Spectrum Sharing via Collaborative RFI Cancellation for Radio Astronomy

Maqsood Careem, Shuvam Chakaraborty, Aveek Dutta, Dola Saha
Department of Electrical and Computer Engineering
University at Albany SUNY
{mabdulcareem, schakraborty, adutta, dsaha}@albany.edu

Gregory Hellbourg
Department of Astronomy
California Institute of Technology
{ghellbourg}@astro.caltech.edu

Abstract—Radio frequency Interference (RFI) from cellular and other communication networks is commonly mitigated at the Radio Telescope without any active collaboration with the interfering sources. The expanding Universe and simultaneous proliferation of Earth based and LEO communication infrastructure is causing unprecedented RFI that require collaborative strategies to maintain the scientific and societal goal of each. In this work, we provide a method of signal characterization and its use in subsequent cancellation, that uses Eigenspaces derived from the astronomical and the RFI signals. This is different from conventional time-frequency domain analysis, which is limited to fixed characterizations (*e.g.*, complex exponential in Fourier methods) that cannot adapt to the changing statistics (*e.g.*, auto-correlation) of the RFI, typically observed from communication systems. Through our analysis and simulation using real-world astronomical signals, we are able to remove RFI from cellular networks by 89.04%, which reduces excision at the Telescope.

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

I. INTRODUCTION

The proliferation of next generation (xG) communication networks [1], [2] increasingly generate unwanted Radio Frequency Interference (RFI), even in bands that are protected for radio astronomy due to out-of-band emissions and intermodulation products. At the same time, technological advances such as wideband and low system temperature receivers are allowing astronomers to observe astronomical emissions over even wider bands. This necessitates stringent mitigation techniques to continue radio astronomy research in presence of challenging RFI. Generally, communication system designers strive to reduce noise from artificially generated signal, whereas radio astronomy focuses on removing communication signals from the astronomical signal. This seemingly opposing requirement is pushing the two communities farther away. Both are equally essential and are designed to overcome a common bottleneck: *RFI*. So, we present a collaborative framework to address this problem, by aggregating concise, yet accurate signal characterization from the RFI source, i.e., the communication networks and *cancel* it at the telescope.

Although radio telescopes are generally located in geographically isolated areas farther from communication networks, it does not prevent RFI from nearby cellular networks. Current RFI mitigation techniques use statistical signal analysis to detect RFI and discard the associated time and frequency bins from the collected data. The excision of corrupted data may at best reduce the sensitivity of the telescope, and at

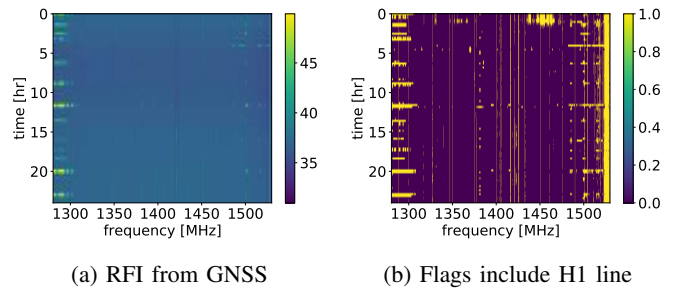(a) RFI from GNSS   (b) Flags include H1 line

Fig. 1: RFI flagging and excision with data from OVRO [3]

worse remove the astronomical signal of interest. In this work, we focus on the sub-6GHz bands due to growing interest in observing the lower frequencies by the radio astronomy community to understand deep space objects. As a source of RFI, we have identified 4G/5G cellular service, which transmits at a high power and can propagate large distances. However, the proposed method is applicable to remove a variety of RFI in other frequency bands as well.

Impact of RFI in Radio Astronomy: Probing the Universe at radio frequencies opens an unexplored window to study matter and energy under extreme conditions that cannot be achieved on Earth. For example, the omnipresent neutral atomic Hydrogen (HI) line, with rest frequency at 1420 MHz, has been extensively used to study the structure of galaxies and the intergalactic medium. As the universe is expanding, more distant objects appear to be moving away from us with increasingly high velocities, a phenomenon termed "redshift", which shifts all spectral lines from their rest frequencies to lower frequencies. The highly redshifted HI line (observed at frequencies below 200 MHz) is also the unique probe available to study the distribution of matter in the early universe and to understand its epoch of reionization. Highly redshifted sources require astronomers to increasingly observe outside the protected bands most of the time [4], [5], [6]. Moreover, astronomical emissions are extremely weak due to the large distances travelled before being detected by highly sensitive radio telescopes. Often, detection requires long integration at ≈40dB below the telescope noise. This is in stark contrast to reception of wireless communications at ≈20dB above the receiver noise level. Hence, these receivers are extremely sensitive to RFI which may lead to false detection of the signal of interest or its masking, lowering the overall sensitivity of the telescope due to excision. Figure 1 illustrates the excision

(a) Galactic H1 line @ 1420 MHz   (b) LTE signal at BS   (c) Signals in time-domain   (d) Composite signal spectrum
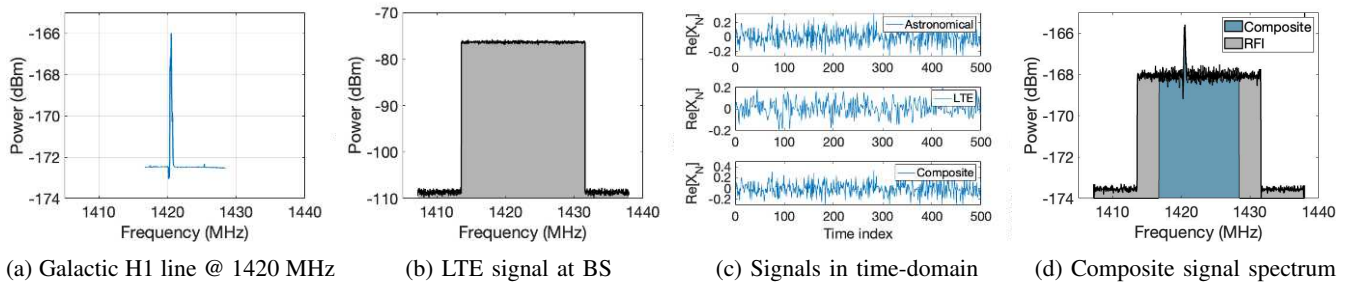
Fig. 2: Spectral and temporal characteristics of Astronomical and RFI signals

problem with data collected with the Deep Synoptic Array DSA-110 [7], [3] located at the Owens Valley Radio Observatory (OVRO), CA, USA (see §II-A). The data in Figure 1a spanning 1280-1350 MHz is corrupted with narrowband, wideband, and impulsive RFI. Figure 1b shows the same data after identification and excision of the RFI-corrupted time-frequency bins. This flagging and excision approach is usually tuned to minimize the probability of non-detection of the RFI resulting in a significant data loss.

Therefore, even with state of the art methods in RFI mitigation, full recovery of an astronomical signal corrupted by RFI cannot be achieved without prior knowledge of the source of 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. This hypothesis has been tested through our analysis and experiments in this paper.

II. MODELS AND PRELIMINARIES

A. Acquisition and processing of astronomical signals

A radio telescope achieves its high sensitivity by maximizing its directivity, collecting areas, and minimizing the system temperature of its receivers. It can vary from large single dish antennas equipped with single or multiple beam receivers, to large arrays of antennas that are either phased together to produce multiple beams in the sky or to perform interferometric synthesis imaging [8], [9]. After signal conditioning (i.e., equalization and filtering), the output of the individual receivers are digitized over hundreds of MHz and channelized into smaller frequency bins of hundreds of kHz width [10]. Channelization is useful for reducing the data rate for real-time processing, share computational resources and excise RFI-corrupted channels before further processing. Subsequent data processing are specific to the observed astronomical object and may include real-time matched-filtered transient searches, spectral integration or data correlation for synthesis imaging.

The OVRO DSA-110 Radio Telescope: The 4s-long data set, utilized throughout in this paper, has been collected with the Deep Synoptic Array being deployed at the OVRO. Figure 2a shows the H1 line at 1420 MHz obtained from the dataset, while the signal spans the telescope bandwidth (BW). At completion, this array will have 110 antennas of 4m-diameter.

At the time of data collection (15 May 2021), only 25 antennas were deployed. The telescope digital receivers channelize the data into 11.7 MHz coarse channels, then further channelizes each of the coarse channels into 30.5 kHz fine channels. After coherently summing the outputs of all antennas, the resulting data are searched in real-time for Fast Radio Bursts (FRB) [11] to produce an image of the field of view of the telescope via synthesis imaging to locate the origin of the FRB .

B. Signal model for astronomical signals

The digitized and channelized output of a single telescope antenna (either single dish or an element of an array) is expressed as (1):

$$x_{f_c}[n] = x_A[n] + x_N[n] + x_R[n] \approx x_N[n] + x_R[n] \quad (1)$$

where $x_{f_c}[n]$ is the channelized baseband signal at a filterbank channel centered around frequency f_c and at time sample n, and follows a stationary (over a short duration) stochastic process, is independently and identically distributed (i.i.d.) with $x_{f_c}[n] \sim \mathcal{NC}(x_R[n], \sigma^2)$. $\mathcal{NC}(\mu, \Gamma)$ 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 with power $\sigma_R^2 = N^{-1} \sum_N |x_R[n]|^2$. However, accurate characterization of $x_{f_c}[n]$ and $x_R[n]$ is non-trivial and is discussed in §III.

C. Signal model for LTE RFI signal

Long term evolution (LTE) 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 (base station (BS) to user equipment) using a carrier f_c is shown in (2).

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

The signal bandwidth depends on the size of the FFT, number of subcarriers, guard bands, etc. as in Table I. An example realization of the LTE signals used as RFI in this

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*

* shows the parameters used to simulate the RFI in this work.

TABLE I: Different LTE parameters produce unique RFI

work is shown in figure 2b, which typically occupies a wider band than the telescope bandwidth. In reality, if the telescope acquires an astronomical signal at or around f_c, this RFI will undergo the same processing as mentioned in §II and will be present in varying strength across multiple channels depending on the telescope aperture, sidelobe gain and the spectral occupancy of the LTE signal itself. We refer to this signal as the *composite signal*, which is defined by (1) and (2) and are shown in figures 2c and 2d in time and frequency domain respectively. Typically, the LTE signal is grouped as radio frames in baseband with each frame containing 10 subframes and each consisting of two slots. The resource block is the smallest unit of an LTE frame allocated to a user. Each resource block consists of 12 subcarriers lasting for a duration of 7 symbols. This arrangement along with the parameters in Table I lend unique characteristics to the RFI. However, propagation over large distances and multipath reflections deteriorates the features of the RFI in time and frequency, hence the need for robust stochastic signal decomposition.

III. COLLABORATIVE RFI CANCELLATION

The literature on cellular RFI mitigation through cancellation in radio astronomy primarily focuses on extracting signal features at the radio telescope via local sensing, without any prior knowledge of the source of the RFI (see §VI for literature review). As a result, the signal information is extremely limited due to equipment (front-end, BW, gain, etc.), signal deterioration due to propagation and multipath, and lack of coherence between the RFI source and the telescope. Another limitation stems from the very method employed to characterize the RFI, which are almost always limited to Fourier methods for frequency domain analysis and temporal statistics like autocovariance, cyclostationarity, density functions and other lower order statistics. All of these methods are sensitive to time-varying RFI from cellular networks and require long observation times to accumulate a steady-state model. Furthermore, cancellation often requires local synthesis (requires sharing of user data) of the RFI signal from the acquired characteristics to either employ time-domain nulling (subtraction) or frequency domain filtering. This has the risk of eliminating the astronomical signal of interest and require high degree of synchrony between the telescope and RFI sources for phase coherent cancellation.

Our method is radically different from the state of the art in three aspects: 1) The RFI is decomposed at the cellular BS into a compact yet accurate eigenspace that is periodically
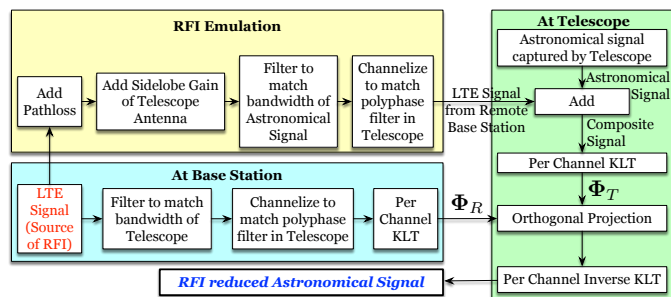

Fig. 3: Block diagram of the RFI cancellation system

shared with the telescope via a shared channel over the Internet. Unlike Fourier methods, which decomposes using complex exponential bases only, our method extracts the bases from the signal itself, which adapts with time-varying cellular RFI. Furthermore, the fidelity of the decomposition is vastly improved at high signal power, which is maximum at the BS. Figure 6 in §V-B discusses this unique feature; 2) At the telescope the composite signal is decomposed using the same method revealing its eigenspace that contain the RFI subspace, ideally orthogonal to the astronomical signal space; and 3) The shared RFI eigenspace is used to cancel the RFI from the composite eigenspace via complimentary orthogonal projections. 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. Figure 3 shows the components and flow of collaborative information between the telescope and the cellular network. For the purposes of this work, the RFI is emulated as a LTE downlink signal and combined with the real astronomical signal as described in §II[1]. Our analysis and simulation is based on this composite signal but adheres to all the channelization and bandpass filtering as employed by the DSA-110 telescope. At the BS, the eigenspace for the LTE RFI signal is shown as Φ_R in figure 3 and that of the composite signal at the telescope is denoted by Φ_T, which are formally defined in §III-A. These two are the key parameters required at the cancellation step along other topological information given in Table II. Each of the parameters have specific roles

Static Parameters	Dynamic Parameters
Observed frequency range [†]	Eigenspace (Φ_R)
Polyphase filter subchannel [†]	KLT window (L)

[†] parameters shared by telescope only. Others are cellular network only.

TABLE II: Shared parameters for collaborative cancellation

in the cancellation apparatus and are explained in subsequent sections where applicable. The static parameters are constant and can be made available via a database lookup. The dynamic parameters change over time and require periodic sharing. The largest update in terms of size is Eigenspace information, Φ_R. For the KLT window length = L and number of time samples = N, each eigen function update for each subchannel of the polyphase filterbank will be $4 \times L \times L$ bytes of data. In our

---

[1]In this work, we consider 1 cellular BS and 1 radio telescope as the foundation for more complex topologies, to be addressed in future work.

1. Input sample sequence
$$\mathbf{x} = [x_0, x_1, x_3, \ldots, x_N]^T$$

2. (L x K = N-L+1) Hankel matrix

$$\mathbf{U} = \begin{bmatrix} x_0 & x_1 & x_2 & \ldots & x_K \\ x_1 & x_2 & & & \vdots \\ x_2 & & & & \vdots \\ \vdots & & & & x_{K+L-3} \\ \vdots & & x_{K+L-3} & x_{K+L-2} \\ x_L & \ldots & \ldots & x_{K+L-2} & x_{K+L-1} \end{bmatrix}$$

3. Covariance matrix
$$\mathbf{R}_{xx} = \mathbb{E}[\mathbf{U}\mathbf{U}^H]$$

4. Eigen-decomposition
$$\mathbf{R}_{xx} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^H$$
$$\Lambda = \text{diag}\{\lambda_0, \lambda_1, \ldots, \lambda_M\}$$

5. SSA implementation of KLT
$$\mathbf{z}_i = \mathbf{\Phi}^H \mathbf{x}_i$$

Time series **x** projected on to orthogonal Eigenfunctions

6. Inverse- SSA (Inverse-KLT)
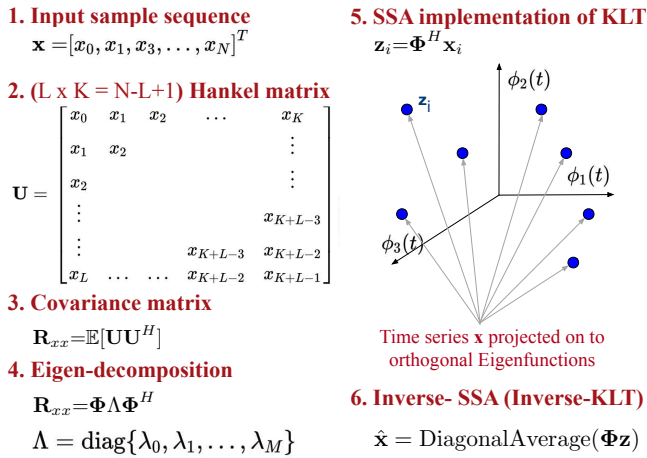$$\hat{\mathbf{x}} = \text{DiagonalAverage}(\mathbf{\Phi}\mathbf{z})$$

Fig. 4: Geometry of KLT based signal characterization

empirical evaluations, this amounts to 1 Mbytes of data per subchannel. Since, this information is shared over the Internet, it does not impose any overhead on the cellular infrastructure. However, it is possible to use the same $\mathbf{\Phi}_R$ for all the channels if the RFI bandwidth is larger than that of the telescope since the RFI is equally incident across the entire telescope bandwidth. It is also possible to limit the frequency of sharing by replicating the averaging window of the telescope at the BS to smooth any temporal variation of $\mathbf{\Phi}_R$. These techniques require further theoretical and experimental evaluation that are currently not possible due to lack of appropriate dataset and experimental facility, and are left as future work.

A. Eigenspace representation of RFI and composite signals

The Karhunen–Loève Transform (KLT) [12], [13], [14] decomposes any stochastic signal like RFI, and has the following advantages: (a) it is suitable for both narrow and wideband signals, (b) applicable to any type of basis and any type of signal (deterministic or stochastic), unlike other transforms like DFT that are limited to sinusoidal basis, (c) able to detect weak signals below the noise floor, and (d) optimal in the minimum mean square error sense. For analysis of sampled signal, KLT is often implemented using Singular Spectrum Analysis (SSA) that relies on the decomposition of the autocorrelation matrix, $\mathbf{R}_{xx}(n_1, n_2)$, which captures temporal correlations of a random signal $x[n]$, into its constituent eigenfunctions and corresponding eigenvalues. Let, $x[n] = x_T[n]$ at the telescope and $x[n] = x_R[n]$ at the BS, and n_1, n_2 are different time indices. $\mathbf{R}_{xx}(n_1, n_2)$ is calculated as in (3) by embedding the signal $x[n]$ in a Hankel matrix, $\mathbf{U}$, of size $L \times K$ [15], [16] as shown in figure 4, where N is the number of time samples in the signal of interest $x[n]$, L is the KLT window size, empirically determined based on the astronomical signal being observed and telescope parameters. Its analytical determination is out of scope of this work.

$$\mathbf{R}_{xx} = \mathbb{E}[\mathbf{U}\mathbf{U}^H], \quad \text{where,} \quad \mathbf{U} = [\mathbf{x}_1, \ldots, \mathbf{x}_K] \quad (3)$$

where $\mathbf{x}_i = [x[n], \ldots, x[n+L-1]]^T$ are lagged vectors of size L, with $i \in [1, K]$, and $\mathbb{E}[.]$ is the expectation operator. Then,
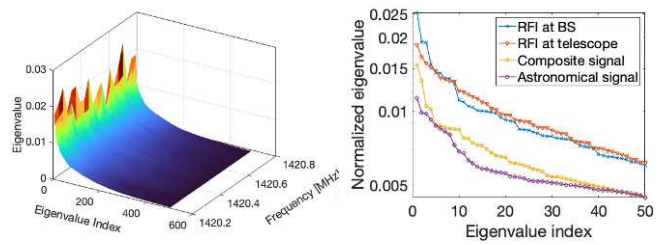
(a) Composite signal at telescope   (b) Eigenvalue comparison

Fig. 5: Derived eigenspectrum from signal decomposition.

the KLT decomposition is obtained by solving the eigenvalue problem in (4),

$$\mathbf{R}_{xx} = \mathbf{\Phi}\mathbf{\Lambda}\mathbf{\Phi}^H, \quad \text{where,} \quad \Lambda = \text{diag}\{\lambda_1, \ldots, \lambda_L\} \quad (4)$$

where λ_j are the eigenvalues, with $j \in [1, L]$, and $\mathbf{\Phi}$ is a unitary matrix containing L eigenvectors as its columns. Since, (4) decomposes the temporal correlations of $x[n]$, each column of $\mathbf{\Phi}$, i.e. ϕ_i, is a time-series, and consequently is referred to as an eigenfunction. Geometrically, the eigenfunctions give an orthogonal set of "directions" (or spatial signatures) present in the autocorrelation matrix, which span the Hilbert space containing the KLT projected time samples, while the eigenvalues represent the power of the signal coming from the corresponding directions, sorted in decreasing order. Consequently, the KLT automatically adapts to the shape of the (signal+noise) irrespective of its behavior in time, by adopting a new reference frame spanned by the eigenfunctions, which makes it appropriate to characterize and subsequently remove RFI. These eigenfunctions project the signal on to the Hilbert space given by, $\mathbf{z}_i = \mathbf{\Phi}^H \mathbf{x}_i$, where $\mathbf{z}_i$ are the columns of the matrix $\mathbf{z}$ and are orthogonal temporal principal components of the input $x[n]$ containing L samples as shown in figure 4. Consequently, $\mathbf{z}$ is given by, $\mathbf{z} = \mathbf{\Phi}\mathbf{U}$ using the definition of $\mathbf{U}$ in (3). Projecting $\mathbf{z}$ onto the $M < L$ eigenfunctions of $\mathbf{\Phi}$ with largest eigenvalues, reconstructs the signal, $\hat{x}[n]$ with minimum noise [17] as shown in figure 4. Depending on where the signal is characterized, at the telescope (includes added RFI and noise) or at the BS (RFI and noise only), the subscripts T and R are appended to the above variables, and the eigenfunctions are henceforth referred to as *Astronomical Kernel* ($\mathbf{\Phi}_T$) and *RFI Kernel* ($\mathbf{\Phi}_R$) respectively. $\mathbf{\Phi}_R$ provides an accurate and compact characterization of the RFI at the BS that is shared with the telescope for RFI cancellation as shown in figure 3. The eigenspectra of the decomposition of the composite signal at the telescope is shown in figure 5a. Figure 5b compares the eigenspectrum for the RFI at the BS, attenuated incident RFI at the telescope and the composite signal. The number of dominant eigenvalues in each plot shows the dimensionality of the associated signal subspace.

B. RFI cancellation with eigenspace projection

The KLT provides L eigenfunctions at the telescope and the BS. Mathematically, RFI in the composite signal is cancelled by projecting the eigenfunctions at the telescope ($\mathbf{\Phi}_T$) onto a

subspace that is *orthogonal* to the subspace spanned by the RFI eigenfunctions ($\mathbf{\Phi}_R$).

1) Orthogonal complement projector: The first step in RFI cancellation is the computation of the orthogonal complement projector using the eigenfunctions at the BS as in (5),

$$\mathbf{P}_{\mathbf{\Phi}_R}^{\perp} = \mathbf{I} - \mathbf{\Phi}_R \left(\mathbf{\Phi}_R^H \mathbf{\Phi}_R \right)^{-1} \mathbf{\Phi}_R^H \qquad (5)$$

where $\mathbf{I}$ is the $L \times L$ identity matrix. $\mathbf{P}_{\mathbf{\Phi}_R}^{\perp}$ is such that $\mathbf{P}_{\mathbf{\Phi}_R}^{\perp} \mathbf{\Phi}_R = 0$ and by extension we can show that, $\mathbf{P}_{\mathbf{\Phi}_R}^{\perp} x_R[n] = 0$ using the definitions in (4) and (3). Therefore, applying this orthogonal projection $\mathbf{P}_{\mathbf{\Phi}_R}^{\perp}$ to the telescope signal has the effect of nulling the RFI component. The unique advantage of collaborative RFI mitigation is that the precision of the estimation of the RFI subspace, $\mathbf{\Phi}_R$, is improved by computing it at the BS where the RFI is received at high SNR, and consequently $\mathbf{P}_{\mathbf{\Phi}_R}^{\perp}$ can be calculated even if the astronomical and RFI signals are not separable at the telescope. Unlike the literature on signal separation or subspace estimation, which typically rely on the assumption of strong or weak signal separability or the existence of orthogonal subspaces of components in composite signals [18], this procedure does not require such assumptions, as the RFI subspace is accurately identified at high SNR at the BS.

2) RFI cancellation: The projection of the eigenfunctions at the telescope using the orthogonal complement projector in (5) is given by (6),

$$\widehat{\mathbf{\Phi}}_T = \mathbf{P}_{\mathbf{\Phi}_R}^{\perp} \mathbf{\Phi}_T \qquad (6)$$

This projects the composite signal subspace at the telescope to the null-space of the RFI. Consequently, $\widehat{\mathbf{\Phi}}_T$ spans the subspace that is orthogonal to the RFI sub-space spanned by $\mathbf{\Phi}_T$. This allows for subspace-based removal of *undesired* eigenfunctions corresponding to any RFI.

Finally, the inverse-KLT is used to reconstruct the RFI-free astronomical signal, i.e., $\hat{x}_T[n]$, which involves two steps. First, the Hankel matrix corresponding to the RFI-free astronomical signal, $\widehat{\mathbf{U}}_T$, is reconstructed by projecting the matrix $\mathbf{z}_T$ onto the projected eigenfunctions $\widehat{\mathbf{\Phi}}_T$ as given by (7).

$$\widehat{\mathbf{U}}_T = \widehat{\mathbf{\Phi}}_T \mathbf{z}_T, \quad \text{where,} \quad \mathbf{z}_T = \mathbf{\Phi}_T^H \mathbf{U}_T \qquad (7)$$

This cancellation method does not require the orthogonality of the astronomical and RFI subspaces in the composite signal at the telescope, since the projection in (7) ensures that any RFI is nulled using the precise estimate of the RFI subspace at the BS. Finally, the cross-diagonal elements of the reconstructed Hankel matrix are averaged [18] using (8), to reconstruct the space-signal time-series, $\hat{x}_T[n]$ from $\widehat{\mathbf{U}}_T$.

$$\hat{x}_T[n] = \begin{cases} \frac{1}{n} \sum_{k=1}^{n} \widehat{\mathbf{U}}_T^{(k,n\text{-}k+1)} & \text{for } 1 \leqslant n < L \\ \frac{1}{L} \sum_{k=1}^{L} \widehat{\mathbf{U}}_T^{(k,n\text{-}k+1)} & \text{for } L \leqslant n \leqslant K \\ \frac{1}{N-n+1} \sum_{k=n\text{-}K+1}^{L} \widehat{\mathbf{U}}_T^{(k,n\text{-}k+1)} & \text{for } K+1 \leqslant n \leqslant N \end{cases} \qquad (8)$$

where the superscript $(m, k\text{-}m\text{+}1)$ indicates the corresponding

element in the matrix $\widehat{\mathbf{U}}_T$. This form of real-time recovery of $\hat{x}_T[n]$ is not possible currently. Therefore, successful deployment of this method at telescope sites, like the DSA-110, will greatly reduce excision and maximize its sensitivity. In order to experimentally evaluate the quality of reconstruction of $\hat{x}_T[n]$, we define a metric that compares the residual interference to the RFI-free astronomical signal.

IV. RQF: A METRIC FOR EVALUATION

We define the Reconstruction Quality Factor (RQF) as the ratio of the power of the RFI-free and the residual RFI after cancellation. This quantity is evaluated empirically and may only be employed under a simulated and controlled framework. The RFI-free signal at the telescope can be represented by (9),

$$\hat{x}_T[n] = x_T[n] - \hat{x}_R[n] \qquad (9)$$

where $\hat{x}_R[n] = x_R[n] + \epsilon_r[n]$ is the estimated RFI contribution, and $\epsilon_r[n] \sim \mathcal{NC}(0, \sigma_{\text{est}}^2)$ captures the cumulative estimation and reconstruction error. Therefore, the RQF over N samples is given by (10),

$$\text{RQF} = \frac{\|x_N\|^2}{\|x_T - \hat{x}_T\|^2} = \frac{\|x_N\|^2}{\|\epsilon_r\|^2} \qquad (10)$$

where $\|x\|^2 = N^{-1} \sum_N |x[n]|^2$. Applying a second order Taylor approximation to the expectation of a ratio of random variables, we get,

$$\mathbb{E}\{\text{RQF}\} = \frac{\mathbb{E}\{\|x_N\|^2\}}{\mathbb{E}\{\|\epsilon_r\|^2\}} - \frac{\text{cov}(\|x_N\|^2, \|\epsilon_R\|^2)}{\mathbb{E}\{\|\epsilon_r\|^2\}^2} + \frac{\text{var}(\|\epsilon_R\|^2) \cdot \mathbb{E}\{\|x_N\|^2\}}{\|\epsilon_R\|^3} \qquad (11)$$

Assuming the independence between time samples, and between both variables $\|x_N\|^2$ and $\|\epsilon_R\|^2$, we have,

$$\mathbb{E}\{\text{RQF}\} = \frac{\sigma_N^2}{\sigma_{\text{est}}^2} \left(1 + 2\frac{N-1}{N^2} \right) \qquad (12)$$

(using here $\text{var}((N-1)^{-1} \sum_N |\epsilon_r|^2) = \frac{2\sigma_{\text{est}}^4}{N-1}$).

An RFI contribution is considered detrimental to an astronomical observation when its power reaches 10% of the telescope system noise power [19]. This level fixes a lower-bound to the RQF where $\sigma_{\text{est}}^2 = \sigma_N^2 / 10$,

$$\text{RQF}_{\text{ref}} = 10 \times \left(1 + 2\frac{N-1}{N^2} \right) \qquad (13)$$

In our evaluations, the RQF in (12) is used as a metric to measure the combined accuracy of the KLT decomposition and the eigenspace based cancellation of the RFI.

V. EVALUATION AND RESULTS

A. Experimental Setup

1) LTE RFI generation: The LTE RFI signal is generated according to the parameters in Table I containing 400 frames (10 milliseconds each) to be comparable in length to the astronomical dataset. Three different modulations: QPSK, QAM-16 and QAM-64 are randomly chosen for each resource block
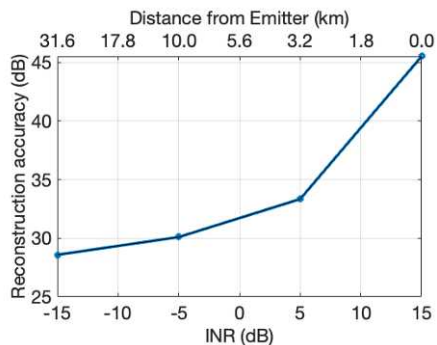
Fig. 6: Reconstruction loss of RFI with varying INR.

Fig. 7: Reconstructed space signal compared to the true astronomical and the composite signal.

data generation. The signal contains the primary (PSS) and secondary (SSS) frame synchronization signals only for simplicity. The LTE signal is then filtered to match the bandwidth of the telescope and attenuated by free space pathloss and side-lobe gains (RFI is primarily acquired via telescope side-lobes, ≈ -40 dBi for OVRO DSA-110). We also assume that the RFI signal is attenuated to a power level such that it is comparable to the astronomical signal as shown in Figure 2d. This requires an aggregate attenuation of ≈ 100 dB, corresponding to 10 km distance to match the pathloss component. The LTE signal is channelized both at the BS and the telescope to match the channels of the telescope, which is a static parameter shared by the telescope as in Table II. The attenuated signal is added to the astronomical signal to simulate the RFI contaminated composite signal. We assume that the collaboration between BS and telescope will start at the same time using standardized GPS clocks at both locations. Our method does not require explicit synchronization and is robust to GPS clock errors.

2) Signal decomposition and reconstruction: The LTE signal at BS and the composite signal at the telescope are decomposed to obtain the eigenfunctions based on the formulation in §III. Theoretically, the KLT produces an infinite dimensional Hilbert space, i.e, infinite number of eigenfunctions. In our SSA based implementation, we limit the KLT window length to L=500 for the Hankel matrix and approximate the eigenspace to contain eigenfunctions whose eigenvalues are within 1% of the maximum eigenvalue. Figure 6 shows the reconstruction accuracy of the RFI using SSA [20], defined as $\epsilon_R = 20 \log_{10}\left(\frac{\|x_R\|}{\|x_R - \hat{x}_R\|}\right)$ as a function of the interference to noise ratio (INR). It is seen that the reconstruction accuracy improves with INR, which is the main motivation for characterizing RFI at the BS instead at the telescope, which receives the RFI at a much lower SNR due to propagation loss. This leads to improper eigenspace representation and erroneous cancellation. We evaluate this in the following section.

B. Experimental Results

1) Reconstruction of RFI-free astronomical signal: Following the signal reconstruction steps, described above, we obtain, the rectified astronomical signal as shown in Figure 7 along with the true astronomical signal and the composite signal. The power levels are relative with measured noise floor at the
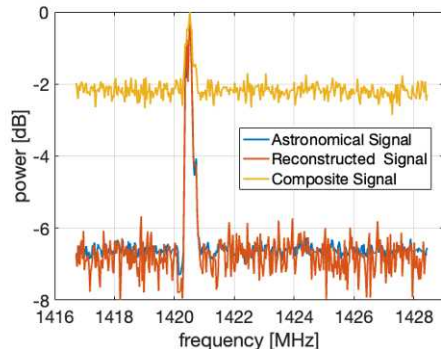
telescope (-174 dBm) as the baseline. We achieved an RQF of 24.6944 dB for this reconstructed signal, which is much higher than the theoretical lower bound of RQF = 10.0007 dB based on the derivation in § IV.

2) Effect of Interference to Noise Ratio (INR): Since RFI is generally acquired by the telescope side-lobes, a low interference to noise ratio (INR) is expected. For the reconstructed astronomical signal shown in Figure 7, INR, i.e, ratio of RFI and and noise power, is set at 5 dB. In most practical cases, this will vary based on distance between telescope and the source BS and side lobe attenuation. We have investigated the effect of INR varying from -10 dB to 15 dB on reconstruction quality using the RQF metric, shown in Figure 8a. RQF improves with the INR as higher interference power results in better eigenspace projection, which in turn facilitates better reconstruction quality [21].

3) Effect of synchronization error: GPS clocks have maximum time-synchronization error of 30 nanoseconds as reported in [22]. Neglecting the transmission time of the RFI characterisation information compensated through the introduction of appropriate delays at the telescope, at most 4 samples of duration 32.826 ns can be out of synchronization. Effect of time synchronization error between BS and telescope, of upto ten samples out of synchronization, which is equivalent to 320 nanoseconds, is shown in Figure 8b. Due to synchronization error, we can observe that the RQF goes down to 20 dB, but it still stays significantly higher than the theoretical lower bound.

4) RQF for varying spectral occupancy of the RFI: In a practical situation, LTE frames are not fully occupied and many resource blocks remain empty. The reconstructed signal in Figure 7 had an RFI consisting of fully occupied LTE frame. We have observed that with varying spectral occupancy, reconstruction quality of the astronomical signal changes as well. With different spectral occupancy, the eigen space of the RF signal gets skewed which produces higher error while projecting the composite signal and reduces RQF. Figure 8c shows that decreasing occupancy reduces the RQF gradually.

5) RQF if RFI is decomposed at the telescope: To prove our claim that characterizing RFI at BS, i.e, at higher power level

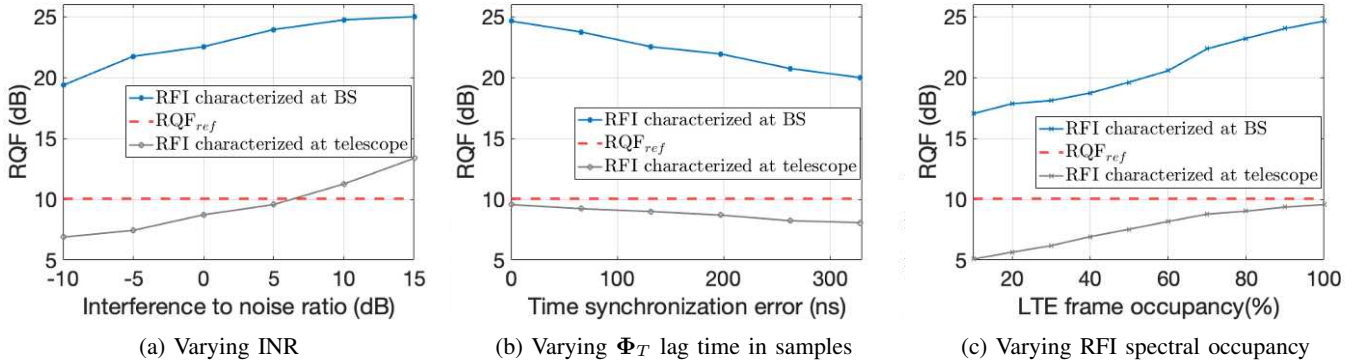| (a) Varying INR | (b) Varying Φ_T lag time in samples | (c) Varying RFI spectral occupancy |

Fig. 8: Reconstruction Quality Factor (RQF) variation in different RFI parameters

leads to superior RFI mitigation, We used the emulated LTE signal at power level same as at the telescope and followed the steps from § V-A and tested all three scenarios mentioned above. As we can observe in Figures 8a, 8b and 8c, the RQF is mostly below the theoretical bound when RFI is characterized at telescope, rendering the proposed apparatus questionable. This proves that, to obtain best results, the RFI characterization must be done at BS.

VI. RELATED WORK

RFI mitigation in radio astronomy: Active RFI mitigation has become a necessary practice in the radio astronomy community. Known persistent and fixed sources of RFI are highly attenuated at the front-end of the receiver using series of analog superconductive filters [23], [24], but frequencies with high RFI density (e.g. FM or Digital Video Broadcast band) are usually simply avoided by design. Fast processors (RFSoC and FPGA) enable the detection and the blanking of impulsive RFI in the baseband digital samples, post-digitization [25], [26]. Data flagging consists of detecting time and frequency data corrupted with RFI, and discard them by replacing these values with zeros or random noise. This process occurs after channelization and time integration, where the intermediate telescope data product has the appropriate time resolution (of the order of 1 ms) to match most RFI duty cycles. This can be applied manually after a careful inspection of the collected data, but conventional data reduction software include automated flagger based on local and global statistics of a given dataset [27], [28]. Research in RFI flagging has also resulted in the development of real-time "on-the-fly" data flaggers [29], and the use of machine learning to automatically recognize and classify detected RFI [30].

Telescope arrays provide spatial information in addition to the time and frequency signatures of the studied objects captured in the telescope correlation matrix. RFI spatial signatures can be extracted from this matrix in order to build an adapted spatial filter, and eventually recover uncorrupted time and frequency data [31]. These methods remain at an experimental level due to their impact on the array calibration. Finally, the subtraction of an estimated and reconstructed RFI waveform from the telescope data have been demonstrated, but

have never been deployed due to their heavy computational complexity [32].

Interference cancellation in communication systems: Collaboration among wireless technologies [33], [34], or avoiding incumbents [35], primarily employ sensing and database management for active users, that cannot be extrapolated to passive users due to an absence of active transmissions. Interference cancellation in wireless communication [36], [37], require decoding the strongest signal first in order to cancel it. These cannot be applied for RFI mitigation or coexistence of active and passive users because: a) the astronomical signal is not a modulated signal with known characteristics, b) the RFI at the passive user is of extremely low power, which cannot be decoded to remove it, and c) sharing active communication signals as digital samples with passive users for cancellation at the telescope, incurs prohibitive bandwidth. Hence, it is essential to accurately characterize RFI in a condensed format such that it can be shared with passive users for cancellation, while preserving user privacy and adhering to cellular standards.

Basis expansion and orthogonal projections: Since, KLT incurs high computational complexity to extract a very large number of eigenfunctions [38], we rely on SSA to implement KLT which extracts only the L eigenfunctions with largest eigenvalues. Spatial filtering of interference or noise projects the signal to the null space of the undesired components [21]. Moreover, methods for signal separation [18], reconstruct the desired signal by identifying correlations in the eigenvectors. These typically assume orthogonality of the subspaces [18], which does not always hold true in practice. Non-orthogonal projections like oblique projection [39] reduces the power distortion of the reconstructed signal even when the orthogonality of astronomical and RFI subspaces are not verified, however requires accurate estimation of null-spaces of RFI which is often challenging when both signals are weak or non-disjoint. In contrast, collaborative RFI cancellation does not require the orthogonality of the astronomical and RFI subspaces at the telescope to nullify the RFI, since the RFI subspace is accurately identified at high SNR at the BS.

ACKNOWLEDGEMENT

VII. CONCLUSION

This work advances the literature on RFI mitigation for radio astronomy by sharing stochastic characterization of the RFI at its source, the cellular base station, with the telescope to cancel the incident RFI. The method has the potential to be deployed in an actual radio telescope, like DSA-110 at OVRO, and promote collaborative spectrum sharing between the active and passive users of the spectrum. The high reconstruction quality of the RFI-free astronomical signal in our evaluations will further motivate both research communities to apply the method to eliminate other forms of RFI in various bands allocated for radio astronomy and other passive services. However, managing computational complexity of large eigenvalue problem remain a challenge for real-time RFI cancellation.

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] U. Rau, R. Selina, and A. Erickson, "Rfi mitigation for the ngvla: A cost-benefit analysis ngvla memo# 70," 2019.

[3] G. Hallinan, V. Ravi, S. Weinreb, J. Kocz, Y. Huang, D. Woody, J. Lamb, L. D'Addario, M. Catha, J. Shi *et al.*, "The dsa-2000–a radio survey camera," *arXiv preprint arXiv:1907.07648*, 2019.

[4] National Academies of Sciences Engineering and Medicine, *Handbook of Frequency Allocations and Spectrum Protection for Scientific Uses: Second Edition*. Washington, DC: The National Academies Press, 2015. [Online]. Available: https://www.nap.edu/catalog/21774/handbook-of-frequency-allocations-and-spectrum-protection-for-scientific-uses

[5] The International Telecommunication Union - Radiocommunication Bureau, *Handbook on Radio Astronomy*. The International Telecommunication Union - Radiocommunication Bureau, 2013.

[6] Committee on Radio Astronomy Frequencies (CRAF), *CRAF Handbook for Radio Astronomy*. EUROPEAN SCIENCE FOUNDATION, 2005. [Online]. Available: https://craf.eu/wp-content/uploads/2015/02/CRAFhandbook3.pdf

[7] J. Kocz, V. Ravi, M. Catha, L. D'Addario, G. Hallinan, R. Hobbs, S. Kulkarni, J. Shi, H. Vedantham, S. Weinreb *et al.*, "Dsa-10: a prototype array for localizing fast radio bursts," *Monthly Notices of the Royal Astronomical Society*, vol. 489, no. 1, pp. 919–927, 2019.

[8] T. L. Wilson, K. Rohlfs, and S. Hüttemeister, *Tools of radio astronomy*. Springer, 2009, vol. 5.

[9] G. B. Taylor, C. L. Carilli, and R. A. Perley, "Synthesis imaging in radio astronomy ii," *Synthesis Imaging in Radio Astronomy II*, vol. 180, 1999.

[10] 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.

[11] E. Petroff, J. Hessels, and D. Lorimer, "Fast radio bursts," *The Astronomy and Astrophysics Review*, vol. 27, no. 1, pp. 1–75, 2019.

[12] 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

[13] 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.

[14] M. Trudu, M. Pilia, G. Hellbourg, P. Pari, N. Antonietti, C. Maccone, A. Melis, D. Perrodin, and A. Trois, "Performance analysis of the karhunen–loève transform for artificial and astrophysical transmissions: denoizing and detection," *Monthly Notices of the Royal Astronomical Society*, vol. 494, no. 1, pp. 69–83, 2020.

[15] A. M. Tomé, D. Malafaia, A. R. Teixeira, and E. W. Lang, "On the use of singular spectrum analysis," 2018.

[16] R. Vautard, P. Yiou, and M. Ghil, "Singular-spectrum analysis: A toolkit for short, noisy chaotic signals," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1, pp. 95–126, 1992. [Online]. Available: https://www.sciencedirect.com/science/article/pii/016727899290103T

[17] A. Szumski, "Karhunen loeve transform as an instrument to detect weak rf signals," InsideGNSS, European Space Agency, Tech. Rep., 2011.

[18] N. Golyandina and A. Zhigljavsky, *Singular Spectrum Analysis for Time Series*, 01 2013.

[19] R. I.-R. RA.769-2, "Protection criteria used for radio astronomical measurements."

[20] J. Harmouche, D. Fourer, F. Auger, P. Borgnat, and P. Flandrin, "The Sliding Singular Spectrum Analysis: a Data-Driven Non-Stationary Signal Decomposition Tool," *IEEE Transactions on Signal Processing*, Sep. 2017. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01589464

[21] A.-J. Boonstra, "Radio frequency interference mitigation in radio astronomy," Ph.D. dissertation, Technische Universiteit Delft, 2005.

[22] "Gps.gov: Gps accuracy," https://www.gps.gov/systems/gps/performance/accuracy/#timing, (Accessed on 10/15/2021).

[23] 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.

[24] 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.

[25] C. Dumez-Viou, R. Weber, and P. Ravier, "Multi-level pre-correlation rfi flagging for real-time implementation on uniboard," *Journal of Astronomical Instrumentation*, vol. 5, no. 04, p. 1641019, 2016.

[26] D. Ait-Allal, C. Dumez-Viou, R. Weber, G. Desvignes, I. Cognard, and G. Theureau, "Rfi mitigation at nançay observatory: Impulsive signal processing," in *Widefield Science and Technology for the SKA SKADS Conference 2009*. ISBN 978-90-805434-5-4, 2010, pp. 201–205.

[27] 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.

[28] 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.

[29] A. Sclocco, D. Vohl, and R. V. van Nieuwpoort, "Real-time rfi mitigation for the apertif radio transient system," in *2019 RFI Workshop-Coexisting with Radio Frequency Interference (RFI)*. IEEE, 2019, pp. 1–8.

[30] C. J. Wolfaardt, "Machine learning approach to radio frequency interference (rfi) classification in radio astronomy," Ph.D. dissertation, Stellenbosch: Stellenbosch University, 2016.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] C. Maccone, *A simple introduction to the KLT and BAM-KLT*, 08 2012, pp. 411–448.

[39] G. Hellbourg, R. Weber, C. Capdessus, and A.-J. Boonstra, "Oblique projection beamforming for rfi mitigation in radio astronomy," in *2012 IEEE Statistical Signal Processing Workshop (SSP)*. IEEE, 2012, pp. 93–96.