Real-time injection of simulated FRBs into SPOTLIGHT

Raghav Wani[1], Ujjwal Panda[2], SPOTLIGHT Team

1. Indian Institute of Science Education and Research (IISER), Pune, India
2 : National Centre for Radio Astrophysics (NCRA-TIFR), Pune, India

Scan to visit SPOTLIGHT website

Background: FRBs and SPOTLIGHT

- Fast Radio Bursts (FRBs): Bright, millisecond radio flashes with dispersion measures (DMs) exceeding Galactic values → extragalactic origin.
- Over 845 discovered, including ~ 58 repeaters. Isotropically distributed, making it challenging to detect.
- SPOTLIGHT@GMRT: A multi-beam, commensal real-time transient survey at the upgraded GMRT, designed to search for FRBs and pulsars.
- Challenge: To study FRBs, pipelines must recover signals across a wide parameter space (DM, width, fluence, scattering) in real-time data streams.

Motivation for Injection in SPOTLIGHT

- Sources like known pulsars and repeaters are not ideal test signals as they don't span the entire FRB parameter space (DM, width, fluence, scattering)
- Controlled injections allow us to:
  - Benchmark the modules within SPOTLIGHT pipelines
  - Generate labeled datasets for training machine learning models
  - Quantify completeness and biases
- Goal: Build a reproducible framework to inject realistic FRB signals into uGMRT/SPOTLIGHT data streams for robust validation.

Methods - Simulation and Injection Framework

- The `simulate_search` package (Luo et al. 2022) generates synthetic FRBs with configurable parameters (DM, width, fluence, scattering, spectral index, arrival times)
- *Arachne*: a lightweight injection tool developed for SPOTLIGHT.
  - Attaches to the SPOTLIGHT's existing shared memory based ring buffer with GMRT data streams.
  - Creates a duplicate shared memory segment that contains the original data with injected pulses. The pipeline can then process data from this injected shared memory. (as shown in the Figure 1)
  - Injection instructions (time of arrival, DM, width, flux) are supplied from the `simulate_search` output.
  - Implements bit-shifting probability to correctly inject into 8-bit quantised data streams in real time. Hence, it does not simply add intensity, rather it probabilistically modifies digitized levels.

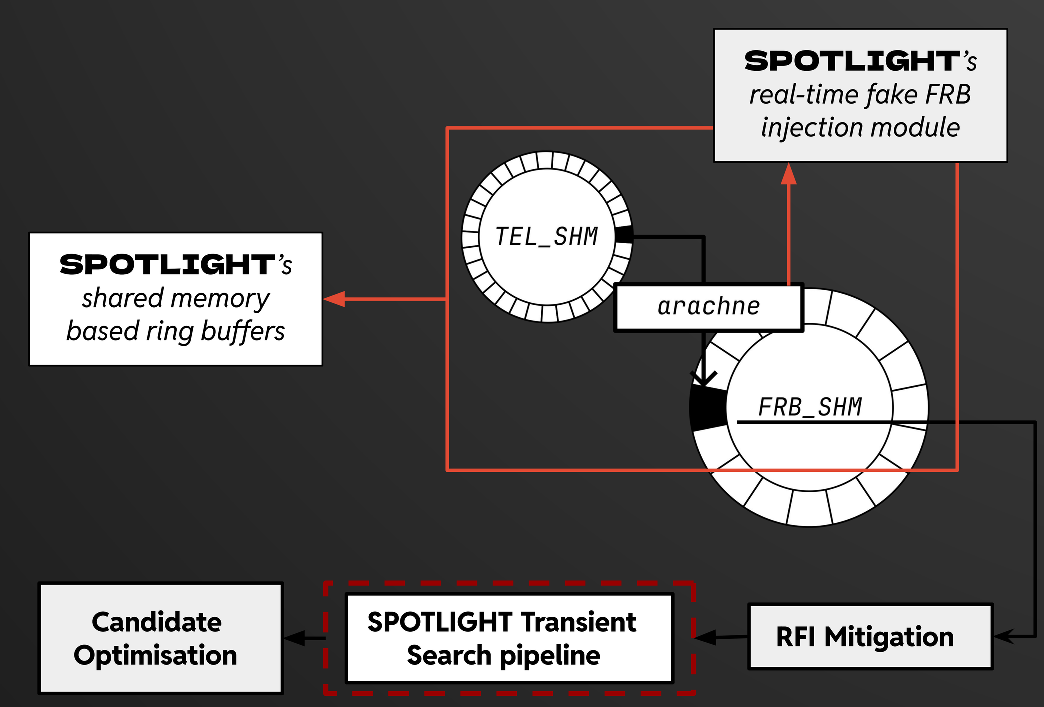

Figure 1: Flowchart of data processing within SPOTLIGHT

Injection Details 1

- Noise model: Assuming the background noise is Gaussian, and the noise level given by the radiometer equation

$$P(\mathcal{N}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\mathcal{N}-\mu)^2}{2\sigma^2}} \qquad \sigma = \frac{T_{sys}}{G\sqrt{N_{\mathrm{pol}}\delta t \Delta\nu}}$$

- Quantisation: For an $\mathcal{M}$-bit digitiser, data samples occupy $2^{\mathcal{M}}$ discrete states.
- Injection principle: Adding a signal of strength S shifts the probability of a sample crossing into a higher/lower quantisation state.

Contacts
Raghav Wani: wani.raghav@students.iiserpune.ac.in
Ujjwal Panda: upanda@ncra.tifr.res.in
For references and other details scan the adjacent QR code

Acknowledgements: This work uses uGMRT data. We thank the staff of the GMRT that made these observations possible. GMRT is run by the National Centre for Radio Astrophysics of the Tata Institute of Fundamental Research.

Injection Details 2

- The sequence of quantized states forms a Markov chain, as the probability of moving to the next state depends only on the current state, not on the sequence of past states.
- Transition probability (general case): Probability of bit shifting from the n^{th} quantized level to the $n+m^{th}$ level depends on the amplitude of the input signal S as:

$$P_{n\to n+m} = \frac{\phi(\min(\mu+(i+m)l\sigma - S, \mu+il\sigma)) - \phi(\max(\mu+(i+m-1)l\sigma - S, \mu+(i-1)l\sigma))}{\phi(\mu+il\sigma) - \phi(\mu+(i-1)l\sigma)}$$

  where ϕ is the Gaussian CDF and l is the level setting number for digitization.
- Transition between special boundary states (0 and $2^{\mathcal{M}} - 1$) are handled with modified formulas (*see references*).
- This framework gives a transition probability matrix Pij describing the probability of each digit changing state when a signal is injected.
- Ensures realistic simulation of how weak astrophysical bursts appear in real telescope data.

Results and Validation

- Consistency with Theory: Injection probabilities ($Pn\to m$) match analytic derivations for injects in 1-bit and 2-bit data mentioned in Luo et al. (2022, MNRAS), Appx B.
- Signal Recovery: Injected bursts are successfully detected by the downstream search in the SPOTLIGHT pipeline
- Bit-level probability framework reproduces the statistical behavior of weak signals in digitized data and avoids artificial "perfect" injections that do not reflect real telescope systematics.
- Dynamic spectrum (before vs. after injection) shows clear but realistic burst signatures with varying widths and flux at different ToA.

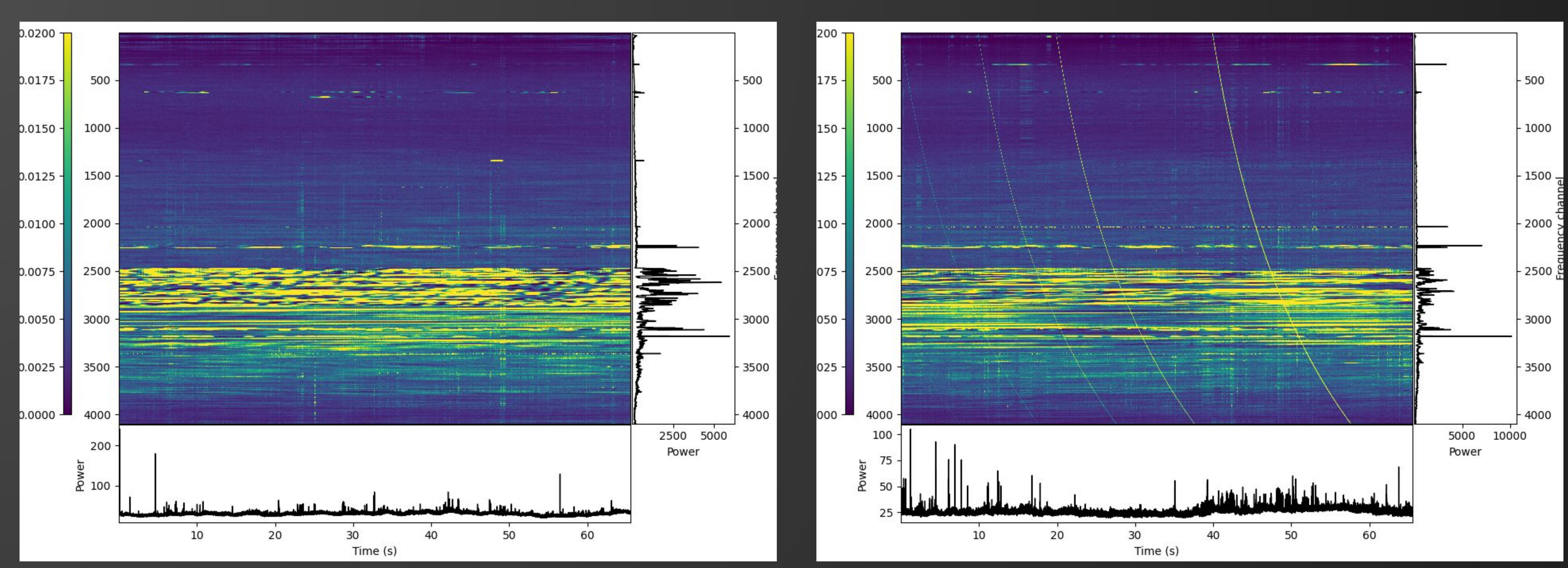

Figure 2 : Before (left) and after (right) injection of simulated burst into real SPOTLIGHT data

- Curated injected FRB datasets for training the FRB classifier within SPOTLIGHT, ensuring exposure to a diverse and realistic FRB parameter space. Current dataset: ~10K candidates covering
  - Band 3 (300-500 MHz); Band 4 (550-750 MHz); Band 5 (1260-1460 MHz and 1060-1460 MHz)
  - SNR: 10 to 55 σ (step = 5σ)
  - DM: 100–1000 pc cm⁻³ (step = 100) for Band 3; 200–2000 pc cm⁻³ (step = 200) for Bands 4 & 5
  - Width: 2, 4, 8, 16, 20, 32, 40, 50, 64 ms

Future Work

- Extend from single-beam to multi-beam injections across SPOTLIGHT's wide FoV. Implement beam-dependent SNR scaling based on primary beam models.

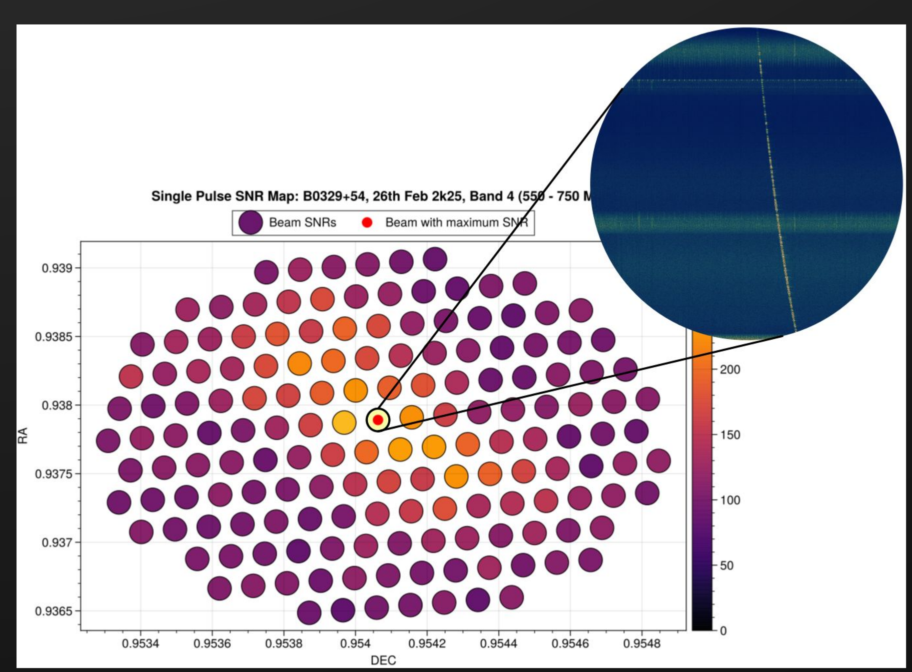

Figure 3: Multi-beam folded SNR Map for B0329+54 with burst visible at the centre

- Benchmark full SPOTLIGHT pipeline performance on a large injected dataset.
- Curate and release a publicly available simulated FRB dataset for community ML training.
- Incorporate injections in real-time observing runs for continuous validation.