The Gravitational waves merger time distribution of binary neutron star systems

Paz Beniamini¹,²* & Tsvi Piran³

¹Department of Physics, The George Washington University, Washington, DC 20052, USA
²Astronomy, Physics and Statistics Institute of Sciences (APSIS)
³Racah Institute of Physics, the Hebrew University, 91904, Jerusalem, Israel

ABSTRACT

Binary neutron stars (BNS) mergers are prime sites for r-process nucleosynthesis. As such their rate determines the chemical evolution of heavy elements in the Milky Way. The merger rate of BNS is a convolution of their birth rate and the gravitational radiation spiral-in delay time. Using the observed population of Galactic BNS we show here that at late times (t ≥ 1 Gyr) the gravitational wave delay time distribution (DTD) follows the expected t⁻¹. However, a significant excess of rapidly merging systems (between 40 – 60% of the entire population) is apparent at shorter times. This population of rapid mergers implies a declining deposition rate of r-process materials that is consistent with several independent observations of heavy element abundances in the Milky Way. At the same time this population that requires initial binary separations of roughly one solar radius clearly indicates that these binaries had common envelope progenitors. Our results suggest that a significant fraction of future LIGO/Virgo BNS mergers would reside in star forming galaxies.

Key words: gravitational waves – stars: neutron – Galaxy: abundances – stars: evolution

1 INTRODUCTION

The merger of binary neutron star (BNS) systems results in a blast of gravitational wave (GW) radiation, in a short gamma-ray burst (GRB) and in synthesis of r-process elements at a rate that can account for the overall formation of those elements in the Universe. The connection between these different aspects of BNS mergers has recently been given very strong support due to the detection of GW from GW170817 which were accompanied by a GRB (albeit an unusually weak one, see e.g. Kasliwal et al. 2017, Matsumoto, Nakar & Piran 2019), a macronova (powered by the radioactive decay of r-process elements) and a multiwavelength afterglow. Interestingly the implied rate and yield of heavy r-process element is consistent with that required for the production of heavy r-process elements in the Galaxy (see e.g. Hotokezaka, Beniamini & Piran 2018 for a recent review).

The rate of BNS mergers depends on the birth rate of these systems as well as on the delay times from birth to merger, commonly known as the ‘delay time distribution’ (DTD). As neutron stars’ progenitors are massive stars their birth rate follows the star formation rate (SFR) with a minimal delay. The DTD is dominated by this GW spiral-in time. As a BNS loses energy due to GW radiation its orbit shrinks and eventually the neutron stars merge. The time until merger t_m ∝ a³(1 – e²)⁷/₂ depends on the initial semi major axis, a, and the eccentricity, e, of the BNS. As such the determination of the DTD and in particular of the minimal time delay, t_m, will provide valuable information on the unknown final stages of evolution of BNS progenitors. Here we explore the DTD using the observed Galactic population of BNS. This approach is complementary to attempts to explore the DTD by comparing the rate of short GRBs (believed to be associated with BNS mergers) and the SFR (see e.g. Guetta & Piran 2005, 2006, Guetta & Stella 2009, Leibler & Berger 2010, Dietz 2011, Coward et al. 2012, Wanderman & Piran 2015). An independent single observation arises from GW170817. The host galaxy of this event is an S0-type galaxy with a very low star formation rate, suggesting a time delay between the BNS formation and merger of ≈ 1 – 10 Gyr.

The GW merger time, t_m, depends strongly on a, and since many BNS exhibit low orbital eccentricities, it is insightful to consider the expected DTD resulting from an initial PL distribution of semi major axes, dN/da ∝ a⁻ⁿ with circular orbits (e = 0). The result is dN/dt_m ∝ t_m⁻ⁿ⁻⁶. In particular, if the semi-major axis is uniformly distributed in log-space, dN/dlog a = const, one gets the well known result dN/dt_m ∝ t_m⁻¹ (see e.g. Piran 1992, Yungelson & Livio 2000, Guetta & Piran 2005, 2006, Totani et al. 2008, Maoz, Mannucci & Nelemans 2014). Note also that due to the strong dependence of t_m on a, the distribution of semi-major axes must be very steep in order to get significantly steeper dN/dt_m profiles.

We use the observed distribution of merger times of the Galactic BNS to derive the DTD in a model-independent way, directly...
from observations. The paper is structured as follows. In §2 we describe the observed sample. We then study analytically the connection between the distribution of delay times at birth and the merger time (and age) distribution viewed at some later time in §3. In §4 we perform a likelihood analysis in order to determine the underlying DTD from the observed population. We examine some implications on the total number of un-merged systems and observable pulsars remaining in BNS in the Galaxy today in §5. We compare our results with other findings in §6 and we summarize and discuss some broader implications in §7.

2 GALACTIC BNS OBSERVATIONS

Our sample includes the 15 observed field Galactic BNS systems. We exclude two Galactic BNS that reside in globular clusters. Table 1 provides the semi-major axes, eccentricities, merger times and spin-down times of these systems. The observed merger time distribution is shown in figure 1.

Clearly this sample does not include all Galactic BNSs, as the observed systems are only those that have an active pulsar pointing towards us. Since Pulsars live a finite time there are many systems that cannot be observed. The BNS pulsars are mostly recycled ones (Beniamini & Piran 2016; Tauris et al. 2017) and their true ages may be considerably younger than their characteristic spin-down times, $P/2\dot{P}$. (Faucher-Giguère & Kaspi 2006; Kiziltan & Thorsett 2010; Oslowski et al. 2011). The latter provide only upper limits on the true ages. For example, in the double pulsar system, J0737-3039, the two pulsars (which are expected to be of a similar age) have spin-down ages of 210 and 50 Myr. This is much longer than what is expected to be their true age difference.

We consider as the true age of the system the time since the formation of the second neutron star.

### Table 1. Parameters of Galactic field binary neutron stars.

| System | $a \times 10^{11}$ cm | $e$ | $t_{\text{ms}}$ Gyr | $P/2\dot{P}$ Gyr | ref. |
|--------|----------------------|-----|------------------|----------------|-----|
| J1946+2052 | 0.73 | 0.064 | 0.05 | 0.3 | 16 |
| J1757-1854 | 1.32 | 0.605 | 0.07 | 0.13 | 14 |
| J0737-3039 | 0.88 | 0.088 | 0.086 | 0.0500.21 | 1 |
| J1906+0746 | 1.22 | 0.085 | 0.31 | $10^{-4}$ | 2.3 |
| B1913+16 | 1.95 | 0.617 | 0.38 | 0.086 | 6 |
| J1913+1102 | 1.45 | 0.089 | 0.47 | 2.7 | 13 |
| J0509+3801 | 2.16 | 0.59 | 0.69 | 0.15 | 17 |
| J1756-2251 | 1.87 | 0.181 | 1.61 | 0.44 | 4.5 |
| J1829+2456 | 4.48 | 0.139 | 55.4 | 12.4 | 8 |
| J1411-2551 | 7.58 | 0.169 | 460 | 10.4 | 15 |
| J0453+1559 | 10.4 | 0.112 | 1400 | 2.6 | 11 |
| J1811-1736 | 28.3 | 0.828 | 6400 | 1.8 | 10 |
| J1518+4904 | 17.2 | 0.249 | 8900 | 10.4 | 9 |
| J1930-1852 | 50.9 | 0.4 | $5 \times 10^3$ | 0.16 | 12 |

[9] References (ordered by table numbers): Kramer et al. (2006); Lorimer et al. (2006); van Leeuwen et al. (2015); Faulkner et al. (2005); Fordman et al. (2012); Weisberg, Nice & Taylor (2010); Stairs et al. (2002); Champion et al. (2006); van Leeuwen et al. (2015); Faulkner et al. (2005); Ferdman et al. (2014); Janssen et al. (2008); Corongiu et al. (2007); Swiggum et al. (2015); Martinez et al. (2015); Lazarus et al. (2016); Martinez et al. (2017); Stovall et al. (2018).

Figure 1. Merger time distributions of observed BNS as compared with a delay time distribution $D(t) \propto t^{-1}$ between $t_{\text{min}} = 0.035$ Gyr and $t_{\text{max}} = 10^6$ Gyr with different values of the maximum pulsar age, $t_a$. A statistically significant excess of systems with merger times $\leq 1$ Gyr is required by observations, whereas the finite age effect can only decrease the number of short merger time systems.

3 GENERAL CONSIDERATIONS

3.1 The delay time distribution - DTD

The merger time distribution, $dN/dt$, is a convolution of the intrinsic DTD, $D(t)$, with the BNS birth rate, $r(t)$, where time here is measured backwards from the present day (i.e. today we have $t = 0$ and at earlier times, $t$ is larger):

$$\frac{dN}{dt} = \int_0^{t_H} r(t') D(t' + t) dt',$$

where $t_H$ is the Hubble time. We will consider DTDs that have a minimal $t_{\text{min}}$ and maximal $t_{\text{max}}$ range.

The distribution $dN/dt$ is different from the observed merger time distribution $dN_{\text{obs}}/dt$ since, as mentioned earlier, the observed distribution depends on detection of the corresponding pulsars and those have a finite lifetime. To account for the finite lifetimes of the pulsars in the BNS systems we introduce a typical lifetime $t_a$ in which one of the pulsars in the BNS system can still be observed. In reality, there is a distribution of lifetimes. As there is not enough information today to reveal this distribution we will take $t_a$ to be a constant. If there is a distribution of lifetimes than $t_{a,\text{min}} < t_a < t_{a,\text{max}}$ the transition between the regimes discussed below will be gradual, but the basic qualitative features will remain. The regime $t << t_a (t \gg t_a)$ should then be interpreted as $t \ll t_{a,\text{min}} (t \gg t_{a,\text{max}}).$ We show later (see §3.3) that the value of $t_a$ is likely small compared to the typical delay times. However, for the sake of generality, we make no assumptions on the value of $t_a$ in the following discussion.

We write the number of observed BNS with a given time to merger, $dN_{\text{obs}}/dt$, as:

$$\frac{dN_{\text{obs}}}{dt} = \int_0^{t_{\text{max}}} r(t') D(t' + t) dt'.$$

For simplicity, we use here and in the following a constant Galactic star formation rate, $r(t) = r_0$. This simple assumption has been suggested in the literature as a good approximation for the star formation rate of the Milky Way (MW) during the majority of its lifetime (see e.g. Snaith et al. 2015). For $r(t) = r_0$ and for $t_a < t_{\text{min}}$.
the observed merger distribution is:

$$\frac{dN_{\text{obs}}}{dt} \propto \left\{ \begin{array}{ll} 0 & \text{for } t < t_{\text{min}} - t_a \\ \text{const} & \text{for } t_{\text{min}} - t_a < t < t_{\text{min}} \\ D(t) & \text{for } t > t_{\text{min}}. \end{array} \right. \tag{3}$$

Note that for $t_a < t_{\text{min}} < t$ the observed distribution is simply proportional to the DTD: $dN_{\text{obs}}/dt \propto D(t)$ regardless of the functional form of $D(t)$. Consider now the opposite limit, $t_a > t_{\text{min}}$ (assuming once more a constant formation rate). For $t > t_a$, that is for merger times much longer than the age of the systems, the pulsars stop shining on a time scale much shorter than their merger time. Hence, $dN_{\text{obs}}/dt \propto D(t)$ as before. But, of course, there are many more ($t_{\text{obs}}/t_a$) BNS systems that we cannot observe. Alternatively, at much shorter times, $t < t_a$ the observed population is reduced compared to an extrapolation from the $t \gg t_a$ regime since only a fraction $t/t_a$ of the systems born in the last $t_a$ years with a delay time $t$ have not merged yet. $dN_{\text{obs}}/dt$ is therefore proportional to $(t/t_a)D(t)$. At delay times even lower than $t_{\text{min}}$, the only systems observable are the tail of the systems born with a delay of $\sim t_{\text{min}}$ (which by construction are the type of systems that dominate the intrinsic population) that were formed $t_{\text{min}} - t$ years ago. For $t < t_{\text{min}}$ the two latter factors are constant, and thus $dN_{\text{obs}}/dt$ is constant.

Consider, for example, a power-law DTD:

$$D(t) = \frac{1-b}{t_{\text{max}} - t_{\text{min}}} t^{-b} \quad \text{for } t_{\text{min}} \leq t \leq t_{\text{max}}. \tag{4}$$

For an index $b > 1$ and $t_a > t_{\text{min}}$ we obtain

$$\frac{dN_{\text{obs}}}{dt} \propto \left\{ \begin{array}{ll} \text{const} & t < t_{\text{min}} \\ t^{-1-b} & t_{\text{min}} \ll t < t_a \\ t^{-b} & t_a \ll t < t_{\text{max}}. \end{array} \right. \tag{5}$$

Overall we find that the finite lifetime of pulsars imply that for merger durations longer than $t_a$ the observed merger time distribution simply follow the DTD. For durations shorter than $t_a$ we observe fewer systems than what would the DTD predict.

### 3.2 The Galactic BNS distribution

The observed distribution of Galactic BNS merger times is shown in figure 1 at times much larger than the typical ages of the pulsar, $dN_{\text{obs}}/dt \propto D(t)$, i.e. it is directly related to the DTD. It is evident from figure 1 that at $t \geq 1$ Gyr, $dN/d\log t$ is approximately constant, suggesting $D(t) \propto t^{-1}$. However, at lower merger times we find an observed excess of systems with $0.05$ Gyr $\leq t_{\text{min}} \leq 1$ Gyr (seven out of fifteen systems, representing a fraction $Z_{<1\text{Gyr}} = 0.47$ of the population), as compared with the $D(t) \propto t^{-1}$ model predictions (for which $Z_{<1\text{Gyr}} = 0.19$, as we discuss later in § 4.2). At the same time, as we argued above, the finite lifetime of the pulsars leads instead to a suppression of the number of systems with $t_{\text{min}} < t_a$.

$D(t) \propto t^{-1}$ is a good description of the observations for $t > 1$ Gyr. The excess of short duration mergers (relative to predictions of the $D(t) \propto t^{-1}$ model) at $t_{\text{min}} \leq 1$ Gyr must be intrinsic feature of the DTD. Note that the possibility that this excess is due to an increase of the BNS formation rate in the last 1 Gyr is highly unlikely. This time-scale is much smaller than the scale on which significant Galactic changes took place.

### 3.3 Ages of observed BNS systems

As mentioned in [2], the ages of BNS systems are less or equal than their spin-down times. The distribution of observed ages of BNS systems, is naturally affected by the DTD (at older ages, a larger fraction of systems have already merged). The number of systems with age $t_a$ is given by

$$\frac{dN}{dt_{\text{age}}} = r(t_{\text{age}}) \int_{t_{\text{age}}}^{\infty} D(t)dt$$

where we have assumed in the second line a constant formation rate and a power-law DTD with $b > 1$.

At the same time, as we argued above, the finite lifetime of the pulsars leads instead to a suppression of the number of systems with $t_{\text{min}} < t_a$. $D(t) \propto t^{-1}$ is a good description of the observations for $t > 1$ Gyr. The excess of short duration mergers (relative to predictions of the $D(t) \propto t^{-1}$ model) at $t_{\text{min}} \leq 1$ Gyr must be intrinsic feature of the DTD. Note that the possibility that this excess is due to an increase of the BNS formation rate in the last 1 Gyr is highly unlikely. This time-scale is much smaller than the scale on which significant Galactic changes took place.

$D(t) \propto t^{-1}$ is a good description of the observations for $t > 1$ Gyr. The excess of short duration mergers (relative to predictions of the $D(t) \propto t^{-1}$ model) at $t_{\text{min}} \leq 1$ Gyr must be intrinsic feature of the DTD. Note that the possibility that this excess is due to an increase of the BNS formation rate in the last 1 Gyr is highly unlikely. This time-scale is much smaller than the scale on which significant Galactic changes took place.

Figure 2. Merger time distributions of observed BNS as compared with four different delay time distribution models. In all cases we take $t_a < t_{\text{min}}$ and DTD model parameters as described in [4]. A statistically significant excess of systems with merger times $< 1$ Gyr is required by observations. Top: the differential distribution, Bottom: the integrated distribution.
servation bias according to which older system may be harder to
detect this ratio is inconsistent with the observations.

To explore the implication of this inconsistency we consider the
P/2P values in table 1 as upper limits on the true ages of the
systems. We assume that the systems live for a fixed time \( t_a \) and
search for a value of \( t_a \) for which a simulated set of \( N \) ages (where
\( N \) is the number of observed BNS systems) is consistent with these
upper limits. For 0.01 Gyr \( \leq t_{\text{min}} \leq 0.05 \) Gyr and 1 \( \leq b \leq 1.5 \), as
suggested by the observed merger times we find \( t_a < 0.03 \) Gyr at a
3\( \sigma \) confidence level. This result, which depends weakly on the
choice of \( t_{\text{min}} \) and \( b \) depends very strongly on the very short \( P/2P \)
of J1906+0746. However, even if the spin-down age of the pulsar
in J1906+0746 is disregarded as an anomaly, the remaining distribu-
tion of spin-down ages, still constrains \( t_a < 0.3 \) Gyr at a 2\( \sigma \)
level. That is even if \( t_a > t_{\text{min}} \), the dynamical range between
the two cannot be large. Therefore, the effect of \( t_a \) on the observed
distribution must be relatively small. This conclusion is consistent
with the estimate provided in §3.1 based on the number of observed
Galactic BNS systems in comparison with estimates of the BNS
formation rate. As discussed in §3.1 this means that the observed
distribution of merger times roughly traces the observed ones, and
that the exact value of \( t_a \) is not required to deduce the shape of the
DTD.

Following these results we assumed in the following that
\( t_{\text{age}} < t_a \). This means that the orbits of the majority of systems
don’t evolve appreciably during the time that their pulsars remain
observable and hence the observed merger time distribution reflects
the intrinsic one. We also verify, later (see §4.5), using a maximum
likelihood analysis this assumption.

4 LIKELIHOOD ANALYSIS

We consider several models that can be fitted to the observed
merger time distribution. Our models involve one or two para-
eters (denoted here as \( X_i \) for the more general case) that describe
the DTD. We search for the values of \( X_i \) that maximize the likeli-
hood of obtaining the observed distribution of Galactic BNS merger
times. To do so, we construct a likelihood function as follows
\[
L(X_i) = \prod_j P(t_{\text{obs,}j}|X_i)
\]
where \( j \) goes over the observed systems in the sample, \( P(t_{\text{obs,}j}|X_i) \)
is the probability that a system was born with a delay time \( t_{\text{obs,}j} \)
given the delay time distribution specified by \( X_i \).

4.1 A \( t^{-1} \) delay time distribution

Our first choice for the DTD is a simple power-law \( t^{-1} \). This is
motivated, as mentioned in §2.5, by the strong dependence of \( t_a \) on
the initial BNS separation (see e.g. Piran 1992; Yungelson & Livio
2000; Guetta & Piran 2005; 2006; Totani et al. 2008; Mazzu, Man-
ucci & Nelemans 2014). Moreover, an inspection by eye of the
long merger time part of the distribution (which is not influenced
by pulsar’s lifetimes and the choice of the edges of the distribu-
tion) suggests that this model is consistent at least for long merger
durations. The observed delay times depend only weakly on \( t_{\text{max}} \)
which we fix here to be \( t_{\text{max}} = 10^6 \) Gyr. The DTD is in this case
a function of a single parameter \( t_{\text{min}} \). The likelihood is maximized
for \( t_{\text{min}} = 0.035 \) Gyr. Using a KS test, we compare the observed
merger time distributions with the model. The two distributions are
shown in figure 2. The KS test rules out the \( t^{-1} \) model at a 2\( \sigma \)
confidence level. As mentioned in §3.1 there is a much larger fraction
of short mergers in observations (\( Z_{\text{3Gyr,1Gyr}} = 0.47 \)) than expected for
the best \( D(t) \propto t^{-1} \) (\( Z_{\text{3Gyr,1Gyr}} = 0.19 \)). This is the main source of
discrepancy between this model and observations. It is important
to stress that, as we have seen in §3.1 introducing longer lifetimes
will only exacerbate the problem as this leads to a decrease in the
number of observed systems with short merger times (see figure 1).
Given this discrepancy we explore in the following several DTD
models that involve a second free parameter.

4.2 A general power-law delay time distribution

The observed merger times clearly span many orders of magnitude
(see table 1). The simplest functional form spanning a wide range
is a power-law. A natural generalization is to use another power-
law \( t^{-b} \) and let \( b \neq 1 \) a free parameter in addition to \( t_{\text{min}} \).
Due to the vast range between \( t_{\text{min}} \) and \( t_{\text{max}} \) (at least seven orders
of magnitude; see table 1), even a small change in \( b \) can have a
significant effect on the observed distribution. Indeed taking \( b = 1.1 \),
\( t_{\text{min}} = 0.035 \) Gyr (and \( t_a < t_{\text{min}} \)) results in a distribution
denoted in figure 2 as \( D_2(t) \) that is consistent with the current
data. The main importance of this possibility is the simplicity of
implementing it in analytic expressions, such as in [3,5].

4.3 A short delay time population

At long merger times \( \gtrsim 1 \) Gyr, the DTD is well described by
\( D(t) \propto t^{-1} \). Motivated by the excess of events with short merger
times, we add a short merger time log-normal component to the \( t^{-1} \)
power-law:
\[
D_1(t) = At^{-b} + B \lognormal(\mu, 1) \text{ for } t_{\text{min}} < t < t_{\text{max}} \, . \tag{8}
\]

The second term is a log-normal distribution with \( \sigma = 1 \) (implying
a width that is of the same order of magnitude as the median). Based
on the observed sample and on the modeling for the \( t^{-1} \) and general
PL DTDs we fix here \( t_{\text{min}} = 0.035 \) Gyr, \( t_{\text{max}} = 10^6 \) Gyr. This model
involves two free parameters: \( \mu, f \), where \( f \) is the ratio of events
resulting from the log-normal distribution out of the total number
of events (and is proportional to \( B/(A + B) \)).

The likelihood function for \( D_1(t) \) is maximized at \( \exp(\mu) = 0.3 \) Gyr
and \( f = 0.4 \). It is depicted in figure 3. The resulting DTD
is now statistically consistent with the data (see also figure 2). The
model is insensitive to the exact value of \( t_{\text{min}} \). Values in the range
0.001 Gyr \( \leq t_{\text{min}} \leq 0.1 \) Gyr are all consistent with observations.

Given the small number of observed systems, the functional
form of the DTD cannot be uniquely determined by the available
data. As an illustration, we consider a broken power-law:
\[
D_2(t) = C \begin{cases} [t/t_0]^{-a} & \text{for } t_{\text{min}} < t < t_0 \ , \\ [t/t_0]^{-1} & \text{for } t_0 < t < t_{\text{max}} \ . \end{cases} \tag{9}
\]

Taking as before \( t_{\text{min}} = 0.035 \) Gyr, \( t_{\text{max}} = 10^6 \) Gyr. The free
parameters of \( D_2(t) \) are \( a, t_0 \). This model is also consistent. The
parameters \( t_0 = 15 \) Gyr, \( a = 1.3 \), \( t_{\text{min}} = 0.035 \) Gyr provide a
statistically consistent fit to the data. Allowing for varying values
of \( t_{\text{min}} \) we find that consistent solutions can be found so long as
0.02 Gyr \( \leq t_{\text{min}} \leq 0.15 \) Gyr.

© 2019 RAS, MNRAS 000, 000–000

2 The results hold as long as there is no significant drop in the age distri-

bution given by Eq. 6 above for age.
4.4 Comparison between the models

To compare $D_1(t), D_2(t), D_3(t)$ to $D \propto t^{-1}$ we perform a likelihood ratio (LR) test. The LR is defined as

$$ LR = 2 \log \left( \frac{L_{max,a}}{L_{max,b}} \right) $$

where $L_{max,a}, L_{max,b}$ are respectively the maximum likelihoods of some model $a$ and of the null hypothesis (model 0). LR follows a $\chi^2$-squared distribution where the number of degrees of freedom is given by the difference in free parameters between the two models (2 between $D_1, D_2$ and $D \propto t^{-1}$ and 1 between $D_3$ and $D \propto t^{-1}$). For all three models, the improvement as compared with a $t^{-1}$ DTD is statistically significant (with a $p$-value for chance coincidence of $p < 0.05$). Inspection by eye of figure 4 suggests that the model with an extra short-time log-normal population looks best. However, this is not statistically significant and it depends on the binning of the data. A comparison of non-nested models with each other (i.e. comparing any two of $D_1, D_2, D_3$) is impossible with the regular LR test and usually requires a larger data-set to be meaningful. However, a shared property of all these consistent models is that they all result in a large fraction, $Z_{\leq 1Gyr} \gtrsim 0.4$, of BNS with merger times of less than 1 Gyr.

4.5 The pulsar’s lifetime

Before concluding, this section, we return once more to the question of the pulsar’s lifetime. To address it we have carried out a likelihood analysis including $t_a$ as a free parameter. Figure 4 depicts the likelihood projections for model $D_3$ but now with $t_a$ as a free parameter. Note that in this analysis we didn’t impose the conditions $t_{age} < P/2P$, as with the very short $P/2P$ of 11906±0746 this would have immediately imposed a very small $t_a$ (see figure 3). Still even without this constraint we find (see figure 4) that $t_a < t_{min}$ is preferred, where the best fit value is $\lesssim 0.02$ Gyr and where in any case the region $t_a > 0.1$ Gyr is practically excluded.

5 TOTAL NUMBER OF SURVIVING AND MERGED GALACTIC BNS

We turn now to explore the unobserved populations: BNS systems that cannot be observed because their pulsars have died or are not pointing towards us and BNS systems that have already merged. The fraction of merged systems depends on the formation rate history of BNS in the MW, and on the DTD:

$$ \Psi_{\text{merg}} = \frac{1}{\tau_{\text{ff}}^b} \int_0^{\tau_{\text{ff}}} r(t) \int_0^{t_H} r(t_d) P(t_d < t) dt \, dt $$

where $P(t_d < t) = \int_0^{t_H} r(t_d) dt$ is the probability of a system born $t$ years ago to have merged by today. A specific case of interest is that of a constant rate $r(t) = \int_0^{t_H} r(t) dt / \tau_{\text{ff}} \equiv r_0$. In this case, for a power-law distribution we have:

$$ \Psi_{\text{merg}} = \frac{t_{\text{H}}^{b-1} + (1+b)t_{\text{H}}^{b-1} - (2-b)t A^{t_{\text{H}}^{b-1}-t_{\text{min}}^{b-1}}}{(2-b)t H(t_{\text{H}}^{b-1} - t_{\text{H}}^{b-1})} $$

$$ \approx 1 - \left( \frac{t_{\text{min}}}{t_{\text{H}}} \right)^{b-1} \approx 0.5 $$

where in the last expression we have used the results for $D_3(t)$ from figure 3. We find a similar fraction of merged systems, $\Psi_{\text{merg}}(D_1) \approx 0.5$, for the distributions, $D_1(t), D_2(t)$. We conclude that about half of the BNS systems formed in the history of the Galaxy have already merged. Using a BNS formation rate within the MW of $\sim 0.04$ Gyr$^{-1}$ [Kochanek & Piran 1993, Burgay et al. 2003, Kalogera et al. 2004, Kim, Perera & McLaughlin 2015, Wanderman & Piran 2015, Belczynski et al. 2018], this translates to a total of $2.7 \times 10^5$ un-merged BNS systems in the MW.

We next turn to estimate the total number of BNS with active pulsars that reside in the Galaxy today. As before we take the formation rate to be constant up to $t_a$ (see figure 3). The total number of active systems in the MW is then estimated as

$$ N_{\text{active}} = \int_0^{t_a} r(t) P(t_d > t) dt \approx $$

$$ \int_0^{t_a} \begin{cases} 1 & t_a < t_{\text{min}} \\ (t_a/t_{\text{min}})^{1-b} & (2-b) t_a > t_{\text{min}} \end{cases} $$

Figure 3. Log-likelihood function for a delay time distribution given by $D(t)$ (see figure 2). The likelihood is shown as a function of $f, \exp(\mu)$. The maximum is depicted by a black circle for clarity.

Figure 4. Projections of the 3D Log-likelihood plots in the case of a power-law DTD, $D$, where the free parameters are $t_a, t_{\text{min}}, b$ and $t_a$ the maximal age of detectable pulsars. The solid line in the bottom left figure marks $t_a = t_{\text{min}}$. One can see that $t_a < t_{\text{min}}$ is preferred by the data, even without taking into account the short spin-down times of the observed pulsars. The likelihood is maximized for $b = 1, t_{\text{min}} = 0.035$ Gyr, $t_a = 0.005$ Gyr.
The observable number of systems in the MW is smaller than the number of active ones due to two factors. The first is $\Phi_{\text{beam}}$. This is the correction that arises since pulsars are detected by their beamed radio emission. For each pulsar with a period $P$ and temporal beam width $W$, there are approximately $P/2W$ actual pulsars in the Galaxy of the same type. Using $W_{50}$ as the beam width (where $W_{50}$ is the width of the pulsar’s pulses at a 50% level), we find that for our sample $\Phi_{\text{beam}} \approx (P/(2W_{50}))^{-1} = 0.1$. The other reduction is due to the incompleteness of the BNS pulsar population, $\Phi_{\text{inc}}$. For example, Keane et al. (2015) have predicted that SKA could increase the number of known radio pulsars by a factor of $\approx 5 - 10$, implying $\Phi_{\text{inc}} < 0.2$. Overall, the observed current number of BNS in the Galaxy, $N_{\text{obs}}$, implies $N_{\text{active}} = N_{\text{obs}} \Phi_{\text{beam}}^{-1} \Phi_{\text{inc}}^{-1} \sim 10^3$. These estimates are consistent with estimates by Kalogera, Kim & Lorimer (2003) that calculated there should be $\approx 300$ Galactic systems similar to each B1534+12 and B1913+16. We can now estimate $t_a$.

$$t_a \approx \frac{N_{\text{obs}}}{\Phi_{\text{beam}} \Phi_{\text{inc}} r} = 0.015 \text{ Gyr} \left( \frac{0.1}{\Phi_{\text{beam}}} \right) \left( \frac{0.03}{\Phi_{\text{inc}}} \right) \left( \frac{0.04}{r \text{ Gyr}} \right)$$

This is consistent with the condition $t_a \lesssim t_{\text{min}}$ suggested by the distribution of spin-down ages (see §3.3).

6 COMPARISON WITH OTHER RESULTS

Our main finding is an excess of rapid mergers over the expected $t^{-1}$ delay time distribution. It turns out that there are other venues that lead to this conclusion. An excess of rapid mergers has also been suggested on completely separate grounds following from observations of r-process element abundances. Specifically, the high abundance of r-process elements in the ultra-faint dwarf (UFD) galaxy Ret-II as compared to the upper limits on those abundances in other UFDs suggests a neutron star merger origin (Beniamini, Hotokezaka & Piran 2016b). At the same time, UFDs are known to have ceased their star formation within the first Gyr of their formation (Brown et al. 2014; Weiss et al. 2015) requiring a rapid merger to enrich the gas before it can be reprocessed into the stars seen in those galaxies today (Beniamini, Hotokezaka & Piran 2016a). More generally, if UFD-like galaxies form the main building blocks from which the MW halo population is composed, then rapid mergers are required to explain the high values of r-process abundance observed in extremely metal poor stars (Argast et al. 2004; Tsujimoto & Shigeyama 2014; Wehmeyer, Pignatari & Thielemann 2015; Beniamini, Hotokezaka & Piran 2016a; Beniamini, Dvorkin & Silk 2018). Furthermore, a significant population of rapid mergers, coupled with a declining star formation rate over the last few Gyr implies a declining rate of mergers over the evolution of the MW. This is consistent with the observed rate of deposition of radioactive $^{244}$Pu on Earth today which is considerably lower than its value 4.6 Gyr ago when the solar system was formed (Hotokezaka, Piran & Paul 2015; Wallner et al. 2015). Finally, a population of rapid mergers was shown by Hotokezaka, Beniamini & Piran (2018) to provide an explanation for the declining rate of [Eu/Fe] as a function of [Fe/H] observed in the MW for [Fe/H] $\gtrsim -1$ which is not trivially explained with a $D \propto t^{-1}$ DTD (see e.g. Matteucci et al. 2014; Wehmeyer, Pignatari & Thielemann 2015; Coté et al. 2016; Komiya & Shigeyama 2016; Simonetti et al. 2019).

As mentioned in §1, using the sGRB redshift and peak flux distribution, Wanderman & Piran (2015) have found that $dN/dt_{\text{m}} \propto t_{\text{m}}^{-0.8 \pm 0.2}$. There is some tension between this result and ours. However, Wanderman & Piran (2015) allow in their analysis for the existence of a second population that follow the SFR. They find that these population comprises of about 1/3 of sGRBs and interpret this as these bursts being “short duration” Collapsars. Theses findings are consistent with our results if a significant fraction of these are genuine sGRBs as genuine sGRBs.

We can compare our results to the implications of the first NS merger observed in gravitational waves, GW170817. As mentioned earlier for GW170817, the time delay is estimated at $t_{170817} \approx 1 - 10$ Gyr. Thus, GW170817, can be used to put an upper limit on the rate of short mergers due to the following argument. The inferred rate of BNS mergers and the amount of r-process mass estimated from this single event, are already large enough to dominate the production of r-process elements in the Universe (see e.g. Hotokezaka, Beniamini & Piran 2018 for a detailed discussion). In fact, they are already somewhat larger than those inferred from the total r-process mass in the MW (albeit with still relatively large errors). If the delay time distribution is significantly steeper than $t^{-1}$, then the inferred rate of overall mergers (and in turn their contribution to the total r-process mass) would have been much larger in the past ($\propto (t_{170817}/t_{\text{min}})^{-1}$). Applying a conservative limit, that the overall merger rate should not be more than 5 times larger than that inferred by GW170817, we find $b < 1 + \log(5)/\log(t_{170817}/t_{\text{min}})$, $b < 1.5$ for $t_{170817} = 1$ Gyr and $b < 1.3$ for $t_{170817} = 10$ Gyr. These limits are consistent with the level of enhancement of short mergers found in the present work.

7 DISCUSSION

We have explored the distribution of delay times (DTD) between formation and merger of binary neutron star (BNS) systems. We have shown that the times since formation of the observed population of BNS systems, that are dictated by the lifetime of the pulsars, are sufficiently short so that the ages of the observed BNS systems have little to no effect on the observed merger time distribution. This distribution then directly traces the intrinsic DTD distribution. Using a maximal likelihood we found several distributions that fit the observed data.

At long times ($t \gtrsim 1$ Gyr), the DTD is consistent with the theoretically expected $D \propto t^{-1}$. At shorter times, however, there is a statistically significant excess of rapidly merging systems (almost half of the systems merge in less than 1 Gyr). We have shown that the finite pulsar’s lifetime cannot explain this excess which must be intrinsic. We explored a few models with this characteristic. One such model that provides a good description of the data is a DTD where 40% of the systems arise from a log-normal distribution with $\mu = \log(0.3)/\text{Gyr}$, $\sigma = 1$ while the other 60% arise from a $t^{-1}$ distribution, with minimum and maximum merger times of $t_{\text{min}} = 0.035$ Gyr, $t_{\text{max}} = 10^4$ Gyr. However there is not enough data to distinguish between this distribution and distributions that have other functional forms that include an excess of short merger time systems.

The observed chemical evolution of heavy r-process elements in the Milky Way is estimated from the ratio of [Eu/Fe] as compared with [Fe/H]. The decrease in [Eu/Fe] for [Fe/H] $\gtrsim -1$ has been puzzling within the context of DNS as sites of r-process nucleosynthesis as it is not trivially explained with a $D \propto t^{-1}$ DTD (see e.g. Matteucci et al. 2014; Wehmeyer, Pignatari & Thielemann 2015; Coté et al. 2016; Komiya & Shigeyama 2016; Simonetti et al. 2019). As mentioned earlier this population of rapid mergers pro-
vide a natural explanation for this puzzle (Hotokezaka, Beniamini & Piran 2018).

The range of merger times, from $t_{\text{min}} \approx 0.035 (0.001 - 0.15)$ Gyr to $t_{\text{max}} \approx 10^6$ Gyr corresponds to initial orbital separations (for a canonical $1.4 - 1.4 M_\odot$ binary) of $7 (3 - 10) \cdot 10^{10}$ cm to $5 \cdot 10^{12}$ cm. The minimal value is most interesting as it provides an indication on the state of the system just before it collapses to form a neutron star. This distance is about a solar radius. Since the progenitor is essentially heavier than that, it implies that the progenitor must have lost a significant fraction of its envelope just before the collapse, or that its companion was within the envelope during the collapse.

Our inferred DTD suggests that roughly half of the BNS systems ever formed have merged by today. This could lead to $\approx 3 \cdot 10^5$ unmerged BNS systems in the MW. That being said, the vast majority of those systems are not expected to host active pulsars.

The departure of the DTD from a simple $t^{-1}$ may point towards different channels of BNS formation. Indeed (Beniamini & Piran 2016) have shown that there are two classes of collapses that lead to the formation of the second NS in a BNS system. In roughly $60 - 70\%$ of the systems, the collapse involves only a small amount of mass ejection ($\Delta M \lesssim 0.5 M_\odot$) and a weak kick ($v_k \lesssim 30 \text{ km/sec}$), while in the rest of the systems the collapse involve larger ejected masses and kicks. The first type of collapse results in low eccentricity ($e \approx 0.1$) systems, while the second can result in much more eccentric orbits. Since there appears to be no correlation between $e$ and $t_m$ (see table [1]), it seems unlikely that the excess of rapidly merging systems is related to the nature of the kicks. An alternative possibility is that the excess is related to the initial separation at birth (see also Belczynski et al. 2006), which in turn may indicate a different stellar evolutionary part involving a larger shrinkage of the orbit in a common envelope phase.

GW detectors are expected to detect many more BNS mergers in the coming years (Abbott et al. 2017). Although the coincident detection of GRBs are predicted to be much less common (Beniamini et al. 2019), we may expect the detection of the isotropic Macronovae in most cases. These will provide us with further data concerning the rate and $r$-process yield of BNS mergers. The SFR history of the host galaxies could reveal information on the DTD. We expect that at least $50\%$ of the hosts will show evidence of recent SFR. Advanced radio telescopes such as SKA would increase the sample of Galactic BNS and this will enable us to improve the statistics explored here.

ACKNOWLEDGMENTS

We thank Ehud Nakar, Bernard Schutz and Ben Shenhav for helpful discussions. TP was partially supported by an ERC advanced grant TReX and by the CHE-ISF center for excellence in Astrophysics.

REFERENCES

Abbott B. P. et al., 2017, Physical Review Letters, 119, 161101
Argast D., Samland M., Thielemann F.-K., Qian Y.-Z., 2004, A&A, 416, 997
Belczynski K. et al., 2018, arXiv e-prints

3 The values in bracket correspond to upper and lower estimates of $t_{\text{min}}$.
Stairs I. H., Thorsett S. E., Taylor J. H., Wolszczan A., 2002, ApJ, 581, 501
Stovall K. et al., 2018, ApJ, 854, L22
Swiggum J. K. et al., 2015, ApJ, 805, 156
Tauris T. M. et al., 2017, ApJ, 846, 170
Totani T., Morokuma T., Oda T., Doi M., Yasuda N., 2008, PASJ, 60, 1327
Tsujimoto T., Shigeyama T., 2014, A&A, 565, L5
van Leeuwen J. et al., 2015, ApJ, 798, 118
Wallner A. et al., 2015, Nature Communications, 6, 5956
Wanderman D., Piran T., 2015, MNRAS, 448, 3026
Wehmeyer B., Pignatari M., Thielemann F.-K., 2015, MNRAS, 452, 1970
Weisberg J. M., Nice D. J., Taylor J. H., 2010, ApJ, 722, 1030
Weisz D. R., Dolphin A. E., Skillman E. D., Holtzman J., Gilbert K. M., Dalcanton J. J., Williams B. F., 2015, ApJ, 804, 136
Yungelson L. R., Livio M., 2000, ApJ, 528, 108

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