

Tight multimessenger constraints on the neutron star equation of state from GW170817 and a forward model for kilonova light-curve synthesis

Nicholl, Matt; Margalit, Ben; Schmidt, Patricia; Smith, Graham P.; Ridley, Evan J.; Nuttall, James

DOI:

[10.1093/mnras/stab1523](https://doi.org/10.1093/mnras/stab1523)

License:

Creative Commons: Attribution (CC BY)

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Nicholl, M, Margalit, B, Schmidt, P, Smith, GP, Ridley, EJ & Nuttall, J 2021, 'Tight multimessenger constraints on the neutron star equation of state from GW170817 and a forward model for kilonova light-curve synthesis', *Monthly Notices of the Royal Astronomical Society*, vol. 505, no. 2, pp. 3016–3032.
<https://doi.org/10.1093/mnras/stab1523>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

This article has been accepted for publication in *Monthly Notices of the Royal Astronomical Society* © 2021 The Author(s). Published by Oxford University Press on behalf of the Royal Astronomical Society. All rights reserved.

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.



Tight multimessenger constraints on the neutron star equation of state from GW170817 and a forward model for kilonova light-curve synthesis

Matt Nicholl^{1,2}*, Ben Margalit^{3,†}, Patricia Schmidt^{1,2}, Graham P. Smith^{1,2}, Evan J. Ridley^{1,2} and James Nuttall¹

¹*School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK*

²*Institute for Gravitational Wave Astronomy, University of Birmingham, Birmingham B15 2TT, UK*

³*Astronomy Department and Theoretical Astrophysics Center, University of California, Berkeley, CA 94720, USA*

Accepted 2021 May 23. Received 2021 April 21; in original form 2021 February 4

ABSTRACT

We present a rapid analytic framework for predicting kilonova light curves following neutron star (NS) mergers, where the main input parameters are binary-based properties measurable by gravitational wave detectors (chirp mass and mass ratio, orbital inclination) and properties dependent on the nuclear equation of state (tidal deformability, maximum NS mass). This enables synthesis of a kilonova sample for any NS source population, or determination of the observing depth needed to detect a live kilonova given gravitational wave source parameters in low latency. We validate this code, implemented in the public MOSFIT package, by fitting it to GW170817. A Bayes factor analysis overwhelmingly ($B > 10^{10}$) favours the inclusion of an additional luminosity source in addition to lanthanide-poor dynamical ejecta during the first day. This is well fit by a shock-heated cocoon model, though differences in the ejecta structure, opacity or nuclear heating rate cannot be ruled out as alternatives. The emission thereafter is dominated by a lanthanide-rich viscous wind. We find the mass ratio of the binary is $q = 0.92 \pm 0.07$ (90 per cent credible interval). We place tight constraints on the maximum stable NS mass, $M_{\text{TOV}} = 2.17^{+0.08}_{-0.11} M_{\odot}$. For a uniform prior in tidal deformability, the radius of a $1.4 M_{\odot}$ NS is $R_{1.4} \sim 10.7$ km. Re-weighting with a prior based on equations of state that support our credible range in M_{TOV} , we derive a final measurement $R_{1.4} = 11.06^{+1.01}_{-0.98}$ km. Applying our code to the second gravitationally detected NS merger, GW190425, we estimate that an associated kilonova would have been fainter (by ~ 0.7 mag at 1 d post-merger) and declined faster than GW170817, underlining the importance of tuning follow-up strategies individually for each GW-detected NS merger.

Key words: gravitational waves – transients: neutron star mergers – methods: data analysis – stars: neutron.

1 INTRODUCTION

The first binary neutron star (NS) merger detected through its gravitational wave emission, GW170817 (Abbott et al. 2017a), was accompanied by both a short gamma-ray burst (GRB; Goldstein et al. 2017; Savchenko et al. 2017) and an optical counterpart (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). Thermal transients from these mergers had long been predicted in the form of a kilonova (Li & Paczyński 1998; Rosswog et al. 1999; Metzger et al. 2010): a non-relativistic outflow heated by the decays of heavy nuclei formed through rapid neutron captures (the so-called ‘r-process’; Lattimer & Schramm 1974; Eichler et al. 1989; Davies et al. 1994; Freiburghaus, Rosswog & Thielemann 1999). Extensive optical and infrared follow-up of GW170817 by many groups showed photometric and spectroscopic properties that were remarkably consistent with the expectations for an r-process kilonova (Andreoni et al. 2017; Arcavi et al. 2017; Chornock et al. 2017;

Cowperthwaite et al. 2017; Díaz et al. 2017; Drout et al. 2017; Evans et al. 2017; Hu et al. 2017; Kasliwal et al. 2017; McCully et al. 2017; Pian et al. 2017; Shappee et al. 2017; Smartt et al. 2017; Tanvir et al. 2017; Troja et al. 2017; Utsumi et al. 2017; Valenti et al. 2017; Nicholl et al. 2017a; Pozanenko et al. 2018).

Studying these sources is crucial for understanding cosmic chemical evolution. The total ejected mass powering the GW170817 kilonova was inferred to be $\sim 0.05 M_{\odot}$, with at least two components of different compositions and hence opacities (e.g. Kasen et al. 2017; Villar et al. 2017; Margutti & Chornock 2020): one containing light r-process elements (atomic mass number $A \lesssim 140$), and one containing the heavier lanthanides, which absorb optical photons in the forest of transitions in their open f -shells (Barnes & Kasen 2013; Kasen, Badnell & Barnes 2013). The large total mass and broad range of mass numbers covering all three r-process abundance peaks (Burbidge et al. 1957) suggested that NS mergers may be the dominant production site of heavy nuclei in the Universe.

Another central goal of multimessenger astronomy is to determine the NS equation of state (EoS), i.e. the relation between pressure and density. This provides unique tests of nuclear physics beyond the saturation density of nuclear matter (e.g. Lattimer & Prakash 2016). Because the EoS determines the NS radius and maximum stable

* E-mail: mnicholl@star.sr.bham.ac.uk

† NASA Einstein Fellow.

mass, measuring these quantities from astrophysical observation directly constrains the EoS. For example, a viable EoS must support the existence of the most massive known pulsars, with mass $\gtrsim 2 M_{\odot}$ (Demorest et al. 2010; Antoniadis et al. 2013; Cromartie et al. 2020). NS mergers provide an excellent laboratory for EoS measurements. The degree of tidal deformation imprinted in the GW signal (Flanagan & Hinderer 2008; Damour, Nagar & Villain 2012) and the strength of shocks at the contact interface (Bauswein, Goriely & Janka 2013b; Hotokezaka et al. 2013) both depend on the radius, typically parametrized as $R_{1.4}$, the radius of a $1.4 M_{\odot}$ NS. Similarly, the lifetime of the massive merger product until collapse to a black hole (BH) depends on the maximum stable NS mass, the Tolman–Oppenheimer–Volkoff mass M_{TOV} (Shapiro & Teukolsky 1986; Margalit & Metzger 2017).

So far, most of the EoS constraints for GW170817 have come from the GW data (e.g. Abbott et al. 2018; De et al. 2018; Raithel, Özel & Psaltis 2018). However, electromagnetic data have also provided important insight. Nicholl et al. (2017a) pointed out that the inferred combination of large mass and high velocity in the low-lanthanide ejecta was predicted only in the case of relatively small radii (Bauswein et al. 2013b; Shibata & Hotokezaka 2019). More quantitatively, Margalit & Metzger (2017) used the observed electromagnetic signatures of GW170817 to infer that the merged remnant must have collapsed to a BH within a short time-scale (~ 10 ms– ~ 1 s), pointing to $M_{\text{TOV}} \lesssim 2.17 M_{\odot}$. Bauswein et al. (2017) related this lifetime to the radius, finding $R_{1.4} > 10.68$ km.

A major caveat in these analyses is our ability to associate a measured ejecta component with a specific physical origin. NS mergers can eject mass dynamically during the merger, through tidal stripping or shock heating, and after the merger, via winds from an accretion disc or long-lived remnant (see the review by Metzger 2019). Each mechanism is expected to produce a different combination of ejecta mass, velocity and composition, and the relative importance of each mechanism depends on the properties of the system before merger (e.g. Bauswein et al. 2013b; Hotokezaka et al. 2013; Sekiguchi et al. 2016; Radice et al. 2018; Shibata & Hotokezaka 2019; Ciolfi & Kalinani 2020).

Most model fits to the electromagnetic data from GW170817 have been specified in terms of the ‘post-merger’ parameters, i.e. the individual masses, velocities and compositions/opacities of some number of ejecta components. This introduces subjectivity into the origin of each component, and neglects physically expected correlations between the properties of the different components. Specifying instead a forward model in terms of ‘pre-merger’ properties – binary masses and mass ratios, and the EoS – would allow one to trace explicitly the ejection mechanism of each component, and exclude models that fit the data well but require unphysical combinations of ejecta parameters. Moreover, analysing the data in terms of pre-merger properties provides a more direct link between the electromagnetic and GW signals.

Progress in combining light-curve models with GW inference was first made by Coughlin et al. (2019a), who fit the light curve of the GW170817 kilonova and related the implied ejecta masses to the predictions of numerical relativity simulations for different system mass ratios and combinations of $R_{1.4}$ and M_{TOV} . Including GW information in their priors, they measured $R_{1.4} = 11.3$ – 13.5 km. However, their model was unable to match the early blue part of the light curve, possibly because a sufficiently high mass of low-lanthanide ejecta was not predicted by merger simulations at the GW-inferred binary mass of GW170817. This may suggest the need to look at alternative luminosity sources for the early emission, such as cooling of matter shock-heated by the GRB jet (Kasliwal et al.

2017; Arcavi 2018; Piro & Kollmeier 2018). While our paper was in preparation, other efforts have built on these techniques. Dietrich et al. (2020) published an updated analysis, using a light-curve model that reasonably reproduced the early flux without shock cooling, and found $R_{1.4} = 11.75 \pm 0.86$ km. An even more recent analysis of GW170817 was published by Breschi et al. (2021), who found $R_{1.4} = 12.2 \pm 0.5$ km using a three-component kilonova model.

A better understanding of NS physics will come with larger samples of kilonovae, but finding these events has proven challenging. Despite subsequent GW detections of binaries hosting possible NSs, such as GW190425 (Abbott et al. 2020c) and GW190814 (Abbott et al. 2020d), and significant effort invested in counterpart searches (Gomez et al. 2019; Hosseinzadeh et al. 2019; Lundquist et al. 2019; Coughlin et al. 2019b; Ackley et al. 2020; Gompertz et al. 2020; Kasliwal et al. 2020; Paterson et al. 2020; Vieira et al. 2020; Andreoni et al. 2020b; Anand et al. 2021), at the end of the third LIGO–Virgo observing run (O3), GW170817 remains the only NS binary detected in both electromagnetic and gravitational waves (but see Graham et al. 2020, for a possible optical counterpart to a BH binary merger). Analysis of short GRB afterglows has yielded several kilonova candidates (Berger, Fong & Chornock 2013; Tanvir et al. 2013; Jin et al. 2016; Gompertz et al. 2018; Troja et al. 2018, 2019; Lamb et al. 2019; Fong et al. 2020) and constraining non-detections (O’Connor et al. 2021; Rastinejad et al. 2021), while optical surveys are also looking for kilonovae in their data streams, though without an unambiguous detection to date (McBrien et al. 2020; Andreoni et al. 2020a). However, a step change in our ability to find kilonovae is anticipated, due to planned improvements in the sensitivity of GW detector networks (Abbott et al. 2020b), and the deepest ever time-domain sky survey: the Vera Rubin Observatory Legacy Survey of Space and Time (LSST; LSST Science Collaboration et al. 2009).

In this paper, we provide a fast, analytic forward model for kilonova light curves that is completely specified by the binary configuration and EoS-dependent properties. Fitting such a model to observed kilonova data allows one to measure fundamental pre-merger properties even in the absence of a GW signal, or to use GW information when available as priors for multimessenger inference. Forward modelling the light curves of kilonovae can increase the efficiency of GW-triggered optical searches, by predicting the electromagnetic light curve directly from the GW signal and providing the required observing depth to detect (or rule out) a counterpart at any time since merger (see also Salafia et al. 2017; Barbieri et al. 2020; Stachie et al. 2021). In the context of all-sky surveys like LSST, such a model will allow for kilonova population synthesis (see also Kashyap, Raman & Ajith 2019; Mochkovitch et al. 2021) from a given NS source population, informing strategies for selecting candidate kilonovae from the data stream or allowing one to constrain the NS merger rate even given non-detections.

GW170817 serves as an ideal testing ground for such a model. By fitting the rich observed data set, we will show the power of using GW results as priors, constrain the physical origins of the different emission components, and present new measurements of the mass ratio of the progenitor system, the viewing angle, and the EoS-dependent quantities $R_{1.4}$ and M_{TOV} .

The structure of this paper is as follows. In Section 2, we describe the model setup and assumptions. Section 3 shows its application to GW170817 and our main results. We briefly demonstrate the utility of our code for aiding future kilonova searches, using the example of GW190425, in Section 4, before concluding in Section 5. A set of kilonova model light curves produced using our code is available for download (see the Data Availability section).

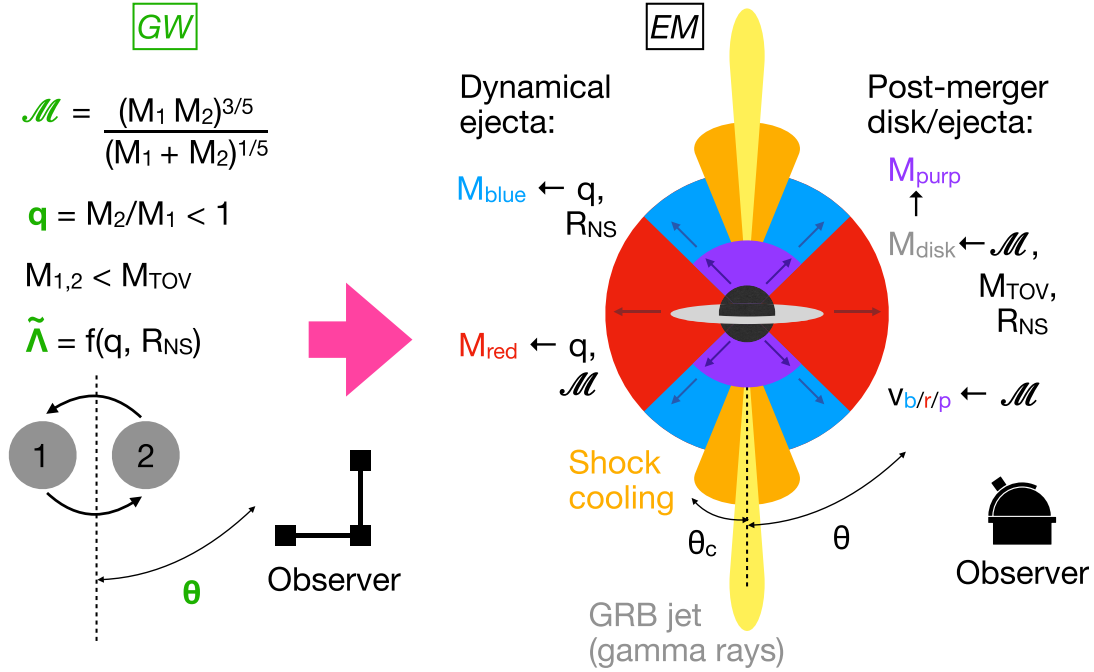


Figure 1. Schematic showing model geometry and relations between parameters. A full description of these parameters is given in Section 2. Left-hand panel: pre-merger properties of an NS binary. Bold, green font is used to indicate properties that can be measured from GW waveforms. Right-hand panel: ejecta properties probed by electromagnetic observations. Left arrows are used to mean ‘depends on’, in order to highlight the most important parameter sensitivities. The merged remnant can be a long-lived or short-lived NS or a BH, depending on the ratio $(M_1 + M_2)/M_{\text{TOV}}$. The blue ejecta are from shocks at the collisional interface and have an opacity $\kappa_{\text{blue}} \sim 0.5 \text{ cm}^2 \text{ g}^{-1}$, the red ejecta are tidally stripped in the orbital plane and have $\kappa_{\text{red}} \sim 10 \text{ cm}^2 \text{ g}^{-1}$, and the purple ejecta are winds from an accretion disc and have an intermediate opacity that depends on the lifetime of the remnant. Low opacity ejecta may not be visible for large viewing angle θ . The shock cooling emission arises from a cocoon of width θ_c heated at its interface with the GRB jet.

2 DESCRIPTION OF MODEL

We have built our model using the Modular Open Source Fitter for Transients (MOSFIT; Guillochon et al. 2018). Modules to calculate a radioactive kilonova light curve for a given set of ejecta masses, velocities and opacities have already been implemented by Cowperthwaite et al. (2017) and Villar et al. (2017). Therefore our main task here is to define modules that provide these masses, velocities and opacities for a given binary configuration. We also provide new modules to take into account the effects of observer viewing angle and shock cooling. We have made these new modules and the parameter files to implement them publicly available via GitHub,¹ as a new model called ‘BNS’. In this section, we will describe the main features. Fig. 1 provides a visual overview that summarizes the geometry and the relations between key parameters. The connection between our modelling framework and the GW observables is discussed in detail in Section 2.6.

2.1 Dynamical ejecta

During the merger process, mass is ejected dynamically both by tidal forces in the orbital plane (Rosswog et al. 1999; Sekiguchi et al. 2015) and by shocks at the contact interface between the two NSs (Bauswein et al. 2013b; Hotokezaka et al. 2013). To estimate the mass of this material, we use a fit by Dietrich & Ujevic (2017) to the ejecta masses of 172 numerical relativity simulations. These simulations, from Hotokezaka et al. (2013), Bauswein et al. (2013b),

Dietrich et al. (2015), Dietrich et al. (2017), Lehner et al. (2016), and Sekiguchi et al. (2016), cover a range of binary masses, though all neglect the effect of NS spin. The catalogue includes both grid-based and smoothed particle hydrodynamics simulations, and simulations with and without neutrino treatment. In total they cover 21 simplified or tabulated EoSs.

Their function takes the form

$$\frac{M_{\text{dyn}}}{10^{-3} M_{\odot}} = \left[a \left(\frac{M_2}{M_1} \right)^{1/3} \left(\frac{1 - 2C_1}{C_1} \right) + b \left(\frac{M_2}{M_1} \right)^n + c \left(1 - M_1/M_1^* \right) \right] M_1^* + (1 \leftrightarrow 2) + d, \quad (1)$$

where M_i is the mass of the i th NS and

$$C_i \equiv GM_i/R_i c^2 \quad (2)$$

is its compactness, given a radius R_i . The term $1 \leftrightarrow 2$ represents exchanging the subscripts, and the best-fitting values of the free parameters a, b, c, d, n can be found in their study. Quantities denoted M_i^* are the baryonic masses (differing from the gravitational mass by the binding energy), which we calculate for a given M_i (in solar masses) following Gao et al. (2020): $M_i^* = M_i + 0.08 M_i^2$. The simulations cover the ranges $1 < M_i/M_{\odot} < 2$ and $0.1 < C_i < 0.23$. The typical fractional uncertainty on M_{dyn} using this parametrization was found by Dietrich & Ujevic (2017) to be ≈ 72 per cent. Similar expressions were provided for the equatorial and polar ejecta velocities, with smaller uncertainties of only 13–33 per cent.

Inspection of equation (1) shows that the dynamical ejecta mass depends on terms of the form M_i/M_j . We adopt the convention that M_1 is the heavier NS (for consistency with GW analyses, e.g. Abbott

¹<https://github.com/guillochon/MOSFiT>.

et al. 2017a, 2018), such that the mass ratio of the system is $q = M_2/M_1 < 1$. The other important sensitivity in equation (1) is to the compactness C_i , an EoS-dependent quantity. This will be discussed further in Section 2.6.

The relationship between binary and ejecta masses for NS mergers was recently revisited by Krüger & Foucart (2020) and Nedora et al. (2020), who used a different set of numerical data but found broadly consistent fits. Coughlin et al. (2019a) provided a fitting formula in terms of $\log(M_{\text{dyn}})$, which performed similarly well. However, we find this logarithmic form can lead to unrealistically large ejecta masses in some regions of parameter space, which is not optimal for population synthesis.

To better capture the multiple physical processes contributing to the dynamical ejecta, we divide it into two components. The tidal ejecta, concentrated in the equatorial plane, are expected to contain a very low electron fraction (the ratio of protons to total nucleons), $Y_e < 0.25$, allowing the complete r-process synthesis of very heavy nuclei. These ejecta have been termed ‘red’ due to the high opacity of the resulting lanthanides at optical wavelengths (Barnes & Kasen 2013). The polar material has a higher $Y_e > 0.25$, due to the importance of shocks and neutrino heating at higher latitudes (e.g. Bauswein et al. 2013b; Sekiguchi et al. 2016; Metzger 2019), and is correspondingly ‘blue’ due to an incomplete r-process only up to mass number $A \sim 140$. Radice et al. (2018) find that the ratio of red to blue ejecta is largely insensitive to the total system mass or the EoS, and depends primarily on the mass ratio of the system.

To estimate the relative mass in the red and blue dynamical ejecta, we use simulations by Sekiguchi et al. (2016). They provide Y_e distributions in the dynamical ejecta as a function of q for two equations of state, one stiff (large radius) and one soft (compact). We integrate these distributions (in their fig. 5) to determine $f_{\text{red}}(q) \equiv M(Y_e < 0.25)/M(\text{total})$ and $f_{\text{blue}} \equiv 1 - f_{\text{red}}$. We find that f_{red} is insensitive to the EoS but is a steep function of q , ranging from 20 per cent for an equal-mass merger ($q = 1$) to 76 per cent for $q = 0.86$. For $q \lesssim 0.8$, $f_{\text{red}} \approx 1$, because the less massive NS is tidally disrupted prior to contact, preventing strong shocks (e.g. Bauswein et al. 2013b; Hotokezaka et al. 2013; Lehner et al. 2016; Dietrich & Ujevic 2017). For $q > 0.8$, we interpolate f_{red} between the available simulations using a smooth polynomial fit (shown in Fig. A1). The predicted ejecta masses in each component, for a range of NS binaries, are shown in Fig. 2. Following Radice et al. (2018), we assume a fixed grey opacity for each component: $\kappa_{\text{blue}} = 0.5$ and $\kappa_{\text{red}} = 10 \text{ cm}^2 \text{ g}^{-1}$ (Villar et al. 2017; Metzger 2019; Tanaka et al. 2020).

2.2 Post-merger ejecta

For most systems, additional mass will be ejected after the merger, driven by neutrino-heated winds from a viscous accretion disc around the remnant or from the surface of a massive remnant NS (Metzger & Fernández 2014). Simulations show that in many cases the mass of post-merger ejecta will exceed that of the dynamical ejecta (Ciolfi et al. 2017; Siegel & Metzger 2017, 2018; Radice et al. 2018; Fernández et al. 2019).

The properties of the wind ejecta depend on the lifetime of the merger remnant, which may be a stable NS, a supramassive (hypermassive) NS temporarily supported by (differential) rotation, or may promptly collapse to a BH. To first order, longer lived remnants produce a larger disc mass with a higher Y_e (Metzger & Fernández 2014; Lippuner et al. 2017).

We use the analytic expression for the disc mass from Coughlin et al. (2019a), who determine the values of their coefficients a , b , c ,

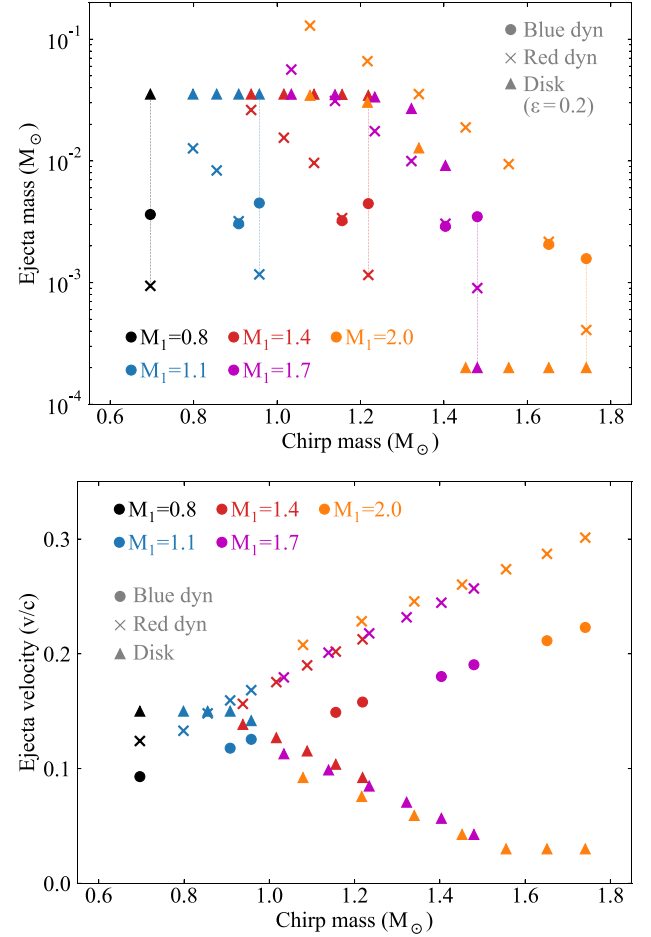


Figure 2. Top panel: predicted ejecta mass in each component (blue and red dynamical, and purple disc ejecta) using relations from Dietrich & Ujevic (2017) and Coughlin et al. (2019a), assuming $R_{1.4} = 11 \text{ km}$, $M_{\text{TOV}} = 2.1 M_\odot$, and that 20 per cent of the remnant disc is ejected. Each colour corresponds to a fixed primary mass, M_1 , with varying mass ratio, $0.4 \leq q \leq 1$. Vertical lines indicate $M_2 = M_1$, with lower mass secondaries (less equal binaries) to the left. Bottom panel: predictions for ejecta velocities. Disc ejecta velocities are estimated using results from Metzger & Fernández (2014).

d by fitting to simulations from Radice et al. (2018):

$$\log_{10}(M_{\text{disk}}) = \max \left(-3, a \left(1 + b \tanh \left(\frac{c - M_{\text{tot}}/M_{\text{thr}}}{d} \right) \right) \right). \quad (3)$$

Here $M_{\text{tot}} = M_1 + M_2$ is the total binary mass, and M_{thr} is the threshold mass for prompt collapse to a BH, parametrized as

$$M_{\text{thr}} = \left(2.38 - 3.606 \frac{M_{\text{TOV}}}{R_{1.4}} \right) M_{\text{TOV}}, \quad (4)$$

i.e. proportional to the maximum stable mass of a non-rotating NS and the compactness (Bauswein, Baumgarte & Janka 2013a). An alternative fit by Radice et al. (2018) was parametrized in terms of the tidal deformability rather than M_{TOV} . Following Coughlin et al. (2019a), we prefer to retain the explicit dependence on M_{TOV} so that we can try to constrain it observationally. Recent fits by Nedora et al. (2020) predict disc masses broadly consistent with the results of Coughlin et al. (2019a). The parametrization above is based on simulations of systems with close to equal-mass ratio, q

≈ 1 (Bauswein et al. 2013b; Radice et al. 2018). We note however that Kiuchi et al. (2019) have shown that the mass ratio can play an important role and enhance the disc mass for small q .

A fraction $\epsilon_{\text{disc}} \sim 0.1\text{--}0.5$ of the disc mass is expected to be ejected by viscously-driven winds, with $\epsilon_{\text{disc}} \approx 0.2$ likely typical (Siegel & Metzger 2017; Radice et al. 2018). This fraction can be fixed in our model for population synthesis, or left free to infer its value from fits to observed KNe (although in principle, ϵ_{disc} may change as a function of disc mass; De & Siegel 2020). Additional post-merger ejecta may arise through neutrino- or magnetically-driven winds (Metzger, Thompson & Quataert 2018; Radice et al. 2018; Ciolfi & Kalinani 2020), which we will discuss in Section 2.5.

To determine the velocity and composition of the disc ejecta, we first estimate the remnant properties as a function of M_{tot} , M_{TOV} , and M_{thr} using table 3 of Metzger (2019). For $M_{\text{tot}} < M_{\text{TOV}}$, the remnant is indefinitely stable and we assume an ejecta velocity of $\approx 0.1c$ based on the simulations of Metzger & Fernández (2014). For $M_{\text{tot}} > M_{\text{thr}}$ (prompt collapse), the same simulations predict a velocity of $\approx 0.03c$. We linearly interpolate between these velocities for intermediate M_{tot} . Fig. 2 shows the predicted disc ejecta masses and velocities alongside the dynamical ejecta from Section 2.1.

To estimate the opacity, we use the simulations from Lippuner et al. (2017), who present Y_e distributions in their fig. 2 for discs with various remnant lifetimes, ranging from indefinitely stable to prompt collapse. We integrate these distributions to determine the mass-weighted Y_e . We then convert this mean Y_e into an effective grey opacity, κ_{purple} , using fits to the Y_e – κ relations from Tanaka et al. (2020) (shown in Fig. A2). The inferred range of $0.25 < \langle Y_e \rangle < 0.38$ corresponds to an opacity range of $5.5 \gtrsim \kappa_{\text{purple}}/(\text{cm}^2 \text{g}^{-1}) \gtrsim 1.5$ for the disc ejecta. The ‘purple’ subscript refers to the fact that this lies intermediate between the red and blue dynamical ejecta.

2.3 Luminosity from r-process decay

A semi-analytic model that captures the essential light-curve physics of kilonovae was developed by Metzger (2019). This model approximates the heating rate at time t , averaged over all r-process radioactive decay chains, as $\dot{Q}(t) \propto \epsilon_{\text{th}} M_r t^{-1.3}$ (Korobkin et al. 2012), where M_r is the r-process mass and ϵ_{th} is a time-dependent thermalization efficiency (Barnes et al. 2016). A module to calculate $\dot{Q}(t)$ in this way was implemented in MOSFIT by Cowperthwaite et al. (2017) and Villar et al. (2017).

To determine the output luminosity, $L(t)$, the ejecta are assumed to have a constant grey opacity and expand homologously with scale velocity v_{ej} , allowing for an analytic radiative diffusion solution to the first law of thermodynamics (Arnett 1982). The temperature of the photosphere is initially calculated using the Stefan–Boltzmann law, with radius $R = v_{\text{ej}}t$, until such time as the cooling ejecta start to recombine. At this point, the photosphere recedes into the ejecta while maintaining the recombination temperature, T_{rec} . The MOSFIT modules to solve for the radiative diffusion and the photospheric temperature and radius are described in Nicholl, Guillochon & Berger (2017b).

Cowperthwaite et al. (2017) and Villar et al. (2017) generalized this model to include the sums of multiple ejecta components: the luminosity scale of each component is proportional to its mass, with a diffusion time-scale determined also by its velocity and opacity. They allowed T_{rec} to vary independently for each of two to three ejecta components, finding best-fitting values in the range $1000 \lesssim T_{\text{rec}} \lesssim 4000 \text{ K}$. To reduce the number of free parameters (and restrict to binary-based quantities of interest) we assume a fixed $T_{\text{rec}} = 2500 \text{ K}$, based on the theoretically predicted recombination

temperature of ionized r-process ejecta (Barnes & Kasen 2013). This falls in the middle of the plausible range found by Villar et al. (2017). We tested the effect of leaving this parameter free, and found a best-fitting value $T_{\text{rec, free}} \approx 2600 \pm 200 \text{ K}$, consistent with theory. For completeness, we also tested the effect of using three independent temperatures in our base model; in this case, they converged to similar values to Villar et al. (2017), and other parameters remained consistent with the fixed temperature results to better than 1σ .

2.4 Geometry

Most analytic light-curve models assume spherical symmetry. However, in kilonovae, the separate ejecta components occupy distinct spatial regions, due to their different ejection mechanisms (as illustrated in Fig. 1). The low-medium Y_e (i.e. blue and purple) ejecta are expected to escape mostly orthogonal to the orbital plane, where shock heating and weak interactions lower the free neutron flux (e.g. Wanajo et al. 2014; Goriely et al. 2015; Foucart et al. 2016). In contrast, the neutron-rich red ejecta escape parallel to the orbital plane since they form tidally. Simulations show that this geometry can be well approximated as a sphere with conical sections of half-opening angle $\theta_{\text{open}} \sim 45^\circ$ around the poles. The blue/purple ejecta reside within the conical sections and the red ejecta outside.

This results in a viewing angle dependence in the observed luminosity, colour, and time-scale of the kilonova (e.g. Kasen, Fernández & Metzger 2015; Kawaguchi, Shibata & Tanaka 2018; Wollaeger et al. 2018; Bulla 2019; Korobkin et al. 2020), which can be understood intuitively: from viewing angles close to the equator, emission from the low-lanthanide polar ejecta is obscured by the curtain of lanthanide-rich ejecta at lower latitudes, meaning that only a slow-evolving red kilonova is observed. Radiative transfer simulations show that this can reduce the peak luminosity by $\approx 1\text{--}2 \text{ mag}$ (Bulla 2019). Neglecting this effect is problematic both for population synthesis and for fitting observed kilonovae, since an offset in the luminosity introduces a systematic error in measuring the ejected mass. Another subtle viewing angle effect is associated with the polar cavity carved out by a GRB jet (whether failed or successful) that can accompany the merger, as recently studied by Nativi et al. (2021) and Klion et al. (2021). This can expose interior regions of the ejecta which are hotter or lower opacity, resulting in brighter or bluer emission.

Villar et al. (2017) approximated the viewing angle effect in MOSFIT by multiplying the luminosity of the equatorial ejecta by $\cos\theta$ and the polar ejecta by $1 - \cos\theta$, with the viewing angle θ a free parameter. Here we implement a new module that treats this more exactly, using the formalism of Darbha & Kasen (2020). This takes a sphere with conical polar caps of half-opening angle θ_{open} , and scales the luminosity of each component using the projected area of the caps (blue/purple ejecta) or remaining sphere (red ejecta) as a function of θ and θ_{open} . We have verified that the change in luminosity with θ is consistent with the radiative transfer results of Bulla (2019) for the same geometry. We note that this treatment does not account for the jet-interaction viewing angle effect discussed by Klion et al. (2021), Nativi et al. (2021), as this would depend on uncertain properties of the jet. For our fiducial model, we assume $\theta_{\text{open}} = 45^\circ$ and vary the observer viewing angle θ .

2.5 Extensions to the base model

We include in our model two optional final ingredients that can contribute additional luminosity. Although they are tied less explicitly in our formalism to the initial binary parameters, neglecting these

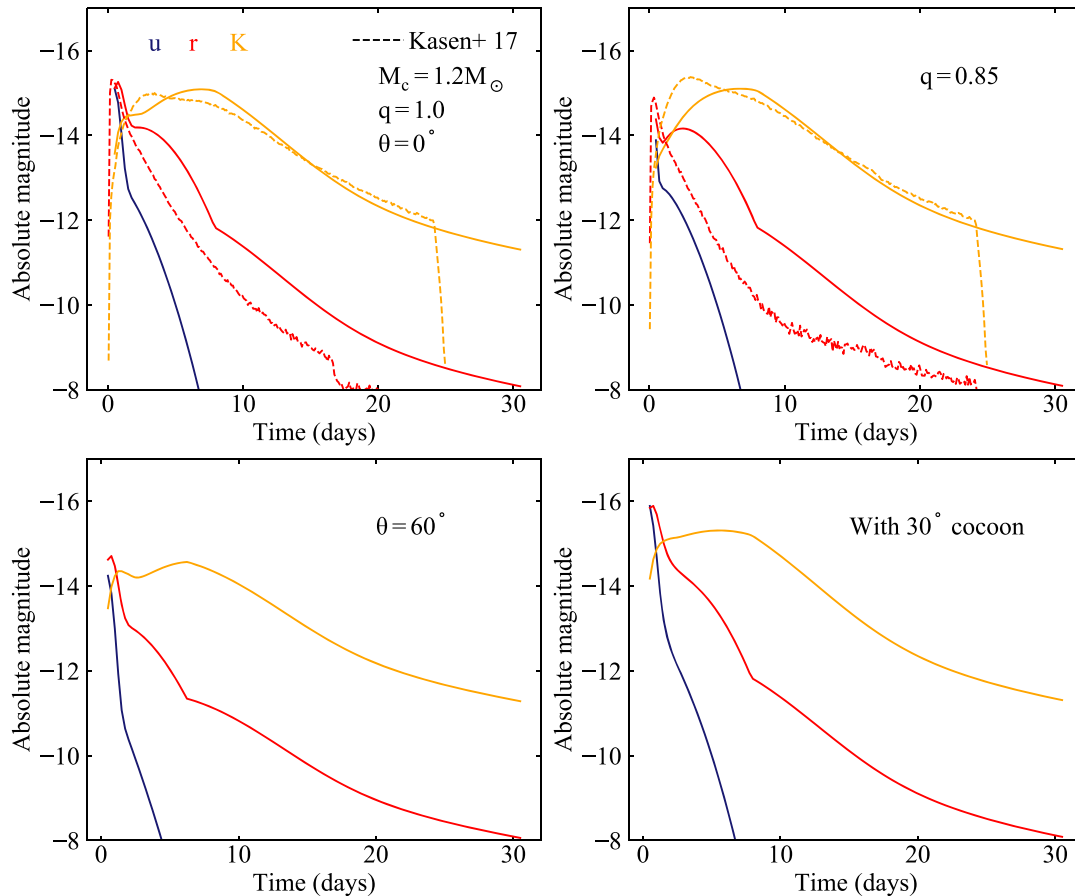


Figure 3. Example light curves from our model in the near-UV (u), optical (r) and near-infrared (K) bands, for a merger with chirp mass $\mathcal{M} = 1.2 M_{\odot}$. Top left-hand panel: fiducial, equal-mass merger viewed along pole. Top right-hand panel: Merger with same chirp mass but asymmetric mass ratio, $q = 0.85$. Lower left-hand panel: Fiducial model viewed at 60° off-axis. Lower right-hand panel: Including shock-heated cocoon with opening angle 30° . The top two panels show light curves calculated with the models of Kasen et al. (2017) for the closest available ejecta masses (assuming lanthanide fractions of 10^{-9} , 10^{-2} , and 10^{-1} for the blue, purple, and red components, respectively). These models have not been tuned to match our simulations and are intended only to guide the eye. The viewing angle effect is very roughly approximated by adding -0.5 mag to the r band and $+0.5$ mag to the K band.

contributions may lead to an underestimate in the model luminosity or to a bias in the inference of other parameters from kilonova data. These components can easily be turned off at run-time, allowing the user to recover the fully predictive baseline model described in the previous sections.

The first is an enhancement in the blue ejecta due to magnetically-driven winds (Metzger et al. 2018; Cioffi & Kalinani 2020). This mechanism is possible only if the remnant avoids prompt collapse, hence we apply this effect only when the merger mass is below M_{thr} (equation 4). We adopt a simple parametrization, inspired by Coughlin et al. (2019a): $M_{\text{blue, tot}} = M_{\text{blue, dyn}}/\alpha$, where $0 < \alpha < 1$. Therefore, fixing $\alpha = 1$ turns off this source of ejecta.

The second effect is shock-heating of the ejecta by a GRB jet. GW170817 was accompanied by a short GRB (as expected in NS mergers; Eichler et al. 1989; Narayan, Paczynski & Piran 1992; Berger 2014), and some studies interpreted the early blue luminosity of the kilonova as arising from cooling of this shocked ‘cocoon’ (Kasliwal et al. 2017; Arcavi 2018). We include this effect using the formulas from Piro & Kollmeier (2018), specifically their equations 38 and 39. The breakout radius is approximated as the time delay between the GW and GRB signals (1.7 s for GW170817; Abbott et al. 2017b) multiplied by the speed of light (Piro & Kollmeier 2018). Following Nakar & Piran (2017), we assume that the shock deposits

constant energy per decade of velocity in the ejecta. The isotropic-equivalent luminosity is proportional to the mass of shocked ejecta, which is a fraction $\theta_c^2/2$ of the total ejecta. The shock can therefore be suppressed by setting the cocoon half-opening angle θ_c equal to zero (or $\cos \theta_c = 1$). We assume that the shocked mass is a fraction of the total dynamical ejecta, as it is unclear whether post-merger ejecta will precede the launching of the GRB jet. The output light curves are only weakly sensitive to this assumption, partially due to degeneracy between M_{ej} and θ_c . Because the cocoon shock occurs primarily in the polar material, we assume this matter has the same opacity as the blue ejecta component. The luminosity of the shock is inversely proportional to this opacity, so we note that a lanthanide-rich polar component could significantly reduce the importance of the shock emission.

We show examples of light curves calculated for different mass ratios, inclinations, and cocoon opening angles in Fig. 3.

2.6 Free parameters and relation to GW observables

After implementing the modules described above, our model depends on the masses of the constituent NSs, their compactness, the maximum stable mass M_{TOV} , and the viewing angle θ . The fraction of the disc mass ejected in the purple component can be varied or fixed

Table 1. Parameters in the MOSFIT BNS model and application to GW170817.

Parameter	Prior distribution		Posterior distribution for fit to GW170817			
	Default prior	GW170817 prior	Base model	Surface ejecta	Shocked cocoon	Agnostic
\mathcal{M}^a (M_\odot)	[0.7, 2.0]	1.188 ± 0.002	1.188 ± 0.002	1.188 ± 0.002	1.188 ± 0.002	1.188 ± 0.002
q^b	[0.4, 1.0]	[0.7, 1.0]	0.99 ± 0.01	$0.88^{+0.04}_{-0.02}$	$0.95^{+0.04}_{-0.07}$	0.92 ± 0.05
$\cos \theta^c$	[0.0, 1.0]	[0.82, 0.97]	0.89 ± 0.04	$0.88^{+0.04}_{-0.02}$	$0.85^{+0.03}_{-0.02}$	$0.84^{+0.03}_{-0.02}$
M_{TOV}^d (M_\odot)	[2.0, 2.5]	[2.0, 2.5]	2.32 ± 0.04	2.26 ± 0.03	$2.14^{+0.06}_{-0.07}$	$2.17^{+0.05}_{-0.06}$
Λ_s^e	[10, 2000]	[10, 2000]*	$\tilde{\Lambda} = 122^{+15}_{-13}$	$\tilde{\Lambda} = 116^{+24}_{-12}$	$\tilde{\Lambda} = 163^{+57}_{-44}$	$\tilde{\Lambda} = 231^{+92}_{-74}$
ϵ_{disc}^f	[0.05, 0.5]	[0.05, 0.5]	0.14 ± 0.01	0.13 ± 0.01	0.12 ± 0.01	0.12 ± 0.01
α^g	[0.1, 1.0]	[0.1, 1.0]	—	$0.37^{+0.17}_{-0.15}$	—	$0.63^{+0.18}_{-0.19}$
$\cos \theta_c^h$	[0.707, 1.0]	[0.707, 1.0]	—	—	$0.89^{+0.04}_{-0.06}$	$0.91^{+0.04}_{-0.07}$
t_0^i (d)	[−1.0, 0.0]	(−0.462)	—	—	—	—
$\log N_{\text{H}}^j$	[19.0, 23.0]	[19.0, 20.3]	$19.35^{+0.31}_{-0.22}$	$19.55^{+0.41}_{-0.34}$	$19.65^{+0.39}_{-0.40}$	$19.67^{+0.39}_{-0.41}$
$\log \sigma^k$	—	[−3.0, 2.0]	-0.34 ± 0.02	-0.39 ± 0.02	-0.44 ± 0.02	-0.44 ± 0.02
Bayesian model evidence (ln Z):			37.1	56.9	80.3	81.1

Notes. Some parameters are relevant only when fitting to data. Priors in square brackets are flat over the specified ranges, otherwise Gaussian. Posterior values for tidal deformability are given in terms of $\tilde{\Lambda}$ to facilitate comparison to GW analyses. The ‘Surface ejecta’ model includes additional blue ejecta from a long-lived remnant; the ‘Shocked cocoon’ model includes emission from a GRB-heated cocoon; and the ‘Agnostic’ model includes both. The stated best-fitting values and uncertainties correspond to the medians and 16th/84th percentiles of the marginalized posteriors.

^aChirp mass. ^bMass ratio. ^cObserver viewing angle. ^dMaximum stable NS mass. ^eSymmetric tidal deformability (specified instead as $R_{1.4}$ in generative mode). ^fFraction of disc ejected. ^gEnhancement ($1/\alpha$) of blue ejecta by surface winds. ^hOpening angle of shocked cocoon. ⁱTime of explosion (rest frame). ^jHydrogen column density in host galaxy (proportional to extinction). ^kWhite noise in likelihood function.

*Additional constraint that derived parameter $\tilde{\Lambda} < 800$; see Section 2.6.

to a fiducial value. Two further parameters allow one to increase the early blue flux through magnetic winds and/or shock cooling. Finally, we include line-of-sight extinction, and a white noise parameter σ in our likelihood function when fitting to data (see Nicholl et al. (see Villar et al. 2017; Nicholl et al. 2017b; Guillochon et al. 2018))

There is a choice to be made about how best to express some of these parameter dependencies. We opt to use the parametrization that most closely matches the quantities constrained by GW observations. This allows the user to quickly predict the electromagnetic light curve for a given GW signal, or to use GW information to inform their priors when fitting to an observed multimessenger source. We therefore parametrize the masses of the system in terms of the ‘chirp’ mass $\mathcal{M} = (M_1 M_2)^{3/5} (M_1 + M_2)^{-1/5}$, the quantity most accurately measured from the GW signal, and $q = M_2/M_1 \leq 1$, to which this signal is also somewhat sensitive.

The dependence on the NS compactness C is more complicated. For population synthesis, it is simplest to specify a radius R , equivalent to implicitly choosing an EoS (since all NSs of comparable mass should have approximately the same radius). However, when fitting to observed data (i.e. when we want to measure, rather than impose, a NS radius), it is more useful to express C in terms of the NS tidal deformability, which measures the responsiveness of the NS to an external tidal field: $\Lambda_i = (2/3)k_2 C_i^{-5}$, where $k_2(M) \approx 0.05\text{--}0.15$ is the quadrupolar tidal Love number (Hinderer 2008; Postnikov, Prakash & Lattimer 2010; Damour et al. 2012).

The reason is that the observable GW phase is to leading order determined by a specific combination of Λ_1 and Λ_2 (Flanagan & Hinderer 2008; Favata 2014; Wade et al. 2014),

$$\tilde{\Lambda} = \frac{16}{13} \frac{(m_1 + 12M_2)M_1^4 \Lambda_1 + (M_2 + 12M_1)M_2^4 \Lambda_2}{(M_1 + M_2)^5}, \quad (5)$$

called the binary or effective tidal deformability, which can therefore be constrained with GW data. The relation between Λ and C allows us to include GW information in our priors to better measure C and hence R . A similar approach was adopted by Coughlin et al. (2019a),

though they used a different relation proposed by De et al. (2018) to relate $\tilde{\Lambda}$ to R without calculating C explicitly.

For consistency with GW analyses (Abbott et al. 2018) and to reduce systematic error, we instead sample Λ in terms of the so-called ‘symmetric’ tidal deformability:

$$\Lambda_s \equiv (\Lambda_1 + \Lambda_2)/2. \quad (6)$$

Using ‘quasi-universal relations’ or QURs (largely independent of the EoS, and hence of complicating factors such as k_2), the antisymmetric tidal deformability $\Lambda_a \equiv (\Lambda_2 - \Lambda_1)/2$ can be derived from Λ_s to ≈ 5 per cent precision (Yagi & Yunes 2016), allowing one to reconstruct Λ_1 , Λ_2 , and $\tilde{\Lambda}$ directly from Λ_s . One may then impose the constraint that $\tilde{\Lambda}$ be consistent with the GW results, using the CONSTRAINTS class in MOSFIT (see Guillochon et al. 2018). The compactness of each NS is then derived from Λ_i using another QUR:

$$C_i = 0.360 - 0.0355 \ln(\Lambda_i) + 0.000705 \ln(\Lambda_i)^2, \quad (7)$$

This relation is accurate to ≈ 6.5 per cent across a variety of equations of state (Yagi & Yunes 2017). These C_i are then used in equation (1) to estimate the dynamical ejecta mass. Following a fit to observed data, the NS radii can be derived from the posterior distribution of Λ_s following the same procedure and using the constituent NS masses in equation 2.

Our final parameter set therefore consists of 5–11 parameters, which differ slightly between the generative and fitting versions of the model. We list all of these parameters and suggest priors in Table 1. For comparison, the parametrization of the original MOSFIT kilonova model by Villar et al. (2017) also had 11 free parameters.

3 APPLICATION TO GW170817

As with any model of NS mergers, GW170817 provides an ideal testing ground for the formalism described above. In this section, we fit the observed optical data to demonstrate (i) how using a binary (rather than ejecta) based parameter set provides additional insight

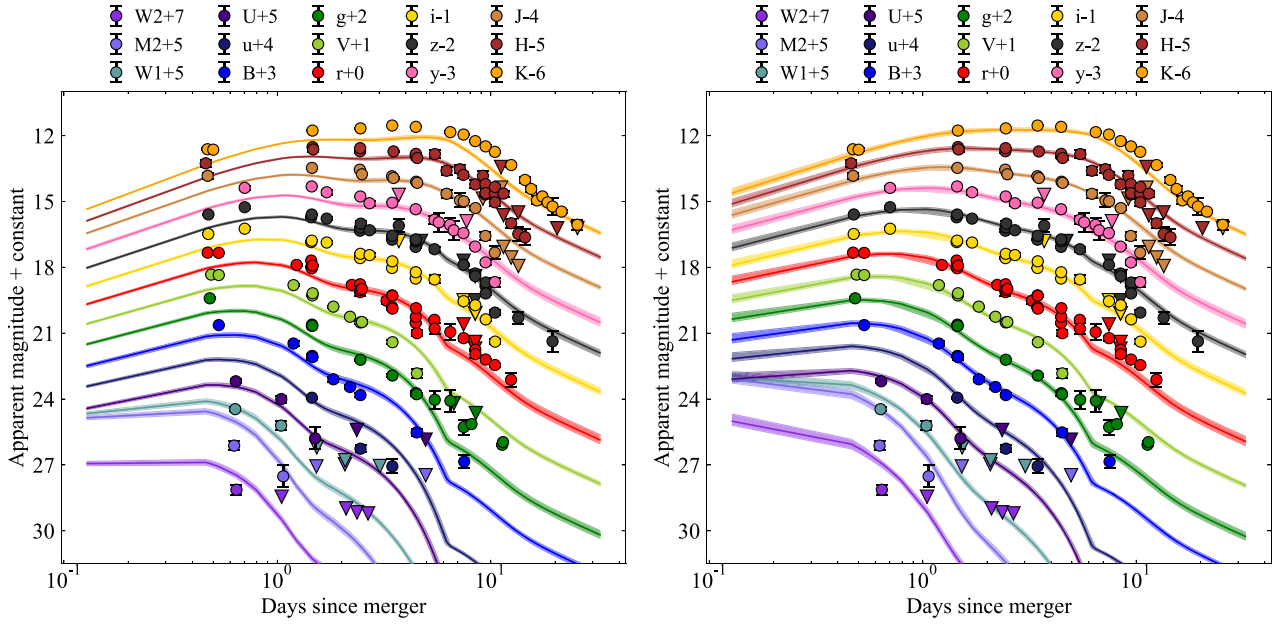


Figure 4. Left-hand panel: best-fitting kilonova model for GW170817 without surface ejecta enhancement or GRB shock cooling. Right-hand panel: best-fitting model with these effects included. Both models provide a good match to the data beyond $t \gtrsim 1.5$ d post-merger, but the fit without shock cooling is too faint by ≈ 0.7 mag (almost a factor of 2) during the first day. The model on the right is preferred to the base model with a Bayes factor $B > 10^{19}$. Light-curve fits for the surface ejecta and shock-only models are shown in the Appendix.

into the physical origin of each luminosity component; (ii) the utility of this model in probing the NS EoS with kilonova data; and (iii) the extent to which the use (or not) of GW information in the model priors affects our posteriors.

We use data from the Open Kilonova Catalog (Guillochon et al. 2017), compiled by Villar et al. (2017). Given the high level of overlap between the observations from different groups, we fit the following subset, chosen to cover the full range of bands with a comparable (approximately nightly) density of sampling in each: *ugrizyH* from Cowperthwaite et al. (2017), *BVgrJH* from Drout et al. (2017), *rizy* from Smartt et al. (2017), *rzK* from Tanvir et al. (2017), *JHK* from Kasliwal et al. (2017), *BV* from Troja et al. (2017), and the *U* and *UV* bands from Evans et al. (2017).

Our priors are given in Table 1. The time of merger is fixed by the GW detection. The priors on \mathcal{M} and q are from Abbott et al. (2017a), while the prior on Λ_s is chosen to match the analysis employed by Abbott et al. (2018). The range of viewing angles, $0.82 < \cos \theta < 0.97$, is from Nakar & Piran (2020), who compiled results from GRB afterglow modelling (Alexander et al. 2017, 2018; Haggard et al. 2017; Margutti et al. 2017, 2018b; Troja et al. 2017, 2018, 2019; D’Avanzo et al. 2018; Dobie et al. 2018; Gill & Granot 2018; Granot et al. 2018; Lazzati et al. 2018; Lyman et al. 2018; Mooley et al. 2018; Fong et al. 2019; Hajela et al. 2019; Lamb et al. 2019; Wu & MacFadyen 2019; Ryan et al. 2020) and from very long baseline interferometry (VLBI; Ghirlanda et al. 2019; Hotokezaka et al. 2019). We assume a distance of 40.7 Mpc (Cantiello et al. 2018) to the host galaxy, NGC 4993.

We run our fits on the University of Birmingham BLUEBEAR cluster, and use DYNesty (Speagle 2020) to integrate the model evidence and sample the posteriors of the model parameter space. We use the likelihood function,

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{i=1}^n \left[\frac{(O_i - M_i)^2}{\sigma_i^2 + \sigma^2} - \ln(2\pi\sigma_i^2) \right] - \frac{n}{2} \ln(2\pi\sigma^2), \quad (8)$$

where O_i and M_i are the set of observed and model magnitudes, σ_i are the errors on the data, and σ is a white noise free parameter to account for additional uncertainty in the data or model; for a good fit within the observed errors, one therefore finds $\sigma \lesssim \langle \sigma_i \rangle$. Speagle (2020) gives an extensive account of how the evidence is evaluated directly in DYNesty; we refer the reader there for details (and a good overview of Bayes’ theorem). In short, random draws from the model priors are used to integrate numerically the hypervolume between shells of constant likelihood. The products of likelihood and volume for these points yield a set of weights that can be summed to determine the total evidence. Points are evolved by replacing those with lowest likelihood; the weights of the final set are proportional to the model posterior (Speagle 2020).

3.1 Model comparison and importance of shock cooling

We begin by fitting GW170817 with four variants of the model to determine the relative importance of different effects. These are (i) the baseline model ($\alpha = \cos \theta_c = 1$), (ii) a model with blue ejecta enhancement from e.g. magnetic surface winds ($\alpha \leq 1$), (iii) a model with shock-cooling of a GRB cocoon ($\cos \theta_c \leq 1$), and (iv) an agnostic model allowing for both effects ($\alpha, \cos \theta_c \leq 1$). The model evidence and marginalized posteriors are given in Table 1. We show resulting light curves compared to the GW170817 data in Fig. 4.

We use the Bayes factor, $B \equiv Z_1/Z_2$, where Z_i is the Bayesian evidence for model i , to compare these models. Generally, $B > 10$ is taken to indicate a strong preference, based on the available data, for one model over another, and $B > 100$ to indicate a decisive preference. Compared to the base model, the model with enhanced surface wind ejecta is preferred with $B > 10^8$.

However, our fits overwhelmingly prefer models that include shock cooling compared to either of the other models, with Bayes factors $B > 10^{10}$ compared to the surface ejecta model and $B > 10^{19}$ compared

to the base model. The agnostic model (with both surface winds and shocks) is weakly favoured over the shock-only model, with $B = 2.2$. The decisive preference for shock cooling can be understood from Fig. 4. Models without shocks, generating their early luminosity only from radioactive decay in the blue ejecta, are too faint by ≈ 0.7 magnitudes (almost a factor of 2 in luminosity) at $t \sim 0.5$ d. Increasing the blue ejecta mass through the parameter α improves the fit at $t \sim 1.5$ d, but is still too faint during the first night (see the Appendix). However, a cocoon with an opening angle $\theta_c \sim 20^\circ$ – 30° provides the excess luminosity at $t = 0.5$ d required to match the data. This supports previous work that has argued for the importance of shock cooling in the early light curve (Kasliwal et al. 2017; Gottlieb, Nakar & Piran 2018; Piro & Kollmeier 2018).

As an important caveat, we note that there are other models that may be degenerate with cocoon emission. Free neutron decays, if present, could heat the outermost ejecta during the first few hours (Kulkarni 2005; Metzger et al. 2015), and produce a signal degenerate with shock heating (Metzger et al. 2018). Nativi et al. (2021) and Klion et al. (2021) recently showed that the bright early light curve can alternatively be explained by accounting for an asymmetric ejecta distribution induced by the GRB jet. Finally, any differences in the nuclear heating rates from those assumed here would also affect the peak luminosity, as recently shown by Zhu et al. (2021). These effects are not modeled within our Bayesian analysis, and therefore remain a potential alternative to the shock-cooling interpretation.

3.2 Fit parameters and the origin of each luminosity component

Taking the agnostic version as our preferred model, we show the two-dimensional posteriors for this fit in Fig. 5. As with all model variants considered, the chirp mass posterior is essentially the prior, since this parameter is constrained to such high precision by the GW data. The electromagnetic data tighten the mass ratio of the system to $q > 0.84$ (about 50 per cent of the GW-inferred prior volume) due to the requirement for blue ejecta. Viewing angles close to $\sim 30^\circ$ are preferred, more in line with afterglow modelling than with VLBI (Hajela et al. 2019; Nakar & Piran 2020). This is also consistent with the GW-inferred viewing angle at the distance of the GW170817 host (Finstad et al. 2018).

The largest degeneracies in the model parameters are between the mass ratio and the surface ejecta enhancement $1/\alpha$, since equal-mass binaries also produce more blue ejecta; between M_{TOV} and ϵ_{disc} , since both affect the mass of purple disc ejecta; and between θ_c and $\tilde{\Lambda}$, with a larger shock heating contribution to the luminosity relaxing the preference for low tidal deformability (compact EoS). We will discuss the EoS-dependent quantities, M_{TOV} and $R_{1.4}(\tilde{\Lambda})$, in more detail in Section 3.3.

The best-fitting models produce typical ejecta masses $M_{\text{blue}} \approx 0.005$, $M_{\text{red}} \approx 0.001$, and $M_{\text{purple}} \approx 0.02 M_\odot$. As we have established, the blue ejecta can only account for ~ 50 per cent of the flux at $t \sim 0.5$ d, and the rest is attributed to cooling of a GRB-shocked cocoon. The large mass of purple ejecta dominates the luminosity at $t \sim 2$ – 10 d, with the low mass of red ejecta contributing significantly only in the extended tail. These results are in agreement with Villar et al. (2017); however in our forward model we can uniquely associate the purple component with remnant disc winds, the blue component with shock-driven dynamical ejecta (possibly but not necessarily boosted by magnetic surface winds), and the red component with tidal dynamical ejecta, while identifying the importance of the cocoon contribution to the luminosity at early times.

3.3 Constraints on the NS EoS

In this section, we examine the posteriors for EoS-dependent quantities in our fit, including estimates of the systematic uncertainty. While M_{TOV} is treated explicitly as a model parameter, the NS radii must be derived from the posterior of the symmetric tidal deformability, $\Lambda_s = 240^{+94}_{-74}$, and the masses of the two NSs, obtained from \mathcal{M} and q .

The major source of systematic uncertainty in $R_{1.4}$ and M_{TOV} is the scatter in relations (1) and (3), which are, respectively, ~ 70 per cent for the dynamical ejecta mass (Dietrich & Ujevic 2017) and ~ 50 per cent of our inferred disc mass (Coughlin et al. 2019a). We test two methods to simultaneously model these systematic errors. First, we run a suite of fits where equations (1) and (3) are modified each time by a random factor drawn from a Gaussian distribution with mean equal to 1 and width equal to the calibration uncertainties. Comparing the posteriors of 50 such runs, we find a scatter of ± 40 in Λ_s , and negligible scatter in M_{TOV} . As a complementary method, we run a single realization of the fit but with two additional free parameters with Gaussian priors, σ_{dyn} and σ_{disc} (where the dynamical ejecta mass is $M'_{\text{dyn}} = \sigma_{\text{dyn}} M_{\text{dyn}}$ and the disc ejecta is treated analogously). In the former case, we broaden the posterior of Λ_s using random draws from a Gaussian of width 40, whereas in the latter case, we can use the posterior distribution of Λ_s directly. We find that the two approaches yield indistinguishable posteriors for Λ_s (and hence $R_{1.4}$) and M_{TOV} .

Since the systematic error is negligible for M_{TOV} , our estimate of this quantity is equal to the posterior distribution shown in Fig. 5. We measure a 90 per cent credible interval $M_{\text{TOV}} = 2.17^{+0.08}_{-0.11} M_\odot$. Models without shock cooling (strongly disfavoured by the Bayes factor analysis) prefer a more massive $M_{\text{TOV}} \sim 2.3 M_\odot$. For comparison, the analysis of Margalit & Metzger (2017) found that $M_{\text{TOV}} \lesssim 2.2 M_\odot$, which is consistent with our preferred shock cooling models but in potential tension with the shock-free models (although Shibata et al. 2019 have extended the Margalit & Metzger 2017 analysis and found $M_{\text{TOV}} \lesssim 2.3 M_\odot$). Other recent works (Rezzolla, Most & Weih 2018; Lucca & Sagunski 2020; Shao et al. 2020) have also favoured an EoS with $M_{\text{TOV}} \lesssim 2.2 M_\odot$. It is worth noting that the M_{TOV} constraints obtained here are complementary to these other works, which were based on arguments that are independent from the kilonova modelling adopted here. It is therefore interesting (and non-trivial) that the different approaches yield consistent results.

We next derive $R_{1.4}$ from the posteriors of Λ_s , \mathcal{M} , and q , using equations (2), (6), and (7). During this conversion, we include the systematic uncertainty introduced by the QURs (Yagi & Yunes 2016, 2017) by adding random scatter to the derived values for C_i (6.5 per cent) and Λ_a (5 per cent), drawn from Gaussian distributions. The masses of the primary and secondary, which enter equation (2), are $M_1 = 1.41 \pm 0.04$ and $M_2 = 1.31 \pm 0.03 M_\odot$; therefore, their measured radii (especially for the primary) can be taken as very close to $R_{1.4}$.

The top panel of Fig. 6 shows the derived probability density functions for the NS radii evaluated in this way. We find the primary has a radius $R_1 = 10.72^{+1.67}_{-1.48}$ km (90 per cent credible interval). The models without shock cooling prefer a smaller radius, ~ 10 km, due to their lower values for Λ_s . The most direct comparison for our radius is with the GW-inferred value; Abbott et al. (2018) find $R_{\text{GW}} = 10.8^{+2.0}_{-1.7}$ km (without imposing any EoS constraints a priori). Our measurement is in excellent agreement with this value, and has a 90 per cent credible bound that is smaller by ~ 0.5 km.

Abbott et al. (2018) find a larger value for the radius and a tighter posterior distribution, $R_{\text{GW}} = 11.9 \pm 1.4$ km, if they impose a constraint that the radius must be compatible with equations of

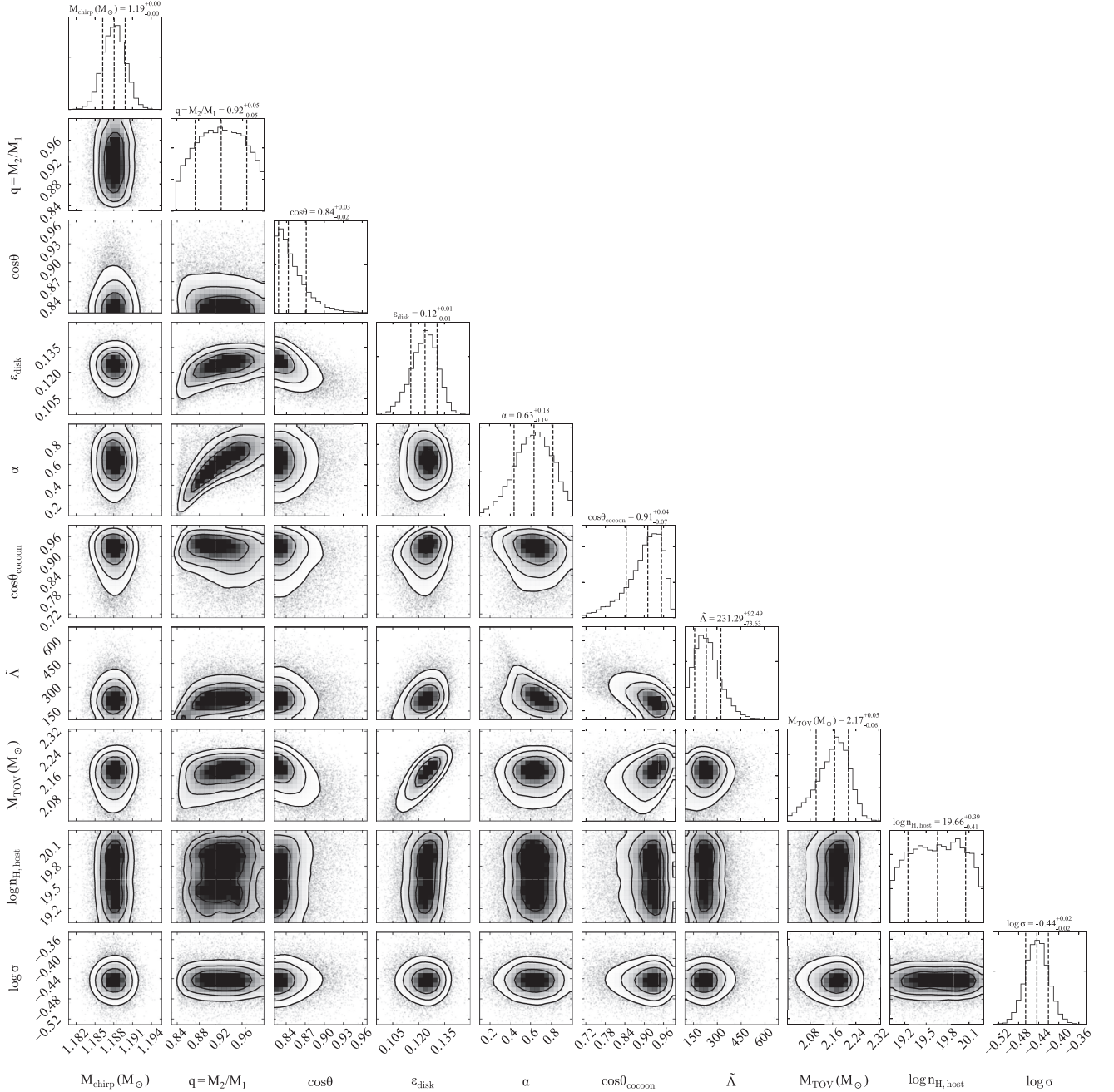


Figure 5. Posterior distributions of free parameters for the preferred agnostic model (including shock cooling and enhanced surface wind ejecta) fit to GW170817, using multimessenger priors (Table 1).

state that support a TOV mass larger than the most massive known NS. In a similar vein, we can use our constraint on M_{TOV} to tighten our posteriors on the NS radii. This assumes that the posteriors of $\tilde{\Lambda}$ and M_{TOV} are uncorrelated (Fig. 5 shows this to be a good approximation). We download the mass–radius curves for a sample of 4000 equations of state from Dietrich et al. (2020), constructed using chiral effective theory (up to 1.5 times nuclear saturation density, and a speed-of-sound parametrization at higher densities) – an approach that is motivated by nuclear theory (Tews, Margueron & Reddy 2018; Capano et al. 2020). As shown in the bottom panel of Fig. 6, we select only those that support a TOV mass within our 90 per cent confidence interval. The inset histogram shows the distribution of

$R_{1.4}$ for this EoS subset. We take this as the a priori probability of a given radius in order to re-weight our posteriors for R_1 and R_2 (after first transforming to a prior flat in $R_{1.4}$ following Raithel, Ozel & Psaltis 2020), resulting in the middle panel of Fig. 6. Our final measurement using Λ_s and ensuring consistency with M_{TOV} is $R_1 = 11.06^{+1.01}_{-0.98}$ km.

Similar to Abbott et al. (2018), imposing a constraint on the EoS results in an increase in $R_{1.4}$ and a tightening of the posterior. Our final value is consistent with theirs at around the 1σ level, and narrower by ~ 0.8 km. We also compare to multimessenger constraints from the literature that used the kilonova data for GW170817. Our median value for the radius lies just below the [11.3, 13.5] km credible region

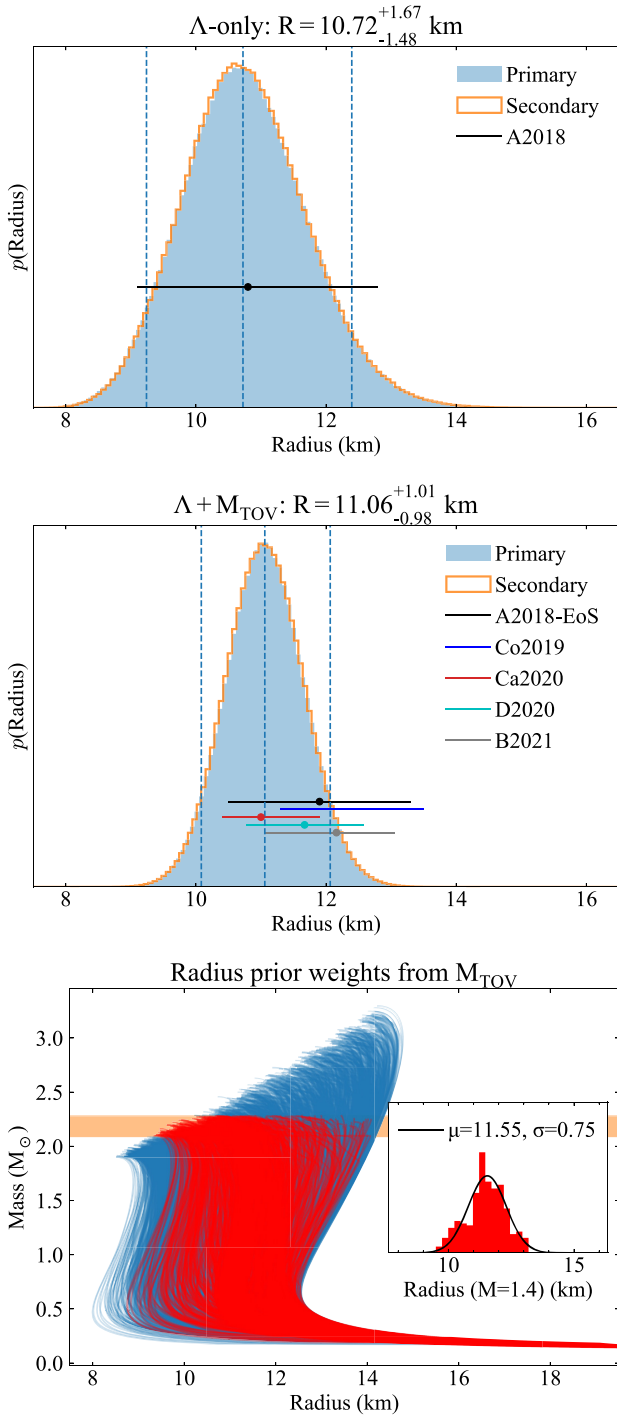


Figure 6. Posterior distributions for the NS radii using the agnostic model, including systematic errors. Vertical lines mark the 5th, 50th, and 95th percentiles for the primary (more massive) NS. Plotted for comparison are the measurements by Abbott et al. (2018) (with and without imposing constraints on the EoS), Coughlin et al. (2019a), Capano et al. (2020), Dietrich et al. (2020), and Breschi et al. (2021). Top panel: NS radii from the posterior of Λ_s . Middle panel: NS radii after re-weighting using equations of state that support an M_{TOV} within our 90 per cent confidence interval. Bottom: construction of these prior weights using mass-radius curves from Dietrich et al. (2020). The red curves satisfy our constraint on M_{TOV} . Inset: Gaussian fit to allowed radii.

obtained by Coughlin et al. (2019a), though there is significant overlap between their bounds and our 90 per cent confidence interval. Our measurement is in excellent agreement with the $11.0^{+0.9}_{-0.6}$ km found by Capano et al. (2020). It is also consistent with the $11.75^{+0.86}_{-0.81}$ km found by Dietrich et al. (2020), at around the 1σ level. Breschi et al. (2021) and Most et al. (2018) report slightly larger values for $R_{1.4}$, finding it to be $12.16^{+0.89}_{-1.11}$ or $12.39^{+1.06}_{-0.39}$ km, respectively, which are consistent with our results only at the $\sim 2\sigma$ level.

3.4 The importance of multimessenger data

In this section we emphasize the importance of the *multimessenger* constraints we used to fit GW170817. Table 1 lists two sets of priors: one set of suggested default priors for fitting an arbitrary NS merger, and one specifically tuned for GW170817. This latter set of priors, which we employed in the previous section, makes use of the GW constraints on \mathcal{M} , q , t_0 , and $\tilde{\Lambda}$, GRB afterglow constraints on θ (the GRB also constrains t_0), and constraints on extinction from analysis of the host galaxy (Blanchard et al. 2017; Levan et al. 2017; Pan et al. 2017).

Taking our preferred model for GW170817 (the agnostic model including both shock cooling and an enhanced surface wind ejecta), we change the priors to the default set and re-run the fit to GW170817. Fig. 7 shows the difference in posteriors when using the default priors, compared to the results using the multimessenger priors. We find a fit of similar quality ($\ln Z = 79.2$), but with much broader posterior distributions for several parameters.

The posterior distribution of \mathcal{M} obtained using the broad priors has a median $\mathcal{M}_{\text{broad}} = 1.2 M_{\odot}$, close to the known value from the GW data. However, the credible region is ~ 30 times wider than when using the very tight GW constraints. Another notable difference is in the viewing angle, where much larger viewing angles $\theta_{\text{broad}} \sim 45^\circ$ are preferred without the GRB constraints. The distribution of mass ratio spans a similar range as in the multimessenger case, but is skewed towards more equal systems ($q_{\text{broad}} \sim 0.94$).

The posteriors of parameters that use the same priors in the default and multimessenger models are (unsurprisingly) affected less severely, e.g. the posteriors of α and $\cos \theta_c$ are very similar in the two cases, as are the posteriors of $\tilde{\Lambda}$. However, the weaker constraints on the total system mass lead to a broadening of the posterior distribution of M_{TOV} , and a skew towards larger values.

The takeaway message is that our model can do a reasonable job of measuring important system properties for a GW170817-like merger even without a GW detection. However, having information from electromagnetic, GW, GRB, and host galaxy analyses leads to tighter constraints, and this effect is perhaps most important when attempting to probe the NS EoS, as in Section 3.3.

4 LIGHT-CURVE PREDICTIONS: EXAMPLE OF GW190425

Having validated our model using GW170817, we now (briefly) demonstrate its utility for generating simulated kilonova data. While the broader applications in terms of population synthesis will be explored in future works, an important use case that we discuss here is to predict the electromagnetic observables of a particular kilonova given a detected GW signal (see e.g. Margalit & Metzger 2019), in order to make follow-up observations more efficient. This is especially relevant given the possibility of target-of-opportunity GW follow-up with the upcoming Vera Rubin Observatory (Margutti et al. 2018a; Cowperthwaite et al. 2019; Smith et al. 2019; Chen et al. 2020).

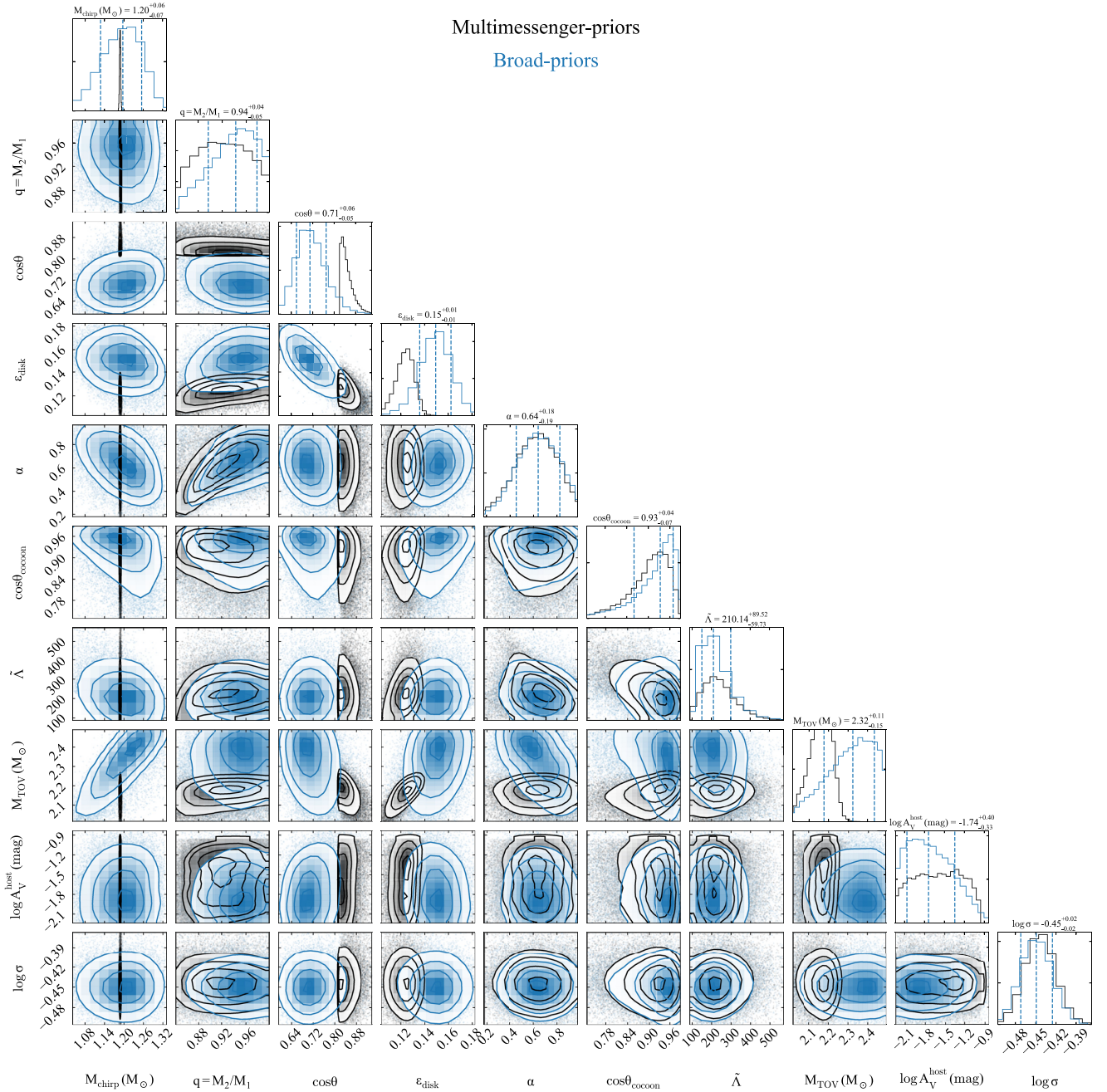


Figure 7. Posterior distributions for model fit to GW170817 when using the broad priors (blue), compared to the posteriors obtained using the multimessenger priors (black; same as Fig. 5). Both sets of priors are listed in Table 1.

During the LIGO-Virgo O3 run, the GW detectors identified several binary NS candidates. One event is contained in the latest GW source catalog, GWTC-2 (Abbott et al. 2020a), with the others discovered in the second half of O3 such that source parameters have not yet been released. The published event is GW190425, which was notable for its large total mass $M_1 + M_2 = 3.4 M_{\odot}$ (Abbott et al. 2020c), implying that (unlike GW170817) at least one of the constituent NSs was more massive than the canonical $\approx 1.4 M_{\odot}$.

Various groups of astronomers followed up the GW discovery of GW190425 with electromagnetic observations in an attempt to detect a counterpart, though without success on this occasion (Coughlin et al. 2019b; Hosseinzadeh et al. 2019; Lundquist et al. 2019; Antier

et al. 2020). This may have been due in part to the large distance to the source, 159^{+69}_{-71} Mpc, and in part because the signal was significant only in one GW detector, leading to a wide sky localization of $> 8000 \text{ deg}^2$ (Abbott et al. 2020c).

Hosseinzadeh et al. (2019) compiled detection limits from both galaxy-targeted and publicly reported wide-field searches, and compared to the light curves of GW170817 (shifted to the distance of GW190425). In Fig. 8, we show these limits compared to our best-fitting model of GW170817 in the left-hand panel. In agreement with Hosseinzadeh et al. (2019), we see that the deep galaxy-targeted searches rule out a GW170817-like counterpart (in those particular galaxies) over most of the 90 per cent plausible distance range.

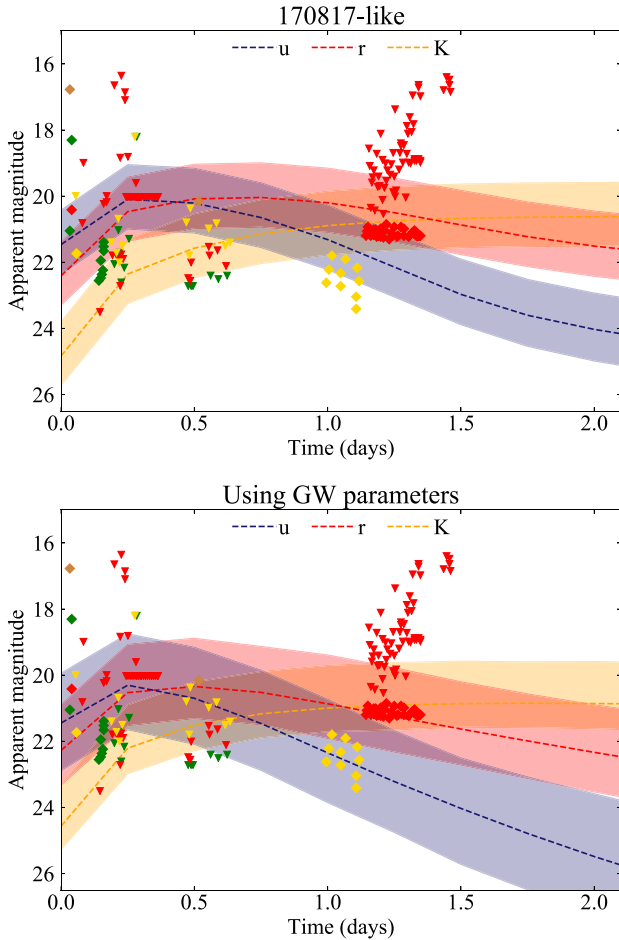


Figure 8. Model light curves for a GW190425 kilonova compared to limits compiled by Hosseinzadeh et al. (2019). The shaded regions correspond to the 90 per cent credible ranges in magnitude for each band. Top panel: assuming a GW170817-like kilonova shifted to the GW-inferred distance of GW190425 (following Hosseinzadeh et al. 2019). Bottom panel: predicting the light curve directly from the GW constraints on the chirp mass, mass ratio, distance, and inclination. The wider uncertainty bands and faster decline in this case show that assuming a GW170817-like event leads to unrealistically tight constraints and observations that may be too shallow to confidently detect this source.

With our new model, we can use the chirp mass and mass ratio constraints obtained by LIGO and Virgo for GW190425 to predict a light curve more directly applicable to this source. A similar approach was adopted by Barbieri et al. (2020) for both NS-NS and NS-BH binaries. This prediction is not only tuned specifically to GW190425, but also folds in the uncertainties on the binary parameters as well as the distance uncertainty. We use a Gaussian distribution for the chirp mass $\mathcal{M} = 1.44 \pm 0.013$, and flat distributions $0.8 < q < 1$, $0 < \cos \theta < 1$, $0.71 < \cos \theta_c < 1$, and a fiducial $\epsilon_{\text{disc}} = 0.15$ based on GW170817. The result is shown in the right-hand panel of Fig. 8.

Two features of this prediction are noteworthy. One is the increased width of the 90 per cent distribution of credible light curves, allowing more space for a faint kilonova to go undetected at the observed depths. More significantly, the median light curve is fainter (by ~ 0.7 mag in the r band at 1 d post-merger) and declines more rapidly than GW170817. The main reason for this is that the more massive remnant in this system is expected to collapse promptly to a BH, reducing the mass of the purple ejecta component by an order of

magnitude compared to GW170817. With this faster decline, even many of the galaxy-targeted observations obtained $\gtrsim 1$ d after merger rule out only ~ 50 per cent of the plausible light curves.

Our goal here is not to calculate precisely the fraction of the localization volume probed by optical observations of GW190425, but simply to make the general point that how constraining a detection limit is depends on the choice of comparison model, and that this in turn depends on the binary source parameters. A better strategy in future would be to tune the observed depths to a level necessary to cover the 90 per cent credible range of kilonova luminosities for a given GW detection. At present, this is not possible, as the GW-inferred binary parameters needed to simulate the light curve are not released at the time of merger. We suggest that greater cooperation between the GW and electromagnetic astronomy communities could be very useful in this regard: If the GW source parameters (\mathcal{M} , q , $\tilde{\Lambda}$) were released in low latency, we could use our model to immediately predict the required observing depth at any wavelength to find or rule out a kilonova counterpart at high significance.

5 CONCLUSIONS

We have presented a simple forward model that combines a range of literature results on ejecta masses, r-process opacities, and the NS EoS, to predict kilonova light curves directly from GW-accessible binary parameters. This model can also be used to derive limits on the kilonova rate from optical surveys, given a model for the BNS source population; this will be the subject of future work. Our code has been developed within the MOSFIT framework and is publicly available.

We validated the model by fitting it to the well-observed kilonova associated with GW170817, incorporating GW and GRB afterglow results in our priors. We found that a model including a shock cooling cocoon component was overwhelmingly favoured by the Bayes factor, compared to models with r-process radioactive decay as the only luminosity source. Data obtained during the first day after merger was essential to differentiating these models (Arcavi 2018). Our approach directly associates each of the r-process ejecta components with a physical origin, where the low opacity material is shock-heated dynamical ejecta, and the higher opacity material that dominates the emission from ~ 2 to 10 d after merger is a viscous disc wind. The very lanthanide-rich tidal ejecta makes a significant contribution only at very late times.

The posteriors of our model fit contain important information on the GW170817 progenitor system. We find that this system likely had a small mass ratio, $q \approx 0.9$, and was viewed $\approx 30^\circ$ off-axis. By including the maximum stable NS mass M_{TOV} and the symmetric tidal deformability Λ_s in our free parameters, we constrained these EoS-dependent parameters. We measure $M_{\text{TOV}} = 2.17^{+0.08}_{-0.11} M_\odot$ (90 per cent credible interval) and derive $R_{1.4} = 11.06^{+1.01}_{-0.98}$ km using the posteriors of Λ , \mathcal{M} , q , and M_{TOV} , carefully taking into account the systematic errors on all relations used. This radius constraint is consistent and competitive with others in the literature.

For high-significance detections in O3, the LIGO-Virgo collaboration provided publicly the distance and source classification probabilities (as NS, BH, NS-BH, or mass-gap binary) along with the sky localization, giving astronomers a rough guide to sources they may wish to follow up electromagnetically. However, even for a source confidently classified as a NS merger, the binary parameters – especially the chirp mass, mass ratio, and viewing angle – have a substantial impact on the likelihood of detecting the counterpart, since they control the peak luminosity, colour, and decline rate

(Figs 2 and 3). If such parameters (which may be estimated even in low latency; Biscoveanu, Vitale & Haster 2019; Finstad & Brown 2020; Krastev et al. 2020) are made available to the astronomical community, this can aid real-time kilonova searches following binary NS merger discoveries from the GW detector network (also advocated for by e.g. Margalit & Metzger 2019; Barbieri et al. 2020). We demonstrated this by synthesizing light curves appropriate to GW190425, for which optical searches did not uncover a counterpart, finding that many were not deep enough to fully rule out a kilonova. For future observing runs with a higher binary NS detection rate, the ability to better predict the required observing depth on a case-by-case basis would enable more efficient allocation of telescope time (which is of course a finite resource) for GW follow-up.

We have not discussed in this work the signature of a merger between a NS and a companion BH; these are also expected to produce kilonovae as long as the mass ratio is modest (though without shock-driven, low-lanthanide ejecta). Our next step will be to implement an analogous model to simulate light curves for this class of events (see also Barbieri et al. 2020).

Finding a larger sample of kilonovae in the coming years, as GW detectors and optical telescopes increase their survey power, will be crucial to improve our understanding of the NS EoS and the source population of NS binaries. Hopefully, the code we present here will be useful to the community both in helping to plan observations that increase this sample and in deriving important physics from the data obtained.

ACKNOWLEDGEMENTS

We thank the referee Rahul Kashyap for valuable suggestions that improved this paper, Jenni Barnes for providing kilonova thermalization data, and Edo Berger, Ryan Chornock, Laurence Datrier, Michael Fulton, Martin Hendry, Albert Kong, Gavin Lamb, Brian Metzger, Christopher Moore, Geraint Pratten, Surojit Saha, Stephen Smartt, Alberto Vecchio, and Ashley Villar for their insights and discussions. MN is supported by a Royal Astronomical Society Research Fellowship. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 948381). BM is supported by NASA through the NASA Hubble Fellowship grant #HST-HF2-51412.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. PS is supported by the Dutch Research Council (NWO) Veni Grant No. 680-47-460. GPS acknowledges support from the Science and Technology Facilities Council through grant number ST/N021702/1.

DATA AVAILABILITY

This paper makes use of existing public data from kilonova.space (Guillochon et al. 2017). The code has been made available at <https://github.com/guillochon/MOSFiT>. A population of kilonova models produced by this code using broad priors, in UV, optical, and NIR filters, are available for download at <https://github.com/mnicholl/kn-models-nicholl2021>.

REFERENCES

Abbott B. P. et al., 2017a, *Phys. Rev. Lett.*, 119, 161101
 Abbott B. P. et al., 2017b, *ApJ*, 848, L12
 Abbott B. P. et al., 2018, *Phys. Rev. Lett.*, 121, 161101

Abbott R. et al., 2020a, preprint ([arXiv:2010.14527](https://arxiv.org/abs/2010.14527))
 Abbott B. P. et al., 2020b, *Living Rev. Relativ.*, 23, 3
 Abbott B. P. et al., 2020c, *ApJ*, 892, L3
 Abbott R. et al., 2020d, *ApJ*, 896, L44
 Ackley K. et al., 2020, *A&A*, 643, A113
 Alexander K. D. et al., 2017, *ApJ*, 848, L21
 Alexander K. D. et al., 2018, *ApJ*, 863, L18
 Anand S. et al., 2021, *Nat. Astron.*, 5, 46
 Andreoni I. et al., 2017, *Publ. Astron. Soc. Aust.*, 34, e069
 Andreoni I. et al., 2020a, *ApJ*, 904, 19
 Andreoni I. et al., 2020b, *ApJ*, 890, 131
 Antier S. et al., 2020, *MNRAS*, 492, 3904
 Antoniadis J. et al., 2013, *Science*, 340, 448
 Arcavi I., 2018, *ApJ*, 855, L23
 Arcavi I. et al., 2017, *ApJ*, 848, L33
 Arnett W. D., 1982, *ApJ*, 253, 785
 Barbieri C., Salafia O. S., Colpi M., Ghirlanda G., Perego A., 2020, preprint ([arXiv:2002.09395](https://arxiv.org/abs/2002.09395))
 Barnes J., Kasen D., 2013, *ApJ*, 775, 18
 Barnes J., Kasen D., Wu M.-R., Martínez-Pinedo G., 2016, *ApJ*, 829, 110
 Bauswein A., Baumgarte T. W., Janka H. T., 2013a, *Phys. Rev. Lett.*, 111, 131101
 Bauswein A., Goriely S., Janka H.-T., 2013b, *ApJ*, 773, 78
 Bauswein A., Just O., Janka H.-T., Stergioulas N., 2017, *ApJ*, 850, L34
 Berger E., 2014, *ARA&A*, 52, 43
 Berger E., Fong W., Chornock R., 2013, *ApJ*, 774, L23
 Biscoveanu S., Vitale S., Haster C.-J., 2019, *ApJ*, 884, L32
 Blanchard P. K. et al., 2017, *ApJ*, 848, L22
 Breschi M., Perego A., Bernuzzi S., Del Pozzo W., Nedora V., Radice D., Vescovi D., 2021, preprint ([arXiv:2101.01201](https://arxiv.org/abs/2101.01201))
 Bulla M., 2019, *MNRAS*, 489, 5037
 Burbidge E. M., Burbidge G. R., Fowler W. A., Hoyle F., 1957, *Rev. Mod. Phys.*, 29, 547
 Cantiello M. et al., 2018, *ApJ*, 854, L31
 Capano C. D. et al., 2020, *Nat. Astron.*, 4, 625
 Chen H.-Y., Cowperthwaite P. S., Metzger B. D., Berger E., 2020, *ApJ*, 908, L4
 Chornock R. et al., 2017, *ApJ*, 848, L19
 Ciolfi R., Kalinani J. V., 2020, *ApJ*, 900, L35
 Ciolfi R., Kastaun W., Giacomazzo B., Endrizzi A., Siegel D. M., Perna R., 2017, *Phys. Rev. D*, 95, 063016
 Coughlin M. W., Dietrich T., Margalit B., Metzger B. D., 2019a, *MNRAS*, 489, L91
 Coughlin M. W. et al., 2019b, *ApJ*, 885, L19
 Coulter D. A. et al., 2017, *Science*, 358, 1556
 Cowperthwaite P. S. et al., 2017, *ApJ*, 848, L17
 Cowperthwaite P. S., Chen H.-Y., Margalit B., Margutti R., May M., Metzger B., Pankow C., 2019, preprint ([arXiv:1904.02718](https://arxiv.org/abs/1904.02718))
 Cromartie H. T. et al., 2020, *Nat. Astron.*, 4, 72
 D'Avanzo P. et al., 2018, *A&A*, 613, L1
 Damour T., Nagar A., Villain L., 2012, *Phys. Rev. D*, 85, 123007
 Darbha S., Kasen D., 2020, *ApJ*, 897, 150
 Davies M. B., Benz W., Piran T., Thielemann F. K., 1994, *ApJ*, 431, 742
 De S., Siegel D., 2020, preprint ([arXiv:2011.07176](https://arxiv.org/abs/2011.07176))
 De S., Finstad D., Lattimer J. M., Brown D. A., Berger E., Biwer C. M., 2018, *Phys. Rev. Lett.*, 121, 091102
 Demorest P. B., Pennucci T., Ransom S. M., Roberts M. S. E., Hessels J. W. T., 2010, *Nature*, 467, 1081
 Díaz M. C. et al., 2017, *ApJ*, 848, L29
 Dietrich T., Ujevic M., 2017, *Class. Quantum Gravity*, 34, 105014
 Dietrich T., Bernuzzi S., Ujevic M., Brüggmann B., 2015, *Phys. Rev. D*, 91, 124041
 Dietrich T., Ujevic M., Tichy W., Bernuzzi S., Brüggmann B., 2017, *Phys. Rev. D*, 95, 024029
 Dietrich T., Coughlin M. W., Pang P. T. H., Bulla M., Heinzel J., Issa L., Tews I., Antier S., 2020, *Science*, 370, 1450
 Dobie D. et al., 2018, *ApJ*, 858, L15
 Drout M. R. et al., 2017, *Science*, 358, 1570

- Eichler D., Livio M., Piran T., Schramm D. N., 1989, *Nature*, 340, 126
- Evans P. A. et al., 2017, *Science*, 358, 1565
- Favata M., 2014, *Phys. Rev. Lett.*, 112, 101101
- Fernández R., Tchekhovskoy A., Quataert E., Foucart F., Kasen D., 2019, *MNRAS*, 482, 3373
- Finstad D., Brown D. A., 2020, *ApJ*, 905, L9
- Finstad D., De S., Brown D. A., Berger E., Biwer C. M., 2018, *ApJ*, 860, L2
- Flanagan É. É., Hinderer T., 2008, *Phys. Rev. D*, 77, 021502
- Fong W. et al., 2019, *ApJ*, 883, L1
- Fong W. et al., 2020, *ApJ*, 906, 26
- Foucart F., O'Connor E., Roberts L., Kidder L. E., Pfeiffer H. P., Scheel M. A., 2016, *Phys. Rev. D*, 94, 123016
- Freiburghaus C., Rosswog S., Thielemann F. K., 1999, *ApJ*, 525, L121
- Gao H., Ai S.-K., Cao Z.-J., Zhang B., Zhu Z.-Y., Li A., Zhang N.-B., Bauswein A., 2020, *Frontiers Phys.*, 15, 24603
- Ghirlanda G. et al., 2019, *Science*, 363, 968
- Gill R., Granot J., 2018, *MNRAS*, 478, 4128
- Goldstein A. et al., 2017, *ApJ*, 848, L14
- Gomez S. et al., 2019, *ApJ*, 884, L55
- Gompertz B. P. et al., 2018, *ApJ*, 860, 62
- Gompertz B. P. et al., 2020, *MNRAS*, 497, 726
- Goriely S., Bauswein A., Just O., Plumbi E., Janka H. T., 2015, *MNRAS*, 452, 3894
- Gottlieb O., Nakar E., Piran T., 2018, *MNRAS*, 473, 576
- Graham M. J. et al., 2020, *Phys. Rev. Lett.*, 124, 251102
- Granot J., Gill R., Guetta D., De Colle F., 2018, *MNRAS*, 481, 1597
- Guillochon J., Parrent J., Kelley L. Z., Margutti R., 2017, *ApJ*, 835, 64
- Guillochon J., Nicholl M., Villar V. A., Mockler B., Narayan G., Mandel K. S., Berger E., Williams P. K. G., 2018, *ApJS*, 236, 6
- Haggard D., Nynka M., Ruan J. J., Kalogera V., Cenko S. B., Evans P., Kennea J. A., 2017, *ApJ*, 848, L25
- Hajela A. et al., 2019, *ApJ*, 886, L17
- Hinderer T., 2008, *ApJ*, 677, 1216
- Hosseinzadeh G. et al., 2019, *ApJ*, 880, L4
- Hotokezaka K., Kiuchi K., Kyutoku K., Okawa H., Sekiguchi Y.-i., Shibata M., Taniguchi K., 2013, *Phys. Rev. D*, 87, 024001
- Hotokezaka K., Nakar E., Gottlieb O., Nissanke S., Masuda K., Hallinan G., Mooley K. P., Deller A. T., 2019, *Nat. Astron.*, 3, 940
- Hu L. et al., 2017, *Sci. Bull.*, 62, 1433
- Jin Z.-P. et al., 2016, *Nat. Commun.*, 7, 12898
- Kasen D., Badnell N. R., Barnes J., 2013, *ApJ*, 774, 25
- Kasen D., Fernández R., Metzger B. D., 2015, *MNRAS*, 450, 1777
- Kasen D., Metzger B., Barnes J., Quataert E., Ramirez-Ruiz E., 2017, *Nature*, 551, 80
- Kashyap R., Raman G., Ajith P., 2019, *ApJ*, 886, L19
- Kasliwal M. M. et al., 2017, *Science*, 358, 1559
- Kasliwal M. M. et al., 2020, *ApJ*, 905, 145
- Kawaguchi K., Shibata M., Tanaka M., 2018, *ApJ*, 865, L21
- Kiuchi K., Kyutoku K., Shibata M., Taniguchi K., 2019, *ApJ*, 876, L31
- Klion H., Duffell P. C., Kasen D., Quataert E., 2021, *MNRAS*, 502, 865
- Korobkin O., Rosswog S., Arcones A., Winteler C., 2012, *MNRAS*, 426, 1940
- Korobkin O. et al., 2020, *ApJ*, 889, 168
- Krastev P. G., Gill K., Villar V. A., Berger E., 2021, *Phys. Lett. B*, 815, 136161
- Krüger C. J., Foucart F., 2020, *Phys. Rev. D*, 101, 103002
- Kulkarni S. R., 2005, preprint (astro-ph/0510256)
- Lamb G. P. et al., 2019, *ApJ*, 883, 48
- Lattimer J. M., Prakash M., 2016, *Phys. Rep.*, 621, 127
- Lattimer J. M., Schramm D. N., 1974, *ApJ*, 192, L145
- Lazzati D., Perna R., Morsony B. J., Lopez-Camara D., Cantiello M., Ciolfi R., Giacomazzo B., Workman J. C., 2018, *Phys. Rev. Lett.*, 120, 241103
- Lehner L., Liebling S. L., Palenzuela C., Caballero O. L., O'Connor E., Anderson M., Neilsen D., 2016, *Class. Quantum Gravity*, 33, 184002
- Levan A. J. et al., 2017, *ApJ*, 848, L28
- Li L.-X., Paczyński B., 1998, *ApJ*, 507, L59
- Lippuner J., Fernández R., Roberts L. F., Foucart F., Kasen D., Metzger B. D., Ott C. D., 2017, *MNRAS*, 472, 904
- Lipunov V. M. et al., 2017, *ApJ*, 850, L1
- LSST Science Collaboration et al., 2009, preprint (arXiv:0912.0201)
- Lucca M., Sagunski L., 2020, *J. High Energy Astrophys.*, 27, 33
- Lundquist M. J. et al., 2019, *ApJ*, 881, L26
- Lyman J. D. et al., 2018, *Nat. Astron.*, 2, 751
- McBrien O. R. et al., 2021, *MNRAS*, 500, 4213
- McCully C. et al., 2017, *ApJ*, 848, L32
- Margalit B., Metzger B. D., 2017, *ApJ*, 850, L19
- Margalit B., Metzger B. D., 2019, *ApJ*, 880, L15
- Margutti R., Chornock R., 2020, preprint (arXiv:2012.04810)
- Margutti R. et al., 2017, *ApJ*, 848, L20
- Margutti R. et al., 2018a, preprint (arXiv:1812.04051)
- Margutti R. et al., 2018b, *ApJ*, 856, L18
- Metzger B. D., 2019, *Living Rev. Relativ.*, 23, 1
- Metzger B. D., Fernández R., 2014, *MNRAS*, 441, 3444
- Metzger B. D. et al., 2010, *MNRAS*, 406, 2650
- Metzger B. D., Bauswein A., Goriely S., Kasen D., 2015, *MNRAS*, 446, 1115
- Metzger B. D., Thompson T. A., Quataert E., 2018, *ApJ*, 856, 101
- Mochkovitch R., Daigne F., Duque R., Zitouni H., 2021, preprint (arXiv:2103.00943)
- Mooley K. P. et al., 2018, *Nature*, 554, 207
- Most E. R., Weih L. R., Rezzolla L., Schaffner-Bielich J., 2018, *Phys. Rev. Lett.*, 120, 261103
- Nakar E., Piran T., 2017, *ApJ*, 834, 28
- Nakar E., Piran T., 2021, *ApJ*, 909, 114
- Narayan R., Paczynski B., Piran T., 1992, *ApJ*, 395, L83
- Nativi L., Bulla M., Rosswog S., Lundman C., Kowal G., Gizzi D., Lamb G. P., Perego A., 2021, *MNRAS*, 500, 1772
- Nedora V. et al., 2020, preprint (arXiv:2011.11110)
- Nicholl M. et al., 2017a, *ApJ*, 848, L18
- Nicholl M., Guillochon J., Berger E., 2017b, *ApJ*, 850, 55
- O'Connor B. et al., 2021, *MNRAS*, 502, 1279
- Pan Y.-C. et al., 2017, *ApJ*, 848, L30
- Paterson K. et al., 2021, *ApJ*, 912, 128
- Pian E. et al., 2017, *Nature*, 551, 67
- Piro A. L., Kollmeier J. A., 2018, *ApJ*, 855, 103
- Postnikov S., Prakash M., Lattimer J. M., 2010, *Phys. Rev. D*, 82, 024016
- Pozanenko A. S. et al., 2018, *ApJ*, 852, L30
- Radice D., Perego A., Hotokezaka K., Fromm S. A., Bernuzzi S., Roberts L. F., 2018, *ApJ*, 869, 130
- Raithel C. A., Özel F., Psaltis D., 2018, *ApJ*, 857, L23
- Raithel C., Özel F., Psaltis D., 2020, *ApJ*, 908, 10
- Rastinejad J. C. et al., 2021, preprint (arXiv:2101.03175)
- Rezzolla L., Most E. R., Weih L. R., 2018, *ApJ*, 852, L25
- Rosswog S., Liebendörfer M., Thielemann F.-K., Davies M. B., Benz W., Piran T., 1999, *A&A*, 341, 499
- Ryan G., van Eerten H., Piro L., Troja E., 2020, *ApJ*, 896, 166
- Salafia O. S., Colpi M., Branchesi M., Chassande-Mottin E., Ghirlanda G., Ghisellini G., Vergani S. D., 2017, *ApJ*, 846, 62
- Savchenko V. et al., 2017, *ApJ*, 848, L15
- Sekiguchi Y., Kiuchi K., Kyutoku K., Shibata M., 2015, *Phys. Rev. D*, 91, 064059
- Sekiguchi Y., Kiuchi K., Kyutoku K., Shibata M., Taniguchi K., 2016, *Phys. Rev. D*, 93, 124046
- Shao D.-S., Tang S.-P., Sheng X., Jiang J.-L., Wang Y.-Z., Jin Z.-P., Fan Y.-Z., Wei D.-M., 2020, *Phys. Rev. D*, 101, 063029
- Shapiro S. L., Teukolsky S. A., 1986, *Black Holes, White Dwarfs and Neutron Stars: The Physics of Compact Objects*. Wiley, New York
- Shappee B. J. et al., 2017, *Science*, 358, 1574
- Shibata M., Hotokezaka K., 2019, *Annu. Rev. Nucl. Part. Sci.*, 69, 41
- Shibata M., Zhou E., Kiuchi K., Fujibayashi S., 2019, *Phys. Rev. D*, 100, 023015
- Siegel D. M., Metzger B. D., 2017, *Phys. Rev. Lett.*, 119, 231102
- Siegel D. M., Metzger B. D., 2018, *ApJ*, 858, 52
- Smartt S. J. et al., 2017, *Nature*, 551, 75
- Smith G. P., Robertson A., Bianconi M., Jauzac M., 2019, preprint (arXiv:1902.05140)
- Soares-Santos M. et al., 2017, *ApJ*, 848, L16
- Speagle J. S., 2020, *MNRAS*, 493, 3132

- Stachie C. et al., 2021, preprint ([arXiv:2103.01733](https://arxiv.org/abs/2103.01733))
- Tanaka M., Kato D., Gaigalas G., Kawaguchi K., 2020, *MNRAS*, 496, 1369
- Tanvir N. R., Levan A. J., Fruchter A. S., Hjorth J., Hounsell R. A., Wiersema K., Tunnicliffe R. L., 2013, *Nature*, 500, 547
- Tanvir N. R. et al., 2017, *ApJ*, 848, L27
- Tews I., Margueron J., Reddy S., 2018, *Phys. Rev. C*, 98, 045804
- Troja E. et al., 2017, *Nature*, 551, 71
- Troja E. et al., 2018, *Nat. Commun.*, 9, 4089
- Troja E. et al., 2019, *MNRAS*, 489, 2104
- Utsumi Y. et al., 2017, *PASJ*, 69, 101
- Valenti S. et al., 2017, *ApJ*, 848, L24
- Vieira N. et al., 2020, *ApJ*, 895, 96
- Villar V. A. et al., 2017, *ApJ*, 851, L21
- Wade L., Creighton J. D. E., Ochsner E., Lackey B. D., Farr B. F., Littenberg T. B., Raymond V., 2014, *Phys. Rev. D*, 89, 103012
- Wanajo S., Sekiguchi Y., Nishimura N., Kiuchi K., Kyutoku K., Shibata M., 2014, *ApJ*, 789, L39
- Wollaeger R. T. et al., 2018, *MNRAS*, 478, 3298
- Wu Y., MacFadyen A., 2019, *ApJ*, 880, L23
- Yagi K., Yunes N., 2016, *Class. Quantum Gravity*, 33, 13LT01
- Yagi K., Yunes N., 2017, *Phys. Rep.*, 681, 1
- Zhu Y. L., Lund K. A., Barnes J., Sprouse T. M., Vassh N., McLaughlin G. C., Mumpower M. R., Surman R., 2021, *ApJ*, 906, 94

APPENDIX A: NEW FITS TO SIMULATION DATA

Fig. A1 shows our quadratic fit to determine f_{red} and $f_{\text{blue}} \equiv 1 - f_{\text{red}}$, the fractions of red (high opacity) and blue (low opacity) dynamical ejecta, using the simulated Y_e distributions from Sekiguchi et al. (2016). Ejecta with $Y_e \leq 0.25$ are assumed to be red.

Fig. A2 shows our broken polynomial fit to the simulations of Tanaka et al. (2020), which we use to estimate the opacity of the purple (post-merger) ejecta as a function of Y_e .

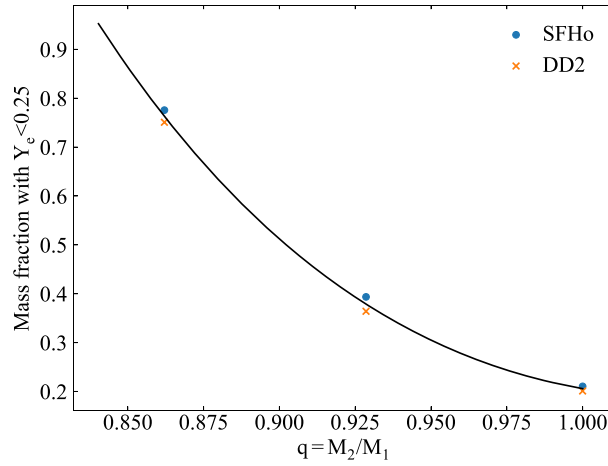


Figure A1. Mass fraction of lanthanide-poor dynamical ejecta as a function of the binary mass ratio, $q = M_2/M_1$, using Y_e distributions from Sekiguchi et al. (2016) for the SFHo (soft) and DD2 (stiff) equations of state. Our polynomial fit predicts essentially no lanthanide-poor ejecta for $q \lesssim 0.8$, in agreement with Dietrich & Ujevic (2017).

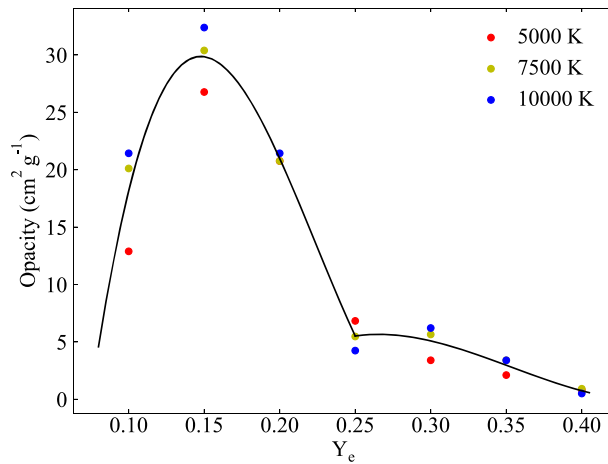


Figure A2. Opacity of r-process ejecta as a function of electron fraction and temperature from Tanaka et al. (2020). The broken polynomial fit is used to estimate the opacity of the disc ejecta, using the average Y_e predicted by Lippuner et al. (2017) for stable, short-lived, or prompt-collapse merger remnants. In practice, the average $Y_e > 0.25$ for all reasonable input parameters.

APPENDIX B: PLOTS OF ADDITIONAL LIGHT-CURVE FITS

Fig. B1 shows the light curve fits for the surface-wind-only and shock-only versions of the model.

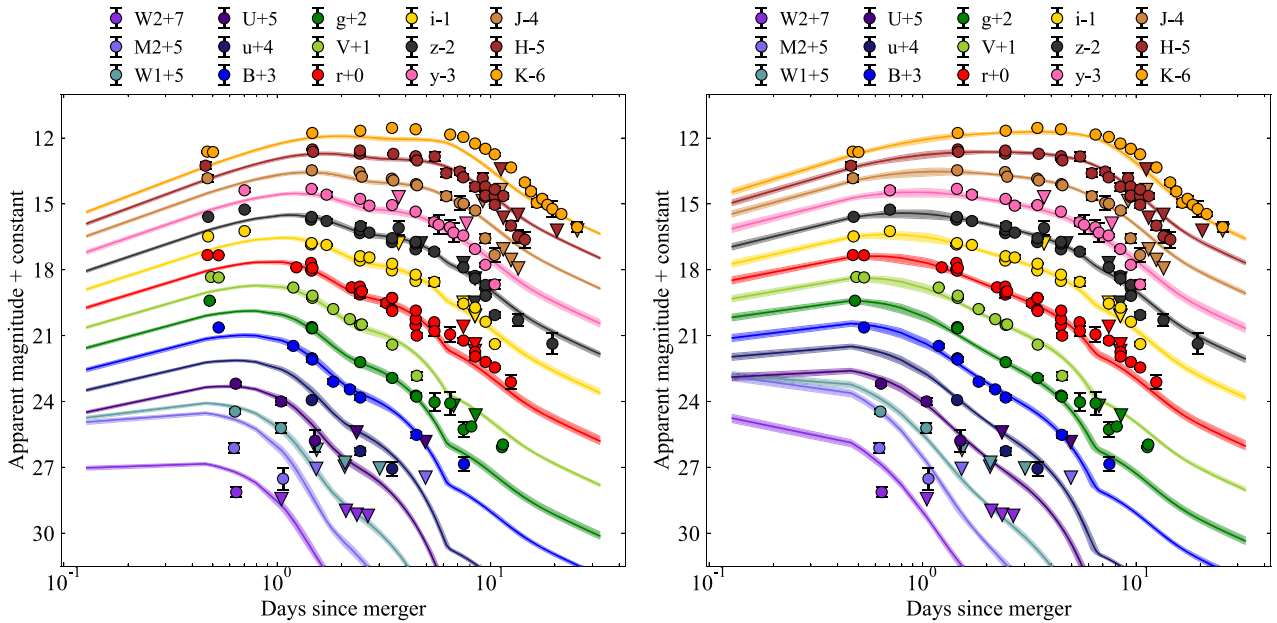


Figure B1. Left-hand panel: best-fitting kilonova model for GW170817 with surface ejecta enhancement. Right-hand panel: best-fitting model with shock cooling. The shock cooling model is preferred with a Bayes factor $B > 10^{10}$. This figure is similar to Fig. 4 in the main text, which shows the best-fitting light curves including either both or neither of these effects.

This paper has been typeset from a \LaTeX file prepared by the author.