Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **483,** 1342–1353 (2019)
Advance Access publication 2018 November 10

doi:10.1093/mnras/sty3031

The slope of the source-count distribution for fast radio bursts

C. W. James [1,2]★ R. D. Ekers,[1,3] J.-P. Macquart [1,2] K. W. Bannister[3] and R. M. Shannon[4]

[1]*International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia*
[2]*ARC Centre of Excellence for All-Sky Astrophysics (CAASTRO), Australia*
[3]*Australia Telescope National Facility, CSIRO Astronomy and Space Science, PO Box 76, Epping, NSW 1710, Australia*
[4]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia*

Accepted 2018 November 5. Received 2018 October 8; in original form 2018 August 20

ABSTRACT

The slope of the source-count distribution of fast radio burst (FRB) fluences, α, has been estimated using a variety of methods. Hampering all attempts have been the low number of detected FRBs, and the difficulty of defining a completeness threshold for FRB surveys. In this work, we extend maximum likelihood methods for estimating α, using detected and threshold signal-to-noise ratios applied to all FRBs in a sample without regard to a completeness threshold. Using this method with FRBs detected by the Parkes radio telescope, we find $\alpha = -1.18 \pm 0.24$ (68 per cent confidence interval, CI), i.e. consistent with a non-evolving Euclidean distribution ($\alpha = -1.5$). Applying these methods to the Australian Square Kilometre Array Pathfinder (ASKAP) Commensal Real-time ASKAP Fast Transients (CRAFT) FRB survey finds $\alpha = -2.2 \pm 0.47$ (68 per cent CI). A full maximum likelihood estimate finds an inconsistency with the Parkes rate with a p-value of 0.86 per cent (2.6σ). If not due to statistical fluctuations or biases in Parkes data, this is the first evidence for deviations from a pure power law in the integral source-count distribution of FRBs. It is consistent with a steepening of the integral source-count distribution in the fluence range 5–40 Jy ms, for instance due to a cosmological population of FRB progenitors evolving more rapidly than the star formation rate, and peaking in the redshift range 1–3.

Key words: methods: data analysis – radio continuum: general – radio continuum: transients.

1 INTRODUCTION

Fast radio bursts (FRBs) are one of the most poorly understood astrophysical phenomena. FRBs are radio pulses with measured durations of milliseconds, arriving with high dispersion measures inconsistent with a Galactic origin. This makes their intrinsic luminosity extreme, with a plethora of models proposed to explain their progenitors.

The nature of FRB progenitors is so poorly constrained because, with the exception of the only known repeating source (FRB 121102; Spitler et al. 2016), FRBs arrive from unpredictable directions at unpredictable times. Despite estimates for the all-sky daily rate being in the hundreds to thousands, the total number of detected FRBs is only approximately 50 (Petroff et al. 2016, www.frbcat.org; Shannon et al. 2018).

Efforts to understand the distribution of FRB fluences tend to assume a power-law distribution for the all-sky rate R of FRBs above a given fluence threshold F of the form

$$R(F) = R_0 \left(\frac{F}{F_0}\right)^{\alpha} \text{ sky}^{-1}\,\text{d}^{-1}, \quad (1)$$

where R_0 is the rate at fluence threshold F_0. In the case that FRBs originate from the local Universe – as suggested by the proximity of FRB 121102 (Chatterjee et al. 2017; Tendulkar et al. 2017) – a value of $\alpha = -1.5$ is expected.

Shannon et al. (2018) have recently shown a dispersion–brightness relation for the FRB population probed by the Parkes and Australian Square Kilometre Array Pathfinder (ASKAP) telescopes, suggesting this population originates at distances up to redshifts of 2–3. For cosmological populations of radio sources, source evolution and k-corrections interact with the intrinsic source luminosity distribution to produce unique features in the measured source-count distribution (see e.g. Wall (1996)), and the same is therefore expected of the observed FRB fluence distribution (Macquart & Ekers 2018b). FRB data have been too sparse however to test more complicated models. The approach taken here is to search

★ E-mail: clancy.james@curtin.edu.au

for the first hints of such structure by finding deviations from the simple power-law model for the observed fluence distribution.

Estimates of the parameters of the simple power-law model vary greatly. Champion et al. (2016) find $R_0 = 6^{+4}_{-3} \times 10^3$ FRBs sky^{-1} d^{-1} (95 per cent confidence interval, CI) based on the nine FRBs detected by the High Time Resolution Universe (HTRU) survey at Parkes above the fluence duration range from 0.13 Jy ms at 0.128 ms to 5.9 Jy ms over 262 ms. However, this estimate did not include the non-detection of the HTRU mid-latitude survey (Petroff et al. 2014). Bhandari et al. (2018) use a larger sample of Parkes FRBs, finding a rate of $1.7^{+1.5}_{-0.9} \times 10^3$ FRBs sky^{-1} d^{-1} above a 'fluence-complete' limit of 2 Jy ms. Oppermann, Connor & Pen (2016) estimate both the FRB rate and α using 15 FRBs from seven surveys, finding $-0.8 \geq \alpha \geq -1.7$ (95 per cent CI), with rates above 10^5 sky^{-1} d^{-1}. Importantly, Oppermann et al. (2016) use the ratio of observed to threshold signal-to-noise ratio (S/N) values in their likelihood estimate, a parameter to which we shall return in Section 2.

These estimates effectively calculate R by dividing the total number of bursts by the total observation time and sensitive solid angle, and hence fluctuate due to the low number of FRBs observed. In addition, as Macquart & Ekers (2018a) have noted, the effective detection threshold $F_{\rm eff}$ and solid angle $\Omega_{\rm eff}$ of an FRB search depends on α, with small negative values of α (e.g. -1) favouring strong bursts detected far from beam centre, and large negative values (e.g. -3) favouring bursts detected near threshold close to beam centre. Without accounting for these effects, estimates of R will be highly biased.

James et al. (2018) account for the interaction between R and α in using 19 FRBs detected in the ASKAP Commensal Real-time ASKAP Fast Transients (CRAFT) survey (Shannon et al. 2018), finding R to be between 6.9 ($\alpha = -1.1$) and 23 FRBs sky^{-1} d^{-1} ($\alpha = -3.0$), above effective thresholds of 92 and 36 Jy ms, respectively, at the survey time resolution of 1.2656 ms.

Lawrence et al. (2017) perform a maximum likelihood fit of equation (1) to FRBs detected by several telescopes, accounting for the effect of idealized beam shapes. They find $R = 587^{+336}_{-315}$ above 1 Jy ms for $\alpha = -0.91 \pm 0.34$ (95 per cent CI). Vedantham et al. (2016) also use maximum likelihoods, estimating α using two methods. The observed fraction of multiple- to single-beam detections of Parkes yields $-0.52 > \alpha > -1.0$ (90 per cent CI), and combining both detections and non-detections of several other instruments finds $-0.32 > \alpha > -0.92$, with a combined constraint of $-0.5 > -\alpha > -0.9$.

The constraints set by the maximum likelihood methods of both Vedantham et al. (2016) and Lawrence et al. (2017), which come from comparing multitelescope data, are sensitive to unmodelled differences in system response to the (unknown) dark matter (DM) distribution, time duration, and frequency dependence of FRBs. Neither include the effects of radio-frequency interference (RFI) on detection efficiency, and use analytic approximations to beam patterns, deviations from which have been shown by Macquart & Ekers (2018a) to greatly affect implied FRB properties. Furthermore, both analyses include the Lorimer burst, which should be excluded from statistical calculations on the grounds of discovery bias (Macquart & Ekers 2018a).

In summary, while the methods of both Vedantham et al. (2016) and Lawrence et al. (2017) are analytically sound, their quantitative results should be subject to revision.

A more robust method to estimate α is to apply the maximum likelihood method of Crawford, Jauncey & Murdoch (1970) to FRB samples from a single telescope. Both Bhandari et al. (2018) and Macquart & Ekers (2018a) do so to different samples of FRBs detected by Parkes, including data from the SUrvey for Pulsars and Extragalactic Radio Bursts (SUPERB), estimating $\alpha = -2.6^{+0.7}_{-1.3}$ and $-2.2^{+0.6}_{-1.2}$, respectively. It is therefore an outstanding question as to whether or not the differences in the values of α obtained with different methods are due to the pitfalls of making comparisons between different telescopes, a true change in the FRB spectral slope at different fluences, a statistical fluctuation, or some other effect.

This paper concerns estimates of α and the maximum likelihood method of Crawford et al. (1970) – estimates of the absolute rate R will be left to a future work. In Section 2, we extend the maximum likelihood method of Crawford et al. (1970), formulating it in terms of measured and threshold S/Ns as per Oppermann et al. (2016). Our formulation avoids all of the aforementioned uncertainties in beam shape and FRB properties, and renders the concept of a 'completeness fluence' irrelevant. We outline the correct method for applying this updated maximum likelihood test to FRB data in Section 3. The new approach is both simpler, and allows the use of a greater proportion of detected FRBs than previous applications. This gives it greater statistical power and less bias. This method is then applied independently to both Parkes and ASKAP FRBs in Section 4, which have sufficient detections to provide independent meaningful estimates of α. A full maximum likelihood optimization, accounting for errors, is performed in Section 5, which allows the likelihood ratios to test whether or not the same power-law distribution can account for both samples. The discussion of Section 6 compares the resulting estimates of α, both to each other and to those of previous authors, and interprets the result in terms of the cosmological source evolution of FRB progenitors.

2 GENERALIZATION OF THE METHOD OF CRAWFORD ET AL.

The application of maximum likelihood methods to source-counting statistics is often referenced to Crawford et al. (1970), who discuss the problem in terms of a population of $N(S)$ sources with flux densities greater than some value S. It is assumed that $N(S)$ has the form of a power law:

$$N(S) = kS^{\alpha}, \tag{2}$$

where k is a constant, and α the index. Note that Crawford et al. (1970) use $-\alpha$ for the index, and a for the estimator of α. Using here α^* as the maximum likelihood estimator of α, α^* is then given by

$$\frac{1}{\alpha^*} = -\frac{1}{M}\sum_i \ln s_i' \tag{3}$$

for an observation of M sources with flux densities S_i, where $s_i = S_i/S_0$ for a detection threshold S_0. The standard deviation of the result for large M, for which $\alpha^* \sim \alpha$, is given by

$$\sigma_\alpha(\alpha) \sim \sigma_\alpha(\alpha^*) = \frac{M\alpha^*}{(M-1)(M-2)^{0.5}}. \tag{4}$$

The distribution of α^* given α, $p(\alpha^*|\alpha)$, follows a gamma distribution:

$$p(\alpha^*|\alpha) = \frac{(-\alpha)^M}{M!}\left(\frac{M}{-\alpha^*}\right)^{M+1} {\rm e}^{-\alpha M/\alpha^*}. \tag{5}$$

Crawford et al. (1970) also note that while the estimate $(\alpha^*)^{-1}$ of α^{-1} obtained from equation (3) is unbiased, the estimate of α, $\alpha*$, is biased. An unbiased estimate, α', can be found using

$$\alpha' = \frac{M-1}{M}\alpha^*. \tag{6}$$

Error estimates for α' will also be modified by the same factor.

Equations (2) and (3) specify both the source distribution and the detection threshold S_0, in terms of a single variable, S. However, the flux density threshold S_0, or equivalently a fluence threshold F_0, of a transient source will vary with event duration. In the case of FRBs, both dispersion measure and source position in the beam are often discussed complicating factors. In the following section, we first demonstrate that equation (3) is applicable to an arbitrarily complicated distribution when s is simply the observed S/N relative to a threshold, $\mathrm{S/N_{th}}$. We also formulate the problem from hereon in terms of fluence F, which is the standard used by the FRB community.

2.1 Maximum likelihood methods in a multidimensional observation space

Consider a set of parameters $\boldsymbol{\theta}$ covering both intrinsic source properties (e.g. dispersion measure) and observational effects (e.g. beam shape). Defining the relative source distribution of events within the space $\boldsymbol{\theta}$ as $k(\boldsymbol{\theta})$, and total K such that

$$K = \int \mathrm{d}\boldsymbol{\theta}\, k(\boldsymbol{\theta}), \tag{7}$$

we assume that k is independent of fluence threshold F, so that the integral source-count distribution still has the form

$$N(F) = K\left(\frac{F}{F_0}\right)^{\alpha}. \tag{8}$$

Note that we now normalize F relative to some threshold F_0 (cf. equation 2). While the use of number counts N, rather than a rate R, is arbitrary, we do so as a reminder that $\boldsymbol{\theta}$ can include time-dependent factors.

For this formulation, the probability of an event $p(\boldsymbol{\theta}, F)$ occurring in the range $\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}F$ given a single observation is

$$\frac{\mathrm{d}p(F,\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}F} = -\frac{k(\boldsymbol{\theta})}{C}\frac{\alpha}{F_0}\left(\frac{F}{F_0}\right)^{\alpha-1}. \tag{9}$$

The normalizing constant C is required to ensure that the probability density, when integrated over the entire parameter space, equates to unity. Considering events between a threshold F_{th} and maximum F_{m}, where (crucially for this method) both depend on $\boldsymbol{\theta}$ in an arbitrary way to account for possible experimental sensitivities, C, becomes

$$\begin{aligned} C &= \int \mathrm{d}\boldsymbol{\theta} \int_{F_{\mathrm{th}}(\boldsymbol{\theta})}^{F_{\mathrm{m}}(\boldsymbol{\theta})} \frac{-\mathrm{d}N(F,\boldsymbol{\theta})}{\mathrm{d}F}\mathrm{d}F \\ &= \int \mathrm{d}\boldsymbol{\theta}\, k(\boldsymbol{\theta})\left[\left(\frac{F_{\mathrm{th}}(\boldsymbol{\theta})}{F_0}\right)^{\alpha} - \left(\frac{F_{\mathrm{m}}(\boldsymbol{\theta})}{F_0}\right)^{\alpha}\right]. \end{aligned} \tag{10}$$

Defining relative fluence threshold s and maximum b similarly to Crawford et al. (1970):

$$s \equiv \frac{F}{F_{\mathrm{th}}(\boldsymbol{\theta})}, \tag{11}$$

$$b \equiv \frac{F_{\mathrm{m}}}{F_{\mathrm{th}}(\boldsymbol{\theta})}, \tag{12}$$

equations (9) and 10 can be written in terms of s, using $\mathrm{d}F = F_{\mathrm{th}}\mathrm{d}s$:

$$\frac{\mathrm{d}p(s,\boldsymbol{\theta})}{\mathrm{d}\boldsymbol{\theta}\,\mathrm{d}s} = -\frac{k(\boldsymbol{\theta})}{C'}\alpha s^{\alpha-1}F_{\mathrm{th}}^{\alpha}(\boldsymbol{\theta}), \tag{13}$$

$$C' = \int \mathrm{d}\boldsymbol{\theta}\, k(\boldsymbol{\theta})F_{\mathrm{th}}^{\alpha}(\boldsymbol{\theta})\left[1 - b^{\alpha}(\boldsymbol{\theta})\right], \tag{14}$$

where a factor of F_0^{α} has disappeared into the new normalization constant, C'. Practically, this means that an experiment observing N events need not be concerned about the absolute scale F_0 of the power-law distribution when calculating the relative probability of observing s – however, the absolute scale is critically important when estimating the absolute rate R, necessary to compare different experiments. A simple example would be two experiments with thresholds F_0 of 1 and 10 Jy ms, where the latter would measure a rate R of 10^{α} less than the former.

In the case where the relative maximum fluence, b, is independent of $\boldsymbol{\theta}$, the total probability of observing s given a single observation becomes

$$\begin{aligned} p(s) &= \int p(s,\boldsymbol{\theta})\mathrm{d}\boldsymbol{\theta} \\ &= -\alpha\frac{s^{\alpha-1}}{1-b^{\alpha}}\frac{1}{C''}\mathrm{d}s\int \mathrm{d}\boldsymbol{\theta}\, k_{\mathrm{obs}}(\boldsymbol{\theta})F_{\mathrm{th}}^{\alpha}(\boldsymbol{\theta}), \\ C'' &= \int \mathrm{d}\boldsymbol{\theta}\, k_{\mathrm{obs}}(\boldsymbol{\theta})F_{\mathrm{th}}^{\alpha}(\boldsymbol{\theta}). \end{aligned} \tag{15}$$

Noting that the normalization constant C'' simply cancels out the integral, $p(s)$ reduces to

$$p(s) = -\alpha\frac{s^{\alpha-1}}{1-b^{\alpha}}\mathrm{d}s. \tag{16}$$

This is exactly the formula derived by Crawford et al. (1970) in their equation (6). All results presented in that work then apply, specifically that in the case where $b = \infty$, α can be calculated as per equation (3). To repeat their analysis, maximizing the log likelihood $\mathcal{L}$ in terms of estimated probabilities p_i for each event i gives

$$\begin{aligned} \mathcal{L} &= \log\prod_{i=1}^{N} p_i(s) \\ &= \sum_{i=1}^{N}\left[\log(-\alpha) + (\alpha-1)\log s_i - \log(1-b^{\alpha})\right]. \end{aligned} \tag{17}$$

Differentiating with respect to α produces

$$\frac{\partial\mathcal{L}}{\partial\alpha} = \sum_{i=1}^{N}\left[\alpha^{-1} + \log s_i + \frac{b^{\alpha}\log b}{1-b^{\alpha}}\right]. \tag{18}$$

Setting this to zero to find the maximum with respect to α defines the estimator α^*. Using $b = \infty$ produces

$$\begin{aligned} \frac{N}{\alpha^*} + \sum_{i=1}^{N}\log s_i &= 0 \\ \Rightarrow \frac{1}{\alpha^*} &= \frac{-1}{N}\sum_{i=1}^{N}\log s_i. \end{aligned} \tag{19}$$

In summary, if one can calculate the ratio s between signal strength and detection threshold $F_{\mathrm{th}}(\boldsymbol{\theta})$ for each event, one can estimate the slope of the cumulative source-count distribution α *without* knowing the complex dependencies of that threshold on event parameters $\boldsymbol{\theta}$, or the distribution of true source events with that space $k(\boldsymbol{\theta})$ – or even what the source space $\boldsymbol{\theta}$ is.

3 APPLICATION TO FRB SEARCHES

FRB searches currently suffer from several complicating effects. Their (generally) once-off nature means that most FRBs are poorly localized, and hence their true fluence F is observed at reduced sensitivity $F_{\rm obs} = BF$, where B is the value of the antenna beam pattern at which they are observed. Their variable duration – itself a function of dispersion measure, scattering, and intrinsic pulse width – means that longer bursts require a higher fluence to be detected. RFI can result in a time-dependent threshold, while different surveys have vastly different sensitivities. The analysis algorithms used by each experiment have their own complicated response patterns. The particular shape of each FRB pulse in both time and frequency – and how this is measured at the experimental time/frequency resolution – also affects the detection probability (Keane & Petroff 2015). While the source distribution $k(\boldsymbol{\theta})$ will obviously be independent of some of these parameters (e.g. local RFI, position in beam), all may affect the fluence sensitivity threshold $F_{\rm th}(\boldsymbol{\theta})$.

It is important to emphasize that all these parameters fall within the definition of the parameter space $\boldsymbol{\theta}$ introduced in an abstract sense in Section 2. Indeed, $\boldsymbol{\theta}$ can extend to include different experiments in the one analysis.

Correctly incorporating these experimental effects into an analysis will result in a powerful probe of the FRB fluence distribution. However, they make it very difficult to reconstruct the true value of F for any given event. Furthermore, the intrinsic FRB distribution is related to the observed properties through equation (10), where in this example, the 'normalizing constant' C is the total number of detected events. Estimating the FRB rate therefore requires integrating over the parameter space $\boldsymbol{\theta}$ of experimental dependencies. This is very difficult, since each telescope's $F_{\rm th}(\boldsymbol{\theta})$ is often not well understood, and the DM, duration, frequency, and signal-shape dependence of $k(\boldsymbol{\theta})$ certainly is not. The result of Section 2.1 is so useful because this complicated integration is cancelled when using s in maximum likelihood estimates of α (the cancellation occurs at equation 15). Furthermore, for FRB searches, the relative fluence s is generally very easy to calculate.

FRB searches regularly report the S/N of events passing a predetermined threshold, $\mathrm{S/N_{th}}$. Calculating both parameters is readily achieved by assuming Gaussian noise statistics for received power in the time–frequency domain, and simply tracking the number of samples added in any given search. In such a case, the parameter s is simply the ratio

$$s = \frac{\mathrm{S/N}}{\mathrm{S/N_{th}}}, \tag{20}$$

i.e. it is the detected significance relative to threshold significance. Combining equations (19) and (20) gives

$$\frac{1}{\alpha^*} = \frac{-1}{N}\sum_{i=1}^{N} \log\left(\frac{\mathrm{S/N}}{\mathrm{S/N_{th}}}\right)_i . \tag{21}$$

It is worthwhile discussing the assumptions in the derivation of Section 2.1 in the context of FRB searches. Certainly, it is possible that the true integral source-counts distribution, $N(F, \boldsymbol{\theta})$, is not separable between 'complicating parameters' $\boldsymbol{\theta}$ and the fluence distribution $(F/F_0)^{\alpha}$. An example would be observationally brighter FRBs having lower DMs. However, should this be the case, the integral source-counts spectrum would almost certainly not be a power law – in this specific example, it would necessarily turn over due to the missing bright, high-DM population. Hence, a power-law model would simply be the wrong model to fit, and more sensitive tests using s will then be more likely to detect deviations from this model.

Assuming that the relative maximum threshold b is constant over all $\boldsymbol{\theta}$ (to derive equation 16), and set to infinity (to derive equation 19), is not a very stringent criterion – b needs only be very large. This is the case for both FRB 010724 (Lorimer et al. 2007) and FRB 180309 (Oslowski et al. 2018a), which demonstrate the possibility of an FRB hitting a limit at which its S/N can no longer be estimated in the primary detection beam. In both cases, and likely future cases, it was due to insufficient dynamic range causing saturation in recorded data. For FRB 010724, the S/N could be recovered using neighbouring beam detections, while for FRB 180309, saturation affected relatively few scintles in the dynamic spectrum. If this effect can be characterized as a moderate linear reduction in an intrinsically high S/N, the resultant effect on $\log s$ for that event would be small. For example, were the true S/N for FRB 180309 to be 421 rather than 411, $\log s$ would have increased by 0.01 for that particular event, and hence α increased for an FRB sample by an even smaller fraction. Equivalently, this means that the calculated values of α will be insensitive to any cut-off in the peak FRB luminosity lying far above experimental threshold.

It is also worthwhile noting that the derivation in Section 2.1 (and of Crawford et al. 1970) ignores errors in s, i.e. it assumes that the measured S/N is unaffected by noise. As discussed in Murdoch, Crawford & Jauncey (1973), when the threshold S/N, $\mathrm{S/N_{th}}$, is 6 or greater, this has negligible effect on the resulting calculations of α. The calculation of Section 5 illustrates how to account for these errors.

Some cautionary notes are needed. The value of S/N must be calculated using the detection algorithm that sets $\mathrm{S/N_{th}}$, not with a more detailed follow-up calculation (e.g. fitting a detailed pulse profile) that boosts S/N. Doing so would create an artificial gap between the original threshold $\mathrm{S/N_{th}}$ and the boosted S/N values, which could not be filled by boosted subthreshold events, since these are missed by the detection algorithm.

If a data set is re-examined using a more sensitive *detection* method however, then the newer updated values of S/N and $\mathrm{S/N_{th}}$ *must* be used in the calculation, even for FRBs detected in the prior analyses. The more sensitive method will only identify new events that were subthreshold in a previous analysis, equivalent to setting a very low upper sensitivity limit b in equation (18). To recover the $b = \infty$ result of equation (19) therefore requires using updated values for all events. This applies even if no new events are detected, since changing statistical methods based on observed outcomes ('flip-flopping') leads to biased results.

3.1 A possible cause of bias

The most likely cause for bias in calculated values of s, however, is the use of human 'by-eye' discretion to identify plausible candidates against RFI background, as noted by Macquart & Ekers (2018a) in section 2.3 of that work. When these inspection methods are used to rule out RFI-only candidates, as is usually the case, this does not present a bias. However, when a true candidate is detected in the presence of RFI, then the chance that it is falsely rejected decreases with increasing signal strength. Mathematically, this implies a time-varying threshold S/N that can be significantly larger than the nominal computed values. If a candidate is accepted this way, then using the nominal $\mathrm{S/N_{th}}$ will give an artificially large value of s, and hence a biased value of α.

Any method to overcome this bias necessarily involves estimating S/N_{th} on an event-by-event basis. The first step requires using a reproducible algorithm (e.g. a machine learning method) to identify FRBs, and associate them with a statistical S/N value. Once an FRB is detected, however, it should be fed back into the detection algorithm with artificially reduced S/N, i.e. after manually subtracting its fitted flux. This should be repeated until the algorithm no longer identifies the event as an FRB, thus determining the actual S/N_{th} for that event. This method could also work with crowdsourced classification methods, where the initial and subsequent identifications are performed by a random, and hence independent (but statistically identical) ensemble of volunteers. However, it could not be used when a small number of experts perform by-eye analysis, both because the initial identification will not be replicable, and subsequent identifications of reduced S/N events will not be independent.

3.2 Equivalence to V/V_{max} test

The V/V_{max} test of Schmidt (1968) has been discussed in the context of FRBs by Oppermann et al. (2016) and Macquart & Ekers (2018a), and applied to a sample of CRAFT FRBs by Shannon et al. (2018). The test calculates for each detection the volume of space over which a source would have been detectable, V_{max}, and the volume enclosed within the actual detected distance, V. For a non-evolving source distribution, the measured values of V/V_{max} will be uniformly distributed between 0 and 1, with a mean of 0.5, and standard deviation of $\frac{1}{2\sqrt{3}}$ (for a single sample).

In the case of FRBs, where the distance to each source is generally unknown, the value of V/V_{max} for a particular event in Euclidean space can be calculated using

$$\begin{aligned} \frac{V}{V_{max}} &= \frac{R^3}{R_{max}^3}, \\ \frac{R}{R_{max}} &= \left(\frac{F}{F_{th}}\right)^{-0.5} \\ \Rightarrow \frac{V}{V_{max}} &\equiv s^{-1.5}. \end{aligned} \tag{22}$$

In other words, the V/V_{max} test uses exactly the same information from each event (s) as the calculation of α in equation (21). This explains why the V/V_{max} test is insensitive to variations in burst properties and variations in sensitivity, as noted by Shannon et al. (2018).

An interesting exercise is to determine for which kinds of true source distributions each test more effectively rejects the Euclidian, non-evolving hypothesis. In the case of pure power-law distributions, Table 1 compares the expected values for each test as a function of α, and the rejection power r, defined here as

$$r = \lim_{N\to\infty} \sqrt{N}\left(\frac{\mu(\alpha) - \mu(\alpha = -1.5)}{\sigma(\alpha = -1.5)}\right), \tag{23}$$

where μ is the expected value of each test, and σ is the standard deviation in the case of $\alpha = -1.5$ and number of samples N. The limit $N \to \infty$ is taken so that σ approaches its large N form, and biases tend to zero (i.e. $\alpha' \to \alpha$), while the factor $\sqrt{N}$ renormalizes this to a single sample. From Table 1, the rejection power of the maximum likelihood test is approximately 30 per cent greater than that of the V/V_{max} test. This is to be expected, since the maximum likelihood estimate has been optimized specifically for a power-law distribution.

Table 1. Rejection power r, defined as per equation (23), for rejecting the $\alpha = -1.5$ hypothesis, in the case where the true FRB fluence distribution is a power law with index given by the leftmost column. In the $N \to \infty$ limit, this will also equal the expectation value μ_l for the maximum likelihood estimation of α from equation (3). The expectation value of the V/V_{max} test is given by μ_v, while r_l and r_v are the rejection powers of the maximum likelihood test (equation 3) and the V/V_{max} test (equation 22), respectively.

$\alpha(=\mu_l)$	μ_v	r_l	r_v
−1.0	0.410	−0.41	−0.31
−1.1	0.430	−0.33	−0.24
−1.2	0.449	−0.25	−0.18
−1.3	0.467	−0.16	−0.11
−1.4	0.484	−0.08	−0.06
−1.5	0.500	0	0
−1.6	0.516	0.08	0.06
−1.7	0.531	0.16	0.11
−1.8	0.544	0.24	0.15
−1.9	0.557	0.33	0.20
−2.0	0.570	0.41	0.24

4 CALCULATIONS WITH PARKES AND ASKAP DATA

The methods described above can be applied simultaneously to combined sets of FRB data from multiple telescopes. Here, we limit this to the two instruments with sufficiently large numbers of detected FRBs such that fits to individual telescope data sets will be meaningful. Currently, this is data from various searches with the Parkes radio telescope, and FRBs detected by the Commensal Real-time ASKAP Fast Transients (CRAFT; Macquart et al. (2010)) survey with ASKAP. In the near future, the upgraded Molonglo Observatory Synthesis Telescope (UTMOST; Bailes et al. 2017) and the Canadian Hydrogen Intensity Mapping Experiment (CHIME; The CHIME/FRB Collaboration et al. 2018) should both have sufficiently high statistics to apply this method.

Applying equation (3) to Parkes and ASKAP data requires only being able to calculate s from equation (20) for each detected FRB, i.e. it requires knowing both the detected and instantaneous threshold S/Ns. As discussed in Section 3, the multitude of potentially complicated effects causing some fraction of FRBs to remain undetected can be ignored.

All FRBs detected by Parkes listed on FRBCAT (Petroff et al. 2016) with published S/N_i and S/N_{th} are given in Table 2. The only exceptions were FRB 150807 (Ravi et al. 2016), for which no published S/N or S/N_{th} value exists; and the Lorimer Burst (FRB 010724; Lorimer et al. 2007), which is excluded on grounds of discovery bias (Macquart & Ekers 2018a). Equivalent values for ASKAP FRBs detected by CRAFT, which all used $S/N_{th} = 9.5$ (after post-processing of initial data searched at $S/N_{th} = 10$), are given in Table 3. Fig. 1 shows the distributions of $N(s)$ for these samples.

Using these values as input to equation (21), we calculate bias-corrected values $\alpha'_p = -1.18 \pm 0.24$ and $\alpha'_a = -2.20 \pm 0.47$ for Parkes and ASKAP/CRAFT, respectively. The errors are 68 per cent CI, calculated using equation (5), and scaled according to equation (6) for α'. They are symmetric at the stated level of precision. Thus the difference corresponds to 1.9σ assuming Gaussian errors.

The analysis allows both data sets to be combined if we assume that the same power law governs both data sets. Doing so produces $\alpha'_c = -1.55 \pm 0.23$.

Table 2. Parkes FRBs with known S/Ns and threshold ratios S/N_{th}, allowing relative detection strength s to be calculated.

FRB	S/N	S/N_{th}	s	Ref.
110220	49	9	5.44	Thornton et al. (2013)
110627	11	9	1.22	
110703	16	9	1.78	
120127	11	9	1.22	
090625	28	10	2.8	Champion et al. (2016)
121002	16	10	1.6	
130626	20	10	2	
130628	29	10	2.9	
130729	14	10	1.4	
150418	39	10	3.9	Keane et al. (2016)
150610	18	10	1.8	Bhandari et al. (2018)
151206	10	10	1	
151230	17	10	1.7	
160102	16	10	1.6	
010125	16.9	7	2.415	Burke-Spolaor & Bannister (2014)
010621	16.3	8	2.04	Keane et al. (2011); Keane (private communication)
131104	30.6	10	3.06	Ravi et al. (2016)
140514	16	10	1.6	Petroff et al. (2015)
150215	19	10	1.9	Petroff et al. (2017)
171209	40	10	4	Shannon et al. (2017)
180301	16	10	1.6	Price et al. (2018)
180309	411	10	41.1	Oslowski et al. (2018a)
180311	11.5	10	1.15	Oslowski et al. (2018b)
180714	22	10	2.2	Oslowski et al. (2018c)

Table 3. CRAFT FRBs showing detected S/Ns and relative ratio s, compared to the detection threshold $S/N_{th} = 9.5$.

FRB	S/N	s
	Bannister et al. (2017)	
170107	16	1.68
	Shannon et al. (2018)	
170416	13.1	1.38
170428	10.5	1.11
170712	12.7	1.34
170707	9.5	1.00
170906	17	1.79
171003	13.8	1.45
171004	10.9	1.15
171019	23.4	2.46
171020	11.	2.05
171116	11.8	1.24
171213	25.1	2.64
180110	35.6	3.75
180119	15.9	1.67
180128.0	12.4	1.31
180128.2	9.6	1.01
180130	10.3	1.08
180131	13.8	1.45
180212	18.3	1.93
	Macquart et al. (2018)	
180315	10.5	1.11
180324	9.8	1.03
180525	27.4	2.88

To check the robustness of the fit, a cut-off in s, s_{cut}, is introduced. Modified values of s, s_i', are used in equation (3), calculated as

$$s_i' = \frac{s_i}{s_{cut}}, \tag{24}$$

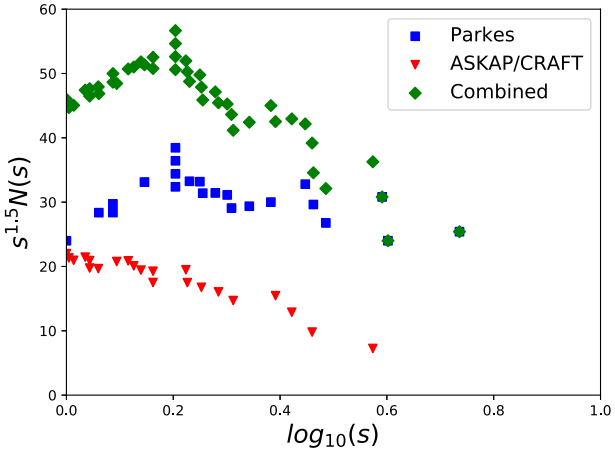

Figure 1. Cumulative source-count distributions $N(s)$ of the Parkes, ASKAP/CRAFT, and combined FRB samples used in this work, as a function of the logarithm of relative detection significance s. The factor $s^{1.5}$ normalizes $N(s)$ to the Euclidean expectation. The point due to FRB 180309 at $\log_{10}(s) = 1.61$ is not shown .

with only data points satisfying $s_i' \geq 1$ included in the sum. Doing so should produce statistically similar (although correlated) values of α', due to the scale invariance of power-law distributions, while α will slowly vary as per the expected bias. Significantly changing values of α' indicate deviations from a power law, e.g. due to near-threshold effects. This procedure is similar to that used by Macquart & Ekers (2018a) to search for a threshold fluence, where s_{cut} is set to each measured value s_i successively.

Fig. 2 shows the change of both corrected α' and uncorrected α as a function of s_{cut}. For each data set, changing s_{cut} produces only small variations in α and α', since neighbouring estimates differ by at most one FRB, and are thus highly correlated. This explains why the error ranges are much larger than the variation in the data. While the errors gradually increase, the calculated value for Parkes data, α_p', stays approximately constant. There is a small hint of a systematic decrease in α_p' for $s_{cut} < 1.6$, which is consistent with the notion that RFI is acting to obscure marginal events, as discussed in Section 3. However, this trend does not appear to be statistically significant, and is followed by a subsequent increase in α_p in the range $1.6 < s_{cut} < 2$, i.e. there is no conclusive evidence for this behaviour in the Parkes data.

This is not the case for ASKAP data however, where the fitted value of α' decreases rapidly with the cut-off fluence, i.e. the FRB integral source-count distribution appears to be steeper at higher fluences. However, the lower bound of the 68 per cent CI stays constant at around -1.7, so this also should not be seen as a significant result.

Fig. 2 illustrates the bias that comes with using detected values of s_i as cut-off fluences. Local minima in α are found when $s_{cut} = s_i$, and local maxima when s_{cut} is slightly above s_i. Since a point with $s_i = s_{cut}$ contributes nothing to the sum in equation (3), but does change the normalization N by 1, the increase in α when moving from using N to $N - 1$ data points is a factor of $(N - 1)/N$. Choosing either extremum can thus be associated with a systematic error of $\pm 0.5\alpha/N$, where the error is negative (positive) for including (excluding) the threshold point s_i.

To test for shape deviations from a pure power law, the Kolmogorov–Smirnov (K–S) test (Kolmogorov 1933; Smirnov 1948) is performed for each of the data sets, and illustrated in Fig. 3.

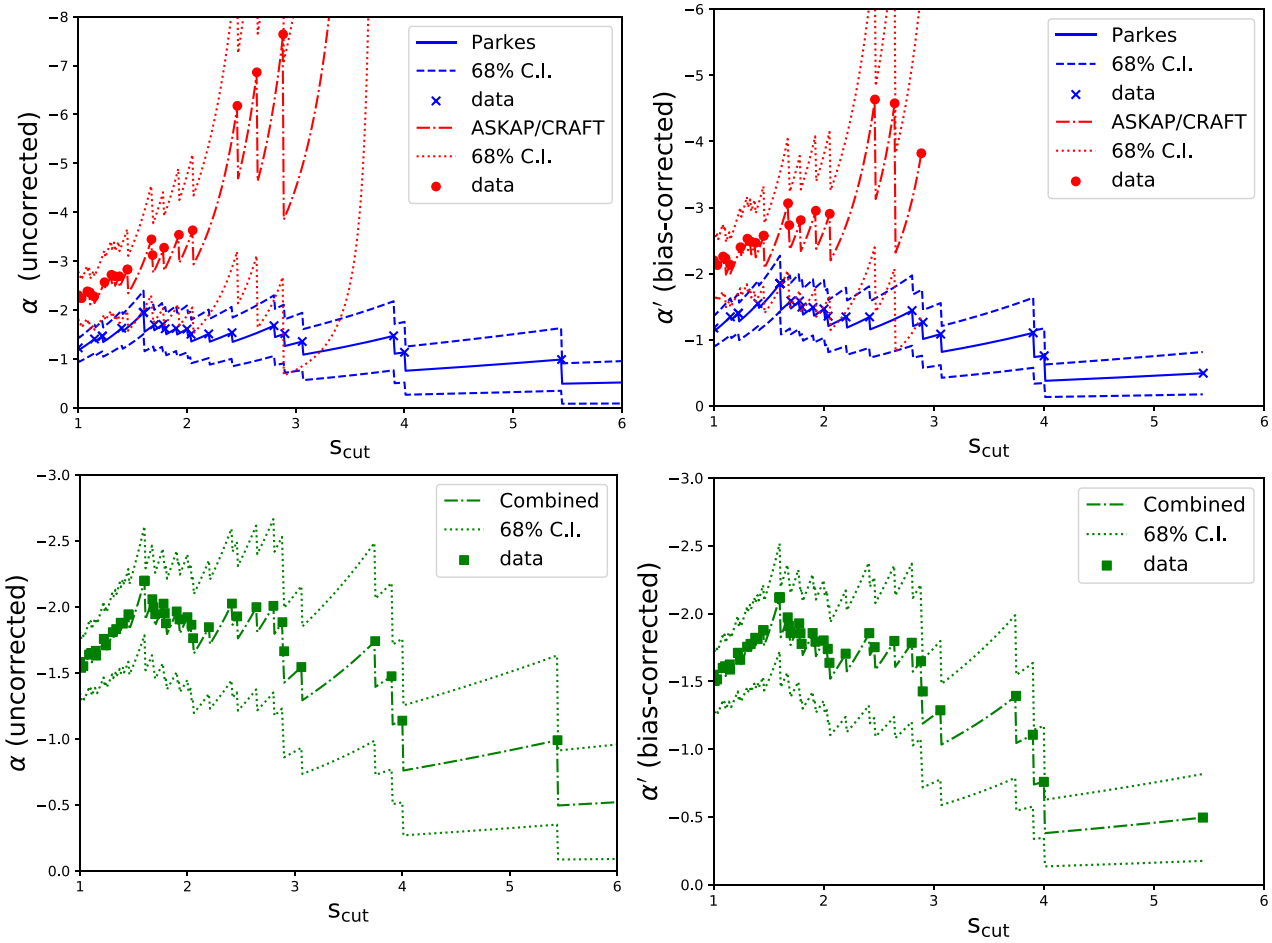

Figure 2. Calculated values of uncorrected α (left) and corrected α' (right), with upper plots using Parkes (blue) and ASKAP/CRAFT (red) data, and lower plots using a combined data set. Calculations used equation (3), as a function of cut-off S/N, $s_{\rm cut}$. This was performed as a continuous function of $s_{\rm cut}$ (lines, with error bars according to equation 5), and using values of detected FRBs s_i as the cut-off (points). The uncorrected plots are shown to illustrate the bias on estimates of α when compared to the corrected values.

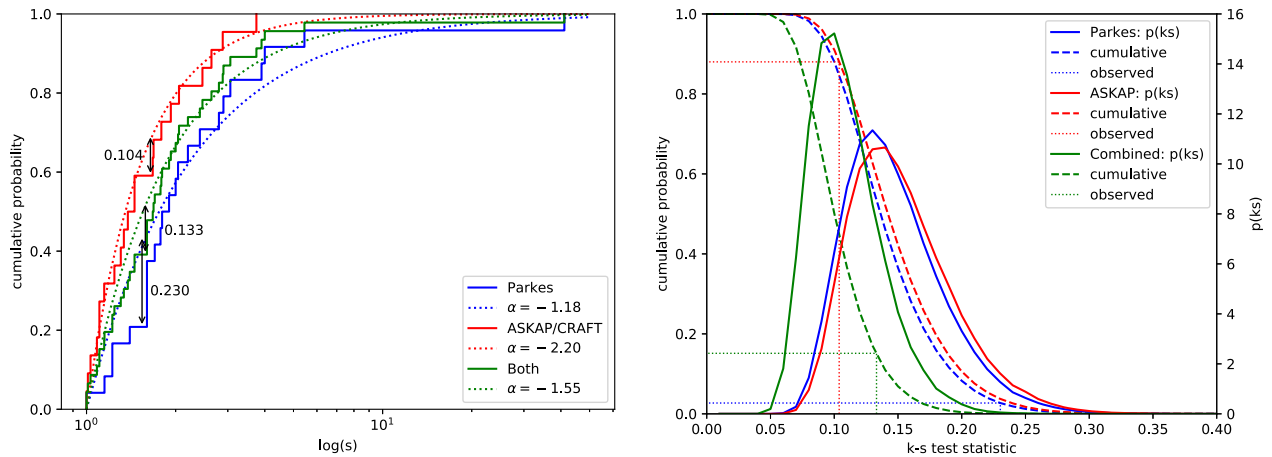

Figure 3. Left: Kolmogorov–Smirnov (K–S) test for compatibility between bias-corrected power-law fits (dotted) to Parkes (blue), ASKAP (red), and combined (green) data, as a function of the natural log of relative S/N, s (equation 20). Maximum deviations, i.e. the K–S statistics, are indicated by black arrows. Right: results of a Monte Carlo estimate of the distributions of the test statistic K–S, assuming true power laws equal to the fitted values α'. Shown are the cumulative (left-hand axis) and differential (right-hand axis) probability distributions for Parkes and ASKAP FRBs, and lines indicating the observed values of the K–S statistic for each.

The maximum deviations for each of the Parkes, ASKAP/CRAFT, and combined samples, i.e. the K–S statistics, are 0.230, 0.104, and 0.133, respectively.

Since the values of α are estimated from the data, the standard tables of the K–S test statistic under the null hypothesis do not apply. Goldstein, Morris & Yen (2004) provide tables of the test statistic for fitted power-law distributions in the case where the estimated value of α is drawn from the data, using the Monte Carlo method. However, their table is sparse, and it is much simpler to perform such a Monte Carlo estimate of the distribution of the K–S statistic under the null hypothesis, which allows a p-value to be assigned. The results of this investigation are given in Fig. 3 (right). The associated p-values (probability of seeing a K–S statistic equal to or greater than that observed, assuming a true power law with values given by α') are 2.7 per cent (Parkes), 88 per cent (ASKAP), and 15 per cent (combined).[1] That is, from the K–S test, there is some evidence that the Parkes sample deviates significantly from an underlying power-law distribution, but none in the case of ASKAP data. The result from the combined case is, as expected, intermediate.

The K–S statistic for Parkes data is a maximum immediately below $s_{\rm cut} = 1.6$. Using data at and above $s_{\rm cut} = 1.6$ only, we find $\alpha' = 1.85$, for which there is no significant discrepancy with the value of α' found for ASKAP data.

That the K–S statistic highlights the Parkes results as being less consistent with a power-law distribution than the ASKAP/CRAFT data, whereas Fig. 2 suggests the opposite, is likely due to the correlated nature of the data presented in Fig. 2, and a reminder of the dangers of by-eye interpretations of it.

The K–S test is a very general, and consequently not very powerful, test for differences in the integral source-count spectrum between the two data sets. A full likelihood ratio test is used in Section 5 to quantify this difference. First, we make some comments on previous results.

4.1 Comments on previous results using completeness fluences

The maximum likelihood method of Crawford et al. (1970) used in Section 2 has been invoked by both Macquart & Ekers (2018a) and Bhandari et al. (2018) on different samples of Parkes FRBs, finding $\alpha = -2.6^{+0.7}_{-1.3}$ and $-2.2^{+0.6}_{-1.2}$ for $N = 9$ and 5 FRBs above threshold, respectively. In both cases, the values of s inserted into equation (3) were the estimated fluences (with no beam correction) relative to a completeness threshold in fluence. Macquart & Ekers (2018a) calculate α using only observed fluences as thresholds, and hence their estimate of α is biased downwards by $0.5\alpha/N = 0.14$, whereas Bhandari et al. (2018) use a pre-determined experimental completeness threshold of 2 Jy ms. Also correcting for the intrinsic bias in α, the adjusted estimates of these authors are $\alpha' = -2.2^{+0.55}_{-1}$ and $-2.0^{+0.55}_{-1.1}$, respectively (errors are scaled identically to the mean).

It is also interesting to comment on the use of a 'completeness threshold' F_0 with which to normalize observed fluences F in calculating s. First, note that in the original historical context (source counts in the Third Cambridge Catalogue of radio galaxies), it is not readily possible to estimate an S/N – see e.g. the discussion in section II of Murdoch et al. (1973). Thus it was more common to try to find a 'completeness flux' (here fluence) $F_{\rm c}$ above which all sources in the survey area would have been detected. Applying this notion to the formalism of Section 2.1, $F_{\rm c}$ should be set such that

$$F_{\rm c} \geq F_{\rm th}(\boldsymbol{\theta}) \, \forall \, \boldsymbol{\theta}. \tag{25}$$

In this case, $F_{\rm c}$ can simply replace $F_{\rm th}$ in equation (9), and the basic method of Macquart & Ekers (2018a) – increasing $F_{\rm c}$ until stability is reached in α – is sound. However, it has several deficiencies in the case of FRB counts.

First, the nominal value of $F_{\rm c} = 2$ Jy ms in the case of most searches at Parkes arises from the *observed* maximum pulse duration of 30 ms (Keane & Petroff 2015). However, if $F_{\rm th}$ exceeds the completeness threshold $F_{\rm c}$ over any part of the parameter space for which FRBs exist (i.e. $k(\boldsymbol{\theta})$ is non-zero in this range), then the survey will not in fact be complete. For instance, FRBs of 32 ms width. The lack of FRBs with observed widths in this range can either be interpreted as validating the $k(\boldsymbol{\theta}) = 0$ requirement, or as being evidence of incompleteness itself.

Secondly, the criteria of equation (25) results in many events being rejected, which is unnecessary when their relative S/N can simply be calculated according to equation (20). In particular, for short FRBs, 2 Jy ms is well above the actual detection threshold, which is closer to 0.5 Jy ms for FRBs of 1 ms duration.

Thirdly, equation (6) shows how the bias in $\alpha(F_{\rm c})$ increases as the number of events in the sample is reduced. The effect is even greater if the events themselves define the threshold. It induces an artificial slope in plots of α as a function of threshold, which can be seen in Fig. 2 (left). Any method that searches for a completeness limit via a flattening in $\alpha(F_{\rm c})$ above a critical value will therefore misidentify the completeness limit.

Fourthly, whereas S/N is readily calculable for a FRB search algorithm, care must be taken not to calculate the observed fluence (which is a derived data product) using a more advanced method than that used for the FRB search, e.g. by the fitting of and integration over a Gaussian profile. Of course, corrected fluences (e.g. beam-corrected values) can yield important information on FRB properties – they just should not be used for this method.

Finally, when considering the impact of human discretion in identifying candidates against an RFI background, the use of a completeness threshold defined in terms of fluence rather than S/N means that some candidates (e.g. those with long durations) with fluence near $F_{\rm c}$ will indeed be marginal and difficult to identify by eye, while others (e.g. those with short durations) will be extremely strong. Artificially increasing s as per equation (24) is thus a much more effective way to search for this effect.

The other aspect to note in the method of Macquart & Ekers (2018a) (and applied by Bhandari et al. 2018) is that beam shape remains a hidden and uncorrected variable. That is, while a sample may be complete in received fluence, it is not complete in true fluence. Hence, while the former authors note that defining thresholds in terms of S/N leads to incompleteness at a given fluence level, in fact a completeness fluence $F_{\rm c}$ cannot be defined in terms of true FRB fluence either, which will always be greater by some amount than the observed fluence.

In Appendix, we show – for completeness – how this case (defining S/N and/or $F_{\rm c}$ in terms of detected fluence) reduces to the results of Section 2.1, and how the threshold method of equation (25) applies (i.e. equally, and with the same caveats). This result can be briefly understood by realizing that a detection off beam centre will affect both threshold and measured S/N (or equivalently fluence) by an equal, if unknown, amount, thus preserving their ratio for statistical purposes.

[1] The trial factor is somewhere between 2 and 3 due to the correlated nature of the combined and individual tests, corresponding to between a 5.5 and 8.2 per cent chance of any result being as inconsistent as the Parkes data under the null hypothesis of a true power-law distribution.

5 FULL MAXIMUM LIKELIHOOD CALCULATION OF PARKES AND ASKAP FRBS IN THE PRESENCE OF NOISE

We now present a traditional maximum likelihood estimate of α, and use a likelihood ratio test to determine whether separate power laws describe ASKAP/CRAFT and Parkes data significantly better than a single power law.

The methods used in Section 2 ignore the contribution of noise to the measured value of S/N, and hence s. Such a case is covered by Murdoch et al. (1973), where the authors note that when the threshold S/N is 6 or greater, no significant effect due to noise is expected when calculating the power-law index α. Here, we explicitly check this using a full likelihood maximization.

The likelihood of observing an S/N, $S_{\rm obs}$, for a true power-law distribution of $S_{\rm true}$ depends on the noise deviate $n = S_{\rm obs} - S_{\rm true}$:

$$P(S_{\rm obs}) = \frac{1}{C}\int_{-\infty}^{+\infty} P(n)\,{\rm d}n \left(\frac{S_{\rm true}}{S_{\rm th}}\right)^{\alpha}, \tag{26}$$

where the normalization constant C is the probability of observing any such event above a threshold $S_{\rm th}$:

$$C = \int_{S_{\rm th}}^{\infty} P(S_{\rm obs}){\rm d}S_{\rm obs}. \tag{27}$$

For standard normal deviates, $P(S_{\rm obs})$ is simply

$$P(S_{\rm obs}) = \frac{1}{C}\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}}\,{\rm e}^{-(0.5n^2)}{\rm d}n \left(\frac{S_{\rm true}}{S_{\rm th}}\right)^{\alpha}. \tag{28}$$

The integral over dn can in practice be limited to a small range – here ± 5 is used. This has been tested to reproduce the values of Murdoch et al. (1973) (table 1) to within the stated precision of four significant figures.

The simplest definition of the likelihood function $\mathcal{L}$ for the Parkes and ASKAP samples of FRBs is therefore,

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} \log P(S^i_{\rm obs}), \tag{29}$$

where the sum proceeds over all i observations. The P_i are calculated according to equation (28), with $S_{\rm obs}$ and $S_{\rm th}$ the measured and threshold S/Ns for that particular observation (the constant C must also be renormalized for each observation). This is only applicable when using Parkes data; the ASKAP data used a constant threshold S/N of 9.5σ.

A more correct estimate in the case of Parkes or combined data would require weighting each observation by the relative fraction of observation time spent observing at that given threshold. However, these values are generally not available, and the expected loss of precision for a small variation in $S_{\rm th}$ is also small.

Calculating $\mathcal{L}$ as a function of α for the ASKAP, Parkes, and combined samples (Tables 2 and 3) produces the values shown in Fig. 4. To estimate the influence of Gaussian noise, calculations were also performed by setting $S_{\rm obs} = S_{\rm true}$, i.e. $P(S_{\rm obs}) = \delta(S_{\rm obs} - S_{\rm true})$. In general, this was found to have negligible effect on the resulting likelihoods, except for large values of α.

The best-fitting values of α are found at the maximum of $\mathcal{L}$, $\mathcal{L}^{\rm max}$. For the Parkes, ASKAP/CRAFT, and combined data sets, these were -1.24, -2.40, and -1.61, respectively. These are consistent with both corrected and uncorrected values calculated in Section 3, although the uncorrected values are the proper comparators (the bias of the analytic method is inherent in the maximum likelihood procedure).

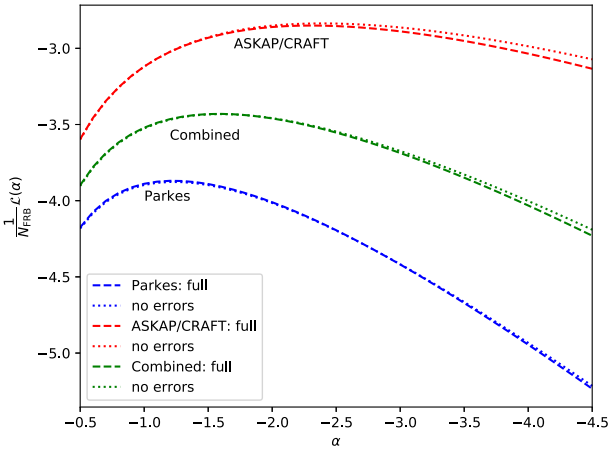

Figure 4. Evolution of the maximum likelihood $\mathcal{L}$ as a function of α for Parkes, ASKAP/CRAFT, and combined data sets. Dashed lines use the full Gaussian error distribution from equation (28), while dotted lines ignore the noise contribution to $S_{\rm obs}$.

While this numerical method does not allow CI to be set for α, it does allow the likelihood ratio test to be performed. This is because the combined fit represents a constricted model of individual fits to ASKAP and Parkes data. The test statistic $\mathcal{D}$ is defined as

$$\mathcal{D} = -2\log\left(\frac{\mathcal{L}_{\rm c}^{\rm max}}{\mathcal{L}_{\rm a}^{\rm max} + \mathcal{L}_{\rm p}^{\rm max}}\right), \tag{30}$$

where 'p', 'a', and 'c' denote Parkes, ASKAP, and combined data, respectively. According to Wilks' theorem (Wilks 1938), as the number of observations tend to infinity, the distribution of $\mathcal{D}$ under the null hypothesis (that both Parkes and ASKAP data come from the same power-law distribution) will approach a χ^2_1 distribution. The number of degrees of freedom is 1 due to the single extra power law being fitted.

For this case, $\mathcal{D} = 6.9$, corresponding to a one-sided p-value of 0.86 per cent for a χ^2_1 distribution. That is, there is evidence at the 2.6σ level that the Parkes and ASKAP data originate from distributions with power laws of different indices, rather than the same index.

Note that this test is complementary to the K–S test presented in Section 3, which tests for deviations from power-law-like behaviour. In the specific case of testing one versus two power laws, however, this test is more sensitive, since it tests a specific deviation from single power-law behaviour, rather than the general case. Hence, greater inconsistency with the null hypothesis of both Parkes and ASKAP data coming from the same power-law distribution is found.

Specifically testing for differences of each sample with a Euclidean distribution ($\alpha = -1.5$), the test statistic becomes

$$\mathcal{D} = -2\log\left(\frac{\mathcal{L}_{\rm p/c}(\alpha = -1.5)}{\mathcal{L}_{\rm p/c}^{\rm max}}\right). \tag{31}$$

This produces $D = 1.01$ for Parkes data (no evidence for deviations from a Euclidean distribution) and $D = 4.00$ for ASKAP/CRAFT data, corresponding to a one-sided 5 per cent probability of a more extreme deviation in the case of a Euclidean distribution.

6 DISCUSSION

6.1 Evidence of a break in the FRB source-count distribution

Using the likelihood ratio test, we find evidence (p-value 0.86 per cent, i.e. 2.6σ) that the Parkes and ASKAP/CRAFT samples arise from distributions with different power-law indices, of -1.22 ± 0.25 and -2.15 ± 0.49, respectively (68 per cent CI). Our ASKAP/CRAFT result agrees with the value of $-2.1^{+0.6}_{-0.5}$ found by Shannon et al. (2018) using an almost identical sample, but with the less sensitive $V/V_{\rm max}$ test.

The nominal Parkes threshold for FRB detection is 0.5 Jy ms for a 1 ms duration FRB (Keane et al. 2018), while for ASKAP, it is 26 Jy ms (Shannon et al. 2018). Therefore, if the FRB source-count distribution observed by ASKAP is steeper, Parkes cannot observe a pure power-law distribution. This is confirmed by our K–S test results, which show that only 2.7 per cent of true power-law distributions would show a worse fit to Parkes data.

Under the null hypothesis of Parkes and ASKAP/CRAFT data coming from the same power-law distribution, both p-values – denoted as $p_1 = 0.86$ per cent and $p_2 = 2.7$ per cent – will be independent and uniformly distributed (this has been confirmed with simulations). Their product p_1p_2 will therefore have a cumulative distribution of $p_1p_2(1 - \ln p_1p_2)$, with low values providing evidence for the alternative hypothesis of deviation from power-law behaviour. This one-sided test provides 0.22 per cent (3.1σ) evidence against the null hypothesis. We have performed this calculation a posteriori however, and therefore consider our a priori probability of 0.86 per cent more reliable.

The K–S test of the ASKAP/CRAFT S/N distribution tests for the steepening occurring above the ASKAP detection threshold. Since we find this data to be well fitted by a power law (p-value 86 per cent), this is consistent with the steepening occurring over the fluence range between the Parkes and ASKAP/CRAFT thresholds.

Another interpretation of our results is the potential bias discussed in Section 3.1, whereby effective experimental thresholds are significantly larger than the claimed thresholds. If this is correct, then Parkes data below $s = 1.6$ should be excluded, i.e. the true effective experimental threshold for most Parkes observations must be near S/N = 16, given for the most common Parkes threshold of 10σ.

Further evidence for a turn-over in the observed Parkes rates, however, can be found in the literature. The values of α found for Parkes data by both Macquart & Ekers (2018a) and Bhandari et al. (2018), after correcting for biases as discussed in Section 4.1, are $\alpha = -2.2^{+0.55}_{-1}$ and $-2.0^{+0.55}_{-1.1}$, respectively. While these estimates use an artificially high fluence threshold due to the use of a completeness limit, the calculations – after bias correction – are still statistically valid. Their results therefore appear to be in tension with our value of $\alpha = -1.22 \pm 0.25$ for Parkes. However, the higher fluence threshold samples a fluence regime more similar to that seen by ASKAP, so that if a break in the source-count spectrum is present, then the used completeness threshold likely lies above it. This then explains the better agreement with the ASKAP value of α found here despite the use of data from Parkes.

The effective threshold at $\alpha = -2.2$ of ASKAP/CRAFT observations is 40 Jy ms to a 1.2656-ms pulse (James et al. 2018). For Parkes, the nominal threshold of 0.5 Jy ms would be increased to an effective threshold of at least 5 Jy ms for $\alpha > -1.5$ and a Gaussian beam shape (James et al. 2018). Therefore, we expect the downturn in the FRB fluence distribution, $F_{\rm b}$, to lie in this range.

6.2 Interpretation

In this section, we consider what our results would imply about the nature of FRBs, should they be verified by further data.

Macquart & Ekers (2018b) discuss how features in the FRB source-count distribution will be intimately related to the cosmological evolution of the FRB progenitors. A value of $\alpha = -2.15$ found here is typical of a flat FRB spectral index s_ν ($F \propto \nu^{s_\nu}$ for frequency ν) and source evolution function that is more strongly peaked than the star formation rate. It is analogous to the high-fluence limit of active galactic nuclei (AGNs) source counts (e.g. Wall, Pearson & Longair 1980).

Macquart & Ekers (2018b) also note that a flat distribution in FRB energy/luminosity up to some maximum $E_{\rm max}/L_{\rm max}$, combined with a downturn in cosmological source evolution (at redshift z_0), will lead to a break in the observed differential source-count distribution. The break is predicted to be at a fluence $F_{\rm b} = E_{\rm max}(1+z_0)^{2+s_\nu}(4\pi D_L^2(z_0)^{-1})$, where D_L is the luminosity distance.

FRB data are currently too sparse to produce a full fit of the FRB luminosity function. However, it is useful to check whether existing constraints on $F_{\rm b}$, $E_{\rm max}$, z_0, and s_ν can be mutually consistent with the model of Macquart & Ekers (2018b). Shannon et al. (2018) observe that ASKAP/CRAFT FRBs are viewed out to approximately $z \sim 1$, and Parkes FRBs out to $z \sim 2$–3 (suggesting $1 \leq z_0 \leq 3$), and that the most energetic FRBs detected by ASKAP/CRAFT, Parkes, and UTMOST have energies $E_{\rm max}$ around 10^{33}–10^{34} erg Hz^{-1}. Macquart et al. (2018) find $s_\nu = -1.5 \pm 0.3$ for the ASKAP/CRAFT sample. These ranges for s_ν, $E_{\rm max}$, and z_0 are indeed generally consistent with $5 \leq F_{\rm b} \leq 40$ Jy ms, although some regions of the parameter space can be excluded. The observed ranges are not hard limits however, and further analyses and/or observations will be required to further constrain these parameters.

7 CONCLUSIONS

We have extended the statistical results of Crawford et al. (1970), allowing surveys with unknown completeness thresholds in physical units such as flux/fluence, but readily definable detection and threshold values of S/N, to calculate the slope of power-law distributions of detected events.

Applied to FRBs, this allows improved estimates of the slope of the integral source-count distribution, α, without the need to consider a completeness limit, beam pattern, or any other confounding factor. Combining detections with the Parkes radio telescope and CRAFT detections using ASKAP, we find a bias-corrected value for α of -1.52 ± 0.24. Fitting these samples individually however, we find -1.18 ± 0.24 and -2.20 ± 0.47, respectively. A likelihood ratio test indicates this is compatible with a single power law at a p-value of 0.86 per cent, and that the ASKAP/CRAFT sample is only compatible with the Euclidean expectation of -1.5 at a 5 per cent level. This implies either a steepening in the FRB luminosity function in the fluence range 5–40 Jy ms, or that the experimental threshold for most Parkes FRBs is significantly higher than reported, being approximately 16σ. This is the first hint of structure in the FRB source-count distribution, which could be the first evidence of a turn-over in a sharply peaked source evolution of FRB progenitors in the redshift range 1–3.

ACKNOWLEDGEMENTS

We thank S. Bhandari for discussion on the analysis of Parkes FRBs detected by SUPERB. RMS acknowledges support through Australian Research Council (ARC) grants FL150100148 and CE170100004. This work was supported by resources provided by The Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. Parts of this research were conducted by the Australian Research Council Centres of Excellence for All-Sky Astrophysics (CAASTRO, CE110001020). This research was also supported by the Australian Research Council through grant DP18010085.

REFERENCES

Bailes M. et al., 2017, Publ. Astron. Soc. Aust., 34, e045
Bannister K. W. et al., 2017, ApJ, 841, L12
Bhandari S. et al., 2018, MNRAS, 475, 1427
Burke-Spolaor S., Bannister K. W., 2014, ApJ, 792, 19
Champion D. J. et al., 2016, MNRAS, 460, L30
Chatterjee S. et al., 2017, Nature, 541, 58
Crawford D. F., Jauncey D. L., Murdoch H. S., 1970, ApJ, 162, 405
Goldstein M. L., Morris S. A., Yen G. G., 2004, European Phys. J. B, 41, 255
James C. W. et al., 2018, PASA, (arXiv:1810.04356)
Katz J. I., 2017, MNRAS, 472, L85
Keane E. F., Petroff E., 2015, MNRAS, 447, 2852
Keane E. F., Kramer M., Lyne A. G., Stappers B. W., McLaughlin M. A., 2011, MNRAS, 415, 3065
Keane E. F. et al., 2016, Nature, 530, 453
Keane E. F. et al., 2018, MNRAS, 473, 116
Kolmogorov A., 1933, G. Ist. Ital. Attuari, 4, 83
Lawrence E., Vander Wiel S., Law C., Burke Spolaor S., Bower G. C., 2017, AJ, 154, 117
Lorimer D. R., Bailes M., McLaughlin M. A., Narkevic D. J., Crawford F., 2007, Science, 318, 777
Macquart J.-P., Ekers R. D., 2018a, MNRAS, 474, 1900
Macquart J.-P., Ekers R. D., 2018b, MNRAS, 480, 4211
Macquart J.-P. et al., 2010, Publ. Astron. Soc. Aust., 27, 272
Macquart J.-P., Shannon R. M., Bannister K. W., James C. W., Ekers R. D., Bunton J. D., 2018, preprint (arXiv:1810.04353)
Murdoch H. S., Crawford D. F., Jauncey D. L., 1973, ApJ, 183, 1
Oppermann N., Connor L. D., Pen U.-L., 2016, MNRAS, 461, 984
Oslowski S. et al., 2018a, Astron. Telegram, 11385
Oslowski S. et al., 2018b, Astron. Telegram, 11396
Oslowski S. et al., 2018c, Astron. Telegram, 11851
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., 2017, MNRAS, 469, 4465
Price D. C. et al., 2018, Astron. Telegram, 11376
Ravi V. et al., 2016, Science, 354, 1249
Schmidt M., 1968, ApJ, 151, 393
Shannon R. M. et al., 2017, Astron. Telegram, 11046
Shannon R. M. et al., 2018, Nature, 562, 386
Smirnov N., 1948, Ann. Math. Stat., 19, 279
Spitler L. G. et al., 2016, Nature, 531, 202
Tendulkar S. P. et al., 2017, ApJ, 834, L7
The CHIME/FRB Collaboration et al., 2018, ApJ, 863, 48
Thornton D. et al., 2013, Science, 341, 53
Vedantham H. K., Ravi V., Hallinan G., Shannon R. M., 2016, ApJ, 830, 75
Wall J. V., 1996, in Ekers R. D., Fanti C., Padrielli L., eds, Proc. IAU Symp. 175, Extragalactic Radio Sources. Kluwer, Dordrecht, p. 547
Wall J. V., Pearson T. J., Longair M. S., 1980, MNRAS, 193, 683
Wilks S. S., 1938, Ann. Math. Stat., 9, 60

APPENDIX: EXPLICIT BEAM DEPENDENCE OF TELESCOPE SENSITIVITY

The abstract definition of the parameter set $\boldsymbol{\theta}$ in Section 2.1 makes it difficult to relate to specific experimental examples. Below is explicitly outlined its application in the context of a 'completeness limit' defined by observable fluence (i.e. reduced by the telescope beam), as discussed in Section 4.1, and used by Macquart & Ekers (2018a) and Bhandari et al. (2018). A similar method of explicitly removing beam-shape effects from observed FRB fluence statistics has been presented by Katz (2017), albeit only in the case where all other confounding parameters are ignored, and without reference or reduction to the statistical methods of Crawford et al. (1970).

Let the beam shape of a telescope be described by the direction-dependent relative sensitivity $B(\Omega)$, such that $B(\Omega) \leq 1$, with (in general) equality at beam centre. Ignoring frequency dependencies (which could be treated within the set $\boldsymbol{\theta}$), an FRB with fluence F would be registered with modified fluence F':

$$F'(B) = BF. \tag{A1}$$

The rate of FRBs, and hence the source term k of equation (7), will not depend on B, and FRBs will arrive uniformly in solid angle Ω. Furthermore, assume that the threshold $F_{\rm th}(\boldsymbol{\theta})$ depends on B only via equation (A1). It is useful to explicitly remove B from the parameter set $\boldsymbol{\theta}$ by defining

$$\boldsymbol{\theta} = \{\boldsymbol{\phi}, B\}, \tag{A2}$$

i.e. $\boldsymbol{\phi}$ is the set of all confounding parameters except B. Hence, we can define

$$F_{\rm th}(\boldsymbol{\theta}) = \frac{F_{\rm th}(\boldsymbol{\phi})}{B}, \tag{A3}$$

and, since the event rate is not dependent on the beam, it can be written as

$$\begin{aligned} R(\boldsymbol{\theta}) &= R(\boldsymbol{\phi}) \\ &= k(\boldsymbol{\phi}) \left(\frac{F}{F_0}\right)^{\alpha}. \end{aligned} \tag{A4}$$

The total integral C over all parameter space (cf. equation 10) can be calculated as

$$C = \int \mathrm{d}\boldsymbol{\phi} \int \mathrm{d}\Omega \int_{F_{\rm th}(\boldsymbol{\theta})}^{F_{\rm m}(\boldsymbol{\theta})} k(\boldsymbol{\phi}) \frac{-\alpha}{F_0} \left(\frac{F}{F_0}\right)^{\alpha-1} \mathrm{d}F. \tag{A5}$$

This can also be written as an integral over beam factor B, by defining the solid angle viewed at each value of B, $\Omega(B)$ (see James et al. 2018). Hence, C becomes

$$\begin{aligned} C = &\int k(\boldsymbol{\phi}) \mathrm{d}\boldsymbol{\phi} \int \Omega(B) \mathrm{d}B \int_{F_{\rm th}(\boldsymbol{\phi})/B}^{F_{\rm m}(\boldsymbol{\phi})/B} \\ &\times \frac{-\alpha}{F_0} \left(\frac{F}{F_0}\right)^{\alpha-1} \mathrm{d}F. \end{aligned} \tag{A6}$$

As with equation (10), the integration over F in equation (A6) can be performed explicitly:

$$\begin{aligned} C = &\int k(\boldsymbol{\phi}) \mathrm{d}\boldsymbol{\phi} \int \Omega(B) \mathrm{d}B \\ &\times \left[\left(\frac{F_{\rm th}(\boldsymbol{\phi})}{BF_0}\right)^{\alpha} - \left(\frac{F_{\rm m}(\boldsymbol{\phi})}{BF_0}\right)^{\alpha} \right]. \end{aligned} \tag{A7}$$

Changing the definitions of relative minimum and maximum fluences s and b in equation (12) to be relative to observed values:

$$s = B\frac{F}{F_{\rm th}(\boldsymbol{\phi})},$$
$$b = B\frac{F}{F_{\rm max}(\boldsymbol{\phi})}, \tag{A8}$$

and inserting these into equation (A7) with $\mathrm{d}s = BF_{\rm th}^{-1}\mathrm{d}F$, produces

$$C = \int k(\boldsymbol{\phi})\mathrm{d}\boldsymbol{\phi}\int \Omega(B)\mathrm{d}B \times \int_{s=1}^{b}\frac{-\alpha}{F_0}\left(\frac{sF_{\rm th}}{BF_0}\right)^{\alpha-1}\frac{F_{\rm th}}{B}\mathrm{d}s. \tag{A9}$$

What we wish to calculate is the normalized probability $P(s)$, since neither B nor $\boldsymbol{\phi}$ is observable. This is defined as

$$P(s) = \int \mathrm{d}\phi\int \mathrm{d}B\,P(s, B, \phi), \tag{A10}$$

where the joint probability $P(s, B, \phi)$ is simply the normalized negative of the integrand in equation (A9):

$$P(s, B, \phi) = \frac{-1}{C}k(\boldsymbol{\phi})\Omega(B)\alpha s^{\alpha-1}\left(\frac{F_{\rm th}}{BF_0}\right)^{\alpha}. \tag{A11}$$

Observe that the integrals over $\boldsymbol{\phi}$ and B can be separated out:

$$P(s) = \frac{-1}{C}\alpha s^{\alpha-1}F_0^{-\alpha} \times\left[\int k(\boldsymbol{\phi})F_{\rm th}^{\alpha}\mathrm{d}\phi\right]\left[\int \Omega(B)B^{-\alpha}\mathrm{d}B\right]. \tag{A12}$$

Returning to the constant C, performing the integration over s in equation (A9) produces

$$C = \int k(\boldsymbol{\phi})\mathrm{d}\boldsymbol{\phi}\int \Omega(B)\mathrm{d}B\left(\frac{F_{\rm th}}{BF_0}\right)^{\alpha}\left[1-b^{\alpha}\right]. \tag{A13}$$

Noting that $F_{\rm th}$ is a function of $\boldsymbol{\phi}$, but not B, and again assuming that b is independent of both parameters, C can be written as

$$C = \left[1-b^{\alpha}\right]F_0^{-\alpha}\int \Omega(B)B^{-\alpha}\mathrm{d}B\int k(\boldsymbol{\phi})F_{\rm th}^{\alpha}(\boldsymbol{\phi})\mathrm{d}\boldsymbol{\phi}. \tag{A14}$$

The integrals over B and $\boldsymbol{\phi}$ in equation (A14) for C are identical to, and cancel with, the same integrals in equation (A12). Removing also the common power of $F_0^{-\alpha}$ produces

$$P(s) = -\alpha\frac{s^{\alpha-1}}{1-b^{\alpha}}. \tag{A15}$$

This is identical to equation (16), with s and b defined in terms of observable (i.e. beam-affected) quantities. All results thus follow, i.e. the results of Crawford et al. (1970) outlined in Section 2 hold.

We also note that applying an artificial cut (or completeness limit) $F_{\rm c} > F_{\rm th}$ in the integrals over $\mathrm{d}\boldsymbol{\phi}$, and defining s and b relative to $F_{\rm c}$ rather than $F_{\rm th}$, produces an identical result. Therefore, the application of Crawford et al. (1970) by Macquart & Ekers (2018a) and Bhandari et al. (2018) to beam-affected FRB fluences above a completeness limit is valid.

This paper has been typeset from a TEX/LATEX file prepared by the author.