

Low Overhead Multi-Source RFI Cancellation

Shuvam Chakraborty*, Dola Saha*, Aveek Dutta* and Gregory Hellbourg†

*Department of Electrical and Computer Engineering, University at Albany, SUNY

†Department of Astronomy, California Institute of Technology

Email: *schakraborty@albany.edu, *dsaha@albany.edu, *adutta@albany.edu, †ghellbourg@astro.caltech.edu

Abstract—Radio Frequency Interference (RFI) from cellular and other communication networks is commonly mitigated at passive user sites (e.g., radio telescope) without any active collaboration with the interfering sources. The expanding Universe and simultaneous proliferation of Earth-based and LEO communication infrastructure are causing unprecedented RFI that require collaborative strategies to maintain the scientific and societal goals of each. However, collaborative cancellation methods carry the baggage of large communication overhead and hence is impractical to apply to multiple sources of RFI appearing simultaneously in time, frequency, or both. In this work, we develop a set of protocols based on the characterization of RFI and astronomical signals in the Eigen-domain independent of temporal and spectral characteristics. This allows us to reduce communication overhead based on incident RFI power and across time and frequency. Furthermore, the second stage of the protocol cancels RFI from multiple sources with the incident power-based priority that reduces redundancy in the cancellation and reconstruction process. Extensive analyses were performed using real-world astronomical signals captured using Deep Synoptic Array-110 (DSA-110) in Owens Valley Radio Observatory (OVRO) and simulated RFI from multiple base stations. Results show up to 63% reduction in communication overhead and computation complexity at the telescope is reduced from n^2 to n^d , where $1 < d < 2$ with astronomical signal recovery accuracy of 81.13%.

Keywords—Radio frequency interference mitigation, Radio astronomy, Passive spectrum sharing.

I. INTRODUCTION

The proliferation of next-generation (xG) communication networks [1] increasingly generates unwanted Radio Frequency Interference (RFI), even in bands that are protected for radio astronomy due to out-of-band emissions and inter-modulation products. At the same time, technological advances such as wideband and low system temperature receivers are allowing astronomers to observe weak astronomical emissions over even wider bands. This requires stringent mitigation techniques to continue to reach the goals of both active and passive spectrum users in the presence of challenging RFI. Although radio telescopes are generally located in geographically isolated areas farther from communication networks, it does not prevent RFI from nearby cellular networks. RFI mitigation techniques in practice use statistical signal analysis to detect RFI and discard the associated time and frequency bins from the collected data known as excision. This may at best reduce the sensitivity of the telescope, and at worse remove the astronomical signal of interest.

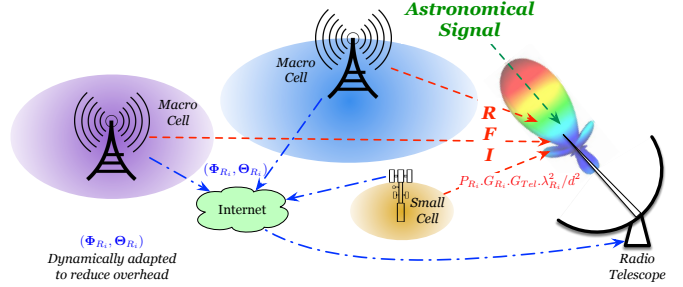


Fig. 1: Multiple Sources of RFI incident on radio telescope from different geographical locations.

Therefore, even with state of the art methods in RFI mitigation, full recovery of an astronomical signal cannot be achieved without prior knowledge of the RFI. Fortunately, communication signals can be *characterized* accurately and made available to the telescope through collaboration, which can be *intelligently cancelled* from the telescope data to reveal the astronomical signal as pioneered in our previous work [2]. We have proposed signal decomposition employing the Karhunen–Loève Transform (KLT) to explore the RFI signal space at its source before sharing the information with the telescope for eventual RFI cancellation.

However, radio telescopes have a wide range of view in time, frequency, and space as they generally have large bandwidth, scan a significant portion of the sky including degrees of freedom in azimuth and elevation, and have a long exposure time, ranging from minutes to hours. This feature along with low system temperature makes them susceptible to RFI contamination from diverse sources. One of the strongest and most prevalent sources of terrestrial RFI is 4G/5G cellular networks and is considered as the prime source of RFI in this work. Figure 1 shows a candidate scenario where multiple macro and small cell base stations can inject RFI in a radio telescope. The RFI generated from all these sources can be starkly different due to differences in transmit power, antenna gains, location, propagation characteristics, active resource blocks, requirements of different service providers etc. To alleviate RFI all of these sources are required to share information with the radio telescope periodically or in a continuous manner leading to enormous data processing needs both at the RFI source and the telescope and a large communication overhead. To provide context, the collaborative RFI cancellation method in [2] leads to communication overhead of several gigabits per second (Gbps), which will be increased manifold with the

number of sources, observation time, and bandwidth.

Although several techniques for information compression have been proposed in different domains of application, they come with their own set of baggage. These methods can be broadly classified in two sets - a) lossless compression, b) lossy compression. Lossless methods maintain high precision but offer little to no benefits in terms of communication overhead and storage requirements. Additionally, to get the original information, a reconstruction step is necessary adding to the computation cost. On the other hand, lossy compression, classified as traditional methods and data-driven methods, has the obvious drawback of permanent information loss. Also, the data-driven machine learning (ML) methods being treated as black box limits the degrees of freedom in data recovery.

In this work, we propose a bidirectional collaboration and communication overhead reduction methodology for RFI cancellation that has several unique benefits - a) minimal information loss due to overhead reduction, b) no requirement of additional data recovery step, and c) reduction in memory and computation cost without affecting the RFI mitigation and signal reconstruction quality, d) no hand-shake among RFI sources required even if they overlap in time or frequency. The proposed method primarily sets a priority for the RFI sources to participate in collaborative cancellation based on the instantaneous incident power at the telescope and its impact on standard operation on the telescope. This prevents blind information sharing and cancellation, which can be subjectively redundant. Furthermore, this method exploits the unique benefit of signal characterization using Karhunen–Loève Transform (KLT) [3], [4]. The characterization being independent of changing temporal and spectral statistics of the signal allows a significant reduction in communication overhead. This method can also be adapted to any interference cancellation problem assuming an active collaboration is possible with the interfering source.

The contributions of this work can be summarized as:

- 1) We have developed a novel communication overhead reduction method in collaborative RFI cancellation from multiple sources.
- 2) This method sets priority for collaboration with RFI sources based on the incident power at the radio telescope leading to reduction in aggregated computation load without compromising the accuracy of signal recovery.
- 3) This method exploits properties of KLT for signal characterization to reduce communication overhead across time and frequency.
- 4) The proposed method is evaluated with real astronomical signals and simulated RFI from multiple downlink LTE transmissions and is compared to state-of-the-art techniques.

II. RELATED WORK

Interference cancellation in communication systems: Interference cancellation in wireless communication [5], [6], require decoding the strongest signal first in order to re-

cover the intended signal. Also, for active spectrum users, collaboration among wireless technologies [7], [8], [9], or avoiding incumbents [10] have been utilized that can not be implemented for the coexistence of active and passive users due to the Gaussian nature of signal captured at the telescope, very low power of RFI preventing identification and decoding to be applied before cancellation, and prohibitive bandwidth due to large communication overhead for sharing raw RFI data. The proposed method is capable of overcoming most of these shortcomings of existing methods.

RFI mitigation in radio astronomy: With proliferating communication infrastructure, active RFI mitigation has become a necessary practice in the radio astronomy community. In parts of the radio spectrum - for known persistent and fixed sources of RFI, strong attenuation is applied at the front-end of the receiver using a series of analog superconductive filters [11], [12], whereas data flagging and discarding is done in practice for unknown sources of RFI. This process is generally done using conventional data reduction software that includes an automated flagger based on local and global statistics of a given dataset [13], [14]. Real-time data flaggers [15]. Machine learning-based methods have also been proposed for RFI detection [?] and excision [16]. Adaptive spatial filters are proposed to extract RFI spatial signature and recover the astronomical signal [17]. Time domain nulling has been demonstrated in [18], however, it suffers from high complexity issues. Finally, collaborative cancellation methods are proposed in [2], [19] for singular sources and in Despite high accuracy, communication overhead and complexity issues may prove challenging for such methods.

Data Compression Methods: Compressing and/or reducing communication overhead is of utmost importance for collaborative cancellation methods considering the scaling of the system across different RFI sources. Data compression techniques can broadly be classified in two ways - lossless and lossy transmission. Lossless compression methods exploit statistical redundancy in data with prime examples of general compression [20], [21], audio compression [22], and image compressions [23]. However, lossless methods prioritize retaining information and communication overhead reduction becomes insignificant. Lossy compression/reduction methods have been applied in a variety of applications as well such as traditional methods [24], [25] and machine learning-based methods [26], [27] in image, video, I-Q data compression etc. However, lossy methods not only suffer from significant information loss, it incurs a significant computation load for data recovery which is not desirable for multi-source RFI cancellation.

III. BACKGROUND

A. Acquisition of astronomical signals

A radio telescope achieves its high sensitivity by maximizing its directivity, collecting areas, and minimizing the system temperature of its receivers. Typically they are equipped with large arrays of antennas either phased together to produce mul-

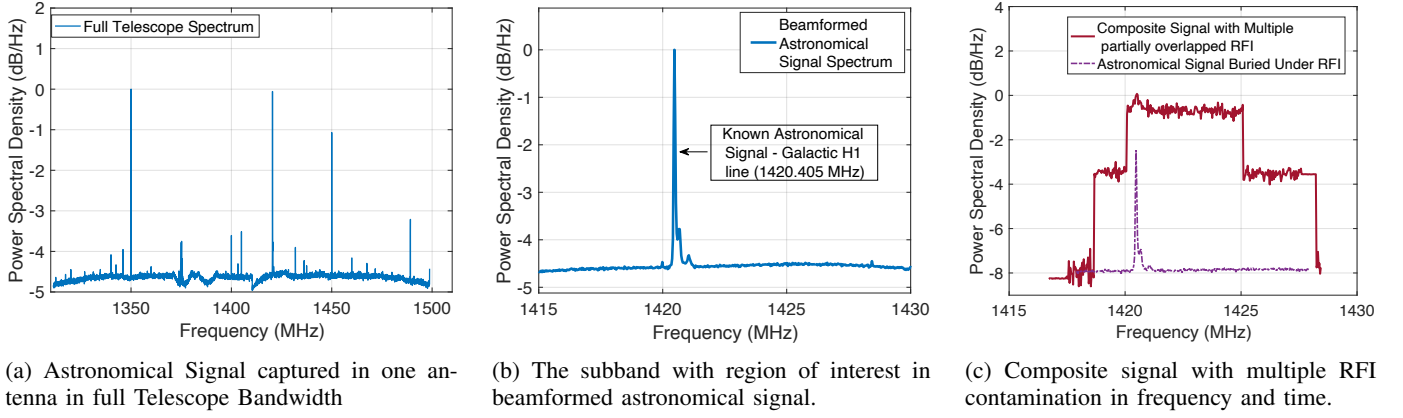


Fig. 2: Spectral characteristics of the astronomical signal and RFI contaminated composite signal.

Bandwidth (MHz)	1.25	2.5	5	10	15	20
Occupied BW (MHz)	1.140	2.265	4.515	9.015	13.515	18.015
Frame (ms)	10					
Subframe (ms)	1					
N_{FFT}	128	256	512	1024	1536	2048
N_{guard}	52	105	211	423	635	847
Resource Blocks	6	12	25	50	75	100

TABLE I: Different LTE parameters produce unique RFI

multiple beams in the sky or to perform interferometric synthesis imaging [28]. After equalization and filtering the captured signal, the output of the individual receivers are digitized over hundreds of MHz and channelized into smaller frequency bins of hundreds of kHz in bandwidth [29]. Channelization is useful for reducing the data rate for real-time processing, share computational resources and excise RFI-corrupted channels before further processing.

B. Signal model for LTE RFI signal

Long term evolution (LTE) and 5G signals employ a multicarrier modulation scheme to maximize spectral efficiency called orthogonal frequency-division multiplexing (OFDM) with a variety of parameters defined by the 3rd generation partnership project (3GPP) standardization body. The general model for an OFDM signal for typical downlink transmissions using a carrier f_c is shown in (1).

$$x_R(t) = \text{Re} \left\{ e^{j2\pi f_c t} \sum_{k=-N_{FFT}/2}^{N_{FFT}/2} \alpha_k e^{j2\pi k(t-t_g)/T_u} \right\} \quad (1)$$

The range of LTE parameters shown in Table I is utilized to generate downlink RFI signals from various sources (BS) for evaluation of the proposed system.

C. Signal model for astronomical signals

The RFI contaminated signal output of a single telescope antenna (referred to as *composite signal*) is expressed as (2):

$$x_T[n] = x_A[n] + x_N[n] + x_R[n] \approx x_N[n] + x_R[n] \quad (2)$$

where $x_T[n]$ is the channelized composite signal at a filterbank channel centered around frequency f_c and at time sample n , and follows a stationary stochastic process, is independently and identically distributed (i.i.d.) with $x_T[n] \sim \mathcal{NC}(x_R[n], \sigma^2)$.

$\mathcal{NC}(\mu, \Gamma)$ indicates the stationary circular complex Gaussian distribution with mean μ and covariance Γ . $x_A[n] \sim \mathcal{NC}(0, \sigma_A^2)$ is i.i.d. and represents the accumulated contribution of all astronomical sources in the field of view of the telescope, $x_N[n] \sim \mathcal{NC}(0, \sigma_N^2)$ is i.i.d. and represents the system noise contribution, and $x_R[n]$ is the deterministic RFI contribution.

IV. SYSTEM DESIGN

A. Collaborative RFI Cancellation

Collaborative RFI cancellation has been shown to be a unique concept [2] where: 1) The RFI is decomposed at the source into a compact yet accurate eigenspace using KLT and extracting the bases from the signal itself, which adapts with time-varying cellular RFI. This is particularly important because this method is independent of time-frequency domain characteristics of the signal, is applicable to both deterministic and stochastic signals and is able to detect weak signals below the noise floor. First, the autocorrelation matrix $\mathbf{R}_{xx}(n_1, n_2) = \mathbb{E}[\mathbf{U}\mathbf{U}^H]$ of the signal of interest $x[n]$ is generated by embedding delayed versions of $x[n]$ in a Hankel matrix $\mathbf{U}$. Decomposition of the autocorrelation matrix generates:

$$\mathbf{R}_{xx} = \mathbf{\Phi} \mathbf{\Lambda} \mathbf{\Phi}^H, \quad \text{where, } \mathbf{\Lambda} = \text{diag} \{ \lambda_1, \dots, \lambda_L \} \quad (3)$$

where λ_j are the eigenvalues, with $j \in [1, L]$, and $\mathbf{\Phi}$ is a unitary matrix containing L eigenvectors as its columns. L is the KLT window length. The necessary and sufficient conditions for the choice of L are shown in [2]. Since (3) decomposes the temporal correlations of $x[n]$, each column of $\mathbf{\Phi}$, i.e. ϕ_i , is a time-series, and is referred to as an eigenfunction. The set of eigenfunctions representing RFI is collectively referred to as RFI kernel $\mathbf{\Phi}_R$. This is periodically shared with the telescope via a shared channel over the Internet. At the telescope, the composite signal $x_T[n]$ is decomposed using the same method revealing its eigenspace $\mathbf{\Phi}_T$ that contains the RFI subspace. The shared RFI eigenspace is used to cancel the RFI from

Static Parameters	Dynamic Parameters
Observed frequency range [†]	Eigenspace (Φ_R)
Polyphase filter subchannel [†]	KLT window (L)
Telescope and BS location *	Transmission time
Antenna Gains *	Telescope and BS Orientation *

[†] parameters shared by telescope only, * shared by both, others are shared by RFI source only.

TABLE II: Shared parameters for collaborative cancellation

the eigenspace of the composite signal via complimentary orthogonal projections as given in (4) and (5)

$$\mathbf{P}_{\Phi_R}^\perp = \mathbf{I} - \Phi_R (\Phi_R^H \Phi_R)^{-1} \Phi_R^H \quad (4)$$

$$\hat{\Phi}_T = \mathbf{P}_{\Phi_R}^\perp \Phi_T \quad (5)$$

where $\hat{\Phi}_T$ is the projected composite signal subspace. Since the cancellation happens in the eigenspace, a final step to convert the eigenspace to the corresponding time-domain signal will reveal the RFI-free astronomical signal $\hat{x}_T[n]$. One advantage of this method is the fidelity of the decomposition is vastly improved at high signal power, which is maximum at the RFI source. However, the issues with large RFI information, communication with multiple sources of RFI, and subsequent cancellation lead to impractical burden of communication overhead and computation resource requirement at both the radio telescope and RFI sources.

B. Information Shared in Collaboration

To streamline the entire RFI cancellation and reconstruction process, a set of information is required to be shared between the radio telescope beyond the RFI kernel. These can be classified into two sets - static parameters (shared once and do not typically change over time), and dynamic parameters (changes instantaneously or over time and are shared periodically.) These include but are not limited to signal characterization information, topological information, transmission time, bandwidth, etc as given in Table II. Roles of these parameters in the RFI cancellation apparatus are explained in subsequent sections where applicable.

C. Evaluation Metric for RFI Cancellation

The quality of RFI cancellation and astronomical signal reconstruction is evaluated using the reconstruction quality factor (RQF). We can express the recovered astronomical signal as:

$$\hat{x}_T[n] = x_A[n] + \epsilon_R[n] \quad (6)$$

where $x_A[n]$ is the true astronomical signal and $\epsilon_R[n]$ is the residual RFI contamination. We express the RQF metric as:

$$\psi = f_{RQF}(\hat{x}_T[n]) = \frac{\sigma_\epsilon^2}{\sigma_A^2} = 1 + \frac{\sigma_A^2}{\hat{\sigma}_T^2} - 2 \cdot \frac{\hat{\sigma}_{T,A}}{\sigma_A^2} \quad (7)$$

where σ_x^2 and $\sigma_{x,y}$ are variance and covariance operators. In practice the variance of astronomical signal is replaced by that of the system noise floor at the telescope. Uniqueness of this metric are: 1) it is directly related to the International Telecommunication Union (ITU) detrimental-level interference criterion [30], 2) it is not susceptible to the appearance of

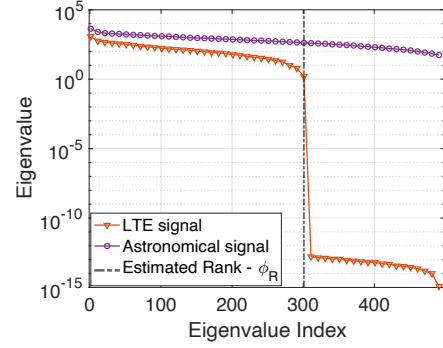


Fig. 3: Rank Indicator of Different Signals

large peaks in the astronomical signal capture and maintains sensitivity to RFI contamination, and 3) this metric is robust to distortion in statistically Gaussian properties of the astronomical signal due to post-processing at the telescope (e.g, quantization, phase rotation etc). The ITU-determined criterion is set to 10% distortion ΔP_H in telescope sensitivity over a 2000-second long integration defined as:

$$\Delta P_H = 0.1 \Delta P \Delta f \quad \text{where,} \quad \Delta P = \frac{P}{\sqrt{\Delta f \cdot t}} \quad (8)$$

where P is the power spectral density (PSD) at the telescope, ΔP is the sensitivity in terms of power spectral density, Δf is the bandwidth, and t is the integration time. We claim without loss of generality $\sigma_A^2 = P_t = P \cdot \Delta f$. If the variation in total power to $P'_t = P_t \pm \delta_\epsilon^2$ changes the PSD to P' and sensitivity to $\Delta P'$ then:

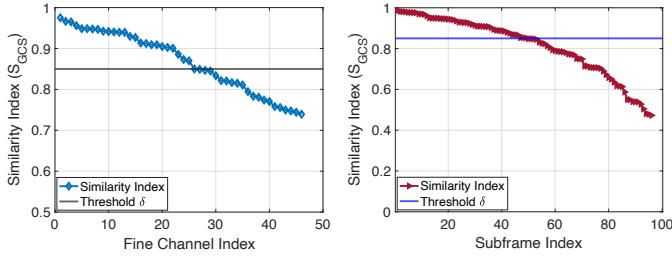
$$\hat{\sigma}_\epsilon^2 = P'_t \mp P_t = (P' \mp P) \cdot \Delta f = (\Delta P' \mp \Delta P) \cdot \Delta f \cdot \sqrt{\Delta f \cdot t} \quad (9)$$

Hence, for a given observation bandwidth Δf and integration time t being constant, $\hat{\sigma}_T^2$ is can be approximated to be linearly related to the change in telescope sensitivity.

Finally, incorporating the ITU recommended criterion (8) in (9), we get the upper bound of RQF (ψ') = 0.1. The lower bound is achieved in ideal astronomical signal capture at RQF = 0. Analysis for this is trivial and is skipped in this discussion. In presence of RFI, we can not find the true system noise temperature. Thus, a calibrated estimate or a recent historical average for the same can be utilized for comparable atmospheric conditions for RQF calculation.

D. Collaboration Based on Incident RFI Power

In this step, we reduce the number of required eigenfunctions shared with radio-telescope to sufficiently represent the RFI kernel leading to a reduction in communication overhead. In [2], the number of eigenfunctions is chosen based on the rank estimate (N_R) of the RFI signal space. Figure 3 shows eigenvalues of a downlink LTE signal and an astronomical signal where each eigenvalue indicates the strength of the signal along the direction of its complement eigenfunction. The sharp drop in eigenvalue for LTE simply indicates the rank as LTE, being a modulated signal, is spanned by a finite number of eigenfunctions. However, the RFI is characterized at the source and hence is at a much higher signal power level compared to the incident RFI at the telescope which is at a



(a) Similarity in Eigenspace across frequency (b) Similarity in Eigenspace across time

Fig. 4: Similarity in RFI eigenspace in time and frequency.

much lower power level depending on several factors such as distance between the RFI source and telescope, antenna gains at the transmitter side and the telescope, and propagation conditions. Thus we can efficiently represent the signal subspace for this incident RFI with fewer eigenfunctions without losing crucial information. We define a driving parameter ρ that is proportional to the incident power of the telescope at a given time:

$$\rho \propto P_R \cdot G_R \cdot G_{Tel} \cdot \lambda_R^2 / d^2 \quad (10)$$

where P_R is the transmit power at RFI source, G_R and G_{Tel} are the directive gains at RFI source and telescope respectively, d is the link distance, and λ_R is the wavelength of incident RFI. We don't consider other small-scale fading effects as due to the geographical isolation of telescopes, free space path loss (FSPL) becomes the key factor in the link budget. Based on the shared parameters in Table II, ρ can be estimated for any given scenario of RFI injection. The maximum possible value of the parameter can be estimated $\rho_{max} = k \cdot P_R \cdot G_R \cdot G_{Tel}^{max} \lambda_R^2 / d^2$ where k is the proportional constant G_{Tel}^{max} is the maximum possible sidelobe gain. λ_R and d typically will remain constant. Other parameters can be subjective to antenna activities of RFI sources. The number of eigenfunctions per fine channel (N_{Φ_R}) for each source is chosen as:

$$N_{\Phi_R} = (\rho / \rho_{max})^x \cdot N_R \quad (11)$$

It is evident that $x \in [0, 1]$ is a fair choice with 0 being the full precision scenario. The optimal value of x is chosen empirically in §VI-A. This method is unique because -a) we discard low-impact eigenfunctions while achieving accurate RFI characterization information, b) the communication overhead is significantly improved by reduction in RFI subspace, c) no additional information is required to compensate for the reduction in overhead, and d) this process leads to higher fidelity of cancellation and reconstruction compared to the obvious alternative of characterizing the RFI at the telescope. Characterization of incident RFI at telescope is not practical as infrastructure (e.g, reference antennas with specific orientations to achieve directive gain) and computation scaling will be needed for proliferating RFI sources.

E. Communication Overhead Reduction in Frequency

The eigenfunction produced from RFI characterization is typically done for a narrowband (tens of kHz) based on the

fine channel width of the telescope. Due to this choice of processing narrowband signals, the spectral properties remain considerably similar for several fine channels. Hence, these channels can be approximated with the same set of eigenfunctions. Since eigenfunctions indicate directions of orthogonal signal components, the cosine similarity is utilized to measure the degree of similarity of eigenfunctions across fine channels given in (12):

$$S_{GCS}(v_i, v_j) = \frac{|v_i v_j|}{\|v_i\| \cdot \|v_j\|}, \text{ for any two vectors } v_i, v_j \quad (12)$$

It represents angular separation among the two sets of eigenfunctions with 1 meaning identical and 0 orthogonal. Figure 4a shows cosine similarity across fine channels averaged over the top N_{Φ_R} eigenfunctions for each fine channel given as:

$$\hat{S}_{GCS}(\Phi_R^i, \Phi_R^j) = \frac{1}{N_{\Phi_R}} \cdot \sum_{k=1}^{N_{\Phi_R}} \frac{|\phi_R^{i,k} \phi_R^{j,k}|}{\|\phi_R^{i,k}\| \cdot \|\phi_R^{j,k}\|} \quad (13)$$

F. Communication Overhead Reduction in Time

If the RFI under consideration does not change significantly in time, a similar overhead reduction method as in frequency can be applied for continuous time RFI signals as well. Figure 4b shows the cosine similarity of a downlink LTE signal. Since LTE is quantized in time in terms of frames (10 ms), the time axis represents number of LTE frames.

However, we cannot indefinitely utilize the same set of eigenfunctions for all fine channels since the similarity reduces both with increasing frequency and in time as shown Figures 4a and 4b and this will diminish the quality of RFI cancellation process. Thus, a cut-off ($\hat{S}_{GCS} = \delta$) is decided upto which, this approximation is valid. In this work, we empirically determine this cut-off value in §VI-A.

G. Updated RFI Characterization at Source

Based on the steps discussed above for reducing communication overhead depending on incident power and similarity in eigenfunctions across both time and frequency, Algorithm 1 is developed that encompasses the RFI characterization system for any given RFI source.

In this algorithm, for eigenfunctions above cosine similarity the cutoff δ , a set of fine channels are approximated by eigenfunctions of one fine channel. Below the cut-off, a new set of eigenfunctions is generated to represent another set of fine channels. This is possible due to a limited number of fine channels spanning the RFI bandwidth. However, in time, this may prove challenging due to long observation times at the telescope. Hence, across time, when the $\hat{S}_{GCS}$ cut-off is reached, we introduce a parameter Θ . Since eigenfunctions are directional components of the signal subspace, we store the angular deviation among the eigenfunctions of specific indices in time given as:

$$\Theta(\Phi_R^i, \Phi_R^j) = \cos^{-1}(S_{GCS}(\Phi_R^i, \Phi_R^j)) \quad (14)$$

Hence the complete set of RFI characterization information from a source is provided by the set $\{\Phi_R, \Theta\}$.

Algorithm 1: RFI Characterization at Source

Data: $x_R[n], \rho, \rho_{max}, W_D^T, W_D^R, W_S^T, \delta$
Result: Φ_R

```

1  $x_R^T[n] = f_{filt}(x_R[n]);$  /*  $f_{filt}$  = filterbank function */
2  $N_{\Phi_R} \leftarrow (\rho/\rho_{max})^x;$ 
3  $f_{KLT}(R_{xx}) = \Phi \Lambda \Phi^H;$ 
4  $\Phi_R \leftarrow \Phi(1:N_{\Phi_R});$ 
5  $j \leftarrow 1;$  /*  $N_f$  = num fine channels */
6 while  $j < N_f$  do
7    $\Phi_R^* \leftarrow \Phi_R^1;$ 
8   Estimate  $\hat{S}_{GCS}(\Phi_R^*, \Phi_R^j);$ 
9   if  $\delta \geq \hat{S}_{GCS}$  then
10     $\Phi_R^j \leftarrow \Phi_R^*;$ 
11  else
12     $\Phi_R^* \leftarrow \Phi_R^j;$ 
13  end
14  while  $t < N_t$  do
15     $\Phi_R^{j,*} \leftarrow \Phi_R^{j,1};$  /*  $N_t$  = num time units */
16    Estimate  $\hat{S}_{GCS}(\Phi_R^{j,*}, \Phi_R^{j,t});$ 
17    if  $\delta \geq \hat{S}_{GCS}$  then
18       $\Phi_R^{j,t} \leftarrow \Phi_R^{j,*};$ 
19    else
20       $\Phi_R^{j,t} \leftarrow \Theta^{j,t};$ 
21    end
22     $t \leftarrow t + 1;$ 
23  end
24   $j \leftarrow j + 1;$ 
25 end

```

H. Successive RFI cancellation and Reconstruction

The obvious next step is canceling RFI from multiple overlapping sources and astronomical signal reconstruction. The shared information $\{\Phi_R, \Theta\}$, provides the complete RFI kernels to the radio-telescope for each source. However, The key idea is RFI injected at different detrimental power levels are not equally prioritized for cancellation. We developed Algorithm 2 based on the relative power level of incident RFI. The incident power can be estimated from the shared static and dynamic parameters in Table II for each contaminating source. Additionally, we know the established stopping criterion based on telescope sensitivity in §IV-C and its relation to the RQF estimate. Algorithm 2 presents the cancellation process for N number of RFI sources.

The concept of prioritizing RFI with higher power is important because in a practical situation due to 1) geographical isolation of radio-telescope, 2) propagation conditions, and 3) RFI appearing at the null or high side-lobe rejection zone of the telescope, the RFI may actually be below the detrimental level requiring no additional processing for astronomical signal recovery. This algorithm avoids such redundant scenarios significantly reducing the computational burden at the telescope.

V. EXPERIMENTS**A. RFI Generation and Injection**

The LTE RFI signal from multiple sources is generated according to the parameters in Table I for duration compa-

Algorithm 2: Multi-source RFI Cancellation

Data: $\rho ([\rho_1, \dots, \rho_N]), \Phi_R^i (i \in [1, N]), \Phi_T, W_S^R, W_D^R, W_S^T$
Result: $\hat{\Phi}_T, \hat{x}_T[n] (1 \leq k \leq N)$

```

1  $\rho \leftarrow \text{sort}(\rho);$  /* In descending order */
2 Rearrange  $\Phi_R^i (i \in [1, N]);$  /* on indices of  $\rho$  */
3  $j \leftarrow 0;$ 
4  $\psi \leftarrow f_{RQF}(x_T[n]);$ 
5 if  $\psi \leq \psi'$  then
6    $\hat{\Phi}_T \leftarrow \Phi_T; \hat{x}_T[n] \leftarrow x_T[n];$ 
7 else
8   while  $j < N$  do
9      $j \leftarrow j + 1;$ 
10     $\mathbf{P}_{\Phi_R^j}^\perp \leftarrow f_{proj}(\Phi_R^j); \hat{\Phi}_T^j \leftarrow \mathbf{P}_{\Phi_R^j}^\perp \Phi_T^j;$ 
11     $x_T^j[n] \leftarrow f(\Phi_T^j);$ 
12     $\psi_j \leftarrow f_{RQF}(x_T^j[n]);$ 
13    if  $\psi_j \leq \psi'$  then
14       $\hat{\Phi}_T \leftarrow \hat{\Phi}_T^j;$ 
15       $\hat{x}_T[n] \leftarrow x_T^j[n];$ 
16      Break;
17    end
18  end
19 end

```

table to the astronomical dataset. Three different modulations: QPSK, QAM-16 and QAM-64 are used for RFI generation. The closest RFI source to the radio telescope is assumed to be same as the closest BS to the DSA-110 (see section V-B) radio telescope ~24 kilometers. FSPL and absorption loss are introduced to the RFI along with small scale fading effects based on channel models in [31]. Next, the RFI is subjected to varying sidelobe gains at the telescope antennas. The side-lobe attenuation depends on different telescope parameters. The value of side-lobe gain for DSA-110 can range between 35 dBi to $-\infty$ dBi, depending on whether the RFI is received through the main lobe or any null in the directivity of the dish. This attenuated RFI is filtered and channelized based on §V-B and quantized. Finally, it is added to the astronomical signal under consideration to generate the composite signal.

B. Capturing Astronomical Signals

The real astronomical datasets that have been utilized in this work, are collected with the Deep Synoptic Array (DSA-110), a radio interferometer made of the 110 4.65 m-antennas operating in the 1280-1530 MHz band, located at the Owens Valley Radio Observatory near Bishop, CA [32]. The DSA-110 operates a real-time data processing pipeline to detect fast astrophysical radio bursts (FRBs) in beamformed data, record raw baseband data associated with them, and produce correlation matrices to localize their origin. The signal from each antenna is first amplified and filtered, then digitized and channelized into 11.7 MHz-wide coarse channels each of which contains 384 equal width fine channels using a polyphase filterbank (PFB) [29]. The digitized coarse channels are then transferred to compute nodes for beamforming and searching these beams for FRBs using an incoherent de-dispersion search algorithm [33].

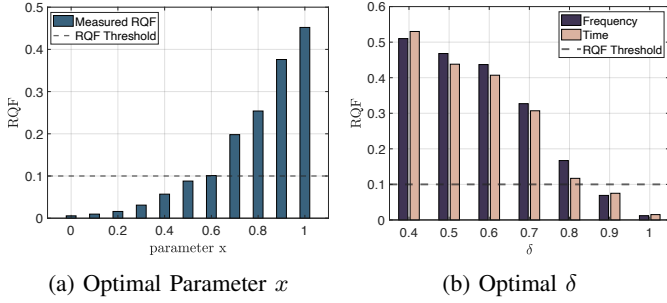


Fig. 5: Finding Optimal Design Parameters.

Dataset 1 is a 4.65-second long signal, focusing on the coarse channel containing the galactic H1 line (1420 MHz) [34], captured in April 2021 with 25 operational DSA-110 antennas. This is a beamformed full-precision dataset.

Dataset 2 is 7.2 seconds long data, containing the galactic H1 line (1420 MHz), captured using DSA-110 in February, 2023 with 63 operational DSA-110 antennas. This dataset includes signals from these individual antennas at a floating point precision of 4 bits. The precision of signals from individual antennas are inherent design specification of DSA-110 and is beyond our control.

VI. EVALUATION

A. Design Parameters

First, we show the impact of the power factor x in (11) to find optimal N_{Φ_R} . Figure 5a shows the quality of reconstruction considering one RFI source and 16-bit precision of eigenfunctions where x is varied from 0 to 1 at steps of 0.1. Based on the RQF threshold optimal $x \leq 0.6$. For, ease of implementation we choose $x = 0.6$.

Next we evaluate the impact of cosine similarity threshold δ as it directly governs both key contributions of this work - overhead reduction and accurate RFI cancellation. Figure 5b shows impact of changing δ across frequency (fine channels) and time (frames). We observe $\delta \geq 0.9$ is a reasonable choice for either case which is equivalent to angular deviation $\leq 23^\circ$.

B. Communication Overhead Analysis

We Consider a given LTE RFI source, fine channel bandwidth (f_t) of the telescope polyphase filter bank and signal duration of one LTE frame ($t_{frame} = 10$ ms) for characterization. For an RFI contaminating the astronomical signal of total bandwidth f_R and for duration t_R , the communication overhead without reduction will be:

$$w_{KLT} = 2 \times N_R \times L \times \frac{f_R}{f_t} \times \frac{t_R}{t_{frame}} \times B \text{ bits} \quad (15)$$

Where L is the KLT window length and B is the floating point precision. Additionally, two at the beginning is used as

the eigenfunctions are complex. Applying Algorithm 2 to the same configuration we find communication overhead:

$$w_{new} = \left(2 \times N_{\Phi_R} \times L \times \frac{f_R}{n_\delta^f \times f_t} \times B \right) + \left(N_{\Phi_R} \times \frac{f_R}{n_\delta^f \times f_t} \times \left(\frac{t_R}{t_{frame}} - n_\delta^t \right) \times B \right) \text{ bits} \quad (16)$$

where n_δ^f and n_δ^t indicate fine channels and frames across which the required eigenfunctions have $S_{GCS} \geq \delta$. The first term corresponds to those eigenfunctions that are fully shared with the telescope and the second term corresponds to those where only the deviation parameter Θ is shared. This formulation is shown for one RFI source only based on Algorithm 1. Plugging in parameters for a 20 MHz RFI signal with $\rho = 0.5$ we observe a 63% reduction in overhead.

C. Communication Overhead Comparison

Next, we compare the communication overhead required for RFI cancellation process with the KLT-based method in [2] and data-driven method in [19] and explore parameter space that impacts the overhead reduction. Overhead estimates are presented in bits with floating point precision of 16 bits for each method as a reasonable choice. Further quantization will yield poor cancellation and is out of scope of this work.

1) *Interference to Noise Ratio*: Interference to noise ratio (INR) is the ratio of incident RFI power and system noise at the telescope. Since, incident power dictates the number of eigenfunctions to be shared from an RFI source, with increasing INR, communication overhead is also increased. Figure 5a shows a case of change in communication overhead with INR for a single RFI source.

2) *Spectral Occupancy*: Spectral occupancy of RFI impacts the communication overhead as with changes in % spectral occupancy, the parameter S_{GCS} changes due to changing properties of the signal across fine channels and in time. We observe that with lower spectral occupancy, reduction in overhead is possible as shown in Figure 5b. It is important to note that although for both the proposed method and [19] overhead changes with spectral occupancy, the net reduction in overhead indicates the proposed method is superior.

3) *Multi-Source Collaboration*: Diverse geographical location of different RFI sources, orientation of telescope antennas, propagation condition and different spectral occupancy causes a significant variation in communication overhead from each RFI source. Figure 5b shows the cumulative overhead for upto five different RFI sources with INR - 20, 16, 12, 8 and 4 dB respectively, and spectral occupancy chosen randomly between 20% and 80%. We observe a *sublinear increase* in cumulative overhead for the proposed method while [19] and [2] have no scaling gain (linear) with the increasing number of sources.

D. Impact of Successive Multi-source Cancellation

The RFI cancellation and signal reconstruction procedure have a high computation cost ($\sim \mathcal{O}(n^2)$) which is reduced by

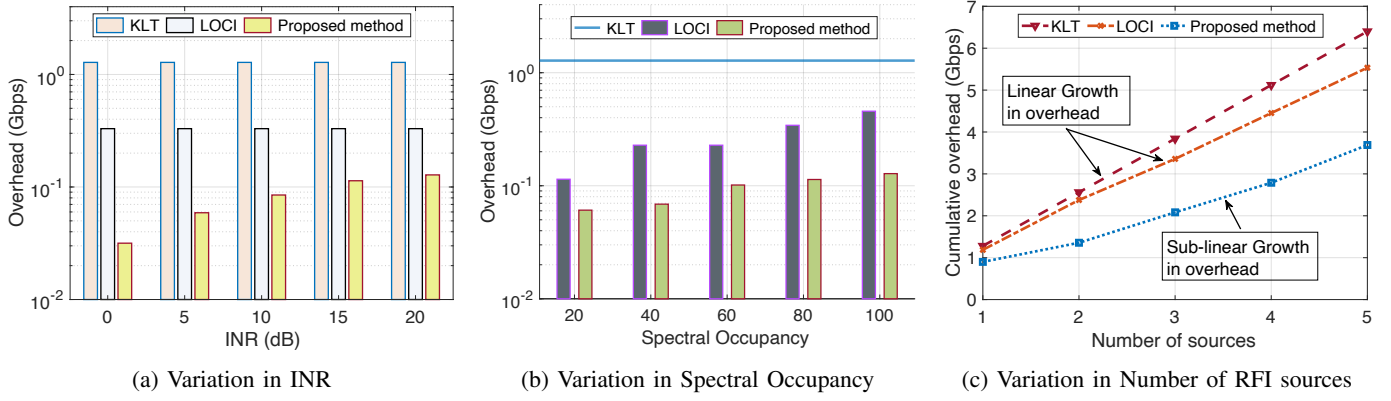


Fig. 6: Overhead Comparison of proposed method evaluated in the parameter space with following fixed parameters for each subfigure - a) 1 source and 80% occupancy, b) 1 source and 15 dB INR, c) 5 sources with varying INR and occupancy (see §VI-C).

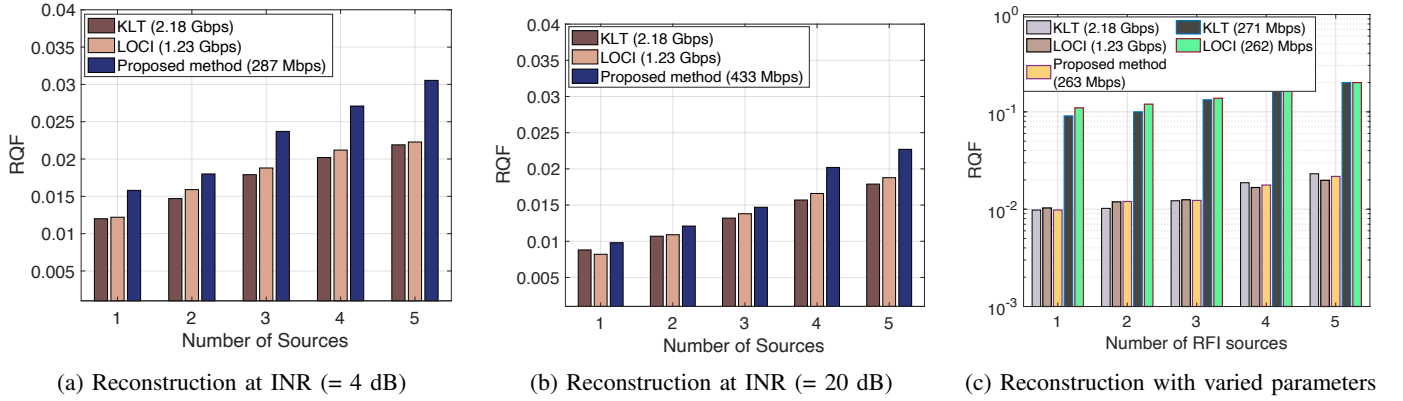


Fig. 7: Reconstruction Quality of proposed method with following fixed parameters for each subfigure - a) INR = 4 dB, b) INR = 20 dB, c) Varying INR and occupancy. For all cases in figures (a) and (b) $RQF < \psi'$. The corresponding communication overheads for each method are added in subfigure legends.

the implementation of multi-source successive RFI cancellation elaborated in Algorithm 2. We provide an experimental example to prove this claim in Table III. Three RFI sources are considered in three different scenarios where the relative power between the strongest source and the rest of the sources vary from < 5 dB to > 15 dB. We observe that when the power of RFI sources is comparable and at a much higher level relative to the system noise floor of the telescope, cancellation for all sources is necessary. However, when the relative power difference between different RFI is large, canceling the strongest or the strongest few sources can be sufficient to reach the sensitivity criterion of the telescope thus reducing the computation load as evident in Table III. Additionally, $N_{\Phi_R} \leq N_R$, further reducing the complexity. Hence without loss of generality, we can claim computation complexity at the telescope side will be $\mathcal{O}(n^d)$ where $1 < d < 2$. We cannot claim $d = 1$ as the complement projector in (4) is nonlinear. Compared to this, the data-driven method [19] has a complexity of $\mathcal{O}(n^2 + n)$. It is important to note that complexity reduction does not involve characterization of either RFI or the composite signal.

Relative Power (dB)	RQF after Successive RFI Cancellation		
	Strongest Source	Source 2	Source 3
≤ 5 dB	0.511	0.277	0.028
> 5 and ≤ 15	0.117	0.024	0.016
> 15	0.026	0.021	0.020

TABLE III: RQF after different cancellation stages in different relative power scenarios among RFI sources. Blue indicates RQF threshold is reached and red indicates it is not.

E. Reconstruction Accuracy

Figures 7a and 7b show the reconstruction accuracy of the proposed method with upto five RFI sources at INR 4 dB and 20 dB respectively. [2] has established that increasing INR can be beneficial in RFI cancellation as the features of RFI are more pronounced and cancellation accuracy is high. However, in the proposed method, reduction in Φ_R based on incident RFI power can lead to compromise in accuracy at low INR. At INR = 20 dB we see comparable RQF for the proposed method and the state-of-the-art. But, with decreasing INR, this effect becomes more prominent as shown in 7a. However, this is not detrimental to the cancellation process as the RQF remains below the threshold (0.1).

Next, we consider a similar setup of five RFI sources with INR as a varying parameter (20, 16, 12, 8 and 4 dB respectively), and spectral occupancy between 20% and 80%. We evaluate two sub-cases - 1) where the communication overhead of the proposed method is the same as state of the art - leading to an order of magnitude gain in RQF and 2) where the accuracy of the proposed method is comparable to the state of the art - which is achieved at only 37% of the overhead in [2]. The results are presented in Figure 7c.

VII. CONCLUSION

This work promotes collaboration between active and passive radio spectrum users in the current growing ubiquitous radio communication paradigm towards sustained coexistence. We have successfully presented a viable system design that exploits the unique benefits of time and frequency independent characterization of RFI at its source (high fidelity). We have also laid out an infrastructure to minimize communication overhead essential for the collaborative RFI cancellation and spectrum sharing regime. In future, we plan to explore other lower complexity decomposition techniques to improve computational burden of signal characterization. The high quality of astronomical signal recovery in our evaluations will serve as motivation to apply such collaborative methods for other forms of RFI as well.

ACKNOWLEDGEMENT

This work is funded by the National Science Foundation SWIFT Program (Award Number - 2128581).

REFERENCES

- [1] Federal Communication Commission, "2020 FCC Broadband Deployment Plan," 2020 *BROADBAND DEPLOYMENT REPORT*, vol. 21, no. 1, pp. 1–9, 2020. [Online]. Available: <http://mpoc.org.my/malaysian-palm-oil-industry/>
- [2] S. Chakraborty, G. Hellbourg, M. Careem, D. Saha, and A. Dutta, "Collaboration with cellular networks for rfi cancellation at radio telescope," *IEEE Transactions on Cognitive Communications and Networking*, vol. 9, no. 3, pp. 765–778, 2023.
- [3] T. Hastie, J. Friedman, and R. Tibshirani, *Basis Expansions and Regularization*. New York, NY: Springer New York, 2001, pp. 115–163. [Online]. Available: https://doi.org/10.1007/978-0-387-21606-5_5
- [4] A. Szumski and G. Hein, "Finding the interference karhunen-loève transform as an instrument to detect weak rf signals," in *InsideGNSS*, 2011, p. 57–64.
- [5] N. I. Miridakis and D. D. Vergados, "A survey on the successive interference cancellation performance for single-antenna and multiple-antenna ofdm systems," *IEEE Communications Surveys Tutorials*, vol. 15, no. 1, pp. 312–335, 2013.
- [6] M. Amjad, F. Akhtar, M. H. Rehmani, M. Reisslein, and T. Umer, "Full-duplex communication in cognitive radio networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 19, no. 4, pp. 2158–2191, 2017.
- [7] S. Sagari, S. Baysting, D. Saha, I. Seskar, W. Trappe, and D. Raychaudhuri, "Coordinated dynamic spectrum management of lte-u and wi-fi networks," in *2015 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2015, pp. 209–220.
- [8] G. Ding, Y. Jiao, J. Wang, Y. Zou, Q. Wu, Y. Yao, and L. Hanzo, "Spectrum inference in cognitive radio networks: Algorithms and applications," *IEEE Communications Surveys Tutorials*, vol. 20, no. 1, pp. 150–182, 2018.
- [9] A. Ali and W. Hamouda, "Advances on spectrum sensing for cognitive radio networks: Theory and applications," *IEEE Communications Surveys Tutorials*, vol. 19, no. 2, pp. 1277–1304, 2017.
- [10] M. R. Souryal and T. T. Nguyen, "Effect of federal incumbent activity on cbrs commercial service," in *2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2019, pp. 1–5.
- [11] A. Soliman, S. Weinreb, G. Rajagopalan, C. Eckert, and L. Hilliard, "Quadruple-ridged flared horn feed with internal rfi band rejection filter," in *2016 Radio Frequency Interference (RFI)*, 2016, pp. 115–116.
- [12] P. Bolli and F. Huang, "Superconducting filter for radio astronomy using interdigitated, capacitively loaded spirals," *Experimental Astronomy*, vol. 33, no. 1, pp. 225–236, 2012.
- [13] R. Urvashi, A. P. Rao, and N. Pune, "Automatic rfi identification and flagging," *The National Centre for Radio Astrophysics of the Tata Institute of Fundamental Research*, 2003.
- [14] A. Offringa, R. Wayth, N. Hurley-Walker, D. Kaplan, N. Barry, A. Beardsley, M. Bell, G. Bernardi, J. Bowman, F. Briggs *et al.*, "The low-frequency environment of the murchison widefield array: radio-frequency interference analysis and mitigation," *Publications of the Astronomical Society of Australia*, vol. 32, 2015.
- [15] A. Sclocco, D. Vohl, and R. V. van Nieuwpoort, "Real-time rfi mitigation for the aperitif radio transient system," in *2019 RFI Workshop-Coexisting with Radio Frequency Interference (RFI)*. IEEE, 2019, pp. 1–8.
- [16] C. J. Wolfaardt, "Machine learning approach to radio frequency interference (rfi) classification in radio astronomy," Ph.D. dissertation, Stellenbosch: Stellenbosch University, 2016.
- [17] G. Hellbourg, K. Bannister, and A. HotarP, "Spatial filtering experiment with the askap beta array," in *2016 Radio Frequency Interference (RFI)*. IEEE, 2016, pp. 37–42.
- [18] S. W. Ellingson, J. D. Bunton, and J. F. Bell, "Removal of the glonass c/a signal from oh spectral line observations using a parametric modeling technique," *The Astrophysical Journal Supplement Series*, vol. 135, no. 1, p. 87, 2001.
- [19] S. Chakraborty, D. Saha, A. Dutta, and G. Hellbourg, "Loc: Learning low overhead collaborative interference cancellation for radio astronomy," in *ICC 2023 - IEEE International Conference on Communications*, 2023, pp. 6391–6396.
- [20] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, 1952.
- [21] A. Robinson and C. Cherry, "Results of a prototype television bandwidth compression scheme," *Proceedings of the IEEE*, vol. 55, no. 3, pp. 356–364, 1967.
- [22] G. Schuller, B. Yu, D. Huang, and B. Edler, "Perceptual audio coding using adaptive pre- and post-filters and lossless compression," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 6, pp. 379–390, 2002.
- [23] X. Wu and N. Memon, "Context-based, adaptive, lossless image coding," *IEEE Transactions on Communications*, vol. 45, no. 4, pp. 437–444, 1997.
- [24] G. Wallace, "The jpeg still picture compression standard," *IEEE Transactions on Consumer Electronics*, vol. 38, no. 1, pp. xviii–xxxiv, 1992.
- [25] V. K. Goyal, A. K. Fletcher, and S. Rangan, "Compressive sampling and lossy compression," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 48–56, 2008.
- [26] Q. Yang, M. B. Mashhadi, and D. Gündüz, "Deep convolutional compression for massive mimo csi feedback," in *2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2019, pp. 1–6.
- [27] S. Lee, J.-B. Jeong, and E.-S. Ryu, "Entropy-constrained implicit neural representations for deep image compression," *IEEE Signal Processing Letters*, vol. 30, pp. 663–667, 2023.
- [28] T. L. Wilson, K. Rohlf, and S. Hüttemeister, *Tools of radio astronomy*. Springer, 2009, vol. 5.
- [29] D. C. Price, "Spectrometers and polyphase filterbanks in radio astronomy," in *The WSPC Handbook of Astronomical Instrumentation: Volume 1: Radio Astronomical Instrumentation*. World Scientific, 2021, pp. 159–179.
- [30] R. I.-R. RA.769-2, "Protection criteria used for radio astronomical measurements."
- [31] 3GPP, "Study on channel model for frequencies from 0.5 to 100 GHz," 3rd Generation Partnership Project (3GPP), TR 38.901.
- [32] [Online]. Available: <https://www.ovro.caltech.edu/>
- [33] G. Walsh and R. Lynch, "Searching for single pulses using heimdall," in *American Astronomical Society Meeting Abstracts# 231*, vol. 231, 2018, pp. 243–04.
- [34] J. M. Dickey and F. J. Lockman, "Hi in the galaxy," *Annual review of astronomy and astrophysics*, vol. 28, pp. 215–261, 1990.