University of Technology Sydney
Faculty of Engineering and Information Technology
School of Electrical and Data Engineering

Wireless Interference Mitigation for Emerging Applications and Systems

Linh Manh Hoang

A Thesis Submitted
in Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

under the supervision of

Dr. Diep N. Nguyen
Dr. J. Andrew Zhang

Sydney, Australia

May 2023

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, LINH MANH HOANG, declare that this thesis is submitted in fulfillment of the requirements for the award of DOCTOR OF PHILOSOPHY in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Student Name: Linh Manh Hoang

Student Signature: Production Note:
Signature removed prior to publication.

Date: May 11, 2023

ABSTRACT

Wireless Interference Mitigation for Emerging Applications and Systems

by

Linh Manh Hoang

Wireless jamming is one of the critical threats to emerging wireless applications, e.g., Ultra-reliable low-latency communication (URLLC). The problem becomes even more challenging when the jamming signals come simultaneously from multiple sources. This thesis aims to investigate and mitigate such a problem by leveraging signal beam-forming and machine learning (ML) techniques. To that end, the impact of the nonzero and varying correlation coefficients between transmitted jamming signals from multiple sources was first studied. It is observed that by varying the correlation coefficients, jammers can "virtually change" the jamming channels hence their nullspace even when these channels do not physically change. That makes most conventional interference suppression techniques that rely on the beam-forming matrix derived from the nullspace no longer applicable. To tackle the problem, we propose techniques to monitor the jamming residual and effectively update the beam-forming matrix.

However, such a jamming residual monitoring process incurs additional system overhead, thus significantly reducing the spectral efficiency. This gives rise to a more challenging problem in optimizing the duration of the nullspace estimation and the data transmission phases. However, even ignoring the unknown strategy of the jammers and the challenging nullspace estimation process, the resulting problem is an integer programming problem, hence intractable to obtain its optimal solution. To deal with such uncertainty and incomplete information, as well as to circumvent the intractability of the above optimization problem, we reformulate it using a partially observable semi-Markov decision process (POSMDP). We then design a deep duel-

ing Q-learning-based technique to quickly obtain the optimal policy for legitimate devices. Simulation results show that the resulting spectral efficiency is much higher than that of other methods and close to that of the perfect jamming nullification case (for moderate jamming power).

Next, we study the methods to deal with jamming signals in the joint radar and communication (JRC) systems. Specifically, novel modulation and demodulation schemes were proposed for a frequency-hopping (FH) JRC system with robustness against jamming. In these schemes, both sub-pulse frequencies and durations are used for information modulation, leading to higher communication data rates. For information demodulation, a novel scheme was proposed by using the time-frequency analysis (TFA) technique and a 'you only look once' (YOLO)-based detection system. As such, the proposed system does not require channel estimation, simplifying the transmission signal frame design. Simulation results demonstrate the effectiveness of the proposed scheme. Moreover, by using the FH signal, the JRC system is robust against the jammers that cannot implement wideband jamming due to power limitations.

Finally, we study jamming mitigation in joint communication and radar (JCR) systems. Specifically, we study how to optimize the durations of the jamming nullspace estimation, the preamble, and the data transmission phases. First, in the jamming nullspace estimation phase, the beam-forming matrix used to mitigate the jamming signal is estimated. Then, the preamble is used to estimate the legitimate channel and also for the radar function. Finally, the data transmission is performed in the data transmission phase. As such, increasing the duration of the nullspace estimation and the preamble phases can increase the radar's performance. However, such an increase also reduces the effective spectral efficiency of the communication function, because the data transmission phase fraction is decreased. The surrounding radio environments of the JRC systems are typically dynamic with high uncertainties due to their high mobility, making the duration optimization problem even more challenging. To deal with such uncertainty, we reformulate the problem using a Markov decision process (MDP). Then, we design a deep dueling

Q-learning-based technique to quickly obtain the optimal policy.

Acknowledgements

First and foremost, I would like to take this opportunity to express my deepest gratitude to my supervisors, Dr. Diep N. Nguyen and Dr. J. Andrew Zhang, for all the support, guidance, and encouragement. Without them, this dissertation would have been impossible. During my study, they not only guided me to pursue great and impactful research but also encouraged me to go beyond my boundary to do things that seemed impossible. I am truly privileged and lucky to be supervised by them.

I would like to thank all my colleagues and friends at the University of Technology Sydney for their support, discussion, and friendship. They made my PhD life more colorful. My thanks also go to the SEDE admin team for handling all the paperwork and forms during my PhD study. I would like to thank the university and the Faculty of Engineering and Information Technology (FEIT) for awarding me the International Research Scholarship and FEIT Scholarship.

I would especially like to thank my dearest family for encouraging and supporting me. Their endless love gives me the strength and power to overcome difficulties in my life. A special thanks to my friends in the HML group, including Minh Van Le and Van Huan Pham, for always being there and having insightful discussions.

Contents

List of Publications

Journal Papers

J-1. [1] **L. M. Hoang**, J. A. Zhang, D. Nguyen, X. Huang, A. Kekirigoda, and K.-P. Hui, "Suppression of multiple spatially correlated jammers," *IEEE Trans. Veh. Tech-nol.*, vol. 70, no. 10, pp. 10489 - 10500, Sep. 2021. (*Corresponding to Chapter 3*)

J-2. [2] **L. M. Hoang**, D. N. Nguyen, J. A. Zhang, and D. T. Hoang, "Multiple Correlated Jammers Nullification using LSTM-based Deep Dueling Neural Network," *submitted to IEEE Trans. Comm.*, **under review.** (*Corresponding to Chapter 4*)

J-3. [3] **L. M. Hoang**, J. A. Zhang, D. N. Nguyen, and D. T. Hoang, "Frequency Hopping Joint Radar-Communications with Hybrid Sub-pulse Frequency and Duration Modulation," *IEEE Wireless Commun. Lett.*, Aug. 2022, **Early Access** (*Corresponding to Chapter 5*)

Conference Papers

C-1. [4] J. A. Zhang, **L. M. Hoang**, D. N. Nguyen, X. Huang, A. Kekirigoda, and K.-P. Hui, "Multi-user MIMO Communications with Interference Mitigation in Time-varying Channels," in *Proc. IEEE Int. Conf. Sig. Process. Comm.*, Adelaide, SA, Australia, Dec. 2020, pp. 1–7. (*Corresponding to Chapter 3)*

C-2. [5] **L. M. Hoang**, J. A. Zhang, D. N. Nguyen, X. Huang, A. Kekirigoda, and K.-P. Hui, "Nullification of Multiple Correlated Jammers," in *Proc. IEEE Veh. Tech-nol Conf. (VTC2021-Fall)*, Norman, OK, USA, Dec. 2021, pp. 1–6. (*Corresponding to Chapter 3)*

C-3. [6] **L. M. Hoang**, D. N. Nguyen, J. A. Zhang, and D. T. Hoang, "Multiple Correlated Jammers Suppression: A Deep Dueling Q-learning Approach," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Austin, TX, USA, Apr. 2022, pp. 998–1003. (*Partly corresponding to Chapter 4)*

List of Figures

Abbreviation

AI	artificial intelligence
AoA	angle of arrival
AWGN	additive white Gaussian noise
CNN	convolutional neural network
CWD–TFI	Choi–Williams distribution time-frequency image
DL	deep learning
DSSS	direct sequence spread spectrum
eMBB	enhanced mobile broadband
ENR	energy-to-noise ratio
FHSS	frequency-hopping spread spectrum
HAP	high altitude platform
JCR	joint communication and radar
JRC	joint radar and communication
MDP	Markov decision process
MIMO	multi-input multi-output
ML	machine learning
NFV	network function virtualization
PAPR	peak-to-average power ratio
PCR	pulse compression ratio
POSMDP	partially observable semi-Markov decision process
QoS	quality of service
RF	radio frequency
R-CNN	region-based CNN
RL	reinforcement learning
SDN	software-defined networking

SINR	signal-to-interference-plus-noise ratio
SMDP	semi-Markov decision processes
STFT	short-time Fourier transform
TFI	time-frequency image
UCA	uniform circular array
ULA	uniform linear array
URLLC	ultra-reliable and low-latency communications
WVD	Wigner-Ville distribution
YOLO	'You Only Look Once' (detection system)

Chapter 1

Introduction and Literature Review

This section describes the threats to a wireless communication system posed by jamming devices, which are the motivation behind the thesis. Then, state-of-the-art solutions to the problems are comprehensively reviewed. Finally, the main contributions and structure of the thesis are provided.

1.1 Motivations

The recent development in wireless communications has enabled various emerging applications, e.g., intelligent transportation systems, industry automation (industry 4.0), smart city, etc. These applications require enhanced mobile broadband (eMBB) and ultra-reliable low latency communication (URLLC), which are made available by using state-of-the-art technologies, such as mmWave, massive multiple input multiple output (MIMO), small cell, software-defined networking (SDN), network function virtualization (NFV), etc.

However, modern wireless telecommunications networks remain susceptible to (both deliberate and inadvertent) jamming attacks due to the exposed nature of wireless links. These jamming attacks can lead to reduced throughput or even complete disruption of communications links between legitimate devices. In addition, recent studies show that jamming is very simple to implement using the software-defined radio (SDR) concept. Specifically, according to [8], a $20 USB dongle with a transmitting power of 100 mW can be used to jam 20 MHz bandwidth, enough to terminate a home or office Wifi service. Therefore, mitigating the jammer is critical

to ensure the operation of wireless networks.

Compared to intra-cell or inter-cell interferences, jamming signals have different characteristics that make them more challenging to be mitigated. This is because legitimate devices have information about interference signals, which is not the case for jamming signals (both deliberate and inadvertent). Moreover, for deliberate jammers, jamming strategies are constantly being improved to overcome existing anti-jamming measures. Therefore, anti-jamming techniques must be constantly updated to ensure adequate protection of communication systems.

Given the above, this thesis aims to propose new anti-interference techniques to deal with jamming strategies that are capable of bypassing existing anti-interference measures. Existing studies are thoroughly analyzed, and new techniques based on the most advanced tools, including machine learning techniques, are proposed to ensure effectiveness against the most advanced jamming strategies.

1.2 Literature Review and Contributions

This section provides a comprehensive review of the existing techniques for the mentioned problem, including their advantages and limitations. Then, the thesis's main contributions are highlighted. Specifically, the first study analyzes the impact of the time-varying correlation between transmitted jamming signals and then proposes a heuristic approach to deal with the impact. However, this heuristic approach creates additional network overhead, reducing the system's spectral efficiency. This limitation leads to the second study, which proposes a deep dueling Q-learning technique capable of dealing with the time-varying correlation and improving the system's spectral efficiency. Then, in the third and fourth studies, jamming mitigation techniques are extended to dual-function radar and communications (DFRC) systems.

1.2.1 Suppression of Multiple Spatially Correlated Jammers

1.2.1.1 Literature Review

One conventional approach for suppressing jamming is nullifying jamming signals using the angle of arrival (AoA) based beam-forming technique. It is realized by steering the nulls in the receiving beam towards the AoAs of jamming signals, which are obtained by AoA estimation techniques [9]. Popular examples of such an approach are multiple signal classification (MUSIC) [10], estimation of signal parameters via rotational invariance techniques (ESPRIT) [11], and the matrix pencil [12]. However, for these techniques, the receiver has to "sacrifice" at least one antenna to nullify each arriving jamming signal [13]. This makes the approach not suitable when the number of jammers is large, or when the signal from each jammer reaches the receiver through a large number of propagation paths (e.g., in urban areas). This is because the receiver does not want to "lose" too many antennas for jamming suppression, as a lower number of remaining antennas generally results in lower or even zero throughput between legitimate devices [14].

Spread spectrum communication is another approach to deal with jamming signals. Specifically, in the frequency-hopping spread spectrum (FHSS) technique [15–18], the legitimate devices try to avoid the jamming signals by first detecting the jamming frequency range and then switching their operating frequency to another channel (within their operating frequency band). The direct sequence spread spectrum (DSSS) technique [15,19,20], on the other hand, uses pseudo-random noise (PRN) codes to encode and spread the legitimate data signal to a much larger bandwidth, hence avoiding narrow-band jamming. Moreover, using the DSSS technique, the encoded data signals become noise-like and have low average power, making it challenging for the jammer to detect and interfere. However, when the operating frequency band of the legitimate devices is known to the jammers, they can perform

wide-band jamming (i.e., given sufficient jamming power) to void the functionality of the spread spectrum techniques.

Alternatively, jamming suppression can be realized by estimating the projection of jamming channels [21–23], the ratios between jamming channels [24], or the nullspace of the jamming channels [25–29], and then designing filters to nullify the jamming signals. These methods require only one receiver's antennas to nullify each spatial stream of a jammer, which is more effective than the AoA-based beam-forming technique mentioned above.

However, the techniques in [21–30] do not consider the impact of the correlations between transmitted jamming signals on the jamming nullification performance. Note that the correlation between the jamming channels does not affect the estimation of the jamming channels' nullspace. On the other hand, the correlations between transmitted jamming signals, as later demonstrated in this chapter, cause a "virtual change" in the jamming channels, making the estimated nullspace incapable of nullifying the jamming signals. In [31] and [32], the authors show that, by choosing a suitable correlation, the jammers can maximize the jamming impact to communications between legitimate devices. Therefore, to ensure the effectiveness of jamming suppression techniques, it is important to consider the correlations between transmitted jamming signals. Nonetheless, the main objective of [31] and [32] is to analyze the impact of the correlation on the jamming outcome. The two papers neither propose any suppression technique for correlated jamming nor discuss the impact of the correlation on the performance of such a jamming nullification method, especially under time-varying channels. Another unanswered question is the impact of varying the non-zero correlation coefficient, if any, on the jamming suppression process.

Instead of "dodging" (i.e., by DSSS) or "escaping" (i.e., by FHSS), legitimate

devices can also leverage jamming signals for data transmission. In [33], using backscattering and energy harvesting, the legitimate transmitter can choose to either adapt its transmission rate, backscatter the jamming signals, harvest energy (from the jamming signals), or stay idle. In this way, the legitimate transmitter can mitigate the impact of the jamming signals and even utilize jamming energy for its communication. However, this approach requires additional hardware component, which is not always available, especially in mobile devices. Moreover, backscattering communication is limited in its transmission range and throughput.

It is worth noting that the AoA-based beam-forming technique mentioned above performs pretty well against the time-varying correlations, except for the weakness of "sacrificing" too many degree-of-freedoms for jamming suppression. Specifically, the spatial smoothing technique [20] can decorrelate the received jamming signals before the AoAs of the jamming signal streams are estimated. However, spatial smoothing can only be applied directly on arrays with Vandermonde-structure steering vectors, such as the uniform linear array (ULA). For other array structures, such as the uniform circular array (UCA), the phase mode excitation (PME) technique is needed to transform the steering vector into the Vandermonde structure, which degrades the AoA estimation performance because of PME approximation. To overcome the limitation of the spatial smoothing technique, the compressive sensing-based AoA estimation technique is developed [34], which shows an acceptable performance against the correlations for all types of the array structure. After the AoAs of the spatial streams are estimated, the jamming signals can be nullified by steering receiving beam nulls toward the estimated AoAs. However, as discussed above, the AoA-based beam-forming technique is not suitable when the environment is rich scattering, because an excessive number of degree-of-freedoms is needed only for jamming suppression.

The jamming suppression has some analogies to interference avoidance in cog-

nitive radio. The secondary user (SU) tries to avoid interfering the primary user (PU) by transmitting signals that lie in the nullspace of the channel between the SU transmitter and PU receiver. In [35–38], the authors propose blind nullspace learning and tracking techniques to estimate the nullspace of the channel between the SU transmitter and the PU receiver, without requiring any SU-PU interaction. This is realized by having the SU continuously adjust the transmitted signal's spatial direction and monitor the impact of the adjustment on the interference output on the PU's receiver (by observing the transmitted power from the corresponding PU transmitter). However, this approach cannot be applied in the jamming suppression scenario, because we do not have control over the transmitted signals from malicious jammers.

1.2.1.2 Contributions

Given the above, in Chapter 3, we study the impact of the non-zero and varying correlations between transmitted jamming signals on the jamming suppression/nullification in multi-user multiple-input multiple-output (MU-MIMO) and joint communications and sensing systems. Although some existing works already exploit channel variations to improve frequency diversity [39] and delay diversity [40], this work uniquely aims to combat the artificial channel variation caused by varying correlation of jamming signals. The work starts with the jamming suppression for uncorrelated jamming signals yet under time-varying channels. Then, the effect of non-zero and varying correlations on the jamming suppression process is proved to be similar to that under the time-varying channels. This finding then leads to the development of jamming nullification techniques that effectively track the jamming nullspace and correspondingly update receiving beams under all correlation levels. The proposed techniques cost only a single degree-of-freedom of receiving antennas to nullify each jammer's spatial stream. Monte Carlo simulations are provided,

showing that the proposed techniques are capable of suppressing the jamming signals for all considered scenarios with non-zero and varying correlations (between transmitted jamming signals). The contributions of this chapter can be summarized as follows.

- Study the impact of the non-zero correlations between transmitted jamming signals on the jamming suppression/nullification process.
- Reveal that by varying the correlation coefficients, jammers can effectively "change" the jamming channels even when these channels do not physically change.
- Characterize the impact of the non-zero and varying correlations on the jamming suppression/nullification process.
- Develop novel jamming suppression algorithms that effectively update the beam-forming matrix to suppress the jamming signals for all considered scenarios with non-zero and varying correlations.
- Propose jamming nullification techniques that cost only one receiver's degree-of-freedom to suppress each spatial stream of the jamming signals, so that the remaining antennas can be used for communication between legitimate devices.

1.2.2 Multiple Correlated Jammers Nullification Using LSTM-based Deep Dueling Neural Network

1.2.2.1 Literature Review

As described in subsection 1.2.1, by deliberately varying the correlations among jamming signals, attackers can create an impact on the nullspace estimation of the jamming channel, similar to having a "virtual change" in the jamming channel, even

when the physical channels remain unchanged. The "virtual change" makes the estimated beam-forming matrix outdated/incapable of nullifying the jamming signals. To deal with this "virtual change" in the jamming channels, one can continuously monitor the residual jamming signals and then heuristically adjust the estimated beam-forming matrix. However, such a jamming residual monitoring process incurs additional system overhead, thus significantly reducing the spectral efficiency. This creates a more challenging problem in optimizing the duration of the nullspace estimation and the data transmission phases. A longer nullspace estimation phase may result in a lower jamming residual but then a shorter data transmission phase.

1.2.2.2 Contributions

Given the above, in Chapter 4, we propose a solution to nullify multiple correlated jammers whose correlation is unknown and time-varying. To this end, the optimization problem of the nullspace estimation and data transmission phases is systematically formulated. Even ignoring the unknown strategy of the jammers and the challenging nullspace estimation process, the resulting problem is an integer programming problem, hence intractable to obtain its optimal solution. In practice, as aforementioned, the jammers can deliberately vary the correlation range, making jamming nullification even more challenging. To deal with such uncertainty and incomplete information, and to circumvent the intractability of the above conventional optimization problem, the problem is reformulated using a partially observable semi-Markov decision process (POSMDP). Then, a deep dueling Q-learning based technique [41, 42] is designed, which improves the training process by using two streams of fully connected hidden layers to concurrently train the Q-learning algorithm, thereby quickly obtaining the optimal policy for the legitimate devices. The proposed technique does not require legitimate devices to constantly monitor the residual jamming signals, and only costs a single degree-of-freedom to nullify

each jammer, even with an unknown and time-varying correlated jamming strategy. Unlike the strategy obtained by the game-theoretic approach, which may not be optimal, by leveraging the latest advances in deep reinforcement learning, this work aims to design an algorithm that converges to the optimal strategy by observing and learning from jammers' strategy. Simulation results show that the resulting spectral efficiency is much higher than that of other methods and close to that of the perfect jamming nullification case. The major contributions of the chapter are as follows.

- We demonstrate that using an incorrectly estimated beam-forming matrix (i.e., caused by the time-varying correlations) can lose the receiver's degree-of-freedoms. This loss corresponds to the lower-bound spectral efficiency that is even less than the spectral efficiency of the system when jamming nullification is not employed.

- We formulate the nullspace estimation and data transmission phase optimization problem. The objective is to efficiently and accurately estimate the beam-forming matrix to nullify the jamming channel, and hence maximizing the spectral efficiency and minimizing the outage probability. The resulting integer programming problem is challenging even with complete information about the jammer's strategy and no uncertainty about the legitimate and jamming channels (that is not the case in reality).

- To address the above problem, we propose a POSMDP framework to capture the dynamics, uncertainty, and unobservability of the jamming strategy.

- To cope with the large space dimension of the POSMDP, we design a deep dueling-based Q-learning technique to obtain the optimal solution for the duration of the nullspace estimation and data transmission phases.

- We carry out extensive simulations to validate the performance of the proposed

framework and compare it with state-of-the-art jamming suppression methods and the upper bound where the jamming is perfectly nullified.

1.2.3 Frequency Hopping Joint Radar and Communications with Hybrid Sub-pulse Frequency and Duration Modulation

In the previous sections, jamming mitigation techniques for communication systems were studied. In the following sections, we extend the study to dual-function radar and communications (DFRC), including joint radar and communication (JRC) and joint communication and radar (JCR) systems.

1.2.3.1 Literature Review

The DFRC systems [43–49] have received significant attention due to their potential to improve spectrum efficiency and reduce device size, cost, and power consumption. According to their primary function, the DFRC systems can be categorized into (communication-centric) joint communication and radar (JCR) and (radar-centric) joint radar and communication (JRC). The JCR systems prioritize their communication performance, hence are most applicable in civil applications that require high transmission rates, such as vehicular networks or perceptive mobile networks [50]. On the other hand, the JRC systems focus on radar sensing performance, and are more prevalent in military applications, such as airborne, shipborne, and ground-based combat systems.

For JRC systems, the primary research topic is the data embedding scheme on the radar signal. The first approach requires modifications to the radar signal waveform, including spatial, code, and frequency domain embedding techniques. In the spatial domain [43], the data bits are embedded by controlling the phase or amplitude of the transmitted signal in the beampattern's sidelobes. This technique, however, restricts communication to the sidelobe direction only, while the communication receiver

can reside in the main lobe. The code-domain embedding technique, on the other hand, uses binary or poly-phase modulation for communication signals and avoids interference between the communication and radar signals by using different types of spread spectrum sequences [50]. However, this technique may lead to leakage outside the permitted bandwidth due to its alteration in the signal spectrum. Finally, the frequency-domain embedding technique uses the upward and downward frequency chirp in a frequency-modulated continuous waveform (FMCW) signal frequency to represent 1's and 0's bits, respectively. Nevertheless, this embedding technique limits its communication rate to the radar signal's pulse repetition frequency (PRF).

The frequency-hopping (FH) JRC systems [43, 50–54] have become more and more prevalent in both military and civil applications, e.g., airborne, shipborne, ground-based combat systems, the connected autonomous vehicle (CAV). This is made possible thanks to the powerful anti-jamming properties of the frequency hopping signal, which spreads the signal spectrum over a large bandwidth. The most popular FH JRC systems include multiple-input multiple-output orthogonal frequency division multiplexing (MIMO-OFDM), multi-Carrier AgilE phaSed Array Radar (CAESAR), and FH-MIMO [50].

Similar to other JRC systems, a key challenge in FH JRC is to embed/modulate data (onto frequency sub-pulses) at the transceiver and demodulate the signal at the receiver. For data embedding, the most dominant technique is index modulation (IM), which utilizes frequency selection/combination and/or antenna selection/permutation for data representation. Specifically, the data bits are embedded by selecting different sets of sub-pulse frequencies (i.e., frequency combination) and allocating them to different antennas (i.e., antennas permutation). For data demodulation, the optimal demodulator can be based on the maximum likelihood principle [51], while sub-optimal methods, which require lower computation complexities, are based on compressive sensing (CS) or discrete Fourier transform (DFT) [50].

However, these methods' performance depends on the channel estimation's accuracy. Unfortunately, a long training sequence necessary for accurate channel estimation is not always feasible. This is because a long training sequence reduces the communication fraction over the whole time frame, thus decreasing the data bit rate. In particular, the data embedding schemes using both sub-pulse frequency combinations and antenna permutations require complex demodulation techniques that are prone to demodulation error. On the other hand, the data embedding scheme using only the sub-pulse frequency combination has a limited data transmission rate.

1.2.3.2 Contributions

In Chapter 5, we propose novel techniques to embed and demodulate data in an FH JRC system to increase the data rate and reduce the demodulation error. For data embedding, both sub-pulse frequency and duration are used, increasing the data transmission rate compared to only using the sub-pulse frequency. For data demodulation, a novel scheme based on the signal's time-frequency image (TFI) and a 'You Only Look Once' (YOLO)-based detection system is proposed. Instead of requiring a channel estimation, this demodulation scheme only requires the estimation of the channel delay spread, thus less prone to estimation error. Moreover, the proposed demodulation technique is more robust to the Doppler shift and the timing offset between the transceiver and the communication receiver than the existing ones. Additionally, the proposed data embedding and demodulating schemes are spatially flexible and not limited to the sidelobe of the transmit beampattern. This is because the data is not embedded by utilizing the phase or amplitude of the beampattern sidelobe.

The contributions of the proposed technique are summarized as follows.

- A novel data modulation technique is proposed by using both sub-pulse frequency and duration, thereby increasing the data transmission rate compared

to only using the sub-pulse frequency.

- A novel data demodulation technique is proposed based on the signal's TFI and a YOLO-based detection system. This demodulation technique is less prone to channel estimation error and is more robust to the Doppler shift and the timing offset between the transceiver and the communication receiver compared to the existing ones.

- The proposed modulation and demodulation techniques are spatially flexible and not limited to the sidelobe of the transmitted beam-pattern, because the data is not modulated by using the beam-pattern sidelobe.

- Extensive simulations were carried out to demonstrate the performance of the proposed technique compared to the existing ones.

1.2.4 Jamming Mitigation in Joint Communication and Radar Systems: A Deep Dueling Q-learning Approach

1.2.4.1 Literature Review

In this subsection, we study the optimization of the durations of the jamming nullspace estimation, preamble, and data transmission phases in a JCR system. This is an extension of the study described in Subsection 1.2.2, where the durations of the nullspace estimation and data transmission phases are optimized. Specifically, the beam-forming matrix, used to mitigate the jamming signal, is estimated during the jamming nullspace estimation phase. On the other hand, the preamble is used for the radar function as well as to estimate the legitimate channel. Therefore, by increasing the duration of the nullspace estimation and the preamble phases the radar function performance can be improved. However, such an increase also reduces the effective spectral efficiency of the communication function, because the data transmission phase fraction is decreased.

There have been multiple studies on the optimization of the signal waveforms to achieve a better joint radar and communication performance [55–59]. Specifically, [56] proposes using the single carrier physical layer (SC-PHY) frame's preamble for the radar function. However, even though the detection and range estimation performances are acceptable, the velocity estimation is poor due to the limited length of the preamble. To address this limitation, [59] use the preamble of the control physical layer (C-PHY) frames that is longer than that of the SC-PHY. To further improve the velocity estimation performance, the multi-frame technique [60–62] is proposed by using the signal from multiple frames for the radar function. However, the studies in [55–59] do not consider the impact of the jamming signal on the JCR performances. More importantly, the dynamic and uncertainty of the surrounding environment can also significantly degrade the performance of the JRC system.

1.2.4.2 Contributions

Given the above, in Chapter 6, we study how to optimize the durations of the jamming nullspace estimation, the preamble, and the data transmission phases in a dynamic environment. To deal with the uncertainty in the environment, we reformulate the problem using a Markov decision process (MDP). Then, we design a deep dueling Q-learning-based technique to quickly obtain the optimal policy. The contributions of the proposed technique are summarized as follows.

- We formulate the nullspace estimation, preamble, and data transmission phase optimization problem. The objective is to maximize the joint radar and communication metric.
- We demonstrate that the dynamics and uncertainty of the surrounding environment make it even more challenging to achieve an optimized solution.
- To address the above problem, we propose a semi-MDP (SMDP) framework

to capture the dynamics, uncertainty, and unobservability of the jamming strategy.

- To cope with the large space dimension of the SMDP, we design a deep dueling-based Q-learning technique to obtain the optimal solution for the duration of the nullspace estimation and data transmission phases.

1.3 Thesis Organization

The rest of this thesis is organized as follows.

- Chapter 2: This chapter provides the fundamental background of reinforcement learning, deep learning, deep reinforcement learning, and the "You Only Look Once" object detection system. Specifically, Section 2.1 describes the fundamentals of reinforcement learning, including the Markov decision process (MDP), semi-MDP (SMDP), and the Q-learning technique. Section 2.3 describes the deep reinforcement technique, which combines the advantages of DL and Q-learning. Finally, Section 2.4 describes the YOLO detection system used for the demodulation technique in the JRC system.

- Chapter 3: This chapter presents a heuristic solution to the non-zero and varying correlations between transmitted jamming signals. Specifically, Section 3.1 formulates the problem. Section 3.3 analyzes the impact of the correlations between transmitted jamming signals on the system performance. Section 3.4 and Section 3.5 describe jamming nullification when the correlations are zero and non-zero, respectively. Section 3.6 provide simulation results. Finally, Section 3.7 concludes the Chapter.

- Chapter 4: This chapter studies the technique to avoid the additional system overhead required by the heuristic solution proposed in Chapter 3 and, especially, to deal with the dynamics, uncertainty and incomplete information

of the system. Specifically, Section 4.1 formulates the problem being investigated. The deep dueling Q-learning technique to solve the stated problem is presented in Section 4.2. The simulation results are given in Section 4.3. Finally, the conclusions are drawn in Section 4.4.

- Chapter 5: This chapter studies jamming mitigation in a JRC system. In particular, Section 5.1 describes the system model. Section 5.2 and Section 5.3 describe the data embedding technique and the sensing function at the radar transceiver. The demodulation technique used to retrieve the embedded data is explained in Section 5.4. The simulation results are given in Section 5.5. Finally, the conclusions are drawn in Section 5.6.

- Chapter 6: This chapter studies jamming mitigation in a JCR system. In particular,

- Chapter 7: This chapter concludes the thesis and provides future research directions.

Chapter 2

Background

This thesis aims to exploit the current advances in AI, including reinforcement learning, deep reinforcement learning, and object detection systems, to mitigate interference and jamming signals in future communication systems. In the following, the fundamentals of reinforcement learning and deep reinforcement learning are first provided. Then, 'You Only Look Once', an object detection system, is discussed in detail.

2.1 Reinforcement Learning

Reinforcement learning is a type of machine learning that enables the system to learn iteratively from interacting with the environment. Specifically, given a state of the environment, the system takes action and observes the immediate reward and the next state of the environment, as shown in Fig. 2.1. In this way, the system can achieve the optimal policy that defines the rule to choose the action given an environment state. Moreover, by continuously learning from interacting with the environment, reinforcement learning helps the system to deal with the environment's dynamics and uncertainties. In the following, we introduce the Markov Decision Process (MDP), Q-learning, deep Q-learning, and deep dueling Q-learning, which are the key concepts in RL.

2.1.1 Markov Decision Process (MDP)

MDP is a mathematical decision-making framework used to deal with the dynamic and uncertainty of the system. An MDP is defined by a tuple $(\mathcal{S}, \mathcal{A}, r)$,

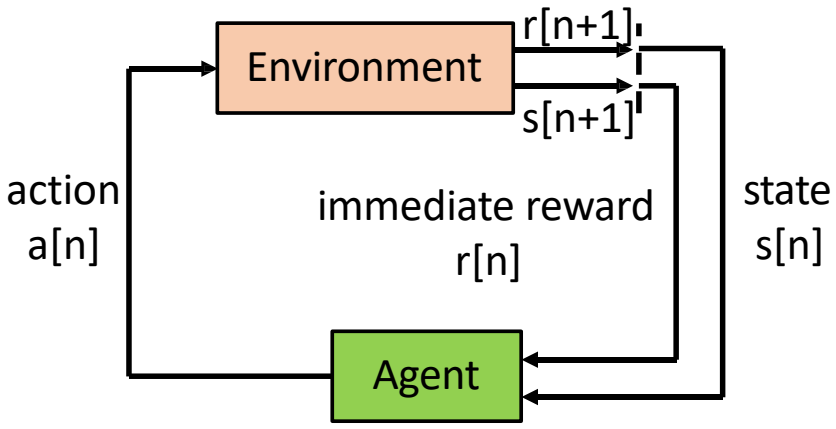


Figure 2.1 : Reinforcement Learning.

where $\mathcal{S}$, $\mathcal{A}$, and r denote the state space, action space, and the reward function, respectively. A semi-MDP (SMDP), an extension of the MDP, retains the three components mentioned above and adds an additional component, that is, the nth decision epoch length, denoted by $t[n]$. In an MDP, the state transition occurs at regular time steps (and hence the decision epoch length $t[n]$ is excluded). On the other hand, the SMDP allows the state transition to occur at irregular time steps (i.e., different $t[n]$ for different epochs). Hence, the SMDP is more suitable for real-time systems because of irregular event occurrences. Details on the MDP and SMDP can be found in [63].

2.1.2 Q-learning Technique

This subsection introduces Q-learning [64], a model-free reinforcement learning technique used to acquire the optimal policy for the decision-making system. Let $\pi : \hat{\mathcal{S}} \rightarrow \mathcal{A}$ denote a policy which is a mapping function from the states to the actions taken by the system. The purpose of Q-learning is to find the optimal value of π, denoted by π^*, that maximizes the average long-term reward of the system. As demonstrated in Fig. 2.2, the Q-learning technique contains a Q-table to store all the state-action values, which is used to select the action based on the current state.

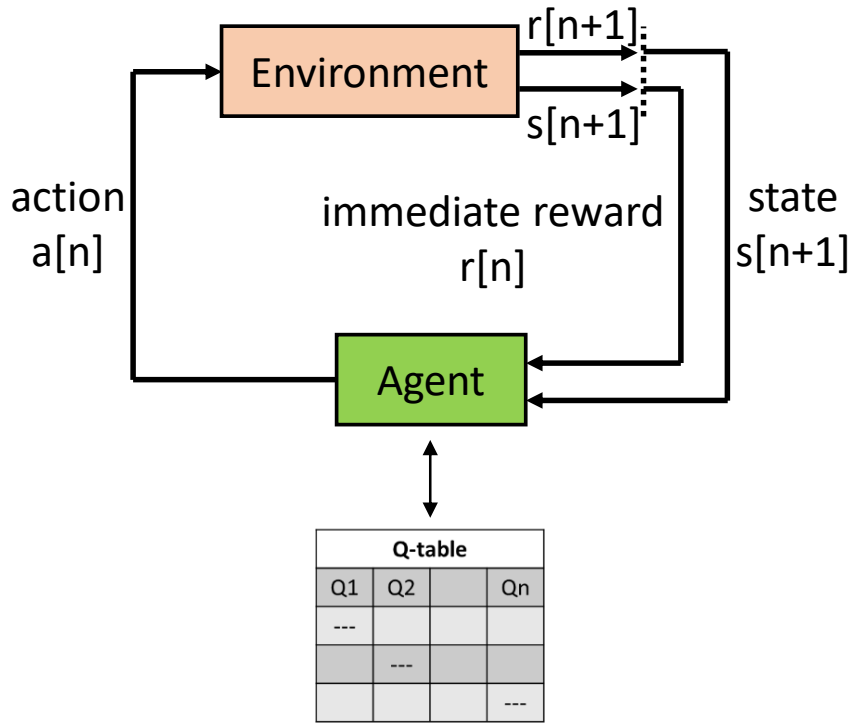


Figure 2.2 : Q-Learning.

After performing the selected action, the system observes the immediate reward and the next state to update the corresponding state-action value in the Q-table. In this way, Q-learning learns iteratively from its interaction with the environment.

Let $\mathcal{V}(s)$, $Q(s,a)$, and $\mathcal{G}(s,a)$ denote the state value function, the state-action value function, and the (state-dependent) action advantages function, respectively. The state value function $\mathcal{V}(s)$ is the expected cumulative reward of the system starting from the state s, illustrating "how good" it is for the system to be in the state s. On the other hand, the state-action value function $Q(s,a)$ demonstrates the expected discounted reward of the system in state s selecting an action a. Finally, $\mathcal{G}(s,a)$ subtracts the state value function $\mathcal{V}(s)$ from the state-action value function $Q(s,a)$ to acquire the importance of each action. Specifically, $\mathcal{V}(s)$ and $Q(s,a)$ can

be expressed by [65, Ch. 3],

$$\mathcal{V}(s) \triangleq \mathbb{E}\Big\{ \sum_{i=0}^{\infty} \gamma^i r[n+i+1] \big| s[n] = s \Big\}, \tag{2.1}$$

$$Q(s,a) \triangleq \mathbb{E}\Big\{ \sum_{i=0}^{\infty} \gamma^i r[n+i+1] \big| s[n] = s, a[n] = a \Big\}, \tag{2.2}$$

$$\mathcal{G}(s,a) \triangleq Q(s,a) - \mathcal{V}(s). \tag{2.3}$$

where $\gamma \in [0,1)$ is the discount factor, illustrating the importance of the long-term reward over the immediate reward. Specifically, a larger value of $\gamma[i]$ means the system is more interested in the long-term reward and vice versa. The objective of the Q-learning is to find the optimal value of the state-action value function $Q(s,a)$, denoted by $Q^*(s,a)$, for all state-action pairs. Then, the optimal policy π^* is obtained by $\pi^* = \text{argmax}_a\, Q^*(s,a)$, meaning at each state, the system selects the action that maximizes the expected discounted reward. The optimal state-action value function $Q^*(s,a)$ can be achieved in an iterative manner as follows. First, the state-action value function $Q(s,a)$ is arbitrarily initialized (e.g., by setting all elements to zero). Then, at each ith iteration, given system state $s[i]$, the Q-learning technique executes an action $a[i]$ and observes the immediate reward $r[i]$ and the next state $s[i+1]$. A set of values $\{s[i], a[i], r[i], s[i+1]\}$ is referred to as a transition. By using the current transition $\{s[i], a[i], r[i], s[i+1]\}$, the system iteratively updates the state-action value function $Q(s,a)$ using the formula

$$\begin{aligned} Q(s[i],a[i]) = (1-\alpha[i])Q(s[i],a[i]) \\ + \alpha[i]\big\{ r[i] + \gamma[i] \max_{a[i+1]} Q(s[i+1],a[i+1]) \big\}, \end{aligned} \tag{2.4}$$

where $\alpha[i]$ is the learning rate at the ith iteration that defines the importance of the current transition $\{s[i], a[i], r[i], s[i+1]\}$ compared to the prior ones. Specifically, increasing the value of $\alpha[i]$ enhances the importance of the current transition and vice versa. Note that, for the Q-learning technique, to guarantee that $Q(s,a)$ converges

to $Q^*(s,a)$, the learning rate $\alpha[i]$ should satisfies the conditions [33, 64]

$$\alpha[i] \in [0,1), \sum_{i=1}^{\infty} \alpha[i] = \infty, \text{ and } \sum_{i=1}^{\infty} (\alpha[i])^2 < \infty. \tag{2.5}$$

By using the updating rule in Eq. (2.4), the system can learn from the experiences (i.e., the sets of transitions $\{s[i], a[i], r[i], s[i+1]\}$) to iteratively derive the optimal state-action value function $Q^*(s,a)$, which corresponds to the optimal policy π^*. However, when the state dimension is large, the Q-learning algorithm converges slowly, which leads to the development of the deep Q-learning technique discussed below.

2.2 Deep Learning

Deep learning (DL), also referred to as deep neural networks, is a subset of machine learning that uses multiple hidden layers in the neural network to learn the patterns within the input data, as demonstrated in Fig. 2.3. By mimicking the human brain, DL algorithms map the new input to already learned data to propose an accurate output. DL has several advantages that make it superior to conventional ML methods, as described in the following.

- *Automatic feature generation:* In traditional ML techniques, the features need to be manually designed by an ML expert. On the other hand, with DL, the features are automatically generated based on the set of training data, enabling its capability to handle large and complex datasets.

- *Supports parallel and distributed algorithms:* this property is beneficial for large datasets. Specifically, instead of spending a huge amount of time training the model, one can train it on multiple machines to leverage higher computing power and speed up the training process.

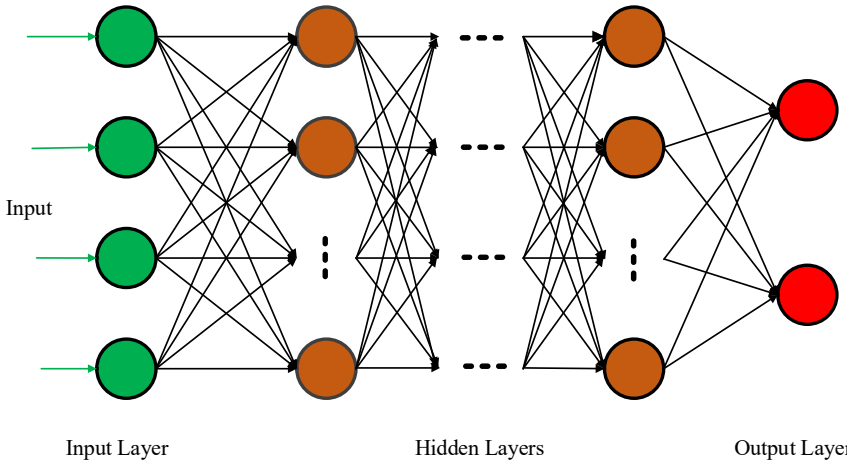


Figure 2.3 : Deep learning neural network.

- *Transfer learning:* this concept refers to the capability to retrain a pre-trained network to use in the new recognition task, thereby significantly reducing the training time.

2.3 Deep Reinforcement Learning

This section introduces deep reinforcement learning (DRL), which combines RL and DL's advantages to quickly achieve the system's optimal policy. Specifically, deep Q-learning and deep dueling Q-learning, two of the most important DRL techniques, are presented.

2.3.1 Deep Q-learning Technique

The deep Q-learning technique is used to resolve the slow-convergence problem of the Q-learning technique described above, especially when the state dimension is large. In the Q-learning approach, $Q^*(s, a)$ is iteratively obtained and stored in a Q-table. However, in the deep Q-learning technique, a neural network, referred to as the Q-network and denoted by $\mathcal{Q}$, is used as a nonlinear function approximator to estimate $Q^*(s, a)$. The input to the Q-network $\mathcal{Q}$ is the approximate state s, and

the output from $\mathcal{Q}$ is the optimal state-action value function $Q^*(s, a)$.

Let θ denote the parameters of the Q-network $\mathcal{Q}$; the problem of finding $Q^*(s, a)$ becomes the problem of finding θ^*, which are the optimal values of θ. Accordingly, θ is included in the state value function, the state-action value function, and the action advantages function notations. For example, the state-action value function is now denoted by $Q(s, a; \theta)$, and its optimum is denoted by $Q^*(s, a; \theta^*)$. The deep Q-learning technique is based on the one in [66], and formed by the following techniques.

- ϵ-greedy action selection policy: At each training iteration, the system implements *exploration* (i.e., by selecting a random action) with a probability of ϵ, or *exploitation* (i.e., by choosing the action that maximizes the current state-action value $Q(s, a; \theta)$) with a probability of $1 - \epsilon$. The value of ϵ is large (e.g., $\epsilon = 1$) at the starting iteration, and decays over the iterations as θ gets closer to the optimal value θ^*.

- Experience replay: Instead of using instant transition at each iteration, the system stores the transitions in a memory pool $\mathbf{M}$ of size M using the first-in-first-out (FIFO) protocol. This technique allows the previous transitions to be used more than once, hence improving the training data's efficiency. More importantly, by randomly selecting the training data from $\mathbf{M}$, the algorithm can remove the correlation between the consecutive training data.

- Target Q-network: This technique uses a separate network to generate the target Q-values $y[j]$. The separate network with parameter $\hat{\theta}$ is named the target Q-network, and denoted by $\hat{\mathcal{Q}}$. Instead of updating at every iteration, the target Q-network $\hat{\mathcal{Q}}$ is only renewed every C steps. As such, the primary Q-network is slowly updated, which helps to reduce the correlations between the estimated and target Q-values, hence improving the stability of the deep

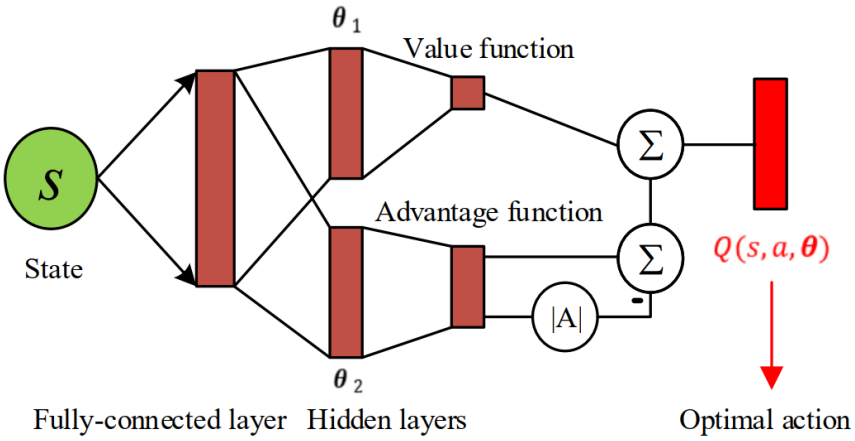


Figure 2.4 : Deep dueling neural network architecture.

Q-learning technique.

- Mini-batch gradient descent (MGD) [67]: At each training iteration of the deep Q-learning technique, the system randomly samples a mini-batch with N_{mb} samples from $\mathbf{M}$, and then performs mini-batch gradient descent on the mini-batch. By setting $N_{\text{mb}} \ll D$, the training time can be reduced dramatically [67].

2.3.2 Deep Dueling Q-learning Technique

As illustrated in [41], due to the overestimation of the optimizer, the convergence rate of the deep Q-learning algorithm is still limited. Therefore, in [41], the author proposed a novel dueling structure for the Q-network, as illustrated in Fig. 2.4. As can be seen, the deep neural network is divided into two streams of fully connected layers to estimate the advantages function $\mathcal{G}(s,a)$ and the state value function $\mathcal{V}(s)$ separately. Then, these two are combined at the output layer to calculate the state-action value function $Q(\hat{s},a;\theta)$ as [41]

$$Q(\hat{s},a;\theta) = \mathcal{V}(\hat{s};\theta) + \big(\mathcal{G}(\hat{s},a;\theta) - \frac{1}{|\mathcal{A}|}\sum_{a}\mathcal{G}(\hat{s},a;\theta)\big), \tag{2.6}$$

where $|\mathcal{A}|$ denotes the dimension of the action space $\mathcal{A}$. A detailed description of the deep dueling Q-learning algorithm is given in [41]. The deep dueling Q-learning algorithm is given in Alg. 2.1.

2.4 'You Only Look Once' (YOLO) Object Detection System

This section introduces the YOLO [7], one of the object detection systems achieving the highest mean average precision (mAP). Unlike conventional convolutional neural networks (CNN), which only classify the whole image, object detection systems can determine the object type and its location in the image. State-of-the-art object detection systems, besides YOLO, include region-based CNN (R-CNN [68]), fast R-CNN [69], faster R-CNN [70], and Single Shot Multi-Box Detector (SSD) [71]. Among these object detection systems, YOLO is well-known for its real-time detection speed, which is critical in tactical applications, such as military JRC systems. The real-time detection speed is achieved by splitting the input image into a grid of cells and predicting an object and its bounding box on each cell, instead of performing region proposal as in other object detection systems. Details of the YOLO detection system can be found in Fig. 2.5 and [7].

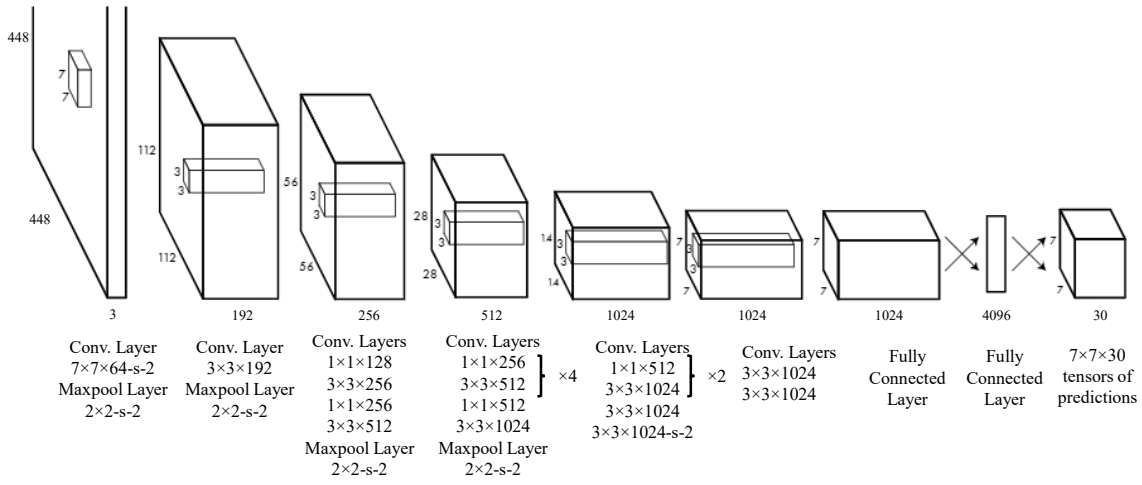


Figure 2.5 : YOLO network contains 24 convolutional layers followed by two fully connected layers [7].

Algorithm 2.1 Deep Dueling Q-learning Algorithm.

1: Initialize a memory $\mathbf{M}$ with capacity M, initialize a value for the learning rate α and the discount factor γ.

2: Initialize $\mathcal{Q}$ and $\hat{\mathcal{Q}}$ with random weights θ.

3: **for** iteration $i = 1$ *to* I **do**

4: Select action

$$a[i] = \begin{cases} \text{random action}, & \text{with probability } \epsilon \\ \text{argmax}_a \, Q(\hat{s}[i], a; \theta), & \text{otherwise.} \end{cases} \tag{2.7}$$

5: Perform $a[i]$, observe reward $r[i]$ and the next approximate state $\hat{s}[i+1]$.

6: Store the transition $\{\hat{s}[i], a[i], r[i], \hat{s}[i+1]\}$ in $\mathbf{M}$.

7: Randomly sample a mini-batch of N_{mb} transitions $\{s[j], a[j], r[j], s[j+1]\}$ from $\mathbf{M}$.

8: Set

$$Q(\hat{s}[j], a[j]; \theta) = \mathcal{V}(\hat{s}[j]; \theta) + \big(\mathcal{G}(\hat{s}[j], a[j]; \theta) - \frac{1}{|\mathcal{A}|} \sum_{a[j]} \mathcal{G}(\hat{s}[j], a[j]; \theta)\big) \tag{2.8}$$

9: Set $y[j] = r[j] + \gamma[i] \max_{a[j+1]} \hat{Q}(\hat{s}[j+1], a[j+1]; \hat{\theta})$

10: Perform MGD [67] with a learning rate $\alpha[i]$ on $\{y[j] - Q(\hat{s}[j], a[j]; \theta)\}^2$ with respect to θ.

11: Set $\hat{\mathcal{Q}} = \mathcal{Q}$ every C iterations.

12: **end for**

Chapter 3

Suppression of Multiple Spatially Correlated Jammers

This chapter studies the impact of the non-zero and varying correlations between transmitted jamming signals on the jamming suppression/nullification in a multi-user multiple-input multiple-output (MU-MIMO) system. Although some existing works already exploit channel variations to improve frequency diversity [39] and delay diversity [40], this work uniquely aims to combat the artificial channel variation caused by varying correlation of jamming signals. The jamming suppression for uncorrelated jamming signals yet under time-varying channels was studied first. Then, it is proved that the effect of non-zero and varying correlations on the jamming suppression process is similar to that under the time-varying channels. This finding then leads to the development of jamming nullification techniques that effectively track the jamming nullspace and correspondingly update receiving beams under all correlation levels. The proposed techniques cost only a single degree-of-freedom of receiving antennas to nullify each jammers' spatial stream. Monte Carlo simulations are provided, showing that the proposed techniques are capable of suppressing the jamming signals for all considered scenarios with non-zero and varying correlations (between transmitted jamming signals).

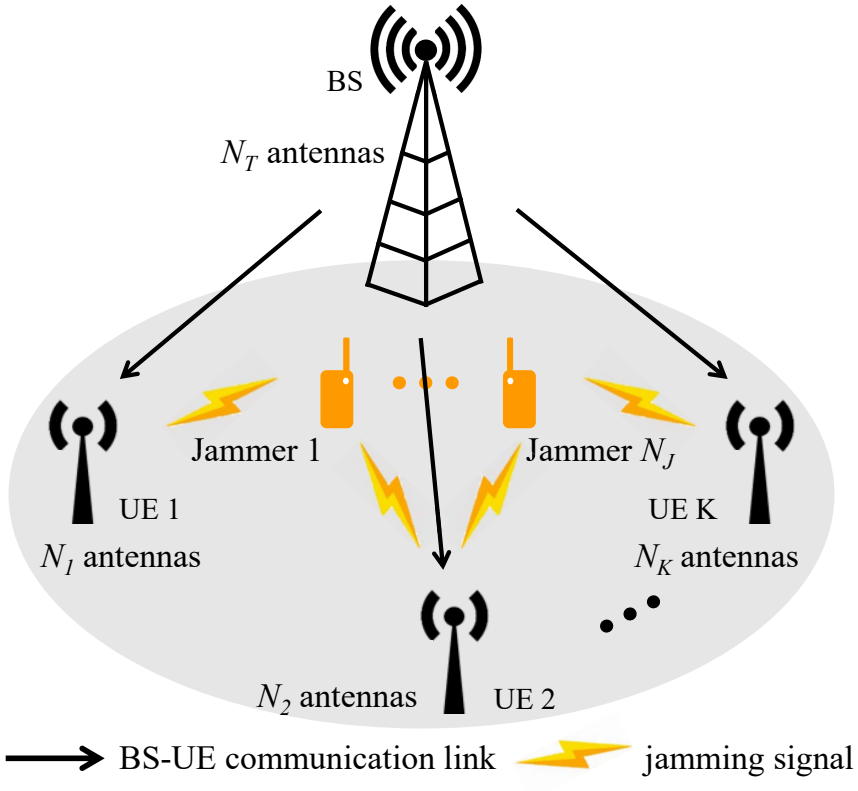


Figure 3.1 : MU-MIMO system with proactive jammers.

3.1 System Model

3.1.1 Network Model

A multi-user multiple-input multiple-output (MU-MIMO) downlink system is considered with one BS and K user equipment (UEs), as demonstrated in Fig. 3.1. The BS and each kth UE have uniform linear array (ULA) structures with N_{T} and N_k antennas, respectively. For each kth UE, M_k independent streams are transmitted from the BS. The BS-UEs communication system is jammed/interfered with by N_{J} single-antenna proactive jammers. Note that the techniques and results obtained in this chapter also capture the case with multi-antenna jammers that can be considered as multiple single-antenna jammers. For either case, it is assumed that the total number of jamming antennas is N_{J}.

The received signal at the kth UE is given by

$$\mathbf{y}_k = \sqrt{P_{\mathrm{T}}}\mathbf{H}_k\mathbf{P}_k\mathbf{x}_k + \sqrt{P_{\mathrm{T}}}\mathbf{H}_k\sum_{l\neq k}^{K}\mathbf{P}_l\mathbf{x}_l + \mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \mathbf{w}, \tag{3.1}$$

where P_{T} is the transmitted power from the BS, $\mathbf{H}_k \in \mathbb{C}^{N_k\times N_{\mathrm{T}}}$ denotes the BS-kth UE channel, $\mathbf{P}_k \in \mathbb{C}^{N_T\times M_k}$ is the precoder applied at the BS for the kth UE, M_k is the number of independent streams for the kth UE, $\mathbf{x}_k \in \mathbb{C}^{M_k\times 1}$ denotes the signal transmitted from the BS to the kth UE, $\mathbf{x}_{\mathrm{J}} = [\mathbf{x}_{\mathrm{J}1}; \mathbf{x}_{\mathrm{J}2}; ...\mathbf{x}_{\mathrm{J}N_{\mathrm{J}}}] \in \mathbb{C}^{N_{\mathrm{J}}\times 1}$ is the transmitted jamming signal, $\mathbf{Z}_k \in \mathbb{C}^{N_k\times N_{\mathrm{J}}}$ is the jammers-kth UE channel, and $\mathbf{w} \in \mathbb{C}^{N_k\times 1}$ is complex noise. The elements of $\mathbf{w}$ are assumed to be zero-mean circularly-symmetric complex Gaussian random variables (i.e., $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \sigma_{\mathrm{w}}^2\mathbf{I}_{N_k})$, where $\mathbf{I}_{N_k}$ denotes the identity matrix of size N_k, and σ_{w}^2 is the noise variance). Likewise, it is assumed that $\mathbf{x}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{M_k})$. The columns of the precoding matrix $\mathbf{P}_k$ are selected from these of the Walsh–Hadamard matrix [72], and are normalized to meet the power constraint, such that $||\mathbf{P}_k||_F = 1$, where $||.||_F$ denotes the Frobenius norm of a matrix. It is assumed that the jammers, if more than one, can coordinate to control the correlation between the transmitted jamming signals.

3.1.2 Jamming Signal Model

The received jamming signal at the kth UE is given by

$$\mathbf{y}_{\mathrm{J}_k} = \mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \mathbf{w}, \tag{3.2}$$

It is assumed that $\mathbf{x}_{\mathrm{J}} \sim \mathcal{CP}(\boldsymbol{\mu}_{\mathrm{J}}, \boldsymbol{\Sigma}_{\mathrm{J}})$, where $\mathcal{P}$ denotes a distribution function, $\boldsymbol{\mu}_{\mathrm{J}}$ and $\boldsymbol{\Sigma}_{\mathrm{J}}$ are the mean and covariance matrix of $\mathbf{x}_{\mathrm{J}}$. Note that the jamming strategy, $\mathcal{P}$, $\boldsymbol{\mu}_{\mathrm{J}}$ and $\boldsymbol{\Sigma}_{\mathrm{J}}$ are unknown to the BS and the UEs. Note also that we consider persistent jammers due to their popularity [73] in wireless networks and the existence of several papers dealing with smart jammers that intermittently transmit the jamming signals [33, 74].

Let ρ_{ij} be the complex Pearson correlation between the transmitted jamming signals from the ith and the jth jammer. The value of ρ_{ij} can be expressed by [75, Ch. 4]

$$\rho_{ij} = \frac{\mathbb{E}(\mathbf{X}_{\mathrm{J}i}\mathbf{X}_{\mathrm{J}j}^{H})}{\sigma_{\mathrm{J}_i}\sigma_{\mathrm{J}_j}}, \tag{3.3}$$

where $\mathbf{X}_{\mathrm{J}j}$ denotes a sample set of $\mathbf{x}_{\mathrm{J}j}$, and $\sigma_{\mathrm{J}_j}^2$ is the variance of the jth transmitted jamming signal. The covariance matrix $\mathbf{\Sigma}_{\mathrm{J}}$ can be expressed by

$$\mathbf{\Sigma}_{\mathrm{J}} = \begin{bmatrix} \sigma_{\mathrm{J}_1}^2 & \rho_{12}\sigma_{\mathrm{J}_1\sigma_2} & ... & \rho_{1N_{\mathrm{J}}}\sigma_{\mathrm{J}_1}\sigma_{\mathrm{J}_{N_{\mathrm{J}}}} \\ \rho_{12}^{*}\sigma_{\mathrm{J}_1}\sigma_{\mathrm{J}_2} & \sigma_{\mathrm{J}_2}^2 & ... & \rho_{2N_{\mathrm{J}}}\sigma_{\mathrm{J}_2}\sigma_{\mathrm{J}_{N_{\mathrm{J}}}} \\ ... & ... & ... & ... \\ \rho_{1N_{\mathrm{J}}}^{*}\sigma_{\mathrm{J}_1}\sigma_{\mathrm{J}_{N_{\mathrm{J}}}} & \rho_{2N_{\mathrm{J}}}^{*}\sigma_{\mathrm{J}_2}\sigma_{\mathrm{J}_{N_{\mathrm{J}}}} & ... & \sigma_{\mathrm{J}_{N_{\mathrm{J}}}}^2 \end{bmatrix}. \tag{3.4}$$

As mentioned in the introduction and will be described in more detail in Section 3.5, the time-varying correlations between transmitted jamming signals create a "virtual change" in the jamming channel, even when the physical channels stay unchanged. Based on the behavior of the "virtual change" in the jamming channel, the jammers can disable the functionality of the conventional jamming suppression techniques by deliberately varying the correlations. Formally, the correlations are controlled by the jammers using the formula

$$\rho_{ij}(t) = \mathcal{J}(i, j, t), \forall i \neq j \in (1, 2, ..., N_{\mathrm{J}}), \tag{3.5}$$

where $\mathcal{J}$ is a function unknown to the UEs and the BS.

3.1.3 Channel Model

The BS-kth UE channel can be given by

$$\mathbf{H}_k = \frac{1}{\sqrt{\eta_k}} \sum_{p=1}^{N_k^{\mathrm{p}}} \alpha_{k,p}\mathbf{a}(\phi_{k,p}^{\mathrm{a}})\mathbf{a}(\phi_{k,p}^{\mathrm{d}})^{T}, \tag{3.6}$$

where η_k denotes the large-scale path-loss of the BS-kth UE channel, $\alpha_{k,p}$ is the complex path gain, N_k^{p} is the total number of propagation paths, $\phi_{k,p}^{\mathrm{a}}$ and $\phi_{k,p}^{\mathrm{d}}$ are

the AoA and angle of departure (AoD) of the pth path for the kth UE, respectively, and $\mathbf{a}(\phi^{\mathrm{a}}_{k,p})$ and $\mathbf{a}(\phi^{\mathrm{d}}_{k,p})$ are the steering vectors corresponding to $\phi^{\mathrm{a}}_{k,p}$ and $\phi^{\mathrm{d}}_{k,p}$, respectively. The steering vectors can be expressed by [9, Ch. 3]

$$\mathbf{a}(\phi^{\mathrm{a}}_{k,p}) = \left[1, e^{-j\frac{2\pi d_k}{\lambda}\sin(\phi^{\mathrm{a}}_{k,p})}, ..., e^{-j\frac{2\pi d_k}{\lambda}(N_k-1)\sin(\phi^{\mathrm{a}}_{k,p})}\right]^T,$$
$$\mathbf{a}(\phi^{\mathrm{d}}_{k,p}) = \left[1, e^{-j\frac{2\pi d_{\mathrm{T}}}{\lambda}\sin(\phi^{\mathrm{d}}_{k,p})}, ..., e^{-j\frac{2\pi d_{\mathrm{T}}}{\lambda}(N_{\mathrm{T}}-1)\sin(\phi^{\mathrm{d}}_{k,p})}\right]^T,$$

where d_k and d_{T} are the antenna element spacing at the kth UE and the BS, respectively, and λ is the carrier's wavelength. The AoA $\phi^{\mathrm{a}}_{k,p}$ is assumed to be uniformly distributed over $[0, 2\pi]$ [76]. It is further assumed that N^{p}_k is sufficiently large (e.g., $N^{\mathrm{p}}_k \geq 8$), such that the elements of $\mathbf{H}_k$ are zero-mean circularly-symmetric complex Gaussian random variables (i.e., $\mathbf{H}_k \sim \mathcal{CN}(\mathbf{0}, 1/\sqrt{\eta_k})$).

Similarly, the jth jammer-kth UE channel is given as

$$\mathbf{Z}_{k,j} = \frac{1}{\sqrt{\eta_{k,j}}} \sum_{p=1}^{N^{\mathrm{p}}_{k,j}} \alpha_{k,j,p}\mathbf{a}(\phi^{\mathrm{a}}_{k,j,p}), \tag{3.7}$$

with $\eta_{k,j}$, $N^{\mathrm{p}}_{k,j}$, $\alpha_{k,j,p}$, $\phi^{\mathrm{a}}_{k,j,p}$, and $\mathbf{a}(\phi^{\mathrm{a}}_{k,j,p})$ defined in the same way to η_k, N^{p}_k, $\alpha_{k,p}$, $\phi^{\mathrm{a}}_{k,p}$, and $\mathbf{a}(\phi^{\mathrm{a}}_{k,p})$ in Eq. (3.6), respectively. The COST 231 Hata model [77, Ch. 4] is used to model the large-scale path-losses η_k and $\eta_{k,j}$. For the multipath fadings (i.e., expressed by the summations in Eq. (3.6) and Eq. (3.7)), without loss of generality, the flat fast fading Rician model in [78] is adopted. The values of related parameters are specified in Section 3.6.

For reference purposes, important notations and symbols are given in Table 5.1. The superscripts "e" and "d" are employed to denote symbols in the nullspace estimation and data transmission phases, respectively. For example, $\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}}$ and $\mathbf{\Sigma}^{\mathrm{d}}_{\mathrm{J}}$ represent $\mathbf{\Sigma}_{\mathrm{J}}$ values in the nullspace estimation and data transmission phases, respectively. On the other hand, $\hat{(.)}$ denotes the estimated value. For example, the estimated value of $\mathbf{G}^{\mathrm{e}}_k$ (introduced in Section 3.2) is denoted by $\hat{\mathbf{G}}^{\mathrm{e}}_k$.

Table 3.1 : Notation and Symbols.

Notation	Description	Notation	Description
$\hat{\mathbf{F}}_k^{\mathrm{e}}$	Estimated beam-forming matrix for the kth UE.	$\lvert.\rvert$	Modulus of complex number or dimension of space.
N^{e}, N^{d}	Number of samples in the nullspace estimation and data transmission phases, respectively.	$\mathcal{N}^{\mathrm{e}}$, $\mathcal{N}^{\mathrm{d}}$	Sets of candidates for N^{e} and N^{d}, respectively.
$\delta_{k,m}$, $\hat{\delta}_{k,m}$, $\bar{\delta}_k$, $\delta_{\min}$	Post-equalization signal-to-interference-plus-noise ratio (SINR) of the mth stream for the kth UE, estimate of $\delta_{k,m}$, average of $\hat{\delta}_{k,m}$, and minimum required post-equalization SINR, respectively.	$\mathcal{S}$, $\hat{\mathcal{S}}$, $\mathcal{A}$, r	State space, approximate state space, action space, and immediate reward, respectively.
$\Lambda_{k,l}$ $\bar{\Lambda}_l$	The lth largest singular value of $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}}$ and the average of $\Lambda_{k,l}$, respectively.	$\mathcal{Q}$, $\hat{\mathcal{Q}}$	Q-network and target Q-network, respectively.
$s[n]$, $\hat{s}[n]$ $a[n]$, $r[n]$	State, approximate state, action, and immediate reward at the nth epoch, respectively.	$(.)^{-1}$, $(.)^*$, $(.)^T$, $(.)^H$	The inverse, transpose, conjugate, and Hermitian transpose matrix operations, respectively.
$\mathcal{V}(s)$, $Q(s,a)$ $\mathcal{G}(s,a)$	Value function, state-action value function, and action advantages function, respectively.	Var(.) Cov(.)	The variance and covariance, respectively.

3.2 Problem Formulation

First, the communication protocol and the estimated beam-forming matrix employed to suppress the jamming signals are described. Then the impacts of the time-varying correlations among jamming signals on the jamming suppression process (using the estimated beam-forming matrix) is briefly analyzed. Next, the upper and lower bounds for the spectral efficiency of each BS-UE communication link that employs the estimated beam-forming matrix to suppress the jamming signals are derived. Finally, the problem is mathematically stated. To highlight the impact of the time-varying correlations on designing the estimated beam-forming matrix, it is

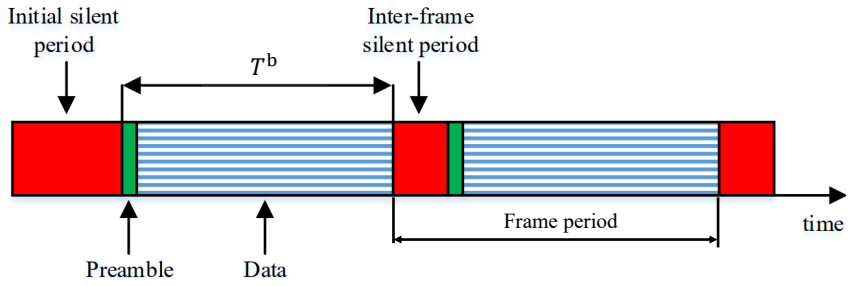


Figure 3.2 : Communication protocol for jamming suppression.

assumed that the BS-UEs and the jamming channels follow a block-fading model with coherence time [37] T^{c}, corresponding to N^{c} samples. It is further assumed that the nullspace estimation, preamble, and data transmission phases of the communication protocol (described below) are performed within the interval T^{c}, such that $N^{e} + N^{p} + N^{d} < N^{c}$, where N^{e}, N^{p}, N^{d} are the number of samples of the nullspace estimation, preamble, and data transmission phase, respectively.

To guarantee BS-UE communication, both interference signals from other UEs and the jamming signals must be suppressed. There exist various techniques for suppressing the interference signals, including orthogonal signal design, precoder, and linear minimum mean square error (LMMSE) equalization. Therefore, in this chapter, the concentration is on techniques for suppressing the jamming signals. In particular, in the subspace-based approach, this is achieved by multiplying the received signal with a beam-forming matrix derived from the left nullspace [79, p. 181] of the received jamming signals.

3.2.1 Communication Protocol

Fig. 3.2 illustrates the communication protocol for the jamming/interference nullification purpose [30]. As shown, each frame is comprised of three phases: nullspace estimation, preamble, and data transmission.

- During the nullspace estimation phase, which lasts for N^{e} samples, the beam-forming matrix that is used to suppress the jamming signals, is estimated. Let $\hat{\mathbf{F}}_k^{\mathrm{e}}$ denote the estimated beam-forming matrix. Let $\mathbf{x}_{\mathrm{J}}^{\mathrm{e}}$ and $\mathbf{y}_{\mathrm{J}_k}^{\mathrm{e}}$ denote the values of $\mathbf{x}_{\mathrm{J}}$ and $\mathbf{y}_{\mathrm{J}_k}$ during the nullspace estimation phase, respectively. Let $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}}$ and $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}}$ be a sample set of N^{e} samples and the corresponding covariance of $\mathbf{y}_{\mathrm{J}_k}^{\mathrm{e}}$. We have,

$$\mathbf{y}_{\mathrm{J}_k}^{\mathrm{e}} = \mathbf{Z}_k \mathbf{x}_{\mathrm{J}}^{\mathrm{e}} + \mathbf{w}, \qquad \mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}} = \frac{1}{N^{\mathrm{e}}} \mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}} (\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}})^H. \tag{3.8}$$

  Let $\mathbf{G}_k^{\mathrm{e}} \in \mathbb{C}^{(N_k - N_{\mathrm{J}}) \times N_k}$ be a matrix whose rows form an orthonormal basis for the left nullspace [80, Ch. 2] of $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}}$, and let $\hat{\mathbf{G}}_k^{\mathrm{e}}$ denote the estimated value of $\mathbf{G}_k^{\mathrm{e}}$. The estimated beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{e}}$ is designed by choosing its rows from the rows of $\hat{\mathbf{G}}_k^{\mathrm{e}}$. Let B_k be the row number of $\hat{\mathbf{F}}_k^{\mathrm{e}}$, one has $B_k \leq (N_k - N_{\mathrm{J}})$. To leverage all $(N_k - N_{\mathrm{J}})$ remaining degree-of-freedoms (after jamming suppression) for BS-UE signal multiplexing, we set $\hat{\mathbf{F}}_k^{\mathrm{e}} = \hat{\mathbf{G}}_k^{\mathrm{e}}$ by letting $B_k = (N_k - N_{\mathrm{J}})$. Note that N_{J} can be estimated by using the minimum description length (MDL) technique [81]. For the kth UE receiver being able to suppress N_{J} jammers then demodulate M_k spatial signal streams, we assume that $N_k \geq N_{\mathrm{J}} + M_k$ (i.e., $B_k \geq M_k$). This is because at least one degree-of-freedom is needed to suppress each jammer or demodulate each stream. Such a limitation on the antenna number does not come from our framework but from the inherent degree-of-freedom constraint in any MIMO system. Similar to [25], $\hat{\mathbf{G}}_k^{\mathrm{e}}$ can be calculated by letting $\hat{\mathbf{G}}_k^{\mathrm{e}} = (\mathbf{U}_{\mathrm{w}}^{\mathrm{e}})^H$, where $\mathbf{U}_{\mathrm{w}}^{\mathrm{e}}$ is extracted from the singular value decomposition (SVD) of $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}}$ as

$$\mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}} = [\mathbf{U}_{\mathrm{s}} \ \mathbf{U}_{\mathrm{w}}^{\mathrm{e}}] \begin{bmatrix} \mathbf{\Lambda}_{\mathrm{s}} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{\mathrm{w}} \end{bmatrix} \begin{bmatrix} (\mathbf{U}_{\mathrm{s}})^H \\ (\mathbf{U}_{\mathrm{w}}^{\mathrm{e}})^H \end{bmatrix}. \tag{3.9}$$

The received signal after nullification can be represented as [82],

$$\mathbf{p}_k = \hat{\mathbf{F}}_k^{\mathrm{e}}(\sqrt{P_{\mathrm{T}}}\mathbf{H}_k\mathbf{P}_k\mathbf{x}_k + \sqrt{P_{\mathrm{T}}}\mathbf{H}_k\sum_{l\neq k}^{K}\mathbf{P}_l\mathbf{x}_l + \mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \mathbf{w}). \tag{3.10}$$

- During the preamble phase, which lasts for N^{p} samples, the jamming signals are nullified by multiplying Eq. (3.1) with $\hat{\mathbf{F}}_k^{\mathrm{e}}$, and the BS-UE equivalent channel $\tilde{\mathbf{H}}_k$ (i.e., $\tilde{\mathbf{H}}_k = \hat{\mathbf{F}}_k^{\mathrm{e}}\mathbf{H}_k\mathbf{P}_k$) is estimated. The estimation of $\tilde{\mathbf{H}}_k$ can be performed using pilot signals and a channel estimator, such as the minimum mean-square error (MMSE) or least-square (LS) technique.

- During the data transmission phase, which lasts for N^{d} samples, the BS sends data to UEs.

3.2.2 Problem Formulation

In fact, the effectiveness of the jamming nullification depends on how close $\hat{\mathbf{G}}_k^{\mathrm{e}}$ is to $\mathbf{G}_k^{\mathrm{d}}$, where $\mathbf{G}_k^{\mathrm{d}}$ denotes the matrix whose rows form an orthonormal basis for the left nullspace of the received jamming signals at the kth UE in the data transmission phase. The closer $\hat{\mathbf{G}}_k^{\mathrm{e}}$ to $\mathbf{G}_k^{\mathrm{d}}$ results in the better jamming nullification. The first factor that affects the similarity between $\hat{\mathbf{G}}_k^{\mathrm{e}}$ and $\mathbf{G}_k^{\mathrm{d}}$ is how relatively frequently $\hat{\mathbf{G}}_k^{\mathrm{e}}$ is updated. In practice, a more frequent update of $\hat{\mathbf{G}}_k^{\mathrm{e}}$ makes it closer to $\mathbf{G}_k^{\mathrm{d}}$, and certainly leads to less residual jamming after the jamming nullification process. On the other hand, the more frequent update of $\hat{\mathbf{G}}_k^{\mathrm{e}}$ also increases the system's overhead, because the BS-UE channel estimation and synchronization have to be performed at the beginning of each frame. Therefore, less time is available for BS-UE communication. The second factor that affects the similarity between $\hat{\mathbf{G}}_k^{\mathrm{e}}$ and $\mathbf{G}_k^{\mathrm{d}}$, as will be demonstrated in this chapter, is how fast the correlations between transmitted jamming signals change. Because of the second factor, the jammers can intentionally vary the correlations to make jamming nullification more challenging.

Given the above relationship between $\hat{\mathbf{G}}_k^{\mathrm{e}}$ and $\mathbf{G}_k^{\mathrm{d}}$. The problems to be solved are as follows.

- Characterize the impact of the correlations between transmitted jamming signals on the estimation of $\hat{\mathbf{G}}_k^{\mathrm{e}}$.
- Characterize the impact of the time-varying correlations on jamming nullification.
- Design jamming nullification techniques for different scenarios of the correlations.

The content of the following sections is as follows. Section 3.3 describes the impact of the correlations on the estimation of $\hat{\mathbf{G}}_k^{\mathrm{e}}$. Then, in Section 3.4 and Section 3.5, the jamming nullification techniques for different scenarios of the correlations are proposed.

3.3 Impact of the Correlations between Transmitted Jamming Signals

In this section, the impact of the correlations between transmitted jamming signals on the relationship between $\hat{\mathbf{G}}_k^{\mathrm{e}}$ and $\mathbf{G}_k^{\mathrm{d}}$ is described. Note that, to emphasize the impact of the correlations on the estimation of $\hat{\mathbf{G}}_k^{\mathrm{e}}$, it is assumed that the jamming channels are under a slow block fading with coherence time $T^{\mathrm{c}} \gg T^{\mathrm{b}}$. Therefore, the jamming channels are considered unchanged over each frame, such that $\mathbf{Z}_k^{\mathrm{e}} = \mathbf{Z}_k^{\mathrm{d}} = \mathbf{Z}_k$. For brevity, in this section, only $\mathbf{Z}_k$ is used to represent the jamming channels. Correspondingly, only $\mathbf{G}_k$ is used to denote the matrix whose rows form an orthonormal basis of the left nullspace of $\mathbf{Z}_k$, instead of $\mathbf{G}_k^{\mathrm{e}}$ and $\mathbf{G}_k^{\mathrm{d}}$. The section is started by an introduction on the algorithm for estimating $\hat{\mathbf{G}}_k^{\mathrm{e}}$ from

the received jamming signals $\mathbf{y}^{\mathrm{e}}_{\mathrm{J}_k}$. Then, the relationship between $\hat{\mathbf{G}}^{\mathrm{e}}_k$ and $\mathbf{G}^{\mathrm{d}}_k$ is illustrated by comparing them with $\mathbf{G}_k$.

3.3.1 Impact of Correlations between Transmitted Jamming Signals

To illustrate the impact of the correlations between transmitted jamming signals on the relationship between $\hat{\mathbf{G}}^{\mathrm{e}}_k$ and $\mathbf{G}^{\mathrm{d}}_k$, the impact of the correlations on the relationship between $\hat{\mathbf{G}}^{\mathrm{e}}_k$ and $\mathbf{G}_k$ is first illustrated. It is done by analyzing the matrix $\hat{\mathbf{G}}^{\mathrm{e}}_k\mathbf{Z}_k$ in relation to the correlations between transmitted jamming signals.

Note that, from Eq. (3.8), $\mathbf{R}^{\mathrm{e}}_{\mathrm{J}_k}$ is a Hermitian matrix. Therefore, the SVD shown in (3.9) is equivalent to an eigen-decomposition (EVD) of $\mathbf{R}^{\mathrm{e}}_{\mathrm{J}_k}$. Consequently, the columns of $\mathbf{U}_{\mathrm{s}}$ and $\mathbf{U}^{\mathrm{e}}_{\mathrm{w}}$ are the eigenvectors of $\mathbf{R}^{\mathrm{e}}_{\mathrm{J}_k}$ with the corresponding eigenvalues in the diagonals of $\mathbf{\Lambda}_{\mathrm{s}}$ and $\mathbf{\Lambda}_{\mathrm{w}}$ in Eq. (3.9), respectively,

$$\begin{aligned}\mathbf{R}^{\mathrm{e}}_{\mathrm{J}_k}\mathbf{u}^{\mathrm{e}}_{\mathrm{w}_i} &= \Lambda_{\mathrm{w}ii}\mathbf{u}^{\mathrm{e}}_{\mathrm{w}_i}\\ \mathbf{R}^{\mathrm{e}}_{\mathrm{J}_k}\mathbf{U}^{\mathrm{e}}_{\mathrm{w}} &= \mathbf{\Lambda}^{\mathrm{e}}.\end{aligned} \tag{3.11}$$

where $\mathbf{u}^{\mathrm{e}}_{\mathrm{w}_i}$ is the ith column of $\mathbf{U}^{\mathrm{e}}_{\mathrm{w}}$, $\Lambda_{\mathrm{w}ii}$ is the ith diagonal element of $\mathbf{\Lambda}_{\mathrm{w}}$, and $\mathbf{\Lambda}^{\mathrm{e}} \in \mathbb{C}^{N_k\times(N_k-N_J)}$ denotes a matrix with the ith column equal to $\Lambda_{\mathrm{w}ii}\mathbf{u}^{\mathrm{e}}_{\mathrm{w}_i}$.

Theorem 3.1. $\mathbf{R}^{\mathrm{e}}_{\mathrm{J}_k}$ *converges in probability [83, p. 175] to* $(\mathbf{Z}_k\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}}\mathbf{Z}^H_k + \sigma^2_{\mathrm{w}}\boldsymbol{I}_{N_k})$.

Proof: The proof is given in Appendix A.1. ■

Assuming N^{e} is sufficiently large such that the law of large number is applicable, from (3.11),

$$\begin{aligned}(\mathbf{Z}_k\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}}\mathbf{Z}^H_k + \sigma^2_{\mathrm{w}}\mathbf{I}_{N_k})\mathbf{U}^{\mathrm{e}}_{\mathrm{w}} &= \mathbf{\Lambda}^{\mathrm{e}}\\ \mathbf{Z}_k\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}}\mathbf{Z}^H_k\mathbf{U}^{\mathrm{e}}_{\mathrm{w}} &= \mathbf{\Lambda}^{\mathrm{e}} - \sigma^2_{\mathrm{w}}\mathbf{U}^{\mathrm{e}}_{\mathrm{w}}\\ \hat{\mathbf{G}}^{\mathrm{e}}_k\mathbf{Z}_k\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}}\mathbf{Z}^H_k &= (\mathbf{\Lambda}^{\mathrm{e}} - \sigma^2_{\mathrm{w}}\mathbf{U}^{\mathrm{e}}_{\mathrm{w}})^H\\ \hat{\mathbf{G}}^{\mathrm{e}}_k\mathbf{Z}_k &= \frac{(\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{Jf}})^T}{\det(\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}})}(\mathbf{\Lambda}^{\mathrm{e}} - \sigma^2_{\mathrm{w}}\mathbf{U}^{\mathrm{e}}_{\mathrm{w}})^H\mathbf{Z}_k(\mathbf{Z}^H_k\mathbf{Z}_k)^{-1},\end{aligned} \tag{3.12}$$

where $\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{Jf}}$ denotes the cofactor matrix [79, p. 275] of $\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}}$. The matrix $\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}}$ can be expressed by the variances of the transmitted jamming signals and the correlations between them,

$$\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}} = \begin{bmatrix} \sigma_1^2 & \rho^{\mathrm{e}}_{12}\sigma_1\sigma_2 & ... & \rho^{\mathrm{e}}_{1N_J}\sigma_1\sigma_{N_J} \\ (\rho^{\mathrm{e}}_{12})^*\sigma_1\sigma_2 & \sigma_2^2 & ... & \rho^{\mathrm{e}}_{2N_J}\sigma_2\sigma_{N_J} \\ ... & ... & ... & ... \\ (\rho^{\mathrm{e}}_{1N_J})^*\sigma_1\sigma_{N_J} & (\rho^{\mathrm{e}}_{2N_J})^*\sigma_2\sigma_{N_J} & ... & \sigma^2_{N_J} \end{bmatrix}, \tag{3.13}$$

where σ_j^2 is the variance of the jth transmitted jamming signal, and ρ^{e}_{ij} is the complex correlation coefficient between the ith and jth transmitted jamming signals in the inter-frame silent period, with $i, j \in \{1, 2, ...N_J\}$. The complex correlation coefficient ρ^{e}_{ij} can be calculated as [75, p. 87]

$$\rho^{\mathrm{e}}_{ij} = \frac{\mathbb{E}(\mathbf{X}^{\mathrm{e}}_{\mathrm{J}i}(\mathbf{X}^{\mathrm{e}}_{\mathrm{J}j})^H)}{\sigma_i\sigma_j}. \tag{3.14}$$

The correlation matrix [83, p. 139] of the transmitted jamming signals can be expressed as

$$\mathbf{C}^{\mathrm{e}} = \begin{bmatrix} 1 & \rho^{\mathrm{e}}_{12} & ... & \rho^{\mathrm{e}}_{1N_J} \\ (\rho^{\mathrm{e}}_{12})^* & 1 & ... & \rho^{\mathrm{e}}_{2N_J} \\ ... & ... & ... & ... \\ (\rho^{\mathrm{e}}_{1N_J})^* & (\rho^{\mathrm{e}}_{2N_J})^* & ... & 1 \end{bmatrix} \tag{3.15}$$

and can be calculated from $\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}}$ by

$$\mathbf{C}^{\mathrm{e}} = \mathbf{D}^{-1}\mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}}\mathbf{D}^{-1}, \tag{3.16}$$

where $\mathbf{D}$ is a $N_J \times N_J$ diagonal matrix with the jth diagonal element being σ_j, and is referred to as the standard deviation matrix of the transmitted jamming signals.

Observation on the elements of $\hat{\mathbf{G}}_k^{\mathrm{e}}\mathbf{Z}_k$*:* from (3.12), it can be seen that the elements of $\hat{\mathbf{G}}_k^{\mathrm{e}}\mathbf{Z}_k$ increase or decrease without bound when $|\rho_{ij}^{\mathrm{e}}| \to 1$, where $|.|$ denotes the modulus of a complex number. It is because when $|\rho_{ij}^{\mathrm{e}}| \to 1$, the correlation matrix $\mathbf{C}^{\mathrm{e}}$ becomes singular. Therefore, from (3.16), $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}}$ becomes a singular matrix. As a result, $\det(\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}}) \to 0$. This makes the elements of $\hat{\mathbf{G}}_k^{\mathrm{e}}\mathbf{Z}_k$ increase or decrease without bound.

This observation is demonstrated in Fig. 3.3 which shows the normalized residual value, $r_{\mathrm{c}} = ||\hat{\mathbf{G}}_k^{\mathrm{e}}\mathbf{Z}_k||^2/||\mathbf{Z}_k||^2$, as a function of $|\rho_{12}^{\mathrm{e}}|$ using different distribution functions for the transmitted jamming signals, including the Gaussian, uniform, and Chi-square (with two degrees-of-freedoms) distribution functions. The number of jammer is $N_J = 2$, the UE antennas is $N_k = 8$, and $||.||$ denotes matrix's Frobenius norm. As can be seen, without noise, $r_{\mathrm{c}} = 0$ for any value of $|\rho_{12}^{\mathrm{e}}|$ except for $|\rho_{12}^{\mathrm{e}}| = 1$. It is because when $\sigma_{\mathrm{w}}^2 = 0$, all the diagonal elements of $\mathbf{\Lambda}_{\mathrm{w}}^{\mathrm{e}}$ are zero, and (3.12) becomes

$$\hat{\mathbf{G}}_k^{\mathrm{e}}\mathbf{Z}_k = \frac{(\mathbf{\Sigma}_{\mathrm{Jf}}^{\mathrm{e}})^T}{\det(\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}})}\mathbf{0}_{(N_k-N_J,N_k)}\mathbf{Z}_k(\mathbf{Z}_k^H\mathbf{Z}_k)^{-1},$$

where $\mathbf{0}_{(N_k-N_J,N_k)} \in \mathbb{C}^{(N_k-N_J)\times N_k}$ denotes a null matrix. On the other hand, in the presence of noise with a received signal-to-noise ratio (SNR) of 14 dB, r_{c} is small and stable for small values of $|\rho_{12}^{\mathrm{e}}|$. However, r_{c} increases dramatically when $|\rho_{12}^{\mathrm{e}}|$ approaches 1, meaning the transmitted jamming signals are highly correlated.

Therefore, when the jamming signals are uncorrelated, the elements of $\hat{\mathbf{G}}_k^{\mathrm{e}}\mathbf{Z}_k$ are small, meaning $\hat{\mathbf{G}}_k^{\mathrm{e}}$ is close to $\mathbf{G}_k$. Similarly, $\mathbf{G}_k^{\mathrm{d}}$ is close to $\mathbf{G}_k$ when the jamming signals are uncorrelated. Therefore, $\hat{\mathbf{G}}_k^{\mathrm{e}}$ is close to $\mathbf{G}_k^{\mathrm{d}}$ and can be used to nullify $\mathbf{y}_{\mathrm{J}_k}^{\mathrm{d}}$. However, when the correlations between transmitted jamming signals are large, there are large values in the elements of $\hat{\mathbf{G}}_k^{\mathrm{e}}\mathbf{Z}_k$, meaning $\hat{\mathbf{G}}_k^{\mathrm{e}}$ is not close to $\mathbf{G}_k$. Therefore, $\hat{\mathbf{G}}_k^{\mathrm{e}}$ may not be close to $\mathbf{G}_k^{\mathrm{d}}$ and using $\hat{\mathbf{F}}_k^{\mathrm{e}}$ derived from $\hat{\mathbf{G}}_k^{\mathrm{e}}$ may not guarantee jamming nullification.

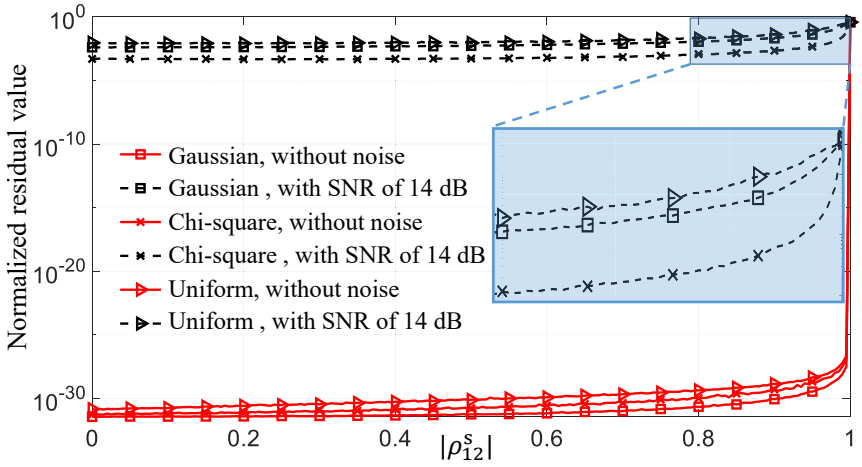


Figure 3.3 : r_{c} values for different $|\rho_{12}^{\mathrm{e}}|$.

Accordingly, different techniques to suppress the jamming signals for different correlations between transmitted jamming signals are needed. In the next section, the case of uncorrelated jamming signals (i.e., the correlations are zero) is considered, which is the assumption in most existing studies [25–29]. Then, in Section 3.5, different scenarios of non-zero correlations are analyzed.

3.4 Jamming Nullification with Uncorrelated Transmitted Jamming Signals and Time-varying Channels

This section describes jamming nullification when the jamming signals are uncorrelated, which is an assumption in most of the existing jamming suppression techniques [25–29]. The technique in this section is developed to deal with time-varying channels.

3.4.1 Protocol for Time-varying Channels

The baseline protocol in Fig. 3.2 is revised into that in Fig. 3.4 and Alg. 3.1 to deal with time-varying BS-UE and jamming channels. Similar to the protocol in Fig. 3.2, the protocol in Fig. 3.4 also consists of an initial silent period for the

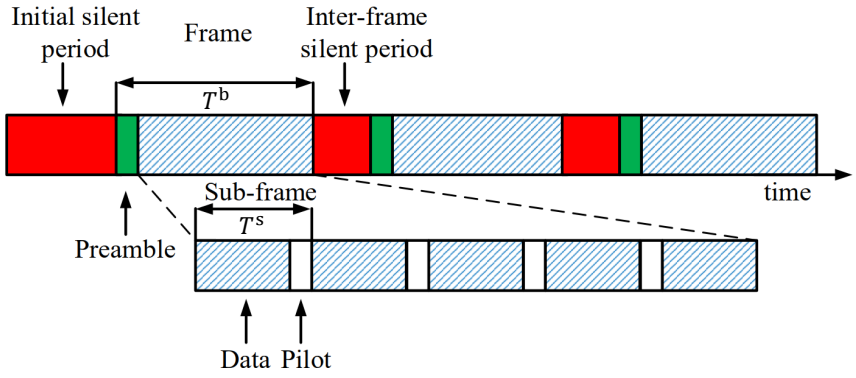


Figure 3.4 : Protocol for time-varying channels.

acquisition of $\hat{\mathbf{G}}_k^{\mathrm{e}}$, and inter-frame silent periods between frames to update $\hat{\mathbf{G}}_k^{\mathrm{e}}$. Each frame starts with a preamble which contains a short training sequence (STS) and a long training sequence (LTS). The STS is used for frame detection and fine timing, and carrier frequency offset (CFO) estimation. The LTS is used for BS-UE channel estimation. However, different from the baseline protocol in Fig. 3.2, each frame in Fig. 3.4 is divided into N^{sf} sub-frames, each of length $T^{\mathrm{s}} = T^{\mathrm{b}}/N^{\mathrm{sf}}$. Each sub-frame consists of data payload and pilot samples $\mathbf{x}_k^{\mathrm{m}}$, which is used for updating equalization coefficients $\mathbf{K}_k$ by the adaptive equalizer described below.

3.4.2 Adaptive Equalization

To deal with time-varying UE channels, it is suggested to use one of several optional adaptive equalizers, whose equalization coefficients can be adapted to channel variations. There have been a few types of adaptive filters being reported in the literature, such as the modified complex gradient-projection-II (MCG-P-II) [84] and the adaptive decision feedback equalizer (DFE) for V-BLAST systems [85]. Based on the DFE, an adaptive equalizer is also developed with the recursive least square (RLS) principle, without using decision/pilot-based feedback. This is realized by removing the feed-backward filter but only keeping the feed-forward filter in the

Algorithm 3.1 Protocol for time-varying channels.

1: Acquire $\hat{\mathbf{G}}_k^{\mathrm{e}}$ during the initial silent period or update $\hat{\mathbf{G}}_k^{\mathrm{e}}$ during inter-frame silent periods.
2: Design $\hat{\mathbf{F}}_k^{\mathrm{e}}$ from $\hat{\mathbf{G}}_k^{\mathrm{e}}$.
3: Estimate $\tilde{\mathbf{H}}_k$ using the preamble.
4: Calculate equalization coefficients $\mathbf{K}_k$.
5: Perform BS-UE data transmission.
6: $n^{\mathrm{sf}} \leftarrow 1$
7: **while** $n^{\mathrm{sf}} \leq N^{\mathrm{sf}} - 1$ **do**
8: Update $\mathbf{K}_k$ using $\mathbf{x}_k^{\mathrm{m}}$.
9: Perform BS-UE data transmission.
10: $n^{\mathrm{sf}} \leftarrow n^{\mathrm{sf}} + 1$
11: **end while**
12: Repeat from 1.

DFE as described in [85]. It is called RLS-FFE. Compared to DFE, the RLS-FFE has lower computational complexity, and is shown to work well using simulation.

In summary, when the transmitted jamming signals are uncorrelated, the communication protocol represented in Fig. 3.4 and Algorithm 3.1 is used to nullify the jamming signals and perform BS-UE data transmission. In the next section, the case with non-zero correlations between transmitted jamming signals is considered.

3.5 Jamming Nullification with Correlated Transmitted Jamming Signals

In this section, jamming nullification when the transmitted jamming signals are correlated is investigated. Similar to Section 3.3, to emphasize the impact of the correlations, it is assumed that the jamming channels are unchanged over each frame, and only $\mathbf{Z}_k$ is used to denote the jamming channels. First, it is demonstrated that the change in the correlations between transmitted jamming signals causes a "virtual change" in the jamming channels. Therefore, even though $\mathbf{Z}_k$ is unchanged, using

$\hat{\mathbf{F}}_k^{\mathrm{e}}$ deriving from $\hat{\mathbf{G}}_k^{\mathrm{e}}$ may not guarantee jamming nullification performance. Then, jamming nullification is analyzed for two schemes of non-zero correlations between transmitted jamming signals.

Theorem 3.2. *Let*

$$\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}} = \mathbf{V}^{\mathrm{e}}\mathbf{S}^{\mathrm{e}}(\mathbf{V}^{\mathrm{e}})^H \text{ and } \mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{d}} = \mathbf{V}^{\mathrm{d}}\mathbf{S}^{\mathrm{d}}(\mathbf{V}^{\mathrm{d}})^H$$

be the SVD of $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}}$ *and* $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{d}}$*, respectively. Let* $\mathbf{D}$ *be the "virtual change" factor given by*

$$\mathbf{D} = \mathbf{V}^{\mathrm{d}}\sqrt{\mathbf{S}^{\mathrm{d}}(\mathbf{S}^{\mathrm{e}})^{-1}}(\mathbf{V}^{\mathrm{e}})^H. \tag{3.17}$$

Then, the change over time from $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}}$ *to* $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{d}}$ *causes a "virtual change" in the jamming channels from* $\mathbf{Z}_k$ *to* $(\mathbf{Z}_k\mathbf{D})$*.*

Proof: The proof is given in Appendix A.2. ■

Examining the behavior of $\mathbf{D}$ provides us interesting insights about the impact of the change in the correlations between transmitted jamming signals on the "virtual change" in the jamming channels. The first observation is for the case of unchanged (i.e., over each frame) correlations, such that $\mathbf{C}^{\mathrm{d}} = \mathbf{C}^{\mathrm{e}}$. In this case, $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{d}} = \mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}}$, $\mathbf{V}^{\mathrm{d}} = \mathbf{V}^{\mathrm{e}}, \mathbf{S}^{\mathrm{d}} = \mathbf{S}^{\mathrm{e}}$, and $\mathbf{D} = \mathbf{I}$. Therefore, $\mathbf{Z}_k\mathbf{D} = \mathbf{Z}_k$, and there is no "virtual change" in the jamming channels. The case of unchanged correlations between transmitted jamming signals will be considered in more detail in Subsection 3.5.1. The second observation is for the case of time-varying correlations, and is described by the following corollary.

Corollary 2.1***:*** When $|\rho_{ij}^{\mathrm{e}}| \to 1$ and $\rho_{ij}^{\mathrm{d}} \neq \rho_{ij}^{\mathrm{e}}$, the elements of the "virtual change" factor $\mathbf{D}$ increase or decrease without bound.

Proof: when $|\rho_{ij}^{\mathrm{e}}| \to 1$, $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}}$ is a singular matrix. Therefore, there is one diagonal element $s_{jj}^{\mathrm{e}} \in \mathbf{S}^{\mathrm{e}}$ approaches 0 [86, p. 261]. From (3.17), when $s_{jj}^{\mathrm{e}} \in \mathbf{S}^{\mathrm{e}}$ approach 0

while $s^{\mathrm{d}}_{jj} \in \mathbf{S}^{\mathrm{d}}$ is not equal to s^{e}_{jj}, the elements of $\mathbf{D}$ increase or decrease without bound. ■

Therefore, there is a "virtual change" in the jamming channels observed by the receiver when the correlations between transmitted jamming signals vary over time. The "virtual change" is significant when $|\rho^{\mathrm{e}}_{ij}| \to 1$ for $i \neq j$. It implies that, from the receiver's observation, the estimated left nullspace becomes ineffective when the correlation is high and largely changed. Therefore, using $\hat{\mathbf{G}}^{\mathrm{e}}_k$ to generate $\hat{\mathbf{F}}^{\mathrm{e}}_k$ may not guarantee jamming nullification in the data transmission phase. The case of time-varying correlations between transmitted jamming signals will be considered in more detail in Subsection 3.5.2.

3.5.1 Jamming Nullification with Unchanged Non-zero Correlations

In this subsection, the case with $T^{\rho} \geq T^{\mathrm{b}}$ is considered, where T^{ρ} denotes the correlation coherence time, over which the change in the correlations is negligible, and T^{b} is the time interval of one frame as illustrated in Fig. 3.4.

As described in the first observation on the behavior of $\mathbf{D}$, when $\mathbf{C}^{\mathrm{d}} = \mathbf{C}^{\mathrm{e}}$, we have $\mathbf{\Sigma}^{\mathrm{d}}_{\mathrm{J}} = \mathbf{\Sigma}^{\mathrm{e}}_{\mathrm{J}}$, $\mathbf{V}^{\mathrm{d}} = \mathbf{V}^{\mathrm{e}}, \mathbf{S}^{\mathrm{d}} = \mathbf{S}^{\mathrm{e}}$, and $\mathbf{D} = \mathbf{I}$. Therefore, there is no "virtual change" in the jamming channels. As a result, the beam-forming matrix $\hat{\mathbf{F}}^{\mathrm{e}}_k$, which is derived from $\hat{\mathbf{G}}^{\mathrm{e}}_k$, can be used to nullify $\mathbf{y}^{\mathrm{d}}_{\mathrm{J}_k}$, regardless of the value of the correlations between transmitted jamming signals.

Therefore, jamming signals with non-zero and unchanged correlations can be suppressed in the same manner as the uncorrelated jamming signals in Section 3.4. In these cases, only the time-varying channels have to be handled, but not the time-varying correlations between transmitted jamming signals. Hence, the protocol in Fig. 3.4 and Algorithm 3.1 can be used to nullify the jamming signals and perform BS-UE communication. In the following sub-section, the case with time-varying non-zero correlations is investigated.

3.5.2 Jamming Nullification with Time-varying Non-zero Correlations

In this sub-section, the case with $T^{\rho} < T^{\mathrm{b}}$ is considered, meaning the correlations between transmitted jamming signals vary rapidly enough, such that during the frame interval T^{b}, the change in the correlations is not negligible. As stated in Theorem 2, the change in the correlations causes a "virtual change" in the jamming channels, characterized by the "virtual change" factor

$$\mathbf{D} = \mathbf{V}^{\mathrm{d}} \sqrt{\mathbf{S}^{\mathrm{d}} (\mathbf{S}^{\mathrm{e}})^{-1}} (\mathbf{V}^{\mathrm{e}})^{H}.$$

Moreover, from Corollary 2.1, when $\mathbf{C}^{\mathrm{d}} \neq \mathbf{C}^{\mathrm{e}}$ and the absolute value of one or more off-diagonal element of $\mathbf{C}^{\mathrm{e}}$ approach 1, there is a significant "virtual change" in the jamming channels. The significant "virtual change" in the jamming channels (now $\mathbf{Z}_k\mathbf{D}$) make it unable to use $\hat{\mathbf{F}}_k^{\mathrm{e}}$, which derived from $\hat{\mathbf{G}}_k^{\mathrm{e}}$, to nullify the jamming signals in the data transmission phase $\mathbf{y}_{\mathrm{J}_k}^{\mathrm{d}}$.

A potential solution to the change in the correlations is to decrease the frame time interval to $T^{\mathrm{b}} \leq T^{\rho}$. However, more frequent tracking of $\hat{\mathbf{G}}_k^{\mathrm{e}}$ means less BS-UE data transmission time. A suitable value of T^{b} should be based on two factors: the jamming channels' coherence time T^{c} and the correlation coherence time T^{ρ}. For example, the longer T^{c} or the longer T^{ρ}, the longer T^{b} should be. While the first factor depends on the nature of the jamming channels environment, and its value can be predicted based on field measurement, we cannot control or predict the second factor.

Given the above, a communication protocol capable of dealing with the change of the correlations between transmitted jamming signals is developed, as shown in Fig. 3.5 and Algorithm 3.2. This protocol has two differences compared to that in Fig. 3.4 and Alg. 3.1. First, the pilot signal $\mathbf{x}_k^{\mathrm{m}}$, demonstrated by the white color in Fig. 3.5, in addition to the first use to update equalization coefficient $\mathbf{K}_k$, is also used to measure the jamming residual. Second, when the measured jamming

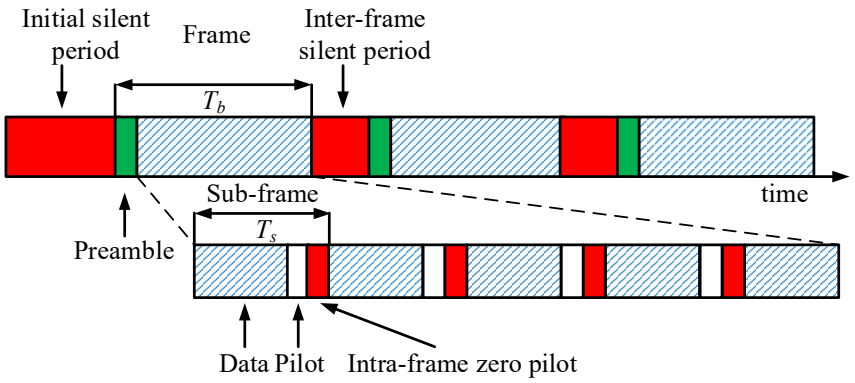


Figure 3.5 : Protocol for time-varying correlations and time-varying channels.

residual is higher than a predefined value, intra-frame zero pilots are used to update the beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{e}}$, as described below. By using the pilot signal $\mathbf{x}_k^{\mathrm{m}}$, the jamming residual can be measured by

$$\mathbf{r}_k = \mathbf{p}_k - \hat{\mathbf{F}}_k^{\mathrm{e}}(\sqrt{P_{\mathrm{T}}}\mathbf{H}_k\mathbf{P}_k\mathbf{x}_k^{\mathrm{m}} + \sqrt{P_{\mathrm{T}}}\mathbf{H}_k\sum_{l\neq k}^{K}\mathbf{P}_l\mathbf{x}_l) = \hat{\mathbf{F}}_k^{\mathrm{e}}(\mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \mathbf{w}).$$

When the normalized residual value $\gamma_{\mathrm{J}} = ||\mathbf{r}_k||^2/||\mathbf{p}_k||^2$ exceeds a pre-defined threshold γ^{r}, the BS starts sending intra-frame zero pilots, and $\hat{\mathbf{G}}_k^{\mathrm{i}}$ is estimated using the SVD as described in Section 3.3. As shown in Table 5.1, $\hat{\mathbf{G}}_k^{\mathrm{i}}$ denotes a matrix whose rows form an orthonormal basis for the left nullspace of the received signal at the kth UE in the intra-frame silent period. The new beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{i}}$ is obtained from $\hat{\mathbf{G}}_k^{\mathrm{i}}$, which is the estimated value of $\hat{\mathbf{G}}_k^{\mathrm{i}}$. Note that a smaller value of γ^{r} results in better jamming nullification, and, therefore, less error in the BS-UE data transmission. However, reducing the normalized residual threshold also results in more intra-frame zero pilot activation, and less time for BS-UE communication. Note also that starting the intra-frame zero pilots only reduces the spectral efficiency of the UEs that do not request it. However, the communication between the UE that needs intra-frame zero pilots and the BS will be completely lost if the intra-frame zero pilots are not activated. Therefore, the BS starts the intra-frame zero pilots

Algorithm 3.2 Protocol for time-varying correlations and time-varying channels.

1: Acquire $\hat{\mathbf{G}}_k^{\mathrm{e}}$ during the initial silent period or update $\hat{\mathbf{G}}_k^{\mathrm{e}}$ during inter-frame silent periods.
2: Design $\hat{\mathbf{F}}_k^{\mathrm{e}}$ from $\hat{\mathbf{G}}_k^{\mathrm{e}}$.
3: Estimate $\tilde{\mathbf{H}}_k$ using the preamble.
4: Calculate equalization coefficients $\mathbf{K}_k$.
5: Perform BS-UE data transmission.
6: $n^{\mathrm{sf}} \leftarrow 1$
7: **while** $n^{\mathrm{sf}} \leq N^{\mathrm{sf}} - 1$ **do**
8: Update $\mathbf{K}_k$ using $\mathbf{x}_k^{\mathrm{m}}$.
9: Measure $\gamma_{\mathrm{J}} = ||\mathbf{r}_k||^2/||\mathbf{p}_k||^2$ using $\mathbf{x}_k^{\mathrm{m}}$.
10: **if** $\gamma_{\mathrm{J}} \geq \gamma^{\mathrm{r}}$ **then**
11: Start intra-frame zero pilots.
12: Estimate $\hat{\mathbf{G}}_k^{\mathrm{i}}$ during intra-frame zero pilot.
13: Design $\hat{\mathbf{F}}_k^{\mathrm{i}}$ from $\hat{\mathbf{G}}_k^{\mathrm{i}}$.
14: Update equivalent channel $\tilde{\mathbf{H}}_k \leftarrow \hat{\mathbf{F}}_k^{\mathrm{i}}(\hat{\mathbf{F}}_k^{\mathrm{e}})^H\tilde{\mathbf{H}}_k$.
15: Update beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{e}} \leftarrow \hat{\mathbf{F}}_k^{\mathrm{i}}$.
16: **end if**
17: Perform BS-UE data transmission.
18: $n^{\mathrm{sf}} \leftarrow n^{\mathrm{sf}} + 1$
19: **end while**
20: Repeat from 1.

whenever requested by one or more UE.

To reduce the system's overhead, the equivalent channel is only estimated using a training signal and an MMSE or LS estimator once per frame. For each sub-frame, when intra-frame zero pilot is used, $\tilde{\mathbf{H}}_k$ is replaced by $\tilde{\mathbf{H}}_k^{\mathrm{i}}$ given by

$$\tilde{\mathbf{H}}_k^{\mathrm{i}} = \hat{\mathbf{F}}_k^{\mathrm{i}}\mathbf{H}_k\mathbf{P}_k = \hat{\mathbf{F}}_k^{\mathrm{i}}(\hat{\mathbf{F}}_k^{\mathrm{e}})^H\hat{\mathbf{F}}_k^{\mathrm{e}}\mathbf{H}_k\mathbf{P}_k = \hat{\mathbf{F}}_k^{\mathrm{i}}(\hat{\mathbf{F}}_k^{\mathrm{e}})^H\tilde{\mathbf{H}}_k.$$

By using the pilots signal $\mathbf{x}_k^{\mathrm{m}}$ and intra-frame zero pilots, each frame with a time interval T^{b} is divided into N^{sf} sub-frames with a time interval T^{s}. This division significantly reduces the change in the correlations between transmitted jamming

signals and reduces the elements of the "virtual change" factor $\mathbf{D}$. Therefore, jamming nullification performance can be significantly improved.

It is worth noting that an inter-frame silent period can be started as soon as γ_{J} exceeds γ^{r}. However, starting a new inter-frame silent period requires the BS-UE synchronization and $\tilde{\mathbf{H}}_k$ estimation to be performed again. Therefore, valuable BS-UE communication time is reduced dramatically when T_ρ decreases. Therefore, using the intra-frame zero pilot as in Fig. 3.5 and Algorithm 3.2 to combat the change in the correlations between transmitted jamming signals is a more suitable approach to reduce the overhead of the system and increase BS-UE communication time.

3.6 Simulation result

This section presents the simulation results to validate the proposed schemes. The performance metrics include the normalized residual value $r_{\mathrm{J}} = \frac{||\hat{\mathbf{G}}_k^{\mathrm{e}} \mathbf{y}_{\mathrm{J}_k}^{\mathrm{d}}||^2}{||\mathbf{y}_{\mathrm{J}_k}^{\mathrm{d}}||^2}$ and the bit error rate (BER). The simulation parameters are as follows unless otherwise stated. A single carrier narrowband MIMO system is considered with a bandwidth of 200 kHz and a center frequency of 447 MHz. The symbol duration is 5 μs. Note that as reported in [87], the typical root mean square (RMS) delay spread value in an urban area is 0.73 μs, corresponding to a coherence bandwidth of about 685 kHz [72, p. 45]. The RMS delay spreads in the suburban and combined areas are smaller than in the urban area, corresponding to larger coherence bandwidths. A 16 Quadrature Amplitude Modulation (QAM) with no forward error correction (FEC) coding is used. Note that we do not use FEC to eliminate its impact on the system's performance. Both BS-UE and jammers-UE channels are modeled using a flat fast-fading Rician model from [78] with a Doppler frequency of $f_{\mathrm{d}} = 24.83$ Hz (i.e., corresponding to BS-UE and jammers-UE relative speeds of about 60 km/h), and a Rician factor of 5. The number of propagation paths for each signal is 5.

The path losses over the BS-UE channel and jammer-UE channel are both modeled using the COST 231 Hata model [77, p. 135] with $h_{\mathrm{BS}} = h_{\mathrm{J}} = 50\mathrm{m}$, $h_{\mathrm{UE}} = 10\mathrm{m}$, and $d_{\mathrm{BS}} = d_{\mathrm{J}} = 1\mathrm{km}$, where h_{BS}, h_{J}, h_{UE}, d_{BS}, and d_{J} denote the BS height, jammer height, UE height, BS-UE distance, and the distance from each jammer to the UE, respectively. The ratio of transmitted jamming power to BS transmitted power is 3. The BS and UE have a uniform linear array (ULA) with 12 and 8 antennas, respectively. The BS is communicating with $K = 4$ UEs. The number of samples in the initial silent periods, inter-frame zero pilots, intra-frame zero pilots, and $\mathbf{x}_k^{\mathrm{m}}$ are 64, 5, 5 and 5, respectively. Note that in order to obtain a fair comparison of the efficiency of the adaptive equalizer and the beam-forming matrix in jamming nullification, the same number of samples for inter-frame zero pilot, intra-frame zero pilots, and $\mathbf{x}_k^{\mathrm{m}}$ are used. All the simulation results are obtained by averaging over 500 independent trials.

3.6.1 Uncorrelated Jamming Signals

In this sub-section, simulation results for the case of uncorrelated transmitted jamming signals is provided. As demonstrated in Section 3.4, the inter-frame silent period is used to update the beam-forming matrix and one of four equalizers (i.e., ZF, MCG-P-II, DFE, and RLS-FFE equalizers) to update the equalizer coefficient $\mathbf{K}_k$. To demonstrate the effectiveness in suppressing the jamming signals of the beam-forming matrix and the adaptive equalizer, the following scenarios are simulated.

- *No jammer*: There is no jammer in the system;
- *Ideal*: Beam-forming matrix is derived directly from the known jamming channels;
- *Inter-frame*: Beam-forming matrix is acquired during the initial silent period and updated using inter-frame zero pilots;

Table 3.2 : BER performance for various equalizers.

Scenarios	**ZF**	**MCG-P-II**	**RLS-FFE**	**DFE**
No jammer	0.0026	0.0028	0.0029	0.0021
Ideal	0.0230	0.0234	0.0256	0.0286
Inter-frame	0.0610	0.0586	0.0499	0.0497
Initial	0.3168	0.1074	0.0811	0.0834

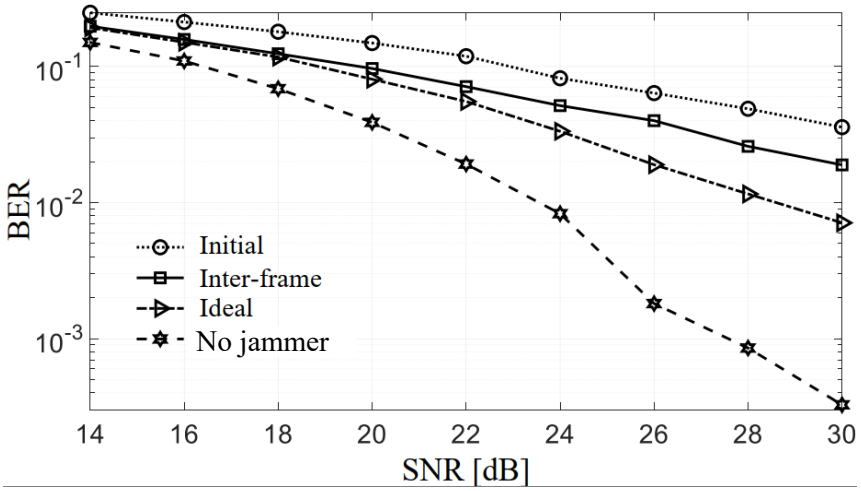


Figure 3.6 : BER performance of MCG-P-II.

- *Initial*: Beam-forming matrix is acquired during the initial silent period, without updating.

Table 3.2 shows the performance of different equalizers for different beam-forming matrix updating scenarios. The SNR used for simulation is 25 dB. As can be seen, in the presence of jammers, the adaptive equalizers (i.e., MCG-P-II, RLS-FFE, and DFE) all achieve better performance than the ZF equalizer for all of the beam-forming matrix updating scenarios. Moreover, the adaptive equalizers using the *initial* beam-forming matrix updating scenario achieve a higher BER compared to that

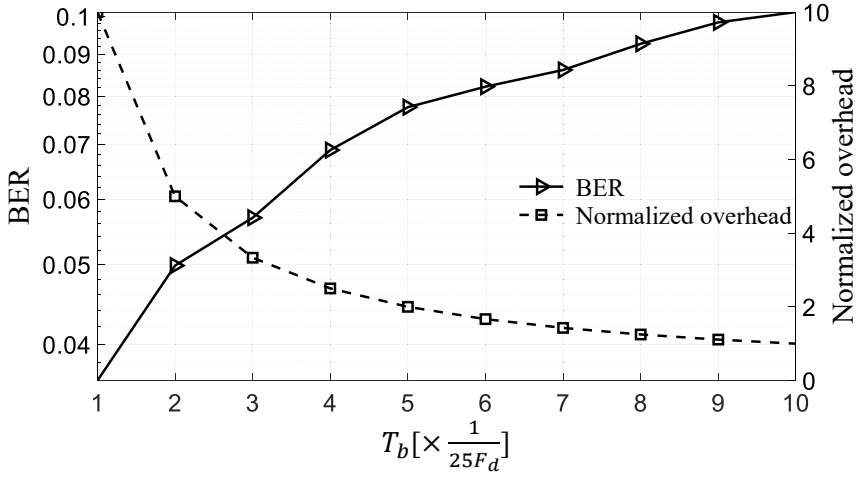


Figure 3.7 : BER and overhead trade-off for different values of T^{b}.

of the ZF equalizer using the *inter-frame* beam-forming matrix updating scenario. It means the beam-forming matrix is more efficient than the adaptive equalizer in dealing with jamming signals. It is because varied jamming nullspace mainly causes increased jamming, similar to noise, while adaptive equalizers are developed based on structured signal models. Among the adaptive equalizers, MCG-P-II achieves only slightly higher BER compared to the other ones. On the other hand, while MCG-P-II has a computational complexity of $\mathcal{O}((N_k - N_J)^2)$ [84], both DFE and RLS-FFE have $\mathcal{O}((N_k - N_J)^3)$ [85]. Therefore, MCG-P-II achieves a better balance between performance and complexity, and is regarded as the best option. Accordingly, from Fig. 3.6 onwards, the MCG-P-II equalized is used to generate simulation results.

Fig. 3.6 illustrates the MCG-P-II's BER performance with different beam-forming matrix updating scenarios and different SNR values. As can be seen, the *inter-frame* beam-forming matrix updating scenario can successfully nullify the jamming signals, resulting in system performance of only about 3 dB performance degradation compared to that of the system with *ideal* knowledge about the jamming channels. Note

that there is a high gap between the *Ideal* and *No jammer* scenario because even though a perfect jamming nullification can remove the jamming signal, it also costs N_{J} degree-of-freedoms.

Fig. 3.7 illustrates the BER performance and the normalized overhead of the system as functions of T^{b}, which quantifies how frequently the beam-forming matrix is updated. For comparison purposes, the overhead is normalized such that the normalized overhead when $T^{\mathrm{b}} = \frac{1}{25f_{\mathrm{d}}}$ is 1. As can be seen, there is a trade-off between the BER and the overhead of the system. When T^{b} increases, the BER of the system increases, whereas the overhead of the system decrease, and vice versa.

3.6.2 Unchanged Non-Zero Correlations

In this sub-section, the simulation results are presented for the case of non-zero correlations between transmitted jamming signals, and with $T^{\rho} \geq T^{\mathrm{b}}$, such that the correlations are considered constant over each frame. As discussed in Subsection 3.5.1, when the correlations between the transmitted jamming signals are non-zero and constant, there is no "virtual change" in the jamming channels. Therefore, the protocol in Fig. 3.4 and Alg. 3.1 can be used to nullify jamming signals and perform BS-UE communication.

Fig. 3.8 illustrates the values of r_{J} as a function of the jamming signals' SNR for different values of $\rho_{12}^{\mathrm{e}} = \rho_{12}^{\mathrm{d}} = \rho_{12}$. As can be seen, r_{J} has similar values for different $|\rho_{12}|$ values, meaning the value of $|\rho_{12}|$ has little impact on the jamming nullification performance. As a result, the system achieves similar BER performance for different values of $|\rho_{12}|$, as illustrated in Fig. 3.9.

3.6.3 Time-varying Correlations

In this sub-section, the simulation result for the case of time-varying correlations with $T^{\rho} < T^{\mathrm{b}}$ is presented. As discussed in Subsection 3.5.2, we have to deal with

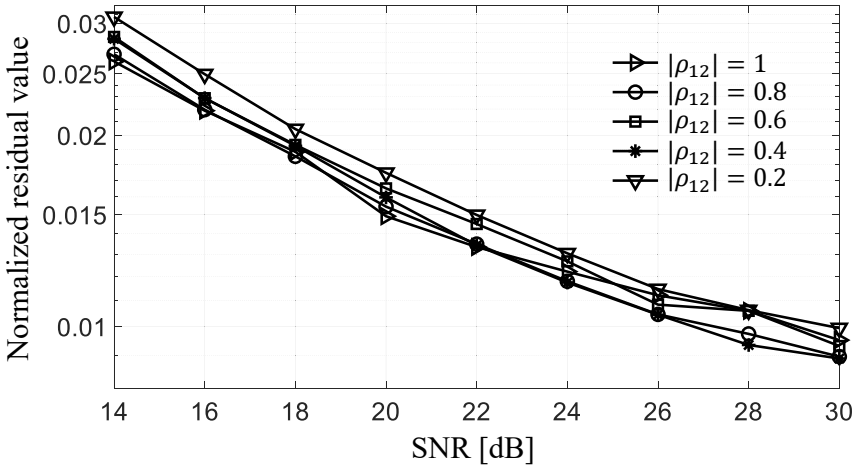


Figure 3.8 : r_{J} values when $|\rho_{12}|$ is unchanged.

both the time-varying channel and the "virtually change" in the jamming channels due to the change over time of the correlations between transmitted jamming signals. To deal with those issues, the protocol illustrated in Fig. 3.5 and Algorithm 3.2 is used with $N^{\mathrm{sf}} = 5$, and $\gamma^{\mathrm{r}} = 0.05$. Note that in this simulation, the correlation between transmitted jamming signals is assumed to change linearly. Therefore, the changes from $|\rho_{12}^{\mathrm{e}}|$ into $|\rho_{12}^{\mathrm{d}}|$ over one frame is equally divided into $N^{\mathrm{sf}} = 5$ changes over each sub-frame.

Fig. 3.10 illustrates the values of r_{J} for different scenarios of the time-varying correlations between transmitted jamming signals. As can be seen, there are large jamming residual values for the system using the protocol in Alg. 3.1 and Fig. 3.4. It means the above protocol is unable to catch "virtually change" in the jamming channels, caused by the change in the correlations between transmitted jamming signals. Another phenomenon is that the jamming residual becomes larger with the larger difference between $|\rho_{12}^{\mathrm{e}}|$ and $|\rho_{12}^{\mathrm{d}}|$. That is because a larger difference between $|\rho_{12}^{\mathrm{e}}|$ and $|\rho_{12}^{\mathrm{d}}|$ results in larger elements of the "virtual change" factor $\mathbf{D}$. On the other hand, the system using the protocol in Alg. 3.2 and Fig. 3.5 has successfully

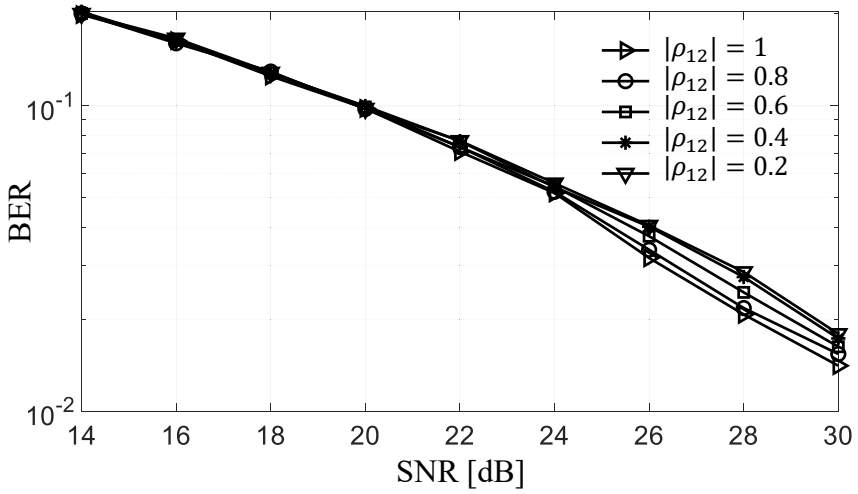


Figure 3.9 : BER performance when $|\rho_{12}|$ is unchanged.

caught up with the "virtually change" in the jamming channels, resulting in a much smaller jamming residual.

Fig. 3.11 compares the BER performance of the system using protocol in Alg. 3.1 and Fig. 3.4 to that of the protocol in Alg. 3.2 and Fig. 3.5 for different scenarios of the time-varying correlations mentioned above. As can be seen, the system using the protocol in Alg. 3.1 and Fig. 3.4 has high BER values, because the jamming signals are not effectively nullified. On the other hand, the system using the protocol in Alg. 3.2 and Fig. 3.5 successfully nullifies the jamming signals, resulting in an improvement of 3 dB.

3.7 Conclusion and Future Work

This chapter investigated the impact of non-zero and time-varying correlations between transmitted jamming signals on the jamming channels' left nullspace estimation and received jamming signals nullification. The technique to effectively suppress the received jamming signals for different levels of correlations between transmitted jamming signals was proposed. Monte Carlo simulation results show

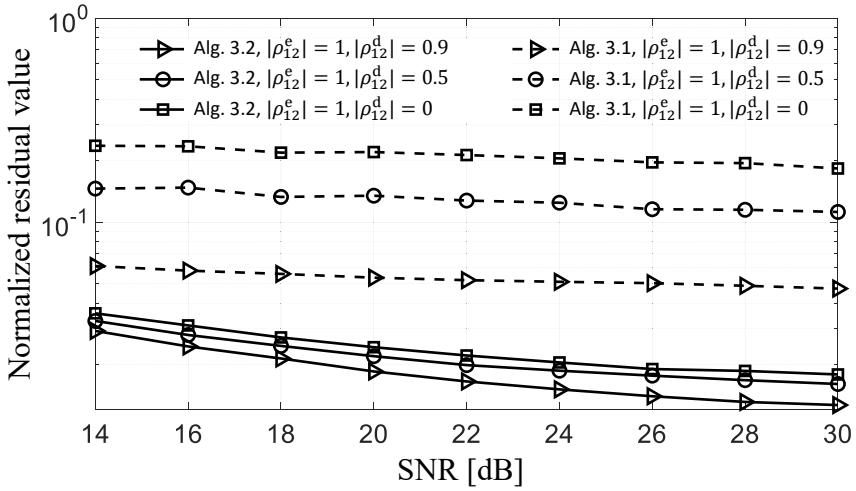


Figure 3.10 : r_{J} values for different values of $|\rho_{12}^{\mathrm{e}}|$ and $|\rho_{12}^{\mathrm{d}}|$.

that the proposed techniques can nullify the jamming signals for different values of the correlations. It is worth noting that the techniques proposed in this thesis can be extended to a wideband OFDM system with frequency-selective multipath channels by applying the technique to, e.g., each sub-carrier of the OFDM system. The future research problem is to design the parameters of the communication protocol analytically.

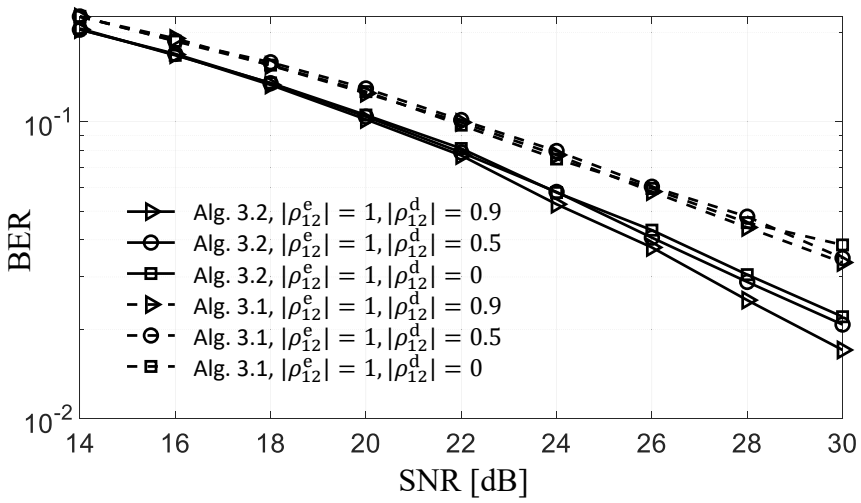


Figure 3.11 : BER for different values of $|\rho_{12}^{e}|$ and $|\rho_{12}^{d}|$.

Chapter 4

Multiple Correlated Jammers Nullification Using LSTM-based Deep Dueling Neural Network

To deal with the "virtual change" in the jamming channels, as shown in Chapter 3, one can continuously monitor the residual jamming signals and then heuristically adjust the estimated beam-forming matrix. However, such a jamming residual monitoring process incurs additional system overhead, thus significantly reducing the spectral efficiency. This gives rise to a more challenging problem in optimizing the duration of the nullspace estimation and the data transmission phases. A longer nullspace estimation phase may result in a lower jamming residual but then a shorter data transmission phase. Therefore, this chapter proposes a solution to nullify multiple correlated jammers whose correlation is unknown and time-varying, without requiring monitoring the jamming residual and then updating the estimated beam-forming matrix. To this end, the optimization problem of the nullspace estimation and data transmission phases is first systematically formulated. Even ignoring the unknown strategy of the jammers and the challenging nullspace estimation process, the resulting problem is an integer programming problem, hence intractable to obtain its optimal solution. In practice, as aforementioned, the jammers can deliberately vary the correlation range, making jamming nullification even more challenging. To deal with such uncertainty and incomplete information, as well as to circumvent the intractability of the above conventional optimization problem, the problem was reformulated using a partially observable semi-Markov decision process (POSMDP). Then, a deep dueling Q-learning-based technique [41, 42] is designed, which improves the training process by using two streams of fully connected

hidden layers to concurrently train the Q-learning algorithm, thereby quickly obtaining the optimal policy for the legitimate devices. The proposed technique does not require legitimate devices to constantly monitor the residual jamming signals, and only costs a single degree-of-freedom to nullify each jammer, even with an unknown and time-varying correlated jamming strategy. Unlike the strategy obtained by the game-theoretic approach [74, 88, 89], which may not be optimal, by leveraging the latest advances in deep reinforcement learning, this work aims to design an algorithm that converges to the optimal strategy by observing and learning from jammers' strategy. Simulation results show that the resulting spectral efficiency is about 4 dB higher than that of other methods and close to that of the perfect jamming nullification case.

4.1 Problem Formulation

In this section, the communication protocol and the estimated beam-forming matrix employed to suppress the jamming signals are described first. Then, the impacts of the time-varying correlations among jamming signals on the jamming suppression process (using the estimated beam-forming matrix) is briefly analyzed. Next, the upper and lower bounds for the spectral efficiency of each BS-UE communication link that employs the estimated beam-forming matrix to suppress the jamming signals are derived. Finally, the problem is mathematically stated. To highlight the impact of time-varying correlations on the estimated beam-forming matrix, it is assumed that the BS-UEs and the jamming channels follow a block-fading model with coherence time [37] T^{c}, corresponding to N^{c} samples. It is further assumed that the nullspace estimation, preamble, and data transmission phases of the communication protocol (described below) are performed within the interval T^{c}, such that $N^{\mathrm{e}} + N^{\mathrm{p}} + N^{\mathrm{d}} < N^{\mathrm{c}}$, where $N^{\mathrm{e}}, N^{\mathrm{p}}, N^{\mathrm{d}}$ are the number of samples of the nullspace estimation, preamble, and data transmission phase, respectively.

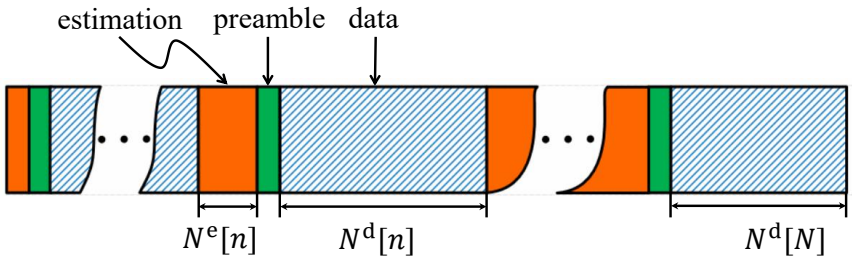


Figure 4.1 : Communication protocol for jamming suppression.

4.1.1 Communication Protocol

Fig. 4.1 illustrates the communication protocol for the jamming/interference nullification purpose. As mentioned in Section 3.2, each frame is comprised of three phases: nullspace estimation, preamble, and data transmission. During the data transmission phase, which lasts for N^{d} samples, the BS sends data to UEs. Let $\delta_{k,m}[n]$ denotes the post-equalization SINR of the mth stream for the kth UE during the data transmission phase of the nth frame. As will be demonstrated in Appendix B, $\delta_{k,m}$ can be given by

$$\delta_{k,m} = \frac{P_{\mathrm{T}}}{\mathbb{E}\{[\mathbf{A}_k^{\mathrm{ZF}}(\delta\mathbf{F}_k\mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \hat{\mathbf{F}}_k^{\mathrm{e}}\mathbf{w})]_m^2\}}. \tag{4.1}$$

where $\mathbf{A}_k^{\mathrm{ZF}}$ denotes the zero-forcing equalizer, $(.)_m$ denotes the mth elements of a vector, $\delta\mathbf{F}_k = (\hat{\mathbf{F}}_k^{\mathrm{e}} - \mathbf{G}_k^{\mathrm{d}})$ is the nullspace estimation error, where $\mathbf{G}_k^{\mathrm{d}}$ denotes a matrix whose rows form an orthonormal basis for the left nullspace of $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{d}}$.

Let $C_{k,m}[n]$, and $C_{k,m}^{\mathrm{eff}}[n]$ denote the corresponding spectral efficiency and effective spectral efficiency [55], respectively. We have

$$C_{k,m}[n] = \log_2\{1 + \delta_{k,m}[n]\}, \qquad C_{k,m}^{\mathrm{eff}}[n] = \mu[n]C_{k,m}[n], \tag{4.2}$$

where

$$\mu[n] = \frac{N^{\mathrm{d}}[n]}{N^{\mathrm{e}}[n] + N^{\mathrm{p}}[n] + N^{\mathrm{d}}[n]} \tag{4.3}$$

is the data transmission phase fraction over the whole frame.

4.1.2 Impact of the Time-varying Correlations on Jamming Suppression

The estimated beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{e}}$ described in the previous subsection is derived from $\hat{\mathbf{G}}_k^{\mathrm{e}}$, whose rows form an orthonormal basis for the left nullspace of $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}}$. Therefore, $\hat{\mathbf{F}}_k^{\mathrm{e}}$ should be capable of nullifying $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{d}}$ when the left nullspace of $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}}$ is the same to that of $\mathbf{Z}_k$. However, when $|\rho_{ij}^{\mathrm{e}}| \to 1$, the nullspace of $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}}$ is not the same to that of $\mathbf{Z}_k$, making the rows of $\hat{\mathbf{G}}_k^{\mathrm{e}}$ do not span the nullspace of $\mathbf{Z}_k$ [1]. In this case, $\hat{\mathbf{F}}_k^{\mathrm{e}}$, which is derived from $\hat{\mathbf{G}}_k^{\mathrm{e}}$, may not be capable to nullify $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{d}}$.

In fact, when the correlations are high but not time-varying (i.e., $\rho_{ij}^{\mathrm{e}} = \rho_{ij}^{\mathrm{d}} = \rho_{ij} \to 1$), the estimated beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{e}}$, despite of not being an accurate estimation of the nullspace of $\mathbf{Z}_k$, can still be used to nullify $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{d}}$. This is because unchanged correlation values lead to the same values of $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}}$ and $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{d}}$, and hence the same values of $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}}$ and $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{d}}$. Therefore, the rows of $\hat{\mathbf{G}}_k^{\mathrm{e}}$ still span the nullspace of $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{d}}$, even though they do not span the nullspace of $\mathbf{Z}_k$. Accordingly, $\hat{\mathbf{F}}_k^{\mathrm{e}}$, derived from $\hat{\mathbf{G}}_k^{\mathrm{e}}$, can be used to nullify $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{d}}$.

However, when the correlations are time-varying, (i.e., $\rho_{ij}^{\mathrm{d}} \neq \rho_{ij}^{\mathrm{e}}$ and $|\rho_{ij}^{\mathrm{e}}| \to 1$), the rows of $\hat{\mathbf{G}}_k^{\mathrm{e}}$ do not span the nullspace of $\mathbf{Z}_k$, and the covariance matrix of the received signal varies over time (i.e., $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{d}} \neq \mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}}$). In this case, the rows of $\hat{\mathbf{G}}_k^{\mathrm{e}}$ do not span the nullspace of $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{d}}$. Therefore, the time-varying correlation degrades the jamming nullification capability of $\hat{\mathbf{F}}_k^{\mathrm{e}}$. We consider this impact as if the correlations produced a "virtual change" in the jamming channel. Hence, when the correlations are time-varying, using $\hat{\mathbf{F}}_k^{\mathrm{e}}$ to suppress $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{d}}$ leads to significant jamming residual and correspondingly a low post-equalization SINR at the receiver in the data transmission phase. Theorem 1 below formally states the effectiveness of $\hat{\mathbf{F}}_k^{\mathrm{e}}$ in nullifying the jamming signals by examining the lower and upper spectral efficiency bounds of each BS-UE communication link. Here, for illustration purposes, we use the zero-forcing equalization at the UE receiver. Moreover, to focus on the impact

of jamming residual after the nullification process on the system performance, we assume that the BS-UE channel can be estimated perfectly. This is because the impact of the BS-UE channel estimation error on system performance has been well-studied [90, 91].

Theorem 4.1. *The spectral efficiency of the mth stream for the kth UE is bounded by*

$$C_{k,m}^{\mathrm{lb}} \leq C_{k,m} \leq C_{k,m}^{\mathrm{ub}}, \tag{4.4}$$

where $C_{k,m}^{\mathrm{lb}}$ and $C_{k,m}^{\mathrm{ub}}$ are the lower and upper bounds of $C_{k,m}$

$$C_{k,m}^{\mathrm{lb}} = \log_2 \big[1 + \frac{P_{\mathrm{T}}(N_k - N_{\mathrm{J}} - M_k)}{\eta_k(\sigma_{\mathrm{w}}^2 + \sum_{j=1}^{N_{\mathrm{J}}} \eta_{k,j}\sigma_{\mathrm{J}_j}^2)}\big], \tag{4.5}$$

$$C_{k,m}^{\mathrm{ub}} = \log_2 \big[1 + \frac{P_{\mathrm{T}}(N_k - N_{\mathrm{J}} - M_k)}{\eta_k\sigma_{\mathrm{w}}^2}\big]. \tag{4.6}$$

Proof: The proof is given in Appendix B.1. ■

Without nullifying/suppressing the jamming signals (or without beam-forming), the corresponding spectral efficiency, referred to as $C_{k,m}^{\mathrm{wbf}}$, is

$$C_{k,m}^{\mathrm{wbf}} = \log_2 \big[1 + \frac{P_{\mathrm{T}}(N_k - M_k)}{\eta_k(\sigma_{\mathrm{w}}^2 + \sum_{j=1}^{N_{\mathrm{J}}} \eta_{k,j}\sigma_{\mathrm{J}_j}^2)}\big]. \tag{4.7}$$

As can be seen, the use of the estimated beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{e}}$ does not always guarantee better spectral efficiency than not using it, i.e., $C_{k,m}^{\mathrm{wbf}} > C_{k,m}^{\mathrm{lb}}$. It is because when using $\hat{\mathbf{F}}_k^{\mathrm{e}}$, the UE receiver has to "sacrifice" N_{J} degree-of-freedoms, as demonstrated by comparing the numerators of Eq. (4.5) and Eq. (4.6) to that of Eq. (4.7). Especially, when the estimated beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{e}}$ is not accurate, its use can lead to the worst case with the lower-bound spectral efficiency. In this case, the UE receiver lost N_{J} degree-of-freedoms without achieving any jamming nullification effect, leaving the denominator of Eq. (4.5) unchanged compared to that of Eq. (4.7). On the other hand, when $\hat{\mathbf{F}}_k^{\mathrm{e}}$ is estimated perfectly (i.e., $\hat{\mathbf{F}}_k^{\mathrm{e}} = \mathbf{G}_k^{\mathrm{d}}$), the UE

receiver can completely nullify the jamming signals, leaving only the noise in the denominator of Eq. (4.6), thereby achieving the upper-bound spectral efficiency. In the sequel, the target is to accurately estimate the beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{e}}$, even when the correlations are time-varying; thereby achieving the perfect beam-forming, with the spectral efficiency of each stream close to the upper bound given in Eq. (4.6).

4.1.3 Problem Formulation

Given the above, one can maximize the spectral efficiency by continuously adapting the length of the nullspace estimation and data transmission phases (i.e., N^{e} and N^{d}, respectively). Specifically, N^{e} and N^{d} can be tuned based on the following principles.

- First, N^{e} and N^{d} are jointly optimized to guarantee that the estimation $\hat{\mathbf{F}}_k^{\mathrm{e}}$ can be obtained when none of $|\rho_{ij}^{\mathrm{e}}|$ is close to 1. As presented in Corollary 1.1, when $|\rho_{ij}^{\mathrm{e}}| \to 1$, the elements of $\mathbf{D}$ approach infinity, making the estimated beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{e}}$ unable to suppress the jamming signals, and hence resulting in a lower spectral efficiency.

- Second, by saving the time spent on monitoring the residual jamming signals as in Chapter 3 to update the estimated beam-forming matrix, the effective spectral efficiency of the system can be significantly improved, as demonstrated in Eq. (4.2). In fact, only when necessary, the system may increase the nullspace estimation time (by increasing N^{e}) to average the correlations between transmitted jamming signals and avoid $|\rho_{ij}^{\mathrm{e}}|$ being close to 1.

- Third, by adjusting N^{d}, the communication system can adapt to the change in the BS-UE channel condition. For example, when the channel coherence time decreases, the value of N^{d} should be decreased to maintain an acceptable post-

equalization SINR level (e.g., above the required minimum post-equalization SINR, below which the UE is considered to be an outage). On the other hand, when the coherence time increases, the system can increase N^{d} to improve the communication phase fraction over the whole frame, hence increasing the effective spectral efficiency.

The optimization of the durations of the nullspace estimation and data transmission phases can be formally stated as follows:

$$\max_{N^{\mathrm{e}},N^{\mathrm{d}}} \lim_{N\to\infty} \Big\{ \frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{K} \sum_{m=1}^{M_k} \mu[n] \log_2\{1+\delta_{k,m}[n]\} \Big\} \tag{4.8}$$

$$\begin{aligned}
\textbf{s.t.} \quad & \mu[n] = \frac{N^{\mathrm{d}}[n]}{N^{\mathrm{e}}[n]+N^{\mathrm{p}}[n]+N^{\mathrm{d}}[n]} \text{ as in (4.3)}, \\
& \delta_{k,m}[n] = \frac{P_{\mathrm{T}}}{\mathrm{Var}\{[\mathbf{A}_k^{\mathrm{ZF}}(\delta\mathbf{F}_k\mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \hat{\mathbf{F}}_k^{\mathrm{e}}\mathbf{w})]_m[n]\}} \text{ as in (4.1)}, \\
& \delta\mathbf{F}_k = \hat{\mathbf{F}}_k^{\mathrm{e}} - \mathbf{G}_k^{\mathrm{d}}, \quad \mathbf{A}_k^{\mathrm{ZF}} = (\tilde{\mathbf{H}}_k^H \tilde{\mathbf{H}}_k)^{-1} \tilde{\mathbf{H}}_k^H, \\
& \tilde{\mathbf{H}}_k = \hat{\mathbf{F}}_k^{\mathrm{e}} \mathbf{H}_k \mathbf{P}_k, \quad \delta_{k,m}[n] \geq \delta_{\min},
\end{aligned}$$

where N is the number of frames and $\delta_{\min}$ is the required minimum post-equalization SINR, below which the UE is considered to be an outage.

There are analogies between the estimation error of $\hat{\mathbf{F}}_k^{\mathrm{e}}$ and that of the BS-UE channel in MIMO communications (e.g., [90, 91]), because both errors lead to unwanted interference to the legitimate signals, thus reducing the post-equalization SINR and consequently the spectral efficiency. However, unlike [90, 91] in which the training signals and the channel estimator can be designed to quantify the mean and variance of the channel estimation error, the jamming signals in this scenario are controlled by the jammers (i.e., described by Eq. (3.5)). Therefore, the BS and UE do not have knowledge of the mean and variance of $\delta\mathbf{F}_k$. More importantly, to make jamming suppression even more challenging, the jammers can deliberately adjust the correlations controlling function $\mathcal{J}$, making the previous measurements

no longer representative of the current jamming strategy. To deal with such incomplete information and uncertainty, in the next section, the deep dueling Q-learning technique is used to solve the problem stated in Eq. (4.8).

4.2 Deep Dueling Q-Learning Technique for Jamming Suppression

This section reformulates the problem (4.8) using a POSMDP. A deep dueling Q-learning-based framework is then designed to tune the durations of the nullspace estimation and data transmission phases by obtaining the optimal policy for the underlying POSMDP process.

4.2.1 POSMDP

A conventional MDP is defined by a tuple $(\mathcal{S}, \mathcal{A}, r)$, where $\mathcal{S}$, $\mathcal{A}$, and r denote the state space, action space, and the reward function, respectively. An SMDP, on the other hand, retains the three components mentioned above and adds an additional component, that is the nth decision epoch length, denoted by $t[n]$. In an MDP, the state transition occurs at regular time steps (and hence the decision epoch length $t[n]$ is excluded). On the other hand, the SMDP allows the state transition to occur at irregular time steps (i.e., different $t[n]$ for different epochs), facilitating the selections of N^{e} and N^{d} at irregular state transition times. Note that the notification from the BS to the UE about N^{e} and N^{d} values only causes a negligible overhead to the system.. A POSMDP combines the observations of the SMDP to generate the approximate state, and is used when the state is only partially observable, as explained below.

4.2.1.1 State Space

There are several essential factors to consider for maximizing the effective spectral efficiency of the system while avoiding outages. The first factor is the post-equalization SINR at the UEs during the previous data transmission phase. This is because the post-equalization SINR implicitly captures the BS-UE channel condition that affects the selection of N^{d}. Specifically, when the post-equalization SINR in the previous data transmission phase is poor (e.g., below or close to $\delta_{\min}$), N^{d} value can be decreased to improve the post-equalization SINR. Hence, the system can avoid outages and increase spectral efficiency during the data transmission phase. However, N^{d} should not be incautiously decreased, as that reduces the fraction of the data transmission phase over the whole frame and thus reduces the effective spectral efficiency of the system. The second factor, as demonstrated in the previous section by Corollary 1, is the correlation between transmitted jamming signals in the nullspace estimation phase. This is because the correlations ρ^{e}_{ij} affect the magnitude of the "virtual change" factor $\mathbf{D}$ in the jamming channel, which directly affects the jamming nullification capability of $\hat{\mathbf{F}}^{\mathrm{e}}_{k}$. Therefore, the system's state space can be defined as

$$\mathcal{S} \triangleq \big\{(\delta_{k,m}, |\rho^{\mathrm{e}}_{ij}|) : \forall k \in (1,2,...,K), \\ \forall m \in (1,2,...,M_k), \forall i \neq j; i,j \in (1,2,...,N_{\mathrm{J}})\big\}. \tag{4.9}$$

4.2.1.2 Observation

Out of the two components of the state space, only the post-equalization SINR can be estimated at the UEs. The correlation coefficients, on the other hand, are controlled by the jammers (i.e., by Eq. (3.5)), and are neither known nor directly observable by the UEs or the BS. Therefore, the problem is formulated as a POS-MDP [65], where the state $\mathcal{S}$ in Eq. (4.9) is replaced by the approximate state [65,66]

$\hat{\mathcal{S}}$ derived from the observation. The observation space of the system is defined as,

$$\mathcal{O} \triangleq \big\{(\bar{\delta}_k, \bar{\Lambda}_l) : \forall k \in (1, 2, ..., K), \\ \forall m \in (1, 2, ..., M_k), \forall l \in (1, 2, ..., N_{\mathrm{J}})\big\}, \tag{4.10}$$

where $\bar{\delta}_k = (1/M_k) \sum_{m=1}^{M_k} \hat{\delta}_{k,m}$ is the average estimated post-equalization SINR, $\hat{\delta}_{k,m}$ is the estimated value of $\delta_{k,m}$, $\Lambda_{k,l}$ is the lth largest singular value of $\mathbf{R}^{\mathrm{e}}_{\mathrm{J}_k}$, and $\bar{\Lambda}_l = (1/K) \sum_{k=1}^{K} \Lambda_{k,l}$. As can be seen, K values of $\bar{\delta}_k$ are used instead of $K \times M_k$ values of $\hat{\delta}_{k,m}$ to reduce the observation size (i.e., and consequently the approximate state size). On the other hand, N_{J} values of $\bar{\Lambda}_l$ are used to represent $|\rho^{\mathrm{e}}_{ij}|$, which is the second component of the state space. Specifically, small $|\rho^{\mathrm{e}}_{ij}|$ results in relatively equal $\bar{\Lambda}_l$ values, while large $|\rho^{\mathrm{e}}_{ij}|$ results in massive differences between the $\bar{\Lambda}_l$ values.

While $\bar{\Lambda}_l$ can be calculated from $\Lambda_{k,l}$ values, which are readily available using the SVD of $\mathbf{R}^{\mathrm{e}}_{\mathrm{J}_k}$, the value of $\hat{\delta}_{k,m}$ can be estimated by [92]

$$\hat{\delta}_{k,m}[n] = 10 \log_{10} \frac{\sum_{i=1}^{N^{\mathrm{d}}[n]} (I^2_{k,m,i} + Q^2_{k,m,i})}{\sum_{i=1}^{N^{\mathrm{d}}[n]} (\delta I^2_{k,m,i} + \delta Q^2_{k,m,i})}, \tag{4.11}$$

where $\delta I_{k,m,i}$ and $\delta Q_{k,m,i}$ are the in-phase and quadrature absolute differences between the ith observed constellation point and its maximum likelihood estimation (MLE), respectively. On the other hand, $I_{k,m,i}$ and $Q_{k,m,i}$ are the in-phase and quadrature components of the MLE of the ith constellation point, respectively. These terms are demonstrated in more detail in Fig. 4.2. For each ith actual constellation point, the $\delta I_{k,m,i}$ and $\delta Q_{k,m,i}$ values are calculated by subtracting $I_{k,m,i}$ and $Q_{k,m,i}$ from the in-phase and quadrature components of the actual constellation point, respectively. Then, $\hat{\delta}_{k,m}$ can be calculated using Eq. (4.11). Note that, similar to the state $\mathcal{S}$ mentioned above, the observation $\mathcal{O}$ is composed of $\bar{\delta}_k$ and $\bar{\Lambda}_l$ values in the most recent frame, such that

$$o[n] \triangleq \big[\bar{\delta}_k[n-1], \bar{\Lambda}_l[n-1]\big]. \tag{4.12}$$

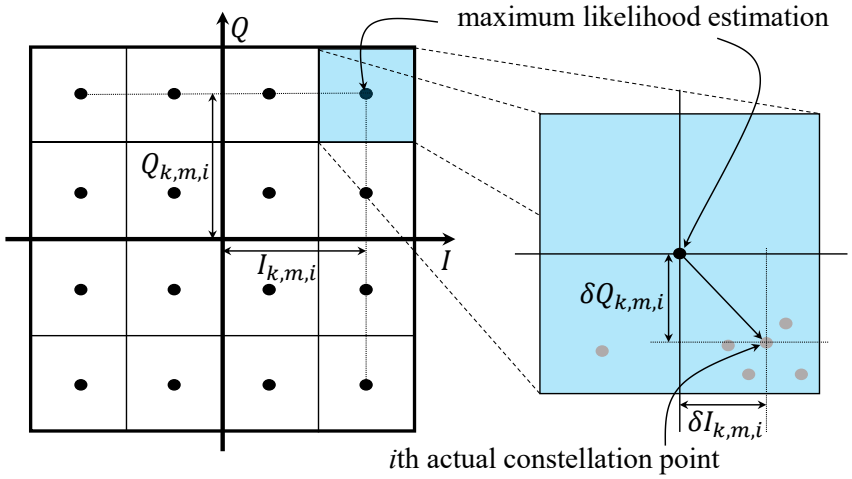


Figure 4.2 : Estimation of the post-equalization SINR.

4.2.1.3 Approximate State [65, 66]

As described above, the state $\mathcal{S}$ is replaced by the approximate state $\hat{\mathcal{S}}$ derived from the observations. Similar to [65, 66], the last H observations and actions are used as the approximate state, i.e., $\hat{s}[n] \triangleq \big[o[n], a[n-1], o[n-1], ..., a[n-H]\big]$, where H denotes the history length. This formalism, referred to as the Hth-order history approach, generates a large but finite POSMDP, in which each sequence is a distinct approximate state. As a result, we can apply standard reinforcement learning techniques used in MDPs or SMDPs to find the optimal action given the current approximate state.

4.2.1.4 Action

At the end of each frame, an action is taken to determine which are the next N^{e} and N^{d} values, given the current approximate state. Let $\mathcal{N}^{\mathrm{e}} \triangleq (N_1^{\mathrm{e}}, N_2^{\mathrm{e}}, ..., N_{L^{\mathrm{e}}}^{\mathrm{e}})$ and $\mathcal{N}^{\mathrm{d}} \triangleq (N_1^{\mathrm{d}}, N_2^{\mathrm{d}}, ..., N_{L^{\mathrm{d}}}^{\mathrm{d}})$ be the sets of L^{e} and L^{d} candidates for N^{e} and N^{d},

respectively. The action space is defined as $\mathcal{A} \triangleq \{a : a \in (1, 2, ..., L^{\mathrm{e}} \times L^{\mathrm{d}})\}$, and

$$a = \begin{cases} 1, & N^{\mathrm{e}} = N_1^{\mathrm{e}} \text{ and } N^{\mathrm{d}} = N_1^{\mathrm{d}}, \\ 2, & N^{\mathrm{e}} = N_2^{\mathrm{e}} \text{ and } N^{\mathrm{d}} = N_1^{\mathrm{d}}, \\ ... & \\ L^{\mathrm{e}} \times L^{\mathrm{d}}, & N^{\mathrm{e}} = N_{L^{\mathrm{e}}}^{\mathrm{e}} \text{ and } N^{\mathrm{d}} = N_{L^{\mathrm{d}}}^{\mathrm{d}}. \end{cases}$$

4.2.1.5 Immediate Reward

The immediate reward is defined as the amount of data transmitted during the data transmission phase, and zero if one (or more) post-equalization SINR value during the data transmission phase is smaller than the minimum required post-equalization SINR. Specifically,

$$r[n] = \begin{cases} \sum_{k=1}^{K} \sum_{m=1}^{M_k} N^{\mathrm{d}}[n] \log_2(1 + \hat{\delta}_{k,m}[n]), & \text{when } \hat{\delta}_{k,m}[n] \geq \delta_{\min}, \\ & \forall k \in (1, 2, ..., K), \forall m \in (1, 2, ..., M_k), \\ 0, & \text{otherwise.} \end{cases} \tag{4.13}$$

4.2.1.6 Optimization Formulation

Let $\pi : \hat{\mathcal{S}} \rightarrow \mathcal{A}$ denote a policy which is a mapping function from the approximate states to the actions taken by the system. The purpose is to find the optimal value of π, denoted by π^*, that maximizes the average long-term reward [42] of the BS-UE communication system. The optimization problem in Eq. (4.8) is then converted into the optimization problem of π^*, expressed by

$$\max_{\pi} \quad \mathcal{R}(\pi) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}\{r[n]\} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}\{r\{\hat{s}[n], \pi\{\hat{s}[n]\}\}\}, \tag{4.14}$$

where $\mathcal{R}(\pi)$ denotes the average long-term reward of the system with the policy π.

Note that, while the convergence of the Q-learning algorithm is guaranteed when the learning rate follows specific criteria [64], there is no such guarantee for the deep

dueling Q-learning technique. Let T_{J} and T_{Q} denote the time after that the jammers change their strategy and the convergence time of the proposed technique. When $T_{\mathrm{J}} \geq T_{\mathrm{Q}}$, the proposed technique converges to the optimal policy. On the other hand, when $T_{\mathrm{J}} < T_{\mathrm{Q}}$, the system does not have enough time to obtain the optimal policy, and hence the system performance is worse than the optimal one. However, because the system performance is improved over the training period (i.e., as demonstrated in the simulation result), the achieved performance is still better than that of the system without the proposed technique. In other words, the proposed technique still effectively improves the system's performance even without obtaining the optimal one.

It is also worth noting that when $T_{\mathrm{J}} < T_{\mathrm{Q}}$ for random jammers in general, the strategic interaction (attack and defense strategies) can be formulated as a non-cooperative game between the jammers and the legitimate devices, which has been well-investigated with a rich literature, e.g., [74, 88, 89, 93, 94]. Under such a model, there often exists a Nash equilibrium (NE). By sticking with the Nash equilibrium strategy, the legitimate devices can ensure that the jammers cannot cause further "damage" by unilaterally deviating from their strategy. In this case, the NE defense strategy of the legitimate nodes (at the NE) is not necessarily their optimal strategy corresponding to a specific jamming strategy of jammers (referred to as a "pure" strategy in game theory), which can be achieved by using the proposed technique).

4.2.2 Network Structure and Complexity Analysis

In this subsection, the Long Short-Term Memory (LSTM) [95]-based deep dueling Q-network used in the deep dueling Q-learning technique is introduced. The network is illustrated in Fig. 4.3. Unlike conventional recurrent neural networks (RNN) that have difficulty learning the long-term dependencies of the inputs [96], the LSTM is capable of learning those dependencies. That is because the LSTM

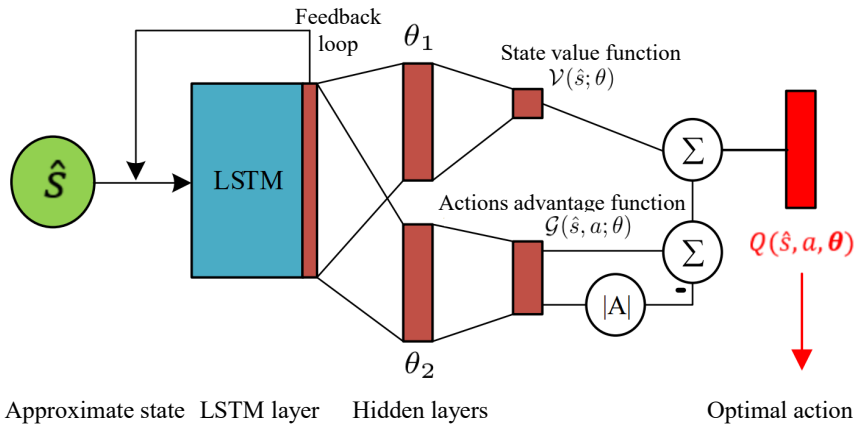


Figure 4.3 : Nullspace estimation and data transmission phases tuning using an LSTM-based deep dueling Q-network.

is designed to avoid the "vanishing gradient" and "exploding gradient", which are the main problems in the training process of the RNN. Therefore, the LSTM is capable of solving sequential processing tasks not solvable by the RNN. On the other hand, the deep dueling network structure, as mentioned in Chapter 2, is developed to improve the convergence rate of the deep Q-learning technique, thanks to its innovative network structure. Specifically, the dueling structure contains two streams that separately estimate the state value and the advantages of actions. As such, the advantage streams of the network can concentrate on learning from only the relevant input (i.e., approximate state), hence improving the convergence rate. The input-output flow of the LSTM-based deep dueling Q-network in Fig. 4.3 is illustrated in Algorithm 2.1 and explained below.

First, the approximate state $\hat{s}$ is fed into the LSTM layer as the input. The LSTM captures the time-varying characteristic of the correlations between transmitted jamming signals (i.e., by observing the average singular values $\bar{\Lambda}_l$) and the change in channels condition (i.e., by monitoring the average post-equalization SINR values $\bar{\delta}_k$). The output from the LSTM is then processed by two separated streams

of fully-connected hidden layers to calculate the state value function $\mathcal{V}(\hat{s};\theta)$ and the advantages of actions $\mathcal{G}(\hat{s},a;\theta)$. The state-action value function $Q(\hat{s},a;\theta)$ is then calculated from $\mathcal{V}(\hat{s};\theta)$ and $\mathcal{G}(\hat{s},a;\theta)$ by [41]

$$Q(\hat{s},a;\theta) = \mathcal{V}(\hat{s};\theta) + \big(\mathcal{G}(\hat{s},a;\theta) - \frac{1}{|\mathcal{A}|}\sum_{a}\mathcal{G}(\hat{s},a;\theta)\big), \tag{4.15}$$

where $|\mathcal{A}|$ denotes the dimension of the action space $\mathcal{A}$ (i.e., $|\mathcal{A}| = L^{\mathrm{e}} \times L^{\mathrm{d}}$).

The training process of the proposed technique has a computational complexity of $\mathcal{O}(W_\theta)$, where W_θ denotes the total number of the Q-network's parameters; W_θ is given by [97]

$$W_\theta = (4N_{\mathrm{i}}N_{\mathrm{c}} + 4N_{\mathrm{c}}^2 + N_{\mathrm{c}}N_{\mathrm{ol}} + 3N_{\mathrm{c}}) + (N_{\mathrm{ol}}N_{\theta_1} + N_{\mathrm{ol}}N_{\theta_2} + N_{\theta_2}|\mathcal{A}| + N_{\theta_1}), \tag{4.16}$$

where N_{i} is the number of input features, which is equal to the size of the 1st-order history observation (i.e., $N_{\mathrm{i}} = K + N_{\mathrm{J}} + 1$), $N_{\mathrm{c}} = H$ is the number of memory cells of the LSTM, N_{ol} is the output size of the LSTM layer, and N_{θ_1} and N_{θ_2} are the neuron number of the upper and lower separated fully-connected hidden layers in Fig. 4.3, respectively. The summation in the first parentheses in Eq. (4.16) shows LSTM's total number of parameters, while the total number of parameters from after the LSTM to the end of the LSTM-based deep dueling Q-network is given in the second parentheses. Note that the training process is performed by the BS, which has more computational power than the UEs. Moreover, when necessary, the training process can be offloaded to a cloud server connected to the BS through a backhaul link.

4.3 Performance Evaluation

4.3.1 Parameter Setting

A 200 kHz bandwidth (i.e., corresponding to a symbol duration of 5 μs) MIMO system is considered, containing a BS and $K = 4$ UEs, each receiving $M_k = 3$ signal streams from the BS. The BS and all the UEs have ULA array structures with 12

Table 4.1 : Parameters for the Deep Dueling Q-learning.

Parameter	**Value**	**Parameter**	**Value**
LSTM's input feature size N_{i}	7	LSTM's number of memory cell N_{c}	6
LSTM's output size N_{ol}	128	Fully connected layers size $N_{\theta_1}, N_{\theta_2}$	(16, 16)
Mini-batch size N_{mb}	32	Memory size M	10,000
Exploration rate ϵ range	[0.10 1.00]	Exploration decay rate	0.99
Target network updating steps C	1,000	Learning rate α	0.01
Discount factor γ	0.90		

and 8 antennas, respectively (i.e., $N_{\mathrm{T}} = 12$ and $N_k = 8, \forall k \in (1, 2, 3, 4)$). The carrier frequency is 447 MHz and the transmitted power is $P_{\mathrm{T}} = 44$ dBm. The received signals at the receivers are sampled with a sampling rate of 400 kHz. The legitimate signals are modulated using the 16-Quadrature Amplitude Modulation (QAM) technique. To evaluate the system's performance more directly, no forward error correction (FEC) coding is used. For small-scale fading, as presented in Section 3.6, a flat fast fading Rician model in [78] is used to simulate the jammers-UEs and BS-UEs channels. The jammers-UE and BS-UE relative velocities are assumed to be around 20 km/h, corresponding to a Doppler frequency of $F_{\mathrm{d}} = 8.28$ Hz. The number of propagation paths for each BS-UE link is $N_k^{\mathrm{p}} = 8, \forall k \in (1, 2, 3, 4)$. For the large-scale path-losses η_k and $\eta_{k,j}$, the COST 231 Hata model [77, Ch. 4] is used, and $h_{\mathrm{T}} = 50$ m, $h_{\mathrm{J}} = h_{\mathrm{UE}} = 2$ m, and $d_{\mathrm{J}} = d_{\mathrm{T}} = 100$ m are assumed, where $h_{\mathrm{T}}, h_{\mathrm{J}}, h_{\mathrm{UE}}, d_{\mathrm{J}}$, and d_{T} denote BS's height, jammers' height, UEs' height, the initial

distance from each jammer to the UEs, and initial BS-UE distance, respectively.

There are $N_{\mathrm{J}} = 2$ jammers with intentionally time-varying correlations between their transmitted signals. Even though the proposed technique can learn and adapt to any jamming strategy, for demonstration, a sawtooth function is used to describe the correlations between transmitted jamming signals. Specifically, the correlation ρ_{12} between the two jammers is assumed to repeatedly and linearly decrease from 1 to 0.8 over 5000 samples as

$$\rho_{12}(p) = \mathcal{J}(p) = 1 - 0.2(p - 5000 \times \lfloor p/5000 \rfloor), \tag{4.17}$$

where $\lfloor . \rfloor$ denotes the floor function, determining the closest smaller integer. The candidate sets for N^{e} and N^{d} are $\mathcal{N}^{\mathrm{e}} = (10, 20, 30, 40)$ and $\mathcal{N}^{\mathrm{d}} = (200, 250, 300, 350)$ samples (i.e., corresponding to $(100, 125, 150, 175)$ symbols per each data transmission phase), respectively. The selections for N^{e} and N^{d} follow the ETSI standard for Terrestrial Trunked Radio (TETRA) [98]. Note that the TETRA standard is used since it allows us to use flexible parameters, suitable for both civil and tactical applications. Note also that, since $\hat{\mathbf{F}}_k^{\mathrm{e}}$ is estimated from $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}}$, it is required that $N^{\mathrm{e}} \geq N_{\mathrm{J}}$, because this is a necessary condition for the nullspace of $\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}}$ to be the same to that of $\mathbf{Z}_k$. The parameters of the LSTM-based deep dueling Q-network and its training parameters are given in Table 4.1, which are similar to those in [33, 66].

To illustrate the advantage of the proposed technique, the following schemes are compared:

- *Upper bound*: The system is assumed to perfectly nullify the jamming signals by using the estimated beam-forming matrix. For illustration purposes, the beam-forming matrix is directly estimated using the jamming signal in the data transmission phase (which is, in reality, unknown to the system).

- *Fixed action*: The system uses a fixed pair of values for N^{e} and N^{d}. The

performance metrics are calculated by averaging the performance of $(L^{\mathrm{e}} \times L^{\mathrm{d}})$ action choices;

- *Heuristic approach*: The system uses the jamming nullification technique in Chapter 3, in which the residual jamming signals are measured, and the estimated beam-forming matrix is updated whenever the residual exceeds a predefined value;

- *Proposed technique*: The values of N^{e} and N^{d} are determined by the optimal policy obtained using the proposed deep dueling Q-learning technique.

4.3.2 Simulation Result

For a fair comparison, the effective spectral efficiency and the outage probability are averaged over $N = 5000$ frames and the K UEs, and denote them by $C_{\mathrm{av}}^{\mathrm{eff}}$ and $p_{\mathrm{av}}^{\mathrm{ot}}$, respectively. We have

$$C_{\mathrm{av}}^{\mathrm{eff}} = \frac{1}{NKM_k} \sum_{n=1}^{N} \sum_{k=1}^{K} \sum_{m=1}^{M_k} C_{k,m}^{\mathrm{eff}}[n],$$
$$p_{\mathrm{av}}^{\mathrm{ot}} = \frac{1}{NKM_k} \sum_{n=1}^{N} \sum_{k=1}^{K} \sum_{i=1}^{M_k} \mathbf{1}_{\hat{\delta}_{k,m}[n] < \delta_{\min}},$$

where

$$\mathbf{1}_{\hat{\delta}_{k,m}[n] < \delta_{\min}} = \begin{cases} 1, & \hat{\delta}_{k,m}[n] < \delta_{\min}, \\ 0, & \text{otherwise}, \end{cases}$$

and $\delta_{\min} = 11.8$ dB is the minimum required SINR.

4.3.2.1 Effective Spectral Efficiency Analysis

Fig. 4.4 shows the average effective spectral efficiency $C_{\mathrm{av}}^{\mathrm{eff}}$ of each BS-UE communication link for different jamming nullification techniques and different jamming power P_{J}. Note that the jamming power is calculated by the variance sum of all the

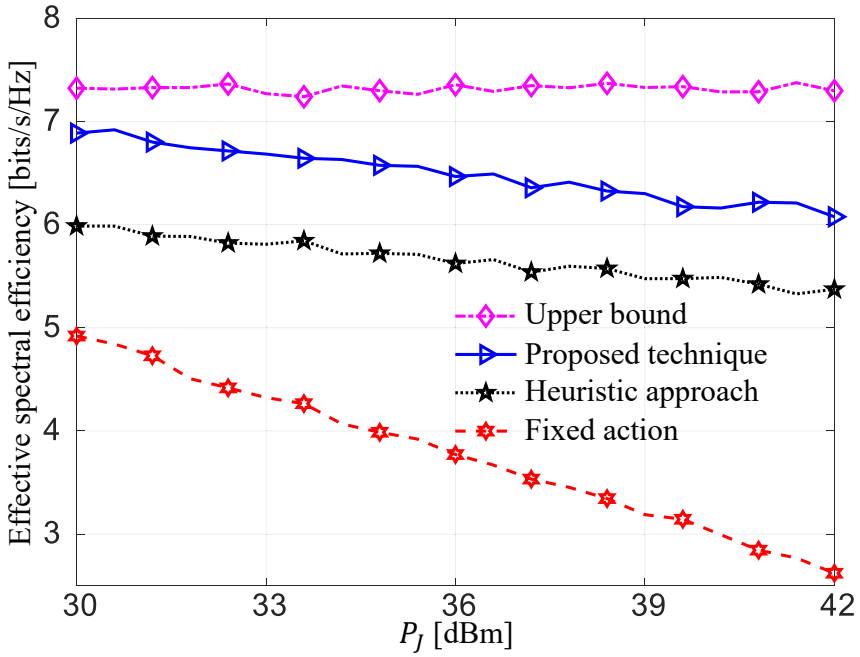


Figure 4.4 : Effective spectral efficiency for different techniques and jamming powers.

jammers (i.e., $P_{\mathrm{J}} = \sum_{j=1}^{N_{\mathrm{J}}} \sigma_{\mathrm{J}_j}^2$). As can be seen, the proposed deep dueling Q-learning technique achieves the highest average effective spectral efficiency for all values of the jamming power, thanks to its ability to efficiently adjust the N^{e} and N^{d} values according to the change in the correlations and channel conditions. On the other hand, the other two techniques have several limitations. The technique in Chapter 3 spends an excessive amount of time monitoring the residual jamming signals and estimating the beam-forming matrix, thus reducing the data transmission time. Meanwhile, the *fixed action* technique cannot adapt to the change in the channel conditions, and more importantly, the change in the correlations between transmitted jamming signals, resulting in an ineffective estimated beam-forming matrix $\hat{\mathbf{F}}_k^{\mathrm{e}}$. These limitations of Chapter 3 and *fixed action* result in lower spectral efficiencies of the communication system. Note that there is a gap between the spectral efficiency of the proposed technique and that of the *upper bound*. This is because, in the *upper bound* case, the jamming signal is completely canceled out without needing to adjust

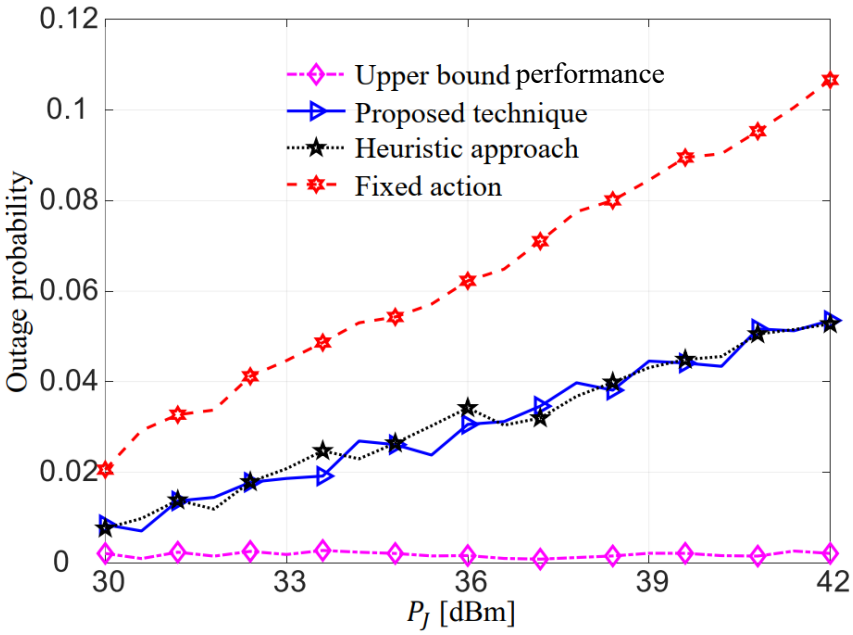


Figure 4.5 : Outage probability for different techniques and jamming powers.

the values of N^{e} and N^{d}. Meanwhile, for the proposed technique, N^{e} may have to be increased to improve the estimation accuracy of $\hat{\mathbf{F}}_k^{\mathrm{e}}$, hence reducing the effective spectral efficiency.

4.3.2.2 Outage Probability Analysis

Fig. 4.5 illustrates the outage probability of the systems using the three mentioned techniques and different values of the jamming power. As can be seen, the proposed deep dueling Q-learning technique and the techniques in Chapter 3 have very similar outage probabilities, and are much lower than that of the *fixed action* technique. This is because the *fixed action* technique cannot adapt to the change of the correlations between transmitted jamming signals and channels condition, resulting in many outage frames because of excessive jamming residuals. On the other hand, both *heuristic approach* and the proposed technique effectively nullify the jamming signals. However, as mentioned above, the *heuristic approach* in Chapter 3

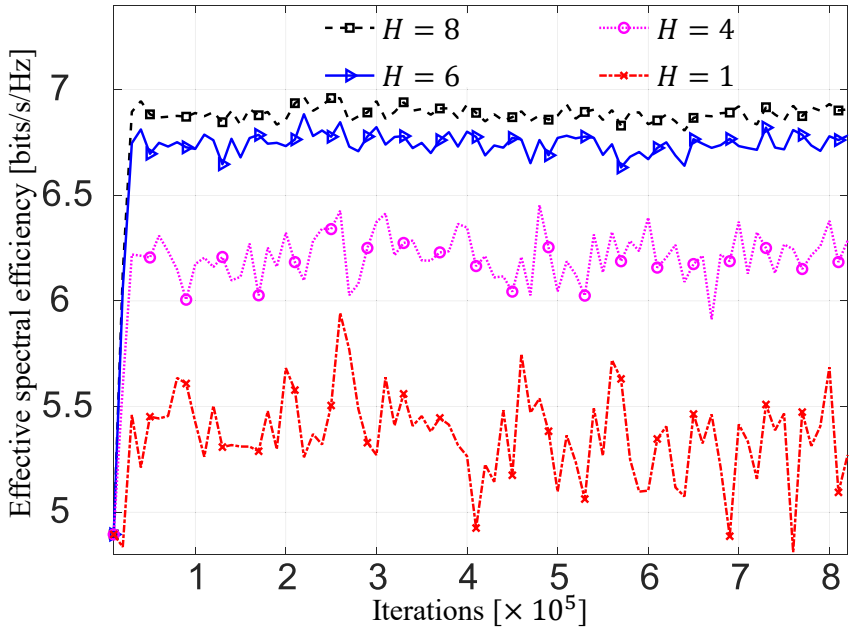


Figure 4.6 : Spectral efficiency convergence rate for different history lengths.

spends an excessive amount of time monitoring the residual jamming signals and estimating the beam-forming matrix, resulting in a much lower spectral efficiency. Note that there is a difference in the outage probability between the proposed technique and the *upper bound*, because the jamming signal, in fact, cannot be entirely canceled out as in the *upper bound* case.

4.3.2.3 Impact of History Length H

Fig. 4.6 and Fig. 4.7 illustrate the impact of the history length H on the convergence of the proposed deep dueling Q-learning technique. The jamming power used to generate these figures is 30 dBm. As can be seen, the deep dueling Q-learning technique converges after around 2×10^4 iterations, corresponding to a minimum T_{Q} value of $T_{\mathrm{Q}}^{\min} = 2 \times 10^4 \times (10 + 200) \times 1/(4 \times 10^5) = 10.5$ seconds to $T_{\mathrm{Q}}^{\min} = 2 \times 10^4 \times (40 + 350) \times 1/(4 \times 10^5) = 19.5$ seconds. Note that, however, T_{Q} also depends on the system's hardware computational capability and the system's bandwidth.

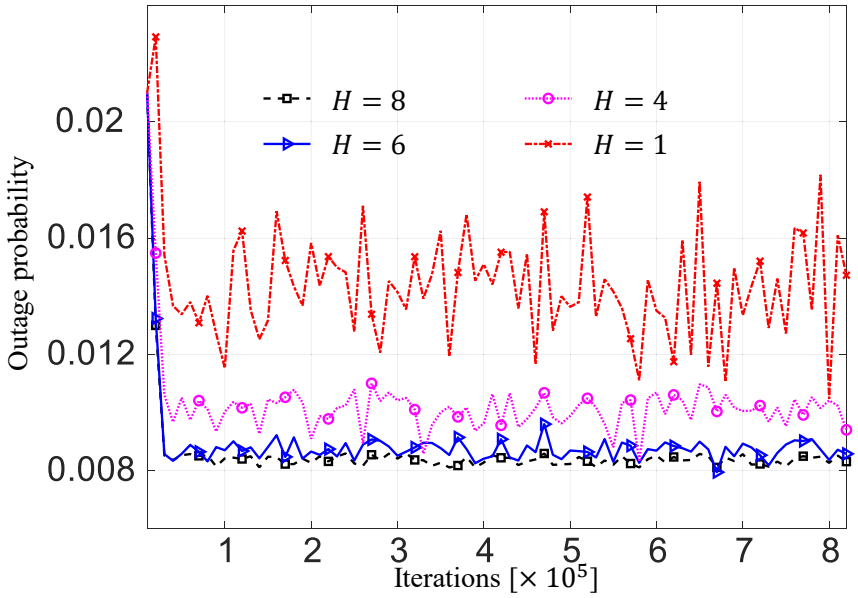


Figure 4.7 : Outage probability convergence rate for different history lengths.

Specifically, the interval of each iteration is the maximum value between the frame interval and the amount of time it takes for the system to perform each training iteration. Additionally, for a system with a higher bandwidth (and correspondingly a higher sampling rate), it would require a much shorter period of time to obtain sufficient training data set. For example, with a sampling rate of 30 MHz (e.g., LTE system), the value range for T_Q can be significantly reduced (i.e., given enough system hardware's computational capability). Note also that a longer history length H results in a higher effective spectral efficiency and a lower outage probability. This is because a longer history H allows the system to better observe the impact of action selection on the system reward. Therefore, the system is able to select better actions that result in better system performance. However, increasing the value of H also increases the computational complexity of the deep dueling Q-learning technique, as demonstrated by Eq. (4.16). In the simulations setting, $H = 6$ is selected, which has similar performance and a lower computational complexity compared to $H = 8$.

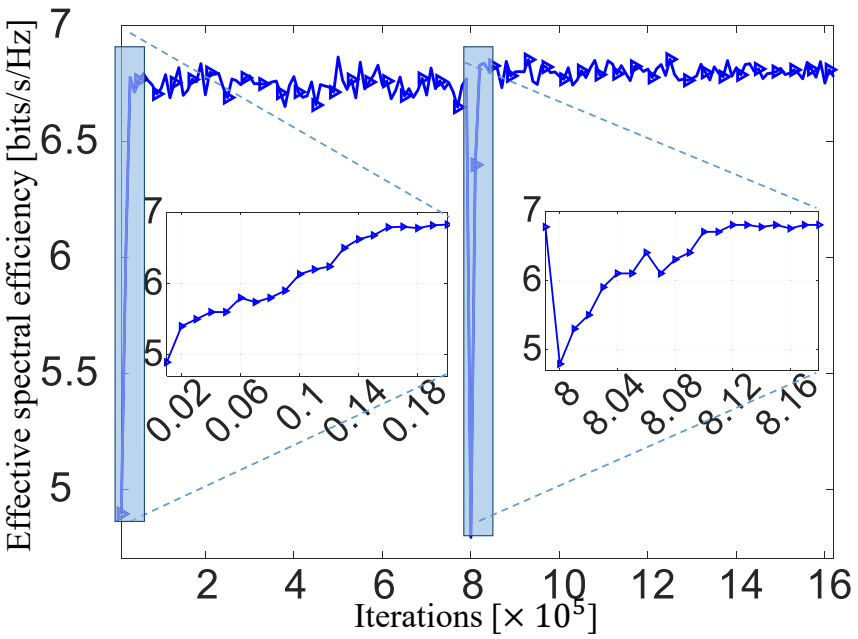


Figure 4.8 : Adaptability to the change in the jamming strategy.

4.3.2.4 Adaptability to the Change in Jamming Strategy

Fig. 4.8 and Fig. 4.9 illustrate the capability of the deep dueling Q-learning technique to adapt the optimal anti-jamming policy when the jamming strategy is changed. Specifically, from the (8×10^5)th iterations, instead of linearly decreasing as in Eq. (4.17), the correlation is linearly increased from 0.8 to 1 as

$$\rho_{12}(p) = \mathcal{J}'(p) = 0.8 + 0.2[(p - 8 \times 10^5) - 5000\lfloor (p - 8 \times 10^5)/5000 \rfloor].$$

As can be seen, the change in the jamming strategy causes a sudden drop in the system performance. When such a drop is observed, the exploration rate is reset to 1 to increase the system's convergence rate. As shown, the deep dueling Q-learning technique can adapt to the change in the jamming strategy by quickly re-obtain the convergence for both effective spectral efficiency and outage probability. In particular, the re-establishment of the convergence status (i.e., from the (8×10^5)th iteration where the jamming strategy changes) is even slightly faster than the first

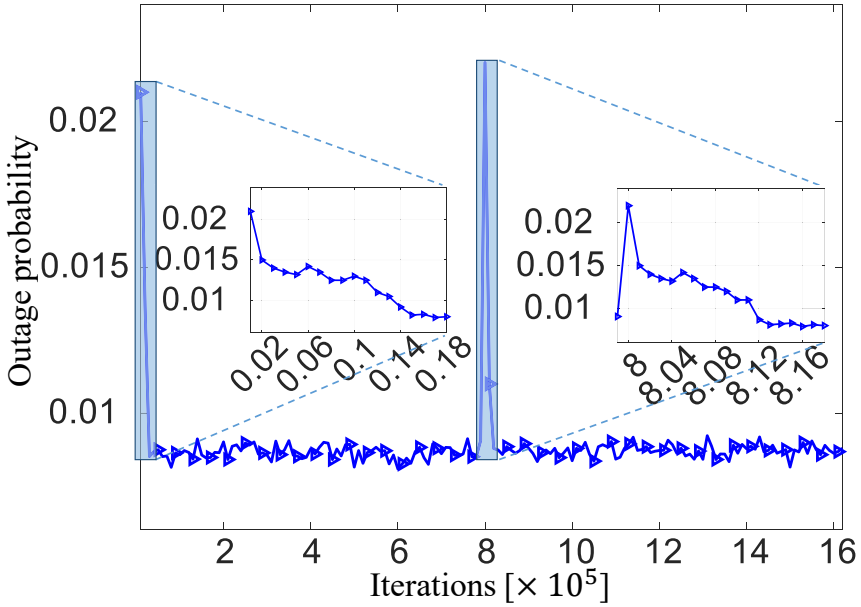


Figure 4.9 : Adaptability to the change in the jamming strategy.

convergence (i.e., from the first iteration). This is because the Q-network at the (8×10^5)th iteration is "initialized" by the network parameters obtained from the previous training process. Therefore, this "initialization" performs better than the random initialization at the first iteration when the training process begins.

4.3.2.5 Convergence Rate Analysis:

Fig. 4.10 illustrates the convergence rates of different Q-learning techniques. As can be seen, the conventional Q-learning technique has not converged after 8×10^5 iteration, which is much larger than 2×10^4 iterations required for the convergence of the proposed deep dueling Q-learning technique. Therefore, the deep dueling Q-learning technique has significantly improved the convergence rate of the system.

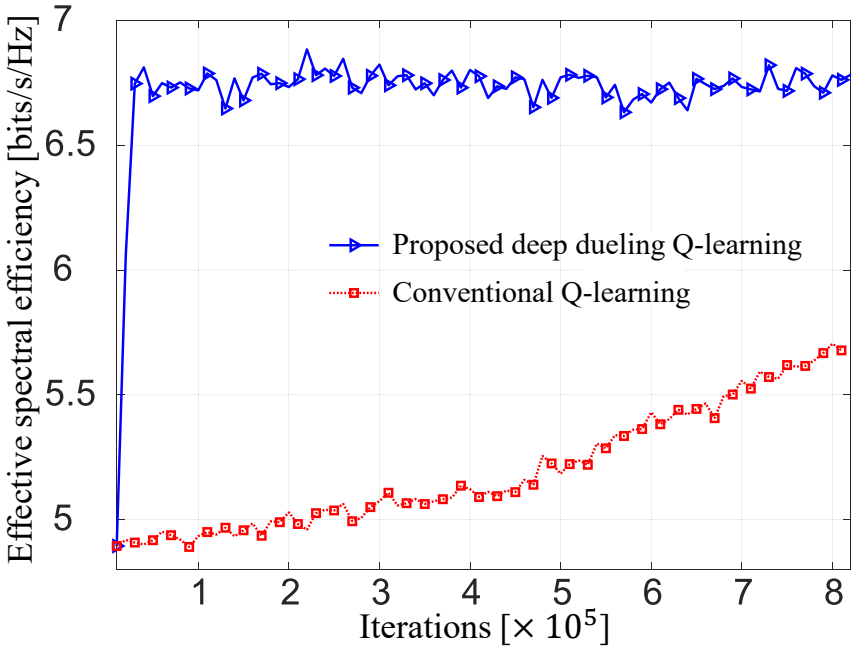


Figure 4.10 : Convergence rate of different Q-learning techniques.

4.3.2.6 Impact of the Number of Jamming Antennas:

Fig. 4.11 illustrates the impact of the number of jamming antennas on the achieved spectral efficiency of the system trained by the proposed deep dueling Q-learning technique. As can be seen, the spectral efficiency decreases dramatically as the number of jamming antennas increases from 2 to 4. This is because when $N_{\mathrm{J}} = 4$, the $k$the UE receiver has only 4 remaining degree-of-freedoms to demodulate $M_k = 3$ spatial streams, resulting in a much lower SINR, and hence a much lower spectral efficiency.

4.4 Conclusions

The impact of time-varying correlations between transmitted jamming signals on jamming nullification has been examined. It has been demonstrated that using an incorrectly estimated beam-forming matrix can lose the receiver's degree-of-freedoms

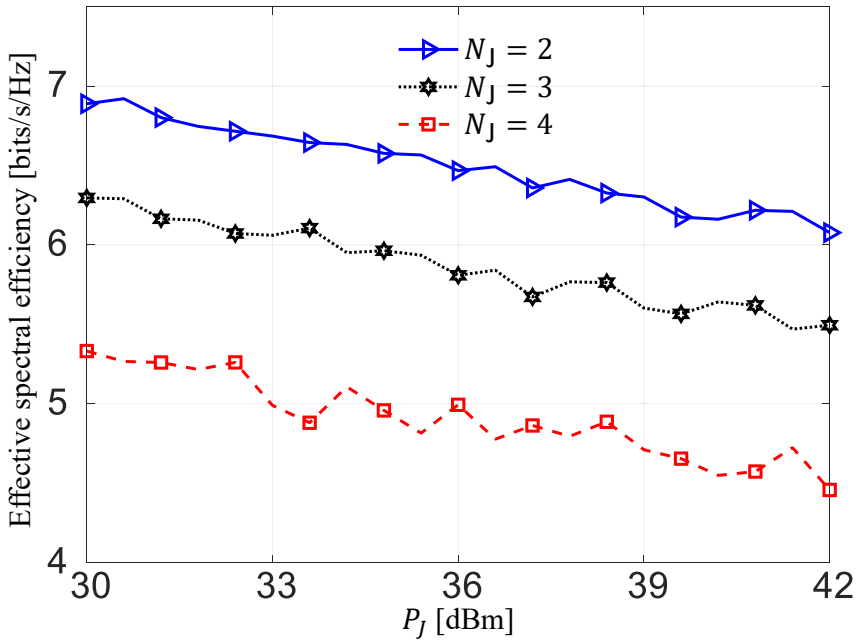


Figure 4.11 : Spectral efficiency for different numbers of jamming antennas.

without achieving effective jamming suppression. The deep dueling Q-learning technique was proposed to effectively estimate the beam-forming matrix and nullify the jamming signals designed to aggravate the jamming impact. The dueling network architecture allows the proposed technique to quickly obtain the optimal solution against the jammers, therefore effectively dealing with unknown and time-varying jamming strategies. The simulation results demonstrate that the proposed deep dueling Q-learning technique achieves a higher effective spectral efficiency and a lower outage probability than the existing techniques.

Chapter 5

Frequency Hopping Joint Radar and Communications with Hybrid Sub-pulse Frequency and Duration Modulation

In the previous sections, jamming mitigation techniques for communication systems were studied. This chapter extends jamming suppression techniques to joint radar and communication (JRC) systems. Specifically, this chapter designs a frequency hopping (FH) signal, which possesses robustness against jamming, for a JRC system. In particular, novel techniques to embed and demodulate data to increase the data rate and reduce the demodulation error are also investigated. For data embedding, both sub-pulse frequency and duration are used, therefore increasing the data transmission rate compared to only using the sub-pulse frequency. For data demodulation, a novel scheme is proposed based on the signal's time-frequency image (TFI) and a 'You Only Look Once' (YOLO)-based detection system. This demodulation scheme, instead of requiring a channel estimation, only requires the estimation of the channel delay spread, thus less prone to estimation error. Moreover, the proposed demodulation technique is more robust to the Doppler shift and the timing offset between the transceiver and the communication receiver compared to the existing ones. Additionally, the proposed data embedding and demodulation schemes are spatially flexible and not limited to the sidelobe of the transmit beampattern, since the data is not embedded by utilizing the phase or amplitude of the beampattern sidelobe.

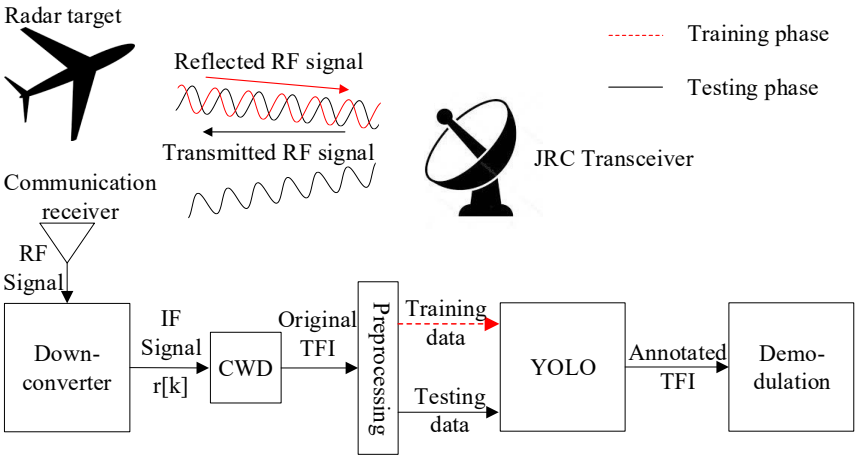


Figure 5.1 : FH joint radar and communication system.

5.1 System and Signals

5.1.1 System Overview

A ground-to-air JRC system is considered, where the JRC transceiver is located on the ground, and the communication receiver is mounted on an aircraft or a high-altitude platform (HAP), as illustrated in Fig. 5.1. For sensing, the JRC transceiver transmits the radio frequency (RF) signal and processes the reflected RF signal to determine the target's position and velocity. For data transmission, the communication receiver obtains the RF signal from the JRC transceiver and demodulates the embedded data. At the transceiver or communication receiver, the received RF signal is down-converted into the intermediate frequency (IF) signal and sampled at sampling frequency f_s to generate a discrete-time signal as

$$r[k] = h[k] * s[k] + w[k]. \tag{5.1}$$

Here, k is the sample index which increases every $T_s = 1/f_s$, $s[k]$ and $r[k]$ are the transmitted and received signal samples, respectively, $h[k]$ denotes a Rician channel model [78], and $w[k]$ is the complex additive white Gaussian noise (AWGN). The JRC system uses a pulse wave (PW) waveform [99] with a duty cycle $D =$

$\triangle t_{\rm w}/\triangle t_{\rm r}, D \ll 1$, where $\triangle t_{\rm w}$ and $\triangle t_{\rm r}$ are the pulse width and pulse repetition intervals, respectively. The received signal at the communication receiver is then used to generate the Choi–Williams distribution time-frequency image (CWD–TFI). Next, the CWD–TFI is preprocessed in the Preprocessing block to generate the training and testing data for the YOLO detection system. The training process of the YOLO detection system is demonstrated by the dashed line in Fig. 5.1. The trained YOLO detection system and the testing CWD–TFIs are used to generate the annotated CWD–TFIs, which are used to demodulate the embedded data.

5.1.2 Signal Model

The JRC system uses FH signals, with each signal pulse being divided into $N_{\rm f}$ sub-pulses (i.e., hops) with different frequencies and durations. Let $\mathbf{F} = (f_1, f_2, ..., f_{N_{\rm f}})$ denote the FH sequence and $\mathbf{T} = (\triangle t_1, \triangle t_2, ..., \triangle t_{N_{\rm f}})$ denote the sub-pulse duration sequence. The JRC signal in each pulse is represented by

$$s[k] = \sum_{n=1}^{N_{\rm f}} A e^{j2\pi f_n k T_{\rm s}} {\rm rect}\Big[\frac{kT_{\rm s} - \sum_{i=0}^{n-1} \triangle t_i}{\triangle t_n}\Big], \tag{5.2}$$

where

$$ {\rm rect}[k] \triangleq \begin{cases} 1, & 0 \le k \le 1, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\triangle t_0 \triangleq 0.$$

5.2 Data Embedding Schemes

In this work, the Costas arrays are used to generate the FH sequence of each signal pulse in one of the data embedding schemes. Using the Costas array, an FH signal possesses a narrow peak at the origin of its ambiguity function (AF) and

low sidelobes elsewhere, which is desirable for radar measurement (i.e., range and speed) accuracy. Specifically, the pulse compression ratio (PCR) of a Costas signal is N_{f}^2. Details of the Costas arrays, their construction, and their characteristics can be found in [100, Ch. 5].

Two data embedding techniques are proposed, namely *Random* and *Costas-based* schemes. Unlike existing studies that only utilize sub-pulse frequencies to convey data, both sub-pulse frequencies and durations are used in the proposed schemes. First, the codebooks $\mathcal{S}$ and $\mathcal{S}_{\mathrm{C}}$ of the *Random* and *Costas-based* schemes, respectively, are designed. Then, for each embedding scheme, data bits are embedded by mapping each element of the corresponding codebook to a symbol. Let $\mathcal{F} \triangleq [(1, 2, ..., N_{\mathrm{f}}) \times f_{\mathrm{f}}]$ be the selection set of the sub-pulse frequency, where f_{f} is the fundamental frequency. Let $\mathcal{D} \triangleq (\triangle_1, \triangle_2, ..., \triangle_{N_{\mathrm{T}}})$ be the selection set of the sub-pulse duration, where N_{T} is the number of sub-pulse duration selections. The codebooks $\mathcal{S}$ and $\mathcal{S}_{\mathrm{C}}$ are designed as follows.

- *Random* scheme: $\mathcal{S}$ is designed by selecting each sub-pulse frequency from the frequency set $\mathcal{F}$, and selecting each sub-pulse duration from the duration set $\mathcal{D}$ as

$$f_i \in \mathcal{F}, \triangle t_i \in \mathcal{D}, \forall i \in (1, 2, ..., N_{\mathrm{f}}), \tag{5.3}$$

$$f_i \neq f_{i+1}, \forall i \in (1, 2, ..., N_{\mathrm{f}} - 1). \tag{5.4}$$

  Here, condition (5.4) means every two consecutive sub-pulse frequencies are not equal, which is necessary for the data demodulation technique that will be presented in Section 5.4.

- *Costas-based* scheme: Unlike the *Random* scheme, for designing $\mathcal{S}_{\mathrm{C}}$, each sub-pulse frequency is not individually selected. Instead, the whole FH sequence of each signal pulse is selected from the Costas-based FH sequence defined as follows.

Let $\mathcal{A}_{\mathrm{C}} \triangleq (\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_{N_{\mathrm{C}}})$ be the set of Costas array of length N_{f}, where $\mathbf{A}_i, \forall i \in (1, 2, ..., N_{\mathrm{C}})$ is an $\mathbb{N}^{N_{\mathrm{f}} \times 1}$ Costas array, and N_{C} is the number of Costas sequence with length of N_{f}. Then, let $\mathcal{F}_{\mathrm{C}} \triangleq [(\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_{N_{\mathrm{C}}}) \times f_{\mathrm{f}}]$ be the selection set of the Costas-based FH sequence. Then, the FH sequence of each signal pulse is selected from $\mathcal{F}_{\mathrm{C}}$. On the other hand, similar to the *Random* scheme, each sub-pulse duration is selected from the sub-pulse duration set $\mathcal{D}$ as

$$\mathbf{F} \in \mathcal{F}_{\mathrm{C}}, \triangle t_i \in \mathcal{D}, \forall i \in (1, 2, ..., N_{\mathrm{f}}). \tag{5.5}$$

Note that for the *Costas-based* scheme, because the FH sequence is generated using the Costas array, condition (5.4) is satisfied. The dimensions of $\mathcal{S}$ and $\mathcal{S}_{\mathrm{C}}$ are given by

$$|\mathcal{S}| = \underbrace{N_{\mathrm{T}}^{N_{\mathrm{f}}}}_{\text{sub-pulse duration selections}} \times \underbrace{N_{\mathrm{f}}(N_{\mathrm{f}}-1)^{(N_{\mathrm{f}}-1)}}_{\text{sub-pulse frequency selections}}, \tag{5.6}$$

$$|\mathcal{S}_{\mathrm{C}}| = \underbrace{N_{\mathrm{T}}^{N_{\mathrm{f}}}}_{\text{sub-pulse duration selections}} \times \underbrace{N_{\mathrm{C}}}_{\text{sub-pulse frequency selections}}. \tag{5.7}$$

As shown in (5.6), the number of selections of sub-pulse frequencies is $N_{\mathrm{f}}(N_{\mathrm{f}} - 1)^{(N_{\mathrm{f}}-1)}$ instead of $N_{\mathrm{f}}^{N_{\mathrm{f}}}$ due to the requirement in (5.4).

It follows from (5.6) and (5.7) that the maximum number of bits C and C_{C} that can be embedded in each pulse of the two schemes are given by

$$\begin{aligned} C &= \lfloor \log_2(|\mathcal{S}|) \rfloor = \lfloor \log_2(N_{\mathrm{T}}^{N_{\mathrm{f}}} N_{\mathrm{f}}(N_{\mathrm{f}}-1)^{(N_{\mathrm{f}}-1)}) \rfloor, \\ C_{\mathrm{C}} &= \lfloor \log_2(|\mathcal{S}_{\mathrm{C}}|) \rfloor = \lfloor \log_2(N_{\mathrm{T}}^{N_{\mathrm{f}}} N_{\mathrm{C}}) \rfloor, \end{aligned} \tag{5.8}$$

where $\lfloor . \rfloor$ denotes the floor function, determining the closest smaller integer. As can be seen, the *Random* scheme achieves a higher transmission rate than the *Costas-based* scheme.

5.3 Sensing at the JRC Transceiver

This section describes the matched filtering technique used at the JRC transceiver to detect the reflected radar signal. It is also demonstrated that by using the sub-pulse frequencies and durations for data embedding, as presented in the previous section, the detection probability of the sensing function is not affected.

Let P_{D} and P_{FA} denote the detection and false alarm probabilities of the JRC transceiver, respectively. Let E_{p} be the signal energy per each pulse of the JRC signal and σ_{w}^2 be the noise variance. The energy-to-noise ratio (ENR) of the JRC signal is given by

$$\mathrm{ENR} = 10\log_{10}(\frac{E_{\mathrm{p}}}{\sigma_{\mathrm{w}}^2}). \tag{5.9}$$

To detect the transmitted signal $s[k]$ buried in the reflected RF signal $r[k]$, the JRC transceiver performs the matched filtering. Let R_{s} denote the output of the matched filter. Based on the Neyman-Pearson criterion, the transceiver decides a detection if R_{s} exceeds a threshold γ

$$R_{\mathrm{s}} = \sum_{k=0}^{N_{\mathrm{p}}-1} r[k]s[k] > \gamma, \tag{5.10}$$

where N_{p} is the number of signal sample per pulse,

$$\gamma = \sqrt{\sigma_{\mathrm{w}}^2 E_{\mathrm{p}}}Q^{-1}(P_{\mathrm{FA}}), \tag{5.11}$$

$$Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}t^2)\, dt \tag{5.12}$$

is the complement of the cumulative distribution function (CDF) of the standard Gaussian distribution, and $Q^{-1}(x)$ denotes the inverse function of $Q(x)$. The characteristic of P_{D} is given in Theorem 1 below.

Theorem 5.1: The value of P_{D} is independent from the JRC signal waveform and only depends on P_{FA} and ENR.

Proof: From [101, Ch. 4], P_{D} can be expressed by P_{FA} and ENR as

$$P_{\mathrm{D}} = Q\Big[Q^{-1}(P_{\mathrm{FA}}) - \sqrt{10^{\mathrm{ENR}/10}}\,\Big]. \tag{5.13}$$

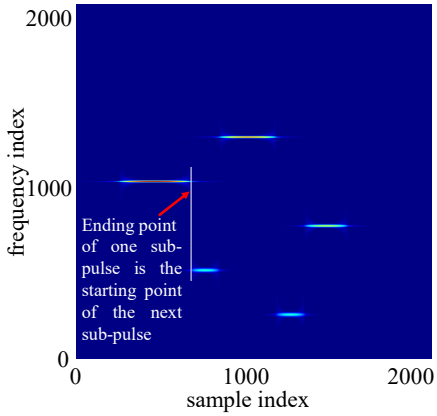
Accordingly, P_{D} only depends on P_{FA} and ENR, and is independent of the JRC signal waveform. ■

Therefore, varying the sub-pulse frequencies and durations does not affect the detection performance of the sensing function, as long as the signal energy per each pulse is fixed. Note that, however, varying the sub-pulse durations may cause some increase in the peak-to-average power ratio (PAPR), which reduces the efficiency of the power amplifier at the front end of the JRC transceiver and communication receiver. Therefore, the sub-pulse selection set $\mathcal{D}$ should be carefully designed to avoid an excessive PAPR value.

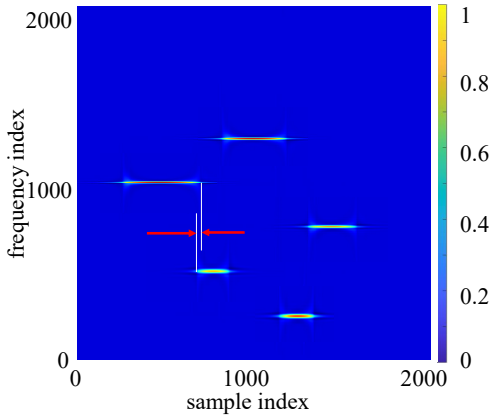
5.4 Demodulation at the Communication Receiver

This section presents the detailed procedure of the proposed data demodulation schemes. First, the CWD is used to generate the received signal's CWD–TFI, which illustrates the sub-pulse frequency and duration of the signal by the signal objects. Then, the 'You Only Look Once' (YOLO) detection system is used to recognize the signal objects on the CWD–TFI. By using the coordinates of the bounding box of the signal objects, the sub-pulse frequencies and durations can be determined. Finally, the embedded symbols are demodulated by comparing the sub-pulse frequencies and durations to the codebook.

There have been multiple linear time-frequency analysis (TFA) techniques that can be used to represent the signal's time-frequency characteristic, such as the short-time Fourier transform (STFT) and the wavelet transform (WT). However, these linear TFA techniques are susceptible to noise, and hence not suitable for processing the (typically) low-SNR radar signals. On the other hand, the bilinear TFA tech-



(a) AWGN channel



(b) Rician fading channel with $\tau_{\mathrm{d}} = 40/f_{\mathrm{s}}$

Figure 5.2 : CWD–TFIs of a Costas signal with $\mathbf{F} = [(4, 2, 5, 1, 3) \times f_{\mathrm{f}}]$, $\mathbf{T} = (5, 2, 4, 2, 3) \times 80/(f_{\mathrm{s}})$, $f_{\mathrm{f}} = f_{\mathrm{s}}/16$, and a starting point of 322.

niques are robust against noise, thanks to the autocorrelation operation in their formulas. Among the bilinear TFA techniques, the most basic form, which also provides the best time-frequency resolution, is the Wigner-Ville distribution (WVD) [102]. However, the WVD has the drawback of the cross-term effect, i.e., the cross-term can obscure essential features of the signal. Therefore, the CWD is developed to address the main limitation of the WVD by using an exponential kernel function in the general time-frequency distribution class introduced by Cohen [103]. In this work, the CWD is used to generate the CWD–TFI, which represents the time-frequency

characteristic of the signal. Specifically, the CWD is mathematically given by [104]

$$\begin{aligned}\mathrm{CW}[k,n] = 2 \sum_{\tau=-\infty}^{\infty} & W_N(\tau) e^{-j2\pi n\tau/N} \\ & \left[\sum_{\mu=-\infty}^{\infty} W_M(\mu) \frac{1}{\sqrt{4\pi\tau^2/\sigma}} e^{-\frac{\mu^2}{4\tau^2/\sigma}} \right. \\ & \left. \times\, y[k+\mu+\tau] y^*[k+\mu-\tau] \right], \end{aligned} \tag{5.14}$$

where k and n are the time index and frequency index, respectively, $W_N(\tau)$ is a symmetrical window function with non-zero values for the range of $-N/2 \leq \tau \leq N/2$, while $W_M(\mu)$ is a rectangular window function with a value of 1 for the range of $-M/2 \leq \mu \leq M/2$, and σ is the scaling factor. In this chapter, similar to [105], rectangular windows are used for both $W_N(\tau)$ and $W_M(\mu)$. Furthermore, $\sigma = 1$ is used to achieve a balance between frequency resolution and the cross-terms reduction in the CWD.

5.4.1 CWD–TFI Preprocessing

To generate the CWD–TFI, similar to [99,105,106], $N_\mathrm{s} = 2048$ consecutive signal samples are captured, and these N_s samples are assumed to contain a complete signal pulse, and the signal pulse starts from an arbitrary point within the N_s samples. Fig. 5.2a and Fig. 5.2b illustrate the CWD–TFIs of a Costas signal travels through an AWGN and a Rician fading channel, respectively. The parameters of the signal includes a hopping sequence of $\mathbf{F} = [(4, 2, 5, 1, 3) \times f_\mathrm{f}]$, a sub-pulse duration sequence of $\mathbf{T} = (5, 2, 4, 2, 3) \times 80/(f_\mathrm{s})$, a fundamental frequency of $f_\mathrm{f} = f_\mathrm{s}/16$, and a starting point of 322. As can be seen in Fig. 5.2a, for the AWGN channel, the ending point of one sub-pulse is the starting point of the next sub-pulse. On the other hand, in Fig. 5.2b, due to the non-line-of-sight (nLoS) paths, each sub-pulse signal object is lengthened by τ_d, the delay spread of the channel h. The procedure to estimate and compensate for this length offset is described in Subsection 5.4.2.

Fig. 5.3 shows the CWD–TFI processing procedure used to generate the input

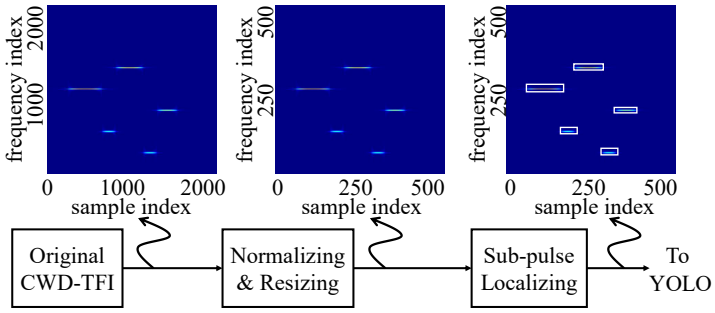


Figure 5.3 : CWD–TFI preprocessing procedure for the Costas signal in Fig. 5.2a.

to the YOLO detection system. First, using the nearest neighbor interpolation, the CWD–TFI of $N_s \times N_s$ pixels is resized to $L \times L$ pixels, which is the input size of the YOLO detection system, where $L = 500$ pixels is the width of the resized CWD–TFIs. Then, the signal sub-pulse locations in the CWD–TFI are specified by the bounding boxes (i.e., white-color rectangles in Fig. 5.3), as a requirement for the training of the YOLO. Note that the localization of the bounding boxes is not needed in the CWD–TFI processing procedure in the testing phase, because the bounding boxes are parts of the output of the trained YOLO detection system.

5.4.2 Data Demodulation Technique

Fig. 5.4 shows the YOLO's output for the CWD–TFI of a Costas signal with an FH sequence of $\mathbf{F} = [(4, 2, 5, 1, 3) \times f_f]$, a sub-pulse duration sequence of $\mathbf{T} = (4, 3, 5, 4, 2) \times 80/(f_s)$, a fundamental frequency of $f_f = f_s/16$, and a starting point of 259. As can be seen, the output contains not only the signal type (i.e., "Costas") but also the locations of the signal objects (i.e., the white-color rectangular bounding boxes) and the probabilities of the detection. The bounding boxes determine the position, horizontal, and vertical width of the signal objects in the CWD–TFI. The

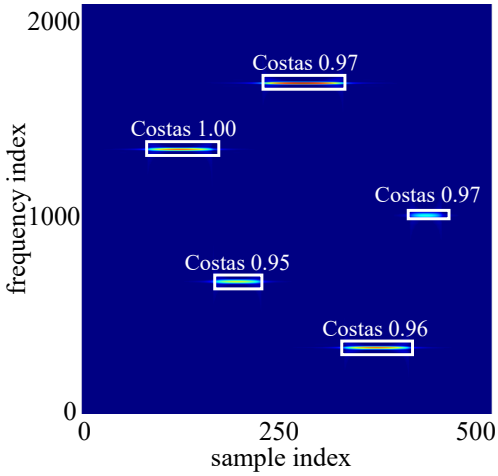


Figure 5.4 : YOLO's output for a Costas signal with $\mathbf{F} = [(4, 2, 5, 1, 3) \times f_{\mathrm{f}}]$, $\mathbf{T} = (4, 3, 5, 4, 2) \times 80/(f_{\mathrm{s}})$, $f_{\mathrm{f}} = f_{\mathrm{s}}/16$, and a starting point of 259.

ith left-most bounding box, corresponds to the ith sub-pulse, is localized by a tuple of $(x_i^{\min}, x_i^{\max}, y_i^{\min}, y_i^{\max})$. Recall that for the CWD–TFI, L pixels in the vertical axis correspond to a frequency range from 0 to $f_{\mathrm{s}}/2$, while L pixels in the horizontal axis correspond to a time duration of $N_{\mathrm{s}} \times (1/fs)$, the frequency and duration of the ith sub-pulse can be calculated by

$$\begin{aligned} f_i &= \frac{f_{\mathrm{s}}(y_i^{\min} + y_i^{\max})}{4L} \\ \triangle t_i &= \frac{N_{\mathrm{s}}(x_i^{\max} - x_i^{\min})}{L f_{\mathrm{s}}} - \tau_{\mathrm{d}}, \end{aligned} \tag{5.15}$$

where τ_{d} is the channel's delay spread. To estimate τ_{d}, a training process is performed, where a signal pulse with known FH sequence $\mathbf{F}$ and sub-pulse duration $\mathbf{T}$ is transmitted by the JRC transceiver and captured by the communication receiver. Then, τ_{d} can be determined by

$$\tau_{\mathrm{d}} = \frac{1}{N_{\mathrm{f}}} \sum_{1}^{N_{\mathrm{f}}} (\triangle t_i - \hat{\triangle} t_i), \tag{5.16}$$

where $\hat{\triangle} t_i$ is the ith sub-pulse duration demonstrated on the CWD–TFI of the training signal. By comparing the combination of f_i and $\triangle t_i$ values with the selections in

Table 5.1 : Parameters values.

Parameter	Description	Value
N_{f}	Number of sub-pulses per pulse	5
N_{C}	Number of Costas array	40
$\mathcal{D}$	Sub-pulse duration selection set	$(1, 1.5, 2, 2.5, 3) \times 80/(f_{\mathrm{s}})$
f_{f}	Fundamental frequency	$f_{\mathrm{s}}/16$
L	Resized CWD–TFI width (pixels)	500
τ_{d}	Delay spread of the Rician channel	$40/f_{\mathrm{s}}$

the codebook $\mathcal{S}$ or $\mathcal{S}_{\mathrm{C}}$, the embedded symbol in each signal pulse can be determined.

It is worth noting that because the proposed data demodulation technique is based on the CWD–TFI, its performance is not affected by the Doppler shift or timing offset between the transceiver and the communication receiver, as long as the signal pulse is fully captured by the $N_{\mathrm{s}} = 2048$ consecutive samples. This is because the Doppler shift or timing offset merely creates vertical or horizontal shifts of the whole set of lines, respectively. On the other hand, the shapes of the lines in the CWD–TFI remain unchanged.

5.5 Performance Evaluation

Table. 5.1 shows the parameter values used to produce the simulation results. All available Costas sequences (i.e., $N_{\mathrm{C}} = 40$ for $N_{\mathrm{f}} = 5$) are used for data modulation. The list of available Costas arrays and their construction can be found in [100, Ch. 5]. To train the YOLO detection system, for each data embedding scheme, $2,700$ signals are generated with SNR ranging from -6 dB to 10 dB and a step size of 2 dB. The $2,700$ signals are divided into a training set of $2,160$ signals (80% of the total) and a validation set of 540 signals (20% of the total). For the testing data, we generate

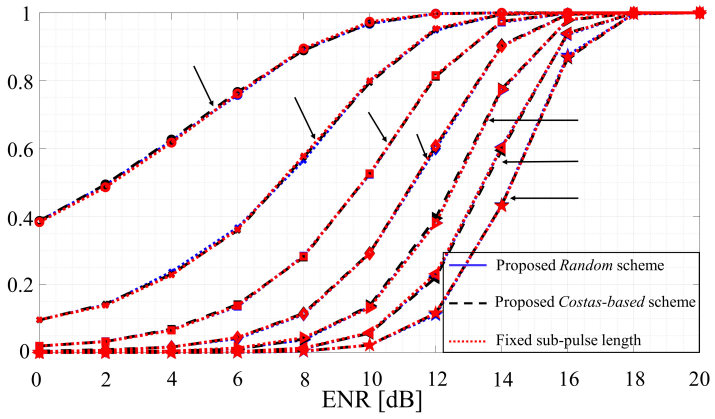


Figure 5.5 : Detection probability of the JRC sensing function for different signal waveforms, P_{FA}, and ENR values.

$3,300$ signals for each embedding scheme, with SNR ranging from -10 dB to 10 dB and a step size of 2 dB.

Fig. 5.5 shows the detection probability of the JRC sensing function for different signal waveforms (i.e., *Random* scheme, *Costas-based* scheme, and fixed sub-pulse length), P_{FA}, and ENR values. The results are obtained by averaging $10,000$ independent trials. As can be seen, for each P_{FA} value, the three curves almost overlap, meaning the detection probability only depends on P_{FA} and ENR but not the signal waveform. Therefore, varying the sub-pulse durations does not affect the detection capability of the sensing function of the FH RJC system, as long as the signal energy per each pulse is fixed. Fig. 5.6 demonstrates the maximum number of bits that can be transmitted over a signal pulse of the two proposed embedding schemes and those in [51, 53]. As can be seen, the *Random* scheme can embed a significantly higher amount of data than the *Costas-based* scheme, while the *Costas-based* scheme can convey a slightly higher number of data bits than the scheme in [51, 53]. That is because the schemes in this chapter use both sub-pulse durations and frequencies to embed signals. On the other hand, the technique in [51] focuses on using antenna

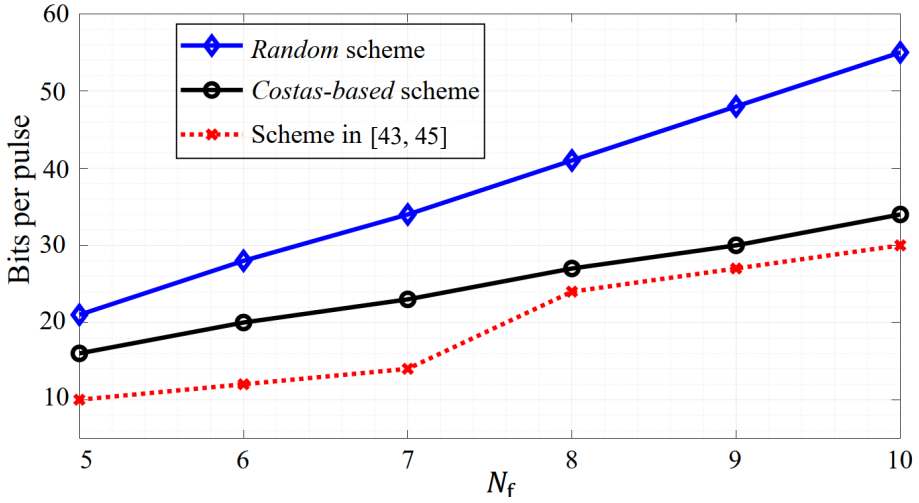


Figure 5.6 : Maximum number of bits per pulse for different data embedding schemes.

permutations for data embedding. Therefore, the proposed techniques are more suitable when the number of antennas is limited, while the technique in [51, 53] is suitable when an extensive number of antennas is available.

Fig. 5.7 illustrates the symbol error rate (SER) of the two proposed schemes and those of the techniques in [53] as functions of the signal-to-noise ratio (SNR). As can be seen, the two proposed schemes have lower SERs compared to those of the techniques in [53] for all SNRs. Distinctively, at low SNRs (i.e., ≤ -2 dB), the two proposed schemes have similar SERs, and are about 4 dB and 10 dB better than those of the FH and BPSK techniques in [53], respectively. This is because the proposed techniques use the CWD, which contains the auto-correlation operation, to generate CWD–TFIs of the signals. As such, the proposed techniques are more robust against noise compared to the other techniques in [53].

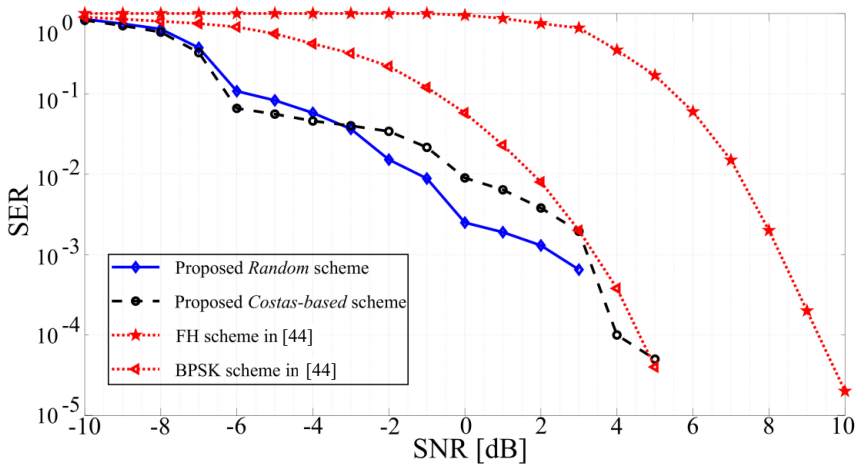


Figure 5.7 : Symbol error rate (SER) of different embedding and demodulation schemes.

5.6 Conclusions

In this chapter, novel approaches for embedding and demodulating data bits in FH JRC systems were proposed. For data embedding, both sub-pulse durations and frequencies are used to increase the data transmission rate. For data demodulation, to reduce the error rate, a CWD–TFI and YOLO-based demodulation scheme that does not require channel estimation and is robust against Doppler shift and timing offset between the JRC transceiver and communication receiver is proposed. Simulation results have shown that the proposed techniques achieve higher data transmission and lower symbol error rates than the existing ones.

Chapter 6

Jamming Mitigation in Joint Communication and Radar Systems: A Deep Dueling Q-learning Approach

This chapter studies the optimization of the durations of the jamming nullspace estimation, the preamble, and the data transmission phases in a dynamic environment. To deal with the uncertainty in the environment, we reformulate the problem using a Markov decision process (MDP). Then, we design a deep dueling Q-learning-based technique to quickly obtain the optimal policy. The contributions of the proposed technique are summarized as follows.

6.1 Problem Formulation

In this section, we first describe the JCR protocol. Then, the signal models and the performance metrics of the JCR system are presented. Finally, we formulate the optimization problem.

6.1.1 Joint Communication and Radar Protocol

Fig. 6.1 illustrates the protocol of the JCR system in the presence of a proactive jammer. As shown, each frame contains three phases: nullspace estimation, preamble, and data transmission.

- During the nullspace estimation phase, which lasts for N^{e} samples, the beam-forming matrices are estimated. Let $\hat{\mathbf{F}}^{\mathrm{e}}_{\mathrm{BS}}$ and $\hat{\mathbf{F}}^{\mathrm{e}}_{\mathrm{UE}}$ denote the beam-forming matrix used to suppress the received jamming signals at the BS and the UE, respectively. The beam-forming matrices $\hat{\mathbf{F}}^{\mathrm{e}}_{\mathrm{BS}}$ and $\hat{\mathbf{F}}^{\mathrm{e}}_{\mathrm{UE}}$ are estimated using the

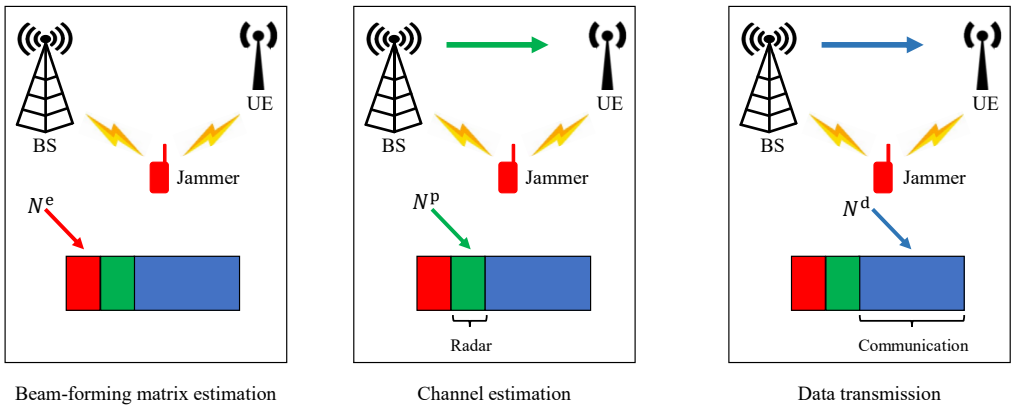


Figure 6.1 : JCR system with proactive jammers.

SVD of the received jamming signals' covariance matrices, using the process similar to that described in Chapter 3.

- During the preamble phase, which lasts for N^{p} samples, the jamming signals that arrive at the UE are nullified by using the beam-forming matrix. Then, the BS-UE equivalent channel is estimated.

- During the data transmission phase, which lasts for N^{d} samples, the BS sends data to UEs. Concurrently, the BS receives the reflected preamble and performs the radar function.

6.1.2 Signal Model

When the optimal beam-forming [107] is applied at the BS for jamming suppression, the radar signal received at the BS after the optimal beam-forming and matched filtering can be represented as

$$\hat{y}_{\mathrm{r}}[n] = G_{\mathrm{r}} x[n - 2df_{\mathrm{s}}/c] e^{j4\pi vt/\lambda} + \mathcal{E}_{\mathrm{r}} \sigma_{\mathrm{J}} x_{\mathrm{J}}[n] + w[n], \tag{6.1}$$

where n is the sample index, $x[n]$ denotes the reflected preamble, G_{r} is the (two-way) radar channel gain, d is the distance between the BS and the target, f_{s} is

the sampling frequency, c denotes the speed of light, λ is the signal wavelength of the preamble, $\mathcal{E}_{\mathrm{r}}$ is the reduction ratio of the optimal beam-forming based jamming suppression [107] at the BS, σ_{J} is the standard deviation of the jamming signal, and w is the AWGN. On the other hand, the received signal at the UE after the optimal beam-forming and time synchronization can be represented as

$$\hat{y}_{\mathrm{c}}[n] = G_{\mathrm{c}} x[n] + \mathcal{E}_{\mathrm{c}} \sigma_{\mathrm{J}} x_{\mathrm{J}}[n] + w[n], \tag{6.2}$$

where G_{c} denotes the (one-way) communication channel gain, and $\mathcal{E}_{\mathrm{c}}$ is the reduction ratio of the optimal beam-forming based jamming suppression at the UE.

It is worth noting that optimal beam-forming requires knowledge of both legitimate and jamming signals. However, in our case, such knowledge is unavailable, and hence the beam-forming matrix is estimated using the orthogonal projection method [107], which design the beam-forming matrix from the SVD of the received jamming signals as demonstrated in Chapter 3. Accordingly, the radar signal and the communication signal are represented as [107],

$$y_{\mathrm{r}}[n] = G_{\mathrm{r}} x\left[n - \frac{2d}{c} f_{\mathrm{s}}\right] e^{j4\pi v t/\lambda} + \frac{N^{\mathrm{e}} + N^{\mathrm{J}}}{N^{\mathrm{e}}} \left\{ \mathcal{E}_{\mathrm{r}} \sigma_{\mathrm{J}} x_{\mathrm{J}}[n] + w[n] \right\}, \tag{6.3}$$

$$y_{\mathrm{c}}[n] = G_{\mathrm{c}} x[n] + \frac{N^{\mathrm{e}} + N^{\mathrm{J}}}{N^{\mathrm{e}}} \left\{ \mathcal{E}_{\mathrm{c}} \sigma_{\mathrm{J}} x_{\mathrm{J}}[n] + w[n] \right\}. \tag{6.4}$$

where N^{e} and N^{J} are the number of samples in the nullspace estimation phase and the number of jammers, respectively.

6.1.3 Performance Metrics

In this section, we present the performance metrics for the radar function and communication functions.

6.1.3.1 Radar Function

For the radar function, the Cramer-Rao lower bounds (CRB) [108] of the range and velocity estimations are adopted. Specifically, the CRB for the range and ve-

locity estimations are given by [108]

$$\mathrm{CRB_d} = \frac{3\lambda^2}{8\pi^2 N^{\mathrm{p}} \mathrm{SINR_r}}, \tag{6.5}$$

$$\mathrm{CRB_v} = \frac{3\lambda^2 B^2}{8\pi^2 (N^{\mathrm{p}})^3 \mathrm{SINR_r}}, \tag{6.6}$$

where

$$\mathrm{SINR_r} = \frac{N^{\mathrm{e}}}{N^{\mathrm{e}} + N^{\mathrm{J}}} \frac{G_{\mathrm{r}}}{\sqrt{\sigma_{\mathrm{J}}^2 + \sigma_{\mathrm{w}}^2}} \tag{6.7}$$

is the SINR of the received radar signal at the BS using the beam-forming matrix, and σ_{w}^2 is the noise variance. As can be seen, by increasing the preamble samples, the CRB for the range and velocity estimation decreases dramatically, especially for the velocity estimation. Similarly, when the number of samples in the estimation phase increase, the received radar signal's $\mathrm{SINR_r}$ is enlarged, resulting in a better radar performance (i.e., for both range and velocity estimations).

6.1.3.2 Communication Function

For the communication function, similar to [57, 60], we adopt the effective minimum mean square error (MMSE) metric as

$$\mathrm{MMSE_{eff}} = \frac{1}{(1 + \mathrm{SINR_c})^{\alpha}}, \tag{6.8}$$

where

$$\alpha = \frac{N^{\mathrm{d}}}{N^{\mathrm{e}} + N^{\mathrm{p}} + N^{\mathrm{d}}}, \tag{6.9}$$

$$\mathrm{SINR_c} = \frac{N^{\mathrm{e}}}{N^{\mathrm{e}} + N^{\mathrm{J}}} \frac{G_{\mathrm{c}}}{\sqrt{\sigma_{\mathrm{J}}^2 + \sigma_{\mathrm{w}}^2}} \tag{6.10}$$

is the SINR of the received communication signal at the UE using the beam-forming matrix. As can be seen, increasing the data transmission samples enlarge the value of $\mathrm{SINR_c}$, hence improving the communication performance.

6.1.4 Problem Formulation

In this section, we develop an optimization function to decide the optimal value for the number of samples in the nullspace estimation, preamble, and data transmission phases. Similar to [55, 58], we use a weighted sum of the MMSE metric of the communication function and the CRB of the radar function as the objective function. Specifically, the optimization of the durations of the nullspace estimation, preamble and data transmission phases can be formally stated as

$$\max_{N^{\mathrm{e}},N^{\mathrm{p}},N^{\mathrm{d}}} \quad \omega_{\mathrm{d}} \log(\mathrm{CRB}_{\mathrm{d}}) + \omega_{\mathrm{v}} \log(\mathrm{CRB}_{\mathrm{v}}) + \omega_{\mathrm{c}} \log(\mathrm{MMSE}_{\mathrm{eff}}) \tag{6.11}$$

$$\begin{aligned}
\textbf{s.t.} \qquad \mathrm{CRB}_{\mathrm{d}} &= \frac{3\lambda^2}{8\pi^2 N^{\mathrm{p}} \mathrm{SINR}_{\mathrm{r}}}, \\
\mathrm{CRB}_{\mathrm{v}} &= \frac{3\lambda^2 B^2}{8\pi^2 (N^{\mathrm{p}})^3 \mathrm{SINR}_{\mathrm{r}}}, \\
\mathrm{MMSE}_{\mathrm{eff}} &= \frac{1}{(1+\mathrm{SINR}_{\mathrm{c}})^{\alpha}}, , \\
\mathrm{SINR}_{\mathrm{r}} &= \frac{N^{\mathrm{e}}}{N^{\mathrm{e}}+N^{\mathrm{J}}} \frac{G_{\mathrm{r}}}{\sqrt{\sigma_{\mathrm{J}}^2+\sigma_{\mathrm{w}}^2}} \\
\alpha &= \frac{N^{\mathrm{d}}}{N^{\mathrm{e}}+N^{\mathrm{p}}+N^{\mathrm{d}}}, \\
\mathrm{SINR}_{\mathrm{c}} &= \frac{N^{\mathrm{e}}}{N^{\mathrm{e}}+N^{\mathrm{J}}} \frac{G_{\mathrm{c}}}{\sqrt{\sigma_{\mathrm{J}}^2+\sigma_{\mathrm{w}}^2}}
\end{aligned}$$

where the weighting factor ω_{d}, ω_{v}, and ω_{c} are used to determine the importance of the range estimation, velocity estimation, and communication function, respectively.

Therefore, the JCR system needs to obtain an optimal policy that optimizes the duration of each sub-phase to achieve the desired JCR performance. However, the JCR's surrounding environment often varies significantly over time, making optimizing the sub-phase duration an intractable problem. In the following section, we describe our proposed MDP framework for the JCR that enables the JCR to quickly and effectively learn the optimal policy without requiring complete information from the surrounding environment, thereby achieving optimal performance compared with traditional solutions.

6.2 Deep Dueling Q-Learning Technique for Jamming Suppression

In this section, we reformulate the problem in (6.11) using an SMDP. A deep dueling Q-learning-based framework is then designed to tune the durations of the nullspace estimation, preamble, and data transmission phases by obtaining the optimal policy for the underlying SMDP process.

6.2.1 SMDP

We use the SMDP [109] to maximize the objective function in Eq. (6.11) while accounting for the changes in the values of N^{e}, N^{p} and N^{d}. A conventional MDP is defined by a tuple $(\mathcal{S}, \mathcal{A}, r)$, where $\mathcal{S}$, $\mathcal{A}$, and r denote the state space, action space, and the reward function, respectively. An SMDP, on the other hand, retains the three components mentioned above and adds an additional component, that is, the nth decision epoch length, denoted by $t[n]$. In an MDP, the state transition occurs at regular time steps (and hence the decision epoch length $t[n]$ is excluded). On the other hand, the SMDP allows the state transition to occur at irregular time steps (i.e., different $t[n]$ for different epochs). Therefore, the SMDP is suitable to formulate the problem of optimizing the selections of N^{e}, N^{p}, and N^{d} at irregular state transition times.

6.2.2 State Space

There are two essential factors to consider for maximizing the objective function in Eq. 6.11. The first factor is the SINR at the UE during the previous data transmission phase. This is because $\mathrm{SINR_c}$ implicitly captures the BS-UE channel condition that affects the selection of N^{d}. Specifically, when the SINR in the previous data transmission phase is poor, N^{d} value can be decreased to improve the value of $\mathrm{SINR_c}$. However, N^{d} should not be incautiously decreased, as that reduces the

fraction of the data transmission phase over the whole frame and thus reduces the effective spectral efficiency of the system. The second factor is the SINR of the reflected signal at the BS. This is because SINR_r affect the CRB of both range estimation and velocity estimation, as demonstrated in Eq. (6.5) and Eq. (6.6), respectively. Therefore, the system's state space can be defined as

$$\mathcal{S} \triangleq \big\{(\text{SINR}_\text{c}, \text{SINR}_\text{r})\big\}. \tag{6.12}$$

The method to calculate SINR_c is described in Chapter 4. On the other hand, the process to determine SINR_r can be found in [108].

6.2.3 Action

An action is taken at the end of each frame to determine which are the next N^e, N^p, and N^d values, given the current state. Considering a fixed number of samples for each frame (i.e., $N^\text{e} + N^\text{p} + N^\text{e} + N^\text{p} = \text{constant}$), the problem of determining N^e, N^p, and N^d becomes the problem of determining N^e and N^p. Let $\mathcal{N}^\text{e} \triangleq (N_1^\text{e}, N_2^\text{e}, ..., N_{L^\text{e}}^\text{e})$ and $\mathcal{N}^\text{p} \triangleq (N_1^\text{p}, N_2^\text{p}, ..., N_{L^\text{p}}^\text{p})$ be the sets of L^e and L^p candidates for N^e and N^p, respectively. The action space is defined as $\mathcal{A} \triangleq \{a : a \in (1, 2, ..., L^\text{e} \times L^\text{p})\}$, and

$$a = \begin{cases} 1, & N^\text{e} = N_1^\text{e} \text{ and } N^\text{p} = N_1^\text{p}, \\ 2, & N^\text{e} = N_2^\text{e} \text{ and } N^\text{p} = N_1^\text{p}, \\ ... & \\ L^\text{e} \times L^\text{p}, & N^\text{e} = N_{L^\text{e}}^\text{e} \text{ and } N^\text{p} = N_{L^\text{p}}^\text{p}. \end{cases}$$

6.2.4 Immediate Reward

The immediate reward is defined as the weighted sum of the SINRs of the signal at the UE and the reflected signal at the BS. Specifically,

$$r[n] = \omega_r \text{SINR}_\text{r}[n] + \omega_c \text{SINR}_\text{c}[n] \tag{6.13}$$

6.2.5 Optimization Formulation

Similar to chapter 4, let $\pi : \hat{\mathcal{S}} \rightarrow \mathcal{A}$ denotes a policy which is a mapping function from the approximate states to the actions taken by the system. Our purpose is to find the optimal value of π, denoted by π^*, that maximizes the average long-term reward [42] of the BS-UE communication system. The optimization problem in Eq. (4.8) is then converted into the optimization problem of π^*, expressed by

$$\max_{\pi} \quad \mathcal{R}(\pi) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{E}\{r[n]\}, \tag{6.15}$$

where $\mathcal{R}(\pi)$ denotes the average long-term reward of the system with the policy π. In fact, unlike the Q-learning algorithm that guarantees its convergence, such a guarantee is not possible for works that apply deep reinforcement learning techniques. However, because the state, action and reward are carefully designed, we anticipate that the proposed technique can converge to the optimal policy, as demonstrated in Chapter 4.

Chapter 7

Conclusions and Future Work

7.1 Conclusion

This thesis designed novel methods to deal with multiple jammers by leveraging signal beam-forming and ML techniques. We have found that by varying the correlation coefficients between transmitted jamming signals, jammers can "virtually change" the jamming channels hence their nullspace even when these channels do not physically change. That makes most conventional interference suppression techniques that rely on the beam-forming matrix derived from the nullspace no longer applicable. To tackle the problem, the techniques to monitor the jamming residual and effectively update the beam-forming matrix were developed.

In Chapter 4, it was found that a jamming residual monitoring process proposed in Chapter 3 increases the system overhead, thus significantly reducing the spectral efficiency. To solve the problem and to deal with the uncertainty and incomplete information, the problem was reformulated using a partially observable semi-Markov decision process (POSMDP). Then, a deep dueling Q-learning-based technique was designed to quickly obtain the optimal policy for legitimate devices. Simulation results showed that the proposed technique achieved a much higher spectral efficiency than that of other methods, close to that of the perfect jamming nullification case.

Next, in Chapter 5, the jamming mitigation study was extended to the joint radar and communication (JRC) systems. Specifically, a frequency hopping (FH) signal was designed, which possesses robustness against jamming, for a JRC system. In particular, novel techniques to embed and demodulate data to increase the data

rate and reduce the demodulation error are also investigated. Specifically, both subpulse frequencies and durations are used for data modulation, resulting in higher communication data rates. A novel scheme for data demodulation was also proposed by using the time-frequency analysis (TFA) technique and a 'You Only Look Once' (YOLO)-based detection system. Simulation results have shown that the proposed techniques achieve higher data rates and lower demodulation errors compared to the existing ones.

Finally, in Chapter 6, we study the optimization of the durations of the jamming nullspace estimation, the preamble, and the data transmission phases. Specifically, in the jamming nullspace estimation phase, the beam-forming matrix used to mitigate the jamming signal is estimated. On the other hand, the preamble is used to estimate the legitimate channel and also for the radar function. As such, increasing the duration of the nullspace estimation and the preamble phases can increase the radar's performance. However, such an increase also reduces the effective spectral efficiency of the communication function, because the data transmission phase fraction is decreased. Moreover, the surrounding radio environments of the JRC systems are typically dynamic with high uncertainties due to their high mobility, making the durations optimization problem even more challenging. To deal with such uncertainty, we reformulate the problem using a Markov decision process (MDP). Then, we design a deep dueling Q-learning-based technique to quickly obtain the optimal policy.

7.2 Future Works

ML in general and RL in particular possess great potential in dealing with problems that are difficult to be handled with conventional techniques. Therefore, ML and RL have a lot of potential to solve a number of outstanding problems, including those outlined below.

- *RL for power allocation in JCR system:* There have been many studies on power allocation in MIMO communication systems. However, optimization for the JCR system will be much more complicated, because the optimization objective is not only communication performance but also the radar function (e.g., detection, range and velocity estimation). The application of RL can assist the system in quickly obtaining the optimal solution.

- *ML for signal detection in JRC system:* One challenge in the JCR system is to differentiate the reflected signal from the target (for the radar function) and the up-link signal from the user under interference from other BS and noise. Supervised ML can be a potential candidate to classify the received signal at the BS, thereby detecting the reflected signal for radar function.

- *ML against eavesdropping in JRC system:* By sharing the spectrum with the communication function, the radar function may unintentionally leak sensitive information to the commercial communication system, or even worse, to the eavesdroppers. ML can be effective in dealing with eavesdroppers, hence securing sensitive data.

Appendix A

Proofs in Chapter 3

A.1 The proof of Theorem 3.1

Because $\mathbf{x}_{\mathrm{J}}^{\mathrm{e}} \sim \mathcal{CF}(\boldsymbol{\mu}_{\mathrm{J}}^{\mathrm{e}}, \boldsymbol{\Sigma}_{\mathrm{J}}^{\mathrm{e}})$, we have $\mathbb{E}(\mathbf{X}_{\mathrm{J}}^{\mathrm{e}}) = \boldsymbol{\mu}_{\mathrm{J}}^{\mathrm{e}}$ and $\mathbb{E}(\mathbf{X}_{\mathrm{J}}^{\mathrm{e}}(\mathbf{X}_{\mathrm{J}}^{\mathrm{e}})^H) = N^{\mathrm{e}}\boldsymbol{\Sigma}_{\mathrm{J}}^{\mathrm{e}}$.

Then, recalling that $\mathbf{Z}_k$ is deterministic, we have

$$\begin{aligned}
\mathbb{E}(\mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}}) &= \mathbb{E}\left[\frac{1}{N^{\mathrm{e}}}\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}}(\mathbf{Y}_{\mathrm{J}_k}^{\mathrm{e}})^H\right] \\
&= \frac{1}{N^{\mathrm{e}}}\mathbb{E}\left[(\mathbf{Z}_k\mathbf{X}_{\mathrm{J}}^{\mathrm{e}} + \mathbf{W})(\mathbf{Z}_k\mathbf{X}_{\mathrm{J}}^{\mathrm{e}} + \mathbf{W})^H\right] \\
&= \frac{1}{N^{\mathrm{e}}}\mathbb{E}\left[\mathbf{Z}_k\mathbf{X}_{\mathrm{J}}^{\mathrm{e}}(\mathbf{X}_{\mathrm{J}}^{\mathrm{e}})^H\mathbf{Z}_k^H + \mathbf{W}(\mathbf{X}_{\mathrm{J}}^{\mathrm{e}})^H\mathbf{Z}_k^H + \mathbf{Z}_k\mathbf{X}_{\mathrm{J}}^{\mathrm{e}}\mathbf{W}^H + \mathbf{W}\mathbf{W}^H\right] \\
&= \frac{1}{N^{\mathrm{e}}}\mathbb{E}\left[\mathbf{Z}_k\mathbf{X}_{\mathrm{J}}^{\mathrm{e}}(\mathbf{X}_{\mathrm{J}}^{\mathrm{e}})^H\mathbf{Z}_k^H\right] + \frac{1}{N^{\mathrm{e}}}\mathbb{E}\left[\mathbf{W}(\mathbf{X}_{\mathrm{J}}^{\mathrm{e}})^H\mathbf{Z}_k^H\right] + \frac{1}{N^{\mathrm{e}}}\mathbb{E}(\mathbf{Z}_k\mathbf{X}_{\mathrm{J}}^{\mathrm{e}}\mathbf{W}^H) \\
&\quad + \frac{1}{N^{\mathrm{e}}}\mathbb{E}(\mathbf{W}\mathbf{W}^H) \\
&= \mathbf{Z}_k\boldsymbol{\Sigma}_{\mathrm{J}}^{\mathrm{e}}\mathbf{Z}_k^H + \sigma_{\mathrm{w}}^2\mathbf{I}_{N_k}. \qquad \text{(A.1)}
\end{aligned}$$

Therefore, by the law of large number, the covariance matrix $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}}$ converges in probability [R12, p. 175] to $(\mathbf{Z}_k\boldsymbol{\Sigma}_{\mathrm{J}}^{\mathrm{e}}\mathbf{Z}_k^H + \sigma_{\mathrm{w}}^2\mathbf{I}_{N_k})$ i.e., $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}} \xrightarrow{p} (\mathbf{Z}_k\boldsymbol{\Sigma}_{\mathrm{J}}^{\mathrm{e}}\mathbf{Z}_k^H + \sigma_{\mathrm{w}}^2\mathbf{I}_{N_k})$.

A.2 The proof of Theorem 3.2

Because the nullspace $\hat{\mathbf{G}}_k$ is calculated from the singular value decomposition (SVD) of $\mathbf{R}_{\mathrm{J}_k}$ as described in Section 3.2, the change of $\mathbf{R}_{\mathrm{J}_k}$ from $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{e}}$ to $\mathbf{R}_{\mathrm{J}_k}^{\mathrm{d}}$ due

to the change of $\mathbf{\Sigma}_{\mathrm{J}}$ from $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}}$ to $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{d}}$ are examined as follows.

$$\begin{aligned}
\mathbf{R}_{\mathrm{J}_k}^{\mathrm{d}} &\xrightarrow{p} (\mathbf{Z}_k \mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{d}} \mathbf{Z}_k^H + \sigma_{\mathrm{w}}^2 \mathbf{I}_{N_k}) \\
&= \mathbf{Z}_k \mathbf{V}^{\mathrm{d}} \mathbf{S}^{\mathrm{d}} (\mathbf{V}^{\mathrm{d}})^H \mathbf{Z}_k^H + \sigma_{\mathrm{w}}^2 \mathbf{I}_{N_k} \\
&= \mathbf{Z}_k \mathbf{V}^{\mathrm{d}} \sqrt{\mathbf{S}^{\mathrm{d}} (\mathbf{S}^{\mathrm{e}})^{-1}} \mathbf{S}^{\mathrm{e}} \sqrt{\mathbf{S}^{\mathrm{d}} (\mathbf{S}^{\mathrm{e}})^{-1}} (\mathbf{V}^{\mathrm{d}})^H \mathbf{Z}_k^H + \sigma_{\mathrm{w}}^2 \mathbf{I}_{N_k} \\
&= \mathbf{Z}_k \mathbf{V}^{\mathrm{d}} \sqrt{\mathbf{S}^{\mathrm{d}} (\mathbf{S}^{\mathrm{e}})^{-1}} (\mathbf{V}^{\mathrm{e}})^H \mathbf{V}^{\mathrm{e}} \mathbf{S}^{\mathrm{e}} (\mathbf{V}^{\mathrm{e}})^H \mathbf{V}^{\mathrm{e}} \sqrt{\mathbf{S}^{\mathrm{d}} (\mathbf{S}^{\mathrm{e}})^{-1}} (\mathbf{V}^{\mathrm{d}})^H \mathbf{Z}_k^H + \sigma_{\mathrm{w}}^2 \mathbf{I}_{N_k} \\
&= (\mathbf{Z}_k \mathbf{D}) \mathbf{V}^{\mathrm{e}} \mathbf{S}^{\mathrm{e}} (\mathbf{V}^{\mathrm{e}})^H (\mathbf{Z}_k \mathbf{D})^H + \sigma_{\mathrm{w}}^2 \mathbf{I}_{N_k} \\
&= (\mathbf{Z}_k \mathbf{D}) \mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}} (\mathbf{Z}_k \mathbf{D})^H + \sigma_{\mathrm{w}}^2 \mathbf{I}_{N_k}, \qquad \text{(A.2)}
\end{aligned}$$

where $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{d}} = \mathbf{V}^{\mathrm{d}} \mathbf{S}^{\mathrm{d}} (\mathbf{V}^{\mathrm{d}})^H$ and $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}} = \mathbf{V}^{\mathrm{e}} \mathbf{S}^{\mathrm{e}} (\mathbf{V}^{\mathrm{e}})^H$ are the SVD of $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{d}}$ and $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}}$, respectively, and $\mathbf{D} = \mathbf{V}^{\mathrm{d}} \sqrt{\mathbf{S}^{\mathrm{d}} (\mathbf{S}^{\mathrm{e}})^{-1}} (\mathbf{V}^{\mathrm{e}})^H$. Therefore, the change of $\mathbf{\Sigma}_{\mathrm{J}}$ from $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{e}}$ to $\mathbf{\Sigma}_{\mathrm{J}}^{\mathrm{d}}$ causes an impact to the nullspace estimation, similar to the jamming channel has been "virtually changed" from $\mathbf{Z}_k$ to $(\mathbf{Z}_k \mathbf{D})$.

Appendix B

Proofs in Chapter 4

B.1 The proof of Theorem 4.1

The received signal at the kth UE after zero-forcing equalization can be expressed by

$$\mathbf{y}_k^{\mathrm{ZF}} = (\tilde{\mathbf{H}}_k^H \tilde{\mathbf{H}}_k)^{-1} \tilde{\mathbf{H}}_k^H (\sqrt{P_{\mathrm{T}}} \tilde{\mathbf{H}}_k \mathbf{x}_k + \tilde{\mathbf{Z}}_k \mathbf{x}_{\mathrm{J}} + \tilde{\mathbf{w}}) = \sqrt{P_{\mathrm{T}}} \mathbf{x}_k + \mathbf{A}_k^{\mathrm{ZF}} (\tilde{\mathbf{Z}}_k \mathbf{x}_{\mathrm{J}} + \tilde{\mathbf{w}}), \quad \text{(B.1)}$$

where $\mathbf{A}_k^{\mathrm{ZF}} \triangleq (\tilde{\mathbf{H}}_k^H \tilde{\mathbf{H}}_k)^{-1} \tilde{\mathbf{H}}_k^H$ is the zero-forcing equalizer, $\tilde{\mathbf{H}}_k$, $\tilde{\mathbf{Z}}_k$, and $\tilde{\mathbf{w}}$ are the BS-kth UE equivalent channel, jammer-kth UE equivalent channel, and equivalent noise, respectively. The values of $\tilde{\mathbf{H}}_k$, $\tilde{\mathbf{Z}}_k$, and $\tilde{\mathbf{w}}$ can be expressed by

$$\tilde{\mathbf{H}}_k = \tilde{\mathbf{H}}_k^{\mathrm{wbf}} = \mathbf{H}_k \mathbf{P}_k, \quad \tilde{\mathbf{Z}}_k = \mathbf{Z}_k, \quad \tilde{\mathbf{w}} = \mathbf{w}, \qquad \text{without beam-forming,}$$

$$\tilde{\mathbf{H}}_k = \tilde{\mathbf{H}}_k^{\mathrm{bf}} = \hat{\mathbf{F}}_k^{\mathrm{e}} \mathbf{H}_k \mathbf{P}_k, \quad \tilde{\mathbf{Z}}_k = \hat{\mathbf{F}}_k^{\mathrm{e}} \mathbf{Z}_k, \quad \tilde{\mathbf{w}} = \hat{\mathbf{F}}_k^{\mathrm{e}} \mathbf{w}, \qquad \text{with beam-forming.}$$

First, considering the case without beam-forming, from Eq. (B.1), the post-equalization SINR of the mth stream for the kth UE can be expressed by

$$\delta_{k,m}^{\mathrm{wbf}} = \frac{P_{\mathrm{T}} \mathrm{Var}(x_{k,m})}{\mathrm{Var}\{[\mathbf{A}_k^{\mathrm{ZF}} (\mathbf{Z}_k \mathbf{x}_{\mathrm{J}} + \mathbf{w})]_m\}} = \frac{P_{\mathrm{T}}}{\mathrm{Var}[(\mathbf{A}_k^{\mathrm{ZF}} \mathbf{w}')_m]} = \frac{P_{\mathrm{T}}}{\mathrm{Var}[(\mathbf{A}_k^{\mathrm{ZF}})_m \mathbf{w}']}, \quad \text{(B.2)}$$

where $(.)_m$ denotes the mth row of a matrix or the mth element of a vector, and

$$\mathbf{w}' \triangleq (\mathbf{Z}_k \mathbf{x}_{\mathrm{J}} + \mathbf{w}) \in \mathbb{C}^{N_k \times 1}. \quad \text{(B.3)}$$

Using the law of total variance [110], we have

$$\mathrm{Var}[(\mathbf{A}_k^{\mathrm{ZF}})_m \mathbf{w}'] = \mathbb{E}\{\mathrm{Var}[(\mathbf{A}_k^{\mathrm{ZF}})_m \mathbf{w}' | \mathbf{w}']\} + \mathrm{Var}\{\mathbb{E}[(\mathbf{A}_k^{\mathrm{ZF}})_m \mathbf{w}' | \mathbf{w}']\}, \quad \text{(B.4)}$$

where (.|.) denotes the conditional operation.

For the first term in the right-hand side (RHS) of Eq. (B.4),

$$\mathbb{E}\{\mathrm{Var}[(\mathbf{A}_k^{\mathrm{ZF}})_m\mathbf{w}'|\mathbf{w}']\} = \sigma_{\mathrm{w}'}^2\mathbb{E}\{[\mathbf{A}_k^{\mathrm{ZF}}(\mathbf{A}_k^{\mathrm{ZF}})^H]_{mm}\}, \tag{B.5}$$

that is because $\mathbf{w}'$ is given and considered constant, and the zero-forcing equalizer $\mathbf{A}_k^{\mathrm{ZF}}$ and $\mathbf{w}'$ are independent. Moreover,

$$\begin{aligned}\mathbb{E}\left\{\mathbf{A}_k^{\mathrm{ZF}}(\mathbf{A}_k^{\mathrm{ZF}})^H\right\} &= \mathbb{E}\{[(\tilde{\mathbf{H}}_k^{\mathrm{wbf}})^H\tilde{\mathbf{H}}_k^{\mathrm{wbf}}]^{-1}(\tilde{\mathbf{H}}_k^{\mathrm{wbf}})^H\tilde{\mathbf{H}}_k^{\mathrm{wbf}}[(\tilde{\mathbf{H}}_k^{\mathrm{wbf}})^H\tilde{\mathbf{H}}_k^{\mathrm{wbf}}]^{-1}\} \\ &= \mathbb{E}\{[(\tilde{\mathbf{H}}_k^{\mathrm{wbf}})^H\tilde{\mathbf{H}}_k^{\mathrm{wbf}}]^{-1}\} \overset{\text{(a)}}{=} \frac{\eta_k\mathbf{I}_{M_k}}{N_k - M_k},\end{aligned} \tag{B.6}$$

where (a) follows because $\tilde{\mathbf{H}}_k^{\mathrm{wbf}} \in \mathbb{C}^{N_k\times M_k} \sim \mathcal{CN}(\mathbf{0}, 1/\sqrt{\eta_k})$ and hence $[(\tilde{\mathbf{H}}_k^{\mathrm{wbf}})^H\tilde{\mathbf{H}}_k^{\mathrm{wbf}}]$ has a complex Wishart distribution with $\mathbb{E}\{[(\tilde{\mathbf{H}}_k^{\mathrm{wbf}})^H\tilde{\mathbf{H}}_k^{\mathrm{wbf}}]^{-1}\} = \eta_k\mathbf{I}_{M_k}/(N_k - M_k)$ [42]. Note that, $\tilde{\mathbf{H}}_k^{\mathrm{wbf}} = \mathbf{H}_k\mathbf{P}_k \sim \mathcal{CN}(\mathbf{0}, 1/\sqrt{\eta_k})$ because the columns of $\mathbf{P}_k$ are selected from these of the Walsh–Hadamard matrix, and hence multiplying the legitimate channel $\mathbf{H}_k$ with $\mathbf{P}_k$ does not change the channel distribution.

Furthermore, from Eq. (B.3), the elements of $\mathbf{w}'$ have the same variance $\sigma_{\mathrm{w}'}^2 = \sigma_{\mathrm{w}}^2 + \sum_{j=1}^{N_{\mathrm{J}}}\eta_{k,j}\sigma_{\mathrm{J}_j}^2$. Therefore, from Eq. (B.5) and Eq. (B.6), the first term in the RHS of Eq. (B.4) equal to

$$\mathbb{E}\{\mathrm{Var}[(\mathbf{A}_k^{\mathrm{ZF}})_m\mathbf{w}'|\mathbf{w}']\} = \frac{\eta_k(\sigma_{\mathrm{w}}^2 + \sum_{j=1}^{N_{\mathrm{J}}}\eta_{k,j}\sigma_{\mathrm{J}_j}^2)}{N_k - M_k}. \tag{B.7}$$

For the second term in the right-hand side of (B.4),

$$\mathrm{Var}\{\mathbb{E}[(\mathbf{A}_k^{\mathrm{ZF}})_m\mathbf{w}'|\mathbf{w}']\} = \sigma_{\mathrm{w}'}^2(\mathbb{E}\{\mathbb{E}[(\mathbf{A}_k^{\mathrm{ZF}})_m]^2\} - \mathbb{E}\{\mathbb{E}[(\mathbf{A}_k^{\mathrm{ZF}})_m]\}^2) = 0. \tag{B.8}$$

The ergodic spectral efficiency of the mth stream for the kth UE is therefore expressed by

$$C_{k,m}^{\mathrm{wbf}} = \log_2\big[1 + \frac{P_{\mathrm{T}}(N_k - M_k)}{\eta_k(\sigma_{\mathrm{w}}^2 + \sum_{j=1}^{N_{\mathrm{J}}}\eta_{k,j}\sigma_{\mathrm{J}_j}^2)}\big]. \tag{B.9}$$

When the estimated beam-forming matrix is applied, the post-equalization SINR of the mth stream for the kth UE is

$$\delta_{k,m} = \frac{P_{\mathrm{T}}}{\mathrm{Var}\{[\mathbf{A}_k^{\mathrm{ZF}}\hat{\mathbf{F}}_k^{\mathrm{e}}(\mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \mathbf{w})]_m\}} \tag{B.10}$$

$$= \frac{P_{\mathrm{T}}}{\mathrm{Var}\{[\mathbf{A}_k^{\mathrm{ZF}}(\delta\mathbf{F}_k\mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \hat{\mathbf{F}}_k^{\mathrm{e}}\mathbf{w})]_m\}}, \tag{B.11}$$

where $\delta\mathbf{F}_k = \hat{\mathbf{F}}_k^{\mathrm{e}} - \mathbf{G}_k^{\mathrm{d}}$ is the nullspace estimation error. Note that for simplicity, the zero-forcing equalizer is still denoted by $\mathbf{A}_k^{\mathrm{ZF}}$, even though by the use of $\hat{\mathbf{F}}_k^{\mathrm{e}}$, the dimension of $\mathbf{A}_k^{\mathrm{ZF}}$ is now changed to $(N_k - N_{\mathrm{J}}) \times M_k$ from $N_k \times M_k$ (i.e., when $\hat{\mathbf{F}}_k^{\mathrm{e}}$ is not applied).

When the estimated beam-forming matrix is estimated perfectly (i.e., $\delta\mathbf{F}_k = \mathbf{0}$), the jamming signals are cancelled totally, Eq. (B.11) becomes

$$\delta_{k,m}^{\mathrm{ub}} = \frac{P_{\mathrm{T}}}{\mathrm{Var}[(\mathbf{A}_k^{\mathrm{ZF}}\hat{\mathbf{F}}_k^{\mathrm{e}}\mathbf{w})_m]} = \frac{P_{\mathrm{T}}}{\mathrm{Var}[(\mathbf{A}_k^{\mathrm{ZF}}\tilde{\mathbf{w}})_m]}, \tag{B.12}$$

where $\tilde{\mathbf{w}} \triangleq \hat{\mathbf{F}}_k^{\mathrm{e}}\mathbf{w}$.

Note that, because $\hat{\mathbf{F}}_k^{\mathrm{e}}$ is estimated using the SVD, all its rows are unit vectors and orthogonal to each other, and hence the elements of $\tilde{\mathbf{w}}$ are independent and identically distributed (i.i.d) with the same variance $\sigma_{\tilde{\mathrm{w}}}^2 = \sigma_{\mathrm{w}}^2$. Following the derivation in Eqs. (B.4)–(B.9), and remembering that multiplying with the unitary matrix $\hat{\mathbf{F}}_k^{\mathrm{e}}$ only reduces the dimension of the equivalent channel from $\tilde{\mathbf{H}}_k^{\mathrm{wbf}} \in \mathbb{C}^{N_k \times M_k}$ into $\tilde{\mathbf{H}}_k^{\mathrm{bf}} \in \mathbb{C}^{(N_k - N_{\mathrm{J}}) \times M_k}$, but does not change the distribution of the equivalent channel, the spectral efficiency when the beam-forming is estimated perfectly is

$$C_{k,m}^{\mathrm{ub}} = \log_2\big[1 + \frac{P_{\mathrm{T}}(N_k - N_{\mathrm{J}} - M_k)}{\eta_k\sigma_{\mathrm{w}}^2}\big]. \tag{B.13}$$

For the spectral efficiency lower bound, considering Eq. (B.10) in the worst case when $\hat{\mathbf{F}}_k^{\mathrm{e}}$ is independent of $(\mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \mathbf{w})$. Let $\tilde{\mathbf{w}}' \triangleq \hat{\mathbf{F}}_k^{\mathrm{e}}(\mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \mathbf{w})$. Because $\hat{\mathbf{F}}_k^{\mathrm{e}}$ is a unitary matrix, multiplying with $\hat{\mathbf{F}}_k^{\mathrm{e}}$ does not change the distribution of $(\mathbf{Z}_k\mathbf{x}_{\mathrm{J}} + \mathbf{w})$.

Therefore, similar to $\mathbf{w}'$, the elements of $\tilde{\mathbf{w}}'$ have the same variance $\sigma^2_{\tilde{\mathrm{w}}'} = \sigma^2_{\mathrm{w}'} = \sigma^2_{\mathrm{w}} + \sum_{j=1}^{N_\mathrm{J}} \eta_{k,j}\sigma^2_{\mathrm{J}_j}$. From Eq. (B.10),

$$\delta^{\mathrm{lb}}_{k,m} = \frac{P_\mathrm{T}}{\mathrm{Var}\{[\mathbf{A}^{\mathrm{ZF}}_k \hat{\mathbf{F}}^{\mathrm{e}}_k(\mathbf{Z}_k\mathbf{x}_\mathrm{J} + \mathbf{w})]_m\}} = \frac{P_\mathrm{T}}{\mathrm{Var}[(\mathbf{A}^{\mathrm{ZF}}_k \tilde{\mathbf{w}}')_m]}. \tag{B.14}$$

Following the derivation in Eqs. (B.4)–(B.9), and remembering the equivalent channel is now $\tilde{\mathbf{H}}^{\mathrm{bf}}_k \in \mathbb{C}^{(N_k - N_\mathrm{J}) \times M_k}$, the lower bound of the spectral efficiency is

$$C^{\mathrm{lb}}_{k,m} = \log_2\big[1 + \frac{P_\mathrm{T}(N_k - N_\mathrm{J} - M_k)}{\eta_k(\sigma^2_\mathrm{w} + \sum_{j=1}^{N_\mathrm{J}} \eta_{k,j}\sigma^2_{\mathrm{J}_j})}\big]. \tag{B.15}$$

Bibliography

[1] L. M. Hoang, J. A. Zhang, D. Nguyen, X. Huang, A. Kekirigoda, and K.-P. Hui, "Suppression of multiple spatially correlated jammers," *IEEE Trans. Veh. Tech-nol.*, vol. 70, no. 10, pp. 10 489–10 500, 2021.

[2] L. M. Hoang, D. N. Nguyen, J. A. Zhang, and D. T. Hoang, "Multiple correlated jammers nullification using LSTM-based deep dueling neural network," Feb. 2022. [Online]. Available: https://arxiv.org/abs/2202.03600

[3] L. M. Hoang, J. A. Zhang, D. N. Nguyen, and D. T. Hoang, "Frequency hopping joint radar-communications with hybrid sub-pulse frequency and duration modulation," *IEEE Wireless Commun. Lett.*, vol. 37, no. 7, pp. 1–8, 2022.

[4] J. A. Zhang, L. Hoang, D. Nguyen, X. Huang, A. Kekirigoda, and K.-P. Hui, "Multi-user mimo communications with interference mitigation in time-varying channels," in *Proc. IEEE Int. Conf. Sig. Process. Comm.,*. IEEE, Dec. 2020, pp. 1–7.

[5] L. Hoang, J. A. Zhang, D. Nguyen, A. Kekirigoda, and K.-P. Hui, "Nullification of multiple correlated jammers," in *Proc. IEEE Veh. Tech-nol Conf. (VTC2021-Fall).* IEEE, Dec. 2021, pp. 1–6.

[6] L. M. Hoang, D. Nguyen, J. A. Zhang, and D. T. Hoang, "Multiple correlated jammers suppression: A deep dueling q-learning approach," in *Proc. IEEE Wireless Commun. Netw. Conf.* IEEE, Apr. 2022, pp. 998–1003.

[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *IEEE CVPR*, 2016, pp. 779–788.

[8] M. Vanhoef and F. Piessens, "Advanced wi-fi attacks using commodity hardware," in *Proc. Comput. Security Appl. Conf*, 2014, pp. 256–265.

[9] Z. Chen, G. Gokeda, and Y. Yu, *Introduction to Direction-of-arrival Estimation.* Norwood, MA, USA: Artech House, 2010.

[10] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[11] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, 1989.

[12] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 5, pp. 814–824, 1990.

[13] A. J. Fenn, *Adaptive antennas and phased arrays for radar and communications.* Boston, MA, USA: Artech House, 2007.

[14] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, 2014.

[15] C. Popper, M. Strasser, and S. Capkun, "Anti-jamming broadcast communication using uncoordinated spread spectrum techniques," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 5, pp. 703–715, 2010.

[16] M. K. Hanawal, D. N. Nguyen, and M. Krunz, "Jamming attack on in-band full-duplex communications: Detection and countermeasures," in *Proc. 35th*

Annu. IEEE Int. Conf. Comput. Commun., Apr. 2016, pp. 1–9.

[17] Y. Gao, Y. Xiao, M. Wu, M. Xiao, and J. Shao, "Game theory-based anti-jamming strategies for frequency hopping wireless communications," *IEEE Trans. Wireless Commun.*, vol. 17, no. 8, pp. 5314–5326, 2018.

[18] B. Wang, Y. Wu, K. R. Liu, and T. C. Clancy, "An anti-jamming stochastic game for cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 4, pp. 877–889, 2011.

[19] Y. Liu, P. Ning, H. Dai, and A. Liu, "Randomized differential DSSS: Jamming-resistant wireless broadcast communication," in *Proc. IEEE INFOCOM*, 2010, pp. 695—703.

[20] M. Simon, J. Omura, R. Scholtz, and B. Levitt, *Spread spectrum communications handbook*. McGraw-Hill Education, 2002.

[21] T. T. Do, E. Björnson, E. G. Larsson, and S. M. Razavizadeh, "Jamming-resistant receivers for the massive MIMO uplink," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 1, pp. 210–223, 2017.

[22] H. Akhlaghpasand, E. Björnson, and S. M. Razavizadeh, "Jamming suppression in massive MIMO systems," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 67, no. 1, pp. 182–186, 2019.

[23] ——, "Jamming-robust uplink transmission for spatially correlated massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3495–3504, 2020.

[24] Q. Yan, H. Zeng, T. Jiang, M. Li, W. Lou, and Y. T. Hou, "Jamming resilient communication using MIMO interference cancellation," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 7, pp. 1486–1499, 2016.

[25] X. G. Doukopoulos and G. V. Moustakides, "Fast and stable subspace tracking," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1452–1465, 2008.

[26] J.-F. Yang and M. Kaveh, "Adaptive eigensubspace algorithms for direction or frequency estimation and tracking," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 2, pp. 241–251, 1988.

[27] S. Attallah and K. Abed-Meraim, "Low-cost adaptive algorithm for noise subspace estimation," *Electron. Lett.*, vol. 38, no. 12, pp. 609–611, 2002.

[28] S. Attallah, "The generalized Rayleigh's quotient adaptive noise subspace algorithm: A Householder transformation-based implementation," *IEEE Trans. Circuits Syst. I*, vol. 53, pp. 3–7, 2006.

[29] S. Bartelmaos and K. Abed-Meraim, "Principal and minor subspace tracking: Algorithms & stability analysis," in *Proc. IEEE ICASSP'06*, vol. III, Toulouse, France, May 2006, pp. 560–563.

[30] Y. Léost, M. Abdi, R. Richter, and M. Jeschke, "Interference rejection combining in LTE networks," *Bell Labs Tech. J.*, vol. 17, no. 1, pp. 25–49, 2012.

[31] M. H. Brady, M. Mohseni, and J. M. Cioffi, "Spatially-correlated jamming in gaussian multiple access and broadcast channels," in *Proc. CISS*, Princeton, NY, USA, Mar. 2006, pp. 1635–1639.

[32] J. Gao, S. A. Vorobyov, H. Jiang, and H. V. Poor, "Worst-case jamming on MIMO Gaussian channels," *IEEE Trans. Signal Process.*, vol. 63, no. 21, pp. 5821–5836, 2015.

[33] N. Van Huynh, D. N. Nguyen, D. T. Hoang, and E. Dutkiewicz, ""Jam me if you can:" Defeating jammer with deep dueling neural network architecture and ambient backscattering augmented communications," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2603–2620, 2019.

[34] J. Yin and T. Chen, "Direction-of-arrival estimation using a sparse representation of array covariance vectors," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4489–4493, 2011.

[35] Y. Noam and A. J. Goldsmith, "Blind null-space learning for MIMO underlay cognitive radio with primary user interference adaptation," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1722–1734, 2013.

[36] ——, "Blind null-space learning for spatial coexistence in MIMO cognitive radios," in *Proc. IEEE Int. Conf. Commun. (ICC)*. IEEE, Jun. 2012, pp. 1726–1731.

[37] A. Manolakos, Y. Noam, K. Dimou, and A. J. Goldsmith, "Blind null-space tracking for MIMO underlay cognitive radio networks," in *Proc. Global Commun. Conf.*, 2012, pp. 1223–1229.

[38] A. Manolakos, Y. Noam, and A. J. Goldsmith, "Null space learning in cooperative MIMO cellular networks using interference feedback," *IEEE Trans.Wireless Commun.*, vol. 14, no. 7, pp. 3961–3977, Oct 2014.

[39] W. Su, Z. Safar, M. Olfat, and K. R. Liu, "Obtaining full-diversity space-frequency codes from space-time codes via mapping," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2905–2916, 2003.

[40] A. Wittneben, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 1993, pp. 1630–1634.

[41] Z. Wang, T. Schaul, M. Hessel, H. Hasselt, M. Lanctot, and N. Freitas, "Dueling network architectures for deep reinforcement learning," Nov 2015. [Online]. Available: https://arxiv.org/abs/1511.06581

[42] N. Van Huynh, D. T. Hoang, D. N. Nguyen, and E. Dutkiewicz, "Optimal and fast real-time resource slicing with deep dueling neural networks," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 6, pp. 1455–1470, 2019.

[43] A. Hassanien, M. G. Amin, E. Aboutanios, and B. Himed, "Dual-function radar communication systems: A solution to the spectrum congestion problem," *IEEE Signal Process. Mag.*, vol. 36, no. 5, pp. 115–126, Sep 2019.

[44] J. Qian, M. Lops, L. Zheng, X. Wang, and Z. He, "Joint system design for coexistence of MIMO radar and MIMO communication," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3504–3519, Jul. 2018.

[45] A. Hassanien, M. G. Amin, Y. D. Zhang, and F. Ahmad, "Signaling strategies for dual-function radar communications: An overview," *IEEE Aerospace and Electron. Syst. Mag.*, vol. 31, no. 10, pp. 36–45, Oct. 2016.

[46] M. Nowak, M. Wicks, Z. Zhang, and Z. Wu, "Co-designed radar-communication using linear frequency modulation waveform," *IEEE Aerospace and Electron. Syst. Mag.*, vol. 31, no. 10, pp. 28–35, Oct. 2016.

[47] A. Hassanien, M. G. Amin, Y. D. Zhang, and F. Ahmad, "Dual-function radar-communications: Information embedding using sidelobe control and waveform diversity," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2168–2181, Apr. 2015.

[48] F. Liu, L. Zhou, C. Masouros, A. Li, W. Luo, and A. Petropulu, "Toward dual-functional radar-communication systems: Optimal waveform design," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4264–4279, Aug 2018.

[49] S. D. Blunt, P. Yatham, and J. Stiles, "Intrapulse radar-embedded communications," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 3, pp. 1185–1200, Jul 2010.

[50] J. A. Zhang, F. Liu, C. Masouros, R. W. Heath, Z. Feng, L. Zheng, and A. Petropulu, "An overview of signal processing techniques for joint communication and radar sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 69, no. 2, pp. 1295–1315, Nov 2021.

[51] T. Huang, N. Shlezinger, X. Xu, Y. Liu, and Y. C. Eldar, "MAJoRCom: A dual-function radar communication system using index modulation," *IEEE Trans. Signal Process*, vol. 68, pp. 3423–3438, May 2020.

[52] K. Wu, J. A. Zhang, X. Huang, Y. J. Guo, and R. W. Heath, "Waveform design and accurate channel estimation for frequency-hopping MIMO radar-based communications," *IEEE Trans. on Commun.*, vol. 69, no. 2, pp. 1244–1258, Oct 2020.

[53] W. Baxter, E. Aboutanios, and A. Hassanien, "Dual-function MIMO radar-communications via frequency-hopping code selection," in *Proc. Asilomar Conf. Signals, Syst., Comput.* IEEE, 2018, pp. 1126–1130.

[54] J. A. Zhang, M. L. Rahman, K. Wu, X. Huang, Y. J. Guo, S. Chen, and J. Yuan, "Enabling joint communication and radar sensing in mobile networks—a survey," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 1, pp. 306–345, 1st Quart. 2021.

[55] P. Kumari, S. A. Vorobyov, and R. W. Heath, "Adaptive virtual waveform design for millimeter-wave joint communication–radar," *IEEE Trans. Signal Process.*, vol. 68, pp. 715–730, Apr. 2020.

[56] P. Kumari, N. Gonzalez-Prelcic, and R. W. Heath, "Investigating the IEEE 802.11 ad standard for millimeter wave automotive radar," in *Proc. Veh. Technol. Conf. (VTC-Fall).* IEEE, Sep. 2015, pp. 1–5.

[57] P. Kumari, R. W. Heath, and S. A. Vorobyov, "Virtual pulse design for ieee 802.11 ad-based joint communication-radar," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process., (ICASSP)*. IEEE, 2018, pp. 3315–3319.

[58] P. Kumari, D. H. Nguyen, and R. W. Heath, "Performance trade-off in an adaptive IEEE 802.11 ad waveform design for a joint automotive radar and communication system," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process., (ICASSP)*. IEEE, 2017, pp. 4281–4285.

[59] E. Grossi, M. Lops, L. Venturino, and A. Zappone, "Opportunistic radar in IEEE 802.11 ad networks," *IEEE Trans. Signal Process.*, vol. 66, no. 9, pp. 2441–2454, Mar. 2018.

[60] P. Kumari, J. Choi, N. González-Prelcic, and R. W. Heath, "IEEE 802.11 ad-based radar: An approach to joint vehicular communication-radar system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 4, pp. 3012–3027, Apr. 2017.

[61] S. H. Dokhanchi, B. S. Mysore, K. V. Mishra, and B. Ottersten, "A mmwave automotive joint radar-communications system," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 3, pp. 1241–1260, 2019.

[62] G. R. Muns, K. V. Mishra, C. B. Guerra, Y. C. Eldar, and K. R. Chowdhury, "Beam alignment and tracking for autonomous vehicular communication using ieee 802.11 ad-based radar," in *Proc. IEEE INFOCOM Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*. IEEE, Apr. 2019, pp. 535–540.

[63] J. Filar and K. Vrieze, *Competitive Markov decision processes*. New York: Springer-Verlag, 1997.

[64] C. J. Watkins and P. Dayan, "Q-learning," *Mach. Learn.*, vol. 8, no. 3-4, pp. 279–292, 1992.

[65] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. Cambridge, MA, USA: MIT Press, 2018.

[66] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[67] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. Cambridge, MA, USA: MIT press, 2016.

[68] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, 2014, pp. 580–587.

[69] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1440–1448.

[70] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[71] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. 14th Eur.Conf. Comput. Vis. (ECCV)*. Springer, 2016, pp. 21–37.

[72] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge, UK: Cambridge University Press, 2005.

[73] K. Grover, A. Lim, and Q. Yang, "Jamming and anti-jamming techniques in wireless networks: a survey," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 17, no. 4, pp. 197–215, 2014.

[74] L. Jia, Y. Xu, Y. Sun, S. Feng, and A. Anpalagan, "Stackelberg game approaches for anti-jamming defence in wireless networks," *IEEE Wireless Communications*, vol. 25, no. 6, pp. 120–128, 2018.

[75] P. J. Schreier and L. L. Scharf, *Statistical signal processing of complex-valued data: the theory of improper and noncircular signals.* Cambridge, UK: Cambridge Univ. Press, 2010.

[76] G. S. Prabhu and P. M. Shankar, "Simulation of flat fading using matlab for classroom instruction," *IEEE Trans. Educ.*, vol. 45, no. 1, pp. 19–25, 2002.

[77] E. Damosso, L. M. Correia *et al.*, "Digital mobile radio towards future generation systems," Eur. Commission, Brussels, Belgium, Tech. Rep., Apr. 1999.

[78] C. Xiao, Y. R. Zheng, and N. C. Beaulieu, "Novel sum-of-sinusoids simulation models for Rayleigh and Rician fading channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3667–3679, 2006.

[79] G. Strang, *Introduction to linear algebra.* Wellesley, MA, USA: Wellesley-Cambridge Press, 2016.

[80] ——, *Introduction to linear algebra.* Wellesley, MA, USA: Wellesley-Cambridge Press, 2016.

[81] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, 2004.

[82] A. Kekirigoda, K.-P. Hui, Q. Cheng, Z. Lin, A. Zhang, J. Zhang, D. Nguyen, and X. Huang, "Massive MIMO for tactical ad-hoc networks in RF contested environments," in *IEEE/AFCEA Mil. Commun. Conf. (MILCOM)*, 2019.

[83] R. A. Johnson and D. W. Wichern, *Applied multivariate statistical analysis*. New Jersey, US: Pearson Prentice Hall, 2007.

[84] A. S. Lalos, V. Kekatos, and K. Berberidis, "Adaptive conjugate gradient dfes for wideband MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2406–2412, 2009.

[85] J. Choi, H. Yu, and Y. H. Lee, "Adaptive MIMO decision feedback equalization for receivers with time-varying channels," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4295–4303, 2005.

[86] G. H. Golub and C. F. Van Loan, *Computation Matrix*. Baltimore, MD: Johns Hopkins Univ. Press, 2013.

[87] E. S. Sousa, V. M. Jovanovic, and C. Daigneault, "Delay spread measurements for the digital cellular channel in toronto," *IEEE Trans. Veh. Tech-nol.*, vol. 43, no. 4, pp. 837–847, 1994.

[88] M. H. Manshaei, Q. Zhu, T. Alpcan, T. Bacşar, and J.-P. Hubaux, "Game theory meets network security and privacy," *ACM Comput. Surv.*, vol. 45, no. 3, pp. 1–39, Jun. 2013.

[89] Z. Shen, K. Xu, and X. Xia, "Beam-domain anti-jamming transmission for downlink massive mimo systems: A Stackelberg game perspective," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 2727–2742, Feb. 2021.

[90] C. Wang, E. K. Au, R. D. Murch, W. H. Mow, R. S. Cheng, and V. Lau, "On the performance of the MIMO zero-forcing receiver in the presence of channel estimation error," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 805–810, Mar. 2007.

[91] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.

[92] E. TR101290, "Digital video broadcasting (DVB) measurement guidelines for DVB systems," May 2001.

[93] L. Xiao, Y. Ding, J. Huang, S. Liu, Y. Tang, and H. Dai, "UAV anti-jamming video transmissions with QoE guarantee: A reinforcement learning-based approach," *IEEE Trans. Comm.*, vol. 69, no. 9, pp. 5933–5947, Sep. 2021.

[94] P. Siyari, M. Krunz, and D. N. Nguyen, "Friendly jamming in a MIMO wiretap interference network: A nonconvex game approach," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 3, pp. 601–614, Mar. 2017.

[95] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[96] Y. Bengio, P. Frasconi, and P. Simard, "The problem of learning long-term dependencies in recurrent networks," in *Proc. IEEE Int. Conf. Neural Netw.* IEEE, 1993, pp. 1183–1188.

[97] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition," 2014. [Online]. Available: https://arxiv.org/abs/1402.1128

[98] E. TR300-1, "Terrestrial trunked radio (TETRA); voice plus data (V+ D); designers' guide; part I: Overview, technical description and radio aspects," May 1997.

[99] S.-H. Kong, M. Kim, L. M. Hoang, and E. Kim, "Automatic LPI radar waveform recognition using CNN," *IEEE Access*, vol. 6, pp. 4207–4219, 2018.

[100] N. Levanon and E. Mozeson, *Radar signals*. Hoboken, NJ, USA: John Wiley & Sons, 2004.

[101] S. M. Kay, *Fundamentals of statistical signal processing: detection theory, volume II*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1998.

[102] P. E. Pace, *Detecting and Classifying Low Probability of Intercept Radar*. Norwood, MA, USA: Artech House, 2009.

[103] L. Cohen, "Time-frequency distributions-a review," *Proc. IEEE*, vol. 77, no. 7, pp. 941–981, Jul. 1989.

[104] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 6, pp. 862–871, Jun 1989.

[105] L. M. Hoang, M. Kim, and S.-H. Kong, "Automatic recognition of general LPI radar waveform using SSD and supplementary classifier," *IEEE Trans. Signal Process.*, vol. 67, no. 13, pp. 3516–3530, May 2019.

[106] L. Hoang, M. Kim, and S.-H. Kong, "Deep learning approach to LPI radar recognition," in *IEEE Radar Conf. (RadarConf)*. IEEE, 2019, pp. 1–8.

[107] H. Subbaram and K. Abend, "Interference suppression via orthogonal projections: a performance analysis," *IEEE Trans. Antennas Propag.*, vol. 41, no. 9, pp. 1187–1194, Sep. 1993.

[108] M. A. Richards, *Fundamentals of radar signal processing*. New York, NY, USA: McGraw-Hill Education, 2014.

[109] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Hoboken, NJ, USA: Wiley, 2014.

[110] J. K. Blitzstein and J. Hwang, *Introduction to Probability*. Boca Raton, FL: USA: CRC Press, 2015.