Electrical and Computer Engineering ETDs          Engineering ETDs

Summer 7-9-2018

# Multirate Frequency Transformations: Wideband AM-FM Demodulation with Applications to Signal Processing and Communications

Wenjing Liu
*UNM*

_____

*Candidate*

_____

*Department*

This dissertation is approved, and it is acceptable in quality and form for publication:

*Approved by the Dissertation Committee:*

_____ , Chairperson

_____

_____

_____

_____

_____

_____

_____

# Multirate Frequency Transformations: Wideband AM–FM Demodulation with Applications to Signal Processing and Communications

by

## Wenjing Liu

B.S., Nankai University, China, 2012

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Engineering

The University of New Mexico

Albuquerque, New Mexico

July, 2018

# Dedication

*To my parents, for their support, encouragement.*

*"In prosperity, the indifferent, meet adversity, take things calmly."*

# Acknowledgments

I would like to thank my advisor, Prof. Balu Santhanam, for his support and guidence. He has been constantly inspiring me with wonderful ideas and catching my mistakes with great patience throughout the years.

My gratitude is also delivered to the members of my committee: Dr. Marios Pattichis, Dr. Manel Martinez-Ramon and Dr. Maria Crisina Pereyra for their valuable suggestions to my research and service on my committee.

I would also like to thank my uncles, who gave me a lot of help during my study at the University of New Mexico over the years. I have several other people I would like to thank, as well.

# Multirate Frequency Transformations: Wideband AM–FM Demodulation with Applications to Signal Processing and Communications

by

**Wenjing Liu**

B.S., Nankai University, China, 2012

Ph.D., Engineering, University of New Mexico, 2018

## Abstract

The AM–FM (amplitude & frequency modulation) signal model finds numerous applications in image processing, communications, and speech processing. The traditional approaches towards demodulation of signals in this category are the analytic signal approach, frequency tracking, or the energy operator approach. These approaches however, assume that the amplitude and frequency components are slowly time-varying, *e.g.*, narrowband and incur significant demodulation error in the wideband scenarios. In this thesis, we extend a two-stage approach towards wideband AM–FM demodulation that combines multirate frequency transformations (MFT) enacted through a combination of multirate systems with traditional demodulation techniques, *e.g.*, the Teager-Kasiser energy operator demodulation (ESA) approach to large wideband to narrowband conversion factors.

The MFT module comprises of multirate interpolation and heterodyning and converts the wideband AM–FM signal into a narrowband signal, while the demodulation module such as ESA demodulates the narrowband signal into constituent amplitude and frequency components that are then transformed back to yield estimates for the wideband signal.

This MFT-ESA approach is then applied to the various problems of: (a) wideband image demodulation and fingerprint demodulation, where multidimensional energy separation is employed, (b) wideband first-formant demodulation in vowels, and (c) wideband CPM demodulation with partial response signaling, to demonstrate its validity in both monocomponent and multicomponent scenarios as an effective multicomponent AM–FM signal demodulation and analysis technique for image processing, speech processing, and communications based applications.

# Contents

*Contents*

*Contents*

*Contents*

# List of Figures

*List of Figures*

*List of Figures*

# List of Tables

# Glossary

AM–FM        Amplitude modulation frequency modulation.

IA        Instantaneous amplitude.

IF        Instantaneous frequency.

GNGD        Generalized normalized gradient descent.

AS        Analytic signal.

HT        Hilbert transform.

ESA        Energy separation algorithm.

CESA        Continuous energy separation algorithm.

NRMSE        Normalized RMS IF demodulation error.

EMD        Empirical Mode Decomposition.

CR/IB        Carrir to information bandwidth ratio.

CR/FD        Carrier to frequency deviation ratio.

GNGD        Generalized normalized gradient descent.

FIR        Finite response.

*Glossary*

| | |
|---|---|
| BPF | Bandpass filter. |
| AWGN | Additive white Gaussian noise. |
| SNR | Signal-to-noise ratio. |
| CRLB | *Cramér Rao* lower bound. |
| MFT | Multirate frequency transformations. |
| DRFM | Digital radio frequency memory. |
| BMFT | Bi-dimensional multirate frequency transformation. |
| DCA | Dominant component analysis. |
| IFGT | Instantaneous frequency gradient tensor. |
| DEO | Diffierential energy operator. |
| DHODA | Discrete higher order demodulation algorithm. |
| LPC | Linear predictive coding. |
| CPM | Continuous phase modulation. |
| VA | Viterbi algorithm |
| FSK | Frequency shift keying. |
| MSK | Minimum shift keying. |
| GSM | Global System for Mobile communications. |
| MA | Moving-average. |
| AR | Auto-regressive. |

*Glossary*

DFE            Decision feedback equalization.

SBS            Symbol by symbol detector.

EDM           Energy demodulation of mixtures.

GDE           Generating differential or difference equation.

SEP            Normalized carrier separation.

# Chapter 1

# Introduction

## 1.1　Overview and Problem Statement

*Amplitude-modulation frequency-modulation* (AM–FM) is a modulation model of heightened interest, with various applications in communications, speech and images. In communications, *frequency modulation* (FM) is employed in mobile radio applications due to its resistance to fading and interference. Besides, AM-FM is well-suited for modeling the nonlinear resonance of the vocal tract in speech production and widely used in texture analysis for modeling images of certain categories such as fingerprints.

As a result, estimation of frequency content in the signals becomes a frequently encountered problem in signal processing. The parameter of interest in the case of AM–FM signal is the the instantaneous frequency. Various demodulation techniques based on Fourier spectrum analysis can be found in literature, which generate excellent performance for stationary signals but require excessive samples of observations, thereby reducing the resolution of the estimates. Other methods based on linear prediction filtering, or more recently adaptive filtering improve the frequency resolution

by continually adjusting the coefficients of the filter for better tracking. In addition, energy operator and its variants are also proposed for instantaneous frequency tracking. *However, these conventional approaches are limited to narrowband signals and fail to track the variation of the instantaneous frequency in the wideband scenario.*

## 1.2   AM–FM Signal Model

Monocomponent AM–FM signals are time-varying sinusoids of the form:

$$s\left(t\right) = a(t)\cos\left(\int_{-\infty}^{t} \omega_i\left(\tau\right) d\tau + \theta\right), \tag{1.1}$$

where *instantaneous amplitude* (IA) is denoted by $a(t)$ and the *instantaneous frequency* (IF) is given by

$$\omega_i\left(t\right) = \omega_c + \omega_m q_i(t). \tag{1.2}$$

Note that $w_c$ is the carrier (or mean) frequency and $q_i(t)$ is the normalized baseband modulated signal.

Specific for sinusoidal FM, where $a(t)$ remains a constant $A$, and $q_i(t)$ becomes a sinusoid, the IF can be further expressed as

$$\omega_i\left(t\right) = \omega_c + \omega_m \cos(\omega_f t + \theta). \tag{1.3}$$

Sinusoidal FM signals can be expressed via Bessel functions via

$$s\left(t\right) = A \sum_{n=-\infty}^{+\infty} J_n\left(\beta\right) \cos\left(\omega_c t + n\omega_m t\right), \tag{1.4}$$

where $J_n$ is the $n^{th}$ order cylindrical Bessel function of the first kind. The modulation index of the sinusoidal FM is defined as the ratio $\beta = \omega_m/\omega_f$ and the associated Carson's bandwidth [1] is given by

$$B = 2(\beta + 1)\omega_f. \tag{1.5}$$

If $\beta \gg 1$, then it corresponds to the wideband FM according to the literature of FM communication systems. In addition, the *carrier-to-information-bandwidth ratio* (CR/IB) and the *carrier-to-frequency-deviation ratio* (CR/FD) are defined respectively as:

$$\frac{CR}{IB} = \frac{\omega_c}{\omega_f}, \quad \frac{CR}{FD} = \frac{\omega_c}{\omega_m}. \tag{1.6}$$

## 1.3 Conventional Demodulation Approaches

In this section, we briefly review three conventional demodulation approaches. Note that the adaptive linear predictive IF tracking are intended for pure FM signals, whereas the *analytic signal* (AS) approach based on the *Hilbert transform* (HT) and the *energy separation algorithm* (ESA) can be applied to AM–FM signals.

### 1.3.1 Adaptive Linear Predictive IF Tracking

For a pure FM signal, the IF demodulation employing adaptive filters is proposed in prior work [2, 3]. According to the Wiener-Hopf equations [4], the optimal coeffcients of a linear predictor are given by

$$\mathbf{w}^{opt} = \mathbf{R}_{xx}^{-1} \mathbf{r}_{dx}, \tag{1.7}$$

where $\mathbf{w}^{opt}$ denotes the optimal tap weight vector, $\mathbf{R}_{xx}$ denotes the input correlation matrix and $\mathbf{r}_{dx}$ denotes the cross-correlation between input vector and desired signal.

As summarized in [2], the prediction error filter is given by:

$$E(z) = 1 - \sum_{l=1}^{L} g_l^{opt} z^{-l}, \tag{1.8}$$

where $\{g_l^{opt}\}_{l=1}^L$ are the coefficients of the corresponding optimal predictor. Conditioned on the small prediction error assumption and the further assumption that the message signal remains essentially invariant over the sampling range of the linear prediction filter, we end up with the approximation in [5] through (1.8) given by

$$\sum_{l=1}^L g_l(k) \exp\{-jl\,[w_c + m(k)]\} \simeq 1, \tag{1.9}$$

where $g_l(k)$ is the weight coefficient of tap $l$ at time index $k$ and $m(k)$ is the sample of the message signal at time index $k$. Then the IF of the signal of interest can be estimated by executing the following steps: 1) Compute the coefficients of the prediction error filter; 2) Obtain the roots of the coefficient polynomial; 3) Calculate the phase argument of the complex conjugate pole location of the corresponding roots.

In the prior work [2], adaptive algorithms such as AS-LMS and AF-RLS have been incorporated into the structure of linear predictor and compared with each other based on the demodulation error. However, for both algorithms, the step-size or the forgetting factor need to be truncated to remain within certain range. Here we choose the *generalized normalized gradient descent* (GNGD) [6] for coefficients update of the linear predictor, which avoids truncation of the adaptively adjusted step-size. The algorithm for this GNGD linear predictive filter is summrized via

$$e(k) = x(k+1) - \sum_{l=1}^L g_l(k)\,x(k-L+1), \tag{1.10}$$

$$\beta(k) = \beta(k-1) - \rho\alpha\frac{e(k)e(k-1)\mathbf{x}^T(k)\mathbf{x}(k-1)}{(\|\mathbf{x}(k)\|_2^2 + \beta(k))^2}, \tag{1.11}$$

$$\eta(k) = \frac{\alpha}{\|\mathbf{x}(k)\|_2^2 + \beta(k)}, \tag{1.12}$$

$$\mathbf{g}(k+1) = \mathbf{g}(k) + \eta(k)e(k)\mathbf{x}(k), \tag{1.13}$$

where $\mathbf{x}(k)$ and $\mathbf{g}(k)$ denote the vectors of input and tap weights at time index $k$ respectively, $\alpha$ is the step-size parameter and $\rho$ is the offset learning rate parameter.

The merit of the GNGD algorithm lies in that the adaptation of its learning rate provides compensation for the assumptions in the derivation of NLMS. Due to its robustness and stability, the GNGD is well-suited for narrowband nonstationary signal environments.

## 1.3.2 Analytic Signal and Hilbert Transform

For a one-dimensional real-valued signal $s(t) : \mathbb{R} \to \mathbb{R}$, its *Hilbert transform* is defined by the convolution of $s(t)$ and the function $h(t) = \frac{1}{\pi t}$ using Cauchy principal value via

$$q(t) = \mathcal{H}\left[s(t)\right] = s(t) * \frac{1}{\pi t} = \frac{1}{\pi} \, \text{p.v.} \int_{\mathbb{R}} \frac{s(t - \xi)}{\xi} \mathrm{d}\xi. \tag{1.14}$$

The *analytic signal* (AS) is defined as the complex-valued signal $z(t) = s(t) + jq(t)$, where the imaginary part $q(t)$ is given by the *Hilbert transform* of $s(t)$. Assume that $s(t)$ is a real-valued AM–FM signal given by

$$s(t) = a(t) \cos\left(\omega_c t + \phi(t)\right), \tag{1.15}$$

where $a(t)$ is the IA, $\omega_c$ is the carrier frequency and $\phi(t)$ is the phase for the IF $\phi'(t)$. Note that the IF defined here does not involve the carrier frequency $\omega_c$. According to the basic property of the *Hilbert transform* and *Bedrosian's theorem*, $q(t)$ is the approximation to the product of the IA and the quadrature of the FM part given by

$$q(t) = \mathcal{H}\left[s(t)\right] \approx a(t) \sin\left(\omega_c t + \phi(t)\right). \tag{1.16}$$

However, this approximation is valid only if the following conditions hold: 1) $a(t)$ is a narrowband lowpass signal that varies slowly with time, 2) and the carrier frequency $\omega_c$ is sufficiently large such that

$$\omega_c \gg \phi'(t). \tag{1.17}$$

As a result, the corresponding analytic signal is of the form:

$$z(t) = s(t) + jq(t) \approx a(t) \exp\left[j\left(\omega_c t + \phi(t)\right)\right]. \tag{1.18}$$

Hence the IA $a(t)$ and the IF $\phi'(t)$ can be estimated respectively by

$$a(t) \approx \|z(t)\|, \tag{1.19}$$

$$\phi'(t) \approx \frac{\partial}{\partial t}\left(\arctan\frac{\Im\left(z(t)\right)}{\Re\left(z(t)\right)}\right) - \omega_c. \tag{1.20}$$

### 1.3.3   Energy Separation Algorithm

The *energy separation algorithm* as summarized in [7, 8], based on the Teager-Kaiser energy operator, is widely used for monocomponent AM–FM demodulation. For continous-time signal $x(t)$, the nonlinear Teager-Kaiser energy operator in the continuous case is given by

$$\Psi_c[x(t)] = \dot{x}^2(t) - x(t)\ddot{x}(t), \tag{1.21}$$

where $\dot{x}(t)$ and $\ddot{x}(t)$ denote the first and second derivatives of $x(t)$ respectively.

It was first introduced by Kaiser and was applied to track the energy of harmonic oscillation. For instance, the energy of the oscillatory signal $x(t) = A\cos(\omega_c t + \theta)$ can be tracked by the operator $\Psi_c$ via

$$\Psi_c[A\cos(\omega_c t + \theta)] = (A\omega_c)^2. \tag{1.22}$$

The energy operator can be further used to analyze the oscillation of signals with time-varying amplitude and frequency. By applying the energy operators $\Psi_c$ on the AM–FM signal, the IA $a(t)$ and IF $\omega_i(t)$ (excluding the carrier frequency $\omega_c$) satisfy the following relation:

$$\Psi_c\left[a(t)\cos\left(\omega_c t + \int_0^t \omega_i(\tau)d\tau + \theta\right)\right] \approx [a(t)(\omega_c + \omega_i(t))]^2. \tag{1.23}$$

Therefore, the IA $a(t)$ and the IF $\omega_i(t)$ of an AM–FM signal $x(t)$ can be estimated via the *continuous energy separation algorithm* (CESA) summarized by

$$\frac{\Psi[x(t)]}{\sqrt{\Psi[\dot{x}(t)]}} \approx |a(t)|. \tag{1.24}$$

$$\sqrt{\frac{\Psi[\dot{x}(t)]}{\Psi[x(t)]}} - \omega_c \approx \omega_i(t) \tag{1.25}$$

assuming the IA $a(t)$ and the IF $\omega_i(t)$ do not vary too fast or too greatly in value compared to the carrier frequency $\omega_c$.

## 1.4   Narrowband Constraint

In general, conventional demodulation techniques only perform well when the input signal is under the narrowband constraint. The IA is required to be slowly-varying and the IF cannot change too fast or too greatly in value compared to the carrier frequency. To quantify the narrowband constraint, assume the IA $a(t)$ and the IF $\omega_i(t)$ are both bandlimited with highest frequencies $\omega_a$ and $\omega_f$ respectively. We further assume a nonnegative IA $a(t) = 1 + \kappa b(t)$ without loss of generality. According to the work by Maragos *et al.* [7], the narrowband constraint can be summarized via

$$\omega_a \ll \omega_c, \tag{1.26}$$

$$\kappa \ll 1, \tag{1.27}$$

$$\omega_f \ll \omega_c, \tag{1.28}$$

$$\lambda = \frac{\omega_m}{\omega_c} \ll 1, \tag{1.29}$$

where $\lambda$ is the inverse of the CR/FD, which is also called the FM *modulation depth.* Conventional demodulation approaches only perform properly under the narrowband condition, in particular, when the two parameters CR/FD and CR/IB are sufficiently

large. However, in the case of wideband scenario, significant errors are incurred due to the wideband nature of the signal, regardless of the choice of demodulation approaches.

## 1.5 Claim of this Work

The objective of this thesis is to develop a general scheme called *multirate frequency transformations* (MFT) that can be combined with conventional demodulation approaches to improve their performances in the wideband scenario. The proposed framework incorporates the multirate operations and heterodyning to transform the wideband signal into a narrowband waveform. The transformed narrowband signal is then demodulated and the instantaneous frequency and amplitude estimates of the original wideband signal are recovered via the corresponding inverse transformation. *The transformed AM–FM signal encounters significantly smaller demodulation errors in comparison to the original wideband signal. An alternative framework that enables large multirate conversion factors is further proposed and practically implemented by employing the multirate identities.*

The proposed MFT framework is then applied to the various problems to demonstrate its validity in both monocomponent and multicomponent scenarios as an effective multicomponent AM–FM signal demodulation and analysis technique for image processing, speech processing, and communications based applications:

- FM image can be locally categorized as wideband in terms of the modulation indexes defined along the gradient directions of the phase, where the notion of a synthetic modulation index has been introduced. The MFT is then extended to two dimensions for wideband AM–FM image demodulation, *e.g.*, fingerprint images, where multidimensional higher order energy operator is also employed.

- The first-formant of many vowels is demonstrated to have significant amount of *frequency modulation.* Henceforth, these large deviation first formants are better tracked by employing the MFT framework and popular AM–FM demodulation techniques such as ESA, than the widely used *linear predictive coding* (LPC) approach that only assumes narrowband AM formants.

- The MFT is applied to wideband *continuous phase modulation* (CPM) demodulation with partial response signaling, which can be employed in satellite communications. It is shown to have significantly less complexity than the Viterbi solution that depends on the modulation index and phase states of the CPM scheme.

## 1.6 Structure of this Work

The thesis is organized as the following: 1) In chapter 2, we derive the MFT framework that aims at relaxing the narrowband constraints encountered by most AM–FM demodulation techniques in the wideband scenario. The possibility to incorporate large conversion factors is also investigated and the corresponsding alternative framework has been proposed. 2) Chapter 3 extends the one-dimensional MFT framework to two-dimensional images. The resultant *bi-dimensional multirate frequency transformations* (BMFT) approach in combination with higher order energy operator is proposed for wideband images demodulation. Simulation results with application to fingerprint images have been illustrated to prove its efficacy. 3) In chapter 4, we apply the MFT-ESA combination to demodulate the instantaneous frequency of the large deviation first formant of human vowels extracted via *empirical mode decomposition* (EMD). Then we demonstrate that the formant estimates based on the instantaneous frequency are more precise than that of the widely used LPC approach. 4) Chapter 5 proposes a solution for wideband partial response CPM demodulation employed in

satellite communications based on the MFT-ESA approach and the *decision feedback equalization* (DFE).

# Chapter 2

# Multirate Frequency Transformations

The *multirate frequency transformations* approach that involves multirate systems as well as heterodyning has been proposed in prior work [3] as a mechanism for wideband FM demodulation that usually has a: 1) large frequency deviation; 2) large information bandwidth. The alternative MFT framework that can accommodate large multirate conversion factors with practical implementation has also been proposed in recent work [9] by the authors. In a nutshell, the MFT can be combined with existing demodulation methods, such as the *analytic signal* approach or energy operators, serving as a general framework for demodulation of wideband FM signals or AM–FM signals whose FM components are categorized as wideband.

## 2.1 Basic MFT System

Using the scaling property of the Fourier transform, compression in frequency domain is equivalent to expansion in the time domain expressed as

$$y(t) = x(at) \iff Y(\omega) = X\left(\frac{\omega}{a}\right), \tag{2.1}$$

where $a = 1/R < 1$ is the factor of frequency compression. The IF of the compressed signal becomes a scaled version of the input IF by a factor $R$ expressed as

$$\tilde{\omega}_i(t) = \frac{\omega_i(t)}{R} = \frac{\omega_c}{R} + \frac{\omega_m}{R} q_i\left(\frac{t}{R}\right). \tag{2.2}$$

Note that for compressed signal, the carrier frequency is also scaled by the same factor $R$, which is undesirable since the ratios CR/FD and CR/IB that we wish to increase still remains invariant. Hence the heterodying operator is cascaded right after the frequency compression module in order to upshift the carrier frequency to a higher level where we can attain larger CR/FD and CR/IB ratios. The compressed signal after frequency translation and bandpass filtering in the heterodying module is given by

$$y_{ush}(t) = [y(t)\cos(\omega_d t)] * h_{BPF}(t), \tag{2.3}$$

where $*$ denotes the convolution, $\omega_d$ refers to the amount of frequency translation, and $h_{BPF}(t)$ represents the impulse response of the bandpass filter. Specific for the case of sinusoidal FM, it can be further simplified as

$$y_{ush}(t) \simeq \frac{1}{2} A \cos\left(\phi\left(\frac{t}{R}\right)\right), \tag{2.4}$$

where $\phi(t)$ denotes the phase of the original FM signal. The resultant signal that has a scaled information bandwidth with a higher CR/IB then passes through the demodulation block for IF extraction. Eventually, the IF estimate of the original signal is evaluated by the inverse MFT relation via

$$\omega_i^{\text{out}}(t) = R\left(\tilde{\omega}_i(Rt) - \omega_d\right), \tag{2.5}$$

where $\tilde{\omega}_i(t)$ is the IF of the compressed and frequency translated signal.

As for the discrete-time signal, compression and expansion can be substituted by the corresponding multirate operations of interpolation and decimation respectively [10] with their properties carried over to the discrete counterparts. The block diagram of the basic MFT framework in prior work [3] is illustrated by Fig. 2.1.
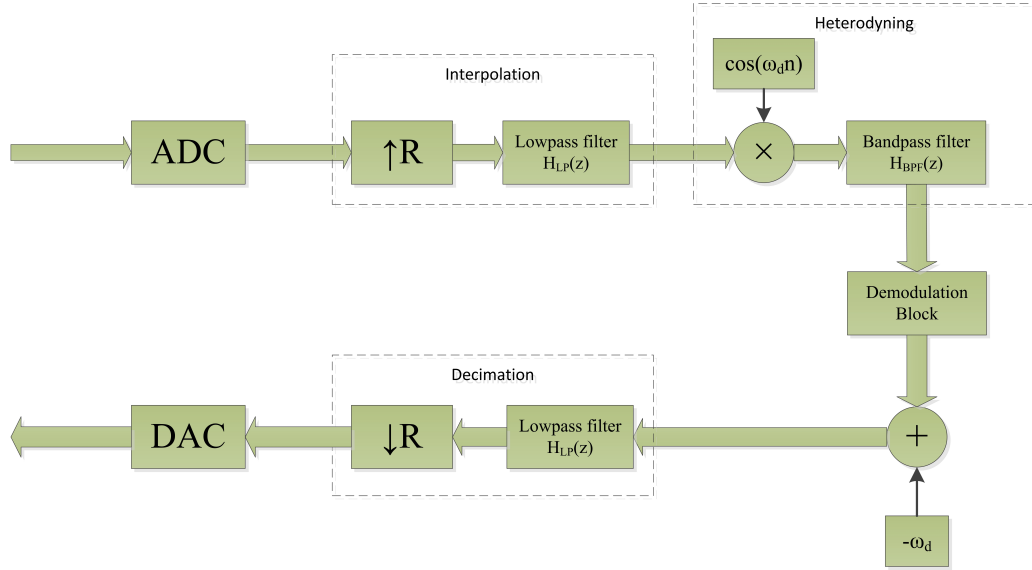
Figure 2.1: Block diagram of the basic MFT framework. The wideband signal is first sufficiently sampled above the Nyquist rate, interpolated by a factor $R$ and then heterodyned via multiplying $\cos(\omega_d n)$, followed by a discrete FIR bandpass filter with a scaling module based on (2.4) to achieve MFT. Then it goes through a demodulation block to obtain the IF estimation of the compressed heterodyned signal. To estimate IF of the original signal, the compressed heterodyned IF is then shifted back by subtracting $\omega_d$, decimated by $R$ and scaled back approapriately according to (2.2), followed by the DAC module.

Interpolating the input signal will result in the reduction of both the frequency deviation and information bandwidth by a factor of $R$. Similar to increasing the sampling rate, the IF of the interpolated signal becomes slow-varying and the assumption that the message signal remains constant over the carrier period is more likely to hold, which in turn boost the performances of conventional demodulation

algorithms. Meanwhile, heterodying serves the purpose of increasing the ratios of CR/FD and CR/IB by compensating for the scaled carrier frequency. By passing through the heterodying process, the CR/FD and the CR/IB of $y_{ush}(t)$ are given by

$$\left[\frac{CR}{FD}\right]_{out} = \left[\frac{CR}{FD}\right]_{in} + \frac{R\omega_d}{\omega_m} \tag{2.6}$$

$$\left[\frac{CR}{IB}\right]_{out} = \left[\frac{CR}{IB}\right]_{in} + \frac{R\omega_d}{\omega_f}. \tag{2.7}$$

## 2.2 Alternative MFT System for Large Conversion Factors

As we look further into the MFT framework, an important question regarding the selection of the conversion factor $R$ arises. Larger factors over hundred or thousand can be supported by current high-speed DSP with large memory, as in the case of *digital radio frequency memory* (DRFM) system design [11]. It is intuitive to expect a further reduction in the demodulation error since the gain brought by frequency compression should be extendable through the use of a larger factor. However, for sufficiently large factors $R$ the passband of the lowpass filter in the multirate operation and that of the heterodyne-BPF operation will be scaled by $R$. For example, if $R = 1000$, we require a lowpass filter with cut-off frequency at $\pi/1000$ and a bandpass filter with a passband edge less than or equal to $\pi/1000$. However, filters with such narrow passbands are unrealistic for direct implementation by any structure[1]. Therefore, the design of the BPF within the basic MFT framework becomes the bottleneck that limits the use of a very large factor.

In general, the MFT framework achieves better demodulation performance when the conversion factor becomes larger, whereas large conversion factors require the

---

[1]Narrow passband implies clustered poles and zeros that result in sensitivity and stability issues of digital filters as described in [12].
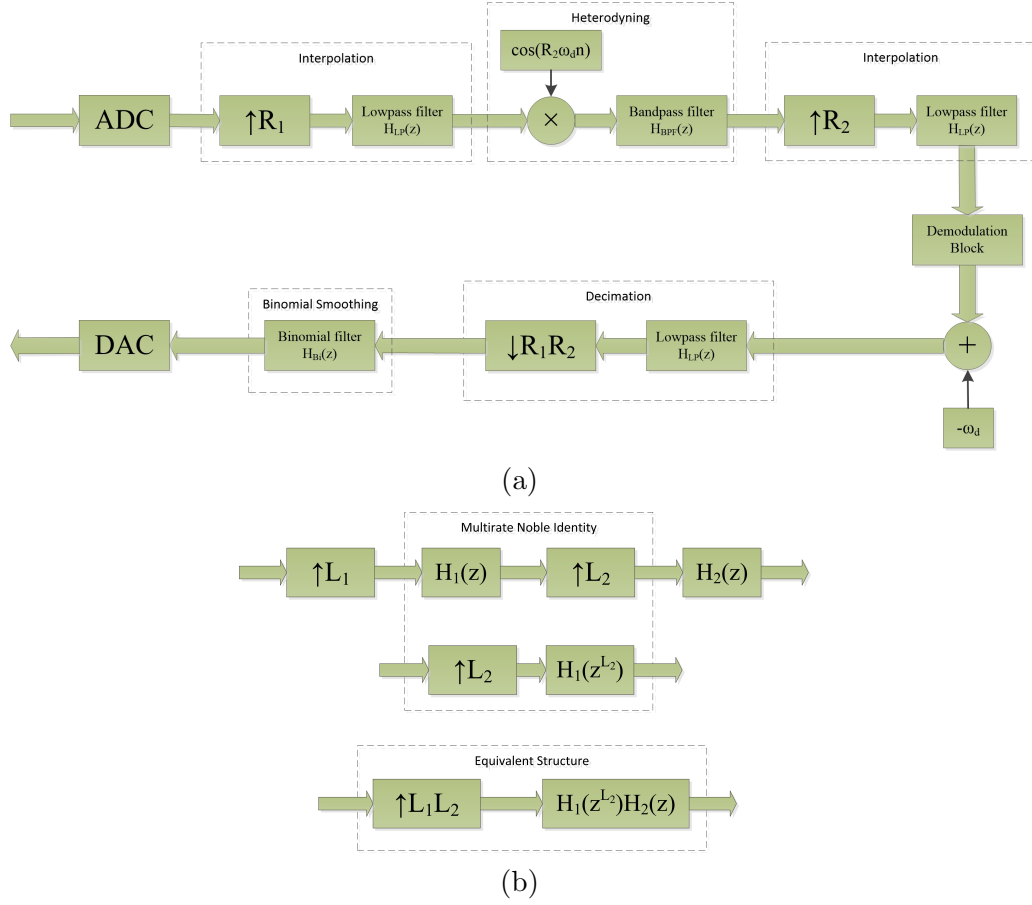
Figure 2.2: Block diagrams of the alternative MFT system for large conversion factors and the Noble identities relation applied in the system. (a) Block diagram of the proposed MFT framework for large wideband to narrowband conversion factors. The interpolation module of the prior framework is separated into two with one in front of the heterodyning module and the other one right after. The first interpolation module has a relatively small upsampling rate of $R_1$ which is appropriately chosen such that the discrete BPF can be implemented within the range of half the sampling rate after heterodyning the signal with a frequency translation of $\omega_d$. The relatively small $R_1$ would result in a wider passband for the discrete bandpass filter, thus reducing its design of complexity. (b) Multistage implementation for interpolation modules in 2.2(a) based on the Noble identities. Note that multistage implementation for the corresponding decimation modules can be realized in a similar fashion.

heterodyning bandpass filter to have significantly narrow passband, which is hardly achieved in practice. In order to reduce the burden placed on the practical imple-

mentation of the bandpass filter, we first consider a different MFT framework where the order of the interpolation operator and the heterodyning operator are exchanged. Due to the switch of interpolation and heterodying, the CR/FD and CR/IB parameters under this MFT framework are given by

$$\left[\frac{CR}{FD}\right]_{out} = \left[\frac{CR}{FD}\right]_{in} + \frac{\omega_d}{\omega_m} \tag{2.8}$$

$$\left[\frac{CR}{IB}\right]_{out} = \left[\frac{CR}{IB}\right]_{in} + \frac{\omega_d}{\omega_f}. \tag{2.9}$$

By comparing these ratios with (2.6) and (2.7), note that the upshift frequency $\omega_d$ in this case needs to be sufficiently large such that the ratios of CR/FD and CR/IB still stay at high level. However, if $\omega_d$ is too large such that the resultant carrier frequency after heterodying exceeds one half of the sampling rate, then we need to interpolate the signal first by an appropriate factor in order to perform discrete-time bandpass filtering after heterodyning. Hence the practical implementation of MFT framework for large conversion factors is not simply exchanging the order of interpolation and heterodyning. In practice, an interpolation operation is required prior to the heterodying with an appropriate factor that depends on the upshift frequency $\omega_d$ and the sampling frequency of the original wideband FM signal. This implies that the overall interpolation factor can be split into two with the first one prior to the frequency translation and the other one right after. Then upshifting by a frequency $\omega_d$ that is not too large would result in a relatively small factor for the first interpolation, thereby reducing the burden of the heterodyne-BPF.

Based on the analysis above, an alternative MFT framework that is capable of incorporating large conversion factors and achieving practical design of the heterodyne-BPF is proposed in this thesis, as illustrated in Fig. 2.2(a). The interpolation module of the prior framework is separated into two with one in front of the heterodyning module and the other one right after. The first interpolation module has a relatively small upsampling rate of $R_1$ which is appropriately chosen such that the discrete BPF

can be implemented within the range of half the sampling rate after heterodyning the signal with a frequency translation of $\omega_d$. The relatively small $R_1$ would result in a wider passband for the discrete bandpass filter, thus reducing complexity of the design. In general, there is a sacrifice in terms of achievable CR/IB and CR/FD ratios for the alternative MFT framework. However, the basic MFT system cannot realize large conversion factors, due to the placement of impractical constraints on BPF design.

In addition, a binomial smoothing module is incorporated into the alternative framework to further reduce the effects of noise. Though the signal is wideband, the IF waveform itself is not necessarily wideband. In many cases, the wideband FM is primarily incured by a large modulation index while the IF waveform remains in the narrowband range. Under this assumption, applying binomial smoothing efficiently filters out the high frequency noise in the IF estimation. When the SNR is high, the improvement becomes clearly evident as we demonstrate later. Usually we would expect a gain between 5 dB and 10 dB in the scenario of relatively high SNR.

## 2.3   Demodulation Results via MFT Approaches

In this section, we present demodulation results using the basic and alternative MFT frameworks respectively under both noise free and noisy environments. Here we first apply the GNGD approach introduced in Subsection 1.3.1 as the demodulation technique to test on wideband FM signals. Note that the demodulation performance is judged by the *normalized RMS IF demodulation error* (NRMSE) throughout this section.

We first look at the case of a wideband sinusoidal FM signal that has a large modulation index $h = 10$ and the CR/IB of 20 dB. Under a noise free environment, the performance of the basic MFT framework is illustrated by Fig. 2.3(a). Note

Figure 2.3: Comparison between performances of both MFT frameworks under noise free environments. (a) Performance of the basic MFT framework with R specifying the multirate factor and L specifying the order of BPF. (b) Performance of the alternative MFT framework with R specifying the multirate factor and L specifying the order of BPF.

that the performance associated with $R = 1$, i.e, the origin of the performance curve corresponds to direct demodulation by GNGD without MFT, while $\omega_d$ is the normalized upshift radian frequency translation in the range of $[0, \pi]$. By applying a large conversion factor of $R = 128$, a reduction of around 40 dB in the demodulation error over direct GNGD demodulation is attained. The result of Fig. 2.3(a) confirms the claim that a large conversion factor strengthens the benefits achieved by frequency compression thus leading to significant reduction in the demodulation error. Moreover, it reflects the fact that the use of a larger factor requires a very high order *finite response* (FIR) *bandpass filter* (BPF) with a satisfactory frequency response. For example, $R = 128$, demands the order of FIR bandpass filter to be as high as 4096, which results in unrealistic parameters for the narrow passband. This constraint seriously limits the implementation in a practical system for large factors.

To relax the constraint, the alternative MFT framework is applied, the demodulation performance of which is illustrated by 2.3(b). In comparison to Fig. 2.3(a), it can be observed that the required order for the FIR bandpass filter is effectively
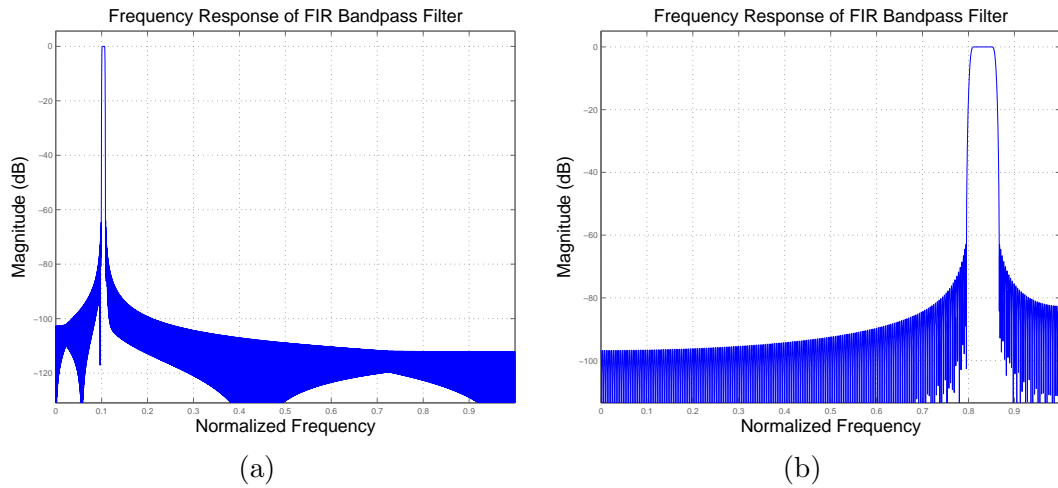
Figure 2.4: Comparison between the basic and alternative MFT frameworks in terms of the requirements on the heterodyne-BPF with a conversion factor $R = 128$. (a) Frequency response of BPF in the previous MFT framework with order $L = 4096$. (b) Frequency response of BPF in the previous MFT framework with order $L = 512$. Note that the proposed framework succesfully reduced the complexity of the design of heterodyne-BPF.

reduced at the cost of sacrificing a small amount of improvement in demodulation error. For instance, when $R = 128$, the order for the FIR bandpass filter drops significantly from 4096 to 512 with just 4 dB loss in performance, suggesting no observable difference in performance between both MFT frameworks except for the dramatic reduction of the order for the heterodyne-BPF. Note that the frequency response of the heterodyne-BPF in the alternative MFT framework for the case $R = 128$ is illustrated by Fig. 2.4(b), which has much wider passband compared to the frequency response of the BPF in the basic MFT framework shown in Fig. 2.4(a) and thus practical for implementation.

For noisy environments, the performance of the basic MFT framework is illustrated by Fig. 2.5(a) for the same sinusoidal FM signal corrupted by *additive white Gaussian noise* (AWGN). Based on this observation, improvement over GNGD demodulation alone varies according to the different SNRs. For example, when the
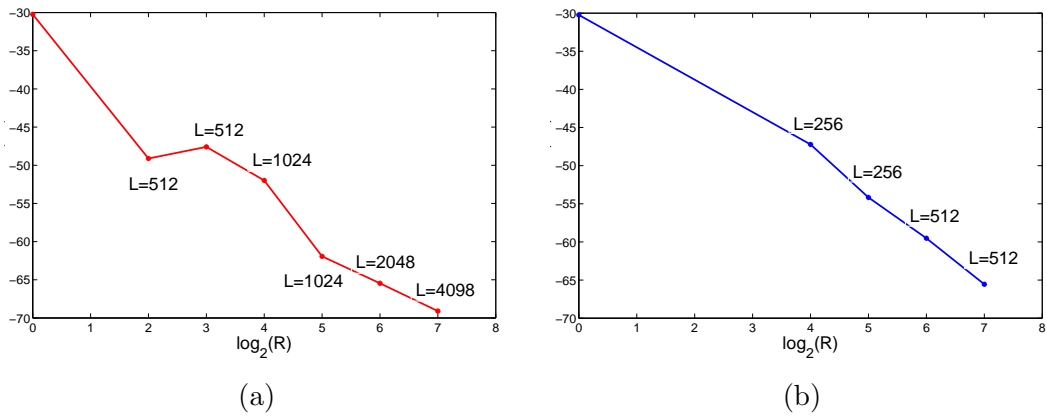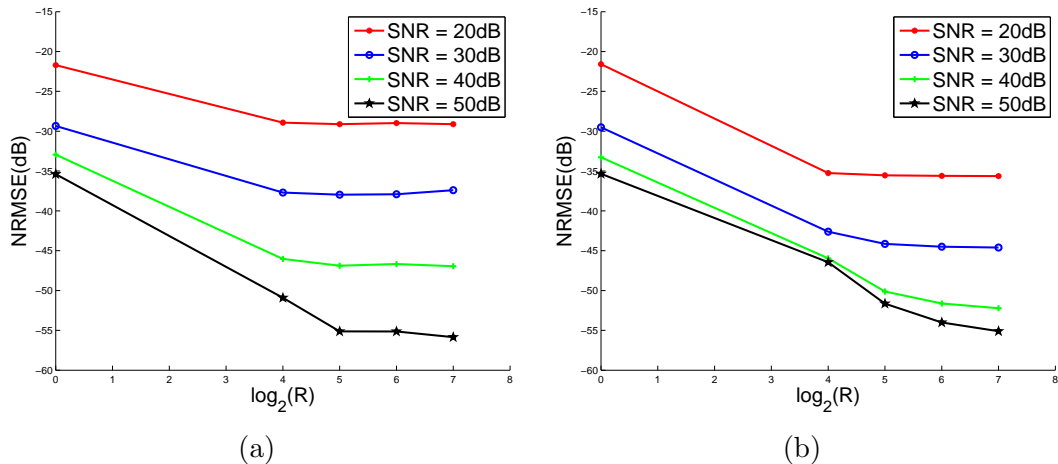
Figure 2.5: Comparison between performances of both MFT frameworks in environments corrupted by AGWN. (a) Performance of the basic MFT framework with R specifying the multirate conversion factor. (b) Performance of the alternative MFT framework with R specifying the multirate conversion factor.

SNR is 20 dB, the improvement is only around 6 dB, and when the SNR increases to 40 dB, improvement increases to around 20 dB. In addition, Fig. 2.5(b) summarizes the results of the alternative MFT framework in the presense of noise, indicating better performance due to the binomial smoothing. For the case of 20 dB SNR, the improvement increases to 15 dB compared with the basic MFT system. Note that the NRMSE gradually becomes saturated as R increases, due to LTI filtering induced harmonic distortion of the FM signals[2]. In contrast, the proposed MFT approach results in a better performance, and offers a more relaxed constraint on the BPF design.

We further investigate an extreme wideband scenario under the noise free environment, where the modulation index $\beta$ is as large as 50 and the frequency deviation is equal to the carrier frequency with the IF varying over the entire carrier range. For the signal of interest, the IF estimates of both the basic and alternative MFT

---

[2]Since they are only approximate eigenfunctions as described in [13].
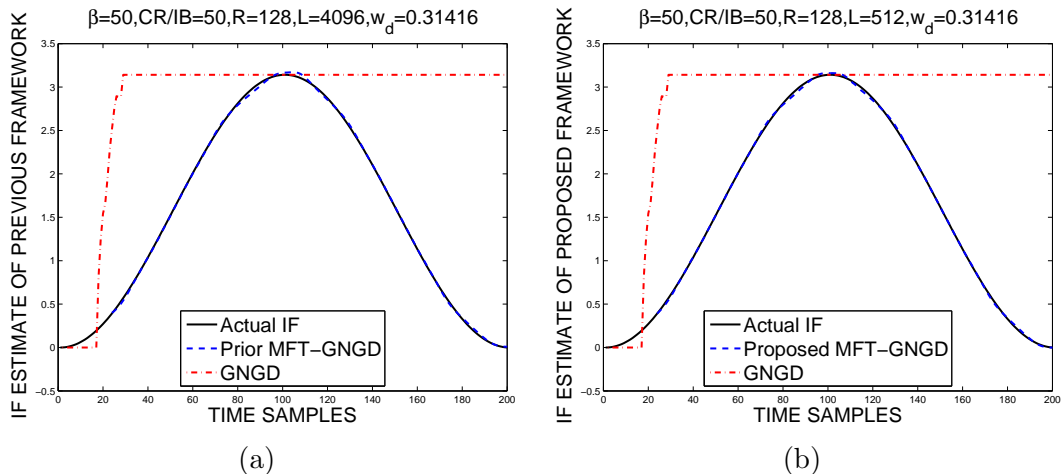
Figure 2.6: Demodulation performances of both MFT frameworks with conversion factor $R = 128$ and normalized radian frequency shift $w_d = 0.1\pi$ under the extreme senario with modulation index $\beta = 50$. (a) IF estimates (dashed line) of GNGD under the basic MFT framework with the heterodyne-BPF order $L = 4096$, (dashed-dotted line) the GNGD, and (solid line) actual IF. (b) IF estimates (dashed line) of the GNGD under the alternative MFT framework with the heterodyne-BPF order $L = 512$, (dashed-dotted line) the GNGD, and (solid line) actual IF. Note that the GNGD alone fails under this large deviation FM signal with a modulation index as large as 50, while the GNGD combined with both MFT frameworks maintain tracking with similar performances.

frameworks are illustrated by Fig. 2.6(a) and Fig. 2.6(b) respectively. It can be observed that the GNGD demodulation alone fails in this extreme wideband scenario, while both MFT frameworks maintain tracking. This observation implies that both MFT frameworks guarantee demodulation with acceptable performance even in the worst case scenarios where conventional algorithms would fail.

To quantify the performance of the MFT approach, we explore another scenario where the signal is a wideband linear chirp instead of sinusoidal FM. The short-time spectrum of the chirp signal is illustrated by Fig. 2.7(a). To validate the performance of the MFT approach , we can compare the variance of error with respect to the chirp rate estimate with its *Cramér Rao lower bound* (CRLB). The chirp rate can be obtained from the demodulated IF followed by a least square estimator. In
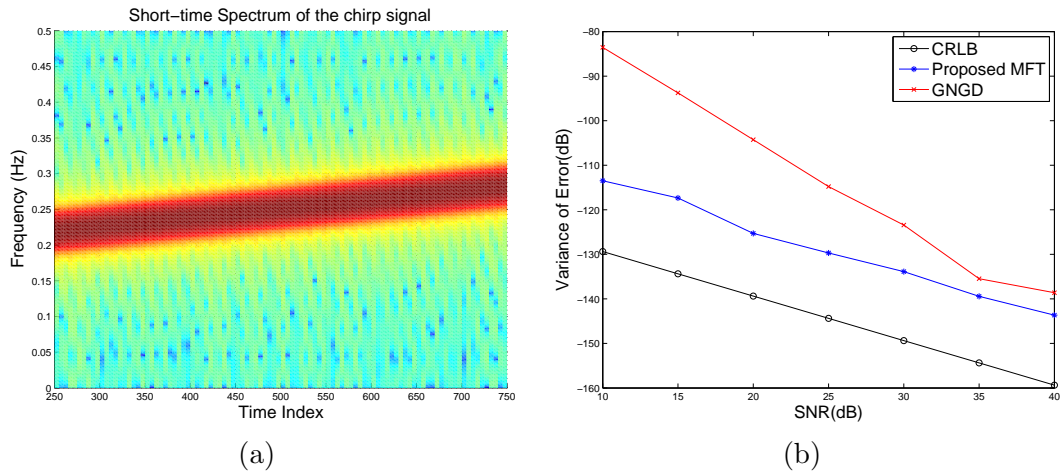
Figure 2.7: Performance of the alternative MFT framework with a multirate conversion factor of 128 in wideband linear chirp scenario. (a) Short-time spectrum of the wideband linear chirp. (b) Comparison of the error variance with respect to the estimate of chirp rate. Note that there is only a small gap less than 20 dB between the error variance estimate of the alternative MFT framework and the corresponding *Cramér Rao lower bound* (CRLB).

the presence of noise with different SNRs, the result is summarized in Fig. 2.7(b). Improvement over the GNGD alone is more apparent with respect to lower SNR, indicating satisfactory performance of the MFT approach in noisy conditions. Also note that gap between the error variance estimate of the MFT approach and the corresponding CRLB is nearly a constant, which can be explained directly via the loss of spectrum incurred due to filtering of the FM signal.

## 2.4 MFT Compatibility with HT and ESA Approaches

In general, the MFT framework can be combined with most demodulation approaches to enhance their performances in the wideband scenario. However, for a given de-

modulation scheme, the gain in error reduction achievable via the MFT approach mainly depends on its inherent sensitivity to parameters such as CR/IB, CR/FD, and the modulation index. Here we refer to the comparison results of the HT approaches and ESA approaches in 1D [14].

The HT demodulation approaches are usually not very sensitive to the CR/IB of the signal. In general, the demodulation performance of the HT is invariant to the CR/IB of the signal. As shown in prior work [14] by Potamianos and Maragos, the error versus CR/IB response of the Hilbert transform is almost flat, suggesting that by increasing the CR/IB via the MFT, the gain achievable through error reduction is limited. On the other hand, the performance of the ESA is shown to be very sensitive to the CR/IB of the signal, where the error decreases sharply as the CR/IB of the signal increases. Hence the gain attained by increasing the CR/IB via the MFT with demodulation schemes based on ESA is expected to be more significant than demodulation schemes using the HT approaches.

As for the CR/FD, both the HT and ESA demodulation methods suffer from the condition where frequency deviation of the IF is large compared to the carrier frequency. Hence the MFT framework produces a benefit to both methods by increasing the CR/FD of the input signal.

Since the modulation index is jointly determined by both the CR/IB and CR/FD, its influence on the demodulation performance is not straightforward. However, the modulation index of the FM signal cannot be too large such that the spectrum of the sidelobes near the origin are significant and begin to incur significant demodulation error as in the case of sinusoidal FM. The theoretical analysis for this phenomenon is provided in the Appendix, where we analyze the demodulation error associated with the Hilbert transform for demonstration. As a result, the error incurred by a large modulation index is unavoidable for any demodulation algorithm and imposes a lower bound on the demodulation error. Even the MFT approach cannot reduce

the error significantly in this extreme wideband scenario.

In conclusion, the proposed MFT approach is more compatible to be combined with the ESA demodulation schemes. It also relaxes the estimation range constraint of the ESA. By compressing and shifting the original input in the frequency domain, we can always choose proper conversion factors and translation frequencies to transform its frequency components into the range between 0 and $\frac{\pi}{2}$, which is required for the ESA to work properly.

In addition, the MFT framework can be extended in a straightforward separable way into multi-dimensional space when combined with the multi-dimensional higher order differential operators [15, 16] recently proposed by Salzenstein *et al.*. Therefore the conclusion also applies to the two-dimensional images as discussed in the next chapter.

## 2.5   Conclusion

The basic MFT system that combines multirate processing and heterodyning to achieve wideband AM–FM demodulation was proposed. Prior work combining these systems was shown to produce impractical designs for large factors, needing bandpass filters of very high orders and very narrow passbands. An alternative MFT framework that interchanges the order of the heterodyne and multirate modules and employs the Noble identities was proposed and demonstrated to reduce the computational burden placed on the bandpass filter for large conversion factors. The MFT serve as the cornerstone for various wideband applications including image demodulation, speech formant estimation, satellite communiations and multicomponent AM-FM demodulation as discussed in subsequent chapters.

# Chapter 3

# Applications to Wideband Image Demodulation

The AM-FM representation model has found various applications with images recently including image analysis, texture processing [17] and fingerprint classification [18]. According to earlier work [19, 20] by Havlicek, Bovik *et al.*, nonstationary images can be modeled as superpositions of multiple AM-FM components,

$$I(x, y) = \sum_{i=1}^{n} a_i(x, y) \cos\left(\phi_i\left(x, y\right)\right). \tag{3.1}$$

The multicomponent AM-FM image is first decomposed by employing a set of bandpass filters such as Garbor filterbanks or via the use of the *bi-dimensional empirical mode decomposition* [21]. Each resulting monocomponent AM-FM image is further demodulated into corresponding IA $a(x, y)$ and IF,

$$\nabla\phi(x, y) = \left[\frac{\partial\phi(x, y)}{\partial x}, \frac{\partial\phi(x, y)}{\partial y}\right]^{T}. \tag{3.2}$$

In particular, the IA depicts the contrast present in the image, while the IF reveals the locally emergent frequency variation. Conventional image demodulation approaches involve 2D extension of the *analytic signal* [22] and multidimensional

*energy separation algorithm* [23] with additional processing techniques such as *dominant component analysis* (DCA) [24, 25]. Recently, the monogenic image approach using the Riesz-Laplace wavelet [26, 27, 28] was also proposed.

However, in most of these approaches, narrowband assumptions were imposed on each AM-FM component of the image. For example, most literature inexplicitly assumes the AM-FM image to be globally wideband yet each of its components to be locally narrowband. In general, both the IA, $a(x, y)$ and the IF, $\nabla\phi(x, y)$ of a single component are assumed to be slowly varying, otherwise the approximations inherent in most demodulation approaches are no longer valid and incur significant error especially under the wideband scenario.

In this chapter, we propose a 2D extension of the MFT approach called the *bi-dimensional multirate frequency transformations* [29] that can be combined with a variety of demodulation techniques to enhance their demodulation performances, traditionally limited by the narrowband constraint on the frequency modulation part of the monocomponent AM-FM image.

## 3.1   Wideband Frequency Modulated Image

Since images are nonseparable in general, we cannot directly extend the related definitions of 1D wideband signal to 2D. For instance, we are not able to define the corresponding modulation index along any specific direction for 2D images in a global sense. However, as proposed by Pattichis and Bovik [30], the complex FM image can be locally approximated by the product of two 1D FM signals. The corresponding 1D signals are defined along the directions of the eigenvectors of the *instantaneous frequency gradient tensor* (IFGT), which is simply the Hessian of the phase. Let $\vec{z} = [z_1, z_2]$ denote the representation under the eigenvector coordinate system, $\widetilde{\phi}(\vec{z})$ denote the phase of the FM image and $\widetilde{\vec{F}}$ denote the IFGT. According to [30], around

a given point $\vec{z}_0 = [a_1, a_2]$ of the image, the local phase is approximated using a Taylor series expansion by

$$\widetilde{\phi}(\vec{z}) \approx \widetilde{\phi}(a_1, a_2) + \widetilde{\phi}_1(z_1) + \widetilde{\phi}_2(z_2), \tag{3.3}$$

where:

$$\widetilde{\phi}_i(z_i) = \frac{\partial \widetilde{\phi}}{\partial z_i}(a_1, a_2)(z_i - a_i) + \frac{\lambda_i(a_1, a_2)}{2}(z_i - a_i)^2, \ \ i = 1, 2. \tag{3.4}$$

Note that $\lambda_1$ and $\lambda_2$ are the eigenvalues of $\widetilde{\vec{F}}$. The complex FM image is then locally approximated by the product of two 1D FM functions defined with respect to the eigenvector coordinate directions:

$$\exp\left[j\widetilde{\phi}(z_1, z_2)\right] \approx \exp\left[j\phi_1(z_1)\right]\exp\left[j\phi_2(z_2)\right]. \tag{3.5}$$

As a result, FM images can be viewed as locally separable, and we can define the wideband FM image locally by deriving the corresponding parameters along the eigenvector coordinate directions. For example, if the modulation index along the directions of the IFGT eigenvectors are sufficiently large around the reference point, the image can be expected to exhibit similar wideband pattern locally, analogous to the 1D case. Assume the modulation index along both directions of the eigenvectors are denoted by $\beta_{z_1}$ and $\beta_{z_2}$ respectively. To locally indicate whether an FM image is wideband or not in the sense similar to the literature in FM communication systems, we can define a local modulation index $\beta_l$ via

$$\beta_l = \left(\beta_{z_1}^2 + \beta_{z_2}^2\right)^{1/2}. \tag{3.6}$$

If $\beta_l$ is sufficiently large, then the corresponding FM image exhibits a wideband pattern locally. This definition of the local modulation index $\beta_l$ is especially effective for the sinusoidal FM images with IF components specified along the horizontal and vertical directions. In this case, the modulation index for each eigenvector direction can be well defined according to the original formulation in the 1D context.

## 3.2   Partial Hilbert Transform Demodulation

The partial or directional *Hilbert transform* based on analytic image as summarized in [31] by Havlicek *et al.*is widely used for monocomponent AM-FM image demodulation.

It can be derived simply from the 1D *Hilbert transfrom*. Let $f(\vec{x}) : \mathbb{R}^n \to \mathbb{R}$ and $\vec{e}_i$ denote the unit vector in the $x_i$ direction. The partial *Hilbert transform* along the direction of $\vec{e}_i$ is then defined by

$$\mathcal{H}_{x_i}\left[f(\vec{x})\right] = \int_{\mathbb{R}} \frac{f(\vec{x} - \xi\vec{e}_i)}{\pi\xi} d\xi. \tag{3.7}$$

Note that the partial *Hilbert transform* defined here is specified along the horizontal and vertical directions.  Thus it is particularly suitable for demodulating images with frequency modulation patterns emerging along the same directions.  Assume that $I(x, y)$ is a monocomponent AM-FM image with frequency modulation patterns emerging along both the horizontal and vertical directions given by

$$I(x, y) = A(x, y)\cos\left(\Omega_x x + \Omega_y y + \varphi(x, y)\right), \tag{3.8}$$

where $\Omega_x$ and $\Omega_y$ denote the carrier (mean) frequencies along the $x$ and $y$ directions respectively.  Under this assumption, according to (1.16), we are able to approximate the partial *Hilbert transform* along the $x$ axis by fixing variable $y$ via

$$\mathcal{H}_x\left[I(x, y)\right] \approx A(x, y)\sin\left(\Omega_x x + (\Omega_y y + \varphi(x, y))\right). \tag{3.9}$$

This approximation holds only under certain conditions similar to the 1D case.  First of all, the $A(x, y)$ should be slow-varying and narrowband.  Analogous to (1.17), $\Omega_x$ is required to be sufficiently large such that

$$\Omega_x \gg \frac{\partial\varphi(x, y)}{\partial x}. \tag{3.10}$$

A similar condition applies to the $y$ axis as well. As a result, we can easily obtain the approximations for the IA and the IF components with respect to both directions according to the 2D extensions of (1.19) and (1.20).

Synthetic examples are provided to illustrate the efficacy of the proposed approach. Since the synthetic images we present are predefined with its frequency modulation patterns along horizontal and vertical directions, the partial *Hilbert transform* is prefered over the monogenic image approach to serve as the comparison technique to the proposed approach due to the following reasons: 1) the demodulation performance of the partial *Hilbert transform* is as competitive as the Riesz transform when the image display frequency modulation patterns along the same specified directions, 2) the partial *Hilbert transform* can be implemented easily via the one-dimensional *Hilbert transform*.

## 3.3   Higher Order Energy Operator

A variety of methods based on the multidimensional energy operator [23] proposed by Maragos and Bovik are also widely used for AM-FM image demodulation. An image demodulation algorithm based on higher order Teager-Kaiser operators is proposed in the recent work [32, 33, 34] by Salzenstein, Diop and Boudraa. They have been reported to provide better performance for narrowband AM-FM images than the classical 2D-ESA [35].

The $k$-order *differential energy operator* (DEO) [35] in 1D for any given signal $s(t)$ is defined by

$$\Psi_k[s(t)] = \frac{\partial s(t)}{\partial t}\frac{\partial^{k-1}s(t)}{\partial t^{k-1}} - s(t)\frac{\partial^k s(t)}{\partial t^k}, \tag{3.11}$$

where $\Psi_2$ refers to the commonly used Teager-Kaiser energy operator.

In the case of discrete-time, for a given image $I(k, l)$, the higher order demodulation algorithm (DHODA) [32, 33, 34] can be summarized via:

$$I_1(k, l) = \frac{1}{2}[I(k+1, l) - I(k-1, l)], \tag{3.12}$$

$$I_2(k, l) = \frac{1}{2}[I(k, l+1) - I(k, l-1)], \tag{3.13}$$

$$I_{12}(k, l) = \frac{1}{2}[I_2(k+1, l) - I_2(k-1, l)], \tag{3.14}$$

$$\Psi_2[I(k, l)] = \{2[I(k, l)]^2 - I(k-1, l)I(k+1, l) - I(k, l-1)I(k, l+1)\}$$
$$+ 2[I_1(k, l)I_2(k, l) - I(k, l)I_{12}(k, l)], \tag{3.15}$$

$$I_{12}^1(k, l) = \frac{1}{2}[I_{12}(k+1, l) - I_{12}(k-1, l)], \tag{3.16}$$

$$I_{12}^2(k, l) = \frac{1}{2}[I_{12}(k, l+1) - I_{12}(k, l-1)], \tag{3.17}$$

$$|\widehat{a}(k, l)| = \left(\frac{\Psi_2[I_1(k, l)]\Psi_2[I_2(k, l)]}{\Psi_2[I_{12}^1(k, l) + I_{12}^2(k, l)]}\right)^{1/2}, \tag{3.18}$$

$$|\widehat{\Omega}_1(k, l)| = \arcsin\left(\left(\frac{\Psi_2[I_{12}(k, l)]}{\Psi_2[I_2(k, l)]}\right)^{1/2}\right), \tag{3.19}$$

$$|\widehat{\Omega}_2(k, l)| = \arcsin\left(\left(\frac{\Psi_2[I_{12}(k, l)]}{\Psi_2[I_1(k, l)]}\right)^{1/2}\right), \tag{3.20}$$

where $\widehat{a}(k, l)$ is the IA estimation while $\widehat{\Omega}_1(k, l)$ and $\widehat{\Omega}_2(k, l)$ are the IF estimations along the spatial axies of the image. Note that the IF estimations are obtained through the inverse sine function, indicating that the demodulation approaches based on the ESA can only estimate IF components that range between 0 and $\frac{\pi}{2}$, or in other words, up to one fourth of the sampling frequency. Moreover, the demodulation approaches based on the ESA also suffer from the narrowband constraint as the *Hilbert transform*. Both the IA and the IF waveforms may not vary too fast or too greatly in value.

## 3.4 Bi-dimensional Multirate Frequency Transformations

The MFT that performs wideband-to-narrowband conversion was proposed in the previous section. It primarily increases the CR/IB and CR/FD ratios of the original signal to improve the demodulation performance of the conventional demodulation approaches. The *bi-dimensional multirate frequency transformations* (BMFT) is derived by generalizing the underlying idea to 2D. For simplicity, assume that the input is a monocomponent wideband FM image of the form

$$J(x, y) = A \cos\left(\phi(x, y)\right). \tag{3.21}$$

It is first compressed in frequency domain by appropriate factors $\vec{R} = diag[R_x, R_y]$, which corresponds to spatial expansion given by

$$J_1(x, y) = A \cos\left(\phi\left(\frac{x}{R_x}, \frac{y}{R_y}\right)\right). \tag{3.22}$$

Then we heterodyne the resultant image by a frequency translation vector $\vec{\Omega} = [\Omega_x, \Omega_y]$ via

$$
\begin{aligned}
J_2(x, y) =& J_1(x, y) \cos(\Omega_x x) \cos(\Omega_y y) \\
=& \frac{A}{2} \cos\left(\Omega_x x + \phi\left(\frac{x}{R_x}, \frac{y}{R_y}\right)\right) \cos(\Omega_y y) \\
& + \frac{A}{2} \cos\left(\Omega_x x - \phi\left(\frac{x}{R_x}, \frac{y}{R_y}\right)\right) \cos(\Omega_y y) \\
=& \frac{A}{4} \cos\left(\Omega_x x + \Omega_y y + \phi\left(\frac{x}{R_x}, \frac{y}{R_y}\right)\right) \\
& + \frac{A}{4} \cos\left(\Omega_x x - \Omega_y y + \phi\left(\frac{x}{R_x}, \frac{y}{R_y}\right)\right) \\
& + \frac{A}{4} \cos\left(\Omega_x x + \Omega_y y - \phi\left(\frac{x}{R_x}, \frac{y}{R_y}\right)\right) \\
& + \frac{A}{4} \cos\left(\Omega_x x - \Omega_y y - \phi\left(\frac{x}{R_x}, \frac{y}{R_y}\right)\right).
\end{aligned} \tag{3.23}
$$

After the heterodyning, a 2D bandpass filter is employed to extract the desired high-frequency term through

$$
\begin{aligned}
\widetilde{J}(x,y) &= J_2(x,y) * h_{BP}(x,y) \\
&\approx \frac{A}{4} \cos\left(\Omega_x x + \Omega_y y + \phi\left(\frac{x}{R_x}, \frac{y}{R_y}\right)\right) \\
&= \frac{A}{4} \cos\left(\widetilde{\phi}(x,y)\right).
\end{aligned}
\tag{3.24}
$$

Let us assume that the support of the original image spectrum $J(\omega_1, \omega_2)$ is within the range $\omega_i \in [-\Omega_i, \Omega_i]$, $i = 1, 2$. The 2D bandpass filter should be carefully designed with its passband range given by

$$
|\omega_1| \in [\Omega_x, \Omega_x + \frac{\Omega_1}{R_x}], \quad |\omega_2| \in [\Omega_y, \Omega_y + \frac{\Omega_2}{R_y}].
\tag{3.25}
$$

On one hand, the carrier (or mean) frequencies of the modulation in both dimensions are increased via the frequency translation vector $\vec{\Omega} = [\Omega_x, \Omega_y]$. On the other hand, the bandwidth of the modulating image is reduced by the appropriate conversion factors $\vec{R} = diag[R_x, R_y]$. These two benefits derived via the BMFT are crucial to improving the IF demodulation due to the following reasons: 1) a majority of demodulation approaches require the input to have high CR/IB in both dimensions, 2) the CR/FD of the input must be sufficiently large in both dimensions such that the deviations of the IF components can be tolerated in accordance with the narrowband assumption of the input.

Then we can recover the IF of the input image from the IF estimation of the transformed image $\widetilde{J}(x,y)$ via the inverse BMFT. Assume that the IF components of $J_(x,y)$ and $\widetilde{J}(x,y)$ are given by

$$
\Omega_1(x,y) = \frac{\partial \phi(x,y)}{\partial x}, \quad \Omega_2(x,y) = \frac{\partial \phi(x,y)}{\partial y},
\tag{3.26}
$$

$$
\widetilde{\Omega}_1(x,y) = \frac{\partial \widetilde{\phi}(x,y)}{\partial x}, \quad \widetilde{\Omega}_2(x,y) = \frac{\partial \widetilde{\phi}(x,y)}{\partial y}.
\tag{3.27}
$$
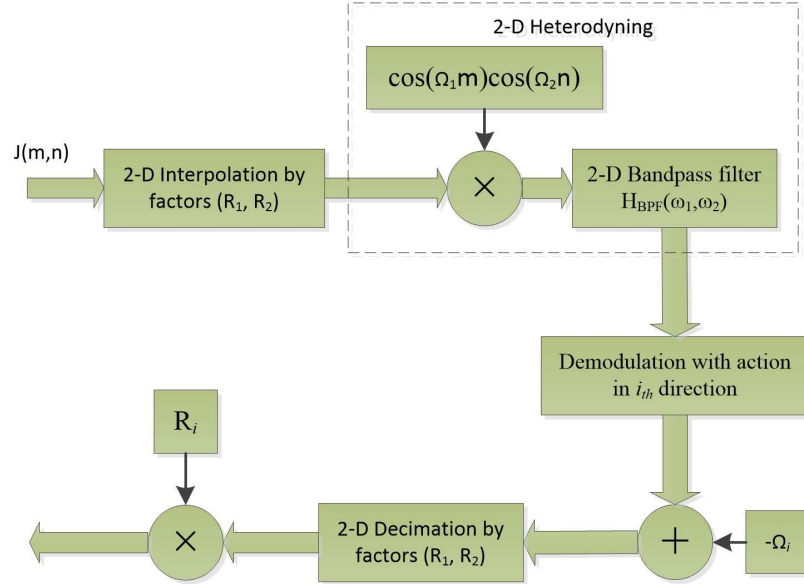
Figure 3.1: Block diagram of the Bi-dimensional Multirate Frequency Transformations.

---

The IF estimation for input image is recovered through the inverse BMFT relation (see Fig. 3.1) via

$$\Omega_1(x, y) = R_x \left( \widetilde{\Omega}_1(R_x x, R_y y) - \Omega_x \right), \tag{3.28}$$

$$\Omega_2(x, y) = R_y \left( \widetilde{\Omega}_2(R_x x, R_y y) - \Omega_y \right), \tag{3.29}$$

where $\widetilde{\Omega}_1(R_x x, R_y y)$ and $\widetilde{\Omega}_2(R_x x, R_y y)$ represent spatial compression (or frequency expansion) of the IF estimation for the transformed image $\widetilde{J}(x, y)$.

In order to implement the BMFT in discrete time, we replace compression and expansion in frequency domain by their discrete equivalences. Note that the compression in the frequency domain corresponds to interpolation while the expansion corresponds to decimation. As a result, the block diagram of the BMFT demodulation approach is depicted in Fig 3.1. The BMFT is implemented through discrete-time operations of interpolation, heterodyning, and bandpass filtering. The transformed FM image is then demodulated via a monocomponent demodulation approach and

the original IF components are recovered from the IF estimation of the transformed image via the inverse BMFT. However, discrete-time operations in 2D may not be extended in a straightforward way from their 1D counterpart. It primarily depends on whether the input image is separable or not. If the image is separable, each operation of the BMFT can be implemented by simply cascading its 1D equivalence with action along each dimension. If this is not the case, we need to pay attention to a few issues associated with the non-separable nature of the image, which are often intractable. Therefore, deriving 2D operations directly from their 1D realizations is favored in terms of its simplicity for practical implementation.

In particular, the frequency upshift in heterodyning is not uniquely defined in 2D. This can be realized either through the product of two separable cosine terms $\cos{(\Omega_1 m)}\cos{(\Omega_2 n)}$ in each dimension as in Fig. 3.1 or via just one cosine term $\cos{(\Omega_1 m + \Omega_2 n)}$ in the diagonal direction. These two different realizations will lead to different designs of the 2D bandpass filters. For the first case, we only need to bandpass the outermost quarter of the frequency spectrum in each quadrant to extract the desired high frequency term expressed in (3.24). This is simply achieved using a separable 2D bandpass filter, by cascading two 1D bandpass filters with action in each dimension. For the latter case, where the frequency upshift is diagonal, it is much more complicated to achieve the same objective with a realizable design of the 2D bandpass filter. The issue is that the 2D frequency upshift in the diagonal direction results in only two copies of the original spectrum in the first and third quadrant respectively. However, the design of a 2D bandpass filter whose passband is only present in two quadrants and also capable of extracting the desired term in (3.24) is computationally complex.

In addition, the BMFT framework can be directly applied on a monocomponent AM-FM image provided that the IA of the given image is slowly varying, which is inherently assumed by most image demodulation algorithms. Under this constraint,

the IA can be approximated via the inverse BMFT relation by

$$A(x, y) = \widetilde{A}\left(R_x x, R_y y\right), \tag{3.30}$$

where $\widetilde{A}(x, y)$ denotes the demodulated IA of the transformed image $\widetilde{J}(x, y)$. Numerical results are given later in this chapter to support this claim.

## 3.5  Example of Wideband Image Demodulation

In this section, we present numerical results for the proposed BMFT-energy approach. We begin with an example of a synthetic sinusoidal AM-FM image as illustrated in Fig. 3.5. The expression of this synthetic image and its corresponding IF components are given by

$$f(m, n) = \left[1 + 0.5 \cos\left(\frac{\pi}{50}m + \frac{\pi}{30}n\right)\right] \cos\left(\frac{\pi}{5}m + \frac{\pi}{3}n + 6\sin\left(\frac{\pi}{50}m + \frac{\pi}{2}\right) + \right. \\ \left. 5\sin\left(\frac{\pi}{30}n\right)\right), \tag{3.31}$$

$$\Omega_1(m, n) = \frac{\partial \phi(m, n)}{\partial m} = \frac{\pi}{5} + \frac{3\pi}{25} \cos\left(\frac{\pi}{50}m + \frac{\pi}{2}\right), \tag{3.32}$$

$$\Omega_2(m, n) = \frac{\partial \phi(m, n)}{\partial n} = \frac{\pi}{3} + \frac{\pi}{6} \cos\left(\frac{\pi}{30}n\right). \tag{3.33}$$

Note that the frequency modulation index is 6 along the horizontal direction and 5 along the vertical direction, both of which are sufficiently large. We can also easily check that the CR/IB and the CR/FD are small along both directions for this wideband sinusoidal example.

The demodulation results via the DHODA combined with the proposed BMFT framework using conversion factors of $\vec{R} = diag[8, 8]$ are given by Fig. 3.3. It is compared with the demodulation via the DHODA alone. Note that the demodulated IF with respect to either the horizontal or the vertical direction exhibits a sinusoidal
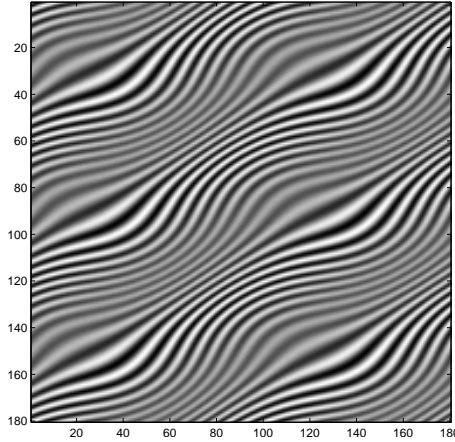
Figure 3.2: Synthetic sinusoidal AM-FM image with wideband FM components of the form:
$$f(m,n) = \left[1 + 0.5\cos(\tfrac{\pi}{50}m + \tfrac{\pi}{30}n)\right]\cos\left(\tfrac{\pi}{5}m + \tfrac{\pi}{3}n + 6\sin(\tfrac{\pi}{50}m + \tfrac{\pi}{2}) + 5\sin(\tfrac{\pi}{30}n)\right).$$

pattern along that direction, which can be inferred from (3.31). We can easily observe that both the IA and the IF obtained via the BMFT-DHODA are smoother than that via the DHODA alone. The MSE for the IA is reduced from 0.8910 to 0.0649 and the *root mean square error* (RMSE) for the IF is reduced significantly from 1.83% to 0.25% through the BMFT framework. Here we define the RMSE as the $l_2$ norm of the difference between the true IF, $\nabla\phi(x,y)$ and the estimated IF, $\nabla\widehat{\phi}(x,y)$ against the $l_2$ norm of the true IF itself via

$$\text{RMSE} = \frac{\|\nabla\phi(x,y) - \nabla\widehat{\phi}(x,y)\|_{l_2}}{\|\nabla\phi(x,y)\|_{l_2}} \times 100\%. \tag{3.34}$$

In fact, the demodulation error can be further reduced if larger multirate conversion factors are applied, similar to the 1D case of MFT. As shown in Table 3.5, the demodulation errors of the proposed BMFT framework using DHODA with conversion factors of $\vec{R} = diag[8,8]$ (BMFT-DHODA-8) and with conversion factors of $\vec{R} = diag[16,16]$ (BMFT-DHODA-16) are compared with the DHODA alone. The RMSE can be further reduced to 0.18% via the BMFT with conversion factors $\vec{R} = diag[16.16]$, achieving an error reduction of 10 times compared with the

36
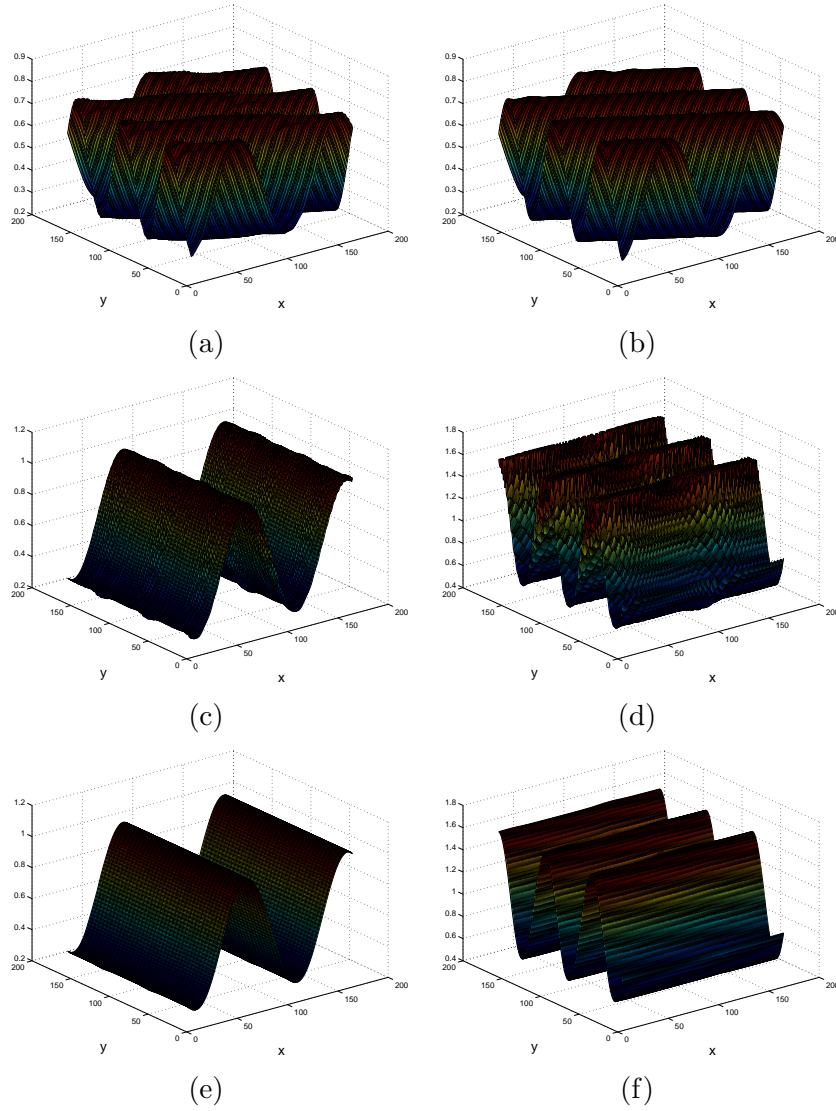
Figure 3.3: Perspective plot of the demodulation results for the wideband sinusoidal example via the DHODA and via the BMFT-DHODA with conversion factors $\vec{R} = diag[8, 8]$, (a) estimation of the IA via the DHODA, (b) estimation of the IA via the BMFT-DHODA, (c) estimation of the IF component along the horizontal direction via the DHODA, (d) estimation of the IF component along the vertical direction via the DHODA, (e) estimation of the IF component along the horizontal direction via the BMFT-DHODA, (f) estimation of the IF component along the vertical direction via the BMFT-DHODA.

| | $\text{var}(\frac{a-\widehat{a}}{a})$ | $\text{MSE}(a,\widehat{a})$ | $\text{var}(\frac{\Omega_1-\widehat{\Omega}_1}{\Omega_1})$ | $\text{var}(\frac{\Omega_2-\widehat{\Omega}_2}{\Omega_2})$ |
|---|---|---|---|---|
| DHODA | 4.7876 | 0.8910 | 2.6688 | 8.0073% |
| BMFT-DHODA-8 | 0.2934 | 0.0649 | 0.3554 | 0.1281% |
| BMFT-DHODA-16 | 0.1677 | 0.0412 | 0.5608 | 0.0813% |
| HTDA | 0.5125 | 0.0942 | 0.3715 | 0.0683% |
| BMFT-HTDA-8 | 0.2930 | 0.0646 | 0.1228 | 0.0796% |
| | $\text{MSE}(\Omega_1,\widehat{\Omega}_1)$ | $\text{MSE}(\Omega_2,\widehat{\Omega}_2)$ | RMSE | |
| DHODA | 0.6948 | 12.9626 | 1.83% | |
| BMFT-DHODA-8 | 0.0427 | 0.1310 | 0.25% | |
| BMFT-DHODA-16 | 0.0419 | 0.0701 | 0.18% | |
| HTDA | 0.1276 | 0.0430 | 0.23% | |
| BMFT-HTDA-8 | 0.0235 | 0.0716 | 0.20% | |

Table 3.1: Comparison of the demodulation errors

DHODA alone. The use of larger factors results in a much narrower passband for digital filters that are difficult to realize in practice due to a sharper transition band. To implement such FIR filters, a very large filter order is required, which may not be acceptable in terms of the desired system complexity. Therefore the choice of multirate conversion factors should be weighted between the tolerance for demodulation error and that of the system complexity. Considering the alternative MFT framework that enables large conversion factors as proposed in chapter 2, we may further reduce the demodulation error by exploring even larger factors.

The demodulation results via the partial *Hilbert transform* alone (HTDA) and the BMFT and the partial *Hilbert transform* combination with conversion factors $\vec{R} = diag[8,8]$ (BMFT-HTDA-8) are also illustrated in Table 3.5. We can observe that the error reduction via the BMFT framework for the HT demodulation is not as obvious as for the ESA demodulation, which justifies our discussion in the previous section.

In the second experiment, we test the BMFT and the DHODA combination on a wideband sinusoidal FM image whose IF components are out of the range $[0, \frac{\pi}{2}]$,

which is the estimation range constraint for demodulation approaches based on the ESA. As shown in Fig. 3.5, the synthetic image for this example and its corresponding IF components are given by:

$$f(m,n) = 0.5\cos(\frac{2\pi}{3}m + \frac{2\pi}{3}n + 6\sin\left(\frac{\pi}{50}m + \frac{\pi}{2}\right) + 5\sin\left(\frac{\pi}{30}n\right)), \tag{3.35}$$

$$\Omega_1(m,n) = \frac{\partial\phi(m,n)}{\partial m} = \frac{2\pi}{3} + \frac{3\pi}{25}\cos\left(\frac{\pi}{50}m + \frac{\pi}{2}\right), \tag{3.36}$$

$$\Omega_2(m,n) = \frac{\partial\phi(m,n)}{\partial n} = \frac{2\pi}{3} + \frac{\pi}{6}\cos\left(\frac{\pi}{30}n\right). \tag{3.37}$$
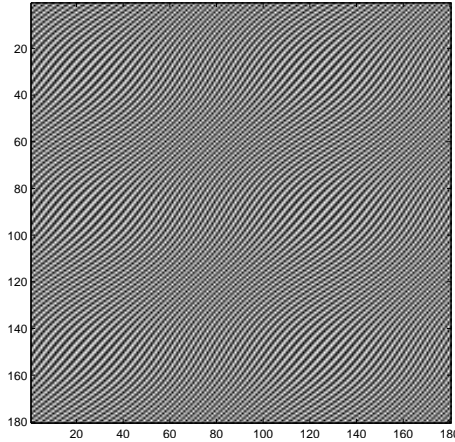


Figure 3.4: Wideband sinusoidal FM image with IF components outside the range of ESA constraint $[0 \ \frac{\pi}{2}]$ in the form:
$f(m,n) = 0.5\cos\left(\frac{2\pi}{3}m + \frac{2\pi}{3}n + 6\sin\left(\frac{\pi}{50}m + \frac{\pi}{2}\right) + 5\sin\left(\frac{\pi}{30}n\right)\right).$

Note that the carrier frequency along each direction is $\frac{2\pi}{3}$. Hence the IF components of the image are out of the range $[0, \frac{\pi}{2}]$. The demodulation results for the DHODA and the BMFT-DHODA are illustrated in Fig. 3.5. As we can observe, the demodulated IF components via the DHODA are seriously distorted in both amplitudes and phases, whereas the demodulated IF components via the BMFT-DHODA are smooth and valid. The RMSE for the DHODA is as large as 56.74% while the RMSE for the BMFT-DHODA is merely 0.9113%. Hence our claim that the BMFT
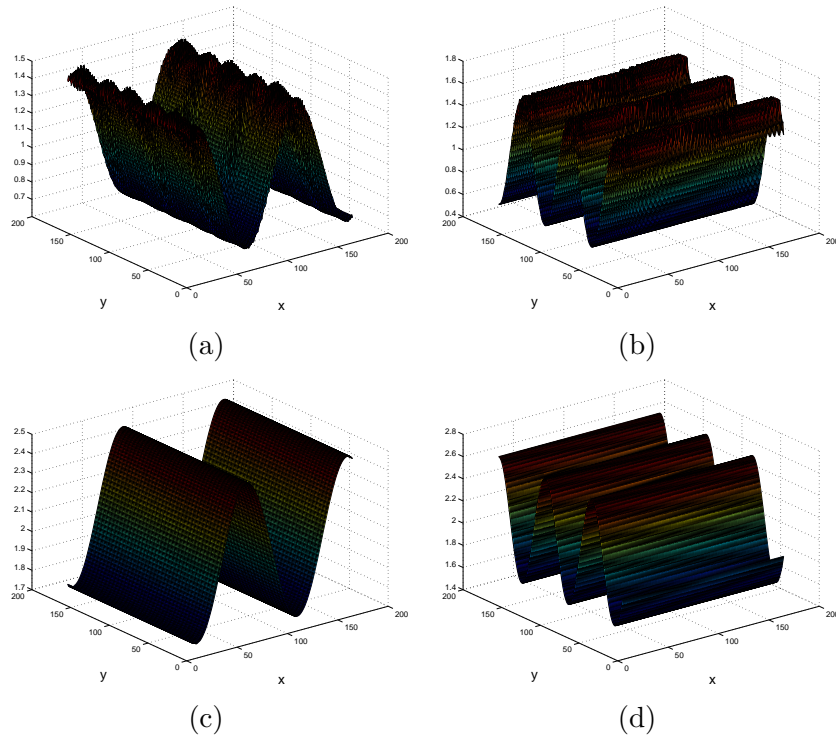
Figure 3.5: Perspective plot of the IF estimation for the wideband sinusoidal example where IF components are out of the range $[0, \frac{\pi}{2}]$ via the DHODA and via the BMFT-DHODA with conversion factors $\vec{R} = diag[8, 8]$, (a) estimation of the IF component along the horizontal direction via the DHODA, (b) estimation of the IF component along the vertical direction via the DHODA, (c) estimation of the IF component along the horizontal direction via the BMFT-DHODA, (d) estimation of the IF component along the vertical direction via the BMFT-DHODA. Note that the IF estimation via the DHODA is seriously distorted in both its amplitude and phase.

helps overcome the range constraint of demodulation algorithms based on ESA is justified.

In the third experiment, we justify the efficacy of the proposed BMFT approach on real images. The real image of an oakring (photo by H.D. Grissino-Mayer) is shown by Fig. 3.6(a). The demodulation results associated with this real oakring image are compared in Fig. 3.6. The estimated IF components via the DHODA, as shown in Fig. 3.6(b) and Fig. 3.6(c), have singular points with significantly large values.

The estimated IF components via the BMFT-DHODA, as shown in Fig. 3.6(d) and Fig. 3.6(e), are much smoother and do not have any singular point with significantly large value. As a result, the estimated IF needle plot of the DHODA (Fig. 3.6(f)) is seriously distorted by those singular points, whereas the estimated IF needle plot of the BMFT-DHODA (Fig. 3.6(g)) reveals the ring pattern corresponding to the real oakring image. Note that the IA estimated via the DHODA (Fig. 3.6(h)) suffers from the same issue whereas the IA estimated via the BMFT-DHODA (Fig. 3.6(i)) does not. Based on this observation, we see that ESA demodulation such as that employed by the DHODA prohibits its direct application to wideband real images due to the narrowband and estimation range constraints, while the proposed BMFT framework overcomes such constraints and allows for direct application of ESA demodulation to wideband real images.

## 3.6   Wideband Fingerprint Demodulation

The joint approach has been further applied to wideband fingerprint images demodulation [36] to demonstrate its efficacy. Since fingerprint is the pattern of furrows and rigdes on the surface of a fingertip, the AM-FM models are well-suited for modeling the fingerprint pattern where the AM component describes the ridge intensity and the FM component accounts for the ridge variation.

Prior to demodulation, the *dominant component analysis* (DCA) is applied to the fingerprint image to obtain its fundamental AM-FM component. The goal of DCA is to estimate the modulating functions at each pixel that dominates the frequency spectrum. It captures the local oscillations of the AM-FM image by pointwise choosing the output with maximum response among the channels based on certain criterion. The procedure of the DCA implementation [37] illustrated in Fig. 3.7 is summarized via

- Filter the fingerprint image $t(\mathbf{x})$ via a set of Gabor channel filters with response $g_i$, we obtain different channel output images $t_i = t * g_i$. The Garbor filterbank is radially arranged in frequency domain with different scales.

- Based on the filter selection criterion:

$$\Psi_i(\mathbf{x}) = \frac{t_i(\mathbf{x})}{\max_{\mathbf{\Omega}} |G_i(\mathbf{\Omega})|}, \tag{3.38}$$

  where $G_i(\mathbf{\Omega})$ is the frequency response of the $i^{th}$ channel filter. According to Havlicek, $\Psi_i(\mathbf{x})$ can be approximated by the crude estimate of the amplitude modulation function of the component that dominates the channel response of the $i^{th}$ channel at the corresponding pixel. Since the instantaneous amplitude $a_i(x_1, x_2)$ for the corresponding channel can be estimated using conventional demodulation techinques such as ESA, we select the channel with the maximum amplitude estimate as the output pixelwise, which is indexed by k via

$$k = \operatorname*{argmax}_{i} a_i(x_1, x_2). \tag{3.39}$$

- Smooth the resultant fundamental AM-FM component using 2D smoothing filters to reduce the noise.

The BMFT approach proposed in this chapter is then applied to demodulate the fundamental AM-FM component extracted from the fingerprint image via the DCA. A separable FIR bandpass filter with 1025 taps in each direction are employed in the BMFT heterodyne module. Separable heterodyning along with separable bandpass filters are chosen to reduce the complexity of the BMFT system. A multichannel Gabor fllterbank with eight rays or orientations and nine radial frequencies per ray are used to isolate the fundamental FM component. A 3 by 3 Gaussian filter is used to smooth the DCA image prior to demodulation. Residual low frequencies appear as a background in the DCA image. Figure 3.8(a) describes a fingerprint image and Fig. 3.8(b) describes the fundamental AM–FM component extracted

using the Gabor filterbank. Figure 3(b,c) depicts the IF needle plot associated with the application of the directional Hilbert transform to the original image and the fundamental FM component. The needle plot of the IF of the dominant AM-FM component shows that the fundamental AMFM component has better defined ridges in comparison to direct demodulation of the original image. Figure 4 (a,b) depict the application of the BMFT approach with directional Hilbert transform demodulation of the dominant component for conversion factors of 8 and 16 respectively. The IF needle plots further depict a significant improvement in the IF needle plots in comparison to the results without the BMFT.

## 3.7 Conclusion

In this chapter, we have formally defined the notion of locally wideband FM images and the corresponding local modulation index. We extended the 1D approach combining multirate frequency transformations and energy demodulation to 2D and images by using their separable counterparts. The proposed algorithm was applied to both synthetic and real images and shown to produce significant reduction in the demodulation errors. In particular, it was demonstrated to be effective in cases where the DHODA is limited in terms of estimation range. Further we apply the BMFT-DHODA combination to wideband fingerprint images by utilizing the fundamental AM-FM component extracted via the DCA to prove its efficacy.

(a)



(b)



(c)



(d)



(e)



(f)



(g)



(h)



(i)
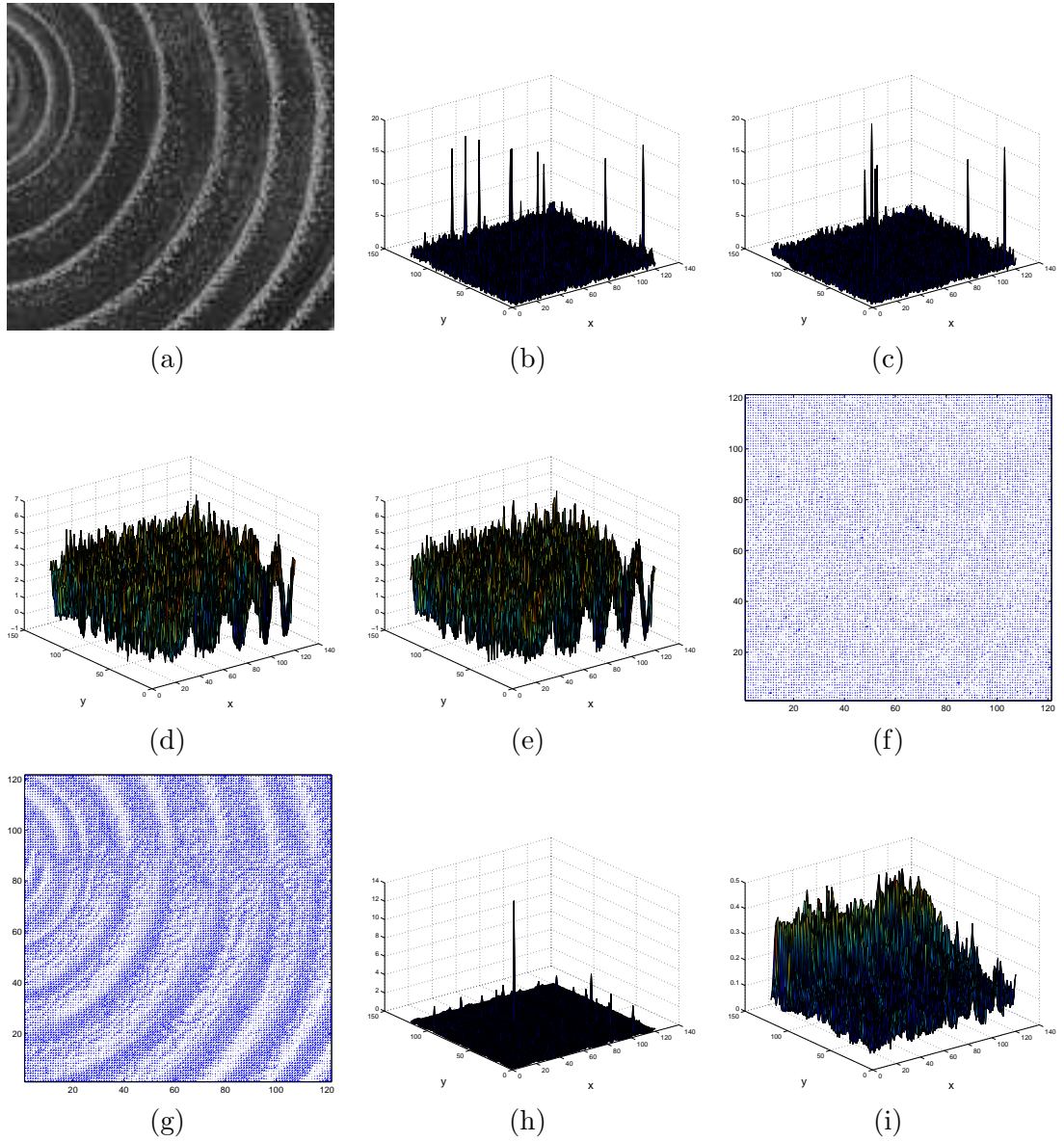
Figure 3.6: Demodulation of the oakring image (photo by H.D. Grissino-Mayer). (a) Real image of the oakring, (b) estimation of the IF component along the horizontal direction via the DHODA, (c) estimation of the IF component along the vertical direction via the DHODA, (d) estimation of the IF component along the horizontal direction via the BMFT-DHODA, (e) estimation of the IF component along the vertical direction via the BMFT-DHODA, (f) estimation of the IF needle plot via the DHODA. (g) estimation of the IF needle plot via the BMFT-DHODA. (h) estimation of the IA via the DHODA. (i) estimation of the IA via the BMFT-DHODA. Note that the conversion factors of the BMFT-DHODA here are $\vec{R} = diag[16, 16]$.
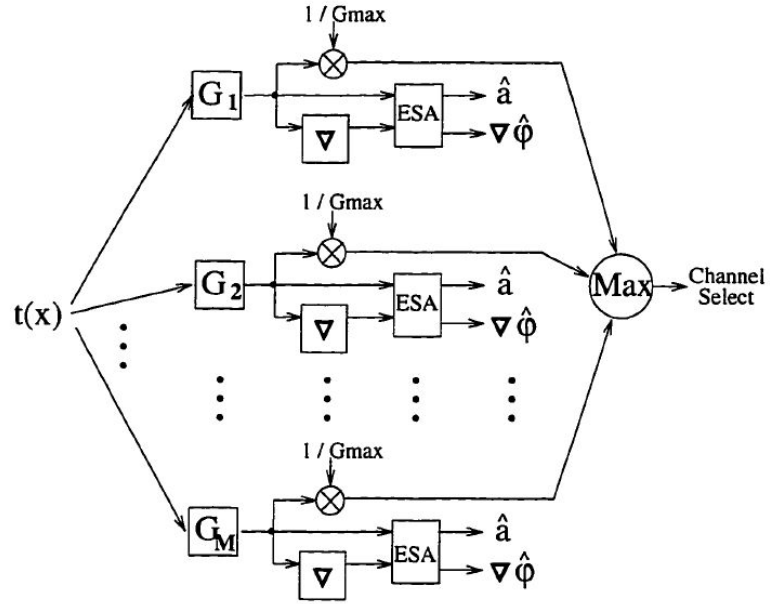
Figure 3.7: Block diagram of DCA.
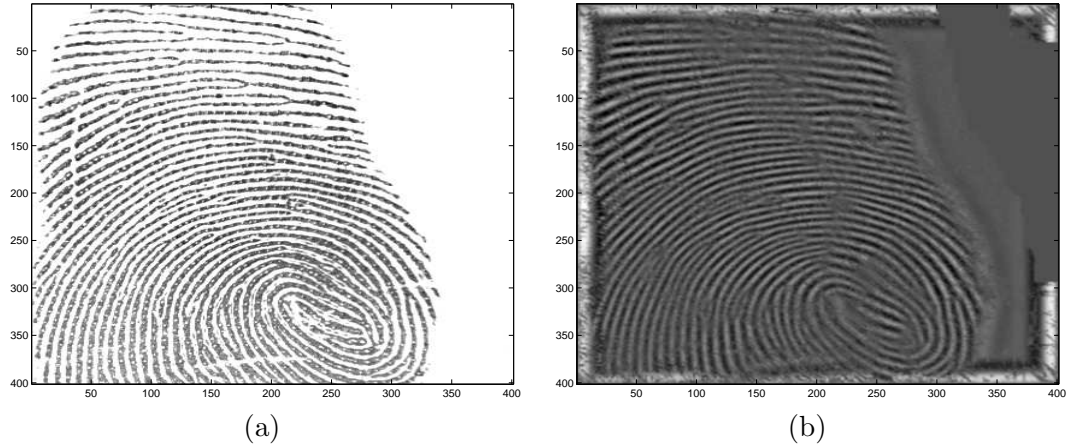


Figure 3.8: Fundamental AM-FM component. (a) Original fingerprint image. (b) Fundamental AM-FM component extracted via DCA.

(a)



(b)



(c)
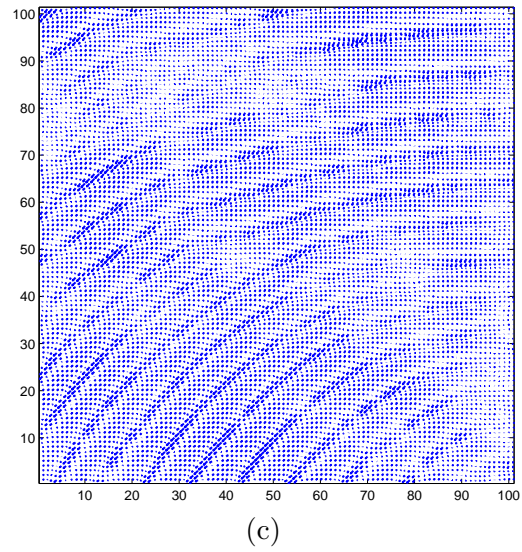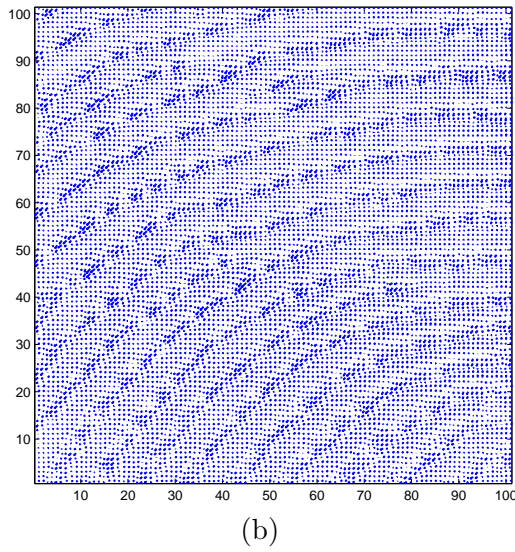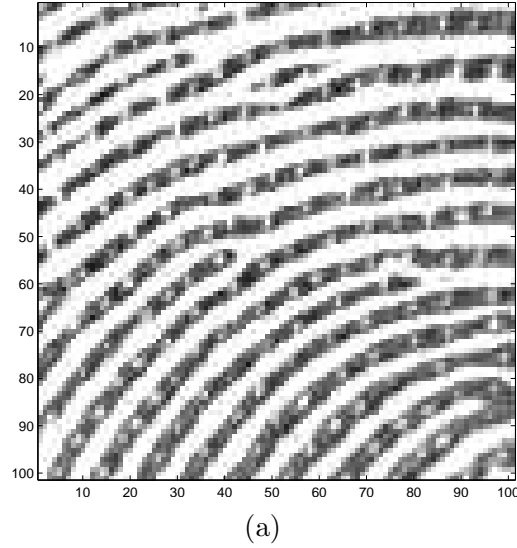
Figure 3.9: (a) Original zoomed in fingerprint image. (b) IF needle plot of the original fingerprint via directional Hilbert transorm. (c) IF needle plot of the estimated fundamental AM-FM component via directional Hilbert transform.
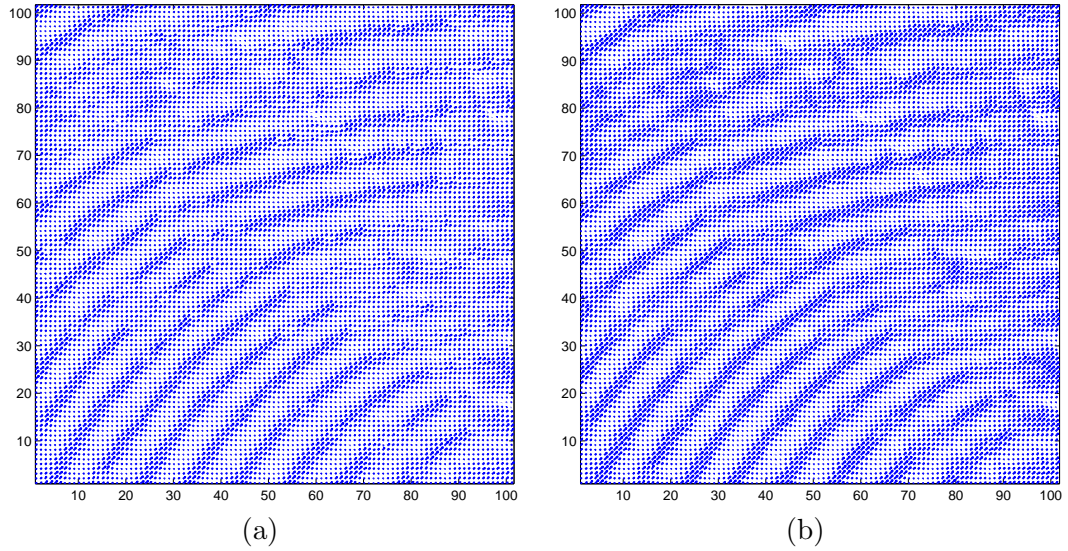
Figure 3.10: (a) IF needle plot of the estimated fundamental AM-FM component via the BMFT with factors $[8, 8]$ and the DHODA. (b) IF needle plot of the estimated fundamental AM-FM component via the BMFT with factors $[16, 16]$ and the DHODA.

# Chapter 4

# Applications to Large Deviation First Formant Estimation

Formants are natural resonances of the vocal tract that are closely related to the vocal tract geometry as a function of the velum, the lips, the jaw and the tongue. They are visually observed as the resonance peaks in the spectrum of the voiced speech. The center-frequency and bandwidth of the formant associated with different vowels differ in a number of ways. Since formant location is an important cue for speech recognition, the estimation of its center-frequency and bandwidth has significant implications in various speech applications. Existing formant estimation approaches are generally based on the LPC assuming that each formant is merely a narrowband AM component. In this regard, LPC is a parametric approach that does not model the spectral valleys properly, hence incapable of handling formants with considerable amount of frequency modulation. For example, the center-frequency of the first formant of many vowels is only around 500 Hz. Such formants are expected to have a large *bandwidth-to-center-frequency ratio* (BW/CF) due to its inherent small CR/IB and CR/FD as introduced in chapter 1, resulting in significant amount of frequency modulation that can be categorized as large deviation formants. Other approaches

using AM-FM representation of speech signals estimate the center-frequency and bandwidth of the formants from their demodulated IF and IA. Hence it is straightforward to improve the demodulation performance for these large deviation formants by employing the proposed MFT framework.

In this chapter, we first apply the *empirical mode decomposition* (EMD) to extract the large deviation first formant, and then perform the demodulation to obtain the IA and IF via the MFT-ESA combination. Finally we conclude that the estimates of the first formant based on the proposed method are more precise than the LPC estimates that only assume narrowband AM for these large deviation formants.

## 4.1 Formant Estimation via LPC

Formant estimation based on *linear predictive coding* [38] is widely used in acoustics and speech processing. The LPC has been the dominant approach for parameter estimation of the discrete-time speech model such as pitch, short-time spectra and formant. Based on the source-filter theory, the basic discrete-time model for speech production is an all-pole filter representing the composite spectrum effects of the vocal tract

$$H(z) = \frac{1}{1 - \sum_{k=1}^{p} a_k z^{-k}},$$

(4.1)

the speech sample $\hat{s}_n$ can be predicted via an autoregressive (AR) predictor given by

$$\hat{s}_n = \sum_{k=1}^{p} a_k s_{n-k},$$

(4.2)

where $p$ is the order of the predictor and $a_k$ denote the coefficients. The prediction error is defined as

$$e_n = s_n - \hat{s}_n = s_n - \sum_{k=1}^{p} a_k s_{n-k}.$$

(4.3)

The goal of the LPC is hence simplified as to find the predictor coefficients that minimize the mean-squared prediction error over a short segment of the speech. By windowing the speech signal over a finite interval given by

$$s_n = \begin{cases} \text{speech sample}, \ 0 \le n \le N - 1 \\ 0, \ \text{otherwise} \end{cases}.$$

$$\text{(4.4)}$$

The MSE to be minimized can be expressed by

$$E = \sum_{n=-\infty}^{\infty} \left( \sum_{k=0}^{p} a_k s_{n-k} \right)^2,$$

$$\text{(4.5)}$$

where $a_0 = 1$. By By setting the partial derivative of $E$ with respect to $a_i$ to zero, we obtain a set of equations

$$\frac{\partial E}{\partial a_i} = 2 \sum_{k=0}^{p} a_k \sum_{n=-\infty}^{\infty} s_{n-k} s_{n-i} = 2 \sum_{k=0}^{p} a_k r_{k,i} = 0,$$

$$\text{(4.6)}$$

where $r_{k,i}$ is defined by

$$r_{k,i} = \sum_{n=-\infty}^{\infty} s_{n-k} s_{n-i} = \sum_{n=-\infty}^{\infty} s_n s_{n-(k-i)}.$$

$$\text{(4.7)}$$

Since the term $r_{k,i}$ only depends on the difference of the index $k-i$, we have $r_{k,i} = r_{k-i}$ as the autocorrelation function of the speech signal $s_n$, which is also symmetric by its definition. By taking account of the windowed speech samples, the summation limit for calculating the autocorrelation is between $0$ and $N + p - 1$. Hence Eq. 4.6 can be written as

$$\sum_{k=1}^{p} a_k r_{k-i} = -r_i.$$

$$\text{(4.8)}$$

Given the computed autocorrelation function, the coefficients $a_k$ of the LPC polynomial can be obtained by solving the set of linear equations via the matrix form

$\mathbf{Ra} = \mathbf{r}$, where the matrix $\mathbf{R}$ is defined as

$$
\mathbf{R} = \begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{p-1} \\ r_1 & r_2 & r_3 & \cdots & r_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & r_{p-3} & \cdots & r_0 \end{bmatrix} \tag{4.9}
$$

Solving for the LPC coefficients requires the inversion of the matrix $\mathbf{R}$, which is in general computationally complex. As for the autocorrelation matrix, since it is symmetric and Toeplitz, the inversion can be simplified using the Levinson-Durbin Recursion method that is more efficient.

Assume that p is an even integer, the $z$-transform of the vocal tract transfer function can be represented by

$$
H(z) = \frac{b_0}{1 - \sum_{k=1}^{p} a_k z^{-k}} = \frac{b_0}{\prod_{k=1}^{p/2} \left(1 - p_k z^{-1}\right)\left(1 - p_k^* z^{-1}\right)}, \tag{4.10}
$$

where $b_0$ is a constant gain factor, $p_k = r_k e^{j\omega_k}$ and $p_k^*$ are the corresponding roots of the denominator or poles of the transfer function that are a pair of complex conjugates.

With a sufficiently large sampling frequency $F_s$, each formant is considered to be a discrete-time sinusoid modulated by a decaying exponential as given by

$$
F_k(n) = e^{-\delta_k n} \cos(\omega_k n), \tag{4.11}
$$

where $k$ denotes the $k^{th}$ formant, $\omega_k$ and $\delta_k$ can be computed from the roots of the LPC polynomials corresponding to the $k^{th}$ formant. The formant frequency and the associated bandwidth are then determined by the computed $\omega_k$ and $\delta_k$ via

$$
f_k = \frac{F_s \cdot \omega_k}{2\pi}, \tag{4.12}
$$

$$
B_k = \frac{F_s \cdot \delta_k}{\pi}. \tag{4.13}
$$

The formant location can also be determined via other methods using the LPC coefficients such as peak-picking on the frequency response of the transfer function.

The limitations of LPC analysis for formant estimation are twofold. First of all, the LPC model assumes merely narrowband AM for each formant that does not take frequency modulation into account. As a result, it is only suitable for estimation of formants with small *bandwidth-to-center-frequency ratio* (BW/CF), such as the formants that lie in the high frequency range. Besides, crude estimation of the LPC coefficients is prone to incur significant error, leading to inaccurate poles location of the spectrum.

## 4.2   Formant Estimation via AM-FM Demodulation

As opposed to LPC analysis, AM-FM representation of speech retains the nonlinear nature of the resonance, which is evident for the first formant with a large deviation IF compared to its formant frequency. According to Potamianos [39, 40, 41], the short-time estimates of the formant frequency and the bandwidth associated with the corresponding formant can be obtained from the IA and IF using squared amplitude as weight given by

$$f_1 = \frac{\int_{t_0}^{t+T} q(t)[a(t)]^2 dt}{\int_{t_0}^{t+T} [a(t)]^2 dt}, \tag{4.14}$$

$$B_1 = \frac{\int_{t_0}^{t+T} \{(a(t)/2\pi)^2 + (q(t) - f_1)^2 [a(t)]^2\} dt}{\int_{t_0}^{t+T} [a(t)]^2 dt}. \tag{4.15}$$

A variety of demodulation techniques such as the *Hilbert transform* and the ESA can be applied to compute the IA and IF of the first formant. However, most conventional demodulation techniques rely on the narrowband assumption of the signal and only work properly under the narrowband constraint.

Due to the significant amount of frequency modulation inherent in the first formant, incorporating the MFT framework is expected to achieve a better formant demodulation result. Therefore we can generate more precise center-frequency and bandwidth estimates for the large deviation first formant than the commonly used LPC estimates that are based on the narrowband AM assumption.

## 4.3   First Formant Extraction via EMD

Prior to demodulating the first formant to obtain the IA and IF estimates, we first need to extract it from the original speech signal by separating out the different formants. This separation of formants is usually achieved by multiband filtering, for example the Garbor filterbanks as proposed by Potamianos [39]. But filtering the wideband first formant may require particularly accurate center-frequency and bandwidth. Instead of multiband filtering, in this paper we propose the use of the empirical mode decomposition (EMD) to extract the first formant, due to the following reasons: 1) EMD does not require precise center-frequency and bandwidth information, which are hardly accessible. 2) EMD allows for more sidelobes of the large deviation first formant since it does not have a fixed passband that will directly cut off the spectral components which locate outside the passband.

Initially proposed in [42], the EMD is an intuitive method that peforms the decomposition process adaptively with an aposteriori defined basis derived from the data itself. It generally involves two constituent procedures, namely the sifting process and decomposition. A function is called an intrinsic mode function (IMF) if the following conditions are satisfied: 1) The number of extremas and the number of zero-crossings equals or differs at most by one; 2) the average of the upper envelope defined by local maximas and the lower envelope defined by local minimas at any point is zero. The IMF reflects the oscillation mode inherent in the signal and can

be modeled as AM-FM.

The sifting process is a systematic way to extract the IMF from the input data $x(t)$ and can be sumarized via

- Initialize $d_0(t) = x(t)$

- Identify the local extremas of $d_n(t)$.

- Interpolate the local maximas and local minimas to form the the upper envelope $u_n(t)$ and lower envelope $v_n(t)$ respetively.

- Determine the local mean of the upper and lower envelopes via $m_n(t) = [u_n(t) + v_n(t)]/2$.

- Extract the detail: $d_{n+1}(t) = d_n(t) - m_n(t)$.

- Repeat from step 2 to step 5 until $d_{n+1}(t)$ is an IMF (zero mean or stopping criterion met).

Assume that the speech signal $S(t)$ is composed of oscillatory modes that can be modeled as IMFs. Decomposition is a procedure that keeps repeating the sifting process to decompose the original signal as the sum of IMFs plus the residue, as given by

$$S(t) = \sum_{k=1}^{n} c_k(t) + r_n(t), \tag{4.16}$$

where $c_k(t)$ denotes the corresponding IMF and $r_n(t)$ denotes the final residue. The decompsition procedure is summarized via

- Initialize $r_0(t) = S(t)$.

- Apply the sifting process on $r_n(t)$ to obtain the corresponding IMF $c_{n+1}(t)$ and the residue $r_{n+1}(t) = r_n(t) - c_{n+1}(t)$.

- Repeat the previous step until the residue $r_{n+1}(t)$ has no more extremas or meets the stopping criterion.

Note that the number of extremas associated with the extracted IMF is gradually reduced iterating from one residue to the next in the decomposition procedure, the EMD thus functions as a filterbank with the subbands changing from high frequency range to low frequency range. However, it is different from any predetermined subband filtering, since the frequency range and resolution associated with each subband is adaptively time-varying. It offers more flexibility than the conventional multiband filtering approach in capturing features that are nonstationary.

To extract the large deviation first formant, we choose a number $m$, ignore the first $m$ consecutive IMFs that oscillate at high-frequency range and sum up the rest of the IMFs and the residue as given by

$$F_1(t) = \sum_{k=m+1}^{n} c_k(t) + r_n(t). \tag{4.17}$$

By adjusting the parameters associated with the stopping criterion and observing the oscillation mode of IMFs, we can determine an appropriate number $m$ to obtain the first formant. In this paper, we adopt the EMD algorithm as proposed in [43] to extract the large deviation first formant of the vowels.

Accroding to the structure of the EMD, the number of extrema associated with the extracted IMF is gradually reduced iterating from one residue to the next in the decomposition procedure. It corresponds to a filterbank structure with subbands changing from high frequency range to low frequency range. However, it is different from any predetermined subband filtering, since the frequency range and resolution associated with each subband are adaptively time-varying. It offers more flexibility than the traditional multiband filtering approach in capturing features that are nonstationary.

There are a few variants of the basic EMD such as the ensemble empirical mode decomposition (EEMD) [44]. The key idea of the EEMD relies on averaging the modes obtained by applying EMD to different realizations of Gaussian white noise added to the original signal. By introducing white Gaussian noise, the mode mixing problem can be solved via populating the whole time-frequency space to take advantage of the dyadic filter bank behavior of the EMD. In addition, a complete ensemble empirical mode decomposition with adaptive noise based on the EEMD is also proposed in [45].

In this thesis, we employ the CEEMD algorithm as proposed in [43] to obtain the large deviation first formants of different vowels.

## 4.4   Simulation Results

The short-time fourier transform spectrum of a women's vowel /i:/ [1] and the first formant extracted via the EMD are illustrated in Fig. 4.1. As we can observe, the first formant has a large BW/CF and non-negligible sidelobes induced by its considerable amount of frequency modulation. The residue and the first three IMFs of the vowel /i:/ after the EMD are also illustrated in Fig. 4.2. Note that the first formant extracted by EMD retains these sidelobes that are usually ignored in the LPC analysis.

The demodulated IF of the large deviation first formant of /i:/ via different approaches are compared in Fig. 4.3(a). The estimated IF of the LPC approach varies slowly like a straight line, since the LPC only picks up the pole and cuts off the sidelobes induced by frequency modulation. The estimated IF by the ESA varies too sharply in some range due to the large deviation nature of the first formant. The

---

[1]The experimental data is based on the source in the vowel database of Hillenbrand, Getty, Clark & Wheeler (1995)

Table 4.1: Comparison of the formant estimates for different female first formants.

|  | /i:/ | | | /uw/ | | |
|---|---|---|---|---|---|---|
|  | Formant | BW | BW/CF | Formant | BW | BW/CF |
| LPC | 433 Hz | 71 Hz | 16.4% | 438 Hz | 12 Hz | 2.7% |
| ESA | 523 Hz | 412 Hz | 78.8% | 453 Hz | 240 Hz | 53.0% |
| MFT-ESA | 387 Hz | 157 Hz | 40.6% | 435 Hz | 90 Hz | 20.7% |
|  | /ei/ | | | /ae/ | | |
|  | Formant | BW | BW/CF | Formant | BW | BW/CF |
| LPC | 458 Hz | 29 Hz | 6.3% | 627 Hz | 75 Hz | 12.0% |
| ESA | 479 Hz | 378 Hz | 78.9% | 682 Hz | 433 Hz | 63.5% |
| MFT-ESA | 451 Hz | 102 Hz | 22.6% | 588 Hz | 167 Hz | 28.4% |

Table 4.2: Comparison of the formant estimates for different male first formants.

|  | /i:/ | | | /uw/ | | |
|---|---|---|---|---|---|---|
|  | Formant | BW | BW/CF | Formant | BW | BW/CF |
| LPC | 290 Hz | 33 Hz | 11.4% | 332 Hz | 34 Hz | 10.2% |
| ESA | 516 Hz | 1012 Hz | 196.1% | 464 Hz | 338 Hz | 72.8% |
| MFT-ESA | 280 Hz | 115 Hz | 41.1% | 338 Hz | 139 Hz | 41.1% |
|  | /ei/ | | | /ae/ | | |
|  | Formant | BW | BW/CF | Formant | BW | BW/CF |
| LPC | 396 Hz | 31 Hz | 7.8% | 633 Hz | 56 Hz | 8.8% |
| ESA | 455 Hz | 482 Hz | 105.9% | 753 Hz | 667 Hz | 88.6% |
| MFT-ESA | 393 Hz | 151 Hz | 38.4% | 622 Hz | 191 Hz | 30.7% |

IF estimated via the MFT-ESA combination varies smoothly within the frequency range of the first formant, which is better than the prior estimates. The demodulated IA estimates via the ESA and the MFT-ESA are also compared in Fig. 4.3(b). from which we can see that the IA estimate by the MFT-ESA is varying slowly and smoother than that of the ESA. From Eq. 4.11 we know that the envelope for the formant modeled by the LPC approach is a decaying exponential function, which is diffiuent from the IA estimates of both the ESA and the MFT-ESA, thus not compared in Fig. 4.3(b). According to Eq. 4.15, the square of the IA estimates serve as the weight for computing the bandwidth for the first formant.

The center-frequency and bandwidth estimates of the first formants extracted via the EMD based on three approaches for different female and male vowels in the database are compared in Table 4.1 and Table 4.2 respectively. The bandwidth estimates of the MFT-ESA turn out to be greater than those of the LPC approach assuming only amplitude modulation, matching the spectrum better as indicated in Fig. 4.1. The estimated BW/CF of each first formant via the MFT-ESA lies in the large-deviation regime while that of the LPC approach is too small to characterize the frequency modulation inherent in the first formant. Without the MFT, the ESA, however, incurs significant error in the large-deviation regime [39], the bandwidth estimates are too large, leading to erroneous BW/CF estimates as well. Therefore we conclude that the formant estimates by the MFT-ESA are more precise than the LPC estimates for large deviation first formants.

## 4.5   Conclusion

In this chapter, we have presented an approach that applies IF demodulation via the MFT-ESA combination to the first formant of vowels extracted by EMD and then computed the formant frequency and bandwidth estimates based on the demodulated IF and IA. By taking into account the wideband nature of the large deviation first formant, which usually has a large BW/CF, the formant estimates via the proposed approach are demonstrated to be more reasonable than the LPC estimates that are based on the narrowband AM formant assumption and the traditional multiband ESA that assumes narrowband AM–FM components.
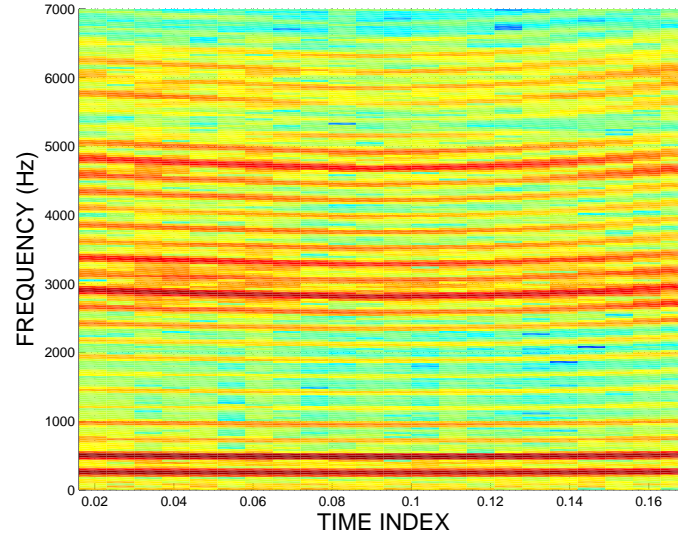
(a)
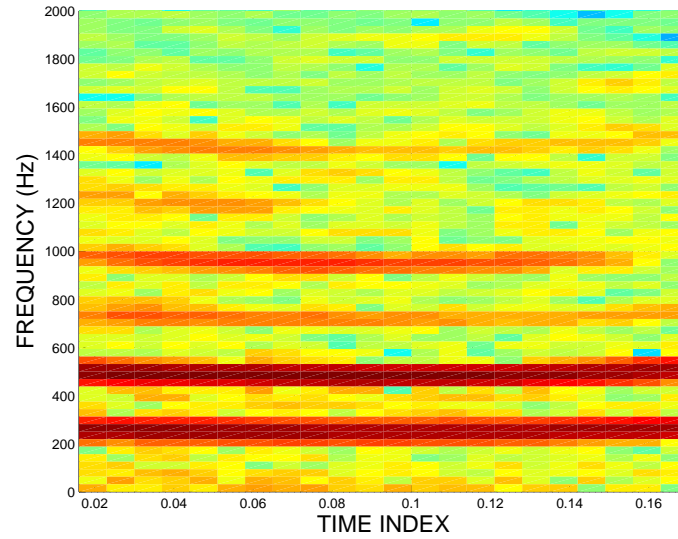


(b)

Figure 4.1: (a) Original speech segment. (b) First formant extracted via the EMD.
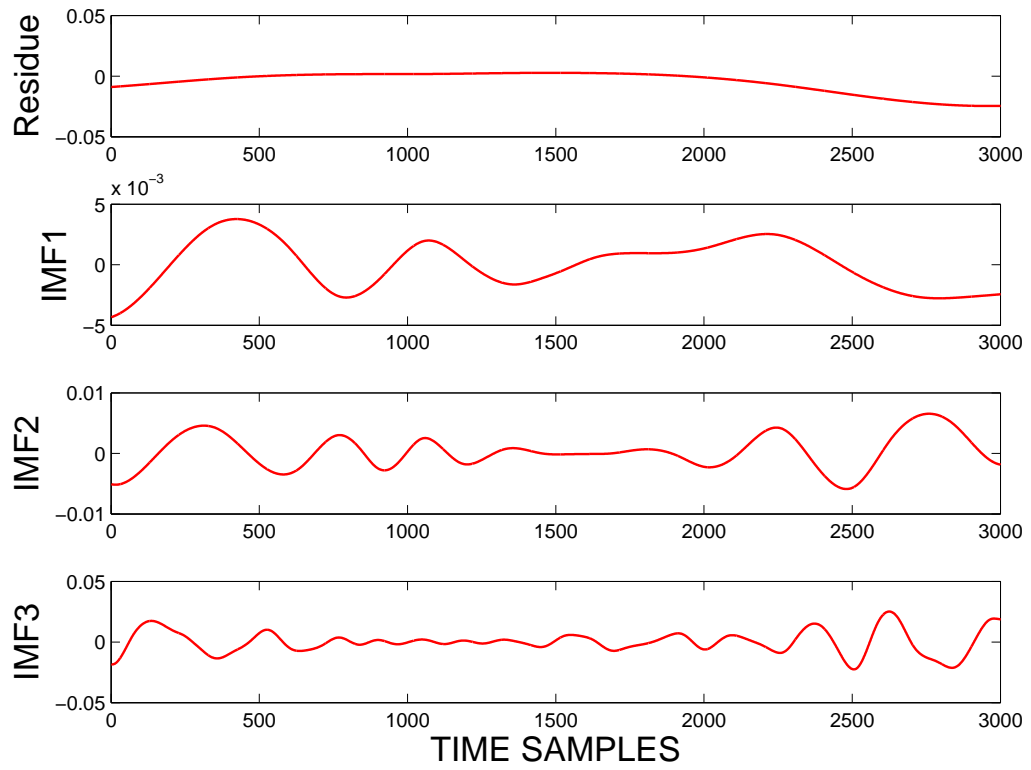
Figure 4.2: IMFs of the vowel /i:/ after the EMD.
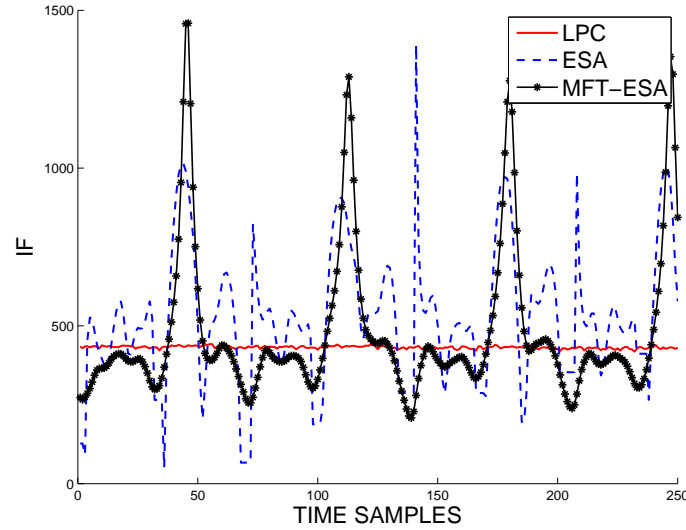
(a)
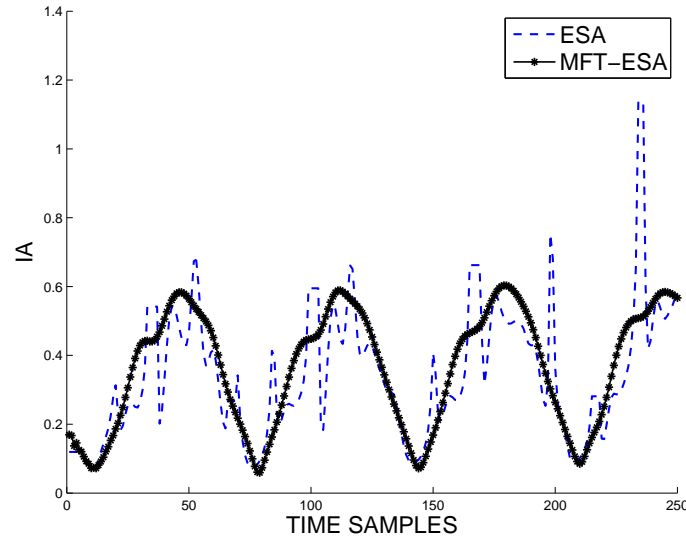


(b)

Figure 4.3: (a) IF estimates via the LPC (red line), ESA (blue dashed line), MFT-ESA (black dotted line). Note that the ESA-2 algorithm with second order binomial smoothing, a large multirate factor $R = 32$ and a normalized heterdyning frequency 0.2 is applied in the MFT-ESA for this example. (b) IA estimates via both ESA (blue dashed line) and MFT-ESA (black dotted line).

# Chapter 5

# Applications to Wideband Partial-response CPM Demodulation

*Continuous phase modulation* [46, 47, 48] is a popular form of frequency modulation employed in mobile communications [49] and has desirable spectral efficiency [50] and constant modulus properties that facilitate use of class-C amplifiers. *Gaussian Minimum Shift Keying* (GMSK), a specific form of CPM is the main ingredient in the *Global System for Mobile communications* (GSM) [47, 48] used in GPS applications. Pragmatic CPM modulation schemes have recently been studied as capacity attaining low-complexity alternatives to serially concatenated CPM [51].

The conventional demodulation technique used for narrowband signals is phase demodulation [52] followed by unwrapping and *maximum likelihood sequence estimation* (MLSE) using the *Viterbi algorithm* (VA) [46, 47, 53]. This approach has a complexity that grows exponentially with the number of phase states and restrictions on the modulation index that needs to be the ratio of incommensurate integers.

Other frequency discrimination approaches [54] rely upon instantaneous frequency (IF) estimation and are not subject to the restrictions required by the Viterbi but have not been investigated for wideband CPM with memory.

In prior work [2, 3] it was shown that frequency discrimination for full response CPM demodulation has the same performance as that of BPSK detection in AWGN. Further in recent work [9], frequency tracking based wideband FM demodulation was extended to large wideband to narrowband conversion factors using the MFT scheme as proposed in chapter 2. Frequency estimation based approaches have the added advantage that they are immune to phase distortions introduced by the channel which would adversely affect the Viterbi approach. In addition, these approaches do not require prior knowledge of the carrier frequency.

In this chapter, we investigate the approach that utilizes MFT-ESA for wideband CPM demodulation with decision feedback equalization (DFE) to effectively eliminate the memory introduced by partial response signaling [55]. The wideband partial response CPM signal is first converted into narrowband via the MFT and then demodulated by the ESA to obtain the IF estimates. By averaging the IF estimates within each symbol period to form the partial response channel, the DFE is eventually applied to equalize the channel in order to remove the CPM memory. Simulation results are used to verify the efficacy of the combined approach for both wideband binary-CPM and multilevel-CPM.

# 5.1   Continuous Phase Modulation

The standard CPM model depends on its pulse-shaping function $p(t)$, with duration length of symbol periods $L$ and modulating symbols, i.e, binary PAM symbols $a[k] \in$

$\{-1, 1\}$. The IF signal takes the form [47]:

$$\omega_i(t) = \omega_c + 2\pi h \sum_{k=-\infty}^{\infty} a[k] p(t - kT_b), \tag{5.1}$$

where $\omega_c$ is the carrier frequency and $h$ is the modulation index of CPM. The carrier phase is given by:

$$\phi_{\text{dev}}(t; \mathbf{a}) = 2\pi h \sum_{k=-\infty}^{\infty} a[k] q(t - kT_b), \tag{5.2}$$

where $p(t)$ is the normalized pulse shaping function defined over the interval $[0, LT]$ and $q(t) = \int_0^t p(\tau) d\tau$ corresponds to the phase pulse shaping function. In general, $p(t)$ and $q(t)$ are assumed to satisfy the following condition:

$$p(t) = p(LT - t), \tag{5.3}$$

$$q(t) = \int_0^t p(\tau) d\tau = \frac{1}{2}, \quad t \geq L. \tag{5.4}$$

If $p(t)$ is a rectangular pulse then this form of CPM is referred to as (L-REC) CPM, and if $p(t)$ is a raised cosine pulse then it is referred to as (L-RAC) CPM. The CPM signal is then obtained via frequency modulation:

$$s(t) = A \cos \left( \int_{-\infty}^t \omega_i(\tau) d\tau + \theta \right). \tag{5.5}$$

Usually the modulation index of a wideband CPM signal is large such that the frequency deviation of its IF is comparable to the carrier frequency.

Using a pulse shaping function of duration larger than a symbol period ($L > 1$), i.e., partial response signaling introduces memory into the modulation scheme.

## 5.2 Optimum Demodulation of CPM

The MLSE implemented via the *Viterbi algorithm* has been proposed as an optimal solution for CPM demodulation and detection. According to Eq. 5.2, the phase

deviation of the CPM signal can be split into two parts:

$$\phi_{\text{dev}}(t; \mathbf{a}) = 2\pi h \sum_{k=-\infty}^{\infty} a[k]q(t - kT_b)$$

$$= \pi h \sum_{k=-\infty}^{n-L} a[k] + 2\pi h \sum_{k=n-L+1}^{n} a[k]q(t - kT) \tag{5.6}$$

$$= \theta_n + \theta(t, \mathbf{a}), \quad nT \leq t \leq (n+1)T$$

where $\theta_n$ is the cumulative phase for prior symbols and $\theta(t, \mathbf{a})$ is the phase influenced by the most recent L symbols up to the $n^{th}$ period. Hence we define the state of the CPM at time index $t = nT$ via

$$S_n = \{\theta_n, a[n-1], a[n-2], a[n-L+1]\}. \tag{5.7}$$

Then the state at next time index $t = (n+1)T$ is represented by

$$S_{n+1} = \{\theta_{n+1}, a[n], a[n-1], a[n-L+2]\}, \tag{5.8}$$

where

$$\theta_{n+1} = \theta_n + \pi h a[n-L+1]. \tag{5.9}$$

Now assume that the modulation index $h$ is rational, which can be represented by

$$h = \frac{m}{p}, \tag{5.10}$$

where $m$ and $p$ are relatively prime positive integers. For memoriless CPM ($L = 1$), the phase trellis of the CPM scheme has only $p$ (when $m$ is even) or $2p$ (when $m$ is odd) states assoicated with $\theta_n$. When memory is introduced into the CPM ($L > 1$), additional states are required by the information symbols $\{a[n-1], a[n-2], ..., a[n-L+1]\}$ contained in the state of the CPM due to the partial response signaling. Therefore, the number of states for partial response CPM is denoted by

$$N_s = \begin{cases} pM^{L-1}, & m \text{ even} \\ 2pM^{L-1}, & m \text{ odd} \end{cases} \tag{5.11}$$

where $M$ is the size of the alphabet for the transmitted symbols. The equivalent baseband received signal is expressed by

$$r(t) = A \cos(\omega_c t + \phi_{dev}(t, \mathbf{a})) + n(t). \tag{5.12}$$

where $n(t)$ is assumed to be AWGN. The MLSE aims at maximizing the the probability of the observation conditioned on the transmitted symbols that is proportional to the cross-correlation metric given by

$$
\begin{aligned}
CM_n(\mathbf{a}) &= \int_{-\infty}^{(n+1)T} r(t) \cos[\omega_c t + \phi_{\text{dev}}(t, \mathbf{a})] dt \\
&= CM_{n-1}(\mathbf{a}) + \int_{nT}^{(n+1)T} r(t) \cos[\omega_c t + \theta(t, \mathbf{a}) + \theta_n] dt \\
&= CM_{n-1}(\mathbf{a}) + v_n(\theta_n, \mathbf{a})
\end{aligned} \tag{5.13}
$$

where $CM_{n-1}(\mathbf{a})$ denotes the metrics for the surviving sequences up to time index $nT$ and $v_n(\theta_n, \mathbf{a})$ denotes the additional increments of the metrics induced by the signal in the interval $[nT, (n+1)T]$. When the VA is implemented, the number of surviving sequences associated with each state is $Ns$. Hence the complexity of the optimal MLSE via the VA is influenced by the chosen alphabet size $M$, rational modulation index $h$ and the duration length $L$ of the CPM pulse shaping function. As a result, the complexity of the VA grows exponentially as $L$ increases, which is especially significant when $M$ is larger than 2. For such cases, the suboptimum demodulation schemes have to be used.

## 5.3 CPM Demodulation via Equalization

The memory introduced by the partial response signaling linearly distort the transmitted signal and results in the *intersymbol interference* (ISI) for the IF of the CPM signal within each symbol period. If precise IF estimates of the CPM signal are

accessible, the recovery of the original transmitted sequence is equivalent to equalization of the ISI channel. The discrete-time model for the ISI channel is described by

$$r[k] = \sum_{l=-\infty}^{\infty} h[l]a[k-l] + z[k]. \tag{5.14}$$

where $h[k]$ denotes the impulse response of the channel, $a[k]$ denotes the original transmit sequence and $z[k]$ denotes the noise averaged from the IF estimation. Since the noise term $z[k]$ is not guaranteed to be AWGN by any demodulation approach, the MLSE implemented via the VA directly applied to the demodulated IF of the CPM is not the optimal solution in this case. For general equalization schemes, the impulse response of the channel $h[k]$ is assumed to be known, but practically $h[k]$ is hardly accessible and needs to be estimated, i.e., using training sequence of pilot symbols.

## 5.3.1 Partial Response Channel

The partial response channel is a special case of the ISI channels that has a finite-length impulse response, and is also causal and monic with $h[0] = 1$. The memory introduced by partial response CPM depends on the type of the pulse shaping function and the corresponsding duration length $L$. For example, the discrete-time partial response channel using L-REC CPM can be modeled as

$$y_{\text{REC}}(t) = \sum_{k=0}^{L-1} a(t-k) + z(t), \tag{5.15}$$

where $t$ is the discrete-time index and $a(t)$ is the original transmitted symbol sequence.

For wideband partial response CPM, the output of the partial response channel can be obtained by demodulating the CPM signals using the MFT-ESA approach as

previously proposed and matched filtering on the estimated IF within each symbol period. Hence our goal of recovering the original transmitted sequence is equivalent to the equalization of the partial response channel.

## 5.3.2 Carrier Frequency and Amplitude Estimation

The carrier frequency and the amplitude of the CPM can be estimated from the MFT-ESA demodulation result, which turns out to be an advantage for the frequency estimation based approaches. The IF estimate of the MFT-ESA in the specific case of CPM takes the form of:

$$\hat{\omega}_i(t) = \omega_c + 2\pi h \sum_{k=-\infty}^{\infty} a[k]h_f(t - kT_b) + \epsilon_\omega(t), \tag{5.16}$$

where $h_f(t)$ corresponds to the pulse shaping function and $\epsilon_w(t)$ corresponds to zero-mean IF noise, which unlike the observation noise is not white. Assuming equiprobable symbols and taking expectations on both sides yields:

$$E\{\hat{\omega}_i(t)\} = \omega_c. \tag{5.17}$$

The carrier frequency and the amplitude of the AM–FM signal can then be estimated from the IF and IA estimates from either algorithm by simple averaging:

$$\begin{aligned} \hat{\omega}_c &= \frac{1}{T}\int_0^T \hat{\omega}_i(t)dt \\ \hat{A} &= \frac{1}{T}\int_0^T \hat{a}_i(t)dt. \end{aligned} \tag{5.18}$$

This is a consequence of the fact that these approaches are bandpass estimation approaches whereas traditional in-phase and quadrature demodulation, employed in narrowband communication systems, is a baseband estimation approach requiring prior knowledge of the carrier frequency.

### 5.3.3 Partial Response Channel Estimation via Recursive Prediction Error Method

Since the channel response is required by general equalization schemes in order to remove the memory of the partial response CPM, in this thesis we perform the channel estimation via the recursive prediction error method [56]. The partial response channel, depicted for example in Eq. (5.15), is actually a monic *moving-average* (MA) process. We are capable of estimating the channel response by fitting the output of the partial response channel to a ARMAX model [56].

In general, the structure of the ARMAX model is described by

$$\sum_{k=0}^{n_a} a_k y[t-k] = \sum_{k=1}^{n_b} b_k u[t-k] + \sum_{k=0}^{n_c} c_k e[t-k], \tag{5.19}$$

where $n_a$, $n_b$ and $n_c$ are the number of coefficients for the *auto-regressive* (AR) part, system input and moving-average (MA) part respectively. It can also be written as

$$A(q)y(t) = B(q)u(t) + C(q)e(t). \tag{5.20}$$

where $q$ is the backward shift operator. Specifically,

$$A(q) = 1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a}, \tag{5.21}$$

$$B(q) = b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b}. \tag{5.22}$$

$$C(q) = 1 + c_1 q^{-1} + \ldots + c_{n_b} q^{-n_c}. \tag{5.23}$$

By assuming $A(q) = 1$ and $B(q) = 0$, the ARMAX model can be simplified to the MA model that exactly fits Eq. (5.15), as described by

$$y(t) = C(q)e(t) = \sum_{k=0}^{n_c} c_k q^{-k} e(t) = \sum_{k=0}^{n_c} c_k e[t-k]. \tag{5.24}$$

By satisfying certain conditions, $C(q)$ is actually invertible, that is, $e(t)$ can be calculated via an inverse operator $\tilde{C}(q)$ via

$$e(t) = \tilde{C}(q)y(t) = \sum_{k=0}^{\infty} \tilde{c}_k y(t-k). \tag{5.25}$$

The coefficient estimation for the ARMAX model can be achieved via an iterative search algorithm that minimizes a more robust quadratic prediction error criterion. The parameter vector $\vec{\theta}$ can be formed by grouping the coefficients of the ARMAX model as

$$\vec{\theta} = [a_1, ..., a_{n_a}, b_1, ..., b_{n_b}, c_1, ..., c_{n_c}]. \tag{5.26}$$

The predictor for the ARMAX model is given by

$$C(q)\hat{y}(t|\vec{\theta}) = B(q)u(t) + [C(q) - A(q)]y(t). \tag{5.27}$$

It can be rewritten as

$$\hat{y}(t|\vec{\theta}) = B(q)u(t) + [1 - A(q)]y(t) + [C(q) - 1]\epsilon(t, \vec{\theta}), \tag{5.28}$$

where $\epsilon(t, \vec{\theta})$ is defined as the prediction error given by

$$\epsilon(t, \vec{\theta}) = y(t) - \hat{y}(t|\vec{\theta}). \tag{5.29}$$

Therefore we can express the predictor in the form of pseudolinear regression via

$$\hat{y}(t|\vec{\theta}) = \vec{\varphi}^T(t, \vec{\theta})\vec{\theta}, \tag{5.30}$$

where we define the data vector $\vec{\varphi}(t, \vec{\theta})$ as

$$\begin{aligned} \vec{\varphi}(t, \vec{\theta}) = & [-y(t-1), ..., -y(t-n_a), u(t-1), ..., \\ & u(t-n_b), \epsilon(t-1, \vec{\theta}), ..., \epsilon(t-n_c, \vec{\theta})]. \end{aligned} \tag{5.31}$$

According to Eq. (5.28), the gradient of the predictor $\vec{\psi}(t, \vec{\theta})$ w.r.t $\vec{\theta}$ can be computed via

$$C(q)\vec{\psi}(t, \vec{\theta}) = \vec{\varphi}(t, \vec{\theta}). \tag{5.32}$$

The gradient $\vec{\psi}(t, \vec{\theta})$ can be obtained by filtering the data vector $\vec{\varphi}(t, \vec{\theta})$ through an inverse filter of $C(q)$.

The cost function for the recursive prediction error method is defined as

$$V_t(\vec{\theta}, \vec{Y}^t) = \gamma(t)\frac{1}{2}\sum_{k=1}^{t}\beta(t,k)\epsilon^2(k,\vec{\theta}), \tag{5.33}$$

where $\beta(t,k)$ and $\gamma(t)$ satisfy the following conditions

$$\beta(t,k) = \prod_{j=k+1}^{t}\lambda(j), \quad \beta(t,t) = 1, \tag{5.34}$$

$$\sum_{k=1}^{t}\gamma(t)\beta(t,k) = 1, \tag{5.35}$$

Note that $\lambda(j)$ is the forgetting factor, which is often set to a constant less than 1.

The algorithm of the recursive prediction error method is then summarized via

$$\epsilon(t) = y(t) - \hat{y}(t), \tag{5.36}$$

$$\hat{\vec{\theta}}(t) = \hat{\vec{\theta}}(t-1) + \gamma(t)R^{-1}(t)\vec{\psi}(t)\epsilon(t), \tag{5.37}$$

$$R(t) = R(t-1) + \gamma(t)[\vec{\psi}(t)\vec{\psi}^T(t) - R(t-1)], \tag{5.38}$$

where $\vec{\psi}(t)$ and $\hat{y}(t)$ are short for the resulting approximations of $\vec{\psi}(t, \hat{\vec{\theta}}(t-1))$ and $\hat{y}(t|\hat{\vec{\theta}}(t-1))$ respectively.

### 5.3.4   Memory Removal via DFE

*Decision feedback equalization* [57] is proposed in this thesis to equalize the partial response channel in order to recover the transmitted sequence. The block diagram of the general DFE is illustrated in Fig. 5.1. The DFE structure involves two filters, the feedforwrd filter whose input is the output sequence of the partial response channel and the feedback filter that feeds the previously determined symbols $\hat{I}[k-l]$, $l \geq 1$ back to the loop. Note that the feedforward filter $F(z)$ can be represented by

$$F(z) = \sum_{k=-\infty}^{\infty} f[k]z^{-k}, \tag{5.39}$$

and the feedback filter $B(z)$ are given by

$$B(z) = \sum_{l=0}^{L_B-1} b[l]z^{-l}, \tag{5.40}$$

where $B(z)$ is causal and monic with $b[0] = 1$. The feedforward filter is only in-
tended for suppresion of the pre-cursor ISI, leading to fewer constraints and noise
enhancement that is much smaller than the linear equalization. The post-cursors are
compensated by the feedback filter, resulting in no additional noise enhancement due
to the *symbol by symbol detector* (SBS). Note that the DFE analysis is based on the
assumption that previous decisions are correct, otherwise the error propagation due
to wrong decisions incurs performance degradation but is usually not catasrophic.
According to the choice of the feedforward filter $F(z)$ and the feedback filter $B(z)$, a
variety of decision feedback equalizers can be implemented, such as the zero-forcing
DFE that directly inverts the channel and the MMSE DFE that employs a minimum
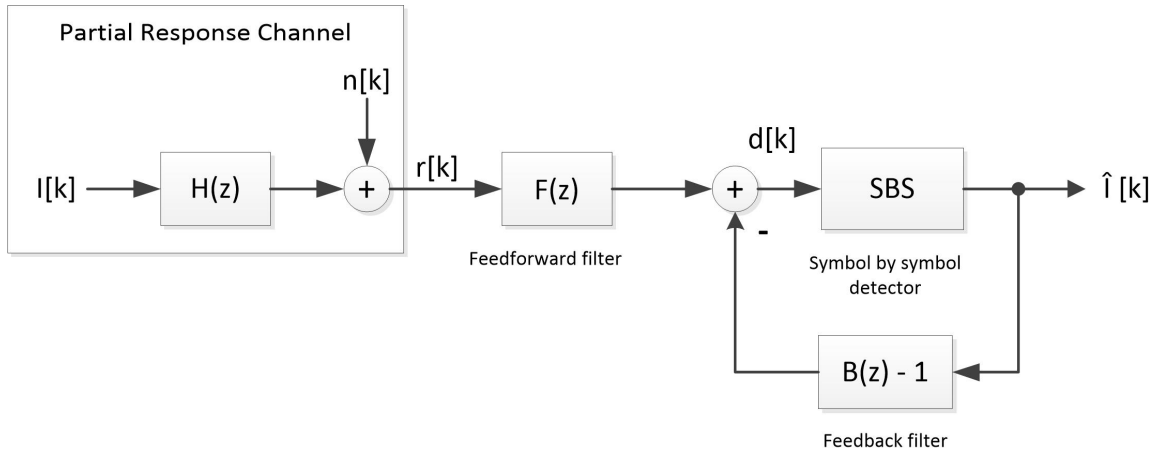mean square error criterion.



Figure 5.1: Block diagram of the general decision-feedback equalizer. The partial
response channel output in this case are the mixed symbols due to memory introduced
by partial response signaling. The mixed symbols can be obtained from the estimated
IF of the MFT-ESA demodulation block.

**ZF-DFE Solution**

The ZF-DFE can be viewed as cascade of the linear ZF equalizer $\frac{1}{H(z)}$ and the prediction error filter $P_e(z) = 1 - z^{-1}P(z)$ as

$$F(z) = \frac{P_e(z)}{H(z)}, \tag{5.41}$$

where $P(z)$ is the noise predictor that predicts the future noise sample using most recent $L_p$ noise samples

$$P(z) = \sum_{m=0}^{L_P-1} p[m]z^{-m}. \tag{5.42}$$

The feedback filter is then given by

$$B(z) = 1 - z^{-1}P(z) = P_e(z). \tag{5.43}$$

If the channel noise is assumed to be AWGN with variance $N_0$, the power spectrum of the error at $d[k]$ is computed by

$$\Phi_{ee}(z) = \frac{P_e(z)P_e^*(1/z^*)N_0}{H(z)H^*(1/z^*)}. \tag{5.44}$$

The optimum noise prediction error filter is a whitening filter given by

$$P_e(z) = \frac{H_{\min}(z)}{h_{\min}[0]}, \tag{5.45}$$

where $H_{\min}(z)$ is the minimum phase equivalent of $H(z)$ that can be obtained via spectral factorization.

For example, in the case of the 3-REC multilevel CPM, the expression for the inverse filter is given by

$$H_{inv}(z) = \frac{1}{H(z)} = \frac{1}{1 + \hat{c}_1 z^{-1} + \hat{c}_2 z^{-2}}. \tag{5.46}$$

Note that the inverse filter $H_{inv}(z)$ is an IIR All-pole filter, which can be implemented via direct recursion of its difference equation. It requires the input, i.e., in our case,

the partial response channel output sequence $y_{\mathrm{REC}}(t)$, to be nearly perfect, otherwise the error propagation due to inaccurate input sequence can be significant. This in turn requires the MFT demodulation section to produce a sufficiently accurate demodulation result. By incorporating the slicer, i.e., the SBS detection device into the inverse filter recursion we can implement the decision feedback version of the zero-forcing equalizer (ZF-DFE) to eliminate memory induced in the partial response CPM signals.

**MMSE-DFE Solution**

Instead of focussing on just removal of the channel between the IF input and the information symbols if we further incorporate a MMSE cost function that balances the task of eliminating memory with the task of reducing symbol distortion we obtain the linear MMSE equalizer that can provide further improvement in the symbol error performance in low SNR environment. The corresponding decision feedback version of the linear MMSE equalizer (MMSE-DFE) incorporates both pre-cursor and post-cursor taps. The feedforward filter coefficients of the MMSE-DFE are obtained from the Wiener solution and then used to solve for the feedback filter coefficients. Details of the MMSE-DFE solution are described in [53, 58]. In this thesis, we implement the MMSE-DFE using FIR filters. Assume that the length of the feedforward filter and the feedback filter are denoted by $L_F$ and $L_B$ respectively, then the input to the SBS is given by

$$d[k] = \sum_{m=0}^{L_F-1} f[m]r[k-m] - \sum_{m=0}^{L_B-1} b[m]I[k-k_0-m], \qquad (5.47)$$

where $k_0 \geq 0$ denotes the decision delay. The goal of the MMSE-DFE is to minimize the variance of the error

$$\begin{aligned}
e[k] &= d[k] - I[k-k_0] \\
&= \mathbf{f}^H\mathbf{r}[k] - \mathbf{b}^H\mathbf{I}[k-k_0-1] - I[k-k_0],
\end{aligned} \qquad (5.48)$$

where

$$\mathbf{f} = [f[0] \ f[1] \ ... \ f[L_F - 1]]^H$$

$$\mathbf{b} = [b[0] \ b[1] \ ... \ b[L_B - 1]]^H$$

$$\mathbf{r}[k] = [r[k] \ r[k-1] \ ... \ r[k-(L_F-1)]]^T$$

$$\mathbf{I}[k - k_0 - 1] = [I[k-k_0-1] \ I[k-k_0-2] \ ... \ I[k-k_0-(L_B-1)]]^T.$$

(5.49)

By computing the error variance $J$ and differentiating it with respect to $\mathbf{f}^*$ and $\mathbf{b}^*$

$$\frac{\partial J}{\partial \mathbf{f}^*} = (\mathbf{\Phi}_{hh} + \sigma_n^2 I)\mathbf{f} + \mathbf{H}\mathbf{b} - \mathbf{h}, \tag{5.50}$$

$$\frac{\partial J}{\partial \mathbf{b}^*} = \mathbf{b} - \mathbf{H}^H \mathbf{f}, \tag{5.51}$$

where $\mathbf{\Phi}_{hh}$ is the autocorrelation matrix of the channel response and

$$\mathbf{h} = \varepsilon\{\mathbf{r}[k]I^*[k-k_0]\} = [h[k_0] \ h[k_0-1] \ ... \ h[k_0-(L_F-1)]]^T$$

$$\mathbf{H} = \varepsilon\{\mathbf{r}[k]\mathbf{I}^*[k-k_0-1]\}.$$

(5.52)

By setting the two equations to zero, we can obtain the optimum coeffients for the feedforward filter and feedback filter respectively given by

$$\mathbf{f}_{\text{opt}} = \left((\mathbf{\Phi}_{hh} - \mathbf{H}\mathbf{H}^H) + \sigma_n^2 I\right)^{-1} \mathbf{h}, \tag{5.53}$$

$$\mathbf{b}_{\text{opt}} = \mathbf{H}^H \mathbf{f}. \tag{5.54}$$

## 5.4 Simulation Results

### 5.4.1 Performance of Carrier Frequency Estimation

In practice, since prior knowledge of distributions is not available, we replace the expectation with a simple time-average as in Eq. 5.18. While in the discrete-time case, we replace the integral with a time-average sum:

$$\hat{\Omega}_c = \frac{1}{L} \sum_{k=0}^{L-1} \Omega_i[k].$$
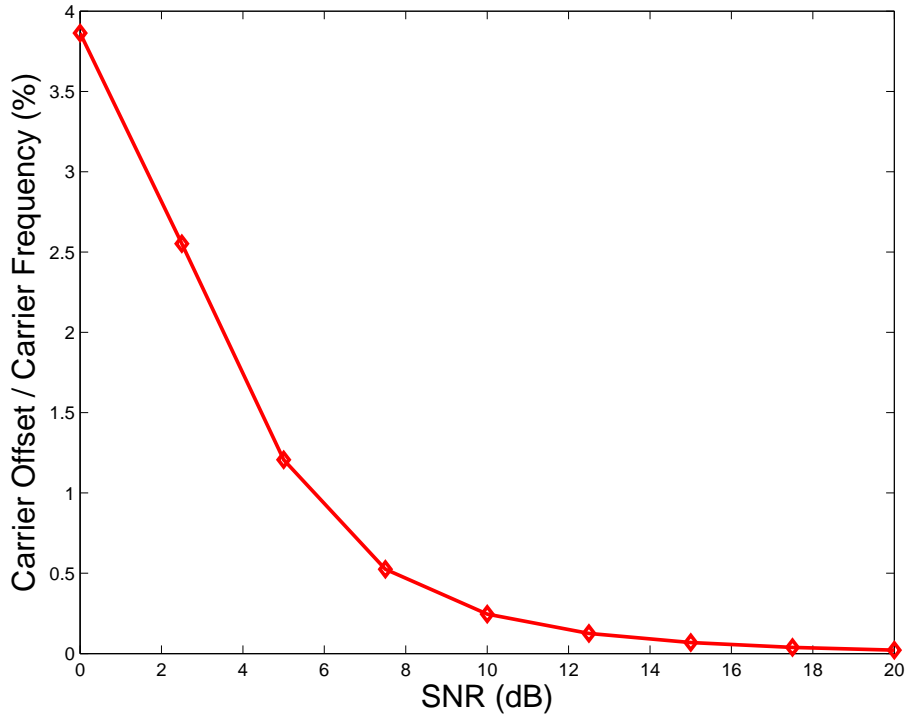
Figure 5.2: Performance of carrier frequency estimation for the MFT-ESA approach.

Fig. 5.2 depicts the carrier frequency estimation error of the MFT-ESA for the case where we have equiprobable symbols versus SNR. For larger SNR values the carrier-frequency estimation error approaches zero indicating that the IF yields a reliable carrier frequency estimate.

## 5.4.2 Performance of Wideband Partial Response Multilevel CPM Demodulation

A wideband 3-REC multilevel CPM signal with original symbols taking values in the alphabet $\{-3, -1, 1, 3\}$ is used for performance test. The frequency deviation of the IF is equal to the carrier frequency in this extreme wideband case with modulation index $h = 4$. The symbol error probability for the proposed MFT-ESA approach is
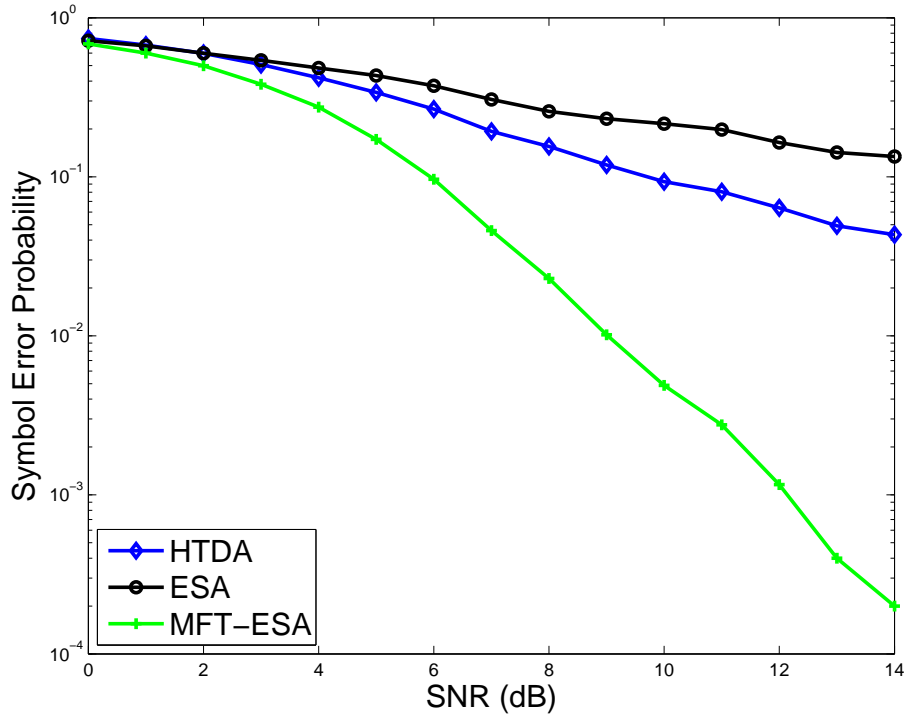
Figure 5.3: Symbol error probability associated with the mixed symbols for wideband 3-REC multilevel CPM. Note that the SEP will drop to zero around 16 dB for the MFT-ESA approach, which is not shown due to limitations of the log scale.

depicted in Fig. 5.3. Since the 3-REC CPM signal has memory, the symbols here refer to the mixed symbols due to the memory effect of partial response signaling as described in Eq. 5.15. As the SNR increases, the MFT and ESA combination reduces the error dramatically. The error eventually drops to zero when the SNR passes certain threshold, while the error performances of the HTDA and the ESA gradually saturate at certain levels due to carry-over effects from incomplete demodulation induced by narrowband constraints.

Unlike the common AWGN channel, the noise imposed on the CPM signal is not linearly added to the modulation signals (or symbols). Since the modulation signals (or symbols) are conveyed in the IF of the CPM signal, the effect of the

noise remains indirect. Therefore the symbol error performance for the CPM signal is different from what is observed with modulation schemes that fit into the AWGN channel analysis, such as the classic Q curve for BPSK modulation. The CPM format provides robustness to noise when the SNR exceeds a certain threshold and if the CPM signal is sufficiently sampled.

### 5.4.3 Performance of Partial Response Channel Estimation via Recursive Error Method



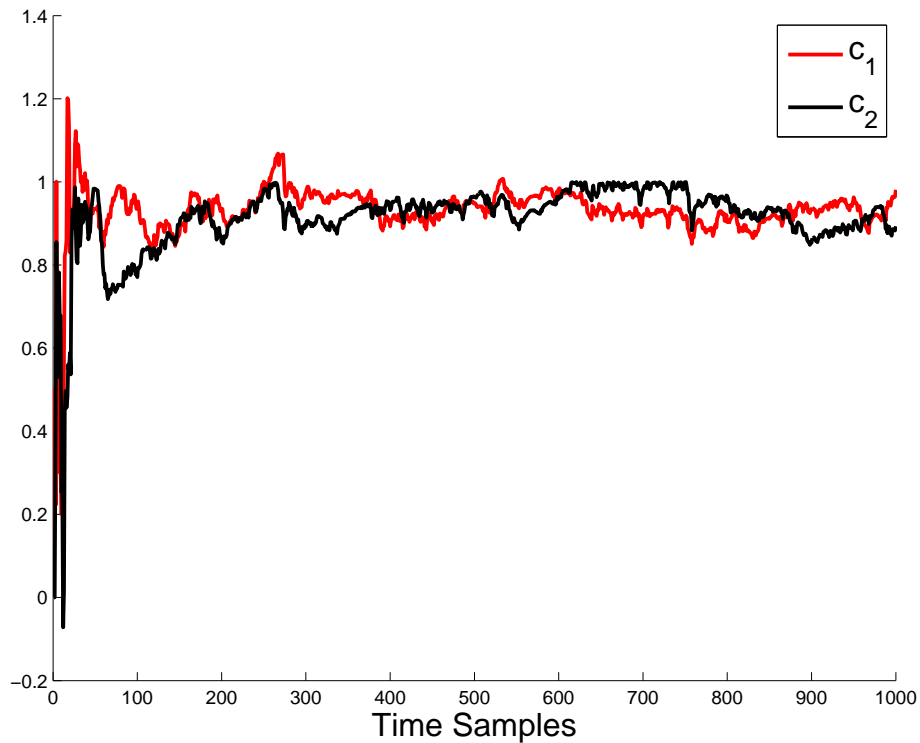Figure 5.4: Partial response channel estimation for 3-REC multilevel CPM via recursive error method. Note that the multilevel symbols take values in $\{-3, -1, 1, 3\}$. The red line indicates the convergence of the estimated coefficient $\hat{c}_1$ and the black line indicates the convergence of the estimated coefficient $\hat{c}_2$.

The experimental result of the partial response channel estimation via the recur-

sive error method for wideband 3-REC multilevel CPM is illustrated by Fig. 5.4. With a memory length $L = 3$, the partial response channel for the 3-REC multilevel CPM can be expressed via

$$y_{\mathrm{REC}}(t) = s[t] + c_1 s(t - 1) + c_2 s(t - 2) + z(t), \tag{5.55}$$

where $c_1 = c_2 = 1$. As we can observe from Fig.5.4, the estimated coefficients converge close to the true value 1 in the case of REC-CPM. They serve as useful estimates when other forms of pulse shaping such as in RAC-COM or SRAC-CPM are employed.

### 5.4.4 Performance of Partial Response CPM Memory Removal via Decision Feedback Equalization

The MFT-ESA demodulation module is then combined with the decision-feedback equalization for memory removal to obtain estimates of the original information symbols. The channel response required by equalization has been estimated via the recursive error method.

The symbol error probability performance of the proposed MFT-ESA demodulation combined with ZF-DFE and MMSE-DFE for memory removal are compared in Fig. 5.5 for both binary and multilevel 3-REC CPM scenarios. Note that the modulation indices of the wideband CPM signals in this example are deliberately chosen such that the implementation of the VA is not practical due to the complexity of its phase states. For binary case, the MMSE-DFE performs slightly better than the zero-forcing DFE in the low SNR region, as shown in Fig. 5.5(a). For multilevel case, the performances of both approaches are almost the same as in Fig. 5.5(b). Since the multilevel signaling compress the decision region for symbol detection, the resolution of the symbol by symbol detector (or the slicer) is reduced, resulting in degraded
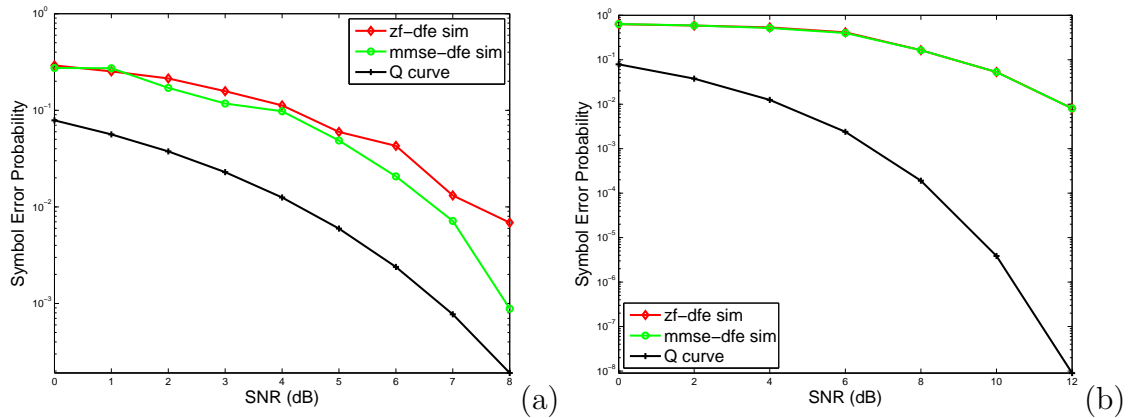
Figure 5.5: Memory removal via decision feedback equalization. (a) symbol error probability of MFT-ESA approach for binary CPFSK with zero-forcing and MMSE decision feedback equalization to remove memory induced due to partial response signaling, and (b) symbol error probability for multilevel CPFSK with zero-forcing and MMSE decision feedback equalization. In both cases 3-REC-CPM with parameters $T_b = 1s, f_s = 50Hz, f_c = 12Hz$ and iterations number of 10 was employed. The MFT conversion factor was $R = 16$ and the modulation indices for the binary and multilevel cases were $h = 97/21$ and $h = 19/15$ respectively. For equalization purposes, 50 pilot symbols were used. Note that the SEP of (a) and (b) will drop to zero around 11 dB and 16 dB respectively, which are not shown due to limitations of the log scale.

performances of both approaches at the same level in the extreme wideband case where the IF deviation is close to the carrier frequency.

From our previous analysis, we know that the error associated with output of the partial response channel extracted from the demodulated IF within each symbol period for low SNR is significant due to the error propagation. The performance of symbol detection for low SNR suffers serious degradation since the proposed memory removal approach is very sensitive to its input. However as the SNR increases, the output of the partial response channel obtained from MFT-ESA demodulation module become more precise, thereby resulting in significant improvement in the ability to recover the original symbols as evident in the symbol error probability. Above a SNR threshold of around 10 dB for the binary case and 12 dB for the

multilevel case, the symbol error probability becomes negligible attributable to the fact that inverse filtering solution becomes nearly perfect after that threshold, which has been verified but not shown in Fig. 5.5 due to limitations of the log scale.

## 5.5    Conclusion

In this chapter, we have presented an approach towards wideband CPM demodulation by extending the MFT-ESA approach. The characteristic features of the proposed approach are: 1) Unlike the *Viterbi algorithm* whose complexity increases with the number of phase states induced by $m$ and $p$ as in Eq. 5.10, the complexity of the proposed approach is independent of the modulation index; 2) The proposed approach does not require prior knowledge of the carrier frequency and this parameter can be computed from the IF estimates; 3) The proposed approach accommodates large modulation indices and multilevel signaling making it conducive to the large bandwidth requirements proposed in the M-ary FSK system for satellite communications [59].

The proposed MFT-ESA approach was then applied to the demodulation of wideband CPM signals with partial response signaling, where memory is introduced into the estimated IF. Subsequent to the MFT-ESA demodulation stage, a recursive prediction approach based on MA signal modeling of the estimated IF, together with decision feedback equalization, were presented to address the problem of removal of memory introduced by partial response signaling. Both the zero-forcing solution based on direct inversion of the memory channel, its corresponding decision feedback version and the MMSE-DFE solution to memory removal were investigated and shown to produce significant reduction in the symbol error probability over no equalization.

# Chapter 6

# Conclusions and Future Work

## 6.1 Conclusions

The objective of the dissertation is to present the MFT system that converts the wideband AM-FM signals to narrowband. Chapter 2 introduces the MFT framework that changes the CR/IB and CR/FD parameters of the signal via frequency compression, heterodyning and bandpass filtering. The alternative MFT framework that incorporates the use of large conversion factors has also been proposed with practical implementation of the bandpass filters by splitting the conversion factors and utilizing the multirate identity. The proposed MFT framework can be combined with various monocomponent AM-FM demodulation methods to improve their performances. Its compatibilities with two commonly used demodulation techniques, the HT and the ESA, have been analyzed in details. Chapter 3 extends the basic MFT to the BMFT in two dimensions and applies it to wideband images demodulation in combination with higher order energy operator. The BMFT-DHODA approach was further applied to wideband fingerprint images to demonstrate its efficacy. Chapter 4 utilizes the MFT-ESA combination for large deviation first formant

demodulation. The large deviation first formants are first extracted via the EMD and then demodulated by the MFT-ESA approach. The formant estimates computed from the demodulated IA and IF are demonstrated to be more reasonable than the widely used LPC approach that only assumes narrowband AM formant. Chapter 5 presented an approach towards the problem of wideband partial response CPM demodulation by combining the MFT-ESA based IF demodulation and decision feedback equalization. Unlike the VA whose complexity increases with the number of phase states, the proposed approach is not affected by the modulation index of the partial response CPM. The removal of the memory introduced by the partial response signaling is then achieved by decision feedback equalization.

## 6.2 Future Work

### 6.2.1 Multicomponent AM-FM Demodulation

We aimed at applying the MFT framework presented in this dissertation to the problem of wideband multicomponent AM-FM demodulation. In general, bandpass filtering is required for isolation of each component present in the multicomponent signal. The monocomponent demodulation scheme is then applied on each isolated AM-FM component to obtain the IA and IF estimates. However, the bandpass filtering may suffer from problems such as spectral overlapping. In prior work [60], the *energy demodulation of mixtures* (EDM) algorithm was proposed for the demodulation of two-component AM-FM signals. The EDM exploits the structural properties of a mixture of two sinusoids in terms of the solution to a *generating differential or difference equation* (GDE). The coeffcients of the GDE are then expressed via generalizations of the higher order differential energy operator to achieve separation and demodulation of the two-component signal. In this chapter, we propose the

combination of the MFT framework and the EDM algorithm to solve the problem of wideband multicomponent AM-FM demodulation.

## 6.2.2 EDM Algorithm

Assume that the two-component AM-FM signals with slowly time-varying IA and IF is instantaneously modeled by

$$s(t) = a_1 \cos(\omega_1 t + \theta_1) + a_2 \cos(\omega_2 t + \theta_2), \tag{6.1}$$

which satisfies the fourth-order GDE given by

$$s^{(4)} + c_1 s^{(2)} + c_2 s = 0, \tag{6.2}$$

where

$$s^{(n)} = \frac{\partial^n s}{\partial t^n}, \quad c_1 = (\omega_1^2 + \omega_2^2), \quad c_2 = \omega_1^2 \omega_2^2. \tag{6.3}$$

The differential energy operator of $k^{\text{th}}$-order is defined by

$$\Phi_k(s) = s^{(1)} s^{(k-1)} - s s^{(k)}. \tag{6.4}$$

The Teager-Kaiser energy operator used in the ESA is the special case when $k = 2$. By solving the resultant linear system of equations, the coefficients for the GDE can be expressed as

$$c_1 = -\frac{\Phi_5(s)}{\Phi_3(s)}, \quad c_2 = \frac{\Phi_3(s^{(2)})}{\Phi_3(s)}. \tag{6.5}$$

Hence the frequencies $\omega_1$ and $\omega_2$ can be estimated via

$$\omega_{1,2} = \sqrt{\frac{c_1 \pm \sqrt{c_1^2 - 4c_2}}{2}}. \tag{6.6}$$

Using the EDM frequency estimates, the amplitude $a_1$ and $a_2$ can be estimated via

$$a_{1,2} = \sqrt{\frac{\omega_{2,1}^4\left(\Phi(s^{(3)}) - \omega_1^2\omega_2^2\Phi(s^{(1)})\right)}{\omega_1^4\omega_2^4(\omega_1^2 - \omega_2^2)^2} - \frac{\omega_{2,1}^6\left(\Phi(s^{(2)}) - \omega_1^2\omega_2^2\Phi(s)\right)}{\omega_1^4\omega_2^4(\omega_1^2 - \omega_2^2)^2}}. \tag{6.7}$$

The estimated frequencies and amplitudes become approximations for the IA and IF estimates of each AM-FM component in the narrowband scenario. The discrete-time EDM algorithm can be easily derived from its continuous-time counterpart. Post-smoothing the output of the EDM yields the *smoothed* EDM (SEDM) algorithm that further reduce the demodulation error by suppressing the noise presented in the IA and IF estimates.

## 6.2.3 Performance of the MFT-SEDM

A wideband two-component example is presented here to demonstrate the efficacy of the proposed MFT-SEDM approach. Consider a discrete-time mixture signal of two sinusoidally modulated FM signals with a normalized sampling frequency $f_s = 1$:

$$s[n] = \sum_{i=1}^{2} a_i \cos\left(\int_0^n \Omega_i[m]dm\right), \tag{6.8}$$

$$\Omega_i[n] = \Omega_{ci} + \Omega_{mi}\cos\left(\Omega_{fi}n + \theta_i\right), \quad i = 1, 2. \tag{6.9}$$

The parameters of the two sinusoidal FM signals are given by

$$\Omega_{ci} = \frac{8\pi}{25}, \ \frac{16\pi}{25}, \quad \Omega_{mi} = \Omega_{fi} = 0.04\pi, \ 0.0432\pi, \quad \theta_i = 0, \ 0. \tag{6.10}$$

The CR/IB and CR/FD ratios for the discrete-time mixture signal is given by

$$\text{CR/IB}_{old,i} = 8, \ \frac{400}{27}, \quad \text{CR/FD}_{old,i} = 8, \ \frac{400}{27}. \tag{6.11}$$

In prior work [60], the *normalized carrier separation* (SEP) parameter that measures the spectral separation between the two component is defined as:

$$\text{SEP} = \frac{|\Omega_{c2} - \Omega_{c1}|}{\sum_i \left(\Omega_{fi} + \Omega_{ai} + \Omega_{mi}\right)}, \tag{6.12}$$

where we have SEP $= 1.923$ in this example. The power spectrum of the mixture signal is illusrated in Fig. 6.1(a). By performing the MFT using a conversion factor $R = 16$ and a translation frequency $w_d = \frac{23\pi}{100}$, the mixture signal is transformed into a narrowband signal with power spectrum illustrated in Fig. 6.1(b). The CR/IB and CR/FD ratios for each transformed component is computed by

$$\text{CR/IB}_{new,i} = 100, 100, \quad \text{CR/FD}_{new,i} = 100, 100. \tag{6.13}$$

Since the MFT transforms each component in a uniform way, the SEP of the transformed mixture signal remains the same.

Then we apply the SEDM algorithm to the transformed mixture signal and recover the IF estimates for each component in the original mixture signal according to the inverse MFT relation. The demodultion results have been shown in Fig. 6.1(c), justifying the validity of the MFT-SEDM approach for solving the wideband multicomponent demodulation problem.

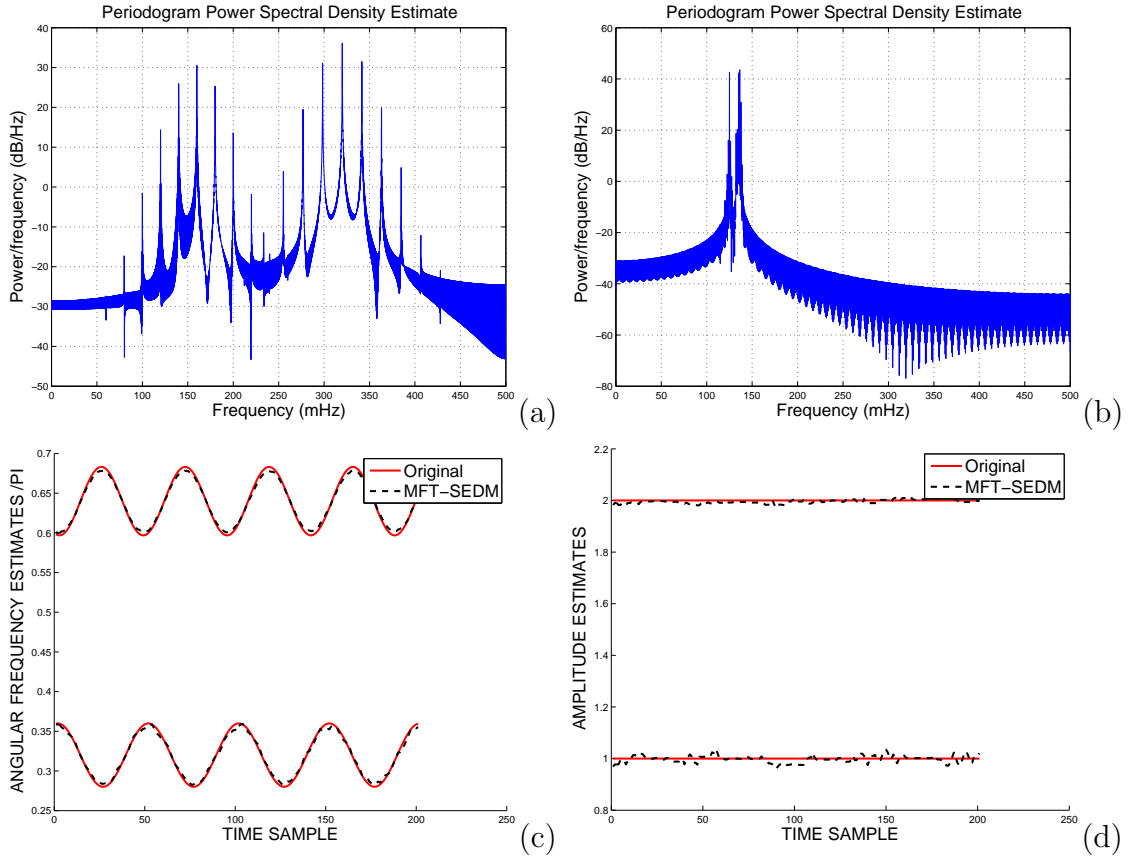Figure 6.1: Two-component AM-FM demodulation via MFT-SEDM approach.  a) Power spectrum of the original mixture signal.  b) Power spectrum of the mixture signal after MFT.  c) IF estimates for each component of the mixture signal via the MFT-SEDM approach.  d) IA estimates for each component of the mixture signal via the MFT-SEDM approach.

# Appendices

**A  Error Analysis for Extremely Wideband FM Signal**

# Appendix A

# Error Analysis for Extremely Wideband FM Signal

Here we present the error analysis for wideband sinusoidal FM demodulation in 1D [1]. According to the handbook of mathematic formula and basic trigonometric identities, we have

$$\cos(\beta \sin \theta) = J_0(\beta) + 2 \sum_{k=1}^{\infty} J_{2k}(\beta) \cos(2k\theta),$$

$$\sin(\beta \sin \theta) = 2 \sum_{k=0}^{\infty} J_{2k+1}(\beta) \sin((2k+1)\theta),$$

$$\cos(u)\cos(v) = \frac{1}{2}[\cos(u-v) + cos(u+v)],$$

$$\sin(u)\sin(v) = \frac{1}{2}[\cos(u-v) - cos(u+v)],$$

where $J_n(\beta)$ denotes the Bessel function of the first kind.

---

[1] Since the partial *Hilbert transform* is a separable combination of the 1D *Hilbert transform*, consequently this analysis is valid for the proposed 2D approach as well.

*Appendix A. Error Analysis for Extremely Wideband FM Signal*

Using above equations, the sinusoidal FM can be expanded as

$$\cos(\omega_c t + \beta \sin(\omega_m t))$$

$$= \cos(\omega_c t) \cos\left(\beta \sin(\omega_m t)\right) - \sin(\omega_c t) \sin\left(\beta \sin(\omega_m t)\right)$$

$$= \cos(\omega_c t) J_0(\beta) + 2 \sum_{k=1}^{\infty} J_{2k}(\beta) \cos(2k\omega_m t) \cos(\omega_c t)$$

$$- 2 \sum_{k=0}^{\infty} J_{2k+1}(\beta) \sin((2k+1)\omega_m t) \sin(\omega_c t)$$

$$= \cos(\omega_c t) J_0(\beta) + 2 \sum_{k=1}^{\infty} J_{2k}(\beta) \cos\left((\omega_c + 2k\omega_m)t\right)$$

$$+ 2 \sum_{k=1}^{\infty} J_{2k}(\beta) \cos\left((\omega_c - 2k\omega_m)t\right)$$

$$+ 2 \sum_{k=0}^{\infty} J_{2k+1}(\beta) \cos\left((\omega_c t + (2k+1)\omega_m)t\right)$$

$$- 2 \sum_{k=0}^{\infty} J_{2k+1}(\beta) \cos\left((\omega_c - (2k+1)\omega_m)t\right)$$

Its quadrature part we desire for demodulation can also be expanded similarly given by

$$\sin\left(\omega_c t + \beta \sin(\omega_m t)\right)$$

$$= \sin(\omega_c t) \cos\left(\beta \sin(\omega_m t)\right) + \cos(\omega_c t) \sin\left(\beta \sin(\omega_m t)\right)$$

$$= \sin(\omega_c t) J_0(\beta) + 2 \sum_{k=1}^{\infty} J_{2k}(\beta) \cos(2k\omega_m t) \sin(\omega_c t)$$

$$+ 2 \sum_{k=0}^{\infty} J_{2k+1}(\beta) \sin((2k+1)\omega_m t) \cos(\omega_c t)$$

$$= \sin(\omega_c t) J_0(\beta) + 2 \sum_{k=1}^{\infty} J_{2k}(\beta) \sin\left((\omega_c + 2k\omega_m)t\right)$$

$$+ 2 \sum_{k=1}^{\infty} J_{2k}(\beta) \left(\sin(\omega_c - 2k\omega_m)t\right) + 2 \sum_{k=0}^{\infty} J_{2k+1}(\beta) \sin\left((\omega_c t + (2k+1)\omega_m)t\right)$$

$$- 2 \sum_{k=0}^{\infty} J_{2k+1}(\beta) \sin\left((\omega_c - (2k+1)\omega_m)t\right)$$

*Appendix A.  Error Analysis for Extremely Wideband FM Signal*

Note that the *Hilbert transform* of the corresponding sinusoidal FM can be easily obtained by performing the *Hilbert transform* at each of its harmonics as

$$\mathcal{H}[\cos\left(\omega_c t + \beta \sin\left(\omega_m t\right)\right)]$$

$$= \sin(\omega_c t) J_0(\beta) + 2\sum_{k=1}^{\infty} J_{2k}(\beta) \sin\left(\omega_c + 2k\omega_m\right)t)$$

$$+ 2\sum_{k=1}^{\infty} J_{2k}(\beta) \sin\left(|\omega_c - 2k\omega_m|t\right)$$

$$+ 2\sum_{k=0}^{\infty} J_{2k+1}(\beta) \sin\left(\left(\omega_c + (2k+1)\omega_m\right)t\right)$$

$$- 2\sum_{k=0}^{\infty} J_{2k+1}(\beta) \sin\left(|\omega_c - (2k+1)\omega_m|t\right)$$

By comparing the quadrature and the *Hilbert transform* of the corresponding sinusoidal FM, we observe that when $\omega_c < n\omega_m$,

$$\mathcal{H}[\cos\left((\omega_c - n\omega_m)t\right)] = \sin(n\omega_m - \omega_c)t,$$

whereas the corresponding part in the quadrature part is of opposite sign $\sin(\omega_c - n\omega_m)t$. For narrowband sinusoidal FM, the Bessel coefficients $J_n(\beta)$ associated with these corresponding sidelobes are very small, and the demodulation error incurred by the *Hilbert transform* due to these opposite sign terms can be neglected. Hence the HT, provides almost perfect reconstruction of the quadrature part we desire for demodulation using the *analytic signal*, for sinusoidal FM, in the narrowband case. For extremely wideband sinusoidal FM, however, due to its large modulation index $\beta$, the Bessel coefficients associated with these sidelobes cannot be neglected anymore. As a result, the error incurred by the HT is much more significant. In general, this unavoidable error imposes a lower bound on the error performance for HT and energy operator in the extremely wideband scenario.

# References

[1] S. Haykin, *Communication Systems*, 5th ed.   Wiley Publishing, 2009.

[2] M. Gupta and B. Santhanam, "Adaptive linear predictive frequency tracking and cpm demodulation," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 1, 2003, pp. 202–206 Vol.1.

[3] B. Santhanam, "Generalized energy demodulation for large frequency deviations and wideband signals," *IEEE Signal Process. Lett.*, vol. 11, no. 3, pp. 341–344, 2004.

[4] S. S. Haykin, *Adaptive Filter Theory*, 4th ed.   Upper Saddle River, NJ, USA: Prentice Hall Press, 2005.

[5] L. J. Griffiths, "Rapid measurement of digital instantaneous frequency," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 23, no. 2, pp. 207–222, 1975.

[6] D. Mandic, "A generalized normalized gradient descent algorithm," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 115–118, 2004.

[7] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with application to speech analysis," *IEEE Transactions on Signal Processing*, vol. 41, no. 10, pp. 3024–3051, Oct 1993.

[8] ——, "On amplitude and frequency demodulation using energy operators," *IEEE Transactions on Signal Processing*, vol. 41, no. 4, pp. 1532–1550, Apr 1993.

[9] W. Liu and B. Santhanam, "Wideband-fm demodulation for large wideband to narrowband conversion factors via multirate frequency transformation," in *Proceedings of Signal Processing and SP Education Workshop*, 2015, pp. 97–102.

*References*

[10] A. V. Oppenhein and R. W. Schafer, *Discrete-time signal processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009.

[11] W. Zongbo, G. Meiguo, L. Yunjie, and J. Haiqing, "Design and application of drfm system based on digital channelized receiver," in *Radar, 2008 International Conference on*, 2008, pp. 375–378.

[12] F. F. Kuo and J. F. Kaiser, *System analysis by digital computer.* Wiley, 1966.

[13] B. Santhanam, "Orthogonal modes of frequency modulation and the sturm-liouville frequency modulation model," *IEEE Transactions on Signal Processing*, vol. 60, no. 7, pp. 3486–3495, 2012.

[14] A. Potamianos and P. Maragos, "A comparison of the energy operator and the hilbert transform approach to signal and speech demodulation," *Signal Processing*, vol. 37, no. 1, pp. 95–120, 1994.

[15] F. Salzenstein and A.-O. Boudraa, "Multi-dimensional higher order differential operators derived from the teager–kaiser energy-tracking function," *Signal Processing*, vol. 89, no. 4, pp. 623–640, 2009.

[16] F. Salzenstein, A.-O. Boudraa, and T. Chonavel, "A new class of multi-dimensional teager-kaiser and higher order operators based on directional derivatives," *Multidimensional Systems and Signal Processing*, pp. 1–30, 2013.

[17] J. G. Daugman and C. J. Downing, "Demodulation, predictive coding, and spatial vision," *J. Opt. Soc. Am. A*, vol. 12, no. 4, pp. 641–660, Apr 1995. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-12-4-641

[18] M. S. Pattichis, G. Panayi, A. C. Bovik, and S.-P. Hsu, "Fingerprint classification using an am-fm model," *IEEE Transactions on Image Processing*, vol. 10, no. 6, pp. 951–954, Jun 2001.

[19] J. Havlicek, D. Harding, and A. Bovik, "Discrete quasi-eigenfunction approximation for am-fm image analysis," in *Image Processing, 1996. Proceedings., International Conference on*, vol. 1, Sep 1996, pp. 633–636 vol.1.

[20] A. Bovik, N. Gopal, T. Emmoth, and A. Restrepo, "Localized measurement of emergent image frequencies by gabor wavelets," *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 691–712, March 1992.

[21] N. E. Huang, *Hilbert-Huang transform and its applications.* World Scientific, 2014, vol. 16.

*References*

[22] J. Havlicek, J. Havlicek, and A. Bovik, "The analytic image," in *Image Processing, 1997. Proceedings., International Conference on*, vol. 2, Oct 1997, pp. 446–449 vol.2.

[23] P. Maragos and A. C. Bovik, "Image demodulation using multidimensional energy separation," *J. Opt. Soc. Am. A*, vol. 12, no. 9, pp. 1867–1876, Sep 1995.

[24] I. Kokkinos, G. Evangelopoulos, and P. Maragos, "Advances in texture analysis: energy dominant component multiple hypothesis testing," in *Image Processing, 2004. ICIP '04. 2004 International Conference on*, vol. 3, Oct 2004, pp. 1509–1512 Vol. 3.

[25] V. Murray, P. Rodriguez, and M. Pattichis, "Multiscale am-fm demodulation and image reconstruction methods with improved accuracy," *Image Processing, IEEE Transactions on*, vol. 19, no. 5, pp. 1138–1152, May 2010.

[26] K. G. Larkin, D. J. Bone, and M. A. Oldfield, "Natural demodulation of two-dimensional fringe patterns. i. general background of the spiral phase quadrature transform," *J. Opt. Soc. Am. A*, vol. 18, no. 8, pp. 1862–1870, Aug 2001. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-18-8-1862

[27] K. G. Larkin, "Natural demodulation of two-dimensional fringe patterns. ii. stationary phase analysis of the spiral phase quadrature transform," *J. Opt. Soc. Am. A*, vol. 18, no. 8, pp. 1871–1881, Aug 2001. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-18-8-1871

[28] M. Unser, D. Sage, and D. Van De Ville, "Multiresolution monogenic signal analysis using the riesz–laplace wavelet transform," *Image Processing, IEEE Transactions on*, vol. 18, no. 11, pp. 2402–2418, 2009.

[29] W. Liu and B. Santhanam, "Wideband image demodulation via bi-dimensional multirate frequency transformations," *J. Opt. Soc. Am. A*, vol. 33, no. 9, pp. 1668–1678, Sep 2016. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-33-9-1668

[30] M. Pattichis and A. Bovik, "Analyzing image structure by multidimensional frequency modulation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 29, no. 5, pp. 753–766, May 2007.

[31] J. Havlicek, J. Havlicek, N. Mamuya, and A. Bovik, "Skewed 2d hilbert transforms and computed am-fm models," in *Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on*, vol. 1, Oct 1998, pp. 602–606 vol.1.

*References*

[32] F. Salzenstein, A.-O. Boudraa, and J.-C. Cexus, "Generalized higher-order nonlinear energy operators," *J. Opt. Soc. Am. A*, vol. 24, no. 12, pp. 3717–3727, Dec 2007. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-24-12-3717

[33] E.-H. Diop, A. Boudraa, and F. Salzenstein, "Higher order teager-kaiser operators for image analysis: Part i - a monocomponent image demodulation," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, April 2009, pp. 1041–1044.

[34] E. H. S. Diop and A. O. Boudraa, "Higher order teager-kaiser operators for image analysis: Part ii - a multicomponent image demodulation," in *2009 16th IEEE International Conference on Image Processing (ICIP)*, Nov 2009, pp. 1357–1360.

[35] P. Maragos and A. Potamianos, "Higher order differential energy operators," *IEEE Signal Processing Letters*, vol. 2, no. 8, pp. 152–154, Aug 1995.

[36] W. Liu and B. Santhanam, "Wideband fingerprint demodulation via bi-dimensional multirate frequency transformations," in *2016 50th Asilomar Conference on Signals, Systems and Computers*, Nov 2016, pp. 576–579.

[37] J. P. Havlicek, "Am-fm image models," 1996.

[38] L. Rabiner and R. Schafer, *Theory and Applications of Digital Speech Processing*, 1st ed. Upper Saddle River, NJ, USA: Prentice Hall Press, 2010.

[39] A. Potamianos and P. Maragos, "Speech formant frequency and bandwidth tracking using multiband energy demodulation," *The Journal of the Acoustical Society of America*, vol. 99, no. 6, pp. 3795–3806, 1996.

[40] ——, "Speech analysis and synthesis using an am–fm modulation model," *Speech Communication*, vol. 28, no. 3, pp. 195–209, 1999.

[41] P. Tsiakoulis and A. Potamianos, "On the effect of fundamental frequency on amplitude and frequency modulation patterns in speech resonances," *Proc. INTERSPEECH*, pp. 649–652, 2010.

[42] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

*References*

[43] G. S. M. A. Colominas and M. Torres, "Improved complete ensemble emd: A suitable tool for biomedical signal processing," *Biomed. Sig. Process. and Control*, vol. 14, pp. 19–29, 2014.

[44] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Advances in adaptive data analysis*, vol. 1, no. 01, pp. 1–41, 2009.

[45] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Acoustics, speech and signal processing (ICASSP), 2011 IEEE international conference on*. IEEE, 2011, pp. 4144–4147.

[46] M. J. Gertsman and J. H. Lodge, "Symbol-by-symbol map demodulation of cpm and psk signals on rayleigh flat-fading channels," *IEEE Transactions on Communications*, vol. 45, no. 7, pp. 788–799, 1997.

[47] T. A. J. B. Anderson and C.-E. W. Sundberg, *Digital Phase Modulation*. New York: Plenum, 1986.

[48] C. E. Sundberg, "Continuous phase modulation," *IEEE Communications Magazine*, vol. 24, pp. 25–38, April 1986.

[49] S. Li, N. Zhang, S. Lin, L. Kong, A. Katangur, M. K. Khan, M. Ni, and G. Zhu, "Joint admission control and resource allocation in edge computing for internet of things," *IEEE Network*, vol. 32, no. 1, pp. 72–79, 2018.

[50] S. Lin, L. Kong, Q. Gao, M. K. Khan, Z. Zhong, X. Jin, and P. Zeng, "Advanced dynamic channel access strategy in spectrum sharing 5g systems," *IEEE Wireless Communications*, vol. 24, no. 5, pp. 74–80, 2017.

[51] A. Perotti, A. Tarable, S. Benedetto, and G. Montorsi, "Capacity-achieving cpm schemes," *IEEE Transactions on Information Theory*, vol. 56, no. 4, pp. 1521–1541, 2010.

[52] W. Xue, W. Shang, S. B. Makarov, and Y. Xu, "A phase trajectories optimization method for cpm signal based on pan-function model," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 1, p. 55, May 2016.

[53] J. G. Proakis and M. Salehi, *Digital Communications, Fifth edition*. New York: McGraw-Hill Publishing Company, 1995.

[54] S. S. Abayesekara, "Robust full response m-ary raised-cosine cpm receiver design via frequency estimation," in *Proceedings of International Conference on DSP*, 2015, pp. 935–939.

*References*

[55] W. Liu and B. Santhanam, "Wideband partial response cpm demodulation via multirate frequency transformations and decision feedback equalization," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, p. 91, May 2018. [Online]. Available: https://doi.org/10.1186/s13638-018-1097-2

[56] L. Ljung, *System Identification: Theory for the User, Second edition.* New Jersey: Prentice Hall, 1999.

[57] J. H. R. Fischer and C. Windpassinger, "Signal processing in decision-feedback equalization of intersymbol-interference and multiple-input/multiple-output channels: a unified view," *Signal Processing*, vol. 83, no. 8, pp. 1633–1642, 2003.

[58] V. Kavitha and V. Sharma, "Optimal mse solution for a decision feedback equalizer," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, p. 172, Aug 2012.

[59] M. Fitch and K. Briggs, "Gaussian multilevel fm for high-bandwidth satellite communications," *University College of London*, 2004.

[60] B. Santhanam and P. Maragos, "Energy demodulation of two-component am-fm signals with application to speaker separation," in *1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings*, vol. 6, May 1996, pp. 3517–3520 vol. 6.