On the Progenitor of Binary Neutron Star Merger GW170817

LIGO Scientific Collaboration and Virgo Collaboration

(See the end matter for the full list of authors.)

Received 2017 October 12; revised 2017 October 16; accepted 2017 October 16; published 2017 December 1

Abstract

On 2017 August 17 the merger of two compact objects with masses consistent with two neutron stars was discovered through gravitational-wave (GW170817), gamma-ray (GRB 170817A), and optical (SSS17a/AT 2017fo) observations. The optical source was associated with the early-type galaxy NGC 4993 at a distance of just \( \sim 40 \) Mpc, consistent with the gravitational-wave measurement, and the merger was localized to be at a projected distance of \( \sim 2 \) kpc away from the galaxy’s center. We use this minimal set of facts and the mass posteriors of the two neutron stars to derive the first constraints on the progenitor of GW170817 at the time of the second supernova (SN). We generate simulated progenitor populations and follow the three-dimensional kinematic evolution from binary neutron star (BNS) birth to the merger time, accounting for pre-SN galactic motion, for considerably different input distributions of the progenitor mass, pre-SN semimajor axis, and SN-kick velocity. Though not considerably tight, we find these constraints to be comparable to those for Galactic BNS progenitors. The derived constraints are very strongly influenced by the requirement of keeping the binary bound after the second SN and having the merger occur relatively close to the center of the galaxy. These constraints are insensitive to the galaxy’s star formation history, provided the stellar populations are older than 1 Gyr.

Key words: binaries: general – gravitational waves – stars: kinematics and dynamics – stars: neutron

1. Introduction

The era of observational gravitational-wave (GW) astronomy was firmly marked by the detection of the first binary black hole coalescence GW150914 (Abbott et al. 2016) by the Advanced LIGO detectors (Aasi et al. 2015). Discovery of a GW source accompanied by coincident electromagnetic (EM) emission, however, remained elusive until now.

On 2017 August 17 the Advanced LIGO (Aasi et al. 2015) and Advanced Virgo (Acernese et al. 2015) interferometer network recorded a transient GW signal consistent with the coalescence of a binary neutron star (BNS) GW170817 (Abbott et al. 2017b). Independently, a gamma-ray signal, classified as a short gamma-ray burst (sGRB), GRB 170817A, coincident in time and sky location with GW170817 was detected by the Fermi-GBM instrument (Abbott et al. 2017a, 2017b). The three-detector GW data analysis led to the smallest sky-localization area ever achieved for a GW source: \( \sim 31 \) deg\(^2\) when initially shared with the astronomy LIGO–Virgo partners (LIGO Scientific Collaboration & Virgo Collaboration 2017) and later improved to \( \sim 28 \) deg\(^2\) with a fully coherent data analysis (Abbott et al. 2017b).

Aided by the tight localization constraints of the three-detector network and the proximity of the GW source, multiple independent surveys across the EM spectrum were launched in search of a counterpart beyond the sGRB (Abbott et al. 2017c). Such a counterpart, SSS17a (later IAU-designated AT 2017fo), was first discovered in the optical less than 11 hours after merger, associated with the galaxy NGC 4993 (Coulter et al. 2017a, 2017b), a nearby early-type E/S0 galaxy (Lauberts 1982). Five other teams made independent detections of the same optical transient and host galaxy all within about one hour and reported their results within about five hours of one another (Allam et al. 2017; Arcavi et al. 2017a, 2017b; Lipunov 2017b; Tanvir & Levan 2017; Yang et al. 2017; Soares-Santos et al. 2017; Lipunov et al. 2017a). The same source was followed up and consistently localized at other wavelengths (e.g., Corsi et al. 2017; Deller et al. 2017a, 2017b, 2017c; Goldstein et al. 2017a, 2017b; Haggard et al. 2017a, 2017b; Mooley et al. 2017; Savchenko et al. 2017; Alexander et al. 2017; Haggard et al. 2017c; Goldstein et al. 2017; Savchenko et al. 2017). The source was reported to be offset from the center of the galaxy by a projected distance of about 10" (e.g., Coulter et al. 2017a, 2017b; Haggard et al. 2017a, 2017b; Kasliwal et al. 2017; Yang et al. 2017; Yu et al. 2017). NGC 4993 has a Tully–Fisher distance of \( \sim 40 \) Mpc (Freedman et al. 2001; NASA/IPAC Extragalactic Database\(^{164}\)), which is consistent with the luminosity distance measurement from gravitational waves (40\(^{+14}_{-8}\) Mpc). Using the Tully–Fisher distance, the \( \sim 10'' \) offset corresponds to a physical offset of \( \sim 2.0 \) kpc. This value is consistent with offset measurements of sGRBs in other galaxies, though below the median value of \( \sim 3–4 \) kpc (Fong et al. 2010; Fong & Berger 2013; Berger 2014).

BNS systems were first revealed with the discovery of PSR B1913+16, the first binary radio pulsar ever detected (Hulse & Taylor 1975). This immediately triggered new ideas for how such close pairs of neutron stars can form in nature (De Loore et al. 1975; Flannery & van den Heuvel 1975; Massevitch et al. 1976; Clark et al. 1979), based on models for the formation of high-mass X-ray binaries (van den Heuvel & Heise 1972; Tutukov & Yungelson 1973) and Wolf–Rayet X-ray binaries, for which strong orbital shrinkage is needed.

\(^{164}\) The NASA/IPAC Extragalactic Database (NED) is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.
(van den Heuvel & De Loore 1973). With years of pulsar-timing observations PSR B1913+16 provided the first firm evidence that GWs existed (Einstein 1916, 1918) and were emitted by close binary compact objects (Taylor & Weisberg 1982). This discovery greatly motivated the efforts to directly detect GWs with laser-interferometric detectors and made BNS coalescence events key targets in GW searches (see Abadie et al. 2010 for an overview).

The formation of close binaries with two neutron stars that will merge within a Hubble time is now understood to require complex evolutionary sequences of massive binaries that involve stable and unstable mass-transfer phases and two core-collapse supernova (SN) explosions through which the binary system survives (for reviews, see, e.g., Kalogera et al. 2007; Postnov & Yungelson 2014; Tauris et al. 2017). In particular, the SN explosions that lead to the formation of neutron stars are expected to develop asymmetries during the collapse, either due to neutrino emission or an anisotropic explosion (e.g., Kusenko and Segrè 1996; Janka et al. 2007; Janka 2013). This anisotropy imparts linear momentum on the stellar remnant, known as an SN kick or natal kick.

Strong evidence for this process comes from observations of Galactic pulsar proper motions, which indicate some neutron stars are moving substantially faster than the inferred speed of their progenitors and must receive a large SN kick of \( \sim 400-500 \text{ km s}^{-1} \) at birth (Lyne & Lorimer 1994; Kaspi et al. 1996; Arzoumanian et al. 2002; Chatterjee et al. 2005; Hobbs et al. 2005; Verbunt et al. 2017). However, comprehensive studies of the known BNS systems in the Milky Way have shown that some neutron stars, particularly those in binary systems, might receive smaller kicks than their isolated counterparts (Podsiadlowski et al. 2004; van den Heuvel 2007).

About a decade after the Hulse-Taylor discovery, mergers of two neutron stars were proposed as a potential source of GRBs (Goodman 1986; Paczynski 1986; Eichler et al. 1989; Narayan et al. 1992), especially those of short duration (Kouveliotou et al. 1993). Since the discovery of host galaxies for short GRBs in 2005 (Berger et al. 2005; Fox et al. 2005; Gehrels et al. 2005; Hjorth et al. 2005; Villasenor et al. 2005), substantial evidence had accumulated in support of this hypothesis. For example, many sGRBs have a significant offset relative to the center of their host galaxy (see, e.g., Troja et al. 2008; Fong et al. 2010; Church et al. 2011; Behroozi et al. 2014); this suggests that the progenitors of these sources have migrated from their birth sites to their eventual explosion sites. Specifically, the offset distribution, together with the locations of sGRBs relative to the stellar light of their hosts, are indicative of systemic kicks (see, e.g., Berger 2014). To date, GW170817 is the strongest observational evidence for an extragalactic BNS system and the first GW signal confidently coincident with an sGRB (Abbott et al. 2017a).

In this study, we focus on constraining the immediate progenitor of GW170817 right before the second SN (SN2) that formed the BNS system. We use (i) SN-kick dynamics and kinematic modeling within the host galaxy from SN2 to merger, and (ii) the GW-measured neutron star masses, the identification of the source host galaxy, and its projected distance offset from the galactic center based on the early optical detections (Section 2). We emphasize that we develop this analysis using the very limited knowledge about the galaxy properties available in the literature prior to the announcement of the GW170817 discovery, as at this time we do not have access to the new analysis of galaxy characteristics and star formation history. We present our main results for constraints on the SN kicks, progenitor masses, pre-SN semimajor axes, and galactic radii of SN2 in Section 3, and we explore the sensitivity of our results to all our input assumptions. We find that the constraints are (i) primarily dictated by the requirement that the progenitor remains bound after SN2 and (ii) insensitive to the star formation history of the host galaxy, provided stellar ages are longer than \( \sim 1 \) Gyr. In Section 4, we use the GW BNS merger rate to estimate a BNS formation efficiency for NGC 4993, comment on the role of NGC 4993’s globular cluster content in BNS formation, and conclude our analysis.

## 2. Analysis Methodology

To investigate the constraints that can be placed on the progenitor of GW170817, we develop a modeling approach comprised of the following elements: (i) assume a gravitational-potential model for the host galaxy, described by a stellar and dark-matter (DM) density profile; (ii) place binary systems in the galaxy according to the stellar profile, and give them a pre-SN orbit in the galaxy; (iii) sample the pre-SN binary properties (pre-SN semimajor axis, progenitor mass of the second neutron star, location of SN2 within the galaxy) and the SN-kick velocity imparted on the binary following from SN2, using multiple assumptions about the underlying distribution of these parameters; (iv) sample the post-SN masses from GW parameter-estimation posterior samples of GW170817; (v) determine if the binary remains bound after SN2 and calculate the post-SN orbital properties, systemic velocity, and inspiral time, assuming two-body orbital mechanics and an instantaneous SN explosion; (vi) evolve the system forward in time, following the trajectory of the binary through the static galactic potential until it merges; (vii) select the systems with a projected offset at merger consistent with the GW170817 measurements, and label them as GW170817-like; (viii) impose constraints based on the age at which the binary formed (thus, its delay time between SN2 and merger) and the true (unprojected) distance from the galactic center, and investigate how such constraints affect our inference on progenitor properties; (ix) repeat the above steps for different input assumptions of the progenitor properties to assess the robustness of our results.

For each set of input assumptions, we evolve 50 million binaries according to the above procedures, which is sufficient to properly sample the distributions of GW170817-like systems. This section provides the model details that are adopted in our analysis.

### 2.1. Source Properties

The orbital-dynamics and kinematic analyses presented here require both GW and EM information. The post-SN orbital characteristics of a binary, such as the semimajor axis, eccentricity, and systemic velocity, depend on the component masses of the binary, which are measured in the GW inspiral. The projected offset of the binary relative to NGC 4993’s center, measured by EM observations, allows us to select GW170817-like systems in the model populations.

The best-measured property of a GW inspiral is a combination of the component masses known as the *chirp*
mass, as it determines the leading-order frequency evolution of a GW signal (Cutler & Flanagan 1994; Blanchet et al. 1995). As the binary orbit shrinks and the orbital period decreases, the GW phase becomes progressively influenced by relativistic effects that are related to the mass ratio. Due to its higher-order contribution, the mass ratio is constrained to a lesser degree than the chirp mass. The measurement of these two parameters are used to extract the component masses of the binary. GW170817 had a measured primary mass of 1.36–1.60 $M_\odot$ and a secondary mass of 1.17–1.36 $M_\odot$, using low-spin priors isotropic in orientation and with $a < 0.05$, where $a$ is the dimensionless spin parameter (see Abbott et al. 2017b for more details). Such low-spin priors are consistent with measured spins in Galactic BNS systems (Brown et al. 2012). We sample posterior distributions of these component mass measurements for each binary realization and assume that the secondary neutron star is the result of SN2.

The location of the source is measured with optical and X-ray observations to an accuracy of $\lesssim0.5''$ (Coulter et al. 2017a, 2017b; Haggard et al. 2017a, 2017b; Kasliwal et al. 2017; Yang et al. 2017; Yu et al. 2017). We combine the information in these references with the range of distances reported in NED and adopt a projected offset distance of $\approx 2.0 \pm 0.2$ kpc for our analysis.

### 2.2. Galactic Model for NGC 4993

To approximate the galactic potential of NGC 4993, we employ the Hernquist density profile (Hernquist 1990) for the stellar component and the Navarro–Frenk–White (NFW) density profile (Navarro et al. 1996) for the DM halo. We use the stellar profile for sampling the location of binaries within the galaxy, and both the stellar and DM profile for calculating the pre-SN circular galactic velocity and evolving the post-SN binaries in the combined static potential.

The Hernquist profile has a density distribution given by

$$\rho_*(r) = \frac{M_* a_{\text{bulge}}}{2\pi r (r + a_{\text{bulge}})^3},$$

where $M_*$ is the total stellar mass and $a_{\text{bulge}}$ is a scale length (Hernquist 1990). This profile satisfies de Vaucouleurs $R^{1/4}$ law, an empirical law for the luminosity as a function of radius for early-type galaxies (de Vaucouleurs 1948). Solving Poisson’s equation for the gravitational potential yields

$$\Phi_*(r) = -\frac{GM_*}{r + a_{\text{bulge}}}.$$  

The value for the scale length can be computed numerically in terms of the half-light radius ($R_{\text{eff}}$) as $a_{\text{bulge}} \approx 0.55R_{\text{eff}}$ (Hernquist 1990).

The NFW profile is one of the most commonly used profiles for representing the density distribution of DM halos:

$$\rho_{\text{DM}}(r) = \frac{\rho_0}{\frac{r}{R_c} \left(1 + \frac{r}{R_c}\right)^2},$$

where $\rho_0$ and the scale radius $R_c$ vary from halo to halo (Navarro et al. 1996). Solving Poisson’s equation leads to the gravitational potential:

$$\Phi_{\text{DM}}(r) = -\frac{4\pi G\rho_0 R_s^3}{r} \ln \left(1 + \frac{r}{R_s}\right).$$  

(4)

Given a measurement of the DM halo mass, $M_{\text{DM}}$, we assume that $M_{\text{DM}} \approx M_{200}$, where $M_{200}$ is the mass of the halo enclosed within radius $R_{200}$ at which the density of the enclosed volume is 200 times the critical density of the universe. To determine the value of the constants, we first find the concentration parameter for this volume, $c$, using the empirical expression from Duffy et al. (2008). The two constants are then calculable: $R_s$ is defined as $R_s = R_{200}/c$ and the density parameter $\rho_0$ is calculated by integrating the mass distribution up to $R_{200}$. Though the gravitational potential energy is dominated by the stellar component at small radii (see Figure 1), we use the combined potential when determining the pre-SN galactic velocity and evolving the binary post-SN: $\Phi(r) = \Phi_*(r) + \Phi_{\text{DM}}(r)$.

NGC 4993 has a stellar mass of $(10^{10.454}/h^2)M_\odot$ (Lim et al. 2017). This stellar mass is derived using K-band luminosity of the galaxy from the 2MASS Redshift Survey (Huchra et al. 2012), and the relationship between stellar mass and K-band luminosity from the EAGLE simulation (Schaye et al. 2015). For our analysis, we use the median value for the Hubble parameter from Planck Collaboration et al. (2016): $h = 0.679$. The DM halo mass for NGC 4993 is $(10^{13.7}/h)M_\odot$ (Lim et al. 2017). In addition to stellar and halo masses, we use measurements of the half-light radius of NGC 4993, $R_{\text{eff}}$, which is used in the Hernquist profile. The measured value of $R_{\text{eff}}$ for NGC 4993 is provided in galaxy surveys (e.g., Lauri & Valentijn 1989) and was recently reported as 2.8 kpc (Yu et al. 2017), indicating that the merger occurred at a projected distance of $\approx 0.71 R_{\text{eff}}$ from the NGC 4993 center. With the above information, we construct a simple model for the galactic potential of NGC 4993 to be used in our kinematic modeling.

### 2.3. Orbital Dynamics with SN Kicks

We consider the effects of the SN explosion on the orbital dynamics, assuming it is an instantaneous event which imparts a SN kick to the newly formed neutron star and a mass-loss...
kick (often referred to as a Blaauw kick; Blaauw 1961) on the companion neutron star in the binary. We ignore the effects of the first SN (SN1) on the trajectory and orbital properties of the system. The primary reason for this is that previous studies have shown that post-SN1 systemic velocities are small (50–100 km s\(^{-1}\)) compared to the galactic-motion velocities (see Figure 6 in Belczynski et al. 2002). This is due to the wide pre-SN orbits and hence low pre-SN binary orbital velocities, which regulate the post-SN systemic velocities and limit them to low values (see limits derived in Kalogera 1996). Also, any eccentricity or high orbital separation imparted by SN1 would likely be mitigated by circularization and inspiral during the common-envelope phase of the companion prior to SN2.

The post-SN orbital properties, assuming the binary has circularized prior to SN2, are derived as in Kalogera (1996):

\[
A_{\text{post}} = G(m_1 + m_2) \left[ \frac{2G(m_1 + m_2)}{A_{\text{pre}}} - V_{\text{kick}}^2 - V_{\text{rel}}^2 - 2V_k V_{\text{rel}} \right]^{-1},
\]

(5)

\[
1 - e_{\text{post}}^2 = \left( \frac{V_{\text{kick}}^2 + V_{\text{rel}}^2 + V_k^2}{G(m_1 + m_2)A_{\text{post}}} \right),
\]

(6)

where \(A_{\text{pre}}\) and \(A_{\text{post}}\) are the pre-SN and post-SN semimajor axes, \(e_{\text{post}}\) is the post-SN eccentricity, \(m_2\) is the mass of the neutron born in SN2, \(m_1\) is the mass of the companion neutron star, \(V_{\text{rel}}\) is the relative velocity between the binary components pre-SN, and \(V_k\) are the components of the SN-kick velocity \(V_{\text{kick}}\) in the frame of the binary, which is centered on the exploding star with the pre-SN objects lying along the \(x\)-axis and orbiting in the \(x\)-\(y\)-plane.

The system is initially set on a circular orbit in a random direction about the center of the model galaxy. As we show in Figure 2, pre-SN orbits are essential to include when calculating the trajectory of the binary and constraining kick velocities, as kicks tangential to the galactic orbital velocity cause a slingshot effect, which is much more efficient at propelling the binary to outer regions of the galaxy than a purely radial kick. In addition, the post-SN systemic velocity of the binary depends heavily on the mass-loss kick as well as the SN kick. Therefore, placing true constraints on the SN kick based on the offset of the merger requires knowledge of the magnitude of this mass-loss kick, which is dependent on the progenitor helium-star mass\(^{165}\) \((M_{\text{He}})\) and pre-SN semimajor axis as well as the final neutron star mass. In Figure 2, we assume an optimally oriented mass-loss kick that is parallel to the galactic velocity to show the true lower limits on the SN kick as a function of SN2 location, for multiple choices of \(M_{\text{He}}\) and \(A_{\text{pre}}\). By comparing the solid lines, we see that the lower limit of SN kicks is strongly dependent on the progenitor properties we assume. Adopting a fiducial value consistent with our constraints \((M_{\text{He}} = 3 M_\odot\) and \(A_{\text{pre}} = 2 R_\odot\)) we find that \(\geq99.95\%\) of BNS systems born within 2 kpc of the galactic center satisfy this lower limit. Furthermore, as all systems above this limit reach the offset of 2 kpc in \(\lesssim10\) Myr, which is about two orders of magnitude smaller than the typical delay time, it is necessary to continue the evolution of the binary as it explores the galaxy and possibly crosses the projected offset many times, as discussed in Section 2.5.

Following the computation of the post-SN orbital properties, the effect of the kick is added to the pre-SN systemic velocity. Due to the SN kick and mass loss, the velocity of the exploding star changes by

\[
V_2 = \left( V_{kx}, V_{ky}, \frac{m_1}{M_{\text{He}} + m_1}, V_{rel}, V_{kz} \right),
\]

(7)

where, again, \(M_{\text{He}}\) is assumed to leave behind the secondary neutron star component \(m_2\). Thus, the contribution of the kicks to the post-SN systemic velocity in the center-of-mass frame of the system becomes

\[
V_{\text{sys}} = \frac{1}{m_1 + m_2} \left( m_2 V_{kx}, m_2 V_{ky}, \frac{(M_{\text{He}} - m_2)m_1}{m_1 + M_{\text{He}}}, V_{rel}, m_2 V_{kz} \right).
\]

(8)

\(^{165}\) Just before SN2 the companion to the first neutron star is expected to be the He-rich core of a massive star, stripped of its H-rich envelope because of a prior unstable mass-transfer episode and common-envelope phase. Without such a phase, the binary orbits remain too wide for a BNS system that will merge within a Hubble time to form.
Given the pre-SN properties of the systems involved, the post-SN systemic velocities are comparable to the galactic-motion velocities (see Figure 1).

Before the systemic velocity is added to the pre-SN galactic velocity at a random angle, we check constraints on the post-SN orbital properties to ensure the system remains bound. First, we require that the post-SN orbit passes through the pre-SN positions of the masses (Flannery & van den Heuvel 1975):

$$ (1 - e_{\text{pre}}) \leq \frac{A_{\text{pre}}}{A_{\text{post}}} \leq (1 + e_{\text{post}}). $$

The mass loss and SN-kick magnitude give upper and lower bounds on the amount of orbital expansion or contraction, imposed as in Kalogera & Lorimer (2000):

$$ 2 - \frac{M_{\text{He}} + m_2}{m_1 + m_2} \left( \frac{V_{\text{kick}}}{V_{\text{rel}}} + 1 \right)^2 \leq \frac{A_{\text{pre}}}{A_{\text{post}}} \leq 2 - \frac{M_{\text{He}} + m_2}{m_1 + m_2} \left( \frac{V_{\text{kick}}}{V_{\text{rel}}} - 1 \right)^2, $$

Finally, the kick velocity is constrained from above by the requirement that the system remains bound, and from below by the minimum kick velocity needed to keep the system intact if more than half the mass of the progenitor is lost in SN2 (Kalogera & Lorimer 2000):

$$ \frac{V_{\text{kick}}}{V_{\text{rel}}} < 1 + \left( \frac{m_1 + m_2}{M_{\text{He}} + m_2} \right)^{1/2}, $$

$$ \frac{V_{\text{kick}}}{V_{\text{rel}}} > 1 - \left( \frac{m_1 + m_2}{M_{\text{He}} + m_2} \right)^{1/2}. $$

### 2.4. Distributions for Pre-SN Progenitor Properties and SN Kicks

The full 13-dimensional input space from which we sample is

$$ \Theta = [m_1, m_2, M_{\text{He}}, A_{\text{pre}}, R_{\text{gal}}, \theta_{\text{gal}}, \phi_{\text{gal}}, \Omega_{\text{gal}}, V_{\text{kick}}, \theta_k, \phi_k, \theta_{\text{sys}}, \phi_{\text{sys}}], $$

where $m_1$ and $m_2$ are sampled from the posterior parameters of GW170817 (Abbott et al. 2017b); $M_{\text{He}}$ is the progenitor helium-star mass; $R_{\text{gal}}, \theta_{\text{gal}},$ and $\phi_{\text{gal}}$ are the spherical coordinates of SN2 in the galactic frame of reference drawn from the Hernquist stellar profile; $\Omega_{\text{gal}}$ indicates the direction of motion of the system about the center of the galaxy just prior to SN2; $V_{\text{kick}}$ is the magnitude of the SN-kick velocity imparted on the newly formed neutron star; $\theta_k$ and $\phi_k$ are the angular direction of the kick relative to the plane of the binary; and $\theta_{\text{sys}}$ and $\phi_{\text{sys}}$ are the orientation of the plane of the binary with respect to the galactic coordinates. All angles are sampled isotropically in the sphere. This leaves $M_{\text{He}}, A_{\text{pre}},$ and $V_{\text{kick}},$ for which we consider various sampling procedures based on either broad assumptions or observationally motivated distributions.

The majority of constraints on SN kicks come from proper motion measurements of pulsars within our galaxy (Gott et al. 1970; Lyne & Lorimer 1994; Kaspi et al. 1996; Arzoumanian et al. 2002; Chatterjee et al. 2005; Hobbs et al. 2005; Verbunt et al. 2017). We adopt the distribution from Hobbs et al. 2005 (Hobbs) as one of our input distributions for $V_{\text{kick}}$: a Maxwellian distribution with a 1D rms $\sigma$ of 265 km s$^{-1}$. However, the mechanisms that impart SN kicks to isolated neutron stars may differ from those imparted to neutron stars that remain bound in BNS systems. There are fewer than 20 known BNS systems in the Milky Way, making inference on SN-kick properties a challenging endeavor. Nonetheless, many studies have been performed to better understand the formation process of these systems, combining the observational data with theoretical modeling (e.g., Willems & Kalogera 2004; Piran & Shaviv 2005; Stairs et al. 2006; Willems et al. 2006; Wong et al. 2010; Osłowski et al. 2011; Beniamini & Piran 2016; Tauris et al. 2017). Comprehensive analyses of observed Galactic BNS systems demonstrate that 3–4 systems

### Table 1: Table of Pertinent Parameters in Our Simulations

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Type</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_1$</td>
<td>Primary NS Mass</td>
<td>Sampled</td>
<td>GW Parameter</td>
</tr>
<tr>
<td>$m_2$</td>
<td>Secondary NS Mass</td>
<td>Sampled</td>
<td>GW Parameter</td>
</tr>
<tr>
<td>$M_{\text{He}}$</td>
<td>Helium-star Mass</td>
<td>Sampled</td>
<td>Uniform, Power Law, BP16</td>
</tr>
<tr>
<td>$A_{\text{pre}}$</td>
<td>Pre-SN Semi-major Axis</td>
<td>Sampled</td>
<td>Uniform, Log Uniform</td>
</tr>
<tr>
<td>$R_{\text{gal}}$</td>
<td>Galactic Radius of SN2</td>
<td>Sampled</td>
<td>Hernquist</td>
</tr>
<tr>
<td>$V_{\text{kick}}$</td>
<td>SN Kick Velocity</td>
<td>Sampled</td>
<td>Uniform, Hobbs, BP16</td>
</tr>
<tr>
<td>$A_{\text{post}}$</td>
<td>Post-SN semi-major Axis</td>
<td>Calculated</td>
<td>Equation (5)</td>
</tr>
<tr>
<td>$e_{\text{post}}$</td>
<td>Post-SN Eccentricity</td>
<td>Calculated</td>
<td>Equation (6)</td>
</tr>
<tr>
<td>$V_{\text{sys}}$</td>
<td>Systemic Velocity</td>
<td>Calculated</td>
<td>Equation (8)</td>
</tr>
<tr>
<td>$T_{\text{delay}}$</td>
<td>Delay Time</td>
<td>Calculated</td>
<td>Equation (14)</td>
</tr>
<tr>
<td>$R_{\text{merger}}$</td>
<td>Galactic Radius of Merger</td>
<td>Simulated</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Note. Each parameter is designated as either “Sampled,” “Calculated,” or “Simulated.” Hernquist (Hernquist 1990) is a stellar profile used for elliptical galaxies (Section 2.2). Hobbs (Hobbs et al. 2005) is a Maxwellian distribution with a scale of 265 km s$^{-1}$ (Section 2.4). BP16 (Beniamini & Piran 2016) fits log-normal models, with different best-fit parameters for low-eccentricity and high-eccentricity binaries, for distributions in $M_{\text{He}}$ and $V_{\text{kick}}$ (Section 2.4).
require small SN kicks ($\lesssim 100$ km s$^{-1}$), while another 3–4 clearly require high SN kicks ($\gtrsim 100–200$ km s$^{-1}$; Wong et al. 2010; Tauris et al. 2017). For the rest, SN-kick constraints are too broad. In addition, theoretical considerations indicate that SN kicks might be smaller for SN2 when progenitors are stripped of their envelopes (Podsiadlowski et al. 2004; van den Heuvel 2007; Janka 2013; Beniamini & Piran 2016). This may suggest bimodality in the SN-kick distribution for neutron stars in binary systems, likely based on the stage of binary evolution the system is in at the time of the SN kick (van den Heuvel 2007).

Beniamini & Piran (2016) present a two-population model for this apparent bimodality, differentiating low-kick and high-kick Galactic BNSs into two groups based on their observed eccentricity and the rotation period of the pulsar in the system. We use the best-fit parameters from this two-population model (BP16) as another kick prescription from which we sample. Beniamini & Piran (2016) also fit a mass-loss model to their two populations, which is tied to the kick model since systems with lower mass loss are expected to have a smaller shell at the time of SN2 and therefore lower SN kicks. We use this two-population model for mass loss as an input distribution for $M_{\text{He}}$, which accompanies the bimodal SN-kick model. Physically, the high-kick model corresponds to SN kicks from a Fe core-collapse SN, whereas the low-kick model is meant to emulate the population of binaries that receive electron-capture SN kicks or SN kicks as an ultra-stripped helium star. For the branching ratio between these two populations, we draw 60% of samples from the low-kick model and 40% from the high-kick model, as this is the proportion of Galactic systems that fall into each of these categories (Beniamini & Piran 2016).

Finally, we consider an input distribution in SN-kick velocities that is not informed by observations: uniform over the range [0, 2500 km s$^{-1}$] (uniform). Figure 3 shows the input distributions of the three SN-kick models described above.

In addition to the various SN-kick velocity input distributions, we consider multiple different sampling procedures for $M_{\text{He}}$ and $A_{\text{pre}}$. For $M_{\text{He}}$, we use a uniform sampling and a power law with an index of $-2.35$ (Salpeter 1955), ranging from $m_2$ (i.e., no mass loss) to the nominal black hole limit of $8 M_\odot$, along with the two-population maximum-likelihood model for mass loss from Beniamini & Piran (2016). We sample $A_{\text{pre}}$ uniform and log uniform from 0.1 $R_\odot$ to 10.0 $R_\odot$. The ranges for both progenitor masses and semimajor axes are motivated by the studies of Galactic BNS systems (e.g., Wong et al. 2010; Tauris et al. 2017).

We summarize the various parameters in our model and sampling procedures in Table 1. To gauge the impact our input distribution on progenitor constraints, we perform runs in which we alter the input distributions of $M_{\text{He}}$, $A_{\text{pre}}$, and $V_{\text{kick}}$ in various ways. We use our least constraining input distribution as our reference: uniform in $V_{\text{kick}}$, uniform in $M_{\text{He}}$, and uniform in $A_{\text{pre}}$. This reference sampling is used for our figures, unless otherwise specified.

### 2.5. Kinematic Modeling

With the above we have all necessary quantities to evolve the binary until merger. We calculate the delay time of the binary as a function of post-SN semimajor axis and eccentricity as in Peters (1964):

$$T_{\text{delay}}(a_0, e_0) = \frac{15c^2l_0^4}{304 G^2 m_1 m_2 (m_1 + m_2)} \times \int_0^{e_{\text{max}}} \frac{e^{29/19} \left[1 + (121/304) e^{21/19} \right]^{1181/2299}}{(1 - e^2)^{3/2}} \, de,$$

where $l_0$ is the orbital angular momentum. The delay time is the difference in time between the two most massive particles' last big-bang time and the merger time. In the absence of mass loss, the delay time is given by this formula. For the case of mass loss, the delay time is defined as the time difference between the two last big-bang times of the two most massive particles, which are given by $t_0$ and $t_1$, respectively. The delay time is calculated using the orbit parameters and the mass distribution at the time of SN2.
where $a_0$ and $e_0$ are the initial (post-SN) semimajor axis, and $k_0$ is determined from the initial semimajor axis and eccentricity of the system.

After the binary has evolved for time $T_{\text{delay}}$, we determine the offset of the binary from the center of the galaxy by projecting the system onto the $x-y$ plane in galactic coordinates (i.e., we assume the observer is looking at NGC 4993 down the galactic $z$-axis). If the binary ends at an offset between 1.8 and 2.2 kpc and merges in less than a Hubble time, it is considered a GW170817-like system.

We initially take a simplistic approach and assume that all binaries with delay times less than a Hubble time are valid GW170817 analogs. We then consider a full range of possible stellar-population ages for NGC 4993, from as old as the age of the universe to as young as the present. Further discussion on the star formation history of NGC 4993 is found in Section 4. We also vary the projected offset of GW170817, as if it were not known, to investigate how constraints on progenitor properties change as systems are discovered further from their host galaxies.

3. Results

Our main results comprise constraints on pre- and post-SN binary properties and SN-kick velocities, which also determine how long each binary lives between SN2 and its GW-driven inspiral and merger. In Figure 4, we show a variety of galactic orbits that potential GW170817 progenitors follow in their host galaxy, depending on post-SN properties and associated delay times. Delay times much longer than the dynamical timescale of the galaxy ($\approx$20 Myr at 2 kpc) typically lead to progenitors exploring most of the galaxy kinematically despite the merger happening relatively close to the galactic center. Shorter delay times typically lead to simple orbits of minimal structure, facilitating nearby BNS birth and merger locations, although not always (see, for example, the bottom middle panel of Figure 4).

The $T_{\text{delay}}$ times are effectively coupled to the star formation history of NGC 4993, which prior to GW170817 was not well studied. These values are indicative of how long ago SNe typically occurred, and therefore mark the ages of the most dominant stellar populations in this galaxy. In the analysis of our results, we consider a range of different $T_{\text{delay}}$ constraints and assess the sensitivity/robustness of derived constraints on progenitor properties to assumptions about the stellar age of NGC 4993, i.e., $T_{\text{delay}}$ of GW170817-like progenitors. Though the projected offset of the optical counterpart to GW170817 was well constrained, we also consider our results' robustness against this location constraint. Last, we explore different assumed distributions for the initial progenitor properties and SN kick, and we assess the robustness of our results against such changes.

The main results are presented in Figure 5, for our fiducial simulation where we assume uniform distributions for all input parameters (see Section 2.4). For the progenitor populations in the top row, we examine probability density functions (PDFs) on GW170817 progenitor properties when we impose the projected distance offset constraint of $2.0 \pm 0.2$ kpc, and different lower limits on the $T_{\text{delay}}$. It is remarkable that, provided the stellar population in NGC 4993 is older than 1 Gyr, the progenitor constraints are highly robust. We also find this insensitivity to the fine details of stellar ages to be true for our other input distributions and when we constrain $T_{\text{delay}}$ to specific ranges rather than imposing lower limits. Only if $T_{\text{delay}}$ values shorter than 1 Gyr are allowed (i.e., recent star formation has persisted in the host galaxy) are the constraints on the SN
kick and the pre-SN semimajor axis strongly affected: shorter time delays imply tighter post-SN BNS systems which allow for tighter pre-SN binaries that can remain bound even with higher SN-kick magnitudes. Delay times shorter than 1 Gyr also produce a very sharp peak in the galactic radius of SN2 around the merger distance, as the progenitor population becomes dominated by binaries that are born as BNSs relatively close to their merger site with short $T_{\text{delay}}$. To summarize, for delay times greater than 1 Gyr, the median values and 90% ranges for our reference sampling are: $\sim3.0^{+1.5}_{-1.3} M_\odot$, for the progenitor mass of the second neutron star at explosion, $\sim3.5^{+2.0}_{-1.5} R_\odot$ for the pre-SN semimajor axis, $\sim300^{+250}_{-200}$ km s$^{-1}$ for the second SN-kick magnitude, and $\sim2.0^{+4.0}_{-1.5}$ kpc for the birth radius away from the galaxy center. More detailed results examining additional parameters and parameter correlations can be found in Figure 8 in the Appendix.

In addition to SN-kick velocities, we examine constraints on the post-SN systemic velocities ($V_{\text{sys}}$). We find somewhat tighter constraints on $V_{\text{sys}}$ for GW170817-like binaries, peaking at $\sim250$ km s$^{-1}$ and with 90% of systems below $\sim400$ km s$^{-1}$ when we constrain the population to $T_{\text{delay}} \geq 1$ Gyr. Tighter constraints are to be expected as the systemic velocities are limited by the requirement that the post-SN binary remains bound; as a result, the systemic velocities saturate at values of about 1.5–2 times the pre-SN relative orbital velocities (see Kalogera 1996 for the analytical derivation of upper limits). We again find that $V_{\text{sys}}$ is robust to age constraints; the PDFs on $V_{\text{sys}}$ are practically identical provided the stellar population is $\geq 1$ Gyr.

In the middle and bottom rows of Figure 5, we examine how significant of a constraint is the knowledge of the merger’s offset from the galaxy’s center. Results in the middle row demonstrate that the primary origin of our constraints on SN kicks and progenitor properties stems from the requirement that systems remain bound after the explosion. Higher kicks, more massive helium-rich progenitors, and wider pre-SN orbits tend to disrupt a larger fraction of systems. We also find that any offset constraint at all differentiates the $R_{\text{SN}}$ distributions between SN survivors and GW170817-like systems the most: remaining bound post-SN is not affected by galactic location and without the offset, of course, progenitors follow the galaxy mass distribution. Imposing an offset constraint limits the birth radius to within a factor of typically $\sim2–3$ from the offset. The relatively small offset from the galaxy center shifts the SN kicks and helium-star masses to smaller values, effectively reducing the BNS post-SN systemic velocities, while it leaves the constraints on $A_{\text{pre}}$ unaffected.

We further explore the robustness of our results on the assumed input distributions for $V_{\text{kick}}$, $M_{\text{He}}$, and $A_{\text{pre}}$, again adopting the merger projected offset constraint of 2.0 ± 0.2 kpc. Specifically in Figure 6, we show our results for the three different SN-kick distributions (see Section 2.4). We choose...
only two cases of $T_{\text{delay}}$ constraints given the robustness of our results demonstrated in Figure 5: (i) no constraint (top row), i.e., star formation has continued in this galaxy up until the present, and (ii) $T_{\text{delay}} > 1$ Gyr (bottom row), i.e., the stellar population in NGC 4993 is older than 1 Gyr. It is evident that the constraints on $M_{\text{He}}$, $A_{\text{pre}}$, and $R_{\text{SN}}$ are robust against these different SN-kick assumptions. However, the robustness against different kick assumptions comes with the corollary that the data from this one observation is not extremely informative on the true underlying SN-kick distribution. In general, we see that GW170817 constraints exhibit mild sensitivity to the input SN-kick distributions. The uniform and Hobbs sampling procedures tend to shift to smaller SN-kick magnitudes but by relatively small amounts. The behavior with the BP16 assumption is different, but not surprising: the BP16 input distributions are extremely narrow and prescriptive, strongly dictating the allowed SN kicks and progenitor masses. The constraints for $A_{\text{pre}}$ and $R_{\text{SN}}$ are slightly stronger than those for $V_{\text{kick}}$ and $M_{\text{He}}$: $A_{\text{pre}}$ is strongly influenced by the limits placed on delay times, and $R_{\text{SN}}$ by the offset of GW170817 with respect to the galactic center.

Lastly, we have performed additional simulations with varying assumptions about the neutron star mass posteriors (Section 2.1) and the galaxy parameters (Section 2.2). To test the sensitivity of our results to neutron star mass measurements, we sampled the high-spin prior ($a < 0.89$) component mass posteriors, which have a much broader range of 1.36–2.26 $M_{\odot}$ and 0.86–1.36 $M_{\odot}$ for the primary and secondary components masses, respectively (Abbott et al. 2017b). We find quantitatively insignificant differences in our progenitor constraints. We also assess the robustness of our results against variations in the measured properties of the galaxy (i.e., stellar and DM halo masses, effective radius) and find insignificant changes for variations up to $\sim 30\%$.

We quantitatively test the robustness of our results against all assumption variations by calculating the Kullback–Leibler (KL) divergence (Kullback & Leibler 1951), described in more detail in the Appendix. The KL results, as well as the median and 90% credible intervals on all progenitor parameters and all input distributions are reported in Table 2 in the Appendix, and quantitatively justify our statements on insensitivity and robustness. The median values for the progenitor masses and semimajor axes are mostly consistent with favored values found with forward population synthesis of binary evolution (K. Breivik 2017, private communication).

Prior to SN2, the helium star may have been overflowing its Roche lobe and transferring mass to its neutron star companion. If so, the helium star could have lost significant amounts of mass ($\gtrsim 1 M_{\odot}$) prior to its explosion (Pols & Dewi 2002; Ivanova et al. 2003). To investigate the possibility of Roche-lobe overflow (RLO) at the time of the SN2, we examine whether the properties of GW170817 progenitors satisfy the conditions for RLO, adopting the analytical fit for the helium-star radius from Kalogera & Webbink (1998). Figure 7 plots successful binaries on progenitor $A_{\text{pre}}/M_{\text{He}}$ space, indicating those that would have been in an RLO phase at SN2 in green. We see that a significant fraction ($\sim 46\%$), assuming uniform input distributions as described in Section 2.4 of the GW170817 progenitor systems may have undergone RLO at the time of the BNS formation. This is not a major surprise, as it is well established by several independent studies that the double pulsar (and other known BNS systems) was also in an RLO phase at the time of SN2 (Willems & Kalogera 2004; Piran & Shaviv 2005; Stairs et al. 2006; Willems et al. 2006; Wong et al. 2010; Beniamini & Piran 2016; Tauris et al. 2017).

4. Discussion and Conclusions

In the modeling analysis presented here, we focus on constraining the immediate progenitor of GW170817, from its actual formation at the time of the second SN to the final merger. We use (i) SN-kick dynamics and kinematic modeling within the host galaxy and (ii) the GW-measured neutron star masses, the identification of the source host galaxy, and its projected distance offset from the galactic center based on the early optical discoveries. We make the most minimal/agnostic assumptions possible and avoid full, high-fidelity population synthesis models, which can account for the complex binary evolution before SN2. We explore the robustness of our results for different input assumptions.

In our analysis, we assume that the GW170817 progenitor evolved as an isolated binary in the galaxy’s field population. There are no reported results regarding observations of globular clusters (GCs) in NGC 4993, so the number of GCs in the galaxy is not known. Given that NGC 4993 could have a sizable population of GCs, a dynamical formation channel for the coalescing BNSs cannot be ruled out a priori. Typically, the number of observed GCs in a galaxy correlates with the luminosity of the galaxy (Barr et al. 2007; Harris 2016). This observed correlation can be used to estimate the number of GCs in NGC 4993. With an apparent V-band magnitude of 12.4 mag (Bellini et al. 2017) and a distance of 40 Mpc, we find an absolute V-band magnitude of $-20.6$ mag, which for an E/S0-type galaxy would correspond to $250^{+750}_{-150}$ GCs. However, since in general GCs comprise only a small fraction of the total mass of the galaxy ($\sim 0.01\%$–0.1%; Harris et al. 2015) and estimated observational merger rates for BNSs originating from
GCs are low (Grindlay et al. 2006; Ivanova et al. 2008), the isolated formation scenario is preferred. Bae et al. (2014), for example, estimated detection rates for LIGO–Virgo of 0.024 to 0.1 events per year for 1.4 $M_\odot$–1.4 $M_\odot$ BNSs coming from GCs (assuming a design-sensitivity BNS range of 200 Mpc for the LIGO detectors), which gives $\sim 10^{-6} - 10^{-2}$ events per year given the sensitivity at the time of detection ($\sim 50, 100$, and 25 Mpc for Hanford, Livingston, and Virgo, respectively). Such a low rate estimate is in contrast with the rate implied by the current GW discovery, and therefore we consider it unlikely that GW170817 was formed in a GC.

We can use the current knowledge from just this one BNS detection in GWs to extract a first estimate of the BNS formation efficiency; the fraction of massive binaries that become merging BNS systems. Our GW data analysis has yielded a measurement of the BNS rate density of 320–4740 Gpc$^{-3}$ yr$^{-1}$ (Abbott et al. 2017b), consistent with the measurements from radio-pulsar observations (e.g., Abadie et al. 2010). Given the volume density of Milky Way–like galaxies in the local universe (i.e., galaxies of comparable mass to the Milky Way) this rate measurement translates to 32–474 Myr$^{-1}$ per Milky Way equivalent galaxy (Abadie et al. 2010). NGC 4993 is of a transitional galaxy type rather than a spiral galaxy. We therefore use the galaxy’s stellar mass (instead of their blue luminosities as traditionally done for star-forming galaxies) to scale the BNS rate volume density to a rate for this specific galaxy. For the Milky Way and NGC 4993 the masses are very comparable: both have approximately 60 billion solar masses in stars (Licquia & Newman 2015; Lim et al. 2017), and therefore the BNS rate estimates based on stellar mass are roughly comparable. Assuming a Salpeter-like initial mass function (Salpeter 1955) we find that NGC 4993 has formed $\sim 4 \times 10^8$ binaries with stars whose initial masses are greater than 5 $M_\odot$. For a typical merger delay time (which dominates the lifetime of a BNS system) of 4 Gyr (see Table 2 in the Appendix) we calculate that NGC 4993’s efficiency in forming BNS merger systems per number of massive binaries is in the range of $(1–50) \times 10^{-4}$.

We model NGC 4993 with a spherically symmetric stellar and dark-matter halo profile. We note that the galaxy type is E/S0, an intermediate morphology between spiral and elliptical galaxies, and it can possibly retain a disk structure component instead of a pure spherical, radial profile (Lambas et al. 1992). Most recently, however, Im et al. (2017) show that NGC 4993 is dominated by its bulge, further supporting our assumption of a spherical gravitational potential. In addition, NGC 4993 has an axis ratio of $a/b \approx 0.9$ (Crook et al. 2007), which is consistent with a nearly spherical elliptical galaxy (though it does not exclude a face-on disk). In our analysis, we also adopt circular orbits for the galactic motion of the progenitors prior to SN2, even though there are more complex orbits allowed in realistic potentials expected for galaxies like NGC 4993 (e.g., box orbits). The key effect of including the galactic orbits is simple: the pre-SN progenitor was already in motion with orbital velocities of hundreds of km s$^{-1}$, which is comparable to the systemic post-SN velocities of the source. The specific shape of the galactic orbits or of the gravitational potential does not appear to be of particular importance, as our constraints are primarily dictated from the necessity that the system remain bound prior to SN2. This assertion is further supported by the fact that our quantitative constraints for progenitor properties are comparable to those found for BNS systems in the Milky Way where the galactic potential of a spiral galaxy is used instead.

In conclusion, we use a minimal set of observational information to constrain GW170817’s immediate progenitor, the SN kick imparted to the second neutron star, and its birth location in NGC 4993 with an appropriate galaxy model and the merger offset both informed by photometry. We obtain relatively robust constraints on the progenitor properties, albeit not always tight, strongly influenced by the requirement of keeping the binary bound after the SN and having the merger occur relatively close to the center of NGC 4993. The GW170817 progenitor constraints derived in this study are in good agreement with the progenitor constraints derived for the Galactic BNS systems as well (e.g., Wong et al. 2010; Tauris et al. 2017).

It is important to note that these constraints are essentially unchanged provided the stellar populations in NGC 4993 are older than 1 Gyr. The current literature on NGC 4993 does not provide quantitative information on the galaxy’s star formation history. Recent observations (e.g., Foley et al. 2017) might indicate some star formation activity, but as an E/S0 galaxy, it is unlikely that GW170817 was the result of very recent star formation (DeGraaff et al. 2007). Additionally, observations from Im et al. (2017) conclude that the stellar population in NGC 4993 is older than 3 Gyr. Our results strongly indicate that, for a small projected offset like that of GW170817, knowledge of the precise star formation history of the host galaxy is not vital in further constraining SN kicks and progenitor properties.

As more EM counterparts to BNS mergers are identified, we will add to the current sample of BNS systems from the Milky Way and inferred from extragalactic sGRB offset measurements to advance our constraints on progenitor properties. We note that larger projected offsets from the host-galaxy center may provide stronger constraints on the SN-kick magnitudes. In such cases, information on the age of the host-galaxy stellar population, and therefore on the BNS inspiral time, may become more useful.

The authors gratefully acknowledge the support of the United States National Science Foundation (NSF) for the construction and operation of the LIGO Laboratory and Advanced LIGO as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. The authors gratefully acknowledge the Italian Istituto Nazionale di Fisica Nucleare (INFN), the French Centre National de la Recherche Scientifique (CNRS) and the Foundation for Fundamental Research on Matter supported by the Netherlands Organisation for Scientific Research, for the construction and operation of the Virgo detector and the creation and support of the EGO consortium. The authors also gratefully acknowledge research support from these agencies as well as by the Council of Scientific and Industrial Research of India, the Department of Science and Technology, India, the Science & Engineering Research Board (SERB), India, the Ministry of Human Resource Development, India, the Spanish Agencia Estatal de Investigación, the Vicepresidência i Conselleria d’Innovació Recerca i Turisme.
and the Conselleria d’Educació i Universitat del Govern de les Illes Balears, the Conselleria d’Educació Investigació Cultura i Esport de la Generalitat Valenciana, the National Science Centre of Poland, the Swiss National Science Foundation (SNSF), the Russian Foundation for Basic Research, the Russian Science Foundation, the European Commission, the European Regional Development Funds (ERDF), the Royal Society, the Scottish Funding Council, the Scottish Universities Physics Alliance, the Hungarian Scientific Research Fund (OTKA), the Lyon Institute of Origins (LIO), the National Research, Development and Innovation Office Hungary (NKFI), the National Research Foundation of Korea, Industry Canada and the Province of Ontario through the Ministry of Economic Development and Innovation, the Natural Science and Engineering Research Council Canada, the Canadian Institute for Advanced Research, the Brazilian Ministry of Science, Technology, Innovations, and Communications, the International Center for Theoretical Physics South American Institute for Fundamental Research (ICTP-SAIFR), the Research Grants Council of Hong Kong, the National Natural Science Foundation of China (NSFC), the Leverhulme Trust, the Research Corporation, the Ministry of Science and Technology (MOST), Taiwan and the Kavli Foundation. The authors gratefully acknowledge the support of the NSF, STFC, MPS, INFN, CNRS, and the State of Niedersachsen/Germany for provision of computational resources.

Appendix
Detailed Constraints and Statistics for Progenitor Properties

We provide more detailed PDFs on progenitor properties inferred from GW170817-like systems with different delay time constraints (Figure 8) and summary statistics for output PDFs (Table 2). In Table 2, we include quantities that measure the degree to which the PDFs change when we constrain the delay times and require that the binaries match the observed offset of GW170817. This PDF comparison is done using the KL divergence (Kullback & Leibler 1951): $KL(P||Q) = \int P(x) \log \frac{P(x)}{Q(x)} dx$, where we take $Q$ to be the samples prior to applying a constraint (i.e., delay time or correct offset) and $P$ to be the samples post-application of the constraint.

![Figure 8](image.png)

Figure 8. Marginalized and joint PDFs on progenitor system properties $V_{\text{kick}}$, $M_{\text{He}}$, $A_{\text{pre}}$, $R_{\text{SN}}$, and $R_{\text{merger}}$. We restrict GW170817-like systems to various lower limits for $T_{\text{delay}}$. The black points show the full population of binaries that correspond to the measured offset of GW170817 (i.e., they have no constraint on $T_{\text{delay}}$). As delay times become $\gtrsim 1$ Gyr, the constraints on GW170817-like samples are significantly tightened, in particular, removing systems with low $A_{\text{pre}}$ and thus short inspiral times, systems with extremely high SN-kick velocities, and systems that are born as BNSs and quickly merge right at the offset of GW170817. The diagonal line in the joint $R_{\text{merger}}-R_{\text{SN}}$ PDF, for example, is an artifact of extremely short inspiral times leading to BNS systems merging at the same location as the second supernova.
Specifically, the KL divergence measures the information gained by updating a prior $Q$ to a posterior $P$. The values are computed by histogramming the samples from $Q$ to approximate $q(x)$ and using the same bin locations to make a histogram of $p(x)$. The integral for KL divergence then becomes analytic.

### References

Acernese, F., Agathos, M., Agatsuma, K., et al. 2015, CGJra, 32, 024001
Arcavi, I., Hossinezadeghi, G., Howell, D., et al. 2017b, Nat, 551, 64
Blauw, A. 1961, BAN, 15, 265

Note. Reported values are the median and 90% confidence interval. The letters in the first three columns indicate the input sampling used: uniform (U), Hobbs (H), log uniform (L), power law (PL), and BP16 (BP); see Section 2.4 for more details. KL divergence scores (in units of Nats) are reported in the fourth-rightmost columns, quantifying information gained by imposing constraints (and by proxy, how much the PDFs change). "Offset" quantifies the amount of information learned by taking all binaries that survive the supernova and imposing the constraint that they merge at the correct projected offset, using the $(V_\text{kick}, M_\text{BH})$ joint PDF. The remaining three KL values take all post-SN binaries with the correct offset and sampling method indicated by the first three columns of a given row and restrict to those with a $T_{\text{delay}}$ range specified by that row. These compare the 2D PDFs indicated by the column headers. The rows with $T_{\text{delay}} \in [0, 14]$ Gyr have no age constraints, and thus their age-restricted KL divergence will always be zero (they are being compared to themselves). We have labeled these as "ref" instead of zero, to make this clear. As a rough rule of thumb, KL values $\lesssim 0.1$ correspond to small differences in the distributions, $\sim 0.4-0.6$ to modest differences, and $\gtrsim 1.0$ to quite large differences. In general, we find that we learn slightly more by imposing the offset constraint if the age constraint is tighter. We also consistently get higher KL values when $\Lambda_{\text{pre}}$ is included in the analysis, which is due to its strong correlation with $T_{\text{delay}}$. Input distributions are described in detail in Section 2.4.