

# The $z$ –DM distribution of fast radio bursts

C. W. James<sup>1</sup>,<sup>\*</sup> J. X. Prochaska,<sup>2,3</sup> J.-P. Macquart<sup>1</sup>,<sup>†</sup> F. O. North-Hickey,<sup>1</sup> K. W. Bannister<sup>4</sup> and A. Dunning<sup>4</sup>

<sup>1</sup>International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia

<sup>2</sup>Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA

<sup>3</sup>Kavli Institute for the Physics and Mathematics of the Universe, 5-1-5 Kashiwanoha, Kashiwa 277-8583, Japan

<sup>4</sup>CSIRO Astronomy and Space Science, PO Box 76, Epping, NSW 1710, Australia

Accepted 2021 October 19. Received 2021 October 19; in original form 2021 January 20

## ABSTRACT

We develop a sophisticated model of fast radio burst (FRB) observations, accounting for the intrinsic cosmological gas distribution and host galaxy contributions, and give the most detailed account yet of observational biases due to burst width, dispersion measure, and the exact telescope beamshape. Our results offer a significant increase in both accuracy and precision beyond those previously obtained. Using results from ASKAP and Parkes, we present our best-fitting FRB population parameters in a companion paper. Here, we consider in detail the expected and fitted distributions in redshift, dispersion measure, and signal to noise. We estimate that the unlocalized ASKAP FRBs arise from  $z < 0.5$ , with between a third and a half within  $z < 0.1$ . Our predicted source-counts ( $\log N$ – $\log S$ ) distribution confirms previous indications of a steepening index near the Parkes detection threshold of 1 Jy ms. We find no evidence for a minimum FRB energy, and rule out  $E_{\min} > 10^{39.0}$  erg at 90 per cent C.L. Importantly, we find that above a certain DM, observational biases cause the Macquart (DM– $z$ ) relation to become inverted, implying that the highest-DM events detected in the unlocalized Parkes and ASKAP samples are unlikely to be the most distant. More localized FRBs will be required to quantitatively estimate this effect, though its cause is a well-understood observational bias. Works assuming a 1–1 DM– $z$  relation may therefore derive erroneous results. Our analysis of errors suggests that limiting factors in our analysis are understanding of FRB spectral behaviour, sensitivity response of search experiments, and the treatment of the repeating population and luminosity function.

**Key words:** methods: statistical – transients: fast radio bursts.

## 1 INTRODUCTION

Fast radio bursts (FRBs) are radio transients of millisecond duration and extragalactic origin (Lorimer et al. 2007; Thornton et al. 2013). Their progenitors are unknown, with very many production mechanisms proposed (Platts et al. 2019). FRB surveys are providing increasingly large statistics with which to study the FRB population (Bhandari et al. 2018; Shannon et al. 2018; The CHIME/FRB Collaboration 2021), including a handful of localized FRBs (Tendulkar et al. 2017; Bannister et al. 2019; Prochaska et al. 2019b; Ravi et al. 2019; Marcote et al. 2020). Furthermore, Macquart et al. (2020) have used localized FRBs as probes of the cosmological distribution of ionized gas, illustrating their utility for cosmological studies (McQuinn 2014; Masui & Sigurdson 2015; Caleb, Flynn & Stappers 2019; Madhavacheril et al. 2019). Of the key questions surrounding FRBs, this work focuses on FRB population statistics.

Studies of the FRB population are important both for understanding the nature of FRBs themselves, and their use as cosmological probes. Typical fitted parameters include the FRB luminosity

function, e.g. minimum and maximum energies, and its shape; FRB spectral properties; and the source evolution. Studies including repeating FRBs must also fit the distribution of repetition rates and allow for time dependence between bursts from a single object.

Linking FRB observations to the underlying true FRB population however is difficult. Connor (2019) review previous methods of studying the FRB population, and emphasize that accurate estimates require accounting for the sensitivity effects of telescope beamshape, intrinsic burst width, and the dispersion measure distribution  $p(\text{DM}|z)$  for a given redshift. In short, one must integrate over known or hypothesized intrinsic distributions of these variables, model observational biases, and then attempt to match observations. Doing so improperly will produce biased results.

In what is usually seen as a different line of inquiry, cosmological studies using FRBs take advantage of their dispersion measure (DM), which integrates the column density of ionized gas along their line of sight. This encodes information on the diffuse gas in voids and galactic haloes which is otherwise difficult to study. Macquart et al. (2020) have recently used the localized FRB population to constrain the total baryon density of the Universe and the degree of feedback. In that work, the authors analyse the probability distribution of observed dispersion measures, DM, given the redshift  $z$  of identified FRB host galaxies,  $p(z, \text{DM})$ . This controls somewhat for the effects of the population of FRBs, which primarily affects the redshift distribution

\* E-mail: [clancy.james@curtin.edu.au](mailto:clancy.james@curtin.edu.au)

† Deceased.

$p(z)$ . The authors do note however the potentially biasing effects of the FRB population as observed by FRB surveys, although a comprehensive treatment of such biasing effects is not performed due to the intrinsic error in using a small sample size (5–7). Many cosmological studies (such as helium reionisation) require many FRBs to both exist and be detectable at a redshift of  $z \sim 4$  (Caleb et al. 2019), which is well beyond the most distant localized FRB to date at  $z \sim 0.6$  (Law et al. 2020).

Fundamentally, both lines of inquiry aim to study the intrinsic distribution of FRBs in  $z$ –DM space,  $p(\text{DM}, z)$ . The only difference is the aspect of interest: population studies try to resolve a redshift distribution  $p(z)$  and treat the distribution  $p(\text{DM}|z)$  as a nuisance distribution, while cosmological studies aim to resolve  $p(\text{DM}|z)$  and attempt to remove the  $p(z)$  factor. They are thus fundamentally coupled problems. Unbiased estimates of the cosmological distribution of ionized gas require knowing the FRB population and the consequent biasing effects on measured dispersion measures; and understanding the FRB population requires knowing the dispersion-measure distribution and its biasing effects on the measured luminosity function.

Caleb et al. (2016) provide the first comprehensive model of observational biases on a simulated burst population, and FRBPOP (Gardenier et al. 2019) is being developed to include such effects. To date however, only Luo et al. (2020) have used this approach to fit population parameters. The authors study a sample of FRBs detected in the  $\sim 1$  GHz band from Parkes, the Upgraded Molonglo Synthesis Telescope (UTMOST), the Australian Square Kilometre Array Pathfinder (ASKAP), Arecibo, and the Greenbank Telescope (GBT). The authors evaluate the validity of their model using MULTINEST (Feroz, Hobson & Bridges 2009), which applies a Bayesian framework, and find a peak FRB luminosity  $L^*$  of  $2.9^{+11.9}_{-1.7} 10^{44} \text{ erg s}^{-1}$ , a differential power-law index of  $-1.79^{0.31}_{-0.35}$ , and a volumetric rate of  $3.7^{+5.7}_{-2.4} 10^4 \text{ Gpc}^{-3} \text{ yr}^{-1}$  above  $10^{42} \text{ erg s}^{-1}$ .

In this work, we significantly improve upon FRB population models in the following manner:

- (i) using an unbiased sample of FRBs from ASKAP and Parkes;
- (ii) using seven localized FRBs detected by ASKAP;
- (iii) correctly accounting for the full telescope beamshape;
- (iv) using the measured signal-to-noise ratio in probability estimates;
- (v) including the intrinsic spread in the cosmological DM contribution due to large-scale structure and galaxy haloes;
- (vi) and allowing for redshift evolution of the FRB event rate per comoving volume.

As with other population models, to make the problem tractable, we assume that the cosmological rate evolution, host galaxy DM contribution, burst width, and burst energy distributions are all independent; and that FRB observations are random and uncorrelated, i.e. we do not model rapidly repeating FRBs.

We begin by describing the ingredients to our model: a model of the DM distribution of FRBs as a function of redshift (Section 2), based on the model of Macquart et al. (2020); the intrinsic properties of FRBs, such as their burst width and the luminosity function, using a standard power-law description (Section 3); and the influence of observational effects such as beamshape and search sensitivity (Section 4). The method to combine these to calculate the expected ‘ $z$ –DM’ distribution for observed FRB surveys is given in Section 5. In Section 6, we describe the data from ASKAP and Parkes to which we fit our model using maximum-likelihood methods. The best-fitting FRB population parameters, and their uncertainties, are given in a companion paper (James et al. 2021a). In this work, we present

detailed comparisons to the observed DM, redshift, and signal-to-noise ratio distributions in Section 8, where we test for goodness of fit and search for deviations from expectations. Section 9 shows our estimates for the expected  $z$ –DM distribution of FRBs detected by ASKAP and Parkes. We summarize our results in Section 10. We attach in appendices a discussion of neglected effects in our modelling, and extra data for alternative source evolution scenarios.

## 2 DISPERSION MEASURE DISTRIBUTION

The distribution of dispersion measure, DM, of FRBs from a given redshift  $z$ ,  $p(\text{DM}|z)$ , is of both intrinsic interest, and is a nuisance factor in calculating the properties of the FRB population itself. Here, we use the method and parameters of Macquart et al. (2020). We model the DM of an FRB as

$$\text{DM} = \text{DM}_{\text{ISM}} + \text{DM}_{\text{halo}} + \text{DM}_{\text{cosmic}} + \text{DM}_{\text{host}}, \quad (1)$$

with respective contributions from the Milky Way’s interstellar medium (ISM), its halo, the cosmological distribution of ionized gas, and the FRB host. In this work, we divide this into an ‘extragalactic’ contribution,

$$\text{DM}_{\text{EG}} \equiv \text{DM}_{\text{cosmic}} + \text{DM}_{\text{host}}, \quad (2)$$

and a ‘local’ contribution,

$$\text{DM}_{\text{local}} \equiv \text{DM}_{\text{ISM}} + \text{DM}_{\text{halo}}. \quad (3)$$

The ‘local’ contribution is subtracted from FRB observations, and thus all comparisons between expectations and measurements are made in terms of  $\text{DM}_{\text{EG}}$ . This model slightly differs from that in Macquart et al. (2020), who model both  $\text{DM}_{\text{host}}$  and  $\text{DM}_{\text{halo}}$  using the same nuisance term,  $\text{DM}_x$ . The distinction becomes important at large redshifts.

### 2.1 $\text{DM}_{\text{ISM}}$

We use the NE2001 model (Cordes & Lazio 2002),<sup>1</sup> to estimate the Galactic contribution to dispersion measure. Since DM is an ingredient in the calculation of detection efficiency (see Section 4.2), the full integral for the FRB rate extends over the pointing direction as a function of Galactic coordinates, as discussed in Section A2 and equation (A1). Since most FRBs and FRB searches have been at high Galactic latitudes, however, we use the mean value  $\overline{\text{DM}}_{\text{ISM}}$  to calculate the sensitivity for each survey, while using the individual values  $\text{DM}_{\text{ISM}}$  when calculating FRB likelihoods.

### 2.2 $\text{DM}_{\text{halo}}$

The exact contribution of the Milky Way halo to DM is uncertain, with estimates of order  $10\text{--}80 \text{ pc cm}^{-3}$  (Prochaska & Zheng 2019; Keating & Pen 2020). FRBs have been observed down to a DM of little more than  $110 \text{ pc cm}^{-3}$  (CHIME/FRB Collaboration et al. 2019a) and  $114 \text{ pc cm}^{-3}$  (Shannon et al. 2018), favouring the middle of this range and consistent with current estimations based on the full set of observed DMs (Platts, Prochaska & Law 2020). We therefore use a value of  $\text{DM}_{\text{halo}} = 50 \text{ pc cm}^{-3}$  in our default model.

Deviations between our assumed values for  $\text{DM}_{\text{ISM}}$  and  $\text{DM}_{\text{halo}}$  will be absorbed into our model for the host galaxy contribution.

<sup>1</sup>Ben Bar-Or, J. Prochaska, available at <https://readthedocs.org/projects/ne2001/>

### 2.3 Cosmological DM

We caution that symbols  $E$ ,  $F$ , and  $\alpha$  have different definitions in this section than in the remainder of this work. The notation regarding  $\text{DM}_{\text{cosmic}}$  is derived from Macquart et al. (2020), and we preserve it for ease of reference to that work.

The ‘cosmological’ contribution to DM,  $\text{DM}_{\text{cosmic}}$ , can be understood as the DM incurred when an FRB is emitted at  $z$  at a random point in the Universe and propagates until the current epoch,  $z = 0$ . A parametrization based on detailed simulations (McQuinn 2014) is given in Macquart et al. (2020), as a function of burst redshift  $z$ , as

$$\text{DM}_{\text{cosmic}} = \langle \text{DM}_{\text{cosmic}} \rangle \Delta_{\text{DM}} \quad (4)$$

$$p(\text{DM}_{\text{cosmic}}|z) = \frac{p(\Delta_{\text{DM}}|z)}{\langle \text{DM}_{\text{cosmic}} \rangle}. \quad (5)$$

The expected value  $\langle \text{DM}_{\text{cosmic}} \rangle$  is calculated as per Ioka (2003) and Inoue (2004):

$$\langle \text{DM}_{\text{cosmic}} \rangle = \int_0^z \frac{c\bar{n}_e(z')dz'}{H_0(1+z')^2 E(z)}, \quad (6)$$

$$E(z) = \sqrt{\Omega_m(1+z')^3 + \Omega_\Lambda}, \quad (7)$$

using the mean density of ions,  $\bar{n}_e$ , and cosmological parameters relevant for the range  $0 \leq z \leq 5$ :  $H_0 = 67.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , and matter and dark energy densities  $\Omega_m = 0.315$  and  $\Omega_\Lambda = 0.685$  for a critical density Universe. See Macquart et al. (2020) and references contained therein for further details – cosmological parameters are taken from Planck Collaboration et al. (2018).

The probability of deviations from the mean,  $p(\Delta_{\text{DM}})$ , is given by

$$p(\Delta_{\text{DM}}) = A \Delta_{\text{DM}}^{-\beta} \exp \left[ -\frac{(\Delta_{\text{DM}}^{-\alpha} - C_0)^2}{2\alpha^2 \sigma_{\text{DM}}^2} \right], \quad (8)$$

with  $\alpha = 3$ ,  $\beta = 3$ , and  $C_0$  being numerically tuned such that the expectation value of the distribution is unity. The degree of feedback  $F$  is reflected in the choice of  $\sigma_{\text{DM}} = Fz^{-0.5}$ . In this work, we fix  $F = 0.32$  based on results of (Macquart et al. 2020).

The resulting distribution of  $\text{DM}_{\text{cosmic}}$ ,  $p(\text{DM}_{\text{cosmic}}|z)$  is shown in Fig. 1.

### 2.4 $\text{DM}_{\text{host}}$

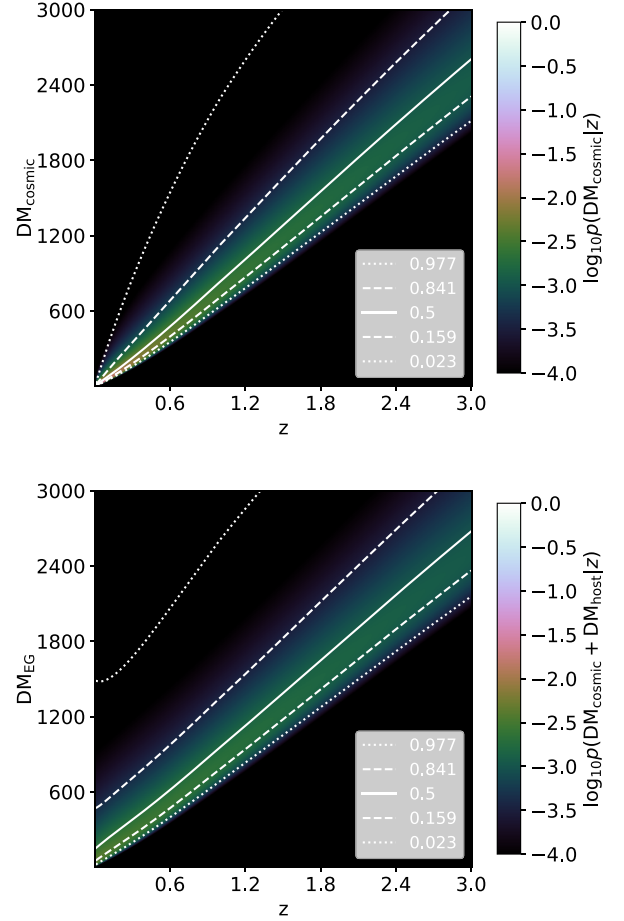
The contribution of the FRB host galaxy (including the local environs of the FRB itself) to DM is highly uncertain. Some FRBs, most notably FRB 121102 and FRB 190608, show a large excess DM beyond what is expected from cosmological and MW contributions, which cannot be explained by passage through an intervening galaxy along the line of sight (Spitler et al. 2014; Chatterjee et al. 2017; Hardy et al. 2017; Tendulkar et al. 2017; Chittidi et al. 2020). Yet as noted in Section 2.2, many FRBs do not allow for a great deal of excess DM. Macquart et al. (2020) generically model this large spread using a lognormal distribution

$$p(\text{DM}'_{\text{host}}) = \frac{1}{\text{DM}'_{\text{host}}} \frac{1}{\sigma_{\text{host}} \sqrt{2\pi}} \exp \left[ -\frac{(\log \text{DM}'_{\text{host}} - \mu_{\text{host}})^2}{2\sigma_{\text{host}}^2} \right]. \quad (9)$$

We also correct the host contribution for redshift via

$$\text{DM}_{\text{host}} = \frac{\text{DM}'_{\text{host}}}{1+z}. \quad (10)$$

In this work, we use  $\mu_{\text{host}}$  and  $\sigma_{\text{host}}$  as free parameters. Thus uncertainties in other DM contributions – including from our assumed value of feedback  $F$  – will be absorbed into these quantities.



**Figure 1.** Distribution  $p(\text{DM}|z)$  of observation-independent DM parameters,  $\text{DM}_{\text{cosmic}}$  only (top), and also including the best-fitting distribution for  $\text{DM}_{\text{host}}$  derived in James et al. (2021a) (bottom), as a function of redshift  $z$ , showing contours.

### 2.5 The intrinsic $z$ -DM grid

The probability distribution of observation-independent factors,  $\text{DM}_{\text{cosmic}} + \text{DM}_{\text{host}}$ , is given in Fig. 1. In this work, a linear grid in both  $z$  and DM space is used, with 1200 DM points spaced from 0–7000 in intervals of  $5 \text{ pc cm}^{-3}$ , and 500 in redshift from 0.01–5. FRBs have their nominal local contributions,  $\text{DM}_{\text{local}}$  subtracted from their observed values of DM prior to evaluating their likelihood on this grid. In this model, only a small fraction of FRBs will have a DM very much larger than the mean. In particular, for  $z < 1.5$ , the majority of the spread in DM comes from the host galaxy, rather than the cosmological contribution.

## 3 FRB POPULATION

### 3.1 Energetics

Our model of the FRB population  $\Phi$  is consistent with that used in the literature. We adopt a power-law distribution of burst energies  $E$  between  $E_{\text{min}}$  and  $E_{\text{max}}$  with integral slope  $\gamma$ . We use ‘burst energy’ as the isotropic equivalent energy at 1.3 GHz, and assume an effective bandwidth of 1 GHz when converting between ‘per Hz’ and total quantities. If FRBs are beamed into a characteristic solid angle  $\Omega_{\text{FRB}}$ , then all energies in this work should be scaled down by a factor

$\Omega_{\text{FRB}}(4\pi)^{-1}$  – but the true FRB rate will increase by an identical factor.

In this model, the probability of a burst occurring above a threshold  $E_{\text{th}}$  is a piecewise function

$$\begin{aligned} p(E > E_{\text{th}}) &= 1 \quad (E_{\text{th}} < E_{\text{min}}) \\ p(E > E_{\text{th}}) &= 0 \quad (E_{\text{th}} > E_{\text{max}}) \\ p(E > E_{\text{th}}) &= \frac{\left(\frac{E_{\text{th}}}{E_{\text{min}}}\right)^\gamma - \left(\frac{E_{\text{max}}}{E_{\text{min}}}\right)^\gamma}{1 - \left(\frac{E_{\text{max}}}{E_{\text{min}}}\right)^\gamma} \quad \text{otherwise.} \end{aligned} \quad (11)$$

Observations show no evidence for a minimum FRB energy, and indeed the event rate is generally insensitive to  $E_{\text{min}}$  for  $\gamma > -1.5$  (Macquart & Ekers 2018b). Thus, we use a very low value of  $E_{\text{min}} = 10^{30}$  erg, which is several orders of magnitude below the minimum detected burst energy of all known FRBs – including SGR 1935+2154 at  $10^{34}$ – $10^{35}$  erg (The Chime/Frb Collaboration 2020; Bochenek et al. 2020) – and treat this as a fixed parameter. We re-examine this assumption in Section 8.3.

Several other authors have used a Schechter function to model the FRB luminosity function (Lu & Piro 2019; Luo et al. 2020), which adds an exponential cut-off of the form  $\exp(-E/E_{\text{max}})$  to the luminosity function of equation (35). This is neither observationally nor theoretically motivated – the form of the Schechter function is used to model galaxy luminosities – and adds computational complexity, but avoids the unphysicality of a sharp cut-off. A preliminary investigation showed that using the Schechter function provided effectively identical fits to the data. We therefore advocate for the simple power-law model out of simplicity.

To calculate  $E_{\text{th}}$ , we convert between FRB energy  $E$  and observable fluence  $F$  using

$$E(F) = \frac{4\pi D_L^2(z)}{(1+z)^{2+\alpha}} \Delta\nu F, \quad (12)$$

where  $\alpha$  is the spectral index ( $F \propto \nu^\alpha$ ), and  $\Delta\nu$  the bandwidth (here we use 1 GHz). Macquart et al. (2019) fit  $\alpha$  to 23 FRBs detected by ASKAP in Fly’s Eye mode, finding  $\alpha = -1.5^{+0.2}_{-0.3}$ . Thus, we use a default value of  $\alpha = -1.5$ . We return to the interpretation of  $\alpha$  shortly.

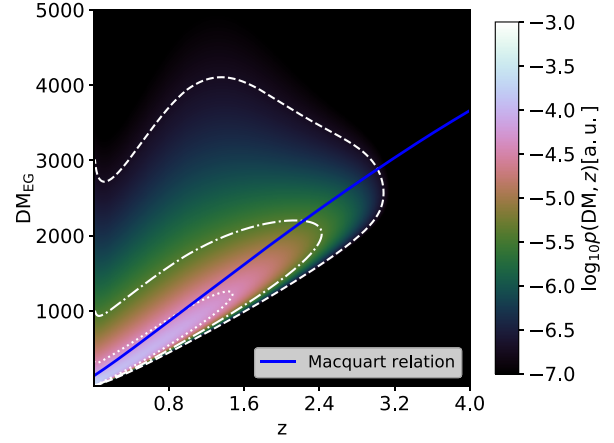
### 3.2 Population evolution

The rate of FRBs per comoving volume will likely be a function of redshift. While FRB host galaxies do not appear to be drawn from a population sampled proportionally to their star-forming activity (Safarzadeh et al. 2020), they certainly are not exclusively associated with very old galaxies in which star-forming activity has ceased (Bhandari et al. 2020a; Heintz et al. 2020). We therefore adopt the approach of Macquart & Ekers (2018b) and generically model the population evolution of FRBs as being to some power of the star-formation rate, i.e.

$$\Phi(z) = \frac{\Phi_0}{1+z} \left( \frac{\text{SFR}(z)}{\text{SFR}(0)} \right)^n, \quad (13)$$

with  $\Phi_0$  – and hence  $\Phi(z)$  – taking the units of bursts per proper time per comoving volume, i.e. bursts  $\text{yr}^{-1} \text{Mpc}^{-3}$ . The factor of  $(1+z)^{-1}$  converts between proper time in the emission and observer frames. We take  $\text{SFR}(z)$  from Madau & Dickinson (2014)

$$\text{SFR}(z) = 1.0025738 \frac{(1+z)^{2.7}}{1 + \left(\frac{1+z}{2.9}\right)^{5.6}}. \quad (14)$$



**Figure 2.** Relative rate of FRB detections in  $z$ – $\text{DM}$  space when ignoring beamshape and burst width effects for the best-fitting FRB population parameters presented in James et al. (2021a). All observational biases from Section 4 have been ignored, and a constant detection threshold of 1 Jy ms is used. The Macquart relation, approximated by  $\text{DM} = \langle \text{DM}_{\text{cosmic}} \rangle + \exp(\mu_{\text{host}})$ , is also shown. The contours represent 50 per cent (dotted), 90 per cent (dash-dotted), and 99 per cent (dashed) of the FRB population.

This model is useful in that  $n$  can be scaled as a smooth parameter. However, it does not accurately model the source evolution should FRB progenitors originate from e.g. binary mergers with long delay times, as investigated by Cao, Yu & Zhou (2018).

The total FRB rate in a given redshift interval  $dz$  and sky area  $d\Omega$  will also be proportional to the total comoving volume  $dV$ ,

$$\frac{dV}{d\Omega dz} = D_H \frac{(1+z)^2 D_A^2(z)}{E(z)}, \quad (15)$$

for angular diameter distance  $D_A$ , Hubble distance  $D_H = c/H_0$ , and scale factor  $E(z)$  from equation (7).

Applying this model with the best-fitting FRB population parameters derived in our accompanying work to the  $\text{DM}$ -distribution of Section 2 with a nominal threshold of 1 Jy ms produces the distribution of FRBs shown in Fig. 2. This ignores the important observational biases to be introduced in Section 4, and hence a quantitative analysis of the implications are left to Section 8. However, it is clear that while at least 90 per cent of FRBs will follow a 1–1  $\text{DM}$ – $z$  relation (the Macquart relation; Macquart et al. 2020), a significant minority will lie well above the  $\text{DM}$ – $z$  curve. Indeed the highest  $\text{DM}$  events in a large sample are not likely to be the most distant. Consider  $\text{DM} \geq 3000 \text{ pc cm}^{-3}$  in Fig. 2. Such events can be produced by  $z \sim 3.2$  lying on the Macquart relation. However, they must be the most intrinsically luminous FRBs to be detectable. At  $z \sim 1.6$ , observations probe a factor of  $\sim 4$  further down the energy distribution, allowing a greater number of events to be visible, and its high- $\text{DM}$  tail may dominate the  $\text{DM} \geq 3000 \text{ pc cm}^{-3}$  event rate. This effect becomes more important for steeper luminosity distributions (large negative  $\gamma$ ) – this plot uses  $\gamma = -1.2$ .

### 3.3 Interpretation of $\alpha$

Many FRBs have a limited band occupancy (originally noted for FRB 121102 by Law et al. 2017), in which case the notion of a spectral index for an individual FRB has little meaning. In this case, the results of Macquart et al. (2019) can be interpreted as meaning either that there are more low-frequency FRBs, or that low-frequency FRBs are stronger. For an experiment with bandwidth similar to or



less than that of the FRB bandwidth (which is the case with the data used here – see Section 6), the latter interpretation behaves identically to that of broad-band bursts defined by a spectral index. However, the interpretation of an FRB population with a frequency-dependent rate does not. We denote this interpretation as the ‘rate interpretation’ of  $\alpha$ , and that of Macquart et al. (2019) as the ‘spectral index’ interpretation.

Under the spectral index interpretation of  $F(\nu) \sim \nu^\alpha$ , a negative  $\alpha$  increases the detection threshold  $E_{\text{th}}$  at high  $z$  due to the  $k$ -correction factor of  $(1+z)^{-\alpha}$  through equation (12). This in turn decreases the rate by a factor  $(1+z)^{\nu\alpha}$  when  $E_{\text{th}} < E_{\text{max}}$  through equation (11). Under the rate interpretation, the FRB population itself behaves as  $\Phi \sim \Phi(z, \nu) = \Phi(z)\nu^\alpha$ , and the  $k$ -correction therefore directly changes the rate, adding an additional factor of  $(1+z)^\alpha$  to equation (13). Therefore, when  $\gamma = -1$ , and  $E_{\text{th}} < E_{\text{max}}$ , the two interpretations are identical. The situation becomes less simple near  $E_{\text{max}}$ , which is frequency dependent under the spectral index interpretation, and constant under the rate interpretation – and the true behaviour may be more complicated than either result.

Ultimately, we expect further observational data to be required to discriminate between the two scenarios, and consider both equally plausible for the time being. In this work, we present results using the spectral-index interpretation, but give additional data for the rate interpretation when constraining FRB population parameters.

#### 4 DETECTION THRESHOLD – OBSERVATIONAL BIASES

FRB surveys usually calculate the fluence threshold above which FRBs would be detected using the radiometer equation, referenced to a 1 ms duration burst, using the sensitivity of the telescope at beam centre. This readily calculable value represents an unrealistic ideal. Bursts of longer duration will be harder to detect due to increased noise, while those viewed away from beam centre will be seen at less sensitivity. Furthermore, incoherent dedispersion searches will not perfectly match the shape of an FRB to the time–frequency resolution of the search, resulting in a lower detection efficiency.

In this work, we model the effective fluence threshold  $F_{\text{th}}$  as a function of nominal fluence threshold at 1 ms  $F_1$ , beam sensitivity  $B$  (normalized to a maximum of 1), and an efficiency factor due to burst duration,  $\eta$ , as

$$F_{\text{th}} = \frac{F_1}{\eta B}. \quad (16)$$

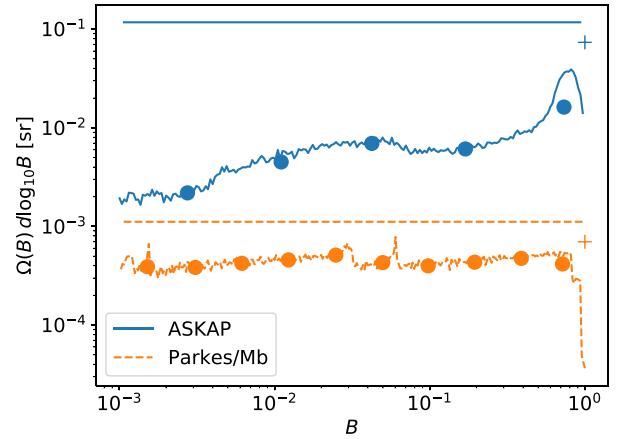
This results in a theoretical distribution of bursts in  $z$ -DM space,  $p(z, \text{DM})$ , as

$$p(z, \text{DM}) = \int dB \int d\eta p(z, \text{DM} | F_{\text{th}}(\eta, B)) \Omega(B) p(\eta), \quad (17)$$

where  $p(z, \text{DM} | F_{\text{th}})$  is the distribution at a fixed threshold (Fig. 2),  $\Omega(B)$  is the region of sky at which the beam sensitivity is  $B$ , and  $p(\eta)$  is the probability that burst properties lead to a total detection efficiency  $\eta$ . The effects of these two factors are investigated in the following sections.

##### 4.1 Beamshape

A telescope’s beamshape is usually represented as the relative sensitivity  $B$  as a function of the sky position  $\Omega$  relative to boresight,  $B(\Omega)$ , such that  $B(0) = 1$ . The beamshape is often approximated as a Gaussian or Airy function, although precise measurements of  $B$  can become important when attempting to localize FRBs detected



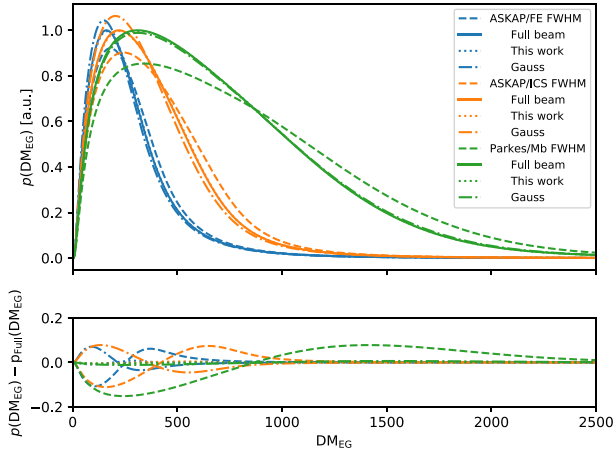
**Figure 3.** Inverse beamshape  $\Omega(B)$  as a function of beam sensitivity  $B$  for ASKAP (closepack36 configuration at 1.296 GHz) and the Parkes Multibeam (1.382 GHz). Shown are the best measurements (lines), FWHM approximation (crosses), the numerical approximations used in this work (circles), and the Gaussian beamshape (flat horizontal lines). The points for the FWHM and numerical approximations represent  $\delta$  functions – in the FWHM case, the true amplitude is shown, while in the numerical approximations, the amplitudes are renormalized by their spacing in  $\log B$  for comparability with the best measurements.

in multiple beams, or estimating the relative rate of single- versus multiple-beam detections (Vedantham et al. 2016; Macquart & Ekers 2018b).

For the purpose of estimating the number of FRBs detected, however, the ‘inverse beamshape’,  $\Omega(B)$ , which describes the amount of sky  $\Omega$  viewed at any given sensitivity  $B$ , becomes more relevant (James et al. 2019a). Most calculations of FRB rates have characterized a telescope beam as viewing out to the FWHM at full sensitivity, i.e.  $\Omega(B) = \Omega_{\text{FWHM}} \delta(B - 1)$  (e.g. Thornton et al. 2013; Bhandari et al. 2018). Others have used a Gaussian approximation for the beamshape (e.g. Lawrence et al. 2017), which is equivalent to  $\Omega(B) d \log B = \text{const}$ . We here analyse the sufficiency of these approximations, using for the Gaussian approximation  $\sigma = (\text{FWHM}/2)(2 \log 2)^{0.5}$ , where the full width at half maximum (FWHM) assumes an airy disc, i.e.  $\text{HPBW} = 1.22 \lambda / D$  for wavelength at central frequency  $\lambda$  and dish diameter  $D$ .

ASKAP FRB observations have varied the observation frequency and configuration of beams formed from ASKAP’s phased array feeds (PAFs). However, the majority of both fly’s eye and incoherent sum observations have used the ‘closepack36’ configuration at a central frequency of 1.296 GHz. We therefore use the beamshape derived in James et al. (2019a). In the case of the Parkes multibeam, we use a central frequency of 1.382 GHz, and the simulations of K. Bannister (published as Vedantham et al. 2016) and A. Dunning (referenced as ‘private communication’ by Macquart & Ekers 2018a), which produce equivalent results for  $\Omega(B)$ . This also allows us to conclude that while the ability to localize FRBs detected in multiple beams may be limited by systematic uncertainties in the beamshape (Macquart & Ekers 2018a), the inverse beamshape  $B(\Omega)$  is robust against such certainties, since it does not care about where on the sky any given patch of sensitivity is located.

Fig. 3 shows the resulting ‘inverse beamshape’ function  $\Omega(B)$ . This is compared to the equivalent  $\Omega(B)$  when using the Gaussian and  $\Omega_{\text{FWHM}}$  approximations. Since implementing the full function  $\Omega(B)$  in the calculation of the  $z$ -DM distribution is numerically expensive, we investigate the accuracy of reducing  $\Omega(B)$  to a small number of values. A set of such values is also shown in Fig. 3. The accuracy



**Figure 4.** Top: dependence of expected FRB DM distributions ( $\text{pc cm}^{-3}$ ) on beamshape models for three different FRB surveys. The beamshape models considered are shown in Fig. 3: the ‘FWHM’ approximation, the ‘full’ beamshape  $\Omega(B)$ , a numerical approximation used in ‘this work’, and a Gaussian beamshape. All curves for each telescope are normalized such that the distribution for the full beamshapes peaks at unity. Bottom: the difference between the ‘full’ beamshape and that found when using the ‘FWHM’, ‘approx.’, and Gaussian approximations.

**Table 1.** Percentage errors in the total FRB rate and mean DM value,  $\overline{\text{DM}}$ , and maximum deviation  $|\delta\text{DM}|_{\text{max}}$ , when using different beam approximations when compared to that found for the full beamshape function  $\Omega(B)$ , for each of three FRB surveys considered.

Survey	Approximation	Rate	$\overline{\text{DM}}$	$ \delta\text{DM} _{\text{max}}$
ASKAP/FE	FWHM	+76	+6	0.1
	This work	+17	+0.2	0.009
	Gauss	+513	−3	0.07
ASKAP/ICS	FWHM	+73	+8	0.11
	This work	+5.8	+0.2	0.005
	Gauss	+440	−4	0.08
Parkes/Mb	FWHM	+20	+14	0.15
	This work	+2.8	+0.2	0.005
	Gauss	+160	+1	0.011

of all approximations is assessed against the full beamshape, by comparing the total predicted number of events and the mean value of DM to those calculated for the full beamshape, and also assessing the maximum difference between the  $p(\text{DM})$  curves, as shown in Fig. 4. Table 1 lists the resulting errors. This is evaluated for the best-fitting set of parameters found in our accompanying paper – however, a brief investigation has shown that results are not sensitive to the assumed parameters within a reasonable range.

We find that using five values of  $B$  for ASKAP, and ten for Parkes, achieves an approximate DM distribution with an error in  $p(\text{DM})$  of less than 1 per cent, and an error in the mean value  $\overline{\text{DM}}$  of only 0.2 per cent. In contrast, using the FWHM approximation, which is the standard in the current literature, results in  $\mathcal{O} \sim 10$  per cent deviations in the DM distribution and its mean value, and pushes the mean towards higher values. The total expected detection rate found when using the FWHM approximation is almost double that when found when using the full ASKAP beam, but there is only a 20 per cent excess for Parkes. When assuming Gaussian beams, a huge excess in the total rate of ASKAP bursts is predicted, since this does not account for the closely packed, and thus overlapping,

beams. We note that uncertainties in the true ASKAP beamshape due to the calibration procedure (see James et al. 2019a) are less than the errors introduced by our numerical approximation.

In the case of Parkes/Mb, the excess rate when using a Gaussian beam is due to outer beams being less sensitive than the central beam at which the sensitivity is usually calculated. However, the Gaussian beam approximation accurately estimates  $\overline{\text{DM}}$  and the shape of  $p(\text{DM})$  distribution. This suggests that even very complex beamshapes, such as that of the Canadian Hydrogen Intensity Mapping Experiment (CHIME), could be included in our model in a relatively simplified manner.

## 4.2 Detection efficiency

We model the threshold at which an FRB of fixed fluence  $F$  can be detected as scaling with the square root of its effective width,  $w_{\text{eff}}$ , relative to the nominal width of 1 ms, using an efficiency factor  $\eta$ :

$$\eta \equiv \frac{F_1}{F_{\text{th}}(w_{\text{eff}})} \quad (18)$$

$$= \left( \frac{w_{\text{eff}}}{1 \text{ ms}} \right)^{-0.5}. \quad (19)$$

The effective width is modelled as per Cordes & McLaughlin (2003), being a function of its intrinsic duration  $w_{\text{int}}$ , scattered width  $w_{\text{scat}}$ , DM smearing within each frequency channel  $w_{\text{DM}}$ , and the time-resolution of the search  $w_{\text{samp}}$ :

$$w_{\text{eff}} = \sqrt{w_{\text{int}}^2 + w_{\text{scat}}^2 + w_{\text{DM}}^2 + w_{\text{samp}}^2}. \quad (20)$$

Often, the scattered width and intrinsic width are indistinguishable, and their separation only becomes important for telescopes observing at different frequencies. We therefore define the ‘incident’ width,  $w_{\text{inc}}$ , as

$$w_{\text{inc}}^2 = w_{\text{int}}^2 + w_{\text{scat}}^2. \quad (21)$$

An alternative model is presented by Arcus et al. (2021), which is based on fits to simulated ASKAP and Parkes FRBs. Since it is not clear how the fit parameters translate to the general case, and because we wish to present a broadly scalable model, we do not use their formulation. We remark rather that the widely used model of equation (19) should be investigated further.

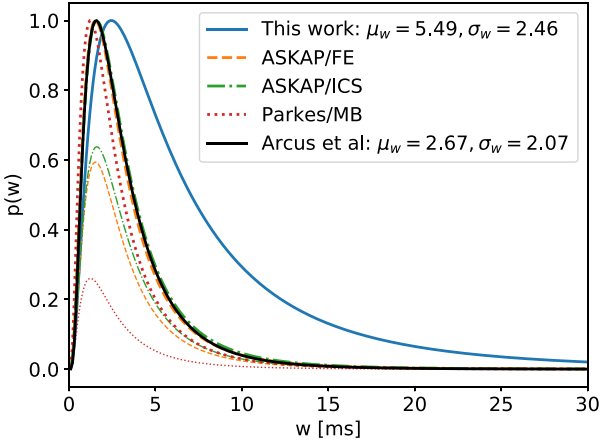
In order to model the distribution of  $w_{\text{inc}}$ ,  $p(w_{\text{inc}})$ , we use a lognormal distribution,

$$p(w_{\text{inc}})dw_{\text{inc}} = \frac{w_{\text{inc}}}{\log \sigma_w (2\pi)^{0.5}} e^{-\frac{1}{2} \left( \frac{\log w_{\text{inc}} - \log \mu_w}{\log \sigma_w} \right)^2}. \quad (22)$$

We do not include any DM or  $z$  dependence in the width distribution – see Appendix A3 for further discussion on this topic.

Unlike Luo et al. (2020), we do not include fits of the model parameters  $\mu_w$  and  $\sigma_w$  as part of our general fitting process. Rather, we use the low correlation between  $\mu_w$ ,  $\sigma_w$ , and other parameters to first fit for these values using a preliminary parameter set, and then check that the fit is still valid for the final parameter set presented in our companion paper (James et al. 2021a).

Arcus et al. (2021) use equation (21) to model the observed width distribution  $p(w_{\text{inc}})$  of ASKAP and Parkes FRBs, finding  $\mu_w = 2.67$  ms and  $\sigma_w = 2.07$ . We instead use equation (21) to model the intrinsic width distribution, and vary  $\mu_w$ ,  $\sigma_w$  until the simulated width distribution of the ASKAP/FE survey matches the parametrization of observed widths by Arcus et al. (2021). We obtain  $\mu_w = 5.49$  and  $\sigma_w = 2.46$ , and then proceed to use these values in further calculations to optimize FRB population parameters. This is shown



**Figure 5.** The effect of intrinsic burst width on survey sensitivity. Blue: intrinsic distribution of burst widths  $w_{\text{int}}$  (peak set to unity). Orange/green/red: the simulated width distributions expected from the three FRB surveys after accounting for observational bias. Two normalizations are shown: the upper curves have been normalized to a peak probability of unity, for comparison with the fit to observed values by Arcus et al. (2021) (shown in black); while the lower curves are normalized relative to the rate for  $w = 0$ , and thus illustrate the relative reduction in event rate as a function of intrinsic width. The upper curves for ASKAP/FE and ASKAP/ICS coincide almost identically with the results of Arcus et al. (2021).

**Table 2.** The effects of different assumptions on the intrinsic burst width distribution at 1.3 GHz for three FRB surveys. Given are the total detection rate (normalized to a rate of unity at  $w = 0$ ), and mean DM  $\overline{\text{DM}}$ . The width distributions are parametrized via equation (21), with given values of  $\mu_w$  and  $\sigma_w$  corresponding to no intrinsic width, the observed width distribution of Arcus et al. (2021), when accounting for observational bias to find the intrinsic distribution, and when numerically approximating ( $\sim$ ) the intrinsic distribution with a few characteristic values.

Parameter	$\mu_w$	$\sigma_w$	ASKAP/FE	ASKAP/ICS	Parkes
Rate	0	0	1	1	1
	2.67	2.07	0.46	0.51	0.20
	5.49	2.46	0.27	0.30	0.11
	$\sim 5.49$	$\sim 2.46$	0.27	0.30	0.11
$\overline{\text{DM}}$	0	0	263	371	488
	2.67	2.07	286	397	724
	5.49	2.46	293	401	726
	$\sim 5.49$	$\sim 2.46$	292	400	724

in Fig. 5. The effect on the predicted DM distributions is given in Fig. 6; the effect on total rates, and mean estimated DM values, given in Table 2. Finally, we re-evaluate the fits using the final optimal set of parameters we present in our companion paper (James et al. 2021a).

Comparing the intrinsic (blue) and observed (black, coloured) distributions in Fig. 5, modelling the intrinsic FRB rate, and accounting for observational bias, correctly reproduces the observed FRB width distribution as estimated by Arcus et al. (2021). The effect of observational bias is clearly reflected in the total expected FRB rate, with high-width bursts much less likely to be detected. The magnitude of this effect, shown in Table 2, depends on the significance of the other terms in equation (19). When the time resolution is poor ( $w_{\text{samp}}$  large), the effect of a large intrinsic width is less – thus the reduction in rate for ASKAP is less than that for Parkes surveys. A second-order effect is that more-sensitive surveys, which probe further into

the Universe and see FRBs with on-average higher DMs, are also less sensitive to  $w_{\text{int}}$ , since  $w_{\text{DM}}$  is larger. Hence, the reduction in rate for ASKAP/ICS is slightly lower than for ASKAP/FE, while the greatest effect is for the Parkes/Mb survey, where the intrinsic FRB width reduces the number of detected FRBs by a factor of 10.

A consequence of this is that the true number of high-width FRBs will be very difficult to estimate, so that our lognormal model is effectively untested beyond  $w = 10$  ms. Thus while we estimate that ASKAP/FE and ASKAP/ICS miss  $\sim 70$  per cent of FRBs due to their intrinsic width, and Parkes/Mb 90 per cent, we do not consider this quantitatively reliable – there may be virtually any number of high-width events remaining to be detected. We note that The CHIME/FRB Collaboration (2021) find evidence for an intrinsically large population of FRBs with scattering width at 600 MHz between 10 and 100 ms (approximately between 4 and 40 ms after scaling to ASKAP/Parkes frequencies using  $\tau_{\text{scat}} \sim \nu^{-4}$ ) which is also poorly constrained due to selection effects.

Whatever the lost rate, losses will preferentially arise from the nearby Universe where  $w_{\text{DM}}$  is low. Including the width distribution therefore increases the mean expected DM,  $\overline{\text{DM}}$ . This effect is small ( $\sim 10$  per cent) for ASKAP observations, but more significant for Parkes ( $\sim 30$  per cent).

Finally, we note that while including the width distribution is clearly very important, the details matter less. Using the parameters of Arcus et al. (2021), or approximating the true distribution with a small number of points for computational efficiency, produces an almost identical value of  $\overline{\text{DM}}$ . This also means that the loss of efficiency to high-width bursts for the HEIMDALL<sup>2</sup> search pipeline found by Gupta et al. (2021) – which is commonly used in Parkes FRB searches – is insignificant to our modelling.

### 4.3 Numerical implementation

The integrals over  $B$  and  $\eta$  in equation (17) are numerically expensive. Furthermore, we have shown above that approximating the beamshape with 5–10 values, and the width distribution with five, allows for a very good approximation to the continuous distributions. Therefore, we approximate these continuous distributions with these discrete distributions, i.e.

$$\Omega(B) \sim \sum_i^{N_B} \Omega_i \delta(B - B_i) \quad (23)$$

$$p(\eta(\text{DM}, w)) \sim \sum_i^{N_w} p_i \delta(\eta(\text{DM}, w_w) - \eta(\text{DM}, w_i)), \quad (24)$$

recalling that  $\eta$  is a function of both DM and  $w$  through equation (18) and (19). Therefore, the  $z$ -DM distribution of equation (17) becomes a weighted sum over individual distributions at fixed observational thresholds,

$$p(z, \text{DM}) = \sum_{i=1}^{N_B} \sum_{j=1}^{N_w} p(z, \text{DM}, F_{\text{th}}(B_i, \eta(\text{DM}, w_j))) \Omega_i p_j. \quad (25)$$

This implementation is available at GITHUB (Prochaska et al. 2019a; James, Prochaska & Ghosh 2021b).

## 5 METHODOLOGY

The ingredients described above are implemented in PYTHON. Here, we describe its effective implementation, which is identical in

<sup>2</sup><https://sourceforge.net/p/heimdall-astro/wiki/Home/>

approach to the method proposed by Connor (2019) and implemented by Luo et al. (2020), even if the exact details differ.

We consider FRB data from some number of independent FRB surveys. The total likelihood  $\mathcal{L}$  of the outcome of multiple independent FRB surveys is simply a product of their individual likelihoods,  $\mathcal{L}_i$ ,

$$\mathcal{L} = \prod_i^{N_{\text{surveys}}} \mathcal{L}_i, \quad (26)$$

for surveys  $i = 1 \dots N_{\text{surveys}}$ . We describe  $\mathcal{L}_i$  as the product of three independent terms:

$$\mathcal{L}_i = p_n(N_i) \prod_{j=1}^{N_i} p_{\text{dmz}}(\text{DM}_j, z_j) p_s(s_j | \text{DM}_j, z_j). \quad (27)$$

Here,  $p_n$  is the probability of detecting  $N_i$  FRBs in survey  $i$ , while  $p_{\text{dmz}}$  is the likelihood of the  $j$ th FRB from survey  $i$  being detected with dispersion measure  $\text{DM}_j$  and, when applicable, at redshift  $z_j$ . We also include the probability  $p_s$  of an FRB being detected with fluence  $F$  a factor  $s$  above the fluence threshold  $F_{\text{th}}$  given it was observed with properties  $\text{DM}_j, z_j$ . Each of these terms is described independently below.

It is also possible to separate further terms in equation (26). As noted by Vedantham et al. (2016), for a multibeam instrument, the relative likelihood of a single- versus multibeam detection, and the relative likelihood of detection in different beams of varying sensitivity, are functions of the FRB fluence distribution. Such measures are only relevant when the true FRB fluence  $F$  cannot be resolved, as is common with FRB detections by the Parkes multibeam. When  $F$  can be reconstructed, as is the case for CRAFT FRB detections with ASKAP (Shannon et al. 2018), then the survey acceptance to that particular FRB can be calculated exactly, and  $p(s)$  in equation (26) can be written in terms of  $p(F)$ . Similarly, we will not include the observed value of an FRB's width when evaluating  $p_{\text{dmz}}$  and  $p_s$ , with only the overall width distribution being accounted for. We consider that adding such terms will yield only a small increase in analytical power for resolving the FRB population, at the cost of a large increase in complexity. Thus they are ignored – however we do acknowledge that we are discarding a small amount of information by doing so.

### 5.1 Probability of $N$ detections, $p_n(N)$

Ignoring the correlations caused by repeating FRBs (see the discussion in Section 5.4), the total number of observed FRBs in survey  $i$ ,  $N_i$  comes from a Poisson distribution,

$$p_n(N_i) = \frac{\langle N_i \rangle^{N_i} \exp(-\langle N_i \rangle)}{N_i!}, \quad (28)$$

where  $\langle N_i \rangle$  is the expectation value of  $N_i$ . The calculation of  $\langle N_i \rangle$  is the heart of the problem that we address in this work, since it must necessarily incorporate all relevant properties that affect the detection rate.

Combining the dependencies in Sections 2–4,  $\langle N_i \rangle$  is calculated as

$$\langle N_i \rangle = T_i R_i, \quad (29)$$

$$R_i = \int dz \Phi(z) \frac{dV(z)}{d\Omega dz} \int d\text{DM} p(\text{DM}|z) \times \int dB \Omega(B) \int dw p(w) p(E > E_{\text{th}}(B, w, z, \text{DM})). \quad (30)$$

Here,  $T_i$  is the survey duration, which multiplies a rate  $R_i$  to produce the total expected number of bursts.  $\Phi(z)$  is the FRB source evolution function (equation (13)),  $dV(z)/d\Omega/dz$  is the comoving volume per steradian per redshift interval from equation (15),  $p(\text{DM}|z)$  is the extragalactic DM distribution found by convolving equation (5) and (9) (shown in Fig. 1);  $\Omega(B)$  is the beamshape discussed in Section 4.1 and approximated as per equation (22);  $p(w)$  is the width distribution of equation (21), discretized as per equation (23); and  $p(E > E_{\text{th}})$  is the cumulative energy function of equation (11). The dependency of this threshold  $E_{\text{th}}$  on the parameters ( $B, w, z, \text{DM}$ ) is encapsulated in equations (12), (16), and (18).

For some surveys, no controlled survey time  $T_i$  is available, and this term is simply set to unity in equation (26). However,  $R_i$  can be calculated regardless of knowledge of  $T_i$ . For those surveys with known  $T_i$ , the most likely value of the lead constant in the population function of equation (14),  $\Phi_0$ , can be estimated without recalculating the integral of equation (29).

### 5.2 Calculating $p_{\text{dmz}}$

The probability of an FRB being observed with a given dispersion measure DM and redshift  $z$  is given by the appropriate integrand of equation (29)

$$p_{\text{dmz}} = R_i^{-1} \Phi(z) \frac{dV(z)}{d\Omega dz} p(\text{DM}|z) \int dB \Omega(B) \int dw p(w) p(E > E_{\text{th}}(B, w, z, \text{DM})). \quad (31)$$

For FRBs with no measured host redshift, the relevant quantity is

$$p_{\text{dm}} = \int p_{\text{dmz}} dz, \quad (32)$$

and  $p_{\text{dm}}$  replaces  $p_{\text{dmz}}$  in equation (26). The rate  $R_i$  is used as a normalizing factor in equation (30), so that

$$\int \int p_{\text{dmz}} d\text{DM} dz = \int p_{\text{dm}} d\text{DM} = 1. \quad (33)$$

The shape of  $p_{\text{dmz}}$  in  $z$ –DM space is a primary quantity of interest in this work.

### 5.3 Calculating $p_s(s)$

The measured fluence  $F$  of an FRB also holds information on the FRB population. However, in many telescope systems – and notably for Parkes (Macquart & Ekers 2018a) –  $F$  is not directly measured, since the location of the FRB in the beam is not known. Furthermore, we are interested in  $p(F|F_{\text{th}})$ , i.e. the probability of measuring  $F$  given an FRB has been detected at threshold  $F_{\text{th}}$ , which itself has complex dependency through equation (16).

This difficulty can be overcome by noting that the signal-to-noise ratio, SNR, is a readily observable parameter for an FRB, and most FRB-hunting systems use a well-defined threshold SNR,  $\text{SNR}_{\text{th}}$ , to distinguish FRBs from noise. As per James et al. (2019b), who base their work on Crawford, Jauncey & Murdoch (1970), we define

$$s \equiv \frac{\text{SNR}}{\text{SNR}_{\text{th}}} \quad (34)$$

$$= \frac{F}{F_{\text{th}}}, \quad (35)$$

where  $F_{\text{th}}$  is the fluence threshold to that FRB. As detailed in Section 4,  $F_{\text{th}}$  is a function of the burst DM, width, and the location in which it is observed by the telescope's beam, so that neither



term in equation (33) is known. However, the ratio is preserved in the measurable quantities of equation (33), making  $s$  a very useful observable.

The probability  $p(s)$  of observing  $s$  in the range  $s$  to  $s + ds$  given that an FRB has already been observed is

$$p_s(s|DM, z) = \frac{R_i^{-1}}{p_{dmz}} \Phi(z) \frac{dV(z)}{d\Omega dz} p(DM|z) \int dB \Omega(B) \int dw p(w) \frac{dp(sF_{th})}{ds},$$

$$p_s(s|DM) = \frac{R_i^{-1}}{p_{dm}} \int \Phi(z) \frac{dV(z)}{d\Omega dz} p(DM|z) dz \int dB \Omega(B) \int dw p(w) \frac{dp(sF_{th})}{ds} \quad (36)$$

for localized and unlocalized FRBs, respectively. In the integrands,  $p(sF_{th})$  is the probability of detecting a fluence  $F = sF_{th}$  in an interval between  $s$  and  $s + ds$ , where  $F_{th}$  depends on  $B$ ,  $w$ , and DM as per equation (16).

The probability  $p(sF_{th})$  can be found from equation (11). Given that such an FRB has been observed at all, the integral distribution of  $E$  given that an FRB has been detected can be found by replacing  $E_{th}$  with  $E$  and  $E_{min}$  with  $E_x$ . Differentiating by  $E$  produces the probability amplitude

$$\frac{dp(E_{obs}|E_{th})}{dE} = \frac{\gamma}{E_{th}} \frac{\left(\frac{E}{E_{th}}\right)^{\gamma-1}}{1 - \left(\frac{E_{max}}{E_{th}}\right)^{\gamma}}$$

$$= \frac{\gamma}{E_{th}} \frac{s^{\gamma-1}}{1 - \left(\frac{E_{max}}{E_{th}}\right)^{\gamma}}. \quad (37)$$

The value of  $E_{th}$  can be found as a function of  $z$  by inserting  $F_{th}$  into equation (12); thus  $p(sF_{th})$  is equivalent to  $p(sE_{th}(z, F_{th}))$ . Relating  $dp/dE$  from equation (35) to the required  $dp/ds$  of equation (34) produces

$$\frac{dp(E_{obs}|E_{th})}{ds} = \frac{dp(E_{obs}|E_{th})}{dE} \frac{dE}{dF} \frac{dF}{ds} \quad (38)$$

$$= \frac{dp(E_{obs}|E_{th})}{dE} \frac{E_{th}}{F_{th}} F_{th} \quad (39)$$

$$= \gamma \frac{s^{\gamma-1}}{1 - \left(\frac{E_{max}}{E_{th}}\right)^{\gamma}}. \quad (40)$$

Inserting equation (36) into equation (34), and integrating over the appropriate dimensions, produces the desired  $p_s$ .

#### 5.4 What about repeating FRBs?

By writing the individual burst probabilities as being independent in equation (26), and assuming that the number of detected bursts follows a Poisson distribution in equation (27), we ignore the potential of FRBs to repeat. While the fraction of the FRB population which is observed to repeat is a current topic of debate, it is certain that many do. Formally, the FRB population described in Section 3 represents all *bursts*, rather than all FRB emitters, and the summation of equation (26) runs over all detected bursts. The distinction becomes irrelevant for distant, rarely repeating sources for which only ever zero or one bursts will be observed. For bright, nearby repeaters, the probability of having such an object in a survey's field of view is rare, especially when burstiness and/or periodicity is included (Oppermann, Yu & Pen 2018; Chime/Frb

Collaboration et al. 2020; Cruces et al. 2020; Rajwade et al. 2020), and an observation of zero bursts will be more likely than that estimated by equation (27). Conversely, the probability of observing many bursts will also be high, with observations of single bursts being much rarer than otherwise expected.

We note that the only FRB survey to ever observe an FRB repeat in an unbiased way are the observations by CHIME (CHIME/FRB Collaboration 2019b,c; Fonseca et al. 2020) – all other repeating FRBs have been discovered in targeted follow-up observations. This suggests that the majority of FRB observations can safely be classified as being in a ‘one burst per progenitor’ regime, regardless of the true fraction which are actually repeating objects. For these, our approach should be valid. We revisit this assumption in Section 8.4.

## 6 SURVEYS

Estimates of the FRB population have been made using data from many telescopes, which are often drawn from FRBCAT (Petroff et al. 2016). Due to the large number of FRBs they have detected and published, results from Parkes and ASKAP remain the most important, and we focus on these instruments here. Other important instruments we wish to examine in future works include the Upgraded Molonglo Observatory Synthesis Telescope (UTMOST), and the Canadian Hydrogen Intensity Mapping Experiment (CHIME).

The sensitivity of an FRB survey – and hence the functions  $p_N$ ,  $p_{dm}$ , and  $p_s$  from Section 5 – depends on the local contribution to DM, and hence varies with the value of  $DM_{ISM}$ . Since this fluctuates pointing-by-pointing, in theory these functions must be recalculated for every single pointing direction, which becomes computationally prohibitive. Evaluating  $p_{dm}$  and  $p_s$  however for the measured  $DM_{EG}$  and  $s$  of an FRB is much quicker. This motivates grouping FRB observations not just by telescope, but also by other observational properties, such as Galactic latitude. Here, we use five groups, as described below.

### 6.1 Parkes

All Parkes FRBs published so-far have used the multibeam (‘Mb’) receiver (Staveley-Smith et al. 1996). However, a new ultra-wideband receiver is now in place (Hobbs et al. 2020), which is being used for FRB searches and follow-up observations. We therefore refer to results from Parkes as ‘Parkes/Mb’ to distinguish this from future works.

For the Parkes multibeam, nominal sensitivity to a burst at beam centre is  $F_1 = 0.5 \text{ Jy ms}$  to a 1 ms duration burst (Keane et al. 2017) – this is an approximation, since different FRB searches used slightly different values of  $SNR_{th}$ . We neglect the effect of 1-bit sampling with early searches for FRBs with Parkes, which would have slightly degraded the sensitivity of these observations. These and other properties are summarized in Table 3.

Of the many Parkes FRB discoveries, we consider only those by the High Time Resolution Universe (HTRU; Keith et al. 2010; Thornton et al. 2013; Petroff et al. 2014; Champion et al. 2016) and Survey for Pulsars and Extragalactic Radio Bursts (SUPERB; Keane et al. 2017; Bhandari et al. 2018) collaborations to have an unbiased estimate of observation time,  $T_i$ . This is because their observation time and pointing directions were pre-determined, and the results published regardless of outcome. Other results suffer from publication bias whereby non-detections are less likely to be published. Thus, while their discovery can contribute to individual FRB likelihoods via  $p_{dmz}$  and  $p_s$ , no well-defined observation time  $T_i$  exists for use in equation (29), and the term  $p_i(N_i)$  must be neglected in equation (26). Thus they cannot contribute to estimates of the total FRB rate.

**Table 3.** Observational properties of follow-up observations, for ASKAP Fly’s Eye (FE) and incoherent sum (ICS) modes (Shannon et al. 2018; Bannister et al. 2019); the Parkes multibeam (MB) receiver (Keane et al. 2017); and the Greenbank Telescope’s (GBT’s) 820 MHz primary focus and L-band receivers (Kumar et al. 2019). From left to right: the telescope and receiver names, the central frequency  $\bar{\nu}$  and total bandwidth  $\Delta\nu$ , time- and frequency- resolutions  $\delta t$  and  $\delta\nu$ , and nominal sensitivity to a 1 ms duration burst.

Telescope	$N_{\text{FRB}}$	$T_{\text{obs}}$ (h) Mode	$\bar{\nu}$ (MHz)	$\Delta\nu$ (MHz)	$\delta t$ (ms)	$\delta\nu$ (MHz)	$F_1$ (Jy ms)
ASKAP/FE	20	26 616	1296	336	1.2565	1	21.9
ASKAP/ICS	7	N/A	1315	336	1.2565	1	4.4
Parkes/Mb	13	3946	1382	337.1	0.064	0.39	0.5

**Table 4.** Properties of fast radio bursts detected by the Parkes multibeam system, and used in this work. Given is the original FRB designation; measured total DM and  $\text{DM}_{\text{ISM}}$  estimated by the NE2001 (Cordes & Lazio 2002) in  $\text{pc cm}^{-3}$ , and ratio of measured to threshold SNR. Entries with a ‘\*’ indicates that this value is approximate; ‘†’ are discussed in text.

FRB	DM	Parkes: total $T_{\text{obs}} = 164.4$ d		Ref.
		$\text{DM}_{\text{ISM}}$	$s$	
110214	168.9	32	1.44	Petroff et al. (2019)
110220	944.4	36	5.44	Thornton et al. (2013)
110627	723	48	1.22	
110703	1103.6	33	1.78	
120127	553.3	33	1.22	
090625	899.55	32	2.8	Champion et al. (2016)
121002	1629.18	72	1.6	
130626	952.4	65	2	
130628	469.88	52	2.9	
130729	861	32	1.4	
151230	960.4	48	1.7	Bhandari et al. (2018)
160102	2596.1	36	1.6	
Parkes: unnormalized observation time				
FRB	DM	$\text{DM}_{\text{ISM}}$	$s$	Ref.
010305	350	44	1.02*	Zhang et al. (2020a)
010312	1187	51	1.1*	Zhang et al. (2019)
010724	375	45	200†	Lorimer et al. (2007)
131104	779	71	3.06	Ravi, Shannon & Jameson (2015)
140514	562.7	36	1.6	Petroff et al. (2015)
150807	266.5	38	5*	Ravi et al. (2016)
180309	263.52	46	241.6†	Osłowski et al. (2019)
180311	1570.9	46	1.15	

Since the distribution of  $\text{DM}_{\text{ISM}}$  affects telescope sensitivity, surveys at low Galactic latitudes have significantly reduced sensitivity compared to those at high latitudes, and the full distribution of time spent at each  $\text{DM}_{\text{ISM}}$  must be accounted for. In particular,  $\text{DM}_{\text{ISM}}$  will vary significantly for each pointing at low latitudes, making estimates numerically taxing. We therefore include only FRBs detected at mid ( $19.5^\circ < |b| < 42^\circ$ ) and high ( $42^\circ \leq |b| \leq 90^\circ$ ) Galactic latitudes. This criteria leaves 12 FRBs detected in a total of 164.4 d by HTRU and SUPERB (Bhandari et al. 2018), and another 8 FRBs by other groups with no reliable observation time. A full list is given in Table 4.

Early searches for FRBs with Parkes used a sparse grid of DMs and arrival times, resulting in sensitivities that would fluctuate by  $\pm 15$  per cent (Keane & Petroff 2015b). This was corrected with the use of HEIMDALL,<sup>3</sup> which has been used to (re)process the data from HTRU and SUPERB. While early HTRU searches extended only to  $\text{DM} = 2000 \text{ pc cm}^{-3}$  (Thornton et al. 2013), latter searches extended this to  $5000 \text{ pc cm}^{-3}$ ; and while the SUPERB ‘F’ pipeline looks for

FRBs with  $\text{DM} \leq 2000 \text{ pc cm}^{-3}$ , the SUPERB ‘T’ pipeline extends the search to  $10\,000 \text{ pc cm}^{-3}$ . Thus we treat all Parkes FRB searches as fully covering DM space.

Two FRBs deserve special mention. Both 010724 and 180309 saturated the primary beam, with their detection SNR being a lower limit on their true SNR in an idealized linear system. In the case of 010724, Lorimer et al. (2007) use the effect of this burst on the absolute normalization of the data in the analog stage to estimate a flux of 40 Jy over the time resolution of 5 ms, i.e. a total fluence of 200 Jy ms; we find their other estimate, of 20 Jy using the multibeam beamshape, to be less reliable for reasons discussed in Macquart & Ekers (2018a). Assuming a 1 Jy ms threshold, this yields  $s = 200$  for a 5 ms burst. For 180309, Osłowski et al. (2019) find an SNR of 2416 using data from a simultaneous pulsar folding system, yielding  $s = 241.6$ . Both these numbers have sizeable errors on them, however, to leave out the most powerful FRBs would skew the data. An alternative would be to allow lower limits on  $s$  to be included in the likelihood analysis, however this has not yet been implemented.

## 6.2 ASKAP

The Commensal Real-time ASKAP Fast Transients (CRAFT) group have performed several FRB surveys with ASKAP. The majority of ASKAP FRBs have been observed in single-antenna (‘Flye’s Eye’, or ‘FE’) mode during the ‘lat50’ survey, i.e. while observing Galactic latitudes of  $|b| = 50^\circ \pm 5^\circ$  (Bannister et al. 2017; Shannon et al. 2018). Twenty FRBs have been initially reported (Bhandari et al. 2019; Macquart et al. 2019), with a total recorded data time of 1274.6 antenna-days duration (James et al. 2019a). A further six FRBs have been detected in a variety of surveys (Bhandari et al. 2019; Macquart et al. 2019; Qiu et al. 2019), of which four satisfy the Galactic latitude requirement. These are listed in Table 5. We err on the side of caution and do not assume a known observation time for this last category, since several other FRB searches outside the lat50 survey have been performed and were not reported. As with Parkes,  $p_i(N_i)$  in equation (26) is only evaluated for the former category.

All ASKAP/FE searches have used the same time/spectral resolutions and almost identical frequency ranges, as given in Table 3. The beamshape and threshold for this survey are given in James et al. (2019a).

ASKAP has recently been observing in incoherent sum mode (ICS), with voltage buffers used in offline analysis to localise FRBs to sub-arcsec precision (Bannister et al. 2019). Follow-up observations with radio and optical instruments have determined the redshifts of the host galaxies of each FRB, allowing the DM– $z$  grid to be directly probed for the first time. This mode has undergone an extended period of commissioning, with the number of telescopes, observation frequency, and time resolution of the search all varying. The total observation time is difficult to estimate, again precluding the use of observation time in this survey’s likelihood calculation. A compre-

<sup>3</sup><http://sourceforge.net/projects/heimdall-astro/>

**Table 5.** As per Table 4, for ASKAP/FE observations.

FRB	DM	ASKAP/FE: $T_{\text{obs}} = 1274.6$ d			Ref.
		DM <sub>ISM</sub>	s		
170107	609.5	37	1.68		Bannister et al. (2017)
170416	523.2	40	1.38		Shannon et al. (2018)
170428	991.7	40	1.11		
170707	235.2	38.5	1.00		
170712	312.8	35.9	1.34		
170906	390.3	38.9	1.79		
171003	463.2	40.5	1.45		
171004	304.0	38.5	1.15		
171019	460.8	37	2.46		
171020	114.1	38.4	2.05		
171116	618.5	35.8	1.24		
171213	158.6	36.8	2.64		
171216	203.1	37.2	1*		
180110	715.7	38.8	3.75		
180119	402.7	35.6	1.67		
180128.0	441.4	32.3	1.31		
180128.2	495.9	40.5	1.01		
180130	343.5	38.7	1.08		
180131	657.7	39.5	1.45		
180212	167.5	30.5	1.93		
FRB	DM	ASKAP/FE: unnormalized time			Ref.
		DM <sub>ISM</sub>	s		
180417	474.8	26.1	1.84		Agarwal et al. (2019)
180515	355.2	32.6	1.27		Bhandari et al. (2019)
180324	431	64	1.03		Macquart et al. (2019)
180525	388.1	30.8	2.88		

**Table 6.** As per Table 4, for ASKAP/ICS observations, with redshift  $z$  included.

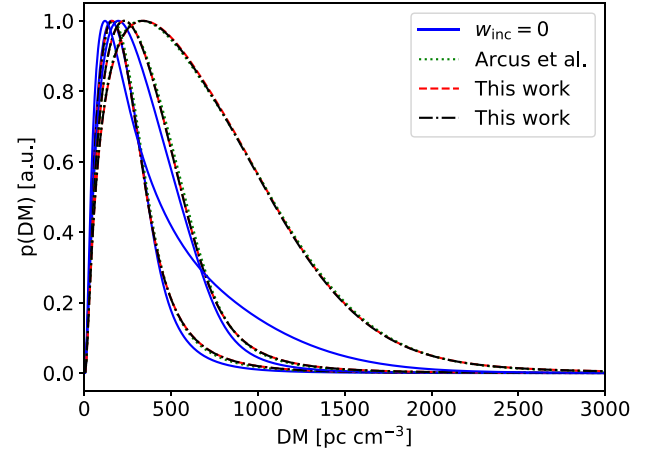
FRB	ASKAP/ICS: unnormalized time					Ref.
	DM	DM <sub>ISM</sub>	s	z		
180924	362.4	40.5	2.34	0.3214		Bannister et al. (2019)
181112	589.0	40.2	2.14	0.4755		Prochaska et al. (2019b)
190102	364.5	57.3	1.38	0.291		Macquart et al. (2020)
190608	339.5	37.2	1.79	0.1178		
190611.2	322.2	57.6	1.03	0.378		
190711	594.6	56.6	2.64	0.522		
190714	504.7	38.5	1.19	0.209		Heintz et al. (2020)

hensive analysis would involve recalculating the  $z$ -DM grid for each and every observed burst. For reasons of computational efficiency, we instead use mean observation parameters to evaluate the likelihood on this grid. This precludes the use of FRB 191001, which was detected at a lower frequency during commensal observations (Bhandari et al. 2020b). The remaining seven FRBs used are given in Table 6.

## 7 INITIAL RESULTS

### 7.1 Calculations

We use a brute-force approach to find the best-fitting values of  $E_{\text{max}}$ ,  $\alpha$ ,  $\gamma$ ,  $n$ ,  $\mu_{\text{host}}$ ,  $\sigma_{\text{host}}$ , and evaluate equation (25) over a multidimensional cube of parameter values. The resulting likelihood dependence on each is calculated for single (pairs of) parameters by marginalizing over the remaining five (four) parameters. That is, the value taken is the maximum likelihood found when the remaining five (four) values are varied over their full range.



**Figure 6.** Effect of using different distributions of the intrinsic width  $w_{\text{int}}$  on the expected DM distribution of FRBs,  $p(\text{DM})$ , from the FRB surveys considered here: ASKAP/FE (left-most), ASKAP/ICS (centre), and Parkes/Mb (right-most). Assuming no intrinsic width (blue solid), using a lognormal distribution of widths with parameters from Arcus et al. (2021) (green dotted), using a lognormal distribution with parameters derived in this work (red dashed), and numerically approximating the latter (black dot-dashed). The distributions are normalized to a peak of unity for illustrative purposes.

In this work, we use a frequentist approach to setting confidence limits. Confidence intervals are determined using Wilks' theorem

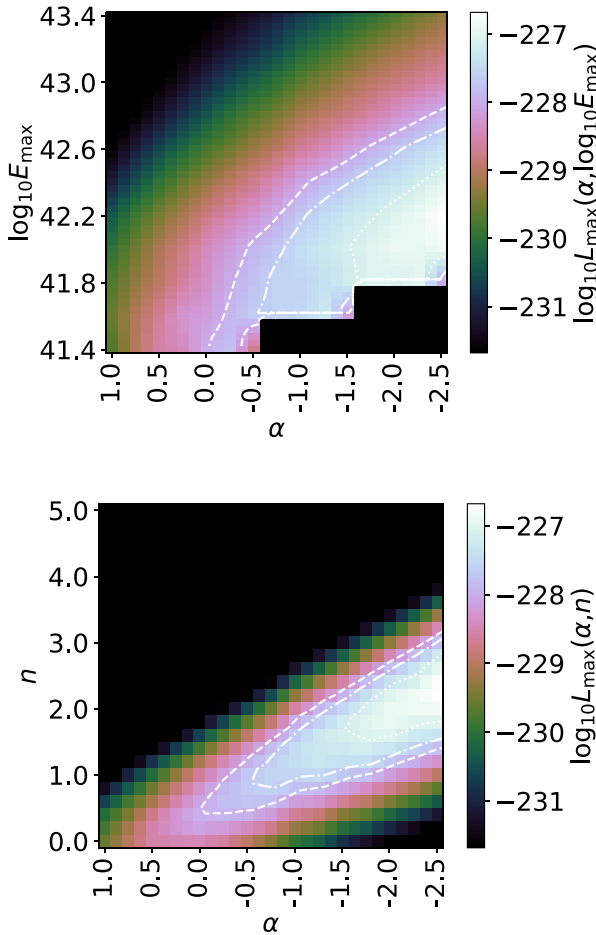
$$2(\log L_{\text{max}} - \log L) \sim \chi^2_{\text{ndf}}, \quad (41)$$

where  $\chi^2_{\text{ndf}}$  is a chi-square distribution with  $\text{ndf}$  degrees of freedom, here equal to the number of parameters which have not been marginalized over (either one or two throughout this work; Wilks 1962; James 2006). The 51 FRBs used in this work should satisfy the large- $N$  limit required for equation (37) to be valid. We do however investigate the validity of this test in an extreme case in Appendix A6, and generate Bayesian confidence intervals in Appendix A7, which turn out to be narrower than those calculated using a frequentist approach.

### 7.2 Parameter degeneracy

Fits to data taken at a single frequency suffer from a degeneracy in the fitting parameters between  $E_{\text{max}}$ ,  $\alpha$ , and  $n$ . This is shown in Fig. 7, which plots the variation of the marginalized likelihood over these parameters. The data shows a small preference for large, negative values of  $\alpha$  – preliminary investigations covered  $-8 \leq \alpha \leq 0$  – with no distinct maximum being found. For these final computations, we restrict  $\alpha$  to the range  $-2.5 \leq \alpha \leq 1$  for reasons which will be discussed below.

The reason for this degeneracy is that these three parameters are strongly related to the high-DM, high- $z$  cut-off in the observed FRB distribution. This is restricted by the lack of ASKAP/ICS-localized FRBs at high redshifts, and by the number of high-DM FRBs detected by ASKAP/FE and Parkes multibeam observations. Having a steep spectral index (very negative value of  $\alpha$ ) provides a mechanism to reduce the expected number of FRBs via the  $k$ -correction, and is both consistent with, and requires, a higher value of  $E_{\text{max}}$ . Similarly, it also allows for a rapidly evolving population with redshift (high  $n$ ), due to suppression by the  $k$ -correction. A very large number of high- $z$  events, or large value of  $E_{\text{max}}$ , is excluded when  $\alpha$  is near zero. This



**Figure 7.** Degeneracy of the marginalized likelihood with  $\alpha$ . Top: dependence of  $E_{\max}$ , bottom: dependence on  $n$ . The contours give the 68, 90, and 95 per cent confidence limits. The step-like behaviour in  $E_{\max}$  is due to coarse gridding.

degeneracy was also noted by Lu & Piro (2019), who by default use  $\alpha = -1.5$ .

Without further data at different frequencies, or a prior on any of these three parameters, this degeneracy cannot be broken. All results from the literature on  $E_{\max}$  and  $n$  derive from similar analyses as in this work, albeit with simpler methods, and are therefore not independent. We therefore investigate other evidence for plausible values of  $\alpha$ , in order to break this near degeneracy.

There are two broad methods for measuring  $\alpha$ : searching for spectral dependence in the FRBs detected by a single experiment, and comparing FRB rates between experiments with different frequencies. We discuss results from both below.

### 7.3 Spectral dependence from ASKAP FRBs

In an analysis of 23 ASKAP FRBs detected in the 1147–1483 MHz band, Macquart et al. (2019) found the summed power was consistent with  $\alpha = -1.5^{+0.2}_{-0.3}$ .<sup>4</sup> This measurement was possible because of the closely packed ASKAP beams, which allowed the FRB location

<sup>4</sup>The error is in fact almost symmetric, with single-digit rounding resulting in the apparently large asymmetry.

within the beam to be determined, and hence spectral corrections to be applied due to the frequency-dependent beamshape.

It has been pointed out by the referee of this work that while Macquart et al. (2019) applied these spectral corrections to observed FRBs, they did not account for such effects to model the FRBs they did not observe. In the spectral index interpretation of  $\alpha$ , there is no such effect – all FRBs are broad-band, and every measured FRB tends to have the same spectral index. Thus, it is impossible to detect a biased fraction, and we use a Gaussian prior of  $\alpha = -1.5 \pm 0.3$  for the spectral interpretation of  $\alpha$ .

In the rate interpretation, however, where every FRB is narrow band, each burst is observed according to the sensitivity at its particular frequency. For experiments with widely spaced beams, the observed sky area usually scales as  $\nu^{-2}$ , so that far-more low-frequency bursts would be observed even for a frequency-independent rate. A similar effect, but less extreme, would be observed for an intermediate case where FRBs have a distribution of spectral indices that average to  $\alpha$ , and redder FRBs would preferentially be detected.

The measurements of Macquart et al. (2019) were made with ASKAP, with closely spaced beams overlapping at the half-power points. To estimate the maximum possible bias from this effect, we perform the following toy calculation. We use a Euclidean source-counts such that rate is proportional to beam power  $B^{1.5}$ , and assume that ASKAP beams are Gaussian and overlap at half-max near band centre, with width scaling inversely with the frequency range of Table 3. We calculate the frequency-dependent observed rate assuming all FRBs have a very narrow bandwidth (i.e. under the rate interpretation), with a true  $\alpha = 0$ . We find this effect would result in 25 per cent more FRBs at the low-frequency end than the high-frequency end of the ASKAP band, so that one would measure  $\alpha = -0.85$ , despite the true value being  $\alpha = 0$  – i.e. the measured value needs to have 0.85 added to it to obtain the true value. Applying this correction to the results of Macquart et al. (2019) in the rate interpretation, we use a Gaussian prior of  $\alpha = -0.65 \pm 0.3$ .

### 7.4 Spectral dependence from cross-experimental rates

FRB spectral dependence can also be derived by comparing detections and non-detections between different instruments, with several results published in the literature.

Sokolowski et al. (2018) performed shadowing observations of ASKAP with the Murchison Widefield Array (MWA) at 170–200 MHz, with their non-detections being inconsistent with the high fluences predicted by  $\alpha = -1.5^{+0.2}_{-0.3}$ . Those authors suggest possible explanations as a higher (less negative) value of  $\alpha$ ; missing the emission due to the small 30 MHz bandwidth and low FRB band occupancy; synchrotron self-absorption in the emission region; and scatter broadening at low frequencies.

Using the then-non-detection of FRBs at 350 MHz by the Robert C. Byrd Green Bank Telescope, Chawla et al. (2017) estimate  $\alpha > -0.3$ . However, the Greenbank result is somewhat dependent on the assumed Parkes rate and FRB scattering distribution, with a constraint of  $\alpha > -0.9$  reached under different assumptions – compatible with the adjusted Macquart et al. (2019) value in the rate interpretation. Farah et al. (2019) find a relatively low FRB rate using the Upgraded Molonglo Observatory Synthesis Telescope (UTMOST) at 843 MHz. A value of  $\alpha > 0$  would be required for compatibility with the Parkes and ASKAP rates. Gupta et al. (2021) have since shown however that the UTMOST pipeline would tend to underestimate the SNR of FRBs, which is a likely explanation for the low rate. The CHIME/FRB Collaboration (2021) present and



**Table 7.** Summary of results and methods used by other authors in estimating FRB population parameters, compared to those of this work using the spectral interpretation of  $\alpha$  with a Gaussian prior. Shown in columns 2–7 are the best-fitting parameter values and corresponding uncertainties, converted to the units and notation used in this work. For the sake of comparability, all limits have been converted to  $1\sigma$  standard deviations assuming that uncertainties are Gaussian distributed. Columns 8–11 summarize the different effects that were (Y) or were not (N) included in the modelling: accounting for instrumental beamshape, a distribution in dispersion measure for a given redshift, a DM-dependent telescope threshold, and a distribution of burst widths  $f(w)$  and its corresponding effects on sensitivity. Note that a ‘Y’ does not necessarily mean a comparable treatment to this work.

Author	$E_{\max}$	$\alpha$	$\gamma$	$n$	$\mu_{\text{host}}$	$\sigma_{\text{host}}$	Beam	DM  $z$	$\eta(\text{DM})$	$f(w)$
Lu & Piro (2019)	$43.1^{+1.1a}_{-0.7}$	$-1.5^b$	$-0.6 \pm 0.3$	$0.3^{+1.0c}_{-1.1}$	N	N	N	N	N	N
Luo et al. (2020)	$41.16^{+0.51a}_{-0.19}$	$0^b$	$-0.79^{+0.16}_{-0.18}$	$0^b$	$30^{b,d,e}$	$0.17^{b,d,e}$	Y	Y	Y	Y
Arcus et al. (2021)	$45.1^b$	$-2.1^{1.1}_{-1.4}$	$-0.7^{+0.2}_{-0.2}$	$1 \pm 1$	N	N	N	N	Y	Y
Caleb et al. (2016)	$41.2^f$	$0^b$	N/A	$0.1^b$	$100^b$	$0^b$	Y	Y	Y	Y
Macquart et al. (2020)	N/A	N/A	N/A	N/A	$65^{+95e}_{-25}$	$0.9^{+0.9e}_{-0.6}$	N	Y	N	N
This work	$41.70^{+0.53}_{-0.06}$	$-1.55^{+0.18g}_{-0.18}$	$-1.09^{+0.14}_{-0.10}$	$1.67^{+0.25}_{-0.40}$	$130^{+66}_{-48}$	$0.53^{+0.15}_{-0.11}$	Y	Y	Y	Y

<sup>a</sup>Indicates the use of a Schechter function, with exponential decay after  $E_{\max}$ .

<sup>b</sup>Parameter held fixed by the authors.

<sup>c</sup>The  $(1+z)^\beta$  scaling used in this work is converted to  $n$  via  $n = 2.7\beta$  from equation (14).

<sup>d</sup>The host galaxy contribution is modelled according to the ‘ALGs(YMW16)’ model from Luo et al. (2018), with approximate mean  $\sim 0.8$  and standard deviation 0.2 in  $\log_{10} \text{DM}$ , to which is added a ‘local’ contribution with a uniform distribution from 0–50, for a total mean of 30 and quadrature-added deviation 15, i.e. 0.17 in  $\log_{10}$ .

<sup>e</sup>These values are not explicitly quoted, and were approximately derived from plots shown in the text. <sup>f</sup>This constraint primarily arises from the prior of Macquart et al. (2019).

<sup>g</sup>A lognormal luminosity function was used; here we quote the mean of  $10^{38.2}$  erg plus three standard deviations of 1 in  $\log_{10}$ .

<sup>h</sup>This constraint primarily arises from the prior of Macquart et al. (2019).

analyse the spectral shapes of more than 500 bursts detected by the Canadian Hydrogen Intensity Mapping Experiment (CHIME) FRB system. Their implied all-sky rate is compatible with that of Parkes and ASKAP when assuming  $\alpha = 0$ , although the range of compatible values is not greatly studied, and effects due to different detection efficiencies are not considered.

These results all point to a less negative value of  $\alpha$  than found by Macquart et al. (2019). We therefore also consider a broad uniform (‘boxcar’) prior of  $-2.5 < \alpha < 1.0$ . However, as will be discussed below, we do not consider this prior as equally valid as the Gaussian prior, and thus by default quote the former.

## 7.5 Validity of the priors

We caution against taking the uniform prior as equally valid as the Gaussian prior from the measurements of Macquart et al. (2019). First, of the Parkes, ASKAP, UTMOST, Greenbank, and CHIME rates used by the studies in Section 7.4, only those by CHIME have accounted for telescope beamshape, although beamshape effects are relatively small in ASKAP due to the closely packed beams. Secondly, different telescopes have different response functions to FRB width and DM, and scattering is greater at lower frequencies than at high frequencies. This effect will, in general, reduce the measured rate for low-frequency systems; for a given measured rate, the true rate will be higher, and hence the true value of  $\alpha$  will be lower (or more negative) than implied by count rates alone. Thirdly, we emphasize that ASKAP FRBs are absolutely not in the narrow-bandwidth limit used to estimate the bias in the result of Macquart et al. (2019). Pleunis et al. (2021) find 30 per cent of FRBs to have broad-band structure, 60 per cent to be band-limited, and 10 per cent to have complicated morphology, indicating that the truth lies in-between the two extreme interpretations of  $\alpha$  – and while the bursts detected by ASKAP do show narrow-band features, these tend to be spread across the spectrum (Shannon et al. 2018). Finally, it is also possible that the FRB spectrum is flat below 1 GHz, and steep above it – and all the results discussed in Section 7.4 are obtained below 1 GHz, and may therefore be considered irrelevant.

Another interesting effect is that all experiments modelled in this work have broader beams at low frequency than at high. Thus, and remaining bias in  $\alpha$  due to low-frequency FRBs being more likely to be detected would, to first order, be cancelled when incorporating that same increased chance of detection into FRB beamshape modelling. Indeed, the effect would go in the opposite direction: since Parkes has a wider beam spacing, the positive detection bias to low-frequency FRBs would be even greater than that in the ASKAP measurement of  $\alpha$ , so that while the true value of  $\alpha$  might be less negative than the one found by Macquart et al. (2019), the applicable value of  $\alpha$  to Parkes measurements may be more negative. Such a difference is clearly a second-order effect however.

## 7.6 Comparison of results

The limits on single FRB population parameters presented in James et al. (2021a) are obtained with this approach. We observe that while our prior on  $\alpha$  shifts the preferred values of the other parameters by small amounts, it is not a large influence, and primarily acts to limit very strong source evolution models with  $n > 3$ , and no to negative evolution with  $n \leq 0$ . Table 7 compares these results with a prior on  $\alpha$  to those of other authors, as well as a brief summary of which effects are included.

## 8 RESULTS 2: REDSHIFT AND DISPERSION MEASURE

The best-fitting parameter set (James et al. 2021a) allows a comparison between the expected and observed distributions of FRBs in DM,  $z$ , and SNR space. However, each combination of allowed parameters produces a unique map of the  $z$ -DM distribution of FRBs for each telescope. Rather than present a very large number of plots, we use the following approach to identify a finite number of reasonable possibilities.

For each parameter, we take its value at both the upper and lower limits of its 1D 90 per cent C.L., and choose the corresponding values of the other parameters. This results in 12 further parameter

sets, which are listed in Table 8. For comparison, we also include the best-fitting parameter set assuming no source evolution, and the upper limit on  $E_{\min}$  as described in Section 8.3.

For each of these parameter sets – including the best-fitting set – we generate the expected FRB distribution in DM– $z$  space, and compare this to the observed values. All the results shown in this section are calculated using the spectral index interpretation of  $\alpha$  with the Gaussian prior.

In interpreting the results in this section, we caution that the data to which we compare expectations have been used to determine the FRB population parameters, so the two are not independent. However, given that only eight free parameters (the standard six, plus two for the FRB width distribution) are used to fit multiple observables from three different FRB surveys, good agreement should not be the result of overfitting, but rather indicate a genuine correspondence between the models used in this work and reality.

### 8.1 Observed and predicted distributions

The predicted DM and  $z$  distributions from all 15 cases – best-fitting, no source evolution,  $E_{\min}$ , and 12 sets reflecting parameter uncertainty – are shown in Figs 8 and 9, respectively. In the case of the DM distribution, data and predictions are shown individually from each survey described in Section 6, and stacked together to allow a better comparison. For the  $z$  distribution, the only available data comes from ASKAP/ICS observations – however, predictions from each individual survey are also shown.

The first question to ask is – are the best fits indeed good fits? Our best-fitting parameter estimates (James et al. 2021a) do not necessarily indicate that the modelled DM,  $z$ , and SNR distributions are good fits to the data – merely that they are the best fits given the form of the model used. A by-eye analysis shows that the predictions from the best-fitting parameter set are indeed a good match to the data. However, there appears to be an overprediction of FRBs at low DM and particularly at low redshift. This is a common feature over all the parameter sets within the 90 per cent error margins, although the degree of peakedness near  $z = 0$  varies greatly.

We consider four possible explanations for this below.

### 8.2 Random fluctuations

The observed deficit of low-DM and low- $z$  FRBs may simply be a statistical fluctuation. To evaluate this, we perform Kolmogorov–Smirnov tests on the ASKAP ICS  $z$  distribution and total DM distribution (Kolmogorov 1933; Smirnov 1948). Predicted and measured cumulative distributions of both DM and  $z$  are shown in Fig. 10 – the KS-statistic is the maximum absolute difference between the two curves. To evaluate the significance of this statistic, we take the best-fitting curves as the truth, and randomly generate 10 000 samples from each. Histogramming the results produces the expected distributions of the KS statistic under the null hypothesis that the best-fitting prediction is true. Comparing this distribution to the observed value of the KS statistic in Fig. 11 shows that our observed values of DM and  $z$  are consistent with predictions, with larger values of the KS statistics observed in 7 and 42 per cent of cases for the DM and  $z$  distributions, respectively. Performing a similar analysis using the  $n = 0$  distribution gives  $p$ -values of 8 and 6 per cent for the DM and  $z$  distributions, respectively – evidently, the DM distributions of the ASKAP/FE and Parkes samples show a much better fit for  $n > 0$  than  $n = 0$ . We therefore conclude that the apparent deficit of FRBs at low DM and redshift compared with predictions is consistent within statistical fluctuations of expectations. None the less, we proceed

with further analysis since the presence or otherwise of a minimum energy, or effects due to a large fraction of the population being repeaters, is of great interest.

### 8.3 Evidence for a minimum energy $E_{\min}$

In our standard modelling, we have set the minimum FRB energy  $E_{\min}$  to an extremely low value of  $10^{30}$  erg – well below the characteristic energies of observed FRBs, effectively making it zero. This is because values of  $\gamma > -1.5$  render FRB observations primarily sensitive to  $E_{\max}$  (Macquart & Ekers 2018b), while bursts from FRB 121102 have been observed at much lower energies than are likely to be probed by Parkes and ASKAP observations (e.g. Law et al. 2017).

However, a clear possible explanation for the apparent deficit of FRBs at low DM- $z$  is a minimum FRB energy. Such has recently been observed for the repeating FRB 121102 by FAST (Li et al. 2021), with the burst rate suppressed below  $4.8 \times 10^{37}$  erg. Without such a cut-off, as telescopes probe ever lower values of  $E$  in the nearby Universe, the predicted number of FRB observations per comoving volume will increase without limit when  $\gamma \leq -1$ . If  $-1 \geq \gamma > -1.5$ , the reducing volume will somewhat compensate, and the total rate will remain finite. This gives rise to the sharp increase in the best fit expected redshift distributions near  $z = 0$  in Fig. 9, although such a peak may not be present within 90 per cent error margins.

To investigate the effect of  $E_{\min}$ , we fix the best-fitting parameters, and vary  $E_{\min}$ . The evolution of the likelihoods is shown in Fig. 12. Interestingly, the total likelihood decreases with increasing  $E_{\min}$ , with  $E_{\min} = 10^{39.0}$  erg the 90 per cent C.L. upper limit. Why? As expected from Figs 8 and 9, the  $p(z, \text{DM})$  contribution increases with  $E_{\min}$ , peaking near  $10^{39}$  erg. This peak is a combination of the ASKAP/ICS observations strongly favouring a large  $E_{\min} \approx 10^{40.5}$  erg, and the Parkes multibeam observations strongly favouring  $E_{\min} < 10^{39}$  erg, with the ASKAP/FE observations being relatively neutral until  $E_{\min} > 10^{40}$  erg.

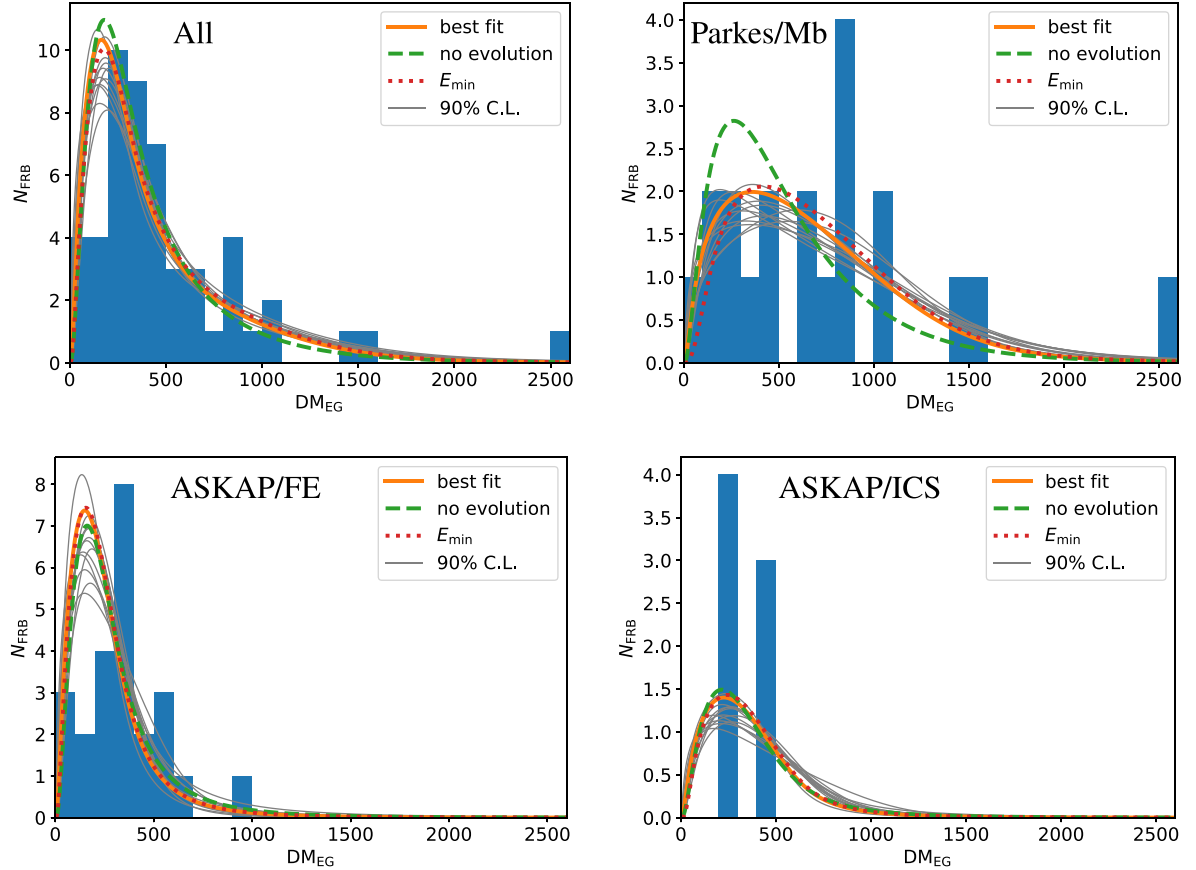
The reason why the Parkes observations are strongly against high values of  $E_{\min}$  is that sensitive telescopes with small fields of view are highly unlikely to observe low-DM bursts, unless the small volume of the near Universe corresponding to such a low DM is populated by many (necessarily low energy) FRBs. The lowest DM burst detected by Parkes during the included surveys, FRB 110214, had a DM of  $168.5 \text{ pc cm}^{-3}$  (Petroff et al. 2019). Setting  $E_{\min} = 10^{39.0}$  implies reduced FRB rates for redshifts closer than  $z = 0.332$ , assuming a limiting fluence of  $0.3 \text{ Jy ms}$ , with those bursts that are detected likely to have an SNR significantly greater than threshold. A redshift of  $z = 0.332$  implies a most likely dispersion measure of approximately  $322 \text{ pc cm}^{-3}$ , being composed of a cosmological contribution of  $\text{DM}_{\text{cosmic}} = 142 \text{ pc cm}^{-3}$ , local contribution of  $82 \text{ pc cm}^{-3}$ , and our model best-fitting value of  $\mu_{\text{host}}/(1+z) = 97.6 \text{ pc cm}^{-3}$ . The observation of FRBs by Parkes with DMs below this value therefore disfavour a significant minimum energy  $E_{\min}$ .

This is illustrated in Fig. 13, where we show model predictions of  $p_s$ , and the observed values of  $s$ , for each Parkes-detected FRB. This is done for the best-fitting model, and when using  $E_{\min} = 10^{39.0}$  erg. While the  $E_{\min}$  model predicts that the high fluence of FRB 180309 with a DM of  $\sim 263 \text{ pc cm}^{-3}$  and  $s = 41$  is slightly more likely, it predicts that FRB 110214, with a DM of  $\sim 169 \text{ pc cm}^{-3}$ , is much less likely to have its observed value of  $s = 1.44$ .

In this work, we have not included  $E_{\min}$  as a global minimization parameter due to computational constraints. Might there be some other combination of parameters for which a significant  $E_{\min}$  is found? To investigate this, we have repeated the  $E_{\min}$  optimization for

**Table 8.** Parameter sets used in this work: the best-fitting values found in our companion paper (James et al. 2021a); sets allowing the parameter in the left-most column to take its minimum and maximum value within 90 per cent C.L.; the best-fitting parameter set when forcing  $n = 0$ , i.e. no source evolution; and a set when setting  $E_{\min} = 10^{38.5}$  erg, but otherwise identical to the best fit.

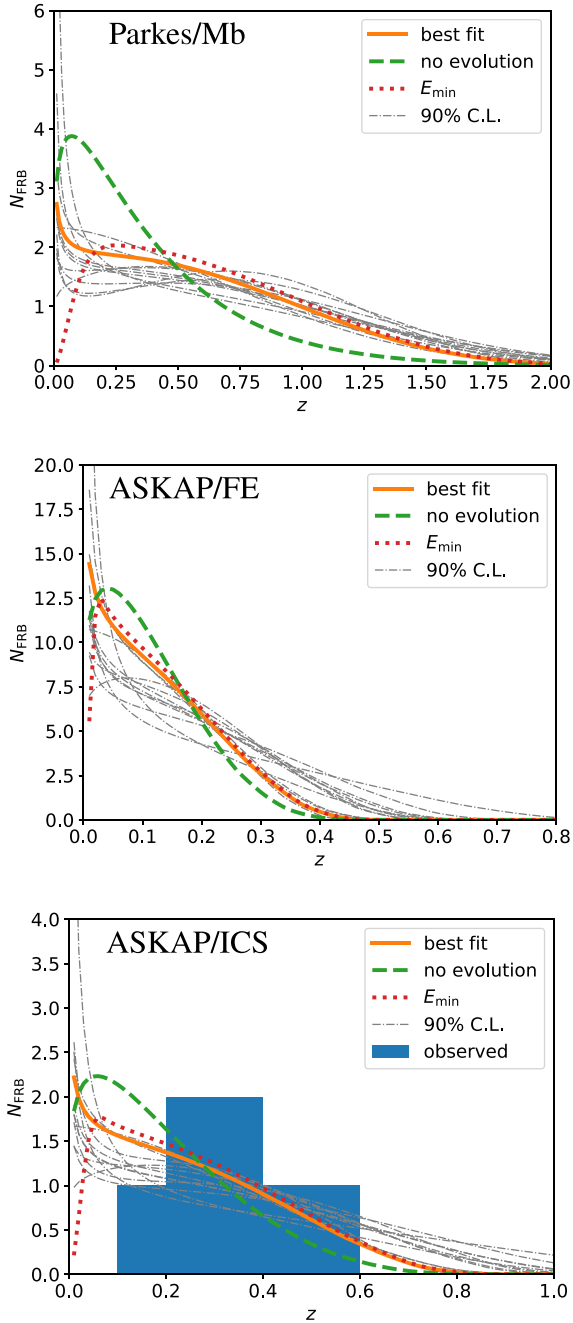
Set	$\log_{10} E_{\min}$	$\log_{10} E_{\max}$	$\alpha$	$\gamma$	$n$	$\mu_x$	$\sigma_x$
Best fit	30	41.7	−1.55	−1.09	1.67	2.11	0.53
$\log_{10} E_{\max}$	30.0	41.6	−1.5	−1.2	2.6	2.0	0.5
	30.0	42.51	−1.5	−1.2	1.69	2.0	0.5
$\alpha$	30.0	42.0	−1.88	−1.05	1.8	2.0	0.5
	30.0	42.0	−1.2	−1.1	1.48	2.0	0.5
$\gamma$	30.0	42.08	−1.5	−1.34	2.02	2.25	0.5
	30.0	41.92	−1.5	−0.96	1.4	2.0	0.54
$n$	30.0	41.8	−1.5	−1.0	1.11	2.25	0.5
	30.0	42.0	−1.75	−1.14	2.28	2.0	0.5
$\mu_{\text{host}}$	30.0	42.18	−1.5	−1.1	1.78	1.77	0.59
	30.0	41.8	−1.5	−1.1	1.47	2.41	0.63
$\sigma_{\text{host}}$	30.0	42.08	−1.5	−1.1	1.6	2.0	0.36
	30.0	42.0	−1.5	−1.1	1.6	2.0	0.81
$n = 0$	30.0	41.6	−1.25	−0.9	0.	2.25	0.5
$E_{\min}$	39	41.7	−1.55	−1.09	1.67	2.11	0.53



**Figure 8.** Observed (histogram) and predicted (lines) distributions of DM for all surveys considered (upper left), and the individual Parkes multibeam (upper right), ASKAP Fly’s Eye (lower left), and ASKAP ICS (lower right) surveys. Predictions show the best fit over the entire parameter space, when constrained to no source evolution, when allowing  $E_{\min}$  to vary, and when varying parameters within their 90 per cent C.L.

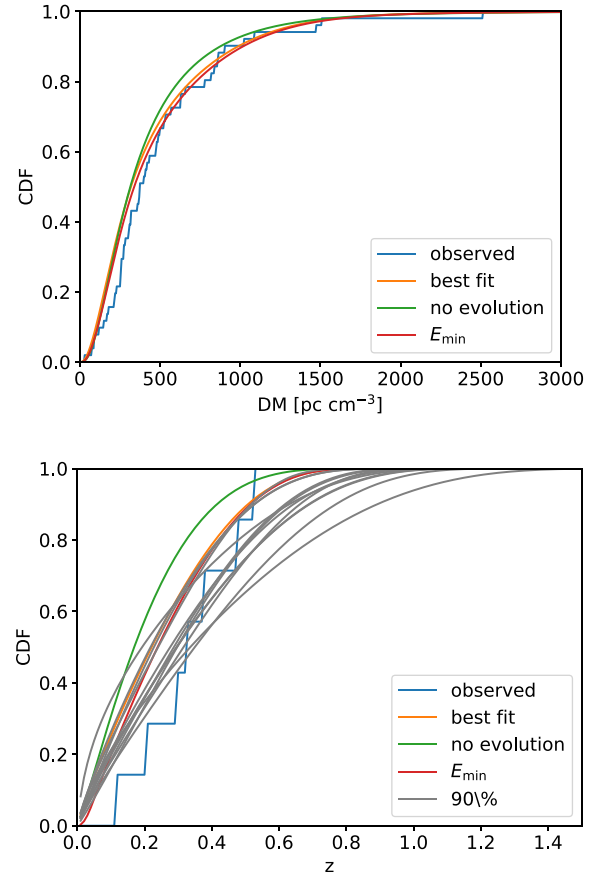
all parameter sets in Table 8. Only for the parameter set minimizing  $\gamma$  do we find a significant value of  $E_{\min}$  to be preferred, with a best-fitting value of  $10^{38.2}$  erg, and 90 per cent upper limit of  $10^{39.1}$  erg. This makes sense, since otherwise a steep energy function would

overpredict the number of near-Universe FRBs. The resulting gain in likelihood acts to weaken our confidence in the lower limits on  $\gamma$ , e.g. the 90 per cent lower limit shifts to 60 per cent confidence.



**Figure 9.** Predicted distributions (lines) of  $z$  for Parkes multibeam (top) and ASKAP Fly’s Eye (middle) surveys, with the ASKAP ICS results (bottom) also showing the observed values (histogram). Predictions show the best fit over the entire parameter space, when constrained to no source evolution, when setting  $E_{\min} = 10^{39}$  erg, and when varying parameters within their 90 per cent C.L.

Our findings clearly do not indicate that there is no minimum FRB energy. Within the confines of our power-law model, the data used appear insensitive to  $E_{\min} < 10^{37}$  erg, and we can only rule out  $E_{\min} > 10^{39.0}$  erg at 90 per cent C.L. – but it does run counter to the findings of Luo et al. (2020). These authors find a most likely minimum luminosity of  $10^{42}$  erg s $^{-1}$ , which is approximately  $10^{39}$  erg assuming a standard 1-ms burst, although they conclude rather that this finding is due to the limit of their sample.



**Figure 10.** Observed and predicted cumulative distribution functions for the DM from all considered surveys (above) and  $z$  from the ASKAP ICS survey (below). Predictions show the best fit over the entire parameter space, when constrained to no source evolution, and when allowing  $E_{\min}$  to vary.

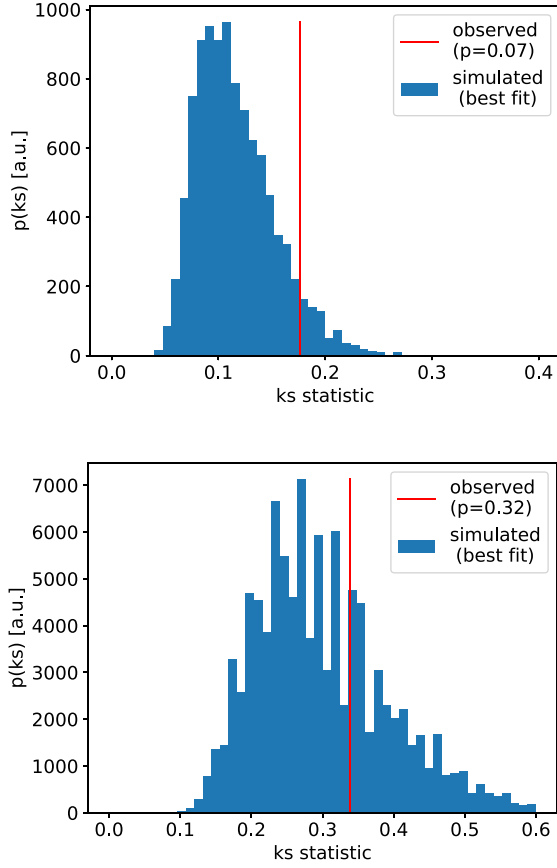
#### 8.4 Influence of repetition

In this study, we have ignored repeating FRBs on the basis that none of the FRBs detected by ASKAP and Parkes have been observed more than once – even though some are known to repeat (Kumar et al. 2019, 2020; Patek & Chime/Frb Collaboration 2019; The CHIME/FRB Collaboration 2021), the surveys were not sensitive enough to probe this. However, the method of Section 5 covers the entirety of  $z$ –DM space, regardless of whether or not an FRB has been detected at that point. Therefore, for any hypothesized repeating FRB, there will always be a sufficiently nearby volume of the Universe where any survey would be expected to detect more than one burst if the FRB landed in its field of view. If such a repeating FRB happens to be located in that volume, the observed burst rate will be greater than expected – but if one does not, it will be less.

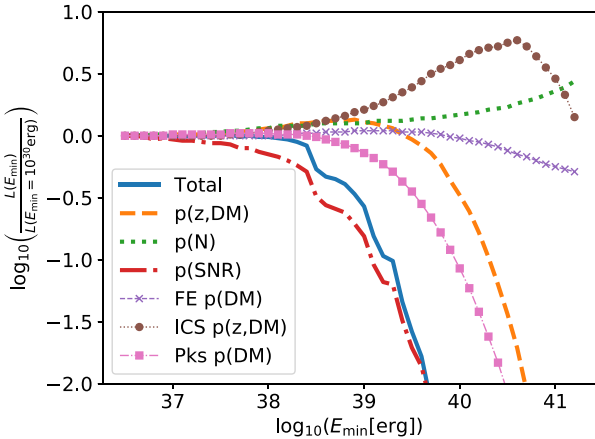
This effect is analysed in the context of Canadian Hydrogen Intensity Mapping Experiment (CHIME) observations by Gardenier et al. (2020), who note that the DM distribution of repeating FRBs should be lower than that of repeaters observed only once. While this has not yet been observed (CHIME/FRB Collaboration 2019a,b,c; Fonseca et al. 2020), this may be due to a very broad distribution of intrinsic repetition rates (James et al. 2020b).

As discussed by James (2019), the ASKAP/FE observations have deeply probed some regions of sky with over 50 antenna-days spent on individual fields. This has allowed strong limits to be placed on FRB repetitions – but also made the observations more susceptible

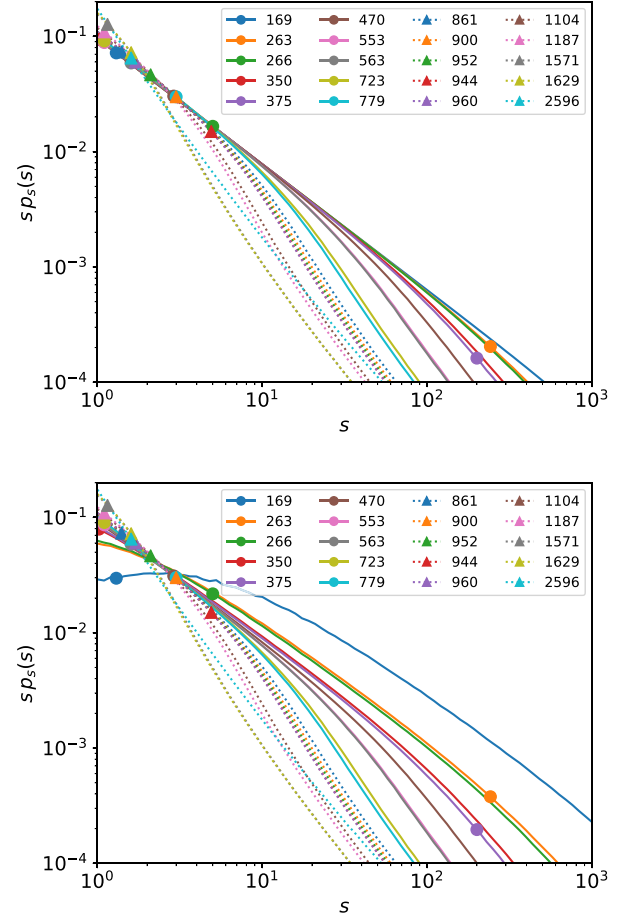




**Figure 11.** Observed value (red line) and simulated distribution (blue histogram) of the KS statistics for the DM (top) and redshift (bottom) distributions under the best-fitting hypothesis. Shown are the corresponding p-values to reject this hypothesis.



**Figure 12.** Evolution of likelihoods as a function of minimum energy  $E_{\min}$ , relative to the value at  $E_{\min} = 10^{30}$  erg. Shown is the total likelihood; the likelihood over all surveys, split into contributions from the  $z$ -DM distribution  $p(z, DM)$ , the number of events  $p(N)$ , and the signal-to-noise probability  $p(SNR)$ . The  $p(z, DM)$  contribution is further split into components from the ASKAP/FE, ASKAP/ICS, and Parkes/Multibeam surveys.



**Figure 13.** Likelihoods of observing FRBs with a given  $s$ ,  $p_s(s)$ , as a function of their dispersion measure, DM, for the Parkes multibeam observations, weighted by  $s$  for clarity. This is calculated for best-fitting parameters using  $E_{\min} = 10^{30}$  erg (top), and  $E_{\min} = 10^{39.0}$  erg (bottom).

to whether or not a strong repeater at  $z < 0.2$  was present in these fields. We observe that the low-DM deficit is greatest in ASKAP/FE observations, and lowest for Parkes observations, consistent with this prediction.

Until more is known about the population of repeating FRBs, we cannot further quantify this effect, except to add that this effect is guaranteed to be present to some degree (some FRBs definitely do repeat), with its importance increasing if all FRBs are explained by strong repeaters, and lessening as a lower fraction of all FRBs are due to repeaters, and those repeaters are weak.

## 8.5 Minimum search DM

Many FRB surveys use a minimum searched DM, either to exclude Galactic FRB candidates, or due to local RFI with intrinsic DM of 0 that none the less contaminates searches. For example, Thornton et al. (2013) reject bursts with  $DM < 100 \text{ pc cm}^{-3}$ , although in our model nearby FRBs could have a DM of as little as  $DM \approx 80 \text{ pc cm}^{-3}$ . This effect will sometimes therefore exclude FRBs in the very nearby Universe, potentially resulting in the observed deficit.

At initial completion of this work, only two bursts – 171020 (Shannon et al. 2018) and 180729.J0558+56 (CHIME/FRB Collaboration 2019a) – had been reported near this limit, however during the referee period, Bhardwaj et al. (2021) reported an FRB with DM  $87.82 \text{ pc cm}^{-3}$ . That CHIME have only reported one such FRB from

several years' worth of observations however implies that the number of these events is low – as are the number of FRBs predicted to be in this range by our model. The observed deficit extends well above  $100 \text{ pc cm}^{-3}$ , so we consider this an unlikely cause.

Could FRB search algorithms be less efficient at low DM? Tests on Galactic pulsars have shown the FREDDA algorithm used in ASKAP FRB observations to be equally efficient ( $\sim 90$  per cent of bursts detected) at  $67.99$  as at  $478.8 \text{ pc cm}^{-3}$  (James et al. 2019a), and follow-up (unpublished) studies have shown that Parkes/MB observations have similarly high levels of efficiency. Analysis of historical Parkes data by Zhang et al. (2020b) has found one new FRB at  $350 \text{ pc cm}^{-3}$  (which was too recent to include in this analysis), but none at lower DM. We therefore do not consider a DM-dependent search efficiency to be a likely cause of the observed deficit.

### 8.6 Summary: redshift and dispersion measure distributions

Having established the robustness of these results, we summarize the predictions for the redshift distributions from Fig. 9. In all scenarios, the redshift distribution of the ASKAP/FE detections lies in the range  $z_{\text{max}} < 0.8$ , and most fits find  $z_{\text{max}} < 0.6$ . Over all parameter sets, between 27 and 48 per cent of ASKAP/FE bursts should originate from within  $z < 0.1$ , confirming that these bursts are ideal follow-up targets due to their likely proximity. This suggests that the limits set on the repeatability of individual FRBs by James et al. (2020a) are significantly stronger than published, since those authors conservatively assume maximal redshifts. It also lends additional weight to the possible association of FRB 171020 with a galaxy at  $40 \text{ Mpc}$  by Mahony et al. (2018).

The predicted  $z$ -distribution of Parkes bursts is significantly broader than that of ASKAP/FE observations, although 3–20 per cent (over all parameter sets) of observed FRBs are predicted to be within  $z < 0.1$ .

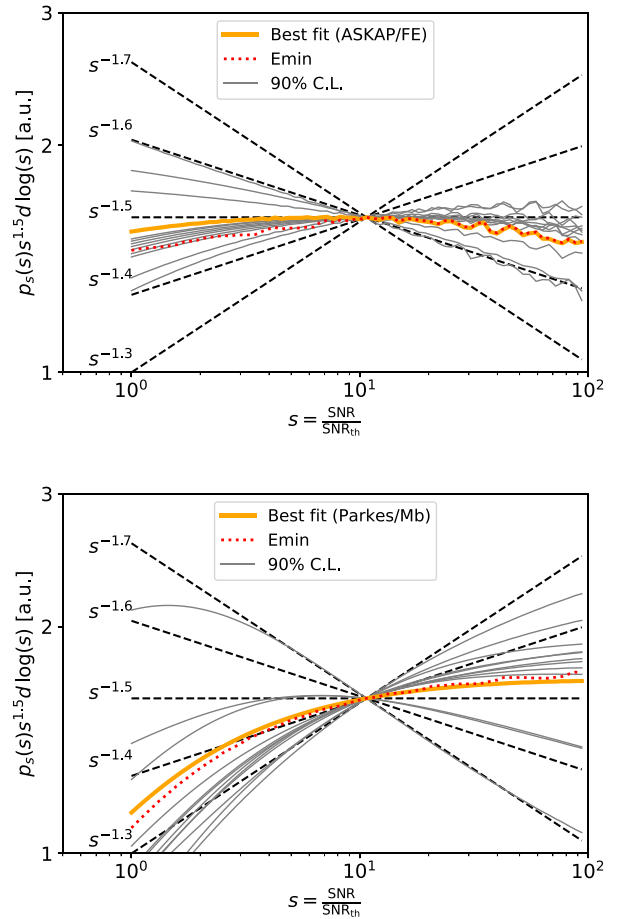
A key test of our prediction of a large number of near-Universe FRBs will be future ASKAP/ICS detections. So far, all bursts detected by ASKAP's ICS mode have been in a limited range of both DM and  $z$ , which seem to be from the central part of all predicted distributions. Our best-fitting model predicts that 23 per cent of ASKAP/ICS localizations should lie within  $z < 0.1$ , with a range over all sets of 14–33 per cent. To date, none have been observed – even amongst the unpublished ones not included here.

### 8.7 Source counts distribution

The slope of the source counts ('logN–logS') distribution was one of the first FRB observables to be analysed. Adapted from its original use in the study of radio galaxies characterized by their flux  $S$  (Ryle 1968), applied to FRBs, the source counts distribution is the number  $N$  of observed FRBs as a function of fluence threshold  $F_{\text{th}}$ . In an infinite Euclidean Universe, the distribution is expected to have a form

$$N(F_{\text{th}}) = C \left( \frac{F_{\text{th}}}{F_0} \right)^a, \quad (42)$$

with  $a = -1.5$ , and  $C$  and  $F_0$  normalizing constants. Studies using a variety of methods, with different treatments – or neglect of – observational biases in  $F_{\text{th}}$ , and using data from telescopes with different detection thresholds, found inconsistent values of  $a$  in the range of  $-0.8 \geq a \geq -2.6$  (Vedantham et al. 2016; Lawrence et al. 2017; Bhandari et al. 2018; Oppermann et al. 2018; Macquart & Ekers 2018a). James et al. (2019b), reviving methods applied to radio galaxy studies by Crawford et al. (1970), argue that  $s$  – defined



**Figure 14.** Predicted source counts distribution,  $p(s)$ , multiplied by  $s^{1.5}$  for clarity, for top: ASKAP/lat50, and bottom: Parkes/Mb, according to the best-fitting parameter set (black solid), when setting  $E_{\text{min}} = 10^{39} \text{ erg}$  (red dashed), and when varying parameters within their 90 per cent C.L. (thin grey). Also shown are lines of constant power-law slopes to guide the eye. The fluctuations above  $s = 10$  are due to finite gridding in  $z$  and DM.

in equation (33) – is a bias-independent measure of the source-counts slope, and find  $a = -1.18 \pm 0.24$  (68 per cent C.L.) for Parkes FRBs, and  $a = -2.2 \pm 0.47$  for ASKAP/FE FRBs, equating to a  $2.6\sigma$  tension. This was qualitatively consistent with Macquart & Ekers (2018b), who argue that at high values of  $F_{\text{th}}$ , the slope should be Euclidean ( $a = -1.5$ ), while the parameters of the FRB population will lead to a flattening at lower thresholds.

Our model can be used to derive the expected distribution of  $s$  using equation (34), by integrating over all values of DM and  $z$ , then converting this differential distribution to a cumulative distribution. The results for ASKAP/FE and Parkes/Mb observations are given in Fig. 14.

We immediately see that in all scenarios, ASKAP/FE observations, with a higher base threshold of  $26 \text{ Jy ms}$ , are expected to follow a Euclidean ( $a \sim -1.5$ ) distribution, while near-threshold ( $1 \leq s \leq 10$ ) Parkes/Mb observations exhibit a flatter source-counts index near  $a = -1.3$ . Different scenarios predict source-counts indices in the range  $-1.3 \leq a \leq -1.7$  above  $s = 10$  for Parkes/Mb observations. The existence or otherwise of a minimum energy does not significantly affect the distribution.

Our results in all scenarios are consistent with the findings of James et al. (2019b), with the greatest tension – about  $1.5\sigma$  – being between the source-counts slope that those authors find for ASKAP

of  $-2.2 \pm 0.47$  (68 per cent C.L.), and the approximate range of  $-1.4$  to  $-1.6$  found here. In particular, we also find that the source-counts index for Parkes/Mb is expected to be flatter than for ASKAP/lat50, and furthermore, the apparent ‘deficit’ of FRBs with low values of  $s$  is potentially attributable to the true behaviour of the FRB population, rather than statistical fluctuations or a measurement bias. This also suggests that the lower values of  $a$  found by previous authors – Vedantham et al. (2016), Oppermann et al. (2018), and Lawrence et al. (2017) – incorporating data from telescopes more sensitive than Parkes may have been correct.

## 9 THE $z$ -DM DISTRIBUTION

We argue in this work that the best representation of the FRB population observable by a telescope is a 2D function of extragalactic (cosmological plus host) dispersion measure, and redshift. Necessarily, the observable fraction of this distribution is a function of survey parameters, and also local DM contribution, which reduces sensitivity as it increases at lower Galactic latitudes. Our best-fitting  $z$ -DM distributions for the three surveys considered are plotted in Fig. 15.

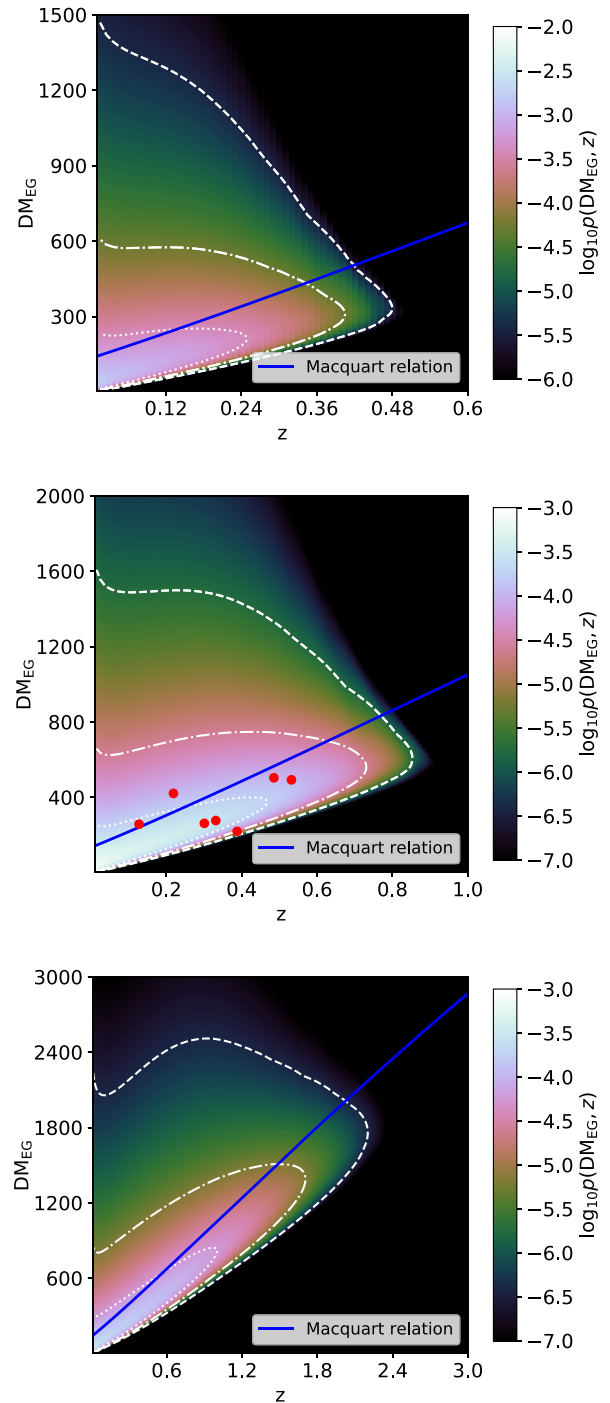
We describe the general features of these plots. Most FRBs are expected to lie near to, or below, the Macquart relation, being the 1–1 correspondence of FRB redshift with DM. This relation, with a slope of approximately  $100 \text{ pc cm}^{-3}$  per 0.1 in redshift, continues up to some maximum detectable distance, being approximately  $z=0.5, 0.85$ , and  $2.2$  for ASKAP/FE, ASKAP/ICS, and Parkes/Mb, respectively. However, the majority of FRBs will not be found near the maximum redshift, simply because there are very many more FRBs with lower energy, and much more of the sky covered at lower sensitivity. This is most evident with Parkes/Mb observations, where the most likely half of all FRB observations will arise on the Macquart relation with  $z < 1$ .

Off the Macquart relation, there is a significant fraction of FRBs expected to be found with higher than expected DMs, due to a combination of their host and cosmological contributions. Only our ASKAP/ICS (localized) FRB sample has provable examples, being FRB 190714 ( $504.7 \text{ pc cm}^{-3}$  from  $z = 0.209$ ) and FRB 190608 ( $339.5 \text{ pc cm}^{-3}$  from  $z = 0.1178$ ), although a third (FRB 191001, with  $506.92 \text{ pc cm}^{-3}$  from  $z = 0.23$ ) was excluded due to being detected at a lower frequency. This effect is less pronounced for more-sensitive surveys, since excess host contributions are dominated by cosmological ones. We emphasize that much of the structure above the Macquart relation – i.e. in the high-DM region – is poorly constrained, since our adopted lognormal distributions may not reflect reality.

At very high DMs, only near-Universe FRBs are observable, since a burst must be observed with very high fluence to overcome the detection bias against high DM. The upper bound of the 99 per cent region (dashed lines) slopes backward, against the Macquart relation, since more low-energy FRBs with large excess DM are predicted than high-energy FRBs lying on the Macquart relation. We do not expect our quantitative estimates in this region to be accurate until it is directly probed with localized FRBs. Since the cause of this effect is a well-understood observational bias; however, it will clearly be present to some degree.

### 9.1 The Macquart relation

The ‘Macquart relation’ is the general one-to-one-ness of the relationship between redshift and DM of FRBs. It is predicted from the distribution of baryonic matter in the Universe (Inoue 2004), and first



**Figure 15.** Normalized probability distribution  $p(z, \text{DM}_{\text{EG}})$  of the redshift  $z$  and extragalactic dispersion measure  $\text{DM}_{\text{EG}}$  of FRBs detected in ASKAP Fly’s Eye (top), ICS (middle), and Parkes multibeam (bottom) surveys. Contours indicate the intervals in which 50 (dotted), 90 (dot-dashed), and 99 per cent (dashed) of all FRBs would be detected. Red dots indicate the values of  $\text{DM}_{\text{EG}}$  and host galaxy redshift  $z$  for ASKAP ICS localized FRBs.

evidence for it was given in Shannon et al. (2018), by comparing the DM distributions of ASKAP/FE and Parkes/Mb populations, where the higher sensitivity of Parkes allowed it to probe more-distant FRBs with higher DMs. The relation was first observed directly by Macquart et al. (2020), who showed that the redshifts of localized FRBs were consistent with the baryonic content of the Universe.

For the purpose of comparing survey results, we propose a useful distinction: the ‘weak’ Macquart relation (which might more accurately be titled the ‘Shannon’ relation), where telescopes with higher sensitivity on average observe more-distant FRBs with higher DM; and the ‘strong’ version (or true Macquart relation), where the DMs of FRBs *within* a survey are a good proxy for their redshift. Several authors use a 1–1  $z$ –DM relation in performing estimates of the FRB population from the DMs of un-localized FRBs (Shannon et al. 2018; Deng, Wei & Wu 2019; Lu & Piro 2019; James et al. 2020a).

Clearly, the weak version of the relation is well-established, and an obvious consequence of the cosmological nature of FRBs. Here, we test the strong version, which can be readily tested by calculating

$$p(z|\text{DM}) = \frac{p(z, \text{DM})}{p(\text{DM})} \quad (43)$$

for a given survey. This is shown for each survey in Fig. 16.

In each survey, the Macquart relation applies up to a maximum redshift  $z_{\text{max}}$ , beyond which it reverses. The reversal can be intuitively understood by realizing that at the maximum redshift at which an FRB can be detected, FRBs with any excess of DM cannot be detected, due to DM smearing reducing sensitivity. Therefore, the only way to detect FRBs with a DM lying above that expected from a burst originating at  $z_{\text{max}}$  is to have the burst originate at a nearer redshift. As noted in Section 3 and Fig. 2, the increased number of FRBs in the local Universe is related to the cumulative luminosity index  $\gamma = -1.09$ , with more negative values leading to more nearby high-DM events.

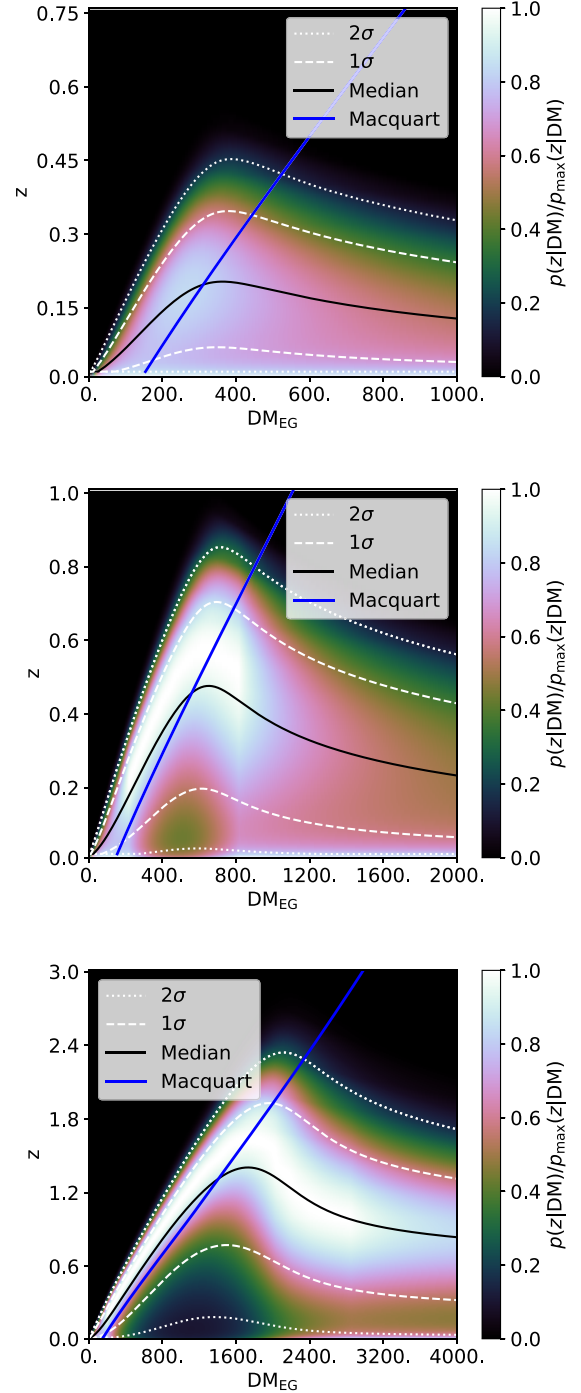
The reversal of the Macquart relation has several practical consequences. First: for surveys with a large sample of FRBs, the burst with the greatest DM will *not* be the most distant. An excellent example of this phenomenon is FRB 170428 (ASKAP; Shannon et al. 2018), which is most likely to originate below  $z = 0.3$ , rather than the value of  $z \sim 1$  expected from the Macquart relation. FRB 160102, observed by Parkes, is another likely candidate. The implication is that works using a 1–1 DM– $z$  relation will vastly overestimate the maximum FRB energy, since they will necessarily attribute a large distance and therefore high energy to the highest DM event, which may in fact be quite local.

A key test of this reversal would be the detection of an FRB with  $\text{DM} \gtrsim 1000 \text{ pc cm}^{-3}$  by ASKAP in ICS mode, and its subsequent localization to a redshift  $z \lesssim 0.6$ . Again, we note that this reversal of the Macquart relation will always be present to some extent, since it is fundamentally due to observational effects which are known and understood – it is merely the extent of this effect, and the DM for any given survey above which it occurs, that is currently uncertain.

Finally, we note that the existence of FRBs with very high DMs has raised the possibility of using FRBs to probe for the signature of Helium reionization (Deng & Zhang 2014; Caleb et al. 2019; Linder 2020). While this is by no means ruled out, it emphasizes that doing so will require FRBs to be localized, since simple measures of FRB properties as a function of DM will yield a very large scatter in redshifts, and hence reduced statistical power.

## 10 CONCLUSION

We have developed a precise model of FRB observations, including observational biases due to the full telescope beamshape, degradation in efficiency due to DM, and intrinsic burst width. None of these effects are fundamentally new; many others should take credit for highlighting their importance, and Luo et al. (2020) should be attributed with a first analysis using these techniques. Here, we have



**Figure 16.** Probability  $p(z|\text{DM})$  of an FRB originating from redshift  $z$  given it has been observed with the specified DM in the ASKAP/FE (top), ASKAP/ICS (middle), and Parkes/Mb (bottom) surveys.

improved upon this method by using an unbiased data sample, adding new observations of localized FRBs, studying the effects of source evolution, including the likelihood of the observed signal strength, and improving the beam model. We show that ignoring, or incorrectly modelling, these factors leads to significant biases in the expected redshifts of observable FRBs.

We have also highlighted how uncertainties in the spectral index  $\alpha$ , and indeed even in the *interpretation* of  $\alpha$  as either a true FRB spectral index or a frequency-dependent rate, is a systematic uncertainty in



the modelling, and resolving this question should be a focus of future research. We suggest treating both interpretations of  $\alpha$  as being equally plausible for the time being.

We have used our approach to model FRB observations with ASKAP in both fly’s eye and incoherent sum mode, and Parkes multibeam observations. We have carefully selected our data to ensure it is not biased due to under-reporting of observation time, or due to large local DM contributions reducing sensitivity. Crucially, we have included a sample of localized FRBs from ASKAP for which the redshift of the host galaxies is measured.

The  $z$ , DM, and SNR distributions of FRBs predicted by our best-fitting population estimates, presented in James et al. (2021a), are tested against observations. We find no evidence for, and some evidence against, a lower bound to the FRB energy distribution, although we only exclude  $E_{\min} \geq 10^{39.0}$  erg (90 per cent C.L.). No such minimum energy is expected for the magnetar-origin hypothesis, which links the observed extragalactic FRB population to radio bursts from magnetar flares in our own Galaxy.

Our model also allows us to make inferences on the redshifts of the un-localized samples of FRBs detected by ASKAP and Parkes. We find these to be somewhat lower than expected from the Macquart relation, and indeed that the highest-DM events are likely not the most distant, due to observational effects causing an inversion of the Macquart relation, and the relatively steep best-fitting value of the FRB energy distribution. The ability of this model to place priors on the expected redshift of FRBs given their measured DMs will also aid FRB localization efforts, especially for those bursts – such as FRB 171020 – which have uncertain, but promising, host associations.

For the first time, we have incorporated the measured signal-to-noise ratio  $s$  into FRB population modelling, allowing the use of this observable to constrain population parameters, and to predict the source-counts (‘logN–logS’) distribution. In all scenarios, we find a steepening of this distribution from Parkes to ASKAP, consistent with the predictions of Macquart & Ekers (2018b) and the observations of James et al. (2019b).

Anomalies between observations and our model include a lack of ASKAP incoherent sum (localized) FRBs within  $z < 0.1$ , an over (under) prediction of FRBs from Parkes multibeam (ASKAP fly’s eye) observations, and an overprediction of low-DM bursts. None of these have a high statistical significance – however, we identify several potential explanations for these discrepancies. These are a spectral break in the FRB luminosity function, and the influence of the repeating FRB population. We recommend future works attempt to improve this modelling, and also improve understanding of the spectral behaviour of FRBs.

Future observations of localized FRBs with very large excess DMs, and/or from the local Universe, would verify predictions from this work. FRB surveys with very sensitive radio telescopes such as the Five-hundred-meter Aperture Spherical Telescope (FAST), or repeating previous Parkes and ASKAP surveys at different frequencies, would help to further constrain FRB population models. In particular, the application of this model to the large sample of bursts observed by CHIME would be particularly useful, although it would then need to be adapted to include repeating FRBs. We also aim to investigate improved numerical/computational methods to speed up calculations to allow the inclusion of further data.

## ACKNOWLEDGEMENTS

This research has made use of NASA’s Astrophysics Data System Bibliographic Services. This research made use of PYTHON libraries

MATPLOTLIB (Hunter 2007), NUMPY (van der Walt, Colbert & Varoquaux 2011), and SCIPY (Virtanen et al. 2020). This work was performed on the GPU Supercomputer for Theoretical Astrophysics Research (gSTAR) national facility at Swinburne University of Technology. gSTAR is funded by Swinburne and the Australian Government’s Education Investment Fund. This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. This research was partially supported by the Australian Government through the Australian Research Council’s Discovery Projects funding scheme (projects DP180100857 and DP210102103).

## DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

- Agarwal D. et al., 2019, *MNRAS*, 490, 1  
 Arcus W. R., Macquart J.-P., Sammons M. W., James C. W., Ekers R. D., 2021, *MNRAS*, 501, 5319  
 Bannister K. W. et al., 2017, *ApJ*, 841, L12  
 Bannister K. W. et al., 2019, *Science*, 365, 565  
 Bhandari S. et al., 2018, *MNRAS*, 475, 1427  
 Bhandari S. et al., 2020a, *ApJ*, 895, L37  
 Bhandari S. et al., 2020b, *ApJ*, 901, L20  
 Bhandari S., Bannister K. W., James C. W., Shannon R. M., Flynn C. M., Caleb M., Buntun J. D., 2019, *MNRAS*, 486, 70  
 Bhardwaj M. et al., 2021, *ApJ*, 910, L18  
 Bochenek C. D., Ravi V., Belov K. V., Hallinan G., Kocz J., Kulkarni S. R., McKenna D. L., 2020, *Nature*, 587, 59  
 Caleb M., Flynn C., Bailes M., Barr E. D., Hunstead R. W., Keane E. F., Ravi V., van Straten W., 2016, *MNRAS*, 458, 708  
 Caleb M., Flynn C., Stappers B. W., 2019, *MNRAS*, 485, 2281  
 Cao X.-F., Yu Y.-W., Zhou X., 2018, *ApJ*, 858, 89  
 Champion D. J. et al., 2016, *MNRAS*, 460, L30  
 Chatterjee S. et al., 2017, *Nature*, 541, 58  
 Chawla P. et al., 2017, *ApJ*, 844, 140  
 CHIME/FRB Collaboration, 2019a, *Nature*, 566, 230  
 CHIME/FRB Collaboration, 2019b, *Nature*, 566, 235  
 CHIME/FRB Collaboration, 2019c, *ApJ*, 885, L24  
 Chime/Frb Collaboration, 2020, *Nature*, 582, 351  
 Chittidi J. S. et al., 2020, preprint (arXiv:2005.13158)  
 Cho H. et al., 2020, *ApJ*, 891, L38  
 Connor L. et al., 2020, *MNRAS*, 499, 4716  
 Connor L., 2019, *MNRAS*, 487, 5753  
 Cordes J. M., Lazio T. J. W., 2002, preprint (arXiv:astro-ph/0207156)  
 Cordes J. M., McLaughlin M. A., 2003, *ApJ*, 596, 1142  
 Crawford D. F., Jauncey D. L., Murdoch H. S., 1970, *ApJ*, 162, 405  
 Cruces M. et al., 2020, *MNRAS*, 500, 448  
 Day C. K. et al., 2020, *MNRAS*, 497, 3335  
 Deng W., Zhang B., 2014, *ApJ*, 783, L35  
 Deng C.-M., Wei J.-J., Wu X.-F., 2019, *J. High Energy Astrophys.*, 23, 1  
 Farah W. et al., 2018, *MNRAS*, 478, 1209  
 Farah W. et al., 2019, *MNRAS*, 488, 2989  
 Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601  
 Fonseca E. et al., 2020, *ApJ*, 891, L6  
 Gardenier D. W., van Leeuwen J., Connor L., Petroff E., 2019, *A&A*, 632, A125  
 Gardenier D. W., Connor L., van Leeuwen J., Oostrum L. C., Petroff E., 2021, *A&A*, 647, A30  
 Gupta V. et al., 2021, *MNRAS*, 501, 2316  
 Hardy L. K. et al., 2017, *MNRAS*, 472, 2800  
 Heintz K. E. et al., 2020, *ApJ*, 903, 152

- Hessels J. W. T. et al., 2019, *ApJ*, 876, L23
- Hobbs G. et al., 2020, *Publ. Astron. Soc. Aust.*, 37, e012
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Inoue S., 2004, *MNRAS*, 348, 999
- Ioka K., 2003, *ApJ*, 598, L79
- James F., 2006, *Statistical Methods in Experimental Physics: 2nd Edition*, World Scientific, Singapore
- James C. W., 2019, *MNRAS*, 486, 5934
- James C. W. et al., 2019a, *PASA*, 36, e009
- James C. W., Ekers R. D., Macquart J.-P., Bannister K. W., Shannon R. M., 2019b, *MNRAS*, 483, 1342
- James C. W. et al., 2020a, *MNRAS*, 495, 2416
- James C. W. et al., 2020b, *ApJ*, 895, L22
- James C., Prochaska J., Macquart J.-P., North-Hickey F., Bannister K., Dunning A., 2021a, *MNRAS*
- James C. W., Prochaska J. X., Ghosh E. M., 2021b, *zdm*, <https://zenodo.org/record/5213780#.YRxh5BMzZKA>
- Keane E. F., Petroff E., 2015a, *MNRAS*, 447, 2852
- Keane E., Petroff E., 2015b, *MNRAS*, 447, 2852
- Keane E. F. et al., 2017, *MNRAS*, 473, 116
- Keating L. C., Pen U.-L., 2020, *MNRAS*, 496, L106
- Keith M. J. et al., 2010, *MNRAS*, 409, 619
- Kolmogorov A., 1933, *G. Ist. Ital. Attuari.*, 4, 83
- Kumar P. et al., 2019, *ApJ*, 887, L30
- Kumar P. et al., 2021, *MNRAS*, 500, 2525
- Lawrence E., Vander Wiel S., Law C., Burke Spolaor S., Bower G. C., 2017, *AJ*, 154, 117
- Law C. J. et al., 2017, *ApJ*, 850, 76
- Law C. J. et al., 2020, *ApJ*, 899, 161
- Li D. et al., 2021, *Nature*, 598, 267
- Linder E. V., 2020, *Phys. Rev. D*, 101, 103019
- Lorimer D. R., Bailes M., McLaughlin M. A., Narkevic D. J., Crawford F., 2007, *Science*, 318, 777
- Lu W., Piro A. L., 2019, *ApJ*, 883, 40
- Luo R., Lee K., Lorimer D. R., Zhang B., 2018, *MNRAS*, 481, 2320
- Luo R., Men Y., Lee K., Wang W., Lorimer D. R., Zhang B., 2020, *MNRAS*, 494, 665
- Macquart J.-P., Ekers R. D., 2018a, *MNRAS*, 474, 1900
- Macquart J.-P., Ekers R., 2018b, *MNRAS*, 480, 4211
- Macquart J.-P., Koay J. Y., 2013, *ApJ*, 776, 125
- Macquart J.-P., Shannon R. M., Bannister K. W., James C. W., Ekers R. D., Bunton J. D., 2019, *ApJ*, 872, L19
- Macquart J. P. et al., 2020, *Nature*, 581, 391
- Madau P., Dickinson M., 2014, *ARA&A*, 52, 415
- Madhavacheril M. S., Battaglia N., Smith K. M., Sievers J. L., 2019, *Phys. Rev. D*, 100, 103532
- Mahony E. K. et al., 2018, *ApJ*, 867, L10
- Marcote B. et al., 2020, *Nature*, 577, 190
- Masui K. W., Sigurdson K., 2015, *Phys. Rev. Lett.*, 115, 121301
- McQuinn M., 2014, *ApJ*, 780, L33
- Michilli D. et al., 2018, *Nature*, 553, 182
- Murdoch H. S., Crawford D. F., Jauncey D. L., 1973, *ApJ*, 183, 1
- Oppermann N., Yu H.-R., Pen U.-L., 2018, *MNRAS*, 475, 5109
- Osłowski S. et al., 2019, *MNRAS*, 488, 868
- Patek C., Chime/Frb Collaboration, 2019, *Astron. Telegram*, 13013, 1
- Petroff E. et al., 2014, *ApJ*, 789, L26
- Petroff E. et al., 2015, *MNRAS*, 447, 246
- Petroff E. et al., 2016, *Publ. Astron. Soc. Aust.*, 33, e045
- Petroff E. et al., 2019, *MNRAS*, 482, 3109
- Planck Collaboration VI, 2018, *A&A*, 641, A6
- Platts E., Weltman A., Walters A., Tendulkar S. P., Gordin J. E. B., Kandhai S., 2019, *Phys. Rep.*, 821, 1
- Platts E., Prochaska J. X., Law C. J., 2020, *ApJ*, 895, L49
- Pleunis Z. et al., 2021, preprint ([arXiv:2106.04356](https://arxiv.org/abs/2106.04356))
- Prochaska J. X., Simha S., Law C., Tejos N., Neeleman M., 2019a, *FRB*, <https://zenodo.org/record/3403651#.YRxkcBMzZKA>
- Prochaska J. X. et al., 2019b, *Science*, 365, aay0073
- Prochaska J. X., Zheng Y., 2019, *MNRAS*, 485, 648
- Qiu H., Bannister K. W., Shannon R. M., Murphy T., Bhandari S., Agarwal D., Lorimer D. R., Bunton J. D., 2019, *MNRAS*, 486, 166
- Qiu H. et al., 2020, *MNRAS*, 497, 1382
- Rajwade K. M. et al., 2020, *MNRAS*, 495, 3551
- Ravi V., Shannon R. M., Jameson A., 2015, *ApJ*, 799, L5
- Ravi V. et al., 2016, *Science*, 354, 1249
- Ravi V. et al., 2019, *Nature*, 572, 352
- Ryle M., 1968, *ARA&A*, 6, 249
- Safarzadeh M., Prochaska J. X., Heintz K. E., Fong W.-fai., 2020, *ApJ*, 905, L30
- Shannon R. M. et al., 2018, *Nature*, 562, 386
- Smirnov N., 1948, *Ann. Math. Stat.*, 19, 279
- Sokolowski M. et al., 2018, *ApJ*, 867, L12
- Spitler L. G. et al., 2014, *ApJ*, 790, 101
- Staveley-Smith L. et al., 1996, *Publ. Astron. Soc. Aust.*, 13, 243
- Tendulkar S. P. et al., 2017, *ApJ*, 834, L7
- The Chime/Frb Collaboration, 2020, *Nature*, 587, 54
- The CHIME/FRB Collaboration, 2021, preprint ([arXiv:2106.04352](https://arxiv.org/abs/2106.04352))
- Thornton D. et al., 2013, *Science*, 341, 53
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput. Sci. Eng.*, 13, 22
- Vedantham H. K., Ravi V., Hallinan G., Shannon R. M., 2016, *ApJ*, 830, 75
- Virtanen P. et al., 2020, *Nat. Methods*, 17, 261
- Wilks S., 1962, *Mathematical Statistics*. John Wiley and Sons Ltd, New York
- Zhang S. B., Hobbs G., Dai S., Toomey L., Staveley-Smith L., Russell C. J., Wu X. F., 2019, *MNRAS*, 484, L147
- Zhang S. B. et al., 2020a, *ApJS*, 249, 14
- Zhang S. B. et al., 2020b, *ApJS*, 249, 14
- Zhu W., Feng L.-L., Zhang F., 2018, *ApJ*, 865, 147

## APPENDIX A: NEGLECTED EFFECTS

Here, we discuss effects that affect search sensitivity which have been neglected in this work.

### A1 Fine sensitivity effects

The precise sensitivity of an incoherent FRB search to an FRB depends not only on the ‘coarse’ effects of effective pulse width, but also the exact alignment of a burst with the time–frequency binning of the data compared to the FRB arrival time, and the DMs which are searched compared to the exact DM of the FRB. This effect is analysed by Keane & Petroff (2015a) for early FRB search methods, finding fluctuations of up to  $\pm 15$  per cent, although in current FRB search methods fluctuations are at the level of a few percent. This was also found to be the case for an internal investigation into the CRAFT FRB searches with ASKAP for Shannon et al. (2018), with fluctuations of  $\pm 3$  per cent. Such effects are thus ignored in this work, and likely should be in all future works.

The presence of radio-frequency interference (RFI) during observations can result in a loss of effective bandwidth – and hence sensitivity – via vetoed frequency channels; loss of effective observing time if the RFI results in FRBs being completely unobservable; and reduced sensitivity to FRBs in a certain parameter space (particularly low-DM or high-width FRBs). The sensitivity of ASKAP to such effects has been studied using pulsar calibration observations by James et al. (2019a), finding a typical 10 per cent loss of effective observation time and  $\sim 10$  per cent fluctuation in sensitivity.

Further details of search pipelines can effect burst sensitivity. The zero-DM subtraction method – i.e. subtracting the mean detected power prior to dedispersion – will reduce sensitivity to low-DM events and very bright bursts, although typically only if the total dispersion sweep across the detection bandwidth is low. For instance, we have found a few percent bias in the estimated SNR for ASKAP

searches when the expected SNR is greater than 100. An analysis of the HEIMDALL software – used in most Parkes FRB searches – using real-time injected FRBs by Gupta et al. (2021) finds a 20 per cent reduction in SNR compared to expectations. This mostly affected bursts with widths above 20 ms however, and we find in Section 4 our estimates of the DM- $z$  distribution to be insensitive to the population of very wide bursts. For ASKAP, the expected SNR of pulsar bursts was found by Shannon et al. (2018) to be within a few percent of that found by the Fast Real-time Engine for Dedispersing Amplitudes (FREDDA) algorithm used by the CRAFT collaboration. Nonetheless, for future precision cosmological studies, the response of FRB detection systems should be quantified in more detail.

We also note that in this work we use SNR as the true SNR of the FRB, when by definition this has a  $1\sigma$  error of unity. Murdoch, Crawford & Jauncey (1973) find that for SNR above six, the measured SNR can be approximated as the true SNR for statistical purposes, and so we have here ignored this effect, rather than marginalising over it. Possible non-linearities between FRB fluence and SNR become important at high SNR, e.g. as found for FRB 180309, which saturated the Parkes digitiser system and may have had a true SNR as high as 2616 (Osłowski et al. 2019). However, these effects can be corrected-for offline; care must be taken however to use the corrected values for model evaluation. Small effects are also possible at lower SNR, e.g. as found for high-width FRBs at UTMOST (Gupta et al. 2021).

Our beamshape model for Parkes uses the mean observation frequency, while for ASKAP it is uniformly weighted over all frequencies. A preference for e.g. low-frequency bursts would result in a slight increase in the survey effective area for both instruments, with only the difference in increase being relevant when estimating population parameters other than the absolute FRB rate. We consider this a third-order effect for the frequencies and bandwidths considered here.

Finally, we have not discussed errors in  $F_0$ . The absolute scale of the mean ASKAP threshold was calculated by James et al. (2019a) by referencing observations to Parkes, and using Hydra A as an absolute flux calibrator. Thus we expect errors in flux calibration for Parkes and ASKAP to be linked, and cancel to first order. However, individual antenna sensitivity for ASKAP was found to vary by  $\sim \pm 5$  per cent, and this is similar to the level of uncertainty found when performing calibration observations for Bannister et al. (2019). Perhaps the largest source of uncertainty in  $F_0$  comes from the quoted threshold of 0.5 Jy ms for Parkes, which has an inherent rounding uncertainty of  $\pm 10$  per cent, and thus a corresponding rate uncertainty of  $\pm 15$  per cent for a cumulative source counts index of  $-1.5$  (see Section 8.7).

## A2 Pointing

Here, we treat a given FRB survey as having a constant local DM component,  $\text{DM}_{\text{local}} = \text{DM}_{\text{MW}} + \text{DM}_{\text{host}}$ . This is because a greater local DM reduces sensitivity to extragalactic FRBs due to DM smearing.

When this effect becomes significant, this requires extending the integral in equation (29) to

$$\begin{aligned} \langle N_i \rangle = & \sum_k T_{i,k} \int dz \Phi(z) \frac{dV(z)}{d\Omega dz} \int d\text{DM}_{\text{EG}} p(\text{DM}_{\text{EG}}|z) \\ & \times \int dB \Omega(B) \int dw p(w) p(E > E_{\text{th}}), \end{aligned} \quad (\text{A1})$$

$$E_{\text{th}} \sim E_{\text{th}}(B, w, z, \text{DM} = \text{DM}_{\text{EG}} + \text{DM}_{\text{local},k}), \quad (\text{A2})$$

where  $T_{i,k}$  is the total time spent observing at  $\text{DM}_{\text{local}} = \text{DM}_{\text{local},k}$ .

## A3 Scattering

In Section 4, we combine the intrinsic burst width  $w_{\text{int}}$  and scatter-broadening width  $w_{\text{scat}}$  into the incident width  $w_{\text{inc}}$ . However, while  $w_{\text{int}}$  can reasonably be assumed to be independent of other FRB properties, it is plausible that  $w_{\text{inc}}$  will be correlated with both DM and/or  $z$ , and it will certainly be frequency dependent. This is due to interstellar scattering, which is implicitly included through modelling of observed FRB widths, but is not explicitly accounted for as per e.g. Caleb et al. (2016). We largely avoid this question in this work by choosing surveys of similar frequency – however, we consider evidence for such a correlation here.

Most FRB searches have time resolutions in the range of  $100 \mu\text{s}$ – $1 \text{ ms}$ , and resolving scattering tails from the intrinsic burst structure is difficult: Qiu et al. (2020) are able to do this for only six of the 33 FRBs observed at the typically  $1.26 \text{ ms}$  real-time resolution of ASKAP FRB searches. The broader time-frequency structure exhibited by many repeating FRBs (CHIME/FRB Collaboration 2019c; Fonseca et al. 2020; The CHIME/FRB Collaboration 2021) however can often be resolved at these resolutions (e.g. Hessels et al. 2019), while observations down to  $10 \mu\text{s}$  have revealed a  $1 \text{ GHz}$  scattering time of only  $24 \mu\text{s}$  in FRB 121102, which is obtained indirectly from the measured scintillation bandwidth (Michilli et al. 2018).

The most reliable way to resolve these two contributions is to use time-frequency data at the Nyquist resolution. ASKAP (Cho et al. 2020; Day et al. 2020) and UTMOST (Farah et al. 2018, 2019) have analysed such data for six FRBs each. Bursts were found to have strong sub-burst structures down to  $50 \mu\text{s}$ . Scattering was conclusively measured for a total of four ASKAP FRBs, being in the range  $40 \mu\text{s}$ – $3.3 \text{ ms}$  at  $1.27 \text{ GHz}$ , while for 181112, it was at most  $20 \mu\text{s}$ . UTMOST FRBs had  $835 \text{ MHz}$  scattering times of  $4 \mu\text{s}$ – $30 \text{ ms}$ , with one upper limit at  $0.2 \text{ ms}$ .

Clearly, there is a broad distribution of FRB intrinsic widths and scattering times. There are general effects that this distribution can have on the DM- $z$  distribution of FRBs.

The most complicated potential interaction of scattering and the intrinsic width is one that is dependent on the exact position of the FRB in DM- $z$  space.

If indeed FRBs do arise from two source populations with different cosmological evolutions or different host galaxy properties, or otherwise one population of objects with properties that age on cosmological time-scales, the intrinsic width distribution may have some DM- $z$  dependence. This possibility should not be ignored. However, the most likely redshift dependence arises from the scattering term.

Theoretical studies have examined expectations for scattering of FRBs during cosmic propagation, with effects attributed to the intergalactic medium (IGM), intracluster medium (ICM), and the haloes and interstellar medium (ISM) of intersected galaxies.

The general form of the redshift dependence is analysed by Macquart & Koay (2013), finding that scattering due to the IGM will increase as  $(1+z)^2$  to  $z \lesssim 1$ , and as  $(1+z)^{0.2-0.5}$  for  $z \gtrsim 1$ . This is in contrast to the contribution from hosts, which scales as  $(1+z)^{-3}$ . Results on the absolute magnitude of the scattering depend on the assumed minimum and maximum length-scales of the turbulence. Macquart & Koay (2013) argue that for realistic turbulence parameters, the total amount of scattering from the IGM and ICM is expected to be low, at  $\lesssim 1$  and  $\lesssim 5 \text{ ms}$  at  $300 \text{ MHz}$ , respectively. Zhu, Feng & Zhang (2018) simulate FRBs propagating in a clumpy IGM. Examining the dependence of the mean scattering



time  $\bar{\tau}$  on DM, they find  $\bar{\tau} \sim \text{DM}^{1-2}$  for voids, clusters, and filaments over a range of simulation parameters. They also conclude that unreasonably high turbulence scales would be required to achieve scattering values comparable to that observed in FRBs.

In some cases, FRBs localizations have been sufficient to identify intersections of the line of sight with Galactic haloes, placing upper limits on the degree of scattering caused by such intersections (Prochaska et al. 2019b; Connor et al. 2020).

Qiu et al. (2020) find no evidence however for a DM-dependence of scattering. Both measurements and expectation therefore suggest that the observed distribution of scattering arises from the host galaxies. We do not however attempt to model this in this work, since the  $(1+z)^2$  reduction in width at high redshift will be in any case insignificant against the DM smearing effect. A full model of scattering will however become important when including observations over a wide frequency range.

#### A4 Influence of specific parameters

This work uses measured FRB detection numbers, dispersion measures, redshifts, and strengths from three different FRB searches. It is useful to probe the influence of each of these on our overall result. This also places an absolute limit on the aforementioned effects discussed in Appendix A – no matter how untrustworthy we consider the number of FRBs detected, the information contained in detection numbers will still be better than no information at all.

To do so, we show in Fig. A1 our confidence limits for each parameter, both with (top) and without (bottom) our prior on  $\alpha$ , when removing the likelihood corresponding to  $p_N$  (equation 27),  $p_{\text{DMZ}}$  (equation 30),  $p_s$  (equation 34), and when removing ASKAP/FE, ASKAP/ICE, and Parkes/Mb observations. We do this under the spectral index interpretation of  $\alpha$  – results are similar under the rate interpretation.

Unsurprisingly, the most dramatic effect occurs when removing information on measured DM and  $z$ . In this limit, we do not find lower bounds on  $E_{\text{max}}$  in the range  $\log_{10} E_{\text{max}} [\text{erg}] > 40$ , and limits on other parameters are very broad. This is not a very interesting, informative, or unexpected result, and we ignore the ‘No  $p_{\text{DMZ}}$ ’ in the following discussion. However, the fact that some constraints are none the less derived is evidence for the source-counts distribution of FRBs containing useful information (Macquart & Ekers 2018b).

For all other parameters, we obtain a variety of modified limits. Clearly, the ASKAP/ICS sample is most constraining to  $E_{\text{max}}$ , while without the ASKAP/FE sample, the constraints are narrower. This is natural, since the ASKAP/FE sample probes the most intrinsically luminous bursts, without which a fit for a narrow population distribution is possible.

The slope of the cumulative luminosity function  $\gamma$  is constrained to be flat by  $p_n$ , without which a steeper (lower) value would be obtained, while both the Parkes/Mb sample, and  $p_{\text{snr}}$ , prevent flat (higher) values. This is a very interesting result. As noted in our companion paper, our best fit underpredicts the number of ASKAP/FE FRBs (13.9 versus 20 in 1274.6 d), and overpredicts those for Parkes/Mb (17.3 versus 12 in 164.4 d). One would expect that removing either Parkes/Mb or  $p_n$  would have a similar effect by removing this tension. However, removing Parkes/Mb also removes the Parkes DM distribution. This then suggests that the Parkes DM distribution is indicative of a steep source-counts spectrum, whereas the number of FRBs detected by Parkes is indicative of a flat spectrum. We do not have a full solution to this quandry, although it is clearly related to the discussion in Section 8. We do not

elaborate further, since this discrepancy is of only marginal statistical significance.

Perhaps our most significant result – strong evidence for source evolution with redshift, i.e.  $n > 0$  – is most strongly disfavoured by  $p_n$ , such that when this is removed, we prefer a strongly evolving population. Removing any of the three FRB surveys, or  $p_{\text{snr}}$ , simply results in mildly less evidence for  $n > 0$  – there is no single result strongly favouring  $n > 0$ . We thus conclude that our conclusion on  $n$  robust, especially considering uncertainties discussed in Appendix A1.

Our host galaxy parameters,  $\mu_{\text{host}}$  and  $\sigma_{\text{host}}$  – recall these are in log DM space – are most affected by the highest and lowest sensitivity surveys, ASKAP/FE and Parkes/Mb. Without these observations, very low values of  $\mu_{\text{host}}$  (and hence very large values of  $\sigma_{\text{host}}$ ) become possible. This validates the use of DM distributions only when studying these parameters. Neither  $p_{\text{snr}}$  nor  $p_n$  have a large effect.

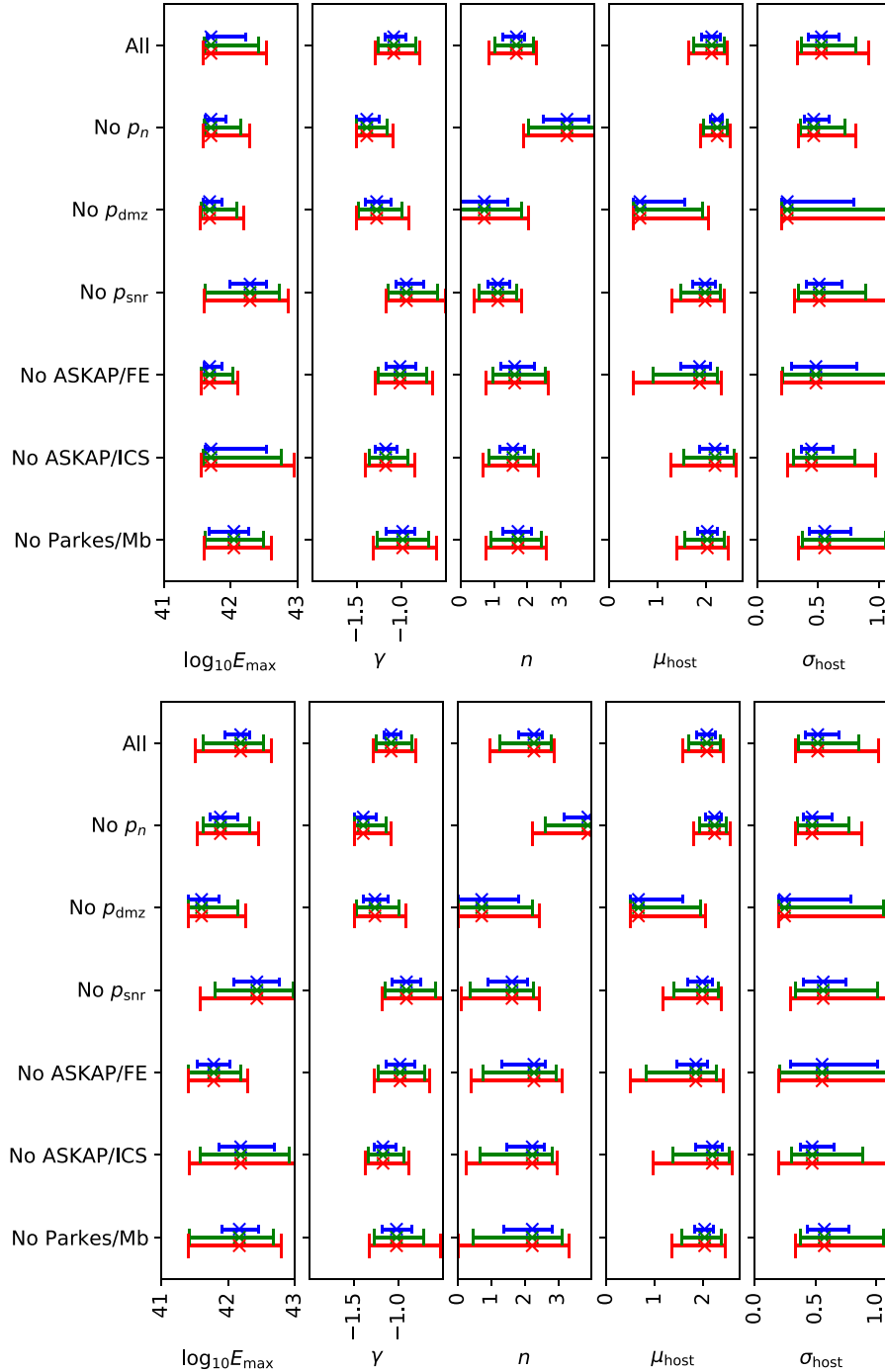
Are the changes when ignoring measurements indicative of statistical robustness? The most statistically significant change in best-fitting values occurs when ignoring  $p_n$ . This, predictably, is related to the low number of bursts observed by Parkes/Mb relative to ASKAP/FE compared to expectations discussed above, which alone argues against source evolution. The best-fitting value when ignoring  $p_n$  predicts that Parkes/Mb surveys should have observed five times as many FRBs as ASKAP/FE, as opposed to the 60 per cent (12/20) observed. We do not consider it possible that any of the fine sensitivity effects discussed in this Appendix could be responsible for a factor of 5 reduction in observable bursts at Parkes. Furthermore, ignoring  $p_n$  favours a very steep luminosity function ( $\gamma = -1.4$ ), which will tend to counteract the effect of the best-fitting strong source evolution ( $n = 3.2$ ). In isolation, these values are inconsistent with those found when including  $p_n$  at approximately the 99 per cent level, although including trial factors due to looking at the effects on five parameters when excluding six subsets of data reduces this significance to the 10 per cent level (since parameters are correlated, it is difficult to quote precise statistical significances). We thus interpret the large changes in best-fitting values of  $\gamma$  and  $n$  when excluding  $p_n$  as being due to the fit trying to account for anomalies in the  $z$ –DM distributions previously discussed in Section 8.1.

#### A5 Influence of uncertainty in Parkes sensitivity

To provide a more realistic estimate of potential systematic effects, we perform two tests. First, we test the effect of a systematic error in FRB detection threshold, as might occur if the reduced SNR found for UTMOST (Gupta et al. 2021) also applied for Parkes. We take a more extreme value, and artificially increase the Parkes/Mb detection threshold  $F_0$  by 60 per cent to 0.8 Jy ms. Secondly, we assume all Parkes measurements below a SNR of 16 to be untrustworthy, as suggested by Macquart & Ekers (2018a) and James et al. (2019b), and discussed in Section 8.7. As well as also raising the detection threshold to 0.8 Jy ms, this removes seven FRBs – six during the normalisable observation time – from the analysis. Interestingly, removing six of twelve normalisable FRBs is exactly what is predicted from unbiased observations of a Euclidean source-counts distribution, where  $N_{\text{FRB}} \propto \text{SNR}_{\text{th}}^{-1.5}$ , since  $(16/10)^{-1.5} = 0.49$ . This suggests that any observational bias was present only in the less systematic FRB searches with Parkes. Re-running the fit optimization under the ‘rate’ assumption on  $\alpha$  (because numerical evaluation is quicker), the best-fitting values of the fit parameters are shown in Table A1.

Increasing the Parkes threshold to 0.8 Jy ms has almost negligible impact on the parameter estimates. We attribute this to the change





**Figure A1.** Resulting 68, 90, and 95 per cent confidence limits (respectively: blue, upper; green, middle; red, lower) on individual parameters under the spectral index interpretation of  $\alpha$ , using in descending order: all available information, removing information on the number of detected FRBs, their DMs and redshifts, their measured signal-to-noise ratios, and data from ASKAP/FE, ASKAP/ICS, and Parkes/MB, respectively. Top: using a prior on  $\alpha$ ; bottom: no prior on  $\alpha$ .

in threshold being small compared to the difference in sensitivity between Parkes and ASKAP observations. This highlights the utility of fitting data from telescopes with very different thresholds. Setting  $\text{SNR}_{\text{th}} = 16$  however has a greater impact: a flatter luminosity function is preferred, as is a reduced  $n$ . With a uniform prior on  $\alpha$ , the source evolution ( $n \sim 0$ ) is preferred, although with a Gaussian prior on  $\alpha$ , evolution consistent with the star formation rate ( $n \sim 1$ ) is favoured.

In general, varying these two systematics produces effects of smaller size than the 68 per cent parameter confidence intervals. Except in the case that the Parkes suffered from missing FRBs below  $\text{SNR}_{\text{th}} = 16$ , and the measurements of Macquart et al. (2019) are somehow flawed, would the our main conclusion on the evolving nature of the FRB population be affected.

Note that confidence intervals will be correlated between columns in Table A1, and are shown to evaluate the importance of these systematics against random variation.

**Table A1.** Comparison of best-fitting parameter values to those calculated assuming an increase in the Parkes FRB detection threshold,  $F_{\text{th}}^{\text{Pks}}$ , to 0.8 Jy ms; and when also excluding Parkes FRBs with  $\text{SNR} < 16$  from the analysis. Shown are results calculated under the ‘rate’ approximation to  $\alpha$ , both with and without a prior on  $\alpha = -1.5 \pm 0.3$ . Errors correspond to 68 per cent confidence intervals calculated using Wilks’ theorem.

Parameter	Best fit	Uniform prior on $\alpha$		Best fit	Gaussian prior on $\alpha$	
		$F_{\text{th}}^{\text{Pks}} = 0.8 \text{ Jy ms}$	$\text{SNR}_{\text{th}}^{\text{Pks}} = 16$		$F_{\text{th}}^{\text{Pks}} = 0.8 \text{ Jy ms}$	$\text{SNR}_{\text{th}}^{\text{Pks}} = 16$
$\log_{10} E_{\text{max}}$	$42.19^{+0.14}_{-0.24}$	$41.44^{+0.28}_{-0.06}$	$41.40^{+0.08}_{-0.02}$	$41.70^{+0.53}_{-0.06}$	$41.40^{+0.30}_{-0.02}$	$41.40^{+0.14}_{-0.02}$
$\gamma$	$-1.09^{+0.11}_{-0.08}$	$-1.11^{+0.14}_{-0.17}$	$-0.92^{+0.16}_{-0.16}$	$-1.09^{+0.14}_{-0.10}$	$-1.11^{+0.13}_{-0.14}$	$-0.95^{+0.20}_{-0.14}$
$n$	$2.27^{+0.25}_{-0.45}$	$2.00^{+0.36}_{-1.64}$	$0.06^{+1.52}_{-0.30}$	$1.67^{+0.25}_{-0.40}$	$1.58^{+0.36}_{-0.30}$	$1.15^{+0.48}_{-0.36}$
$\mu_{\text{host}}$	$2.07^{+0.18}_{-0.20}$	$2.20^{+0.16}_{-0.18}$	$2.23^{+0.11}_{-0.16}$	$2.11^{+0.18}_{-0.20}$	$2.20^{+0.16}_{-0.16}$	$2.20^{+0.14}_{-0.18}$
$\sigma_{\text{host}}$	$0.52^{+0.17}_{-0.10}$	$0.48^{+0.14}_{-0.09}$	$0.46^{+0.17}_{-0.07}$	$0.53^{+0.15}_{-0.11}$	$0.48^{+0.14}_{-0.09}$	$0.51^{+0.15}_{-0.11}$

In concluding this section, we remind readers that the  $F_{\text{th}}^{\text{Pks}} = 0.8 \text{ Jy ms}$  case is larger than we consider plausible, while the  $\text{SNR}_{\text{th}}^{\text{Pks}} = 16$  case discards a significant fraction of data from the most sensitive FRB sample; thus we expect the true systematic errors in our analysis to be less than the variation evident in Table A1. None the less, since  $p_n$  is an important constraint on FRB parameters, improved modelling may be needed of the Parkes experiment to avoid uncertainties limiting the accuracy of future, otherwise more-precise calculations.

#### A6 Validity of statistical assumptions

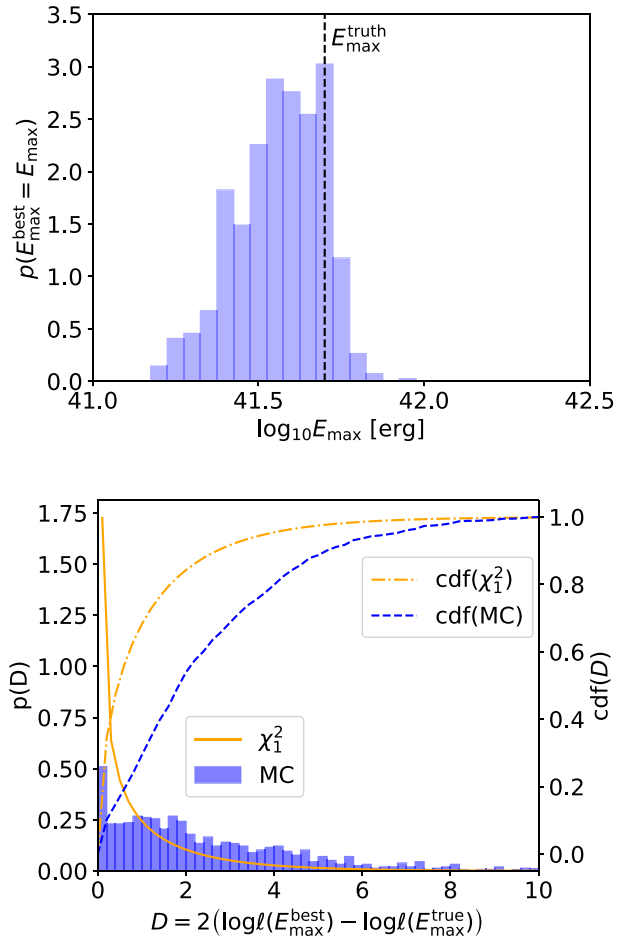
In this work, We have by default used a frequentist approach to setting confidence intervals. This is more typical of the particle physics community, where the validity of models are evaluated using maximum likelihood methods. Here, we have taken this approach, and invoked Wilks’ theorem in setting confidence intervals.

A more robust approach would be to use Monte Carlo methods to set these intervals. This would involve, for each combination of parameters:

- (i) For each parameter set  $p_{\text{set}}$ , generating ‘pseudo-experiments’ by sampling FRBs from the simulated distributions in  $z$ , DM,  $s$  space for each of the ASKAP/FE, ASKAP/ICS, and Parkes/Mb experiments;
- (ii) evaluating the likelihood of each FRB sample using the parameter set it was generated with, and the parameter set  $p_{\text{max}}$  maximizing the likelihood for the actual, observed FRBs;
- (iii) include  $p_{\text{set}}$  in a  $C$  per cent confidence interval if at least  $C$  per cent of all Monte Carlo FRB samples from  $p_{\text{set}}$  had a likelihood difference between  $p_{\text{max}}$  and  $p_{\text{set}}$  less than or equal to the observed value.

Such a procedure would, however, be too computationally complex for this multidimensional space. Rather, we test the validity of Wilks’ theorem for this particular data sample in one dimension, by varying  $E_{\text{max}}$  only. We choose  $E_{\text{max}}$  because the sharp drop in the likelihood distribution below the observed energy of localized FRBs makes this the least well-behaved parameter. To further simplify the problem, we only generate pseudo-experiments at one simulated truth, being the set  $p_{\text{max}}$  reported for the best-fitting with a Gaussian prior on  $\alpha$  given in Table A1 (and in our companion paper). We do not vary the number of observed FRBs  $N_{\text{FRB}}$ , and simply re-generate the properties  $z$ , DM, and  $s$  for each FRB. We then evaluate these while varying  $E_{\text{max}}$ , and measure the test statistic

$$D = 2 \log \left( \ell(E_{\text{max}}^{\text{best}}) / \ell(E_{\text{max}}^{\text{true}}) \right). \quad (\text{A3})$$



**Figure A2.** Upper: histogram of the values of  $E_{\text{max}}$  which maximise the likelihood,  $E_{\text{max}}^{\text{best}}$ , compared to the simulated true  $E_{\text{max}}^{\text{true}}$ , when varying  $E_{\text{max}}$  only, and using a Monte Carlo sample of FRBs as per Appendix A6. Lower: differential and cumulative distributions of the test statistic  $D$  compared to the predicted  $\chi^2_1$  distribution from Wilks’ theorem.

The resulting distribution of best-fitting  $E_{\text{max}}$  values,  $E_{\text{max}}^{\text{best}}$ , and the distribution of  $D$ , are given in Fig. A2, upper and lower respectively, using  $10^4$  pseudo-experiments.

From Fig. A2(upper),  $E_{\text{max}}^{\text{best}}$  is clearly biased towards low values. This is expected, since the fit naturally prefers values of  $E_{\text{max}}$  only marginally above the highest observed FRB energy. The MC never generates FRBs with energy above  $E_{\text{max}}^{\text{true}}$ , but can generate samples with FRBs with lower energy. Thus, the distribution of  $E_{\text{max}}^{\text{best}}$  is

**Table A2.** Ranges for the uniform priors on parameters used in the calculation of Bayesian confidence intervals. In all cases other than  $\alpha$ , the likelihood is negligible outside the interval.

Parameter	Min	Max
$\log_{10} E_{\max}$	41	43.4
$\alpha$	-2.50	1.0
$\gamma$	-0.5	-1.5
$n$	-2	5
$\mu_{\text{host}}$	0.5	2.75
$\sigma_{\text{host}}$	0.2	1.1

similarly shaped to the likelihood function  $\ell(E_{\max})$  when evaluated on actual data, but skewed in the opposite direction.

According to Wilks' theorem,  $D$  should follow a  $\chi^2_1$  distribution. It is evident from Fig. A2 (right) that while the shapes are similar, the true distribution of  $D$  is significantly skewed to the right compared to the expectation. The cumulative distributions illustrate that the confidence intervals calculated using Wilks' theorem may therefore suffer from undercoverage.

We note that we have tested Wilks' theorem for the most extreme case. Our confidence intervals, when marginalized over other parameters, should be more trustworthy than indicated by Fig. A2. None the less, we proceed to calculate Bayesian confidence intervals to assess the sensitivity of our conclusions to the statistical method used.

## A7 Bayesian confidence intervals

A Bayesian posterior probability distribution on a single parameter  $\theta_i$  in a parameter set  $\vec{\theta}$  can be calculated by integrating the likelihood function  $\ell$ , weighted by appropriate priors, over all dimensions except  $i$ :

$$p(\theta_i) = \int \ell(\vec{\theta}) p(\vec{\theta}) d\vec{\theta}_{\neq i}. \quad (\text{A4})$$

We choose as our priors uniform distributions in  $\mu_{\text{host}}$  and  $\sigma_{\text{host}}$  (which are already defined in log-space),  $\gamma$ , and  $n$ , and log-uniform priors in  $E_{\max}$ , in the ranges shown in Table A2. As previously, we take both uniform and Gaussian priors on  $\alpha$ .

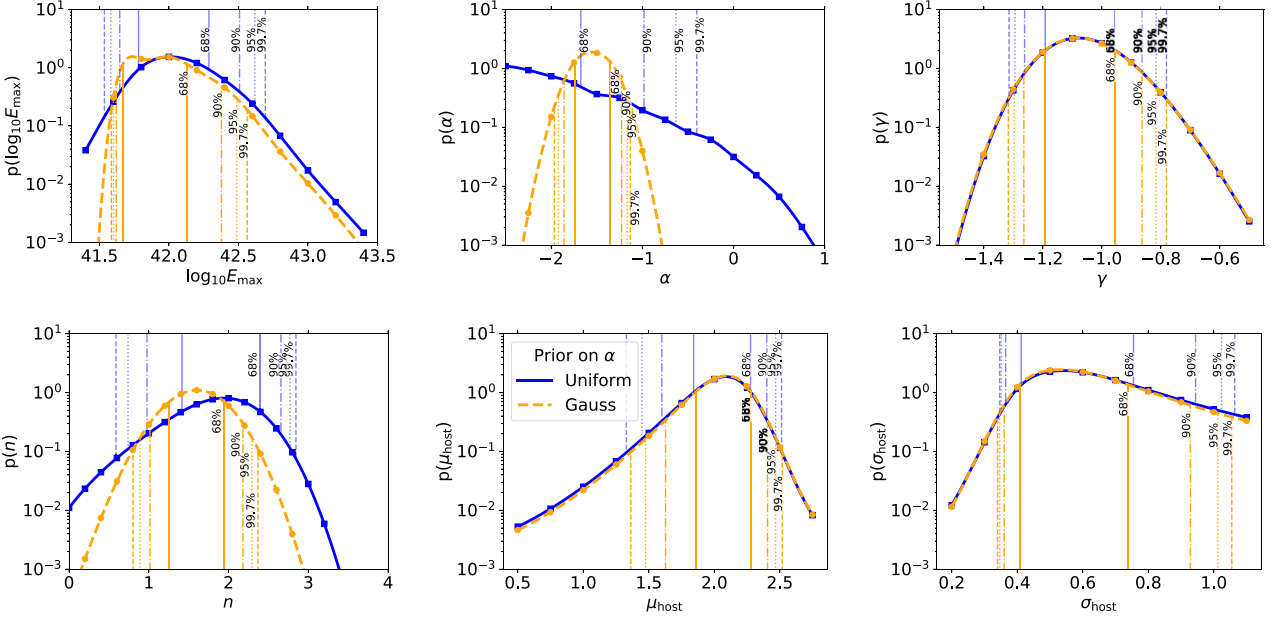
The result using the spectral (rate) interpretation of  $\alpha$  is plotted in Fig. A3 (Fig. A4). Confidence intervals are constructed by including regions of parameter space with the greatest  $p(\theta_i)$  until the desired level of confidence is reached. These intervals are also listed in Table A2. We observe that Bayesian confidence intervals are generally more constraining than those calculated using Wilks' theorem in the frequentist approach. Thus, while it is possible that the latter intervals do suffer from undercoverage, they are none the less more conservative than the intervals which would be generated using the more standard Bayesian approach, as used by e.g. Luo et al. (2020).

**Table A3.** Bayesian confidence limits on single parameters, with both a uniform prior and Gaussian prior on  $\alpha$ , calculated assuming the spectral interpretation of  $\alpha$  (see Section 3). The best-fitting value is quoted such that the posterior probability is maximized.

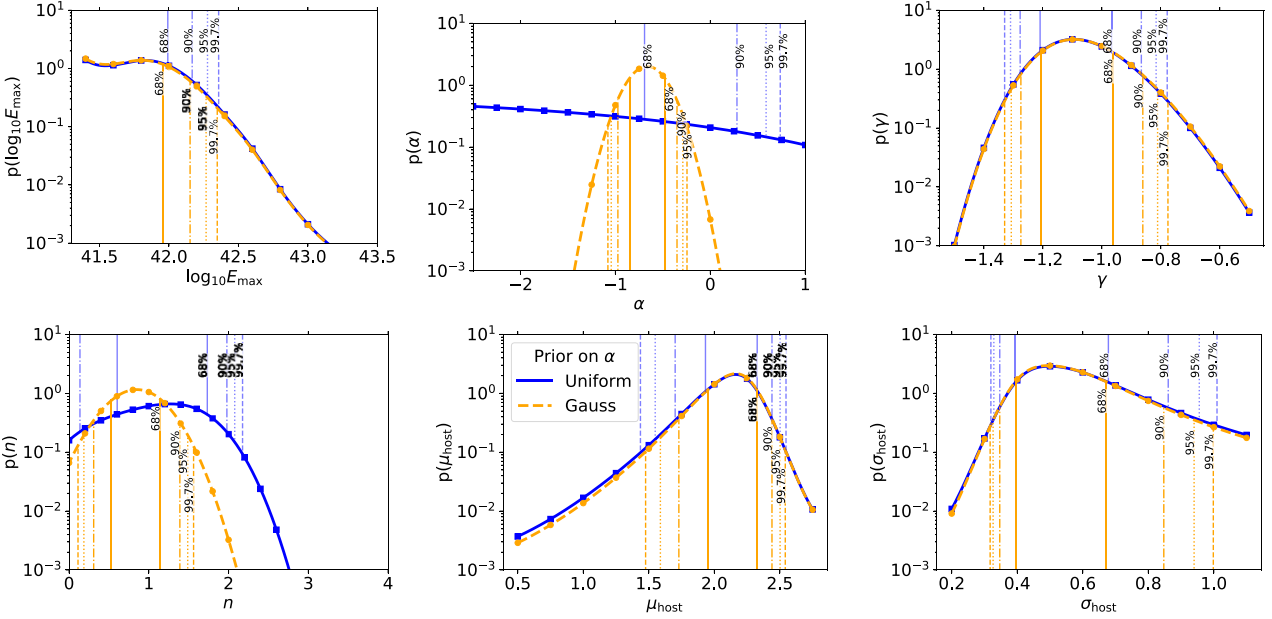
Parameter	Best fit	Uniform prior on $\alpha$				Best fit	Gaussian prior on $\alpha$			
		68 per cent C.L.	90 per cent C.L.	95 per cent C.L.	99.7 per cent C.L.		68 per cent C.L.	90 per cent C.L.	95 per cent C.L.	99.7 per cent C.L.
$\log_{10} E_{\max}$	42.01	+0.28 -0.23	+0.50 -0.36	+0.61 -0.43	+0.69 -0.47	41.74	+0.40 -0.07	+0.64 -0.12	+0.75 -0.14	+0.83 -0.15
$\gamma$	-1.08	+0.12 -0.11	+0.22 -0.18	+0.26 -0.22	+0.30 -0.24	-1.08	+0.13 -0.11	+0.22 -0.18	+0.27 -0.22	+0.30 -0.24
$n$	1.94	+0.45 -0.53	+0.71 -0.96	+0.83 -1.20	+0.90 -1.35	1.60	+0.34 -0.35	+0.58 -0.59	+0.69 -0.71	+0.76 -0.80
$\mu_{\text{host}}$	2.08	+0.19 -0.24	+0.32 -0.48	+0.38 -0.64	+0.43 -0.76	2.10	+0.18 -0.24	+0.31 -0.47	+0.37 -0.62	+0.42 -0.73
$\sigma_{\text{host}}$	0.54	+0.21 -0.13	+0.40 -0.18	+0.48 -0.19	+0.52 -0.19	0.53	+0.20 -0.13	+0.39 -0.17	+0.48 -0.19	+0.52 -0.19

**Table A4.** As per Table A3, assuming the rate interpretation of  $\alpha$  (see Section 3). A lower limit on  $\log_{10} E_{\max}$  [erg] of 41.2 has been assumed.

Parameter	Best fit	Uniform prior on $\alpha$				Best fit	Gaussian prior on $\alpha$			
		68 per cent C.L.	90 per cent C.L.	95 per cent C.L.	99.7 per cent C.L.		68 per cent C.L.	90 per cent C.L.	95 per cent C.L.	99.7 per cent C.L.
$\log_{10} E_{\max}$	41.40	+0.59 -0.20	+0.77 -0.20	+0.88 -0.20	+0.96 -0.20	41.40	+0.56 -0.20	+0.75 -0.20	+0.87 -0.20	+0.95 -0.20
$\gamma$	-1.09	+0.13 -0.12	+0.23 -0.18	+0.28 -0.22	+0.32 -0.24	-1.09	+0.13 -0.11	+0.23 -0.18	+0.28 -0.21	+0.32 -0.24
$n$	1.27	+0.47 -0.66	+0.71 -1.13	+0.81 -1.33	+0.91 -1.42	0.85	+0.29 -0.32	+0.54 -0.54	+0.64 -0.66	+0.71 -0.74
$\mu_{\text{host}}$	2.16	+0.16 -0.23	+0.28 -0.46	+0.34 -0.61	+0.38 -0.73	2.16	+0.16 -0.21	+0.28 -0.43	+0.34 -0.58	+0.38 -0.69
$\sigma_{\text{host}}$	0.50	+0.18 -0.10	+0.36 -0.15	+0.46 -0.17	+0.51 -0.18	0.49	+0.18 -0.10	+0.35 -0.15	+0.45 -0.17	+0.51 -0.18



**Figure A3.** Bayesian confidence intervals on parameters  $E_{\max}$ ,  $\alpha$ ,  $\gamma$ ,  $n$ ,  $\mu_{\text{host}}$ ,  $\sigma_{\text{host}}$  when marginalized over the other five, under the spectral index interpretation of  $\alpha$ , using a Gaussian (orange, lower) and uniform (blue, upper) prior on the spectral index  $\alpha$ . Calculation results are given by points, with lines drawn using cubic spline smoothing. Vertical lines are single-parameter intervals at the labelled degree of confidence calculated.



**Figure A4.** As per Fig. A3, except calculated using the rate interpretation of  $\alpha$ .

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.