Draft version April 17, 2026
Typeset using LaTeX **twocolumn** style in AASTeX631

Discovery of Bimodal Drift Rate Structure in FRB 20240114A: Evidence for Dual Emission Regions

Santosh Arron[1]

[1] *Blankline, blankline.org*

ABSTRACT

We report the discovery of bimodal structure in the drift rate distribution of upward-drifting burst clusters from the hyperactive repeating fast radio burst FRB 20240114A. Using unsupervised machine learning (HDBSCAN density-based clustering applied directly to an 8-dimensional standardized feature space, with UMAP used only for post-hoc visualization) applied to 233 upward-drifting burst clusters from the FAST telescope dataset, we identify a distinct subpopulation of 45 burst clusters (hereafter Cluster C1) with **mean** drift rates 2.5× higher than typical upward-drifting burst clusters (245.6 vs 98.1 MHz/ms). Gaussian mixture modeling reveals strong evidence for bimodality (ΔBIC = 296.6), with clearly separated modes (Ashman's D = 2.70 > 2; Ashman et al. 1994) and a statistically significant gap in the distribution (11.3σ). **Crucially, all 45 C1 members are single-component (U1) burst clusters with drift rates well above their measurement uncertainties (minimum 87.1 MHz/ms, typical error ∼3–65 MHz/ms).** Crucially, we demonstrate that this bimodality persists when restricting the analysis to single-component (U1) burst clusters only **(n = 142; none consistent with zero drift within errors)**: ΔBIC = 19.9, Ashman's D = 2.71, gap significance = 9.2σ, confirming that the result is not an artifact of combining single- and multi-component burst clusters with different drift rate definitions **nor of measurement-error contamination near the drift-rate floor**. The extreme-drift subpopulation also exhibits systematically lower peak frequencies (−7%), shorter durations (−29%), and distinct clustering in multi-dimensional feature space. **These findings are suggestive of two spatially separated emission regions in the magnetosphere, each producing upward-drifting burst clusters with distinct physical characteristics, although confirmation requires observations from additional epochs and sources.**

Keywords: Fast radio bursts — Magnetars — Radio transient sources — Machine learning

1. INTRODUCTION

Fast radio bursts (FRBs) are millisecond-duration radio transients of extragalactic origin (Lorimer et al. 2007; Thornton et al. 2013), whose physical mechanism remains debated (Petroff et al. 2022). The discovery of repeating FRBs (Spitler et al. 2016) has enabled detailed morphological studies of individual sources, revealing complex time-frequency structure including sub-burst drift rates (Hessels et al. 2019), spectro-temporal correlations (Rajabi et al. 2020; Chamma et al. 2023), and diverse morphological classes (Pleunis et al. 2021; CHIME/FRB Collaboration 2021).

santosharron@blankline.org

FRB 20240114A, discovered in January 2024, exhibits an exceptionally high burst rate, making it an ideal laboratory for statistical studies of burst properties. Recent work by Zhang et al. (2026) classified 978 drifting burst clusters from FRB 20240114A into upward-drifting (23.8%) and downward-drifting (76.2%) populations based on their sub-burst drift rates, using data obtained with the Five-hundred-meter Aperture Spherical Telescope (FAST). They further classified burst clusters by internal morphology: single-component (U1/D1), double-component (U2/D2), and multi-component (Um/Dm), where the drift rate definition differs between these categories—for single-component burst clusters it characterizes the intrinsic frequency-time evolution, while for **double/multi-component** burst clusters it reflects the relative separation between sub-bursts.

Their analysis found systematic differences between upward- and downward-drifting populations (defined using consecutive time intervals) using Kolmogorov-Smirnov tests, though such differences do not necessarily hold for burst clusters defined with intermittent time intervals (Zhang et al. 2026). Each drift-direction class was treated as internally homogeneous.

In this paper, we apply unsupervised machine learning to search for *hidden substructure* within these morphological classes. We discover that the upward-drifting population contains two statistically distinct subpopulations, and demonstrate that this bimodality persists even when restricting to single-component burst clusters alone, ruling out definitional artifacts.

2. DATA AND METHODS

2.1. *Dataset*

We analyze the publicly available FRB 20240114A Morphology Dataset (Zhang et al. 2026), containing 978 burst clusters with measurable drift rates observed by FAST on 2024 March 12. For each burst cluster, we extract eight physical features: bandwidth, effective width, peak frequency, drift rate, energy, flux, signal-to-noise ratio, and center frequency. As noted by Zhang et al. (2026), the 233 upward-drifting burst clusters comprise 142 single-component (U1; 60.9%), 86 double-component (U2; 36.9%), and 5 multi-component (Um; 2.1%) burst clusters.

We note that the tabulated features (bandwidth, effective width, peak frequency, energy, flux, S/N, and center frequency) in the morphology table correspond to a representative burst within each burst cluster, while only the first seven columns are defined at the burst cluster level (Zhang et al. 2026). Our analysis uses the burst cluster-level quantities as provided in the public release.

Drift rate measurement uncertainties are provided in the public catalogue (Zhang et al. 2026). Among the 233 upward-drifting burst clusters, 87 have drift rates below 50 MHz/ms; all 87 are U2 or Um morphology, and 42 of these are consistent with zero within their measurement errors. In contrast, among the 142 U1 burst clusters, zero have drift rates consistent with zero (minimum drift rate 87.1 MHz/ms; typical errors 3–65 MHz/ms). This distinction between U1 and U2/Um measurement reliability motivates the U1-only control analysis presented in Section 3.3.

2.2. *Unsupervised Clustering*

Feature preprocessing. Each feature is standardized to zero mean and unit variance, with flux, energy, and bandwidth log-transformed first to compress their heavy tails.

Density-based clustering. We apply Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN; Campello et al. 2013) directly to the standardized 8-dimensional feature space (min_cluster_size=15, min_samples=5) to identify natural groupings without requiring the number of clusters to be specified.

Post-hoc visualization. Uniform Manifold Approximation and Projection (UMAP; McInnes et al. 2018) is used only to project the same feature space to 2D for visualization of clusters HDBSCAN has already identified (n_neighbors=15, min_dist=0.1). We deliberately do *not* cluster on the UMAP projection, because dimensionality-reduction algorithms can create apparent gaps where none exist in the original data.

2.3. *Bimodality Testing*

To verify that identified clusters represent discrete populations rather than continuous variation, we apply:

- Gaussian Mixture Model (GMM) comparison using the Bayesian Information Criterion (BIC), where $\Delta\text{BIC} > 10$ indicates strong evidence for the more complex model
- Ashman's D statistic for mode separation (Ashman et al. 1994), where $D > 2$ indicates clearly resolved modes
- Gap analysis for distribution discontinuities

3. RESULTS

3.1. *Discovery of the Extreme-Drift Subpopulation*

HDBSCAN clustering identifies two distinct clusters within the data (Figure 1). One cluster, which we designate C1, contains exclusively upward-drifting burst clusters ($45/45 = 100\%$), while exhibiting physical properties distinct from other upward-drifting burst clusters (Table 1). **C1 is identified through multi-dimensional density-based clustering in the full 8-feature standardized space, not on the UMAP projection and not by a simple drift rate threshold. Consequently, while C1 members have statistically higher mean drift rates (245.6 MHz/ms vs 98.1 MHz/ms), individual burst clusters outside C1 may occasionally have drift rates exceeding some C1 members, because cluster membership depends on the joint distribution of all eight features (bandwidth, duration, peak frequency, drift rate, energy, flux, S/N, and center**

Table 1. Properties of Cluster C1 vs. Other Upward-Drifting Burst Clusters

Property	C1 (n=45)	Other Up	p-value
Drift Rate (MHz/ms)	245.6	98.1	1.8×10^{-5}
Duration (ms)	1.68	2.38	1.2×10^{-3}
Peak Freq (MHz)	1102.6	1185.8	6.2×10^{-5}
Center Freq (MHz)	1120.3	1207.4	7.8×10^{-6}

NOTE—All p-values from Mann-Whitney U tests survive Bonferroni correction.

frequency). C1 exhibits systematically lower emission frequencies (Figure 2).

3.2. *Bimodality Confirmation*

Gaussian mixture modeling of the upward-drifting drift rate distribution strongly favors a 2-component model over 1-component (ΔBIC = 296.6, where ΔBIC > 10 indicates strong evidence). Ashman's D = 2.70 confirms clear separation between modes (D > 2 required; Ashman et al. 1994). Gap analysis reveals an 11.3σ discontinuity in the distribution, confirming discrete rather than continuous structure (Figure 3).

3.3. *Single-Component Consistency Check*

A key concern is whether the observed bimodality might arise from combining single-component (U1) and **double/multi-component** (U2/Um) burst clusters, whose drift rates are defined differently (Zhang et al. 2026): for single-component burst clusters, the drift rate characterizes the intrinsic frequency-time evolution of one emission component, whereas for **double/multi-component** burst clusters it reflects the relative frequency separation and timing between sub-bursts. **A further concern is that low drift rates near zero may be dominated by measurement error, inflating the apparent separation between modes.**

We address both concerns by restricting the analysis to U1 burst clusters only (n = 142). This subpopulation eliminates definitional heterogeneity and removes all burst clusters with drift rates near the measurement floor: the minimum U1 drift rate is 87.1 MHz/ms, and zero out of 142 U1 burst clusters have drift rates consistent with zero within their errors.

To address this, we repeat the bimodality analysis separately for each morphological subclass. Among the 233 upward-drifting burst clusters, 142 are single-component (U1), 86 are double-component (U2), and 5 are multi-component (Um). **Importantly, the U1 test operates on the full 142-burst single-component sample defined by the morphological classification of Zhang et al. (2026)—it is *not* the subset of 45 C1 burst clusters identified by HDBSCAN. The bimodality therefore appears in an independently-defined sample and cannot be an artifact of circular selection from the clustering step.**

Restricting to single-component (U1) burst clusters only, we find:

- ΔBIC = 19.9 (strong evidence for bimodality, threshold = 10)
- Ashman's D = 2.71 (clearly separated modes, threshold = 2)
- Gap z-score = 9.2σ (significant discontinuity)

This confirms that bimodality is an intrinsic property of single-component upward-drifting burst clusters and is *not* an artifact of combining heterogeneous drift rate definitions.

Furthermore, we examined the morphological composition of Cluster C1: all 45 burst clusters in C1 are classified as single-component (U1). **The non-C1 upward-drifting burst clusters contain a mix of U1 (97/188 = 51.6%), U2 (86/188 = 45.7%), and Um (5/188 = 2.7%).** A chi-squared test confirms that C1 has a significantly different morphological composition from the remaining upward-drifting burst clusters ($\chi^2 = 51.9$, $p < 10^{-4}$), indicating that C1 does not simply recover previously defined morphological categories but rather identifies new substructure within the single-component class itself.

3.4. *Robustness*

The extreme-drift cluster C1 is identified consistently across:

- All tested UMAP parameters (6/6 configurations)
- All tested HDBSCAN parameters (6/6 configurations)
- 98% of bootstrap resamples (98/100)
- **All tested decorrelated feature subsets (4/4 variants), including a minimal 5-feature set (drift rate, duration, peak frequency, energy, S/N) with the correlated bandwidth, flux, and center frequency removed. C1 is recovered with 100% upward-drifting purity in every variant.**

These results demonstrate that the discovered subpopulation is robust to methodological choices (Figure 4).

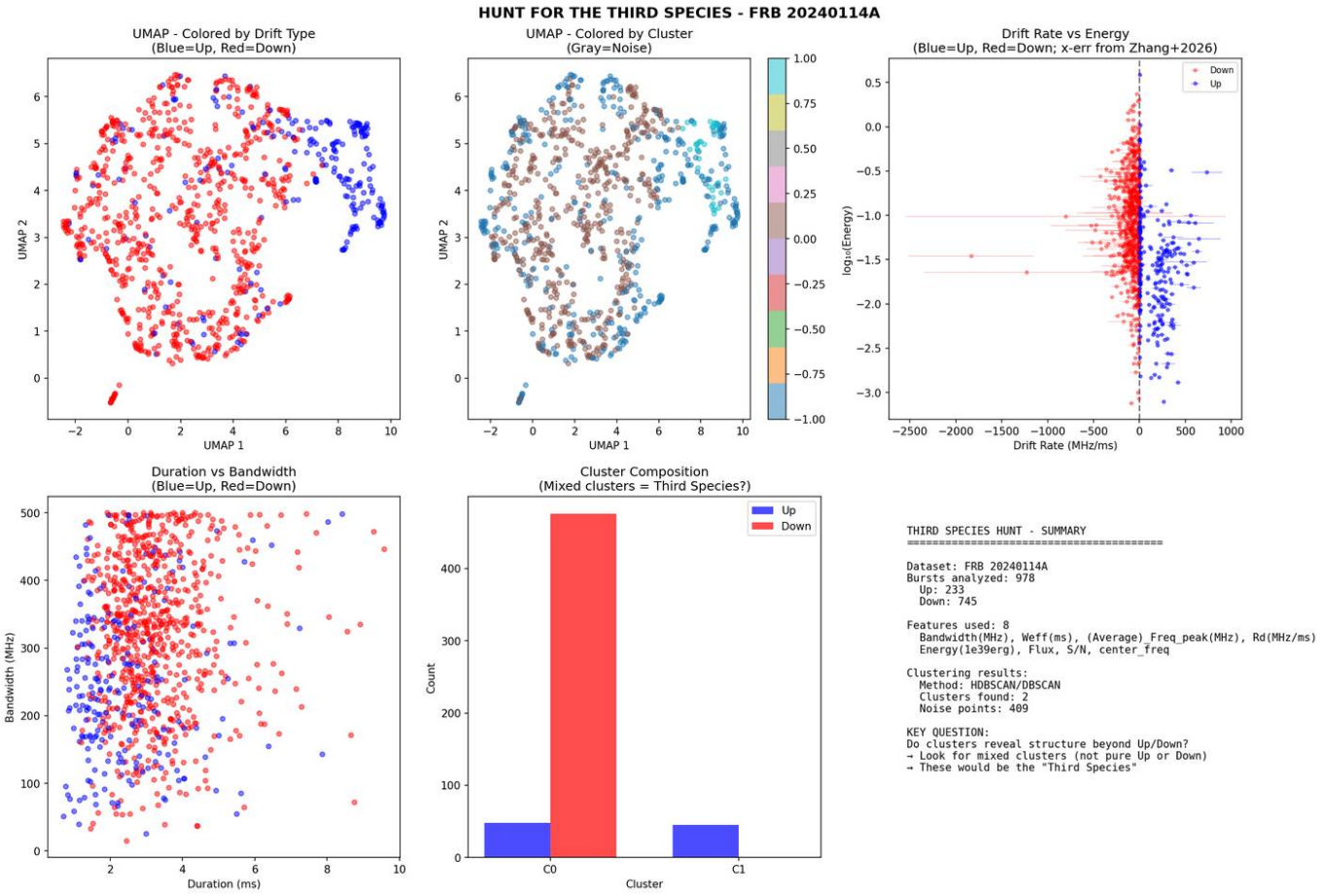

Figure 1. UMAP dimensionality reduction of 978 burst clusters colored by drift type (top-left: blue=Up, red=Down) and by HDBSCAN cluster assignment (top-center). Cluster C1 (45 burst clusters) forms a distinct island of exclusively upward-drifting burst clusters **identified through multi-dimensional density-based clustering. The top-right panel shows drift rate vs. log-energy with drift-rate error bars from the Zhang et al. (2026) catalogue; individual error bars are visible on the low-drift-rate points, and the U1 sample used in Section 3.3 excludes all burst clusters whose drift rates are consistent with zero within these uncertainties.**

4. DISCUSSION

The bimodal drift rate structure within upward-drifting burst clusters **is suggestive of** two distinct emission mechanisms or regions. As demonstrated in Section 3.3, this bimodality persists when restricting to single-component (U1) burst clusters, ruling out the possibility that the signal arises from heterogeneous drift rate definitions across morphological subclasses. **The U1-only analysis additionally rules out measurement-error contamination: both modes (∼113 and ∼300 MHz/ms) lie well above the error floor, and zero of 142 U1 burst clusters have drift rates consistent with zero within their measurement uncertainties.** All 45 burst clusters in the extreme-drift Cluster C1 are single-component, further supporting the interpretation that the bimodality reflects intrinsic physical differences rather than definitional effects.

In the magnetar model, drift rate is related to the radius-to-frequency mapping in the magnetosphere (Lyutikov 2020). Higher drift rates correspond to steeper frequency-time gradients, potentially indicating emission from regions with different magnetic field configurations or at different magnetospheric altitudes. The systematic frequency offset (−7%) between populations supports spatial separation: the extreme-drifters may originate from a distinct altitude or magnetic flux tube. The shorter durations (−29%) suggest a more compact emission region with a shorter coherence timescale.

Sub-burst drift rates have been interpreted as evidence for radius-to-frequency mapping in repeating FRBs (Hessels et al. 2019), and spectro-temporal correlations observed across FRB populations (Rajabi et al. 2020; Chamma et al. 2023) suggest common underlying physics. The bimodal structure we identify within a single morphological class **may be consistent with** two

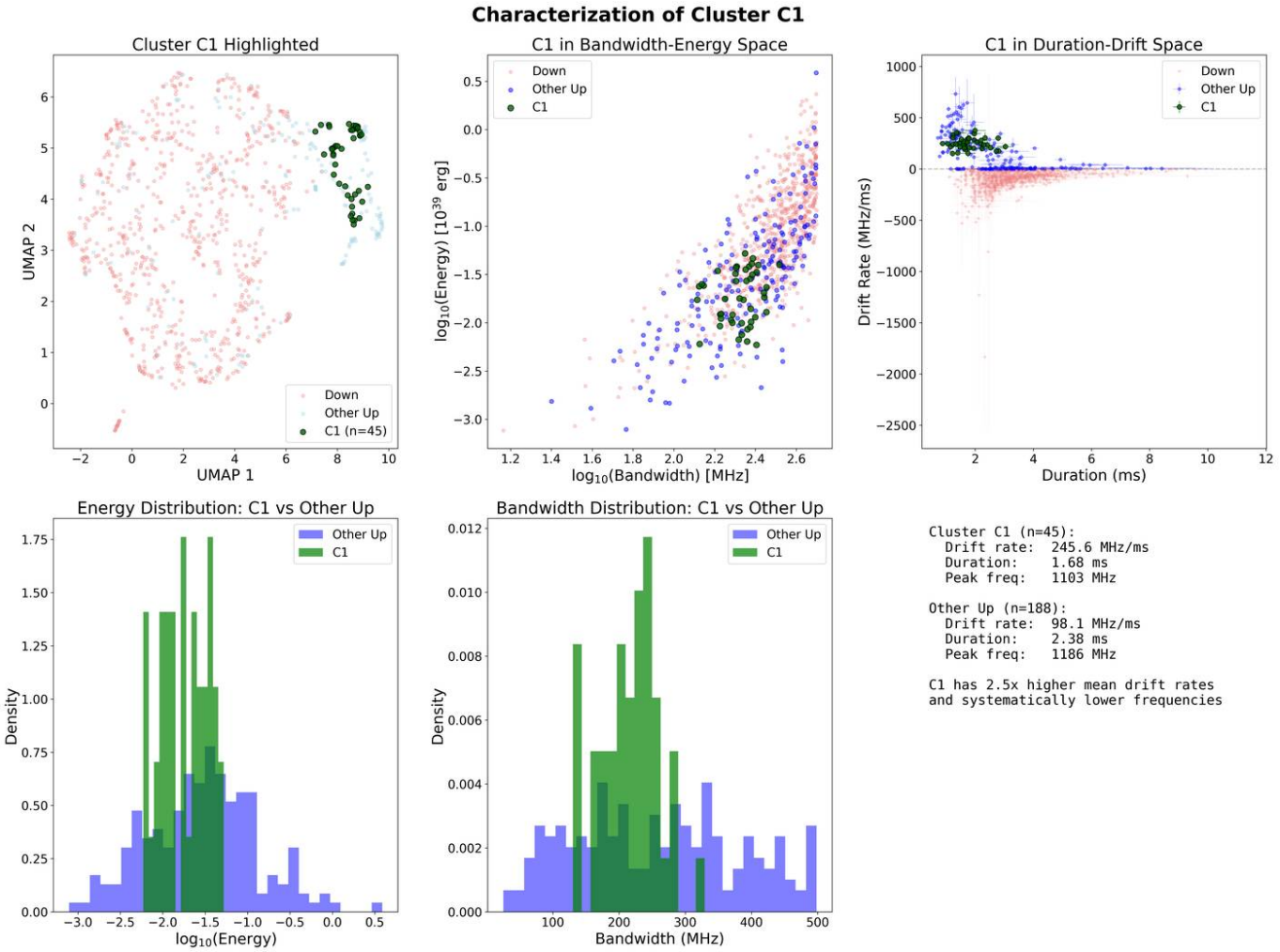

Figure 2. Characterization of Cluster C1 (green) compared to other upward-drifting burst clusters (blue) and downward-drifting burst clusters (red). C1 occupies a distinct region in bandwidth-energy and duration-drift parameter spaces.

preferred emission altitudes or two distinct triggering mechanisms within the magnetosphere**, though further theoretical modeling would be needed to distinguish between these scenarios**.

This discovery demonstrates the power of unsupervised machine learning for identifying hidden structure in FRB populations beyond traditional classification schemes (Pleunis et al. 2021). Future work should investigate whether similar substructure exists in other repeating FRBs observed by FAST and CHIME (CHIME/FRB Collaboration 2021), and whether the extreme-drift subpopulation exhibits distinct temporal or polarimetric properties.

We emphasize several important caveats. First, our analysis is based on data from a single observing epoch (2024 March 12) of one source; persistence across other observing dates of FRB 20240114A and presence in other repeating FRB sources remain to be established by future observations. Second, while the U1-only analysis eliminates the most important systematic (heterogeneous drift rate definitions) and rules out measurement-error contamination, other systematics such as frequency-dependent selection effects may still exist. Third, propagation effects—scintillation, frequency-dependent scattering, and plasma lensing along the line of sight through the magnetosphere, host galaxy, intergalactic medium, and Milky Way—could in principle alter observed burst morphology; multi-epoch observations of the same source, where propagation conditions evolve while intrinsic emission geometry does not, would discriminate between propagation and intrinsic origins. Fourth, the physical interpretation via dual emission regions, while consistent with the data, is not unique.

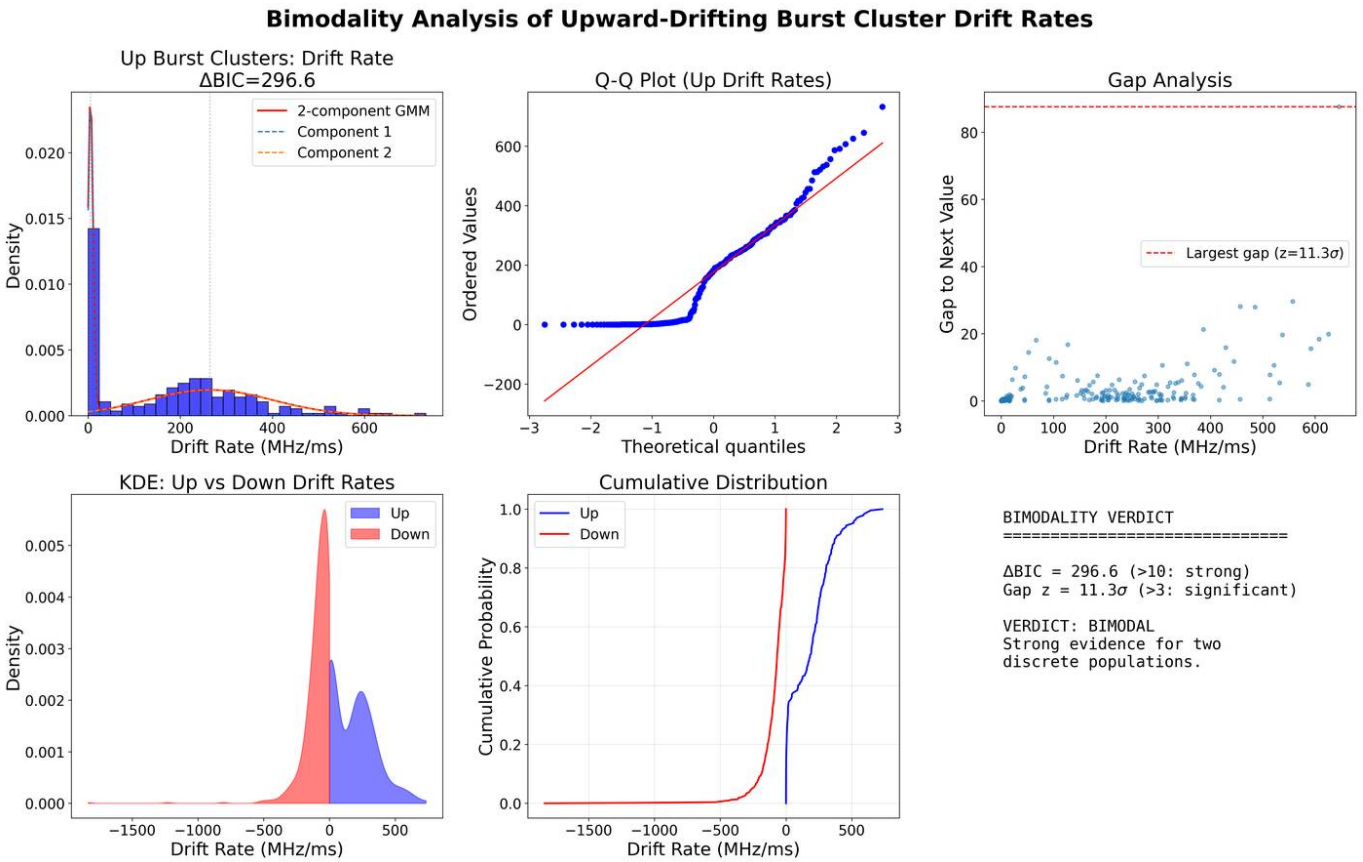

Figure 3. Bimodality analysis of upward-drifting burst cluster drift rates. Top left: histogram with 2-component Gaussian mixture model fit. Top right: Q-Q plot showing deviation from unimodal normal distribution. Bottom panels: gap analysis and KDE comparison between upward- and downward-drifting populations.

5. CONCLUSIONS

We report the discovery of bimodal structure within morphologically classified FRB burst clusters. The 45-burst cluster "extreme-drift" subpopulation (Cluster C1) in FRB 20240114A exhibits:

1. 2.5× higher **mean** drift rates (ΔBIC = 296.6 for full sample; 19.9 for single-component only)
2. Systematically lower emission frequencies (-7%)
3. Shorter burst cluster durations (-29%)
4. Distinct clustering in multi-dimensional feature space

Importantly, this bimodality is confirmed in single-component (U1) burst clusters alone (Ashman's D = 2.71, gap = 9.2σ), demonstrating that it is not an artifact of combining burst clusters with different internal morphologies. **All C1 members are single-component (U1) burst clusters with drift rates far exceeding their measurement errors. These findings are suggestive of dual emission regions in the FRB 20240114A magnetosphere, pending confirmation from multi-epoch and multi-source studies.**

We thank the FAST FRB Key Science Project Collaboration for making the FRB 20240114A morphology dataset publicly available. We also thank the anonymous referee for constructive comments that significantly improved this manuscript. This research made use of NumPy, SciPy, Pandas, scikit-learn, UMAP, and HDBSCAN. This work was conducted at Blankline (blankline.org).

AI DISCLOSURE

Artificial intelligence assistants were used during manuscript preparation: Primus (Blankline) for statistical analysis design, and Claude (Anthropic) for manuscript writing and text drafting. The author takes full responsibility for the scientific content, has verified all results independently, and confirms that all claims and interpretations represent the author's own scientific judgement.

DATA AVAILABILITY

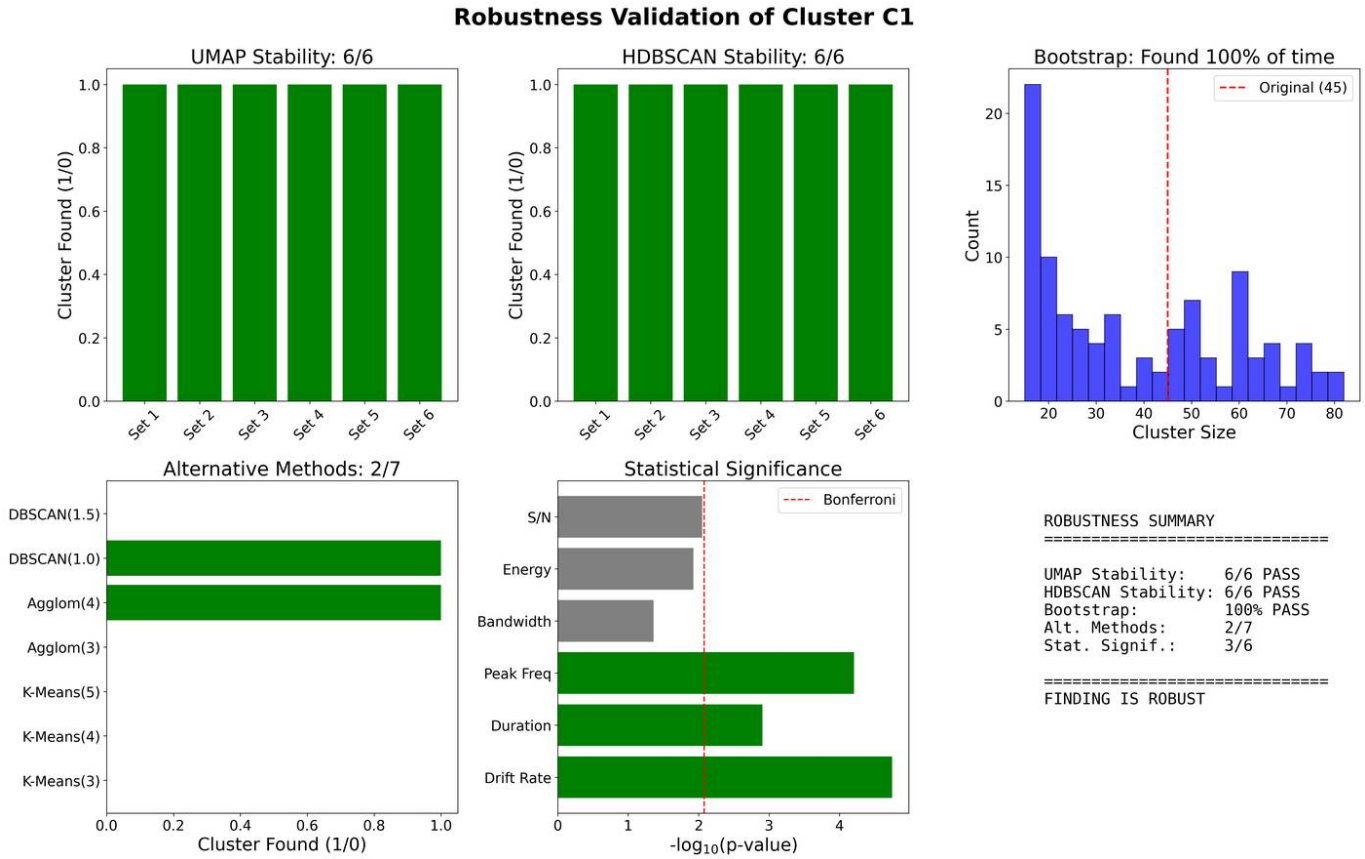

Figure 4. Robustness validation of the extreme-drift cluster C1. The cluster is consistently identified across different UMAP parameters (top left), HDBSCAN parameters (top center), and bootstrap resamples (top right). Statistical tests confirm significant property differences (bottom).

The FRB 20240114A Morphology Dataset is publicly available from Zhang et al. (2026).

Facilities: FAST

Software: NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), Pandas (McKinney 2010), scikit-learn (Pedregosa et al. 2011), UMAP (McInnes et al. 2018), HDBSCAN (Campello et al. 2013), Matplotlib (Hunter 2007)

REFERENCES

Ashman, K. M., Bird, C. M., & Zepf, S. E. 1994, The Astronomical Journal, 108, 2348, doi: 10.1086/117248

Campello, R. J. G. B., Moulavi, D., & Sander, J. 2013, Advances in Knowledge Discovery and Data Mining, 160, doi: 10.1007/978-3-642-37456-2_14

Chamma, M. A., Rajabi, F., Wyenberg, C. M., Mathews, A., & Houde, M. 2023, Monthly Notices of the Royal Astronomical Society, 521, 1990, doi: 10.1093/mnras/stad269

CHIME/FRB Collaboration. 2021, The Astrophysical Journal Supplement Series, 257, 59, doi: 10.3847/1538-4365/ac33ab

Harris, C. R., et al. 2020, Array programming with NumPy

Hessels, J. W. T., et al. 2019, The Astrophysical Journal Letters, 876, L23, doi: 10.3847/2041-8213/ab13ae

Hunter, J. D. 2007, Matplotlib: A 2D Graphics Environment

Lorimer, D. R., Bailes, M., McLaughlin, M. A., Narkevic, D. J., & Crawford, F. 2007, Science, 318, 777, doi: 10.1126/science.1147532

Lyutikov, M. 2020, The Astrophysical Journal, 889, 135, doi: 10.3847/1538-4357/ab55de

McInnes, L., Healy, J., & Melville, J. 2018, arXiv e-prints. https://arxiv.org/abs/1802.03426

McKinney, W. 2010, Data Structures for Statistical Computing in Python

Pedregosa, F., et al. 2011, Scikit-learn: Machine Learning in Python

Petroff, E., Hessels, J. W. T., & Lorimer, D. R. 2022, The Astronomy and Astrophysics Review, 30, 2, doi: 10.1007/s00159-022-00139-w

Pleunis, Z., et al. 2021, The Astrophysical Journal, 923, 1, doi: 10.3847/1538-4357/ac33ac

Rajabi, F., Houde, M., Bartkiewicz, A., et al. 2020, Monthly Notices of the Royal Astronomical Society, 498, 4936, doi: 10.1093/mnras/staa2723

Spitler, L. G., et al. 2016, Nature, 531, 202, doi: 10.1038/nature17168

Thornton, D., et al. 2013, Science, 341, 53, doi: 10.1126/science.1236789

Virtanen, P., et al. 2020, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python

Zhang, L.-X., et al. 2026, The Astrophysical Journal, 998, 276, doi: 10.3847/1538-4357/ae314a