The Astrophysical Journal, 944:105 (22pp), 2023 February 10 https://doi.org/10.3847/1538-4357/acaf06

Inferring the Energy and Distance Distributions of Fast Radio Bursts Using the First CHIME/FRB Catalog

Kaitlyn Shin[1,2], Kiyoshi W. Masui[1,2], Mohit Bhardwaj[3,4], Tomas Cassanelli[5], Pragya Chawla[6], Matt Dobbs[3,4], Fengqiu Adam Dong[7], Emmanuel Fonseca[8,9], B. M. Gaensler[10,11], Antonio Herrera-Martín[11,12], Jane Kaczmarek[13], Victoria Kaspi[3,4], Calvin Leung[1,2], Marcus Merryfield[3,4], Daniele Michilli[1,2], Moritz Münchmeyer[14], Aaron B. Pearlman[3,4], Masoud Rafiei-Ravandi[3,4], Kendrick Smith[15], Ingrid Stairs[7], and Shriharsh P. Tendulkar[16,17]
[1] MIT Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA, 02139, USA
kshin@mit.edu
[2] Department of Physics, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA, 02139, USA
[3] McGill Space Institute, McGill University, 3550 rue University, Montréal, QC, H3A 2A7, Canada
[4] Department of Physics, McGill University, 3600 rue University, Montréal, QC, H3A 2T8, Canada
[5] Department of Electrical Engineering, Universidad de Chile, Av. Tupper 2007, Santiago, 8370451, Chile
[6] Anton Pannekoek Institute for Astronomy, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands
[7] Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC, V6T 1Z1, Canada
[8] Department of Physics and Astronomy, West Virginia University, P.O. Box 6315, Morgantown, WV, 26506, USA
[9] Center for Gravitational Waves and Cosmology, West Virginia University, Chestnut Ridge Research Building, Morgantown, WV, 26505, USA
[10] Dunlap Institute for Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON, M5S 3H4, Canada
[11] David A. Dunlap Department of Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON, M5S 3H4, Canada
[12] Department of Statistical Sciences, University of Toronto, 700 University Ave., Toronto, ON, M5G 1Z5, Canada
[13] Dominion Radio Astrophysical Observatory, Herzberg Research Centre for Astronomy and Astrophysics, National Research Council Canada, P.O. Box 248, Penticton, BC, V2A 6J9, Canada
[14] Department of Physics, University of Wisconsin-Madison, 1150 University Ave, Madison, WI, 53706, USA
[15] Perimeter Institute of Theoretical Physics, 31 Caroline Street North, Waterloo, ON, N2L 2Y5, Canada
[16] Department of Astronomy and Astrophysics, Tata Institute of Fundamental Research, Mumbai, 400005, India
[17] National Centre for Radio Astrophysics, Post Bag 3, Ganeshkhind, Pune, 411007, India
Received 2022 July 28; revised 2022 December 19; accepted 2022 December 24; published 2023 February 15

Abstract

Fast radio bursts (FRBs) are brief, energetic, typically extragalactic flashes of radio emission whose progenitors are largely unknown. Although studying the FRB population is essential for understanding how these astrophysical phenomena occur, such studies have been difficult to conduct without large numbers of FRBs and characterizable observational biases. Using the recently released catalog of 536 FRBs published by the Canadian Hydrogen Intensity Mapping Experiment/Fast Radio Burst (CHIME/FRB) collaboration, we present a study of the FRB population that also calibrates for selection effects. Assuming a Schechter function, we infer a characteristic energy cut-off of $E_{\rm char} = 2.38^{+5.35}_{-1.64} \times 10^{41}$ erg and a differential power-law index of $\gamma = -1.3^{+0.7}_{-0.4}$. Simultaneously, we infer a volumetric rate of $[7.3^{+8.8}_{-3.8}(\text{stat.})^{+2.0}_{-1.8}(\text{sys.})] \times 10^4\ \text{Gpc}^{-3}\ \text{yr}^{-1}$ above a pivot energy of 10^{39} erg and below a scattering timescale of 10 ms at 600 MHz, and find we cannot significantly constrain the cosmic evolution of the FRB population with star-formation rate. Modeling the host's dispersion measure (DM) contribution as a log-normal distribution and assuming a total Galactic contribution of 80 pc cm^{-3}, we find a median value of $\text{DM}_{\rm host} = 84^{+69}_{-49}$ pc cm^{-3}, comparable with values typically used in the literature. Proposed models for FRB progenitors should be consistent with the energetics and abundances of the full FRB population predicted by our results. Finally, we infer the redshift distribution of FRBs detected with CHIME, which will be tested with the localizations and redshifts enabled by the upcoming CHIME/FRB Outriggers project.

Unified Astronomy Thesaurus concepts: Radio transient sources (2008); Radio bursts (1339)

1. Introduction

Ever since fast radio bursts (FRBs) were first discovered (Lorimer et al. 2007; Thornton et al. 2013), their mysterious origin has been an open question with no simple resolution. Studying FRBs as a population can provide valuable insight into their nature, as well as their utility for cosmological applications (e.g., McQuinn 2014; Zheng et al. 2014; Zhou et al. 2014; Masui & Sigurdson 2015; Akahori et al. 2016; Madhavacheril et al. 2019). A natural way to study FRBs is through inferring their luminosity function, which can tell us about observables such as volumetric rate and maximum FRB energy (Luo et al. 2020). The inferred parameters can be compared with those of other known astrophysical events to contextualize the question of FRB origins better. However, populations studies of FRBs are challenging—as a population, FRBs have been observed in a variety of observational contexts; it is yet unknown how many distinct populations there may be, and whether there is a dominant progenitor channel (James et al. 2022a). For example, some FRBs have been observed to repeat, while many more have only been observed as apparent one-off bursts. Indeed, recent studies have put forth the possibility, on the basis of observational properties such as pulse widths and observed frequency bandwidths, that repeaters and non-repeaters may comprise two distinct populations of FRBs (e.g., Cui et al. 2021; Pleunis et al. 2021).

Furthermore, FRBs have been found to be associated with a variety of types of galaxies, and in some cases, localized to Galactic regions with $\lesssim$arcsecond precision. Repeaters have been shown to originate both within star-forming regions (e.g., Marcote et al. 2020; Nimmo et al. 2022) and offset by $\sim$250 pc from the closest knot of local star formation (e.g., Tendulkar et al. 2021). By contrast, the repeating FRB 20200120E was localized to a globular cluster associated with M81 (Bhardwaj et al. 2021; Kirsten et al. 2022), surprising given that globular clusters host a much older stellar population than do star-forming regions. FRB-like bursts have also been observed from the Galactic magnetar SGR 1935+2154 (Bochenek et al. 2020; CHIME/FRB Collaboration et al. 2020b; Kirsten et al. 2021). It is possible that more such FRB-like events occur in the Milky Way, but remain undetected due to interstellar medium (ISM) scattering (Gohar & Flynn 2022). FRB localizations are not limited to repeating bursts; apparent one-offs have also been localized to host galaxies, and in some cases, the localizations are precise enough to be associated with regions of low star-formation rate as well (e.g., Bannister et al. 2019; Bhandari et al. 2020; Heintz et al. 2020; Li & Zhang 2020; Mannings et al. 2021). A comprehensive overview of the breadth of such localization results can be found in Petroff et al. (2022).

Nonetheless, there are many more FRBs that have been detected than just the $\sim$20 FRBs robustly associated with host galaxies. By 2020, there were over 100 verified FRBs[18] (Petroff et al. 2016), and as of writing, the number of verified FRB events is $\sim$800; a full catalog of FRBs is maintained at the Transient Name Server.[19] Based on these numbers alone, it is clear that studying FRBs as a population based only on the subsample of FRBs with precise localizations has strong limitations—that is, optical follow-up is time-intensive and expensive, and cannot currently be feasibly pursued for the hundreds of detected FRBs. There is also valuable information possessed by the many FRBs that do not have host galaxy associations, e.g., FRB energetics and abundances. Using a larger sample of FRBs adds valuable information, and should be very beneficial for statistically constraining how FRB energies are distributed. However, the central challenge to measuring the FRB luminosity function is, again, the lack of precise distance (localization) information for the majority of FRBs. Without a good handle on distance, the FRB energetics, distances, and volumetric abundances are all degenerate.

For FRBs, the dispersion measure (DM) is an (imperfect) distance proxy that allows us to break this modeling degeneracy. Since we expect DM to correlate with distance, while brightness will anticorrelate, any correlation/anticorrelation in modeling between DM and brightness should be a distance effect, thus providing a distance scale. As the uncertainties when modeling FRB energetics, distances, and abundances all affect each other, it is important to work with these properties simultaneously to minimize bias. Luo et al. (2020) and James et al. (2022b) have recently implemented this approach to studying the FRB population.

Previously, in order to conduct large-number statistical studies of the FRB population, it was necessary to combine burst detections from multiple surveys or telescopes such as the Australian Square Kilometre Array Pathfinder (ASKAP) or the Parkes radio telescope (e.g., Luo et al. 2020; James et al. 2022b). Using a heterogeneously observed sample of FRBs introduces differing, nonuniform selection effects that are difficult to calibrate. This motivates the use of FRBs observed from a single observing campaign to study the FRB population as a whole. Recently, the Canadian Hydrogen Intensity Mapping Experiment/Fast Radio Burst (CHIME/FRB) collaboration released a catalog of 536 FRBs, the largest sample of bursts detected thus far in a single survey, and is hereafter referred to as "Catalog 1" (CHIME/FRB Collaboration et al. 2021). With Catalog 1, we now possess the ability to conduct a statistical study with uniform selection effects and probe the intrinsic FRB population with greater numbers and logistical simplicity. Crucially, the selection effects are carefully characterizable with the use of the CHIME/FRB injections system (Merryfield et al. 2022). With 536 bursts, the Catalog 1 sample dominates the currently available sample of observed FRBs.

In this paper, we present a study of FRB population parameters, including the intrinsic luminosity function, derived from the sample of Catalog 1 bursts. We fit the joint brightness–DM model developed by James et al. (2022b), and use formalism that accounts for selection effects detailed by CHIME/FRB Collaboration et al. (2021). The outline of this paper is as follows. In Section 2, we describe the data sets used for this analysis. Section 3 outlines key components of the model we are fitting to the Catalog 1 data set, using the methodology described in Section 4. Results are presented in Section 5 and discussed in Section 6. Implications of our results for the CHIME/FRB Outriggers project are presented in Section 7, and we conclude in Section 8. We adopt Planck cosmological parameters (Planck Collaboration et al. 2020) throughout this analysis.

2. Data Set

The data used in this work consist of observed bursts from the CHIME/FRB Catalog 1 sample, corrected for selection effects in order for our analyses to reflect accurately the properties of the intrinsic population of FRBs as much as possible. The following subsections will detail the data sets used for this work.

2.1. CHIME/FRB Catalog 1 Observations

CHIME/FRB is a collaboration that uses the Canadian Hydrogen Intensity Mapping Experiment (CHIME; The CHIME Collaboration et al. 2022), located at the Dominion Radio Astronomical Observatory (DRAO). The telescope consists of four 20 m $\times$ 100 m cylindrical paraboloid reflectors, each oriented N–S and populated with 256 equispaced dual-linear-polarization antennae. This telescope is sensitive to bursts in the frequency range 400–800 MHz. The CHIME/FRB instrument is designed such that it observes the entire sky at decl. $> -11^\circ$ as it transits above each day; such a field of view is larger than other radio surveys or telescopes such as ASKAP or Parkes. Thus, CHIME/FRB is uniquely equipped to observe a unparalleled number of FRBs. Further technical details of the CHIME/FRB system can be found by CHIME/FRB Collaboration et al. (2018).

The Catalog 1 sample released by CHIME/FRB contains 536 bursts observed between from 2018 July 25 to 2019 July 1.[20] In the work presented by CHIME/FRB Collaboration

[18] https://www.frbcat.org
[19] https://wis-tns.weizmann.ac.il

[20] This sample has been made publicly available at https://chime-frb.ca/catalog.

et al. (2021), the population analysis of Catalog 1 bursts considered six observed FRB properties measured for each burst: fluence (F), DM, scattering timescale (τ), pulse width (w), spectral index (denoted as γ by CHIME/FRB Collaboration et al. 2021, but denoted as α in this paper), and spectral running (r). These observed properties were modeled by a least-squares fitting routine called `fitburst`[21] which processes the intensity data of FRBs offline and fits the 2D dynamic spectrum with a 2D analytic model. For bursts with one or more components i, `fitburst` models the DM, burst time of arrival ($t_{\mathrm{arr},i}$), signal amplitude (A_i), temporal width (w_i), power-law spectral index (α_i), spectral running (r_i), and scattering timescale (τ). The mathematical definitions of these components can be found in Section 3.3 of CHIME/FRB Collaboration et al. (2021). The same DM and τ are assumed for all sub-bursts, and we also assume $\tau \propto \nu^{-4}$ (Lang 1971; Lorimer & Kramer 2012), with 600 MHz as the reference frequency for scattering. For the small minority of bursts with complex structure in Catalog 1, if there was high enough signal-to-noise ratio (S/N), `fitburst` was used to optimize the structure. Dimmer bursts may have been over-dedispersed, especially for the bursts with downward-drifting structure, but the effect is likely minimal (i.e., no more than a few DM units). For each FRB, `fitburst` was run twice: once with τ fixed to 0, and once allowing τ to vary. In order to choose the preferred model, an F-test[22] was used to compare the χ^2 values from both models, adopting a threshold of $p < 0.001$ to determine the significance of scattering. Electron density models of the Milky Way DM, NE2001 (Cordes & Lazio 2002) and YMW2016 (Yao et al. 2017), were used in combination with the DM value determined by `fitburst` to validate the extragalactic nature of each FRB. Further details about `fitburst` can be found in Section 3.3 of CHIME/FRB Collaboration et al. (2021).

One further thing to note about the CHIME/FRB observations is that the fluence measurements are uncertain estimates of the true fluence, limited primarily by burst localization uncertainty and, to a lesser extent, by beam model uncertainty. While these fluence measurements are biased low (Section 3.4 in CHIME/FRB Collaboration et al. 2021), our analysis is not affected by this systematic because we use the S/N as a more reliable observable proxy to study the fluence. Further discussion of our treatment of the fluence values is in Sections 2.3 and 3.1.

2.2. CHIME/FRB Catalog 1 Injections

In order to properly characterize and account for instrumental effects when inferring the intrinsic distributions of FRB properties, a synthetic signal injection system was designed and implemented (Merryfield et al. 2022). Synthetically generated pulses were injected into the software detection system in order to allow for the accounting of real-time detection effects such as the RFI environment. These pulses will be hereafter referred to as "injections."

There are some important properties to note for the population of synthetic bursts injected into the real-time detection pipeline, which are briefly noted here but elaborated on by CHIME/FRB Collaboration et al. (2021). Of the $N_{\mathrm{inj}} = 5 \times 10^6$ total injected FRBs assigned random locations on the sky, some locations were automatically discarded: locations below the horizon, where the band-averaged primary beam response was $<10^{-2}$, and where the response was $<10^{-3}$ for all 1024 synthesized beams. The fraction of surviving sky locations is $f_{\mathrm{sky}} = 0.0277$. This "forward-modeling" method, using the beam models to simulate f_{sky}, was employed to account for the beam response of the telescope on the sky. Each of these FRBs had properties drawn from initial probability density functions $P_{\mathrm{init}}(F)$, $P_{\mathrm{init}}(\mathrm{DM})$, $P_{\mathrm{init}}(\tau)$, $P_{\mathrm{init}}(w)$, and $P_{\mathrm{init}}(\alpha, r)$, designed to sample, fully and densely, the phase space of observed properties in the catalog. Of these 5×10^6 injected FRBs, a cut was applied to events with little chance of being detected, and 96,942 events were ultimately scheduled for injection. Due to system errors that affected an effectively random subset of injections, 84,697 bursts were successfully injected with an injection efficiency $\epsilon_{\mathrm{inj}} = 0.874$ in 2020 August, with 39,638 events detected and assigned a S/N by the CHIME/FRB detection pipeline. Full technical details on how the injections are detected by the real-time pipeline can be found in CHIME/FRB Collaboration et al. (2021) and Merryfield et al. (2022).

2.3. Determining the Fiducial Property Distributions

A primary goal of the initial Catalog 1 populations analysis was to fit a fiducial model to the intrinsic FRB property distributions after correcting for selection effects. Such a model would be an imperfect, but still reasonable, match to the data. However, although the Catalog 1 sample makes up the first large sample of FRBs observed in a single survey with uniform selection effects, a number of bursts are affected by certain selection effects that are unquantifiable (e.g., non-nominal telescope operation). If we want to characterize the population while measuring and compensating for selection effects, we must only use bursts with the most robustly known statistics. Therefore, bursts affected by the following criteria were excluded from both the Catalog 1 sample as well as from the corresponding injections campaign sample:

1. Events detected during pre-commissioning, during epochs of low sensitivity, or on days with software upgrades. These events were correspondingly excluded from the Catalog 1 survey duration time, Δt.
2. Far-sidelobe events, for which the primary beam response is poorly understood.
3. Events with a detected S/N < 12, to avoid human inspection error affecting the completeness of identifying low-S/N FRBs.
4. Events with DM $< 1.5 \max(\mathrm{DM}_{\mathrm{NE2001}}, \mathrm{DM}_{\mathrm{YMW16}})$, to avoid mistakenly classifying extragalactic bursts as Galactic or vice versa.
5. Events with DM < 100 pc cm^{-3}.
6. Highly scattered events ($\tau > 10$ ms at 600 MHz), which are poorly constrained observationally due to a strong selection bias against them from the CHIME instrument, which would dominate uncertainties.
7. Repeat bursts after the first-detected burst from any known repeating FRBs in the sample.

The first exclusion criterion (requiring S/N $\geqslant 12$) dominates the 271 total excluded bursts in the populations analysis

[21] The codebase for is not yet open-source but based on the model formalism used by Masui et al. (2015). A description of `fitburst` as developed and used by the CHIME/FRB collaboration will be presented by E. Fonseca et al. (2023, in preparation).

[22] F-tests are good for comparing the statistics of nested models, and the model for bursts with scattering can be interpreted as nested from the model for bursts without scattering.

conducted by CHIME/FRB Collaboration et al. (2021). All of the above exclusion criteria are used in this analysis as well; for more information about these cuts, refer to Sections 5 and 6.1 in CHIME/FRB Collaboration et al. (2021). One further exclusion criterion than was used by CHIME/FRB Collaboration et al. (2021) is used in this analysis. In addition to excluding events with DM < 100 pc cm^{-3}, in order to minimize the variance in our results involved with the Milky Way contribution to the DM, we also exclude events with $\mathrm{DM_{NE2001}} > 100$ pc cm^{-3}. This was done because in our model, we take a fixed value of $\overline{\mathrm{DM}}_{\mathrm{MW}}$ to be 80 pc cm^{-3}; further discussion on this can be found in Section 5.3.1. It is worth noting that this extra exclusion criteria brings the total Catalog 1 sample to 225 bursts, down from the sample of 265 bursts used by CHIME/FRB Collaboration et al. (2021). Thus, hereafter, when referring to the Catalog 1 sample used for populations studies in this analysis, we are referring to the 225 bursts that survived all of these cuts.

To elaborate on exclusion criterion #7, simply excising all known repeating bursts would not work as there is no way to know if an apparent one-off burst will later turn out to be a repeater. Choosing the highest S/N burst will bias the sample to higher fluences, and choosing "random" bursts of the repeat bursts also introduce a subtle bias as the detection trigger threshold is set to be more sensitive to possible repeat bursts, i.e., bursts detected in the same directions and DMs as previously detected FRBs. Thus, to minimize bias (especially with respect to trigger thresholds), we include only the first-detected burst from any known FRB source that otherwise passes other cuts (i.e., S/N > 12). As detailed by Pleunis et al. (2021), there are differences in burst widths and bandwidths for observed repeaters and apparent non-repeaters, raising the question of whether it is fair to assume identical detection biases for these bursts. However, applying all the exclusion criteria except for the exclusion criterion for repeaters excludes 306 bursts instead of 311. Thus, even though the selection biases may differ for repeaters and apparent non-repeaters, the difference is unlikely to affect this analysis strongly. As the sample of observed repeaters grows, further investigation into the differing selection biases for repeaters and apparent non-repeaters will be crucial; further discussion is in Section 6.6.

Using the response of the real-time detection pipeline to the injected bursts, in CHIME/FRB Collaboration et al. (2021) we were able to determine the selection functions for fluence, DM, scattering, and pulse width. Also, as detailed by CHIME/FRB Collaboration et al. (2021), the injections system does not have the ability to forward-model fluence measurement processes in a robust manner, and measurements of the observed fluence are currently limited in their certainty. Thus, the intrinsic fluence distribution was studied by using the detection S/N as a proxy. The key assumptions that go into using the S/N as a proxy for fluence are (1) the detection S/N is strongly correlated with intrinsic fluence, and (2) we can statistically and accurately forward-model how fluence maps to the detection S/N with the injections system. As detailed in Section 3.1, the same methodology is employed in this analysis. Additionally, although the spectral index and spectral running were two properties also assigned to each injected pulse and observed in each real burst, due to the correlated nature of those properties, a functional form was not fit to the distribution $P(\alpha, r)$.

Alongside the selection functions, the following selection-corrected fiducial distributions were determined: $P_{\mathrm{fid}}(F)$, $P_{\mathrm{fid}}(\mathrm{DM})$, $P_{\mathrm{fid}}(\tau)$, and $P_{\mathrm{fid}}(w)$. For each of these respective properties, these fiducial distributions provide a reasonable description to the Catalog 1 data. For details on their derivation, refer to Section 6.1 and Appendix C of CHIME/FRB Collaboration et al. (2021). Here we note that to obtain these distributions, a number of simplifying assumptions were made in the Catalog 1 analysis, notably that the F–DM distribution factorizes into a power law in F and a free function of DM. In this paper, we improve upon this F–DM relation and use a more realistic one following the model detailed by James et al. (2022b). Consistent with the framework we used to obtain the fiducial distributions in Catalog 1, we hold the distributions of τ and w to the obtained fiducial models $P_{\mathrm{fid}}(\tau)$ and $P_{\mathrm{fid}}(w)$ while fitting for a new joint distribution of F and DM here. For simplicity of notation, we define the parameter set $\boldsymbol{\Theta} = \{\tau, w\}$ such that $P(\boldsymbol{\Theta}) = P(\tau)P(w)$, and $\int d\boldsymbol{\Theta} P(\boldsymbol{\Theta}) = 1$.

3. Fluence–DM Distribution Model

The model detailed by James et al. (2022b) contains components that detail how both fluence (i.e., brightness) and DM depend on redshift (i.e., distance). Contained within this populations model is information about FRB energetics, how the FRB population may evolve over cosmic time, as well as the distance distribution as inferred from DM observations.

We define the joint *rate* distribution of F and DM as

$$R(F, \mathrm{DM}|\boldsymbol{\lambda}) = \int dz R(F, z|\boldsymbol{\lambda}_1) P(\mathrm{DM}|z, \boldsymbol{\lambda}_2), \tag{1}$$

where the parameters $\boldsymbol{\lambda}$ for the fluence DM distribution considered will be detailed shortly. Inside the integrand, we can see two specific components: the joint rate distribution of fluence and redshift, and the probability distribution of DM given redshift. Appendix A details how to derive $R(F, \mathrm{DM}|\boldsymbol{\lambda})$, closely following James et al. (2022b). The key components of the model are as follows.

We model the FRB energy distribution with a Schechter (1976) function

$$P(E)dE \propto \frac{1}{E_{\mathrm{char}}}\left(\frac{E}{E_{\mathrm{char}}}\right)^{\gamma} \exp\left[-\frac{E}{E_{\mathrm{char}}}\right] dE, \tag{2}$$

where E_{char} is the characteristic exponential cutoff energy and γ is the differential power-law index. Such a form for the energy/luminosity distribution of FRBs has also been considered by previous works (e.g., Fialkov et al. 2018; Luo et al. 2018, 2020; Niu et al. 2021b). Note that we are assuming $P(E)$ does not evolve with redshift.

We also model the evolution of the FRB population by smoothly scaling the star formation rate (SFR) with a power-law index n,

$$\Phi(z) = \frac{\Phi_0}{1+z}\left(\frac{\mathrm{SFR}(z)}{\mathrm{SFR}(0)}\right)^{n}, \tag{3}$$

where $\Phi(z)$ represents the rate of FRBs per comoving volume above a fixed pivot energy E_{pivot}, and SFR (z) (Equation (A9)) is from Madau & Dickinson (2014). A pivot energy is required for us to be able to quote a rate of FRBs; we choose a pivot energy of 10^{39} erg under the assumption of spectral bandwidth at 1 GHz (see also Appendix A.1). This pivot energy is above the threshold of ruled out minimum FRB energies (i.e., James et al. 2022b rules out $E_{\mathrm{min}} > 10^{38.5}$ erg at 90% C.L.). The

parameter Φ_0 is the volumetric rate in units $\mathrm{Gpc}^{-3}\ \mathrm{yr}^{-1}$ at $z = 0$. The population redshift evolution with the star formation history (SFH) is parameterized by n, where $n = 0$ would imply no evolution with cosmic SFH and $n = 1$ would imply evolution that perfectly traces the cosmic SFH. This model also allows for the possibility of FRB progenitors that evolve faster than the SFH of the universe (e.g., $n > 1$). Thus, the free parameters in $\boldsymbol{\lambda}_1$ are Φ_0, $E_{\rm char}$, γ, the spectral index α, and n. It is worth noting that α comes into the model in the conversion from E to F in Equation (A4), such that this model assumes the intrinsic α is the same for every FRB. The "rate interpretation" of α, further discussed in Section 5.4, Appendix B, and by James et al. (2022b), takes the other extreme where every FRB has a unique spectral index. As FRBs have been observed to have a wide range of spectral properties, the treatment of α is a systematic uncertainty that is model-dependent.

The DM distribution accounts for contributions from the disk and halo of the Milky Way, cosmological contributions through the intergalactic medium (IGM; refer to Equation (A12) in Appendix A.2), and contributions from the host galaxy of the FRB. The Milky Way contribution to the DM is fixed to be $\overline{\mathrm{DM}}_{\rm MW} = 80\ \mathrm{pc\ cm}^{-3}$ (discussed in Section 5.3.1), and the host DM contribution is modeled with a log-normal distribution

$$P(\mathrm{DM}'_{\rm host}) = \frac{1}{\mathrm{DM}'_{\rm host}} \frac{1}{\sigma_{\rm host}\sqrt{2\pi}} e^{-\frac{(\log \mathrm{DM}'_{\rm host} - \mu_{\rm host})^2}{2\sigma^2_{\rm host}}}, \quad (4)$$

where the host DM contribution is parameterized by $\mu_{\rm host}$ and $\sigma_{\rm host}$ and corrected for redshift with $\mathrm{DM}_{\rm host} = \mathrm{DM}'_{\rm host}/(1+z)$. (Note, the parameters as presented in the exponential of Equation (4) are in natural log space.) One thing that is not considered in this model is the fact that the ionized gas content of a galaxy and galaxy ensembles evolves over the cosmic SFH. Thus, the DM distribution would also evolve with redshift such that the observed DM distribution differs from the rest-frame one. While a simple scaling relation could be applied to model the cosmic evolution of $\mathrm{DM}_{\rm host}$ (e.g., Luo et al. 2018), we choose not to in order to minimize the complexity of the model while keeping the physical parameters necessary to encapsulate the dependence of fluence and DM on distance. Thus, the free parameters in $\boldsymbol{\lambda}_2$ are $\mu_{\rm host}$ and $\sigma_{\rm host}$.

Note, $\mathrm{DM}'_{\rm host}$ is the log-normal variate parameterized by $\mu_{\rm host}$ and $\sigma_{\rm host}$. Thus, $\mu_{\rm host}$ and $\sigma_{\rm host}$ are the expected value and standard deviation of $\log(\mathrm{DM}'_{\rm host})$, not of $\mathrm{DM}'_{\rm host}$ itself. The median and standard deviation of $\mathrm{DM}'_{\rm host}$ can be obtained with the parameters $\mu_{\rm host}$ and $\sigma_{\rm host}$ as follows:[23]

$$\mathrm{Median} P(\mathrm{DM}'_{\rm host}) = \exp(\mu_{\rm host}), \quad (5)$$

$$\text{Std. dev. } P(\mathrm{DM}'_{\rm host}) = \sqrt{(\exp(\sigma^2_{\rm host}) - 1)\exp(2\mu_{\rm host} + \sigma^2_{\rm host})}. \quad (6)$$

Throughout this work, to avoid confusion, we consider the log-normal variate $\mathrm{DM}'_{\rm host}$ so as to keep all values in physical units of $\mathrm{pc\ cm}^{-3}$. We also quote the median of $P(\mathrm{DM}'_{\rm host})$ rather than the mean, as it is a log-normal distribution that we expect to be skewed by the existence of atypical, but plausible, sources with high host DM contributions, and the median is more robust against outliers in a skewed distribution.

[23] https://sites.google.com/site/probonto (Swat et al. 2016).

3.1. Subtleties Involved with the CHIME/FRB Fluence

The model we aim to fit to the Catalog 1 data is a joint rate distribution of F and DM, $R(F, \mathrm{DM}|\boldsymbol{\lambda})$. However, as noted in Sections 2.1 and 2.3, the observed fluence measurements are uncertain estimates of the true fluence. While the synthetic bursts are assigned a "true" fluence value before injection, there is no good way to forward-model how their fluence values would be "measured" by the real-time detection pipeline. Thus, directly comparing the "true" fluence values of the injected bursts with the "observed" fluence values of the Catalog 1 bursts would be subject to systematics not yet fully characterized. The better alternative is to use the S/N as a proxy for the fluence when fitting for $R(F, \mathrm{DM}|\boldsymbol{\lambda})$, as the fluence is strongly correlated with the S/N with known calibration factors, with the beam response also taken into account (Merryfield et al. 2022).

Thus, we wish to make a prediction for the observed (as opposed to intrinsic) CHIME/FRB Catalog 1 DM–S/N distribution using a given fluence–DM model. The next section will detail how we fit for the S/N–DM distribution, and by extension, obtain the parameters $\boldsymbol{\lambda}$ in $R(F, \mathrm{DM})$.

4. Methodology

The observed CHIME/FRB Catalog 1 S/N–DM distribution is denoted with n_{ij}, where i indexes the S/N bin and j indexes the DM bin. We want to compare the observable n_{ij} to our model prediction ζ_{ij}, which depends on the parameters $\boldsymbol{\lambda} = \{\Phi_0, E_{\rm char}, \gamma, \alpha, n, \mu_{\rm host}, \sigma_{\rm host}\}$.

The candidate population model ζ_{ij} is parameter-dependent —and thus model-dependent—because we construct weights to convert the original single injected population model to any other population model. The alternative approach would have been to conduct an injections campaign for every candidate population model we wished to test, which is logistically infeasible (the full duration of the injections campaign used in this analysis was over a month). Whereas n_{ij} is an unweighted histogram of the observed Catalog 1 data in S/N and DM bins, ζ_{ij} is a weighted histogram of the synthetic data "detected" by the real-time pipeline. Each synthetic burst m in a given bin n_{ij} has a weight $W_m(F, \mathrm{DM}, \boldsymbol{\Theta}, \boldsymbol{\lambda})$ assigned to it, i.e.,

$$\zeta_{ij} = \sum_{m=1}^{n_{ij}} W_m(F, \mathrm{DM}, \boldsymbol{\Theta}, \boldsymbol{\lambda}). \quad (7)$$

Note also that the units of ζ_{ij} are counts per n_{ij} bin. If we divide each bin in ζ_{ij} by the bin area ($\Delta\mathrm{S/N}\,\Delta\mathrm{DM}$), then we would obtain an observed rate of bursts per S/N per DM, e.g., $R_{\rm obs}(\mathrm{S/N}, \mathrm{DM}|\lambda)$. This quantity is the prediction for the observed (as opposed to intrinsic) CHIME/FRB Catalog 1 DM–S/N distribution.

We employ Markov Chain Monte Carlo (MCMC) techniques to sample the likelihood while simultaneously fitting for all seven parameters in $R(F, \mathrm{DM}|\boldsymbol{\lambda})$. We use the logarithmic binned Poisson likelihood (Zyla et al. 2020)

$$\log \mathcal{L}(\boldsymbol{\lambda}) \propto \sum_{ij} n_{ij} \log(\zeta_{ij}) - \zeta_{ij}. \quad (8)$$

The likelihood was evaluated with 15 logarithmically spaced bins in S/N, ranging from 12 to 200, and 20 logarithmically spaced bins in DM, ranging from 100 to 3500 pc cm^{-3}.

The evaluation of the different candidate population models ζ_{ij} is possible through the construction of weights $W_m(F, \mathrm{DM}, \boldsymbol{\Theta}, \boldsymbol{\lambda})$. The formalism is as follows. Each injected event m has its own parameter-dependent weight $W_m(F, \mathrm{DM}, \boldsymbol{\Theta}, \boldsymbol{\lambda})$, which means each weight depends on the properties of each synthetic burst. These weights are constructed to be the ratio of how many bursts were injected to how many occurred on the sky during the survey for a given set of model parameters $\boldsymbol{\lambda}$. Thus, assuming the parameters $\boldsymbol{\lambda}$ are true, then for every injected burst with properties F and DM, we should have instead injected a number equal to $W_m(F, \mathrm{DM}, \boldsymbol{\Theta}, \boldsymbol{\lambda})$ in order to match the observed sky population exactly.

To obtain the property-dependent weights $W_m(F, \mathrm{DM}, \boldsymbol{\Theta}, \boldsymbol{\lambda})$, for each synthetic burst m with properties F and DM, using the model parameters $\boldsymbol{\Theta}$ and $\boldsymbol{\lambda}$, we divide the full injected population of bursts on the sky by the initial injected population of bursts

$$W(F, \mathrm{DM}, \boldsymbol{\Theta}, \boldsymbol{\lambda}) = \frac{R(F, \mathrm{DM}, \boldsymbol{\Theta}|\boldsymbol{\lambda})\Delta t}{R_{\mathrm{init}}(F, \mathrm{DM}, \boldsymbol{\Theta})}, \tag{9}$$

$$= \frac{\Delta t f_{\mathrm{sky}}}{N_{\mathrm{inj}}\epsilon_{\mathrm{inj}}} \frac{R(F, \mathrm{DM}, \boldsymbol{\Theta}|\boldsymbol{\lambda})}{P_{\mathrm{init}}(F, \mathrm{DM}, \boldsymbol{\Theta})}. \tag{10}$$

The quantity $R_{\mathrm{init}}(F, \mathrm{DM}, \boldsymbol{\Theta})$ is the full initial population of injected bursts, with units of bursts per F per DM per injections campaign. It is equivalent to the normalized initial population distribution of the injected synthetic bursts, $P_{\mathrm{init}}(F, \mathrm{DM}, \boldsymbol{\Theta})$, multiplied by the total number of FRBs injected (N_{inj}) corrected for the efficiency of injections (ϵ_{inj}) and the sky fraction (f_{sky}). That is

$$R_{\mathrm{init}}(F, \mathrm{DM}, \boldsymbol{\Theta}) = \frac{N_{\mathrm{inj}}\epsilon_{\mathrm{inj}}}{f_{\mathrm{sky}}} P_{\mathrm{init}}(F, \mathrm{DM}, \boldsymbol{\Theta}). \tag{11}$$

The model $R(F, \mathrm{DM}, \boldsymbol{\Theta}|\boldsymbol{\lambda})$ gives a rate in bursts per year. When multiplied by Δt, the survey duration in days ($\Delta t = 214.8$ days), we can then get a rate of observed bursts per survey, which can be directly compared against $R_{\mathrm{init}}(F, \mathrm{DM}, \boldsymbol{\Theta})$.

Also, as mentioned at the end of Section 2.3, the distributions of $\boldsymbol{\Theta}$ were held to the fiducial models found in Catalog 1. Therefore

$$W(F, \mathrm{DM}, \boldsymbol{\Theta}, \boldsymbol{\lambda}) = \frac{\Delta t f_{\mathrm{sky}}}{N_{\mathrm{inj}}\epsilon_{\mathrm{inj}}} \frac{R(F, \mathrm{DM}|\boldsymbol{\lambda})}{P_{\mathrm{init}}(F, \mathrm{DM})} \frac{P_{\mathrm{fid}}(\boldsymbol{\Theta})}{P_{\mathrm{init}}(\boldsymbol{\Theta})}, \tag{12}$$

where the fiducial distributions and initial distributions are all factorizable.

Each synthetic burst can have an associated weight because each burst has "intrinsic" F and DM information associated with it. Because each burst also has "observed" S/N information, each weighted burst can be assigned a bin in ζ_{ij}, thus allowing for ζ_{ij} to be model-dependent (i.e., dependent on $\boldsymbol{\lambda}$) and directly comparable to n_{ij} in the likelihood function (Equation (8)).

Note that the equations in this section are simply a more explicit definition of the weights also defined by CHIME/FRB Collaboration et al. (2021), wherein the weights were used only for obtaining the model parameters for $P_{\mathrm{fid}}(F)$, $P_{\mathrm{fid}}(\mathrm{DM})$ and $P_{\mathrm{fid}}(\boldsymbol{\Theta})$. For these distributions, the prefactors used to determine accurately the rate ($\Delta t f_{\mathrm{sky}}/N_{\mathrm{inj}}\epsilon_{\mathrm{inj}}$) do not matter.

5. Results

Using the methodology described by CHIME/FRB Collaboration et al. (2021) and elaborated on in Section 4, we use the injections system to calibrate out selection effects while fitting this full population model, $R(F, \mathrm{DM}|\boldsymbol{\lambda})$, to the Catalog 1 data. We use the `emcee` package (Foreman-Mackey et al. 2013) to generate MCMC samples of the posterior, using `scipy.optimize.curve_fit` to obtain the initial parameters for the MCMC chains. In order to ensure the posterior distribution is proportional to the likelihood in Equation (8), all of our parameters are drawn from uniform priors. We then run the MCMC sampler with 40 walkers, 2000 burn-in steps, and then 25,000 steps after that. While there is no simple way to ensure convergence, the first and second halves of the chains show similar posterior distribution morphologies to each other, heuristically indicating confidence our chains have converged (Hogg & Foreman-Mackey 2018). Additionally, the chains are longer than 75 times the integrated autocorrelation time for all parameters, with an acceptance fraction of ≈0.26—within the range of recommended values (Foreman-Mackey et al. 2013). The results can be seen in Figure 1, generated using the `corner` package (Foreman-Mackey 2016).[24] We quote the median of the posterior distributions, along with the 16% quantile as the lower 1σ error bar and the 84% quantile as the upper 1σ error bar, as generally recommended by Hogg & Foreman-Mackey (2018). These values are tabulated in Table 1.

In order to verify the robustness of the MCMC sampling, we also redo our analysis on "mock" data. In order to get a mock CHIME/FRB Catalog 1 data set, we took the full initial population of injected FRBs and resampled them with weights according to the best-fit model parameter values. We then ran the MCMC sampler with this mock Catalog 1 data and recovered all the best-fit model parameters to within 2σ (most parameters to within 1σ).

Figure 2 shows slices through the S/N–DM distribution of the observed Catalog 1 data n_{ij} compared against the best-fit model prediction from the MCMC run, ζ_{ij}, in each parameter space.

In both the low- and high-DM slices, the observed S/N distribution peaks at lower values closer to the S/N cutoff threshold, with negligibly few sources at the high-end of the S/N distribution. Additionally, when viewing the data in S/N slices, there are more sources in the lower S/N ranges than in the higher ones, with the DM distribution appearing to peak at ∼500 pc cm^{-3} for the slice most representative of the Catalog 1 data (the lowest S/N sample). Note the rarity of detecting exceptionally high S/N events at any DM range. In each view of the slices through the data, the best-fit model appears to be consistent with the observed data.

In Figure 3, one can see a side-by-side comparison of 2D views of the observed S/N–DM distribution and model prediction. Qualitatively, it is clear that the data appear to be consistent with the model and added Poissonian noise. Both the observed data and the model prediction show that the majority

[24] Data products are also made available at https://github.com/kaitshin/CHIMEFRB-Cat1-Energy-Dist-Distrs.

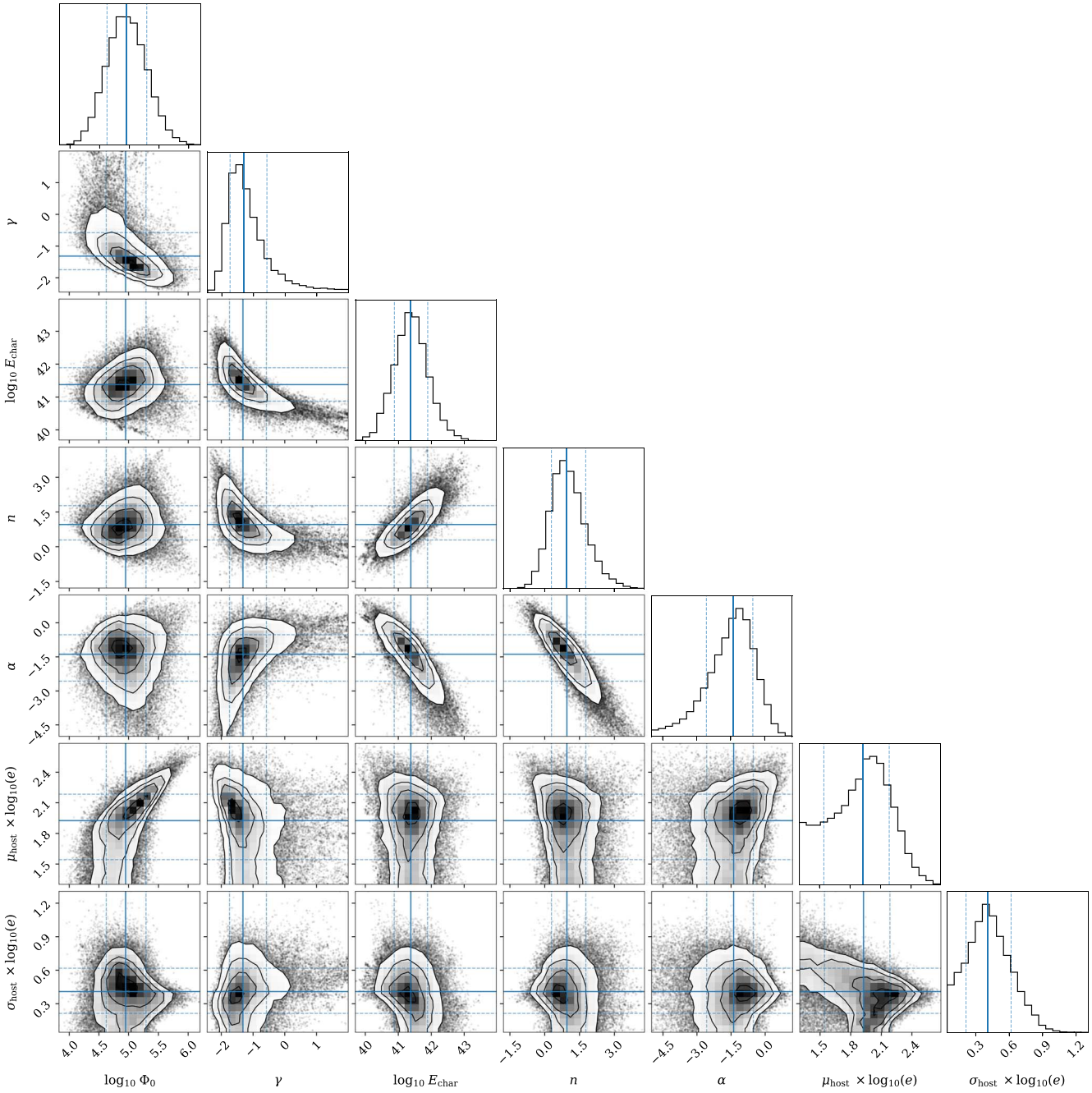

Figure 1. Corner plot of the results of the MCMC run, thinned by a factor of 20 for visualization purposes. Overlaid solid blue lines denote the medians of the posterior distributions, with the dashed blue lines enclosing the central 68% of the samples.

of FRBs are detected at lower S/N ranges, peaking around DM values of hundreds of pc cm^{-3}.

5.1. Goodness-of-fit

One advantage of using maximum likelihood methods with binned data, as was mentioned in Section 4, is that a goodness-of-fit statistic can be simultaneously obtained (Zyla et al. 2020). Maximizing the likelihood $\mathcal{L}(\boldsymbol{\lambda})$ is equivalent to maximizing the likelihood ratio $\xi(\boldsymbol{\lambda})$, and thus is also equivalent to minimizing the quantity $-2\log\xi(\boldsymbol{\lambda})$. We define the quantity $T = -2\log\xi(\boldsymbol{\lambda})$. For Poissonian data, this expression takes the form

$$T = 2\sum_{ij}\left[\zeta_{ij} - n_{ij} + n_{ij}\log\left(\frac{n_{ij}}{\zeta_{ij}}\right)\right]. \tag{13}$$

The smaller the value of T, the better the data n_{ij} agree with the model ζ_{ij}. If Wilks' theorem holds, then the minimum of T follows a χ^2 distribution (Wilks 1938), and the value can be interpreted as a p-value. However, one of the conditions of Wilks' theorem is that the sample size must be large, which our binned data do not satisfy. Thus, to quantify the goodness-of-fit, we aim to estimate the null distribution of T by Monte Carlo simulations. We generate 10,000 Monte Carlo samples of

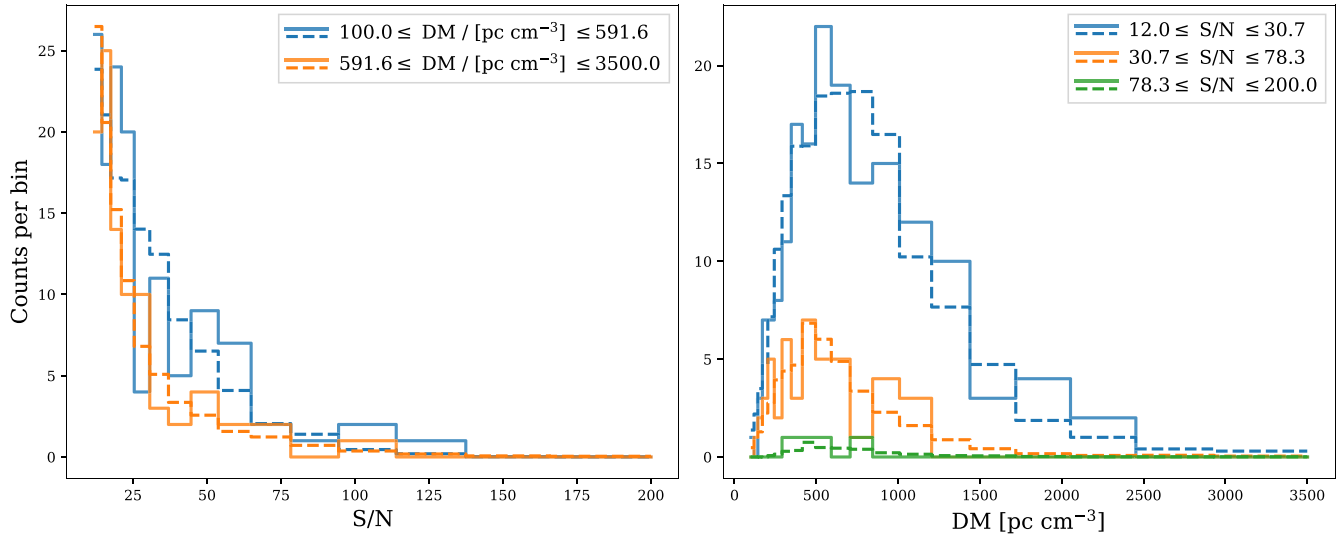

Figure 2. Observed distributions, denoted with solid lines, compared with the best-fit model predictions, denoted with dashed lines. Each sample range is determined by the logarithmic binning described in Section 4, and not by the number of events that fall in each bin. Left panel: the S/N distribution is shown for a low-DM sample and a high-DM sample. Each shown DM range spans ten bins. Right panel: the DM distribution is shown in three S/N samples. Each shown S/N range spans five bins. The best-fit model predictions use the median of the posterior distributions of each parameter after the MCMC run. In every viewed slice, the model appears consistent with the observed data.

Table 1
Table of Results of the Parameters Fitted in $R(F, \mathrm{DM})$ (see Equation (1))

Parameter	Uniform Prior Range	Best-fit Result
$\log_{10}\Phi_0$[a]	[−0.96, 6.43]	$4.86^{+0.34}_{-0.33}$
Φ_0	…	$7.3^{+8.8}_{-3.8}\times10^4$ Gpc^{-3} yr^{-1}
γ	[−2.50, 2.00]	$-1.3^{+0.7}_{-0.4}$
$\log_{10} E_{\rm char}$[a]	[38.00, 49.00]	$41.38^{+0.51}_{-0.50}$
$E_{\rm char}$	…	$2.38^{+5.35}_{-1.64}\times10^{41}$ erg
n	[−2.00, 8.00]	$0.96^{+0.81}_{-0.67}$
α	[−5.00, 5.00]	$-1.39^{+0.86}_{-1.19}$
$\mu_{\rm host}\times\log_{10}(e)$[a]	[1.30, 2.70]	$1.93^{+0.26}_{-0.38}$
Median $P(\mathrm{DM}')$	…	84^{+69}_{-49} pc cm^{-3}
$\sigma_{\rm host}\times\log_{10}(e)$[a]	[0.04, 1.74]	$0.41^{+0.21}_{-0.20}$
Std.dev. $P(\mathrm{DM}')$	…	174^{+319}_{-128} pc cm^{-3}

Notes. The best-fit results quote the median values of the posterior distributions, with the error bars containing the central 68% of the samples. Where illustrative, fits are paired with the corresponding derived physically meaningful quantities.
[a] These parameters were fitted in natural log space but are presented in $\log_{10}$ space for ease of interpretation.

Poissonian-distributed n_{ij} from our ζ_{ij} and calculate the test statistic T. The results of the Monte Carlo simulations are in Figure 4. Based on this statistic, for over 100 runs, ≈25% of the simulations have a worse fit than our best-fit model prediction. In other words, we have estimated a p-value of 0.25 where in 10,000 different realizations of the data, the model $R(F, \mathrm{DM}|\boldsymbol{\lambda})$ was a worse fit. To reiterate, this p-value represents the fraction of data realizations drawn from the best-fit model that are a better fit to the data. Therefore, a p-value of 0.25 indicates that the model is neither an anomalously bad nor good fit (i.e., overfit) to the data. Thus, this estimated p-value is consistent with the conclusion that our model matches the data reasonably well, albeit with room for improvement in some combination of the model details, the collected data, and our treatment of observed selection effects.

5.2. R(F, z) Parameters

The parameters in $\boldsymbol{\lambda}_1$ are Φ_0, $E_{\rm char}$, γ, α, and n. Contained within this component of the fit model are predictions about the energetics and abundances of FRBs based on the CHIME/FRB Catalog 1 sample.

We find a volumetric rate at $z=0$ above $E_{\rm pivot}=10^{39}$ erg, with our statistical uncertainties given by the MCMC. In the work presented by CHIME/FRB Collaboration et al. (2021), the net systematic uncertainty for the sky rate was $+27\%/-25\%$; we assume the systematics propagate forward in a similar manner for the volumetric rate derived here. Thus, the volumetric rate is $\Phi_0=[7.3^{+8.8}_{-3.8}(\mathrm{stat.})^{+2.0}_{-1.8}(\mathrm{sys.})]\times10^4$ Gpc^{-3} yr^{-1}. In CHIME/FRB Collaboration et al. (2021), the systematic errors are by far the largest for factors that affect the overall rate. Given that in this analysis, the systematic errors are subdominant to the statistical errors for the rate calculation, we assume that the remainder of the parameters are dominated by statistical errors as well. Thus, for all other parameters we quote only statistical errors on our values.

The energy distribution within $R(F, z)$ is modeled with a Schechter function (Equation (2)), and we find best-fitting values of $E_{\rm char}=2.38^{+5.35}_{-1.64}\times10^{41}$ erg and $\gamma=-1.3^{+0.7}_{-0.4}$, where γ is the differential power-law index.

Our best-fitting value for the spectral index in $F\propto\nu^{\alpha}$ is $\alpha=-1.39^{+0.86}_{-1.19}$. However, one can also interpret α as a frequency-dependent rate; see Section 5.4 for details of this interpretation. For the parameter n which parameterizes evolution with the cosmic SFH, we find a best-fitting value of $n=0.96^{+0.81}_{-0.67}$. As can be seen in Figure 1, the parameters α and n are strongly anticorrelated—intuitively, the more strongly FRBs appear to evolve with (or faster than) the cosmic SFH (higher n), the more sources would originate from further away and appear fainter, thus leading to a steeper observed "true spectral index." This degeneracy is stronger under the "rate interpretation" of α. The parameter for the cutoff energy $E_{\rm char}$ is also correlated (anticorrelated) with n (α), as the more energetic we expect FRBs can be, the more FRBs we should be able to observe further away (fainter).

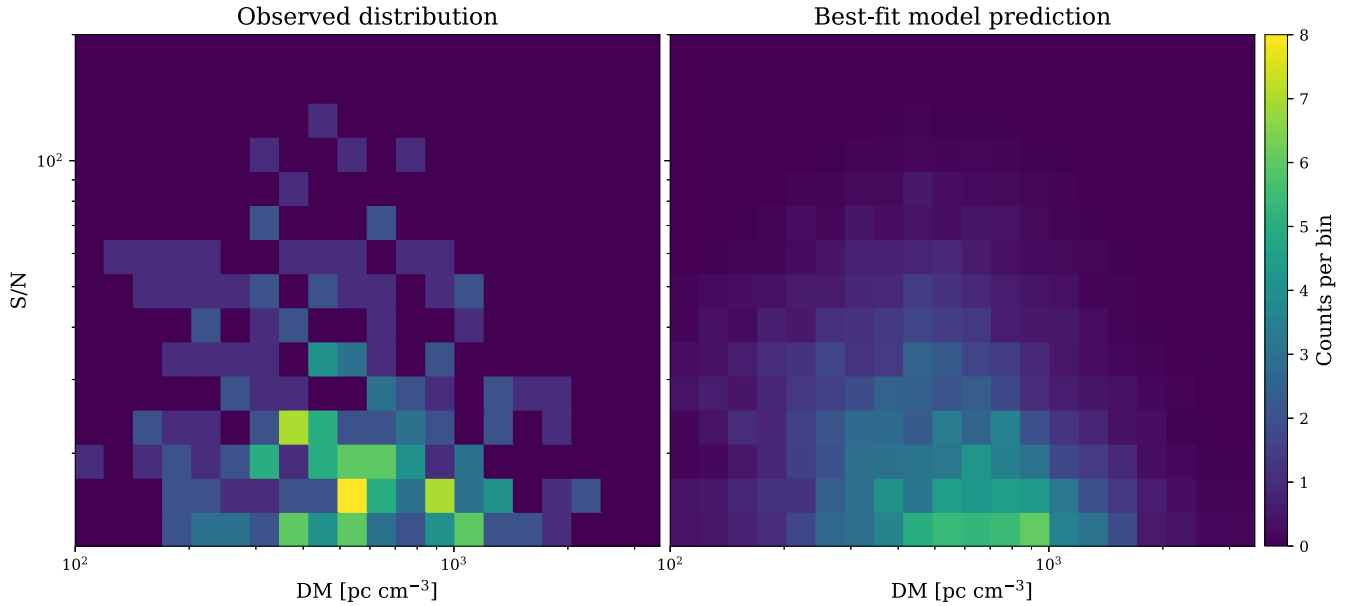

Figure 3. Comparison of the S/N–DM distribution between the observed data (left) and the best-fit model prediction (right). The best-fit model predictions use the median of the posterior distributions of each parameter after the MCMC run. The two appear to behave qualitatively similarly, in-line with expectations from Figure 2.

5.3. P(*DM*|z) Parameters

The parameters in $\boldsymbol{\lambda}_2$ are $\mu_{\rm host}$ and $\sigma_{\rm host}$. Modeling the host galaxy contribution of the FRB to its DM as a log-normal distribution, our best-fit values are $\mu_{\rm host} \times \log_{10}(e) = 1.93^{+0.26}_{-0.38}$ and $\sigma_{\rm host} \times \log_{10}(e) = 0.41^{+0.21}_{-0.20}$. These values correspond to a median $\rm DM'_{host}$ value of 84^{+69}_{-49} pc cm^{-3}, with a standard deviation of 174^{+319}_{-128} pc cm^{-3}.

As can be seen in Figure 1, the allowed parameter space for $\mu_{\rm host}$ and the volumetric rate Φ_0 have complex morphologies. To explain some of this morphology intuitively, $\mu_{\rm host}$ sets the overall distance scale for FRBs—the higher $\mu_{\rm host}$ is, the more we expect the DM contribution to be from the host galaxy rather than the IGM, and thus FRBs would be closer, less energetic, and more volumetrically abundant.

5.3.1. Varying Assumed $\overline{\rm DM}_{\rm MW}$

When fitting for $P(\rm DM|\it z)$, we adopted $\overline{\rm DM}_{\rm MW} = 80$ pc cm^{-3} as our fiducial value. However, the Galactic contribution to the DM of an FRB has non-negligible uncertainties, especially when considering the Galactic halo. Although we currently aim to minimize the uncertainty associated with the halo contribution by excluding sources with $\rm DM_{NE2001} > 100$ pc cm^{-3} (Section 2.3), a more sophisticated analysis would take into account updated models of the Milky Way halo DM (e.g., Cook et al. 2023). Nonetheless, we wish to test how sensitive our analysis may be to different values of $\overline{\rm DM}_{\rm MW} = 80$ pc cm^{-3}. Thus, we performed multiple more iterations of the MCMC analysis, varying $\overline{\rm DM}_{\rm MW}$ to between be 50 pc cm^{-3} and 100 pc cm^{-3} in increments of 10 pc cm^{-3}. We do these discrete iterations because, otherwise, $\overline{\rm DM}_{\rm MW}$ as a free parameter would be degenerate with $\mu_{\rm host}$ (see, e.g., Macquart et al. 2020). In fact, modifying the formalism to include $\overline{\rm DM}_{\rm MW}$ as a free parameter would likely have constraining power only after the sample of FRBs grows to reduce the already large statistical uncertainties already associated with $\mu_{\rm host}$ and $\sigma_{\rm host}$. Having $\rm DM_{MW}$ as a direction-dependent free parameter, and also having it be independent of $\mu_{\rm host}$, would require significant changes to the model formalism, and is beyond the scope of this analysis.

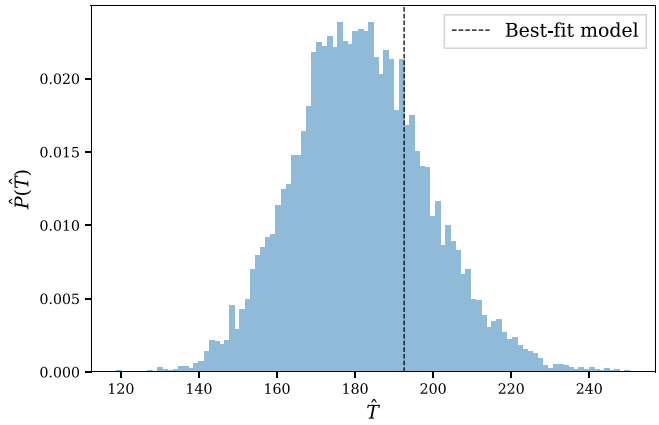

Figure 4. The estimated probability distribution $\hat{P}(\hat{T})$ of the test statistic T, defined in Equation (13), from Monte Carlo sampling. The value of this test statistic computed for our best-fit model prediction is overlaid in a vertical dashed line. Approximately 25% of $\hat{P}(\hat{T})$ lies to the right of the dashed line, giving an estimated p-value of 0.25.

Predictably, the varying of this $\overline{\rm DM}_{\rm MW}$ assumption most affects the posterior distributions of the parameters that go into the log-normal distribution model of $\rm DM_{host}$, e.g., $\boldsymbol{\lambda}_2 = \{\mu_{\rm host}, \sigma_{\rm host}\}$. None of the five other posterior distributions from the MCMC run appreciably changed. Shown in Figure 5 are the posterior distributions for the derived physically meaningful quantities corresponding to $\mu_{\rm host}$ and $\sigma_{\rm host}$—i.e., the median and the standard distribution of $P(\rm DM'_{host})$, which is parameterized by $\mu_{\rm host}$ and $\sigma_{\rm host}$—from the runs where $\overline{\rm DM}_{\rm MW}$ was varied.

Comparing Figure 5 to the results presented in Table 1, the variation of where the posterior distributions peak appears to be within the 1σ statistical uncertainties. Thus, while we know our assumption of $\overline{\rm DM}_{\rm MW}$ is imperfect, its effect on the modeling of $\rm DM_{host}$ is subdominant to statistical uncertainties, and it also does not appreciably affect the other fit parameters in $R(F, \rm DM|\boldsymbol{\lambda})$. As

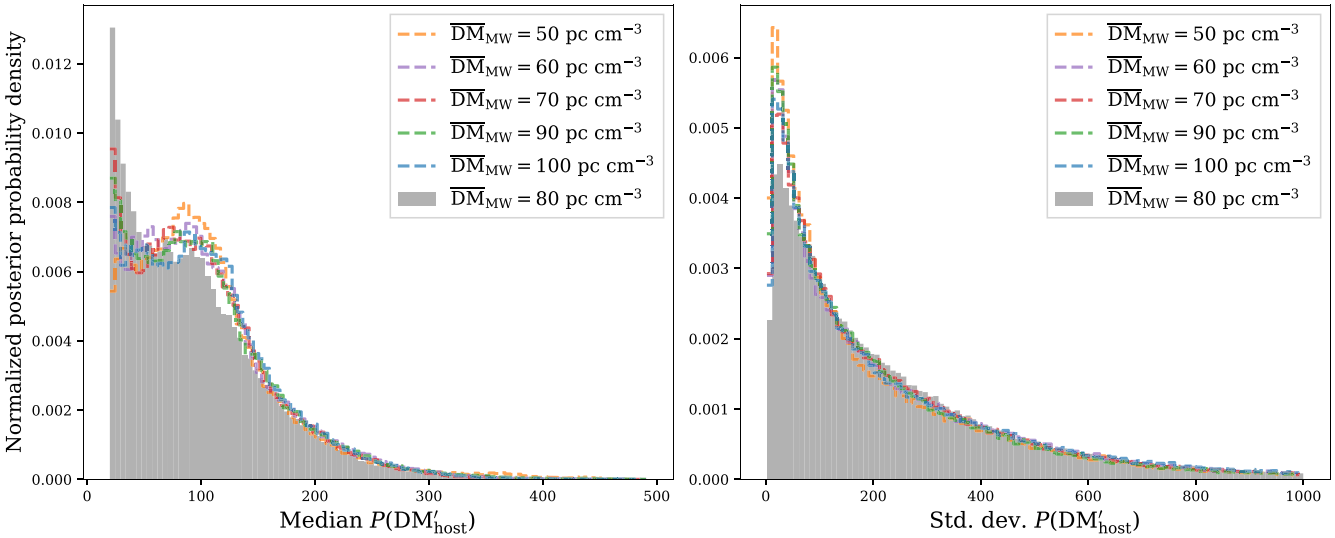

Figure 5. Overlaid posterior distributions from the MCMC runs run with $\overline{\mathrm{DM}}_{\mathrm{MW}} = 50$ pc cm^{-3} (orange dotted–dashed lines), 80 pc cm^{-3} (gray shaded region), and 100 pc cm^{-3} (blue dashed lines). On the left is the normalized posterior distribution for μ_{host} and on the right is the normalized posterior distribution for σ_{host}. When $\overline{\mathrm{DM}}_{\mathrm{MW}}$ is varied away from 80 pc cm^{-3}, the resulting posterior distributions of μ_{host} and of σ_{host} appear insignificantly changed. Thus the parameters we find assuming a log-normal distribution for $\mathrm{DM}_{\mathrm{host}}$ are relatively robust to reasonable assumptions of $\overline{\mathrm{DM}}_{\mathrm{MW}}$.

such, we continue with our fiducial value of $\overline{\mathrm{DM}}_{\mathrm{MW}} = 80$ pc cm^{-3}.

5.4. "Rate Interpretation" of α

Throughout this work, we have treated α as a true FRB spectral index, i.e., a spectral index for an individual FRB (observed as a broadband burst) that also applies to the full population of studied FRBs. The flux behaves as $F \sim \nu^{\alpha}$. However, as James et al. (2022b) noted, α can also be interpreted as a frequency-dependent rate, particularly when considering experiments with a similar bandwidth to that of the detected FRB. As FRBs can be band-limited (e.g., Law et al. 2017; Pleunis et al. 2021), this means that there are either a larger number of low-frequency FRBs, or that FRBs are stronger at low frequencies. Under such an interpretation, the rate now behaves as $\Phi \sim \Phi(z)\nu^{\alpha}$, and through the changed k-correction, the factor of $(1+z)^{\alpha}$ is directly added to the source evolution (Equation (A8)), thus most directly affecting the parameter n. It is likely that the most realistic treatment of α is more complicated than either single interpretation—for example, it is reasonable to think FRBs may be both broadband (as per the "true spectral index" interpretation) and have unique spectral indices (as per the "rate interpretation"). Thus the treatment of α remains a model-dependent systematic uncertainty.

For completeness, we redo the analysis, this time treating α as a frequency-dependent rate. The complete results of the parameters can be found in Figure B1 and Table B1. Of the parameters, α and n had the most significant changes; they become far more correlated under the "rate interpretation" of α. Under the "true spectral index interpretation" of α, we obtain $\alpha = -1.39^{+0.86}_{-1.19}$ and $n = 0.96^{+0.81}_{-0.67}$; under the "rate interpretation" of α, we obtain $\alpha' = -1.10^{+0.67}_{-0.99}$ and $n' = 1.72^{+1.48}_{-1.10}$.

6. Discussion

Using the population model developed by James et al. (2022b), this analysis simultaneously constrains the redshift, energy, and host DM distributions of FRBs using CHIME/FRB Catalog 1 data. This is the first time such a population study has been conducted on a large sample of bursts from the same uniform survey while explicitly accounting for selection effects; previous works have had to combine bursts from different samples, while attempting to account for all the heterogeneous selection effects involved (e.g., Luo et al. 2020; James et al. 2022b). Furthermore, it is the injections system of the CHIME/FRB instrument that crucially allows for the careful calibration of selection biases in a comprehensive end-to-end way that no other radio survey is currently able to conduct. There are a total of seven free parameters fitted in this population model to the CHIME/FRB data; when interpreting α as a frequency-dependent rate, another set of fit parameters can be obtained. In this section, we discuss our results and the property distributions they describe. Notably, in Section 6.6, we discuss the implications for considering repeating and non-repeating FRBs for this fitted one-population model.

6.1. FRB Volumetric Rate

We find a local rate of bursts above 10^{39} erg and below a scattering of 10 ms at 600 MHz of $7.3^{+8.8}_{-3.8} \times 10^4$ Gpc^{-3} yr^{-1}. This volumetric rate appears consistent with the rate presented by James et al. (2022a), who found a rate of bursts above 10^{39} erg of $8.7^{+1.7}_{-3.9} \times 10^4$ bursts Gpc^{-3} yr^{-1}, and is also broadly consistent with other studies as well (e.g., Lu & Piro 2019; Ravi 2019; Luo et al. 2020). On the surface, the consistent volumetric rates are interesting to note when considering the different frequency bands of the respective samples of FRBs. However, great caution should be employed when directly comparing these values, as the rates presented by other studies do not employ a scattering cut. As noted by CHIME/FRB Collaboration et al. (2021), there is evidence that there is a notable population of FRBs with a scattering timescale at 600 MHz larger than 10 ms that is missed by the CHIME/FRB observations. Although highly scattered bursts were excluded from this analysis due to the poorly constrained nature of their population, including such highly scattered bursts may greatly increase our volumetric rate, possibly

making the rate found in this analysis less comparable with other derived rates.

This volumetric rate of $7.3^{+8.8}_{-3.8} \times 10^4$ Gpc^{-3} yr^{-1} exceeds that of known populations of cataclysmic phenomena, especially when considering how the volumetric rate would increase when including highly scattered bursts. As summarized by Oguri (2019), Type Ia and core-collapse supernovae each have a local volumetric rate of $\mathcal{O}(10^4)$ Gpc^{-3} yr^{-1}, with superluminous supernovae having a lower volumetric rate of $\mathcal{O}(10^2)$ Gpc^{-3} yr^{-1}. Both long and short gamma-ray bursts have much lower local volumetric rates as well, with $\mathcal{O}(\lesssim 10^2)$ Gpc^{-3} yr^{-1}. The inferred rate of binary compact object mergers (e.g., binary black hole, binary neutron star, and neutron star–black hole) with the cumulative Gravitational Wave Transient Catalog 3 (GWTC-3) are all $\mathcal{O}(10^2)$ Gpc^{-3} yr^{-1} (The LIGO Scientific Collaboration et al. 2021). It has also been proposed that white dwarf mergers may occur in globular clusters in the local universe at $\mathcal{O}(10)$ Gpc^{-3} yr^{-1} (Kremer et al. 2021).

It is worth noting that when applying all the exclusion criteria in Section 2.3 except for the cut on repeat bursts, the Catalog 1 sample for populations analysis would have been 230 bursts instead of 225. However, given the total net systemic uncertainty on the sky rate found by CHIME/FRB Collaboration et al. (2021) was $^{+27\%}_{-25\%}$, this systemic uncertainty of treatment of repeaters is almost certainly a subdominant effect. Thus, from a practical point of view, this volumetric rate can be interpreted as either a rate of bursts or of sources. This reasoning is undeniably complicated by the fact that the observed rates are also strong functions of telescope sensitivity and the FRB repetition luminosity function.

Nonetheless, since it is possible that one-off bursts in our sample can actually be bursts from repeating sources, and the cut on repeaters is a subdominant effect in this analysis, if one looks purely at the obtained volumetric rate above 10^{39} erg, it may be physically sensible to interpret it as a rate of bursts. Extrapolating below the pivot energy introduces a large amount of uncertainty, and our model does not fit for a minimum energy, but it is reasonable to expect that the volumetric rate of FRBs is larger when including energies below $E_{\rm pivot}$. If our observed volumetric rate is interpreted as a rate of bursts, as this rate exceeds that of known one-off event populations (e.g., Nicholl et al. 2017), such a scenario could suggest that FRBs are predominantly from repeaters, most of which instruments such as CHIME/FRB may miss. This in turn could imply that many apparent one-off FRBs are in fact repeat bursts from progenitors with long stretches of burst quiescence (Caleb & Keane 2021), though the notion that all—or even most—FRBs are repeaters is one that has fallen out of relative favor (e.g., Ravi 2019; James et al. 2020; Ai et al. 2021). There is also the possibility that repeaters and one-off FRBs are distinct populations with distinct populations parameter values; further discussion on this consideration can be found in Section 6.6. Interpreting our observed volumetric rate as a rate of bursts can also suggest that FRBs predominantly come from a yet-unknown population of one-off events. Nonetheless, it is important to emphasize that this discussion is driven purely by consideration of the numerical volumetric rate result; our model $R(F, \mathrm{DM}|\boldsymbol{\lambda})$ does not directly make any claims as to which progenitor origin scenario is preferred. (The model does allow for source evolution, which could more directly argue for preferred progenitor scenarios based on evolution with SFR, but as discussed in Section 6.4, our results there are inconclusive.)

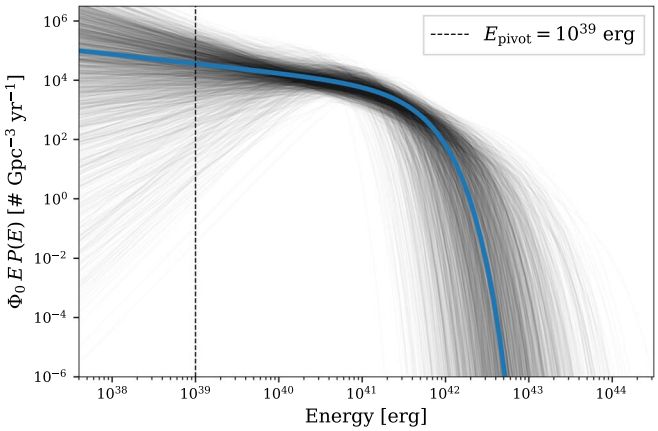

Figure 6. Plotted in gray are the results of the allowable Schechter function values scaled by energy and the volumetric rate from the MCMC run, thinned by a factor of 500 for visualization purposes. The volumetric rate is quoted above the pivot energy, $E_{\rm pivot} = 10^{39}$ erg, which is overlaid with a dashed black line. The best-fit Schechter function from the MCMC run is overlaid in solid blue. Abundances at a given energy, above $E_{\rm pivot}$, can be directly read off the y-axis.

6.2. FRB Energy Function

Constraining the energy function of FRBs is particularly illuminating for better understanding FRB origins. Although we model the energy distribution as a Schechter function for all energies, we only quote a volumetric rate above a pivot energy of $E_{\rm pivot} = 10^{39}$ erg. We do not test for a minimum energy. An allowable range of energy distributions is plotted in Figure 6. We see that the shape of the function is quite well-constrained—the characteristic energy after which the distribution exponentially falls off appears to be visually consistently on the order of $\approx 10^{41}$ erg. We find $\log_{10} E_{\rm char}$ (erg) to be $41.38^{+0.51}_{-0.50}$, in fairly good agreement with the James et al. (2022a) result of $\log_{10} E_{\rm max}$ (erg) of $41.70^{+0.53}_{-0.06}$. (Note James et al. 2022a preferred to model the energy distribution with a power law with a maximum energy cutoff rather than as a Schechter function, although both models were tested with comparable results.)

We also find a differential power-law energy distribution index of $\gamma = -1.3^{+0.7}_{-0.4}$. This result is in broad agreement with the results of Hashimoto et al. (2022), who also used Catalog 1 data to explore the energy function of FRBs, albeit employing a different methodology (e.g., they do not fit for a distance distribution simultaneously) and using a different subset of bursts released by CHIME/FRB Collaboration et al. (2021). Within the error bars, this power-law index is also in agreement with what Luo et al. (2020) found ($-1.79^{+0.31}_{-0.35}$) using a heterogenous sample of FRBs, including bursts from UTMOST-SS (Caleb et al. 2017) and from CRAFT (Shannon et al. 2018). Interesting to note is that James et al. (2022a), who developed the model $R(F, \mathrm{DM}|\boldsymbol{\lambda})$ we use in this analysis, find a steeper power-law index value ($-2.09^{+0.14}_{-0.10}$) using a sample of ASKAP and Parkes FRBs, though it is still consistent with our results within the 2σ statistical uncertainties. Further data and larger numbers of FRB observations are needed to explore the consistency of FRB energy distributions across different frequency bands.

6.3. FRB Spectral Dependence with α

When interpreting α as a true spectral index, every FRB is observed as a broadband burst with the same spectral index. Using a sample of 23 FRBs detected with ASKAP, Macquart et al. (2019) find $\alpha = -1.5$, a default value adopted in other studies (e.g., Lu & Piro 2019). Using a Gaussian prior on $\alpha = -1.5 \pm 0.3$, James et al. (2022b) find a consistent value of $\alpha = -1.55^{+0.18}_{-0.18}$. Our result of $\alpha = -1.39^{+0.86}_{-1.19}$ is consistent with $\alpha \approx -1.5$, although our statistical uncertainties are larger than those of the aforementioned quoted values.

The treatment of α is far from standardized in the literature. Multiple studies fix $\alpha = 0$ (e.g., Caleb et al. 2016; Luo et al. 2020; Chawla et al. 2022) when studying their sample. In particular, Chawla et al. (2022) set $\alpha = 0$ for the Catalog 1 sample to be consistent with the implicit assumptions made when measuring fluences for Catalog 1 bursts (CHIME/FRB Collaboration et al. 2021). Also, as has been detailed by James et al. (2022b) and mentioned by Law et al. (2017), there is little meaning of a "spectral index" for an individual FRB with a limited band occupancy. Instead, when interpreting α as a frequency-dependent rate, every FRB is observed as a narrowband burst, depending on the frequency-dependent survey sensitivity. Our rate interpretation result is $\alpha' = -1.10^{+0.67}_{-0.99}$, which is slightly flatter than our "spectral index" interpretation of $\alpha = -1.39^{+0.86}_{-1.19}$ but again, not significantly so—especially when considering the uncertainties. Certainly more work remains to better constrain the spectral index of FRBs, and how best to treat the systemic uncertainty of how to interpret α.

6.4. FRB Evolution with SFH

The progenitors of FRBs are a widely debated topic, and one way to probe their origins as a population is to track their redshift distribution with respect to SFH. If FRBs arise predominantly from magnetars, as the association of the FRB-like burst with the Galactic magnetar SGR 1935+2154 might suggest (Bochenek et al. 2020; CHIME/FRB Collaboration et al. 2020b), then they would closely track the SFH of the universe. On the other hand, if FRBs arise predominantly from older stellar populations, as the association of FRB 20200120E with a globular cluster in M81 might suggest (Bhardwaj et al. 2021; Kirsten et al. 2022), then they would not appear to trace the SFH. There is also the possibility that FRBs arise from younger and older stellar populations in comparable proportions (e.g., Guo & Wei 2022), but given the limited information regarding FRB host environments and the risk of over-interpreting available data, we do not consider this scenario in this work.

Instead, the formalism of this population model simply aims to quantify how closely the FRB population evolves with SFR. James et al. (2022a) claim that the FRB population does evolve with SFR, a result that is not incompatible with our results of $n = 0.96^{+0.81}_{-0.67}$ when considering the different samples of FRBs used. The frequency coverage differs between Parkes (1157–1546 MHz) and ASKAP (1129–1465 MHz) compared to CHIME (400–800 MHz), raising the possibility that there may be frequency-dependent evolution behavior. However, there is no concrete evidence in our analysis to support this claim, and again, any related investigation is outside the scope of this work. Other studies have investigated the cosmic evolution of FRBs (e.g., Hashimoto et al. 2022; Qiang et al. 2022; Zhang & Zhang 2022) using Catalog 1 data, and have concluded that FRBs from the CHIME/FRB Catalog 1 sample do not evolve with SFR. Our result of $n = 0.96^{+0.81}_{-0.67}$ is much less conclusive; the large uncertainties on our results are also visualized in Figure 7. Of our 10^6 chains, $\approx$7.2% give $n \leqslant 0$ (the rest of the parameter results vary across chains; there is no real physical interpretation for $n < 0$, but we allowed the range for completeness). Our result under the "rate interpretation" of α also gives a poorly constrained $n' = 1.72^{+1.48}_{-1.10}$; of those chains, $\approx$4.5% give $n \leqslant 0$. We note that under the "rate interpretation" of α, our results tend toward stronger evolution with SFR—consistent with what James et al. (2022b) found. However, the statistical uncertainties on both parameters, in both interpretations, are too large to draw any significant conclusions about how the interpretation of α may affect the modeling of FRB evolution with SFR.

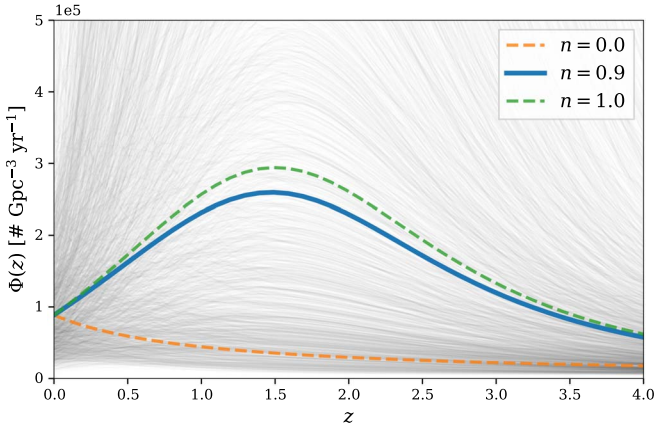

Figure 7. Results of how the Catalog 1 FRBs evolve with SFR, $\Phi(z)$. The gray lines denote the allowable behavior of $\Phi(z)$ using the parameters from the MCMC run, thinned by a factor of 500 for visualization purposes. The best-fit $\Phi(z)$ from the MCMC is overlaid in solid blue. The lower dashed orange curve represents no evolution with SFR, and the upper dashed green curve represents perfect evolution with SFR.

The difference between the results of this study and of previous studies using Catalog 1 data can be explained in a few ways. Using the full injections sample, this analysis fits a population model which simultaneously models both the energy distributions and the redshift distributions of FRBs. In the work presented by Hashimoto et al. (2022), redshifts are directly inferred from DM measurements rather than treating them as separate parameters, and they thus explored the redshift evolution of the derived FRB energy functions. Both Qiang et al. (2022) and Zhang & Zhang (2022) generate Monte Carlo distributions from varied redshift and energy distributions, and then test them against the publicly available CHIME/FRB Catalog 1 data. While this approach allows for explicit testing of different evolutionary models (i.e., including time delays), it does not fit specific parameters to the data—it can only test how consistent a data set may be with a tested distribution. Additionally, these previous studies did not have access to the full injections sample to calibrate out observational selection effects with their methodology. We do test one more toy model, elaborated on in Appendix C, to explore the potential constraining power of our data for FRBs from older stellar populations. Ultimately, though, because our result of $n = 0.96^{+0.81}_{-0.67}$ has such large uncertainties, we conclude that our data cannot meaningfully constrain how the FRB population evolves with the SFH of the universe. However, localizations to host galaxies and direct redshift measurements

of FRBs may enable better constraints in the future. They would enable determination of information such as stellar masses, which one can then use to also test whether FRB host galaxies track the stellar mass function (e.g., as Bhandari et al. 2022 did with ASKAP FRBs).

6.5. Host Galaxy DM Contribution

The rest-frame host galaxy contributions to the DMs of the FRBs were modeled with a log-normal distribution, parameterized by $\mu_{\rm host}$ and $\sigma_{\rm host}$. However, the values of $\mu_{\rm host}$ and $\sigma_{\rm host}$ do not directly correspond to physical values. Instead, the median and standard deviation of the log-normal rest-frame $\rm DM_{host}$ distribution described by those parameters are 84^{+69}_{-49} pc cm^{-3} and 174^{+319}_{-128} pc cm^{-3}, respectively. As illustrated by the statistical uncertainties, these parameters are not well-constrained; how poorly constrained these parameters are can also be visualized with the allowable $P(\rm DM'_{host})$ distributions plotted in Figure 8. As such, caution must be employed when comparing these values against those found in the literature. Nonetheless, we highlight a few observations.

Macquart et al. (2020) and James et al. (2022a, 2022b) also generically modeled the intrinsic host galaxy contribution to the DM with a log-normal distribution, with only the latter study explicitly quoting derived values. In the work presented by James et al. (2022a), they quote a log-mean host DM contribution of 129^{+66}_{-48} pc cm^{-3} on top of a typical Galactic contribution of $\rm DM_{disk} + DM_{halo} = 50 + 35 = 85$ pc cm^{-3}. It is important to note that James et al. (2022a) used a different sample of FRBs from different surveys (at different observing frequencies), did not apply a scattering cut for $\tau > 10$ ms to their sample as we did, and also quote a typical Galactic contribution of 85 pc cm^{-3} instead of 80 pc cm^{-3}, although this latter difference is likely to be minimal (Section 5.3.1). However, in Figure 1, we can see that our MCMC chains also strongly disfavor low $\sigma_{\rm host}$ and high $\mu_{\rm host}$ values, a finding consistent with James et al. (2022a).

Both James et al. (2022a) and our analysis find "central" values of $\rm DM'_{host}$ well within the statistical errors of each other, as well as within "typical" host galaxy contributions which are often quoted as $\sim$50 pc cm^{-3} (e.g., Shannon et al. 2018) or $\sim$100 pc cm^{-3} (e.g., Cordes & Chatterjee 2019). One should also note that the log-normal nature of the distribution allows for an occasional extraordinarily high host galaxy DM, such as the estimated $\rm DM_{host} \approx 903^{+72}_{-111}$ pc cm^{-3} found for FRB 20190520B by Niu et al. (2021a). This is also in agreement with Rafiei-Ravandi et al. (2021), who find statistical evidence for some FRBs having a large host DM (defined therein as $\sim$400 pc cm^{-3}). Overall, we note that the poor constraints we obtain with the Catalog 1 data are perhaps unsurprising when one considers that no redshift information was used for any FRB in this study.

6.6. Considering Repeaters Versus Non-repeaters

A growing number of FRBs has been observed to repeat (e.g., Scholz et al. 2016; Spitler et al. 2016; CHIME/FRB Collaboration et al. 2019, 2020a; Fonseca et al. 2020; Bhardwaj et al. 2021; Lanman et al. 2021). There is currently no reliable method to identify which FRBs are from repeaters, but the growing sample of observed repeat bursts has led to some promising methods. Pleunis et al. (2021) found that repeaters could be a distinct population from apparent one-offs

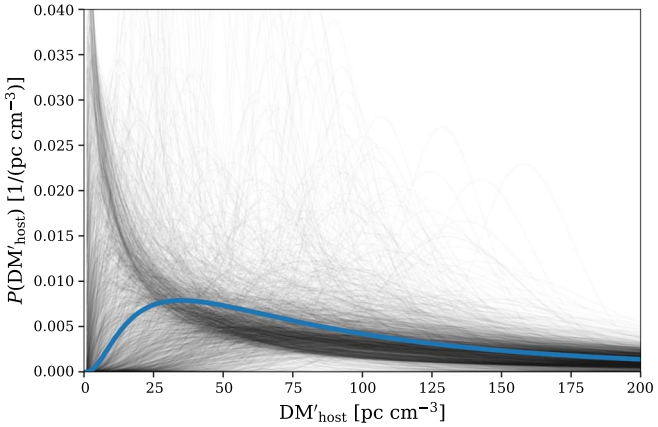

Figure 8. Plotted in gray are the results of the allowable distributions of the rest-frame contribution to the DM from the host galaxies of FRBs, $P(\rm DM'_{host})$, from the MCMC run, thinned by a factor of 500 for visualization purposes. The best-fit distribution from the MCMC run is overlaid in solid blue. Due to the extended tail nature of this distribution, only values up to $\rm DM'_{host} = 200$ pc cm^{-3} are plotted. However, higher $\rm DM'_{host}$ values are certainly allowed.

based on the temporal and spectral studies of the Catalog 1 sample. Also using the Catalog 1 sample, Chen et al. (2022) have recently used unsupervised machine learning to create a catalog of repeater source candidates. Of course, all this is assuming that there are two distinct populations of FRB sources: those that repeat, and those that do not. Although observationally such a classification scheme is intuitive, it is also possible that all FRBs repeat, or all FRBs represent distinct phases of the same class of progenitors (e.g., Caleb et al. 2019; Cui et al. 2021; Pleunis et al. 2021).

If most FRBs are from repeaters, the results presented here are likely to be affected to a stronger degree than if a lower fraction of FRBs is from repeaters. As mentioned in Section 2.3, to minimize unquantifiable biases, we include only the first-detected burst from any observed FRB source. However, one should expect the repetition rate of FRBs to affect the distributions of observed parameters. Naively, the presence of increasing numbers of repeating FRBs would cause the burst rate within an observed volume of the universe to be increasingly greater than expected. Likewise, Gardenier et al. (2021) noted that repeating FRBs should have a lower mean DM than apparent one-off FRBs, although CHIME/FRB Collaboration et al. (2021) found that the distributions of repeating FRBs and apparent one-off FRBs are consistent with being drawn from the same underlying DM distribution. Pleunis et al. (2021) show that there are observable differences in the spectral parameters (e.g., α) of repeaters and non-repeaters, and of course the differences could extend to physically different values of $\rm DM_{host}$ contributions (e.g., $\mu_{\rm host}$ and $\sigma_{\rm host}$), evolution with SFH (e.g., n), and more.

Although it is tempting to conduct a populations study considering only apparent non-repeaters from the Catalog 1 sample, it is likely that some apparent one-off bursts are actually from repeater sources, with both Pleunis et al. (2021) and Chen et al. (2022) proposing candidate apparent one-off bursts that may actually be from repeater sources. In fact, FRB20190304A was identified as an apparent non-repeater source in Catalog 1, but since then another burst has been observed from this source (CHIME/FRB Collaboration et al., private communication), establishing its nature as a repeater instead. There is also the fact that CHIME/FRB has only

observed one burst from the first repeating source FRB 121102 (Josephy et al. 2019), despite there being many more additional observed bursts present in the literature (e.g., Rajwade et al. 2020; Li et al. 2021). Furthermore, Good et al. (2022) demonstrate that the possibility of a low-repetition subpopulation may further complicate the observational identification of repeater FRBs. The presence of such as-yet-unidentified repeat bursts could represent a not insignificant "contamination" of the nominal non-repeating sample of FRBs. Chen et al. (2022) suggest a large potential "contamination" of repeaters, but Pleunis et al. (2021) warn that instrumental effects, such as chromatic sensitivities away from the detection beam centers, make it difficult to isolate the true repeaters. This is because such instrumental effects would result in observed bursts with biased low bandwidths that also cause mixing of bursts with spectra that are intrinsically broad and narrow. Further complicating this unquantifiable effect of as-yet unidentified repeaters in a sample of nominally one-off bursts is that the injections framework of CHIME/FRB currently only injects bursts from a single population distribution of FRB parameters. Nonetheless, as larger samples of well-defined repeaters are identified, we expect our ability to do comparative populations analyses between observed repeat bursts and one-off bursts will improve.

7. CHIME/FRB Outrigger Predictions

For the majority of FRBs detected by CHIME/FRB, the localization precision is insufficient for identifying the host galaxies for these bursts. Host galaxies, and thus redshifts, for FRBs are highly desirable for both understanding the nature of FRB origins (Bhandari et al. 2022) and using them as cosmological probes (Masui & Sigurdson 2015). Enabling larger-scale localization abilities is a high-priority science goal for FRBs, and one that will be addressed with the CHIME/FRB Outriggers project (e.g., Cary et al. 2021; Leung et al. 2021; Cassanelli et al. 2022; Mena-Parra et al. 2022). Through CHIME/FRB Outriggers, CHIME-like outrigger telescopes will be deployed at continental baselines to form a very-long-baseline interferometry (VLBI) network. While CHIME/FRB Outriggers will trigger on a subset of sources observed by CHIME/FRB (e.g., not on low-S/N bursts), the VLBI network will still obtain near real-time, subarcsecond localization capability for hundreds of FRBs a year (Mena-Parra et al. 2022). Although it is unrealistic to expect that every FRB will have its VLBI position followed-up with optical observations to obtain its spectroscopic redshift, it would certainly be possible for spectroscopic redshifts to be obtained for a much more significant fraction of FRBs than is currently possible. Additionally, a larger fraction of FRBs may have reasonably reliable photometric redshifts obtained by cross-matching the VLBI localizations with photometric galaxy surveys such as the Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Surveys (Dey et al. 2019). While photometric redshifts would not be as reliable as spectroscopic redshifts, they would be more informative than redshifts inferred from DM values. In any event, further additional redshift information would in turn allow for much more sensitive constraints on the intrinsic $\mathrm{DM_{host}}$ contribution.

With the methodology used in this paper, we can obtain a prediction for the observed FRB redshift distribution that CHIME/FRB Outriggers may see. Recall Equation (1)

$$R(F,\,\mathrm{DM}|\boldsymbol{\lambda}) = \int dz R(F,\,z|\boldsymbol{\lambda}_1) P(\mathrm{DM}|z,\,\boldsymbol{\lambda}_2).$$

It is here that the model $R(F, \mathrm{DM}|\boldsymbol{\lambda})$ is marginalized over z. As described in Section 4, $R(F, \mathrm{DM}|\boldsymbol{\lambda})$ is used to construct parameter-dependent weights. These weights are then used to create the model prediction ζ_{ij}, which is a weighted histogram of the synthetic data in S/N and DM (Equation (7)). Properly normalized, this earlier obtained ζ_{ij} can be translated to observed rate distributions of S/N and DM, given parameters $\boldsymbol{\lambda}$, over the redshift range from 1×10^{-3} to 4.

Instead of integrating over z and ending up with a 2D model prediction (histogram) in S/N and DM, we can construct a 3D model prediction that depends on S/N, DM, and z

$$\zeta_{ij}|_{z_{\mathrm{val}}} = \sum_{m=1}^{n_{ij}} W_m(F,\,\mathrm{DM},\,z_{\mathrm{val}},\,\boldsymbol{\Theta},\,\boldsymbol{\lambda})\,, \tag{14}$$

$$= \sum_{m=1}^{n_{ij}} \frac{R(F,\,\mathrm{DM},\,z_{\mathrm{val}}|\boldsymbol{\lambda})\Delta t}{R_{\mathrm{init}}(F,\,\mathrm{DM})}\,, \tag{15}$$

where

$$R(F,\,\mathrm{DM},\,z_{\mathrm{val}}|\boldsymbol{\lambda}) = \Delta z\; R(F,\,z_{\mathrm{val}}|\boldsymbol{\lambda}_1) P(\mathrm{DM}|z_{\mathrm{val}},\,\boldsymbol{\lambda}_2), \tag{16}$$

and Δz is estimated by the midpoint spacing between the evaluated z_{val} points across the full redshift range range from 1×10^{-3} to 4. (Note that for simplicity of notation, $\boldsymbol{\Theta}$ was dropped as their distributions were held to the fiducial models found in Catalog 1 for all evaluated models, for each burst.)

This 3D histogram, after dividing by the bin widths in z, gives a prediction for the observed CHIME/FRB Catalog 1 DM–S/N–z distribution per z for a given model $R(F, \mathrm{DM}|\boldsymbol{\lambda})$. We can thus sum over the S/N bins (which span the values 12–200) and the DM bins (which span the values 100–3500 pc $\mathrm{cm^{-3}}$) to obtain a prediction for the observed redshift distribution, $P_{\mathrm{obs}}(z)$. The range of predicted redshift distributions our MCMC run allows appears reasonably well-constrained, and is shown in Figure 9.

The $P_{\mathrm{obs}}(z)$ distribution predicted by the best-fit model peaks at around a redshift of $z \approx 0.36$. This result is consistent with the DM distribution observed by CHIME/FRB shown in CHIME/FRB Collaboration et al. (2021), where the peak of that distribution is around DM ≈ 500 pc $\mathrm{cm^{-3}}$. That the redshift peak is below $z \sim 0.5$ suggests that there is a noticeable fraction of low- to intermediate-z FRBs that have a sizable DM contribution from astrophysical environments other than the cosmological IGM (e.g., local environments, host galaxies, or intervening galaxies). This $P_{\mathrm{obs}}(z)$ prediction also contains $\approx$99.4% of its area within $z \leqslant 2$, suggesting that CHIME/FRB Outriggers will see a negligible number FRBs with $z > 2$. Such a result implies that the sample of FRBs observed by CHIME/FRB Outriggers will be limited in their ability to be used as high-z cosmological probes (e.g., Linder 2020). It is worth noting this prediction is based on the sample that survived the statistical cuts detailed in Section 2.3; therefore, it is likely that the true $R_{\mathrm{obs}}(z)$ distribution will differ when including observed bursts that do not pass these cuts (such as bursts with DM < 1.5 max ($\mathrm{DM_{NE2001}}$, $\mathrm{DM_{YMW16}}$) and S/N < 12). Nevertheless, this prediction supports the notion that low-z cosmological studies such as lensing from foreground structures will likely be a promising scientific application of

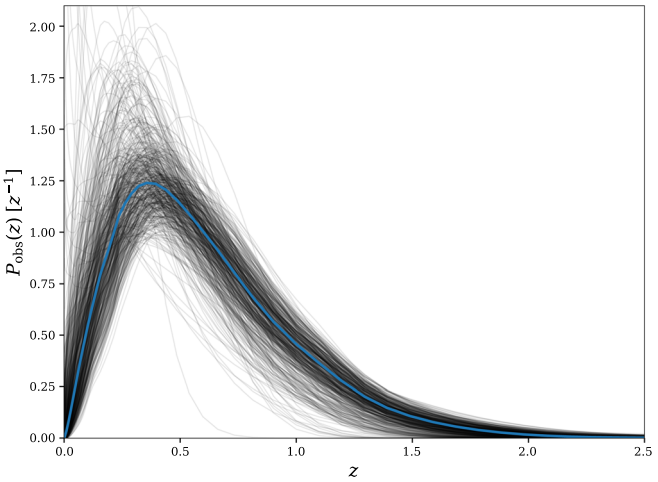

Figure 9. Plotted in gray are the results of the allowable observed FRB redshift distributions $R_{\rm obs}(z)$ from the MCMC run, thinned by a factor of 2000 for visualization purposes. The $R_{\rm obs}(z)$ prediction using the best-fit distribution from the MCMC run is overlaid in solid blue. Although the model was fit up to $z = 4$, only redshifts up to $z = 2.5$ are plotted.

FRBs in the near future (e.g., Kader et al. 2022; Leung et al. 2022).

We can also take the 3D histogram that gives the prediction for the observed CHIME/FRB Catalog 1 DM-S/N-z distribution, given by Equations (14) and (16), and divide by the bin widths in z and DM. Then, we can sum over the S/N bins and normalize the distribution such that we obtain a prediction for the observed joint redshift and DM distribution, per z and per DM: $P_{\rm obs}(z|{\rm DM})$. We show a visualization of what $P_{\rm obs}(z|{\rm DM})$ predictions are allowed by the MCMC run in Figure 10.

In the left-hand panel of Figure 10, the $P_{\rm obs}(z|{\rm DM})$ result shows that, although the relation between the redshift and DM of FRBs is a reasonably tight, single-modal distribution (e.g., Shannon et al. 2018; Macquart et al. 2020), for a survey sample as large as Catalog 1, there is a "turnover" at high DM where the true redshift of an FRB is much lower than would be expected by applying a naive ∼linear DM–z relation (first shown by James et al. 2022b). This can be seen the apparent excess of probability in the overlapping region of $0.5 \lesssim z \lesssim 1.0$ and ${\rm DM} \gtrsim 2000$ pc cm^{-3}. There, the plot is saying that, for an FRB observed by CHIME/FRB Outriggers with a measured ${\rm DM} \gtrsim 2000$ pc cm^{-3}, it is more likely for the source to have a redshift of $z < 1$ with a high ${\rm DM_{host}}$ contribution than a redshift of $z > 2$ with a large ${\rm DM_{IGM}}$ contribution. This finding is consistent with Rafiei-Ravandi et al. (2021), who find statistical evidence for a nonzero correlation between high-DM FRBs ($\geqslant$785 pc cm^{-3}) and galaxies at $z \sim 0.4$, with a ${\rm DM_{host}}$ contribution ∼greater than the ${\rm DM_{IGM}}$ contribution. This result once again emphasizes the importance of actually obtaining redshifts for FRBs rather than relying on inferred redshift values when doing cosmological studies, especially for high-DM FRBs. Of course, one must keep in mind that these predictions are for an overall population, and that an individual detected FRB can certainly be a cosmologically distant source with a high intrinsic luminosity. In fact, Ryder et al. (2022) recently reported the discovery of FRB 20220610A, the first FRB observed at $z \gtrsim 1$.

The right-hand panel of this figure shows the fractional uncertainties associated with our prediction for $P_{\rm obs}(z|{\rm DM})$ across the chains in our MCMC run. This plot illustrates that, along the relation between z and DM, the predicted $P_{\rm obs}(z|{\rm DM})$ is quite consistent across realizations. The further the deviation from that relation, the more "scatter" there is between predictions in the run. It is also worth noting that "fractional uncertainty" can also mean "scatter surrounding how close this value is to zero across runs," as appears to be the case in the region with ${\rm DM} < 500$ pc cm^{-3} and $0.5 < z < 1.0$—a region that appears to have negligible probability in the left-hand panel of this figure, but can have scatter due to variations in large scale structure noise.

8. Conclusion

Using CHIME/FRB Catalog 1 data, we improve upon the initial Catalog 1 populations analysis by fitting a more physical model to the fluence and DMs of the FRBs, a model initially developed in previous works (e.g., Luo et al. 2020; James et al. 2022b). We use the weighting formalism developed by CHIME/FRB Collaboration et al. (2021), which makes use of injected synthetic pulses to calibrate out selection function effects when fitting the model to the observed FRB data. The model fits for seven parameters: a volumetric rate Φ_0, a spectral index α, a differential power-law energy index γ, a characteristic exponential cutoff energy $E_{\rm char}$, a power-law index n for smooth-scaling evolution with SFR, and parameters $\mu_{\rm host}$ and $\sigma_{\rm host}$ that parameterize the log-normal distribution of the host DM contribution. We also explore the "rate interpretation" of α.

We find a volumetric rate of FRBs above 10^{39} erg and below a scattering of 10 ms at 600 MHz of $\Phi_0 = [7.3^{+8.8}_{-3.8}({\rm stat.})^{+2.0}_{-1.8}({\rm sys.})] \times 10^4$ Gpc^{-3} yr^{-1}. We also find tight constraints on the energy distribution observed by CHIME/FRB, modeled as a Schechter function, with $E_{\rm char} = 2.38^{+5.35}_{-1.64} \times 10^{41}$ erg and $\gamma = -1.3^{+0.7}_{-0.4}$. However, the constraints on the remainder of the parameters are much weaker. In particular, the constraints on n, $\mu_{\rm host}$, and $\sigma_{\rm host}$ may significantly improve with the use of redshift information for many FRBs. Such a sample of redshift information would be enabled by the upcoming CHIME/FRB Outriggers project, for which we predict the observed FRB redshift distribution. The distribution shows that the application of CHIME-detected FRBs to cosmological studies may be limited to the $z < 2$ regime, which seems to limit the possibility of using these FRBs to probe events such as helium reionization, but does allow for FRB lensing studies.

As the amount of data seems to be the limiting factor for this FRB populations study, it is promising that CHIME/FRB has continued to detect a large number of FRBs since Catalog 1 (which only included bursts up to 2019 July 1). These observed FRBs, combined with knowledge about survey selection effects provided by the injections system, and possibly redshift information with the commissioning of CHIME/FRB Outriggers, suggest much more powerful populations-level constraints in the near future. With more data comes more responsibility to handle subtle effects—namely, how best to treat repeating sources versus apparent one-off bursts when modeling the FRB distribution. Such a question would be of particular importance when exploring population behavior in a parameter space where repeater bursts would dominate, such as events with large pulse widths. In addition, proper modeling of repeater and non-repeater sources—whether apparently or intrinsically non-repeating—would enable future large-sample population studies to investigate to what extent all FRBs may repeat. Further in the future, care must be taken to incorporate redshift information into populations

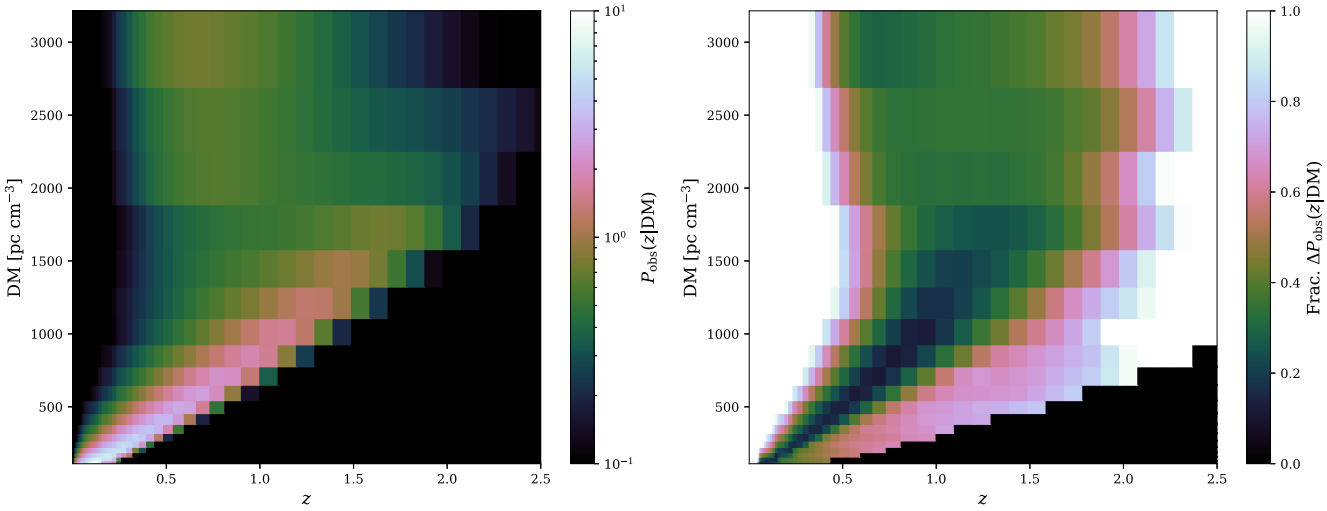

Figure 10. Left panel: the median $P_{\rm obs}(z|{\rm DM})$ distribution for the results of the MCMC run, thinned by a factor of 2000 for computational speed purposes. Evidence for a "turnover" of the DM–z relation can be seen at higher DMs and lower redshifts, although most of the FRBs appear to occur at DM < 2000 pc cm^{-3} and $z < 2$. Right panel: the standard deviation of the of the $P_{\rm obs}({\rm DM}, z)$ distributions at each point divided by the median $P_{\rm obs}({\rm DM}, z)$ (shown on the left) to visualize the fractional uncertainty of $P_{\rm obs}(z|{\rm DM})$ values across MCMC realizations.

studies properly while accounting for the selection biases (e.g., repeater or high-S/N) that lead to some, but not all, FRBs having associated redshifts.

We acknowledge that CHIME is located on the traditional, ancestral, and unceded territory of the Syilx/Okanagan people. We thank the anonymous referee for comments that have improved the quality of this manuscript. We are grateful to the staff of the Dominion Radio Astrophysical Observatory, which is operated by the National Research Council of Canada. CHIME is funded by a grant from the Canada Foundation for Innovation (CFI) 2012 Leading Edge Fund (Project 31170) and by contributions from the provinces of British Columbia, Québec and Ontario. The CHIME/FRB Project is funded by a grant from the CFI 2015 Innovation Fund (Project 33213) and by contributions from the provinces of British Columbia and Québec, and by the Dunlap Institute for Astronomy and Astrophysics at the University of Toronto. Additional support was provided by the Canadian Institute for Advanced Research (CIFAR), McGill University and the McGill Space Institute thanks to the Trottier Family Foundation, and the University of British Columbia. A.B.P. is a McGill Space Institute (MSI) Fellow and a Fonds de Recherche du Quebec—Nature et Technologies (FRQNT) postdoctoral fellow.

A.B.P. is a McGill Space Institute (MSI) Fellow and a Fonds de Recherche du Quebec—Nature et Technologies (FRQNT) postdoctoral fellow. A.H.M. is supported by the CANSSI CRT program. B.M.G. is supported by an NSERC Discovery Grant (RGPIN-2022-03163) and by the Canada Research Chairs (CRC) program. C.L. was supported by the U.S. Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program. F.A.D. is supported by the U.B.C Four Year Fellowship. FRB research at UBC is supported by an NSERC Discovery Grant and by the Canadian Institute for Advanced Research. K.S. is supported by the NSF Graduate Research Fellowship Program. K.W.M. is supported by NSF grants 2008031 and 2018490. M.B. is supported by an FRQNT Doctoral Research Award. M.D. is supported by a Canada Research Chair, NSERC Discovery Grant, CIFAR, and by the FRQNT Centre de Recherche en Astrophysique du Québec (CRAQ). M.Me. is supported by an NSERC PGS-D award. M.M. is supported by DOE grant DE-SC0022342. S.P.T. is a CIFAR Azrieli Global Scholar in the Gravity and Extreme Universe Program. V.M.K. holds the Lorne Trottier Chair in Astrophysics & Cosmology, a Distinguished James McGill Professorship, and receives support from an NSERC Discovery grant (RGPIN 228738-13), from an R. Howard Webster Foundation Fellowship from CIFAR, and from the FRQNT CRAQ.

Appendix A
Details of the Fluence–DM Distribution Model

A.1. Joint Fluence and Redshift Distribution

In this section, we focus on how to obtain $R(F, z|\boldsymbol{\lambda}_1)$. The parameter vector $\boldsymbol{\lambda}_1$ contains information related to FRB energetics.

First, we define the energy function $P(E)$. We model $P(E)$ with a Schechter function with a characteristic exponential cutoff energy $E_{\rm char}$ and power-law index γ, e.g.,

$$P(E)dE \propto \frac{1}{E_{\rm char}}\left(\frac{E}{E_{\rm char}}\right)^{\gamma} \exp\left[-\frac{E}{E_{\rm char}}\right]dE. \quad \text{(A1)}$$

In order for $\int E\,P(E)\,dE$ in Equation (A1) to not diverge, we must have $\gamma < -2$. However, this function is not truly normalizable, i.e., we cannot guarantee $\int_0^{\infty} P(E)\,dE = 1$ for all redshifts. Thus, we shall choose to normalize the function such that $\int_{E_{\rm pivot}}^{\infty} P(E)\,dE = 1$ while assuming the energy itself extrapolates to follow the Schechter function. This allows us to assume this energy distribution is valid for all FRB energies down to a minimum (but unknown) energy $E_{\rm min}$, while quoting a rate of FRBs above $E_{\rm pivot}$. For a reasonable choice of $E_{\rm pivot}$, Φ_0 should be weakly correlated with γ. We chose a pivot energy of 10^{39} erg as it is above the threshold of ruled-out minimum FRB energies (e.g., James et al. 2022b); in our results in Figure 1, we can see that Φ_0 and γ are clearly correlated, but because their correlation is not strong, we did not fine tune $E_{\rm pivot}$ further so as to have Φ_0 and γ be more weakly correlated.

To normalize this energy distribution, recall that the integral of a Schechter function is an upper incomplete gamma function, defined as

$$\Gamma\left(\gamma+1, \frac{E_{\rm pivot}}{E_{\rm char}}\right) = \int_{E_{\rm pivot}}^{\infty} \frac{1}{E_{\rm char}} \times \left(\frac{E}{E_{\rm char}}\right)^{\gamma} \exp\left[-\frac{E}{E_{\rm char}}\right] dE. \quad \text{(A2)}$$

Thus, the "normalized" Schechter function above a pivot energy $E_{\rm pivot}$ is

$$P(E)dE = \frac{1}{\Gamma(\gamma+1, E_{\rm pivot}/E_{\rm char})} \frac{1}{E_{\rm char}} \times \left(\frac{E}{E_{\rm char}}\right)^{\gamma} \exp\left[-\frac{E}{E_{\rm char}}\right] dE. \quad \text{(A3)}$$

With this normalization, all rates should be interpreted as a rate above the pivot energy $E_{\rm pivot}$.

We can convert between the FRB energy and fluence with

$$E = \frac{4\pi D_L^2(z)}{(1+z)^{2+\alpha}} \Delta\nu F, \quad \text{(A4)}$$

where E is the FRB energy, $D_L(z)$ is the luminosity distance, α is the spectral index ($F \propto \nu^{\alpha}$), and $\Delta\nu$ is the frequency bandwidth (taken as 1 GHz).

Using the fact that $\int P(F, z)dF = \int P(E, z)dE$, $R(F, z)$ and $P(E)$ are related by

$$R(F, z) = P(F|z)R(z) = R(E, z)\frac{dE}{dF}, \quad \text{(A5)}$$

$$= P(E)R(z)\frac{dE}{dF}. \quad \text{(A6)}$$

$R(z)$ is a rate per unit redshift, and contains the comoving volume element scaled by the SFR at that given redshift

$$R(z) = \chi(z)^2 \frac{d\chi}{dz} \Phi(z). \quad \text{(A7)}$$

Since $\chi(z)$ is the comoving distance, the component $\chi(z)^2 \frac{d\chi}{dz}$ is equivalent to the volume element $\frac{dV}{dz}$. $\Phi(z)$ is the rate of FRBs per comoving volume, which can then be generically modeled as

$$\Phi(z) = \frac{\Phi_0}{1+z}\left(\frac{\rm SFR(z)}{\rm SFR(0)}\right)^n, \quad \text{(A8)}$$

with Φ_0 representing the rate of FRBs per comoving volume at $z = 0$ and $E > E_{\rm pivot}$, taking the units of bursts per proper time per comoving volume, i.e., bursts yr^{-1} Gpc^{-1}.

The SFR(z) is given by Madau & Dickinson (2014) as

$$\rm SFR(z) = 1.0025738 \frac{(1+z)^{2.7}}{1+\left(\frac{1+z}{2.9}\right)^{5.6}}. \quad \text{(A9)}$$

This means the distribution of FRB energy as a function of distance is proportional to the distribution of energies at a given volume element and scaled by the SFR.

We can obtain $\frac{dE}{dF}$ from Equation (A4), which we can then combine with Equations (A1) and (A7) to get

$$R(F, z) = P(E)\chi(z)^2 \frac{d\chi}{dz}\Phi(z)\frac{4\pi D_L^2(z)}{(1+z)^{2+\alpha}}\Delta\nu. \quad \text{(A10)}$$

Thus, the parameters in $\boldsymbol{\lambda}_1$ are α, the spectral index, n, the power-law scaling with SFR, γ, the power-law index, and $E_{\rm char}$, the exponential cutoff energy, the latter two of which together characterize the Schechter energy function.

This aspect of the model was evaluated on an intrinsic grid in F and z space. There were 200 F points spaced logarithmically from $F = 1$–5×10^5 Jy ms, encompassing the range of burst fluences from the Catalog 1 sample. There were 100 points in z space, spaced logarithmically from $z = 1 \times 10^{-3}$–1 and then linearly from $z = 1$–4. This sampling was done in order to sample densely and reduce the numerical error for the relatively low-z region, where we would expect most FRBs to be.

A.2. DM Distribution

We now have the first component of the integrand needed to find $R(F, {\rm DM}|\boldsymbol{\lambda})$; the second component of the integrand is given by

$$P({\rm DM}|z, \boldsymbol{\lambda}_2) = P({\rm DM} = {\rm DM_{EG}} + \overline{\rm DM}_{\rm MW}|z, \boldsymbol{\lambda}_2), \quad \text{(A11)}$$

where the observed DM has an extragalactic contribution $\rm DM_{EG}$ and an average Galactic contribution $\overline{\rm DM}_{\rm MW}$ consisting of contributions from the disk and halo of the Milky Way. To obtain $\rm DM_{disk}$, the ISM contribution to the DM in the Milky Way disk, we consider the NE2001 model of electron density (Cordes & Lazio 2002) and the YMW2016 model of electron density (Yao et al. 2017) for the observed Catalog 1 bursts. The former gives a median value of $\rm DM_{disk} \approx 50$ pc cm^{-3}, while the latter gives a median value of $\rm DM_{disk} \approx 45$ pc cm^{-3}. The halo contribution is much more uncertain, with estimates ranging from ≈10–80 pc cm^{-3} (e.g., Dolag et al. 2015; Prochaska & Zheng 2019; Keating & Pen 2020). As done by Chawla et al. (2022), we start with the NE2001 model for estimating $\rm DM_{disk}$, and based on Dolag et al. (2015), adopt $\rm DM_{halo} = 30$ pc cm^{-3}. Thus, $\overline{\rm DM}_{\rm MW}$ is originally fixed to be 80 pc cm^{-3} in our model, where deviations from the assumption will be absorbed into contribution from the host galaxy of the FRB. We test the sensitivity of our results to the adopted $\overline{\rm DM}_{\rm MW}$ in Section 5.3.1. Since $\overline{\rm DM}_{\rm MW}$ is fixed, obtaining $P({\rm DM_{EG}}|z, \boldsymbol{\lambda}_2)$ will then allow us to obtain $P({\rm DM}|z, \boldsymbol{\lambda}_2)$.

$\rm DM_{EG}$ has contributions from the cosmological ionized gas distribution and host galaxy of the FRB, e.g., $\rm DM_{EG} = DM_{cosmic} + DM_{host}$. Using the mean density of ions $\bar{n}_e$ and relevant cosmological parameters, the expected cosmological contribution to the DM can be written as

$$\langle {\rm DM_{cosmic}} \rangle (z) = \int_0^z \frac{c\bar{n}_e(z')dz'}{H_0(1+z')^2\sqrt{\Omega_m(1+z')^3+\Omega_\Lambda}}. \quad \text{(A12)}$$

Then

$$\rm DM_{cosmic} = \Delta_{DM}\langle DM_{cosmic}\rangle(z), \quad \text{(A13)}$$

where $\Delta_{\rm DM}$ is the deviation from the mean. Thus

$$P(\mathrm{DM}_{\rm cosmic}|z) = P(\Delta_{\rm DM})\frac{d\Delta_{\rm DM}}{d\mathrm{DM}_{\rm cosmic}}, \quad \text{(A14)}$$

$$= P(\Delta_{\rm DM})\frac{1}{\langle \mathrm{DM}_{\rm cosmic}\rangle(z)}. \quad \text{(A15)}$$

The expression $P(\Delta_{\rm DM})$ is given by

$$P(\Delta_{\rm DM}|z) = \kappa\Delta_{\rm DM}^{-B}\exp\left[-\frac{(\Delta_{\rm DM}^{-A} - C_0)^2}{2A^2\sigma_{\rm DM}^2}\right], \quad \text{(A16)}$$

with $A = 3$, $B = 3$, κ the normalization factor, C_0 set by the requirement of $\langle\Delta_{\rm DM}\rangle = 1$, and $\sigma_{\rm DM}$ parameterized by feedback with $Fz^{-0.5} = 0.32z^{-0.5}$ (James et al. 2022b). Note that these parameters are fixed in all the fits.

The host galaxy contribution can be generically modeled as a log-normal distribution

$$P(\mathrm{DM}'_{\rm host}) = \frac{1}{\mathrm{DM}'_{\rm host}}\frac{1}{\sigma_{\rm host}\sqrt{2\pi}} \times \exp\left[-\frac{(\log\mathrm{DM}'_{\rm host} - \mu_{\rm host})^2}{2\sigma_{\rm host}^2}\right], \quad \text{(A17)}$$

where $\mu_{\rm host}$ and $\sigma_{\rm host}$ are free parameters that make up the parameter vector $\boldsymbol{\lambda}_2$. The median of this distribution is $\exp(\mu_{\rm host})$, and the variance is $[\exp(\sigma_{\rm host}^2) - 1]\exp(2\mu_{\rm host} + \sigma_{\rm host}^2)$. Thus, $\boldsymbol{\lambda}_2$ contains information related to the host galaxy contribution to the extragalactic DM.

The host contribution should also be corrected for redshift via

$$\mathrm{DM}_{\rm host} = \frac{\mathrm{DM}'_{\rm host}}{1+z}. \quad \text{(A18)}$$

A more sophisticated model, using a data set with more constraining power, would also take into consideration the evolution of the DM distribution with the cosmic SFH (e.g., Luo et al. 2018).

Similarly to the cosmological DM

$$P(\mathrm{DM}_{\rm host}|z) = P(\mathrm{DM}'_{\rm host})\frac{d\mathrm{DM}'_{\rm host}}{\mathrm{DM}_{\rm host}}, \quad \text{(A19)}$$

$$= \frac{1}{\mathrm{DM}'_{\rm host}}\frac{1}{\sigma_{\rm host}\sqrt{2\pi}} \times \exp\left[-\frac{(\log\mathrm{DM}'_{\rm host} - \mu_{\rm host})^2}{2\sigma_{\rm host}^2}\right](1+z). \quad \text{(A20)}$$

With expressions for $P(\mathrm{DM}_{\rm cosmic}|z)$ and $P(\mathrm{DM}_{\rm host}|z)$, the expression for the extragalactic contribution to the DM becomes a convolution

$$P(\mathrm{DM}_{\rm EG}|z) = \int d\mathrm{DM}_{\rm host}P(\mathrm{DM}_{\rm host}|z)P \times (\mathrm{DM}_{\rm cosmic} = \mathrm{DM}_{\rm EG} - \mathrm{DM}_{\rm host}|z). \quad \text{(A21)}$$

However, our expression for $P(\mathrm{DM}_{\rm host}|z)$ is analytical, while our expression for $P(\mathrm{DM}_{\rm cosmic}|z)$ is numerical. Thus, a change of variables in the convolution is better for ease of evaluation

$$P(\mathrm{DM}_{\rm EG}|z) = \int d\mathrm{DM}_{\rm cosmic}P(\mathrm{DM}_{\rm host} = \mathrm{DM}_{\rm EG} - \mathrm{DM}_{\rm cosmic}|z)P(\mathrm{DM}_{\rm cosmic}|z). \quad \text{(A22)}$$

Note that while Chawla et al. (2022) modeled host galaxy contributions to the DM using the Catalog 1 sample, the results of the simulations were predictions with noise. While it was the more complete way to model host galaxy contribution compared to this model, this methodology needs a more precise analytical form.

By Equation (A11), we can obtain $P(\mathrm{DM}|z)$. Combining that with the earlier expression for $R(F, z)$ gives us the brightness distribution $R(F, \mathrm{DM})$.

This aspect of the model was evaluated on an intrinsic grid in z and DM space. The points in z space are the same as outlined in Appendix A.1, whereas there are 150 points in DM space linearly spaced from $\mathrm{DM} = 100$–6500 pc cm^{-3}. When evaluating on the $\mathrm{DM}_{\rm EG}$ grid, $\overline{\mathrm{DM}}_{\rm MW}$ is subtracted from the DM.

Appendix B
Supplemental Results for the "Rate Interpretation" of α

This section contains the equivalent results of Figure 1 and Table 1, except with the analysis done under the "rate interpretation" of α. They are presented in Figure B1 and Table B1. The logic in the following few paragraphs was originally presented by Macquart & Ekers (2018), and are presented again here to emphasize how the equations change between the "rate interpretation" of α and the "true spectral index interpretation."

The flux density in the frame of the observer is related to the luminosity in the frame of the emitting source by the following relation

$$S_\nu = (1+z)\frac{L_{(1+z)\nu}}{L_\nu}\frac{L_\nu}{4\pi D_L^2}, \quad \text{(B1)}$$

where D_L is the luminosity distance and the ratio $L_{(1+z)\nu}/L_\nu$ represents the k-correction, necessary because the radiation was

Table B1
Table of Results of the Parameters Fit in $R(F, \mathrm{DM})$ Assuming the "Rate Interpretation" of α

Parameter	Uniform Prior Range	Best-fit Result
$\log_{10}\Phi'_0$[a]	$[-0.96, 6.43]$	$4.83^{+0.35}_{-0.33}$
Φ'_0	…	$6.8^{+8.3}_{-3.6}\times10^4$ Gpc^{-3} yr^{-1}
γ'	$[-2.50, 2.00]$	$-1.3^{+0.8}_{-0.5}$
$\log_{10} E_{\rm char}{}'$[a]	[38.00, 49.00]	$41.29^{+0.47}_{-0.46}$
$E'_{\rm char}$	…	$1.95^{+3.75}_{-1.28}\times10^{41}$ erg
n'	$[-2.00, 8.00]$	$1.72^{+1.48}_{-1.10}$
α'	$[-5.00, 5.00]$	$-1.10^{+0.67}_{-0.99}$
$\mu'_{\rm host}\times\log_{10}(e)$[a]	[1.30, 2.70]	$1.91^{+0.27}_{-0.38}$
Median $P(\mathrm{DM}')$	…	81^{+69}_{-48} pc cm^{-3}
$\sigma'_{\rm host}\times\log_{10}(e)$[a]	[0.04, 1.74]	$0.41^{+0.22}_{-0.21}$
Std. dev. $P(\mathrm{DM}')$	…	174^{+347}_{-130} pc cm^{-3}

Notes. Best-fit results quote the median values of the posterior distributions, with the error bars containing the central 68% of the samples. Where illustrative, the fits are paired with the corresponding physical values.
[a] These parameters were fit in natural log space but are presented in $\log_{10}$ space for ease of interpretation.

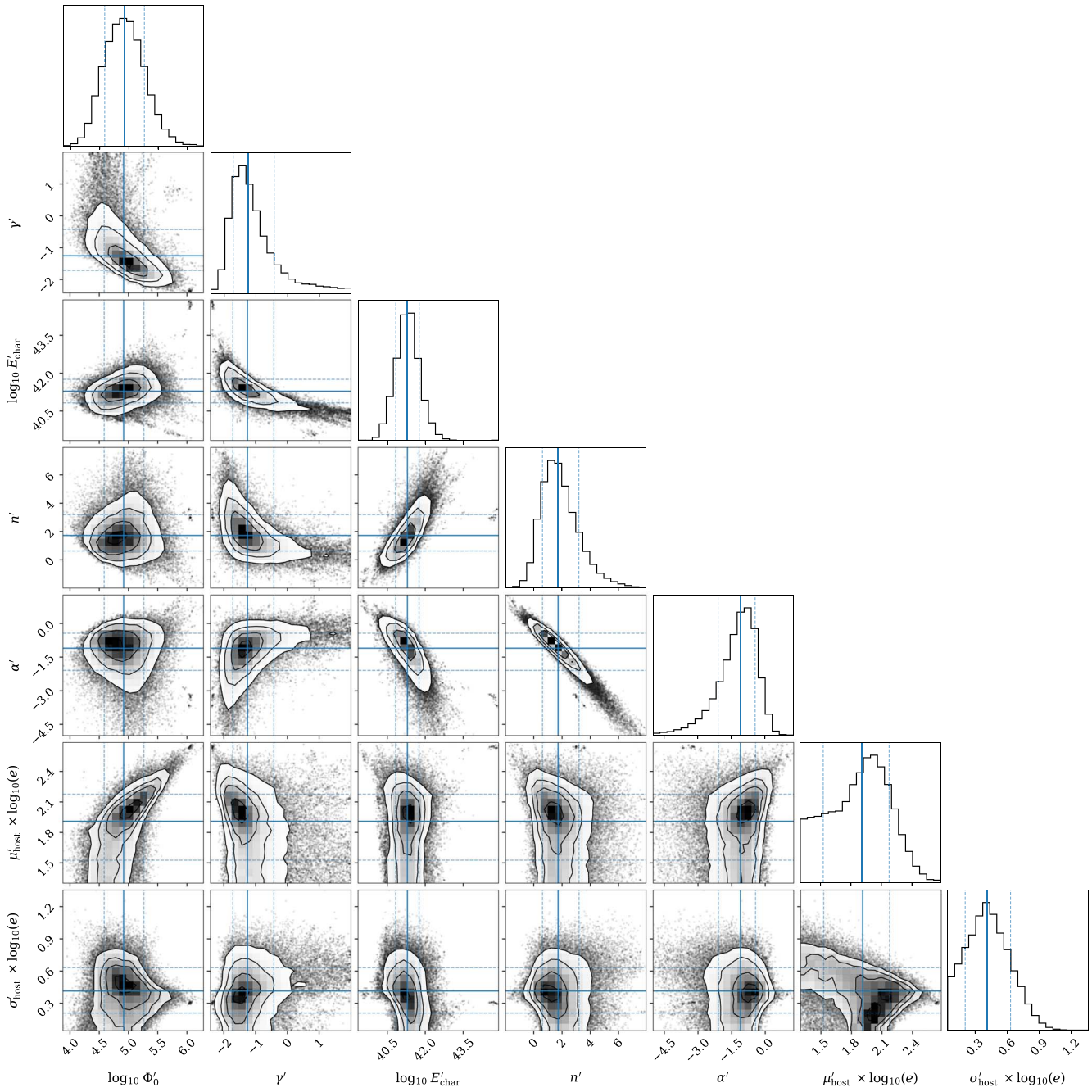

Figure B1. Corner plot of the results of the MCMC run under the "rate interpretation" of α, thinned by a factor of 20 for visualization purposes. The overlaid solid blue lines denote the median of the posterior distributions, with the dashed blue lines enclosing the central 68% of the samples.

emitted in a different band than in which it was observed. Since $S_\nu \sim \nu^\alpha$, this k-correction factor is $(1+z)^\alpha$.

The observed fluence and emitted luminosity are related via the following relation (Marani & Nemiroff 1996)

$$F_\nu = \Delta t_e (1+z)^2 \frac{L_{(1+z)\nu}}{L_\nu} \frac{L_\nu}{4\pi D_L^2}, \tag{B2}$$

where Δt_e represents the burst duration in the emission frame. Given that $E_\nu = \int L_\nu(t) dt$, and we are only considering E and F within a frequency bandwidth $\Delta\nu$, we can rewrite the above relation as

$$F = (1+z)^2 (1+z)^\alpha \frac{E}{4\pi D_L^2 \Delta\nu}. \tag{B3}$$

With some rearranging, this equation is equivalent to Equation (A4). This factor of α is explicitly due to the k-correction from the "true spectral index interpretation" of α, i.e., $F \sim \nu^\alpha$.

Under the "rate interpretation" of α, the k-correction instead comes into play by directly modifying the rate. As mentioned

in Section 5.4, Equation (A8) is modified such that it becomes

$$\Phi(z) = \frac{\Phi_0}{1+z}\left(\frac{\mathrm{SFR}(z)}{\mathrm{SFR}(0)}\right)^n (1+z)^{\alpha}. \tag{B4}$$

Additionally, Equation (A4) is now

$$E = \frac{4\pi D_L^2(z)}{(1+z)^2}\Delta\nu F. \tag{B5}$$

It is worth noting that although the parameter values of n and α changed under this interpretation, the constraints on both values were not very powerful to begin with. We present these results for completeness regarding systematic modeling uncertainties.

Appendix C
Testing an FRB Evolution Model Delayed with Respect to the SFR

In our fiducial $R(F, \mathrm{DM})$ model, the evolution of the FRB population smoothly scales with the SFR as a power-law parameterized by n

$$\Phi(z) = \frac{\Phi_0}{1+z}\left(\frac{\mathrm{SFR}(z)}{\mathrm{SFR}(0)}\right)^n. \tag{C1}$$

Despite the poor constraints on n, we decide to explore whether our data can constrain the FRB population in a toy model where the rate of FRBs strictly lags behind the SFR by a delay parameter τ (Gyr). Such a model is represented by

$$\Phi(z) = \frac{\Phi_0}{1+z}\frac{\mathrm{SFR}(z_a(t_{\mathrm{MW}}(z) - \tau))}{\mathrm{SFR}(0)}, \tag{C2}$$

where z_a is a function that obtains the redshift at a given lookback time. Of course, it is not physical to expect that for each star that forms, an FRB occurs exactly τ Gyr later. This toy model is intended to be for illustrative purposes only, as its key purpose is to investigate whether our data have any constraining power for a more complex FRB population evolution model. Such a more realistic model would include a distribution of delay times, as in Zhang & Zhang (2022). If this toy model shows no constraining power, then a more complex model would almost certainly have no constraining power as well. Thus, we run this toy model first with the same procedure as in Sections 4 and 5, with uniform priors on τ from 0 to 14 Gyr. The results of the MCMC are plotted in Figure C1.

We immediately see that τ is even more poorly constrained than n; in other words, with this toy model, our data still hold no constraining power for the FRB evolution with SFR. The posterior distribution for τ appears bimodal, with peaks at both a negligible ($\sim$0 Gyr) time delay and a significant ($\sim$12 Gyr) time delay. Of our 10^6 chains, $\approx$25.8% have $\tau < 2$ Gyr and $\approx$13.3% have 12 Gyr $<$ 14 Gyr (the rest of the parameter results vary across the chains). Although it is tempting to ascribe significance to this bimodality, we caution against overinterpretation for a number of reasons. First, the middle of the posterior distribution does not dip down to negligible probabilities, mitigating the bimodal appearance of this distribution. Second, and more importantly, this is an unphysical toy model whose primary purpose was to explore whether the Catalog 1 data could constrain a model where FRBs as a population are delayed with respect to the SFH of the universe. We conclude that the Catalog 1 data cannot constrain such a model.

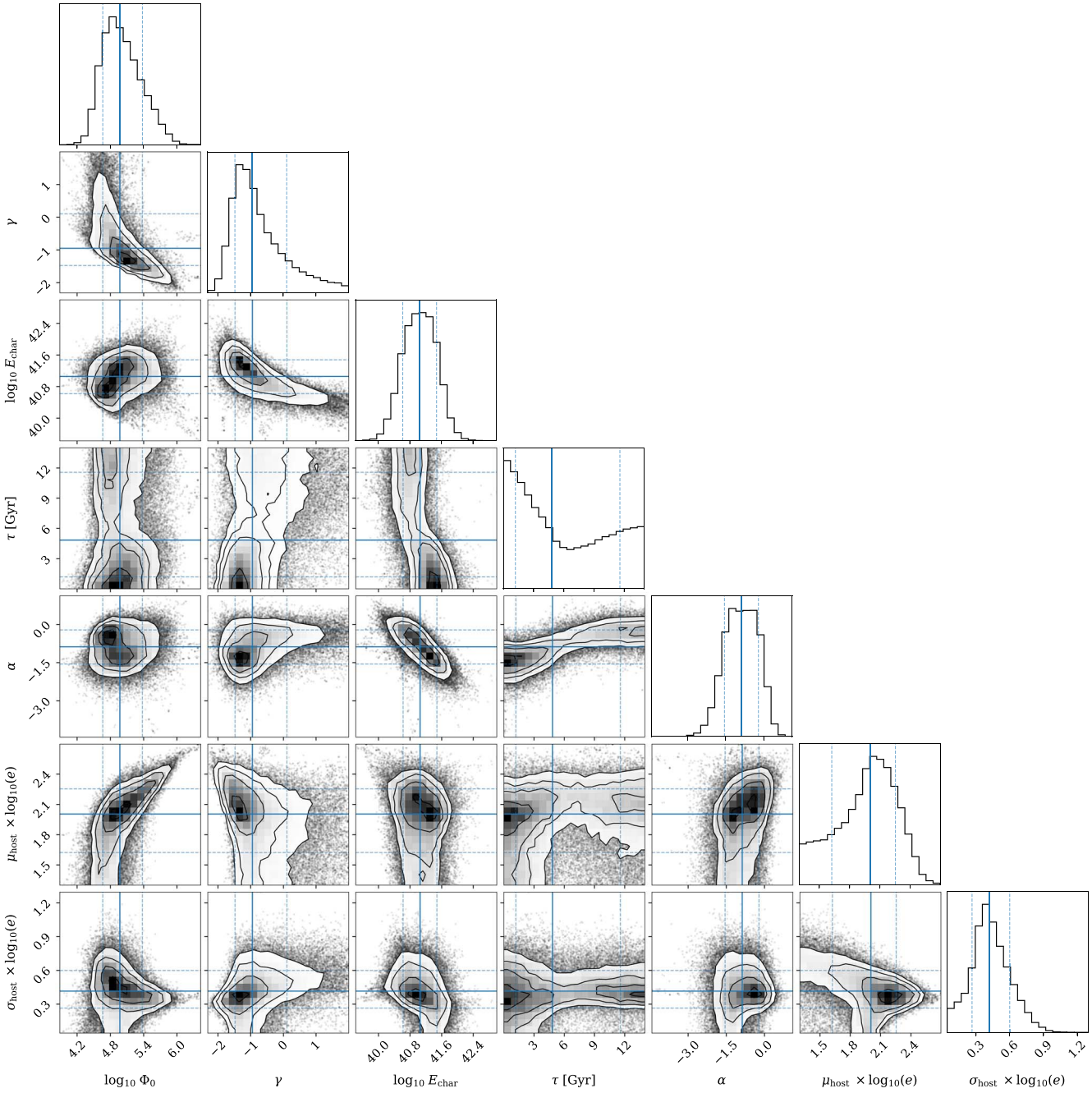

Figure C1. Corner plot of the results of the MCMC run using a time-delay toy model, thinned by a factor of 20 for visualization purposes. The overlaid solid blue lines denote the median of the posterior distributions, with the dashed blue lines enclosing the central 68% of the samples.

ORCID iDs

Kaitlyn Shin https://orcid.org/0000-0002-6823-2073
Kiyoshi W. Masui https://orcid.org/0000-0002-4279-6946
Mohit Bhardwaj https://orcid.org/0000-0002-3615-3514
Tomas Cassanelli https://orcid.org/0000-0003-2047-5276
Pragya Chawla https://orcid.org/0000-0002-3426-7606
Matt Dobbs https://orcid.org/0000-0001-7166-6422
Fengqiu Adam Dong https://orcid.org/0000-0003-4098-5222
Emmanuel Fonseca https://orcid.org/0000-0001-8384-5049
B. M. Gaensler https://orcid.org/0000-0002-3382-9558
Antonio Herrera-Martín https://orcid.org/0000-0002-3654-4662
Jane Kaczmarek https://orcid.org/0000-0003-4810-7803
Victoria Kaspi https://orcid.org/0000-0001-9345-0307
Calvin Leung https://orcid.org/0000-0002-4209-7408
Marcus Merryfield https://orcid.org/0000-0003-2095-0380
Daniele Michilli https://orcid.org/0000-0002-2551-7554
Moritz Münchmeyer https://orcid.org/0000-0002-3777-7791
Aaron B. Pearlman https://orcid.org/0000-0002-8912-0732
Masoud Rafiei-Ravandi https://orcid.org/0000-0001-7694-6650
Kendrick Smith https://orcid.org/0000-0002-2088-3125
Ingrid Stairs https://orcid.org/0000-0001-9784-8670
Shriharsh P. Tendulkar https://orcid.org/0000-0003-2548-2926

References

Ai, S., Gao, H., & Zhang, B. 2021, ApJL, 906, L5
Akahori, T., Ryu, D., & Gaensler, B. M. 2016, ApJ, 824, 105
Bannister, K. W., Deller, A. T., Phillips, C., et al. 2019, Sci, 365, 565
Bhandari, S., Heintz, K. E., Aggarwal, K., et al. 2022, AJ, 163, 69
Bhandari, S., Sadler, E. M., Prochaska, J. X., et al. 2020, ApJL, 895, L37
Bhardwaj, M., Gaensler, B. M., Kaspi, V. M., et al. 2021, ApJL, 910, L18
Bochenek, C. D., Ravi, V., Belov, K. V., et al. 2020, Natur, 587, 59
Caleb, M., Flynn, C., Bailes, M., et al. 2016, MNRAS, 458, 708
Caleb, M., Flynn, C., Bailes, M., et al. 2017, MNRAS, 468, 3746
Caleb, M., & Keane, E. 2021, Univ., 7, 453
Caleb, M., Stappers, B. W., Rajwade, K., & Flynn, C. 2019, MNRAS, 484, 5500
Cary, S., Mena-Parra, J., Leung, C., et al. 2021, RNAAS, 5, 216
Cassanelli, T., Leung, C., Rahman, M., et al. 2022, AJ, 163, 65
Chawla, P., Kaspi, V. M., Ransom, S. M., et al. 2022, ApJ, 927, 35
Chen, B. H., Hashimoto, T., Goto, T., et al. 2022, MNRAS, 509, 1227
CHIME Collaboration, Amiri, M., Bandura, K., et al. 2022, ApJS, 261, 29
CHIME/FRB Collaboration, Amiri, M., Andersen, B. C., et al. 2020a, Natur, 582, 351
CHIME/FRB Collaboration, Amiri, M., Andersen, B. C., et al. 2021, ApJS, 257, 59
CHIME/FRB Collaboration, Amiri, M., Bandura, K., et al. 2018, ApJ, 863, 48
CHIME/FRB Collaboration, Amiri, M., Bandura, K., et al. 2019, ApJL, 885, L24
CHIME/FRB Collaboration, Andersen, B. C., Bandura, K. M., et al. 2020b, Natur, 587, 54
Cordes, J. M., & Chatterjee, S. 2019, ARA&A, 57, 417
Cook, A. M., Bhardwaj, M., Gaensler, B. M., et al. 2023, arXiv:2301.03502
Cordes, J. M., & Lazio, T. J. W. 2002, arXiv:astro-ph/0207156
Cui, X.-H., Zhang, C.-M., Wang, S.-Q., et al. 2021, MNRAS, 500, 3275
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
Dolag, K., Gaensler, B. M., Beck, A. M., & Beck, M. C. 2015, MNRAS, 451, 4277
Fialkov, A., Loeb, A., & Lorimer, D. R. 2018, ApJ, 863, 132
Fonseca, E., Andersen, B. C., Bhardwaj, M., et al. 2020, ApJL, 891, L6
Foreman-Mackey, D. 2016, JOSS, 1, 24
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Gardenier, D. W., Connor, L., van Leeuwen, J., Oostrum, L. C., & Petroff, E. 2021, A&A, 647, A30
Gohar, N., & Flynn, C. 2022, MNRAS, 509, 5265
Good, D. C., Chawla, P., Fonseca, E., et al. 2022, arXiv:2204.09090
Guo, H.-Y., & Wei, H. 2022, JCAP, 2022, 010
Hashimoto, T., Goto, T., Chen, B. H., et al. 2022, MNRAS, 511, 1961
Heintz, K. E., Prochaska, J. X., Simha, S., et al. 2020, ApJ, 903, 152
Hogg, D. W., & Foreman-Mackey, D. 2018, ApJS, 236, 11
James, C. W., Osłowski, S., Flynn, C., et al. 2020, MNRAS, 495, 2416
James, C. W., Prochaska, J. X., Macquart, J. P., et al. 2022a, MNRAS, 510, L18
James, C. W., Prochaska, J. X., Macquart, J. P., et al. 2022b, MNRAS, 509, 4775
Josephy, A., Chawla, P., Fonseca, E., et al. 2019, ApJL, 882, L18
Kader, Z., Leung, C., Dobbs, M., et al. 2022, PhRvD, 106, 043016
Keating, L. C., & Pen, U.-L. 2020, MNRAS, 496, L106
Kirsten, F., Marcote, B., Nimmo, K., et al. 2022, Natur, 602, 585
Kirsten, F., Snelders, M. P., Jenkins, M., et al. 2021, NatAs, 5, 414
Kremer, K., Piro, A. L., & Li, D. 2021, ApJL, 917, L11
Lang, K. R. 1971, ApJ, 164, 249
Lanman, A. E., Andersen, B. C., Chawla, P., et al. 2022, ApJ, 927, 59
Law, C. J., Abruzzo, M. W., Bassa, C. G., et al. 2017, ApJ, 850, 76
Leung, C., Kader, Z., Masui, K. W., et al. 2022, PhRvD, 106, 043017
Leung, C., Mena-Parra, J., Masui, K., et al. 2021, AJ, 161, 81
Li, D., Wang, P., Zhu, W. W., et al. 2021, Natur, 598, 267
Li, Y., & Zhang, B. 2020, ApJL, 899, L6
Linder, E. V. 2020, PhRvD, 101, 103019
Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, Sci, 318, 777
Lorimer, D. R., & Kramer, M. 2012, Handbook of Pulsar Astronomy (Cambridge: Cambridge Univ. Press)
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., et al. 2020, MNRAS, 494, 665
Macquart, J. P., & Ekers, R. 2018, MNRAS, 480, 4211
Macquart, J. P., Prochaska, J. X., McQuinn, M., et al. 2020, Natur, 581, 391
Macquart, J. P., Shannon, R. M., Bannister, K. W., et al. 2019, ApJL, 872, L19
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Madhavacheril, M. S., Battaglia, N., Smith, K. M., & Sievers, J. L. 2019, PhRvD, 100, 103532
Mannings, A. G., Fong, W.-f., Simha, S., et al. 2021, ApJ, 917, 75
Marani, G. F., & Nemiroff, R. S. 1996, in ASP Conf. Ser. 88, Clusters, Lensing, and the Future of the Universe, ed. V. Trimble & A. Reisenegger (San Francisco, CA: ASP), 107
Marcote, B., Nimmo, K., Hessels, J. W. T., et al. 2020, Natur, 577, 190
Masui, K., Lin, H.-H., Sievers, J., et al. 2015, Natur, 528, 523
Masui, K. W., & Sigurdson, K. 2015, PhRvL, 115, 121301
McQuinn, M. 2014, ApJL, 780, L33
Mena-Parra, J., Leung, C., Cary, S., et al. 2022, AJ, 163, 48
Merryfield, M., Tendulkar, S. P., Shin, K., et al. 2022, arXiv:2206.14079
Nicholl, M., Williams, P. K. G., Berger, E., et al. 2017, ApJ, 843, 84
Nimmo, K., Hewitt, D. M., Hessels, J. W. T., et al. 2022, ApJL, 927, L3
Niu, C. H., Aggarwal, K., Li, D., et al. 2022, Natur, 606, 873
Niu, C.-H., Li, D., Luo, R., et al. 2021b, ApJL, 909, L8
Oguri, M. 2019, RPPh, 82, 126901
Petroff, E., Barr, E. D., Jameson, A., et al. 2016, PASA, 33, e045
Petroff, E., Hessels, J. W. T., & Lorimer, D. R. 2022, A&ARv, 30, 2
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Pleunis, Z., Good, D. C., Kaspi, V. M., et al. 2021, ApJ, 923, 1
Prochaska, J. X., & Zheng, Y. 2019, MNRAS, 485, 648
Qiang, D.-C., Li, S.-L., & Wei, H. 2022, JCAP, 2022, 040
Rafiei-Ravandi, M., Smith, K. M., Li, D., et al. 2021, ApJ, 922, 42
Rajwade, K. M., Mickaliger, M. B., Stappers, B. W., et al. 2020, MNRAS, 495, 3551
Ravi, V. 2019, NatAs, 3, 928
Ryder, S. D., Bannister, K. W., Bhandari, S., et al. 2022, arXiv:2210.04680
Schechter, P. 1976, ApJ, 203, 297
Scholz, P., Spitler, L. G., Hessels, J. W. T., et al. 2016, ApJ, 833, 177
Shannon, R. M., Macquart, J. P., Bannister, K. W., et al. 2018, Natur, 562, 386
Spitler, L. G., Scholz, P., Hessels, J. W. T., et al. 2016, Natur, 531, 202
Swat, M. J., Grenon, P., & Wimalaratne, S. 2016, Bioinf., 32, 2719
Tendulkar, S. P., Gil de Paz, A., Kirichenko, A. Y., et al. 2021, ApJL, 908, L12
The LIGO Scientific Collaboration, the Virgo Collaboration, the KAGRA Collaboration, et al. 2021, arXiv:2111.03634
Thornton, D., Stappers, B., Bailes, M., et al. 2013, Sci, 341, 53
Wilks, S. S. 1938, Ann. Math. Stat., 9, 60
Yao, J. M., Manchester, R. N., & Wang, N. 2017, ApJ, 835, 29
Zhang, R. C., & Zhang, B. 2022, ApJL, 924, L14
Zheng, Z., Ofek, E. O., Kulkarni, S. R., Neill, J. D., & Juric, M. 2014, ApJ, 797, 71
Zhou, B., Li, X., Wang, T., Fan, Y.-Z., & Wei, D.-M. 2014, PhRvD, 89, 107303
Zyla, P. A., Barnett, R. M., Beringer, J., et al. 2020, PTEP, 2020, 083C01