

1-1-2003

# Time-frequency analysis of spread spectrum based communication and audio watermarking systems

Serhat Erküçük  
*Ryerson University*

Follow this and additional works at: <http://digitalcommons.ryerson.ca/dissertations>

 Part of the [Electrical and Computer Engineering Commons](#)

---

## Recommended Citation

Erküçük, Serhat, "Time-frequency analysis of spread spectrum based communication and audio watermarking systems" (2003). *Theses and dissertations*. Paper 28.

This Thesis is brought to you for free and open access by Digital Commons @ Ryerson. It has been accepted for inclusion in Theses and dissertations by an authorized administrator of Digital Commons @ Ryerson. For more information, please contact [bcameron@ryerson.ca](mailto:bcameron@ryerson.ca).

In compliance with the  
Canadian Privacy Legislation  
some supporting forms  
may have been removed from  
this dissertation.

While these forms may be included  
in the document page count,  
their removal does not represent  
any loss of content from the dissertation.



# Time-Frequency Analysis of Spread Spectrum Based Communication and Audio Watermarking Systems

by

Serhat Erküçük

B.S., Middle East Technical University, Turkey, 2001

A thesis

presented to Ryerson University

in partial fulfillment of the  
requirement for the degree of

Master of Applied Science

in the Program of

Electrical and Computer Engineering

Toronto, Ontario, Canada, 2003

©Serhat Erküçük 2003



National Library  
of Canada

Bibliothèque nationale  
du Canada

Acquisitions and  
Bibliographic Services

Acquisitions et  
services bibliographiques

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file    Votre référence*

*ISBN: 0-612-87154-1*

*Our file    Notre référence*

*ISBN: 0-612-87154-1*

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

**Canada**

## Author's Declaration

I hereby declare that I am the sole author of this thesis.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Author's signature:

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Author's signature:

## Borrower's Page

Ryerson University requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

[illegible]

# Abstract

## Time-Frequency Analysis of Spread Spectrum Based Communication and Audio Watermarking Systems

©Serhat Erküçük 2003

Master of Applied Science  
Department of Electrical and Computer Engineering  
Ryerson University

In this study, we present novel applications of time-frequency analysis to spread spectrum based communication and audio watermarking systems. Our objective is to detect and estimate nonstationary signals, such as chirps, that are characterized by directional elements in the time-frequency plane. Towards this goal, we model nonstationary signals using the matching pursuit decomposition algorithm, generate a positive time-frequency representation of the signal model using the Wigner-Ville distribution and estimate the energy varying directional elements using a line detection algorithm based on the Hough-Radon transform.

Spread spectrum communication systems frequently encounter nonstationary signals with energy varying directional elements as hostile jamming signals. In this thesis, we develop a new interference excision algorithm for spread spectrum communication systems based on the directional element estimation algorithm. At the receiver, we first excise the interference from the spread spectrum signal before despreading and data symbol detection. The new algorithm can excise single and multi-component interferences such that the spread spectrum system can reliably detect the transmitted message symbols even when the interference power exceeds the jamming margin of the system. We verify the effectiveness of the interference excision algorithm using simulation studies.

Watermarking is the process of embedding imperceptible data into the host signal for marking the copyright ownership. The embedded data should be extractable to prove ownership. Watermarking systems face problems similar to those in spread spectrum communication systems, namely, intentional attacks by the adversaries. In watermarking, the adversaries try to obliterate the embedded watermark in order to prevent its detection by authorized parties. In this thesis, we develop a spread spectrum audio watermarking scheme, where we embed perceptually shaped linear chirps as watermark messages. The directional elements of the chirp signals represent different watermark messages. We extract the watermark by first detecting the transmitted message symbols in the spread spectrum signal. We then use the directional element estimation algorithm based on the time-frequency analysis as a post-processing tool to minimize the effects of hostile attacks on the extractability of the embedded watermark. We demonstrate the robustness of the algorithm by extracting the watermark correctly after common signal processing operations representing hostile attacks by adversaries.



## Acknowledgement

I initially would like to thank my supervisor Dr. Sridhar Krishnan, and my co-supervisor Dr. Mehmet Zeytinoglu for their excellent guidance, helpful feedback, and continuous support throughout my research work.

I also would like to thank the Department of Electrical and Computer Engineering and the School of Graduate Studies of Ryerson University for their support in terms of departmental facilities, research stipend, and scholarship.

I would like to acknowledge my supervisors' funding resources National Sciences and Engineering Research Council of Canada (NSERC), and Micronet R&D, Canada for providing me financial support for my research work.

I finally would like to thank IBM T.J. Watson Research Center for recognizing our research work in audio watermarking, and awarding me a grant to attend and present our proposed watermarking algorithm at the Student Workshop of IEEE International Conference on Multimedia and Expo in Switzerland in August 2002.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Spread Spectrum Systems</b>	<b>6</b>
2.1	Characteristics of Spread Spectrum Signals . . . . .	7
2.2	Applications of Spread Spectrum Techniques . . . . .	12
2.2.1	Spread Spectrum Communications . . . . .	12
2.2.2	Spread Spectrum Audio Watermarking . . . . .	16
<b>3</b>	<b>Time-Frequency Analysis</b>	<b>18</b>
3.1	Signal Localization . . . . .	20
3.1.1	Spectrogram . . . . .	21
3.1.2	Wigner-Ville Distribution . . . . .	22
3.1.3	Adaptive Signal Decomposition Techniques . . . . .	24
3.2	Detection of Directional Elements . . . . .	28
3.2.1	The Hough Transform . . . . .	29
3.2.2	The Radon Transform . . . . .	30
3.2.3	Combined Hough and Radon Transform . . . . .	31
<b>4</b>	<b>Spread Spectrum Communications</b>	<b>36</b>
4.1	Review . . . . .	36
4.2	A New Interference Excision Algorithm . . . . .	38
4.3	Simulation Results and Discussion . . . . .	44
<b>5</b>	<b>Spread Spectrum Watermarking</b>	<b>49</b>
5.1	Review . . . . .	50
5.2	Watermark Embedding . . . . .	51
5.3	Watermark Extraction . . . . .	58
5.4	Simulation Results and Discussion . . . . .	63
<b>6</b>	<b>Conclusions and Future Research</b>	<b>68</b>
6.1	Conclusions . . . . .	68
6.2	Future Research . . . . .	72
	<b>Bibliography</b>	<b>74</b>
<b>A</b>	<b>List of Publications</b>	<b>79</b>

# List of Figures

1.1	Overall block diagram for the TF analysis of SS systems proposed in the thesis. . . . .	4
2.1	Signals at the spread spectrum generator. . . . .	7
2.2	Spreading a narrowband signal with a PN sequence of $L = 128$ . . . .	12
2.3	Spread spectrum communication system. . . . .	13
2.4	Spectrograms of signals at various points of a SS system. <b>Top:</b> audio signal $x(t)$ , <b>Middle:</b> SS signal $w(t)$ , <b>Bottom:</b> SS signal in the presence of an interfering chirp. . . . .	15
2.5	Block diagram for SS audio watermarking. . . . .	16
3.1	Representation of two different chirps. <b>Top:</b> time-domain, <b>Middle:</b> FT magnitude spectrum, <b>Bottom:</b> joint TF representation. . . . .	19
3.2	Illustration of a TF tiling. . . . .	21
3.3	TF tilings of STFT. <b>Left:</b> short-window, <b>Right:</b> long-window. . . . .	22
3.4	Spectrogram of a linear chirp with increasing frequency. . . . .	23
3.5	WVD of a linear chirp with increasing frequency. . . . .	24
3.6	TF plane tiling for a two-band wavelet transform. . . . .	25
3.7	TF plane tiling from an adaptive signal decomposition. . . . .	26
3.8	MP TFD of a linear chirp with increasing frequency. . . . .	28
3.9	Line detection using RT. . . . .	31
3.10	Quantization values for $\theta$ and $\rho$ for $K = N = 176$ . . . . .	34
3.11	HRT space for two straight lines. . . . .	35
4.1	Block diagram of a DSSS system. . . . .	38
4.2	Despread waveforms at the receiver with a synchronized PN sequence. <b>Top:</b> reference with no jammer, <b>Middle:</b> JSR = 10 dB, <b>Bottom:</b> JSR = 30 dB. . . . .	40
4.3	BER vs. JSR results for a self-excised SS system. . . . .	40
4.4	Interference excision. . . . .	42
4.5	MP TFD of a multi-component interference consisting of linear and quadratic chirps. . . . .	46
4.6	Chip error rate vs. SNR for JSR = 40 dB. . . . .	46
4.7	Chip error rate vs. SNR for JSR = 5 dB. . . . .	47

4.8	TFDs of <b>Left:</b> SS signal with a linear interference at JSR = 5 dB, <b>Middle:</b> estimate of the interference, <b>Right:</b> interference excised SS signal. . . . .	48
5.1	Watermark embedding scheme. . . . .	52
5.2	Perceptual shaping. . . . .	52
5.3	Time-domain and TF representation of chirp signals. <b>Top:</b> continuous amplitude chirp, <b>Bottom:</b> 1-bit quantized chirp. . . . .	54
5.4	Perceptual shaping of watermark information. . . . .	56
5.5	Spectrograms of <b>Top:</b> the original signal, <b>Middle:</b> perceptually shaped watermark, <b>Bottom:</b> the watermarked signal. . . . .	57
5.6	Watermark extraction. . . . .	58
5.7	Watermark bit detection scheme. . . . .	59
5.8	Postprocessing of extracted bits. . . . .	61
5.9	Changes of the WVD and the HRT with increasing BER. . . . .	67

# List of Tables

5.1	Bit error rate (in percentage) for $N_x = 10,000$ . . . . .	65
5.2	Bit error rate (in percentage) for $N_x = 5,000$ . . . . .	65

# Chapter 1

## Introduction

**S**PREAD spectrum (SS) systems are the main components of today's communication systems based on code division multiple access (CDMA) technology. SS systems allow multiple users with user-specific codes to occupy the same frequency band simultaneously. The user-codes have low cross-correlation values to minimize interferences among users, or interferences resulting from intentional jammers. Commercial implementation of SS techniques such as CDMA has become increasingly wide-spread after the 80's. The first SS systems were developed during the World War-II to support military and diplomatic communications. The motivation was to achieve resistance to intentional interferences and to ensure message privacy. Today, most military communication systems continue to use SS techniques which can be considered as a form of cryptography, which is the science of coding and decoding secret messages. It scrambles message contents with a user-known code rendering the transmitted message undecodable to unauthorized users.

The 90's was the era which saw the emergence of digital coding as the norm of representing multimedia signals, such as the digital audio format on compact disks and more recently the digital video recording format on DVDs. The representation of multimedia signals in digital formats allows users to make copies which are indistinguishable from the original. The increasing popularity of the Internet in parallel with the digital coding of multimedia signals paved the way to unauthorized copying and redistribution of multimedia data. Today, widely available software tools allow an unprecetended number of users to exchange multimedia files over the Internet.

As a result, the video and music industries have been experiencing a very significant drop in the legitimate sales of CDs and DVDs. By all accounts, the worldwide loss of revenue to the music and video industries is in the order of several billion dollars. To address this problem, a number of industry technology groups were established including the Copy Protection Technical Working Group (CPTWG) for the video industry, and the Secure Digital Music Initiative (SDMI) for the music industry [1]. While cryptography can provide copyright protection of multimedia data, it can only protect the data at the source. Once a customer decrypts the data, it is no longer protected. Consequently, these groups recommend watermarking as the means of copyright protection since the watermark remains always part of the multimedia data.

Watermarking is the process of embedding imperceptible data into multimedia signals. Upon the detection of the watermark, the information retrieved can be used to take action, for example, permit or deny playing the multimedia data. Watermarking consists of encoding (embedding) and decoding (detecting) stages and can be seen as a form of communications. The embedding and detecting terms used in watermarking are analogous to modulation (transmitting) and demodulation (receiving) terms used in communications. The transmission channel in a watermarking system is the multimedia data, where the watermark is embedded (transmitted). The main difference between a traditional communication channel and a watermarking channel is the requirement that the quality of multimedia data must be preserved. In addition, a watermark has to satisfy a set of requirements including imperceptibility and robustness. These constraints have led many watermark embedding schemes to utilize SS techniques: SS techniques code the signals with less power, and spread over a wide frequency spectrum—promising imperceptibility, and show resistance to interferences—ensuring robustness.

SS based military communication and watermarking systems face similar problems, namely, intentional attacks by the adversaries. In a communication system, the adversaries try to prevent the authorized parties from correctly detecting the message

signal by transmitting high power frequency modulated (FM) interferences within the same spectral band as the SS signal. In watermarking, the adversaries try to obliterate the embedded watermark in order to prevent its detection by authorized parties. To achieve this goal, the adversaries perform signal manipulations on the multimedia signal. These intentional attacks degrade the performance of both systems, and the message received may potentially be in error.

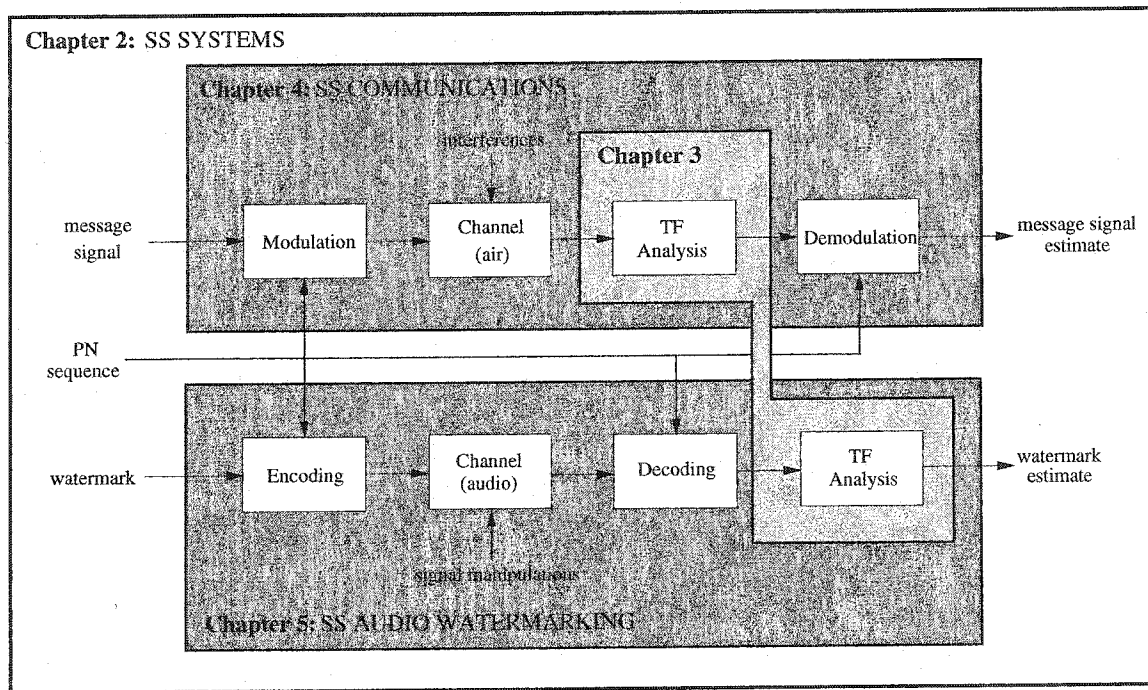
Intentional attacks in military communications are characterized by the injection of nonstationary signals into the communication channel. Therefore, SS systems are designed to identify nonstationary signals with the goal of excising them for improved system performance. Conversely, nonstationary signals can be used as embedded watermarks. The same interference excision tools then can be used for robust identification of the embedded watermarks in the presence of hostile signal manipulations. In this study, we will first consider the problem of analyzing and detecting nonstationary signals. The best way to study a nonstationary signal is by joint time-frequency (TF) analysis, which evaluates the time-varying power spectral density of the signal. TF analysis extracts important information by localizing the signal both in time and frequency. Signal processing techniques can then be used to process the extracted information for enhanced identification of nonstationary signals. One of the signal processing techniques is a line detection technique based on the Hough-Radon transform (HRT) that can detect energy varying directional components in the TF plane. These directional components can represent interferences in a military communication channel or signals we want to estimate in the post-processing stage of a watermarking scheme.

In this study, we develop a novel approach of TF analysis with applications in SS communication and watermarking systems. In a military communication system, the received signal typically includes a jamming signal in the form of a high power FM signal. These FM signals are usually chirps, which exhibit directional characteristics in the TF plane. We model the nonstationary interference using a signal decomposition algorithm based on the matching pursuit (MP) algorithm. We then estimate



the interference using the HRT and excise it from the received signal. We show that the combination of the MP and the HRT overcomes several drawbacks such as poor TF resolution, presence of cross-terms, and limited interference suppression.

In the context of watermarking, we use TF analysis as a post-processing tool to minimize the effects of hostile attacks on the extractability of the embedded watermark. We embed linear chirps as watermark messages after perceptual shaping, which maintains imperceptibility while providing resistance against signal manipulations.



**Figure 1.1:** Overall block diagram for the TF analysis of SS systems proposed in the thesis.

Figure 1.1 provides an overview of TF analysis of SS systems as presented in this thesis. The organization of the thesis is as follows. The next two chapters provide the background information for spread spectrum systems and joint time-frequency analysis. In **Chapter 2**, we introduce SS systems and discuss their characteristics. We present applications of SS techniques in communications and audio watermarking. In **Chapter 3**, we explain the need for joint TF analysis. We introduce the theory of TF analysis and describe the widely used TF analysis techniques including

a line detection algorithm based on the HRT. We study the HRT based line detection algorithm for detecting directional elements extracted from TF distributions (TFDs).

**Chapter 4** presents a SS communication system, where the transmitted signal is jammed with FM-type interferences. We review earlier works in the area of interference suppression, point out their characteristics and limitations. We develop a new interference excision algorithm based on the MP and the HRT algorithms. The new algorithm excises the interferences at the receiver to increase the probability of detection. Simulation results demonstrating the capabilities of the new algorithm conclude the chapter.

**Chapter 5** presents a novel SS audio watermarking algorithm. We review previous watermarking algorithms, and justify the need for a new watermarking algorithm. We explain the watermark embedding process, which includes the perceptual shaping of the SS watermark information. We delineate the watermark detection algorithm based on the TFDs and the HRT introduced in earlier chapters. We subject the new algorithm to a variety of the signal manipulations representing typical hostile attacks and verify its robustness by measuring its success in extracting the embedded watermark.

We conclude with **Chapter 6**, where we discuss the main contributions of this thesis, and point out directions for the future research.

## Chapter 2

# Spread Spectrum Systems

SPREAD spectrum signals are used for the transmission of digital information. A distinguishing characteristic of SS signals is that their bandwidth is much greater than the information rate [2]. In SS techniques, the message signal is spread over a wider bandwidth with a pseudo-noise (PN) code also known by the receiver. The resulting signal is modulated and transmitted over the channel in the form of a low-power wideband signal [3]. These characteristics make it difficult to sense the presence of a SS signal. Even when the adversaries know the presence of the SS signal, they cannot detect the transmitted message without access to the PN code. As a result, adversaries may attempt to jam the signal in order to prevent the authorized party from detecting the message signal. However, the increased bandwidth of the SS signal resulting from modulation by the wideband PN signal provides a high degree of interference suppression at the receiver, and increases the probability of delivery of the message signal. Consequently, SS signals can be used for secure communications.

In this chapter, we present a simple model for the SS systems and explain the system characteristics. We focus our discussion on the characteristics that make the system secure. We then introduce two applications of SS systems that are based on private messaging, and discuss possible jamming techniques used by adversaries to degrade the performance of the system.

## 2.1 Characteristics of Spread Spectrum Signals

SS systems are composed of three main components: the message signal  $m(t)$ , the spreading signal  $p(t)$ , and the SS signal  $w(t)$ . Let the message signal  $m(t)$ , be represented as

$$m(t) = \sum_k b_k \text{rect}_{T_m}(t - kT_m), \quad (2.1)$$

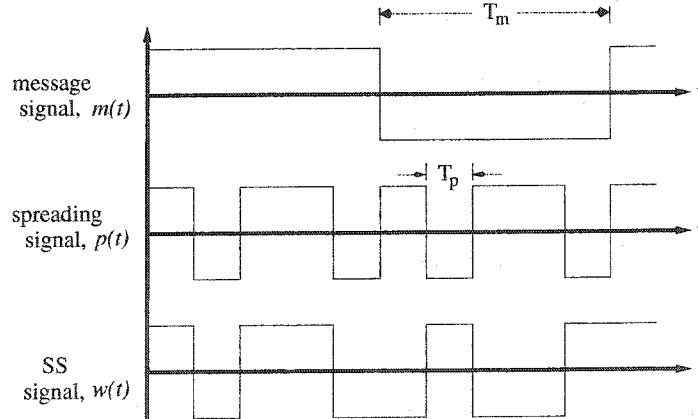
where  $b_k = \pm 1$ , for  $k \in \mathcal{Z}$ , and  $\text{rect}_\tau(t)$  is a rectangular pulse of duration  $\tau$ . Let  $p(t)$  be the spreading signal expressed as

$$p(t) = \sum_{n=0}^{L-1} c_n \text{rect}_{T_p}(t - nT_p), \quad (2.2)$$

where  $c_n = \pm 1$  is the  $n$ th chip of the  $L$ -element PN sequence with  $T_m/T_p = L \gg 1$ . Let  $w(t)$  be the SS signal such that

$$\begin{aligned} w(t) &= m(t)p(t), \\ &= \sum_k b_k p(t - kT_m). \end{aligned} \quad (2.3)$$

Figure 2.1 shows  $m(t)$ ,  $p(t)$  and  $w(t)$  for the case of  $T_m = 5T_p$ .



**Figure 2.1:** Signals at the spread spectrum generator.

The PN sequence  $\{c_n\}$  is periodic with period  $L$  such that:

$$\sum_{n=0}^{L-1} c_n c_{n+kL} = L, \quad (2.4)$$

for  $k \in \mathcal{Z}$ . Furthermore, the PN sequence  $\{c_n\}$  also satisfies the properties [5]:

$$\mathbb{E} \left[ \sum_{n=0}^{L-1} c_n \right] = 0, \quad (2.5)$$

and

$$R_c(i) = \mathbb{E}[c_n c_{n+i}] = L \delta(i), \quad (2.6)$$

where  $\mathbb{E}[\cdot]$  is the expected value operator and  $R_c$  is the autocorrelation function of the PN sequence  $\{c_n\}$ .

The SS signal  $w(t)$  exhibits the following characteristics:

- The SS modulation scheme increases the bandwidth of the message signal by a factor called the processing gain. If the message signal bandwidth is  $B_m$  Hz and the corresponding SS signal bandwidth is  $B_w$  Hz, then

$$\text{Processing Gain} = B_w / B_m. \quad (2.7)$$

In addition, the SS signal is transmitted as a low-power signal. Therefore, the presence of SS signals are difficult to detect.

- The adversaries cannot decode the SS signal without access to the PN sequence.
- SS signals are resistant to intentional jamming, unintentional interferences from other users and self interference due to multipath propagation.

Comparison of SS and binary discrete-time systems demonstrates the advantages of SS systems [4]. Consider a binary discrete-time communication system which transmits the elements of a binary information sequence  $\{b_k\}$  over a channel at  $T_m$  second intervals. The  $k$ th element<sup>1</sup> of the received sequence is;

$$r_k = \mathcal{E} b_k + n_k, \quad (2.8)$$

where  $\mathcal{E} > 0$  is the energy of the pulse representing each message symbol.  $n_k$  is the  $k$ th element of a zero-mean, additive white Gaussian noise (AWGN) sequence such that

---

<sup>1</sup>In the discussion that follows, we use the simplified notation without time arguments such that  $x_k$  represents the value of the signal  $x(t)$  taken at the  $k$ th sampling instant.

$n_k \sim \mathcal{N}(0, \sigma^2)$  with  $\sigma^2 = \mathbb{E}[n_k^2]$ . Furthermore,  $\{n_k\}$  is assumed to be uncorrelated with the PN sequence  $\{c_k\}$ . Under the AWGN assumption, we define the test statistic  $\Lambda = r_k$ , where  $\Lambda \sim \mathcal{N}(\mathcal{E}b_k, \sigma^2)$ . At the receiver, we use  $\Lambda$  to estimate the transmitted message elements as:

$$\hat{b}_k = \begin{cases} +1, & \text{if } \Lambda \geq 0, \\ -1, & \text{if } \Lambda < 0. \end{cases} \quad (2.9)$$

In the case of a SS system, each element  $b_k$  of the message sequence  $\{b_k\}$  is transmitted after multiplied by the spreading signal  $\mathbf{p}_k = [c_0 \dots c_{L-1}]^T$ , where  $\mathbf{p}_k$  represents the  $L$ -element PN sequence  $\{c_n\}$ . The resulting signal is

$$\mathbf{w}_k = \mathcal{E}_L b_k \mathbf{p}_k, \quad (2.10)$$

where  $\mathcal{E}_L = \mathcal{E}/L$  is the pulse energy of the SS signal. At the receiver the AWGN corrupted signal<sup>2</sup>  $\mathbf{r}_k = \mathbf{w}_k + \mathbf{n}_k$ , is correlated with  $\mathbf{p}_k$  to determine the test statistic:

$$\begin{aligned} \Lambda &= \langle \mathbf{r}_k, \mathbf{p}_k \rangle = \mathbf{p}_k^T \mathbf{r}_k, \\ &= \mathcal{E} b_k + \mathbf{p}_k^T \mathbf{n}_k. \end{aligned} \quad (2.11)$$

Using the statistical properties of the noise and PN sequences, we can show that  $\Lambda \sim \mathcal{N}(\mathcal{E}b_k, \sigma^2)$  [4]. Therefore,  $\Lambda$  can again be used as a test statistic to estimate the transmitted message symbols using the decision rule given in Equation (2.9). A comparison of this result with the binary discrete-time communication system shows that spreading the message and despreading it at the receiver with a properly synchronized PN sequence results in identical detection performance under the AWGN assumption.

## Interference Suppression

Let  $I$  be a constant amplitude interference added to the transmitted signal. If we use a binary discrete-time communication system, the test statistic  $\Lambda$  will have a mean value of  $\mathcal{E}b_k + I$  and variance  $\sigma^2$ . The estimated value  $\hat{b}_k$  will be incorrect if  $|I|$  is

---

<sup>2</sup>Each noise sample in  $\mathbf{n}_k$  has a variance of  $\sigma^2/L$ . This observation ensures that the power of the noise term  $n_k$  in Equation (2.8) and of  $\mathbf{n}_k$  in Equation (2.11) remain the same. Since both  $r_k$  and  $\mathbf{r}_k$  are observed over the same time duration  $T_m = LT_p$ , the total noise power levels in both cases are equal.

sufficiently large. If the same constant amplitude interference component interferes with the SS system, the received signal will be

$$\mathbf{r}_k = \mathcal{E}_L b_k \mathbf{p}_k + \mathbf{i}_k + \mathbf{n}_k, \quad (2.12)$$

where  $\mathbf{i}_k = I$ . We construct the decision variable, i.e., the test statistic, by correlating the received signal with the PN sequence as before:

$$\Lambda = \langle \mathcal{E}_L b_k \mathbf{p}_k + \mathbf{i}_k + \mathbf{n}_k, \mathbf{p}_k \rangle, \quad (2.13)$$

which results in:

$$\Lambda = \mathcal{E} b_k + I \sum_{n=0}^{L-1} p_k(n) + \mathbf{p}_k^T \mathbf{n}_k, \quad (2.14)$$

or

$$\Lambda \approx \mathcal{E} b_k + 0 + \mathbf{p}_k^T \mathbf{n}_k. \quad (2.15)$$

Since  $\Lambda \sim \mathcal{N}(\mathcal{E} b_k, \sigma^2)$ , the SS system can successfully suppress the interference as a result of the despreading operation at the receiver.

### Multiple access

Let the number of users in a SS channel be  $M$ , where the  $m$ th user has the despreading signal  $\mathbf{p}_k^{(m)}$  with the PN sequence  $\{c_n^{(m)}\}$ , which satisfies the cross-correlation property

$$\frac{1}{L} \sum_{n=0}^{L-1} c_n^{(m)} c_{n+i}^{(l)} \approx \begin{cases} 1, & m = l, i = 0; \\ 0, & m = l, 0 < |i| < L; \\ 0, & m \neq l. \end{cases} \quad (2.16)$$

As a result of the cross-correlation property, the PN sequences have zero cross-correlation value. Assuming synchronization among the signals, the received is

$$\mathbf{r}_k = \sum_{m=1}^M \mathcal{E}_L^{(m)} b_k^{(m)} \mathbf{p}_k^{(m)} + \mathbf{n}_k, \quad (2.17)$$

where  $\mathcal{E}_L^{(m)} = \mathcal{E}^{(m)}/L$ ,  $m = 1, \dots, M$ . To receive the signal of the  $m_0$ th user, we use the corresponding despreading signal  $\mathbf{p}_k^{(m_0)}$  at the receiver and construct the decision variable:

$$\Lambda^{(m_0)} = \mathcal{E}_L^{(m_0)} b_k^{(m_0)} [\mathbf{p}_k^{(m_0)}]^T \mathbf{p}_k^{(m_0)} + \sum_{\substack{m=1, \\ m \neq m_0}}^M b_k^{(m)} [\mathbf{p}_k^{(m)}]^T \mathbf{p}_k^{(m_0)} + [\mathbf{p}_k^{(m_0)}]^T \mathbf{n}_k. \quad (2.18)$$

Using the cross-correlation property, we simplify the decision variable to the form:

$$\Lambda^{(m_0)} \approx \mathcal{E}^{(m_0)} b_k^{(m_0)} + 0 + [\mathbf{p}_k^{(m_0)}]^T \mathbf{n}_k. \quad (2.19)$$

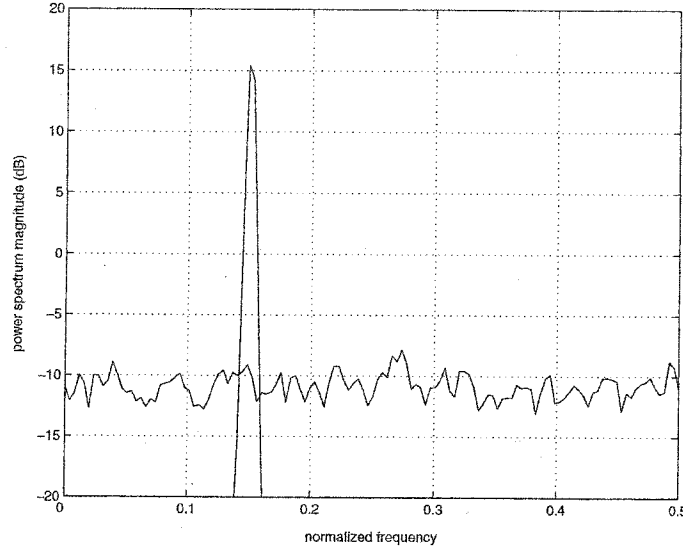
Hence, the presence of other users does not degrade the performance of the system. The cross-correlation property of the PN sequences allows simultaneous transmission over the same channel. Such systems are usually called CDMA systems. Using the cross-correlation property of the PN sequences, we can also show that a well designed SS system can effectively suppress time delayed versions of other transmitted signals, delayed versions of the signal itself resulting from multipath propagation, and channel dispersive effects [4].

In modeling the SS system, we assumed the perfect synchronization of the received signal and the PN sequence. If the signals are not synchronized at the receiver, the correlation of the signals will yield low or zero correlation value, and the transmitted information will be lost. Therefore, synchronization is one of the most important requirements in SS systems. Since the main focus of this thesis is signal processing in SS systems, we assume perfect synchronization, and also do not consider multipath and multiuser effects.

### Message privacy

A properly synchronized and structured PN sequence is essential for the proper decoding of a SS signal generated by the same PN sequence. Therefore, unauthorized users who do not know the PN sequence used in the generation of a SS signal, cannot decode the transmitted message. Besides being code-dependent, SS signals are transmitted as low-power wideband noise, which is a property that renders SS signal indistinguishable from background noise. Consequently, the presence of a SS signal is difficult to detect. Figure 2.2 shows the power spectral densities of a narrowband signal  $m(t)$ , and the SS signal  $w(t)$ , where  $w(t) = m(t)p(t)$ .  $m(t)$  and  $p(t)$  are defined by Equations (2.1) and (2.2), respectively. The chip length of  $p(t)$  is  $L = 128$ . The resulting signal  $w(t)$  has flat frequency spectrum at a low power level.





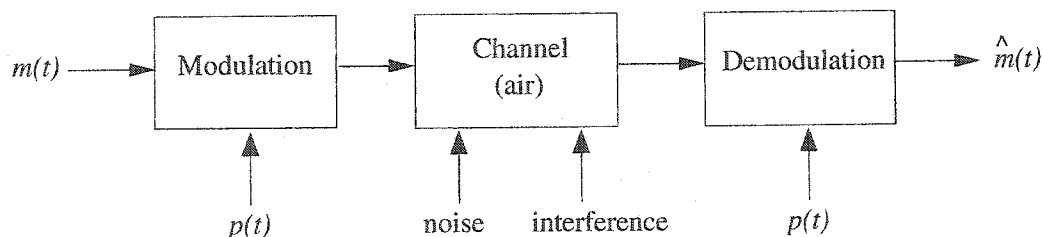
**Figure 2.2:** Spreading a narrowband signal with a PN sequence of  $L = 128$ .

## 2.2 Applications of Spread Spectrum Techniques

Applications of SS systems are very diverse: they are widely used in space systems, avionics systems, test systems, position localization systems, mobile communication systems as well as in military communications and watermarking [3]. Combination of privacy, and the transmission of multiple messages with high resistance to interferences, makes SS systems an important tool for secure message transmission in military communications and watermarking [3, 6]. In military communications and watermarking, a completely different class of interference, namely, intentional interferences and/or jamming, becomes the primary concern. These interferences can substantially degrade the system performance and potentially render the entire communication link unusable. Signal processing techniques can be used to increase the robustness of the SS systems against intentional interferences. In the remainder of this chapter, we will provide two examples, where signal processing techniques are applied to *spread-spectrum communications* and *spread-spectrum audio watermarking*.

### 2.2.1 Spread Spectrum Communications

Figure 2.3 provides an overview of a SS communication system. There are two main approaches to SS communications [2]. In the first approach, the message signal  $m(t)$



**Figure 2.3:** Spread spectrum communication system.

is multiplied by a wideband PN waveform  $p(t)$  prior to modulation, resulting in the SS signal  $w(t) = m(t)p(t)$ . The SS signal is then modulated to produce a double-sideband (DSB) signal. If  $m(t)$  is a polar signal taking the values  $\pm 1$  representing a digital message, the output from the DSB modulator would be a binary phase-shift keying (BPSK) signal. This system is called the *direct-sequence spread spectrum* (DSSS) system.

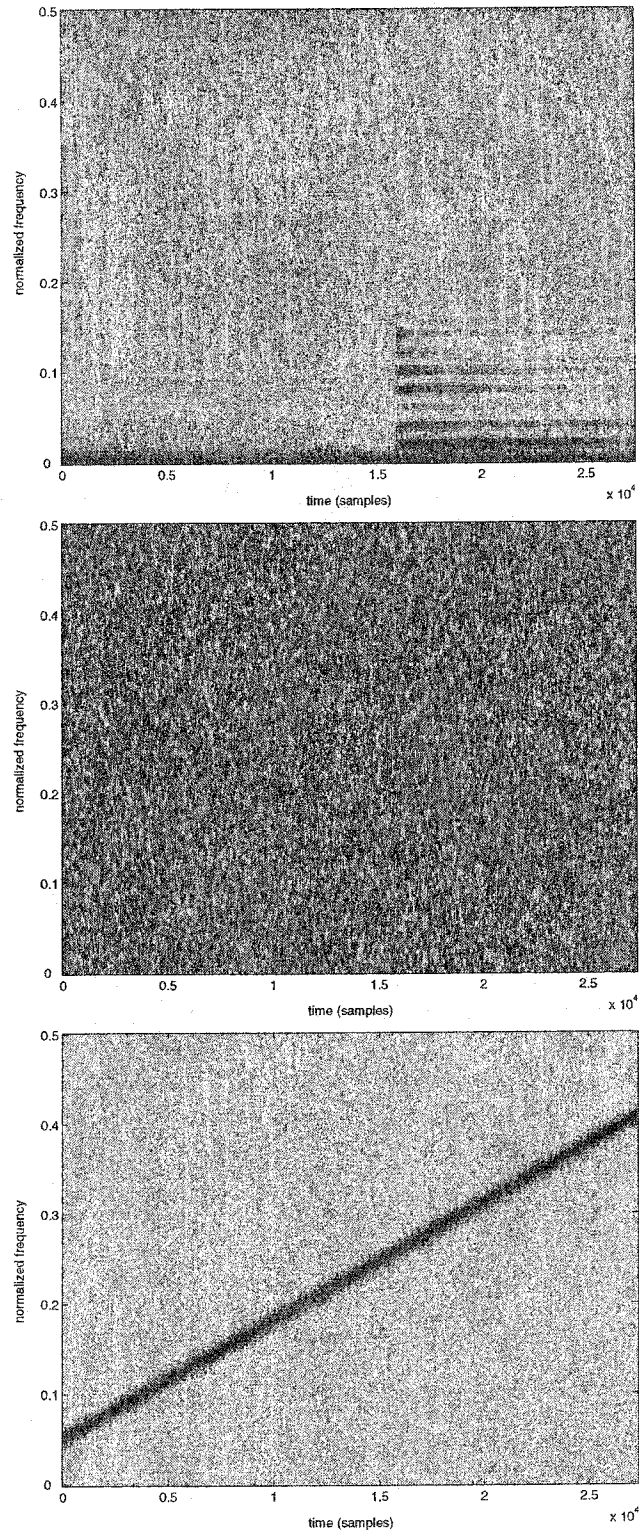
Alternatively, the PN generator can drive a frequency synthesizer that produces a wideband sequence of frequencies that can cause the data-modulated carrier to hop from one frequency to another. This system is called the *frequency hopping spread spectrum* (FHSS) system. A FHSS system is analogous to a frequency shift keying (FSK) system, with a greater range for frequency choices. While there are other modulation methods, the DSSS and FHSS systems represent by far the most frequently used SS techniques in signal transmission [3].

During the transmission of the SS signals, noise and jamming signals may interfere. As we demonstrated in Section 2.1, SS systems are intrinsically capable of suppressing interferences. These interference suppression properties, however, have been derived under a strict set of assumptions such as additive and uncorrelated channel noise and perfect synchronization between the transmitter and the receiver. Deviation from these assumptions is likely to place practical limits on the interference excision capability, thus limiting the expected performance of real-world SS systems. Consequently, we may consider approaches such as increasing the processing gain, increasing the power of the transmitted signal or decreasing the number of users within the same channel to overcome the practical limitations of SS systems. Such methods,

however, are costly and the resources may not be available [7]. Alternatively, signal processing techniques can assist in interference suppression.

Most interference suppression techniques are designed to deal with narrowband interferences. One possible approach to interference suppression includes time-domain methods including adaptive notch and prediction/estimation type filtering. A second approach uses transform-domain filtering, which exploits the time-frequency distributions of the SS signals to excise the interference. Due to the dynamic nature of typical interference signals, all interference suppression techniques are inherently adaptive.

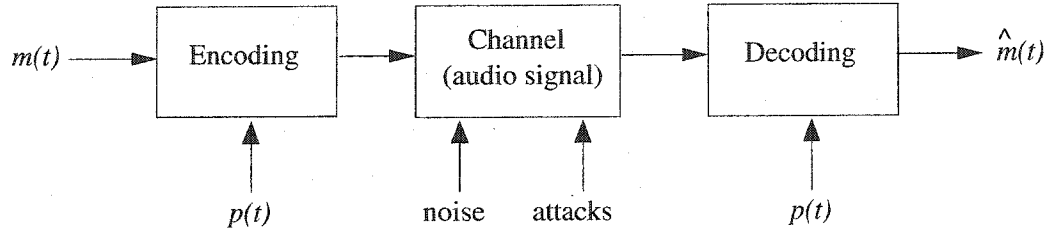
As a result of spreading with the PN sequence, the SS signal is a wideband signal with a time-frequency distribution that covers the entire TF plane. Conversely, jamming signals from hostile sources exhibit narrowband characteristics such as a straight line representation (chirp) in the TF plane. Figure 2.4 shows the spectrograms of a message signal (600 ms segment from a music file sampled at 44.1 kHz) before and after spreading with a PN signal. The figure also shows the spectrogram of the SS signal with an interfering chirp signal. In Chapter 4, we will present an interference excision algorithm designed to improve the detection performance of spread spectrum systems, which experience hostile attacks in the form of chirp signals.



**Figure 2.4:** Spectrograms of signals at various points of a SS system. **Top:** audio signal  $x(t)$ , **Middle:** SS signal  $w(t)$ , **Bottom:** SS signal in the presence of an interfering chirp.

### 2.2.2 Spread Spectrum Audio Watermarking

Watermarking is the process of embedding additional data into the host signal—multimedia data such as audio, image or video—for marking copyright ownership. Watermarking is a widely pursued topic in the field of multimedia signal processing. *Spread spectrum watermarking* is a type of watermarking which uses the fundamentals of SS techniques. In the context of watermarking, the host, i.e., the multimedia signal, functions as the “channel”, which we attempt to securely transmit the embedded watermark message through. To ensure the secure and reliable transmission of the watermark signal, we use spread spectrum techniques. Figure 2.5 presents the block diagram for an audio watermarking system.



**Figure 2.5:** Block diagram for SS audio watermarking.

SS audio watermarking techniques are analogous to those used in SS communications. In SS communications, the SS signal remains imperceptible during transmission over a communication channel due to its wideband characteristics resulting from modulation by the PN waveform. In the case of audio watermarking, we first convert the message, i.e., the watermark signal  $m(t)$ , into a wideband signal by multiplying with the spreading signal  $p(t)$  to generate the SS signal  $w(t) = m(t)p(t)$ . Next, we perceptually shape the SS signal to ensure that it will not degrade the quality of the host audio signal. Finally, we embed the perceptually shaped SS signal into the host audio signal, a process which in the context of SS watermarking is analogous to transmission of a SS signal over a communications channel.

Audio signals are highly nonstationary. The top spectrogram in Figure 2.4 clearly demonstrates the nonstationary characteristics of the audio signal which has most of its energy concentrated in the lower frequency band. Its energy in the higher

frequency bands changes with the dynamics of the music. These characteristics of the audio signals should be considered for perceptually shaping the SS watermark signal.

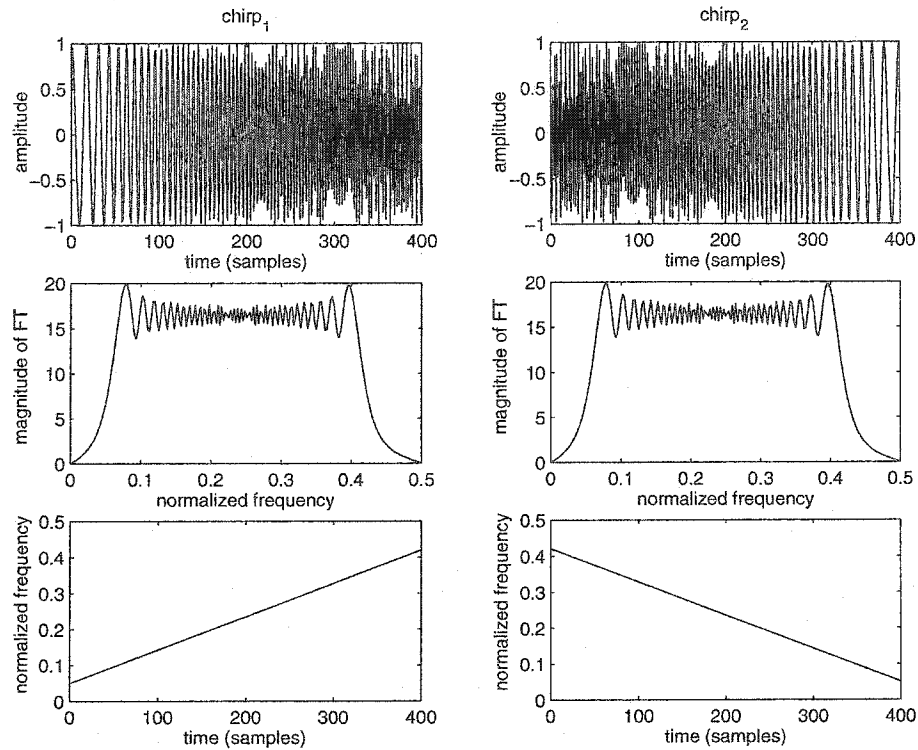
In the context of audio watermarking, the “adversaries” are the parties that will attempt to obliterate the watermark embedded in the audio signal in order to prevent the tracing of the copyright back to its rightful owner(s). Let us consider an audio signal with an embedded SS watermark. We assume that the adversaries know the SS watermarking algorithm used to hide the watermark information—a common scenario as the watermark embedding algorithms are typically made public to ensure their widespread use. Since the adversaries do not have access to the watermark embedding key, i.e., the PN code, they cannot monitor, detect and excise the watermark from the audio signal. It is unwise for the adversaries to embed high power interferences which will degrade the quality of the audio signal. Instead, they are likely to perform signal processing operations to make the embedded SS watermark information irrecoverable while retaining the quality of the signal. In other words, the adversaries will function as the jammer in the SS communications example. Common signal processing operations to jam the watermark message are lowpass filtering, band-pass filtering, compression, resampling, amplitude scaling and adding noise. These operations may alter the signal characteristics without degrading the audio quality. Yet, they may also introduce sufficient distortion to the host signal, such that when we attempt to recover the watermark, some bits representing the watermark would be incorrectly detected. In this case, it may be necessary to postprocess the extracted bits to increase the probability of detection of the embedded watermark. In Chapter 5, we will introduce a robust watermarking scheme that embeds perceptually shaped chirps, and extracts them by utilizing TF analysis in the post-processing stage.

## Chapter 3

# Time-Frequency Analysis

A SIGNAL is the physical representation of information, such as a waveform representing voltage, temperature, intensity, pressure, etc. Signal processing is concerned with the representation, transformation and manipulation of signals, and the information they contain. Frequency and energy of a signal are two important parameters used in signal analysis. The Fourier transform (FT) is a widely used technique to analyze the frequency and energy content of the signal. FT assumes the signal to be stationary, and evaluates the frequency content of the entire signal. However, most of the real-world signals are nonstationary, where the signal's frequency content varies with time. Therefore, analyzing only the frequency content of the signal without considering the time would be inadequate to extract the necessary information from the signal.

Joint time-frequency (TF) analysis focuses on the localization of the energy of a signal in both time and frequency domains. To demonstrate the need for joint TF analysis, let us consider chirps, which are nonstationary signals with frequency information varying as a function of time. Figure 3.1 depicts two different chirps. The left column plots represent a chirp with linearly increasing frequency, whereas the right column plots represent a second chirp with linearly decreasing frequency. As the magnitude plots of their respective FTs in the middle row show, both chirp signals appear to have the same frequency composition. Since the changes in the frequency content of each chirp signal as a function of time are different, the FT magnitude spectrum is not sufficient to differentiate between the two signals. While the inclusion



**Figure 3.1:** Representation of two different chirps. **Top:** time-domain, **Middle:** FT magnitude spectrum, **Bottom:** joint TF representation.

of the phase spectrum of the chirps would allow us to differentiate between the two chirps, the information provided by the FT alone will not be useful for determining their joint TF characteristics. On the other hand, the joint TF representations of the two chirp signals as depicted in the lower row of plots in Figure 3.1, clearly show how the signals' instantaneous frequencies change with respect to time, and allow monitoring of the signal dynamics.

In the above example, the joint TF representation of two chirp signals exhibit a very fine time and frequency resolution. However, there are limitations on the resolution that can be achieved by different TF representations. In the next section, we discuss the theoretical and practical limits on TF resolution, and introduce some of the common TF analysis techniques.



### 3.1 Signal Localization

Time-frequency distributions (TFDs) are two-dimensional (2D) energy distributions of a signal as a function of time and frequency. The TFD of a signal can be seen as an image, where the intensity of each pixel corresponds to its energy. There are different TFD methods. We can classify these methods into two main categories: (i) representations, and (ii) decompositions. Representations obtained by using linear or quadratic transforms refer to computing the frequency content of a signal using different windows. Decompositions refer to decomposing and representing the signal as a linear combination of TF functions. Let us consider the decomposition of the signal  $x(t)$  in terms of a orthonormal basis as:

$$x(t) = \sum_{j,k} a_{j,k} w_{j,k}(t), \quad (3.1)$$

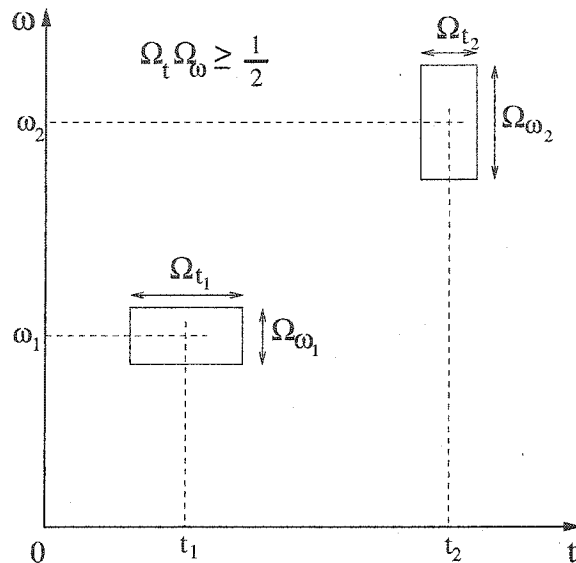
where the basis functions are of the form:

$$w_{j,k} = w(t - t_j) e^{-i\omega_k t}, \quad (3.2)$$

and  $w(t)$  is the window function. If  $\Omega_t$  and  $\Omega_w$  are the time and frequency resolutions of  $w(t)$ , then the uncertainty principle demands [8]:

$$\Omega_t \Omega_w \geq \frac{1}{2}. \quad (3.3)$$

Each basis function in the expansion of  $x(t)$  can be considered schematically as a tile in the TF plane. Each tile graphically illustrates the energy concentration of the basis functions where the time and frequency resolution parameters,  $\Omega_t$  and  $\Omega_w$ , represent the dimensions of this tile. The uncertainty principle allows us to draw two conclusions: (1) we cannot achieve an infinitesimal tile (perfect resolution); (2) we can increase the frequency resolution at the expense of losing time resolution, or vice versa. The equality condition holds only for the Gaussian functions, which optimize the concentration of signal's energy in the joint TF domain [8]. Figure 3.2 shows two TF tiles centered at points  $(t_1, \omega_1)$  and  $(t_2, \omega_2)$ , with dimensions  $\Omega_{t_1} \times \Omega_{\omega_1}$  and  $\Omega_{t_2} \times \Omega_{\omega_2}$ , respectively.



**Figure 3.2:** Illustration of a TF tiling.

### 3.1.1 Spectrogram

The most basic and the widely used TF analysis technique is the short-time Fourier transform (STFT). In STFT, a sliding window divides the signal into overlapping time segments, and evaluates the Fourier spectrum for each segment. The discrete STFT of signal<sup>1</sup>  $x(n)$  is given by

$$\mathcal{X}(n, w) = \sum_{m=-\infty}^{\infty} x(n+m)w(m)e^{-jwm}, \quad (3.4)$$

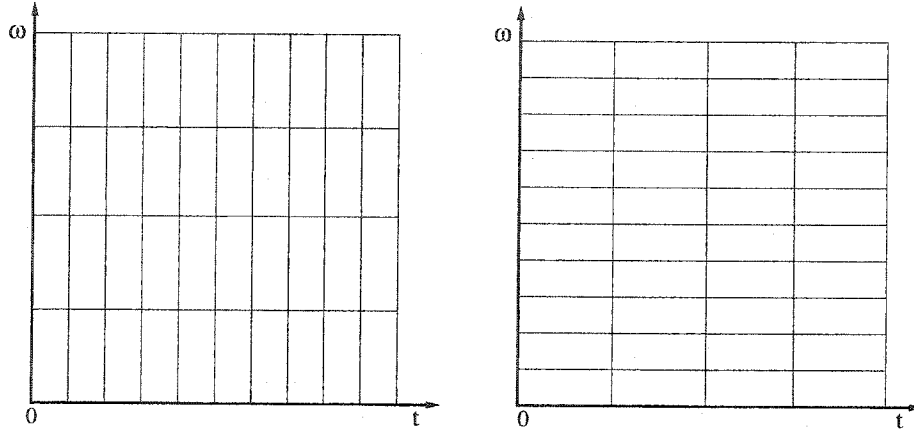
where  $\{w(n)\}$  is the window sequence.  $\mathcal{X}(n, w)$  is a two-dimensional function of the discrete time variable  $n$  and the continuous frequency variable  $w$ . Spectrogram is the squared modulus of the STFT defined as

$$\mathcal{S}_x(n, w) = |\mathcal{X}(n, w)|^2. \quad (3.5)$$

In the spectrogram, the selection of the window function  $w(n)$  is important. A window function with a long time spread can achieve a high frequency resolution. However, one must note that for a long-window, the FT will consider each segment as stationary. If the signal is highly nonstationary, then the spectrogram will not provide a good

<sup>1</sup>We will be processing discrete-time signals in the following chapters. Therefore, we introduce the signals used in the TF analysis with their discrete-time representations.

representation of the TFD of the signal. On the other hand, choosing a short-window will result in coarse frequency resolution. Figure 3.3 shows TF tilings corresponding to a short-window and a long-window based spectrogram. As the length of the window size approaches infinity, the width,  $\Omega_w$ , of the tile will approach zero, and the TF representation will become the FT representation without time localization.



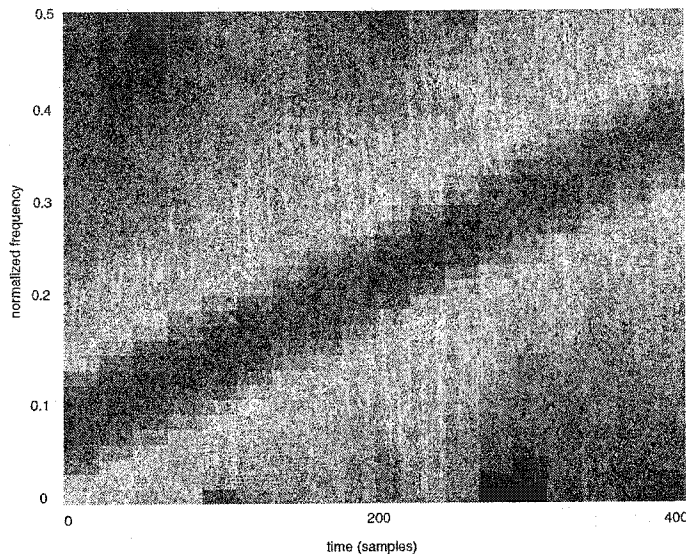
**Figure 3.3:** TF tilings of STFT. **Left:** short-window, **Right:** long-window.

The fixed length window function used by the spectrogram can be a significant disadvantage in dealing with signals that exhibit stationary and nonstationary characteristics in different time segments, as the TF resolution will not be optimal. Despite its shortcomings and limitations, the spectrogram is still a popular TF analysis tool mainly due to efficient and easy methods of computation. Figure 3.4 shows the TFD<sup>2</sup> of a chirp using spectrogram. The spectrogram exhibits poor frequency resolution due to the fixed length window function. However, the spectrogram conveys the information that the signal is a chirp with a linearly increasing frequency modulation.

### 3.1.2 Wigner-Ville Distribution

The Wigner-Ville distribution (WVD) was originally developed by the quantum physicist E.P. Wigner, and later adapted to signal processing by J. Ville. It is a quadratic

<sup>2</sup>The frequency axis is normalized with reference to the sampling frequency of the signal. We will use this normalization throughout this thesis.



**Figure 3.4:** Spectrogram of a linear chirp with increasing frequency.

distribution and it forms the prototype for Cohen's class of bilinear TFDs [8]. Let  $\mathcal{W}(n, w)$  be the discrete WVD of  $x(n)$  given by

$$\mathcal{W}(n, w) = \sum_{k=-\infty}^{\infty} x(n+k)x^*(n-k)e^{-jwk}, \quad (3.6)$$

where  $x^*(n)$  is the complex conjugate of the signal.  $\mathcal{W}(n, w)$  can be seen as the FT of the autocorrelation of  $x(n)$ . This provides a very high TF resolution. Figure 3.5 shows the WVD of a linear chirp. The main drawback of  $\mathcal{W}(n, w)$  is that it suffers from cross-terms in the presence of multi-component signals. Cross-terms are highly oscillating components that appear in TFDs. They occur due to the superposition of different spectral components. For a multi-component signal, if the signal has a frequency of  $w_1$  at time  $n_1$ , and a frequency of  $w_2$  at time  $n_2$ , then a cross-term appears with a frequency of  $w_{12} = (w_1 + w_2)/2$ , at time  $n_{12} = (n_1 + n_2)/2$  [8]. For a multi-component signal, the TF representation constructed using the WVD may lead to incorrect TF information, although the cross-terms in the TF plane may not have much energy as the main spectral component. In the category of Cohen's class of bilinear TFDs, there have been several works done to minimize the effects of cross-terms [8, 9, 10].

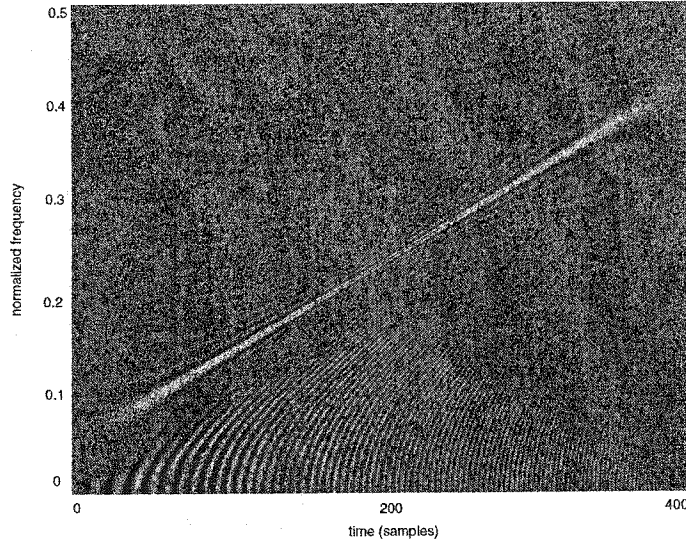


Figure 3.5: WVD of a linear chirp with increasing frequency.

### 3.1.3 Adaptive Signal Decomposition Techniques

#### Wavelet Decomposition

The shortcomings of fixed windowing techniques and Cohen's class of bilinear TFDs are respectively, limited TF resolution and existence of cross-terms. The wavelet analysis [11] provides a solution to the fixed TF resolution. A wavelet is a zero-mean waveform with a finite time-domain support. A discrete-time wavelet is represented with the function  $\psi(n)$  and referred as the mother wavelet. In wavelet transform, the mother wavelet is scaled (stretched and compressed to match the high and low frequency components of the signal) and slid across the signal. The wavelet transform results in the decomposition of the signal  $x(n)$  as

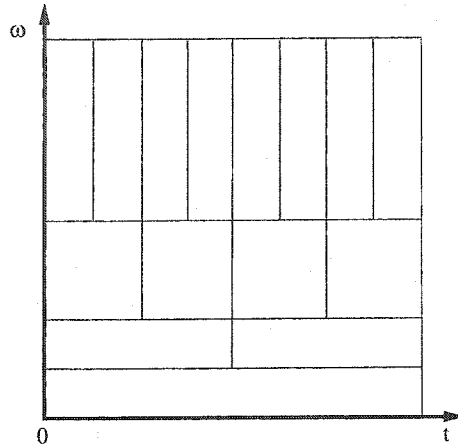
$$x(n) = \sum_{j,k} d_{j,k} \psi(2^j n - k). \quad (3.7)$$

where the expansion coefficients  $d_{j,k}$  are found as

$$\begin{aligned} d_{j,k} &= \langle x(n), \psi(2^j n - k) \rangle \\ &= \sum_n x(n) \psi(2^j n - k). \end{aligned} \quad (3.8)$$

The inner product in Equation (3.8) calculates the similarity measure between the  $\psi(n)$  and  $x(n)$  at different scales  $j$ , and time instants  $k$ . Figure 3.6 shows the TF

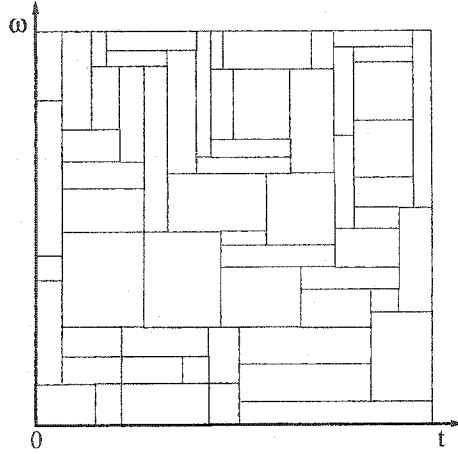
plane tiling resulting from a two-band wavelet transform at the third decomposition level. The partitioning resulting from a two-band wavelet transform can be compared



**Figure 3.6:** TF plane tiling for a two-band wavelet transform.

with the subband decomposition of a signal using a dyadic filterbank structure [11]. Another important characteristic of the wavelet transform observed from the TF tiling diagram is that at high frequencies, wavelets provide good time but poor frequency resolution, which is essential for accurate localization for the onset of high frequency signal components.

We can have a non-uniform TF tiling by searching for different frequency components within each TF tile. In this approach, the signal is decomposed into functions from the over-complete dictionary of orthogonal functions. The corresponding TF tiles can take various sizes by adapting to the energy of the signal for joint TF information. Figure 3.7 shows the tiling of the TF plane resulting from such an adaptive signal decomposition. TFDs based on adaptive signal decomposition have two advantages. First, the signal can be represented with TF functions that have different time and frequency supports, hence achieving a good TF resolution. Second, in case of a multi-component signal, decomposing the signal into its constituent TF functions allows the prevention and removal of cross-terms in the TFD.



**Figure 3.7:** TF plane tiling from an adaptive signal decomposition.

### Matching Pursuit Algorithm

The matching pursuit (MP) algorithm [12] is an adaptive signal decomposition technique that can decompose the signal into its TF functions. In MP, the signal  $x(n)$  of length  $N$  is decomposed into a linear combination of TF functions in  $\{g_{\gamma_m}(n)\}$ , and can be represented as

$$x(n) = \sum_{m=0}^{\infty} a_m g_{\gamma_m}(n), \quad (3.9)$$

where

$$g_{\gamma_m}(n) = \frac{K_{s_m}}{\sqrt{s_m}} g\left(\frac{n - p_m}{s_m}\right) e^{j\left(\frac{2\pi k_m}{N}n + \phi_m\right)}. \quad (3.10)$$

The set  $\{a_m\}$  are the expansion coefficients, and  $\{g(n)\}$  is the window function.  $K_{s_m}$  normalizes  $g(n)$ . The scale factor  $s_m$ , and the temporal placement parameter  $p_m$ , control the width and the displacement of the window function, respectively. The parameters  $k_m$  and  $\phi_m$  represent the frequency and the phase of the exponential function, respectively.  $k_m$  allows the search for different frequencies at each scale. The discrete dictionary is limited with the set  $\{\gamma_m\} = \{(s_m, p_m, 2\pi k_m/N), 1 < s_m < N, 0 \leq p_m < N, \text{ and } 0 \leq k_m < N\}$ . One possible set of functions to be used in the dictionary is the set of Gaussian functions, where

$$g(\lambda) = e^{-\lambda^2}. \quad (3.11)$$

The equality in the uncertainty principle expressed in Equation (3.3) holds for Gaussian signals resulting in an optimal TF resolution [8].

In MP, the signal  $x(n)$  is projected onto the dictionary  $\{g_{\gamma_m}(n)\}$  of TF functions with all possible window sizes, frequencies and temporal placements. At each iteration, the best-correlated function  $g_{\gamma_m}$  is selected from the dictionary and the remainder of the signal, which is called the residue, is further decomposed using the same iteration procedure. After  $M$  iterations, the signal  $x(n)$  can be represented as

$$x(n) = \sum_{m=0}^{M-1} \langle R^m x, g_{\gamma_m} \rangle g_{\gamma_m}(n) + R^M x, \quad (3.12)$$

where  $R^m x$  represents the residue of the signal  $x(n)$  after  $m$  iterations with  $R^0 x = x$ , such that

$$a_m = \langle R^m x, g_{\gamma_m} \rangle. \quad (3.13)$$

The first term in Equation (3.12) represents the first  $M$  Gaussian functions best matching the signal (we will refer to the first term as  $x'(n)$ ) and the second term (referred as  $x''(n)$ ) represents the residue of the signal  $x(n)$ . In order for the signal to be fully decomposed, the iteration process continues until all the energy in the residue signal is consumed. However, for some applications such as denoising, the signal does not need to be fully decomposed.

After the signal decomposition is achieved, the TFD  $\mathcal{W}(n, w)$ , may be constructed by taking the WVD [12] of the Gaussian functions represented in  $x'(n)$ :

$$\mathcal{W}(n, w) = \sum_{m=0}^{M-1} |a_m|^2 \mathcal{W}_{g_{\gamma_m}}(n, w) + \sum_{m=0}^{M-1} \sum_{\substack{l=0 \\ l \neq m}}^{M-1} a_m a_l \mathcal{W}_{g_{\gamma_m}, g_{\gamma_l}}(n, w), \quad (3.14)$$

where  $\mathcal{W}_{g_{\gamma_m}}(n, w)$  is the WVD of the Gaussian function  $g_{\gamma_m}(n)$ , and

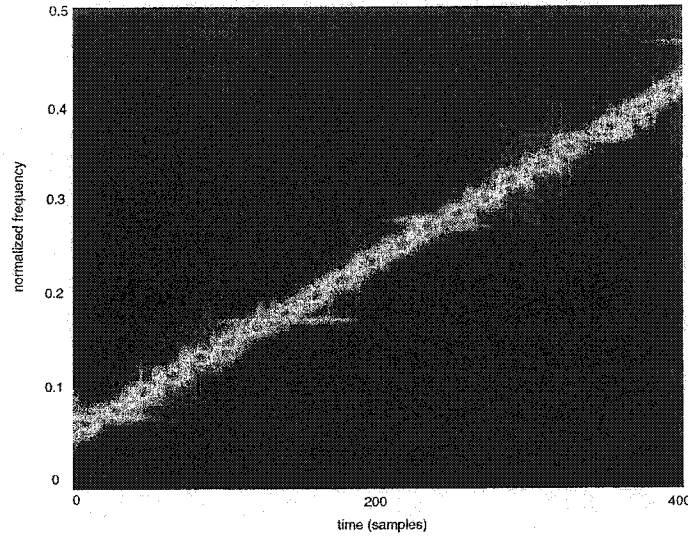
$$\mathcal{W}_{g_{\gamma_m}, g_{\gamma_l}}(n, w) = \sum_{k=-\infty}^{\infty} g_{\gamma_m}(n+k) g_{\gamma_l}^*(n-k) e^{-jwk}. \quad (3.15)$$

$\mathcal{W}_{g_{\gamma_m}}(n, w)$  takes discrete time and frequency values since  $\{\gamma_m\}$  is a set of integers. The second term in Equation (3.14) corresponds to the cross-terms of the WVD and should be rejected in order to obtain a cross-term free energy distribution of  $x'(n)$  in the TF plane [12]. Therefore, the MP TFD which we denote with the symbol  $\mathcal{W}'(n, w)$  is given as:

$$\mathcal{W}'(n, w) = \sum_{m=0}^{M-1} |a_m|^2 \mathcal{W}_{g_{\gamma_m}}(n, w). \quad (3.16)$$



The MP TFD is a cross-term free distribution with high resolution. Figure 3.8 shows the MP TFD of a linear chirp with increasing frequency.



**Figure 3.8:** MP TFD of a linear chirp with increasing frequency.

## 3.2 Detection of Directional Elements

As we demonstrated in the previous section, TFDs are techniques that are appropriate for monitoring and localization of nonstationary signals in the TF plane. We use the term *information extraction* to refer to the process of localization of nonstationary signals in the TF plane. The information extracted is the energy variation of the signal. Once the information is localized, we can treat the TFD as a gray level image and process it by using image processing techniques. One useful way of processing the TFD is to detect directional elements localized in the TF plane. In the context of this study, we will consider the detection of FM signals as directional elements of interest. These FM signals can be lines (linear chirps) or curves (quadratic chirps) in the 2D TF plane. FM signals are commonly found in synthetic aperture radars, multipath communication channels, whale sounds, helicopter sounds and sonars. Moreover, FM signals can be the result of hostile jamming designed to interfere with the signal of interest. This is a typical case in SS communications. Under this hostile interference scenario, we would like to detect and suppress the FM interference to enhance the

probability of detection of the transmitted signal. On the other hand, FM signals can be transmitted messages, such as a secret or public messages transmitted in a noisy environment. Under this alternate scenario, we attempt to recover the transmitted message by estimating the partially missing directional components of the signal. Therefore, we need a directional element detector that can detect time-varying energy values. The line detector that can satisfy our needs is a detector that uses the combination of Hough and Radon transforms proposed in [13]. This detector has been mathematically proven to be an optimal detector as it provides the maximum likelihood identification of a chirp signal [14]. The combined Hough and Radon transform (HRT) is an efficient tool to detect directional and time-varying energy components in the TF plane. In the following sections, we will first discuss the Hough transform and the Radon transform, and then continue to discuss the advantages of using the combined HRT for TFDs.

### 3.2.1 The Hough Transform

The Hough Transform (HT) is a pattern recognition method for calculating the number of points that satisfy a parametric constraint. The HT was developed by Paul Hough in 1962, and patented by IBM [15]. It is used in image processing applications such as object detection, texture analysis, character recognition, directional image analysis, and image compression. Although HT is mainly applied to straight line detection, it can also be applied to other curves that can be described by equations [16]. Let the parametric constraint be represented as

$$f(U, \Theta) = 0, \quad (3.17)$$

where  $U = (u_1, u_2, \dots, u_K)$  is a point in the space of possible features and  $\Theta = (\theta_1, \theta_2, \dots, \theta_L)$  is a point in the space of parameters. The parameter space is commonly referred to as Hough space. The constraint may represent a curve, a line or a surface depending on the interpretation of the feature point. Each point  $\Theta_0$  in the parameter space represents a constraint that is a particular instance of a curve, line or a surface.

The constraint may be mapped into the Hough space by evaluating

$$\{U : f(U, \Theta_0) = 0\}. \quad (3.18)$$

The parameter values consistent with the existence of a given feature point  $U_0$  are curves that the particular point may lie on, and are given by

$$\{\Theta : f(U_0, \Theta) = 0\}. \quad (3.19)$$

Given a number of feature points that satisfy a constraint specified by the parameter  $\Theta_0$ , the sets generated by Equation (3.19) for each feature will contain the point  $\Theta_0$ . Furthermore, Equation (3.19) may be viewed as a hypersurface in a continuous space of parameters. The curves of features satisfying a particular constraint will intersect at the common point  $\Theta_0$  in the parameter space.

The implementation of the HT algorithm is based on the observation that a particular curve defined by a parametric constraint can be identified if all the elements in a group of points, i.e., the points that lie on the curve, satisfy the same parametric constraint. The grouping process is carried out by using a search grid that covers the parameter space. Each cell in the search grid is associated with a counter. For a given feature point  $U_0$ , we solve Equation (3.19) to find the number of points that are on the parameter-space curve, and record it in the counter associated with that parameter point. A cell in the search grid that contains the intersection of many curves will represent a large number of points. At the completion of this process for every element in the search grid, the parameter point corresponding to the counter with the highest value will provide the estimate of the parametric constraint.

The size of the search grid has a significant effect on the computational time required for the implementation of the HT. The limitation of HT is that it can be applied to binary images only.

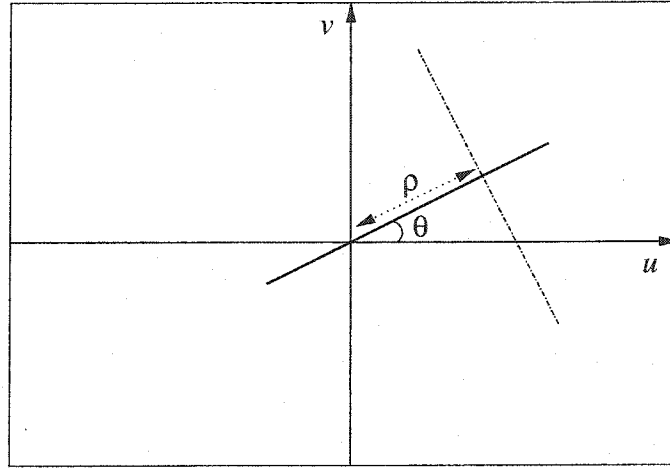
### 3.2.2 The Radon Transform

The Radon transform (RT) is a commonly used line detection technique in computer tomography [17]. The RT computes the projections of different angles of an image

(TFD) or 2D data distribution  $i_f(u, v)$  measured as line integrals along ray paths [17]:

$$\mathcal{R}(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i_f(u, v) \delta(\rho - (u \cos \theta + v \sin \theta)) du dv, \quad (3.20)$$

where  $\theta$  is the angle of the ray path of integration,  $\rho$  is the distance of the ray path from the center of the image, and  $\delta$  is the Dirac delta function. Equation (3.20) represents integration of  $i_f(u, v)$  along the line  $\rho = u \cos \theta + v \sin \theta$  as illustrated in Figure 3.9.  $\rho$  denotes the distance of the perpendicular bisect from the origin, and  $\theta$  denotes the angle spanned by the line. The RT adds up the pixel values in the given



**Figure 3.9:** Line detection using RT.

image along a straight line in a particular direction and at a specific displacement. The RT can be applied to both binary and gray-level images.

### 3.2.3 Combined Hough and Radon Transform

The Hough and the Radon transforms are individually not adequate to detect directional elements with varying energy levels. The underlying principle of the HT is that it is a process for counting the number of pixels that satisfy parametric constraints in a binary image. This property may result in misdetection of some energy varying components. The RT may be seen as a special case of the HT for straight line detection. While the RT can be applied to gray level images, it does not encompass all possible variations of the HT. Considering the advantages and disadvantages of each transform, we use the combined HRT as proposed in [13]. Using the combined

HRT, we can detect the pixels that form a parametric constraint in a gray-level image. These constraints can be straight lines or curves in the image of the TF plane. We will consider the TF plane as an image matrix which will replace the 2D data distribution  $i_f(u, v)$  in the formulation of the RT given in Equation (3.20). Let  $\mathcal{I}$  be an  $K \times N$  image matrix representation of the TF plane, where its elements  $\mathcal{I}(k, n)$  represent the gray-level intensities.  $K$  is the number of rows corresponding to the number of frequency slots, and  $N$  is the number of columns that correspond to the number of time slots in the TFD.  $K$  and  $N$  vary according to the resolution of the TFD, the time duration and the signal bandwidth.

The formulation of the HRT for discrete data sets is given as follows.

$$\mathcal{R}(\Theta) = \sum_{(k,n) \in \mathcal{A}(\Theta)} \mathcal{I}(k, n), \quad (3.21)$$

where

$$\mathcal{A}(\Theta) = \{ (k, n) : (k, n) \in [1, K] \times [1, N] : f(k, n, \Theta) = 0 \}, \quad (3.22)$$

and

$$f(k, n, \Theta) = 0 \quad (3.23)$$

is the parametric constraint equation in the image plane.

In general, the implementation of the HRT would require that we first determine a sufficiently fine search grid  $\Theta_s$ , for the parameter space, which will allow us to differentiate all parametric curves of the form given in Equation (3.23) within the resolution limitations of the image matrix. This search grid functions as a quantized parameter space.

In the implementation of the HRT, the transform value  $\mathcal{R}(\Theta_0)$  at some  $\Theta_0 \in \Theta_s$  contains the total energy in the pixels that satisfy the parametric constraint equation. Therefore, we can devise an HRT based system to detect directional elements defined by parametric equations: the peak values of the HRT will yield the most likely parameter values.

### An Example

In this example we will consider the parametric constraint equation corresponding to a straight line equation expressed in terms of the parameter vector  $\Theta = (\rho, \theta)$ :

$$\rho - (n \cos \theta + k \sin \theta) = 0. \quad (3.24)$$

With the above constraint equation, Equation (3.21) becomes equivalent to:

$$\mathcal{R}(\Theta) = \sum_{n=1}^N \sum_{k=1}^K \mathcal{I}(k, n) \delta(\rho - (n \cos \theta + k \sin \theta)). \quad (3.25)$$

Equation (3.25) represents the summation of  $\mathcal{I}(k, n)$  along the line  $\rho = n \cos \theta + k \sin \theta$ . Equivalently, Equation (3.25) becomes the special case of the HRT resulting the discretized version of the RT in Equation (3.20).

**Quantized parameter space  $\Theta_s$ :** In this example, we will consider the HRT based line detection process based on a  $176 \times 176$  image matrix, i.e.,  $K = N = 176$ . Let  $L = \max(K, N)/\sqrt{2}$ . The parameter space is given as:

$$\Theta = \{ (\theta, \rho) : \theta \in [-\pi/2, \pi/2], \rho \in [-L, L] \}. \quad (3.26)$$

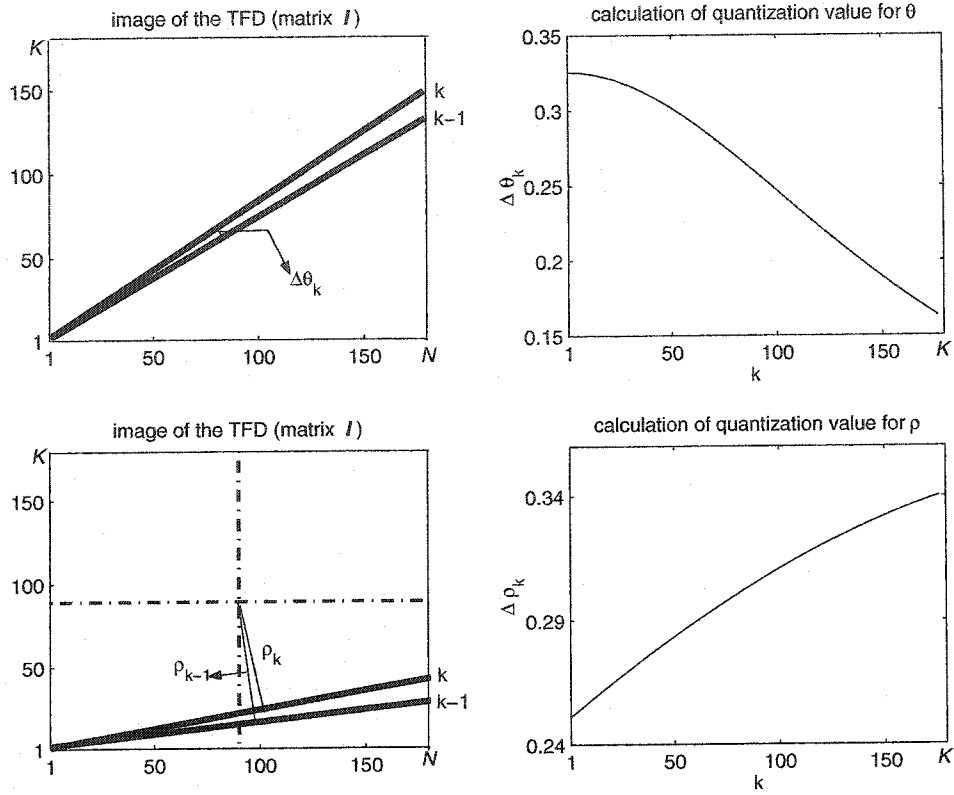
To calculate the quantized parameter space  $\Theta_s$ , we first quantize  $\theta$  and  $\rho$  to appropriate values in order to be able to evaluate all possible lines within the  $K \times N$  image plane. The quantization process for these parameters is as follows:

Let  $\theta_k$  be the angle occurring from the image of a line with its initial and end points corresponding to the first and the  $k$ th frequency slots, respectively. To find the minimum step size for  $\theta$ , we evaluate  $\Delta\theta_k = \theta_k - \theta_{k-1}$  as a function of  $k$  such that

$$\Delta\theta_k = \tan^{-1} \left( \frac{k}{N} \right) - \tan^{-1} \left( \frac{k-1}{N} \right), \quad (3.27)$$

where  $1 \leq k \leq K$ . The minimum value of  $\Delta\theta_k$  is achieved at  $\Delta\theta_K$  as shown in Figure 3.10. Let  $\rho_k$  be the distance of the line with initial and end points corresponding to the first and  $k$ th frequency slots, measured relative to the center of the image. We calculate  $\Delta\rho_k = \rho_k - \rho_{k-1}$  for each  $1 \leq k \leq K$  as

$$\Delta\rho_k = \left[ \left( \frac{K}{2} \cos \theta_k - \frac{N}{2} \sin \theta_k \right) - \left( \frac{K}{2} \cos \theta_{k-1} - \frac{N}{2} \sin \theta_{k-1} \right) \right]. \quad (3.28)$$

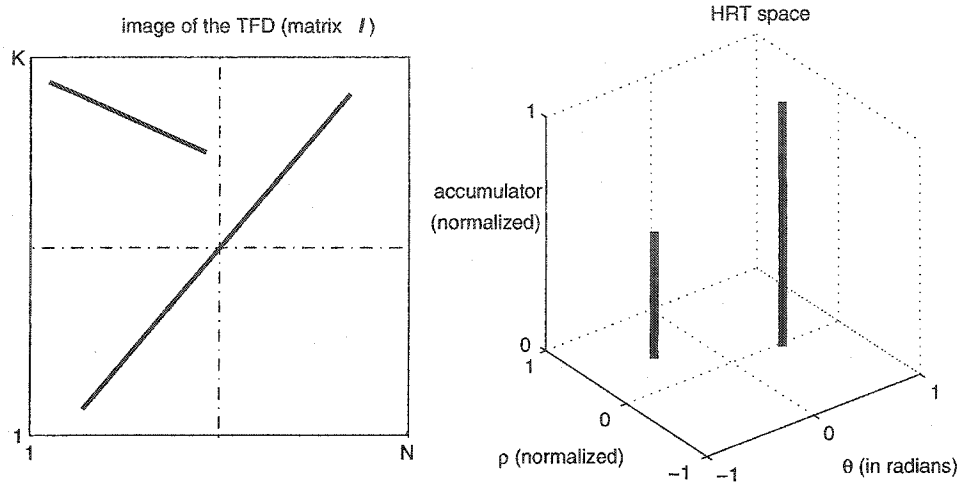


**Figure 3.10:** Quantization values for  $\theta$  and  $\rho$  for  $K = N = 176$ .

The minimum value of  $\Delta\rho_k$  is achieved at  $\Delta\rho_1$  as shown in Figure 3.10.  $\Delta\theta_K$  and  $\Delta\rho_1$  are the maximum possible values to be able to evaluate all possible lines in the  $K \times N$  image plane. We can use smaller quantization steps at the expense of more computation time.

The quantization step values  $\Delta\theta_K$  and  $\Delta\rho_1$  allow us to determine the quantized parameter space  $\Theta_s$ . We then proceed to compute  $\mathcal{R}(\Theta)$  for all elements  $\Theta \in \Theta_s$ . In particular, let us consider the two lines shown in Figure 3.11. The shorter line is represented in the HRT space by a peak with small amplitude ( $\rho > 0$  and  $\theta < 0$ ), and the longer line is represented by another peak with larger amplitude ( $\rho = 0$  and  $\theta > 0$ ). In the calculation of the HRT of the straight lines, it is assumed that the TF resolutions of the chirps are perfect.

The procedure for detecting linear chirp signals, i.e., straight lines in the TF plane, can be extended to include nonlinear FM signals [13]. The parameters  $\rho$  and



**Figure 3.11:** HRT space for two straight lines.

$\theta$  used in the detection of linear chirps can be replaced by any set of parameters that can satisfy a parametric equation. Some of the parametric equations that can be represented are a sinusoidal FM, a hyperbolic FM, or an FM with a second order equation. These equations can be represented respectively as  $f = m \sin(2\pi f_0 t + \theta) + b$ ,  $f = m/t + b$ , and  $f = mt^2 + b$ , where the  $(t, f)$  pairs can be solved for  $m$ , the amplitude,  $f_0$ , the number of cycles of the sinusoidal FM,  $\theta$ , the phase, and  $b$ , the frequency shift in the TF plane. The sinusoidal equations will have four parameters, whereas the hyperbolic and the second order equation will have two parameters each. Each parameter has to be quantized to appropriate values to evaluate the corresponding parametric equation. Solution of these parameters will yield the information about the presence or absence of these components in the TF plane. This information can be used for the recognition, detection, or removal of these components.



## Chapter 4

# Spread Spectrum Communications

**S**PREAD spectrum techniques represent an important class of coding methods frequently used in secure communication systems such as those used by the military. During the transmission of the SS signals, adversaries may deliberately attempt to jam the transmission. Jamming signals, or interferences are narrowband signals or wideband signals with narrowband instantaneous frequency elements, such as chirps. Signal processing techniques allow interference suppression and excision capabilities resulting in improved performance of the SS systems. Among signal processing based techniques, TFD-based methods can localize the interference both in time and frequency domains.

In this chapter, we propose a new method based on TFDs and the HRT to excise chirp-type interferences from SS signals. First, we present a brief review of interference suppression techniques based on signal processing. We provide the motivation for the proposed algorithm, then proceed to develop and state the full algorithm. We evaluate the performance of the new interference excision algorithm in terms of bit error rate and chip error rate measured as functions of jammer-to-signal-power ratio and signal-to-noise-power ratio values.

### 4.1 Review

Most interference suppression techniques are designed to deal with narrowband interferences [18, 19, 20]. Among time-domain approaches to narrowband interference

excision, the most notable methods include adaptive notch filtering and decision-directed adaptive filtering techniques [7].

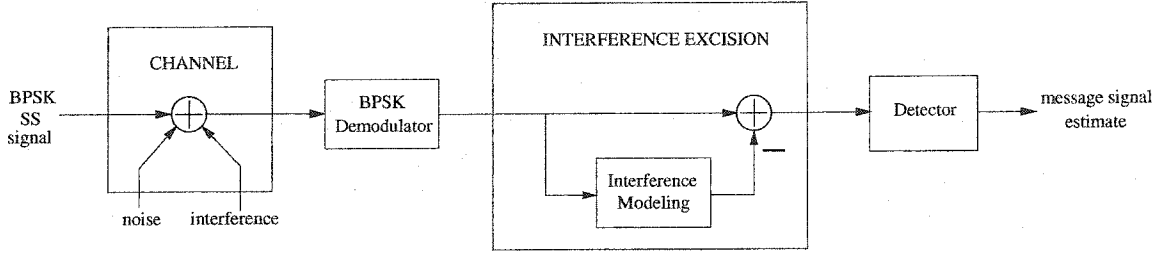
While SS systems can successfully reject narrowband interferences, their performance in rejecting wideband interferences is limited. In practical systems, adversaries are not likely to transmit high power wideband jamming signals due to the power limitations of the interference source. Additive Gaussian noise can be considered as the only realizable wideband interference, which is very challenging to predict and excise. Therefore, substantial amount of research has been conducted on wideband interferences with narrowband instantaneous frequency elements such as FM signals. Most of these methods focus on suppressing the interference using TFDs to localize the interference signals [8]. However, commonly used TFDs suffer from a trade-off between the TF resolution and cross-term suppression. In [21], Amin proposed a method based on the WVD of the signal, which represents the signal with precise TF localization; yet, the method is shown to suffer from cross-terms in the presence of multi-component interferences. In the extension of this work [22], the authors use the Wigner-Hough transform (WHT) to reduce the crossterms [22]; however, the system is shown to be sensitive to the signal model.

In [23] and [24], different window length STFTs are used to localize the interference. In [25], the authors use a signal decomposition algorithm consisting of a chirp-based dictionary to represent linear chirp interferences on the TF plane. The chirp interferences can be modeled with few coefficients and the proposed method performs well with linear chirp interferences. However, the generalization of the system to include quadratic, hyperbolic or sinusoidal FM interferences is not possible. In [26], the instantaneous frequency of the interference is recursively estimated using the discrete evolutionary and Hough transforms, and the interference is subtracted from the signal by using the singular value decomposition of the de-chirped signal. In [27], the authors propose an adaptive TF exciser that decides the domain of the excision by evaluating both the time and frequency properties. This system performs well in the case of narrowband interferences. There are also the TF projection fil-

tering techniques as proposed in [28, 29]. A common characteristic of most of the interference excision algorithms is the continuing presence of bit errors even after the interference is suppressed.

## 4.2 A New Interference Excision Algorithm

Let us consider a DSSS system as shown in Figure 4.1. In this system, the transmitter generates a SS signal which is in turn transmitted over a communications channel as a BPSK modulated signal. Additive channel noise as well as jamming signal act on the transmitted signal. At the receiver, the noise and interference corrupted signal is first demodulated. The “standard” SS receiver correlates the baseband SS signal



**Figure 4.1:** Block diagram of a DSSS system.

with the synchronized PN sequence, and the resulting signal is processed and input into a threshold detector to estimate the transmitted binary data sequence.

Let  $b_k = \pm 1$  be the  $k$ th message symbol transmitted in a DSSS system such that

$$\mathbf{w}_k = b_k \mathbf{p}_k, \quad (4.1)$$

where  $\mathbf{p}_k = [c_0, \dots, c_{L-1}]^T$  for  $\{k = 1, 2, \dots\}$  is a PN sequence with a chip length  $L$ , and  $\mathbf{w}_k$  is the SS signal<sup>1</sup>. The received signal  $\mathbf{r}_k$  at the output of the channel will consist of the SS signal  $\mathbf{w}_k$ , the additive white Gaussian noise term  $\mathbf{n}_k$ , and the interference  $\mathbf{i}_k$  such that

$$\mathbf{r}_k = \mathbf{w}_k + \mathbf{n}_k + \mathbf{i}_k. \quad (4.2)$$

<sup>1</sup>For notational convenience, we assume that the pulse energy of the SS signal,  $\mathcal{E}_L$  given in Equation (2.10) equals unity.

We will use the notation  $\mathbf{r}$  to refer to the received signal sequence:

$$\mathbf{r} = \{r_1(0), \dots, r_1(L-1), r_2(0), \dots\} \quad (4.3)$$

Similarly, we will use the notation  $\mathbf{w}$ ,  $\mathbf{n}$  and  $\mathbf{i}$  to refer to respectively, the complete SS signal, noise and interference sequences before they are separated into  $L$ -element vectors in the form  $\mathbf{w}_k$ ,  $\mathbf{n}_k$  and  $\mathbf{i}_k$ , for  $k = 1, 2, \dots$ .

Let  $\mathcal{E}_w = \mathbf{w}^T \mathbf{w}$ ,  $\mathcal{E}_n = \mathbf{n}^T \mathbf{n}$  and  $\mathcal{E}_i = \mathbf{i}^T \mathbf{i}$ . We define signal-to-noise-power ratio (SNR) and jammer-to-signal-power ratio (JSR) as:

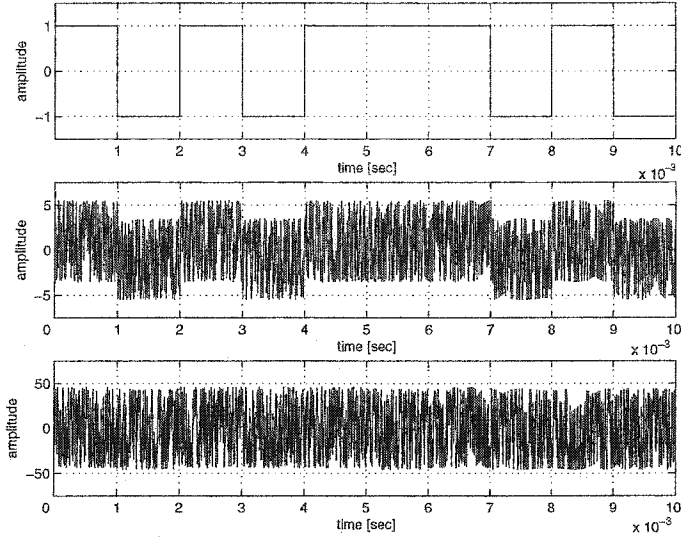
$$SNR = \mathcal{E}_w / \mathcal{E}_n, \quad (4.4)$$

$$JSR = \mathcal{E}_i / \mathcal{E}_w, \quad (4.5)$$

### Motivation

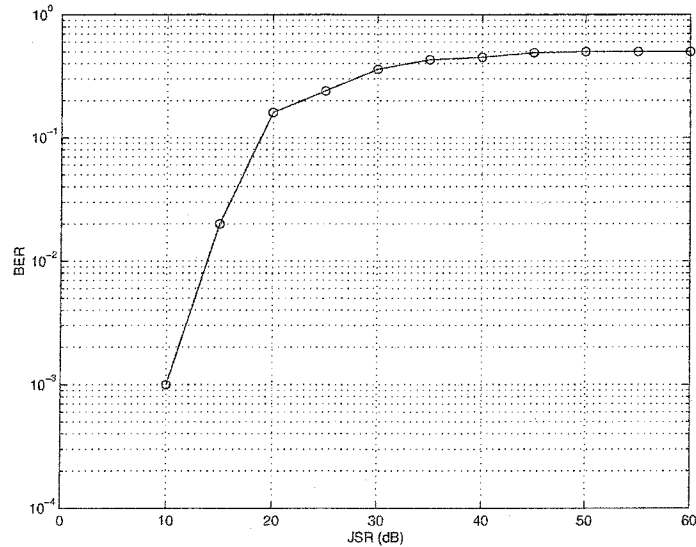
The results presented in Section 2.1 show that SS systems can inherently suppress interferences as a result of the spreading gain. In particular, when the power of the jamming signals that we assume are part of the received signal coming from the channel increase relative to the power of the SS signal, the performance of the detector and therefore of the SS system as a whole starts to degrade. Figure 4.2 shows three different waveforms as they appear in the receiver after the baseband SS signal has been despread by the PN sequence. These plots were generated with  $L = 128$ ,  $\mathcal{E}_n = 0$  and a linear chirp as the jamming signal at different JSR values. The assumption of zero noise power was made to better illustrate the effects of the jammer on the performance of the SS system. The middle plot in Figure 4.2 shows a waveform that will result in the correct estimation of the transmitted symbols. However, the bottom plot depicts a case when the jamming signal overwhelms the SS signal such that its interference suppression capacity is exceeded.

To further illustrate the effects of increasing jammer power, we simulated the channel output with  $\mathcal{E}_n = 0$ , and a linear chirp as the jamming signal which sweeps the entire frequency spectrum of  $\mathbf{w}$ . We changed the JSR values from 0 to 60 dB in 5 dB steps. To measure the performance of the SS signal, we despread the received signal



**Figure 4.2:** Despread waveforms at the receiver with a synchronized PN sequence. **Top:** reference with no jammer, **Middle:** JSR = 10 dB, **Bottom:** JSR = 30 dB.

$\mathbf{r} = \mathbf{w} + \mathbf{i}$ , with the PN sequence  $\mathbf{p}$ , integrate the resulting sequence and compare the result with a threshold to estimate the transmitted message symbols. The bit-error-rate (BER) results obtained from this simulation provides a measure of the built-in interference suppression, i.e., self-excision, capability of the SS system. Figure 4.3 shows the BER values at different JSR levels. The results presented in Figure 4.3 show



**Figure 4.3:** BER vs. JSR results for a self-excised SS system.

that the SS system was able to completely self-excise the interference for JSR < 10 dB,

as manifested with  $\text{BER} = 0$ . The resistance of the system to interference decreased with increasing JSR. For  $\text{JSR} > 40$  dB, we observed  $\text{BER} \approx 50\%$  indicating that the SS system cannot suppress any part of the interference.

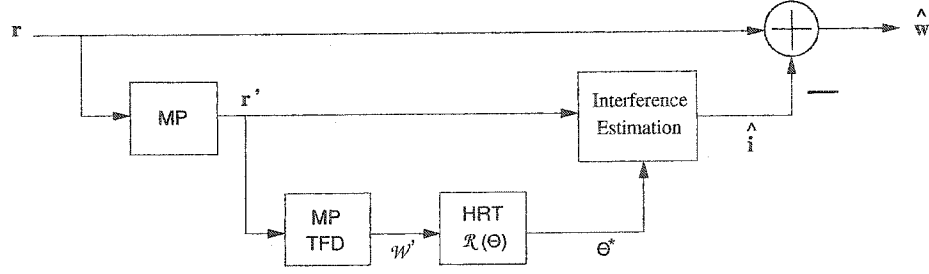
From these observations we conclude that pre-processing of the SS signals is an essential step in expanding the operating range of SS systems to high JSR environments. In particular, the pre-processing operations take the form of modeling the interference and excising from the SS signal before the despreading and detection steps.

### Overview of the Underlying Techniques

In this study, we consider the interference excision problem with single-component linear or quadratic FM signals, and multi-component FM signals—linear and quadratic FM signals interfering simultaneously. Considering the directional nature of FM signals in the TF plane, the first step of the proposed algorithm is the modeling of the interference by localizing the FM type interference signals in the TF plane.

Earlier interference excision methods based on TFDs suffer from a trade-off between the TF resolution and the TFD cross-terms [14, 30, 31]. Therefore, we propose a new excision method based on constructing a positive TFD of the received SS signal using an adaptive signal decomposition technique, the MP algorithm [12]. By decomposing a signal into its components, the interaction between components can be kept under control and possibly eliminated. The decomposition will allow the construction of a cross-term free TFD by combining the TFDs of the individual components generated by the decomposition. Also, by using Gaussian functions as bases for decomposition, we can achieve a high TF resolution, since the Gaussian functions satisfy the equality in the uncertainty principle explained in Section 3.1. We construct the TFD of the TF functions resulting from the MP, treat the TFD as an image, and detect the interfering signals using the HRT. We then reconstruct a model the interfering chirps using the TF functions and excise the reconstructed interference from the received signal. Figure 4.4 provides an overview of the proposed algorithm.

After the interference excision, the “interference free” SS signal is processed as be-



**Figure 4.4:** Interference excision.

fore by first correlating with the synchronized PN sequence, integrating the resulting sequence and estimating the transmitted data symbols using a threshold detector.

### The Algorithm

We assume that the information on the number and type of interference signals is available. In particular, we assume that the interference signals are linear or quadratic FM signals which can be present simultaneously. Let  $\tau \in \{\text{linear, quadratic}\}$  be the type of interference, and  $M_\tau$  be the number of interference signals of type  $\tau$ .

**Step 1:** The received signal  $\mathbf{r}$  is modeled as a linear combination of Gaussian functions using the MP algorithm given in Section 3.1.3. Let  $\mathbf{r}'$  be the model generated by MP algorithm as in

$$r'(n) = \sum_{m=0}^{M-1} a_m g_{\gamma_m}(n). \quad (4.6)$$

where  $g_{\gamma_m}(n)$  are the Gaussian TF functions given in Equation (3.10). The model order  $M$  is determined as the smallest positive integer which will make  $\mathbf{r}' = \{r'(0), r'(1), \dots\}$  satisfy the condition:

$$\sum_{n=0}^{N-1} |r(n) - r'(n)|^2 \leq N, \quad (4.7)$$

where  $N$  is the length of  $\mathbf{r}$ .

**Step 2:** Formulate the parameter set  $\mathcal{G}$  using the parameters of the Gaussian TF functions  $g_{\gamma_m}$ , such that

$$\mathcal{G} = \{(k_m, p_m), m = 0, \dots, M-1\}, \quad (4.8)$$

where  $k_m$  and  $p_m$  are the frequency and temporal placement parameters of the TF function  $g_{\gamma_m}$ , respectively.

**Step 3:** Compute the cross-term free TFD of  $\mathbf{r}'$  using the WVD:

$$\mathcal{W}'(n, w) = \sum_{m=0}^{M-1} |a_m|^2 \mathcal{W}_{g_{\gamma_m}}(n, w), \quad (4.9)$$

where  $\mathcal{W}_{g_{\gamma_m}}(n, w)$  is the WVD of the Gaussian function  $g_{\gamma_m}(n)$ .

**Step 4:** Let  $\mathcal{I}(k, n)$  be the  $K \times N$  image matrix representation of  $\mathcal{W}'(n, w)$ . For each interference type  $\tau$  known to be present in the received signal, determine the corresponding quantized parameter space  $\Theta_\tau$  and evaluate the HRT given in Section 3.2.3 and  $\mathcal{R}(\Theta_\tau)$  using Equations (3.21–3.23).

**Step 5:** For each interference type  $\tau$  known to be present in the received signal, determine the  $M_\tau$  parameters  $\{\Theta_\tau^{(1)}, \dots, \Theta_\tau^{(M_\tau)}\}$  from the quantized parameter space  $\Theta_\tau$  corresponding to the first  $M_\tau$  maxima of  $\mathcal{R}(\Theta_\tau)$ . Let

$$\Theta_\tau^* = \{\Theta_\tau^{(1)}, \dots, \Theta_\tau^{(M_\tau)}\}. \quad (4.10)$$

**Step 6:** For each interference type  $\tau$  known to be present in the received signal and for each  $\Theta_\tau^{(m)} \in \Theta_\tau^*$ , determine the index set  $\mathcal{L}_\tau^{(m)} \subset \{0, \dots, M-1\}$  defined as:

$$\mathcal{L}_\tau^{(m)} = \{l : l \in \{0, \dots, M-1\}, (k_l, p_l) \in \mathcal{G}, f_\tau(k_l, p_l, \Theta_\tau^{(m)} \pm \Delta\Theta) = 0\}, \quad (4.11)$$

where  $f_\tau(k, p, \Theta) = 0$  is the parametric constraint describing the interference of type  $\tau$  and  $\Delta\Theta$  is the empirically determined confidence measure.

**Step 7:** For each interference type  $\tau$  known to be present in the received signal and for each  $m \in \{1, \dots, M_\tau\}$ , construct the corresponding interference model as:

$$\hat{i}_\tau^{(m)}(n) = \sum_{l \in \mathcal{L}_\tau^{(m)}} a_l g_{\gamma_l}(n). \quad (4.12)$$

**Step 8:** Determine the interference excised SS signal by subtracting the interference models generated in **Step 7** from the received signal:

$$\hat{w}(n) = r(n) - \sum_{\tau} \sum_{m=1}^{M_\tau} \hat{i}_\tau^{(m)}(n). \quad (4.13)$$



## Detector

At the receiver, the interference excised signal  $\hat{\mathbf{w}}$  is first synchronized and correlated with the same spreading signal  $\mathbf{p}$ . Let  $\hat{\mathbf{w}}_k$  be the  $k$ th block of  $\hat{\mathbf{w}}$ . To estimate  $b_k$ , we use the PN sequence  $\mathbf{p}_k$  to despread  $\hat{\mathbf{w}}_k$ , and integrate the result to generate the test statistic  $\Lambda$ :

$$\begin{aligned}\Lambda &= \langle \hat{\mathbf{w}}_k, \mathbf{p}_k \rangle, \\ &= \mathbf{p}_k^T \hat{\mathbf{w}}_k, \\ &= \sum_{n=0}^{L-1} p(n) \hat{w}_k(n).\end{aligned}\tag{4.14}$$

Using the test statistic  $\Lambda$ , we estimate the message symbols as:

$$\hat{b}_k = \begin{cases} +1, & \text{if } \Lambda \geq 0, \\ -1, & \text{if } \Lambda < 0; \end{cases}\tag{4.15}$$

## 4.3 Simulation Results and Discussion

The simulation results presented in this section are based on a DSSS system with  $L = 128$  chips per message symbol  $b_k$ . The transmitted message contained 100 message symbols. We assumed that the channel was non-dispersive, and the received signal and the PN sequence were synchronized.

### Performance Measures

**Bit error rate (BER)** For the DSSS model used in this study, we process the received signal using the interference excision algorithm, and estimate the transmitted message symbols using the detector structure presented in Section 4.2. A comparison of the estimated message symbols  $\{\hat{b}_k\}$  with  $\{b_k\}$ , and expressing the number of erroneous estimates as a percentage of the total number of message symbols yields the bit error rate.

**Chip error rate** We define the chip error as

$$\text{sign}[p_k(n)\hat{w}_k(n)] \neq \text{sign}[p_k(n)w_k(n)],\tag{4.16}$$

for  $n \in \{0, \dots, L-1\}$  and  $k \in \{1, 2, \dots\}$ .

## BER Performance

To measure the performance of the DSSS system using the new interference excision algorithm developed in this chapter, we evaluated the BER results for the following three interference scenarios, where we assumed the presence of:

- a single-component linear chirp,
- a single-component quadratic chirp, and
- a multi-component interference with linear and quadratic chirps.

The interferences were measured with JSR values in the range of 0 to 50 dB at 10 dB steps. We assumed the SNR to be 10 dB in each case. We suppressed the interference before despreading, using the proposed interference excision algorithm. *We observed zero bit errors in all cases after the excision of single-component and multi-component interferences.* We repeated the same process for different SNR values in the range of -10 dB to 10 dB, and also recorded zero bit errors.

One of the main reasons for having zero BER in these simulation runs is the accurate TF representation of interferences, and the successful detection by the HRT. A similar observation was made by Bultan et.al. in [25], where they represent linear interferences with good TF localization using adaptive chirplet decomposition. However, they do not report any results on the excision of quadratic and/or multi-component interferences. Other TFD based methods reported bit errors for similar excision conditions [21, 22, 24].

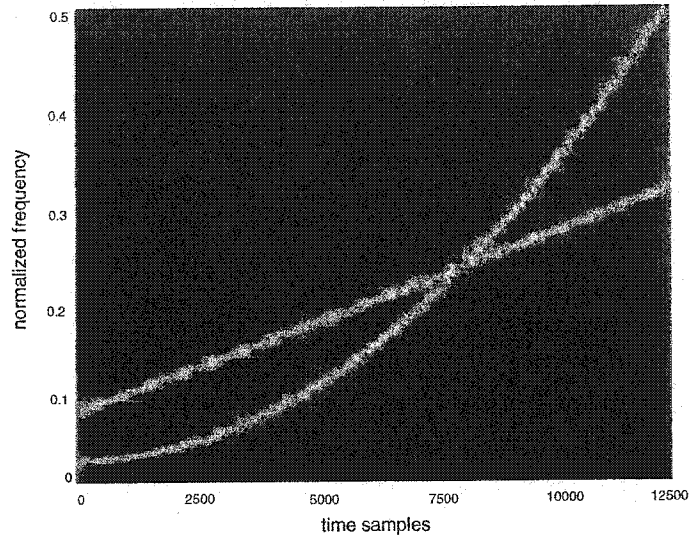
As an example we plot in Figure 4.5 the MP TFD<sup>2</sup> of a multi-component interference consisting of linear and quadratic chirps at JSR = 40 dB.

## Chip Error Rate Performance

We evaluated the DSSS system by calculating the percentage of chips received in error at various SNR values. Figures 4.6 and 4.7 show the simulation results for calculating the chip error rates for the JSR values 40 dB and 5 dB, respectively. The JSR value

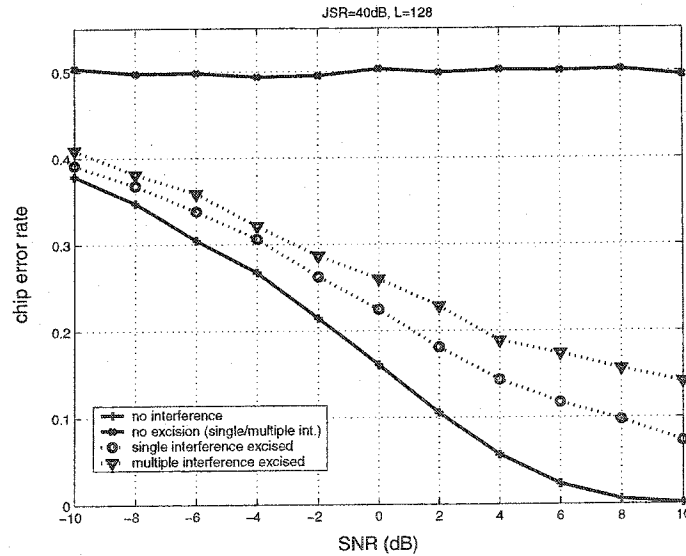
---

<sup>2</sup>Although the SS signal is also partly decomposed and represented on the TF plane, its lower energy compared to interferences makes it invisible on the MP TFD obtained by the WVD.



**Figure 4.5:** MP TFD of a multi-component interference consisting of linear and quadratic chirps.

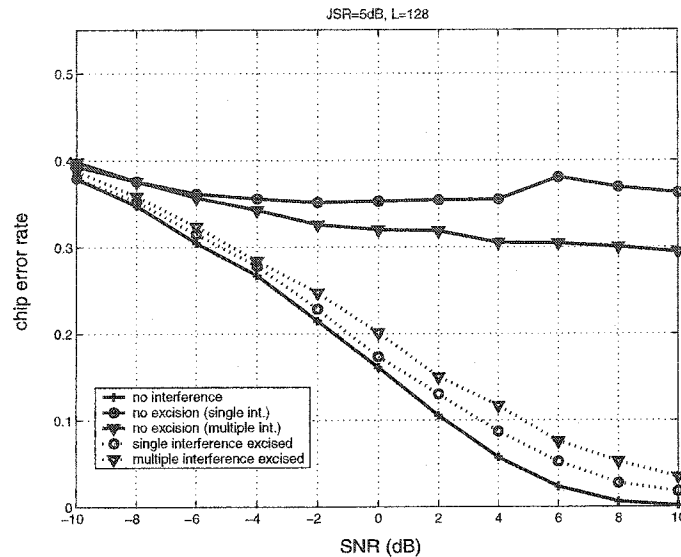
of 40 dB was used because at this JSR level with no interference excision, the system BER is approximately 50 per cent indicating that the system cannot suppress any part of the interference. At JSR = 5 dB, the system can suppress the interference partly



**Figure 4.6:** Chip error rate vs. SNR for JSR = 40 dB.

without interference excision prior to despreading. In systems proposed by other researchers [21], the excision of the low power interference degrades the performance of the system. In the case of the new interference excision algorithm we developed in

this chapter, the proposed system has substantially improved the chip error rate.

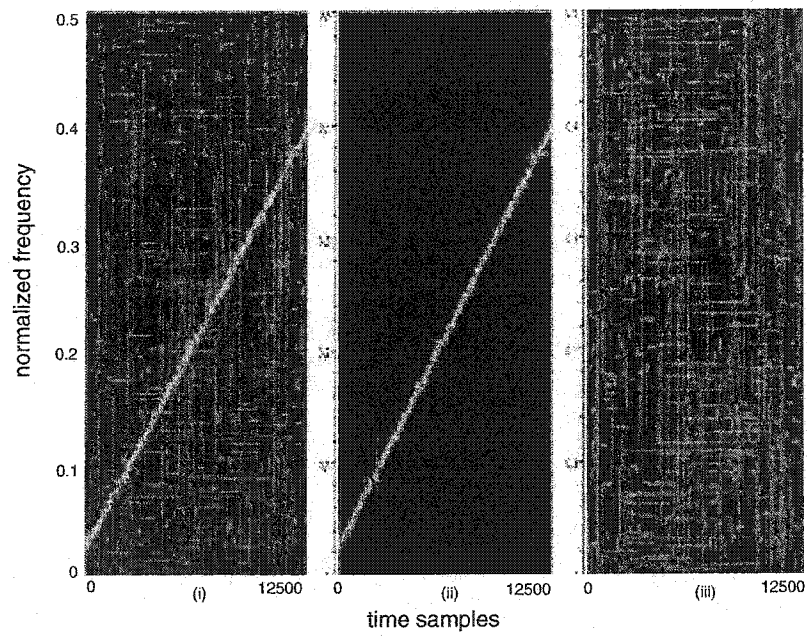


**Figure 4.7:** Chip error rate vs. SNR for JSR = 5 dB.

Attempts to excise multi-component interferences may introduce more noise to the system than the excision attempts for single-component interferences at the same power level. This observation follows from the final step of the interference excision algorithm. When the estimates of the interferences are excised from the SS signal, part of the SS signal in the vicinity of the interference localization will also be suppressed. Therefore, multi-component interferences are likely to introduce more residual noise.

For illustration purposes, in Figure 4.8 we provide the TFDs of the SS signal with a single interference (JSR = 5 dB) (left plot), the detected interference (middle plot), the interference excised SS signal (right plot).

The simulation results show that the proposed technique can be successfully used for excision of single-component and multiple-component chirp-like interferences using adaptive TFDs and the HRT. The new algorithm suppresses the interferences, while introducing an acceptable magnitude of noise, which can be overcome by the spreading gain. This results in zero bits in error after interference excision.



**Figure 4.8:** TFDs of **Left:** SS signal with a linear interference at JSR = 5 dB, **Middle:** estimate of the interference, **Right:** interference excised SS signal.

## Chapter 5

# Spread Spectrum Watermarking

IN recent years, the digital format has become the standard for the representation of multimedia content. Today's technology allows the copying and redistribution of multimedia content over the Internet at a very low or no cost. This has become a serious threat for multimedia content owners. Therefore, there is significant interest to protect copyright ownership of multimedia content (audio, image, and video). Watermarking is the process of embedding additional data into the host signal for identifying the copyright ownership. The embedded data characterizes the owner of the data and should be extracted to prove ownership. Besides copyright protection, watermarking may be used for data monitoring, fingerprinting, and observing content manipulations. All watermarking techniques should satisfy a set of requirements [32]. In particular, the embedded watermark should be:

- imperceptible,
- undetectable to prevent unauthorized removal,
- resistant to all signal manipulations, and
- extractable to prove ownership.

Before the proposed technique is made public, all the above requirements should be met.

Persons and/or institutions who propose watermarking algorithms are expected to publish a test criteria and the measured performance of their respective watermarking algorithms with respect to such criteria. As all watermark designers are free to develop

their own test criteria, an objective comparison of watermark algorithms with respect to a common reference criteria is difficult to achieve. In [33], the authors explain the need for a third party evaluation of existing schemes, and present the architecture of a public automated evaluation service for still images, audio and video. While there are some public image watermarking evaluation schemes, none currently exists for audio watermarking algorithms to the best of author's knowledge.

## 5.1 Review

The watermarking literature describes two classes of watermarking schemes. The first class of techniques is the *one-bit watermarks* [6], which detects only the presence of the watermark. In [34], the authors propose an algorithm that embeds a narrowband sequence into cepstral coefficients. In [35], a spread spectrum image watermarking algorithm is proposed. Bassia et.al. embed a chaotic sequence in the time domain [36]. In [37], the authors embed a spread spectrum sequence in the modulated complex lapped transform domain. All these algorithms use correlation-based detectors to determine the presence of the watermark. Another common characteristic of these algorithms is the increase of their robustness through the use of longer audio blocks and the repetition of embedded watermark sequences.

The second class of techniques not only detects but also extracts the embedded watermark message. In [38], the proposed watermark embedding algorithm employs a projection of an audio signal's frequency subbands onto a secret key. In [39], the authors present an algorithm which embeds watermarks in time-domain using the energy of the consecutive audio blocks. An algorithm embedding the watermark as noise into Fourier coefficients is proposed in [40]. Gang et.al. proposed an algorithm considering the effects of MP3 compression [41]. All these algorithms embed and extract multiple watermark bits. As a result of signal manipulations some message bits extracted by the detector may be in error, potentially resulting in the detection of the wrong watermark message.

## 5.2 Watermark Embedding

In this chapter, we develop an audio watermarking algorithm which detects the presence of the watermark, and extracts the embedded watermark message in the presence of bit errors. To achieve this objective, we embed linear chirps as watermark messages. Different chirp rates, i.e., slopes on the TF plane, represent watermark messages such that each slope corresponds to a different message. The narrowband watermark messages are first spread with a watermark key (a binary PN sequence). The resulting wideband noise is perceptually shaped and added to the original signal. The original and watermarked signals exhibit no perceptual differences. At the receiver a line detection algorithm based on the HRT detects the slope of the extracted chirp in the image of the TF plane, even at discontinuities corresponding to bit errors.

Our first attempt was to embed an arbitrary TF function as a watermark message into the TF functions obtained from the decomposition of the audio signal. Embedding a single TF function as the watermark message resulted in an algorithm that was not robust with respect to signal manipulations. We continued our work by embedding multiple TF functions, where the TF functions would satisfy a parametric constraint thus allowing detection by the HRT algorithm. There are two major concerns with embedding TF functions as watermarks: (i) the trade-off between imperceptibility and recovery, and (ii) the detectability of the parametric constraint on the TF plane. We then used a key-dependent modulation scheme, based on the multiplication of each bit of the watermark message with one sample of the PN sequence, where the length of the PN sequence was assumed to equal the number of bits in the watermark message. Such a scheme not having any spreading gain, significantly compromised the detectability of the embedded watermark. We proceeded with a SS modulation approach with a spreading gain resulting from *embedding one watermark bit in each audio block after spreading with a PN sequence of length  $L$  where  $L \gg 1$ .*

Figure 5.1 provides an overview of the proposed watermark embedding scheme. The watermark embedding scheme is a block process, where we first divide the audio signals into finite length blocks and then embed one watermark message bit into each



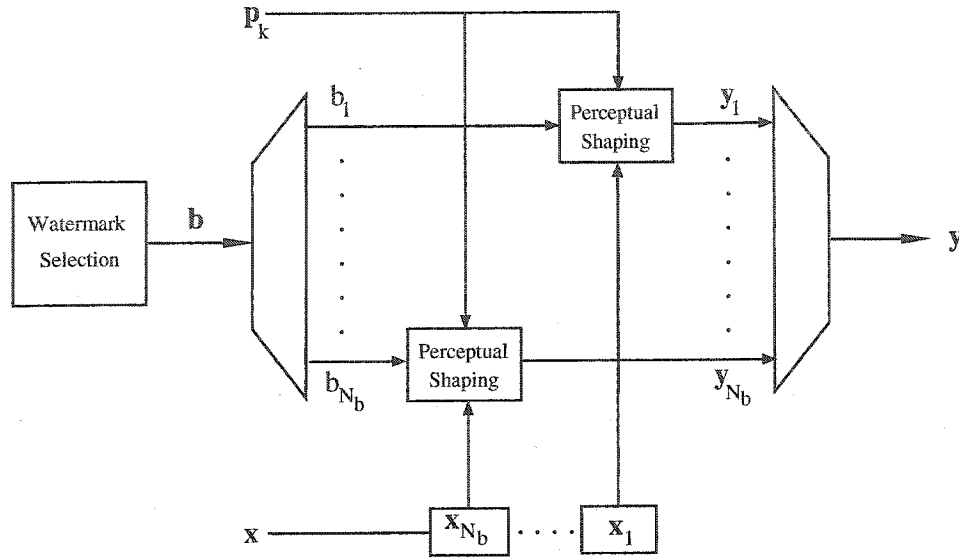


Figure 5.1: Watermark embedding scheme.

audio block. Figure 5.2 shows the main elements of the perceptual shaping process.

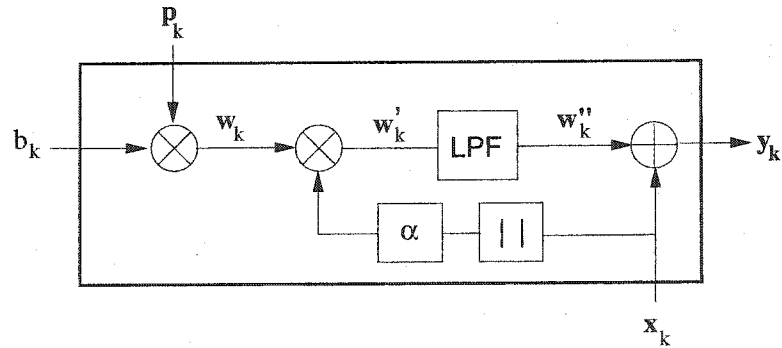


Figure 5.2: Perceptual shaping.

The main steps of the proposed algorithm are as follows.

1. Selection of the watermark message.
2. Perceptual shaping which includes spreading each bit of the watermark message with a binary PN sequence to generate a spread spectrum signal.
3. Embedding the watermark message.

We now proceed to provide the details of these steps which are designed to achieve the imperceptibility, security, robustness and detectability of the watermark message.

### Selection of the watermark message

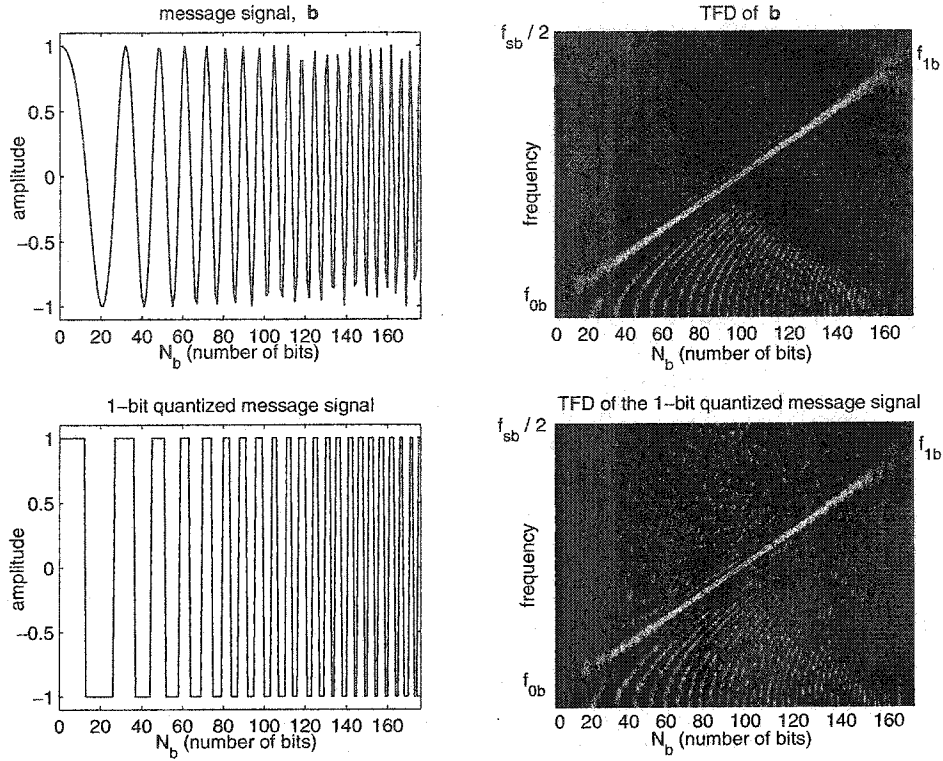
Let the watermark message  $\mathbf{b}$  be an  $N_b$ -sample long normalized linear chirp sequence sampled at the sampling frequency  $f_{sb}$ .  $f_{0b}$  and  $f_{1b}$  are the initial and final frequencies of  $\mathbf{b}$ , which are chosen to satisfy the Nyquist criteria with respect to the sampling frequency  $f_{sb}$ . Distinct pairs of initial and final frequencies form different watermark messages, each represented by a different slope in the TF plane.

The amplitudes of the samples in the normalized chirp sequence  $\mathbf{b}$  take continuous values in the interval  $[-1, 1]$ . Therefore, we first need to select a suitable finite-length representation for the sample amplitudes. A  $K$ -bit representation of sample amplitudes with  $K > 1$  will yield an accurate representation of the chirp sequence. However, such a representation format will also result in a  $K$ -fold reduction of the data payload<sup>1</sup> as a result of our earlier decision to embed only one watermark bit in each audio block in order to maximize the spreading gain. We choose a 1-bit representation of the sample amplitudes such that each sample in the chirp sequence equals  $\pm 1$ . If we evaluate the TF distributions of the continuous and the 1-bit quantized amplitude versions of  $\mathbf{b}$ , for both signals we observe a line in the TF plane corresponding to the linearly changing frequencies of the chirp signal. Figure 5.3 shows the TF distributions of the continuous and the 1-bit quantized amplitude versions of  $\mathbf{b}$  computed using the Wigner-Ville distribution.

While the main slope in the TF distributions of both chirp signals are the same, the quantization process introduces harmonics in the TF plane. Since the watermark information we want to embed is contained in the slope of the linear chirp in the TF plane, the harmonics can be eliminated by suitable manipulations of the TF distribution and/or by appropriate thresholding of the image representing the TF distribution at the receiver. In fact, the HRT based line detection technique we

---

<sup>1</sup>We define the data payload as the embedding capacity of the watermark embedding scheme.



**Figure 5.3:** Time-domain and TF representation of chirp signals. **Top:** continuous amplitude chirp, **Bottom:** 1-bit quantized chirp.

introduced in Section 3.2, identifies the parameters corresponding to the dominant line in the image of the TF plane, in effect suppressing all the spurious components resulting from harmonic terms. Therefore, the 1-bit representation format chosen for the chirp sequence  $\mathbf{b}$  will maximize the data payload.

### Perceptual shaping.

Let  $\mathbf{x} = \{x(0) x(1) \dots\}$  be the audio signal which we first divide into  $N_x$ -sample long blocks. We use the notation

$$\mathbf{x}_k = [x(kN_x) \dots x((k+1)N_x - 1)]^T \quad (5.1)$$

to represent the samples for the  $k$ th audio block. We will also use the notation  $x_k(n)$  to refer to the  $n$ th element of  $\mathbf{x}_k$ . The watermark message is the  $N_b$ -sample long normalized linear chirp sequence  $\mathbf{b}$  with its elements in 1-bit quantized format. Let  $b_k = \pm 1$  be the  $k$ th element of  $\mathbf{b}$  representing the watermark message bit to be

embedded into the  $k$ th audio block. Each  $b_k$  is spread with a binary PN sequence  $\mathbf{p}_k$  with a chip length  $N_x$  to generate the wideband noise vector

$$\mathbf{w}_k = b_k \mathbf{p}_k. \quad (5.2)$$

The spread spectrum signal  $\mathbf{w}_k$  appears as wideband noise and occupies the entire frequency spectrum spanned by the audio signal  $\mathbf{x}$ . We need to process  $\mathbf{w}_k$  to ensure its imperceptibility once it is embedded into the audio signal. The first step of this perceptual shaping process is amplitude scaling of  $\mathbf{w}_k$ . Let  $\mathbf{w}'_k$  be signal adaptive scaled wideband noise

$$\mathbf{w}'_k = \mathbf{D}_k \mathbf{w}_k, \quad (5.3)$$

where

$$\mathbf{D}_k = \alpha \mathbf{diag}[|x_k(1)|, \dots, |x_k(N_x)|], \quad (5.4)$$

and  $\alpha > 0$  is the embedding strength factor. An appropriately chosen  $\alpha$  and the fact that each element of  $\mathbf{w}'_k$  is scaled relative to the magnitude of the corresponding audio sample would ensure that the embedded watermark will be imperceptible.

Audio signals have most of their energies limited to low to middle frequencies. Conversely, typical signal manipulations including MP3 compression, lowpass filtering and resampling are likely to distort the higher frequency components of the audio signal. Therefore, if we embed  $\mathbf{w}'_k$  as the watermark message, signal manipulations will likely distort its high frequency components rendering the watermark message potentially irrecoverable. To increase the robustness of the embedded watermark message, we generate the frequency-limited noise by lowpass filtering  $\mathbf{w}'_k$  to limit its bandwidth to the frequency band  $[0, 0.05f_{sx}]$ , where  $f_{sx}$  is the sampling frequency of the audio signal. The upper frequency limit equals to 10 per cent of the maximum audio frequency, which represents part of the signal spectrum with significant energy content. Therefore, for the embedding strength factor  $\alpha$ , we can use larger values than 0.005 reported in [42]. The advantage of having larger  $\alpha$  is an improvement in the detectability of the embedded watermark message as a result of increased robustness with respect to signal manipulations. Let  $\{h_0, \dots, h_{L-1}\}$  be the impulse response

sequence of the  $L$ th order FIR lowpass filter with cutoff frequency  $0.05f_{sx}$ . We define  $\mathbf{w}_k''$  as the bandlimited, signal adaptive noise:

$$\mathbf{w}_k'' = \mathbf{H}\mathbf{w}_k', \quad (5.5)$$

where

$$\mathbf{H} = \begin{bmatrix} h_0 & & & & & \\ h_1 & h_0 & & & & \\ h_2 & h_1 & h_0 & & & \\ \vdots & \vdots & & \ddots & & \\ h_{L-1} & h_{L-2} & \dots & & h_0 & \\ 0 & h_{L-1} & \dots & & h_1 & h_0 \\ \vdots & \vdots & & & \vdots & \vdots & \ddots \\ 0 & 0 & & & & & h_0 \end{bmatrix} \quad (5.6)$$

is a lower triangular Toeplitz matrix representing the convolution of the finite length vector  $\mathbf{w}_k'$  with the FIR filter. Figure 5.4 shows the power spectral densities of various signal and noise components, and displays the effects of the perceptual shaping process.

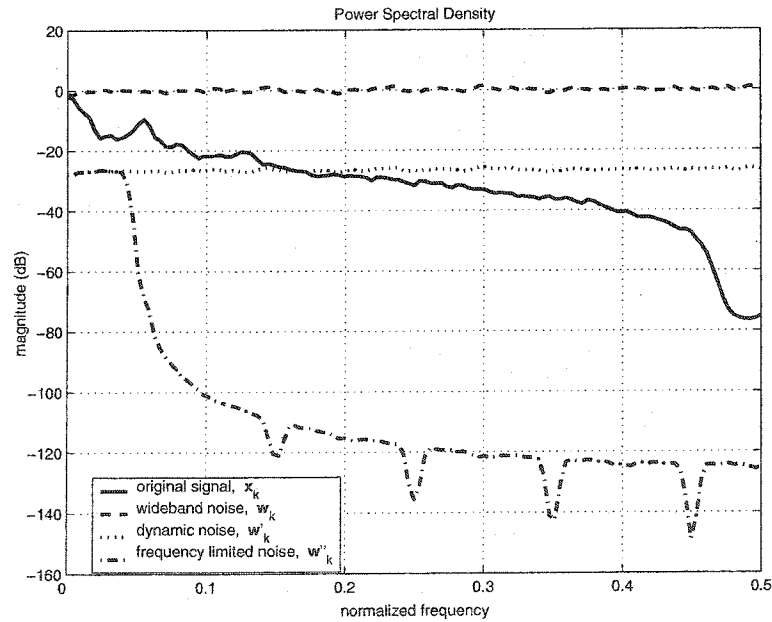


Figure 5.4: Perceptual shaping of watermark information.

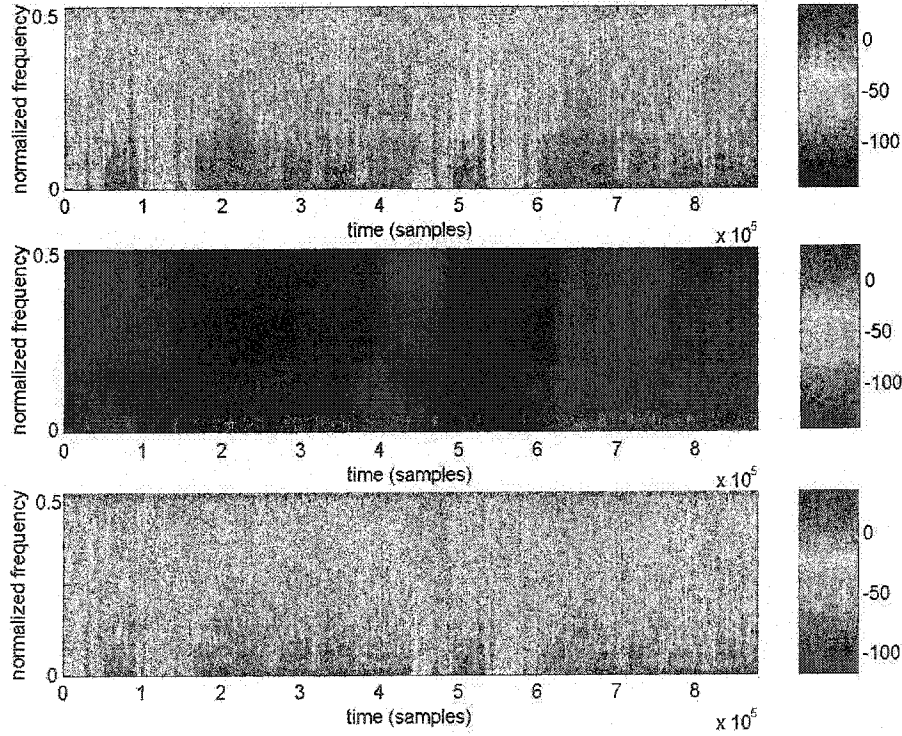
### Embedding the watermark message.

For each processing block, after we compute the perceptually shaped bandlimited noise sequence  $w_k''$ , we generate  $y_k$  as

$$y_k = x_k + w_k'', \quad (5.7)$$

which represents the  $k$ th audio block with the perceptually embedded  $k$ th bit of the watermark message  $\mathbf{b}$ . We repeat this three step process for each block until we embed all the bits in  $\mathbf{b}$ . Let  $\mathbf{y}$  be the augmented vector generated by concatenating the individual block outputs.

Figure 5.5 shows the spectrograms of the unprocessed audio signal  $\mathbf{x}$ , perceptually shaped watermark  $\mathbf{w}''$  and the watermarked signal  $\mathbf{y}$  for a 20 second long music file sampled at 44.1 kHz. A comparison of the spectrograms of  $\mathbf{x}$  and  $\mathbf{y}$  results in no



**Figure 5.5:** Spectrograms of **Top:** the original signal, **Middle:** perceptually shaped watermark, **Bottom:** the watermarked signal.

visually identifiable differences between the the two plots. Listening tests performed with this particular file and other music files confirm the imperceptibility of the watermarked audio files.

### 5.3 Watermark Extraction

The watermark embedded signal  $\mathbf{y}$  is transmitted over a communications channel to its destination. During the transmission, signal manipulations resulting from deliberate attempts to render the watermark irrecoverable will distort the signal. We will model the distortion introduced by such hostile signal manipulations as the additive channel noise  $\mathbf{n}$ . At the receiver, the watermark extraction process starts as a block processing operation since the individual bits of the watermark message  $\mathbf{b}$  are embedded in non-overlapping,  $N_x$ -sample long blocks. We will perform the following operations on the received signal to extract the embedded watermark.

**Bit Estimation:** Separate the received signal into non-overlapping,  $N_x$ -sample long blocks. Process each block of samples to detect the embedded watermark message bit.

**Post-processing of estimated bits for watermark extraction:** Post-process the estimated watermark message vector using a TFD and the HRT based line detection process for enhanced chirp detection as a form of error correction.

Figure 5.6 provides an overview of the watermark extraction process.

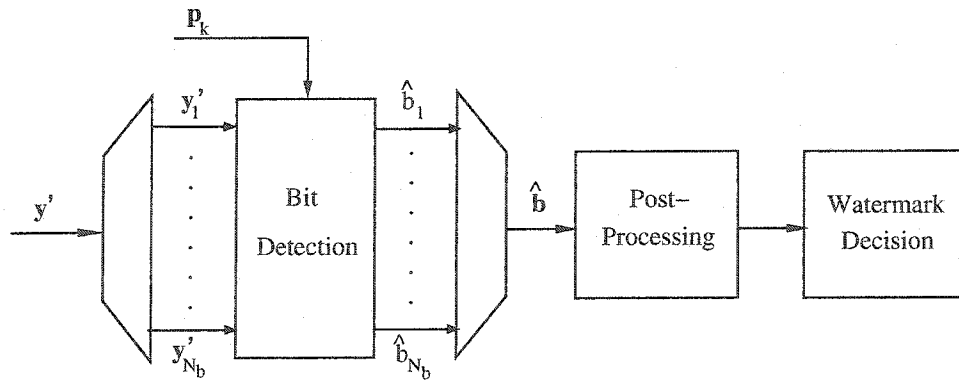


Figure 5.6: Watermark extraction.

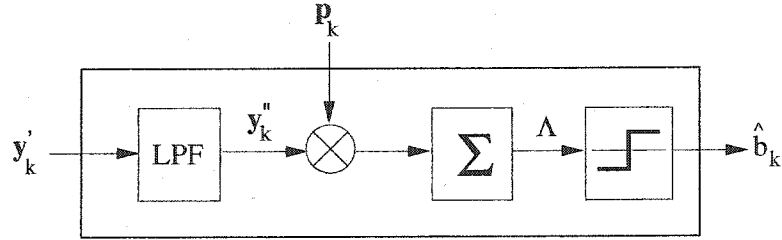
### Watermark Bit Estimation.

Let  $\mathbf{y}'_k$  be the  $k$ th block of the received signal at the channel output:

$$\mathbf{y}'_k = \mathbf{y}_k + \mathbf{n}_k, \quad (5.8)$$

where  $\mathbf{n}_k$  is the distortion component in the  $k$ th block resulting from hostile signal manipulations. We assume that  $\mathbf{n}_k$  is a zero-mean random vector uncorrelated with the PN sequence. In Section 5.4, we will further discuss the effects of the signal manipulations on the watermarked signal.

Since the watermark bit is embedded in the low frequency bands of transmitted signal, we extract the watermark bit by processing the low frequency bands of the received signal. Assuming perfect synchronization of the PN sequences used at the transmitter and the receiver, we reverse the embedding process [6] and construct the test statistic to estimate the watermark bit. Figure 5.7 shows the bit detection process. At the receiver we use the same  $L$ th order FIR lowpass filter with impulse



**Figure 5.7:** Watermark bit detection scheme.

response coefficients  $\{h_0, \dots, h_{L-1}\}$  as the transmitter. Let  $\mathbf{y}''_k$  be the lowpass filtered signal

$$\mathbf{y}''_k = \mathbf{H}\mathbf{y}'_k, \quad (5.9)$$

where  $\mathbf{H}$  defined in Equation (5.6) is the convolution matrix.  $\mathbf{y}''_k$  thus formed has its bandwidth limited to  $0.05f_{sx}$ , which is the frequency band where the signal containing the information about the watermark bit is embedded. We use the watermark key, i.e., the PN sequence  $\mathbf{p}_k$ , to despread  $\mathbf{y}''_k$  and integrate the resulting sequence to generate the test statistic  $\Lambda$ :

$$\Lambda = \langle \mathbf{y}''_k, \mathbf{p}_k \rangle. \quad (5.10)$$



Using Equations (5.5)–(5.7) we can express  $\Lambda$  as

$$\begin{aligned}
 \Lambda &= \langle \mathbf{y}_k'', \mathbf{p}_k \rangle \\
 &= \langle \mathbf{H}(\mathbf{y}_k + \mathbf{n}_k), \mathbf{p}_k \rangle \\
 &= \langle \mathbf{H}(\mathbf{x}_k + \mathbf{H}\mathbf{w}_k' + \mathbf{n}_k), \mathbf{p}_k \rangle \\
 &= \mathbf{p}_k^T \mathbf{H}\mathbf{x}_k + \mathbf{p}_k^T \mathbf{H}^2 \mathbf{w}_k' + \mathbf{p}_k^T \mathbf{H}\mathbf{n}_k.
 \end{aligned} \tag{5.11}$$

A properly chosen and normalized PN sequence  $\mathbf{p}_k$  will have the properties [43]:

$$\mathbf{E}[\mathbf{p}_k] = 0, \tag{5.12}$$

$$\mathbf{E}[\mathbf{p}_k \mathbf{p}_k^T] = \Sigma_p = \mathbf{I}, \tag{5.13}$$

where  $\mathbf{I}$  is the identity matrix. Furthermore,  $\mathbf{p}_k$  will be uncorrelated with  $\mathbf{x}_k$  and  $\mathbf{n}_k$ . In Equation (5.11), the first term represents the despreading of the lowpass filtered audio signal, and the third term represents the despreading of the lowpass filtered noise introduced by signal manipulations. Using the properties of the PN sequence we can show that the expected value of these two terms are:

$$\mathbf{E}[\mathbf{p}_k^T \mathbf{H}\mathbf{x}_k + \mathbf{p}_k^T \mathbf{H}\mathbf{n}_k] = 0. \tag{5.14}$$

We will use the property,

$$\mathbf{E}[\mathbf{p}_k^T \mathbf{Q} \mathbf{p}_k] = \text{trace}[\mathbf{Q} \Sigma_p], \tag{5.15}$$

where  $\mathbf{Q}$  is a square matrix that defines the quadratic form in the argument of the expectation operation in Equation (5.15), and the “trace” operator applied to an  $N \times N$  square matrix  $\mathbf{A}$  with elements  $a_{ij}$  is:

$$\text{trace}[\mathbf{A}] = \sum_{i=1}^N a_{ii}. \tag{5.16}$$

Using these results we can determine the expected value of the second term in the expansion of  $\Lambda$  in Equation (5.11):

$$\begin{aligned}
 \mathbf{E}[\mathbf{p}_k^T \mathbf{H}^2 \mathbf{w}_k'] &= \alpha b_k \mathbf{E}[\mathbf{p}_k^T \mathbf{H}^2 \mathbf{D}_k \mathbf{p}_k], \\
 &= \alpha b_k \text{trace}[\mathbf{H}^2 \mathbf{D}_k \Sigma_p], \\
 &= \alpha b_k \text{trace}[\mathbf{H}^2 \mathbf{D}_k], \\
 &= [\alpha h_0^2 \sum_{n=1}^{N_x} |x_k(n)|] b_k.
 \end{aligned} \tag{5.17}$$

Combining the results given in Equations (5.14) and (5.17) we obtain

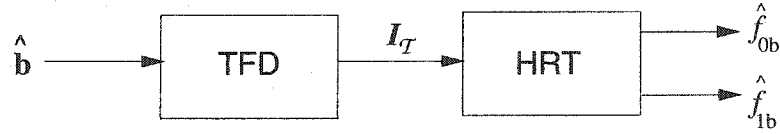
$$E[\Lambda] = [\alpha h_0^2 \sum_{n=1}^{N_x} |x_k(n)|] b_k. \quad (5.18)$$

We observe that the sign of  $E[\Lambda]$  is solely determined by  $b_k$  as the expression in the square brackets on the right side of Equation (5.18) is always positive. This observation justifies that we can estimate the  $k$ th watermark message bit using the decision rule:

$$\hat{b}_k = \begin{cases} +1, & \text{if } \Lambda \geq 0, \\ -1, & \text{if } \Lambda < 0. \end{cases} \quad (5.19)$$

### Post-processing of the estimated bits for watermark extraction.

We repeat the bit estimation process outlined above for each input block, until we have an estimate of all the transmitted message bits. While it is possible to combine the estimated message bits  $\{\hat{b}_k, k = 1, \dots, N_b\}$  to form an estimate of the chirp sequence, we can improve the performance of the watermark extraction algorithm by post-processing the estimated message bits using TFD and HRT. This second stage



**Figure 5.8:** Postprocessing of extracted bits.

of the watermark extraction algorithm, consisting of TFD and HRT, functions as an error-correcting technique and significantly increases the robustness of the proposed watermarking scheme. After all message bits are extracted, we first construct the TFD of the elements in  $\{\hat{b}_k, k = 1, \dots, N_b\}$ . The TFD can be generated by using one of the many possible TF representations such as STFT, WVD, wavelets, Gabor functions, which we introduced in Chapter 3. Each TFD is characterized by a different resolution of the TF plane [44]. Let

$$\mathcal{I}_{\mathcal{T}} = \mathcal{T}(\hat{b}_1, \dots, \hat{b}_{N_b}) \quad (5.20)$$

be the  $K \times N$  image matrix of the TF plane resulting from the TFD of the estimated message bits using the TF representation  $\mathcal{T}$ .  $K$  is the number of rows in  $\mathcal{I}_{\mathcal{T}}$  that

corresponds to the number of frequency slots, and  $N$  is the number of columns that corresponds to the number of time slots. The choice of the TF representation  $\mathcal{T}$ , determines the values of  $K$  and  $N$ , i.e., the frequency and time resolution of  $\mathcal{T}$ . Let  $\Omega_f$  be the frequency resolution of the TF representation given as:

$$\Omega_f = \frac{f_{sb}}{2K} \text{ Hz}, \quad (5.21)$$

where  $f_{sb}$  is the sampling rate used to generate the chirp sequence. Once we generate the image of the TF plane, HRT operates on  $\mathcal{I}_{\mathcal{T}}$  and functions as a line detection algorithm that searches for the presence of the line and estimates its parameters. There are three considerations that play an important role in the post-processing of the extracted bits:

**Quantization of the HRT parameter space:** Prior to the start of the search for chirp parameters, we need to select an appropriate search grid for the HRT parameter space. This is equivalent to determining the proper quantization level of the HRT parameters that will allow the potential detection of all possible chirp parameters.

**Selection of the TFD:** As the HRT will operate on  $\mathcal{I}_{\mathcal{T}}$ , the parameter search will be dependent on the TFD method used. For example, the STFT may have limited TF resolution, while the WVD may exhibit cross-terms, which may result in false detection of the watermark.

**Threshold level selection in the HRT space:** As a result of hostile signal manipulations, a number of message bits may be incorrectly estimated, which may result in uncertain localization of the linear chirp in the TF plane. In accordance with earlier watermarking research, we search for the global maximum in the HRT space to estimate the transmitted watermark parameters. However, the success of the HRT as a line detection algorithm will depend on the selection of a proper detection threshold which minimizes the probability of incorrect line parameter estimation. This is an area of further research.

In light of the consideration of the above points we design and implement a line detection algorithm based on the HRT as presented in Chapter 3. Let  $(\rho^*, \theta^*)$  be the line parameter values determined by the HRT. We construct the corresponding line  $\rho^* = n \cos \theta^* + k \sin \theta^*$  in  $\mathcal{I}_T$ . We identify  $(k_0, 1)$ ,  $1 \leq k_0 \leq K$  and  $(k_1, N)$ ,  $1 \leq k_1 \leq K$  as the coordinates of the start and end points of the line in the TF plane. Let  $\hat{f}_{0b}$  and  $\hat{f}_{1b}$  be the estimates of the initial and final frequencies of the linear chirp, which we calculate as:

$$\hat{f}_{0b} = (k_0 - 1) \Omega_f \quad (5.22)$$

$$\hat{f}_{1b} = (k_1 - 1) \Omega_f \quad (5.23)$$

Finally, using  $\hat{f}_{0b}$  and  $\hat{f}_{1b}$  we identify the embedded watermark message from an alphabet of  $K^2$  watermark messages.

## 5.4 Simulation Results and Discussion

We evaluated the proposed scheme using five different audio signals  $\{S_1, \dots, S_5\}$  sampled at  $f_{sx} = 44.1$  kHz. These sample audio files represent respectively rock, classical, harp, piano, and pop music. The sampling frequency  $f_{sb}$  of the watermark messages equals 1 kHz. Therefore, the initial and final frequencies,  $f_{0b}$  and  $f_{1b}$  of the linear chirps representing all watermark messages are constrained to [0–500] Hz. We embedded these messages into audio signals of 40 second duration for a chip length of 10,000 samples per message bit, and into audio signals of 20 second duration for a chip length of 5,000 samples per message bit. Having a data rate of 4.41 bits per second (bps) with the 40 second duration audio signals and 8.82 bps with the 20 second long audio signals, we embedded 176 message bits into each audio signal.

During the embedding process, we limited the maximum frequency of the lowpass noise to 10 per cent of the maximum audio frequency, i.e., 2.205 kHz. We used the empirically determined value of 0.3 for the embedding strength parameter  $\alpha$ , which ensured that the embedded watermark signal remained imperceptible with all the tested audio signals. This value of  $\alpha$  is different than what is used in [42] because we embed a lowpass noise rather than a wideband noise.

To measure the robustness of the watermarking algorithm, we performed the following signal manipulation tests, which represent commonly used signal processing techniques.

- $T_0$ : no signal manipulation.
- $T_1$ : MP3 compression with output bit rate at 128 kbps,
- $T_2$ : MP3 compression with output bit rate at 80 kbps,
- $T_3$ : lowpass filtering to 4 kHz,
- $T_4$ : resampling at different sampling rates (22.05 kHz and 11.025 kHz),
- $T_5$ : amplitude scaling,
- $T_6$ : inversion of audio samples,
- $T_7$ : addition of delayed signal,
- $T_8$ : additive noise,
- $T_9$ : embedding multiple (two) watermarks.

As  $T_0$  corresponds to watermark detection without signal manipulation, it will serve as a reference. Tests  $T_1$ – $T_4$  affect the frequency spectrum of the signal,  $T_5$ – $T_6$  measure the signal dependency of the watermark, and tests  $T_7$ – $T_9$  introduce interferences such as additive noise.

### Watermark Bit Detection: Results and Discussion

For each test, we process the watermark embedded signal  $y$  to generate the received signal  $y'$  resulting from the signal manipulation operations described in that particular test. We apply the watermark extraction algorithm described in Section 5.3 on the received signal  $y'$ . In all the tests performed, we assumed that the PN sequences used in the watermark embedding and extraction processes are synchronized.

Tables 5.1 and 5.2 show the bit error rate (BER) results expressed as a percentage of the total number of message bits for the two chip lengths and for each signal manipulation operation. From the results presented in these tables, we observe that the filtering operation  $T_3$  introduced the highest BER among all signal manipulations, whereas the amplitude scaling ( $T_5$ ) and the sample inversion ( $T_6$ ) manipulations did not introduce any additional bit errors relative to the no signal manipulation case

Audio Sample	Robustness Test						
	$T_0/T_5/T_6$	$T_1$	$T_2$	$T_3$	$T_4$	$T_7$	$T_8/T_9$
$S_1$	1.14	1.14	1.14	3.42	3.98	1.14	2.27
$S_2$	0.57	0.57	0.57	3.42	3.42	0.57	2.84
$S_3$	0.00	0.00	0.00	1.14	2.27	0.00	1.70
$S_4$	0.57	1.70	1.70	5.68	3.98	1.14	2.27
$S_5$	0.00	0.00	0.00	1.70	1.14	0.57	1.14

**Table 5.1:** Bit error rate (in percentage) for  $N_x = 10,000$ .

Audio Sample	Robustness Test						
	$T_0/T_5/T_6$	$T_1$	$T_2$	$T_3$	$T_4$	$T_7$	$T_8/T_9$
$S_1$	3.98	4.55	5.68	9.66	5.68	4.55	5.11
$S_2$	2.27	2.27	2.27	9.09	5.68	2.84	3.42
$S_3$	4.55	4.55	5.11	11.36	6.25	5.68	6.25
$S_4$	3.98	3.98	3.98	10.80	4.55	5.11	5.68
$S_5$	3.98	3.98	4.55	9.66	5.11	5.68	5.68

**Table 5.2:** Bit error rate (in percentage) for  $N_x = 5,000$ .

( $T_0$ ). This is due to the embedded information being signal dependent.

In all tests, the BER increased with decreasing chip length  $N_x$  from 10,000 to 5,000. This is due to the decrease in the spreading gain of the SS signal. Using chip lengths greater than 10,000 will lead to a more robust system at the expense of reduced data payload.

In Equation (5.8), we introduced the additive noise term  $\mathbf{n}_k$  as the distortion component in the  $k$ th block to represent hostile signal manipulations. We assumed that  $\mathbf{n}_k$  was a zero-mean random vector with a symmetric distribution. To justify this assumption, we evaluated the signal differences after each manipulation. All the manipulations except  $T_5$  and  $T_6$  resulted in additive distortion terms with symmetric, zero-mean distributions, which in many instances were comparable to a Laplacian distribution. The signal manipulations in  $T_5$  and  $T_6$  are amplitude scaling and inversion of audio samples. These signal processing operations function directly at the sample level and are therefore not suitable candidates for modeling as an additive noise term. However, due to the signal adaptive watermark embedding procedure, such

signal manipulations have very limited effect on the recoverability of the watermark, a fact that is verified by identical bit error rates in tests  $T_0$ ,  $T_5$  and  $T_6$ .

### Watermark Extraction: Results and Discussion.

After bit extraction, we localized the extracted bits on the TF plane using a spectrogram generated by a fixed window length STFT. We used a Hanning window of length 32 with 30 samples overlap in each block. As a result of the limited resolution of the spectrogram, we formed a  $17 \times 73$  image matrix  $\mathcal{I}_T$ . The initial and final frequencies of the linear chirp could take one of the 17 frequency values. This approach allowed us to use a message alphabet with  $\{17^2\}$  possible watermark messages. Although some bits were received in error (even in the case of no signal manipulation), the HRT correctly detected the slope of the chirp functions in the image of the TF plane and successfully extracted the embedded watermark messages thus providing error-correction capability.

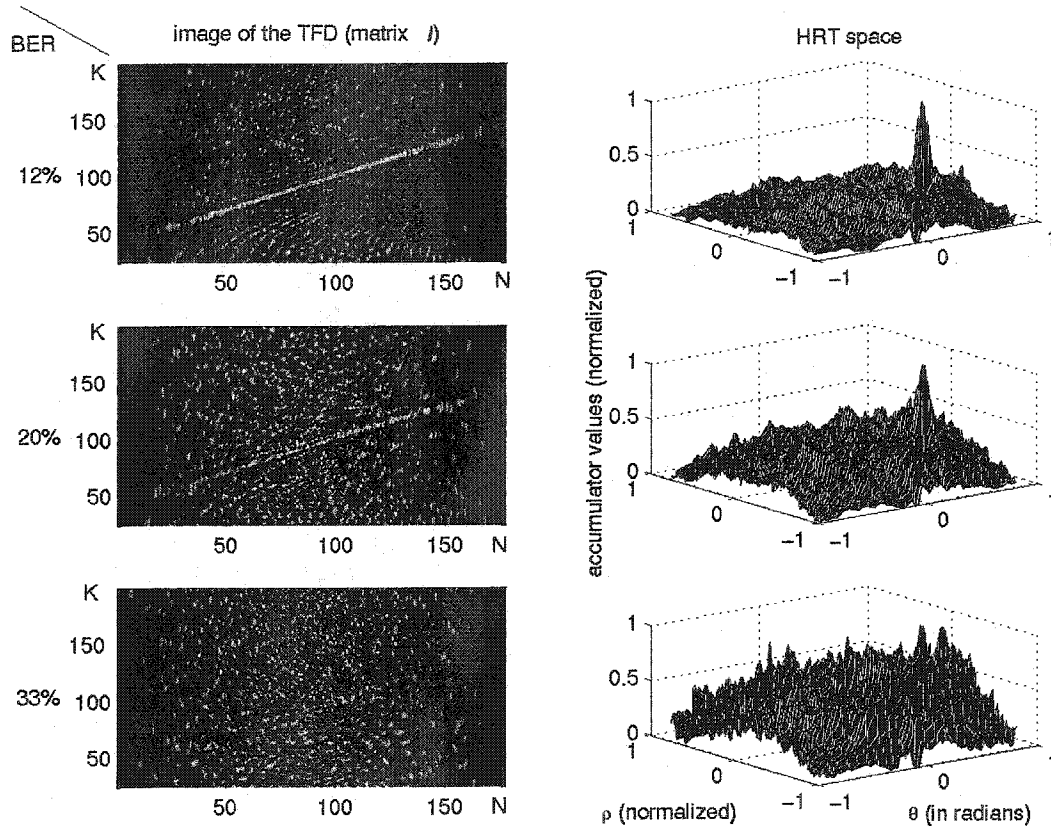
As a second evaluation, we post-processed the same extracted bits using WVD which gives high TF resolution [8]. We formed the  $176 \times 176$  image matrix  $\mathcal{I}_T$ . Therefore, we had a message alphabet of  $\{176^2\}$  messages. Due to the size of  $\mathcal{I}_T$ ,  $\theta$  was bounded<sup>2</sup> by  $\pm\pi/4$  (i.e.,  $\theta_{max} = \theta_{176} = \pi/4$ ) and  $\rho$  by  $\pm 88$  (i.e.,  $\rho_{max} = \rho_1 = 88$ ). We were able to successfully detect and extract all the watermark messages.

In all the robustness tests performed, the maximum BER was 11.36 per cent. The HRT was able to detect the watermark message and extract the parameters correctly even in the worst case scenario. After the extracted chirp is localized using a TFD, the HRT calculated the line parameters ( $\theta$  and  $\rho$ ). The global maximum in the HRT space is used to calculate the initial and final frequencies used in the chirp. The initial and final frequencies resulting from the spectrogram and WVD were found to be different. This is a result of the differences between frequency resolutions and the allocated frequency slots of each TFD. WVD based method extracted the  $f_{0b}$  and  $f_{1b}$  parameters more accurately than the spectrogram.

---

<sup>2</sup>Since the initial and final frequencies of the linear chirp are constrained to the frequency range  $[0, f_{sb}/2]$ , the  $\pm\pi/4$  bounds on  $\theta$  delineate the part of the HRT parameter space where the linear chirps are allowed.

As the next step, we decreased the value of the embedding strength  $\alpha$  and evaluated watermark message detectability with increasing BER. Since  $\alpha$  is the same for each block, bit errors occurred randomly. Using WVD, we were able to detect the embedded chirp up to 20 per cent BER. Figure 5.9 illustrates how the WVD and the HRT space change with increasing BER. The prominence of the global maximum in the HRT space provides an indication of watermark detectability. We were able to



**Figure 5.9:** Changes of the WVD and the HRT with increasing BER.

detect the presence of watermark using windowing based TFDs (spectrogram) up to 30 per cent BER. However, the TF resolution was not sufficient to calculate the initial and final frequencies accurately. To test the robustness of the HRT with respect to large discontinuities, we assumed that the bits in first 20 seconds of a 40 second long watermarked audio signal are extracted in error with 50 per cent probability, and yet, the HRT successfully detected the slope of the linear chirp at the large discontinuity, extracted the message and determined  $f_{0b}$  and  $f_{1b}$ .



# Chapter 6

## Conclusions and Future Research

### 6.1 Conclusions

#### SS Communications

In this study, we evaluated the performance of the new interference excision algorithm presented in Chapter 4 in terms of the BER and chip error rates. The most striking observation resulting from the simulation studies is that *there were no bit errors after the excision of single and multi-component interferences at all JSR levels tested, i.e.  $JSR \leq 50\text{ dB}$* . Under similar test conditions, the algorithms developed in earlier studies reported bit errors with the notable exception of [25]. This highly desirable characteristic is the result of three factors.

1. The model of the interference uses Gaussian functions, which provide optimal TF resolution within the limits of the uncertainty principle.
2. The MP TFD uses WVD, which also provides a high TF resolution. The modeling of the interferences as a linear combination of basis functions eliminated the cross-terms in the construction of the TFD for multi-component interferences. Lack of cross-terms prevents undesired peaks in the HRT space, which may lead to incorrect parameter estimates.
3. The HRT algorithm acts as an adaptive thresholding mechanism successfully determining the functions that model the interference.

We can draw the following conclusions from the analysis of the interference excision algorithm and the simulation results.

- SS communication systems can reject the interferences if the jamming margin exceeds the interference power. When the interference power increases beyond the jamming margin, the performance of the SS decreases. If the jammer power is high, the interference should be excised from the SS signal prior to despreading.
- In simulation runs when we recorded no bit errors, we assessed the performance of the excision algorithm in terms of chip errors. The performance of the system is improved by reducing the number of chips received in error at all SNR levels. Some algorithms have reported performance degradation after the excision of low power interferences [21], while the interference excision algorithm presented in this study does not suffer from such side-effects.
- The excision of a multi-component interference introduces more noise than the excision of a single-component interference at the same power level. When the estimates of the interferences are excised from the SS signal, part of the SS signal in the vicinity of the interference localization will also be suppressed. Therefore, multi-component interferences are likely to introduce more residual noise.
- The proposed algorithm can excise any interference that can be modeled using a parametric equation. This is a consequence of the HRT being able to detect any directional element defined by parametric constraints. Other algorithms focus only on linear, or sinusoidal interferences, and break down, if there are nonlinear and/or multi-component interferences.
- The proposed algorithm does not operate in real-time due to the decomposition stage. However, this condition does not introduce undue limitations, since the received signal will be time-limited.

- We assumed that the received SS signal and the PN sequence are synchronized. All other interference excision schemes make the same assumption as the focus of these studies is interference modeling and suppression. However, the implementation of any of these algorithms must address the important synchronization issue.

### Audio Watermarking

In this study, we developed a novel SS audio watermarking scheme, which includes a post-processing stage based on TF analysis. We evaluated the performance of the new watermarking scheme in terms of the BER and the detection of the embedded watermark. *We were able to detect the watermark correctly after common signal processing operations*, demonstrating the robustness of the algorithm with respect to hostile signal manipulations. Most SS watermark embedding schemes detect only the presence of the watermark. Although currently there is no prevailing watermarking standard, it is expected that successful watermarking algorithms should have moderate data payload and should be able to identify the embedded watermark correctly. The watermark embedding algorithm we developed in Chapter 5 satisfies both conditions.

We state the following conclusions resulting from the analysis of the watermarking algorithm, simulation results and the observations we made in developing the algorithm.

- In all robustness tests performed, the worst case scenario resulted in the maximum BER of 11.36 per cent. The simulation results show that the HRT applied to a TFD generated by WVD can successfully detect the watermark correctly at bit error rates up to 20 per cent. The HRT applied to a TFD generated by spectrogram can detect the *presence* of the watermark at bit error rates up to 30 per cent.
- Using the HRT, we were able to suppress the spurious low-power components resulting from the quantized representation of the chirp. We observe that WVD

generated TFDs achieve high resolution and therefore represent the best TFD choice for the detection of watermarks.

- The BER depends on the embedding strength, the lowpass filter used, and the length of the PN sequence. Increasing the embedding strength will result in a more robust system at the expense of imperceptibility. Increasing the length of the PN sequence will also result in a more robust system without adversely affecting the imperceptibility of the watermark. However, a longer PN sequence will reduce the data payload.
- In all the time-domain watermark embedding schemes, the synchronization problem that occurs as a result of cropping attacks is an outstanding issue. This problem may be addressed by redundancy coding or embedding synchronization bits. In our simulations studies, we assumed that the watermarked signal and the PN sequence were synchronized, since our main focus was the introduction of TF analysis as a post-processing tool.
- Embedding chirps without modulating by SS techniques is an example of a fragile watermarking scheme [6]. The embedded watermark can be seen as a line in the TF domain of the watermarked signals. This can be used to observe the signal manipulations made by adversaries. It can also be used for the synchronization of the watermarked signal and the PN sequence. Knowing the exact location of the fragile watermark will allow the user to understand if any cropping has occurred. One important issue is that embedding chirps directly into the signal may result in degrading the quality of the signals. On the other hand, low amplitude chirps may not be localized on the TF plane.
- Time-domain embedding of the watermarking scheme results in a fast algorithm suitable for real-time implementation.

## 6.2 Future Research

### SS Communications

In the case of the interference excision algorithm of Chapter 4, there are potential areas of development that would greatly improve the implementation speed of the algorithm.

- The current interference excision algorithm decomposes the received signal until most of the energy of the interference is modeled. If the power of the interference is large, then the modeling algorithm may require a long time to converge. Therefore, we may consider an alternate approach, where the decomposition process continues until the unexcised interference power falls below the jamming margin of the system. Such a system is expected to operate at a performance level comparable to a full interference excised system.
- In the current algorithm, the interferences are modeled using Gaussian functions. An alternate interference modeling approach would be to use a set of chirp-based functions. Such an approach is expected to model chirp-type interferences more accurately with fewer coefficients. In the implementation of the set of chirp-based functions, all possible chirps should be included not to mismodel the interference.
- A further improvement would be synthesizing the chirp directly from the TFD and subtracting it from the received signal. However, phase issues have to be dealt carefully.
- In the current implementation of the interference modeling routine, we use an empirically determined confidence measure. As a potential improvement to the algorithm, we should consider an adaptive process to determine the value of the confidence measure possibly as a function of the Gaussian window scale factors  $\{s_m\}$ , as they determine the TF resolution of the chirp.

Successful solutions to the above problems would make the proposed interference excision algorithm much faster and therefore potentially more suitable for real-time implementation.

### Audio Watermarking

To increase the data payload of the watermarking system, we can consider embedding other watermark messages that are represented by parametric equations defined in terms of a larger parameter set. We observe that data payload capacity is inversely proportional to the length of the PN sequence. However, the length of the PN sequence directly affects the BER such that a BER less than 20 per cent—the empirical limit for the HRT to correctly identify the directional features in the image of the TFD—is required for the correct detection of the watermark. Therefore, the shortest PN sequence that would keep the BER below 20 per cent would maximize the data payload. The proposed improvements for increasing the data payload are likely to increase the implementation time of the watermarking algorithm. However, real-time implementation issues are less of a concern for watermarking.

The SS watermarking algorithm determines the presence of a watermark by comparing the test statistic with an empirically determined threshold. Similarly, in the post-processing stage, the HRT identifies the peak values in the HRT space which yield the most likely parameter values representing the directional elements. Selection of the peaks is again a detection problem which would require a careful analysis of the HRT space particularly at high BER values. Therefore, any insight gained from further study of how the threshold values are to be selected, would enhance the robustness of the watermarking algorithm.

Finally, we can consider applying the innovative features and the post-processing power of the TF analysis techniques to other watermarking algorithms to enhance their robustness and watermark detection/extraction performance.

# Bibliography

- [1] I.J. Cox, and M.L. Miller, "The first 50 years of electronic watermarking" *Eurasip Journal on Applied Signal Processing*, vol. 2002, no. 2, Feb. 2002.
- [2] J.G. Proakis, *Digital Communications*, McGraw-Hill Companies, New York, NY, 2001.
- [3] R.C. Dixon, *Spread Spectrum Systems with Commercial Applications*, John Wiley and Sons, New York, NY, 1994.
- [4] P.G. Flikkema, "Spread spectrum techniques for wireless communication," *IEEE Signal Processing Magazine*, pp. 26–36, May 1997.
- [5] D.V. Sarwate and M.B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE*, vol. 68, pp. 593–619, May 1980.
- [6] I.J. Cox, M.L. Miller and J.A. Bloom, *Digital Watermarking*, San Diego, CA, Academic Press, 2002.
- [7] J.D. Laster and J.H. Reed, "Interference rejection in digital wireless communication," *IEEE Signal Processing Magazine*, pp. 37–62, May 1997.
- [8] L. Cohen, "Time-frequency distributions—A review," *Proc. IEEE*, vol. 77, pp. 941–981, 1989.
- [9] R.G. Baraniuk and D.L. Jones, "A Signal-Dependent Time-Frequency Representation: Optimal Kernel Design," *IEEE Trans. Signal Processing*, vol. 41, no. 4, pp. 1589–1602, April 1993.

- [10] F. Auger, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Trans. Signal Processing*, vol. 43, no. 5, pp. 1068–1089, May 1995.
- [11] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA, 1998.
- [12] S.G. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Trans. Signal Proc.*, 41(12): 3397–3415, 1993.
- [13] R.M. Rangayyan and S. Krishnan, "Feature identification in the time-frequency plane by using the Hough-Radon transform," *Pattern Recognition*, vol. 34, pp. 1147–1158, 2001.
- [14] S. Krishnan, "Adaptive Signal Processing Techniques for Analysis of Knee Joint Vibroarthrographic Signals," Ph.D. Thesis, University of Calgary, June 1999.
- [15] P.V.C. Hough, "Method and Means for Recognizing Complex Patterns," *US Patent no. 3069654*, 1962.
- [16] R.O. Duda and P.E. Hart, "Use of Hough transform to detect lines and curves in pictures," *Communications of the ACM*, 15(1): 11–15, January 1972.
- [17] G.T. Herman, *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography*, Academic Press, New York, NY, 1980.
- [18] W. Yang and G. Bi, "Adaptive wavelet packet transform-based narrowband interference canceller in DSSS systems," *IEE Electronics Letters*, pp. 1189–1190, July 1997.
- [19] A. Ranheim, "Narrowband interference rejection in direct-sequence spread-spectrum system using time-frequency decomposition," *IEE Proc. Comm.*, vol. 142, no. 6, pp. 393–400, Dec. 1995.



- [20] H.V. Poor and X. Wang, "Adaptive suppression of narrowband digital interferers from spread spectrum signals," *Proc. IEEE Intl. Conf. Acoustics, Speech and Signal Processing*, vol. 2, pp. 1061–1064, May 1996.
- [21] M.G. Amin, "Interference mitigation in spread spectrum communication systems using time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 90–101, January 1997.
- [22] S. Barbarossa and A. Scaglione, "Adaptive time-varying cancellation of wide-band interferences in spread-spectrum communications based on time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 47, no. 4, pp. 957–965, Apr. 1999.
- [23] B.S. Krongold, M.L. Kramer, K. Ramchandran and D.L. Jones, "Spread spectrum interference suppression using adaptive time-frequency tilings," *Proc. IEEE Intl. Conf. Acoustics, Speech and Signal Processing*, pp. 1881–1884, Apr. 1997.
- [24] X. Ouyang and M.G. Amin, "Short-time Fourier transform receiver for non-stationary interference excision in direct sequence spread spectrum communications," *IEEE Trans. Signal Processing*, vol. 49, no. 4, pp. 851–863, Apr. 2001.
- [25] A. Bultan and A.N. Akansu, "A novel time-frequency exciser in spread spectrum communications for chirp-like interference," *Proc. IEEE Intl. Conf. Acoustics, Speech and Signal Processing*, pp. 3265–3268, 1998.
- [26] R. Suleesathira and L.F. Chaparro, "Jammer excision in spread spectrum using discrete evolutionary-Hough transform and singular value decomposition," *Proc. 10th IEEE Workshop Statistical Signal and Array Processing*, pp. 519–523, 2000.
- [27] M.V. Tazebay and A.N. Akansu, "Adaptive subband transforms in time-frequency excisers for DSSS communications systems," *IEEE Trans. Signal Processing*, vol. 43, no. 11, Nov. 1995.

- [28] G. Matz and F. Hlawatsch, "Time-frequency projection filters: online implementation, subspace tracking, and application to interference excision," *Proc. IEEE Intl. Conf. Acoustics, Speech and Signal Processing*, May 2002.
- [29] M.G. Amin and G.R. Mandapati, "Nonstationary interference excision in spread spectrum communications using projection filtering methods," *Proc. 32nd Asilomar Conf. Signals, Systems and Computers*, vol. 1, pp. 827–831, Nov. 1998.
- [30] L. Cohen and T. Posch, "Positive time-frequency distribution functions," *IEEE Trans. Acoust. Speech Signal Processing*, vol. ASSP-33, no. 1, pp. 31–38, 1985.
- [31] P.J. Loughlin, J.W. Pitton and L.E. Atlas, "Construction of positive time-frequency distributions," *IEEE Trans. Signal Processing*, vol. 42, no. 10, pp. 2697–2705, Oct. 1994.
- [32] M. Arnold, "Audio watermarking: Features, applications and algorithms," *Proc. IEEE Intl. Conf. Multimedia and Expo*, vol. 2, pp. 1013–1016, 2000.
- [33] F.A.P. Petitcolas, M. Steinebach, F. Raynal, J. Dittmannb, C. Fontaine, N. Fates "A public automated web-based evaluation service for watermarking schemes: StirMark Benchmark" *Proc. Electronic Imaging 2001*, vol. 4314, pp. 22–26, Jan 2001.
- [34] S. Lee and Y. Ho, "Digital audio watermarking in the cepstrum domain," *IEEE Trans. Consumer Electronics*, vol. 46, no. 3, pp. 744–750, August 2000.
- [35] I.J. Cox, J. Killian, F.T. Leighton and T. Shamoan, "Secure spread-spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol. 6, no. 12, pp. 1673–1687, Dec. 1997.
- [36] P. Bassia, I. Pitas and N. Nikolaidis, "Robust audio watermarking in the time domain," *IEEE Trans. Multimedia*, vol. 3, no. 2, June 2001.
- [37] D. Kirovski and H. Malvar, "Spread-spectrum audio watermarking: Requirements, applications, and limitations," *IEEE Fourth Workshop Multimedia Signal Processing*, pp. 219–224, October 2001.

- [38] M.D. Swanson, B. Zhu and A.H. Tewfik, "Current state of the art, challenges and future directions for audio watermarking," *IEEE Intl. Conf. Multimedia Computing and Systems*, pp. 19–24, vol. 1, 1999.
- [39] W.N. Lie and L.C. Chang, "Robust high quality time-domain audio watermarking subject to psychoacoustic masking," *IEEE Intl. Symp. Circuits and Systems*, pp. 45–48, vol. 2, 2001.
- [40] J.W. Seok and J.W. Hong, "Audio watermarking for copyright protection of digital audio data," *Electronics Letters*, pp. 60–61, vol. 37, no. 1; Jan. 2001.
- [41] L. Gang, A.N. Akansu and M. Ramkumar, "MP3 resistant oblivious steganography," *IEEE Intl. Conf. Acoustics, Speech and Signal Processing*, vol. 3 pp. 1365–1368, 2001.
- [42] W. Bender, D. Gruhl, N. Morimoto and A. Lu "Techniques for data hiding," *IBM Systems Journal*, vol. 35, nos. 3 & 4, pp. 313–336, 1996.
- [43] S. Katzenbeisser, and F.A.P. Petitcolas, *Information hiding techniques for steganography and digital watermarking*, Artech House, Norwood, MA, 2000.
- [44] S. Erköçük, S. Krishnan and M. Zeytinoglu, "Robust audio watermarking using a chirp based technique," *IEEE Intl. Conf. Multimedia and Expo*, vol. 2, pp. 513–516, July 2003.

# Appendix A

## List of Publications

In this section, we list the publications resulted from our research work for the thesis.

### Publications and Presentations Related to Communications

- S. Erkucuk and S. Krishnan, "Interference excision in spread spectrum communications using adaptive positive time frequency distributions," *IEEE Intl. Conf. Acoustics, Speech and Signal Proc.*, vol. 4, pp. 4180, May 2002 (Student Forum).
- S. Krishnan and S. Erkucuk, "Interference excision in spread spectrum communications using adaptive positive time frequency distributions and Hough-Radon transform," submitted to *IEEE Trans. Wireless Communications*, Dec. 2002.
- S. Erkucuk and S. Krishnan, "Time frequency filtering of interferences in spread spectrum communications," *Intl. Symp. Signal Proc. and its Appl.*, vol. 2, pp. 323–326, July 2003.

### Publications and Presentations Related to Watermarking

- S. Erkucuk, S. Krishnan and M. Zeytinoglu, "A novel technique for digital audio watermarking," *Intl. Conf. for Upcoming Engineers*, vol. 1, pp. 123–127, May 2002.
- S. Erkucuk, "A novel technique for digital audio watermarking," *IEEE Intl. Conf. Multimedia and Expo / IBM Student Workshop presentation*, Aug. 2002.

- S. Erkucuk, S. Krishnan and M. Zeytinoglu, "Robust audio watermarking using a chirp based technique," *IEEE Intl. Conf. Multimedia and Expo*, vol. 2, pp. 513–516, July 2003.
- S. Erkucuk, K. Umapathy and S. Krishnan, "Digital audio processing for the next generation Internet," to appear in the proceedings of *Micronet Annual Workshop*, Sept. 2003.