Uncertainties in kilonova modeling

C. L. Fryer*, C. J. Fontes, O. Korobkin, M. Mumpower and R. Wollaeger

Center for Theoretical Astrophysics, Los Alamos National Laboratory,
Los Alamos, NM 87545, USA
* E-mail: fryer@lanl.gov
https://ccsweb.lanl.gov/astro/index.html

E.M. Holmbeck
The Observatories of the Carnegie Institute for Science,
Pasadena, CA, 91101, USA

R. O’Shaugnessy
Rochester Institute of Technology,
Rochester, NY 14623

The detection of the merger of a neutron star binary in both gravitational waves and a broad spectrum of electromagnetic waves (GW170817) provided the most compelling evidence to date that such mergers produce heavy r-process elements. The inferred rate of these mergers coupled to the estimated r-process production suggests that these mergers could produce nearly all of the r-process elements in the universe. However, uncertainties in the merger rate and the amount of r-process production per merger means that scientists can not constrain the fraction of the merger r-process contribution to better than 1–100% of the total amount in the universe. The total r-process mass synthesized is best constrained by the observations themselves and uncertainties in the inferred production quantity follows from the uncertainties in modeling the emission from the NSM ejecta. In this paper, we review these modeling uncertainties.

Keywords: Neutron Stars; Gravitational Waves, r-Process

1. Kilonovae, a Key to Understanding r-Process

Many of the heaviest elements in the universe are produced through a process of rapid neutron capture (r-process), building isotopes far from stability that then decay to the long-lived heavy isotopes observed in the universe. This rapid neutron capture requires extreme conditions with high densities and large fractions of free neutrons. Scientists identified neutron star mergers (NSMs) as a potential source of these r-process elements over 40 years ago\(^1,2\) and evidence has been growing steadily to argue that these mergers could dominate the r-process production.

Although the evolution of the ejecta from these mergers reaches ideal conditions (high densities, temperatures and neutron fractions) for the production of r-process elements, observational validation of r-process production in NSMs has been much more difficult to obtain. Until recently, theoretical and observational evidence relied on galactic chemical evolution arguments and their associated uncertainties.\(^3,4\)
This situation changed with the joint electromagnetic- and gravitational-wave observations of NSM GW170817.\textsuperscript{5} Between the gamma-, x-, optical, infra-red, and radio emission, it became clear that the NSM in GW170817 produced both a relativistic jet and a lower-velocity mass ejecta that argued for both neutron-rich dynamical ejecta and higher electron-fraction wind ejecta components.

Although initial analyses argued strongly that GW170817 produced a large fraction of r-process (0.04M\(_\odot\)), as more studies were completed, uncertainties in the inferred r-process production from GW170817 grew.\textsuperscript{6} Many of these differences arose from different assumptions about the properties of the outflow material and a difference in which observations were used to constrain the data. But these are just some of the uncertainties in the analysis of the optical and infra-red emission from NSMs. In this paper, we review these uncertainties. In section 2, we review the uncertainties in the properties (distributions of the velocity, density, entropy and neutron fraction) of the outflow from these mergers. Section 3 reviews some of the microphysics uncertainties in these calculations. Section 4 details alternative power sources that can alter the mass inferred from observations.

2. Uncertainties in the Outflows

The emission from astrophysical transients depends upon the ejecta properties (e.g. composition, velocity and angular distributions). Unlike many transients that are produced by a single outburst, the ejecta from NSMs arises from multiple components including ejecta from the initial tidal tails produced in the dynamical merger and a outflow arising from a disk. The disk outflow is driven by viscous forces and neutrino emission from a post-disruption torus of material (Figure 1). We will refer to the former as dynamical ejecta and the latter as wind ejecta. The bulk of the dynamical material is ejected along the orbital plane. This material has a high-neutron fraction, ideally suited for the production of the most massive elements (heavy r-process). In contrast, the wind ejecta is more isotropic and consists of material spanning a broad range of neutron fractions that depend on both the properties of the disk and the fate of the merged core.

Many of the initial studies of GW170817 used spherically-symmetric models, varying the fraction of heavy r-process to lighter elements.\textsuperscript{6} Some of these initial models used two components to differentiate between the dynamical ejecta primarily flowing along the orbital plane and a more symmetric or axis-aligned wind outflow\textsuperscript{7} shown in Figure 1. However, even these two-component models made simplifying assumptions using a single composition for each component and a fixed velocity distribution of the ejecta. These assumptions oversimplify the nature of the ejecta and can have order-of-magnitude effects on the inferred mass of r-process ejecta. Before we study how these two components fit together, we first review some of the outflow uncertainties in the individual dynamical and wind components.

Although models of the dynamical ejecta have produced a range of ejecta masses, e.g. compare the results in,\textsuperscript{8,9} these numerical differences are slowly diminishing as
compositional methods improve. Even with this convergence, the ejecta mass will vary with the individual masses of the merging compact objects. The composition of this ejecta is much better constrained than the ejecta mass. At high neutron fractions (above 0.8), the final composition does not depend sensitively on the neutron fraction. Because of this, the small variation in the neutron fraction does not alter the Lanthanide composition considerably and the composition of this ejecta is fairly stable. In addition, the final composition is also less sensitive to the details of the outflow (density and temperature evolution). For this reason, the composition of this ejecta is well-understood with the dominant uncertainties depending on our lack of understanding of the nuclear physics (Sec. 3). Finally, due to the dense forest of absorption lines in the opacity produced by the heavy r-process lanthanides and actinides, the light curves do not depend sensitively on the exact relative abundance...
of these elements. The primary uncertainties in models of the dynamical ejecta emission lie in our understanding of the total ejecta mass and outflow properties (velocity distribution).

The properties of the wind ejecta are much more uncertain. The mass, velocity distribution and composition of this ejecta varies considerably from model to model. Whereas the neutron fraction is high and exhibits only small variations for dynamical ejecta, the neutron fraction ranges across a broad range in the wind ejecta. Further, at the lower neutron fractions (between 0.5-0.8), the yields become much more sensitive to the exact value of the neutron fraction, so the composition of the wind ejecta can vary wildly. To further complicate these light-curve models, disk models of this wind predict a range of compositions that vary with angle and time. This composition variation is not taken into account in most of the two-component models currently in the literature. In addition, current wind models typically only follow the wind out to, at most, 1 second and much of the ejecta is below the escape velocity (but still accelerating) at this time. Although models are being constructed to follow this outflow to later times, we still do not have accurate models of this ejecta’s velocity distribution.

Fig. 2. Bolometric Light curves versus time for a broad range of models varying the morphology of the two components in a dynamical plus wind model. Varying the morphology alone can change the bolometric luminosity by over an order of magnitude.
Although we have mostly discussed the role of the composition, the morphology of the ejecta (velocity and angular distribution) may be even more important. Figure 2 shows a broad set of models with the same explosion energy and mass, but varying the angular morphology of the ejecta. In this study, two-component models were developed, each with a different morphology. In these models, although the energy of the ejecta is the same, the velocity distribution of the ejecta can vary. Changing the morphology alters both the peak luminosity and the time of that peak, varying both by an order of magnitude.

At this time, a complete study of all the outflow uncertainties has not been done. We end this section with the discussion of a final study varying ejecta mass and velocity of a two-component model. This grid of models used a single dynamical ejecta composition and one of two wind compositions. It focused on a primary pair of morphologies for these two components, using a default r-process composition for the dynamical ejecta and two different compositions for the wind. But the compositions were fixed for each component. Although the peak velocity was changed, the velocity distribution was unaltered. Depending on the relative velocities of the two components, the dynamical ejecta, with its high-opacity lanthanides, can obscure the wind ejecta. In these cases, the luminosity can depend sensitively on the viewing angle. In other cases, the wind ejecta is sufficiently fast that the light-curve is similar, regardless of viewing angle. Although more massive ejecta models are, on average, brighter, lower mass models can be brighter than higher mass models based on this viewing angle.

3. Uncertainties in the Microphysics and its Implementation

The outflow models depend on detailed general-relativistic, magneto-hydrodynamic models including neutrino transport (and the related neutrino microphysics). We do not have enough space to discuss all of this microphysics in detail. Instead we focus on the uncertainties in the nuclear physics and atomic physics with a brief discussion on each.

Unstable neutron-rich nuclei participating in the r-process have many unmeasured properties that consequently influence kilonova signals. Uncertainties in nuclear binding energies and reaction rates impact the flow of material to heavier mass regions as well as the subsequent flow back towards lighter mass regions in ejecta with low electron fraction. Properties that depend on excited nuclear states, including half-lives and branching ratios, additionally control the timescale over which energy is released as well as potentially stored. The creation of the heaviest elements in merger events may produce distinct observable signals, e.g. via the production of long-lived species like $^{254}$Cf or via unique electromagnetic signatures. Central to the reduction of large nuclear uncertainties that impact kilonovae are experimental studies undertaken at radioactive beam facilities around the world, see e.g.
A major uncertainty in the application of atomic physics to the modeling of kilonovae concerns the bound-bound (line absorption) contribution to the radiative opacity. While photons traveling through a plasma can be absorbed in a number of atomic processes, line absorption is the dominant mechanism in the dynamical ejecta due to the dense forest of lines associated with near-neutral lanthanide and actinide elements. Complete sets of lanthanide opacities have been recently produced for the purpose of kilonova modeling\textsuperscript{25–27} using different atomic physics methods, which is another potential source of uncertainty. In an effort to mitigate this issue, we have made available to the public our lanthanide opacities\textsuperscript{28} so that kilonova modelers can use a consistent set of opacities, thereby ruling out emission differences that could arise from atomic physics implementations. As a specific illustration of these concepts, we present in Figure 3 the opacity of Nd, calculated under the assumption of local thermodynamic equilibrium (LTE), at a temperature of $T = 0.5$ eV and mass density $\rho = 10^{-13}$ g/cm$^3$, using four different models (see\textsuperscript{26} for details). Note that the line features enhance the opacity by six to eight orders of magnitude compared to what would occur if only free-electron processes were included. The Planck mean opacity, $\kappa^P$, is also displayed for each model, indicating a maximum discrepancy of about 40% between the various models for this integrated quantity. Detailed light-curve and spectral comparisons were carried out with these four models.\textsuperscript{26} Only

![Fig. 3. The LTE line-binned opacity for neodymium at $T = 0.5$ eV and $\rho = 10^{-13}$ g/cm$^3$ using four different models described in: a) FR, b) FR-SCNR, c) FR-SCR, and d) SR. The Planck mean opacity, obtained via integration of the line-binned opacity, is also listed in each panel.](image)
modest differences were noted in the light-curves, i.e. differences of 10–20% in the peak of the light curve and a maximum shift of half a day in the time of peak luminosity. Similar, modest differences were observed in the spectra produced with these models, i.e. certain spectral features are shifted redward are blueward, but the overall spectral characteristics are similar.

From a more general perspective, we note that opacities are typically calculated from first principles, rather than measured. The line contribution depends on quantities such as transition energies, line strengths (quantum mechanical matrix elements) and atomic level populations. While it is beyond the scope of this article to provide a detailed discussion of these concepts, we mention that recent work\textsuperscript{29, 30} has been performed to compare calculated energies with benchmark values in the NIST Atomic Spectra Database\textsuperscript{31} for a number of important, low-lying levels of the lanthanides. Also, most simulations employ the LTE approximation when calculating the level populations. This assumption is expected to be valid up to about one week post merger. But, as with many astrophysical transients, the LTE conditions become less and less valid as the expansion continues and the density becomes so low that the thermalization timescale from collisions becomes long compared to evolutionary timescales of the ejecta. A recent study\textsuperscript{32} explores the possibility that non-thermal, $\beta$-decay electrons, produced from the radioactive decay of r-process nuclei, are primarily responsible for the heating and collisional ionization in the ejecta.

With the large number of atomic levels, the number of line features is enormous (in the 10s of millions). Implementing these opacities into a numerical calculation is intractable even from a computational memory standpoint. In addition, because of the high velocities in the ejecta, modeling radiation transport becomes even more difficult. For a single strong line, the Sobolev approximation can be used to determine the likelihood of a photon to both be absorbed as it expands out of a medium and its likelihood of escaping this line when re-emitted. Transient explosions exhibit homologous outflow conditions where the velocity of the ejecta is proportional to the radius. As a photon moves out through this ejecta, its energy spans a broad range with respect to the rest frame of the outflow. Even if the photon does not interact with a line feature at one radius, it may interact with this feature at a different radius.

This treatment includes the fact that the velocity of the ejecta spans a range such that, even if a photon does not interact with an absorption feature in the rest frame of the outflowing material at one position, it may, as it transports outward, interact with other material. Hence, a single line feature can affect a larger energy band of photons. On the other hand, if most of the re-emitted photons arise from a direct de-excitation of the absorption feature, the photons remain trapped in a line feature in material moving at a constant velocity. For the homologous outflow of astrophysical transients, the photon can escape more quickly, effectively reducing the opacity. Implementing the Sobolev approximation is already difficult, but in
Fig. 4. A number of prescriptions have been proposed to model the expansion opacities. This figure shows the total optical depth using three different prescriptions for different numbers of dominant line features as a function of the optical depth of the line feature. There is some convergence with large number of lines, showing why a simple averaged opacity can work nearly as well as a more complicated expansion procedure.

In a medium where a forest of lines exists, producing an accurate model for spectra becomes increasingly difficult. The deficiencies of the Pinto & Eastman expansion opacity were outlined in its inception and a number of alternative approaches have been developed (for a review, see). A number of methods have been developed to calculate the optical depth and these different models produce a range of results based on these prescriptions (Figure 4). For a simplified model with pure neodymium, Figure 5 compares the bolometric luminosity for different approaches: Sobolev, expansion opacity and a simple binned treatment. Compared to many of our other uncertainties, the opacity implementation errors seem low.
4. Alternative Energy Sources

Most calculations assume that the power source behind the kilonova emission arises solely from radioactive decay. But, just as with core-collapse supernovae, the light-curves and spectra can be powered by a variety of sources. For kilonovae, a number of additional power sources have been proposed to drive the observed emission: magnetar emission, shock heating in the ejecta, and fallback accretion. In this section, we will discuss each of these energy sources in turn, reviewing both their basic physics and expected features.

The black hole accretion disk engine paradigm argued that long-duration bursts are produced in stellar collapse, while short-duration GRBs are produced by the merger of compact remnants. Under this paradigm, scientists predicted that short-duration GRBs should be offset from their formation region due to the momentum imparted on compact binaries at birth.\textsuperscript{35, 36} Although other power sources do not differentiate between short- and long-duration bursts in this manner, magnetar engines for short-duration bursts have remained a competitive model for these bursts. The basic idea behind this engine is that the merged core of a NSM is expected to
be rapidly rotating. If strong, dipole magnetic fields are generated in this merged core, then the rotational energy in the core can power the GRB. Whether or not the magnetar drives the initial burst of gamma-rays, magnetars do provide a natural explanation for the roughly half of all short-duration bursts that exhibit a long-lived (“plateau”) phase in the X-ray that can last up to $10^5$ s.\textsuperscript{37} The magnetar emission, if reprocessed by the expanding ejecta, can power the optical and even infra-red emission,\textsuperscript{38, 39} and can dominate the observed emission, altering the inferred r-process yield in the merger.

In supernovae, shock heating is believed to dominate the heating in type II supernovae, shock breakout events, and many superluminous supernovae. Shock heating may also play a role in kilonova emission. Shocks between the jet and the wind ejecta have been invoked to explain additional emission at early times (first day) in the kilonova emission.\textsuperscript{40} At later times (beyond a day), material falls back onto the compact remnant, likely forming a disk and further outflows. The potential energy released in this fallback can be tapped to provide an additional power source. Depending upon the timing of the fallback, this accretion emission

![Fig. 6. Bolometric luminosity versus time for a 0.05 $M_\odot$ 1D spherical wind model (blue), a 0.001 $M_\odot$ 1D spherical model with a $\dot{M} = 0.003$ $M_\odot$/s central fallback luminosity source (dashed orange), and the EM counterpart of GW170817, AT 2017gfo (dashed black, with error bars).]
can play a major role in the emission after the first day (Figure 6). The accretion luminosity is proportional to the potential energy released:

\[ L_{\text{acc}} \propto GM_{\text{core}} \dot{M}_{\text{fallback}}/r_{\text{ISCO}} \]  

(1)

where \( G \) is the gravitational constant, \( M_{\text{core}} \) is the mass of the merged core, \( r_{\text{ISCO}} \) is its innermost stable circular orbit if the core is a black hole or the neutron star radius if it is a neutron star, and \( \dot{M}_{\text{fallback}} \) is the fallback accretion rate. Whether the merged core is a black hole or a neutron star, the fallback material is likely to have sufficient angular momentum to initially be centrifugally supported in an accretion disk, providing a mechanism by which a fraction of the potential energy released in the infall can be converted to both energy and mass ejection.

5. Conclusions

In addition to being the primary mechanism behind short-duration gamma-ray bursts, NSMs have the potential to be the dominant source of r-process elements in the universe. To validate this claim, scientists must be able to accurately infer the mass and abundances from observations of the ejecta-driven transient from these mergers (a.k.a. kilonovae). In this paper, we reviewed a broad set of uncertainties affecting the modeling of kilonova light-curves and the current status of our efforts to characterize and, hopefully constrain, them. Much more work must be done to infer accurate r-process masses from these observations.

In this short paper, we focused mostly on broad-band light-curves. Spectra are key to constraining the different models and a number of studies have focused on trying to find line features in the kilonova observations e.g.\(^{41}\). Gamma-ray decay lines, particularly in the remnants of the NSM ejecta, also have the potential to probe detailed yields.\(^{21,42}\) So despite the difficulties that face astronomers in understanding the yields from these mergers, the future is looking bright.

References


