Lessons learned from CHIME repeating FRBs

Wenbin Lu\textsuperscript{1⋆}, Anthony L. Piro\textsuperscript{2}, and Eli Waxman\textsuperscript{3}

\textsuperscript{1}Theoretical Astrophysics, Walter Burke Institute for Theoretical Physics, Mail Code 350-17, Caltech, Pasadena, CA 91125, USA
\textsuperscript{2}The Observatories of the Carnegie Institution for Science, 813 Santa Barbara Street, Pasadena, CA 91101, USA
\textsuperscript{3}Department of Particle Physics & Astrophysics, Weizmann Institute of Science, Rehovot 76100, Israel

31 March 2020

ABSTRACT
CHIME has now detected 18 repeating fast radio bursts (FRBs). We explore what can be learned about the energy distribution and activity of the repeaters by constructing a realistic FRB population model, taking into account wait-time clustering and cosmological effects. For a power-law energy distribution $dN/dE \propto E^{-\gamma}$ for the repeating bursts, a steep energy distribution means that most repeaters should be found in the local Universe with low dispersion measure (DM), whereas a shallower distribution means some repeaters may be detected at large distances with high DM. It is especially interesting that there are two high-DM repeaters (FRB 181017 and 190417) with $\text{DM} \sim 10^3 \text{pc cm}^{-3}$. These can be understood if: (i) the energy distribution is shallow $\gamma = 1.7^{+0.3}_{-0.1}$ (68% confidence) or (ii) a small fraction of sources are extremely active. In the second scenario, these high-DM sources should be repeating more than 100 times more frequently than FRB 121102, and the energy index is constrained to be $\gamma = 1.9^{+0.3}_{-0.2}$ (68% confidence). In either case, this $\gamma$ is consistent with the energy dependence of the non-repeating ASKAP sample, which suggests that they are drawn from the same population. Finally, we show that the CHIME repeating fraction can be used to infer the distribution of activity level in the whole population.

Key words: fast radio bursts; general

1 INTRODUCTION

Since the discovery of the first repeater FRB 121102 (Spitler et al. 2016), it is clear that a significant fraction of fast radio bursts (FRBs) are from non-cataclysmic sources. This is supported by the detection of 18 more repeaters by the Canadian Hydrogen Intensity Mapping Experiment (CHIME/FRB Collaboration et al. 2019a,b; Fonseca et al. 2020; The CHIME/FRB Collaboration et al. 2020, hereafter C1, C2, C3, C4). A common property shared by these repeaters (at least the ones with $>3$ bursts) is that fainter bursts are more common than brighter ones (e.g., Scholz et al. 2016; Law et al. 2017; Gourdji et al. 2019; Kumar et al. 2019; Oostrum et al. 2019, C1-C4). For FRB 121102, the energy distribution of bursts can be modeled by a power-law $dN/dE \propto E^{-\gamma}$, but the index $\gamma$ is debated due to the lack of a homogeneously selected sample that spans a sufficiently wide range of burst energy\textsuperscript{1}. Such a distribution means that it is generally more difficult to detect a repeating source if it is located at a larger distance. For instance, if FRB 121102 were at a redshift $z = 1$, then the CHIME fluence threshold of a few Jy ms corresponds to an energy threshold of $\sim 10^{33}$ erg Hz$^{-1}$. However, none of the observed bursts from FRB 121102 are sufficiently bright to exceed this threshold. We see that the observed rate of such bright bursts should depend on $\gamma$ and that this should control whether repeaters can be detected at high redshift.

Generally, a very steep (or soft) energy distribution means that luminous bursts are extremely rare and hence most repeaters should be found in the local Universe. Conversely, a very shallow (or hard) energy distribution means that most bursts are common and hence should be detected far away. This can be illustrated with a simple toy model. For the case of a Euclidean Universe (without redshift factor), a survey with fluence threshold $F_{\text{th}}$ will only be able to detect sources above the energy threshold $E_{\text{th}} = 4\pi D^2 F_{\text{th}}$, where $D$ is the distance. Hence the cumulative detection rate is $N(> E_{\text{th}}) \propto E_{\text{th}}^{1-\gamma} \propto D^{2-2\gamma}$. If we take the Poisson waiting time distribution in the limit of small detection probability (appropriate at sufficiently large distances $D$), then the probability of detecting two bursts from the same source is $P_{\text{rep}} \propto [N(> E_{\text{th}})]^2 \propto D^{4-4\gamma}$. The distance distribution of repeaters from a given survey is $dN_{\text{rep}}/dD \propto D^2 P_{\text{rep}}(D) \propto D^{6-4\gamma}$. For this particular example, we see that most bursts

\textsuperscript{1}wenbinlu@caltech.edu

\textsuperscript{1} This is further complicated by the well-known temporal clustering of repetitions, which means a large number of observing sessions are needed to statistically analyze the rate of the brightest bursts.
will be detected at large distances if \( \gamma < 7/4 \), and for very steep energy distribution \( \gamma > 7/4 \), nearly all repeaters should be found nearby.

In this paper, we follow similar arguments to construct a more realistic model of the population of repeating FRBs. We take into account the non-Poissonian waiting time distribution (Oppermann et al. 2018), the large number of CHIME observing sessions, the redshift evolution of source densities, the FRB frequency spectrum, and the stochastic DM contributions from the host galaxy and the inhomogeneous intergalactic medium (IGM). After accounting for the uncertainties in these aspects, our model predicts the DM distribution of repeaters, which is compared with observations to constrain \( \gamma \).

The paper is organized as follows. In §2, we calculate the probabilities of detecting single and multiple bursts from a given source. Then in §3, we consider two different populations, one of which assumes that all repeaters have the same energy distribution function and the other one assumes a broad range of activity levels (some are more active than others). The results for these two cases are presented in §3.1 and §3.2, respectively. We discuss the implications of the inferred value of \( \gamma \) and the link between repeaters and the apparent non-repeaters in §4.1. Then in §4.2, we show that, in the future, the CHIME repeating fraction can be used to infer the distribution of activity levels among different sources. Various caveats in our modeling are discussed in §4.3. We provide a summary of our findings in §5. We adopt the latest Planck ΛCDM cosmology (Planck Collaboration et al. 2016).

## 2 Model for a Single Repeater

To study the distribution of bursts from a repeating source, we use the Weibull probability density function (PDF) which has been used to model bursts from FRB 121102 (Oppermann et al. 2018). For a time interval \( \delta \) between adjacent bursts of isotropic energy above \( E \), the PDF is

\[
W(\delta; \lambda(E), k) = \lambda k (\lambda \delta)^{k-1} \exp[-(\lambda \delta)^k],
\]

where \( \lambda(E) \) is related to the mean repeating rate above energy \( E \) and \( k \) is a shape parameter. The cumulative density function (CDF) and mean interval are given by

\[
\text{CDF}(\delta) = \int_0^\delta W(\delta)d\delta = 1 - \exp[-(\lambda \delta)^k], \tag{2}
\]

\[
\langle \delta \rangle = \int_0^\infty \delta W(\delta)d\delta = r^{-1}, \quad r = \frac{\lambda}{\Gamma(1+1/k)}, \tag{3}
\]

where \( \Gamma(x) \) is the gamma function and \( r \) is the mean repeating rate. When \( k < 1 \), \( W(\delta; \lambda, k) \) describes that the bursts are clustered in that small intervals are favored compared to the Poissonian case and that the presence of one burst makes the detection of an of additional burst in the near future more likely. The \( k = 1 \) case recovers the Poisson distribution, because \( r(k = 1) = \lambda \). When \( k > 1 \), the Weibull distribution can approximately describe skewed or symmetric normal distributions.

For a single continuous observing run lasting for \( T \), one can derive that the probability of seeing zero events is

\[
P_0(\lambda, k) = \frac{1}{k !} \Gamma(1+1/k) \Gamma_{\text{in}}(1/k, (\lambda T)^k), \tag{4}
\]

the probability of seeing at least one event is

\[
1 - P_0(\lambda, k) = \frac{1}{k !} \Gamma(1+1/k) \gamma_{\text{in}}(1/k, (\lambda T)^k), \tag{5}
\]

and the probability of seeing exactly one event is

\[
P_1(\lambda, k) = \frac{\lambda T}{\Gamma(1+1/k)} \int_0^1 \exp[-(\lambda T)^k x^k + (1 - x)^k] \, dx. \tag{6}
\]

In these expressions, \( \Gamma_{\text{in}}(s, x) = \int_0^\infty t^{s-1}e^{-t}dt \) and \( \gamma_{\text{in}}(s, x) = \int_0^x t^{s-1}e^{-t}dt \) are the upper and lower incomplete gamma functions respectively. These probabilities are shown in Figure 1. In the limit \( (\lambda T)^k \ll 1 \), the lowest order expansion of \( \exp(-x) \approx 1 - x \) gives

\[
1 - P_0 \approx r T \left[ 1 - \frac{(\lambda T)^k}{k + 1} \right], \tag{7}
\]

\[
P_1 \approx r T \left[ 1 - \frac{2(\lambda T)^k}{k + 1} \right], \tag{8}
\]

and the probability of detecting two or more bursts is

\[
1 - P_0 - P_1 \approx r T \frac{(\lambda T)^k}{k + 1}. \tag{9}
\]

Therefore, for a single observing session, if the chance of detection is small \( 1 - P_0 \approx r T \ll 1 \), the probability of identifying the source as a repeater is even smaller by another factor of \( (\lambda T)^k/(k + 1) \).

If there are \( n \) independent observing runs of identical durations \( T \) (so that the total duration is \( nT \)), the probability of detecting at least one burst is

\[
P_{\text{det}}(\lambda, k) = 1 - P_0^n. \tag{10}
\]

The probability of single detection is

\[
P_{\text{sig}}(\lambda, k) = n P_1 P_0^{n-1}. \tag{11}
\]

The probability of repetitive detection (of at least two events) is

\[
P_{\text{rep}}(\lambda, k) = P_{\text{det}} - n P_1 P_0^{n-1}. \tag{12}
\]

The fraction of repetition is

\[
f_{\text{rep}} = \frac{P_{\text{rep}}}{P_{\text{det}}} = \frac{1 - P_0^n (1 + n P_1/P_0)}{1 - P_0^n}. \tag{13}
\]

These probabilities are shown in Figure 2. The formalism can be generalized into \( n \) runs of non-equal durations, but the set-up of identical sessions is reasonable since the CHIME beams sweep across the position of each source on a regular basis. Our numerical results are based on the exact expressions above.

For the purpose of intuitive understanding, we consider the limit \( n r T \ll 1 \) or \( P_{\text{det}} \ll 1 \) (such that simultaneous observation of a large number of \( P_{\text{det}} \) sources is needed to yield detection), and the probabilities can be simplified to the first order

\[
P_{\text{det}} \approx n r T \left[ 1 - \frac{(\lambda T)^k}{k + 1} - \frac{n - 1}{2} r T \right], \tag{14}
\]

\[© 0000 RAS, MNRAS 000, 000–000\]
Figure 1. Probabilities of detecting at least one (left) or at least two (middle) bursts, and the ratio of these two probabilities (right panel), for a single observing session of duration $T$. In the limit of small Weibull shape parameter $k \ll 1$ (the highly clustered case), the detection probability $1 - P_0$ is typically much less than the Poissonian value of $rT$. Another consequence of strong temporal clustering ($k \ll 1$) is that the repeating fraction $(1 - P_0 - P_1)/(1 - P_0)$ is much higher than the corresponding Poissonian case.

Figure 2. Probabilities of detecting at least one (left) or at least two (middle) bursts, and the ratio of these two probabilities (right panel), for $n = 150$ independent identical observing sessions of duration $T$. With a large number of observing runs, the detection probability $P_{\text{det}}$ is close to the Poissonian value $nT$ for $k \gtrsim 1$ and is much higher for $k \ll 1$. Strong temporal clustering ($k \ll 1$) also increases the fraction of repeating bursts out of all the detected ones, $P_{\text{rep}}/P_{\text{det}}$. 
For a given source at redshift $z$ and luminosity distance $D_L(z)$, the survey is only sensitive to bursts above energy

$$E_{\text{th}} = 4 \pi D_L^2 F_{\text{th}} (1+z)^{\alpha-1},$$

where we have used k-correction by adopting an intrinsic spectrum of $E_o \propto \nu^{-\alpha}$. We use $\alpha = 1.5$ as motivated by statistical studies of the ASKAP sample (Macquart et al. 2019). Our results are insensitive to the spectral slope, as long as it is not extremely steep, $\alpha \lesssim 3$ (e.g., Sokolowski et al. 2018). We integrate equation (17) to obtain the rate of events above $E_{\text{th}}$,

$$r(E_{\text{th}}) = r_0 \int_{E_{\text{th}}}^{\infty} \frac{dE}{E} \left( \frac{E}{E_0} \right)^{-\gamma} \exp \left( -\frac{E}{E_{\max}} \right),$$

For a given source above a specific energy $E$ (in units of erg Hz$^{-1}$) to be in the Schechter form with $\gamma > 1$,

$$\frac{d\nu}{dE} = \frac{r_0}{E_0} \left( \frac{E}{E_0} \right)^{-\gamma} \exp \left( -\frac{E}{E_{\max}} \right),$$

which means that fainter bursts occur more frequently than brighter ones and that there is a maximum energy $E_{\max}$ above which the rate cuts off exponentially. The rate of individual sources is normalized at energy $E_0 = 10^{39}$ erg Hz$^{-1}$, and $r_0$ is in units of hr$^{-1}$. The comoving number density of FRB sources as a function of redshift $z$ (for $z \gtrsim 1$) is

$$n_\star(z) = n_0 (1+z)^\beta,$$

where $n_0$ (in Gpc$^{-3}$) is the local number density and $\beta$ describes the cosmological evolution of FRB sources (e.g., $\beta \approx 2.7$ if FRB sources trace cosmic star formation history, Madau & Dickinson 2014).

For a given source, the probability for at least one detection $P_{\text{det}}[r(\nu) > E_{\text{th}}]$, $k$, $n$, $T$] and the probability for repeating detection is $P_{\text{rep}}[r(\nu) > E_{\text{th}}], k$, $n$, $T]$, as given by equations (10) and (12), respectively. If we let $z_c$ be the critical redshift below which more than half of the detected sources are repeaters, which is found by setting $r(\nu > E_{\text{th}}) n T \approx 1$ (corresponding to $P_{\text{det}} \sim 1$ and $P_{\text{rep}} \sim 0.5$, see Figure 2), we estimate

$$z_c \approx 0.1 \left[ \frac{r_0 n T}{(\gamma - 1)} \right]^{1/(1 - \gamma)}.$$
For a given set of parameters \( p \), the difference compared to the overall statistical error.

The number of sources in the peak bin near \( \log \text{DM} \) is dominated by where the error is the largest. This has been discussed in detail in the Appendix where we take into account the Milky Way contribution and IGM.

The posterior distribution for \( \log \text{DM} \) is shown in Figure 3. We constrain the power-law index for the energy distribution to be \( 1.6 \leq \gamma \leq 3 \) (68% confidence interval), whereas the two other parameters \( k \) (waiting time) and \( \beta \) (redshift evolution) are essentially unconstrained due to small number statistics. Note that the constraints given by our likelihood function are conservative because we only make use of the detected number of repeaters in two bins at \( \log \text{DM}_\text{ex} \) between 4 and 7, and the resulting difference is smaller than the statistical error. We also find that directly using \( n_{\text{peak}} \) instead of \( \lambda_{\text{peak}} \) only makes a small difference compared to the overall statistical error.

\[ n_{\text{obs}} = 2 \text{ in this bin. Since the number of sources in the field of view is large, the number of detections is time independent and the likelihood function for this set of parameters is Poissonian,} \]

\[ L(p) = n_{\text{exp}}^{n_{\text{obs}}} e^{-n_{\text{exp}}}/n_{\text{obs}}!, \]

which is then used to calculate the posterior distribution of the parameters \( p \) using the Bayesian theorem. We use the following priors: \( 0.2 < k < 1.3, 1.2 < \gamma < 3, 0 < \beta < 3 \). The lower limit of \( k \) is motivated by the studies of FRB 121102 by Oppermann et al. (2018) and by the fact that the detections of repetitions are often (but not always) spread over multiple observing sessions (C1-C4, Kumar et al. 2019; Oostrum et al. 2019). The motivation for the prior on redshift evolution is that the source number density is assumed to be somewhere between non-evolving \((\beta = 0)\) and tracing star-formation history \((\beta \approx 3)\). The final result of the marginalized PDF for \( \gamma \) depends weakly on the \( \beta \) prior.

The posterior distribution for \( p = (k, \gamma, \beta) \) is sampled by a Markov-Chain Monte Carlo (MCMC) simulation and is shown in Figure 3. We constrain the power-law index for the energy distribution to be \( 1.6 \leq \gamma \leq 3 \) (68% confidence interval), whereas the two other parameters \( k \) (waiting time clustering) and \( \beta \) (redshift evolution) are essentially unconstrained due to small number statistics. Note that the constraints given by our likelihood function are conservative because we only make use of the detected number of repeaters in two bins at \( \log \text{DM}_\text{ex} \) between 4 and 7 (to minimize possible systematic errors due to selection biases).
Our method can be generalized to include the full $\text{DM}_{\text{ex}}$ distribution when more bursts are available and possible selection biases are reasonably understood.

For the purpose of gaining analytical insight, in the $z \gg z_c$ limit, we have $r(>E_{\text{th}})nT \ll 1$ and the repeating probability of each source is roughly given by equation (16). For the simplest case of Euclidean universe (without redshift factor) and Poisson distribution ($k = 1$), the cumulative distance ($D$) distribution of repeaters is given by

$$N_{\text{rep}}(< D) \simeq \frac{n_* \Omega(nT)^2}{2} \int_0^D dD D^2 [r(>E_{\text{th}})]^2 \simeq \frac{n_* \Omega(r_0 nT)^2}{2(\gamma - 1)^2} \int_0^D dD D^2 \left( \frac{E_{\text{th}}}{E_0} \right)^{2-2\gamma},$$

which means $N_{\text{rep}}(< D) \propto dN_{\text{rep}}/d\log D \propto D^{7-4\gamma}$. The $\text{DM}_{\text{ex}}$ distribution in Figure 3 is not much steeper than $\sim D^{-1}$, implying $\gamma \lesssim 2$. Properly including comoving volume in a $\Lambda$CDM Universe will predict less high-$z$ repeaters and strengthen the upper limit on $\gamma$. On the other hand, for strong waiting time clustering $k \lesssim 0.5$, we predict more high-$z$ repeaters because $P_{\text{rep}} \propto r^{1+k}$ instead of $r^2$ so the larger $\gamma$ is allowed.

### 3.2 Broad Distribution of $r_0$

We now relax the assumption of a single population of FRB sources and consider a broad distribution of $r_0$. Such a scenario is applicable if some sources are more active than others (as suggested by the non-detection of repeaters by ASKAP and follow-ups, e.g., James et al. 2019). We adopt a power-law distribution as follows

$$\frac{dn_*}{dr_0} = \frac{n_*}{r_{\text{max}}} \frac{q - 1}{R^{q-1} - 1} (r_0/r_{\text{max}})^{-q},$$

where $n_*$ is the total number density, $r_{\text{max}}$ is the maximum repeating rate normalization, $R = r_{\text{max}}/r_{\text{min}}$ is the ratio between the maximum and minimum repeating rates, and $q$ is a power-law index. Currently, it is unclear whether most sources are near the most active end ($r_0 \sim r_{\text{max}}$, $q < 1$) or the least active end ($r_0 \sim r_{\text{max}}/R$, $q > 1$). A possible physical scenario is the following. For each source, the re-
peating rate drops as a power-law function of age $r_0 \propto t^{-p}$ and, if the source birth rate per unit volume is constant, the number of sources is proportional to the age $t \propto r_0^{-1/p}$, so we obtain $dn_*/dr_0 \propto r_0^{-1-1/p}$ or $q = 1 + 1/p > 1$.

The cumulative redshift distribution of repeaters in equation (23) is now modified to

$$N_{\text{rep}}(<z) = \frac{\Omega}{4\pi} \int_0^z dz \frac{dV}{dz} (1+z)^{\beta} \int_{z_{\text{max}}}^{r_{\text{max}}/r_{\text{th}}} \frac{dr_0}{r_0} \frac{dn_*}{dr_0}$$

$$\times P_{\text{rep}}[r(>E_{\text{th}})]_k n, T/(1+z)].$$

A broad distribution of $r_0$ strongly impacts the redshift distribution of repeaters at $z \ll 1$, because the critical redshift $z_c$, below which $P_{\text{rep}} \sim 0.5$, depends on $r_0$ through equation (22). However, the shape of the redshift distribution at $z \gg z_c(r_{\text{max}})$ is not affected, because $P_{\text{rep}} \propto r_0^t + k \propto r_0^{1+\kappa} E_{\text{th}}$ and hence the $dr_0$ integral separates from the $\int dz$ integral. If the most active repeaters are similar to FRB 121102, i.e., $r_{\text{max}} \sim 0.1 h^{-1}$ (Law et al. 2017; James 2019), which corresponds to $z_c(r_{\text{max}}) \sim 0.1 h^{-1}$, then $z \gg z_c(r_{\text{max}})$ roughly holds for the majority of CHIME repeaters with $DM_{ex} \gtrsim 300$ pc cm$^{-3}$. Thus, the constraints on $\gamma$ are similar to the single population case as discussed in §3.1.

It is also possible that a small fraction of sources are extremely active such that the high-$z$ repeaters are dominated by the most active ones. We include two additional free parameters $\log r_{\text{max}}[h^{-1}]$ and $q$ in the likelihood analysis, with sufficiently wide priors of $-1 < \log r_{\text{max}}[h^{-1}] < 2$ and $0 < q < 3$ to account for this. The maximum $r_{\text{max}} = 10^2 h^{-1}$ roughly corresponds to repeating sources that are $10^3$ times more active than FRB 121102 (although no such hyper-active sources have been identified observationally).

We have tested that the final constraints on $\gamma$ are not sensitive to the ratio $R = r_{\text{max}}/r_{\text{min}}$ as long as it is sufficiently large (we adopt $R = 10^6$ in practice).

The MCMC-sampled posterior distribution for $p = (k, \gamma, \beta, \log r_{\text{max}}[h^{-1}], q)$ is shown in Figure 4. We constrain the energy distribution of repeating bursts to be $\gamma = 1.9^{+0.3}_{-0.2}$ (68% confidence interval), whereas the other parameters are unconstrained due to small number statistics (see §4.2 for a discussion on how $r_{\text{max}}$ and $q$ may be constrained by the CHIME repeating fraction). It is interesting to look at the covariance between $r_{\text{max}}$ and $\gamma$. As anticipated, we see that very large $r_{\text{max}} \gtrsim 10 h^{-1}$ (for sources more than $10^3$ times more active than FRB 121102) makes it possible to detect high-$DM_{ex}$ repeaters without requiring a shallow energy distribution for each source, so the best-fit energy distribution index $\gamma$ is pushed to higher values (in this case, small $\gamma$ tend to over-produce the number of high-$DM_{ex}$ repeaters). This can be directly tested by future monitoring of high-$z$ repeaters FRB 181017 and 190417, just to see whether they are hyper active with $r_0 \gtrsim 10 h^{-1}$.

4 DISCUSSION

In this section, we first discuss the implications of our constraints on $\gamma$ and the link between repeaters and the apparent non-repeaters. Then, we show that the CHIME repeating fraction can be used to infer the distribution of activity levels among different sources. Finally, we point out the caveats of our modeling with suggestions for improvement in the future.

4.1 Implications of $\gamma$

To summarize, we find that CHIME detection of high-$DM_{ex}$ repeaters constrains $1.6 \lesssim \gamma \lesssim 2$, if all sources are from the same population similar to FRB 121102 as assumed in §3.1. In the more general case considered in §3.2, where some sources are allowed to be much more active than others, the constraints become $1.7 \lesssim \gamma \lesssim 2.2$, slightly steeper because the detected high-$DM_{ex}$ repeaters could simply be the most active sources (and hence do not require a shallow energy distribution for each source). If the total volumetric rate of FRBs is dominated by the sources we model (with the underlying assumption that they all repeat with the same $\gamma$), then the differential volumetric rate per energy also has the same power-law behavior

$$\frac{d\Phi}{dE} = \int dr_0 \frac{dn_*}{dr_0} r_0 \frac{dN}{dE} = n_* \langle r_0 \rangle \frac{dN}{dE} \propto E^{-\gamma},$$

where $\langle r_0 \rangle \equiv n_*^{-1} \int dr_0 (dn_* / dr_0) r_0$ is the mean rate normalized and $\Phi$ is in units of Gpc$^{-3}$ yr$^{-1}$.

Independent analysis of the ASKAP sample of apparent non-repeating FRBs by the authors (Lu & Piro 2019) constrains $1.3 \lesssim \gamma \lesssim 1.9$ (68% confidence interval), which is in rough agreement with the constraints from the CHIME sample of repeaters$^4$, as shown in Figure 5. This suggests, although does not prove, that all FRB sources are repeating$^5$ with the same $\gamma$. The errors in both constraints are

$^4$ It is worth noting that the CRAFT survey has relatively poor threshold ($F_{\text{th}} \sim 50$ Jy ms) and is only capable of detecting the brightest bursts (Shannon et al. 2018). The repeating rate above the threshold energy for each source $r(E_{\text{th}})$ is small in that $P_{\text{repl}} \ll 1$ is well satisfied, so the detection probability is linearly proportional to the total observing time spent on each source, independent of the arrangement of observing runs as shown by equation (14).

$^5$ This is also supported by two other observations: (1) One of the
dominated by the small number of sources, which will be dramatically improved by future observations.

We also note that direct measurements of $\gamma$ from individual repeaters may have large uncertainties (Law et al. 2017; Wang & Yu 2017; Gourdji et al. 2019; James 2019, C4); see Table 2 of Oostrum et al. (2019). This can be understood if the monitoring time is not long enough to capture the more luminous bursts. For a given source, the Weibull waiting time distribution with $k < 0.5$ has the property that a survey either detects a large number of bursts or no burst at all (see Figure 4 of Oppermann et al. 2018). This effect leads to a bias towards larger $\gamma$ or steeper energy distribution.

### 4.2 Future constraints by repeating fraction

In this subsection, we show how the CHIME repeating fraction can be used to constrain the unknown parameters other than $\gamma$ and reveal the distribution of activity levels among different FRB sources (as defined in eq. 26). Generally, a steep activity-level distribution $q > 2$ means that most sources are very active with $r_0 \ll r_{\text{max}}$ and most CHIME repeaters should be the just repeating frequently enough to give repetitive detection. A very shallow distribution $q < 1$ means that most sources are very active with $r_0 \sim r_{\text{max}}$ and hence we roughly recover the single population case as discussed in 3.1. The intermediate region of $1 < q < 2$ is more complex in that, although most sources are not very active, the detected repeaters may or may not be dominated by the most active sources near $r_{\text{max}}$.

We calculate the total repeating fraction for the CHIME survey $f_{\text{rep, tot}} = N_{\text{rep}}/N_{\text{det}}$ within redshift of 2, as a function of $\log r_{\text{max}}$ and $q$. We fix $k = 1/3$ (as indicated by FRB 121102, Oppermann et al. 2018), $\gamma = 1.8$ (as discussed §4.1), and $\beta = 1.5$ for mild redshift evolution, and the results are qualitatively similar for other choices of these parameters. As shown in Figure 6, the CHIME repeating fraction increases towards larger $r_{\text{max}}$ and smaller $q$. If most sources are similar to or more active than FRB 121102 ($\log r_{\text{max}}[\text{hr}^{-1}] \gtrsim -1$, $q < 1$), then $\gtrsim 10\%$ of all CHIME sources should be repeaters. In the other extreme limit of $q > 2.5$ and $\log r_{\text{max}}[\text{hr}^{-1}] \simeq -1$ (FRB 121102 represents the most active sources), the repeating fraction is much less than 1%.

More information on the population properties can be obtained by studying how the repeating fraction depends on redshift. We define the cumulative repeating fraction as $f_{\text{rep}}(\leq z) = N_{\text{rep}}(< z)/N_{\text{det}}(< z)$, which is shown in Figure 7, for a number of population models with different maximum repeating rate normalizations $r_{\text{max}}$ and activity-level distribution slopes $q$. Again, as a representative example, we fix $k = 1/3$, $\gamma = 1.8$, and $\beta = 1.5$. For this example, we find that, if most sources are similar to FRB 121102 ($r_{\text{max}} \sim 0.1\text{hr}^{-1}$, $q < 1$), then more than 60% sources at $z \lesssim 0.1$ should be repeaters. If FRB 121102 represents the most active sources ($r_{\text{max}} \sim 0.1\text{hr}^{-1}$) but $q > 2.5$, then the repeating fraction is less than a few percent even at very low redshift $z \sim 0.01$. The overall repeating fraction is much higher for the $r_{\text{max}} = 10\text{hr}^{-1}$ cases.
4.3 Potential Limitations

In an effort to be as clear as possible about potential limitations to our analysis, we mention a number of caveats in our current modeling. These will be improved in the future with better statistics and a deeper understanding of CHIME’s selection biases.

1. Our constraints on $\gamma$ may be subjected to CHIME selection biases if high-DM ($\sim 10^3$ pc cm$^{-3}$) bursts are more difficult to detect than the ones with the same fluence but at lower DM ($\sim 300$ pc cm$^{-3}$). The intra-channel dispersion smearing for the CHIME survey with spectral resolution $\Delta \nu$ at frequency $\nu$ is given by

$$\Delta t_{DM} \approx 0.8 \, \text{ms} \, \frac{\text{DM}}{10^3 \, \text{pc cm}^{-3}} \, \frac{\Delta \nu}{0.02 \, \text{MHz}} \left(\frac{\nu}{600 \, \text{MHz}}\right)^{-3},$$

so it is likely that some narrow bursts near the lower end of the frequency band $\nu \sim 400$ MHz are missed. It is also possible if high-DM bursts are preferentially scattering broadened. If such biases exist, then the true fraction of high-DM$_{ex}$ repeaters is larger and hence $\gamma$ should be slightly reduced (giving a shallower energy distribution).

2. The sky positions of different sources have different exposure time and different fluence threshold. Additionally, as the CHIME beams regularly sweep across the position of a given source, the differential exposure time under the instantaneous fluence threshold may not be well modeled by a top-hat function as in our model. These complications can be included in a generalized version of our model in the future when a better understanding of the CHIME beams is available. At the current moment, since most bursts are not detected far away from the beam centers (see Figure 1’s of C2, C3), the effects of beam biases should be weak. Also, the exposure time and fluence threshold for the locations of the two highest DM sources FRBs 181017 and 190417 are close to the median in the CHIME repeater sample (see C2 and C3), so the potential biases due to non-uniform sky coverage should not be strong.

3. Another possible complication is that the host DM contribution for FRB 181017 and 190417 may be close to $10^3$ pc cm$^{-3}$ such that they are actually located at much lower redshifts $z \ll 1$. This is possible given the uncertainties on local ($\lesssim$ pc) environment of FRB progenitors and their host galaxy properties. However, the ($\sim$10) known examples of FRB host galaxies have low to modest DM$_{host}$ of a few 10’s up to about 200 pc cm$^{-3}$. A low local contribution is also expected for young neutron star scenarios if the observed DM is not changing appreciably over time (e.g., Piro 2016; Piro & Gaensler 2018). Future host localizations will test this possibility.

4. Note that our constraints on $\gamma$ are conservative because we only make use of the detected number of repeaters in two bins at log DM$_{ex}$pc cm$^{-3}] \sim 2.5$ and 3. This is to minimize possible systematic errors due to CHIME selection biases in other bins (at DM$_{ex} \lesssim 10^2$ or $\gg 10^3$ pc cm$^{-3}$). Our method can be generalized to include the full DM$_{ex}$ distribution when more bursts are available and possible selection biases are understood. It is also possible to extend our model to predict the DM$_{ex}$ distribution of the sources with more than 2 or 3 detected bursts and then compare it with observations. These additional constraints will provide information on other parameters such as the source number density $n_s$ and the maximum repeating rate $r_{max}$.

5 SUMMARY

In this work we have explored how the redshift (or DM$_{ex}$) distribution of repeating FRBs in a given survey depends strongly on the energy distribution of repeaters and hence can be used to constrain the important property of the sources. We constructed a model for the whole FRB population based on the Weibull waiting time distribution with arbitrary clustering, properly taking into account realistic cosmological effects and that some sources may repeat more frequently than others. The model-predicted DM$_{ex}$ distribution was then compared to the CHIME repeaters to constrain the energy distribution index $\gamma$ in a Bayesian way. Our findings are summarized as follows.

1. Figures 1 and 2 provide a sense for whether single or multiple observing sessions are expected to find repeating sources if all FRBs repeat. This can roughly be compared with future surveys and different strategies to get a better idea if they are able to rule out repetition or not.

2. CHIME’s detection of two high-DM$_{ex}$ repeaters can be understood if either a small fraction of sources are intrinsically much more active than FRB 121102 or the energy distribution for repetitions is shallow. In the first explanation, FRBs 181017 and 190417 should be at least $\sim 10^2$ times more active than FRB 121102. If such extremely active sources dominate high-DM$_{ex}$ repeaters, then the energy distribution index is constrained to be $1.7 \lesssim \gamma \lesssim 2.2$. On the other hand, the second explanation gives shallower power-law index of $1.6 \lesssim \gamma \lesssim 2$. This can also be tested by future monitoring of nearby repeaters.

3. The hypothesis that all FRB sources are repeating with a universal $\gamma \sim 1.8$ is consistent with all observations$^6$, including the CHIME repeaters, the apparent non-repeaters found by the CRAFT survey (Lu & Piro 2019), and FRB 121102 (Law et al. 2017; James 2019; Oostrum et al. 2019). This power-law index is shallower than that of the Crab giant pulses ($\beta \sim 2.1$–3.5, Mickaliger et al. 2012) but consistent with magnetar X-ray bursts ($\beta \sim 1.4$–2.0, Turolla et al. 2015) and other systems displaying self-organized criticality (Katz 1986; Bak et al. 1987). This lends indirect support to the magnetar nature of FRB progenitors (as pointed out earlier by Lu & Kumar 2016; Wang & Yu 2017; Cheng et al. 2020).

4. Our model predicts the repeating fraction $f_{rep}(<z)$ for sources within redshift $z$, which depends on the distribution of activity levels among different FRB sources and is generally a decreasing function of redshift, as shown in Figures 6 and 7. This can be applied once we know the DM$_{ex}$ distributions of both repeaters and the apparent non-repeaters in the CHIME sample. For instance, if most sources are similar to FRB 121102, then we predict $r_{max}$

$^6$ We also tried constraining other parameters (especially $r_{max}$ and $q$) by imposing a prior of $\gamma = 1.8$ or other similar values between 1.6 and 2. We only found that $r_{max} \lesssim 10 \text{hr}^{-1}$ is favored but the current repeater data is insufficient to rule out larger $r_{max}$ at high confidence.
(i) more than 10% of all CHIME sources should be repeaters, and (ii) at sufficiently low redshifts \( z \lesssim 0.1 \) (or \( \text{DM}_{\text{ex}} \lesssim 100 \text{ pc cm}^{-3} \)) nearly all sources should be observed as repeaters by CHIME. Violation of either of them means that most sources are repeating much less frequently than FRB 121102.

6 ACKNOWLEDGMENTS
We thank Paz Beniamini for useful discussions. We thank the organizers for the 2019 Toronto FRB Day workshop and the 2020 Flaton Institute FRB workshop where the current work was initiated. WL is supported by the David and Ellen Lee Fellowship at Caltech.

REFERENCES
Bak P., Tang C., Wiesenfeld K., 1987, PRL, 59, 381
Bannister K. W., et al., 2019, Science, 365, 565
CHIME/FRB Collaboration et al., 2019a, Nature, 566, 235
CHIME/FRB Collaboration et al., 2019b, ApJL, 885, L24
Cheng Y., Zhang G. Q., Wang F. Y., 2020, MNRAS, 491, 1498
Deng W., Zhang B., 2014, ApJL, 783, L35
Fonseca E., et al., 2020, arXiv e-prints, p. arXiv:2001.05959
Foreman-Mackey D., 2016, The Journal of Open Source Software, 24
Gourdji K., Michilli D., Spitler L. G., Hessels J. W. T., Seymour A., Cordes J. M., Chatterjee S., 2019, ApJL, 877, L19
James C. W., 2019, MNRAS, 486, 5934
Marcote B., et al., 2020, Nature, 577, 190
McQuinn M., 2014, ApJL, 780, L33
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Mickaliger M. B., et al., 2012, ApJ, 760, 64
Mount A., Bautz M. A., 2019, Science, 365, 565
Oppermann N., Yu H.-R., Pen U.-L., 2018, MNRAS, 475, 5109
Piro A. L., Gaensler B. M., 2018, ApJ, 861, 150
Ravi V., et al., 2019, Nat, 572, 352
Shannon R. M., et al., 2018, Science, 366, 231
Sokolowski M., et al., 2018, ApJL, 852, L11
Spitler L. G., et al., 2016, Nature, 531, 202
Tendulkar S. P., et al., 2017, ApJL, 834, L7
The CHIME/FRB Collaboration et al., 2020, arXiv e-prints, p. arXiv:2001.10275
Turolla R., Zane S., Watts A. L., 2015, Reports on Progress in Physics, 78, 116901
Wang F. Y., Yu H., 2017, JCAP, 2017, 023
Xu J., Han J. L., 2015, Research in Astronomy and Astrophysics, 15, 1629
Yamasaki S., Totani T., 2020, ApJ, 888, 105
Yao J. M., Manchester R. N., Wang N., 2017, ApJ, 835, 29

APPENDIX A: THE STOCHASTIC RELATION BETWEEN REDSHIFT AND DM_{ex}

An important aspect of comparing our models with the CHIME observations is converting from \( N_{\text{rep}}(< z) \) to \( dN/d\text{dlogDM}_{\text{ex}} \). We describe our approach to this in more detail below.

The free electrons within the host galaxy (ISM and halo) and in the IGM along the line of sight both contribute to \( \text{DM}_{\text{ex}} \) beyond the Milky Way. Each of these two components may have stochastic fluctuations, depending on the host galaxy properties and the number of intervening halos. For a given source at redshift \( z \), we approximate these two components as random Gaussian variables with mean and standard deviation: \( \mu_{\text{host}}, \sigma_{\text{host}}; \mu_{\text{IGM}}, \sigma_{\text{IGM}} \). Then, the PDF of \( \text{DM}_{\text{ex}} \) is given by

\[
\frac{dP}{d\text{DM}_{\text{ex}}}(\text{DM}_{\text{ex}}, z) = \frac{1}{\sqrt{2\pi(\sigma_{\text{host}}^2 + \sigma_{\text{IGM}}^2)}} \times \exp \left[ -\frac{(\text{DM}_{\text{ex}} - \mu_{\text{host}} - \mu_{\text{IGM}})^2}{2(\sigma_{\text{host}}^2 + \sigma_{\text{IGM}}^2)} \right]. (A1)
\]

We adopt \( \mu_{\text{host}} = 100/(1+z) \), \( \sigma_{\text{host}} = 30/(1+z) \), \( \mu_{\text{IGM}}(z) = 900z \), \( \sigma_{\text{IGM}} = 200\sqrt{z} \), all in units of \( \text{pc cm}^{-3} \). These values are motivated by observational and theoretical studies of (potential) FRB host galaxies (Tendulkar et al. 2017; Bannister et al. 2019; Ravi et al. 2019; Prochaska et al. 2019; Marcote et al. 2020; Xu & Han 2015; Luo et al. 2018) and IGM electron density distribution (McQuinn 2014; Deng & Zhang 2014; Shull & Danforth 2018; Ravi 2019; Prochaska & Zheng 2019; Kumar & Linder 2019). The \( \sigma_{\text{IGM}} \propto \sqrt{z} \) behavior can be roughly understood if one divides the line of sight into many short segments (of e.g., ~ 50 Mpc) each of which has fractional DM fluctuation of order unity due to the on average one intervening massive halo, so the sum of \( n \) segments gives fractional fluctuation of \( n^{-1/2} \). In reality, the \( \sigma_{\text{IGM}} \) fluctuation is non-Gaussian with a long tail at high DM given by non-zero probability of an intervening galaxy cluster. Our analysis is only weakly affected by the choice of the above parameters, which may change as our understanding of FRB host galaxies improves.

The above normal distribution allows (unphysical) negative \( \text{DM}_{\text{ex}} \) but at a very low probability, which we ignore. Then the following convolution

\[
\frac{dN_{\text{rep}}}{d\text{DM}_{\text{ex}}}(\text{DM}_{\text{ex}}) = \int_0^\infty dz \frac{dN_{\text{rep}}}{dz}(z) \frac{dP}{d\text{DM}_{\text{ex}}}(\text{DM}_{\text{ex}}, z). (A2)
\]

provides the relation between \( dN_{\text{rep}}/dz \) in equation (23) and the desired distribution of \( dN_{\text{rep}}/d\text{DM}_{\text{ex}} \).