TERAHERTZ BAND COMMUNICATIONS:
SYSTEM ARCHITECTURE AND PHYSICAL
LAYER SOLUTIONS

by

Ngwe Thawdar

May 15, 2018

A dissertation submitted to the
Faculty of the Graduate School of
the University at Buffalo, State University of New York
in partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

Department of Electrical Engineering
Acknowledgements

I owe a great debt of gratitude to my PhD advisor, Dr. Dimitris Pados for completion of this work. I would not have been able to do it without his invaluable guidance, encouragement and understanding over the years. As a non-traditional student who is working and raising a young family, I am very grateful for his consistent support throughout the program. I have always regarded my research training with Dr. Pados as a privilege and an honor, which will have a lasting effect on me for the rest of my life.

I would also like to express my deepest thanks to my committee members—Dr. Stella Batalma for her guidance and being my female role model, Dr. Weifeng Su for many summers spent at Air Force Research Laboratory (AFRL) Rome, NY guiding and collaborating with me on cooperative relaying project, and Dr. Nick Mastronarde for his excellent knowledge in video communications which helped me a great deal in data embedding project. This work also would not have been accomplished without the encouragement and collaboration of Dr. Josep Miquel Jornet on plasmonic nano-antenna arrays. I would also like to thank my colleagues, Drs. Ming Li, Ying Liu, Sandipan Kundu, Kanke Gao, Michel Kulhandjian, Panos Markopoulos, George Sklivanitis and Nikolaos Tsagkarakis for their help, friendship and meaningful discussions during my time at University at Bu alo.

This dissertation would not have been completed without the support of my
AFRL supervisor, Mr. Todd Humiston. I am forever grateful for his advocacy and support as I was pursuing the long and winding path towards the PhD degree. I am very fortunate to have him as my supervisor, mentor and friend over the past five years. I would also like to thank my mentors at AFRL—chief scientist Dr. Paul Antonik, chief engineer Dr. Bryant Wysocki and technical advisor, Dr. John Matyjas whose exemplary careers have inspired me and whose guidance has pushed me to reach my potential. My sincere thanks also go to Dr. Michael Medley, my undergraduate advisor and mentor at AFRL for advising me, putting me on the path to PhD program and most of all, never doubting in my ability to accomplish it. I very much appreciate the collaboration of Ulyssess Lee for co-operative relaying experiments and Michael Andrello for COMSOL Multiphysics simulations. Without their help and input, this dissertation would not have been complete.

I would also like to thank my mother who raised me to stand on my own feet and never give up on my dreams. My husband Andrew earned very special thanks as he had to put up with my ups and downs through this journey, yet never stopped believing that I can do it. Finally, this dissertation is dedicated to my son Henry who showed me that pursuits of knowledge and happiness do not have to be separate paths.
Contents

Acknowledgements iii

List of Figures x

List of Tables xi

Abstract xii

1 Introduction 1
  1.1 Benefits and Challenges of Terahertz Band .............. 1
  1.2 Potential Applications for THz Band Communications .... 4
  1.3 Research Objectives and Contributions .................. 5

2 Nano-antenna Array Architecture for THz Band Communica-
  tions 8
  2.1 Graphene-based Plasmonic Devices ..................... 9
  2.2 Plasmonic Nano-antenna Array Architecture ............ 11
    2.2.1 Antenna Design ................................... 12
  2.3 Modulator Design ..................................... 15
  2.4 Beamforming Performance ............................... 16
  2.5 Conclusion and Future Work ............................ 17
### 3 Maximum SINR Phased Arrays for THz Communications

3.1 Background on Array Signal Processing ........................................ 20

3.2 Maximum-SINR Phased Array Problem Formulation .................. 22

3.2.1 System Model .......................................................... 22

3.2.2 Maximum-SINR Phased Array Problem ....................... 24

3.3 Phased-array Optimization via Complex L1-norm PCA .......... 27

3.4 Conclusions and Future Work .................................. 29

### 4 Cooperative Relaying for THz Band

4.1 Background on Cooperative Communications ....................... 31

4.2 System Model .......................................................... 34

4.3 System Model .......................................................... 34

4.4 System Architecture .................................................. 35

4.4.1 Frame Structure .................................................. 36

4.4.2 Transmitter ....................................................... 37

4.4.3 Relay ............................................................... 38

4.4.4 Receiver .......................................................... 39

4.5 Experimental Study ................................................... 41

4.5.1 Indoor Test ....................................................... 42

4.5.2 Indoor Test ....................................................... 42

4.5.3 Outdoor Test .................................................... 45

4.6 Conclusions ......................................................... 46

### 5 Data Embedding in Multimedia for THz Band

5.1 Overview on Data Embedding Technology .......................... 49

5.1.1 Current Methods in Data Embedding in Video Streams ... 51

5.2 Optimal Data Embedding in Video Streams ....................... 53

5.2.1 Signal Model and Notation ..................................... 54

5.2.2 Data Embedding and Carrier Optimization .................. 55
List of Figures

2.1 Schematic of the proposed plasmonic nano-antenna array architecture. ................................................. 11
2.2 Graphene-based nano-antenna resonant at 1.03 THz. ....... 14
2.3 Reflection coefficient of graphene-based plasmonic nano-antenna as a function of frequency. ......................... 15
2.4 Nano-antenna front end with modulator set to (left) $E_F = 0.15 \text{ eV}$ and (right) $E_F = 0.65 \text{ eV}$. ......................... 16
2.5 2D far field of xy-plane for 2x2 array, all modulators (left) $E_F = 0.15 \text{ eV}$ and (right) $E_F = 0.65 \text{ eV}$. ......................... 17
2.6 3D far field of yz-plane for 2x2 array, incremental $E_F$ change by 0.1 eV from 0.15 eV to 0.65 eV (increasing left to right, top to bottom). 18
2.7 (Left) Electromagnetic fields on the nano-antennas at different phase modulation states in 2x2 array, (middle) resulting 2D far field, and (right) resulting 3D far field plots. ......................... 18
4.1 Auto correlation function of a length-63 m-sequence. ........... 37
4.2 Transmitter architecture. .................................................. 38
4.3 Cooperative relay architecture. ................................. 38
4.4 Cooperative receiver architecture. ................................. 39
4.5 Performance of approximated LRT [79] compared to conventional cross-correlation metric using live transmission data from source and relay. Approximated LRT captured the beginning of every frame when conventional cross-correlation metric did not. 40

4.6 Anechoic chamber testbed. 43

4.7 Anechoic chamber testbed. 44

4.8 Aerial relay node with an SDR onboard. 45

4.9 A screenshot from the outdoor test video with real-time BER evaluation. 46

4.10 A snapshot of the BER performance of different transmission schemes. 47

5.1 Uncompressed domain approach 52

5.2 Compressed domain approach 53

5.3 BER as a function of allowable per-message distortion $D_i$ (Foreman CIF, $K = 16$, $L = 63$, $\sigma_n^2 = 3$ dB, embedded data delivery rate $\sim 760$ kbps). 63

5.4 BER as a function of allowable per-message distortion 63

5.5 (a) PSNR as a function of frame index, (b) BER as a function of frame index (Foreman CIF, $K = 16$, $L = 63$, $D_i = 18$ dB, $i = 1, \ldots , K$) 64

5.6 (a) PSNR as a function of frame index, (b) BER as a function of frame index (Bus CIF, $K = 16$, $L = 63$, $D_i = 18$ dB, $i = 1, \ldots , 16$). 64

5.7 BER as a function of PSNR loss (Foreman CIF, $K = 16$, $L = 63$, compressed by H.264/AVC [JM reference software], quantization parameters (QP) 18, 24, and 30). 65

5.8 BER as a function of PSNR loss (Bus CIF, $K = 16$, $L = 63$, compressed by H.264/AVC [JM reference software], quantization parameters (QP) 18, 24, and 30). 65
5.9 (a)-(c) The 1st, 200th, and 300th frames of Foreman video after compression, (d)-(f) same frames after \((s_i^{\text{opt}}, c_i^{\text{opt}})\) embedding and compression, (g)-(i) same frames after sub-optimal Frame-1-based embedding and compression, \((QP=24, K = 16, L = 63, D_i = 22 \text{ dB}, i = 1, \ldots, 16)\). ........................................ 66

5.10 (a)-(c) The 1st, 75th, and 150th frames of Bus video after compression, (d)-(f) same frames after \((s_i^{\text{opt}}, c_i^{\text{opt}})\) embedding and compression, (g)-(i) same frames after sub-optimal Frame-1-based embedding and compression, \((QP=24, K = 16, L = 63, D_i = 22 \text{ dB}, i = 1, \ldots, 16)\). ............................................................ 66


List of Tables

4.1 Indoor test parameters. ........................................... 43
4.2 Outdoor test parameters. ........................................ 45
Abstract

Growing demand for wireless spectrum has motivated research and development in underutilized or unused frequency bands such as the Terahertz frequencies (0.1-10 THz). With tens to hundreds of GHz-wide available bandwidth, the THz band can theoretically enable ultra high-bandwidth applications, such as virtual-reality video. The main challenges in THz-band communications are severe path loss caused by atmospheric absorption and low transmission power output due to small antenna size. In this dissertation, we propose a novel system architecture and algorithms to overcome these challenges and extend the effective transmission range of THz-band communications. First, we present a novel graphene-based plasmonic phased-array architecture where THz signals are directly generated and manipulated to form sharp beams. Next, we describe and propose a phase-only beamformer that maximizes the signal-to-interference ratio (SINR) and present a state-of-the-art polynomial-time approximate optimization algorithm. Then, we shift our focus to the problem of beam blockage and present a novel cooperative relaying system to extend transmission range as well as improve signal reception quality. Finally, for multimedia applications of THz-band communications (e.g., 3-D virtual reality content) we present a data embedding scheme that can augment the data rate further, while providing real-time media multiplexing capabilities and data confidentiality/security.
CHAPTER 1

Introduction

1.1 Benefits and Challenges of Terahertz Band

Overcrowded wireless spectrum has been a prevailing challenge to wireless communications and an impediment to delivering spectrum agility in both commercial and defense sectors. Recently, demand signals for spectrum availability and agility have motivated research and development in traditionally underutilized bands such as mmWave (30-100 GHz) and terahertz frequencies (0.1-10 THz). With strong commercial interest at 24-28 GHz and 37-40 GHz for 5G cellular systems, and at 77 GHz for automotive radar as well as already-assigned ISM band at 60 GHz, mmWave bands are already becoming congested [1].

Currently, neither the U.S. Federal Communications Commission (FCC) nor the United Nations International Telecommunications Union (ITU) have any allocations above 275 GHz. In the U.S., the FCC has no licensed or unlicensed service rules above 95 GHz currently but it is actively seeking comments to open up hundreds of GHz of spectrum above 95 GHz [2]. With large transmission windows that are tens to hundreds of GHz wide, THz band offers a new spectrum fron-
tier that can enable ultrahigh-bandwidth applications for future communication systems.

While the THz band offers a large bandwidth and promises to support tens of gigabit-per-second (Gbps) links, high path loss caused by atmospheric absorption remains a major challenge for communications [3],[4]. The high atmospheric absorption in THz frequencies is beneficial for secure communications but at the same time, it limits the effective transmission range the available bandwidth for long-range distances. For this reason, THz band was deemed unsuitable for applications other than deep space sensing in the past [5]. Moreover, with higher frequencies come higher antenna gains and narrower beamwidths, which in turn makes antenna pointing, acquisition and tracking more challenging. When compared to free-space infrared (IR) based systems, however, THz systems has lower attenuation under certain atmospheric conditions such as fog and rain. THz systems are also less susceptible to scintillation effects compared to IR systems [6].

To overcome path loss and increase transmission range in THz frequencies, massive MIMO techniques utilizing dynamic antenna arrays have been proposed in the literature [7, 3]. At ultra-high frequencies such as THz band, the footprint required for antennas become very small (nm to mm range) enabling massive MIMO schemes with hundreds or thousands of antennas. With a few square centimeters, hundreds of antennas can be used to beamform and track pencil-sharp THz beams. However, novel system architecture and algorithms to dynamically control multiple beams in massive MIMO arrays are still in the development. Even with the right control mechanisms, beam blockage is a challenge at THz frequencies where the wavelengths can be as small as hundreds of micrometers. Novel physical and link layer solutions are still needed to avoid beam blockage and extend transmission range.

Traditionally, one of the main challenges with THz band communication has been the lack of compact high-power signal sources and high-sensitivity detectors.
which limited the feasibility of long range communications. Recently, advances in material and device technology have enabled components such as high power amplifiers, low noise amplifiers, frequency multipliers and mixers that operate in THz frequencies [8],[9].

Today’s device technology in THz band can be categorized into photonics approach and electronics approach. In photonics approach, visible or near IR laser beams are converted to THz signals by employing optoelectronics devices. One common way is to use optical frequency combs which shine sub-THz rays onto unitraveling-carrier photodiode (UTC-PD) and it is usually followed by a high electron-mobility transistor (HEMT) amplifier. However, even with the HEMT amplifier, the maximum output power of UTC-PD method is only about 6 mW for sub-THz range and 10 $\mu$W for 0.4-1 THz range [10, 5]. Another approach is to use quantum cascade lasers (QCLs), however, it is very difficult to realize THz QCLs operating below 1.5 THz and at room temperature [10, 11, 12]. Therefore, photonics approach is not viable for long-range THz communications with current technology but can support short-range applications such as sensor networks.

The electronics approach can be categorized into silicon-based and III-V semiconductor-based device technologies. While SiGe Heterojunction Bipolar-junction Transistor (HBT) technology provides easy integration including on-chip antennas and reduces costs, the low output power has restricted the system transmission range. The state of the art all-electronic SiGe technology that has demonstrated to achieve 50 Gbps at 240 GHz outputs 8 mW of power [13] and at higher THz frequencies such as 860 GHz, 1.9 $\mu$W output power is demonstrated [14]. When integrated SiGe HBT technology has high promise for mid-range applications, current technology is not suitable for long-range applications.

III-V semiconductor based technologies such as InP/GaAS/GaN High Electron Mobility Transistors (HEMT) are more expensive and harder to integrate than Silicon-based components but provides superior output power. In particu-
lar, InP HEMT technology has provided components and integrated transceivers in 190, 235 and 670 GHz bands with tens to hundreds of mW power output and currently leading the field of power amplifiers, making them attractive for long-range applications.

Recently, the use of nanomaterials such as graphene has attracted attention in the research community in developing novel plasmonic devices which intrinsically operate in the THz band. Graphene supports the propagation of Surface Plasmon Polariton (SPP) waves at THz frequencies and by tuning its Fermi energy, we can control the propagation speed of the SPP waves. This property of Graphene has been exploited to design and develop novel plasmonic nano-antennas, modulators and transceivers [15, 16].

Building on the success of the device technologies, past decade has seen demonstration of wireless links at 220, 240, 300 and 670 GHz in US, Asia and Europe [17, 18, 19, 20] at distances ranging from <1m to 850m. To break the barrier of communications at 1 THz or higher, however, significant research and development in the components as well as physical layer solutions and signal processing is still required.

1.2 Potential Applications for THz Band Communications

Wideband applications: Despite the atmospheric loss, THz band offers transmission windows of tens to hundreds GHz that can satisfy the demand for high data rates in commercial and defense applications. For example, 5G and future cellular systems would require wideband data rates for delivery of 4K high definition video and virtual reality (VR) videos while military ISR relays requires delivery of video-rate synthetic aperture radar. Long-range high-bandwidth applications such as multimedia delivery to commercial aircrafts are also possible
utilizing air-to-space and space-to-space THz cross links.

**Secure communications:** Anti-jam (AJ), low probability of intercept (LPI) and low probability of detection (LPD) capabilities are organically built into THz systems through highly directional beams, less scattering of radiation compared to IR systems, high isolation from eavesdroppers provided by atmospheric loss and large channel bandwidth for spread spectrum techniques. Inherent characteristics of the THz band channel can be exploited for extremely robust and secure communications for long range air-to-air links, secure space cross links and secure short range communications among swarms of unmanned vehicles as well as Internet of Things (IoT) applications.

**Personal Area Networks:** THz communication systems can, in fact, make use of high attenuation of the channel for small cell networks and particularly, personal area networks. With bandwidths large enough to support delivery of virtual reality and augmented reality media in short transmission range, THz communication systems can help enhance gaming experience for civilians as well as training experience of law enforcement. Ultrahigh-bandwidth and high-attenuation THz systems can also provide secure personal area networks such as sensor networks between the wearables, sensors and control system for biomedical applications.

### 1.3 Research Objectives and Contributions

The main challenges in THz band communications are to overcome path loss and extend transmission range as well as avoid beam blockage for mid to long-range applications. With hundred of GHz of spectrum to exploit, wideband applications targeted for high data rate multimedia delivery are anticipated in five to ten years. With that in mind, research objectives and contributions of this thesis are outlined
• Antenna arrays consisting of hundreds or thousands of antenna elements have been proposed in the literature for THz band communications, however, novel architecture and dynamic beamforming solutions are still needed to enable robust THz communications. Our research objective is to develop a novel graphene-based plasmonic phased array architecture where THz signals are directly generated and manipulated in phase to form beams in the desired direction. In Chapter 2, we present modeling and performance analysis of, for the first time ever, integrated plasmonic antenna and phase shifter. We give the detailed description of antenna and modulator design and we end the chapter with beam patterns of a 2x2 plasmonic array resonating at >1 THz.

• In practical massive MIMO systems, challenges in terms of complexity, energy consumption and cost have prohibited implementing full digital beamforming system where both amplitude and phase of the signals are adjusted to form robust beams. While phased-only algorithms enable a simpler front-end design, they sacrifice in ability to form precise beams. In Chapter 3, we propose a phased-only beamformer that maximizes the signal-to-interference-ratio (SINR) and present a polynomial-time approximation algorithm to solve maximum-SINR phased array problem by leveraging the recent advances in L1-norm principal component analysis for complex data.

• In Chapter 4, we focus on the problem of beam blockage and present a novel cooperative relaying system to extend transmission range. We design, implement and demonstrate differential amplify-and-forward cooperative relaying scheme as a novel diversity technique to decrease bit error rate as well as promote range extension. We implement our novel physical layer design using software-defined radios and demonstrate it using an air-borne relay
node.

- In Chapter 5, we introduce the concept of steganography as a mean to augment achievable bandwidth in multimedia streams. With wideband multimedia applications in THz band, we present a novel data embedding scheme that can augment the data rate further while providing data confidentiality and security. We present a SINR-optimal spread spectrum embedding algorithm to video host data as well as a suboptimal embedding scheme for low computation complexity. Numerical simulation studies are presented for both raw and H.264 compressed video streams.

Our ultimate goal in this work is to present physical layer solutions to the challenges in THz band communications. In addition, we highlight a wide array of potential applications that THz band has to offer and demonstrate its benefits.

The work in chapter 2 is submitted to present at the 5th ACM/IEEE International Conference on Nanoscale Computing and Communication (September 2018). The research in chapter 3 is being prepared for IEEE Transactions on Signal Processing submission. The work in chapter 4 is presented at 2017 Military Communications conference in Baltimore, Maryland. The work in chapter 5 was presented at 2014 IEEE International Conference on Communications in Sydney, Australia.
Extending the effective transmission range is a major challenge in THz band communications, attributed by low transmission power of the devices and severe path loss of the channel caused by molecular absorption. Dynamic massive MIMO arrays have been proposed to overcome this challenge and increase the transmission range as well as spectrum efficiency. In this chapter, a new antenna array architecture that leverages the properties of graphene-based plasmonic devices is proposed. In this array architecture, each element consists of a plasmonic front-end integrated by a THz plasmonic signal source, a THz plasmonic direct signal modulator, and a THz plasmonic nano-antenna. The possibility to directly modulate the signal without using frequency up-converters or sub-harmonic mixers leads to very compact front-ends, which can be much more densely packed than with traditional THz technologies. After presenting the THz plasmonic nano-antenna and THz plasmonic modulator models, the performance of an integrated front-end is numerically investigated. In addition, the beamforming and beamsteering
capabilities of a 2x2 array are numerically investigated and discussed.

2.1 Graphene-based Plasmonic Devices

Graphene is a one atom thick layer of carbon atoms in a honeycomb lattice, which has attracted a lot of attention in the scientific community due to its very unique electronic properties. Among others, it has been shown that its conductivity at THz-band frequencies drastically changes with the dimensions and chemical potential, which can be utilized to create frequency tunable devices [21, 22]. Furthermore, graphene supports the propagation of surface plasmon polariton (SPP) waves at THz-band frequencies [23, 24]. SPP waves are confined electromagnetic (EM) waves coupled to the surface electric charges at the interface between a metal and a dielectric, and have been shown to have a wavelength orders of magnitude shorter than that of free-space EM waves. By leveraging the unique plasmonic properties of graphene (i.e., tunability and highly confined wavelength), novel graphene-based communication devices can be developed which operate at THz-band frequencies.

Among others, a THz-band nano-transceiver has been proposed which tackles the issue of a lack of compact high-power signal sources [16]. This device, which is based on a high electron mobility transistor built with III-V semiconductors and enhanced with graphene, is able to generate SPP waves at THz frequencies by taking advantage of the Dyakonov-Shur instability principle [25, 26]. A plasmonic phase modulator has also been developed, which exploits the electrical tunability of graphene to modulate a propagating SPP wave without the need of sub-harmonic mixers prior to being radiated [27, 28]. Finally, a graphene-based plasmonic nano-antenna, which can be thought of as a plasmonic waveguide with lossy ends, has been developed to effectively launch SPP waves in free space [15, 29]. Used in conjunction, these devices form a complete THz front-end for
Despite their efficiency, the low transmission power of THz sources and the small size of THz antennas turn increasing the transmission distance into a challenging problem for THz communications. However, recent work has demonstrated that graphene-based plasmonic nanoantennas can be packed into very dense arrays, allowing for higher energy output through beamforming gain [30] and are at the basis of ultra-massive MIMO systems [7]. This work showed that graphene-based plasmonic nano-antennas can be placed as close as their plasmonic wavelength without any significant mutual coupling, as opposed to the free-space wavelength separation required of classical antennas.

Existing analysis have been focused on looking only at individual elements (i.e. transceiver, modulator or antenna) or antenna array. This paper numerically investigates the performance of a fully-plasmonic antenna array architecture, in which each radiating element is modeled as an integrated THz modulator and THz antenna. Full-wave simulations with COMSOL Multi-physics are utilized to design and illustrate the performance of a binary plasmonic phase shifter, a plasmonic nano-antenna, and the plasmonic front-end resulting from their integration. Then, an array of front-ends is simulated to investigated the beamforming capabilities of the resulting system.

The remainder of the chapter is organized as follows. In Section 2.2, we introduce our novel graphene-based plasmonic antenna array architecture. We describe the detailed antenna design and modulator design in Section 2.2.1 and Section 2.3, respectively. In Section 2.4, we investigate the beamforming performance and generated radiation patterns of an array of plasmonic front-ends.
2.2 Plasmonic Nano-antenna Array Architecture

A new antenna array architecture has been developed that leverages the key features of THz plasmonic devices (Figure 2.1). The proposed architecture differs from traditional digital and analog architectures in many ways. In analog architectures, the modulated baseband signal is up-converted to the target center frequency and split into N transmission lines, each with its phase controller, power amplifier and antenna. In digital architectures, the amplitude and phase control is applied to each of the N baseband signals, before being up-converted. This approach provides the maximum control on the transmitted signals, but requires N up-conversion chains. Hybrid architectures, in which some of the phase and amplitude control is done in baseband and others are done after up-conversion, have also been proposed.

Instead, in this novel architecture, the modulation, amplitude and phase control are directly applied to the carrier signal, i.e., there is no up-conversion but direct modulation of the carrier signal. As illustrated in Figure 2.1, the modulation and phase control can be done in sequential blocks, or, ideally, with a
single plasmonic phase modulator [31]. Such devices allow the continuous control of phase, but does not allow the control of the amplitude. Given the lack of power amplifiers at THz frequencies, the only way to control the amplitude of the signal radiated by each antenna is by controlling the THz source itself. The Dyakonov-Shur principle by which the THz plasmonic source is governed, is a non-linear process which can only be switch on or off, i.e., it does not allow continuous support of the amplitude [16].

To dynamically operate the array, the derivation of the array weights that maximize the gain in the intended direction while minimizing the side lobes can be formulated as an optimization problem with architecture specific constraints. To create a highly directional beam pattern, the expected total received beam power will be minimized while ensuring unity gain in the desired beam direction so that the signal of interest is not eliminated. The solution to this optimization problem is certainly non-trivial, and is the subject of next chapter. In this chapter, we focus on the design and performance analysis of the architecture in light of the plasmonic device physics.

2.2.1 Antenna Design

The proposed graphene-based plasmonic nano-antenna resembles a nano-patch antenna, and it is composed of a graphene-layer (the active layer), a metallic ground plane and a dielectric layer in between. As with traditional patch antennas, the plasmonic nano-antenna can be thought of as a plasmonic waveguide with lossy ends. The performance of the plasmonic nano-antenna depends on the propagation properties of the SPP waves on the graphene layer, which in turn depend on the conductivity of the graphene layer. By using the surface conductivity model for infinitely large graphene sheet obtained using the Kubo formalism
[32, 33], the graphene's conductivity can be written as:

\[ \sigma^g = \sigma^g_{\text{intra}} + \sigma^g_{\text{inter}}, \quad (2.1) \]

\[ \sigma^g_{\text{intra}} = \frac{2e^2}{\pi \hbar^2} \frac{k_B T}{\omega + i\tau_g} \ln \left( 2 \cosh \left( \frac{E_F}{2k_B T} \right) \right), \quad (2.2) \]

\[ \sigma^g_{\text{inter}} = \frac{e^2}{4\pi} \frac{H(\omega/2)}{\hbar} + \frac{4\omega}{\pi} \int_0^\infty \frac{H(\epsilon) - H(\omega/2)}{\omega^2 - 4\epsilon^2} \, d\epsilon \quad (2.3) \]

with

\[ G(a) = \frac{\sinh (\hbar a/k_B T)}{\cosh (E_F/k_B T) + \cosh (\hbar a/k_B T)}, \quad (2.4) \]

where \( \omega = 2\pi f \) is the angular frequency, \( h = h/2\pi \) is the reduced Planck's constant, \( e \) is the electron charge, \( k_B \) is the Boltzmann constant, \( T \) is temperature, \( \tau_g \) is the relaxation time of electrons in graphene, and \( E_F \) refers to the Fermi energy of the graphene sheet.

As is the case in the following design, this model is accurate for graphene strips larger than 50 nm in each direction, and within the long wavelength limit (i.e. \( \omega \gg k_{\text{SPP}} v_f \), where \( k_{\text{SPP}} \) is the SPP wave number and \( v_f \sim 8 \times 10^5 \text{m/s} \) is graphene's Fermi velocity) [34]. With this model, the propagation properties of SPP waves on graphene can be determined by solving the following dispersion equation:

\[ -i \frac{\sigma^g}{\omega \varepsilon_0} = \varepsilon_1 + \varepsilon_2 \cot \left( \frac{k_{\text{SPP}} d}{k_{\text{SPP}}} \right), \quad (2.5) \]

where \( \varepsilon_1 \) is the relative permittivity of the dielectric above graphene, \( \varepsilon_2 \) is the relative permittivity of the dielectric between graphene and the ground plane, and \( d \) is the separation distance between graphene and the ground plane.

The reference nano-antenna design in our analysis is shown in Figure 2.2. COMSOL Multi-physics is utilized to simulate the behavior of a graphene-based plasmonic nano-antenna by modeling the graphene layer as a transition boundary.
condition having the conductivity defined above. The design is a 9 µm by 15 µm patch of graphene with $E_F = 1.25$ eV and $\tau = 0.5$ ps (based on analysis of Raman spectra for CVD-grown graphene) [35], on top of a 90 nm-thick $SiO_2$ layer, resting on a metallic ground plane with a 9.9 µm feedline, all at room temperature (300 K). Figure 2.3 illustrates the $S11$ parameter of this nano-antenna as a function of frequency. The result shows that this antenna is resonant at 1.03 THz, which corresponds to the center of the first absorption-defined transmission window above 1 THz, and has a confinement factor $\gamma \sim 20$, in line with what is expected for realistic graphene-based nano-antennas. The reflection coefficient seen in Figure 2.3 is not fully optimized, as this antenna was designed under the assumption that a feed network could be made to match the antennas impedance, but this in fact set by the output impedance of the modulator, described in the next section.
2.3 Modulator Design

As was first introduced in [27], the working principle of the modulator relies on the fact that, by tuning the chemical potential of the graphene layer, or Fermi energy $E_F$, the SPP wave velocity can be modified. By this mechanism, the modulator acts as a delay line to change the phase of SPP waves entering the graphene-based plasmonic nano-antenna. The graphene layer in the modulator is modeled in COMSOL Multiphysics the same way as the antenna, only now with a variable Fermi energy. In our reference design, the size of the modulator is set to 6 $\mu$m long and 9.9 $\mu$m wide to fit with the dimensions of the graphene-based nano-antenna described above. This design allows continuous control of the SPP phase radiated by a single antenna. As an example, Figure 2.4 illustrates a phase change of $\pi$ on the antenna by switching the modulator between Fermi energies of 0.15 and 0.65 eV.

Given the nature of the device, a feeding network cannot be perfectly impedance
matched to the modulator. For the values shown in Figure 2.4 (0.15 and 0.65 eV), the antenna front end has a characteristic impedance of 33.4 and 18.7 \( \Omega \), respectively. Therefore, the feed network should have a characteristic impedance within this range that is sufficient for all values that may be used in operation. For instance, using a feedline impedance equal to 22.6 \( \Omega \) yielded acceptable radiation behavior for both values, with S11 equal to approximately 15 dB and radiated power equal to in the order of 0.1 \( \mu \)W. While these values can be used to create orthogonal symbols, continuous phase control using non-discrete Fermi energies is also possible, and will allow beamforming when used in an array.

## 2.4 Beamforming Performance

In this section, we numerically analyze and discuss preliminary results related to the beamforming abilities of the proposed antenna array architecture. Although graphene-based plasmonic nano-antennas can be operated in an extremely dense array (spacing equal to the plasmonic wavelength on graphene [30]), due to computational complexity, we focus our analysis on a sparse array (with spacing equal to the 1.03 THz free-space wavelength) with 2x2 elements. The results can be extrapolated to an array with the same footprint but more elements to display
similar beamforming patterns with higher radiated power. Figures 2.5 and 2.6 illustrate the directionality of a 2x2 array, as well as how beam steering is enabled by the modulators ability to perform continuous phase control.

The 2x2 array is further illustrated in Figure 2.7 (left), which shows what the EM wave would look like on each antenna for energy states that may be used during continuous phase modulation (arbitrarily chosen as follows: top left $E_F = 0.65$ eV, top right $E_F = 0.35$, bottom left $E_F = 0.25$ eV, bottom right $E_F = 0.15$ eV). Using these carefully selected energy states, one can see from the radiation patterns in Figure 2.7 (middle and right) how the amplitude of side lobes is altered by varying the plasmonic phase modulators Fermi energy. These results demonstrate the possibility to perform beamforming with co-designed plasmonic phase modulators and antennas, without the need for separate phase shifters.

### 2.5 Conclusion and Future Work

In this chapter, we proposed a novel antenna array architecture where modulation is applied directly to the signal source without the need for up-converters or sub-harmonic mixers, and front-ends can be packed much more densely than classical antennas. Antenna and modulator has been designed and simulated with
Figure 2.6: 3D far field of $yz$-plane for $2\times2$ array, incremental $E_F$ change by 0.1 eV from 0.15 eV to 0.65 eV (increasing left to right, top to bottom).

Figure 2.7: (Left) Electromagnetic fields on the nano-antennas at different phase modulation states in $2\times2$ array, (middle) resulting 2D far field, and (right) resulting 3D far field plots.
COMSOL Multiphysics to operate in the THz-band, and allow continuous phase control. An assembly of 2x2 elements is utilized to demonstrate the beamforming and beamsteering capabilities. Future work will focus on incorporating the THz signal source to complete the fully-plasmonic front-end model, in addition to investigating the time delay associated with beamforming and steering.
CHAPTER 3

Maximum SINR Phased Arrays for THz Communications

In the previous chapter, we proposed a massive MIMO antenna array architecture to form directional THz beams. In this chapter, we focus on the robustness of these beams by deriving an adaptive signal processing algorithm that can separate the desired signal from background noise and interference. We exploit the spatial diversity provided by the antenna array and formulate a beamformer where the signal-to-interference ratio (SINR) is maximized at the output of the phased-only beamformer. We leverage the recent advances in L1-norm principal component analysis on complex data and provide a polynomial time algorithm that achieves a fast approximate solution.

3.1 Background on Array Signal Processing

Antenna arrays have a long history of applications in radar [36], [37], sonar [38], communications [39], [40] and space systems [41], [42]. In phased array systems, the phase of signals from each antenna is varied so that the signals constructively
combine to protect the desired signal and destructively combine to suppress the unwanted interference. Traditionally, this beamforming operation is realized by the use of phase-shifters and/or time delay circuits [43] although recent advances in analog-to-digital converter (ADC) technology had given rise to digital beam-forming architectures [44], [45]. From signal processing perspective, there are two approaches to combat interference—array tapering where static weights are employed to lower sidelobe levels [46],[47] and adaptive array processing where beamforming weights are dynamically adapted to steer nulls in the direction of interference [48], [49].

In this article, we focus on adaptive signal processing techniques that enable the beamforming operation to respond to statistical variations of the signal environment [50]. Next-generation antenna array applications such as space-time adaptive processing (STAP) [51], commercial cellular systems [52] and MIMO radar [53] rely heavily on adaptive array processing algorithms.

Recently, there is a renewed interest in large-scale antenna arrays for multi-user massive MIMO systems which is identified as one of the key technologies for 5G wireless systems [54], [55]. In this scenario, a base station with a large number of antennas serves a fixed number of single-antenna terminals and the number of base station antennas is much larger than the number of active users [56]. In a practical massive MIMO system, challenges in terms of complexity, energy consumption and cost arise when implementing a full digital beamforming system with a radio frequency (RF) and ADC chain for each antenna. In order to reduce the number of digital transceivers, hybrid analog and digital beamforming architectures are proposed, in which phased arrays are used as high dimensional RF precoders to be followed by a low dimensional digital baseband processor [57], [58], [59].

Regardless of the application, the objective of every phased array system is to separate the desired signal from background noise and interference by exploiting
the spatial diversity provided by the antenna array. To that end, the antenna array together with its linear beamformer is treated as a spatial filter and signal-to-interference-and-noise-ratio (SINR) at the output of the filter is the parameter of interest to optimize. Many adaptive algorithms for maximum-SINR beamforming problem are developed where both amplitude and phase of the complex weights of the beamformer (or filter taps) are adjusted according to the signal environment [48], [60]. When the degree of freedom is limited to the phase of the beamformer weights only, the problem becomes non-convex and combinatorial [57] and only approximated solutions are available.

In the following sections, we show that solving maximum SINR phased array problem is equivalent to solving a quadratic optimization problem over a non-convex unimodular set. This optimization problem is well known in the literature as unimodular quadratic program (UQP) [61] and approximated solutions are derived using semidefinite relaxation (SDR) techniques [62]. In this chapter, we propose a polynomial-time approximation algorithm to solve maximum-SINR phased array problem by leveraging the recent advances in L1-norm principal component analysis for complex data [63], [64].

3.2 Maximum-SINR Phased Array Problem Formulation

3.2.1 System Model

Consider the case of a narrowband plane wave impinging on a D-element antenna array with an arbitrary geometry. We define the position of an antenna element relative to a reference point in three dimensional space as \( \mathbf{p}_n = [x_n, y_n, z_n]^T \) for \( n = 0, \ldots, D - 1 \). Then the array response vector for a plane wave propagating in the direction \( \mathbf{a} = [\cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta]^T \) incident on the array is
where \( \phi \) is the azimuth angle, \( \theta \) is the elevation angle and \( \lambda \) is the wavelength of the propagating signal.

The signals at \( D \) antenna output are collected in a space-time snapshot vector 
\[
x(t) = [x_0(t), \ldots, x_{D-1}(t)]^T
\]
which can be modeled as:
\[
x(t) = u(t) + d(t) \tag{3.2}
\]
where \( u(t) \in \mathbb{C}^{D \times 1} \) is the signal of interest and \( d(t) \in \mathbb{C}^{D \times 1} \) is the disturbance signal consisting of mutually uncorrelated interference and noise. Without loss of generality, we assume that \( u(t) \) has zero mean and covariance matrix \( R_u = E\{u(t)u^H(t)\} \). The covariance matrix of the disturbance signal can be modeled as \( R_d = E\{d(t)d^H(t)\} \).

In narrowband systems, the space-time snapshots are usually processed by a beamformer with complex weights, \( w \in \mathbb{C}^{D \times 1} \). The combined signal at the beamformer output can be written as:
\[
y(t) = w^H x(t) = w^H [u(t) + d(t)]. \tag{3.3}
\]

Signal to interference ratio (SINR) of \( y(t) \) is a typical metric of to measure the performance of beamformer and is given by
\[
\text{SINR}(w) = \frac{w^H R_u w}{w^H R_d w}. \tag{3.4}
\]

For any signal arriving from direction \( \mathbf{a} \), finding the vector \( w \in \mathbb{C}^{D \times 1} \) that
maximizes the output SINR can be formulated as

$$w^{\text{opt}} = \arg \max_{w \in \mathbb{C}^{D \times 1}} \text{SINR}(w) = \arg \max_{w \in \mathbb{C}^{D \times 1}} \frac{w^{H} R_{a} w}{w^{H} R_{d} w}.$$  \hspace{1cm} (3.5)$$

A known closed form solution for problem in (3.5) exists such that $w^{\text{opt}} = R_{a}^{-1} s_{a}$ and maximum SINR value is $\text{SINR}^{\text{opt}} = P s_{a}^{H} R_{a}^{-1} s_{a}$, where $P$ is the power of signal of interest [60].

We are interested in the case of phase-only beamformers where the hardware implementation limits $w$ such that $w \in \mathbb{U}^{D \times 1}$ where $\mathbb{U} = \{ w \in \mathbb{C} : |w| = 1 \}$. Then maximum SINR phased array optimization problem can be described as

$$\mathcal{P}_{1}: \quad w^{\text{opt}} = \arg \max_{w \in \mathbb{U}^{D \times 1}} \text{SINR}(w) = \arg \max_{w \in \mathbb{U}^{D \times 1}} \frac{w^{H} R_{a} w}{w^{H} R_{d} w}. \hspace{1cm} (3.6)$$

Note that if $d(t)$ consists only of a zero-mean additive white complex noise vector with covariance matrix of $\sigma^{2} I$, then the problem $\mathcal{P}_{1}$ in (3.6) reduces to

$$w^{\text{opt}} = \arg \max_{w \in \mathbb{U}^{D \times 1}} \frac{1}{\sigma^{2}} w^{H} R_{a} w$$

and the optimal beamformer that maximizes the signal to noise ratio (SNR) is a matched filter to the array response vector $s_{a}$ [65]. The maximum SNR value is given by $\frac{P}{\sigma^{2}} D$.

In general, the disturbance term $d(t)$ consists of thermal white noise, and colored noise due to the presence of clutter and other interference sources. Therefore, the covariance matrix $R_{d}$ is in general not diagonal but can be assumed as positive definite because of the presence of receiver noise [65].

### 3.2.2 Maximum SINR Phased Array Problem

Let us rewrite $\mathcal{P}_{1}$ in the epigraph problem form [66] by introducing a new variable $\gamma$:  

24
\( \mathcal{P}_2 : \maximize_{w \in \mathbb{U}^D \times 1, \in \mathbb{R}} \gamma \) 
\begin{align*}
\text{subject to } \quad & w^H R_u w - \gamma \geq 0. 
\end{align*}
(3.7)

Note that \((w^{\text{opt}}, \gamma^{\text{opt}})\) is optimal for \( \mathcal{P}_2 \) if and only if \( w^{\text{opt}} \) is optimal for \( \mathcal{P}_1 \) and \( \gamma^{\text{opt}} = \frac{(w^{\text{opt}})^H R_u w^{\text{opt}}}{(w^{\text{opt}})^H R_d w^{\text{opt}}} \). In other words, optimality for \( \mathcal{P}_2 \) is achieved when the constraint in (3.7) achieves equality. It follows that

\[
\frac{(w^{\text{opt}})^H R_u w^{\text{opt}}}{(w^{\text{opt}})^H R_d w^{\text{opt}}} - \gamma^{\text{opt}} = 0 \\
\iff (w^{\text{opt}})^H (R_u - \gamma^{\text{opt}} R_d) w^{\text{opt}} = 0.
\] (3.8)

Therefore, for a known \( \gamma^{\text{opt}} \), the optimal beamformer that achieves maximum SINR is given by:

\[ w^{\text{opt}}(\gamma^{\text{opt}}) = \arg \max_{w \in \mathbb{U}^D \times 1} w^H (R_u - \gamma^{\text{opt}} R_d) w. \] (3.9)

\( R_u - \gamma^{\text{opt}} R_d \) is a Hermitian matrix and the objective function in (3.9) is quadratic since the covariance matrices \( R_u \) and \( R_d \) are Hermitian. Therefore, the problem in (3.9) is a quadratic optimization problem over the set \( \mathbb{U}^D \) and is shown to be NP-hard in [67]. Consequently, no known closed-form solution exists for (3.9), even for a given \( \gamma^{\text{opt}} \). In general, \( \gamma^{\text{opt}} \) is not known a priori, however, we can pursue a desired SINR, namely \( \gamma_0 \) and approximately solve for a beamformer that achieves the desired SINR as follows:

\[ \mathcal{P}_3 : \quad w_0(\gamma_0) = \arg \max_{w \in \mathbb{U}^D \times 1} w^H (R_u - \gamma_0 R_d) w \] (3.10)

Note that \( \mathcal{P}_3 \) is equivalent to \( \mathcal{P}_2 \) if \( \gamma_0 \) happens to be \( \gamma^{\text{opt}} \). Without loss of generality, \((w_0, \gamma_0)\) pair obtained from solving \( \mathcal{P}_3 \) provides a number of insights into the relationship of the desired SINR \( \gamma_0 \) and maximum SINR \( \gamma^{\text{opt}} \).
1. For a given \((w_0, \gamma_0)\) pair, if equality in (3.8) is achieved and \(w_0^H(R_u - \gamma_0 R_d)w_0 = 0\), then \(\gamma_0 = \gamma_{\text{opt}}\) and \(w_0 = w_{\text{opt}}\).

2. If \(w_0^H(R_u - \gamma_0 R_d)w_0 \geq 0\), then \(\gamma_0 < \gamma_{\text{opt}}\) and we can proceed with more iterations, if we wish, possibly by employing the classical bisection method.

3. If \(w_0^H(R_u - \gamma_0 R_d)w_0 < 0\), then \(\gamma_0 > \gamma_{\text{opt}}\) and there is no \(w\) that satisfies the constraint in \(P_2\) for the chosen \(\gamma_0\). We may solve \(P_3\) again with a lower \(\gamma_0\) value.

Without loss of generality, we can cast \(P_3\) as a quadratic optimization problem over a nonconvex set \(\mathbb{U}^D\) as follows.

\[
\begin{align*}
    w_0(\gamma_0) &= \arg \max_{w \in \mathbb{U}^D} w^H(R_u - \gamma_0 R_d)w \\
    \iff& \arg \max_{w \in \mathbb{U}^D} w^H S w \\
    \iff& \arg \max_{w \in \mathbb{U}^D} w^H X^H X w \quad (3.11)
\end{align*}
\]

where \(S = R_u - \gamma_0 R_d\) when \(R_u - \gamma_0 R_d\) is positive semidefinite and when it is not, we can ensure decomposability of \(S\) by diagonal loading with \(S = R_u - \gamma_0 R_d + \lambda I\) where \(I\) is the identity matrix of size \(D\) and \(\lambda \geq \text{eig}_i(R_u - \gamma_0 R_d)\) for all \(i\).

The optimization problem in (3.11) is well-studied in as Unimodular Quadratic Program (UQP)[61] and algorithms for an approximated solution are usually derived via semidefinite relaxation method. In the next section, we present a polynomial-time approximation algorithm to solve \(P_3\) by leveraging the recent advances in L1-norm principal component analysis for complex data [63], [64].
3.3 Phased-array Optimization via Complex L1-norm PCA

Consider the following problem:

\[
\max_{Q \in \mathbb{C}^{D \times K}, Q^H Q = I_K} \|Q^H X\|_1. \tag{3.12}
\]

The problem in (3.12) is the principal component analysis problem of complex data using L1-norm. Principal component analysis (PCA) in general is a process to find a new orthogonal basis that preserves the maximum energy content of a data matrix. In the literature, L2-norm is commonly used in measuring the distance of a data point from the center of the coordinate system. While L2-norm PCA celebrates computationally efficient algorithms, it has been known to be error prone when the data matrices include outliers. Recently, L1-norm PCA has attracted attention of the research community as an outlier-resistant processing compared to L1-norm PCA. In the case of a data matrix without outliers, the solutions to L1-norm PCA and L2-norm PCA describe an almost identical subspace.

Recent breakthrough in L1-norm PCA on complex data [64] showed that L1-norm PCA can be cast as an optimization problem over the set of unimodular matrices. In particular, consider a single component case of L1-norm maximization problem in the following:

\[
\max_{q \in \mathbb{C}^{D \times 1}, \|q\|_2 = 1} \|X^H q\|_1. \tag{3.13}
\]

The solution to (3.13) is given by

\[
q^{\text{opt}} = \frac{X w^{\text{opt}}}{X w^{\text{opt}}}, \tag{3.14}
\]
where
\[ w^{opt} = \text{sgn}(X^H q^{opt}). \] (3.15)

Now we show that solving the L1-norm PCA problem in (3.13) is equivalent to solving the UQP in (3.9) as follows.

\[
\begin{align*}
\max_{q \in \mathbb{C}^{D \times 1}, \|q\|_2 = 1} \|Xq\|_1 & \iff \max_{q \in \mathbb{C}^{D \times 1}, \|q\|_2 = 1} \max_{w \in \mathbb{U}^{D \times 1}} \Re\{w^H Xq\} \\
& \iff \max_{w \in \mathbb{U}^{D \times 1}} \max_{q \in \mathbb{C}^{D \times 1}, \|q\|_2 = 1} \Re\{q^H Xw\} \\
& \iff \max_{w \in \mathbb{U}^{D \times 1}} \|Xw\|_* \\
& \iff \max_{w \in \mathbb{U}^{D \times 1}} \|Xw\|_2^2 \\
& \iff \max_{w \in \mathbb{U}^{D \times 1}} w^H X^H Xw
\end{align*}
\] (3.16)

(3.16) follows from that \[ \max_{b \in \mathbb{U}^{D \times 1}} \Re\{b^H A\} = \|A\|_1 \] and we obtain (3.17) using Procrustes Theorem [68]. Therefore, UQP in \( P_3 \) can be solved by the polynomial-time algorithms for complex L1-PCA in [64] and we can obtain \( w_0 = \text{sgn}(X^H q^{opt}) \) for a given \( \gamma_0 \). We summarize the steps to solving maximum-SINR phased array problem as follows.

1. For a desired SINR \( \gamma_0 \), compute \( X^H X = R_u - \gamma_0 R_d \).

2. Solve single-component L1-norm PCA in (3.13) and obtain \( w_0(\gamma_0) = \text{sgn}(X^H q^{opt}) \).

3. Compute \( w_0^H (R_u - \gamma_0 R_d)w_0 \).

   (a) If \( w_0^H (R_u - \gamma_0 R_d)w_0 = 0 \), then we achieve optimality, and \( w_0 = w^{opt} \) and \( \gamma_0 = \gamma^{opt} \).

   (b) If \( w_0^H (R_u - \gamma_0 R_d)w_0 \geq 0 \), then \( \gamma_0 < \gamma^{opt} \) and may proceed with another iteration of \( \gamma_1 > \gamma_0 \).
(c) If \( w_0^H (R_a - \gamma_0 R_d) w_0 < 0 \), then \( \gamma_0 > \gamma_{\text{opt}} \). There is no solution since the desired SINR is higher than maximum achievable SINR.

Since we can solve (refl1pca) in polynomial-time approximately according to [64], we can solve the maximum-SINR phased array problem in polynomial time as well.

### 3.4 Conclusions and Future Work

In this chapter, we derived a phased-only beamformer and presented a polynomial time algorithm for maximum SINR phased array problem by leveraging the recent advances in L1-norm principal component analysis on complex data. First, we show that the maximum SINR phased array problem is equivalent to unimodular quadratic optimization problem. Then, we show that solving single-component L1-norm PCA is equivalent to unimodular quadratic optimization problem. Future work will focus on numerical simulations and analysis of computational complexity as compared to semidefinite relaxation-based approaches.
At high frequencies such as THz band, the propagation characteristics of the channel is drastically different from lower radio frequencies, resulting in high attenuation of the signal and easily blocked beams. In particular, mobile and highly dynamic applications such as 5G cellular networks or airborne networks present challenges with signal distortion and bit errors. In this chapter, we employ cooperative relaying as a technique to provide alternate propagation paths and exploit spatial diversity of to enhance the robustness of the signal.

In this chapter, we present differential amplify-and-forward (DAF) cooperative relaying as a novel diversity technique to combat bit errors and outages caused by severe path loss and blockage. By using differential modulation and demodulation, we reduce the system complexity by avoiding channel estimation which is often unreliable or computationally costly in dynamic environments with mobile nodes. We further simplify the requirements of the relay node by employing amplify-and-forward relaying where the relay amplifies the incoming signal and retransmits at a different frequency assignment.

To demonstrate DAF relaying, we designed and implemented a robust physical
layer using the GNU Radio software platform and Universal Software Radio P
eripherals (USRPs) in Wi-Fi band. Our physical layer design can be easily applied
to a THz front-end. We designed a pilot sequence specifically for differential mod-
ulation and implemented frame level synchronization which can combine frames
properly at the destination node for joint decoding even when some frames are
dropped either from the direct link or relay link. We verified simulation results
of DAF relaying in a controlled test in an anechoic chamber and showed that
coopera tive relaying achieves significantly lower bit error rate compared to direct
link transmission and two-hop relaying schemes. We then validated our system
outdoors with an airborne relay node and demonstrated our results.

4.1 Background on Cooperative Communications

We consider the use of cooperative diversity, in which remote nodes are employed
as relaying transceivers to facilitate a form of spatial diversity. The destination
node applies diversity combining techniques [69] to the signals received from relay
nodes and the source node, in contrast to traditional relaying techniques where
only the signal from the last relay node is used to decode the message.

The concept for cooperative relaying can be traced back to [70], where the
achievable information theoretic capacity was determined for general relay chan-
nels. Cooperative relaying protocols are generally categorized by the functions
of the relay node as amplify-and-forward (where the relay simply amplifies the sig-
nal and retransmits), decode-and-forward (where the relay decodes and re-encodes
the signal before retransmitting), and other variants such as compress-and-forward
relaying [71],[72]. The performance benefits of amplify-and-forward and decode-
and-forward relaying schemes in terms of outage probability and transmit diver-
sity bounds were provided in [71]. Detailed symbol-error-rate (SER) analysis and
optimum power allocation schemes for both decode-and-forward and amplify-and-forward protocols were derived in [73].

In this work, we consider a scenario where a pair of nodes with an impaired communication link, whether by blockage or path loss, employs a relay to create cooperative relaying diversity. In this scenario, communication links between the source and relay or the destination and relay have higher link quality than the link between source and destination nodes. To avoid channel estimation which is often unreliable or computationally costly in dynamic environments with highly mobile nodes, we consider non-coherent transceiving where we differentially encode and decode information symbols [74].

We consider a transparent relaying technique, amplify-and-forward relaying to trade-off system simplicity with slight decrease in performance of about 2.4 dB for BPSK and 1.2 dB for QPSK compared to decode-and-forward relaying [73]. Since the relay is not required to decode the message from the source in order to retransmit, it does not require knowledge of detailed parameters for demodulation, decryption, etc. By reducing the requirements at the relaying node, we also protect the message integrity and limits error propagation. The BER performance of differential amplify-and-forward (DAF) relaying was shown in [75] to be superior to either differential detection or coherent detection in transmissions that do not exploit cooperative diversity.

To verify and validate the benefits of DAF relaying, we designed and implemented a Software-Defined Radio (SDR) based system using the GNU Radio software platform and Universal Software Radio Peripherals (USRPs). To decode the received symbols, cooperative relaying requires received signals from the source and relay to be aligned and combined before differentially decoding the symbols. Therefore, we investigated methods to efficiently achieve symbol level synchronization and perform diversity combining in a software-defined radio framework. This is particularly challenging in scenarios involving aerial relays because signal
paths may be highly dynamic, resulting in frequent changes in weighting coefficients, and the latency from the relayed signal paths will vary due to spatial geometry.

We note that a SDR testbed with a medium-access-control (MAC) protocol was developed for cooperative communications with coherent transceiving signals in [76], however this model requires a common timing source and assumes small variation in latency between different signal paths. In [77], cooperative relaying with maximum ratio combining was implemented for direct sequence spread spectrum signals on an FPGA platform. Their architecture relies on a correlator to synchronize the received signals, without considering the case of frame loss when one of the links suffer from severe path loss. In our system, we proposed a robust correlator and a frame synchronization algorithm to properly align the signals even when frames are dropped from ground-to-ground link or air-to-ground link. We verified our implemented system and the algorithms in a controlled anechoic chamber and showed that cooperative relaying achieves significantly lower bit error rate compared to direct link and two-hop relaying. We also developed an airborne relay node with a small Unmanned Aerial System (sUAS) and demonstrated the implemented system in an outdoor testing facility.

The rest of the chapter is organized as follows. In Section 4.3, we briefly describe the differential amplify-and-forward system model that our implementation is based on and in Section 4.4, we describe the design architecture of our implementation in detail. In Section 4.5, we describe both our controlled indoor testbed in an anechoic chamber and outdoor testbed with an airborne relay node and we discuss our findings.


4.2 System Model

4.3 System Model

For our implementation, we follow the DAF system model in [74] and [75] closely where cooperative transmission occurs in two phases either by frequency or time division multiplexing. In Phase 1, information symbol \( v_m = e^{j \phi_m} \), where \( \{ \phi_m \}_{0}^{M-1} \) is a set of M phases, is differentially encoded into \( x^\tau \), the symbol to be transmitted by source \( s \) at time \( \tau \), by the following recursion

\[
x^\tau = v_m x^{\tau-1}.
\]  

(4.1)

The source transmits \( x^\tau \) with power \( P_1 \). The received signals at relay \( r \) and destination \( d \) is

\[
y_{s,r}^\tau = \sqrt{P_1} h_{s,r}^\tau x^\tau + w_{s,r}^\tau
\]  

(4.2)

and

\[
y_{s,d}^\tau = \sqrt{P_1} h_{s,d}^\tau x^\tau + w_{s,d}^\tau,
\]  

(4.3)

respectively, where \( h_{i,j}^\tau \) represents instantaneous channel coefficients and \( w_{i,j}^\tau \) represents additive noise between node pairs \( i,j \). \( h_{i,j} \) and \( w_{i,j} \) are assumed to be independent zero mean complex Gaussian random variables with variance \( \sigma_{i,j}^2 \) for the channel coefficients and \( N_0 \) for the additive noise.

During Phase 2, the relay retransmits the received signal with power \( P_2 \). Then, the signal received at the destination from the relay is effectively modeled by

\[
y_{r,d}^\tau = \alpha h_{r,d}^\tau y_{s,r}^\tau + w_{r,d}^\tau
\]  

(4.4)
where

\[ \alpha = \sqrt{\frac{P_2}{\frac{1}{2} \sigma_{s,r}^2 + N_0}} \]

and \( w_{r,d} \) is again additive Gaussian noise of mean 0 and variance \( N_0 \).

After receiving both signal copies from the source and the relay, the destination combines the signal from (4.3) and (4.4) to differentially decode the information symbols. The combined signal is given by

\[ y^r = a_1 \left( y_{s,d}^{-1} \right)^* y_{s,d} + a_2 \left( y_{r,d}^{-1} \right)^* y_{r,d} \]

where \( a_1 \) and \( a_2 \) are the diversity combining coefficients. In [75] SNR-optimal maximum ratio combining (MRC) was computed with long-term average statistics such as channel variance, instead of instantaneous channel state information,

\[ \hat{a}_1 = \frac{1}{N_0}, \quad \hat{a}_2 = \frac{P_1 \sigma_{s,r}^2 + N_0}{N_0 \left( P_1 \sigma_{s,r}^2 + P_2 \sigma_{r,d}^2 + N_0 \right)}. \]

In our implementation, we performed maximum ratio combining when the channel statistics are available and we performed equal-weight combining otherwise. The destination then performs final symbol detection by

\[ \hat{\nu}_m = \arg \max_{m=0,1,\ldots,M-1} Re \{ \nu_m^* y^r \}. \]

### 4.4 System Architecture

In this section, we describe the system architecture of our software-defined radio implementation based on the popular GNU Radio software and Universal Soft-
ware Radio Peripheral (USRP) platform. We used differential binary shift keying (DBPSK) modulation to transmit symbols over a frequency division multiplexing (FDM) scheme as USRP front-ends and GNU Radio interface enable rapid development of FDM.

The main challenge in implementation of cooperative relaying is the proper synchronization of the received signals from the source and the relay at the destination node in order to decode the symbol using (4.6). If the signals, $y^r_{s,d}$ and $y^r_{r,d}$, are not perfectly aligned in (4.6), detected symbols are perceptible to errors caused by mismatch of source and relay signals. In the following sections, we address these challenges by designing a frame structure with unique pilot signals (called access codes in GNU Radio literature), implementing a robust correlator to synchronize the frames and use of queues. We used available GNU Radio software functions or blocks whenever appropriate and built our own blocks for the custom functionalities in cooperative transceivers.

### 4.4.1 Frame Structure

Since the cooperative receiver is required by (4.6) to combine the signals before they are decoded and detected into bits, the receiver must properly synchronize the signals using undecoded IQ symbols. To facilitate this symbol level synchronization, we frame the payload periodically with an access code that generates pilot IQ symbols known at the receiver. Traditionally, pilot sequences are chosen to have good auto-correlation properties so that the receiver can use a correlator to detect the sequence. For our work, we choose maximum-length sequences or m-sequences which are balanced pseudo-noise sequences generated using linear feedback shift registers. M-sequences have a two-valued periodic autocorrelation function with a 1 at 0-shift and $\frac{1}{L}$ elsewhere, where $L$ is the length of the sequence [78]. When an access code framing the payload is modulated into DBPSK symbols it generates two sign-opposite m-sequences based on the leading bit due
to differential encoding. Since the sequences are sign-opposite, the receiver can use only one correlator to detect both sequences and look for positive and negative peaks. Detailed discussion of the correlator design is given in section 4.4.4.

A 32-bit header is also added to each frame after the access code with three fields that describe the frame index number (or frame id), frame length (payload length in symbols), and an 8-bit cyclic redundancy check (CRC) to ensure the header information is correct. This structure allows for the receiver to synchronize the received signals from the source and relays at frame level with an appropriate queueing before maximum ratio combining and differential decoding. More details on queueing and decoding are given in Section 4.4.4.

### 4.4.2 Transmitter

The transmitter at the source node is designed and implemented by modifying existing GNU Radio software for DBPSK modulation. Each the functional blocks of the transmitter are shown in Figure 4.2 and described as follows:

- **Data Sourcing**: Data bits from higher layers are parsed into appropriately sized payloads. The payload size is selected so that it accounts for the coherence time of transmission environment.

- **Frame Generation**: Access code and appropriate headers as described in
Figure 4.2: Transmitter architecture.

Figure 4.3: Cooperative relay architecture.

Section 4.4.1 are generated and appended at the start of the payload.

- **Differential Symbol Mapping**: After the frames are generated, each frame is differentially encoded into DBPSK symbols.

- **Pulse Shaping**: The IQ symbols are pulse-shaped by using root-raised cosine filter.

- **Transmission**: The pulse shaped symbols are passed to the RF front-end on USRP through the Universal Hardware Driver (UHD) for upconversion to the carrier frequency and transmission.

### 4.4.3 Relay

As (4.4) requires the relay to simply amplify and retransmit, the architecture of the relay consists of three GNU Radio function blocks shown in Figure 4.3. This simplistic structure given by amplify-and-forward relaying allows the ability to recruit available assets for relaying without intrusively modifying their existing architecture. For example, the relay node is not required to have the same decoding hardware as the destination node to facilitate cooperative communications.
### 4.4.4 Receiver

The concept of our receiver architecture is shown in Figure 4.4. First, the received signals from source and relay are down-converted and sampled by two USRPs tuned to the source transmit frequency and relay transmit frequency. Next, each signal is processed through matched-filtering and sampling by the root-raised cosine filter to remove the signal envelope. Then the access code is detected from the IQ symbols using a correlator as described below.

\[
\Lambda = \sum_{i=0}^{L} r_i c_i - \max \left\{ \sum_{i=0}^{L} r_i^I, \sum_{i=0}^{L} r_i^Q \right\},
\]

where \( r_i \) are received complex symbols, and \( c_i \) are access code symbols. (4.9) is an approximation of an optimum likelihood ratio test (LRT) [79] given by

Figure 4.4: Cooperative receiver architecture.

For noncoherent frame detection, we adapted the use of a correlator described in [79] to detect the access code at the beginning of the frame. The noncoherent detection metric of the access code is given by

\[
\Lambda = \frac{1}{\sqrt{2}} \sum_{i=0}^{L} r_i^I + r_i^Q, \frac{1}{\sqrt{2}} \sum_{i=0}^{L} r_i^I - r_i^Q \right\}
\]

(4.9)
Figure 4.5: Performance of approximated LRT [79] compared to conventional cross-correlation metric using live transmission data from source and relay. Approximated LRT captured the beginning of every frame when conventional cross-correlation metric did not.

\[
\Lambda(r) \triangleq \frac{I_0\left(\sum_{i=0}^{N_{SW}-1} \bar{r}_i^* c_i\right)}{\int_0^\pi \prod_{i=0}^{N_{SW}-1} \cosh(\bar{r}_i \cos \varphi + \bar{r}_{i}^Q \sin \varphi) d\varphi} \geq \lambda. \tag{4.10}
\]

The superior performance of approximated LRT compared to conventional cross-correlation metric \( |\sum_{i=0}^{N_{SW}-1} \bar{r}_i^* c_i| \) is shown in Figure 4.5 using live transmission data.

After the access code has been detected, the receiver performs partial decoding of the header information, while leaving the payload symbols untouched. Only the first 32 symbols after the access code are decoded into bits, and these header bits are checked against the corresponding CRC field. If the header bits pass the CRC evaluation, the subsequent payload length symbols (determined from the header information) will be delivered to the next step with an accompanying GNU
Radio tag indicating the frame ID and payload length. If the header bits do not successfully pass the CRC evaluation, these header bits will be ignored. A simple method of doing this is by discarding the first sample of the header symbols, as this will also discard the corresponding GNU Radio ‘access code detected’ tag. The corresponding frame will not be delivered to the next step and will be indicated as 'missing’ during the combining step.

**Frame Queueing, Combining and Decoding**

The combiner has individual queues designated for the storage of frames from each incoming signal path and a combining ID (CID) counter. The CID counter is a non-negative integer that assumes values over the range of our frame IDs and is used to maintain the current ID of frames to combine (For our implementation, frame ID assumes values from 0 to 4095). Upon receiving frames from the previous step, the combiner will store the corresponding symbols, frame ID, and payload lengths into queues. After all the incoming symbols have been stored, the combiner will search each of its queues for frames with the current CID counter. The combiner will use the discovered frames and the weighting coefficients to compute (4.6). If frames are missing, the corresponding $y_{i,j}$ values in (4.6) are replaced by zeros. Symbol detection is performed on the resulting (4.6) to retrieve the original data bits. The data bits are delivered to higher layer applications and the CID counter is incremented by one. This process repeats until the CID counter reaches the maximum frame ID, in which it repeats from the minimum frame ID once again.

### 4.5 Experimental Study

To verify the benefits of DAF relaying and to evaluate our DAF implementation, we carried out a controlled indoor test in an anechoic chamber where we examined
the BER performance of our system against simulation results. To demonstrate operation in aerial relaying environments, we tested our system with an aerial relay node onboard a small unmanned aerial system (sUAS) at AFRL’s Stockbridge test site. in aerial relaying environments, we tested our system with an aerial relay node onboard a small unmanned aerial system (sUAS) at AFRL’s Stockbridge test site.

4.5.1 Indoor Test

4.5.2 Indoor Test

In our indoor setup, the source, relay and destination nodes were laptops equipped with Ettus N210 USRPs with SBX400 RF daughter boards placed in a triangular topology inside of an RF anechoic chamber as shown in Figure 4.6. The USRP nodes are placed without any obstruction between source and relay transmission path and relay and destination transmission path to emulate unobstructed air-to-ground line-of-sight links. Radiation Absorbent Material (RAM) was placed directly between the source and destination USRPs so that the direct source to destination transmission path was degraded by an additional ~15dB. To facilitate frequency division multiplexing, the destination node was outfitted with 2 receiving USRPs tuned to the source and relay transmit frequencies respectively.

In this test, we compared the BER performance of DAF relaying with the BER of direct transmission between source and destination and the BER of two-hop relaying without diversity combining. For a fair comparison, power in each test was preserved by a half-power allocation scheme, which means that in two-hop relaying and DAF relaying the source and relay each transmits at half the power of what the source transmits in direct transmission. The list of parameters used to conduct the experiments is provided in Table 4.1.

In this setup, we assumed that channel characteristics for fixed position nodes
inside the anechoic chamber stay constant for a reasonably long time. Therefore, we measured channel coefficients and calculated channel variances for each links prior to the experiment to be used during the test to calculate the MRC coefficients and amplification weight.

Figure 4.7 shows the BER performance by direct link transmission, two-hop relaying and DAF cooperative relaying schemes. The test results were against the simulation results with Rayleigh channel coefficients and additive white Gaussian noise. We found that our test results followed the simulation results closely and as expected, we found that cooperative relaying achieves lower bit error rate compared to direct line-of-sight transmission and two-hop relaying schemes.
Figure 4.7: Anechoic chamber testbed.
4.5.3 Outdoor Test

In our outdoor setup, source and destination nodes were equipped with Ettus N210 USRPs as in our indoor setup and 150 feet apart. The relay node USRP was replaced with a lightweight Ettus E310 to address the Size, Weight, and Power (SWAP) constraints of an airborne hex-rotor node as shown in Figure 4.8. As GNU Radio is compatible with different Ettus USRPs, the amplify-and-forward function of the relay was implemented onto the onboard CPU of the E310. The list of parameters used to conduct the experiments are provided in Table 4.2. For the duration of the experiment, the aerial node relayed received signals while patrolling the airspace between the source and destination.

![Figure 4.8: Aerial relay node with an SDR onboard.](image)

Table 4.2: Outdoor test parameters.

<table>
<thead>
<tr>
<th>RF Front Ends</th>
<th>Ettus N210 USRP w/SBX400, Relay: Ettus E310</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Frequency</td>
<td>2.7 GHz</td>
</tr>
<tr>
<td>Relay Frequency</td>
<td>2.75 GHz</td>
</tr>
<tr>
<td>Sample Rate</td>
<td>200,000 sample-per-second</td>
</tr>
<tr>
<td>Payload Size</td>
<td>800 symbols</td>
</tr>
<tr>
<td>Transmission Power</td>
<td>15 dBm</td>
</tr>
</tbody>
</table>

In the outdoor test with an airborne node, we modified the receiver so that it can effectively switch between standard receiver mode and cooperative receiver mode.
mode depending on the received frames. When the receiver has frames from both source and relay in the corresponding queues, it combines the two signals and differentially decodes the frame. When the receiver finds a certain frame only in one queue, it operates as a standard differential BPSK receiver and decodes the message. We performed equal-weight combining for this test and found that DAF cooperative relaying has better performance over single path (direct transmission) or two-hop relaying schemes even without MRC.

In particular, the BER of the individual links and the DAF relaying link were averaged with a sliding window of 100 payloads and plotted using LiveGraph during the experiment runtime. Figure 4.9 shows a screenshot from the video of our outdoor test and Figure 4.10 shows an instantaneous snapshot of the BER performance of different transmission schemes. We observed that during the course of our experiment, DAF relaying had about an order of magnitude BER reduction compared to direct and two-hop relaying schemes. Two-hop relaying performance depended heavily on the position of the relay.

4.6 Conclusions

In this chapter, we designed and implemented differential amplify-and-forward cooperative relaying with GNU Radio software and USRP hardware. We specifically
designed the physical layer frames so that we can detect the beginning of each frame using only IQ symbols before the frames are decoded. We designed a cooperative receiver architecture that enables detecting frames from raw IQ symbols and then decodes the header information to queue up the payload symbols from source and relay with their respective frame IDs. The method allows us proper alignment of the payload symbols even when a frame is dropped from the direct transmission path or the relay path. We verified the benefits of DAF cooperative relaying in an anechoic chamber test which showed that cooperative relaying outperforms direct link transmission and two-hop relaying in terms of bit-error rate. We then validated our implementation in an outdoor testbed with an aerial relay node and demonstrates in real-time the benefits of cooperative relaying.
CHAPTER 5

Data Embedding in Multimedia for THz Band

Multimedia data embedding refers to the process of inserting information bearing messages into a digital host medium without -ideally- introducing perceptible distortion to the host or cover medium. Applications of data embedding range from annotation, copyright marking, and watermarking, to single-stream media merging (text, audio, image) and covert or secure communications.

In this chapter, we develop data hiding technology for a variety of signaling environments including THz band communications where wideband multimedia applications such as 4K video, virtual reality and augmented reality videos delivery are expected. In congested environments with bandwidth constraints, additional information can be embedded in video sequences without adversely affecting original video quality or hidden message recovery. An example application of such bandwidth augmentation would be embedding sensor data from a UAV in its downlink surveillance video.

In this chapter, we present our SINR-optimal spread spectrum embedding al-
algorithm to video host data by treating uncompressed video frames as a sequence of images. We propose an optimal data embedding system with minimum video mean-square distortion for any required data recovery error rate in [80]. As a practical consideration, a sub-optimal computationally efficient embedding algorithm with comparable results is also proposed. Extensive experimental studies on data embedding in raw and compressed video sequences are conducted using MATLAB software simulations. Our studies also demonstrate the robustness of the optimal (and sub-optimal) embedding schemes to H.264 compliant encoding.

5.1 Overview on Data Embedding Technology

Data embedding schemes proposed in the literature aim at a wide variety of applications—from perceptually hiding annotation, watermarking for authentication, and ownership protection to covert communications [81, 82]. In annotation, secondary data is embedded into digital media to deliver side information without increasing the file size and subsequent bandwidth for transport [82, 83]. In fragile watermarking, watermarks are used to detect tampering and indicate the authenticity of the data. Robust watermarks are embedded for copyright protection and fingerprinting to trace customers who break license agreements or insiders who leak classified information [81, 82]. In covert communication or steganography, secret private messages are embedded in a cover medium (or host medium) in such a way that the existence of the message is known only to the communication parties [84].

Depending on the specific application, the objectives and tradeo parameters of an embedding algorithm may vary. For watermarking, visual imperceptibility is not the only objective [85], the watermark structure and embedding algorithm need to be designed in such a way that removing the watermark degrades the
fidelity of the cover object [86][87]. In other words, the embedded information should be resilient to malicious or unintentional signal processing such as quantization and geometric processing (cropping, arbitrary row removal, frame dropping) [88], [83], [86, 89]. When embedding secondary information for bandwidth augmentation, the capacity of the embedding system (payload size) is a parameter to optimize [83, 84]. In steganographic application, the embedding algorithm should also provide statistical undetectability for covertness [84]. The one objective that applies to all applications mentioned above is the reliable recovery of the watermark/secret message on the receiver end of communication. While each data embedding application has its own individual requirements, the broad common objective of most applications is a satisfactory tradeoff between the following four basic attributes [90]-payload, transparency (fidelity), detectability (security) and robustness (decoding reliability).

The first step in the design of a data embedding system is to determine the embedding process. This is a crucial task since host-and-data-carrier properties, message detector design and performance depend directly on the way data is inserted in the host signal. Data embedding is performed either directly in the spatial (image) domain [91, 92] or in a transform domain (for example, full-frame discrete Fourier transform (DFT) [93]-[94], full-frame discrete cosine transform (DCT) [95], block DFT or DCT [96, 97], or wavelet transforms [98, 99]). Direct embedding in the original host signal domain may be desirable for system complexity purposes, while embedding in a transform domain may take advantage of the particular transform domain properties.

In our research efforts, we focus our attention on transform domain spread-spectrum (SS) embedding methods for image and video hosts. In a broad sense, any data embedding system for which the secret signal is spread over a wide range of host image frequencies can be referred to as an SS embedding system. Once the transform embedding domain is selected, the hidden message can be applied to
the host data through an additive [93, 94, 96, 100, 101] or multiplicative [102, 103] rule. In the literature, additive SS embedding methods either directly apply (add) the message/watermark to several host coefficients [104] or in direct analogy to SS digital communications systems use an amplitude modulated signature to deposit one information symbol across a group of host data coefficients [93, 94, 96, 105] or a linearly transformed version of the host data coefficients [101].

In direct analogy to SS digital communications systems, SS embedding algorithms have been based on the understanding that the host signal acts as a source of interference to the secret message of interest. Yet, it should also be understood that this interference is known to the message embedding. Such knowledge can be exploited appropriately to facilitate the task of the blind receiver at the other end and minimize the recovery error rate for a given host distortion level, minimize host distortion for a given target recovery error rate, maximize the Shannon capacity of the covert channel, etc. In an information theoretic context, data embedding is viewed as a communications-with-side-information problem [106, 107]. Optimized embedding methods can facilitate host interference suppression at the receiver side when knowledge of the host signal is adequately exploited in system design.

5.1.1 Current Methods in Data Embedding in Video Streams

Current techniques proposed in the literature for data embedding in video can be categorized into two approaches—uncompressed-domain embedding and compressed-domain embedding. In uncompressed-domain approach (Fig. 5.1), data embedding occurs in raw video sequences before any type of compression. In this approach, message embedding/extraction and video compression processes are treated separately. In [108], the secret message is embedded in raw video files by
manipulating the least significant bit (LSB) of pixels in pseudo random locations derived from a secret key. [109, 110] embed the data in transform domain coefficients by quantization index modulation (QIM)-type schemes. [88] and [83] study the embedding capacity problem in data hiding and propose a multi-level embedding algorithm which employs both QIM-type and spread spectrum techniques.

Compressed-domain approaches (Fig. 5.2) combine the embedding and video-encoding procedures such that data is embedded in partially or fully encoded (compressed) video data. The embedding and encoding are treated as a joint process and access to the video codec by the embedding process is assumed. [111] employs bit-plane complexity segmentation steganography on video data compressed by wavelet-based methods such as 3-D set partitioning in hierarchical trees (SPIHT). [86]-[89] embed data in DCT coefficients of MPEG or H.264-encoded video. In [112] and [113], the motion vectors of the MPEG/H.26x inter-coded macroblocks are utilized as the data carriers of the embedding schemes. [85] and [114] propose to embed data in MPEG and H.26x coded bitstreams by manipu-
5.2 Optimal Data Embedding in Video Streams

In this project, we deal with the general problem of transform-domain vector-carrier data embedding. Spatial domain embedding is included as an identity-transform subcase. As with embedding in still images [115],[101], for any required data recovery error rate we aim at minimizing the video mean-squared distortion due to the embedding operation. In particular, for any given video frame sequence and any (block)transform domain of interest, we find the optimal carrier and linear operator parameter that maximize the output signal-to-interference-plus-noise ratio (SINR) of the maximum-SINR data receiver filter. Equivalently, we minimize the average embedding distortion for any target message extraction error rate. The procedure is extended from single-carrier to multi-carrier video embedding.
that allows delivery of multiple messages to multiple intended recipients. Our emphasis is directed primarily toward low complexity, conditional multi-carrier optimization. As a practical consideration, a sub-optimal computationally efficient embedding algorithm is proposed. Experimental results demonstrate that for any given fixed average embedding distortion to the video, our sub-optimal embedding scheme performs closely (in terms of extraction error rate) to optimal embedding. Our studies also demonstrate that the optimal and sub-optimal embedding schemes are robust to H.264 compliant encoding.

The following notation is used throughout the report. Boldface lower-case letters indicate column vectors and boldface upper-case letters indicate matrices; \( \mathbb{R} \) denotes the set of all real numbers; \( ()^T \) denotes matrix transpose; \( \mathbf{I}_L \) is the \( L \times L \) identity matrix; \( \text{sgn}\{\cdot\} \) denotes zero-threshold quantization; \( \mathbb{E}\{\cdot\} \) represents statistical expectation and \( \| \cdot \| \) is vector norm.

### 5.2.1 Signal Model and Notation

Typically, the color space in video frames consists of three components: \( Y \) component to represent brightness or luminance, \( C_b \) and \( C_r \) (blue-difference and red-difference) components to represent color or chrominance. The human visual system is less sensitive to color than to brightness and usually chrominance (\( C_b \) and \( C_r \)) is represented with a lower resolution than luminance (\( Y \)).

We consider a raw video sequence of \( N \) frames in 8-bit \( Y'\text{CbCr} \) form and \( w \times h \) luma component (\( Y' \)) where \( w \) is the width and \( h \) is the height of the frame. Without loss of generality, data embedding is performed only in the luma component of each frame. Let \( \mathbf{H}_n \in \{0, 1, \ldots, 255\}^{w \times h} \) represent the luma component of the \( n \)th frame of the raw video, \( n = 1, \ldots, N \). \( \mathbf{H}_n \) is partitioned into \( M \) local non-overlapping small blocks of size \( \frac{w \times h}{M} \). Each block, \( \mathbf{H}_{n,1}, \mathbf{H}_{n,2}, \ldots, \mathbf{H}_{n,M} \), is to carry one hidden information bit (\( M \) bits payload per frame). Embedding is performed
in a 2D transform domain $\mathcal{T}$ (such as DCT, wavelet transform, etc.). After transform calculation and vectorization (for example by conventional zig-zag scanning), we obtain $\mathcal{T}(\mathbf{H}_{n,m}) \in \mathbb{R}^{w \times h}$, $n = 1, \ldots, N$, $m = 1, \ldots, M$. From the transform domain vectors $\mathcal{T}(\mathbf{H}_{n,m})$ we choose a fixed subset of $L \leq \frac{w \times h}{M}$ coefficients (bins) to form the final host vectors $\mathbf{x}_{n,m} \in \mathbb{R}^{L}$, $n = 1, \ldots, N$, $m = 1, 2, \ldots, M$. It is common and appropriate to avoid the dc coefficient (if applicable) due to high perceptual sensitivity in changes of the dc value.

The autocorrelation matrix of the transform domain host data of each frame is defined as

$$ R_{x,n} \triangleq \frac{1}{M} \sum_{m=1}^{M} \mathbf{x}_{n,m} \mathbf{x}_{n,m}^{T}, \quad n = 1, \ldots, N. $$

The transform domain host data autocorrelation matrix over all frames is defined as

$$ R_{x} \triangleq \frac{1}{MN} \sum_{n=1}^{N} \sum_{m=1}^{M} \mathbf{x}_{n,m} \mathbf{x}_{n,m}^{T} = \frac{1}{N} \sum_{n=1}^{N} R_{x,n}. $$

It is easy to verify that in general $R_{x,n} \neq \alpha \mathbf{I}_L, \alpha > 0$; that is, $R_{x,n}$ (and $R_{x}$) is not constant-value diagonal or “white” in field language.

### 5.2.2 Data Embedding and Carrier Optimization

Conventional direct additive carrier embedding has the form

$$ y_{n,m} = A b_{n,m} \mathbf{s} + \mathbf{x}_{n,m} + \mathbf{n}_{n,m} $$

where $b_{n,m} \in \{\pm 1\}$ is the message bit embedded in the host vector $\mathbf{x}_{n,m}$, $n = 1, \ldots, N$, $m = 1, \ldots, M$, $A > 0$ is the bit amplitude, $\mathbf{s} \in \mathbb{R}^{L}$, $\|\mathbf{s}\| = 1$, is the (normalized) embedding carrier (signature), and $\mathbf{n}_{n,m}$ represents potential external white Gaussian noise\(^1\) of mean 0 and autocorrelation matrix $\sigma_n^2 \mathbf{I}_L$, $\sigma_n^2 > 0$.

\(^1\)Additive white Gaussian noise is frequently viewed as a suitable (maximum entropy) model for general quantization errors, channel transmission disturbances, and/or image processing attacks.
In an effort to reduce the interference effect of the host signal to the carrier $s \in \mathbb{R}^L$, the host vectors $x_{n,m}$, $n = 1, \ldots, N$, $m = 1, \ldots, M$, can be steered away from the embedding carrier using a linear operator of the form $(I_L - css^T)$ with parameter $c \in \mathbb{R}$. In parallel to (5.3), the composite signal after additive embedding in linearly modified host data is

$$y_{n,m} = A b_{n,m} s + (I_L - css^T) x_{n,m} + n_{n,m} \quad (5.4)$$

where the information symbol bit $b_{n,m} \in \{ \pm 1 \}$ is embedded with amplitude $A > 0$ and (normalized) carrier $s \in \mathbb{R}^L$, $\|s\| = 1$, in the linearly modified host data vector $(I_L - css^T) x_{n,m}$, $n = 1, \ldots, N$, $m = 1, \ldots, M$.

For embedding in linearly modified hosts by (5.4), the mean-square-error (MSE) distortion per-block\(^2\) to the $n$th frame due to the embedding operation only is

$$D(n) = \mathbb{E} \left\{ \left( A b_{n,m} s + (I_L - css^T) x_{n,m} \right)^2 \right\} = A^2 + c^2 s^T R_x s$$

and the total-squared-error (TSE) distortion to the $n$th frame is $M \times D(n)$. Consequently, the MSE distortion per-block over the whole video sequence is

$$\mathcal{D} = \frac{1}{N} \sum_{n=1}^{N} D(n) = A^2 + c^2 s^T R_x s$$

and the TSE distortion to the whole video sequence is $N M \mathcal{D}$. We observe that the distortion level is controlled not only by $A > 0$ but by $c \in \mathbb{R}$ and $s \in \mathbb{R}^L$, $\|s\| = 1$, as well; $R_x (R_{x,n})$ is intrinsic to the video sequence (frame).

The intended recipient of the message will perform embedded bit detection by

\(^2\)Herein, MSE distortion refers to average distortion per block. Pixel-wise MSE distortion is $D(n)/P$ where $P$ is the number of pixels in each block.
looking at the sign of the output of a linear filter \( w \in \mathbb{R}^L \)

\[
\hat{b}_{n,m} = \text{sgn} \left\{ w^T y_{n,m} \right\}.
\]  \( (5.7) \)

With signal of interest \( A_b, m \) and total disturbance \( (I_L - c s s^T)x_{n,m} + n_{n,m} \) in (5.4), the output SINR of filter \( w \) is

\[
\text{SINR} = \frac{\mathbb{E}\{\|w^T(A_b, m s)\|^2\}}{\mathbb{E}\{\|w^T((I_L - c s s^T)x_{n,m} + n_{n,m})\|^2\}}.
\]  \( (5.8) \)

The probability of error of the detector in (5.7) is, in general, a monotonically decreasing function of SINR. The linear filter \( w \) that offers maximum SINR at its output is

\[
w_{\text{max SINR}} = \left( (I_L - c s s^T)R_x(I_L - c s s^T)^T + \sigma_n^2 I_L \right)^{-1} s.
\]  \( (5.9) \)

The maximum output SINR value attained by \( w_{\text{max SINR}} \) is

\[
\text{SINR}_{\text{max}}(s, c) = A^2 s^T \left( (I_L - c s s^T)R_x(I_L - c s s^T)^T + \sigma_n^2 I_L \right)^{-1} s.
\]  \( (5.10) \)

At this point, we view \( \text{SINR}_{\text{max}}(s, c) \) as a function of the embedding carrier \( s \) and the linear-operator parameter \( c \). If \( q_1, \ldots, q_L \) denote the eigenvectors of \( R_x \) with corresponding eigenvalues \( \lambda_1 \geq \ldots \geq \lambda_L \), then, for any given target per-block MSE distortion value \( D \) over the whole video sequence, the pair (carrier \( s \), parameter \( c \)) that maximizes the output SINR of the maximum SINR filter is \( s_{\text{opt}} = q_L \) and \( c_{\text{opt}} = \frac{\lambda_L + \sigma_n^2 + D - \sqrt{(\lambda_L + \sigma_n^2 + D)^2 - 4\lambda_L D}}{2\lambda_L} \). Maximum SINR data filtering simplifies to plain matched filtering, \( w_{\text{max SINR}} \equiv s_{\text{opt}} = q_L, \hat{b}_{n,m} = \text{sgn} \left\{ q_L^T y_{n,m} \right\} \) and the output SINR is maximized to \( \text{SINR}_{\text{max}}(s_{\text{opt}}, c_{\text{opt}}) = \frac{D - c_{\text{opt}}^2 \lambda_L}{\lambda_L (1 - c_{\text{opt}})^2 + \sigma_n^2} \).

The target distortion \( D \) is achieved when the embedding amplitude is set at

\[
A = \sqrt{\frac{D - c_{\text{opt}}^2 \lambda_L}{\lambda_L (1 - c_{\text{opt}})^2 + \sigma_n^2}}.
\]
5.2.3 Multi-carrier Data Embedding.

Multi-carrier embedding can deliver multiple messages to multiple intended recipients. We can generalize the signal model in (5.4) to cover multi-carrier/multi-message embedding of the form

\[
y_{n,m} = \sum_{i=1}^{K} A_i b_{i,n,m} s_i + (I_L - \sum_{i=1}^{K} i_s i_s^T) x_{n,m} + n_{n,m} \tag{5.11}
\]

where bits \( \{b_{1,n,m}, b_{2,n,m}, \ldots, b_{K,n,m}\} \), belonging potentially to \( K \) distinct messages, are embedded simultaneously in the linearly modified host \((I_L - \sum_{i=1}^{K} c_i s_i s_i^T)x_{n,m}\) of frame \( n = 1, \ldots, N \), block \( m = 1, \ldots, M \), with corresponding amplitudes \( A_i > 0 \) and embedding carriers \( s_i \in \mathbb{R}^L, \|s_i\| = 1 \), \( i = 1, 2, \ldots, K \).

Under statistical independence across messages, the MSE distortion per block to the \( n \)th frame is

\[
D(n) = \sum_{i=1}^{K} A_i^2 + c_i^2 s_i^T R_x s_i
\]

and the MSE distortion per block over the whole video sequence induced by each individual message is

\[
D_i = A_i^2 + c_i^2 s_i^T R_x s_i, \quad i = 1, 2, \ldots, K.
\]

The TSE distortion to the whole video sequence from all messages is \( NM \sum_{i=1}^{K} D_i \).

The intended receiver of the \( j \)th-message bits will use a linear filter \( w_j \) to recover the corresponding embedded bits

\[
\hat{b}_{j,n,m} = \text{sgn} \left\{ w_j^T y_{n,m} \right\}
\]

\[
= \text{sgn} \left\{ \left( A_j b_{j,n,m} s_j + \sum_{i=1, i \neq j}^{K} A_i b_{i,n,m} s_i + (I_L - \sum_{i=1}^{K} c_i s_i s_i^T)x_{n,m} + n_{n,m} \right)^T \left( I_L - \sum_{i=1}^{K} c_i s_i s_i^T \right)^{-1} \right\} \tag{5.13}
\]
With signal of interest $A_j b_{j,n,m} s_j$ and total disturbance $\sum_{i=1, i \neq j}^K A_i b_{i,n,m} s_i + (I_L - \sum_{i=1}^K c_i s_i s_i^T)x_{n,m} + n_{n,m}$, the linear filter that operates on $y_{n,m}$ and offers maximum SINR at its output to the $j$th-message recipient is

$$w_{\text{maxSINR},j} = R_{j}^{-1} s_j.$$  \hspace{1cm} (5.14)

In (5.14), $R_{j}$ denotes the “exclude-$j$” data autocorrelation matrix, that is the autocorrelation matrix of the disturbance to message-$j$ defined as

$$R_{j} \triangleq \sum_{i=1, i \neq j}^K A_i^2 s_i s_i^T + (I_L - \sum_{i=1}^K c_i s_i s_i^T)R_c(I_L - \sum_{i=1}^K c_i s_i s_i^T) + \sigma_n^2 I_L.$$  \hspace{1cm} (5.15)

The output SINR value attained by $w_{\text{maxSINR},j}$ is

$$\text{SINR}_{\text{max},j} = A_j^2 s_j^T R_{j}^{-1} s_j.$$  \hspace{1cm} (5.16)

Similar to single-carrier embedding, if we view $\text{SINR}_{\text{max}}(s_j, c_j)$ as a function of the embedding carrier $s_j$ and the transformation parameter $c_j$, then we can identify $s_j$ and $c_j$ that maximize the SINR value. Unfortunately, $R_{j}$ remains a function of $s_j$ as well as $c_j$ and the multiuser sum-SINR globally optimal solution is intractable.

A per-user conditionally optimal solution, however, is attainable. For $i = 1, 2, \ldots, K \leq L$, we can design sequentially the carriers $s_i$ and the parameters $c_i$ such that the $i$th-user output SINR is conditionally maximized given all past $i-1$ fixed embeddings. In particular, if $q_1, q_2, \ldots, q_L$ denote the eigenvectors of $R_c$ with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_L$, then for any given $i$th-message-induced distortion level $D_i$, the optimal carriers and transformation parameters $(s_{i}^{\text{opt}}, c_{i}^{\text{opt}}), i = 1, \ldots, K$, that conditionally maximize the output SINR

$^3$In contrast, if no host data linear transformation is attempted ($c_i = 0$ in (5.11), $i = 1, \ldots, K$) and the carriers $s_1, s_2, \ldots, s_K, K \leq L$, are to be orthonormal, then the eigen-carrier design in (5.17) is multiuser sum-SINR optimal per [116].
of the maximum SINR filters given the previous $i - 1$ embeddings are

$$s_i^{opt} = q_{L-i+1},$$

$$c_i^{opt} = \lambda_{L-i+1} + \sigma_n^2D_i - \sqrt{(\lambda_{L-i+1} + \sigma_n^2 + D_i)^2 - 4\lambda_{L-i+1}D_i}, i = 1, \ldots, K.$$  

Therefore, maximum SINR data filtering per user again simplifies to plain matched filtering, $q_{L-i+1}^T y_{n,m}$. The target per message distortion $D_i$ is achieved when the embedding amplitude is set to $A_i = \sqrt{D_i - c_i^{opt}^2 \lambda_i}, i = 1, \ldots, K$. When $s_i = q_{L-i+1}$ and $c_i = c_i^{opt}$, the output SINR is individually (conditionally) maximized to

$$\text{SINR}_{\text{max},i} = \frac{D_i - c_i^{opt}^2 \lambda_{L-i+1}}{\lambda_{L-i+1}(1 + c_i^{opt})^2 + \sigma_n^2}, i = 1, \ldots, K.$$  

### 5.2.4 Sub-optimal Carrier Design

The optimal carrier and transformation parameter designs in Sections 5.2.2 and 5.2.3 aim at global video sequence optimization which is based on second-order block statistics $R_x = \sum_{n=1}^{N} R_{x,n}$ over all $N$ frames. We understand that real-time applications may not be able to afford either the storage (buffer) needs to temporarily store all $N$ frames before embedding or the delay associated with the collection of all $N$ frames.

The time dimension of the video medium makes imperceptibility of data embedding more challenging compared to still images. Maintaining the distortion or peak signal-to-noise-ratio (PSNR) at a consistent level across all frames of a video stream to avoid flickering is a concern when video data embedding is considered. To that respect, we recall that the MSE distortion per block to the $n$th frame is $D(n) = \sum_{i=1}^{K} A_i^2 + c_i^2 s_i^T R_{x,n} s_i, n = 1, \ldots, N$. We understand that when we
minimize the average distortion over all $N$ frames

$$D = \frac{1}{N} \sum_{n=1}^{N} D(n) = \sum_{i=1}^{K} A_i^2 + c_i^2 s_i^T R_x s_i,$$

the distortion pertinent to each frame may vary.

To address the above technical concern, we consider a sub-optimal embedding approach that exploits the strong similarity of adjoint frames. For such frames, the transform domain hosts have similar statistical properties. Therefore, instead of using $R_x$, we calculate $(s_i, c_i)$, $i = 1, \ldots, K$, based on $R_{x,1}$, the autocorrelation of the host vectors of the first frame, and embed data in all frames using $(s_i, c_i)$, $i = 1, \ldots, K$. Thus, real-time operation can be supported since data embedding starts right after collecting the first frame and data-encoded video can be delivered immediately. To provide consistent distortion across frames, the embedding amplitudes $A_{i,n}$ for the $i$th message in the $n$th frame is scaled/adjusted by $A_{i,n} = \sqrt{1 - c_i^2 s_i^T R_{x,n} s_i}$, $i = 1, \ldots, K$, $n = 1, \ldots, N$, which assures equal distortion per-block to each frame. We note that, if frames are similar in second-order block statistics, then $R_{x,1} \approx R_{x,2} \approx \ldots \approx R_{x,N} \approx R_x$ and sub-optimal embedding approaches global video optimal embedding.

### 5.3 Experimental Studies for Optimal Data Embedding

In this section, we evaluate the performance of the proposed data embedding schemes using two CIF uncompressed video sequence examples as hosts, the Foreman and the Bus video sequence. The luma frames consists of $352 \times 288$ pixels. Data is embedded in $8 \times 8$ DCT luma frame blocks with $K = 16$ carriers. Embedding is over all bins except dc. Hence, our carrier length is $L = 63$ and there is a
total of $1584 \times 16 = 25,344$ bits hidden in each raw frame. With frame rate $\sim 30$ frames per second, hidden-data delivery rate is about 760 Kbps. For the sake of generality, we also incorporate in our studies white Gaussian noise of variance $\sigma_n^2 = 3$ dB.

Figs. 5.3 and 5.4 show the average recovery BER for the Foreman and Bus video sequence, respectively, as a function of the distortion created by the embedded message over the 8 dB to 18 dB range and under carrier matched-filter detection. Three different embedding schemes are examined: i) Arbitrary carrier embedding, which serves as performance baseline; ii) optimal $(s_i, c_i)$ embedding as presented in Section 5.2.3; and iii) sub-optimal frame-1-based embedding as introduced in Section 5.2.4. We see that sub-optimal embedding can provide extraction error rate that is close to optimal embedding. Comparing Figs. 5.3 and 5.4, the observed suboptimal-embedding performance degradation of the Foreman video sequence with respect to the Bus video sequence is due to the higher variation of the second-order block statistics $R_{x,n}, n = 1, 2, \ldots, N$, across frames in the Foreman sequence. On the other hand, the scene similarity observed across frames in the Bus sequence makes the performance curve of the frame-1-based scheme nearly indistinguishable from the global-frame-optimal $(s_i, c_i)$ performance curve.

Figs. 5.5 and 5.6 show BER and PSNR (distortion) per frame for the Foreman and Bus video sequence, respectively. Frame-1-based embedding induces BER and PSNR per frame that are comparable to the BER and PSNR, respectively, of the optimal $(s_i, c_i)$ embedding scheme for all Bus frames and most Foreman frames (BER and PSNR per frame worsen when there is a significant change of scene in the Foreman sequence).

Figs. 5.7 and 5.8 evaluate the robustness of the proposed embedding schemes to video compression. In our studies, data are embedded in the raw video sequence with $K = 16$, $L = 63$, and a certain distortion budget $D_i, i = 1, 2, \ldots, 16$. The
Figure 5.3: BER as a function of allowable per-message distortion $D_t$ (Foreman CIF, $K = 16$, $L = 63$, $\sigma_n^2 = 3$ dB, embedded data delivery rate $\sim 760$ kbps).

Figure 5.4: BER as a function of allowable per-message distortion
Figure 5.5: (a) PSNR as a function of frame index, (b) BER as a function of frame index (Foreman CIF, $K = 16, L = 63, D_t = 18$ dB, $i = 1, \ldots, K$)

Figure 5.6: (a) PSNR as a function of frame index, (b) BER as a function of frame index (Bus CIF, $K = 16, L = 63, D_t = 18$ dB, $i = 1, \ldots, 16$).
Figure 5.7: BER as a function of PSNR loss (Foreman CIF, $K = 16$, $L = 63$, compressed by H.264/AVC [JM reference software], quantization parameters (QP) 18, 24, and 30).

Figure 5.8: BER as a function of PSNR loss (Bus CIF, $K = 16$, $L = 63$, compressed by H.264/AVC [JM reference software], quantization parameters (QP) 18, 24, and 30).
Figure 5.9: (a)-(c) The 1st, 200th, and 300th frames of Foreman video after compression, (d)-(f) same frames after $(s_{i}^{opt}, c_{i}^{opt})$ embedding and compression, (g)-(i) same frames after sub-optimal Frame-1-based embedding and compression, (QP=24, $K = 16$, $L = 63$, $D_i = 22$ dB, $i = 1,\ldots, 16$).

Figure 5.10: (a)-(c) The 1st, 75th, and 150th frames of Bus video after compression, (d)-(f) same frames after $(s_{i}^{opt}, c_{i}^{opt})$ embedding and compression, (g)-(i) same frames after sub-optimal Frame-1-based embedding and compression, (QP=24, $K = 16$, $L = 63$, $D_i = 22$ dB, $i = 1,\ldots, 16$).
data-embedded video is then compressed by H.264/AVC-compliant JM reference software [117]. After decompressing the encoded video, the embedded data are retrieved and BER is evaluated. Since both compression and data embedding induce distortion to video, we use PSNR-loss as a data embedding quality metric that captures distortion caused only by embedding:

\[
PSNR_{loss} = PSNR_{clean} - PSNR_{embed}
\]  

(5.18)

where \(PSNR_{clean}\) is the PSNR of compressed clean video and \(PSNR_{embed}\) is the PSNR of compressed data-embedded video. In Figs. 5.7 and 5.8, we plot the average BER as function of \(PSNR_{loss}\) of the \((s_{i}^{opt}, c_{i}^{opt})\) optimal and frame-1-based embedding schemes for the Foreman and Bus video sequence, respectively, using three different quantization parameter (QP) values in the H.264/AVC compression standard (QP = 18, 24, and 30).

In Figs. 5.9 and 5.10, we set QP = 24 during H.264/AVC compression and show three different frames of the Foreman and Bus video sequence, respectively, for the following scenario: (i) no data embedding (a)-(c); (ii) optimal \((s_{i}^{opt}, c_{i}^{opt})\) embedding (d)-(f); and (iii) frame-1-based embedding (g)-(i). We observe no perceptual distortion/difference between the compressed clean frames and the compressed data-embedded frames. The embedded-data payload rate is still 760 Kbps for both sequences, which at the H.264/AVC Q=24 setting corresponds - rather impressively- to about 20% of the whole Foreman video rate and 15% of the Bus video rate.

### 5.4 Conclusions

In this chapter, we considered the problem of single-carrier and multi-carrier (multiple sessions) data embedding in linearly modified (transform) domains of raw
video sequences with minimum video mean-square distortion for any required data recovery error rate. In particular, for any given video frame sequence and any (block) transform domain of interest, we find the optimal carrier and scalar parameterized linear operator on the video data that maximize the output signal-to-interference-plus-noise ratio (SINR) of the maximum-SINR data receiver filter or, equivalently, minimize the average embedding distortion for any target message extraction error rate. As a practical consideration, a sub-optimal computationally efficient single-frame-only designed embedding was also considered. Experimental results demonstrated high embedded data delivery rates over H.264 compliant video streams with imperceptible video distortion.
Conclusions

Terahertz band is poised to be the next spectrum frontier for secure wideband communications in military and commercial applications. With a large unoccupied bandwidth, THz band channel comes with its unique peculiarities and challenges. The main challenges include severe path loss caused by atmospheric absorption that limits the effective transmission range as well as beam blockage due to very short wavelengths of transmission. This dissertation proposes two approaches for range extension—massive MIMO phased arrays and cooperative relaying. We also proposed a novel data embedding algorithm to augment the bandwidth in delivery of multimedia streams that are expected to be a leading application for THz band. We summarize our findings and contributions below.

- We proposed a novel antenna array architecture where modulation is applied directly to the signal source without the need for up-converters or sub-harmonic mixers, and front-ends can be packed much more densely than classical antennas. Antenna and modulator has been designed and simulated with COMSOL Multiphysics to operate in the THz-band, and allow continuous phase control. An assembly of 2x2 elements is utilized to
demonstrate the beamforming and beamsteering capabilities.

- We then focused on the robustness of the phase-only beamformer that maximizes the SINR of the signal presented a polynomial time algorithm for maximum SINR phased array problem by leveraging the recent advances in L1-norm principal component analysis on complex data.

- To prevent beam blockage and improve robustness of the THz signal, we proposed a novel diversity scheme known as cooperative relaying. We designed and implemented differential amplify-and-forward (DAF) cooperative relaying with GNU Radio software and USRP hardware. We verified the benefits of DAF cooperative relaying in an anechoic chamber as well as in an outdoor testbed with an aerial relay node and demonstrates in real-time the benefits of cooperative relaying.

- Finally, we developed a data embedding scheme for a wideband signaling environment such as THz band with multimedia applications. We find the SINR optimal embedding algorithms for single and multi-carrier embedding schemes both for raw and H.264 compressed video data. We also proposed a sub-optimal embedding algorithm for computational efficiency trade-off.
Bibliography


