ANALYSIS OF A RADIO FREQUENCY CLASS D AMPLIFIER ARCHITECTURE WITH BANDPASS SIGMA-DELTA MODULATION

by

Thomas Johnson
B.A.Sc., University of British Columbia, 1987
M.A.Sc., Simon Fraser University, 2001

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in the School
of
Engineering Science

SIMON FRASER UNIVERSITY
Fall 2006

APPROVAL

Name: Thomas Johnson

Degree: Doctor of Philosophy

Title of thesis: Analysis of a Radio Frequency Class D Amplifier Architecture with Bandpass Sigma-Delta Modulation

Examining Committee: Dr. Albert Leung, Chair
Professor of Engineering Science

Dr. Shawn Stapleton, Senior Supervisor
Professor of Engineering Science

Dr. James Cavers, Supervisor
Professor of Engineering Science

Dr. Rodney Vaughan, Supervisor
Professor of Engineering Science

Dr. Daniel Lee, SFU Examiner
Associate Professor of Engineering Science

Dr. Yuanxun Wang, External Examiner
Professor of Electrical Engineering,
University of California, Los Angeles

Date Approved: _____

Abstract

This thesis analyzes an amplifier architecture that combines a RF class D amplifier with a bandpass sigma-delta modulator, broadening the utility of class D amplification to include signals with envelope variation. An integrated design methodology is presented that incorporates the coding efficiency and average pulse transition frequency of the encoded pulse train into classical RF class D amplifier design equations. The equations are used to predict the power efficiency of a complementary voltage switched class D amplifier design with CMOS, pHEMT, and MESFET switches. Simulated results are compared with the analysis and verify the design methodology.

The power efficiency analysis shows a direct link between modulator coding efficiency and the output power of the amplifier; therefore, a modulator with high coding efficiency is desirable. It is shown that coding efficiency depends significantly on the order of the modulator loop filter as well as the carrier oversample ratio employed in the design. The variation with carrier oversample ratio is not monotonic for second and fourth order modulators, and some oversample ratios are more optimal than others.

Bandpass $\Sigma\Delta$ modulation synthesizes a pulse train with synchronous zero-crossings, and the coding efficiency limitations of encoding a binary amplitude pulse train with constrained zero-crossings is analyzed. The analysis and characterization of other encoder designs shows that bandpass $\Sigma\Delta$ modulation is remarkably efficient. The analysis is extended to pulse train upconversion employing Manchester encoding. Upconversion reduces the difficulty of implementing highly selective noise shaping resonators at RF frequencies, and the impact of upconversion in terms of coding efficiency and average transition frequency is shown.

Dedicated to my mother, Gwendolyn Johnson
1929–2005

We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.

— *T. S. Eliot "Little Gidding"*

Acknowledgments

My senior supervisor, Dr. Shawn Stapleton, has been visionary in terms of promoting switching RF amplifier technology for wireless system applications starting in the mid-1990's. He has said that there will be many opportunities in this emerging research area, and I agree. My journey over the last four years has lead to the discovery of many interesting problems, and this thesis is a summary of the insights I have gained. I sincerely thank Shawn for supporting this research, and giving me this opportunity to gain experience in what I am sure will continue to be an active research area.

I would also like to thank my other two supervisors, Dr. James Cavers and Dr. Rodney Vaughan. My meetings with them have been infrequent over the years, but their support has been invaluable and assisted me on my journey.

The Ph.D. is an enormous commitment, and without a doubt I have had unwavering support from my family as I have traveled the ups and downs of research. I have an incredible family, and I have had the loving support of a wife, mother, father, and brother. My two boys have also traveled with me on my journey through graduate school, and I look forward to the new adventures which lie ahead.

Finally, I would like to thank my friends, colleagues in the wireless lab, and funding sponsors for supporting my work. Telus Mobility has been the primary sponsor for this research, and I am grateful for a scholarship from the Canadian Wireless and Telecommunications Association, as well as graduate fellowships, and stipends from Simon Fraser University. The funding has provided a unique opportunity for me, and I expect to continue the research which they have sponsored.

Contents

List of Tables

List of Figures

List of Abbreviations

ADC	Analog-to-digital converter
APP	Average Pulse Period
ATF	Average Transition Frequency
BW	Bandwidth
CDMA	Code Division Multiple Access
CMOS	Complementary Metal Oxide Semiconductor
CSCD	Current Switched Class D
CVSCD	Complementary Voltage Switched Class D
CT	Continuous-time
DAC	Digital-to-analog converter
DC	Direct Current
DFT	Discrete Fourier Transform
DT	Discrete-time
FFT	Fast Fourier transformation
GmC	Transconductance capacitor
HBT	Heterojunction Bipolar Transistor
IV	Current-voltage device characteristics
MESFET	Metal Semiconductor Field Effect Transistor
MSE	Mean Square Error
NGN	Narrowband Gaussian Noise
NRZ	Non-return to zero
NTF	Noise Transfer Function

OSR	Oversample ratio
PA	Power Amplifier
PAR	Peak-to-average power ratio
PDF	Probability Distribution Function
pHEMT	Pseudomorphic High Electron Mobility Transistor
PLR	Power Loss Ratio
PMF	Probability Mass Function
PSD	Power Spectral Density
PWM	Pulse Width Modulation
RBW	Resolution Bandwidth
RF	Radio Frequency
RMS	Root Mean Square
$\Sigma\Delta$M	Sigma-Delta Modulation
SNR	Signal-to-noise ratio
STF	Signal Transfer Function
UCR	Upconversion Ratio
VSCD	Voltage Switched Class D
W-CDMA	Wideband CDMA
WSS	Wide-sense Stationary
3GPP	3rd Generation Partnership Project

List of Symbols

η_a	overall amplifier power efficiency
η_d	drain power efficiency
η_p	coding efficiency
ρ	peak-to-average power ratio
σ_u^2	modulator drive power
$b[k]$	discrete-time binary amplitude modulator pulse train
$b_c(t)$	continuous-time binary level *impulse* train
$D_f(s)$	feedback DAC transfer function
$D_o(s)$	output DAC transfer function
$e[k]$	discrete-time quantization error signal
f_{av}	average transition frequency
f_c	carrier frequency of source signal
f_b	bandwidth of source signal
f_o	carrier frequency after upconversion
f_s	sample frequency (or clock rate)
$g[k]$	discrete-time reconstruction filter impulse response
$G(s)$	continuous-time reconstruction filter transfer function
$H(s)$	continuous-time modulator loop filter transfer function
$H(z)$	discrete-time modulator loop filter transfer function
$H_i(z)$	input path discrete-time filter transfer function
$H_f(z)$	feedback path discrete-time filter transfer function
$i_A(t)$	current signal into a VSCD amplifier reconstruction filter

$n[k]$	noise component in pulse train $b[k]$
$p(t)$	continuous-time modulator output pulse train
$p_a(t)$	polar class D amplifier pulse train with levels of $\pm\Delta_a$ (no DC offset)
$p_{dr}(t)$	driver output binary amplitude pulse train
P_{DC}	DC supply power
P_L	reconstructed signal (load) power
P_n	reconstructed signal noise power
R_c	carrier oversample ratio: $f_s/(2f_c)$
R_e	envelope oversample ratio: $f_s/(2f_b)$
R_L	load resistance in amplifier
Q	resonator quality factor
$s[k]$	signal component of pulse train $b[k]$
$\mathrm{sinc}(x)$	$\sin(\pi x)/(\pi x)$
SNR_t	signal-to-noise ratio threshold
$S_p(f)$	power spectral density of $p(t)$
T	sample period
T_{av}	average pulse period
$u(t)$	source signal
$v[k]$	discrete-time quantizer input signal
$v_A(t)$	class D amplifier pulse train with DC offset
$v_L(t)$	load signal
$y[k]$	discrete-time binary amplitude modulator pulse train
$y_c(t)$	continuous-time binary level *impulse* train
$y(t)$	continuous-time modulator pulse train before upconversion

Chapter 1

Introduction

The first description of the RF class D amplifier circuit was given by Baxandall [1] in 1959, and still remains as one of the few amplifier topologies with an ideal DC to RF conversion efficiency of 100%. The high power efficiency is obtained by operating the active devices as switches. The switching action in the amplifier generates a binary amplitude output pulse train preserving the zero-crossings of the input signal only. Hence, the primary limitation of the class D amplifier topology is the elimination of the signal envelope after amplification, which limits the utility of the amplifier to constant envelope signals.

The class of source signals which can be amplified by a class D circuit topology is broadened to include time varying envelope signals, providing the amplifier is preceded by a source encoder. The source encoder maps the time varying envelope input signal to a binary amplitude pulse train for subsequent amplification in the class D stage. The switches in the class D amplifier circuit replicate and amplify the encoded pulse train, and a time varying envelope load signal is reconstructed by an output bandpass filter inherent in the class D circuit topology. Therefore, the source encoder for a class D amplifier circuit is constrained by signal reconstruction in the amplifier. Bandpass sigma-delta ($\Sigma\Delta$) modulation is an example of a compatible source encoder for RF class D amplification which is consistent with signal reconstruction in the class D amplifier.

The focus of this thesis is the analysis of the augmented configuration of a bandpass $\Sigma\Delta$ modulator and RF class D amplifier. By preceding the RF class D amplifier with a bandpass $\Sigma\Delta$ modulator, the utility of the class D amplifier is broadened and opens up many new applications. The efficient amplification of wireless signals such as W-CDMA is challenging, and the augmented class D amplifier architecture has the potential to boost power efficiency to levels well beyond conventional class AB designs.

Although the augmented bandpass $\Sigma\Delta$ modulator and RF class D amplifier architecture has been proposed by other authors [2], the published research has been primarily conceptual.[1] The motivation for this research project is to extend the conceptual ideas into a concrete theoretical framework identifying the key parameters which control power efficiency in the bandpass $\Sigma\Delta$ modulator and RF class D amplifier architecture. In this introductory chapter, an overview of the amplifier is given, and important terms are defined. The overview includes a summary of related work, and the chapter concludes with a discussion of research goals and contributions.

1.1 Amplifier Architecture Overview

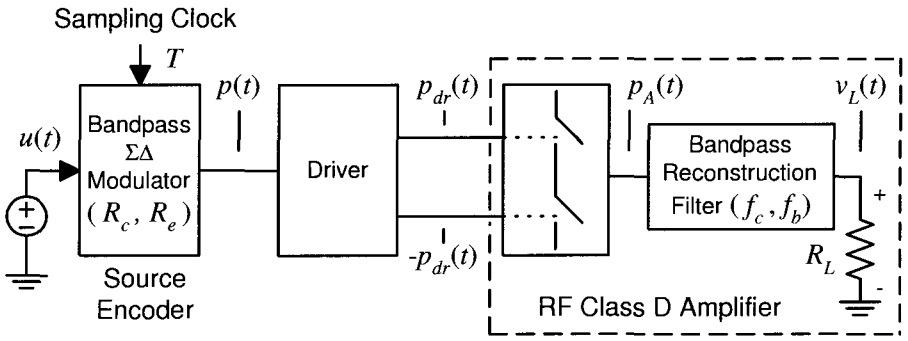
A simplified model of the bandpass $\Sigma\Delta$ modulator and RF class D amplifier architecture is given in Fig. 1.1. The amplifier has three main parts: 1) a source encoder to map the modulated source signal $u(t)$ into a binary amplitude pulse train $p(t)$, 2) a driver which level shifts and amplifies the modulator pulse train to drive the class D switches, and 3) a RF class D power amplifier (PA) stage which generates a load signal $v_L(t)$.

The input signal $u(t)$ to the amplifier, also called the source signal, is a modulated RF signal with carrier frequency f_c and bandwidth f_b. The amplifier output signal is $v_L(t)$, and the overall input-output response of the amplifier is ideally equivalent to a linear amplifier with gain G_a. The amplifier introduces distortion, and the distortion is measured by comparing the signal-to-noise ratio (SNR) of the source and load signals. The amplifier also introduces signal delay τ_a as shown in Fig. 1.1(b). In the following sections, the encoder, power amplifier, and driver stages are discussed in more detail.

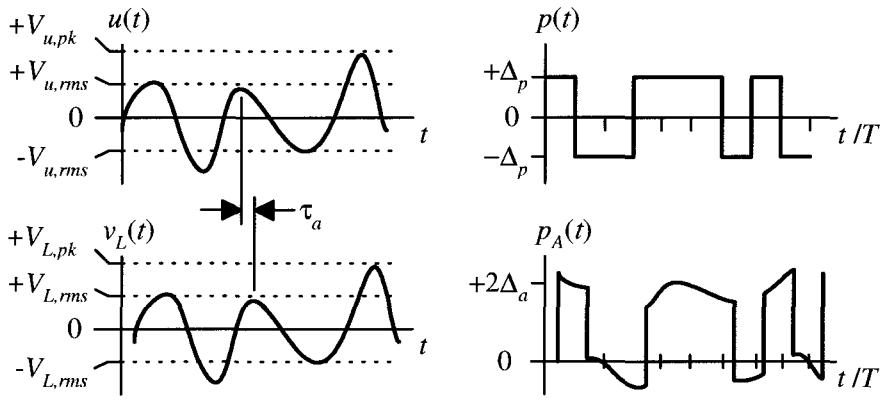
1.1.1 Source Encoder

The function of the source encoder block is to convert a continuous-time (CT) source signal $u(t)$ into a CT pulse train $p(t)$. The subsequent amplification of $p(t)$, and the method of signal reconstruction in the RF class D amplifier, impose two constraints on the class of potential source encoders for the augmented amplifier. First, the switches in the class D amplifier limit pulse train amplitude states to two levels. Consequently, a source signal

[1] The citation uses the term class S. The distinction between class S and class D is not universal, and the author prefers the term class D based on the recommendation of F. Raab, co-author of [3], a frequently cited reference on switching amplifier theory.



(a) Amplifier architecture.



(b) Amplifier signals: 1) source signal $u(t)$, 2) modulator pulse train $p(t)$, amplifier pulse train $p_A(t)$, and 4) load signal $v_L(t)$.

Figure 1.1: RF class D amplifier with a bandpass $\Sigma\Delta$ modulator.

with a time varying envelope must first be quantized prior to amplification. The second encoder constraint is imposed by the implementation of the decoder, which in this case is an analog bandpass filter. Linear signal reconstruction by a bandpass filter significantly limits potential encoding methods, and yet the filter is an integral part of the class D amplifier and is required to realize high power efficiency.

Three possible source coding methods compatible with the class D amplifier constraints are: 1) pulse width modulation (PWM), 2) click modulation [4] (also known as zero-position coding [5,6]), and 3) bandpass $\Sigma\Delta$ modulation. In low frequency applications such as audio amplifiers, all three different methods have been applied [7–9]. However, when the extension of these methods to RF frequencies is considered, the implementation of PWM and click modulation becomes more difficult. PWM is intrinsically nonlinear and usually requires a high oversample ratio to reduce in-band distortion products, while click modulation requires significant digital signal processing [6]. On the other hand, lowpass $\Sigma\Delta$ modulation used in audio frequency amplifiers can also be implemented in a bandpass configuration, and experimental implementations have been realized with sample rates of several GHz [10–12]. The implementation considerations make bandpass $\Sigma\Delta$ modulation attractive for RF amplifier applications; thus, bandpass $\Sigma\Delta$ modulation is the principle focus of this work.

As may be expected, the quantization and signal reconstruction constraints have system costs. One of the primary goals in this research work has been to develop a methodology to evaluate these system costs, determine how efficient bandpass $\Sigma\Delta$ modulation is as a source encoder, and explore how to improve source encoding.

1.1.2 RF Class D Amplifier

The RF class D amplifier is composed of two switches which close alternately depending on the state of the drive signals $\pm p_{dr}(t)$. The switching action generates a large amplitude binary level pulse train $p_A(t)$, and a load signal $v_L(t)$ is reconstructed from $p_A(t)$ by a bandpass filter. Examples of the switching signal and load signal are shown in Fig. 1.1(b). RF class D amplifiers circuits are subdivided into two main circuit topologies depending on whether the switched pulse train $p_A(t)$ is a current or voltage signal; the corresponding circuit topologies are called current switched class D (CSCD) and voltage switched class D (VSCD).

The out-of-band response of the reconstruction filter, $g(t)$, is important, and affects the efficiency of the RF class D amplifier. For example, assume $p_A(t)$ is a voltage switched

signal, and assume the reconstruction filter is a series resonator. The resonator permits the circulation of load current at in-band frequencies, while preventing out-of-band current flow through the load and switches. Although the spectrum of the voltage signal is broadband, no power is dissipated in the switches at out-of-band frequencies because out-of-band current is significantly attenuated by the resonator. The simultaneous co-existence of a voltage signal with a broadband frequency spectrum, and a narrowband current signal with only in-band spectral components are fundamental to the high efficiency operation of the VSCD circuit. A similar condition exists in the CSCD amplifier except current and voltage are interchanged with respect to a VSCD amplifier. A simple resonator is usually sufficient for periodic drive signals, since the spectrum is discrete with widely spaced harmonics, but for a broadband power spectrum such as bandpass $\Sigma\Delta$ modulation, a higher order filter is required.

In principle, the class D amplifier circuit amplifies any binary amplitude pulse train $p_{dr}(t)$, and there are no restrictions on the timing of switch closures. Since the timing of switch closures is unrestricted, periodic and aperiodic pulse trains are amplified equally well. The characteristic that the class D amplifier amplifies any binary amplitude pulse train is attractive, and distinguishes it from class E which is also a switch-mode amplifier topology that has restrictions on the timing of switch closures.

Although the class D amplifier circuit can amplify both periodic and aperiodic binary amplitude pulse trains, the type of drive signal affects the implementation of the class D circuit. Periodic drive signals simplify circuit implementation, and most RF amplifier implementations reported in the literature have been designed and tested with periodic drive signals [13–15]. If the input pulse train is periodic with a 50% duty cycle, the circuit implementation advantages include zero voltage switching in a CSCD, zero current switching in a VSCD amplifier, low order reconstruction filters where a second order resonator usually suffices, and orthogonal harmonic components that are compatible with frequency periodic transmission line structures. When the pulse train is aperiodic, like bandpass $\Sigma\Delta$ modulation, the phase of the load current is random with respect to switch closures, the pulse train power spectrum is continuous and broadband, and the reconstruction filter has tight attenuation requirements. The term 'conventional' class D circuit shall sometimes be used in reference to circuits designed for periodic square wave pulse trains, and these circuits are distinguished from more general class D circuits which amplify both periodic and aperiodic signals.

1.1.3 Driver

The driver stage shown in Fig. 1.1(a) is the interface stage between the source encoder and the RF class D amplifier stage. The purpose of the driver stage is to level shift and amplify the modulator pulse train $p(t)$ to provide sufficient voltage swing and current to control the switches in the class D circuit. Depending on the circuit topology of the class D circuit, the driver may have a single output $p_{dr}(t)$, or two opposite phase output signals $+p_{dr}(t)$ and $-p_{dr}(t)$. A single-ended driver output is used when the RF class D amplifier has a transformer coupled input to generate switch control signals [3,16], or when complementary (n and p type) devices are employed as switches [3]. Otherwise, two independent drive signals from the driver are required for controlling the two switches in the amplifier.

The implementation of a driver is a difficult design problem, and frequently an ideal driver is assumed when reporting results [2, 17–19]. Driver issues include edge timing to prevent significant shoot-through current when both switches are transitioning, generating sufficient drive current to have short rise and fall times with a large capacitive load, and tracking floating node voltages in the class D circuit. Although the implementation of the driver is not the primary focus of this work, a limited analysis of driver issues is given for a CMOS amplifier design in Chapter 3.

1.2 Definitions

There are several new terms that have been introduced in this work to support the analysis of the integrated bandpass $\Sigma\Delta$ modulator and RF class D amplifier architecture. The new terms are the distinction between envelope and carrier oversample ratio in a bandpass $\Sigma\Delta$ modulator, the coding efficiency of the encoded pulse train, and the average transition frequency of the pulse train. Definitions are also given for drain efficiency and overall power efficiency.

1.2.1 Drain Efficiency

Drain efficiency (η_d) is a common measure of power efficiency in RF power amplifiers. Drain efficiency is defined as

$$\eta_d = P_L / P_{DC,pa} \tag{1.1}$$

where P_L is the average load power and $P_{DC,pa}$ is the DC supply power provided to the class D stage. The total DC power in a VSCD stage is approximately equal to the sum of three terms: 1) load power P_L, 2) ohmic conduction losses in the switches, P_c, and 3) capacitive switching losses, P_Q. Therefore, in a VSCD amplifier

$$\eta_d \approx \frac{P_L}{P_L + P_c + P_Q}. \tag{1.2}$$

1.2.2 Overall Power Efficiency

Overall power efficiency is more encompassing than drain efficiency and includes the driver power as well as the modulator power. Let η_a be the overall amplifier power efficiency, then

$$\eta_a = \frac{P_L}{P_{DC,\Sigma\Delta} + P_{DC,dr} + P_{DC,pa}}. \tag{1.3}$$

where two additional DC power terms are added, $P_{DC,\Sigma\Delta}$ and $P_{DC,dr}$, to include the DC power consumption of the modulator and driver stages, respectively. Specific hardware implementations of the modulator are beyond the scope of this work, and overall efficiency is constrained to a consideration of driver and class D stage power only.

1.2.3 Envelope and Carrier Oversample Ratios

When $\Sigma\Delta$ modulation is used in an analog-to-digital converter (ADC), oversampling is employed to synthesize an encoded output signal with high SNR over a narrow output bandwidth. Oversampling distributes the noise power from a low resolution quantizer over a wide bandwidth, and the modulator has a noise shaping loop to create a noise well where the source signal is encoded with high SNR. The term 'noise', in the context of a bandpass $\Sigma\Delta$ modulator, refers to the quantization error signal generated by a low resolution quantizer. A binary quantizer is assumed throughout this work. An example of the encoded power spectrum at the output of a bandpass $\Sigma\Delta$ modulator is shown in Fig. 1.2. The quantization noise is spread uniformly over a wide bandwidth depending on the sample rate in the modulator, and a signal is encoded in a narrow notch shaped in the noise spectrum.

There are many factors which affect the SNR of the encoded source signal. One of the most important variables is the envelope oversample ratio defined as

$$R_e = f_s/(2f_b). \tag{1.4}$$

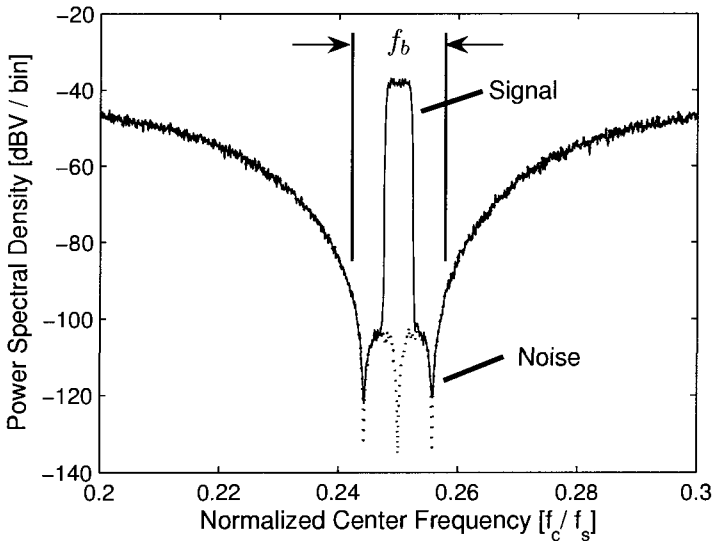


Figure 1.2: Power spectral density of a sixth order bandpass ΣΔ modulator pulse train; the spectrum of the quantization noise is superimposed on the pulse train power.

The term 'envelope oversample ratio' is usually called 'oversample ratio' by most authors, and is applicable to lowpass ΣΔ modulators where there is only one oversample ratio. However, in bandpass ΣΔ modulators, there is a second oversample ratio called the carrier oversample ratio defined as:

$$R_c = f_s/(2f_c). \tag{1.5}$$

Both the envelope and carrier oversample ratios are defined such that an oversample ratio of unity corresponds to Nyquist sampling.

The terms 'envelope' and 'carrier' are adopted to distinguish between the two oversample ratios, reflecting the communication system application. Other authors prefer not to explicitly identify the carrier oversample ratio, and usually describe their designs in terms of the ratio f_s/f_c [20]. However, the distinction of the carrier oversample ratio as a design variable is important in this work, and a contribution of this research is to show how the carrier oversample ratio affects the power efficiency of the class D amplifier. As a design parameter, the carrier oversample ratio impacts both the coding efficiency and average transition frequency of the encoded pulse train $p(t)$.

1.2.4 Coding Efficiency

Although out-of-band spectral components do not dissipate any power in an ideal class D circuit, a high level of out-of-band spectral power is wasteful and reduces the utility of the pulse train to generate load power. Significant out-of-band power is generated by one bit quantization, and coding efficiency is introduced as a figure of merit to evaluate encoder performance.

Coding efficiency is defined as the ratio of the reconstructed load power relative to the total pulse train power [19], and a coding efficiency measurement is illustrated in Fig. 1.3. A polar pulse train $p_a(t)$ with amplitude levels $\pm\Delta_a$ is filtered with a reconstruction filter, and the load power P_L is measured across a load R_L. The total pulse train power is measured without the filter and is Δ_a^2/R_L for a voltage pulse train and $\Delta_p^2 R_L$ for a current pulse train. With these definitions, coding efficiency (η_p) is expressed as

$$\eta_p = \begin{cases} P_L R_L/\Delta_a^2 & \text{switched voltage} \\ P_L/(\Delta_a^2 R_L) & \text{switched current} \end{cases} . \tag{1.6}$$

A simple rearrangement of equation (1.6) leads to

$$V_{L,rms} = \sqrt{\eta_p}\,\Delta_a \qquad \text{switched voltage} \tag{1.7}$$

$$I_{L,rms} = \sqrt{\eta_p}\,\Delta_a \qquad \text{switched current} \tag{1.8}$$

which shows that coding efficiency can also be used as an effective gain to relate the root mean square (RMS) amplitude of the load signal, $V_{L,rms}$ or $I_{L,rms}$, to the amplitude of the switched pulse train in the circuit.

Two additional remarks are made about coding efficiency. First, coding efficiency is always less than one for bandlimited bandpass source and load signals. In other words, the source signal cannot be encoded into a binary amplitude pulse train without generating quantization noise, as in the case of bandpass $\Sigma\Delta$ modulation, or distortion, as in the case of PWM. Coding efficiency is unity in the unquantized case only, and any out-of-band noise or distortion generated by the encoder means that the coding efficiency of the pulse train is necessarily less than one. The second remark is that coding efficiency is distinct from drain efficiency in the class D power amplifier. The coding efficiency of the pulse train depends on the encoder, while drain efficiency depends on the circuit design in the amplifier. Hence, the coding efficiency of the encoder could be low, while amplifier power efficiency is high.

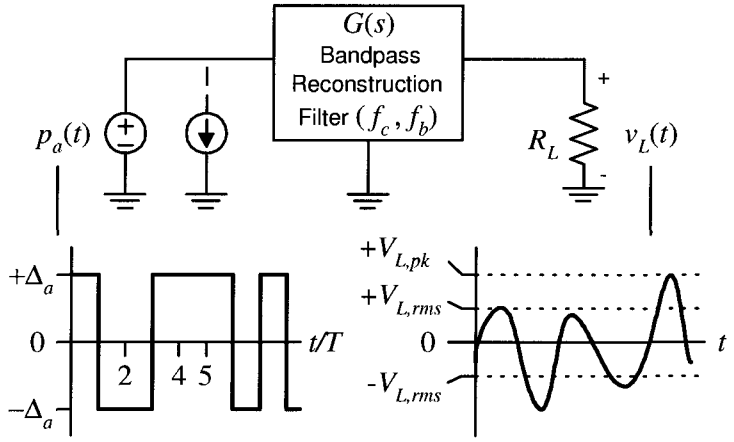


Figure 1.3: Coding efficiency measurement.

The definition of coding efficiency in (1.6) assumes the reconstructed signal power is measured with some reconstruction filter $G(s)$, and the definition stems from the amplifier application. The definition indirectly includes assumptions about the reconstruction filter bandwidth and center frequency f_c. However, another method to estimate coding efficiency is from a direct measurement of the power spectral density (PSD) of the modulator pulse train $p(t)$.

Let the one-sided power spectral density of $p(t)$ be $S_p(f)$. Coding efficiency is then expressed as

$$\eta_p = \frac{1}{\Delta_p^2} \int_{f_c - f_b/2}^{f_c + f_b/2} S_p(f) df \tag{1.9}$$

where $p(t)$ is a polar signal with amplitude levels $\pm\Delta_p$, and f_b is the equivalent noise bandwidth of the reconstruction filter. The PSD method is also frequently used to estimate the reconstructed SNR for $\Sigma\Delta$ modulators [21] assuming an ideal bandpass reconstruction filter.

If $p(t)$ is zero mean, which it usually is assuming a bandpass source signal, then the total pulse train power can also be written as the variance of $p(t)$ [22]. Therefore coding efficiency is unaffected by DC level shifts in the amplifier circuit and is strictly a function of the encoding method

1.2.5 Average Transition Frequency

One of the most significant power loss mechanisms in RF class D power amplifiers is the energy dissipated when a switch changes state [3, 7, 23]. The stored energy in parasitic capacitance and inductance changes during a switch transition, and the current to effect the change in stored energy dissipates power in the switches. The frequency of switch transitions therefore has a significant impact on the overall power efficiency of the amplifier architecture.

The average transition frequency (ATF) of the modulator pulse train $p(t)$ is defined as

$$f_{av} = 1/T_{av} \tag{1.10}$$

where T_{av} is the average pulse period (APP) in $p(t)$. A pulse period is defined as the time between two consecutive amplitude changes in $p(t)$. Since $p(t)$ is polar, the ATF is the same as the average zero-crossing rate of the pulse train. The definition of ATF in (1.10) is similar to other definitions in the literature except that a pulse width is sometimes defined as the time between two rising edges [7].

The bandpass $\Sigma\Delta$ modulator pulse train generates a synchronous pulse train, and the pulse widths in $p(t)$ are integer multiples of the clock period T. Let k be a positive integer, then

$$T_{av} = T \sum_{k=1}^{K} k \Pr\{kT\} \tag{1.11}$$

where $\Pr\{kT\}$ is the probability of a pulse of width kT in $p(t)$. Examples of the pulse width probability mass function are given in Chapter 2 for the bandpass $\Sigma\Delta$ modulator, and in Chapter 3, the ATF is incorporated in design equations to predict amplifier power efficiency.

1.3 Motivation

The interest in high efficiency amplifier solutions for wireless systems is significant both in terms of infrastructure and handsets applications. A reduction in power consumption provides many advantages including cost, size, and battery lifetime. These factors make the potential benefit of the bandpass $\Sigma\Delta$ modulator and class D RF amplifier worthy of further consideration.

The research areas associated with bandpass $\Sigma\Delta$ modulation and RF class D amplifiers are relatively separate disciplines, and a primary goal of this research project is to bridge

these research areas into a cohesive system analysis of the composite amplifier system. As a potential high efficiency amplifier solution, implementation issues are paramount, and an important question is: how well can we expect the bandpass $\Sigma\Delta$ modulator and RF class D amplifier configuration to work at RF frequencies? The question of overall system performance is general, and a number of specific questions are posed:

1. Given the bandwidth of the bandpass $\Sigma\Delta$ modulator signal, what RF class D circuit topology should be used?
2. What design methodology should be used to determine the size (gate width) of the switches in the amplifier?
3. What technology parameters are required to make first order predictions about the power efficiency of the RF class amplifier stage?
4. What power efficiency performance is expected using typical monolithic technologies such as GaAs and CMOS?
5. How should the driver be implemented, and what is the overall power efficiency including driver power losses?
6. What is the coding efficiency of bandpass $\Sigma\Delta$ modulation?
7. What parameters affect the coding efficiency of bandpass $\Sigma\Delta$ modulation?
8. How can the coding efficiency of the modulator be increased?
9. What is the average transition frequency of a bandpass $\Sigma\Delta$ modulator pulse train?
10. What parameters affect the average transition frequency of a modulator pulse train?
11. What architecture changes can be made to reduce the implementation challenges in the modulator?

The methodologies applied to investigate these problems are analysis and simulation. As the following literature summary shows, there are relatively few papers which have focused on an integrated analysis of the bandpass $\Sigma\Delta$ modulator and RF class D amplifier architecture. By exploring the design issues through analysis and simulation, important groundwork is laid for evaluating the potential power efficiency of the design at RF frequencies.

1.4 Literature Review

The RF class D amplifier and bandpass $\Sigma\Delta$ modulator architecture has evolved from the merging of three distinct bodies of research. The research areas are: 1) conventional RF class D amplifiers with square wave switching signals; 2) bandpass $\Sigma\Delta$ modulation, and 3) audio class D power amplifiers with lowpass $\Sigma\Delta$ modulation. A brief summary of relevant contributions in these research areas is given next.

1.4.1 RF Class D Amplifiers

The RF class D amplifier topology dates back over 40 years and was initially developed for highly efficient sinusoidal sources in high frequency (HF) transmitters [1] and HF heating applications [24]. Throughout the 1970's and 1980's a key contributor to RF class D development was Raab [25, 26] who implemented a number of different HF class D designs. As ways to overcome the envelope limitations of class D amplifiers, he proposed a RF PWM scheme in 1977 [25]. Most of his circuits either used the complementary VSCD configuration or transformer coupled VSCD circuits [3, 26]. In recent years, the frequency limits of RF class D amplifiers has been pushed up to several GHz based on CSCD topologies with transmission line output structures [13–15]. However, the experimental work with CSCD amplifiers has been constrained to periodic switching. Aperiodic switching in a CMOS VSCD amplifier has been demonstrated recently by Wagh et. al. [27].

1.4.2 Audio Class D Power Amplifiers

Audio applications of the class D circuit topology with lowpass $\Sigma\Delta$ modulation have been around for at least ten years [7], when $\Sigma\Delta$ modulation was introduced as an alternate to PWM. PWM modulation intrinsically generates harmonic distortion components, and the noise shaping characteristics of $\Sigma\Delta$ modulation offered attractive solutions to mitigate distortion mechanisms. Magrath and Sandler [7] wrote a paper on the analysis of lowpass $\Sigma\Delta$ modulation and class D amplification. In their paper, they show simulated results for the average transition frequency of a lowpass modulator, and describe a way to reduce the ATF of the pulse train through pulse grouping.

Research into improved binary amplitude encoders for class D amplification has been limited to audio applications. In 2002, Kato [28] presented a paper at an Audio Engineering Society convention, and reported on a modified sigma-delta architecture which replaced the

binary quantizer with a trellis search using the Viterbi algorithm. A few other papers have reported on similar architectures for lowpass $\Sigma\Delta$ modulators [29,30], and in this work some of these ideas are applied to boosting the coding efficiency of bandpass $\Sigma\Delta$ modulation. Although these papers report on improved encoding performance, it appears that a bound on the coding efficiency of binary amplitude pulse trains with constrained zero-crossings is still an open research problem.

1.4.3 Bandpass $\Sigma\Delta$ Modulation

Since the early 1990's, the first implementations of bandpass $\Sigma\Delta$ modulators started to appear with Thurston et al. [31] showing a continuous-time bandpass $\Sigma\Delta$ modulator at 2.5 MHz. Modulator designs have continued to improve and designs with clock rates as high as 3 GHz have been reported [12]. The circuit implementation considerations of bandpass $\Sigma\Delta$ modulators are beyond the scope of this project with the exception that it is recognized that the implementation of integrated high Q resonators on a monolithic silicon substrate is very difficult at RF frequencies [32]. Additional references on the design methodology of bandpass $\Sigma\Delta$ modulators are given in Chapter 2.

1.4.4 RF Class D Amplifiers With Bandpass $\Sigma\Delta$ Modulation

In 1989, Schreier and Snelgrove [33] made a reference to potential applications of bandpass $\Sigma\Delta$ modulation for amplifier applications. Stapleton [34] wrote an amplifier product note for Agilent in 1996 which included a bandpass $\Sigma\Delta$ modulator and class D RF amplifier, and in 1997, Jayaraman et al. [2] published the first simulated results for a bandpass $\Sigma\Delta$ modulator and RF class D amplifier. As part of his Ph.D. work, Jayaraman implemented a fourth order continuous-time bandpass $\Sigma\Delta$ modulator, and the paper summarizes simulation results shown in the last chapter of his thesis for a VSCD circuit with HBT switches. The primary contribution of his work was to show by simulation the feasibility of the amplifier architecture a radio frequencies.

Other system design papers include Ketola et al. [18] and Dupuy and Wang [35]. Ketola et al. propose using images from a bandpass $\Sigma\Delta$ modulator output with a binary amplitude digital-to-analog converter (DAC) to drive a class D RF amplifier. The advantage is a reduction in modulator sample rate, but the disadvantage of this scheme is a significant reduction in coding efficiency, because the carrier oversample rate is less than one. Dupuy and

Wang have defined a different system architecture called envelope sigma-delta modulation. In their scheme, a lowpass $\Sigma\Delta$ modulator encodes the envelope into a binary amplitude pulse train, and the pulse train directly modulates a RF carrier which is then amplified by a class E amplifier. The resultant pulse train after modulation is not a binary amplitude pulse train, as in the case of bandpass $\Sigma\Delta$ modulation, and an alternate way of upconverting a binary amplitude pulse which preserves binary amplitude levels is proposed in Chapter 6.

In summary, switching RF power amplifiers for wireless applications is an emerging research area with relatively few publications, and many research problems still remained unsolved.

1.5 Contributions

An important theme in this research project is linking the encoder and RF class D amplifier through the concepts of coding efficiency and average transition frequency. The synthesis of a binary amplitude pulse train with high coding efficiency and low average transition frequency is essential to optimize the utility of the RF class D amplifier. The goal of this research project has been to evaluate the coding efficiency and average transition frequency of bandpass $\Sigma\Delta$ modulation, and incorporate these concepts in the analysis of the power efficiency of the RF class D amplifier.

The main contributions of this work are:

1. RF class D circuit topologies are reviewed and the complementary voltage switched class D amplifier circuit is considered to be the best topology.

2. The concept of coding efficiency is introduced as a link between the modulator and the reconstructed load power at the amplifier output.

3. Conventional RF class D design equations which assume a periodic square wave drive signal are generalized to include aperiodic drive signals. The design equations are generalized by incorporating coding efficiency, average transition frequency, and peak-to-average power ratio terms.

4. A design methodology is presented that estimates the gate width of the switches for the RF class D amplifier given basic device parameters.

5. The drain efficiency of the complementary VSCD circuit is analyzed for pHEMT, MESFET, and CMOS devices. In the case of CMOS, the overall power efficiency analysis includes a driver.

6. Coding efficiency and average transition frequency characteristics are shown for six different bandpass modulator designs. The coding efficiency of second and fourth order modulators is significantly higher than sixth order designs; therefore, low order modulators are recommended for RF class D amplifiers.

7. Extensive simulation results demonstrate that coding efficiency does not vary monotonically with carrier oversample ratio for second and fourth order modulators. For sinusoidal and narrowband gaussian source signals, it is shown that some carrier oversample ratios are more optimal than others. For example, although $R_c = 2$ has been used in bandpass $\Sigma\Delta$ modulator implementations [11, 20, 31, 36, 37], a better carrier oversample ratio is $R_c = 1.7$, both in terms of coding efficiency and average transition frequency.

8. The coding efficiency limitations of a binary amplitude pulse train with constrained zero-crossings at integer multiples of a clock are analyzed for periodic load signals. The model gives insight into how coding efficiency can vary with carrier oversample ratio. It is also proven that the maximum amplitude bound of an undistorted sinusoidal load signal is $4/\pi$ for pulse trains with constrained zero-crossings and unity amplitude levels.

9. A numerical search algorithm is used to explore the coding efficiency limits of binary amplitude pulse train signals with constrained zero-crossings. The results of the numerical search demonstrate that the coding efficiency of bandpass $\Sigma\Delta$ modulation is remarkably good.

10. A tree search, instead of a binary quantizer, is added in the bandpass $\Sigma\Delta$ modulator loop to boost coding efficiency. Simulation results show that the enhanced modulator can improve the coding efficiency of a sixth order modulator.

11. The coding efficiency and ATF of Manchester encoding is analyzed as a method of upconverting a low frequency bandpass $\Sigma\Delta$ modulator pulse train.

1.6 Supporting Publications

Some of this research work has been previously published in conference [19, 22, 38, 39] proceedings and journals [40, 41]. These papers are listed below.

1. T. Johnson and S. Stapleton, "Coding Efficiency of Sigma-Delta Modulation for Class D Amplification in W-CDMA Applications," in *Wireless 2004 Proceedings*, Calgary, Alberta, July 12–14 2004, 288–295.

2. T. Johnson and S. Stapleton, "Available load power in a RF class D amplifier with a sigma-delta modulator driver," in *IEEE Radio and Wireless Conference*, Atlanta, Georgia, Sept. 19–22 2004, pp. 439–442.

3. T. Johnson, S. Stapleton, and J. Cavers, "Binary coding for RF class D amplifier applications," in *The Ninth Canadian Workshop on Information Theory*, Montréal, Québec, June 5–8 2005, pp. 74–77.

4. T. Johnson, R. Sobot, and S Stapleton, "Measurement of bandpass sigma-delta modulator coding efficiency and pulse transition frequency for RF class D power amplifier applications," in *Canadian Conference of Computer and Electrical Engineering*," Ottawa, Ontario, May 7–10 2006, pp. 2279–2282.

5. T. Johnson, R. Sobot, and S. Stapleton, "Manchester encoded bandpass sigma-delta modulation for RF class D amplifiers," accepted for publication in *IEE Proc.—Circuits Devices Syst.*, August 2006.

6. T. Johnson and S. Stapleton, "RF class D amplification with bandpass sigma-delta modulator drive signals," accepted for publication in *IEEE Trans. Circuits Syst. I*, August 2006.

Also, two other journal submissions are currently under review:

1. T. Johnson and S. Stapleton, "Comparison of Bandpass $\Sigma\Delta$ Modulator Coding Efficiency with a Periodic Signal Model," submitted to *IEEE Trans. Circuits Syst. I*, Jan. 2006.

2. T. Johnson, R. Sobot, S. Stapleton, "CMOS RF Class D Power Amplifier with a Bandpass Sigma-Delta Modulator," submitted to Microelectronics Journal, Aug. 2006.

1.7 Organization of Results

In this chapter, an overview of the bandpass $\Sigma\Delta$ modulator and RF class D amplifier has been given. Subsequent chapters focus on analyzing and simulating the coding efficiency, average transition frequency, and power efficiency of the amplifier configuration. The remaining chapters are organized as follows:

- Chapter 2 focuses exclusively on evaluating the coding efficiency and average transition frequency of bandpass $\Sigma\Delta$ modulation. Six different modulator designs are shown, and simulation results are presented which show how coding efficiency and average transition frequency vary with carrier oversample ratio. The principle conclusion is that for second and fourth order modulators there are some carrier oversample ratios which are more optimal than others.

- In Chapter 3, the focus is shifted to the RF class D amplifier and driver stage. Classical design equations for class D amplifiers are modified to include coding efficiency and average transition frequency parameters. The equations show how the modulator and amplifier are linked, and the power efficiency of a VSCD design is evaluated with pHEMT, MESFET, and CMOS switches.

- The importance of modulator coding efficiency, and its relationship to load power at the output of the amplifier, motivates questions about the limits of coding efficiency for a binary amplitude pulse train with constrained zero-crossings. In Chapter 4, a sinusoidal load signal is assumed, and periodic binary amplitude pulse trains are systematically evaluated to draw conclusions about the limits of coding efficiency for sinusoidal source signals. At large input amplitudes, the periodic model mimics the behavior of the modulator, and both the model and modulator show variation in coding efficiency with carrier oversample. Conclusions about selecting an 'optimal' carrier oversample ratio are given.

- Chapter 5 continues the exploration of modulator coding efficiency, and alternate encoding structures are considered to improve coding efficiency. The alternate structures replace the binary quantizer with a tree search to look-ahead and select the best sequence which minimizes the mean square error of the reconstructed signal. The most significant improvement with tree look-ahead appears to be a boost in the coding efficiency of high order modulators, and an example is given for a sixth order design.

Otherwise, the coding efficiency of second and fourth order modulators appears to be remarkably good.

- Chapter 6 analyzes Manchester encoding as a means of upconverting a low frequency pulse train to a high frequency pulse train for driving a RF class D amplifier. The analysis is motivated by the difficulty of implementing highly selective resonators at RF frequencies. The coding efficiency loss and impact on average transition frequency are analyzed and conclusions are made on the best upconversion sequence.
- Chapter 7 presents conclusions and recommendations for future research topics.

Chapter 2

Bandpass Sigma-Delta Modulation

The purpose of this chapter is to identify what parameters affect coding efficiency and average transition frequency in a bandpass $\Sigma\Delta$ modulator. In a modulator design, coding efficiency and reconstructed signal to noise ratio (SNR) are trade-offs. Therefore, coding efficiency cannot be analyzed in isolation, and coding efficiency has meaning when it is coupled with a minimum acceptable SNR threshold. Factors which affect coding efficiency include the peak-to-average power ratio of the source signal, the carrier oversample ratio, and the order of the noise shaping filter. SNR also depends on the type of source signal and filter order, as well as the envelope oversample ratio. The average transition frequency depends primarily on the carrier oversample ratio and to a lesser degree on the type of source signal.

We begin the chapter by first presenting a high level model of a bandpass $\Sigma\Delta$ modulator. After the model, a summary of the transfer functions for six different discrete and continuous-time modulator noise shaping filters is given. The six designs include second, fourth, and sixth order filters which are used as examples for demonstrating coding efficiency and average transition frequency characteristics. A discussion of SNR measurement methods for sinusoidal and wideband source signals follows, and the chapter concludes with simulated results of the modulator designs.

The simulated results show how SNR varies with coding efficiency as a function of modulator order, source signal peak-to-average power ratio, and with carrier oversample ratio. Similar results are shown for the average transition frequency. An important conclusion in this chapter is the observation that coding efficiency does not vary monotonically with carrier oversample ratio, and certain carrier oversample ratios should be avoided. For example, a carrier oversample ratio of 2.0 has been used in the literature [11, 20, 31, 36, 37], and this leads to inferior coding efficiency relative to a carrier oversample ratio of 1.7.

2.1 Bandpass ΣΔ Modulator Models

ΣΔ modulators are broadly classified into discrete-time (DT) and continuous-time (CT) designs depending on the implementation of the noise shaping filter. At low frequencies, DT designs are common and often implemented with switched capacitor techniques [20]. The switched capacitor designs are difficult to implement at RF frequencies, and instead, DT designs are converted to CT designs [32]. In a CT design, the noise shaping filters are most often implemented with passive high Q resonators [31] or transconductors [36] which can be integrated into a monolithic design.

2.1.1 Continuous-Time Model

A block diagram of a CT bandpass ΣΔM is shown in Fig. 2.1(a). The modulator consists of a high Q resonator $H(s)$, followed by an ideal impulse sampler, and a binary level quantizer in a loop with a feedback DAC $D_f(s)$. The loop has high gain over a very narrow signal bandwidth f_b centered at f_c, and very low gain outside the signal band. The action of the loop is to shape the quantization noise generated by the low resolution quantizer away from the signal band. When the ΣΔ modulator is used in an ADC configuration, the impulse train $b_c(t)$ is decimated and filtered to yield a high resolution, low bandwidth, digital signal. In the amplifier application, the impulse train $b_c(t)$ is converted to a CT pulse train by an output DAC $D_o(s)$, and the bandpass ΣΔ modulator is configured as an analog-to-analog converter or encoder. The encoder converts the CT source signal $u(t)$ to a pulse train $p(t)$ with quantized amplitude levels and quantized zero-crossings.

The model in Fig. 2.1(a) is conceptual, and the sampling, quantization, feedback DAC, and output DAC are shown as different blocks. In an implementation, several of these blocks are frequently combined. For example, all four blocks can be combined into a single block with a latching comparator where the feedback DAC and output DAC are both NRZ. Often there are implementation advantages to have different feedback and output DAC's, and several feedback DAC's may be employed to synthesize the required transfer function. For instance, Shoaei and Snelgrove [42] use two feedback DAC's, a return-to-zero and half return-to-zero DAC, to increase the degrees of freedom for implementing different loop filter designs.

The details of how the resonator functions $H(s)$ are implemented, and how to choose the feedback DAC's in a design, has been covered by other authors [20, 36, 42–44]. Instead,

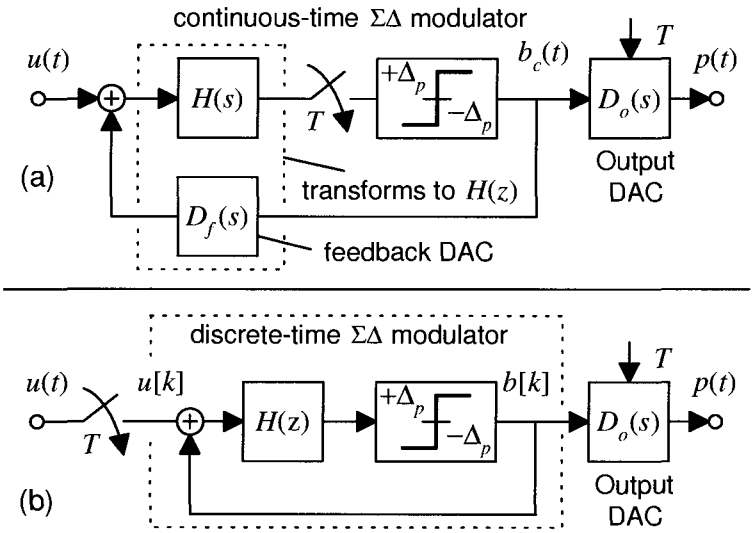


Figure 2.1: ΣΔ modulator models: (a) continuous-time, and (b) discrete-time.

we revert to the underlying discrete-time equivalent models that are frequently used as a prototype for designing CT transfer functions. The synthesis and analysis of ΣΔ modulators is easier in the discrete-time domain, and the most common methodology applied to CT designs is to transform a DT prototype to an equivalent CT design using the pulse transform, a Laplace transform which assumes a specific DAC pulse shape [45]. A few examples of simulations with equivalent CT modulator designs are shown throughout this work, but for the most part, conclusions are derived from DT equivalent models.

2.1.2 Discrete-Time Model

Fig. 2.1(b) shows an equivalent DT model for the CT modulator in Fig. 2.1(a). The primary difference in the DT model is the absence of the feedback DAC, and the sampler now precedes the modulator loop. A strictly bandlimited source signal is assumed, and the input sampler converts the CT input signal $u(t)$ to a DT signal $u[k]$ at a sample rate of f_s, the same as the sample rate in the CT model. The DT model has two blocks: a DT resonator $H(z)$ with at least one sample of delay, and a binary quantizer. The loop works in the same way as the CT modulator with high gain in the signal band to shape the quantization noise. The binary amplitude DT pulse train $b[k]$ is converted to a CT pulse train $p(t)$ by a NRZ DAC with transfer function $D_o(s)$, which is identical to the transfer function of the output DAC

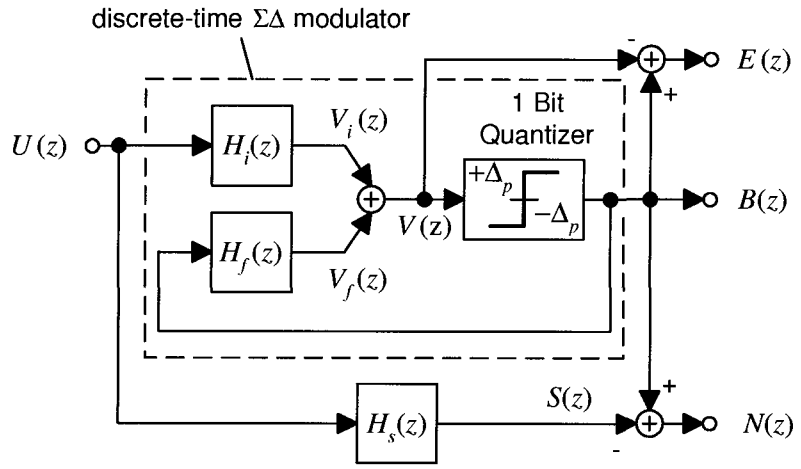


Figure 2.2: Discrete-time ΣΔ modulator signals.

in the CT model.

A more general representation of a DT modulator is shown in Fig. 2.2 where the input signal transfer function $H_i(z)$ and feedback path transfer function $H_f(z)$ are explicitly shown. The two transfer functions are usually implemented as a two input, single output, transfer function; in a CT design, the equivalent input transfer function $H_i(s)$ is often designed as an anti-alias filter. Expressions for the signal and noise shaping loops are derived next.

2.1.3 Signal and Noise Transfer Functions

Define a quantization error signal $E(z)$ as the difference between the quantizer output $B(z)$ and quantizer input signal $V(z)$:

$$E(z) = B(z) - V(z). \tag{2.1}$$

The output signal $B(z)$, shown in Fig. 2.2, is then defined in terms of the input signal $U(z)$ and the quantization noise $E(z)$ as

$$B(z) = U(z)\frac{H_i(z)}{1 - H_f(z)} + E(z)\frac{1}{1 - H_f(z)}. \tag{2.2}$$

Let $H_n(z)$ be defined as the noise transfer function (NTF)

$$H_n(z) = \frac{1}{1 - H_f(z)}, \tag{2.3}$$

and $H_s(z)$ as the signal transfer function (STF)

$$H_s(z) = \frac{H_i(z)}{1 - H_f(z)}, \tag{2.4}$$

then

$$B(z) = \underbrace{U(z)H_s(z)}_{S(z)} + \underbrace{E(z)H_n(z)}_{N(z)}. \tag{2.5}$$

Equation (2.5) shows that the output signal $B(z)$ is composed of a signal component $S(z)$ and a noise component $N(z)$ where:

$$S(z) = U(z)H_s(z) \tag{2.6}$$

$$N(z) = E(z)H_n(z) \tag{2.7}$$

$$B(z) = S(z) + N(z). \tag{2.8}$$

Although equation (2.5) appears linear, the quantization noise $e[k]$ depends on the quantizer input signal $v[k]$ as shown in (2.1). The quantizer input signal $v[k]$ in turn depends on the input signal $u[k]$; therefore, the quantization error signal $e[k]$ is dependent on the input signal. Often the quantization noise is assumed to be independent of the input signal, which leads to a linear approximation for the modulator operation [21]. As shown later, this approximation is good for moderate level input signals, but at high input levels the noise component is correlated with the signal component.

2.2 Bandpass ΣΔ Modulator Designs

A number of different modulator designs are benchmarked for evaluating SNR and coding efficiency performance. For most of this work, a tunable fourth order DT bandpass modulator design called modulator A is used. Several other designs are also used for comparative testing, and are included to verify whether observations about modulator A are consistent over different designs. These designs are described next and summarized in Table 2.1.

2.2.1 Discrete-time Second Order Bandpass Modulator

A DT bandpass ΣΔ modulator design with an adjustable carrier oversample ratio is designed by transforming a first order lowpass modulator to a bandpass design. This modulator is called design B.

Table 2.1: Bandpass ΣΔ Modulator Designs

Name	Order	Description
A	4	discrete-time; coincident NTF zeros; tunable R_c
B	2	discrete-time; coincident NTF zeros; tunable R_c
C	4	discrete-time; spread NTF zeros; specific designs for $R_c = 1.7$ and 2.0 in Table 2.2
D	4	continuous-time; coincident NTF zeros; tunable R_c
E	2	continuous-time; coincident NTF zeros; tunable R_c
F	6	discrete-time; spread NTF zeros; specific designs for $R_c = 1.7$ and 2.0 in Table 2.2

The first order lowpass modulator [21] has transfer functions

$$H_i(z) = \frac{1}{z-1} \qquad \text{and} \qquad H_f(z) = \frac{-1}{z-1}. \tag{2.9}$$

In this case, $H_f(z) = -H_i(z)$. The lowpass to bandpass transformation [21, Eqn. 9.4] is

$$z \rightarrow \frac{-z(z+\alpha)}{\alpha z + 1} \tag{2.10}$$

where

$$\alpha = -\cos(\pi/R_c). \tag{2.11}$$

The parameter α determines the position of the noise well relative to the sample frequency. With the transformation, the bandpass modulator is

$$H_i(z) = -H_f(z) = \frac{-(\alpha z + 1)}{z^2 + 2\alpha z + 1}. \tag{2.12}$$

$$\tag{2.13}$$

Using the expressions $H_i(z)$ and $H_n(z)$, the signal transfer function $H_s(z)$ and the noise transfer function $H_n(z)$ are:

$$H_s(z) = \frac{-(\alpha z + 1)}{z(z+\alpha)} \tag{2.14}$$

$$H_n(z) = \frac{(z - e^{j\omega_c})(z - e^{-j\omega_c})}{z(z - \cos\omega_c)}. \tag{2.15}$$

The signal transfer function gain $|H(e^{j\omega}|$ is unity for all values of ω and α, while the noise transfer function has single zeros at $\pm\omega_c$ for creating a noise well.

2.2.2 Discrete-time Fourth Order Bandpass Modulator

A similar lowpass to bandpass transformation of the second order double loop modulator [21] yields the fourth order tunable bandpass $\Sigma\Delta$ modulator with:

$$H_i(z) = \frac{-z(z+\alpha)(\alpha z+1)}{(z^2+2\alpha z+1)^2} \tag{2.16}$$

$$H_f(z) = \frac{(\alpha z+1)(2z^2+3\alpha z+1)}{(z^2+2\alpha z+1)^2}. \tag{2.17}$$

$$H_s(z) = \frac{\alpha z+1}{-z(z+\alpha)} \tag{2.18}$$

$$H_n(z) = \frac{(z-e^{j\omega_c})^2(z-e^{-j\omega_c})^2}{z^2(z-\sqrt{\cos\omega_c})(z+\sqrt{\cos\omega_c})}. \tag{2.19}$$

The NTF has two double order complex zeros at $\omega_c = 2\pi f_c/f_s$, and the STF has unity gain, independent of α or R_c. The maximum out-of-band NTF gain for this modulator is 4 and is independent of R_c. This modulator is called design A.

A value of $R_c = 2$ is a classic bandpass $\Sigma\Delta$ modulator configuration [11, 20, 31, 36, 37], corresponding to a modulator with a noise well at $f_s/4$. When $R_c = 2$, the lowpass to transformation (2.10) is simple: $z \rightarrow -z^2$. Despite the simplicity of the transformation, we will show later that a carrier oversample ratio of two has low coding efficiency, and other carrier oversample ratios are better for class D amplifier applications.

2.2.3 Discrete-time Bandpass Modulators — Spread Zeros

Although the coincident zero designs for the modulator NTF described so far have compact analytic forms, they are not optimal in terms of SNR [46]. By spreading the zeros, the SNR can be improved by several dB for a signal bandwidth of f_b. The design of modulator transfer functions with spread zeros is usually done with numerical optimization algorithms [47] assuming bandwidth and NTF out-of-band gain constraints. For this work, a method proposed by Schreier [46] was used to design the modulators in Table 2.1 with spread zero NTF's.

The method starts with Butterworth zeros and poles, and moves the zeros to optimal positions assuming the quantization noise is independent, white, and uniformly distributed. The Butterworth poles are then adjusted to control the maximum NTF out-of-band gain. The maximum out-of-band gain affects the stability of the feedback loop in the modulator. As a guideline, the maximum NTF gain is usually less than 2, a rule of thumb first established

Table 2.2: DT Bandpass $\Sigma\Delta$ Modulators With NTF Spread Zeros ($R_e = 128$)

Order	R_c	NTF/STF Poles	STF Zeros	NTF Zeros	$\max\{\|H_n(z)\|\}$
4	1.7	Butterworth Poles for $BW = 14f_b$	Butterworth Zeros for $BW = 14f_b$	$-0.280471 \pm j0.959863$ $-0.266841 \pm j0.963740$	1.72
4	2.0	Butterworth Poles for $BW = 16f_b$	Butterworth Zeros for $BW = 16f_b$	$\pm 0.007085 \pm j0.999975$	1.76
6	1.7	Butterworth Poles for $BW = 9.5f_b$	Butterworth Zeros for $BW = 9.5f_b$	$-0.282793 \pm j0.959181$ $-0.273663 \pm j0.961826$ $-0.264508 \pm j0.964384$	1.69
6	2.0	Butterworth Poles for $BW = 9.5f_b$	Butterworth Zeros for $BW = 9.5f_b$	$\pm 0.009506 \pm j0.999955$ $\pm j1.0$	1.6

by Lee [48]. A maximum gain of approximately 1.7 is selected as a compromise between stability and SNR, and two modulator designs are summarized in Table 2.2. The table shows the pole/zero locations for a fourth order modulator design called modulator C, and a sixth order modulator called modulator F. Also two variants of each design are shown corresponding to carrier oversample ratios of 1.7 and 2.0.

2.2.4 Continuous-time Bandpass Modulators

Continuous-time modulator designs are frequently designed by transforming a discrete-time prototype into an equivalent continuous-time design [36, 49]. The transformation methods are well established and synthesize equivalent DT and CT filter impulse responses. An example of this design procedure is given in Appendix A where the second order DT modulator B is transformed to an equivalent second order CT modulator, called modulator E.

In general, the transformation can be extended to higher order filters, although closed form expressions can become unwieldy. Some special cases of modulator A, the DT fourth order design, have been transformed to equivalent CT designs. Shoaei [36, 42] and Sobot [44] both have designs which match the DT modulator prototype for $R_c = 2$. At other carrier oversample ratios, the CT designs of Shoaei and Sobot do not transform directly to modulator A. The second order modulator (design E) and Sobot's fourth order modulator [44, Eqn. 29], which shall be called modulator D, are used for comparative testing.

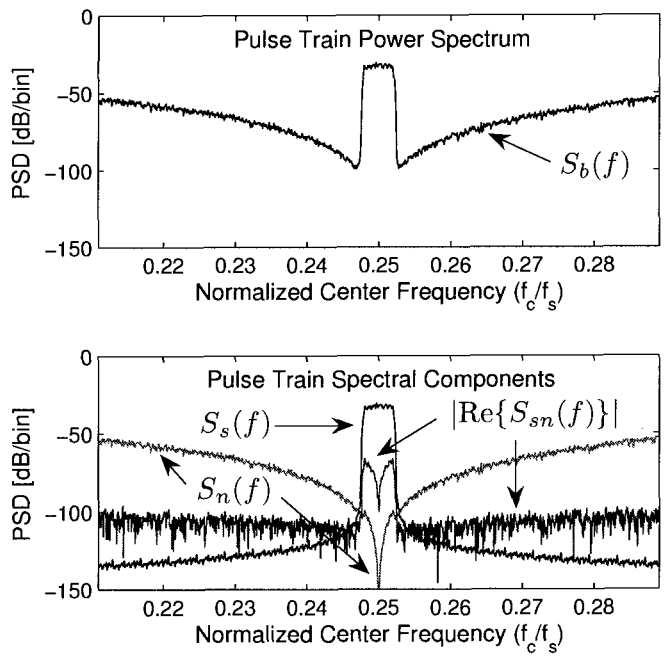


Figure 2.3: Spectral components of a discrete-time pulse train $b[k]$ for modulator A (see section 2.2.2) with a W-CDMA source signal; the source power (σ_u^2) is 0.034.

2.3 Modulator Pulse Train Power Spectral Density

The coding efficiency and SNR of bandpass $\Sigma\Delta$ modulation are derived from the power spectral density (PSD) of the output pulse train. An example power spectrum of a discrete pulse train $b[k]$ is shown in Fig. 2.3, and a number of important results are obtained by decomposing the spectrum into signal and noise components.

2.3.1 Signal and Noise Power Spectrums

Let $S_b(f)$ be the one-sided PSD of $b[k]$ where the substitutions $z = e^{j\omega}$ and $\omega = 2\pi f/f_s$ are used to express the discrete spectral density in terms of the frequency variable f. Then, with (2.8), the PSD of the pulse train is expanded into the following components:

$$S_b(f) = S_s(f) + S_n(f) + 2\mathrm{Re}\{S_{sn}(f)\} \quad 0 \leq f \leq f_s/2. \tag{2.20}$$

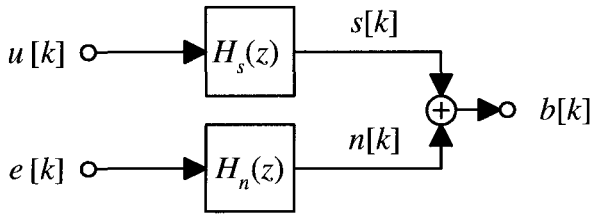


Figure 2.4: Signal and noise components in ΣΔ modulator pulse train.

The term $S_n(f)$ is the PSD of the noise component in the pulse train $b[k]$, $S_s(f)$ is the PSD of the signal component, and $\mathrm{Re}\{S_{sn}(f)\}$ is the real part of the cross power spectral density of the signal and noise components. Since the pulse train is polar with amplitude levels of $\pm\Delta_p$, the total pulse train power is Δ_p^2, and we have the constraint:

$$\int_0^{f_s/2} S_b(f)\,df = \Delta_p^2. \tag{2.21}$$

In other words, the power of the signal and noise components in the pulse train must always equal a constant.

The power balance between noise and signal components gives insight into the modulator. With reference to Fig. 2.4, the noise and signal components are expanded into spectral densities for the source signal $S_u(f)$ and the quantization error signal $S_e(f)$. The output pulse train $b[k]$ is the sum of a signal component $s[k]$ generated by the signal transfer function, and a noise component $n[k]$ generated by the noise transfer function. Therefore,

$$\begin{aligned}\Delta_p^2 = &\int_0^{f_s/2} S_u(f)|H_s(f)|^2\,df + \int_0^{f_s/2} S_e(f)|H_n(f)|^2\,df \\ &+2\int_0^{f_s/2} \mathrm{Re}\{\,S_{ue}(f)\,H_s^*(f)H_n(f)\}\,df.\end{aligned} \tag{2.22}$$

The signal transfer function $H_s(f)$ usually has constant gain A_s and linear phase over the signal bandwidth, and the first term on the right hand side of (2.22) is equal to $A_s^2\sigma_u^2$ where σ_u^2 is the power of the source signal.

The power balance equation is simplified further by assuming the quantization error $e[k]$ is white, uniformly distributed over $\pm\Delta_p$, and independent of the source signal $u[k]$. The noise assumptions are commonly employed in linear modulator models [21,33]. With these

assumptions, the quantization error PSD is

$$S_e(f) = \frac{\Delta_p^2}{3} \times \frac{1}{f_s/2}. \tag{2.23}$$

Since $e[k]$ is independent and zero mean, the cross power term in (2.22) is zero, and the power balance equation simplifies to:

$$\Delta_p^2 = A^2\sigma_u^2 + \int_0^{f_s/2} S_e(f)|H_n(f)|^2\,df. \tag{2.24}$$

If the signal power is small such that $A^2\sigma_u^2 \ll \Delta_p^2$, then

$$\Delta_p^2 \approx \int_0^{f_s/2} \frac{2\Delta_p^2}{3f_s}|H_n(f)|^2\,df. \tag{2.25}$$

Equation (2.25) is equivalently written as a constraint on the NTF gain characteristic:

$$\frac{2}{f_s}\int_0^{f_s/2} |H_n(f)|^2\,df = 3. \tag{2.26}$$

If the NTF is constant for all frequencies outside the signal bandwidth f_b, and zero in the signal bandwidth, then (2.26) leads to the conclusion that the out-of-band NTF gain is approximately $\sqrt{3} \approx 1.7$ assuming a high envelope oversample ratio. This conclusion is consistent with empirical design methodology that usually constrains the maximum NTF gain to be less than 2 [46,48], and provides another guideline for NTF gain characteristics.

When the large signal behavior of the modulator is analyzed, the independence assumption is no longer valid, and the quantization error signal becomes correlated with the input signal. Large signals generate pulse trains with high coding efficiency; therefore, large signal modulator characteristics are important in the amplifier application.

As with small signal behavior, the power balance equation gives insight into large signal behavior. Since the pulse train power is constant, the correlation between $e[k]$ and $u[k]$ increases as the source level increases, and the cross power term eventually limits the reconstructed signal power. Fig. 2.3 shows the noise, signal, and cross-power spectral components of the pulse train power spectrum for a modulator with a W-CDMA signal. The source level is typical for an amplifier application, and the cross-correlation between the noise and signal is evident.

2.3.2 Power Spectral Density Measurements

For SNR measurements it is important to separately identify the signal and noise components in the pulse train power spectrum. A common SNR measurement procedure for sinusoidal source signals is to capture a time series of the output pulse train and estimate the PSD with periodogram techniques. The following excerpt from a book by Norsworthy, Schreier and Temes summarizes the procedure [21, pg. 290]:

> Output noise can be estimated by taking a Hann-weighted discrete Fourier transform (DFT) of the output sequence and summing the power in the in-band "bins," excluding those containing the input tone. The input should be chosen at a frequency centered in a bin but not at the exact band center because in that case symmetries occasionally cause anomalously good SNR readings.

The above SNR measurement procedure shall be referred to as the 'bin delete' method. The main advantage of the bin delete method is that the output pulse train does not need to be decomposed into signal and noise components, providing the DFT resolution over the modulator bandwidth f_b is large. A limitation of the bin delete method is that it cannot be used effectively to estimate the SNR of wideband signals if the noise is not flat over the signal bandwidth. If results are published for wideband signals, they usually resort to a worst case measurement. The noise level adjacent to the signal spectrum is measured, and then assumed to be uniform across the signal bandwidth.

As an alternative to the bin delete method, the noise component of the pulse train is extracted from the pulse train $b[k]$ by generating $s[k]$ in a simulation. A second modulator loop is configured without a quantizer, and generates the signal component of the output signal; with reference to (2.5), the quantization error signal is set to zero, and the loop without the quantizer yields $B(z) \equiv S(z)$. By simultaneously generating $b[k]$ and $s[k]$ during a simulation, the noise component $n[k] = b[k] - s[k]$ can be extracted during post processing. The method is especially useful for calculating the SNR of wideband signals, and was used to generate the power spectra shown in Fig. 2.3 for a W-CDMA source signal. Although the method of separating signal and noise is simple, it has not been reported before, and some authors like Ardalan [50] were unable to verify analytic models for wideband gaussian source signals.

For continuous-time modulators, the output signal component $s(t)$ is not easily synthesized in a simulation because of mixed continuous and discrete time signals, and the bin

delete method is simpler in this case. Other post simulation methods can be applied to continuous-time pulse trains to separate signal and noise by using an adaptive algorithm to cancel the signal component, but the method is not as elegant as the discrete-time method with real time signal cancellation.

Whether the bin delete or signal and noise separation method is used, all power spectral density plots shown in this work are estimated using a modified periodogram technique. The periodograms are constructed with Hann windowed overlapping blocks of 2^{14} samples with 7/8 overlap.

2.4 Signal-to-Noise Ratio

For the purposes of benchmarking modulator performance, it is very common in the literature to assume an ideal bandpass reconstruction filter, and measure SNR directly from the power spectrum of the output pulse train $b[k]$. The advantage of this method is that it focuses exclusively on evaluating modulator performance providing a common reconstruction filter for comparing the SNR of different modulator designs. The ideal bandpass filter is not necessarily the optimum reconstruction filter, because the noise floor varies over the signal bandwidth. However, as will be shown, the ideal bandpass filter is nearly optimum when the SNR is high, and a realizable filter approximating the ideal bandpass filter is the 'best' reconstruction filter for the RF amplifier system. SNR also depends on the peak-to-average power ratio of the source signal, and examples of SNR responses are given in the following sections for sinusoidal, W-CDMA, and narrowband gaussian sources.

2.4.1 SNR With An Ideal Bandpass Reconstruction Filter

A complete discrete-time equivalent system model shown in Fig. 2.5 is used for SNR analysis. The discrete-time model is justified even when the sequence $b[k]$ is converted to a continuous-time pulse train $p(t)$ with an output DAC $D_o(s)$ as shown earlier in Fig. 2.1(b). Since the envelope oversample ratio is large, the NRZ DAC attenuation is approximately constant over a narrow signal bandwidth f_b, and the SNR of the discrete signal $b[k]$ and the SNR of the continuous signal $p(t)$ are approximately the same.

Let the reconstructed signal be $\hat{u}[k]$, and let $\tilde{u}[k]$ be a delayed and amplitude matched source signal, that minimizes the mean square error (MSE) of the reconstructed error signal

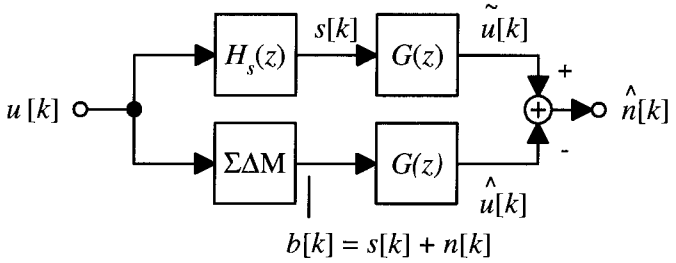


Figure 2.5: Signal reconstruction model for SNR analysis.

$\hat{n}[k]$. The reconstructed signal-to-noise ratio is then defined as

$$SNR = \frac{\mathrm{E}\{|\hat{u}[k]|^2\}}{\mathrm{E}\{|\hat{u}[k] - \tilde{u}[k]|^2\}} = \frac{\mathrm{E}\{|b[k] * g[k]|^2\}}{\mathrm{E}\{|n[k] * g[k]|^2\}} \tag{2.27}$$

where $\mathrm{E}\{\cdot\}$ denotes the expected valued value and $*$ denotes convolution.

The expectations in the SNR expression (2.27) are not easily evaluated directly, and it is much more convenient to express SNR in terms of power spectral densities associated with the modulator signals $b[k]$ and $n[k]$. Assuming the reconstruction filter $g[k]$ is ideal with frequency response

$$G(f) = \begin{cases} 1 & |f - f_c| < f_b/2 \\ 0 & \text{otherwise} \end{cases}, \tag{2.28}$$

and assuming the source signal lies completely within the filter bandwidth, then

$$SNR = \frac{\int_{f_c - f_b/2}^{f_c + f_b/2} S_b(f) df}{\int_{f_c - f_b/2}^{f_c + f_b/2} S_n(f) df}. \tag{2.29}$$

Therefore, an SNR calculation of the reconstructed signal with an ideal bandpass filter requires knowledge of both the pulse train power spectral density $S_b(f)$ and the noise component spectral density $S_n(f)$.

The SNR can also be expressed in terms of the source signal $u[k]$ by expanding the numerator term. After using (2.6), and assuming the STF gain is unity, we get

$$SNR = \frac{\sigma_u^2 + \int_{f_c - f_b/2}^{f_c + f_b/2} [S_n(f) + 2\mathrm{Re}\{S_{un}(f)\, H_s^*(f)\}]\, df}{\int_{f_c - f_b/2}^{f_c + f_b/2} S_n(f)\, df} \approx \frac{\sigma_u^2}{\int_{f_c - f_b/2}^{f_c + f_b/2} S_n(f)\, df}. \tag{2.30}$$

The approximation is valid up to moderate input powers providing $\sigma_u^2 < \Delta_p^2$, and shows the direct relationship between the input signal power and SNR. Eventually the SNR of

the modulator reaches a peak value, and once the modulator input power is large, the cross power term is no longer insignificant. The noise and signal spectrum start to become correlated at high source levels, and the reconstructed signal power is limited to maintain constant pulse train power.

2.4.2 SNR With An Optimum Reconstruction Filter

The ideal bandpass filter is a common reference filter for characterizing the SNR of a modulator, but it is not necessarily the optimum reconstruction filter. Optimizing the reconstructed SNR is analogous to a signal estimation problem where the signal is assumed to be corrupted by an additive noise process. Both the source signal and noise are stochastic, and the optimum filter which minimizes the MSE is called the Wiener filter. The design of ideal and causal Wiener filters is covered in many references such as [51, 52]. Using the non-causal form of the Wiener filter for analytic simplicity, we show by way of an example, that the optimum reconstruction filter $g(t)$ is well approximated by an ideal bandpass filter when the SNR is high.

The input signal to the reconstruction filter $g(t)$ is a pulse train $p(t)$, and it consists of the desired signal $s(t)$ with an additive noise component $n(t)$:

$$p(t) = s(t) + n(t). \tag{2.31}$$

The frequency response of the optimum non-causal Wiener filter is [51]

$$G(f) = \frac{S_{ps}(f)}{S_p(f)} = \frac{S_s(f)}{S_s(f) + S_n(f)} \tag{2.32}$$

where the cross power spectral density $S_{ps}(f)$ and the power spectral density $S_p(f)$ are expanded in terms of the signal PSD $S_s(f)$ and the noise PSD $S_n(f)$. An immediate result from (2.32) is that if the PSD of the signal is bandlimited, then the optimum filter is an attenuator outside the signal bandwidth. However, within the signal bandwidth, the optimum filter depends on the signal and noise PSD's.

As described earlier, the attenuation of the NRZ DAC which converts $b[k]$ to $p(t)$ is assumed to be uniform over the narrow signal bandwidth f_b. Hence signal and noise are attenuated equally, and the relative power spectral densities are unchanged by the DAC. It is convenient to first derive results with discrete-time spectral densities, and then renormalize for equivalent continuous-time densities. The variable ω shall denote the discrete frequency

and has values in the interval $[0, 2\pi)$; the equivalent continuous-time expressions use the substitution $\omega = 2\pi f/f_s$.

The source signal is a bandlimited bandpass signal with bandwidth ω_b, carrier frequency ω_c, and total power σ_u^2. Assume the source signal PSD is uniformly distributed over the signal bandwidth similar to W-CDMA, then the one-sided PSD is:

$$S_s(\omega) = \begin{cases} \sigma_u^2 \, |\omega - \omega_c|/\omega_b & |\omega - \omega_c| < \omega_b/2 \\ 0 & \text{otherwise} \end{cases} . \tag{2.33}$$

The noise PSD over the signal bandwidth is found by approximating the noise transfer function $H_n(e^{j\omega})$ with a Taylor series expansion about ω_c, and assuming the quantization noise power is uniform. Schreier used this technique for estimating the SNR of a bandpass $\Sigma\Delta$ modulator [33]. If the noise transfer function has coincident zeros, then a first term Taylor expansion leads to the expression

$$|H_n(e^{j\omega})|^2 = K^2 \, (\omega - \omega_c)^M \tag{2.34}$$

where K is a gain constant and M is the order of the NTF. For example, the NTF for modulator A was given in (2.19), and the corresponding values for K and M are both 4. The approximation and the exact equation (2.19) are compared in Fig. 2.6, and shows that the approximation is very accurate in the region of the noise well.

The quantization error signal $e[k]$ from the binary quantizer is assumed to independent, white, and uniformly distributed over $\pm\Delta_p$ with variance $\Delta_p^2/3$. The quantization error power is spread over 2π, and the one-sided spectral density is

$$S_e(\omega) = \frac{\Delta_p^2}{3\pi}. \tag{2.35}$$

The PSD of the noise component in the pulse train is then

$$S_n(\omega) = S_e(\omega)|H(e^{j\omega})|^2 = \frac{\Delta_p^2 \, K^2}{3\pi} \, (\omega - \omega_c)^M. \tag{2.36}$$

Substituting expressions for the signal and noise PSD's in (2.32) yields

$$G(\omega) = \frac{\eta_p}{\eta_p + K^2/3 \, R_e^{-1} \, |\omega - \omega_c|^{M-1}} \tag{2.37}$$

where the relations $\eta_p = \sigma_u^2/\Delta_p^2$ and $R_e = \pi/\omega_b$ have been used in the simplification. When the frequency is equal to ω_c, the second term in the denominator is zero, and the

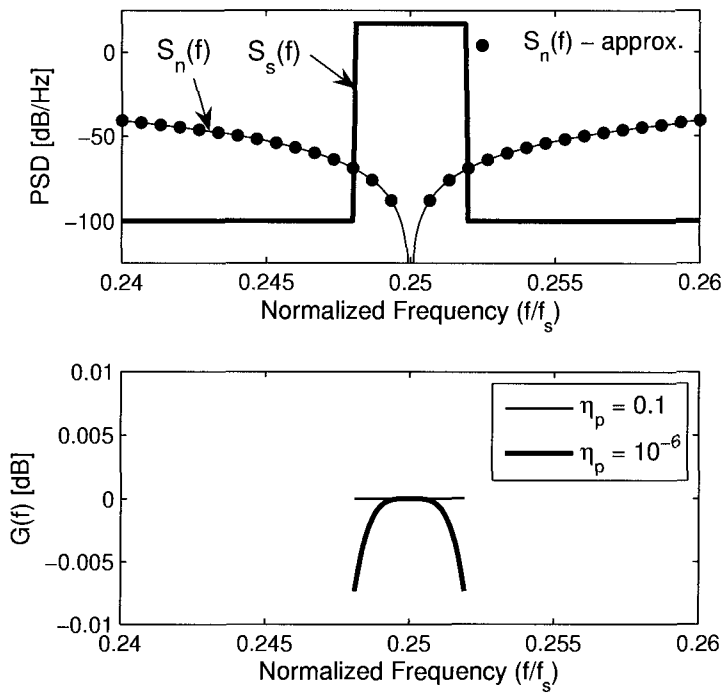


Figure 2.6: Top figure: power spectral density of the signal and noise components for modulator A with $R_c = 2$ and $R_e = 128$. Bottom figure: Optimal Wiener filter frequency responses for different pulse train coding efficiencies.

reconstruction filter has unity gain. For frequencies away from the center frequency, the second term affects the frequency response, but as shown next, this term is very small when the coding efficiency of the signal is high.

The maximum deviation of the optimum reconstruction filter from unity gain, occurs at the band edges when $\omega = \omega_c \pm \omega_b/2$. At these frequencies, the maximum deviation is

$$\max(1 - G(\omega)) = \frac{1}{1 + 3\eta_p R_e^M K^{-2}(2/\pi)^{M-1}}. \tag{2.38}$$

Since the envelope oversample ratio is large, the denominator is large, and the deviation from unity is very small. The impact on the noise spectral density starts to become significant when the coding efficiency is very low, or equivalently the SNR of the reconstructed signal is very low. An example of the optimum filter response for modulator A is given for $\eta_p = 0.1$ and $\eta_p = 10^{-4}$ in Fig. 2.6. Since the coding efficiency is always large in an amplifier application, the analysis shows that the optimum reconstruction filter is well approximated by an ideal bandpass filter.

2.4.3 Sinusoidal Source

SNR measurements with a sinusoidal source use a test signal of the form

$$u[k] = \sigma_u\sqrt{2}\cos[2\pi(f_c + \Delta f)k + \phi] \qquad 0 < f_c < \frac{1}{2} \tag{2.39}$$

where ϕ is a random variable with uniform distribution over $[0, 2\pi)$, and Δf is an offset satisfying $\Delta f < f_b/2$. The definition of f_c and f_b are consistent with the definitions of carrier and envelope oversample ratios. For example, a tone with $f_c = 1/3$ and $\Delta f = \pi/3140 \approx 0.001$ has a carrier oversample ratio of 1.5, and the offset is within a bandwidth defined by an envelope oversample ratio of 128. Depending on the values of the carrier oversample ratio and Δf, the large signal behavior of the modulator may be phase sensitive and ensemble averages are required to estimate SNR. However, for most cases where the source frequency and sample frequency are non-commensurate, a time average of a single sample function is sufficient to estimate SNR. The distinction between ensemble and time averages is important in Chapter 4, otherwise all SNR measurements with a sinusoidal signal are derived from time averages with a frequency offset of $\pi/3140$. A noise bandwidth corresponding to $R_e = 128$ is consistently applied throughout this work for SNR measurements.

As an example of a sinusoidal SNR response, Fig. 2.7 shows the SNR for modulator A with a carrier oversample ratio of 1.5. The SNR increases linearly over most of the dynamic

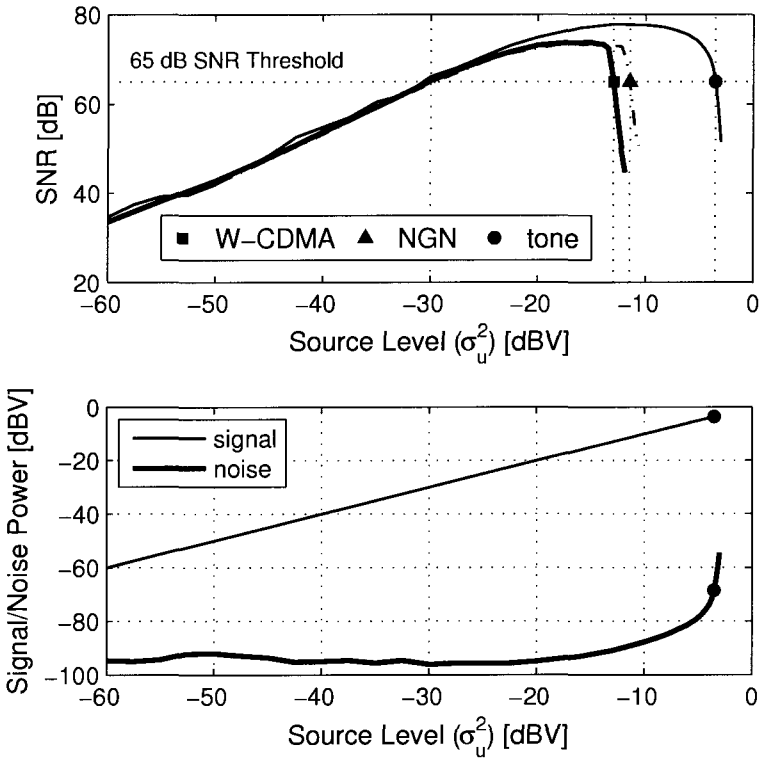


Figure 2.7: Top figure: SNR characteristics for modulator A with sinusoidal, NGN, and W-CDMA source signals; the source levels for a SNR threshold of 65 dB are -3.5 dBV, -11.5 dBV, and -13.0 dBV, respectively. Lower figure: signal and noise power for a sinusoidal source.

range of the modulator, and then it eventually reaches a peak value. With sufficiently large input amplitudes, the modulator SNR collapses very rapidly, and in this region the modulator may eventually be unstable depending on the design. The SNR response is consistent with equation (2.30), and shows the linear range as well as the nonlinear range when signal and noise components start to become correlated. For comparison with the SNR, the power of the reconstructed signal ($\sigma_{\hat{u}}^2$) and reconstructed noise components ($\sigma_{\hat{n}}^2$) are shown in Fig. 2.7(b). Signal power varies linearly with source power, while noise power is constant over most of the dynamic range until it abruptly increases around the point of SNR collapse.

2.4.4 W-CDMA Source

When the input signal to the modulator has a time varying signal envelope like a W-CDMA carrier, the SNR response of the modulator changes. Fig. 2.7 shows the SNR response for modulator A with a W-CDMA source signal. Over a large dynamic range, the SNR increases linearly similar to a sinusoidal source signal, but the peak SNR and the point of SNR collapse are backed off relative to the sinusoidal source. The drive characteristics of the modulator are therefore similar to a class A drive stage, and the peak-to-average power ratio (PAR) of the source signal limits both SNR and coding efficiency.

The envelope distribution of the W-CDMA signal is shown in Fig. 2.8 where the distribution is shown as a complementary cumulative distribution function (CCDF). Using the 65 dB SNR threshold line in Fig. 2.7 as a reference, the maximum W-CDMA drive level is -13.0 dBV, or 9.5 dB less than the sinusoidal SNR drive level. From the CCDF, a PAR of 9.5 dB corresponds to peak amplitude excursions of less than 10^{-3}, very close to a peak PAR of 10.2 dB.

The sensitivity of modulator SNR relative to peak amplitude levels is not surprising; the modulator is an oversampling converter, and the envelope is quasi-static over a large number of clock cycles. Therefore, the maximum drive level for a modulated signal is approximately equal to the sinusoidal drive level backed off by the envelope PAR.

2.4.5 Narrowband Gaussian Source

As an alternate signal source to W-CDMA, narrowband gaussian noise (NGN) can also be used to generate a source signal with similar PAR. NGN is easily generated for simulations,

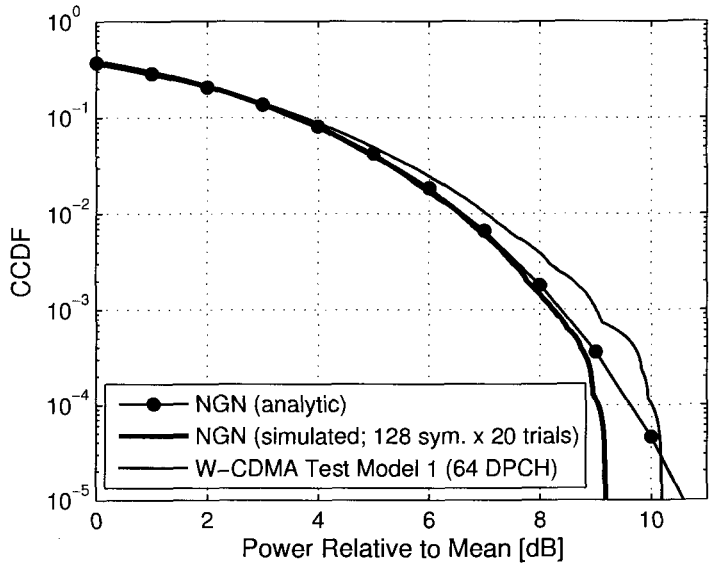


Figure 2.8: Complementary cumulative distribution function of envelope power relative to mean for narrowband gaussian noise and W-CDMA sources. NGN has a Rayleigh envelope distribution, and the analytic CCDF is $e^{-\rho}$ where ρ is the PAR level.

and the gaussian probability distribution has the maximum entropy of any source signal. The SNR response for an NGN source is shown in Fig. 2.7, and the maximum drive level for an SNR threshold of 65 dB is approximately 1.5 dB higher than W-CDMA. The slightly higher drive level for NGN corresponds to a slighter lower peak-to-average power ratio of the NGN test signal.

The NGN source signal is generated with a W-CDMA equivalent root-raised cosine (RRC) filter with a roll-off factor of 0.22. The simulated and analytic PAR for the signal is shown in Fig. 2.8. The simulated CCDF corresponds to the ensemble average of 20 sample functions of 2^{15} points, or a total of 2056 symbols with an envelope oversample ratio of 128. The W-CDMA response corresponds to a single sample function of 2048 symbols of test model 1 in the W-CDMA standard [53]. As the data show, although NGN has a higher theoretical PAR, the simulated result has a PAR of about 9.2 dB, and is 1 dB less than W-CDMA. Higher PAR is generated by more symbols, but the simulation times are very long. The primary reason for using short sample functions and ensemble averaging is to limit numerical computation times which are not necessarily linear with the length of sample sequences. In Chapter 5, numerical complexity is especially significant where tree

search methods are used to evaluate modulator coding efficiency.

2.5 Coding Efficiency

As shown in the introductory chapter, the output power of the RF class D amplifier depends on the coding efficiency of the modulator pulse train. A primary objective in power amplifier design is to maximize load power, and consequently we seek to maximize modulator coding efficiency. As with SNR, it is convenient to measure coding efficiency directly from the modulator pulse train, and the measurement method using the PSD of $p(t)$ is employed. Providing the reconstructed SNR is large, the reconstructed signal power is

$$P_r = \int_{f_c - f_b/2}^{f_c + f_b/2} S_p(f) df. \tag{2.40}$$

Unlike reconstructed SNR, the attenuation of the output DAC is significant and affects coding efficiency; therefore, the coding efficiency of the DT pulse train $b[k]$ must be adjusted for DAC attenuation. Assuming the ZOH DAC pulse is a trapezoidal pulse with rise and fall times of T_e and pulse width T, the CT pulse train $p(t)$ has a coding efficiency of [41]

$$\eta_p = \eta_p^{DT} \big(\text{sinc}[1/(2R_c)]\,\text{sinc}(f_c T_e)\big)^2. \tag{2.41}$$

The superscript DT denotes the coding efficiency of the DT pulse train $b[k]$. If the rise and fall times are less than $T/8$, the term with T_e is negligible. We assume the DAC circuit satisfies this condition, and CT coding efficiency is then a function of only the carrier oversample ratio and the DT modulator coding efficiency.

The ZOH DAC pulse attenuates the DT coding efficiency, and a large carrier oversample ratio is desirable to minimize the reduction in coding efficiency by the $\text{sinc}[1/(2R_c)]$ term. On the other hand, a high sample rate requires higher bandwidth circuitry, consumes more power, and as shown later generates more transitions in the pulse train reducing amplifier power efficiency.

If the STF gain in the modulator is unity as in modulator A, then the signal power in $b[k]$ is equal to the input power. Under this condition, coding efficiency increases with increasing input power. However, coding efficiency is eventually bounded by correlation between the input signal and the quantization noise. As shown earlier in Fig. 2.7 (lower figure), if the input signal power is restricted to a range around the point of SNR collapse

where the SNR is high, no significant signal compression occurs. Therefore, with unity STF gain, and unit quantization levels $\pm\Delta_p$,

$$\eta_p^{DT} = \sigma_u^2 \qquad \sigma_u^2 < \sigma_{u,max}^2 < 1 \tag{2.42}$$

where $\sigma_{u,max}^2$ is the maximum input power for which the approximation is valid.

2.5.1 Coding Efficiency and SNR Trade-offs

If coding efficiency is to be maximized, then the modulator drive level should be set to $\sigma_{u,max}^2$. However, this is unconstrained and does not consider the trade-off with SNR. There is a minimum acceptable SNR threshold SNR_t which must be met while maximizing coding efficiency. A SNR threshold line at 65 dB is shown in Fig. 2.7 and intersects a SNR response at two points. Clearly, the optimal operating point corresponds to the larger drive level, and the figure shows how important the region near SNR collapse is. It is within this region that the coding efficiency becomes significant and where reasonable power can be extracted from the encoded signal.

2.5.2 Coding Efficiency and Carrier Oversample Ratio Trade-offs

One of the main distinctions between lowpass $\Sigma\Delta$ modulation and bandpass $\Sigma\Delta$ modulation is the carrier oversample ratio. From an information theoretic perspective, the SNR of the encoded pulse train depends only on the signal bandwidth and is independent of the carrier frequency. On the other hand, as (2.41) shows, coding efficiency does depend on carrier oversample ratio, and coding efficiency is expected to increase monotonically with the carrier oversample ratio. It is therefore interesting to observe that the large signal amplitude characteristics of bandpass modulators show variations in coding efficiency that are not monotonic, and that some carrier oversample ratios appear more optimal than others. These conclusions are based on extensive simulations based on the modulator designs in section 2.2, as well as supporting theory from other encoding models presented later in Chapters 4 and 5.

As examples of coding efficiency variation which are not monotonic with carrier oversample ratio, Fig. 2.9 shows the SNR versus coding efficiency response for six different modulator designs at $R_c = 1.7$ and $R_c = 2.0$. All the responses shown in this figure are consistent with stable modulators, and the input to the quantizer is bounded. The variance of the quantizer input signal $v[k]$ is less than 130 for the shown operating range.

The NRZ DAC attenuation is larger for a carrier oversample ratio of 1.7, and yet the coding efficiency is substantially better than at 2.0 for all the second and fourth order modulator designs. This observation is consistent over designs including spread zero designs and DT or CT transfer functions. Since the output power delivered by the amplifier depends on coding efficiency, as will be shown in Chapter 3, it is advantageous to select a carrier oversample ratio that maximizes coding efficiency. Based on these data, $R_c = 1.7$ is a better choice than $R_c = 2$.

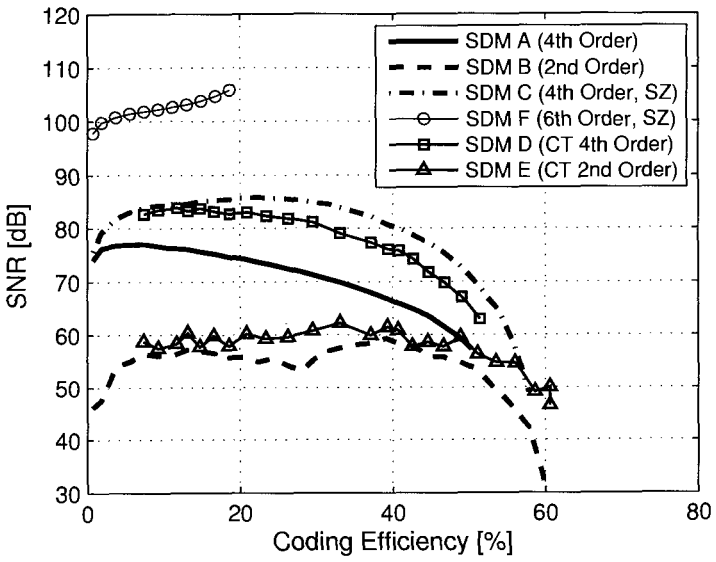
Figure 2.9 also includes the SNR and coding efficiency response of a sixth order modulator design, called design F. Unlike second and fourth order designs, the sixth order modulator performance does not vary significantly with carrier oversample rate. The stability of the sixth order modulator becomes the overriding limitation in the design. As the results show, although the SNR is about 10 to 15 dB better than a fourth order design, the peak coding efficiency is approximately 18%, well below the peak coding efficiency of the second and fourth order designs. The stability issues with high order modulators are well known [21], and this limits their usefulness for high drive level applications. For this reason, most of this research is limited to second and fourth order designs.

Since the variation of coding efficiency with carrier oversample ratio is not monotonic for second and fourth order designs, a characterization of the modulator design should include a sweep over carrier oversample ratio. In the following sections, the coding efficiency variation of modulator A is shown for sinusoidal and W-CDMA sources.

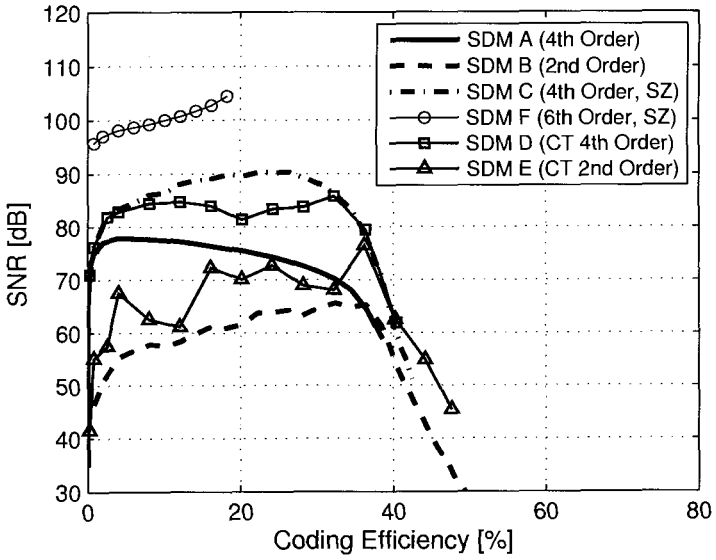
2.5.2.1 Coding Efficiency with a Sinusoidal Source

Fig. 2.10 shows a detailed characterization of modulator A where the coding efficiency at different SNR thresholds is shown as a function of the carrier oversample ratio. The coding efficiency is for a CT pulse train and includes the sinc response of the NRZ DAC. For these data, a sinusoidal source signal with 2^{18} samples is used with an irrational frequency offset of $\pi/3140$. The SNR versus coding efficiency is measured at each carrier oversample ratio, then the coding efficiency corresponding to an SNR threshold is found.

The data clearly show dips in coding efficiency that are not predicted by a simple overlay of the NRZ DAC frequency response. For example, at an SNR threshold of 65 dB, there are dips in coding efficiency at carrier oversample ratios of 1.5, 2, and 2.5. As the SNR threshold is increased, the dips become less pronounced, and eventually the coding efficiency varies monotonically with the NRZ DAC frequency response. Therefore, the coding



(a) $R_c = 1.7$



(b) $R_c = 2.0$

Figure 2.9: SNR and coding efficiency variation with carrier oversample ratio ($R_e = 128$; $\sigma_v^2 < 130$).

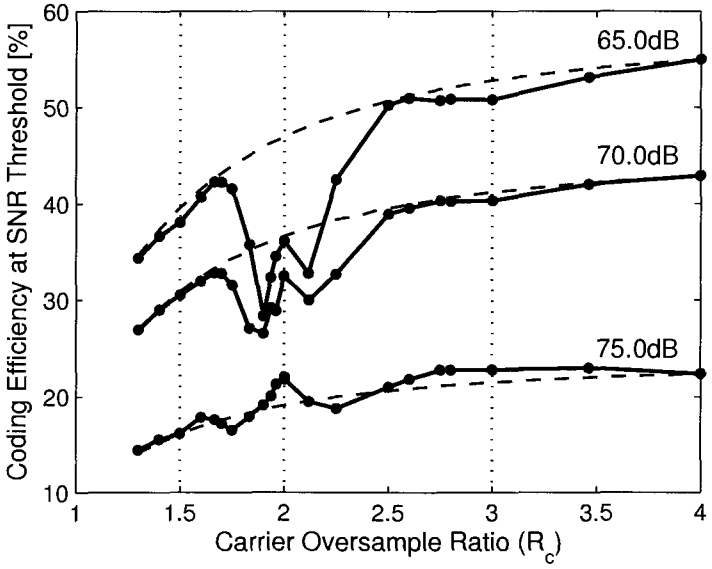


Figure 2.10: Continuous-time pulse train coding efficiency of modulator A at different SNR thresholds for a sinusoidal source. The dashed line is the expected response assuming a NRZ DAC.

efficiency variation is only observed with large source signals, exactly where coding efficiency is maximized.

The variation in coding efficiency with carrier oversample ratio and the SNR threshold is substantial. For example, at an SNR threshold of 65 dB, a carrier oversample ratio of 1.7 has as much as 10% more coding efficiency compared to a modulator encoding at a carrier oversample ratio of 2.0. Assuming DAC attenuation characteristics only, the exact opposite response is expected — coding efficiency should be less at a lower carrier oversample ratio. SNR threshold also has a significant impact on coding efficiency. For a carrier oversample ratio of 1.7, the coding efficiency is approximately 18% at 75 dB, 32% at 70 dB, and 42% at 65 dB. A 10 dB reduction in the SNR threshold results in more than double the reconstructed signal power.

As another representation of the same data, Fig. 2.11 shows the same data set for the DT pulse train $b[k]$ without the DAC response. Also, instead of plotting the coding efficiency versus carrier oversample ratio, the x-axis is shown in terms of a normalized center frequency. The dips in coding efficiency occur around low fractional ratios of $\frac{1}{3}$ ($R_c = 1.5$), $\frac{1}{4}$ ($R_c = 2.0$), and $\frac{1}{6}$ ($R_c = 2.5$). The periodic model in Chapter 4 predicts the dips at these fractional

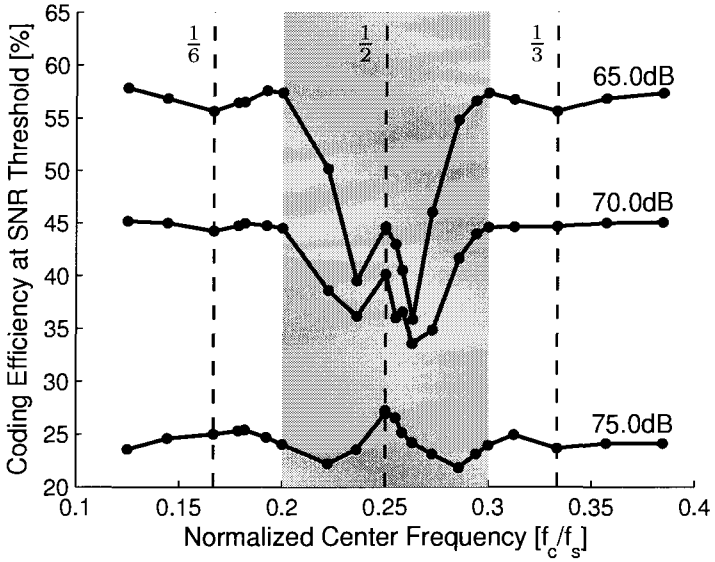


Figure 2.11: Discrete-time pulse train coding efficiency with a sinusoidal source for modulator A.

ratios, and the large signal characteristics of modulators shows similar variation in the neighborhood of these specific carrier oversample ratios.

2.5.2.2 Coding Efficiency with a W-CDMA Source

The coding efficiency variation with carrier oversample ratio is consistent with more complex source signals such as W-CDMA. As an example of this behavior, Fig. 2.12(a) shows the SNR versus coding efficiency responses for six different carrier oversample ratios. A 65 dB threshold line is shown for reference, and the carrier oversample ratios in order of increasing coding efficiency are 2.0, 1.3, 1.5, 1.7, 3, and 2.6. As with the sinusoidal source signal, the variation is not monotonic, and some carrier oversample ratios are more optimal than others.

In Fig. 2.12(b), a series of responses are shown in the neighborhood of a carrier oversample ratio of 2.0. The responses show that the coding efficiency at $R_c = 2$ is not a singularity — there is a narrow range about $R_c = 2$ where coding efficiency deviates significantly from the NRZ DAC response. The behavior is centered at a carrier oversample ratio of 2, and shows that the coding efficiency dip is not anomalous.

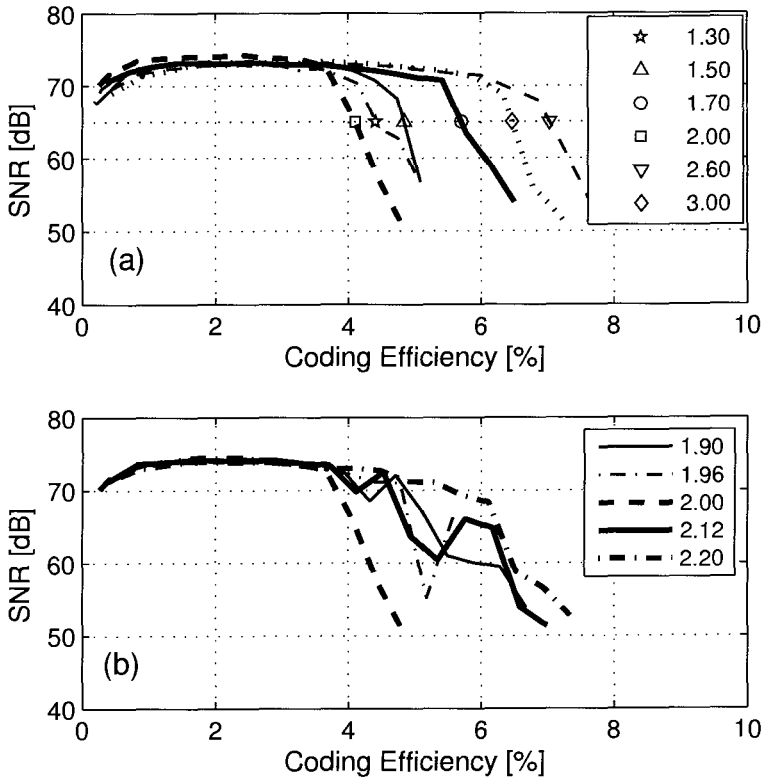


Figure 2.12: SNR versus coding efficiency for modulator A with a W-CDMA source signal. The marked points in (a) correspond to an SNR threshold of 65 dB with coding efficiencies of: 4.4%, 4.8%, 5.7%, 4.1%, 6.5%, and 7.0% for $R_c = 1.3, 1.5, 1.7, 2, 3$ and 2.6, respectively.

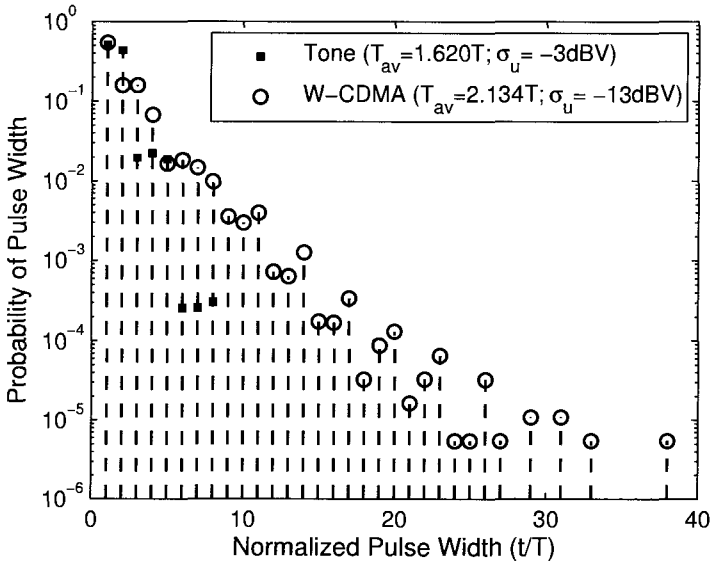


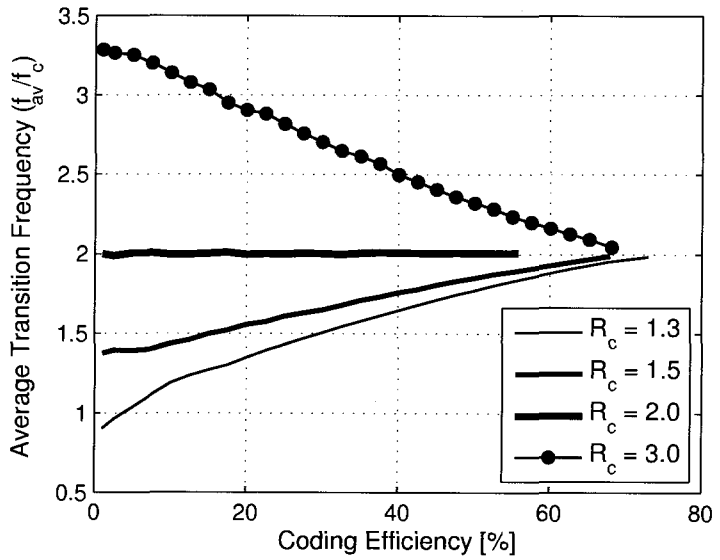
Figure 2.13: Probability mass function of pulse widths for sinusoidal and W-CDMA source signals. The data are for modulator A with $R_c = 1.5$.

2.6 Average Transition Frequency

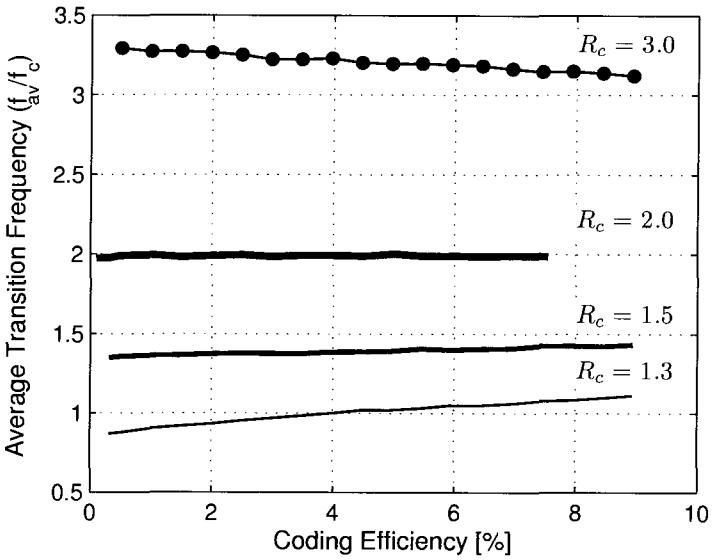
A third modulator parameter, in addition to SNR and coding efficiency, that is important in RF class D amplifier design is the average transition frequency (ATF) of the pulse train. The definition of ATF was given in Chapter 1, equation (1.10), and is defined as the inverse of the average pulse period (T_{av}). The average pulse period is the expected value of the pulse widths in the modulator pulse train $p(t)$. Estimates of the probability mass function for the pulse widths are made by capturing long sample functions of the modulator pulse train, and counting the distribution of pulse widths. An example of a pulse width probability mass function is given in Fig. 2.13.

The pulse width distribution changes with source level and source type. The ATF of modulator A for sinusoidal and W-CDMA sources is shown in Fig. 2.14. Unlike coding efficiency variation with respect to carrier oversample ratio, ATF varies monotonically with changes in carrier oversample ratio.

For W-CDMA source signals, the ATF is nearly independent of the modulator drive level and is primarily a function of carrier oversample ratio. As shown in Fig. 2.14(b), an



(a) Sinusoidal source signal.



(b) W-CDMA source signal.

Figure 2.14: Average transition frequency for modulator A with a sinusoidal and W-CDMA source signal.

empirical approximation for ATF is

$$f_{av} \approx R_c f_c = f_s/2 \qquad \text{(W-CDMA)}. \tag{2.43}$$

The coding efficiency of the pulse train with a W-CDMA source is low, and most of the pulse train power is quantization noise. The noise has a mean frequency of $f_s/2$; hence the approximation (2.43). From this, we conclude that it is best to choose a low carrier oversample ratio to minimize ATF for W-CDMA source signals.

The ATF for a sinusoidal source signal is more interesting, and Fig. 2.14(a) shows that ATF asymptotically approaches $2f_c$ for all carrier oversample ratios as the modulator input levels increases. As will be shown later in Chapter 4, the maximum coding efficiency for a sinusoidal load signal at frequency f_c corresponds to the fundamental frequency component of a square wave signal with a 50% duty cycle. The ATF of the square wave pulse train is $2f_c$, and the ATF convergence of the modulator to $2f_c$ shows that the modulator eventually attempts to generate a square wave signal as the amplitude of the source signal gets large. For low amplitude sinusoidal signals, the ATF depends on carrier oversample ratios, similar to W-CDMA source signals.

2.7 Conclusions

The SNR, coding efficiency, and average transition frequency characteristics of six different bandpass $\Sigma\Delta$ modulators have been shown. Coding efficiency increases with source amplitude, and this means the large amplitude behavior of the modulator is important in RF class D amplifier applications. The NRZ DAC attenuation characteristics have a significant impact on coding efficiency, and high oversample ratios minimize DAC attenuation. In addition to the DAC frequency response, all the second and fourth order modulator designs show a significant dip in coding efficiency in the neighborhood of $R_c = 2$. The simulated data show that it is better to choose a carrier oversample ratio of less than 1.7, or more than 2.2 to avoid the drop in coding efficiency. Since ATF increases with carrier oversample ratio, switching power losses in the amplifier also increase; therefore, the selection of a carrier oversample ratio involves trade-offs between coding efficiency and ATF. In the next chapter, amplifier design equations are derived which incorporate coding efficiency and ATF parameters, and the equations provide a way to select a carrier oversample ratio that balances load power and power efficiency.

Chapter 3

RF Class D Amplifier Power Efficiency

In this chapter, we analyze the power efficiency of the augmented bandpass $\Sigma\Delta$ modulator and RF class D amplifier configuration. The concepts of coding efficiency and average transition frequency developed in the previous chapter are parameters that link the modulator with the analysis of amplifier power efficiency.

We begin by first selecting an RF class D amplifier topology which has a broadband frequency response for amplifying a bandpass $\Sigma\Delta$ modulator pulse train. The complementary voltage switched amplifier topology is considered to be the best choice for bandpass $\Sigma\Delta$ modulation. The power efficiency of the circuit is then analyzed. Design and analysis equations are derived which estimate the optimum load resistance, load power, conduction losses, and capacitive switching losses. The equations are applied to three different designs with pHEMT, MESFET, and CMOS switches. In the latter case, a design with a CMOS driver stage is included in the power efficiency analysis. The analysis is based on first order device models and verified by simulation.

3.1 Complementary Voltage Switched Class D Amplifier

The complementary voltage switched class D (CVSCD) amplifier is considered to be the best circuit topology for amplifying a broadband bandpass $\Sigma\Delta$ modulator pulse train. The advantage of the complementary voltage switched class D amplifier (CVSCD) configuration [3], also called the half bridge circuit in audio class D amplifier literature [54], is that it does not require transformers or baluns. Other RF class D amplifier configurations such as the transformer coupled current switched amplifier (CSCD) [3, 24], or modifications of the

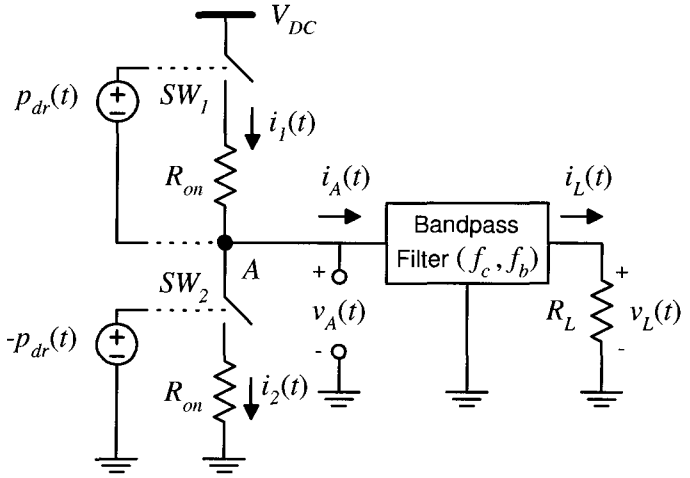
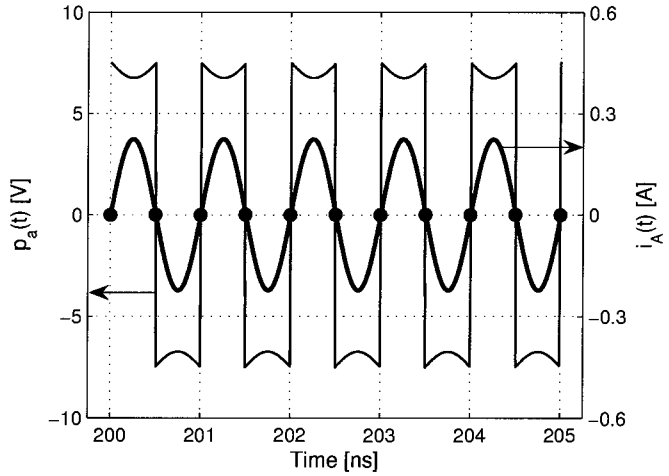


Figure 3.1: Complementary voltage switched class D (CVSCD) amplifier.

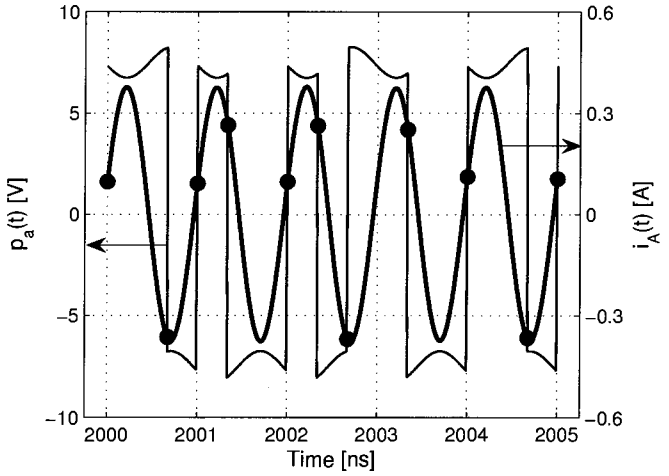
CSCD amplifier such as a shunt current filter [13–15], require transformers or baluns, and are likely to be more challenging to implement given the GHz bandwidth of $\Sigma\Delta$ modulated pulse trains at wireless frequencies. The CVSCD circuit also has the advantage that it is easily implemented in CMOS technology, and advances in RF CMOS technology open potential low power applications for the amplifier architecture.

A simplified model of the CVSCD circuit is shown in Fig. 3.1. Two switches are arranged in a totem pole configuration, and the reconstruction filter is connected to the switches at node A. The switches in the amplifier are modeled as an ideal switch in series with a linear 'on' resistance denoted as R_{on}. The simplified switch model represents a semiconductor device such as a pHEMT, and the equivalent on resistance is derived from device technology parameters.

An elegant feature of class D amplification with a conventional square wave drive signal is that the current is ideally zero at switching instants. The amplifier generates a periodic switched voltage pulse train $v_A(t)$ with a 50% duty cycle square wave, and the load current is the first harmonic of the pulse train. When the input drive signal $p_{dr}(t)$ is changed to an aperiodic pulse train, the timing of pulse train transitions varies, and the phase of the load signal no longer has a fixed relationship with the voltage transitions. A comparison of the current $i_A(t)$ and the polar signal $p_a(t)$, defined as $v_A(t) - V_{DC}/2$ in the CVSCD circuit, are shown in Fig. 3.2 for periodic and aperiodic switching.



(a) Zero Current Switching - 50% Duty Cycle Square Wave Drive



(b) Non-zero Current Switching - Bandpass ΣΔ Modulator Drive

Figure 3.2: Switching conditions in a VSCD amplifier with non-zero switch resistance. The current at the switching transitions is marked by a dot, and the voltage pulse train has ripple generated by non-zero switch resistance. If switch resistance is zero, then the amplitude levels of $p_a(t)$ would be exactly ± 7.5 V.

Although the principal operation of the class D amplifier remains unchanged whether the input pulse train is a periodic or aperiodic, the switch requirements are different in the two cases. When $p_{dr}(t)$ is a 50% duty cycle square wave, the switch currents $i_1(t)$ and $i_2(t)$ are always positive for the current directions shown in Fig. 3.2. On the other hand, when $p_{dr}(t)$ is aperiodic, the switch currents are no longer unidirectional, and the sign of $i_1(t)$ and $i_2(t)$ can change while the switch is on. Bidirectional current switches are therefore required when the CVSCD amplifier is driven by aperiodic pulse trains. Switches are discussed more in section 3.1.3.

3.1.1 Ideal Circuit Analysis - Zero Switch Resistance

We start the analysis of the CVSCD circuit assuming zero switch resistance, and then adjust design equations to compensate for non-zero switch resistance. The ideal load resistance, assuming zero switch resistance, is indicated with a prime, R'_L, and compensated load resistance for non–zero switch resistance is indicated without a prime, R_L.

The amplifier generates an average load power P_L in R'_L, and the load signal has a time varying envelope with a peak-to-average power ratio of ρ. The peak load power is

$$P_{L,pk} = \frac{V_{L,pk}^2}{2R'_L} = \frac{1}{2} I_{L,pk}^2 R'_L \tag{3.1}$$

the average load power is

$$P_L = \frac{V_{L,rms}^2}{R'_L} = I_{L,rms}^2 R'_L \tag{3.2}$$

and the load PAR is

$$\rho = \frac{P_{L,pk}}{P_L} = \frac{V_{L,pk}^2}{2P_L R'_L} = \frac{I_{L,pk}^2 R'_L}{2P_L}. \tag{3.3}$$

The average load power (3.2) can also be expressed in terms of the coding efficiency η_p of the pulse train $p_a(t)$. With reference to Fig. 3.1, the switch closures generate a voltage switched pulse train $v_A(t)$ with amplitude levels of either 0V or V_{DC}. The switched pulse train $v_A(t)$ has a DC offset of $V_{DC}/2$, and the equivalent zero mean pulse train $p_a(t)$ is

$$p_a(t) = v_A(t) - V_{DC}/2. \tag{3.4}$$

Let $\Delta_a = V_{DC}/2$, then with the coding efficiency equation (1.6), the required load resistance given load power P_L is

$$R'_L = \frac{\eta_p \Delta_a^2}{P_L}. \tag{3.5}$$

If (3.5) is substituted for R'_L in (3.3), then an expression for the peak load current is derived.

$$I_{L,pk} = \frac{P_L}{V_{DC}/2}\sqrt{\frac{2\rho}{\eta_p}} \tag{3.6}$$

Equation (3.6) is consistent with conventional square wave drive signal equations [3] assuming coding efficiency is set to $8/\pi^2$ and ρ is set to 1. The coding efficiency value of $8/\pi^2$ is derived by using equation (1.7) from Chapter 1, where the 'signal component' corresponds to the fundamental frequency component of a 50% duty square wave pulse train.

Assuming a switch has a gate (or base) width of W_{sw} mm, (3.6) is rearranged to give an expression for normalized load power. Let $\bar{P}_L$ be a normalized load power defined as P_L/W_{sw} with units of W/mm, where the overbar denotes a normalized variable. Typical semiconductor devices are often characterized by a maximum drain current called I_{max} [55] which can also be normalized, and the normalized current, $\bar{I}_{max}$, is a figure of merit for device technology. For example, a typical pHEMT may have an $\bar{I}_{max}$ of 0.4 A/mm. With these definitions, (3.6) is rewritten as

$$\bar{P}_L = \frac{P_L}{W_{sw}} = \bar{I}_{max} V_{DC}\sqrt{\frac{\eta_p}{8\rho}} \tag{3.7}$$

where it is assumed that $I_{L,pk} \equiv I_{max}$ and the device is fully utilized.

Equation (3.7) shows the fundamental parameters that affect the utility of both the device and circuit topology to generate load power. Both $\bar{I}_{max}$ and V_{DC} are determined by the device technology, coding efficiency is a function of the modulator, and ρ is a function of the source signal. The utilization of a specific device technology can therefore be boosted either by increasing coding efficiency or by reducing the PAR of the source signal.

3.1.2 Compensation for Switch Resistance

The load current, through switches with non-zero resistance, generates a voltage drop which adds a small ripple component to the switched voltage waveform $v_A(t)$ as shown in Fig. 3.2. The voltage drop across the switches reduces the amplitude of the desired signal component in the switched pulse train, and the loss is compensated by decreasing the ideal load resistance R'_L. The compensated load resistance, R_L, assuming a switch resistance of R_{on} is calculated next.

Since the switch resistance is not zero, power is dissipated in both switches and the load. The power dissipated in the switches (conduction losses) P_c and the power dissipated in the

load P_L must equal the total signal power P_s available in the switched pulsed train.

$$P_c + P_L = P_s \tag{3.8}$$

The conduction power loss is $I_{L,rms}^2 R_{on}$, the load power is $I_{L,rms}^2 R_L$, and the available signal power in the switched voltage pulse train for an effective load of $R_{on}+R_L$ is $\eta_p \Delta_a^2/(R_{on}+R_L)$. With these relations, (3.8) is written as

$$P_L + P_L\left(\frac{R_{on}}{R_L}\right) = \frac{\eta_p \Delta_a^2}{R_{on}+R_L}. \tag{3.9}$$

Equation (3.9) is rearranged as a quadratic equation in R_L,

$$P_L {R_L}^2 + \left(2P_L R_{on} - \eta_p \Delta_a^2\right) R_L + P_L R_{on}^2 = 0 \tag{3.10}$$

and R_L is

$$R_L = \frac{\eta_p \Delta_a^2}{2P_L}\left(1 \pm \sqrt{1 - 4R_{on}\frac{P_L}{\eta_p \Delta_a^2}}\right) - R_{on}. \tag{3.11}$$

The solution corresponds to the positive root such that $\lim_{R_{on}\to 0} R_L = R'_L$ and collapses to (3.5), the zero resistance solution. Using (3.5), R_L is expressed as

$$R_L = R'_L\left(\frac{1+\sqrt{1-4R_{on}/R'_L}}{2}\right) - R_{on}. \tag{3.12}$$

Assuming $R_{on} \ll R'_L$, the root is expanded in a Taylor series and

$$R_L \approx R'_L - 2R_{on} + O(R_{on}^2). \tag{3.13}$$

The design equations are consistent with conventional equations [26] assuming the coding efficiency and PAR for a 50% duty cycle square wave are used.

3.1.3 Current Utilization Margin

A FET like switch, such as a MESFET or pHEMT, is well suited for voltage switched class D circuits, since it supports bidirectional current flow when the switch is turned on. The main precaution in designing the switch is to ensure that the gate-drain junction does not forward bias significantly, and in low frequency circuits the source to drain terminals are often shunted by a diode for additional protection [54]. At RF frequencies, the addition of a shunt diode increases parasitic capacitance, and a design without a diode is considered by

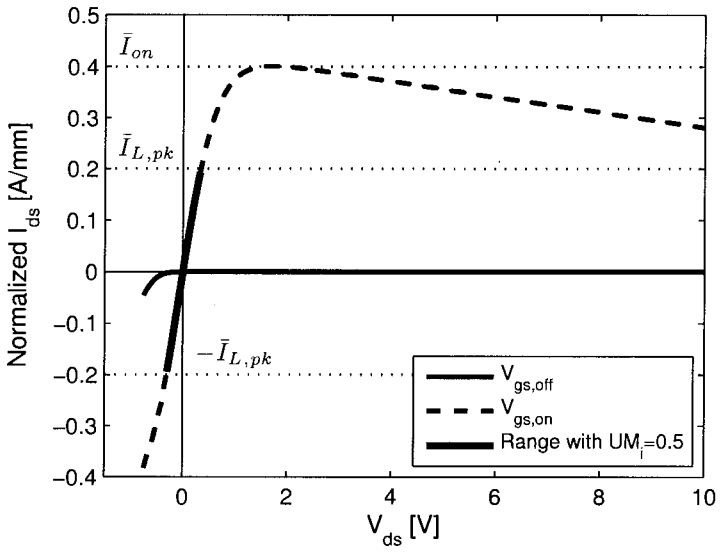


Figure 3.3: Normalized IV characteristics for a typical 0.5 μm pHEMT in the on and off states. Angelov model values for equation (3.30): $I_{pk} = 0.23$ A, $V_{pk} = -0.3$ V, $P_1 = 1.6$, $P_3 = 0.7$, $\lambda = -0.035$ and $\alpha = 1.5$.

limiting load current to a safe operating range where negative drain-source voltage swings do not forward bias the gate-drain junction.

The typical IV characteristics for a n-type depletion pHEMT are shown in Fig. 3.3. The drain current I_{ds} has been normalized to the gate width W_{sw} and shown in A/mm. The load resistance in the amplifier is selected to operate the switch in the linear region where the on resistance is approximately constant, and the load current swings from $-\bar{I}_{L,pk}$ to $+\bar{I}_{L,pk}$.

The gate to source voltage, $V_{gs,on}$, is selected to minimize the on resistance of the device, while at the same time maximize the peak current swing $\pm I_{L,pk}$ without forward biasing the gate-drain junction at a drain current of $-I_{L,pk}$. Depending on the device IV characteristics, the maximum drain current at $V_{gs,on}$, called I_{on}, will probably lie in the range of I_{DSS} to I_{max} for a typical microwave power FET. For example, in Fig. 3.3, $\bar{I}_{on}$ is approximately 0.4 A/mm when the gate source voltage is 0.4 V and is typical for pHEMT devices. At this gate bias, negative drain current up to a maximum level of $-\bar{I}_{on}$ A/mm does not forward bias the gate drain junction, and gate current is kept to a level of less than 1 mA/mm. Assuming a current utilization margin UM_i is applied in the design, $I_{L,pk}$ is selected to be

less than I_{on} where

$$I_{L,pk} = UM_i\bar{I}_{on}W_{sw}. \tag{3.14}$$

There are several reasons for incorporating a current utilization margin in the design. First, the on resistance becomes much larger if signal peaks drive current to a maximum level of I_{on}. When on resistance becomes highly nonlinear, conduction losses are higher, and distortion is larger. Thus, a current margin constrains current swing to a more linear range. A second reason for applying a current margin is to have sufficient overhead to source capacitive switching current during switch transitions. During switch transitions, the phase of the load current could be near a peak, and sufficient current must be supplied to prevent current starvation during a transition. Current is also required during a transition to change the stored charge in parasitic capacitances; therefore, the utilization margin affects switching time by balancing the competing current loads during transitions. An example of a 50% current utilization margin is shown by the heavy line in Fig. 3.3; over this range, the normalized on resistance, $\bar{R}_{on}$, is approximately 2 Ω mm.

Another switch configuration is to shunt a bipolar device such as a HBT with a diode [2]. The main advantage of bipolar switches is the nearly constant saturation voltage which is independent of switch current; however, the advantage of a constant saturation voltage is diminished when the load current changes polarity. Both the HBT and diode have saturation voltages which create an abrupt transition region in load current as the switch voltage changes polarity. The abrupt current transition generates distortion and given the advantage of the linear on resistance in FET's, the bipolar switch configuration is not analyzed further.

3.2 Conduction Losses

Conduction losses are defined as the dissipated power in the switches arising from load current. Assuming instantaneous switching, the current $i_A(t)$ passes through either SW_1 or SW_2, and the conduction loss, denoted as P_c, is $I^2_{L,rms}R_{on}$. The drain efficiency of the amplifier assuming only conduction losses is

$$\eta_d = \frac{P_L}{P_L + P_c} = \frac{R_L}{R_L + R_{on}} \tag{3.15}$$

where R_L is given by (3.11).

The expression for R_L in (3.11) is written in terms of the load power P_L, and it is desirable to have an efficiency expression in terms of fundamental device parameters only.

With equations (3.1) and (3.3), P_L is

$$P_L = I_{L,pk}^2 R_L/(2\rho). \tag{3.16}$$

Substituting (3.16) in (3.10) and solving the revised quadratic equation yields:

$$R_L = \frac{\Delta_a}{I_{L,pk}}\sqrt{2\eta_p\rho} - R_{on}. \tag{3.17}$$

Equation (3.14) is now substituted for $I_{L,pk}$ in (3.17) and

$$R_L = \frac{\Delta_a}{UM_i\,\bar{I}_{on}W_{sw}}\sqrt{2\eta_p\rho} - R_{on}. \tag{3.18}$$

By multiplying (3.18) by the switch gate width W_{sw}, a normalized load resistance $\bar{R}_L$ is computed and substituted in (3.15). In terms of normalized device parameters $\bar{I}_{on}$ and $\bar{R}_{on}$, the drain efficiency is:

$$\eta_d = 1 - \frac{UM_i\bar{I}_{on}\bar{R}_{on}}{\Delta_a\sqrt{2\rho\eta_p}} \tag{3.19}$$

The peak-to-peak amplitude of the switched voltage pulse train is $2\Delta_a$, and the voltage swing is limited by the gate-drain breakdown voltage $|BV_{gd}|$ of the device. Assuming a voltage utilization margin UM_v is assigned to provide margin for peak gate-drain voltages which are at least $|V_{gs,off} - 2\Delta_a|$, then

$$\Delta_a = \frac{UM_v\,|BV_{gd}|}{2} \tag{3.20}$$

and

$$\eta_d = 1 - \frac{2UM_i\,\bar{I}_{on}\,\bar{R}_{on}}{UM_v\,|BV_{gd}|\sqrt{2\rho\,\eta_p}}. \tag{3.21}$$

Equation (3.21) shows how drain efficiency and coding efficiency are linked. If the amplifier is ideal, $\bar{R}_{on} \equiv 0$, and drain efficiency is unity. On the other hand, if $\bar{R}_{on} \neq 0$, coding efficiency affects the conduction losses in the amplifier.

3.3 Capacitive Switching Losses

In this section, the analysis of the class D amplifier is expanded to include switching losses generated by the change in stored energy of parasitic switch capacitance. The analysis of

capacitive switching losses for conventional switching with a square wave is well known [3,23], and the intent here is to expand the analysis to include switching losses with a bandpass $\Sigma\Delta$ modulator drive signal.

The analysis of switching losses for square wave signals is typically done by computing the change in stored energy after a switch transition, and then combining the net change in stored energy with the frequency of transitions to compute dissipated power. The analysis assumes that the amount of dissipated energy is equal to the change in stored energy, and is strictly valid for linear capacitance models only where conservation of energy holds.

The same energy loss approach is considered for the bandpass $\Sigma\Delta$ modulator driven class D amplifier configuration, except the validity of the energy loss assumptions has to be re-evaluated. Since switch current is non-zero during transitions, there is a voltage drop across the switch, and the stored energy varies with load current. Device capacitances such as C_{gs} and C_{gd} are typically nonlinear, and the combined variation in stored energy with load current is expected to be greater than if device capacitances are assumed to be linear. These effects make the analysis of switching loss more complicated with bandpass $\Sigma\Delta$ modulation. In the following sections, we evaluate the potential errors in calculating switching loss with storage energy methods by considering a typical pHEMT, and show that these errors are within 20% for a first order analysis.

3.3.1 Charge and Discharge Current Paths

In Fig. 3.4, the CVSCD circuit is shown with switches modeled by an ideal switch, a current source, and three intrinsic capacitances: C_{gs}, C_{gd} and C_{ds}. The circuit model is set up to analyze the charge and discharge current paths in the various capacitors during an instantaneous switching transition when SW_1 is switched off and SW_2 is switched on. A pulse is shown next to the current sources which are active during the transition, and the peak current available for charging is limited by the device, in this case no more than I_{on} A.

Prior to the change in switch states, $i_A(t)$ is flowing through SW_1, and the capacitances C_{gs1}, C_{gd1}, and C_{ds1} are all charged to low voltages, while the capacitances C_{gs2}, C_{gd2}, and C_{ds2} are all charged to high voltages. After SW_1 opens and SW_2 closes, current is sourced by I_{ds2} to change the capacitor voltages to the alternate state. The capacitor currents i_{Cgd1}, i_{Cds1}, i_{Cgd2} and i_{Cds2} all pass through SW_2 and dissipate power. Power is also dissipated in the drivers I_{dr1} and I_{dr2}, by currents i_{Cgd1}, i_{Cgs1}, i_{Cgs2} and i_{Cgd2}.

Let the change in stored energy of the device capacitances for a single switch be ΔW_{gs},

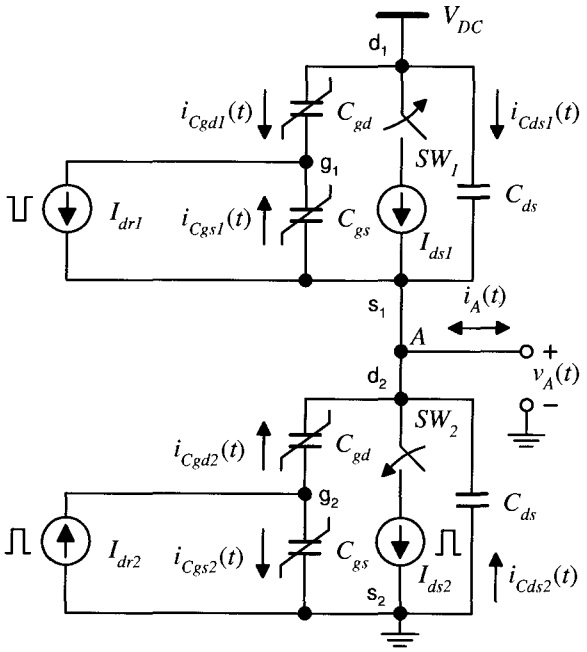


Figure 3.4: Capacitive currents when SW_1 opens and SW_2 closes.

ΔW_{gd}, and ΔW_{gd}, then the total energy dissipated per voltage transition is

$$W_Q = \Delta W_{gs} + \Delta W_{gd} + \Delta W_{gd}. \tag{3.22}$$

For a periodic drive signal, W_Q is deterministic, but for a bandpass $\Sigma\Delta$ modulator W_Q is a random variable and depends on the distribution of the current $i_A(t)$ at the voltage transitions. The average power dissipated by capacitive switching losses is dependent further on the average frequency of transitions determined from the random process $p_a(t)$. If the distribution of the current $i_A(t)$ at transitions is assumed to be independent of when a transition occurs, then the total switching loss is

$$P_Q = 2\, f_{av}\, \mathrm{E}\{W_Q\} \tag{3.23}$$

where $\mathrm{E}\{\cdot\}$ is the expectation taken over $i_A(t)$, f_{av} is the average transition frequency (ATF) of the pulse train $p_a(t)$, and the factor of two accounts for two switches.

The amplitude probability distribution function (PDF) of the current $i_A(t)$ is the same as the source PDF, and the transitions in the pulse train $p_a(t)$ occur at integer multiples of the sampling clock (see Fig. 3.2(b)). Since sampling is independent of the source signal,

and the modulator oversamples the envelope by a large ratio, the amplitude distribution of $i_A(t)$ is expected to be independent of when a transition occurs. This is readily verified by simulation; for example, when the source is a modulated W-CDMA signal, we find the amplitude distribution of the current is approximately gaussian, which is expected for this type of source signal. Proceeding with the independence assumption, we now focus on evaluating (3.23).

3.3.2 pHEMT/MESFET Capacitance Models

A simplified large signal device model is considered for the analysis of switching loss in the CVSCD amplifier. Both the gate-source C_{gs} and gate-drain C_{gd} capacitances are considered as nonlinear capacitances, while the drain-source capacitance is modeled as a linear capacitor. The intrinsic capacitances are typical of large signal models for pHEMT's and MESFET's such as the Curtice [56], Angelov [57], Statz [58], and COBRA [59] models.

In a modulator driven CVSCD circuit, the device drain-source voltage V_{ds} is reverse biased when the switch is on, and the switch currents $i_1(t)$ and $i_2(t)$ are negative. Consequently, devices must be characterized over reverse bias conditions for class D amplifier design. A literature survey was made to find examples of modeled and experimental measurements for pHEMT's and MESFET's under reverse bias conditions and this yielded relatively few papers. Example data from two papers is used for illustrative purposes. The first, is a paper by Angelov [57] in which he provides a complete model for a 0.35 μm $\times$ 200 μm pHEMT which is valid over a V_{ds} range from -1V to +5V. A second paper, by Cojocaru and Brazil [59], includes measured characteristics for a 0.2 μm $\times$ 120 μm pHEMT.

In Angelov's paper he gave nonlinear model equations for C_{gs} and C_{gd} with dependencies on the gate-to-source voltage V_{gs} and the drain-to-source voltage V_{ds}. The model equations are shown in Fig. 3.5 for two gate source conditions: $V_{gs,off} = -1.2$ V and $V_{gs,on} = +0.4$ V. The capacitances are normalized to the gate width of the device, in this case 200 μm, for extrapolation.

Two additional data points indicated by the triangles in Fig. 3.5 correspond to normalized capacitances measured at $V_{ds} = 0$ V for an 0.2 μm pHEMT in Cojocaru and Brazil's paper. The normalized data are relatively consistent, and Angelov's model equations are used for subsequent examples.

For analyzing switching loss, the data in Fig. 3.5 is simplified and only the capacitance ranges corresponding to the on and off device states are considered. When the switch is on,

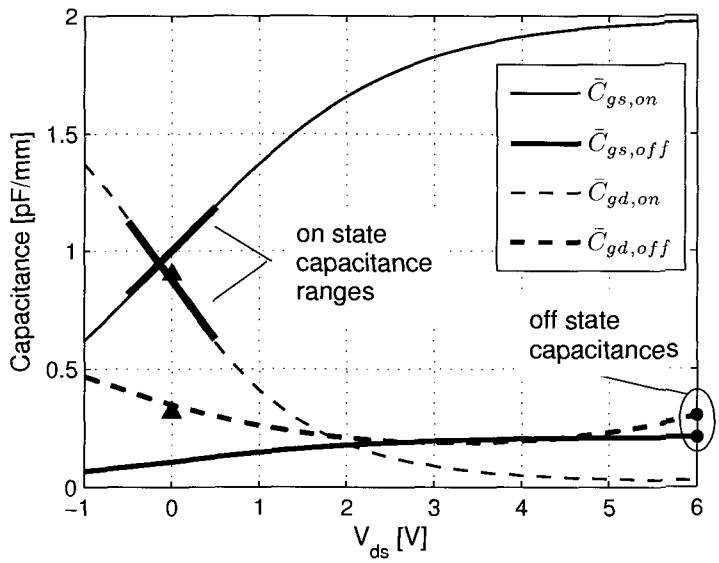


Figure 3.5: Angelov's C_{gs} and C_{gd} functions normalized to the gate width for $V_{gs,on} = 0.4$ V and $V_{gs,off} = -1.2$ V. The device is a 0.35 μm $\times$ 200 μm pHEMT. The triangles are data for a 0.2 μm $\times$ 120 μm pHEMT from Cojocaru and Brazil.

$v_{ds,on}(t)$ is low and equals

$$v_{ds,on}(t) = \pm i_{ds,on}(t) R_{on}. \tag{3.24}$$

The range over which $v_{ds,on}(t)$ swings is $[-I_{L,pk}R_{on}, +I_{L,pk}R_{on}]$, and the capacitances vary over the range indicated by the heavy lines in Fig. 3.5. Both C_{gs} and C_{gd} are highly nonlinear when the switch is on, but well approximated by simple equations:

$$\bar{C}_{gs,on} = \frac{3}{8} V_{ds,on} + 1\,\text{pF/mm} \quad |V_{ds,on}| < 1\text{V} \tag{3.25}$$

$$\bar{C}_{gd,on} = \frac{-1}{2} V_{ds,on} + \frac{7}{8}\,\text{pF/mm} \quad |V_{ds,on}| < 1\text{V}. \tag{3.26}$$

On the other hand, when the device is off, V_{ds} is large and equal to V_{DC}. In this region, the capacitances are relatively constant and, from Fig. 3.5, $\bar{C}_{gs,off}$ and $\bar{C}_{gd,off}$ are approximately equal to 0.2 pF/mm and 0.3 pF/mm, respectively. The data is now analyzed to determine the change in stored energy for each capacitance in the CVSCD circuit.

Table 3.1: Change in Stored Energy vs. Switch Current

	Energy change when $\bar{I}_{ds,on}$ is:			Simplified	Approx.
	0 A/mm	−0.2 A/mm	+0.2 A/mm	Approx.	Error
ΔW_{gs}	0.22 pJ	0.21 pJ	0.24 pJ	0.26 pJ	±17 %
ΔW_{ds}	2.5 pJ	2.3 pJ	2.7 pJ	2.5 pJ	±8 %
ΔW_{gd}	18.9 pJ	17.8 pJ	20.2 pJ	18.8 pJ	±7 %

3.3.3 Stored Energy

Assume an amplifier generates a switched voltage pulse train $v_A(t)$ of 10 V, $V_{gs,on} = 0.4$ V, and $V_{gs,off} = -1.2$ V. Let the switch resistance $\bar{R}_{on}$ be 2 Ω mm and $\bar{I}_{max} = 0.4$ A/mm. If a current utilization margin of 50% is used for the peak load current, then $\bar{I}_{ds,on}$ is in the range of ±0.2 A/mm. Table 3.1 shows the corresponding change in stored energy for C_{gs}, C_{ds}, and C_{gd} versus different load current conditions during a switching interval. Equations (3.25) and (3.26) are used for $\bar{C}_{gs,on}$ and $\bar{C}_{gd,on}$, while $\bar{C}_{gs,off} = 0.2$ pF/mm, $\bar{C}_{gd,off} = 0.3$ pF/mm, and $\bar{C}_{ds} = 0.05$ pF/mm.

The data leads to several simplifications for the analysis of parasitic capacitance energy loss in the CVSCD circuit. First, C_{gd} is the dominant parasitic and has the largest change in stored energy. Second, with reference to columns 3, 4, and 5 in Table 3.1, the sensitivity in stored energy to different load current conditions is small, and within 8% of the stored energy when $I_{ds,on}$ is equal to zero. Third, the net change in energy is dominated by the stored energy when the switch is off. The tabulated data does not show this explicitly, but this is easily understood in the context of the energy storage equation $\frac{1}{2}CV^2$. When the switch is on, the voltage across the capacitances is relatively small and the stored energy is small, even though the nonlinear capacitance of C_{gd} and C_{gs} are large in the on state. In the off state, the nonlinear capacitances are small while the voltage is large, and since the stored energy is proportional to the square of the voltage, the stored energy is much larger.

Based on the relative magnitudes of stored energy in the on and off states, the analytic equations for energy storage are simplified in the following ways. First, the capacitor voltages are assumed to be constant and independent of the load current during the transition interval. By making this assumption, the expected value of energy loss $\mathrm{E}\{W_Q\}$ in (3.23) is reduced to a deterministic quantity. Second, C_{gs} and C_{gd} are assumed to be constant and equal to the off state capacitances, and third, the stored energy when the switch is on is

assumed to be zero for C_{gd} and C_{ds}. Based on these assumptions, the following equations are used as analytic estimates of energy dissipation during a transition:

$$\Delta W_{ds} = \frac{1}{2} C_{ds} V_{DC}^2 \tag{3.27}$$

$$\Delta W_{gs} = \frac{1}{2} C_{gs,off} (V_{gs,on} - V_{gs,off})^2 \tag{3.28}$$

$$\Delta W_{gd} = \frac{1}{2} C_{gd,off} (V_{DC} - V_{gs,off})^2. \tag{3.29}$$

The energy equations (3.27)-(3.29) for the parasitic capacitances are independent of the input drive signal, and apply equally well to conventional square wave drive signals as well as $\Sigma\Delta$ modulator drive signals. The drive signal independence results from a low on resistance which leads to the approximation that capacitances are fully discharged in the on state, and fully charged to static voltages in the off state.

If the simplified equations are used to approximate the change in stored energy, the worst case errors are 17%, 8%, and 8% for ΔW_{gs}, ΔW_{ds}, and ΔW_{gd}, respectively. The data is tabulated in columns 4 and 5 of Table 3.1 for comparison with the example. Since ΔW_{gd} is much larger than the other two stored energy components, the error in ΔW_{gd} is the most important term in power loss calculations and the equations are used for first order approximations.

3.4 pHEMT/MESFET Amplifier Design

The design equations are now applied to predict the load power and drain efficiency of a CVSCD design. The design is evaluated for pHEMT and MESFET switches with a gate width of 1 mm assuming normalized device parameters shown in Table 3.2. The input pulse train is assumed to be generated by modulator A, and we begin by first selecting a modulator operating point that balances load power and power efficiency.

3.4.1 Modulator Operating Point

In chapter 2, it was shown how coding efficiency and ATF vary with the carrier oversample ratio of modulator A. The coding efficiency data show that there is a local maximum around $R_c = 1.7$. The data also show that coding efficiencies greater than the local maximum are obtained for higher carrier oversample ratios in excess of 2.2. However, high carrier oversample ratios increase ATF, and power efficiency is expected to decrease.

Table 3.2: pHEMT and MESFET Normalized Device Parameters

Normalized Parameter	pHEMT Value	MESFET
$\bar{I}_{on}$	0.4 A/mm	0.26 A/mm
$\bar{R}_{on}$	2 Ω mm	4 Ω mm
$\bar{C}_{gs}$	0.2 pF/mm	1 pF/mm
$\bar{C}_{gd}$	0.3 pF/mm	0.125 pF/mm
$\bar{C}_{ds}$	0.050 pF/mm	0.05 pF/mm
$V_{gs,on}$	0.4 V	0.4 V
$V_{gs,off}$	-1.6 V	-1.8 V
BV_{gd}	-15 V	-15 V

The relationship between load power and drain efficiency as a function of the modulator carrier oversample ratio are shown in Fig. 3.6. For this figure, the source is a W-CDMA signal, the SNR threshold is 65dB, the device is a 1 mm pHEMT, the current utilization margin is 1, and the voltage utilization margin is 0.9. As expected, the general trends are that load power increases with carrier oversample ratio, while power efficiency decreases with oversample ratio. Both load power and power efficiency drop in the neighborhood of $R_c = 2$, which is consistent with the observation that coding efficiency dips significantly for this range of carrier oversample ratios. The local maximums for load power around $R_c = 1.7$ and $R_c = 2.6$ are similar to the variation in modulator coding efficiency with carrier oversample ratio, and either point represents a good operating point. Since power efficiency is about 6% better at $R_c = 1.7$, it is selected as the operating point for the design. At this operating point, the predicted load power is 128 mW and the drain efficiency is 60.6%.

3.4.2 Optimum Load

Using design equation (3.18), the optimum load resistance is calculated for three different pulse trains assuming a pHEMT switch. The pulse trains are: 1) a periodic square wave with 50% duty cycle, 2) a bandpass $\Sigma\Delta$M pulse train with a sinusoidal source, and 3) a bandpass $\Sigma\Delta$M pulse train with a 10 dB PAR W-CDMA source. Table 3.3 summarizes the different designs for $W_{sw} = 1$ mm.

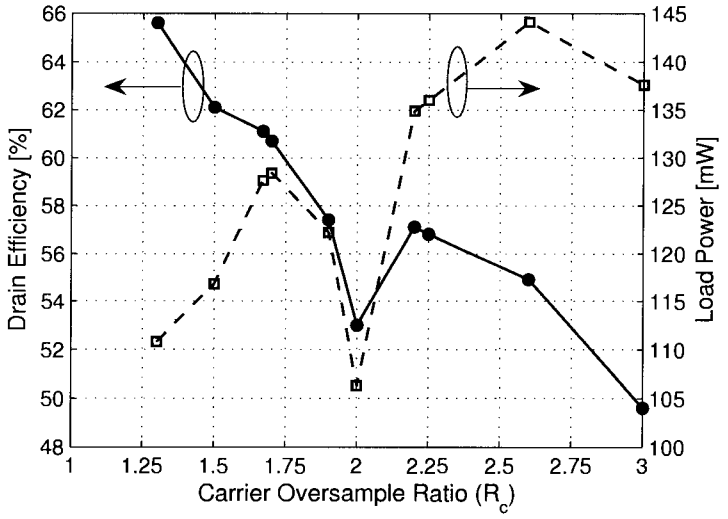


Figure 3.6: Power efficiency and load power for a 1 mm pHEMT as a function of carrier oversample ratio with $f_c = 500$ MHz. The coding efficiency and ATF correspond to a SNR threshold of 65 dB ($R_e = 128$) for a W-CDMA source signal.

Table 3.3: Design Values for CVSCD Amplifier with pHEMT Switch

	Sq. Wave	ΣΔ tone	ΣΔ W-CDMA	ΣΔ W-CDMA
PAR [dB]	0	0	10	7.1
σ_u [dBV]	-	-2.5	-11.2	-8.3
η_p [%]	$8/\pi^2$	41.9	5.7	11.0
f_{av}/f_c	2	1.937	1.706	1.706
UM_i	0.8	0.5	1	1
UM_v	0.9	0.9	0.9	0.9
R_L [Ω]	24.9	28.9	15.9	15.9

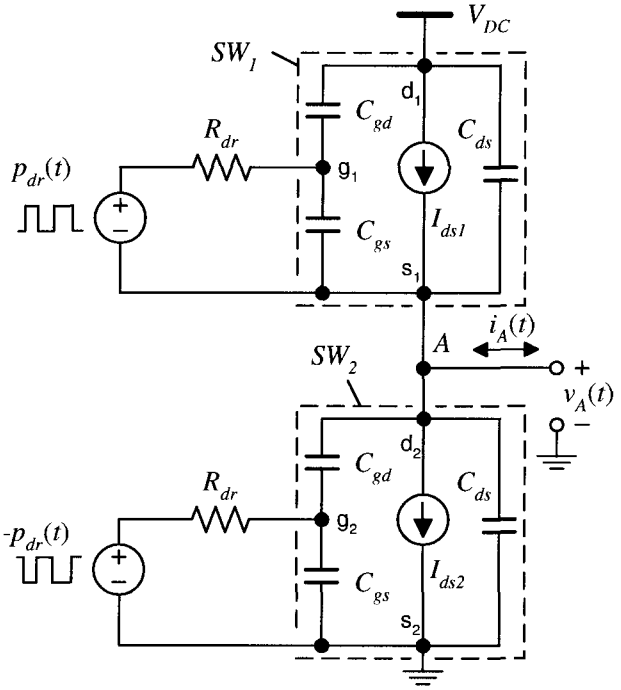


Figure 3.7: CVSCD circuit model for analysis and simulation.

3.4.3 Reconstruction Filter

After the optimum load is calculated, the remaining step in the design procedure is to calculate the filter component values. The design is evaluated at a carrier frequency of 500 MHz. For modulator pulse trains, a 10MHz sixth order Butterworth reconstruction filter is used. The filter is designed to match the nominal on resistance, 2 ohms for the pHEMT switch, to the optimum load resistance. In the case of the square wave pulse train, a relaxed filter specification is adequate, because the spectra is discrete, and a series LC resonator is sufficient.

3.4.4 Analytic and Simulated Results

The comparison between analytic and simulated results is made with the circuit model shown in Fig. 3.7. The simulation circuit is similar to the circuit in Fig. 3.4 except that the capacitances are linear, the current sources are nonlinear, and the circuit no longer switches instantaneously.

The voltage controlled current sources in Fig. 3.7 are modeled by

$$I_{ds} = I_{pk}\left[1 + \tanh\left(P_1(V_{gx} - V_{pk}) + P_3(V_{gx} - V_{pk})^3\right)\right](1 + \lambda V_{ds})\tanh(\alpha V_{ds}) \quad (3.30)$$

where

$$V_{gx} = \begin{cases} V_{gs} & V_{ds} \geq 0 \\ V_{gd} & V_{ds} < 0 \end{cases} . \quad (3.31)$$

Equation (3.30) is a modified Angelov model and generates the IV curve shown earlier in Fig. 3.3. The drivers in Fig. 3.7 are modeled as voltage sources and generate a pulse train $p_{dr}(t)$ with amplitude levels of $V_{gs,on}$ and $V_{gs,off}$. A normalized driver resistance of 5 ohms is assumed and corresponds to a device with a gate width which is 40% of the switch gate width.

The analytic and simulated results for the circuit in Fig. 3.7 are compared in Table 3.4 for three different drive signals at a carrier frequency of 500 MHz and modulator sample rate of 1700 MHz. For each source signal, the simulated and analytic results (shown in brackets) are tabulated for two model levels. The first model level includes only conduction losses (P_c) and the capacitances are omitted, while the second model level includes both conduction and capacitive switching (P_Q) losses. The total dissipated power in the switch, either P_c or $P_c + P_Q$, is shown in the column labeled P_{sw}. Most of the analytic results are within 10% of the simulated results, verifying the design methodology which incorporates the concepts of coding efficiency, signal PAR, and ATF.

The rise and fall times of the pulse train $p_{dr}(t)$ are important and affect power efficiency. In the simulation, switching times have a duration of $T/32$. If the rate is reduced, then the simultaneous conduction of the switches during amplitude transitions starts to generate a short circuit current path. The short circuit current generates a loss, called shoot-through loss, that reduces power efficiency. It is noted, that if the instantaneous switching model in Fig. 3.1 is simulated, the conduction losses, load power, and efficiency are within a few percent of the analytic values. Therefore, the switching transition time of the pulse train is significant, and fast edges are required for good power efficiency.

Another significant factor which controls the accuracy of the design equations is the current margin. The allocation of current margin depends on the probability distribution of the current distribution at switching intervals, since the peak switch current is limited to I_{on}, and during a transition, capacitive switching current and load current compete. The

Table 3.4: Simulated and (Analytic) Results for a pHEMT CVSCD Amplifier

Source	Configuration	P_{sw} [mW]	P_Q [mW]	P_L [mW]	η_d [%]
Square wave	P_c	101.4 (102.4)	-	1297.0 (1272.7)	92.8 (92.6)
	$P_c + P_Q$	172.6 (180.7)	71.2 (78.3)	1295.0 (1272.7)	88.5 (87.6)
$\Sigma\Delta$ tone	P_c	43.2 (40.0)	-	589.7 (578.2)	93.2 (93.5)
	$P_c + P_Q$	117.1 (115.8)	73.9 (75.8)	554.0 (578.2)	83.1 (83.3)
$\Sigma\Delta$ W-CDMA	P_c	23.8 (16.0)	-	127.9 (127.6)	84.3 (88.9)
	$P_c + P_Q$	87.2 (82.8)	63.4 (66.8)	115.2 (127.6)	57.0 (60.6)

probability of large current swings near $\pm I_{L,pk}$ is much larger with a sinusoidal load signal than with a W-CDMA signal. For this reason, the current margin is set to 0.5 for the design with the tone source, while no margin is assigned for the W-CDMA source as shown in Table 3.3. Clearly the case of the W-CDMA source signal is more interesting, because sinusoidal load signals are generated more efficiently without encoding, and the simulation results validate the selected margins.

In Fig. 3.8, power spectral density plots are shown for signals in the W-CDMA simulation. The source $u(t)$, the modulator pulse train $p(t)$, and the load signal $v_L(t)$ are shown. The power normalizations used to generate the overlays were calculated directly from the analytic equations and compare well. The data also shows the degradation in SNR of the source signal, first by the quantization in the modulator, and then by distortion in the amplifier. The reconstructed SNR is degraded and shows that distortion cancellation techniques may be required to improve SNR. For example, predistortion techniques could be used, or the power amplifier could be configured in a closed loop with the modulator.

3.4.5 Power Efficiency

The analytic equations for power efficiency are very useful for initial estimates of performance and are easily applied to explore design trade-offs. Two trade-offs are explored in the following sections: 1) a reduction in PAR assuming some kind of crest factor reduction method, and 2) MESFET versus pHEMT switches.

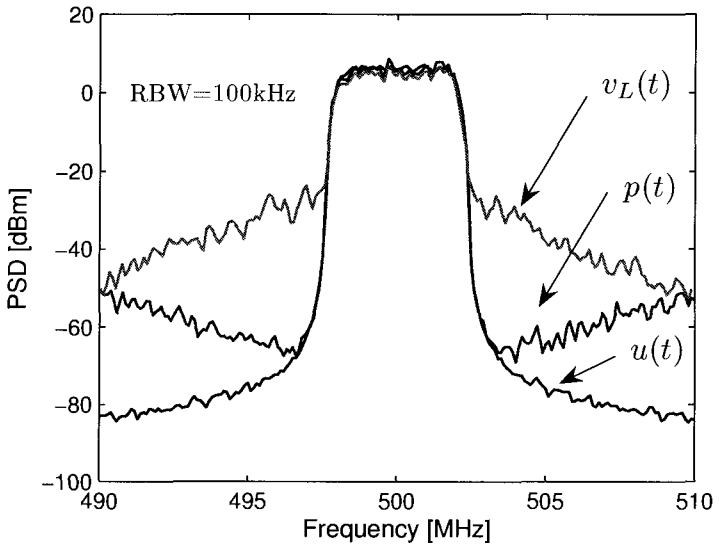


Figure 3.8: PSD of the W-CDMA source $u(t)$, modulator pulse train $p(t)$, and the load signal $v_L(t)$.

3.4.5.1 Low Peak-to-Average Power Ratio W-CDMA

There has been significant work in reducing the PAR of a W-CDMA signal envelope through crest factor reduction techniques to improve the power efficiency of linear amplifiers [60]. In a similar way, the power efficiency of the bandpass $\Sigma\Delta$ modulator class D configuration also improves from reductions in envelope PAR. By reducing the source PAR, the modulator drive level is increased, and coding efficiency is improved as explained in section 2.4.4. W-CDMA with PAR values as low as 7.1 dB have been reported [60], and if this source signal is used, coding efficiency is boosted from 5.7% to 11%. The boost in coding efficiency increases the reconstructed load power, and since switching losses are approximately constant, there is an overall increase in power efficiency. Fig. 3.9 shows the predicted efficiency response for the low PAR signal. The plot includes a reconstruction filter insertion loss of 0.5 dB. At a frequency of 500 MHz, the power efficiency is boosted by nearly 10% from 54.1% to 63.9% when the PAR of the W-CDMA source signal is reduced from 10 dB to 7.1 dB.

3.4.5.2 Comparison of pHEMT and MESFET Switches

In section 3.3, it was noted that capacitive switching losses are dominated by the off state gate to drain capacitance. A normalized value of 0.3 pF/mm has been used for pHEMT

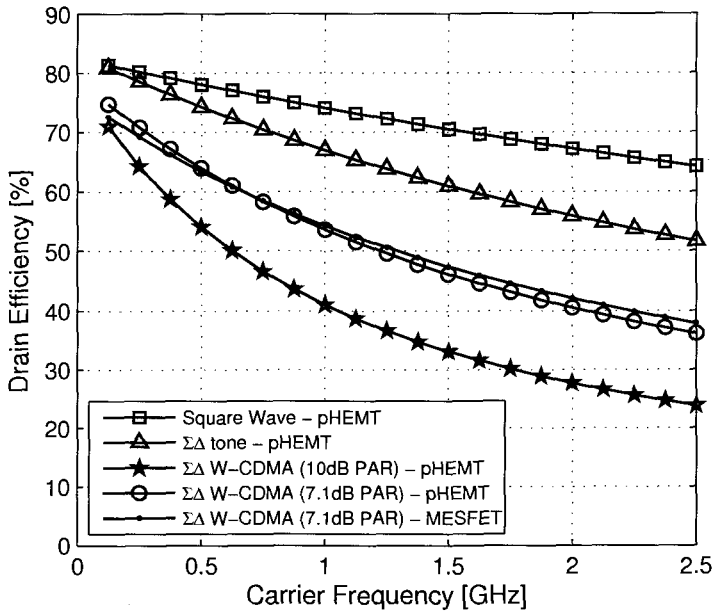


Figure 3.9: Drain efficiency including conductive losses, switching losses, and filter insertion loss (0.5 dB) with a modulator pulse train ($R_c = 1.7$).

switches, but a smaller value is typical of MESFET switches. For example, Wei reports on a 2 mm MESFET [61], and a normalized $\bar{C}_{gd,off}$ of approximately 0.125 pF/mm is obtained for this device. The gate-to-source capacitance is significantly larger than C_{gd} and is approximately 1 pF/mm; the normalized on resistance is also larger than a pHEMT and approximately 4 Ωmm for $V_{gs,on} = 0.4$ V. For comparison, Fig. 3.9 includes the predicted efficiency for a MESFET with values tabulated in Table 3.2. The advantage of the MESFET's lower $C_{gd,off}$ becomes apparent at high frequencies where the power efficiency is slightly better than the pHEMT. However, at low frequencies, the conduction losses dominate and the power efficiency is less than the pHEMT. The efficiencies of the two devices cross over at a carrier frequency of about 750 MHz.

The data in Fig. 3.9 are useful for making preliminary estimates on drain efficiency for potential W-CDMA applications. There are three main bands of frequencies specified in the W-CDMA standard centered around 850 MHz, 1.8 GHz, and 2.1 GHz. At these frequencies, assuming a PAR of 7.1 dB and MESFET devices, the corresponding power efficiency is approximately 56%, 42%, and 41%, respectively.

3.5 CMOS Amplifier Design with Driver

A disadvantage of using two depletion mode n-channel devices in the CVSCD configuration is the construction of a driver for the upper switch in Fig. 3.7. The source of SW_1 swings with the switched voltage at node A, and the driver pulse train $p_{dr}(t)$ must track this node voltage to turn the device on and off. Two possible ways to drive the upper device include a full swing driver that matches the required gate swing with the source voltage, or a bootstrap [62] or level-shifter [8] circuit which has a floating reference to track the voltage at node A.

The design of a driver for the CVSCD is greatly simplified if low voltage complementary devices are employed in the circuit. For this reason, CMOS technology has significant advantages for the circuit topology. An example of a design in 0.5 μm CMOS is shown next including the driver stage. Although 0.5 μm is not state of the art technology, public technology parameters are available for this technology through MOSIS [63] for analytic calculations. The results of the analysis are confirmed with SPICE simulations based on a commercial foundry model. The SPICE simulations were run by a colleague, Robert Sobot, using proprietary BSIM3 models for IBM's 5HP SiGe BiCMOS process.

3.5.1 CMOS Amplifier Overview

A CMOS amplifier design with six stages is shown in Fig. 3.10. The first five stages of the amplifier are the driver stages. The first driver stage consists of a 0.5 μm $\times$ 10 μm NMOS device (N_1) with a 0.5 μm $\times$ 20 μm PMOS device (P_1). Subsequent stages are scaled by a factor of three relative to the preceding stage to minimize propagation delay, similar to device scaling in digital CMOS design for output buffers [64, pg. 267]. The last driver stage (N_5 and P_5) provides enough current to drive the gate capacitance of the switches (N_6 and P_6) in the class D output stage.

The six stages generate approximately 100 mW of output power at 181 MHz with a 3.3 V DC supply voltage (V_{DD}). A sixth order Butterworth bandpass filter with a 3 dB bandwidth of 10 MHz is used for reconstruction. The component values for the filter are given in Table 3.5, and assume a load impedance of 2.0 ohms. Matching to 50 ohms is not considered in this design.

The encoded pulse train $p(t)$ in this design is synthesized by a fourth order continuous-time modulator called modulator D (see section 2.2.4). The carrier oversample ratio is 2.21

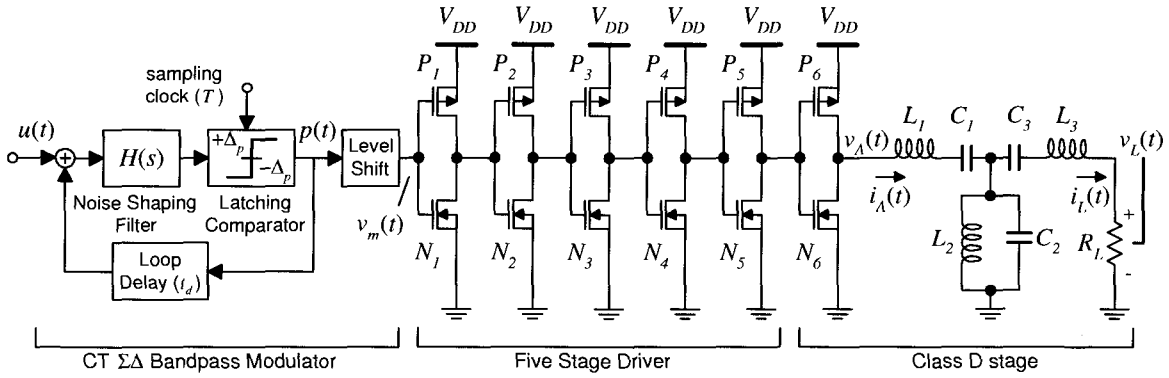


Figure 3.10: CMOS driver and class D amplifier.

Table 3.5: Bandpass Filter Values (f_c = 181 MHz, BW = 10 MHz)

L_1	31.83nH	C_1	24.29pF
L_2	48.58pH	C_2	15.92nF
L_3	31.83nH	C_3	24.29pF

with a center frequency of 181 MHz, and corresponds to test conditions that Sobot used in an implementation of this modulator [65]. Simulations were made with both a two tone and W-CDMA source signal. The corresponding coding efficiency and ATF's for the pulse trains are given in Table 3.6.

Examples of the circuit signals are shown in Fig. 3.11 for a modulator pulse train with a two tone source signal. The level shifted modulator pulse train is called $v_m(t)$, and the switched voltage pulse train at the output of the amplifier (node A) is shown as $v_A(t)$. The delay between $v_A(t)$ and $v_m(t)$ is approximately 700 ps. The switched voltage pulse train at node A also has distortion caused by non-zero switch resistance. The load current $i_A(t)$ creates a voltage drop across the switch and superimposes the reconstructed load signal on

Table 3.6: Bandpass ΣΔ Modulator Parameters for $R_c = 2.21$

Source	PAR [dB]	f_c [MHz]	η_p [%]	f_{av} [MHz]
Two tone	3.0	181	25.04	397.8
W-CDMA	8.7	181	6.58	398.9

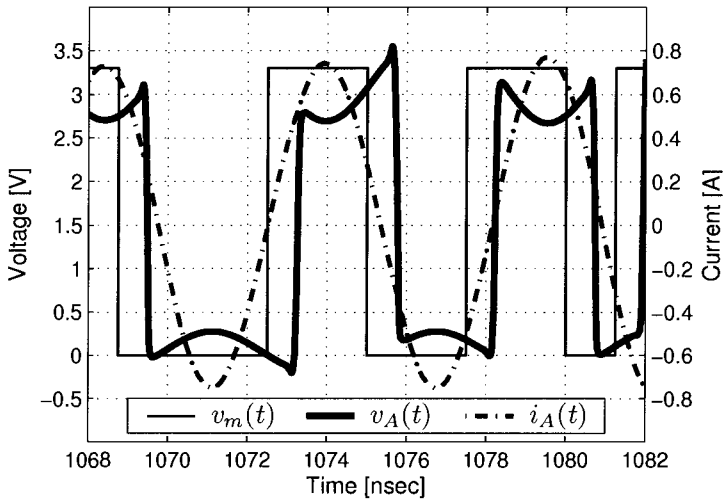


Figure 3.11: CMOS amplifier signals with a two tone source: $v_m(t)$ the level shifted modulator pulse train; $v_A(t)$ the class D switched voltage pulse train; and $i_A(t)$ the load current signal into the bandpass filter.

the switched voltage pulse train.

The corresponding power spectrum for the reconstructed load signal $v_L(t)$ is shown in Fig. 3.12. The spectrum shows an intermodulation distortion level of approximately -40 dBc. A second PSD plot for the W-CDMA source signal is shown in Fig. 3.13, and the SNR of the W-CDMA signal is approximately 37 dB. The figures also include an overlay of the modulator pulse train $p(t)$ to show the distortion introduced by the driver and class D amplifier stages.

3.5.2 Switch Design

Since load current flows in either direction during a transition, the drain current of N_6 and P_6 can have either polarity. Similar to pHEMT and MESFET switches, MOSFET switch current is continuous even under reverse bias drain conditions. Negative drain current is limited to a small reverse bias range, and the voltage drop under reverse bias must be designed to prevent the substrate-drain junction from conducting. In the circuit shown in Fig. 3.10, the substrate is biased to the source potential; additional protection for reverse bias voltage swings can be provided by applying a separate substrate voltage. A gate oxide breakdown voltage of 4 V is assumed, and has been cited by others [66] as a conservative

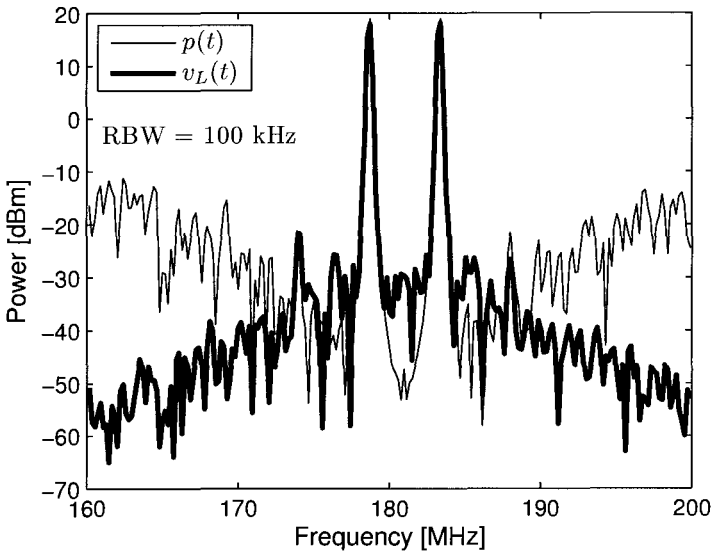


Figure 3.12: PSD of a reconstructed load signal with a two tone source signal.

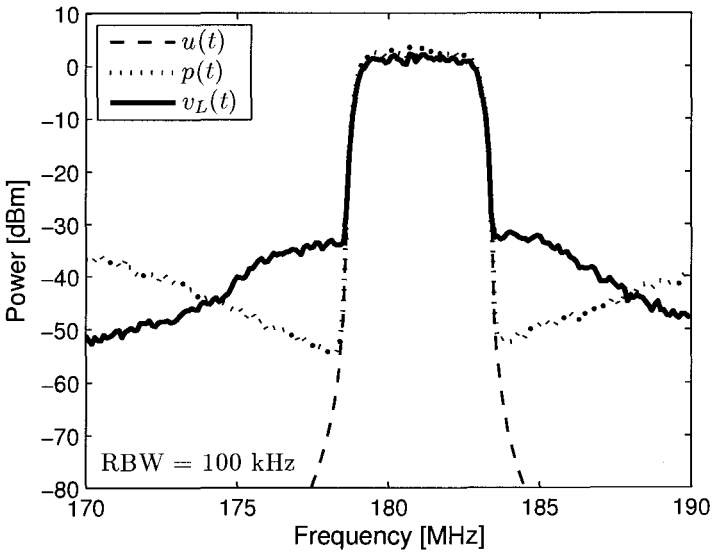


Figure 3.13: PSD of a reconstructed load signal with a W-CDMA source signal.

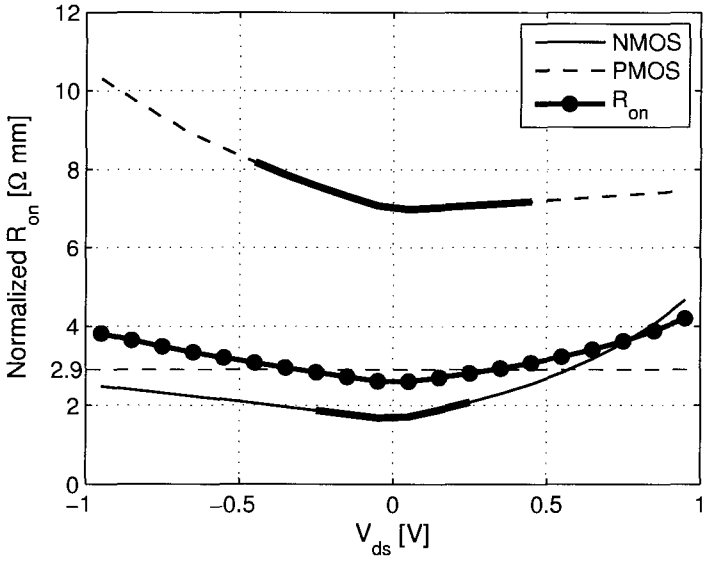


Figure 3.14: Channel resistance for 0.5 μm $\times$ 1 mm NMOS and PMOS devices. The effective switch resistance for a 2 mm PMOS and 1 mm NMOS are labeled as R_{on}.

estimate of maximum voltages in a 3.3 V CMOS process with a typical oxide thickness of 9.5 nm. The corresponding voltage margin is then 0.825 for the design.

When N_6 or P_6 is on, the channel resistance is nonlinear, and the load current swing is restricted to a range where the resistance variation is within an acceptable range. Figure 3.14 shows the normalized on resistance, $\bar{R}_{on}$, for a NMOS and PMOS device with 1 mm of gate width. The resistance characteristics are obtained by differentiating the IV device characteristics when they are turned on ($|V_{gs}| = 3.3$ V).

The PMOS device has a much larger channel resistance, and the PMOS gate width is scaled relative to the NMOS device width to match on resistances. In the design, a PMOS to NMOS gate width of two is used. The gate ratio provides a balance between DC on resistance characteristics, as well as equalizing the rise and fall times under dynamic switching conditions. A third curve in Fig. 3.14 shows the effective switch resistance R_{on} for a scaled switch pair assuming each device is active for half the time. A dashed line at 2.9 Ω-mm corresponds to a nominal linear on resistance used for power loss calculations.

With a 2 Ω load in the amplifier design, the peak current swings over the range shown by the heavy lines in Fig. 3.14. Even with the device scaling factor of two, the PMOS channel resistance is larger than the NMOS channel resistance, and the voltage swing across the

Table 3.7: CMOS Capacitance Equations

Description	Parameter	Equation	NMOS	PMOS
Gate Length	L_{eff}	$L - 2\text{LINT}$	0.324μm	0.284μm
Gate Capacitance	C_g	$C_{ox}L_{eff} + \text{CGSO}$	1.579 fF/μm	1.439 fF/μm
Gate-drain Capacitance	C_{gd}	CGDO	0.356 fF/μm	0.365 fF/μm
Drain Capacitance	C_d	$C_{jb} + C_{jswg}$	0.906 fF/μm	0.950 fF/μm

PMOS switch is larger than the NMOS switch. The current utilization margin shown by the heavy lines is 0.37 for the NMOS device and 0.59 for the PMOS device. Maximum normalized switch currents of $\bar{I}_{on,n} = 0.384$ A/mm and $\bar{I}_{on,p} = -0.120$ A/mm at $|V_{ds} = 1|$ V and $|V_{gs} = 3.3|$ V are assumed for the current utilization margin, and correspond to the x-axis limits of Fig. 3.14.

3.5.3 Capacitive Switching Losses

Capacitive switching losses are significant in the CMOS amplifier. Similar to the pHEMT switch analysis, energy is dissipated each time the voltage changes across parasitic capacitance during a level transition in the pulse train. The capacitance losses for the driver and the class D RF stage are analyzed separately in anticipation of calculating drain efficiency for the class D stage only, and overall power efficiency which includes both driver and amplifier losses.

A first order physical capacitor model is used for analyzing switching losses in the circuit [67]. The analytic model includes gate oxide capacitance, gate-drain overlap capacitance, and drain/source diffusion capacitances. The capacitance model equations are summarized in Table 3.7, and capacitance values are derived from a selected set of BSIM3 model parameters given in Table 3.8. The complete BSIM3 model provided by MOSIS [63] is given in Appendix B.

The nonlinear junction capacitances C_{jb} and C_{jswg} are significant. The equivalent linear capacitance given in Table 3.7 is calculated assuming a voltage transition of V_{DD} and includes the bottom and gate sidewall junction areas. For the calculations, the drain diffusion area is 1.4 μm $\times$ 10 μm for a 10 μm cell, and the corresponding bottom area of the junction is $14\,\mu\text{m}^2$.

Using the normalized capacitance values in Table 3.7, the total parasitic capacitance for the driver and the PA are calculated. For the calculation, C_{gd} is modeled as a Miller

Table 3.8: CMOS Capacitance Model Values

Parameter	NMOS	PMOS
TOX [nm]	9.3	9.3
LINT [μm]	0.088	0.108
CGDO [fF/μm]	0.356	0.365
CGSO [fF/μm]	0.356	0.365
CJ [fF/μm^2]	0.871	0.844
PB [V]	0.8	0.6
MJ	0.38	0.204
CJSWG [fF/μm]	0.164	0.064
PBSWG [V]	0.8	0.7
MJSWG	0.91	0.99

capacitance and contributes $2C_{gd}$ to the gate and drain terminals. The sum of the effective capacitance at the driver drain nodes N_1 through N_5 is 60.2pF, and the total class D capacitance at node A is 24.2pF. These capacitance values, with the corresponding ATF's in Table 3.6, yield the driver ($P_{Q,dr}$) and class D ($P_{Q,pa}$) switching losses shown in Table 3.9.

3.5.4 Power Efficiency

Expressions for load power, conduction loss, driver and class D switching loss, and DC power are combined to predict the drain efficiency and overall efficiency of the amplifier. Table 3.9 summarizes the analytic and simulated results. The simulated results use a foundry BSIM3 model, while the analysis results use a simplified model based on extracted model parameters from MOSIS.

The estimated drain efficiency (η_d) of the amplifier is 64.2% for the two tone source and 43.7% with the W-CDMA source, compared to simulated results of 55.4% and 31.7%, respectively. The overall efficiency of the amplifier (η_a) includes driver power losses, and was defined in Chapter 1, equation (1.3). The corresponding overall power efficiency for the two tone and W-CDMA sources are 45.5% and 21.1%, compared with simulated results of 40.1% and 16.6%. The analytic results are optimistic and obviously do not consider all the complexity in a BSIM3 model. On the other hand, the simplified equations provide an initial estimate of performance, and confirm the proposed design methodology for making an initial CMOS design.

Table 3.9: Analytic and Simulated Results for a RF CMOS Amplifier

Parameter	Two Tone		W-CDMA	
	Analytic	Simulated	Analytic	Simulated
$i_{L,rms}$ [A]	0.318	0.330	0.163	0.162
$i_{L,pk}$ [A]	0.635	0.704	0.628	0.626
P_L [mW]	202.2	217.3	53.2	52.8
P_c [mW]	60.3	-	15.9	-
$P_{Q,pa}$ [mW]	52.4	-	52.6	-
$P_{DC,pa}$ [mW]	314.9	392.5	121.6	166.5
η_d [%]	64.2	55.4	43.7	31.7
$P_{Q,dr}$ [mW]	129.6	149.1	130.0	152.1
η_a [%]	45.5	40.1	21.1	16.6

3.6 Conclusions

An analysis of a bandpass $\Sigma\Delta$ modulator driven class D RF amplifier has been presented for the complementary voltage switched configuration. First order design equations were developed for calculating optimum load resistance, load power, and power efficiency. As a link between the amplifier design and the modulator design, coding efficiency and average transition frequency parameters were incorporated into the design equations. Both coding efficiency and average transition frequency depend on the carrier oversample ratio of the modulator, and the choice of the oversample ratio impacts the amplifier output power and power efficiency. The analysis was confirmed with simulations of a typical pHEMT device and showed a drain efficiency of over 50% with W-CDMA at 500 MHz. An example of a CMOS class D RF amplifier with a driver was also presented. Improvements in amplifier power efficiency can be expected if modulator coding efficiency is improved, or if better devices with low on resistance, low gate-to-drain capacitance, or high breakdown voltages are used for the switches.

Chapter 4

Coding Efficiency of a Periodic Signal Model

The power efficiency analysis of the RF class D amplifier demonstrates the importance of coding efficiency. It has been shown that the pulse train coding efficiency has a direct relationship with the scaling of the devices in the class D circuit for a given load power. Since coding efficiency is so important, a relevant question is: what are the coding efficiency limits for a binary amplitude pulse train with constrained zero-crossings? A coding efficiency bound would show how much margin is available if better encoding methods were used. It appears that the answer to this question is quite difficult for general source signals, and in this chapter we consider the problem of encoding a sinusoidal source signal into a pulse train.

A bandpass $\Sigma\Delta$ modulator synthesizes a pulse train with synchronous zero-crossings that are determined by the sample period T. Since the timing of zero-crossings are constrained, a sequence of L bits, $y[0] \ldots y[L-1]$, generates a finite number of reconstructed signals. We consider the periodic extension of the L bit sequence, and construct a Fourier series expansion for the continuous-time pulse train corresponding to the sequence. The reconstructed signal space, defined as the amplitude and phase of a harmonic component at f_c, is then examined as a function of L.

It is shown that the reconstructed signal constellation is bounded by a regular n-sided polygon for rational carrier oversample ratios. Within the bounding polygon, the SNR is determined by the length of the sequence L, and SNR can be large by choosing L to be sufficiently large. There is also a circular region within the constellation, defined by a maximum amplitude r_{max}, where the reconstructed signal is phase insensitive. The value of r_{max} varies with carrier oversample ratio, and the periodic model shows that coding

efficiency does not vary monotonically with an increase in carrier oversample ratio. The model shows dips in coding efficiency at specific carrier oversample ratios, similar to the dips shown in Chapter 2 for bandpass $\Sigma\Delta$ modulators.

Although the model is constrained to the case where the source and load frequency are commensurate with the sample rate, the model shows that the frequency error for non-commensurate frequencies is not uniformly distributed for a fixed sequence length. The frequency resolution around carrier oversample ratios such as $R_c = 2$ is much coarser than, for example, a sample ratio of 1.7. By attempting to mimic the synthesis of a periodic load signal, and imposing a constraint on zero-crossings as in a bandpass $\Sigma\Delta$ modulator, insight is gained into why large amplitude behavior can exhibit variation with carrier oversample ratio. The model also shows that the maximum amplitude of the reconstructed signal constellation is bounded by $(4/\pi)\Delta_p$ as $L \rightarrow \infty$ for a pulse train with amplitude levels of $\Delta_p = \pm 1$.

4.1 Periodic Binary Amplitude Pulse Trains

A Fourier series (FS) expansion is calculated for a binary level periodic signal $p(t)$ which has a period of LT where L is a positive integer and T is the clock period. The construction of $p(t)$ is constrained to be similar to a bandpass $\Sigma\Delta$ modulator pulse train. The pulse train has only two amplitude levels ± 1, and zero-crossings at integer multiples of T as shown in Fig. 4.1. After calculating the FS for $p(t)$, we map the reconstruction signal space (amplitude and phase) for sequences with a harmonic component corresponding to a carrier oversample ratio of R_c.

4.1.1 Fourier Series

Let f be a family of functions synthesized by an L bit sequence from a zero-order hold digital-to-analog converter. The ZOH DAC pulse is defined as $d(t)$

$$d(t) = \begin{cases} 1 & 0 \leq t \leq T \\ 0 & \text{otherwise} \end{cases} \tag{4.1}$$

where T is the clock period and $T > 0$. The functions are defined as

$$f : t \rightarrow \sum_{k=0}^{L-1} b_k \, d(t - kT) \tag{4.2}$$

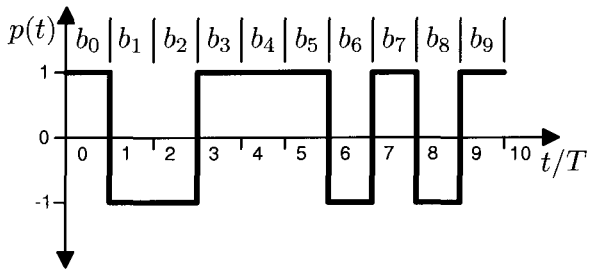


Figure 4.1: Binary level pulse train with amplitude levels of ± 1 and zero-crossings at integer multiples of T.

over the interval $[0, LT]$ and the bits are binary with values $b_k \in \{-1, +1\}$.

For L bits there are 2^L possible functions and these functions are expanded in a Fourier series. Let

$$\begin{aligned} p(t) &= \sum_{n=-\infty}^{+\infty} f(t - nT) = \sum_{n=-\infty}^{\infty} c_n e^{j\omega_p nt} \\ &= \frac{r_0}{2} + \sum_{n=1}^{\infty} r_n \cos(\omega_p nt + \phi_n) \end{aligned} \tag{4.3}$$

where:

$$c_n = \frac{r_n}{2} e^{j\phi_n} \tag{4.4}$$

$$\omega_p = 2\pi/(LT) \tag{4.5}$$

$$c_n = \frac{1}{LT} \int_0^{LT} f(t) e^{-j\omega_p nt}\, dt \quad n = 0, \pm 1, \pm 2, \ldots \tag{4.6}$$

After evaluating (4.6) with (4.1) and (4.2), the coefficients are:

$$c_n = \frac{\sin(n\pi/L)}{n\pi} e^{-jn\pi/L} \sum_{k=0}^{L-1} b_k\, e^{-j2\pi nk/L}. \tag{4.7}$$

A change of variables is made to express the coefficients in terms of the carrier oversample ratio R_c. Assume we want to synthesize a periodic pulse train with a harmonic frequency component at $f_c = 1/(2R_cT)$. Since the clock period is fixed, f_c must be a fractional ratio of the clock frequency f_s, and R_c is a rational number. Let R_c be defined as

$$R_c = M/(2N) \tag{4.8}$$

where M and N are positive integers, then $f_c = (N/M)f_s$.

Using (4.5), the frequency of the nth harmonic is $(n/L)f_s$; therefore, any sequence of length $L = mM$, and corresponding harmonic $n = mN$ where m is a positive integer, has a frequency component that satisfies (4.8). A sequence length of $L = M$ is the shortest sequence satisfying (4.8). Longer sequences, whose length is a multiple of M, $2M, 3M, \ldots$, also have frequency components at $f_c = (N/M)f_s$ corresponding to harmonics $n = 2N, 3N, \ldots$, respectively.

The FS coefficient c_n in (4.7) is rewritten as c_{mN}

$$c_{mN} = \text{sinc}[1/(2R_c)]\, W^{N/2} \frac{1}{mM} \sum_{k=0}^{mM-1} b_k\, W^{kN} \tag{4.9}$$

where $W = e^{-j2\pi/M}$, the Mth root of unity.

Since W is periodic over M bits, (4.9) is written as a double sum:

$$c_{mN} = \text{sinc}[1/(2R_c)]\, W^{N/2} \frac{1}{M} \sum_{k=0}^{M-1} W^k \frac{1}{m} \sum_{l=0}^{m-1} \tilde{b}_{k+lM}. \tag{4.10}$$

The bit sequence $\{\tilde{b}_0, \tilde{b}_1, \ldots, \tilde{b}_{L-1}\}$ corresponds to a reordered bit sequence of $\{b_0, b_1, \ldots, b_{L-1}\}$ and is consistent with the expansion of (4.9) into a double sum. Let

$$a_k = \frac{1}{m} \sum_{l=0}^{m-1} \tilde{b}_{k+lM} \tag{4.11}$$

and

$$g_M = \frac{1}{M} \sum_{k=0}^{M-1} a_k W^k \tag{4.12}$$

then (4.10) is finally written as:

$$c_{mN} = \text{sinc}[1/(2R_c)] \times W^{N/2} \times g_M. \tag{4.13}$$

Three terms are identified in (4.10). The first term $\text{sinc}[1/(2R_c)]$ corresponds to the ZOH response of a DAC with a hold period of T. The DAC attenuation is expressed in terms of the carrier oversample ratio R_c, and shows that the attenuation is reduced as the carrier oversample ratio is increased. The second term is a constant phase rotation $W^{N/2}$ and has no impact on the magnitude of the harmonic. The third term g_M is a weighted sum

of the M roots of unity and shows how the bit sequence affects the amplitude and phase of the FS coefficient. The weights a_k are bounded and take on rational values in the interval $[-1, +1]$. The resolution of a_k depends on the sequence length factor m as shown in (4.11).

Since our goal is to compare the periodic signal model with bandpass $\Sigma\Delta$ modulation, the range of R_c is restricted to values where f_c is less than the Nyquist rate $f_s/2$. The modulators are sampled systems and aliasing is avoided by restricting $f_c < f_s/2$. A subsampling modulator design could also be implemented if the input is a bandlimited bandpass signal, but the coding efficiency is less than an oversampled design. In an oversampled design, R_c is restricted to a range greater than 1 which implies M and N satisfy $M > 2N$.

4.1.2 Signal Reconstruction Space Examples

In this section, the amplitude and phase of the FS coefficients c_{mN} are systematically evaluated for all combinations of bits in a sequence of length mM. Equation (4.13) is written in terms of the carrier oversample ratio variable R_c and provides a way to compare the periodic pulse train model with bandpass $\Sigma\Delta$ modulation. In the model, the harmonic at f_c is the reconstructed signal $\hat{u}(t)$, and the bits are selected to minimize the reconstruction error for a source signal $u(t) = r\cos(2\pi f_c t + \phi)$. The reconstruction error $|u(t) - \hat{u}(t)|$ then depends on how close the amplitude and phase of the reconstructed signal match the source signal.

The term signal reconstruction space describes the set of all possible amplitudes and phases generated by a bit sequence of length mM at a frequency f_c. As subsequent examples demonstrate, the reconstruction space for a periodic pulse train $p(t)$ is easily visualized as a constellation of discrete reconstruction points plotted in the complex plane. The examples also show:

1. the signal constellations lie in an equal sided polygon and the number of edges depends on M;

2. the density (resolution) in the constellation depends on m, and as m gets large, the resolution increases;

3. the length of the vertices that defines the signal constellation polygon depends on M and N.

The examples are arranged in terms of increasing M, and we begin with $M = 3$, as this is the shortest sequence satisfying $M > 2N$.

4.1.2.1 $M = 3$

For $M = 3$, there is one value of N, $N = 1$, which satisfies $M > 2N$. This configuration is analogous to a bandpass $\Sigma\Delta$ modulator with a noise notch at a center frequency of $f_s/3$ or $R_c = 1.5$.

In this case, $W = e^{-j2\pi/3}$ and there are three vectors, W^0, W^1, and W^2 as shown in Fig. 4.2(a), which establish the shape of the signal constellation. Let $\alpha_k = \cos[\pi k/(2M)]$ then

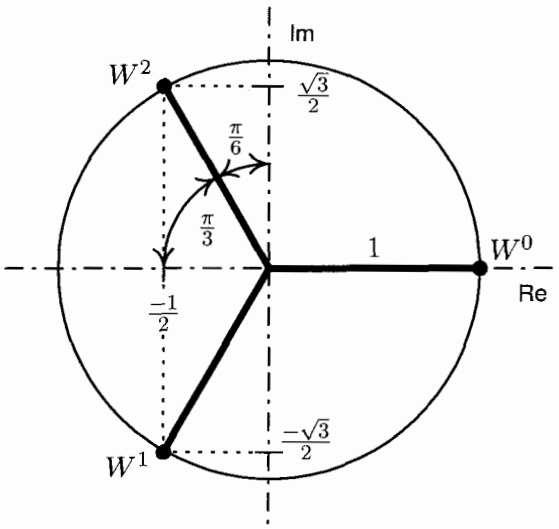
$$\begin{aligned} c_m &= \text{sinc}(1/3) \times W^{1/2} \times \frac{1}{3}(a_0 W^0 + a_1 W^1 + a_2 W^2) \\ &= \frac{\sqrt{3}}{2\pi} \times e^{-j\pi/3} \times [a_0\alpha_0 - (a_1 + a_2)\alpha_2 - j(a_1 - a_2)\alpha_1] \end{aligned} \tag{4.14}$$

When $m = 1$, and using (4.11), a_k has values of $\{-1, 1\}$. Evaluating the third term of (4.14) for all combinations of a_k, then yields a hexagonal constellation as shown in Fig. 4.2(b). The vertices of the hexagon are generated by all three bit sequences that have a zero-crossing, and the constellation point at the origin corresponds to a sequence where all the bits are the same sign (DC), hence no harmonic at f_c.

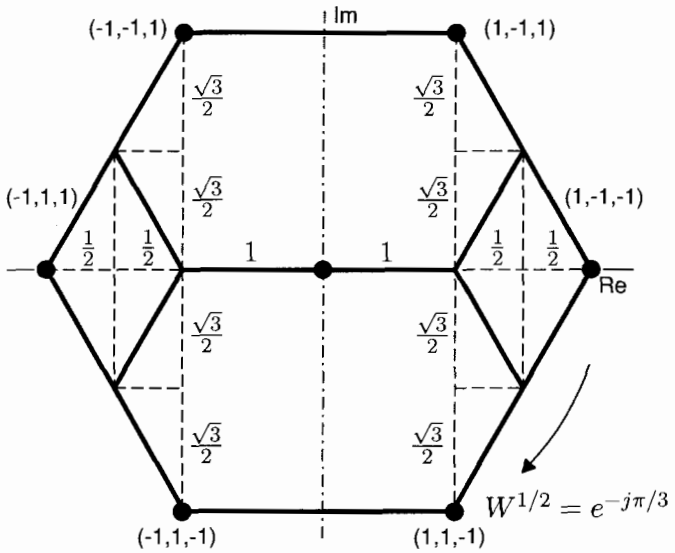
The symmetry of the bounding hexagon makes the polar form of the FS coefficient convenient. Adopting the notation (r_n, ϕ_n) for the FS coefficient c_n, we have $r_n = 2|c_n|$ and $\phi_n = \arg(c_n)$ from (4.4). Let r_v be the radius of the vertices of the bounding hexagon, then $r_v = \max(r_n)$. Since the bounding hexagon is invariant to the phase rotation term $e^{-j\pi/3}$ in (4.14), two vertices lie on the real axis as shown in Fig. 4.2(b). Choosing the vertex on the positive real axis yields the following expression for r_v:

$$r_v = \frac{\sqrt{3}}{\pi} \times (\alpha_0 + 2\alpha_2) = \frac{2\sqrt{3}}{\pi}. \tag{4.15}$$

Longer bit sequences generate additional reconstruction points in the constellation by increasing the resolution of the values a_k. For $m = 2$, the sequence is 6 bits long, and the values of a_k are from the set $\{-1, 0, 1\}$. For $m = 3$, the sequence is 9 bits long and the values of a_k are from the set $\{-1, -\frac{1}{3}, \frac{1}{3}, 1\}$. By induction, it is clear that the amplitude and phase resolution within the hexagonal signal constellation continues to grow as the sequence length increases. The increase in resolution is illustrated in Fig. 4.4 for m equal to 1, 2, and 3.



(a) Vectors W^k for $M = 3$.



(b) Addition of signal vectors with $M = 3$ and $m = 1$.

Figure 4.2: $M = 3$ signal constellation construction ($\alpha_1 = \sqrt{3}/2$; $\alpha_2 = 1/2$).

4.1.2.2 $M = 4$

When $M = 4$ and $N = 1$, the harmonic has a period of $4T$ or frequency of $f_s/4$. This corresponds to a bandpass modulator with $R_c = 2$ and is a classic modulator example which is generated by transforming a lowpass DT modulator with the $z \rightarrow -z^2$ transformation [31].

When $M = 4$, W^k takes on one of four values: ± 1 or $\pm j$. If $v_1 = 1$ and $v_2 = j$, then c_m is constructed from two orthogonal basis vectors and the bits in the sequence determine the weights of the two basis vectors. Evaluating the signal constellations for sequences lengths of 4, 8, and 12 bits yields the constellations shown in Fig. 4.4.

The constellation is enclosed within a square with vertices that have a distance of $4/\pi$ from the origin. The vertices correspond to the following ordered sets of a_k values: $(1, 1, -1, -1)$, $(-1, 1, 1, -1)$, $(-1, -1, 1, 1)$, and $(1, -1, -1, 1)$. Each set of values synthesizes a square wave pulse train with a 50% duty cycle, and the sets are generated by a cyclic shift of any one of the four sequences. We expand on this point later in the context of a computing a bound on the maximum coding efficiency for a periodic sequence.

Also, observe that the number of edges of the constellation polygon is equal to 4 and this equals the minimum sequence length M. In the previous example with M odd, the constellation was enclosed in a six sided polygon or $2M$. In general, if M is odd, then the polygon has $2M$ edges, and if M is even then the polygon has M edges.

4.1.2.3 $M = 5$

As a final example, the constellation for $M = 5$ is shown in Fig. 4.4. In this case, there are two values of N which satisfy $M > 2N$, and the example demonstrates that the only effect of N on the signal constellation is to scale the vertex length r_v.

With reference to Fig. 4.3, the real and imaginary components of vectors W^k are expressed in terms of $\alpha_k = \cos(k\pi/(2M))$. The term g_M defined in (4.12) is then

$$g_M = \frac{1}{5}\Big\{a_0\alpha_0 - (a_2 + a_3)\alpha_2 + (a_1 + a_5)\alpha_4 + j[-(a_1 - a_5)\alpha_1 - (a_2 - a_3)\alpha_3]\Big\} \tag{4.16}$$

and the vertex radius of the constellation is:

$$r_v = 2\,\mathrm{sinc}(1/(2R_c)) \times \max(|g_M|) \tag{4.17}$$

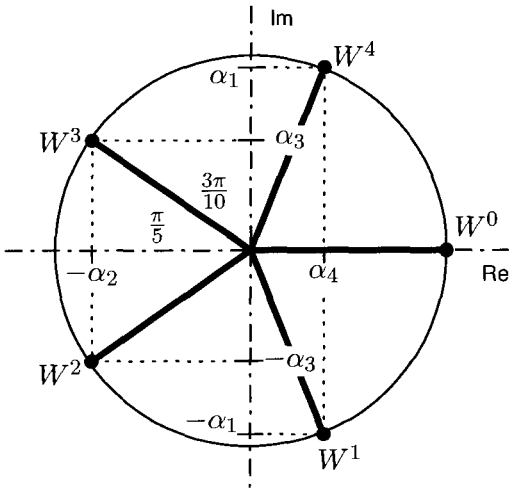


Figure 4.3: W^k vectors for $M = 5$.

There are ten different bit sequences which maximize $|g_M|$ and the bounding polygon for the signal constellation is a decagon. The vertices correspond to bit sequences which generate an ordered set $(a_0, a_1, a_2, a_3, a_4)$ which is a cyclic shift of the sequence $(1, 1, -1, -1, -1)$ or a cyclic shift of the same sequence with opposite sign: $(-1, -1, 1, 1, 1)$.

One of the vertices lies on the positive real axis, and choosing this vertex we get

$$r_v = \text{sinc}(1/(2R_c)) \times \frac{2}{5}(\alpha_0 + 2\alpha_2 + 2\alpha_4) \tag{4.18}$$

For $N = 1$, $r_v \approx 1.21$, and for $N = 2$ $r_v \approx 0.98$. The example shows how the term g_M determines the shape and number of points in the constellation, while the term $\text{sinc}(1/(2R_c))$ is a global scale factor and depends on both M and N.

4.1.2.4 General M

The preceding examples show some of the characteristics of reconstructed signals obtained from Fourier series expansions of periodic pulse trains. In general, for positive integers M and N where $M > 2N$, the reconstruction constellations have the following properties:

1. for even M the signal constellation is bounded by a polygon with M edges and each face has m points excluding the vertices,

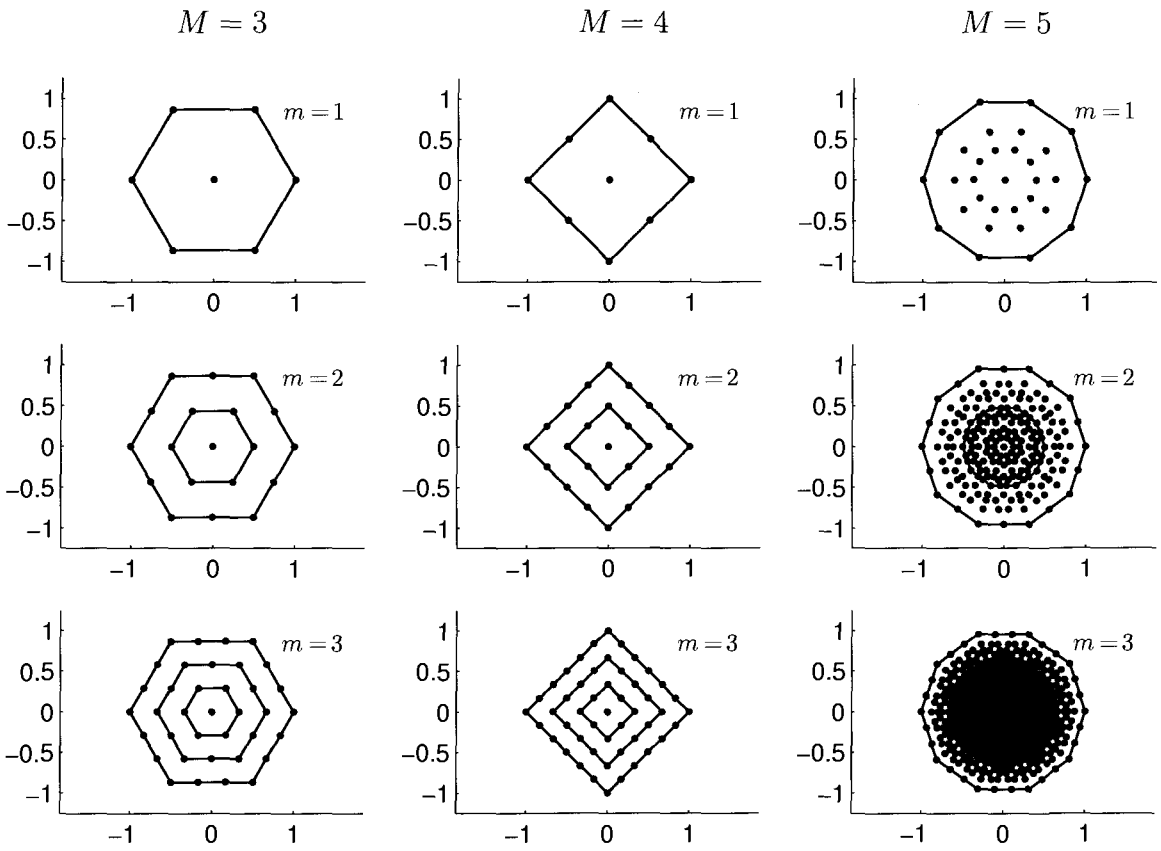


Figure 4.4: Normalized signal constellations for different values of M and m; the amplitude is normalized to r_v.

2. for odd M the signal constellation is bounded by a polygon with $2M$ edges and each face has $m-1$ points excluding the vertices,

3. for even M the maximum vertex distance r_v is $4/\pi$

4. for odd M the maximum vertex distance r_v is $4/\pi \cos[\pi/(2M)]$

5. the resolution of the reconstruction points in the signal constellation increases with the length of the sequence $L = mM$.

Points 1, 2, and 5 are derived by induction by continuing the examples to higher M. A proof of 3 and 4 is given in Appendix C.

4.2 Maximum Amplitude Boundary Model

The periodic signal model derived for binary pulse trains with a harmonic at f_c and carrier oversample ratio R_c is now used as a model for analyzing SNR and coding efficiency. The analysis so far has shown that the reconstructed signal lies within a bounded region defined by an equal sided polygon. The bounding polygon is called the maximum amplitude boundary, and within the boundary, the reconstructed SNR depends on the sequence length $L = mM$.

As shown in Fig. 4.5, a second bounding region is defined by a circle of radius r_{max} that is tangent to the normals of each edge. For source amplitude levels $r \leq r_{max}$, the reconstruction error is phase insensitive, and the reconstruction error $|u(t) - \hat{u}(t)|$ is made arbitrarily small for any source phase angle ϕ by choosing L to be sufficiently long. For amplitude levels which exceed r_{max}, the reconstruction error is phase sensitive and may or may not depend on the sequence length L. If the source vector (r, ϕ) falls within the maximum amplitude boundary, then the reconstruction error can be arbitrarily small. On the other hand, if (r, ϕ) falls outside the bounding polygon, the reconstruction becomes a function of the displacement between the maximum amplitude boundary and the source vector. The term overload shall be used to refer to the condition when (r, ϕ) falls outside the maximum amplitude boundary.

As a general model for subsequent analysis, a polygon with $2K$ edges is assumed for the maximum amplitude boundary as shown in Fig. 4.5. If M is even, then $K = M/2$; if M is odd, then $K = M$. The polygon is centered at the origin O of the complex plane, and the

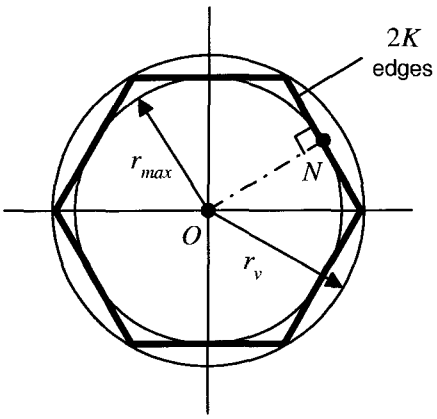


Figure 4.5: Maximum amplitude boundary for the periodic model.

vertices of the polygon lie on a circle with radius r_v. Each face of the polygon subtends an angle

$$\theta = \frac{2\pi}{2K} = \frac{\pi}{K} \tag{4.19}$$

and $r_{max} = r_v \cos(\theta/2)$.

4.3 Overload SNR

The maximum amplitude boundary model derived from the analysis of periodic pulse trains is now used to model SNR in overload for sinusoidal signals. The set of all possible reconstructed signals are confined to periodic signal constellations which are bounded by a $2K$-sided polygon. In the model, we use the amplitude r and phase of the source signal ϕ to decide on the best reconstructed signal which minimizes the mean square error. The magnitude of the reconstructed signal, denoted as $\hat{r}$, and the magnitude of the reconstructed error, denoted as r_e, are calculated as functions of the source signal (r, ϕ). Model equations for the reconstructed signal power ($\mathrm{E}\{\hat{r}^2\}/2$) and reconstructed SNR ($\mathrm{E}\{\hat{r}^2/r_e^2\}$) are then derived by computing ensemble averages assuming the source phase is uniformly distributed over $[0, 2\pi]$ radians.

4.3.1 Overload Regions

Overload is divided into two distinct amplitude ranges depending on the magnitude of the source signal: $r_{max} < r < r_v$ and $r > r_v$. The different overload conditions are shown in Fig. 4.6. Since the source is assumed to have a uniform phase distribution, and the bounding polygon is symmetrical with $2K$ edges, we limit the analysis of overload to a single edge of the polygon. Ensemble averages over all phase angles are computed later and exploit the symmetry.

For $r_{max} < r < r_v$, the overload error vector is always orthogonal to the maximum amplitude boundary as shown in Fig. 4.6(a). Depending on the phase of the source, two subregions are distinguished: R_0 and R_1. The region R_0 is not an overload region and corresponds to the region within the periodic signal constellation where reconstruction error is a function of sequence length L. The magnitude of the reconstruction error r_e is small in R_0 and assumed to be less than a value called ϵ. In R_1, the phase of the source signal is in a region where the source vector falls outside the maximum amplitude boundary resulting in a reconstruction error vector that is normal to the nearest face. The boundary between R_0 and R_1 is denoted by the phase angle α.

For source amplitudes which exceed r_v, the source vector always falls outside the bounding polygon for all source phase angles; however, the reconstructed error vector that minimizes the mean square error may or may not be orthogonal to a bounding edge. As shown in Fig. 4.6(b), two subregions, R_1 and R_2, are defined to distinguish the conditions when the source phase generates an orthogonal error vector and a non-orthogonal error vector, respectively. Similar to the range $r_{max} < r < r_v$, R_1 is a region where the source phase angle generates an error vector which is normal to the maximum amplitude boundary. In R_2, the source phase angle lies close to a vertex of the bounding polygon and the best reconstructed signal is a signal that corresponds to a vertex of the polygon. The resulting reconstruction error vector is no longer orthogonal to a bounding edge. The boundary between R_1 and R_2 is denoted by the phase angle β.

4.3.2 Model Equations

With reference to Fig. 4.7, the vector OA depicts the source signal (r, ϕ) where ϕ is measured relative to the normal ON. The reconstructed signal $(\hat{r}, \hat{\phi})$ is vector OB and the error vector is BA. In Fig. 4.7(a) the error vector is normal to the face, while in Fig. 4.7(b) the error

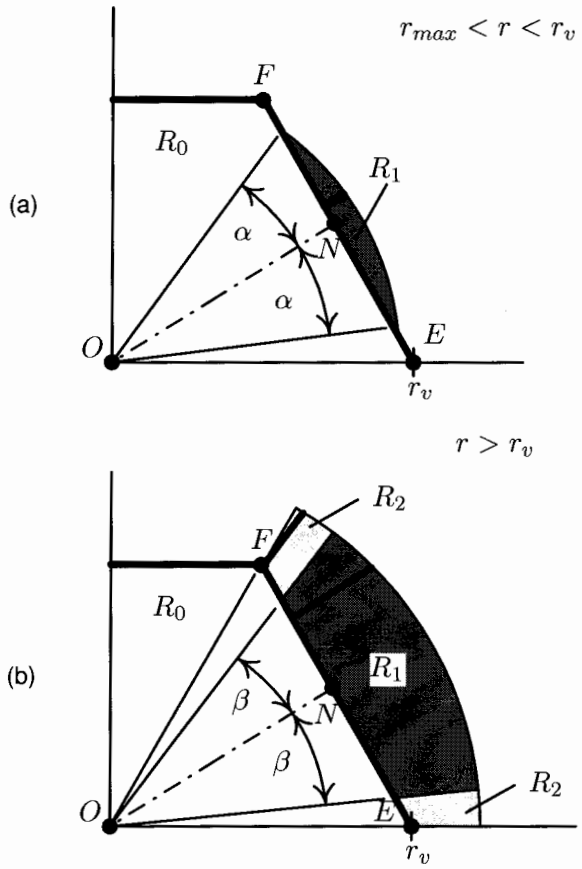


Figure 4.6: Signal reconstruction regions for a sinusoidal source.

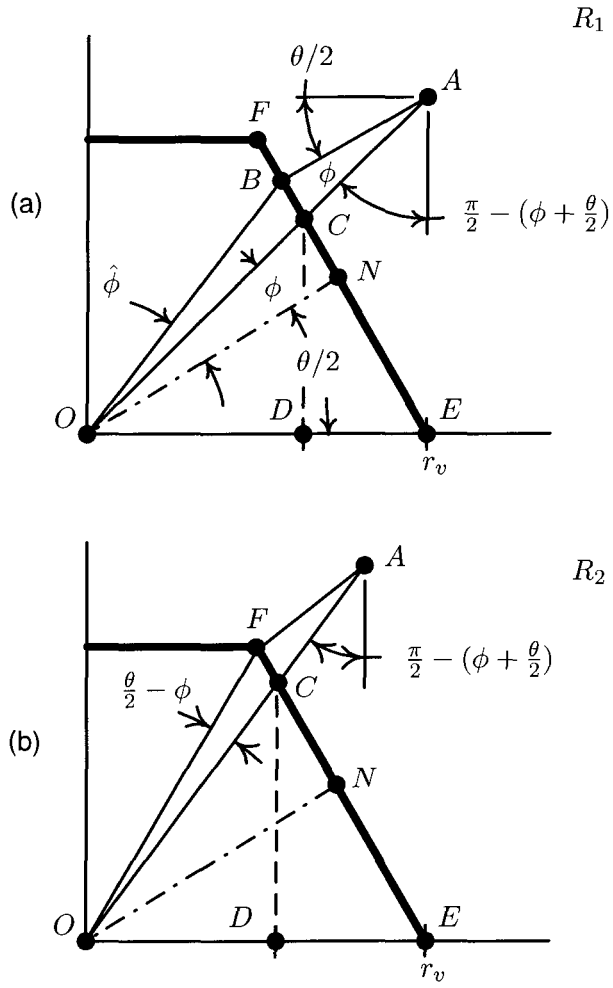


Figure 4.7: Signal reconstruction in overload for a sinusoidal source.

vector is non-orthogonal. We want to find expressions for magnitude of the reconstructed signal $\hat{r}$ and the magnitude of the reconstruction error r_e as functions of the source (r, ϕ).

4.3.2.1 $r_{max} < r < r_v$

The reconstruction vector OB and the error vector BA are found from the components, OC and CA, of the source vector. Using basic trigonometry, we find the length of OC, denoted as $|OC|$ is:

$$|OC| = \frac{r_{max}}{\cos(\phi)}. \tag{4.20}$$

The magnitude of the error vector BA is $|CA|\cos(\phi)$ where $|CA|$ is $r - |OC|$, and with (4.20) we get:

$$r_e = r\cos(\phi) - r_{max} \quad |\phi| < \alpha. \tag{4.21}$$

r_e is nonzero when $|\phi| < \alpha$ where

$$\alpha = \arccos(r_{max}/r) \quad r_{max} < r \leq r_v. \tag{4.22}$$

With $|OA|$ and $|BA|$ known, the cosine law is used to compute the squared magnitude of the reconstruction vector OB.

$$|OB|^2 = |OA|^2 + |BA|^2 - 2|OA||BA|\cos(\phi) \tag{4.23}$$

Substituting (4.21) into (4.23) and changing variables ($|OB| = \hat{r}$ and $|OA| = r$), $\hat{r}$ is

$$\hat{r}^2 = r_{max}^2 + r^2\sin^2(\phi) \quad 0 \leq \phi \leq \alpha. \tag{4.24}$$

The derivation of expressions for r_e^2 and $\hat{r}^2$ is completed for $r_{max} < r < r_v$ by considering the boundary conditions between R_0 and R_1. In R_0, the magnitude of the reconstruction error is bounded and less than or equal to ϵ. From the triangle inequality, $r^2 \leq \hat{r}^2 + \epsilon^2$, and for small ϵ, $\hat{r}^2 \approx r^2$. Therefore, in R_0

$$r_e^2 = \epsilon^2 \quad \alpha \leq \phi \leq \theta/2 \tag{4.25}$$

$$\hat{r}^2 = r^2 \quad \alpha \leq \phi \leq \theta/2. \tag{4.26}$$

At the boundary of R_0 and R_1, when $\phi = \alpha$, the expressions for r^2 and r_e^2 should match. The expression in R_1 given by (4.21) is zero for $\phi = \alpha$, and the derivation assumes

that if a source vector falls on the maximum amplitude boundary, then there is perfect reconstruction. The expression (4.21) is modified to include the constant ϵ and provide a consistent boundary condition at $\phi = \alpha$:

$$r_e^2 = (r\ \cos(\phi) - r_{max} + \epsilon)^2 \quad 0 \leq \phi \leq \alpha. \tag{4.27}$$

4.3.2.2 $r > r_v$

For the source range $r > r_v$, the error spans two regions, R_1 and R_2. The boundary between R_1 and R_2 occurs at $\phi = \beta$ and is defined as the point when $\hat{r} = r_v$ in R_1. From (4.24) we have

$$r_v^2 = r_{max}^2 + r^2 \sin^2(\beta) \tag{4.28}$$

which is rearranged to find β:

$$\beta = \arcsin\left(\sqrt{(r_v^2 - r_{max}^2)/r^2}\right). \tag{4.29}$$

In R_1, the equations in the previous section are applicable except α is replaced with β (4.29). In R_2, the reconstruction vector OF is constant and has a magnitude equal to the vertex radius r_v as shown in Fig 4.7. The reconstruction error vector AF is non-orthogonal, and the magnitude of the reconstruction error r_e is found using the cosine law to get:

$$r_e^2 = r^2 + r_v^2 - 2rr_v \cos(\theta/2 - \phi). \tag{4.30}$$

4.3.3 Reconstructed Signal Power

For a sinusoidal source, the reconstructed signal power is equal to $\frac{1}{2}\mathrm{E}\{\hat{r}^2\}$. The probability density function of the source phase is uniformly distributed over 2π, and by the symmetry of the maximum amplitude boundary, the expected value over the phase range $[0, 2\pi]$ is equal to $4K$ times the expected value over $[0, \theta/2]$.

For $r_{max} < r < r_v$, $\hat{r}^2$ is given by (4.26) in R_0 and given by (4.24) in R_1. Let $\hat{P}_{s0}$ and $\hat{P}_{s1}$ be the reconstructed power contributions in region R_0 and R_1, respectively. Then,

$$\hat{P}_{s0} = 4K \int_{\alpha}^{\theta/2} \frac{r^2}{2} \frac{d\phi}{2\pi} = \frac{\theta/2 - \alpha}{\theta} r^2 \tag{4.31}$$

and

$$\begin{aligned}\hat{P}_{s1} &= 4K \int_0^{\alpha} \frac{r_{max}^2 + r^2 \sin^2(\phi)}{2} \frac{d\phi}{2\pi} \\ &= \frac{\alpha}{2\theta}(2r_{max}^2 + r^2) - \frac{r^2}{4\theta} \sin(2\alpha). \end{aligned} \tag{4.32}$$

For $r > r_v$, the integration is over regions R_1 and R_2 with a boundary angle of β. The power in R_1 is the same as (4.32) except α is replaced with β (4.29). In R_2, the reconstruction amplitude is fixed and corresponds to the amplitude of the vertex r_v. The reconstruction power contributed in region R_2 is

$$\hat{P}_{s2} = 4K \int_{\beta}^{\theta/2} \frac{r_v^2}{2} \frac{d\phi}{2\pi} = \frac{\theta/2 - \beta}{\theta} r_v^2. \tag{4.33}$$

Together (4.31), (4.32), and (4.33) give expressions for reconstruction power $\hat{P}_s$ for all source amplitudes. For $r \leq r_{max}$, $\alpha = 0$, and (4.31) gives $\hat{P}_s = r^2/2$. For $r_{max} < r < r_v$, $\hat{P}_s$ is equal to $\hat{P}_{s0} + \hat{P}_{s1}$ where α is given by (4.22). For $r > r_v$, $\hat{P}_s$ is equal to $\hat{P}_{s1} + \hat{P}_{s2}$ where β is given by (4.29). Asymptotically, the reconstruction power is bounded and approaches $r_v^2/2$ as the source amplitude gets large.

4.3.4 Reconstructed SNR

The SNR of the reconstructed signal is defined as $\mathrm{E}\{\hat{r}^2/r_e^2\}$. In R_0 the SNR is

$$SNR_0 = 4K \int_{\alpha}^{\theta/2} \frac{r^2}{\epsilon^2} \frac{d\phi}{2\pi} = \frac{\theta - 2\alpha}{\theta} \frac{r^2}{\epsilon^2}. \tag{4.34}$$

In R_1 the SNR is given by

$$SNR_1 = 4K \int_0^{\alpha} \frac{r_{max}^2 + r^2 \sin^2(\phi)}{(r\cos(\phi) - r_{max} + \epsilon)^2} \frac{d\phi}{2\pi}. \tag{4.35}$$

Let

$$a = \frac{r}{r_{max} - \epsilon}. \tag{4.36}$$

then SNR_1 is written as

$$SNR_1 = \frac{a^2\left[2(r_{max}/r)^2 + 1\right]}{\theta} I_1 - \frac{a^2}{\theta} I_2 \tag{4.37}$$

where

$$I_1 = \int_0^{\alpha} \frac{d\phi}{\left[a\cos(\phi) - 1\right]^2} \tag{4.38}$$

and

$$I_2 = \int_0^{\alpha} \frac{\cos(2\phi)\, d\phi}{\left[a\cos(\phi) - 1\right]^2}. \tag{4.39}$$

The integrals (4.38) and (4.39) have closed form solutions [68]. Let

$$\begin{aligned} b &= a\cos(\alpha) - 1 \\ d &= \operatorname{arctanh}\left[\sqrt{\frac{a+1}{a-1}}\tan(\alpha/2)\right] \end{aligned} \tag{4.40}$$

then

$$I_1 = \frac{a(a^2-1)^{1/2}\sin(\alpha) + 2bd}{b(a^2-1)^{3/2}} \tag{4.41}$$

and

$$I_2 = \frac{bd(6a^2-4) - a(a^2-2)(a^2-1)^{1/2}\sin(\alpha)}{a^2 b(a^2-1)^{3/2}} + \frac{2\alpha}{a^2}. \tag{4.42}$$

The SNR in R_2 is

$$SNR_2 = 4K\int_{\beta}^{\theta/2} \frac{r_v^2}{r^2 + r_v^2 - 2rr_v\cos(\theta/2 - \phi)} \frac{d\phi}{2\pi}. \tag{4.43}$$

Let $k = (r^2 + r_v^2)/2rr_v$ and $\psi = \theta/2 - \phi$, then

$$\begin{aligned} SNR_2 &= \frac{r_v}{r\theta}\int_0^{\theta/2-\beta} \frac{d\psi}{k - \cos(\psi)} \\ &= \frac{2r_v}{r\theta\sqrt{k^2-1}} \arctan\left[\sqrt{\frac{k+1}{k-1}}\tan(\gamma/2)\right] \end{aligned} \tag{4.44}$$

where $\gamma = \theta/2 - \beta$.

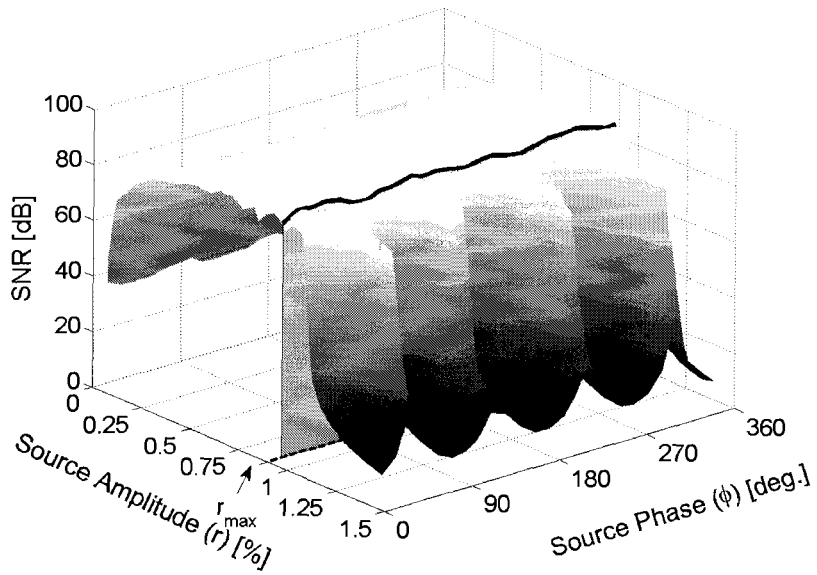
The SNR as a function of source amplitude r is summarized as follows. For $r \le r_{max}$, the source vector falls entirely within the polygon for all phases and the SNR is r^2/ϵ^2. For $r_{max} < r < r_v$, the SNR is equal to $SNR_0 + SNR_1$ where α is given by (4.22). For $r > r_v$, the SNR is equal to $SNR_1 + SNR_2$ where β is given by (4.29), and equations (4.40) to (4.42) for SNR_1 are modified by changing α to β (4.29).

4.4 Comparison of the Periodic Model with Bandpass ΣΔ Modulation

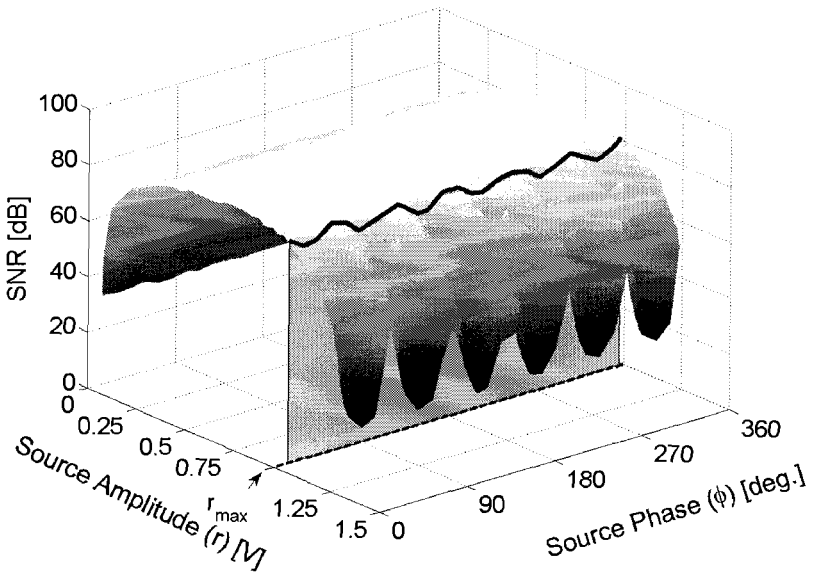
The analytic expressions developed for the periodic signal model are now compared with simulated SNR results for bandpass ΣΔ modulator pulse trains. Before discussing the comparison, we remark that the analysis of the periodic pulse train and the resulting structure of the reconstruction signal constellations leads to interesting large signal behavior. The maximum source amplitude r_{max} that defines the region where the reconstructed SNR is phase insensitive depends on the carrier oversample ratio. Also, within this region, the reconstruction error is constant (ϵ) and SNR increases with an increase in source amplitude r. For source amplitude levels larger than r_{max}, the reconstructed SNR become phase sensitive. Regions of high SNR occur if the source phase and amplitude lie within the maximum amplitude boundary. If the source phase and amplitude exceed the maximum amplitude boundary, then the SNR becomes dependent on the distance between the source vector and the maximum amplitude boundary. When $r > r_v$, the SNR is expected to drop off very rapidly, since the source vector falls outside the maximum amplitude boundary for all phases.

Simulation results for modulator A in section 2.2.2 are compared with the periodic model for a sinusoidal source signal, $r\cos[2\pi(f_c+\Delta f)t+\phi]$. Figure 4.8(a) shows simulation results for $R_c = 2$ ($f_c = f_s/4$) versus source amplitude and phase. The $f_s/4$ modulator design is a classic bandpass ΣΔ modulator configuration [11, 20, 31, 36, 37]. When the modulator is driven with large amplitude signals and $\Delta f = 0$, the SNR eventually collapses and enters a region of phase sensitivity. Most interesting is the occurrence of four dips in SNR at source phase angles of $0, \pi/2, \pi$ and $3\pi/2$ which is exactly what would be expected from the corresponding square constellation in the periodic model (Fig. 4.4) except for a phase offset of $\pi/4$. In Fig. 4.8(b), a similar plot is shown for the case when $R_c = 1.5$. Again, the presence of six dips in SNR is consistent with the hexagonal constellation for the periodic model. Because the large amplitude modulator behavior is consistent with the periodic model, it is likely that the modulator is limit cycling [69], and the data shows how the limit cycle behavior is phase sensitive.

In Fig. 4.9, the simulated SNR response of modulator A is compared with the SNR equations derived for the periodic model. Unlike other SNR responses shown earlier in Chapter 2, the SNR for each random trial is shown explicitly, and the ensemble average



(a) $R_c = 2$.



(b) $R_c = 1.5$.

Figure 4.8: SNR response of modulator A versus amplitude and phase for a sinusoidal source ($\Delta f = 0$).

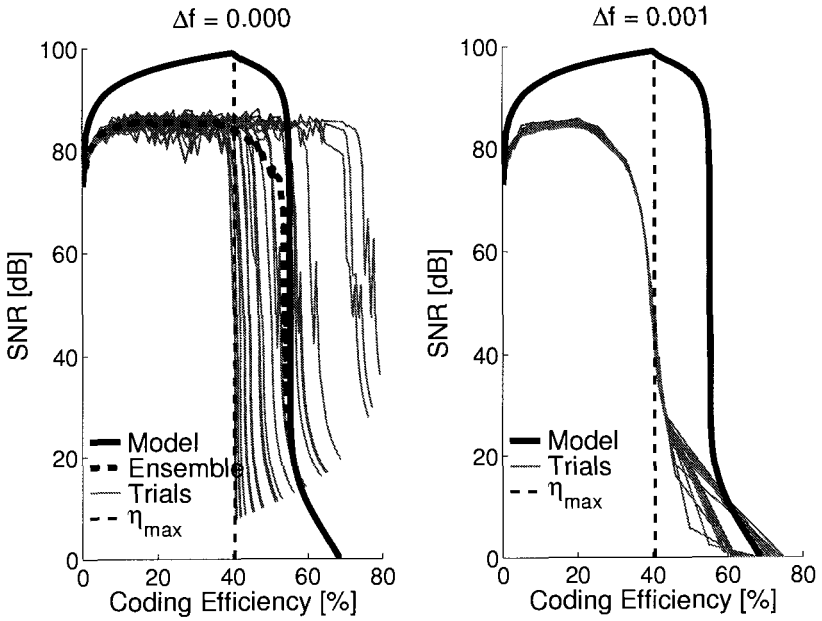
corresponds to the thick dashed line. For region R_0 in the periodic model, a value of $1 \cdot 10^{-5}$ is used for ϵ, and the value is derived from a linear analysis of the noise power in a fourth order bandpass modulator [33]. The figures show that the ensemble SNR averages of the simulation and model converge at low amplitudes and high amplitudes.

A vertical reference line denoted η_{max} is added to the figures. η_{max} corresponds to $r_{max}^2/2$, the maximum coding efficiency of a reconstructed signal in the phase insensitive region of the periodic model. Although the simulated results show some minor phase sensitivity for $r < r_{max}$, the phase sensitivity is much greater when $r > r_{max}$, and the coding efficiency in this region is tightly bounded to the right of the η_{max} reference line.

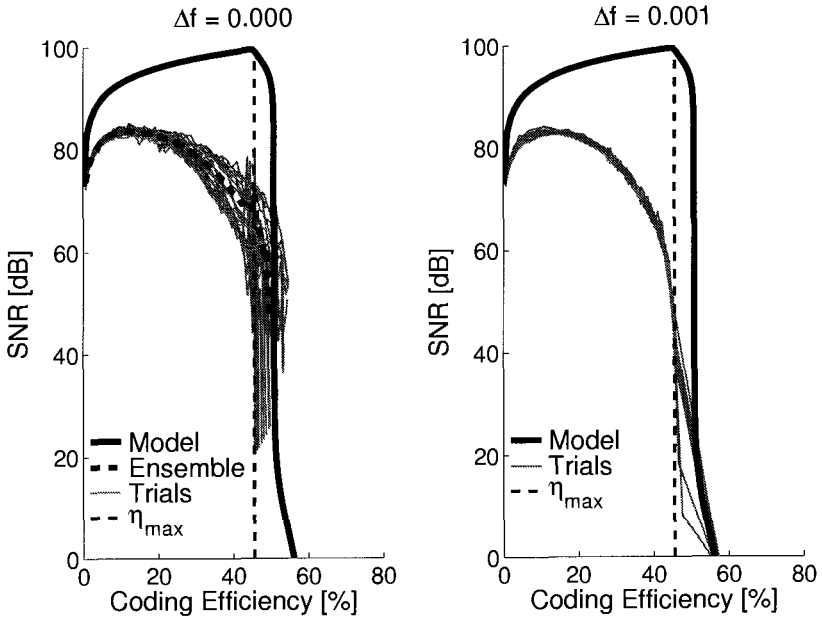
The modulator SNR measurements are made with an envelope oversample ratio of 128, and the SNR depends on the offset frequency Δf. When Δf is zero and the carrier oversample ratio is rational, the source frequency and sample rate are commensurate. Under this condition, the modulator SNR is sensitive to the source phase as shown in the left hand graphs of Fig. 4.9(a) and Fig. 4.9(b). On the other hand, if Δf is not zero, the phase sensitivity is significantly reduced as shown in the right hand graphs. From the periodic model, the reduction in phase sensitivity is expected providing the offset still generates a rational carrier oversample ratio. A small offset would significantly increase the number of edges on the reconstructed signal boundary reducing the overload region where source phase affects SNR. If the polygon has a large number of edges, then the boundary is nearly circular and phase sensitivity becomes small.

Another result derived from the periodic model is the relationship between the sequence length L and the frequency resolution of the reconstructed sinusoidal signal. As an example of the frequency resolution, consider all possible frequencies which are synthesized by periodic pulse trains with values of L up to 64 bits. After computing the difference between the discrete reconstruction frequencies, we get a graph of frequency resolution Δf_c versus f_c as shown in Fig. 4.10. The figure is very interesting and shows that frequency resolution is not uniform, and frequencies in the neighborhood of low integer ratios such as $f_c = 1/4$ and $f_c = 1/3$ have coarse frequency resolution. The distribution of resolution with center frequency is independent of L, and therefore the resolution is consistently poorer in the neighborhood of frequencies with low integer ratios. Thus, reconstruction frequencies with low integer ratios have both a maximum amplitude boundary with a small number of edges as well as coarse frequency resolution.

The relationship between frequency resolution and sequence length has implications in



(a) $R_c = 2$.



(b) $R_c = 1.5$.

Figure 4.9: SNR versus coding efficiency for modulator A with a sinusoidal source.

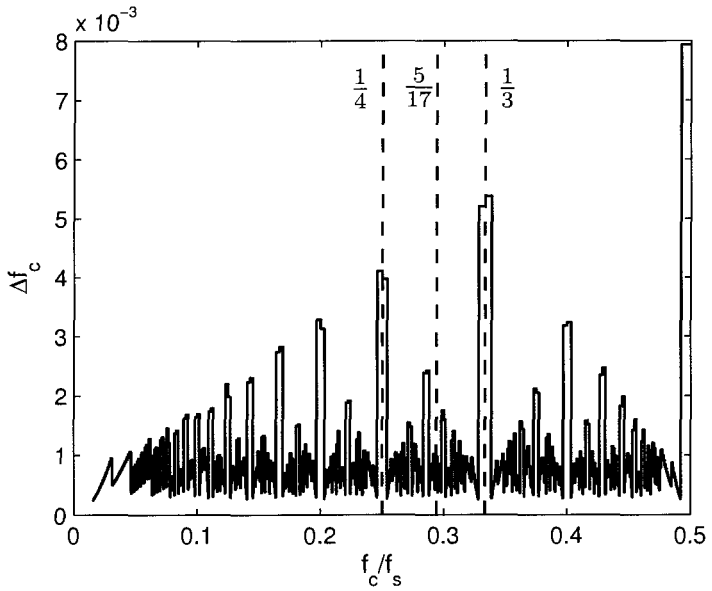


Figure 4.10: Reconstructed signal frequency resolution for sequences up to 64 bits.

terms of the pulse train synthesized by a modulator. Since the instantaneous frequency of the sinusoidal reconstruction signal varies in cases where the source frequency and modulator sample rate are not commensurate, the modulator attempts to find the best pulse train which minimizes the reconstructed signal error. If the source frequency is near a low integer ratio, then the pulse train must have a very long period to construct a signal with high SNR, and the modulator must track the error over a very long time window. On the other hand, if the source signal frequency is in the neighborhood of a frequency with fine resolution such as $f_c = 5/17$ ($R_c = 1.7$), then shorter sequences are expected to generate the same SNR relative to less optimal center frequencies.

The relationship between frequency resolution and SNR is tested in two ways. First, the SNR of the reconstructed signal drops when the source frequency is offset by $0.001 f_c$ relative to source frequencies of $1/4 f_s$ and $1/3 f_s$. This is shown by comparing the ensemble SNR of the left hand graphs in Fig. 4.9(a) and Fig. 4.9(b) with the SNR of the right hand graphs with offset. The peak gain in the noise shaping loop occurs at frequencies corresponding to the zeros in the noise shaping filter. As the source frequency shifts away from the zeros, the gain is reduced. Therefore, we expect that the reconstructed SNR will be less for source signals which are offset relative to the frequency of zeros in the noise shaping loop.

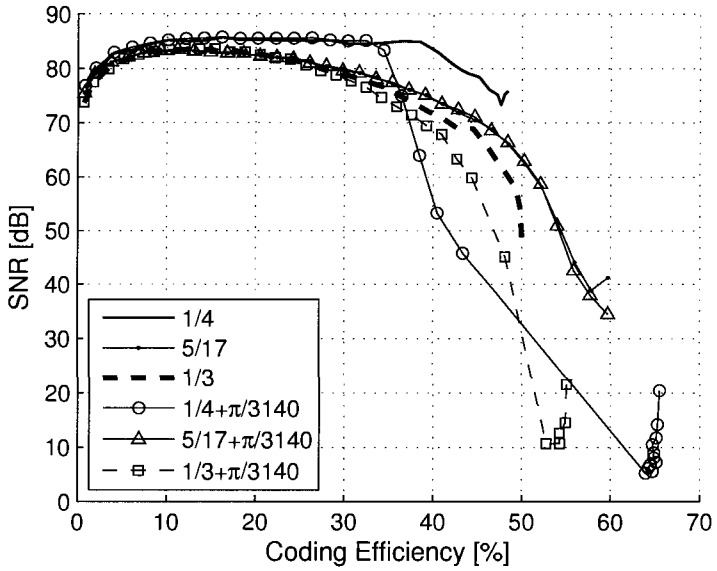


Figure 4.11: SNR versus coding efficiency for sinusoidal source signals with $\Delta f = 0$; the figure legend values correspond to f_c.

Instead of offsetting the source frequency relative to zeros in the modulator noise shaping loop, a second test is made where the modulator and source frequencies are locked and simultaneously tuned in the neighborhood of frequencies where the periodic model predicts coarse frequency resolution. The simulation results are summarized in Fig 4.11. SNR is first shown for $f_c = 1/4, 1/3$, and $5/17$ with $\Delta f = 0$, and then for an irrational offset of $\pi/3140 \approx 0.001$ with $\Delta f = 0$ again. The simulations show that when the frequency is offset near $f_c = 1/4$ and $f_c = 1/3$, the SNR drops, while the SNR remains approximately the same around $f_c = 5/17$. In all cases, the loop filter gain at the source frequency is the same, yet the modulator has difficulty synthesizing pulse trains with constant SNR in the neighborhood of $f_c = 1/4$ and $f_c = 1/3$. The observation is consistent with predictions from the periodic model which show that the length of the sequences changes depending on the frequency of the source signal.

The other modulator behavior worth noting in Fig. 4.11 is the apparent decrease in stability of the loop as the frequency is shifted away from $f_c = 1/4$ and $f_c = 1/3$. The SNR versus coding efficiency responses shown in the figure correspond to identical source level ranges which range from $\sigma_u^2 = 0.01$ to $\sigma_u^2 = 0.8$. Although the SNR is phase sensitive for $f_c = 1/4$ and $f_c = 1/3$ without offset, the modulator is stable and the SNR does not

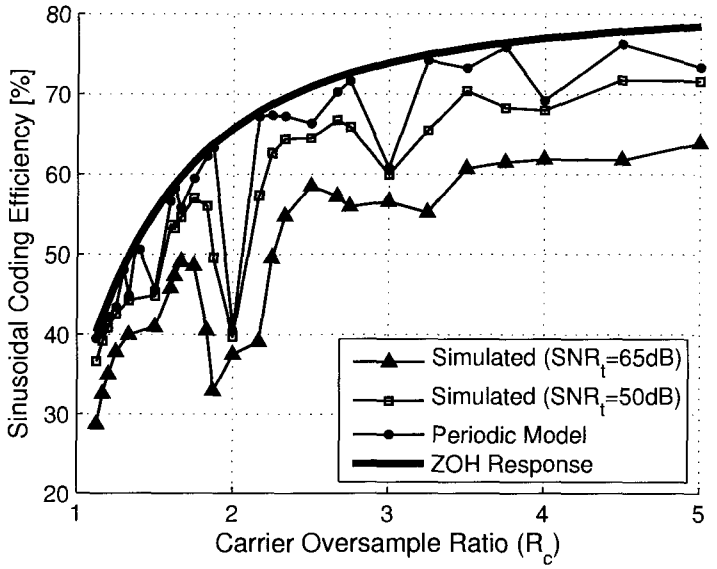


Figure 4.12: Coding efficiency versus carrier oversample ratio for a sinusoidal source signal.

collapse. However, when the source and noise shaping loop are offset, the SNR collapses sooner and the same range generates unstable operation as shown by the tails of the SNR responses near 10 dB. On the other hand, the robustness of the modulator at $f_c = 5/17$ is apparent. The conclusion from both the periodic model and modulator simulations is that SNR and coding efficiency vary depending on the carrier oversample ratio.

As a final result, the coding efficiency predicted by the periodic model is compared with the coding efficiency of bandpass $\Sigma\Delta$ modulation over a range of carrier oversample ratios. For the comparison, the model values of η_{max} are used for all rational carrier oversample ratios satisfying $M \leq 16$ and $M > 2N$. The model values are compared with simulation results for coding efficiency at SNR thresholds of 50 dB and 65 dB, where the noise shaping loop is set to the corresponding carrier oversample ratio, and the input signal has an offset of $0.001 f_c$. With reference to Fig. 4.9(a) and Fig. 4.9(b) again, the 50 dB threshold is expected to correspond closely with η_{max}, while the 65 dB threshold is expected to have lower coding efficiency than the model. The comparative results are shown in Fig. 4.12.

Both the model and modulator show similar variation in coding efficiency and the coding efficiency dip around $R_c = 2$ is significant. $R_c = 2$ corresponds to the square reconstruction signal constellation, and r_{max} is significantly less than r_v in this case. Other significant dips

in coding efficiency occur when M has a low value such as $M = 6$ for carrier oversample ratios of 1.5 and 3. As expected, the correlation of the model is better at an SNR threshold of 50 dB than at a threshold of 65 dB. There are also differences between the model and modulator, and a more comprehensive model would integrate both the amplitude and frequency resolution aspects which were analyzed separately.

4.5 Conclusions

As an approach to understanding how coding efficiency varies in binary amplitude pulse trains with synchronous zero-crossings, we analyzed the class of periodic pulse trains. The analysis was motivated by considering how to encode a sinusoidal source signal into a pulse train with constraints similar to a bandpass $\Sigma\Delta$ modulator. A comparison of the periodic model with a fourth order bandpass $\Sigma\Delta$ modulator shows that coding efficiency dips at certain common carrier oversample ratios including 1.5, 2.0, and 2.5, and carrier oversample ratios in the neighborhood of these values should be avoided in a RF class D amplifier design.

Chapter 5

Signal Reconstruction and Coding Efficiency

In the previous chapters, we have evaluated the coding efficiency of different bandpass $\Sigma\Delta$ modulators, shown how coding efficiency is related to load power in a RF class D amplifier, and compared modulator coding efficiency with a periodic model. Since the coding efficiency of the binary amplitude pulse train is so important in the amplifier application, a natural question to ask is: are there better source coding methods than bandpass $\Sigma\Delta$ modulation? And if there are better methods, how much better are they?

Linear signal reconstruction in the RF class D amplifier imposes significant constraints on the encoding method. The distinction of a linear reconstruction filter is important, and differentiates the problem from classic rate-distortion theorems. For example, there is a source coding theorem for gaussian sources with memory [70], and the application of this theorem to a bandlimited source signal shows that far superior encoders and decoders exist. An example of this theorem is given in Appendix D. Although the theorem shows the existence of better encoders and decoders, it does not consider system constraints which are consistent with class D amplification.

In this chapter, several different methods are taken to evaluate the coding efficiency limitations of linear signal reconstruction. The first method is to examine the reconstruction signal space assuming a second order bandpass filter. An elementary analysis of the second order bandpass filter shows that the maximum reconstructed signal is bounded by $4/\pi$ for any binary amplitude pulse train without any constraint on the zero-crossing positions. From this, we show that the maximum coding efficiency for an undistorted sinusoidal signal is approximately 81%. The result is consistent with the periodic signal model in Chapter 4, except that it is derived from the assumption of a linear reconstruction filter only.

In the second and third parts of this chapter, we attempt to synthesize more optimal binary amplitude pulse trains with higher coding efficiency than bandpass $\Sigma\Delta$ modulation. Two alternate encoding methods are considered. First, a numerical search algorithm is used to synthesize a semi-infinite binary sequence which minimizes the reconstruction signal error power. The high envelope oversample ratio in the encoder makes the computational complexity of the search very challenging, and after trying different algorithms, the M-algorithm was selected for linear numerical complexity. The results of the numerical search suggest that bandpass $\Sigma\Delta$ modulation is remarkably efficient and difficult to exceed. The numerical search results also show similar coding efficiency variation with changes in carrier oversample ratio as observed with bandpass $\Sigma\Delta$ modulation.

The second alternate encoder replaces the binary quantizer in a bandpass $\Sigma\Delta$ modulator with a more general decision unit. The concept is recent, and began with a paper by Kato [28]. He proposed a trellis search with the Viterbi algorithm in place of the binary quantizer in a lowpass modulator. We build on Kato's concept, and define a decision unit which 'looks ahead' to make a better decision than a binary quantizer. An encoder design with a decision unit in a bandpass modulator loop is evaluated, and the results show that a few bits of look-ahead can improve the robustness of the modulator especially for high order loop filters.

5.1 Linear Signal Reconstruction

Assume the input signal to the reconstruction filter $g(t)$ is limited to two amplitude states of ± 1 V. The output (reconstructed) signal is then a superposition of step responses for the reconstruction filter. Let the input to the filter be a pulse train $p(t)$ with zero-crossings at times t_i. The input pulse train is

$$p(t) = \pm\left(u_s(t) + 2\sum_{i=1}^{\infty}(-1)^i\, u_s(t-t_i)\right) \tag{5.1}$$

where $u_s(t)$ is the unit step input and the sign depends on whether the pulse train starts with a $+1$ at $t = 0$ or a -1 at $t = 0$. The reconstructed signal, called $\hat{u}(t)$, is then a superposition of the filter step response $g_s(t)$ and equal to

$$\hat{u}(t) = \pm\left(g_s(t) + 2\sum_{i=1}^{\infty}(-1)^i\, g_s(t-t_i)\right). \tag{5.2}$$

Equations (5.1) and (5.2) are very general, and they describe every possible type of encoder pulse train, and every possible reconstructed signal at the output of a RF class D amplifier. The source signal is encoded entirely in the zero-crossings at times t_i of the input pulse train, and signal reconstruction depends on the step response $g_s(t)$ of the reconstruction filter.

The zero-crossing times t_i can be synchronous with a periodic clock signal as in $\Sigma\Delta$ modulation or asynchronous as in PWM. Synchronous zero-crossing times are assumed, and t_i is restricted to an integer multiple of the clock period T. Under these conditions, the number of output amplitude levels at $t = kT$, where k is an integer, is countable, and the set of all possible output levels lies on a discrete lattice.

5.1.1 Signal Reconstruction with a Second Order Filter

A second order bandpass filter has a transfer function

$$H(s) = \frac{2\zeta s}{(s+\zeta)^2 + \nu^2} \tag{5.3}$$

where ζ is the damping factor and ν is the natural frequency of the resonator. The step response is a damped oscillation with a single harmonic component at a frequency of ν rad/s:

$$g_s(t) = 2\frac{\zeta}{\nu} e^{-\zeta t} \sin(\nu t)\, u_s(t). \tag{5.4}$$

Although the step response is an elementary signal with a single sinusoidal frequency component, the superposition of step responses can generate signals with time varying amplitude and phase.

The step response of the second order bandpass filter is shown in Fig. 5.1 for different damping values ζ and a fixed natural frequency ν of 10 rad/s. A high damping value or low Q resonator generates an initial large amplitude sinusoidal step and the amplitude decreases rapidly with time. From a signal reconstruction perspective, this step function has coarse amplitude resolution and short memory. On the other hand, a low damping value or high Q resonator generates a low amplitude sinusoid with a slowly drooping envelope; the step function provides fine amplitude resolution with long memory. In $\Sigma\Delta$ modulators, second order resonator sections with very high Q are used, and consequently fine amplitude resolution with very long memory is obtained. It also shows why a large envelope oversample ratio is required to obtain good SNR.

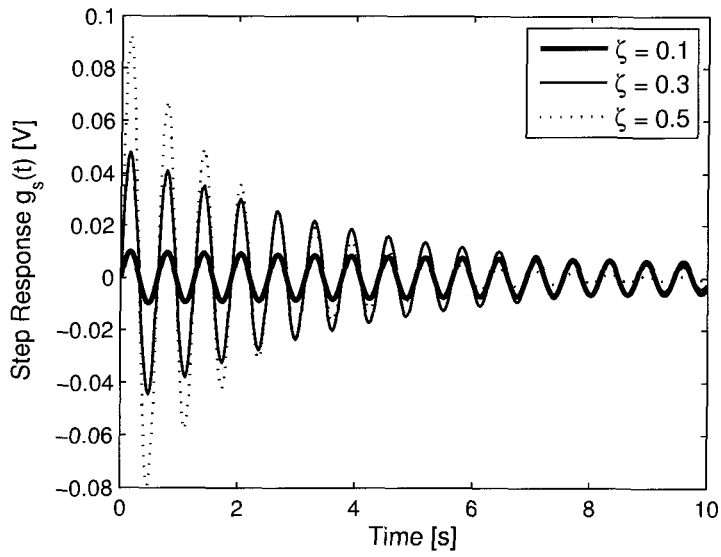


Figure 5.1: Second order bandpass filter step response for $\omega_c = 10$ rad/s.

Focusing on high Q resonators and assuming the peak amplitude of the oscillation is nearly constant for many cycles, it is easy to construct a sinusoid with quantized amplitude values. Let T_ϵ be the duration of the step response where the peak amplitude of the oscillation is within ϵ of some peak amplitude $A = 2\zeta/\nu$:

$$|\max(|g_s(t)|) - A| < \epsilon \quad t \in [0, T_\epsilon]. \tag{5.5}$$

If $T_\epsilon \gg T_\nu$, then the peak amplitude at time $t = T_\nu/4$ is a good approximation of the future peak amplitudes for $0 < t < T_\epsilon$. An example based on this assumption is shown in Fig. 5.2 where the amplitude is stepped up and then down in amplitude increments of $2A$. A crude approximation based on this assumption is that the amplitude error is $\pm A$, and a reconstructed signal consists of lattice points spaced by $2A$.

5.1.2 Maximum Output Amplitude with a Second Order Filter

The maximum output amplitude from the filter is generated by an infinite sequence of step responses that add constructively at the output of the filter. Hence, the zero-crossings of the input pulse train must be synchronized and periodic. The input pulse train $p(t)$ that maximizes the superposition of step responses at the filter output is a pulse train that changes sign every $T_\nu/2$ where $T_\nu = 2\pi/\nu$. Note that the pulse train is selected without any

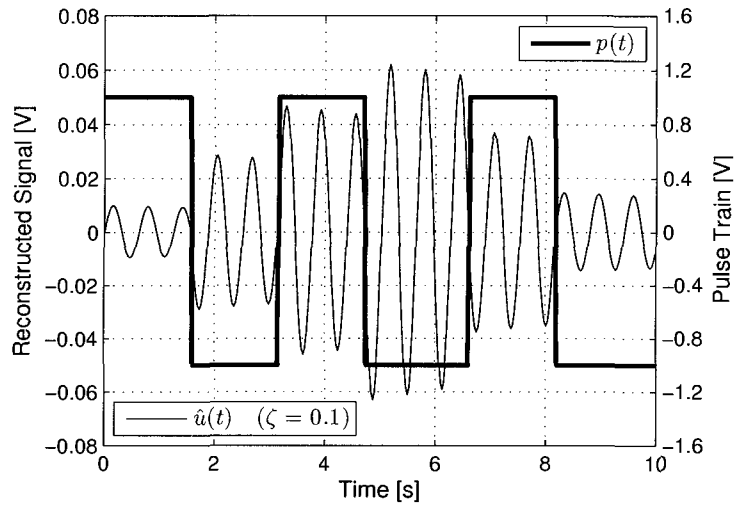


Figure 5.2: Example of a reconstructed load signal $\hat{u}(t)$ with a second order bandpass filter for an input pulse train $p(t)$.

prior constraint on where the zero-crossings are located, and is

$$p(t) = \text{sgn}[\sin(\nu t)] \qquad t \geq 0. \tag{5.6}$$

The maximum amplitude of the step response occurs at $t = T_\nu/4$, and

$$g_s(T_\nu/4) \;=\; 2\frac{\zeta}{\nu}e^{-\frac{\zeta\pi}{2\nu}} = 2\alpha a \tag{5.7}$$

where $\alpha = \zeta/\nu$ and $a = e^{-\frac{\zeta\pi}{2\nu}}$. After k time steps, the peak amplitude is

$$g_{max} \;=\; 4\alpha(a + a^3 + a^5 + \ldots + a^{k-1}) + 2\alpha a^k \tag{5.8}$$

and the limit as $k \rightarrow \infty$ is

$$g_{max} \;=\; 4\alpha \sum_{i=0}^{\infty} a^{(2i+1)} = 4\alpha a \sum_{i=0}^{\infty} (a^2)^k = \frac{2\alpha}{\sinh(\pi\,\alpha/2)}. \tag{5.9}$$

For a high Q resonator, the damping factor, ζ, is small. Consequently, $\alpha = \zeta/\nu$ is small, and in the limit as $\zeta \rightarrow 0$ we get

$$\lim_{\alpha \rightarrow 0} g_{max} = \lim_{\alpha \rightarrow 0} \frac{4}{\pi \cosh(\pi\alpha/2)} = \frac{4}{\pi} \tag{5.10}$$

where l'Hopital's rule is used to evaluate the limit. The maximum output amplitude is $4/\pi$ for a perfect resonator; otherwise it is (5.9), and less than $4/\pi$.

The maximum output amplitude signal is sinusoidal and generated by input pulse train (5.6) which is a square wave with a 50% duty cycle. If $p(t)$ is extended to a bi-infinite interval $t \in (-\infty, \infty)$, then the Fourier series for $p(t)$ has a fundamental frequency component with amplitude $4/\pi$, and is identical to a causal resonator with infinite Q as $t \to \infty$. The implication of (5.10) is that the coding efficiency of the pulse train for an undistorted sinusoidal load signal is limited to approximately 81%. The example also shows that any attempt to encode a more complex source signal like W-CDMA with signal peaks exceeding $4/\pi$ will have a clipped reconstructed signal.

5.2 Binary Sequence Search

In this section, the constraint of linear reconstruction filters is explored further for higher order filters. Analytic results are more difficult to generate for filter orders greater than two, and conclusions on coding efficiency for more complex signals such as narrowband gaussian noise (NGN) are desirable. Therefore, we revert to simulation to gain further insight. A Butterworth filter, with a 3 dB bandwidth of f_b and center frequency f_c, is selected where f_b and f_c correspond to a specific envelope and carrier oversample ratio, respectively. A numerical search algorithm is then used to synthesize the best binary sequence which minimizes the MSE of the reconstructed signal.

A discrete-time equivalent model, shown in the top of Fig. 5.3, is constructed for the simulation. The impulse response of the reconstruction filter is $g[k]$, and a search algorithm synthesizes a binary pulse train $b[k]$ which minimizes the MSE of the reconstructed signal $\hat{u}[k]$. The SNR and coding efficiency of the synthesized pulse train are then computed. The coding efficiency of the discrete-time pulse train is adjusted for a NRZ DAC response to compare with bandpass $\Sigma\Delta$ modulation. The search algorithm implemented in the encoder and simulation results are described next.

5.2.1 M-Algorithm

Since the encoder synthesizes a semi-infinite binary sequence $b[k]$, a numerically efficient search algorithm is required. After experimenting with different numerical algorithms, the sub-optimal M-algorithm by Anderson [71] was selected for its linear computation complexity. Other algorithms tested on the problem were sphere decoding, the T-algorithm, and semi-definite programming. The sphere decoding algorithm finds optimal binary sequences,

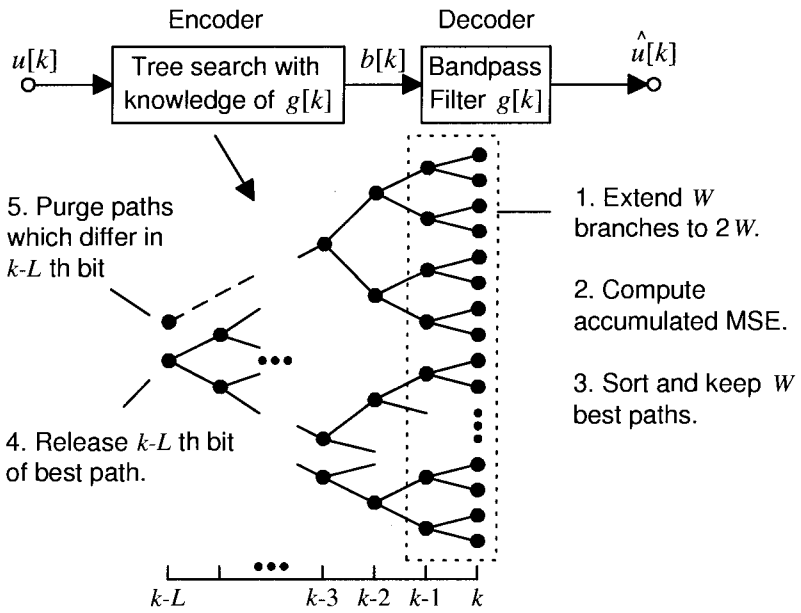


Figure 5.3: Binary sequence search with the M-algorithm.

but numerical simulation times were exponential with sequence length, and Matlab simulations beyond 30 bits were very long. The T-algorithm has a threshold which limits the number of retained branches in a tree search. It was found that even a very small threshold accumulated a large number of branches, and hard limits were required to limit tree growth effectively reducing the algorithm to the M-algorithm. In joint work with Amin Mobasher, semi-definite programming methods were tested on the problem, and again we jointly concluded that this method was no better than the M-algorithm. [1]

The M-algorithm is briefly summarized below and reference is made to Fig. 5.3.

1. There are two key variables: W is the number (breadth) of contenders retained at each step in the tree search, and L is the depth (delay) after which an output bit is released.

2. At each iteration, the W paths are extended by 1 bit to generate $2W$ contenders. The $2W$ contenders are sorted in terms of a MSE and W paths with the lowest MSE are retained.

[1] Mobasher et al [72] have successively applied SDP methods to lattice search problems.

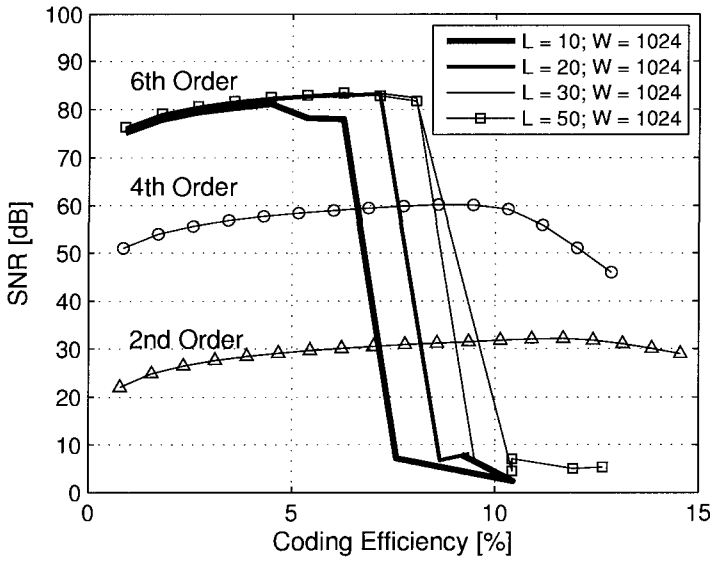


Figure 5.4: Comparison of SNR versus coding efficiency for different Butterworth reconstruction filters for a NGN source; the 3 dB filter bandwidths are $f_s/256$ and $R_c = 2.0$. The second order parameters are $L = 2$ and $W = 4$, while the fourth order parameters are $L = 8$ and $W = 128$.

3. At each iteration k, the $k - L$ th bit is released for the best path. After releasing the bit, any remaining contenders with a different bit in the $k - L$ th position are purged.

5.2.2 Selecting Algorithm Parameters

A number of different numerical experiments were conducted to determine what algorithm parameters should be selected for L and W. The experimental results show asymptotic behavior, and the increase in performance diminishes with increasing parameter values. Parameters values which balance numerical simulation time and performance were selected, and the results of the numerical experiments are summarized in Figures 5.4 and 5.5.

Fig. 5.4 shows SNR versus coding efficiency responses for a binary sequence of 32768 bits encoding a NGN source with different filter orders and different algorithm parameters. Starting with filter order, there is a direct relationship between the selectivity of the filter which increases with filter order, and the complexity of the numerical search. For a second order filter, the reconstructed SNR is low, and a binary sequence is easily found with two bits of look ahead ($L = 2$; $W = 4$). As the filter order increases, the impulse response of

the filter spans a larger time window, and the parameters need to be much larger in order to find sequences with large SNR and high coding efficiency. For example, typical fourth order filter search parameters are $L = 8$ and $W = 128$, while typical sixth order parameters are $L = 40$ and $W = 1024$. Also, with the sixth order filter, the experiments show that even with large parameter values the algorithm fails to find binary sequences with large coding efficiency. For convergence, a large number of contender paths must be retained and memory starts to be a limiting resource in the search. Since robust results are obtained with the fourth order filter, it is selected for comparison with bandpass $\Sigma\Delta$ modulation.

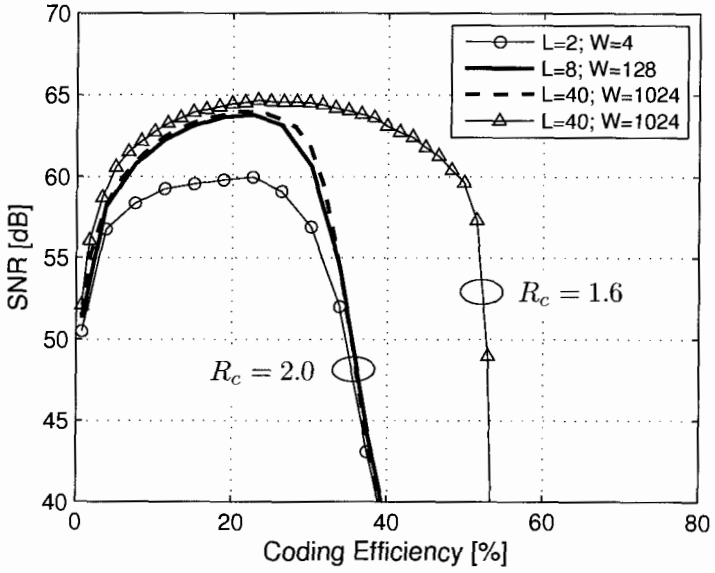
A detailed characterization of the algorithm parameters with the fourth order reconstruction filter is shown in Fig. 5.5 for sinusoidal and NGN sources. The data show there are significant improvements as L is increased to 8, and as W is increased to 128. Beyond these values, the SNR and coding efficiency improvements diminish rapidly for both sinusoidal and NGN sources. Therefore the algorithm parameters are set to $L = 8$ and $W = 128$, and the binary sequence search is employed to characterize the variation in SNR and coding efficiency with carrier oversample ratio.

5.2.3 Simulation Results

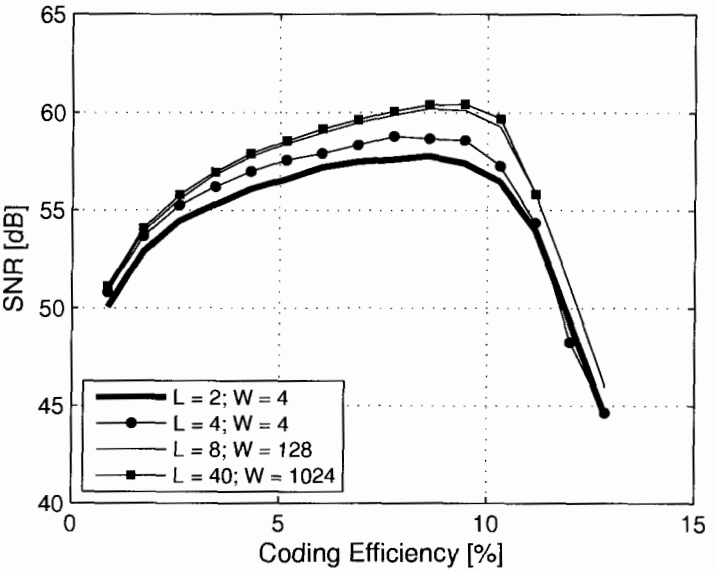
For both sinusoidal and NGN sources, the M-algorithm sequence search generates results very similar to bandpass $\Sigma\Delta$ modulation. These results are summarized in Figures 5.6–5.8.

Fig. 5.6 compares the binary sequence search with different modulator designs for a carrier oversample ratio of 2.0. The $\Sigma\Delta$ modulator (SDM) SNR and coding efficiency are measured with a sixth order (BPF6) Butterworth reconstruction filter, while binary search (BS) results are presented for both fourth order (BPF4) and sixth order (BPF6) Butterworth filters. The tails of the SNR responses converge for fourth order $\Sigma\Delta$ modulator designs and for the binary search with a fourth order reconstruction filter (BS; BPF4). On the other hand, both the sixth order modulator (SDM F) and the binary sequence search with a sixth order reconstruction filter (BS; BPF6) show abrupt collapses in the SNR responses.

The sixth order modulator provides superior SNR performance for low amplitude input signals; however, the maximum coding efficiency is limited by stability — a well known problem with high order modulators [21]. A similar collapse in SNR is observed for the binary search when a sixth order reconstruction filter is used. For large input amplitudes, the sequence search is unable to retain a low MSE branch, and much more memory (L and W) is required to synthesize a sequence with comparable performance to bandpass $\Sigma\Delta$



(a) Sinusoidal source.



(b) NGN source for $R_c = 2.0$.

Figure 5.5: Comparison of SNR versus coding efficiency for different M-algorithm parameters with a fourth order reconstruction filter; the 3 dB filter bandwidth is $f_s/256$.

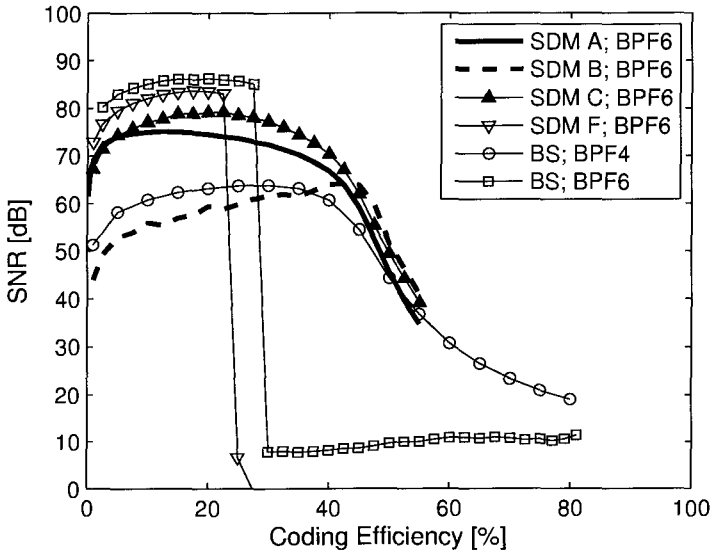


Figure 5.6: Coding efficiency of sequences generated by the M-algorithm binary sequence search for sinusoidal source signals; $R_c = 2$ and the 3 dB filter bandwidth is $f_s/256$.

modulation. For example, modulators A, B, and C all synthesize more optimal sequences with better coding efficiency. Since the sequence search with a fourth order filter is much better for large source amplitudes, a better search algorithm might include an adaptive bandwidth or adaptive order reconstruction filter. The filter order could be adjusted in response to the source amplitude, resulting in a smooth transition from high SNR pulse trains at low source amplitudes, to moderate SNR pulse trains at high source amplitudes. This is not explored further in this research project, and it could be an area for future research.

A more extensive sweep of SNR versus coding efficiency is used to generate the coding efficiency plots shown in Fig. 5.7. Two SNR thresholds of 52.5 dB and 62.5 dB are selected, and the coding efficiency of the synthesized pulse train is measured versus carrier oversample ratio. Similar to results shown in Chapter 2 for bandpass $\Sigma\Delta$ modulation (cf. Fig. 2.10) and for the periodic model (cf. Fig. 4.12), the coding efficiency dips significantly at carrier oversample ratios of 1.5, 2.0, and 3.0. The data also shows the coding efficiency dips in the neighborhood of these carrier oversample ratios.

The SNR versus coding efficiency for a NGN source is shown in Fig. 5.8. Data are shown for four carrier oversample ratios of 1.5, 1.7, 2.0, and 2.2. In terms of increasing

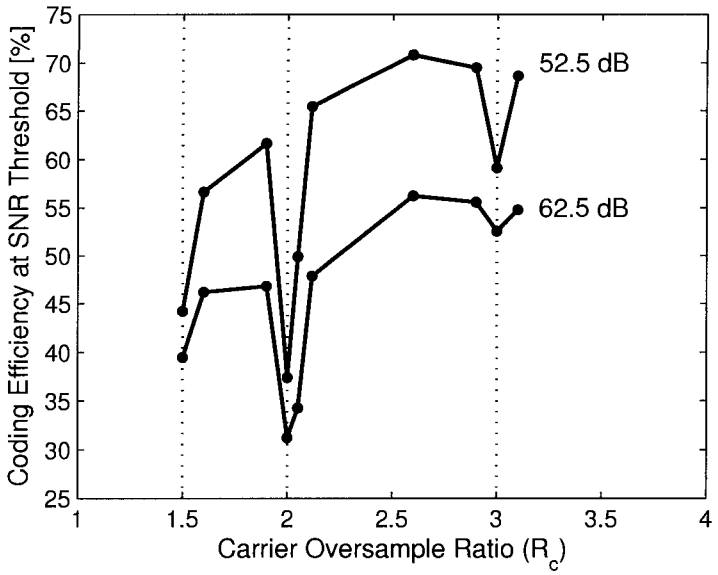


Figure 5.7: Coding efficiency of sequences generated by the M-algorithm binary sequence search with a fourth order Butterworth bandpass reconstruction filter. The source signal is sinusoidal and the 3 dB filter bandwidth is $f_s/256$.

coding efficiency, the carrier oversample ratios are ordered as 2.0, 1.5, 1.7, and 2.2. Similar results are obtained for bandpass $\Sigma\Delta$ modulation (cf. Fig. 2.12).

From these simulations we make several conclusions. First, a numerical search does not generate substantially better pulse trains with higher coding efficiency than bandpass $\Sigma\Delta$ modulation. Since the search parameters in the algorithm are selected to be large enough where the gains in performance diminish, the data suggests that the coding efficiency of bandpass $\Sigma\Delta$ modulation appears difficult to exceed. Second, there is a region before SNR collapse that SNR can be boosted by higher order reconstruction filters or higher order modulators, as illustrated by the sixth order filters. However, higher order filters have long impulse responses, and the memory starts to limit the efficient synthesis of pulse trains at high amplitudes where the SNR starts to collapse. Third, as with bandpass $\Sigma\Delta$ modulation, the binary sequence search confirms that the coding efficiency of some carrier oversample ratios are more optimal than others.

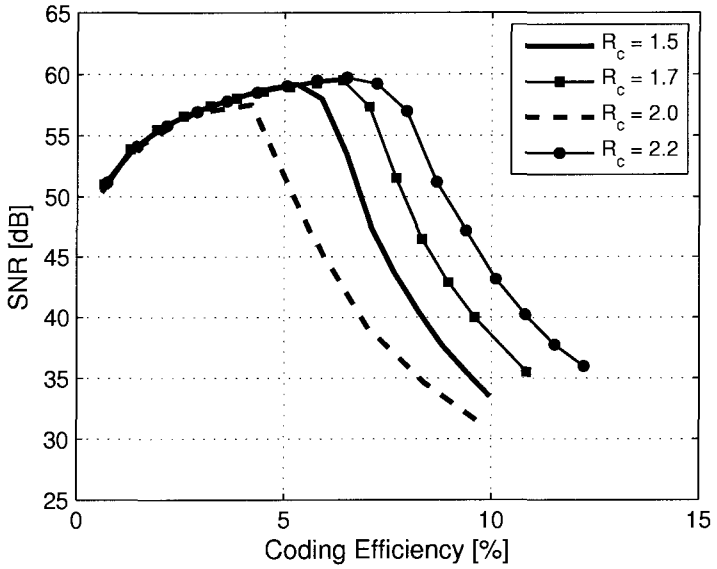


Figure 5.8: Coding efficiency of sequences generated by the M-algorithm binary sequence search with a fourth order Butterworth bandpass reconstruction filter. The source signal is NGN and the 3 dB filter bandwidth is $f_s/256$.

5.3 Bandpass ΣΔ Modulation with a Decision Unit

The results of the M-algorithm binary sequence search suggests that the bandpass ΣΔ modulator is efficient. One of the reasons the modulator is so efficient has to do with the long memory of high Q resonators, and the large envelope oversample ratio. The high Q resonator $h[k]$ is in a feedback loop as shown in Fig. 5.9, and stores a long history of the error signal $w[k]$, defined as the difference between the input $u[k]$ and output $b[k]$ signals. The high Q resonator creates a long search window over which the error signal is integrated, and the sign of the weighted error signal $v[k]$ is used as a decision in the circuit. The quantizer is a simple and an easily implemented decision unit in the modulator loop, but are there better ways to make a decision with the error signal $v[k]$ that improve coding efficiency?

5.3.1 L Bit Look-Ahead

There has been recent work in lowpass audio modulators [28–30] where alternate decision units have been considered in the modulator loop instead of a quantizer. A similar idea is applied here to bandpass ΣΔ modulators. The bandpass ΣΔ modulator is modified

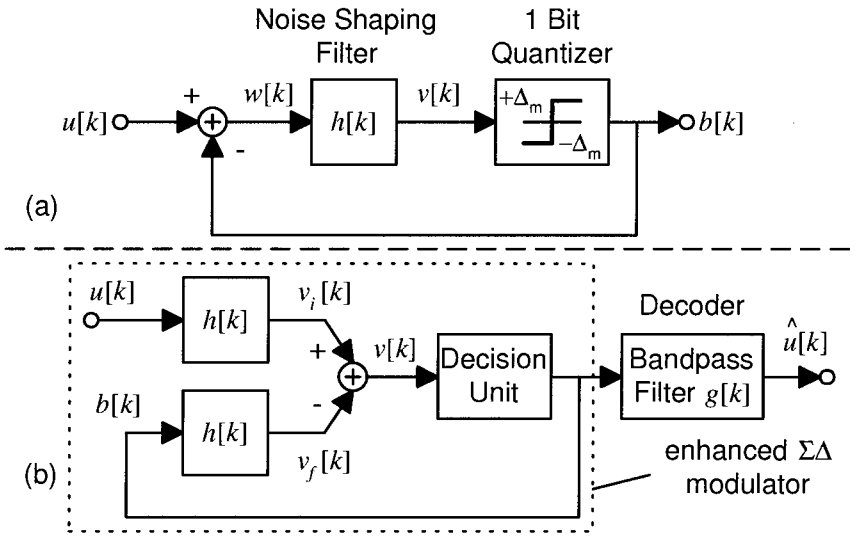


Figure 5.9: ΣΔ modulator (a) and equivalent model (b) where the quantizer is replaced by a decision unit.

to include a short length look-ahead tree search of L bits, and a bit is released from the path with the lowest MSE. The depth of the tree search is limited to eight bits, and the accumulated MSE of all $W = 2^L$ paths are evaluated before releasing a bit from the path with the minimum MSE. The modified modulator architecture is shown in Fig. 5.9.

When a decision unit with at least one cycle of delay is included in the modulator loop, the loop filter no longer requires delay for the loop to be causal. Usually the noise shaping filter $h[k]$ has a least one delay, and the delay can now be moved or shared with the decision unit. As an example, the transfer function of modulator A is modified to have zero delay, and the corresponding transfer function is

$$H(z) = \frac{z(z+\alpha)(2z^2+3\alpha z+1)}{(z^2+2\alpha z+1)^2}. \tag{5.11}$$

Equation (5.11) can be compared with the transfer functions (2.16) and (2.17) shown earlier in Chapter 2. Note that in this case, $H(z)$ is equivalent to $H_i(z)$, the input transfer function, and $-H(z)$ is equivalent to $H_f(z)$, the feedback transfer function. Simulation results for this transfer function, as well as a sixth order design using a delayed transfer, modulator F, are shown next.

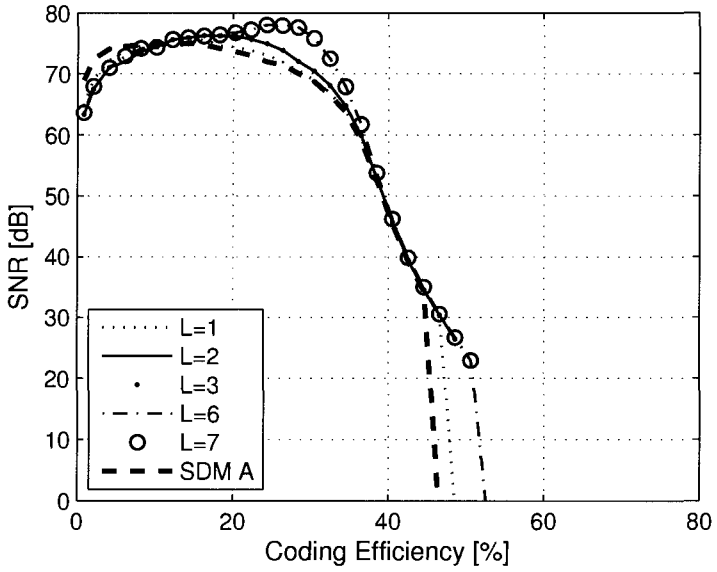
5.3.2 Simulation Results

It has already been shown that a bandpass $\Sigma\Delta$ modulator, composed of a high Q resonator and a binary quantizer in a feedback loop, is highly efficient. The memory of a high Q resonator is analogous to a long tree search; therefore, it might be expected that long tree searches in the decision unit after the resonator are unlikely to yield much improvement. As the simulation data show, the improvements in coding efficiency depend on the SNR threshold, the carrier oversample ratio, and on the modulator design. The greatest benefit from a tree search decision unit appears to be improved robustness in the modulator by improving loop stability.

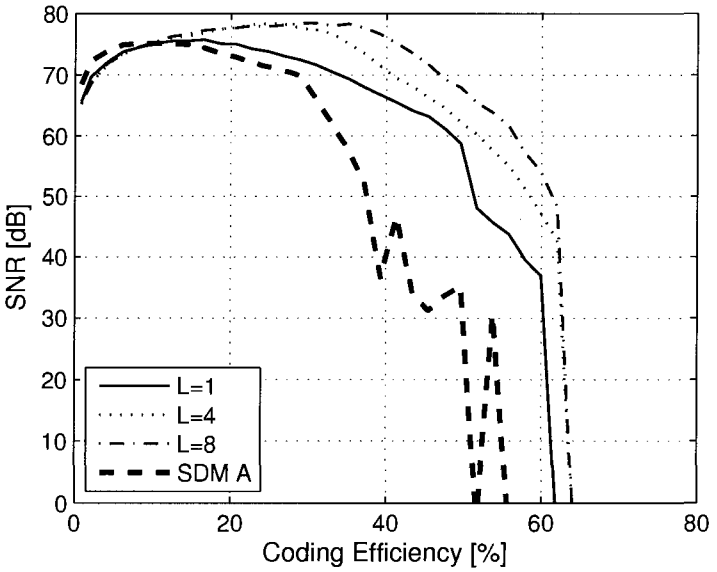
Figure 5.10(a) shows the SNR versus coding efficiency of the enhanced fourth order modulator design using the loop filter in (5.11) with look-ahead. The results are compared with modulator A for a carrier oversample ratio of 2.0, and show that look-ahead can boost the SNR near the point of SNR collapse. All the tails of the different designs converge around an SNR of 50dB, and the tails of the designs with a decision unit extend farther to the right. For $L = 7$, the SNR results have the appearance of pushing up against a boundary layer, and the limits of coding efficiency appear to be close to the simulation results.

A similar comparison between the enhanced modulator and modulator A is shown in Figure 5.10(b) for a carrier oversample ratio of 2.1. The carrier oversample ratio is near 2.0, and the SNR response for modulator A is jagged showing signs of instability at high source levels. The response is typical near the coding efficiency dip at a carrier oversample ratio of 2.0. The effectiveness of look-ahead is substantially better in this case, and significant improvements in coding efficiency are obtained even with one bit of look-ahead. This shows that the breadth of the dip in coding efficiency near 2.0 seen in bandpass $\Sigma\Delta$ modulators with a binary quantizer, could be narrowed significantly with an improved decision unit.

Since the robustness of the modulator appears to improve with a tree search, experiments were also made with a sixth order modulator design, modulator F. Sixth order modulators are notoriously unstable at high amplitudes, yet they have significant advantages in terms of SNR. The results of the simulations are shown in Fig. 5.11. The SNR collapses very rapidly in sixth order designs. Since the SNR of the modulators does not collapse gracefully as the source amplitude gets large, some form of clipping is required at the input to insure the modulator remains stable. If this is done, then the addition of look-ahead has substantial benefits in the sixth order design, and clearly pulse trains with higher coding efficiency can



(a) $R_c = 2.0$



(b) $R_c = 2.1$

Figure 5.10: Coding efficiency of a fourth order bandpass $\Sigma\Delta$ modulator with a tree search decision unit. The source signal is sinusoidal and the reconstruction filter is a sixth order Butterworth filter with a 3 dB bandwidth of $f_s/256$.

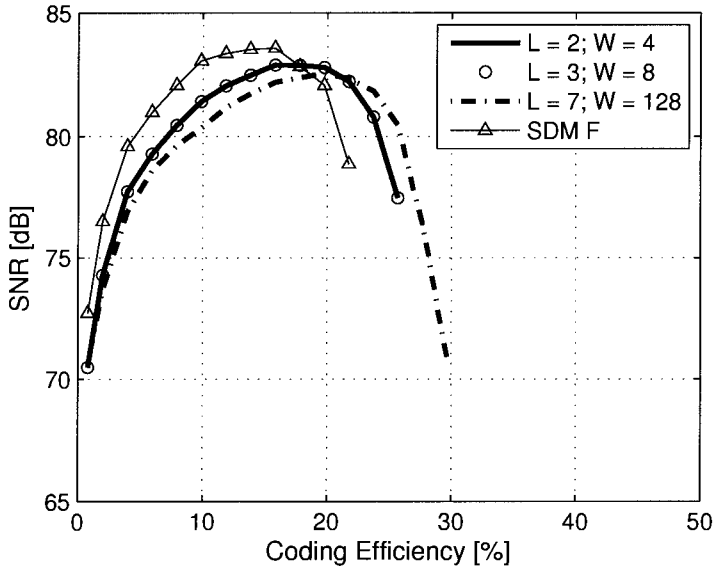


Figure 5.11: Coding efficiency of a sixth order bandpass $\Sigma\Delta$ modulator with a tree search decision unit. The source signal is sinusoidal, the reconstruction filter is a sixth order Butterworth filter with a 3 dB bandwidth of $f_s/256$, and $R_c = 2.0$.

be synthesized. For example, at an SNR threshold of 80 dB, 2 bits of look-ahead could boost coding efficiency by 3%, and 7 bits could boost coding efficiency by 5%.

5.4 Conclusions

The goal of this chapter has been to explore limitations in coding efficiency for the synthesis of binary amplitude pulse trains with constrained zero-crossings. A fundamental constraint in the RF class D amplifier stage is linear signal reconstruction, and different approaches were taken to evaluate coding efficiency limits assuming a bandpass reconstruction filter. The first approach was to assume the filter is a second order resonator, and then ask what kind of signal reconstruction is possible with this filter. It was proven for the second order filter, that the maximum amplitude of any reconstructed signal is $4/\pi$ for a pulse train with unit amplitude levels. The amplitude bound corresponds to a periodic excitation, and consequently the maximum coding efficiency of a sinusoidal load signal without distortion is 81%.

Signal reconstruction was explored more generally for other filters and source signals

using a binary sequence search. The simulation results with the M-algorithm show that bandpass $\Sigma\Delta$ modulation is an efficient search engine, and it is difficult to exceed the performance by employing numerical methods for generating binary sequences. As a final approach to evaluating coding efficiency limitations, the efficiency of bandpass $\Sigma\Delta$ modulation was combined with a tree search, and together modest improvements in coding efficiency were obtained. The main advantages of the enhanced modulator architecture were: 1) a boost in SNR near the point of SNR collapse, 2) a reduction in the range of carrier oversample ratios where coding efficiency dips, and 3) improved stability in a sixth order modulator.

Chapter 6

Manchester Encoded Bandpass Sigma-Delta Modulation

A challenge in high frequency continuous-time modulator implementations is the design of RF resonators with sufficient Q to obtain good signal-to-noise ratio (SNR) [32]. A discrete component design can take advantage of high Q inductors and capacitors [31]. In monolithic implementations, on chip inductors have low Q, and Q enhancement techniques are required to improve resonator selectivity [43,73]. Other monolithic resonator structures such as Gm-C filters are also feasible, but high frequency designs are still challenging.

As an alternative to the implementation of RF resonators, frequency conversion [74] can be employed in the modulator to reduce the resonator frequency in the noise shaping loop. In this chapter, a general form of Manchester encoding is analyzed as a way to upconvert the output of a low frequency bandpass $\Sigma\Delta$M to a RF output, and thereby mitigate issues with RF resonator implementations. The proposed concept is to multiply the output pulse train $b[k]$ with a synchronized clock signal which effectively maps each modulator bit to an integer multiple m of a $(+1,-1)$ or $(-1,+1)$ pattern depending on the sign of the modulator bit [40,75]. Although $m = 1$ is usually defined as Manchester encoding, the term Manchester encoding shall be used generically for any $m > 0$ as used by the authors in [75]. The objective of the analysis is to evaluate the coding efficiency and average transition frequency of the upconverted pulse train, and assess the cost of the upconversion relative to a direct synthesis of the pulse train.

The analysis yields interesting conclusions. When a power loss ratio (PLR) term is constructed to compare the coding efficiency of a bandpass CT $\Sigma\Delta$M with and without upconversion, a relatively simple equation results which depends only on the upconversion frequency ratio. The PLR equation is independent of m, and shows that the reduction

in coding efficiency with upconversion does not exceed 40.5%. Coding efficiency losses are high when the upconversion frequency ratio is large, and low when the upconversion ratio is small. The analysis of ATF, on the other hand, does depend on m; ATF is minimized, hence power efficiency is maximized, when m is unity. The overall conclusion of the analysis is that $m = 1$ is the best configuration for minimizing the upconversion power loss ratio and ATF. The analysis is verified by simulations with modulator A.

The chapter is organized as follows. An overview is given of a Manchester encoded $\Sigma\Delta$M in a RF class D amplifier application. Then the coding efficiency of Manchester encoding is analyzed, and the chapter concludes with an analysis of ATF.

6.1 Upconversion and Manchester Encoding

As an alternate to direct synthesis, upconversion of a low frequency pulse train $y(t)$ to a RF frequency pulse train $p(t)$ is considered. With upconversion, the load signal has a carrier frequency f_o, while the source signal has a carrier frequency f_c, and the ratio f_o/f_c is defined as the upconversion ratio:

$$\text{UCR} = f_o/f_c. \tag{6.1}$$

In the direct path $f_o \equiv f_c$ and $p(t) \equiv y(t)$.

One way to upconvert is to multiply the binary amplitude signal $y(t)$ with a clock signal $c(t)$ as shown in Fig. 6.1 when the switches are in position (b). The bit period (sample period) from the modulator is T, and the clock period in the upconversion path is T/m. When $m = 1$, the method is equivalent to Manchester encoding [76]; each bit in $y(t)$ is mapped to a high-low or low-high pulse. In general, m is any positive integer, and a designer may want to choose $m > 1$ because the modulator and upconverter circuitry might be implemented in different device technologies. An example of upconversion with $m = 2$ is shown in Fig. 6.2, and in [75], an example with $m = 5$ is given.

In Fig. 6.2(a), the power spectral density of the modulator pulse train $S_y(f)$ is shown for a source signal with $f_c = f_s/3$. The upconverted PSD, $S_p(f)$, is shown in Fig. 6.2(b) for a UCR of 5. The dashed lines in Fig. 6.2(a)-(b) outline the spectral shaping by the modulator: NRZ spectral shaping without upconversion, and Manchester spectral shaping from $c(t)$ with upconversion. The equations for spectral shaping are given later in section 6.2. Besides upconverting the modulator signal, the upconverter reduces low frequency power which has

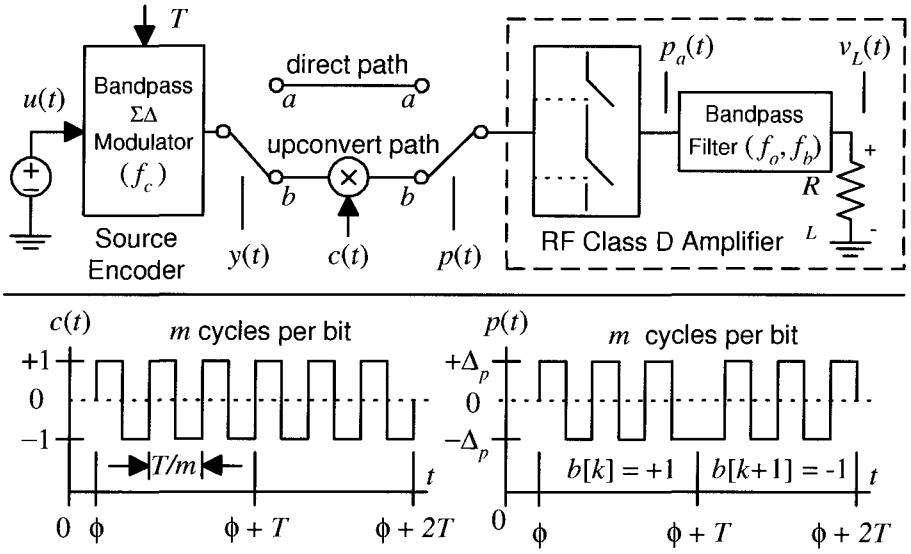


Figure 6.1: ΣΔ modulator, upconverter, and RF class D amplifier model.

the additional benefit that AC coupling can be used in the RF amplifier. Fig. 6.2(c) shows a narrow span of $S_p(f)$ around the upconverted signal. After upconversion, a pulse train with high coding efficiency and low ATF are desirable for subsequent amplification.

6.2 Power Loss Ratio

Upconversion changes the ΣΔM pulse train power spectrum and reduces coding efficiency. The reduction in coding efficiency is analyzed by constructing a ratio of the upconverted pulse train coding efficiency over the direct conversion pulse train coding efficiency. The ratio turns out to be a ratio of reconstructed signal powers if upconversion does not alter the amplitude (Δ_p) of the modulator pulse train. The ratio is called the power loss ratio (PLR). The PLR is a measure of the relative change in load power at the output of the amplifier after changing the switch from the direct path, switch position (a), to the upconvert path, switch position (b), in Fig. 6.1.

Let the coding efficiency of the direct path and upconverter path be η_y and η_p, respectively. The amplitude of the modulator pulse train $y(t)$ is $\pm\Delta_p$, and the amplitude of the

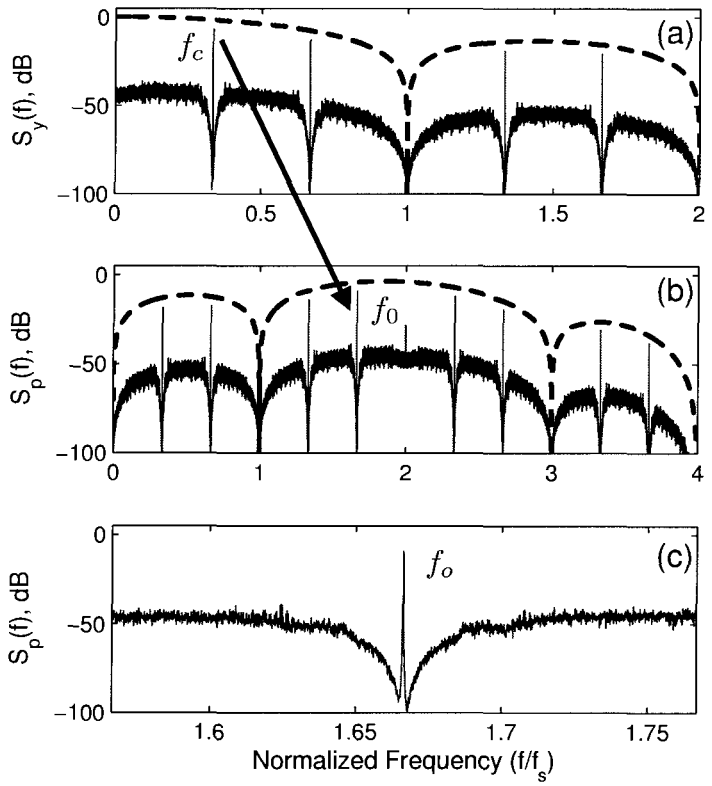


Figure 6.2: PSD of modulator pulse train $y(t)$ and upconverted pulse train $p(t)$.

clock signal is ± 1. Then,

$$\text{PLR} = \frac{\eta_p}{\eta_y} = \frac{P_L^{uc}/\Delta_p^2}{P_L/\Delta_p^2} = \frac{P_L^{uc}}{P_L} \tag{6.2}$$

where P_L is the reconstructed signal power obtained from the $\Sigma\Delta$M pulse train $y(t)$, and P_L^{uc} is the reconstructed signal power obtained from the upconverted pulse train $p(t)$.

6.2.1 Power Loss Ratio Analysis

Rather than using the multiplier upconverter model in Fig. 6.1, an impulse train model shown in Fig. 6.3 is used for analysis. The approach is similar to the PSD analysis of pulse amplitude modulated signals [76,77]. In the impulse train model, the upconverter has an impulse response $d_m(t)$, and generates a burst of m pulses per modulator bit period T. Each pulse in the burst is defined as a rectangular pulse of width $T/(2m)$ and amplitude $+1$, followed by a second rectangular pulse of equal duration and amplitude -1. An equation for the impulse response is

$$d_m(t + T/2) = \sum_{n=1}^{m} \text{rect}\left[\frac{t - a_n T}{T/(2m)}\right] - \text{rect}\left[\frac{t + a_n T}{T/(2m)}\right] \tag{6.3}$$

where $\text{rect}(t)$ is a rectangular pulse function defined as

$$\text{rect}(t) = \begin{cases} 1 & |t| < 1/2 \\ 0 & \text{otherwise} \end{cases} \tag{6.4}$$

and a_n is a normalized delay function:

$$a_n = (4n - 3 - 2m)/(4m). \tag{6.5}$$

The impulse response in (6.3) is shown with a $T/2$ time advance, and results in an odd function; the time advance is compensated for later in the Fourier transform domain.

The impulse train from the modulator is defined as a wide-sense stationary random process $b_c(t)$

$$b_c(t) = \sum_{k=-\infty}^{\infty} b[k]\delta(t - kT - \phi) \tag{6.6}$$

with binary weights $b[k] \in \{-\Delta_p, \Delta_p\}$ and sample period T. The polarity of each pulse sequence at the upconverter output depends on the sign of the modulator bit $b[k]$. An

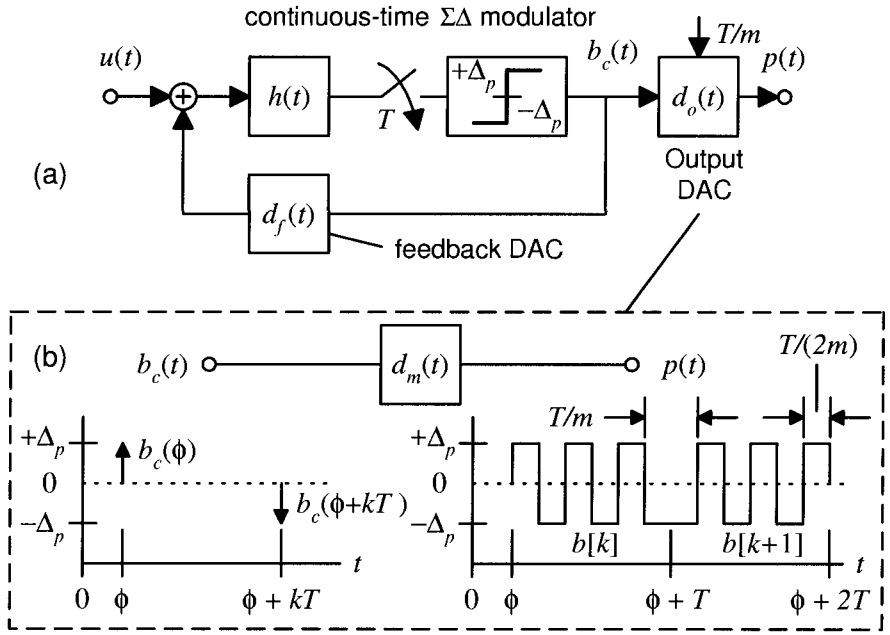


Figure 6.3: Bandpass ΣΔM model: direct conversion when $d_o(t) \equiv d_{nrz}(t)$ and upconversion when $d_o(t) \equiv d_m(t)$.

arbitrary initial delay ϕ is assumed, and ϕ is a random variable with a uniform distribution over the interval $[-T/2, T/2]$. The impulse train $b_c(t)$ is derived by either differentiating the modulator pulse train $y(t)$, or, as shown next, it is a signal which is directly available in a CT ΣΔM behavioral model.

A CT ΣΔM model is shown in Fig. 6.3. In the model, a noise shaping filter $h(t)$ is followed by an ideal sampler and quantizer with a feedback DAC $d_f(t)$. Since the model has an ideal sampler in the loop, an impulse train $b_c(t)$ already exists in the model. In a direct conversion configuration, $b_c(t)$ drives an NRZ output DAC: $d_o(t) \equiv d_{nrz}(t)$; in an upconverter configuration it drives a Manchester pulse: $d_o(t) \equiv d_m(t)$. The bandpass ΣΔM is assumed to be tunable and the center frequency of the noise notch f_c is set by the carrier oversample ratio R_c. The carrier oversample ratio and upconversion frequency ratio are related as shown below:

$$\text{UCR} = \frac{f_o}{f_c} = \frac{mf_s - f_c}{f_c} = 2mR_c - 1. \tag{6.7}$$

Since $R_c > 1$, equation (6.7) imposes a constraint on the relationship between UCR and m:

$$\text{UCR} > 2m - 1 \qquad \text{for } m = 1, 2, 3, \ldots . \tag{6.8}$$

The PSD of $b_c(t)$ is [77]

$$S_{b_c}(f) = \frac{1}{T} S_b(e^{j\omega}) \tag{6.9}$$

where $\omega = 2\pi f/fs$, and $S_b(e^{j\omega})$ is the discrete-time Fourier transform of the autocorrelation function of the bit sequence $b[k]$. Since $S_b(e^{j\omega})$ is periodic with period f_s, $S_{b_c}(f)$ is periodic. In the upconversion path, the mth image of the principal spectral part $|f| < f_s/2$ is selected by spectral shaping in the upconverter. The index m corresponds to which image is selected.

The output PSD after upconversion is

$$S_p(f) = |D_m(f)|^2 S_{b_c}(f) \tag{6.10}$$

where $D_m(f)$ is the Fourier transform of $d_m(t)$. Similarly, the output PSD for the direct path is

$$S_y(f) = |D_{nrz}(f)|^2 S_{b_c}(f) \tag{6.11}$$

where $D_{nrz}(f)$ is the transform of the NRZ impulse response $d_{nrz}(t)$. With (6.2)-(6.11), the PLR is

$$\text{PLR} = \frac{\int_{-f_b/2}^{+f_b/2} |D_m(f + f_o)|^2 S_b(f + mf_s - f_c) df}{\int_{-f_b/2}^{+f_b/2} |D_{nrz}(f + f_c)|^2 S_b(f + f_c) df} \tag{6.12}$$

where the substitution $f_o = mf_s - f_c$ has been used.

$\Sigma\Delta$M's are oversampling converters and the signal bandwidth of the carrier f_b is much less than the sample rate f_s. Since $f_b \ll f_s$, the attenuation over the signal bandwidth is approximately constant for both the encoder $D_m(f)$ and the NRZ pulse $D_{nrz}(f)$. The pulse shaping terms in the integrals of (6.12) are approximately constant, and

$$\text{PLR} \approx \frac{|D_m(f_o)|^2}{|D_{nrz}(f_c)|^2} = \frac{|D_m(\text{UCR} f_c)|^2}{|D_{nrz}(f_c)|^2}. \tag{6.13}$$

We begin the evaluation of the PLR by first computing the transforms of the impulse responses. The Fourier transforms of the NRZ pulse is

$$D_{nrz}(f) = Te^{j\pi fT/2} \text{sinc}(fT) \tag{6.14}$$

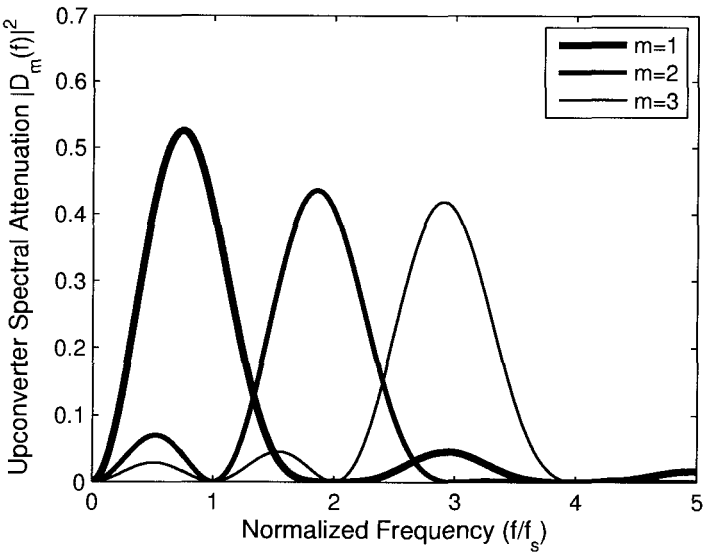


Figure 6.4: Upconverter spectral shaping from $|D_m(f)|^2$ ($T = 1$).

and the transform of the upconverter pulse (6.3) is

$$D_m(f) = \frac{T}{m} e^{j\pi(fT-\frac{1}{2})} \text{sinc}\left(\frac{fT}{2m}\right) \sum_{n=1}^{m} \sin(2\pi f T a_n). \tag{6.15}$$

At this point it is useful to plot $|D_m(f)|^2$ and it is shown in Fig. 6.4 for values of $m = 1, 2, 3$. The spectral shaping has a global maximum for $f > 0$ at each value of m in the interval $0 < (mf_s - f) < 1$. If coding efficiency is to be maximized, then it is clearly desirable to have an image of the encoded source signal fall near a maxima. Also, spectral attenuation is minimum for $m = 1$; however, as we show next, when PLR is calculated, there is no longer a dependence on m, and this results because of the construction of a ratio where terms cancel out. Instead, the spectral shaping manifests itself in terms of the UCR.

Continuing with the analysis of PLR, let $\nu = fT$ be a normalized frequency, then $\nu_c = f_c T$ and $f_o T = m - \nu_c$. With these variable changes

$$\text{PLR} = \left[\frac{\text{sinc}[(m-\nu_c)/(2m)]}{m\,\text{sinc}(\nu_c)} \sum_{n=1}^{m} \sin[2\pi(m-\nu_c)a_n]\right]^2. \tag{6.16}$$

By converting the sum to a closed form solution, (6.16) can be reduced to

$$\text{PLR} = \left[\frac{\nu_c}{m-\nu_c}\cot[\pi\nu_c/(2m)]\right]^2. \tag{6.17}$$

The derivation of (6.17) is given in Appendix E. The ratio ν_c/m is equal to UCR + 1 and (6.17) is finally written as

$$\text{PLR} = \left[\frac{\cot[\pi/(2\text{UCR}+2)]}{\text{UCR}}\right]^2 \tag{6.18}$$

with the constraint UCR $> 2m-1$ as was given earlier in (6.8). The PLR equation (6.18) is shown in Fig. 6.5.

From the PLR equation, several important conclusions are made. First, the equation collapses to a PLR of unity when the UCR is unity, and hence is consistent with the trivial case when there is no upconversion. Second, the lower bound on the PLR is $4/\pi^2 \approx 0.405$ and found by letting UCR $\rightarrow \infty$. Third, and most important, is the observation that the PLR is independent of the number of clock cycles m in the upconverter pulse shape. The independence of PLR in terms of m has resulted from the normalization of coding efficiency to a configuration without upconversion (ie: an NRZ DAC), and shows that there is no advantage in using $m > 1$. Later, when the average pulse period is analyzed for upconversion, it is shown that the average period is proportional to m; therefore, $m = 1$ is the best overall choice to maximize coding efficiency and minimize switching losses.

6.2.2 Power Loss Ratio Simulations

A simulation of a fourth order continuous-time modulator, design D, is used to verify the power loss ratio equation. The simulated power loss ratio is found by measuring the power of the encoded signal in the upconverted pulse train $p(t)$ relative to the power of the encoded signal in the pulse train $y(t)$. The source signal is a sinusoid with random phase, and the signal power is measured both with an ideal bandpass filter and with a sixth order Butterworth filter. The reconstruction filter bandwidth is $f_b/256$, corresponding to an envelope oversample ratio of 128. With both methods, an ensemble of 25 random trials is calculated for the shown data and the tone has an input level of -4.5 dBV relative to the quantizer step size Δ_p.

The analytic and simulated results are shown in Fig. 6.5. As the simulated data in Fig. 6.5 show, the measured power loss ratio tracks the analytic results well, and deviates

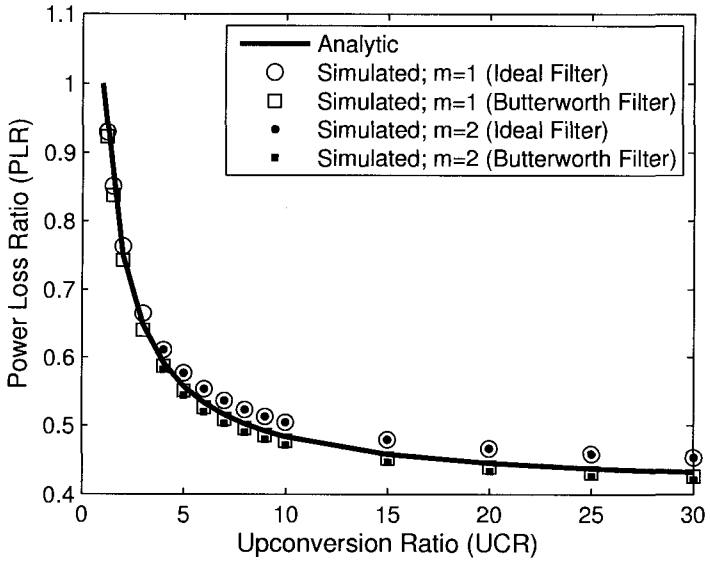


Figure 6.5: Upconverter power loss ratio calculated over a pulse train duration of $2^{15}T$.

slightly at high UCR's. The power loss ratio is slightly lower than the analytic result when the power is measured with a Butterworth filter, while the simulated results are slightly higher than analytic results when the reconstruction filter is ideal. Overall, the variability between simulated and analytic results is small, and confirms that (6.18) is a suitable design equation for evaluating the impact of upconversion on coding efficiency.

6.3 Average Pulse Period

The ATF (f_{av}) of the upconverted pulse train is analyzed by first computing the average pulse period (APP) of the upconverted pulse train. Unlike the $\Sigma\Delta$M pulse train which has a range of different pulse widths corresponding to integer multiples of T, the upconverted pulse train has only two pulse widths: T/m or $T/(2m)$. An example of the modulator pulse period distribution was given in Chapter 2 (see Fig. 2.13), and shows pulse widths ranging from T to $38T$. An upconverted pulse train on the other hand, has pulses of width $T/(2m)$ generated by each transition in $c(t)$, and pulses of width T/m when two successive modulator bits are of opposite sign. The two different pulse widths are easily identified in the

pulse train $p(t)$ shown in Fig. 6.3. The effect of upconversion on APP is to therefore reduce what was formerly a probability mass function of different pulse widths corresponding to multiples of the clock period to only two pulse widths.

An estimate of APP is obtained by assuming the bits in the modulator pulse train are independent with equal probability. Of course, the modulator bits are not independent, but the analysis is simplified significantly by the independence assumption and useful conclusions are obtained.

With reference to Fig. 6.3, the APP is calculated for a period spanning two successive bits. All four bit sequences are equally probable and two sequences, $(+1,-1)$ and $(-1,+1)$, create a transition, while the other two do not. When there is no transition, the pulse burst has only one pulse width, $T/(2m)$, and there are $4m$ pulses per $2T$. On the other hand, when there is a transition, there are $4m-1$ pulses in the burst of which $4m-2$ have a width $T/(2m)$, and one pulse has a width of T/m. The average pulse width in a two bit sequence with a transition is

$$T_t = \frac{4m-2}{4m-1}\left(\frac{T}{2m}\right) + \frac{1}{4m-1}\left(\frac{T}{2}\right) = \frac{2T}{4m-1}. \tag{6.19}$$

The average pulse width for any two bit sequence is then

$$T_{av} = \frac{1}{2}\left(\frac{T}{2m}\right) + \frac{1}{2}\left(\frac{2T}{4m-1}\right) = \frac{(8m-1)T}{4m(4m-1)}. \tag{6.20}$$

Equation (6.20) shows that for large m the APP converges to $T/(2m)$ which is expected. For $m=1$, T_{av} is $\frac{7}{12}T$, and for $m=2$, the APP is nearly halved and is $\frac{15}{56}T$. An important observation about (6.20) is that T_{av} is independent of UCR, or equivalently R_c, while the $\Sigma\Delta$M without upconversion does depend on R_c.

The effect of upconversion on APP is to reduce what was formerly a probability mass function of different pulse widths corresponding to multiples of the clock period to only two pulse widths. An example of the modulator pulse period distribution was given earlier in Fig. 2.13, Chapter 2. Simulated measurements of the APP for the upconverter configuration are compared with the direct configuration in Fig. 6.6. The analytic approximation for APP (6.20) is also shown for the upconversion cases with $m=1$ and $m=2$. The relative insensitivity to UCR is evident in the simulated results for the configuration with upconversion, and the data are well approximated by (6.20).

Since T_{av} is approximately equal to $T/(2m)$, the average pulse period is largest when $m=1$. If the analysis of pulse period is combined with the analysis of coding efficiency

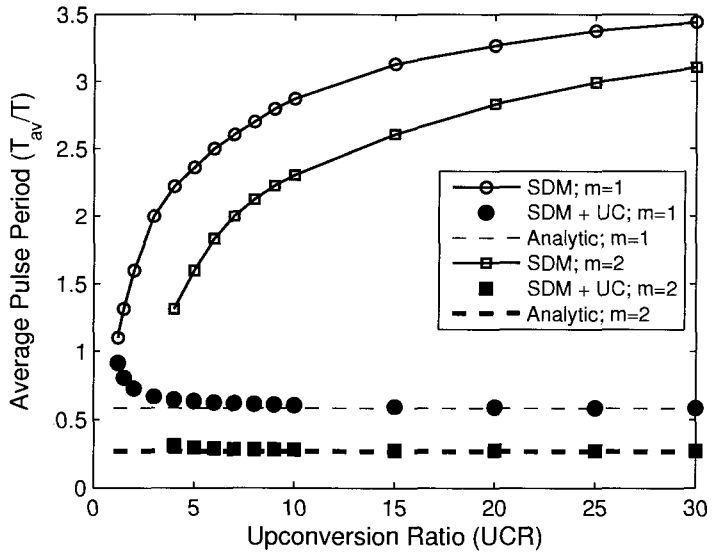


Figure 6.6: Average pulse period (APP) of a fourth order CT bandpass ΣΔM (modulator D) with and without upconversion (UC); the source is a tone at -4.5 dBV.

and PLR, the conclusion is that $m = 1$ is the best overall choice for a RF class D amplifier application.

6.4 Conclusions

A general form of Manchester coding for upconverting a bandpass ΣΔ modulator pulse train was analyzed in this chapter. The analysis focused on evaluating two parameters: coding efficiency and average pulse period. Coding efficiency was analyzed by comparing the coding efficiency of an upconverted pulse train with a direct path configuration. A term called the power loss ratio was constructed, and the power loss ratio was shown to depend only on the upconversion ratio and was independent of m, the number of clock periods used to encode each modulator bit. On the other hand, the analysis showed that the average pulse period depends only on m, and is independent of the upconversion ratio. Considering both factors, it is concluded that $m = 1$ maximizes coding efficiency and minimizes the average pulse period. The equations provide a designer a way to evaluate changes in coding efficiency and average transition frequency depending on the upconversion frequency ratio in the encoder.

Chapter 7

Conclusions and Future Work

An integrated analysis of the bandpass $\Sigma\Delta$ modulator and RF class D amplifier has been presented. The analysis links modulator coding efficiency and the average transition frequency of the synthesized pulse train, with power amplifier efficiency and output power. Load power depends directly on the coding efficiency of the encoded pulse train, and a highly efficient encoder is desirable as this directly affects the size of the class D switches for a specific target load power. The average transition frequency of the pulse train is equally important and directly affects switching losses in the amplifier. Design equations were derived for the RF class D amplifier which integrate coding efficiency and average transition frequency factors. The equations provide first order estimates of the expected amplifier performance, and they are consistent with both periodic and aperiodic pulse trains.

Several chapters in this thesis were devoted to the characterization and analysis of binary amplitude pulse train coding efficiency assuming constrained zero-crossings. The bandpass $\Sigma\Delta$ modulated pulse train transitions are synchronous and timed with a clock edge. The restriction on zero-crossing timing impacts coding efficiency; this was illustrated by the Fourier series analysis of periodic pulse trains. Although the periodic pulse train analysis is a restricted case and applicable to only sinusoidal load signals, it revealed behavior that is observed in bandpass $\Sigma\Delta$ modulators. It is noted that the analysis of the periodic pulse train model motivated the extensive characterization of bandpass $\Sigma\Delta$ modulator coding efficiency, and the characterization shows that some carrier oversample ratios are more optimal than others. Very distinct variation is observed in a variety of second and fourth order modulators, and a carrier oversample ratio of 1.7 was found to be a good compromise between coding efficiency and average transition frequency. One of the next steps in future work shall be to build an experimental modulator and see if the selection of a 'sweet spot' has benefits in an experimental modulator with many other non-ideal characteristics.

Coding efficiency and SNR are fundamental trade-offs in an encoder design. An important example of this is the sixth order modulator design which has significantly better SNR than a fourth order modulator, but the stability of the loop limits coding efficiency, and the reconstructed SNR immediately collapses beyond a certain modulator drive level. The difference in coding efficiency is substantial and the coding efficiency of a fourth order modulator was twice as large as a sixth order modulator. Simulation experiments with a look-ahead tree search in the modulator loop, instead of a binary quantizer, showed that coding efficiency could be boosted in a sixth order design. Furthermore, a sixth order design with significantly better stability should have a smooth SNR collapse similar to second and fourth order designs. An overall conclusion from characterizing bandpass $\Sigma\Delta$ modulators, binary sequence searches, and enhanced modulator designs with tree searches, suggests that a better encoder design might include an adaptive filter which relaxes the selectivity of the noise shaping filter for large signal levels.

An analytic bound on SNR versus coding efficiency assuming linear reconstruction filters is highly desirable. The overload model generated from the periodic signal analysis was taken as a step towards evaluating a bound for sinusoidal load signals. From this model, we were able to prove that the peak amplitude of any sinusoidal load signal cannot exceed $4/\pi$ for a unity amplitude pulse train. An even more general result was proved for a second order bandpass reconstruction filter, where it was shown that the maximum amplitude bound is $4/\pi$ for any pulse train. A comprehensive and rigorous model is desirable that includes a bound for narrowband gaussian sources, and integrates a general class of linear bandpass reconstruction filters. In spite of a comprehensive analytic model, a comparison of bandpass $\Sigma\Delta$ modulation coding efficiency with alternate encoders, such as a binary sequence search, and an enhanced modulator with a tree search, shows that it may be difficult to substantially exceed the performance of bandpass $\Sigma\Delta$ modulation assuming constrained zero-crossings. Better coding efficiency may be obtained by relaxing the zero-crossing constraint, although the implementation of the encoder and filter reconstruction would need to be considered.

The concluding work in this research project was an evaluation of Manchester encoding as a means of upconverting a pulse train from a low frequency bandpass $\Sigma\Delta$ modulator. The benefits of employing Manchester encoding are a reduction in the implementation frequency of noise shaping filters, and spectral shaping that enables AC coupling. The analysis focused on evaluating the cost of upconversion in terms of coding efficiency and average transition frequency. The analysis showed that an upconversion sequence of exactly two bits per

modulator bit was most optimal. Also, the modulator should have a large carrier oversample ratio to maximize the upconversion frequency ratio, and thereby minimize upconversion coding efficiency losses.

Groundwork has been laid for experimental work. A CMOS amplifier design incorporating a bandpass $\Sigma\Delta$ modulator and RF class D amplifier in a complementary voltage switched configuration is recommended. An implementation in small scale technology, such as 0.18 μm, could increase the operating frequency of the design compared to simulation results which were shown for 0.5 μm technology. The maximum operating frequency depends significantly on the load power requirement, and low power designs at RF frequencies may be possible. Other RF circuit topologies may be viable by employing Manchester encoding or spectral shaping, and an implementation with pHEMT or MESFET devices would significantly boost the operating range of the amplifier architecture. Since there are many applications for highly efficient RF amplifiers, it is expected that as implementation challenges are resolved, interest in the bandpass $\Sigma\Delta$ modulator and RF class D amplifier architecture will continue to grow.

Appendix A

Second Order Continuous-Time Bandpass $\Sigma\Delta$ Modulator

The design of a second order continuous-time bandpass $\Sigma\Delta$ modulator is given in this appendix. The modulator is called design E, and transforms to an equivalent second order discrete-time modulator called design B (cf. section 2.2.1). The modulator has a tunable carrier oversample ratio, and includes delay compensation for a general rectangular feedback DAC pulse $D_f(s)$ (cf. Fig. 2.1).

A.1 s to z Plane Mapping

One way to define equivalence between the continuous-time modulator and the discrete-time modulator is to match impulse responses at sampling instants. With reference to Fig. 2.1, the CT feedback loop consists of the DAC transfer function $D_f(s)$ and the loop filter $H(s)$, and the impulse response of the cascade must match the DT transfer function $H(z)$. The equivalence is expressed as

$$\mathcal{L}^{-1}\{D_f(s)H(s)\} \stackrel{t=kT}{\longrightarrow} \mathcal{Z}^{-1}\{H(z)\} \tag{A.1}$$

where the inverse Laplace transform of the CT transfer function is equal to the inverse z-transform at $t = kT$. The transformation, also called the pulse transformation [45], assumes an ideal sampler with period T follows the cascade. The complex mapping from the s plane to z plane is obtained with the transformation $z = e^{sT}$.

Equation (A.1) can also be written as

$$\int_0^\infty d_f(\tau)h(\tau - t)\delta(t - kT)d\tau = h[k] \tag{A.2}$$

and shows that the sampling must follow the convolution; in other words, $d_f(kT)h(kT) \neq h[k]$, where the discrete variable is indicated with square brackets. Consequently, the equivalence mapping must be applied to the cascade of both the feedback DAC and loop filter.

A.2 Transformation with a Rectangular DAC Pulse

Assume the feedback DAC is a rectangular pulse of width τ_w and delayed by τ_d. The delayed DAC pulse is

$$d_f = u_s(t - \tau_d) - u_s(t - \tau_d - \tau_w) \tag{A.3}$$

and the Laplace transform of the pulse is

$$D_f(s) = e^{-s\tau_d} \frac{1 - e^{-s\tau_w}}{s}. \tag{A.4}$$

If the continuous-time loop filter has single order zeros with a partial fraction expansion of

$$H(s) = \sum_{i=1}^{N} \frac{\gamma_i}{s - \lambda_i} \tag{A.5}$$

then

$$D_f(s)H(s) = \sum_{i=1}^{N} \frac{e^{-s\tau_d}(1 - e^{-s\tau_w})\gamma_i}{(s - \lambda_i)s} = \sum_{i=1}^{N} \frac{\gamma_i}{\lambda_i} e^{-s\tau_d}(1 - e^{-s\tau_w})\left(\frac{1}{s - \lambda_i} - \frac{1}{s}\right). \tag{A.6}$$

Focusing on the transformation of a single pole λ_i, the contribution to the continuous-time impulse response is

$$[d_f(t) * h(t)]_i = \begin{cases} 0 & t < \tau_d \\ \frac{\gamma}{\lambda_i}(e^{\lambda_i(t-\tau_d)} - 1) & \tau_d \leq t < \tau_d + \tau_w \\ \frac{\gamma}{\lambda} e^{-\lambda\tau_d}(1 - e^{-\lambda\tau})e^{\lambda t} & t \geq \tau_d + \tau_w \end{cases} \tag{A.7}$$

Assuming $\tau_d + \tau_w \leq T$, the equivalent discrete-time impulse function corresponding to the i'th pole is

$$h_i[n] = \zeta_i \, p_i^{n-1} \, u_s[n - 1] \tag{A.8}$$

where $p_i = e^{\lambda_i T}$ and ζ_i is equal to

$$\zeta_i = \frac{\gamma_i}{\lambda_i} e^{\lambda_i T}(1 - e^{-\lambda_i \tau_w}) e^{-\lambda_i \tau_d}. \tag{A.9}$$

Equation (A.9) shows how a discrete-time pole p_i with residue ζ_i is transformed to a continuous-time pole λ_i with residue γ_i. The transform is now used to map the second order tunable discrete-time modulator, modulator B, to a continuous-time equivalent.

A.3 Second Order Modulator Transfer Function

The discrete-time second order transfer functions were given in equation (2.12). For the second order modulator we focus on transforming only the feedback loop function $H_f(z)$. A special input transfer function $H_i(s)$ could also be implemented to bandlimit the source signal, but in this design we set $H_i(s) = -H_f(s)$.

The discrete-time transfer function $H_f(z)$ is rewritten as a partial fraction expansion

$$H_f(z) = \frac{\zeta}{z - p} + \frac{\zeta^*}{z - p^*} \tag{A.10}$$

where $p = e^{j\pi/R_c}$ and $\zeta = -p/2$. The expansion has a pair of conjugate poles, since $h_f[k]$ is real. The corresponding impulse response is

$$h_f[k] = -\cos(k\pi/R_c) u_s[n-1]. \tag{A.11}$$

With (A.9), and expressions for p and ζ, we get:

$$\lambda = j\pi/(R_c T) \tag{A.12}$$

$$\gamma = \frac{-\pi \exp\left[j\pi(\tau_d + \tau_w/2)/(R_c T)\right]}{4R_c T \sin\left[\pi\tau_w/(R_c T)\right]}. \tag{A.13}$$

Combining the conjugate poles and residues, yields the desired continuous-time function

$$H_f(s) = \frac{-2\Omega_c}{4\sin(\Omega_c \tau_w/2)} \cdot \frac{\cos[\Omega_c(\tau_d + \tau_w/2)]\, s - \Omega_c \sin[\Omega_c(\tau_d + \tau_w/2)]}{s^2 + \Omega_c^2}. \tag{A.14}$$

where $\Omega_c = \pi/(R_c T)$. The expression is valid for $\tau_d + \tau_w \leq T$ and covers feedback configurations such as NRZ ($\tau_d = 0$, $\tau_w = T$), return-to-zero ($\tau_d = 0$, $\tau_w = T/2$), and delayed return-to-zero ($0 \leq \tau_d \leq T/2$, $\tau_w = T/2$) configurations. The delay compensation is important in implementations, and compensates for physical delay arising from the layout of the circuit. The results shown in this work use the NRZ configuration, but any configuration consistent with the design assumptions matches the underlying discrete-time prototype.

Appendix B

MOSIS BSIM3 Model Parameters

The following BSIM3 model files for IBM's 5HP SiGe BiCMOS process are publicly available from the MOSIS website [63]. The following model data has been reprinted with permission from MOSIS.

The NMOS BSIM3v3 model:

```
.MODEL CMOSN NMOS (                                                 LEVEL   = 49
+VERSION = 3.1            TNOM    = 27              TOX     = 9.3E-9
+XJ      = 1.5E-7         NCH     = 1.7E17          VTH0    = 0.5711857
+K1      = 0.8580719      K2      = -0.0532947      K3      = -1.9862763
+K3B     = 0.4939298      W0      = 1E-8            NLX     = 2.203657E-7
+DVT0W   = 0              DVT1W   = 0               DVT2W   = 0
+DVT0    = 0.5812673      DVT1    = 0.2558206       DVT2    = -0.2635266
+U0      = 430.1140074    UA      = 1.412819E-13    UB      = 1.348015E-18
+UC      = 1.810789E-11   VSAT    = 1.164741E5      A0      = 0.7106364
+AGS     = 0              B0      = 1.595763E-8     B1      = 7.54285E-7
+KETA    = 5.662343E-3    A1      = 3.046233E-3     A2      = 0.518668
+RDSW    = 886.6087295    PRWG    = -2.48676E-14    PRWB    = -0.0536069
+WR      = 1              WINT    = 1.007498E-7     LINT    = 8.816408E-8
+DWG     = -5.98056E-9    DWB     = -5.160179E-9    VOFF    = -0.15
+NFACTOR = 2.5            CIT     = 0               CDSC    = 2.4E-4
+CDSCD   = 0              CDSCB   = 0               ETA0    = 0.5
+ETAB    = -0.2           DSUB    = 1.3234467       PCLM    = 1.4594414
+PDIBLC1 = 1.888868E-3    PDIBLC2 = 3.814985E-3     PDIBLCB = -0.1003047
+DROUT   = 0.0781941      PSCBE1  = 3.553249E8      PSCBE2  = 7.276963E-6
+PVAG    = 5.731754E-3    DELTA   = 0.01            RSH     = 4.9
+MOBMOD  = 1              PRT     = 0               UTE     = -1.5
+KT1     = -0.11          KT1L    = 0               KT2     = 0.022
+UA1     = 4.31E-9        UB1     = -7.61E-18       UC1     = -5.6E-11
+AT      = 3.3E4          WL      = 0               WLN     = 1
+WW      = 0              WWN     = 1               WWL     = 0
+LL      = 0              LLN     = 1               LW      = 0
+LWN     = 1              LWL     = 0               CAPMOD  = 2
+XPART   = 0.5            CGDO    = 3.56E-10        CGSO    = 3.56E-10
+CGBO    = 8E-9           CJ      = 8.705838E-4     PB      = 0.8
+MJ      = 0.380292       CJSW    = 1E-12           PBSW    = 0.8
+MJSW    = 0.9100091      CJSWG   = 1.64E-10        PBSWG   = 0.8
+MJSWG   = 0.9100091      CF      = 0               PVTH0   = 2.750475E-3
+PRDSW   = -113.3047521   PK2     = -7.091507E-3    WKETA   = -6.114503E-3
+LKETA   = -0.0360384      )
```

The PMOS BSIM3v3 model:

```
.MODEL CMOSP PMOS (                                LEVEL   = 49
+VERSION = 3.1          TNOM    = 27               TOX     = 9.3E-9
+XJ      = 1.5E-7       NCH     = 1.7E17           VTH0    = -0.517194
+K1      = 0.8352946    K2      = -0.1             K3      = 0
+K3B     = 3.5528102    W0      = 1E-8             NLX     = 5.686989E-9
+DVTOW   = 0            DVT1W   = 0                DVT2W   = 0
+DVT0    = 1.4102179    DVT1    = 0.5535835        DVT2    = -0.3
+U0      = 136.9497274  UA      = 1.72401E-9       UB      = 1E-21
+UC      = 2.926207E-12 VSAT    = 2E5              A0      = 0.4829056
+AGS     = 0.1826775    B0      = 1.952907E-6      B1      = 5E-6
+KETA    = -0.0195473   A1      = 0                A2      = 0.6016027
+RDSW    = 3E3          PRWG    = -0.0417495       PRWB    = -0.1055256
+WR      = 1            WINT    = 8.2492E-8        LINT    = 1.078283E-7
+DWG     = -1.237077E-8 DWB     = 1.429735E-8      VOFF    = -0.0361606
+NFACTOR = 0.8753188    CIT     = 0                CDSC    = 2.4E-4
+CDSCD   = 0            CDSCB   = 0                ETA0    = 0
+ETAB    = -0.0685902   DSUB    = 1                PCLM    = 1.3089385
+PDIBLC1 = 0.037445     PDIBLC2 = 3.79828E-3       PDIBLCB = -0.1
+DROUT   = 0.2789967    PSCBE1  = 5.066296E9       PSCBE2  = 5E-10
+PVAG    = 0.014985     DELTA   = 0.01             RSH     = 2.9
+MOBMOD  = 1            PRT     = 0                UTE     = -1.5
+KT1     = -0.11        KT1L    = 0                KT2     = 0.022
+UA1     = 4.31E-9      UB1     = -7.61E-18        UC1     = -5.6E-11
+AT      = 3.3E4        WL      = 0                WLN     = 1
+WW      = 0            WWN     = 1                WWL     = 0
+LL      = 0            LLN     = 1                LW      = 0
+LWN     = 1            LWL     = 0                CAPMOD  = 2
+XPART   = 0.5          CGDO    = 3.65E-10         CGSO    = 3.65E-10
+CGBO    = 8E-9         CJ      = 8.435277E-4      PB      = 0.6
+MJ      = 0.2044972    CJSW    = 8E-13            PBSW    = 0.7
+MJSW    = 0.9900098    CJSWG   = 6.4E-11          PBSWG   = 0.7
+MJSWG   = 0.9900098    CF      = 0                PVTH0   = 5.98016E-3
+PRDSW   = 14.8598424   PK2     = 3.73981E-3       WKETA   = 7.455348E-5
+LKETA   = -5.006694E-4  )
```

Appendix C

Sinusoidal Amplitude Bound

We want to show that the amplitude of a harmonic at $f_c = N/M$ is no greater than $4/\pi$ assuming $\Delta_p = 1$. The FS coefficient is given in (4.13) and the goal is to find the sequences $\{a_k\}$ that maximize $r_{mN} = 2|c_{mN}|$. We first focus on maximizing the term $|g_M|$, since this is the only term which is a function of a_k.

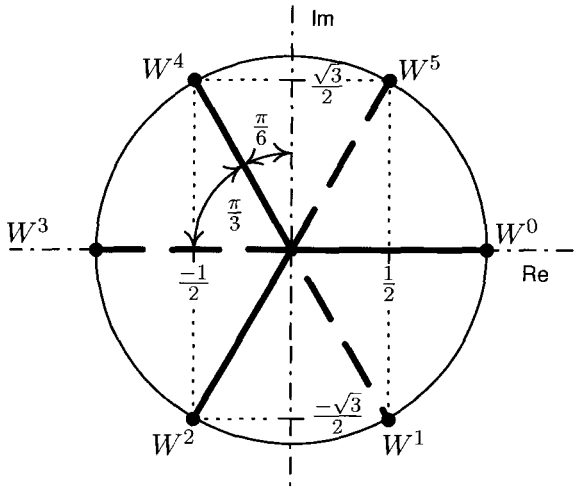
As the examples in section 4.1.2 show, the bounding polygon of the signal constellation is defined completely by the shortest sequence $m = 1$, and longer sequences increase the resolution of reconstruction points either in or on the bounding polygon. Therefore, without loss of generality, assume $m = 1$ which limits a_k to the binary set $\{-1, +1\}$, and the maxima of $|c_{mM}|$ correspond to the vertices of the bounding polygon.

As stated earlier, if $f_c < f_s/2$, then $M > 2N$ where M and N are positive integers and M takes values from the set $S = \{3, 4, 5, \ldots\}$. Divide S into three sets:

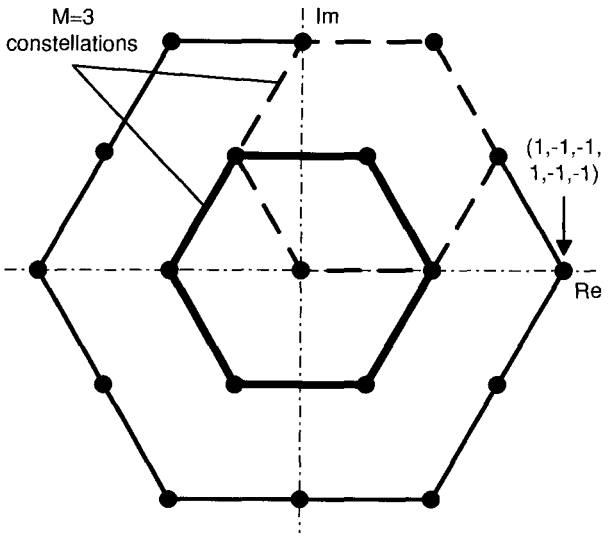
1. $S_1 = \{M \in S : \text{mod}(M, 2) \neq 0\}$
2. $S_2 = \{M \in S : \text{mod}(M, 4) = 0\}$
3. $S_3 = \{M \in S : S \cap (S_1 \cup S_2)\}$.

The set partitioning simplifies the evaluation of the maxima of $|g_M|$ and distinguishes two features. First, the set S_1 has only odd values of M, while sets S_2 and S_3 have even values of M. Second, the roots of unity (W^k) for $M \in S_2$ include $\pm j$, while $M \in S_1$ or $M \in S_3$ do not.

When $M \in S_1$, M is odd and the bounding polygon for all sequences a_k has $2M$ vertices. Examples for $M = 3$ and $M = 5$ where given earlier and show this feature. The $2M$ vertices are generated because there are M base vectors W^k and a sign reversal of each vector generates an additional M vectors. The other feature of $M \in S_1$ is that a vertex lies on



(a) Vectors W^k for $M = 6$.



(b) $M = 6$ reconstruction signal constellation for $m = 1$.

Figure C.1: The $M = 6$ signal constellation generated from the superposition of scaled and translated $M = 3$ constellations shown in Fig. 4.2.

the positive real axis which is exploited for computing the vertex radius r_v. By choosing a sequence a_k which maximizes the projection of the vectors W^k on the positive real axis, we get

$$\begin{aligned}\max|g_M| &= \frac{1}{M}\left[2\sum_{k=1}^{n}\cos(k\pi/M) - 1\right] \\ &= \frac{1}{M}\left\{\frac{\sin(n\pi/M)[1+\cos(\pi/M)]}{\sin(\pi/M)} - \cos(n\pi/M)\right\} \qquad \text{(C.1)}\end{aligned}$$

where $n = (M+1)/2$.

For $M \in S_3$, M has the characteristic that $\max|g_M| = \max|g_{M/2}|$. The members of S_3 are even and have the property that the base vectors are the union of two sets $\tilde{W}^k$ and $-\tilde{W}^k$ where $\tilde{W}^k$ are the base vectors for $M/2 \in S_1$. The implication of this composition is that the bounding polygon of $M \in S_3$ has M vertices, the same number of vertices as $M/2 \in S_1$.

An example is shown in Fig. C.1(a) for $M = 6$; the six base vectors W^k are equal to the base vectors of $M = 3$ plus the reflections of the $M = 3$ vectors. The bounding polygon for $M = 3$ and $M = 6$ is a hexagon and the difference is the density of the reconstruction points in the constellation. As shown in Fig. C.1(b), the signal constellation of $M = 6$ can be visualized as the superposition and translation of $M = 3$ constellations. The $M = 6$ bounding polygon is generated by $\sum_{k=0}^{5} a_k W^k$ and has a vertex radius twice as long as the $M = 3$ bounding polygon generated by $\sum_{k=0}^{2} \tilde{a}_k \tilde{W}^k$. However, the function g_M as defined in (4.12) includes a normalizing factor of $1/M$ and the normalized vertex lengths of the bounding polygons for $M = 3$ and $M = 6$ are identical. We therefore conclude that the maximum of $|g_M|$ for $M \in S_3$ is given by (C.1) with the substitution of $M \rightarrow M/2$.

Unlike S_1 and S_3, when $M \in S_2$ the bounding polygon for g_M does not have vertices on the real axis. Instead, the positive real axis bisects a edge of the polygon. The maximum of $|g_M|$ is computed by projecting the base vectors onto the line $f(r) = \{re^{j\theta/2} : 0 \geq r \leq \infty\}$ as shown in Fig. C.2. θ is defined as the angle subtended by an edge of the polygon given by (4.19). By choosing a sequence a_k which maximizes the projection onto this line we get

$$\begin{aligned}\max|g_M| &= \frac{4}{M}\sum_{k=1}^{M/4}\cos\left[\frac{\pi(2k-1)}{M}\right] \\ &= \frac{2}{\sin(\pi/M)}. \qquad \text{(C.2)}\end{aligned}$$

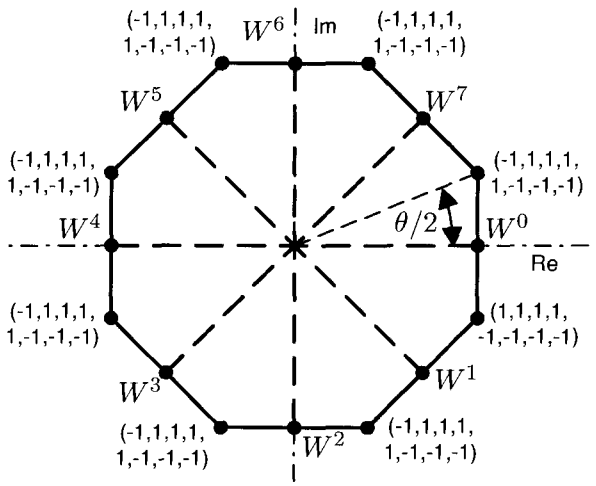


Figure C.2: $M = 8$ signal constellation; a vertex passes through a line at $\theta/2$.

All that remains to find $\max|c_{mN}|$ is to combine (C.1) and (C.2) with the $\text{sinc}(N/M)$ term. Since $M > 2N$, $N/M < \frac{1}{2}$ and the range where a maximum is computed is confined to the main lobe of the $\text{sinc}(\cdot)$ function. Therefore $\text{sinc}(N/M)$ is maximized when $N = 1$, and the maximum vertex radius of the bounding polygon is

$$\max r_v = 2\,\text{sinc}(1/M)\max|g_M|. \tag{C.3}$$

Combining (C.1)-(C.3) yields the expressions:

$$\max r_v = \begin{cases} 4/\pi \cdot \cos\left[\pi/(2M)\right] & M \in S_1 \\ 4/\pi & M \in (S_2 \cup S_3) \end{cases}. \tag{C.4}$$

Appendix D

Unconstrained Coding Theorem for Gaussian Sources

Encoding a source signal into a binary amplitude pulse train with synchronous zero crossings and linear reconstruction filters is a constrained case of more general source coding problems. An example of a more general result is given in the following theorem where there any no constraints on the implementation of either the encoder or decoder. The theorem and a proof can be found in [70].

Theorem 1 *Suppose a discrete-time gaussian source has a two-sided power spectral density given by*

$$S_u(f) = \sum_{k=-\infty}^{+\infty} R_u[k]e^{-j2\pi kf/f_s} \qquad |f| < f_s/2, \tag{D.1}$$

then the information content of the source with respect to a quadratic distortion measure has the following parametric representation:

$$R_b(\theta) = \frac{1}{2}\int_{-f_s/2}^{f_s/2} \max\left[0, \log_2\left(\frac{S_u(f)}{\theta}\right)\right] df \tag{D.2}$$

$$D(\theta) = \int_{-f_s/2}^{f_s/2} \min[0, S_u(f)]df. \tag{D.3}$$

R_b is the bit rate of the encoder and D is the corresponding distortion generated by the encoder.

As an example of the bound which the theorem predicts, consider a bandlimited gaussian source signal with power σ_u^2 distributed uniformly over a bandwidth f_b with a carrier frequency f_c. Let D^* be the required distortion level for a target SNR threshold, then

$$\theta = \frac{D^*}{2f_b}. \tag{D.4}$$

From (D.2) the minimum encoder rate to generate a distortion level D^* is

$$R_b = f_b \log_2 \left(\frac{\sigma_u^2}{D^*} \right). \tag{D.5}$$

The ratio σ_u^2/D^* is the SNR threshold SNR_t of the reconstructed signal and the minimum rate is then expressed as

$$R_b = f_b \log_2(SNR_t). \tag{D.6}$$

Assuming the sample rate R_b is equal to f_s, (D.6) is rearranged to yield

$$R_e \geq \frac{1}{2} \log_2(SNR_t). \tag{D.7}$$

The rate-distortion theorem shows that in an unconstrained case, the reconstructed signal depends only on the envelope oversample ratio — an intuitively satisfying result since the source information is in the lowpass envelope only.

For comparison with bandpass $\Sigma\Delta$ modulation, consider the comparative envelope oversample ratio of the unconstrained case for an SNR threshold of 65 dB. Evaluating (D.7) shows that the theoretical minimum envelope oversample ratio is 10.8 much less than 128 which is typical for a bandpass $\Sigma\Delta$ modulator. As stated earlier, the theorem shows that much more efficient encoding/decoding methods exist, but it does not consider constraints on pulse train amplitude levels, the timing of transitions, or linear signal reconstruction filters.

Appendix E

Derivation of the Manchester Power Loss Equation

The principal steps in the derivation of the PLR equation (6.17) are shown below. The derivation begins with (6.16) and the following steps are taken to convert the sum into a closed form expression.

The term

$$\sin[2\pi(m-\nu_c)a_n] \tag{E.1}$$

is rewritten as

$$(-1)^{m+n+1}\sin[\pi(m-\nu_c)(2n-1)/(2m)]. \tag{E.2}$$

The equivalence of (E.1) and (E.2) is found by enumerating the delay a_n over $n = 1, 2, \ldots m$. The term (E.2) is of the form $\sin(\alpha - \beta)$ and after expanding reduces to

$$(-1)^m \cos[\pi\nu_c(2n-1)/(2m)]. \tag{E.3}$$

The sum term in (6.16) is now

$$\sum_{n=1}^{m}\sin[2\pi(m-\nu_c)a_n] \;=\; (-1)^m\sum_{n=1}^{m}\cos\left[\frac{\pi\nu_c(2n-1)}{2m}\right] \tag{E.4}$$

and has a well known closed form solution of [78]

$$(-1)^m\frac{\sin(\pi\nu_c/2)\cos(\pi\nu_c/2)}{\sin[\pi\nu_c/(2m])} \;=\; \frac{(-1)^m\sin(\pi\nu_c)}{2\sin[\pi\nu_c/(2m)]}. \tag{E.5}$$

With a closed form expression for the sum, the remaining steps are easy. The expression for $|D_m(\nu)|^2$ is

$$|D_m(\nu)|^2 = \left[\frac{T\sin(\pi\nu_c)\cot(\pi\nu_c/(2m))}{\pi(m-\nu_c)}\right]^2 \tag{E.6}$$

and finally the PLR is given by

$$\text{PLR} = \frac{|D_m(\nu)|^2}{|D_{nrz}(\nu)|^2} = \left[\frac{\nu_c}{m-\nu_c}\cot[\pi\nu_c/(2m)]\right]^2. \tag{E.7}$$

Bibliography

[1] P. J. Baxandall, "Transistor sinewave oscillators," *Proc. Inst. Elec. Eng.*, vol. 106, part B, suppl. 16, pp. 748–758, May 1959.

[2] A. Jayaraman, P. F. Chen, G. Hanington, L. Larson, and P. Asbeck, "Linear high-efficiency microwave power amplifiers using bandpass delta-sigma modulators," *IEEE Microwave Guided Wave Lett.*, vol. 8, no. 3, pp. 121–123, Mar. 1998.

[3] H. L. Krauss, C. W. Bostian, and F. H. Raab, *Solid-State Radio Engineering.* John Wiley & Sons, Inc., 1980.

[4] B. F. Logan Jr., "Click modulation," *AT&T Bell Lab Tech. J.*, vol. 63, no. 3, pp. 401–423, Apr. 1984.

[5] M. Streitenberger, F. Felgenhauer, H. Bresch, and W. Mathis, "Zero-position coding (ZePoC) — a generalized concept of pulse-length modulated signals and its application to class-D audio power amplifiers," in *Audio Engineering Society 110th Convention, Preprint (5365)*, May 12–15 2001.

[6] M. Streitenberger and W. Mathis, "A novel coding topology for digital class-D audio power amplifiers with very low pulse-repetition rate," in *Proceedings of the 28th European Solid-State Circuits Conference (ESSCIRC)*, Sept. 24–26 2002, pp. 515–518.

[7] A. J. Magrath and M. B. Sandler, "Hybrid pulse width modulation/sigma-delta modulation power digital-to-analogue converter," *IEE Proc. -Circuits Devices Syst.*, vol. 143, no. 3, pp. 149–156, June 1996.

[8] H. Ballan and M. Declercq, "12V $\Sigma - \Delta$ class-D amplifier in 5V CMOS technology," in *IEEE Custom Integrated Circuits Conference*, May 1–4 1995, pp. 559–562.

[9] V. M. Tousi, F. Sahandi, M. Atarodi, and M. Shojaei, "A 3.3V/1W class D audio power amplifier with 103dB DR and 90% efficiency," in *IEEE International Conference on Microelectronics*, vol. 2, May 12–15 2002, pp. 581–584.

[10] J. F. Jensen, G. Raghavan, A. E. Cosand, and R. H. Walden, "A 3.2-GHz second-order delta-sigma modulator implemented in InP HBT technology," *IEEE J. Solid-State Circuits*, vol. 30, no. 10, pp. 1119–1127, Oct. 1995.

[11] A. Jayaraman, P. Asbeck, K. Nary, S. Beccue, and K.-C. Wang, "Bandpass delta-sigma modulator with 800 MHz center frequency," in *IEEE GaAs IC Symposium Tech. Dig.*, Oct. 12–15 1997, pp. 95–98.

[12] J. F. Jensen and G. Raghavan, "Bandpass delta-sigma modulators for direct IF and RF sampling digital receivers implemented in InP HBT IC technology," in *IEE Conf. on Advanced A/D and D/A Conversion Techniques and their Applications*, July 27–28 1999, pp. 34–37.

[13] H. Kobayashi, J. M. Hinrichs, and P. Asbeck, "Current-mode class-D power amplifiers for high-efficiency RF applications," *IEEE Trans. Microwave Theory Tech.*, vol. 49, no. 12, pp. 2480–2485, Dec. 2001.

[14] A. Long, J. Yao, and S. I. Long, "A 13W current mode class D high efficiency 1 GHz power amplifier," in *Midwest Symposium on Circuits and Systems*, vol. 1, Aug. 4–7 2002, pp. 33–36.

[15] T.-P. Hung, A. G. Metzger, P. J. Zampardi, M. Iwamoto, and P. M. Asbeck, "High efficiency current-mode class-D amplifier with integrated resonator," in *IEEE MTT-S Digest*, June 2004, pp. 2035–2038.

[16] M. Iwamoto, J. Hinrichs, A. Jayaraman, L. E. Larson, and P. M. Asbeck, "A push-pull bandpass $\Delta - \Sigma$ class-S amplifier," *IEEE Topical Workshop on Power Amplifiers for Wireless Communications*, Sept. 11–12 2000.

[17] M. Iwamoto, A. Jayaraman, G. Hanington, P. F. Chen, A. Bellora, W. Thornton, L. E. Larson, and P. M. Asbeck, "Bandpass delta-sigma class-S amplifier," *Electronic Letters*, vol. 36, no. 12, pp. 1010–1011, June 8th 2000.

[18] J. Ketola, J. Sommarek, J. Vankka, and K. Halonen, "Transmitter utilising bandpass delta-sigma modulator and switching mode power amplifier," in *IEEE International Symposium on Circuits and Systems*, Vancouver, Canada, May 23–26 2004, pp. 633–636.

[19] T. Johnson and S. Stapleton, "Available load power in a RF class D amplifier with a sigma-delta modulator driver," in *IEEE Radio and Wireless Conference*, Atlanta, Georgia, Sept. 19–22 2004, pp. 439–442.

[20] S. A. Jantzi, W. M. Snelgrove, and J. Paul F. Ferguson, "A fourth-order bandpass sigma-delta modulator," *IEEE J. Solid-State Circuits*, vol. 28, no. 3, pp. 282–291, Mar. 1993.

[21] S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta-Sigma Data Converters - Theory, Design, and Simulation.* New York: IEEE Press, 1997.

[22] T. Johnson, S. Stapleton, and J. Cavers, "Binary coding for RF class D amplifier applications," in *The Ninth Canadian Workshop on Information Theory*, Montréal, Québec, June 5–8 2005, pp. 74–77.

[23] S. El-Hamamsy, "Design of high-efficiency RF class-D power amplifier," *IEEE Trans. Power Electron.*, vol. 9, no. 3, pp. 297–308, May 1994.

[24] W. J. Chudobiak and D. F. Page, "Frequency and power limitations of class-D transistor amplifiers," *IEEE J. Solid-State Circuits*, vol. SC-4, no. 1, pp. 25–37, Feb. 1969.

[25] F. H. Raab, "Radio frequency pulsewidth modulation," *IEEE Trans. Commun.*, vol. 21, pp. 958–966, Aug. 1973.

[26] ——, "Class-D power amplifier load impedance for maximum efficiency," in *Proc. RF Technology Expo*, Anaheim, CA, Jan. 1985, pp. 287–295.

[27] P. Wagh, P. Midya, P. Rakers, J. Caldwell, and T. Schooler, "An all-digital universal RF transmitter," in *IEEE Custom Integrated Circuits Conference*, Oct. 3–6 2004, pp. 549–552.

[28] H. Kato, "Trellis noise-shaping converters and 1-bit digital audio," *Audio Engineering Society 112th Convention*, no. 5615, May 10–13 2002.

[29] J. A. S. Angus, "Tree based lookahead sigma delta modulators," *Audio Engineering Society 114th Convention*, no. 5825, Mar. 22–25 2003.

[30] E. Janssen and D. Reefman, "Advances in trellis based SDM structures," *Audio Engineering Society 115th Convention*, no. 5993, Oct. 10–13 2003.

[31] A. Thurston, T. H. Pearce, and M. J. Hawksford, "Bandpass implementation of the sigma-delta A-D conversion technique," in *IEE International Conference on Analogue to Digital and Digital to Analogue Conversion*, Sept. 17–19 1991, pp. 81–86.

[32] O. Shoaei and W. M. Snelgrove, "Optimal (bandpass) continuous-time $\Sigma\Delta$ modulator," in *IEEE International Symposium on Circuits and Systems*, vol. 5, May 30–June 2 1994, pp. 489–492.

[33] R. Schreier and M. Snelgrove, "Bandpass sigma-delta modulation," *Electronics Letters*, vol. 25, no. 23, pp. 1560 – 1561, Nov. 9 1989.

[34] S. P. Stapleton, "Class S power amplifiers for RF systems using bandpass delta-sigma modulators," Agilent EEsof Design Seminar, Product Note N3508A, 1996.

[35] A. Dupuy and Y. E. Wang, "High efficiency power transmitter based on envelope delta-sigma modulation (EDSM)," in *IEEE Vehicular Technology Conf. (VTC2004-Fall)*, vol. 3, Sept. 26–29 2004, pp. 2092–2095.

[36] O. Shoaei and M. Snelgrove, "Design and implementation of a tunable 40MHz - 70MHz Gm-C bandpass $\Sigma\Delta$ modulator," *IEEE Trans. Circuits Syst. II*, vol. 44, no. 7, pp. 521–530, July 1997.

[37] J. A. Cherry and W. M. Snelgrove, "On the design of a fourth-order continuous-time LC delta-sigma modulator for UHF A/D conversion," *IEEE Trans. Circuits Syst. II*, vol. 47, no. 6, pp. 518 – 530, June 2000.

[38] T. Johnson and S. Stapleton, "Coding efficiency of sigma-delta modulation for class D amplification in W-CDMA applications," in *Wireless 2004 Proceedings*, Calgary, Alberta, Canada, July 12–14 2004, pp. 288–295.

[39] T. Johnson, R. Sobot, and S. Stapleton, "Measurement of bandpass sigma-delta modulator coding efficiency and pulse transition frequency for RF class D power amplifier applications," in *Canadian Conference of Computer and Electrical Engineering*, Ottawa, Ontario, May 7–10 2006, pp. 2279–2282.

[40] ——, "Manchester encoded bandpass sigma-delta modulation for RF class D amplifiers," *IEE Proc. -Circuits Devices Syst.*, Aug. 2006, (accepted for publication).

[41] T. Johnson and S. Stapleton, "RF class D amplification with bandpass sigma-delta modulator drive signals," *IEEE Trans. Circuits Syst. I*, Aug. 2006, (accepted for publication).

[42] O. Shoaei and W. M. Snelgrove, "A multi-feedback design for LC bandpass Delta-Sigma modulators," in *IEEE International Symposium on Circuits and Systems*, vol. 1, Apr. 28–May 3 1995, pp. 171–174.

[43] W. Gao and W. M. Snelgrove, "A 950-MHz IF second-order integrated LC bandpass delta-sigma modulator," *IEEE J. Solid-State Circuits*, vol. 33, no. 5, pp. 723 – 732, May 1998.

[44] R. Sobot, S. Stapleton, and M. Syrzycki, "Tunable Continuous Time Bandpass Sigma-Delta Modulators with Fractional Delays," *IEEE Trans. Circuits Syst. I*, vol. 53, no. 2, pp. 264–273, Feb. 2006.

[45] B. C. Kuo, *Digital Control Systems*, 2nd ed. Saunders College Publishing, 1992.

[46] R. Schreier, "An empirical study of high-order single-bit delta-sigma modulators," *IEEE Trans. Circuits Syst. II*, vol. 40, no. 8, pp. 461–466, Aug. 1993.

[47] S. A. Jantzi, C. Ouslis, and A. Sedra, "Transfer function design for $\Delta\Sigma$ converters," in *IEEE International Symposium on Circuits and Systems*, vol. 5, May 30–June2 1994, pp. 433 – 436.

[48] W. L. Lee, "A novel higher order interpolative modulator topology for high resolution oversampling A/D converters," Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, 1987.

[49] J. A. Cherry and W. M. Snelgrove, "Excess loop delay in continuous-time delta-sigma modulators," *IEEE Trans. Circuits Syst. II*, vol. 46, no. 4, pp. 376 – 389, Apr. 1999.

[50] S. H. Ardalan, "Analysis of delta-sigma modulators with bandlimited gaussian inputs," in *International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, 1988, pp. 1866–1869.

[51] W. A. Gardner, *Introduction to Random Processes With Applications to Signals and Systems*, 2nd ed. McGraw-Hill Inc., 1990.

[52] K. S. Shanmugan and A. M. Breipohl, *Random Signals: Detection, Estimation and Data Analysis.* John Wiley & Sons, Inc., 1988.

[53] (2003) Technical Specification 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Base Station (BS) conformance testing (FDD) (release 6). 25141-640.doc. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/25-series.htm

[54] M. Berkhout, "An integrated 200-W class-D audio amplifier," *IEEE J. Solid-State Circuits*, vol. 38, no. 7, pp. 1198–1206, July 2003.

[55] S. C. Cripps, *RF Power Amplifiers for Wireless Communications.* Artech House, Inc., 1999.

[56] W. R. Curtice, "GaAs MESFET modeling and nonlinear CAD," *IEEE Trans. Microwave Theory Tech.*, vol. 36, no. 2, pp. 220–230, Feb. 1988.

[57] I. Angelov, H. Zirath, and N. Rorsman, "A new empirical nonlinear model for HEMT and MESFET devices," *IEEE Trans. Microwave Theory Tech.*, vol. 40, no. 12, pp. 2258–2266, Dec. 1992.

[58] H. Statz, P. Newman, I. W. Smith, R. A. Pucel, and H. A. Haus, "GaAs FET device and circuit simulation in SPICE," *IEEE Trans. Electron Devices*, vol. ED-34, no. 2, pp. 160–169, Feb. 1987.

[59] V. I. Cojocaru and T. J. Brazil, "A scalable general-purpose model for microwave FET's including DC/AC dispersion effects," *IEEE Trans. Microwave Theory Tech.*, vol. 45, no. 12, pp. 2248–2255, Dec. 1997.

[60] (2002) Multi-carrier WCDMA basestation design considerations - amplifier linearization and crest factor control. 2021396_006227.pdf. [Online]. Available: http://www.pmc-sierra.com/pressRoom/whitePapers.html

[61] C.-J. Wei, Y. A. Tkachenko, and D. Bartle, "An accurate large-signal model of GaAs MESFET which accounts for charge conservation, dispersion, and self-heating," *IEEE Trans. Microwave Theory Tech.*, vol. 46, no. 11, pp. 1638–1644, Nov. 1998.

[62] B. E. Taylor, Ed., *Power MOSFET Design.* W. Sussex, England: Chichester, 1993.

[63] (2005) IBM 5HP wafer electrical test data and SPICE model parameters. t51f_5hp-params.txt. [Online]. Available: http://www.mosis.org

[64] S.-M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits: Analysis and Design*, 3rd ed. McGraw-Hill Companies, Inc., 2003.

[65] R. Sobot, "Design methodology for continuous-time bandpass sigma-delta modulators," Ph.D. dissertation, Simon Fraser University, Burnaby, BC, Canada, 2005.

[66] F.-J. Huang and K. O, "A 0.5-μm CMOS T/R switch for 900-MHz wireless applications," *IEEE J. Solid-State Circuits*, vol. 36, no. 3, pp. 486–492, Mar. 2001.

[67] D. A. Hodges, H. G. Jackson, and R. A. Saleh, *Analysis and Design of Digital Integrated Circuits In Deep Submicron Technology*, 3rd ed. New York, NY: The McGraw-Hill Companies, Inc., 2004.

[68] MAPLE (version 9). 615 Kumpf Drive, Waterloo, Ontario, Canada N2V 1K8. [Online]. Available: http://www.maplesoft.com

[69] D. Reefman, J. Reiss, E. Janssen, and M. Sandler, "Description of limit cycles in sigma-delta modulators," *IEEE Trans. Circuits Syst. I*, vol. 52, no. 6, pp. 1211–1223, June 2005.

[70] R. E. Blahut, *Principles and Practice of Information Theory.* Addison-Wesley Publishing Company, 1987.

[71] J. B. Anderson and S. Mohan, "Sequential coding algorithms: A survey and cost analysis," *IEEE Trans. Commun.*, vol. COM-32, no. 2, pp. 169–176, Feb. 1984.

[72] A. Mobasher, M. Taherzadeh, R. Sotirov, and A. K. Khandani, "A near maximum likelihood decoding algorithm for MIMO systems based on semi-definite programming," in *IEEE International Symposium on Information Theory (ISIT'05)*, Sept. 4–9 2005, pp. 1686–1690.

[73] K. Thomas, R. S. Rana, and L. Yong, "A 1GHz CMOS fourth-order continuous-time bandpass sigma delta modulator for RF receiver front end A/D conversion," in *Asia and South Pacific Design Automation Conference (ASP-DAC)*, vol. 2, no. 12, Jan. 18–21 2005, pp. 665 – 670.

[74] H. Tao and J. M. Khoury, "A 400-Ms/s frequency translating bandpass sigma-delta modulator," *IEEE Trans. Circuits Syst. I*, vol. 34, no. 12, pp. 1741 – 1752, Dec. 1999.

[75] R. Sobot, S. Stapleton, and M. Syrzycki, "Fractional delay sigma-delta upconverter," *Electronics Letters*, vol. 41, no. 23, pp. 1267–1268, 10 Nov. 2005.

[76] L. W. Couch, *Digital and Analog Communication Systems.* Prentice-Hall Inc., New Jersey, 2001.

[77] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, Inc., 1991.

[78] H. F. Davis, *Fourier Series and Orthogonal Functions.* New York: Dover Publications, Inc., 1963.