

PAPER • OPEN ACCESS

## Interference Mitigation using Data-driven Sparse Component Analysis

To cite this article: Wen Deng *et al* 2024 *J. Phys.: Conf. Ser.* **2722** 012010

View the [article online](#) for updates and enhancements.

You may also like

- [An evaluation of three commercially available metal artifact reduction methods for CT imaging](#)

Jessie Y Huang, James R Kerns, Jessica L Nute et al.

- [Analysis of radon mitigation methods: 10-year review](#)

E Kouroukla, T D Gooding and H S Fonseca

- [Crosstalk in DAQ-based measurement platforms for structural health monitoring: effects on damage detection and mitigation methods](#)

Guilherme Marconato Rezende and Fabricio Guimaraes Baptista

# Interference Mitigation using Data-driven Sparse Component Analysis

Wen Deng, Xiang Wang, Zhitao Huang, Qiang Xu

National University of Defense Technology, Changsha, China

E-mail: dw0907@nudt.edu.cn

**Abstract.** To address the challenge of mitigating asynchronous and non-stationary interference under single-channel conditions, we propose a sparse component analysis interference mitigation method based on the recurrent neural network. This method aims to recover the desired signal from the received time-frequency over-lapped co-channel signal. Unlike previous interference mitigation methods, our proposed method achieves "end-to-end" signal recovery in the time domain without any priori requirements on the received signals, making it more universal than existing methods. Numerical results validate the effectiveness of our proposed method and demonstrate its significantly superior mitigation performance compared to existing schemes under different environmental noises, intensities of interfering signals, and generalization test conditions.

## 1. Introduction

Mitigation interference is crucial for enhancing the performance of cooperative/non-cooperative communication systems under single-channel conditions and has drawn attention for decades.

Existing methods under single-channel conditions primarily include methods based on generalized spectral wiener filtering and signal subspace projection. The former assumes that there is a certain spectral domain in which the signals do not exist aliasing; thus, the target signals can be separated based on linear Wiener filtering after the received signals are converted to the spectral domain. The linear-conjugate-linear frequency shift filter (LCL-FRESH) [1] and its variant, the blind adaptive FRESH filter (BA-FRESH) [2], are a concrete implementation of this idea. The disadvantage of such methods is that their performance significantly depends on the apriority of the target signal, such as cycle frequency and waveform. The principle of another scheme is to first construct a base set of subspace to ensure that the target and interfering signals are orthogonal in the subspace formed by the base set and finally achieve the separation of a single-signal waveform by projecting the received signals into this signal subspace. However, its implementation mostly requires that the signal meet specific requirements, such as the baseband signal [3] and the presence of single-signal region (SSR) [4, 5].

Unlike the above methods based on pre-established mathematical models of target signals and channels, a deep convolutional network demodulator (DCND) for signal demodulator based on multilayer convolutional neural network (CNN) was proposed in [6], in which received signals were considered input and the output of network classification layer was considered the estimation result of signal symbol sequence. However, it is susceptible to the power of interfering signals and cannot recover the waveform of the desired signal, making it impractical for applications such as spectrum monitoring and specific signal detection.



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

To overcome the limitations of existing methods, we propose a sparse component analysis interference mitigation method based on a recurrent neural network (RNN) under strong interference conditions. The proposed method utilizes the autoregressive (AR) characteristic modeling ability of RNN to design an interference mitigation framework consisting of "sparse domain encoding-sparse domain representation estimation-sparse domain decoding". First, we perform sparse transformation of the time-frequency over-lapped co-channel signal (TFOCS) using a sparse domain encoder. Next, we complete the estimation and separation recovery of the target signal representation in the sparse domain. Finally, we convert the target signal in the sparse domain to the time domain using sparse decoding to recover the target signal. Unlike existing model-based signal recovery methods, our proposed method establishes a signal recovery model based on the self-adaptation of RNN. This breaks the restriction on the a priori knowledge of received signals, making it more practical and versatile.

The numerical results demonstrate the high-quality recovery of the target signal achieved by our proposed method, even under various interference types and intensities. Furthermore, our data-driven approach exhibits commendable performance in generalization ability tests.

## 2. System Model

We consider including  $N$  pairs of "transmitting-receiving" channels [5],  $N$  main communication links, and  $N(N - 1)$  interference links, where each node uses a single antenna. The TFOCS at receiver  $i$  ( $i = 1, 2, \dots, N$ ) is defined as:

$$x_i(n) = a_{ii}s_i(n - m_i) + \sum_{j=1, j \neq i}^N a_{ji}s_j(n - m_j) + v_i(n). \quad (1)$$

where  $n \in [1, T]$ , for which  $T$  denotes the number of sampling points;  $a_{ji}$  is the channel coefficient between transmitting node  $j$  ( $j = 1, 2, \dots, N$ ) and receiving node  $i$ ;  $s_j(n - m_j)$  denotes the signal transmitted by transmitting node  $j$  with a transmission delay of  $m_j$  sampling intervals; and  $v_i(n)$  denotes AWGN at receiving node  $i$ . Here, the mixed model at the receiving node is considered a linear time-delay instantaneous mixed model, Eq. (1) can also be expressed vectorially as:

$$\mathbf{X}_i = \mathbf{A}_i \mathbf{S}_i + \mathbf{V}_i. \quad (2)$$

where  $\mathbf{X}_i \in \mathbb{R}^{1 \times T}$  denotes the TFOCS at the receiving node  $i$ ;  $\mathbf{A}_i = [a_{1i}, a_{2i}, \dots, a_{ji}, \dots, a_{Ni}] \in \mathbb{R}^{1 \times N}$  denotes the mixed matrix, where each column contains the channel coefficient between the transmitting node  $j$  and receiving node  $i$ ;  $\mathbf{S}_i = [s_{1i}, s_{2i}, \dots, s_{ji}, \dots, s_{Ni}]^\top \in \mathbb{R}^{N \times T}$  denotes the signal matrix, where each row denotes the signal transmitted by transmitting node  $j$  and received at receiving node  $i$ ; and  $\mathbf{V}_i \in \mathbb{R}^{1 \times T}$  denotes AWGN at the receiving node  $i$ . In this paper, asynchronous non-stationarity refers to the dynamic change in the quantity and variety of current interfering signals as well as their arrival time at the receiving node. Fig. 1 shows a schematic of the received TFOCS considered here.

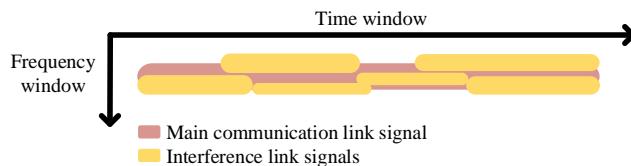


Figure 1: Schematic of TFOCS.

In this case, the purpose of interference mitigation is to achieve signal recovery of the main communication link component (i.e., the target signal  $a_{ii}$ ) contained in  $\mathbf{X}_i$ .

Specifically, signals  $s_i(n)$  ( $i = 1, 2, \dots, N$ ) considered here can be expressed as:

$$s_i(n) = \sqrt{E_i} \sum_{q_i=1}^{Q_i} A_{q_i}^i g_i[nT_{is} - (q_i - 1)N_{ic}T_{is}] e^{j2\pi f_{ic}nT_s}. \quad (3)$$

where  $E_i$  denotes the signal power, and  $A_{q_i}^i$  corresponds to the  $q_i$ -th modulation symbol of the  $i$ -th signal;  $g_i(\cdot)$  is the pulse-forming function, which proceeds for  $lN_{ic}T_{is}$  sampling points from  $(1-l)N_{ic}T_{is}$ ;  $l$  is the length of the intersymbol interference;  $N_{ic}$  is the number of oversampling points of the symbols;  $T_{is}$  is the sampling interval;  $f_{ic}$  and  $Q_i$  are the carrier frequency and total number of symbols for the  $i$ -th signal, respectively; and  $j$  denotes an imaginary number.

In our recent research [7], we have demonstrated that by applying  $r$  ( $r \in N$ )-times forward iterations to Eq. (3), we obtain  $s_i(n+r) \approx e^{(j2\pi f_{ic}rT_s)} s_i(n)$ ,  $n \in [1, Q_i N_{ic} - r]$  (the bias is derived from the amplitude change of the pulse shaping function), which describes a typical  $r$ -order centralized AR process in which the AR coefficient is  $e^{j2\pi f_{ic}rT_s}$ . When the modulation process is based on a continuous carrier, the communication signal can be well modeled via an AR process. Therefore, we model the autoregressive properties of the signal using RNNs, which contain state vectors that can store contextual information in hidden layer nodes and have the ability to model structured (backward and forward correlated) input/output sequences [8].

### 3. Proposed Interference Mitigation Method

Inspired by extant studies on SCA [9], which are based on the assumption of signal sparsity to achieve signal separation, we propose an SCA-based interference mitigation method based on a recurrent neural network to achieve the adaptive selection of the sparse transform domain of the target signal, adaptive learning of the mixed matrix in the transform domain, and automatic recovery of the target signal.

#### 3.1. system overview

The schematic of the proposed scheme is illustrated in Fig. 2, which is inspired by the "encoder–separator–decoder" structure as described in [10]. It comprises three components: a sparse domain encoder (SDE), a sparse domain representation estimator (SDRE), and a sparse domain decoder (SDD).

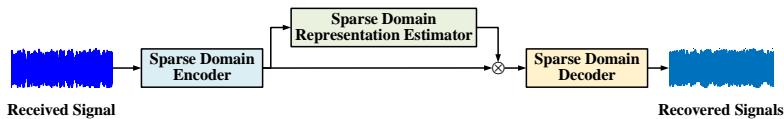


Figure 2: Schematic of the proposed interference mitigation scheme.

First, the TFOCS is inputted to the encoder to obtain the sparse domain representation; the transformation process is defined as

$$\mathbf{W}_i = f_{Encoder}(\mathbf{X}_i) \quad (4)$$

where  $\mathbf{W}_i \in \mathbb{R}^{K \times L}$  is the high-dimensional representation of  $\mathbf{X}_i$  in the sparse domain,  $K$  is the number of encoder filters,  $L$  is the output length of a single filter, and  $f_{Encoder}(\cdot)$  denotes the mapping function of the encoder.

Second, the SDRE estimates the target signal mixed matrix  $\mathbf{M}_j \in \mathbb{R}^{K \times L}$  in the sparse domain space. This matrix is then multiplied by  $\mathbf{W}_i$  to obtain an estimate  $a_{ii}\tilde{s}_{ii} \in \mathbb{R}^{K \times L}$  of the target signal in the sparse domain. The conversion process is defined as

$$a_{ii}\tilde{s}_{ii} = \tilde{\mathbf{X}}_i \otimes \mathbf{M}_j \quad (5)$$

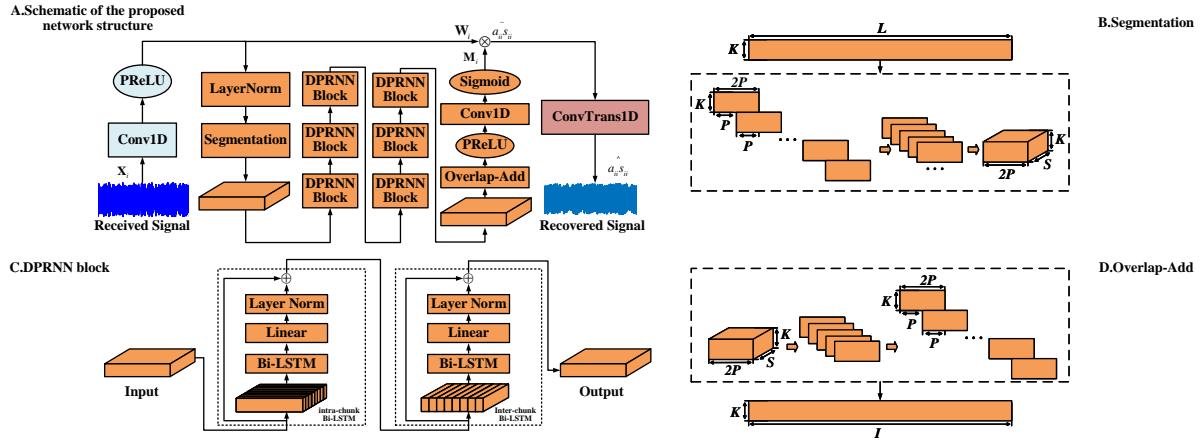


Figure 3: System flowchart of the proposed interference mitigation scheme. A. Schematic of the proposed network structure. B. Segmentation processing. C. DPRNN block processing D. Overlap-add processing.

where  $\otimes$  is the matrix dot multiplication operation.

Finally, the decoder converts  $a_{ii}\tilde{s}_{ii}$  to the time domain to obtain the estimate  $\hat{a}_{ii}s_{ii} \in \mathbb{R}^{1 \times T}$  of the target signal. The conversion process is defined as

$$a_{ij}\hat{s}_{ij} = f_{Decoder}(a_{ij}\tilde{s}_{ij}) \quad (6)$$

where  $f_{Decoder}(\cdot)$  is the mapping function of the decoder.

### 3.2. Network Design

Fig. 3(A) shows the flowchart of the proposed interference mitigation system. The design of the framework is inspired by the end-to-end audio recovery framework proposed in [10]. Moreover, the parameters of the network structure are set based on the parameters of the optimal performance case.

**3.2.1. Sparse domain encoder and decoder** Sparse transformation is achieved by a single one-dimensional convolutional layer and a parametric rectified linear unit (PReLU) [11], with the number of filters in the convolutional layer set to 64, a convolutional kernel size of 8, and a step size of 4. Sparse inverse transformation is performed by a single one-dimensional transposed convolutional layer.

**3.2.2. Segmentation and overlap-add processing** Segmentation processing splits a continuous input into overlapping blocks with an overlapping ratio of 50% and connects them to form a three-dimensional tensor, where the block step  $P$  is set to 64; the block length is 128, and finally  $S = \lceil L/P \rceil + 1$  blocks are obtained.  $\lceil \cdot \rceil$  is a ceiling function. Overlap-add processing is the reverse operation of segmentation processing. Fig. 3(B) and (D) show flowcharts of segmentation and overlap-add processing, respectively.

**3.2.3. Sparse domain representation estimator** The adaptive learning of the mixed matrix in the transform domain is mainly achieved by the RNN stacking block, which is composed of six dual-path recurrent neural network (DPRNN) blocks. The output of the RNN stacking block sequentially passes through an overlap-add process, a convolutional layer with a convolutional

kernel size of 1, and a sigmoid activation function, thereby completing the estimation of the sparse representation matrix of the target signal in the sparse domain.

The DPRNN block structure is shown in Fig. 3(C), where we use bidirectional long short-term Memory (Bi-LSTM) to model the sequence, which contains two modules: intra-chunk Bi-LSTM and inter-chunk Bi-LSTM. These modules are used for local and global modeling respectively. The intra-block Bi-LSTM processes the sequence-local blocks of each block independently, and the inter-block RNN summarizes the information from all blocks for full sequence-level processing. Considering that the intra-block LSTM is bidirectional, each time step in its output contains all the information of the block it belongs to, which enables the inter-block LSTM to perform full sequence-level modeling. Here the number of hidden nodes of LSTM is set to 128.

#### 4. Numerical Results

All experiments are implemented based on PyTorch 1.11.0 environment, and its optimization objective is to minimize the loss of mean square error (MSE) between the recovered and source signals. The training optimizer and learning rate are set to Adam [12] and 0.001, respectively; the batch size is set to 10; and the number of training rounds is 100. When evaluating the signal recovery performance, the symbol error rate (SER) of the recovered target signal is used as the performance evaluation index. The signal-to-noise ratio (SNR) and signal-to-interference ratio (SIR) in received signal are used as measures of disturbance intensity.

##### 4.1. Experiment 1: Performance Test and Analysis

This experiment considers the TFOCS at receiving node 1 in the multi-user interference channel of  $N = 4$ , which consists of the target signal and three interfering signals. The modulation style of the target signal is BPSK, whereas the modulation styles of the three interfering signals are 8PSK, PAM8, and 16QAM. The sampling duration is 10 ms, and sampling points is 32064, where the time interval of the target signal is 0–10 ms, the carrier frequency of the target signal is 100 kHz. The time intervals of the three interfering signals are 0–5 ms, 4–8 ms, and 6–10 ms respectively, and the carrier frequencies of the three interfering signals are randomly taken in the three ranges of 101–103 kHz, 97–99 kHz, and 104–106 kHz (considering a random carrier frequency offset of  $\pm[0,0.5]$  kHz). The amount of training/verification/testing data of the model is  $7 \times 10^6$ ,  $2 \times 10^6$ , and  $1 \times 10^6$  under a single SNR or SIR, respectively, with a single target signal sample containing 334 symbols and a symbol rate of 31.25 kB (Baud) for both the target and interfering signals.

Table 1 shows the performance of the proposed method under different disturbance intensities, with  $SER_P$  and  $SER_E$  denoting the signal demodulation SER before and after recovery respectively. Here,  $SER_P$  is the SER that demodulates the target signal directly based on TFOCS. Table 1(a) indicates that the presence of interfering signals before recovery results in extremely high SER for demodulation of the target signal and significantly poor communication quality. After recovery by the proposed method,  $SER_E$  decreases significantly. When  $SNR = 0$  dB,  $SER_E$  was reduced to  $3 \times 10^{-3}$ , which was two orders of magnitude lower than  $SER_P$ . When  $SNR \geq 10$ , SER is even reduced by three orders of magnitude. Table 1(b) suggests that the proposed method can ensure the high-quality recovery of the target signal waveform under different interference signal strength SIR. Although  $SER_E$  increases with increasing interference intensity, it maintains an SER of 10-4 orders of magnitude, which is an average of three orders of magnitude lower compared to  $SER_P$  before recovery. To test the adaptability of the proposed method to changes in the interference type, the interfering signal is changed to QPSK, PAM4, and 2FSK, whose result is similar to the above experiment and is not repeated herein.

Table 1: (a)SERs at different SNR, with SIR =  $-4.77$  dB. (b)SERs at different SIR, with SNR = 20 dB.

SNR (dB)	SER <sub>P</sub>	SER <sub>E</sub>	SIR (dB)	SER <sub>P</sub>	SER <sub>E</sub>
0	$1.54 \times 10^{-1}$	$3.19 \times 10^{-3}$	-4.77	$1.52 \times 10^{-1}$	$4.56 \times 10^{-5}$
5	$1.53 \times 10^{-1}$	$1.07 \times 10^{-3}$	-7.77	$2.35 \times 10^{-1}$	$7.32 \times 10^{-5}$
10	$1.52 \times 10^{-1}$	$4.54 \times 10^{-4}$	-9.54	$2.78 \times 10^{-1}$	$9.41 \times 10^{-5}$
15	$1.52 \times 10^{-1}$	$1.95 \times 10^{-4}$	-10.79	$3.06 \times 10^{-1}$	$1.18 \times 10^{-4}$
20	$1.52 \times 10^{-1}$	$1.86 \times 10^{-4}$	-11.76	$3.24 \times 10^{-1}$	$1.32 \times 10^{-4}$
25	$1.52 \times 10^{-1}$	$1.79 \times 10^{-4}$	-12.55	$3.39 \times 10^{-1}$	$1.83 \times 10^{-4}$
30	$1.52 \times 10^{-1}$	$1.41 \times 10^{-4}$	-13.22	$3.51 \times 10^{-1}$	$2.71 \times 10^{-4}$
			-13.80	$3.61 \times 10^{-1}$	$3.25 \times 10^{-4}$
			-14.31	$3.68 \times 10^{-1}$	$4.36 \times 10^{-4}$

(a)

(b)

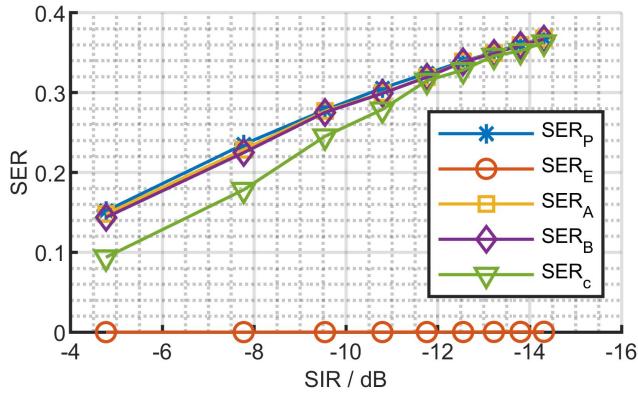


Figure 4: Comparison of the performance of different interference mitigation methods under different interference intensities with SNR being 20 dB, where SER<sub>A</sub>, SER<sub>B</sub> and SER<sub>C</sub> indicate the performance of the method in [4, 5] and [6], respectively. The overall change of SER<sub>E</sub> shows an increasing trend, and its detailed values are listed in Table 1(b).

#### 4.2. Experiment 2: Comparison with Existing methods

The existing methods involved in the comparison include a single signal recovery method based on sparse coding and template matching proposed in [4], a single signal recovery method based on dictionary learning and sparse coding under constraints proposed in [5], and a multilayer CNN-based signal demodulator (DCND) proposed in [6], among which the weight coefficient of the signal property of the method in [5] is set to 50, and for DCND, the number of convolution kernels contained in the two convolutional layers is set to 128 and 64, respectively. The experimental data and parameter settings are the same as those in Experiment 1.

As shown in Fig. 4, because the methods in [4] and [5] are based on the assumption of the presence of SSR, their performance deteriorates sharply in the absence of SSR of the target signal in TFOCS, and gradually loses the ability to mitigate interference as the intensity of the interfering signal increases. Furthermore, when the intensity of the interfering signal is low, the performance of the proposed algorithm in [5] is slightly superior to that of the algorithm proposed in [4], which is because [5] introduces the characteristics of the target signal as a constraint condition, making it possible to extract some information of the target signal when the intensity of the interfering signal is low. However, as the interference strength increases, the

power of the target signal in the TFOCS will gradually decrease, which eventually leads to the failure of the algorithm. The algorithm in [6] is based directly on the TFOCS as input. Without prior preprocessing that has interference mitigation capability, its performance deteriorates at a high intensity of the interfering signal ( $SIR \geq -9$  dB). The method proposed can maintain robust interference mitigation performance in low SIR scenarios. Meanwhile, considering that the duration of interfering signals in practice is mostly uncontrollable and the possibility of having a completely clean SSR is low, the communication scenarios tackled by the method proposed are more realistic and have wider application prospects.

#### 4.3. Experiment 3: Generalization Test

The ability of the proposed scheme to cope with generalized test conditions is further tested here. The training data parameters are the same as those in simulation experiment 1, while the setting of the test data has two changes: one is to consider the different durations of the three interfering signals, which are set to 0–4.5 ms, 4–7.5 ms, and 5.5–10 ms; the second is to consider the change in the intensity of the environmental noise interference, with the initial SNR set to 15 dB. Upon comparing the results with those of the closed-set test (i.e., a test in which the parameters of the test dataset are consistent with those of the training dataset), as depicted in Fig. 5, it can be observed that the proposed method copes well with all the aforementioned generalization conditions. This excellent generalization ability can be mainly attributed to the method of constructing the sparse domain of the target signal, which makes the target signal sufficiently sparse on the data-driven adaptive construction of the sparse domain. Notably, if different SNRs are considered during the training, the performance of the interference mitigation method proposed in this paper will be predictably improved at lower test SNRs, which can be achieved simply by generating training data over a wider range of SNRs.

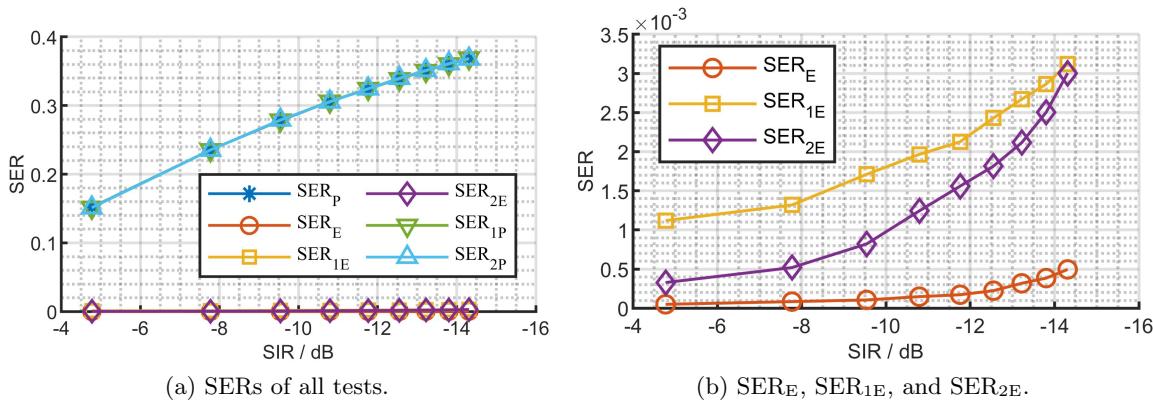


Figure 5: SER of the proposed interference mitigation method under generalization test conditions. SER<sub>1E</sub> represents the test performance in the scenario of duration change, SER<sub>2E</sub> represents the test performance at SNR = 15 dB, and SER<sub>1P</sub> and SER<sub>2P</sub> represent the performance calculated directly from their corresponding raw data, respectively.

## 5. Conclusions

This paper presents an SCA interference mitigation method based on the recurrent neural network to suppress asynchronous and non-stationary interference in multi-user wireless communication systems. Numerical results indicate that the proposed method can recover the target signal with high quality under different interference types and interference intensities and has good adaptability to strong interference situations with large differences between interference and signal power levels.

## References

- [1] W. A. Gardner, "Cyclic Wiener filtering: Theory and method," IEEE Trans. Commun., vol. 41, no. 1, pp. 151–163, Jan. 1993.
- [2] J. Zhang, "Wong KM, Luo Z", et al. Blind adaptive FRESH filtering for signal extraction [J], IEEE Transactions on Signal Processing, vol. 47, pp. 1397–1402, May 1999.
- [3] S. Heidari and C. L. Nikias, "Co-channel interference mitigation in the time-scale domain: The CIMTS algorithm," IEEE Trans. Signal Process., vol. 44, no. 9, pp. 2151–2162, 1996.
- [4] Z. Huang and X. Cai, "Interference mitigation via sparse coding in K-user interference channels," IEEE Wirel. Commun. Lett., vol. 8, no. 6, pp. 1596–1599, Dec. 2019.
- [5] X. Cai et al., "Asynchronous and non-stationary interference cancellation in multiuser interference channels," Asynchronous and Non-stationary Interference Cancellation in Multiuser Interference Channels, vol. 20, no. 8, p. 976-4989, Aug. 2021.
- [6] X. Lin, et al., "A deep convolutional network demodulator for mixed signals with different modulation types" in Computing IEEE international conference on Big Data Intelligence. Cyber Science and Technology Congress, 2017, pp. 893–896.
- [7] X. Cai, X. Wang, Z. Huang, and F. Wang, "Single-channel blind source separation of communication signals using pseudo-mimo observations," IEEE Communications Letters, vol. 22, no. 8, pp. 1616–1619, 2018.
- [8] Y. LeCun et al., "Deep learning." Nature, vol. 7553:436, p. 521, 2015.
- [9] P. Georgiev et al., "Sparse component analysis and blind source separation of underdetermined mixtures," IEEE Trans. Neural Netw., vol. 16, no. 4, Jul., 992–996, 2005, doi: 10.1109/TNN.2005.849840.
- [10] Y. Luo et al., "Dual-Path RNN: Efficient long sequence modeling for time-domain single-channel speech separation," ICASSP IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 46-50.
- [11] K. He et al., "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," pp. 1026-1034.
- [12] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," arXiv preprint arXiv:1412.6980, 2014.