Lecture Notes in Electrical Engineering 145

Abbas Mohammadi
Fadhel M. Ghannouchi

RF Transceiver Design for MIMO Wireless Communications

Springer

Lecture Notes in Electrical Engineering

Volume 145

Abbas Mohammadi and Fadhel M. Ghannouchi

RF Transceiver Design for MIMO Wireless Communications

Springer

Authors
Prof. Abbas Mohammadi
Amirkabir University
Electrical Engineering Department
Tehran
Iran

Prof. Fadhel M. Ghannouchi
University of Calgary
Electrical and Computer Engineering
Intelligent RF Radio Laboratory
Alberta
Canada

ISSN 1876-1100 e-ISSN 1876-1119
ISBN 978-3-642-27634-7 e-ISBN 978-3-642-27635-4
DOI 10.1007/978-3-642-27635-4
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2011945157

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The multiple-input multiple-output (MIMO) technique provides higher bit rates and better reliability in wireless systems. The efficient design of RF transceivers has a vital impact on the implementation of this technique. This first book is completely devoted to RF transceiver design for MIMO communications. The book covers the most recent research in practical design and applications and can be an important resource for graduate students, wireless designers, and practical engineers.

The book opens with an introduction to MIMO wireless communications, where the main advantages of using MIMO technique are described. This is followed by a discussion on the implementation techniques for MIMO modulators. After describing the fundamental concepts for RF transceivers and power amplifiers, the design and analysis methods for the RF section of MIMO transmitters and receivers are presented. Furthermore, the RF impairments in MIMO and OFDM systems, including nonlinearity, phase noise, I/Q imbalance and DC offset, are discussed; and, their compensation methods are presented. Finally, the design techniques for single RF front-end MIMO systems are described.

Audience

The book can be used by graduate students, researchers and design engineers in microwave and wireless design areas. It is assumed that the reader has a fundamental knowledge of communication circuit design and communication systems theory. The book may be used as a textbook for a graduate course on wireless transceiver design techniques.

Acknowledgments

We would like to acknowledge the Informatics Circle of Research Excellence (iCORE), Alberta, Canada, for their financial contribution to support the visiting professor stay of the first author at iRadio Lab, University of Calgary, which helped the completion of this book. We are grateful to our great students and researchers in Microwave/mm-wave and Wireless Communications Lab of Amirkabir University and iRadio Lab of the University of Calgary. This book could not have been completed without their fruitful research. Their useful comments, discussions, collaborations and help in producing many of the results presented in this book over the years are appreciated. In particular, we would like to

thank A. Abdipour, M. Heloui, R. Mirzavand, S.A. Bassam, M.R. Keshavarzi, M.H. Madani, and I. D'Adamo. We also gratefully appreciate the careful and thorough editing by Ms. Cathie Heys.

Finally, we would like to profoundly thank our spouses for their understanding and patience throughout the many evenings and weekends taken to prepare this book, as well as our parents for their encouragement and valuable support in our early professional years as graduate students and young researchers.

Contents

Chapter 1
Introduction

Multiple input multiple output (MIMO) wireless systems provide many advantages by using more than a single antenna. As such, it is considered in the current and future wireless standards. However, radio frequency (RF) transceiver design for MIMO wireless communications is a challenging task. This subject has been attracting research attention in both academia and industry. This chapter provides a general overview of MIMO systems and transceiver implementation.

1.1 Multiantenna Wireless Communications

Wireless communication systems are realized by using a transmitter, receiver and wireless channel. The classic implementation of these systems is usually a transmitter and its single antenna on the transmitting side and a receiver and its single antenna on the receiving side. The concept of using multiple antennas at receiver was introduced in the early 1960s [1]. The basic idea is the provision of multiple copies from a transmitted signal on the receiving side and their combination to obtain a signal with better performance. This technique is called the space diversity technique. To realize the receiver diversity technique, multiple antennas are employed on the receiver side. The antennas' output signals are then combined by applying different weighting coefficients, according to different diversity techniques.

Figure 1.1 shows the general diagram of the receiver diversity method. The receiver diversity may be realized using three different methods, selection combining, equal gain combining, and maximal ratio combining. In the selection combining method, the receiver selects the strongest signals among the antennas' outputs. This signal is processed in the single receiver. Maximal ratio combining is realized by weighting both the gain and phase of the received signals to maximize the signal-to-noise ratio (SNR) at the output of the receiver.

On the other hand, in the equal gain combining method, only the phases of the channel using weighting coefficients are cancelled. The maximum ratio combiner diversity technique provides the best performance to mitigate the wireless channel impact, while the selection diversity offers the simplest implementation technique among the different diversity methods [1].

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 1–8.
springerlink.com

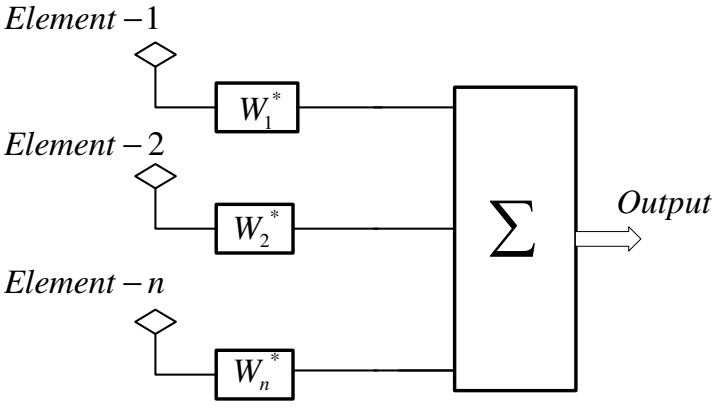


Fig. 1.1 Diagram of the receiver diversity technique

The receiver diversity technique is a special implementation for the use of multiple antennas. The multiple antenna technique can be used in both the transmitter and the receiver. This technique is generally known as a multiple input, multiple output (MIMO) system. The various realization techniques for MIMO systems are shown in Figure 1.2. A system that uses a single antenna on both the transmitting and receiving sides is called a single input single output (SISO) system.

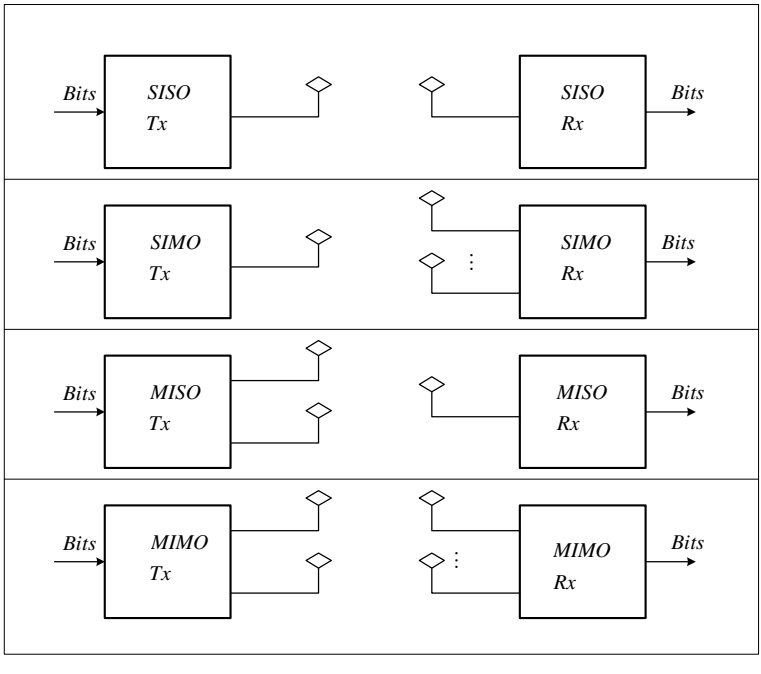


Fig. 1.2 Block diagram of multiple antennas systems

On the other hand, a system that employs a single antenna in the transmitter and multiple antennas in the receiver is called a single input multiple output (SIMO) system. Indeed, this system realizes the receiver diversity scenario. The idea of exploring the transmitter diversity was introduced in 1990s [2]. By using a similar notation, a system that uses multiple antennas at the transmitter and a single antenna in the receiver is called a multiple input single output (MISO) system. Likewise, the system that employs multiple antennas in the transmitter and multiple antennas in the receiver edges is the MIMO system.

A MIMO system with M_T transmitter antennas and M_R receiver antennas is shown in Figure 1.3. The input data is transmitted through M_T antennas after processing on the transmitter side. The processing includes channel coding, modulation, space-time encoding, spatial mapping, and RF up-conversion.

Each antenna transmits a signal through a wireless channel. Accordingly, all M_T simultaneous radiators operate as a transmitter.

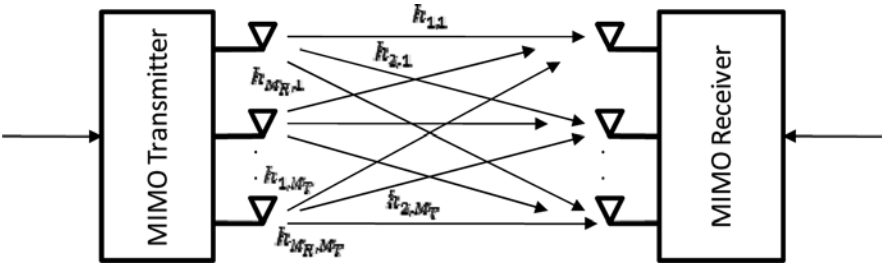


Fig. 1.3 General block diagram of MIMO systems

1.2 MIMO Wireless Transceivers

The art of designing modern wireless transceivers has seen extensive progress during the last two decades. The design of these transceivers generally must follow three main factors. They should be high performance, low cost, and able to handle the complex objectives of advanced wireless communication systems. A general block diagram of a wireless transceiver using multiple antennas in both the transmitter and the receiver is shown in Figure 1.4.

In recent transceivers, a great amount of signal processing is achieved in the baseband using digital signal processors (DSPs), application specific integrated circuits (ASICs), and field-programmable gate arrays (FPGAs). A signal at the output of the baseband section is up-converted to RF frequency using a heterodyne technique or a direct conversion method. A frequency synthesizer provides the carrier signal for up-conversion. The power amplifiers play a crucial role in providing the performance of the system. The RF front-end delivers the signal to the transmitter antennas. After passing through a wireless channel on the receiver side, the signals of the multiple antennas must be down-converted. This process includes RF signal processing using low noise amplifiers (LNAs), frequency synthesizing, and down conversion. Signals must then be delivered to the receiver baseband.

The excessive signal processing is performed in the receiver to detect the signals. Likewise, on the transmitter side, digital signal processing is carried by receiver digital circuits.

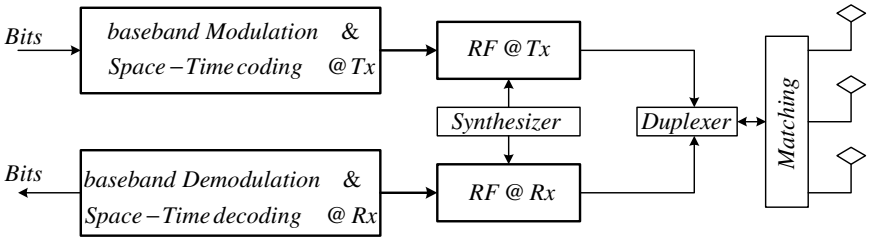


Fig. 1.4 Block diagram for MIMO wireless transceivers

1.3 MIMO Techniques in Commercial Wireless Systems

In a fading environment, the performance of a wireless link can be greatly improved by using multiple antennas on both the transmitting and receiving sides. These benefits include increased reliability as well as high data rates. A MIMO wireless communication system can be designed to take advantage of reliability improvement or increasing the data rate. The first improvement is called a diversity order improvement, and the latter is called spatial multiplexing. However, it has been shown that both types of gains can be simultaneously obtained using a fundamental tradeoff between them [4-5]. The application of the MIMO technique in wireless communications is aimed for both spatial multiplexing and diversity improvement.

Most wireless standards have been reconsidered regarding the usage of MIMO systems. In the case of wireless networking standards (IEEE 802.11x), the IEEE 802.11n amendment introduces MIMO in the physical layer of the network to improve the network throughput over previous standards. This increases the maximum data rate from 54 Mbps for the IEEE 802.11a/g standards to 600 Mbps for the IEEE 802.11n standard, using a 40 MHz bandwidth with MIMO capability. The MIMO technique is used in IEEE 802.11n wireless local area network (LAN) to provide spatial multiplexing.

The increased throughput comes from a mixture of changes to the way data packets are sent, along with the use of sophisticated radio techniques that demand high performance from the RF hardware. Multiple RF channels may be implemented with integrated transceivers and the same local oscillator (LO) with a separate front-end module. Figure 1.5 shows the main components for the IEEE 802.11n wireless LAN [3].

In addition, the World Interoperability for Microwave Access (WiMAX) standard also supports MIMO for both fixed and mobile WiMAX (IEEE 802.16e and IEEE 802.16m). Its MIMO feature is intended for high-speed applications, and it covers both space-time and spatial multiplexing codes [10]. The functional block diagram of a WiMAX IEEE802.16e from NXP Semiconductors (Philips) is shown in Figure 1.6 [11]. By using two antennas, this system provides a better performance for a wireless system.

The MIMO technique is also considered in the 3G (third generation) Long Term Evolution (LTE) standard, in order to increase the data rate to up to 277 Mbps in the downlink for a single user in an ideal radio link condition using a 4x4 MIMO system with a 20 MHz bandwidth [6]. Figure 1.7 shows a simplified block diagram of a LTE transceiver using two-channel reception [9].

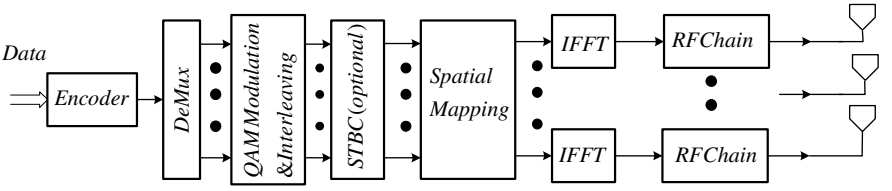


Fig. 1.5 Simplified block diagram of WLAN IEEE802.11n using MIMO

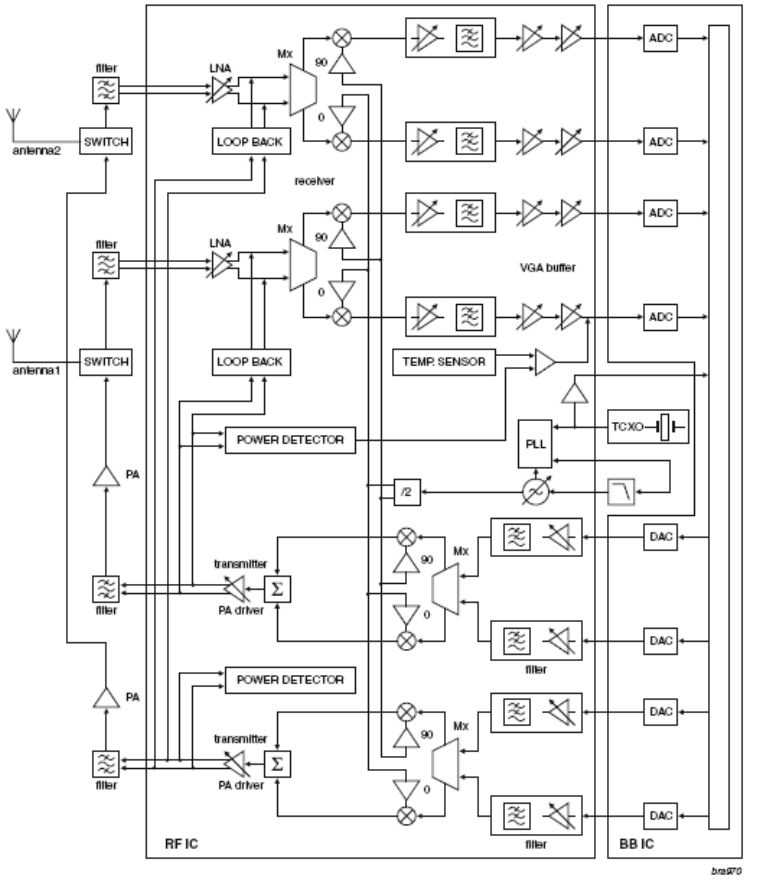


Fig. 1.6 Block diagram of WiMAX RF transceiver IEEE802.16e using MIMO from NXP [11]

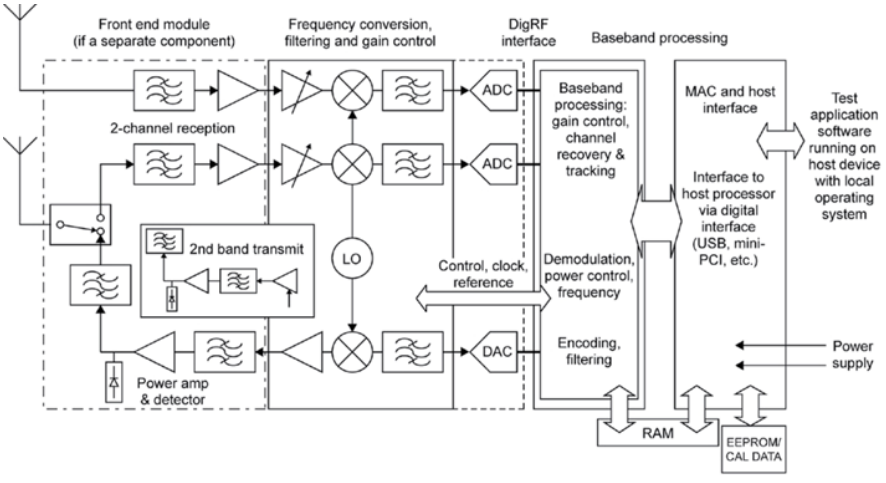


Fig. 1.7 Simplified transceiver block diagram of LTE [9]

1.4 Organization of the Book and Future Challenges

The book addresses some of the key design techniques for RF transceivers of MIMO wireless systems. Chapter 2 presents the fundamentals for understanding the multiple-antenna theory and advantages. This begins by presenting the MIMO system model, and MIMO channel models are the briefly discussed. This is followed by an examination of the MIMO capacity concept. The diversity gain and spatial multiplexing concepts are also introduced.

Chapter 3 deals with digital modulation techniques and orthogonal frequency-division multiplexing (OFDM) systems. After introducing the power efficiency and bandwidth efficiency concepts in digital modulation, the impact of fading channels on these modulators are discussed. The MIMO-OFDM technique and its implementation process are also discussed in this chapter. Chapter 4 provides a comprehensive overview of the fundamental concepts in wireless transceivers design. Chapter 5 covers the different subjects on RF power amplifiers and linearization techniques. The various power amplifier design techniques, including linear, high-efficiency and broadband techniques, are described. This is followed by a discussion on power amplifier linearization techniques. RF transceiver design for MIMO wireless communications is an active research area where more suitable architecture of MIMO transceiver is still an ongoing research topic [12-14].

Chapter 6 deals with transmitter design issues in MIMO wireless transceivers. The various transmitter architectures, such as heterodyne and direct conversion, are investigated. This is followed by the introduction of the transmitter design technique for MIMO applications. Chapter 7 presents the receiver design procedure for wireless MIMO communications. This chapter includes the commercial receiver design techniques for MIMO wireless communications. RF

impairments in MIMO systems and their compensation techniques have been attracting the research attention in both academia and industry [15-18].

Chapter 8 discusses RF impairments and compensation techniques for OFDM wireless transceivers. The impairments include phase noise and power amplifier nonlinearities. The impacts of the impairments are also examined in this chapter, and the compensation techniques are presented. Chapter 9 discusses RF impairments and compensation techniques for MIMO wireless transceivers. The impairments include phase noise, DC offset, I/Q imbalance, and nonlinearities. The last chapter addresses the different alternatives in designing single RF front-end MIMO transceivers. The cost and complexity of the RF transceiver of a MIMO system increases linearly with an increasing number of antennas. The simple RF front-end design for MIMO systems is actively researched in both academia and industry [18-24].

References

[1] Janaswamy, R.: Radiowave Propagation and Smart Antennas for Wireless Communications. Kluwer Academic Publishers (2001)

[2] Paulraj, A., Nabar, R., Gore, D.: Introduction to Space-Time Wireless Communications. Cambridge University Press (2003)

[3] Agilent Technologies, MIMO Wireless LAN PHY Layer RF Operation & Measurement. Agilent Application Note #1509 (April 2008)

[4] Biglieri, E., Calderbank, R., Goldsmith, A., Paulraj, A., Vincentpoor, H.: MIMO Wireless Communications. Cambridge University Press (2007)

[5] Ebrahimzad, H., Mohammadi, A.: Diversity-Multiplexing Tradeoff in MIMO Systems with Finite SNR. In: European Conference on Wireless Technology, Munich, pp. 146–149 (October 2007)

[6] Qualcomm Incorporated, LTE—A Well-Designed Mobile OFDMA IP Solution, Qualcomm Incorporated Report (January 2008)

[7] Tsoulos, G.: MIMO System Technology for Wireless Communications. CRC Press (2006)

[8] Kaiser, T., Bourdoux, A., Boche, H., Fonollosa, J.R., Andersen, J.B., Utschick, W.: Smart Antennas—State of the Art. Hindawi Publishing Corporation (2005)

[9] Rumbey, M.: LTE and the Evolution to 4G Wireless: Design and Measurement Challenges. Wiley (2009)

[10] Hanzo, L., Akhtitman, Y., Wang, L., Jiamg, W.: MIMO-OFDM for LTE, WiFi and Wimax. Wiley (2010)

[11] Locher, M., Tomesen, M., Kuenen, J., Daanen, A., Visser, H., Essink, B., Vervoort, P.P., Nijrolder, M., Kopmeiners, R., Redman-White, W., Balmford, R., El Waffaoui, R.: A Low Power, High Performance BiCMOS MIMO/Diversity Direct Conversion Transceiver IC for WiBro/WiMAX (802.16e). In: IEEE 2007 Custom Intergrated Circuits Conference, CICC (2007)

[12] Domizioli, C.P., Hughes, B.L., Gard, K.G., Lazzi, C.: Optimal front-end design for MIMO receivers. In: IEEE Global Communications Conference, Globcom 2008, New Orleans (2008)

[13] Rafati, H., Razavi, B.: Receiver Architecture for Dual-Antenna Systems. IEEE Journal Solid-State Circuits 42(6), 1291–1299 (2007)

[14] Eickhoff, R., Kraemer, R., Santamaria, I., Gonzalez, L.: Developing energy-efficient MIMO radios. IEEE Vehicular Technology Magazine 4(1), 34–41 (2009)
[15] Madani, M.H., Abdipour, A., Mohammadi, A.: Analysis of performance degradation due to non-linearity and phase noise in orthogonal frequency division multiplexing systems. IET Communications Journal 4(10) (2010)
[16] Keshavarzi, M.R., Mohammadi, A., Abdipour, A., Ghannouchi, F.M.: Characterization and Compensation of DC Offset on Adaptive MIMO Direct Conversion Transceivers. IEICE Transactions on Communications (January 2011)
[17] Schenk, T.: RF Imperfections in High-rate Wireless Systems. Springer, Heidelberg (2008)
[18] Horlin, F., Bourdoux, A.: Digital compensation for analog front-ends: a new approach to wireless transceiver design. Wiley (2008)
[19] Kalis, A., Kanatas, A.G., Papadias, C.B.: A novel approach to MIMO transmission using a single RF front end. IEEE Journal of Selected Areas of Communication 26(6) (August 2008)
[20] Duman, T.M., Ghrayeb, A.: Coding for MIMO Communication Systems. Wiley (2008)
[21] Molisch, A., Win, M., Winters, J.: Reduced-complexity transmit/receive-diversity systems. IEEE Transactions on Signal Processing 51(11), 2729–2738 (2003)
[22] Alrabadi, O.N., Papadias, C.B., Kalis, A., Prasad, R.: A Universal Encoding Scheme for MIMO Transmission Using a Single Active Element for PSK Modulation Schemes. IEEE Transactions on Wireless Communications 8(10), 5133–5143 (2009)
[23] Tzeng, F., Jahanian, A., Pi, D., Heydari, P.: A CMOS Code-Modulated Path-Sharing Multi-Antenna Receiver Front-End. IEEE Journal of Solid-State Circuits 44(5), 1321–1335 (2009)
[24] Lari, M., Bassam, A., Mohammadi, A., Ghannouchi, F.M.: Time-Multiplexed Single Front-End MIMO Receivers with Preserved Diversity Gain, vol. 5(6), pp. 789–796 (2011)

Chapter 2
MIMO Wireless Communications

The multiple input multiple output (MIMO) technique provides the higher bit rate and the better reliability in wireless systems. These advantages are achieved by designing appropriate apace-time codes that provide diversity improvement, spatial multiplexing gain, or a trade-off between diversity order and spatial multiplexing. This chapter provides an overoview on MIMO wireless system concept and its performance. Moreover, the MIMO channel models are discussed.

2.1 MIMO System

A multiple input multiple out (MIMO) system with M_T transmitting antennas and M_R receiving antennas is shown in Figure 2.1. The input data are transmitted through M_T antennas after processing on the transmitter side. The processing includes channel coding, modulation, space-time encoding, spatial mapping, and radio frequency (RF) up-conversion. Each antenna transmits a signal through a wireless channel. Accordingly, the M_T antennas simultaneously operate as an entire transmitter. The radiated signals are represented by a column vector ($\mathbf{x}$) that has $M_T \times 1$ dimensions. These signals, after passing through the wireless channel, are received by M_R receiving antennas.

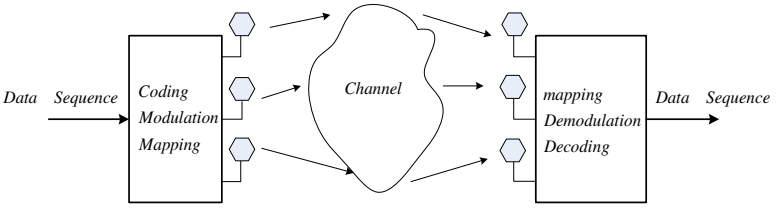


Fig. 2.1 Block diagram of a MIMO system

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 9–26.
springerlink.com

2.2 MIMO Channel

The impulse response of a linear time-varying communication system between a transmitter and a receiver is used to describe the effects of a linear transmission channel.

2.2.1 SISO Channel Model

In the first step, a narrowband system using a single antenna at both the transmitter and the receiver is assumed (single input single out – SISO). The symbol period is assumed to be *T*. Moreover, the digital signal in discrete time may be represented by the complex time series $\{x_k\}$.

In this case, the transmitted signal is represented by:

$$x(t) = \sum_{m=-\infty}^{\infty} \sqrt{E_s}\, x_m \delta(t - mT) \tag{2.1}$$

where E_s is the transmitted symbol energy, assuming that the average energy constellation is normalized to unity. In a linear time-invariant (LTI) system, a function *h(t)* as the time-invariant impulse response of the channel can be considered [1]. If the signal $x(t)$ is transmitted, the received signal $r(t)$ is given by

$$r(t) = h(t) * x(t) + n(t) \tag{2.2}$$

where * denotes the convolution product, and *n(t)* is the additive noise of the system. Therefore, the input-output relation is represented as:

$$r(t) = \sum_{m=-\infty}^{\infty} \sqrt{E_s}\, x_m h(t - mT) + n(t) \tag{2.3}$$

One obtains the discrete representation of the received signal by sampling the received signal at the rate of *T,* (*r(kT)*) as:

$$\begin{aligned} r(kT) &= \sum_{m=-\infty}^{\infty} \sqrt{E_s}\, x_m h[(k-m)T] + n(kT) \\ r[k] &= \sum_{m=-\infty}^{\infty} \sqrt{E_s}\, x_m h[k-m] + n[k] \end{aligned} \tag{2.4}$$

As may be seen, a time-invariant channel can be represented as an LTI system and its sampled representation *h[k]*. The extension of this representation to a time-varying channel is completely straightforward [2], [4]. Moreover, there are

different methods to extract the channel impulse response both in narrowband and broadband transmissions [3]. The channel impulse response generally depends on attenuation of the path loss term, shadowing, and multipath fading.

2.2.2 MIMO Channel Modeling

In MIMO systems, both the transmitter and receiver have several antennas. A MIMO system with M_T transmitting antennas and M_R receiving antennas is shown in Figure 2.2. In this system, the channel between each transmitting-receiving antenna pair can be modeled as a SISO channel. Accordingly, the channel matrix (**H**) for a MIMO system with M_T transmitting antennas and M_R receiving antennas may be obtained. By arranging all inputs and outputs in vectors, $\mathbf{x}[k]=[x_{1,k},...,x_{M_T,k}]^T$ and $\mathbf{r}[k]=[r_{1,k},...,r_{M_R,k}]^T$, the input-output relationship at any given time instant (k) is obtained as:

$$\mathbf{r}[k]=\sqrt{E_s}\,\mathbf{H}[k]\mathbf{x}[k]+\mathbf{n}[k] \tag{2.5}$$

where $\mathbf{H}[k]$ is defined as the $M_R \times M_T$ MIMO channel matrix in sampling time kT, and $\mathbf{n}[k]=[n_{1,k},...,n_{M_R,k}]^T$ is the sampled noise vector, containing the noise contribution at each receive antenna, such that the noise is white in both time and spatial dimensions as:

$$E\{\mathbf{n}[k]\mathbf{n}[m]\}=\sigma_n^2\mathbf{I}_{M_R}\delta[k-m] \tag{2.6}$$

where σ_n^2 is the Gaussian noise variance and $\mathbf{I}_{M_R}$ is the identity matrix. By normalizing the transmit symbol energy to unity, the received signal is shown to be:

$$\mathbf{r}=\mathbf{H}\mathbf{x}+\mathbf{n} \tag{2.7}$$

where $\mathbf{r}$ is the received column vector signal with dimensions of $M_R \times 1$ composed of the received signal, r_j; $\mathbf{n}$ is an $M_R \times 1$ column vector composed of the noise components, n_j; and, $\mathbf{H}$ is the $M_R \times M_T$ channel matrix with jith component being the channel coefficient, $h_{j,i}$. Accordingly, a MIMO system can be represented using the following matrix equation:

$$\begin{bmatrix} r_1 \\ \vdots \\ r_{M_R} \end{bmatrix}=\begin{bmatrix} h_{1,1} & \cdots & h_{1,M_T} \\ \vdots & \ddots & \vdots \\ h_{M_R,1} & \cdots & h_{M_R,M_T} \end{bmatrix}\begin{bmatrix} x_1 \\ \vdots \\ x_{M_T} \end{bmatrix}+\begin{bmatrix} n_1 \\ \vdots \\ n_{M_R} \end{bmatrix} \tag{2.8}$$

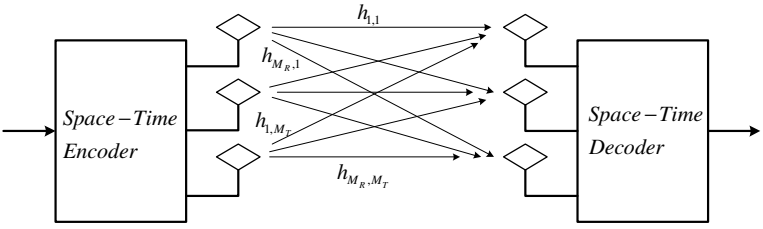


Fig. 2.2 MIMO channel

■ ***Example 2.1:*** *Independent and Identically distributed Rayleigh MIMO Channel*

If each individual channel coefficients is a zero-mean complex, circularly symmetric Gaussian variable or, equivalently, a complex variable whose amplitude and phase are Rayleigh and uniformly distributed, respectively, it is called a wide-sense stationary uncorrelated scattering homogeneous (WSSUSH) Rayleigh fading channel [4]. When the antenna spacing on both sides of the link is large enough, the various channel correlations become very small and can be assumed to be equal to zero [2]. The typical antenna spacing for negligible correlation is about $\lambda/2$, where λ is the wavelength. Furthermore, if all individual channels coefficients are characterized by the same average power, the channel matrix is represented as $\mathbf{H}_w$. This is a random fading matrix with unit variance and circularly symmetric complex Gaussian entries. This channel, which uses the independent and identically distributed (i.i.d.) assumption, is represented by as $CN(0,1)$.

2.3 MIMO Capacity

The channel capacity is a fundamental limit on the rate of error-free messages that can be transmitted through a communication channel. In this section, the capacity of SISO and MIMO channels under fading are discussed.

2.3.1 SISO Capacity

The channel capacity for SISO communication systems over additive white Gaussian noise (AWGN) channels has been extracted by Shannon as:

$$C = B\log_2(1+\frac{S}{N}) \tag{2.9}$$

where C is the channel capacity in bit per second, B is the channel bandwidth in Hz, S is the transmitted power, and N is the noise power in the channel bandwidth.

Defining $\gamma \triangleq S / N$, the channel capacity for fading channels can be obtained as [5], [6]:

$$C = \int_0^\infty B \log_2(1+\gamma) p_\gamma(\gamma) d\gamma \tag{2.10}$$

where $p_\gamma(\gamma)$ is the probability density function of fading channel.

2.3.2 MIMO Capacity

MIMO systems with multiple antennas on both the transmitting and receiving sides have been considered. The capacity of MIMO systems can be expressed in bits per second per hertz (bps/Hz) as [4], [7]:

$$\begin{aligned} &C = \max\{\log_2 \det(I_{M_R} + \frac{\gamma}{M_T} H R_{ss} H^H)\} \\ &T_r(R_{ss}) = M_T \end{aligned} \tag{2.11}$$

where I_{M_R} is the identity matrix, H is the channel matrix with H^H being its transpose conjugate, $R_{ss} = E\{ss^H\}$ is covariance matrix of the transmit signal, γ gives the average signal-to-noise ratio (SNR) per receiver branch, and $T_r(.)$ represents the trace of a matrix. As can be seen, the channel capacity may be reached by choosing the optimal covariance structure for the transmitted signals. If the channel is unknown on the transmitter side, the transmitted signal can be considered spatially white, e.g., $R_{ss} = I_{M_T}$. Accordingly, the MIMO channel capacity for a sample deterministic realization is given by:

$$C = \log_2 \det(I_{M_R} + \frac{\gamma}{M_T} H H^H) \tag{2.12}$$

If the eigendecomposition of HH^H is represented as $Q \Lambda Q^H$, where Q is an $M_R \times M_R$ matrix satisfying $QQ^H = Q^H Q = I_{M_R}$ and $\Lambda = diag\{\lambda_1, \lambda_2, ..., \lambda_{M_R}\}$ with $\lambda_i \geq 0$, equation (2.12) can be written as:

$$C = \log_2 \det(I_{M_R} + \frac{\gamma}{M_T} Q \Lambda Q^H) \tag{2.13}$$

By using the identity of $\det(I_m + AB) = \det(I_n + BA)$ for A_{mn}, B_{nm} and using $QQ^H = Q^H Q = I_{M_R}$, the capacity relation may be written as:

$$C = \log_2 \det(I_{M_R} + \frac{\gamma}{M_T}\Lambda) \tag{2.14}$$

This relation equivalently may be represented as:

$$C = \sum_{i=1}^{r} \log_2(1 + \frac{\gamma}{M_T}\lambda_i) \tag{2.15}$$

where r is the rank of channel and $\lambda_i \geq 0$ are the eigenvalues of HH^H [4].

Equation (2.12) can be also used to obtain the MIMO capacity in the fading channel. Extending equation (2.10) for an MIMO channel, the capacity of the MIMO channel under fading can be represented by:

$$C = E\{\log_2 \det(I_{M_R} + \frac{\gamma}{M_T} HH^H)\} \tag{2.16}$$

where the average is over the distribution of the elements of H.

In order to provide a clear idea about MIMO capacity, the asymptotic MIMO capacity should be examined. Assuming the spatially white Guassian channel ($CN(0,1)$) and using the law of large numbers considering large $M (M_T = M_R = M)$ [14], [7] yields:

$$\frac{1}{M} H_W H_W^H \to I_M \qquad M \to \infty \tag{2.17}$$

where H_W is the spatially white Gaussian channel. Accordingly, the capacity can be written as:

$$C = M \log_2(1 + \gamma) \tag{2.18}$$

This interesting result shows that the capacity increases linearly with an increasing number of antennas.

The capacity of a single input multiple out (SIMO) system, which is also known as the receiver diversity, can be expressed as [7]:

$$C_{SIMO} = \log_2(1 + \gamma M_R) \tag{2.19}$$

As can be seen, only the logarithmic improvement in capacity with increasing M_R is achievable in the SIMO system. For a random channel, this relation is valid for both the known channel state information (CSI) at the transmitter and the unknown CSI. This is due to the implementation of a single transmitting antenna. The capacity of the multiple input single output (MISO) channel when no CSI is provided at transmitter is obtained as:

$$C_{SIMO} = \log_2(1+\gamma) \tag{2.20}$$

The capacity of the MISO system with a known CSI at transmitter is expressed as:

$$C_{MISO} = \log_2(1+\gamma M_T) \tag{2.21}$$

■ ***Example 2.2:*** *Capacity Estimation for 2X2 MIMO System*

The capacity of a 2X2 MIMO system is presented in this example [25]. As shown in Figure 2.3, a MIMO system with polarized matched transmitting and receiving antennas is assumed. The antennas gains are 0 dBm.

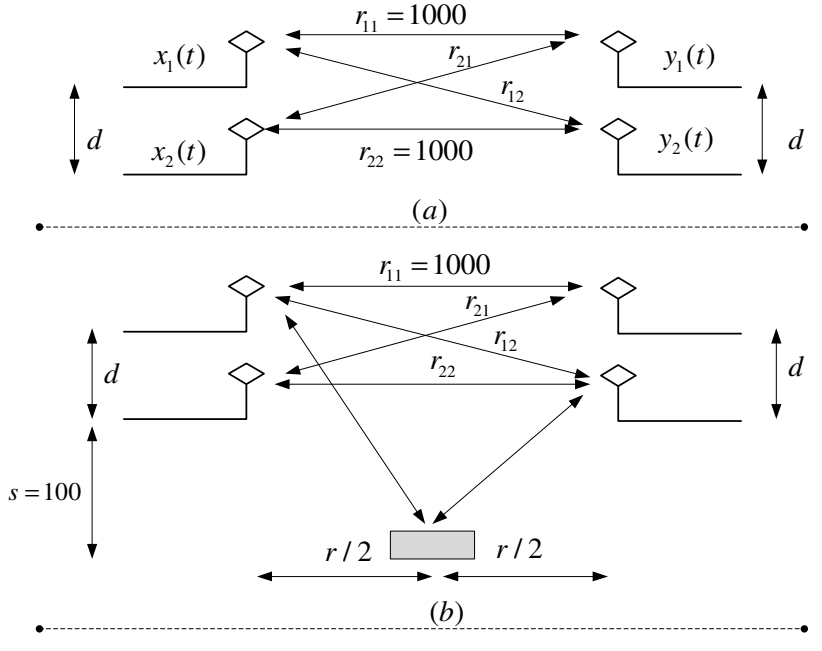


Fig. 2.3 Capacity estimation for 2X2 MIMO channel: a) without scattering b) with single scattering [25]

The communications system operates at 1.9 GHz with a 200 KHz bandwidth and a transmitting power of 1 mW. The noise temperature is 300 K. The capacity is calculated in the following cases:

Case I: *SISO Capacity*

The received power due to the propagation path loss is a function of frequency operation and is obtained as:

$$P_r = P_t G_t G_r \left(\frac{16\pi^2 d^n}{\lambda^2}\right)$$

where P_r, P_t are the receiving and transmitting powers; G_r, G_t are the receiving and transmitting antenna gains; d is the distance between the transmitter and the receiver; n is the path loss exponent, which is usually between 2 and 6; and, λ is the free space path length. In this example, it is assumed that $n = 2$. Using the above relations and equation (2.20):

$$\begin{aligned}
&P_r = 1.58\times10^{-13} W\,(-98 dBm)\\
&N_o = kT_e B = 8.28\times10^{-16} W\,(-120.8 dBm)\\
&\gamma = 22.8 dBm\\
&C_{1\times1} = \log_2(1+\gamma) = 7.6 \qquad bps/Hz
\end{aligned}$$

Case II: *MIMO Capacity without Scattering*

As illustrated in Figure 2.3(a), by calculating **H** and using equation (2.15):

$$\mathbf{H} = \begin{pmatrix} e^{-jkr_{11}} & e^{-jkr_{12}} \\ e^{-jkr_{21}} & e^{-jk_{11}r_{22}} \end{pmatrix}$$

where $k = \omega\sqrt{\mu_0\varepsilon_0}$.

The capacity is obtained as:

$$\text{if} \qquad d = .5 \rightarrow \lambda_1 = 2.0, \lambda_2 = 0 \rightarrow C_{2\times2} = \sum_{i=1}^{2} \log_2\left(1+\frac{\gamma}{2}\lambda_i\right) = 7.6 \quad bps/Hz$$

$$\text{if} \qquad d = 10 \rightarrow \lambda_1 = 1.41, \lambda_2 = .6 \rightarrow C_{2\times2} = \sum_{i=1}^{2} \log_2\left(1+\frac{\gamma}{2}\lambda_i\right) = 12.95 \quad bps/Hz$$

Case III: *MIMO Capacity with Single Scattering*

Similarly, as shown in Figure 2.3(b), by calculating **H** and using equation (2.15), the capacity is obtained as:

$$if \qquad d = .5 \rightarrow \lambda_1 = 2.14, \lambda_2 = .2 \rightarrow C_{2\times 2} = \sum_{i=1}^{2} \log_2(1+\frac{\gamma}{2}\lambda_i) = 12.01 \quad bps/Hz$$

Case IV: *MIMO Capacity in a Rich Scattering Environment*

By using equation (2.18) and assuming a scenario of rich scattering, the capacity is obtained as *15.2 bps/Hz*.

2.4 MIMO Design Advantages

MIMO systems can be designed either to provide maximal diversity to increase transmission reliability or to achieve maximal multiplexing gain to support high data rates [8]. Equation (2.15) shows that the channel capacity of a MIMO system can be much higher than that of a SISO system. This performance of a MIMO system is quantified with spatial multiplexing gain. On the other hand, if the signal copies are transmitted from multiple antennas or received at more than one antenna, the multiple antenna systems can provide a gain that can improve reliability of a wireless link. This gain is called diversity gain.

Furthermore, MIMO systems can be used to simultaneously provide both diversity and multiplexing gain. However, there is a fundamental tradeoff between them, either in large SNR values [9] or finite SNR values [11].

2.4.1 Space-Time Codes for Diversity

In this design, the signal copies are transmitted from multiple antennas or received at more than one antenna in space-time multi-antenna systems. The average symbol error probability of a MIMO communications system for maximum likelihood (ML) detection is has an upper boundary in high SNR values as [4]:

$$p_e \le \overline{N_e}(\frac{\gamma d_{\min}}{4M})^{-M} \tag{2.22}$$

where $\overline{N_e}$ is the number of nearest neighbors in a scalar constellation, $d_{\min}$ is the minimum distance of separation of the scalar constellation, and $M=min\{M_R,M_T\}$.

As can be seen in Figure 2.4, increasing the number of antennas increases the slope of the bit error rate (BER) curves and improves the reliability of wireless communications. Two general techniques to generate the space-time codes for

diversity are space-time block codes and space-time trellis codes [4], [7]. Although the space-time trellis codes were introduced earlier [12], the space-time block codes have been the preferred coding techniques in practice, due to the ease of their implementation.

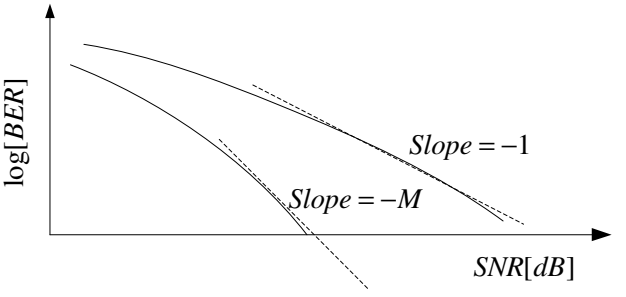


Fig. 2.4 Space-time diversity gain in MIMO systems

■ ***Example 2.3:*** *Alamouti Space-Time Coding*

The Alamouti code is an orthogonal space-time block code (O-STBC). It is implemented by using two antennas at the transmitter and an arbitrary number of the antennas at the receiver [10]. The code words for multiple antennas are written as:

$$\mathrm{X} = \frac{1}{\sqrt{2}} \begin{pmatrix} \chi_1 & -\chi_2^* \\ \chi_2 & \chi_1^* \end{pmatrix} \tag{2.23}$$

The Alamouti transmitter is shown in Figure 2.5. The Alamouti code has a spatial multiplexing rate equal to one, as two symbols are transmitted over two symbol durations. The performance of the Alamouti code is presented in Figure 2.6.

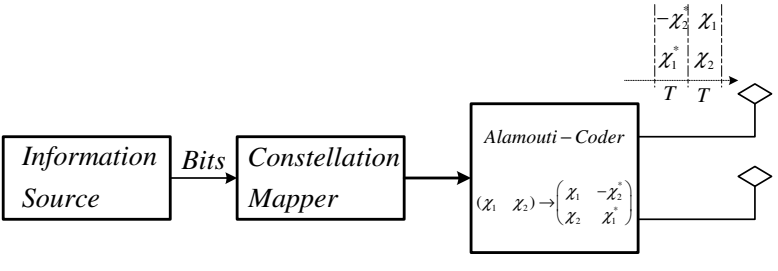


Fig. 2.5 The Alamouti space-time coding

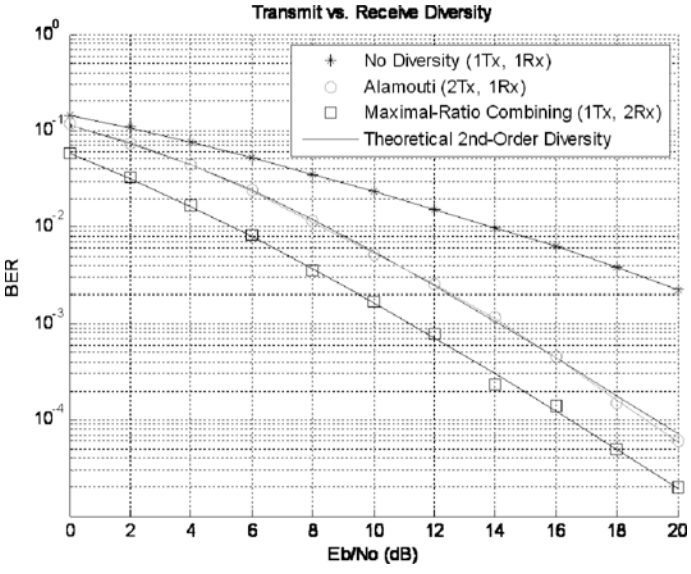


Fig. 2.6 The performance of the Alamouti space-time code with different numbers of antennas at the receiver

2.4.2 Spatial Multiplexing

In the layered space-time codes for spatial multiplexing, the sequence of information bits is divided into certain sets of substreams. The substream where the signal processing is conducted is referred as a layer. Hence, this is known as the layer space-time (LAST) technique. As shown in Figure 2.7, M_T independent substreams are transmitted through M_T antennas. In this technique, the number of receiving antennas must be equal to or larger than the number of transmitting antennas [7].

The process of dividing the sequence of information into substreams is done by a demultiplexer. The demultiplexing process may be applied to bits or symbols. Hence, the encoding can be realized in three different ways according to the demultiplexer position in the transmitter chain and the direction of the layer [1]. These encoding processes are referred to as horizontal, vertical and diagonal encoding. The different encoding techniques are shown in Figure 2. 8.

In horizontal encoding, the data bits are demultiplexed into M_T substreams that are independently encoded, interleaved and modulated. On the other hand, in the vertical realization, the data stream is encoded, interleaved and modulated; and, the resulting symbols are then demultiplexed into M_T substreams. The process to realize the diagonal spatial multiplexing is similar to horizontal encoding with the only difference being that, after the final stage, the frames of symbols undergo a stream interleaver, which rotates the transmitted frames [7], [8].

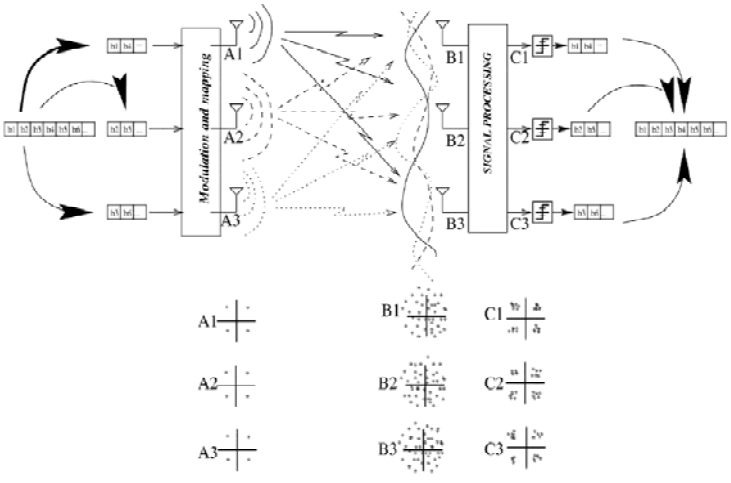


Fig. 2.7 Spatial multiplexing with three antennas at the transmitter and three antennas at the receiver [8]

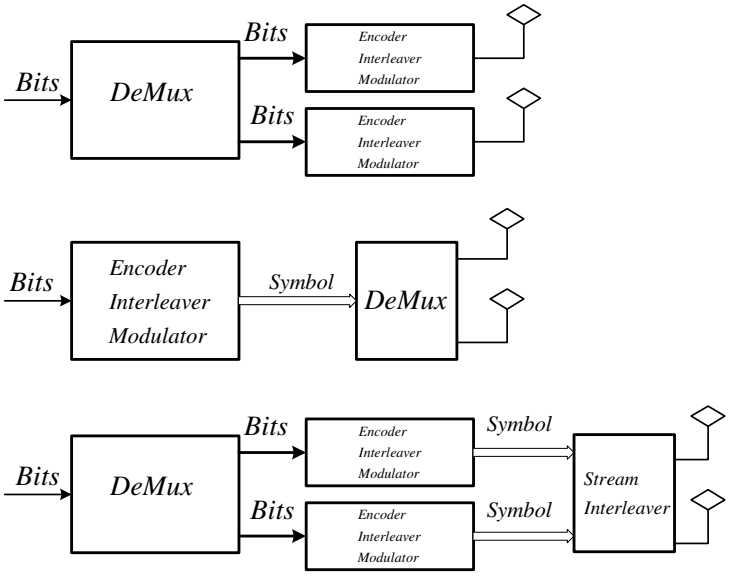


Fig. 2.8 Spatial multiplexing encoding: a) horizontal encoding, b) vertical encoding, c) diagonal encoding

2.4.3 Diversity-Multiplexing Tradeoff

The channel capacity of a MIMO system can be considerably higher than that of a SISO system using layered space-time coding. This performance of a MIMO channel provides spatial multiplexing gain. On the other hand, MIMO systems can improve the reliability of a wireless link; and, this performance provides diversity gain. Traditionally, MIMO systems have been designed to extract either maximal diversity or spatial multiplexing gain.

MIMO systems, however, can provide both diversity gain and multiplexing gain. In [9], an optimal tradeoff curve between the asymptotic diversity and the multiplexing gain has been derived at infinite SNR. The diversity gain is obtained as:

$$d = \lim_{\rho \to \infty} -\frac{\log P_{out}}{\log \rho} \tag{2.24}$$

The spatial multiplexing gain is defined as:

$$r = \lim_{\rho \to \infty} \frac{R}{\log \rho} \tag{2.25}$$

where d is the diversity gain, r is the multiplexing gain, P_{out} is the outage probability, R is the data rate in bps/Hz, and ρ is average SNR per antenna. It has been shown that the most famous space-time coding schemes, such as space-time block code (STBC) and Bell layered space-time (BLAST), are not optimal with respect to the diversity-multiplexing tradeoff (DMT) criteria. The asymptotic diversity multiplexing tradeoff and the finite SNR diversity multiplexing tradeoff for MISO/SIMO and MIMO are presented in Figures 2.9 and 2.10 [14].

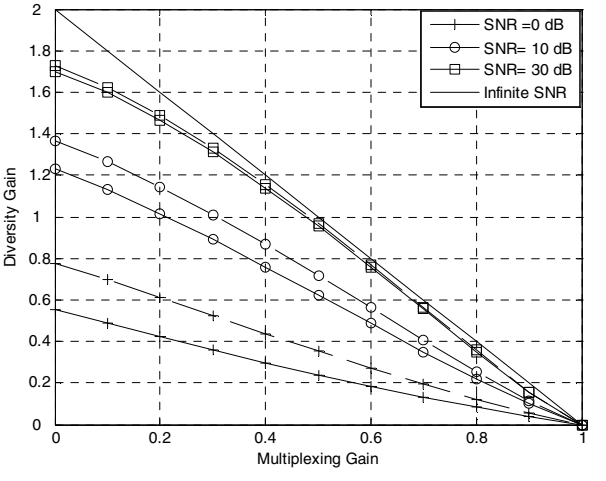


Fig. 2.9 Diversity-multiplexing tradeoff curve for SIMO (dashed lines) and MISO (solid lines) using two antennas

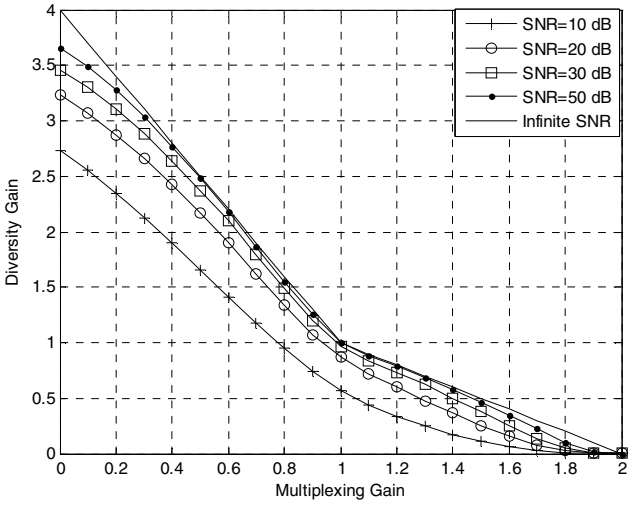


Fig. 2.10 Diversity-multiplexing tradeoff curve for MIMO (2 X 2)

2.5 MIMO Channel Models

There are different classifications for MIMO channels [2], [16]. As shown in Table 2.1, a classification based on the modeling approach divides the models into physical models and analytical models. Physical channel models characterize a channel on the basis of electromagnetic wave propagation. They explicitly model wave propagation parameters, such as the complex amplitude and delay. More sophisticated models also incorporate polarization and time variation. Physical models are independent of antenna configurations (antenna pattern, number of antennas, array geometry, polarization, mutual coupling) and system bandwidth [16].

Physical MIMO channel models can further be split into deterministic models, geometry-based stochastic models, and non-geometric stochastic models. Deterministic models characterize the physical propagation parameters in a completely deterministic manner, e.g. ray tracing. Using geometry-based stochastic channel models, the impulse response is characterized by the laws of wave propagation applied to specific transmitters (Tx) and receivers (Rx) and scatterer geometries, which are chosen in a stochastic (random) manner. In contrast, non-geometric stochastic models describe and determine physical parameters in a completely stochastic way by prescribing underlying probability distribution functions without assuming an underlying geometry [17].

In contrast to physical models, analytical channel models characterize the impulse response of the channel between the individual transmitting and receiving antennas in a mathematical/analytical way without explicitly accounting for wave

propagation. The individual impulse responses are subsumed in a MIMO channel matrix. Analytical models are very popular for synthesizing MIMO matrices in the context of system and algorithm development and verification.

Table 2.1 MIMO Channel Modeling Classification

Physical Models	Analytical Models
Deterministic, e.g. ray tracing	*Correlation based*
Geometry-based Stochastic	*Propagation motivated*
Non-geometrical stochastic, e.g. Saleh-Valenzuela's method [17], [19]	

Analytical models can be further subdivided into correlation-based models and propagation-motivated models. Correlation-based models characterize the MIMO channel matrix statistically, in terms of the correlations between the matrix entries. The propagation-motivated models characterize the channel matrix via propagation parameters.

On the other hand, empirical methods can be extracted from analytical and physical models based on experimental results [18], [19]. They either generalize the tap-delay line concept to include the directional domains or use a combination of tap-delay lines and geometry-based models with prescribed parameters. These models are very useful for simulation purposes. They also constitute the background of many standardized channel models, such as 3GPP, COST 259, COST 273, IEEE 802.16a, IEEE 802.16e, and IEEE 802.11n.

2.5.1 Stanford University Interim Channel Models

In this section, the Stanford University Interim (SUI) channel models of MIMO systems are briefly discussed [2], [19], [20], [24]. SUI channel models have been developed for fixed wireless access networks at 2.5 GHz [16]. The general architecture of the SUI model is presented in Figure 2.11.

The input-mixing matrix models transmitter antenna correlations, and the output-mixing matrix presents the receiver antenna correlations. The SUI channel models basically use a three tap-delay line to model the multipath fading [4], [20]. Antennas are assumed to be omnidirectional on both the transmitting and receiving sides. The Doppler spectrum is assumed to be given by:

$$S(\upsilon) = 1 - 1.720\left(\frac{\upsilon}{\upsilon_m}\right)^2 + 0.785\left(\frac{\upsilon}{\upsilon_m}\right)^4 \qquad \upsilon \le \upsilon_m \tag{2.26}$$

where υ_m is the maximum Doppler frequency.

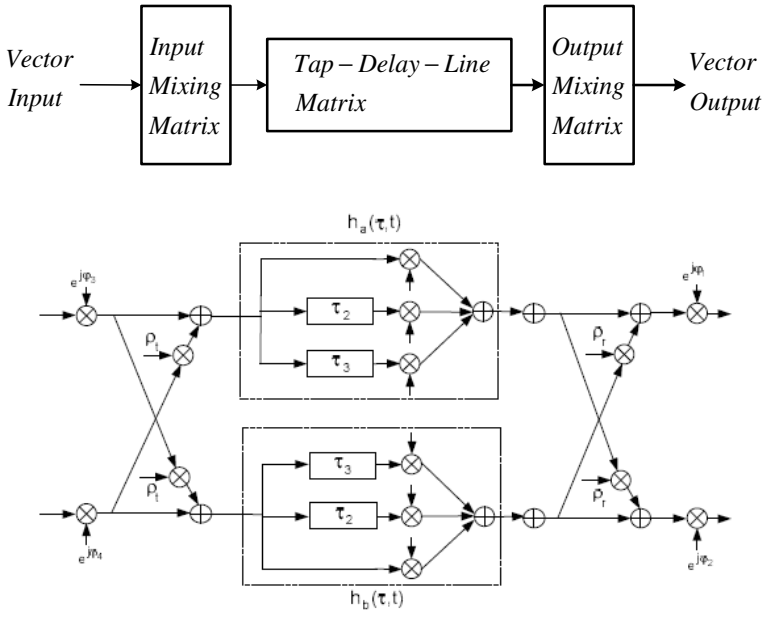


Fig. 2.11 Schematic of a SUI channel [4]

The models describe six types of channels, numbered from 1 to 6 [21]. As an example, the parameters of SUI channel 1 are presented in Table 2.2.

Table 2.2 The channel parameters for SUI-1 model

SUI-1 Model	Tap 1	Tap 2	Tap 3
Delay [microsec]	0.0	0.4	0.9
Power [dB]	0	-15	-20
99% K-factor	4	0	0
Doppler Frequency [Hz]	0.4	0.3	0.5
Envelope Antenna Correlation	0.7	0.7	0.7

■ ***Example 2.4:*** *IEEE 802.16d/e Models*

These models are intended for macro-cellular fixed wireless access [24]. The targeted scenario is as follows:

- The cell size is less than 10 km in radius,
- The user's antenna is installed under the eaves or on the rooftop, and
- The base station height is 15 to 40 m.

In fact, the IEEE 802.16 model is an improved version of the SUI channel models and is valid for both omnidirectional and directional antennas. The use of directional antennas cause the global K-factor to increase, while the delay spread decreases. As an example, Table 2.3 indicates how SUI channel 4 is modified when the terminal antenna has a 30-degree beamwidth. An additional feature of the IEEE 802.16 standard is a model for the narrowband Ricean K-factor:

$$K = K_0 F_s F_h F_b R^{\gamma} u \tag{2.26}$$

where F_s is a seasonal factor, $F_s = 1.0$ in summer (leaves) and 2.5 in winter (no leaves); F_h is the receiving antenna height factor, $F_h = 0.46(h/3)$ (h is the receiving antenna height in meters); F_b is the beamwidth factor, $F_b = -0.62(b/17)$ (b in degrees); K_0 and γ are regression coefficients, $K_0 = 10$ and $\gamma = -0.5$; and, u is a log-normal variable, i.e. $10\log_{10}(u)$ is a zero-mean normal variable with a standard deviation of 8 dB.

Table 2.3 IEEE 802.16 Model

IEEE 802.16 Model	Tap 1	Tap 2	Tap 3
Delay [microsec]	0.0	1.5	4
Power [dB]	0	-10	-20
99% K-factor	1	0	0
Doppler Frequency [Hz]	0.2	0.15	0.25
Envelope Antenna Correlation	0.3	0.3	0.3

References

[1] Carlson, A.B., Crilly, P.B., Rutledge, J.C.: Communication Systems: An Introduction to Signal and Noise in Electrical Communications, 4th edn. McGraw Hill (2001)

[2] Oestges, C., Clerckx, B.: MIMO Wireless Communications: From Real World Propagation to Space Time Code Design. Academic Press (2007)

[3] Pahlavan, K., Levesque, A.: Wireless Information Networks, 2nd edn. John Wiley and Sons (2005)

[4] Paulraj, A., Nabar, R., Gore, D.: Introduction to Space-Time Wireless Communications. Cambridge University Press (2003)

[5] Lee, W.C.Y.: Estimation of channel capacity in Rayleigh fading environment. IEEE Transactions on Vehicular Technology 39(3), 187–189 (1990)

[6] Mohammadi, A., Kumar, S.: Characterization of Adaptive Modulators in Fixed Wireless ATM Networks. IEEE/KICS Journal of Communications and Networks 6(2), 123–132 (2004)

[7] Tsoulos, G.: MIMO System Technology for Wireless Communications. CRC Press (2006)

[8] Gesbert, D., Shafi, M., Shiu, D., Smith, P.J., Naguib, A.: From Theory to Practice: An Overview of MIMO Space–Time Coded Wireless Systems. IEEE Journal on Selected Areas in Communications 21(3), 281–302 (2003)

[9] Zheng, L., Tse, D.N.C.: Diversity and multiplexing: a fundamental tradeoff in multiple antenna channels. IEEE Transactions on Information Theory 49, 1073–1096 (2003)
[10] Alamouti, S.M.: A simple transmit diversity technique for wireless communications. IEEE Journal on Selected Areas in Communications 16(10), 1451–1458 (1998)
[11] Ebrahimzad, H., Mohammadi, A.: On Diversity-Multiplexing Tradeoff in MIMO channel at Finite SNR. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences E93-A(11), 2057–2064 (2010)
[12] Tarokh, V., Seshadri, N., Calderbank, A.R.: Space-time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction. IEEE Transactions on Information Theory 44(2), 744–765 (1998)
[13] Gershman, A.B., Sidiropoulos, N.D.: Space-Time Processing for MIMO Communications. Wiley (2005)
[14] Ebrahimzad, H., Mohammadi, A.: Diversity-Multiplexing Tradeoff in MIMO Systems with Finite SNR. In: European Conference on Wireless Technology, Munich, pp. 146–149 (October 2007)
[15] Papoulis, A., Pillai, S.U.: Random Variable Variables and Stochastic Process, 4th edn. McGraw Hill (2002)
[16] Almers, P., Bonek, E., Burr, A., Czink, N., Debbah, M., Degli-Esposti, V., Hofstetter, H., Kyosti, P., Laurenson, D., Matz, G., Molisch, A.F., Oestges, C., Ozcelik, H.: Survey of Channel and Radio Propagation Models for Wireless MIMO Systems. EURASIP Journal on Wireless Communications and Networking 2007, article ID 19070, 19 pages (2007)
[17] Saleh, A.M., Valenzuela, R.A.: A statistical model for indoor multipath propagation. Journal on Selected Areas in Communications 5(2), 128–137 (1987)
[18] Rappaport, T.S.: Wireless Communications: Principles and Practice, 2nd edn. Prentice Hall (2002)
[19] Wallace, J.W., Jensen, M.A.: Modeling the indoor MIMO wireless channel. IEEE Transactions on Antennas and Propagation 50(5), 591–599 (2002)
[20] Baum, D.S., Gore, D.A., Nabar, R.U., Panchanathan, S., Hari, K.V.S., Erceg, V., Paulraj, A.J.: Measurement and characterization of broadband MIMO fixed wireless channels at 2.5 GHz. In: Proceedings of the International Conference on Personal Wireless Communications (ICPWC 2000), India (December 2000)
[21] Erceg, V., et al.: IEEE p802.16 – channel models for fixed wireless applications (ieee802.16.3c-01/29r4) (2001)
[22] Correia, L.M.: COST 259 – Wireless flexible personalized communications. Wiley, London (2001)
[23] Correia, L.M.: COST 273 – Towards mobile broadband multimedia networks. Elsevier, London (2006)
[24] IEEE P802.16 e /D12, IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment for Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands (October 2005)
[25] Flaviis, F.D., Jofre, L., Romeu, J., Grau, A.: Multiantenna Systems for MIMO Communications. Morgan & Craypool Publishers (2008)

Chapter 3
Digital Modulation Techniques in MIMO Systems

The basic work related to digital communication techniques was carried out by Shannon [1]. Since then, the theory and applications of digital communication systems have been greatly advanced [2-6]. With the introduction of mobile and cellular communications, digital communication techniques have been investigated for wireless systems [7-10]. In addition, a new stream of research on and implementation of digital communication techniques has been established with the advent of multiple input multiple output (MIMO) communication systems during the last decade, [11-14].

This chapter provides an overview of different digital modulation techniques and their applications in single input single output (SISO) and MIMO wireless systems with an emphasis on their implementation techniques. The fundamentals of digital modulation techniques are presented. Then, the analysis and the modeling of the various modulation schemes both in single carrier, multicarrier and orthogonal frequency-division multiplexing (OFDM) systems are described. The last section of the chapter is devoted to the accuracy metrics of digital modulation implementation.

3.1 Criteria to Design Digital Modulators

The design of any digital communication system is initiated with a description of the channel (received power, available bandwidth, attenuation, noise statistics, and other impairments such as fading) and a definition of the system requirements (data rate and error performance). Design choices that match the channel and meet the performance requirements then need to be determined [4], [15]. The general criteria to select a digital modulation method are spectral efficiency, bandwidth efficiency, and implementation complexity. These criteria are even more critical in wireless communications, due to spectrum limitations. In addition, the proper operation of a digital modulation scheme in a wireless channel is an important factor that must be considered.

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 27–54.
springerlink.com

3.1.1 Spectral Efficiency

The spectral efficiency of a digital modulator is defined as the number of bits per second that can be transmitted in one Hertz of system bandwidth. A suitable starting point to describe the spectral efficiency is the Shannon capacity theorem, which can be stated as:

$$C = W \log_2(1 + \frac{S}{N}) \tag{3.1}$$

where S/N is the ratio of the average received signal power to the noise power, and W is the channel bandwidth. The capacity, *C*, is given in bit/sec. The capacity of a channel defines the maximum number of bits that can be reliably sent per second over the channel. Accordingly, the maximum spectral efficiency, $\eta_{\max}$, is defined as:

$$\eta_{\max} \triangleq \frac{C}{W} = \log_2(1 + \frac{S}{N}) \tag{3.2}$$

Relationship (3.1) can be written as:

$$\frac{C}{W} = \log_2(1 + \frac{S}{N_o W}) \tag{3.3}$$

where N_o is the power spectral density of white noise.

On the other hand, for any digital communication system, the relationship between the received S/N and the received bit energy to noise-power spectral density, E_b/N_o, is as follows [4],[15]:

$$\frac{S}{N} = \frac{E_b}{N_o} r_b \tag{3.4}$$

where r_b is the data rate in bits per second.

Figure 3.1 shows the relationship between r_b/W and E_b/N_o for the case where the data (information) rate, r_b, is equal to *C*. As can be seen, a curve separates a region of practical communication systems from a region where such communication systems cannot operate reliably. In the spectral efficient region, a modulation scheme can be designed to transmit a number of bits in one hertz using a practical E_b/N_o.

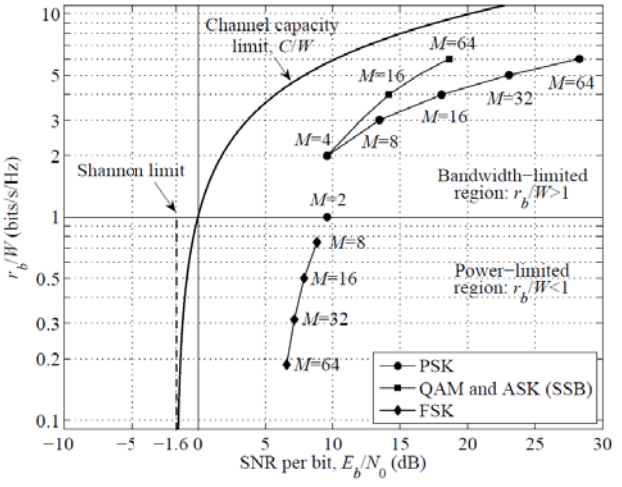


Fig. 3.1 The bandwidth efficiency of various digital modulation systems[15]

3.1.2 Power Efficiency

The power efficiency of a digital modulator is defined as the required signal-to-noise ratio (SNR) for a certain bit error rate (BER) probability over an additive white Gaussian noise (AWGN) channel. The power spectral efficiency can be explained with the use of Figure 3.2. This figure shows the waterfall curves for

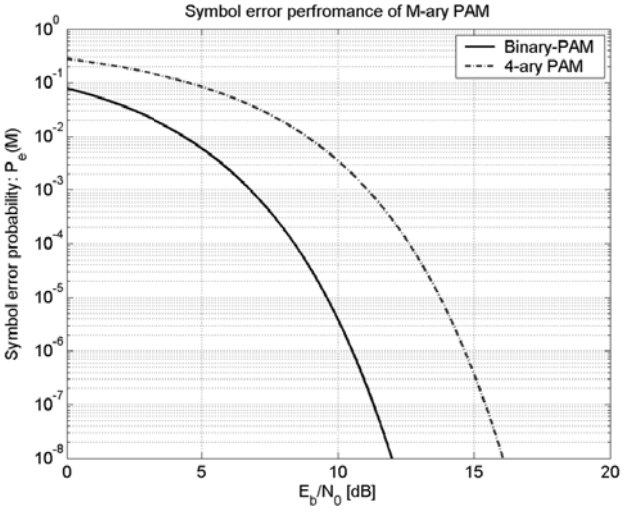


Fig. 3.2 The BER curves versus E_b / N_o for Binary PAM and 4-ary PAM modulation in AWGN channel

two modulation schemes – binary pulse amplitude modulation (B-PAM) and 4-ary PAM. The B-PAM modulation scheme requires less E_b / N_o to provide a certain BER probability compared to the 4-ary PAM modulation scheme. Therefore, B-PAM is more power efficient than 4-ary PAM.

3.1.3 Implementation Complexity

The complexity of the modulator implementation has a considerable impact on the performance and cost. As can be seen from Figure 3.2, the M-ary modulation techniques provide either better bandwidth efficiency, e.g., MQAM (M-ary quadrature amplitude modulation) and MPSK (M-ary phase-shift keying), or better power efficiency, e.g., MFSK (multiple frequency-shift keying). However, their implementations are complex and costly. On the other hand, some constant envelope modulation techniques, such as Guassian minimum-shift keying (GMSK), can be easily implemented, but their bandwidth efficiencies are usually low. In wireless applications, although multicarrier modulations, such as OFDM, require complex and costly implementations, they provide better performance in multipath fading channels [7], [9].

3.1.4 Performance Study in Wireless Channel

Multipath channels are characterized by coherence bandwidth. The coherence bandwidth is the maximum channel width that can be used and is extracted from the power delay profile of a wireless channel [7], [8]. To estimate the capacity in a multipath fading channel, the Shannon capacity is examined first. The Shannon capacity for an AWGN channel is obtained using (3.1). For a fading channel, the Shannon limit can be modified as [5], [16]:

$$\frac{\prec C \succ}{W} = \int_{\gamma} \log_2(1+\gamma) p(\gamma) d\gamma \tag{3.5}$$

where $\gamma = S / N$ and $p_\gamma(\gamma)$ is the probability density functions of the received signal.

In a fading channel, the received signal power varies randomly over distance or time in a fading channel. Accordingly, γ is0020a random variable with a distribution of $p_\gamma(\gamma)$. Therefore, the capacity in a fading channel decreases by a factor:

$$\xi_c = \frac{\prec C \succ}{C} = \frac{\int_{\gamma} \log_2(1+\gamma) p(\gamma) d\gamma}{\log_2(1+\gamma)} \tag{3.6}$$

where ξ_c is called the capacity reduction factor.

On the other hand, if the signal fade level is constant at about one symbol time, the average error probability is a suitable performance metric [8]. The average error probability is computed by integration of the error probability in the AWGN channel over the fading distribution. It is obtained as:

$$\overline{P}_s = \int_{\gamma} P_s(\gamma)\, p_{\gamma}(\gamma) d\gamma \tag{3.7}$$

where $P_s(\gamma)$ is the probability of the symbol error in AWGN with a SNR equal to γ. The symbol error rate for coherent modulation can be shown as a generic relationship [8]:

$$P_s(\gamma) \approx \alpha_M Q(\sqrt{\beta_M \gamma}) \tag{3.8}$$

where α_M and β_M depend on the modulation and its approximation and $Q(z) = \int_z^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$. This generic relationship is depicted in Table 3.1 for QAM modulation.

Table 3.1 Approximate Symbol and Bit Error Probabilities for Bandwidth Efficient Modulation Schemes [8]

Modulation	$P_s(\gamma_s)$	$P_b(\gamma_b)$
MPAM	$P_s = \frac{2(M-1)}{M} Q\left(\sqrt{\frac{6\overline{\gamma}_s}{M^2-1}}\right)$	$P_b \approx \frac{2(M-1)}{M \log_2 M} Q\left(\sqrt{\frac{6\overline{\gamma}_b \log_2 M}{M^2-1}}\right)$
MPSK	$P_s \approx 2Q\left(\sqrt{2\overline{\gamma}_s} \sin\left(\frac{\pi}{M}\right)\right)$	$P_b \approx \frac{2}{\log_2 M} Q\left(\sqrt{2\gamma_b \log_2 M} \sin\left(\frac{\pi}{M}\right)\right)$
MQAM	$P_s \approx 4Q\left(\sqrt{\frac{3\overline{\gamma}_s}{M-1}}\right)$	$P_b \approx \frac{4}{\log_2 M} Q\left(\sqrt{\frac{3\overline{\gamma}_b \log_2 M}{M-1}}\right)$

On the other hand, for a Rayleigh fading channel, the probability density function of the received SNR is given by:

$$p_{\gamma}(\gamma) = \frac{1}{\Gamma} e^{-\gamma/\Gamma} \tag{3.9}$$

where Γ is the average power of γ.

Thus, according to (3.7), the average probability of error in Rayleigh fading can be approximated as [8]:

$$\overline{P}_s = \int_{\gamma} P_s(\gamma) p_{\gamma}(\gamma) d\gamma = \int_{\gamma} \alpha_M Q(\sqrt{\beta_M \gamma}) \frac{1}{\Gamma} e^{-\gamma/\Gamma} d\gamma \approx \frac{\alpha_M}{2\beta_M \Gamma} \tag{3.10}$$

The last approximation is more accurate for higher SNRs. The symbol error rate in a Rayleigh fading channel decreases linearly with an increasing average SNR. However, the symbol error rate in an AWGN channel decreases exponentially with increasing SNR.

Similarly, the average capacity concept can be extended to MIMO channels. The average capacity of a MIMO channel in Rayleigh fading can be represented as [17], [18]:

$$\frac{\prec C \succ}{W} = \frac{1}{Ln(2)} \int_0^{\infty} Ln(1+\rho y) e^{-y} \sum_{k=0}^{s-1} \frac{k! y^{t-s}}{(k+t-s)!} [L_k^{t-s}(y)]^2 dy \tag{3.11}$$

where $t = \max(L_t, L_r), s = \min(L_t, L_r), \rho$ is the normalized transmitting SNR per branch, and $L_n^a(.)$ is the generalized Laguerre polynomial which is expressed as:

$$L_n^{\alpha}(x) = \frac{x^{-\alpha} e^x}{n!} \frac{d^n}{dx^n} (e^{-x} x^{n+\alpha}) \tag{3.12}$$

3.2 Single-Carrier SISO Systems

3.2.1 Single-Carrier Modulation

In digital communication systems, each symbol in an M-ary alphabet is related to a unique sequence of *k* bits, expressed as:

$$M = 2^k \qquad k = \log_2 M \tag{3.13}$$

where *M* is the size of the alphabet. In the case of digital transmissions, the term "symbol" refers to the member of the M-ary alphabet that is transmitted during each symbol duration, T_s. In order to transmit the symbol, it must be mapped onto an electrical voltage or current waveform. Since one of *M* symbols or waveforms is transmitted during each symbol duration, T_s , the data rate, r_b, in b/s, can be expressed as [4], [6]:

$$r_b = \frac{1}{T_b} = \frac{k}{T_s} = \frac{\log_2 M}{T_s} \qquad bit/\sec \tag{3.14}$$

From (3.10), one can write that the effective time duration, T_b, of each bit in terms of the symbol duration, T_s, or the symbol rate, R_s, is:

$$T_b = \frac{1}{r_b} = \frac{T_s}{k} = \frac{1}{kR_s} \tag{3.15}$$

Using (3.9) and (3.11) .we can express the symbol rate, R_s, in terms of the bit rate, r_b, as follows:

$$R_s = \frac{r_b}{\log_2 M} \qquad Symbol/\sec \tag{3.16}$$

From (3.10) and (3.11), one can see that any digital scheme that transmits $k = log_2 M$ bits in T_s seconds using a bandwidth of W Hertz, operates at a bandwidth efficiency of:

$$\eta = \frac{r_b}{W} = \frac{\log_2 M}{WT_s} = \frac{1}{WT_b} \tag{3.17}$$

where T_b is the effective time duration of each data bit [4].

From (3.13), the smaller the WT_b product, the more bandwidth efficient is any digital communication system. Thus, signals with small WT_b products are often used in systems of limited bandwidth. MQAM and MPSK modulation schemes are two examples of bandwidth efficient modulation techniques that are often used in practical radio systems. The symbol and bit error probabilities for the bandwidth efficient modulation schemes based on $\gamma_b = E_b / N_o$ and $\gamma_s = E_s / N_o$ are presented in Table 3.1 [8]. The $\overline{\gamma}_s$ is equal to the average energy per symbol.

In the following subsections, the MQAM and MPSK modulation techniques are explained in more detail. The different spectral efficient modulation techniques are discussed extensively in literature [2-10].

3.2.2 MQAM

M-ary quadrature amplitude modulation (MQAM) is a two-dimensional linear modulation scheme and it may be represented as:

$$s_m(t) = A_{mc} g(t) cos 2\pi f_c t - A_{ms} g(t) sin 2\pi f_c t, \qquad m = 1, 2, \ldots, L \tag{3.18}$$

where $g(t)$ is a signal pulse shape, f_c is carrier frequency, A_{mc} and A_{ms} are the in-phase (I) and quadrature (Q) signal amplitudes, and $L = \sqrt{M}$.

The constellation diagram of MQAM modulation in the I-Q plane is illustrated in Figure 3.3 [6].

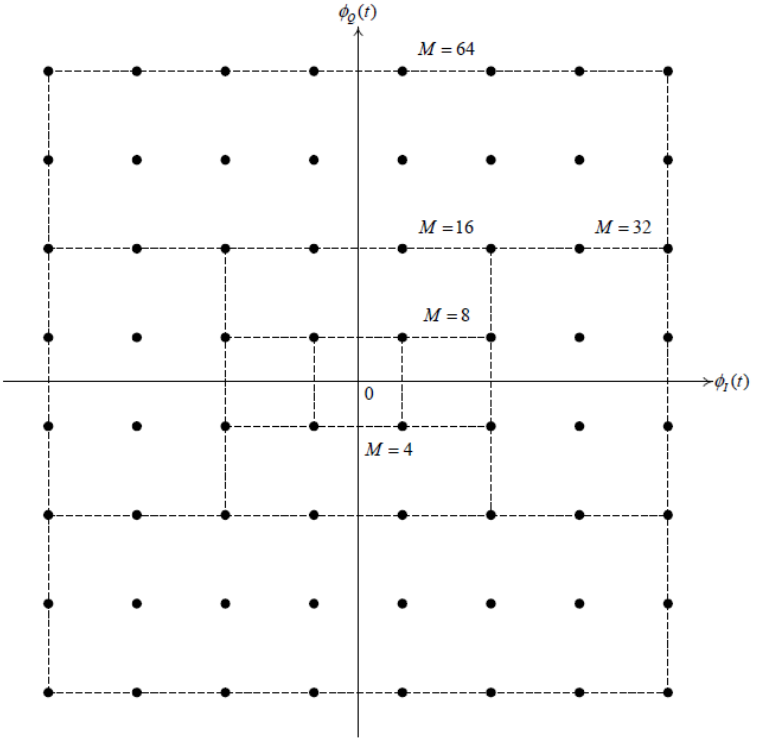


Fig. 3.3 The constellation diagram of MQAM modulator with different constellation sizes

3.2.2.1 MQAM Spectral Efficiency

The low-pass equivalent of an MQAM signal is given by:

$$v(t) = \sum_n (a_n + jb_n) g(t - nT_s); \quad a_n, b_n = \pm 1, \pm 3, \ldots, \pm L - 1 \tag{3.19}$$

where T_s is the symbol rate, and $g(t)$ is the pulse shape. The band pass signal, $s_m(t)$, can be related to the low-pass signal, $v(t)$, through:

$$s_m(t) = Re\left[v(t) e^{j2\pi f_c t} \right] \tag{3.20}$$

The autocorrelation function of $s_m(t)$ is given by:

$$\varphi_{ss}(\tau) = Re\left[\phi_{vv}(\tau) e^{j2\pi f_c t} \right] \tag{3.21}$$

where $\varphi_{vv}(\tau)$ is the autocorrelation function of the equivalent low-pass signal, $v(t)$.

The power spectral density of $s_m(t)$ is then given by:

$$\phi_{ss}(f) = \frac{1}{2}\left[\phi_{vv}(f - f_c) + \phi_{vv}(-f - f_c)\right] \tag{3.22}$$

where $\phi_{vv}(f)$ is the power spectral density of $v(t)$. It can be shown [2] that the power spectral density of $v(t)$ is given by:

$$\phi_{vv}(f) = \frac{1}{2}|G(f)|^2 \phi_{ii}(f) \tag{3.23}$$

where $G(f)$ is the Fourier transform of *g(t)* and $\phi_{ii}(f)$ is the power spectral density of the information sequence.

The QAM signal may be considered to consist of two pulse amplitude modulation (PAM) signals conveyed by the cosine and sine carriers. The spectral efficiency of the QAM will be twice that of its PAM components in the I and Q channels. Substitution for $\phi_{ii}(f)$ in the above equation [2], gives the power spectral density signal carriers by the cosine carrier as:

$$\phi_{vv}(f) = \frac{\sigma^2}{T_s}|G(f)|^2 + \frac{\mu^2}{T^2}\sum_{m=-\infty}^{\infty}\left|G\left(\frac{m}{T_s}\right)\right|^2 \delta(f - \frac{m}{T}), \tag{3.24}$$

where μ and σ^2 are the mean and variance of the information sequence.

It is known that the mean value of the PAM symbols is zero ($\mu = 0$), and, thus, this equation can be simplified as:

$$\phi_{vv}(f) = \frac{\sigma^2}{T_s}|G(f)|^2 \tag{3.25}$$

Hence, the spectral efficiency of the PAM signal is controlled by the pulse shape, $g(t)$. The raised cosine pulse is commonly used as a pulse shape for the digital modulations in bandwidth-limited channels. The signal $g(t)$ for this pulse shape is given by:

$$g(t) = \frac{\sin(\frac{\pi t}{T_s})}{\frac{\pi t}{T_s}} \frac{\cos(\frac{\pi \alpha t}{T_s})}{1 - (\frac{2\alpha t}{T_s})^2} \tag{3.26}$$

where α is called the roll of factor.

The Fourier transform $g(t)$ is given by:

$$G(f)=\begin{cases} T \quad if \quad 0\le|f|\le\dfrac{1-\alpha}{2T_s} \\ \dfrac{T_s}{2}[1+\cos(\dfrac{\pi T_s}{\alpha}(|f|-\dfrac{1-\alpha}{2T_s}))] \quad if \quad \dfrac{1-\alpha}{2_s}\le|f|\le\dfrac{1+\alpha}{2T_s} \\ 0 \quad if \quad |f|\ge\dfrac{1+\alpha}{2T_s} \end{cases} \tag{3.27}$$

The normalized raised cosine filter in the frequency and time domains are illustrated in Figure 3.4. Thus, the bandwidth efficiency of a QAM modulator can be adjusted by parameter α.

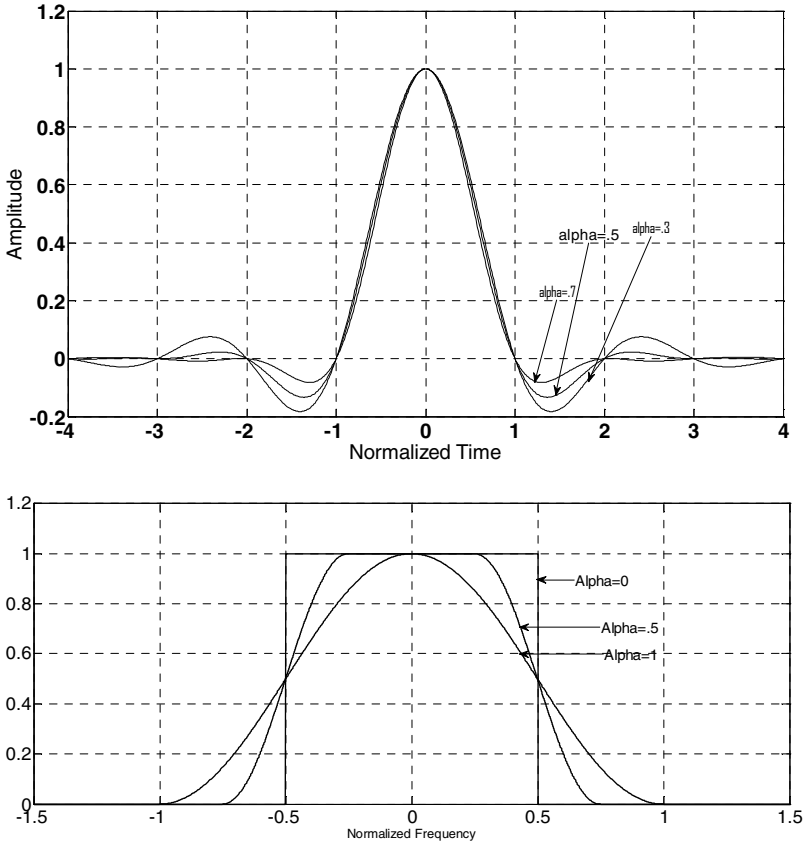


Fig. 3.4 The pulse shape and the spectrum of a raised cosine pulse shape

As can be seen from Figure 3.4, the channel bandwidth, *W*, is approximately equal to $\frac{1+\alpha}{T_s}$; and, since $\frac{1}{T_s} = r_b / \log_2(M)$, we obtain the result of:

$$\eta = \frac{r_b}{W} = \frac{\log_2 M}{1+\alpha}, \tag{3.28}$$

where η is the QAM spectral efficiency, and *R* is the QAM transmission rate.

3.2.2.2 MQAM Power Efficiency Performance

As explained, the power efficiency is a measure of the received power needed to achieve a specified BER (bit error rate). For an AWGN channel, the probability of bit error in the MQAM system for the coherent modulation is obtained as [19]:

$$\begin{aligned} P_{b,MQAM} = \frac{4(1-\frac{1}{\sqrt{M}})}{\log_2 M}\Bigg\{ & Q\left(\sqrt{\frac{3\gamma_s}{M-1}}\right)+ \\ & +Q\left(3*\sqrt{\frac{3\gamma_s}{M-1}}\right)+Q\left(5*\sqrt{\frac{3\gamma_s}{M-1}}\right)+\ldots\Bigg\} \end{aligned} \tag{3.29}$$

It is usually a good approximation to compare the first terms, if one ignores the other terms. Accordingly, the BER of QAM modulation can be written as [20]:

$$P_b \approx \frac{4(\sqrt{M}-1)}{\sqrt{M}\log_2(M)} Q\left(\sqrt{\frac{3\gamma_s}{M-1}}\right) \tag{3.30}$$

The BER of MQAM modulation schemes, assuming perfect clock and carrier recovery, are shown in Figure 3.5 [20]. The symbol error rate of an ideal rectangular QAM modulation can be obtained as [2]:

$$\begin{aligned} P_M &= 1-\left(1-2P_{\sqrt{M}}+(P_{\sqrt{M}})^2\right) \\ &= 2P_{\sqrt{M}}-(P_{\sqrt{M}})^2 \end{aligned} \tag{3.31}$$

where

$$P_{\sqrt{M}} = 2\left(1-\frac{1}{\sqrt{M}}\right)\times Q\left(\sqrt{\frac{3\gamma_s}{M-1}}\right) \tag{3.32}$$

Similarly, if one can ignore the second term, in order compare to the first terms in (3.31), the symbol error rate of an ideal rectangular MQAM modulation can be obtained as:

$$P_M \simeq 2P_{\sqrt{M}} = 4\left(1-\frac{1}{\sqrt{M}}\right)\times Q\left(\sqrt{\frac{3\gamma_s}{M-1}}\right) \tag{3.33}$$

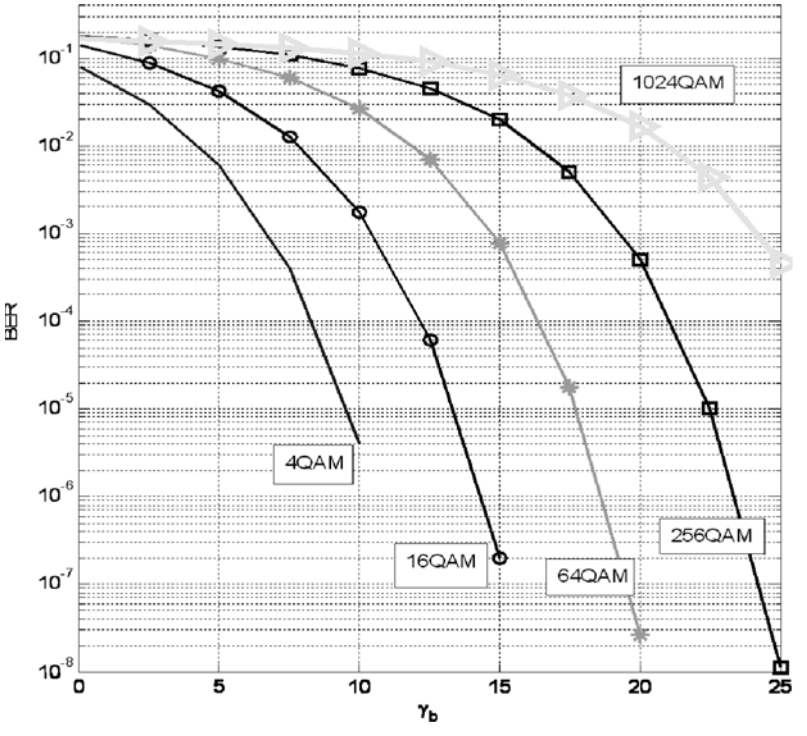


Fig. 3.5 Bit error rate versus Eb/No for ideal MQAM modulation in an AWGN channel

3.2.2.3 MQAM Modulator Implementation

The QAM modulator has been traditionally realized using the heterodyne method shown in Figure 3.6. The QAM mapper in the figure divides a binary sequence with a rate of f_b into two binary symbol streams, each with a rate of $f_b/2$.

A 2-to-$L = \sqrt{M}$ level baseband converter is used to convert the streams into $L = \sqrt{M}$ -level baseband PAM signals in the I or Q paths. These symbols are shaped by raised cosine filters. The filtered I and Q baseband signals are used to modulate the I and Q outputs of an intermediate frequency (IF) oscillator. The modulated signal at the IF is up-converted to the desired transmitting frequency in one or more steps. Mixers are commonly used in this realization. A power amplifier is used to boost the modulated signal at the transmitting frequency to the required power.

It should be emphasized that, if the IF frequency is low, the classic heterodyne design described above suffers from disadvantages, such as design complexities, high costs, radio frequency (RF) filter requirements, and data rate limitations.

Direct implementation methods are often preferred to the classic design in wireless systems, where these limitations can be effectively overcome. The basic idea of direct microwave implementation is the use of the baseband signal to modulate a carrier at the desired transmitted frequency. This technique is examined in detail in the following chapters.

The MQAM receiver can be also implemented using either superheterodyne or direct conversion methods. A general block diagram of an MQAM receiver using a superheterodyne method is shown in Figure 3.7.

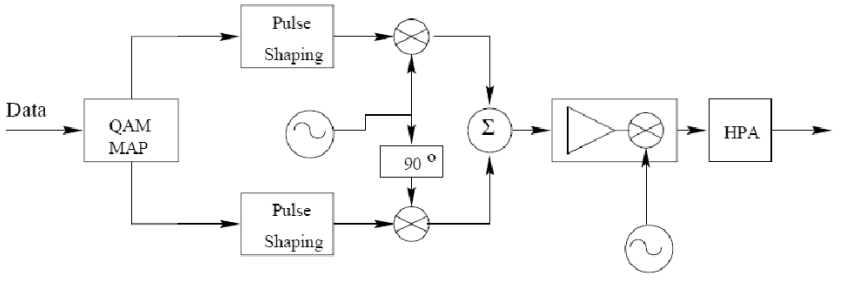


Fig. 3.6 A superheterodyne implementation of an MQAM transmitter

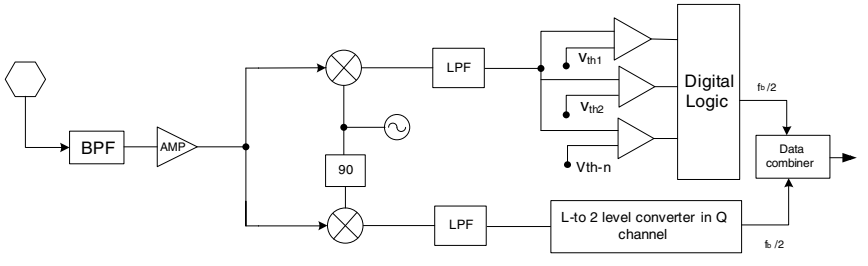


Fig. 3.7 A superhetrodyne implememntation of an MQAM receiver

3.2.3 MPSK

In the MPSK modulation scheme, all of the information is encoded in the phase of the transmitted signal. Therefore, the transmitted signal over symbol period, T_s ,is [23], [8]:

$$s_m(t)=\sqrt{\frac{2E_s}{T_s}}\cos\left(2\pi f_c t+\frac{2\pi(m-1)}{M}\right) \qquad 0\le t\le T_s \qquad m=1,2,....,M \tag{3.34}$$

where $E_s=\int_{-\infty}^{\infty}x_m^2(t)dt$.

The most common MPSK technique is quadrature phase-shift keying (QPSK) where M = 4. In this modulation, the symbols can have four different values. The modulated signals are:

$$\begin{aligned}x_m(t)&=\sqrt{\frac{2E_s}{T_s}}\cos\left[(2m-1)\frac{\pi}{4}\right]\cos(2\pi f_c t)\\&-\sqrt{\frac{2E_s}{T_s}}\sin\left[(2m-1)\frac{\pi}{4}\right]\sin(2\pi f_c t) \qquad 0\le t\le T_s \qquad ,m=1,2,3,4\end{aligned} \tag{3.35}$$

Hence, the modulated signal can be shown as:

$$x_m(t)=\sqrt{E_s}\cos\left[(2m-1)\frac{\pi}{4}\right]\phi_1(t)-\sqrt{E_s}\sin\left[(2m-1)\frac{\pi}{4}\right]\phi_2(t) \tag{3.36}$$

where

$$\begin{cases}\phi_1(t)=\sqrt{\dfrac{2}{T_s}}\cos(2\pi f_c t) & 0\le t\le T_s\\ \phi_2(t)=\sqrt{\dfrac{2}{T_s}}\sin(2\pi f_c t) & 0\le t\le T_s\end{cases}$$

The constellation diagrams for QPSK and 8PSK modulation schemes are illustrated in Figure 3.8. Figure 3.8(a) shows the QPSK constellation diagram in two-dimensional signal space. The decision regions are selected as [23]:

$$\begin{aligned}&\hat{s}="00"\ if && \vec{r}\in D_1;\\ &\hat{s}="01"\ if && \vec{r}\in D_2;\\ &\hat{s}="11"\ if && \vec{r}\in D_3;\\ &\hat{s}="10"\ if && \vec{r}\in D_4;\end{aligned}$$

The BER and symbol error rate (SER) of QPSK modulation based on *Eb/No* is presented in Figure 3.9. As can be seen from Figures 3.5 and 3.9, the BER of QPSK modulation is similar to that of 4QAM modulation.

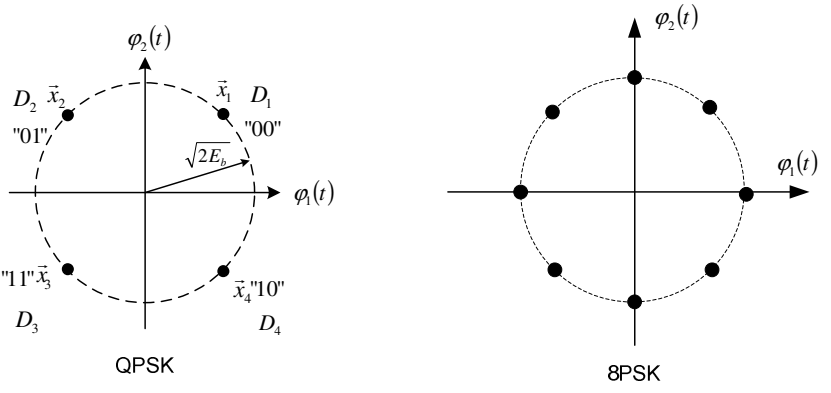


Fig. 3.8 a) QPSK modulation and decision regions [23], b) 8PSK modulation

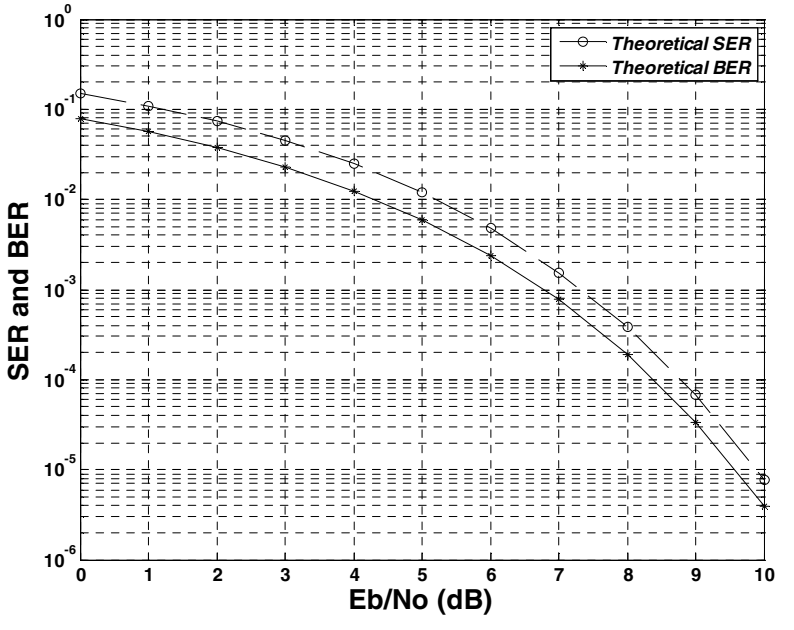


Fig. 3.9 BER and SER of QPSK modulation

3.3 Single-Carrier MIMO Systems

A flat-fading MIMO channel with N_T transmitter (Tx) antennas and N_R receiver (Rx) antennas is assumed. In a single-carrier MIMO system considering perfect synchronization, the input-output relationship is given by:

$$\mathbf{y=Hx+w} \tag{3.37}$$

where **y** is an $N_R \times 1$ vector of the received signal, **H** denotes the $N_R \times N_T$ channel matrix, **x** is an $N_T \times 1$ vector of the transmitted symbols, and **w** is an $N_R \times 1$ AWGN vector. The channel coefficients, $\left\{h_{\mu\upsilon}\right\}_{\mu=1,\upsilon=1}^{N_R,N_T}$, are statistically independent and identically distributed complex-valued Gaussian random variables with a variance of 1. The components of the noise vector have the same distribution with a variance of N_0.

3.3.1 QAM-MIMO

As discussed in Chapter 2, the objective of MIMO system design can be either improvement of the diversity order using space-time codes or increasing the data rate using layered space-time codes for spatial multiplexing. In the first design, the space-time codes are used to obtain a SER in a fading channel similar to the SER in an AWGN channel. It can be proven that, by increasing the number of antennas in the transmitter and receiver and using a suitable space-time code, the SER based on SNR curves in a fading channel get closer to this type of curve in an AWGN channel [14], [11]. Space-time block codes are often used for this application. However, the theory of the space-time trellis codes are also well developed [11], [12].Figure 3.10 illustrates a general block diagram of QAM-MIMO implementation using a space-time coder.

On the other hand, in a spatial multiplexing scenario, the objective is the increase of data. In this application, the main target is the sending of more bits per second per hertz compared to a SISO system. The SER (or BER) performance of this design is similar to the waterfall curves in a fading channel. Figure 3.11 shows a BER curve based on SNR for QAM-MIMO using four antennas in both the transmitter and receiver in a Rayleigh fading channel. The measurement results are also included to evaluate the analytical results.

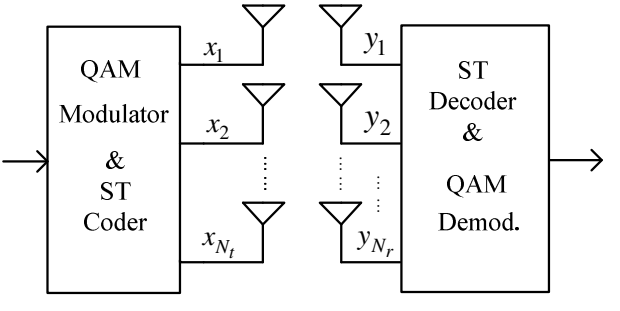


Fig. 3.10 Block diagram of a QAM-MIMO system

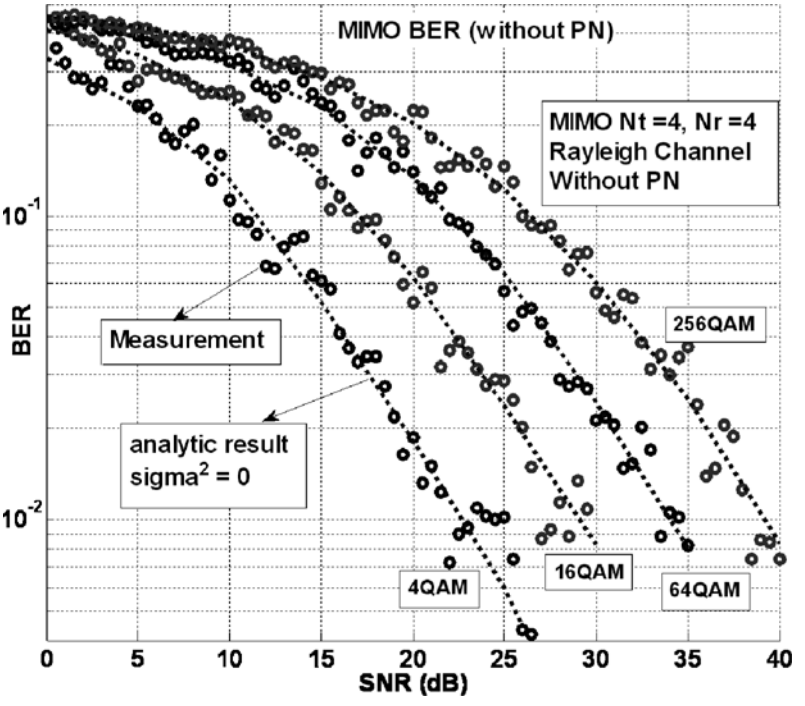


Fig. 3.11 QAM-MIMO spatial multiplexing in Rayleigh fading

3.3.2 MPSK-MIMO

MPSK-MIMO is also used in design scenarios, either for the diversity order improvement or spatial multiplexing. A general block diagram of an MPSK-MIMO system is presented in Figure 3.12. This system also uses space-time

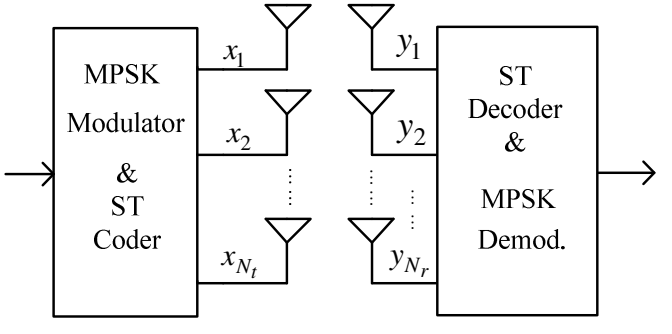


Fig. 3.12 Block diagram of an MPSK-MIMO system

codes for diversity order improvement and layered space-time codes for spatial multiplexing purposes. As discussed in Chapter 2, the simplest space-time codes are Alamouti codes [23]. The SER performance of the Alamouti scheme with 8PSK modulation is presented in Figure 3.13 [14].

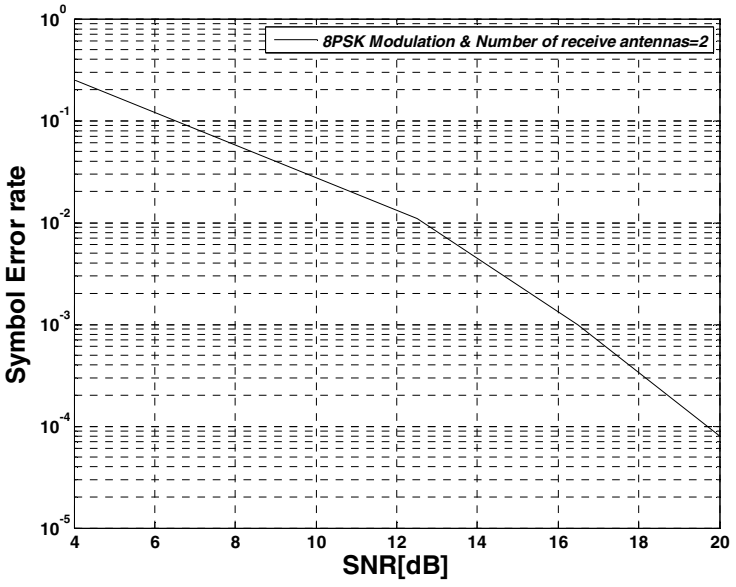


Fig. 3.13 Symbol error rate of Alamouti space-time codes with 8PSK modulation

3.4 Multicarrier SISO Modulation Systems

In a wireless multipath channel, the coherence bandwidths may be smaller than the channel bandwidth. If the channel bandwidth is smaller than the coherence bandwidth, the channel can be assumed a flat-fading channel [7]. This results in a frequency selective channel response. The main idea of a multicarrier modulation technique is the division of the channel bandwidth to a number of orthogonal subchannels, where the number of subchannels is chosen to make the subchannel bandwidth to be smaller than the coherence bandwidth. Each subchannel carries a fraction of data rate. Indeed, multicarrier modulation is conceptually designed based on frequency-division multiplexing (FDM) [8]. However, unlike a conventional FDM system, the overlapping between adjacent subcarriers is possible. A generic diagram of a multicarrier modulator and demodulator is shown in Figure 3.14. The high-speed data stream, $x(n)$, is divided into N slower streams and modulated using a prescribed signal constellation [22]. The modulated streams, $d^{(k)}(n)$, are then up-sampled by a factor. This results in signals $y^{(k)}(n)$. They are then filtered by a bank of synthesis filters, $g^{(k)}(n)$, and the filtered signals are summed. The transmitted signal, $s(n)$, has the form:

$$s(n) = \sum_{k=0}^{N-1} \sum_{l=-\infty}^{\infty} g^{(k)}(l) y^{(k)}(n-l) \tag{3.38}$$

At the receiver, a bank of filters, $\hat{f}^{(k)}(n)$, are used to separate the composite signal, $r(n)$, into N individual signals, $\hat{y}^{(k)}(n)$. These signals are down-sampled by factor N, yielding $\hat{d}^{(k)}(n)$. To remove the distortion introduced by the channel, equalizers, $\hat{w}^{(k)}(n)$, k = 0,1,2, …, N-1, are employed on a per subcarrier basis.

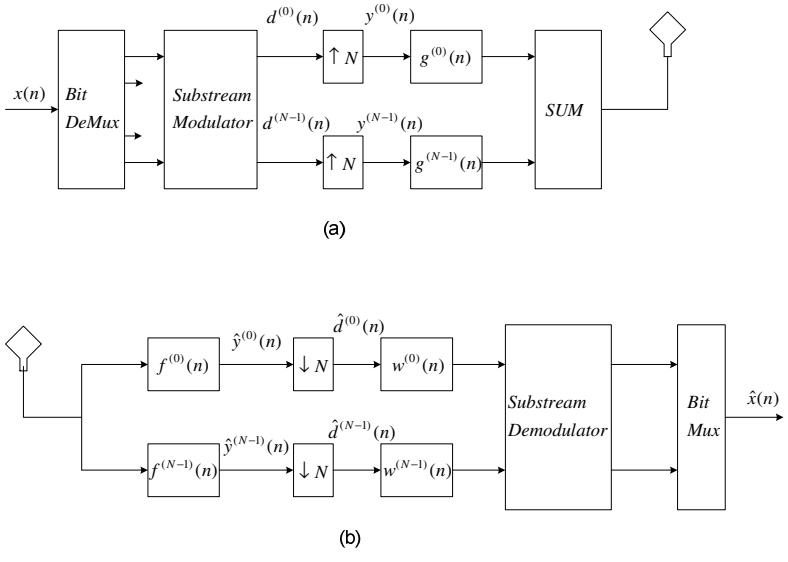


Fig. 3.14 Multicarrier modulation systems: a) transmitter b) receiver

3.5 OFDM Modulation Systems

Orthogonal frequency-division multiplexing (OFDM) is realized using the substream spectrum overlapping concept, where each substream is transmitted through a subchannel; and, the substreams are still orthogonal, and they can be separated in a demodulator. Moreover, it is implemented using a fast Fourier transformation (FFT) and inverse FFT (IFFT) technique and contains a cyclic prefix. OFDM is basically a multicarrier modulation technique, and its operation can be described based on the concept of the overlapping multicarrier modulation technique [8].

Assume a communication system with a baseband bandwidth, B, (passband bandwidth, $2B$) and a data rate, R. Moreover, it is assumed that the coherence bandwidth for the channel is $B_c < B$ and that the whole bandwidth is divided into N subchannels, where $B_N = B/N << Bc$. Using these assumptions, flat fading on each subchannel can be considered [8]. The input data is also divided into N substreams. The center frequency of the overlapping channels are:

$$f_n = f_c + nB_N \qquad n = 0,...,N-1 \tag{3.39}$$

where f_c is the carrier frequency.

The data rate for each substream is $R_N = R/N$. If a raised cosine pulse shape is assumed, the symbol time for each substream is obtained as:

$$T_N = \frac{1+\alpha}{B_N} \tag{3.40}$$

where α is the roll-off factor.

The modulated signal related to all subcarriers is presented as:

$$s(t) = \sum_{i=0}^{N-1} s_i g(t) \cos(2\pi f_i t + \varphi_i) \tag{3.41}$$

where $g(t)$ is the pulse-shaping filter, s_i is the complex symbol associated with ith subcarrier, and φ_i is the phase of the ith carrier. The subcarriers, $\cos(2\pi(f_c + i/T_N)t + \varphi_i)$, provide a set of orthonormal basis functions over the interval [0, T_N]. It is a strightforward exercise that shows that the minimum frequency separation required for subcarriers to remain orthogonal over the symbol interval [0, T_N] is $1/T_N$. [8].

The passband bandwidth of each subchannel is $2B_N$; thus, the passband subchannels in this system would overlap. Therefore, these overlapping subchannels are demodulated using the receiver structure shown in Figure 3.15.

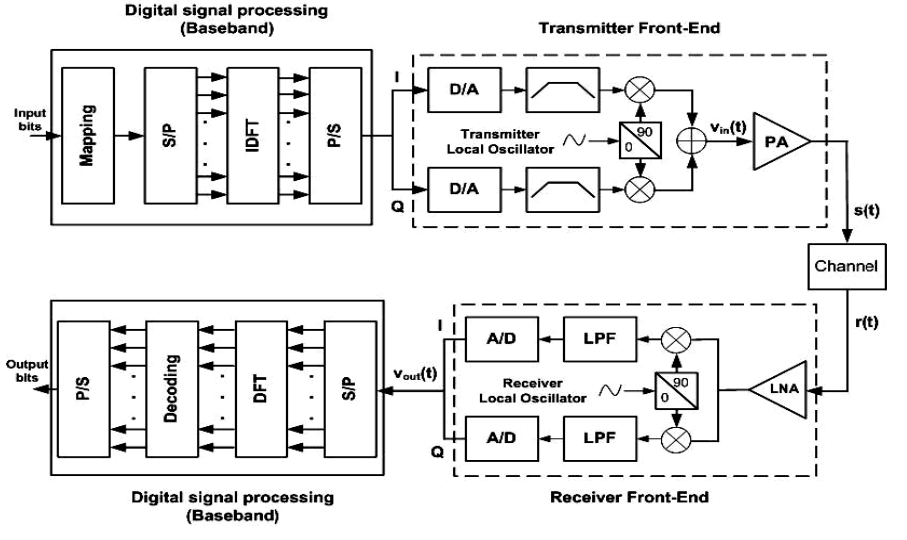


Fig. 3.15 Block diagram of OFDM communication system

3.5.1 OFDM Implementation Using IFFT/FFT

OFDM is usually implemented using FFT and IFFT. This realization of an OFDM system is illustrated in Figure 3.15 and can be described as follows. The input data is first de-multiplexed into N parallel substreams, and each one is mapped to a real or complex symbol stream using some modulation constellation, e.g., QAM or PSK [35], [26]. An IFFT is computed on each set of symbols, giving a set of complex time-domain samples. These samples are then quadrature mixed to passband in the standard way. The real and imaginary components are first converted to the analog domain using digital-to-analog converters (DACs); and, the analog signals are then used to modulate cosine and sine waves at the carrier frequency, f_c, respectively. These signals are then summed to give the transmission signal, $s(t)$.

The receiver picks up the signal $r(t)$, which is then quadrature mixed down to baseband using cosine and sine waves at the carrier frequency. This also creates signals centered on $2f_c$, so low-pass filters are used to reject these. The baseband signals are then sampled and digitized using analog-to-digital converters (ADCs), and a forward FFT is used to convert back to the frequency domain. This returns N parallel streams, each of which is converted to a binary stream using an appropriate symbol detector. These streams are then re-combined into a serial stream, which is an estimate of the original binary stream at the transmitter.

3.5.2 Cyclic Prefix in OFDM Modulation

The filters of OFDM have a non-ideal spectral transfer function. Due to the dispersive effects of the wireless channel, intersymbol interference (ISI) of the adjacent OFDM symbol results and degrades the error performance of the system. Therefore, a cyclic extension is appended to the OFDM symbol to compensate for these effects. Although a buffer of zeros would suffice in preventing the ISI, the use of cyclic extension has more advantages [22].

A representation of the use of a cyclic extension is shown in Figure 3.16. This cyclic extension in called the cyclic prefix (CP). As can be seen, the CP is created by copying the end of the OFDM symbol and placing that copy at the beginning of the symbol, for every symbol. The length of the CP is selected according to the impulse response of the wireless channel. As result of the channel dispersion, the samples of symbol M-1 spreads to samples of symbol M, while the samples of symbol M are spreading to the samples of symbol M+1. The CP must have enough length to capture the interference due to channel dispersion.

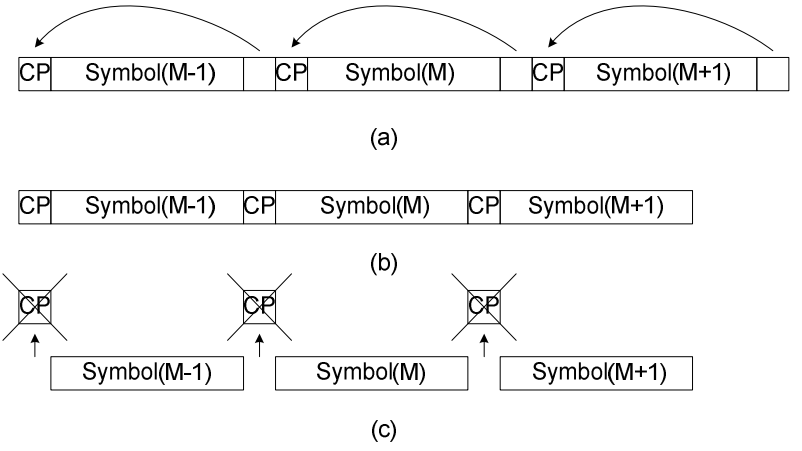


Fig. 3.16 Cyclic prefix in OFDM modulation: a) addition of CP to OFDM symbols, b) transmission through the channel, c) removal of CP at receiver

3.5.3 RF Impairment in OFDM Systems

3.5.3.1 Peak-to-Average Power Ratio in OFDM Modulation

The ratio between the peak and average powers of a signal is called the peak-to-average power ratio (PAPR). Here, the PAPR is extracted for multicarrier modulation using N carriers [24]. Hence, this calculation may also be used to obtain PAPR for OFDM modulation.

A multicarrier signal may be considered as the summation of N independent random voltages (as random variables), each with an average voltage, V_{rms}, and average power, V_{rms}^2. The total normalized average power of a multicarrier signal is given by:

$$P_{ave} = \sum_{N} V_{rms}^2 \tag{3.42}$$

Assuming unmodulated sinusoidal carriers with same amplitude, this equation may be written as:

$$P_{ave} = \sum_{N} \left(\frac{V_{max}}{\sqrt{2}} \right)^2 = \frac{N}{2} V_{max}^2 \tag{3.43}$$

On the other hand, the peak power is the power of the sum of the in-phase voltage of all the carriers. It can be represents as:

$$P_{peak} = (V_{1max} + V_{2max+\ldots+} V_{Nmax})^2 = N^2 V_{max}^2 \tag{3.44}$$

The ratio of the signal peak power to its average power is obtained as:

$$\frac{P_{peak}}{P_{ave}} = 2N \tag{3.45}$$

The average power of the envelope of the signal is more common in multicarrier modulation. A multicarrier signal with N unmodulated carriers with uniform frequency spacing, $\Delta\omega$, each having unity amplitude; and, the zero phase offset can be written as:

$$V = \sum_{n=0}^{N-1} V_{\max} Cos(\omega t + n\Delta\omega t) \tag{3.46}$$

This equation can be also wriiten as [24]:

$$V = V_{\max} Cos(\omega t + \frac{(N-1)\Delta\omega t}{2}] \frac{\sin(\frac{N\Delta\omega t}{2})}{\sin(\frac{\Delta\omega t}{2})} \tag{3.47}$$

Therefore, the envelope of a multicarrier modulation is obtained as:

$$V_{env} = \left| \frac{Sin(\frac{N\Delta\omega t}{2})}{Sin(\frac{\Delta\omega t}{2})} \right| V_{max} \tag{3.48}$$

The average power of the envelope is obtained by integration over the envelope period:

$$P_{env-ave} = \frac{1}{T} \int_0^T \left[\left| \frac{Sin(\frac{N.\Delta\omega.t}{2})}{Sin(\frac{\Delta\omega.t}{2})} \right| V_{max} \right]^2 dt = NV_{max}^2 \tag{3.49}$$

Hence, the envelope peak-to-average ratio is obtained as:

$$PAPR = \frac{P_{peak}}{P_{env-ave}} = N \tag{3.50}$$

The signal peak-to-average ratio (dB) and envelope peak-to-average ratio (dB) as functions of the number of unmodulated carriers are illustrated in Figure 3.17.

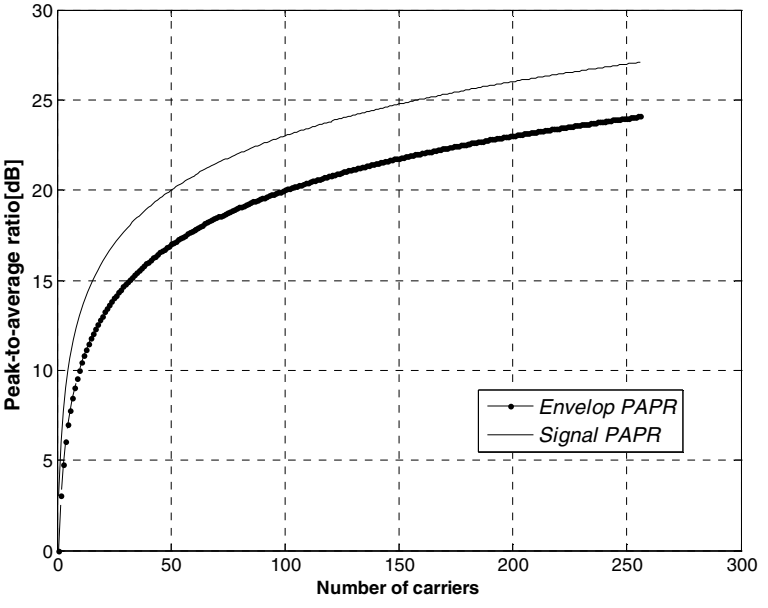


Fig. 3.17 Signal and envelope peak-to-avearge ratios as functions of the number of unmodulated carriers.

3.5.3.2 Phase Noise in OFDM Modulation

An oscillator is a system that generates a periodic signal with a specified or controllable frequency. There is phase noise in the output of an RF oscillator. The output voltage of an oscillator can be represented by [28]:

$$V_{out}(t) = (A_0 + a_N(t)).Cos(2\pi f_0 t + \varphi_N(t)) \tag{3.51}$$

where A_0 is the average amplitude of the output signal, f_0 is nominal frequency of oscillation, and $a_N(t)$ and $\varphi_N(t)$ represent the time-varying components of the amplitude and phase, respectively. These components are considered as noise, because an ideal oscillator would have a constant amplitude, A_0, and a phase that varied at a constant rate, $2\pi f_0$.

If the amplitude noise can be neglected, the main noise contribution in the OFDM signal is due to phase noise. Therefore, the SNR performances in an OFDM system with a noisy oscillator that can be calculated by integrating its phase noise power spectral density (PSD) function over the channel bandwidth. The phase noise PDF of an oscillator is approximated by a Lorenzian function with uniform phase distribution [29], [30]. It is parameterized by its total

integrated phase noise, K, and a -3dB bandwidth, B, to which we superimpose a noise floor, L_0 (set at 35dB above the thermal noise) [27]:

$$L(f)^2 = \frac{1}{\pi}.\frac{K.B}{f^2 + B^2} + L_0 \tag{3.52}$$

Figure 3.18 shows phase noise PSDs for different K and a -3dB bandwidth of 100 Hz. Therefore, the phase noise variance is:

$$\sigma_\varphi^2 = 2\int_0^{B_s/2} L(f)^2\, df = 2\frac{1}{\pi}\int_0^{B_s/2} (\frac{K\,B}{f^2 + B^2} + L_0)\, df = \frac{2K}{\pi} arctg(\frac{B_s}{2B}) + \frac{1}{\pi} L_0 B_s \tag{3.53}$$

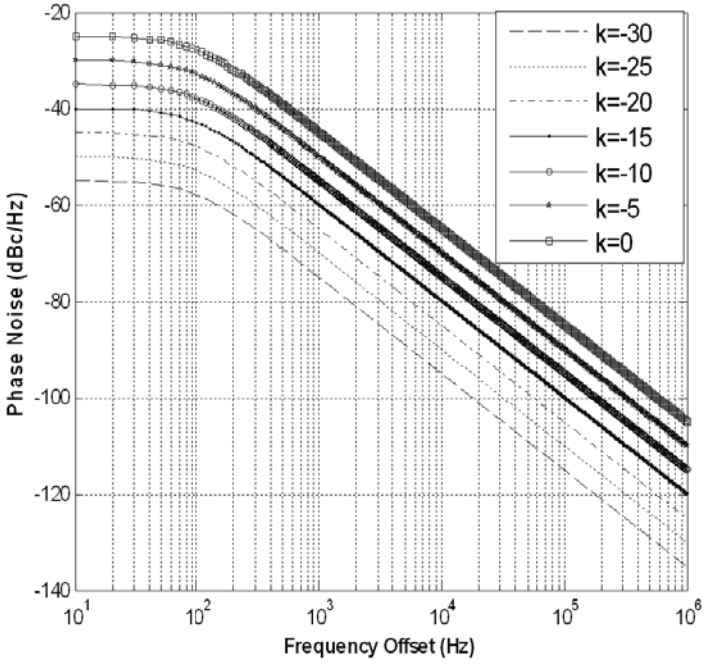


Fig. 3.18 Phase noise PDF for different K, B = 100Hz

Noise includes components at many frequencies; therefore, its phase, with respect to the main carrier, is random. Figure 3.19 shows noise added to the carrier phasor. Random amplitude and phase variation due to noise changes the constellation of an MQAM modulated signal and degrades the performance of the systems [26].

Phase noise degrades the performance of OFDM signals. Indeed, the OFDM signal is impaired by non-ideal oscillators in both the transmitter and the receiver. The first impact of this impairment is the common phase error. Moreover, phase noise also results in inter-subchannel interference (ICI). This interference develops from the mutual effects of adjacent channels. This leads to loss of the orthogonality between subchannels. These effects are studied in detail in chapter 8.

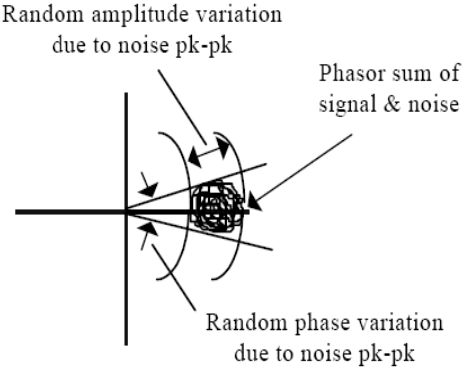


Fig. 3.19 Noise added to the carrier phasor

3.6 MIMO OFDM Systems

OFDM is effectively used to deal with transmission over frequency-selective fading channels and to avoid relatively complex signal processing in the receiver. Accordingly, the idea of multicarrier modulation, or OFDM, can be applied to MIMO frequency-selective channels as shown in Figure 3.20.

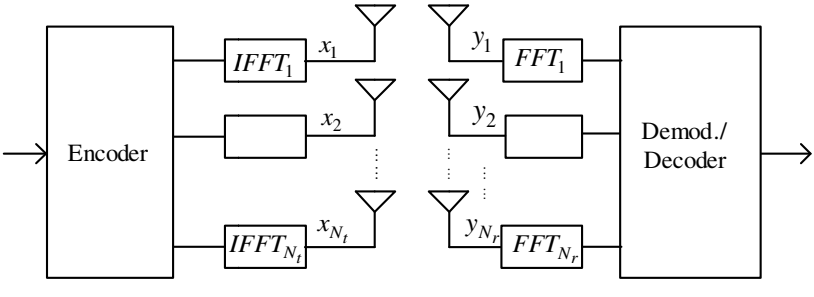


Fig. 3.20 Block diagram of a MIMO-OFDM communication system

For each element of the transmitting antenna element, the IFFT of M-ary symbols must be obtained. This is followed by adding a cyclic prefix to each of these sequences, and the resulting signal is transmitted using a suitable RF front-

end. For each of the received signals, the FFT of the aggregate received signal (superposition of all the OFDM words from each of the transmitted antennas) is calculated; and, the cyclic prefix is removed. The resulting set of signals is then used for demodulation/decoding [14].

As mentioned, the main advantage of OFDM system is the avoidance of expensive equalization algorithms and achievement of signal transmission at a reasonable complexity. However, there are some challenges with the use of MIMO-OFDM. It should be emphasized that MIMO-OFDM symbols are typically long. This leads to some channel variation over one OFDM word. Hence, this may destroy the orthogonality of the subcarriers at the receiver and may cause inter-channel interference, degrading the system performance [14]. Moreover, MIMO-OFDM suffers from all RF impairments as discussed in the previous section.

References

[1] Shannon, C.E.: A Mathematical Theory of Communication. Bell System Technical Journal 27, 379–423, 623–656 (1948)
[2] Proakis, J.G., Salehi, M.: Digital Communications, 5th edn. McGraw-Hill, New York (2008)
[3] Wozencraft, J.M., Jacobs, I.M.: Principles of Communication Engineering (1965)
[4] Sklar, B.: Digital Communications: Fundamentals and Applications. Prentice-Hall (2001)
[5] Wolfowitz, J.: Coding Theorem of Information Theory, 2nd edn. Springer, Heidelberg (1964)
[6] Nguyen, H.N., Shwedyk, E.: A First Couse in Digital Communications. Cambridge University Press (2009)
[7] Rappaport, T.: Wireless Communication, Principles and Practice, 2nd edn. Prentice-Hall (2002)
[8] Goldsmith, A.: Wireless Communications. Cambridge University Press (2005)
[9] Pahlevan, K., Leveque, A.H.: Wireless Information Networks, 2nd edn. (2005)
[10] Stuber, G.L.: Principle of Mobile Communications. Kluwer Academic Press (2001)
[11] Paulraj, A., Nabar, R., Gore, D.: Introduction to Space-Time Wireless Communications. Cambridge University Press (2003)
[12] Biglieri, E., Constantinides, A., Goldsmith, A., Paulraj, A., Vincentpoor, H.: MIMO Wireless Communications. Cambridge University Press (2007)
[13] Tsoulos, G.: MIMO System Technology for Wireless Communications. CRC Press (2006)
[14] Duman, T.M., Ghrayeb, A.: Coding for MIMO Communication Systems. Wiley (2008)
[15] Sklar, B.: Defining, designing, and evaluating digital communication systems. IEEE Communications Magazine 31(11), 91–101 (1993)
[16] Mohammadi, A., Kumar, S.: Characterization of Adaptive Modulators in Fixed Wireless ATM Networks. KICS/IEEE Journal of Communications and Networks (JCN) 6(2), 123–132 (2004)
[17] Foschini, G.J., Gans, M.J.: On limits of wireless communications in a fading environment when using multiple antennas. Wireless Personal Communications 6(3), 311–335 (1998)

[18] Simon, M.K., Alouini, M.S.: Digital Communication over Fading Channels, 2nd edn. John Wiley & Sons (2005)
[19] Lu, J., Letaief, K.B., Chuang, J.C.-I., Liou, M.L.: M-PSK and M-QAM BER Computation Using Signal-Space Concepts. IEEE Transactions on Communications 47, 181–184 (1999)
[20] Keshavarzi, M.R., Mohammadi, A., Abdipour, A., Ghannouchi, F.M.: Characterization of DC Offset on Adaptive MIMO Direct Conversion Transceivers. IEICE Transactions on Communications (1) (2011)
[21] Mark, J.W., Zhuang, W.: Wireless Communications and Networking. Prentice-Hall (2003)
[22] Wyglinski, A.M., Nekovee, A., Hous, Y.T.: Cognitive Radio Communications and Networks, Principle and Practice. Academic Press (2010)
[23] Alamouti, S.M.: A simple transmit diversity technique for wireless communications. IEEE Journal of Selected Areas of Communications 16(10), 1451–1458 (1998)
[24] Pothecary, N.: Feedforward Linear Power Amplifiers. Artech House (1999)
[25] http://en.wikipedia.org/wiki/Orthogonal_frequency-division_multiplexing
[26] Madani, M.H., Abdipour, A., Mohammadi, A.: Analysis of performance degradation due to nonlinearity and phase noise in orthogonal frequency division multiplexing system. IET Communications 4(10), 1226–1237 (2010)
[27] Madani, M.H., Abdipour, A., Mohammadi, A.: Analytical Performance Evaluation of the OFDM Systems in the Presence of Jointly Fifth Order Nonlinearity and Phase Noise. Journal of Analog Integrated Circuits and Signal Processing (2010)
[28] Keshavarzi, M.R., Mohammadi, A., Abdipour, A.: Characterization of Adaptive Modulation MIMO Systems in the Presence of Phase Noise. In: Wireless Communications & Mobile Computing Conference, Leipzig, Germany (2009)
[29] Come, B., Ness, R., Donnay, S., Van der Perre, L.: Impact of front-end non-idealities on bit error rate performance of WLAN-OFDM transceivers. In: IEEE Radio and Wireless Conf., Rawcon, pp. 91–94 (September 2000)
[30] Schiek, B., Rolfes, I., Siweris, H.J.: Noise in highfrequency circuits and oscillators. John Wiley & Sons, NY (2006)

Chapter 4
Fundamental Concepts in Wireless Transceiver Design

The fundamentals related to transceiver design for wireless communications are presented in this chapter. The main parameters in the transceiver design are discussed in order to provide a performance metric in the evaluation of wireless communication systems. The modulation accuracy is evaluated using error vector magnitude (EVM) to measure the modulation and demodulation accuracy. Some practical considerations are also discussed.

4.1 Linear Systems and Transformations

A linear time-invariant (LTI) system is characterized by its impulse response. The output of the LTI system can be expressed in terms of input and impulse response as a convolution relation:

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(t-\tau)h(\tau)d\tau \tag{4.1}$$

where $h(t)$ is the impulse response of an LTI system.

The Fourier transform of the impulse response is called the frequency response, which can be shown as:

$$H(f) = F\{h(t)\} = \int_{-\infty}^{\infty} h(t)e^{-j2\pi f_c t}dt \tag{4.2}$$

As shown in Figure 4.1, the relation between the input and the output of an LTI system in the frequency domain is presented as:

$$Y(f) = H(f)X(f) \tag{4.3}$$

where $Y(f), X(f)$ are the spectra of the input and output signals, respectively.

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 55–76.
springerlink.com

If the system is considered linear time-variant (LTV), the output is obtained as:

$$y(t) = x(t) * h(t,\tau) = \int_{-\infty}^{\infty} x(t-\tau)h(t,\tau)d\tau \tag{4.4}$$

where $h(t,\tau)$ is the impulse response of an LTV system.

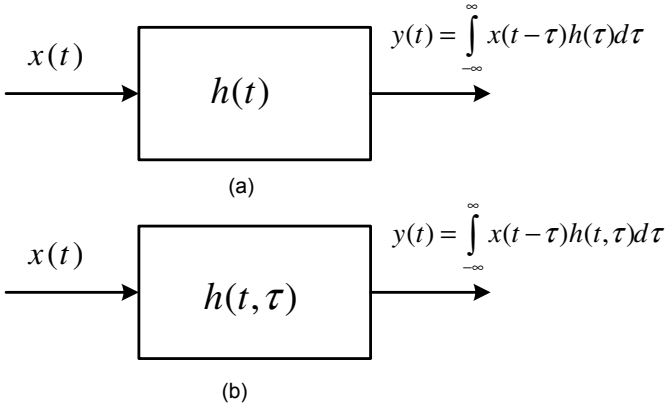


Fig. 4.1 Impulse responses of (a) an LTI system, (b) an LTV system

4.2 Nonlinear Systems

The output of a linear and memoryless system is expressed as:

$$y(t) = ax(t) \tag{4.5}$$

where a is the linear system gain, $x(t)$ and $y(t)$ are the input and output, respectively.

For memoryless nonlinear systems, the input-output relation is:

$$y(t) = a_0 + a_1 x(t) + a_2 x^2(t) + a_3 x^3(t) + ... \tag{4.6}$$

where parameters a_i are time dependent for time-varying systems.

If a sinusoidal signal, $(x(t) = A\cos(\omega_o t))$, is applied to a nonlinear system, the output is expressed as:

$$y(t) = a_0 + \frac{a_2 A^2}{2} + (a_1 A + \frac{3}{4} a_3 A^3)\cos(\omega_o t) + \frac{a_2 A^2}{2}\cos(2\omega_o t) + ... \tag{4.7}$$

4.2.1 One-dB Compression Point

If higher order terms are neglected rather than the first three terms, we may define the gain of a nonlinear system as [1]:

$$g \triangleq a_1 + \frac{3}{4} a_3 A^2 \tag{4.8}$$

The gain is reduced as a function of A, if a_3 is negative. Figure 4.2 shows the output of a typical nonlinear system.

The 1-dB compression point quantifies this gain reduction factor. This point is either defined as the input signal level at which g is reduced by 1-dB compared to the linear gain (input 1-dB compression point, P_{1-1dB}) or the point at which the linear output power and the output power of the nonlinear amplifier differ by 1 dB (output 1dB compression point, P_{O-1dB}).

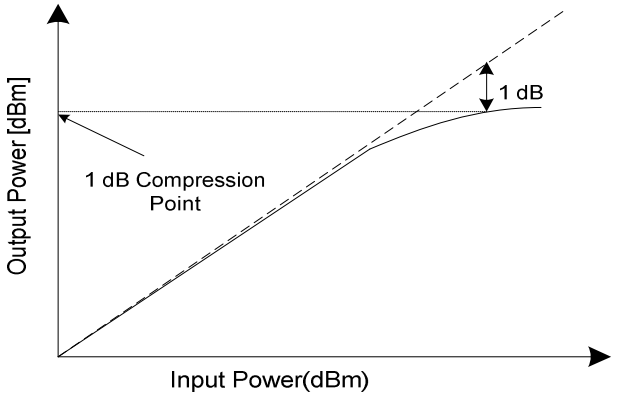


Fig. 4.2 One-dB compression point of a power amplifier

According to Figure 4.2, this point is obtained as:

$$A_{1dB} = .38\sqrt{\left|\frac{a_1}{a_3}\right|} \tag{4.9}$$

If one defines the root mean square (RMS) of the input power, with respect to a source resistor, R_S, as [2]:

$$P_{I-1dB} = 10\log(\frac{1}{2}\frac{A^2}{R_s}) \quad dBm \tag{4.10}$$

The input power at which the 1-dB input compression point occurs for a single-tone input is then given as:

$$P_{I-1dB} = 10\log(\frac{.0725}{R_s}\left|\frac{a_1}{a_3}\right|) \quad dBm \tag{4.11}$$

4.2.2 Third-Order Intercept Point

When signals of different frequencies are applied to the input of a nonlinear system, the output contains the combination terms, where are called intermodulation. If we consider the input signal as:

$$x(t) = A\cos(\omega_1 t) + B\cos(\omega_2 t) \tag{4.12}$$

the desired and intermodulation terms are generated at the output of the nonlinear system (4.6) as:

$$\begin{aligned} y(t) &= a_2[\frac{A^2}{2}+\frac{B^2}{2}]+(a_1+\frac{3}{4}a_3A^2+\frac{3}{2}a_3B^2)A\cos(\omega_1 t) \\ &+(a_1+\frac{3}{4}a_3B^2+\frac{3}{2}a_3A^2)B\cos(\omega_2 t)+a_2[\frac{A^2}{2}]\cos(2\omega_1 t)+a_2[\frac{B^2}{2}]\cos(2\omega_2 t) \\ &+a_2AB\cos((\omega_1\pm\omega_2)t)+\frac{3}{4}a_3A^2B\cos((2\omega_1\pm\omega_2)t)+ \\ &\frac{3}{4}a_3B^2A\cos((\omega_1\pm 2\omega_2)t)+\frac{a_3A^3}{4}\cos(3\omega_1 t)+\frac{a_3B^3}{4}\cos(3\omega_2 t) \end{aligned} \tag{4.13}$$

The term $\omega_1 \pm \omega_2$ is second-order intermodulation, and the terms $2\omega_1 \pm \omega_2$ and $\omega_1 \pm 2\omega_2$ are third-order intermodulation terms. If we assume, $A=B$ in (4.11), it can be seen that the output power of the third-order products increases with the cube of the input signal, whereas the desired output is proportional to the input signal. A hypothetical intersection point when the first-order power product (P_0) and the third-order power product (P_{0IM3}) are equal is called the third order intercept point. This is shown in Figure 4.3.

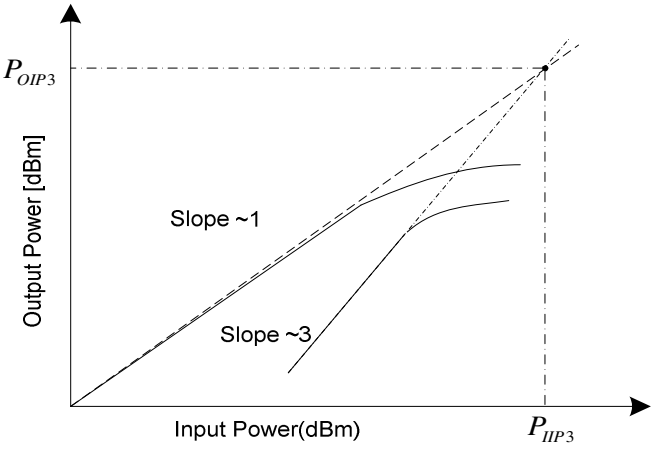


Fig. 4.3 Input and output third-order intercept point

The input amplitude of the third-order intercept point is extracted as:

$$A_{IIP3} = 2\sqrt{\left|\frac{a_1}{3a_3}\right|} \tag{4.14}$$

Using A_{IIP3} in relation to A_{1-dB}, a relation between these parameters can be extracted as:

$$\frac{A_{IIP3}}{A_{1-dB}} = \frac{2\sqrt{\left|\frac{a_1}{3a_3}\right|}}{.38\sqrt{\left|\frac{a_1}{a_3}\right|}} = 3.04 \rightarrow 9.44dB \tag{4.15}$$

On the other hand, if two tones are applied to obtain the 1-dB compression point, it can be shown that the A_{IIP3} and A_{1-dB} relation is expressed as [2]:

$$\frac{A_{IIP3}}{A_{1-dB}} = \frac{2\sqrt{\left|\frac{a_1}{3a_3}\right|}}{.22\sqrt{\left|\frac{a_1}{a_3}\right|}} = 5.25 \rightarrow 14.4dB \tag{4.16}$$

The equivalent of the input third-order intercept point (*IIP3)* of an n-stage cascaded network is obtained as:

$$\frac{1}{A_{IIP3}^2} = \frac{1}{A_{IIP3,1}^2} + \frac{a_1^2}{A_{IIP3,2}^2} + \frac{a_1^2 b_1^2}{A_{IIP3,3}^2} + ... \tag{4.17}$$

where $A_{IIP3,1}$,… are the third-order input intercept amplitude and a_1, and b_1 are linear gain coefficients of the corresponding block.

4.2.3 Second-Order Intercept Point

Second-order intermodulation distortion is important in differential circuits [1], [3]. This parameter is obtained from a hypothetical intersection point of the first-order product, a_1A, and the second-order product, a_2A^2 , of a nonlinear system. The amplitude of the second-order distortion is expressed as:

$$A_{IIP2} = \frac{a_1}{a_2} \tag{4.18}$$

The cascaded input second-order intercept point *(IIP2)* can be written as:

$$\frac{1}{A_{IIP2}^2} = \frac{1}{A_{IIP2,1}^2} + \frac{a_1}{A_{IIP2,2}^2} + \frac{a_1 b_1}{A_{IIP2,3}^2} + ... \tag{4.19}$$

4.2.4 Cross Modulation

Suppose a single-tone sinusoidal signal is added to an amplitude-modulated signal with modulation index *m* as:

$$V(t) = A\cos(\omega_1 t) + B(1 + m(t))\cos(\omega_2 t) \tag{4.20}$$

If $V(t)$ is applied to the input of a nonlinear system, the output contains the combination terms. One of the third-order terms can be presented as:

$$\frac{3}{2} AB^2 (1 + 2m(t) + m^2(t))\cos(\omega_1 t) \tag{4.21}$$

The distorted version of the modulation signal of the carrier of ω_2 has been transferred to the carrier of ω_1. This is called a cross modulation effect and is due to a nonlinear circuit.

4.2.5 AM/AM and AM/PM Distortion

The relation between the input and the output of a linear system in time domain is represented as [4]:

$$y(t) = ax(t-\tau) \tag{4.22}$$

where a is a constant and τ is a time delay.

The output amplitude has a linear relation with the input amplitude, and the delay is constant to all inputs. On the other hand, for memoryless nonlinear systems, the input-output relation is:

$$y(t) = a_0 + a_1 x(t) + a_2 x^2(t) + a_3 x^3(t) + \ldots \tag{4.23}$$

where parameters a_i are time dependent for time-varying systems.

For a third-order nonlinearity, the output signal for an input signal, $(x(t) = A\cos(\omega_o t))$, is:

$$y(t) = (a_1 A + \frac{3}{4} a_3 A^3)\cos(\omega_o t) \tag{4.24}$$

So far, it has been assumed that the coefficients of the memoryless system are real. In reality, most systems exhibit a complex behavior with complex coefficients, where the nonlinear circuits create nonlinear distortions in both amplitude (AM/AM conversion) and phase (AM/PM conversion). It is straightforward to show that a nonlinear circuit causes a deviation in linear characteristics, both in amplitude and phase. Deviations in amplitude create AM/AM distortion, and deviation in phase result in AM/PM distortion. Figure 4.4 shows typical AM/AM and AM/PM conversions.

■ ***Example 4.1:*** *Nonlinear Model for PA*

Memoryless power amplifier nonlinearity are commonly modeled using amplitude-dependent AM/AM and AM/PM distortion models. In a memoryless amplifier, the results of the above conversion only depend on the current value of a signal. Saleh's model is a simple model that characterizes the nonlinearity of a power amplifier [5]. The model is experimentally verified for traveling wave tube amplifiers (TWTA). In this model, the input and output of an amplifier are:

$$x(t) = \rho e^{j\varphi}$$
$$y(t) = F(\rho)e^{j(\theta(\rho)+\varphi)}$$

where ρ is the input amplitude, φ is the input phase, $F(\rho)$ is the output amplitude, and $\theta(\rho)+\varphi$ is the output phase. According to Saleh's model, $F(\rho)$ and $\theta(\rho)$ are obtained as:

$$F(\rho) = \frac{\alpha_1 \rho}{1+\alpha_2 \rho^2} \qquad \theta(\rho) = \frac{\beta_1 \rho^2}{1+\beta_2 \rho^2}$$

where $\alpha_1 = 2.1587$, $\alpha_2 = 1.1517$, $\beta_1 = 4.0330$, and $\beta_2 = 9.1040$.

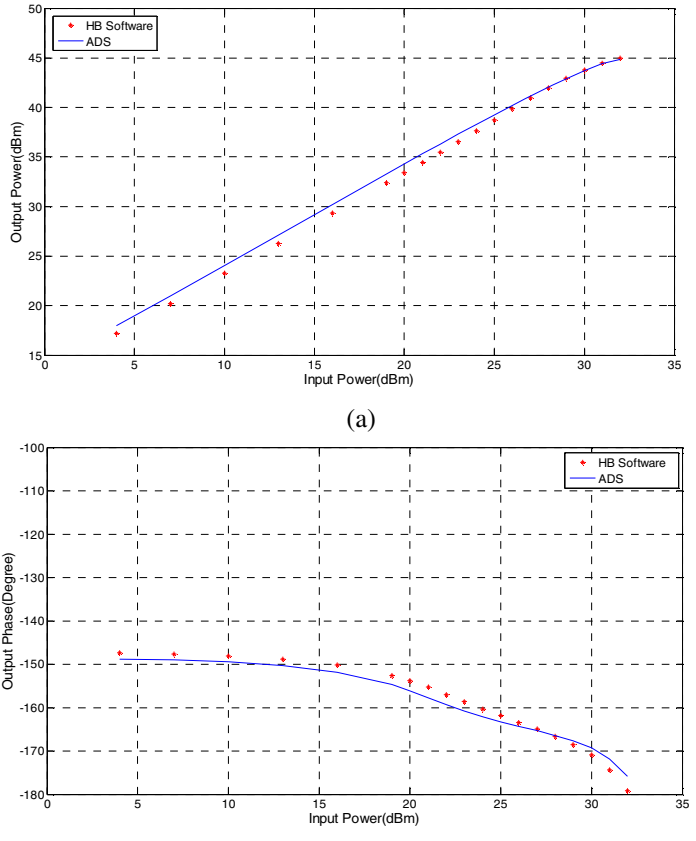


Fig. 4.4 (a) AM/AM conversion and (b) AM/PM conversion of a typical Doherty power amplifier obtained by ADS software from Agilent and the in-house harmonic balance method

4.2.6 Spectral Regrowth

If a power amplifier exhibits nonlinearity, the shape of the input signal at the output of the transmitter is not preserved. In this case, the spectrum of the output signal is not limited to the desired bandwidth. This effect is called spectral regrowth, which can be quantified by the adjacent channel power ratio (ACPR).

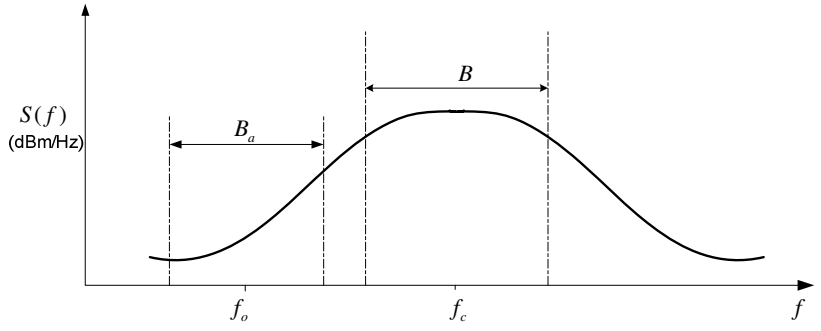


Fig. 4.5 Input and output third-order intercept point

In wireless communications, each user is assigned a specific channel width in the system bandwidth. The interference power from the adjacent channels in the main channel causes adjacent interference. Interference produced in the adjacent channel is characterized by ACPR, which is the power in the main channel divided by the power in the lower or upper adjacent channels.

Suppose $S(f)$ denotes the power spectral density of the transmitted signal, B is the bandwidth of the main signal, B_a is the bandwidth of the adjacent channel, f_c is the center frequency of the main signal, and f_o is the offset frequency of the adjacent signal. According to Figure 4.5, the ACPR is obtained as:

$$ACPR = 10\log\left\{\frac{\int_{f_c-f_o-B_a/2}^{f_c-f_o+B_a/2} S(f)df}{\int_{f_c-B/2}^{f_c+B/2} S(f)df}\right\} \tag{4.25}$$

■ ***Example 4.2:*** *ACPR of TETRA*

The channel width of the Terrestrial Trunked Radio (TETRA) standard is 25 kHz. The specifications of the ACPR at the adjacent channel is -60 dBc, measured at 25 kHz off from the transmitter center frequency, -70 dBc at 50 kHz, and -74 dBc at 1 MHz. The power is integrated using a bandpass root raised cosine filter with $\alpha = 0.35$ and a bandwidth of 18 kHz [6].

4.3 Noise

All signals are contaminated by noise, which degrades the system performance. The noisiness of a signal is usually specified in terms of the signal-to-noise ratio, $SNR = S/N$, where S is the signal power and N is the noise power. This parameter is generally a function of frequency. The reduction in *SNR* throughout a two-port network is characterized by the noise factor.

$$F = \frac{SNR_I}{SNR_o} \tag{4.26}$$

where SNR_I and SNR_O are the input and output signal-to-noise ratios, respectively.

Figure 4.6 shows a noisy two-port network. An expression for noise factor of a two-port network, which is related to the reflection coefficients of a two-port network, is [1]:

$$F = F_{\min} + 4\frac{R_N}{Z_o}\frac{\left|\Gamma_s - \Gamma_{opt}\right|^2}{(1-\left|\Gamma_s\right|^2)\left|1+\Gamma_{opt}\right|^2} \tag{4.27}$$

where $F_{\min}$ is the minimum noise factor, R_N is the equivalent noise resistance, and Γ_{opt} is the optimal reflection coefficient corresponding to the optimal source admittance that provides the minimum noise factor. These parameters are the device characteristics. In addition, Γ_s is the source reflection coefficient.

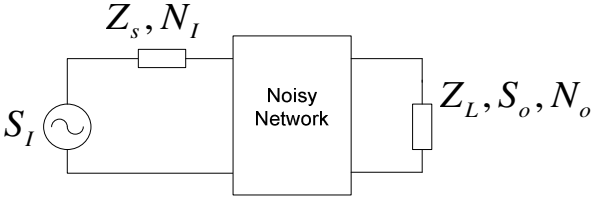


Fig. 4.6 A noisy two-port network

4.3.1 Noise Factor of Cascaded Networks

The block diagram of a cascaded connection of the system stages is shown in Figure 4.7. In this diagram, g_i and F_i are the power gain and noise factor, respectively, of the corresponding stages. The equivalent noise factor for the cascaded connection of the stages is given by [5]:

$$F_T = F_1 + \frac{F_2 - 1}{g_1} + \frac{F_3 - 1}{g_1 g_2} + \ldots = F_1 + \sum_{l=2}^{L}\left(\frac{F_l - 1}{\prod_{n=1}^{l-1} g_n}\right) \tag{4.28}$$

where the noise factor is often expressed in decibels. In this case, it is called the noise figure (NF) and is defined as $NF = 10\log(F)$. Since the minimum value for the noise factor is equal to one, the noise figure of an ideal noiseless network is 0 dB.

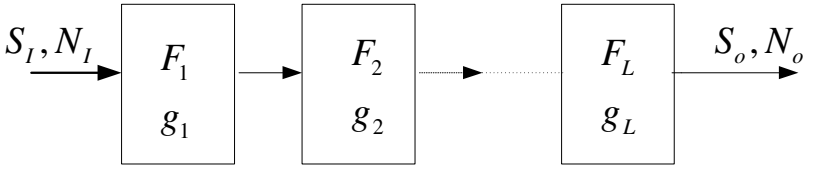


Fig. 4.7 A cascaded connection of the stages with specific power gains and noise factors

On the other hand, the noise temperature is also a measure of the noise of a system. The parameters of a two-port network are the power, gain g, bandwidth B, and noise temperature, T_E. The noise temperature of the source is considered to be T_0. By assuming power matching at the input of the network, the input noise power at temperature T_0 may be shown as $N_I = kT_0B$, where k is Boltzmann's constant. At a room temperature, $T_0 = 298K, kT_0 = 4\times10^{-21}$ W/Hz or $SNR_{d_{\min}}$. Therefore, the output noise power is obtained as:

$$N_o = kB(T_0 + T_E)g \tag{4.29}$$

Accordingly, the relationship between the noise factor, *F,* and the equivalent noise temperature can be obtained as:

$$T_E = (F-1)T_0 \tag{4.30}$$

Similarly, the equivalent noise temperature of T_E of an n-stage cascaded system is obtained as:

$$T_E = T_{E1} + \frac{T_{E2}}{g_1} + \frac{T_{E3}}{g_1 g_2} + \ldots \tag{4.31}$$

■ ***Example 4.3:*** *Cascaded Noise Figure of an LTE System*

The components and their gain and noise figures of an LTE direct conversion receiver is shown in Table 4.1. By using (4.28), the cascaded system noise figure was obtained to be 9.48 dB [2].

Table 4.1 The Parameters of a Typical LTE Receiver

Parameter	Band-switch	Duplexer	LNA	Mixer	Roofing/ Filter	VGA
Gain(dB)	-1.47	-1.8	16	2.7	32.2	10
NF(dB)	1.47	1.8	2.4	6.8	22.43	19.35

4.3.2 Capacity Reduction due to Noise Factor of an RF Receiver

A typical radio frequency (RF) receiver, including an antenna, an RF front-end and a load, is illustrated in Figure 4.8 [11], [12]. The antenna is modeled by a Thevenin equivalent circuit with impedance $z_A = r_A + jx_A$. The front-end may include amplifiers, mixers, and filters. In this model, *u* is the antenna open-circuit voltage containing both signal and noise. It can be represented as:

$$u = hx + n_o \tag{4.32}$$

where *x* is the transmitted signal, *h* is the channel coefficient, and n_o is the thermal noise as $n_o \sim CN(0, 4kT_oB.r_A)$, *B* is bandwidth in Hertz, and $r_A = \mathrm{Re}[z_A]$. The output signal may be related to input signal *u* as:

$$r \propto hx + n$$

where $n = n_o + z$, and z is the noise contribution from the front-end. The noise factor of the front-end can be defined as:

$$F = \frac{E[|n|^2]}{E[|n|^2]\Big|_{z=0}} = 1 + \frac{E[|z|^2]}{4kT_oB.r_A} \tag{4.33}$$

The SNR at the output of the front-end is reduced by factor *F*. The SNR at the output is obtained as:

$$SNR_o = \frac{S}{F.N_I} \tag{4.34}$$

where $N_I = 4kT_oB.r_A$.

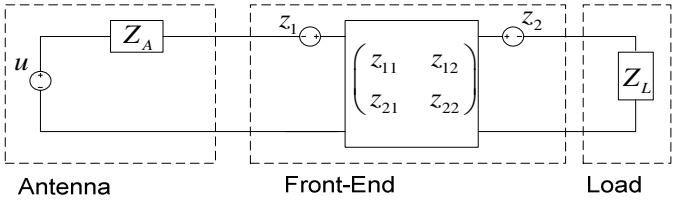


Fig. 4.8 A circuit model for a receiver front-end

This relation can also be presented as:

$$SNR_o = SNR_I - NF \tag{4.35}$$

where NF is the noise figure of the front-end system.

According to (3.1), $C = W\log_2(1+\frac{S}{N})$, which is the capacity is related to the SNR ratio. Therefore, the noisy front-end reduces the system capacity.

In designing a front-end, the general objective is usually to minimize the noise figure of the front-end.

4.4 RF Systems Design Parameters

A wireless link analysis is performed in the design process to determine some link parameter, e.g., received power, considering the limitations from signal specifications and the standardization bodies. For instance, one may obtain the range of a specific radio using the parameters of transmitter powers, antenna gains, path loss and required SNR.

4.4.1 Path Loss

The propagation path loss is a function of the frequency operation and is obtained as:

$$L = 10\log_{10}(\frac{16\pi^2 d^n}{\lambda^2}) \tag{4.36}$$

where L is the propagation path loss; d is distance between the transmitter and the receiver; n is the path loss exponent, which is usually between 2 and 6; and, λ is the free space path length.

4.4.2 Link Budget

Certain values for the different parameters are required in a radio link, in order to achieve a specific quality. The link budget calculation provides an initial level of estimation of link performance. The link budget equation is expressed as follows:

$$SNR_{d_{\min}}[dB] = P_T + G_T + G_R - L - 10\log_{10}(kT) - 10\log_{10}(B) - NF \tag{4.37}$$

where P_T is the transmit power, G_T is the transmitter antenna gain, G_R is the receiver antenna gain, L is the path loss, $k = 1.38\times10^{-23}$ and is Boltzmann's constant, T is room temperature in Kelvin, B is the receiver bandwidth, and NF is the noise figure of the receiver.

4.4.3 Receiver Sensitivity

The receiver sensitivity is defined as the minimum received signal power needed to successfully demodulate the received signal for a certain bit error rate (BER)

requirement. To be more specific, the minimum signal level at the receiver's input, S_{sense}, which can be detected with an acceptable, $SNR_{d_{\min}}$ is called the sensitivity of a system. According to the definition of noise factor:

$$SNR_{I_{\min}}[dB] = SNR_{d_{\min}}[dB] + NF \tag{4.38}$$

where $SNR_{I_{\min}}[dB] = S_{sense}[dB] - N_I[dB]$.

Assuming input power matching and having $N_I[dB] = -174 + 10\log_{10}(B)$, the sensitivity can be written as:

$$S_{sense}[dB] = -174 + 10\log_{10}(B) + NF + SNR_{d_{\min}} \tag{4.39}$$

■ ***Example 4.4****: Receiver Sensitivity for Digital Demodulation*

According to (3.4), the follwing relationship governs a digital modulation system:

$$\frac{S}{N} = \frac{E_b}{N_o}\frac{R}{B}$$

Using this relation in (4.35), the receiver sensitivity for a specific digital modulation system is obtained as:

$$S_{sense}[dB] = -174 + 10\log_{10}(R_b) + NF + \left.\frac{E_b}{N_o}\right|_{\min}$$

where $\left.\frac{E_b}{N_o}\right|_{\min}$ is the minimum value of energy of bit per power spectral density of noise that provides an acceptable BER for a digital system.

4.4.4 Dynamic Range

The difference between the maximum and minimum signals that can be processed by a receiver is called the dynamic range. The linear dynamic range is expressed as a difference between the input signal level that results in 1-dB compression and the minimum signal level that can be distinguished from the noise. This definition is more suitable for power amplifier applications. However, for low-noise amplifiers and mixers, the minimum noise ends at a boundary; and, the maximum power level for which distortion becomes unacceptable is the other boundary.

In general, the spurious free dynamic range (SFDR) definition is used, which is the range between the receiver sensitivity and the maximum power level for which the output third-order intermodulation product is equal to the output noise level.

The lower level depends on the minimum detectable signal. The minimum level is a signal power that allows for detection with a desired SNR, in order to obtain a desired BER. This is shown in Figure 4.9. This parameter can be expressed as [1]:

$$SFDR[dB] = \frac{2}{3}(IIP3[dB] + 174 - 10\log_{10}(B) - NF) - SNR_{d_{\min}} \tag{4.40}$$

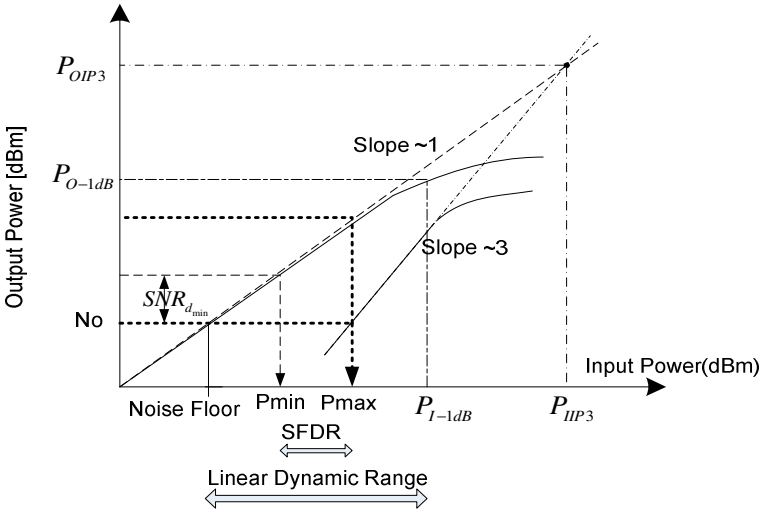


Fig. 4.9 Dynamic range analysis

4.4.5 ADC Selection

Analog-to-digital converters (ADCs) and digital-to-analog converters (DACs) are the main blocks of a wireless transceiver. The SNR of an ideal N-bit ADC is given by [7], [8]:

$$SNR = 6.02N + 1.76 \quad dB \tag{4.41}$$

For practical applications, the difference between ideal and practical SNR values may be pronounced [10]. If one considers the measured SNR values, the effective number of bits for a practical ADC is obtained as:

$$N = \frac{SNR_{measured} - 1.76}{6.02} \tag{4.42}$$

4.4.6 Phase Noise

An oscillator is a system that generates a periodic signal with a specified or controllable frequency. In a practical oscillator, the power distributed in frequency bands on both sides of the oscillator frequency, rather than all power being concentrated at the oscillation frequency. These unwanted frequency components are called oscillator noise. There are the two kinds of noise in an oscillator output, namely amplitude noise and phase noise. The output voltage of an oscillator can generally be represented by[7], [8]:

$$V_{out}(t) = (A_0 + a_N(t)).Cos(2\pi f_0 t + \phi_N(t)) \tag{4.43}$$

where A_0 is the average amplitude of the output signal; f_0 is the nominal frequency of oscillation; and, $a_N(t)$ and $\phi_N(t)$ represent the time varying components of the amplitude and phase, respectively.

Oscillator noise has negative impacts on both the transmitter and receiver sides. In the transmitter, oscillator noise causes the signal to be transmitted below or above the desired frequencies. In the receiver, the oscillator phase noise can mix with unwanted interferences and generate unwanted signals in the intermediate frequencies. This process is called reciprocal mixing [7].

Reciprocal mixing is shown in Figure 4.10. As can be seen, the desired signal, A, and the adjacent signal, A_a, are both converted to the same frequency. It is usually possible to keep the amplitude of the oscillator relatively constant. A simple phase noise model that agrees with measurements is Leeson's model [9]. Based on this model, the normalized single-sideband noise spectral density is obtained as:

$$L(f_m) = 10\log_{10}\{\frac{FkT}{P_S}[(\frac{f_o}{2Q_L f_m})^2 + 1](\frac{f_c}{f_m} + 1)\} \tag{4.44}$$

where f_m is the offset frequency in Hz, $L(f_m)$ is the noise level at f_m in dB_c/Hz, f_0 is the center frequency in Hz, Q_L is loaded *quality factor* of the oscillator, F is the noise factor, T is the absolute temperature, P_S is the carrier power in dBm, and f_c is the corner frequency for flicker noise in Hz. The $L(f_m)$ is specified at a particular offset frequency from the carrier frequency.

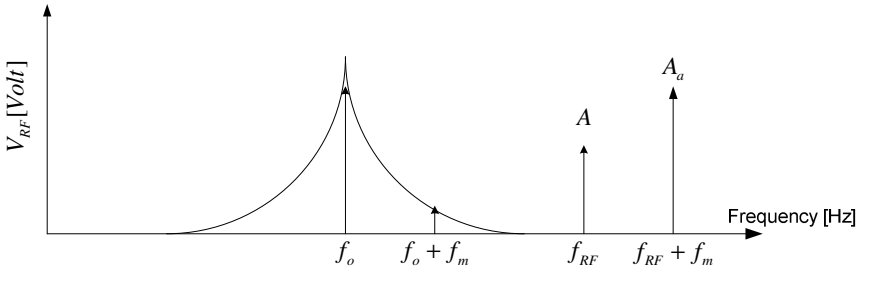


Fig. 4.10 Reciprocal mixing due to oscillator phase noise

Leeson's model tells us that phase noise (at a given offset) improves as both the carrier power and Q increase. These dependencies make sense. Increasing the signal power improves the ratio simply because the thermal noise is fixed, while increasing Q improves the ratio quadratically because the tank's impedance falls off as $1/Q_L f_m$.

4.5 Modulation Accuracy

4.5.1 Error Vector Magnitude (EVM)

The error vector magnitude (EVM) is a metric that measures the modulation and demodulation accuracy. It is defined as:

$$EVM \triangleq \sqrt{\frac{\int \left| y_I(t) - y(t) \right|^2}{\int \left| y_I(t) \right|^2}} \tag{4.45}$$

where $y(t)$ is the actual signal, and $y_I(t)$ is the ideal signal. If the difference between actual signal and ideal signal is assumed due to noise, it is obvious that the EVM can be expressed as:

$$SNR_{dB} = 10log_{10}(\frac{1}{EVM^2}) \tag{4.46}$$

It should be emphasized that EVM is an overall measure of accuracy including in-phase / quadrature (I/Q) imbalance, phase noise, carrier leakage, ripple in gain, nonlinear degradation, and frequency error [13]. The modern wireless systems have EVM specifications for specific modulation schemes. For example, the required EVM for wideband code division multiple access (WCDMA) using quadrature phase-shift keying (QPSK) is less than 17.5% [13].

4.5.2 Effect of I/Q Imbalance on EVM

I/Q imbalance can be modeled using a local oscillator (LO) signal. In this model, the LO signal is as shown in Figure 4.11. The output signal in this model can be represented as:

$$r_{RF}(t) r_{LO}(t) = \frac{1}{2} r_{RF}(t) \left[e^{j2\pi f_{LO} t}(1 + \alpha e^{j\theta}) + e^{-j2\pi f_{LO} t}(1 - \alpha e^{-j\theta}) \right] \tag{4.47}$$

where

$$r_{RF}(t) = I(t)Cos(2\pi f_{RF} t) - Q(t)Sin(2\pi f_{RF} t) \tag{4.48}$$

$$r_{LO}(t) = Cos(2\pi f_{LO} t) + j\alpha Sin(2\pi f_{LO} t + \theta) \tag{4.49}$$

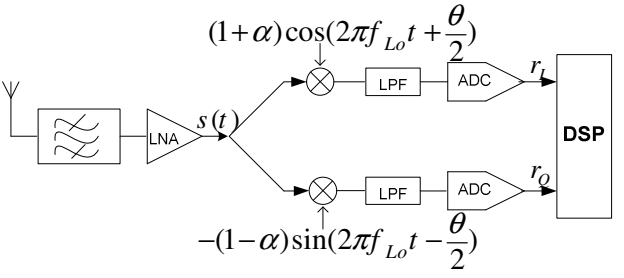


Fig. 4.11 I/Q imbalance modeling

I/Q imbalance results in image signals. Accordingly, the image suppression, *IS*, (or its inverse, image rejection, *IRR*) is related to EVM as:

$$EVM = \sqrt{10^{IS/10}} \tag{4.50}$$

where

$$\frac{1}{IS} = IRR = \frac{\left|1-\alpha e^{-j\theta}\right|}{\left|1+\alpha e^{j\theta}\right|} = \sqrt{\frac{1+\alpha^2-2\alpha Cos\theta}{1+\alpha^2+2\alpha Cos\theta}} \tag{4.51}$$

4.5.3 Effect of Phase Noise on EVM

Assuming the desired output signal is $y_I(t) = Acos(\omega t)$ and the actual output signal is $y(t) = Acos(\omega t + \varphi(t))$, it is straightforward to show that EVM due to phase noise [14], [15] is:

$$EVM_{rms} = \sqrt{\frac{1}{SNR} + 2 - 2\exp(-\frac{\sigma^2}{2})} \tag{4.52}$$

where σ is the standard deviation of phase noise.

Approximation of $\exp(-\frac{\sigma^2}{2})$ with Taylor series, (3.61) can be written as:

$$EVM_{rms} = \sqrt{\frac{1}{SNR} + \sigma^2} \tag{4.53}$$

EVM variation with phase noise is presented in Figure 4.12.

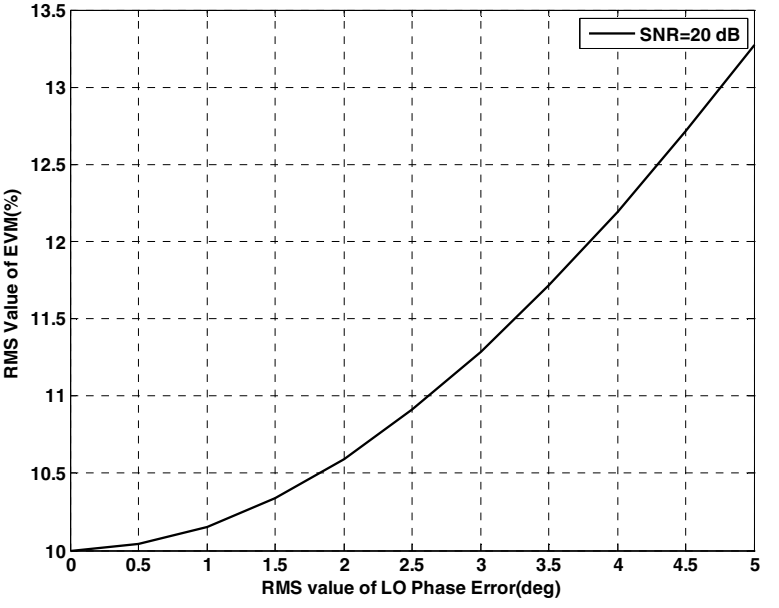


Fig. 4.12 EVM variation with LO phase noise

4.5.4 Effect of Nonlinearity on EVM

To estimate the effect of nonlinearity on EVM, a histogram of the transmitted power is usually needed [13l]. However, if the power amplifier of the transmitter operates well below its 1-dB compression point, the third-order intermodulation interference will have a major contribution to the EVM among the nonlinear effects. Accordingly, third-order intermodulation distortion (IMD3) may be used to calculate the EVM using the approximate expression of [15]:

$$EVM_{IM3} \simeq \sqrt{10^{(P_{IM3}-P_{RF})/10}} \tag{4.54}$$

where P_{RF} is the desired RF power in dBm.

4.5.5 Effect of Carrier Leakage on EVM

Carrier leakage results in DC offset in receiver and can be shown as:

$$s(t) = A_I(t)\cos(\theta(t))\cos(\omega_c t) - A_Q(t)\sin(\theta(t))\sin(\omega_c t) \tag{4.55}$$
$$+ d_I(t)\cos(\omega_c t) - d_Q(t)\sin(\omega_c t)$$

where $A_I(t)$ and $A_Q(t)$ the quadrature amplitudes of the I and Q signals, and $\theta(t)$ is the corresponding phase angle. The $d_I(t)$ and $d_Q(t)$ are the DC offsets at the output of the I and Q DACs, respectively. The carrier suppression can be defined as the average carrier leakage to the average power of signal:

$$R = 10log_{10}(\frac{P_{C-Leak}}{P_s}) \tag{4.56}$$

The degradation in terms of EVM is given by:

$$EVM = \sqrt{10^{R/10}} \tag{4.57}$$

4.5.6 Total EVM

If the different factors that contribute to the degradation of the modulation accuracy can be considered uncorrelated, the total EVM of the transmission signal can then be obtained as:

$$EVM_{Total} = \sqrt{\sum_k EVM_k^2} \tag{4.58}$$

where EVM_k is the individual EVM due to any impairment factor that degrades the signal performance.

■ ***Example 4.5:*** *EVM and RCE in WiMAX Systems*

The relative constellation error (RCE) is often used instead of EVM in WiMAX (Worldwide Interoperability of Microwave Access) systems [16]. In general, RCE and EVM are nearly interchangeable terms. However, RCE describes an EVM measurement that is calculated over an entire Fixed WiMAX frame. There are two WiMAX classifications, Fixed and Mobile, based on a subset of the IEEE 802.16 standards and are defined by the WiMAX forum. More specifically, Fixed WiMAX is based on the orthogonal frequency-division multiplexing (OFDM) physical layer of the 802.16-2004 specifications, which are sometimes called IEEE 802.16d [17].

Mobile WiMAX is based on the orthogonal frequency-division multiplexing access (OFDMA) physical layer of the 802.16e-2005 standard, which is a revision of the original Fixed WiMAX standard [18]. Mobile WiMAX provides added functionality, such as base station handoffs, multiple input multiple output (MIMO) transmit/receive diversity, and scalable fast Fourier transform (FFT) sizes. Table 4.2 shows a high-level side-by-side comparison of the Fixed and Mobile WiMAX standards.

The IEEE 802.16-2004 standard prescribes that a Fixed WiMAX transmitter must have a minimum RCE per each modulation scheme for either a subscriber unit or a base station, as shown in Table 4.3. The IEEE 802.16e-2005 standard (8.4.12.3) prescribes that a Mobile WiMAX transmitter must have a minimum RCE (for either a subscriber unit or a base station) per each modulation scheme, as shown in Table 4.4.

Table 4.2 Fixed Versus Mobile WiMAX

	Fixed WiMAX	**Mobile WiMAX**
Standard	IEEE 802.16	IEEE 802.16e-2005
Multiplexing	OFDM	OFDMA
FFT size	256	Scalable (512, 1024, …)
Duplexing mode	TDD, FDD	TDD
Modulation scheme	BPSK, QPSK, 16-QAM, and 64-QAM	QPSK, 16-QAM, and 64-QAM
Subcarrier spacing	15.625, 31.25, 45 kHz	10.94 kHz
Signal bandwidths	3.5, 7, and 10 MHz	5, 7, 8.75, and 10 MHz
Spectrum	3.5 and 5.8 GHz	2.3, 2.5, and 3.5 GHz

Table 4.3 802.16d Minimum RCE for Various Modulation Types [3], [4]

Burst Type	**SS RCE (dB)**	**BS (RCE)**
BPSK-1/2	-13	-13
QPSK-1/2	-16	-16
QPSK-3/4	-18.5	-18.5
16-QAM-1/2	-21.5	-21.5
16-QAM-3/4	-25.0	-25.0
64-QAM-2/3	-29.0	-29.0
64-QAM-3/4	-30.0	-31.0

Table 4.4 802.16e-2005 Minimum RCE for Various Modulation Types [4]

Burst Type	**SS RCE (dB)**	**BS (RCE)**
QPSK-1/2	-15.0	-15.0
QPSK-3/4	-18.0	-18.0
16-QAM-1/2	-20.5	-20.5
16-QAM-3/4	-24.0	-24.0
64-QAM-1/2	-26.0	-26.0
64-QAM-2/3	-28.0	-28.0
64-QAM-3/4	-30.0	-30.0

References

[1] Tasic, A., Serdijn, W.A., Long, J.R.: Adaptive Low Power Circuits for Wireless Communications. Springer, Heidelberg (2006)
[2] Rouphael, T.J.: RF and Digital Signal Processing for Software Defined Radio. Newnes Publisher (2009)
[3] Shaeffer, D.K., Lee, T.: The Design and Implementation of Low-Power CMOS Radio Receivers. Kluwer Academic Press (1999)
[4] Giannini, F., Leuzzi, G.: Nonlinear Microwave Circuit Design. John Wiley (2004)

[5] Saleh, A.A.M.: Frequency-independent and frequency-dependent nonlinear models of TWTA amplifiers. IEEE Transactions on Communications 29(11), 1715–1720 (1981)
[6] Luzzatto, A., Shirazi, G.: Wireless Transceiver Design. Wiley (2007)
[7] Lee, T.H.: The Design of CMOS Radio Frequency Integrated Circuits, 2nd edn. Cambridge University Press, New York (2004)
[8] Smith, J.R.: Modern Communication Circuits, 2nd edn. McGraw Hill (1998)
[9] Leeson, D.B.: A simple model of feedback oscillator noise spectrum. Proceedings of IEEE 54, 136–154 (1966)
[10] Kenington, P.: RF and Baseband Techniques for Software Defined Radio. Artech House (2005)
[11] Rothe, H., Dahlke, W.: Theory of noisy fourpoles. In: Proceedings IRE, vol. 44, pp. 811–818 (1956)
[12] Domizioli, C.P., Hughes, B.L., Gard, K.G., Lazzi, G.: Optimal Front-End Design for MIMO Receivers. In: IEEE Global Telecommunications Conference (2008)
[13] Rouphael, T.J.: RF and Digital Signal Processing for Software Defined Radio, A Multi-Standard Multi-Mode Approch. Elsevier Newnes Publications (2009)
[14] Georgiadis, A.: Gain, phase imbalance, and phase noise effects on error vector magnitude. IEEE Transactions on Vehicular Technology 53(2), 443–449 (2004)
[15] Liu, R.F., Li, Y.M., Chen, H.Y., et al.: EVM estimation by analyzing transmitter imperfections mathematically and graphically. Analog Integrated Circuits and Signal Processing 48(3), 257–262 (2006)
[16] National Instruments, Introduction to WiMAX Transmitter Measurements (2009)
[17] IEEE Standard for Local and Metropolitan Area Networks: IEEE Std. 802.16-2004
[18] IEEE Standard for Local and Metropolitan Area Networks: IEEE Std. 802.16e-2005

Chapter 5
RF Power Amplifier and Linearization Techniques

The radio frequency (RF) power amplifier (PA) is one of the most critical components in designing transmitters in wireless communication systems, and its efficiency dominates the overall efficiency of the transmitter. On one hand, the PA is most power consuming ; for example, in a cellular phone, the battery life is largely determined by the power efficiency of the PA. On the other hand, it is desirable to have the ability to transmit data at the highest possible rate for a given channel bandwidth, i.e., to have high spectral efficiency [1].

5.1 Introduction

There is considerable industrial interest in producing RF PAs with good linearity and power efficiency. These two contradictory requirements can be achieved by using external circuitry to linearize an efficient amplifier [2]. Linearity is the ability of an amplifier to amplify all signals' parts by the same amount, so that all signals are amplified equally and is the most important characteristic of a PA. PAs must be linear to minimize interference and spectral regrowth. However, PAs generally have nonlinear behavior and are basically the main sources of distortion and nonlinearities in the RF transmitter.

Researchers have been focusing on designing more efficient power amplification techniques. It has been shown that the power efficiency of traditional amplification techniques could be improved at a cost of linearity degradation, which may not be tolerable by the standards' requirements. The current state of the art is the design of a moderately linear PA with the additional implementation of a linearization technique [3]. To maximize power efficiency, the amplifier should operate as close to saturation as possible, with the linearization system maximizing the spectral efficiency in this near-saturated region.

In this context, behavioral modeling and linearization of PAs are critical steps in designing high-performance power amplification systems for modern wireless communications infrastructure. Indeed, behavioral modeling of PAs is vital for performance estimation and system level simulation of the transmitter, which provides a time and computationally efficient alternative to the physics based modeling approach [4].

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 77–128.
springerlink.com

5.2 Transmitter Systems Parameters

The quality of a PA can be characterized by a number of specifications, many of which are provided in detail in the following sections.

5.2.1 Gain

The gain of an amplifier is the ratio of the output amplitude or power to the input amplitude or power, which is usually measured in decibels (dB). The power gain of a PA in dB is given by:

$$G(dB) = 10\log\left(\frac{P_{out}}{P_{in}}\right) \tag{5.1}$$

where P_{in} is the input power and P_{out} is the output power.

5.2.2 Bandwidth

The bandwidth of an amplifier is the range of frequencies for which the amplifier delivers acceptable performance. A well-accepted metric for performance is the half power points (i.e., frequencies where the power gain is half of its peak value). Hence, the bandwidth can be defined as the difference between the lower and upper half power points and is also called a 3-dB bandwidth.

5.2.3 Noise Figure

The noise factor, *F*, is a metric that gives an indication of noise added by the circuit and is defined as the input signal-to-noise ratio (SNR) divided by the output SNR, which is given by:

$$F = \frac{SNR_{in}}{SNR_{out}} \tag{5.2}$$

Since the amplifier always introduces noise, the SNR at the output is always less than that at the input; therefore, the noise factor is always greater than unity.

If the noise factor is expressed in dB, it is called a noise figure (NF) for the circuit, which is obtained as:

$$NF(dB) = 10\log(F) = SNR_{in,dB} - SNR_{out,dB} \tag{5.3}$$

5.2.4 Power Efficiency

Power efficiency in PAs is a metric to quantify the ability of a system to transform the given input power to useful output power. The most power-consuming component of

a wireless transmitter is the PA; and, in fact, the PA is the dominant component in determining the overall power efficiency of a wireless transmitter [5].

There are two forms of input power to the PA: direct current (DC) via the biasing circuits to the gate and drain of the transistor, and RF power at the input of the PA. The RF input power is amplified by transforming the DC power into RF power at the drain level. Part of the DC energy is dissipated as heat, which reduces the PA power efficiency.

There are three different definitions of power efficiency in the literature, which are the total efficiency, the drain efficiency, and the power-added efficiency [6].

5.2.4.1 Total Efficiency

The total efficiency can be obtained as the ratio of the RF power delivered at the output of the PA to the summation of the RF and DC powers that entered the PA. The total efficiency is obtained by:

$$\eta_t = \frac{P_{out}}{P_{dc} + P_{in}} \tag{5.4}$$

where η_t is the total efficiency of the PA; and, P_{dc}, P_{in}, and P_{out} are the DC and RF powers at the input and output of the PA, respectively.

5.2.4.2 Drain Efficiency

Drain efficiency, $\eta_{D\,trans.}$, is defined as the ratio of the RF power to the DC power at the drain level, which is specific to field-effect transistors (FETs). This efficiency is called collector efficiency in bipolar junction transistors (BJTs). Drain efficiency is given by:

$$\eta_{D\,trans.} = \frac{P_{RF\ Drain}}{P_{dc\ Drain}} \tag{5.5}$$

where $P_{dc\ Drain}$ and $P_{RF\ Drain}$ are the DC and RF powers, respectively, at the drain level of the transistor.

The DC power at the gate level of a transistor is usually very small compared to the DC power at the drain level. Hence, the DC power at the drain level of the transistor, $P_{dc\ Drain}$, can be approximated as the total DC power entering the PA, P_{dc}. By considering losses in the matching networks, the drain efficiency, η_D, is given as the ratio of the RF output power, P_{out}, to the DC power that entered the PA, P_{dc}, which is expressed as:

$$\eta_D = \frac{P_{out}}{P_{dc}} \tag{5.6}$$

5.2.4.3 Power-Added Efficiency

The power-added efficiency (PAE), η_{PAE}, demonstrates the ability of the PA to transform DC power to RF power. PAE can be obtained as the ratio of the added power, which is the difference between the output RF power, P_{out}, and the input RF power, P_{in}, to the DC power, P_{dc}. PAE is given by:

$$\eta_{PAE} = \frac{P_{out} - P_{in}}{P_{dc}} \tag{5.7}$$

5.2.5 P1dB

The 1 dB compression point (P1dB) is a measure of amplitude linearity. The gain of an amplifier compresses when the output signal level enters the compression region before it reaches saturation. Higher output power corresponds to a higher compression point. It is an input (or output) power for which the gain of the PA is 1 dB less than the small-signal gain (ideal linear gain).

The maximum saturation point corresponds to the point where the PA reaches its output maximum power. This maximum power is called the saturation power $P_{sat[\max]}$. The 3 dB saturation power $P_{sat[dB]}$ corresponds to the power for which the gain of the PA is 3 dB less than the small-signal gain. Figure 5.1 illustrates the P1dB, maximum saturation and 3 dB saturation points for a typical PA. To avoid intermodulation (IM) problems and distortion, the output power needs to be reduced below the P1dB. Manufacturers usually back off about 10 dB from the P1dB point.

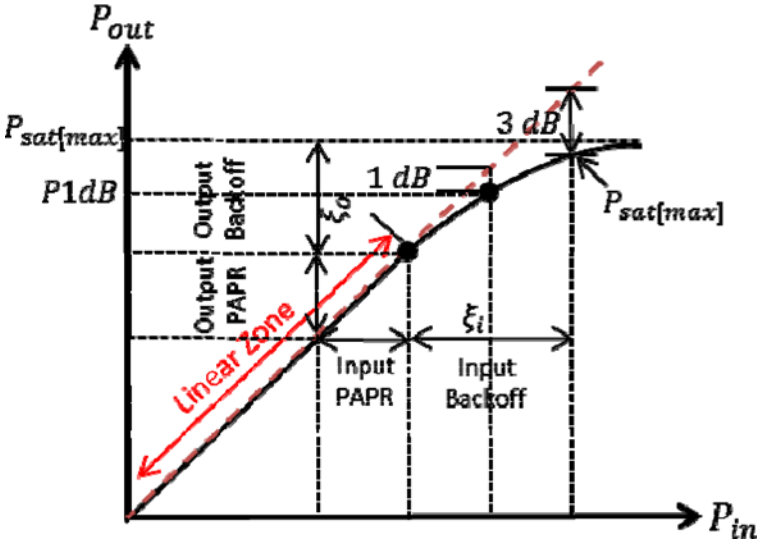


Fig. 5.1 P1dB and saturation power definition for PAs

5.2.6 IP3

A third-order intercept point (IP3) is a widely used metric in PAs, which gives information about the linearity of an amplifier. A higher IP3 means better linearity and lower distortion generation. It is a theoretical point at which the desired output signal and the third-order IM (undesired) signal are equal in levels considering an ideal linear gain for the PA. The theoretical input point is the input IP3 (IIP3) and the output power is the output IP3 (OIP3).

The two-tone test, as shown in Figure 5.2, is commonly used to determine IP3, which uses two closely spaced frequencies. When two signals at frequencies f_1 and f_2 are incident on an amplifier, the output of the amplifier contains these two signals, as well as IM products at frequencies $mf_1 \pm nf_2$, where $m+n$ is known as the order of the IM product. The ratio of third-order IM products and the carrier is known as third-order IM (IM3). IM3 products are important since their frequencies, $2f_1 - f_2$ and $2f_2 - f_1$, fall close to the desired signal, which makes filtering of IM3 an issue. IM3 for $2f_1 - f_2$ is given by:

$$IM3(dBc) = 10\log\left(\frac{P_{2f_1-f_2}}{P_{f_1}}\right) \tag{5.8}$$

and IM3 for $2f_2 - f_1$ is expressed by:

$$IM3(dBc) = 10\log\left(\frac{P_{2f_2-f_1}}{P_{f_2}}\right) \tag{5.9}$$

where p_{f_1}, p_{f_2}, $p_{2f_1-f_2}$, and $p_{2f_2-f_1}$ are the power outputs at frequencies f_1, f_2, $2f_1 - f_2$, and $2f_2 - f_1$, respectively. IM3 is measured in units of dBc, because it is calculated relative to the main tone power output.

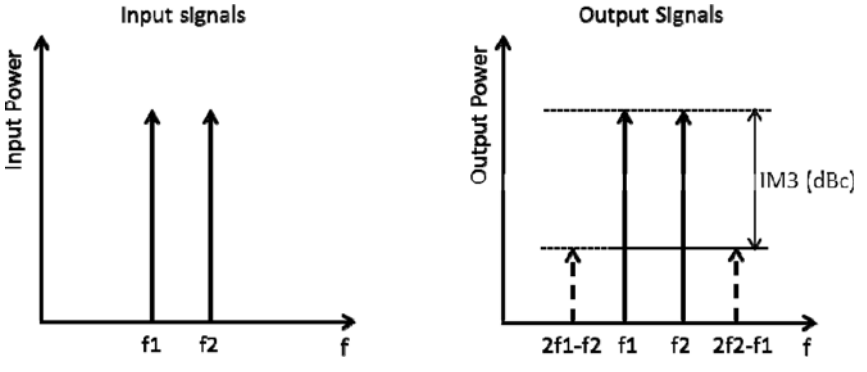


Fig. 5.2 Third-order intermodulation distortion (IM3)

5.2.7 PAPR

A metric used to quantify required linearity is called the peak-to-average power ratio (PAPR), which is the ratio between the peak power P_{peak} (related to peak amplitude) and the average power P_{avg} (related to mean amplitude) of a signal. It is also called the crest factor and is given by:

$$PAR(dB) = 10\log\left(\frac{\max(|x(t)|^2)}{mean(|x(t)|^2)}\right) = 10\log\left(\frac{P_{peak}}{P_{avg}}\right) \tag{5.10}$$

5.2.8 Power Back-Off

An amplifier appears linear for sufficiently small departures from its bias conditions. The power back-off is defined as the ratio between the PA's saturation power to the RF signal's mean power. There are two types of power back-off: input power back-off and output power back-off. The back-off at the input of the PA, ξ_i, is obtained by:

$$\xi_i = P_{i,sat} - P_{i,mean} \quad (dB) \tag{5.11}$$

where $P_{i,sat}$ and $P_{i,mean}$ are the saturation power and mean signal power at the input of the PA, respectively.

Similarly, the back-off at the output of the PA, ξ_o, is given by:

$$\xi_o = P_{o,sat} - P_{o,mean} \quad (dB) \tag{5.12}$$

where $P_{o,sat}$ and $P_{o,mean}$ are the saturation power and mean signal power at the output of the PA, respectively.

The input power back-off and output power back-off are illustrated in Figure 5.1.

5.2.9 ACPR

The adjacent channel power ratio (ACPR) is a critical figure of merit in the evaluation of the IM distortion performance of RF PAs. It is a measure of spectral regrowth and appears in the signal sidebands. ACPR is defined as the ratio of power in a bandwidth adjacent to the main channel to the power within the main signal bandwidth.

The ACPR for the right side of the power spectral density (PSD) can be defined as:

$$ACPR(Right) = \frac{\int_{f_c+\Delta f-\frac{B}{2}}^{f_c+\Delta f+\frac{B}{2}} PSD(f)df}{\int_{f_c-\frac{B}{2}}^{f_c+\frac{B}{2}} PSD(f)df} \tag{5.13}$$

Similarly, the ACPR for the left side of the PSD can be obtained as:

$$ACPR(Left) = \frac{\int_{f_c-\Delta f-\frac{B}{2}}^{f_c-\Delta f+\frac{B}{2}} PSD(f)df}{\int_{f_c-\frac{B}{2}}^{f_c+\frac{B}{2}} PSD(f)df} \tag{5.14}$$

where f_c is the carrier frequency, B is the bandwidth of the modulated signal, and $PSD(f)$ is the power spectral density at frequency f. Figure 5.3 illustrates the way that ACPR is calculated.

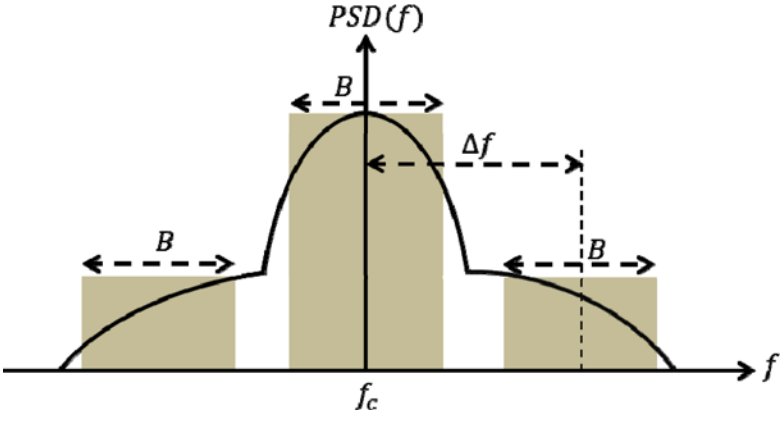


Fig. 5.3 Adjacent channel power ratio (ACPR)

5.2.10 EVM

The effect of noise, distortion, and in-phase / quadrature (I/Q) imbalance on signal quality can be analyzed based on its error vector magnitude (EVM) measurement. The EVM is a common metric for the fidelity of the symbol constellation. Ideally, a signal sent by a transmitter or received by a receiver would have all constellation points precisely at the ideal locations. However, several deficiencies in implementation, such as I/Q imbalance, noise and carrier leakage, cause the deviation of constellation points from their ideal locations. In other words, EVM is a measure of how far the actual (measured) points are from the ideal locations. The discrepancy between actual and ideal symbol vectors is quantified through an error vector, as shown in Figure 5.4.

The EVM is the ratio of the power of the error vector to the power of the reference vector related to the ideal constellation. The EVM can be defined in decibels (dB) or percentage (%) as:

$$EVM\,(dB) = 10\log_{10}\left(\frac{P_{error}}{P_{ideal}}\right) = 10\log_{10}\left(mean\left(\frac{(I_{actual} - I_{ideal})^2 + (Q_{actual} - Q_{ideal})^2}{I_{ideal}^2 + Q_{ideal}^2}\right)\right) \tag{5.15}$$

$$EVM\,(\%) = \sqrt{\frac{P_{error}}{P_{ideal}}} \times 100 \tag{5.16}$$

where P_{error} and P_{ideal} are the root mean square (RMS) power of the error vector and the outermost (highest power) point in the reference signal constellation, I_{ideal} and Q_{ideal} are the ideal in-phase and quadrature signals, and I_{actual} and Q_{actual} are the transmitted in-phase and quadrature signals.

EVM can be also represented as:

$$EVM = \sqrt{\frac{\frac{1}{N}\sum_{n=1}^{N}\left|S_{ideal,n} - S_{actual,n}\right|^2}{\frac{1}{N}\sum_{n=1}^{N}\left|S_{ideal,n}\right|^2}} \tag{5.17}$$

where N is the number of unique symbols in the constellation.

EVM is useful in quantifying the amount of noise and distortion in a transmitter. The random distribution of the error vectors is caused by noise, while patterns in the error vector measurements correlated with symbol amplitude and phase may be caused by distortion.

It is common to characterize AM/AM (amplitude modulation to amplitude modulation) distortion and AM/PM (amplitude modulation to phase modulation) distortion, which is shown in Figure 5.5. Variation in the transmitter gain causes AM/AM distortion. Phase shifting of the carrier correlated to the signal amplitude introduces AM/PM distortion [7].

Undesired expansion of the output spectrum is also caused by distortion. Nonlinearity of a PA that can be quantified as AM/AM and AM/PM distortion generates counterfeit energy near the fundamental and harmonics of the carrier frequency. This broadening in the output spectrum is referred to as spectral regrowth or out-of-band distortion, which deteriorates the ACPR. The transmitter output spectrum showing spectral regrowth is illustrated in Figure 5.6. Spectral regrowth is a potentially significant problem and demands special attention during the transmitter design process. The nonlinearity in the device can be compensated for using linearization techniques [4], which are covered in Section 5.4.

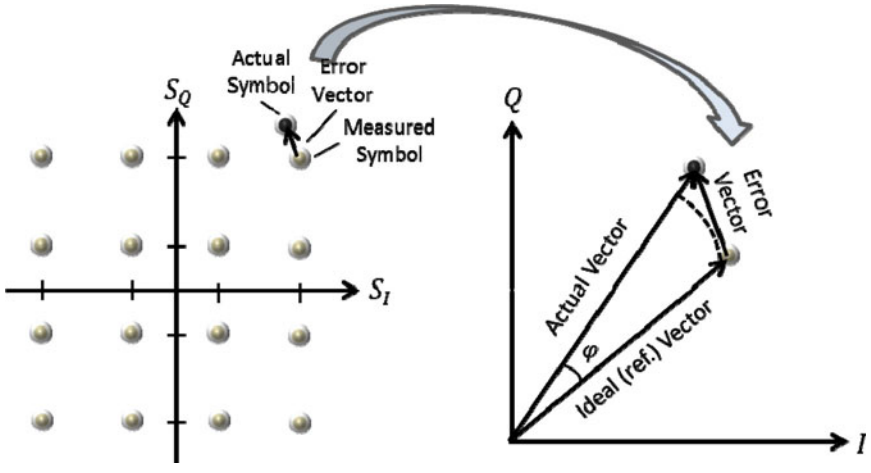


Fig. 5.4 Error vector is the difference between the actual and ideal symbol vectors

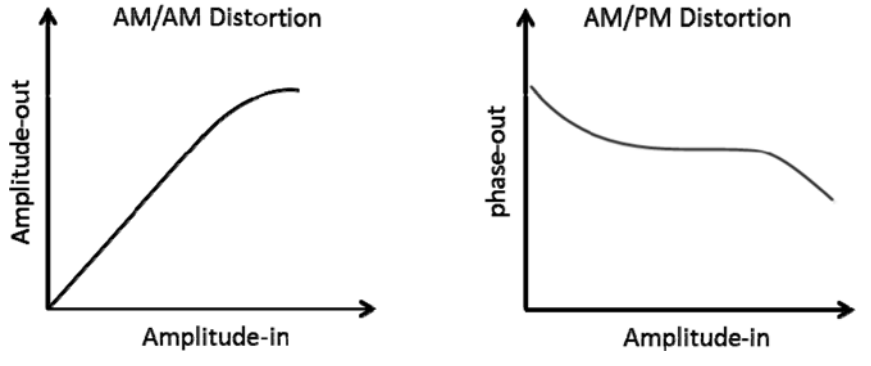


Fig. 5.5 AM/AM distortion and AM/PM distortion in transmitters

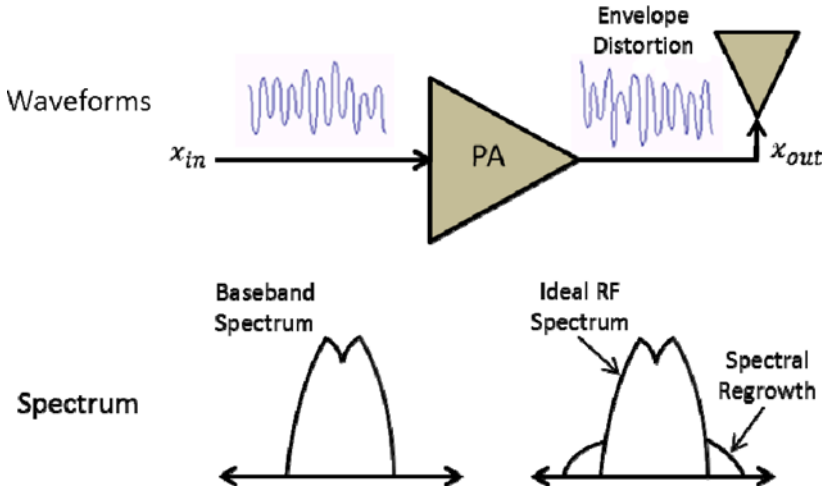


Fig. 5.6 Spectral regrowth in transmitter output spectrum

5.2.11 Memory Effect

In a two- or multi-tone IM test, if the amplitude or phase of the IM signals is affected by the tone frequency spacing, the PA exposes memory effects. Memory effects can be explained as time lags between the AM/AM and AM/PM responses of the PA. Electrical and thermal are two types of memory effects.

The electrical memory effect is produced by poor gate and drain decoupling in FET and base and collector decoupling in BJT at low frequencies. Electrical memory effects cause the distortion of the envelope currents and result in IM asymmetry. The memory effect is more significant for class AB PAs than class A, with a reduced conduction angle.

Electrical memory effects are generated by non-constant node impedances within frequency bands. Most of them are produced by frequency-dependent envelope impedances. The thermal memory effect is generated by the junction temperature that is modulated by the applied signal envelope.

5.3 RF Power Amplifiers

The wireless and satellite communications communities are always searching for radios that are more power efficient. However, the power efficiency is greatly dominated by the efficiency of the RF PA in the transmitting path. For a transistor, the output power, gain and power efficiency depend on the biasing and matching conditions, which define the PA's class of operation. A diagram showing the conduction angles, as well as the input drive levels, is presented in Figure 5.7 for the different classes of power amplifiers: (i) continuously driven class A, AB, B and C PAs, (ii) saturated PAs and (iii) class D, E and F switching-modes PAs [1].

By reducing the conduction angle, the efficiency of the PAs is increased. The efficiency is even higher when operating the PA in a switching mode; however, efficiency enhancement is accompanied with linearity degradation. The higher the efficiency when passing from one class to another, the poorer is the linearity performance.

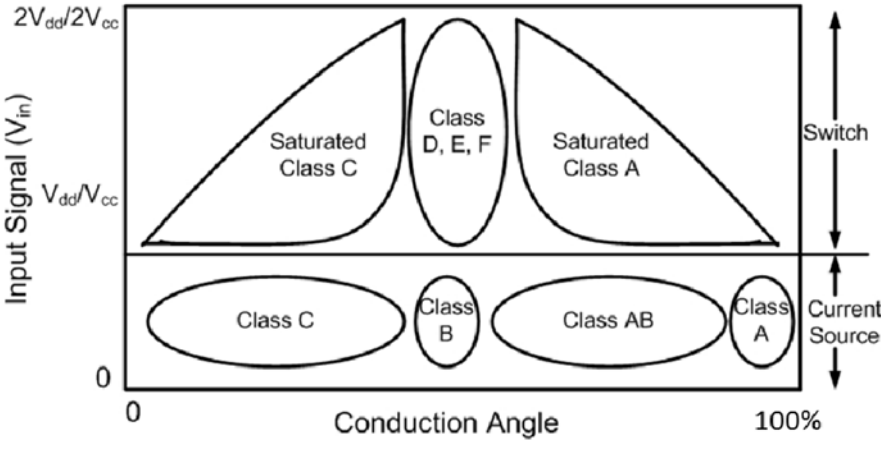


Fig. 5.7 Classes of power amplifiers [1]

Time-varying envelope signals are implemented in most communication systems that have recently been deployed or are to be deployed in the near future. PAs are required to operate in their back-off region with such signals to meet the required linearity. This linearity is defined either by the ACPR or EVM in the context of wireless applications. The back-off level setting is generally a function of the input signal's PAPR. Unfavorably, the power efficiency of the PA decreases as the back-off level increases. Therefore, this leads to the design of very low efficiency amplifiers that need large DC power modules for fixed terminals or shorten battery life for mobile terminals. In fact, linearization techniques are being used to improve achievable power efficiency [1].

5.3.1 Linear PAs (Classes A, AB, B, C)

This section discusses several types of classic RF PAs, which are called classes A, AB, B and C. Power amplifiers, depending on the class of operation, behave with different levels of nonlinearity and provide different levels of efficiency. These PAs, in terms of levels of linearity and power efficiency, are described; and, the tradeoff between the linearity and efficiency is discussed.

A general block diagram of a PA is shown in Figure 5.8, which consists of a transistor with its biasing circuits and the input and matching networks. The required current and voltage to bias the transistors are provided by the biasing circuits. The input and output matching networks are designed to transfer the maximum power through the PA.

The class of operation of the PA relies upon the biasing level of the transistor. For instance, the DC current consumption variation of the FET transistor is shown in Figure 5.9, as a function of the gate voltage for a given drain voltage. The DC current in this figure cuts into two regions. In the first region, the current is zero; therefore, the transistor is off. In the second region, the current changes almost linearly with the gate voltage. The gate voltage at the edge of these two regions is described as the pinch-off voltage. The gate bias regions corresponding to class A, AB, B and C PAs are also illustrated in Figure 5.9.

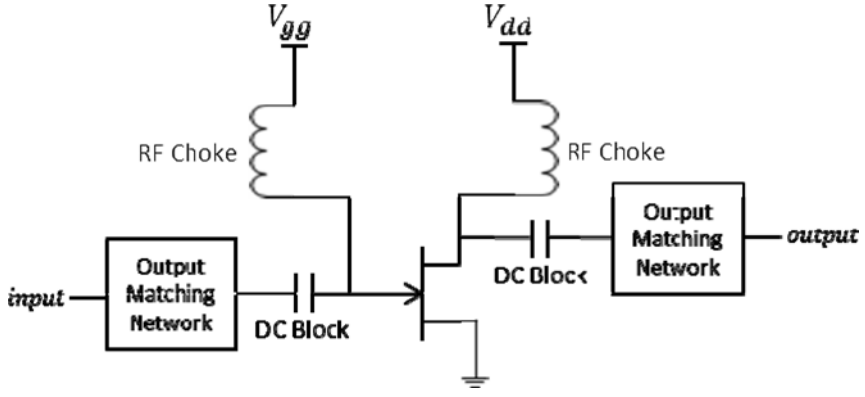


Fig. 5.8 A general block diagram of a PA

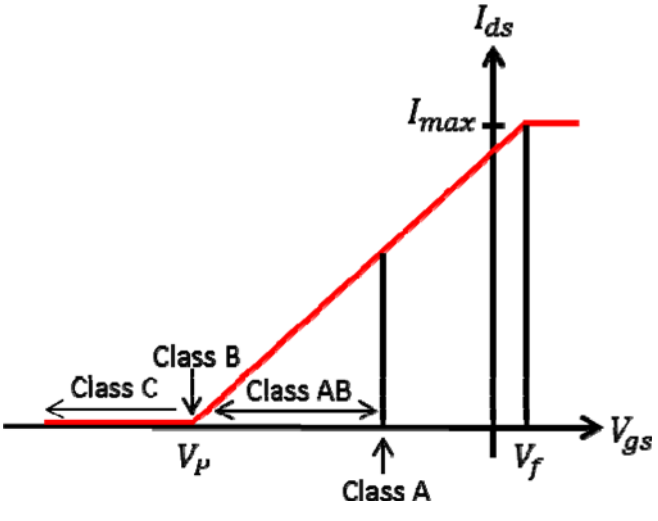


Fig. 5.9 The DC current consumption versus the gate voltage of the FET transistor

In class B operation, the transistor starts conducting as soon as the input RF signal is presented at the input of the PA; in fact, the PA consumes power as long as the RF signal is present at the input. Operation of the PA in classes A, AB and C can be achieved by biasing the transistor at the gate level, below or above the pinch-off voltage. In class A, the transistor is on, even when there is no input RF signal. Hence, the transistor conducts the input RF signal with no limitation on the input power level up to the maximum level that the transistor can handle. In contrast, in class C, the transistor is off as long as the input RF signal amplitude does not exceed a power threshold. Class A PAs have a linear power transfer function, while it becomes nonlinear moving toward class C.

The conduction angle, α, is defined as the angular period of the input RF signal cycle in which the transistor conducts. It varies between 0 to 2π, and $\alpha = \pi$ corresponds to class B operation. The conduction angle is calculated by:

$$\alpha = 2\cos^{-1}\left(1 - \frac{V_P}{V_{gs}}\right) \tag{5.18}$$

where V_P is the pinch-off voltage, and V_{gs} is the gate-source voltage.

It was theoretically proven [8] that the maximum drain efficiency of the PA can be related to the conduction angle by the following equation:

$$\eta_{\max} = \frac{\alpha - \sin(\alpha)}{4\sin\left(\frac{\alpha}{2}\right) - 2\alpha\cos\left(\frac{\alpha}{2}\right)} \tag{5.19}$$

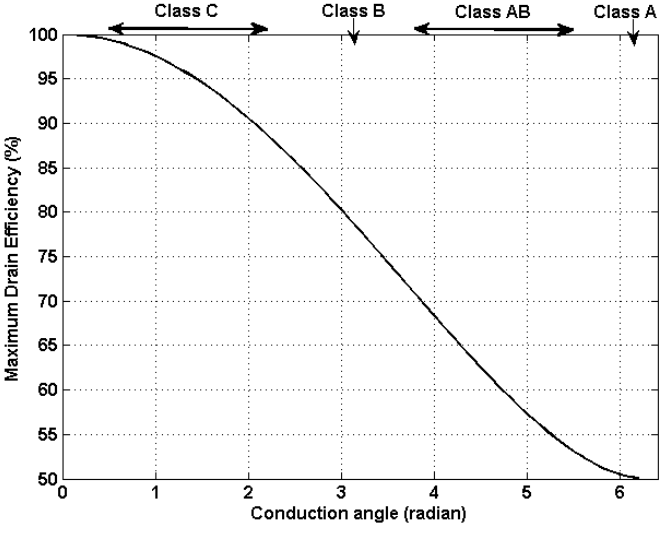


Fig. 5.10 The maximum drain efficiency versus the conduction angle

Based on the previous equation, the maximum drain efficiency of the PA increases from 50% for a class A PA with a conduction angle of 2π to 100% for a class C PA with a conduction angle of 0. This is achieved by decreasing the conduction angle. The theoretical maximum drain efficiency of the PA for a conduction angle between 0 to 2π is demonstrated in Figure 5.10. It can be seen that the maximum efficiency of the PA increases from class A to class C, where the class A PA has a linear power transfer function and the class C PA has a nonlinear power transfer function. Indeed, there should be a tradeoff between the efficiency and linearity of the PA.

Considering the preceding discussion, the choice of a proper class of operation for a PA in wireless transmitters is fairly challenging. Signal quality degradation can be avoided by highly linear PAs, but it results in poor power efficiency. The overall linearity and efficiency of wireless transmitters are highly affected by the PA, since it is the dominant source of nonlinearities and power consumption. A class AB PA supplies an acceptable tradeoff between linearity and power consumption.

5.3.1.1 Class A PA

The class A PA has the highest linearity over the other classes of operation. In the class A amplifier, the conduction angle is set to $\alpha = 360^\circ$; and, current flows constantly and is not cut off during any part of the cycle, acting as a current source. To achieve high linearity and gain, the amplifier's base and drain DC voltage has to be chosen properly so that the amplifier can operate in a linear region. There is continuous loss of power in the amplifier, since it is constantly carrying current. The class of operation A is distinguished owing to the fact that it is, in theory, perfectly linear, but inefficient.

In a Class A PA, the gate and drain of a FET transistor are biased according to:

$$V_{gg} = \frac{V_P + V_f}{2} \tag{5.20}$$

$$V_{dd} = \frac{V_{\max} + V_{knee}}{2} \tag{5.21}$$

where V_P is the gate pinch-off voltage, V_f is the gate diode forward voltage, $V_{\max}$ is the maximum tolerable drain voltage of the transistor, and V_{knee} is the drain knee voltage. V_P, V_f, $V_{\max}$, and V_{knee} are inherent parameters of a FET transistor and can be collected from the device data sheet.

To achieve the maximum drain current swing without clipping, the maximum amplitude of the RF gate voltage should be chosen as $V_f - V_{gg}$, which is illustrated in Figure 5.11 (a). The input power corresponding to this gate voltage is called the maximum linear input power. Consequently, the device conducts for the entire RF cycle. The maximum drain current swing, and thus maximum linear output power, can be obtained by proper selection of the load impedance. The gate and drain voltages and drain current waveforms for a class A PA are illustrated in Figure 5.11, as a function of the angular phase, θ.

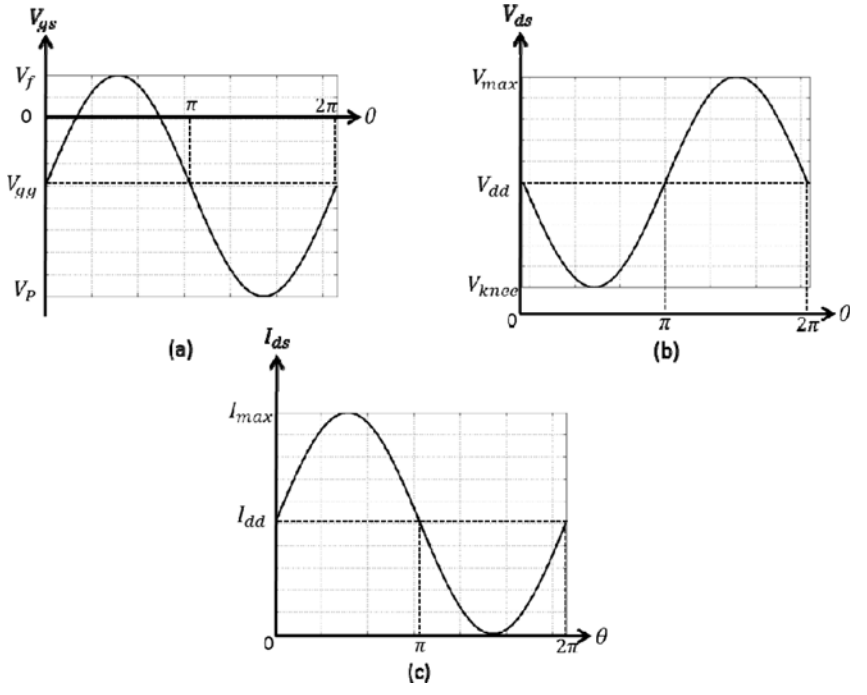


Fig. 5.11 Class A: (a) gate voltage, (b) drain voltage and (c) drain current waveforms

The drain current shown in Figure 5.11 (c) is the response of the transistor, which is assumed to be an ideal voltage-dependent current source, to the applied gate voltage shown in Figure 5.11 (a). The drain voltage is also represented in Figure 5.11 (b) with the assumption of providing the optimal load for the maximum linear output power. The gate and drain voltages and current corresponding to the waveforms in Figure 5.11 can be described as:

$$V_{gs}(\theta) = V_{gg} + v_g \cdot \sin\theta \text{, where } v_g = V_f - V_{gg} \tag{5.22}$$

$$V_{ds}(\theta) = V_{dd} - v_d \cdot \sin\theta \text{, where } v_d = V_{dd} - V_{knee} \tag{5.23}$$

$$I_{ds}(\theta) = I_{dd} + i_d \cdot \sin\theta \text{, where } i_d = I_{dd} = I_{\max} / 2 \tag{5.24}$$

Thus, the optimal load, $R^A_{opt.}$, for the maximum drain current swing can be calculated as:

$$R^A_{opt.} = \frac{v_d}{i_d} = \frac{V_{dd} - V_{knee}}{I_{\max} / 2} = 2\frac{V_{dd} - V_{knee}}{I_{\max}} \tag{5.25}$$

The maximum output power at peak drain efficiency is obtained if the transistor works on a load line of $R^A_{opt.}$. The required DC power, $P^A_{dc,\max}$, maximum RF output power, $P^A_{RF,\max}$, and peak drain efficiency, $\eta^A_{D,\max}$, for class A operation can be obtained using the following equations:

$$P^A_{dc,\max} = V_{dd} I_{dd} = \frac{V_{dd} \cdot I_{\max}}{2} \tag{5.26}$$

$$P^A_{RF,\max} = \frac{v_d i_d}{2} = \frac{(V_{dd} - V_{knee}) \cdot I_{\max}}{4} \tag{5.27}$$

$$\eta^A_{D,\max} = \frac{P^A_{RF,\max}}{P^A_{dc,\max}} = \frac{(V_{dd} - V_{knee})}{2V_{dd}} \tag{5.28}$$

if $V_{knee} = 0$, then $\eta^A_{D,\max} = \frac{1}{2}$.

The dissipated power in the transistor, which is the difference between the injected DC power to the amplifier and the RF output power produced by the PA, can be obtained as:

$$P^A_{diss} = \frac{1}{2\pi} \int_0^{2\pi} V_{ds}(\theta) \cdot I_{ds}(\theta) d\theta = P_{dc} - P_{RF} = \frac{V_{\max} \cdot I_{\max}}{8} \tag{5.29}$$

5.3.1.2 Class B PA

As shown in Figure 5.9, in the class B PA, the gate of the transistor is biased at the pinch-off voltage, and the conduction angle is set to $\alpha = 180^\circ$. Hence, the transistor remains in the active region during half of the RF cycle. A class B PA operates at zero quiescent current. This half-sine waveform of the drain current reduces the corresponding DC power; therefore, the power efficiency is increased.

In the class B mode of operation, the gate and drain of a FET transistor are biased using:

$$V_{gg} = V_P \tag{5.30}$$

$$V_{dd} = \frac{V_{\max} + V_{knee}}{2} \tag{5.31}$$

The gate and drain voltages and drain current waveforms for a class B PA are depicted in Figure 5.12. As in the class A mode of operation, the load-line impedance at the fundamental frequency is chosen to obtain the maximum drain voltage and current swing. The harmonics of the fundamental frequency are ideally shortened to ground at the output of the PA.

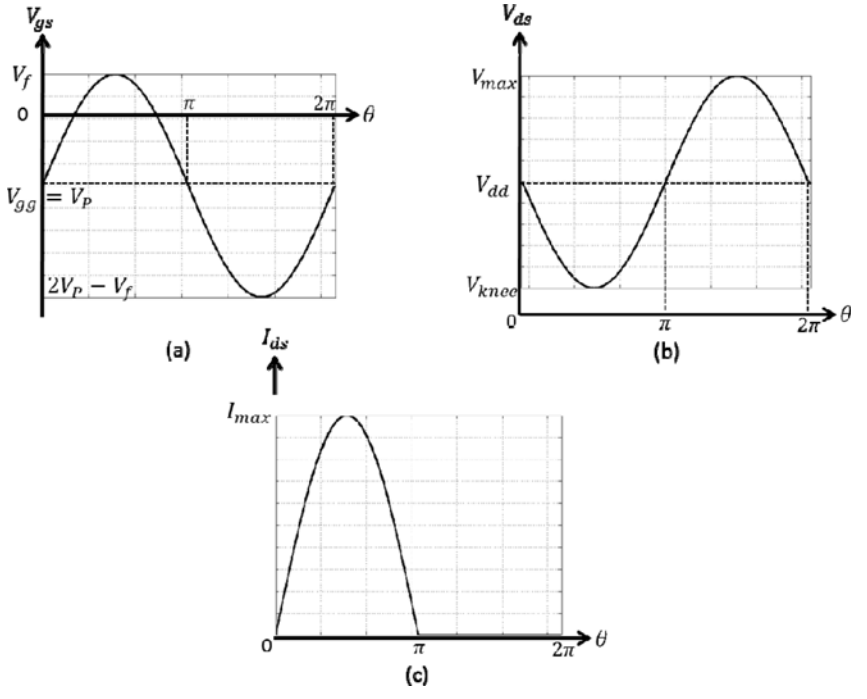


Fig. 5.12 Class B: (a) gate voltage, (b) drain voltage and (c) drain current waveforms

In the class B mode of operation, injecting a sinusoidal waveform to a transistor with the gate biased at the pinch-off voltage results in a half-sine waveform of the drain current that consists of fundamental and even harmonic frequencies. The drain voltage only includes the fundamental frequency, since all higher harmonic voltage components are suppressed by proper harmonic termination. The gate and drain voltage equations are exactly the same as those in (5.22) and (5.23), respectively, but with different V_{gg}. The drain current is calculated as:

$$I_{ds}(\theta) = \begin{cases} I_{\max} \cdot \sin\theta & 0 \le \theta \le \pi \\ 0 & \pi < \theta \le 2\pi \end{cases} \tag{5.32}$$

The amplitude of the DC and the fundamental frequency content of the drain current of a class B PA can be calculated with the following equations:

$$I_{dd} = \frac{1}{2\pi}\int_0^{2\pi} I_{ds}(\theta)d\theta = \frac{I_{\max}}{\pi} \tag{5.33}$$

$$i_{d,1} = \frac{1}{\pi}\int_0^{2\pi} I_{ds}(\theta)\cdot\sin(\theta)d\theta = \frac{I_{\max}}{2} \tag{5.34}$$

Assuming the same maximum drain voltage and current as in the class A mode of operation can be obtained, the peak output power of class B is equal to that of the class A PA, but at the cost of 6 dB more input power requirement. This results in a gain that is lower by 6 dB. The reason is that only half of the input RF signal is amplified by the PA. However, class B mode of operation achieves higher peak efficiency than class A, due to the lower average DC current, and, hence, the lower dissipated DC power.

A class B PA yields maximum power when offered the same optimal load line as that of a class A PA, which is calculated in (5.25). The required DC power, $P^B_{dc,\max}$, maximum fundamental RF output power, $P^A_{RF,\max}$, and maximum power efficiency, $\eta^A_{D,\max}$, for the class B operation mode can be calculated based on the following equations:

$$P^B_{dc,\max} = V_{dd} I_{dd} = \frac{V_{dd} \cdot I_{\max}}{\pi} \tag{5.35}$$

$$P^B_{RF,f_0,\max} = \frac{v_{d,1} i_{d,1}}{2} = \frac{(V_{dd} - V_{knee}) \cdot I_{\max}}{4} \tag{5.36}$$

$$\eta^A_{D,\max} = \frac{P^B_{RF,f_0,\max}}{P^B_{dc,\max}} = \frac{\pi (V_{dd} - V_{knee})}{4 V_{dd}} \tag{5.37}$$

if $V_{knee} = 0$, then $\eta^B_{D,\max} = \frac{\pi}{4} = 78.5\%$

The theoretical optimal efficiency of the class B PA is $\eta^B_{D,\max} = 78.5\%$ higher than that of the class A PA $\eta^A_{D,\max} = 50\%$, with the cost of reduction in the linearity of the device.

The dissipated power in the transistor for a class B mode of operation can be calculated as:

$$P^B_{diss} = \frac{1}{2\pi} \int_0^{2\pi} V_{ds}(\theta) \cdot I_{ds}(\theta) d\theta = P_{dc} - P_{RF} = \frac{4 - \pi}{8\pi} V_{\max} \cdot I_{\max} \tag{5.38}$$

In comparison to (5.29), the dissipated power in the transistor in class B PA is around 27% of the power dissipation in class A PA for the same maximum output power.

In conclusion, a class B PA yields the same maximum output power as class A PA, while producing higher efficiency, but 6 dB less gain. Furthermore, a more complex load network is required, due to the requirement of proper harmonic frequency termination.

5.3.1.3 Class AB PA

A class AB PA compromises between class A and class B operation in terms of efficiency and linearity. The conduction angle is $180^\circ < \alpha < 360^\circ$, so that the transistor remains in the active region for more than half of but less than the full cycle of the input RF signal. Indeed, the efficiency in the class AB mode of operation is between 50% and 78.5%. Class AB PAs offer a wider dynamic range than either class A or B PAs. The linearity of class AB is quite similar to a class A PA before its input power becomes large enough to cause the drain current clipping from the bottom, which is the time when nonlinearity starts and the gain decreases.

5.3.1.4 Class C PA

In the class C mode of operation, the gate bias point is chosen below the pinch-off voltage of the device. In a class C PA, the conduction angle is $\alpha < 180^\circ$, and the transistor remains in the active region for less than half of the RF cycle. Hence, in order to achieve the maximum transistor drain current, $I_{\max}$, more input power than in the class B is needed. The class C PA has higher efficiency than classes A, AB and B, but it is highly nonlinear. It is biased so that the output current is zero for more than one half of an input sinusoidal signal cycle.

In the class C operation mode, the gate and drain of a FET transistor are biased according to:

$$V_{gg} < V_P \tag{5.39}$$

$$V_{dd} = \frac{V_{\max} + V_{knee}}{2} \tag{5.40}$$

Figure 5.13 shows the gate and drain voltages and drain current waveforms for a class C PA. The gate and drain voltage equations are the same as those in (5.22) and (5.23), respectively, but with a different V_{gg}. The conduction angle in Figure 5.13 (c) can be calculated by:

$$V_P = V_{gg} + v_g \cdot \sin\gamma \tag{5.41}$$

where, $\gamma = \frac{\pi}{2} - \theta_0$, then,

$$\theta_0 = \cos^{-1}\left(\frac{V_P - V_{gg}}{v_g}\right) \tag{5.42}$$

Assuming an ideal linear relationship between the drain current, I_{ds}, and the gate voltage, V_{gs}, as demonstrated in Figure 5.9, the drain current can be specified as:

$$I_{ds}(\theta) = \frac{I_{\max}}{V_f - V_P} V_{gs}(\theta) - \frac{I_{\max} \cdot V_P}{V_f - V_P} \tag{5.43}$$

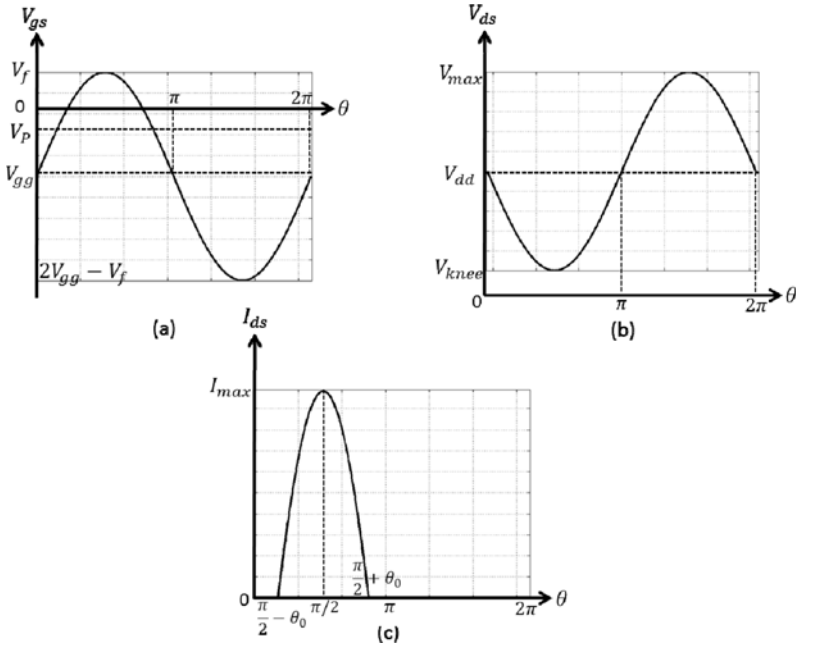


Fig. 5.13 Class C: (a) gate voltage, (b) drain voltage and (c) drain current waveforms

Based on the class C conduction angle and applying (5.22) to (5.43), the drain current can be calculated as:

$$I_{ds}(\theta) = \begin{cases} \dfrac{I_{\max} v_g}{V_f - V_P} \sin\theta + \dfrac{I_{\max}(V_{gg} - V_P)}{V_f - V_P} & \dfrac{\pi}{2} - \theta_0 \le \theta \le \dfrac{\pi}{2} + \theta_0 \\ 0 & otherwise \end{cases} \tag{5.44}$$

This equation is valid for all the classes of operation, if v_g, V_{gg} and θ_0 are suitably chosen. In the class C PA, the drain current waveform includes both odd and even harmonic frequency content. Hence, all higher frequency components need to be short-circuited at the output to have zero drain voltage at all harmonics except at the fundamental frequency.

Using (5.44), the drain current at DC and the fundamental frequency can be derived in a few steps with the following equations:

$$I_{dd} = \frac{1}{2\pi}\int_0^{2\pi} I_{ds}(\theta)d\theta = \frac{I_{\max} \cdot v_g}{\pi(V_f - V_p)}(\sin\theta_0 - \theta_0\cos\theta_0) \tag{5.45}$$

$$i_{d,1} = \frac{1}{2\pi}\int_0^{2\pi} I_{ds}(\theta) \cdot \sin(\theta)d\theta = \frac{I_{\max} \cdot v_g}{\pi(V_f - V_p)}(\theta_0 - \sin\theta_0\cos\theta_0) \tag{5.46}$$

To achieve peak drain current and, accordingly, maximum output power, the required RF gate voltage amplitude of a class C PA should be selected as:

$$v_g = V_f - V_{gg} \tag{5.47}$$

Using (5.41) and (5.47), v_g can be calculated as a function of θ_0 as:

$$v_g = \frac{V_f - V_P}{1-\cos\theta_0} \tag{5.48}$$

By substituting (5.48) in (5.45) and (5.46), the following equations can be obtained:

$$I_{dd} = \frac{I_{\max}}{\pi}\frac{(\sin\theta_0 - \theta_0\cos\theta_0)}{1-\cos\theta_0} \tag{5.49}$$

$$i_{d,1} = \frac{I_{\max}}{\pi}\frac{(\theta_0 - \sin\theta_0\cos\theta_0)}{1-\cos\theta_0} \tag{5.50}$$

The class C optimal load line, $R^C_{opt.}$, for the maximum drain voltage swing is calculated as:

$$R^C_{opt.} = \frac{v^C_{d,1}}{i^C_{d,1}} = \frac{\pi(V_{dd} - V_{knee})}{I_{\max}} \cdot \frac{1-\cos\theta_0}{(\theta_0 - \sin\theta_0 \cdot \cos\theta_0)} \tag{5.51}$$

Biasing the gate deeper in class C decreases the drain current; hence, a larger load resistor is needed to achieve the maximum voltage swing.

The required DC power, $P^C_{dc,\max}$, maximum fundamental RF output power, $P^C_{RF,\max}$, and maximum drain power efficiency, $\eta^C_{D,\max}$, for the class C operation mode can be calculated based on the following equations:

$$P^C_{dc,\max} = V_{dd} I_{dd} = \frac{V_{dd} \cdot I_{\max}}{\pi} \cdot \frac{(\sin\theta_0 - \theta_0\cos\theta_0)}{1-\cos\theta_0} \tag{5.52}$$

$$P^C_{RF,f_0,\max} = \frac{v_{d,1} i_{d,1}}{2} = \frac{I_{\max} \cdot (V_{dd} - V_{knee})}{2\pi} \cdot \frac{(\theta_0 - \sin\theta_0\cos\theta_0)}{1-\cos\theta_0} \tag{5.53}$$

$$\eta^C_{D,\max} = \frac{P^C_{RF,f_0,\max}}{P^C_{dc,\max}} = \frac{(V_{dd} - V_{knee})}{2V_{dd}} \cdot \frac{(\theta_0 - \sin\theta_0\cos\theta_0)}{(\sin\theta_0 - \theta_0\cos\theta_0)} \tag{5.54}$$

if $V_{knee} = 0$, then $\eta^C_{D,\max} = \frac{1}{2}\frac{(\theta_0 - \sin\theta_0\cos\theta_0)}{(\sin\theta_0 - \theta_0\cos\theta_0)}$

The dissipated power in the transistor for class C operation mode can be calculated as:

$$P_{diss}^{B} = \frac{1}{2\pi}\int_{0}^{2\pi} V_{ds}(\theta)\cdot I_{ds}(\theta)d\theta = P_{dc} - P_{RF} = \frac{V_{\max}\cdot I_{\max}}{4\pi}\cdot\frac{2\sin\theta_0 + \sin\theta_0\cos\theta_0 - 2\theta_0\cos\theta_0 - \theta_0}{1-\cos\theta_0} \quad (5.55)$$

5.3.2 Switching-Mode PAs (Classes D, E, F)

In contrast to class A, AB, B and C PAs, where operation in the triode region should be avoided, class D, E, and F PAs rely on operation in the triode region for optimal efficiency and output power. In these amplifiers, which are also called switching-mode amplifiers, the output device is driven by a large square wave signal. For signals that have mainly phase and frequency modulation, such as quadrature phase-shift keying (QPSK) or Gaussian minimum-shift keying (GMSK), the envelope is constant. This means that a nonlinear high-efficiency PA can be used to amplify such signals.

The current and voltage waveforms at the drain level of the transistor can be shaped to reduce the power dissipation and enhance the efficiency of PAs. This reduction can be achieved by avoiding simultaneous presence of high current and high voltage at the drain level of the transistor, since the power dissipation is equal to the product of both current and voltage.

In the ideal case, no power dissipation occurs if the transistor behaves as a lossless switch, which makes the transistor operate in two states. In the first state, when the transistor is on, the voltage across it is zero and the current is high, i.e., the transistor behaves as a closed switch during this part of a cycle. In the second state, when the transistor is off, its current is zero and its voltage is high, i.e., the transistor behaves as an open switch during this part of a cycle.

The gate and drain voltages and drain current waveforms for an ideal switch-like transistor are shown in Figure 5.14. One way to achieve the switching behavior is proper termination of the device at the fundamental and harmonic frequencies, in order to shape the current and voltage waveforms to the switching-mode operation [9].

The drain voltage and current corresponding to the waveforms in Figure 5.14 can be described as:

$$I_{ds}(\theta) = \frac{I_{\max}}{2} + \sum_{n=0}^{\infty}\frac{2I_{\max}}{(2n+1)\pi}\sin(2n+1)\theta \quad (5.56)$$

$$V_{ds}(\theta) = \frac{V_{\max}}{2} + \sum_{n=0}^{\infty}\frac{2V_{\max}}{(2n+1)\pi}\sin(2n+1)\theta \quad (5.57)$$

where $\frac{V_{\max}}{2} = V_{dd}$.

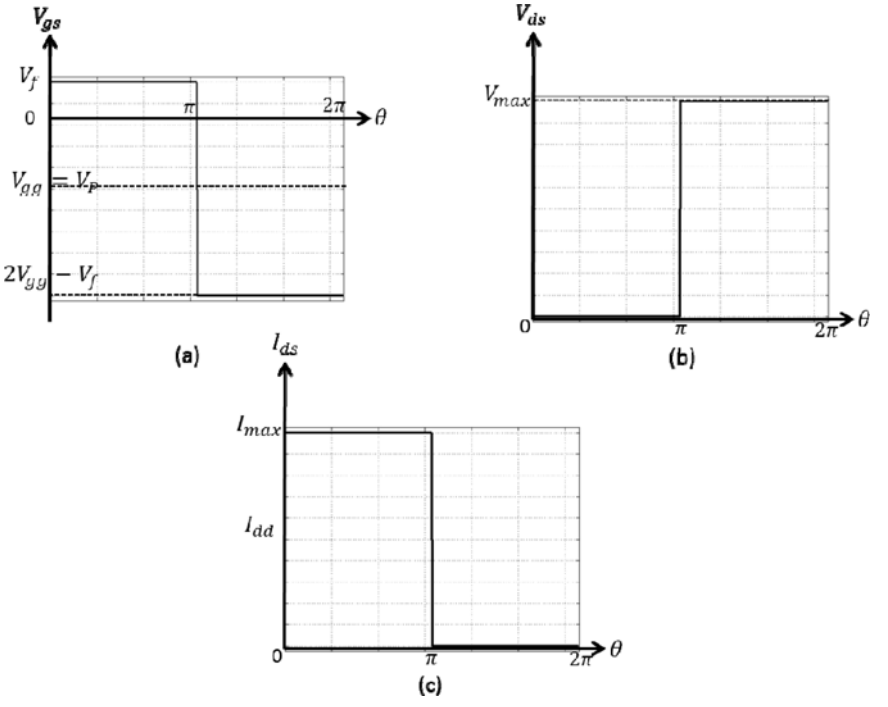


Fig. 5.14 Ideal switch-like transistor: (a) gate voltage, (b) drain voltage and (c) drain current waveforms

The dissipated power in the transistor for ideal switching-mode operation is calculated as:

$$P_{diss}^{B} = \frac{1}{2\pi}\int_{0}^{2\pi} V_{ds}(\theta)\cdot I_{ds}(\theta)d\theta = P_{dc} - P_{RF} = 0 \tag{5.58}$$

This is because the multiplication of the drain voltage and drain current is zero at each angular phase. Hence, all the applied DC power is transfer to RF output power. The DC power is calculated as:

$$P_{dc} = \frac{V_{\max} I_{\max}}{4} \tag{5.59}$$

However, practically, the RF output power of the transistor consists of the fundamental frequency power, desired output power, and unwanted harmonic frequencies power. The output RF power for different harmonic frequencies can be obtained as:

$$P_{RF} = \sum_{n=1}^{\infty}\frac{1}{2}\left|V_{ds,nf_0} I_{ds,nf_0}\right| = \sum_{n=0}^{\infty}\frac{2V_{\max} I_{\max}}{(2n+1)^2\pi^2} \tag{5.60}$$

where V_{ds,nf_0} and I_{ds,nf_0} are the drain voltage and current amplitudes of the harmonic signals. Simply, the following equation can be obtained:

$$P_{RF} = \frac{V_{\max} I_{\max}}{4} \tag{5.60}$$

As a result, the load harmonic impedance network should be organized such that unwanted harmonic power is rejected, while the drain waveforms are controlled to keep the dissipated power in the transistor at zero. Consequently, all the DC power is transferred to the fundamental frequency output power, and the output power efficiency theoretically becomes 100%.

The switch-like behavior of the PA is administered by the gate biasing and the RF input signal. The fundamental and harmonic load impedances supplied for the drain of the transistor shape the drain waveforms, which define the different classes of switching-mode PAs. Switching-mode PAs can be categorized into two groups as:

- Single branch configurations, such as classes E and F, and inverse class F PAs; and,
- Balanced or push-pull configurations, such as voltage mode class D and current mode class D PAs.

5.3.2.1 Class D PA

The increased efficiency of class D PAs is a result of exploiting transistors as switches. In most cases, a class D PA implements a pair of active devices operating in a push-pull mode and a tuned output circuit. Switching-mode PAs with an output filter tuned to the fundamental frequency ideally transform all the DC power to fundamental frequency power, which can be delivered to the load without power losses at the harmonics. This results in a power output of $(8/\pi^2)V_{dd}^2/R$ for the transformer-coupled configuration. The output circuit is tuned to the switching frequency and ideally removes its harmonic components, which results in a completely sinusoidal signal that can be delivered to the load. Current is drained only through the transistor that is on, resulting in a 100% efficiency for an ideal class D PA [6], [10].

Class D amplifiers can be categorized into two groups as:

- Class D voltage-switching PAs (also designated as voltage mode class D, VMCD); and,
- Class D current-switching PAs (also designated as current mode class D, CMCD).

Figure 5.15 shows a circuit schematic of a voltage-switching push-pull class D PA, where C_b is the bypass capacitor, C_o is the blocking capacitor, and R is the load resistance. By using the input transformer, both transistors can be driven with currents that are 180° out of phase. This can be done by reversing one secondary winding on the transformer. The RF connection of the transistor output is in parallel configuration as a result of the grounding effect of a bypass capacitor, C_b. Hence, the equivalent load resistance is equal to $2R$ for each device.

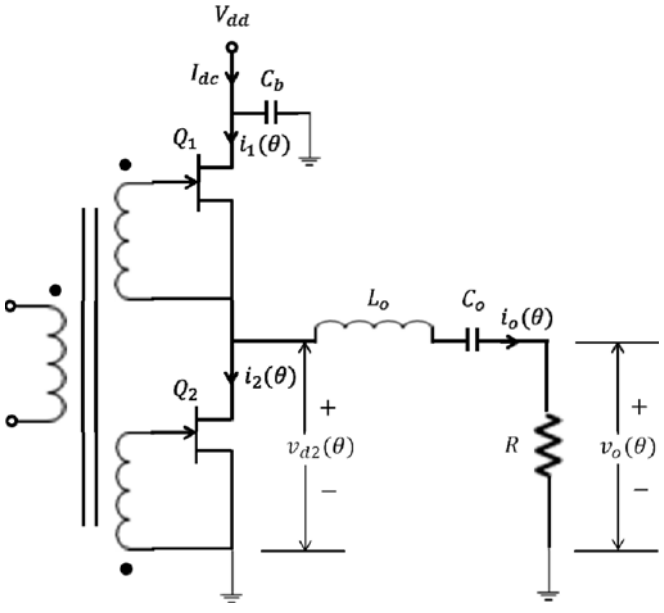


Fig. 5.15 Voltage-switching push-pull class D PA

The voltage and current waveforms are shown in Figure 5.16 for a voltage-switching push-pull class D PA. The square voltage waveform enables good utilization of the device's breakdown capability. The maximum amplitude of the half sinusoidal current through each transistor is $I_{\max} = \pi \cdot I_{dc}$.

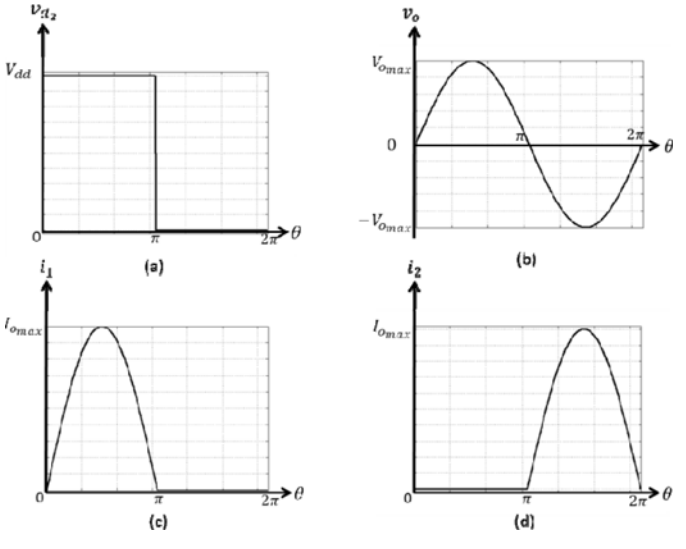


Fig. 5.16 Voltage-switching class D voltage and current waveforms

The Fourier series of current waveforms, $i_1(\omega t)$ and $i_2(\omega t)$, can be obtained as:

$$i_1(\omega t) = I_{dc}\left(1+\frac{\pi}{2}\sin(\omega t)-2\sum_{n=2,4,6,\ldots}^{\infty}\frac{\cos n\omega t}{n^2-1}\right) \tag{5.61}$$

$$i_2(\omega t) = I_{dc}\left(1-\frac{\pi}{2}\sin(\omega t)-2\sum_{n=2,4,6,\ldots}^{\infty}\frac{\cos n\omega t}{n^2-1}\right) \tag{5.62}$$

The Fourier series of the voltage at a connecting node can be obtained as:

$$v_{tot}(\omega t) = V_{dd}\left(\frac{4}{\pi}\right)\left(\sum_{n=1,3,5,\ldots}^{\infty}\frac{\sin n\omega t}{n}\right) \tag{5.63}$$

The current passing through the load is given by:

$$i_R(\omega t) = i_1(\omega t) - i_2(\omega t) = \pi \cdot I_{dc} \cdot \sin(\omega t) \tag{5.64}$$

Voltage-switching class D PAs are widely used in audio applications. However, they have been rarely used at frequencies in the range of 1 GHz or higher. This is because of the losses associated with the transition time, parasitic reactance and turn-on resistance of the transistors. At high frequencies, the losses in the output parasitic reactance, mainly the shunt capacitance, C_{ds}, dominates the loss due to the turn-on resistance. The energy, E, dissipated within C_{ds} per switching cycle can be obtained by:

$$E = \frac{1}{2}C_{ds} \cdot V_{dd}^2 \tag{5.65}$$

where V_{dd} is the voltage across the transistor at switch closure [11].

To overcome this limitation, current-switching class D PAs have been proposed. In this configuration, a shunt LCR (inductor, resistor, capacitor) output filter is implemented. In fact, the current waveform becomes a voltage waveform and vice versa. A series connection transforms to a parallel connection and vice versa. Furthermore, a voltage-source supply to transformer center-tap converts to a RF-choke feed (current source) to transformer center-tap.

The use of the current-switching class D PA cancels out the losses in the shunt capacitance by achieving a zero voltage switching. Therefore, current-switching class D amplifiers are more appropriate for high-frequency applications resulting in higher power efficiency.

Figure 5.17 shows a current-switching class D PA, in which the two devices are connected in parallel. An input transformer is required to drive the devices out-of-phase, and also an output transformer is required to extract the differential load

voltage. As demonstrated in Figure 5.18, the currents through each transistor have the shape of square waveforms. In order to have proper switching, large input power is required. Voltages v_1 and v_2 have the shape of half-sinusoidal waveforms. The maximum amplitude of the half–sinusoidal voltage waveforms is $V_{\max} = \pi \cdot V_{dd}$, and high breakdown voltage is required.

The Fourier series of voltage waveforms, $v_1(\omega t)$ and $v_2(\omega t)$, can be calculated as:

$$v_1(\omega t) = V_{dd}\left(1 + \frac{\pi}{2}\sin(\omega t) - 2\sum_{n=2,4,6,\ldots}^{\infty}\frac{\cos n\omega t}{n^2 - 1}\right) \tag{5.66}$$

$$v_2(\omega t) = V_{dd}\left(1 - \frac{\pi}{2}\sin(\omega t) - 2\sum_{n=2,4,6,\ldots}^{\infty}\frac{\cos n\omega t}{n^2 - 1}\right) \tag{5.67}$$

The Fourier series of the RLC tank current waveform is given by:

$$I_{tot}(\omega t) = I_{dc}\left(\frac{4}{\pi}\right)\left(\sum_{n=1,3,5,\ldots}^{\infty}\frac{\sin n\omega t}{n}\right) \tag{5.68}$$

The voltage across the load is obtained as:

$$v_{load}(\omega t) = v_1(\omega t) - v_2(\omega t) = \pi \cdot V_{dd} \cdot \sin(\omega t) \tag{5.69}$$

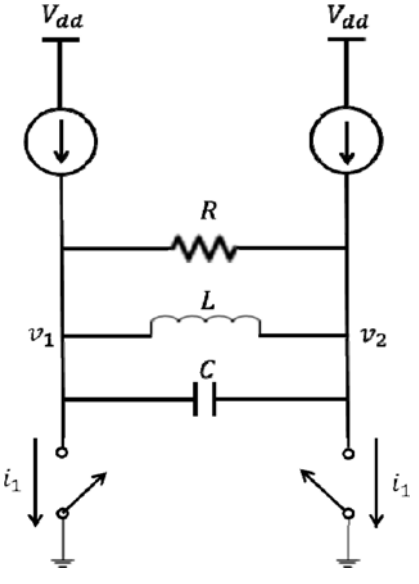


Fig. 5.17 Current-switching class D PA

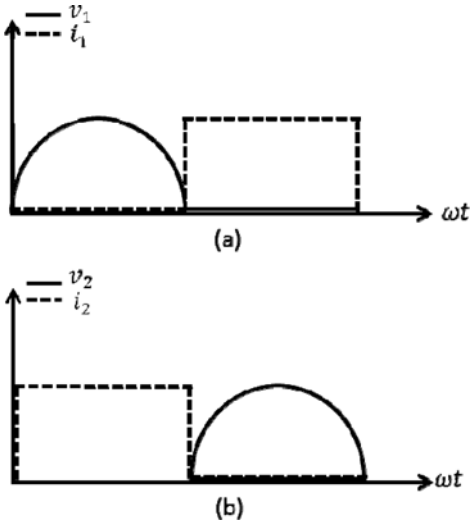


Fig. 5.18 Current-switching class D voltage and current waveforms

5.3.2.2 Class E PA

In a class E PA, a single transistor operates as a switch. The drain voltage waveform results from the sum of the DC and RF currents charging the drain current capacitance. In an ideal class E PA, the drain voltage drops to zero and has a zero slope when the transistor turns on (zero voltage switching). This eliminates the losses associated with charging the drain capacitance in class D PAs and considerably decreases the switching losses [12].

In order to optimize the efficiency, the transient response of the load network has to be controlled even at the time the switching speed is a considerable fraction of the RF cycle. Indeed, class E PAs can be designed to operate at frequencies higher than those of class D PAs. Class E PAs can be designed at frequency bands up to 3.5 GHz, while attaining high power efficiency.

In a switching-mode class E PA, it is possible to have simultaneous high voltage and high current during the switching between on and off. Therefore, a switching-mode PA with resistive load has power dissipation during switching transitions. Class E can have good efficiency only if switching times are much smaller than the waveform period. Figure 5.19 shows current and voltage waveforms in a square-wave switching-mode class E amplifier with resistive load.

Switching power dissipation at every instant of time can be obtained as the multiplication of the voltage and current in the active part of the transistor. The current and voltage must rise and fall, but not at the same time. In other words, when the voltage rises, the current must fall and vice versa. Furthermore, time-displace voltage and current transitions never have high current and high voltage at the same time. In a class E PA, switching has to be performed as fast as possible with a reasonable input drive. A class E PA has been utilized for high-efficiency amplification at frequencies up to K-band [6].

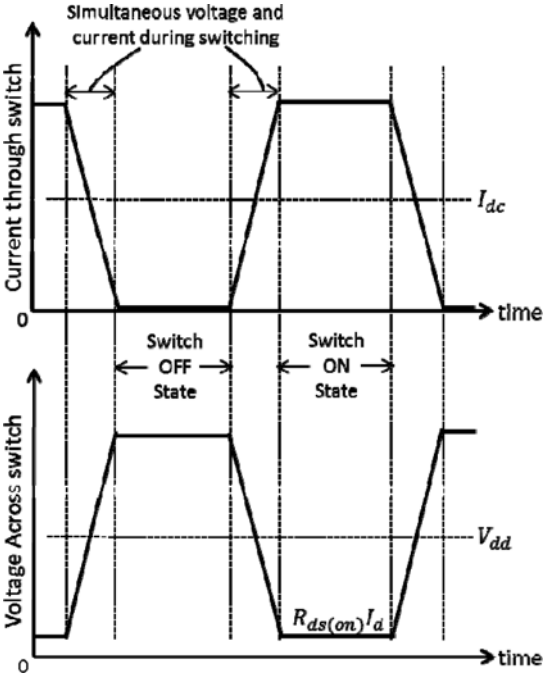


Fig. 5.19 Current and voltage waveforms in a square-wave switching-mode class E PA

5.3.2.3 Class F and Inverse Class F PAs

A class F PA is characterized by a load network that has resonances at one or more harmonic frequencies as well as at the carrier frequency. Class F amplifiers are capable of high efficiency (88.4% for traditionally defined class F or 100% if infinite harmonic tuning is used), which results from a low DC voltage current product. In other words, the drain voltage and current are shaped to minimize their overlap region.

Class F amplifier design is challenging mainly due to the complex design of the output-matching network [6]. At microwave frequencies, it becomes difficult to fabricate class F amplifiers, as capacitors and inductors function poorly; and, planar structures becomes hard to realize, especially when tuning multiple harmonics.

If the even harmonics are terminated to short circuits and the odd harmonics are terminated to open circuits, the current and voltage waveforms have the shapes of half-sinusoidal and square waveforms, respectively. The Fourier series of half-sinusoidal current and square voltage waveforms, both at the same angular velocity, ω, are given by:

$$i(\omega t) = I_{dc}\left(1 - \frac{\pi}{2}\sin(\omega t) - 2\sum_{n=2,4,6,\ldots}^{\infty}\frac{\cos n\omega t}{n^2 - 1}\right) \tag{5.70}$$

$$v(\omega t) = V_{dc}\left(1+\frac{4}{\pi}\sin(\omega t)+\frac{4}{\pi}\sum_{n=3,5,7,\ldots}^{\infty}\frac{\sin n\omega t}{n}\right) \tag{5.71}$$

where I_{dc} and V_{dc} are the DC current and voltage components, respectively [13].

This class of operation is called class F or voltage mode class F. The voltage and current waveforms of the theoretical class F PA is demonstrated in Figure 5.20 (a). The power dissipation for this class of operation, which is equal to the product of the current and voltage, is zero. Therefore, a class F PA is able to achieve 100% efficiency in theory.

Inverse class F or current mode class F PA operation is obtained by terminating the odd harmonics to short circuits and the even harmonics to open circuits. In this case, the current and voltage waveforms have the shapes of square and half-sinusoidal waveforms, respectively. The Fourier series of current and voltage waveforms are in the form of:

$$i(\omega t) = I_{dc}\left(1+\frac{4}{\pi}\sin(\omega t)+\frac{4}{\pi}\sum_{n=3,5,7,\ldots}\frac{\sin n\omega t}{n}\right) \tag{5.72}$$

$$v(\omega t) = V_{dc}\left(1-\frac{\pi}{2}\sin(\omega t)-2\sum_{n=2,4,6,\ldots}\frac{\cos n\omega t}{n^2-1}\right) \tag{5.73}$$

The voltage and current waveforms of the theoretical inverse class F are shown in Figure 5.20 (b). As in class F operation, the power dissipation of inverse class F operation is zero; therefore, this class is also capable of achieving 100% efficiency in theory [13].

However, due to some limitations, the theoretical 100% efficiency of class F and inverse class F PAs is not possible in practice. In general, an infinite number of serial odd-harmonic tank resonators can provide a half-sinusoidal current waveform and a square voltage waveform. Figure 5.21 (a) shows such an output-matching network with a multiple-resonator output filter that tunes the harmonics to the class F operational conditions. All the even harmonics are tuned with one parallel tank resonating at the fundamental frequency. Likewise, the harmonic matching for the inverse class F operation needs an infinite number of serial even-harmonic tank resonators and one parallel tank resonating at the fundamental frequency. The harmonic matching network for the inverse class F is shown in Figure 5.21 (b). Indeed, such a harmonic matching circuit is not feasible in practice.

Berini et al. [14] and Raab [15] showed that the significant effect on the waveform shaping and, in fact power efficiency, is only related to the first few harmonics. Tuning more harmonics slightly enhances the efficiency of the PA at the cost of increasing design complexity and output matching losses. In addition to the matching networks' complexity and losses, the intrinsic and extrinsic effects of the transistor impact the maximum efficiency that can be achieved.

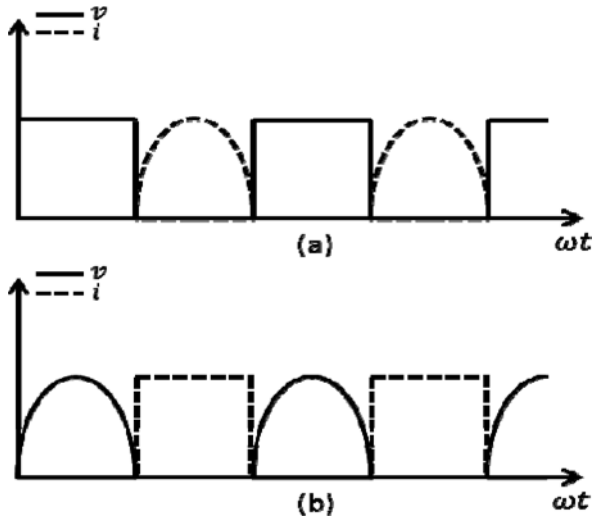


Fig. 5.20 Ideal waveforms in switching mode PAs: a) class F and b) inverse class F

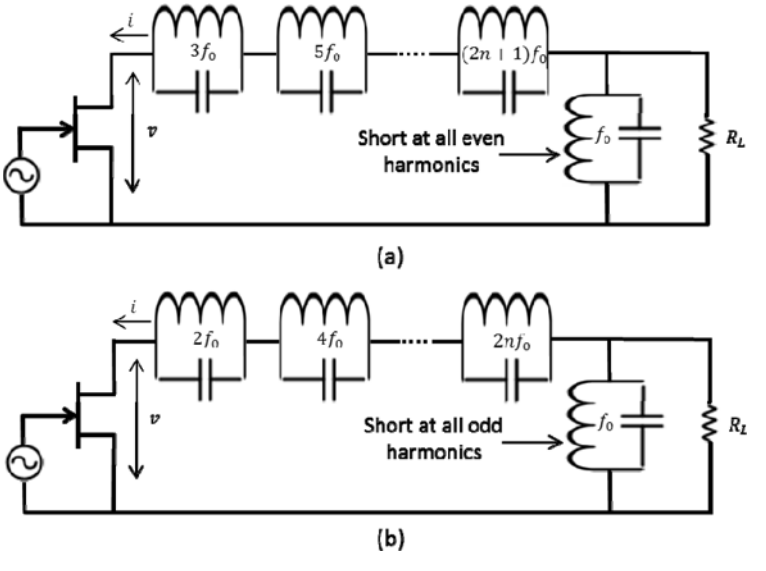


Fig. 5.21 Basic circuits of PAs: a) class F and b) inverse class F

5.3.3 Comparison of the Classes of Operation in RF PAs

Table 5.1 compares the performance of the different classes of operations described in the preceding sections, in terms of output power, gain, efficiency, and linearity. Table 5.1 shows that the efficiency of the linear amplifiers decreases from class A to class C. However, the high linearity of the class A PA trends to the high nonlinearity of the class C PA when moving from class A to class C. In the case of switching-mode PAs, class D and E PAs have very high efficiencies, but they are strongly nonlinear. The class F PA also has very high efficiency and is highly nonlinear.

Table 5.1 Performance Comparison for Different Class of Operations of PAs

	Class A	Class AB	Class B	Class C	Class D	Class E	Class F
Output power	++++	+++	++	+	++	++	+++
Gain	++++	+++	++	+	++	++	+++
Efficiency	+	++	+++	++++	++++	++++	+++
Linearity	++++	+++	++	+	+	+	++

5.4 Linearization of RF Power Amplifiers

The nonlinear behavior of the RF front-end, especially RF transmitters, can significantly degrade the overall performance of wireless systems. The power efficiency of an RF amplifier is optimal when it is operated near saturation [16]. An amplifier operating in this nonlinear range generates IM distortion that interferes with neighboring channels. Therefore, there should be compensation for the nonlinearities and distortions of the RF transmitter.

An efficient PA design reduces the cost of power consumption and increases the battery life of wireless mobile transmitters. However, the increase in efficiency is usually accompanied by linearity deterioration, which requires a trade-off between efficiency and linearity.

The output power, gain and efficiency variations for a typical PA as functions of the input power are shown in Figure 5.22. As can be observed from this figure, there are two major operational regions for the amplifier. In the first region, the amplifier has a linear gain, but the efficiency is low. In the second region, the amplifier has high efficiency, but the gain is nonlinear. Linearity deterioration is caused by the distortions introduced to the signal by the compression of the PA in the saturation region. As a result, the linearity/nonlinearity of the PA depends on the input signal power [17].

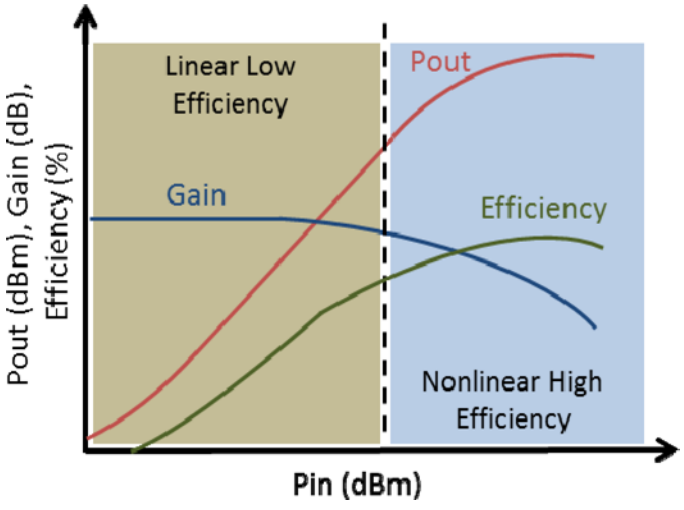


Fig. 5.22 Output power, gain and efficiency of a PA versus input power

Linearization is a systematic approach to reduce an amplifier's distortion and is inevitable for enhancing the linearity of an amplifier to the high input power drive levels and achieving linearity requirements when operating the device over its entire power range [18]. Linearization allows a PA to generate more power and operate with higher efficiency for a given level of distortion. There are different methods for linearizing an RF amplifier. The three major linearization techniques are feedback, feedforward, and predistortion [19].

5.4.1 Feedback Linearization

The feedback linearization technique is a closed-loop system that can provide high levels of linearization. It is based on the concept of the feedback loop, which is widely used in control theory. As shown in Figure 5.23, the amplifier's output signal is fed back and subtracted from the amplifier's input signal to force the output signal to be a linear replica of the input signal. The use of feedback linearization has been widely investigated, but has received little use at RF frequencies. The most important reason is probably related to issues with amplifier stability and the difficulty in production of networks with non-ideal components that function over wide frequency ranges [19].

Indirect feedback techniques are reported to be more widely used. The correction in two alternative indirect feedback strategies, Cartesian feedback and polar feedback, are processed in a baseband domain. Two loops are needed for the amplitude and phase. If only one loop is used for the amplitude at lower frequencies, the technique is also called an automatic gain control (AGC).

There are some limitations with feedback linearization techniques that make them unsuitable for wireless communication applications. The gain of the linearized amplifier is reduced to the gain of feedback loop. Furthermore, the delays associated with the feedback loop must be small enough to ensure stability, which limits the use of feedback linearization techniques to narrowband signals [20].

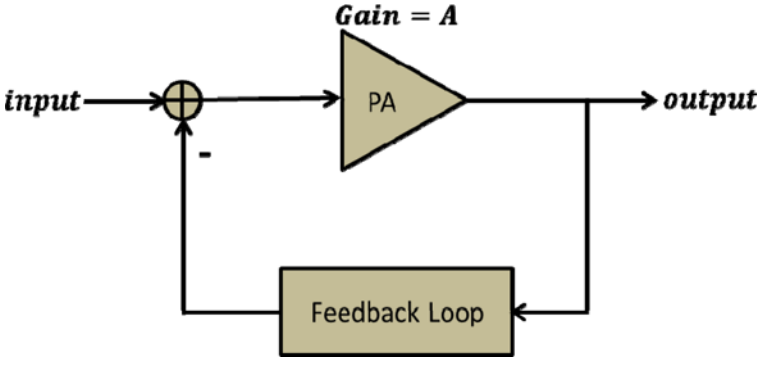


Fig. 5.23 Feedback linearization block diagram

The frequency bandwidth of feedback linearization is limited by the inverse of the propagation delay of the feedback loop. For third-order distortion correction, the bandwidth of the loop must be three times the bandwidth of the baseband envelope. Moreover, at small signals, the system becomes inefficient, and the modulator/demodulator has a limited dynamic range. Operational amplifiers and demodulators introduce noise in the system. Stability, noise and orthogonality of the modulator/demodulator can be improved if the loop is digitized with a digital signal processor (DSP). However, current DSPs cannot reach the speed of the equivalent analog systems.

Different approaches have been proposed to address the limitations, including envelope feedback, polar feedback, Cartesian feedback, adaptive double envelope feedback, and digital Cartesian feedback. In envelope feedback linearization, input and output samples are fed into two envelope detectors. As illustrated in Figure 5.24, the error between the outputs of the envelope detectors is injected to the main signal path via a vector modulator. The use of this error signal to control the amplifier's output only allows the compensation for AM/AM nonlinearities of the PA, not its AM/PM nonlinearities.

Hence, a second feedback loop is added in polar feedback linearization, which is shown in Figure 5.25. In polar feedback linearization technique, the addition of phase lock loop (PLL) to the envelope feedback linearizer allows the compensation for both AM/AM and AM/PM nonlinearities of the PA.

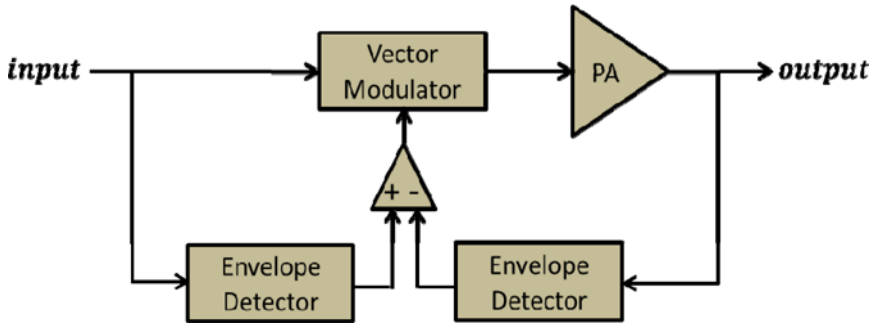


Fig. 5.24 The envelope feedback linearization technique

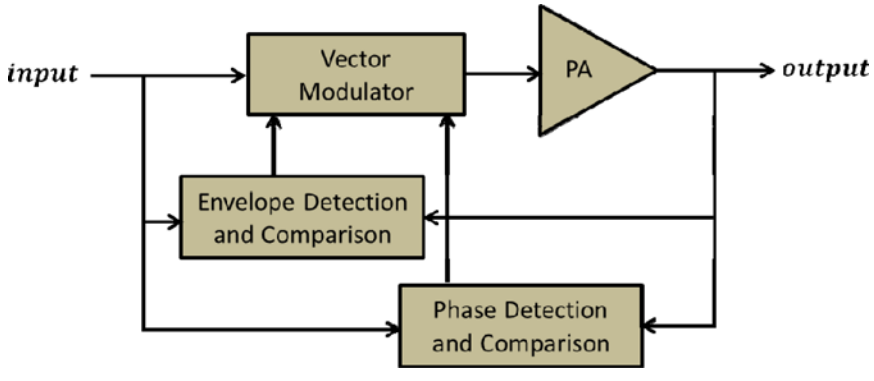


Fig. 5.25 The polar feedback linearization technique

An alternative approach is Cartesian feedback, which separates the signal into in-phase and quadrature components. In this technique, the need for phase-shift components is eliminated; and, correction of gain and phase is still possible by tuning the amplitudes of two orthogonal components. The baseband in-phase and quadrature components are compared to control the attenuators in the vector demodulator [19].

5.4.2 Feedforward Linearization

Feedforward linearization is a powerful technique that has the same fundamental error-correcting operation as feedback, but removes the inherent bandwidth and stability problems of a conventional feedback loop at RF frequencies. The feedforward technique is theoretically correct for all nonlinearity orders.

In feedforward linearization, a correction signal is injected at the output of the PA to cancel its nonlinearity distortions. Figure 5.26 demonstrates the block

diagram of the feedforward linearization technique. As can be observed from this figure, there are two loops in the feedforward system, which are the signal cancellation loop and the distortion cancellation loop [6]. In the signal cancellation loop, a portion of the PA's output signal is compared to a time-, amplitude- and phase-aligned replica of the input signal. The signal at the output of the comparator represents the additive distortion products at the output of the PA. These products, after being amplified to their original magnitudes, are then subtracted from the time-aligned version of the PA output, which theoretically leads to perfect cancellation of the distortions generated by the PA [2].

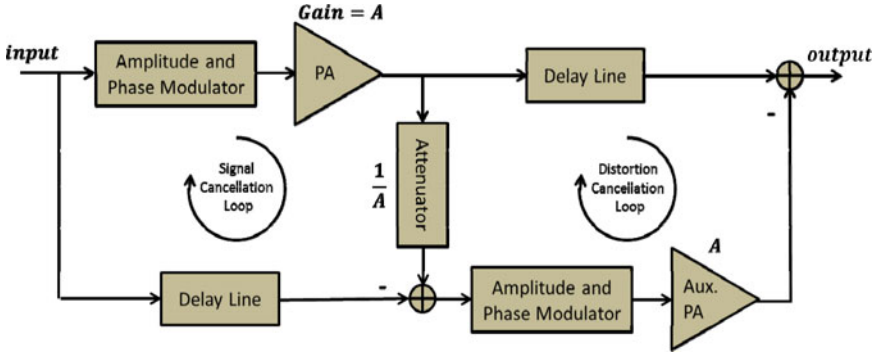


Fig. 5.26 Feedforward linearization block diagram

The feedforward linearization technique is naturally sensitive to changes in operating conditions. However, the development of adaptation methods to compensate for such changes has increased the interest in the technique. Due to the phase shifts introduced in the carrier by the amplifiers, the feedforward linearizers are very sensitive to the alignment in both loops [16]. Indeed, delay lines must be precisely tuned in order to achieve desirable performance. Furthermore, feedforward linearization is intrinsically non-adaptive and requires adaptive control techniques in both loops to retain the proper alignment. The gains of the attenuator and auxiliary PA must be very well matched to the gain of the PA. Furthermore, the delay line and subtractor in the output path of the PA have to be low loss to achieve an effective linearization.

Feedforward linearizers are commonly used in wireless communication base stations. In these linearizers, very high linearity can be achieved (50 dBc or higher), and wide bandwidth can be supported (40-60 MHz or more). Power efficiency is the major drawback of this technique. Although the main PA is operating at high efficiency, the overall efficiency of the linearized amplifier is considerably decreased by the error amplifier, which needs to be perfectly linear. The overall efficiencies of third-generation (3G) multicarrier PAs linearized by the feedforward technique are in the range of 10%-15%.

5.4.3 Digital Predistortion Linearization

The predistortion can be performed in the analog or digital domain and on baseband, intermediate frequency (IF) or RF signals. Digital baseband predistortion, which is an open-loop technique, is currently the preferred linearization technique and is widely used for applications up to 20 MHz bandwidth. In a digital predistortion (DPD) technique, a complementary nonlinearity upstream of the PA is applied, so that the cascade of the digital predistorter and the PA behaves as a linear amplification system. Figure 5.27 presents a simplified block diagram of the DPD linearization technique [4]. DPD achieves high linearity performance with high power efficiency. The efficiencies of 3G PAs linearized by digital predistorters are reported in the range of 30-45%.

Behavioral modeling of RF PAs is an essential task in the design of high-performance wireless transmitters that combine high power efficiency and spectrum compliant linearity performance. This is even more important in modern communication systems that employ envelope-varying signals with high PAPRs. Indeed, such signals set rough linearity requirements on the PA and usually lead to the forfeit of the system's power efficiency, in order to meet the linearity requirements [1].

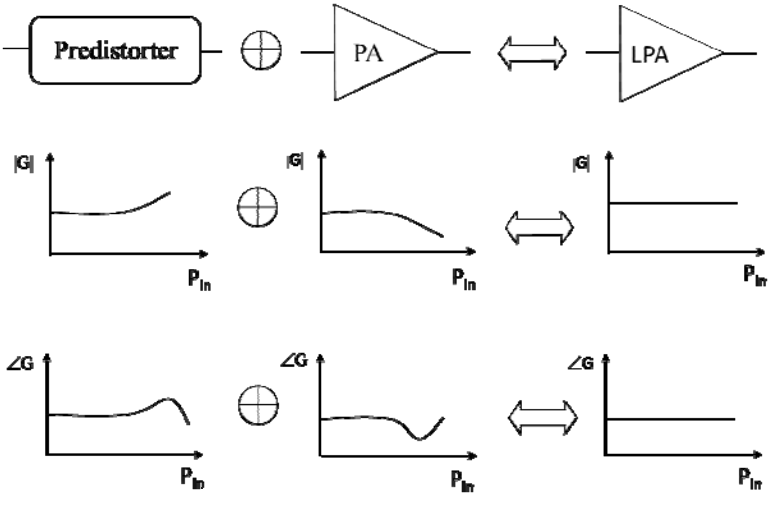


Fig. 5.27 Predistortion linearization technique

The most important advantage of behavioral modeling is that it does not require extensive knowledge of the RF circuit physics and functionality. As demonstrated in Figure 5.28, behavioral modeling simplifies the modeling of the RF circuit to the identification of a mathematical formulation that relates the input and output of the device under test (DUT) that can be considered as a black box. Consequently,

behavioral modeling appears as a time and resource efficient process for transmitter performance evaluation and digital predistorter design [4].

Due to current broadband and highly varying signals, such as code-division multiple-access (CDMA), wideband CDMA (WCDMA), andorthogonal frequency-division multiplexing (OFDM), PAs or wireless transmitters have to be considered as dynamic nonlinear systems. Therefore, an appropriate architecture has to be chosen to extract and identify an accurate and robust forward or reverse model for such systems [1].

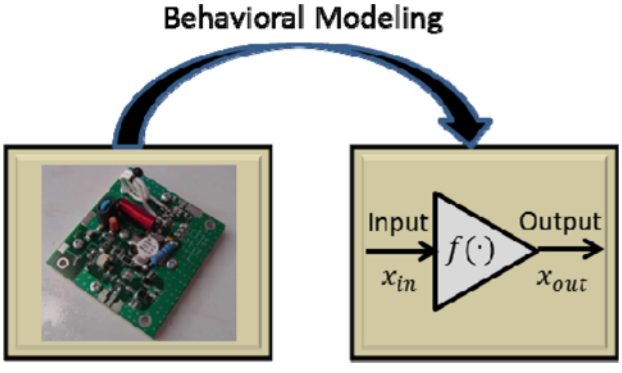


Fig. 5.28 Black-box based behavioral modeling

The block diagram of a DPD-based linear amplification system is shown in Figure 5.29. The signal at the input of the digital predistorer ($x_{in\_DPD}(n)$) can be derived from the signal at the output of the PA block according to:

$$x_{in\_DPD}(n) = \frac{x_{out}(n)}{G} \tag{5.74}$$

The behavioral modeling of the DUT identifies function f_{DUT} that satisfies $f_{DUT}(x_{in}(n)) = x_{out}(n)$. The synthesis of the DPD function is equivalent to the estimation of function f_{DPD} , such that $f_{DPD}(x_{out}(n)/G) = x_{in}(n)$.

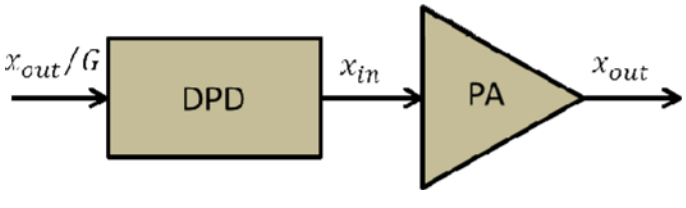


Fig. 5.29 Block diagram of digital predistortion (DPD) based linear PA

Predistorters potentially can handle much wider modulation bandwidths, including multicarrier signals. Unlike closed-loop systems, they do not have stability problems. Predistorters can be integrated in monolithic microwave integrated circuits (MMICs). They provide good linearity improvement in traveling tube wave amplifiers (TWTAs), but less significant improvement for metal semiconductor field-effect transistor (MESFET) classes A and AB. Predistortion linearization techniques have a smaller dynamic range than feedforward linearization techniques. In addition, predistorters require temperature compensation or an adaptive control mechanism.

Numerous formulations have been proposed for behavioral modeling and DPD of RF PAs and transmitters. Some commonly used formulations are memoryless look-up table, nested look-up table, Volterra, memory polynomial, forward, reverse and parallel twin nonlinear two-box, envelope memory polynomial, Wiener and augmented Wiener, and Hammerstein and augmented Hammerstein models. The look-up table and memory polynomial models are described in the following sections. More information about the other models can be found in [4].

5.4.3.1 Look-Up-Table Model

The static look-up table (LUT) model is the basic behavioral model for memoryless AM/AM and AM/PM nonlinearities. The complex gain of the DUT is stored in two look-up tables. The output waveform can be shown as:

$$x_{out}(n) = G\left(\left|x_{in}(n)\right|\right) \cdot x_{in}(n) \tag{5.75}$$

where $G\left(\left|x_{in}(n)\right|\right)$ is the instantaneous complex gain of the DUT.

The AM/AM and AM/PM characteristics of the DUT can be obtained from the raw measured data using averaging or polynomial fitting techniques.

5.4.3.2 Memory Polynomial Model

The memory polynomial model is widely used for behavioral modeling and DPD of PAs and transmitters exhibiting memory effects. The output waveform of memory polynomial model is given by:

$$x_{out}(n) = \sum_{j=0}^{M} \sum_{i=1}^{N} a_{ji} \cdot x_{in}(n-j) \cdot \left|x_{in}(n-j)\right|^{i-1} \tag{5.76}$$

where N and M are the nonlinearity order and the memory depth of the DUT, respectively; and, a_{ji} are the model coefficients.

Figure 5.30 demonstrates typical spectrum results for a linearized Doherty amplifier using a memory polynomial based digital predistorter. This figure clearly presents the perfect cancellation of the nonlinearities introduced by the DUT.

There are other variations of the memory polynomial model in the literature, including the orthogonal memory polynomial model and the memory polynomial model with cross-terms, which is also called as the generalized memory polynomial model [4].

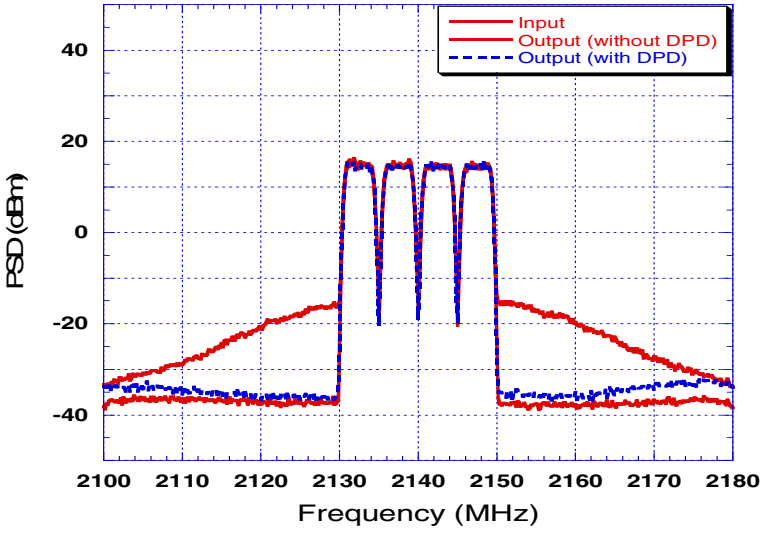


Fig. 5.30 Performance of a DPD linearized 30-Watt GaN Doherty PA with four-carrier WCDMA signal (PAPR = 11.25 dB) [4]

5.4.4 Analog Predistortion Linearization

RF linearizers based on predistortion have been used for several years in satellite transponder PAs, either as solid-state power amplifiers (SSPAs) or TTWTAs, to correct the nonlinear characteristics. The use of predistorters together with the PAs on board a satellite allows for the operation of its transponders at reduced distortion levels and higher power efficiency, while carrying multicarrier traffic [21]. Analog predistorters (PDs) are still in use for high power applications in the upper GHz frequency bandwidths, where medium linearization performance is targeted. PDs have less complexity, low implementation cost, wide bandwidth and the capability to be added to existing PAs as a separate stand-alone component. Applications include mobile and handset PAs, which usually requires a simple PD network, typically a diode gain amplifier [7].

A linearizer based on a series diode with a parallel capacitor is a very simple structure that utilizes the nonlinearity of the series resistance of the diode, which produces a characteristic positive gain and negative phase with increasing input power. This is a simple linearizer, but requires an additional isolation mechanism between the linearization circuit and the PA. Furthermore, it has very limited control on the achieved characteristics and, hence, has very limited practical applications [22].

The most basic linearization circuit with diodes is shown in Figure 5.31 (a) in a head-tail configuration, which consists of semiconductor diodes as a nonlinear resistor. This setup suffers from linear phase distortions in the AM/PM characteristic

of a PA. The linearizer can be tuned to resonate at a center frequency to mitigate this effect, but the linearizer is a nonlinear circuit. Therefore, the resonance cannot be achieved for the full range of input signals. Furthermore, at microwave frequencies, the problem is even more severe, as the capacitive effects of the junction capacitance of the Schottky diodes limit the application of this linearizer to a very moderate range of input signals. This has been addressed in [22].

To overcome this problem, the resonance can be attained by utilizing a suitable inductance, L_s in shunt, as illustrated in Figure 5.31 (b). By utilizing this shunt inductance, the problem of phase distortion can be reduced for any input power and any bias voltage, V_{sch}, of Schuttky diodes D_1 and D_2. The actual bias voltage V_{PK} across the Schottky diodes D_1 and D_2 is regulated by the bias feed resistance, R_b, and the input RF power. The PIN diode, D_3, delivers a variable resistance in shunt to the biased Schottky diodes D_1 and D_2. D_3 provides a dynamic control to achieve the required gain and phase characteristics and also provides an improved dynamic range in the gain and phase responses, V_{pin}. The actual bias of the PIN diode is regulated by the bias feed resistance, R_{b1}, and the input RF power.

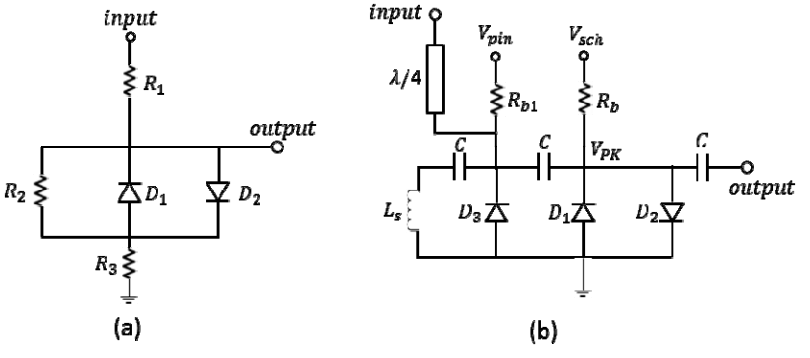


Fig. 5.31 (a) The basic linearization circuit with diodes, (b) an analog linearization circuit

PAs may have single or dual inflection points in their distortion characteristics based on device technology or bias conditions, as shown in Figure 5.32. A diode-based RF PD linearizer that can be used to linearize PAs with dual inflection points in their distortion characteristics is proposed in [23]. The block diagram of the suggested PD is shown in Figure 5.33. Two linearizer branches connected to through and coupled ports of a branch line hybrid coupler. Each linearizer is composed of two Schottky diodes connected in an anti-parallel configuration as well as a PIN diode in parallel with them. The key point in the design of the PD is the cancellation of the imaginary part of the admittance seen from connection points of the diodes in small-signal operation.

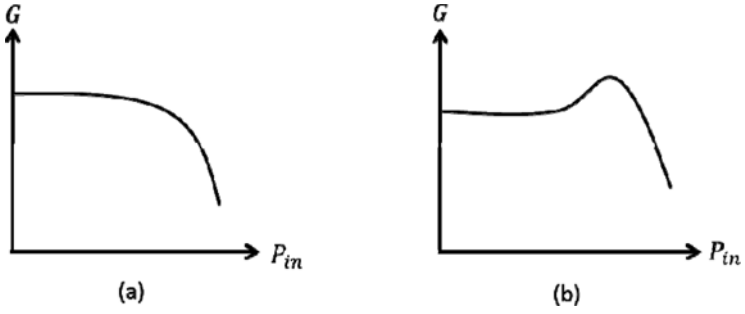


Fig. 5.32 (a) Single and (b) dual inflection amplitude characteristics

Each linearizer generates a nonlinear reflection coefficient, Γ, that relates the input signal, a_1, to the output signal, b_4, using the matrix of the hybrid coupler as follows:

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & -j & 0 \\ 1 & 0 & 0 & -j \\ -j & 0 & 0 & 1 \\ 0 & -j & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ \Gamma \cdot b_2 \\ \Gamma \cdot b_3 \\ 0 \end{bmatrix} \tag{5.77}$$

where $a_2 = \Gamma \cdot b_2$ and $a_3 = \Gamma \cdot b_3$.

Further manipulation of the above matrix results in:

$$b_4 = -\Gamma \cdot a_1 \tag{5.78}$$

Considering the input power $|a_1|^2$ and output power $|b_4|^2$, the following relationship can be derived:

$$P_{out} = |\Gamma|^2 \cdot P_{in} \tag{5.79}$$

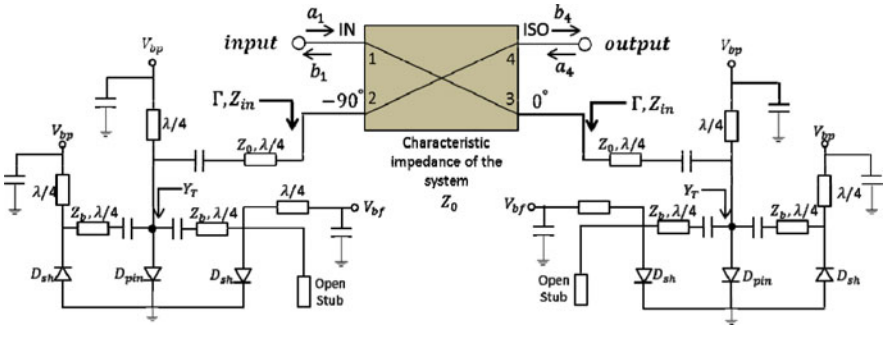


Fig. 5.33 Analog predistorter architecture for linearization of the distortion with dual inflection point

This equation shows that the PD gain is equal to the square value of the reflection coefficient seen from both linearizer branches.

Figure 5.34 shows the performance of the PD in suppressing the distortion of the PA and flattening of the nonlinear part of the AM/AM characteristic of the PA. It is obvious that the distortion is considerably compensated for by the PD.

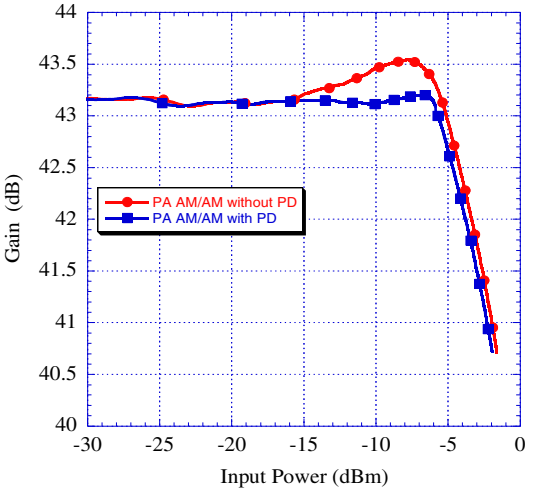


Fig. 5.34 Measured AM/AM of AP603 with PD and without PD versus the input power [23]

5.4.5 Comparison of Linearization Techniques

Table 5.2 compares the performance of feedback, feedforward, and predistortion linearization techniques, in terms of frequency bandwidth, linearity, complexity, power efficiency, and adaptation.

Table 5.2 Performance Comparison of Different Linearization Methods

	Feedback	Feedforward	Analog / Digital Predistortion
Frequency Bandwidth	Narrow	Wide	Ultra/Medium
Linearity	Good	Ultra	Medium/Ultra
Complexity	Medium	High	Medium/Medium
Power Efficiency	High	Low	High/High
Adaptation	Intrinsically adaptive	Intrinsically non-adaptive	Intrinsically non-adaptive

5.5 RF Transmitter Architectures

Power efficient and linear and linear transmitters, such as polar, linear amplification with nonlinear components (LINC), envelope elimination and restoration (EER) and delta-sigma transmitters have received increased attention. These advanced transmitter architectures are theoretically highly efficient and linear, even for sophisticated digital modulation signals with non-constant envelopes and large PAPRs.

5.5.1 Polar Transmitter Architecture

In the polar transmitter, the in-phase (I) and quadrature (Q) baseband signals are transformed from the Cartesian representation to the polar representation. The Cartesian domain signal representation, $S(t) = I(t)\cos\omega_c t + Q(t)\sin\omega_c t$, can be transformed to the polar domain as $S(t) = A(t)\cos(\omega_c t + \varphi(t))$. The transformation can be done with the following equations:

$$A(t) = \sqrt{I^2(t) + Q^2(t)} \tag{5.80}$$

$$\varphi(t) = arct\tan(Q(t)/I(t)) \tag{5.81}$$

A major difference between the Cartesian and polar representations is that the polar basis vectors (amplitude and phase) have widely differing spectral properties compared to the I/Q basis vectors and the modulated RF output. This is clear in (5.80) and (5.81) as $A(t)$ and $\varphi(t)$ are derived from $I(t)$ and $Q(t)$ through nonlinear operations. The nonlinearities in (5.80) and (5.81) cause the polar basis vectors to lose the bandlimited property of the Cartesian I/Q representation. To design the system for high integrity, low EVM and high spectral fidelity, the wideband nature of the amplitude and phase paths has to be considered in the architecture and circuit-level design [24].

In a polar transmitter, the phase, $\varphi(t)$, and amplitude, $A(t)$, components of the signal are processed separately. Figure 5.35 shows a general structure of the polar transmitter. The phase of the carrier is modulated, and it then passes through an amplifier, in which the value of the power supply varies proportionally to the envelope. The PA is shown as a variable-gain amplifier (VGA) in Figure 5.35.

Generally, any signal that modulates the amplitude of the carrier can be employed in a polar transmitter. The phase information extracted from the original signal with a constant or non-constant envelope is converted to a constant envelope signal. This can be achieved by phase modulation with the help of a phase lock loop (PLL) or digital phase lock loop (DPLL) to output the desired transmitting frequencies [25]. The output signal can now be amplified without any concern of distorting the amplitude information.

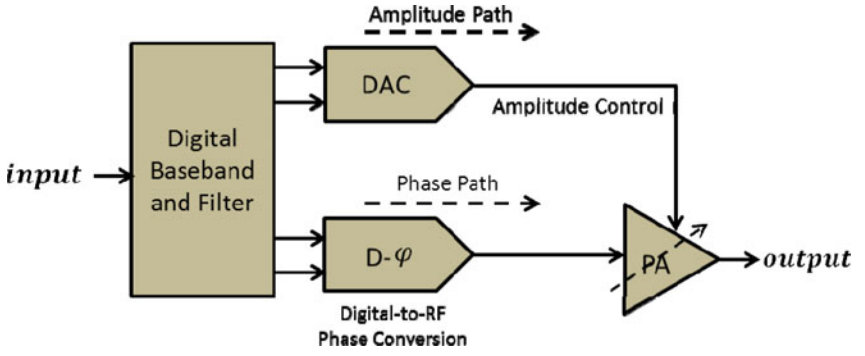


Fig. 5.35 General structure of the polar transmitter

5.5.2 LINC Transmitter Architecture

The principle of the linear amplification with nonlinear components (LINC) amplification system is based on converting the highly varying amplitude modulation of the input signal, $S_{in}(t)$, into phase modulations of two constant envelope signals, $S_1(t)$ and $S_2(t)$, which can be obtained using the following equations:

$$S(t) = r(t) \cdot e^{j \cdot \phi(t)} = S_1(t) + S_2(t) \tag{5.82}$$

$$r(t) = r_{\max} \cdot \cos(\theta(t)) \tag{5.83}$$

hence,

$$\begin{cases} S_1(t) = \dfrac{r_{\max}}{2} \cdot e^{j(\phi(t)+\theta(t))} \\ S_2(t) = \dfrac{r_{\max}}{2} \cdot e^{j(\phi(t)-\theta(t))} \end{cases} \tag{5.84}$$

where $r_{\max}$ represents the maximum of $r(t)$, $\phi(t)$ is the phase of the baseband signal, and $\theta(t)$ is the additional phase modulation angle related to the amplitude of signal, $r(t)$, which is given by:

$$\theta(t) = \arccos\left(\frac{r(t)}{r_{\max}}\right) \tag{5.85}$$

The resultant signal, $S_{out}(t)$, is a linearly amplified version of the input signal, $S_{in}(t)$:

$$S_{out}(t) = \sqrt{2} \cdot G \cdot S_{in}(t) \tag{5.86}$$

where G is the gain of the branch amplifier.

Another way to compute $S_1(t)$ and $S_2(t)$ can illustrated as:

$$\begin{cases} S_1(t) = \dfrac{1}{2} S(t)\left[1 + j \cdot e(t)\right] \\ S_2(t) = \dfrac{1}{2} S(t)\left[1 - j \cdot e(t)\right] \end{cases} \tag{5.87}$$

where $e(t)$ is obtained by:

$$e(t) = \sqrt{\frac{r_{\max}^2}{r^2(t)} - 1} \tag{5.88}$$

Considering that $S_1(t)$ and $S_2(t)$ have constant envelopes, they can be efficiently amplified by means of power efficient or switching-mode nonlinear PAs. The two amplified signals are then combined to retrieve a linearly amplified version of the original amplitude modulated input signal. Therefore, RF PAs can be operated at saturation, which results in maximum power efficiency.

The structure of a LINC amplification system is shown in Figure 5.36. A vector representation of the separated baseband components is also given in Figure 5.37. There are three main components in the LINC transmitter: the signal separator, the nonlinear amplifiers, and the signal combiner [5].

Successful implementation of a LINC transmitter is highly dependent on the precision and control of signal separation, because the LINC architecture is sensitive to amplitude and phase balance. Although there are some analog signal component separation implementations, digital signal component separation implementation offers the flexibility required to apply a control over the precision of the calculations. A digital signal separator also allows implementation of correction algorithms to compensate for any residual imbalance in the analog sections. It is possible to implement the signal separation by implementation of (5.85) or (5.88).

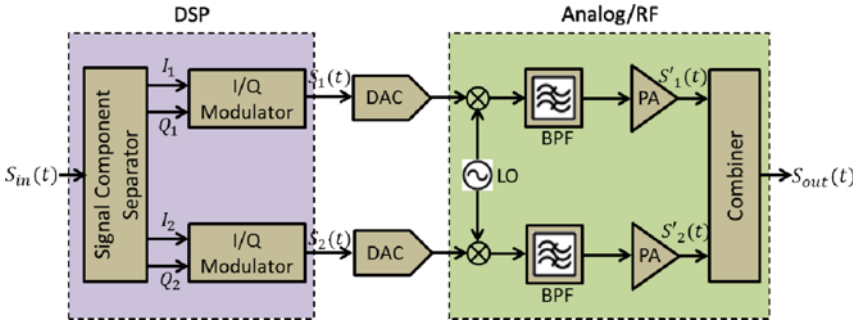


Fig. 5.36 LINC transmitter structure

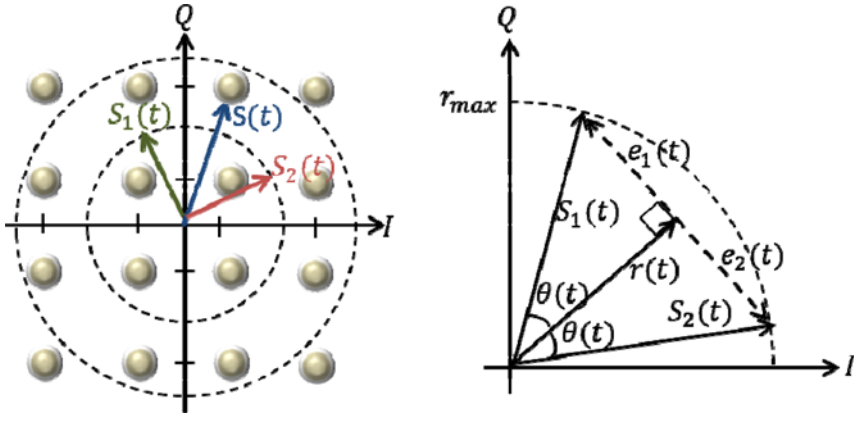


Fig. 5.37 LINC vector decomposition

There are a number of issues that must be addressed in digital implementation of the signal component separation block, such as the use of a look-up table technique, the bit resolution of the computations, the required memory resources, the required computation speed, and the choice of digital-to-analog converters [5].

The PAs should be designed for the highest possible efficiency at saturation, based on the selection of the PA's biasing and impedance matching circuits. For this reason, class F and inverse class F PA designs can be used, which have power efficiencies of around 80%. While the load changes from matched to very high impedance, the PAs should provide a rapid drop-off in DC power consumption in an ideal case.

The bandwidth response of the PA needs to be carefully selected, as the LINC separated signals have larger bandwidth than the original envelope-modulated signal, due to added phase modulation. Additionally, the power amplification block should be made of two identical or quasi-identical amplifiers, in order to preserve the amplitude and phase balance between the two branches [5].

The use of two highly efficient PAs operated with constant envelope signals in the LINC system does not guarantee that the overall efficiency of the LINC transmitter will be high, but the available power at the output of the combiner determines the overall efficiency of the LINC transmitter. In [26], it is shown that the average efficiency of the LINC system also depends on the power distribution function (PDF) of the signal and the type of combiner used. Hence, the combining structure is a key factor in the overall power efficiency and in the linearity of the LINC transmitter. The linearity depends on how the combining structure affects the impedances seen by the amplifiers and how it combines the output signals.

In this framework, the combining structures can be categorized into two classes: 1) matched and isolated combiners, which are also referred to as hybrid or resistive combiners; and, 2) non-matched and non-isolated combiners, which are

also referred to as out-phasing or Chireix combiners [5]. The overall system efficiency, η_{LINC}, of a LINC transmitter using an isolated combiner is given by:

$$\eta_{LINC} = \eta_{PA}^{\max} \cdot \eta_c \tag{5.89}$$

where $\eta_{PA}^{\max}$ is the maximum amplifier efficiency, and η_c is the combiner efficiency. This equation concludes that the LINC system efficiency drops rapidly for low-level signals. The LINC system efficiency can drop as low as $\eta_{PA}^{\max}/10$ at power back-off operation of 10 dB.

In order to make the LINC transmitter applicable for implementation in base stations, some limitations in the LINC transmitter need to be addressed: the gain and phase imbalance between the two branches; and, the bandwidth of the constant envelope signals feeding the PAs. The bandwidths of the phase-modulated signals at the output of the signal separator are generally five times wider than that of the original input signal.

5.5.3 EER Transmitter Architecture

The unmatched efficiency of switching-mode PAs has motivated attempts to use them for linear amplification. Envelope elimination and restoration (EER), proposed by L. R. Kahn in 1952, is one such attempt and is presented in Figure 5.38. A modulated RF signal is split into its polar components, the amplitude signal and the constant amplitude phase-modulated signal by an envelope detector and a limiter, respectively. The limiter output is a constant envelope signal that can be amplified by a highly nonlinear but power efficient PA, usually operating in switching mode, ideally without adding significant AM/AM and AM/PM distortion.

In the envelope path, the amplitude information is extracted by utilizing an envelope detector, which is used to modulate the supply voltage of the PA using the amplitude amplifier / bias modulator. In the phase path, a limiter is utilized to eliminate the amplitude signal variation, which generates the constant amplitude phase-modulated signal to be applied to the PA [6]. The amplitude information is restored to the envelope of the transmitted signal by modulating the supply voltage of the PA, which leads to the EER distinction.

The most significant challenges the EER transmitter faces are mainly related to the bandwidth of the signals to be amplified. The bandwidths of the phase and envelope signals are five times wider than that of the input signal. In fact, both the amplitude and RF power amplifiers should operate over a wide bandwidth, which inevitably reduces their efficiency. The performance of the EER transmitter is significantly affected by the synchronization between the envelope and the RF signal paths [27]. The current state of the art is minimization of the worst-case differential delay between the envelope and RF signal paths.

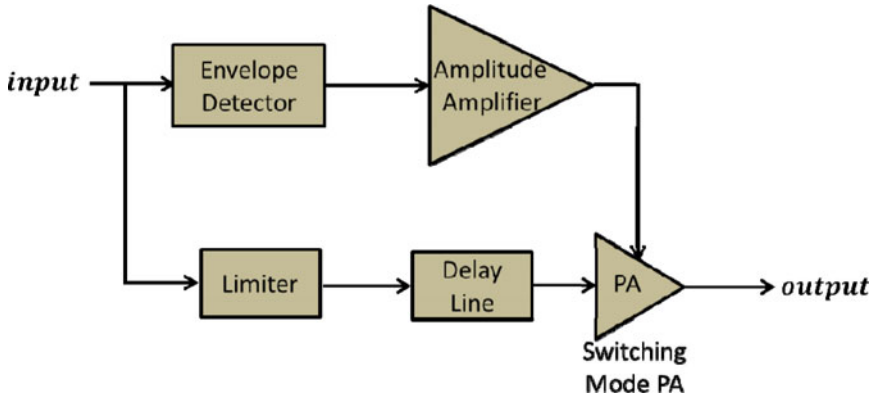


Fig. 5.38 EER transmitter structure

5.5.4 Delta-Sigma Transmitter Architecture

A delta-sigma ($\Delta\Sigma$) modulation based transmitter is another advanced amplification system that transforms the envelope-varying signal to a constant envelope signal and is used with high-efficiency PAs [28]. Figure 5.39 shows the block diagram of an RF delta-sigma based transmitter. Two delta-sigma modulators (DSMs), working at frequency, f_s, are implemented to produce bi-level signals for I and Q signals; and, two high-frequency multiplexers, working at frequency, Nf_s, are used to up-convert the baseband signals to the carrier frequency, $f_c = Nf_s$.

With the use of a third multiplexer, working at $2f_s$, the modulated signals produced by the two multiplexers are combined to generate I and Q signals at carrier frequency f_c. At the output of the third multiplexer, a switching-mode PA is employed to amplify the bi-level I/Q signal. Prior to transmitting the signal through the antenna, a bandpass filter is used to suppress all out-of-band distortion and also to recover the modulated signal around the carrier frequency [1].

The general structure of a DSM is demonstrated in Figure 5.40, in which a quantizer embedded in a loop with a digital-to-analog converter (DAC) in the feedback path is shown. An integrator is in the forward path of the modulator. The input signal to the integrator is the difference between the input signal, $x(t)$, and the quantized output value, $y(t)$, converted back to the predicted analog signal, $\hat{y}(t)$.

Considering that the DAC is ideal and signal delays are negligible, the difference between the input signal, $x(t)$, and the fed back signal, $\hat{y}(t)$, at the integrator input is equal to the quantization error. The quantization noise is denoted by the additive term, $E(t)$. This error is added up in the integrator and then quantized by a 1-bit analog-to-digital converter (ADC).

The output of a DSM is illustrated in the z-domain by:

$$Y(z) = H(z)X(z) + G(z)E(z) \tag{5.90}$$

where $X(z)$, $Y(z)$ and $E(z)$ denote the z-transforms of the input signal, the output signal and the quantization error, respectively.

The signal transfer function, $H(z)$, transforms the signal at the desired frequency band; and, the noise transform function, $G(z)$, suppresses the quantization noise in this band [29].

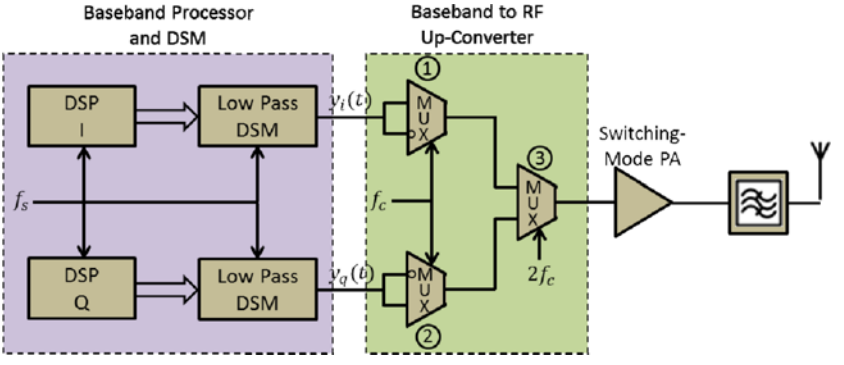


Fig. 5.39 Block diagram of the all-digital delta-sigma transmitter

In the case of a first-order DSM, which is shown in Figure 5.40, the z-domains of the signal and noise transfer functions can be obtained by:

$$H(z) = z^{-1} \tag{5.91}$$

$$G(z) = (1 - z^{-1}) \tag{5.92}$$

There are two DSM-based transmitter architectures in the literature to transform baseband pulsed signals to the desired RF frequency band: the low-pass (LP) DSM-based and the bandpass (BP) DSM-based transmitter architectures [30].

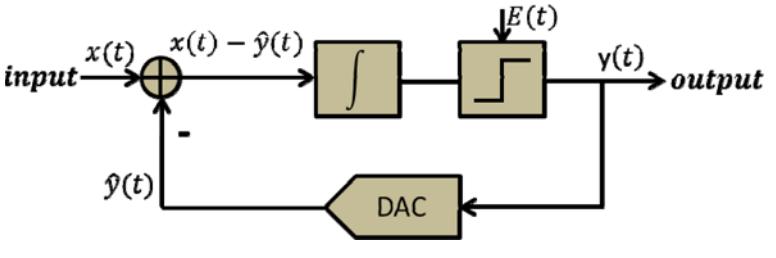


Fig. 5.40 A first-order low-pass DSM general structure

References

[1] Ghannouchi, F.M.: Power Amplifier and Transmitter Architectures for Software Defined Radio Systems. IEEE Circuits and Systems Magazine 10(4), 56–63 (2010)

[2] Larose, C.L., Ghannouchi, F.M.: Optimal Adaptation Methods and Class of Operation: Keys to Improving Feedforward Amplifier Power Efficiency. IEEE Transactions on Vehicular Technology 54(2), 456–467 (2005)

[3] Taijun, L., Bumaiza, S., Ghannouchi, F.M.: Augmented Hammerstein Predistorter for Linearization of Broad-Band Wireless Transmitters. IEEE Transactions on Microwave Theory and Techniques 54(4), 1340–1349 (2006)

[4] Ghannouchi, F.M., Hammi, O.: Behavioural Modeling and Predistortion. IEEE Microwave Magazine 10(7), 52–64 (2009)

[5] Birafane, A., El-Asmar, M., Kouki, A.B., Helaoui, M., Ghannouchi, F.M.: Analyzing LINC Systems. IEEE Microwave Magazine 11(5), 59–71 (2010)

[6] Raab, F.H., Asbeck, P., Cripps, S., Kenington, P.B., Popovic, Z.B., Pothecary, N., Sevic, J.F., Sokal, N.O.: Power Amplifiers and Transmitters for RF and Microwave. IEEE Transactions on Microwave Theory and Techniques 50(3), 814–826 (2002)

[7] Cripps, S.C.: RF Power Amplifiers for Wireless Communications, 2nd edn. Artech House, Norwood (2006)

[8] Kenington, P.B.: High-Linearity RF Amplifier Design. Artech House, Norwood (2000)

[9] Ebrahimi, M.M., Helaoui, M., Ghannouchi, F.M.: Trading-off Stability for Efficiency in Designing Switching-Mode GaN PAs for WiMAX Applications. In: Proc. IEEE Microwave Conference 2009 (APMC 2009), Asia Pacific, pp. 2348–2351 (December 2009)

[10] Grebennikov, A.: RF and Microwave Transmitter Design. John Wiley & Sons (2011)

[11] Hung, T.P.: High Efficiency Switching-Mode Amplifiers for Wireless Communication Systems: ProQuest (2008)

[12] Raab, F.H.: Class-E, Class-C, and Class-F power amplifiers based upon a finite number of harmonics. IEEE Transactions on Microwave Theory and Techniques 49(8), 1462–1468 (2001)

[13] Grebennikov, A., Sokal, N.O.: Switchmode RF Power Amplifiers: Newnes (2007)

[14] Berini, P., Desgagne, M., Ghannouchi, F.M., Bosisio, R.G.: An Experimental Study of the Effects of Harmonic Loading on Microwave MESFET Oscillators and Amplifiers. IEEE Transactions on Microwave Theory and Techniques 42(6), 943–950 (1994)

[15] Raab, F.H.: Maximum Efficiency and Output of Class-F Power Amplifiers. IEEE Transactions on Microwave Theory and Techniques 49(6), 1162–1166 (2011)

[16] Larose, C.L., Ghannouchi, F.M.: Optimization of Feedforward Amplifier Power Efficiency on the Basis of Drive Statistics. IEEE Transactions on Microwave Theory and Techniques 51(1), 41–54 (2003)

[17] Bassam, S.A., Helaoui, M., Ghannouchi, F.M.: Crossover Digital Predistorter for the Compensation of Crosstalk and Nonlinearity in MIMO Transmitters. IEEE Transactions on Microwave Theory and Techniques 57(5), 1119–1128 (2009)

[18] Hammi, O., Ghannouchi, F.M.: Power Alignment of Digital Predistorters for Power Amplifiers Linearity Optimization. IEEE Transactions on Broadcasting 55(1), 109–114 (2009)

[19] Katz, A.: Linearization: Reducing Distortion in Power Amplifiers. IEEE Microwave Magazine 2(4), 37–49 (2001)

[20] Cardinal, J.S., Ghannouchi, F.M.: A New Adaptive Double Envelope Feedback (ADEF) Linearizer for Solid State Power Amplifiers. IEEE Transactions on Microwave Theory and Techniques 43(7), 1508–1515 (1995)

[21] Ghannouchi, F.M.: An S Band RF Digital Linearizer for TWTAs and SSPAs. In: European Conference on Circuit Theory and Design (ECCTD 2009), Antalya, Turkey, pp. 735–738 (August 2009)

[22] Hashmi, M.S., Rogojan, Z.S., Ghannouchi, F.M.: A Flexible Dual-Inflection Point RF Predistortion Linearizer for Microwave Power Amplifiers. Progress in Electromagnetics Research C 13, 1–18 (2010)

[23] Rezaei, S., Hashmi, M.S., Dehlaghi, B., Ghannouchi, F.M.: A Systematic Methodology to Design Analog Predistortion Linearizer for Dual Inflection Power Amplifiers. In: International Microwave Symposium (IMS 2011), Baltimore, Maryland, USA (August 2011)

[24] Nagle, P., Burton, P., Heaney, E., McGrath, F.: A wide-band linear amplitude modulator for polar transmitters based on the concept of interleaving delta modulation. IEEE Journal of Solid-State Circuits 37(12), 1748–1756 (2002)

[25] Staszewski, R.B., Wallberg, J.L., Rezeq, S., Hung, C.M., Eliezer, O.E., Vemulapalli, S.K., Fernando, C., Maggio, K., Staszewski, R., Barton, N., Lee, M.C., Cruise, P., Entezari, M., Muhammad, K., Leipold, D.: All-digital PLL and transmitter for mobile phones. IEEE Journal of Solid-State Circuits 40(12), 2469–2482 (2005)

[26] Birafane, A., Kouki, A.: On the Linearity and Efficiency of Outphasing Microwave Amplifiers. IEEE Transactions on Microwave Theory and Techniques 52(7), 1702–1708 (2004)

[27] Hammi, O., Helaoui, M., Ghannouchi, F.M.: Green Power Amplification Systems for 3G+ Wireless Communication Infrastructure. In: Proc. IEEE Vehicular Technology Conference, pp. 1–5 (September 2010)

[28] Schreier, R., Temes, G.C.: Understanding Delta-Sigma Data Converters. Wiley-IEEE Press (2004)

[29] Helaoui, M., Hatami, S., Negra, R., Ghannouchi, F.M.: A Novel Architecture of Delta-Sigma Modulator Enabling All-Digital Multiband Multistandard RF Transmitters Design. IEEE Transactions on Circuit and Systems II: Express Briefs 55(11), 1129–1133 (2008)

[30] Ghannouchi, F.M., Hatami, S., Aflaki, P., Helaoui, M., Negra, R.: Accurate Power Efficiency Estimation of GHz Wireless Delta-Sigma Transmitters for Different Classes of Switching Mode Power Amplifiers. IEEE Transactions on Microwave Theory and Techniques 58(11), 2812–2819 (2010)

Chapter 6
Transmitter Design for MIMO Wireless Communications

The high demand for broadband multimedia Internet access and wireless connections has increased the need for more advanced and sophisticated wireless communication systems. However, wireless channels usually provide limited bandwidth and lower quality links.

The next generation of wireless technologies is targeting two essential goals in their design and development. One is the provision of high-speed data rates up to 100 megabits per second (Mb/s) for mobile users and 1 gigabit per second (Gb/s) for stationary users. Achieving the goal of improving the data rate and increasing the system capacity is feasible through the use of more advanced signal processing and coding techniques, such as spectral efficient 64-QAM (quadrature amplitude modulation), orthogonal frequency-division multiplexing (OFDM), and multiple input multiple output (MIMO) topology.

6.1 Complexity and Cost in MIMO Systems

MIMO wireless communication systems are designed to deliver either maximal diversity to enhance transmission reliability or to increase maximal multiplexing gain to support high data rate [1]. The channel capacity and spectral efficiency of a MIMO system is usually much higher than that of a single-input single-output (SISO) system. This can be achieved by employing multiple antennas at the transmitter and receiver.

This performance of a MIMO system can be quantified by the spatial multiplexing gain. Therefore, it is possible to send parallel independent data streams to achieve overall system capacities scaled with $\min(N_t, N_r)$, where N_t and N_r are the number of receiving and transmitting antennas [2]. On the other hand, if the signal copies are transmitted from or received at multiple antennas, this multiantenna system can provide a gain that enhances the reliability of a wireless link. This gain is known as diversity gain, which is achieved by space-time coding. Both diversity and multiplexing gain can be supported by MIMO systems at the same time, but there is a fundamental trade-off between them [1].

In multi-branch transceivers, multiple radio frequency (RF) front-ends are integrated on the same platform. Multiple modulated signals are simultaneously

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 129–149.
springerlink.com

transmitted and received using these RF front-ends. The multi-branch MIMO transceiver is a simple case of MIMO transceiver architecture, which can improve the power and diversity gains of the system. Transmitter and receiver nonlinearity, modulator imbalance and other impairments should be considered in the design of each branch of a multi-branch transceiver. In addition, when multiple transmission/reception paths are realized on the same chipset, new issues are generated, such as RF crosstalk between the multiple paths, which are due to the proximity of the different circuits [3].

A multi-branch transceiver with N antennas at the transmitter and N antennas at the receiver is demonstrated in Figure 6.1. This system uses multiple parallel RF front-ends, and the number of RF front-ends is equal to the number of antennas. At the receiver, the baseband processing unit decodes N received baseband paths to recover the signal and also to obtain the diversity gain.

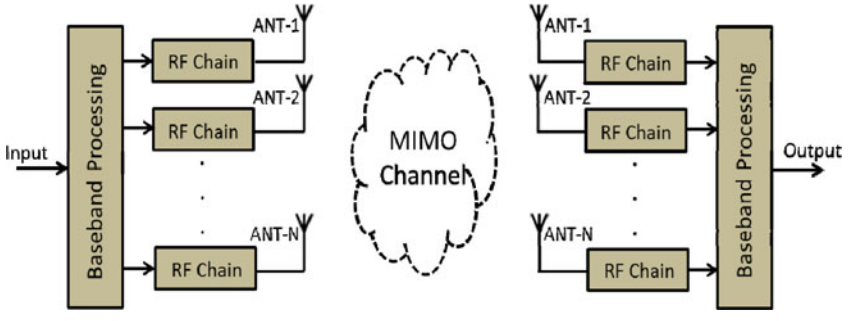


Fig. 6.1 A conventional multiantenna receiver

The multiple antenna RF front-end architecture design usually leads to higher complexity and hardware costs in the RF section. Furthermore, increasing the number of antennas causes growth in RF circuit mismatches and coupling. In fact, these issues limit the application of a high number of antennas at the transceiver. One option to overcome such a problem is the adoption of a single RF front-end in a MIMO system, where a single RF path is implemented instead of multiple parallel RF paths [2], [4]. This reduces the complexity and cost of an RF section and also enables a compact design with lower power consumption.

An orthogonal transmission of multiple RF streams over a single front-end must be identified for the realization of a single RF front-end path. Common designs that may be used for a single receiver front-end realization are the antenna selection technique [5], frequency-division multiplexing, time-division multiplexing [1], and code-division multiplexing [6].

6.2 Transmitters Architectures

Several architectures have been proposed for transmitters in wireless communications, which are dependent on the performance requirements.

6.2.1 Direct Conversion Transmitter

The direct conversion transmitter architecture has received considerable attention for multi-standard, multi-band applications, due to its smaller number of components and usability. This architecture consists of a digital baseband processing unit, digital-to-analog converters (DACs), a direct up-converter, a PA, and an antenna. The digital baseband processing unit has the responsibilities of digital modulation, coding and filtering, in order to develop and synthesize an appropriate transmission signal from the digital input data.

The DACs convert the digital in-phase (I) and quadrature (Q) baseband signals to analog signals, which feed the quadrature modulator for direct up-conversion. Analog signals are directly up-converted to the RF with I and Q carriers. The block diagram of a direct conversion transmitter is shown in Figure 6.2. The bandpass filter after the signal summation is used to suppress the out-of-band signals produced by the harmonic distortion of the carrier. The RF signal power is increased by the PA for transmission through the antenna.

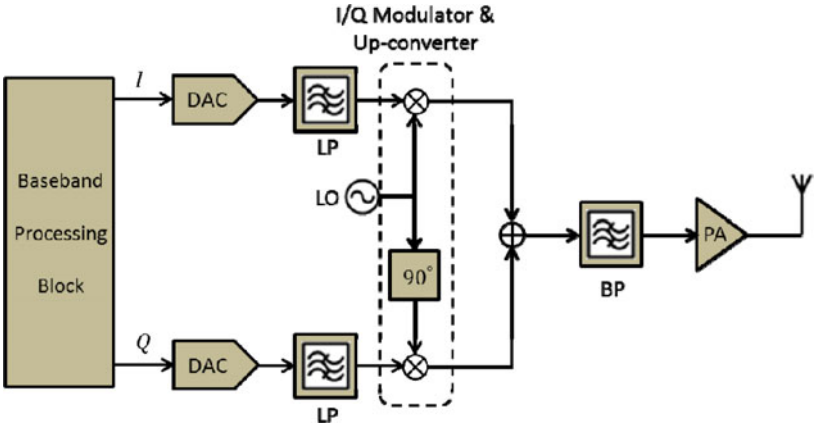


Fig. 6.2 Block diagram of a direct conversion transmitter

The direct conversion transmitter is theoretically simple, and there are no intermediate frequency (IF) components. The drawbacks of this architecture are local oscillator (LO) leakage at RF frequencies, voltage controlled oscillator (VCO) pulling, and the requirement of an I/Q mixer at RF frequencies [7]. The direct conversion transmitter suffers from unequal complex gains of the I and Q paths. This depends on the frequency and operating temperature. The quality of the output RF signal is strongly affected by the I/Q imbalance caused by the modulators.

In addition, the PA stage has nonlinear behavior, which deteriorates the output signal quality and produces out-of-band power emission. Digital predistortion (DPD) techniques are suitable candidates to enhance the signal quality and compensate for the out-of-band power. However, the I/Q imbalance at the transmitter can considerably reduce the linearization capacity of the digital predistorter [8].

6.2.2 *Superheterodyne Transmitter*

The most common double conversion approach is the superheterodyne architecture. A schematic of a superheterodyne transmitter is shown in Figure 6.3. In the first stage, the baseband input is up-converted to IF by an I/Q mixer. The first bandpass filter, BP1, shown in this figure reduces any spurious out-of-band power prior to the second up-conversion. In the second stage, the IF signal goes through a second up-conversion to achieve the desired transmit frequency. A second bandpass filter, BP2, can be utilized to filter additional spurious out-of-band power caused by the second mixing operation [9]. The IF frequency is determined by the first LO frequency, f_1. Two frequencies are generated by the second mixer stage, $f_2 + f_1$ and $f_2 - f_1$. One of the frequencies is selected by the second bandpass filter and the unwanted frequency (image frequency) is rejected.

The superheterodyne architecture was developed to overcome the inherent problems in the direct conversion or homodyne architecture. In this architecture, some of the unwanted frequency components originating from nonlinearities (e.g., in the mixers) can be removed by filters. Figure 6.4 shows the frequency plan of the superheterodyne transmitter with one IF frequency.

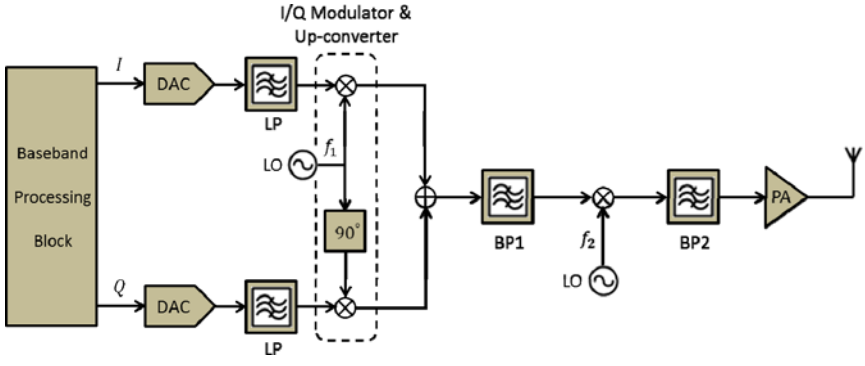


Fig. 6.3 Block diagram of a superheterodyne transmitter

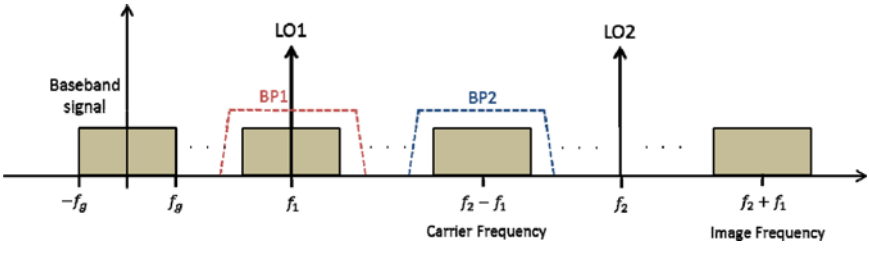


Fig. 6.4 Frequency plan of the superheterodyne transmitter

6.3 Brief Overview of MIMO Transmission Schemes

In telecommunications, multiplexing is a technique by which multiple analog signals or digital data streams are combined into a single signal to be transmitted over a shared medium. The basis of the multiplexing technique is the division of the entire signaling dimensions into parts or channels and the allocation of these parts or channels to different users. The most common techniques of dividing signal space are along the frequency, time and code axes, which are called frequency-division multiplexing (FDM), time-division multiplexing (TDM), and code-division multiplexing (CDM).

6.3.1 FDM Technique

FDM is a technique that combines several signals into one medium by sending signals in several distinct frequency ranges over that medium, as shown in Figure 6.5 (a). This is a familiar technique in many industries, such as telephone and commercial radio and television broadcasting. Although the FDM technique can be applied to both analog and digital systems, it has been widely used in the analog communication systems.

Guard bands are used in the FDM technique to avoid interference between channels. The FDM produces a spectral efficiency, which is due to the transmission rate that is quite close to the maximum rate needed by the user. FDM needs frequencies that can provide different carriers with different channels, because the transmission in communication systems is continuous.

6.3.2 TDM Technique

TDM is a multiplexing technique that allows more than one user to access RF channels without any interference between them. Figure 6.5 (b) shows how the frequency channels are shared at different times. It involves the sequencing of groups of signals from each individual input, so that they can be associated with the appropriate receiver. TDM has been used in digital communication technologies, including significant application in cellular phone technologies. The guard times are used in the TDM technique to reduce the transmission impairments, such as delay propagation, the transient response of the pulse signal, and other impairments.

6.3.3 CDM Technique

CDM is a multiplexing technique where several channels share the same frequency spectrum at the same time, as shown in Figure 6.5 (c). CDM employs spread-spectrum technology and a special coding method in which each transmitter is

assigned a code, allowing multiple users to be multiplexed over a shared physical medium. Direct sequence spread spectrum (DSSS) and frequency hopping are two forms of the CDM technique. In CDM, the modulated coded signal has a much higher bandwidth than the data being communicated. One of the early applications for CDM is in Global Positioning Systems (GPSs). CDM techniques are also used in wireless networks due to the advantage of reducing the interference among different users.

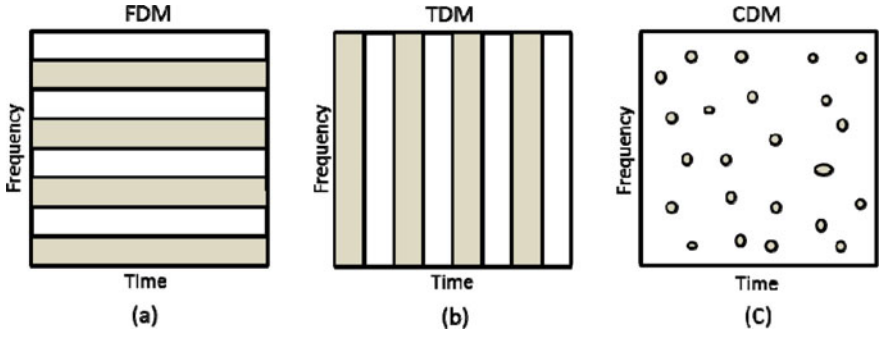


Fig. 6.5 Time and/or frequency sharing in (a) TDM, (b) FDM and (c) CDM techniques

6.4 MIMO Transceiver Architectures

MIMO refers to a system where multiple inputs have interaction with multiple outputs. In the context of wireless communications, MIMO refers to topologies in which multiple modulated signals, separated in the space, time or frequency domain are simultaneously transmitted through an RF front-end. A MIMO system with modulated signals separated in the space domain is are already known as a MIMO system in wireless communication theory. This topology has multiple branches of RF front-ends that are involved in simultaneous transmission of signals. A MIMO system with modulated signals separated in the frequency domain is a system in which multiple signals modulated in different carrier frequencies are transmitted all together through a single-branch RF front-end.

6.4.1 Antenna Selection Architecture

Antenna selection techniques in spatial multiplexing at the transmitter and/or receiver can lead to simpler RF front-end MIMO systems. Based on antenna selection techniques, some of the available antennas are chosen; therefore, fewer RF chains than the number of transmitter and/or receivers antennas are used. Consequently, this reduces cost and complexity; and, at the same time, the system performance is maintained. The antenna selection technique, both in the transmitter

and receiver of a MIMO system, is demonstrated in Figure 6.6. Although the implementation of the antenna selection technique in the receiver is straightforward, the implementation in the transmitter needs a feedback path from the transmitter to the receiver [2].

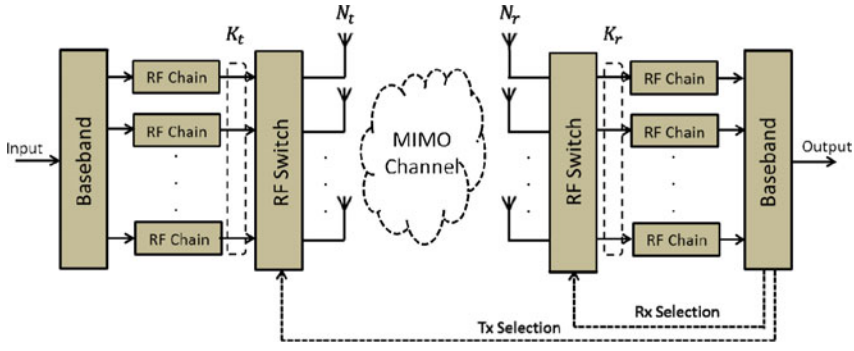


Fig. 6.6 A MIMO system with antenna selection technique

The performance of the antenna selection algorithm is largely dependent on the objective function or the selection criterion utilized to construct the subset of antennas. Several criteria have been introduced. A number of common selection criteria are briefly presented in the following subsections.

6.4.1.1 Maximum Capacity Criterion

Channel capacity maximization is a typical criterion for antenna selection. An analytical bound for the channel capacity of MIMO systems with antenna selection [1], [10] can be given as:

$$C \le \sum_{i=1}^{K_r} \log_2\left(1 + \frac{\rho}{N_t}\gamma_i\right) \tag{6.1}$$

where ρ is the mean signal-to-noise ration (SNR), K_r is the number of selected antennas in the receiver, and γ_i represents the squared norm of the i th row of the MIMO channel matrix, H, which is ordered from the smallest to the largest. Based on the capacity formula in (6.1), the capacities for all possible subsets of a receiver's antennas, $p \in P$, are calculated; and, the subset with the largest capacity C_p is selected.

6.4.1.2 Maximum Minimum Singular Value Criterion

For each subset of receiver antennas, $p \in P$, the minimum singular value ($\lambda_{\min}$) corresponding to different H_p is calculated. Then, the subset with the largest $\lambda_{\min}$ is selected.

6.4.1.3 Norm-Based Selection (NBS) Criterion

In the norm-based selection (NBS) criterion, the subset of receiver antennas related to the rows of H with the largest Euclidean norm is chosen. Although this technique is not optimal when the number of RF front-ends is greater than one [11], it remains a popular criterion because of its simplicity.

6.4.2 Frequency-Division Multiplexing (FDM) Architecture

In FDM architecture, the signals of different antennas are shifted in frequency with mixing by different LOs, added together, and transmitted through a single-branch RF frond-end. Figure 6.7 shows the general block diagram of FDM architecture. The multiple signals from different antennas are separated in the frequency domain. A single RF front-end is used to down-convert the signals, and the frequency shifts of the multiple streams are removed in the baseband block. The diversity gain is extracted by further processing [4]. Examples of this type of system are concurrent dual-band transmitters and multicarrier transmitters.

A system-level performance of a scenario with two antennas at the receiving side was studied in [4]. Considering $s(t)$ as the passband signal with a center frequency, f_c, and a bandwidth, ω, the received signal of the first antenna is $h_1 \cdot s(t) + n_1(t)$; and, the received signal from the second antenna is $h_2 \cdot s(t) + n_2(t)$, where $n_1(t)$ and $n_2(t)$ are the passband noise, and h_1 and h_2 are the Rayleigh flat fading channel coefficients. The signal of second antenna is combined with a low-frequency oscillator at the same frequency (ω), then filtered to remove the IF term.

The deficiency of this technique is the requirement of a narrowband filter in the RF frequency. The simulation results for binary phase shift keying (BPSK) modulation using two antennas are demonstrated in Figure 6.8. The variation between the ideal and simulated curves for diversities 1 and 2 is related to the filter design and its quality factor.

6.4.3 Time-Division Multiplexing (TDM) Architecture

The general block diagram of TDM architecture is shown in Figure 6.9. In this architecture, the same number of antennas as in the conventional topology is used,

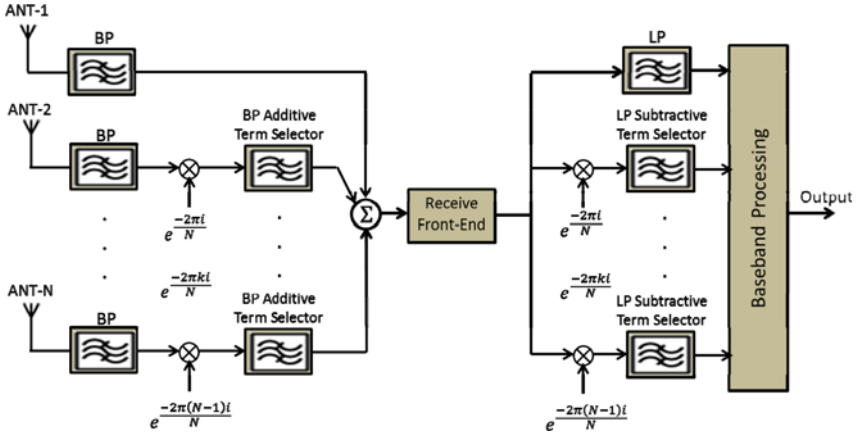


Fig. 6.7 Realization of single-branch multiantenna receiver based on FDM

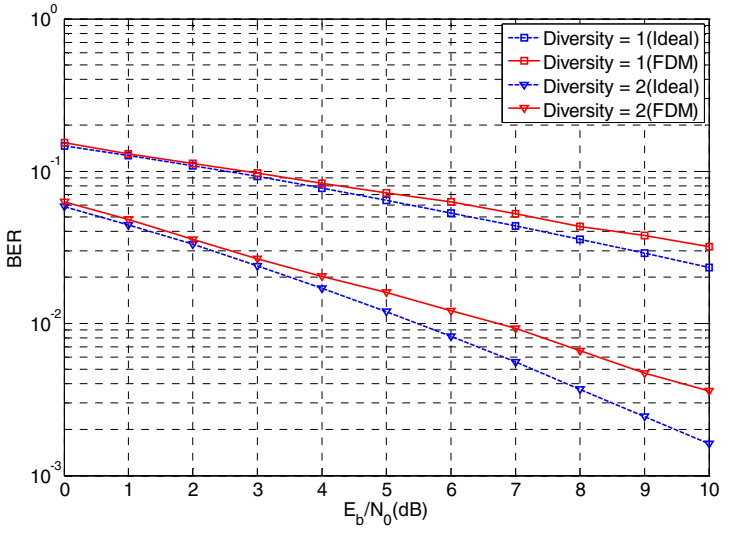


Fig. 6.8 Bit error rate (BER) versus E_b / N_0 for a single frond-end receiver based on FDM using BPSK modulation [4]

but instead of having multiple RF front-ends, a single pole, multiple-throw RF switch along with a single RF front-end are used to down-convert the RF signals to baseband. Furthermore, the signals are carried to the baseband processing unit using a de-multiplexer. The switch is used to capture the signals of all the antennas for every symbol time interval of the modulated signal. This offers some constraints on the switching speed. In addition, careful alignment of the data before

and after the multiplexing and de-multiplexing of the multiple signals in the MIMO receivers is desired.

A single front-end MIMO receiver using the TDM technique can be realized as a single RF architecture by time-multiplexing the antennas' signals using a single-pole multiple-throw RF switch. The received signals are de-multiplexed after RF processing according to a diversity algorithm. Although this architecture implements the same number of antennas as in the conventional architecture, it reduces the number of RF front-ends from N to 1. Consequently, the overall cost and size of the multiantenna receiver and also RF design mismatch are significantly reduced [2].

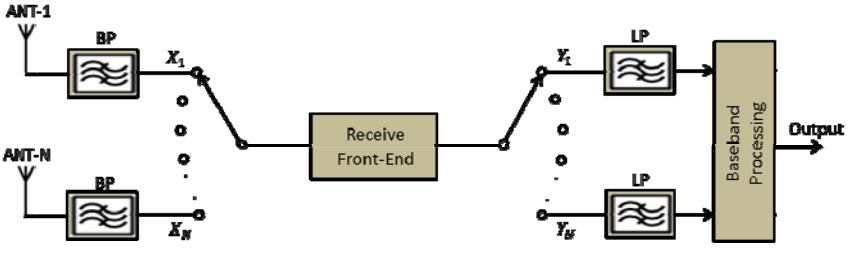


Fig. 6.9 Realization of single-branch multiantenna receiver based on TDM

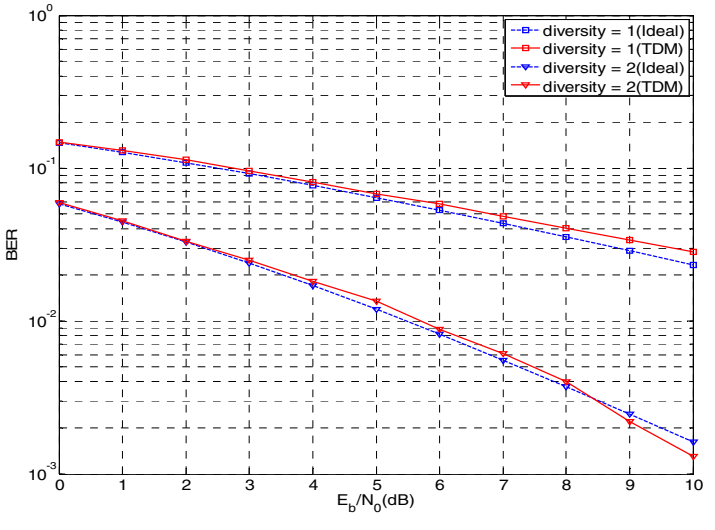


Fig. 6.10 BER versus E_b/N_0 for a single frond-end receiver based on TDM using BPSK modulation [4]

Figure 6.10 shows the results of a time-multiplexed receiver using a raised cosine pulse shape filter with a roll-off factor of 0.5 for single antenna and two-antenna cases using BPSK modulation. Zero forcing (ZF) receivers are implemented. To compare the TDM architecture with a conventional architecture, the bit error rate (BER) of the multiple frond-end systems is demonstrated as well. The diversity gain using a single time-multiplexed RF front-end is equal to the diversity gain of the multiple RF front-end, as shown in Figure 6.10. This architecture demands a wider bandwidth than the one used in the conventional MIMO receiving front-end.

6.4.4 Code-Division Multiplexing (CDM) Architecture

In code-division multiplexing (CDM), the signals from different antennas are multiplied by orthogonal codes and added together. At the receiver, a single RF front-end is used to down-convert the resultant RF signals to baseband, where the signals are again multiplied by the orthogonal codes, integrated, and de-multiplexed. This technique is demonstrated in Figure 6.11. In the case of two receiver antennas, the first and second signals are multiplied by $C_1(t)$ and $C_2(t)$ codes, respectively, in the symbol duration, and then summed together. These codes should be orthogonal, i.e., $C_1(t)$ is equal to 1 between 0 and T_s, and $C_2(t)$ is equal to 1 between 0 and $T_s/2$ and equal to -1 between $T_s/2$ and T_s. T_s is the symbol duration. After summing the signals, the total signal, $(h_1 \cdot s(t) + n_1(t)) \cdot c_1(t) + (h_2 \cdot s(t) + n_2(t)) \cdot c_2(t)$, is down-converted using a single RF front-end. The baseband signal is expressed as $(h_1 \cdot \tilde{s}(t) + \tilde{n}_1(t)) \cdot c_1(t) + (h_2 \cdot \tilde{s}(t) + \tilde{n}_2(t)) \cdot c_2(t)$.

Followed by down-conversion in the baseband processing unit block, the signal is multiplied by $C_1(t)$ and $C_2(t)$. This separates the signal into different paths. An integrator removes the effect of the other signals in each path. However, the effect of the noise of both antennas is preserved in each path. The output of the integrator in the first path can be obtained as:

$$\begin{aligned} & h_1 \cdot \tilde{s} + \tilde{n} = \\ & \frac{1}{T_s} \int_0^{T_s} \left[(h_1 \tilde{s}(t) + \tilde{n}_1(t)) \cdot c_1(t) \cdot c_1(t) + (h_2 \cdot \tilde{s}(t) + \tilde{n}_2(t)) \cdot c_2(t) \cdot c_1(t) \right] dt \\ & = \frac{1}{T_s} \int_0^{T_s} h_1 \tilde{s}(t) dt + \frac{1}{T_s} \int_0^{T_s} \tilde{n}_1(t) dt + 0 + \frac{1}{T_s} \int_0^{T_s} \tilde{n}_2(t) \cdot c_2(t) \cdot c_1(t) dt \end{aligned} \tag{6.2}$$

where $\tilde{s}(t)$ is invariable in the symbol duration, T_s, and the codes have unit energy. The third term is zero, which is due to the orthogonality of $C_1(t)$ and $C_2(t)$; and, $\tilde{s}(t)$ is constant in the symbol duration, T_s. The fourth term is non-zero, because $\tilde{n}_2(t)$ is stochastic and variable in the symbol duration.

The second and fourth terms of the integral have similar power; therefore, the output noise power using CDM is twice that of conventional design. The noise level is increased by $10\log(N)$ dB in each branch, when N antennas are utilized and the signals are down-converted with the CDM technique [12]. Figure 6.12 demonstrates the simulation results for BPSK modulation using two receiving antennas. A separation of 3 dB can be noticed between the CDM technique and the conventional diversity system [4].

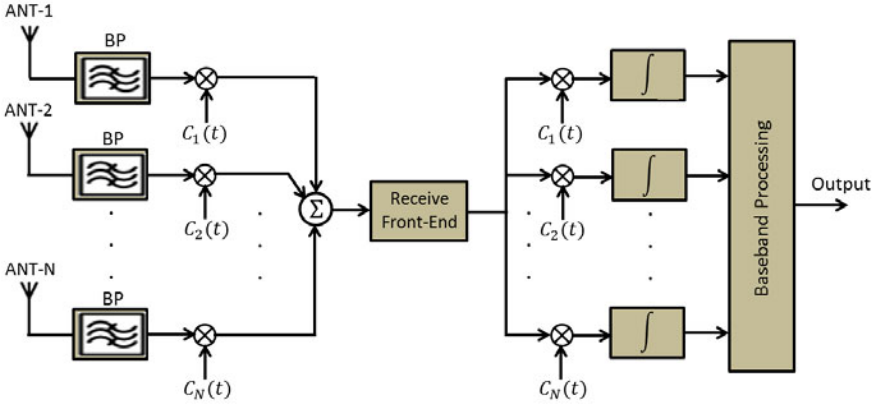


Fig. 6.11 Realization of single-branch multiantenna receiver based on CDM

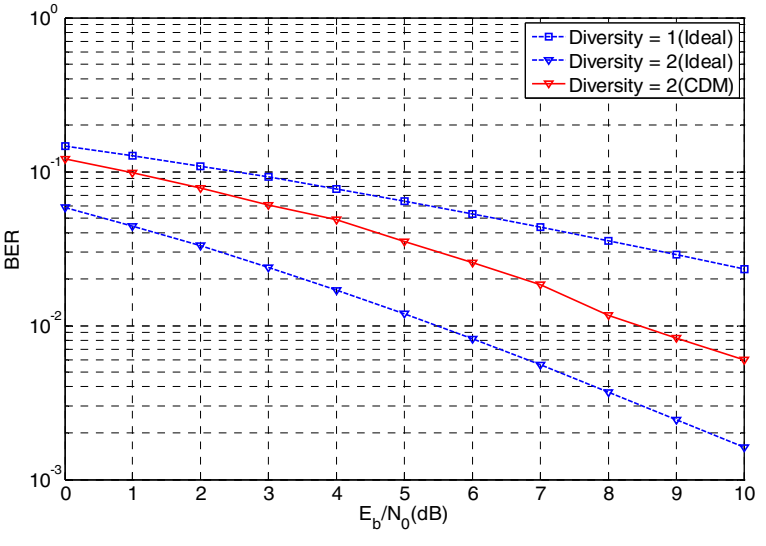


Fig. 6.12 BER versus E_b/N_0 for a single frond-end receiver based on CDM using BPSK modulation [4]

6.5 Distortion and Impairment Compensation in MIMO Transmitters

MIMO topology is a highly efficient solution for improving the spectral efficiency of wireless systems. Indeed, moving from SISO transceivers to MIMO transceivers could theoretically multiply the system capacity or data rate by the number of outputs integrated in the MIMO transceiver. However, MIMO topology faces various implementations issues that can be classified into two major categories. The first category consists of issues related to the general transceiver design, such as transmitter linearity, receiver dynamic range, and imbalance and leakage in mixers [3]. This group is not unique to MIMO systems.

The second category is specific to MIMO transceivers. In multi-branch MIMO transceivers, crosstalk results from the coupling and interference between the signals of different branches. Crosstalk is more likely to take place between the branches, because the signals in different branches use the same operating frequency and have equal transmission power [13]. This is more significant in an integrated circuit (IC) design, especially when the actual footprint of the circuit is small [14].

RF crosstalk can be categorized as linear and nonlinear. Linear crosstalk occurs beyond the output of the transmitter, i.e., at the antenna, and can be modeled as a linear function of the interference and desired signals. This means the signal affected by linear crosstalk does not pass through nonlinear components. Conversely, nonlinear crosstalk affects the signal before it passes through nonlinear components. This crosstalk that takes place in the transmitter circuit prior the PA is the main source of nonlinear crosstalk, since the PA is the main source of nonlinearity [3]. The origin and nature of both types of crosstalk are illustrated in Figure 6.13.

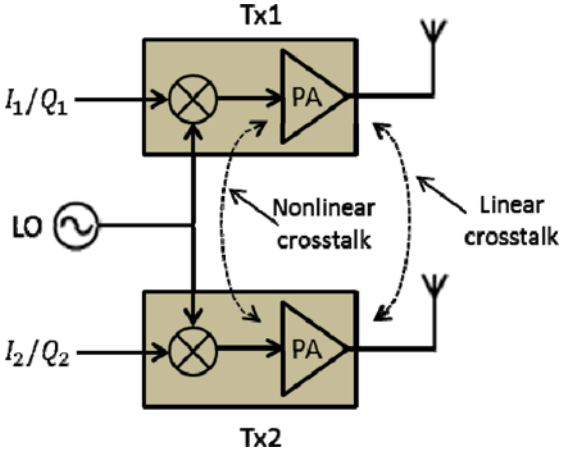


Fig. 6.13 Linear and nonlinear crosstalk in dual-branch MIMO transmitters

6.5.1 Antenna Crosstalk

Antenna crosstalk is the effect of the transmitted signal from one antenna element to another antenna element, which can be modeled as:

$$\vec{Y} = \mathrm{A}\vec{X} \tag{6.3}$$

where $\vec{X} = [x_1 \quad x_2]^T$ and $\vec{Y} = [y_1 \quad y_2]^T$ are the signals at the input and output of the antennas, respectively; and, $\mathrm{A} = \begin{bmatrix} 1 & \alpha \\ \beta & 1 \end{bmatrix}$ is the antenna crosstalk matrix. For symmetric antenna crosstalk, α and β are equal [3].

With the expansion of (6.3), the signals at the output of the antennas after crosstalk can be shown as:

$$\begin{cases} y_1 = x_1 + \alpha x_2 \\ y_2 = \beta x_1 + x_2 \end{cases} \tag{6.4}$$

From this equation, it can be observed that y_1 and y_2 are linear functions of the signals at the antennas' input.

Linear crosstalk has been addressed in several research studies [15], [16]; and, several techniques have been proposed to compensate for it. The compensation is simultaneously performed mostly at the receiver side for the composite linear crosstalk generated at the transmitter and receiver antennas and also by the channel.

A 2×2 MIMO system with three types of linear crosstalk is shown in Figure 6.14. The total linear crosstalk can be modeled as:

$$\vec{Y} = (\mathrm{BHA})\vec{X} + \mathrm{B}\vec{N} \tag{6.5}$$

where A and B are crosstalk matrices for the transmitting and receiving antennas, respectively; and, $\mathrm{H} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$ is a crosstalk matrix for the channel; and, $\vec{N}$ is an additive white Gaussian channel noise (AWGN) vector.

The uncorrelated received signal using the matrix inversion algorithm is given by:

$$\vec{Y}_{uncorr.} = (\mathrm{BHA})^{-1}\vec{Y} \tag{6.6}$$

Substituting (6.5) into (6.6):

$$\vec{Y}_{uncorr.} = \vec{X} + (\mathrm{HA})^{-1}\vec{N} \tag{6.7}$$

where $\vec{Y}_{uncorr.}$ is the uncorrelated component of the received signal without the noise component [3].

From (6.7), it can be observed that the matrix inversion method is practical if the (BHA) matrix is invertible. Furthermore, the crosstalk in the transmitter antenna and MIMO channel may deteriorate the performance of the MIMO system due to the noise enhancement, as the last term in (6.7). The performance of the MIMO system is not sensitive to the receiver antenna crosstalk as long as matrix **B** is invertible.

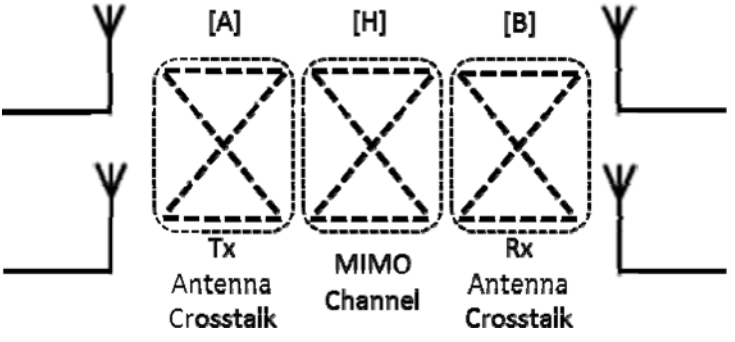


Fig. 6.14 A 2×2 MIMO channel and antenna crosstalk

6.5.2 Nonlinear RF Crosstalk

MIMO transmitters need to maintain certain levels of efficiency and linearity in the system, similar to those of single-branch transmitters. Actually, the power efficiency and linearity requirements of MIMO transmitters are more restrained than those of single-branch transmitters. As demonstrated in Figure 6.13, crosstalk that occurs before the PA is recognized as nonlinear. Some sources of this type of crosstalk can be the leakage of the RF signal through the common LO [14] and interference in the chipset.

Considering RF nonlinear crosstalk and transmitters' nonlinearities, the transmitters' output in a dual-branch MIMO transmitter can be modeled as:

$$y_1 = f_1(x_1 + \alpha x_2) \tag{6.8}$$

$$y_2 = f_2(\beta x_1 + x_2) \tag{6.9}$$

where $f_1(\cdot)$ and $f_2(\cdot)$ are nonlinear functions that represent the transmitter responses of each branch; x_i and y_i are the input and output signals, respectively; and, α and β are the RF nonlinear crosstalk. Since $f_1(\cdot)$ and $f_2(\cdot)$ are nonlinear, a simple matrix inversion is not sufficient to compensate for the effect of nonlinear crosstalk.

Due to unavoidable RF crosstalk between branches of a MIMO transmitter, independent modeling of each branch is not adequate to include the effects of RF nonlinear crosstalk. Hence, the traditional DPD techniques developed for SISO systems are not suitable for MIMO systems. In fact, a new DPD architecture with the capability of having more than one input signal is required [3].

6.5.3 Effects of Nonlinear Crosstalk on DPD Extraction

In Figure 6.15 (a), a MIMO transmitter is shown as a black box with two inputs and two outputs. One way to model the dual-branch MIMO transmitter is independent modeling of each branch of the MIMO transmitter, which is illustrated in Figure 6.15 (b). The independent modeling is reliable if there are no interactions between the two branches. However, as discussed previously, nonlinear crosstalk is an inevitable incident in dual-branch MIMO transmitters. As a matter of the fact, the transmitters' output with the effect of crosstalk can be modeled as in (6.8) and (6.9). It can be observed from the expressions in (6.8) and (6.9) that the signal at the output of each branch is a function of both input signals, in the presence of RF nonlinear crosstalk. Hence, independent modeling of each branch, according to the input and output signals of that branch, leads in an inefficient model for the dual-branch transmitter. Similarly, the use of DPD linearization techniques leads to a similar scenario when linearization blocks are realized separately for each branch of the dual-branch transmitter.

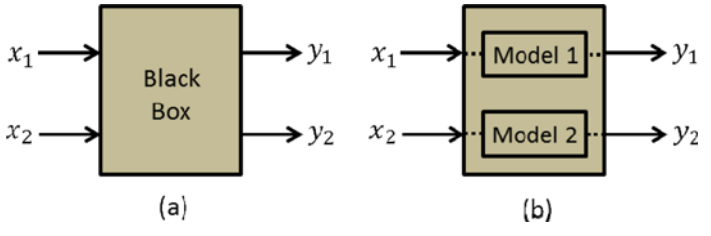


Fig. 6.15 A MIMO transmitter as (a) a black box, (b) two independent parallel models

The cascade of DPD and the RF frond-end complex transfer function in an ideal transmitter without nonlinear crosstalk results in a linear complex transfer function. However, in MIMO transmitters, the quality of the DPD extraction degrades when nonlinear crosstalk modifies the complex envelope of the RF signal at the input of the PA of a transmitter [3]. As shown in Figure 6.16, the DPD coefficients of the upper branch transmitter are extracted using the digital baseband signal to be transmitted (z_1) and the equivalent complex envelope of the PA RF output (y_1). The coupled signal from the second path (z_2) is also added to signal

z_1, and the combined signal ($z_1 + \alpha z_2$) is amplified by the PA. Hence, the output PA signal can be expressed as:

$$y_1 = f_1(z_1 + \alpha z_2) \tag{6.10}$$

In the DPD extraction process, the DPD function, $g_1(\cdot)$, is the inverse function of $f_1(\cdot)$, which depends on both input signals, x_1 and x_2, through z_1 and z_2 terms. Considering that $g_1(\cdot)$ is only a function of the x_1 input signal, the PA output signal is:

$$y_1 = f_1(g_1(x_1) + \alpha z_2) \neq G_0 \cdot x_1 \tag{6.11}$$

where G_0 is the linear or small-signal gain of the PA.

Hence, the nonlinear crosstalk influences the extraction and effectiveness of the DPD function.

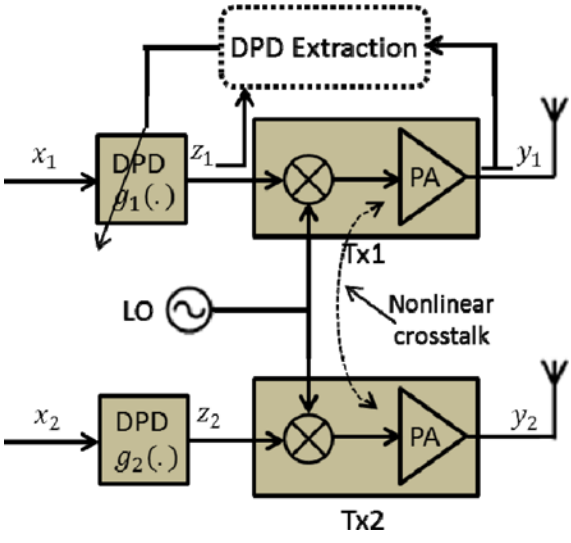


Fig. 6.16 DPD and PA with nonlinear crosstalk in a MIMO system

6.5.4 Impairment and Distortion Compensation

Most of the techniques proposed so far have addressed either nonlinear distortion, the I/Q imbalance of the modulator, or the coupling effects in MIMO transmitters. However, there are not many proposed solutions that jointly address all the issues at the same time. In order to characterize the static and dynamic (memory effect) nonlinear behavior of the transmitter, the multi-branch polynomial model can be

used [17]. The output complex envelope signal at the output of the nonlinear transmitter can be expressed as:

$$y(n) = \sum_{j=0}^{M} \sum_{i=1}^{N} b_{i,j} \cdot x(n-j) \cdot \left| x(n-j) \right|^{i-1} \tag{6.12}$$

where $x(n)$ and $y(n)$ are the input and output complex signal envelopes, respectively; $b_{i,j}$ are the model coefficients of the jth filter tap; and, N and M are the maximum polynomial order and memory depth, respectively.

In a MIMO transmitter, for the modeling of the coupling effects in addition to the PA nonlinearity, each output of the transmitter should contain cross terms between the input signals [3]. The memory polynomial model can be extended for a dual-input dual-output transmitter with crosstalk effects. The signal at each output of the transmitter is given by:

$$\begin{aligned} y_k(n) = & \sum_{j=0}^{M} \sum_{i=1}^{N} b_{k,1,i,j} \cdot x_1(n-j) \cdot \left| x_1(n-j) \right|^{i-1} + \\ & \sum_{j=0}^{M} \sum_{i=1}^{N} b_{k,2,i,j} \cdot x_2(n-j) \cdot \left| x_2(n-j) \right|^{i-1} \end{aligned} \tag{6.13}$$

where $x_1(n)$ and $x_2(n)$ are the input complex signals envelopes, and $y_k(n)$ is the output complex signal envelop at the k -branch of the MIMO transmitter. The modeling of the nonlinear crosstalk (memoryless case) in a MIMO transmitter is illustrated in Figure 6.17.

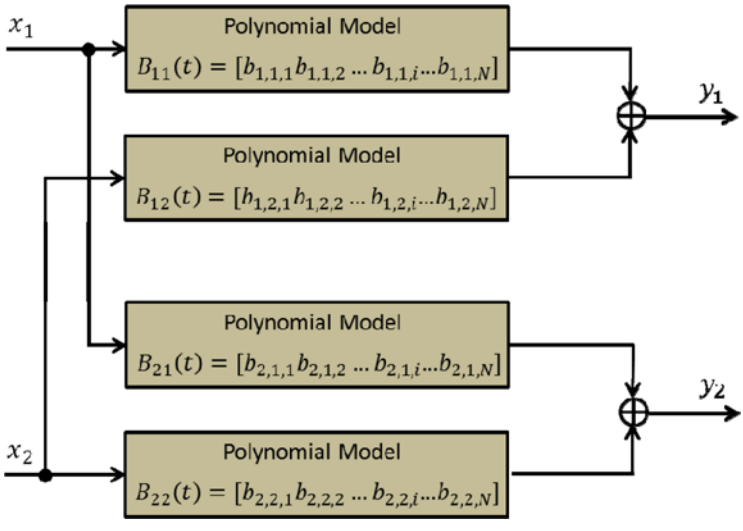


Fig. 6.17 Modeling of a 2×2 MIMO transmitter with nonlinear crosstalk

In addition to crosstalk, the I/Q imbalance of the modulator affects the performance of the MIMO transmitter. The modeling of the I/Q imbalance of the modulator is based on the modeling of the cross coupling channels between the I and Q components of the modulator signal input. The cross coupling terms can be presented in the model by utilizing the conjugate of the input signal [18]. I/Q imperfections are generally gain and phase imbalance and DC offset. The signal at the output of a SISO transmitter suffering from such impairments can be expressed as:

$$\begin{aligned} y(n) &= \sum_{j=0}^{M}\sum_{i=1}^{N} b_{i,j}\cdot x(n-j)\cdot\left|x(n-j)\right|^{i-1} \\ &+\sum_{j=0}^{M}\sum_{i=1}^{N} b'_{i,j}\cdot x^{*}(n-j)\cdot\left|x(n-j)\right|^{i-1} + b_{dc} \end{aligned} \tag{6.14}$$

where $x(n)$, $x^{*}(n)$ and $y(n)$ are the input, input conjugate and output complex signals envelopes, respectively; and, b_{dc} is a DC term used to estimate the DC offset of the modulator.

Figure 6.18 demonstrates the modeling of the modulator's I/Q imbalance for a SISO transmitter for the memoryless case.

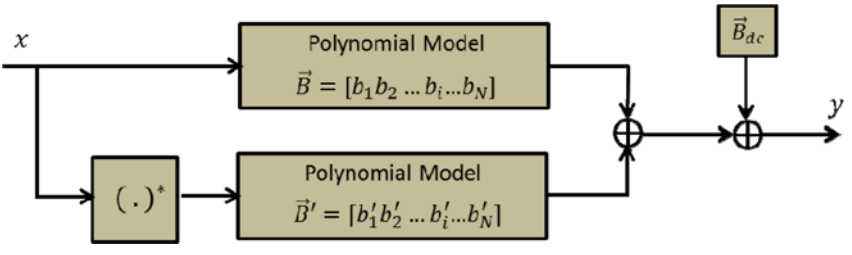


Fig. 6.18 Modeling of modulator I/Q imbalance for a SISO transmitter

Based on the prior models, a dual-input dual-output MIMO transmitter that suffers from PA nonlinearity, modulator I/Q imbalance and coupling effects can be counted as a nonlinear system with four inputs (x_1, x_2, x_1^{*}, x_2^{*}) and two outputs (y_1, y_2). The outputs can be related to the inputs as:

$$\begin{bmatrix} y_1 & y_2 \end{bmatrix}^T = \mathbf{W}\begin{bmatrix} x_1 & x_2 & x_1^{*} & x_2^{*} \end{bmatrix}^T \tag{6.15}$$

where $\mathbf{W}$ is a nonlinear matrix function that characterizes the behavior of the transmitter. The signal at the output of the MIMO transmitter suffering from PAs nonlinearity, nonlinear crosstalk, modulator I/Q imbalance and DC offset can be expressed as:

$$y_k(n) = \sum_{j=0}^{M}\sum_{i=1}^{N} b_{k,1,i,j} \cdot x_1(n-j) \cdot \left|x_1(n-j)\right|^{i-1} +$$
$$\sum_{j=0}^{M}\sum_{i=1}^{N} b'_{k,1,i,j} \cdot x_1^*(n-j) \cdot \left|x_1(n-j)\right|^{i-1} +$$
$$\sum_{j=0}^{M}\sum_{i=1}^{N} b_{k,2,i,j} \cdot x_2(n-j) \cdot \left|x_2(n-j)\right|^{i-1} +$$
$$\sum_{j=0}^{M}\sum_{i=1}^{N} b'_{k,2,i,j} \cdot x_2^*(n-j) \cdot \left|x_2(n-j)\right|^{i-1} + b_{dc}$$

Digital baseband preprocessing techniques can be used to compensate for all the effects previously mentioned. Hence, identification needs to be performed to extract the coefficients of the inverse model. The inverse model is used to generate the MIMO transmitter input when fed with the MIMO transmitter output.

References

[1] Duman, T.M., Ghrayeb, A.: Coding for MIMO Communication Systems. Wiley (2008)

[2] Mohammadi, A., Ghannouchi, F.M.: Single RF Front-End MIMO Transceivers. IEEE Communications Magazine (December 2011)

[3] Bassam, S.A., Helaoui, M., Ghannouchi, F.M.: Crossover Digital Predis-torter for the Compensation of Crosstalk and Nonlinearity in MIMO Transmitters. IEEE Transactions on Microwave Theory and Techniques 57(5), 1119–1128 (2009)

[4] Lari, M., Bassam, S.A., Mohammadi, A., Ghannouchi, F.M.: Time-Multiplexed Single Front-End Multiple-Input Multiple-Output Receivers with Preserved Diversity Gain. IET Communications 5(6), 789–796 (2011)

[5] Molisch, A., Win, M., Winters, J.: Reduced-Complexity Trans-mit/Receive-Diversity Systems. IEEE Transactions on Signal Processing 51(11), 2729–2738 (2003)

[6] Tzeng, F., Jahanian, A., Pi, D., Heydari, P.: A CMOS code-modulated path-sharing multi-antenna receiver front-end. IEEE Journal of Solid-State Circuits 44(5), 1321–1335 (2009)

[7] Vankka, J.: Digital Synthesizers and Transmitters for Software Radio. Springer, Heidelberg (2005)

[8] Bassam, S.A., Boumaiza, S., Ghannouchi, F.M.: Block-Wise Estimation of and Compensation for I/Q Imbalance in Direct-Conversion Transmitters. IEEE Transactions on Signal Processing 57(12), 4970–4973 (2009)

[9] Wiser, R.: Tunable Bandpass RF Filters for CMOS Wireless Transmitters. ProQuest (2008)

[10] Gesbert, D., Shafi, M., Shiu, D.-S., Smith, P.J., Naguib, A.: From Theory to Practice: An Overview of MIMO Space-Time Coded Wireless Systems. IEEE Journal on Selected Areas in Communications 21(3) (April 2003)

[11] Molisch, A.F., Win, M.Z., Winters, J.H.: Reduced-Complexity Transmit/Receive Diversity Systems. In: Proc. IEEE Vehicular Technology Conference, pp. 1996–2000 (May 2001)

[12] Jahanian, A., Tzeng, F., Heydari, P.: Code-Modulated Path-Sharing Multi-Antenna Receivers: Theory and Analysis. IEEE Transactions on Wireless Communications 8(5), 2193–2201 (2009)

[13] Bassam, S.A., Helaoui, M., Boumaiza, S., Ghannouchi, F.M.: Experimental Study of the Effects of RF Front-End Imperfection on the MIMO Transmitter Performance. In: Proc. IEEE MTT-S International Symposium Digest, pp. 1187–1190 (June 2008)

[14] Palaskas, Y., Ravi, A., Pellerano, S., Carlton, B.R., Elmala, M.A., Bishop, R., Banerjee, G., Nicholls, R.B., Ling, S.K., Dinur, N., Taylor, S.S., Soumyanath, K.: A 5-GHz 108-Mb/s 2 ×2 MIMO Transceiver RFIC with Fully In-tegrated 20.5-dBm P_{1dB} Power Amplifiers in 90-nm CMOS. IEEE Journal of Solid-State Circuits 41(12), 2746–2756 (2006)

[15] Tse, D., Viswanath, P.: Fundamentals of Wireless Communication. Cambridge University Press, Cambridge (2005)

[16] Medvedev, I., Bjerke, B.A., Walton, R., Ketchum, J., Wallace, M., Howard, S.: A Comparison of MIMO Receiver Structures for 802.11n WLAN—Performance and Complexity. In: The 17th Annual IEEE International Symposium on Personal, Indoor, and Mobile Radio Communications, Helsinki, Finland (September 2006)

[17] Ghannouchi, F.M., Hammi, O.: Behavioural Modeling and Predistortion. IEEE Microwave Magazine 10(7), 52–64 (2009)

[18] Anttila, L., Handel, P., Valkama, M.: Joint Mitigation of Power Amplifier and I/Q Modulator Impairments in Broadband Direct-Conversion Transmitters. IEEE Transactions on Microwave Theory and Techniques 58(4) (April 2010)

Chapter 7
Receiver Design for MIMO Wireless Communications

In the previous chapters, it has been shown that multiple input multiple out (MIMO) systems can provide either diversity order improvement or spatial multiplexing advantages. In this chapter, the receiver front-end architectures in MIMO systems are discussed. Starting with the traditional single input single output (SISO) receiver front-ends, their extensions are presented for MIMO applications. Moreover, the capacity reduction due to a noisy front-end and radio frequency interference are examined. In addition, a MIMO testbed realization technique is presented using a MIMO testbed; and, the more advanced MIMO transceiver implementation techniques are described under the commercial MIMO transceivers section.

7.1 Receiver Architectures

7.1.1 Superheterodyne Receiver

The superheterodyne receiver architecture, which is widely used in wireless communications, includes two mixers. The receiver is shown in Figure 7.1. The first mixer down-converts a radio frequency (RF) modulated signal to an intermediate frequency (IF) using a local oscillator (LO). The IF signal is generated either above or below the LO frequency. The undesired signal is called an image signal. The frequency interval between the desired and image signals is 2IF.

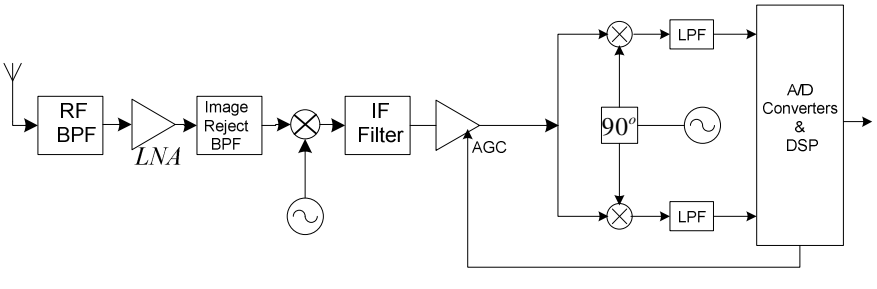


Fig. 7.1 The architecture of a superheterodyne receiver

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 151–180.
springerlink.com

It is likely that an unwanted signal is allocated in the 2IF distance of an RF signal. In this case, the unwanted signal interferes with the desired IF signal. An image rejection filter before the down-converter is used to prevent from this problem. Moreover, to remove the unwanted components, a highly selective bandpass IF filter is used at receiver. The next stage is the automatic gain control (AGC) section, which should keep the input level of the analog-to-digital converter (ADC) constant by adjusting the gain. The IF signal is then down-converted to the baseband signal using an IF oscillator. Finally, the baseband signal is processed in the digital signal processor (DSP) after analog-to-digital conversion. The DSP can demodulate any modulation format within its processing and data conversion bandwidth [1], [2].

7.1.2 Direct Conversion Receiver

The direct conversion, or zero IF, receiver is proposed to overcome the need to use of an IF section at the receiver. This architecture down-converts the RF modulated digital signal directly to a baseband signal. A typical architecture of this receiver is shown in Figure 7.2. The oscillator frequency in this receiver is the same as the carrier frequency of input signal. The channels are selected by tuning the LO and removing unwanted signals using a low-pass filter [2].

The architecture of a direct conversion receiver is simpler than that of a superheterodyne receiver. This results in a smaller sized and lower cost receiver. However, there are more design challenges in these receivers. The main design issues are due to the RF impairments, such as DC offset, I/Q imbalance, and second-order distortion. On the other hand, this architecture provides less spurious power, making it very attractive for the receiver applications.

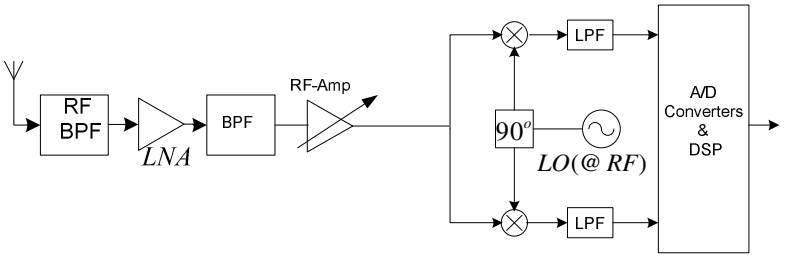


Fig. 7.2 The architecture of a direct conversion receiver

7.1.3 Low IF Receiver

The architecture of a low IF receiver is shown in Figure 7.3. This architecture has been proposed in order to take advantage of both direct conversion and superheterodyne receivers [3]. The lack of DC offset is the main advantage of a low IF receiver over the direct conversion structure. However, the image rejection poses greater difficulty compared to the superheterodyne receiver, because the image frequency is closer to the RF frequency.

The IF frequency in this architecture is usually selected from half the channel bandwidth to several times the channel bandwidth. This creates more challenges in removing the image signal. If the low IF frequency is selected to be equal to at least half the channel bandwidth, the adjacent channel interference and image interference can be filtered out by a bandpass filter known as a polyphase filter [3].

7.1.4 Image Rejection Receiver

The image reject receiver was developed to overcome the image problem in superheterodyne receivers. The most common image reject receiver architectures are Hartley and Weaver. The architecture of image reject receivers is shown in Figure 7.4.

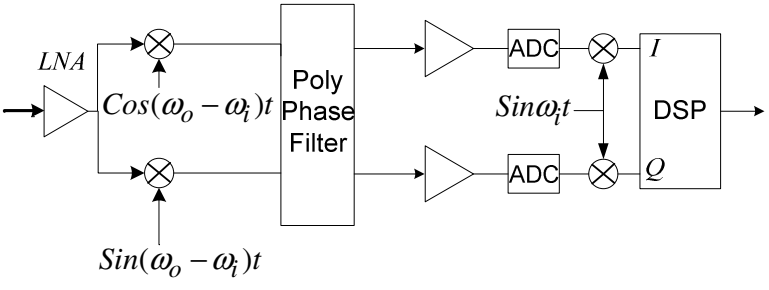


Fig. 7.3 The architecture of a low IF receiver

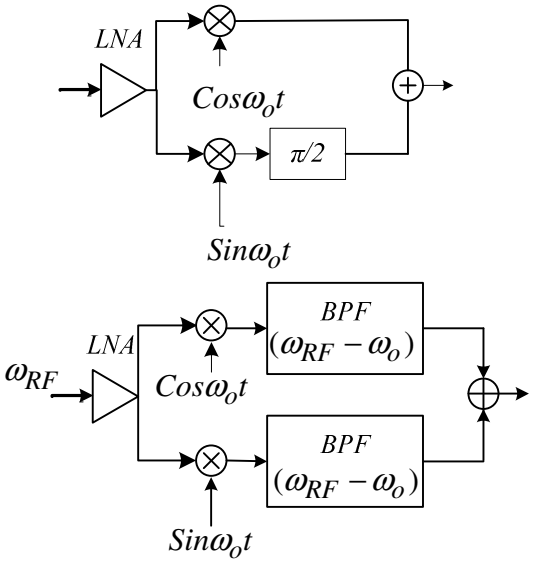


Fig. 7.4 The architecture of image rejection receivers: (top) Hartley architecture and (bottom) Weaver architecture

7.1.5 Six-Port Receiver

A six-port receiver consists of two inputs and four outputs. In direct conversion receiver applications, the RF and LO signals are applied to the two input ports, as shown in Figure 7.5. The six-port structure produces four various combinations of RF and LO signals and injects them to four power detectors, so that the outputs can be obtained [4]. If the six-port junction is a linear network, the four outputs of the multi-port structure are as follows:

$$P_i = \left| S_{i1}\overline{a}_{LO} + S_{i2}\overline{a}_{RF} \right|^2 \qquad i = 1, 2, ..., 4 \tag{7.3a}$$

$$S_{i1} = \left| S_{i1} \right| e^{j\phi_{i1}} \tag{7.3b}$$

$$S_{i2} = \left| S_{i2} \right| e^{j\phi_{i2}} \tag{7.3c}$$

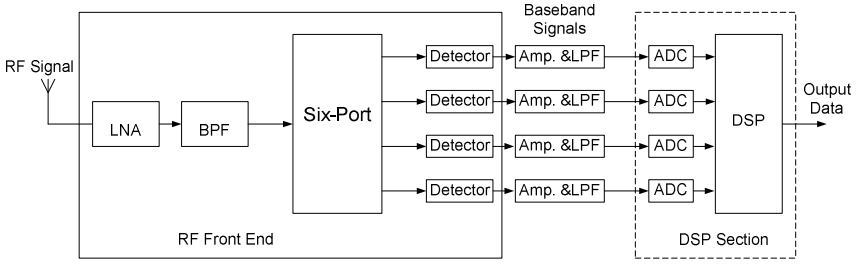


Fig. 7.5 Block diagram of a direct conversion based six-port receiver

where S_{ij} is the scattering parameter of the multi-port structure.

Assuming the RF signal with digital modulation has the same frequency as the LO signal, the RF and LO signals can be represented by:

$$a_{LO} = \frac{1}{\sqrt{2}} \left| a_{LO} \right| e^{j\varphi_{LO}} \tag{7.4}$$

$$a_{RF} = \frac{1}{\sqrt{2}} \left| a_{RF} \right| \sqrt{I^2 + Q^2}\, e^{j\varphi_{RF}} \tag{7.5}$$

where I and Q are in-phase and quadrature-phase modulating signals.

By substituting (7.4) and (7.5) into (7.3) and using the relations of:

$$I = \sqrt{I^2 + Q^2}\,Cos\left(\varphi_{RF}\right) \tag{7.6a}$$

$$Q = -\sqrt{I^2 + Q^2}\,Sin\left(\varphi_{RF}\right) \tag{7.6b}$$

the expressions of the output powers can be derived as:

$$\begin{aligned} P_i &= \frac{1}{2}\left|S_{i1}\right|^2\left|a_{LO}\right|^2 + \frac{1}{2}\left|S_{i2}\right|^2\left|a_{RF}\right|^2\left(I^2 + Q^2\right) \\ &+ \left|S_{i1}\right|\left|S_{i2}\right|\left|a_{RF}\right|\left|a_{LO}\right|\left(ICos\left(\varphi_i\right) + QSin\left(\varphi_i\right)\right) \end{aligned} \tag{7.7}$$

where $\varphi = \varphi_{RF} - \varphi_{LO}$ and $\varphi_i = \varphi_{i1} - \varphi_{i2}$.

Assuming that the LO power is known, the unknown parameters are (I^2+Q^2), I, and Q. Thus, (7.7) can be written as:

$$\begin{aligned} &P_i = M_i + L_i\left(I^2 + Q^2\right) + N_i Cos\left(\varphi_i\right) I + N_i Sin\left(\varphi_i\right) Q \\ &M_i = \frac{1}{2}\left|S_{i1}\right|^2\left|a_{LO}\right|^2,\ L_i = \frac{1}{2}\left|S_{i2}\right|^2\left|a_{RF}\right|^2,\ N_i = \left|S_{i1}\right|\left|S_{i2}\right|\left|a_{RF}\right|\left|a_{LO}\right| \\ &\varphi_i = \varphi_{i1} - \varphi_{i2},\ for \qquad i = 1, 2, \ldots, N \end{aligned} \tag{7.8}$$

where M_i, L_i and N_i are constant governing parameters for given operating conditions, which are functions of scattering parameters and the power of LO and received RF signals.

Equation (7.8) presents N equations with three unknown parameters. From (7.8), the six-port equations can be written in matrix form as:

$$\begin{bmatrix} 1 \\ \left(I^2 + Q^2\right) \\ I \\ Q \end{bmatrix} = D_T^{-1} \cdot \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix} \tag{7.9}$$

where D_T is a matrix defined as:

$$D_T = \begin{bmatrix} M_1 & L_1 & N_1 Cos(\phi_1) & N_1 Sin(\phi_1) \\ M_2 & L_2 & N_2 Cos(\phi_2) & N_2 Sin(\phi_2) \\ M_3 & L_3 & N_3 Cos(\phi_3) & N_3 Sin(\phi_3) \\ M_4 & L_4 & N_4 Cos(\phi_4) & N_4 Sin(\phi_4) \end{bmatrix} \tag{7.10}$$

If the determinant of the D_T matrix is non-zero, (7.9) can be solved with respect to unknown parameters:

$$I = \alpha_1' P_1 + \alpha_2' P_2 + \alpha_3' P_3 + \alpha_4' P_4 \tag{7.11.1}$$

$$Q = \beta_1' P_1 + \beta_2' P_2 + \beta_3' P_3 + \beta_4' P_4 \tag{7.11.2}$$

Equation (7.9) shows the general demodulation equations in six-port structures.

7.1.6 Five-Port Receiver

As seen in (7.9), there is a dependent equation that leads to a fixed value in the first element of the calculated matrix [4], [5]. This dependent equation is the result of the known LO power. This assumption is reasonable, because the LO power is known to the receiver designer and can be removed from the unknown parameters.

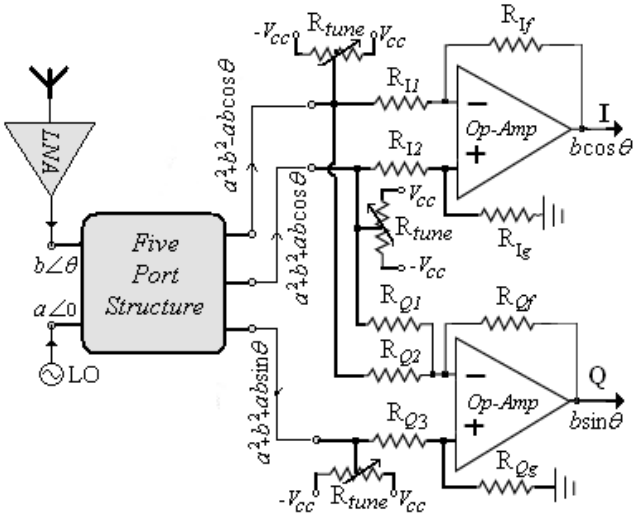


Fig. 7.6 A five-port receiver

Nevertheless, if the LO power is so unstable that it cannot be treated as a constant, the DC terms of each output can be easily removed by DC cancellation methods (in a simple case, by a capacitor), which are always used in direct conversion receivers for the DC offset cancellation. Therefore, the slow variation of the LO signal with the other DC terms is removed, and it is not necessary to know

or calibrate the LO power exactly. Note that blindly removing LO power cannot cause any problems, since the information needed for *I*/*Q* detection exists in terms of the alternating current (AC) outputs. Note that the (I^2+Q^2) parameter is variable and cannot be removed from demodulation calculations.

This is the reason for reducing the number of outputs from four to three; therefore, (7.9) may be changed to:

$$\begin{bmatrix} \left(I^2+Q^2\right) \\ I \\ Q \end{bmatrix} = D^{-1}.\begin{bmatrix} P_1 - M_1 \\ P_2 - M_2 \\ P_3 - M_3 \end{bmatrix} \tag{7.12}$$

where the matrix *D* is defined as:

$$D = \begin{bmatrix} L_1 & N_1 Cos\left(\phi_1\right) & N_1 Sin\left(\phi_1\right) \\ L_2 & N_2 Cos\left(\phi_2\right) & N_2 Sin\left(\phi_2\right) \\ L_3 & N_3 Cos\left(\phi_3\right) & N_3 Sin\left(\phi_3\right) \end{bmatrix} \tag{7.13}$$

where the determinant of matrix *D* is non-zero. Therefore, the demodulation equations will be:

$$I = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 - c_1 \tag{7.14.1}$$

$$Q = \beta_1 P_1 + \beta_2 P_2 + \beta_3 P_3 - c_2 \tag{7.14.2}$$

It should be mentioned that the additional equation in the common six-port structure is only used to simplify *I*/*Q* calculations. In this case, it is sometimes better to reduce the number of ports from six to five, if the complexity of the calculation remains the same. In regards to (7.9) and (7.12), the additional equation in the six-port architecture leads to more calibration parameters, which may increase the errors of the *I*/*Q* calculations due to errors in the calibration procedure. A low-cost generation of an *I*/*Q* signal, as shown in Figure 7.6, uses conventional Op-Amp (operational amplifier) circuits [5], although the accuracy may be slightly reduced.

7.2 Smart Antenna Receiver Architectures

The principle of operation of a smart antenna is presented in Figure 7.7 [6]. In this model, a set of antenna elements is arranged in space; and, the output of each element is multiplied by a complex weight and summed. New patterns result from combining the radiation patterns of the individual elements with phases and

amplitudes, depending on the weights applied. The weights may be allowed to vary in time, and the adaptive array is then realized. This array can be used to improve the performance of a mobile communication system by choosing the weights so as to optimize some measure of the system performance [1], [6].

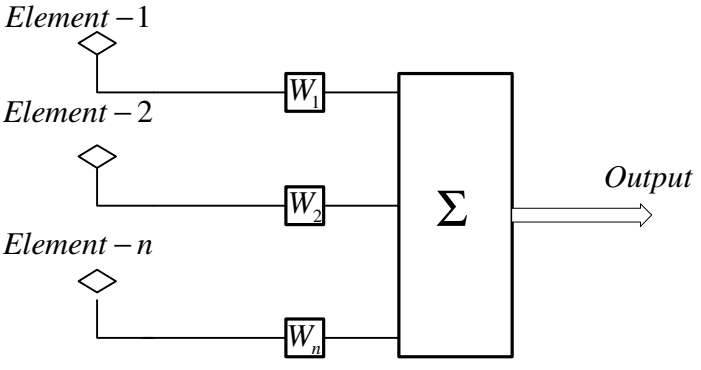


Fig. 7.7 Principle of operation of smart antennas

■ ***Example 7.1:***

This example shows how adaptive antennas at a receiver can be used to minimize the impact of interference [6]. Figure 7.8 shows a desired radiator and an interfering antenna transmitting co-channel signals s_1 and s_2, respectively. The signals are received at a receiver with two independent antenna elements. The signals received by the antenna elements are multiplied by complex weights, *w*, and summed, yielding the output, *y*, which would then be demodulated.

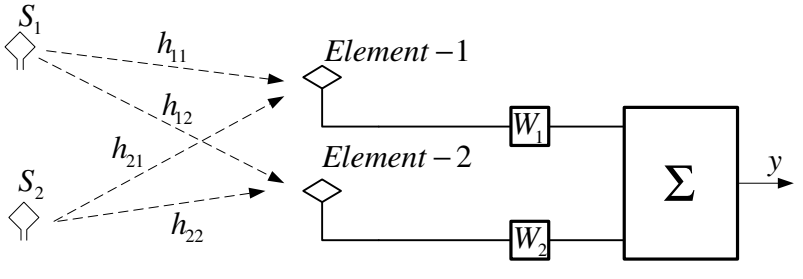


Fig. 7.8 Two elements smart antenna

It is assumed that there is complete channel state information (CSI) at the receiver, where the channel coefficients, h_{ij}, are known. One can set the weights, w_j, in such way the output y minimizes the output due to the interferer, while leaving the desired signal unaffected. The signals received at the antenna elements are:

$$x_1 = h_{11}s_1 + h_{21}s_2$$
$$x_2 = h_{12}s_1 + h_{22}s_2$$

where x_1 and x_2 are the received signals at element 1 and element 2. The output of the combiner is:

$$y = x_1w_1 + x_2w_2$$

The output can be written as:

$$y = (h_{11}s_1 + h_{21}s_2)w_1 + (h_{12}s_1 + h_{22}s_2)w_2$$
$$y = s_1(h_{11}w_1 + h_{12}w_2) + s_2(h_{21}w_1 + h_{22}w_2)$$

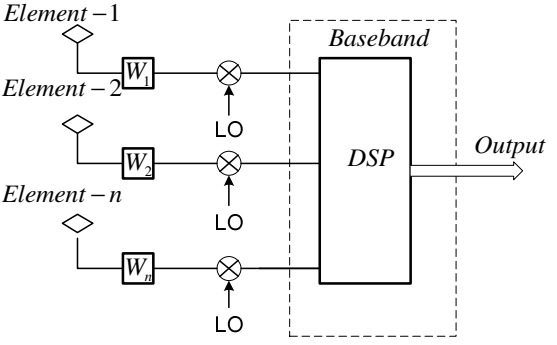
If the output must be the same as the signal from the desired mobile, the term multiplying s_1 should be set to 1 and the term multiplying s_2 must be set to 0. Accordingly, the weighting functions are obtained as [6]:

$$w_1 = \frac{h_{22}}{h_{11}h_{22} - h_{12}h_{21}}$$
$$w_2 = \frac{-h_{21}}{h_{11}h_{22} - h_{12}h_{21}}$$

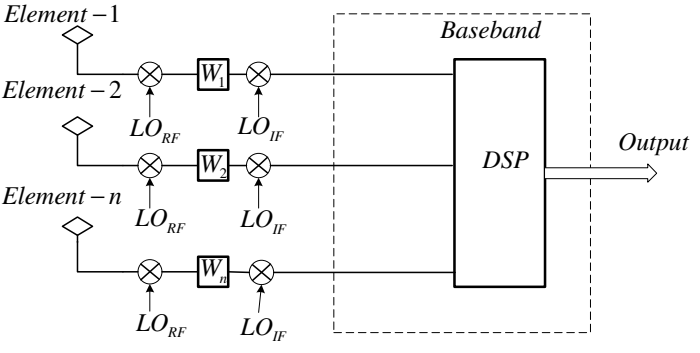
As may be seen from this example, the selection of the weighting functions is the main task in the realization of smart antenna architecture. There are different methods to select and adaptively adjust the weighting functions using DSP. The common techniques to implement the weighting functions are shown in Figure 7.9.

Figure 7.9 (a) shows the implementation of the weighting functions in RF. The weighting functions are usually realized using the variable attenuator and the variable phase shifter. This technique has also found some applications for uncompressed high-definition (HD) video transmission in a 60 GHz link and is being supported by IEEE 802.15 [7].

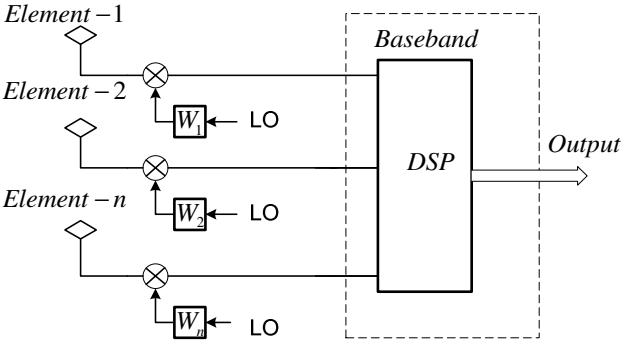
Figure 7.9 (b) shows the implementation of the weighting functions in the IF section. Figure 7.9 (c) shows the weighting functions implementation in LOs: this technique is more suitable for superheterydyne receivers. In this design, the variable gain modules and variable phase shifter are implemented in IF. The weighting function implementation in baseband is illustrated in Figure 7.9 (d). This technique is very common, due to its ease of implementation.



(a)

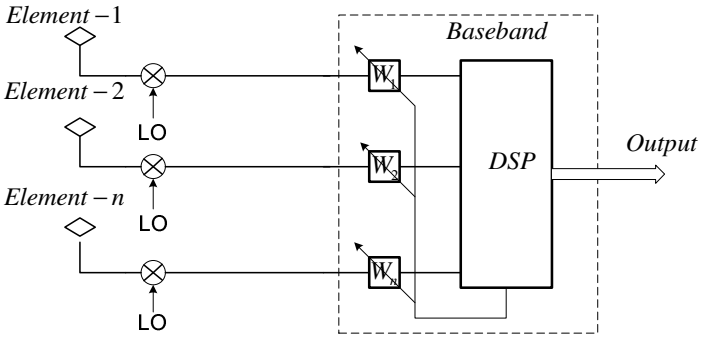


(b)



(c)

Fig. 7.9 The realization of the smart antenna receiver: (a) weighting function in RF, (b) IF weighting function in IF, (c) weighting function in LO, and d) weighting function in baseband



(d)

Fig. 7.9 *(continued)*

Figure 7.10 illustrates a simple implementation of the smart antenna and beamforming structure using weighting functions in baseband and having only an RF front-end [8]. This structure uses parasitic antenna concepts and is very promising for the provision of low-cost smart antenna and beamforming systems. The design technique for this receiver is discussed in Chapter 10.

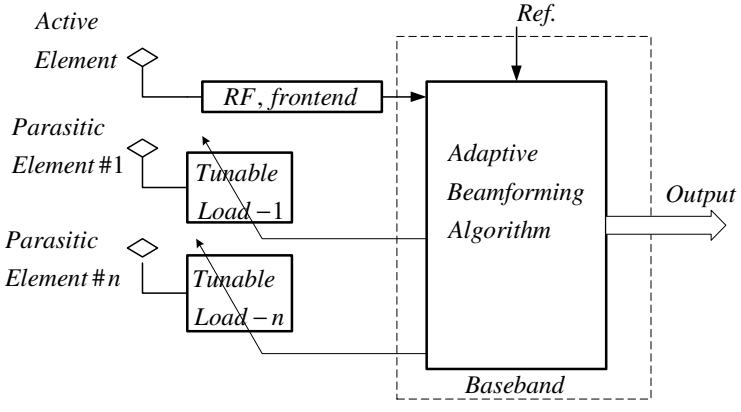


Fig. 7.10 Adaptive smart antenna receiver using parasitic elements

7.3 MIMO Receiver Architectures

The MIMO receiver front-end is traditionally realized using parallel SISO receivers. In baseband, it is usually assumed that the spatial and temporal properties of the MIMO impulse response, $H(t)$, can be separated as:

$$H(t) = u(t)H \tag{7.15}$$

where $u(t)$ is the pulse shape.

The matched filter, $H^*(-t)$, in the MIMO baseband receiver can, therefore, be decomposed into a cascade of a space-only column matrix, H^*, and followed by a bank of time-only matched filters, $u(-t)$. This is a common assumption in narrowband models [9]. It is also usually considered that the channel is known on the receiver side.

In this section, three architectures are discussed to implement an RF front-end for MIMO receivers. All these architectures use parallel SISO front-ends. There are also alternative designs that use a single RF front-end for MIMO receivers [10]. This technique is discussed in details in the last chapter.

7.3.1 Superheterodyne MIMO Front-End

A MIMO superheterodyne receiver is illustrated in Figure 7.11. As can be seen, the architecture is basically designed based on N parallel SISO receivers. However, common LOs, both in RF and IF, are used. The common LO minimizes the voltage controlled oscillator (VCO) / synthesizer power dissipation and chip area and prevents pulling between multiple VCOs on the same die [39]. The parallel down-converted streams are delivered to the baseband section using analog-to-digital sections. The signals are then detected by decomposing the spatial and temporal properties of received signals using (7.15). Furthermore, the detected signals are combined depending on their encoding format. The encoding can be used to increase the reliability, known as space-time coding, or can be used to increase the rate, referred to as spatial multiplexing.

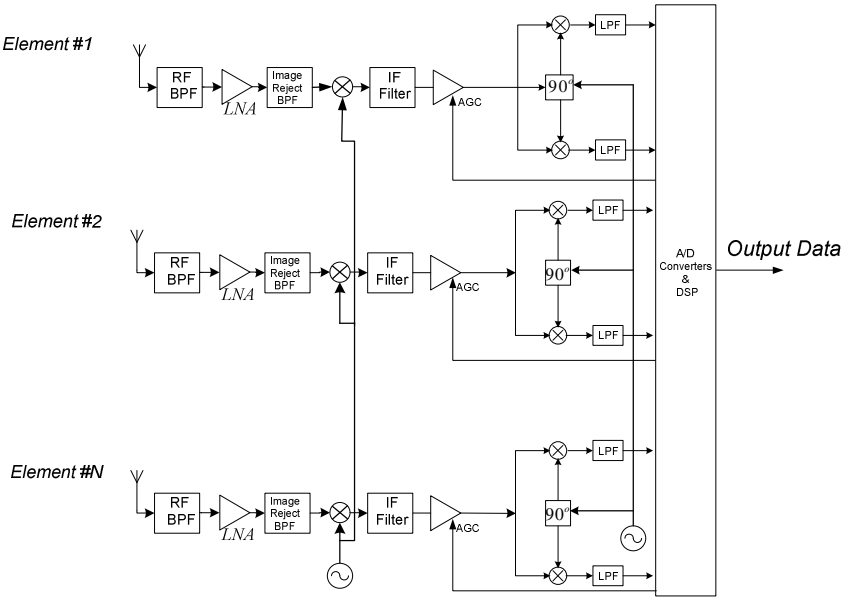


Fig. 7.11 The architecture of a MIMO superheterodyne receiver

Having N antenna elements in this architecture, N SISO front-ends are required. This makes the receiver expensive and increases the power consumption of the receiver. The RF impairments, such as phase noise and I/Q imbalance, also impact on the performance of this receiver. On the positive side, this design has the advantages of the SISO superheterodyne front-end.

7.3.2 Direct Conversion MIMO Front-End

The architecture of a MIMO direct conversion receiver is shown in Figure 7.12. This design uses an LO in the same frequency as the input RF signal. Accordingly, it removes the IF section and uses fewer components compared to superheterodyne architecture.

Moreover, the size and cost of this architecture are usually less than those of the superheterodyne design. However, the direct conversion design suffers from DC offset impairment [2]. The modeling and compensation of the DC offset impairment in direct conversion receiver is discussed in the future chapters.

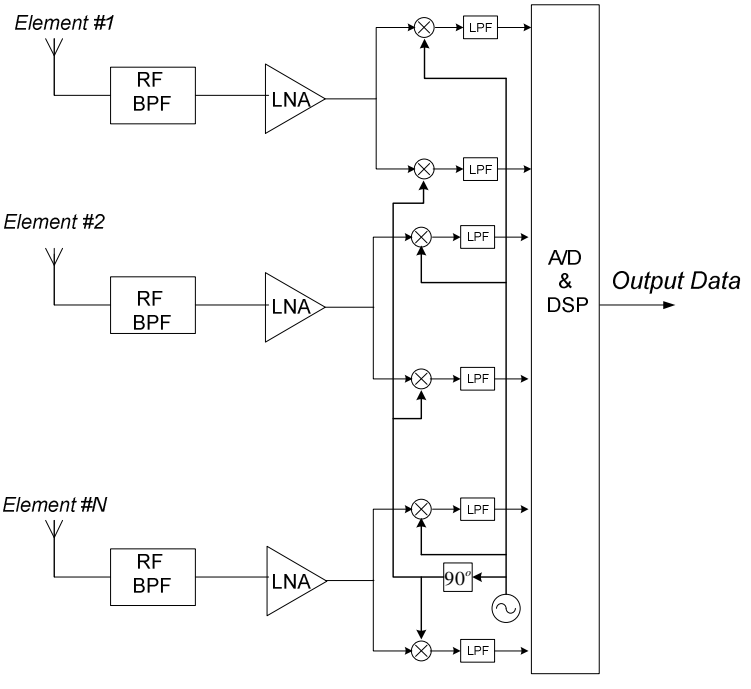


Fig. 7.12 The architecture of a MIMO direct conversion receiver

7.3.3 Six-Port MIMO Front-End

The architecture of a MIMO six-port receiver is presented in Figure 7.13. This architecture essentially belongs to the direct conversion class, where only an LO is used to convert the RF signal to baseband signals. However, due to the use of passive circuits and detector diodes instead of mixers, it promises simple and low-cost implementation, particularly at high microwave and millimeter-wave frequencies. Moreover, this technique provides lower power consumption compared to the other designs. Meanwhile, the RF impairments in this technique are similar to those in the direct conversion receivers.

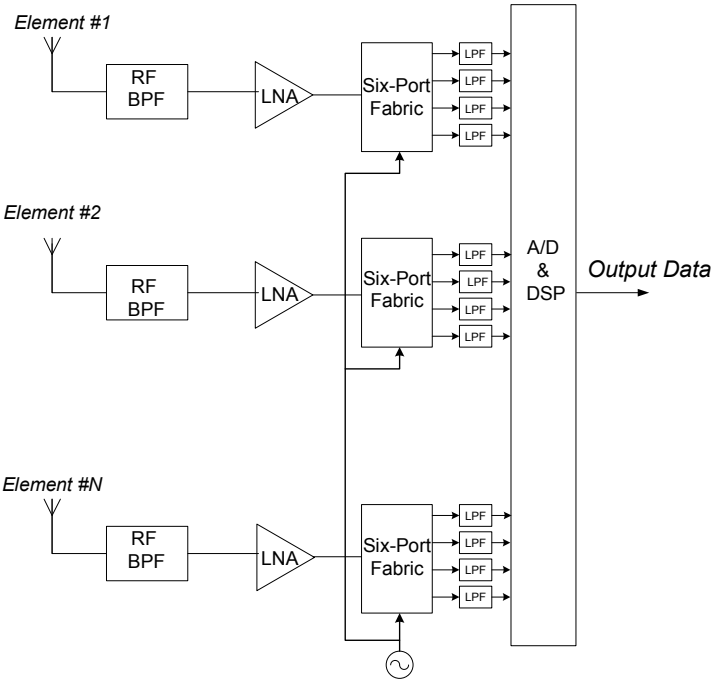


Fig. 7.13 The architecture of a MIMO six-port receiver

7.4 Capacity Reduction of MIMO System due to the Front-End

As seen in (4.32), an antenna can be modeled by a Thevenin equivalent circuit with the open circuit voltage as u and the impedance, $z_A = r_A + jx_A$. This model can be readily extended to an antenna array [11]. The open circuit array voltage is obtained as:

$$\mathbf{u} = \sum_i \mathbf{h_i} x_i + \mathbf{n_o} \tag{7.16}$$

where x_i is the transmitted signal, $\mathbf{h}_i$ is the spatial channel vector, and $\mathbf{n_o}$ is the thermal noise, and $\mathbf{n_o} \sim CN(\mathbf{0}, 4kT_oB.\mathrm{Re}[\mathbf{Z_A}])$, where B is bandwidth in Hertz and $\mathbf{Z_A}$ is the impedance matrix of the antenna array.

The front-end may include parallel sets of amplifiers, mixers and filters [13-15]. The impedance matrix of the frond-end is represented as $\mathbf{Z}$. Figure 7.14 also shows that the front-end contributes noise to the delivered signal to the load [12], [16]. Accordingly, the output signal can be related to input signal, **u**, as:

$$\mathbf{r} = \sum_{i=1}^{N} \mathbf{h}_i x_i + \mathbf{n} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{7.17}$$

where $\mathbf{n}=\mathbf{n_o}+\mathbf{z}$, $\mathbf{z}$ is the noise contribution from the front-end, $\mathbf{n} \sim CN(0, \mathbf{R_{nn}})$ and, $\mathbf{h_i} \sim CN(0, \mathbf{R_{hh}})$ where $\mathbf{R_{nn}}$ is the covariance matrix of noise, and $\mathbf{R_{hh}}$ is the covariance matrix of the channel.

For the correlated antenna, the Ergodic capacity with complete channel state information (CSI) at the receiver and power constraint $\sum_i E|x_i|^2 < P$ can be obtained as [15]:

$$C = E[\log_2 \det(1 + \mathbf{H_W^H R H_W})] \tag{7.18}$$

where

$$\mathbf{R} = \frac{P}{N} \mathbf{R_h^{1/2}} \mathbf{R_n^{-1}} \mathbf{R_h^{1/2}} \tag{7.19}$$

This matrix (7.19) is called the signal-to-noise ratio (SNR) matrix as it is analogous to the SNR of a SISO system [16]. By extending relation (4.33) of the SISO case to the MIMO system, the noise factor matrix for the MIMO front-end can be defined as:

$$\mathbf{F} = (\mathbf{R_n}|_{\mathbf{z=0}})^{-1/2} \mathbf{R_n} (\mathbf{R_n}|_{\mathbf{z=0}})^{1/2} \tag{7.20}$$

The SNR matrix may be also presented, in terms of noise factor matrix, as [16]:

$$\mathbf{R} = \frac{P}{4kT_0BN} \mathbf{R_h^{1/2}} \,\mathrm{Re}\{\mathbf{Z_A}\}^{-1/2} \mathbf{F}^{-1} \,\mathrm{Re}\{\mathbf{Z_A}\}^{-1/2} \mathbf{R_h^{1/2}} \tag{7.21}$$

It is derived that two otherwise identical MIMO systems, having front-ends with noise factor matrices of $\mathbf{F_1} \le \mathbf{F_2}$, result in corresponding capacities of $C_1 \ge C_2$, with equality if and only if $\mathbf{F_1} = \mathbf{F_2}$ [16].

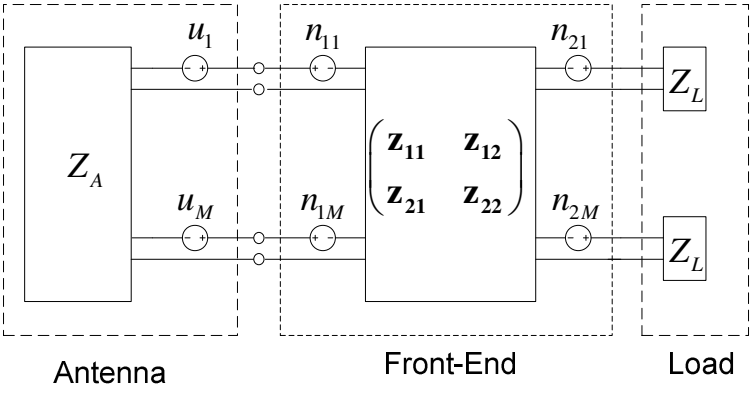


Fig. 7.14 A circuit model for a receiver front-end

7.5 Radio Frequency Interference on MIMO Receivers

Radio frequency interference (RFI) is a combination of independent radiation events that usually have non-Gaussian statistics. As an example, a computer has emissions of varying intensities that interfere with wireless transceivers. Some interfering sources from computer subsystems are shown in Table 7. 1.

Table 7.1 Example of Computer Subsystems Interfering with Wireless Standards [18]

Standard	Carrier (GHz)	Wireless Networking	Example Interfering Computer Subsystems
Bluetooth	2.4	Personal Area Network	Gigabit Ethernet, PCI Express Bus, Memory, Processor, LCD
IEEE 802.11b/g	2.4	Wireless LAN (Wi-Fi)	Gigabit Ethernet, PCI Express Bus, Memory, Processor, LCD
IEEE 802.11n	2.4	High-Speed Wireless LAN	Gigabit Ethernet, PCI Express Bus, Memory, Processor, LCD
IEEE 802.16e	2.5–2.69 3.3–3.8 5.725–5.85	Mobile Broadband (Wi-Max)	PCI Express Bus, Memory, Processor, LCD
IEEE 802.11a	5.2	Wireless LAN (Wi-Fi)	PCI Express Bus, Memory, Processor, LCD

The impact of computational platform generated RFI on wireless local area network (WLAN, IEEE 802.11b/g) data communication transceivers has been studied [19]. It was shown that computational platform generated RFI may increase the receiver noise floor ten-fold, causing a more than 50% loss in range and a significant impact on the throughput performance. As an illustration, the impact of LCD (liquid crystal display) noise on the throughput of an IEEE802.11g embedded receiver is presented in Figure 7.15 [18]. As may be seen, the LCD interference may highly reduce the throughput of the wireless system.

The common RFI model for single antenna systems are Middleton class A, B and C models [17]. RFI may have an even larger impact on MIMO transceiver performance. The extension of the single antenna RFI models to MIMO system is not a very straightforward task. Figure 7.16 shows the performance degradation of a typical MIMO receiver in the presence of RFI.The RFI model is an additive bivariate Middleton class A noise model [20]. In this model, the energy of the Middleton Class A model is a sum of the Gaussian and impulsive components of the noise. At high SNRs, the receivers become sensitive to the impulsive component of RFI, thereby causing severe degradation of performance due to RFI.

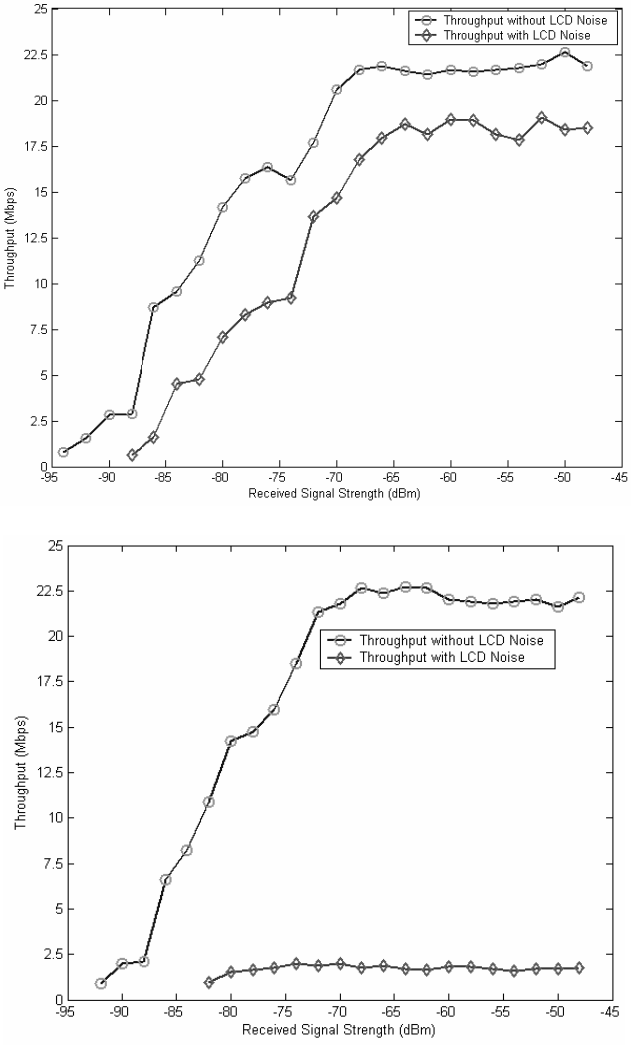


Fig. 7.15 The impact of LCD interference on throughput of an IEEE 802.11g receiver: (top) channel 1 (bottom) channel 7 © IEEE [19]

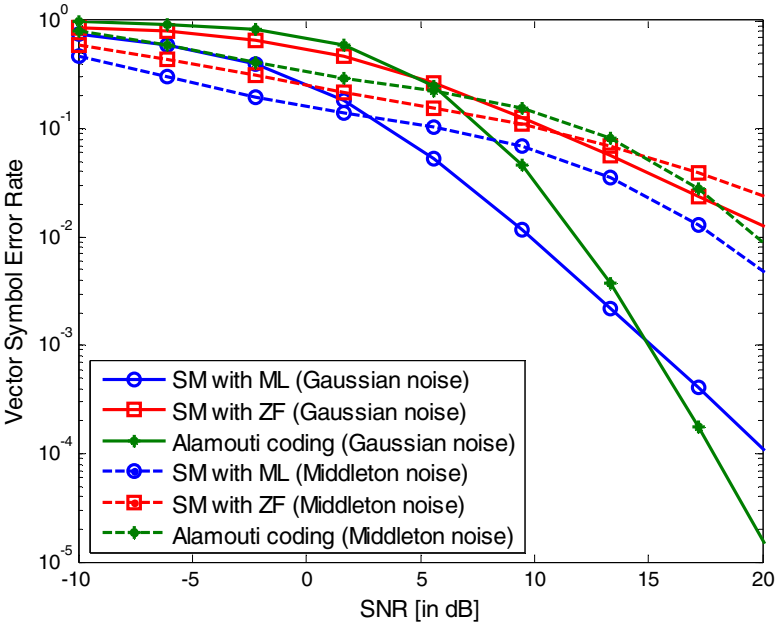


Fig. 7.16 The performance of spatial multiplexing (SM) and Alamouti coding MIMO receivers in presence of Gaussian and bivariate Middleton class A noise © Springer [20]

7.6 MIMO Testbed Design

Wireless MIMO communications have necessitated a variety of performance examinations; therefore, a MIMO testbed, which shows the system performance and examines the analytical results, is required in research laboratories in both academia and industry. Moreover, the performance degradation of the system over realistic and imperfect channels needs to be accurately characterized. The first narrowband MIMO prototype was vertical BLAST (Bell Laboratories Layered Space-Time) or V-BLAST (Vertical BLAST) architecture [21]. This was followed by several MIMO testbeds that have been reported in the literature, e.g. [22]-[32]. A modem MIMO testbed design needs flexibility in software, since firmware implementation of different transmission schemes, such as space-time processing, is typically tedious work. In this section, a typical MIMO test-bed is described.

7.6.1 Testbed Architecture

Figure 7.17 shows a functional block diagram of a testbed. For a given data stream, a program in a PC builds up frames by using a desired transmission scheme. These frames are then transmitted to the field-programmable gate array

(FPGA) board and are buffered in its internal RAM (random access memory). After some processing, these signals are converted to analog signals by four digital-to-analog converters (DACs). The analog converted signals are fed to the RF transmitters.

On the receiver side, the received signals are down-converted by means of four synchronous superheterodyne receivers. The digitized signals are fed to the FPGA. An energy detector is implemented in the FPGA to detect the frame position [33]. The received frames are sent to the PC for processing.

The testbed specifications are presented in Table 7.2, and the various sections of the testbed are described in the following subsections.

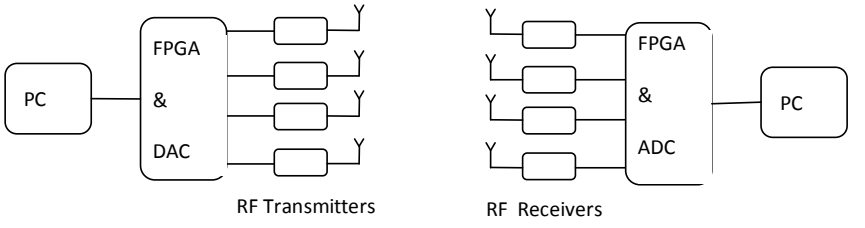


Fig. 7.17 Functional block diagram of the testbed

Table 7.2 Testbed specifications

Parameter	Description
Radio Frequency	2.4~2.5 GHz
Bandwidth	<4 MHz
Modulation	Any
Channel	Up to 4×4
TX DAC Resolution	12 bit
RX ADC Resolution	12 bit
TX Output power	<18 dBm
TX IF	21.4 MHz
RX IF1	268 MHz
RX IF2	10.7 MHz
Connection type	USB

7.6.1.1 RF Transmitter

The RF transmitter module for a path is shown in Figure 7.18. The RF transmitter consists of an I/Q mixer, a common oscillator for all modulators and a power amplifier with 18 dBm maximum output power. The input signal to the transmitter block is fed by a DAC. A common LO with a power level of -7 dBm is used to provide the oscillator signals for all paths. It is shown in Figure 7.19.

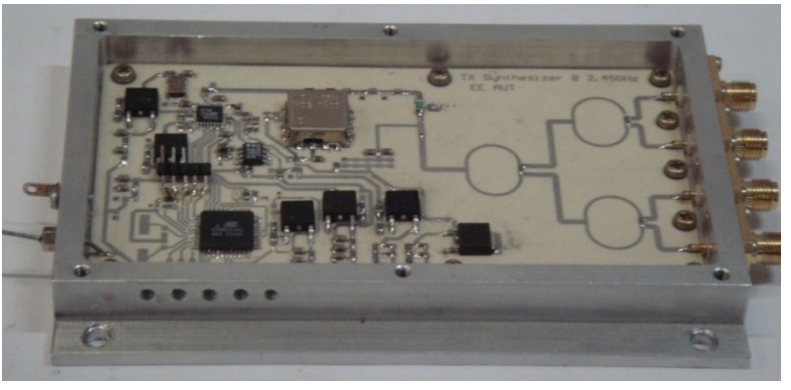


Fig. 7.18 The transmitter module

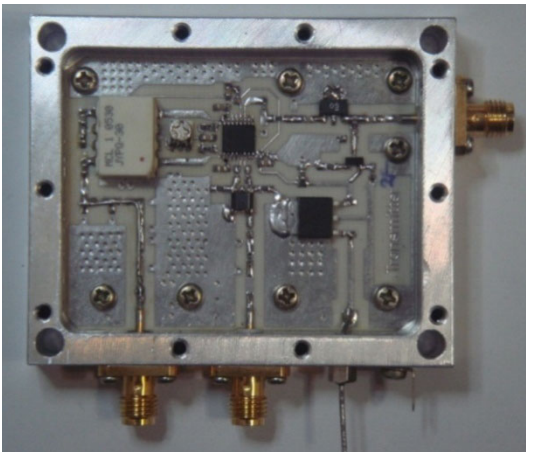


Fig. 7.19 Transmitter oscillator in 2358.6 MHz

7.6.1.2 RF Receiver

The RF receiver diagram is shown in Figure 7.20 The RF receiver consists of a band-pass filter (BPF), a low-noise amplifier (LNA), an image rejection (IMR) filter, an RF to IF down-converter, two common oscillators for all down-converters and an automatic gain control (AGC) amplifier. The first stage of the receiver is a micro strip BPF filter to protect the following LNA from blocking by other radio applications. An IMR filter is used to additionally suppress image frequency signals.

The down-conversion to both the first and second IF frequencies is performed by a double-balanced mixer. The LO signals with a power level of +12 and +7 dBm are applied for the first and second mixers, respectively. A bandpass SAW (surface acoustic wave) filter is applied at the first IF stage to limit the noise bandwidth of the RF system to 4 MHz and to filter undesired spectral components. An AGC amplifier with 68 dB dynamic range is in the last stage of the receiver. The 10.7 MHz IF signal feeds the ADC. The RF receiver module is shown in Figure 7.21.

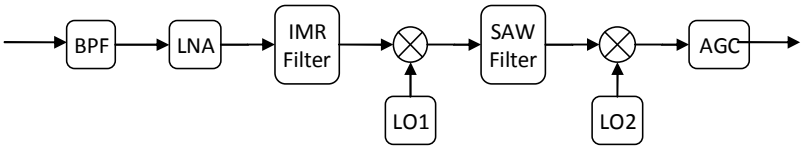


Fig. 7.20 Receiver functional diagram

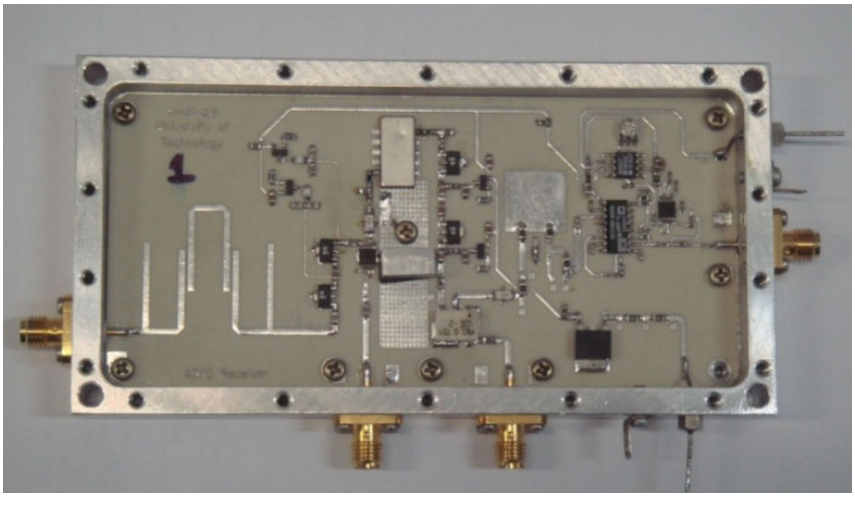
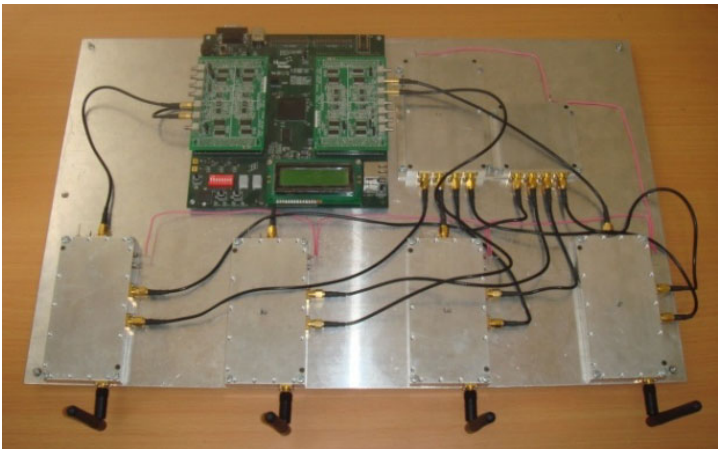
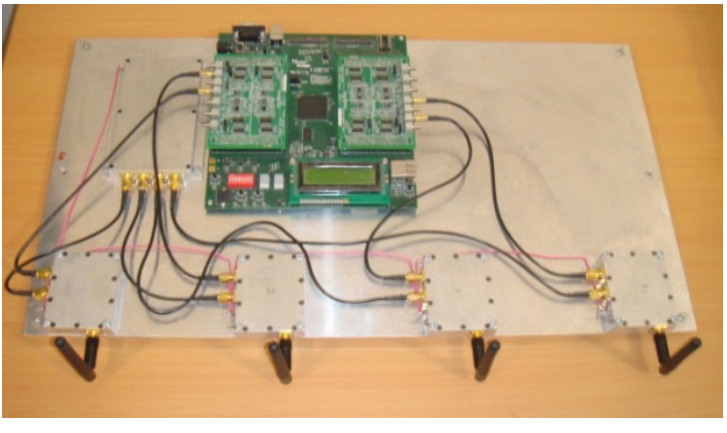


Fig. 7.21 Receiver module

The developed testbed is shown in Figure 7.22 In addition, a channel simulator can be implemented in MATLAB software and is placed after MIMO transmitter. The overall system is shown in Figure 7.23. This testbed is used in the future chapters to investigate the performance of MIMO transceivers.



(a)



(b)

Fig. 7.22 Block diagrams of (a) the RF receiver, and (b) the RF transmitter

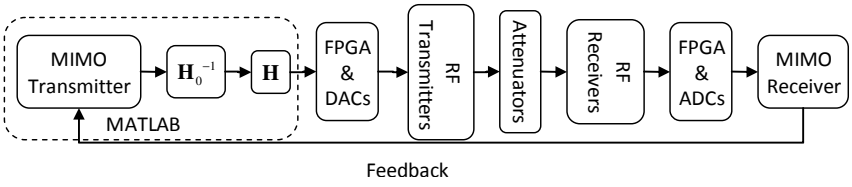


Fig. 7.23 Block diagram of the overall testbed

7.7 Commercial MIMO Transceivers

Some manufacturers have recently introduced a few MIMO transceivers to the wireless market. These transceivers usually target commercial applications, including WiMAX (Worldwide Interoperability for Microwave Access) and LTE (Long Term Evolution).

7.7.1 UXA234 MIMO Transceiver from NXP-Philips

The functional block diagram of the UXA234 MIMO transceiver from NXP-Philips is shown in Figure 7.24. The UXA234 covers the frequencies from 2.3 GHz to 3.8 GHz using conversion transceiver structures. A direct conversion architecture is used to achieve the lowest power consumption and the least number of external components [34]. On the receiver side, the signal enters at the LNA, is mixed down in an I/Q mixer to a zero-IF. It is then low-pass filtered and buffered before being fed to the ADC (located in the baseband chip). On the transmitter side, I/Q signals are created by the current mode DAC of the baseband integrated circuit (IC) and converted on to the transceiver's IC to a voltage. The signal is then passed through reconstruction filters before being mixed up to the RF and amplified. The LO signal is shared between the receiver and transmitter.

On the receiver side, the high- and mid-gain modes of the LNA are implemented with active circuitry, while the low-gain mode is passive. Over the whole receiver band, the RF front-end requires a very low noise figure of 3 dB. The circuit has been fabricated with SiGe (silicon-germanium) BiCMOS technology [3]. Without calibration, the receiver's error vector magnitude (EVM) is reported at about -32dB and, after calibration, is about - 38dB [35].

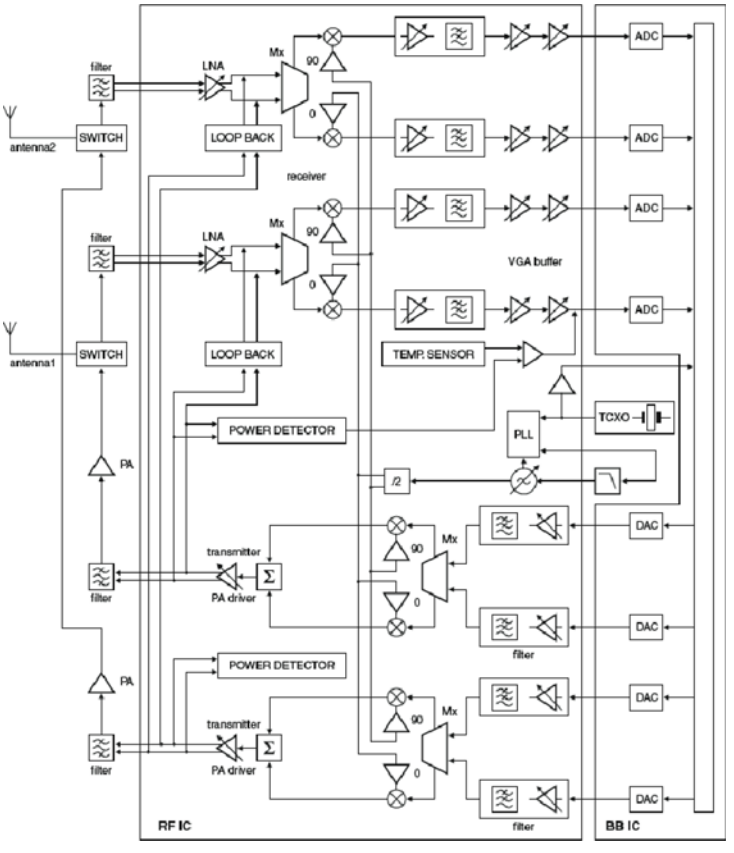


Fig. 7.24 Functional block diagram for the UXA234 MIMO transceiver © NXP-Philips [35]

7.7.2 MAX2842 MIMO Transceiver from MAXIM

The functional block diagram of the MAX2842 MIMO transceiver is shown in Figure 7.25 [36]. The MAX2842 is a direct conversion zero-IF RF transceiver that operates from 3.3 GHz to 3.9 GHz. It has two transmitters and two receivers and may be used to realize a 2x2 MIMO system. The receiver noise figure is 3.8 dB. This transceiver supports channel bandwidths of 3.5 MHz, 5 MHz, 7 MHz, and 10 MHz.

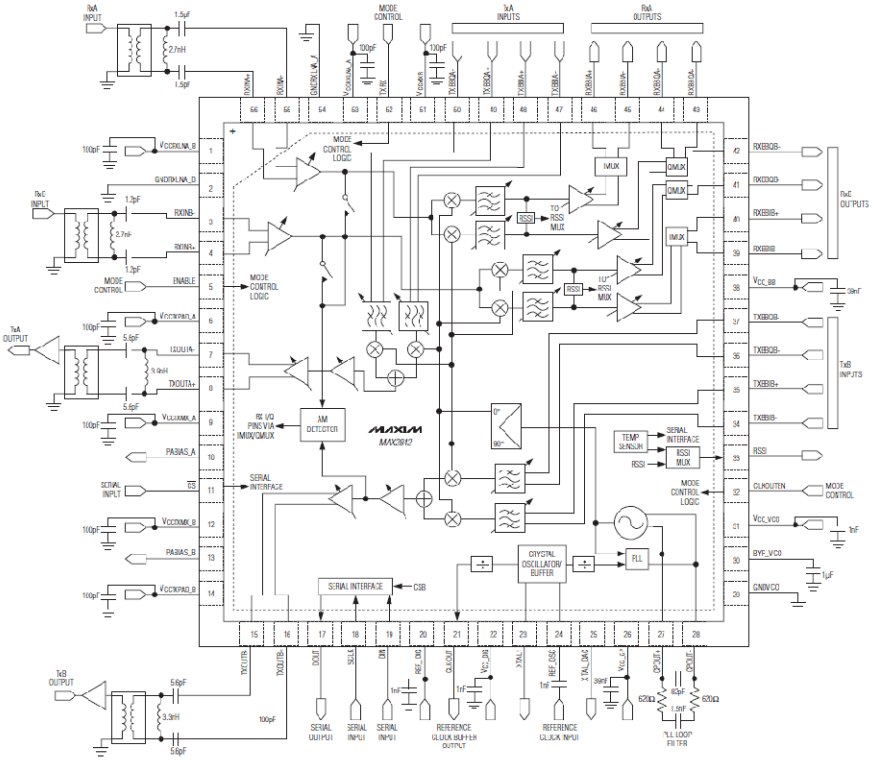


Fig. 7.25 Functional block diagram for the MAX2842 MIMO transceiver © MAXIM [36]

The IC integrates an on-chip AM (amplitude modulation) detector for measuring the transmitter's I/Q imbalance and LO leakage. An internal transmit-to-receive loopback mode allows for receiver I/Q imbalance calibration. The MAX2842 completely eliminates the need for external SAW filters by implementing on-chip programmable monolithic filters for both the receiver and transmitter for channel bandwidths from 3.5MHz to 10MHz [36].

7.7.3 PM8800 MIMO Transceiver from PMC-Seirra

The functional block diagram of the PM8800 MIMO transceiver is shown in Figure 7.26 [37]. The PM8800 2x2 is a zero-IF MIMO transceiver. It has two independent radio paths (2 transmitters and 2 receivers) and supports operation in the 2.3 GHz to 2.9 GHz band or the 3.3 GHz to 3.8 GHz band.

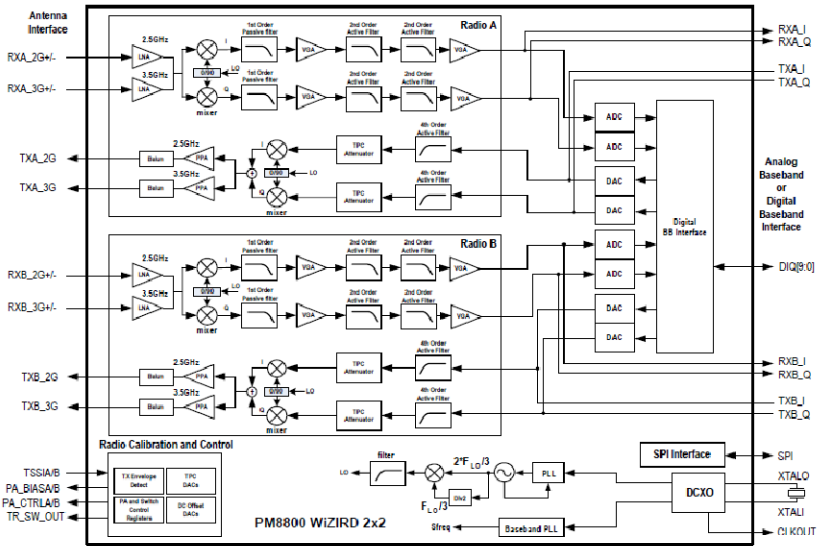


Fig. 7.26 Functional block diagram for the PM8800 MIMO transceiver © PMC-Seirra [37]

The EVM in the transmitting path is about -36 dB at output power of 0 dBm (2.5 GHz) and -34 dB at output power of 0 dBm (3.5 GHz). Moreover, the EVM in the receiver path is about -36 dB at 2.5 GHz and -34 dB at 3.5 GHz. The noise figure of the receiver is about 3 dB. Moreover, the adjacent channel power ratio (ACPR) for 64QAM (quadrature amplitude modultion) is about 26 dB [37].

7.7.4 AD9356 MIMO Transceiver from Analog Devices

A functional block diagram of the ADS9356 MIMO transceiver is shown in Figure 7.27 [38]. The AD9356 is an RF transceiver with dual receivers and transmitters and operates in the 2.3 GHz to 2.7 GHz range. This transceiver supports channel bandwidths of 3.5 MHz, 4.375 MHz, 5 MHz, 7 MHz, 8.75 MHz, and 10 MHz. The receiver uses direct conversion architecture. The transmitting path has excellent spectral purity with sideband noise of less than -130 dBc/Hz at an 8 MHz frequency offset and offers an EVM of -40 dB. Two high dynamic range ADCs, followed by decimation and channel filters, digitize the received signals and produce 12-bit output signals at a sample rate determined by the bandwidth mode.

The transmitting path takes 12-bit input data and interpolates before converting to the analog domain and up-converting to the carrier frequency [38]. This chip can be used in MIMO base station design by adding suitable power amplifiers and duplexers. This is presented in Figure 7.28.

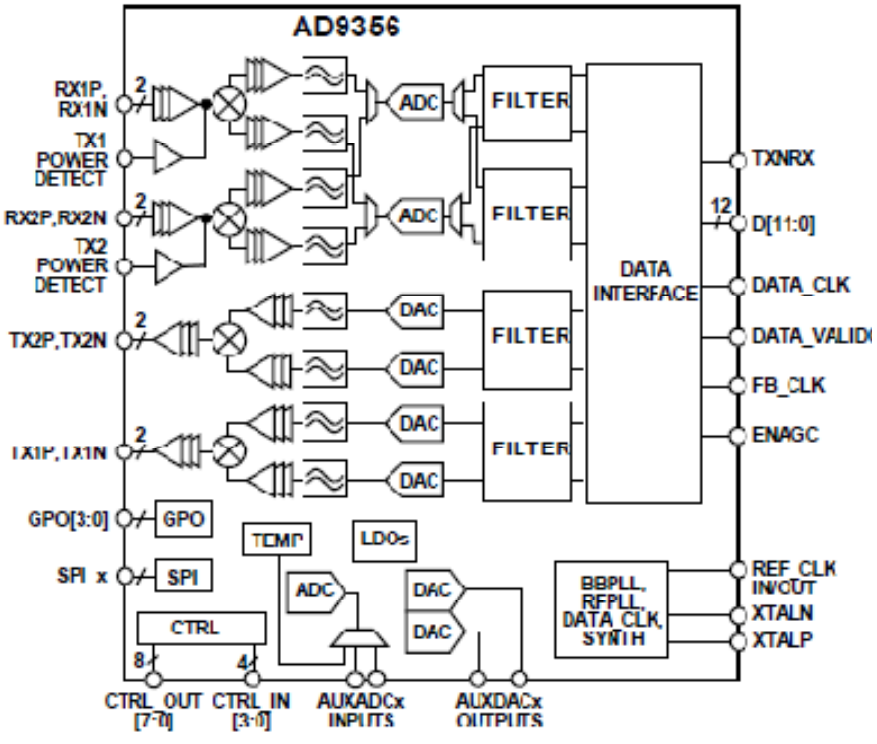


Fig. 7.27 Functional block diagram for the AD9356 MIMO transceiver © Analog Devices [38]

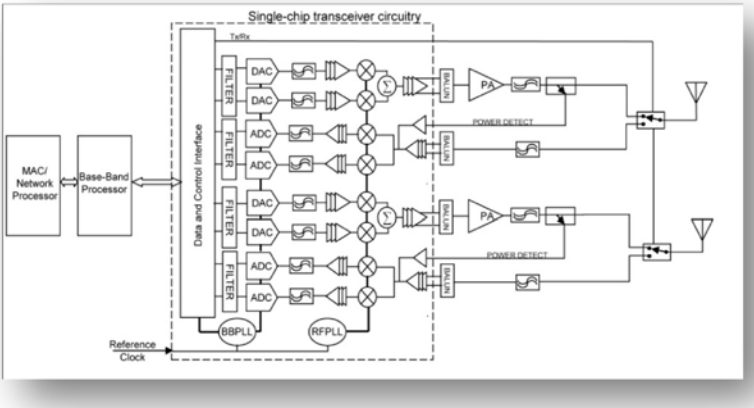


Fig. 7.28 2×2 MIMO base station transceiver architecture using the AD9356 MIMO transceiver ©Analog Device [40].

References

[1] Ellinger, F.: Radio Frequency Integrated Circuits Technology, 2nd edn. Springer, Heidelberg (2008)
[2] Razavi, B.: RF Microelectronics. Prentice Hall (1998)
[3] Crols, J., Steyaert, M.S.J.: Low-IF Topologies for High-Performance Ana-log Front Ends of Fully Integrated Receivers. IEEE Transactions on Circuits and Systems II 45, 269–282 (1998)
[4] Ghannouchi, F.M., Mohammadi, A.: Six-port Technique with Microwave and Wireless Applications. Artech House (2009)
[5] Mirzavand, R., Mohammadi, A., Ghannouchi, F.M.: Five-port receivers, architectures and applications. IEEE Communications Magazine 48(6), 30–36 (2010)
[6] Saunders, S.R., Aragon-Zavala, A.: Antenna and Propagation for Wireless Communications, 2nd edn. John Wiley & Sons (2007)
[7] Kraemer, R., Katz, M.D.: Short Range Wireless Communications, Emerging Technologies and Applications. Wiley (2009)
[8] Chandran, S.: Advances in Direction-of-Arrival Estimation. Artech House (2006)
[9] Barry, J., Lee, E.A., Messerschmitt, D.G.: Digital Communications, 3rd edn. Kluwer Academic (2004)
[10] Lari, M., Bassam, A., Mohammadi, A., Ghannouchi, F.M.: Time-Multiplexed Single Front-End MIMO Receivers with Preserved Diversity Gain. IET Communications (2010)
[11] Balanis, C.A.: Antenna Theory: Analysis and Design, 3rd edn. John Wiley (2005)
[12] Rothe, H., Dahlke, W.: Theory of noisy four poles. In: Proceeding of IRE, vol. 44, pp. 811–818 (June 1956)
[13] Gans, M.J.: Channel capacity between arrays – Part I: Sky noise dominates. IEEE Transactions on Communications 54, 1586–1592 (2006)
[14] Gans, M.J.: Channel capacity between arrays – Part II: Amplifier noise dominates. IEEE Transactions on Communications 54, 1983–1992 (2006)
[15] Shiu, D., et al.: Fading correlation and its effect on the capacity of multielement antenna systems. IEEE Transactions on Communications 48, 502–513 (2000)
[16] Domizioli, C.P., Hughes, B.L., Gard, K.G., Lazzi, C.: Optimal front-end design for MIMO receivers. In: IEEE Global Communications Conference, Globcom 2008, New Orleans (2008)
[17] Middleton, D.: Non-Gaussian noise models in signal processing for telecommunications: New methods and results for Class A and Class B noise models. IEEE Transactions on Information Theory 45(4), 1129–1149 (1999)
[18] Nassar, M., Gulati, K., DeYoung, M.R., Evans, B.L., Tinsley, K.R.: Mitigating near-field interference in laptop embedded wireless transceivers. Journal of Signal Processing Systems (March 2009)
[19] Shi, J., Bettner, A., Chinn, G., Slattery, K., Dong, X.: A study of platform EMI from LCD panels - impact on wireless, root causes and mitigation methods. In: IEEE International Symposium on Electromagnetic Compatibility, August 14-18, vol. 3, pp. 626–631 (2006)

[20] Gulati, K., Chopra, A., Heath Jr., R.W., Evans, B.L., Tinsley, K.R., Lin, X.E.: MIMO Receiver Design in the Presence of Radio Frequency Interference. In: Proc. IEEE Global Communications Conference, November 30 - December 4, New Orleans (2008)

[21] Wolniansky, P.W., Foschini, G.L., Golden, G.D., Valenzuela, R.A.: V-Blast: an architecture for realizing very high data rates over the richscattering wireless channel. In: Proceedings of International Symposium on Signals, Systems, and Electronics (ISSSE 1998), Pisa, Italy, pp. 295–300 (September- October 1998)

[22] Sampath, H., Talwar, S., Tellado, L., Erceg, V., Paulraj, A.: A fourth-generation MIMO-OFDM broadband wireless system: Design, performance, and field trial results. IEEE Communications Magazine 40(9), 143–149 (2002)

[23] Murphy, P., Lou, F., Sabharwal, A., Frantz, L.P.: An FPGA based rapid prototyping platform for MIMO systems. In: Conference Record of the 37th Asilomar Conference on Signals, Systems, and Computers, ASILOMAR 2003 (November 2003)

[24] Adjoudani, A., Beck, E.C., Burg, A.P., et al.: Prototype experience for MIMO BLAST over third-generation wireless system. IEEE Journal on Selected Areas in Communications 21(3), 440–451 (2003)

[25] Aschbacher, E., Langwieser, R., Caban, S., Mehlfuehrer, S., Caban, G., Keirn, W., Scholtz, A.L., Rupp, M.: Design and subsystem verification of a flexible and scalable 4X4 MIMO testbed. In: IEEE Radio and Wireless Conference (RAWCON 2004), Workshop on MIMO Implementation Aspects (Septembeer 2004)

[26] Garrett, D., Woodward, G., Davis, L., Knagge, G., Nicol, C.: A 28.8 Mb/s 4x4 MIMO 3G highspeed downlink packet access receiver with normalized least mean square equalization. In: Proceedings of IEEE International Solid-State Circuits Conference (ISSCC 2004), San Francisco, vol. 1, pp. 420–536 (February 2004)

[27] Borkowski, D., Bruhl, L.: Hardware implementation for real-time multi-user MIMO systems. In: Proceedings of IEEE Radio and Wireless Conference (RAWCON 2004) Workshop 2 on MIMO Implementation Aspects, Atlanta (September 2004)

[28] Bialkowski, K.S., Zagriatski, S., Postula, A., Bialkowski, M.E.: MIMO Test-bed with an insight into signal strength distribution around transmitter/receiver sites. In: Proceedings of EuMW 2005, Paris, pp. 117–120 (October 2005)

[29] Liberti, J.C., Koshy, J.C., Hoerning, T.R., Martin, C.C., Dixon, J.L., Triolo, A.A., Murray, R.R., McGiffen, T.G., Military, T.: Experimental results using a MIMO test bed for wideband, high spectral efficiency tactical communications. In: IEEE Communications Conference, MILCOM, October 17-20, vol. 3, pp. 1340–1345 (2005)

[30] Bialkowski, K.S., Postula, A., Abbosh, A., Bialkowski, M.E.: 2x2 MIMO Testbed for Dual 2.4GHz/5GHz Band. In: International Conference on Electromagnetics in Advanced Applications, ICEAA 2007, September 17-21, pp. 1–4 (2007)

[31] Chiu, C.Y., Cheng, C.H., Wan, Y.S., Rowell, C.R., Murch, R.D.: Design of a Flat Fading 4 x 4 MIMO Testbed for Antenna Characterization using a Modular Approach. In: IEEE Conference on Wireless Communications and Networking, Kowloon, pp. 2913–2918. Dept. of Electron. Comput. Eng, Hong Kong Univ. of Sci. Technol. (March 2007)

[32] Spring, R., Zhou, L., Gogate, N., Daryoush, A.S.: 4x4 MIMO Experimental Test-bed using COTS at ISM Band. In: IEEE Radio and Wireless Symposium, pp. 173–176. Dept. of ECE, Drexel Univ., Philadelphia, PA (2007)

[33] Azami, F., Ghorssi, A., Hemesi, H., Mohammadi, A., Abdipour, A.: Design and Implementation of a Flexible 4x4 MIMO Testbed. In: IEEE International Symposium on Telecommunications, IST 2008, Isfehan (2008)

[34] Locher, M., Tomesen, M., Kuenen, J., Daanen, A., Visser, H., Essink, B., Vervoort, P.P., Nijrolder, M., Kopmeiners, R., Redman-White, W., Balmford, R., El Waffaoui, R.: A Low Power, High Performance BiCMOS MIMO/Diversity Direct Conversion Transceiver IC for WiBro/WiMAX (802.16e). In: IEEE 2007 Custom Intergrated Circuits Conference, CICC (2007)
[35] NXP Philips Inc., WiMAX 802.16e MIMO transceivers UXA234xx (2007)
[36] Maxim Integrated Product, 3.3GHz to 3.9GHz MIMO Wireless Broadband RF Transceiver, 2842 data sheet (2010)
[37] PMC Seirra Inc., PM8800 WiZIRD 2x2, Released Product Brief (2008)
[38] Analog Devices Inc., WiMAX/BWA/WiBRO/LTE RF MxFE 2 × 2 MIMO Transceiver, AD9356 data sheet (2010)
[39] Palaskas, Y., Ravi, A., Pellerano, S., Carlton, B.R., Elmala, M.A., Bishop, R., Banerjee, G., Nicholls, R.B., Ling, S.K., Dinur, N., Taylor, S.S., Soumyanath, K.: A 5-GHz 108-Mb/s 2x2 MIMO Transceiver RFIC With Fully Integrated 20.5-dBm P1dB Power Amplifiers in 90-nm CMOS. IEEE Journal of Solid-State Circuits 41(12), 2746–2756 (2006)
[40] Wier, P.: Synchronizing NxN MIMO Basestations to an External Timing Reference. Analog Devices Inc. (2010)

Chapter 8
RF Impairments in OFDM Transceivers

Orthogonal frequency-division multiplexing (OFDM) is a very attractive transmission technique for future high data rate wireless multimedia communication systems. OFDM has been developed for wideband wireless digital communication and is used in applications such as Digital Video and Audio Broadcasting (DVB, DAB, respectively), wireless networking and broadband Internet access [1]-[3].

OFDM has advantages, such as high spectral efficiency, easy adaptation to severe channel conditions without complex equalization, and efficient implementation using fast Fourier transform (FFT). However, it also has some disadvantages, such as sensitivity to frequency offset, a high peak-to-average-power ratio (PAPR) and sensitivity to noises [1], [4], [5].

As discussed in Chapter 3, OFDM is successfully used in MIMO systems. This chapter is devoted to investigate the various RF impairments, specifically noise and nonlinearity, and their mutual impacts on OFDM transceivers.

8.1 OFDM Transceivers

Orthogonal frequency-division multiplexing is a frequency-division multiplexing (FDM) scheme utilized as a digital multicarrier modulation method. A large number of closely spaced orthogonal subcarriers are used to carry data, which are divided into several parallel data streams or channels, one for each subcarrier. Each sub-carrier is modulated with a conventional modulation scheme, such as quadrature amplitude modulation (QAM) or phase-shift keying (PSK) at lower symbol rates, maintaining total data rates similar to conventional single-carrier modulation schemes in the same bandwidth [2], [20]. Figure 8.1 shows a block diagram of an OFDM transceiver system.

The input data is first de-multiplexed into N parallel streams, and each one is mapped to a (possibly complex) symbol stream using some modulation constellation (QAM, PSK, etc.). An inverse FFT is computed on each set of symbols, giving a set of complex time-domain samples. These samples are then quadrature-mixed to passband in the standard way. The real and imaginary components are first converted to the analog domain using digital-to-analog converters (DACs); and, the analog signals are then used to modulate cosine (real) and sine (imaginary) waves at the carrier frequency, f_c. These signals are then summed to give the transmission signal, $S(t)$.

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 181–226.
springerlink.com

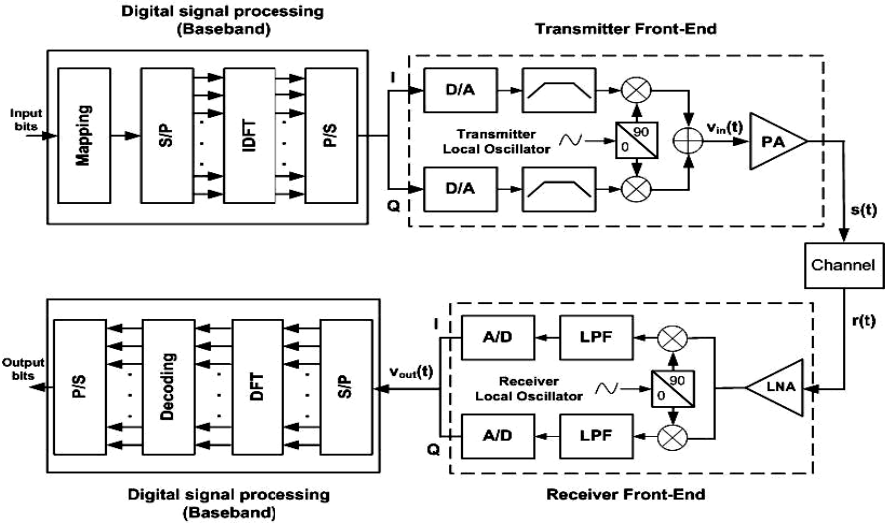


Fig. 8.1 Block diagram of OFDM communication system

The receiver picks up the signal, $r(t)$, which is then quadrature-mixed down to baseband using cosine and sine waves at the carrier frequency. This also creates signals centered on $2f_c$, so low-pass filters are used to reject these. The baseband signals are then sampled and digitized using analog-to-digital converters (ADCs), and a forward FFT is used to convert back to the frequency domain. This returns N parallel streams, each of which is converted to a binary stream using an appropriate symbol detector. These streams are then re-combined into a serial stream, which is an estimate of the original binary stream at the transmitter [1], [20].

8.2 Noise in OFDM Transceivers

It is very important to exactly predict and analyze the oscillator noise influences in OFDM communication systems and quantify the tolerable level of noise. Phase noise effects in OFDM have been analyzed in many papers by several authors [4], [5], [7]-[19]. However the effect of amplitude noise was not considered and always ignored relative to phase noise. In this section, the theoretical analysis of the impact of the oscillator phase and amplitude noise on the signal-to-noise ratio (SNR) and, hence, the bit error rate (BER) performance of MQAM-OFDM signals over an additive white Gaussian noise (AWGN) channel are presented; and, the tolerable levels of phase and amplitude noise are extracted.

8.2.1 Phase and Amplitude Noise

An oscillator is a system that generates a periodic signal with a specified or controllable frequency. There are the two kinds of noise in radio frequency (RF) oscillator output, namely amplitude noise and phase noise. The output voltage of an oscillator can be represented by [6]:

$$V_{out}(t) = (A_0 + a_N(t)).Cos(2\pi f_0 t + \varphi_N(t)) \tag{8.1}$$

$$V_{out}(t) = A(t).Cos(2\pi f_0 t + \varphi(t)) \tag{8.2}$$

where A_0 is the average amplitude of the output signal, f_0 is nominal frequency of oscillation, and $a_N(t)$ and $\varphi_N(t)$ represent the time-varying components of the amplitude and phase, respectively. These components are considered noise, because an ideal oscillator would have constant amplitude A_0 and phase varying at a constant rate of $2\pi f_0$.

It can be shown that the white noise power of the amplifier electronics is equally partitioned into amplitude and phase noise components with a zero mean and variances of σ_A^2 and σ_φ^2, respectively [13]. The SNR performances that can be achieved in an OFDM system with a noisy oscillator can be approximated by integrating its phase and amplitude noise power density functions (PDFs) over the channel bandwidth. The phase noise PDF of an oscillator is approximated by a Lorenzian function with uniform phase distribution. It is parameterized by its total integrated phase noise, *K*, and -3dB bandwidth, *B*, to which we superimpose a noise floor (set at 35dB above the thermal noise), as shown in (8.3) [5]:

$$L(f)^2 = \frac{1}{\pi}.\frac{K.B}{f^2 + B^2} + L_0 \tag{8.3}$$

$$L(f)^2 = \frac{1}{\pi}.\frac{K.B}{f^2 + B^2} + L_0 \tag{8.3}$$

Figure 8.2 shows phase noise PDFs for different *K* and a -3dB bandwidth of 100 Hz. The amplitude noise PDF of an oscillator is similar to the phase noise PDF, but is, on average, *n* dB less than it is in spectrum bandwidth, where *n* for a practical parameter of an oscillator is normally between 10 to 20 dB relative to the oscillator's quality [21]. Therefore, the phase and amplitude noise variances are:

$$\sigma_\varphi^2 = 2\int_0^{B_s/2} L(f)^2\, df = 2\frac{1}{\pi}\int_0^{B_s/2} (\frac{K\,B}{f^2 + B^2} + L_0)\, df = \frac{2K}{\pi} arctg(\frac{B_s}{2B}) + \frac{1}{\pi} L_0 B_s \tag{8.4a}$$

$$\sigma_A^2 = \sigma_\varphi^2 \times 10^{-\frac{n}{10}} \tag{8.4b}$$

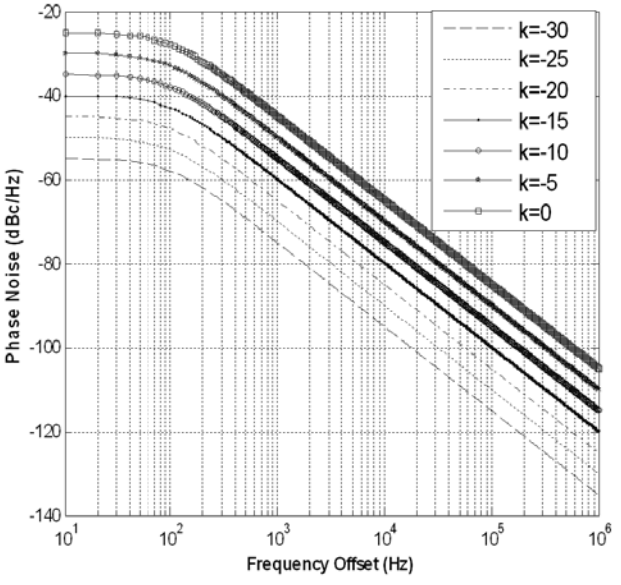


Fig. 8.2 Phase noise PDFs for different K, B = 100 Hz.

Table 8.1 shows the phase and amplitude noise variances of an oscillator for different K and n. In the next section, we use these values for calculating MQAM-OFDM degradation caused by phase and amplitude noise in simulation. Noise includes components at many frequencies, so its phase with respect to the main carrier is random, and its amplitude is also random.

Figure 8.3 shows noise added to the carrier phasor. Random amplitude and phase variation due to noise changes the constellation of M-QAM modulated signal and degrade the performance of the systems.

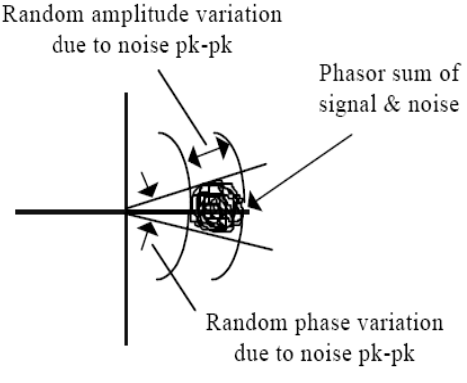


Fig. 8.3 Noise added to the carrier phasor [22]

Table 8.1 Phase and Amplitude Noise Variances for Different K, n

B_s(Hz)		10^7				
K		**-30**	**-20**	**-10**	**-5**	**0**
σ_φ^2		**0.0010**	**0.0100**	**0.1**	**0.3158**	**1**
σ_A^2	***n* = 100**	1e-13	1e-12	3.158e-11	1e-10	1e-10
	***n* = 20**	0.00001	0.0001	0.0032	0.0100	0.0100
	***n* = 15**	0.000032	0.00032	0.0100	0.0316	0.0316
	***n* = 10**	0.0001	0.0010	0.0316	0.1000	0.1000
	***n* = 5**	0.00032	0.0032	0.1000	0.3162	0.3162

8.2.2 Analysis of Phase and Amplitude Noise Impacts in OFDM

As shown in Figure 8.1, if N subcarriers are used, and each subcarrier is modulated using M alternative symbols, the OFDM symbol alphabet consists of M^N combined symbols. The low-pass equivalent OFDM signal is expressed as [1]:

$$S(t) = \sum_{k=0}^{N-1} X_k e^{j2\pi kt/T} \qquad 0 \le t < T \tag{8.5}$$

where $j = \sqrt{-1}$, X_k is the data symbol for the kth subcarrier, N is the number of subcarriers, and T is the OFDM symbol time.

The subcarrier spacing of $1/T$ makes them orthogonal over each symbol period. $S(t)$ is corrupted by the phase and amplitude noise of the transmitter's local oscillator (LO). The received signal is also influenced by the phase and amplitude noise of the receiver's LO. It can be assumed that the phase and amplitude noises of the LOs at the transmitter and the receiver are independent and identical. Therefore, it is expressed as follows:

$$r(t) = \{S(t).A_{TX}(t).e^{j\varphi_{TX}(t)} + n(t)\}.A_{RX}(t).e^{j\varphi_{RX}(t)} \tag{8.6}$$

where $n(t)$ is the complex Gaussian noise, and $\varphi_{TX}(t), A_{TX}(t)$ & $\varphi_{RX}(t), A_{RX}(t)$ are the time-varying phase and amplitude noise processes generated by the RF LO at the transmitter (TX) and the receiver (RX), respectively.

The sampled signal for the kth subcarrier after the FFT processing stage in the receiver can be written as:

$$\begin{aligned} Y_k &= \frac{1}{N}\sum_{m=0}^{N-1} r(m)e^{-j\frac{2\pi}{N}km} \\ &= \frac{1}{N}\sum_{m=0}^{N-1} S(m)A_{TX}(m)A_{RX}(m)e^{j(\varphi_{TX}(m)+\varphi_{RX}(m))}e^{-j\frac{2\pi}{N}km} \\ &\quad + \frac{1}{N}\sum_{m=0}^{N-1} n(m)A_{RX}(m).e^{j\varphi_{RX}(m)}e^{-j\frac{2\pi}{N}km} \\ &= \frac{1}{N}\sum_{m=0}^{N-1}(\sum_{l=0}^{N-1} X_l e^{j\frac{2\pi}{N}lm}).A(m).e^{j\varphi(m)}e^{-j\frac{2\pi}{N}km} + N_K \\ &= \frac{1}{N}\sum_{l=0}^{N-1} X_l \sum_{m=0}^{N-1} A(m)e^{j\varphi(m)}e^{-j\frac{2\pi}{N}(l-k)m} + N_K \\ &= \sum_{l=0}^{N-1} X_l Q_{l-k} + N_K \end{aligned} \tag{8.7}$$

where N_K is a sampled FFT version of the AWGN noise multiplied by the phase and amplitude noise of the RX LO with a variance of σ_n^2. $A(m)$ and $\varphi(m)$ are defined as :

$$\begin{cases} A(m) \triangleq A_{TX}(m)A_{RX}(m) \\ \varphi(m) \triangleq \varphi_{TX}(m) + \varphi_{RX}(m) \end{cases} \tag{8.8}$$

and Q_k is given by :

$$Q_k = \frac{1}{N}\sum_{m=0}^{N-1} A(m)e^{j\varphi(m)}e^{j\frac{2\pi}{N}km} \tag{8.9}$$

The received signal is composed of three contributions: Y_{k_1} is the kth desired subcarrier, Y_{k_2} are the other subcarriers, and N_k is the AWGN [7], [9], [10], [13]-[15]:

$$Y_k = Y_{K_1} + Y_{k_2} + N_k \tag{8.10}$$

or:

$$Y_k = X_k.Q_0 + \sum_{l=0,l\neq k}^{N-1} X_l.Q_{l-k} + N_k \tag{8.11}$$

The expected value of $X_k.Q_0$ is the desired signal component; therefore, the desired signal power can be found as:

$$P_{s-desired} = [E[X_k Q_0]]^2 = P_s[E[Q_0]]^2 \tag{8.12}$$

where:

$$E[Q_0] = \frac{1}{N}\sum_{m=0}^{N-1} EA(m)e^{j\varphi(m)} = \frac{1}{N}.N.EA(m)e^{j\varphi(m)} = EA(m)Ee^{j\varphi(m)} = e^{-\frac{\sigma_\varphi^2}{2}} \tag{8.13}$$

$$Ee^{j\varphi(m)} = \int_{-\infty}^{+\infty} e^{j\varphi}.\frac{1}{\sqrt{2\pi}.\sigma_\varphi}.e^{-\frac{\varphi^2}{2\sigma_\varphi^2}} d\varphi = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}.\sigma_\varphi}.e^{\frac{\varphi^2}{2\sigma_\varphi^2}+j\varphi} d\varphi = e^{-\frac{\sigma_\varphi^2}{2}} \tag{8.14}$$

Therefore:

$$P_{s-desired} = P_s.e^{-\sigma_\varphi^2} \tag{8.15}$$

The variance of $X_k.Q_0$ is the phase and amplitude noise power caused by common phase error (CPE); and, P_{CPE} is calculated as follows:

$$P_{CPE} = Var[X_k Q_0] = P_s.Var[Q_0] \tag{8.16}$$

where:

$$\begin{aligned} Var[Q_0] &= Var[\frac{1}{N}\sum_{m=0}^{N-1} A(m)e^{j\varphi(m)}] = \frac{1}{N^2}\sum_{m=0}^{N-1} Var(A(m)e^{j\varphi(m)}) \\ &= \frac{1}{N}(E\mid A(m)\mid^2 .\mid e^{j\varphi}\mid^2 - A_0^2.e^{-\sigma_\varphi^2}) \\ &= \frac{1}{N}((1+\sigma_A^2) - e^{-\sigma_\varphi^2}) \end{aligned} \tag{8.17}$$

Therefore:

$$P_{CPE} = P_s \times \frac{1}{N}((1+\sigma_A^2) - e^{-\sigma_\varphi^2}) \tag{8.18}$$

Next, $\sum_{l=0,l\neq k}^{N-1} X_l.Q_{l-k}$ is the interference component caused by ICI (inter-subcarrier interference). So, P_{ICI} can be calculated as:

$$P_{ICI} = P_s.\{Var[\sum_{l=0}^{N-1} Q_l] - Var[Q_0]\} \tag{8.19}$$

where:

$$\begin{aligned} Var[\sum_{l=0}^{N-1} Q_l] &= Var[\sum_{l=0}^{N-1}\{\frac{1}{N}\sum_{m=0}^{N-1} A(m)e^{j\varphi(m)}e^{j\frac{2\pi}{N}lm}\}] \\ &= \frac{1}{N^2}Var[\sum_{l=0}^{N-1}\{\sum_{m=0}^{N-1} A(m)e^{j\varphi(m)}e^{j\frac{2\pi}{N}lm}\}] \\ &= (1+\sigma_A^2 - e^{-\sigma_\varphi^2}) \end{aligned} \tag{8.20}$$

Therefore:

$$P_{ICI} = P_s.\frac{N-1}{N}.(1+\sigma_A^2 - e^{-\sigma_\varphi^2}) \tag{8.21}$$

N_k is a sampled FFT version of the AWGN noise multiplied by the phase and amplitude noise of the RX LO, as seen in (7). Hence, for the sake of simple analysis, we consider $Var[N_k]$ as that of the AWGN [38]:

$$P_{with\{\varphi,A\}} = Var[N_k] = P_n \tag{8.22}$$

From (8.12) to (8.22), the SNR, including the phase and amplitude noise, can be arranged as :

$$\begin{aligned} (\frac{S}{N})_{With\{\varphi,A\}} &= \frac{P_S.e^{-\sigma_\varphi^2}}{(P_{CPE}+P_{ICI})+P_n} = \frac{P_S.e^{-\sigma_\varphi^2}}{P_S(1+\sigma_A^2-e^{-\sigma_\varphi^2})+P_n} \\ &= \frac{(\frac{S}{N})_{Without\{\varphi,A\}}e^{-\sigma_\varphi^2}}{1+(\frac{S}{N})_{Without\{\varphi,A\}}(1+\sigma_A^2-e^{-\sigma_\varphi^2})} \end{aligned} \tag{8.23}$$

The above-mentioned powers can be used to evaluate the performance degradation of the OFDM system in the presence of phase and amplitude noise of LOs. Therefore, the degradation factor (DF) is calculated as follows:

$$DF = 10\log\frac{(\frac{S}{N})_{Without\{\phi,A\}}}{(\frac{S}{N})_{With\{\phi,A\}}} = 10\log\{(1+(1+\sigma_A^2 - e^{-\sigma_\phi^2}).SNR_{Without\{\phi,A\}}) / e^{-\sigma_\phi^2}\} \tag{8.24}$$

In (8.24), σ_φ^2 is the variance of random variables φ_{TX} and φ_{RX}, and σ_A^2 is the variance of random variables A_{TX} and A_{RX}. These are calculated by integrating the power spectrum density of the total output phase and amplitude noise.

8.2.3 Impacts of Phase and Amplitude Noise on OFDM Systems

This section presents a performance analysis of an OFDM communication system influenced by the phase and amplitude noise of the RF LO in an AWGN channel. It is difficult to compensate for the phase and amplitude noise in the OFDM system. Therefore, it is important to quantify the permissible levels of phase and amplitude noise in the LO to maintain system performance and to design the LO with limited maximum phase and amplitude noise.

The BER formula for a system performance analysis can be found by the SNR with phase and amplitude noise, instead of the SNR without phase and amplitude noise. Equations (8.25) and (8.26) shows the probability of the bit error per carrier and the probability of the bit error for QAM modulation, respectively.

$$P_{bc} = \frac{2}{k}(1-\frac{1}{\sqrt{M}})Q(\sqrt{\frac{3k}{(M-1)}\cdot\frac{E_b}{N_0}}) \tag{8.25}$$

where

$$Q(x) = \frac{1}{2}erfc(\frac{x}{\sqrt{2}})$$

$$P_b = 1-(1-P_{bc})^2 = 2P_{bc} - {P_{bc}}^2 \tag{8.26}$$

where k is the number of bits per symbol, M is the number of symbols in the modulation constellation, E_b is the energy per bit, and N_0 is the noise power spectral density. Equations (8.27) to (8.30) are the fundamental BER formulas that can be used in the system using the Gray-encoded 4,16,64,256QAM modulation method and the coherent detection technique in the AWGN channel [23].

$$P_{4-QAM} = \frac{1}{2} erfc(\sqrt{SNR_{with(\varphi,A)}}) - \frac{1}{16} erfc^2(\sqrt{SNR_{with(\varphi,A)}}) \tag{8.27}$$

$$P_{16-QAM} = \frac{3}{8} erfc(\sqrt{\frac{2}{5} SNR_{with(\varphi,A)}}) - \frac{9}{256} erfc^2(\sqrt{\frac{2}{5} SNR_{with(\varphi,A)}}) \tag{8.28}$$

$$P_{64-QAM} = \frac{7}{24} erfc(\sqrt{\frac{4}{21} SNR_{with(\varphi,A)}}) - \frac{49}{4096} erfc^2(\sqrt{\frac{4}{21} SNR_{with(\varphi,A)}}) \tag{8.29}$$

$$P_{256-QAM} = \frac{15}{64} erfc(\sqrt{\frac{4}{85} SNR_{with(\varphi,A)}}) - \frac{225}{16384} erfc^2(\sqrt{\frac{4}{85} SNR_{with(\varphi,A)}}) \tag{8.30}$$

where $SNR_{with\{\phi,A\}}$ is the SNR with phase and amplitude noise in the OFDM system as in (8.23).

The phase and amplitude noise variances, $\sigma_\phi^2, \sigma_A^2$, achieved in the previous section can be used to analyze the BER performance in OFDM system with phase and amplitude noise. Figures 8.4 to 8.7 show the theoretical and simulation BER performances of an OFDM communication system employing 4,16,64,256QAM modulations in the AWGN channel [38]. IER stands for irreducible error rate.

Detailed results for a BER of 10^{-6}, every phase noise variance and different amplitude noise order relative to it are shown in Table 8.2. As shown in Figures 8.4 to 8.7, the relative SNR penalty (degradation) to attain a BER of 10^{-6} for different amplitude noise orders, n, are presented in Table 8.3.

The results from Tables 8.2 and 8.3 shows that, if the amplitude noise order relative to phase noise, $\boldsymbol{n}$, is higher than 15 dB (15-20 dB), we can ignore the amplitude noise effect in the performance analysis. However, if it is less than 15 dB (10-15 dB), we must consider the amplitude noise for an exact analysis and determination of the tolerable level of amplitude noise.

Figure 8.8 shows the combination of two effects (CPE, ICI) on a 4QAM-OFDM signal after demodulation caused by phase and amplitude noise. The first scatter plot (a) corresponds to the demodulated signal when the variance of the phase noise that affects the OFDM signal is zero. The second scatter plot (b) corresponds to the demodulated signal when the variance of the phase noise is 0.0316 and the amplitude noise order relative to the phase noise is n = 100 dB (ideal). The third scatter plot (c) corresponds to the demodulated signal when the variance of

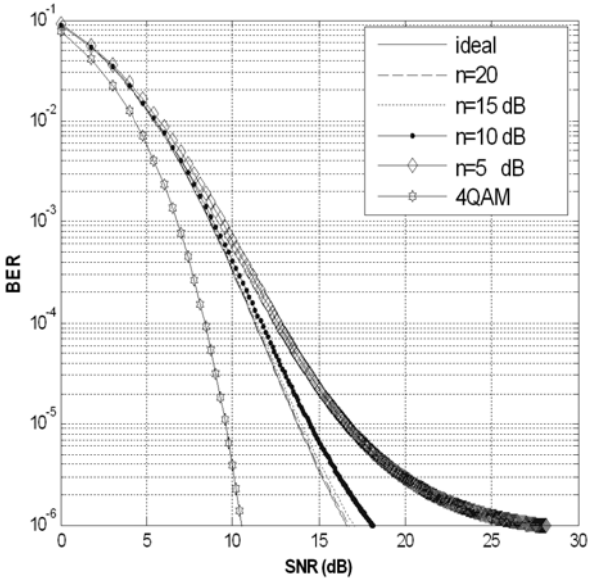


Fig. 8.4 4QAM-OFDM system with ($\sigma_{\varphi}^2 = 0.0316$)

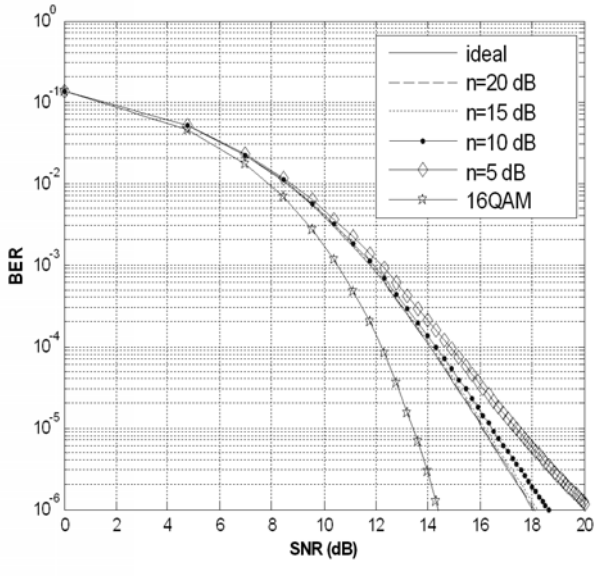


Fig. 8.5 16QAM-OFDM system with amplitude and phase noise ($\sigma_{\varphi}^2 = 0.01$)

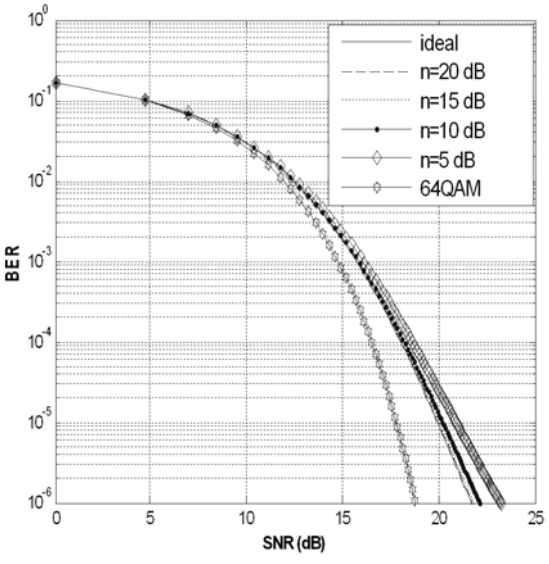


Fig. 8.6 64QAM-OFDM system with amplitude and phase noise($\sigma_{\varphi}^2 = 0.0032$)

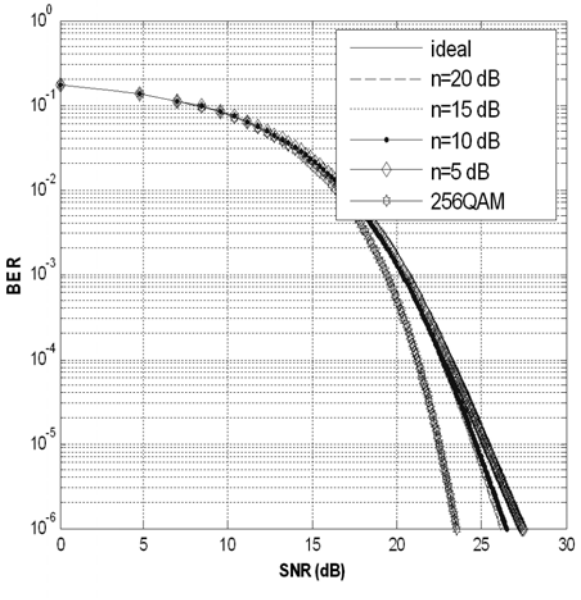


Fig. 8.7 256QAM-OFDM system with amplitude and phase noise ($\sigma_{\varphi}^2 = 0.001$)

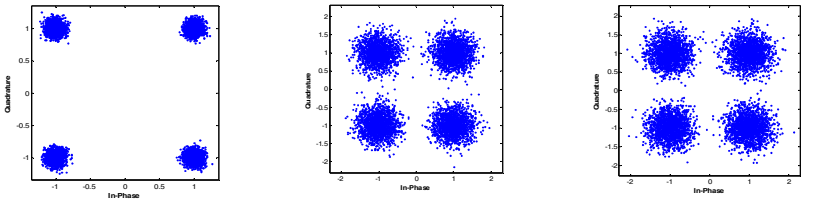
Table 8.2 Phase and Amplitude Noise Degredation in 4,16,64,256 QAM-OFDM for BER=10^{-6}.

σ_{φ}^{2}			**0.0316**	**0.01**	**0.0032**	**0.001**
P_e	**Modulation**	**AM-Noise**	**Degradation (dB)**			
10^{-6}	**4QAM**	**n=100 (ideal)**	6.083	1.218	0.351	0.104
		n=20	6.211	1.232	0.354	0.105
		n=15	6.500	1.260	0.362	0.107
		n=10	7.568	1.348	0.386	0.113
		n=5	17.01	1.648	0.464	0.135
	16QAM	**n=100**	IER	3.619	0.872	0.254
		n=20	IER	3.674	0.881	0.256
		n=15	IER	3.796	0.901	0.262
		n=10	IER	4.208	0.967	0.279
		n=5	IER	5.849	1.179	0.334
	64QAM	**n=100**	IER	IER	2.904	0.720
		n=20	IER	IER	2.945	0.727
		n=15	IER	IER	3.035	0.744
		n=10	IER	IER	3.334	0.797
		n=5	IER	IER	4.442	0.971
	256QAM	**n=100**	IER	IER	IER	2.599
		n=20	IER	IER	IER	2.635
		n=15	IER	IER	IER	2.713
		n=10	IER	IER	IER	2.969
		n=5	IER	IER	IER	3.897

the phase noise is 0.0316 and the amplitude noise order relative to the phase noise is n = 10 dB. As shown in Figure 8.8, clouds in the constellation of (c) have the higher spreading tendency and cause higher performance degradation.

Table 8.3 Relative Degradation for Amplitude Noise Order in 4,16,64,256 QAM-OFDM for BER=10^{-6} and Specified Phase Noise

σ_φ^2		0.0316	0.01	0.0032	0.001
Modulation		**4QAM**	**16QAM**	**64QAM**	**256QAM**
P_e	**AM-Noise**	**Relative Degradation (dB)**			
10^{-6}	**n=100 (ideal)**	0	0	0	0
	n=20	0.1272	0.0554	0.0412	0.0356
	n=15	0.4168	0.1778	0.1314	0.1136
	n=10	1.4854	0.5892	0.4300	0.3700
	n=5	10.9285	2.2301	1.5376	1.2977



(a) (b) (c)

SNR=20dB, $\sigma_\varphi^2 = 0$, n=100 dB (ideal)

SNR=20dB, $\sigma_\varphi^2 = 0.0316$, n=100 dB (ideal)

SNR=20dB, $\sigma_\varphi^2 = 0.0316$, n=10dB

Fig. 8.8 Constellation point of 4QAM signal

8.3 Nonlinearity in OFDM Transceivers

According to the theory of smoothly nonlinear systems, we can represent a component's transfer function by a truncated power series around the DC operating point. The power series law can be used to describe the transfer characteristics of any two-port component, as long as a frequency independent model can represent it [24]. All the internal nonlinearities, such as parasitic capacitances and

inductances, are assumed to have insignificant frequency dependence within the frequency band of operation. Moreover, it is assumed that the circuit is smoothly nonlinear and the circuit output, $V_{out}(t) = f_{NL}[V_{in}(t)]$, can be approximated by the first five terms of its Taylor series expansion:

$$f[x(t)] \cong \sum_{n=0}^{5} a_n x^n = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_5 x^5 \tag{8.31}$$

The above approximation is realistic for a wide range of RF devices, such as amplifiers and mixers [25]-[30]. It is important to note that even-order distortion (*n* even) is mapped to bands that are far from the original OFDM signal passband. It is assumed that these components are filtered out and do not influence the BER. Therefore, only odd-order distortion is of importance. Without loss of generality, we may neglect the DC term, a_0, and express the output of the nonlinear circuit as:

$$v_{out}(t) = a_1 v_{in}(t) + a_3 v_{in}^3(t) + a_5 v_{in}^5(t) \tag{8.32}$$

where a_1, a_3 and a_5 are the linear gain, third-order and fifth-order nonlinearity coefficients, respectively, while v_{in} is the input OFDM voltage.

In the case of OFDM and other multi-tone signals, it can be inferred from the above description that a large number of intermodulation distortion (IMD) products are generated. We calculate the number of IMD products using the generating function technique [17]-[19]. As discussed in Chapter 4, in order to assess nonlinear characteristics, circuit designers make use of either the 1-dB compression point or the input/output inferred third-order intercept points (IIP3/OIP3) [31]. Three coefficients, a_1, a_3 and a_5, can be derived from the electrical amplifier parameters, such as the gain of amplifier (G_0), the third-order output intercept point ($OIP3$) and the 1dB compression point (P_{1dB}), as below [32]:

$$a_1 = 10^{\frac{G_0}{20}} \tag{8.33}$$

$$a_3 = -\frac{2}{3} 10^{\frac{3G_0}{20} - \frac{OIP_3}{10}} \tag{8.34}$$

$$a_5 = \frac{8a_1(1-10^{0.05}) - 6a_3\alpha 10^{0.05}}{5\alpha^2 10^{0.05}} \tag{8.35}$$

where $\alpha = 2 \times 10^{\frac{P_{1dB}+1-G_0}{10}}$

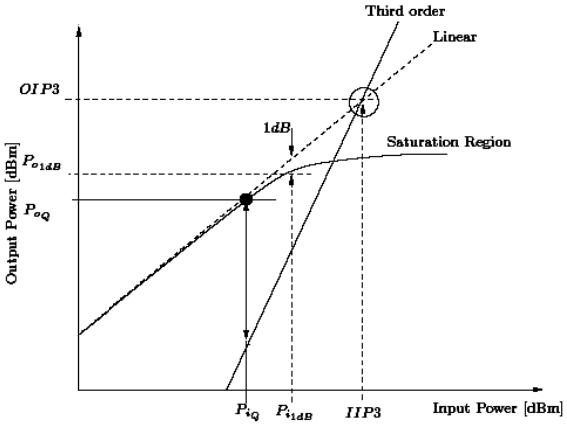


Fig. 8.9 Illustration of 1-dB compression point and IIP3/OIP3

8.3.1 Analysis of Nonlinear Circuit Impact in OFDM

If we consider N modulated data symbols from a particular signaling constellation, $X_k = (X_0, X_1, ..., X_{N-1})$, over a time interval $[0,T]$, the OFDM symbol can be written as:

$$x(t) = \sum_{k=0}^{N-1} X_k e^{j\frac{2\pi kt}{T}} \qquad 0 \le t < T \tag{8.36}$$

where $j = \sqrt{-1}$, X_k is the data symbol for the *k*th subcarrier, ***N*** is the number of subcarriers, and ***T*** is the OFDM symbol time.

The subcarrier spacing of **1/*T*** makes them orthogonal over each symbol period. The output voltage can be expressed as in (8.31), obtained by passing a **low-pass equivalent** OFDM signal through nonlinear circuit and substitution of (8.36) into (8.31) [32].

$$\begin{aligned} f[x(t)] = v_{in}(t) = a_1 \sum_{k=0}^{N-1} X_k e^{j[k\frac{2\pi}{T}t]} + \frac{3a_3}{4} \sum_{k_1,k_2,k_3} X_{k_1} X_{k_2} X_{k_3} e^{j[(k_1+k_2-k_3)\frac{2\pi}{T}t]} \\ + \frac{5a_5}{8} \sum_{k_1,k_2,k_3,k_4,k_5} X_{k_1} X_{k_2} X_{k_3} X_{k_4} X_{k_5} e^{j[(k_1+k_2+k_3-k_4-k_5)\frac{2\pi}{T}t]} \end{aligned} \tag{8.37}$$

The received signal is expressed as (8.38), where $n(t)$ is the complex Gaussian noise:

$$v_{out}(t) = v_{in}(t) + a_1 n(t) \tag{8.38}$$

The sampled signal for the kth subcarrier after the FFT processing stage in the receiver can be written as:

$$\begin{aligned} Y_k &= \frac{1}{N}\sum_{m=0}^{N-1} v_{out}(m).e^{-j\frac{2\pi}{N}km} = \frac{1}{N}\sum_{m=0}^{N-1} a_1 \sum_{l=0}^{N-1} X_l e^{j[l\frac{2\pi}{N}m]} e^{-j\frac{2\pi}{N}km} \\ &+ \frac{1}{N}\sum_{m=0}^{N-1} \frac{3a_3}{4} \sum_{k_1,k_2,k_3} X_{k_1} X_{k_2} X_{k_3} e^{j[(k_1+k_2-k_3)\frac{2\pi}{N}m]} e^{-j\frac{2\pi}{N}km} \\ &+ \frac{1}{N}\sum_{m=0}^{N-1} \frac{5a_5}{8} \sum_{k_1,k_2,k_3,k_4,k_5} X_{k_1} X_{k_2} X_{k_3} X_{k_4} X_{k_5} e^{j[(k_1+k_2+k_3-k_4-k_5)\frac{2\pi}{N}m]} e^{-j\frac{2\pi}{N}km} \\ &+ \frac{a_1}{N}\sum_{m=0}^{N-1} n(m).e^{-j\frac{2\pi}{N}km} \end{aligned} \tag{8.39}$$

or:

$$\begin{aligned} Y_k &= \sum_{l=0}^{N-1} X_l Q_{l-k} + \sum_{k_1,k_2,k_3}^{N-1} X_{k_1} X_{k_2} X_{k_3} Q'_{k_1+k_2-k_3-k} \\ &+ \sum_{k_1,k_2,k_3,k_4,k_5}^{N-1} X_{k_1} X_{k_2} X_{k_3} X_{k_4} X_{k_5} Q''_{k_1+k_2+k_3-k_4-k_5-k} + N_k \end{aligned} \tag{8.40}$$

where N_k is a sampled FFT version of the AWGN noise, Q_k , Q'_k and Q''_k are given by :

$$Q_k = \frac{a_1}{N}\sum_{m=0}^{N-1} e^{j\frac{2\pi}{N}km} \tag{8.41-a}$$

$$Q'_k = \frac{3a_3}{4N}\sum_{m=0}^{N-1} e^{j\frac{2\pi}{N}km} \tag{8.41-b}$$

$$Q''_k = \frac{5a_5}{8N}\sum_{m=0}^{N-1} e^{j\frac{2\pi}{N}km} \tag{8.41-c}$$

The received signal is composed of four contributions: Y_{k_1} is the kth desired subcarrier, Y_{k_2} is the interference component caused by third-order intermodulation (IM) of the subcarriers, Y_{k_3} is the interference component caused by fifth-order IM of the subcarriers, and N_k is the AWGN:

$$Y_k = Y_{k_1} + Y_{k_2} + Y_{k_3} + N_k \tag{8.42}$$

or:

$$\begin{aligned} Y_k &= X_k Q_0 + U_3(N,k).X_{k_1}X_{k_2}X_{k_3}\Big|_{k_1+k_2-k_3=k}.Q_0' \\ &+ U_5(N,k).X_{k_1}X_{k_2}X_{k_3}X_{k_4}X_{k_5}\Big|_{k_1+k_2+k_3-k_4-k_5=k}.Q_0'' \\ &+ \sum_{\substack{k_1,k_2,k_3 \\ ,k_1+k_2-k_3\neq k}}^{N-1} X_{k_1}X_{k_2}X_{k_3}Q'_{k_1+k_2-k_3-k} \\ &+ \sum_{\substack{k_1,k_2,k_3 \\ ,k_1+k_2+k_3-k_4-k_5\neq k}}^{N-1} X_{k_1}X_{k_2}X_{k_3}X_{k_4}X_{k_5}.Q''_{k_1+k_2+k_3-k_4-k_5-k} \\ &+ N_k \end{aligned} \tag{8.43}$$

where $U_3(N,k)$ and $U_5(N,k)$ are the number of integer solutions of the equations:

$$k_1+k_2-k_3=k \quad \& \quad k_1+k_2+k_3-k_4-k_5=k \tag{8.44}$$

with condition:

$$0\le k_1,k_2,k_3,k_4,k_5\le N-1\ ,\ \ k\in\{0,...,N-1\}\ .$$

The number $|U_3|$ and $|U_5|$ are derived as below [30].

$$U_3(N,b)=\frac{3}{2}(2bN-2b^2+2b+N^2-N) \tag{8.45}$$

$$\begin{aligned} U_5(N,b)=\frac{5}{12}[&6(2b-b^2-2b^3+b^4)+6(-1+b+3b^2-2b^3)N \\ &+(1+6b-6b^2)N^2+6(2b-1)N^3+11N^4] \end{aligned} \tag{8.46}$$

The desired signal component can be found as:

$$\begin{aligned} P_{s-desired} &= [E[a_1X_k]]^2+[E[\frac{3a_3}{4}U_3(N,k)X_{k_1}X_{k_2}X_{k_3}]]^2 \\ &+[E[\frac{5a_5}{8}U_5(N,k)X_{k_1}X_{k_2}X_{k_3}X_{k_4}X_{k_5}]]^2 \end{aligned} \tag{8.47}$$

Therefore:

$$P_{s-desired}=a_1^2P_s+\frac{U_3^2(N,k)\times 9\times {a_3}^2\times P_s^3}{16}+\frac{U_5^2(N,k)\times 25\times {a_5}^2\times P_s^5}{64} \tag{8.48}$$

If we consider a normalized OFDM signal as in (8.49), the amount of P_s can be calculated as in (8.50):

$$P_{OFDM-Signal} = E\left|\sum_{k=0}^{N-1} X_k e^{j[\frac{2\pi kt}{T}+\phi_{TX}(t)]}\right|^2_{\phi_{TX}(t)=0} = N.E\left|X_k\right|^2 = N.P_s = 1 \quad (8.49)$$

$$P_s = \frac{1}{N} \quad (8.50)$$

Therefore:

$$P_{s-desired} = a_1^2 P_s + \frac{U_3^2(N,k)\times 9\times {a_3}^2\times P_s}{16N^2} + \frac{U_5^2(N,k)\times 25\times {a_5}^2\times P_s}{64N^4} \quad (8.51)$$

The power of interference component caused by third- and fifth-order IM of the subcarriers, P_{IM3} and P_{IM5} can be calculated respectively as :

$$P_{IM3} = [E[\sum_{\substack{k_1,k_2,k_3 \\ ,k_1+k_2-k_3\neq k}}^{N-1} X_{k_1}X_{k_2}X_{k_3}Q'_{k_1+k_2-k_3-k}]]^2 = \frac{9\times a_3^2}{16N^2}.P_s^3.(\sum_{b\neq k}U_3(N,b))^2 \quad (8.52)$$

Therefore:

$$P_{IM3} = \frac{9\times a_3^2}{16N^4}.P_s.(\sum_{b\neq k}U_3(N,b))^2 \quad (8.53)$$

and

$$P_{IM5} = [E[\sum_{\substack{k_1,k_2,k_3,k_4,k_5 \\ ,k_1+k_2+k_3-k_4-k_5\neq k}}^{N-1} X_{k_1}X_{k_2}X_{k_3}X_{k_4}X_{k_5}Q'_{k_1+k_2+k_3-k_4-k_5-k}]]^2$$
$$= \frac{25\times a_5^2}{64N^2}.P_s^5.(\sum_{b\neq k}U_5(N,b))^2 \quad (8.54)$$

Therefore:

$$P_{IM5} = \frac{25\times a_5^2}{64N^6}.P_s.(\sum_{b\neq k}U_5(N,b))^2 \quad (8.55)$$

The power of noise, P_{n_1}, can be calculated as:

$$P_{n_1} = P_{n_{With\{Nonlinearity\}}} = Var[N_k] = a_1^2.P_n \tag{8.56}$$

From (8.48), (8.53), (8.55) and (8.56), the SNR including the nonlinearity can be arranged as:

$$(\frac{S}{N})_{With\{Nonlinearity\}} = \frac{P_{S,desired}}{(P_{IM3} + P_{IM5}) + P_{n_1}} \tag{8.57}$$

$$(\frac{S}{N})_{With\{Nonlinearity\}} = \frac{(\frac{S}{N})_{Without\{Nonlinearity\}}(64a_1^2N^6 + 36a_3{}^2N^4 \times U_3^2(N,k) + 25a_5{}^2N^2 \times U_5^2(N,k))}{(\frac{S}{N})_{Without\{Nonlinearity\}} \times (36a_3^2N^2.(\sum_{b\neq k} U_3(N,b))^2 + 25a_5^2 \times (\sum_{b\neq k} U_5(N,b))^2) + 64a_1^2N^6} \tag{8.58}$$

The expected value of (8.58) for $1 < k < N$ has been used in simulation as the SNR. Next, the above-mentioned powers are used to evaluate the performance degradation of the OFDM system in the presence of the nonlinearity factor. Therefore, the degradation factor (DF) is calculated as follows:

$$DF = 10\log\frac{(\frac{S}{N})_{Without\{Nonlinearity\}}}{(\frac{S}{N})_{With\{Nonlinearity\}}} \tag{8.59}$$

$$DF = 10\log \{\frac{(36a_3^2N^2.(\sum_{b\neq k} U_3(N,b))^2 + 25a_5^2 \times (\sum_{b\neq k} U_5(N,b))^2) \times (\frac{S}{N})_{Without\{Nonlinearity\}} + 64a_1^2N^6}{64a_1^2N^6 + 36a_3{}^2 \times N^4 \times U_3^2(N,k) + 25a_5{}^2 \times N^2 \times U_5^2(N,k)}\} \tag{8.60}$$

■ ***Example 8.1:*** *Impact of Nonlinearity in OFDM Systems*

OFDM modulation has a high peak-to-average power ratio (PAPR), necessitating high linearity active devices to be able to process the signal with low distortion.

For validation of the analytical results, we shall consider one linear high power amplifier with nonlinear coefficients. The power amplifier product is WPS-495922-02. It operates over a 4.9 GHz to 5.9 GHz band with 11 dB gain and 32 dBm output power at the 1dB compression point.

Thia product uses high linearity device technology, which allows us to address the important 802.11 Wi-Fi and the emerging 802.16 WiMAX (Worldwide Interoperability for Microwave Access) infrastructure markets [34]. The electrical parameters and corresponding polynomial coefficients of the amplifier are shown in Tables 8.4 and 8.5 by considering the datasheet of the amplifier and the above equations.

Table 8.4 Electrical Parameters of Amplifier WPS-495922-02

Electrical Parameter	WPS-495922-02	Application
G_0	11.0 dB	802.16 WiMAX 256 carriers 64 QAM
OIP3 (dBm)	47 dBm	
P_{1dB}(dBm)	32 dBm	

Table 8.5 Polynomial Coefficients of Amplifier WPS-495922-02

Polynomial Coefficients	WPS-495922-02
a_1	3.55
a_3	$-5.94*10^{-4}$
a_5	$-8.39*10^{-6}$

Figure 8.10 shows the theoretical and simulation BER performances of an OFDM communication system employing 64QAM modulation and 256 subcarriers with the electrical parameters of the WPS-495922-02 power amplifier. Figure 8.11 shows the theoretical and simulation performance degradation of an OFDM communication system with the electrical parameters of the WPS-495922-02 power amplifier and different numbers of subcarriers. It is shown that increasing the number of subcarriers causes higher performance degradation.

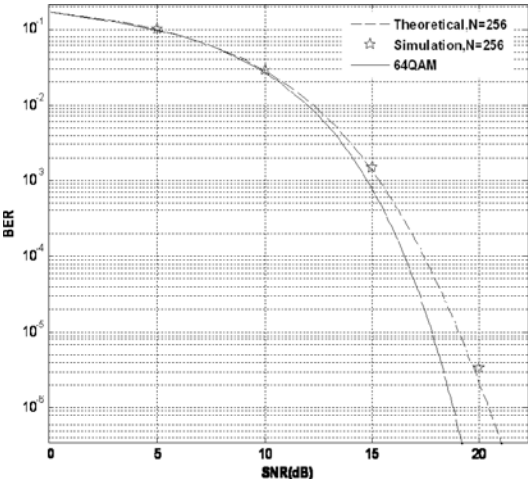


Fig. 8.10 64QAM-OFDM systems with specified electrical parameter and 256 subcarriers

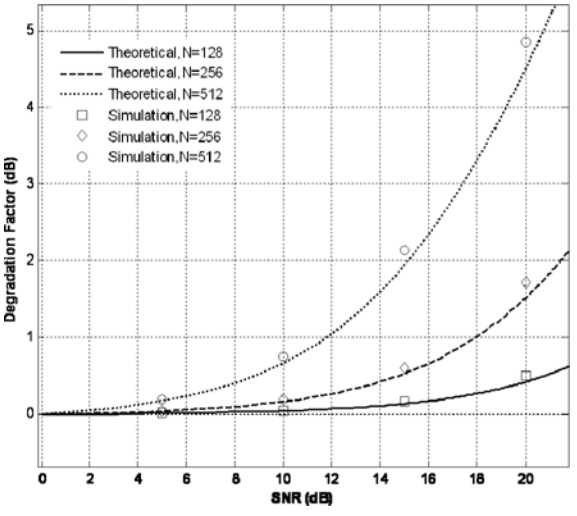


Fig. 8.11 Performance degradation of OFDM system with specified electrical parameters and different subcarriers

8.4 Concurrent Analysis of Nonlinearity and Phase Noise in OFDM Transceivers

This section investigates the concurrent circuit nonlinearity and phase noise effects on the OFDM signal. Their effects are investigated analytically, and the closed form SNR expressions are derived. By using a truncated power series to represent the nonlinear behavior of an RF circuit and a Lorentzian function to model phase noise, the effects of nonlinearity and phase noise on OFDM signals are examined. A closed form expression is obtained for the SNR degradation due to third-order nonlinearity and phase noise effects.

Considering linear and cubic terms alone, OIP3 is defined as [35]:

$$OIP_3 = \left| \frac{2}{3} \frac{a_1^3}{a_3} \right| \tag{8.61}$$

Also, the OIP3 of N stages in cascade is as:

$$\frac{1}{OIP_3} = \frac{1}{(OIP_3)_1 G_2 \ldots G_N} + \ldots \frac{1}{(OIP_3)_{N-2} G_{N-1} G_N} + \frac{1}{(OIP_3)_{N-1} G_N} + \frac{1}{(OIP_N)_N} \tag{8.62}$$

In the following studies, we consider $a_1 = 1$ and the typical range of OIP3 to be 30 dBM to 50 dBm. On the other hand, According to (8.4), the phase noise variance is:

$$\sigma_\varphi^2 = 2\int_0^{B_s/2} L(f)^2\, df = \frac{2K}{\pi} arctg\left(\frac{B_s}{2B}\right) + \frac{1}{\pi} L_0 B_s \tag{8.63}$$

where B_s is the OFDM bandwidth.

8.4.1 Theoretical Analysis

As shown, the low-pass equivalent OFDM signal corrupted by transmitter oscillator phase noise is expressed as:

$$v_{in}(t) = \sum_{k=0}^{N-1} X_k e^{j\left[\frac{2\pi kt}{T} + \phi_{TX}(t)\right]} \qquad 0 \le t < T \tag{8.64}$$

where X_k is the data symbol for the *k*th subcarrier, *N* is the number of subcarriers, $\phi_{TX}(t)$ is the transmitter LO phase noise, and *T* is the OFDM symbol time. The subcarrier spacing of 1/*T* makes them orthogonal over each symbol period.

The output voltage can be obtained by passing a low-pass equivalent OFDM signal through a nonlinear circuit and the substitution of (8.64) into (8.65). The derivation calculus can be found in [38].

$$s(t) = a_1 v_{in}(t) + a_3 v_{in}(t)^3 \tag{8.65}$$

$$s(t) = a_1 \sum_{k=0}^{N-1} X_k e^{j\left[k\frac{2\pi}{T}t + \phi_{TX}(t)\right]} + \frac{3a_3}{4} \sum_{k_1,k_2,k_3} X_{k_1} X_{k_2} X_{k_3} e^{j\left[(k_1+k_2-k_3)\frac{2\pi}{T}t + \phi_{TX}(t)\right]} \tag{8.66}$$

The received signal is:

$$r(t) = s(t) + a_1 n(t) \tag{8.67}$$

where $n(t)$ is the complex Gaussian noise. Moreover, the received signal is influenced by the phase noise of RX LO ($\phi_{RX}(t)$) as:

$$v_{out}(t) = r(t)e^{j\phi_{RX}(t)} = s(t)e^{j\phi_{RX}(t)} + a_1 n(t)e^{j\phi_{RX}(t)} \tag{8.68}$$

$$v_{out}(t) = a_1 \sum_{k=0}^{N-1} X_k e^{j[k\frac{2\pi}{T}t+\phi(t)]} + \frac{3a_3}{4} \sum_{k_1,k_2,k_3} X_{k_1} X_{k_2} X_{k_3} e^{j[(k_1+k_2-k_3)\frac{2\pi}{T}t+\phi(t)]} + a_1 n(t) e^{j\phi_{RX}(t)} \quad (8.69)$$

where $\phi(t) = \phi_{TX}(t) + \phi_{RX}(t)$.

It can be assumed that the phase noise of the local oscillators at the transmitter and receiver are independent and identically distributed (i.i.d.) with a zero mean and variance of σ_ϕ^2; hence, $Var(\phi(t)) = 4\sigma_\phi^2$ [38].

The sampled signal for the *k*th subcarrier after the FFT processing stage in the receiver can be written as:

$$\begin{aligned} Y_k &= \frac{1}{N}\sum_{m=0}^{N-1} v_{out}(m) e^{-j\frac{2\pi}{N}km} = \frac{1}{N}\sum_{m=0}^{N-1} a_1 \sum_{l=0}^{N-1} X_l e^{j[l\frac{2\pi}{N}m+\phi(m)]} e^{-j\frac{2\pi}{N}km} \\ &+ \frac{1}{N}\sum_{m=0}^{N-1} \frac{3a_3}{4} \sum_{k_1,k_2,k_3} X_{k_1} X_{k_2} X_{k_3} e^{j[(k_1+k_2-k_3)\frac{2\pi}{N}m+\phi(m)]} e^{-j\frac{2\pi}{N}km} \\ &+ \frac{1}{N}\sum_{m=0}^{N-1} n(m) e^{j\phi_{RX}(m)} . e^{-j\frac{2\pi}{N}km} \\ &= \sum_{l=0}^{N-1} X_l Q_{l-k} + \sum_{k_1,k_2,k_3}^{N-1} X_{k_1} X_{k_2} X_{k_3} Q'_{k_1+k_2-k_3-k} + N_k \end{aligned} \quad (8.70)$$

where N_k is a sampled FFT of the AWGN noise multiplied by the phase noise of the receiver LO with a variance of σ_n^2, $\phi(m) = \phi_{TX}(m) + \phi_{RX}(m)$, and Q_k and Q'_k are given by:

$$Q_k = \frac{a_1}{N}\sum_{m=0}^{N-1} e^{j\phi(m)} e^{j\frac{2\pi}{N}km} \quad (8.71)$$

$$Q'_k = \frac{3a_3}{4N}\sum_{m=0}^{N-1} e^{j\phi(m)} e^{j\frac{2\pi}{N}km} \quad (8.72)$$

The received signal is composed of four contributions: Y_{k_1} is the *k*th desired subcarrier, Y_{k_2} is the interference component caused by inter-subcarrier interference, Y_{k_3} is the interference component caused by intermodulation of the subcarriers, and N_k is the AWGN, i.e.:

$$Y_k = Y_{k_1} + Y_{k_2} + Y_{k_3} + N_k$$

or:

$$Y_k = X_k Q_0 + U_3(N,k) X_{k_1} X_{k_2} X_{k_3} \Big|_{k_1+k_2-k_3=k} Q_0' + \sum_{l=0, l\neq k}^{N-1} X_l Q_{l-k} + \sum_{\substack{k_1,k_2,k_3 \\ ,k_1+k_2-k_3\neq k}}^{N-1} X_{k_1} X_{k_2} X_{k_3} Q'_{k_1+k_2-k_3-k} + N_k \tag{8.73}$$

where $U_3(N,k)$ is the number of integer solutions of the equation $k_1 + k_2 - k_3 = k$ with condition $0 \le k_1, k_2, k_3 \le N-1$, $k \in \{0,...,N-1\}$ [38].

The expected value of $X_k Q_0 + U_3(N,k) X_{k_1} X_{k_2} X_{k_3} \Big|_{k_1+k_2-k_3=k} Q_0'$ is the desired signal component. Hence, the desired signal power can be obtained as:

$$P_{s-desired} = [E[X_k Q_0]]^2 + [E[U_3(N,k) X_{k_1} X_{k_2} X_{k_3} Q_0']]^2 \tag{8.74}$$

where $E[Q_0] = a_1 e^{-2\sigma_\phi^2}$ and $E[Q_0'] = \frac{3a_3}{4} e^{-2\sigma_\phi^2}$. Accordingly:

$$P_{s-desired} = a_1^2 P_s e^{-4\sigma_\phi^2} + \frac{9U_3^2(N,k) a_3{}^2 P_s{}^3 e^{-4\sigma_\phi^2}}{16} \tag{8.75}$$

Using of a normalized OFDM signal, the amount of P_s can be calculated as:

$$P_{OFDM-Signal} = E\left| \sum_{k=0}^{N-1} X_k e^{j[\frac{2\pi kt}{T} + \phi_{TX}(t)]} \right|^2_{\phi_{TX}(t)=0} = NE\left|X_k\right|^2 = NP_s = 1 \tag{8.76}$$

Therefore:

$$P_{s-desired} = a_1^2 P_s e^{-4\sigma_\phi^2} + \frac{9U_3^2(N,k) a_3{}^2 P_s e^{-4\sigma_\phi^2}}{16N^2} \tag{8.77}$$

The variance of $X_k Q_0 + U_3(N,k) X_{k_1} X_{k_2} X_{k_3} \Big|_{k_1+k_2-k_3=k} Q_0'$ is the phase noise power caused by the common phase error (CPE) and P_{CPE} is calculated as:

$$P_{CPE} = Var[X_k Q_0] + Var[U_3(N,k) X_{k_1} X_{k_2} X_{k_3} Q_0'] = P_s Var[Q_0] + P_s^3 U_3{}^2(N,k) Var[Q_0'] \tag{8.78}$$

where $Var[Q_0] = \frac{a_1^2}{N}(1-e^{-4\sigma_\phi^2})$ and $Var[Q_0'] = \frac{9a_3{}^2}{16N}(1-e^{-4\sigma_\phi^2})$. Hence, one may obtain P_{CPE} as:

$$P_{CPE} = P_s \frac{a_1^2}{N}(1-e^{-4\sigma_\phi^2}) + P_s U_3{}^2(N,k)\frac{9a_3{}^2}{16N^3}(1-e^{-4\sigma_\phi^2}) \tag{8.79}$$

On the other hand, in (8.78), $\sum_{l=0,l\neq k}^{N-1} X_l Q_{l-k}$ is the interference component caused by the inter-subcarrier interference (ICI). So, P_{ICI} can be calculated as:

$$P_{ICI} = P_s\{Var[\sum_{l=0}^{N-1} Q_l] - Var[Q_0]\} \tag{8.80}$$

where $Var[\sum_{l=0}^{N-1} Q_l] = a_1^2(1-e^{-4\sigma_\phi^2})$ and $Var[\sum_{l=0}^{N-1} Q_l'] = \frac{9a_3^2}{16}(1-e^{-4\sigma_\phi^2})$. The derivation calculus of $Var[\sum_{l=0}^{N-1} Q_l]$ and $Var[\sum_{l=0}^{N-1} Q_l']$ is described in detail [38],[38]. Accordingly:

$$P_{ICI} = a_1^2 P_s \frac{N-1}{N}(1-e^{-4\sigma_\phi^2}) \tag{8.81}$$

Moreover, in (8.71), the term $\sum_{\substack{k_1,k_2,k_3 \\ ,k_1+k_2-k_3\neq k}}^{N-1} X_{k_1} X_{k_2} X_{k_3} Q'_{k_1+k_2-k_3-k}$ is the interference component caused by intermodulation of the subcarriers. Hence, P_{IM} can be calculated as:

$$P_{IM} = [E[\sum_{\substack{k_1,k_2,k_3 \\ ,k_1+k_2-k_3\neq k}}^{N-1} X_{k_1} X_{k_2} X_{k_3} Q'_{k_1+k_2-k_3-k}]]^2 = \frac{9a_3^2}{16N^2} P_s^3 e^{-4\sigma_\phi^2}(\sum_{b\neq k} U_3(N,b))^2 \tag{8.82}$$

$$P_{IM} = \frac{9a_3^2}{16N^4} P_s e^{-4\sigma_\phi^2}(\sum_{b\neq k} U_3(N,b))^2 \tag{8.83}$$

As mentioned, N_k is a sampled FFT of the AWGN noise multiplied by the phase noise of the receiver LO. Therefore, for the sake of simple analysis, we consider $Var(N_k)$ as that of the AWGN, since it has a zero mean, the variance is

unchanged on the basis that the FFT process is linear, and the phase noise corrupting the AWGN works as the phase rotation factor.

$$P_{n_1} = P_{n_{With\{\phi,Nonlinearity\}}} = Var[N_k] = a_1^2 P_n \tag{8.84}$$

From the above relations, the SNR, including the nonlinearity and phase noise, can be arranged as:

$$(\frac{S}{N})_{With\{\phi,Nonlinearity\}} = \frac{P_{S,desired}}{(P_{CPE} + P_{ICI} + P_{IM}) + P_{n_1}} \tag{8.85}$$

$$(\frac{S}{N})_{With\{\phi,Nonlinearity\}} = \frac{(\frac{S}{N})_{Without\{\phi,Nonlinearity\}}(16N^4 a_1^2 e^{-4\sigma_\phi^2} + 9{a_3}^2 N^2 U_3^2(N,k)e^{-4\sigma_\phi^2})}{A(\frac{S}{N})_{Without\{\phi,Nonlinearity\}} + 9a_3^2(\sum_{b\neq k} U_3(N,b))^2 e^{-4\sigma_\phi^2} + 16a_1^2 N^4} \tag{8.86}$$

where $A \triangleq (1-e^{-4\sigma_\phi^2})(16N^3 a_1^2 + 9a_3^2 N{U_3}^2(N,k) + 16N^3 a_1^2(N-1))$.

The expected value of (8.84) for $1 < k < N$ has been used in simulations as the SNR. In addition, the above-mentioned powers are used to evaluate the performance degradation of the OFDM system in the presence of nonlinearity and phase noise factors. Therefore, the degradation factor (*DF*) is calculated as follows:

$$DF = 10\log\frac{(\frac{S}{N})_{Without\{\phi,Nonlinearity\}}}{(\frac{S}{N})_{With\{\phi,Nonlinearity\}}} \tag{8.87}$$

$$DF = 10\log\frac{(A + 9a_3^2(\sum_{b\neq k} U_3(N,b))^2)e^{-4\sigma_\phi^2}(\frac{S}{N})_{Without\{\phi,Nonlinearity\}} + 16a_1^2 N^4}{16N^4 a_1^2 e^{-4\sigma_\phi^2} + 9a_3^2 N^2 U_3^2(N,k)e^{-4\sigma_\phi^2}} \tag{8.88}$$

If we ignore from nonlinearity and suppose that OIP3 = ∞ ($a_3 = 0$), the SNR becomes equal to equation (8.87), which is consistent with the literature [38], [39]:

$$(\frac{S}{N})_{With\{\phi\}} = \frac{(\frac{S}{N})_{Without\ \{\phi\}} e^{-4\sigma_\phi^2}}{1 + (\frac{S}{N})_{Without\ \{\phi\}}(1-e^{-4\sigma_\phi^2})} \tag{8.89}$$

Figure 8.12 shows simulation results of (8.87) for the BER of a 4-16-64-256QAM-OFDM system with different phase noise variances, which has been compared with available results.

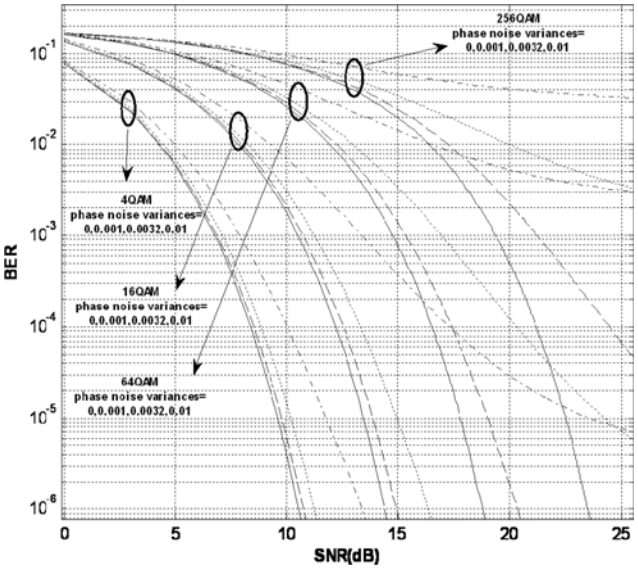


Fig. 8.12 QAM-OFDM systems with phase noise variances

8.4.2 Performance Analysis

The OFDM system performance is degraded by the nonlinear power amplifier and noisy local oscillator. The simulation model to investigate this degradation is presented in Figure 8.13. In this simulation, the behavioral model for RF circuits has been used. As mentioned, the phase noise power density function of the oscillator is modeled with a Lorentzian function with specified parameters and the nonlinear transfer function is modeled with a third-order memoryless polynomial model.

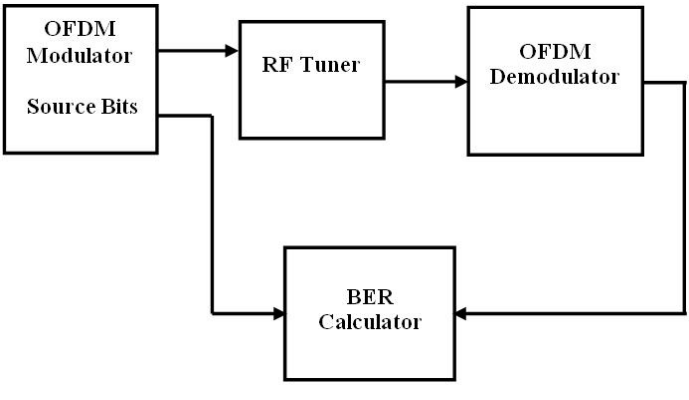


Fig. 8.13 OFDM simulation model

It is generally difficult to compensate for the circuit nonlinearity and phase noise degradation in the OFDM system. Therefore, it is important to quantify their permissible levels in RF tuners, inorder to maintain the system performance. This leads to design RF tuners with limited maximum nonlinearity and phase noise. The BER formula for system performance with phase noise and nonlinearity can be found based on SNR. By use of the Gray-encoded 4,16,64,256QAM modulation method and the coherent detection technique in the AWGN channel, the BER in QAM can be obtained as [38].

$$P_{4-QAM} = \frac{1}{2} erfc(\sqrt{SNR_{with\{\phi,Nonlinearity\}}}) - \frac{1}{16} erfc^2(\sqrt{SNR_{with\{\phi,Nonlinearity\}}}) \quad (8.90\text{-a})$$

$$P_{16-QAM} = \frac{3}{8} erfc(\sqrt{\frac{2}{5} SNR_{with\{\phi,Nonlinearity\}}}) - \frac{9}{256} erfc^2(\sqrt{\frac{2}{5} SNR_{with\{\phi,Nonlinearity\}}}) \quad (8.90\text{-b})$$

$$P_{64-QAM} = \frac{7}{24} erfc(\sqrt{\frac{4}{21} SNR_{with\{\phi,Nonlinearity\}}}) - \frac{49}{4096} erfc^2(\sqrt{\frac{4}{21} SNR_{with\{\phi,Nonlinearity\}}}) \quad (8.90\text{-c})$$

$$P_{256-QAM} = \frac{15}{64} erfc(\sqrt{\frac{4}{85} SNR_{with\{\phi,Nonlinearity\}}}) - \frac{225}{16384} erfc^2(\sqrt{\frac{4}{85} SNR_{with\{\phi,Nonlinearity\}}}) \quad (8.90\text{-d})$$

where $SNR_{with(\varphi,Nonlinearity)}$ is the SNR with phase noise and nonlinearity in an OFDM system. This can be used to analyze the BER performance in the OFDM system with nonlinearity and phase noise.

Figure 8.14 shows the BER performance of the OFDM system employing 4,16,64,256QAM modulations in the AWGN channel. For each modulation scheme, we only considered the nonlinearity with OIP3 in the range of 37 dBm to 41dBm for 256 and 1024 subcarriers in an OFDM system. Figure 8.15 shows the BER performance of the OFDM system employing 4,16,64,256QAM modulations with different subcarriers and a specified OIP3; and, Figure 8.16 shows the BER performance of the same system with different OIP3 and specified subcarriers.

Figure 8.17 shows the combination effect of phase noise and nonlinearity for different phase noise variances and OIP3 on the BER performance employing 4QAM modulation with different subcarriers. In addition, Figure 8.18 shows the constellation of a 16QAM-OFDM system with OIP3 = 40dBm and $\sigma_\phi^2 = 0.001$ for 64,256,1024,2048 subcarriers in OFDM system.

Using these results, engineers are given a trade-off for choosing the best values for the parameters of an OFDM transceiver system, such as the constellation size of modulation, number of subcarriers, phase noise variance of an oscillator and third-order intercept point of a power amplifier, to meet the acceptable BER in a specified SNR.

For example, if RF devices with a specified phase noise and nonlinearity exist, designers can choose a suitable scheme for modulation and an appropriate number of subcarriers; or, if an OFDM system with a specified scheme of modulation and

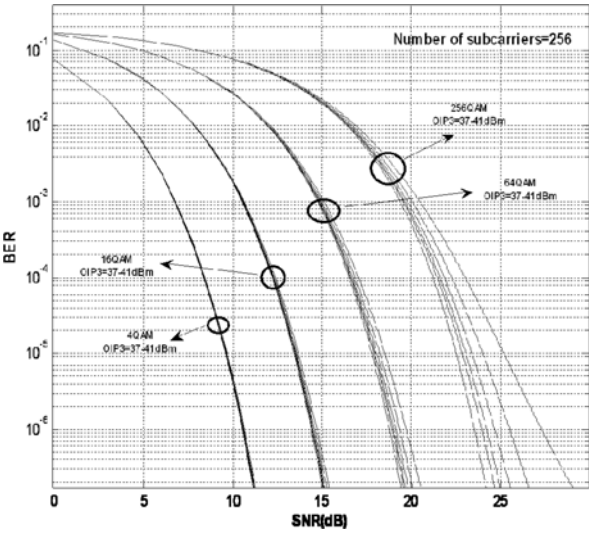


Fig. 8.14 QAM-OFDM systems with different OIP3 and 256 subcarriers

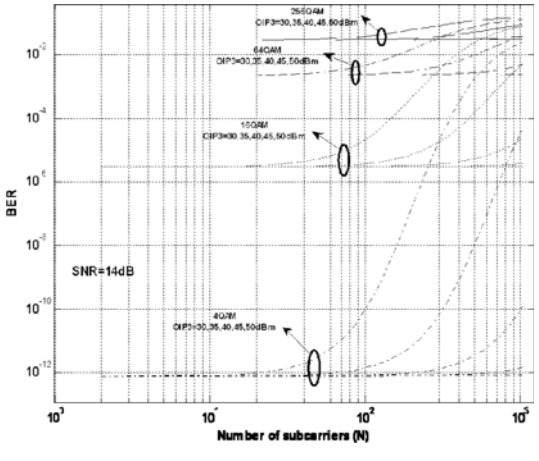


Fig. 8.15 QAM-OFDM systems with different subcarriers and specified OIP3 (SNR = 14dB)

.

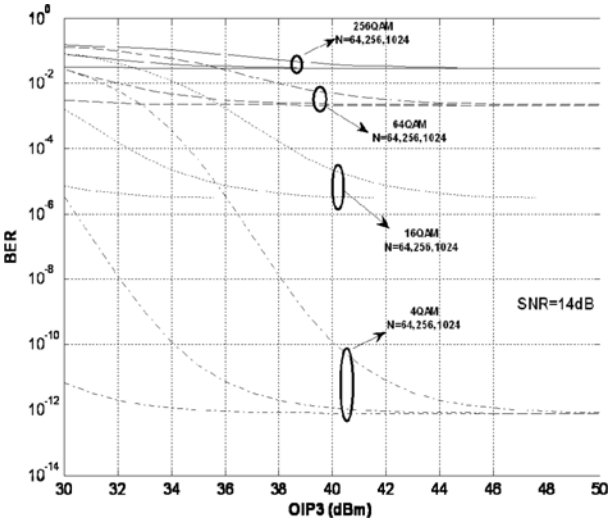


Fig. 8.16 QAM-OFDM systems with different OIP3 and specified subcarriers (SNR = 14dB)

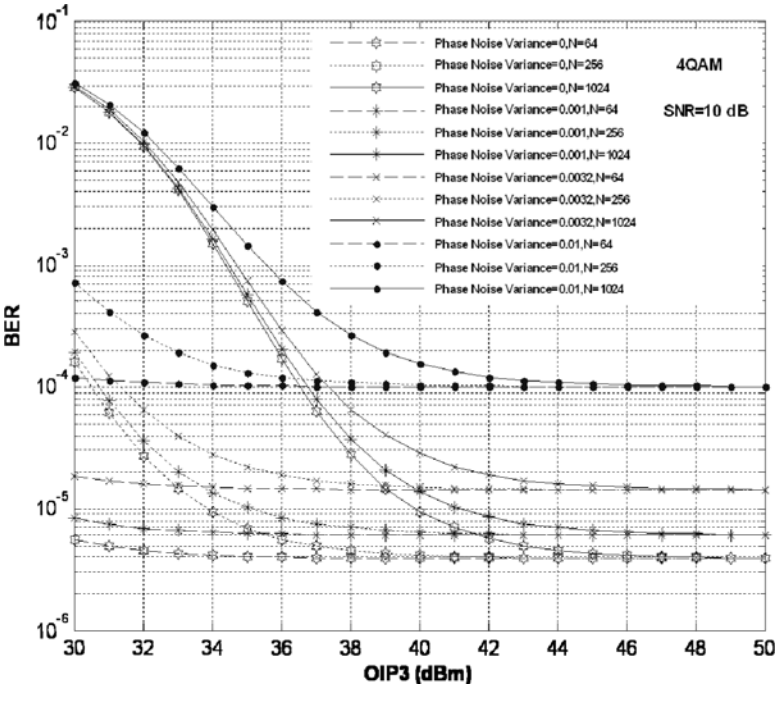


Fig. 8.17 QAM-OFDM system with different OIP3 and specified subcarriers and phase noise variances (SNR = 10dB)

number of subcarriers exists, engineers can choose RF devices with acceptable phase noise and nonlinearity to meet the acceptable BER in a specified SNR

8.4.3 Fifth-Order Nonlinear Circuit and Phase Noise Modeling

Without loss of generality, we may neglect the DC term, a_0, and express the output of the fifth0order nonlinear circuit as:

$$v_{out}(t) = a_1 v_{in}(t) + a_3 v_{in}^3(t) + a_5 v_{in}^5(t) \tag{8.91}$$

where a_1, a_3 and a_5 are the linear gain and third- and fifth-order nonlinearity coefficients, respectively, while v_{in} is the input OFDM voltage. The coefficients are as determined in (8.33), (8.34), and (8.35).

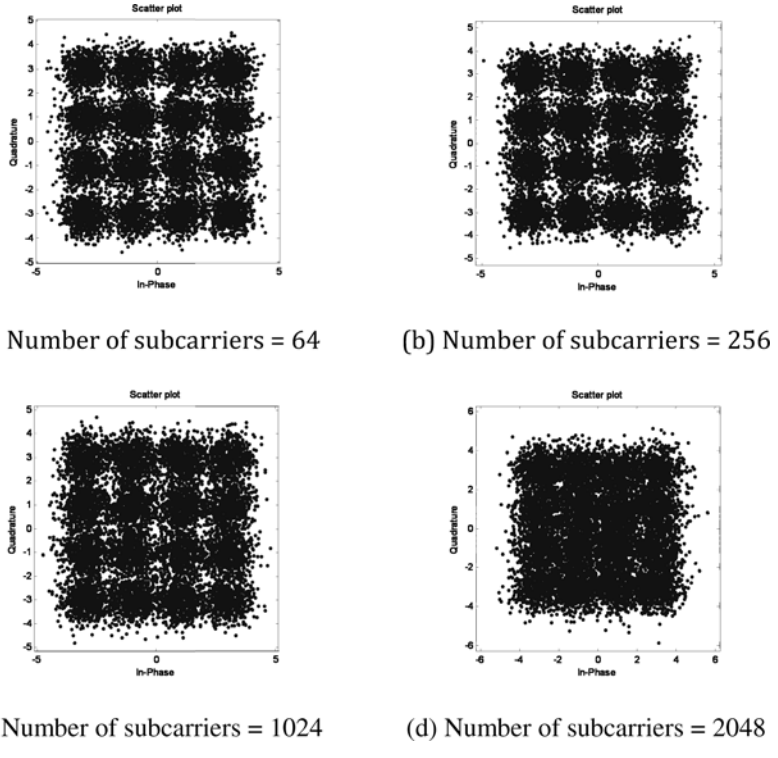


Fig. 8.18 Constellation points of 16QAM signal for different numbers of subcarriers (SNR = 14dB, OIP3 = 40dBm,)

Here, the AM/AM effects on the performance of the OFDM system are studied. The AM/PM effects are excluded because the AM/AM effects on the OFDM performance have been found to be much more significant than those of AM/PM.

Similar to the third-order nonlinearity, the SNR with the nonlinearity and phase noise for the fifth-order nonlinearity can be arranged as:

$$(\frac{S}{N})_{With\{\phi,Nonlinearity\}} = \frac{P_{S,desired}}{(P_{CPE} + P_{ICI} + P_{IM3} + P_{IM5}) + P_{n_1}} \tag{8.92}$$

$$\begin{aligned}(\frac{S}{N})_{With\{\varphi,Nonlinearity\}} &= \frac{(\frac{S}{N})_{Without\{\varphi,Nonlinearity\}}}{(\frac{S}{N})_{Without\{\varphi,Nonlinearity\}}}\times \\ &\frac{e^{-4\sigma_\varphi^2}(64a_1^2\times N^6 + 36{a_3}^2\times N^4\times U_3^2(N,k) + 25{a_5}^2\times N^2\times U_5^2(N,k))}{(B + e^{-4\sigma_\varphi^2}\times(36a_3^2\times N^2\times(\sum_{b\neq k}U_3(N,b))^2 + 25a_5^2\times(\sum_{b\neq k}U_5(N,b))^2)) + 64a_1^2N^6}\end{aligned} \tag{8.93}$$

where:

$$B \triangleq (1-e^{-4\sigma_\varphi^2})\times(64a_1^2\times N^5 + 36{a_3}^2\times N^3\times {U_3}^2(N,k) + 25{a_5}^2\times N\times {U_5}^2(N,k))$$

The expected value for $1<k<N$ has been used in simulations as the SNR. In addition, the above-mentioned powers are used to evaluate the performance degradation of the OFDM system in the presence of nonlinearity and phase noise factors. Therefore, the degradation factor (*DF*) is calculated as follows:

$$DF = 10\log\frac{(\frac{S}{N})_{Without\{\phi,Nonlinearity\}}}{(\frac{S}{N})_{With\{\phi,Nonlinearity\}}} \tag{8.94}$$

$$\begin{aligned}DF = 10\log\{&\frac{1}{e^{-4\sigma_\varphi^2}}\times \\ &(B + e^{-4\sigma_\varphi^2}\times(36a_3^2\times N^2\times(\sum_{b\neq k}U_3(N,b))^2 + 25a_5^2\times(\sum_{b\neq k}U_5(N,b))^2)) \\ &\frac{(\frac{S}{N})_{Without\{\varphi,Nonlinearity\}} + 64a_1^2N^6}{(64a_1^2\times N^6 + 36{a_3}^2\times N^4\times U_3^2(N,k) + 25{a_5}^2\times N^2\times U_5^2(N,k))}\}\end{aligned} \tag{8.95}$$

In this analysis, if we ignore the nonlinearity and suppose that $a_3 = a_5 = 0$, the SNR becomes consistent with previous works [23], [26]:

$$\left(\frac{S}{N}\right)_{With\{\phi\}} = \frac{\left(\frac{S}{N}\right)_{Without\ \{\phi\}} . e^{-4\sigma_\phi^2}}{1+\left(\frac{S}{N}\right)_{Without\ \{\phi\}} \times (1-e^{-4\sigma_\phi^2})} \tag{8.96}$$

Figure 8.19 shows simulation results for the BER of a 4-16-64-256QAM-OFDM system with different phase noise variances that have been compared with available results.

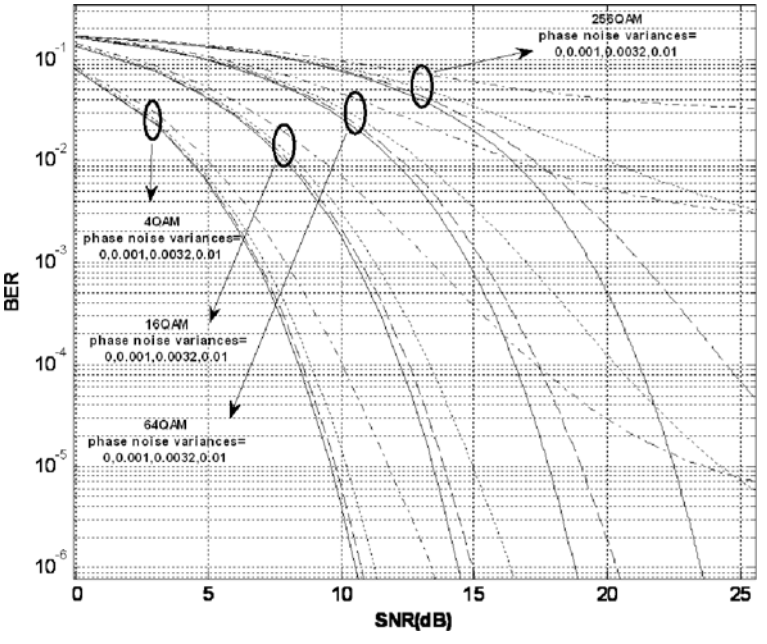


Fig. 8.19 QAM-OFDM systems with phase noise variances

■ ***Example 8.2****: Performance Analysis*

We consider the parameters of one actual and applicable high power amplifier with fifth-order nonlinear coefficients. This power amplifier product is WPS-495922-02. It operates over a 4.9 to 5.9 GHz band with an 11 dB gain and a 32 dBm output power at the 1dB compression point. The electrical parameters, corresponding polynomial coefficients and application of amplifier are shown in Table 8.6, by considering the datasheet of the amplifier and the above equations.

Figure 8.20 shows the combined effect of phase noise and nonlinearity with a specified parameter and different output back-offs (OBOs) on the BER performance of an OFDM communication system employing 4QAM with different subcarriers. Figure 8.21 shows the constellation of a 4QAM-OFDM system with an

OBO = 10dB and $\sigma_\phi^2 = 0.001$ for 64,256,1024,2048 subcarriers in an OFDM system. The SNR is 14 dB. Figure 8.22 shows the BER performance of an OFDM communication system employing 64QAM and 256 subcarriers with specified parameters for different OBOs.

Table 8.6 Electrical Parameters of WPS-495922-02

Electrical Parameter		Polynomial Coefficients		Application
G_0	11.0 dB	a_1	3.55	802.16 Wi-MAX 256 carriers 64 QAM
OIP_3 (dBm)	47 dBm	a_3	$-5.94*10^{-4}$	
P_{1dB} (dBm)	32 dBm	a_5	$-3.89*10^{-6}$	

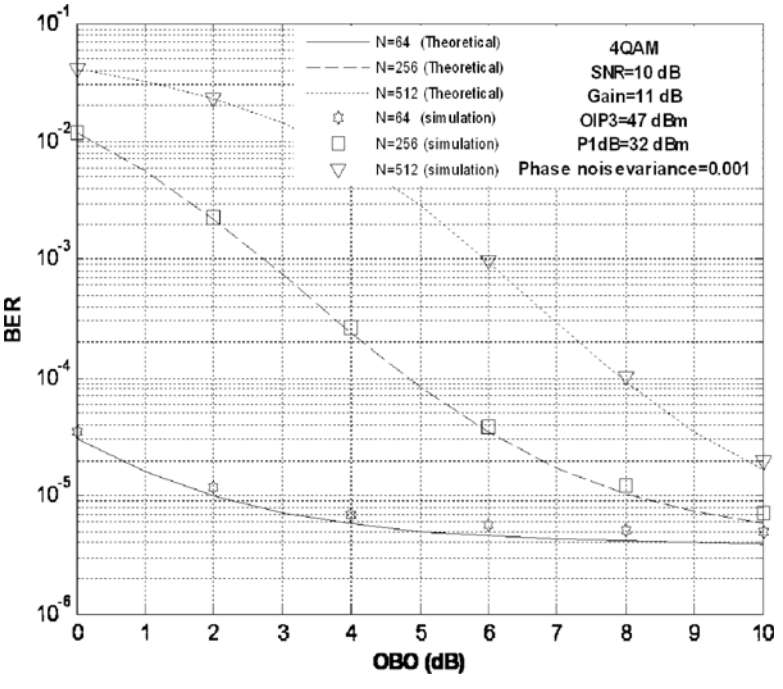


Fig. 8.20 QAM-OFDM system with specified parameters and different OBOs and subcarriers (SNR = 10dB)

Figure 8.23 shows the performance degradation of an OFDM communication system with specified parameters for different subcarriers and OBO = 10 dB. Using these results, one has a metric to select the suitable values of parameters in OFDM transceivers. These parameters can be the constellation size, number of

subcarriers, phase noise variance of an oscillator, and the OBO of a power amplifier. These are used to meet the acceptable BER in a specified SNR. For instance, if RF devices with specified phase noise and nonlinearity are assumed, one can choose a suitable scheme of modulation and an appropriate number of subcarriers. On the other hand, if an OFDM system with a specified scheme of modulation and number of subcarriers is considered, a device with acceptable phase noise and nonlinearity must be selected.

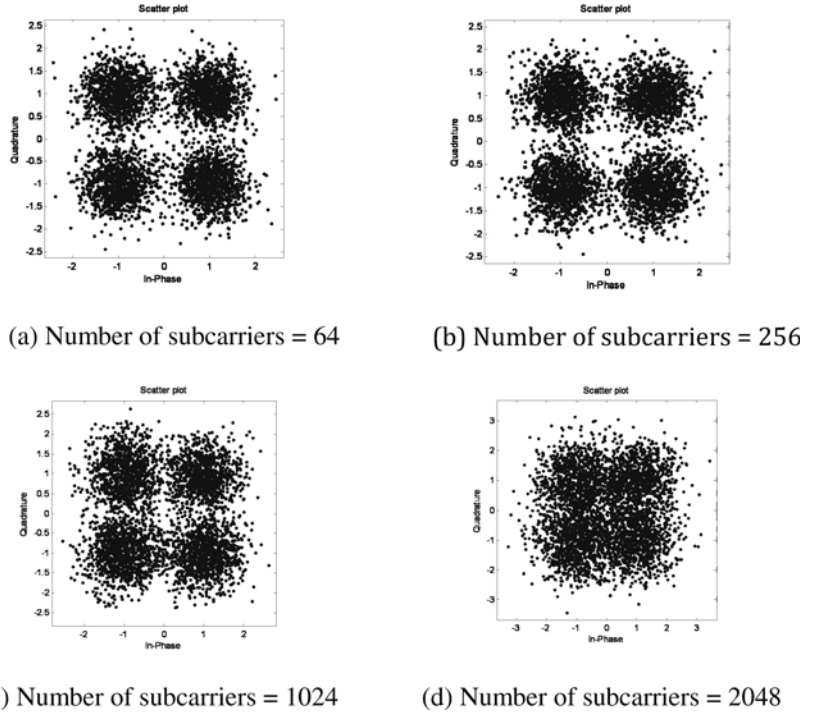


Fig. 8.21 The constellations of 4QAM-OFDM system with different numbers of subcarriers, OBO = 10dB and $\sigma_{\phi}^{2} = 0.001$

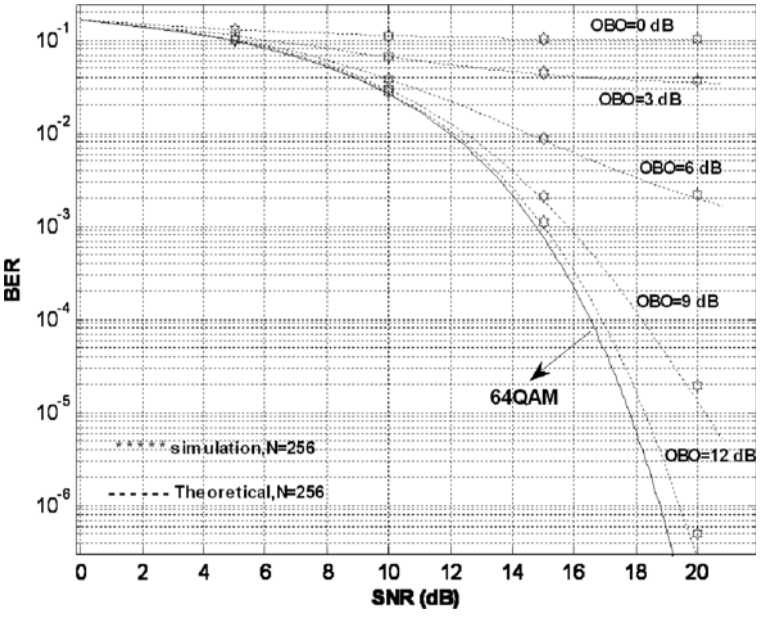


Fig. 8.22 64QAM-OFDM system with specified parameters and different OBOs

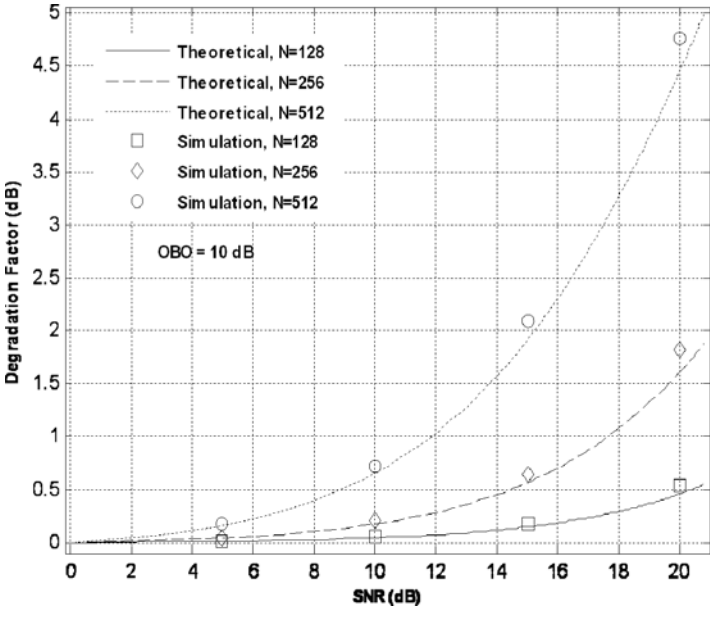


Fig. 8.23 Performance degradation of an OFDM system with specified parameters for different subcarriers and OBO = 10 dB

8.5 OFDM Signal with Phase Noise Passing through Dynamic Nonlinear Circuits

In this section, the SNR degradation due to the effects of dynamic nonlinearity and phase noise is investigated.

8.5.1 Dynamic Nonlinear Circuits

In reality, the output of a nonlinear dynamic circuit, such as a high power amplifier, depends on the previous inputs as well as the current input. This phenomenon is called memory effects or simply temporal dynamics. These memory effects are due to thermal effects and long-time constants in DC bias circuits [30].

The Volterra series is a general nonlinear model with memory, but its practical application is strongly restricted due to its complexity. A simpler and particular case of the Volterra series is the memory polynomial model (MPM), which is less complicated and more analytically tractable [40]-[45]. The MPM consists of several delay taps and nonlinear static functions. This model is a truncation of the general Volterra series, consisting of only the diagonal terms in the Volterra kernels. The number of parameters is, therefore, significantly reduced compared to the general Volterra series. The model is shown in Figure 8.24.

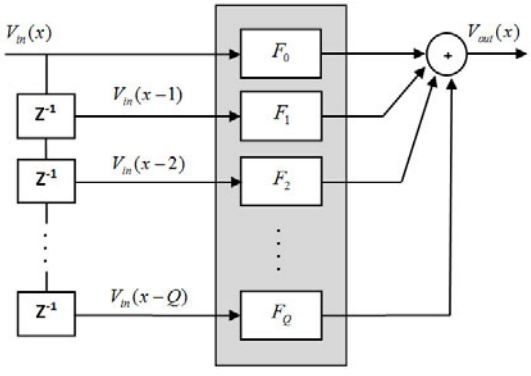


Fig. 8.24 Memory polynomial model [28]

The baseband MPM is now widely used to describe nonlinear effects in a power amplifier with memory effects. As can be seen in Figure 8.24, the general form of a baseband MPM can be written as [40]:

$$V_{Out}(x) = \sum_{q=0}^{Q} \sum_{k=1}^{K} \tilde{a}_{kq} . V_{in}(x-q)^{2k-1} \tag{8.97}$$

where $\tilde{a}_{kq}$ are complex memory polynomial coefficients; k is an integer number; $V_{in}(x)$ and $V_{out}(x)$ are the discrete input and output complex envelope signals of x^{th} sample, respectively; $q = 0,1,...,Q$ is the memory interval and equal to the sampling interval, T; and, Q and K are the maximum memory and polynomial order, respectively. Equation (8.97) only contains odd-order terms, because the signals obtained from the even-order terms are far from the carrier frequency. Equation (8.97) can also be rewritten in a compact form if $F_{kq}(x-q) = V_{in}(x-q)$:

$$V_{out}(x) = \sum_{q=0}^{Q}\sum_{k=1}^{K} \tilde{a}_{kq}.F_{kq}(x-q) = \sum_{q=0}^{Q} F_q(x-q) = F_0 + F_1 + ... + F_Q \quad (8.98)$$

where:

$$F_q = \sum_{k=1}^{K} \tilde{a}_{kq}.V_{in}(x-q)^{2k-1} \quad (8.99)$$

The least mean square error method can be used to find the complex polynomial coefficients, $\tilde{a}_{kq}$. The least mean square solution for (8.98) is $A = (F^{*T}.F)^{-1}.F^{*T}.V_{out}$.

where:

$$V_{out} = [V_{out}(0),...,V_{out}(X-1)]^T, \qquad A = [\tilde{a}_{10},...,\tilde{a}_{K0},\tilde{a}_{11},...,\tilde{a}_{K1},...,\tilde{a}_{1Q},...,\tilde{a}_{KQ}]^T,$$
$$F = [F_{10},...,F_{K0},F_{11},...,F_{K1},...,F_{1Q},...,F_{KQ}].$$

In a memoryless polynomial model, the coefficients can be calculated relative to the electrical parameters as in (8.35), (8.36), and (8.37). Moreover, the phase noise power density function of a noisy oscillator is approximated by a Lorentzian function with a uniform phase distribution. It is parameterized by its total integrated phase noise, *K*, and the -3dB bandwidth, *B*.

8.5.2 Modeling and Analysis

When the low-pass equivalent OFDM signal corrupted by transmitter oscillator phase noise passes through the third-order memory nonlinear circuit that is modeled as:

$$s(t) = \sum_{q=0}^{Q}\sum_{k=1}^{2} \tilde{a}_{kq} v_{in}(t - qt_s)^{2k-1} \quad (8.100)$$

The output voltage, $s(t)$, by using the third-order nonlinearity model can be obtained as:

$$\begin{aligned} s(t) &= \sum_{q=0}^{Q} a_{1q} \sum_{k=0}^{N-1} X_k e^{j[k\frac{2\pi}{T}(t-qt_s)+\phi_{TX}(t-qt_s)]} \\ &+ \sum_{q=0}^{Q} \frac{3a_{3q}}{4} \sum_{k_1,k_2,k_3} X_{k_1} X_{k_2} X_{k_3} e^{j[(k_1+k_2-k_3)\frac{2\pi}{T}(t-qt_s)+\phi_{TX}(t-qt_s)]} \end{aligned} \tag{8.101}$$

The approximated received signal is expressed as (8.101), where $n(t)$ is the complex Gaussian noise.

$$r(t) \simeq s(t) + \sum_{q=0}^{Q} a_{1q} . n(t) \tag{8.102}$$

Moreover, the received signal is influenced by the phase noise of the receiver (RX) LO ($\phi_{RX}(t)$) as:

$$v_{out}(t) = r(t)e^{j\phi_{RX}(t)} = s(t)e^{j\phi_{RX}(t)} + \sum_{q=0}^{Q} a_{1q} . n(t)e^{j\phi_{RX}(t)} \tag{8.103}$$

or, alternatively:

$$\begin{aligned} v_{out}(t) &= \sum_{q=0}^{Q} a_{1q} \sum_{k=0}^{N-1} X_k e^{j[k\frac{2\pi}{T}t - \frac{2\pi kqt_s}{T} + \phi_{TX}(t-qt_s) + \phi_{RX}(t)]} \\ &+ \sum_{q=0}^{Q} \frac{3a_{3q}}{4} \sum_{k_1,k_2,k_3} X_{k_1} X_{k_2} X_{k_3} e^{j[(k_1+k_2-k_3)\frac{2\pi}{T}t - \frac{2\pi kqt_s}{T} + \phi_{TX}(t-qt_s) + \phi_{RX}(t)]} + \sum_{q=0}^{Q} a_{1q} . n(t)e^{j\phi_{RX}(t)} \end{aligned} \tag{8.104}$$

Similar to the previous calculation [39], the performance degradation of the OFDM system is determined in the presence of dynamic nonlinearity and phase noise factors; therefore, the degradation factor (DF) is calculated as follows:

$$DF = 10\log \frac{(\frac{S}{N})_{Without\{\phi, Nonlinearity\}}}{(\frac{S}{N})_{With\{\phi, Nonlinearity\}}} \tag{8.105}$$

$$DF = 10\log\{\frac{(\Psi.(1-e^{-(Q+2)^2\sigma_\phi^2}).\sum_{q=0}^{Q} a_{1q}^2 + \Omega}{e^{-(Q+2)^2\sigma_\phi^2}(16N^4\times(\sum_{q=0}^{Q} a_{1q})^2 + 9N^2\times(\sum_{q=0}^{Q} a_{3q})^2\times U_3^2(N,k))}\} \tag{8.106}$$

where:

$$\Psi \triangleq 16N^3\sum_{q=0}^{Q} a_{1q}^2 + 9N.\sum_{q=0}^{Q} a_{3q}^2\times U_3{}^2(N,k)) + 16N^3(N-1).(1-e^{-(Q+2)^2\sigma_\phi^2})$$

$$\Omega \triangleq 9\times(\sum_{q=0}^{Q} a_{3q})^2.e^{-(Q+2)^2\sigma_\phi^2}(\sum_{b\neq k} U_3(N,b))^2)\times(\frac{S}{N})_{Without\{\phi,Nonlinearity\}} + 16N^4\sum_{q=0}^{Q} a_{1q}^2$$

In this analysis, if we ignore the memory and suppose that $Q=0$, the SNR is consistent with the memoryless models previously extracted. Figure 8.25 shows the simulation results of (8.106) for the BER of a 4QAM-OFDM system with different OIP3 and specified subcarriers and phase noise variances.

If we ignore the memory and nonlinearity simultaneously, the SNR is:

$$(\frac{S}{N})_{With\{\phi\}} = \frac{(\frac{S}{N})_{Without\ \{\phi\}}.e^{-4\sigma_\phi^2}}{1+(\frac{S}{N})_{Without\ \{\phi\}}\times(1-e^{-4\sigma_\phi^2})} \tag{8.107}$$

Figure 8.26 shows the simulation result of (8.107) for the BER of a 4-16-64-256QAM-OFDM system with different phase noise variances that has been compared with available results. Figure 8.27 shows the theoretical and simulation BER performances of an OFDM communication system with phase noise and memory and without nonlinearity, employing 4-16-64-256QAM modulations. For each modulation, we only considered $\sigma_\varphi^2 = 0.001$ and $Q = 0,1,2,3,4,5$.

Figure 8.28 shows the theoretical and simulation BER performances of an OFDM communication system employing 4QAM modulation with different subcarriers and memory length and specified OIP3 and phase noise. Figure 8.29 shows the BER performance of the same system with different OIP3 and memory length and specified subcarriers and phase noise.

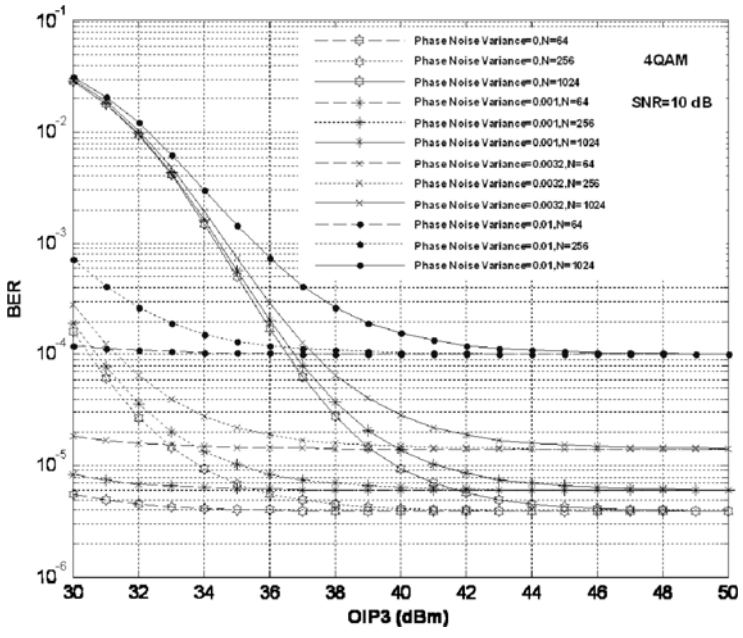


Fig. 8.25 4QAM-OFDM system with different OIP3 and specified subcarriers and phase noise variances (SNR=10dB)

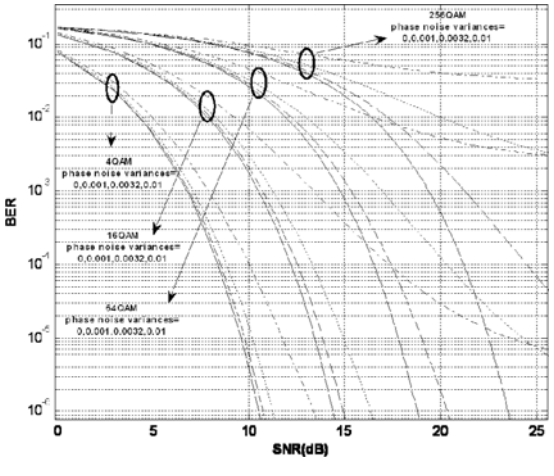


Fig. 8.26 QAM-OFDM systems with phase noise variances

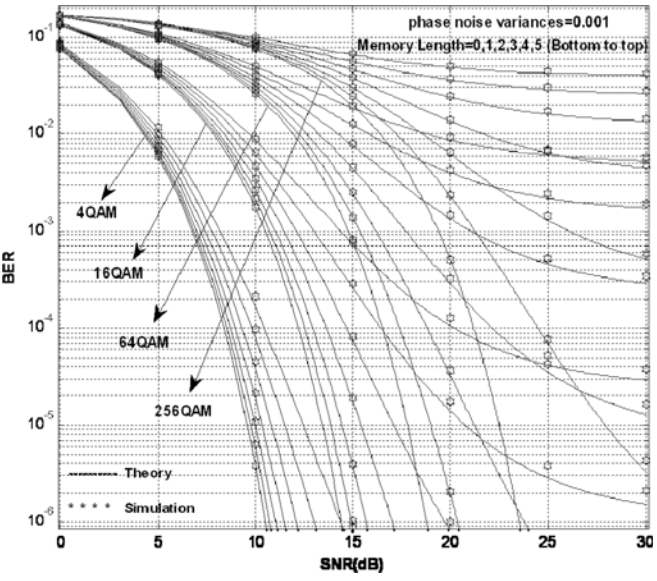


Fig. 8.27 QAM-OFDM systems with different memory lengths and specified phase noise variances

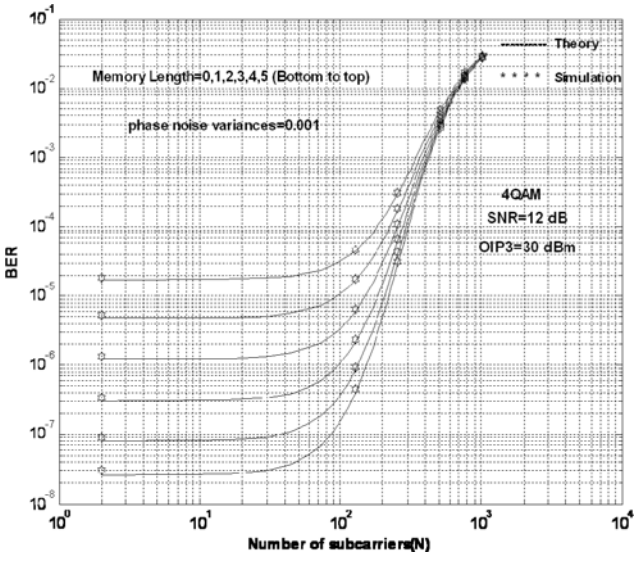


Fig. 8.28 4QAM-OFDM system with different subcarriers and memory length and specified OIP3 and phase noise variances (SNR = 12dB)

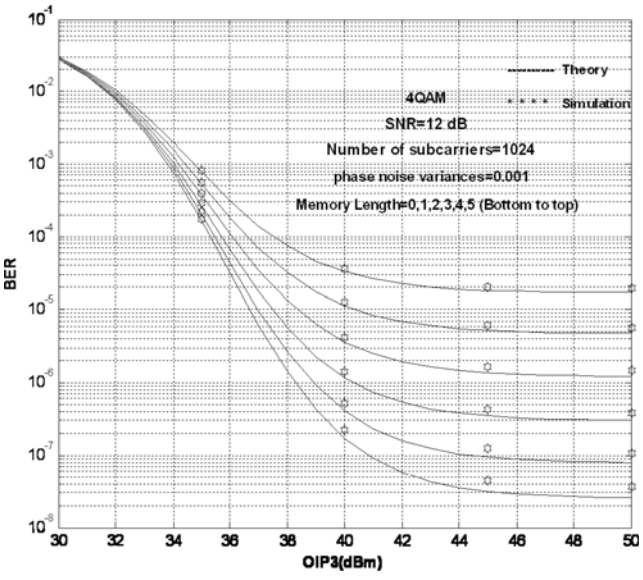


Fig. 8.29 4QAM-OFDM system with different OIP3 and memory lengths and specified subcarriers and phase noise variances (SNR = 12dB).

References

[1] Nasser, C.R., Natarajan, B., Wu, Z., Weigandt, D., Zekavat, S.A., Shattil, S.: Multi Carrier Technologies for Wireless Communication. Kluwer Academic Publishers (2002)

[2] Riaz Ahamed, S.S.: Performance Analysis of OFDM. Journal of Theoretical and Applied Information Technology (2008)

[3] Armada, A.G., Calvo, M.: Rapid Prototyping of a Test Modem for Terrestrial Broadcasting of Digital Television. IEEE Transactions on Consumer Electronics 43(4), 1100–1109 (1997)

[4] Costa, E., Pupolin, S.: M-QAM-OFDM System Performance in the Presence of a Nonlinear Amplifier and Phase Noise. IEEE Transactions on Communications 50(3) (March 2002)

[5] Come, B., Ness, R., Donnay, S., Van der Perre, L.: Impact of Front-End Non-Idealities on Bit Error Rate Performance of WLAN-OFDM Transceivers. In: IEEE Radio and Wireless Conference, September 10-13, pp. 91–94 (2000)

[6] Lee, T.H.: The Design of CMOS Radio Frequency Integrated Circuits, 2nd edn., pp. 664–667. Cambridge University Press, New York (2004)

[7] Liu, P., Bar-Ness, Y., Zhu, J.: Effects of Phase Noise at Both Transmitter and Receiver on the Performance of OFDM Systems. In: IEEE 40th Annual Conference on Information Sciences and Systems (March 2006)

[8] Wu, S., Bar-Ness, Y.: OFDM Systems in The Presence of Phase Noise: Consequences and Solutions. IEEE Transactions on Communications 52(11) (November 2004)

[9] Xu, L., Chen, Y., Feng, Z.: Effects of RF Oscillator Phase Noise on Performance of Communication System. In: 4th International Conference on Microwave and Millimeter Wave Technology, pp. 429–432 (August 2004)
[10] Ryu, H.G., Li, Y.S., Park, J.S.: Nonlinear Analysis of the Phase Noise in the OFDM Communication System. IEEE Transactions on Consumer Electronics 50(1) (February 2004)
[11] Petrovic, D., Rave, W., Fettweis, G.: Performance Degradation of Coded-OFDM Due to Phase Noise. In: 57th IEEE Vehicular Technology Conference, vol. 2(22-25), pp. 1168–1172 (April 2003)
[12] Shentu, J., Panta, K., Armstrong, J.: Effects of Phase Noise on Performance of OFDM Systems Using an ICI Cancellation Scheme. IEEE Transactions on Broadcasting 49(2), 221–224 (2003)
[13] Ryu, H.G., Lee, Y.S.: Phase Noise Analysis of the OFDM Communication System By the Standard Frequency Deviation. IEEE Transactions on Consumer Electronics 49(1), 41–47 (2003)
[14] Wu, S., Bar-Ness, Y.: Performance Analysis on Effect of Phase Noise in OFDM Systems. In: IEEE 7th International Symposium On Spread-Spectrum Techniques and Applications, Prague (September 2002)
[15] Armada, A.G.: Understanding The Effects of Phase Noise in OFDM. IEEE Transactions on Broadcasting 47(2), 153–159 (2001)
[16] Sathananthan, K., Tellambura, C.: Performance Analysis of an OFDM System With Carrier Frequency Offset and Phase Noise. In: 54th IEEE Vehicular Technical Conference, pp. 2329–2332 (2001)
[17] Tomba, L.: On The Effect of Wiener Phase Noise in OFDM Systems. IEEE Transactions on Communications 46(5), 580–583 (1998)
[18] Muschallik, C., Multimedia, T.: Influence of RF Oscillators On An OFDM Signal. IEEE Transactions on Consumer Electronics 41(3) (August 1995)
[19] Pollet, T., Blade, M.V., Moeneclaey, M.: BER Sensitivity of OFDM Systems to Carrier Frequency Offset and Wiener Phase Noise. IEEE Transactions on Communications 43(2/3/4) (February, March, April 1995)
[20] http://en.wikipedia.org/wiki/OFDM
[21] Schiek, B., Rolfes, I., Siweris, H.J.: Noise in High-Frequency Circuits and Oscillators. John Wiley & Sons (2006)
[22] http://www.zen118213.zen.co.uk/Systems_And_Devices_Files/PhaseNoise.pdf
[23] Sampei, S.: Applications of Digital Wireless Technologies to Global Wireless Communications. Prentice-Hall (1997)
[24] Deutsch, R.: Non-linear Transformations of Random Processes. Prentice-Hall International (1962)
[25] Cripps, S.C.: RF Power Amplifiers for Wireless Communications. Artech House, Norwood (1999)
[26] Chien, C.: Digital Radio on a Chip. Kluwer Academic Publishers (2001)
[27] Fager, C., et al.: A comprehensive analysis of IMD behavior in RF CMOS power amplifiers. IEEE Journal of Solid-State Circuits 39(1), 24–34 (2004)
[28] Bradley, J.: Applied Combinatorics With Problem Solving. Addison-Wesley, Reading (1990)
[29] Madani, M.H., Abdipour, A., Mohammadi, A.: Analytical Performance Evaluation of the OFDM Systems Passing Through Nonlinear Circuits. IEICE Electronics Express 7(3) (2010)

[30] Boulejfen, N., Harguem, A., Ghannouchi, F.M.: New Closed-Form Expressions for the Prediction of Multitone Inter-Modulation Distortion in Fifth-Order Nonlinear RF Circuits/Systems. IEEE Transactions on Microwave Theory and Techniques 52(1) (January 2004)

[31] Gregorio, F.H.: Analysis and Compensation of Nonlinear Power Amplifier Effects in Multi Antenna OFDM Systems, PhD dissertation, Helsinki University of Technology, Electrical and Communications Engineering Department (2007)

[32] Hieu, N.T., Jianxue, J., Ryu, H.G.: Polynomial Modeling for Efficient Predistortion of HPA. Journal of the Research Institute for Computer and Information Communication 12(3) (December 2004)

[33] Sakran, H.Y., Shokair, M., Elazm, A.A.: An Efficient Technique for Reducing PAPR of OFDM Sysyem in the Presence of Nonlinear High Power Amplifier. Progress in Electromagnetics Research C 2, 233–241 (2008)

[34] http://www.mwtinc.com/cat/wireless/pdf/WPS-495922-02.pdf

[35] Rodhe, U.: RF/Microwave Circuit Design for Wireless Applications. Wiley Interscience (2000)

[36] Lee, T.H.: The Design of CMOS Radio Frequency Integrated Circuits, 2nd edn. Cambridge University Press (2004)

[37] Schiek, B., Rolfes, I., Siweris, H.J.: Noise in High-Frequency Circuits and Oscillators. John Wiley & Sons (2006)

[38] Madani, M.H., Abdipour, A., Mohammadi, A.: Analysis of performance degradation due to non-linearity and phase noise in orthogonal frequency division multiplexing systems. IET Communications Journal 4(10), 1226–1237 (2010)

[39] Madani, M.H., Abdipour, A., Mohammadi, A.: Analytical Performance Evaluation of the OFDM Systems in the Presence of Jointly Fifth Order Nonlinearity and Phase Noise. Analog Integrated Circuits Journal (January 2011)

[40] Salman Ahmed, A.M.: Analysis, Modeling and linearization of Nonlinearity and memory effects in power amplifiers used for microwave and mobile communications, PhD dissertation, Kassel University (2005)

[41] Sakran, H.Y., Shokair, M., Elazm, A.A.: An Efficient Technique for Reducing PAPR of OFDM Sysyem in the Presence of Nonlinear High Power Amplifier. Progress in Electromagnetics Research C 2, 233–241 (2008)

[42] Ding, L., Zhou, G.T., Morgan, D.R., Ma, Z., Kenney, J.S., Kim, J., Giardina, C.R.: A robust digital baseband predistorter constructed using memory polynomials. IEEE Transactions on Wireless Communications 52(1), 159–165 (2004)

[43] Kim, J., Konstantinou, K.: Digital predistortion of wideband signals based on power amplifier model with memory. Electronic Letters 37(23), 1417–1418 (2001)

[44] Ku, H., Kenney, J.S.: Behavioral modeling of nonlinear RF power amplifiers considering memory effects. IEEE Transactions on Microwave Theory and Techniques 51(12), 2495–2504 (2003)

[45] Bohara, V.A., Ting, S.H.: Analysis of OFDM signals in nonlinear high power amplifier with memory. In: IEEE International Conference On Communications, Beijing, China, pp. 3653–3657 (May 2008)

Chapter 9
RF Impairments in MIMO Transceivers

Direct conversion transceivers are very attractive in wireless communications, due to their simplicity and ease of integration. However, their performances are degraded with radio frequency (RF) impairments. On the other hand, multiple input multiple output (MIMO) wireless communication systems promise higher spectral efficiency than what is possible with single input single output (SISO) wireless communications [1], [2]. However, the significant improvement in wireless system performance using MIMO is achieved by increasing the system complexity, which may result in higher sensitivity and lower performance in practical systems. Therefore, the degradations due to system impairments must be studied to evaluate the realistic behavior of the MIMO systems. Moreover, compensation techniques can be applied by understanding the effects of these impairment factors. These impairments may be noticed as phase noise, DC offset, in-phase / quadrature (I/Q) imbalance, and power amplifier nonlinearity [3]. In this chapter, the impairments due to phase noise, DC offset and I/Q imbalance are analyzed and the different techniques for their prevention and compensation are discussed. The power amplifier nonlinearity effects are investigated in the following chapter.

9.1 Phase Noise in MIMO Transceivers

Among various RF impairment factors, phase noise is one of the major sources of performance reduction in wireless communications systems. This section is mainly focused on the characterization and evaluation of phase noise effects on MIMO systems. By modeling the performance degradation due to the phase noise, a clear understanding of the system and compensation techniques are provided. On the other hand, by providing analytical relations to model the phase noise effects, the performance can be evaluated without the requirement of conducting very time-consuming simulation studies.

9.1.1 Phase Noise Model

Two types of phase noise models exist. When the system is phase locked, the resulting phase noise is low; and, it is modeled as a zero-mean, stationary, finite-power

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 227–263.
springerlink.com

random process [4]. On the other hand, when the system is only frequency locked, the resulting phase noise slowly varies, but is not limited; and, it is modeled as a zero-mean, non-stationary, infinite-power Wiener process [5]. We assume the system employs a phase-locked loop (PLL) for its local oscillators (LO); hence, we use the first model. For simplicity, it is assumed for simplicity that the phase noise (θ_n), is a zero-mean Gaussian distributed random variable [6], [7]:

$$f_{\Theta_n}(\theta_n) = \frac{1}{\sqrt{2\pi}\sigma_{\theta_n}} e^{\frac{-\theta_n^2}{2\sigma_{\theta_n}^2}} \tag{9.1}$$

where $\sigma_{\theta_n}^2$ is the phase noise power.

In the first step, we extract the bit error rate (BER) of a SISO system with phase noise. If a SISO system has phase noise in the receiver, the received signal is expressed as:

$$r = ha_k e^{j\theta_n} + w \tag{9.2}$$

where r is the received signal, h denotes the channel gain, a_k is the transmitted symbol, θ_n is the random phase caused by the phase noise, and w is the additive white Gaussian noise (AWGN), which is a statistically independent and identically distributed (i.i.d.) complex-valued Gaussian random variable with a variance of N_0. It is assumed that h is a complex random variable.

The BER relation for MQAM (M-ary quadrature amplitude modulation) using the signal-space concept is extracted as [8]:

$$P_{e,MQAM} \cong \frac{4}{\log_2 M}\left(1-\frac{1}{\sqrt{M}}\right)\sum_{i=1}^{\sqrt{M}/2} Q\left((2i-1)\sqrt{\frac{3E_b \log_2 M}{(M-1)N_o}}\right) \tag{9.3}$$

It is shown that, using only the first term in (9.3), the BER can be closely approximated as [8]:

$$P_{e,MQAM} \cong \frac{4}{\log_2 M}\left(1-\frac{1}{\sqrt{M}}\right)Q\left(\sqrt{\frac{3E_b \log_2 M}{(M-1)N_o}}\right) \tag{9.4}$$

It is a straightforward task to extract the BER of MQAM due to a sample function of θ_n using same procedure as in [8], [9]:

$$p_{pn}(M,\theta_n) = \frac{4(\sqrt{M}-1)}{M\log_2(M)} \sum_{i=\frac{-\sqrt{M}}{2}}^{\frac{\sqrt{M}}{2}-1} Q\left(\left(1-(2i+1)\sin\theta_n\right)\sqrt{\frac{3\gamma_s}{M-1}}\right) \tag{9.5}$$

where γ_s denotes the received signal-to-noise ratio (SNR) (i.e. $\gamma_s = received\ signal\ power / \sigma_w^2$) and M is the constellation size. Hence, the

BER can be approximated as $p_b(\theta_n) \cong \frac{SER(\theta_n)}{\log_2(M)}$. This relation provides the same results as in [8], [9], if $\theta_n = 0$. The average BER is obtained by taking the expected value of (9.5) with respect to the probability density function (pdf) of θ_n ; therefore:

$$P_e(M) = \int_{-\infty}^{\infty} p_{pn}(M, \theta_n) f_{\Theta_n}(\theta_n) d_{\theta_n} . \tag{9.6}$$

SISO Case Study

In a SISO scenario, one transmitter and one receiver are connected in an AWGN channel. The analytic, simulation and measurement results of the SISO system over the AWGN channel using a very low phase noise oscillator are presented in Figure 9.1. The measurement setup is discussed in the appendix. As can be seen, the simulation agrees very well with the analytic results. However, there is a gap of about 1 dB between the analytical and measurement results. This is due to implementation loss and the phase noiseless assumption of the signal generators that operate as LOs. Nevertheless, there is good agreement between analytical and measurement results. Figure 9.2 shows the results of the SISO system over the AWGN channel with phase noise. As shown, the measurement and analytical results are well matched.

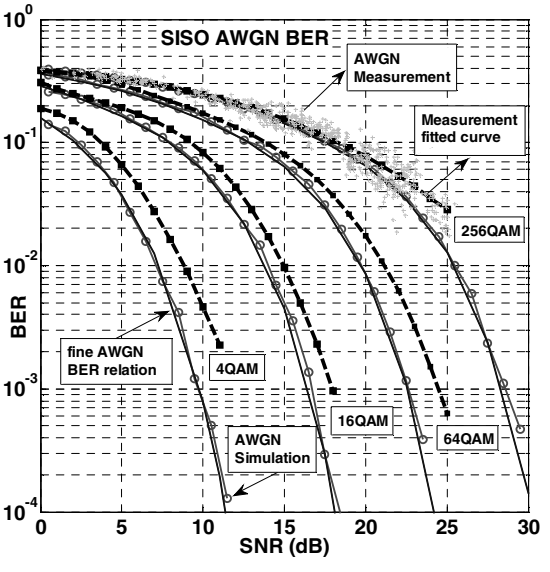


Fig. 9.1 BER of AWGN SISO without phase noise

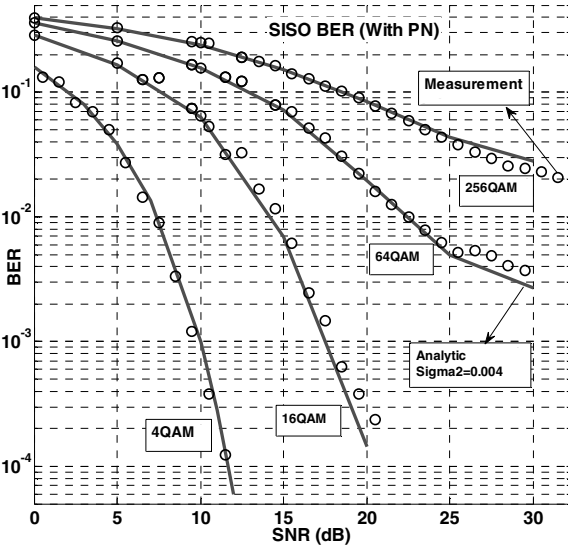


Fig. 9.2 BER of AWGN SISO with phase noise (PN)

9.1.2 Impact of Phase Noise on MIMO Systems

A flat-fading MIMO channel with N_T transmitting antennas and N_R receiving antennas is considered. Assuming perfect synchronization, the input-output relationship is given by:

$$\mathrm{y}=\mathrm{Hx}+\mathrm{w} \tag{9.7}$$

where y is an $N_R \times 1$ vector of the received signal, H denotes the $N_R \times N_T$ channel matrix, x is an $N_T \times 1$ vector of transmitted symbols, and w is an $N_R \times 1$ additive white Gaussian noise vector. The channel coefficients $\left\{h_{\mu\upsilon}\right\}_{\mu=1,\upsilon=1}^{N_R,N_T}$ are statistically i.i.d. complex-valued Gaussian random variables with a variance of 1. Components of the noise vector have the same distribution with a variance of N_0.

If singular value decomposition (SVD) is applied to H, it can be expressed as

$$\mathrm{H}=\mathrm{UDV}^H \tag{9.8}$$

where $(.)^H$ denotes the conjugate transpose; D is an $N_R \times N_T$ matrix with singular values of H, and $\left\{\sqrt{\lambda_i}\right\}_{i=1}^{m}$ and $m \overset{\Delta}{=} \min(N_t, N_r)$, are its main diagonal elements;

U and V are $N_R \times N_R$ and $N_T \times N_T$ unitary matrices with left and right singular vectors of H as their columns, respectively. Substituting (9.8) into (9.7) we have:

$$y' = Dx' + w' \tag{9.9}$$

where:

$$y' \overset{\Delta}{=} U^H y \,,\; x' \overset{\Delta}{=} V^H x \,,\; w' \overset{\Delta}{=} U^H w \tag{9.10}$$

Since $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices, the powers of x and x' are the same, as well y and y', w and w'. The equivalent model of the system is depicted in Figure 9.3, which shows that the MIMO channel is converted into m parallel SISO subchannels by SVD.

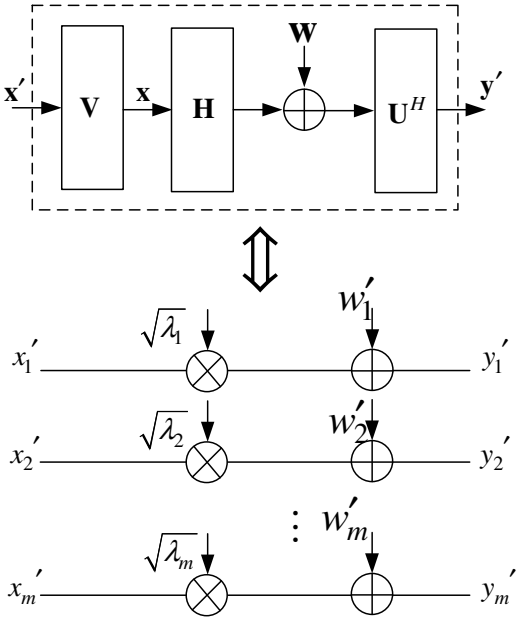


Fig. 9.3 MIMO system and equivalent model

Due to the randomness of the entries of H, the subchannel power gain (λ) is a random process. The marginal pdf of the unordered eigenvalues is [10]

$$Pdf_{\lambda_1}(\lambda_1) = \frac{1}{m}\sum_{k=0}^{m-1} \frac{k!}{(k+d)!}\left[L_k^d(\lambda_1)\right]^2 \lambda_1^d e^{-\lambda_1} \tag{9.11}$$

where $L_k^d(\lambda_1)$ is the associated Laguerre polynomial of order k, i.e.:

$$L_k^d(\lambda_1) = \sum_{l=0}^{k} (-1)^l \frac{(k+d)!}{(k-l)!(d+l)!l!} \lambda_1^l \tag{9.12}$$

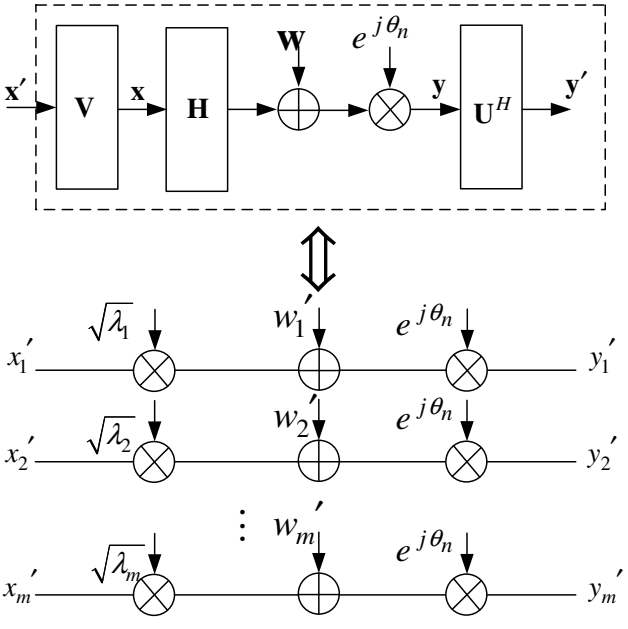


Fig. 9.4 MIMO system and equivalent model with phase noise

The MIMO system model with phase noise is presented in Figure 9.4. The input-output relationship is given by:

$$\mathbf{y}=\mathbf{\Theta}\mathbf{H}\mathbf{x}+\mathbf{\Theta}\mathbf{w} \tag{9.13}$$

where $\mathbf{\Theta}$ is an $N_R \times N_R$ matrix of phase noise and is defined as follows:

$$\Theta = \begin{pmatrix} e^{j\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{j\theta_1} \end{pmatrix}_{N_R \times N_R} \tag{9.14}$$

Equation (9.13) can be rewritten as:

$$y' = \mathrm{U}^{\mathrm{H}}\boldsymbol{\Theta}\mathrm{H}Vx' + \mathrm{U}^{\mathrm{H}}\boldsymbol{\Theta}w \tag{9.15}$$

Since the $\boldsymbol{\Theta}$ is diagonal, (9.13) can be written as (9.15); and, since the $\boldsymbol{\Theta}$ is diagonal and its entries have unit norm and random phase, the entries of $\boldsymbol{\Theta}w'$ are complex Gaussian random variables with a zero mean and a variance of N_0. It is easy to see that the received signal can be again converted into m parallel SISO subchannels by SVD. The difference from the previous situation is that each subchannel is affected by similar phase noise.

$$y' = \boldsymbol{\Theta}\mathrm{U}^{\mathrm{H}}\mathrm{H}Vx' + \boldsymbol{\Theta}w' = \boldsymbol{\Theta}Dx' + \boldsymbol{\Theta}w' \tag{9.16}$$

MIMO Case Study

The results of the MIMO system with adaptive modulation using $N_T = N_R = 4$ are presented. The Rayleigh fading channel is used in the channel simulator. Moreover, we used zero forcing receivers for MIMO detection. Figure 9.5 shows the results. As shown, there is good agreement between the analytic and measurement results.

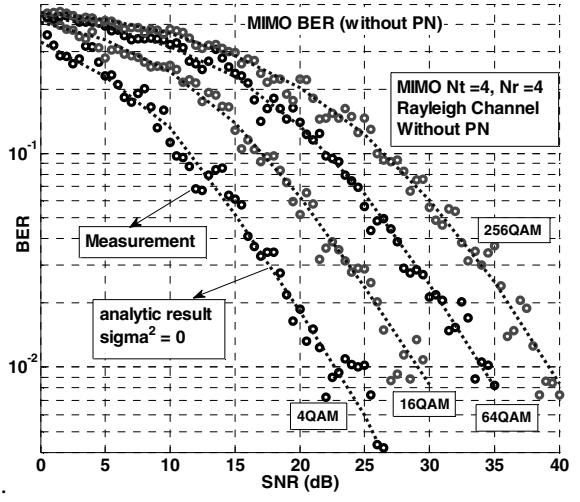
9.1.3 Adaptive Modulation MIMO System

In the adaptive modulation scheme, the channel is estimated in the receiver and fed back to the transmitter. The transmitter adapts the transmitting signal, considering the feedback information, in order to maximize the spectral efficiency. We assume that the feedback path does not introduce any errors and delay. The availability of channel information at the transmitter allows it to adapt its transmission power and rate relative to the channel variation [9].

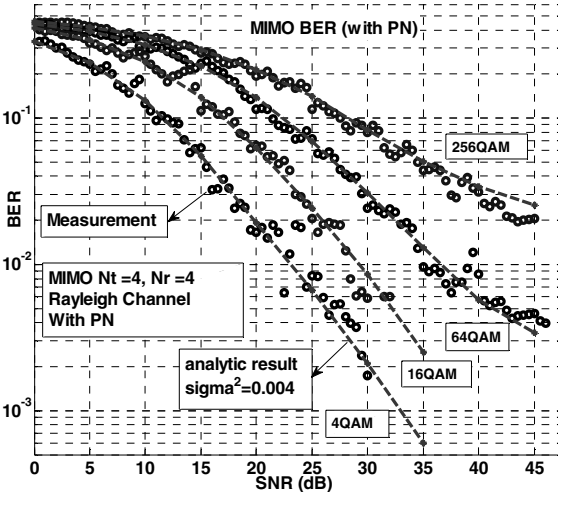
A variable-rate variable-power (VRVP) adaptation scheme is considered here. According to the above decomposition, the adaptive modulation MIMO problem can be considered as an adaptive modulation SISO problem using the unordered eigenvalue distribution [12]. Hence, the following problem should be solved by maximizing the average spectral efficiency (ASE) [13]:

$$\begin{aligned} &\underset{P_1,k_1}{\text{maximize}} \quad ASE = m\varepsilon_{\lambda_1}\left[k_1(\lambda_1)\right] \\ &\text{subject to } \varepsilon_{\lambda_1}\left[P_1(\lambda_1)\right] = P/m \\ &BER_1(\lambda_1) = BER_{tgt} \end{aligned} \tag{9.17}$$

where λ_1 is the first unordered subchannel power gain, and P_1, k_1 and BER_1 are the power, rate and instantaneous BER in the first unordered subchannel, respectively. We use only square MQAMs; therefore, the available rates are:



(a)



(b)

Fig. 9.5 Rayleigh MIMO: (a) without phase noise, (b) with phase noise

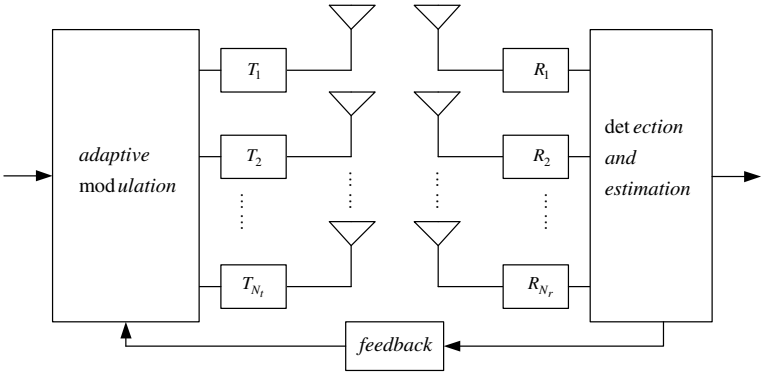


Fig. 9.6 An adaptive modulation MIMO structure

$$\begin{aligned} &k_1(\lambda_1) \in \{r_0, r_1, \ldots, r_N\} \\ &r_i = 2i \quad i = 0,1,\ldots,N \end{aligned} \tag{9.18}$$

The solution to the problem in (9.17) has been given in [3]. Therefore, the optimal power and rate policy is:

$$\begin{cases} k_1(\lambda_1) = r_i, & \upsilon_i \le \lambda_1 < \upsilon_{i+1} \\ P_1(\lambda_1) = \dfrac{M_i - 1}{\lambda_1} \dfrac{\sigma_w^2}{K}, & \upsilon_i \le \lambda_1 < \upsilon_{i+1} \end{cases} \tag{9.19}$$

where $M_i = 2^{r_i}$, $K = \dfrac{-1.5}{\ln(5BER_{tgt})}$ and υ_i denotes the power gain boundary and equals to

$$\upsilon_i = \frac{M_i - M_{i-1}}{r_i - r_{i-1}} \frac{\mu}{K} \tag{9.20}$$

where $\mu > 0$ and is the Lagrangian multiplier that is determined from the average power constraint equation:

$$\sum_{i=1}^{N} \frac{M_i - 1}{K} \left[\int_{\upsilon_i}^{\infty} \frac{Pdf_{\lambda_1}(\lambda_1)}{\lambda_1} d\lambda_1 - \int_{\upsilon_{i+1}}^{\infty} \frac{Pdf_{\lambda_1}(\lambda_1)}{\lambda_1} d\lambda_1 \right] = \frac{P}{m\sigma_w^2} \tag{9.21}$$

After obtaining μ, the average spectral efficiency can be achieved from:

$$R = m \sum_{i=1}^{N} r_i \left[\int_{\upsilon_i}^{\infty} Pdf_{\lambda_1}(\lambda_1) d\lambda_1 - \int_{\upsilon_{i+1}}^{\infty} Pdf_{\lambda_1}(\lambda_1) d\lambda_1 \right] \tag{9.22}$$

9.1.4 BER of Adaptive Modulation MIMO System with Phase Noise

The average BER of the system can be expressed as:

$$Average\ BER = \frac{\sum_{j=1}^{N} P_e(M_j) \times \log_2(M_j) P(\upsilon_j \le \lambda_1 < \upsilon_{j+1})}{\sum_{j=1}^{N} \log_2(M_j) P(\upsilon_j \le \lambda_1 < \upsilon_{j+1})} \tag{9.23}$$

where $P_e(M_j)$ is the BER corresponding to M_j, and $P(\upsilon_j \le \lambda_1 < \upsilon_{j+1})$ is the probability that λ_1 falls in the *j*th region, so that:

$$p(\upsilon_j \le \lambda_1 < \upsilon_{j+1}) = \int_{\upsilon_j}^{\upsilon_{j+1}} Pdf_{\lambda_1}(\lambda_1) d\lambda_1 \tag{9.24}$$

From (9.17), the received signal power (P_{rec}) due to adaptive modulation can be expressed as:

$$P_{rec} = \lambda_1 P_1(\lambda_1) = \frac{-2\ln(5BER_{tgt})\sigma_w^2}{3}\left(M_i - 1\right) \tag{9.25}$$

From (9.5) and (9.25), for a sample function of θ_n, $P_e(M_j)$ can be written:

$$p_{pn}(M_j, \theta_n) = \frac{4(\sqrt{M_j} - 1)}{M_j \log_2(M_j)} \sum_{i=\frac{-\sqrt{M_j}}{2}}^{\frac{\sqrt{M_j}}{2}-1} Q\left(\left(1-(2i+1)\sin\theta_n\right)\sqrt{-2\ln(5BER_{tgt})}\right) \tag{9.26}$$

The BER of the system for a constellation size of M_j is obtained by taking the expected value of (9.26) with respect to θ_n , so that:

$$P_e(M_j) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_{\theta_n}} p_{pn}(M_j, \theta_n) e^{\frac{-\theta_n^2}{2\sigma_{\theta_n}^2}} d_{\theta_n} \tag{9.27}$$

Therefore, the total BER of the adaptive modulation MIMO system under the impact of phase noise can be expressed in a closed form as:

$$BER_{PN} = \frac{\sum_{j=1}^{N}\left[P_e(M_j) \times \log_2(M_j) \times \int_{\upsilon_j}^{\upsilon_{j+1}} Pdf_{\lambda_1}(\lambda_1) d\lambda_1 \right]}{\sum_{j=1}^{N}\left[\log_2(M_j) \times \int_{\upsilon_j}^{\upsilon_{j+1}} Pdf_{\lambda_1}(\lambda_1) d\lambda_1 \right]} \tag{9.28}$$

Case Study: Adaptive Modulation MIMO

The performance of the adaptive modulation MIMO system under the impact of phase noise is presented in this section. The measurement results of the adaptive modulation MIMO system are presented for a set of average SNRs, $N_T = N_R = 4$ and $k_1(\lambda_1) \in \{0,2,4,6,8\}$. The Rayleigh fading channel is used in the channel simulator. A BER_{tgt} of 10^{-2} is selected for our measurements. Perfect channel state information (P-CSI) is assumed at both the transmitter and receiver [14], [15]. However, the phase noise affects the channel estimation. The channel is estimated in the receiver and fed back to the transmitter. To determine the training sequence length, we use the available algorithm [16]. Due to P-CSI, the total average spectral efficiency, which is only related to channel model and average SNR, is the same as the case without phase noise. This is validated in Figure 9.7 using simulation and measurement results.

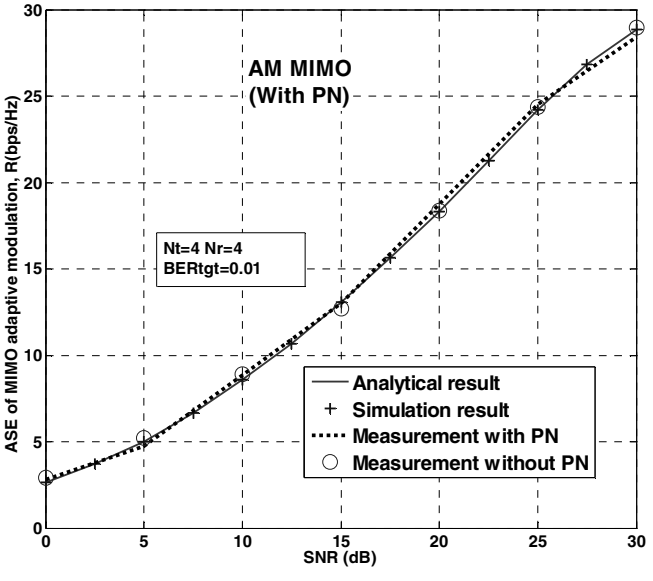
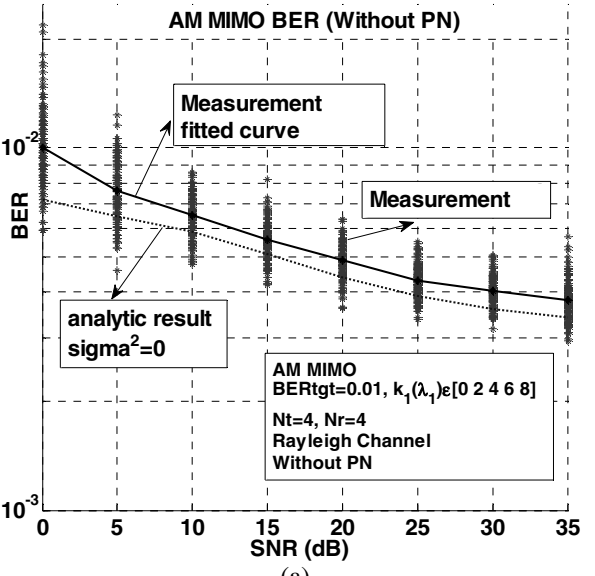


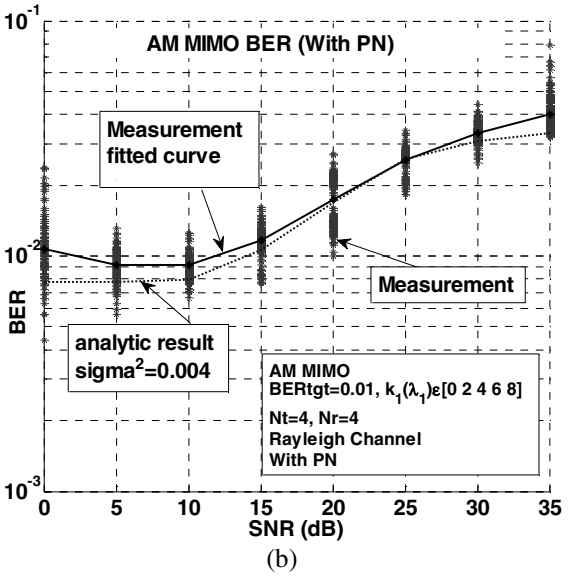
Fig. 9.7 ASE of adaptive modulation MIMO system under impact of phase noise ($BER_{tgt} = 10^{-2}$, $SNR = \frac{P}{\sigma_w^2}$)

In Figure 9.8-a, the measurement results of an adaptive modulation MIMO system in a Rayleigh channel without phase noise are depicted. As shown, the measurement results are close to the analytical results.The measurement results of an adaptive modulation MIMO system over a Rayleigh channel with phase noise are shown in Figure 9.8-b. This figure shows an unusual behavior, where by increasing the average SNR, the BER is more degraded. This is due to the use of higher constellation sizes in larger SNRs by the adaptive modulator. The higher constellation

sizes are more sensitive to the phase noise. Hence, the total impact of higher constellation sizes and greater sensitivity to phase noise results in BER degradation. As can be seen in this figure, there is good agreement between the measurement and analytical results.



(a)



(b)

Fig. 9.8 BER of an adaptive MIMO system: (a) without phase noise, (b) with phase noise

The average spectral efficiency (ASE) of an adaptive modulation (AM) MIMO system for six various antennas configurations are shown in Figure 9.9. The results are obtained with $BER_{tgt} = 10^{-3}$ and $k_1(\lambda_1) \in \{0,2,4,6,8,10\}$,and agree with the available literature [12].

The average BER rate versus the variance of phase noise, $\sigma_{\theta_n}^2$, for different SNR values are shown in Figure 9.10. The average BERs are presented for a BER_{tgt} equal to .01 and .001. As expected, by increasing the phase noise power, the BER is more degraded.

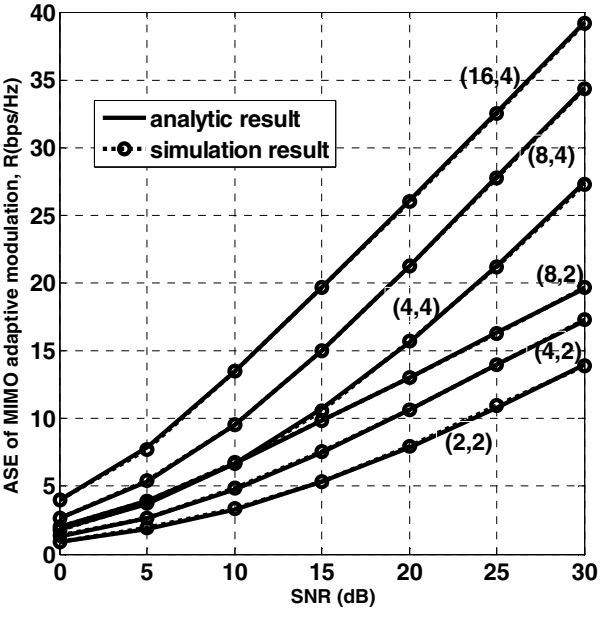


Fig. 9.9 ASE of an AM MIMO system ($BER_{tgt} = 10^{-3}$, $k_1(\lambda_1) \in \{0,2,4,6,8,10\}$, $SNR = \frac{P}{\sigma_n^2}$)

9.2 DC Offset in MIMO Transceivers

The DC offset is the other impairment that reduces the performance of MIMO transceivers. In this section, the DC offset impairment is investigated. The complete removal of the DC offset in the receiver can be very difficult; therefore, the baseband compensations are usually required [19].

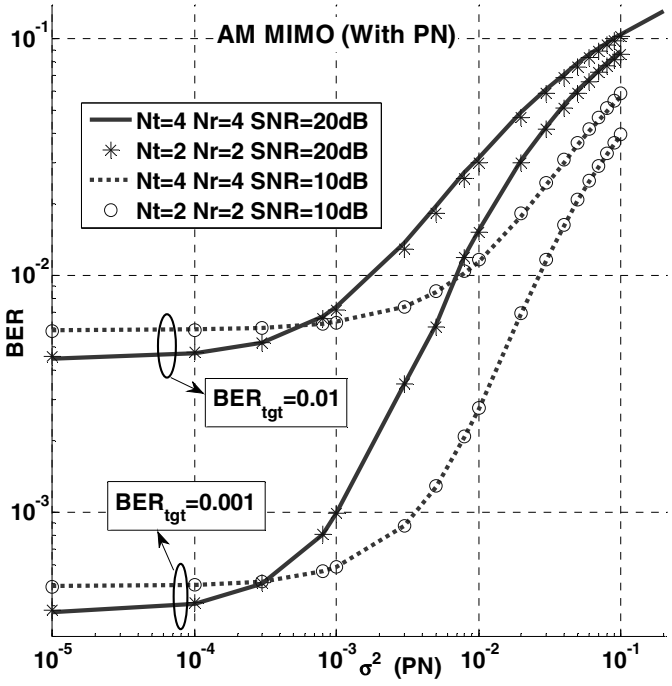


Fig. 9.10 Average BER of an AM MIMO system versus $\sigma_{\theta_n}^2$ (analytic results)

9.2.1 DC Offset

The DC offset sources can be broken down into static DC error and dynamic DC [20]. Static DC errors are generally caused by self-mixing of an LO within the receiver itself. On the other hand, dynamic DC errors are caused by time-varying effects within the receiver environment. Some examples of the dynamic DC offset errors are [21], [22]:

- Reflections of the receiver LO, which is radiated from the receiver antenna and are picked up by receiver and down-converted to DC.
- Rapid changes in signal strength (such as those caused by fading), which are not tracked quickly enough by the receiver automatic gain control (AGC). The receiver is overloaded for a short period of time, and second-order nonlinearity causes the DC offset.

DC offset degrades the BER of the receiver. It may also saturate the baseband analog-to-digital converters (ADCs), which dramatically reduces their dynamic range. Therefore, the DC offset must be removed by means of a calibration method.

Sometimes a capacitive coupling can be used, although it removes some of the wanted signal energy. When the signal has significant energy at or close to DC, capacitive coupling is not an option; however, it is possible to perform DC calibration, which is done by injecting an appropriate DC level to cancel the DC offset.

The measurement process is typically performed in the digital domain by long-term averaging. The calculated DC offset is subtracted from the received signal (typically in the analog domain). DC calibration has a disadvantage in that it is not capable of completely compensating for dynamic DC offsets, whereas static DC offsets can just be removed. In the remainder of this section, wherever DC offset mentioned, it means the uncompensated part of the DC offset. Since the nature of a dynamic DC offset is random and is best modeled by a zero-mean complex Gaussian distributed random variable [23], we assume the same model for the uncompensated part.

9.2.2 BER OF MQAM Modulation under Impact of DC Offset

In this section, a closed form expression for the MQAM BER under the impact of DC offset is extracted. The relation between the input and output of a digital communication system can be written as:

$$s' = s + c + w \tag{9.29}$$

where $s = s_I + js_Q$ is the transmitted signal, $c = c_I + jc_Q$ is the DC offset, s' is the received signal, and w is the AWGN.

It is assumed that static DC offsets have been removed entirely and that the remaining DC offset is modeled as a zero-mean, complex Gaussian distributed random variable. It is considered that the channel varies very slowly, as well as the DC offset; and, the channel is constant in the transmission interval of a frame. (Since the sources of dynamic DC offset are related to the time-varying effects within the receiver environment, the same flat model as the channel model is acceptable for DC offset.)

The constellation diagram of a square QAM signal is depicted in Figure 9.9. This figure shows the s and s' points with Gray-encoded bit mapping. If we use the method of signal-space concepts or the method of [25], the BER for a specific realization of c is:

$$P_b^{MQAM}(d,c) \cong \frac{\left(1-{}^{1}\!/\!_{\sqrt{M}}\right)}{\log_2 M} \times \sum_{i=1}^{\sqrt{M}/2} \left\{ Q\left(\frac{(2i-1)d + c_I}{{}^{\sigma_w}\!/\!_{\sqrt{2}}} \right) + Q\left(\frac{(2i-1)d + c_Q}{{}^{\sigma_w}\!/\!_{\sqrt{2}}} \right) + Q\left(\frac{(2i-1)d - c_I}{{}^{\sigma_w}\!/\!_{\sqrt{2}}} \right) + Q\left(\frac{(2i-1)d - c_Q}{{}^{\sigma_w}\!/\!_{\sqrt{2}}} \right) \right\} \tag{9.30}$$

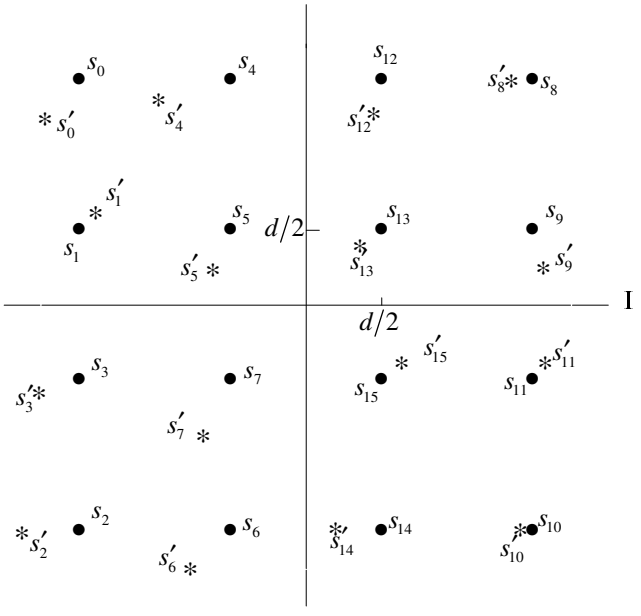


Fig. 9.9 The constellation diagram of a 16-QAM signal

where d is the Euclidean distance of two adjacent points. The Q function is very low for a large argument, so we can approximate relation (9.30) to:

$$P_b^{MQAM}(d,c) \cong \frac{\left(1-1/\sqrt{M}\right)}{\log_2 M} \times \left\{ Q\left(\frac{d+c_I}{\sigma_w/\sqrt{2}}\right) + Q\left(\frac{d+c_Q}{\sigma_w/\sqrt{2}}\right) + Q\left(\frac{d-c_I}{\sigma_w/\sqrt{2}}\right) + Q\left(\frac{d-c_Q}{\sigma_w/\sqrt{2}}\right) \right\} \tag{9.31}$$

It is known that c_I and c_Q are Gaussian with a zero-mean value and a variance of $\sigma_c^2/2$ and that:

$$\frac{d}{\sigma_w} = \sqrt{\frac{3\gamma_s}{2(M-1)}},$$

where $\gamma_s = \frac{\textit{received signal power}}{\sigma_w^2}$ is the received symbol SNR, and σ_w^2 is the AWGN variance.

After some calculations, the MQAM BER under the impact of a DC offset is approximated as:

$$P_{DC}(M) = P_b^{MQAM}(\gamma_s) \cong \frac{4\left(1-{}^{1}\!/\!_{\sqrt{M}}\right)}{\sqrt{\frac{2\pi P}{m\sigma_w^2 SDCR}}\log_2 M} \times \int_0^{\infty}\left[Q\left(\sqrt{\frac{3\gamma_s}{M-1}}+\beta\right)+Q\left(\sqrt{\frac{3\gamma_s}{M-1}}-\beta\right)\right] e^{\frac{-m\sigma_w^2\beta^2}{{}^{2P}\!/\!_{SDCR}}}\, d\beta \tag{9.32}$$

where $SDCR \overset{\Delta}{=} {}^{P}\!/\!_{m\sigma_c^2}$ (signal-to-DC ratio).

9.2.3 MIMO System Model

Each receiving path is impaired by a DC offset, which causes BER degradation in comparison with the ideal case. The input signal, x ($N_t \times 1$), and the output signal, y ($N_r \times 1$), are related by:

$$\mathrm{y} = \mathrm{H}\mathrm{x} + \mathrm{w} + \mathrm{c} \tag{9.33}$$

where w ($N_r \times 1$) is the AWGN and its entries are independent and $w_i \sim \mathcal{CN}(0, \sigma_w^2)$; and, c ($N_r \times 1$) is the DC offset.

The following definitions are supposed:

$$m \overset{\Delta}{=} \min(N_t, N_r), \quad n \overset{\Delta}{=} \max(N_t, N_r), \quad d \overset{\Delta}{=} n - m \tag{9.34}$$

If singular value decomposition (SVD) is applied to H, (9.33) is converted to:

$$\mathrm{y}' = \mathrm{D}\mathrm{x}' + \mathrm{w}' + \mathrm{c}' \tag{9.35}$$

where:

$$\mathrm{y}' \overset{\Delta}{=} \mathrm{U}^H\mathrm{y}, \ \mathrm{x}' \overset{\Delta}{=} \mathrm{V}^H\mathrm{x}, \ \mathrm{w}' \overset{\Delta}{=} \mathrm{U}^H\mathrm{w}, \ \mathrm{c}' \overset{\Delta}{=} \mathrm{U}^H\mathrm{c} \tag{9.36}$$

and U, V and D are the decomposing elements of H, which means $\mathrm{H}=\mathrm{U}\mathrm{D}\mathrm{V}^H$.

U and V are unitary matrices, so the powers of x and x', y and y', w and w', c and c' are the same, respectively. D is a diagonal matrix with the singular values of H, with $\left\{\sqrt{\lambda_i}\right\}_{i=1}^{m}$ as its main diagonal elements. As discussed in the previous section, a MIMO channel by SVD is converted into m parallel SISO subchannels.

9.2.4 BER of Adaptive Modulation MIMO under the Impact of DC Offset

In this section we investigate the impact of DC offset on adaptive modulation in MIMO systems. It is assumed that all subchannels have been impaired by the DC offset with identical variances, so the total bit error probability is equal to the average BER in each subchannel. The average BER in each subchannel can be expressed as:

$$Average\ BER = \frac{\sum_{j=1}^{N} P_{DC}(M_j) \times \log_2(M_j) P(\upsilon_j \le \lambda_1 < \upsilon_{j+1})}{\sum_{j=1}^{N} \log_2(M_j) P(\upsilon_j \le \lambda_1 < \upsilon_{j+1})} \tag{9.37}$$

where $P_{DC}(M_j)$ is the BER that corresponds to M_j, and $P(\upsilon_j \le \lambda_1 < \upsilon_{j+1})$ is the probability that λ_1 falls in the jth region, so that:

$$p(\upsilon_j \le \lambda_1 < \upsilon_{j+1}) = \int_{\upsilon_j}^{\upsilon_{j+1}} P df_{\lambda_1}(\lambda_1) d\lambda_1 \tag{9.38}$$

The adaptation scheme is

$$received\ signal\ power = \lambda_1 P_1(\lambda_1) = \frac{-2\ln(5BER_{tgt})\sigma_w^2}{3}(M_i - 1) \tag{9.39}$$

From (9.32), the $P_{DC}(M_j)$, under the impact of DC offset in the adaptation scheme, can be written as:

$$P_{DC}(M_j) \cong \frac{4\left(1 - 1/\sqrt{M_j}\right)}{\sqrt{\frac{2\pi P}{m\sigma_w^2\, SDCR}}\log_2 M_j} \times \int_0^{\infty} \left[Q\left(\sqrt{-2\ln(5BER_{tgt})} + \beta\right) + Q\left(\sqrt{-2\ln(5BER_{tgt})} - \beta\right) \right] e^{\frac{-m\sigma_w^2\beta^2}{2P/SDCR}} d\beta \tag{9.40}$$

Therefore, the total BER of MIMO adaptive modulation under the impact of DC offset can be achieved as:

$$BER_T = Average\ BER \tag{9.41}$$

where BER_T is the total BER of the adaptive MIMO system, including the effect of the DC offset.

9.2.5 BER Upper Bound of Adaptive Modulation under the Impact of DC Offset

In [26], the following upper bound relation is used for computation of the AM power gain boundaries:

$$BER \le 0.2\exp\left(\frac{-1.5\gamma_s}{M-1}\right) \tag{9.42}$$

where γ_s is the symbol SNR.

To improve the DC-affected AM system, adaptation should be done with respect to the DC offset, because a DC offset degrades the AM MIMO BER, allowing it to become greater than BER_{tgt}. If we change the mechanism of adaptation by considering DC offset, we may have a BER better than BER_{tgt}.

In order to obtain the optimal power and rate adaptation for different modulation schemes, we need an expression for each modulation technique's BER in AWGN that is easily inverted, with respect to rate and power.

Unfortunately, the approximated DC affected BERis neither easily invertible nor easily differentiable in its argument, but these properties are needed for adaptive modulation design. Therefore, we now introduce a new tight BER approximation for MQAM affected by DC offset in AWGN, which can be easily differentiated and inverted. The following upper bound for (9.32) is an acceptable approximation for optimal adaptation:

$$P_{DC}(M) \le 0.2\exp\left(\frac{-1.5\gamma_s}{(M-1)\left(1+\dfrac{P}{m\sigma_w^2 SDCR}\right)}\right) \tag{9.43}$$

As can be seen, the variance of noise (σ_w^2) has been added to the variance of the DC offset (σ_c^2); and, an upper bound relation for the BER in the presence of DC offset has been calculated from (9.42) (replacing σ_w^2 by $\sigma_w^2+\sigma_c^2$). It is a logical approximation due to the nature of the dynamic DC offset, as explained in Section 4. The accuracy of this approximation has been verified by simulation and is shown in Figure 9.12. Based on this upper bound, a new adaptation mechanism is used:

$$\begin{cases} k_1(\lambda_1) = r_i, & \upsilon_i' \le \lambda_1 < \upsilon_{i+1}' \\ P_1(\lambda_1) = \dfrac{M_i - 1}{\lambda_1}\dfrac{\sigma_w^2}{K}\left(1+\dfrac{P}{m\sigma_w^2 SDCR}\right), & \upsilon_i' \le \lambda_1 < \upsilon_{i+1}' \end{cases} \tag{9.44}$$

where

$$\upsilon_i' = \frac{M_i - M_{i-1}}{r_i - r_{i-1}} \frac{\mu}{K}\left(1 + \frac{P}{m\sigma_w^2\, SDCR}\right) \tag{9.45}$$

and μ is calculated from:

$$\sum_{i=1}^{N} \frac{M_i - 1}{K}\left[\int_{\upsilon_i'}^{\infty} \frac{Pdf_{\lambda_1}(\lambda_1)}{\lambda_1} d\lambda_1 - \int_{\upsilon_{i+1}'}^{\infty} \frac{Pdf_{\lambda_1}(\lambda_1)}{\lambda_1} d\lambda_1\right] = \frac{P}{m\sigma_w^2\left(1 + \frac{P}{m\sigma_w^2\, SDCR}\right)} \tag{9.46}$$

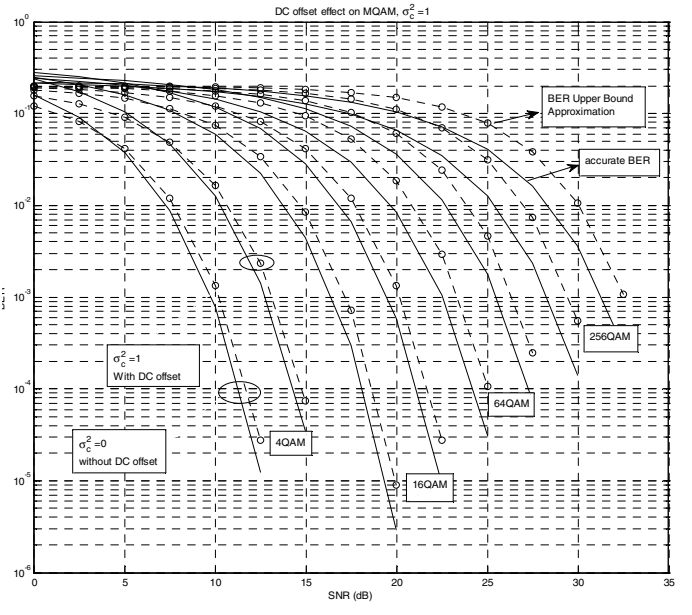


Fig. 9.12 The accurate and approximate BER under the impact of DC offset

In this case, the BER is calculated as:

$$P_{DC}(M_j) \cong \frac{4\left(1-1/\sqrt{M_j}\right)}{\sqrt{\frac{2\pi P}{m\sigma_w^2 SDCR}\log_2 M_j}} \times \int_0^{\infty}\left[Q\left(\sqrt{-2\ln(5BER_{tgt})\left(1+\frac{P}{m\sigma_w^2 SDCR}\right)}+\beta\right) + Q\left(\sqrt{-2\ln(5BER_{tgt})\left(1+\frac{P}{m\sigma_w^2 SDCR}\right)}-\beta\right)\right] e^{\frac{-m\sigma_w^2\beta^2}{2P/SDCR}}\, d\beta \tag{9.47}$$

The ASE is calculated from (9.22) with v_i' instead of v_i.

9.2.6 Throughput Analysis

The throughput is defined as the ratio of the average number of bits in packets that are successfully transmitted in any given time interval, divided by the number of attempted transmissions in that interval [29]. It is a key measure of quality of service (QoS) for wireless data transmission systems. The transmitter constructs the packet and transmits it through the air. The receiver processes the received packet. Upon detecting the packet, the receiver sends an acknowledgment, either positive or negative, back to the transmitter. For ease of analysis we assume this feedback packet goes through a separate control channel and arrives at the transmitter instantaneously and without error. If the receiver detects any error and issues a negative acknowledgement, the transmitter uses a selective repeat protocol to resend the packet. It repeats the process until the packet is successfully delivered.

We define the throughput of a system as the number of payload bits received correctly per second:

$$T = \frac{L-C}{L} R_b P_{sp} \tag{9.48}$$

where T is the throughput, L is the packet length, C is the non-information bits in a packet, R_b is the bit rate, and P_{sp} is the packet success probability. It is assumed that $L >> C$, so $\frac{L-C}{L}$ is ignored; therefore, $T = R_b P_{sp}$.

R_b is the bit rate, so it is equal to $R_b = R_s r_{i_1} + R_s r_{i_2} + \ldots + R_s r_{i_m} = R_s(r_{i_1} + \ldots + r_{i_m})$, where r_{i_1} is the number of bits per each symbol in the first subchannel, r_{i_2} is the number of bits per each symbol in the second subchannel … to r_{i_m}, which is the number of bits per each symbol in the last subchannel.

R_s is the symbol rate in each subchannel. Each packet contains N_s symbols in the first subchannel, N_s symbols in the second subchannel ... to N_s symbols in the last subchannel. The BER in the first subchannel is $P_b(r_{i_1})$ and $P_b(r_{i_2})$ in the second subchannel ... to $P_b(r_{i_m})$ in the last subchannel. The probability that all the bits are transmitted successfully in the first subchannel is $\left(1-P_b(r_{i_1})\right)^{N_s r_{i_1}}$ and … to $\left(1-P_b(r_{i_1})\right)^{N_s r_{i_m}}$, which is the probability that all the bits are transmitted successfully in the last subchannel .

Since any bit error in the packet results in a loss of the packet, the P_{sp} in our VRVP AM system is calculated as:

$$P_{sp} = \left(1-P_b(r_{i_1})\right)^{N_s r_{i_1}} \ldots\ldots \left(1-P_b(r_{i_m})\right)^{N_s r_{i_m}} \tag{9.49}$$

where N_s is the number of symbols in each subchannel, r_{i_j} , $j=1,...,m$ is the number of bits per each symbol of the *j*th subchannel, and $P_b(r_{i_j})$ is the error probability of a bit. Now the throughput can be calculated as:

$$\begin{aligned} T(r_{i_1},\ldots,r_{i_m}) &= R_s(r_{i_1}+\ldots+r_{i_m}) \\ &\times\left(1-P_b(r_{i_1})\right)^{N_s r_{i_1}} \ldots\ldots \left(1-P_b(r_{i_m})\right)^{N_s r_{i_m}} \end{aligned} \tag{9.50}$$

where R_s is the symbol rate of the system and is the same for all subchannels. Finally, the average throughput is obtained as:

$$\begin{aligned} AT = \sum_{i_1=1}^{N}\ldots\sum_{i_m=1}^{N} & T(r_{i_1},\ldots,r_{i_m})\times \\ & p(\upsilon_{i_1}\le\lambda_1<\upsilon_{i_1+1},\ldots,\upsilon_{i_m}\le\lambda_m<\upsilon_{i_m+1}) \end{aligned} \tag{9.51}$$

where $p(\upsilon_{i_1}\le\lambda_1<\upsilon_{i_1+1},\ldots,\upsilon_{i_m}\le\lambda_m<\upsilon_{i_m+1})$ is the probability that the constellation size in the first subchannel is $2^{r_{i_1}}$ and in the second subchannel is $2^{r_{i_2}}$ and … to $2^{r_{i_m}}$ in the last subchannel. The probability is calculated using the joint probability density function (pdf) of the unordered eigenvalues [28].

Relation (9.51) is very complicated, so we may approximate it with a simpler relation. For an approximation of (9.51), we may replace the $P_b(r_{i_j})$, $j=1,...,m$, with the average BER (BER_T) by using the average value approximation. On the other hand, the average of $(r_{i_1}+\ldots+r_{i_m})$ is equal to ASE, defined as *R*. Equation (9.51) can be approximated as:

$$AT \approx ASE \times R_s \times \sum_{i_1=1}^{N} \cdots \sum_{i_m=1}^{N} \left(1-P_b(r_{i_1})\right)^{N_s r_{i_1}} \cdots\cdots$$
$$\left(1-P_b(r_{i_m})\right)^{N_s r_{i_m}} \times p(\upsilon_{i_1} \le \lambda_1 < \upsilon_{i_1+1}, \ldots, \upsilon_{i_m} \le \lambda_m < \upsilon_{i_m+1})$$

$$\Rightarrow AT \approx ASE \times R_s \times \sum_{i_1=1}^{N} \cdots \sum_{i_m=1}^{N} \left(1-BER_T\right)^{N_s (r_{i_1}+r_{i_2}+\ldots+r_{i_m})} \times$$
$$p(\upsilon_{i_1} \le \lambda_1 < \upsilon_{i_1+1}, \ldots, \upsilon_{i_m} \le \lambda_m < \upsilon_{i_m+1})$$

$$AT \approx ASE \times R_s \left(1-BER_T\right)^{ASE \times N_s} \tag{9.52}$$

where BER_T is calculated from (9.41).

Equation (9.52) is very simple and easy to detect, but it causes approximation when calculating the average throughput (*AT*). Accordingly, the BER degradation of AM MIMO systems under the impact of DC offset can be evaluated. The component modulation schemes are uncoded quadrature amplitude modulations (QAMs) with $k_1(\lambda_1) \in \{0, 2, 4, 6, 8, 10\}$.

We set $BER_{tgt} = 10^{-3}$. It is assumed that the channel is known perfectly. At first, the effect of the DC offset on rgw adaptation design is not considered; hence, the ASE remains unchanged (similar to corresponding curves in [28], but later, in the new design of AM, ASE alters). Six transmit-receive antenna configurations are achieved and depicted in Figure 9.13. The simulation results are also depicted in Figure 9.13. For all receiving branches, the DC offset is set to $SDCR = 20\,dB$. The BER_T curve can be achieved from (9.41).

Figure 9.14 shows the BER_T curve for 6 different configurations of transmit-receive antenna in the non-ideal case ($SDCR = 20\,dB$) and the ideal case ($SDCR = Inf$). As can be seen in Figure 9.14, by increasing the number of antennas, the MIMO adaptive modulation BER is less degraded. The SISO system is more degraded in comparison with the MIMO systems. The (2,2) configuration has the worst BER performance, and the (16,4) has the best BER performance among the MIMO systems. This is expected, since according to the definition, σ_c^2 is proportionate to the reverse of m.

The BER of MIMO adaptive modulation under various SDCR values for (2,2) and (4,4) structures is depicted in Figure 9.15. Moreover, Monte Carlo simulation results are shown in Figure 9.15. As expected, by decreasing the $SDCR$, the MIMO adaptive modulation BER is more degraded. Due to the adaptation mechanism in the ideal case, the actual instantaneous BER lies below the BER_{tgt}.

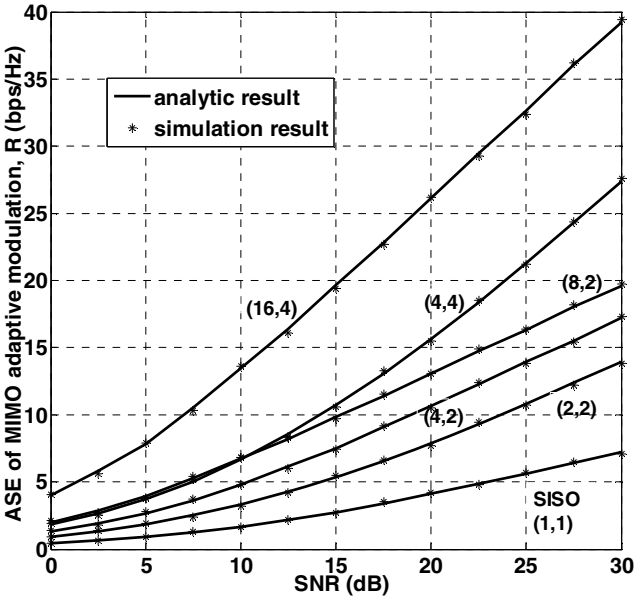


Fig. 9.13 ASE of AM MIMO in 6 various system configurations by supposing $BER_{tgt} = 10^{-3}$, versus $SNR = P / \sigma_w^2$

When the SNR is low or the DC offset is not critical, the actual BER is as good as the ideal case; however, when DC offset gets more severe, the BER becomes greater than BER_{tgt}.

Figures 9.14 and 9.15 also show that the impact of DC offset on the BER is higher for higher SNRs, meaning that the BER is more degraded in the high SNR region. This is because the AM system uses high order component modulations in this region, which are more sensitive to the DC offset. This result can also be seen easily with considering that $\frac{\gamma_s}{(M-1)}$ is constant in VRVP AM systems. It can also be seen that, in low SNRs (low P), the effect of the DC offset is very small, which is also shown in Figures 9.14 and 9.15. The BER of a new design for an AM MIMO system for a (2,2) configuration is depicted in Figure 9.16 ($SDCR = 15\,dB$, $25\,dB$). As shown in this figure, in the new design of AM, the BER remains below $BER_{tgt} = 10^{-3}$.

In Figure 9.17, the ASE of the new design of an AM MIMO system for a (2,2) configuration is depicted ($SDCR = 15\,dB, 25\,dB$) and compared to the ASE of the traditional design of an AM MIMO system. It is shown that the ASE of the new design of an AM MIMO system is decreased. Thus, improvement of the BER leads to a lower ASE. The proposed AM mechanism helps to avoid BER degradation by adjusting the spectral efficiency.

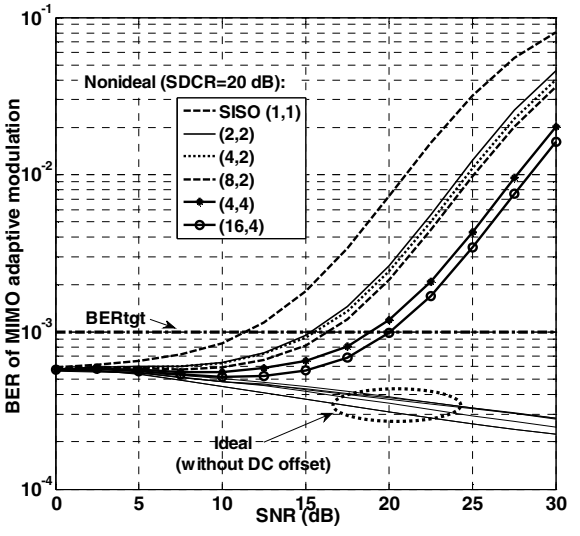


Fig. 9.14 Average BER of AM MIMO systems under the impact of DC offsets (analytic result), $BER_{tgt} = 10^{-3}$, $SDCR = 20\,dB$ $SNR = P/\sigma_w^2$

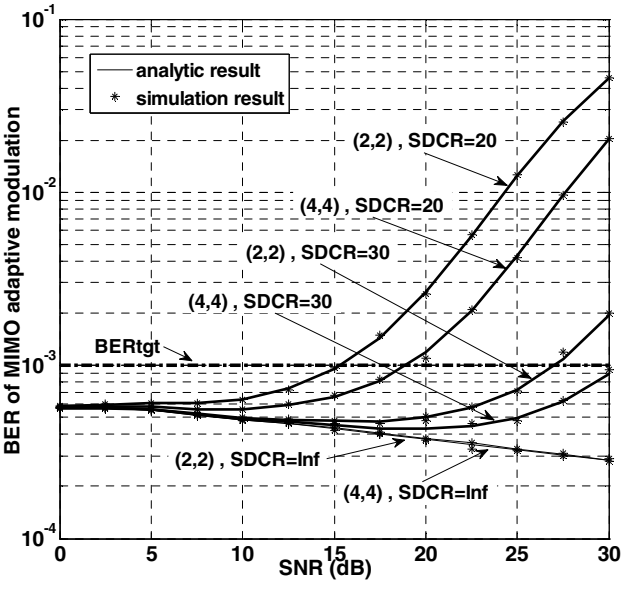


Fig. 9.15 Average BER of AM MIMO systems under the impact of DC offsets under various $SDCR$ values, $BER_{tgt} = 10^{-3}$, $SNR = P/\sigma_w^2$

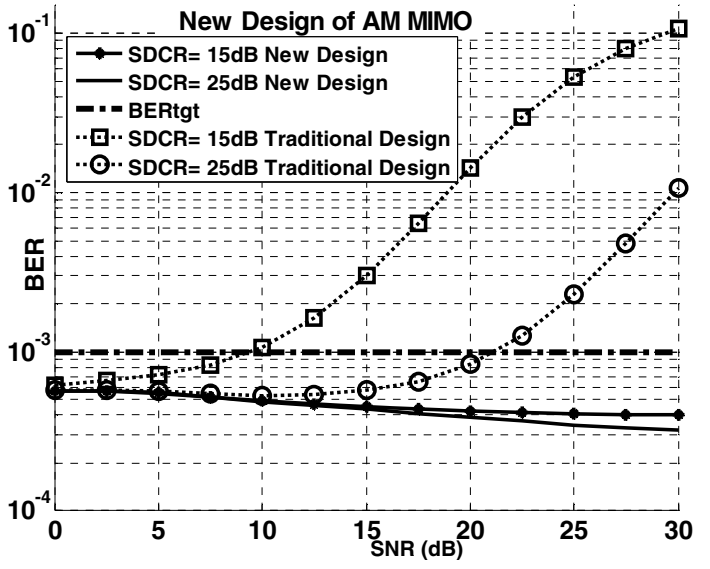


Fig. 9.16 Average BER of a new design of AM MIMO system, $N_t = 2$, $N_r = 2$ $BER_{tgt} = 10^{-3}$, $SNR = P/\sigma_w^2(2,2)$

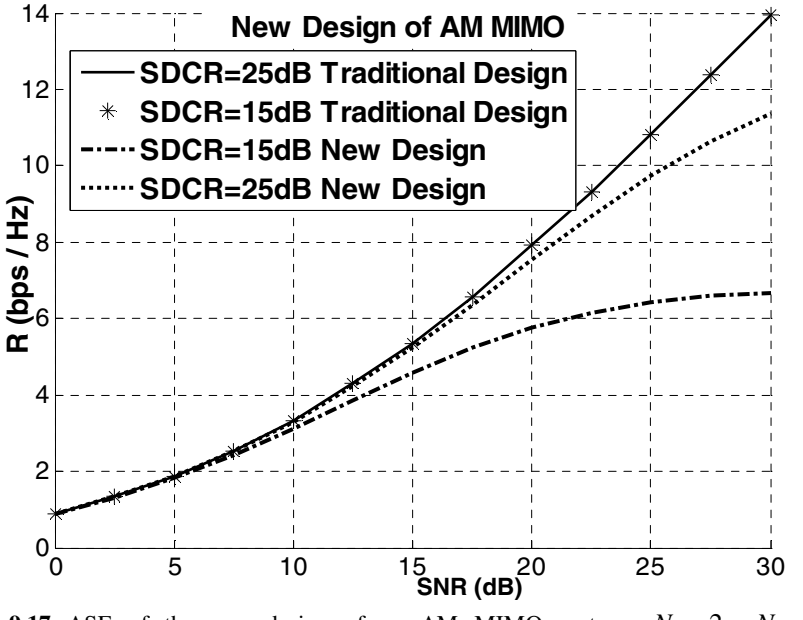


Fig. 9.17 ASE of the new design of an AM MIMO system, $N_t = 2$, $N_r = 2$, $BER_{tgt} = 10^{-3}$, $SNR = P/\sigma_w^2(2,2)$

The throughput results for the (2,2) configuration are shown in Fig. 9.18. It has been supposed that $R_s = 1 M\ Symbols/S$ and $N_s = 50$. In this figure, the throughput of an AM MIMO system under the impact of a DC offset ($SDCR = 25\,dB$) is compared to the throughput of ideal case (i.e. without DC offset) and the throughput of the new design of an AM MIMO system. As can be seen in this figure, the DC offset degrades the throughput dramatically, and the new design of AM improves it. Due to the BER degradation, the throughput is decreased in the high SNR region. There is an acceptable agreement between simulation results and analytic results in most situations.As may be seen, the difference between the analytical and simulation results is higher with larger SNRs. However, this is predictable, due to the use of larger constellation sizes by the system in higher SNRs. As discussed, the analytical relation only predicts the performance degradation due to the DC offset imperfection.

The imperfections due to the other parameters, such as phase noise and I/Q imbalance, also have impacts that are more noticeable in larger constellation sizes, which correspond to larger SNRs. Accordingly, the difference between the simulation and analytical results will be greater with high SNRs than those with small SNRs.

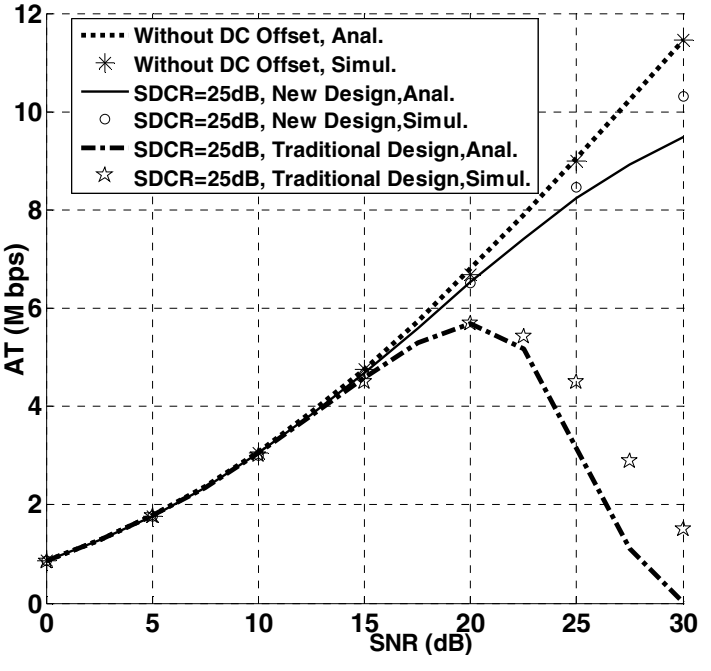


Fig. 9.18 Average throughput of AM MIMO in the presence of DC offset. $N_s = 50$ $R_s = 1 M\ Symbol/S$ $N_t = 2$, $N_r = 2$, $BER_{tgt} = 10^{-3}$, $SNR = P/\sigma_w^2$ (2,2)

The calculated closed form expressions for BER and throughput give us a useful tool to determine whether an ideally designed VRVP AM MIMO system can work in the presence of DC offset.

9.3 I/Q Imbalance in MIMO Transceivers

The I/Q (in-phase/quadrature) imbalance in MIMO systems is also a limiting factor. In this section, the impact of I/Q imbalance on adaptive modulation MQAM MIMO systems is investigated.

9.3.1 I/Q Imbalance Model

The distortion parameters, μ_i and υ_i, are related to the amplitude and phase imbalances between the I and Q branches of each receiving path in the RF/analog demodulation process.

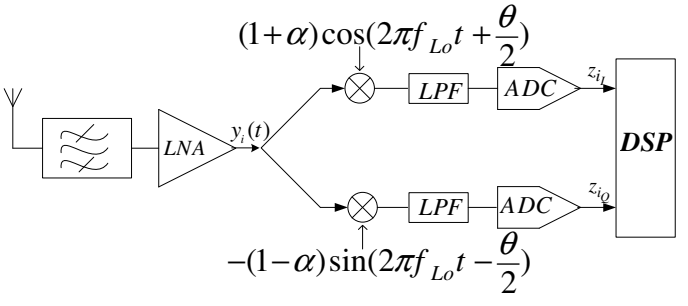


Fig. 9.19 I/Q imbalance in a direct conversion receiver

A simplified model for this distortion is the assumption that there is an amplitude imbalance, α, and a phase imbalance, θ, between the I and Q paths of the mixer, as depicted in Figure 9.19. For such a case, the μ_i and υ_i parameters can be written as [31]:

$$\begin{aligned}\mu &= \cos(\frac{\theta}{2}) + j\alpha\sin(\frac{\theta}{2}) \\ \upsilon &= \alpha\cos(\frac{\theta}{2}) - j\sin(\frac{\theta}{2})\end{aligned} \tag{9.53}$$

9.3.2 MIMO System Model

A flat-fading MIMO channel with N_t transmitting antennas and N_r receiving antennas is considered. Assuming perfect synchronization and perfect channel state information (CSI) at the receiver, the input-output relationship is given by:

$$\mathbf{y}=\mathbf{H}\mathbf{x}+\mathbf{w} \tag{9.54}$$

where y is an $N_r\times1$ vector of the received signal; H denotes the $N_r\times N_t$ channel matrix; x is an $N_t\times1$ vector of the transmitted symbols; and, w is an $N_r\times1$ additive white Gaussian noise (AWGN) vector. The entries of w are also assumed to be independent and $w_i \sim \mathcal{CN}(0,\sigma^2)$. The channel coefficients, $\{h_{\mu\upsilon}\}_{\mu=1,\upsilon=1}^{N_r,N_t}$, are statistically independent and identically distributed (i.i.d.) complex-valued Gaussian random variables with a variance of 1.

Components of the noise vector have the same distribution with a variance of N_0. The received signal, y, after distortion by I/Q imbalances becomes [32]:

$$\mathbf{z}=\boldsymbol{\mu}\mathbf{y}+\boldsymbol{\upsilon}\mathbf{y}^{*} \tag{9.55}$$

where $\boldsymbol{\mu}=\begin{pmatrix}\mu_1 & 0 & \dots & 0\\ 0 & \mu_2 & \dots & \vdots\\ \vdots & \vdots & \ddots & 0\\ 0 & 0 & \dots & \mu_{N_r}\end{pmatrix}$ and $\boldsymbol{\upsilon}=\begin{pmatrix}\upsilon_1 & 0 & \dots & 0\\ 0 & \upsilon_2 & \dots & \vdots\\ \vdots & \vdots & \ddots & 0\\ 0 & 0 & \dots & \upsilon_{N_r}\end{pmatrix}$.

The distortion parameters, μ_i and υ_i, are related to the amplitude and phase imbalances between the I and Q branches of each receiving path in the RF/analog demodulation process. A simplified model for this distortion is the assumption that there is an amplitude imbalance, α, and a phase imbalance, θ, between the I and Q paths of the mixer, as depicted in Figure 9.20.

For such a case, the μ_i and υ_i parameters can be written as [31]:

$$\mu_i=\cos(\frac{\theta_i}{2})+j\alpha_i\sin(\frac{\theta_i}{2})$$

$$\upsilon_i=\alpha_i\cos(\frac{\theta_i}{2})-j\sin(\frac{\theta_i}{2}) \tag{9.56}$$

We define $m\overset{\Delta}{=}\min(N_t,N_r)$, $n\overset{\Delta}{=}\max(N_t,N_r)$, $d\overset{\Delta}{=}n-m$. If the singular value decomposition (SVD) is applied to $\mathbf{H}$, it can be expressed as:

$$\mathbf{H}=\mathbf{U}\mathbf{D}\mathbf{V}^{H} \tag{9.57}$$

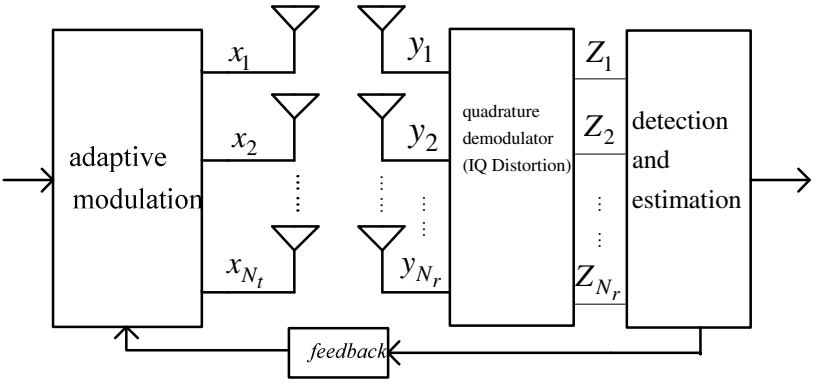


Fig. 9.20 An adaptive modulation MIMO structure

where $(.)^H$ denotes the conjugate transpose; D is an $N_r \times N_t$ matrix with singular values of H, and $\left\{\sqrt{\lambda_i}\right\}_{i=1}^{m}$ are its main diagonal elements; and, U and V are $N_r \times N_r$ and $N_t \times N_t$ unitary matrices with left and right singular vectors of H as their columns, respectively. For the detection and estimation process, the MIMO system should be converted to an m parallel SISO system. Hence, z should be multiplied by U^{H} [33]:

$$\mathrm{z}'=\mathrm{U}^{\mathrm{H}}\mu\mathrm{U}\mathrm{D}\mathbf{x}'+\mathrm{U}^{\mathrm{H}}\upsilon\mathrm{U}^{*}\mathrm{D}\mathbf{x}'^{*}+\mathrm{U}^{\mathrm{H}}\mu\mathbf{w}+\mathrm{U}^{\mathrm{H}}\upsilon\mathbf{w}^{*} \tag{9.58}$$

where $\mathbf{z}'\overset{\Delta}{=}\mathbf{U}^{\mathbf{H}}\mathbf{z}$ and $\mathbf{x}'\overset{\Delta}{=}\mathbf{V}^{\mathbf{H}}\mathbf{x}$. Since U and V are unitary matrices, the powers of x and x' are the same, as well z and z'. If we assume that $\mu' = I - \mu$, where I is the identity matrix; then, (9.58) is written as:

$$\mathrm{z}'=\mathrm{D}\mathbf{x}'-\mathbf{U}^{\mathbf{H}}\boldsymbol{\mu}'\mathbf{U}\mathbf{D}\mathbf{x}'+\mathbf{U}^{\mathbf{H}}\boldsymbol{\upsilon}\mathbf{U}^{*}\mathbf{D}\mathbf{x}'^{*}+\mathbf{U}^{\mathbf{H}}\boldsymbol{\mu}\mathbf{w}+\mathbf{U}^{\mathbf{H}}\boldsymbol{\upsilon}\mathbf{w}^{*} \tag{9.59}$$

Because D is a diagonal matrix, the elements of z' can be written as:

$$z_i' = \sqrt{\lambda_i}\,x_i' + n_{1i}' + n_{2i}' + n_{3i}' \tag{9.60}$$

where $\mathbf{n}_1' = -\mathbf{u}_\mathbf{i}^{\mathbf{H}}\boldsymbol{\mu}'\mathbf{U}\mathbf{D}\mathbf{x}' + \mathbf{u}_\mathbf{i}^{\mathbf{H}}\boldsymbol{\upsilon}\mathbf{U}^{*}\mathbf{D}\mathbf{x}'^{*}$, $\mathbf{n}_2'=\mathrm{u}_\mathrm{i}^{\mathrm{H}}\boldsymbol{\mu}\mathbf{w}$ and $\mathbf{n}_3'=\mathrm{u}_\mathrm{i}^{\mathrm{H}}\boldsymbol{\upsilon}\mathbf{w}^{*}$.

From (9.60), we can see that the I/Q imbalance introduces n_{1i}' as the cross channel interference; therefore, the subchannels are no longer parallel. Although in reality the elements of x' usually have values chosen from a set of finite symbols, we assume the entries of $\mathbf{x}'$ are i.i.d. Gaussian variables, i.e. [33]:

$$x_i' \sim \mathcal{CN}(0, \frac{P}{m}) \tag{9.61}$$

where $\frac{P}{m}$ is the average transmitting power constraint of each subchannel. Under this assumption, n' is a zero-mean complex Gaussian random variable with the variance of $\sigma_{n_i'}^2$. We can calculate $\sigma_{n'}^2$ as:

$$\begin{aligned}\sigma_{n_i}'^2 &= E[\mathbf{n_1'n_1'^H}] + E[\mathbf{n_2'n_2'^H}] + E[\mathbf{n_3'n_3'^H}] \\ &= \frac{P}{m}\mathbf{u_i}^{\mathrm{H}}\left(\mu'\mathrm{UDD^H U^H}\mu'^{*} + \boldsymbol{\upsilon}\mathrm{U^*DD^H U^T}\boldsymbol{\upsilon}^{*}\right)u_i + \sigma_w^2\mathbf{u_i}^{\mathrm{H}}\left(\mu\mu^{*} + \boldsymbol{\upsilon}\boldsymbol{\upsilon}^{*}\right)u_i\end{aligned} \tag{9.62}$$

9.3.3 Impact of I/Q Imbalance on BER of Adaptive Modulation MIMO

The BER of an MQAM signal in an AWGN channel can be approximated as:

$$BER(M_i) \simeq \frac{4(1-1/\sqrt{M_i})}{\log_2(M_i)} Q\left(\sqrt{\frac{3\gamma_s}{M_i - 1}}\right) \tag{9.63}$$

where γ_s is the symbol of the received SNR.

As seen, the received signal power (P_{rec}) due to adaptive modulation can be expressed as:

$$P_{rec} = \lambda_1 P_1(\lambda_1) = \frac{-2\ln(5BER_{tgt})\sigma_w^2}{3}(M_i - 1) \tag{9.64}$$

One can approximate the BER of a subchannel by considering (9.63) as:

$$BER_{IQ}(M_i, H) \simeq \frac{4(1-1/\sqrt{M_i})}{\log_2(M_i)} Q\left(\sqrt{\frac{-2\ln(5BER_{tgt})\sigma_w^2}{\sigma_{n'}^2}}\right) \tag{9.65}$$

$BER(M_i, H)$ has the joint pdf of unordered eigenvalues [35]. We can approximate the subchannel BER by averaging the $\sigma_{n'}^2$ using:

$$E_H[\sigma_{n'}^2] = \bar{\lambda}_1 \times \frac{P}{m} u_i^{\mathrm{H}}\left(\mu'\mu'^{*} + \upsilon\upsilon^{*}\right)u_i + \sigma_w^2 u_i^{\mathrm{H}}\left(\mu\mu^{*} + \upsilon\upsilon^{*}\right)u_i \tag{9.66}$$

where $\bar{\lambda}_1 = \int_0^{\infty} \lambda_1 P df_{\lambda_1}(\lambda_1) d\lambda_1$.

If we consider:

$$\Delta = \max\left(\bar{\lambda}_1 \frac{P}{m}(2+\alpha_j^2-2\cos(\tfrac{\theta_j}{2}))+\sigma_w^2(1+\alpha_j^2)\right) \quad j=1,...,N_r \tag{9.67}$$

An approximated upper bound for the BER can then be achieved:

$$BER_{IQ}(M_i) \le \frac{4(1-1/\sqrt{M_i})}{\log_2(M_i)} Q\left(\sqrt{\frac{-2\ln(5BER_{tgt})\sigma_w^2}{\Delta}}\right) \tag{9.68}$$

The average BER of the adaptive modulation MIMO system impaired by I/Q imbalance can be expressed as:

$$BER_{ave} = \frac{\sum_{j=1}^{N} BER_{IQ}(M_j)\times\log_2(M_j)P(\eta_j \le \lambda_1 < \eta_{j+1})}{\sum_{j=1}^{N}\log_2(M_j)P(\eta_j \le \lambda_1 < \eta_{j+1})} \tag{9.69}$$

where $P(\eta_j \le \lambda_1 < \eta_{j+1})$ is the probability that λ_1 falls in the jth region as:

$$p(\eta_j \le \lambda_1 < \eta_{j+1}) = \int_{\eta_j}^{\eta_{j+1}} Pdf_{\lambda_1}(\lambda_1)d\lambda_1 \tag{9.70}$$

9.3.4 I/Q Imbalance Compensation in Adaptive Modulation MIMO Systems

To compensate for a distorted adaptive modulation MIMO system, adaptation should be done with respect to the I/Q imbalance. In order to obtain the optimal power and rate adaptation for different modulation schemes, we need an expression for the BER in AWGN that is easily inverted, with respect to rate and power. Accordingly, we now introduce a new BER approximation for the MQAM distorted by an I/Q imbalance in AWGN.

From the results of the last section, the following upper bound can be introduced for optimal adaptation as:

$$BER_{IQ}(M) \le 0.2\exp\left(\frac{-1.5\gamma_s}{(M-1)\Delta/\sigma_w^2}\right) \tag{9.71}$$

The values of α and θ are not known at the receiver; therefore, for the implementation of a new adaptation scheme, we use the average of Δ, in terms of α and θ statistics:

$$E[\frac{\Delta}{\sigma_w^2}] = \bar{\lambda}_1 \frac{P}{m\sigma_w^2}(2 + E[\alpha^2 - 2\cos(\frac{\theta}{2})]) + (1 + E[\alpha^2]) \quad (9.72)$$

Based on the definitions of α and θ, $E[\alpha] = E[\theta] = 0$. If we assume a small θ:

$$E[\cos(\frac{\theta}{2})] \simeq 1 - \frac{1}{8}E[\theta^2] \quad (9.73)$$

Accordingly,

$$E[\frac{\Delta}{\sigma_w^2}] = \bar{\lambda}_1 \frac{P}{m\sigma_w^2}(\sigma_\alpha^2 + \frac{1}{4}\sigma_\theta^2) + 1 + \sigma_\alpha^2 \quad (9.74)$$

Now the BER relation is approximated as:

$$BER_{IQ}(M) \le 0.2\exp\left(\frac{-1.5\gamma_s}{(M-1)(\bar{\lambda}_1 \frac{P}{m\sigma_w^2}(\sigma_\alpha^2 + \frac{1}{4}\sigma_\theta^2) + 1 + \sigma_\alpha^2)}\right) \quad (9.75)$$

Based on this approximation, we introduce a new adaptation mechanism as:

$$\begin{cases} k_1(\lambda_1) = r_i, & \eta_i' \le \lambda_1 < \eta_{i+1}' \\ P_1(\lambda_1) = \dfrac{M_i - 1}{\lambda_1}\dfrac{\sigma_w^2}{K}CF, & \eta_i' \le \lambda_1 < \eta_{i+1}' \end{cases} \quad (9.76)$$

where the compensation factor (CF) is defined as:

$$CF = \bar{\lambda}_1 \frac{P}{m\sigma_w^2}(\sigma_\alpha^2 + \frac{1}{4}\sigma_\theta^2) + 1 + \sigma_\alpha^2 \quad (9.77)$$

and $\eta_i' = \frac{M_i - M_{i-1}}{r_i - r_{i-1}}\frac{\mu}{K}CF$.

The Lagrangian multiplier, β, is calculated as:

$$\sum_{i=1}^{N} \frac{M_i - 1}{K}\left[\int_{\eta_i'}^{\infty} \frac{Pdf_{\lambda_1}(\lambda_1)}{\lambda_1} d\lambda_1 - \int_{\eta_{i+1}'}^{\infty} \frac{Pdf_{\lambda_1}(\lambda_1)}{\lambda_1} d\lambda_1\right] = \frac{P}{m\sigma_w^2 CF} \quad (9.78)$$

and the ASE is calculated as:

$$R = m\sum_{i=1}^{N} r_i\left[\int_{\eta_i'}^{\infty} Pdf_{\lambda_1}(\lambda_1)d\lambda_1 - \int_{\eta_{i+1}'}^{\infty} Pdf_{\lambda_1}(\lambda_1)d\lambda_1\right] \quad (9.79)$$

9.3.5 BER Analysis

In this section, an experimental study is presented; and, the analytical relations are compared with measurement results. The effect of an I/Q imbalance (for two sets of phase and amplitude errors) on the BER of adaptive modulation MIMO is depicted in Figure 9.21. The BER of an ideal adaptive modulation MIMO is also illustrated in this figure. The comparison between the ideal and impaired systems shows the extreme degradation of the adaptive MIMO system due to the I/Q imbalance. Good agreement between the measurement results and the simulation results can be observed in this figure. It can be seen that the approximated analytic BER upper bound follows the simulation and measurement results very well. The impact of different parameters on BER is determined from this analytic expression. Figure 9.21, however, shows an unusual behavior: by increasing the average SNR, the BER became more degraded. This response is expected, because the adaptive modulator uses high-order components, which are more sensitive to I/Q imbalance, in the modulation of high SNR regions.

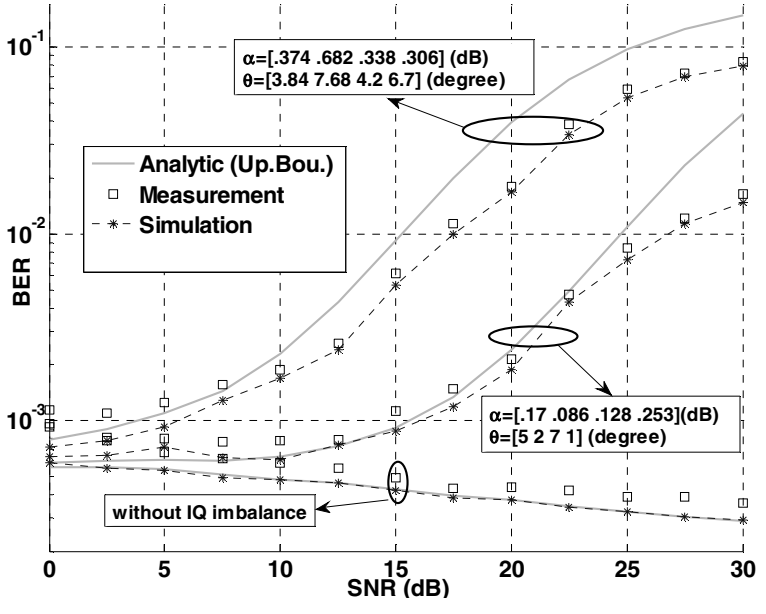


Fig. 9.21 BER of an adaptive modulation MIMO system with I/Q imbalance

Figure 9.22 (top) shows the BER of an adaptive modulation MIMO system impaired by an I/Q imbalance before and after the compensation operation by supposing $\sigma_\alpha = 0.3dB$ and $\sigma_\theta = 5^\circ$.

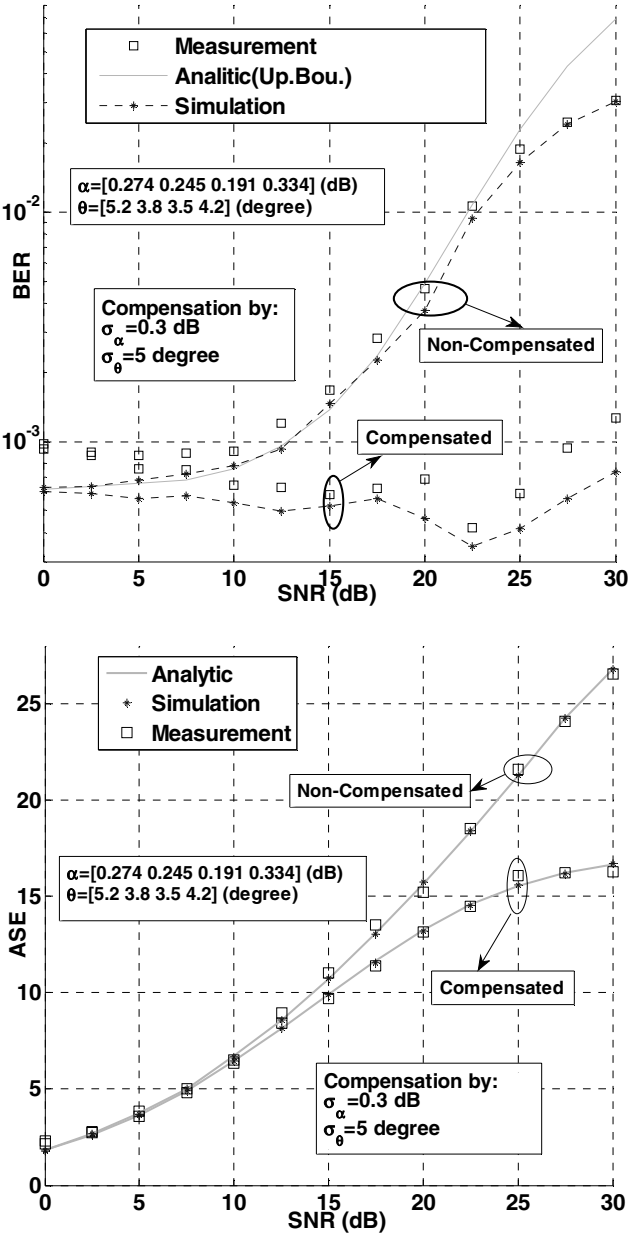


Fig. 9.22 (Top) BER of an I/Q imbalance compensated adaptive modulation MIMO system, and (bottom) ASE of an I/Q imbalance compensated adaptive modulation MIMO system

References

[1] Paulraj, A., Nabar, R., Gore, D.: Introduction to Space-Time Wireless Communications. Cambridge University Press (2003)
[2] Bolcskei, H., Gesbert, D., Papadias, C.B., Veen, A.: Space-Time Wireless System From Array Processing to MIMO Communications. Cambridge University Press (2006)
[3] Schenk, T., Linnartz, J.: RF Imperfections in High-Rate Wireless Systems, Impact and Digital Compensation. Springer, Heidelberg (2008)
[4] Piazzo, L., Mandarini, P.: Analysis of phase noise effects in OFDM modems. IEEE Transactions on Communications 50(10), 1696–1705 (2002)
[5] Tomba, L.: On the effect of wiener phase noise in OFDM systems. IEEE Transactions on Communications 46(5), 580–583 (1998)
[6] Howald, R., Kesler, S., Kam, M.: BER performance of MQAM using OFDM with RF carrier phase noise. In: Proceedings of Southeastern Symposium on System Theory, pp. 419–423 (1998)
[7] Liu, P., Bar-Ness, Y.: Closed-form expressions for BER performance in OFDM systems with phase noise. In: Proceedings of International Conference on Communications (ICC), Turkey, pp. 5366–5370 (2006)
[8] Lu, J., Letaief, K.B., Chuang, J., Liou, M.L.: M-PSK and M-QAM BER computation using signal-space concepts. IEEE Transactions on Communications 47(2), 181–184 (1999)
[9] Proakis, J.G., Salehi, M.: Digital Communications, 5th edn. McGraw-Hill (2007)
[10] Dohler, M.: Virtual antenna arrays, Ph.D. dissertation, King's College, London, U.K (2003)
[11] Chung, S.T., Goldsmith, A.J.: Degrees of freedom in adaptive modulation: A unified view. IEEE Transactions on Communications 49(9), 1561–1571 (2001)
[12] Keshavarzi, M.R., Mohammadi, A., Abdipour, A.: Characterization of Adaptive Modulation MIMO Systems in the Presence of Phase Noise. In: International Conference on Wireless Communications & Mobile Computing, IWCMC, Leipzig, Germany, pp. 1243–1247 (2009)
[13] Muschallik, C.: Influence of RF oscillators on an OFDM signal. IEEE Transactions on Consumer Electronics 41, 592–603 (1995)
[14] Gershman, A.B., Sidiropoulos, N.D.: Space-Time Processing for MIMO Communications. Wiley (2005)
[15] Tsoulos, G.: MIMO System Technology for Wireless Communications. CRC Press (2006)
[16] Zhou, S., Giannakis, G.B.: How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels? IEEE Transactions on Wireless Communications 3(4), 1285–1294 (2004)
[17] Leung, B.: VLSI for Wireless Communications. Prentice Hall (2002)
[18] Hemesi, H., Azami, F., Ghorssi, A., Abdipour, A., Mohammadi, A.: Low-complexity channel and frequency offset estimation in MIMO systems. Electronic Letters 44(9), 599–600 (2008)
[19] Larson, E.: RF and Microwave Circuit Design for Wireless Communications. Artech House, Norwood (1996)

[20] Bateman, A., Haines, D.M.: Direct conversion transceiver design for compact low-cost portable mobile radio terminals. In: Proceedings of the 43th IEEE Vehicular Technology Conference, pp. 55–62 (1989)
[21] Keshavarzi, M.R., Mohammadi, A., Abdipour, A., Ghannouchi, F.M.: Characterization and Compensation of DC Offset on Adaptive MIMO Direct Conversion Transceivers. IEICE Transactions on Communications (1), 253–261 (2011)
[22] Kenington, P.B.: RF and Baseband Techniques for Software Defined Radio. Artech House, Norwood (2005)
[23] Lindoff, B., Malm, P.: BER performance analysis of a direct conversion receiver. IEEE Transactions on Communications 50(5), 856–865 (2002)
[24] Lu, J., Letaief, K.B., Chuang, J.C.-I., Liou, M.L.: M-PSK and M-QAM BER Computation Using Signal-Space Concepts. IEEE Transactions on Communications 47, 181–184 (1999)
[25] Proakis, J.G., Salehi, M.: Digital Communications, 5th edn. McGraw-Hill (2007)
[26] Mohammadi, A., Kumar, S.: Characterization of adaptive modulators in fixed wireless ATM networks. KICS/IEEE Journal of Communications and Networks 6(2), 123–132 (2004)
[27] Dohler, M., Aghvami, H.: On the approximation of MIMO capacity. IEEE Transactions on Wireless Communications 4(1), 30–34 (2005)
[28] Zhou, Z., Vucetic, B., Dohler, M., Li, Y.: MIMO systems with adaptive modulation. IEEE Transactions on Vehicular Technology 54(5), 1828–1842 (2005)
[29] Goldsmith, A.: Wireless Communications. Cambridge University Press (2005)
[30] Pursley, M.B., Shea, J.M.: Adaptive nonuniform phase-shift-key modulation for multimedia traffic in wireless networks. IEEE Journal on Selected Areas in Communications 18, 1394–1407 (2000)
[31] Liu, C.L.: Impacts of I/Q imbalance on QPSK-OFDM-QAM detection. IEEE Transactions on Consumer Electronics 44(3), 984–989 (1998)
[32] Tarighat, A., Sayed, A.H.: MIMO OFDM receivers for systems with IQ imbalances. IEEE Transactions on Signal Processing 53(9), 3583–3596 (2005)
[33] Zhou, Z., Vucetic, B., Dohler, M., Li, Y.: MIMO systems with adaptive modulation. IEEE Transactions on Vehicular Technology 54(5), 1828–1842 (2005)
[34] Dohler, M.: Virtual antenna arrays, Ph.D. dissertation, King's College, London (2003)
[35] Telatar, I.E.: Capacity of multi-antenna Gaussian channels. European Transactions on Telecommunications 10, 585–595 (1999)
[36] Hemesi, H., Azami, F., Ghorssi, A., Mohammadi, A., Abdipour, A.: Design and Implementation of a Flexible 4x4 MIMO Testbed. In: Proceedings of the International Symposium on Telecommunications (IST 2008), pp. 266–272 (2008)

Chapter 10
Single RF Front-End MIMO Transceivers

The main limitation of using a multiple antenna architecture is the complexity and high cost of the hardware in the radio frequency (RF) section, which rise with an increase in the number of antennas. In addition, RF circuit mismatches [1-3] and coupling [4] grow with an increase in the number of antennas; consequently, these factors also limit the use of high numbers of antennas at the transceiver.

One solution to compensate for the extra hardware cost and RF circuit imperfection is the utilization of a single RF front-end in a multiple input multiple output (MIMO) system, where a single RF path is used instead of multiple parallel RF paths. This results in an RF section that has lower complexity and cost, a simpler RF design, a compact size, and lower power consumption.

A conventional MIMO receiver with N antennas at the receiver is shown in Figure 10.1. The receiver uses multiple parallel RF front-ends, where the number of RF front-ends is equal to the number of antennas. In this architecture, the baseband processing section decodes N received baseband paths to recover the signal and obtain the diversity or spatial multiplexing gain. To realize a single RF front-end path, orthogonal transmission of multiple RF streams over a single front-end must be recognized. In the following subsections, the different techniques to realize a single RF front-end in a MIMO system are discussed.

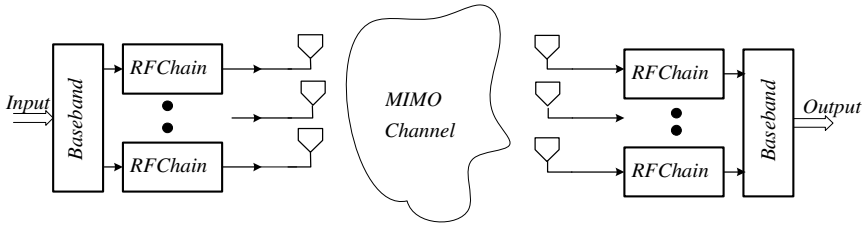


Fig. 10.1 A conventional MIMO system

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 265–288.
springerlink.com

10.1 RF Front-End MIMO Using Antenna Selection

MIMO systems increase spatial diversity; and, in proper channel conditions, the data rate may be increased using spatial multiplexing implementation. As discussed in Chapter 2, diversity methods improve the robustness of the communication system, in terms of BER, by exploiting the multiple paths between the transmitting and receiving antennas. This is achieved by space-time coding. The space-time codes are capable of delivering a diversity order of N_rN_t, where N_r and N_t are the number of receiving and transmitting antennas, respectively. A MIMO system can be also designed to provide spatial multiplexing. In this implementation, the receiver descrambles signals that are transmitted simultaneously from multiple antennas. Accordingly, it is possible to send parallel independent data streams and achieve overall system capacities that scale with $\min(N_r, N_t)$.

The performance enhancement of MIMO systems comes with an increased RF hardware complexity and cost. This problem can be mitigated using antenna selection technique at the transmitter and/or receiver to realize a simpler RF front-end MIMO system [5]. By using such a technique, some of the available antennas are selected, and the MIMO system uses fewer RF chains than the number of transmitter and/or receiver antennas. Therefore, the complexity and cost are reduced while maintaining the system performance.

The antenna selection system, both in the receiver and the transmitter, is presented in Figure 10.2. In this technique, the received signal-to-noise ratio (SNR) values must be known on both the receiver and transmitter sides. Although implementation of antenna selection in the receiver is very straightforward, antenna selection in the transmitter requires a feedback path from the transmitter to the receiver.

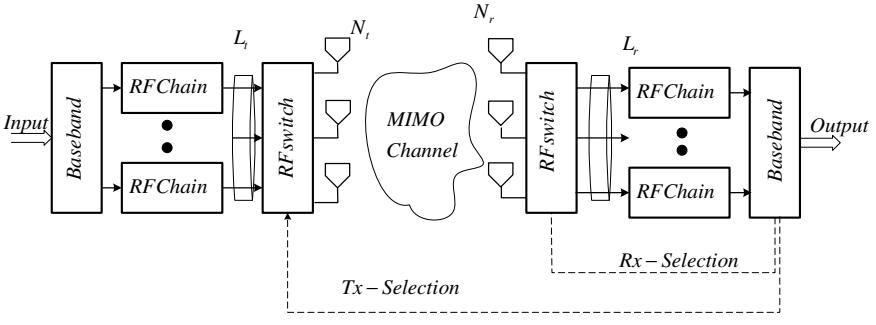


Fig. 10.2 A single antenna selection system

A common criterion for antenna selection is channel capacity maximization. An analytical bound for the channel capacity of MIMO systems with antenna selection is [6]:

$$C_{select} \le \sum_{i=1}^{L_r} \log_2\left(1+\frac{\rho}{N_t}\gamma_i\right) \tag{10.1}$$

where ρ is the mean SNR, L_r is the number of selected antennas in the receiver, and $\gamma(i)$ represents the squared norm of the ith row of H after ordering from the smallest to the largest. They are obtained by ordering a set of N_r independent and identically distributed (i.i.d.) chi-square random variables with $2N_t$ degrees of freedom (DOF). Without loss of generality, when the receiver selects the best L_r antennas that maximize capacity, the resulting channel capacity can be upper bounded [6]. It has been shown that most of the capacity of the MIMO system is retained with antenna selection, provided that the number of selected antennas on one end should be at least as many as the number of available antennas on the other end.

Some suboptimal antenna selection algorithms have also developed [8], [9]. The fastest algorithm starts with zero antennas and adds one antenna per step, where the added antenna is the greatest contribution to the channel capacity [9], [7].

The other criterion in the selection of antennas is outage probability [7], [10]. When the information transmission rate is greater than the instantaneous mutual information, an outage event occurs. In quasi-static fading, since the fading coefficients are constant over the whole frame, we cannot average them with an ergodic measure. In such an event, channel capacity does not exist in the ergodic sense. The probability of such an event is normally referred to as outage probability.

The outage probability, when the best L_r receive antennas are selected, can be upper bounded as [7]:

$$P_{select} \le P[\frac{2^R - 1}{R}\binom{N_t}{L_r}\frac{N_t N_r}{\gamma}, N_t N_r] \to k_o \gamma^{-N_t N_r} \tag{10.2}$$

where R is the information transmission rate in bits per second per hertz, $P(x,a)$ is the normalized in complete gamma function, and:

$$k_o = (\frac{2^R - 1}{R}\binom{N_t}{L_r} N_t N_r)^{N_t N_r} / (N_t N_r)! \tag{10.3}$$

Expression (10.2) suggests that the diversity order is maintained with antenna selection for any N_t, N_r and L_r. If one selects $L_t = L_r = 1$, a single RF transceiver can be realized for a MIMO system. The single antenna selection diagram is shown in Figure 10.3.

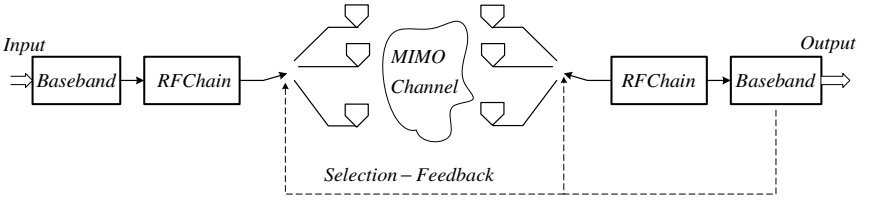


Fig. 10.3 A single antenna selection system

10.2 Single RF Front-End MIMO Using FDM

The frequency-division multiplexing (FDM) technique is presented in Figure 10.4. In this method, the signals of different antennas are shifted in frequency with mixing by different local oscillators (LOs) and added together; therefore, the multiple signals of different antennas are separated in the frequency domain. After down-conversion using a single RF front-end, the frequency shifts of the multiple streams are removed in the baseband section.

By further processing, the diversity gain is extracted. The method is also applicable when different signals are transmitted, e.g., Vertical Bell Laboratories Layered Space-Time (V-BLAST) or other spatial multiplexing MIMO transmission.

The system-level performance of a scenario with two antennas at the receiving side is studied. Assuming $s(t)$ as the passband signal with a center frequency, f_c, and a bandwidth, W. The received signal of the first antenna is $h_1 s(t) + n_1(t)$ and that of the second antenna is $h_2 s(t) + n_2(t)$, where $n_1(t)$ and $n_2(t)$ are the passband noise and h_1, h_2 are the Rayleigh flat-fading channel coefficients. The signal of the second antenna is mixed with a low-frequency oscillator at the same frequency (W) and filtered to removed the unwanted terms. The drawback of this technique is that a narrowband RF filter and extra oscillators are required.

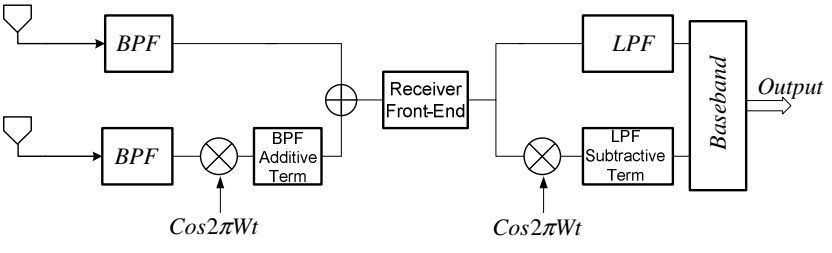


Fig. 10.4 Realization of single-branch multiantenna receiver based on FDM

The performance of this structure has been investigated using binary phase-shift keying (BPSK) modulation over two receiver antennas and is depicted in Figure 10.5. The difference between the ideal and simulated curves for diversities 1 and 2 is related to filter design and its quality factor.

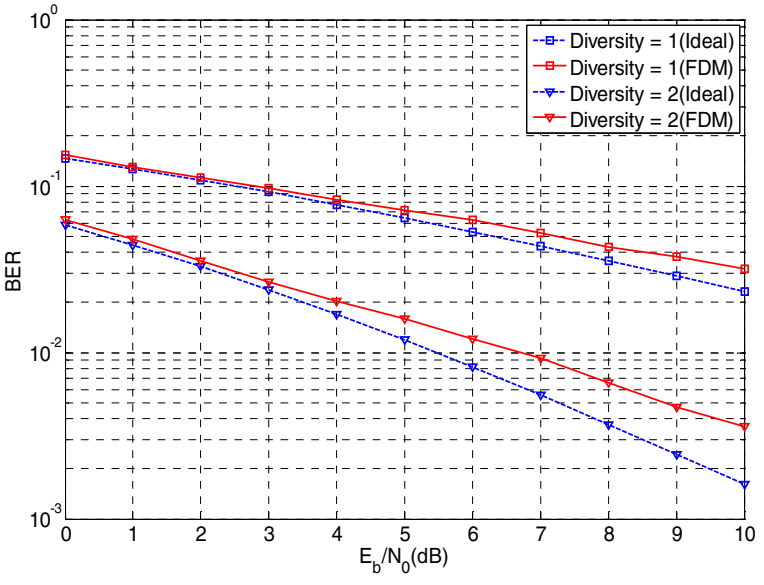


Fig. 10.5 The BER versus E_b/N_0 for single front-end receiver FDM

10.3 Single RF Front-End MIMO Using TDM

The idea of the time-division multiplexing (TDM) technique in MIMO receivers is the use of an RF switch and a single receiver, where signals from different antennas were selected, captured and digitally processed separately. In this regard, the purpose of the switch is the capture of the signals of all the antennas for every symbol time interval of the modulated signal. This introduces constraints on the switching speed and also requires a careful alignment of the data before and after the multiplexing and de-multiplexing of the multiple signals in the MIMO receivers.

In this section, we investigate the possibility of using the TDM technique to simultaneously transfer and extract the data of multiple antennas through a single RF front-end. This receiver can be realized as a single RF architecture by time-multiplexing the antennas' signals using a single-pole, multiple-throw RF switch. The received signals are de-multiplexed after RF processing according to a diversity algorithm. Figure 10.6 shows a general block diagram of this architecture. This architecture uses the same number of antennas as the conventional topology; however, instead of having multiple parallel RF front-ends, a single-pole, multiple-throw RF switch along with a single RF front-end is used to down-convert the RF signals to baseband. Finally, the signals are conveyed to the baseband section using a de-multiplexer. The proposed architecture reduces the number of RF front-ends from N to 1, which significantly reduces the overall cost and size of the multiantenna receiver and also decreases the RF design mismatch.

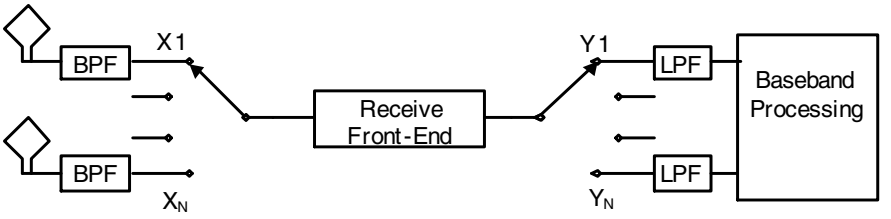


Fig. 10.6 Realization of single-branch multiantenna receiver based on TDM

The results of the time-multiplexed receiver using a raised cosine pulse shape filter with a roll-off factor of 0.5 for single antenna and two-antenna cases using BPSK modulation are shown in Figure 10.7. The zero-forcing (ZF) receivers are used [11], [16], [21]. To compare the proposed architecture with a conventional receiver, the bit error rate (BER) of the multiple front-end receivers is plotted as well. As can be seen, the diversity gain using a single time-multiplexed RF front-end was equal to the diversity gain of the multiple RF front-end.

In the following subsections, the performance of a single front-end MIMO receiver is investigated both theoretically and experimentally.

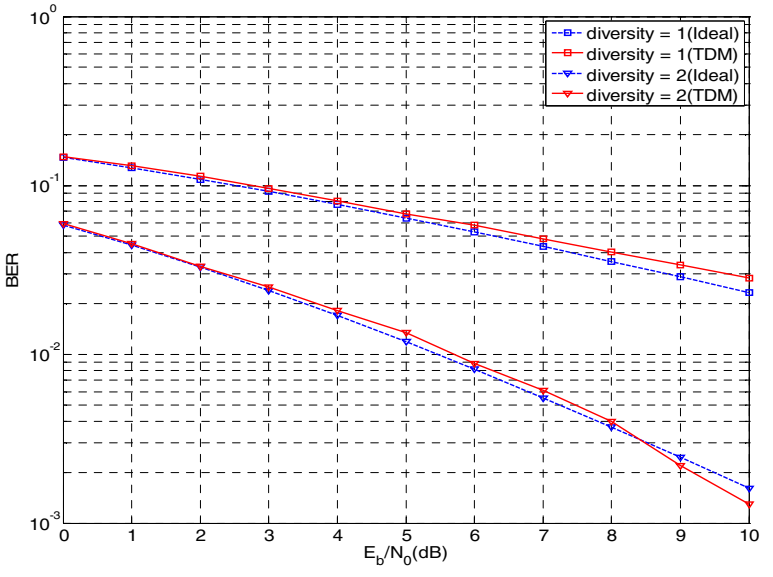


Fig. 10.7 The BER versus E_b/N_0 for single front-end receiver using TDM

10.3.1 Input Signal

The received signal considered is an MQAM (multiple or M-ary quadrature amplitude modulation) signal with a carrier frequency of f_c. The MQAM modulated signal, $s_m(t)$, can be represented as:

$$s_m(t) = A_{mc}u(t)\cos 2\pi f_c t - A_{ms}u(t)\sin 2\pi f_c t \qquad m = 1, 2, \ldots, L \tag{10.4}$$

where $u(t)$ is a signal pulse shape, f_c is the carrier frequency, A_{mc} and A_{ms} are the in-phase and quadrature signal amplitudes, M is the constellation size, and $L = \sqrt{M}$.

The nature of the received signal can be assumed to be deterministic, in order to simplify the representation of the incoming signal in the frequency domain. The spectrum of $s_m(t)$ can, therefore, be represented as:

$$S_m(f) = \frac{A_{mc}}{2}\left[U(f - f_c) + U(f + f_c)\right] - \frac{A_{ms}}{2}\left[-jU(f - f_c) + jU(f + f_c)\right] \tag{10.5}$$

where $S_m(f)$ and $U(f)$ are the Fourier transforms of $s_m(t)$ and $u(t)$, respectively.

Meanwhile, it can be easily shown that the power spectral density of an MQAM modulated signal is [12]:

$$\phi_{ss}(f) = \frac{\sigma^2}{T}\left[\left|U(f - f_c)\right|^2 + \left|U(-f - f_c)\right|^2\right] \tag{10.6}$$

where σ^2 is the variance of the information sequence in the in-phase and quadrature paths. The spectral efficiency of the MQAM signal is controlled by the baseband pulse shape, $u(t)$.

10.3.2 Switch Model

According to the block diagram of the architecture in Figure 10.6, the received MQAM signals are sampled using a single-pole, N-throw (SPNT) RF switch. It is assumed herein that the switch takes K samples from each antenna signal during the symbol period, T. The switch function, which obtains a sample from the first antenna, can be shown as:

$$z(t) = rep_{T/K}\left[rect\left(\frac{t}{\tau}\right)\right] \tag{10.7}$$

The switch waveform and its spectrum are shown in Figure 10.8. The Fourier transform of this switch waveform can be shown as:

$$Z(f) = \frac{K}{T} rep_{\frac{K}{T}}\left[\delta(f)\right]\tau \; Sinc(f\tau) \tag{10.8}$$

The number of antennas is considered to be equal to N; therefore, the switch waveform can be written as:

$$z_N(t) = rep_{T/K}\left[rect\left(\frac{t}{\tau}\right) + rect\left(\frac{t-\frac{T/K}{N}}{\tau}\right) + \ldots + rect\left(\frac{t-\frac{(N-1)T/K}{N}}{\tau}\right)\right] \tag{10.9}$$

The sampling of different signals must use the concept of orthogonal signals in the time domain to prevent overlapping among the different samples. According to (10.9), in an ideal condition, the relation between the duty factor and the period is obtained as $\tau \le \frac{T/K}{N}$.

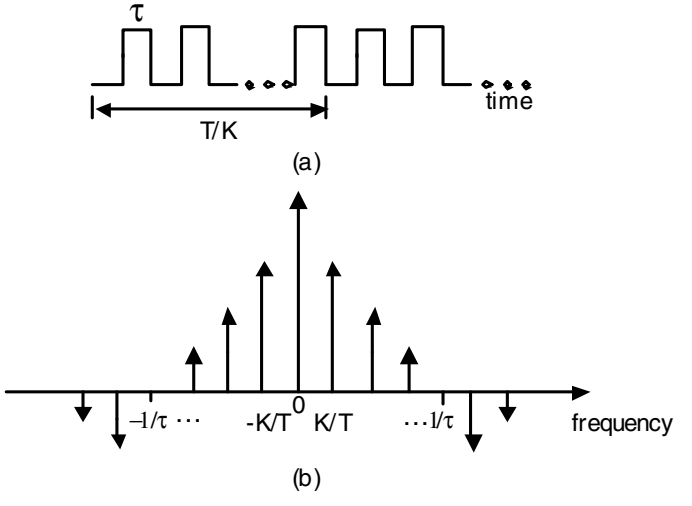


Fig. 10.8 (a) the time domain switch waveform and (b) its spectrum

10.3.3 Time Sampling of Received Signals

According to Figure 10.6, the switch takes the sample from a particular antenna output, where the received signal arrives through a Rayleigh fading channel and is corrupted by AWGN. Let us concentrate on a sample from the first antenna. This

antenna signal is called X1 in Figure 10.6. The signal, after passing through the filter, can be represented as:

$$x_1(t) = [h_1^I(t).A_{mc}u(t) + n_1^I(t)]\cos(2\pi f_c t) - [h_1^Q(t).A_{ms}u(t) + n_1^Q(t)]\sin(2\pi f_c t) \quad (10.10)$$

where $h_1^I(t)$ and $h_1^Q(t)$ are the real and imaginary parts of the channel coefficient corresponding to the first antenna, respectively, and are assumed constant over the symbol duration, T; and, $n_1^I(t)$ and $n_1^Q(t)$ are the real and imaginary parts of the AWGN at the output of band-pass filter (BPF), respectively. This signal is sampled by the RF switch and is represented as:

$$\begin{aligned} x_1^s(t) &= \left\{\left[h_1^I(t).A_{mc}u(t) + n_1^I(t)\right]\right\}\cos(2\pi f_c t)z(t) \\ &-[h_1^Q(t).A_{ms}u(t) + n_1^Q(t)]\sin(2\pi f_c t)z(t) \end{aligned} \quad (10.11)$$

where the superscript 's' represents the sampled signal.

These samples then pass through the single front-end receiver. It is assumed herein that the receiver has enough bandwidth that an RF pulse can pass through it without any distortion. The other RF samples also pass through the receiver at the different time intervals, and the samples from different antennas are orthogonal in each symbol duration. Eventually, the sample of the first antenna is delivered to the corresponding port in the baseband processing section using a low-pass filter (LPF). This port is called Y1 in Figure 10.6.

Complete synchronization between the RF switch and the de-multiplexer is assumed. If a coherent LO is used at the receiver (i.e., $x_{LO}(t)$), the output signal, $y_1(t)$, is expressed as:

$$y_1(t) = x_1^s(t)\,x_{LO}(t) \quad (10.13)$$

where the LO signals in the receiver in the in-phase and quadrature paths are $x_{LO}^I(t) = \cos(2\pi f_C t)$ and $x_{LO}^Q(t) = \sin(2\pi f_C t)$, respectively. The signal $y_1(t)$ also encompasses the second harmonic frequency component, $2f_C$, which is eventually eliminated using a LPF in the baseband section. Accordingly, the output signals can be shown as:

$$y_D^I(t) = \frac{1}{2}[h_1^I A_{mc}u(t)\,z(t) + n_1^I(t)\,z(t)] \quad (10.14\text{-a})$$

$$y_D^Q(t) = \frac{1}{2}[h_1^Q A_{ms}u(t)\,z(t) + n_1^Q(t)\,z(t)] \quad (10.14\text{-b})$$

where $y_D^I(t)$ and $y_D^Q(t)$ are the in-phase and quadrature components, respectively, of the received signals.

The switch effect is also considered by $z(t)$. If $z(t)=1$ in (10.14-a) and (10.14-b), these two equations are true for in-phase and quadrature components of the first antenna's baseband received signals. In baseband, the receiver samples each antenna's signals and performs matched filtering and maximum ratio combining. If the sampler takes *K* samples in each symbol duration, *T*, (10.14-a) and (10.14-b) are accurate for the output of the sampler when $z(t)$ is the same as (10.7) and $\tau << T$. Therefore, these equations are the same for the common receiver and the time-multiplexed receiver.

10.3.4 Baseband Processing

According to Figure 10.6, the received signals are delivered using a demultiplexer to the baseband processing section. The LPFs are used to extract the sampled signals, which must be processed in the baseband section to extract the spatial diversity or spatial multiplexing gain. This is a common assumption in narrowband models [13], where the spatial and temporal properties of the MIMO impulse response, *H(t)*, can be separated as:

$$H(t)=u(t)H \tag{10.15}$$

where *u(t)* is the pulse shape. The matched filter, $H^*(-t)$, in the MIMO baseband receiver can, therefore, be decomposed into a cascade of a space-only column matrix, *H**, and followed by a bank of time-only matched filters, $u(-t)$. It is also assumed that the channel is known on the receiver side.

In a classic multiantenna receiver with *N* antennas and *N* RF front-ends, *K* samples are taken from the baseband signal of each antenna after down-conversion. The matched filter then combines the samples and improves the SNR from the samples of each antenna. Consequently, the transmitted symbol is estimated with the information of each fading channel coefficient. The same procedure can be followed in this system, where there is a corresponding baseband signal for each antenna signal. A zero-forcing receiver to recover the transmitted symbols [20,21] can be used. When the channel coefficients are known, the symbols are estimated as:

$$y_D^I(t)=\frac{1}{2}[h_1^I A_{mc}u(t)z(t)+n_1^I(t)z(t)] \tag{10.16a}$$

$$y_D^Q(t)=\frac{1}{2}[h_1^Q A_{ms}u(t)z(t)+n_1^Q(t)z(t)] \tag{10.17a}$$

where A_{mc} and A_{ms} are the estimated in-phase and quadrature signal amplitudes, respectively.

10.3.5 Experimental Studies

Experimental studies can evaluate the performance of the proposed time-multiplexed single RF front-end receiver and validate the concept of the topology in practical scenarios. The experimental studies were conducted for a 16-QAM modulation scheme, where the number of receiving antennas was selected as $N = 2$. The 16-QAM signaling scheme was eight times oversampled and passed through a raised cosine filter with a roll-off factor of 0.3 [14]. The channel model was a Rayleigh fading channel with $f_d T_s = 0.3$. The maximum ratio combining method was used at the receiver, in order to achieve the maximum diversity.

The objective was a demonstration of the potential of using the proposed time-multiplexed single-branch receiver instead of conventional multiantenna receivers. The measurement setup is shown in Figure 10.9. It consisted of two identical signal generators, which were used to emulate the signal received from the two antennas through uncorrelated fading channels. These two signal generators were connected through a general purpose interface bus (GPIB) connection and controlled with a computer. An evaluation board (MAX2830) was used as the receiver.

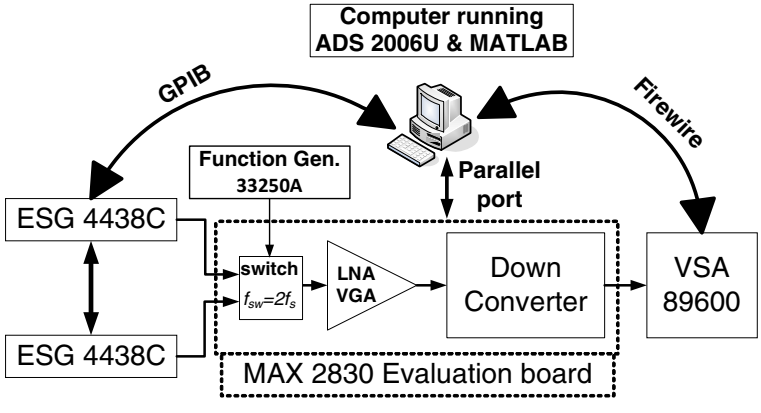


Fig. 10.9 Block diagram of the measurement setup

The MAX2830 is a completely integrated solution for the implementation of RF transceivers: it contains two RF inputs, an antenna diversity switch, a low noise amplifier (LNA), a programmable voltage-gain amplifier (VGA), an RF-to-baseband down-converter and a programmable LPF. Finally, the output of the receiver board was connected to a baseband vector signal analyzer (VSA), which was controlled with the computer [14].

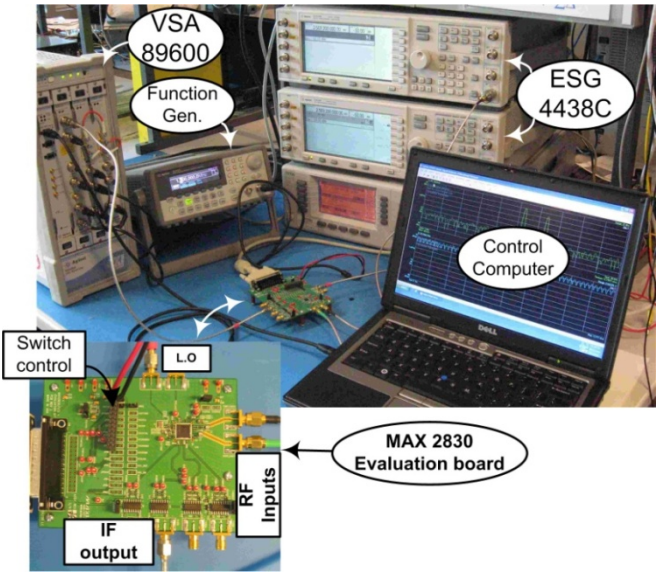


Fig. 10.10 Block diagram of the measurement setup

Two signal sources transmitted the RF signal with the same carrier frequency. The faded 16-QAM modulated signals were downloaded to each source. The outputs of the two sources were connected to the two RF inputs of the receiver board. The received signals through the RF inputs were time-multiplexed using the integrated antenna diversity switch on the receiver board. The switch was controlled by a function generator, where the switch speed was adjusted based on the sampling rate of the modulated signal. In this case, the sampling rate was $f_s = 1.6\ MSPS$; therefore, the switch speed was set to $f_{sw} = 3.2\ MHz$. The receiver board was programmed to down-convert the RF signal to a low interemediate frequency (IF) signal with $f_{IF} = 3.2\ MHz$. The IF signal was then captured with the VSA.

Finally, an off-line process was carried out for down-conversion, de-multiplexing, maximal ratio combining and low-pass filtering of the captured signal, in order to evaluate the performance of the proposed topology based on the measured BER. A picture of the measurement setup is shown in Figure 10.10. Figure 10.11 shows the measured power spectra of the received signal after time-multiplexing and RF down-conversion, but before de-multiplexing. The signal power spectra is split into three lobes: the main one is at $f_{IF} = 3.2\ MHz$, and the two side lobes are at $f_{IF} \pm \frac{f_{sw}}{2} = 3.2 \pm 1.6\ MHz$. This spectrum was exactly as expected from the analytical explanation in the previous section.

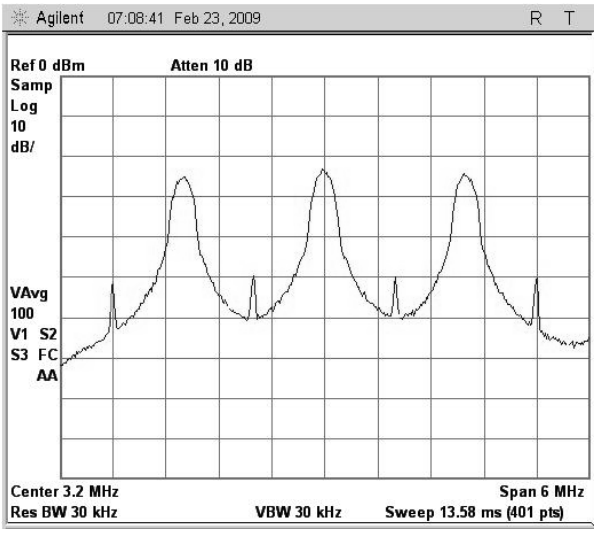
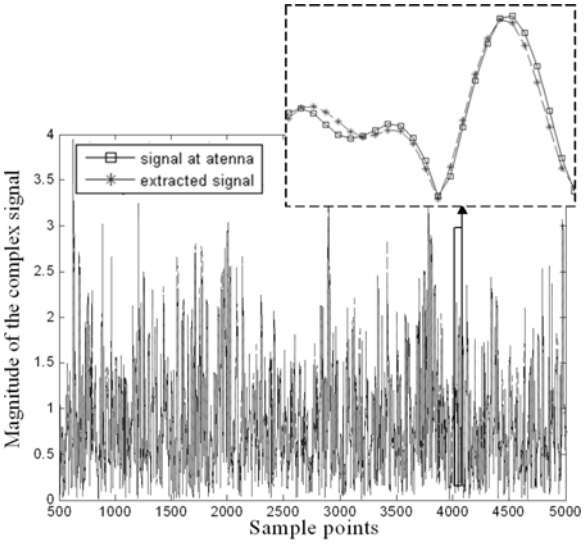


Fig. 10.11 Measured power spectra of the time-multiplexed signal before baseband demultiplexing

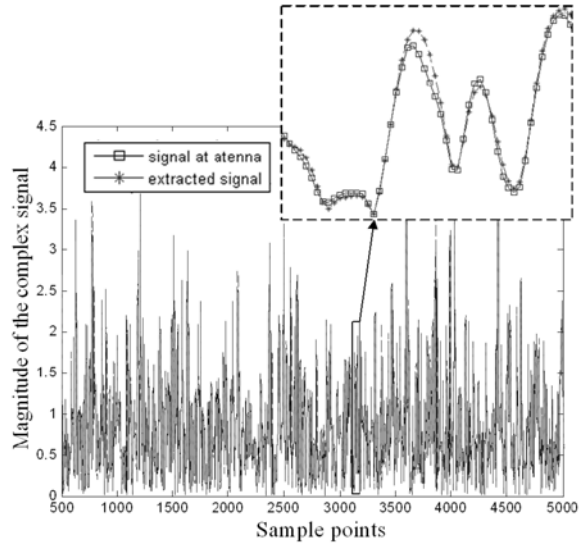
It should be noted that the proposed topology requires wider bandwidth than conventional multiantenna receivers. This results in an extra cost; however, the overall cost is reduced, as the number of RF front-ends is decreased from N in conventional multiantenna receivers to 1 in the proposed single RF front-end receiver.

Figures 10.12(a) and 10.12(b) compare the envelopes of the RF signals and the extracted baseband signals related to each of the antennas after passing through the proposed receiver topology for SNR = 18 dB. The results show very good agreement between the two signals, thereby proving the functionality of the proposed architecture [14].

The measurement BER of the proposed time-multiplexed single RF front-end receiver is shown in Figure 10.13. The measured BER results follow the analytical results of an ideal SIMO (single input multiple out) system with two antennas. The small deviation between the simulation and measurement results can be attributed to the unavoidable errors between practical measurement and theoretical simulation. For comparison, the simulation and measurement BER results of a single-input single-output (SISO) case is also plotted in Figure 10.11. As can be seen, having multiple antennas (in this case, two antennas) improved the BER performance of the communication systems due to the diversity gain. Moreover, Figure 10.13 shows that the diversity gain using the proposed time-multiplexed single branch receiver was equal to the diversity gain of the multiple RF front-ends. This validates the idea of using a time-multiplexed RF front-end in multiantenna systems. It should be noted that this topology requires wider bandwidth than conventional receivers. Accordingly, it is most suitable for the single user applications or multi-user applications when the adjacent channel interference can be neglected.



(a)



(b)

Fig. 10.12 Comparing the time domain signals before and after the proposed time-multiplexed single branch receiver: (a) antenna 1 and (b) antenna 2

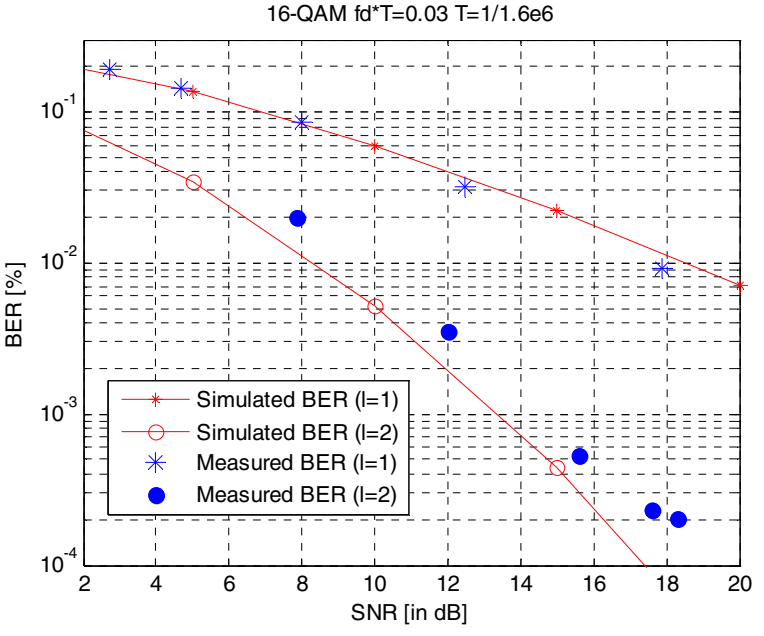


Fig. 10.13 BER comparison of the simulation and measurement for 16-QAM modulated signal over a Rayleigh fading channel.

10.4 Single RF Front-End MIMO Using CDM

In code-division multiplexing (CDM), e.g. [15], the signals of different antennas must be multiplied by orthogonal codes and then added together, where a single RF front-end is used to down-convert the summed RF signals to baseband. In the baseband section, the signals are multiplied by the orthogonal codes, integrated and de-multiplexed. This technique is depicted in Figure 10.14.

In this method, the first signal is multiplied by the $c_1(t)$ code and the second signal by the $c_2(t)$ code in the symbol duration and then added together. These two codes are orthogonal, i.e. $c_1(t)$ is equal to 1 between 0 and T_s, and $c_2(t)$ is equal to 1 between 0 and $T_s/2$ and to -1 between $T_s/2$ and T_s, where T_s is the symbol duration. The signal $\left(h_1 s(t)+n_1(t)\right)c_1(t)+\left(h_2 s(t)+n_2(t)\right)c_2(t)$ is then down-converted using a single RF front-end. The baseband signal is represented as $\left(h_1 \tilde{s}(t)+\tilde{n}_1(t)\right)c_1(t)+\left(h_2 \tilde{s}(t)+\tilde{n}_2(t)\right)c_2(t)$.

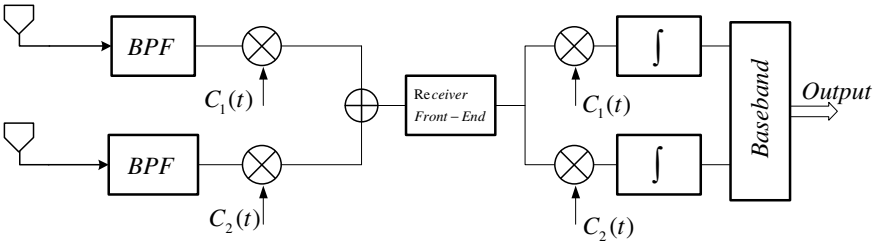


Fig. 10.14 Realization of single-branch multiantenna receiver based on CDM

In the baseband section, the signal is multiplied by $c_1(t)$ and $c_2(t)$ and separated into different paths. In each path, an integrator can remove the effect of the other signal; however, the effect of the noise of both antennas is maintained in each path.

$$\begin{aligned} h_1\tilde{s}+\tilde{n} &= 1/T_s\int_0^{T_s}[(h_1\tilde{s}(t)+\tilde{n}_1(t))c_1(t)c_1(t)+(h_2\tilde{s}(t)+\tilde{n}_2(t))c_2(t)c_1(t)]dt \\ &= 1/T_s\int_0^{T_s}h_1\tilde{s}(t)dt+1/T_s\int_0^{T_s}\tilde{n}_1(t)dt+0+1/T_s\int_0^{T_s}\tilde{n}_2(t)c_1(t)c_2(t)dt \end{aligned} \tag{10.17}$$

This is the output of the integrator in the first path, where $\tilde{s}(t)$ is constant in the symbol duration, T_s, and the codes have unit energy. The first term of the integral is $h_1\tilde{s}$ and the second part of integral is the integration of the first antenna noise. The third term is zero, because the two codes are orthogonal and $\tilde{s}(t)$ is invariable in the symbol duration, T_s.

The fourth term of the integral is not zero, although $c_1(t)$ and $c_2(t)$ are orthogonal, because $\tilde{n}_2(t)$ is stochastic and is not constant in the symbol duration. This term has the same power as the second part. Hence, in each subchannel in baseband, the output noise power using CDM is twice of that of a traditional design. Indeed, the noise level is increased by *10log(N)dB* in each path when *N* antennas are used and the signals are down-converted with CDM method.

The simulation result for BPSK modulation using two receiving antennas is illustrated in Figure 10.15. A distance of 3 dB can be observed between the conventional diversity system and this technique.

An integrated circuit is designed and fabricated using CDM technique to realize a single front-end receiver for two antennas [16]. Figure 10.16(a) illustrates the different sections in the integrated circuit (IC) and Figure 10.16(b) shows the die photo of the receiver in 5 GHz.

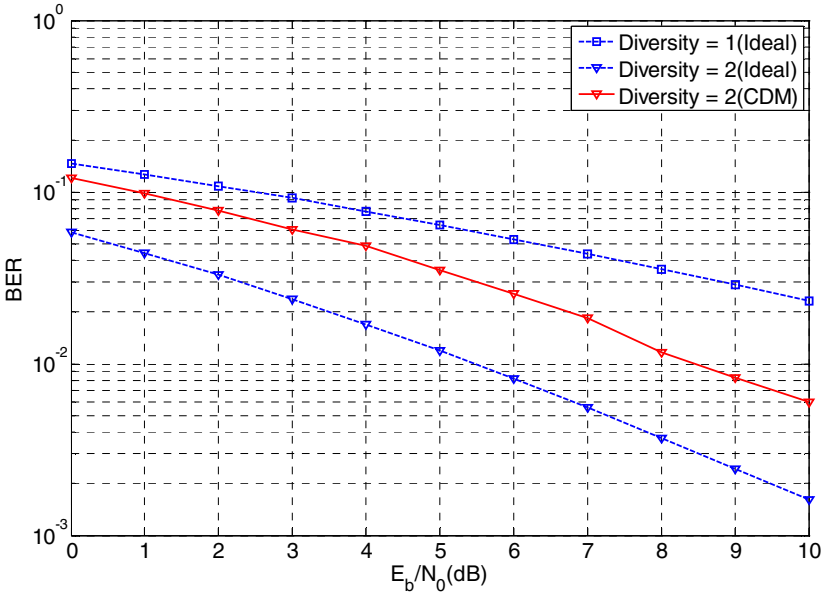
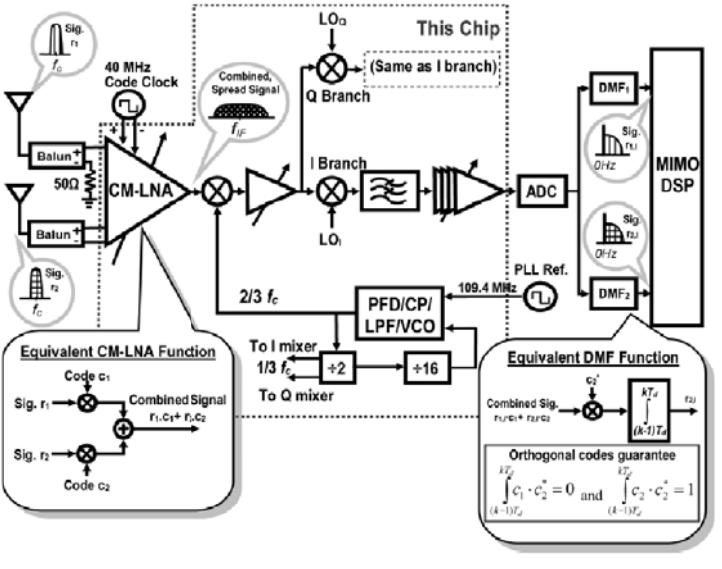
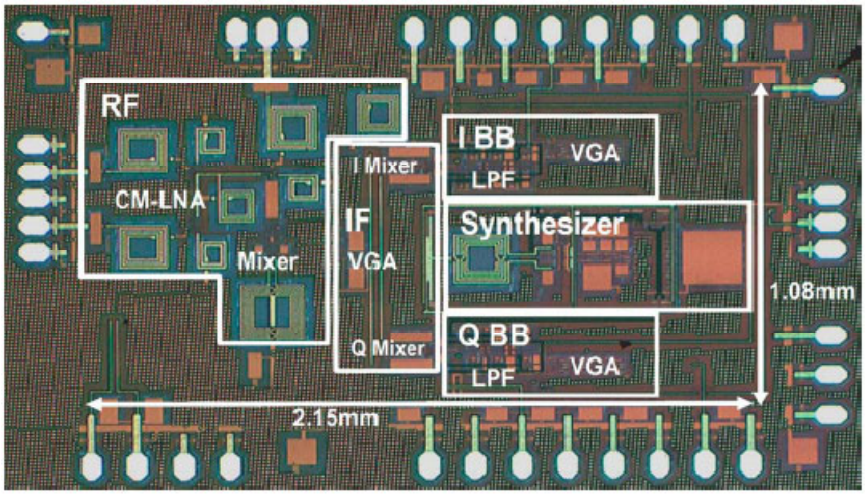


Fig. 10.15 The BER versus E_b/N_0 for single front-end receiver using CDM



(a)

Fig. 10.16 The 5 GHz IC for two-antenna single front-end receiver: a) circuits b) die photo [16]



(b)

Fig. 10.16 *(continued)*

10.5 Single RF Front-End MIMO Using a Parasitic Antenna

The concept of digital beamforming (DBF) antenna arrays is based on the classic phase array method. This technique may be used in different wireless communications systems. According to Figure 10.17, signals received by individual antenna elements are down-converted into baseband signals. These signals are digitized and fed into the digital signal processing (DSP) chip in which the algorithms reside. However, RF circuit branches connected to the array elements, analog-to-digital converters (ADCs) and the baseband DSP chip consume a considerable amount of DC power. Furthermore, each channel connected to the array sensor has the same structure, so the cost of fabrication increases with the number of array elements [22], [23]. These factors limit DBF applications in the wireless communications systems. The parasitic array has been proposed to provide the solutions for the disadvantages of DBF technique [24].

The functional block diagram of a parasitic antenna digital beamformer is shown in Figure 10.18. In this structure, one central element is connected to the single RF front-end, and the surrounding parasitic elements realize the array. Beamsteering is achieved by tuning the load reactances at the parasitic elements surrounding the central active element. The loads may be realized with the use of varactor diodes. The power consumption of this architecture is very small. Furthermore, the system has only one RF front-end, and the small inter-element spacing can be as small as 0.05λ [24].

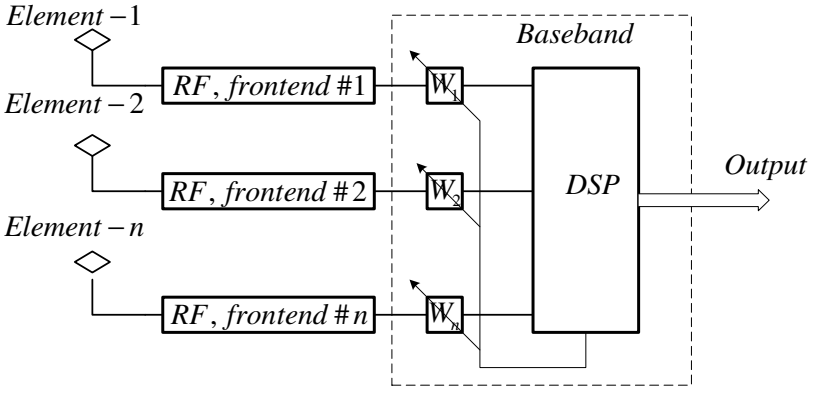


Fig. 10.17 Block diagram of digital beam forming (DBF) antenna array

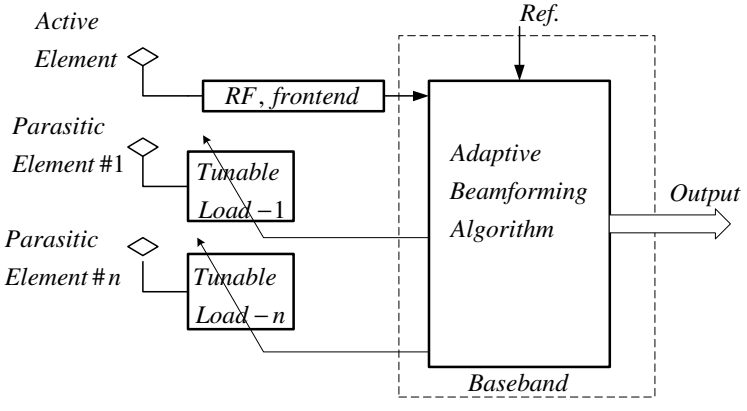


Fig. 10.18 Block diagram of steerable parasitic array radiator (ESPAR) antennas

To analyze the operation of the architecture, we assume that the parasitic array operates in the transmitting mode. The antenna generates a directional beam based on tuning load reactances $(x_1, x_2, \cdots, x_n)$ on the parasitic monopoles. Signals transmitted from the central RF front-end excite the parasitic monopoles with substantial induced mutual currents. Vectors I and V represent the currents and the voltages on the monopoles, respectively:

$$I = (i_0, i_1, \cdots, i_n)^T \tag{10.18}$$

$$V = (v_0, v_1, \cdots, v_n)^T \tag{10.19}$$

where i_0, v_0 represent the current and voltage on the central element, respectively.

One may show the mutual admittances by using a Y matrix, where y_{ij} is the mutual admittance between the i*th* element and the j*th* element. The induced currents are represented with mutual admittances as follows:

$$I = YV = \begin{bmatrix} y_{00} & y_{01} & \cdots & y_{0n} \\ y_{10} & & & y_{1n} \\ \vdots & & & \vdots \\ y_{n0} & y_{n1} & \cdots & y_{nn} \end{bmatrix} \begin{bmatrix} v_0 \\ \vdots \\ v_n \end{bmatrix} \tag{10.20}$$

The parasitic array has a symmetrical structure; therefore:

$$y_{11} = y_{22} = \ldots = y_{nn}$$

$$y_{01} = y_{02} = \ldots = y_{0n}$$

$$y_{12} = y_{23} = \ldots = y_{(n-1)n} = y_{n1}$$

$$y_{13} = y_{24} = \ldots = y_{(n-2)1} = y_{n2}$$

$$y_{1j} = y_{2(j+1)} = \ldots$$

The admittance matrix can thus be determined only by *n* parameters. The voltages on the active central monopole and the m*th* parasitic monopole are obtained as:

$$v_0 = v_s - z_0 i_0 \tag{10.21}$$

$$v_m = -jx_m i_m \tag{10.22}$$

where $m = 1, 2, \ldots, n.$ z_0 and is the characteristic impedance of 50 at the RF port, and v_s represents the transmitted voltage signal source with the amplitude and the phase from the driving RF port at the central element. These equations can be represented in a matrix form as:

$$V = \begin{bmatrix} v_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - XI = v_s U - XI \tag{10.23}$$

where

$$X = \begin{bmatrix} z_0 & 0 & \ldots & 0 \\ 0 & jx_1 & & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \ldots & jx_n \end{bmatrix} \tag{10.24}$$

$$U = [1, 0, \cdots, 0]^T$$

The matrix induced current is then obtained as:

$$I = YV = Y(v_s U - XI) \tag{10.25}$$

Or, similarly, we can have:

$$W \triangleq (Y^{-1} + X)^{-1} U \tag{10.26}$$

This is called an equivalent weight vector [22].

The far-field radiation pattern is the superposition of all monopoles' radiation patterns. Therefore, the far-field current signal in the azimuth direction with its amplitude, and the phase is represented as:

$$y(\theta)_{far-field} = I^T \alpha(\theta) = W^T \alpha(\theta) v_s \tag{10.27}$$

The steering vector, $\alpha(\theta)$, is defined based on the array geometry:

$$\alpha(\theta) = \begin{bmatrix} 1 \\ e^{j\frac{\pi}{2}\cos(\theta)} \\ e^{j\frac{\pi}{2}\cos(\theta - \frac{2\pi}{n})} \\ \\ e^{j\frac{\pi}{2}\cos(\theta - \frac{2(n-1)\pi}{n})} \end{bmatrix} \tag{10.28}$$

According to the reciprocity theory for radiation patterns, if the antenna is working in the receiving mode, the voltage signals, *u(t)*, at the RF port are:

$$u(t) = W^T \alpha(\theta) s(t) \tag{10.29}$$

where $s(t)$ represents the far-field incident current waves with the amplitude and phase in the azimuthal direction (θ), W is dependent on the reactance at each parasitic monopole.

The desired beam patterns can be formed by tuning the reactances $(x_1, x_2, \cdots, x_n)$. This is achieved by changing the control voltages of the diodes, which are connected to the parasitic elements. This relation is the same form as the beamforming equation in the array processing literature [23]. Signals received at the antenna receiver are the sum of the weighted samples at individual antenna sensor elements. This technique is also known as steerable parasitic array radiator (ESPAR) antennas.

The ESPAR technique is used on the transmitter side for implementing MIMO/MISO transmitters for BPSK and QPSK (quadrature phase-shift keying) modulation [25]. Moreover, it has been implemented to realize a single RF front-end MIMO system for MPSK (multiple phase-shift keying) modulation [26].

A simplified schematic diagram of a three-element system is shown in Figure 10.19. As can be seen, an antenna element is connected to the RF front-end, and the other antennas are controlled using two reactances. The idea is to map the spatially multiplexed signals to the spatial basis function where the desired beam pattern in the far field becomes a combination of the individual signals. The reactive loads are used to create linear combination of spatial basis function. The probability of bit error based on E_b/N_o is shown in Figure 10.20 for a three-element antenna using BPSK modulation [27],[28].

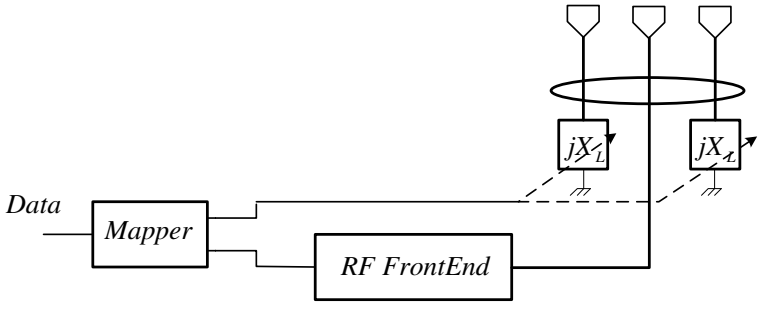


Fig. 10.19 Block diagram of a three-element parasitic antenna MIMO transmitter [26]

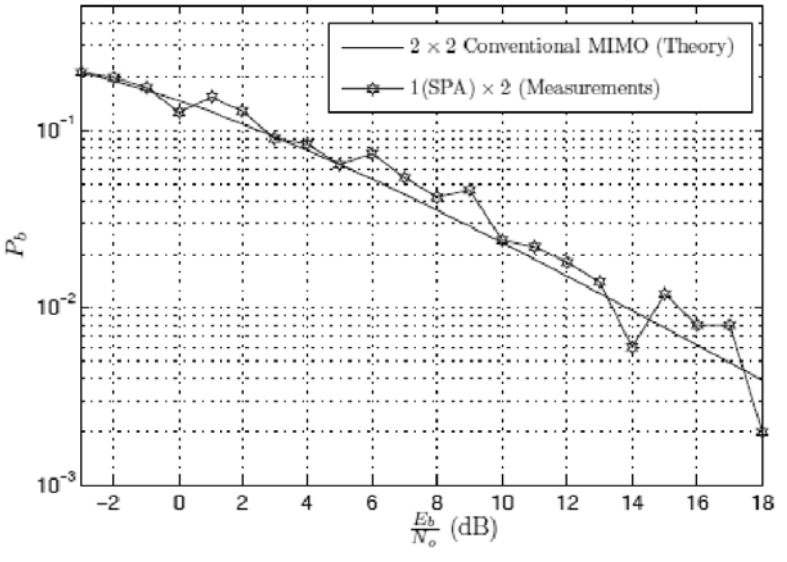


Fig. 10.20 Bit error probability of BPSK parasitic antenna MIMO transmitter [27]

References

[1] Janaswamy, R.: Radiowave Propagation and Smart Antennas for Wireless Communications. Kluwer Academic Publishers (2001)

[2] Oestges, C., Clerckx, B.: MIMO Wireless Communications: From Real World Propagation to Space Time Code Design. Academic Press (2007)

[3] Liu, J., Khaled, N., Petre, F., Bourdoux, A., Barel, A.: Impact and mitigation of multiantenna analog front-end mismatch in transmit maximum ration combining. EURASIP Journal on Applied Signal Processing 2006, 1–14 (2006)

[4] Morris, M.L., Jensen, M.A.: Impact of receive amplifier signal coupling on MIMO system performance. IEEE Transactions on Vehicular Technology 54, 1678–1683 (2005)

[5] Sanayei, S., Nosratinia, A.: Antenna selection in MIMO systems. IEEE Communications Magazine 42(10) (October 2004)

[6] Molisch, A., Win, M., Winters, J.: Reduced-complexity transmit/receive-diversity systems. IEEE Transactions on Signal Processing 51(11), 2729–2738 (2003)

[7] Duman, T.M., Ghrayeb, A.: Coding for MIMO Communication Systems. Wiley (2008)

[8] Gorokhov, A., Gore, D., Paulraj, A.: Receiver Antenna Selection for Spatial Multiplexing: Theory and Algorithem. IEEE Transactions on Signal Processing 51, 2796–2807 (2003)

[9] Gharavi-Alkhansari, M., Greshman, A.: Fast antenna selection in MIMO systems. IEEE Transactions on Signal Processing 52(2), 339–347 (2004)

[10] Gucluoglu, T., Duman, T.M.: Performance of Space-Time Coded Systems with Transmit Antenna Selection. In: Proceedings of the Conference on Information Sciences and Systems (CISS), Baltimore, MD, pp. 863–868 (March 2007)

[11] Greshman, A., Sidiropoulos, N.D.: Space Time Processing for MIMO Communications. John Wiley & Sons (2005)

[12] Proakis, J.G., Salehi, M.: Digital Communications, 5th edn. McGraw-Hill, New York (2008)

[13] Barry, J., Lee, E.A., Messerschmitt, D.G.: Digital Communications, 3rd edn. Kluwer Academic Publishers (2004)

[14] Lari, M., Bassam, A., Mohammadi, A., Ghannouchi, F.M.: Time-Multiplexed Single Front-End MIMO Receivers with Preserved Diversity Gain. IET Communications 5(6), 789–796 (2011)

[15] Jahanian, A., Tzeng, F., Heydari, P.: Code-modulated path-sharing multi-antenna receivers: theory and analysis. IEEE Transactions on Wireless Communications 8(5), 2193–2201 (2009)

[16] Tzeng, F., Jahanian, A., Pi, D., Heydari, P.: A CMOS Code-Modulated Path-Sharing Multi-Antenna Receiver Front-End. IEEE Journal on Solid-State Circuits 44(5), 1321–1335 (2009)

[17] Golden, G.D., Foschini, G.J., Valenzuela, R.A., Wolniansky, P.W.: Detection algorithm and initial laboratory results using the V-BLAST space time communication architecture. Electronics Letters 35(1), 14–16 (1999)

[18] Xiong, F.: Digital Modulation Techniques, 2nd edn. Artech House (2006)

[19] Paulraj, A., Nabar, R., Gore, D.: Introduction to Space-Time Wireless Communications. Cambridge University Press (2003)

[20] Kwan, M., Kok, C.: FIR zero-forcing equalizer for ISI MIMO channels using space-time modulated codes. IEEE Transactions on Wireless Communications 5, 123–132 (2006)
[21] Wang, C., Au, E.K.S., Murch, R.D., Mow, W.H., Cheng, R.S., Lau, V.: On the Performance of the MIMO Zero-Forcing Receiver in the Presence of Channel Estimation Error. IEEE Transactions on Wireless Communications 6, 805–810 (2007)
[22] Sun, C., Hirata, A., Ohira, T., Karmakar, N.C.: Fast Beamforming of Elec-tronically Steerable Parasitic Array Radiator Antennas: Theory and Experiment. IEEE Transactions on Antennas & Propagation 52(7), 1819–1832 (2004)
[23] Johnson, D.H., Dudgeon, D.E.: Array Signal Processing: Concepts and Techniques, pp. 91–365. Prentice-Hall, Englewood Cliffs (1993)
[24] Bains, R., Muller, R.: Using parasitic elements for implementing the rotat-ing antenna for MIMO receivers. IEEE Transactions on Wireless Communications 7(11), 4522–4533 (2008)
[25] Kalis, A., Kanatas, A.G., Papadias, C.B.: A novel approach to MIMO transmission using a single RF front end. IEEE Journal on Selected Areas in Communications 26(6) (August 2008)
[26] Alrabadi, O.N., Papadias, C.B., Kalis, A., Prasad, R.: A Universal Encoding Scheme for MIMO Transmission Using a Single Active Element for PSK Modulation Schemes. IEEE Transactions on Wireless Communications 8(10), 5133–5143 (2009)
[27] Alrabadi, O.N., Divarathne, C., Tragas, P., Kalis, A., Marchetti, N., Kalis, A., Papadias, C.B., Prasad, R.: Spatial Multiplexing with a Single Radio: Proof-of-Concept Experiments in an Indoor Environment with a 2.6-GHz Proto-type. IEEE Communications Letters (2011)
[28] Mohammadi, A., Ghannouchi, F.M.: Single RF Front-End MIMO Transceivers. IEEE Communications Magazine (December 2011)

Index

About the Authors

Abbas Mohammadi is a professor at Electrical Engineering Department of Amirkabir University of Technology where he is head of Radio Communications center of excellence and Director of Microwave/Millimeter-wave and Wireless Communications Research Lab. He has held an invited professor position at the University of Calgary. Prof. Mohammadi authored two books, awarded four patents, and published more than 150 peer-reviewed papers. His research interests are in the areas of microwave and wireless communications, software defined radio, adaptive transceivers and MIMO systems. He is a senior member of Institute of Electrical and Electronics Engineers (IEEE).

Fadhel M. Ghannouchi, a Senior Canada Research Chair in Intelligent RF Radio Technology, iCORE Professor, founder and Director of iRadio Laboratory, University of Calgary is regarded as one of the foremost authorities in the fields of RF Electronics and Radio technology. His current research interests are in their applications in broadband wireless, satellite and biomedical sciences. Other recognitions received by Prof. Ghannouchi include: The Alberta Professional Engineers Geologists and Geophysicist (APEGGA) Alberta Ingenuity Fund Research Excellence Award, The Research Excellence Award of Schulich school of Engineering, University of Calgary and, being elected Distinguished Lecturer of the IEEE Microwave Theory and Technique. For his scientific and technical accomplishments, Dr. Ghannouchi's peers have elected him to Fellow in The Canadian Academy of Engineering, Fellow of The Engineering Institute of Canada, Fellow of The Institution of Engineering and Technology (IET formally IEE), and Fellow of Institute of Electrical and Electronics Engineers (IEEE). Prof. Ghannouchi held several invited positions at several academic and research institutions in Europe, North America and Japan. He has provided consulting services to a number of microwave and wireless communications companies worldwide. His research interests are in the areas of microwave instrumentation and measurements, nonlinear modeling of microwave devices and communications systems, design of power and spectrum efficient microwave amplification systems, design of intelligent RF transceivers for wireless and satellite communications and software defined and cognitive Radio systems. Prof. Ghannouchi has published extensively. He has authored two books, was awarded 7 patents, and published more that 400 peer-reviewed articles, given over 50 invited/keynote/plenary conference presentations, trained over 60 Master and PhD graduates and founded three Spin-off companies.