



HAL
open science

RFI-DRUnet: Restoring dynamic spectra corrupted by radio frequency interference-Application to pulsar observations

Xiao Zhang, Ismaël Cognard, Nicolas Dobigeon

► **To cite this version:**

Xiao Zhang, Ismaël Cognard, Nicolas Dobigeon. RFI-DRUnet: Restoring dynamic spectra corrupted by radio frequency interference-Application to pulsar observations. *Astronomy and Computing*, 2024, 47, pp.100822. 10.1016/j.ascom.2024.100822 . hal-04537955

HAL Id: hal-04537955

<https://hal.science/hal-04537955>

Submitted on 8 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

RFI-DRUnet: Restoring dynamic spectra corrupted by radio frequency interference – Application to pulsar observations

Xiao Zhang^{a,c}, Ismaël Cognard^{a,b}, Nicolas Dobigeon^c

^aLaboratoire de Physique et Chimie de l'Environnement et de l'Espace (LPC2E), CNRS, Université d'Orléans, Orléans, 45071, France

^bStation de Radioastronomie de Nançay, Observatoire de Paris, PSL Research University, CNRS/INSU, Nançay, 18330, France

^cInstitut de Recherche en Informatique de Toulouse (IRIT), Toulouse INP, CNRS, University of Toulouse, Toulouse, 31000, France

Abstract

Radio frequency interference (RFI) have been an enduring concern in radio astronomy, particularly for the observations of pulsars which require high timing precision and data sensitivity. In most works of the literature, RFI mitigation has been formulated as a detection task that consists of localizing possible RFI in dynamic spectra. This strategy inevitably leads to a potential loss of information since parts of the signal identified as possibly RFI-corrupted are generally not considered in the subsequent data processing pipeline. Conversely, this work proposes to tackle RFI mitigation as a joint detection and restoration that allows parts of the dynamic spectrum affected by RFI to be not only identified but also recovered. The proposed supervised method relies on a deep convolutional network whose architecture inherits the performance reached by a recent yet popular image-denoising network. To train this network, a whole simulation framework is built to generate large data sets according to physics-inspired and statistical models of the pulsar signals and of the RFI. The relevance of the proposed approach is quantitatively assessed by conducting extensive experiments. In particular, the results show that the restored dynamic spectra are sufficiently reliable to estimate pulsar times-of-arrivals with an accuracy close to the one that would be obtained from RFI-free signals.

Keywords: Radio astronomy, pulsar, RFI mitigation, dynamic spectrum restoration, deep learning.

1. Introduction

Endpoint of a massive star evolution, a pulsar is a highly magnetized, rapidly rotating neutron star, which emits beams of radiation. Their primary interest lies in the fact that it is an extremely dense and compact object with a remarkably regular rotation period. Large decimetric radio telescopes are used to study pulsars and to time the most stable ones. This intrinsic extreme regularity can then be used to test gravitational theories (Kramer et al., 2021) or detect very low frequency gravitational waves (Agazie et al., 2023; Antoniadis et al., 2023; Reardon et al., 2023).

Low frequency telescopes, such as the dutch Low-Frequency Array (LOFAR) in the Netherlands (Stappers et al., 2011) or the New extension in Nançay upgrading LOFAR (NenuFAR) (Bondonno et al., 2021) can be used to extend the pulsar population or to better understand the effects of interstellar propagation, critical to get a reliable timing.

In the realm of pulsar observations, especially at low frequency, radio frequency interference (RFI) pose a formidable challenge since they may significantly degrade the quality of astronomical data. They are emitted by man-made sources such

as cell phones, Wi-Fi, communication satellites, and radar systems. This source diversity is reflected in the temporal and spectral ranges of the measurements that may be affected by interference, which makes their handling complex. Besides, RFI signals typically exhibit higher amplitudes and distinct distributions compared to astronomical signals.

These RFI corruptions can dramatically impair the study of celestial objects, especially pulsar timing which requires extremely high precision and sensitivity of the data. Thus the problem of RFI mitigation has already received considerable attention in the literature. Researchers have explored a variety of approaches, including traditional thresholding-based methods and more recent data-driven techniques capitalizing on recent advances in machine (deep) learning. However, existing RFI mitigation approaches face two main challenges. The first obstacle is a lack of ground truth, i.e., the absence of properly labeled data required to train models within supervised learning. Since manually labeling large data sets is infeasible, thresholding-based methods (Lazarus et al., 2016) are generally employed to identify RFI in real measurements, which is then considered as the labeled ground truth. However, following such a naive strategy, the most advanced RFI mitigation techniques will certainly not be able to perform better than the unsupervised methods resorted to building the training sets, as already highlighted by Berthereau (2023). The second issue is the subsequent loss of information imposed by most of the existing methods, mainly due to the way the mitigation task is generally formulated. Indeed, conventional thresholding-based

*Part of this work has been supported by the ANR-3IA Artificial, Natural Intelligence Toulouse Institute (ANITI) under grant agreement ANITI ANR-19-P3IA-0004 and by the CNRS through the 80PRIME Program.

Email addresses: xiao.zhang@cnrs-orleans.fr (Xiao Zhang),
icognard@cnrs-orleans.fr (Ismaël Cognard),
nicolas.dobigeon@enseeiht.fr (Nicolas Dobigeon)

and data-driven RFI mitigation methods cast RFI mitigation as a binary classification or segmentation problem. In other words, they only aim at identifying and labeling parts of the dynamic spectra corrupted by RFI, also referred to as RFI flagging. However, once the parts of the dynamic spectrum corrupted by RFI have been identified, they are generally discarded from any subsequent analysis in the posterior data processing pipeline. This leads to a possibly significant loss of information, depending on the spectral and temporal ranges of the measurements affected by interference.

This paper takes up the challenge of overcoming the two aforementioned issues. It proposes a new supervised RFI mitigation method preceded by a dedicated data generation framework that can be used to train this model. Indeed, to compensate for the difficulty of having access to accurately labeled ground truth data, this work first introduces a general framework to easily generate realistic dynamic spectra corrupted by RFI. Even if previous works have considered similar strategies of data generation, none of them have been specifically designed for pulsar observations (Akeret et al., 2017b; Asad et al., 2021; DeBoer et al., 2017). This framework leverages a model-based generation of pulsar and RFI signals separately. Thanks to its versatility, this framework can be instantiated to produce realistic measurements that would have been made in various observational setups. In particular, this paper implements this general framework to produce large data sets that mimic observations by the NenuFAR telescope.

These data sets are subsequently used to train, validate and test a new dedicated RFI mitigation method within a supervised framework. To go beyond simple RFI tagging, this paper draws a parallel between RFI mitigation and a ubiquitous task encountered in image processing, namely image denoising. Contrary to most of the alternatives proposed in the literature, this method formulates this task as a supervised restoration problem. Benefiting from the generation framework previously introduced, the proposed approach has the ambition not only to detect and remove the RFI from the dynamic spectra but also to recover plausible signal values in place of the corrupted ones. The model is chosen as a deep neural network whose architecture inherits the performance reached by a recent yet popular image-denoising model (Zhang et al., 2021).

The contributions reported in this paper can be summarized as follows: *i*) the problem of RFI mitigation is envisioned as a joint detection and restoration task, which opens the door to capitalize on a recently proposed denoising-oriented deep neural architecture, *ii*) to overcome the difficulty of accessing to large labeled training data set, a versatile framework to generate RFI-corrupted dynamic spectra is proposed and instantiated to produce simulated signals compatible with real-world pulsar observations performed by NenuFAR, *iii*) once trained on data sets generated following the aforementioned simulation protocol, the proposed deep network coined as *RFI denoising residual U-net* (RFI-DRUNet) is tested through an extensive set of numerical experiments to quantitatively assess its performance and *iv*) one shows that the adopted strategy allows pulsar time-of-arrival estimation to be efficiently conducted on dynamic spectra restored by RFI-DRUNet, reaching an accu-

racy close to the one obtained on RFI-free signals.

The remaining of the paper is organized as follows. Existing RFI mitigation methods are reviewed and discussed in Section 2. Section 3 presents a versatile framework to generate RFI-corrupted pulsar observation data. The problem of RFI mitigation is envisioned from a restoration perspective in Section 4. The architecture of the proposed deep neural network designed to restore dynamic spectra is also detailed. Section 5 describes the experimental setup followed to instantiate the generation framework within the context of pulsar observed by NenuFAR. The simulation parameters are specified, eight simulation scenarios are introduced and some implementation details regarding the network training are given. The experimental results are reported in Section 6, with respect to two different objectives, namely dynamic spectrum restoration and RFI detection. In particular, the performance of RFI-DRUNet in terms of RFI flagging is compared to those reached by state-of-the-art algorithms. Finally, in Section 7, an application of the proposed method to the estimation of pulsar time-of-arrival is presented. Section 8 concludes the paper. For the sake of reproducibility and to promote open science, the data and the codes associated with this work are freely available online¹.

2. Related works

In radio astronomy, numerous approaches have been proposed to mitigate the RFI during the post-correlation stage. These methods can be divided into two main categories, namely the parametric methods, and the data-driven methods.

Regarding the first type of approaches, Fridman and Baan (2001) draw an overview of various types of RFI and RFI mitigation methods. Some RFI identification methods relying on statistical analysis tools have been proposed by Fridman (2008), Bhat et al. (2005), and Winkel et al. (2007). In an earlier work, Maslakovic et al. (1996) applied a thresholding technique after modeling the temporal waveform signal thanks to a discrete wavelet transform. Besides, thresholding has also motivated the development of simple algorithms based on the general assumption that RFI are characterized by higher amplitudes than astronomical data. The popular cumulative sum method (CUSUM) initially proposed by Page (1954) in the context of statistical process control was first applied by Baan et al. (2004) to the RFI detection task. Subsequently Offringa et al. (2010a) and Offringa et al. (2012) have proposed several improvements referred to as VarThreshold, SumThreshold and AOFlagger. Notably, SumThreshold is a widely deployed algorithm for RFI removal in various current pipelines of radio telescopes because of its reliability and its efficiency (Offringa et al., 2010b; Peck and Fenech, 2013; Akeret et al., 2017b). Athreya (2009) has exploited the particular behavior of the fringe-stopped correlator output of an interferometer baseline in presence of RFI to remove spatially and temporally constant RFI sources. When dealing with multibeam receiver systems, Kocz et al. (2010) have applied spatial filtering to effectively identify and remove

¹<https://github.com/llxzhang/RFI-DRUNet>

RFI from the temporal signals. The technique relies on a singular value decomposition (SVD) of the empirical covariance matrix computed from the Fourier representations of the input signals. [Pen et al. \(2009\)](#) have also attempted to identify RFI with the help of SVD, as well as [Zhao et al. \(2013\)](#) using a principal component analysis. More recently, [Finlay et al. \(2023\)](#) have exploited the expected trajectories followed by RFI to tackle the problem of their removal jointly with the calibration task. Finally, it is worth noting that Coastguard ([Lazarus et al., 2016](#)) and Clfg ([Morello et al., 2019](#)) are two popular RFI mitigation algorithms specifically designed to remove RFI from pulsar data. They exploit the expected pulsar characteristics in combination with various common statistical tools.

Conversely, data-driven approaches attempt to learn the main characteristics or features of RFI from existing data sets. Once adjusted, these models are deployed to identify the RFI. Such techniques rely on conventional machine learning tools such as K-nearest neighbors, Gaussian mixture models, and random forest, as employed by [Mosiane et al. \(2017\)](#) and [Wolfaardt \(2016\)](#). More recently, the last decade has been marked by the advent of deep neural networks, initially to perform vision-oriented tasks and then extended in various application domains. In this context, [Akeret et al. \(2017a\)](#) first designed a specific deep convolutional architecture, namely U-Net, to formulate RFI identification as an image segmentation task. To improve the model capability, [Yang et al. \(2020\)](#) introduce residual blocks and batch normalization, leading to the so-called RFI-Net specifically designed to identify RFI in data provided by the FAST radio telescope. [Yan et al. \(2021\)](#) have investigated the relevance of *atrous* convolution by proposing AC-UNet. R-Net proposed by [Vafaei Sadr et al. \(2020\)](#) have demonstrated some robustness on simulated and real data by using transfer learning. [Chang et al. \(2023\)](#) propose DAARE, a stacked autoencoder model, to remove RFI from auroral kilometric radiation (AKR) spectrograms and to restore RFI-free astronomical data. It is worth noting that the methods discussed above have been proposed in a supervised framework, i.e., they rely on available labeled data sets to train the designed networks. For instance [Hamid et al. \(2022\)](#) have used the predictions provided by the parametric methods Coastguard and Clfg as ground truth to train PSRFINET for detection of the RFI in pulsar data. Other approaches have been developed to tackle the RFI mitigation problem in a semi- or non-supervised framework, i.e., when the data sets are not (or only partially) accompanied by labels. For instance, [Ghanney and Ajib \(2020\)](#) have compared the performance of the supervised method YOLO3 to those reached by an unsupervised method based on a convolutional autoencoder. [Mesarcik et al. \(2020\)](#) have adopted a convolution variational autoencoder (VAE) and a naive support vector machine classifier to project high-dimensional time-frequency data into a low-dimensional prescriptive space. [Mesarcik et al. \(2022\)](#) have introduced an unsupervised method coined as nearest latent neighbors (NLN) which relies on a generative adversarial model to detect and identify RFI without direct observation of the interference. Motivated by a reduction of the computational cost, [Kerrigan et al. \(2019\)](#) have proposed a network referred to as DFCN with comprehensible amplitude and phase informa-

tion. Further, [Li et al. \(2021\)](#) have proposed RFI-GAN based on a generative adversarial network (GAN) and [Vos et al. \(2019\)](#) have attempted to separate RFI signals from astronomical signals by using a source separation technique also based on a GAN architecture. [Wang et al. \(2020\)](#) have implemented pseudo-inverse learning autoencoders not only to remove RFI from pulsar data but also to restore the pulsar data possibly masked by RFI. Finally, [Saliwanchik and Slosar \(2022\)](#) have exploited the expected differences between RFI and astronomical signal with respect to their statistical representations and their resulting compressibility to propose a self-learning network able to remove RFI.

3. A framework for generating RFI-corrupted dynamic spectra

3.1. Observation model for dynamic spectra

NenuFAR (New extension in Nancay upgrading LOFAR) is a new radio telescope built at the Nancay Radio Observatory, which is designed to observe the largely unexpected frequency window from 10 to 85MHz. NenuFAR currently operates according to several modes for the observation of pulsars, including folded mode, single-pulse mode, waveform mode, and dynamic spectra mode. In this work, we propose to identify and remove RFI from data acquired in the dynamic spectra mode, which basically consists of the magnitude of the discrete Fourier transform (DFT) of the raw data. Adopting a discretized, windowed analysis, dynamic spectra can be represented as a time-frequency plane, whose time and frequency resolutions are directly determined by the parameters of the DFT. The signal $S(n, k)$ recorded in the n th frequency channel at the k th temporal bin is assumed to be decomposed as

$$S(n, k) = P(n, k) + R(n, k) + E(n, k) \quad (1)$$

where $P(n, k)$ is the pulsar signal, $R(n, k)$ is a possible RFI component and $E(n, k)$ stands for the system noise and any mismodeling, for $n \in \{1, \dots, N\}$ and $k \in \{1, \dots, K\}$. Within the particular applicative context of the work reported in this manuscript (i.e., NenuFAR-like tied-array observations), the model derived in Eq. (1) is expected to be a reliable approximation of the dynamic spectra. For other particular contexts, this model could be enriched to account for various instrumental and acquisition parameters. For instance, the simulator specifically developed for the Hydrogen Epoch of Reionization Array (HERA) and considered by [Mesarcik et al. \(2022\)](#) in the context of RFI mitigation offers the possibility of handling antenna cross-coupling ([Hera-Team, 2024](#)). An example of dynamic spectrum generated according to the model in Eq. (1) is depicted in Fig. 1, where the pulsar signature $P(n, k)$ appears as exponentially decreasing curves and RFI $R(n, k)$ take the form of vertical or horizontal lines and small clustered-dots. The remaining of this section will detail the procedures to generate each of these three components of the recorded signals. The notations used throughout this manuscript are summarized in Table 1.

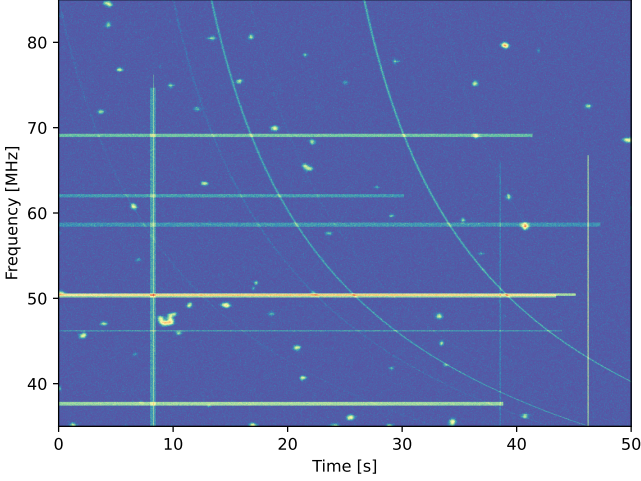


Figure 1: An example of the dynamic spectrum generated according to the proposed model in Eq. (1).

3.2. Simulation of the pulsar signal

3.2.1. A template-based model

The signal $P(n, k)$ associated with the pulsar is mainly characterized by the so-called integrated pulse profile denoted $A(n, k)$ hereafter. This profile can be considered as the "fingerprint" of a pulsar and is of primary interest to astronomers. Apart from possible variations in terms of energy, this signal generally shows fairly high stability in its shape along the observation at the same radio frequency but may exhibit small variations across the observation frequency (Lorimer and Kramer, 2004). To illustrate, Fig. 2 depicts profiles associated with the pulsar B1919+21 at different frequencies as observed by the NenuFAR telescope. It shows that the integrated pulse profile only slightly varies across frequencies

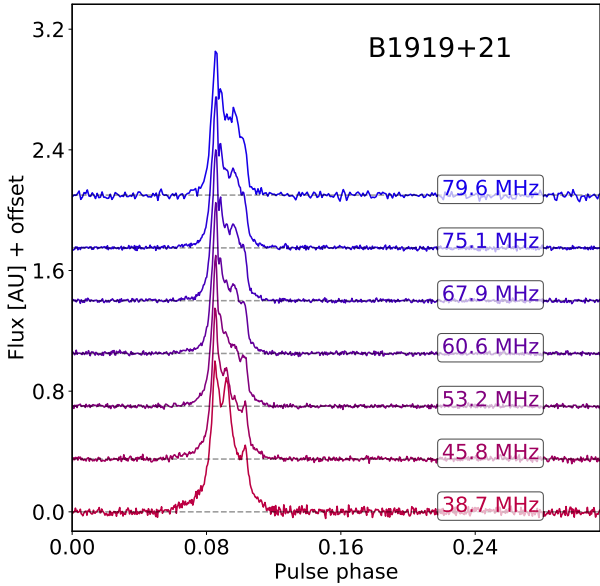


Figure 2: Pulse profiles of the pulsar B1919+21 recorded at different frequencies by NenuFAR. The total integration time for the observation is 19.5 minutes.

Time-frequency representation	
N	Number of spectral bins (channels)
n	Index of the spectral bin (channel)
K	Number of temporal bins
k	Index of the temporal bin
δt	Temporal resolution
δf	Spectral resolution
$S(n, k)$	Dynamic spectrum (full signal)
$P(n, k)$	Pulsar signal
$R(n, k)$	RFI signals
$E(n, k)$	System noise
Pulsar modeling	
D	Number of periods in the observation window
d	Index of the period in the observation window
ρ	Pulsar period
$A(n, k)$	Integrated profile
$\gamma_P(\cdot; \sigma^2)$	1D Gaussian kernel
SNR_d	Signal-to-noise ratio in the d th period
τ_n	Dispersion delay
DM	Dispersion measure
L	Number of Gaussian kernels
a_ℓ	Amplitude of the ℓ th Gaussian kernel
σ_ℓ	Width of the ℓ th Gaussian kernel
μ_ℓ	Location of the ℓ th Gaussian kernel
RFI modeling	
J	Number of RFI
SNR_j	Signal-to-noise ratio of the j th RFI
$M_j(n, k)$	Binary mask to locate the j th RFI
$\gamma_R(\cdot, \cdot; \sigma_T^2, \sigma_F^2)$	Separable 2D Gaussian kernel
α	Probability of occurrence of nbct and bbt RFI
β	Granularity parameter to generate nbt RFI

Table 1: Notations used to describe the RFI-corrupted dynamic spectra generated by the simulation protocol.

Without loss of generality of the method developed throughout this manuscript, the pulsar profile is assumed to be fully described by a unique template that does not vary across frequencies. This template is chosen as a weighted linear combination of L Gaussian-shaped components

$$A(n, k) = \sum_{\ell=1}^L a_\ell \gamma_P(k - \mu_\ell; \sigma_\ell^2) \quad (2)$$

where L is the number of Gaussian components composing the profile, a_ℓ , μ_ℓ , and σ_ℓ^2 stand for the amplitude, location and width of the ℓ th component, respectively, and $\gamma_P(t; \sigma^2) = \exp\left(-\frac{t^2}{2\sigma^2}\right)$. Observing the pulsar over D periods leads to a time-periodic pulsar signal written as

$$P(n, k) = \sum_{d=1}^D \text{SNR}_d \times A(n, k - \tau_n - d\rho) \quad (3)$$

where SNR_d adjusts the signal-to-noise ratio in the d th period, ρ is the pulsar period and τ_n is the frequency-varying delay resulting from the dispersion phenomenon. This quantity is discussed in the following paragraph.

3.2.2. Dispersion measurement

As radio pulses propagate through the interstellar medium and possibly the intergalactic medium, they are affected by a dispersive delay. The dynamic spectra mode operated by the radiotelescope can follow a so-called de-dispersion process to mitigate the impact of this delay through the frequency. However, for the sake of generality, the proposed simulated model for the pulsar signal is designed such that it accounts for this phenomenon. In this case, the dispersion measurement should be adjusted and, for frequency bin n , is given by

$$\tau_n = 4.15 \times \frac{\text{DM}}{n^2 \delta f^2 \delta t} \quad (4)$$

where δf and δt are the spectral and temporal resolutions of the dynamic spectrum. The dispersion measure (DM) is the electron column density through which the pulse has propagated.

3.3. Simulation of RFI

Within the framework of acquisitions operated by NenuFAR, RFI can be roughly classified into three distinct types according to their respective shape along the frequency and the time domains: *i*) narrow-band transient (nbt) RFI, *ii*) narrow-band continuous-time (nbct) RFI and *iii*) broad-band transient (bbt) RFI. To reflect this diversity in terms of RFI spatial and spectral patterns, the approach adopted in this work consists of decomposing the whole RFI signal $R(n, k)$ as the superimposition of J individual RFI signatures. Each signature is described by a unique 2-dimensional Gaussian-shaped template $\gamma_R(t, f; \sigma_T^2, \sigma_F^2) \triangleq \exp\left(-\frac{t^2}{2\sigma_T^2}\right) \exp\left(-\frac{f^2}{2\sigma_F^2}\right)$ whose variances σ_T^2 and σ_F^2 adjust the temporal and the spectral spread of the pattern, respectively. RFI is assumed to be not affected by the dispersion effect, i.e., $\text{DM} = 0 \text{ pc cm}^{-3}$. The whole RFI signal can be written as

$$R(n, k) = \sum_{j=1}^J \text{SNR}_j \times \gamma_R(n - n_j, k - k_j; \sigma_{n_j}^2, \sigma_{k_j}^2) \quad (5)$$

where SNR_j adjusts the power of the j th RFI, n_j and k_j locate the spectral and temporal positions of the center of the j th RFI whose spectral and temporal spreads are driven by $\sigma_{n_j}^2$ and $\sigma_{k_j}^2$, respectively. To randomly locate the RFI over the dynamic spectra, the RFI model in Eq. (5) can be conveniently rewritten by explicitly introducing binary masks that are randomly drawn with prescribed statistical characteristics to mimic the diversity of the time-frequency shapes of the RFI. It yields

$$R(n, k) = \sum_{j=1}^J \text{SNR}_j \times M_j(n, k) * \gamma_R(n, k; \sigma_{n_j}^2, \sigma_{k_j}^2) \quad (6)$$

where $*$ stands for the 2-dimensional convolution operator and the binary mask

$$M_j(n, k) = \delta(n - n_j, k - k_j) = \begin{cases} 1, & \text{if } n = n_j \text{ and } k = k_j \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

takes the value 1 in case of a RFI centered at the time instant k_j in the frequency bin n_j , and 0 otherwise. Two different approaches have been followed to randomly generate these masks, each associated with particular types of RFI. These generation procedures are discussed in what follows.

3.3.1. Spectrally and temporally extended RFI

Narrow-band continuous-time and broad-band transient RFI is two kinds of RFI that are ubiquitous in real observations. For instance, in observations made by NenuFAR, the 36-37MHz frequency band is often affected by such types of interference. Since these instances of RFI are generally mutually independent, it seems reasonable to assume that the entries of the corresponding masks $M_j(n, k)$ can be randomly generated according to a Bernoulli distribution with probability α , i.e.,

$$\mathbb{P}[M_j(n, k) = \epsilon] = \alpha^\epsilon (1 - \alpha)^{1-\epsilon}$$

with $\epsilon \in \{0, 1\}$. It is worth noting that $\alpha = \mathbb{E}[M(n, k)]$ is the probability of RFI occurrence in the dynamic spectra. Thus, this parameter directly adjusts the average number $\alpha N K$ of such spectrally and temporally extended (nbct and bbt) RFI.

3.3.2. Narrow-band transient RFI

To allow nbt RFI to affect the dynamic spectra in a clustered manner, the entries of the corresponding masks are not independently generated. Instead, they are assigned a Markov random field (MRF) which introduces structure correlation across the dynamic spectra. More precisely, this MRF is a multilevel logistic model (also known as an Ising model) and is defined as (Li, 2009)

$$\mathbb{P}[M_j(n, k) = \epsilon \mid \mathcal{M}_j(n, k)] \propto \alpha^\epsilon (1 - \alpha)^{1-\epsilon} \exp\left[\beta \sum_{m \in \mathcal{M}_j(n, k)} \delta(m - \epsilon)\right] \quad (8)$$

where $\epsilon \in \{0, 1\}$ and

$$\mathcal{M}_j(n, k) = \left\{ M_j(n-1, k), M_j(n+1, k), M_j(n, k-1), M_j(n, k+1) \right\}$$

denotes the set of neighbors of the mask entry $M_j(n, k)$ in the time-frequency plane according to a 4-order neighboring structure. In Eq. (8), the so-called granularity parameter β adjusts the correlation between neighboring entries in the mask. In particular, when $\beta = 0$, no correlation is imposed and the model in Eq. (8) reduces to the independent Bernoulli distributed model adopted for spectrally and temporally extended RFI (see Section 3.3.1). Simulating masks according to this MRF can be easily conducted thanks to Gibbs sampling (Li, 2009, Chap. 2).

3.4. Generation of simulated data sets

Thanks to the particular decomposed form of the proposed dynamic spectrum model in Eq. (1), the two main components associated with the pulsar signature and the RFI can be first generated independently and then combined to form the simulated

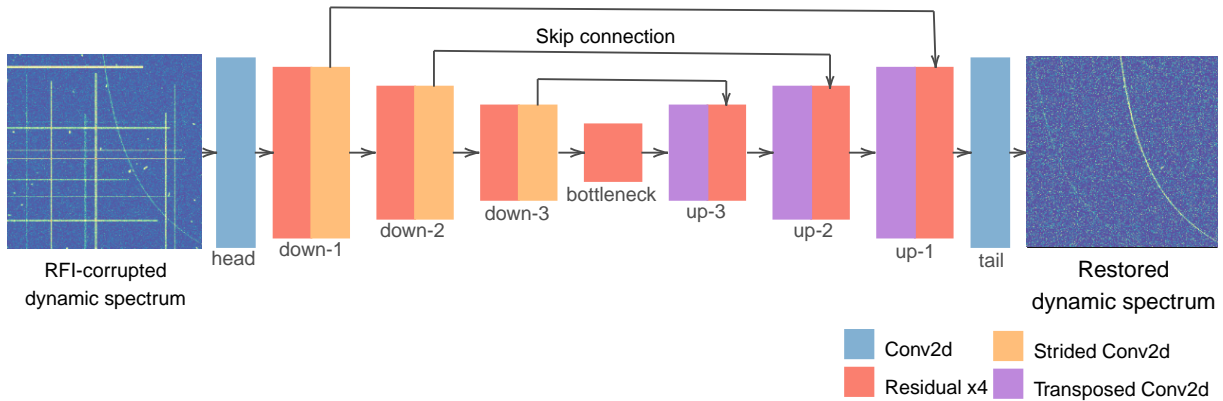


Figure 3: Architecture of the proposed RFI-DRUNet network. It takes as inputs RFI-corrupted dynamic spectra and provides as output restored (i.e., RFI-free) dynamic spectra. Details about the layers are provided in Table 2.

signals. For instance, a given generated pulsar signal $P(n, k)$ can be combined with several RFI signals $R(n, k)$, or reciprocally, to produce an extended set of dynamic spectra $S(n, k)$. Thus, in a nutshell, the proposed simulation framework consists in *i*) generating a pulsar database \mathcal{P} composed of pulsar signals, *ii*) generating an RFI database \mathcal{R} composed of RFI signals from three subsets, $\mathcal{R}^{\text{nb}}t$, $\mathcal{R}^{\text{nbct}}$ and \mathcal{R}^{bbt} , corresponding to three types of RFI, *iii*) randomly selecting a pulsar signal from the database \mathcal{P} and a RFI signal from the database \mathcal{R} and *iv*) combining them according to Eq. (1) to build one RFI-corrupted dynamic spectra of data set \mathcal{S} . The main benefit of this simulation strategy is to significantly reduce the amount of required computing. Indeed, let $|\mathcal{P}|$ and $|\mathcal{R}|$ denote the sizes of the pulsar and RFI databases, respectively. By coupling these two databases, the possible number of distinct generated dynamic spectra is $|\mathcal{S}| = |\mathcal{R}| \times |\mathcal{P}|$. In other words, only $|\mathcal{R}| + |\mathcal{P}|$ independent computations are necessary to generate $|\mathcal{S}|$ dynamic spectra, with $|\mathcal{R}| + |\mathcal{P}| \ll |\mathcal{S}|$.

4. Proposed RFI-DRUNet network for RFI mitigation

4.1. Formulating RFI mitigation as image restoration

As stated in Section 1, most of the methods proposed in the literature have formulated the RFI mitigation problem as a detection or segmentation/classification task. In other words, they only aim at localizing the time-frequency bins of the dynamic spectra possibly affected by RFI, distinguishing RFI-corrupted bins from RFI-free bins. Conversely, we propose to go beyond this crude RFI detection by formulating the problem of RFI mitigation as a restoration task. Our main rationale is that recent advances in machine learning offer the possibility of recovering clear (i.e., RFI-free) dynamic spectra from the RFI-corrupted measurements directly. To do so, we interpret the RFI mitigation objective as image denoising enounced in the time-frequency plane. Image denoising, a particular instance of image restoration, consists of recovering a clean image X from the degraded image Y corrupted by a specific (measurement) noise. When this noise is assumed to be additive and denoted

\tilde{E} , the clean and corrupted images are related through the observation model

$$Y = X + \tilde{E}. \quad (9)$$

It is worth noting that this generic formulation of the denoising task under the model in Eq. (9) perfectly matches the objective of recovering RFI-free spectra from dynamic spectra under the model in Eq. (1). To draw this connection, the observed image Y (resp. clean image X) in Eq. (9) can be associated with the measured dynamic spectra S (resp. RFI-free spectra $P + E$) in Eq. (1) while the noise term \tilde{E} writes $\tilde{E} = R$ in the context of RFI mitigation. In other words, RFI mitigation consists of restoring the dynamic spectrum corrupted by a particularly type of structured noise. As a benefit of translating RFI mitigation into image denoising, one can easily capitalize on recently proposed powerful deep convolutional neural networks such as those proposed by Zhang et al. (2017), Jiang et al. (2018) and Zhang et al. (2020). The architecture of the proposed network, hereafter referred to as RFI-DRUNet, is detailed in what follows.

4.2. RFI-DRUNet architecture

Image denoising is an archetypal image-to-image translation task, which can be efficiently addressed using deep networks with encoder-decoder architectures since they consistently deliver exceptional results. The encoder of the network is able to extract input image features at various levels while reducing the data size, and the decoder reconstructs data with the aid of features at different levels provided by so-called skip connections. In this work, we have customized the popular network referred to as DRUNet (Zhang et al., 2021) to be in agreement with the targeted RFI mitigation task. DRUNet is a deep convolutional network that follows an encoder-decoder architecture and utilizes a residual module that notably enhances the network capacity of feature extraction. Conventional DRUNet implementations are able to handle different noise levels in the data. Indeed, during the training stage, the noisy image and a corresponding noise map are jointly provided as the inputs of the network. For the RFI mitigation task considered in this

work, since interference to be removed from the dynamic spectra can be of any intensity level, such a flexibility is not required. Thus the RFI-DRUNet architecture is designed such that only RFI-corrupted dynamic spectra are provided as inputs during the training stage. In other words, no noise map is jointly provided as input to the proposed network during training.

More specifically, the architecture of the proposed RFI-DRUNet is sketched in Fig. 3. Its backbone consists of three parts, namely the encoder, the decoder, and the middle connection layer. The encoder and decoder networks are mirrored, i.e., each module in the encoder has an associated counterpart in the decoder and is connected by skip-connection. After passing a 64-channel convolutional (head) layer, the encoder comprises three modules, each consisting of 4 residual blocks and a stridden convolution (stride 2×2) as the downsampling layer. The middle connection layer is stacked by 4 residual blocks, which are followed by a decoder structurally symmetrical to the encoder whose modules are sequentially composed of a transposed convolution (stride 2×2 , padding 2×2) as the upsampling layer and 4 residual blocks. After the decoder, the network ends with a convolutional (tail) layer with one channel.

All residual blocks in the network are composed of 2 residual layers connected by a ReLU activation function. Furthermore, as in the conventional implementation of DRUNet, there is no activation function except in the residual blocks. From an overall perspective, the number of channels starts from 1 and increases to 64 after the first convolutional layer, then it is doubled with each downsampling layer and halved for each upsampling layer until 64 then becomes 1 through the last convolutional layer. The details of each layer of the network are summarized in Table 2.

5. Experimental framework

Extensive experiments have been conducted to assess the performance of the proposed RFI-DRUNet method. More precisely, the network designed in Section 4 will be trained on several synthetic data sets generated following the simulation framework outlined in Section 3. The parameters of the model are specified in Section 5.1 and the simulation scenarios as well as associated data sets are described in Section 5.2. Finally, implementation details are provided in Section 5.3.

5.1. Simulation parameters

Parameters of the time-frequency representation – The parameters of the dynamic spectra are chosen to match the main characteristics of the signals observed by NenuFAR. The dynamic spectra is assumed to be characterized by $N = 1024$ frequency channels ranging from 35MHz and 85MHz with a spectral resolution of $\delta f = 48.828\text{kHz}$ and $K = 1024$ temporal bins of resolution of $\delta t = 0.05\text{s}$.

Parameters of the pulsars – The signal of the pulsar $P(n, k)$ is fully described in Eq. (3) by its integrated profile $A(n, k)$, defined in Eq. (2) by the template $\gamma_P(\cdot, \cdot)$, and its periodization over the observation time. The values of the pulsar period

ρ and the dispersion factor DM are uniformly and randomly drawn over specific ranges chosen to mimic realistic signals. The power levels SNR_d ($d = 1, \dots, D$) of the pulsar defined on the D periods are selected according to a log-uniform rule. The number L of Gaussian defining the template is at most equal to 2, localization μ_ℓ and width σ_ℓ^2 of each Gaussian shape are uniformly drawn over pre-defined sets. The admissible ranges of the parameters are reported in Table 3.

Parameter	Notation	Value range
Period	ρ	(20, 40) [bins]
Dispersion	DM	(10, 40) [pc cm^{-3}]
Number of components	L	{1, 2}
Amplitude	a_ℓ	(0.2, 1)
Localization	μ_ℓ	(0, ρ)
Width	σ_ℓ^2	(0.01, 0.04)
Power	SNR_d	(0.01, 20)

Table 3: Parameters of the simulation associated with the pulsar signal.

Parameters of the RFI – The three types of RFI, namely nbt, bbt and nbct, are generated following the same simulation protocol. It consists in drawing binary masks according to a Bernoulli distribution (bbt and nbct RFI) or a Markov random field (nbt RFI). These binary masks are then convolved with a 2-dimensional separable Gaussian kernel whose spectral σ_F^2 and temporal and σ_T^2 spreading are randomly drawn over predefined sets chosen according to the type of generated RFI. The admissible ranges of the parameters involved in the Gaussian kernel as well as the parameters adjusting the statistical properties of the binary masks are reported in Table 4 for the three types of RFI.

Type of RFI	SNR_j	σ_F^2	σ_T^2	α	β
bbt	(1, 10)	(600, 1024)	(1, 10)	(0, 0.01)	N/A
nbct	(1, 10)	(1, 10)	(600, 1024)	(0, 0.01)	N/A
nbt	(0, 1)	(1, 11)	(1, 11)	0.8	40

Table 4: Parameters of the simulation associated with the RFI signal.

5.2. Simulation scenarios

As described in Section 3.4, the generated dynamic spectra are combinations of pulsar signals from data set \mathcal{P} and multiple RFI signals from data set \mathcal{R} . This simulation protocol is instantiated to produce three distinct data sets, namely the training set, the validation and the testing set. The training and validation sets share the same pulsar and RFI databases, but differ by the size of \mathcal{S} . The testing set is built from different pulsar and RFI sets. The sizes of these sets are reported in Table 5, where the RFI set \mathcal{R} contains three subsets matching three types of RFI, each of the same size.

According to this generation protocol, two experimental scenarios are considered to assess the robustness of the proposed method with respect to system noise. More precisely, so-called Scenario 1 (shortened as \mathcal{S}_1 hereafter) considers dynamic spectra composed of pulsar signal and RFI, i.e., free of system noise,

Layer	Block	Operation	Kernel size	Stride	Padding	Input size	Output Size
head	Conv2d	Conv2d	$3 \times 3 \times 64$	(1, 1)	(1, 1)	(64, 64, 1)	(64, 64, 64)
down-1	Residual $\times 4$	(Conv2d+ReLu+Conv2d) $\times 4$	$3 \times 3 \times 64$	(1, 1)	(1, 1)	(64, 64, 64)	(64, 64, 64)
	Strided Conv2d	Conv2d	$2 \times 2 \times 128$	(2, 2)	(1, 1)	(64, 64, 64)	(32, 32, 128)
down-2	Residual $\times 4$	(Conv2d+ReLu+Conv2d) $\times 4$	$3 \times 3 \times 128$	(1, 1)	(1, 1)	(32, 32, 128)	(32, 32, 128)
	Strided Conv2d	Conv2d	$2 \times 2 \times 256$	(2, 2)	(1, 1)	(32, 32, 128)	(16, 16, 256)
down-3	Residual $\times 4$	(Conv2d+ReLu+Conv2d) $\times 4$	$3 \times 3 \times 256$	(1, 1)	(1, 1)	(16, 16, 256)	(16, 16, 256)
	Strided Conv2d	Conv2d	$2 \times 2 \times 512$	(2, 2)	(1, 1)	(16, 16, 256)	(8, 8, 512)
bottleneck	Residual $\times 4$	(Conv2d+ReLu+Conv2d) $\times 4$	$3 \times 3 \times 512$	(1, 1)	(1, 1)	(8, 8, 512)	(8, 8, 512)
up-3	Transposed Conv2d	Conv2d	$3 \times 3 \times 256$	(2, 2)	(2, 2)	(8, 8, 512)	(16, 16, 256)
	Residual $\times 4$	(Conv2d+ReLu+Conv2d) $\times 4$	$3 \times 3 \times 256$	(1, 1)	(1, 1)	(16, 16, 256)	(16, 16, 256)
up-2	Transposed Conv2d	Conv2d	$3 \times 3 \times 128$	(2, 2)	(2, 2)	(16, 16, 256)	(32, 32, 128)
	Residual $\times 4$	(Conv2d+ReLu+Conv2d) $\times 4$	$3 \times 3 \times 128$	(1, 1)	(1, 1)	(32, 32, 128)	(32, 32, 128)
up-1	Transposed Conv2d	Conv2d	$3 \times 3 \times 64$	(2, 2)	(2, 2)	(32, 32, 128)	(64, 64, 64)
	Residual $\times 4$	(Conv2d+ReLu+Conv2d) $\times 4$	$3 \times 3 \times 64$	(1, 1)	(1, 1)	(64, 64, 64)	(64, 64, 64)
tail	Conv	Conv	$3 \times 3 \times 1$	(1, 1)	(1, 1)	(64, 64, 1)	(64, 64, 1)

Table 2: Details of the layers of the proposed RFI-DRUNet network. The size of input data and output data are $(H \times W \times C)$, where H , W and C stand for height, width and the number of channels, respectively. Sizes of the input can vary depending on the needs; an input data size of 64×64 is chosen here as an example.

	Size of \mathcal{P}	Size of \mathcal{R}	Size of \mathcal{S}
Training set	20	300	1800
Validation set	20	300	200
Testing set	10	60	200

Table 5: Size of the generated training, validation and testing sets.

$E = 0$ in Eq. (1). Conversely, Scenario 2 (shortened as S_2 hereafter) considers dynamic spectra generated according to the full model in Eq. (1), i.e., composed of a pulsar signal, RFI and a system noise modeled as an additive white Gaussian variable with variance $\sigma_E^2 = 1$, corresponding to an average value of signal-to-noise ratio (SNR) of -5.6dB over the test set. It is worth noting that Scenario 1 has been considered in addition to Scenario 2 to assess the impact of the system noise on the performance of the proposed restoration method. This scenario will be also of interest to evaluate the best restoration performance that could be reached by an oracle RFI detector (see Section 6.2). Moreover, to further investigate the model capacity to manage various types of RFI, for each scenario, we consider 4 cases that differ by the composition of the RFI set \mathcal{R} . These cases, denoted as C_A to C_D , are defined as follows

- C_A : nbt RFI (pulse-like RFI)
- C_B : nbt RFI + nbct RFI (narrow-band RFI)
- C_C : nbt RFI + bbt RFI (transient RFI)
- C_D : nbt RFI + nbct RFI + bbt RFI (all RFI types)

To summarize, a total of 8 simulation scenarios are considered, denoted as $S_{\square}C_{\Delta}$, with $\square \in \{1, 2\}$ and $\Delta \in \{A, B, C, D, \}$, depending on the presence/absence of system noise and depending on the type of RFI corrupting the pulsar signal. To illustrate, Fig. 4 displays one generated dynamic spectrum for each case in scenario S_2 .

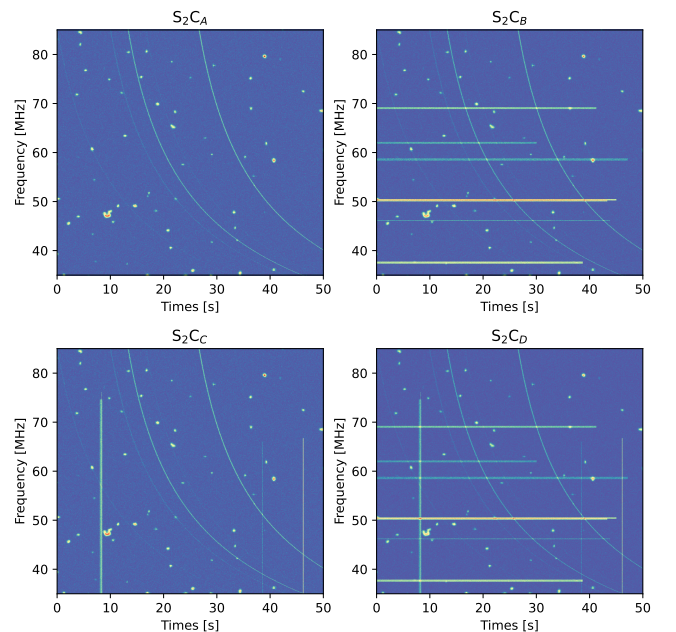


Figure 4: Examples of the dynamic spectrum generated according to the proposed protocol for the 4 cases of S_2

5.3. Implementation and training details

A distinct RFI-DRUNet model has been trained on each data set associated with the 8 simulation scenarios, leading to 8 instances of the proposed network. For all models, the training parameters are the same. During training, the generated dynamic spectrum from the data set \mathcal{S} has been randomly cropped to a patch of size 64×64 . Data augmentation is then applied, including flipping and rotating. The ℓ_1 -norm has been used to define the loss function. The Adam optimizer with a mini-batch size of 64 has been employed with a learning rate starting from 10^{-4} and halved after every 100k iterations until it reaches the value 5×10^{-7} . The training takes about 40 hours with 10000 epochs for an implementation in Pytorch and equipped with an Nvidia RTX 3080 GPU. This choice in terms of batch size and

number of epochs has been driven by technical constraints imposed by the available computing resources. A similar strategy has been adopted by Zhang et al. (2017).

6. Experimental results

The experiments conducted following the framework described in the previous section mainly follow two objectives. Firstly, they have been designed to demonstrate the feasibility of restoring RFI-corrupted dynamic spectra. Appropriate denoising-inspired quantitative performance measures will thus be detailed in Section 6.1. Secondly, we will show that the proposed restoration method can be easily simplified to perform only an RFI detection task. It consists of converting the outputs of the proposed RFI-DRUNet method into binary masks locating the RFI-corrupted time-frequency bins. This allows the proposed approach to be compared to state-of-the-art methods such as U-Net (Akeret et al., 2017a) and RFI-Net (Zhang et al., 2020) which have formulated RFI mitigation as a segmentation/classification task. When interpreted in this way, the methods can be compared with respect to detection-oriented figures-of-merit also defined in Section 6.1. The quantitative results are then reported and discussed in Sections 6.2 and 6.3 with respect to the two aforementioned tasks, namely restoration and detection. An illustration on a real observation is presented in Section 6.4. Finally, in light of these results, Section 6.5 discusses the relevance of the restoration paradigm with respect to the conventional RFI mitigation approaches, namely detection/flagging. It also mitigates these encouraging results by highlighting some cases of recovery failures experienced by RFI-DRUNet.

6.1. Quantitative figures-of-merit

As suggested above, the performance of the proposed method will be evaluated with respect to two main objectives, namely dynamic spectrum restoration and RFI detection (also referred to as RFI flagging in the literature). Regarding the restoration task, a conventional figure-of-merit encountered in image processing is the peak signal-to-noise ratio (PSNR) defined as

$$\text{PSNR} = 10 \log_{10} \frac{\max X^2}{\text{MSE}(X, \hat{X})} \quad (10)$$

where $\max X^2$ is the (squared) maximum value of the RFI-free signal, and $\text{MSE}(X, \hat{X})$ quantifies the mean square error between the ground truth RFI-free signal $X = S - R$ and its restored counterpart estimated by the algorithm \hat{X} . In particular, in the noise free case ($E = 0$, Scenario 1), the observation model in Eq. (1) simplifies to $S = P + R$ and this metric boils down to measuring the quality of the restored pulsar signal, i.e., $\hat{X} = \hat{P}$.

Regarding the detection task, since it can be formulated as a binary classification problem, the performance of the mitigation method can be evaluated by resorting to conventional classification-oriented figures of merit. It requires to first computing the so-called confusion matrix which summarizes the numbers of correct and bad classifications with respect to the

presence of a given target. For the RFI mitigation task of interest, we define the presence of an RFI as a positive instance and, conversely, an RFI-free signal as a negative instance. The confusion matrix reports the estimated probabilities of positive and negative samples being classified correctly or incorrectly, which leads to the definition of four indicators named *true positive* (TP), *true negative* (TN), *false positive* (FP) and *false negative* (FN). Once the confusion matrix has been computed, standard metrics can be derived, which include precision, recall, and F1 score.

Precision measures the proportion of predicted positives that are true positives

$$\text{prec} = \frac{\text{TP}}{\text{TP} + \text{FP}}.$$

Recall is the percentage of correctly identified RFI

$$\text{rec} = \frac{\text{TP}}{\text{TP} + \text{FN}}.$$

The F1 score, which strikes a balance between precision and recall, is calculated as the harmonic mean. This metric is especially valuable when analyzing data sets with unbalanced classes, i.e., when the number of samples in a given class is significantly larger than the number of samples in the other class, which is expected to be the case for moderately corrupted dynamic spectra. It is defined as

$$\text{F1} = \frac{2 \times \text{prec} \times \text{rec}}{\text{prec} + \text{rec}}.$$

Following the evaluation protocol also adopted by Mesarcik et al. (2022), the performance has been also evaluated in term of the area under the precision-recall curve (AUPRC), which quantifies the overall discriminatory ability of the compared models. This metric is also particularly valuable in scenarios with imbalanced datasets, where the number of negative instances significantly outweighs the positive ones. AUPRC emphasises the ability of the model in correctly identifying positive examples while maintaining accuracy. Finally, the evaluation of the model performance has been conducted in light of the receiver operating characteristics (ROC), which provides a comprehensive description of the performance of any detector faced to a binary hypothesis testing. ROC curves plot the true positive rate (TPR or probability of detection) as a function of the false positive rate (FPR or probability of false alarm). The area under the ROC (AUROC) sometimes referred to as the c-statistic (Hastie et al., 2009) has been considered as a figure-of-merit.

6.2. Restoration results

As described in Section 5.2, 8 simulation scenarios have been considered depending on the presence or absence of system noise and depending on the type of RFI corrupting the dynamic spectra. The proposed RFI-DRUNet model has been trained separately on 8 data sets associated with each of these scenarios. For each scenario S_1 or S_2 , the restoration performances of the RFI-DRUNet trained for a given case (C_A to C_D)

are evaluated not only on a testing data set generated according to the same case but also on data sets corresponding to the other cases. This will help to understand the impact of the type of RFI on the model performance and its possible limitations. With a slight abuse of notations, the models will be denoted as $S_{\square}C_{\Delta}$ by shortening the denomination of the data set they have been trained on, where the two indices $\square \in \{1, 2\}$ and $\Delta \in \{A, B, C, D, \}$ refer to the considered scenario and case, respectively.

Noise-free data set – Table 6 reports the results of the restoration in terms of average PSNR and standard deviations computed over the test set for the 4 models trained for the scenario S_1 . To appreciate the performance gain in term of restoration, this table also reports the PSNR computed from the data itself (first row). This is the cheapest strategy that would consist in not performing any restoration of the data. It also provides a quantitative proxy of the difficulty of the RFI-mitigation task (the lower SNR, the more difficult task). Finally, the second row of the table reports the PSNR computed from the noise-free pulsar signal recovered by an oracle detector that would be able to perfectly identify the RFI and would replace the corrupted time-frequency bins by zeros. These results show that the model S_1C_A performs quite differently, since it provides very good results only when tested on the data set S_1C_A . This can be explained by the limited variety of RFI (only of type nbt) in the training set. Other types of RFI are hardly identified and corrected by the algorithm. The models S_1C_B to S_1C_D perform well not only on the test sets corresponding to their training set but also show comparable restoration ability when handling the other cases. These three later cases contain pulse-like (i.e., nbt) RFI but differ substantially since they contain either nbct RFI or bbt RFI, or both. These two types of RFI share some shape similarities but differ in the direction of spreading. The data augmentation used during the training phase, which consists of rotations and flips, can explain this robustness to handle both nbct and bbt RFI when only one of them is present in the training set.

		Data set			
		S_1C_A	S_1C_B	S_1C_C	S_1C_D
Model					
Data		43.17 ± 0.93	35.02 ± 3.14	35.87 ± 2.99	31.92 ± 1.33
Oracle		59.27 ± 6.37	54.43 ± 6.11	54.24 ± 5.61	51.94 ± 5.51
RFI-DRUNet	S_1C_A	70.58 ± 10.86	56.59 ± 10.04	59.75 ± 8.96	53.82 ± 8.10
	S_1C_B	73.33 ± 8.70	72.38 ± 8.24	71.74 ± 8.30	70.80 ± 7.80
	S_1C_C	72.65 ± 9.00	70.94 ± 8.32	71.45 ± 8.65	69.85 ± 7.93
	S_1C_D	72.09 ± 9.23	71.44 ± 8.87	71.31 ± 8.92	70.72 ± 8.61

Table 6: Scenario 1: restoration performance in terms of average PSNR and standard deviations computed over the test data sets.

Noisy data sets – Similar findings can be drawn when consider-

ing the 4 models trained and tested on noisy data sets (scenario S_2), as shown in Table 7. Note that, in this case, the time-frequency bins identified by the oracle have been replaced by random values drawn according to the noise statistical model, i.e., $\mathcal{N}(0, \sigma_E^2)$. The overall values of PSNR are significantly lower than obtained with scenarios S_1 , owing to the presence of system noise. However, the PSNR values are sufficiently high to guarantee correctly restored dynamic spectra, in regard to standard restoration measures encountered in the image processing literature but also to the performance reached by the oracle detector.

		Data set			
		S_2C_A	S_2C_B	S_2C_C	S_2C_D
Model					
Data		43.17 ± 0.93	35.02 ± 3.14	35.87 ± 2.99	31.92 ± 1.33
Oracle		48.89 ± 0.97	44.60 ± 1.98	44.48 ± 1.68	42.27 ± 1.52
RFI-DRUNet	S_2C_A	59.43 ± 6.54	36.38 ± 4.88	37.77 ± 4.27	32.89 ± 1.69
	S_2C_B	60.15 ± 5.08	59.81 ± 4.91	59.88 ± 5.02	59.22 ± 4.73
	S_2C_C	60.36 ± 4.34	60.01 ± 4.40	60.12 ± 4.27	59.51 ± 4.25
	S_2C_D	60.41 ± 4.49	60.18 ± 4.55	60.19 ± 4.47	59.95 ± 4.53

Table 7: Scenario 2: restoration performance in terms of average PSNR and standard deviations computed over the test data sets.

Validation – The restoration performance (in terms of PSNR) as a function of the number of epochs during the validation stage is depicted in Fig. 5. This figure shows that the validation results are consistent with the results obtained during the testing stages. The model S_2C_A reaches a higher validation result compared to the other models since it uses less variety of RFI during the training stage. Conversely, the models S_2C_B and S_2C_C have almost identical validation results, and the model S_2C_D shows slightly lower validation results due to the fact that it has to handle with all types of RFI. In the sequel of the paper, for brevity and unless otherwise specified, only the model S_2C_D , which is expected to be more robust to any type of RFI, will be considered.

Out-of-distribution data sets – To assess the robustness of the proposed RFI-DRUNet to experimental conditions that go beyond those encountered during the training stage, we now investigate the generalization ability of the model when analyzing out-of-distribution (OOD) data. This model mismatch is envisioned with respect to two aspects. First, the performance of the model S_2C_D , trained on dynamic spectra corrupted by an instrumental noise of $\text{SNR} \approx -5\text{dB}$ ($\sigma_E^2 = 1$), is evaluated when restoring data corrupted by noises of higher levels, i.e., $\text{SNR} \in \{-8.6, -10.4, -12.6, -15.6\}$ [dB]. Second, since RFI emitters may produce interference of more complex temporal and spectral profiles than those prescribed in Section 3.3, instances of RFI with sinusoidal shapes are also considered. More precisely, instead of generating RFI signals with the Gaussian

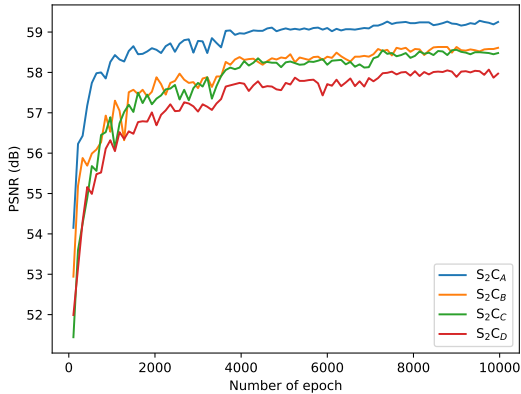


Figure 5: Validation: restoration performance (in terms of PSNR) as a function of the number of epochs.

		Simulation parameters		PSNR
OOD with respect to noise level	$\sigma_E^2 = 1$	SNR = -5.6dB	59.95	± 4.53
	$\sigma_E^2 = 2$	SNR = -8.6dB	59.24	± 0.81
	$\sigma_E^2 = 3$	SNR = -10.4dB	57.81	± 0.61
	$\sigma_E^2 = 5$	SNR = -12.6dB	54.86	± 1.56
	$\sigma_E^2 = 10$	SNR = -15.6dB	50.84	± 2.15
OOD with respect to RFI profile	nbct-sin RFI		59.90	± 3.22
	bbt-sin RFI		59.91	± 3.17
	nbct-sin RFI + bbt-sin RFI		59.59	± 3.11

Table 8: Out-of-distribution data sets: restoration performance in terms of average PSNR and standard deviations over the test data sets.

shapes described in Eq. (5), nbct and/or bbt RFI signals are generated according to an oscillating profile and included into the testing sets. These additional RFI profiles are denoted nbct-sin and bbt-sin, depending on their temporal and spectral spreading. These complementary experiments aim at enriching the diversity of interference patterns and the range of noise levels during testing. Note that the model S_2C_D evaluated in what follows has not been re-trained on these newly generated data sets. Thus this experimental protocol ensures a more thorough assessment of the model adaptability to unforeseen interference conditions. Table 8 reports the restoration results of the model when faced to these OOD data sets. It clearly appears that the restoration performance is weakly impacted by the noise level, up to a reasonable mismatch. Similarly, the presence of sinusoidal shaped RFI in the test sets barely affect the restoration performance, which confirms the robustness of the proposed method.

6.3. Detection results

The proposed method is now simplified to turn it into a simple RFI detector whose performance can be compared to those reached by competitive methods from the literature. To do so, a binary mask $\hat{M}(\cdot, \cdot)$ deciding the presence of possible RFI can

Method	prec	rec	F1-score	AUROC	AUPRC
RFI-DRUNet	0.972 ± 0.398	0.961 ± 0.0157	0.966 ± 0.024	0.995 ± 0.002	0.986 ± 0.016
U-Net	0.858 ± 0.048	0.987 ± 0.005	0.917 ± 0.029	0.991 ± 0.003	0.926 ± 0.027
RFI-Net	0.864 ± 0.047	0.988 ± 0.005	0.921 ± 0.029	0.987 ± 0.004	0.926 ± 0.026

Table 9: Detection performance of compared algorithms in terms of precision, recall, F1-score, AUROC and AUPRC. The results are reported with mean and standard deviation computed over the test data sets.

be easily computed as

$$\hat{M}(n, k) = \begin{cases} 1, & \text{if } |\hat{S}(n, k) - S(n, k)| > \eta \quad (\text{presence of RFI}) \\ 0, & \text{otherwise} \quad (\text{absence of RFI}) \end{cases}$$

where $\hat{S}(\cdot, \cdot)$ and $S(\cdot, \cdot)$ are the restored dynamic spectrum provided by RFI-DRUNet and the input dynamic spectrum, respectively, and η is a threshold balancing the probability of detection and the probability of false alarm. This threshold has been fixed to $\eta = 0.15$ such that it maximizes the F1-score of RFI-DRUNet to ease the comparisons to the other methods, as also done by Kerrigan et al. (2019) and Mesarcik et al. (2022). In particular, the performance of the proposed method is compared to the detection ability of two alternative deep learning-based RFI mitigation methods. The first considered network architecture is a U-Net, initially proposed to perform medical image segmentation tasks and later adapted for RFI mitigation tasks (Akeret et al., 2017a). The second compared model is RFI-Net (Yang et al., 2020), based on an encoder-decoder architecture with residual blocks and batch normalization. These two models have been trained on the data sets corresponding to scenario S_2C_D . More precisely, during the training stage, the binary masks defined by Eq. (7) and randomly drawn during the data generation process are given as output labels to the models.

As explained in Section 6.1, the methods are compared in terms of standard classification scores, namely precision, recall, F1-score, AUROC and AUPRC since they provide a comprehensive performance assessment of a binary classification model. These metrics are reported in Table 9. Although the primary objective of the proposed RFI-DRUNet model is not to detect RFI but rather to restore dynamic spectra, these results demonstrate that its ability to flag RFI is comparable to those of state-of-the-art methods specifically designed to perform this task. It is worth noting that, contrary to RFI-DRUNet, U-Net and RFI-Net are unable to restore RFI-free dynamic spectra.

To conduct a qualitative comparison of the results, outputs provided by the compared methods are depicted in Fig. 6, as well as the spectrum restored and the RFI identified by RFI-DRUNet, for a particular signal.

6.4. Illustration on a real observation

Since the proposed method has already demonstrated outstanding restoration and detection performance on the simulated data sets, we now illustrate this performance on a real observation data from the NenuFAR telescope. A dynamic spec-

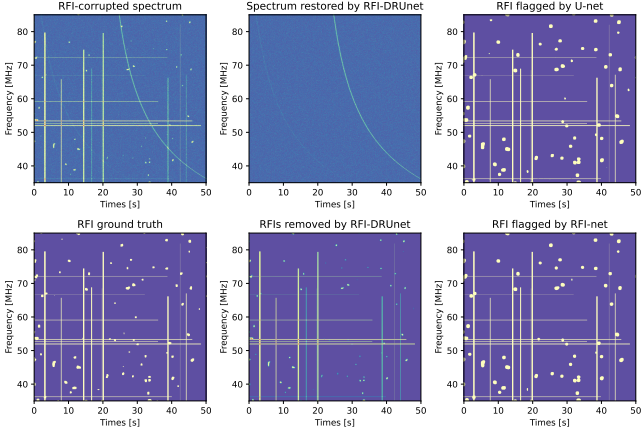


Figure 6: Visual comparison of the results for simulated spectrum from S_2C_D provided by the compared methods.

trum over the spectral range 74 – 80MHz of duration of 5s extracted from an observation for pulsar B1919+21 is considered, as depicted in Fig. 7 (top left panel). The compared models RFI-DRUNet, RFI-DRUNet and U-Net have been trained as in Section 6.3, i.e., using the synthetic data set S_2C_D generated following the protocol described in Section 5.2. This experiment somehow challenges the generalization ability of the compared models. Indeed here they are tested on a real signal whose content is expected to substantially depart from the simplifying modeling assumptions underlying the simulation framework described in Section 3. Figure 7 (top right panel) shows the restored dynamic spectrum achieved by RFI-DRUNet, as well as a comparison results of the RFI detected by RFI-DRUNet and U-Net (bottom panels). The results provided by RFI-Net are not reproduced in this manuscript since they are not of sufficient quality to be informative. The poor results exhibited by RFI-Net may be explained by its weak generalization ability. In these panels, RFI mostly appear as a form of nbct corruptions around 78MHz and persist from 0.2s to 1.5s. Both RFI-DRUNet and U-Net are proficient in detecting most of the RFI, but clearly, RFI-DRUNet provides higher accuracy, especially for the RFI around 78MHz between 0.2s and 1s. RFI-DRUNet also detects several instances of pulse-like RFI, while U-Net hardly detects any. This result confirms the ability of RFI-DRUNet to detect RFI with accuracy, whatever their shapes, and also to restore the corrupted dynamic spectra while preserving most of the signal of interest.

6.5. Discussion

Restoration vs. detection – As stated earlier, it is worth noting that U-Net and RFI-Net have been designed to achieve an RFI-flagging task, i.e., to identify and locate the time-frequency bins possibly affected by RFI in the dynamic spectra. The experimental results demonstrate exceptionally good performance achieved by these methods, which can be explained by the well-documented ability of convolutional neural networks to extract relevant features from 2-dimensional data, i.e., images in general and dynamic spectra in particular. Once RFI signals have been identified, several strategies can be envisioned to handle

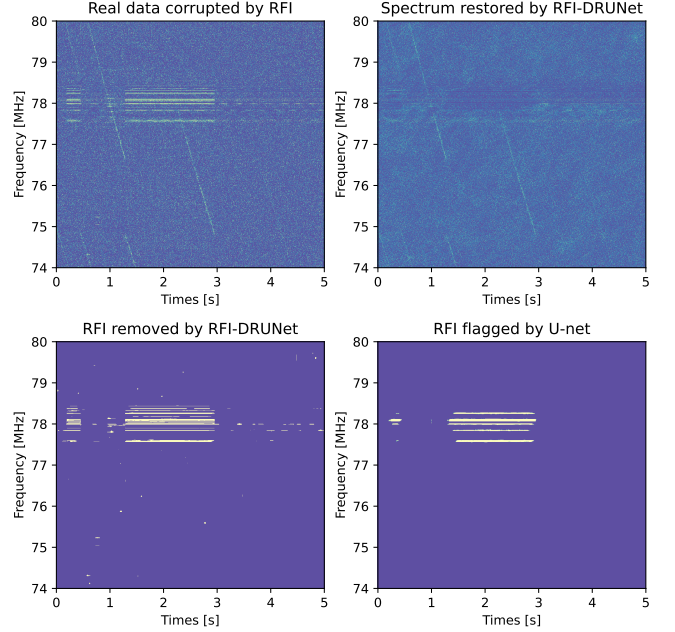


Figure 7: Visual comparison of the results for a real data set provided by RFI-DRUNet and U-Net.

them in a subsequent pipeline of astronomical data processing. The most common strategy consists of throwing away the measurements corresponding to the pulsar period corrupted by interference. Hence these pixel-level segmentation methods appear to be suboptimal since they conduct to the loss of significant information. Conversely, the proposed RFI-DRUNet method is able not only to correctly flag the RFI but also to restore the corrupted time-frequency bins with reliable signal values.

To illustrate this loss of information and also the relevance of restoring plausible values, one considers the strategy that would consist in replacing time-frequency bins corrupted by RFI by plausible values. Such a strategy has been already considered in Section 6.2 while reporting the restoration performance of the so-called oracle, assumed to perfectly know the RFI locations. Hereafter, the analysis goes one step further by considering realistic detectors, namely U-Net and RFI-Net, in addition to the oracle detector. In the noisy scenario (S_2C_D), these bins are filled with random values drawn according to the instrumental noise statistical model, $\mathcal{N}(0, \sigma_E^2)$. Table 10 reports the restoration results and also recalls the performance reached by the proposed RFI-DRUNet model. These results show that there is a significant gap in the quality of the data between the compared strategies.

Interestingly, while it has not been specifically designed to perform this task, RFI-DRUNet also reaches results comparable to those obtained by U-Net and RFI-net when conducting RFI detection/flagging. These significant improvements may come from several technical aspects adopted in this work. First, the architecture of RFI-DRUNet is different from those of the compared methods. Inherited from DRUNet, it is known to provide better results when facing a denoising task. This improvement does not systematically result from a higher complexity since

	Method		PSNR
	Detection	Filling value	
S_2C_D	Data		31.91 \pm 1.33
	RFI-DRUNet		59.95 \pm 4.53
	Oracle	$\mathcal{N}(0, \sigma_E^2)$	42.27 \pm 1.52
	U-Net	$\mathcal{N}(0, \sigma_E^2)$	41.64 \pm 1.33
	RFI-Net	$\mathcal{N}(0, \sigma_E^2)$	41.69 \pm 1.32

Table 10: Restoration performance of compared algorithms in terms of average PSNR and standard deviation computed over the test data sets.

RFI-DRUNet embeds a smaller number of network parameters to be adjusted when compared to the RFI-Net, which significantly reduces the computational burden during the training and testing stages. To illustrate, the number of parameters defining the compared models are reported in Table 11. Second, the pretext (denoising-like) task chosen to design the training loss function minimized by RFI-DRUNet is much more demanding than the ones adopted by U-Net and RFI-Net. When finally simplified, the proposed method is able to solve a easier task. Conversely, U-Net and RFI-Net may reach even better RFI flagging performance after adapting their architecture and their loss functions to be trained on a restoration task. In other words, the capacity of those models may be under-exploited by a too simple training strategy.

	RFI-DRUNet	RFI-Net	U-Net
# parameters ($\times 10^6$)	32.65	48.21	17.26

Table 11: Number of parameters of the compared models.

Failure situation – To provide a fair analysis of the potential brought by the proposed method, one finally wants to point out the fact that RFI-DRUNet may behave quite poorly in some particular situations. Indeed, a careful (empirical) inspection of the restoration results obtained by the network shows that some restoration scores are significantly lower than the overall score averaged over the whole data set. Figure 8 shows one archetypal example of failure. It clearly appears that the method erroneously treats the part of the pulsar signal as a possible RFI signal and tends to decrease its amplitude. This phenomenon typically arises at high frequencies where the pulsar signal is slightly affected by dispersion. The pulsar signal is then almost vertical and has a shape that is very similar to nbt RFI, i.e., temporally localized but spectrally spread interference.

7. Application to estimation of pulsar TOAs

Pulsars have a highly stable period of rotation, so the signals received from their emissions have a consistent period as well. The study of pulsars by precisely measuring their so-called times-of-arrival (TOAs) is known as chronometry. By investigating the pulsar TOAs, one can deduce various physical parameters intrinsic to pulsars and other astronomical quantities associated with various problems such as cosmic clocks and the

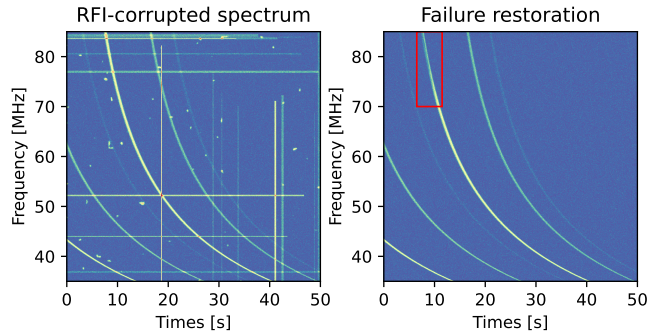


Figure 8: Example of a failure situation from data set corresponding to scenario S_2C_D . The amplitude of the part of the pulsar signal (framed in the red box) has been incorrectly reduced by RFI-DRUNet.

detection of gravitational waves. In what follows, we show that, thanks to its ability to detect RFI and restore dynamic spectra, the proposed RFI-DRUNet method is able to contribute to the improvement of the TOA estimation.

To estimate the pulsar TOA, the profile of a pulsar observed over a short time observation interval is generally assumed to deviate only slightly from the shape computed by integrating over a long period of time, hereafter referred to as the pulsar template. Leveraging this assumption, the observed signal pulsar $P(n, k)$ can be modeled as a noisy, scaled, and temporally shifted version of the template signal denoted $\tilde{P}(n, k)$, i.e.,

$$P(n, k) = u + v\tilde{P}(n, k - \Delta t) + E(n, k)$$

where u is an amplitude offset, b is a scaling factor, Δt is the temporal shift between the observed signal and the template, and $E(k, n)$ is a term accounting for modeling errors. Taylor (1992) introduced a least squares method to estimate the offset Δt between the observed signal and the template signal. The optimization problem is formulated in the Fourier domain. It provides an estimation with an accuracy of $0.1\delta t$, which is better than what can be obtained in the time domain. To illustrate the relevance of the proposed RFI-DRUNet method, we consider synthetic signals generated according to the scenario S_2C_D with a pulsar template $\tilde{P}(\cdot, \cdot)$ fully described by Eq. (2) and a pulsar period fixed to $\rho = 64$ [bins]. Then Taylor’s estimation method was applied to the RFI-free signal, the RFI-corrupted signal and the signal restored by RFI-DRUNet. In this controlled experimental setup, the true value of Δt is perfectly known. Thus the accuracy of the estimation conducted on the 3 types of signals can be qualitatively assessed by computing a mean square error (MSE) and a mean absolute deviation (MAE).

It is worth noting that Taylor’s method grants its TOA estimation with an uncertainty measure denoted $\sigma_{\Delta t}$. This uncertainty is related to the difficulty of the estimation task and, for instance, is driven by the SNR level: lower the SNR, higher the uncertainty. This finding is illustrated in Fig. 9 where each dot is associated to a particular signal generated according to the scenario S_2C_D for a wider range of SNR (from 10^{-1} to 10^3). Each signal is characterized by its SNR (x-axis) and the uncertainty measure provided by the estimation method (y-axis). In this log-log scatter plot, the uncertainty is clearly (in-

versely) proportional to the SNR. As reported in Section 5.1, the maximum SNR value in the generated data set S_2C_D is 20, which empirically corresponds to an uncertainty measure lower-bounded by $\sigma_{\Delta t} = 0.05$. This complementary information provided by Taylor’s method can be used to throw away all estimates accompanied by an uncertainty measure higher than a given threshold since they are considered as unreliable.

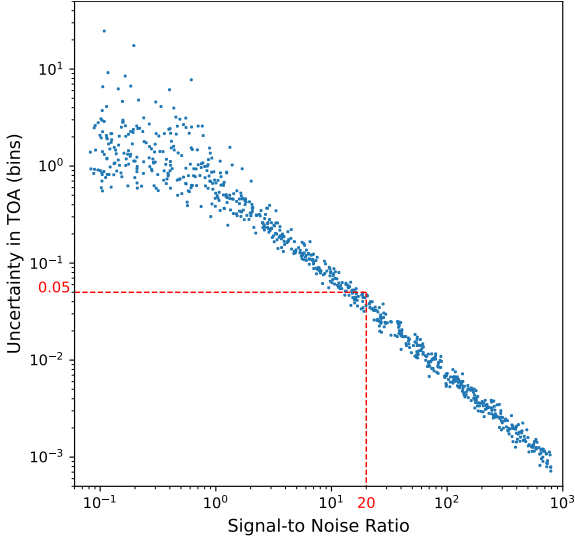


Figure 9: Scatter plot of the TOA estimation uncertainty vs. SNR. The red dashed lines point out the maximum SNR value which is 20, considered in Scenario S_2C_D and the corresponding uncertainty value adjusted by empirical linear regression (in the log-log space).

Figure 10 depicts the estimation errors (in terms of MSE and MAE) as functions of the uncertainty threshold. They have been obtained when the Taylor’s method has been applied to the RFI-free signals, the RFI-corrupted signals and the signals restored by RFI-DRUNet. As a follow-up of the discussion in Section 6.5, the Taylor’s method has been applied on the signals retrieved by three alternative approaches: RFI signals are located/flagged by the oracle detector, U-Net or RFI-Net, and the time-frequency bins identified as corrupted are replaced by random values. As expected, the lowest errors are obtained when estimating the TOAs from the RFI-free dynamic spectra. When the estimation is conducted on RFI-corrupted dynamic spectra, it yields a large discrepancy with the theoretical TOA, which is quite understandable since the pulsar signal is exposed to RFI contamination. The errors obtained from the dynamic spectra restored by RFI-DRUNet are close to those obtained by the RFI-free spectrum. This confirms that the proposed method significantly improves the estimation of the pulsar TOA when RFI-corrupted dynamic spectra have been restored beforehand. U-Net and RFI-Net both exhibited errors higher than those of the the oracle method. It is also important to highlight that the three alternative methods result in significantly larger errors compared to the TOA estimations derived from the RFI-corrupted signal. Again, this finding supports the idea that a substantial loss of information occurs when flagging and replacing corrupted bins with random values. On the contrary, the proposed RFI-DRUNet restoration model consistently

provides reliable estimations.

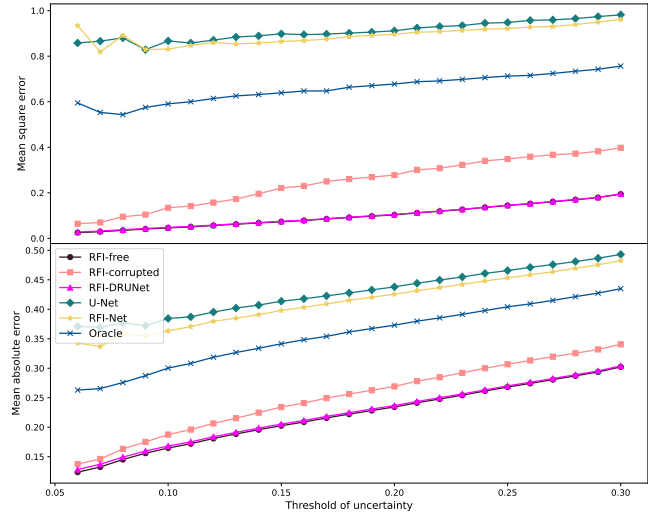


Figure 10: TOA estimation errors in terms of MSE (top) and MAE (bottom) as functions of the uncertainty threshold obtained from RFI-free (blue line), RFI-corrupted (orange line) and restored (green line) dynamic spectra.

8. Summary and conclusions

This paper formulated the problem of radio frequency interference mitigation (RFI) as a restoration task, to go beyond conventional approaches which aimed only at detecting and localizing RFI in dynamic spectra. To conduct this task, a new deep neural network, coined as RFI-DRUNet, was designed by leveraging and customizing a popular network proposed in the computer vision literature. To train this new model, a whole framework was designed to produce simulated RFI-free and corresponding RFI-corrupted dynamic spectra. This framework was instantiated in the specific context of pulsar observations and relied on physics-inspired and statistical models of the pulsar signals and of the RFI. The relevance of the approach adopted in this paper was assessed thanks to an extensive set of numerical experiments which demonstrated the ability of RFI-DRUNet not only to identify RFI but also to restore the corrupted dynamic spectra. The interest of the method was illustrated by monitoring the expected gain in the accuracy reached when estimating pulsar time-of-arrivals. Future works include taking the phase information into account during the restoration process. They will be also devoted to the compression of the deep network to reduce its computational complexity, with the aim of its integration into a real-time processing chain.

Acknowledgements

The Nançay Radio Observatory is operated by the Paris Observatory, associated with the French Centre National de la Recherche Scientifique (CNRS), and partially supported by the Région Centre, France. The development of NenuFAR has been supported by staff and funding from Station de Radioastronomie de Nançay, CNRS-INSU, Observatoire de Paris-PSL,

Université d'Orléans, Observatoire des Sciences de l'Univers en Région Centre, Région Centre-Val de Loire, DIM-ACAV and DIM-ACAV+ of Région Ile-de-France, Agence Nationale de la Recherche. The authors warmly acknowledge Dr. Louis Bondonneau for his help to produce the plots in Figure 2.

References

- Agazie, G., Anumalapudi, A., Archibald, A.M., Arzoumanian, Z., Baier, J., Baker, P.T., Bécsy, B., Blecha, L., Brazier, A., Brook, P.R., Burke-Spolaor, S., Burnette, R., Case, R., Casey-Clyde, J.A., Charisi, M., Chatterjee, S., Cohen, T., Cordes, J.M., Cornish, N.J., Crawford, F., Cromartie, H.T., Crowter, K., DeCesar, M.E., DeGan, D., Demorest, P.B., Dolch, T., Drachler, B., Ferrara, E.C., Fiore, W., Fonseca, E., Freedman, G.E., Garver-Daniels, N., Gentile, P.A., Glaser, J., Good, D.C., Gültekin, K., Hazboun, J.S., Jennings, R.J., Johnson, A.D., Jones, M.L., Kaiser, A.R., Kaplan, D.L., Kelley, L.Z., Kerr, M., Key, J.S., Laal, N., Lam, M.T., Lamb, W.G., Lazio, T.J.W., Lewandowska, N., Liu, T., Lorimer, D.R., Luo, J., Lynch, R.S., Ma, C.P., Madison, D.R., McEwen, A., McKee, J.W., McLaughlin, M.A., McMann, N., Meyers, B.W., Mingarelli, C.M.F., Mitridate, A., Natarajan, P., Ng, C., Nice, D.J., Ocker, S.K., Olum, K.D., Pennucci, T.T., Perera, B.B.P., Pol, N.S., Radovan, H.A., Ransom, S.M., Ray, P.S., Romano, J.D., Saffer, A., Sardesai, S.C., Schmiedekamp, A., Schmiedekamp, C., Schmitz, K., Shapiro-Albert, B.J., Siemens, X., Simon, J., Siwek, M.S., Stairs, I.H., Stinebring, D.R., Stovall, K., Sun, J.P., Susobhanan, A., Swiggum, J.K., Taylor, J.A., Taylor, S.R., Turner, E., Unal, C., Vallisneri, M., Vigeland, S.J., Wahl, H.M., Witt, C.A., Young, O., 2023. The NANOGrav 15-year data set: Search for Transverse Polarization Modes in the Gravitational-Wave Background. *arXiv e-prints* [2310.12138](#).
- Akeret, J., Chang, C., Lucchi, A., Refregier, A., 2017a. Radio frequency interference mitigation using deep convolutional neural networks. *Astronomy and Computing* 18, 35–39.
- Akeret, J., Seehars, S., Chang, C., Monstein, C., Amara, A., Refregier, A., 2017b. Hide & seek: End-to-end packages to simulate and process radio survey data. *Astronomy and Computing* 18, 8–17.
- Antoniadis, J., Arumugam, P., Arumugam, S., Babak, S., Bagchi, M., Nielsen, A.S.B., Bassa, C.G., Bathula, A., Berthereau, A., Bonetti, M., Bortolas, E., Brook, P.R., Burgay, M., Caballero, R.N., Chalumeau, A., Champion, D.J., Chanlaridis, S., Chen, S., Cognard, I., Dandapat, S., Deb, D., Desai, S., Desvignes, G., Dhanda-Batra, N., Dwivedi, C., Falxa, M., Ferdman, R.D., Franchini, A., Gair, J.R., Goncharov, B., Gopakumar, A., Graikou, E., Griefmeier, J.M., Guillemot, L., Guo, Y.J., Gupta, Y., Hisano, S., Hu, H., Iraci, F., Izquierdo-Villalba, D., Jang, J., Jawor, J., Janssen, G.H., Jessner, A., Joshi, B.C., Kareem, F., Karuppusamy, R., Keane, E.F., Keith, M.J., Kharbanda, D., Kikunaga, T., Kolhe, N., Kramer, M., Krishnakumar, M.A., Lackeos, K., Lee, K.J., Liu, K., Liu, Y., Lyne, A.G., McKee, J.W., Maan, Y., Main, R.A., Mickaliger, M.B., Nițu, I.C., Nobleson, K., Paladi, A.K., Parthasarathy, A., Perera, B.B.P., Perrodin, D., Petiteau, A., Porayko, N.K., Possenti, A., Prabu, T., Leclere, H.Q., Rana, P., Samajdar, A., Sanidas, S.A., Sesana, A., Shaifullah, G., Singha, J., Speri, L., Spiewak, R., Srivastava, A., Stappers, B.W., Surnis, M., Susarla, S.C., Susobhanan, A., Takahashi, K., Tarafdar, P., Theureau, G., Tiburzi, C., van der Wateren, E., Vecchio, A., Krishnan, V.V., Verbiest, J.P.W., Wang, J., Wang, L., Wu, Z., 2023. The second data release from the european pulsar timing array. *Astronomy & Astrophysics* 678, A50.
- Asad, K., Girard, J., De Villiers, M., Ansah-Narh, T., Iheanetu, K., Smirnov, O., Santos, M., Lehmsiek, R., Jonas, J., De Villiers, D., et al., 2021. Primary beam effects of radio astronomy antennas – II. modelling MeerKAT L-band beams. *Monthly Notices of the Royal Astronomical Society* 502, 2970–2983.
- Athreya, R., 2009. A new approach to mitigation of radio frequency interference in interferometric data. *The Astrophysical Journal* 696, 885.
- Baan, W., Fridman, P., Millenaar, P., 2004. Radio frequency interference mitigation at the westerbork synthesis radio telescope: Algorithms, test observations, and system implementation. *The Astronomical Journal* 128, 933.
- Berthereau, A., 2023. Méthodes avancées en chronométrie des pulsars : réjection d'interférences radio par apprentissage profond et étude chronométrique du système binaire relativiste J1528-3146. Ph.D. thesis. Université d'Orléans, France.
- Bhat, N., Cordes, J., Chatterjee, S., Lazio, T., 2005. Radio frequency interference identification and mitigation using simultaneous dual-station observations. *Radio Science* 40, 1–16.
- Bondonneau, L., Griefmeier, J.M., Theureau, G., Cognard, I., Brionne, M., Kondratiev, V., Bilous, A., McKee, J., Zarka, P., Viou, C., et al., 2021. Pulsars with NenuFAR: Backend and pipelines. *Astronomy & Astrophysics* 652, A34.
- Chang, A., Knapp, M., LaBelle, J., Swoboda, J., Volz, R., Erickson, P.J., 2023. Removing radio frequency interference from auroral kilometric radiation with stacked autoencoders, in: *Proc. IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, IEEE, pp. 1–5.
- DeBoer, D.R., Parsons, A.R., Aguirre, J.E., Alexander, P., Ali, Z.S., Beardsley, A.P., Bernardi, G., Bowman, J.D., Bradley, R.F., Carilli, C.L., et al., 2017. Hydrogen epoch of reionization array (HERA). *Publications of the Astronomical Society of the Pacific* 129, 045001.
- Finlay, C., Bassett, B.A., Kunz, M., Oozeer, N., 2023. Trajectory based RFI subtraction and calibration for radio interferometry. *arXiv e-prints* [2301.04188](#).
- Fridman, P., 2008. Statistically stable estimates of variance in radio-astronomy observations as tools for radio-frequency interference mitigation. *The Astronomical Journal* 135, 1810.
- Fridman, P., Baan, W., 2001. RFI mitigation methods in radio astronomy. *Astronomy & Astrophysics* 378, 327–344.
- Ghanney, Y., Ajib, W., 2020. Radio frequency interference detection using deep learning, in: *Proc. IEEE Vehicular Technology Conference (VTC-Spring)*, IEEE, pp. 1–5.
- Hamid, A., van Straten, W., Griffin, A., 2022. PSRFINET: Radio frequency interference detection in pulsar data with deep residual networks, in: *Proc. RFI Workshop*.
- Hera-Team, 2024. hera_sim Documentation. Technical Report.
- Jiang, J., Zheng, L., Luo, F., Zhang, Z., 2018. Rednet: Residual encoder-decoder network for indoor rgb-d semantic segmentation. *arXiv e-prints* [1806.01054](#).
- Kerrigan, J., Plante, P.L., Kohn, S., Pober, J.C., Aguirre, J., Abdurashidova, Z., Alexander, P., Ali, Z.S., Balfour, Y., Beardsley, A.P., et al., 2019. Optimizing sparse RFI prediction using deep learning. *Monthly Notices of the Royal Astronomical Society* 488, 2605–2615.
- Kocz, J., Briggs, F., Reynolds, J., 2010. Radio frequency interference removal through the application of spatial filtering techniques on the Parkes multi-beam receiver. *The Astronomical Journal* 140, 2086.
- Kramer, M., Stairs, I.H., Manchester, R.N., Wex, N., Deller, A.T., Coles, W.A., Ali, M., Burgay, M., Camilo, F., Cognard, I., Damour, T., Desvignes, G., Ferdman, R.D., Freire, P.C.C., Grondin, S., Guillemot, L., Hobbs, G.B., Janssen, G., Karuppusamy, R., Lorimer, D.R., Lyne, A.G., McKee, J.W., McLaughlin, M., Münch, L.E., Perera, B.B.P., Pol, N., Possenti, A., Sarkissian, J., Stappers, B.W., Theureau, G., 2021. Strong-Field Gravity Tests with the Double Pulsar. *Physical Review X* 11, 041050.
- Lazarus, P., Karuppusamy, R., Graikou, E., Caballero, R., Champion, D., Lee, K., Verbiest, J., Kramer, M., 2016. Prospects for high-precision pulsar timing with the new Effelsberg PSRIX backend. *Monthly Notices of the Royal Astronomical Society* 458, 868–880.
- Li, S.Z., 2009. Markov random field modeling in image analysis. *Springer Science & Business Media*.
- Li, Z., Yu, C., Xiao, J., Long, M., Cui, C., 2021. Detection of radio frequency interference using an improved generative adversarial network. *Astronomy and Computing* 36, 100482.
- Lorimer, D.R., Kramer, M., 2004. *Handbook of Pulsar Astronomy*. volume 4.
- Maslakovic, S., Linscott, I., Oslick, M., Twicken, J., 1996. Excising radio frequency interference using the discrete wavelet transform, in: *Proc. IEEE Int. Symp. Time-Frequency Time-Scale Analysis (TFTS)*, IEEE, pp. 349–352.
- Mesarcik, M., Boonstra, A.J., Meijer, C., Jansen, W., Rangelova, E., van Nieuwpoort, R.V., 2020. Deep learning assisted data inspection for radio astronomy. *Monthly Notices of the Royal Astronomical Society* 496, 1517–1529.
- Mesarcik, M., Boonstra, A.J., Rangelova, E., van Nieuwpoort, R.V., 2022. Learning to detect radio frequency interference in radio astronomy without seeing it. *Monthly Notices of the Royal Astronomical Society* 516, 5367–5378.
- Morello, V., Barr, E., Cooper, S., Bailes, M., Bates, S., Bhat, N., Burgay, M., Burke-Spolaor, S., Cameron, A., Champion, D., et al., 2019. The high time

- resolution universe survey – XIV. Discovery of 23 pulsars through GPU-accelerated reprocessing. *Monthly Notices of the Royal Astronomical Society* 483, 3673–3685.
- Mosiane, O., Oozeer, N., Aniyan, A., Bassett, B.A., 2017. Radio frequency interference detection using machine learning., in: *IOP Conference Series: Materials science and engineering*, IOP Publishing. p. 012012.
- Offringa, A., De Bruyn, A., Biehl, M., Zaroubi, S., Bernardi, G., Pandey, V., 2010a. Post-correlation radio frequency interference classification methods. *Monthly Notices of the Royal Astronomical Society* 405, 155–167.
- Offringa, A., Van De Gronde, J., Roerdink, J., 2012. A morphological algorithm for improving radio-frequency interference detection. *Astronomy & Astrophysics* 539, A95.
- Offringa, A.R., de Bruyn, A.G., Zaroubi, S., Biehl, M., 2010b. A LOFAR RFI detection pipeline and its first results, in proceedings of rfi mitigation workshop, in: *RFI mitigation workshop - RFI2010*.
- Page, E.S., 1954. Continuous inspection schemes. *Biometrika* 41, 100–115.
- Peck, L.W., Fenech, D.M., 2013. SERPent: Automated reduction and rfi-mitigation software for e-merlin. *Astronomy and Computing* 2, 54–66.
- Pen, U.L., Chang, T.C., Hirata, C.M., Peterson, J.B., Roy, J., Gupta, Y., Odegova, J., Sigurdson, K., 2009. The GMRT EoR experiment: limits on polarized sky brightness at 150 MHz. *Monthly Notices of the Royal Astronomical Society* 399, 181–194.
- Reardon, D.J., Zic, A., Shannon, R.M., Hobbs, G.B., Bailes, M., Marco, V.D., Kapur, A., Rogers, A.F., Thrane, E., Askew, J., Bhat, N.D.R., Cameron, A., Curylo, M., Coles, W.A., Dai, S., Goncharov, B., Kerr, M., Kulkarni, A., Levin, Y., Lower, M.E., Manchester, R.N., Mandow, R., Miles, M.T., Nathan, R.S., Osłowski, S., Russell, C.J., Spiewak, R., Zhang, S., Zhu, X.J., 2023. Search for an isotropic gravitational-wave background with the parkes pulsar timing array. *The Astrophysical Journal Letters* 951, L6.
- Saliwanchik, B.R., Slosar, A., 2022. A self-learning neural network approach for radio frequency interference detection and removal in radio astronomy. *Publications of the Astronomical Society of the Pacific* 134, 114503.
- Stappers, B., Hessels, J., Alexov, A., Anderson, K., Coenen, T., Hassall, T., Karastergiou, A., Kondratiev, V., Kramer, M., Van Leeuwen, J., et al., 2011. Observing pulsars and fast transients with LOFAR. *Astronomy & Astrophysics* 530, A80.
- Taylor, J.H., 1992. Pulsar timing and relativistic gravity. *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences* 341, 117–134.
- Vafaei Sadr, A., Bassett, B.A., Oozeer, N., Fantaye, Y., Finlay, C., 2020. Deep learning improves identification of radio frequency interference. *Monthly Notices of the Royal Astronomical Society* 499, 379–390.
- Vos, E.E., Luus, P.F., Finlay, C.J., Bassett, B.A., 2019. A generative machine learning approach to RFI mitigation for radio astronomy, in: *Proc. IEEE Int. Workshop Machine Learning Signal Processing (MLSP)*, IEEE. pp. 1–6.
- Wang, H.F., Yuan, M., Yin, Q., Guo, P., Zhu, W.W., Li, D., Feng, S.B., 2020. Radio frequency interference mitigation using pseudoinverse learning autoencoders. *Research in Astronomy and Astrophysics* 20, 114.
- Winkel, B., Kerp, J., Stanko, S., 2007. RFI detection by automated feature extraction and statistical analysis. *Astronomische Nachrichten: Astronomical Notes* 328, 68–79.
- Wolfaardt, C.J., 2016. Machine learning approach to radio frequency interference (RFI) classification in radio astronomy. Ph.D. thesis. Stellenbosch: Stellenbosch University.
- Yan, R.Q., Dai, C., Liu, W., Li, J.X., Chen, S.Y., Yu, X.C., Zuo, S.F., Chen, X.L., 2021. Radio frequency interference detection based on the AC-UNet model. *Research in Astronomy and Astrophysics* 21, 119.
- Yang, Z., Yu, C., Xiao, J., Zhang, B., 2020. Deep residual detection of radio frequency interference for FAST. *Monthly Notices of the Royal Astronomical Society* 492, 1421–1431.
- Zhang, K., Li, Y., Zuo, W., Zhang, L., Van Gool, L., Timofte, R., 2021. Plug-and-play image restoration with deep denoiser prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 6360–6376.
- Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L., 2017. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing* 26, 3142–3155.
- Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y., 2020. Residual dense network for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 2480–2495.
- Zhao, J., Zou, X., Weng, F., 2013. WindSat radio-frequency interference signature and its identification over Greenland and Antarctic. *IEEE Transactions on Geoscience and Remote Sensing* 51, 4830–4839.