Frequency Estimation for Single-Carrier and OFDM Signals in Communication and Radar Systems

Vom Fachbereich Elektrotechnik und Informatik der
Universität Siegen
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
(**Dr.-Ing.**)

genehmigte Dissertation

von

M.Sc. Pakorn Ubolkosold

1. Gutachter: Prof. Dr.-Ing. habil. O. Loffeld
2. Gutachter: Prof. Dr. techn. Dr. h.c. B. Hofmann-Wellenhof
Vorsitzender: Prof. Dr.-Ing. H. Roth

Tag der mündlichen Prüfung: 15.04.2009

Contents

List of Figures

List of Tables

Acknowledgments

I would like to take this opportunity to express my greatest thanks to all of you who have supported me in various ways. It is a great pleasure for me to express my sincere gratitude to Prof. Dr.-Ing. habil. Otmar Loffeld, who supervised this work with discretion. During the last three years, I have been nourished with his broad knowledge of applied estimation theory, inspired by his brilliant insights, and guided by his scientific rigorousness. At the same time, he endowed me with maximal freedom for developing ideas. I am also indebted to him for the helps with various non-scientific matters, to which he has never been indifferent. Above all, I enjoyed a lot working with him.

I am grateful to the thesis committee: Prof. Dipl.-Ing. Dr. techn. Dr. h.c. Bernhard Hofmann-Wellenhof and Prof. Dr.-Ing. Hubert Roth, for their careful reading of the manuscript. I thank to Dr.-Ing. Stefan Knedlik for his continuous encouragement and for letting me involved in a project "Attitude and Position Determination (AtPos) for Bistatic SAR Experiments" funded by the German Research Foundation (DFG). I thank Dr. Iyad Abuhadrous for the intuitive discussion on the integration of GPS and INS sensors. I would like to thank all AtPos team members for sharing with me valuable ideas during regular AtPos meetings.

I would like to thank to all colleagues in the Center for Sensor Systems (ZESS) and members of International Postgraduate Programme (IPP). Many thanks to M.Sc. Gustave F. Tchere and M.Sc. Miao Zhang for numerous discussions on both scientific and non-scientific matters and their friendly helps with various problems. I thank to M.Sc. Amaya Medrano-Ortiz for organizing many interesting social events. I thank to Arne Stadermann for his support with computer related matters. I am grateful to Renate Szabó, Ira Dexling, and Silvia Niet-Wunram for their excellent administrative services.

Finally, I dedicate this thesis to my parents, whose love is more than I can describe.

Abstract

Estimating the frequency of a signal embedded in additive white Gaussian noise is one of the classical problems in signal processing. It is of fundamental importance in various applications such as in communications, Doppler radar, synthetic aperture radar (SAR), array processing, radio frequency identification (RFID), resonance sensor, etc.

The requirement on the performance of the frequency estimator varies with the application. The performance is often defined using four indexes: *i*). estimation accuracy, *ii*). estimation range, *iii*). estimation threshold, and *iv*). implementation complexity. These indexes may be in contrast with each other. For example, achieving a low threshold usually implies a high complexity. Likewise, good estimation accuracy is often obtained at the price of a narrow estimation range. The estimation becomes even more difficult in the presence of fading-induced multiplicative noise which is considered to be the general case of the frequency estimation problem. There have been a lot of efforts in deriving the estimator for the general case, however, a generalized estimator that fulfills all indexes can be hardly obtained.

Focusing on communications and radar applications, this thesis proposes a new generalized closed-form frequency estimator that compromises all performance indexes. The derivation of the proposed estimator relies on the nonlinear least-squares principle in conjunction with the well known summation-by-parts formula. In addition to this, several modified frequency estimators suitable for non-fading or very slow fading scenarios, are also introduced in this thesis.

Kurzfassung

Eine der klassischen Problemstellungen in der Signalverarbeitung ist die Schätzung der Frequenz eines Signals, das von weißem Rauschen additiv überlagert ist. Diese bedeutende Aufgabe stellt sich in vielen verschiedenen Anwendungsbereichen wie der Kommunikationstechnik, beim Doppler-Radar, beim Radar mit synthetischer Apertur (SAR), beim Array Processing, bei Radio-Frequency-IDentification (RFID), bei Resonanz-Sensoren usw.

Die Anforderungen bezüglich der Leistungsfähigkeit des Frequenzschätzers hängen von der Anwendung ab. Die Leistungsfähigkeit ist dabei oft unter Berücksichtigung der folgenden 4 Punkte definiert: *i*) Genauigkeit, Richtigkeit der Schätzung, *ii*) Arbeitsbereich (estimation range), *iii*) Grenzwerte der Schätzung (im Vergleich zu einer theoretisch möglichen Schwelle) und *iv*) Komplexität der Implementierung. Diese Anforderungen können nicht unabhängig voneinander betrachtet werden und stehen sich teilweise gegenüber. Beispielsweise erfordert die Erzielung von Ergebnissen nahe an der theoretisch möglichen Schwelle eine hohe Komplexität. Ebenso kann ein Schätzergebnis von hoher Genauigkeit oftmals nur für einen stark eingeschränkten Arbeitsbereich erzielt werden. Die Frequenzschätzung ist im Falle von durch Fading hervorgerufenem multiplikativem Rauschen noch herausfordernder. Es handelt sich dann um den allgemeinen Fall der Frequenzschätzung. Bisher hat man bereits viel Arbeit in die Ableitung eines Schätzers fr diesen allgemeinen Fall investiert. Ein Schätzer, der optimal bezüglich aller oben genannten Kriterien ist, dürfte allerdings nur schwer zu finden sein.

In dieser Dissertation wird mit Blick auf Kommunikationstechnik und Radaranwendungen ein verallgemeinerter, in geschlossener Form vorliegender, Frequenzschätzer eingeführt, der alle genannten Kriterien der Leistungsfähigkeit berücksichtigt. Die Herleitung des Schätzers beruht auf dem Prinzip der kleinsten Fehlerquadrate für den nichtlinearen Fall in Verbindung mit der Abelschen partiellen Summation. Zudem werden verschiedene modifizierte Frequenzschätzer vorgestellt, die sich für Fälle in denen kein Fading oder nur sehr geringes Fading auftritt, eignen.

List of Abbreviations

ACF	Auto Correlation Function
ANLS	Approximated Nonlinear Least Squares
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BLUE	Best Linear Unbiased Estimator
CFO	Carrier Frequency Offset
CMF	Channel Match Filter
CP	Cyclic Prefix
CSI	Channel State Information
CRLB	Cramer Rao Lower Bound
DAB	Digital Audio Broadcast
DFE	Decision Feedback Equalizer
DFT	Discrete Fourier Transform
DOA	Direction of Arrival
DVB	Digital Video Broadcast
EGC	Equal Gain Combining
EKF	Extended Kalman Filter
FFT	Fast Fourier Transform
FIM	Fisher Information Matrix
GPS	Global Positioning System
GRV	Gaussian Random Variable
GSM	Global System for Mobile Communications
ICI	Inter Carrier Interference
IFFT	Inverse Fast Fourier Transform
i.i.d.	Independent Identically Distributed
INS	Inertial Navigation System
ISI	Inter Symbol Interference
KF	Kalman Filter
LKF	Linearized Kalman Filter
LOS	Line of Sight

MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimator
MPSK	M-ary Phase Shift Keying
MRC	Maximum Ratio Combining
MSE	Mean Squared Error
NLS	Nonlinear Least Squares
NSC	Null Subcarrier
OFDM	Orthogonal Frequency Division Multiplexing
PAPR	Maximum Ratio Combining
PSD	Power Spectral Density
PSK	Phase Shift Keying
PN	Pseudonoise
QAM	Quadrature Amplitude modulation
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
RFID	Radio Frequency Identification
SAR	Synthetic Aperture Radar
SAV	Simple Average
SC	Selection Combining
SCA	Schmidl and Cox Algorithm
SNLS	Simplified Nonlinear Least Squares
SNR	Signal to Noise Ratio
SIMO	Single Input Multiple Output
SISO	Single Input Single Output
UAV	Unweighted Average
UMTS	Universal Mobile Telecommunications System
UKF	Unscented Kalman Filter
ULA	Uniform Linear Array
VA	Viterbi Algorithm
WLAN	Wireless Local Area Network
ZF	Zero Forcing

Chapter 1

Introduction

The subject of this thesis is frequency estimation of signals embedded in additive white Gaussian noise for communication and radar related applications. These two applications may at the first glance seem rather different in the goals they wish to accomplish. Communication systems transmit information from one place to another, while radar systems are sensing devices. Despite this difference of purpose, their similarities for certain aspects are greater, and that is the reason they can be studied together in a unified manner. Radar shares many common properties and technologies with communication systems. Both systems use signals for transmission, so signal theory becomes the common background. Both must convert this signal to electromagnetic waves by using devices that operate in the same manner. The waves travel through media that are similar in both cases. On the receiving side, both systems must receive a signal, usually contaminated by noise, and the information that it carries must be extracted.

The estimation of the frequency of a signal embedded in additive white Gaussian noise is one of the classical problems in signal processing. Frequency estimation has been continuously explored for decades and still increasingly gained attention in the recent years. One reason for this is that the problem is relatively easy to understand, but difficult to solve. Another reason, certainly, is the large number of applications that involves frequency estimation, e.g., in *communications*, the frequency offset due to mismatch between the received signal carrier and the local oscillator frequencies, shall be estimated and compensated for; in *Doppler radar*, the Doppler frequency that contains the information about the range/velocity of the target is of interest; in *array processing*, the spatial frequency that is related to the direction of the arrival

of the source is to be determined; in *radio frequency identification* (RFID) systems, the frequency modulation is used in the communications link; and in *resonance sensor* systems, the output signal is given by the frequency displacement from a nominal frequency, etc.

The requirement on the performance of the frequency estimator varies with the application. The performance is often defined using four indexes: *estimation accuracy*, *estimation range*, *estimation threshold*, and *implementation complexity*. These indexes are often in contrast with each other. For example, achieving a low threshold usually implies a high complexity. Likewise, good estimation accuracy is generally obtained at the price of a narrow estimation range.

The fast evolving in the field of digital wireless communications, has opened up several new challenges in frequency estimation (carrier frequency offset estimation). Signal models that have been used as the air-interface for each phase of the evolution are basically different e.g., GSM uses single-carrier signals, UMTS uses spread-spectrum signals, and next generation UMTS uses multiple-carrier OFDM signals. Therefore, new techniques and algorithms have been developed to provide optimal performance. Most of the existing accurate frequency estimators require either high computational complexity, additional phase unwrapping, or information about the propagation channel. These requirements make most of them unattractive.

This thesis focuses on the frequency estimation problems in wireless communication and radar systems. The first part of this thesis is devoted to the development of carrier frequency offset (CFO) estimators that can fulfill the four performance indexes mentioned above. The CFO estimators are developed based upon the application in wireless communications. Later, it will be also shown that the developed concepts are applicable for the radar related areas such as Doppler centroid estimation in Synthetic Aperture Radar (SAR) and Direction-Of-Arrival (DOA) estimation in array processing.

The organization of this thesis is as follows:

- **Chapter 2** describes the basic principle of signal transmission schemes used in communication and radar systems. This includes single-carrier and multiple-carrier signals. Single-carrier signals are often used in radar and narrowband communications applications, while multiple-carrier signals, i.e., orthogonal frequency division multiplexing (OFDM), are employed in modern wireless communication systems. The effect of carrier frequency synchronization errors of both type of signals are also formulated. Antenna diversity combining techniques that can be used to reduce the effect of multipath fading are introduced at the end of this chapter.

- **Chapter 3** addresses the problem of CFO estimation for single-carrier signals. The classical state of the art methods are first described with the detailed derivations given in the appendix. CFO estimation for constant and time-varying envelope models of the signals are treated separately. In the constant envelope case, four new estimators have been proposed. The first one has an improved threshold as well as lower computational complexity as compared to Kay estimator. The second estimator extends the estimation range of Fitz estimator to it's maximum without any additional phase unwrapping algorithm. The third is an approximated maximum likelihood estimator based on the absolute phase of the correlation estimates. The fourth is obtained by applying the weight transformation formula to the third estimator as done for the second estimator. For the time-varying envelope case, two new estimators have been proposed. The first estimator relies on the nonlinear least squares principle in conjunction with the summation-by-parts formula which is simple and does not require the knowledge of the form of the fading correlation as required for most of the existing estimators. The second estimator is designed to track time-varying CFO which is based on the Kalman filter. The contributions to this chapter can be found in [1–5].

- **Chapter 4** deals with the CFO problem for OFDM signals. The state of the art techniques of time-domain (pre-FFT) CFO estimation are briefly discussed. In this chapter, a new CFO estimator for single-input single-output (SISO) OFDM, which is based on the nonlinear least squares estimation concept in conjunction with the summation-by-part formula, is developed. The proposed estimator has been extended to the case of single-input multiple-output (SIMO) in order to resolve the error-floor of the estimation variance in the SISO case. The contributions to this chapter can be found in [6].

- **Chapter 5** formulates the problem of frequency estimation in the context of Doppler centroid estimation in SAR system and DOA estimation in array processing. Moreover, a new integrated GPS/INS for DOA estimation is also proposed. The contribution to this chapter can be found in [7].

- In **Chapter 6** conclusions are provided which summarize the major results obtained in this thesis and outline possible future research work in this field.

Chapter 2

Transmission Schemes

Digital bandpass modulation techniques can be broadly classified into two categories. The first is *single-carrier* modulation, where data is transmitted by using a single radio frequency carrier. The other is *multi-carrier* modulation, where data is transmitted by simultaneously modulating multiple RF carriers. The transmission of high data rates generally implies a small symbol duration T_s. Due to multipath propagation in the radio channel, distortions are observed in the received signal, which appear as *inter-symbol-interference* (ISI) of the successive modulation symbols. This situation is technically very critical when the maximum delay τ_{max} is very large, compared to the symbol duration T_s. In this case, the ISI affects many adjacent transmitted symbols. The radio channel properties resulting from multipath propagation result in a frequency-selective behavior of the channel transfer function.

For the demodulation of the received signal, the impulse response of the channel has to be measured and the signal must be equalized. The complexity of such a time-domain equalizer is at least proportional to the maximum propagation delay, $i = \frac{\tau_{max}}{T_s} \propto$ Equalizer complexity. A narrowband channel corresponds to a high symbol duration $T_s \gg \tau_{max}$. Small ISI are generated that affect only fractions of adjacent symbols. These distortions can be compensated by a simple equalizer. In GSM, a maximum delay of $\tau_{max} = 20\mu s$ is expected. With a symbol duration of $T_s = 4\mu s$, the processing cost in an equalizer can be implemented by $i = 5$ coefficients. For high data rate systems, a broadband channel with small symbol duration $T_s \ll \tau_{max}$ is needed. Thus, ISI span many symbols. As a result, the time-domain equalizer becomes complex or is even not realizable. As an example, in Digital Audio Broadcasting (DAB) the maximum delay is typically $\tau_{max} = 50\mu s$

and the symbol duration (for single-carrier system) would be $T_s = 0.5\mu s$, consequently an equalizer with 100 coefficients would be required.

From this description, the concept of multi-carrier modulation in particular orthogonal frequency division multiplexing (OFDM) is derived. In the frequency domain, a broadband channel is divided into many parallel narrowband subchannels. Each subchannel will then be seen as a frequency flat-fading rather than frequency selective-fading channel which significantly simplifies the channel equalizer. In the time domain, the symbol duration of each subchannel is increased and subsequently the ISI problem is reduced. OFDM, however, suffers from RF impairments such as the high sensitivity of the carrier frequency offset and high peak-to-average power ratio (PAPR). These disadvantages play a less important role in the single-carrier case. Moreover, investigations also suggest that single-carrier systems with properly designed frequency-domain equalizer have similar performance, efficiency, and low signal processing complexity advantages as OFDM.

This chapter introduces the fundamentals of single- and multi-carrier modulations in wireless communications. The carrier frequency offset (CFO) in both systems is also described. The receiving antenna diversity techniques which can be used to reduce the effect of multipath fading in the received signal are introduced.

2.1 Single-Carrier Transmission

In this section, the conventional baseband single-carrier transmission is described. The complete transmission block diagram can be seen in Fig. 2.1.

2.1.1 Basic principle

The equivalent lowpass transmitted signal $s_T(t)$ has the following description

$$s_T(t) = \sum_{k=-\infty}^{\infty} x(k) \cdot e_T(t - kT_s), \quad x(k) \in \mathcal{A}_x \subset \mathbb{C}, \quad e_T(t) \in \mathbb{C} \qquad (2.1)$$

where $\mathcal{A}_x \in \{A_i = e^{j\left(\frac{2\pi}{M} i + \varphi_M\right)}, \quad i = 0, \ldots, M-1\}$ for any $\varphi_M \in \mathbb{R}$, is the transmitted symbol alphabet, taken for example from the M-PSK constellation, and $e_T(t)$ is the basic waveform (e.g., rectangular or raised cosine

shape). The bandpass transmitted signal can be obtained by lowpass-to-bandpass (LP-BP) transform

$$s(t) = \mathrm{Re}\left\{s_T(t)e^{j2\pi f_0 t}\right\} = \mathrm{Re}\left\{\sum_{k=-\infty}^{\infty} x(k)\cdot e_T(t-kT_s)e^{j2\pi f_0 t}\right\} \tag{2.2}$$

where f_0 is the carrier frequency. If $x(k) = |x(k)|e^{j\psi(k)}$ with $|x(k)| = \sqrt{2}$ where $\psi(k)$ is the phase of the transmitted symbol, and $e_T(t) = \mathrm{rect}\left(\frac{t}{T_s}\right)$ is the rectangular basic waveform[1], the transmitted signal $s(t)$ can be rewritten as

$$s(t) = \sqrt{2}\sum_k \cos(2\pi f_0 t + \psi(k))\mathrm{rect}\left(\frac{t-kT_s}{T_s}\right). \tag{2.3}$$

The received signal, $g(t) = s(t) * h_c(t) + n(t)$, is converted into the equivalent lowpass domain in the first part of the receiver, the bandpass-to-lowpass (BP-LP) transform, to obtain the general complex-valued signal in the lowpass domain, $g_T(t)$. The operator $*$ denotes the convolution. The channel impulse response, $h_c(t)$, is assumed now to be time-invariant, and $n(t)$ is the additive white Gaussian noise. The mathematical expression for this BP-LP transform is

$$g_T(t) = \left(g(t)e^{-j2\pi f_0 t}\right) * 2h_{LP}(t) \tag{2.4}$$

where $h_{LP}(t)$ is the impulse response of the ideal lowpass filter. The channel impulse response in the equivalent lowpass domain, $h_{cT}(t)$, is introduced with the relation

$$h_c(t) = \mathrm{Re}\left\{h_{cT}(t)e^{j2\pi f_0 t}\right\}. \tag{2.5}$$

The following relation in the lowpass domain, equivalent to (2.4), can be obtained by

$$g_T(t) = \frac{1}{2}s_T(t) * h_{cT}(t) + n_T(t). \tag{2.6}$$

[1] $\mathrm{rect}(t) = \begin{cases} 0, & \text{if } |t| > 1/2 \\ \frac{1}{2}, & \text{if } |t| = \frac{1}{2} \\ 1, & \text{if } |t| < 1/2. \end{cases}$

The relation between $n(t)$ and $n_T(t)$ is similar as for $h_c(t)$ and $h_{cT}(t)$. By substituting (2.1) into (2.6), the following expression is obtained as

$$\begin{aligned} g_T(t) &= \frac{1}{2}\sum_{k=-\infty}^{\infty} x(k)\cdot e_T(t-kT_s) * h_{cT}(t) + n_T(t) \\ &= \frac{1}{2}\sum_{k=-\infty}^{\infty} x(k)\cdot h_T(t-kT_s) + n_T(t) \end{aligned} \tag{2.7}$$

where $h_T(t) \triangleq \frac{1}{2}e_T(t) * h_{cT}(t)$ can be considered as the impulse response of a transmission system, in the equivalent lowpass domain which includes also influence of the channel. It is referred to as *channel impulse response*.

The first term in (2.7) can be considered as *virtual* transmitted signal that is obtained in a transmission with linear modulation with $h_T(t)$ as basic waveform. A straightforward approach on the receiver side consists of using a correlation filter for $h_T(t)$ at the input of the receiver, a so-called channel match filter (CMF), $h_T^*(-t)$ where $(\cdot)^*$ is the complex conjugate operator. The CMF corresponds to the matched filter for the basic waveform, $e_T(t)$ in an AWGN channel. For AWGN channel $h_{cT}(t) = \delta_T(t)$, using matched filter of the basic waveform, $e_T^*(-t)$ followed by two-dimensional decision device is known to be optimum. However, in general $h_T(t)$ does not fulfill the first Nyquist criterion since $h_{cT}(t)$ differs from $\delta_T(t)$ and consequently the ISI appears. Therefore, in this case, the CMF followed by an appropriate equalizer is the suboptimal solution (see Fig. 2.2).

The sampling values $\tilde{x}_0(t)$ at the output of the CMF become

$$\begin{aligned} \tilde{x}_0(k) = y_T(t)|_{t=kT_s} &= \frac{1}{2}g_T(t) * h_T^*(-t)|_{t=kT_s} \\ &= \frac{1}{2}\left[\sum_{l=-\infty}^{\infty} x(l)h_T(t-l\cdot T_s) + n_T(t)\right] * h_T^*(-t)|_{t=kT_s} \\ &= \sum_{l=-\infty}^{\infty} x(l)\varphi_{h_Th_T}((k-l)\cdot T_s) + n_{Te}(k) \end{aligned} \tag{2.8}$$

where $\varphi_{h_Th_T}(\tau) = h_T^*(-t) * h_T(t)|_{t=\tau} = \int_{-\infty}^{\infty} h_T^*(t)h_T(t+\tau)dt$, is the auto-correlation function (ACF) of the channel impulse response $h_T(t)$, and $n_{Te}(k) = \frac{1}{2}n_T(t) * h_T^*(-t)|_{t=kT_s}$ is the filtered noise.

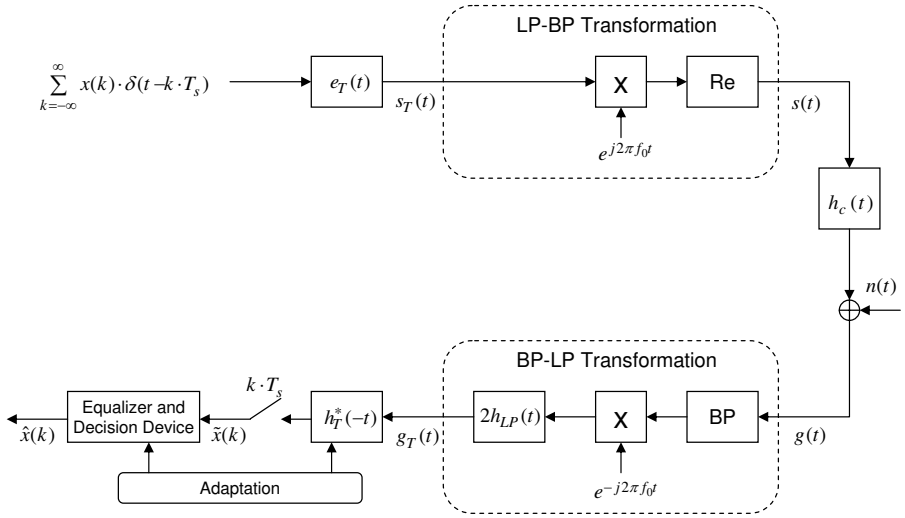

Figure 2.1: Digital transmission over a bandpass channel with AWGN.

According to the Wiener-Lee theorem it applies

$$\varphi_{h_T h_T}(t) = \frac{1}{2}\varphi_{e_T e_T}(t) * \varphi_{h_{cT} h_{cT}}(t). \tag{2.9}$$

For the ideal channel with $n_T(t) = 0$, we have

$$\varphi_{h_{cT} h_{cT}}(t) = K \cdot 2 \cdot \delta_T(t), \quad K \in \mathbb{R},\ K \neq 0. \tag{2.10}$$

Eq. (2.8) reduces to

$$\tilde{x}_0(k) = \sum_{l=-\infty}^{\infty} K \cdot x(l)\varphi_{e_T e_T}((k-l)T_s). \tag{2.11}$$

If $e_T(t)$ fulfills the first Nyquist criterion, that is

$$\varphi_{e_T e_T}(kT_s) = \begin{cases} 1, & \text{for } k = 0; \\ 0, & \text{otherwise.} \end{cases} \tag{2.12}$$

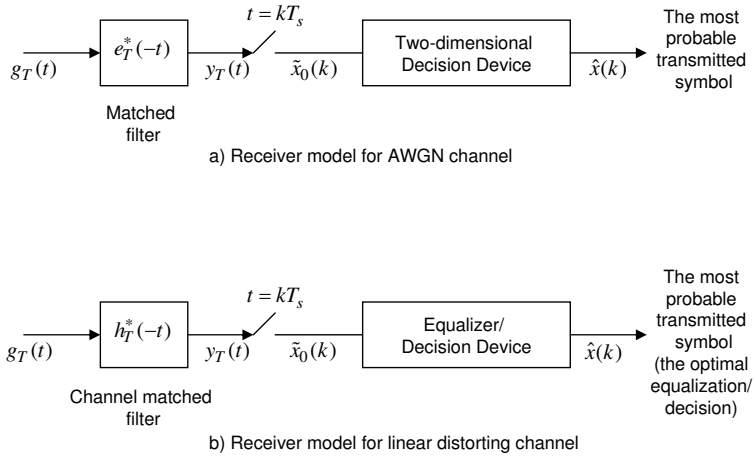

Figure 2.2: Receiver models for AWGN and linear distorting channels.

Then we would obtain

$$\tilde{x}_0(k) = K \cdot x(k). \tag{2.13}$$

By this way, a simple symbol-by-symbol decision would be possible. However, in general, the channel is not ideal. This means that (2.12) does not hold. This causes ISI in the received signal, in other words $\tilde{x}_0(k)$ depends not only on $x(k)$, but also the previous and later symbols.

Eq. (2.8) can be rewritten as

$$\tilde{x}_0(k) = \sum_{l=-\infty}^{\infty} \varphi_{h_T h_T}(lT_s) \cdot x(k-l) + n_{Te}(k). \tag{2.14}$$

By defining $r(k) \triangleq \varphi_{h_T h_T}(kT_s)$, (2.14) becomes

$$\begin{aligned} \tilde{x}_0(k) &= \sum_{l=-\infty}^{\infty} r(l) \cdot x(k-l) + n_{Te}(k) \\ &= r(k) * x(k) + n_{Te}(k) \end{aligned} \tag{2.15}$$

where $r(k)$ can be interpreted as the impulse response of a discrete-time filter. From (2.15) a simple method for digital transmission over a linearly distorting channel is obtained. In general, $r(k)$ has a finite length, and therefore (2.15)

can be rewritten as

$$\tilde{x}_0(k) = \underbrace{r(0)\cdot x(k)}_{\text{desired present symbol}} + \underbrace{\sum_{l=-L}^{-1} r(l)\cdot x(k-l)}_{\text{ISI from future symbols}} + \underbrace{\sum_{l=1}^{L} r(l)\cdot x(k-l)}_{\text{ISI from previous symbols}} + n_{Te}(k). \tag{2.16}$$

The subsequent equalizer aims to minimize (or eliminate if possible) ISI with suitable algorithms. To achieve this there exist different optimal (e.g., Viterbi algorithm or VA) and suboptimal (e.g., Decision Feedback Equalizer or DFE) solutions, see [8] for more details on equalizers.

2.1.2 Stochastic time-variant channel

This subsection considers a special case of the stochastic time-variant channel, known as Rayleigh fading channel. The model of such channel is given in Fig. 2.3. The complex additive white Gaussian noise, $n_T(t)$, is defined as in the previous section. The multiplicative noise, $a(t) = |a(t)|e^{j\varphi_a(t)}$, is a realization of a complex zero-mean Gaussian process (in the equivalent lowpass domain). The absolute value $|a(t)|$ has a Rayleigh probability distribution function. The multiplicative noise can be seen as a stochastic amplitude and phase modulation. The time-variant impulse response of the channel can be written as

$$h_{cT}(\tau, t) = \delta_T(\tau)\cdot a(t) = 2h_{LP}(\tau)\cdot a(t). \tag{2.17}$$

The time-variant transfer function is obtained by applying Fourier transformation w.r.t. τ

$$H_{cT}(f, t) = \int_{-\infty}^{\infty} h_{cT}(\tau, t)e^{j2\pi f\tau}d\tau = 2a(t). \tag{2.18}$$

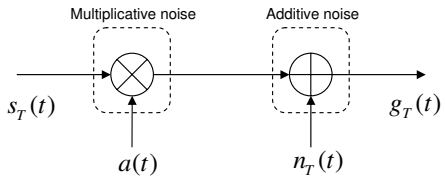

Figure 2.3: Channel model for the Rayleigh fading channel in the equivalent lowpass domain.

The received signal, $g_T(t)$, can be written as

$$\begin{aligned} g_T(t) &= \frac{1}{2} s_T(t) * h_{cT}(\tau, t) \\ &= \frac{1}{2} \int_{-\infty}^{\infty} h_{cT}(\tau, t) \cdot s_T(t-\tau) d\tau \\ &= s_T(t) \cdot a(t) = s_T(t) \cdot |a(t)| e^{j\varphi_a(t)} \\ &= \sum_k x(k) e_T(t - kT_s) \cdot |a(t)| e^{j\varphi_a(t)}. \end{aligned} \tag{2.19}$$

Typically, the time variation of $a(t)$ is slow compared to the symbol duration T_s. It can be assumed that $a(t)$ is constant for (at least) one symbol duration. With this assumption the received signal reduces to

$$\begin{aligned} g_T(t) &\approx \sum_k x(k) \cdot |a(k)| \cdot e^{j\varphi_a(k)} \cdot e_T(t - kT_s) \\ &\approx \sum_k x_a(k) \cdot e_T(t - kT_s) \end{aligned} \tag{2.20}$$

where $x_a(k) = x(k) \cdot |a(k)| \cdot e^{j\varphi_a(k)}$, $|a(k)| = |a(t = kT_s)|$, and $\varphi_a(k) = \varphi_a(t = kT_s)$.

The output of the matched filter is

$$y_T(t) \approx \sum_k x_a(k) \cdot \varphi_{e_T e_T}(t - kT_s). \tag{2.21}$$

After the sampling device, the received symbol $\tilde{x}_0(k)$ is

$$\tilde{x}_0(k) = y_T(t)|_{t=kT_s} \approx \sum_l x_a(l) \cdot \varphi_{e_T e_T}((k-l)T_s). \tag{2.22}$$

If the first Nyquist criterion is fulfilled, then

$$\tilde{x}_0(k) = E_e \cdot x_a(k) = E_e \cdot |a(k)| \cdot x(k) \cdot e^{j\varphi_a(k)}. \tag{2.23}$$

In Rayleigh fading channels, the transmitted symbol, $x(k)$, is corrupted not only by the additive noise, $n_T(t)$, but also by the multiplicative noise, $a(k)$.

2.1.3 Doppler frequency

The frequency shift caused by the relative motion between the transmitter and the receiver is known as *Doppler frequency*. The Doppler frequency depends on the speed of the mobile station and the angle of the received signal. The frequency shift will be maximum when the receiver moves directly towards or away from the transmitter.

The occurrence of the Doppler frequency can be described with a model of a real-valued signal with a single fixed carrier frequency f_0 of the transmitted signal. To simplify the analysis, it is assumed that there exists only a direct line-of-sight (LOS) between transmitter and receiver without multipath propagation. The bandpass transmitted signal may be defined as

$$s(t) = \mathrm{Re}\{s_T(t)e^{j2\pi f_0 t}\} \tag{2.24}$$

where $s_T(t) = \sum_n c_n g(t - nT)$ is the equivalent lowpass or baseband signal. The signal delay τ changes systematically with time due to the motion of the mobile station. A linear motion with time is assumed so that the distance $R(t)$ and hence the signal delay $\tau(t)$ between transmitter and receiver can be expressed as

$$R(t) = R_0 - v_r \cdot t \tag{2.25}$$

where R_0 is the initial distance between transmitter and receiver at time $t = 0$, which varies linearly and continuously with time. A linear motion model is sufficient for a short period consideration. The radial velocity is denoted by v_r. Because of the relative motion, the signal delay $\tau(t)$ changes

continuously

$$\tau(t) = \frac{R(t)}{c} = \frac{(R_0 - v_r \cdot t)}{c}. \tag{2.26}$$

Consequently, the received bandpass signal (noiseless) is

$$\begin{aligned} g(t) &= s(t - \tau(t)) \\ &= \mathrm{Re}\left\{ s_T(t - \tau(t)) e^{j2\pi f_0 (t - \tau(t))} \right\} \\ &= \mathrm{Re}\left\{ s_T(t - \tau(t)) e^{j2\pi f_0 \left(t - \frac{(R_0 - v_r \cdot t)}{c}\right)} \right\} \\ &= \mathrm{Re}\left\{ s_T(t - \tau(t)) e^{j\left(2\pi f_0 t - \frac{2\pi f_0 R_0}{c} + \frac{2\pi f_0 v_r \cdot t}{c}\right)} \right\} \\ &= \mathrm{Re}\left\{ s_T(t - \tau(t)) e^{j\left(2\pi (f_0 + f_0 \frac{v_r}{c}) t - \frac{2\pi f_0 R_0}{c}\right)} \right\}. \end{aligned} \tag{2.27}$$

It can be observed that the received signal frequency is increased as compared to the transmitted signal frequency if the mobile station moves towards transmitter and is decreased if the mobile station moves away from the transmitter.

The Doppler frequency is defined as the frequency difference between the transmitted and the received signals

$$f_D = f_0 - \left(f_0 + f_0 \frac{v_r}{c}\right) = -f_0 \frac{v_r}{c} = -f_0 \frac{v \cdot \cos(\alpha)}{c}. \tag{2.28}$$

The absolute velocity v of the mobile station and the angle α of the arrival signal determine the radial velocity as $v_r = v \cdot \cos(\alpha)$. The baseband received signal $g_T(t)$ can be derived directly from the above equation. The baseband transmitted signal $s_T(t)$ is delayed by $\tau(t)$ and in the case of moving mobile station, a phase shift is additionally observed in the received signal constellation diagram. The direction of rotation and the velocity are determined by the Doppler frequency f_D. The initial phase offset $\phi = \frac{2\pi f_0 R_0}{c}$ is dependent upon the distance R_0 between transmitter and receiver at time $t = 0$ and has no impact on further analysis. The baseband received signal can be expressed as

$$g_T(t) = s_T(t - \tau(t)) \cdot e^{-j(2\pi f_D t + \phi)}. \tag{2.29}$$

The received signal $g_T(t)$ is shifted in frequency with respect to (w.r.t.) the

transmitted signal, but its envelope remains constant

$$|g_T(t)| = |s_T(t - \tau(t))| = \text{const.} \tag{2.30}$$

Thus, the demodulation process in the receiver is disturbed by a frequency shift due to the relative motion, even such ideal conditions are assumed.

2.1.4 Effect of CFO

Assume that there exists LOS with no multipath propagation and perfect time synchronization[2], the received signal at the matched filter input is

$$g_T(t) = \sum_k x(k) e_T(t - kT_s) e^{j(\omega_D t + \phi)} + n_T(t) \tag{2.31}$$

where ϕ is the initial phase offset, and $f_D = \omega_D / 2\pi$ is CFO normalized to symbol rate $1/T_s$. Before sampling, the matched filter output is

$$\begin{aligned} y_T(t) &= g_T(t) * e_T^*(-t) = \int g_T(\lambda) e_T^*(\lambda - t) d\lambda \\ &= e^{j\phi} \sum_l x(l) \int e_T(\lambda - lT_s) e_T^*(\lambda - t) e^{j\omega_D \lambda} d\lambda + n_{Te}(t) \\ &= e^{j\phi} \sum_l x(l) e^{j\omega_D l T_s} \int e_T(\zeta) e_T^*(\zeta - (t - lT_s)) e^{j\omega_D \zeta} d\zeta + n_{Te}(t). \end{aligned} \tag{2.32}$$

Assuming Nyquist pulses, we have

$$\varphi_{e_T e_T}(t)|_{t=kT_s} = e_T(t) * e_T^*(-t)|_{t=kT_s} = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0. \end{cases} \tag{2.33}$$

[2]Indeed, excellent timing information can be normally derived even with frequency errors on the order of $10 - 20\%$ of the symbol rate.

where $n_{Te}(t) = n_T(t) * e_T^*(-t)$ is the filtered noise waveform. The signal sampled at the output of the matched filter is then

$$\begin{aligned}\tilde{x}_0(k) &= y_T(t = kT_s) \\ &= e^{j\phi} \sum_l x(l) e^{j\omega_D l T_s} \int e_T(\zeta) e_T^*(\zeta - (k-l)T_s) e^{j\omega_D \zeta} d\zeta + n_{Te}(k).\end{aligned} \tag{2.34}$$

The noise samples in (2.34) are independent complex Gaussian random variables. The typical assumption to simplify (2.34) is that f_D is small enough so that the complex exponential term inside the integral can be simply equated to one. Eq. (2.34), for $l = k$, becomes

$$\tilde{x}_0(k) = x(k) e^{j(\omega_D k T_s + \phi)} + n_{Te}(k) \tag{2.35}$$

which is an often used signal model for frequency estimation problem. For large frequency offsets, the complex exponential term inside the integral in (2.34) induces ISI. This can be shown by expanding $e^{j\omega_D \zeta}$ in Taylor series as

$$e^{j\omega_D \zeta} = \sum_{m=0}^{\infty} \frac{(j\omega_D \zeta)^m}{m!} = 1 + j\omega_D \zeta - \frac{1}{2}\omega_D^2 \zeta^2 + \dots \tag{2.36}$$

and by defining

$$p_m(\tau) = \int \zeta^m g(\zeta) g^*(\zeta - \tau) d\zeta. \tag{2.37}$$

Eq. (2.34) can then be rewritten as

$$\tilde{x}_0(k) = e^{j\phi} e^{j\omega_D k T_s} \sum_l x(k-l) h_l(\omega_D) + n_{Te}(k) \tag{2.38}$$

where

$$h_l(\omega_D) = e^{-j\omega_D l T_s} \sum_{m=0}^{\infty} \frac{(j\omega_D)^m}{m!} p_m(l T_s) \tag{2.39}$$

is the ISI coefficient. Eq. (2.38) and (2.39) show explicitly the frequency offset dependent ISI. Note that frequency offset causes ISI even if a Nyquist basic waveform is assumed.

2.2 Multiple-Carrier Transmission

This section is concerned with a particular type of multi-carrier modulation, known as orthogonal frequency multiplexing (OFDM). OFDM has gained popularity in a number of applications including digital audio/video broadcasting (DAB/DVB), high-speed digital subscriber line (DSL) modems or wireless local area network (WLAN). It is also used in the new generation mobile communication systems like Long Term Evolution (LTE) and Worldwide Inter-operability for Microwave Access (WiMAX). High data-rate is desired in many applications. However, as the symbol duration reduces with the increase of data-rate, the systems using single-carrier modulation suffer from more severe ISI caused by the dispersive fading of wireless channels, thereby needing more complex equalizers. OFDM modulation divides the entire frequency selective fading channel into many narrow band flat fading subchannels in which high-bit-rate data are transmitted in parallel and do not undergo ISI due to the long symbol duration. The subcarriers have the minimum frequency separation required to maintain orthogonality of their corresponding time domain waveforms, with the signal spectrum corresponding to the different subcarriers overlap in frequency. Hence the available bandwidth is used very efficiently. If knowledge of the channel is available at the transmitter, then the OFDM transmitter can adapt its signaling strategy to match the channel. Due to the fact that OFDM uses a large collection of narrowly spaced subchannels, these adaptive strategies can approach the ideal water pouring[3] capacity of a frequency-selective channel. In practice this is achieved by using adaptive bit loading techniques, where different sized signal constellations are transmitted on subcarriers.

Although OFDM has become widely used recently, the concept dates back some 40 years. Chang's work [9] published in 1966 shows that multi-carrier modulation can solve the multipath problem without reducing data rate. His work is generally considered as the first official publication on multi-carrier modulation. Some early work was Holsinger's 1964 MIT dissertation [10] and some of Gallager's early work on waterfalling [11]. In 1971, Weinstein and Ebert [12] show that multi-carrier modulation can be accomplished using DFT. Cimini at Bell Labs identifies many of the key issues in OFDM transmission and does proof-of-concept design [13].

[3]Water pouring or water filling is a power loading technique, which allocates more power on the subcarriers with high SNR and less power on the subcarrier with low SNR.

2.2.1 OFDM signal

In OFDM, data are transmitted blockwise. A sequence of complex data symbols is split into blocks and allocated to different subcarriers. Let $\{x_{i,k}\}_{k=0}^{N-1}$ be the complex symbols belonging to the i-th OFDM block, the i-th block of the OFDM signal can be expressed as

$$s_i(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_{i,k} e^{j2\pi f_k t} = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_{i,k} \varphi_k(t), \quad 0 \le t \le T_s \tag{2.40}$$

where $f_k = f_o + k\Delta f$, $\frac{1}{\sqrt{N}}$ is the normalizing factor, and

$$\varphi_k(t) = \begin{cases} e^{j2\pi f_k t} & \text{if } 0 \le t \le T_s \\ 0 & \text{otherwise} \end{cases} \tag{2.41}$$

for $k = 0, 1, \cdots, N-1$. T_s and Δf are the symbol duration and subcarrier spacing of OFDM, respectively. For simplicity, the index i can be omitted. Despite the spectrum of the OFDM subcarriers overlap, they do not interfere after demodulated since they are orthogonal with each other, so that x_k can be extracted independently of each other in the receiver. The orthogonality condition yields

$$\begin{aligned} \langle \varphi_k(t), \varphi_l(t) \rangle &= \frac{1}{T_s} \int_0^{T_s} \varphi_k(t) \varphi_l^*(t) dt \\ &= \frac{1}{T_s} \int_0^{T_s} e^{j2\pi(f_k - f_l)t} dt = \frac{1}{T_s} \int_0^{T_s} e^{j2\pi(k-f)\Delta f t} dt \\ &= \frac{1}{j2\pi(k-l)\Delta f T_s} \left[e^{j2\pi(k-l)\Delta f t} \right]_{t=0}^{T_s} \\ &= \frac{1}{j2\pi(k-l)} \left[e^{j2\pi(k-l)} - 1 \right] \\ &= \frac{\sin[\pi(k-l)]}{\pi(k-l)} e^{j\pi(k-l)} = \delta(k-l) \end{aligned} \tag{2.42}$$

where $\delta(\cdot)$ is the discrete delta function.

Eq. (2.42) shows that $\{\varphi_k(t)\}_{k=0}^{N-1}$ is a set of orthogonal functions. Using

this property, the OFDM signal can be demodulated by

$$\begin{aligned}\sqrt{N}\langle s(t), \varphi_k(t)\rangle &= \frac{\sqrt{N}}{T_s}\int_0^{T_s} s(t)\varphi_k^*(t)dt \\ &= \frac{\sqrt{N}}{T_s}\int_0^{T_s}\left(\frac{1}{\sqrt{N}}\sum_{l=0}^{N-1} x_l\varphi_l(t)\right)\varphi_k^*(t)dt \\ &= \sum_{l=0}^{N-1} x_l\delta(l-k) = x_k.\end{aligned} \tag{2.43}$$

By dividing the total bandwidth into narrow subbands, the symbol duration is now N times larger than the original symbol duration. Therefore, the condition that $T_s \gg \tau_{max}$ is fulfilled as in a narrowband channel. The ISI is thus reduced considerably.

2.2.2 FFT implementation

From (2.43), an integral is used for demodulation of OFDM signals. Here we describe the relationship between OFDM and DFT, which can be efficiently implemented by low complexity fast Fourier transform (FFT). Recall the OFDM signal model

$$s(t) = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} x_k e^{j2\pi f_k t}. \tag{2.44}$$

The sampling space of an OFDM signal of bandwidth B is

$$\Delta t = \frac{1}{B} = \frac{1}{N\Delta f}. \tag{2.45}$$

The discrete-time transmit signal, $s_n \triangleq s(n\Delta t)$ for $0 \le n \le N-1$, is

$$s_n = \frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} x_k e^{j2\pi n\Delta t k\Delta f} \tag{2.46}$$

which is known as the DFT of the sequence x_k if $\Delta t \Delta f = \frac{1}{N}$. Hence, the IFFT of the data block is

$$s_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k \exp\left\{ j \frac{2\pi nk}{N} \right\}, \quad n = 0, 1, \cdots, N-1 \tag{2.47}$$

yielding the time-domain sequence s_n. For the same reason, the receiver can be also implemented using FFT.

2.2.3 Cyclic extension

To mitigate the effects of ISI caused by channel delay spread, each block of N-IFFT coefficients is typically preceded by a cyclic prefix (CP) or a guard interval consisting of N_g samples, such that the length of the CP is at least equal to the channel length N_h in samples, where $\mu = \frac{T_h}{T_s} N$, T_h is the length of (continuous) channel, and T_s is the duration of a OFDM block. The cyclic prefix is simply a repetition of the last N_g IFFT coefficients. Alternatively, a cyclic suffix can be appended to the end of the block of N IFFT coefficients, that is a repetition of the first N_g IFFT coefficients. The guard interval of length N_g is an overhead that results in a power and bandwidth penalty, since it consists of redundant symbols. However, the guard interval is useful for implementing time and frequency synchronization functions in the receiver, since it contains repeated symbols at a known sample spacing. The time duration of an OFDM symbol is $N + N_g$ times larger than the modulated symbol in a single-carrier system.

2.2.4 Generalized representation

Let $x_{i,k} = x_k(i)$ represent the complex data symbols belonging to the i-th OFDM symbol block. Eq. (2.47) can be rewritten as

$$s_n(i) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N_u-1} x_k(i) \exp\left\{ j \frac{2\pi nk}{N} \right\}, \quad n = 0, 1, \cdots, N-1 \tag{2.48}$$

where $N_u \leq N$ are now the used subcarriers, N is the IFFT size, and $x_k(i)$ is the k-th subcarrier of the i-th OFDM symbol. If $N_u < N$, the residual $N - N_u$ subcarriers referred to as null or virtual subcarriers are filled with

zeros. The column vectors of $s_n(i)$ and $x_n(i)$ are defined as

$$\begin{aligned} \mathbf{x}(i) &= [x_0(i),\ x_1(i),\ \cdots,\ x_{N_u-1}(i)]^T \\ \mathbf{s}(i) &= [s_0(i),\ s_1(i),\ \cdots,\ s_{N_u-1}(i),\ \cdots\ s_{N-1}(i)]^T \end{aligned} \tag{2.49}$$

and the IFFT matrix is

$$\mathbf{W} = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{j\frac{2\pi}{N}} & \cdots & e^{j\frac{2\pi}{N}(N_u-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{j\frac{2\pi}{N}(N-1)} & \cdots & e^{j\frac{2\pi}{N}(N-1)(N_u-1)} \end{pmatrix}_{N\times N_u} \tag{2.50}$$

the i-th OFDM block can be written in a more compact form as

$$\mathbf{s}(i) = \mathbf{W}\mathbf{x}(i). \tag{2.51}$$

A cyclic prefix (CP) is preceded after the IFFT modulation and its length N_g is assumed to be longer than the maximum delay spread of the channel to completely avoid the ISI. This can be expressed as

$$\mathbf{s}_{\text{cp}}(i) = [\underbrace{s_{N-N_g}(i), \cdots, s_{N-1}(i)}_{\text{cyclic prefix}}, s_0(i), \cdots, s_{N-1}(i)]^T \tag{2.52}$$

this operation can be realized by a matrix-vector multiplication

$$\mathbf{s}_{\text{cp}}(i) = \mathbf{G}\mathbf{s}(i) \tag{2.53}$$

where $\mathbf{G}$ has the form of

$$\mathbf{G} = \begin{pmatrix} \mathbf{0}_{N_g\times(N-N_g)} & \mathbf{I}_{N_g\times N_g} \\ \multicolumn{2}{c}{\mathbf{I}_{N\times N}} \end{pmatrix} \tag{2.54}$$

where $\mathbf{0}$ denotes a matrix with all zero entries, and $\mathbf{I}$ denotes the identity matrix. After parallel to serial conversion, the Dirac-sampled time-domain transmit symbols are filtered by he transmit filter $h_T(t)$, which is a bandlim-

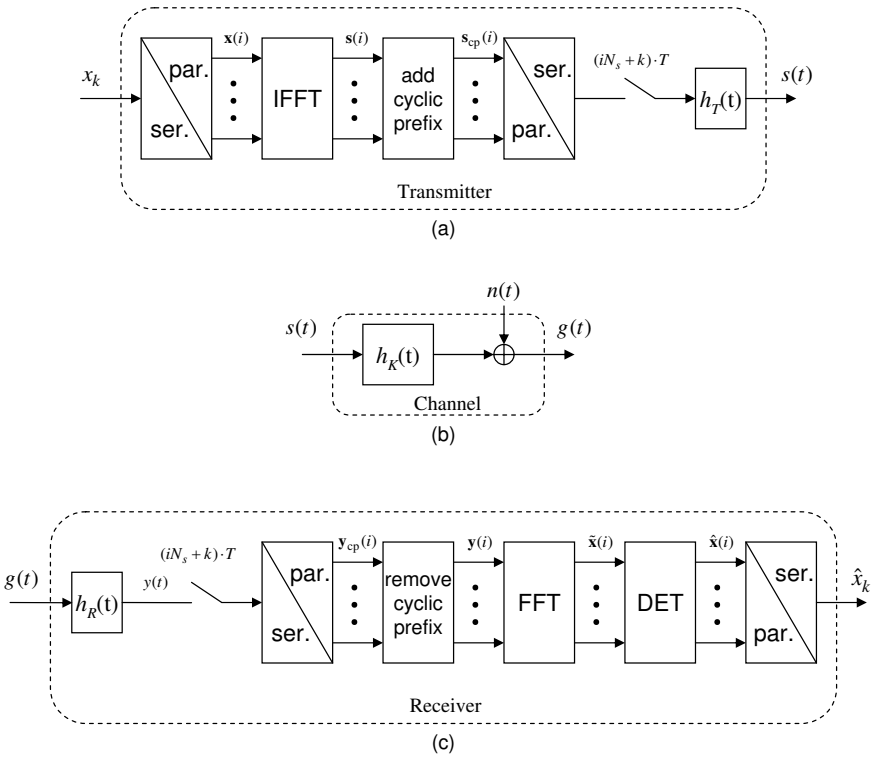

Figure 2.4: Equivalent lowpass domain OFDM transmission (SISO case).

ited low-pass filter with cutoff frequency of $f_g = \frac{1}{2\Delta t}$. $\Delta t = \frac{T_s}{N+N_g}$, which denotes also the single symbol duration T. The resulting signal $s(t)$ can be written as

$$s(t) = \sum_{i=0}^{\infty} \sum_{k=0}^{N_s-1} s_k(i) h_T(t - (iN_s + k)T) \tag{2.55}$$

$s(t)$ is then up-converted and transmitted over the time-invariant frequency-selective channel with impulse response $h_K(t)$. In Fig. 2.4, $h_K(t)$ represents the impulse response of the lowpass equivalence of the physical channel which is generally complex-valued. The output of the physical channel is corrupted by additive white Gaussian noise (AWGN) with two side power spectral density (PSD), $\frac{N_0}{2}$. For high frequency radio channel, bandpass filtering is applied at the receiver before down-conversion. The noise is therefore bandpass white Gaussian. Letting $n(t)$ denote the equivalent lowpass complex-valued

noise with a PSD [8] of

$$\Phi_{nn}(f) = \begin{cases} N_0, & \text{for } |f| \leq \frac{1}{2}B \\ 0, & \text{for } |f| > \frac{1}{2}B \end{cases} \tag{2.56}$$

and its autocorrelation function is

$$\phi_{nn}(\tau) = N_0 \frac{\sin(\pi B\tau)}{\pi\tau} = N_0 B \frac{\sin(\pi B\tau)}{\pi B\tau} = N_0 B \text{sinc}(\pi B\tau) \tag{2.57}$$

where B indicates the bandwidth of the bandpass filter, and $\text{sinc}(x) \triangleq \frac{\sin(x)}{x}$. The limiting form of $\phi_{nn}(\tau)$ as B approaches infinity is $\phi_{nn}(\tau) = N_0\delta(\tau)$. The variance of $n(t)$ is defined as $\sigma_n^2 = \phi_{nn}(0) = N_0 B$.

Assuming that the receiver is perfectly synchronized with the transmitter, the down-converted received signal $g(t)$ at the input of the receiver can be expressed as

$$\begin{aligned} g(t) &= s(t) * h_K(t) + n(t) \\ &= \int_{-\infty}^{\infty} h_K(\tau) s(t-\tau) d\tau + n(t) \end{aligned} \tag{2.58}$$

which is then filtered by the receiver filter $h_R(t)$ and sampled at $t = (iN_s + k)T$. It is assumed that $h_R(t)$ is an ideal lowpass filter having cutoff frequency $f_g = \frac{1}{2T}$. The discrete-time equivalent lowpass channel impulse response from the concatenation of $h_T(t)$, $h_K(t)$, and $h_R(t)$ is

$$h_l \triangleq h(l \cdot T) = h_T(t) * h_K(t) * h_R(t)|_{t=l\cdot T} \tag{2.59}$$

the time index i is ignored since the channel is assumed to be time-invariant. We denote the channel with a vector $\mathbf{h} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{L-1} \end{bmatrix}^T$. If $L > 1$, the channel is time dispersive and thus frequency selective. Consequently, the received time-domain discrete-time symbol $y_k(i)$ is obtained as

$$y_k(i) = \sum_{l=0}^{L-1} h_l s_{k,l}(i) + n'_k(i) \tag{2.60}$$

where $n'_k(i) = h_R(t) * n(t)|_{t=(iN_s+k)T}$. Assume that $B \geq \frac{1}{2T}$. Since $h_R(t)$ is an ideal lowpass filter, $n'_k(i)$ is therefore a complex-valued white Gaussian random variable with variance $\sigma_n^2 = N_0 B$. The received symbol vector after serial to parallel conversion is denoted by $\mathbf{y}(i) = \left[\begin{array}{cccc} y_0(i) & y_1(i) & \cdots & y_{N+N_g-1}(i) \end{array} \right]^T$. Eq. (2.60) can also be formulated as a matrix-vector convolution

$$\mathbf{y}(i) = \sum_{j=-\infty}^{\infty} \mathbf{H}(j)\mathbf{s}(i-j) + \mathbf{n}'(i) \tag{2.61}$$

where $\mathbf{H}(i)$ is $(N+N_g) \times (N+N_g)$ matrix. If the length of $\mathbf{h}$ is not greater than $N+1$, the sequence of $\mathbf{H}(i)$ matrices consist of nonzero matrices for $i=0$ and $i=1$:

$$\mathbf{H}(0) = \begin{pmatrix} h_0 & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & h_0 & & \vdots & & \\ h_{L-1} & \vdots & \ddots & 0 & & \vdots \\ 0 & h_{L-1} & & h_0 & & \\ \vdots & & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h_{L-1} & \cdots & h_0 \end{pmatrix} \tag{2.62}$$

and

$$\mathbf{H}(1) = \begin{pmatrix} 0 & \cdots & 0 & h_{L-1} & \cdots & h_1 \\ \vdots & & \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & \vdots & & h_{L-1} \\ 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{2.63}$$

At the receiver, after the removal of the cyclic prefix for $\mathbf{y}(i)$ and FFT demodulation, the frequency-domain received symbol vector $\tilde{\mathbf{x}}(i)$ is given by

$$\tilde{\mathbf{x}}(i) = \mathbf{W}^H \mathbf{G}_r \mathbf{y}(i) + \mathbf{n}(i) \tag{2.64}$$

with

$$\mathbf{G}_r = \left(\begin{array}{cc} \mathbf{0}_{N \times N_g} & \mathbf{I}_{N \times N} \end{array} \right) \tag{2.65}$$

and $\mathbf{n}(i) = \mathbf{W}^H \mathbf{G}_r \mathbf{n}'(i)$, where $(\cdot)^H$ is the conjugate transpose (or Hermitian

transpose) operator. The noise sample $n_k(i) \in \mathbf{n}(i)$ are also complex-valued white Gaussian random variables with variance σ_n^2. $\mathbf{W}^H$ denotes the normalized Fourier matrix with $W_{mn}^H = \frac{1}{\sqrt{N}} e^{-j\frac{2\pi}{N}(m-1)(n-1)}$ and $\mathbf{W}^H\mathbf{W} = \mathbf{W}^{-1}\mathbf{W} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. Substituting (2.51), (2.53) and (2.60) in (2.64) yields

$$\tilde{\mathbf{x}}(i) = \mathbf{W}^H\mathbf{G}_r\mathbf{H}(i)\mathbf{G}\mathbf{W} * \mathbf{x}(i) + \mathbf{n}(i). \tag{2.66}$$

If L is not greater than $N_g - 1$, then $\mathbf{G}_r\mathbf{H}(1)\mathbf{G}$ yields a zero matrix and

$$\mathbf{G}_r\mathbf{H}(0)\mathbf{G} = \begin{pmatrix} h_0 & 0 & \cdots & 0 & h_{L-1} & \cdots & h_1 \\ h_1 & h_0 & & \vdots & & \ddots & \vdots \\ \vdots & h_1 & \ddots & 0 & & & h_{L-1} \\ h_{L-1} & \vdots & \ddots & h_0 & \ddots & & 0 \\ 0 & h_{L-1} & & h_1 & \ddots & 0 & \vdots \\ \vdots & & \ddots & \vdots & & h_0 & 0 \\ 0 & \cdots & 0 & h_{L-1} & \cdots & h_1 & h_0 \end{pmatrix} \tag{2.67}$$

is circular. According to the properties of a circular matrix, every circular matrix can be diagonalized by the Fourier matrix $\mathbf{W}$

$$\mathbf{G}_r\mathbf{H}(0)\mathbf{G} = \mathbf{W}\boldsymbol{\mathcal{H}}\mathbf{W}^H \tag{2.68}$$

where

$$\begin{aligned} \boldsymbol{\mathcal{H}} &= \sqrt{N} \cdot \operatorname{diag}\left\{\mathbf{W} \cdot [h_0, h_1, \cdots, h_{L-1}, \mathbf{0}_{1\times N-L-1}]^T\right\} \\ &= \operatorname{diag}\left\{[H_0, H_1, \cdots, H_{N-1}]\right\} \end{aligned} \tag{2.69}$$

is a diagonal matrix. Clearly, the elements on the diagonal of $\boldsymbol{\mathcal{H}}$ are the discrete channel transfer functions over N subcarriers. Hence,

$$\mathbf{W}^H\mathbf{G}_r\mathbf{H}(i)\mathbf{G}\mathbf{W} = \boldsymbol{\mathcal{H}}. \tag{2.70}$$

Accordingly, (2.66) reduces to

$$\tilde{\mathbf{x}}(i) = \mathbf{W}^H\mathbf{G}_r\mathbf{H}(i)\mathbf{G}\mathbf{W} \cdot \delta(i) * \mathbf{x}(i) + \mathbf{n}(i) = \boldsymbol{\mathcal{H}}\mathbf{x}(i) + \mathbf{n}(i). \tag{2.71}$$

The diagonal channel transfer function matrix implies that each subcarrier undergoes frequency flat-fading. Data detection can be therefore, realized by a bank of adaptive one-tap equalizers to combat the phase and amplitude distortions resulting from fading. In such a case, the equalization matrix $\boldsymbol{\mathcal{G}}$ is also a diagonal matrix with the weighting factors of these equalizers on the main diagonal. This can be seen in the case of the simple zero forcing (ZF) equalizer. The symbol-by-symbol ZF equalizer applies the inverted channel coefficients

$$g_m = \frac{H_m^*}{|H_m|^2}. \tag{2.72}$$

The ZF equalizer totally recovers the desired symbols, excepted for the case where $|H_m| = 0$. The covariance matrix of the equalized noise $\tilde{\mathbf{n}}$ has the form

$$\mathbf{C}_{\tilde{\mathbf{n}}\tilde{\mathbf{n}}} = \sigma_n^2 \cdot \begin{pmatrix} \frac{1}{|H_0|^2} & & & \mathbf{0} \\ & \frac{1}{|H_1|^2} & & \\ & & \ddots & \\ \mathbf{0} & & & \frac{1}{|H_{N-1}|^2} \end{pmatrix} \tag{2.73}$$

and therefore the output SNR of the ZF is given by

$$\mathrm{SNR} = |H_m|^2 \frac{\sigma_x^2}{\sigma_n^2}. \tag{2.74}$$

Since $\boldsymbol{\mathcal{H}}$ is a diagonal matrix, no noise enhancement occurs in this case. However, it can be observed that the SNR is subcarrier dependent, provided that the channel is frequency selective. Note that the energy loss due to the insertion of the cyclic prefix has not been included in the SNR. As this energy loss is a constant factor and will be reflected as horizontal shift of the BER curve.

2.2.5 Effect of CFO

The CFO can be several times larger than the subcarrier spacing. It is usually divided into an integer part and a fractional part. If the CFO is an integer n multiple of subcarrier spacing Δf, then the received frequency domain subcarriers are shifted by n subcarrier positions. The subcarriers are still mutually orthogonal but the received data symbols, which were mapped

to the OFDM spectrum are now in the wrong position in the demodulated spectrum, resulting in a BER of 0.5. If the CFO is a fraction of the subcarrier spacing, then energy is spilling over between the subcarriers, resulting in loss in their mutual orthogonality. The ICI is then observed which deteriorates the BER of the system. The impact of ICI on the system performance of an OFDM system can be evaluated by investigating the spectrum of the OFDM symbol (see [14]).

For simplicity, the analysis of the effect of CFO will be carried out in the absence of the multipath fading and the additive noise. Consider the transmitted OFDM signal

$$s_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k e^{j\frac{2\pi nk}{N}}, \quad n = 0, 1, \ldots, N-1.$$

where x_k is the transmitted symbol over the k-th subcarrier. If there is a multiplicative time-varying distortion, $\gamma_n \triangleq \gamma(nT) = e^{j\frac{2\pi \upsilon nT}{N}}$, that is caused by frequency offset, the noiseless received signal is

$$r_n = \underbrace{e^{j\frac{2\pi\epsilon n}{N}}}_{\gamma_n} \cdot s_n, \quad \text{with } \epsilon \triangleq \upsilon T \tag{2.75}$$

where a CFO ϵ is a fractional of the subcarrier spacing.

Taking DFT to obtain the OFDM symbol on the m-th subcarrier, gives

$$\begin{aligned}
\tilde{x}_m &= \sum_{k=0}^{N-1} r_k \cdot \frac{1}{\sqrt{N}} e^{-j\frac{2\pi km}{N}} \\
&= \frac{1}{N} \sum_{k=0}^{N-1} e^{j\frac{2\pi\epsilon k}{N}} \sum_{l=0}^{N-1} x_l e^{j\frac{2\pi kl}{N}} e^{-j\frac{2\pi km}{N}} \\
&= \frac{e^{j\pi\epsilon\left(\frac{N-1}{N}\right)}}{N} \frac{\sin(\pi\epsilon)}{\sin\left(\frac{\pi\epsilon}{N}\right)} \cdot x_m + \underbrace{\frac{1}{N} \sum_{l=0,\ l\neq m}^{N-1} x_l \sum_{k=0}^{N-1} e^{j\frac{2\pi k(l-m+\epsilon)}{N}}}_{ICI} \\
&= I_0(\epsilon) \cdot x_m + \underbrace{\sum_{l=0,\ l\neq m}^{N-1} x_l I_{l-m}(\epsilon)}_{ICI}
\end{aligned} \tag{2.76}$$

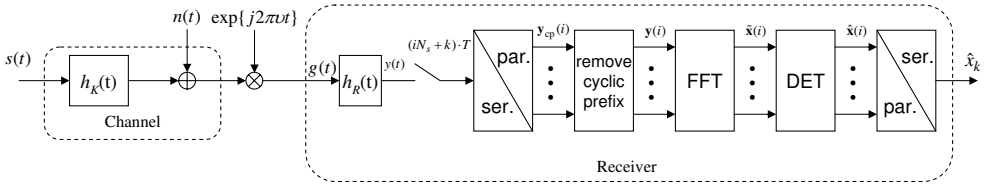

Figure 2.5: Equivalent lowpass domain OFDM reception with CFO.

where

$$I_p(\epsilon) \triangleq \frac{1}{N}\sum_{k=0}^{N-1} e^{j\frac{2\pi k(\epsilon+p)}{N}} = e^{j\left(\pi\left(\frac{N-1}{N}\right)(\epsilon+p)\right)}\frac{\sin\left(\pi(\epsilon+p)\right)}{N\cdot\sin\left(\frac{\pi}{N}(\epsilon+p)\right)}$$

where $p \triangleq l-m$. It can be seen that a fractional CFO ($|\epsilon| < 0.5$) causes a reduction in signal amplitude and ICI.

2.3 Receive Diversity

Diversity combining devotes the entire resources of the array to service a single user (see Fig. 2.6). Specifically, diversity schemes enhance reliability by minimizing the channel fluctuations due to fading. The main idea in diversity[4] is that different antennas receive different versions of the same signal. The probability that all these copies being in a deep fad is small. These schemes therefore make most sense when the fading is independent from element to element and are limited use (beyond increasing SNR) if perfectly correlated (such as in LOS conditions). Independent fading would arise in a dense urban environment where several multipath components add up very differently at each element. Diversity combining is specifically targeted to counteract small scale fading. It is therefore suitable for the assumptions that have been made in this thesis that the received signal experiences Rayleigh slow flat-fading. The physical model assumes the fading to be independent from one element to the next. Each element, therefore, acts as an independent sample of the random fading process (e.g., Rayleigh), i.e., each element of the array receives an independent copy of the transmitted signal. The main of receive antenna diversity is to combine these independent samples

[4]Diversity arises in various forms - space, angle, frequency, time, polarization diversity.

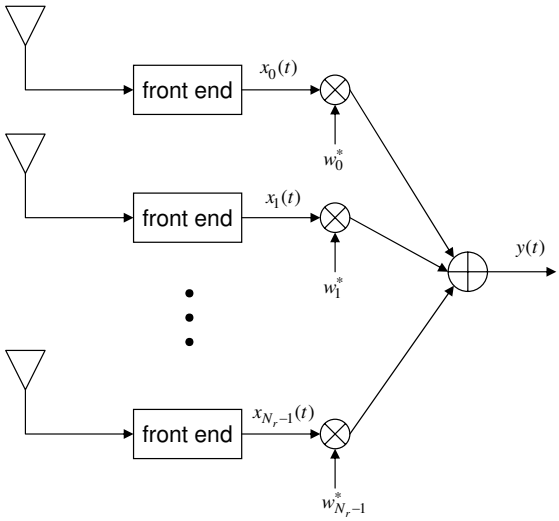

Figure 2.6: Receive antenna diversity.

to achieve the desired goal of increasing the SNR and reducing the BER. Diversity works because given N elements in the receiving antenna array we receive N independent copies of the same signal. It is assumed here that the received signal copies must be uncorrelated[5] or weakly-correlated (i.e., correlation coefficient $\rho \leq 0.5$). The models developed in the following subsections are for single user.

2.3.1 Signal model

Consider a single-user system wherein the received signal is a sum of the desired signal and noise

$$\mathbf{x}(t) = \mathbf{h}(t)u(t) + \mathbf{n}(t) \tag{2.77}$$

where $u(t)$ is the unit power signal transmitted, $\mathbf{h}(t)$ represents the channel (including the signal power) and $\mathbf{n}(t)$ is the noise vector. The power in the

[5]For Gaussian fading (Rayleigh fading is complex Gaussian) uncorrelated fading implies independent fading.

signal over a single symbol period, T_s, at element n, is

$$P_n = \frac{1}{T_s}\int_0^{T_s} |h_n(t)|^2|u(t)|^2dt = |h_n(t)|^2\frac{1}{T_s}\int_0^{T_s}|u(t)|^2dt \simeq |h_n(t)|^2 \quad (2.78)$$

where, if the channel is assumed to be slow fading, the term $|h_n(t)|^2$ remains constant over a symbol period ($h_n(t) \simeq h_n$) and can be brought out of the integral and $u(t)$ is assumed to have unit power. Setting $E\{|n_n(t)|^2\} = \sigma^2$, the instantaneous SNR is obtained as

$$\gamma_n = \frac{|h_n|^2}{\sigma^2}. \quad (2.79)$$

The instantaneous SNR is a random variable with a specific realization given the channel realization h_n.

For Rayleigh fading, $h_n = |h_n|e^{j\measuredangle h_n}$, where $\measuredangle h_n$ is uniform in $[0, 2\pi)$ and $|h_n|$ has a Rayleigh pdf, implying $|h_n(t)|^2$ (and γ_n) has an exponential pdf

$$\begin{aligned} |h_n| &\sim \frac{2|h_n|}{P_0}e^{|h_n|^2/P_0} \\ \gamma_n &\sim \frac{1}{\Gamma}e^{-\gamma_n/\Gamma} \\ \Gamma &= E\{\gamma_n\} = \frac{E\{|h_n|^2\}}{\sigma^2} = \frac{P_0}{\sigma^2}. \end{aligned} \quad (2.80)$$

The instantaneous SNR of each antenna element is an exponentially distributed random variable. Γ represents the average SNR at each element. This is also the SNR of a single antenna, i.e., the SNR if there were no array, Γ will therefore serve as a baseline for the improvement in SNR.

2.3.2 Selection combining

Selection combining (SC) is the simplest diversity technique. It is easy to implement because all that is needed is a side monitoring station and an antenna switch at the receiver. However, it is not an optimal diversity technique since it does not use all of the possible branches simultaneously. The selection combining selects the signal from the element that has greatest SNR

for further processing. The weights in SC is then

$$w_k = \begin{cases} 1, & \gamma_k = \max_n\{\gamma_n\} \\ 0, & \text{otherwise} \end{cases} \tag{2.81}$$

Since the element chosen is the one with the maximum SNR, the output SNR of the selection diversity scheme is $\gamma = \max_n\{\gamma_n\}$. Such scheme would need only a measurement of signal power, phases and variable gains are not required. The average output SNR is obtained as (see also [15] for the analysis of $\mathrm{E}\{\gamma\}$ and the corresponding BER)

$$\mathrm{E}\{\gamma\} \approx \Gamma\left(C + \ln N_r + \frac{1}{2N_r}\right) \tag{2.82}$$

where the final approximation is valid for relatively large value of receive antennas, N_r, and C is Euler's constant. The improvement in SNR over that of a single element is of order of $\ln N_r$.

2.3.3 Maximum ratio combining

Maximum ratio combining (MRC) uses each of the N_r branches in a co-phased and weighted manner such that the highest achievable SNR is available in the receiver at all time. MRC obtains the weighting vector $\mathbf{w}$ that maximizes the SNR of the output signal $y(t)$, i.e., MRC uses linear coherent combining of branch signals with such weights so that the ouput SNR is maximized.

Writing a snapshot of the received signal at the array elements as a vector $\mathbf{x}(t)$, and the output of MRC as a scalar $y(t)$ which may be expressed as

$$\begin{aligned} y(t) &= \mathbf{w}^H\mathbf{x}(t) \\ &= \mathbf{w}^H\mathbf{h}(t)s(t) + \mathbf{w}^H\mathbf{n}(t) \end{aligned} \tag{2.83}$$

where $\mathbf{h}(t) = [h_0(t), h_1(t), \cdots, h_{Nr-1}(t)]^T$ and $\mathbf{n}(t) = [n_0(t), n_1(t), \cdots, n_{Nr-1}(t)]^T$.

Since the signal $s(t)$ has unit average power, the instantaneous output SNR is

$$\gamma = \frac{|\mathbf{w}^H\mathbf{h}|^2}{\mathrm{E}\{|\mathbf{w}^H\mathbf{n}|^2\}}. \tag{2.84}$$

The noise power in the denominator is given by

$$\begin{aligned} P_n &= \mathrm{E}\{|\mathbf{w}^H\mathbf{n}|^2\} = \mathrm{E}\{|\mathbf{w}^H\mathbf{n}\mathbf{n}^H\mathbf{w}|\} = \mathbf{w}^H\mathrm{E}\{\mathbf{n}\mathbf{n}^H\}\mathbf{w} = \sigma_n^2\mathbf{w}^H\mathbf{I}\mathbf{w} \\ &= \sigma_n^2\mathbf{w}^H\mathbf{w} = \sigma_n^2\|\mathbf{w}\|^2. \end{aligned} \tag{2.85}$$

Since constants do not matter, one could scale $\mathbf{w}$ such that $\|\mathbf{w}\| = 1$. The SNR is therefore given by $\gamma = |\mathbf{w}^H\mathbf{h}|^2/\sigma_n^2$. By using Cauchy-Schwarz inequality, γ is maximized if $\mathbf{w}$ is linearly proportional to $\mathbf{h}$, i.e.,

$$\begin{aligned} &\mathbf{w} \triangleq \mathbf{h}, \\ \Rightarrow \gamma &= \frac{|\mathbf{h}^H\mathbf{h}|^2}{\sigma_n^2\mathbf{h}^H\mathbf{h}} = \frac{\mathbf{h}^H\mathbf{h}}{\sigma_n^2} \\ &= \sum_{n=0}^{N_r-1}\gamma_n. \end{aligned} \tag{2.86}$$

The best that a diversity combiner can do, is to choose the weights to be equal to the fading of each element. In some sense this answer is expected since the solution is effectively the matched filter for fading signal. We know that the matched filter is optimal in the single user case.

The expected value of the output SNR is therefore N times the average SNR at each element, i.e.,

$$\mathrm{E}\{\gamma\} = N_r\Gamma \tag{2.87}$$

which indicates that on average, the SNR improves by a factor of N_r. This is significantly better than the factor of $\ln N_r$ improvement in the case of selection diversity case. The BER improvement analysis for MRC can be found in [15].

Note that in order to perform MRC, the receiver has to know the fading or has to have access to the channel state information (CSI). This is usually achieved by sending known symbols (pilot) through the channel and measuring the channel's response. Clearly, such a procedure does not allow for having perfect CSI, but rather approximate CSI which results in suboptimal solutions, in practice.

2.3.4 Equal gain combining

In certain cases, it is not convenient to provide for the variable weighting capability required for true maximum ratio combining. In such cases, the branch weights are all set to unity but the signals from all branches are co-phased to provide equal gain combing (EGC) diversity. This allows the receiver to exploit signals that are simultaneously received on each branch. The possibility of producing an acceptable signal from a number of unacceptable inputs is still retained, and performance is only marginally inferior to MRC and superior to SC.

In equal gain combining, each signal branch weighted with the same factor, irrespective of the signal amplitude

$$\begin{aligned} w_n &= \exp\{j\angle h_n\} \\ \Rightarrow w_n^* h_n &= |h_n| \\ \Rightarrow \mathbf{w}^H \mathbf{h} &= \sum_{n=0}^{N_r-1} |h_n|. \end{aligned} \tag{2.88}$$

The noise and instantaneous SNR are given by

$$\begin{aligned} P_n &= \mathbf{w}^H \mathbf{w} \sigma_n^2 = N_r \sigma_n^2 \\ \gamma &= \frac{\left(\sum_{n=0}^{N_r-1} |h_n|\right)^2}{N_r \sigma_n^2}. \end{aligned} \tag{2.89}$$

The average output SNR is obtained as (see also [15] for the analysis of $\mathrm{E}\{\gamma\}$ and the corresponding BER)

$$\mathrm{E}\{\gamma\} = \Gamma\left[1 + (N_r - 1)\frac{\pi}{4}\right]. \tag{2.90}$$

Despite being much simpler to implement than MRC, the EGC results in an improvement in SNR that is comparable to that of the optimal MRC. The SNR of both combiners increases linearly with N_r.

2.4 Summary

In this chapter, the basic principle of single-carrier transmission is introduced. Doppler shift caused by relative motion, the effect of carrier frequency offset caused by the non-synchronized oscillators between transmitter and receiver, and fading-induced multiplicative noise are described. The analysis of CFO effect shows that for large CFO errors, the inter-symbol-interference (ISI) is induced. As a result, the simple and most frequently used signal model is insufficient.

In the second part of this chapter, the concept of OFDM which is known to be more sensitive to the frequency synchronization than the single-carrier one, is introduced. The matrix notation is also used to describe the system in a more compact way. The effect of CFO of the demodulated OFDM signal is presented. The inter-carrier-interference (ICI) induced by CFO appears after the signal demodulation.

Finally, three receiving diversity techniques, which can be used to reduce the effect of multipath fading in the received signal, are introduced.

Chapter 3

Frequency Offset Estimation: Single-Carrier Case

Estimation of the frequency of a single sinusoid embedded in additive white Gaussian noise is one of the classical problems in signal processing. It is of fundamental importance in many applications, i.e., wireless communications, radar/sonar, measurements, and geophysical exploration, among others. Two signal models, namely constant envelope and time-varying envelope, are often used in the frequency estimation problem. The latter is seen as the more general case of the former. However, treating the constant envelope separately has provided significant insights into the problem, moreover many techniques including the analytical methods developed for the constant envelope case could be efficiently extended to the time-varying envelope model.

3.1 Constant Envelope

In communications context, the constant envelope model is often used in the data-aided approach carrier frequency offset (CFO) estimation problem. The channel is assumed to be a time-invariant AWGN channel with perfect time synchronization. The discrete time constant envelope signal model is often defined as

$$y(k) = a\exp\{j2\pi f_0 k\} + n(k), \qquad 0 \leq k \leq N-1 \tag{3.1}$$

where $a = |a|\exp\{j\phi\}$ is a constant complex envelope, $f_0 = F_0 T$ is the normalized frequency to be estimated, $n(k) = n_I(k) + jn_Q(k)$ is a zero-mean complex-valued white Gaussian process with variance σ^2. The SNR is defined as $|a|^2/\sigma^2$. Note that $|a|$, ϕ, f_0 are deterministic, but unknown.

The theoretical Cramer-Rao lower bound (CRLB) for the estimation variance is derived in the Appendix A, and is given as

$$\mathrm{CRLB}(\hat{f}_0) = \frac{6 \cdot \mathrm{SNR}^{-1}}{(2\pi)^2 N(N^2-1)}. \tag{3.2}$$

3.1.1 Maximum likelihood estimator

It is well known that in most applications, excellent estimates of sought frequencies are easily obtained by peak-picking the periodogram of the data, ie., the magnitude square of the discrete-time Fourier transform, [16]. Besides the fact that the periodogram is an excellent frequency estimator, it can be efficiently implemented using the fast Fourier transform (FFT) of the observation followed by a search, or interpolation, for the spectral maxima. The Maximum Likelihood Estimator (MLE) for f_0 has the form of (see Appendix A.1.2)

$$\begin{aligned} \hat{f}_0 &= \arg\max_{\tilde{f}_0} \frac{1}{N} \left| \sum_{k=0}^{N-1} y(k) \exp\{-j2\pi \tilde{f}_0 k\} \right|^2 \\ &= \arg\max_{\tilde{f}_0} P_y(\tilde{f}_0) \end{aligned} \tag{3.3}$$

where $P_y(\cdot)$ is known as the periodogram of $y(k)$, $\tilde{f}_0$ is the tentative value of f_0. The MLE produces the estimate of f_0 which is consistent and asymptotically efficient (as $N \to \infty$, its variance equals to the CRLB) [17].

In practice, (3.3) can be realized by means of Fast Fourier Transform (FFT) followed by a search for spectral maxima. The frequency variable is sampled $\tilde{f}_0 \triangleq \frac{n}{N}$, $n = 0, \cdots, N-1$. If N is a power of 2, the N-point radix-2 FFT implementation can be directly used. The computation of the N-point radix-2 FFT requires about $\frac{1}{2} N \log_2 N$ flops (1 flop = 1 complex multiplication plus 1 complex addition). If not, the length may be increased by means of zeros padding until the length is a power of 2. Zero padding provides a better representation of the continuous-frequency estimated spectrum when the frequency sampling is too sparse. Applying the FFT to the data sequence

with zero padding reveals finer details in the spectrum which were not visible without zero padding. It must be, however, noted that zero padding can not improve the spectral resolution. A method to further improve the peak localization involves fitting the peak of the periodogram and its two nearest neighbors (see Fig. 3.1) to the quadratic function. A closed-form fine estimate is then obtained as (see Appendix A.1.3)

$$\hat{f}_{\text{fine}} = -\frac{1}{2}\frac{(f_3^2 - f_2^2)p_1 + (f_1^2 - f_3^2)p_2 + (f_2^2 - f_1^2)p_3}{(f_2 - f_3)p_1 + (f_3 - f_1)p_2 + (f_1 - f_2)p_3} \tag{3.4}$$

where $(f_1, p_1), (f_2, p_2), (f_3, p_3)$ are the point left to the peak, the peak, and the point right to the peak, respectively.

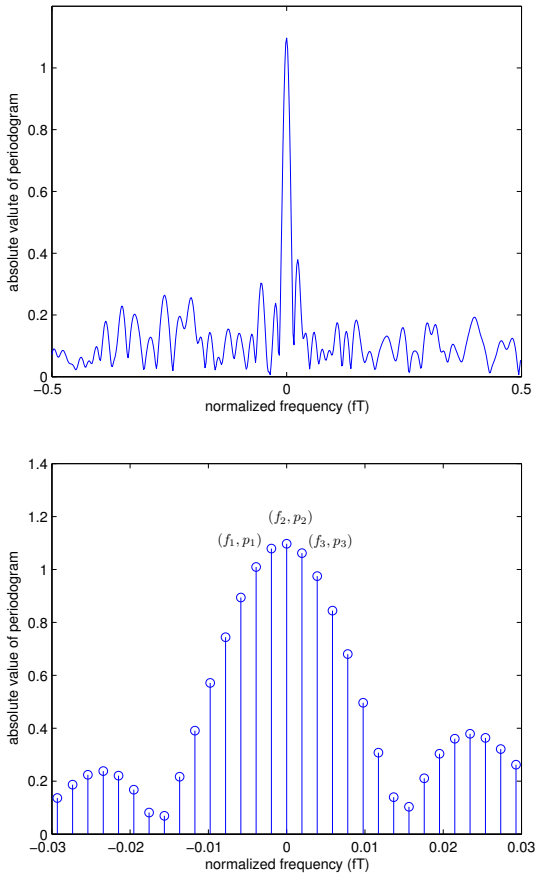

Figure 3.1: A typical periodogram and its peak's vicinity points for SNR = 0 dB and N = 64.

3.1.2 Approximated maximum likelihood estimators

The MLE is known to provide the optimal estimate of f_0, however, it is prohibited from several applications because of it's high computational requirement. In the past years, there are many significant contributions in approximated MLE which yield simpler form of estimators. The approximated maximum likelihood estimators presented here represent the state-of-the-art in frequency estimation. The main ideas and useful derivations for each estimator are therefore summarized in Appendix A.

Tretter estimator

Tretter [18] proposed an approach that provides the significant insightful result that frequency and phase estimation can be equivalently seen as the linear regression of the phase data. Tretter first proved that for a sufficiently high SNR, the complex noise, $n(k)$, can be transformed to the phase noise, $v_Q(k)$, with the variance reduced by factor of 2 (see Appendix A.1.5)

$$\arg\{y(k)\} \approx [2\pi k f_0 + \phi + v_Q(k)]_{-\pi}^{\pi} \tag{3.5}$$

where $y(k)$ is the received signal defined in (3.1) with $\arg\{x\}$ representing the angle of x, $v_Q(k)$ is a real-valued zero-mean white Gaussian noise with variance $\sigma_n^2/(2|a|^2)$, and $[x]_{-\pi}^{\pi}$ is the modulo-2π operation.

He further suggested estimating f_0 and ϕ by linear regression on the signal phase. Tretter estimator has the form of (see Appendix A.1.6)

$$\hat{f}_0 = \frac{12}{2\pi N(N^2-1)} \sum_{k=0}^{N-1} \left[k - \frac{(N-1)}{2} \right] \arg\{y(k)\}. \tag{3.6}$$

Note that a phase unwrapping procedure is needed for $\arg\{y(k)\}$.

A simple one-dimension phase unwrapping procedure based on Itoh's analysis [19] is summarized in the following steps. This procedure unwraps the phase in the array $\psi(i) \in (-\pi, \pi]$ for $0 \leq i < M-1$. In this case $\psi(i) \triangleq \arg\{y(i)\}$.

- Compute phase differences: $D(i) = \psi(i+1) - \psi(i)$ for $i = 0, \ldots, M-2$.
- Compute the wrapped phase differences: $\Delta_{wp}(i) = \arctan\{\frac{\sin D(i)}{\cos D(i)}\}$ for $i = 0, \ldots, M-2$.
- Initialize the first unwrapped value: $\hat{\phi}(0) = \psi(0)$.

- Unwrap by summing the wrapped phase differences: $\hat{\phi}(i) = \hat{\phi}(i-1) + \Delta_{wp}(i-1)$ for $i = 1, \ldots, M-1$.

Kay estimator

The only difficulty with Tretter estimator is that the phase needs to be unwrapped in computing f_0 and ϕ. This phase unwrapping, besides adding to the computation, may prove to be difficult at lower SNR's. To avoid phase unwrapping, Kay [20] derived an estimator based on the phase differences between two consecutive samples. The phase differences take the form of

$$\begin{aligned} \Delta\varphi(k) &= \arg\{y(k)y^*(k-1)\} \\ &\approx 2\pi f_0 + v_Q(k) - v_Q(k-1) \end{aligned} \tag{3.7}$$

provided that $|2\pi f_0| < \pi$ and $v_Q(k)$ is sufficiently small. It is clear from (3.7) that the problem now is to estimate the mean f_0 of a colored Gaussian noise The maximum-likelihood estimate of f_0 is found as (see Appendix A.1.7)

$$\hat{f}_0 = \frac{1}{2\pi} \sum_{k=1}^{N-1} \frac{6k(N-k)}{N(N^2-1)} \Delta\varphi(k). \tag{3.8}$$

The estimation variance of Kay estimator is shown to attain the CRLB at high SNR. Kay also suggested to interchange the summation and the argument operations in (3.8), however resulting in an inferior alternative to his original one. The analysis of this comment from Kay has been carried out in [21]. The generalized Kay estimator with arbitrary lag greater than one, is derived in [22].

Fitz estimator

A promising approach for single frequency estimation relying on calculating or approximating the autocorrelation of the received signal, is known to provide a good threshold. This concept was first introduced by Fitz [23] [24] and later improved by Luise and Reggiannini (L&R estimator) [25]. Fitz

estimator has the form of (see Appendix A.1.8)

$$\hat{f}_0 = \frac{1}{2\pi} \sum_{m=1}^{L} \frac{6m}{N(N+1)(2N+1)} \arg\{\hat{r}(m)\}. \tag{3.9}$$

where $\hat{r}(m)$ is the estimated autocorrelation of the samples, $y(k)$, defined as

$$\hat{r}(m) = \frac{1}{N-m} \sum_{k=m}^{N-1} y(k)y^*(k-m), \quad 1 \le m \le L \tag{3.10}$$

where L is the design parameter not greater than $\frac{N}{2}$. Note that the estimation range of Fitz estimator is limited to $|f_0| \le \frac{1}{2L}$. To enlarge the range, phase unwrapping algorithm such as the one presented previously can be used. The optimal value of L is found to be $\sim 17N/20$.

Mengali estimator

Though Fitz estimator is accurate (the estimation variance attains the CRLB) even at low SNR, it suffers from the relatively small estimation range. Mengali and Morelli [26] derived an estimator based on the phase differences of sample correlations which is conceptually similar to Kay's approach. The estimation range of the Mengali estimator is about $\pm 50\%$ of the symbol rate at high SNR. Mengali estimator has the form of (see Appendix A.1.9)

$$\hat{f}_0 = \frac{3}{2\pi} \sum_{m=1}^{L} \frac{(N-m)(N-m+1) - L(N-L)}{L(4L^2 - 6LN + 3N^2 - 1)} \Delta\varphi(m) \tag{3.11}$$

where $\Delta\varphi(m) = \arg\{\hat{r}(m)\hat{r}^*(m-1)\}$ and $\hat{r}(m)$ is similar to (3.10). The estimation variance of Mengali estimator attains the CRLB when $L = N/2$.

3.1.3 Proposed estimators

In comparing the estimators, the following four performance figures are frequently referred to: *Accuracy*: estimation error variance, *Estimation range*: unambiguous estimation region, *Threshold*: the SNR below which large estimation errors begin to occur, and *Implementation complexity*: number of

operations.

These figures are usually contradicting each other. For example, achieving a low threshold implies a high complexity. Likewise, good estimation accuracy is often paid at the price of a narrow estimation range. So a trade off is called for between various options which can only be made by carefully specifying the actual operation conditions.

Modified Kay estimator

To reduce the threshold of Kay estimator, an estimator based on averaging noise over samples is developed [1]. The idea is to equally divide the observation vector $\mathbf{y} = [y(0), \cdots, y(N-1)]^T$ into segments, each segment contains M samples. The samples within the same segment are added together to produce a new sample $z(k)$ which can be expressed as (see Fig. 3.2)

$$z(k) = \frac{1}{M} \sum_{m=0}^{M-1} y(kM + m), \qquad 0 \leq k \leq P-1 \tag{3.12}$$

where $P = \frac{N}{M}$, and $y(k)$ is defined as in (3.1).

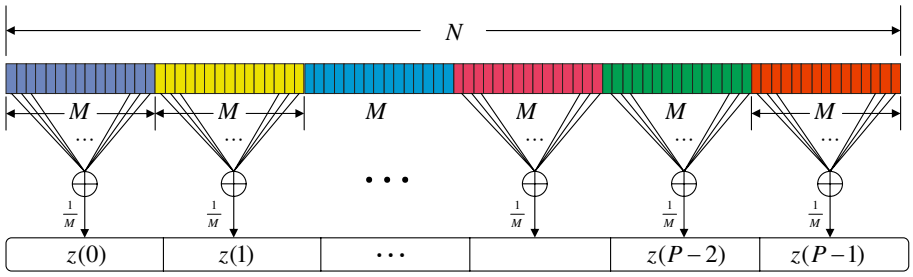

Figure 3.2: Producing new observation sequence by segmenting and adding.

Substituting (3.1) into (3.12), one obtains

$$\begin{aligned} z(k) &= \frac{|a|}{M}\sum_{m=0}^{M-1}\exp\{j[2\pi f_0(kM+m)+\phi]\}+n(kM+m) \\ &= \exp\{j[2\pi f_0 kM+\phi]\}\cdot\frac{|a|}{M}\sum_{m=0}^{M-1}\exp\{j2\pi f_0 m\}+\frac{|a|}{M}\sum_{m=0}^{M-1}n(kM+m) \\ &= \exp\{j2\pi f_0 kM+\phi\}\cdot\left(\frac{|a|}{M}\sum_{m=0}^{M-1}\exp\{j2\pi f_0 m\}+v(k)\right) \end{aligned} \tag{3.13}$$

where

$$v(k) = -\exp\{j[2\pi f_0 kM+\phi]\}\frac{|a|}{M}\sum_{m=0}^{M-1}n(kM+m) \tag{3.14}$$

are independent zero mean couple Gaussian random variables with variance $\sigma_v^2 = \frac{1}{M}\frac{\sigma_n^2}{|a|^2} = \frac{(SNR)^{-1}}{M}$.

To simplify the derivation, f_0 is assumed to be sufficiently small so that $\exp\{j2\pi f_0 m\} \approx 1$. As a result, (3.13) becomes

$$z(k) = \exp\{j[2\pi f_0 kM+\phi]\}(|a|+v(k)) \tag{3.15}$$

as shown in [18] for high SNRs, the complex noise $v(k)$ can be transformed into the phase noise

$$z(k) = |a|\exp\{j[2\pi f_0 kM+\phi+v_Q(k)]\} \tag{3.16}$$

where $v_Q(k)$ is the imaginary part of $v(k)$ with $\sigma_{v_Q}^2 = \frac{\sigma_v^2}{2}$.

We define the angular frequency offset as $\omega_\Delta = 2\pi\Delta f$ and the instantaneous phase of $z(k)$ as $\varphi(k) \triangleq \arg\{z(k)\} = [\omega_\Delta k+\phi+v_Q(k)]_{-\pi}^{\pi}$. To estimate the frequency offset, as suggested in [20], the phase differences of

two consecutive samples are considered

$$\begin{aligned}\Delta(k) &= \arg\{z(k)z^*(k-1)\}, \qquad 1 \le k \le P-1 \\ &= [\varphi(k) - \varphi(k-1)]_{-\pi}^{\pi} \\ &= \omega_\Delta M + v_Q(k) - v_Q(k-1) - 2\pi m(k)\end{aligned} \tag{3.17}$$

where $m(k)$ is an integer representing the modulo operation. For high SNR, $m(k) = 0$ but for small SNR, $m(k) \neq 0$ which leads to an ambiguity in the estimation of Δf. However, by adding up the samples, the noise variance in (3.17) reduces by factor of $\frac{1}{M}$ as compared to the original approach without adding. The reduced noise variance makes the estimation of Δf possible in the lower SNR regions since the chances that $m(k) \neq 0$ decrease.

In (3.17), Δf and $\Delta(k)$ are related in a linear fashion and the problem is to estimate a constant, ω_Δ, from noisy observations, $\Delta(k)$. By performing the maximum-likelihood estimation over $\Delta(k)$, the maximum-likelihood estimator is found by minimizing the log-likelihood function

$$J = (\mathbf{\Delta} - \omega_\Delta M\mathbf{l})^T \mathbf{C}^{-1} (\mathbf{\Delta} - \omega_\Delta M\mathbf{l}) \tag{3.18}$$

where $\mathbf{\Delta} = [\Delta(1), \Delta(2), \cdots, \Delta(P-1)]^T$, $\mathbf{l} = [1, 1, \cdots, 1]^T$, and $\mathbf{C}$ is $(P-1)\times(P-1)$ covariance matrix of $\Delta(k)$. The solution is obtained by $\frac{d}{d\omega_\Delta} J = 0$ which results in

$$\hat{\omega}_\Delta = \frac{1}{M} \frac{\mathbf{l}^T \mathbf{C}^{-1} \mathbf{\Delta}}{\mathbf{l}^T \mathbf{C}^{-1} \mathbf{l}}. \tag{3.19}$$

The covariance matrix $\mathbf{C}$ is found to be

$$\mathbf{C} = \frac{(SNR)^{-1}}{2M} \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}. \tag{3.20}$$

The inverse of (3.20) with the (i, j) element is given by

$$\left[\mathbf{C}^{-1}\right]_{ij} = \frac{2M}{(SNR)^{-1}} \left[\min(i, j) - \frac{ij}{P}\right], \quad i \ge 1,\ j \le P-1 \tag{3.21}$$

where $\min(i,j)$ denotes the minimum between i and j.

Performing some algebra based on (3.21), a closed-form estimator can be obtained. The explicit form of $\mathbf{C}^{-1}$ is

$$\mathbf{C}^{-1} = \frac{2M}{(SNR)^{-1}} \cdot \begin{pmatrix} 1-\frac{1}{P} & 1-\frac{2}{P} & 1-\frac{3}{P} & \cdots & 1-\frac{1(P-1)}{P} \\ 1-\frac{2}{P} & 2-\frac{4}{P} & 2-\frac{6}{P} & \cdots & 2-\frac{2(P-1)}{P} \\ 1-\frac{3}{P} & 2-\frac{6}{P} & 3-\frac{9}{P} & \cdots & 3-\frac{3(P-1)}{P} \\ \vdots & \cdots & \cdots & \cdots & \vdots \\ 1-\frac{1(P-1)}{P} & 2-\frac{2(P-1)}{P} & 3-\frac{3(P-1)}{P} & \cdots & (P-1)-\frac{(P-1)^2}{P} \end{pmatrix}. \tag{3.22}$$

First, $\mathbf{l}^T\mathbf{C}^{-1}\mathbf{l}$ is calculated as follows:

$$\mathbf{l}^T\mathbf{C}^{-1} = \frac{2M}{(SNR)^{-1}} \cdot [1s-0,\ 2s-1,\ 3s-3,\ 4s-6, \cdots] \tag{3.23}$$

where $s = \sum_{i=1}^{P-1}\left(1-\frac{i}{P}\right) = \frac{(P-1)}{2}$. The result yields

$$\begin{aligned} \mathbf{l}^T\mathbf{C}^{-1}\mathbf{l} &= \frac{2M}{(SNR)^{-1}} \sum_{k=1}^{P-1}\left(sk - \frac{k}{2}(k-1)\right) \\ &= \frac{2M}{(SNR)^{-1}} \left(\left(s+\frac{1}{2}\right)\sum_{k=1}^{P-1} k - \frac{1}{2}\sum_{k=1}^{P-1} k^2\right) \\ &= \frac{2M}{(SNR)^{-1}} \cdot \frac{P(P^2-1)}{12}. \end{aligned} \tag{3.24}$$

Using (3.23), $\mathbf{l}^T\mathbf{C}^{-1}\mathbf{\Delta}$ can then be calculated

$$\begin{aligned} \mathbf{l}^T\mathbf{C}^{-1}\mathbf{\Delta} &= \frac{2M}{(SNR)^{-1}} \sum_{k=1}^{P-1}\left(sk - \frac{k}{2}(k-1)\right)\Delta(k) \\ &= \frac{2M}{(SNR)^{-1}} \cdot \frac{1}{2}\sum_{k=1}^{P-1} k(P-k)\Delta(k). \end{aligned} \tag{3.25}$$

Substituting (3.24) and (3.25) into (3.19), results in a closed-form estimator

$$\begin{aligned}\hat{\omega}_{\Delta} &= \frac{1}{M} \cdot \frac{\mathbf{l}^T \mathbf{C}^{-1} \boldsymbol{\Delta}}{\mathbf{l}^T \mathbf{C}^{-1} \mathbf{l}} \\ &= \frac{1}{M} \cdot \sum_{k=1}^{P-1} \frac{6k(P-k)}{P(P^2-1)} \Delta(k).\end{aligned} \tag{3.26}$$

The corresponding estimation variance is

$$\begin{aligned}\mathrm{var}(\hat{\omega}_{\Delta}) &= \frac{1}{M^2} \cdot \frac{1}{\mathbf{l}^T \mathbf{C}^{-1} \mathbf{l}} \\ &= \frac{1}{M^2} \cdot \frac{6}{P(P^2-1)} \cdot \frac{(SNR)^{-1}}{M} \\ &= \frac{6}{N(N^2-M^2)} \cdot (SNR)^{-1}.\end{aligned} \tag{3.27}$$

If $M = 1$, the estimation variance attains the CRLB and the proposed estimator is identical to the original Kay estimator.

Improved Fitz estimator

Fitz frequency estimator [23, 24] relies on the estimated sample autocorrelation and is known to provide efficient estimates even at low SNR. However, it suffers from the narrow estimation range mainly caused by the phase ambiguities of the estimated sample autocorrelation. Kay [20] showed that the estimation range of an estimator which is based only on the phase of the received signal can be extended if the phase difference, $\arg\{y(m)y^*(m-1)\}$, is exploited, instead of the single phase, $\arg\{y(k)\}$. Mengali and Morelli [26] used the same concept, but with the autocorrelation estimates, $\arg\{\hat{r}(m)\hat{r}^*(m-1)\}$, of the received signal sample rather than the received signal itself. The weighting function for such a phase difference based estimator is derived using maximum likelihood estimation.

To explain the benefit of using the phase difference of the estimated correlation, the noiseless condition is assumed. The phase of $\hat{r}(m)$ then reads

$$\phi(m) = \arg\{\hat{r}(m)\} = [2\pi f_0 m]_{-\pi}^{\pi} \tag{3.28}$$

where $[x]_{-\pi}^{\pi}$ means that x is confined within the interval $(-\pi, \pi]$. As m varies from 1 to $N-1$, eq. (3.28) establishes a linear relation between f_0 and $\hat{r}(m)$, only when f_0 lies within the interval about $\pm\frac{1}{2N}$. If f_0 lies outside the interval, eq. (3.28) becomes a highly nonlinear function and introduces ambiguities.

On the other hand, the phase difference of one lag reads

$$\Delta\phi(m) = \arg\{\hat{r}(m)\hat{r}^*(m-1)\} = [2\pi f_0]_{-\pi}^{\pi}\,. \tag{3.29}$$

Eq. (3.29) is independent from m which means that, as long as f_0 lies within $\pm\frac{1}{2}$, $\phi(m)$ keeps equal to $2\pi f_0$ for $1 \leq m \leq N-1$. Clearly, (3.29) relates f_0 and $\phi(m)$ in a linear fashion and the problem is to estimate a constant, f_0, from the noisy measurements, $\phi(m)$. It can be seen that the estimation range is enlarged roughly by a factor of N using phase differences as the observations.

The objective is to transform Fitz estimator (3.9) of the form

$$\hat{f}_0 = \frac{1}{2\pi}\sum_{m=1}^{L} w_m \arg\{\hat{r}(m)\} \tag{3.30}$$

into the the estimator of the following form

$$\hat{f}_0 = \frac{1}{2\pi}\sum_{m=1}^{L} \tilde{w}_m \arg\{\hat{r}(m)\hat{r}^*(m-1)\}. \tag{3.31}$$

The direct derivation as considered by Mengali and Morelli is promising, but it can be lengthy. Alternatively, the well known summation-by-parts formula can be used for this purpose [6, 7]. This method assumes that the weighting function w_m is not a constant and $w_0 = 0$. Consider the following summation-by-parts formula

$$\sum_{m=0}^{L} \Delta b_m a_m = b_L a_L - b_{-1}a_0 - \sum_{m=1}^{L} b_{m-1}\Delta a_m \tag{3.32}$$

where $\Delta b_m = b_m - b_{m-1}$ and $\Delta a_m = a_m - a_{m-1}$. Inspecting (3.32), and the

summation terms in (3.30) and (3.31) yields

$$\begin{aligned}\sum_{m=1}^{L} \underbrace{w_m}_{\Delta b_m} \cdot \underbrace{\arg\{\hat{r}(m)\}}_{a_m} \\ &= b_L \arg\{r(L)\} - \sum_{m=1}^{L} b_{m-1}\Delta\phi(m) \\ &= \sum_{m=1}^{L} \underbrace{(b_L - b_{m-1})}_{\tilde{w}_m} \Delta\phi(m)\end{aligned} \tag{3.33}$$

where $\Delta b_0 = 0$, $b_{-1} = 0$, $\Delta\phi(m) = \arg\{\hat{r}(m)\hat{r}^*(m-1)\}$, $\arg\{r(N-1)\} = \sum_{m=1}^{N-1} \Delta\phi(m)$. Since $b_{-1} = 0$ and $\Delta b_0 = 0$, one obtains

$$b_m = \sum_{k=1}^{m} \Delta b_k. \tag{3.34}$$

Inserting (3.34) into (3.33), the new weighting function $\tilde{w}_m$ can be calculated from w_m by

$$\tilde{w}_m = \sum_{k=1}^{L} \Delta b_k - \sum_{k=1}^{m-1} \Delta b_k \tag{3.35}$$

where $\Delta b_k = \{w_m\}_{m=k}$. Therefore, given w_m and $r(m)$, one can transform (3.30) to the new estimator, having the form of (3.31):

$$\begin{aligned}\hat{f}_0 &= \frac{1}{2\pi} \sum_{m=1}^{L} \left[\sum_{k=1}^{L} w_k - \sum_{k=1}^{m-1} w_k \right] \Delta\phi(m) \\ &= \frac{1}{2\pi} \sum_{m=1}^{L} \underbrace{\left[\sum_{k=1}^{L-m+1} w_{L+m-1} \right]}_{\tilde{w}_m} \Delta\phi(m)\end{aligned} \tag{3.36}$$

The expression (3.36) can be used to transform a given estimator with the form of (3.30) to an estimator of the form (3.31).

The transformed Fitz estimator with the extended estimation range is

$$\hat{f}_0 = \frac{1}{2\pi} \sum_{m=1}^{L} \underbrace{3\frac{N(N+1) - m(m-1)}{N(N+1)(2N+1)}}_{\tilde{w}_m} \arg\left\{\hat{r}(m)\hat{r}^*(m-1)\right\}. \tag{3.37}$$

Note that by comparing $\tilde{w}_m$ in (3.37) with w_m in (3.9) it can be noticed that

$$w_m = \tilde{w}_m - \tilde{w}_{m+1} \tag{3.38}$$

as a result, for a given $\tilde{w}_m$, one can also reversely obtain w_m.

Approximated ML estimators

Two correlation-based estimators are derived in this subsection. First, the absolute phase-based estimator is derived, then the weight transformation formula (3.36) is used to transform the weighting function of the proposed absolute phase based into differential phase based estimator.

An approximated ML estimator may be derived from a periodogram of the form

$$J(\omega_o) = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} y(n)y^*(m)e^{-j\omega_o(n-m)}. \tag{3.39}$$

The derivative of (3.39) with respect to ω_o equated to zero may be rearranged in the following form

$$\mathrm{Im}\left\{\sum_{m=1}^{N-1} m(N-m)|\hat{r}(m)|e^{j(\varphi(m)-\omega_o m)}\right\} = 0. \tag{3.40}$$

By assuming that the accurate phase estimates of $\hat{r}(m)$ are available, and

$|\hat{r}(m)|$ is a constant, (3.40) reduces to

$$\begin{aligned}\hat{\omega}_o &= \frac{\sum_{m=1}^{L} m(N-m)|\hat{r}(m)|\phi(m)}{\sum_{m=1}^{L} m^2(N-m)|\hat{r}(m)|} \\ &= \sum_{m=1}^{L} \underbrace{\frac{12m(N-m)}{L(L+1)(4LN+2N-3L^2-3L)}}_{w_m} \phi(m)\end{aligned} \tag{3.41}$$

where $L < N$ is a design parameter.

To extend the estimation range of (3.41), eq. (3.36) is then applied to w_m in (3.41) resulting in the following transformed weighting function

$$\begin{aligned}\tilde{w}_m &= \frac{12\sum_{k=1}^{L-m+1}(k+m-1)(N-k-m+1)}{L(L+1)(4LN+2N-3L^2-3L)} \\ &= 2\frac{L(L+1)(3N-2L-1)-m(m-1)(3N-2m+1)}{L(L+1)(4LN+2N-3L^2-3L)}.\end{aligned} \tag{3.42}$$

Estimation variance of correlation-based estimators

Absolute phase The absolute phase based estimator reads

$$\begin{aligned}\hat{\omega}_o &= \sum_{m=1}^{L} w_m \phi(m) \\ &\approx \omega_o \sum_{m=1}^{L} m w_m + \sum_{m=1}^{L} w_m \gamma_I(m)\end{aligned} \tag{3.43}$$

where w_m is the weighting function. This class of estimators is unbiased when $\sum_{m=1}^{L} m w_m = 1$ and $\omega_o \in [-\frac{\pi}{L}, \frac{\pi}{L})$. Thus, its estimation range is approximately $\pm\pi/L$. Outside this range, (3.43) becomes highly nonlinear and awkward to handle. To transform the nonlinearity into a linear equation over the entire range an additional phase unwrapping procedure is needed, however this would be difficult at low SNR.

The estimation variance, $\mathrm{var}(\hat{\omega}_o) \triangleq E\left\{(\hat{\omega}_o - \omega_o)^2\right\}$, of (3.43) reads

$$\mathrm{var}(\hat{\omega}_o) = \sum_{m=1}^{L}\sum_{n=1}^{L} w_m w_n \underbrace{E\left\{\gamma_I(m)\gamma_I(n)\right\}}_{\mu(m,n)} \tag{3.44}$$

where

$$\mu(m,n) = \frac{\min(m,n,N-m,N-n)}{\mathrm{SNR}(N-m)(N-n)} + \frac{\delta_{m,n}}{2\mathrm{SNR}^2(N-n)}. \tag{3.45}$$

The exact expression of $\gamma_I(m)$ and the derivation of (3.45) can be found in Appendix A.1.10.

Differential phase The differential phase-based estimator reads

$$\begin{aligned}\hat{\omega}_o &= \sum_{m=1}^{L} \tilde{w}_m \Delta\phi(m) \\ &\approx \omega_o \sum_{m=1}^{L} \tilde{w}_m + \sum_{m=1}^{L} \tilde{w}_m(\gamma_I(m) - \gamma_I(m-1))\end{aligned} \tag{3.46}$$

where $\tilde{w}_m$ is the weighting function. This class of estimators is unbiased when $\sum_{m=1}^{L} \tilde{w}_m = 1$ and ω_o lies within $[-\pi, \pi)$. Thus, its estimation range is about $\pm\pi$ or L times larger than that of (3.43) without additional phase unwrapping algorithm.

The estimation variance of (3.46) is

$$\mathrm{var}(\hat{\omega}_o) = \sum_{m=1}^{L}\sum_{n=1}^{L} \tilde{w}_m \tilde{w}_n \underbrace{E\left\{[\gamma_I(m) - \gamma_I(m-1)][\gamma_I(n) - \gamma_I(n-1)]\right\}}_{\rho(m,n)} \tag{3.47}$$

where $\rho(m,n) = \mu(m,n) + \mu(m-1,n-1) - \mu(m-1,n) - \mu(m,n-1)$.

Efficiency at high SNR The statistical efficiency at high SNR is

$$\text{eff} \triangleq \lim_{\text{SNR}\to\infty} \frac{\text{var}(\hat{\omega}_o)}{\text{CRLB}} \geq 1. \tag{3.48}$$

Substituting (3.44) and (3.47) in (3.48) yield the efficiency measures which depend on L and N. The minimization of the (3.48) with respect to L seems not to be analytically tractable. Thus the numerical minimization is carried out to find the optimum values of L that minimize (3.48).

As a result of the numerical minimization of (3.48) for differential phase based estimators, the efficiencies at high SNR of Mengali and Morelli, transformed proposed approximated ML, and modified Fitz estimators are shown in Fig. 3.3. The corresponding optimal number of lags of these estimators are also identified. It is found that the optimal number of lags (M_{opt}) for M&M is $N/2$, for modified Fitz is $0.86N$, and for proposed approximated ML estimator is $N-1$. It shall be noted that for complexity reason, the lag values in the vicinity of the second optimum (local optimum) may be used which will significantly reduce the implementation cost.

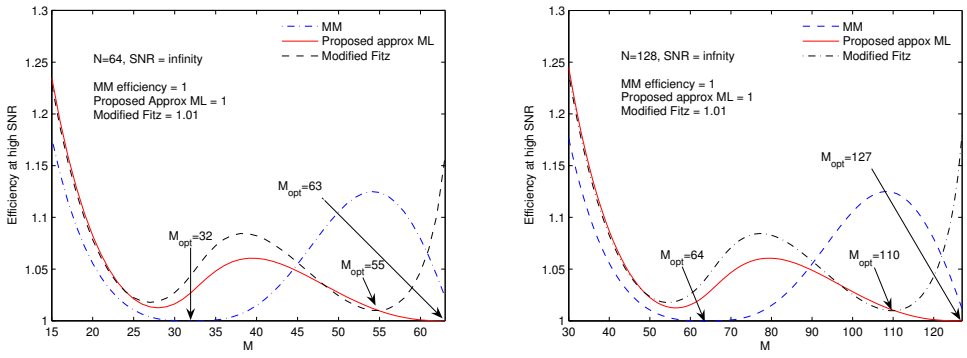

Figure 3.3: Efficiencies at high SNR of M&M, transformed proposed approximated ML, modified Fitz estimators for N=64, and 128.

3.2 Time-Varying Envelope

The traditional methods used in analysis and development of frequency estimators employ an additive white Gaussian noise (AWGN) model with the

signal. However, the additive noise model is not a sufficiently realistic representative in most of the practical applications. In practice, the received signal is likely to exhibit fading behavior caused by time-varying multipath propagation or the amplifier's nonlinearity. This fading characteristic can be observed in the form of random time-varying envelope (also known as the multiplicative noise). For radar applications, a real-valued time-varying envelope is frequently defined [27–29]. In wireless communications, on the other hand, the complex-valued time-varying envelope is usually considered which can be subdivided into two categories i.e., frequency flat fading [30–34] and frequency selective fading [35–38]. In this thesis, the only frequency flat fading case is considered. An example of such a flat fading is in global satellite communication systems employing low and medium earth orbit satellites. In these systems, the Doppler shift of the carrier frequency varies randomly and should be estimated and compensated for.

A general complex sinusoid that has been corrupted by fading and additive noise may be modeled as

$$y(k) = a(k)\exp\{j[2\pi\nu k + \theta]\} + n(k), \qquad 0 \le k \le N-1 \tag{3.49}$$

where the multiplicative noise $a(k) = |a(k)|\exp\{j\phi(k)\}$ is, in general, defined as a complex-valued Gaussian process, circular symmetric around its mean $\mu\exp\{j\varphi\}$ with $\mu \ge 0$ and $\varphi \in [-\pi, \pi)$. ν is the normalized frequency offset to be estimated. The correlation function of the multiplicative noise, $r_a(m) = \mathrm{E}\{a(k)a^*(k-m)\}$, is parameterized by a real-valued[1] parameter vector $\boldsymbol{\alpha} = [\alpha_0, \cdots, \alpha_{M-1}]^T$. Let $\sigma_a^2 \triangleq r_a(0)$. Additive noise $n(k)$ is a zero-mean circular complex white Gaussian process with variance σ_n^2. $a(k)$ and $n(k)$ are mutually independent. The SNR is defined as σ_a^2/σ_n^2.

The correlation of the multiplicative noise, $r_a(m)$, is often modelled by the following functions (see Fig. 3.4):

- Jakes model: $r_a(m) = \sigma_a^2 J_0(2\pi B_D T|m|)$
- Sinc model: $r_a(m) = \sigma_a^2 \mathrm{sinc}(2\pi B_D T|m|)$
- Exponential model: $r_a(m) = \sigma_a^2 \exp\{-2\pi B_D T|m|\}$
- Gaussian model: $r_a(m) = \sigma_a^2 \exp\{-2\pi (B_D T)^2 m^2/4\}$

[1]This assumption has been made implicitly by many authors to ensure a consistent frequency estimate via the correlation-based estimation approach. Otherwise, the phase of $r_a(m)$ has to be estimated prior to the frequency estimation.

where B_D is the Doppler bandwidth, T is the sampling time, $J_0(\cdot)$ is the 0-th order Bessel function of the first kind, and B_DT is the normalized Doppler spread[2]. For mobile cellular communications [30,31,39,40], it is customary to follow *Jakes* and *Sinc* models. In Doppler-radar and ionospheric communications [34,41], the correlation function with *Exponential* and *Gaussian* shapes are often assumed. However, it is widely recognized that the actual spectrum of $r_a(m)$ has little impact on the overall system performance. The sole influential parameter is the Doppler spread, B_D. The correlation function of the multiplicative noise reflects the variability of the channel over time. This second-order statistic generally depends on the propagation geometry, velocity of the platform, and the antenna characteristics.

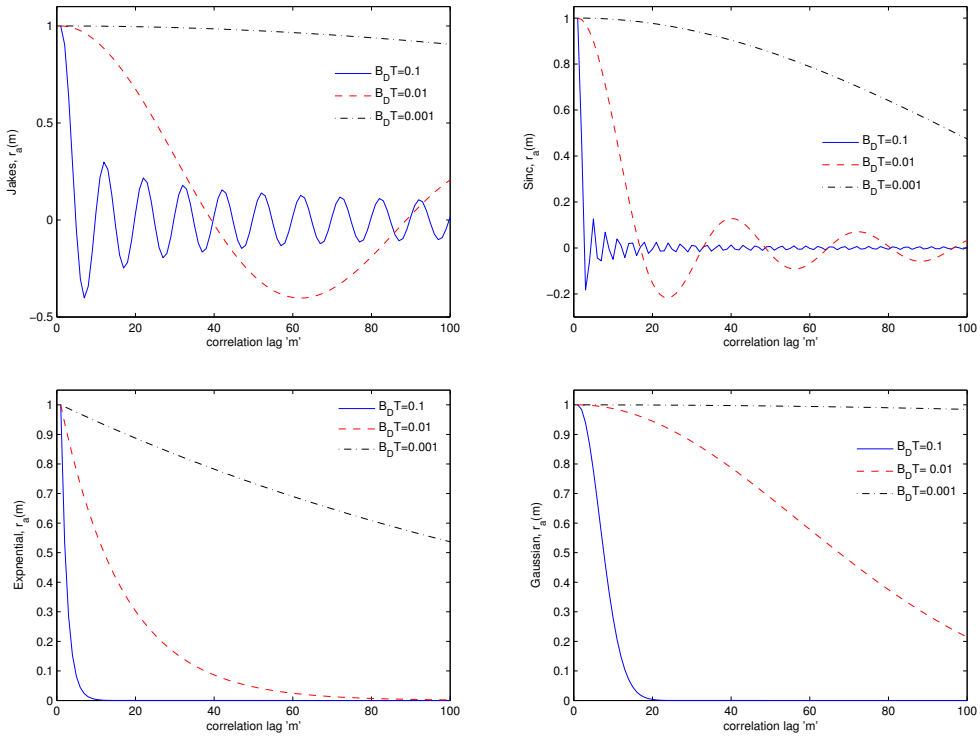

Figure 3.4: Four correlation models of the multiplicative noise with slow ($B_DT = 0.001$), intermediate ($B_DT = 0.01$), and fast fading ($B_DT = 0.1$).

As (3.49) represents a general signal model for CFO estimation problem in single-carrier systems, three special cases can be categorized as follows:

[2]Typical values for practical systems [33] range from $B_DT = 0.001$ (very slow fading) to $B_DT = 0.1$ (very fast fading).

Rayleigh fading: the amplitude of the fading process will have a Rayleigh distribution if $\mu = 0$ and $\sigma_a^2 \neq 0$.

Rician fading: the amplitude of the fading process will have a Rician (dominant+Rayleigh) distribution if $\mu \neq 0$ and $\sigma_a^2 \neq 0$.

Non-fading: or the classical constant-amplitude with only additive Gaussian noise (discussed in the previous section) is obtained when $\mu \neq 0$ and $\sigma_a^2 = 0$.

The theoretical Cramer-Rao lower bound (CRLB) is derived in the Appendix A.2.1 and is given as the corresponding diagonal element of the inverse of the Fisher information matrix (FIM), $J^{-1}_{\omega_o,\omega_o}$. The CRLB is then

$$\text{CRLB}(\hat{\omega}_o) = J^{-1}_{\omega_o,\omega_o} \tag{3.50}$$

where $\hat{\omega}_o = 2\pi\hat{\nu}$. The J_{ω_o,ω_o} is found to be

$$J_{\omega_o,\omega_o} = 2\mu^2 \mathbf{1}^T \mathbf{D}\mathbf{R}^{-1}\mathbf{D}\mathbf{1} + 2\text{tr}\{\mathbf{R}^{-1}\mathbf{D}\mathbf{R}\mathbf{D} - \mathbf{D}^2\} \tag{3.51}$$

where tr is the matrix trace, $\mathbf{R} = \mathbf{R}_a + \sigma_n^2\mathbf{I}$, $\mathbf{1}$ is a $(N \times 1)$ vector of one, $\mathbf{R}_a(i,j) = r_a(|i-j|)$ with $i,j = 1, \cdots, N$, and $\mathbf{D} = \text{diag}\{0, \cdots, N-1\}$. For the correlated Rayleigh channel $\mu = 0$, and $J_{\omega_o,\omega_o} = 2\text{tr}\{\mathbf{R}^{-1}\mathbf{D}\mathbf{R}\mathbf{D} - \mathbf{D}^2\}$.

For high SNR or $\sigma_n^2 \to 0$,

$$J^{\sigma_n^2 \to 0}_{\omega_o,\omega_o} = 2\text{tr}\{\mathbf{R}_a^{-1}\mathbf{D}\mathbf{R}_a\mathbf{D} - \mathbf{D}^2\} \tag{3.52}$$

is a constant which leads to the so called *error floor effect*.

3.2.1 Proposed estimator for complex-valued envelope

The CFO estimators proposed in the past [30–32,34] are associated with one or more of the following limitations/drawbacks: 1). Require a priori knowledge on the form of the fading correlation [30, 31], 2). High computational complexity [32], and 3). Short estimation range [34].

Kuo and Fitz [30] proposed to estimate the normalized (angular) frequency as

$$\hat{\omega}_o = \boldsymbol{\alpha}^T\hat{\boldsymbol{\theta}} \tag{3.53}$$

where $\hat{\boldsymbol{\theta}} = [\hat{\theta}(1), \ldots, \hat{\theta}(M)]^T$, and $\hat{\theta}(m) = m^{-1} \arg\{\hat{r}(m)\}$, $\hat{r}(m)$ is the m-th lag correlation estimate defined in (3.56). The optimal weight vector, $\boldsymbol{\alpha}$, is chosen to minimize the variance of (3.53) and is given by

$$\boldsymbol{\alpha} = \frac{\mathbf{C}_\theta^{-1}(B_D T, SNR)\mathbf{1}}{\mathbf{1}^T \mathbf{C}_\theta^{-1}(B_D T, SNR)\mathbf{1}} \tag{3.54}$$

where $\mathbf{C}_\theta(B_D T, SNR) = E\{(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T\}$, and $\mathbf{1} = [1, \ldots, 1]^T$. It shall be noted that $\mathbf{C}_\theta(B_D T, SNR)$ depends on $B_D T$ and SNR. This covariance matrix of phase estimation errors is assumed to be known in the work of Kuo and Fitz in [30] which is seldom available in practical situations. Moreover, phase unwrapping is required which may be delicate in small samples or in low SNR environments. To circumvent these problems, [31] proposed a similar scheme based on the difference between the phases of the correlation, $\arg\{\hat{r}(m)\hat{r}^*(m-1)\}$. Additionally, the covariance matrix of the phase estimation errors was estimated in [31] prior to performing weighted linear regression. However, this covariance matrix estimate strongly relies on an assumed form for the correlation of the fading process. Therefore methods in [30, 31] are sensitive to mismodeling of the fading correlation. A low-complexity CFO estimator, which does not assume nor estimate the channel parameter, was proposed in [34]. Its estimation range is, however, limited. The nonlinear least-squares estimator proposed in [32], is robust to the lack of information on the form of the correlation of the fading process and is able to operate at the maximum estimation range but at the price of computational complexity.

The complexity of the estimator proposed in [32] can be reduced, but at the cost of a decreased estimation range [3]. However, It will be shown in the following sections that by using the summation-by-parts formula the computational complexity can be reduced without shortening the estimation range.

Observation model

The true correlation sequence corresponding to the data model (3.49) is

$$\begin{aligned} r(m) &= E\{y(k)y^*(k-m)\} \\ &= r_a(m)e^{j2\pi m\nu} + \sigma_n^2 \delta(m). \end{aligned} \tag{3.55}$$

The estimated correlation sequence can be calculated by two standard ways, the first is

$$\hat{r}_1(m) = \frac{1}{N-m} \sum_{k=m}^{N-1} y(k)y^*(k-m), \quad m = 0, 1, \ldots, M-1 \tag{3.56}$$

and the second is

$$\hat{r}_2(m) = \frac{1}{N} \sum_{k=m}^{N-1} y(k)y^*(k-m) \tag{3.57}$$

where the former is known as the unbiased estimate of $r(m)$ since $E\{\hat{r}_1(m)\} = r(m)$, and the latter is known as the biased estimate of $r(m)$, since $E\{\hat{r}_2(m)\} = (1 - |m|/N) \cdot r(m)$. The statistical properties of these two approaches are given in Appendix A.2.2.

Correlation-based nonlinear least-squares estimator

The NLS estimates of σ_n^2, $d(m) \triangleq r_a(m)$, and ν are obtained by solving the following minimization problem

$$\{\hat{\sigma}_n^2, \hat{\mathbf{d}}, \hat{\nu}\} = \arg \min_{\sigma_n^2, \mathbf{d}, \nu} \frac{1}{M} \underbrace{\sum_{m=0}^{M} |\hat{r}(m) - d(m)e^{j2\pi m\nu} - \sigma_n^2 \delta(m)|^2}_{J(\sigma_n^2, \nu, \mathbf{d})}. \tag{3.58}$$

where $\hat{r}(m)$ can be chosen from either (3.56) or (3.57), $d(m) \triangleq r_a(m)$, and $\mathbf{d} = [d(0), \cdots, d(M)]^T$. Setting the derivative of the criterion $J(\sigma_n^2, \nu, \mathbf{d})$ w.r.t. σ_n^2 to zero, we obtain

$$\hat{\sigma}_n^2 = \hat{r}(0) - d(0). \tag{3.59}$$

Substitute this solution into $J(\sigma_n^2, \nu, \mathbf{d})$, yields the reduced criterion

$$\{\hat{\mathbf{d}}, \hat{\nu}\} = \arg \min_{\mathbf{d},\ \nu} \frac{1}{M} \underbrace{\sum_{m=1}^{M} |\hat{r}(m) - d(m)e^{j2\pi m\nu}|^2}_{J(\nu, \mathbf{d})}. \tag{3.60}$$

Since (3.56) is a consistent estimate of (3.55), it follows that $J_\infty(\nu, \mathbf{d}) \triangleq \lim_{N\to\infty} J(\nu, \mathbf{d})$ achieves a global minimum at the true parameters [32, 34]. Consequently, $\hat{\nu}$ is a consistent estimate of ν whatever form of $d(m)$. It will be shown that the criterion in (3.60) can be concentrated w.r.t. $\mathbf{d}$ so as to leave one-dimensional search over ν only. Using the fact that $d(m) \in \mathbb{R}$, $J(\nu, \mathbf{d})$ can be rewritten as

$$\sum_{m=1}^{M} \left\{ |\hat{r}(m)|^2 - 2d(m)\mathrm{Re}\left\{\hat{r}(m)e^{-j2\pi m\nu}\right\} + d^2(m) \right\}. \tag{3.61}$$

Taking the derivative of (3.61) w.r.t. $\mathbf{d}$ and equating the result to zero, yields

$$\hat{d}(m) = \mathrm{Re}\left\{\hat{r}(m)e^{-j2\pi m\nu}\right\} \tag{3.62}$$

where the estimate of the frequency offset has yet to be determined. Inserting (3.62) into (3.61), the NLS estimate of the parameter ν is obtained as

$$\hat{\nu}_{\mathrm{NLS}} = \arg \max_{\nu} \mathrm{Re}\left\{ \sum_{m=1}^{M} \hat{r}^2(m)e^{-j4\pi m\nu} \right\} \tag{3.63}$$

which can be efficiently realized by means of Fast Fourier Transform (FFT) with zero padding. This can be done by $O(N_f \log_2 N_f)$ operations, where N_f is the Fourier bins. A better localization of the peak can be obtained by using the quadratic interpolation as introduced in (3.4).

However, a much simpler form of (3.63) is obtained by assuming the availability of the phase of the correlation estimate, $\phi_m = \arg\{\hat{r}(m)\}$. The NLS criterion (3.63) can be rewritten as

$$\sum_{m=1}^{M} |\hat{r}(m)|^2 \cos(2\phi_m - 4\pi m\nu). \tag{3.64}$$

Setting the derivative of (3.64) w.r.t. ν to zero, we obtain

$$\sum_{m=1}^{M} m|\hat{r}(m)|^2 \sin(2\phi_m - 4\pi m\nu) = 0. \tag{3.65}$$

Under the small error approximation, i.e., $\sin(2\phi_m - 4\pi m\nu) \simeq (2\phi_m - 4\pi m\nu)$, the simplified NLS (SNLS) estimator is obtained as

$$\begin{aligned}\hat{\nu}_{\mathrm{SNLS}} &= \frac{1}{2\pi}\frac{\sum_{m=1}^{M} m|\hat{r}(m)|^2\phi_m}{\sum_{m=1}^{M} m^2|\hat{r}(m)|^2} \\ &= \sum_{m=1}^{M}\alpha_m\phi_m\end{aligned} \tag{3.66}$$

where α_m is the weighting function of the absolute phase-based estimator.

Note that without an additional phase unwrapping algorithm, e.g., [19], the estimation range of the SNLS is limited to about $\pm 1/2M$ the symbol rate.

Weight transformation formula

Using $A_m = \sum_{k=0}^{m}\alpha_k$, $\phi_M = \sum_{m=1}^{M}(\phi_m - \phi_{m-1})$, and $A_{-1} = 0$, the relationship between the absolute phase based and phase difference based estimators is derived as

$$\begin{aligned}\sum_{m=0}^{M}\alpha_m\phi_m &= \sum_{m=0}^{M}(A_m - A_{m-1})\phi_m \\ &= \sum_{m=0}^{M}A_m\phi_m - \sum_{m=-1}^{M-1}A_m\phi_{m+1} \\ &= A_M\phi_M - A_{-1}\phi_{-1} + \sum_{m=0}^{M-1}A_m(\phi_m - \phi_{m+1}) \\ &= A_M\phi_M - \sum_{m=1}^{M}A_{m-1}(\phi_m - \phi_{m-1}) \\ &= \sum_{m=1}^{M}(A_M - A_{m-1})(\phi_m - \phi_{m-1}) \\ &= \sum_{m=1}^{M}\underbrace{\left(\sum_{k=1}^{M}\alpha_k - \sum_{k=1}^{m-1}\alpha_k\right)}_{\beta_m}\Delta\phi_m.\end{aligned} \tag{3.67}$$

Thus, the weight transformation formula reads

$$\beta_m = \sum_{k=1}^{M-m+1} \alpha_{k+m-1} \tag{3.68}$$

or reversely $\alpha_m = \beta_m - \beta_{m+1}$.

It can be shown that if $\sum_{m=1}^{M} m\alpha_m = 1$, then $\sum_{m=1}^{M}\sum_{k=1}^{M-m+1} \alpha_{k+m-1} = 1$. Thus, if original estimator is unbiased, the transformed estimator is also unbiased.

Note that the weight transformation formula (3.68) can also be obtained using the summation-by-parts formula introduced in the previous chapter.

Estimation range extension by weight transformation formula

To avoid an additional phase unwrapping procedure, the phase difference of the estimated correlation, $\Delta\phi_m = \arg\{\hat{r}(m)\hat{r}^*(m-1)\}$, shall be exploited instead of $\phi_m = \arg\{\hat{r}(m)\}$. For this purpose, the weight transformation formula (3.68) is applied to (3.66) to properly transform the SNLS estimator into a new phase-difference-based estimator with a proper weighting function. Thus, the transformed SNLS (or approximated NLS, ANLS) is

$$\hat{\nu}_{\text{ANLS}} = \frac{1}{2\pi} \cdot \frac{\sum_{k=1}^{M} k|\hat{r}(k)|^2 \sum_{m=1}^{M} \Delta\phi_m - \sum_{m=2}^{M}\sum_{k=1}^{m-1} k|\hat{r}(k)|^2 \Delta\phi_m}{\sum_{m=1}^{M} m^2|\hat{r}(m)|^2}. \tag{3.69}$$

For an AWGN channel, or $|\hat{r}(k)|$ is a constant, (3.69) reduces to

$$\hat{\nu}_{\text{AWGN}} = \frac{3}{2\pi} \sum_{m=1}^{M} \frac{(M+1)M - (m-1)m}{M(2M+1)(M+1)} \Delta\phi_m \tag{3.70}$$

which has a similar form as the modified Fitz estimator (3.37) in the constant envelope case.

The implementation complexities of different versions of NLS estimators are presented in Table 3.1.

Estimators	Number of complex multiplications
FFT-based NLS	$N_f \log_2 N_f + \frac{M(2N-M-1)}{2}$
SNLS	$\frac{M(2N-M-1)}{2}$ + Phase unwrapping
ANLS	$\frac{M(2N-M-1)}{2} + (M-2)$
AWGN	$\frac{M(2N-M-1)}{2} + (M-2)$

Table 3.1: Complexity of different versions of NLS estimators.

3.2.2 Proposed EKF for real-valued envelope

As a special case of the complex-valued envelope, the real-valued envelope is also frequently used in radar applications, e.g., [27–29]. The signal envelope, $a(t)$, is real-valued and can be modelled as random (multiplicative noise) or deterministic but time varying. If the batch processing is considered, good examples can be found in [27, 42]. This model has been applied to a real application in estimating particle's velocity in the vicinity of an aircraft by means of a laser velocimeter [43, 44]. In the following, a recursive frequency tracker known as Kalman filter, is proposed. The signal envelope is assumed to be real-valued and deterministic. Since the frequency to be estimated is governed by the nonlinear complex exponential function, therefore the extended Kalman filter (EKF) shall be formulated [4, 45].

Extended Kalman filter

In this subsection, the system state and observation models are formulated. The Kalman filtering algorithm is given. The stochastic observability and the alternative of the Kalman filter are briefly discussed.

System state model

The continuous system state is defined by

$$\mathbf{x}(t) = [a(t) \;\; \dot{a}(t) \;\; \psi(t) \;\; \dot{\psi}(t)]^T \tag{3.71}$$

where $\psi(t)$ is the instantaneous phase of $y(k)$ with $\dot{\psi}(t)$ as its derivative. $\dot{a}(t)$ is the derivative of $a(t)$.

The discrete system state is obtained by letting $\mathbf{x}_k = \mathbf{x}(t)|_{t=kT}$

$$\mathbf{x}_k = [x_{1,a}(k) \;\; x_{2,a}(k) \;\; x_{1,p}(k) \;\; x_{2,p}(k)]^T. \tag{3.72}$$

The difference equations of the time-varying amplitude are defined as

$$\begin{aligned} x_{1,a}(k+1) &= x_{1,a}(k) + x_{2,a}(k) \\ x_{2,a}(k+1) &= x_{2,a}(k) + w_a(k). \end{aligned} \tag{3.73}$$

The time-varying instantaneous phase can also be modelled in the same way

$$\begin{aligned} x_{1,p}(k+1) &= x_{1,p}(k) + 2\pi \cdot x_{2,p}(k) \\ x_{2,p}(k+1) &= x_{2,p}(k) + w_p(k). \end{aligned} \tag{3.74}$$

The corresponding discrete time state-space model can then be expressed as

$$\begin{aligned} \mathbf{x}_{k+1} &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2\pi \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{x}_k + \begin{bmatrix} 0 \\ w_a(k) \\ 0 \\ w_p(k) \end{bmatrix} \\ &= \mathbf{A} \cdot \mathbf{x}_k + \mathbf{w}_k \end{aligned} \tag{3.75}$$

where $\mathbf{A}$ is the state transition matrix, $\mathrm{E}\{\mathbf{w}_k\} = \mathbf{0}$ and $\mathrm{E}\{\mathbf{w}_k \mathbf{w}_j^T\} = \mathbf{Q}(k) \cdot \delta(k, j)$.

Observation model

We define the first two elements of the observation vector from the inphase and quadrature component of (3.49)

$$\begin{aligned} y_I(k) &= a(k) \cos(2\pi\nu k + \theta) + v_I(k) \\ y_Q(k) &= a(k) \sin(2\pi\nu k + \theta) + v_Q(k). \end{aligned} \tag{3.76}$$

We introduce a further observation, $y_p(k) = \arg\{r(k)\}$, which can be approximated at high SNR [18] as

$$y_p(k) \approx 2\pi\nu k + \varphi_k + v_p(k)|_{\mathrm{mod}2\pi} \tag{3.77}$$

$v_I(k)$, $v_Q(k)$ and $v_p(k)$ are regarded as the components of a vectorial zero mean white Gaussian measurement noise vector $\mathbf{v}_k$ with diagonal covariance

matrix $\mathbf{R}(k) = \mathrm{E}\{\mathbf{v}_k\mathbf{v}_k^T\} = \sigma^2\mathrm{diag}\{1, 1, \frac{1}{\mathbf{E}\{a^2(k)\}}\}$. Summarizing we have a vectorial nonlinear observation model

$$\begin{aligned} \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \\ \begin{bmatrix} y_I(k) \\ y_Q(k) \\ y_p(k) \end{bmatrix} &= \begin{bmatrix} a(k)\cos(\psi(k)) \\ a(k)\sin(\psi(k)) \\ \psi(k) \end{bmatrix} + \begin{bmatrix} v_I(k) \\ v_Q(k) \\ v_p(k) \end{bmatrix} \end{aligned} \tag{3.78}$$

where $\psi(k) = 2\pi\nu k + \theta$, $\mathbf{x}_k$ is the state vector, $\mathbf{h}(\cdot)$ is the nonlinear mapping between $\mathbf{y}_k$ and $\mathbf{x}_k$.

Note that on the first sight $y_p(k)$ appears redundant, as one might argue that any deterministic combination of the existing measurements does not provide any new (innovating) information. This would be true if we were dealing with linear models but for this nonlinear case, the linearizations ignore the information that would be present if higher order approximations of the nonlinear observation model were used. Hence the additional formulation of a nonlinear relationship will provide a different aspect of information contained in the other observations and thus make the filter operation more stable. Therefore, we can expect to improve the observability by introducing such an additional observation [46].

Kalman filtering algorithm

Once having found the state and observation model, the Kalman filter is readily formulated [46, 47]. The specific points of interest arise from the given nonlinear observation model. A simple but effective solution is the extended linearized Kalman filter which linearizes the nonlinear observation mapping around the prediction estimate, $\hat{\mathbf{x}}_k^-$. The filtering algorithm is

- Initialization

$$\begin{aligned} \hat{\mathbf{x}}_0^+ &= \mathrm{E}\{\hat{\mathbf{x}}_0\} = \bar{\mathbf{x}}_0 \\ \mathbf{P}^+(0) &= \mathrm{E}\{(\mathbf{x}_0 - \bar{\mathbf{x}}_0)(\mathbf{x}_0 - \bar{\mathbf{x}}_0)^{\mathrm{T}}\} = \mathbf{P}(0) \end{aligned}$$

- Prediction

$$\begin{aligned} \hat{\mathbf{x}}_{k+1}^- &= \mathbf{A}\hat{\mathbf{x}}_{k+1}^+ \\ \mathbf{P}^-(k+1) &= \mathbf{A}\mathbf{P}^+(k)\mathbf{A}^T + \mathbf{Q} \end{aligned}$$

- Correction

$$
\begin{aligned}
\mathbf{K}(k+1) &= \mathbf{P}^{-}(k+1)\mathbf{H}^{T}(k+1) \\
&\quad \cdot[\mathbf{H}(k+1)\mathbf{P}^{-}(k+1)\mathbf{H}^{T}(k+1)+\mathbf{R}]^{-1} \\
\mathbf{res}_{k+1} &= \mathbf{y}_{k+1} - \mathbf{h}(\hat{\mathbf{x}}^{-}_{k+1}) \\
\text{Do while } |\mathbf{res}_{k+1}| &> \pi : \\
\{\mathrm{res}_{3,k+1} &= \mathrm{res}_{3,k+1} - \mathrm{sing}(\mathrm{res}_{3,k+1}) \cdot 2\pi\} \\
\hat{\mathbf{x}}^{+}_{k+1} &= \hat{\mathbf{x}}^{-}_{k+1} + \mathbf{K}(k+1) \cdot \mathbf{res}_{k+1} \\
\mathbf{P}^{+}(k+1) &= [\mathbf{I} - \mathbf{K}(k+1)\mathbf{H}(k+1)]\mathbf{P}^{-}(k+1)
\end{aligned}
$$

The linearized observation matrix (Jacobian matrix) is given by

$$
\mathbf{H}(k) = \frac{\partial}{\partial \mathbf{x}_k}\mathbf{h}(\mathbf{x}_k)|_{\mathbf{x}_k=\hat{\mathbf{x}}^{-}_{k}} = \left[\begin{array}{cccc} \frac{\partial y_I(k)}{\partial x_{1,a}} & \frac{\partial y_I(k)}{\partial x_{2,a}} & \frac{\partial y_I(k)}{\partial x_{1,p}} & \frac{\partial y_I(k)}{\partial x_{2,p}} \\ \frac{\partial y_Q(k)}{\partial x_{1,a}} & \frac{\partial y_Q(k)}{\partial x_{2,a}} & \frac{\partial y_Q(k)}{\partial x_{1,p}} & \frac{\partial y_Q(k)}{\partial x_{2,p}} \\ \frac{\partial y_p(k)}{\partial x_{1,a}} & \frac{\partial y_p(k)}{\partial x_{2,a}} & \frac{\partial y_p(k)}{\partial x_{1,p}} & \frac{\partial y_p(k)}{\partial x_{2,p}} \end{array}\right]_{\mathbf{x}_k=\hat{\mathbf{x}}^{-}_{k}} . \tag{3.79}
$$

Stochastic observability

It is useful to consider stochastic observability of the modeling since it is closely related to the stability of the filter. Stochastic observability can be proven by considering Fisher's Information [47] defined as

$$
\mathbf{F}(k, k-L+1) = \sum_{i=k-L+1}^{k} \mathbf{A}^{T}(i,k)\mathbf{H}^{T}(i)\mathbf{R}^{-1}\mathbf{H}(i)\mathbf{A}(i,k). \tag{3.80}
$$

The parameter is said to be stochastically observable if the diagonal element of (3.80) increases monotonically. It is numerically shown in Fig. 3.5 that the EKF in which the instantaneous phase is incorporated as an additional measurement, acquires more information than the one without incorporating.

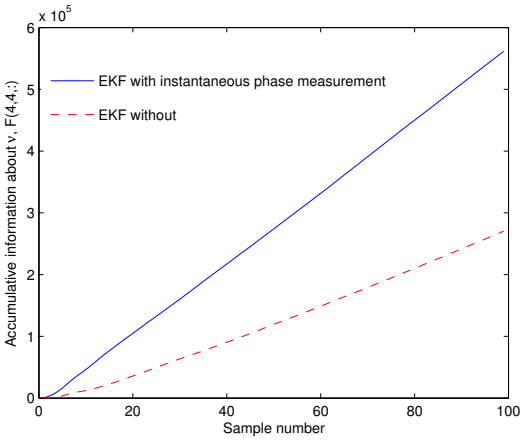

Figure 3.5: Accumulative information about ν at SNR = 10 dB, $N = 100$.

Alternatives of Kalman filters

Although EKF is a frequently used filtering technique, there exists other improved versions such as Unscented Kalman filter (UKF) [48] (see also in Appendix B). The UKF propagates a set of carefully chosen sample points (complectly capture the true mean and covariance of a Gaussian random variable) through the true nonlinear system in which the posterior mean and covariance are captured accurately to the 3rd order (Taylor series expansion) for any given nonlinearity. The EKF, in contrast, only achieves first-order accuracy. To summarize, the UKF effectively evaluates both Jacobian and Hessian[3] precisely through its sigma point propagation, without the need to perform any analytic differentiation. However, it has been observed in [49] that as the state-transition function is linear, the improvement is not obvious.

Frequency estimation examples of EKF-based CFO estimator

Two examples are chosen to show the capability of the proposed EKF frequency estimator. The proposed EKF-based CFO estimator is run in fast and slow fading scenarios over $N = 100$ samples of the received signal. A moderate SNR of 10 dB is assumed. The time-varying envelopes of the fast and slow fading are shown in Fig. 3.6 and Fig. 3.7, respectively. In both

[3]Hessian matrix is the square matrix of second-order partial derivatives of a function; that is, it describes the local curvature of a function of many variables. Hessian matrices are the coefficient of the quadratic term of a local Taylor expansion of a function. The full Hessian matrix is generally difficult to compute in practice.

cases, the phase can be successfully unwrapped via EKF algorithm. The corresponding frequency estimates in fast and slow fading are also presented in Fig. 3.6 and Fig. 3.7. However, the main attraction of EKF-based estimator is that it can be easily extended to track time-varying frequencies which would be relatively difficult in the case of batch processing.

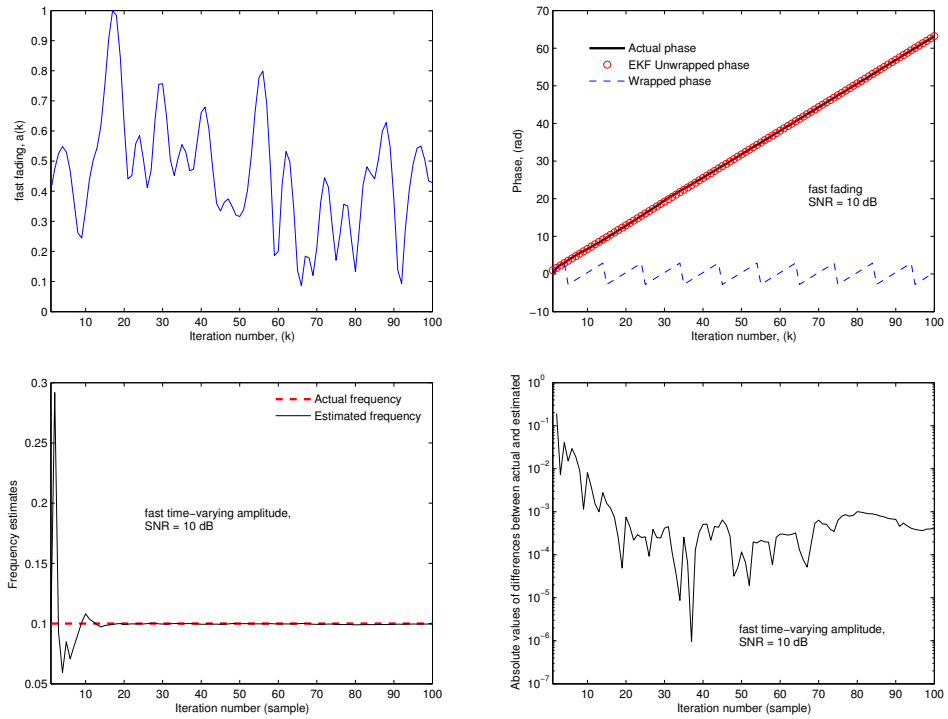

Figure 3.6: Frequency estimates in fast fading real-valued amplitude.

3.3 Simulation Results

3.3.1 Non-fading channel

In the non-fading channels, the estimation ranges and the mean-squares error (MSE) of the ML, Kay, Modified Kay, Fitz, Modified Fitz, proposed approximated ML, and Mengali estimators are shown. The advantages and disadvantages of these estimators can be seen from the simulation results

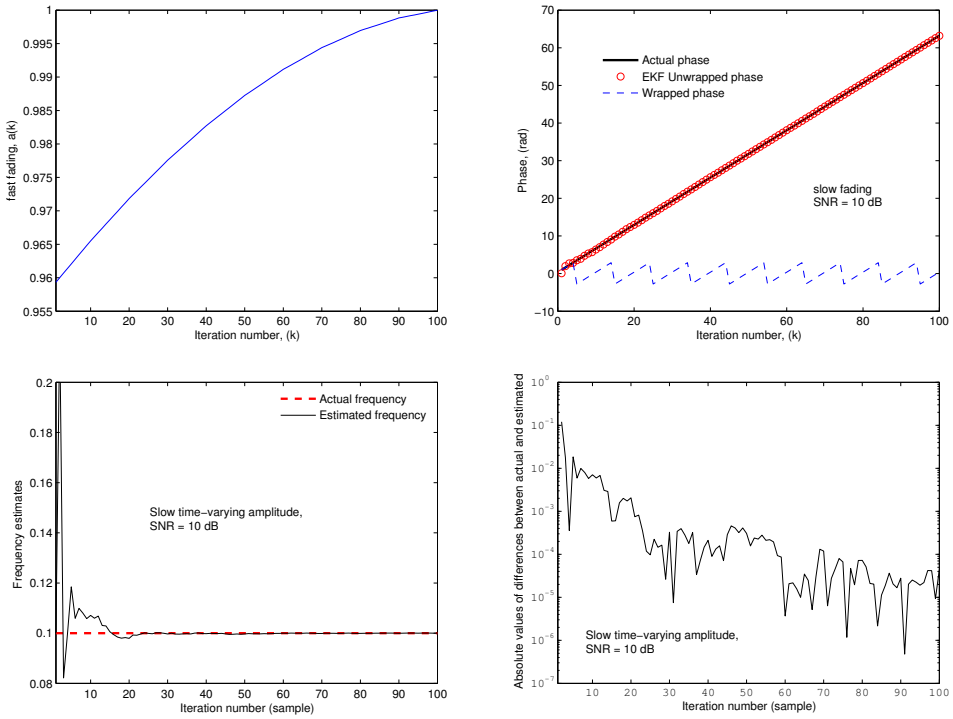

Figure 3.7: Frequency estimates in slow fading real-valued amplitude.

presented in Fig. 3.8 and Fig. 3.9. The data model of (3.1) with $N = 128$ is used in all simulations.

Although the ML estimator is attractive as it attains the CRLB even at a very low SNR and has the maximum estimation range of ± 0.5 the symbol rate, it's complexity is prohibitive. This can be observed in Fig. 3.9 that the required number of Fourier bin (N_f) followed by the interpolation is $N_f = 512$ for ML estimator to attain the CRLB. Without at additional interpolation method, the ML estimator requires at least $N_f = 2^{20} = 1048576$.

Kay estimator is much simpler than the ML estimator and has also the maximum estimation range of about ± 0.5 the symbol rate but its threshold is much higher than that of the ML estimator. This problem has been reduced by the modified Kay estimator in which the threshold can be improved as the length of the segment is increased. The accuracy of the modified Kay estimator reduces, but insignificantly. Moreover, the complexity of the modified Kay estimator reduces as the size of the segment increases.

Fitz estimator (with the optimal lag of $N/2$) attains the CRLB even at SNR as low as zero dB. However, it's estimation range with the optimal lag

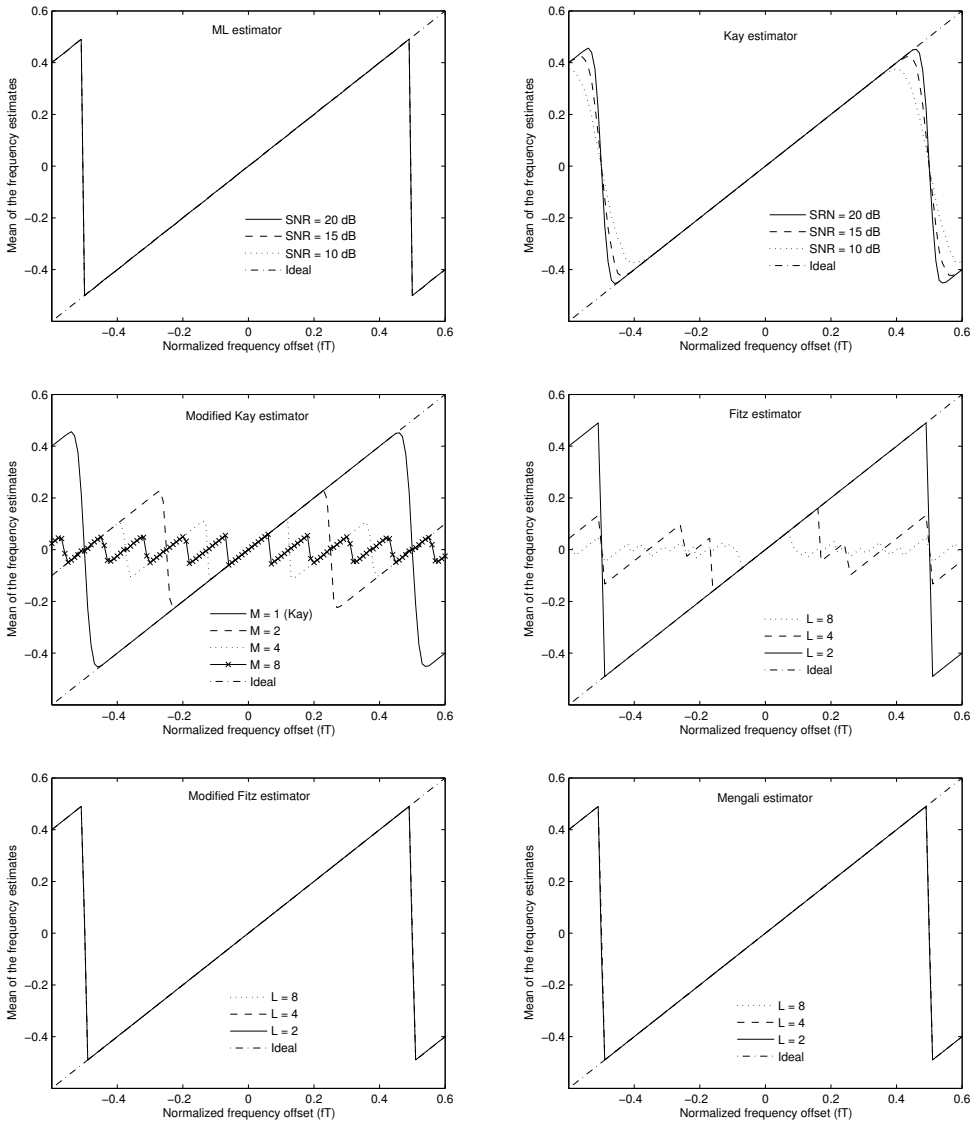

Figure 3.8: Estimation ranges of different CFO estimators for single carrier signal with constant amplitude, with $N = 128$, SNR = 20 dB.

($L = N/2$) is very small and limited to about $1/2L = 1/N$. To solve the phase ambiguity problem directly, an additional phase unwrapping algorithm [19] is required.

With the use of the summation-by-parts formula, the modified Fitz estimator is able to maintain a similar accuracy as Fitz estimator, but the estimation range is extended to about ± 0.5 the symbol rate. The accuracy

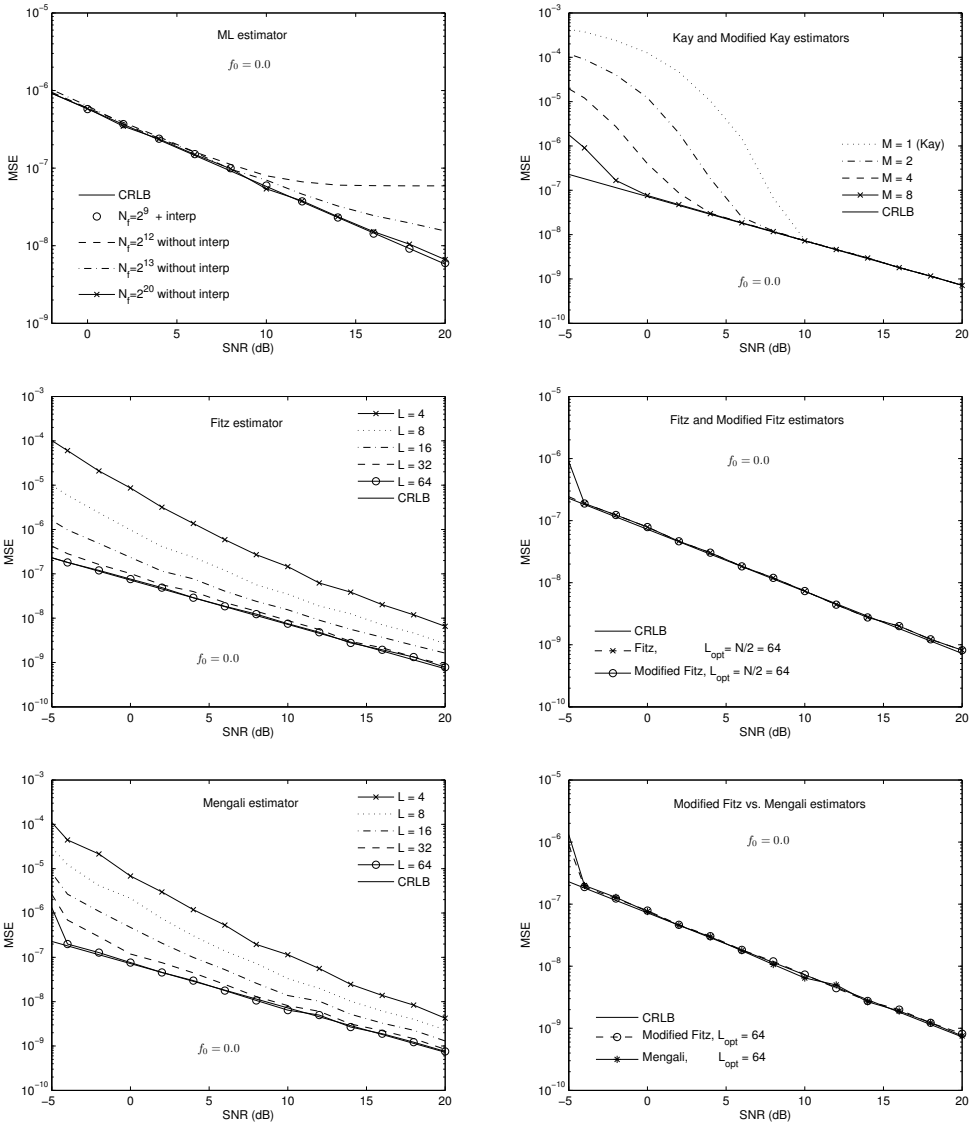

Figure 3.9: MSE versus SNR of different CFO estimators for single carrier signal with constant amplitude, with $N = 128$, $f_0 = 0.0$.

and the estimation range of the modified Fitz are comparable to the Mengali estimator, however the derivation of the modified Fitz is somewhat less complicated than the Mengali estimator.

Fig. 3.10 (left) compares the ratio between the estimation variance and the CRLB of Mengali, modified Fitz, transformed proposed estimators versus SNR, with $N = 64$, $M = M_{opt}$, and $\omega_o = 0.1 \times 2\pi$. It can be seen that the

efficiency of Mengali is better than that of the modified Fitz at high SNR but worse in lower SNR regions. The transformed proposed approximated ML estimator performs better than Mengali, and modified Fitz estimators especially at moderate/low SNRs.

Fig. 3.10 (right) compares the MSEs of the transformed proposed approximated ML and Mengali estimators versus SNRs with $N = 64$, $M = 4, 8, 16$, and $\omega_o = 0.1 \times 2\pi$. This is to show the performances when the complexity is more important and the accuracy can be relaxed. In all cases, the transformed proposed approximated ML performs better than Mengali estimator at moderate/low SNRs.

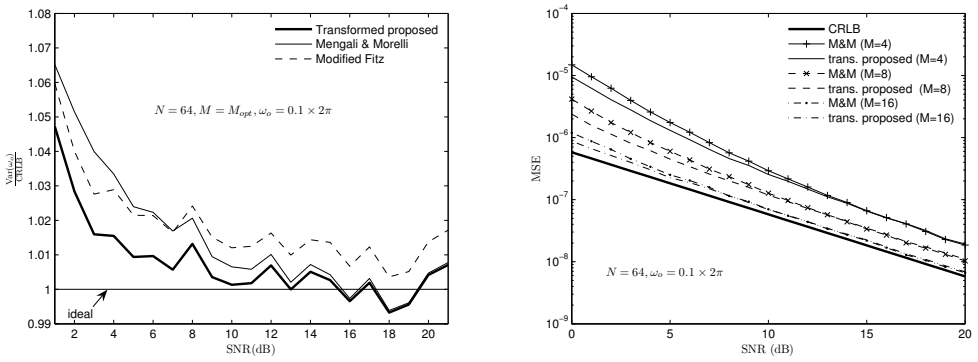

Figure 3.10: Performance comparisons of Mengali, transformed proposed approximated ML, modified Fitz estimators.

3.3.2 Fading channel

For simulation purposes, a reliable and widely accepted autoregressive-based method proposed in [40] is used to generate the realizations of the complex-valued multiplicative noise, $a(k)$, for any arbitrarily given autocorrelation model, $r_a(m)$, (e.g., Jakes, Sinc, Exponential, or Gaussian models).

The data sequence of length $N = 128$ were generated according to (3.49) with the correlation of the multiplicative noise that obeys an exponential-shaped correlation, i.e., $r_a(m) = \sigma_a^2 \exp\{-2\pi|m|B_D T\}$. The performance of the proposed estimator is compared to an efficient estimator proposed in [50] [41] which is known as the Single-Lag (SL) estimator and given as $\hat{\nu}_{\text{SL}} = \arg\{\hat{r}(m_{\text{opt}})\}/m_{\text{opt}}$, for $m_{\text{opt}} \neq 1$ the ambiguities of the frequency estimates are introduced and if not resolved the estimation range is limited to about $\pm 1/2m_{\text{opt}}$. We also compare our proposed ANLS estimator with the well

known CFO estimator introduced by Mengali and Morelli [26] (MM_{AWGN}) which was originally designed to be optimum for the AWGN channel. The mean-square-error (MSE) is obtained by 20000 Monte Carlo trials. In all simulations, the performances between ANLS and SNLS are similar, but the estimation range of SNLS is limited to $\pm 1/2M$ the symbol rate while ANLS can operate at the maximum possible range of about ± 0.5 the symbol rate.

Fig. 3.11 shows that in the non-fading channels the performances of the proposed ANLS and MM_{AWGN} estimators using $M = 127$ attain the CRLB even at SNR as low as zero dB, while the SL estimator, with the optimal lag of $m_{\text{opt}} = 2N/3$ (see [50] [41]) is only close to the CRLB. Moreover with the

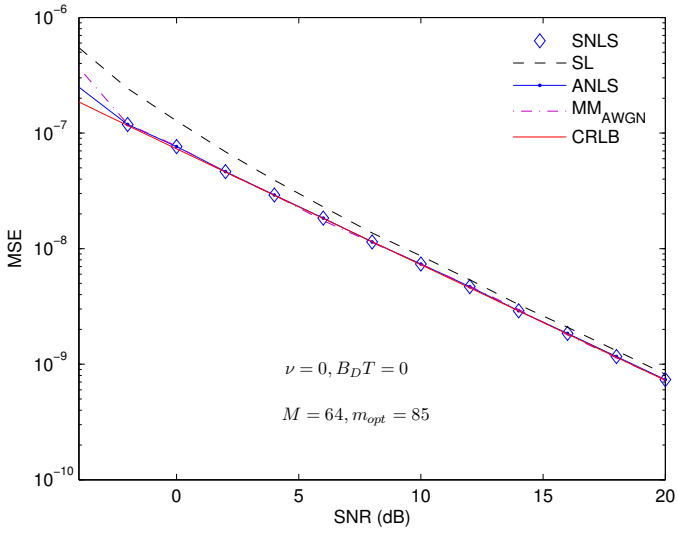

Figure 3.11: MSE versus SNR of the ANLS, MM_{AWGN}, SNLS and SL estimators for $\nu = 0$, $B_D T = 0$ (non-fading channel) with the optimal lag values.

optimal lag, the estimation range of SL estimator is limited to only about $\pm 1/2M$ the symbol rate as can be seen in Fig. 3.12. On the other hand, the estimation ranges of ANLS and MM_{AWGN} are shown to be independent of the lag number and covers about ± 0.5 the symbol rate.

Fig. 3.13 compares the proposed ANLS and MM_{AWGN} estimators in a slow fading channel ($B_D T = 0.0001$). The lag values (M) are set to 16 and 32 for both estimators. For both lag values, the proposed ANLS estimator performs better than the MM_{AWGN} estimator, especially in the low SNR region. However, their performances in the AWGN channel are similar as can be seen in Fig. 3.11.

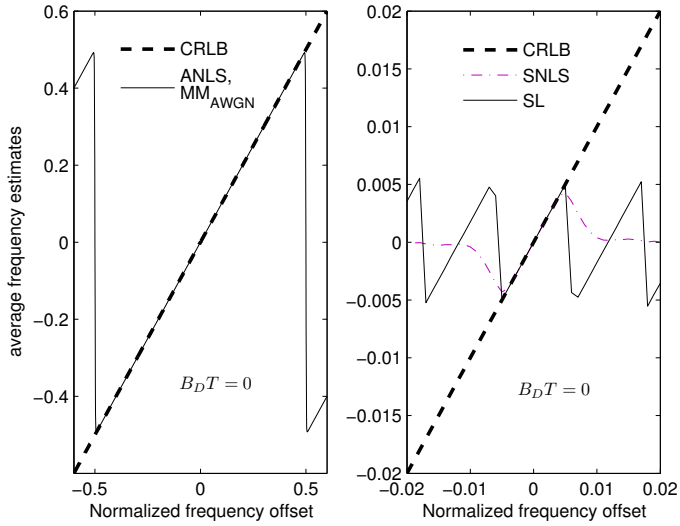

Figure 3.12: Estimation range of the ANLS, MM_{AWGN}, SL, and SNLS estimators for $B_DT = 0$ (non-fading), with SNR = 10 dB.

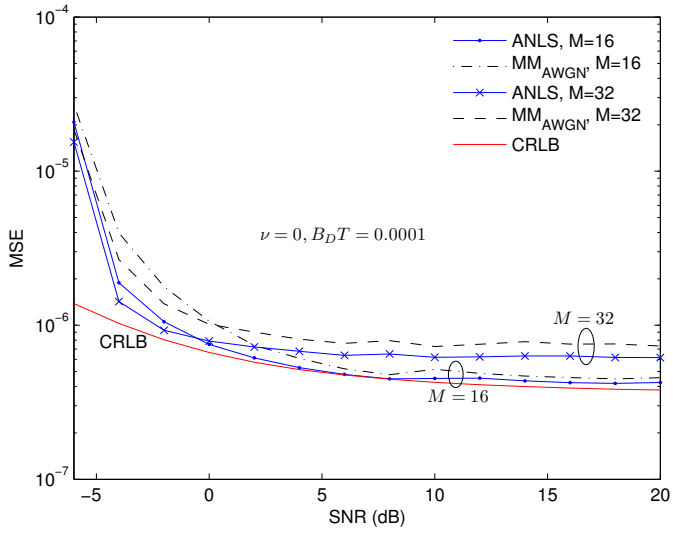

Figure 3.13: MSE versus SNR of the ANLS and MM_{AWGN} estimators for $\nu = 0$, $B_DT = 0.0001$ (slow fading channel) for lag values of M = 16 and 32.

Fig. 3.14 shows the optimal choice of the correlation lag for the SL, SNLS,

and ANLS estimators in fast and moderate fading. In moderate fading, the same optimal lag of 5 is observed for SL, SNLS, and ANLS. In fast fading, the optimal lag for SL is 2 and for SNLS and ANLS is 3. It is seen that the proposed ANLS estimator is less sensitive to the wrong choice of the lag numbers than the SL estimator.

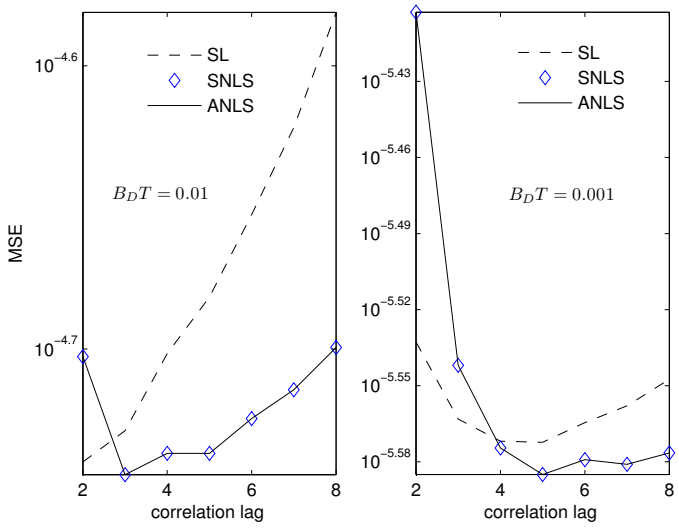

Figure 3.14: MSE versus correlation lag in fast ($B_DT = 0.01$) and moderate fading ($B_DT = 0.001$), with SNR = 10 dB, $\nu = 0.01$, $N = 128$.

Fig. 3.15 shows the estimation ranges of the SL, SNLS, ANLS estimators in moderate fading, however, we observed the same results also in fast fading. It is confirmed again that the estimation range of the proposed ANLS is independent of number of lag and covers the maximum possible range of about 50% the symbol rate, which is not achievable for SL, and SNLS without applying the additional phase unwrapping procedure.

Fig. 3.16 shows the performances of the SL, SNLS, and ANLS estimators in moderate and fast fading. The performances of ANLS and SNLS are slightly better than that of the SL for both scenarios. All attain the CRLB at high SNR. It can be observed that the estimation variance does not vanish as $\sigma_n^2 \to 0$. This behavior is known as the error floor effect which can be explained by (3.52). As the exponential model for the autocorrelation of the multiplicative noise is adopted, the closed-form expression of CRLB at high

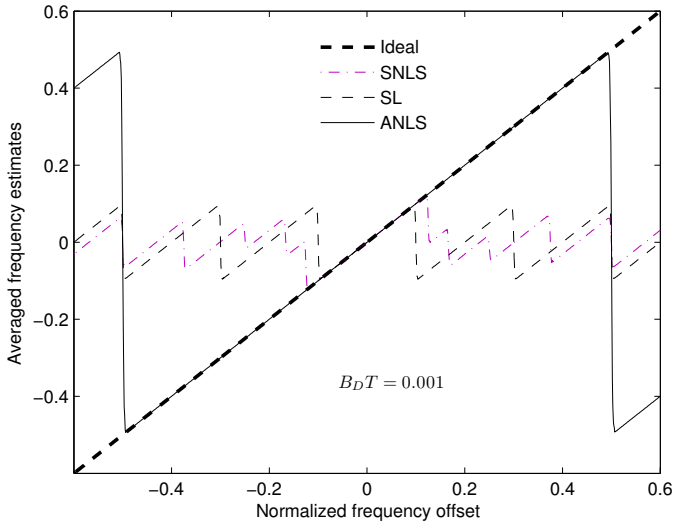

Figure 3.15: Estimation ranges of SL, SNLS, ANLS estimators for $B_DT = 0.001$ with $M = 5$, SNR = 10 dB, $N = 128$.

SNR [34] can be expressed as

$$\mathrm{CRLB}(\hat{\omega}_o)_{\sigma_n^2 \to 0} = \frac{1 - \exp\{-2\pi B_D T\}}{2(N-1)\exp\{-4\pi B_D T\}}. \tag{3.81}$$

Thus the potential accuracy of frequency estimation decreases with increasing Doppler spread.

3.4 Summary

For constant amplitude case, the modified Kay estimator is shown to be simpler and more robust than the original Kay estimator. The modified Fitz estimator uses the well known summation-by-parts formula (weight transformation formula) to overcome the phase ambiguity problem in Fitz estimator. At low SNRs, the performance of the modified Fitz estimator is slightly better than that of the Mengali estimator but worse at higher SNR regions. However, the modified Fitz estimator was derived with somewhat less effort than Mengali estimator. The proposed approximated ML estimator is trans-

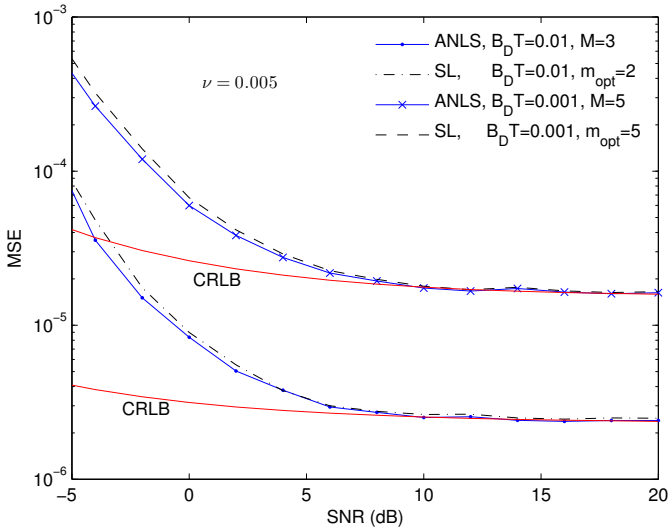

Figure 3.16: MSE versus SNR of ANLS and SL estimators for $B_DT = 0.01$ and $B_DT = 0.001$ with the optimal lag values.

formed using the weight transformation formula to obtain its phase difference version. The performance of the transformed proposed approximated ML estimator with optimal lag is better than that of Mengali estimator. Even using the same lags for low complexity, the transformed proposed approximated ML estimator exhibits better performance at intermediate/low SNRs.

For time-varying amplitude especially in the case of complex multiplicative noise, a new closed-form correlation-based frequency estimator for signals with fading-induced multiplicative noise is presented. It is shown that the proposed estimator can operate at the maximum estimation range of about ± 0.5 the symbol rate without sacrificing its accuracy and does not rely on any assumed form of the fading correlation as needed for most of the existing estimators. The features of the proposed ANLS estimator, as compared to the previously proposed estimators, are summarized in Table 3.2.

For real-valued time-varying amplitude, the extended Kalman filter-based recursive CFO estimator is proposed which can be easily extended for frequency tracking problems.

Estimator	Corr. form	Doppler spread	Unwrapping	Complexity
K&F [30]	required	assumed	required	high
M&M [31]	required	estimated	not needed	high
SL [50]	not needed	not needed	required	low
BS [32]	not needed	not needed	not needed	high
ANLS	not needed	not needed	not needed	low

Table 3.2: Features of different CFO estimators for flat fading channels.

Chapter 4

Frequency Offset Estimation: OFDM Case

OFDM systems are known to be very sensitive to the carrier frequency synchronization errors which lead to ICI and severely degrade the overall system performance [14]. Frequency offset estimation for OFDM systems may be classified as time-domain (pre-FFT) or frequency domain (post-FFT) approaches. The latter are usually used to estimate the integer part (coarse synchronization) of the carrier frequency offset after the fractional part has been identified and corrected. Time-domain methods are used to estimate the fractional part (fine synchronization), although some of these techniques can also estimate the integer part, this thesis focuses on the time-domain methods.

In literature on synchronization and channel estimation, it is typical to classify algorithms as pilot-based (or data-aided), blind, and semiblind methods. In pilot-based methods, a known pilot symbol is transmitted and channel parameters are estimated given the channel model and the known input. In blind techniques, the input is unknown, but some statistical properties such as independent and identically distributed (iid) inputs, may be known. The semiblind algorithms refer both to those that use both pilots and statistical properties of the unknown data and those that exploit additional features (such as finite alphabet) of the unknown symbol stream.

Some CFO estimation methods are classified as being data aided, although they do not use the known pilot block. This is the case for some estimation methods that are based on structuring the OFDM symbol as a

repetition of $L \geq 2$ identical slots [51–53]. These methods, in fact, belong to Null Subcarrier (NSC)-based techniques [54] since a repetitive of $L \geq 2$ identical slots can be generated by nulling all subcarriers whose frequencies are not multiples of $L\Delta f$. In [55], two identical OFDM symbols were used to estimate CFO. Even though this method is not NSC based, it can be mathematically described as such by considering the two symbols as a double-size OFDM symbol (i.e., $2M$ subcarriers) generated by setting the subcarriers with odd frequencies to zeros. These repetitive slots-based techniques require the number of NSCs to be larger than or equal to half the total number of subcarriers. Since the number of training zeros is large, these methods have been called data-aided, although no data are transmitted on those carriers.

In this chapter, the correlation-based with repetitive slots techniques for fine CFO estimation are discussed. In OFDM systems, pilots blocks are usually transmitted prior to the information frame. For instance, the IEEE 802.11 standard for WLANs employs series of identical slots in time domain as a preamble. Moose's [55], Schmidl and Cox's [51], and Morelli and Mengali's [52] estimators as well as numerous variants rely on repetitive pilot structure (see Fig. 4.1). These estimators are either limited in their estimation range or difficult to derive. The proposed estimator makes use of the nonlinear least-squares principle in conjunction with the well known summation-by-parts formula to obtain a new simple estimator with the maximum possible estimation range. With single receive and transmit antenna pair, the estimation variance exhibits the error floor effect which can be reduced by employing the receive antenna diversity concept.

4.1 OFDM Signal Model

As defined in the previous chapters, an OFDM signal is generated by taking the inverse FFT (IFFT) of a block of complex symbols $\{c_n\}$ belonging to a QAM or PSK constellation. The useful part of each block lasts T seconds and is preceded by a cyclic prefix, longer than the length of the channel impulse response, in order to avoid intersymbol interference (ISI). At the receiver, the baseband output of the matched filter after sampled at $T_s = T/N$ where N is the size of the IFFT, may be expressed as

$$y(k) = e^{j2\pi\nu k/N} s(k) + n(k) \tag{4.1}$$

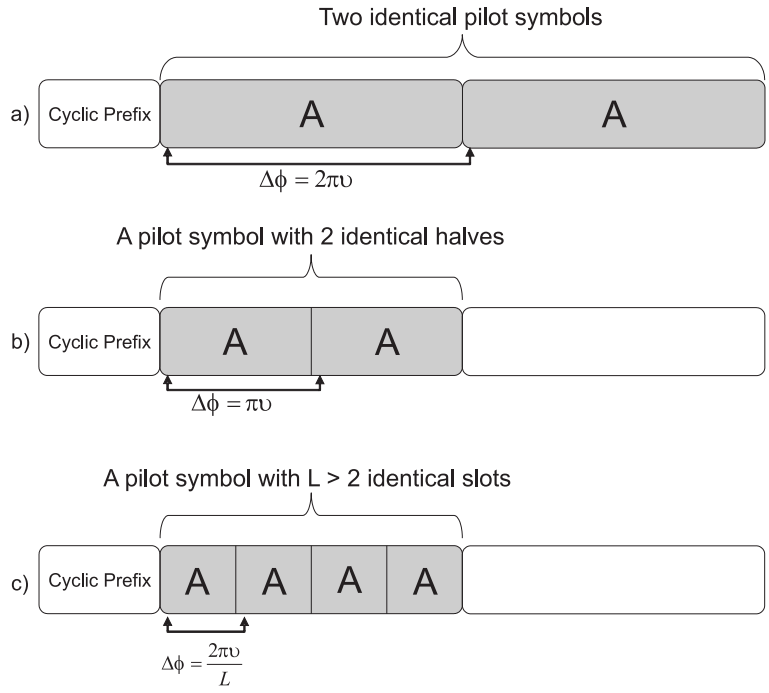

Figure 4.1: Three classical OFDM pilot structures proposed for frequency estimation, a). Moose's pilot structure, b). Schmidl and Cox's pilot structure, and c). Morelli and Mengali's pilot structure.

where $s(k)$ is the signal component, $n(k)$ is a zero-mean complex white Gaussian noise process with variance $\sigma_n^2 = \mathrm{E}\{|n(k)|^2\}$, and ν is the frequency offset normalized to $1/T$. Also assuming $2N_u + 1$ modulated subcarriers, the signal component is in the form

$$s(k) = \frac{1}{\sqrt{T}} \sum_{n=-N_u}^{N_u} c_n \mathcal{H}_n e^{j2\pi f_n k T_s}, \quad 0 \leq k \leq N-1 \tag{4.2}$$

where c_n is the pilot symbol, f_n is the frequency of the nth subcarrier and $\mathcal{H}_n$ is the complex channel response at $f = f_n$. The signal-to-noise ratio is defined as SNR $\triangleq \sigma_s^2/\sigma_n^2$ with $\sigma_s^2 \triangleq \mathrm{E}\{|s(k)|^2\}$.

4.2 Estimators with Repetitive Structure

4.2.1 Moose estimator

The seminal work by Moose [55] proposed to use two identical OFDM symbols without guard interval in between to perform CFO estimation. Moose estimator was originally derived in the frequency domain as

$$\hat{\nu}_{\text{Moose}} = \frac{1}{2\pi} \arg \left\{ \sum_{n \in \mathcal{K}} Y_{1,n}^* Y_{2,n} \right\} \tag{4.3}$$

where $\mathcal{K}$ is the index of the used subcarriers, $Y_{1,n}$ and $Y_{2,n}$ are the FFT outputs of the first and second N samples, respectively. It can be shown that Eq. (4.3) is equivalent to

$$\hat{\nu}_{\text{Moose}} = \frac{1}{2\pi} \arg \left\{ \sum_{k=0}^{N-1} y^*(k) y(k+N) \right\}. \tag{4.4}$$

When there is no timing offset,

$$\begin{aligned} y^*(k)y(k+N) &= \left[s(k)^* e^{-j\frac{2\pi\nu k}{N}} + n^*(k) \right] \left[s(k+N) e^{j\frac{2\pi\nu(k+N)}{N}} + n(k+N) \right] \\ &= |s(k)|^2 e^{j\frac{2\pi\nu}{N}} + \eta(k) \end{aligned} \tag{4.5}$$

where

$$\eta(k) = s(k)^* n(k+N) e^{-j\frac{2\pi\nu k}{N}} + s(k+N) n(k)^* e^{j\frac{2\pi\nu(k+N)}{N}} + n(k)^* n(k+N)$$

is the noise term.

Eq. (4.4) resolves the normalized CFO without any ambiguity in the range $\pm 1/2$, i.e., half of the subcarrier spacing of the repeated symbol. Otherwise, ambiguities are encountered by the $\arg\{\cdot\}$ exceeds $\pm\pi$. A straightforward solution to enlarge the estimation range is to use the shorter OFDM symbols, this is because the subcarrier distance is increased and thus the estimation range is extended. However, shortening the symbol interval results in the

reduced estimation accuracy.

It is shown in [55] that both time-domain (4.4) and frequency-domain (4.3) estimators form a ML estimator for AWGN channels. It should be noted that Eq. (4.4) is similar to the one proposed by Schmidl and Cox [51] explained in the next subsection.

4.2.2 Schmidl and Cox estimator

Two OFDM pilot symbols are also employed by Schmidl and Cox [51]. The first has two identical halves ($N/2$ samples each) and serves to measure the frequency offset with an ambiguity equal to the subcarrier spacing. The second symbol contains a pseudonoise (PN) sequence that and is used to resolve the ambiguity. The OFDM symbol with two identical halves is generated by transmitting PN sequence on the even subcarriers, while the odd ones are set to zeros.

Since the first and the second half of the pilot symbol are identical, $y(k)$ and $y(k+N/2)$ are ideally the same in that part. This facilitates the set up of two quantities, first quantity is defined as

$$\begin{aligned}\mathcal{P}(\tau) &= \sum_{k=0}^{N/2-1} y^*(\tau+k)y(\tau+k+N/2) \\ &= \mathcal{P}(\tau-1) + y^*(\tau+N/2-1)y(\tau+N-1) - y^*(\tau-1)y(\tau-1+N/2)\end{aligned} \tag{4.6}$$

where $\tau \in \mathbb{N}$ is the time index corresponding to the first sample in an observation window of N samples. This sum will be maximized for $\tau = \tau_{\text{opt}}$ corresponding to the first sample of the first OFDM pilot symbol since all elements will line up in-phase. If there exists a guard interval, it will also be maximized for all positions within the guard interval that are not effected by the time diversity of the channel. The received energy $\mathcal{R}(\tau)$ of the second half is also calculated. This quantity is defined as

$$\begin{aligned}\mathcal{R}(\tau) &= \sum_{k=0}^{N/2-1} |y(\tau+k+N/2)|^2 \\ &= \mathcal{R}(\tau-1) + |y(\tau+N-1)|^2 - |y(\tau-1+N/2)|^2 .\end{aligned} \tag{4.7}$$

The observation window slides over time as the receiver searches for the pilot block by looking for the two identical halves in the time-domain. The method proceeds in two steps. The timing offset is first determined as

$$\tau_{\text{opt}} = \arg \max_{\tau} \mathcal{M}(\tau), \quad \mathcal{M}(\tau) = \frac{|\mathcal{P}(\tau)|^2}{(\mathcal{R}(\tau))^2} \tag{4.8}$$

where $\mathcal{M}(\tau)$ is the timing criterion. Then assuming the timing offset τ_{opt} has been correctly estimated, there exists the phase difference of $\pi\nu$ between two symbol halves. Consequently, the CFO is estimated as

$$\hat{\epsilon} = \frac{1}{\pi} \arg\{\mathcal{P}(\tau_{\text{opt}})\}. \tag{4.9}$$

Symbol timing and carrier frequency synchronization are often performed jointly in the literature. With repetitive slots (REP) based estimation, the CFO is usually obtained as the argument of the timing criterion taken at the optimal timing instant. The estimation range of the Schmidl and Cox's algorithm (SCA) is ± 1 the subcarrier spacing. This range is doubled as compared to Moose's estimator, and allows estimating the fractional CFO.

However, there might remain an uncompensated frequency shift of $2g/T_s$ [Hz], where $g \in \mathcal{G} = \{-W, -W+1, \ldots, -1, 0, 1, \ldots, W-2, W-1\}$ with $W_{max} = N/4$, is an integer. To resolve the integer shift $k_\nu = 2g$, another pilot block is required. First, the fractional CFO is corrected before calculating the FFT for both symbols resulting in the frequency domain representations $\mathbf{Y}_1$ and $\mathbf{Y}_2$. The even subcarriers of OFDM symbols $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are then element-wise differentially demodulated. The result is compared to the pre-calculated PN sequence

$$V_n = \sqrt{2} \cdot \frac{X_{2,n}}{X_{1,n}}, \quad n \in \left\{ -\frac{N}{2}, -\frac{N}{2}+2, \ldots, -2, 2, \ldots, \frac{N}{2}-4, \frac{N}{2}-2 \right\} \tag{4.10}$$

resulting from the corresponding transmitted symbols $\mathbf{X}_1$ and $\mathbf{X}_2$. From these values a correlation

$$\hat{k}_\nu = 2 \arg \max_{g} \mathcal{B}(g), \quad \mathcal{B}(g) = \frac{|\sum_n Y^*_{1,n+2g} \cdot V^*_n \cdot Y_{2,n+2g}|^2}{2(\sum_n |Y_{2,n}|^2)^2} \tag{4.11}$$

that also partially eliminates possible influences from transmission channel, is calculated. The $\hat{g}$ that maximizes $\mathcal{B}(g)$ delivers the integer frequency offset, so that the total estimated frequency offset is $\hat{k}_\nu + \hat{\epsilon}$.

4.2.3 Morelli and Mengali estimator

Morelli and Mengali [52] proposed an extension of Schmidl and Cox algorithm (SCA) with a pilot symbol composed of $L > 2$ identical parts in the time domain. They are generated by transmitting a pilot sequence on the frequency multiples of L/T and setting zeros for the remaining frequencies (see Fig. 4.2). In the sequel, it is referred to as M&M estimator. The main benefit over the conventional SCA is an increase of the estimation range which becomes $\pm L/2$ the subcarrier spacing. The M&M exploits the correlation

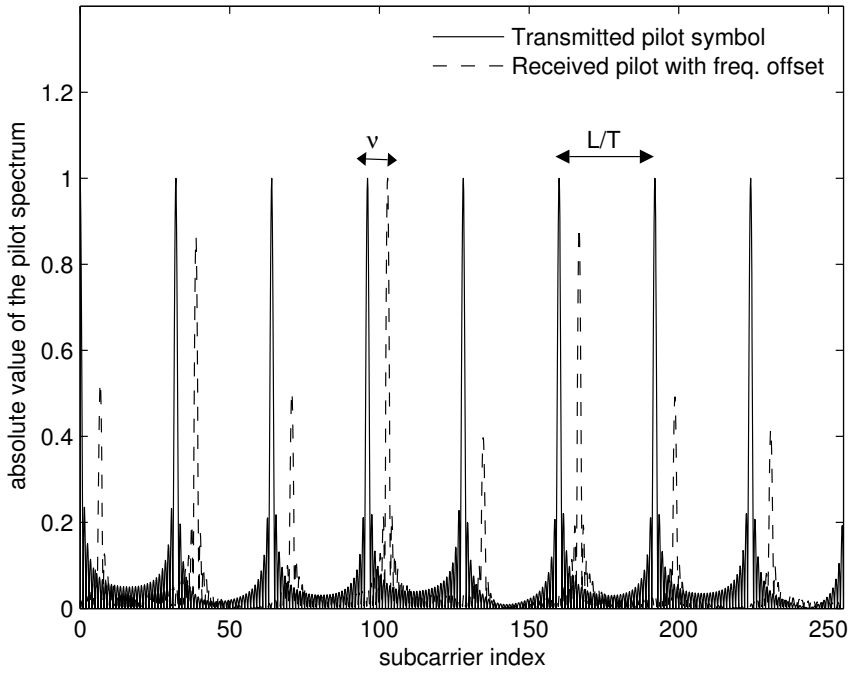

Figure 4.2: Frequency domain of the transmitted and received pilot symbol.

between identical slots of the pilot symbol defined as

$$r(m) = \frac{1}{N - mM} \sum_{k=mM}^{N-1} y(k)y^*(k - mM), \quad 0 \leq m \leq P - 1 \tag{4.12}$$

where $M = N/L$ is the length in sampling interval of each section of the training block and $P \in \mathbb{N}$ is a design parameter. Substituting (4.1) into (4.12) yields

$$r(m) = e^{j2\pi m\nu/L} d(m) + \gamma(m) \tag{4.13}$$

where $d(m) \triangleq \frac{1}{N-mM} \sum_{k=mM}^{N-1} |s(k)|^2$ is the real-valued random envelope, and $\gamma(m)$ is the noise term defined in [52].

By considering the phase differences of the correlation estimate

$$\begin{aligned} \varphi(m) &= \arg\{r(m)r^*(m-1)\}, \quad 1 \leq m \leq P - 1 \\ &\approx 2\pi\nu/L + \gamma_I(m) - \gamma_I(m-1) \end{aligned} \tag{4.14}$$

where $\gamma_I(m)$ is the imaginary part of $\gamma(m)$, the best linear unbias estimator (BLUE) of the CFO can be expressed as

$$\hat{\nu}_{\text{M\&M}} = \frac{L}{2\pi} \sum_{m=1}^{P-1} \omega_{\text{M\&M}}(m)\varphi(m) \tag{4.15}$$

where the weighting function $\omega_{\text{M\&M}}(m)$ is given by

$$\omega_{\text{M\&M}}(m) = 3\frac{(L-m)(L-m+1) - P(L-P)}{P(4P^2 - 6PL + 3L^2 - 1)}. \tag{4.16}$$

The direct derivation of $\omega_{\text{M\&M}}(m)$ using BLUE principle can be found in [35]. The weighting function (4.16) is channel independent. Thus the M&M is the BLUE for any channel. It is shown in [52] that the variance of the M&M estimator achieves its minimum when $P = L/2$. It is interesting to notice that M&M estimator reduces to SCA for $L = 2$ and $P = 1$. The M&M estimates have slightly lower variance than those of SCA. The price to be paid is an increase in computational complexity, as P angles have to be computed instead of one for the SCA.

4.2.4 Proposed nonlinear least-squares estimator

In the derivation of the proposed NLS estimator, we adopt the pilot symbol having a repetitive structure similar to that proposed in [52]. While [52] applies the BLUE concept to the phase differences of the estimated correlation samples, we use a different approach [6] that relies on the NLS principle in conjunction with the summation-by-parts formula. As a result, we obtain two new estimators, the first is the general form for fading channels and the other is its special case for AWGN channel.

Derivation of the generalized NLS estimator

The NLS estimates of $\{d(m)\}_{m=0}^{P-1}$ and ν are obtained by solving the following minimization problem

$$\{\hat{\mathbf{d}}, \hat{\nu}\} = \arg \min_{\mathbf{d},\ \nu} \frac{1}{P} \sum_{m=0}^{P-1} |r(m) - d(m)e^{j2\pi m\nu/L}|^2. \tag{4.17}$$

Let

$$\begin{aligned} \mathbf{r} &= [r(0), \cdots, r(P-1)]^T \\ \mathbf{d} &= [d(0), \cdots, d(P-1)]^T \\ \mathbf{W} &= \text{diag}\left(1, e^{j2\pi\nu/L}, \cdots, e^{j2\pi(P-1)\nu/L}\right) \end{aligned} \tag{4.18}$$

so that the inner term of (4.17) can be formulated as

$$J(\mathbf{d}, \nu) = \frac{1}{P}\|\mathbf{r} - \mathbf{W}\mathbf{d}\|^2 \tag{4.19}$$

where $\|\mathbf{z}\|^2 \triangleq \mathbf{z}^H\mathbf{z}$. Taking the derivative of $J(\mathbf{d}, \nu)$ with respect to $\mathbf{d}$, yields

$$\frac{\partial}{\partial \mathbf{d}} J(\mathbf{d}, \nu) = \frac{1}{P}\left(-\mathbf{W}\mathbf{r}^* - \mathbf{W}^H\mathbf{r} + 2\mathbf{d}\right) \tag{4.20}$$

where $(\cdot)^T$ and $(\cdot)^H$ denote transpose and complex conjugate transpose, respectively. Hence, for any given value of ν, the vector $\mathbf{d}$ that minimizes $\frac{\partial}{\partial \mathbf{d}} J(\mathbf{d}, \nu)$ is $\mathbf{d} = \frac{1}{2}\{\mathbf{W}\mathbf{r}^* + \mathbf{W}^H\mathbf{r}\}$, where $(\cdot)^*$ denotes complex conjugate.

Substituting $\mathbf{d}$ into $J(\mathbf{d}, \nu)$, we need to minimize

$$\begin{aligned}\tilde{J}(\nu) =& \frac{1}{P}\|\mathbf{r} - \frac{1}{2}\mathbf{W}^2\mathbf{r}^* - \frac{1}{2}\mathbf{r}\|^2 \\ =& \frac{1}{2P}\mathbf{r}^H\mathbf{r} - \frac{1}{2P}\mathrm{Re}\{\mathbf{r}^T\mathbf{W}^{2*}\mathbf{r}\}\end{aligned} \tag{4.21}$$

or, equivalently, to maximize the last term of (4.21). Hence, the NLS estimate of the parameter ν can be expressed as

$$\hat{\nu}_{\mathrm{NLS}} = \arg \max_{\nu} \mathrm{Re}\left\{\sum_{m=0}^{P-1} r^2(m) e^{-j4\pi m\nu/L}\right\} \tag{4.22}$$

which can be efficiently realized by means of Fast Fourier Transform (FFT) with zero padding. This can be done by $O(N_f \log_2 N_f)$ operations, where N_f is the number of Fourier bins.

Simplified NLS estimator

A simpler form of (4.22) is obtained by assuming the availability of the phase of the estimated correlation samples. Let $\phi_m = \arg\{r(m)\}$ denotes the phase of the correlation estimate. The NLS criterion (4.22) can then be rewritten as

$$\sum_{m=0}^{P-1} |r(m)|^2 \cos(2\phi_m - 4\pi m\nu/L). \tag{4.23}$$

Setting the derivative of this criterion with respect to ν to zero, we obtain $\sum_{m=0}^{P-1} m|r(m)|^2 \sin(2\phi_m - 4\pi m\nu/L) = 0$. Under the small error approximation, i.e., $\sin(2\phi_m - 4\pi m\nu/L) \simeq (2\phi_m - 4\pi m\nu/L)$, the simplified NLS (SNLS) estimator is obtained as

$$\hat{\nu}_{\mathrm{SNLS}} = \frac{L}{2\pi} \frac{\sum_{m=1}^{P-1} m|r(m)|^2\phi_m}{\sum_{m=1}^{P-1} m^2|r(m)|^2}. \tag{4.24}$$

Note that without applying a proper phase unwrapping algorithm, the estimation range of the SNLS is limited to $\pm L/2P$.

Approximated NLS estimator using summation-by-parts

To make the SNLS independent of any phase unwrapping procedure, the phase increments of the estimated correlation, $\Delta\phi_m = \arg\{\hat{r}(m)\hat{r}^*(m-1)\}$, shall be exploited instead of ϕ_m. For this purpose, the summation-by-parts formula is proposed. Consider the well known summation-by-parts formula

$$\sum_{m=0}^{P-1} \Delta b_m a_m = b_{P-1}a_{P-1} - b_{-1}a_0 - \sum_{m=1}^{P-1} b_{m-1}\Delta a_m \tag{4.25}$$

where $\Delta b_m = b_m - b_{m-1}$ and $\Delta a_m = a_m - a_{m-1}$. Inspecting (4.25), and the nominator term in (4.24) yields

$$\begin{aligned}\sum_{m=1}^{P-1} \underbrace{m|r(m)|^2}_{\Delta b_m} \cdot \underbrace{\arg\{r(m)\}}_{a_m} \\ = b_{P-1}\arg\{r(P-1)\} - \sum_{m=1}^{P-1} b_{m-1}\Delta\phi_m \\ = \sum_{m=1}^{P-1}(b_{P-1} - b_{m-1})\Delta\phi_m\end{aligned} \tag{4.26}$$

where $\Delta b_0 = 0$, $b_{-1} = 0$, $\arg\{r(P-1)\} = \sum_{m=1}^{P-1}\Delta\phi_m$. Since $b_{-1} = 0$ and $\Delta b_0 = 0$, one obtains:

$$b_m = \sum_{k=1}^{m} \Delta b_k = \sum_{k=1}^{m} k|r(k)|^2. \tag{4.27}$$

Inserting (4.27) into (4.26), and then into (4.24), yields

$$\hat{\nu}_{\mathrm{ANLS}} = \frac{L}{2\pi} \cdot \frac{\sum_{k=1}^{P-1} k|r(k)|^2 \sum_{m=1}^{P-1}\Delta\phi_m - \sum_{m=2}^{P-1}\sum_{k=1}^{m-1} k|r(k)|^2\Delta\phi_m}{\sum_{m=1}^{P-1} m^2|r(m)|^2}. \tag{4.28}$$

Eq. (4.28) can be seen as an approximated version of the NLS (ANLS), which does not require any phase unwrapping procedure as needed for the SNLS.

Interesting enough, for the AWGN channel, (4.28) reduces to

$$\hat{\nu}_{\text{AWGN}} = \frac{3L}{2\pi} \sum_{m=1}^{P-1} \frac{(P-1)P-(m-1)m}{P(2P-1)(P-1)} \Delta\phi_m. \tag{4.29}$$

Remark I: For the AWGN channel case, the SNLS estimator (4.24) is equivalent to the well known estimator proposed by Fitz [24] and if the summation-by-parts formula is applied to it, the resulting estimator yields a similar form as (4.29).

Remark II: Although the ANLS estimator (4.28) is derived based on a single pilot symbol with L identical parts, the ANLS (with few modifications) can also be used in the case where the multiple identical pilot symbols are employed.

4.3 Enhancement with Multiple Antennas

Recently, CFO estimation with multiple antennas has increasingly gained attention since the spatial diversity can be exploited in order to enhance the estimation performance [56–59]. The diversity combining techniques for the time-domain CFO problem (pre-FFT) have been explored in [56] for single-input multiple-output (SIMO) systems using the repeated OFDM symbols [55]. To utilize the information from all branches, [56] proposed to perform a nonlinear operation before summing up all the signals (see Fig. 4.3). This method was also used in the work of [57] and [58]. The accuracy of the CFO estimators using such a combining technique can be considerably enhanced. However, these estimators rely on the absolute phase of the correlation sum, $\phi_m = \arg\{r_s(m)\}$, and not on the phase difference, $\Delta\phi_m = \arg\{r_s(m)r_s^*(m-1)\}$, thus they have a limited estimation range. This range can be simply widened using shorter symbols but at the cost of reduced estimation accuracy. In this section, it will be shown that the performance of the ANLS estimator derived in the previous section can be enhanced with antenna diversity combining strategy proposed by [56].

4.3.1 Correlation sum observation

The estimated correlation of the i-th receive antenna is calculated as in (4.12). As the observation, the sum of $r_i(m)$ from all receive antennas can be ex-

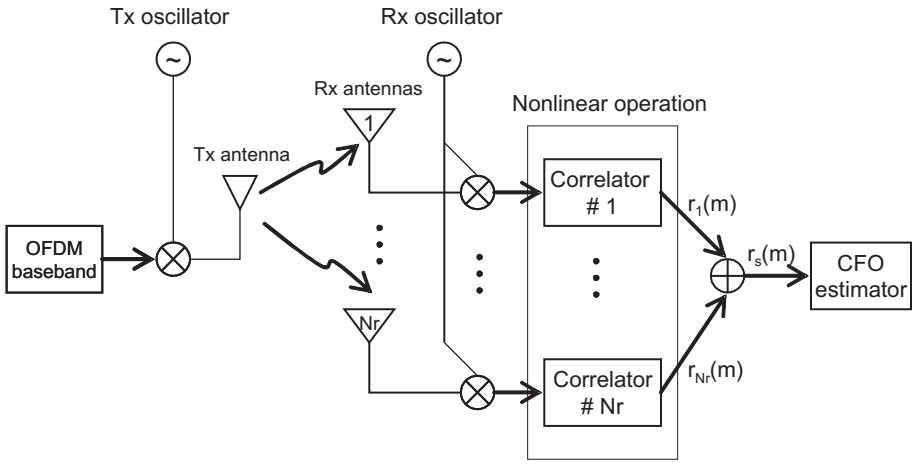

Figure 4.3: Configuration of CFO estimation with multiple receive antennas.

pressed as

$$\underbrace{\sum_{i=1}^{Nr} r_i(m)}_{r_s(m)} = e^{j\frac{2\pi m\nu}{L}} \underbrace{\sum_{i=1}^{Nr} d_i(m)}_{d_s(m)} + \underbrace{\sum_{i=1}^{Nr} \gamma_i(m)}_{\gamma_s(m)} \tag{4.30}$$

where Nr is the number of receive antennas, and $r_s(m) = |r_s(m)|e^{j\arg\{r_s(m)\}}$ is the sum of the estimated correlation. Without considering the probability density functions of $d_s(m)$ and $\gamma_s(m)$, two simple and straightforward estimators can be obtained either by performing the simple average (SAV) on $\phi_m = \arg\{r_s(m)\}$

$$\hat{\nu}_{\text{SAV}} = \frac{L}{2\pi P} \sum_{m=1}^{P-1} \frac{\arg\{r_s(m)\}}{m} \tag{4.31}$$

or applying the unweighted average (UAV) on $\Delta\phi_m = \arg\{r_s(m)r_s^*(m-1)\}$

$$\hat{\nu}_{\text{UAV}} = \frac{L}{2\pi(P-1)} \sum_{m=1}^{P-1} \arg\{r_s(m)r_s^*(m-1)\}. \tag{4.32}$$

The SAV estimator has an estimation range limited to $\pm L/2P$. This range can be extended by increasing L (or reducing M), but at the cost of the reduced accuracy. On the other hand, if the UAV is used the range can be enlarged. However, there is no guarantee regarding its estimation accuracy since the $\Delta\phi_m$ are not properly weighted.

The ANLS (4.28) and AWGN (4.29) estimators derived in the previous section, can be used in this context (see [6]) by simply replacing the $|r(m)|$ with $|r_s(m)|$ and the $\arg\{r(m)r^*m-1\}$ with $\arg\{r_s(m)r_s^*m-1\}$.

4.4 Simulation Results

We evaluate the proposed estimator in the frequency-selective channels and assume that time synchronization is perfectly achieved. The simulations are divided into two parts. The first set of simulations are carried out for OFDM with single transmit and single receive antenna in both quasi-static (static over one OFDM symbol) and time-varying channels. The second set of simulations is performed for single transmit and multiple receive antennas. In both set of simulations, the following common conditions are used:

1. Only one single OFDM pilot symbol, with repetitive slots ($L > 2$), is used.
2. The pilot symbols are taken from a QPSK complex constellation.
3. The useful part of an OFDM symbol contains 256 samples (N=256).
4. A cyclic prefix of 30 samples is used (chosen to be longer than the channel impulse response).
5. A channel bandwidth of 5 MHz and a carrier frequency of 1 GHz are assumed. The sampling interval $T_s = T/N$ is 0.2 μs.

4.4.1 SISO-OFDM

The frequency-selective channels for SISO-OFDM are set up such that the channel has 15 paths, with path delays of $0, 1, 2, \ldots, 14$ samples. The amplitude A_i of the i-th path varies independently of the others according to a Rayleigh distribution with exponential power delay profile, i.e., $E\{A_i^2\} = \exp\{-i/5\}$. The phase of each path is uniformly distributed on the interval $[0, 2\pi)$.

Fig. 4.4 illustrates the average estimate $E\{\hat{\nu}\}$ of different estimators as a function of the true offset ν for SNR = 20 dB. The parameters $P = 16$ and $L = 15$ are used. The ideal curve is given as a reference. The channel is static ($B_D T = 0$). It can be seen that the estimation range of the ANLS is about $\pm L/2$ the subcarrier spacing which is similar to that of the M&M

and NLS estimators. On the other hand, without phase unwrapping, the estimation range of SNLS is limited to ± 1. Similar results are obtained with time-varying channel $B_D T = 0.01$.

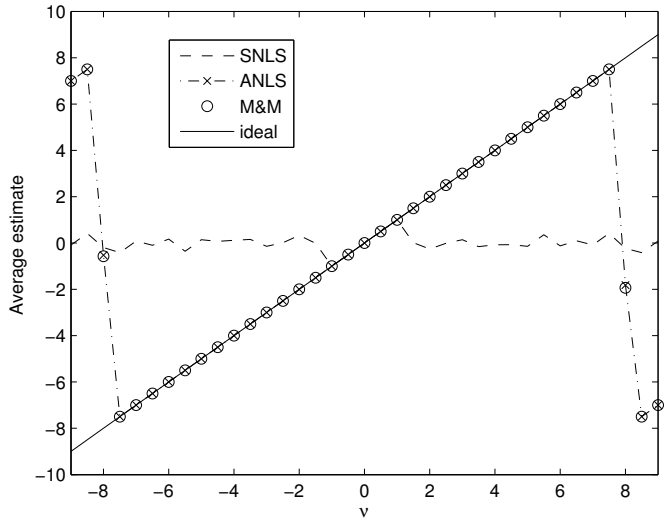

Figure 4.4: Average estimate $E\{\hat{\nu}\}$ vs. ν of the ANLS, SNLS, M&M estimators, SISO case, with L = 16, P = 15, SNR = 20 dB.

Fig. 4.5 shows the mean square estimation error (MSE) as a function of SNR. The carrier frequency offset is set to $\nu = 0.01$ such that it falls within the estimation range of SNLS which is limited to only $\pm L/2P$. L varies from 8, 16, and 32. The design parameter P is set to $P = L/2$ and $P = L - 1$ for each value of L. The CRLB for SISO case is given as

$$\text{CRLB}(\hat{\nu}) = \frac{1}{2\pi^2} \frac{3 \cdot \text{SNR}^{-1}}{N(1 - 1/N^2)} \tag{4.33}$$

for comparison. It can be seen that with the static channels the performances of the ANLS and M&M estimators are similar and close to the CRLB. The performance of both estimators increase as P and L increase. In time-varying scenarios, the curves exhibit floors due to the coherence loss between the segments of the pilot symbol. The performances tend to become worse as m increases. This implies that optimal value of m should be decreased with the Doppler spread of the fading channel.

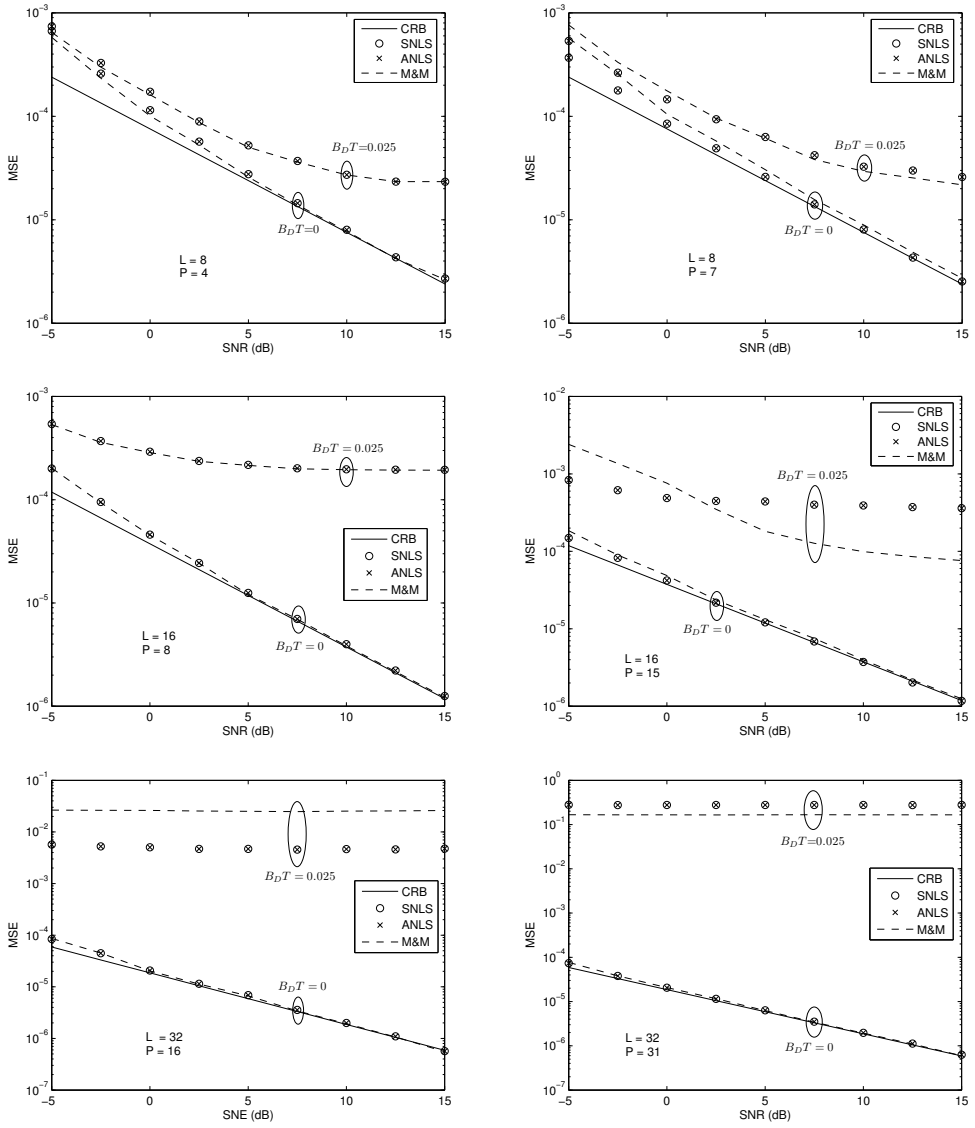

Figure 4.5: MSE vs. SNR of the ANLS, SNLS, M&M estimators, L = 16 and 32, with P = L/2 and L−1, for $\nu = 0.01$.

4.4.2 SIMO-OFDM

Fig. 4.6 illustrates the average estimate $E\{\hat{\nu}\}$ of different estimators as a function of the true offset ν for SNR = 20 dB and $B_DT = 0$ (static channel). The ideal curve is given as a reference. It is confirmed that the estimation ranges of those approaches which exploit the phase difference (ANLS, AWGN,

UAV) are of about $\pm L/2$ the subcarrier spacing. On the other hand, the estimation ranges of SNLS and SAV which use only the absolute phase, are limited to $\pm L/2P$. Similar results are also obtained for time-varying channels, $B_D T = 0.01$.

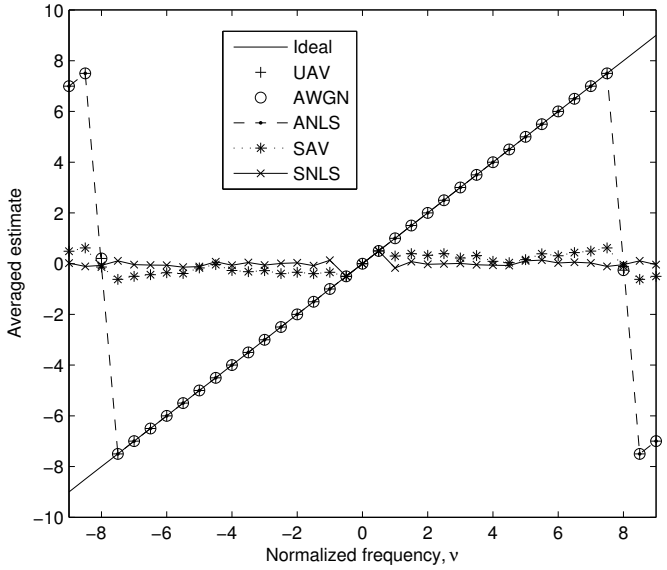

Figure 4.6: Average estimate $E\{\hat{\nu}\}$ vs. ν of ANLS, SNLS, AWGN, UAV, SAV estimators with Nr = 4, L = 16, P = 15, SNR = 20 dB.

The performances of the ANLS estimator with $Nr = 1$ to 4, for large CFO ($\nu = 7.3$ the subcarrier spacing) and $B_D T = 0$, are shown in Fig. 4.7. It can be seen that the performance is enhanced as the number of receive antennas is increased. Even for the high CFO value, the ANLS estimator attains the CRLB at an SNR as low as 0 dB. The performance in low SNR can be considerably improved by using a longer pilot sequence $L = 32$ and $P = 31$ (see Fig. 4.8).

In Fig. 4.9, the performances of the ANLS, SNLS, UAV, and SAV estimators with $Nr = 1$ and 4 for small CFO, $\nu = 0.01$ and $B_D T = 0$, are compared. Clearly in both cases, the UAV estimator which is not properly weighted, has the highest estimation variance. Because $\nu = 0.01$ is in the estimation range of SNLS, the ANLS and SNLS perform similarly. It is also shown that the performance of the ANLS estimator is better than that of the SAV, especially at low SNR. However, it should be noted that the estimation ranges of SNLS and SAV are limited to about $\pm L/2P$, while ANLS has the

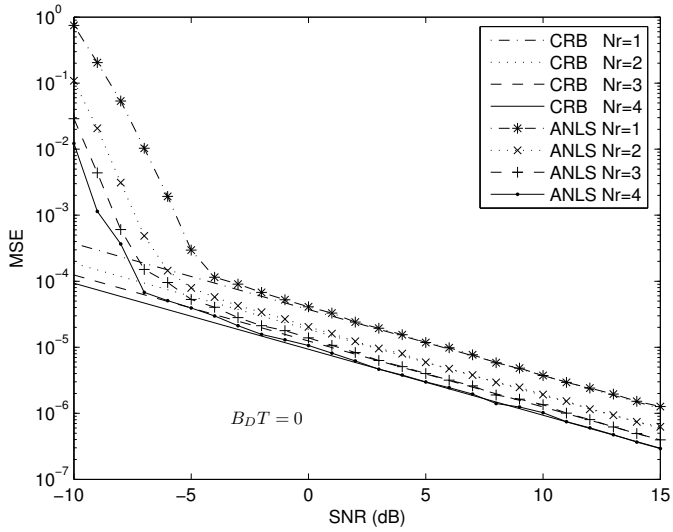

Figure 4.7: MSE vs. SNR of ANLS estimator with Nr = 1 to 4, L = 16, P = 15, and $\nu = 7.3$ ($B_DT = 0$).

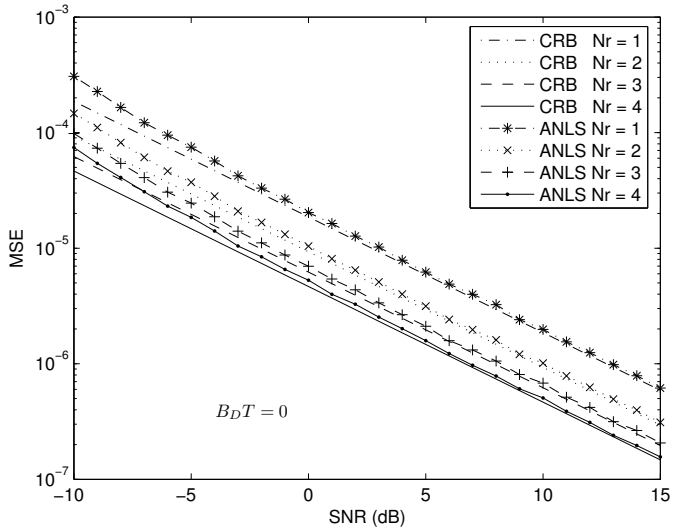

Figure 4.8: MSE vs. SNR of ANLS estimator with Nr = 1 to 4, L = 32, P = 31, and $\nu = 7.3$ ($B_DT = 0$).

maximum range of $\pm L/2$.

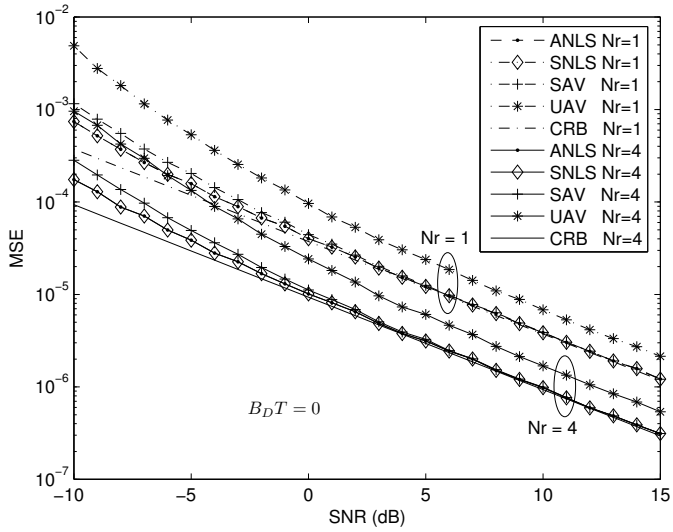

Figure 4.9: MSE vs. SNR of ANLS, SNLS, SAV, UAV estimators with Nr = 1 and 4, and $\nu = 0.01$ ($B_DT = 0$).

Fig. 4.10 shows the error floor effect caused by the Doppler spread of the ANLS and UAV estimators with $Nr = 1$ and 4 for $\nu = 7.3$. It can be observed that the error floor effect reduces as the number of receive antennas increases. However, the unweighted UAV still performs poorer than the proposed ANLS estimator. For $Nr = 4$, the estimation variance of the ANLS estimator is already close to the CRLB.

Fig. 4.11 compares the performances of the ANLS, SNLS, and SAV estimators in time-varying channel for small CFO, $\nu = 0.01$. Since $|\nu| < L/2P$, it can be seen that the performances of the ANLS and SNLS are similar and better than that of the SAV estimator especially at low SNR. Similar to the large CFO case (see Fig. 4.10), the error floor effect reduces as the number of receive antennas increases.

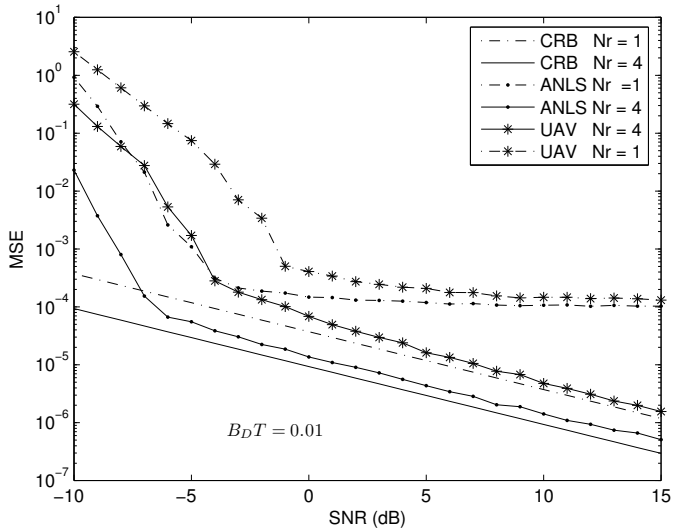

Figure 4.10: MSE vs. SNR of ANLS and UAV estimators with Nr = 1, 4 and $\nu = 7.3$ ($B_DT = 0.01$).

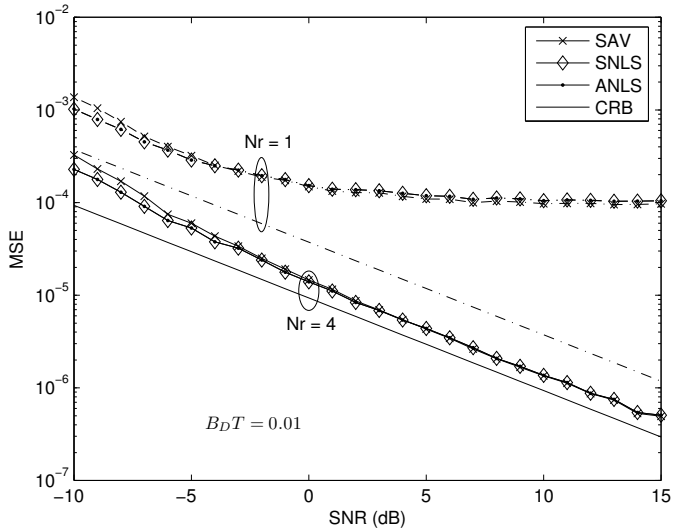

Figure 4.11: MSE vs. SNR of ANLS, SNLS, SAV estimators with Nr = 1, 4 and $\nu = 0.01$ ($B_DT = 0.01$).

4.5 Summary

The classical CFO estimators for OFDM signals have been discussed. The proposed estimator, which relies on the nonlinear least-squares principle in conjunction with the summation-by-parts formula, has shown to be much simpler to derive (as compared to M&M estimator) and has the maximum possible estimation range of about $\pm L/2$ without an additional phase unwrapping algorithm which is needed for most of the classical estimators. The estimation variance of the proposed estimator however suffers from the error floor effect caused mainly by Doppler spread. This effect can be reduced by introducing the receive antenna diversity concept.

Chapter 5

Frequency Estimation: Radar and Array Processing

This chapter relates the problem of frequency estimation in the synthetic aperture radar (SAR) and array processing contexts to the problem of carrier frequency offset estimation introduced in the previous chapters. It will be seen that the signal models for Doppler centroid estimation in SAR and the direction-of-arrival (DOA) estimation in array processing applications are similar to that of the classical one, hence most of the already existing including the proposed frequency estimation algorithms can be applied. A new concept of the integrated GPS/INS for DOA estimation is also introduced in this chapter.

5.1 Doppler Centroid Estimation: SAR

Estimation of the Doppler frequency is a fundamental operation in radar data processing. Weather radar and moving-target indication radar exploit the Doppler shift of each radar return to measure the velocity of scatterers. In the image formation algorithm for SAR, two-dimensional matched filtering (the range compression and azimuth compression) is the fundamental operation. This can be done either by two one-dimensional processing steps, Range-Doppler approach, or implementing a truly two-dimensional processing step usually implemented in frequency domain, where the matched fil-

tering performs the azimuth compression as well as the range migration correction.

For the range-Doppler approach, during the azimuth compression the range compressed data is correlated with the replica of the azimuth chirp, which depends on the phase history of the SAR sensor relative to the terrain being imaged during the interval of observation. The phase history can be approximated by a truncated Taylor series expansion which contains the first and second order derivatives of phase at the center of the synthetic aperture. For many applications, especially when the SAR carrier platform exhibits time varying attitude errors, the first order derivative, known as *Doppler centroid* frequency, f_{DC} must be estimated from the measurement data.

The Doppler centroid can also be seen as a measure of the effective antenna squint angle. It is used to adjust the bandpass characteristic of the azimuth compression filter to the location of the signal spectrum. An inaccurate Doppler centroid not only affects resolution and the signal-to-noise ratio, but also allows aliased azimuth frequency components to fall within the passband of the compression filter, and thus reduces the signal-to-ambiguity ratio [60, 61].

The Doppler centroid (the Doppler shift of a target positioned in the antenna boresight direction) is an important parameter in relation to azimuth (along track) SAR processing. For an airborne radar where the antenna is pointed perpendicular to the flight line, the Doppler centroid is ideally zero. However, if the antenna is off-set in angle (squinted), which might be due to desired or undesired yaw or if a satellite SAR orbiting a rotating earth is considered, then f_{DC} will be different from zero [62].

In principle, it is possible to calculate the Doppler centroid from orbit and attitude data, but measurement uncertainties on these parameters (primarily attitude) will limit the accuracy. These errors will eventually degrade the performance of the processing, especially w.r.t. signal-to-noise ratio, sidelobe, and ambiguity levels. Alternatively, Doppler centroid can be estimated from the received complex echo data. This has been done in the along-track dimension [61–63]. The time domain correlation-based Doppler centroid estimator has been used by Madsen [62] which was originally applied to spectral estimation in other field [64]. In [65, 66], various type of Kalman filters and smoothers were used to improve the Doppler centroid estimates which can be determined by correlation-based techniques.

If the azimuth data $y(k) = y(kT)$ is a stochastic process, its correlation function of two samples separated by m-lag is given by

$$r_y(m) = E\{y(k)y^*(k-m)\} \tag{5.1}$$

and the power spectrum is given by

$$S_y(f) = F\{r_y(m)\} = \sum_{m=-\infty}^{\infty} r(m)e^{-j2\pi mTf} \tag{5.2}$$

or

$$r_y(m) = F^{-1}\{S_y(f)\} = T\int_{-1/2T}^{1/2T} S(f)e^{j2\pi mTf}df. \tag{5.3}$$

If the azimuth power spectrum is shifted by f_{DC}, we have

$$S(f) = S_y(f - f_{DC}). \tag{5.4}$$

For frequency domain approaches, this shift is estimated directly in the azimuth frequency domain from an estimated azimuth power spectrum. However, for the time domain interpretation the corresponding correlation function can be written as

$$r(m) = e^{j2\pi mTf_{DC}} \cdot r_y(m). \tag{5.5}$$

This result is the basis of the time-domain estimation of the Doppler centroid. It is seen that the phase of the correlation function is directly related to the Doppler centroid. Therefore, one can by estimation of the correlation function $r(m)$, arrive at an estimate of the Doppler centroid, which can be calculated using either unbiased (3.56) or biased (3.57) form of the correlation estimates. Madsen used azimuth window size of $m = 1$ as the primary Doppler centroid estimator.

Notice that (5.5) has a similar form as (3.55), and hence the nonlinear least-squares concept (4.17)-(4.24) followed by the summation-by-parts formula can be again directly applied. This yields an estimator of a similar form as (3.69). The CRLB can be obtained in the same way as derived in the Appendix (A.2.1).

5.2 DOA Estimation: Array Processing

Consider the array of M identical sensors uniformly spaced on a line, depicted in Fig. 5.1. Such an array is commonly referred as a uniform linear array

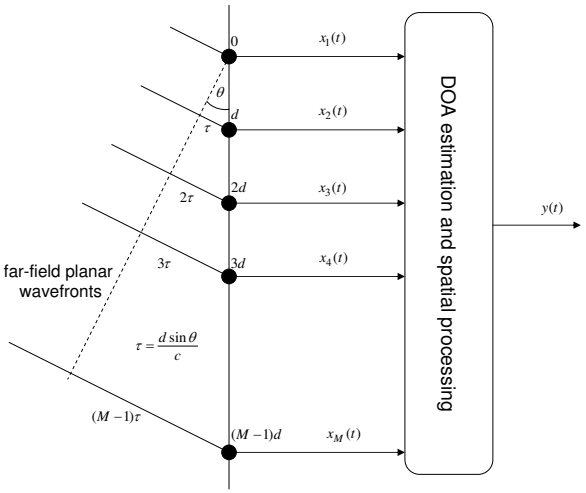

Figure 5.1: Basic geometry of a uniform linear array (ULA) of a single source.

(ULA). Let d denote the distance between two consecutive sensors, and let θ denotes the direction of arrival (DOA) of the narrowband signal illuminating the array, as measured with respect to the normal line of sensors. Then under the planar wave hypothesis (far-field), no angular spread, and the assumption that the first sensor in the array is chosen as the reference point, the array's output vector may be written as

$$\mathbf{x}(t) = \mathbf{a}(\theta)s(t) + \mathbf{n}(t) \tag{5.6}$$

where $\mathbf{x}(t) = [x_1(t), \cdots, x_M(t)]^T$, $\mathbf{a}(\theta) = [1, e^{-j\omega_c\tau_2}, \cdots, e^{-j\omega_c\tau_M}]^T$, and $\mathbf{n}(t) = [n_1(t), \cdots, n_M(t)]^T$. The delay of the wavefront to the m-th sensor is

$$\tau_m = (m-1)\frac{d\sin\theta}{c}, \quad \text{for } \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \tag{5.7}$$

where c is the speed of light, and m is the sensor's index. Inserting (5.7) into $\mathbf{a}(\theta)$ results in

$$\mathbf{a}(\theta) = \left[1, e^{-j\omega_c d\sin\theta/c}, \cdots, e^{-j(M-1)\omega_c d\sin\theta/c}\right]^T. \tag{5.8}$$

The restriction of θ to lie in the interval $[-\pi/2, \pi/2]$ is a limitation of ULAs: two sources at locations symmetric with respect to the array line yield iden-

tical sets of delays and hence cannot be distinguished from another. In practice, this ambiguity is eliminated by using sensors that only pass signals whose DOAs are in $[-\pi/2, \pi/2]$.

Let λ denote the signal wavelength

$$\lambda = c/f_c, \qquad f_c = \omega_c/2\pi. \tag{5.9}$$

By defining

$$f_s = f_c \frac{d \sin\theta}{c} = \frac{d \sin\theta}{\lambda} \tag{5.10}$$

and

$$\omega_s = 2\pi f_s = \omega_c \frac{d \sin\theta}{c} \tag{5.11}$$

yields

$$\mathbf{a}(\theta) = \left[1, e^{-j\omega_s}, \cdots, e^{-j(M-1)\omega_s}\right]^T \tag{5.12}$$

where ω_s is known as the *spatial frequency*. Note that $\mathbf{a}(\theta)$ is a Vandermonde vector which is completely analogous with the vector made from the uniform samples of the sinusoidal signal $\{e^{-j\omega_s t}\}$. The vector $\mathbf{a}(\theta)$ can be uniquely defined (i.e., there is no spatial aliasing) if and only if ω_s is constrained as

$$|\omega_s| \leq \pi \quad \text{or} \quad |f_s| \leq \frac{1}{2} \Leftrightarrow d|\sin\theta| \leq \frac{\lambda}{2}. \tag{5.13}$$

This condition is satisfied if $d \leq \lambda/2$, which implies that the spatial sampling period d should be smaller than half of the signal wavelength.

The extension of (5.6) to the case of multiple sources is straightforward. Since the sensors in the array were assumed to be linear elements, a direct application of the superposition principle leads to the following model of the array

$$\begin{aligned} \mathbf{x}(t) &= \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \\ &= [\mathbf{a}(\theta_1), \cdots, \mathbf{a}(\theta_K)] \begin{bmatrix} s_1(t) \\ \vdots \\ s_K(t) \end{bmatrix} + \mathbf{n}(t) \end{aligned} \tag{5.14}$$

where θ_k is the DOA of the k-th source, and $s_k(t)$ is the signal corresponding to the k-th source.

Eq. (5.6) and (5.14) are similar to the models used in the non-fading sinusoidal signal cases, therefore most of the (temporal) frequency estimation methods for sinusoidal signals can be used for the DOA estimation problem.

In practice, estimating the DOA in the presence of local scattering is more complicated. This is often seen in wireless mobile communications since the sources may appear to be spatially distributed due to local scattering around the mobile unit. Local scattering at the source causes signals from a single source to arrive via different paths and at different angles. The point source is now said to exhibit *angular spread*. Angular spread manifests itself via loss of spatial coherence of the signal across the receiver array. Hence, this scenario is also referred to as imperfect spatial coherence [67]. Imperfect spatial coherence is also encountered in sonar systems when the propagation medium is randomly inhomogeneous [67].

Model formulation

In this section, a single scattered source model for DOA estimation problem is developed. The earlier works for single scattered source can be found in [68–70] and [71] for multiple scattered sources. In order to obtain a proper model, it is assumed that the delay spread caused by the multipath propagation is small compared to the inverse bandwidth of the signal so that the narrowband assumption still holds even in the presence of scattering. The noise free narrowband signal arriving at the m-th sensor of an M-element ULA can be written as

$$x_m(t) = \frac{s(t)}{\sqrt{L}(t)} \sum_{i=1}^{L(t)} g_i(t) \exp\left\{-j2\pi\Delta(m-1)\sin(\theta+\phi_i(t))\right\} \tag{5.15}$$

where $L(t)$ is the number of scattered rays, Δ is the sensor spacing in wavelengths, θ is the angle to be estimated, $g_i(t)$ is the complex amplitude of the i-th arrival, and $\phi_i(t)$ is its angular spread. If the angular spread is small (the scattering occurs near the source and the array is in the far-field), $\sin(\theta+\phi_i(t)) \approx \sin(\theta) + \cos(\theta)\phi_i(t)$. Hence, the noisy array output can be approximated by

$$x_m(t) = s(t)u_m(t)\exp\left\{-j2\pi\Delta(m-1)\sin\theta\right\} + n_m(t) \tag{5.16}$$

where $n_m(t)$ is the additive noise at the k-th sensor, and $u_m(t)$ models the perturbations due to the spatial scattering, and is given by

$$u_m(t) = \mu + \sum_n g_n(t) \exp\{j2\pi\Delta(m-1)\cos(\theta)\phi_n(t)\} \tag{5.17}$$

where μ denotes the complex gain of the dominant path (at the nominal angle), if it is exist. The following assumptions are made in order to obtain a proper model

AS1) The amplitudes, $g_n(t)$, and the angular deviation, $\phi_n(t)$, are iid in both n and t, i.e., spatially and temporally white, zero-mean and mutually independent.

AS2) The angular deviations are stationary and symmetrically distributed around zero.

AS3) The narrowband signal $s(t)$ is deterministic and unknown.

AS4) Additive noise is temporally and spatially iid, zero-mean complex Gaussian and circular.

Under assumptions AS1)-AS2) and assuming that the number of rays, $L(t)$, is large, $[u_1(t), \cdots, u_M(t)]$ can be modeled (using central limit theorem) as a Gaussian M-variate process. The spatial correlation is given by

$$R_u(k,l) = E\{u_k(t)u_l^*(t)\} = |\mu|^2 + \sigma_g^2 R_\phi(k,l) \tag{5.18}$$

where σ_g^2 is the power of the scattered paths and

$$R_\phi(k,l) = E\{\exp(j2\pi\Delta(k-l)\cos(\theta)\phi)\}. \tag{5.19}$$

The covariance of the array outputs is stationary in space but not in time (since $s(t)$ is not assumed to be a constant modulus), and is given by

$$\begin{aligned} R_x(k,l;t) &= E\{x_k(t)x_l^*(t)\} \\ &= |s(t)|^2[|\mu|^2 + \sigma_g^2 R_\phi(k,l)]e^{j\omega_o(k-l)} + \sigma_n^2\delta(k-l) \end{aligned} \tag{5.20}$$

where $\omega_o = 2\pi\Delta\sin\theta$ is the nominal frequency, and σ_n^2 is the variance of the additive noise. Under AS2), $R_\phi(k,l)$ is real-valued.

For a Rayleigh channel ($\mu = 0$), the correlation model reduces to

$$R_x(k,l;t) = |s(t)|^2 \sigma_g^2 R_\phi(k,l) e^{j\omega_o(k-l)} + \sigma_n^2 \delta(k-l). \tag{5.21}$$

If ϕ is uniformly distributed, then $R_\phi(k,l) = \mathrm{sinc}(2\pi\Delta(k-l)\cos\theta)$, if ϕ has Gaussian distribution, then $R_\phi(k,l) = \exp(-[2\pi\Delta(k-l)\cos(\theta)\sigma_\phi]^2/2)$.

The averaged (over N the snapshots) spatial correlation matrix is estimated as

$$\hat{R}_x(k,l) = \frac{1}{N}\sum_{t=1}^{N} x_k(t) x_l^*(t), \quad k,l = 1,\ldots,M. \tag{5.22}$$

Using the following statistic yields

$$\hat{r}(m) = \frac{1}{M-m}\sum_{k=1}^{M-m} \hat{R}_x(k+m,k), \quad m = 0,\ldots,M-1. \tag{5.23}$$

These correlation coefficients are thus obtained after applying the so-called redundancy averaging technique to the covariance matrix estimate. It has been shown in [71] that $\hat{r}(m)$ converges in mean-square to

$$r(m) = \underbrace{P \cdot r_u(m)}_{d(m)} e^{j\omega_o m} + \sigma_n^2 \delta(m) \tag{5.24}$$

where $P = \frac{1}{N}\sum_{t=1}^{N} |s(t)|^2$, and $r_u(m) \triangleq r_u(k,l)$.

Notice that (5.24) has a similar form as (3.55), and hence the nonlinear least-squares concept (4.17)-(4.24) followed by the summation-by-parts formula can be again directly applied. This yields an estimator of a similar form as (3.69). The CRLB can be obtained in the same way as derived in the Appendix (A.2.1).

5.3 DOA Tracker: GPS/INS Integration

In this section, a new concept for DOA tracking via GPS/INS integration is introduced. Typically, algorithms for tracking of the DOA for a particular satellite shall properly handle the interferences from other satellites, which can be complicated. Ignoring these interferences will lead to a significant

DOA estimation error. The proposed concept however, can cope with this problem by incorporating the additional DOA information obtained from the inertial-based navigation unit. This concept can either be seen as a deeply-integration since the inphase and quadrature phase (I,Q) are processed, or as a loosely-integration since it has a decentralized integration architecture (see Fig. 5.2).

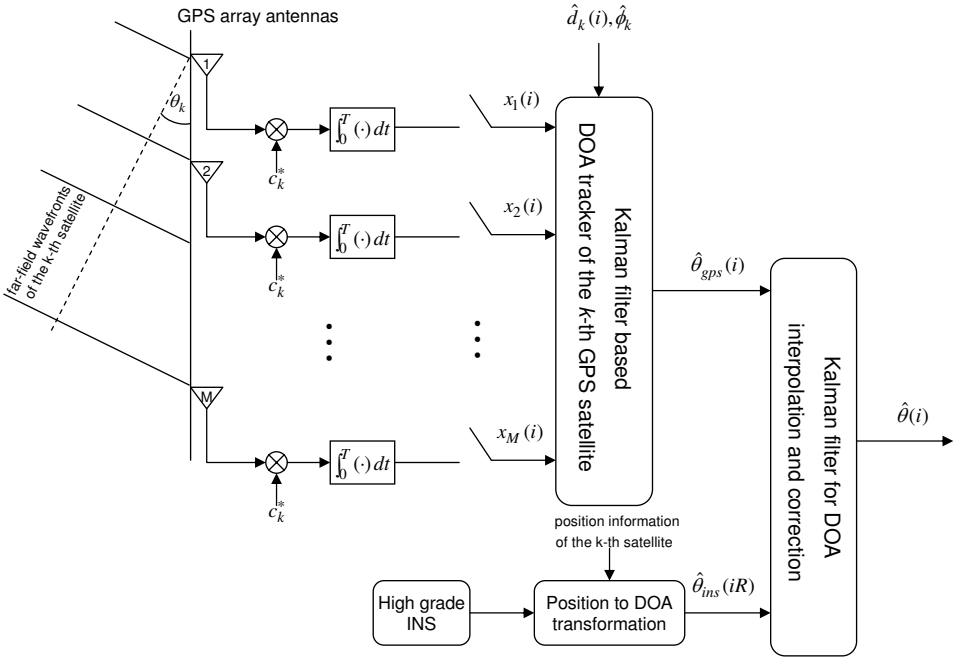

Figure 5.2: A simple GPS/INS integration architecture for DOA tracking.

5.3.1 Signal model

Consider a uniform linear array (ULA) with an M-element array with half-wavelength ($d = \lambda/2$) spaced antenna elements. Thus, the received signal at the m-th antenna may be expressed as

$$y_m(t) = \sum_{k=1}^{K} s_k(t) \exp\{j(m-1)\pi \sin\theta_k(t) + \phi_k\} + n_m(t) \tag{5.25}$$

where K is the number of sources (satellites), $\theta_k(t)$ is the azimuthal DOA of the k-th sources, ϕ_k is the carrier phase which is uniformly distributed in $[0, 2\pi]$, $s_k(t) = \sum_{i=-\infty}^{\infty} d_k(i)c_k(t-iT-\tau_k)$, $d_k(t)$ is the transmitted data, $c_k(t)$ denotes the spreading sequence waveform, $n_k(t)$ is the additive white Gaussian noise, τ and T are the time delay and the symbol interval, respectively. The received signal may be written in a vector form as

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta}(t))\mathbf{S}(t) + \mathbf{n}(t) \tag{5.26}$$

where

- $\mathbf{y}(t) = [y_1(t), y_2(t), \cdots, y_M(t)]^T$ is an $M \times 1$ vector of the received signals at time t.
- $\boldsymbol{\theta}(t) = [\theta_1(t), \theta_2(t), \cdots, \theta_K]^T$is the source DOA parameter vector.
- $\mathbf{A}(\boldsymbol{\theta}(t))$ is the array response matrix which is determined by the DOA of signals. The k-th column of $\mathbf{A}(\boldsymbol{\theta}(t))$ is defined as the array response vector associated with the k-th source and is given by $\mathbf{a}(\theta_k(t)) = [1, e^{-j\pi \sin\theta_k(t)}, \cdots, e^{-j\pi(M-1)\sin\theta_k(t)}]^T$.
- $\mathbf{S}(t) = \mathrm{diag}[s_1(t)e^{j\phi_1}, s_2(t)e^{j\phi_2}, \cdots, s_K(t)e^{j\phi_K}]$.
- $\mathbf{n}(t) = [n_1(t), n_2(t), \cdots, n_M(t)]^T$ is an $M \times 1$ additive noise vector, which is assumed to be spatially and temporally white Gaussian.

The received signals are despread by the users' own spreading sequences for each of the antenna elements. Therefore, the k-th users' despread and sampled array vector signal, $\mathbf{x}_k(i)$ may be expressed as

$$\mathbf{y}_k(i) = d_k(i)\mathbf{a}(\theta_k(i))e^{j\phi_k} + \sum_{l=1,\ l\neq k}^{K} \psi_{kl} d_l(i)\mathbf{a}(\theta_l(i))e^{j\phi_l} + \mathbf{n}_k(i) \tag{5.27}$$

where the cross-correlation of the spread sequence, ψ_{kl} may be defined as $\psi_{kl} \triangleq \int_0^T c_k(t)c_l(t)dt$ when the time delay of all users are zero and $\mathbf{n}_k(i)$ is the despread noise signal vector whose covariance matrix is $\sigma_n^2\mathbf{I}$. The outputs are then multiplied by a conjugate of the transmitted data estimate $\hat{d}_k^*(i)$ and carrier phase estimate $e^{-j\hat{\phi}_k}$ to remove the effect of $d_k(i)$ and ϕ_k. With the assumption that $\hat{d}_k^*(i) = d_k(i)$ and $\hat{\phi}_k = \phi_k$, $\mathbf{y}_k(i)\hat{d}_k^*(i)e^{-j\hat{\phi}_k}$ can be written

as

$$\begin{aligned}\tilde{\mathbf{y}}_k(i) &= \mathbf{y}_k(i)\hat{d}_k^*(i)e^{-j\hat{\phi}_k} \\ &= \mathbf{a}(\theta_k(i)) + \hat{d}_k^*(i)e^{-j\hat{\phi}_k}\sum_{l=1,\ l\neq k}^{K}\psi_{kl}d_l(i)\mathbf{a}(\theta_l(i))e^{j\phi_l} + \hat{d}_k^*(i)e^{-j\hat{\phi}_k}\mathbf{n}_k(i) \\ &= \mathbf{a}(\theta_k(i)) + \boldsymbol{\alpha}_k(i) + \tilde{\mathbf{n}}_k(i)\end{aligned} \tag{5.28}$$

where $\boldsymbol{\alpha}_k(i)$ is a vector containing the interference of the despread signal of the k-th satellite, and $\tilde{\mathbf{n}}_k(i)$ represents the noise vector with the unchanged statistical properties (see Appendix A.1.4).

5.3.2 GPS DOA tracking via extended Kalman filter

In order to track the DOA of the k-th satellite, $\theta_k(i)$, an extended Kalman filter (EKF) is used. The Kalman filtering algorithm is readily formulated in Appendix B.

State space model

For simplicity, the index k indicating the satellite number is omitted. The $\theta(i)$ is modeled as a random walk process.

$$\begin{aligned}\mathbf{x}_{i+1} &= \mathbf{A}\mathbf{x}_i + \mathbf{w}_i \\ \begin{bmatrix} \theta_{gps}(i+1) \\ \dot{\theta}_{gps}(i+1) \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \theta_{gps}(i) \\ \dot{\theta}_{gps}(i) \end{bmatrix} + \begin{bmatrix} 0 \\ w_{\theta_{gps}}(i) \end{bmatrix}.\end{aligned} \tag{5.29}$$

Observation model

The cross-correlation (interference) term, $\boldsymbol{\alpha}_k(i)$, contributed to the despread signal is ignored and only the measurement noise is considered. A simple

observation model can be written as

$$\mathbf{y}_i = \mathbf{h}(\mathbf{x}_i) + \mathbf{v}_i$$

$$\begin{bmatrix} \mathrm{Re}\{\tilde{y}_1(i)\} \\ \mathrm{Re}\{\tilde{y}_2(i)\} \\ \vdots \\ \mathrm{Re}\{\tilde{y}_M(i)\} \\ \mathrm{Im}\{\tilde{y}_1(i)\} \\ \mathrm{Im}\{\tilde{y}_2(i)\} \\ \vdots \\ \mathrm{Im}\{\tilde{y}_M(i)\} \end{bmatrix} = \begin{bmatrix} \cos(0) \\ \cos(-1\pi \sin\theta_{gps}(i)) \\ \vdots \\ \cos(-(M-1)\pi \sin\theta_{gps}(i)) \\ \sin(0) \\ \sin(-1\pi \sin\theta_{gps}(i)) \\ \vdots \\ \sin(-(M-1)\sin\theta_{gps}(i)) \end{bmatrix} + \begin{bmatrix} \mathrm{Re}\{\tilde{n}_1(i)\} \\ \mathrm{Re}\{\tilde{n}_2(i)\} \\ \vdots \\ \mathrm{Re}\{\tilde{n}_M(i)\} \\ \mathrm{Im}\{\tilde{n}_1(i)\} \\ \mathrm{Im}\{\tilde{n}_2(i)\} \\ \vdots \\ \mathrm{Im}\{\tilde{n}_M(i)\} \end{bmatrix} \tag{5.30}$$

where $\mathrm{Re}\{\tilde{y}_m(i)\}$ and $\mathrm{Im}\{\tilde{y}_m(i)\}$ are respectively the real and imaginary part of $\tilde{y}_m(i)$ at the m-th antenna. $\mathrm{Re}\{\tilde{n}_m(i)\}$ and $\mathrm{Im}\{\tilde{n}_m(i)\}$ are respectively the real and imaginary part of $\tilde{n}_m(i)$ at the m-th antenna. $\mathbf{h}(\cdot)$ is the nonlinear mapping between $\mathbf{y}_i$ and $\mathbf{x}_i$.

The EKF requires the linearized observation matrix (Jacobian matrix) which can be obtained by

$$\begin{aligned} \mathbf{H}(i) &= \frac{\partial}{\partial \mathbf{x}_i}\mathbf{h}(\mathbf{x}_i)|_{\mathbf{x}_i=\hat{\mathbf{x}}_i^-} \\ &= \begin{bmatrix} 0 & 0 \\ 1\pi\cos\hat{\theta}_{gps}^-(i)\sin(-1\pi\sin\hat{\theta}_{gps}^-(i)) & 0 \\ \vdots & \vdots \\ (M-1)\pi\cos\hat{\theta}_{gps}^-(i)\sin(-(M-1)\pi\sin\hat{\theta}_{gps}^-(i)) & 0 \\ 0 & 0 \\ -1\pi\cos\hat{\theta}_{gps}^-(i)\cos(-1\pi\sin\hat{\theta}_{gps}^-(i)) & 0 \\ \vdots & \vdots \\ -(M-1)\pi\cos\hat{\theta}_{gps}^-(i)\cos(-(M-1)\pi\sin\hat{\theta}_{gps}^-(i)) & 0 \end{bmatrix}. \end{aligned} \tag{5.31}$$

The linearized observation model is then

$$\mathbf{y}_i = \mathbf{H}(i)\mathbf{x}_i + \mathbf{v}_i. \tag{5.32}$$

Since the cross-correlation matrix is unknown and ignored, therefore, one can

expect bias errors embedded in the output estimates, $\hat{\theta}(i)$.

5.3.3 GPS/INS integration for DOA tracking

The outputs from the array-GPS DOA tracker, $\hat{\theta}_{gps}(i)$, will be fused with the DOA estimates obtained from the INS unit, $\hat{\theta}_{ins}(iR)$, via an additional Kalman filter. This approach may be comparable to the well known loosely-coupled GPS/INS integration [72,73]. Assuming that the INS provides accurate position estimates for a sufficiently long period of time and the satellites' positions are available to the INS so that the INS can use this information to transform its position into the DOA. The array-GPS DOA tracker outputs the DOA estimates at a much higher rate than that of the INS ($\sim 100-200$ Hz), therefore, the integration filter (Kalman filter) will complete the whole iteration (prediction and correction loops) only when the DOA from INS is available otherwise it will only run in the prediction loop. The outputs of the array-GPS DOA tracker are erroneous since the cross-correlations are not considered. However, by fusing the erroneous estimates with the DOA from a high grade INS unit, the error can be significantly reduced.

The state space model is defined as

$$\begin{aligned} \mathbf{x}_{i+1} &= \mathbf{A}\mathbf{x}_i + \mathbf{w}_i \\ \begin{bmatrix} \Delta\theta(i+1) \\ \Delta\dot{\theta}(i+1) \end{bmatrix} &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta\theta(i) \\ \Delta\dot{\theta}(i) \end{bmatrix} + \begin{bmatrix} 0 \\ w_{\Delta\theta}(i) \end{bmatrix}. \end{aligned} \tag{5.33}$$

Since the difference between the reliable DOA obtained from INS and DOA estimates obtained from the array-GPS, is used as the measurement. The simple linear observation model can be obtained as

$$\begin{aligned} \mathbf{y}_i &= \mathbf{H}\mathbf{x}_i + \mathbf{v}_i \\ \begin{bmatrix} \theta_{ins}(i) - \theta_{gps}(i) \end{bmatrix} &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \Delta\theta(i) \\ \Delta\dot{\theta}(i) \end{bmatrix} + \begin{bmatrix} v(i) \end{bmatrix}. \end{aligned} \tag{5.34}$$

The integrated DOA estimate, $\hat{\theta}(i)$, is obtained by (see also Fig. 5.3)

$$\hat{\theta}(i) = \theta_{gps}(i) + \Delta\hat{\theta}(i). \tag{5.35}$$

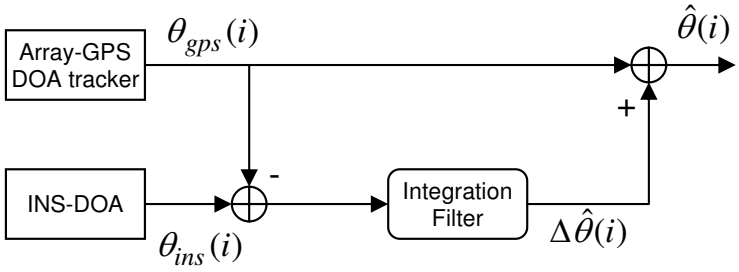

Figure 5.3: GPS/INS integration scheme.

5.3.4 Simulation results

For simulation purpose, the direction-of-arrivals of the GPS signals are generated randomly following the sinusoidal functions. The PN-sequences are used to generate the transmitted and received GPS signals. The following assumptions are also assumed in the simulations.

- 8 GPS linear array antennas are used.
- The data bits are known and can be removed from the received signal.
- Maximum of 12 GPS satellites are seen by all antennas.
- The array-GPS outputs DOA estimates 100 times faster than INS.
- High quality INS is available.
- Calibration and synchronization are assumed to be perfect.

Fig. 5.4 shows the DOA tracking ability of the array-GPS based DOA tracker with and without INS aiding. It can be seen that if the cross-correlation or interference term is ignored and not compensated for, the DOA estimation error is large. However, by using the proposed GPS/INS integration concept for DOA tracking, this error can be significantly reduced. The quality of the INS will of course play an important roll in reducing the estimation error.

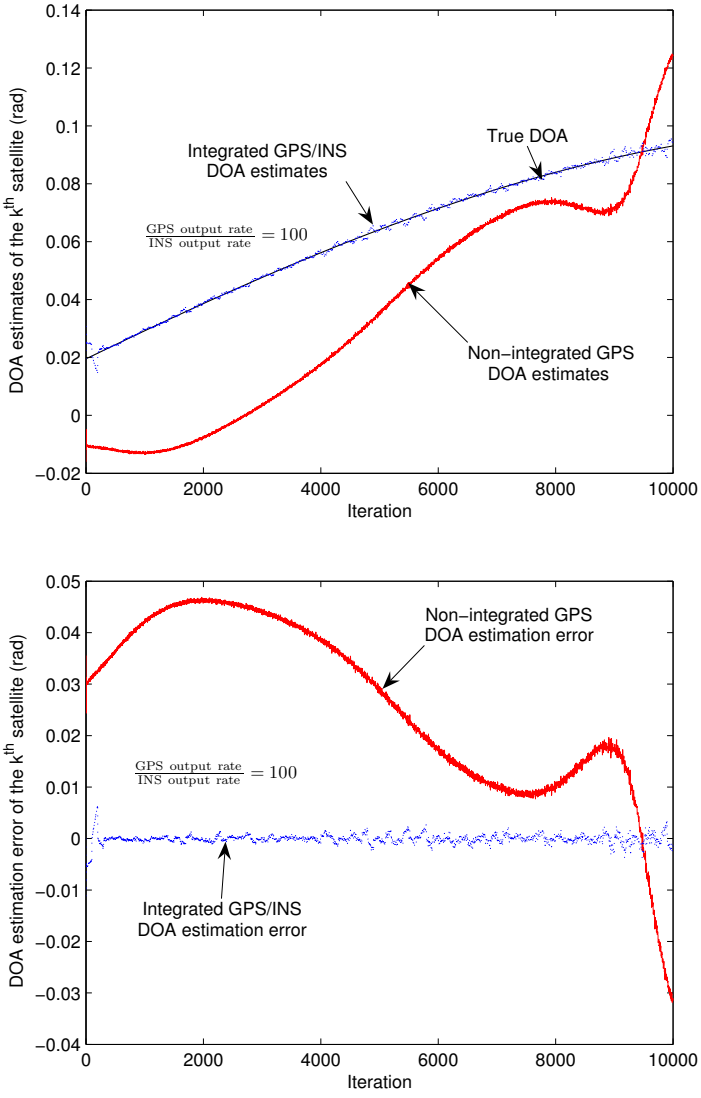

Figure 5.4: DOA tracking via GPS/INS integration.

5.4 Summary

This chapter shows that the problem of frequency estimation can also be seen from SAR (Doppler centroid estimation) and array processing (direction-of-arrival estimation) point of view. The formulations of the problem in both cases present finally similar expression as obtained from the carrier frequency offset estimation problem in communication systems.

The integrated GPS/INS DOA estimation problem is also considered in this chapter. It is shown that an additional DOA information can be used to reduce the GPS-alone DOA estimation errors caused by the uncompensated cross-correlation of GPS signals arriving from other directions.

Chapter 6

Conclusions and Future Works

6.1 Conclusions

The performance of communication systems for both single-carrier and multi-carrier (i.e., OFDM) transmission schemes, depend on the signal quality seen by the receiver. The ideal signal integrity is only obtained when the system impairment such as CFO is eliminated from the system. On the other hand, the performance of Doppler radar systems depend on how accurately and precisely the Doppler frequency is extracted from the echo's signal. These requirements demand properly designed frequency estimation algorithms. In this thesis, issues related to frequency estimation are addressed. In particular, after introduction of the single-carrier, multi-carrier OFDM systems, and the analysis of the effects of CFO errors, new data-aided frequency estimation algorithms have been derived and analyzed. It has been shown in the later chapter that these proposed algorithms can be easily extended to use in the Doppler radar and array processing applications. The main contributions of this thesis include:

Single-carrier case

- Closed-form quadratic interpolation function for the frequency estimation using periodogram has been derived. It significantly reduces the amount of zero-padding required by the FFT-periodogram in order to attain the CRLB.

- The modified Kay's estimator, which solves the problem of high threshold behavior in the original Kay's estimator, has been proposed. This comes at the price of a narrower estimation range, however, for a given estimation range the modified Kay's estimator can provide a much lower threshold than the original one.

- The summation-by-parts formula (weight transformation formula) is applied to extend the estimation range of the original Fitz's estimator which leads to the modified Fitz's estimator. The original Fitz's estimator has a very narrow estimation range when operated in the optimal mode. This problem is usually overcome by applying a proper phase unwrapping algorithm, however at the price of high computational complexity.

- An approximated maximum likelihood frequency estimator based on the absolute phase information of the autocorrelation estimates, is proposed. The weight transformation formula is applied to the proposed estimator resulting in the transformed proposed estimator which has a similar performance as the original one but with the estimation range extended to the maximum. The performance of the transformed proposed estimator is also shown to be better than that of the well known M&M estimator.

- By using the nonlinear least-squares principle in conjunction with the summation-by-parts formula, a new closed-form frequency estimator which can be used to estimate the frequency of signals with fading-induced multiplicative noise has been derived. The estimator requires neither phase unwrapping algorithm nor additional information about the channel as needed for most of the existing estimators.

- Designed a Kalman filter for frequency estimation of signals with real-valued time-varying amplitude. It can be easily extended for frequency tracking problem.

OFDM case

- The nonlinear least-squares estimator derived for the single-carrier case has been applied to multi-carrier OFDM signal. The improvement can be obtained using multiple receive antennas. The error floor effect of the estimation variance introduced by the non-delta function of the channel correlation function can be reduced by using an array of uncorrelated receive antennas.

Array processing

- The integrated GPS/INS DOA tracker has been proposed. By using the additional DOA information obtained from INS, the DOA estimation errors of the GPS-alone tracker can be reduced.

6.2 Future Works

In this work, the problem of frequency estimation in communication and radar related applications has been addressed. Several improved frequency estimators including a novel closed-form NLS-based frequency estimator have been proposed. For future works, a closed-form estimator for the MIMO case may be derived in the same manner as in the case of SISO and SIMO. The performance analysis of these estimators could also be carried out. Moreover, for frequency tracking purposes, various types of Kalman filters and smoothers can be applied on the correlation estimates of the received signals.

Since the proposed NLS-based estimator can also be applied to the Doppler centroid estimation problem in SAR applications, therefore a justification of the proposed estimator by the real SAR data could be of interest.

Appendix A

Derivations

A.1 Constant Envelope

A.1.1 Cramer-Rao lower bound

The complex Gaussian pdf is defined as

$$p(\mathbf{x};\boldsymbol{\xi}) = \frac{1}{\pi^N \det(\mathbf{C}_x(\boldsymbol{\xi}))} \exp\left[-(\mathbf{x}-\boldsymbol{\mu}(\boldsymbol{\xi}))^H \mathbf{C}_x^{-1}(\boldsymbol{\xi})(\mathbf{x}-\boldsymbol{\mu}(\boldsymbol{\xi}))\right] \tag{A.1}$$

where the parameter vector $\boldsymbol{\xi}$ is to be estimated based on the complex data $\mathbf{x} = [x[0], x[1], \ldots, x[N-1]]^T$ and may have components that are real and/or complex. However, in this thesis only the real-valued case is considered.

The Cramer-Rao lower bound (CRLB) for a complex Gaussian pdf is obtained from the inverse of the Fisher information matrix defined as (see eq.(15.52) in [74])

$$[\mathbf{I}(\boldsymbol{\xi})]_{ij} = \mathrm{tr}\left\{\mathbf{C}_x^{-1}(\boldsymbol{\xi})\frac{\partial \mathbf{C}_x(\boldsymbol{\xi})}{\partial \xi_i}\mathbf{C}_x^{-1}(\boldsymbol{\xi})\frac{\partial \mathbf{C}_x(\boldsymbol{\xi})}{\partial \xi_j}\right\} + 2\mathrm{Re}\left\{\frac{\partial \boldsymbol{\mu}^H(\boldsymbol{\xi})}{\partial \xi_i}\mathbf{C}_x^{-1}(\boldsymbol{\xi})\frac{\partial \boldsymbol{\mu}(\boldsymbol{\xi})}{\partial \xi_j}\right\} \tag{A.2}$$

for $i, j = 1, 2, \ldots, p$. The derivatives are defined as the matrix or vector of partial derivatives, where we differentiate each complex element $g = g_R + jg_I$

as

$$\frac{\partial g}{\partial \xi_i} = \frac{\partial g_R}{\partial \xi_i} + j\frac{\partial g_I}{\partial \xi_i}.$$

The equality condition for the CRLB to be efficient is,

$$\frac{\partial \ln p(\mathbf{x};\boldsymbol{\xi})}{\partial \boldsymbol{\xi}} = \mathbf{I}(\boldsymbol{\xi})(g(\mathbf{x}) - \boldsymbol{\xi})$$

where $\hat{\boldsymbol{\xi}} = g(\mathbf{x})$ is the efficient estimate of $\boldsymbol{\xi}$.

Consider the signal model defined in (3.1)

$$y(k) = \tilde{a}\exp\{j2\pi f_0 k\} + n(k), \qquad 0 \le k \le N-1 \tag{A.3}$$

where $\tilde{a} = a\exp\{j\phi\}$ is a constant complex envelope, $f_0 = F_0 T$ is the normalized frequency to be estimated, $n(k) = n_I(k) + jn_Q(k)$ is a zero-mean complex-valued white Gaussian process with variance σ^2. The SNR is defined as a^2/σ^2. Note that a, ϕ, f_0 are deterministic but unknown.

From (A.2), we have

$$\begin{aligned}[\mathbf{I}(\boldsymbol{\xi})]_{ij} &= 2\mathrm{Re}\left\{\frac{\partial \boldsymbol{\mu}^H(\boldsymbol{\xi})}{\partial \xi_i}\frac{1}{\sigma^2}\mathbf{I}\frac{\partial \boldsymbol{\mu}(\boldsymbol{\xi})}{\partial \xi_j}\right\} \\ &= \frac{2}{\sigma^2}\mathrm{Re}\left\{\sum_{k=0}^{N-1}\frac{\partial s^*(k)}{\partial \xi_i}\frac{\partial s(k)}{\partial \xi_j}\right\}\end{aligned} \tag{A.4}$$

where $s(k) = \tilde{a}\exp\{j2\pi f_0 k\} = a\exp\{j(2\pi f_0 k + \phi)\}$. The partial derivatives are now

$$\begin{aligned}\frac{\partial s(k)}{\partial a} &= \exp\{j(2\pi f_0 k + \phi)\} \\ \frac{\partial s(k)}{\partial f_0} &= j2\pi k a\exp\{j(2\pi f_0 k + \phi)\} \\ \frac{\partial s(k)}{\partial \phi} &= ja\exp\{j(2\pi f_0 k + \phi)\}\end{aligned} \tag{A.5}$$

so that

$$\mathbf{I}(\boldsymbol{\xi}) = \frac{2}{\sigma^2}\begin{pmatrix} N & 0 & 0 \\ 0 & a^2\sum_{k=0}^{N-1}(2\pi k)^2 & a^2\sum_{k=0}^{N-1}2\pi k \\ 0 & a^2\sum_{k=0}^{N-1}2\pi k & Na^2 \end{pmatrix}. \tag{A.6}$$

After inversion we have

$$\begin{aligned} \mathrm{CRLB}(\hat{a}) &\geq \frac{\sigma^2}{2N} \\ \mathrm{CRLB}(\hat{f}_0) &\geq \frac{6\sigma^2}{(2\pi)^2 a^2 N(N^2-1)} = \frac{6\mathrm{SNR}^{-1}}{(2\pi)^2 N(N^2-1)} \\ \mathrm{CRLB}(\hat{\phi}) &\geq \frac{\sigma^2(2N-1)}{a^2 N(N+1)} = \frac{(2N-1)\mathrm{SNR}^{-1}}{N(N+1)}. \end{aligned} \tag{A.7}$$

A.1.2 Maximum likelihood estimator

The Maximum Likelihood Estimator (MLE) can be derived by maximizing

$$\begin{aligned} p(\mathbf{y};\mathbf{x}) &= \frac{1}{\pi^N \det(\mathbf{C})}\exp\left\{-[\mathbf{y}-\mathbf{h}(\mathbf{x})]^H\mathbf{C}^{-1}[\mathbf{y}-\mathbf{h}(\mathbf{x})]\right\} \\ &= \frac{1}{\pi^N \det(\sigma^2\mathbf{I})}\exp\left\{-\frac{1}{\sigma^2}[\mathbf{y}-a\mathbf{e}]^H[\mathbf{y}-a\mathbf{e}]\right\} \end{aligned} \tag{A.8}$$

or equivalently, we must minimize the log-likelihood function

$$\begin{aligned} J(a,f_0) &= [\mathbf{y}-a\mathbf{e}]^H[\mathbf{y}-a\mathbf{e}] \\ &= \mathbf{y}^H\mathbf{y} - \mathbf{y}^H\mathbf{e}a - a^H\mathbf{e}^H\mathbf{y} + a^H\mathbf{e}^H\mathbf{e}a \end{aligned} \tag{A.9}$$

where

$$\begin{aligned} \mathbf{y} &= [y(0), y(1), \cdots, y(N-1)]^T \\ \mathbf{e} &= [1, \exp\{j2\pi f_0\}, \cdots, \exp\{j2\pi f_0(N-1)\}]^T. \end{aligned}$$

Minimizing (A.9) with respect to a, provided that f_0 is fixed, can be seen as a complex linear least-squares problem. The derivative of (A.9) is

$$\begin{aligned}\frac{\partial J}{\partial a} &= \mathbf{0} - \left(\mathbf{e}^H\mathbf{y}\right)^* - \mathbf{0} + \left(\mathbf{e}^H\mathbf{e}a\right)^* \\ &= -[\mathbf{e}^H(\mathbf{y} - \mathbf{e}a)]^*.\end{aligned} \tag{A.10}$$

Setting it to zero and solving for a produces

$$\hat{a} = \frac{\mathbf{e}^H\mathbf{y}}{\mathbf{e}^H\mathbf{e}} = \frac{1}{N}\sum_{k=0}^{N-1} y(k)\exp\{-j2\pi f_0 k\}. \tag{A.11}$$

Substitute this back into J yields

$$\begin{aligned}J(\hat{a}, f_0) &= \mathbf{y}^H(\mathbf{y} - \mathbf{e}\hat{a}) \\ &= \mathbf{y}^H\mathbf{y} - \frac{\mathbf{y}^H\mathbf{e}\mathbf{e}^H\mathbf{y}}{\mathbf{e}^H\mathbf{e}} \\ &= \mathbf{y}^H\mathbf{y} - \frac{|\mathbf{e}^H\mathbf{y}|^2}{\mathbf{e}^H\mathbf{e}}.\end{aligned} \tag{A.12}$$

To minimize J over f_0 one needs to maximize $\frac{|\mathbf{e}^H\mathbf{y}|^2}{\mathbf{e}^H\mathbf{e}}$. The MLE for f_0 reads

$$\hat{f}_0 = \arg\max_{\tilde{f}_0} \frac{1}{N}\left|\sum_{k=0}^{N-1} y(k)\exp\{-j2\pi\tilde{f}_0 k\}\right|^2 \tag{A.13}$$

where $\tilde{f}_0$ is the tentative value of f_0. The MLE produces the estimate of f_0 which is consistent and asymptotically efficient (as $N \to \infty$, its variance equals to the CRLB) [17]. The corresponding MLE of $|a|$ and ϕ are found from

$$\hat{\tilde{a}} = \widehat{|a|} = \left|\frac{1}{N}\sum_{k=0}^{N-1} y(k)\exp\{-j2\pi\hat{f}_0 k\}\right| \tag{A.14}$$

$$\hat{\phi} = \arctan \frac{\mathrm{Im}\left\{\hat{\hat{a}}\right\}}{\mathrm{Re}\left\{\hat{\hat{a}}\right\}} = \arctan \frac{\mathrm{Im}\left\{\sum_{k=0}^{N-1} y(k) \exp\{-j2\pi \hat{f}_0 k\}\right\}}{\mathrm{Re}\left\{\sum_{k=0}^{N-1} y(k) \exp\{-j2\pi \hat{f}_0 k\}\right\}}. \tag{A.15}$$

A.1.3 Closed-form quadratic interpolation

It is assumed that the periodogram is approximately quadratic in the vicinity of the peak, and thus can be written as

$$p(f) = c_1 f^2 + c_2 f + c_3. \tag{A.16}$$

Taking the peak (f_2, p_2) obtained from the coarse search and its left (f_1, p_1) and right (f_3, p_3) nearest neighbors, the three coefficients can be found using the following algebra

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} f_1^2 & f_1 & 1 \\ f_2^2 & f_2 & 1 \\ f_3^2 & f_3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} p_1 \\ p_2 \\ P_3 \end{pmatrix} \tag{A.17}$$

where $p_i \equiv p(f_i)$. In order to obtain closed-forms for c_1, c_2, c_3, (A.17) is rewritten as

$$\begin{aligned} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} =& \frac{1}{f_1^2(f_2 - f_3) - f_1(f_2^2 - f_3^2) + (f_2^2 f_3 - f_2 f_3^2)} \\ & \cdot \begin{pmatrix} f_2 - f_3 & f_3 - f_1 & f_1 - f_2 \\ f_3 - f_2 & f_1^2 - f_3^2 & f_2^2 - f_1^2 \\ f_2^2 f_3 - f_2 f_3^2 & f_1 f_3^2 - f_1^2 f_3 & f_1^2 f_2 - f_1 f_2^2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}. \end{aligned} \tag{A.18}$$

The closed-forms for c_1, c_2, c_3 are then found to be

$$\begin{aligned} c_1 &= \frac{(f_2 - f_3)p_1 + (f_3 - f_1)p_2 + (f_1 - f_2)p_3}{f_1^2(f_2 - f_3) - f_1(f_2^2 - f_3^2) + (f_2^2 f_3 - f_2 f_3^2)} \\ c_2 &= \frac{(f_3^2 - f_2^2)p_1 + (f_1^2 - f_3^2)p_2 + (f_2^2 - f_1^2)p_3}{f_1^2(f_2 - f_3) - f_1(f_2^2 - f_3^2) + (f_2^2 f_3 - f_2 f_3^2)} \\ c_3 &= \frac{f_2 f_3(f_2 - f_3)p_1 + f_3 f_1(f_3 - f_1)p_2 + f_1 f_2(f_1 - f_2)p_3}{f_1^2(f_2 - f_3) - f_1(f_2^2 - f_3^2) + (f_2^2 f_3 - f_2 f_3^2)}. \end{aligned} \tag{A.19}$$

The estimated frequency is the frequency that maximizes (A.16) and can be found by taking the derivative of (A.16) and setting to zero, or

$$\hat{f} = -\frac{c_2}{2c_1} = -\frac{1}{2}\frac{(f_3^2 - f_2^2)p_1 + (f_1^2 - f_3^2)p_2 + (f_2^2 - f_1^2)p_3}{(f_2 - f_3)p_1 + (f_3 - f_1)p_2 + (f_1 - f_2)p_3}. \tag{A.20}$$

A.1.4 Statistically equivalent of a complex white Gaussian noise

Let $v(k) = n(k) \cdot \exp\{j\theta(k)\}$ be a product of a complex white Gaussian noise, $n(k)$, and a complex exponential function, $\exp\{j\theta(k)\}$. The statistical properties of $v(k)$ will be shown to be equivalent to that of $n(k)$. The quantity $v(k)$ may be written as

$$v(k) = [n_I(k) + jn_Q(k)] \cdot [\cos(\theta(k)) + j\sin(\theta(k))]$$

where $n(k) = n_I(k) + jn_Q(k)$ with $\mathrm{E}\{n_I(k)\} = \mathrm{E}\{n_Q(k)\} = 0$, $\mathrm{E}\{n_I^2(k)\} = \mathrm{E}\{n_Q^2(k)\} = \frac{\sigma_n^2}{2}$, and $\mathrm{E}\{n_I(k)n_Q(k+m)\} = 0$. In order to show that $n(k)$ and $v(k)$ are statistically equivalent, we start with

$$\begin{aligned} v_I(k) &= n_I(k)\cos(\theta(k)) + n_Q(k)\sin(\theta(k)) \\ v_Q(k) &= -n_I(k)\sin(\theta(k)) + n_Q(k)\cos(\theta(k)). \end{aligned}$$

The means of $v_I(k)$ and $v_Q(k)$ are clearly zeros.

The correlation of $v_I(k)$ is

$$\begin{aligned}\mathrm{E}\{v_I(k)v_I(k+m)\} &= \mathrm{E}\{n_I(k)n_I(k+m)\}\cos(\theta(k))\cos(\theta(k+m)) \\ &+ \mathrm{E}\{n_Q(k)n_Q(k+m)\}\sin(\theta(k))\sin(\theta(k+m))\end{aligned}$$

where

$$\begin{aligned}\mathrm{E}\{n_I(k)n_I(k+m)\} &= \frac{\sigma_n^2}{2}\delta(m) \\ \mathrm{E}\{n_Q(k)n_Q(k+m)\} &= \frac{\sigma_n^2}{2}\delta(m).\end{aligned}$$

From the relation $\cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b)$, we have

$$\begin{aligned}&\mathrm{E}\{v_I(k)v_I(k+m)\} = \frac{\sigma_n^2}{2}\delta(m)\cos(\theta(k)-\theta(k+m)) \\ &= \begin{cases} \mathrm{E}\{v_I^2(k)\} = \frac{\sigma_n^2}{2}, & \text{for } m = 0 \rightarrow \delta(m) = 1, \cos(\cdot) = 1; \\ \mathrm{E}\{v_I(k)v_I(k+m)\} = \frac{\sigma_n^2}{2}\delta(m) = 0, & \text{for } \tau \neq 0 \rightarrow \delta(m \neq 0) = 0. \end{cases}\end{aligned}$$

The correlation of $v_Q(k)$ can be obtained similarly

$$\begin{aligned}\mathrm{E}\{v_Q(k)v_Q(k+m)\} &= \mathrm{E}\{n_I(k)n_I(k+m)\}\sin(\theta(k))\sin(\theta(k+m)) \\ &+ \mathrm{E}\{n_Q(k)n_Q(k+m)\}\cos(\theta(k))\cos(\theta(k+m)) \\ &= \frac{\sigma_n^2}{2}\delta(m)\cos(\theta(k)-\theta(k+m)) \\ &= \frac{\sigma_n^2}{2}\delta(m).\end{aligned}$$

The cross-correlation, $\mathrm{E}\{v_I(k)v_Q(k+m)\}$, is

$$\begin{aligned}\mathrm{E}\{v_I(k)v_Q(k+m)\} &= -\mathrm{E}\{n_I(k)n_I(k+m)\}\cos(\theta(k))\sin(\theta(k+m)) \\ &+ \mathrm{E}\{n_Q(k)n_Q(k+m)\}\sin(\theta(k))\cos(\theta(k+m)) \\ &= \frac{\sigma_n^2}{2}\delta(m)\sin(\theta(k)-\theta(k+m)) \\ &= 0, \quad \forall m.\end{aligned}$$

This proofs that the first and second order statistics of $n(k)$ and $v(k)$ are equivalent.

A.1.5 Complex noise to phase noise transformation

The observed signal may be expressed as

$$\begin{aligned} y(k) &= |a| \exp\{j(\omega_o k + \phi)\} + n(k) \\ &= |a|[1 + v(k)] \exp\{j(\omega_o k + \phi)\} \end{aligned} \tag{A.21}$$

where $\omega_o = 2\pi f_o$, and

$$v(k) = \frac{1}{|a|} n(k) \exp\{-j(\omega_o k + \phi)\} \tag{A.22}$$

is a complex white noise sequence with

$$\text{var}\{v(k)\} = \frac{\sigma_n^2}{|a|^2} = \frac{1}{\text{SNR}}. \tag{A.23}$$

Let $v(k) = v_I(k) + j v_Q(k)$, we have

$$\begin{aligned} 1 + v(k) &= |1 + v(k)| \cdot \exp\{\arg\{1 + v(k)\}\} \\ &= \sqrt{[1 + v_I(k)]^2 + v_Q^2(k)} \cdot \exp\left\{ j \tan^{-1} \frac{v_Q(k)}{1 + v_I(k)} \right\}. \end{aligned} \tag{A.24}$$

For high SNR or $\frac{|a|^2}{\sigma_n^2} \gg 1$,

$$1 + v(k) \approx \exp\{j \tan^{-1} v_Q(k)\} \approx \exp\{j v_Q(k)\} \tag{A.25}$$

therefore, $y(k)$ can be written as

$$y(k) = |a| \exp\{j(\omega_o k + \phi + v_Q(k))\}. \tag{A.26}$$

The additive noise has been transformed into an equivalent phase noise $v_Q(k)$ with

$$\text{var}\{v_Q(k)\} = \frac{1}{2}\text{var}\{v(k)\} = \frac{1}{2}\frac{\sigma_n^2}{|a|^2}. \tag{A.27}$$

Note that this approximation holds true only for high SNR assumption.

A.1.6 Tretter estimator

For high SNR, Tretter showed that (3.1) can be approximated as

$$y(k) \approx |a| \exp\{j(2\pi f_0 k + \phi + v_Q(k))\} \tag{A.28}$$

where $v_Q(k)$ is the equivalent phase noise with variance $\frac{1}{2}\frac{\sigma_n^2}{|a|^2}$ and is defined in Appendix A.1.5. The phase of $y(k)$ is given by

$$\arg\{y(k)\} \approx [2\pi f_0 k + \phi + v_Q(k)]_{-\pi}^{\pi} \tag{A.29}$$

where $[x]_{-\pi}^{\pi}$ is the value of x reduced to the interval $[-\pi, \pi)$. This equation relates f_0 and ϕ to $\arg\{y(k)\}$. Unfortunately, (A.29) is highly nonlinear and awkward to handel. Therefore, phase unwrapping procedure [19] is needed but this would be difficult at low SNR. Assume that phase unwrapping is successfully applied, the unwrapped phase function

$$\text{unwrap}\left[\arg\{y(k)\}\right] = \varphi(k) = 2\pi f_0 k + \phi + v_Q(k) \tag{A.30}$$

is a linear function. The parameters f_0 and ϕ can be estimated by means of least-squares or linear regression. In the case of Gaussian noise, least-squares estimates are equivalent to maximum likelihood estimates. The minimization criterion is the square error

$$J = \sum_{k=0}^{N-1} \left[\varphi(k) - 2\pi \hat{f}_0 k - \hat{\phi}\right]^2. \tag{A.31}$$

The explicit least-squares solution of $\hat{f}_0$ is evaluated in [18] and obtained as

$$\begin{aligned}\hat{f}_0 &= \frac{12}{2\pi N(N^2-1)}\sum_{k=0}^{N-1} k \arg\{y(k)\} - \frac{6}{2\pi N(N+1)}\sum_{k=0}^{N-1}\arg\{y(k)\} \\ &= \frac{12}{2\pi N(N^2-1)}\sum_{k=0}^{N-1}\left[k - \frac{(N-1)}{2}\right]\arg\{y(k)\} \\ &= \frac{1}{2\pi}\sum_{k=0}^{N-1}\frac{6(2k-N+1)}{N(N^2-1)}\arg\{y(k)\}\,.\end{aligned} \tag{A.32}$$

It is also shown in [18] that $\hat{f}_0$ is unbias and the variance achieves the CRLB.

A.1.7 Kay estimator

To avoid phase unwrapping, Kay [20] derived an estimator based on the phase differences between two consecutive samples. The phase differences take the form of

$$\begin{aligned}\Delta\varphi(k) &= \arg\{y(k)y^*(k-1)\} \\ &= [2\pi f_0 + v_Q(k) - v_Q(k-1)]_{-\pi}^{\pi}\end{aligned} \tag{A.33}$$

provided that $|2\pi f_0| < \pi$ and $v_Q(k)$ is sufficiently small, (A.33) can be approximated as

$$\Delta\varphi(k) \approx 2\pi f_0 + v_Q(k) - v_Q(k-1). \tag{A.34}$$

It is clear from (A.34) that the problem now is to estimate the mean f_0 of a colored Gaussian noise process. The maximum-likelihood estimate of f_0 is found by minimizing

$$J = (\mathbf{\Delta} - 2\pi f_0\mathbf{1})^T\mathbf{C}^{-1}(\mathbf{\Delta} - 2\pi f_0\mathbf{1}) \tag{A.35}$$

where $\mathbf{\Delta} = [\Delta\varphi(1), \Delta\varphi(2), \cdots, \Delta\varphi(N-1)]^T$, $\mathbf{1} = [1, 1, \cdots, 1]^T$, and $\mathbf{C}$ is the $(N-1)\times(N-1)$ covariance matrix of $\Delta\varphi(k)$, and can be explicitly

expressed as

$$\mathbf{C} = \frac{\sigma^2}{2|a|^2} \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}. \tag{A.36}$$

The solution to this problem is well known and given by

$$\hat{f}_0 = \frac{1}{2\pi} \frac{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{\Delta}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}}. \tag{A.37}$$

The estimation variance is

$$\mathrm{var}(\hat{f}_0) = \frac{1}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}}. \tag{A.38}$$

The inverse $\mathbf{C}^{-1}$ with the (i, j) element is well known [75] and given by

$$[\mathbf{C}^{-1}]_{ij} = \frac{2|a|^2}{\sigma^2} \left[\min(i, j) - \frac{i \cdot j}{N} \right] \qquad 1 \leq i, j \leq N - 1. \tag{A.39}$$

After some algebra, using (A.37) and (A.39), one obtains

$$\hat{f}_0 = \frac{1}{2\pi} \sum_{k=1}^{N-1} \underbrace{\frac{6k(N-k)}{N(N^2-1)}}_{w_k} \Delta\varphi(k) \tag{A.40}$$

where w_k is Kay's weighting function. The estimation variance of Kay estimator is shown to be identical to the CRLB.

A.1.8 Fitz estimator

The equivalent likelihood function or the periodogram is given as

$$\begin{aligned} P(\tilde{f}_0) &\triangleq \left| \sum_{k=0}^{N-1} y(k) \exp\{-j2\pi f_0 k\} \right|^2 \\ &= \sum_{k=1}^{N-1} \sum_{n=1}^{N-1} y(k) y^*(n) \exp\{-j2\pi f_0 (k-n)\}. \end{aligned} \tag{A.41}$$

Taking the derivative of (A.41) with respect to $\tilde{f}_0$ and equating it to zero yields

$$\sum_{k=1}^{N-1} \sum_{n=1}^{N-1} (k-n) y(k) y^*(n) \exp\{-j2\pi f_0 (k-n)\} = 0 \tag{A.42}$$

after rearranging and grouping terms

$$\mathrm{Im} \left\{ \sum_{m=1}^{N-1} m(N-m) \hat{r}(m) \exp\{-j2\pi f_0 m\} \right\} = 0 \tag{A.43}$$

where $\hat{r}(m)$ is the estimated autocorrelation of the samples, $y(k)$, defined as

$$\hat{r}(m) = \frac{1}{N-m} \sum_{k=m}^{N-1} y(k) y^*(k-m). \tag{A.44}$$

The term in braces in (A.43) can be seen as the Fourier transformation of the estimated autocorrelation, $\hat{r}(m)$, weighted by the parabolic windowing function $w(m) = m(N-m)$. This weighting function accounts for the fact that, in the vicinity of $m = 0$, $\hat{r}(m)$ bears intrinsically little or no information on the frequency offset because it is derived from closely spaced signal samples. On the other hand, when $m \approx N$, $\hat{r}(m)$ becomes a poor estimate of the autocorrelation of $y(k)$, since the number of terms in the summation in (A.44) building up such an estimate is small. The estimated autocorrelation can be expressed as

$$\hat{r}(m) = A \exp\{j2\pi f_0 m\} + \gamma(m). \tag{A.45}$$

where $\gamma(m) = \gamma_I(m) + j\gamma_Q(m)$.

In combination with the phase increments concept of Kay's approach, this model was considered and used to derive the estimator in [26]. Since $\arg\{E\{\hat{r}(m)\}\} = 2\pi f_0 m$. For large N (e.g. small noise in (A.45))

$$\begin{aligned} \operatorname{Im}\{\hat{r}(m)\exp\{-j2\pi f_0 m\}\} &\approx A\sin\left(\arg\{\hat{r}(m)\} - 2\pi f_0 m\right) \\ &\approx A\left(\arg\{\hat{r}(m)\} - 2\pi f_0 m\right), \end{aligned} \tag{A.46}$$

then (A.43) is heuristically approximated for large N as

$$\sum_{m=1}^{N-1} m\left(\arg\{\hat{r}(m)\} - 2\pi f_0 m\right) = 0 \tag{A.47}$$

where

$$\arg\{\hat{r}(m)\} = [2\pi f_0 m + \gamma_I(m)]_{2\pi}. \tag{A.48}$$

After some algebra, (A.47) reduces to

$$\hat{f}_0 = \frac{1}{2\pi}\sum_{m=1}^{N-1} \frac{6m}{N(N-1)(2N-1)} \arg\{\hat{r}(m)\}. \tag{A.49}$$

If the summation in (A.49) is truncated at values less than N, a more practical estimator has the form

$$\hat{f}_0 = \frac{1}{2\pi}\frac{\sum_{m=1}^{L} m\arg\{\hat{r}(m)\}}{\sum_{m=1}^{L} m^2}. \tag{A.50}$$

Note that as m varies from 1 to L, eq. (A.48) will only establish a linear relation between f_0 and $\arg\{\hat{r}(m)\}$, if f_0 lies within the interval $\pm\frac{1}{2L}$. Otherwise, eq. (A.48) becomes highly nonlinear and introduces ambiguities. Hence, phase unwrapping is required. The estimation range of Fitz estimator is, therefore, $|f_0| \leq \frac{1}{2L}$.

A.1.9 Mengali estimator

In deriving Mengali estimator, one shall start with the general signal model

$$\begin{aligned} y(k) &= a(k) \cdot \exp\{j2\pi k f_0\} + n(k), \quad k = 0, 1, \ldots, N-1 \\ &= \exp\{j2\pi k f_0\} \cdot [a(k) + \tilde{n}(k)] \end{aligned} \tag{A.51}$$

where $\tilde{n}(k) = n(k) \exp\{-j2\pi k f_0\}$.

The estimated autocorrelation can be written as

$$\begin{aligned} \hat{r}(m) &= \frac{1}{N-m} \sum_{k=m}^{N-1} y(k) y^*(k-m), \quad m = 0, 1, \ldots, L \\ &= \sigma_a^2 \exp\{j2\pi f_0 m\}[1 + \gamma(m)] \end{aligned} \tag{A.52}$$

where L is a design parameter less than $N-1$, $\sigma_a^2 = \frac{1}{N} \sum_{k=0}^{N-1} |a(k)|^2$, and

$$\begin{aligned} \gamma(m) &= \gamma_R(m) + j\gamma_I(m) \\ &= \frac{1}{\sigma_a^2 (N-m)} \sum_{k=m}^{N-1} [a(k)\tilde{n}^*(k-m) + a^*(k)\tilde{n}(k) + \tilde{n}(k)\tilde{n}^*(k-m)] \\ &\approx \frac{1}{\sigma_a^2 (N-m)} \sum_{k=m}^{N-1} [a(k)\tilde{n}^*(k-m) + a^*(k)\tilde{n}(k)] \end{aligned} \tag{A.53}$$

where $\gamma_R(m)$ and $\gamma_I(m)$ are real and imaginary part of $\gamma(m)$. For SNR $\gg 1$, the amplitude of $n(k)$ is less than unity with high probability and the term $\tilde{n}(k)\tilde{n}^*(k-m)$ can be neglected.

Using phase difference, (A.52), (A.53), and Tretter and Kay approaches [18] [20], one obtains

$$\Delta(m) = \arg\{\hat{r}(m)\hat{r}^*(m-1)\} \approx 2\pi f_0 + \gamma_I(m) - \gamma_I(m-1) \tag{A.54}$$

provided that $|2\pi f_0| < \pi$ and the $\gamma_I(m)$ are sufficient small. Clearly, (A.54) relates f_0 to $\Delta(m)$ in a linear fashion. The maximum likelihood estimate of f_0 is found by

$$\hat{f}_0 = \frac{1}{2\pi} \frac{\mathbf{1}^T \mathbf{C}^{-1} \boldsymbol{\Delta}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} \tag{A.55}$$

where

$$\begin{aligned}\mathbf{C}_{m,l} &= E\left\{[\gamma_I(m)-\gamma_I(m-1)][\gamma_I(l)-\gamma_I(l-1)]\right\} \\ &= \rho(m,l)+\rho(m-1,l-1)-\rho(m-1,l)-\rho(m,l-1)\end{aligned} \tag{A.56}$$

where

$$\rho(m,l) = E\left\{\gamma_I(m)\gamma_I(l)\right\}. \tag{A.57}$$

Inserting (A.53) into (A.57), yields

$$\rho(m,l) = \frac{(\mathrm{SNR})^{-1}}{(N-m)(N-l)}\cdot[N-\max(m,l)-(N-m-l)u(N-m-l)] \tag{A.58}$$

where $u(n)$ is the unit-step function [1]. From (A.58) and (A.56), it can be observed that $\mathbf{C}$ is singular for $L > N/2$. In fact, the angles, $\Delta(m)$, with $m > N/2$ are linearly dependent on those with $1 \leq m \leq N/2$. This means that the information borne by $\mathbf{\Delta}$ is contained in its first $N/2$ components. For this reason, it is meaningful to assume $1 \leq L \leq N/2$.

For $1 \leq L \leq N/2$, (A.58) reduces to

$$\rho(m,l) = \frac{(\mathrm{SNR})^{-1}}{(N-m)(N-l)}\min(m,l). \tag{A.59}$$

Following [26] and using (A.59) and (A.56), the element-wise closed-form expression of $\mathbf{C}^{-1}$ is

$$[\mathbf{C}^{-1}]_{ij} = \begin{cases} (L-N)\times\mathrm{SNR}, & i \neq j \\ (L-N+(N-i)(N-i+1))\times\mathrm{SNR}, & i = j \end{cases} \tag{A.60}$$

[1] The unit-step function is defined as: $u(n) = \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases}$

or in a matrix form

$$\mathbf{C}^{-1} = \mathrm{SNR}\cdot \begin{pmatrix} L+N^2-2N & L-N & L-N & \cdots & L-N \\ L-N & L+N^2-4N+2 & L-N & \cdots & L-N \\ L-N & L-N & L+N^2-6N+6 & \cdots & L-N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ L-N & L-N & L-N & \cdots & \end{pmatrix}_{L\times L}.$$

First explicitly computing $\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1}$, one obtains

$$\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1} = (a+b)\cdot \mathrm{SNR} \tag{A.61}$$

where

$$\begin{aligned} a &= \sum_{m=1}^{L} L+N^2-2mN+m(m-1) \\ &= L^2+LN^2-2N\sum_{m=1}^{L} m+\sum_{m=1}^{L} m^2-\sum_{m=1}^{L} m \\ &= (2L^3+6L^2-2L-6L^2N+6LN^2-6LN)/6 \end{aligned}$$

and

$$b = L(L-1)(L-N) = L^3-L^2-LN^2+LN$$

therefore,

$$\mathbf{1}^T\mathbf{C}^{-1}\mathbf{1} = \frac{L}{3}(4L^2-6LN+3N^2-1)\cdot \mathrm{SNR}. \tag{A.62}$$

Next $\mathbf{1}^T\mathbf{C}^{-1}\mathbf{\Delta}$ is computed as

$$\begin{aligned} \mathbf{1}^T\mathbf{C}^{-1}\mathbf{\Delta} &= \sum_{m=1}^{L} \left[L+N^2-m(2N+1)+m^2+(L-N)(L-1)\right]\cdot \Delta(m) \\ &= \sum_{m=1}^{L} \left[(N-m)(N-m+1)-L(N-L)\right]\cdot \Delta(m). \end{aligned} \tag{A.63}$$

Then Mengali's estimator reads as follows

$$\hat{f}_0 = \frac{1}{2\pi} \sum_{m=1}^{L} \frac{3\left[(N-m)(N-m+1) - L(N-L)\right]}{L(4L^2 - 6LN + 3N^2 - 1)} \cdot \Delta(m). \tag{A.64}$$

Its estimation variance can be obtained as

$$\mathrm{var}(\hat{f}_0) = \frac{1}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} = \frac{3 \cdot \mathrm{SNR}^{-1}}{L(4L^2 - 6LN + 3N^2 - 1)}. \tag{A.65}$$

For $L = N/2$, the estimation variance is equivalent to the CRLB.

A.1.10 Variance of correlation-based estimators

Recall the signal model of a constant amplitude

$$y(n) = ae^{j\omega_o n} + v(n), \quad n = 0, \ldots, N-1 \tag{A.66}$$

where $a = |a|e^{j\phi}$ is a constant complex-valued amplitude, $\omega_o \in [-\pi, \pi)$ is the (angular) frequency to be estimated, and $v(n)$ is a zero-mean complex-valued circular white Gaussian noise with variance σ_v^2. The signal-to-noise ratio (SNR) is defined as $|a|^2/\sigma_n^2$. The Cramer-Rao Lower Bound (CRLB) is $\mathrm{CRLB}(\hat{\omega}_o) = 6/(\mathrm{SNR} \cdot N(N^2-1))$.

The theoretical correlation sequence of (A.66) is given by

$$r(m) = E\{y(n)y^*(n-m)\} = |a|^2 e^{j\omega_o m} + \sigma_v^2 \delta_{m,0} \tag{A.67}$$

where $E\{\cdot\}$ denotes the expectation, $\delta_{m,0}$ is the Kronecker delta, and $(\cdot)^*$ denotes the complex conjugate.

The consistent and unbiased estimate of $r(m)$ [76] is

$$\hat{r}(m) = \frac{1}{N-m} \sum_{n=m}^{N-1} y(n)y^*(n-m), \quad m = 0, 1, \ldots, M \tag{A.68}$$

where M is a design parameter not greater than N.

To associate $\hat{r}(m)$ with ω_o, (A.66) is re-written in the form

$$y(n) = e^{j\omega_o n}[a + \tilde{v}(n)] \tag{A.69}$$

where $\tilde{v}(n) = v(n)e^{-j\omega_o n}$, $\tilde{v}(n)$ and $v(n)$ have the same statistical characteristics. Inserting (A.69) into (A.68) yields

$$\hat{r}(m) = |a|^2 e^{j\omega_o m}[1 + \gamma(m)] \tag{A.70}$$

with

$$\begin{aligned}
\gamma(m) &= \frac{1}{(N-m)} \sum_{n=m}^{N-1} \left[\frac{\tilde{v}^*(n-m)}{a^*} + \frac{\tilde{v}(n)}{a} + \frac{\tilde{v}(n)\tilde{v}^*(n-m)}{|a|^2} \right] \\
&= K \sum_{n=m}^{N-1} [e_R(n) + e_R(n-m) + e_R(n)e_R(n-m) + e_I(n)e_I(n-m)] \\
&+ jK \sum_{n=m}^{N-1} [e_I(n) - e_I(n-m) + e_I(n)e_R(n-m) - e_R(n)e_I(n-m)] \\
&= \gamma_R(m) + j\gamma_I(m)
\end{aligned} \tag{A.71}$$

where $K = 1/(N-m)$, $e(n) = a^{-1}\tilde{v}(n) = e_R(n) + je_I(n)$, and $\gamma_R(m)$ and $\gamma_I(m)$ are zero-mean independent, identically distributed, real Gaussian random variables.

By using the additive noise to phase noise conversion [18], the correlation estimate is approximated, at high SNR, as

$$\hat{r}(m) \approx \sigma_a^2 e^{j(\omega_o m + \gamma_I(m))} \tag{A.72}$$

where

$$\gamma_I(m) = K \sum_{n=m}^{N-1} \underbrace{[e_I(n) - e_I(n-m) + e_I(n)e_R(n-m) - e_R(n)e_I(n-m)]}_{g(n)}. \tag{A.73}$$

From (A.72), one can relate ω_o to $\hat{r}(m)$ either by using the absolute phase

$$\begin{aligned} \varphi(m) &= \arg\{\hat{r}(m)\} \\ &\approx [\omega_o m + \gamma_I(m)]_{-\pi}^{\pi} \end{aligned} \tag{A.74}$$

or the phase difference

$$\begin{aligned} \Delta\varphi(m) &= \arg\{\hat{r}(m)\hat{r}^*(m-1)\} \\ &\approx [\omega_o + \gamma_I(m) - \gamma_I(m-1)]_{-\pi}^{\pi} \end{aligned} \tag{A.75}$$

where $\arg\{\cdot\}$ is the complex argument, and $[x]_{-\pi}^{\pi} \in [-\pi, \pi)$. Equation (A.74) and (A.75) indicate that $\varphi(m)$ and $\Delta\varphi(m)$ consist of deterministic components (linearly proportional to ω_o) and zero-mean random processes. Thus, the sets $\{\varphi(m)\}$ and $\{\Delta\varphi(m)\}$ can be used to derive the estimators of ω_o.

Absolute phase based estimator

The general form of the absolute phase based estimator reads

$$\begin{aligned} \hat{\omega}_o &= \sum_{m=1}^{M} \alpha(m)\varphi(m) \\ &\approx \omega_o \sum_{m=1}^{M} m\alpha(m) + \sum_{m=1}^{M} \alpha(m)\gamma_I(m) \end{aligned} \tag{A.76}$$

where $\alpha(m)$ is the weighting function. This class of estimators is unbiased when $\sum_{m=1}^{M} m\alpha(m) = 1$ and $\omega_o \in [-\frac{\pi}{M}, \frac{\pi}{M})$. Thus, its estimation range is approximately $\pm\pi/M$. Outside this range, (A.76) becomes highly nonlinear and awkward to handle. To transform the nonlinearity into a linear equation over the entire range an additional phase unwrapping procedure is needed, however this would be difficult at low SNR.

The estimation variance, $\mathrm{var}(\hat{\omega}_o) \triangleq E\{(\hat{\omega}_o - E\{\hat{\omega}_o\})^2\}$, of (A.76) for

SNR $\gg 1$ reads

$$\begin{aligned}\mathrm{var}(\hat{\omega}_o) &= \sum_{m=1}^{M}\sum_{n=1}^{M}\alpha(m)\alpha(n)\underbrace{E\left\{\gamma_I(m)\gamma_I(n)\right\}}_{\mu(m,n)} \\ &\approx \sum_{m=1}^{M}\sum_{n=1}^{M}\frac{\alpha(m)\alpha(n)\min(m,n,N-m,N-n)}{\mathrm{SNR}(N-m)(N-n)}.\end{aligned} \tag{A.77}$$

where

$$\begin{aligned}\mu(m,n) &= E\left\{\gamma_I(m)\gamma_I(n)\right\} \\ &= \frac{1}{(N-m)}\frac{1}{(N-n)}\sum_{k=m}^{M}\sum_{l=n}^{M}E\left\{g(k)g(l)\right\}\end{aligned} \tag{A.78}$$

with

$$\begin{aligned}E\left\{g(k)g(l)\right\} &= E\{\left[e_I(k)-e_I(k-m)+e_I(k)e_R(k-m)-e_R(k)e_I(k-m)\right] \\ &\qquad \cdot\left[e_I(l)-e_I(l-m)+e_I(l)e_R(l-m)-e_R(l)e_I(l-m)\right]\} \\ &= \frac{1}{2\mathrm{SNR}}(\delta_{k-m,l-n}+\delta_{k-m,l}+\delta_{k,l-n}+\delta_{k,l}) \\ &+\frac{1}{4\mathrm{SNR}^2}(\delta_{k,l}\delta_{k-m,l-n}-\delta_{k,l-n}\delta_{l,k-m}-\delta_{k,l-n}\delta_{l,k-m}+\delta_{k,l}\delta_{k-m,l-n}) \\ &= \frac{1}{2\mathrm{SNR}}(\delta_{k-m,l-n}+\delta_{k-m,l}+\delta_{k,l-n}+\delta_{k,l})+\frac{1}{2\mathrm{SNR}^2}(\delta_{k,l}\delta_{k-m,l-n}).\end{aligned} \tag{A.79}$$

Inserting (A.79) into (A.78) yields

$$\begin{aligned}\mu(m,n) &= \frac{1}{2(N-m)(N-n)}\times \\ &\frac{1}{\mathrm{SNR}}\underbrace{\sum_{k=m}^{M}\sum_{l=n}^{M}(\delta_{k-m,l-n}+\delta_{k-m,l}+\delta_{k,l-n}+\delta_{k,l})}_{2\min(m,n,N-m,N-n)}+\frac{1}{\mathrm{SNR}^2}\underbrace{\sum_{k=m}^{M}\sum_{l=n}^{M}(\delta_{k,l}\delta_{k-m,l-n})}_{(N-m)\delta_{m,n}} \\ &= \frac{\mathrm{SNR}^{-1}}{2(N-m)(N-n)}\left\{2\min(m,n,N-m,N-n)+\frac{1}{\mathrm{SNR}}(N-m)\delta_{m,n}\right\} \\ &\approx \frac{\min(m,n,N-m,N-n)}{\mathrm{SNR}(N-m)(N-n)}, \quad \text{for SNR} \gg 1 \\ &= \frac{\min(m,n)}{\mathrm{SNR}(N-m)(N-n)}, \qquad \text{if } M \le N/2.\end{aligned} \tag{A.80}$$

Phase difference based estimator

The general form of estimators derived from (A.75) can be written as

$$\begin{aligned}\hat{\omega}_o &= \sum_{m=1}^{M}\beta(m)\Delta\varphi(m) \\ &\approx \omega_o\sum_{m=1}^{M}\beta(m)+\sum_{m=1}^{M}\beta(m)(\gamma_I(m)-\gamma_I(m-1))\end{aligned} \tag{A.81}$$

where $\beta(m)$ is the weighting function. This class of estimators is unbiased when $\sum_{m=1}^{M}\beta(m)=1$ and ω_o lies within $[-\pi,\pi)$. Thus, its estimation range is about $\pm\pi$ or M times larger than that of (A.76) without the need of phase unwrapping algorithm.

The estimation variance of (A.81) is

$$\begin{aligned}\mathrm{var}(\hat{\omega}_o) &= \sum_{m=1}^{M}\sum_{n=1}^{M}\beta(m)\beta(n)\times \\ &\underbrace{E\left\{[\gamma_I(m)-\gamma_I(m-1)][\gamma_I(n)-\gamma_I(n-1)]\right\}}_{\rho(m,n)}\end{aligned} \tag{A.82}$$

where $\rho(m,n) = \mu(m,n) + \mu(m-1,n-1) - \mu(m-1,n) - \mu(m,n-1)$.

A.2 Complex-Valued Envelope

A.2.1 Generalized Cramer-Rao lower bound

In the Rician case, the unknown parameter vector is

$$\boldsymbol{\theta} = [\omega_o, \phi, \mu, \boldsymbol{\alpha}^T, \sigma_v^2]^T.$$

In the Rayleigh case, $\mu = 0$, the parameter vector reduces to

$$\boldsymbol{\theta} = [\omega_o, \boldsymbol{\alpha}^T, \sigma_v^2]^T.$$

The covariance matrix of any unbiased estimate cannot be lower than the CRLB which is given as the inverse of the Fisher information matrix (FIM)

$$\mathrm{E}\left\{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\right\} \geq \mathrm{CRLB}(\boldsymbol{\theta}) = \mathbf{J}_{\boldsymbol{\theta},\boldsymbol{\theta}}^{-1}$$

where $A \geq B$ means that the matrix $A - B$ is positive semi-definite.

For a finite-sample, the fading envelope and observed data vector may be defined as

$$\mathbf{a} = [a(0), \ldots, a(N-1)]^T, \quad \mathbf{y} = [y(0), \ldots, y(N-1)]^T.$$

The vector $\mathbf{y}$ is Gaussian with mean

$$\mathbf{m}_y = \mu e^{j\varphi} \Gamma \mathbf{1}$$

where $\mathbf{1}$ is a $(N \times 1)$ vector of one, and

$$\Gamma = \mathrm{diag}\{e^{j\omega_o n}, \quad n = 0, \ldots, N-1\}.$$

Let $\tilde{\mathbf{y}} = \mathbf{y} - \mathbf{m}_y$; since $\tilde{\mathbf{y}}$ is zero-mean and complex Gaussian, its statistics are completely characterized by the two correlation matrices

$$\mathbf{R}_y \triangleq \mathrm{E}\{\tilde{\mathbf{y}}\tilde{\mathbf{y}}^H\}; \quad \mathbf{U}_y \triangleq \mathrm{E}\{\tilde{\mathbf{y}}\tilde{\mathbf{y}}^T\}.$$

Since $a(k)$ and $n(k)$ are mutually independent, these matrices are given by

$$\mathbf{R}_y = \Gamma\mathbf{R}_a\Gamma^H + \sigma_n^2\mathbf{I}, \quad \mathbf{U}_y = \Gamma\mathbf{U}_a\Gamma^H + \sigma_n^2\mathbf{I}$$

where $\mathbf{R}_a = \mathrm{E}\{\tilde{\mathbf{a}}\tilde{\mathbf{a}}^H\}$, $\mathbf{U}_a = \mathrm{E}\{\tilde{\mathbf{a}}\tilde{\mathbf{a}}^T\}$, and $\mathbf{I}$ is the $(N \times N)$ identity matrix. under the assumption that $a(k)$ is circular around its mean so that $\mathbf{U}_a$ vanishes, and the matrix $\mathbf{U}_y$ consists of then identically zeros. This implies that $\tilde{\mathbf{y}}$ is circularly symmetric.

Finite-sample CRLB

The FIM elements can be expressed as [74] (see also [34])

$$\begin{aligned} J_{\theta_i,\theta_j} &= \mathrm{tr}\left\{\mathbf{R}_y^{-1}\frac{\partial\mathbf{R}_y}{\partial\theta_i}\mathbf{R}_y^{-1}\frac{\partial\mathbf{R}_y}{\partial\theta_j}\right\} + 2\mathrm{Re}\left\{\frac{\partial\mathbf{m}_y^H}{\partial\theta_i}\mathbf{R}_y^{-1}\frac{\partial\mathbf{m}_y}{\partial\theta_j}\right\} \\ &= J_R(\theta_i,\theta_j) + J_m(\theta_i,\theta_j) \end{aligned} \tag{A.83}$$

where $\mathrm{Re}\{u\}$ is the real part of u, and $\mathrm{tr}\{\cdot\}$ is the trace operator.

The partial derivatives of $\mathbf{m}_y$ and $\mathbf{R}_y$ are given by

$$\begin{aligned} \frac{\partial\mathbf{m}_y}{\partial\omega_o} &= j2\mu e^{j\varphi}\mathbf{D}\Gamma\mathbf{1} \\ \frac{\partial\mathbf{m}_y}{\partial\mu} &= e^{j\varphi}\Gamma\mathbf{1} \\ \frac{\partial\mathbf{m}_y}{\partial\varphi} &= j\mu e^{j\varphi}\Gamma\mathbf{1} \\ \frac{\partial\mathbf{R}_y}{\partial\omega_o} &= j2[\mathbf{D}\Gamma\mathbf{R}_a\Gamma^H - \Gamma\mathbf{R}_a\Gamma^H\mathbf{D}] \\ \frac{\partial\mathbf{R}_y}{\partial\sigma_n^2} &= \mathbf{I} \\ \frac{\partial\mathbf{R}_y}{\partial\alpha_l} &= \frac{\partial\mathbf{R}_a}{\partial\alpha_l}\Gamma^H, \quad l = 1,\ldots,p \end{aligned} \tag{A.84}$$

where $\mathbf{D} = \text{diag}\{0, 1, \ldots, N-1\}$. Note that $\mathbf{R}_a$ is not a function of μ and φ, and similarly $\mathbf{m}_y$ is not a function of $\boldsymbol{\alpha}$ and σ_n^2.

Computation of J_{ω_o,ω_o}: Using (A.84), $J_m(\omega_o, \omega_o)$ can be written as

$$\begin{aligned} J_m(\omega_o, \omega_o) &= 2\mu^2 \mathbf{1}^T \mathbf{D} \Gamma^H \mathbf{R}_y^{-1} \Gamma \mathbf{D} \mathbf{1} \\ &= 2\mu^2 \mathbf{1}^T \mathbf{D} (\mathbf{R}_a + \sigma_n^2 \mathbf{I})^{-1} \mathbf{D} \mathbf{1} \end{aligned}$$

since $\Gamma^H \Gamma = \mathbf{I}$. Let $\mathbf{R} = \mathbf{R}_a + \sigma_n^2 \mathbf{I}$, using (A.84) yields

$$\begin{aligned} J_R(\omega_o, \omega_o) = \text{tr}\{&\Gamma \mathbf{R}^{-1} \Gamma^H (j\Gamma \mathbf{D} \mathbf{R}_a \Gamma^H - j\Gamma \mathbf{R}_a \mathbf{D} \Gamma^H) \\ &\cdot \Gamma \mathbf{R}^{-1} \Gamma^H (j\Gamma \mathbf{D} \mathbf{R}_a \Gamma^H - j\Gamma \mathbf{R}_a \mathbf{D} \Gamma^H)\}. \end{aligned}$$

Since $\mathbf{R}_a \mathbf{R}^{-1} = \mathbf{R}^{-1} \mathbf{R}_a$ and $\text{tr}(AB) = \text{tr}(BA)$, we obtain

$$J_R(\omega_o, \omega_o) = 2\text{tr}\{\mathbf{R}^{-1} \mathbf{D} \mathbf{R} \mathbf{D}\} - 2\text{tr}\{\mathbf{D}^2\}.$$

Computation of $J_{\omega_o,\boldsymbol{\alpha}}$: Since $\mathbf{m}_y$ is independent of $\boldsymbol{\alpha}$, $J_m(\omega_o, \alpha_l) = 0$, and $J_{\omega_o,\alpha_l} = J_R(\omega_o, \alpha_l)$, $l = 1, \ldots, p$. Using (A.84), we have

$$J_{\omega_o,\alpha_l} = j\text{tr}\left\{\mathbf{R}^{-1} \mathbf{D} \mathbf{R}_a \mathbf{R}^{-1} \frac{\partial \mathbf{R}_a}{\partial \alpha_l}\right\} - j\text{tr}\left\{\mathbf{R}^{-1} \mathbf{R}_a \mathbf{D} \mathbf{R}^{-1} \frac{\partial \mathbf{R}_a}{\partial \alpha_l}\right\}, \; l = 1, \ldots, p.$$

Since the trance operator is commutative, $\mathbf{R}_a \mathbf{R}^{-1} = \mathbf{R}^{-1} \mathbf{R}_a$, and $\mathbf{D}$ is real-valued, we obtain

$$J_{\omega_o,\alpha_l} = -2\text{Im}\left\{\text{tr}\left\{\mathbf{D} \frac{\partial \mathbf{R}_a}{\partial \alpha_l} \mathbf{R}^{-1}\right\}\right\}, \quad l = 1, \ldots, p.$$

The derivation of the other FIM entries is straightforward and, hence omitted.

However, the results are summarized as follows

$$
\begin{aligned}
J_{\omega_o,\omega_o} &= 2\mu^2 \mathbf{1}^T \mathbf{D}\mathbf{R}^{-1}\mathbf{D}\mathbf{1} + 2\mathrm{tr}\{\mathbf{R}^{-1}\mathbf{D}\mathbf{R}\mathbf{D} - \mathbf{D}^2\} \\
J_{\omega_o,\varphi} &= 2\mu^2 \mathrm{Re}\left\{\mathbf{1}^T\mathbf{R}^{-1}\mathbf{D}\mathbf{1}\right\} \\
J_{\omega_o,\alpha_l} &= -2\mathrm{Im}\left\{\mathrm{tr}\left\{\mathbf{D}\frac{\partial \mathbf{R}_a}{\partial \alpha_l}\mathbf{R}^{-1}\right\}\right\}, \quad l = 1,\ldots,p \\
J_{\varphi,\varphi} &= 2\mu^2 \mathbf{1}^T\mathbf{R}^{-1}\mathbf{1} \\
J_{\mu,\mu} &= 2\mathbf{1}^T\mathbf{R}^{-1}\mathbf{1} \\
J_{\alpha_k,\alpha_l} &= \mathrm{tr}\left\{\mathbf{R}^{-1}\frac{\partial \mathbf{R}_a}{\partial \alpha_k}\mathbf{R}^{-1}\frac{\partial \mathbf{R}_a}{\partial \alpha_l}\right\}, \quad k,l = 1,\ldots,p \\
J_{\alpha_k,\sigma_n^2} &= \mathrm{tr}\left\{\mathbf{R}^{-1}\frac{\partial \mathbf{R}_a}{\partial \alpha_k}\mathbf{R}^{-1}\right\}, \quad k = 1,\ldots,p \\
J_{\sigma_n^2,\sigma_n^2} &= \mathrm{tr}\left\{\mathbf{R}^{-2}\right\}
\end{aligned}
$$

where $\mathbf{R} = \mathbf{R}_a + \sigma_n^2\mathbf{I}$, $\mathbf{D} = \mathrm{diag}\{0,\ldots,N-1\}$, and $\mathrm{Im}\{u\}$ is the imaginary part of u. The other elements are zeros.

Large-sample FIM

For large N, the first term of J_{ω_o,ω_o} is given by

$$
J_m(\omega_o,\omega_o) = 2\mu^2\mathbf{1}^T\mathbf{D}\mathbf{R}^{-1}\mathbf{D}\mathbf{1} \approx \frac{2\mu^2 N^3/3}{S_a(0)+\sigma_n^2} \tag{A.85}
$$

where $S_a(\omega)$ is the Power Spectral Density (PSD) of $a(k)$. Since $y(k)$ is Gaussian and circular around its mean, the first term of the FIM expression in (A.83) is given bz Whittle's formula for large N

$$
\mathrm{tr}\left\{\mathbf{R}_y^{-1}\frac{\partial \mathbf{R}_y}{\partial \theta_i}\mathbf{R}_y^{-1}\frac{\partial \mathbf{R}_y}{\partial \theta_j}\right\} \approx N\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\frac{\partial S_y(\omega)}{\partial \theta_i}\frac{\partial S_y(\omega)}{\partial \theta_j}}{S_y^2(\omega)}d\omega
$$

where $S_y(\omega) = S_a(\omega-\omega_o) + \sigma_n^2$ is the PSD of $y(k)$. Therefore, for large N, we have that

$$
J_{\omega_o,\omega_o} = \frac{2\mu^2 N^3/3}{S_a(0)+\sigma_n^2} + \frac{N}{2\pi}\int_{-\pi}^{\pi}\left(\frac{S_a'(\omega)}{S_a(\omega)+\sigma_n^2}\right)^2 d\omega
$$

where $S'_a(\omega) = \frac{dS_a(\omega)}{d\omega}$. The other FIM entries are obtained similarly.

A.2.2 Estimates of Correlation Sequence

For a given finite record of a random signal $x(k)$, this sequence may be represented by

$$v(k) = \begin{cases} x(k) & \text{for } 1 \leq k \leq N \\ 0 & \text{otherwise.} \end{cases} \tag{A.86}$$

Consider an estimate of the correlation sequence as

$$\hat{r}(m) = \frac{1}{N} c(m) \tag{A.87}$$

where, since $c(-m) = c^*(m)$

$$\begin{aligned} c(m) &= \sum_{k=1}^{N} v^*(k) v(k+m) \\ &= \begin{cases} \sum_{k=1}^{N-|m|} x^*(k) x(k+|m|), & |m| \leq N \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} \sum_{k=1+|m|}^{N} x(k) x^*(k-|m|), & |m| \leq N \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{A.88}$$

corresponding to the aperiodic correlation of a rectangularly windowed segment of $x(k)$.

To determine the properties of this estimate of the correlation sequence, the mean and variance of the random variable $\hat{r}(m)$ are considered. From the above equations, it follows that

$$\begin{aligned} E\{\hat{r}(m)\} &= \frac{1}{N} \sum_{k=1+|m|}^{N} E\{x(k) x^*(k-|m|)\} \\ &= \frac{1}{N} \sum_{k=1+|m|}^{N} r(m) \end{aligned} \tag{A.89}$$

and since $r(m)$ does not depend on k for a stationary random process, then

$$E\{\hat{r}(m)\} = \begin{cases} \frac{N-|m|}{N} r(m), & |m| \leq N \\ 0, & \text{otherwise.} \end{cases} \tag{A.90}$$

From (A.90), it can be seen that $\hat{r}(m)$ is a biased estimate of $r(m)$, since $E\{\hat{r}(m)\} \neq r(m)$, but the bias is small if $|m| \ll N$. It can be directly seen that an unbiased estimator of the correlation sequence for $|m| \leq N-1$ is

$$\breve{r}(m) = \frac{1}{N-|m|} c(m) \tag{A.91}$$

i.e., the estimator is unbiased if we divide by the number of nonzero terms in the sum of lagged products rather than by the total number of samples in the data record.

The variance of the correlation function estimates is difficult to derive, even if using simplifying assumptions. However, approximate formulas for the variance of both biased and unbiased correlation estimates can be found in [77] and are given here without proof. The variance of the unbiased estimate is

$$\text{var}\{\hat{r}(m)\} \approx \frac{N}{(N-|m|)^2} \sum_{k=-\infty}^{\infty} (r^2(m) + r(k+m)r(k-m)) \tag{A.92}$$

and the variance of the biased estimate is

$$\text{var}\{\breve{r}(m)\} \approx \frac{1}{N} \sum_{k=-\infty}^{\infty} (r^2(m) + r(k+m)r(k-m)). \tag{A.93}$$

It can be seen that the variance of $\breve{r}(m)$ for larger lags does not increase as much as the variance of $\hat{r}(m)$. The mean of both types of correlation estimators are asymptotically unbiased (for fixed lags) but it was suggested in several references that for many practical applications the biased correlation estimator is more desirable because of the following reasons:

- For most stationary signals, the correlation function decays rather rapidly, so that $r(m)$ is quite small for larger lags m. Comparing the definitions (A.87) and (A.91), it can be seen that $\hat{r}(m)$ will be smaller

for large m (provided that N is reasonably large), whereas $\breve{r}(m)$ may take large and erratic values for large m, as it is obtained by averaging only few products in such a case (in particular, only one product for $m = N-1$). This observation implies that the biased estimator $\hat{r}(m)$ is likely to be a more accurate estimator of $r(m)$, than $\breve{r}(m)$, for medium and large values of m (compared to N). For small values of m, the two estimators can be expected to behave in a similar manner.

- The biased correlation sequence $\{\hat{r}(m),\ m = 0, \pm 1, \pm 2, \dots\}$ is guaranteed to be positive semidefinite, while this is not the case for $\breve{r}(m)$. This condition is particularly important for spectral estimation. The non positive definiteness of a correlation sequence may lead to a negative spectral estimates which is undesirable in most applications.

Appendix B

Kalman Filters

The Kalman filter in its various forms has become a fundamental tool for analyzing and solving broad class of estimation problems. It was originally derived for linear systems. The Kalman filter was developed by Rudolph Kalman [78] in 1960, although Peter Swerling developed a similar algorithm in 1958. The filter is named after Kalman because he published his results in a more prestigious journal and his work was more general and complete. The standard Kalman filter is limited to linear systems. Most of the real-world systems are nonlinear in which linear Kalman filters cannot be directly applied. However, early 1960s after the discovery of the Kalman filter, the first attempt to extend the linear Kalman filter to the nonlinear problems was successfully carried out at NASA during feasibility studies for circumlinear navigation and control of the Apollo space capsule. The Kalman filter is also referred to as the Kalman-Bucy filter because of Richard Bucy's early work on the topic, conducted jointly with Kalman. The roots of the Kalman filter can be traced all the way back to the 18-year old Karl Gauss's method of least squares in 1795.

This section summarizes the use of the linear Kalman filter and its extensions. To derive the Kalman filter, there is more than one way, each which its own criteria of optimality. The derivation of the Kalman filter based on the recursive Bayes estimation is one of the most classical ways. The Bayesian approach attempts to construct the posterior probability function of the state based on all available information, including the sequence of measurements. Since the probability density function contains all available statistical information, it may be regarded as the complete solution to the estimation problem.

B.1 Linear Kalman filter

B.1.1 State space model

The state evolves according to the following discrete-time stochastic model

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}_{k-1}(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}) \\ &= \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_{k-1} \end{aligned} \tag{B.1}$$

where $\mathbf{f}_{k-1}$ is a known function ($\equiv \mathbf{F}_{k-1}$ if linear) of the state $\mathbf{x}_{k-1}$ and $\mathbf{v}_{k-1}$ is regarded as a process noise sequence. The recursive filter estimates $\mathbf{x}_k$ from the measurements $\mathbf{z}_k$ which are related to the state via the measurement equations

$$\begin{aligned} \mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k, \mathbf{w}_k) \\ &= \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k \end{aligned} \tag{B.2}$$

where $\mathbf{h}_k$ is a known function ($\equiv \mathbf{H}_k$ if linear) and $\mathbf{w}_k$ is a measurement noise. Random sequences $\mathbf{v}_{k-1}$ and $\mathbf{w}_k$ are mutually independent zero-mean white Gaussian, with covariances $\mathbf{Q}_{k-1}$ and $\mathbf{R}_k$, respectively.

B.1.2 Recursive Bayes estimation

The recursive Bayes estimation considers the problem of estimating the state $\mathbf{x}_k$ of a system using observations $\mathbf{y}_k$ of the systems, where the state evolves in the presence of noise and the observation are made sequentially in the presence of noise. The development of the recursive Bayes estimation is based upon the assumptions that the state sequence $\mathbf{x}_k$ is a Markov random process of order one, that is

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}, \cdots, \mathbf{x}_0) = p(\mathbf{x}_k|\mathbf{x}_{k-1}) \tag{B.3}$$

and the observation $\mathbf{y}_k$ depends upon $\mathbf{x}_k$ and possibly also on random noise that is independent[1] from sample to sample, but conditionally independent of prior observations, given $\mathbf{x}_k$; that is $\mathbf{x}_k$ is a Markov random process of

[1]Not necessary, also correlated noise can be handled by introducing form filters expressed in state space and then including that noise state in the overall augmented state

order one, that is

$$p(\mathbf{y}_k|\mathbf{x}_k, \mathbf{y}_{k-1}, \cdots, \mathbf{y}_0) = p(\mathbf{y}_k|\mathbf{x}_k). \tag{B.4}$$

Using Bayes theorem and the above assumptions results in two useful recursive equations. They are known as the *predict* and *update* equations of a probability density function. The update equation is

$$\underbrace{p(\mathbf{x}_k|\boldsymbol{\Upsilon}_k)}_{posterior} = \frac{p(\mathbf{y}_k|\mathbf{x}_k)}{p(\mathbf{y}_k|\boldsymbol{\Upsilon}_{k-1})} \underbrace{p(\mathbf{x}_k|\boldsymbol{\Upsilon}_{k-1})}_{prior} \tag{B.5}$$

where $\boldsymbol{\Upsilon}_k = \mathbf{y}_k, \mathbf{y}_{k-1}, \cdots, \mathbf{y}_0$. The computation of (B.5) requires finding $p(\mathbf{x}_k|\boldsymbol{\Upsilon}_{k-1})$ which can be obtained from the prediction equation below

$$p(\mathbf{x}_k|\boldsymbol{\Upsilon}_{k-1}) = \int p(\mathbf{x}_k|\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1}|\boldsymbol{\Upsilon}_{k-1}) d\mathbf{x}_{k-1}. \tag{B.6}$$

In going from one stage to the next, the conditional $p(\mathbf{x}_k|\boldsymbol{\Upsilon}_{k-1})$ becomes the prior to the next stage. To preserve the computational structure from one stage to the next, it is expedient to have the conditional density be the same type as the prior density.

The recurrence relations (B.5) and (B.6) form the basis for the optimal Bayesian solutions. Knowledge of posterior density $p(\mathbf{x}_k|\boldsymbol{\Upsilon}_k)$ enables one to compute an optimal state estimate with respect to any criterion. For example, the minimum mean-square error (MMSE) estimate is the conditional mean of $\mathbf{x}_k$, while the maximum a posteriori (MAP) estimate is the maximum of $p(\mathbf{x}_k|\boldsymbol{\Upsilon}_k)$.

B.1.3 Linear Kalman filter algorithm

Deriving the Kalman filter from the Bayesian point of view [46] [47] which is a natural outgrowth of the results in the above sections, yields the following steps

Initialization:

$$\begin{aligned} \hat{\mathbf{x}}_0^- &= \mathrm{E}\{\mathbf{x}_0\} \\ \mathbf{P}_0^- &= \mathrm{E}\{(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T\} \end{aligned}$$

Time propagation (Prediction):

$$\begin{aligned}\hat{\mathbf{x}}_{k+1}^{-} &= \mathbf{F}_k \hat{\mathbf{x}}_k^{+} + \mathbf{u}_k \\ \mathbf{P}_{k+1}^{-} &= \mathbf{F}_k \mathbf{P}_k^{+} \mathbf{F}_k^{T} + \mathbf{Q}_k\end{aligned}$$

Measurement update (Correction):

$$\begin{aligned}\mathbf{K}_{k+1} &= \mathbf{P}_{k+1}^{-} \mathbf{H}_{k+1}^{T} [\mathbf{R}_{k+1} + \mathbf{H}_{k+1} \mathbf{P}_{k+1}^{-} \mathbf{H}_{k+1}^{T}]^{-1} \\ \hat{\mathbf{x}}_{k+1}^{+} &= \hat{\mathbf{x}}_{k+1}^{-} + \mathbf{K}_{k+1} [\mathbf{y}_{k+1} - \mathbf{H}_{k+1} \hat{\mathbf{x}}_{k+1}^{-}] \\ \mathbf{P}_{k+1}^{+} &= [\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1}] \mathbf{P}_{k+1}^{-}\end{aligned}$$

The equivalent form of $\hat{\mathbf{x}}_{k+1}^{+}$ and $\mathbf{P}_{k+1}^{+}$ can be written as

$$\begin{aligned}\hat{\mathbf{x}}_{k+1}^{+} &= \mathbf{P}_{k+1}^{+} (\mathbf{P}_{k+1}^{-})^{-1} \hat{\mathbf{x}}_{k+1}^{-} + \mathbf{P}_{k+1}^{+} \mathbf{H}_{k+1}^{T} (\mathbf{R}_{k+1})^{-1} \mathbf{y}_{k+1} \\ \mathbf{P}_{k+1}^{+} &= \left[(\mathbf{P}_{k+1}^{-})^{-1} + \mathbf{H}_{k+1}^{T} (\mathbf{R}_{k+1})^{-1} \mathbf{H}_{k+1} \right]^{-1}.\end{aligned}$$

The above equations are useful for the initialization of the state vector and state error covariance matrix when the a priori information is not available. Let $k = 0$ and since there is information on $(\mathbf{P}_1^{-})^{-1} = 0$

$$\begin{aligned}\mathbf{P}_1^{+} &= \left[\mathbf{H}_1^{T} (\mathbf{R}_1)^{-1} \mathbf{H}_1 \right]^{-1} \\ \hat{\mathbf{x}}_1^{+} &= \mathbf{P}_1^{+} \mathbf{H}_1^{T} (\mathbf{R}_1)^{-1} \mathbf{y}_1.\end{aligned}$$

For initialization purpose $\hat{\mathbf{x}}_0^{+} \triangleq \hat{\mathbf{x}}_1^{+}$ and $\mathbf{P}_0^{+} \triangleq \mathbf{P}_1^{+}$.

B.2 Nonlinear Kalman filters

The key element to nonlinear Kalman filtering is the expansion of the nonlinear terms of the system equation using a Taylor series expansion around a nominal point $\bar{x}$. A Taylor series expansion of a nonlinear function can be written as

$$f(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + f''(\bar{x}) \frac{(x - \bar{x})^2}{2!} + f'''(\bar{x}) \frac{(x - \bar{x})^3}{3!} + \cdots \quad \text{(B.7)}$$

Linearizing a function means expanding it in a *first-order* Taylor series around some nominal point. In other words, the first-order approximation of a function $f(x)$ is $f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x})$. The first-order approximation holds as long as $(x - \bar{x})$ is small enough.

The nominal values may be derived from a predetermined trajectory or from a prediction of the actual trajectory. The first method significantly reduces the computational requirements for real-time applications since most of the influential matrices can be pre-computed prior to the mission. This method is known as *linearized Kalman filter* (LKF). On the other hand, using predicted state is typically closer to the actual state than using a predetermined state as the nominal state. This method is referred to as *extended Kalman filter* (EKF).

B.2.1 Linearized Kalman Filter

The idea of the linearized Kalman filter is to start with the nonlinear system and then find a linear system whose states represent the *deviations* from a nominal state (or trajectory) of the nonlinear system. The nominal trajectory refers to some predetermined dynamics of the system and must be known in advance. For some systems we might know a nominal trajectory ahead of time (for example, an aircraft flight with a predetermined flight plan, or a robot arm moving in a manufacturing environment). The actual trajectory may differ from the nominal trajectory due to mismodeling, disturbances, and other unforseen effects. Nevertheless, the actual trajectory should be close to the nominal trajectory so that the Taylor series approximation is reasonably accurate. The Taylor series linearization of the nonlinear system and measurement equations are

$$\begin{aligned}
\mathbf{x}_{k+1} &= \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k \\
&\approx \mathbf{f}(\bar{\mathbf{x}}_k, \mathbf{u}_k) + \mathbf{f}'(\bar{\mathbf{x}}_k, \mathbf{u}_k)\delta\mathbf{x}_k + \mathbf{w}_k \\
\mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k \\
&\approx \mathbf{h}(\bar{\mathbf{x}}_k) + \mathbf{h}'(\bar{\mathbf{x}}_k)\delta\mathbf{x}_k + \mathbf{v}_k
\end{aligned}$$

The deviations from the nominal trajectory can be written as

$$\begin{aligned}\delta\mathbf{x}_{k+1} &= \mathbf{x}_{k+1} - \mathbf{f}(\bar{\mathbf{x}}_k, \mathbf{u}_k) \\ &= \mathbf{x}_{k+1} - \bar{\mathbf{x}}_{k+1} \\ \delta\mathbf{y}_k &= \mathbf{y}_k - \mathbf{h}(\bar{\mathbf{x}}_k) \\ &= \mathbf{y}_k - \bar{\mathbf{y}}_k \triangleq \mathbf{z}_k\end{aligned}$$

Combining the above equations with the previous equations results in the following state space equations that are linear functions of $\delta\mathbf{x}$, $\delta\mathbf{y}$

$$\begin{aligned}\delta\mathbf{x}_{k+1} &= \mathbf{f}'(\bar{\mathbf{x}}_k, \mathbf{u}_k)\delta\mathbf{x}_k + \mathbf{w}_k \\ \delta\mathbf{y}_k &= \mathbf{h}(\bar{\mathbf{x}}_k)\delta\mathbf{x}_k + \mathbf{v}_k\end{aligned} \tag{B.8}$$

Two important points need to be considered when using linearized Kalman filter:

- After the Kalman filter is used to estimate $\delta\mathbf{x}$, the nominal state vector $\bar{\mathbf{x}}$ must be added to the state estimate of $\delta\mathbf{x}$. This is because $\delta\mathbf{x} = \mathbf{x} - \bar{\mathbf{x}}$.
- If the true state $\mathbf{x}$ gets too far away from the nominal state $\bar{\mathbf{x}}$, then the linearized Kalman filter will not provide good results. This is because the Taylor approximation is inaccurate if the nominal and the true states are too far apart.

B.2.2 Extended Kalman filter

For many applications, the knowledge about the nominal trajectory is not available prior to the mission. It is therefore reasonable to use the predicted state for the linearized process equations and the corrected state for the linearized measurement equations. The EKF algorithm is summarized below

Initialization:

$$\begin{aligned}\hat{\mathbf{x}}_0^- &= \mathrm{E}\{\mathbf{x}_0\} \\ \mathbf{P}_0^- &= \mathrm{E}\{(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T\}\end{aligned}$$

Time propagation (Prediction):

$$\hat{\mathbf{x}}_{k+1}^{-} = \mathbf{f}(\hat{\mathbf{x}}_{k}^{+}, \mathbf{u}_k)$$
$$\mathbf{P}_{k+1}^{-} = \mathbf{\Phi}_k \mathbf{P}_{k}^{+} \mathbf{\Phi}_{k}^{T} + \mathbf{Q}_k$$

Measurement update (Correction):

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1}^{-} \mathbf{H}_{k+1}^{T} [\mathbf{R}_{k+1} + \mathbf{H}_{k+1} \mathbf{P}_{k+1}^{-} \mathbf{H}_{k+1}^{T}]^{-1}$$
$$\hat{\mathbf{x}}_{k+1}^{+} = \hat{\mathbf{x}}_{k+1}^{-} + \mathbf{K}_{k+1} [\mathbf{y}_{k+1} - \mathbf{h}(\hat{\mathbf{x}}_{k+1}^{-})]$$
$$\mathbf{P}_{k+1}^{+} = [\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1}] \mathbf{P}_{k+1}^{-}$$

Definitions:

$$\mathbf{\Phi}_k = \frac{\partial}{\partial \mathbf{x}_k} \mathbf{f}(\mathbf{x}_k)\big|_{\mathbf{x}_k = \hat{\mathbf{x}}_k^{+}} = \mathbf{f}'(\hat{\mathbf{x}}_k^{+})$$
$$\mathbf{H}_k = \frac{\partial}{\partial \mathbf{x}_k} \mathbf{h}(\mathbf{x}_k)\big|_{\mathbf{x}_k = \hat{\mathbf{x}}_k^{-}} = \mathbf{h}'(\hat{\mathbf{x}}_k^{-})$$

Typically, the estimated state (under good observability conditions) is closer to the actual state than a predefined nominal trajectory. Hence, the EKF usually produces a better performance than the LKF.

B.3 Unscented Kalman Filter

In the EKF the state distribution is approximated by a Gaussian random variable (GRV) which is propagated analytically through the first-order linearization of the nonlinear system. As such the EKF can be viewed as providing first-order approximations to the optimal terms (the second-order versions exist but with prohibitive complexity). These approximations, however, can introduce large errors in the true posterior mean and covariance of the transformed random variable, which leads to sub-optimal performance and sometimes divergence of the filter.

The Unscented Kalman filter (UKF), developed by Julier and Uhlman [79], addresses the approximation issues of the EKF. The state distribution is again represented by a GRV, but is now specified using a minimal set of carefully chosen sample points. These sample points completely capture the

true mean and covariance of the GRV, and when propagated through the true nonlinear system, captures the posterior mean and covariance accurately to the 3rd order (Taylor series expansion) for any nonlinearity. The Unscented transformation (UT) is, therefore, the core element to the UKF.

B.3.1 Unscented transformation

The Unscented transformation (UT) is a method for calculating the statistics of a random variable which undergoes a nonlinear transformation. Consider propagating a random variable $\mathbf{x}$ (dimension L) through a nonlinear function, $\mathbf{y} = g(\mathbf{x})$. Assume $\mathbf{x}$ has mean $\bar{x}$ and covariance $\mathbf{P}_{\mathbf{x}}$. To calculate the statistics of $\mathbf{y}$, we form the matrix $\boldsymbol{\mathcal{X}}$ of $2L+1$ *sigma* vectors $\mathcal{X}_i$ (with corresponding weights W_i), according to the following:

$$\begin{aligned}
\mathcal{X}_0 &= \bar{\mathbf{x}} \\
\mathcal{X}_i &= \bar{\mathbf{x}} + \left(\sqrt{(L+\lambda)\mathbf{P}_{\mathbf{x}}}\right)_i \qquad i = 1, \ldots, L \\
\mathcal{X}_i &= \bar{\mathbf{x}} - \left(\sqrt{(L+\lambda)\mathbf{P}_{\mathbf{x}}}\right)_{i-L} \qquad i = L+1, \ldots, 2L \\
W_0^m &= \frac{\lambda}{L+\lambda} \\
W_0^c &= \frac{\lambda}{L+\lambda} + (1 - \alpha^2 + \beta) \\
W_i^m &= W_i^c = \frac{1}{2(L+\lambda)} \qquad i = 1, \ldots, 2L
\end{aligned} \tag{B.9}$$

where $\lambda = \alpha^2(L+\kappa) - L$ is a scaling parameter. α determines the spread of the sigma points around $\bar{\mathbf{x}}$ and is usually set to a small positive value (e.g., 0.001). κ is a secondary scaling parameter whic is usually set to 0, and β is used to incorporate prior knowledge of the distribution of $\mathbf{x}$ (for Gaussian distributions, $\beta = 2$ is optimal). $\left(\sqrt{(L+\lambda)\mathbf{P}_{\mathbf{x}}}\right)_i$ is the ith row of the matrix square root. These sigma vectors are propagated through the nonlinear function,

$$\mathcal{Y}_i = \mathbf{h}(\mathcal{X}_i), \qquad i = 0, \ldots, 2L \tag{B.10}$$

and the mean and covariance for $\mathbf{y}$ are approximated using a weighted sample mean and covariance of the posterior sigma points

$$\begin{aligned} \bar{\mathbf{y}} &= \sum_{i=0}^{2L} W_i^m \mathcal{Y}_i \\ \mathbf{P}_{\mathbf{y}} &= \sum_{i=0}^{2L} W_i^c [\mathcal{Y}_i - \bar{\mathbf{y}}][\mathcal{Y}_i - \bar{\mathbf{y}}]^T \end{aligned} \tag{B.11}$$

Note that this method differs substantially from general Monte-Carlo methods such as particle filters which require orders of magnitude more sample points in an attempt to propagate an accurate (possibly non-Gaussian) distribution of the state. The deceptively simple approach taken with the UT results in approximations that accurate to the third order for Gaussian inputs for all nonlinearities. For non-Gaussian inputs, approximations are accurate to at least second-order, with accuracy of third and higher order moments determined by the choice of α and β (see [80] for more detail of the UT).

B.3.2 Unscented Kalman filter algorithm

Initialization:

$$\begin{aligned} \hat{\mathbf{x}}_0 &= \mathrm{E}\{\mathbf{x}_0\} \\ \mathbf{P}_0 &= \mathrm{E}\{(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T\} \end{aligned}$$

Calculating sigma points:

$$\mathcal{X}_{k-1} = \left[\hat{\mathbf{x}}_{k-1} \quad \hat{\mathbf{x}}_{k-1} + \sqrt{(L+\lambda)\mathbf{P}_{k-1}} \quad \hat{\mathbf{x}}_{k-1} - \sqrt{(L+\lambda)\mathbf{P}_{k-1}}\right]$$

Time update (Prediction):

$$
\begin{aligned}
\boldsymbol{\mathcal{X}}_{k|k-1} &= \mathbf{f}\left(\boldsymbol{\mathcal{X}}_{k-1}, \mathbf{u}_{k-1}\right) \\
\hat{\mathbf{x}}_k^- &= \sum_{i=0}^{2L} W_i^m \mathcal{X}_{i,k|k-1} \\
\mathbf{P}_k^- &= \sum_{i=0}^{2L} W_i^c [\mathcal{X}_{i,k|k-1} - \hat{\mathbf{x}}_k^-][\mathcal{X}_{i,k|k-1} - \hat{\mathbf{x}}_k^-]^T + \mathbf{R}^{\mathbf{v}} \\
\boldsymbol{\mathcal{Y}}_{k|k-1} &= \mathbf{h}\left(\boldsymbol{\mathcal{X}}_{k-1}\right) \\
\hat{\mathbf{y}}_k^- &= \sum_{i=0}^{2L} W_i^m \mathcal{Y}_{i,k|k-1}
\end{aligned}
$$

Measurement update (Correction):

$$
\begin{aligned}
\mathbf{P}_{\mathbf{y}_k\mathbf{y}_k} &= \sum_{i=0}^{2L} W_i^c [\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_k^-][\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_k^-]^T + \mathbf{R}^{\mathbf{n}} \\
\mathbf{P}_{\mathbf{x}_k\mathbf{y}_k} &= \sum_{i=0}^{2L} W_i^c [\mathcal{X}_{i,k|k-1} - \hat{\mathbf{x}}_k^-][\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_k^-]^T \\
\mathbf{K}_k &= \mathbf{P}_{\mathbf{x}_k\mathbf{y}_k} \mathbf{P}_{\mathbf{y}_k\mathbf{y}_k}^{-1} \\
\hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k) \\
\mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k \mathbf{P}_{\mathbf{y}_k\mathbf{y}_k} \mathbf{K}_k^T
\end{aligned}
$$

where $\mathbf{R}^{\mathbf{v}}$ is process noise covariance matrix, and $\mathbf{R}^{\mathbf{n}}$ is measurement noise covariance matrix.

B.3.3 Advantages over EKF

It is known that the advantages of the UKF over the EKF will not be obvious as the state-transition function is linear [49]. However, for the nonlinear observation, the EKF effectively builds up an approximation to the expected Hessian by taking outer products of the gradient. The UKF, however, may provide a more accurate estimate through direct approximation of the expectation of the Hessian. Another distinct advantage of the UKF occurs when either the architecture or the error matric is such that differentiation with respect to the parameters is not easily derived as necessary for EKF. The

UKF effectively evaluates both the Jacobian and Hessian precisely through its sigma point propagation, without the need to perform any analytic differentiation.

Appendix C

Some Useful Identities

C.1 Sums of Powers

$$
\begin{aligned}
\sum_{n=1}^{N} n &= \frac{1}{2}N(N+1) \\
\sum_{n=1}^{N} n^2 &= \frac{1}{6}N(N+1)(2N+1) \\
\sum_{n=1}^{N} n^3 &= \frac{1}{4}N^2(N+1)^2 \\
\sum_{n=1}^{N} n^4 &= \frac{1}{30}N(N+1)(2N+1)(3N^2+3N-1) \\
\sum_{n=1}^{N} n^5 &= \frac{1}{12}N^2(N+1)^2(2N^2+2N-1) \\
\sum_{n=1}^{N} n^6 &= \frac{1}{42}N(N+1)(2N+1)(3N^4+6N^3-3N+1) \\
\sum_{n=1}^{N} n^7 &= \frac{1}{24}N^2(N+1)^2(3N^4+6N^3-N^2-4N+2)
\end{aligned}
\tag{C.1}
$$

C.2 Trigonometry

$$
\begin{aligned}
\cos x \cos y &= \frac{1}{2}\cos(x+y) + \frac{1}{2}\cos(x-y) \\
\sin x \sin y &= -\frac{1}{2}\cos(x+y) + \frac{1}{2}\cos(x-y) \\
\sin x \cos y &= \frac{1}{2}\sin(x+y) + \frac{1}{2}\sin(x-y) \\
\cos x \sin y &= \frac{1}{2}\sin(x+y) - \frac{1}{2}\sin(x-y) \\
\sin(x \pm y) &= \sin x \cos y \pm \cos x \sin y \\
\cos(x \pm y) &= \cos x \cos y \pm \sin x \sin y \\
\cos^2 x &= \frac{1}{2}(1 + \cos 2x) \\
\sin^2 x &= \frac{1}{2}(1 - \cos 2x) \\
\sin x \cos x &= \frac{1}{2}\sin 2x \\
\sin x + \sin y &= 2\sin\frac{x+y}{2}\cos\frac{x-y}{2} \\
\sin x - \sin y &= 2\cos\frac{x+y}{2}\sin\frac{x-y}{2} \\
\cos x + \cos y &= 2\cos\frac{x+y}{2}\cos\frac{x-y}{2} \\
\cos x - \cos y &= -2\sin\frac{x+y}{2}\sin\frac{x-y}{2} \\
\sin x &= \frac{e^{jx} - e^{-jx}}{2j} \\
\cos x &= \frac{e^{jx} + e^{-jx}}{2}
\end{aligned}
\tag{C.2}
$$

Bibliography

[1] P. Ubolkosold, O. Loffeld, S. Knedlik, and F. G. Tchere, "A robust frequency offset estimator for single carrier signals," in *Proceedings of IEEE Sarnoff Symposium*, Princeton, USA, Mar. 2006.

[2] P. Ubolkosold, F. G. Tchere, S. Knedlik, and O. Loffeld, "Nonlinear least-squares frequency offset estimator and its simplified versions for flat-fading channels," in *Proceedings of IEEE International Symposium on Communications and Information Technologies (ISCIT)*, Bangkok, Thailand, Oct. 2006.

[3] ——, "Simple carrier frequency offset estimators for flat-fading channels," in *Proceedings of IEEE International Conferene on Communications (ICC)*, Glasgow, Scotland, June 2007.

[4] ——, "Nonlinear frequency offset estimators for single sinusoidal signals with time-varying envelope," in *Proceedings of IASTED Wireless and Optical Commuications (WOC)*, Monreal, Canada, June 2007.

[5] ——, "A new closed-form frequency estimator in the presence of fading-induced multiplicative noise," in *Proceedings of IEEE Vehicular Technology Conference (VTC)*, Singapore, May 2008.

[6] ——, "An improved frequency offset estimator for SIMO-OFDM with maximum estimation range," in *Proceedings of IEEE Vehicular Technology Conference (VTC)*, Singapore, May 2008.

[7] P. Ubolkosold, S. Knedlik, and O. Loffeld, "A method to extend the estimation range of the exisitng time-domain Doppler centroid estimators," *IEEE Geosci. Remote Sensing Lett.*, vol. 2, pp. 185–188, Apr. 2008.

[8] J. G. Proakis, *Digital communications*. McGraw-Hill, 2001.

[9] R. W. Chang, "Synthesis of band-limited orthogonal signals for multichannel data transmission," *Bell Systems Technical Journal*, vol. 45, pp. 1775–1796, Dec. 1966.

[10] J. L. Holsinger, "Digital communication over fixed time-continuous channels with memory, with special application to telephone channels," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, 1964.

[11] R. G. Gallager, *Information theory and reliable communications*. Wiley, 1968.

[12] S. B. Weinstein and P. M. Ebert, "Data transmission by frequency division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun.*, vol. 19, pp. 628–634, Oct. 1971.

[13] L. J. Cimini, "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. 33, pp. 665–675, Jul. 1985.

[14] T. Pollet, M. Van Bladel, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, pp. 191–199, Apr. 1995.

[15] R. Janaswamy, *Radiowave propagation and smart antennas for wireless communications*. Kluwer Academic Publishers, 2000.

[16] D. C. Rife and R. R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. Inform. Theory*, vol. 20, pp. 591–598, Sept. 1974.

[17] H. L. Van Trees, *Detection, estimation and modulation theory*. John Wiley & Sons, 1968.

[18] S. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression," *IEEE Trans. Inform. Theory*, vol. 31, pp. 832–835, Dec. 1985.

[19] K. Itoh, "Analysis of the phase unwrapping problem," in *Applied Optics*, 1982, p. 2470.

[20] S. Kay, "A fast and accurate single frequency estimator," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 1987–1990, Dec. 1989.

[21] V. Clarkson, P. J. Kootsookos, and B. G. Quinn, "Analysis of the variance threshold of Kay's weighted linear predictor frequency estimator," *IEEE Trans. Signal Processing*, vol. 42, pp. 2370–2379, Sep. 1994.

[22] E. Rosnes and A. Vahlin, "Frequency estimator of a single complex sinusoidal using a generalized Kay estimator," *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 407–415, Mar. 2006.

[23] M. P. Fitz, "Planar filtered techniques for burst mode carrier synchronization," in *Proceedings of IEEE Global Telecommunications Conference (Globecom)*, Phoenix, Arizona, USA, Dec. 1991.

[24] ——, "Further results in the fast estimation of a single frequency," *IEEE Trans. Commun.*, vol. 42, pp. 862–864, Feb.-Apr. 1994.

[25] M. Luis and R. Reggiannini, "Carrier frequency recovery in all-digital modems for burst-mode transmissions," *IEEE Trans. Commun.*, vol. 43, pp. 1169–1178, Feb.-Apr. 1995.

[26] U. Mengali and M. Morelli, "Data-aided frequency estimation of burst digital transmission," *IEEE Trans. Commun.*, vol. 45, pp. 23–25, Jan. 1997.

[27] O. Besson, M. Ghogho, and A. Swami, "Parameter estimation for random amplitude chirp signals," *IEEE Trans. Signal Processing*, vol. 47, pp. 3208–3219, Dec. 1999.

[28] O. Loffeld, "Phase unwrapping for SAR interferometry - A data fusion approach by Kalman filtering," in *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Hamburg, Germany, 1999.

[29] ——, "Demodulation of noisy phase or frequency modulated signals with Kalman filters," in *Proceedings of IEEE International Conference on Accoustics, Speach and Signal Processing (ICASSP)*, Adelaide, Australia, 1994.

[30] W. Kuo and M. P. Fitz, "Frequency offset compensation of pilot symbol assisted modulation in frequency flat fading channels," *IEEE Trans. Commun.*, vol. 45, pp. 1412–1416, Nov. 1997.

[31] M. Morelli, U. Mengali, and G. M. Vitetta, "Further results in carrier frequency estimation for transmissions over flat fading channels," *IEEE Commun. Lett.*, vol. 2, pp. 327–330, Dec. 1998.

[32] O. Besson and P. Stoica, "On frequency offset estimation for flat-fading channels," *IEEE Commun. Lett.*, vol. 5, pp. 402–404, Oct. 2001.

[33] F. Gini, M. Luise, and R. Reggiannini, "Cramer-Rao bounds in parametric estimation of fading radiotransmission channels," *IEEE Trans. Commun.*, vol. 46, pp. 1390–1398, Oct. 1998.

[34] M. Ghogho, A. Swami, and T. Durrani, "Frequency estimation in the presence of Doppler spread," *IEEE Trans. Signal Processing*, vol. 49, pp. 777–789, Apr. 2001.

[35] M. Morelli and U. Mengali, "Carrier-frequency estimation for transmissions over selective channels," *IEEE Trans. Commun.*, vol. 48, pp. 1580–1589, Sep. 2000.

[36] H. Viswanathan and R. Krishnamoorthy, "A frequency offset estimation technique for frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 5, pp. 166–168, Apr. 2001.

[37] E. Jeong, G. Choi, and H. Lee, "Data-aided frequency estimation fpr PSK signaling in frequency-selective fading," *IEEE Trans. Commun.*, vol. 19, pp. 1408–1419, Jul. 2001.

[38] V. Baronkin, Y. Zakharov, and T. Tozer, "Frequency estimation in slowly fading multipath channels," *IEEE Trans. Commun.*, vol. 50, pp. 1848–1859, Apr. 2002.

[39] R. H. Clarke and W. L. Khoo, "3-D mobile radio channel statistics," *IEEE Trans. Veh. Technol.*, vol. 46, pp. 798–799, May 1997.

[40] K. E. Baddour and N. C. Beaulieu, "Autoregressive models for fading channel simulation," in *Proceedings of IEEE Global Telecommunications Conference (Globecom)*, San Antonio, USA, Nov. 2001.

[41] Z. Wang and A. Abeysekera, "Performance of correlation-based frequency estimation methods in the presence of multiplicative noise," *IEEE Trans. Veh. Technol.*, vol. 55, pp. 1281–1290, Jul. 2006.

[42] O. Besson and P. Stoica, "Sinusoidal signals with random amplitude: least-squares estimation and their statistical analysis," *IEEE Trans. Signal Processing*, vol. 473, pp. 2733–2744, Nov. 1995.

[43] O. Besson and F. Galtier, "Estimating particles velocity from laser measurements: maximum liklihood and Cramer-Rao bounds," *IEEE Trans. Signal Processing*, vol. 44, pp. 3056–3068, Dec. 1996.

[44] F. Galtier, "Frequency estimation of laser signals with time-varying amplitude from phase measurements," in *Proceedings of IEEE International Conference on Accoustics, Speach and Signal Processing (ICASSP)*, Munich, Germany, 1997.

[45] P. Ubolkosold, S. Knedlik, and O. Loffeld, "Estimation of oscillators phase offset, frequency offset, and rate of change for bistatic interferometric SAR," in *Proceedings of European Conference on Synthetic Aperture Radar (EUSAR)*, Dresden, Germany, Apr. 2006.

[46] O. Loffeld, *Estimationstheorie I,II.* Oldenbourg-Verlag, 1999.

[47] P. S. Maybeck, *Stochastic models, estimation and control: volume I.* Academic Press, INC, 1979.

[48] S. J. Julier, J. K. Uhlmann, and H. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proceedings of the American Control Conference*, 1995, pp. 1628–1632.

[49] E. A. Wan and R. Merwe, "The unscented Kalman filter for nonlinear estimation," in *Proceedings of IEEE 2000 Adaptive Systems for Signal Processing, Communications and Control Symposium (ASSPCC)*, 2000, pp. 153–158.

[50] S. Abeysekera, "Performance of pulse-pair method of Doppler estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, pp. 520–531, Apr. 1998.

[51] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, pp. 1613–1621, Dec. 1997.

[52] M. Morelli and U. Mengali, "An improved frequency offset estimator for OFDM applications," *IEEE Commun. Lett.*, vol. 3, pp. 75–77, Mar. 1999.

[53] H. K. Song, Y. H. You, J. H. Paik, and Y. S. Cho, "Frequency offset synchronization and channel estimation for OFDM-based transmission," *IEEE Commun. Lett.*, vol. 4, pp. 95–97, Mar. 2000.

[54] M. Ghogho, A. Swami, and G. B. Giannakis, "Optimized null-subcarrier selection for CFO estimation in OFDM over frequency-selective fading channels," *Proceedings of IEEE Globecom (Globecom)*, Nov. 2001.

[55] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, pp. 2908–2914, Oct. 1994.

[56] A. Czylwik, "Synchronization for systems with antenna diversity," in *Proceedings of IEEE Vehicular Technology Conference (VTC)*, vol. 2, 1999, pp. 728–732.

[57] T. Schenk and A. Van Zelst, "Frequency synchronization for MIMO OFDM wireless LAN systems," in *Proceedings of IEEE Vehicular Technology Conference (VTC)*, 2003, pp. 509–513.

[58] A. Van Zelst and T. Schenk, "Implementation of a MIMO OFDM-based wireless LAN system," *IEEE Trans. Signal Processing*, vol. 52, pp. 484–494, Feb. 2004.

[59] ——, "Low-complexity nonlinear least squares carrier offset estimator for OFDM: identifiability, diversity, and performance," *IEEE Trans. Signal Processing*, vol. 52, pp. 2441–2452, Sep. 2004.

[60] F. K. Li and W. T. K. Johnson, "Ambiguities in spaceborne synthetic aperture radar systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 19, pp. 389–397, May 1983.

[61] F. K. Li, D. N. Held, J. C. Curlander, and C. Wu, "Doppler parameter estimation for spaceborne synthetic-aperture radars," *IEEE Trans. Geosci. Remote Sensing*, vol. 1, pp. 47–56, Jan. 1985.

[62] S. N. Madsen, "Estimating the Doppler centroid of SAR data," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 25, pp. 134–140, Mar. 1989.

[63] M. Y. Jin, "Optimal Doppler centroid estimation for SAR data from a quasi-homogeneous source," *IEEE Trans. Geosci. Remote Sensing*, vol. 6, pp. 1022–1025, Nov. 1986.

[64] K. Kristoffersen, "Time-domain estimation on the center frequency and spread of Doppler spectra in diagonostic ultrasound," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 35, pp. 685–700, Nov. 1988.

[65] O. Loffeld, "Doppler centroid estimation with Kalman filters," in *Proceedings of IEEE International Geosciences and Remote Sensing Symposium (IGARSS)*, Washington, USA, May 1990.

[66] ——, "Estimating time varying Doppler centroids with Kalman filters," in *Proceedings of IEEE International Geosciences and Remote Sensing Symposium (IGARSS)*, Helsinki, Finland, May 1991.

[67] A. Paulraj and T. Kailath, "Direction-of-arrival estimation by eigenstructure methods with imperfect spatial coherence of wave fronts," *J. Acoust. Soc. Amer.*, vol. 83, pp. 1034–1040, Mar. 1988.

[68] O. Besson and P. Stoica, "Decoupled estimation of DOA and angular spread for a spatially distributed source," *IEEE Trans. Signal Processing*, vol. 48, pp. 1872–1882, Jul. 2000.

[69] O. Besson, F. Vincent, P. Stoica, and B. A. Gershman, "Approximate maximum likelihood estimators for array processing in multiplicative noise environments," *IEEE Trans. Signal Processing*, vol. 48, pp. 2506–2518, Sep. 2000.

[70] A. Swami and M. Ghogho, "Estimating of direction of arrival in the presence of angular spread and Rician fading," in *Proceedings of Conference on Information Sciences and Systems*, Princeton, NJ, Mar. 2000.

[71] M. Ghogho, O. Besson, and A. Swami, "Estimation of directions of arrival of multiple scattered sources," *IEEE Trans. Signal Processing*, vol. 49, pp. 2467–2480, Nov. 2001.

[72] B. Hofmann-Wellenhof, K. Legat, and M. Wieser, *Navigation: principles of positioning and guidance*. SpringerWienNewYork, 2003.

[73] P. Ubolkosold, S. Knedlik, and O. Loffeld, "Constrained-trajectory based GPS/INS integration for reliable position and attitude determination," in *Proceedings of IEEE International Geosciences and Remote Sensing Symposium (IGARSS)*, Barcelona, Spain, June 2007.

[74] S. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice Hall, 1993.

[75] M. Marcus, *Basic theorems in matrix theory*. National Bureau of Standards, Appl. Math Series, 1960.

[76] P. Stoica and R. Moses, *Introduction to spectral analysis*. Prentice Hall, 1997.

[77] G. M. Jenkins and D. G. Watts, *Spectral Analysis and Its Applications*. Holden-Day, 1968.

[78] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, pp. 35–45, Mar. 1960.

[79] S. J. Julier and J. K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," in *Proceedings of International Symposium on Aerospace/Defence Sensing, Simulation and Controls (AeroSense)*, 1997, pp. 182–193.

[80] ——, "The scaled unscented transformation," in *Proceedings of American Control Conference (ACC)*, 2002, pp. 4555–4559.