

Collaboration with Cellular Networks for RFI Cancellation at Radio Telescope

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

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

†Department of Astronomy, California Institute of Technology

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

Abstract—The growing need for electromagnetic spectrum to support the next generation (xG) communication networks increasingly generate unwanted radio frequency interference (RFI) in protected bands for radio astronomy. RFI is commonly mitigated at the Radio Telescope without any active collaboration with the interfering sources. In this work, we provide a method of signal characterization and its use in subsequent cancellation, that uses Eigenspaces derived from the telescope and the transmitter 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.*, autocorrelation) of the RFI, typically observed in communication systems. We have presented effectiveness of this collaborative method using real-world astronomical signals and practical simulated LTE signals (downlink and uplink) as source of RFI along with propagation conditions based on preset benchmarks and standards. Through our analysis and simulation using these signals, we are able to remove 89.04% of the RFI from cellular networks, which reduces excision at the Telescope and is capable of significantly improving throughput as corrupted time-frequency bins of data become usable.

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

I. INTRODUCTION

Radio Astronomy is a discovery-based science, which has revolutionized our understanding of the Universe through scientific observations across the electromagnetic (EM) spectrum. However, only 1-2% of the spectrum is allocated for science below 50 GHz where almost all of the commercial radio communication occurs. It has become increasingly essential to use spectrum other than the current allocation for astronomical observations for two main reasons: 1) red-shifting of spectral lines due to the expanding Universe and 2) broad bandwidth radio continuum observations can increase the signal-to-noise ratio of weak radio sources. Radio telescopes are generally located in geographically isolated areas to avoid radio frequency interference (RFI) from human generated electromagnetic waves. However, no matter how remote, all radio astronomy sites across the world are vulnerable to growing RFI from terrestrial networks used to extend coverage to the increasing human population [1], [2]. Generally, communication system designers strive to reduce noise from artificially generated signals, whereas radio astronomy focuses on removing communication signals from the astronomical

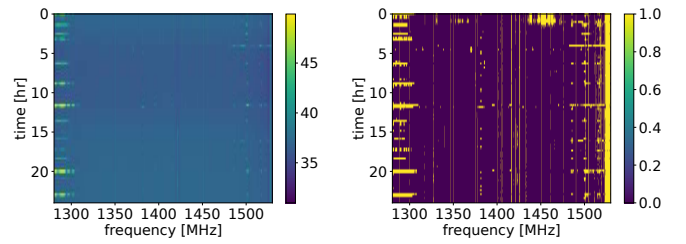
signal. This seemingly opposing requirement is pushing the two communities farther away. Both are equally essential and in essence are designed to overcome a common bottleneck: *RFI*.

Generally, for the purpose of interference cancellation, collaboration among wireless technologies [3], [4], or avoiding incumbents [5], primarily employ sensing and database management for active users, which cannot be extrapolated to passive users such as radio telescopes due to absence of active transmissions. Interference cancellation in wireless communication [6], [7], 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 signals are not modulated signals with known characteristics, b) the RFI at the passive user is of extremely low power, which cannot be decoded to remove it, c) sharing active communication signals as digital samples with passive users for cancellation at the telescope, incurs prohibitive bandwidth, and d) other interference cancellation techniques require perfect synchronization between RFI source and the telescope receiver for accurate cancellation.

Impact of RFI in Radio Astronomy: Probing the Universe at radio frequencies offers the possibility 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, astronomical objects are moving away from us with increasingly high velocities, which creates a Doppler effect, shifting all spectral lines from their rest frequencies to lower frequencies, a phenomenon termed “redshift”. As a result, the HI line is now observed at frequencies around 1415 MHz from the Virgo cluster, at 1400 MHz from the Perseus supercluster, and at 1388 MHz from the Coma cluster of galaxies. The highly redshifted HI line (observed even 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 [8]. Redshifted sources require astronomers to increasingly observe outside the protected bands [9], [10], [11]. Moreover, astronomical emissions are extremely weak due to the large distances travelled before being detected by radio telescopes. Often, detection requires long integration

at ≈ 40 dB below the telescope noise. This is in stark contrast to reception of a typical wireless communication signal at ≈ 20 dB 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. The RFI situation is worsening, mostly driven by two factors: 1) population growth, increasing the chances of RFI polluting telescope sites, and the proliferation of next-generation (xG) wireless technologies, like 5G and rural wireless broadband [12], [13], [14], 2) at the same time, technological advances enable the development of wideband and low system temperature receivers, resulting in dramatic improvements to the sensitivity of modern radio telescopes to faint signals of astronomical origin [15], [16], [17].

Active RFI mitigation in current systems is achieved in different methods. Known persistent and fixed sources of RFI are highly attenuated at the front-end of the receiver using series of analog superconductive filters [18], [19], 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 [20], [21]. Data flagging consists of detecting time and frequency data corrupted with RFI, and discarding them by replacing these values with zeros or random noise - defined as excision. 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 or using automated flaggers in data reduction softwares based on local and global statistics of a given dataset [22], [23]. Research in RFI flagging has also resulted in the development of real-time “on-the-fly” data flaggers [24], and the use of machine learning to automatically recognize and classify detected RFI [25]. Figure 1 illustrates the excision problem using data collected with the Deep Synoptic Array DSA-110 [26], [27] located at the Owens Valley Radio Observatory (OVRO), CA, USA (§ II-A). Despite being geographically remote, it is affected by telemetry, communication, and RADAR RFI. The data in Figure 1a spanning 1280-1350 MHz is corrupted with three types of RFI encountered in radio astronomy: continuous in time and narrow in frequency; intermittent in time and narrow in frequency; and impulsive in time and wide in frequency. Figure 1b shows the same data after identification and excision of the RFI-corrupted time-frequency bins. Here, the RFI detection consists of correcting the baseline of the averaged spectrum or time series using a median filter, then identifying the power excesses due to RFI on both time and frequency axis. The contaminated ranges of frequencies and times are then discarded. This flagging and excision approach is usually tuned to minimize the probability of non-detection of the RFI resulting in a significant data loss, sometimes as high as 40% at L-band (1-2 GHz), 30% at S-band (2-4 GHz) and 20% at X-band (4-8 GHz) [13], impacting the sensitivity of the telescope and recovery of the astronomical



(a) DSA-110 spectrogram over 24 (b) A large fraction of the data are hours. The data are affected with flagged and discarded for further various types of RFI, including astronomical processing. telemetry and radar systems.

Fig. 1: RFI flagging and excision with data from the DSA-110 at Owens Valley Radio Observatory (OVRO) [27]

signal of interest. Telescope arrays provide spatial information in addition to the time and frequency signatures of the studied objects captured in the telescope correlation matrix, which can be extracted in order to build an adapted spatial filter, and recover uncorrupted time and frequency data [28]. These methods remain at an experimental level due to their impact on the array calibration. Subtraction of an estimated and reconstructed RFI waveform from the telescope data has been demonstrated, but has never been deployed due to their heavy computational complexity [29].

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. So, we present a collaborative framework by aggregating concise, yet accurate signal *characterization* from the RFI source, that can be *intelligently cancelled* from the telescope data to reveal the astronomical signal. This method requires limited collaboration overhead, preserves user privacy, and adheres to cellular standards. In this work, we focus on the sub-6 GHz bands as it is among the busiest spectral windows, heavily exploited by commercial wireless applications, while offering a unique opportunity to observe astronomical emissions like continuum synchrotron emissions (e.g. pulsars show strongest emissions below 600 MHz), or at-rest and redshifted spectral lines like Neutral Atomic Hydrogen or Hydroxyl [30]. We have identified 4G/5G cellular service as a source of RFI, which transmits at high power, covers a large bandwidth (~ 2 GHz), is most abundant, and can propagate large distances. Baseline results for such RFI cancellation with simple radio signals as contamination have been presented in our previous work [31]. However, this proposed method is applicable to remove a variety of RFI in other frequency bands as well. The primary contributions of this paper can be enlisted as:

- 1) We present a novel collaborative method of RFI mitigation in radio astronomy by characterizing RFI at its source and sharing concise information with the radio-telescope for RFI cancellation, which is able to remove 89.04% of RFI power from the corrupted astronomical signal.
- 2) This method utilizes signal characterization that is robust to signal to noise ratio (SNR) and adaptable to changing signal bases, which is highly desirable in low power and passive

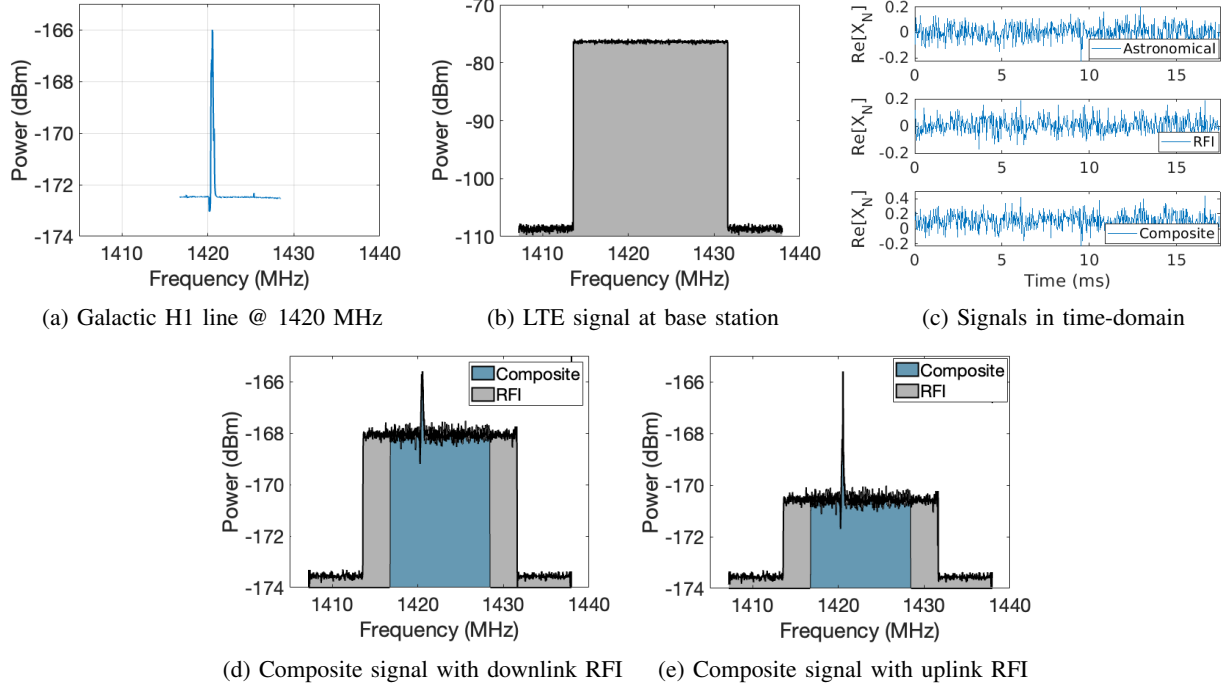


Fig. 2: Spectral and temporal characteristics of Astronomical, RFI and Composite signals. RFI signal parameters for the downlink sample signal: 20 MHz bandwidth, 60% frame occupancy, RFI signal parameters for uplink sample signal: 20 MHz bandwidth, 60% frame occupancy, contribution from 10 UEs.

spectrum usage regime.

3) The proposed model is analyzed and evaluated using real astronomical signals captured at Owens Valley Radio Observatory (OVRO) and simulated cellular RFI.

4) We present system evaluation for practical transmission conditions of RFI and its impact on the proposed cancellation apparatus.

5) Finally, we evaluate the impact of the design parameters of the proposed system and the overhead created by the collaborative mechanism contingent on practicality of the system.

To the best of our knowledge, no prior work applies a priori information of RFI to continuously cancel the interference at the telescope. The rest of the paper is organized as: §II presents the signal models of RFI and astronomical signals, §III elaborates the complete collaborative RFI mitigation method, in §IV we introduce a novel evaluation metric for the proposed method based on benchmarked accuracy thresholds of astronomical signal recovery. §V expands in two parts: In §V-A, experimentation and evaluation method of the RFI cancellation apparatus is detailed, §V-B expands on the system performance, impact of design parameters, different RFI transmission scenarios, and collaboration overhead analysis. Finally, we conclude the paper in §VI.

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 [32], [33]. After signal conditioning (i.e., amplification, equalization, and filtering), the output of the individual receivers is digitized over hundreds of MHz and channelized into smaller frequency bins of hundreds of kHz width [34]. 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 astronomical dataset, utilized throughout in this paper, has been collected with the Deep Synoptic Array being deployed at the OVRO. Figure 3 depicts the processing steps involved in data collection and post-processing at OVRO and are elaborated in § V-A. Figure 2a shows the galactic H1 line at 1420 MHz, collected at OVRO, which is identified as the astronomical signal of interest in this work. It is important to note that the H1 line visible at 1420 MHz falls within a protected band dedicated to radio astronomy and does not contain real RFI. We have injected simulated LTE signals as RFI to the astronomical signal for experimentation purposes.

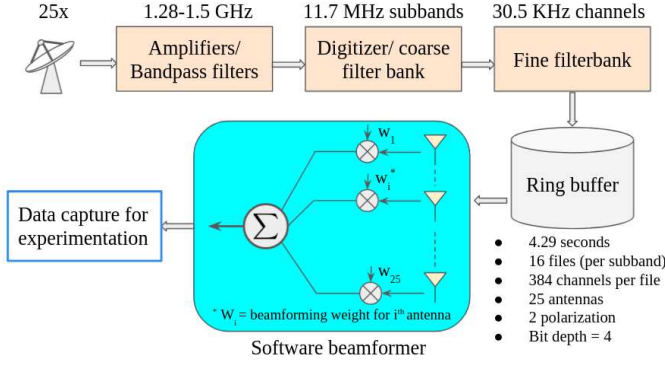


Fig. 3: Block diagram of signal capture and post-processing at telescope

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) [32], [35], [36] is expressed as:

$$x_T[n] = x_A[n] + x_N[n] + x_R[n] \approx x_N[n] + x_R[n] \quad (1)$$

where $x_T[n]$ is the channelized baseband signal at a filterbank channel centered around frequency f_c and at time sample n , and follows a stationary (assumed over a short duration) stochastic process, and 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 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 $\sigma_R^2 = N^{-1} \sum_N |x_R[n]|^2$. We assume $\sigma_A^2 \ll \{\sigma_N^2, \sigma_R^2\}$, which allows the approximation in equation (1), as, in single antenna output, the astronomical contribution is suppressed under system noise [35], [32]. However, accurate characterization of $x_T[n]$ 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 transmissions using a carrier at frequency f_0 is shown in (2):

$$x_R(t) = \text{Re} \left\{ e^{j2\pi f_0 t} \sum_{k=-N_{\text{FFT}}/2}^{N_{\text{FFT}}/2} \alpha_k e^{j2\pi k(t-t_g)/T_u} \right\} \quad (2)$$

Where α_k is the m-ary modulated symbol which is placed on the k_{th} subcarrier, T_u is the symbol duration and k is the subcarrier index. For downlink transmission, orthogonal frequency division multiple access (OFDMA) is employed, whereas, for uplink transmission purposes, single carrier frequency division multiple access (SC-FDMA) is used. Both of these methods are variations of the frequency division

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					
Sampling Frequency (MHz)	1.92	3.84	7.68	15.36	23.04	30.72
N_{FFT}	128	256	512	1024	1536	2048
$N_{\text{subcarrier}}$ (Including DC subcarrier)	76	151	301	601	901	1201
N_{guard}	52	105	211	423	635	847
Resource Blocks	6	12	25	50	75	100
OFDM Symbols/Frame	7/6 (short CP/long CP)					

TABLE I: Different LTE parameters used for simulation

multiplexing scheme tuned based on the requirements and limitations in uplink and downlink. Figure 4a presents the subcarrier assignment scheme of modulated symbols for both OFDMA and SC-FDMA with a comparable reference frame including four symbols. In Figure 4b we observe the difference in the transmitter-receiver chain of both modulation schemes being the requirement of an additional DFT/IDFT stage in symbol generation/detection for SC-FDMA. It dictates the symbol assignment to subcarriers represented in Figure 4a mitigating the issue of high peak to average power ratio due to parallel transmission of different symbols as it can be detrimental for power-limited User equipment (UE), while maintaining the other benefits of OFDMA.

The signal bandwidth depends on the size of the FFT, number of subcarriers, guard bands, etc. as described in Table I. These parameters are identical for OFDMA and SC-FDMA used in downlink and uplink. Typically, the LTE signal is grouped as radio frames in baseband with each frame containing 10 subframes and each subframe 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 characterization. An example realization of the downlink LTE signals used as RFI in this work is shown in figure 2b. In reality, if the telescope acquires an astronomical signal at or around f_0 , this RFI will undergo the same processing as mentioned in §II and will be present in varying strength across multiple telescope channels depending on the telescope aperture, side-lobe gain and the spectral occupancy of the LTE signal. We refer to this signal as the *composite signal*, which is defined by (1) and (2). An example time-domain signal representation of the RFI and composite signal is depicted in Figure 2c. The RFI appears stochastic in time due to the polyphase channelization of the telescope data. The power spectral density of the composite signal in presence of downlink RFI and uplink RFI from identical cell topology and distance from telescope but varying traffic patterns are presented in Figures 2d and 2e. Additional information regarding the LTE signal generation for experimentation and validation purposes of the proposed method is provided in §V-A.

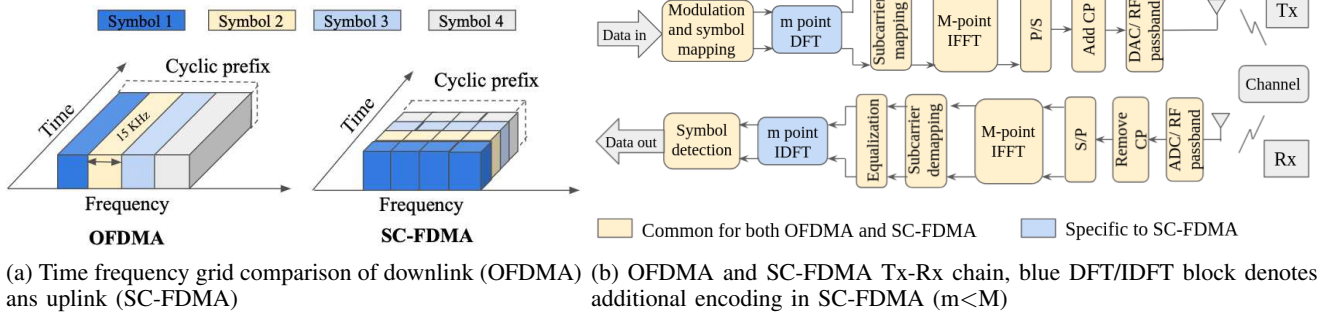


Fig. 4: Uplink and Downlink LTE signal generation

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. As a result, the RFI signal information is extremely limited due to equipment constraints (front-end, BW, gain, etc.), signal deterioration due to propagation, and lack of coherence between the RFI source and the telescope. Another limitation stems from the very methods employed to characterize the RFI, which are almost always limited to Fourier methods for frequency domain analysis and temporal statistics like autocovariance, cyclostationarity, or higher-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 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 requires a high degree of synchrony between the telescope and RFI sources for phase coherent cancellation. Moreover, methods for signal separation [37] reconstruct the desired signal by identifying correlations in the eigenvectors. These typically assume orthogonality of the RFI and astronomical signal subspaces [37], which does not always hold true in practice. Non-orthogonal projections like oblique projection [38] 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 and column-spaces of astronomical signals which is often challenging when both signals are weak or non-disjoint.

Our method is radically different from the state of the art in three aspects: 1) The RFI is characterized at the cellular base station (BS) into a compact yet accurate eigenspace that is periodically shared with the telescope via a shared channel over the Internet. Unlike Fourier methods, which decompose 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 10 in §V-B discusses this unique feature; 2) At the telescope the composite signal is decomposed using

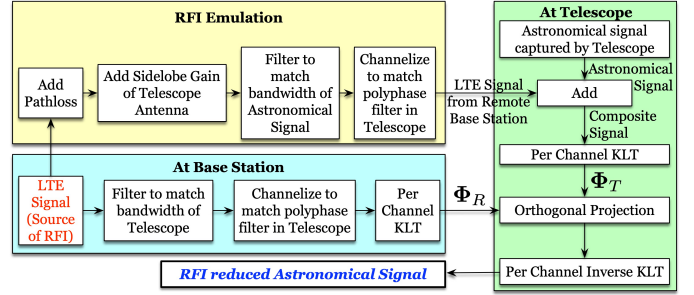


Fig. 5: Block diagram of the RFI cancellation system

the same method revealing its eigenspace that contains 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 orthogonal projections. Since the cancellation happens in the eigenspace, a final step to convert the eigenspace to the time-domain signal will reveal the RFI-free astronomical signal. Figure 5 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 an LTE signal and is combined with the real astronomical signal as described in §II¹. Our analysis and simulation are based on this composite signal but adhere 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 5 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 with other topological information given in Table II.

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.
^{*} see §III-A.

TABLE II: Shared parameters for collaborative cancellation

Each of the parameters have specific roles in the cancellation apparatus and are explained in subsequent sections where applicable. The static parameters are constant and can be

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

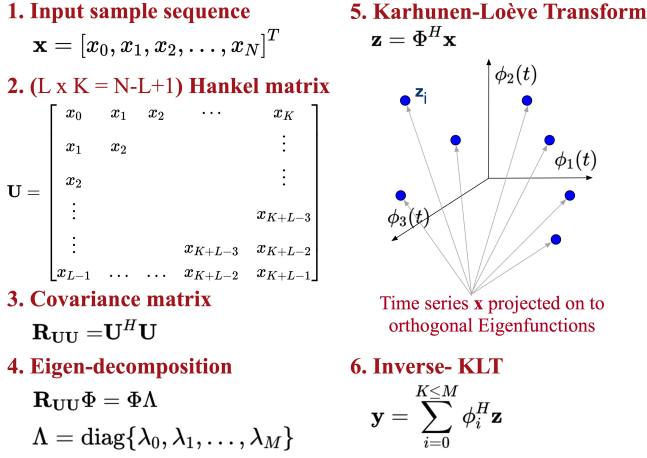


Fig. 6: Geometry of KLT based signal characterization

made available via a database lookup. The dynamic parameters change over time and require periodic sharing. The largest update in terms of size is RFI Eigenspace information, Φ_R . For the 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. However, it is possible to use the same Φ_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. Although, this may not be applicable for RFI sources with lower bandwidth than the telescope's field of view or if there are interference signals present with partial overlap across the telescope bandwidth causing frequency span mismatch between RFI and fine channels of the telescope. 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 Φ_R . Determination of L is discussed later in this section and a brief overhead analysis is provided in §V-B8.

A. Eigenspace representation of RFI and composite signals

The Karhunen-Loève Transform (KLT) [39], [40], [41], [42] decomposes any signal, 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. Additionally, KLT is applicable to a signal irrespective of its stationarity [39], [40] as basis expansion in KLT is not limited to time and frequency domain. Thus there is no stringent requirement for the RFI signal stationarity. 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 are time series obtained per fine channel, 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 $L \times K$ [43], [44] as shown in figure 6, where N is the number of time samples in the signal of interest $x[n]$, L is the KLT window size, determined at the source of RFI (BS) based on the RFI signal being observed ($L < K$).

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

where $\mathbf{x}_i = [x[n], \dots, x[n+L-1]]^T$ are lagged vectors of size L , with $i \in [1, K]$, and $\mathbb{E}[\cdot]$ is the expectation operator. As a bottleneck of the collaborative apparatus, L directly impacts the amount of data sharing and data rate required for success of the RFI cancellation apparatus. Determination of optimal L is accomplished in two steps: 1) Necessary condition - lower bound of L that is required to prevent under-estimation of RFI subspace, elaborated in §III-B, 2) Sufficient condition - lower bound of L that satisfies the construction accuracy criterion described in §IV. This is determined empirically, as described in §V-B7.

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

$$\mathbf{R}_{xx} = \Phi \Lambda \Phi^H, \quad \text{where, } \Lambda = \text{diag}\{\lambda_0, \dots, \lambda_{L-1}\} \quad (4)$$

where λ_j are the eigenvalues, with $j \in [0, L-1]$, and Φ is a unitary matrix containing L eigenvectors as its columns. Since, (4) decomposes the temporal correlations of $x[n]$, each column of Φ , 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. Astronomical signals are generally modeled as stationary signals [45], [35]. Cellular signals, although typically cyclostationary, channelization of cellular signals to a narrowband (30.5 KHz) generates a stationary component on which KLT is performed. KLT, being independent of stationarity of the signal, is not impacted by this changing property and is able to accurately characterize the signals. The eigenfunctions project the signal on to the Hilbert space given by, $\mathbf{z}_i = \Phi^H \mathbf{x}_i$, 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 6. Consequently, $\mathbf{z}$ is given by, $\mathbf{z} = \Phi \mathbf{U}$ using the definition of $\mathbf{U}$ in (3). Projecting $\mathbf{z}$ onto the $M (\leq L)$ eigenfunctions of Φ with largest eigenvalues, reconstructs the signal, $\hat{x}[n]$ with minimum noise [46] as shown in figure 6. 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* (Φ_T) and *RFI Kernel* (Φ_R) respectively.

The eigenfunctions of decomposition of the composite sig-

nal at the telescope are shown in figure 7a. Figure 7b compares the 50 largest eigenvalues for the RFI at the BS, attenuated incident RFI at the telescope and the composite signal.

1) *Additional measures for uplink RFI decomposition:* One of the crucial components of the proposed RFI cancellation apparatus is the characterization of RFI signal at its source. It is done so for downlink signals in baseband at the BS. Unfortunately, that is not possible in case of uplink RFI for several reasons but not limited to: 1) UEs (source of uplink signal) do not have information about simultaneous uplink transmission, 2) UEs are power limited and lack the computation capability for decomposition and sharing the parameter sets of eigenfunctions, 3) efficient combining individual KLT outputs from UEs is hard to achieve. Thus, we propose the uplink signal decomposition to be performed at the respective BS the UEs are reporting to at a given time. Thus characterization is performed on the received signal at BS that is likely to include channel distortions and additive noise. During evaluation, we have explored the consequences of such uplink RFI characterization.

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 (Φ_T) onto a subspace that is *orthogonal* to the subspace spanned by the RFI eigenfunctions (Φ_R).

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

$$\mathbf{P}_{\Phi_R}^\perp = \mathbf{I} - \Phi_R (\Phi_R^H \Phi_R)^{-1} \Phi_R^H \quad (5)$$

where $\mathbf{I}$ is the $L \times L$ identity matrix. $\mathbf{P}_{\Phi_R}^\perp$ is such that $\mathbf{P}_{\Phi_R}^\perp \Phi_R = 0$ and by extension we can show that, $\mathbf{P}_{\Phi_R}^\perp x_R[n] = 0$ using the definitions in (4) and (3). Therefore, applying this orthogonal projection $\mathbf{P}_{\Phi_R}^\perp$ 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, Φ_R , is improved by estimating it at the BS where the RFI is received at high SNR, and consequently $\mathbf{P}_{\Phi_R}^\perp$ can be calculated even if the astronomical and RFI signals are not separable at the telescope. Figure 7c shows a typical downlink RFI subspace compared to the astronomical signal subspace and the system noise. The RFI subspace Φ_R is defined by $M \approx 300$ dimensional subspace whereas astronomical signal and system noise, being white processes, are defined on an infinite dimensional space [35], [45]. As Φ_R (with dimension) characterizes RFI subspace accurately at its source, cancelling RFI using orthogonal projection in (7) ensures that any RFI is nulled with precision. In case of non-orthogonality (non-separable RFI and astronomical subspaces), this cancellation method will cost some astronomical signal energy, but with a carefully chosen $L (> M)$, we can have enough eigenfunctions to reconstruct the astronomical signal with high accuracy. So, 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 [37], this procedure does not require such assumptions. $L > M$ defines the necessary condition for the choice of L .

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

$$\hat{\Phi}_T = \mathbf{P}_{\Phi_R}^\perp \Phi_T \quad (6)$$

This projects the composite signal subspace at the telescope to the null-space of the RFI. Consequently, $\hat{\Phi}_T$ spans the subspace that is orthogonal to the RFI subspace spanned by Φ_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, $\hat{\mathbf{U}}_T$, is reconstructed by projecting the matrix $\mathbf{z}_T$ onto the projected eigenfunctions $\hat{\Phi}_T$ as given by (7).

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

Finally, the cross-diagonal elements of the reconstructed Hankel matrix are averaged [37] using (8), to reconstruct the space-signal time-series, $\hat{x}_T[n]$ from $\hat{\mathbf{U}}_T$.

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

where the superscript $(n, k-n+1)$ indicates the corresponding element in the matrix $\hat{\mathbf{U}}_T$. This form of real-time recovery of $\hat{x}_T[n]$ is not possible with present methods in practice. 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

The performance of the proposed RFI mitigation approach is evaluated empirically using the Reconstruction Quality Factor (RQF) which measures the distortion in the recovered power after mitigating the interfering signal.

The RFI-mitigated signal at the telescope can be expressed as:

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

where $x_T[n]$ is the channelized baseband signal at telescope filterbank channel at time sample n as described in §II-B, $\hat{x}_R[n] = x_R[n] + \epsilon_r[n]$ is the estimated RFI contribution, and $\epsilon_r[n] \sim \mathcal{NC}(0, \sigma_{\text{est}}^2)$ captures the cumulative estimation and reconstruction error. The RQF over N samples is defined as:

$$\text{RQF} = \frac{\|\hat{x}_T - \hat{x}_T\|^2}{\|\hat{x}_T\|^2} \quad (10)$$

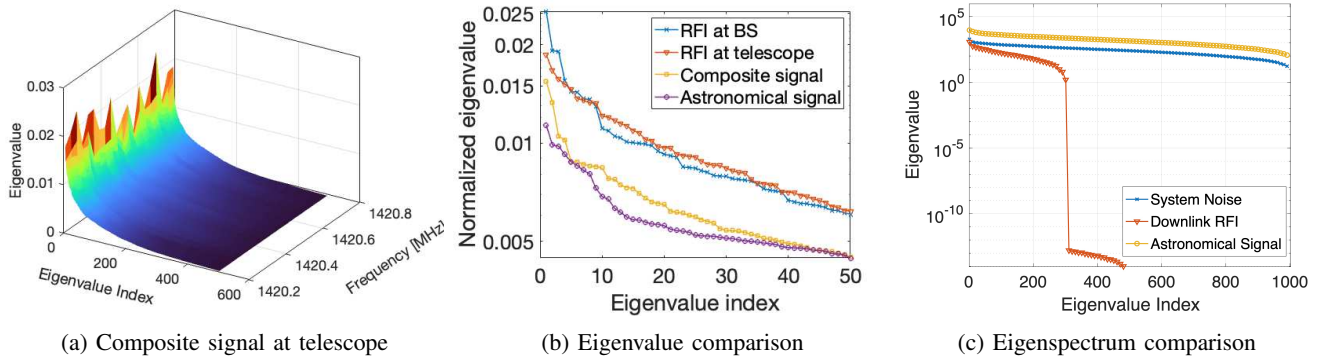


Fig. 7: Derived eigenspace from signal decomposition.

where $\tilde{x}_T = x_A[n] + x_N[n]$ is the true RFI-free astronomical signal, and $\|x\|^2 = \frac{1}{N} \sum_N x^2[n]$ is the mean square error operator.

The advantage of defining the RQF this way is that it directly relates to the International Telecommunication Union (ITU) detrimental-level interference criterion [47], which is set to 10% power distortion of a 2000-seconds long integration. Assuming the independence of the RFI-free telescope time samples, this detrimental level can be adapted to the 4-seconds data sets produced by the telescope back-end. In that case, our proposed RFI mitigation approach reaches the ITU level when:

$$\text{RQF} \leq \frac{10\%}{2000s} \times 4s = 0.02\% \quad (11)$$

V. EVALUATION AND RESULTS

A. Experimental Setup

1) *Astronomical Signal Acquisition:* The astronomical data used in this work have been collected on May 15 2021 with the Deep Synoptic Array (DSA-110, see Fig. 8), a radio interferometer made of the 110 4.65 m-antennas operating in the 1280 - 1530 MHz band, and located at the Owens Valley Radio Observatory near Bishop, CA [48]. At the time of data collection, only 25 antennas were built and operational. The DSA-110 is dedicated to the search of Fast Radio Bursts (FRB) [49], and operates a real-time data processing pipeline to detect the bursts 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 using a polyphase filterbank (PFB) [34]. The digitized coarse channels are then transferred over an Ethernet network to compute nodes in charge of forming beams within the field of view of the instrument, and searching these beams for FRBs using an incoherent de-dispersion search algorithm [50]. This pipeline is implemented using a ring buffer architecture [51], and the baseband data for all antennas and coarse channels for a duration of 2 seconds contained in the buffers is written to disk in case of detection for further processing.

We manually triggered the download of baseband data and formed an arbitrary beam to simulate the output of a



Fig. 8: The DSA-110 at OVRO.

Properties	Uplink	Downlink
Channel Bandwidth (MHz)	{10, 15, 20}	
Distance from BS to the telescope (km)	20	
Power control	Yes	No
Frame occupancy (%)	30 - 80	30 - 80
No. of UEs	5 (within a given cell radius)	

TABLE III: Different LTE parameters to produce unique RFI.

highly-directional single dish radio telescope for this work. We particularly focused on the coarse channel centered at 1420 MHz containing the galactic HI spectral line [52].

2) *LTE Frame generation:* The LTE RFI signal is generated for both uplink and downlink transmission according to the parameters in Table I and the specific transmission schemes laid out in §II, containing 400 frames (10 milliseconds each) to be comparable in duration to the astronomical dataset. Three different modulations: QPSK, QAM-16 and QAM-64 are used in the LTE signals. The LTE frames contain primary (PSS) and secondary (SSS) frame synchronization signals. Other additional components (e.g. control channels) for LTE signal are not addressed for simplicity. Figure 9 presents a structure of an LTE cell of radius r kilometers with uplink and downlink transmission along with the transmission links observed between the BS and the UEs and the geographically isolated radio telescope located at $R (> r)$ kilometers distance from the BS. For both downlink and uplink, UEs are uniformly distributed across the LTE cell with identical topology, but with different traffic patterns and UE density.

3) *LTE transmission:* LTE signal transmission is considered in three different scenarios: a) AWGN channel, b) flat fading channel, and c) realistic MIMO channel for both downlink and uplink. Investigating downlink transmission is a higher priority here as it is more likely to contaminate radio telescope data

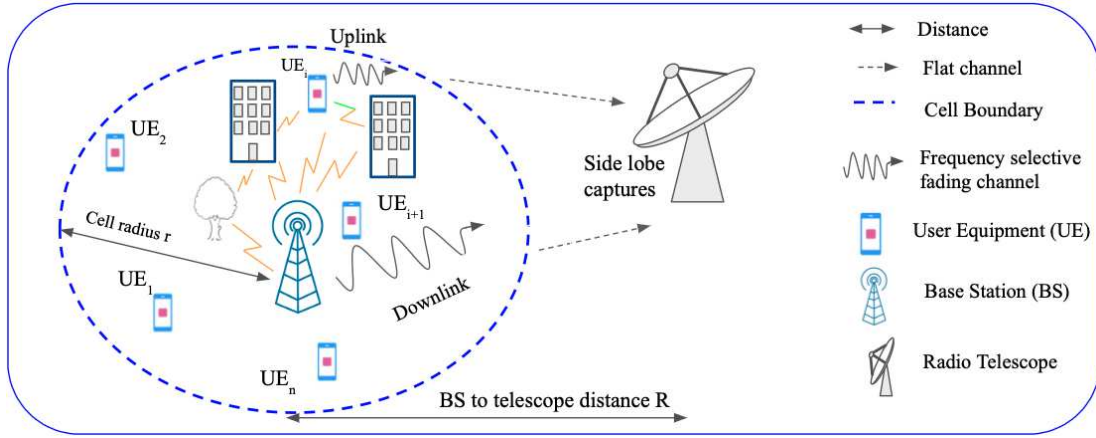


Fig. 9: LTE downlink and uplink RFI from a cell

Simulation scenario	Microcell	Macrocell
Carrier frequency	1420 MHz	1420 MHz
Bandwidth	20 MHz	20 MHz
Max. Cell radius	1 km	5 km*
BS antenna height	25m	25m
BS Num. Antenna	16	16
UE speed	3 km/h (pedestrian)	3km/h (pedestrian), 25 km/h (vehicle)
Num. UEs per cell sector	5	50 **

* radius is 10 km and UE drops are half per sector for rural macro-cell.

TABLE IV: Parameters used for geometry based stochastic LTE channel simulation that uses multiple large and small scale fading effects (Delay spread, Doppler, AoD, AoA, Rician K factor, Shadow fading etc.).

due to higher transmission power and directed beam patterns. As per 3GPP guidelines, Downlink transmit power is fixed at 46 dBm, equally spread across the transmission band and maximum uplink power is fixed at 23 dBm. We have implemented a simpler version of the power control protocol in uplink for validating the proposed system in AWGN and flat fading channel scenarios. In this protocol maximum UE transmit power (P_{max}^{UE}) = 23 dBm at cell boundary with transmit power of n_{th} UE (P_n^{UE}) defined as: $P_n^{UE} = P_{max}^{UE} \cdot (\frac{d_n}{d_{max}})^2$, where d_n, d_{max} are the distances of the n_{th} UE from the BS and the cell radius respectively. Simulations with the realistic MIMO channel models [53] in uplink and downlink are performed to verify the practical feasibility of the proposed apparatus for the following scenarios: 1) urban micro-cell, 2) bad urban macro-cell, 3) indoor to outdoor bchannel, 4) urban macro-cell, 5) bad urban macro-cell, 6) suburban macro-cell, 7) rural macro-cell (Nomenclature are from [53]). This method considers several large and small scale fading parameters e.g., frequency selective fading, delay spread, Doppler, angle of arrival (AoA), angle of departure (AoD), shadow fading, Rician K - factor distributions from which a geometry-based stochastic channel model is developed. Parameters used for this modeling purpose are enlisted in Table IV. Due to changing proximity of UEs to the BS, the uplink RFI typically experiences the LTE channel for a different propagation distance. Additionally, uplink RFI is characterized at BS where although at high power the signal includes some error due to propagation and frequency selective

fading. These elements introduce variation in accuracy of RFI characterization for uplink. But, due to low transmit power of uplink RFI and large propagation loss due to geographical isolation of radio-telescopes impact of such variation is limited at the telescope.

4) *RFI injection at Telescope*: The LTE signal is filtered post introduction of channel and propagation loss and side-lobe gains (RFI is primarily acquired via telescope side-lobes) to match the bandwidth of the telescope. The side-lobe attenuation depends on the angle of elevation of the antenna (The telescope under consideration only have a degree of freedom in elevation angle). The value of side-lobe gain can thus be anywhere between $10 \cdot \log_{10}(Ae * (pi * D\lambda)^2) = 15.1$ dBi to $-\infty$ dBi, depending on whether the RFI is received through the main lobe or any null in the directivity of the dish ($Ae = 0.7$, $D = 4.65$, $\lambda = 21$ cm). An arbitrary value of side-lobe gain of the telescope is chosen as -40 dBi at OVRO DSA-110 for simulation purposes. The BS location for the simulation is determined to be at 20 km (~ 92.46 dB propagation loss) linear distance from the telescope, which is very close to the estimated distance of the closest BS to DSA-110. This brings the LTE signal to a power level comparable to the astronomical signal as shown in Figure 2d. The composite signal spectrum from uplink RFI in Figure 2e, depicts overlap of RFI with different center frequencies, thus covering the guard bands of each other. The spectral plot is limited to the same bandwidth as the downlink rfi spectral plot to provide visual comparability. The signal is channelized (subchannels of telescope back-end, not to be confused with the wireless channel) both at the BS and the telescope to match the channels of the telescope - a shared static parameter shared as in Table II. The attenuated and noisy signal is added to the astronomical signal to simulate the RFI contaminated composite signal.

5) *Signal decomposition and reconstruction*: The LTE signal at BS (both uplink and downlink as discussed before) and the composite signal at the telescope are decomposed to obtain the eigenfunctions based on the formulation in §III. In our SSA based implementation of KLT, the window length is empirically estimated to be $L = 500$ as it satisfies the

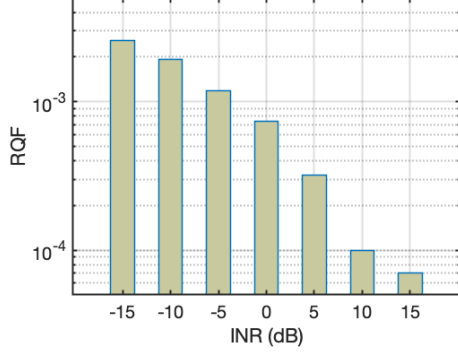


Fig. 10: Reconstruction loss of RFI with varying INR.

necessary condition $L > M$ where $M \approx 300$ as shown in Figure 7c. RFI cancellation and reconstruction with each set of wireless channel scenarios mentioned in §V-A3 is performed 100 times and the reconstruction quality is averaged over the 100 simulations to smooth out any outlier behavior. Figure 10 shows the reconstruction accuracy of a sample LTE signal with finite eigenspace [54] as a function of the interference to noise ratio (INR). We observe that the reconstruction accuracy improves with INR, meeting the RQF criterion presented in §IV at 10 dB INR. This is the primary motivation for characterizing RFI at the BS instead of at the telescope, as the telescope receives the RFI at a much lower INR due to propagation loss leading to improper eigenspace representation and erroneous cancellation. Continuous time synchronization between the source of RFI and the telescope is not required for the proposed collaborative apparatus. From Figure 9, arrival time of a downlink RFI at the telescope is $(t_p + R/c)$, t_p is the processing time for characterization of the RFI signal. With $R = 20$ km, the arrival time = $(66 \mu s + t_p) \mu s$. The same for a UE at a distance r from the BS becomes $((R+r)/c + t_p)$. partial synchronization is performed using standardized GPS clocks at both locations and appropriate delays are to be introduced at the telescope accordingly. Later we have shown in §V-B3, in presence of significant time synchronization error, the proposed model performance deteriorates to only a small extent.

B. Experimental Results

In this section, first, we have reported our findings using downlink signal as RFI propagated through AWGN channel that enables us to establish a baseline for the measure of accuracy of the proposed RFI cancellation method. We have additionally performed parameter space exploration with this downlink RFI signal to investigate the effects of interference to noise ratio (INR) at telescope, frame occupancy in LTE RFI and time synchronization error between RFI source and telescope. The results for these are presented in Figure 12 in purple. The constant parameters (2 out of 3 in each scenario) are set as follows - 1) INR (ratio of RFI and noise power): 5 dB, 2) frame occupancy: 70%, 3) Synchronization error: 0 second. The scale is marked as Φ_R at BS, which indicates Φ_R - the RFI eigenspace is obtained after characterization at BS. Later, we have reported impact of different LTE propagation

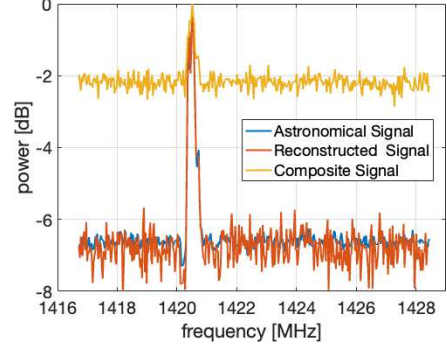


Fig. 11: Reconstructed space signal compared to the true astronomical and the composite signal.

scenarios and system design parameter KLT window-length L on reconstruction quality for both downlink and uplink RFI.

1) *Reconstruction of RFI-free astronomical signal*: we obtain, the rectified astronomical signal as shown in Figure 11 along with the true astronomical signal and the composite signal, following the signal reconstruction steps described in §III. The power levels are relative to the measured noise floor at the telescope (-174 dBm) as the baseline (-8 dB marker). We achieved an RQF of 4.132×10^{-4} for this reconstructed astronomical signal at an estimated interference to noise ratio (INR) of 5 dB at the telescope.

2) *Effect of Interference to Noise Ratio (INR)*: A low INR is expected at telescope since RFI is generally acquired by the telescope side-lobes. For the reconstructed astronomical signal shown in Figure 11, the INR is set at 5 dB as reported previously. In most practical cases, this will change based on the distance between the telescope and the source BS, sensitivity of the telescope front-end, and the varying side lobe attenuation due to changes in elevation angle. We have investigated the effect of INR varying from -10 dB to 20 dB on reconstruction quality using the RQF metric, shown in Figure 12a. RQF decreases with increasing INR as higher RFI power results in better characterization of the RFI facilitating better reconstruction quality [36]. RQF reaches the ITU-mandated threshold at 20 dB INR.

3) *Effect of synchronization error*: GPS clocks have maximum time-synchronization error of 30 nanoseconds as reported in [55]. Assuming appropriate transmission and processing delays are introduced 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 up to 320 nanoseconds is shown in Figure 12b. With increasing synchronization error, RQF increases upto 4.6×10^{-4} at 320 ns delay.

4) *Effect of varying spectral occupancy of RFI*: In a practical situation, LTE frames are not fully occupied and many resource blocks in a frame remain empty. We observe in Figure 12c that RQF changes with variation in spectral occupancy of the RFI. With lower spectral occupancy, the eigenspace of the RFI signal gets skewed due to sparsity in the spectrum which produces higher error while projecting the composite signal and increases RQF indicating a poor

reconstruction quality.

5) *RQF if RFI is decomposed at the telescope*: Accurate Characterization RFI at its source (BS for our use case), i.e at a higher power level, is the most crucial and unique component of the proposed method which leads to improved RFI mitigation and astronomical signal reconstruction. To confirm this claim, we attenuated the emulated LTE signal to power levels comparable to the RFI at the telescope and followed the steps from § V-A to analyse all three scenarios mentioned above. These are expected to provide results equivalent to RFI signals captured locally at the telescope site using a reference antenna for characterization and cancellation purposes. This method of locally characterizing RFI at telescope has been incorporated into several state of the art RFI mitigation techniques. We can observe the red plots in Figures 12a, 12b and 12c with the scales and axes marked as Φ_R at Telescope indicating Φ_R - the RFI eigenspace is obtained after characterization at telescope. The RQF is going up to 3×10^{-3} when RFI is characterized at telescope. Due to strong attenuation, the RFI characterization becomes erroneous leading to poor separability of RFI and astronomical signal subspaces. Thus the error margins are higher in orders of magnitude than the threshold calculated in §IV, rendering the cancellation apparatus questionable. This demonstrates that accurate RFI characterization at its source is paramount to the success of the proposed design.

6) *Effect of LTE practical channels*: In this section variations of RQF in different practical channel conditions are presented for both downlink and uplink RFI. Additionally, effects of varying design parameters of the RFI cancellation and reconstruction apparatus on them are analysed.

a) *Downlink*: Figure 13a shows the reconstruction quality for all eight different channel scenarios mentioned in §V-A4 with the previously described topology of the LTE cell, UEs and the radio-telescope in Figure 9. We observe, micro-cells with a smaller cell radius and macro-cells with less severe channels like suburban or rural scenarios have RQF closer to the ITU-mandated threshold, whereas RQF goes up to 6×10^{-4} for their counterparts with more severe channel and larger cell radius. In case of rural macro-cell, though channel effect and non-linearities are less severe, those are predominant in majority of the propagation path due to its much bigger cell radius. This leads to comparatively poor reconstruction quality with respect to other scenarios.

b) *Uplink*: Effect of practical channels in case of uplink RFI and corresponding RQFs are shown in Figure 13. The propagation scenarios are primarily grouped based on comparable cell radii and the reconstruction quality is presented with respect to the average distance from the BS in Figures 13b and 13c. This is relevant because uplink RFI is characterized at BS which is influenced by the signal propagation dependent on distance of corresponding UE from BS and the channel. This effect is limited in micro-cell scenarios as change in RQF is minimal with changing distance from BS. Although, different propagation scenarios have a varying impact as we observe RQF changes across different scenarios for both micro and macro cell propagation. Distance of UE from BS has a greater influence on the macro-cell propagation scenarios where a

visible change in RQF is observed with changing distance due to the higher order of magnitude in BS to UE distances compared to micro-cell.

7) *Impact of KLT window length*: In §III, we have emphasized the importance of accurate characterization of RFI and carefully choosing L for the success of the proposed method. Figure 14c shows a comparison of eigenspaces (eigenvalues without normalization) of downlink RFI without wireless channel effects (only propagation loss and noise), downlink RFI with channels effects, uplink RFI and astronomical signal. Both downlink scenarios have comparable eigenspace whereas uplink RFI has a similar trend as downlink with a lower magnitude up to 300th eigenvalue with a sharp drop in magnitude beyond 300. For all the LTE parameter variations evaluated in this work, we observed the number of strong eigenvalues of RFI lies in the range [272, 300]. This clearly indicates that RFI subspace has a dimension $M \approx 300$ and this range of values is unlikely to vary, as LTE is a modulated deterministic signal. Hence the necessary condition on L is $L > 300$. Figures 14a and 14b show the variation in reconstruction quality for different channel conditions with respect to KLT window length starting from 300 upto 600. We observe a visible improvement in RQF with increasing L , as more eigenfunctions are available after the orthogonal projection for reconstruction of the astronomical signal. The RQF reaches within an order of magnitude comparable to the prescribed value in §IV at $L = 500$. Thus we can safely declare $L \geq 500$ is the sufficient condition for choice of L . As L determines the overhead for collaboration, $L = 500$ is a fair choice as it is the smallest L satisfying both necessary and sufficient conditions. In an ideal scenario, where RFI and astronomical subspaces are perfectly separable, this trend of improvement in RQF is expected to continue with increasing L . But, in a practical scenario, where the subspaces are not perfectly separable, some signal energy will be lost due to projection. In that case, the RQF eventually saturates at some finite L .

8) *Impact on real-time collaboration*: The number of dominant eigenvalues or the number of eigenfunctions needed for an accurate representation of RFI signals drives the amount of overhead - data that is required to be shared between radio telescope and RFI source.

a) *Rate budget*: Assuming M eigenfunctions are necessary for characterization of RFI subspace, Φ_R will be a complex valued matrix of shape $(L \times M)$. Each complex entry will occupy at least 32 bits as a minimum of signed 16 bit depth for both real and imaginary components is necessary to maintain precision of shared eigenspace. As characterization is performed for individual fine channel (30.5 kHz) as described in §III, This contributes to an estimated overhead (r_d):

$$r_d = \frac{1}{\text{signal duration}} \times L \times M \times 32 \times \frac{B}{\text{Fine channel BW}} \text{bps.} \quad (12)$$

Plugging values from our simulation and signal captures, we get:

$$r_d = \frac{1}{4} \times 500 \times 300 \times 32 \times \frac{20 \text{ MHz}}{30.5 \text{ kHz}} \text{bps} = 786.67 \text{ Mbps.}$$

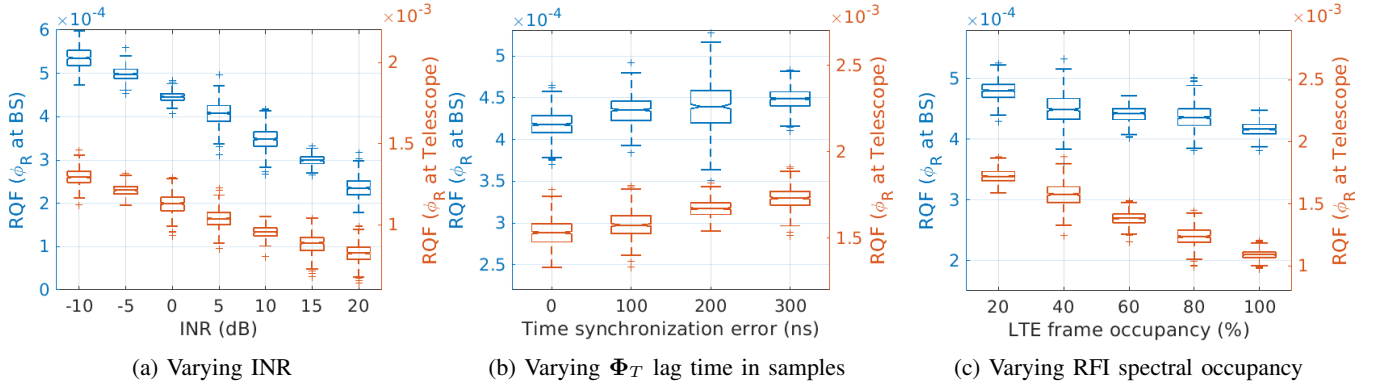


Fig. 12: Reconstruction Quality Factor (RQF) variation caused by different RFI parameters

The data rate in reality can be much lower as LTE signal is a modulated deterministic signal, its eigenspace does not change significantly across the fine channel bandwidth. Thus, it may suffice to share characterization of LTE across one channel or a few channels instead of all of them. Additionally, this can be on demand or periodic based on the use case. Assuming one fine channel characterization is shared for the entire bandwidth, the minimum data rate will be (~ 1.2 Mbps) for real-time processing and storage. This may necessitate some upgrades in software, hardware, and storage at the telescope. This data rate may increase with complex cell topology contributing to the RFI, different traffic patterns, and variation with time and distance. To alleviate the total overhead, the parameter set needs to be more concise and that pushes us to look at non-linear decomposition of signals, generating a set of smaller number non-linear basis functions that are able to accurately characterize the signal. We expect to address such issues in our future work.

VI. CONCLUSION

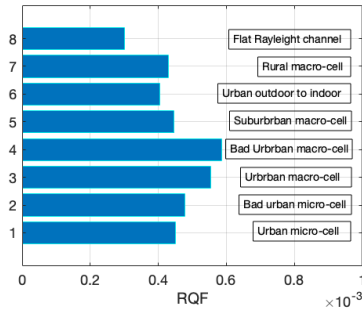
This work advances the literature on RFI mitigation for radio astronomy by sharing stochastic characterization of the RFI at its source, the cellular base station, with the telescope to cancel the incident RFI. The method has the potential to be deployed at an actual radio telescope, like DSA-110 at 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.

ACKNOWLEDGEMENT

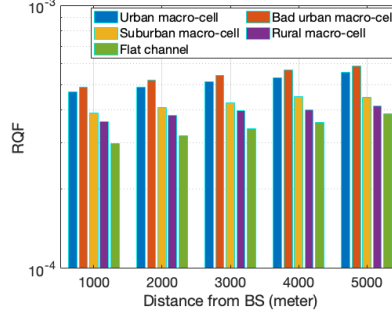
This work is funded by the National Science Foundation SWIFT Program (Award Number - 2128581).

REFERENCES

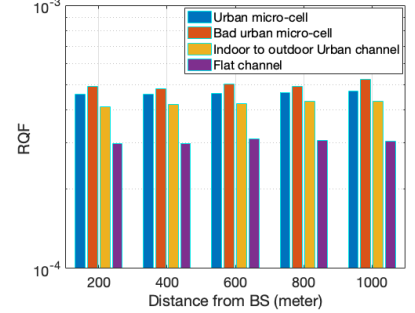
- [1] I. A. Organisation, *Dark and Quiet Skies for Science and Society*, 2020. [Online]. Available: <https://www.iau.org/static/publications/dqsk-ies-book-29-12-20.pdf>
- [2] National Academies of Sciences, Engineering, and Medicine, *A Strategy for Active Remote Sensing Amid Increased Demand for Radio Spectrum*. Washington, DC: The National Academies Press, 2015. [Online]. Available: <https://www.nap.edu/catalog/21729/a-strategy-for-active-remote-sensing-amid-increased-demand-for-radio-spectrum>
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] S. Zaroubi, "The epoch of reionization," in *The First Galaxies*. Springer Berlin Heidelberg, sep 2012, pp. 45–101. [Online]. Available: https://doi.org/10.1007%2F978-3-642-32362-1_2
- [9] 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>
- [10] The International Telecommunication Union - Radiocommunication Bureau, *Handbook on Radio Astronomy*. The International Telecommunication Union - Radiocommunication Bureau, 2013.
- [11] 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/CRAFh_andbook3.pdf
- [12] Federal Communication Commission, "2020 FCC Broadband Deployment Plan," *2020 BROADBAND DEPLOYMENT REPORT*, vol. 21, no. 1, pp. 1–9, 2020. [Online]. Available: <https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2020-broadband-deployment-report>
- [13] U. Rau, R. Selina, and A. Erickson, "Rfi mitigation for the ngvla: A cost-benefit analysis ngvla memo# 70," 2019.
- [14] "National Science Foundation-Funded Wireless Testbed on Rural Broadband," <https://advancedwireless.org/national-science-foundation-funde-d-pawr-program-selects-two-finalists-for-fourth-wireless-testbed/>
- [15] J. E. Velazco, L. Ledezma, J. Bowen, L. Samoska, M. Soriano, A. Akgiray, S. Weinreb, and J. Lazio, "Ultra-wideband low noise amplifiers for the next generation very large array," in *2019 IEEE Aerospace Conference*, 2019, pp. 1–6.
- [16] J. Shi and S. Weinreb, "Temperature compensated internal lna noise calibration source," *IEEE Microwave and Wireless Components Letters*, vol. 31, no. 8, pp. 1016–1019, 2021.



(a) Downlink Scenarios

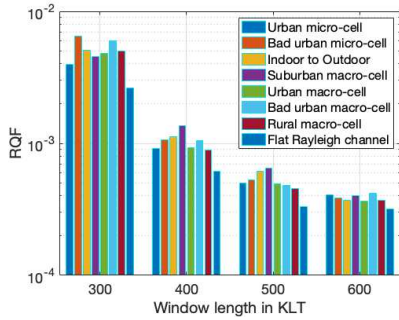


(b) Uplink Macro-cell Scenarios

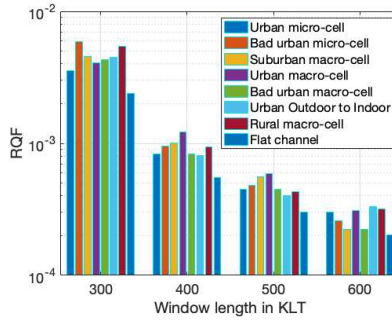


(c) Uplink Micro-cell Scenarios

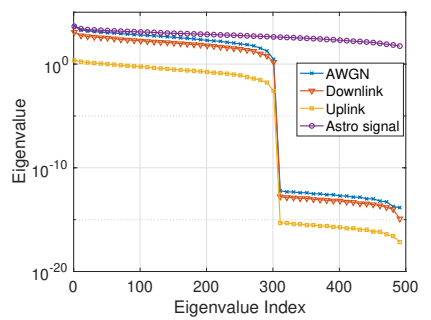
Fig. 13: Reconstruction Quality in presence of RFI from Downlink and Uplink Signal - Different Channel scenarios



(a) Reconstruction quality estimate over KLT window length - Uplink



(b) Reconstruction quality estimate over KLT window length - Downlink



(c) Eigenvalues - Astro, Downlink, Uplink, (AWGN+RFI)

Fig. 14: Effect of window length on Reconstruction Quality Factor

- [17] K. Jeganathan, A. Dunning, Y. S. Chung, M. Bourne, M. Bowen, S. Castillo, N. Carter, P. Doherty, D. George, D. Hayman, S. Mackay, L. Reilly, P. Roberts, P. Roush, S. Severs, K. Smart, R. Shaw, S. Smith, and J. Tuthill, "Ultra wideband (uwb) receiver for radio astronomy," in *2019 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, 2019, pp. 0343–0346.
- [18] 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.
- [19] 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.
- [20] C. Dumez-Viou, R. Weber, and P. Ravier, "Multi-level pre-correlation rfi flagging for real-time implementation on unibord," *Journal of Astronomical Instrumentation*, vol. 5, no. 04, p. 1641019, 2016.
- [21] 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.
- [22] 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.
- [23] 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.
- [24] 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.
- [25] C. J. Wolfaardt, "Machine learning approach to radio frequency interference (rfi) classification in radio astronomy," Ph.D. dissertation, Stellenbosch: Stellenbosch University, 2016.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] N. R. Council *et al.*, *Handbook of Frequency Allocations and Spectrum Protection for Scientific Uses*. National Academies Press, 2007.
- [31] M. Careem, S. Chakarabarty, A. Dutta, D. Saha, and G. Hellbourg, "Spectrum sharing via collaborative rfi cancellation for radio astronomy," in *2021 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2021, pp. 97–104.
- [32] T. L. Wilson, K. Rohlf, and S. Hüttemeister, *Tools of radio astronomy*. Springer, 2009, vol. 5.
- [33] 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.
- [34] 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.
- [35] G. Hellbourg, "Radio Frequency Interference spatial processing for modern radio telescopes," Theses, Université d'Orléans, Jan. 2014. [Online]. Available: <https://theses.hal.science/tel-01069394>
- [36] A.-J. Boonstra, "Radio frequency interference mitigation in radio astronomy," Ph.D. dissertation, Technische Universiteit Delft, 2005.
- [37] N. Golyandina and A. Zhigljavsky, *Singular Spectrum Analysis for Time Series*, 01 2013.
- [38] 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.

- [39] C. Maccone, *A simple introduction to the KLT and BAM-KLT*, 08 2012, pp. 411–448.
- [40] 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
- [41] 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.
- [42] 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.
- [43] A. M. Tomé, D. Malafaia, A. R. Teixeira, and E. W. Lang, “On the use of singular spectrum analysis,” 2018.
- [44] R. Vautard, P. Yiu, 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>
- [45] T. L. Wilson, K. Rohlf, and S. Hüttemeister, *Tools of Radio Astronomy*, 2013.
- [46] A. Szumski, “Karhunen loeve transform as an instrument to detect weak rf signals,” *InsideGNSS*, European Space Agency, Tech. Rep., 2011.
- [47] R. I.-R. RA.769-2, “Protection criteria used for radio astronomical measurements.”
- [48] [Online]. Available: <https://www.ovro.caltech.edu/>
- [49] E. Petroff, J. Hessels, and D. Lorimer, “Fast radio bursts,” *The Astronomy and Astrophysics Review*, vol. 27, no. 1, pp. 1–75, 2019.
- [50] G. Walsh and R. Lynch, “Searching for single pulses using heimdall,” in *American Astronomical Society Meeting Abstracts# 231*, vol. 231, 2018, pp. 243–04.
- [51] W. van Straten, A. Jameson, and S. Osłowski, “Psrdata: Distributed acquisition and data analysis for radio astronomy,” *Astrophysics Source Code Library*, pp. ascl–2110, 2021.
- [52] J. M. Dickey and F. J. Lockman, “Hi in the galaxy,” *Annual review of astronomy and astrophysics*, vol. 28, pp. 215–261, 1990.
- [53] P. Kyösti, J. Meinilä, L. Hentilä, X. Zhao, T. Jämsä, C. Schneider, M. Narandzić, M. Milojević, A. Hong, J. Ylitalo, V.-M. Holappa, M. Alattosava, R. Bultitude, Y. Jong, and T. Rautiainen, “Ist-4-027756 winner ii d1.1.2 v1.2 winner ii channel models,” *Inf. Soc. Technol.*, vol. 11, 02 2008.
- [54] 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>
- [55] “Gps.gov: Gps accuracy,” <https://www.gps.gov/systems/gps/performance/accuracy/#timing>, (Accessed on 10/15/2021).



Shuvam Chakraborty received the MS degree in 2022 at the Electrical & Computer Engineering from University at Albany and the BE degree in Electronics & Telecommunication Engineering from Jadavpur University, India, in 2018. He is pursuing a Ph.D. degree with the Department of Electrical & Computer Engineering, University at Albany SUNY, Albany, NY, USA, where he is a member of the Mobile Emerging Systems and Applications (MESA) Lab. His research interests lie at the crossroads of wireless communications, signal processing and

machine learning with a focus on interpretable ML for wireless systems and spectrum sharing and coexistence.



interference detection and mitigation, and system design for radio astronomy.



CA. His interests include learning based optimization of wireless systems, heterogeneous wireless networks, spectrum management and distributed wireless systems.



Dr. Dola Saha is an Associate Professor in the Department of Electrical & Computer Engineering at University at Albany, SUNY. She co-directs the Mobile Emerging Systems and Applications (MESA) Lab at UAlbany. She was a faculty fellow at Jet Propulsion Laboratory, Caltech, NASA in summer of 2022. She was a visiting faculty at the Air Force Research Laboratory in summers of 2020 and 2021. She is the Vice Chair of the IEEE ComSoc TCCN SIG for AI and Machine Learning in Security and has been appointed a member of the SUNY Innovations Policy Board. Prior to that, she was a Research Assistant Professor in the Department of Electrical & Computer Engineering at Rutgers University. Before that, she was a Researcher in the Mobile Communications and Networking group at NEC Laboratories America. She received the Masters and Doctorate degrees from the Department of Computer Science at the University of Colorado Boulder. She is the recipient of Google Anita Borg Scholarship for her academic credentials. Her research interests lie in the crossroads of Machine Learning for Wireless Communication, Wireless Security, Digital Communication, Wireless Networks, Wireless Signal Processing, and Architecture of Software Defined Radios with focus on systems design and practical evaluation.



Heterogeneous signal processing in high-speed Cloud-RAN and applied machine learning for interference cancellation in radio astronomy.

Dr. Gregory Hellbourg is Staff Scientist at the Cahill Center for Astronomy and Astrophysics at the California Institute of Technology (California, USA) and spectrum manager of the Owens Valley Radio Observatory (California, USA). Gregory will oversee the main decisions and act as project manager. He received his Engineering and M.Sc. degrees in Signal and Image Processing, and Ph.D. degree in Signal Processing and Automation, from the University of Orléans (France) in 2010 and 2014 respectively. Dr Hellbourg’s expertise lies in signal processing, radio

Maqsood Careem is pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University at Albany SUNY, Albany, NY, USA, where he is a member of the Mobile Emerging Systems and Applications (MESA) Lab. He received his MS.c degree in 2019 at the Electrical & Computer Engineering from University at Albany and the BS.c degree in Electrical and Electronic engineering from the University of Peradeniya, Sri Lanka, in 2014. He is currently a System Development Engineer II at Amazon Lab 126, in Sunnyvale,

Dr. Ayeek Dutta is currently an Assistant Professor in the Department of Electrical and Computer Engineering at University at Albany, SUNY. Previously, he was an Assistant Professor at the University of Kansas and Postdoctoral Research Associate at Princeton University and has PhD and MS degrees in Electrical Engineering from the University of Colorado Boulder. His research spans across various topics in the areas of generalization of deep learning for wireless communication, distributed and trustworthy spectrum sensing and policy enforcement,