Draft version March 28, 2024
Typeset using LaTeX default style in AASTeX631

# Evidence of a Common Emission Mechanism for Fast Radio Bursts

Michael Gutierrez
*2023 Dave L. Glackin Memorial SURF Fellow*
*California Institute of Technology, 1200 E. California Boulevard, Pasadena, CA 91125, USA*

Project Mentor: Charles Steinhardt
*Cosmic Dawn Center (DAWN)*
*Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, DK-2100 Copenhagen Ø*

## ABSTRACT

Fast radio bursts (FRBs) are bright, millisecond-duration radio signals of extragalactic origin. The physical nature of their progenitors and emission mechanisms is unclear. UMAP (Uniform Manifold Approximation and Projection) is a machine learning algorithm for dimensionality reduction that has been used to classify gamma-ray bursts into two archetypes suggestive of distinct astrophysical origins. In this work, we apply UMAP to search for archetypes in the CHIME/FRB Catalog, a population of 700 FRBs including about 200 from more than 50 known repeating FRB sources. Consistent with other studies on the catalog, we find evidence for distinct archetypes of FRBs based on their light curves. However, we also find that individual FRB sources can produce repeat bursts of multiple archetypes. We argue this is consistent with a single emission mechanism producing FRBs with a continuum of properties, in contrast to gamma-ray bursts. This suggests that all FRBs have the potential to repeat. We evaluate predictions of previous works on repeating vs. non-repeating FRB sources and assess the implications of a unified FRB model.

## 1. INTRODUCTION

Since their first identification in 2007 (Lorimer et al. 2007), transient astronomical events known as Fast Radio Bursts (FRBs) have been the subject of scrutiny for their unusual properties, their origins, and their potential use as astronomical probes. They are extremely fast ($10^{-4}$ to $10^{-2}$ s) and stochastically produced at a rate of ≈1,000 per day at a fluence of order Jy ms (Petroff et al. 2019). The frequency dispersion measure (DM) and Faraday rotation imparted on FRBs by intervening interstellar and intergalactic media make it possible to probe those otherwise difficult-to-observe environments (Bhandari & Flynn 2021; Prochaska et al. 2019). With sufficient data, FRBs may even help to constrain cosmological model parameters (James et al. 2021).

It is unclear whether all FRBs are a result of a single emission mechanism or whether there are several classes of FRB-emitting phenomena, as with gamma-ray bursts. Some FRBs originate from regions of active star-formation, while others have been localized within globular clusters (Bhandari et al. 2020). There is growing evidence that neutron stars, particularly magnetars, represent a portion of the FRB population (Bochenek et al. 2020; Andersen et al. 2020). Progenitor pathways thus may include the death of massive stars as well as binary formation channels.

A subset of FRB sources are known to periodically repeat (Caleb & Keane 2021). The classification of repeating vs. non-repeating FRB sources has been the focus of significant research in the past five years (Niu et al. 2022; Hilmarsson et al. 2020; Kirsten et al. 2023). A few models have found statistical differences between the populations of repeaters and non-repeaters, e.g., Zhu-Ge et al. (2022). Pleunis et al. (2021) found that FRBs from repeating sources have significantly higher durations and spectral running parameters on average. However, it is unknown whether all FRB sources have the potential to repeat over a range of timescales not yet probed by current observational facilities.

Pleunis et al. (2021) found that the dedispersed light curves of most FRBs can be visually classified into one of four archetypes based on their duration, number of sub-bursts, and scattering timescale. Chaikova et al. (2022) obtained a different set of archetypes by cross-correlating light curves. In both cases, each archetype contained bursts from repeating and non-repeating sources, suggesting that the repeat likelihood of FRBs cannot be determined from their light curves alone.

As suggested in Pleunis et al. (2021), the focus of this work will be creating a predictive model to attempt to reproduce these archetypes. We will use unsupervised machine learning to exclude any existing biases and evaluate the extent to which the light curves truly support a classification of FRBs into distinct categories and, indeed, whether those categories correspond to repeater vs. non-repeater bursts. The success of this model should elucidate whether any observed archetypes of FRBs truly represent distinct astrophysical origins, or instead one continuum of properties.

In Sections 2.1-2.2, we give an overview of the CHIME telescope and the CHIME/FRB public database. In Section 2.3, we introduce the technique of dimensionality reduction and the Uniform Manifold Approximation and Projection (UMAP) algorithm (McInnes et al. 2018). We describe our data pre-processing and UMAP pipeline for FRB light curves in Section 2.4, then present the resulting classifications in Section 2.5. In Section 3, we corroborate the classifications from UMAP with findings of other works (Zhu-Ge et al. 2022; Pleunis et al. 2021; Chaikova et al. 2022) and demonstrate that individual sources can produce repeating bursts with light curves spanning multiple classes. We discuss future work to search for other physical differences between sources of bursts in different classes and evaluate the implications for a theory of a common FRB emission mechanism.

## 2. METHODS AND RESULTS

### 2.1. *CHIME Telescope*

The Canadian Hydrogen Intensity Mapping Experiment (CHIME) is an array of four 100-meter-long cylindrical radio reflectors located in the Canadian province of British Colombia. It has a field of view of more than 200 square degrees, bandwidth spanning 400-800 MHz, and 1 ms sampling time. Its field of view can be broken up into 1,024 individual beams approximately 20-40 arcminutes across (Collaboration et al. 2021).

CHIME has a real-time detection pipeline designed specifically for FRBs. Downsampled data is dedispersed in real time at DM values up to 13,000 pc $\mathrm{cm}^{-3}$ and searched for peaks with a signal-to-noise ratio of $10\sigma$ or higher. If such a signal is found, then the full-resolution data is recalled from a ring buffer and catalogued (Amiri et al. 2018).

This system is very effective at detecting FRBs, with over 1,000 discovered since the start of the project (Collaboration et al. 2023). The instrument itself has a few shortcomings. For instance, many FRBs have durations or microstructures much shorter than CHIME's 1 ms resolution (Farah et al. 2018), meaning that some FRBs within the detectable frequency and intensity range are missed. FRBs often span frequency ranges much larger than CHIME's 400-800 MHz bandwidth, so it is impossible for CHIME to determine their true extent in frequency space. The 20-40 arcminute beam widths are typically unable to constrain the positions of FRBs sufficiently to localize them to a host galaxy (Heintz et al. 2020).

### 2.2. *CHIME/FRB Catalogs*

The CHIME/FRB Collaboration (Collaboration et al. 2021) have published two catalogs of FRB observations, with a total of more than 700 individual bursts, including around 200 from more than 50 known repeating sources. They represent the biggest collection of FRBs from a single site and enable analysis of a large portion of FRB parameter space. Each catalog entry includes a dedispersed light curve and a variety of summary data including burst duration, intensity, and dispersion measure. These data are freely accessible via the CHIME/FRB Public Database.

Summary data are determined for each FRB by a least-squares fit to a dispersed dynamic spectrum profile:

$$\frac{A}{2\tau}\exp\left(\frac{w^2}{2\tau^2}\right)\exp\left(-\frac{t-k\mathrm{DM}[f^{-2}-f_0^{-2}]-t_{arr}}{\tau}\right)\left(1+\mathrm{erf}\left(\frac{t-k\mathrm{DM}[f^{-2}-f_0^{-2}]-\left(t_{arr}+\frac{w^2}{\tau}\right)}{\sqrt{2}w}\right)\right)\left(\frac{f}{f_0}\right)^{\left(-\gamma+r\ln\left(\frac{f}{f_0}\right)\right)}$$

where $f_0 = 400.1953125$ MHz, $k = (2.41\times10^{-4})^{-1}$ s $\mathrm{pc}^{-1}$ $\mathrm{cm}^3$ $\mathrm{MHz}^2$, and $A, \tau, w, \mathrm{DM}, t_{arr}, \gamma, r$ are fit parameters (Collaboration et al. 2021). It is essentially a gaussian pulse convolved with an exponential tail in the time axis and a running power law in the frequency axis. It is worth noting that this is an approximation, so some FRBs with more complex structure may not have good fit results.

### 2.3. *UMAP Algorithm*

Uniform Manifold Approximation and Projection (UMAP) is an unsupervised machine learning algorithm used primarily for dimensionality reduction. Inputs are highly-dimensional datasets and outputs are 2D projections thereof. The quality of the output depends on the choice of the "hyperparameters", `n_neighbors` and `min_dist`, which control

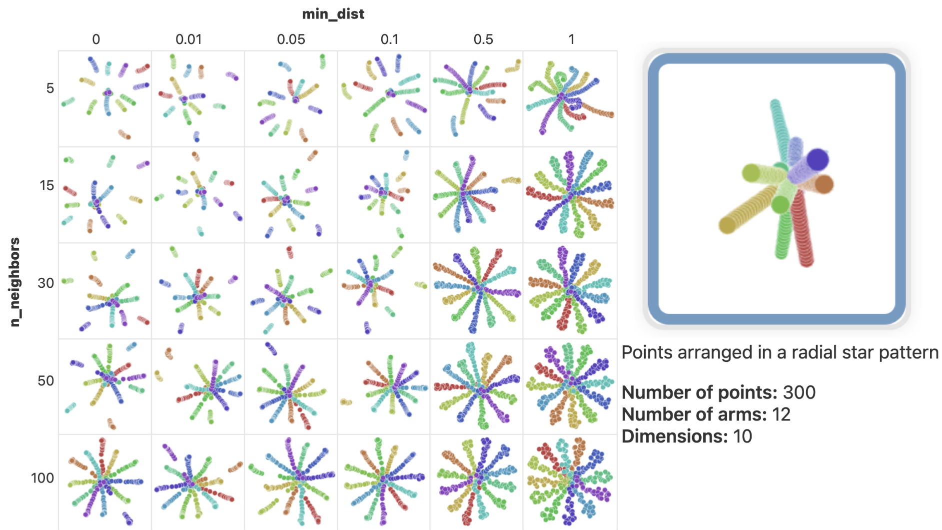



**Figure 1**: UMAP projections (left) of a toy dataset (right) with a variety of common values for the `n_neighbors` and `min_dist` parameters. Figure reproduced from Understanding UMAP by Andy Coenen, Adam Pearce under Apache 2.0 license.

the amount of global vs. local structure preserved from the data and emphasized in the projection, respectively (McInnes et al. 2018).

Figure 1 shows UMAP projections of a toy dataset, demonstrating how global structure of a highly-dimensional dataset can be projected onto a 2D map. The axes of the individual projections are meaningless. Only the relative distance between two points is indicative of their "similarity" as judged by UMAP. The original dataset (the radial star pattern) is essentially continuous on a global scale, so projections using higher values of `n_neighbors` and `min_dist` perform best while those using lower values are overwhelmed by local variations.

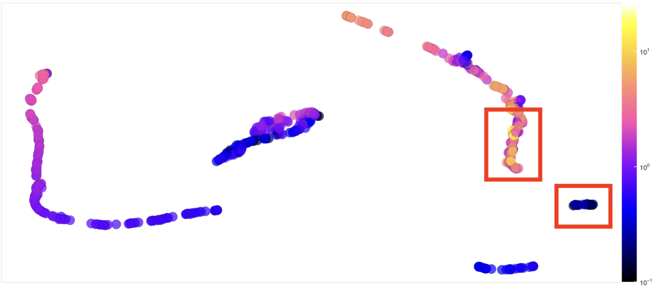
### 2.4. *UMAP Analysis of CHIME/FRB Catalog*

UMAP is unsupervised, meaning it does not consider existing labels or categories of data points in its learning process, so classifying astronomical observations is a natural application. Unsupervised dimensionality reduction algorithms have previously identified distinct classes of gamma-ray bursts in the the Swift/BAT catalog based on their light curves alone (Jespersen et al. 2020). Here, we apply UMAP to light curves in the CHIME/FRB catalog in attempt to reveal a physically-motivated classification of FRBs and their emission mechanisms. We choose a range of hyperparameters that will emphasize the global structure of the FRB population.

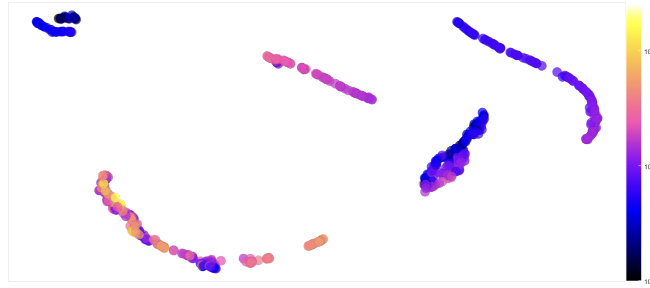
All light curves are available as 0.983 ms resolution arrays of intensity vs. time. They are not of equal length and the peak intensities are not aligned to a common index, so pre-processing is required to make the light curves comparable using UMAP's Euclidean "pixel by pixel" metric. We utilize the pre-processing method in Jespersen et al. (2020) of zero-padding and Discrete Fourier Transforming (DFT) the light curves before running UMAP on the resulting Fourier amplitudes. It was also found that normalizing each light curve to amplitude 1 before the DFT promoted structure-based clustering of FRBs in the projection while de-emphasizing noise and relative brightness. This pipeline was implemented in Python using the `numpy` and `umap_learn` libraries. The code can be accessed on GitHub at https://github.com/guutz/dawn2023.

### 2.5. *Results*

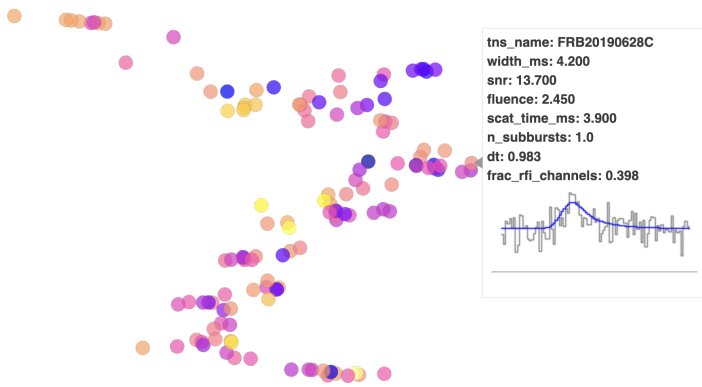
Figure 2 shows UMAP projections of the pre-processed light curves at selected values of `n_neighbors`. Distinct clusters of bursts are present, each roughly sorted by duration and shape. The algorithm was not provided any metadata for the light curves; it was able to identify burst archetypes solely by analyzing the DFTs. Note that the same clusters appear in the two separate projections even though they used different hyperparameters, which confirms some degree of underlying structure within the dataset.



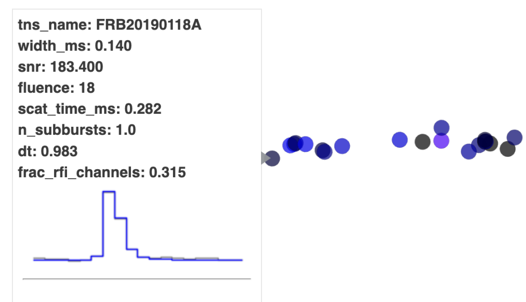
(a) `n_neighbors=15`



(b) `n_neighbors=30`




(c) Expanded view of boxed region in (a). This cluster is composed of longer, more scattered bursts.




(d) Expanded view of boxed region in (a). This cluster is composed of the shortest bursts in the catalog.

**Figure 2**: UMAP projections of zero-padded normalized DFT light curves from CHIME/FRB catalog. Colorbar indicates burst duration in ms on a log scale. Inset graphs display catalog metadata, raw time series (gray trace), and best-fit model time series (blue trace).

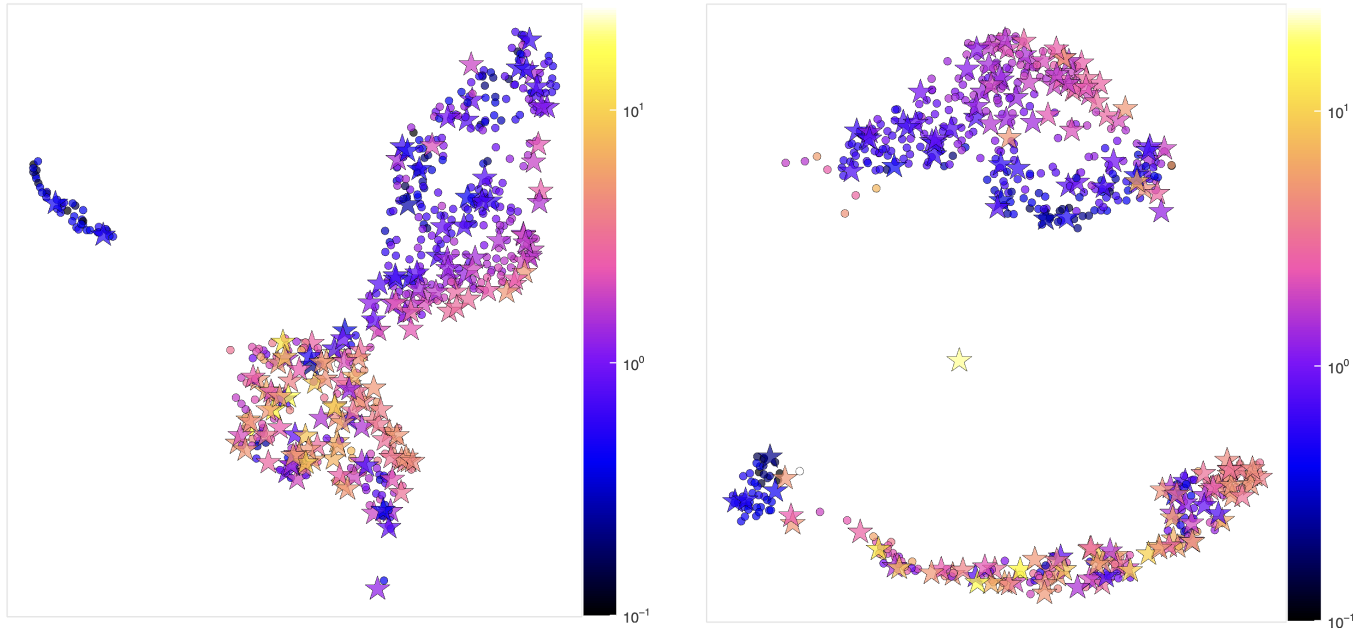
Projections in Figure 3 were made with "parametric UMAP", a variant of UMAP which uses a convolutional neural network to learn the structure of the original dataset instead of the usual "nearest neighbors" method (Sainburg et al. 2020). The parametric UMAP projections have fewer clusters with less distinct boundaries compared to those of non-parametric UMAP, which is an expected consequence of better preservation of global structure. The remaining clusters are still roughly divided between longer- and shorter-duration FRBs, consistent with the non-parametric projections.

Most notable out of these results is the lack of majority-repeater or majority-non-repeater clusters. It can be seen in Figure 3 that repeated FRBs, marked by star shapes, are present in all clusters. Figure 4 highlights two prolific repeaters within the projections from Figure 2. Each has produced a diversity of bursts spanning multiple of the UMAP-identified archetypes. In other words, UMAP has failed to find any statistical differences between light curves of repeating and non-repeating FRBs.

## 3. ANALYSIS AND DISCUSSION

The clusters of burst properties from Section 2.5 confirm several features of the CHIME catalog found by previous studies. Consistent with Pleunis et al. (2021)'s discovery of a correlation between FRB duration and repeat likelihood, we found (e.g., Figure 3a) that regions of the UMAP projections with the longest FRBs also had relatively high densities of repeated FRBs. Our UMAP projections also independently identified four or five distinct archetypes of light curve shapes similar to those in Pleunis et al. (2021) (e.g., Fig. 2). Consistent with Zhu-Ge et al. (2022) and Chaikova et al. (2022), each of whom used different machine learning and/or data pre-processing methods, we found that repeated FRBs appear in every cluster of every projection found. Also in agreement with the latter two works, we found that many individual FRB sources have produced bursts in every cluster of every projection found.

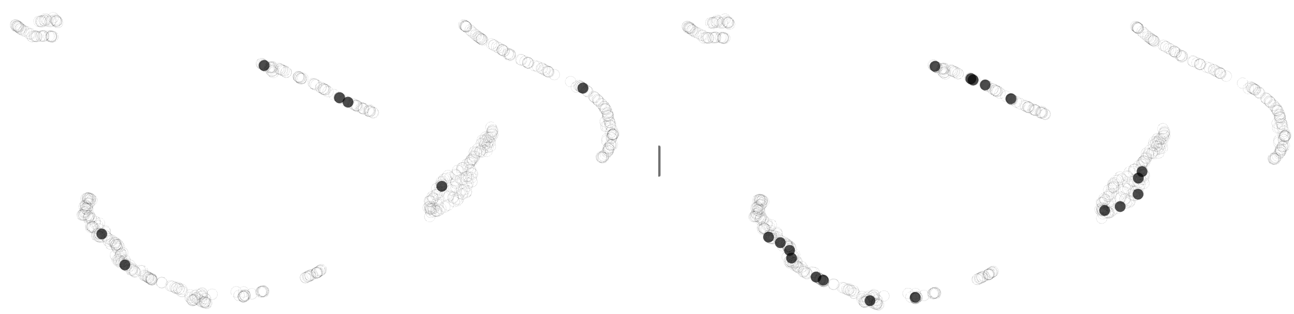
In other words, it is clear there are distinct archetypes of FRBs based on their light curves, and it has been shown that a single repeating FRB source can exhibit multiple archetypes over time. This has also been demonstrated by




(a) Projection of time series data before normalizing. The small cluster of bursts on the left side all have durations $\leq 1$ ms, the resolution limit of CHIME, so it may be a spurious classification.

(b) Projection of time series data after normalizing. The lonely burst in the middle of the graph is highly scattered and had a poor SNR, likely causing it to be set aside from the rest.

**Figure 3**: Parametric UMAP projections of zero-padded DFT light curves from CHIME/FRB catalog. Colorbar indicates burst duration in ms. Star shapes indicate a burst from a known repeating source.



(a) Fig. 2b with all bursts from repeating source FRB20200929C colored black.

(b) Fig. 2b with all bursts from repeating source FRB20180916B colored black.

**Figure 4**: Individual repeating sources can produce FRBs in multiple distinct categories, as defined by the UMAP projections.

observation of the same source over several weeks (Farah et al. 2018). These results suggest that all the FRBs observed so far could have been produced by the same emission mechanism. This would also mean that so-called non-repeating FRBs actually come from repeating sources which have only been seen once, perhaps due to selection effects or a longer repeat timescale.

As discussed in Section 1, the diversity of astrophysical environments in which FRBs have been localized suggests that FRBs can originate from several types of progenitors. This does not rule out a common FRB emission mechanism which could proceed in the environments of all those progenitors, e.g. a highly-magnetized plasma field. Some possible emission mechanisms of this type are discussed in Lyubarsky (2021); Connor et al. (2020); Zhang (2020).

Given that some FRBs are known to originate from the same environments of other high-energy transients like gamma-ray bursts from magnetars (Andersen et al. 2020), an implication of the unified FRB theory is that all FRBs should be accompanied by other high-energy phenomena. Thus, a key part of testing the theory is multi-wavelength observations. Bailes et al. (2021) compiled a wealth of observations of galactic magnetar SGR J1935+2154 in the years leading up to its emission of FRB20200420. Nicastro et al. (2021) suggested a reversed FRB search in which a few small radio telescopes follow the gaze of x-ray or gamma-ray telescopes. Another option could be to focus several different observatories on nearby galaxies and supernova remnants, especially ones with known repeating FRB sources. This would have higher sensitivity and higher reward for a detection. These strategies also have the natural consequence of localizing the sources of detected FRBs to a much more precise position than CHIME is capable of. Localizing FRB sources will thus be one of the best ways to constrain the range of plausible theories.

## 4. ACKNOWLEDGEMENTS

This project was funded by a Summer Undergraduate Research Fellowship (SURF) awarded by the Caltech Student-Faculty Programs Office. This SURF was specifically funded by Kathryn McEachern's *Dave Glackin Memorial SURF Fund*. Special thanks to Profs. Rob Phillips and Lee McCuller for their help with the application, and to Prof. Gregg Hallinan for writing mentorship. Finally, of course, Prof. Charles Steinhardt and everyone in the DAWN Group at *Københavns Universitet* were incredibly welcoming and helpful throughout the summer. Prof. Steinhardt was a fantastic mentor, and provided invaluable advice and different perspectives to consider throughout the project.

## REFERENCES

Amiri, M., Bandura, K., Berger, P., et al. 2018, The Astrophysical Journal, 863, 48, doi: 10.3847/1538-4357/aad188

Andersen, B. C., Bandura, K. M., Bhardwaj, M., et al. 2020, Nature, 587, 54, doi: 10.1038/s41586-020-2863-y

Bailes, M., Bassa, C., Bernardi, G., et al. 2021, Monthly Notices of the Royal Astronomical Society, 503, 5367, doi: 10.1093/mnras/stab749

Bhandari, S., & Flynn, C. 2021, Probing the universe with fast radio bursts, MDPI AG, doi: 10.3390/universe7040085

Bhandari, S., Sadler, E. M., Prochaska, J. X., et al. 2020, The Astrophysical Journal, 895, L37, doi: 10.3847/2041-8213/ab672e

Bochenek, C. D., Ravi, V., Belov, K. V., et al. 2020, Nature, 587, 59, doi: 10.1038/s41586-020-2872-x

Caleb, M., & Keane, E. 2021, A decade and a half of fast radio burst observations, MDPI, doi: 10.3390/universe7110453

Chaikova, A., Kostunin, D., & Popov, S. B. 2022. http://arxiv.org/abs/2202.10076

Collaboration, T. C., :, Amiri, M., et al. 2021, doi: 10.3847/1538-4365/ac33ab

Collaboration, T. C., :, Andersen, B. C., et al. 2023, doi: 10.3847/1538-4357/acc6c1

Connor, L., Miller, M. C., & Gardenier, D. W. 2020, Monthly Notices of the Royal Astronomical Society, 497, 3076, doi: 10.1093/mnras/staa2074

Farah, W., Flynn, C., Bailes, M., et al. 2018, Monthly Notices of the Royal Astronomical Society, 478, 1209, doi: 10.1093/mnras/sty1122

Heintz, K. E., Prochaska, J. X., Simha, S., et al. 2020, doi: 10.3847/1538-4357/abb6fb

Hilmarsson, G. H., Michilli, D., Spitler, L. G., et al. 2020, doi: 10.3847/2041-8213/abdec0

James, C. W., Prochaska, J. X., Macquart, J. P., et al. 2021, doi: 10.1093/mnras/stab3051

Jespersen, C. K., Severin, J. B., Steinhardt, C. L., et al. 2020, The Astrophysical Journal, 896, L20, doi: 10.3847/2041-8213/ab964d

Kirsten, F., Ould-Boukattine, O., Herrmann, W., et al. 2023. http://arxiv.org/abs/2306.15505

Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, doi: 10.1126/science.1147532

Lyubarsky, Y. 2021, Emission mechanisms of fast radio bursts, MDPI AG, doi: 10.3390/universe7030056

McInnes, L., Healy, J., & Melville, J. 2018. http://arxiv.org/abs/1802.03426

Nicastro, L., Guidorzi, C., Palazzi, E., et al. 2021, Multiwavelength observations of fast radio bursts, MDPI AG, doi: 10.3390/universe7030076

Niu, C. H., Aggarwal, K., Li, D., et al. 2022, Nature, 606, 873, doi: 10.1038/s41586-022-04755-5

Petroff, E., Hessels, J. W., & Lorimer, D. R. 2019, Fast radio bursts, Springer Verlag, doi: 10.1007/s00159-019-0116-6

Pleunis, Z., Good, D. C., Kaspi, V. M., et al. 2021, The Astrophysical Journal, 923, 1, doi: 10.3847/1538-4357/ac33ac

Prochaska, J. X., Macquart, J.-P., Mcquinn, M., et al. 2019, The low density and magnetization of a massive galaxy halo exposed by a fast radio burst, Tech. rep. https://www.science.org

Sainburg, T., McInnes, L., & Gentner, T. Q. 2020. http://arxiv.org/abs/2009.12981

Zhang, B. 2020, doi: 10.1038/s41586-020-2828-1

Zhu-Ge, J.-M., Luo, J.-W., & Zhang, B. 2022, Monthly Notices of the Royal Astronomical Society, 519, 1823, doi: 10.1093/mnras/stac3599