

*This article has been accepted for publication in Monthly Notices of the Royal Astronomical Society ©: 2020 The Authors. Published by Oxford University Press on behalf of the Royal Astronomical Society. All rights reserved.*

# Joint inference on the redshift distribution of fast radio burst and on the intergalactic baryon content

S. Hackstein,<sup>1</sup>★ M. Brüggen<sup>1</sup> and F. Vazza<sup>1,2,3</sup>

<sup>1</sup>*Hamburger Sternwarte, University of Hamburg, Gojenbergsweg 112, D-21029 Hamburg, Germany*

<sup>2</sup>*Department of Physics and Astronomy, University of Bologna, Via Gobetti 93/2, I-40129 Bologna, Italy*

<sup>3</sup>*Istituto di Radioastronomia, INAF, Via Gobetti 101, I-40129 Bologna, Italy*

Accepted 2020 December 18. Received 2020 December 17; in original form 2020 August 25

## ABSTRACT

Fast radio bursts are transient radio pulses of extragalactic origin. Their dispersion measure is indicative of the baryon content in the ionized intergalactic medium between the source and the observer. However, inference using unlocalized fast radio bursts is degenerate to the distribution of redshifts of host galaxies. We perform a joint inference of the intergalactic baryon content and the fast radio burst redshift distribution with the use of Bayesian statistics by comparing the likelihood of different models to reproduce the observed statistics in order to infer the most likely models. In addition to two models of the intergalactic medium, we consider contributions from the local environment of the source, assumed to be a magnetar, as well as a representative ensemble of host and intervening galaxies. Assuming that the missing baryons reside in the ionized intergalactic medium, our results suggest that the redshift distribution of observed fast radio bursts peaks at  $z \lesssim 0.6$ . However, conclusions from different instruments regarding the intergalactic baryon content diverge and thus require additional changes to the observed distribution of host redshifts, beyond those caused by telescope selection effects.

**Key words:** galaxies: distances and redshifts – galaxies: intergalactic medium – cosmology: large-scale structure of Universe – cosmology: observations – radio continuum: transients – transients: fast radio bursts .

## 1 INTRODUCTION

Observations of the cosmic microwave background (CMB) show that  $\approx 5$  per cent of energy density in the Universe exists as ordinary matter (Planck Collaboration I 2014). However, in the  $z \leq 2$  range, observations of stars and gas in galaxies, of the hot intracluster medium, and of the Ly  $\alpha$  forest can only account for about half of that amount (e.g. Nicastro, Mathur & Elvis 2008). The ‘missing’ baryons are believed to be hidden in the warm–hot intergalactic medium (WHIM; Cen & Ostriker 1999; Dave et al. 2001), with temperatures of  $10^5$  to  $10^7$  K and low baryon densities of  $10^{-6}$  to  $10^{-5}$  cm $^{-3}$ . This medium is difficult to observe directly as it only couples with radiation through electronic transitions. However, correlation of the distribution of galaxies with CMB distortions due to the Sunyaev–Zel’dovich effect has revealed the likely presence of filaments of warm–hot gas between galaxies, possibly consistent with the missing baryons (e.g. Tanimura et al. 2018; de Graaff et al. 2019).

Several studies have suggested that fast radio bursts (FRBs) can be used to detect the missing baryons. Their signals propagate across cosmological distances and get dispersed by diffuse ionized gas, quantified by the dispersion measure (DM), defined as the column density of free electrons:

$$DM = \int n_e dl. \quad (1)$$

Wei et al. (2019) proposed to compare the redshift evolution of the DMs using  $\approx 3000$  FRBs with known redshifts to that of the Hubble parameter to measure the baryon content  $f_{\text{IGM}}$  in the intergalactic medium (IGM), i.e. the amount of all baryons that reside in the ionized IGM. Deng & Zhang (2014) argued that the redshift observed for  $\gamma$ -ray bursts associated with FRBs can be used to infer  $f_{\text{IGM}}$ . McQuinn (2013) modelled the distribution of DM expected for FRBs from different redshifts and found that  $\sim 100$  DMs of FRBs from redshift  $z > 0.5$ , localized with subarcminute precision around identified galaxies, can be used to infer the baryon profile of galaxies in order to constrain the localization of cosmic baryons. Muñoz & Loeb (2018) proposed to cross-correlate the DM of arcminute-localized FRBs with the thermal Sunyaev–Zel’dovich effect in the CMB. The latter depends on the temperature of the WHIM and the amount of baryons localized in the WHIM, which can be constrained using  $\sim 1000$  FRBs with arcminute localization. Qiang & Wei (2020) suggest to use model-independent Gaussian processes to investigate the evolution of  $f_{\text{IGM}}$  with DM of FRBs with identified redshift.

All of the aforementioned papers require a large number of well-localized FRBs or cross-correlation with associated observations. However, Yang, Li & Zhang (2020) argue that the majority of FRBs are not supposed to have strong associated persistent sources. Furthermore, the exact localization of sources of short-duration signals without known redshift is not trivial to constrain and neither is the identification of the host galaxies of FRBs (Eftekhari & Berger 2017; Mahony et al. 2018; Marcote & Paragi 2019; Prochaska et al. 2019a). The host can be identified using interferometry or by observing persistent counterparts to the FRB. So far, FRBs have been located

\* E-mail: shackste@physnet.uni-hamburg.de

to a multitude of galaxy types, from star-forming dwarf galaxies (Tendulkar et al. 2017) to very massive galaxies with old stellar population (Bannister et al. 2019; Ravi et al. 2019). Unluckily, the small sample of localized FRBs cannot provide enough information to arrive at reasonable conclusions regarding cosmological questions. However, in this work, we show how it is possible to use DM of a large sample of unlocalized FRBs, in order to infer their host redshift distribution in a statistical way, which also allows us to constrain the amount of ionized baryons located in the cosmic web.

Because of the unknown distance to unlocalized FRBs, the assumed distribution of source redshifts can help interpret the distribution of observed DM. Several papers have tried to infer the intrinsic redshift distribution of FRBs, either by modelling the distribution of DM and other FRB properties with analytical or Monte Carlo methods (Bera et al. 2016; Caleb et al. 2016; Gardenier et al. 2019) or by performing a luminosity–volume test (Locatelli et al. 2018). So far, there are inconsistencies in implications of data from different instruments, e.g. ASKAP requiring faster change in FRB density with redshift than Parkes.

Macquart & Ekers (2018) use the flux density and fluence of FRBs to infer the luminosity distance and evolutionary history, as well as the redshift distribution, and show how the history of the ionized IGM affects the distribution of observed DM, and thus influences our inference of the host redshift distribution.

Here, we propose to use unlocalized FRBs to perform a joint analysis of the FRB redshift distribution and the IGM baryon content,  $f_{\text{IGM}}$ , by comparing the expected distribution of DM to the available observed values reported by Parkes, CHIME, and ASKAP observatories (Staveley-Smith et al. 1996; Macquart et al. 2010; The CHIME/FRB Collaboration 2018), using the PREFRBLE<sup>1</sup> software (Hackstein 2020), presented in Hackstein et al. (2020). In Section 2, we explain how we model different values of  $f_{\text{IGM}}$ . The resulting predictions are presented and compared in Section 3. We discuss our results in Section 4 and conclude in Section 5.

## 2 METHOD

Investigation of the fraction of baryons in the ionized IGM,  $f_{\text{IGM}} = \Omega_{\text{IGM}}/\Omega_{\text{baryons}}$ , with DM requires comparison of the observed DM with expectations of  $\text{DM}(z|f_{\text{IGM}})$ , according to source redshift  $z$  (Keane et al. 2016):

$$\langle \text{DM}(z) \rangle = \frac{c\rho_{\text{crit}}\Omega_{\text{b}}}{m_{\text{p}}\mu_{\text{e}}H_0} \int f_{\text{IGM}}(z) \frac{(1+z)}{H(z)} dz. \quad (2)$$

However, FRBs do not come with a direct measure of redshift and for the majority of FRBs, which could not be localized by other means, the DM is the best indicator for the source distance (Dolag et al. 2015; Luo et al. 2018; Niino 2018; Walker, Ma & Breton 2018; Pol et al. 2019). Still, by assuming different plausible redshift distributions of FRBs,  $\pi(z)$ , we can estimate the distribution of DM to be observed by instruments, which is determined by  $f_{\text{IGM}}$ .

By comparing the expected distribution to observations, we can quantify the likelihood of different combinations of  $\pi(z)$  and  $f_{\text{IGM}}$ . Consequently, this allows us to put constraints on the WHIM density without the need to localize FRBs. Furthermore, through a detailed investigation of the redshift evolution of  $f_{\text{IGM}}$  in cosmological simulations, the inference presented here can be used to investigate the helium-reionization history (e.g. Dai & Xia 2020; Linder 2020).

In Hackstein et al. (2020), we obtained likelihood estimates,  $L(\text{DM}|z)$ , to observe extragalactic DM from source at redshift  $z$ . These expectations consider contributions from all regions along the lines-of-sight in our benchmark scenario, considering magnetars as the source of FRBs (Piro & Gaensler 2018) as well as a realistic ensemble of different host and intervening galaxies (Lacey et al. 2016; Rodrigues et al. 2018). Predictions for the IGM were obtained using a constrained cosmological simulation of the local Universe, produced using the cosmological magnetohydrodynamic code ENZO (Bryan et al. 2014) together with initial conditions obtained following Sorce et al. (2016) and cosmological parameters of Planck (Ade et al. 2016). The constrained volume of  $250 (\text{Mpc h}^{-1})^3$  was embedded in a  $500 (\text{Mpc h}^{-1})^3$  grid of  $1024^3$  cells. The use of five adaptive mesh refinement levels allowed for a maximum resolution of 30 kpc in the most dense environments. The simulation starts at redshift  $z = 60$ , where all baryons are in the IGM. However, the limited resolution does not allow us to properly resolve galaxy formation and the condensation of cold gas out of the IGM thus implies  $f_{\text{IGM}} = 1$ , always. Further information on this model can be found in Hackstein et al. (2018, 2019). A reduced version of this model can be found on [crpropa.desy.de](http://crpropa.desy.de) under ‘additional resources’, together with the other models presented in Hackstein et al. (2018).

The contribution of the circumburst environment of magnetars is modelled using the analytical predictions of the wind model in Piro & Gaensler (2018) in a Monte Carlo simulation. More details on this simulation and the assumed priors can be found in Hackstein et al. (2019). The contribution of host and intervening galaxies is estimated by computing line-of-sight integrals in a realistic ensemble of galaxy models that recreate a wide variety of properties of the population of galaxies in the Universe (Lacey et al. 2016; Rodrigues et al. 2018). More details on this Monte Carlo simulation and the resulting likelihoods can be found in Hackstein et al. (2020).

We further use FRBPOPPY<sup>2</sup> (Gardenier et al. 2019) to model the distribution of host redshift of observed FRBs,  $\pi(z)$ . We assume three physically motivated intrinsic redshift distributions, i.e. following the comoving volume (coV), stellar mass density (SMD), or star formation rate (SFR), together with the selection effects of ASKAP, CHIME, and Parkes instruments. We use the parameters of their *complex* model, which best reproduces observations.

Together, these can be used to quantify the distribution of expected DMs,

$$L(\text{DM}) = \int L(\text{DM}, z) \times \pi(z) dz. \quad (3)$$

For more details on the derivation of likelihood  $L(\text{DM}|z)$  and prior  $\pi(z)$ , see Hackstein et al. (2020).

Numerical simulations suggest that up to  $28 \pm 11$  per cent of baryons are located in neutral hydrogen clouds in the IGM, observable in the photoionized Ly $\alpha$  forest, as well as  $25 \pm 8$  per cent in the WHIM, leaving a baryon shortfall of  $29 \pm 12$  per cent (Shull, Smith & Danforth 2012). This leaves a plausible range of  $0.3 \lesssim f_{\text{IGM}} \lesssim 0.9$  (cf. Li et al. 2019).

As explained above, in the constrained simulation  $f_{\text{IGM}}$  is assumed to be constant with redshift, thus to be a global factor to  $\langle \text{DM}(z) \rangle$  (equation 2). This allows us to obtain estimates for different values of  $f_{\text{IGM}}$  in post-processing, allowing us to test how many missing baryons can be located in the WHIM (Deng & Zhang 2014; Keane et al. 2016). To this end, we modify the likelihood of intergalactic contribution,  $L(\text{DM}_{\text{IGM}}|z)$ , obtained from the constrained simulation

<sup>1</sup>[github.com/FRBs/PreFRBLE](https://github.com/FRBs/PreFRBLE)

<sup>2</sup>[github.com/davidgardenier/frbpoppy](https://github.com/davidgardenier/frbpoppy)

of the IGM, such that the result returns the required average value according to equation (2), while shape and norm are conserved:

$$L(\text{DM}_{\text{IGM}}|f_{\text{IGM}}) = f_{\text{IGM}} \times L(f_{\text{IGM}} \times \text{DM}_{\text{IGM}}). \quad (4)$$

The fraction of baryons in the IGM,  $f_{\text{IGM}}$ , is believed to grow with redshift (McQuinn 2013; Prochaska & Zheng 2019). Simulations suggest that  $f_{\text{IGM}} \approx 0.9$  prior to  $z = 1.5$  (Meiksin 2009). To account for the redshift evolution of  $f_{\text{IGM}}$ , we consider a model for the IGM following Pshirkov, Tinyakov & Urban (2016) (called **Pshirkov16** hereafter). We perform a Monte Carlo simulation in which we divide the line-of-sight into segments of Jeans-length size  $\lambda_J(z) = 2.3 (1+z)^{-1.5}$  Mpc. Assuming that the electron density  $n_e$  is well described by the distribution of neutral hydrogen deduced from the observed Lyman- $\alpha$  forest, we pick random values for comoving  $n_e$  drawn from a lognormal distribution of overdensity according to equation (2) in Pshirkov et al. (2016). Finally,  $n_e$  is multiplied by the value of  $f_{\text{IGM}}$  at corresponding redshift. We assume  $f_{\text{IGM}}$  to increase linearly with redshift from  $f_{\text{IGM}} = f_{\text{IGM},0}$  at  $z = 0$  to  $f_{\text{IGM}} = 0.9$  at  $z = 1.5$ , while it stays at this constant value at higher redshifts.

In order to quantify the likelihood of a scenario of  $f_{\text{IGM}}$  and FRB redshift distribution  $\pi(z)$  to account for the observation of an individual FRB with  $\text{DM}_{\text{obs}} = \text{DM}_{\text{EG}} + \text{DM}_{\text{MW}}$ , we take the value of  $L(\text{DM}_{\text{EG}}|f_{\text{IGM}})$  for the benchmark scenario, using the estimates for  $\text{DM}_{\text{MW}}$  according to NE2001 model (Cordes & Lazio 2002) as listed in FRBcat (Petroff et al. 2016). These estimated values of  $\text{DM}_{\text{EG}}$  were shown to be correct to  $\approx 30 \text{ pc cm}^{-3}$  (Manchester et al. 2005).

In order to compare different choices of  $f_{\text{IGM}}$  and host redshift distribution, we compute the Bayes factor

$$\mathcal{B}(\text{DM}_{\text{EG}}|f_{\text{IGM},i}, f_{\text{IGM,ref}}) = \frac{L(\text{DM}_{\text{EG}}|f_{\text{IGM},i})}{L(\text{DM}_{\text{EG}}|f_{\text{IGM,ref}})} \quad (5)$$

by renormalizing all likelihoods of individual FRBs to the same reference choice of  $f_{\text{IGM,ref}}$ .  $\mathcal{B}$  quantifies how much the observation of  $\text{DM}_{\text{EG}}$  changes our corroboration in favour of  $f_{\text{IGM},i}$  compared to  $f_{\text{IGM,ref}}$ . By renormalizing all  $\mathcal{B}$  to the best-fitting scenario  $\mathcal{B}_{\text{max}}$ , the presented values quantify how much the observations change our corroboration in disfavour of the corresponding scenario.

For FRBs with localized hosts, the information on  $f_{\text{IGM}}$  and  $\pi(z)$  can be distinguished, thus allowing for separate likelihoods. For the events listed with host redshift  $z_{\text{FRB}}$  in FRBcat, instead of the procedure explained above, we obtain the likelihood of  $f_{\text{IGM}}$  by comparing values of  $L(\text{DM}_{\text{EG}}|z_{\text{FRB}}, f_{\text{IGM}})$  and compute the Bayes factor  $\mathcal{B}$  by renormalizing the likelihood to a particular choice of  $f_{\text{IGM},i}$ . Likewise, we compare values of  $\pi(z_{\text{FRB}})$  for their implication on the redshift distribution.

$\mathcal{B}$  obtained for individual events and different instruments can simply be multiplied in order to interpret the combined results. Note that  $\mathcal{B}$  quantifies how much the current observations change our corroboration between the considered scenarios. However, according to Bayes theorem

$$P(f_{\text{IGM}}|\text{DM}_{\text{EG}}) \propto L(\text{DM}_{\text{EG}}|f_{\text{IGM}})\pi(f_{\text{IGM}}), \quad (6)$$

the ratio of likelihoods  $L, \mathcal{B}$ , has to be multiplied by the ratio of priors  $\pi$  in order to arrive at the ratio of posteriors  $P$ .  $\pi(f_{\text{IGM}})$  quantifies how likely different values of  $f_{\text{IGM}}$  are prior to investigation of  $\text{DM}_{\text{EG}}$ , while  $P$  expresses our conclusion on the likelihood of values of  $f_{\text{IGM}}$  after consideration of the data. However, we consider no preference for any value of  $f_{\text{IGM}}$ ,  $\pi(f_{\text{IGM}}) = \text{const.}$ ; thus,  $\mathcal{B}$  is identical to the ratio of  $P$ .

### 3 RESULTS

#### 3.1 Expected distribution versus observation

In Fig. 1, we show the results for all combinations of instruments and redshift distributions for two extreme constant values of  $f_{\text{IGM}} = 0.3$  and 0.9, using the constrained simulation for the IGM. These can be compared directly to observations by the corresponding instrument, which are indicated by the dash-dotted line.

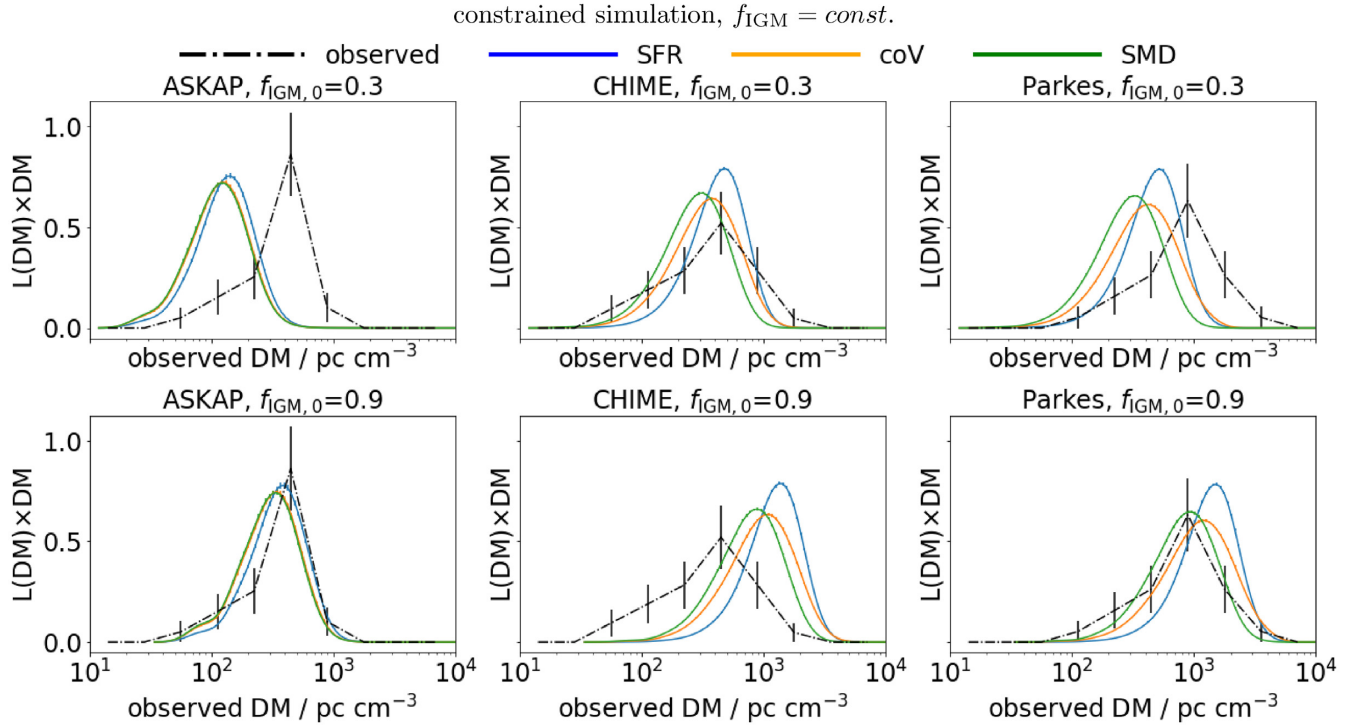
The majority of FRBs observed by ASKAP have  $\text{DM}_{\text{EG}} \approx 500 \text{ pc cm}^{-3}$ . The expected peak values of  $\text{DM}_{\text{EG}}$  for  $f_{\text{IGM}}$  between 0.3 and 0.9 range from 100 to  $440 \text{ pc cm}^{-3}$ . ASKAP results are thus in favour of high values of  $f_{\text{IGM}} \approx 0.9$ . However, for the values of DM observed with ASKAP that are generally higher than expectations, other foregrounds such as the local contributions need to be considered more carefully in order to arrive at reasonable conclusions. The assumed intrinsic redshift distributions do not differ much in their predictions at low  $z$  (cf. Hackstein et al. 2020), thus rendering ASKAP results incapable of distinguishing FRB populations with a small sample of FRBs.

While the distribution of DM observed by CHIME peaks at roughly the same value as for ASKAP  $\text{DM}_{\text{EG}} \approx 500 \text{ pc cm}^{-3}$ , the peak is less pronounced and the upper  $1\sigma$  deviation is more than 1.5 times higher than that for ASKAP, reaching values beyond  $1000 \text{ pc cm}^{-3}$ . However, for CHIME, the expected peak values of  $\text{DM}_{\text{EG}}$  range from 300 to  $1500 \text{ pc cm}^{-3}$  for  $f_{\text{IGM}} = 0.3$  and 0.9, respectively, far above the observed value, hence expecting many more high values of DM. Thus, the CHIME results favour low  $f_{\text{IGM}} \gtrsim 0.3$ , depending on the assumed distribution of the host redshifts. Also, CHIME observes more FRBs with  $50 < \text{DM}_{\text{EG}}/\text{pc cm}^{-3} < 200$  than ASKAP, which is not expected according to results in Fig. 1. This hints on an imprecise assumption on the redshift distribution of FRB sources visible to CHIME, which observes at different wavelengths than the other two instruments.

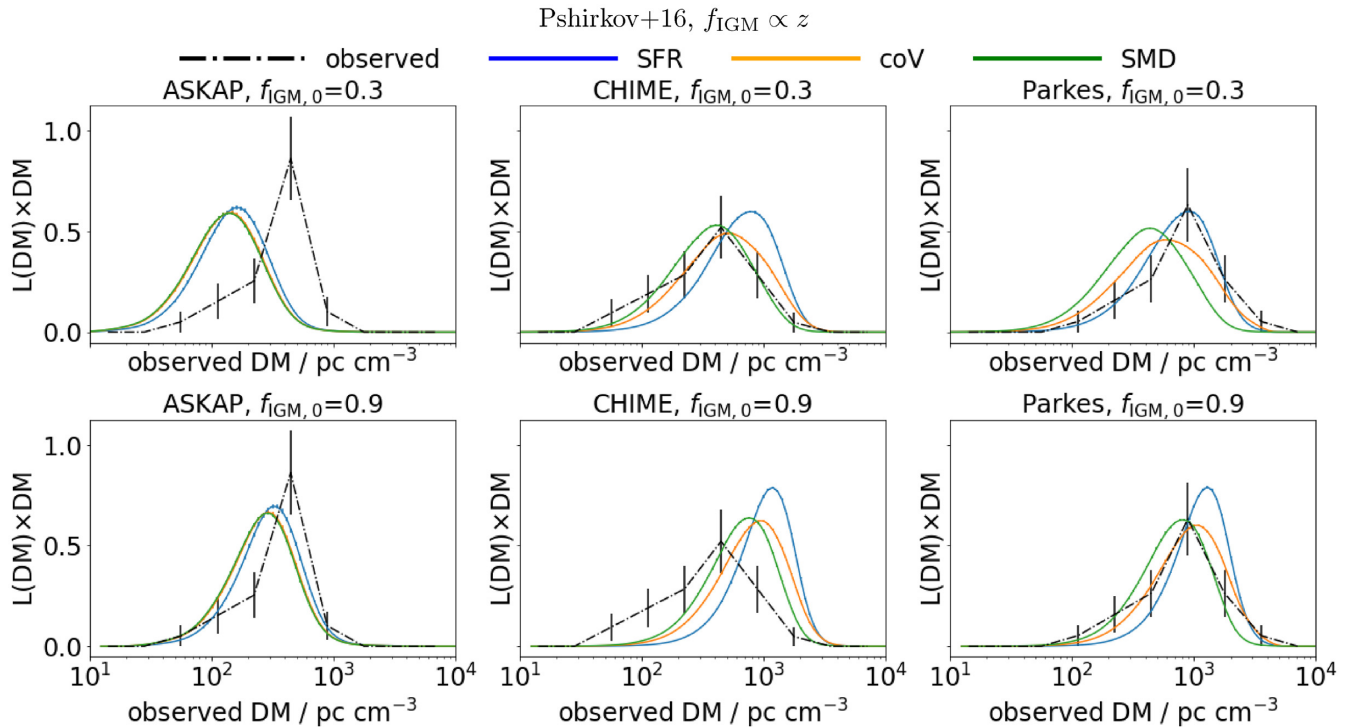
For Parkes, the expected peak value of  $\text{DM}_{\text{EG}}$  ranges from 300 to  $1800 \text{ pc cm}^{-3}$  between  $f_{\text{IGM}} = 0.3$  and 0.9, similar as expected for CHIME. However, the observed sample peaks at  $\text{DM}_{\text{EG}} \approx 1000 \text{ pc cm}^{-3}$  – about twice the value observed with CHIME – which is well within the expected range. Hence, different assumed redshift distributions will favour other values of  $f_{\text{IGM}}$ ; e.g. a population concentrated at low redshifts, such as SMD, will need a high value of  $f_{\text{IGM}} \lesssim 0.9$  to explain the observed DMs, while populations with greater redshifts, such as SFR, will require lower values of  $f_{\text{IGM}}$ . The Parkes sample is dominated by high values of DMs, with  $\approx 30$  per cent of FRBs observed with  $\text{DM} > 1000 \text{ pc cm}^{-3}$ . The majority of these are thus likely located at  $z \gtrsim 1$ , such that local contributions to the DMs are less significant. This makes the Parkes results overall the most reliable for investigation of  $f_{\text{IGM}}$  and other cosmological properties.

In Fig. 2, we present the results for all combinations of instruments and redshift distributions for two extreme values of  $f_{\text{IGM},0} = 0.3$  and 0.9 at  $z = 0$ , by modelling the IGM with the **Pshirkov16** model, including the linearly increasing  $f_{\text{IGM}}$ . For the case that  $f_{\text{IGM},0} = 0.9$ ,  $f_{\text{IGM}} = \text{const.}$  at all redshifts, similar to the constrained simulation in the lower row of Fig. 1. The fact that predictions are hardly distinguishable verifies the similarity of both models.<sup>3</sup> However, for lower values of  $f_{\text{IGM},0}$  there is stronger contribution for DM at large redshift; thus, larger values of DM are predicted in this case. This reduces the variance of predictions with different  $f_{\text{IGM},0}$ , when compared to varying constant  $f_{\text{IGM}}$ .

<sup>3</sup>A comparison of average  $\langle \text{DM} \rangle(z)$  can be found in Hackstein et al. (2020).



**Figure 1.** Distribution of extragalactic  $\text{DM}_{\text{EG}} = \text{DM}_{\text{obs}} - \text{DM}_{\text{MW}}$  as observed (dash-dot) by ASKAP (left), CHIME (centre), and Parkes (right) minus estimate of MW contribution based on NE2001 model, as listed in FRBcat. The solid lines show the expected distribution to be observed by the corresponding instrument according to equation (3), assuming FRBs from magnetars in our benchmark scenario using the constrained simulation to model the IGM and extreme constant values for the baryon content,  $f_{\text{IGM}} = 0.3$  (top) or  $f_{\text{IGM}} = 0.9$  (bottom), as well as FRB redshift distribution to follow SFR (blue), coV (orange), or SMD (green). The error bars show the shot noise of the observed data. For the expected distribution, the barely visible error bars are the shot noise of the Monte Carlo samples used to obtain the likelihood function. The product of Bayes factors (equation 5) for individual FRBs, shown in Fig. 3, quantifies how well different scenarios recreate the observed distribution.



**Figure 2.** Same as Fig. 1, using the Pshirkov16 model with redshift-dependent  $f_{\text{IGM}}$  for the IGM and extreme values for its present value  $f_{\text{IGM},0}$ . Models for the redshift distribution, circumburst environment, host, and intervening galaxies remain the same.

### 3.2 Inference of $f_{\text{IGM}}$ and $\pi(z)$

#### 3.2.1 $f_{\text{IGM}}$ from localized FRBs

Recently, Li et al. (2020) estimated  $f_{\text{IGM}} = 0.84^{+0.16}_{-0.22}$  using DM of the five localized FRBs together with the corresponding luminosity distance of Ia supernovae. Shortly after that, Bhandari et al. (2020) report the localization of two additional FRBs. Using their reported values of  $\text{DM}_{\text{EG}} = \text{DM}_{\text{obs}} - \text{DM}_{\text{MW}}$  for all six FRBs localized at decent redshift,  $z \geq 0.1$ , we estimate  $f_{\text{IGM}} = 0.84^{+0.11}_{-0.39}$  within  $3\sigma$  using our more sophisticated approach with our benchmark scenario and the Pshirkov16 model for the IGM. The lower limit suggests that  $f_{\text{IGM}} \gtrsim 0.5$ . The current sample of localized FRBs is thus not yet sufficient to solve the missing baryons problem unambiguously (cf. Macquart et al. 2020).

Though our inference depends on the chosen cosmology (Ade et al. 2016), as the models for IGM as well as host and intervening galaxies in our benchmark scenario are obtained from a specific cosmological simulation, we derive the same estimated value, with a more conservative lower limit. Still, the conclusion that this confirms the location of missing baryons in the WHIM should be taken with caution, at least because  $f_{\text{IGM}}$  is considered to be constant with redshift. Note that the likelihood function used in this estimate for  $f_{\text{IGM}}$ , i.e.  $\prod_i L(f_{\text{IGM}}|\text{DM}_i, z_i)$ , can be used as prior to infer the host redshift distribution using unlocalized FRBs, replacing assumptions on  $f_{\text{IGM}}$ .

#### 3.2.2 $f_{\text{IGM}}$ from individual instruments

In Fig. 3, we show the Bayes factors  $\mathcal{B}$  resulting from the joint analysis of the FRB redshift distribution and of the IGM baryon content  $f_{\text{IGM}}$ , assumed to be constant in our constrained simulation. For the 28 FRBs observed by ASKAP, high values of  $0.8 \leq f_{\text{IGM}}$  are strongly favoured, regardless of the assumed redshift distribution. Hence, ASKAP values seem to confirm the location of missing baryons in the ionized IGM, most probably the WHIM. Due to the generally low distance of ASKAP events,  $\mathcal{B}$  shows the lowest variability between the assumed host redshift distributions. However, estimates using ASKAP results are more vulnerable to the unknown contribution of the host galaxy and local environment of the source and should be taken with great caution. This is further stressed by the fact that ASKAP likely prefers too high  $f_{\text{IGM}} \gtrsim 0.9$ .

The 30 FRBs observed by CHIME, of which 18 have been found to repeat (Andersen et al. 2019; CHIME/FRB Collaboration 2019; Fonseca et al. 2020), clearly favour low values of  $f_{\text{IGM}} < 0.5$  as the observed distribution of DM concentrates on lower values than expected. Since SFR redshift distribution peaks at a higher redshift,  $z \approx 1$ , it is disfavoured by CHIME results, except for the case of  $f_{\text{IGM}} \approx 0.3$ . Instead, the SMD distribution, which peaks at lowest redshift  $z \approx 0.5$ , is preferred for all probed values of  $f_{\text{IGM}} > 0.4$  and favours  $f_{\text{IGM}} = 0.4$ . Only for  $f_{\text{IGM}} \lesssim 0.4$ , coV and SFR show higher values of  $\mathcal{B}$ . However, the parameters for CHIME selection effects available in FRBPOPPY are an early version (see Gardenier et al. 2019); hence, results for this instrument should be considered with caution. Still, if the current parameters are proven to be at least roughly correct, the present results either indicate a wrong assumption entering the expected distribution of FRB redshifts, which is likely as CHIME observes a different frequency band than ASKAP and Parkes, or suggest a very low baryon content in the IGM, which we deem as implausible, as it would contradict the result of any cosmological simulation. However, CHIME results so far seem to disfavour the ‘WHIM solution’ to the missing baryons problem.

The 27 FRBs observed by Parkes show a high variation among the assumed host redshift distributions. The SFR redshift distribution peaks at highest redshift  $z \approx 1.1$  and favours  $0.4 \lesssim f_{\text{IGM}} < 0.7$ , while it cannot be brought into agreement with the WHIM solution, as  $\mathcal{B}(f_{\text{IGM}} > 0.7) < 10^{-2}$  indicates that these values can be ruled out with more than 99 per cent certainty. The SMD distribution, peaking at  $z \approx 0.6$ , prefers a higher  $f_{\text{IGM}} \geq 0.7$  value. The coV distribution peaks at  $z \approx 0.7$  and favours  $f_{\text{IGM}} \approx 0.6$ , however, with relatively low significance.

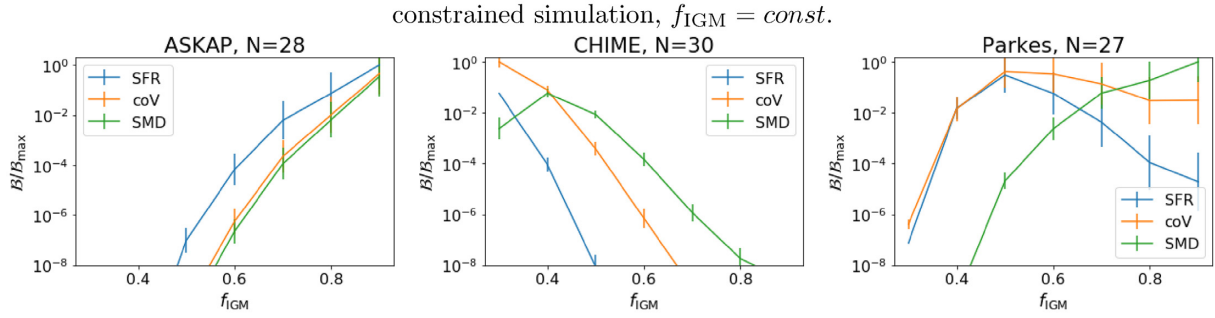
In Fig. 4, we show the Bayes factors  $\mathcal{B}$  resulting from the joint analysis of the FRB redshift distribution and of the IGM baryon content  $f_{\text{IGM},0}$  at  $z = 0$ , assumed to increase linearly until  $f_{\text{IGM}} = 0.9$  at  $z = 1.5$  in the Pshirkov16 model. Since predictions show less variance in  $f_{\text{IGM},0}$ , the graphs are shallower than those in Fig. 3, while the general trends remain the same. ASKAP results are in favour of the WHIM solution to the missing baryon problem. However, CHIME results strongly disfavour high densities of baryons in the local IGM, for all assumed redshift distributions. In particular, there is no value of  $f_{\text{IGM},0}$  for which  $\mathcal{B} > 10^{-2}$  with both ASKAP and CHIME. Such results can thus not be interpreted consistently, without considering further features that differ between the telescopes, e.g. systematic differences in the local contribution or the population of potentially visible FRBs. However, investigation of possible solutions is beyond the scope of this paper.

Not only does the inference of the different data sets differs dramatically, but also the assumed redshift distribution of FRBs strongly affects our interpretation of the data. Thus, the degenerate problem of inferring  $f_{\text{IGM}}$  with unlocalized FRBs is unlikely to be solved simply by increasing the sample size. This stresses the importance of careful investigation of all assumptions entering the interpretation of measures of FRBs in order to arrive at reasonable conclusions. As for now, we cannot derive strong or unambiguous conclusions on  $f_{\text{IGM}}$  and on the intrinsic redshift distribution from unlocalized FRBs. Only for the CHIME results, we can conclude that the underlying redshift distribution has to peak at  $z \lesssim 0.6$ , in order to explain the high number of sources with low DMs. This conclusion is independent of the details of the assumed intrinsic redshift distributions and selection effects.

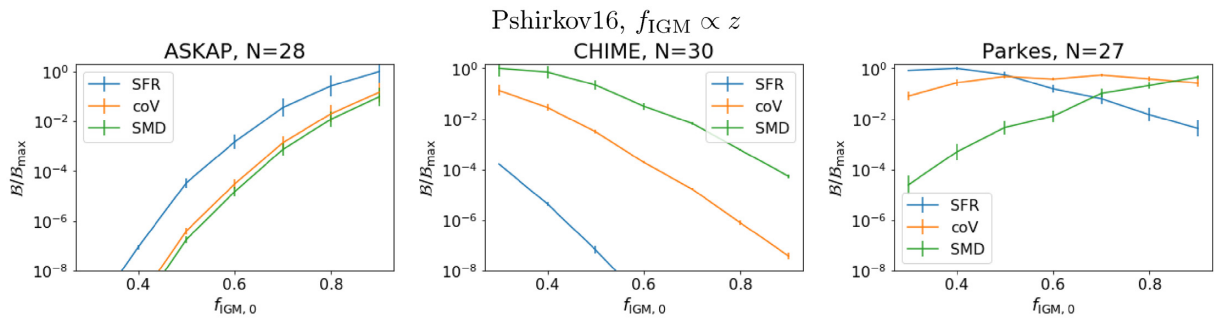
## 4 DISCUSSION

Our work highlights the fact that the distributions of FRBs obtained by different radio telescopes lead to very different inferences about the cosmological distribution of such events. The disagreement between ASKAP and CHIME is to be explained by a strong local contribution of DM to the bright low-distance FRBs observed by ASKAP. These would account for systematically increased values of the observed DM. By not accounting for such a strong local contribution in probability estimates, the increase in DM would be misinterpreted as IGM contribution, calling for a high  $f_{\text{IGM}}$ . If the high DMs observed with ASKAP can be attributed to a stronger local contribution, lower values of  $f_{\text{IGM}}$  would be favoured by FRB observations. This would reduce the fraction of missing baryons found in the WHIM.

However, the fact that CHIME sees a higher amount of FRBs with  $50 < \text{DM}/\text{pc cm}^{-3} < 200$  than ASKAP cannot be explained by low baryon content in the IGM or the general FRB population to be concentrated at low redshifts, which would affect CHIME and ASKAP in the same way. If this feature persists in larger samples, it provides a strong hint towards CHIME observing a different subset of FRB sources than ASKAP. The more than 10 times higher gain of CHIME allows it to potentially measure a different set of fainter



**Figure 3.** Bayes factor  $\mathcal{B}$  of FRBs found in FRBcat comparing scenarios for different baryon contents in IGM  $f_{\text{IGM}}$  and redshift distributions of FRBs, following either SFR, coV or SMD, assuming that FRBs are produced by magnetars in our benchmark scenario. To model the IGM, we use the constrained simulation described in detail in Hackstein et al. (2020). The title indicates the instrument that observed the considered number of  $N$  FRBs. The estimated value and deviation are obtained via jackknife resampling. All  $\mathcal{B}$  are renormalized to the best-fitting  $\mathcal{B}_{\text{max}}$  to indicate how much less likely a given scenario reproduces the observed data (cf. Fig. 1).



**Figure 4.** Same as Fig. 3, using the Pshirkov16 model for the IGM contribution. Models for the redshift distribution, circumburst environment, host, and intervening galaxies remain the same.

sources (Amiri et al. 2018), e.g. magnetars of older age (cf. Metzger, Berger & Margalit 2017; Marcote et al. 2020). Since these fainter sources are only visible at low redshift, this would increase the likelihood for CHIME to observe FRBs with low DM. Furthermore, the local environment of older magnetars contributes much less to observed DM (Piro & Gaensler 2018). Thus, old magnetars could account for the increased number of FRBs observed with low DM compared to our expectations that consider the same population of magnetars to be observed by CHIME and ASKAP. Such an increase in the likelihood to observe FRBs from low redshift would also make the CHIME results favour higher values of  $f_{\text{IGM}}$ . However, since CHIME is also able to observe FRBs visible with Parkes (e.g. FRB121102; see Josephy et al. 2019), we require CHIME to detect more events with  $\text{DM} > 800 \text{ pc cm}^{-3}$  and any above  $> 2000 \text{ pc cm}^{-3}$  in the future. If there is an extended set of sources available, this should reflect in a higher rate of FRBs observed by CHIME than extrapolated from other surveys (e.g. Chawla et al. 2017).

Note that 18 of the 30 FRBs observed by CHIME we considered here are repeaters (Andersen et al. 2019; CHIME/FRB Collaboration 2019; Fonseca et al. 2020). The high number of repeaters already observed by CHIME would suggest an increased chance to observe fainter bursts from repeating sources, e.g. from old magnetars. For these young neutron stars, this implies a more advanced dissipation of supernova remnants, thus a lower local contribution to DM. If ASKAP can only observe magnetars in an early stage, say  $< 1000 \text{ yr}$ , this not only could account for the high local contribution to DM, required earlier, but would also lower the inferred value of  $f_{\text{IGM}}$ . Older magnetars potentially observed by CHIME would also cause a lower contribution to Faraday rotation measure (RM), which could

be identified by statistical investigation of the distribution of RM observed by the different instruments. However, the current sample of reported RM is too small to allow for any firm conclusions. It might also be possible that separate frequency bands are dominated by completely different sources of FRBs.

Still, even with negligible local contributions, the CHIME results suggest a low IGM baryon content  $f_{\text{IGM}} \lesssim 0.5$ . If this cannot be explained by false assumptions about the FRB redshift distribution and selection effects, the CHIME results would constitute strong evidence against missing baryons in the WHIM or, in fact, anywhere in the ionized IGM.

*Caveats:* Our conclusions are based on a number of assumptions regarding the IGM baryon content, the FRB redshift distribution, and local contributions and foregrounds. Furthermore, in our benchmark scenario we assumed all FRBs to be caused by magnetars in axisymmetric galaxies. Alternative progenitor models beyond our assumptions might lead to different conclusions. For example, magnetars that only produce FRBs at an early stage, say less than  $10^3 \text{ yr}$  after their formation, would produce much stronger dispersion than older ones ( $\geq 10^5 \text{ yr}$ ), which we allowed in our benchmark scenario. This would make lower values of  $f_{\text{IGM}}$  more plausible, especially for the low-distance sample observed by ASKAP. In future work, we will consider a wider range of assumptions in order to provide more significant results.

Note that FRBPOPPY uses estimates, e.g. of  $\text{DM}(z)$ , in order to decide how many FRBs will be observed at a given redshift. These estimates have been produced using slightly different assumptions on the contributing regions. However, the DM is dominated by the IGM and the analytical description used in FRBPOPPY provides a

good match to our estimates. Hence, we argue that this does not alter the general conclusions of this work.

Finally, in our benchmark scenario, we do not account for clustering or galaxy haloes (e.g. Prochaska et al. 2019b; Connor et al. 2020). These are especially important for the interpretation of FRBs located at low redshifts. However, FRBs from higher redshift, for which these contributions are less important, are much more informative regarding  $\pi(z)$  as well as  $f_{\text{IGM}}$ . Our results are therefore reasonably robust against the additional presence of haloes along the line of sight.

## 5 CONCLUSIONS

FRBs are an important probe of baryons in the Universe, as their DM is sensitive to all phases of the intergalactic medium, including the WHIM. However, inference of the location of missing baryons with unlocalized FRBs is degenerate to the distribution of their host redshifts. In this work, we use the estimated extragalactic component of the DM to jointly infer the redshift distribution of unlocalized FRBs and of the IGM baryon content with a Bayesian statistical approach. We use an analytical model for the local environment of the source, assumed to be a magnetar, semi-analytical models for the ensemble of host and intervening galaxies, and models for the IGM from cosmological simulations. Comparing expectations to observations of unlocalized FRBs, we investigate the implications of events observed by ASKAP, CHIME, and Parkes, listed in the FRBcat, using the open-source PYTHON software package PREFRBLE<sup>4</sup> (Hackstein 2020). Our main conclusions can be summarized as follows:

(i) From six localized FRBs beyond redshift  $z \geq 0.1$ , as listed in Bhandari et al. (2020), we infer the intergalactic baryon content with  $3\sigma$  limits  $f_{\text{IGM}} = 0.84^{+0.11}_{-0.39}$ . FRBs thus potentially confirm the location of missing baryons in the WHIM. This value agrees well with results of Li et al. (2020). However, the error margin does not yet allow us to confirm the WHIM solution beyond doubt.

(ii) In order to unambiguously infer  $f_{\text{IGM}}$  from unlocalized FRBs, the discrepancy between different instruments needs to be resolved, e.g. by identifying different subsets of the FRB population. A mere increase in the number of events will likely not solve this discrepancy.

(iii) In case that the WHIM solution to the missing baryon problem is correct, the Parkes sample is in conflict with an FRB population following the star formation.

(iv) Limited to the CHIME sample, we find that the distribution that tracks the SFR is the least likely to explain observations. A better agreement is achieved for a distribution that peaks at rather low redshift  $z \lesssim 0.6$ , e.g. following the SMD.

Though the small sample of localized FRBs suggests that the missing baryons are in the ionized IGM, the error margin does not yet allow for unambiguous confirmation of this hypothesis.

## ACKNOWLEDGEMENTS

This work was funded by the Deutsche Forschungsgemeinschaft (DFG) under grant BR2026/25.

SH thanks Rainer Beck, Laura Spitler, and Charles Walker for interesting and fruitful discussions on many aspects of this work.

Our cosmological simulations were performed with the ENZO code (<http://enzo-project.org>), under projects HHH38 and HHH42 at the

Jülich Supercomputing Centre (P.I.: F. Vazza). FV acknowledges financial support from the ERC Starting Grant ‘MAGCOW’, no. 714196.

We further acknowledge the use of computational resources at the Rechenzentrum of the University of Hamburg.

## DATA AVAILABILITY

The data underlying this article are available in the PREFRBLE repository, at <http://doi.org/10.5281/zenodo.3862636>.

## REFERENCES

- Ade P. A. et al., 2016, *A&A*, 594, A13  
 Amiri M. et al., 2018, *ApJ*, 863, 48  
 Andersen B. et al., 2019, *ApJ*, 885, L24  
 Bannister K. W. et al., 2019, *Science*, 365, 565  
 Bera A., Bhattacharyya S., Bharadwaj S., Bhat N. R., Chengalur J. N., 2016, *MNRAS*, 457, 2530  
 Bhandari S. et al., 2020, *ApJ*, 895, L37  
 Bryan G. L. et al., 2014, *ApJS*, 211, 19  
 Caleb M., Flynn C., Bailes M., Barr E., Hunstead R., Keane E., Ravi V., van Straten W., 2016, *MNRAS*, 458, 708  
 Cen R., Ostriker J. P., 1999, *ApJ*, 519, L109  
 Chawla P. et al., 2017, *ApJ*, 844, 140  
 CHIME/FRB Collaboration, 2019, *Nature*, 566, 235  
 Connor L. et al., 2020, *MNRAS*, 499(4), 416  
 Cordes J. M., Lazio T. J. W., 2002, preprint(astro-ph/0207156)  
 Dai J.-P., Xia J.-Q., 2020, preprint (arXiv:2004.11276)  
 Dave R. et al., 2001, *ApJ*, 552, 473  
 de Graaff A., Cai Y.-C., Heymans C., Peacock J. A., 2019, *A&A*, 624, A48  
 Deng W., Zhang B., 2014, *ApJ*, 783, L35  
 Dolag K., Gaensler B. M., Beck A. M., Beck M. C., 2015, *MNRAS*, 451, 4277  
 Eftekhari T., Berger E., 2017, *ApJ*, 849, 162  
 Fonseca E. et al., 2020, *ApJ*, 891, L6  
 Gardener D., van Leeuwen J., Connor L., Petroff E., 2019, *A&A*, 632, A125  
 Hackstein S., , 2020, PrEFRBLE, <https://zenodo.org/record/3862636>, 12.10.2020  
 Hackstein S., Vazza F., Brüggem M., Sorce J. G., Gottlöber S., 2018, *MNRAS*, 475, 2519  
 Hackstein S., Brüggem M., Vazza F., Gaensler B. M., Heesen V., 2019, *MNRAS*, 488, 4220  
 Hackstein S., Brüggem M., Vazza F., Rodrigues L. F. S., 2020, *MNRAS*, 498(4), 4811  
 Josephy A. et al., 2019, *ApJ*, 882, L18  
 Keane E. et al., 2016, *Nature*, 530, 453  
 Lacey C. G. et al., 2016, *MNRAS*, 462, 3854  
 Li B., Li L.-B., Zhang Z.-B., Geng J.-J., Song L.-M., Huang Y.-F., Yang Y.-P., 2019, preprint (arXiv:1901.03484)  
 Li Z., Gao H., Wei J.-J., Yang Y.-P., Zhang B., Zhu Z.-H., 2020, *MNRAS*, 496, L28  
 Linder E. V., 2020, *Physical Review D*, 101(10), 103019  
 Locatelli N., Ronchi M., Ghirlanda G., Ghisellini G., 2018, *A&A*, 625, A109  
 Luo R., Lee K., Lorimer D. R., Zhang B., 2018, *MNRAS*, 481, 2320  
 Macquart J.-P., Ekers R., 2018, *MNRAS*, 480, 4211  
 Macquart J.-P. et al., 2010, *Publ. Astron. Soc. Aust.*, 27, 272  
 Macquart J.-P. et al., 2020, *Nature*, 581, 391  
 McQuinn M., 2013, *ApJ*, 780, L33  
 Mahony E. K. et al., 2018, *ApJ*, 867, L10  
 Manchester R. N., Hobbs G. B., Teoh A., Hobbs M., 2005, *AJ*, 129, 1993  
 Marcote B., Paragi Z., 2019, preprint (arXiv:1901.08541)  
 Marcote B. et al., 2020, *Nature*, 577, 190  
 Meiksin A., 2009, *Rev. Mod. Phys.*, 81, 1405

<sup>4</sup>[github.com/FRBs/PreFRBLE](https://github.com/FRBs/PreFRBLE)



- Metzger B. D., Berger E., Margalit B., 2017, *ApJ*, 841, 14
- Muñoz J. B., Loeb A., 2018, *Phys. Rev. D*, 98, 103518
- Nicastro F., Mathur S., Elvis M., 2008, *Science*, 319, 55
- Niino Y., 2018, *ApJ*, 858, 4
- Petroff E. et al., 2016, preprint (arXiv:1601.03547)
- Piro A. L., Gaensler B. M., 2018, *ApJ*, 861, 150
- Planck Collaboration I, 2014, *A&A*, 571, A1
- Pol N., Lam M. T., McLaughlin M. A., Lazio T. J. W., Cordes J. M., 2019, *ApJ*, 886, 135
- Prochaska J. X., Zheng Y., 2019, *MNRAS*, 485, 648
- Prochaska J. X., et al., 15878, , 2019a, HST.PROP,
- Prochaska J. X. et al., 2019b, *Science*, 366, 231
- Pshirkov M. S., Tinyakov P. G., Urban F. R., 2016, *Phys. Rev. Lett.*, 116, 191302
- Qiang D.-C., Wei H., 2020, *J. Cosmol. Astropart. Phys.*, 2020, 023
- Ravi V. et al., 2019, *Nature*, 572, 352
- Rodrigues L. F. S., Chamandy L., Shukurov A., Baugh C. M., Taylor A. R., 2018, *MNRAS*, 483, 2424
- Shull J. M., Smith B. D., Danforth C. W., 2012, *ApJ*, 759, 23
- Sorce J. G. et al., 2016, *MNRAS*, 455, 2078
- Staveley-Smith L. et al., 1996, *Publ. Astron. Soc. Aust.*, 13, 243
- Tanimura H. et al., 2018, *MNRAS*, 483, 223
- Tendulkar S. P. et al., 2017, *ApJ*, 834, L7
- The CHIME/FRB Collaboration, 2018, *ApJ*, 863, 48
- Walker C. R. H., Ma Y.-Z., Breton R. P., 2018, *A&A*, 638, A37
- Wei J.-J., Li Z., Gao H., Wu X.-F., 2019, *J. Cosmol. Astropart. Phys.*, 2019, 039
- Yang Y.-P., Li Q.-C., Zhang B., 2020, *ApJ*, 895, 7

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