., et al. 2019, BAAS, 51, 436
Chang, S.-W., Onken, C. A., Wolf, C., et al. 2021, PASA, 38, e024
Chase, E. A., O'Connor, B., Fryer, C. L., et al. 2021, arXiv:2105.12268
Christensen, E., Africano, B., Farneth, G., et al. 2018, AAS/DPS Meeting Abstracts, 310, 10
Cook, D. O., Kasliwal, M. M., Van Sistine, A., et al. 2019, ApJ, 880, 7
Coppejans, D. L., Margutti, R., Terreran, G., et al. 2020, ApJL, 895, L23
Coughlin, M. W., Ahumada, T., Anand, S., et al. 2019, ApJL, 885, L19
Coulter, D. A., Foley, R. J., Kilpatrick, C. D., et al. 2017, Sci, 358, 1556
Cowperthwaite, P. S., & Berger, E. 2015, ApJ, 814, 25
Cowperthwaite, P. S., Berger, E., Villar, V. A., et al. 2017, ApJL, 848, L17
Dahiwale, A., & Fremling, C. 2019, TNSCR, 2019-2407, 1
Dalya, G., Galgoczi, G., Dobos, L., et al. 2018, MNRAS, 479, 2374
Darbha, S., Metzger, B. D., Quataert, E., et al. 2010, MNRAS, 409, 846
de Wet, S., Groot, P. J., Bloemen, S., et al. 2021, A&A, 649, A72
deJaeger, T., Shappee, B. J., Kochanek, C. S., et al. 2022, MNRAS, 509, 3427
DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 1611, arXiv:1611.00036
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168

Dichiara, S., Becerra, R. L., Chase, E. A., et al. 2021, ApJL, 923, L32
Dimitriadis, G., Jones, D. O., Siebert, M. R., et al. 2019, GCN, 24358, 1
Dobie, D., Stewart, A., Hotokezaka, K., et al. 2022, MNRAS, 510, 3794
Dobie, D., Stewart, A., Murphy, T., et al. 2019, ApJL, 887, L13
Doctor, Z., Kessler, R., Chen, H. Y., et al. 2017, ApJ, 837, 57
Drout, M. R., Chornock, R., Soderberg, A. M., et al. 2014, ApJ, 794, 23
Drout, M. R., Piro, A. L., Shappee, B. J., et al. 2017, Sci, 358, 1570
Evans, P. A., Cenko, S. B., Kennea, J. A., et al. 2017, Sci, 358, 1565
Flesch, E. W. 2015, PASA, 32, e010
Flesch, E. W. 2021, yCat, VII/290
Fong, W., & Berger, E. 2013, ApJ, 776, 18
Fong, W., Berger, E., Margutti, R., & Zauderer, B. A. 2015, ApJ, 815, 102
Fong, W., Laskar, T., Rastinejad, J., et al. 2021, ApJ, 906, 127
Foucart, F., Deaton, M. B., Duez, M. D., et al. 2013, PhRvD, 87, 084006
Fremling, C., Dahiwale, A., & Dugas, A. 2019, TNSCR, 2019-1923, 1
Fryer, C. L., Belczynski, K., Ramirez-Ruiz, E., et al. 2015, ApJ, 812, 24
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, A&A, 649, A1
Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1
Garcia, A., Morgan, R., Herner, K., et al. 2020, ApJ, 903, 75
Goldstein, D. A., Andreoni, I., Nugent, P. E., et al. 2019, ApJL, 881, L7
Gomez, S., Hosseinzadeh, G., Cowperthwaite, P. S., et al. 2019, ApJL, 884, L55
Gompertz, B. P., Cutter, R., Steeghs, D., et al. 2020, MNRAS, 497, 726
Gompertz, B. P., Levan, A. J., Tanvir, N. R., et al. 2018, ApJ, 860, 62
Graham, M. J., Ford, K. E. S., McKernan, B., et al. 2020, PhRvL, 124, 251102
Gromadzki, M. 2019, TNSCR, 2019-2193, 1
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hartley, W. G., Choi, A., Amon, A., et al. 2022, MNRAS, 509, 3547
Hosseinzadeh, G., Cowperthwaite, P. S., Gomez, S., et al. 2019, ApJL, 880, L4
Hu, L., Wu, X., Andreoni, I., et al. 2017, SciBu, 62, 1433
Jayasinghe, T., Stanek, K. Z., Kochanek, C. S., et al. 2019, MNRAS, 485, 961
Jencson, J., de, K., Anand, S., et al. 2019, GCN, 24233, 1
Kalogera, V., & Baym, G. 1996, ApJL, 470, L61
Kasen, D., Fernandez, R., & Metzger, B. D. 2015, MNRAS, 450, 1777
Kasen, D., Metzger, B., Barnes, J., Quataert, E., & Ramirez-Ruiz, E. 2017, Natur, 551, 80
Kasliwal, M. M. 2012, PASA, 29, 482
Kasliwal, M. M., Anand, S., Ahumada, T., et al. 2020, ApJ, 905, 145
Kasliwal, M. M., Nakar, E., Singer, L. P., et al. 2017, Sci, 358, 1559
Kawaguchi, K., Kyutoku, K., Shibata, M., & Tanaka, M. 2016, ApJ, 825, 52
Kawaguchi, K., Shibata, M., & Tanaka, M. 2020a, ApJ, 889, 171
Kawaguchi, K., Shibata, M., & Tanaka, M. 2020b, ApJ, 893, 153
Kilpatrick, C. D., Coulter, D. A., Arcavi, I., et al. 2021, ApJ, 923, 258
Korobkin, O., Wollaeger, R. T., Fryer, C. L., et al. 2021, ApJ, 910, 116
Landsman, W. B. 1993, in ASP Conf. Ser. 52, Astronomical Data Analysis Software and Systems II, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes (San Francisco, CA: ASP), 246
Li, L.-X., & Paczyński, B. 1998, ApJL, 507, L59
Li, W., Chornock, R., Leaman, J., et al. 2011, MNRAS, 412, 1473
LIGO Scientific Collaboration, & Virgo Collaboration 2019a, GCN, 25324, 1
LIGO Scientific Collaboration, & Virgo Collaboration 2019b, GCN, 25333, 1
The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration et al. 2021, arXiv:2111.03606
Lippuner, J., Fernandez, R., Roberts, L. F., et al. 2017, MNRAS, 472, 904
Lipunov, V. M., Gorbovskoy, E., Kornilov, V. G., et al. 2017, ApJL, 850, L1
Lundquist, M. J., Paterson, K., Fong, W., et al. 2019, ApJL, 881, L26
Lyke, B. W., Higley, A. N., McLane, J. N., et al. 2020, ApJS, 250, 8
Margutti, R., Metzger, B. D., Chornock, R., et al. 2019, ApJ, 872, 18
Masci, F. J., Laher, R. R., Rusholme, B., et al. 2019, PASP, 131, 018003
McCully, C., Hiramatsu, D., Hiramatsu, D., et al. 2019, GCN, 24295, 1
Metcalfe, N., Farrow, D. J., Cole, S., et al. 2013, MNRAS, 435, 1825
Metzger, B. D. 2019, LRR, 23, 1
Metzger, B. D., & Fernandez, R. 2014, MNRAS, 441, 3444
Metzger, B. D., Martinez-Pinedo, G., Darbha, S., et al. 2010, MNRAS, 406, 2650
Morgan, R., Soares-Santos, M., Annis, J., et al. 2020, ApJ, 901, 83
Morokuma, T., Ohta, K., Yoshida, M., et al. 2019, GCN, 24230, 1
Nicholl, M., Margalit, B., Schmidt, P., et al. 2021, MNRAS, 505, 3016
Oates, S. R., Marshall, F. E., Breeveld, A. A., et al. 2021, MNRAS, 507, 1296
Ohgami, T., Tominaga, N., Utsumi, Y., et al. 2021, PASJ, 73, 350
Paterson, K., Lundquist, M. J., Rastinejad, J. C., et al. 2021, ApJ, 912, 128
Perna, R., Chruslinska, M., Corsi, A., & Belczynski, K. 2018, MNRAS, 477, 4228
Pian, E., D'Avanzo, P., Benetti, S., et al. 2017, Natur, 551, 67
Pozanenko, A. S., Minaev, P. Y., Grebenev, S. A., & Chelovekov, I. V. 2020, AstL, 45, 710
Rastinejad, J. C., Fong, W., Kilpatrick, C. D., et al. 2021, ApJ, 916, 89
Rhoades, C. E., & Ruffini, R. 1974, PhRvL, 32, 324
Rossi, A., Stratta, G., Maiorano, E., et al. 2020, MNRAS, 493, 3379
Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103
Shappee, B. J., Simon, J. D., Drout, M. R., et al. 2017, Sci, 358, 1574
Shen, K. J., Kasen, D., Weinberg, N. N., Bildsten, L., & Scannapieco, E. 2010, ApJ, 715, 767
Shibata, M., & Hotokezaka, K. 2019, ARNPS, 69, 41
Silva, D. R., Blum, R. D., Allen, L., et al. 2016, AAS Meeting Abstracts, 228, 317.02
Singer, L. P., Chen, H.-Y., Holz, D. E., et al. 2016, ApJS, 226, 10
Smartt, S. J., Chen, T. W., Jerkstrand, A., et al. 2017, Natur, 551, 75
Smith, K. W., Smartt, S. J., Young, D. R., et al. 2020, PASP, 132, 085002
Soares-Santos, M., Holz, D. E., Annis, J., et al. 2017, ApJL, 848, L16
Strader, J. 2019, TNSCR, 2019-2125, 1
Tachibana, Y., & Miller, A. A. 2018, PASP, 130, 128001
Tak, D., Gibb, M., McGlynn, T., et al. 2021, GCN, 31036, 1
Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ, 66, R1
Tanaka, M., & Hotokezaka, K. 2013, ApJ, 775, 113
Tanvir, N. R., Levan, A. J., González-Fernández, C., et al. 2017, ApJL, 848, L27
Thakur, A. L., Dichiara, S., Troja, E., et al. 2020, MNRAS, 499, 3868
Tody, D. 1986, Proc. SPIE, 627, 733
Tody, D. 1993, in ASP Conf. Ser. 52, Astronomical Data Analysis Software and Systems II, ed. R. J. Hanisch, R. J. V. Brissenden, & J. Barnes (San Francisco, CA: ASP), 173
Tonry, J. L., Denneau, L., Heinze, A. N., et al. 2018, PASP, 130, 064505
Tucker, D., Wiesner, M., Allam, S., et al. 2021, arXiv:2109.13351
Valenti, S., Sand, D. J., Yang, S., et al. 2017, ApJL, 848, L24
Vieira, N., Ruan, J. J., Haggard, D., et al. 2020, ApJ, 895, 96
Villar, V. A., Guillochon, J., Berger, E., et al. 2017, ApJL, 851, L21
Watson, A. M., Butler, N. R., Lee, W. H., et al. 2020, MNRAS, 492, 5916
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Wyatt, S. D., Tohuvavohu, A., Arcavi, I., et al. 2020, ApJ, 894, 127
Yang, S., Valenti, S., Cappellaro, E., et al. 2017, ApJL, 851, L48
Zhou, R., Newman, J. A., Mao, Y., et al. 2021, MNRAS, 501, 3309
Zimmerman, E., Irani, I., Schulze, S., Bruch, R., & Yaron, O. 2020, TNSCR, 2020-168, 1
Zou, H., Zhou, X., Fan, X., et al. 2017, PASP, 129, 064101