A High-Speed CMOS A/D Converter Employing Variable Nonuniform Quantization

by

Rod Zavari

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science
Department of Electrical and Computer Engineering
University of Toronto

© Copyright by Rod Zavari 1998
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.
A High-Speed CMOS A/D Converter Employing Variable Nonuniform Quantization

Rod Zavari
Department of Electrical and Computer Engineering
University of Toronto
Master of Applied Science, 1998

ABSTRACT

In the receiver of a digital communication system, the use of an analog-to-digital converter (ADC) allows many useful functions be implemented digitally. As the speeds of digital communication systems rise, so does the speed associated with the corresponding ADC. The quantization noise introduced by the ADC is dependent on the number of bits of the ADC and needs to be kept within system specifications. However, every additional bit added to a high-speed ADC complicates its design and increases its power, area and input capacitance. In this thesis, nonuniform quantization is investigated to try to lower the quantization noise introduced. An ADC employing variable nonuniform quantization has been implemented in a 0.5 μm digital CMOS process.
I would like to thank my thesis supervisor, Professor David Johns, for his guidance and advice. His suggestions and comments for this thesis were invaluable. His willingness to always discuss problems and answer questions is greatly appreciated.

Thanks to my peers in EA104. Thanks to Jasmine Cheng, Kasra Ardalan and Khoman Phang for their technical advice and assistance. Their clever and bright comments were very useful. Thanks to Marcus van Ierssel for his help with the design of the test board. Also, thanks to Bob Richens for all his help in the lab.

Thanks to my family and friends for their support and encouragement. Special thanks to Joanne Papanicolaou for her understanding and comfort in all the challenging tasks met during the course of my graduate studies.

Also, thanks to Micronet for their financial support and Canadian Microelectronics Corporation for providing fabrication services for my design.
# TABLE OF CONTENTS

ABSTRACT .................................................................................................................................................. ii

ACKNOWLEDGEMENTS ............................................................................................................................ iii

TABLE OF CONTENTS .................................................................................................................................. iv

CHAPTER 1: Introduction ............................................................................................................................ 1
  1.1 Motivation ........................................................................................................................................... 1
  1.2 A PAM Digital Communication System .......................................................................................... 2
  1.3 Summary and Thesis Outline ........................................................................................................... 7

CHAPTER 2: Nonuniform Quantization ..................................................................................................... 8
  2.1 The Analog-to-Digital Converter ....................................................................................................... 8
  2.2 Uniform Quantization ......................................................................................................................... 9
  2.3 Quantization Power ............................................................................................................................ 11
  2.4 Optimal Quantization ....................................................................................................................... 13
  2.5 Nonuniform Quantization in a 4-PAM System ............................................................................... 15
    2.5.1 Possible Optimal Quantization ............................................................................................... 15
    2.5.2 Fine-Coarse Quantization ........................................................................................................ 16
  2.6 Simulations Involving the Nonuniform ADC .................................................................................. 20
    2.6.1 Simulation Setup ....................................................................................................................... 20
    2.6.2 Simulation Results ................................................................................................................... 23
  2.7 Variable Nonuniform Quantization .................................................................................................. 27
  2.8 Comparison to Lloyd-Max Quantizer ............................................................................................... 27
  2.9 Digital Linearization ......................................................................................................................... 29
  2.10 Summary ......................................................................................................................................... 31

CHAPTER 3: Circuit Design ....................................................................................................................... 32
  3.1 High-Speed ADC Architectures ...................................................................................................... 32
  3.2 Interpolation ....................................................................................................................................... 34
  3.3 Overall Circuit Architecture ......................................................................................................... 39
1.1 Motivation

The use of an analog-to-digital converter (ADC) in the receiver end of a digital communication system provides certain advantages. It allows the incoming signal to be digitized, allowing the equalization, clock recovery and echo cancellation to be implemented digitally. However, as data rates continue to rise, the sampling speeds of the ADC also continues to rise.

The design of high-speed ADCs is non-trivial. First, due to the parallel structure of high-speed ADCs, every additional bit of resolution added to the ADC, roughly doubles the size, power and input capacitance of the ADC. Complex circuit architectures have been proposed to try to reduce some of these disadvantages. Nevertheless, every additional bit of resolution added to the ADC will lower its quantization noise and thus improve its dynamic
range. Also, as the sampling speeds and input bandwidths are continuously increased, the circuits within the ADC are required to perform faster and maintain certain performance. This becomes difficult to achieve since their performance degrades at high-speeds and affects the dynamic range of the ADC.

Also, the integration of the ADC with the remainder of the system requires it to be implemented in a standard digital CMOS process. Implementing the ADC in a digital CMOS technology provides an economic advantage and a more compact system. However, this further complicates the design of the ADC since traditionally, most high-speed ADCs were implemented in a bipolar process. Recently, proposed high-speed CMOS ADCs operating above 100MS/s have been reported in [Steyaert, 1993][Flynn, 1996] with the fastest reported in [Roovers, 1996] at 175 MS/s.

Thus, we see that any improvement in the dynamic range of a high-speed CMOS ADC can be rather beneficial. There are many advantages to try to find techniques to improve the performance of an ADC used in the receiver of a digital communication system. This thesis explores the use of nonuniform quantization in the ADC of a digital communication system to see how it affects its dynamic range within the system. A 5-bit, 150 MS/s ADC employing nonuniform quantization was implemented in a 0.5 \( \mu \)m digital CMOS technology.

1.2 A PAM Digital Communication System

Let us consider a PAM digital communication system. We will be primarily concerned with
the transmission and reception of symbols. First, consider the transmitter in Figure 1.1. The

![Transmitter Diagram]

Figure 1.1 Transmitter of a digital communication system.

input symbols, $A_k$, represent one of $N$ levels. This symbol represents $\log_2 N$ bits. The symbol is shaped by a transmit filter to generate the pulse amplitude modulated (PAM) representation of the symbols. We will call the transmit signal the PAM signal and the input symbols, the PAM symbols. The PAM signal is sent for transmission down the channel. A conceptual representation of a 4-PAM transmitter with square pulses is shown in Figure 1.2.

![4-PAM Diagram]

Figure 1.2 A transmit filter with a rectangular impulse response and a possible signal constellation corresponding to the 4-PAM system.

Three possible front-end receiver configurations for a PAM digital communication
system are shown in Figure 1.3. The channel adds noise to the PAM signal sent and can be

Figure 1.3 Three possible PAM receiver architectures: (a) analog implementation, (b) digital implementation and (c) mixed-signal implementation.

modeled as additive white Gaussian noise (AWGN). The task of our receiver is to determine
which symbol was transmitted every symbol period by the transmitter with a minimum amount of error. At the input of the receiver, the PAM signal has two sources of error on it, AWGN and intersymbol interference (ISI). The ISI is due to the transfer response of the channel. The channel has finite bandwidth and will alter the shape of the PAM signal causing it to have ISI. The receiver will attempt to eliminate ISI through equalization while trying to keep the noise small during this process.

There are three receiver architectures shown in Figure 1.3. Figure 1.3(a) shows a full analog implementation. In this case, an analog equalizer will correct for the distortion introduced by the channel. This approach requires equalizing for a large bandwidth and then slicing to determine the sent symbol. Unfortunately, this approach causes a lot of noise enhancement due to the boost of the equalizer. This could seriously degrade the signal-to-noise ratio (SNR) of the received signal and thus the symbol error rate (SER). Another disadvantage of this architecture is the difficulty in designing an analog equalizer as well as its susceptibility to DC offsets [Johns, 1997.2]. However, analog equalization has reduced power and size over its digital counterpart.

The second architecture is a full digital implementation and is shown in Figure 1.3(b). Here, the signal is immediately passed through an ADC and all equalization is done in the digital domain. The digital equalization is broken into a feed-forward equalizer (FFE) and a decision-feedback equalizer (DFE). The FFE deals with pre-cursor ISI and the DFE removes post-cursor ISI.

The DFE has simpler multiplications since, although not shown here, it deals with the estimated symbols as its input signal (the DFE is in a feedback loop). The estimated symbols will usually have a low number of bits (2-bits for the 4-PAM case). However, the FFE requires full multiplications since it deals with digital signals with more bits, depending on the resolution of the ADC [Johns, 1997.2]. Thus, the FFE's implementation is inefficient and power hungry, particularly for a large number of taps and higher sampling rates. It should also
be noted that the FFE boosts noise as does the analog equalizer. However, a DFE is advantageous in that it reduces noise enhancement since the input to the DFE is noise free (assuming the output decisions are correct) [Johns, 1997.2].

The third architecture, shown in Figure 1.3(c), is based on a mixed-signal implementation. Here, we partially equalize the incoming signal in the analog domain. The analog equalizer will remove pre-cursor ISI and partially reduce the post-cursor ISI. Any remaining ISI is removed in the digital domain by a DFE. To bridge these two domains, an ADC is needed.

This method has numerous advantages. In comparison to the first architecture, there is less noise boost, since we are partially equalizing with the analog equalizer and we leave some equalization for a DFE, which reduces noise enhancement. Also, since the analog equalizer will operate for a lower range of frequencies, it makes the design of the analog equalizer simpler. This will be particularly important in a CMOS process where performance for analog circuits at high-speeds is difficult to achieve. The magnitude response of possible analog equalizers for the first architecture and third architecture is shown in Figure 1.4.

![Figure 1.4 The magnitude response of possible analog equalizers.](image)

Compared with the second architecture, the third architecture will eliminate the need
for an FFE which can be power hungry as explained above. We will also greatly reduce the numbered taps required in the DFE by removing some of the post-cursor ISI, causing additional savings in silicon area and power.

Also, since we will be in the digital domain due to the ADC, we can perform digital functions. This could include clock recovery and an echo canceler for a full-duplex system.

The system that we will be concerned with in this thesis will be the third architecture. We will concentrate our efforts on the ADC in this architecture. We will assume we have a 4-PAM signal with a symbol rate of 155.5 MS/s corresponding to 311 Mb/s. This means that the ADC needs to operate at 155.5 MS/s.

1.3 Summary and Thesis Outline

For the remainder of the thesis, we will be considering the ADC in the third architecture of the 4-PAM system. We will explore nonuniform quantization within the ADC to try to improve its performance and dynamic range within the system. Improved dynamic range of the ADC can benefit the system in many ways. It could improve the performance of the system as a whole, it could ease the requirements of the ADC or could allow for a lower resolution ADC to be designed.

An outline of this thesis is as follows. Chapter 1 has provided an overview of possible receiver architectures and how the ADC fits into the picture. Chapter 2 presents the idea of nonuniform quantization, how it can be used for our application and simulation results. Chapter 3 describes the circuits used to design the ADC. Chapter 4 discusses the layout and implementation of the ADC circuit. Chapter 5 concludes the thesis with discussions and recommendations for future work in the area.
CHAPTER 2

Nonuniform Quantization

2.1 The Analog-to-Digital Converter

The inclusion of the ADC in the receiver of the digital communication system is to convert our signal to the digital domain. If the symbol rate of the system is at 155.5 MS/s, then the ADC also needs to operate at this speed. The ADC will take one sample each symbol period and process this digital signal in the digital domain before making a decision on the transmitted symbol.

Sampling is the process of taking a continuous-time signal and converting it to a discrete-time signal, or a sequence of samples. Quantization is the process of converting a sample with continuous-valued amplitude into a sample with an amplitude that is from a finite set of values. That is, we approximate the value of the sample by choosing the closest value in
the finite set. This is where we obtain an N-bit digital signal, which represents one of the $2^N$ values of the finite set. A quantizer is the system that performs the quantization. Essentially, an ADC performs the sampling and quantization of the input analog signal and outputs a digital signal.

A quantizer rounds off the amplitude of an input sample to the appropriate value from its finite set. The process of quantization produces a quantization error on each sample since the input is now rounded to a value in order to go into the digital domain. This can be understood by observing Figure 2.1. We can think of the quantization error as an additional noise source in the system presented in Figure 1.3(c). Thus, we do not want the ADC to introduce a large amount of quantization error because it would degrade the performance of the system. Note that this quantization error is inherent in a quantizer and an ideal ADC also introduces quantization noise.

2.2 Uniform Quantization

An example of a possible 4-level quantizer transfer curve is shown in Figure 2.2. Here, the transitions on the horizontal axis are called the decision levels and the values on the vertical axis are called the output values or quantization levels. The spacing between adjacent decision levels will be called a quantization step. In this example, we have a uniform quantizer since the quantization steps are uniformly spread between the limits of the quantizer. The limits are
-4 and 4 and $V_{ref}$ is defined as the input range of the converter. Since, we have an uniform quantizer, we can define a quantization step as $V_{LSB}$,

$$V_{LSB} = \frac{V_{ref}}{2^N}$$  \hspace{1cm} (2.1)

where $N$ is the number of bits of the converter. In Figure 2.2, this becomes $V_{LSB} = \frac{8}{4} = 2$ and each change in the least significant bit in the digital output produces a change of $V_{LSB}$ in the value it represents. This means that if we were to go into a digital-to-analog converter, each time we change the least significant bit in the digital input, we would get an change of $V_{LSB}$ in the analog output.

![Figure 2.2](image)

**Figure 2.2** The transfer curve of a possible 4-level quantizer or 2-bit ADC.

If we assume that the input to the quantizer is uniformly distributed or varying rapidly [Johns, 1997.1], such that over a large number of samples, the input is quantized with all possible errors with approximately the same rate of occurrence, then the quantization error will be a uniformly distributed random variable between $\pm \frac{V_{LSB}}{2}$. The limits are $\pm \frac{V_{LSB}}{2}$ since this is the maximum error realized by the quantizer, as long as the input lies within the quantizer range such that it doesn’t become overloaded. If the quantization error is distributed
in this manner, then the mean of \( V_q \) is zero and its variance is \( \frac{V_{\text{LSB}}^2}{12} \). Using (2.1), the power of the quantization error or quantization noise is

\[
N_{dB} = 10 \log \left( \frac{V_{\text{LSB}}^2}{12} \right) = 10 \log \left( \frac{V_{\text{ref}}^2}{(12)2^{2N}} \right)
\]

(2.2) simplifies to

\[
N_{dB} = -6.02N - 10.79 + 20 \log(V_{\text{ref}})
\]

Usually, \( V_{\text{ref}} \) is limited by the circuitry and we have limited control of it. The primary variable we can control is the number of bits of the ADC. Every additional bit we add to our quantizer reduces the quantization noise by 6 dB. This should make intuitive sense, since every additional bit doubles the number of output levels used to represent the input signal. This means we have smaller quantization steps and thus more accurate quantization. This implies less quantization noise added onto our signal. Recall that this result was derived for the uniform quantization case.

Uniform quantizers are simple conceptually, appeal to our intuitive senses and are easily implemented. Uniform quantization is commonly used in all kinds of systems and make up the vast majority of ADCs on the market. However, uniform quantization is not necessarily the most effective. As will be shown shortly, there are ways to achieve smaller quantization noise power for a given input signal and number of quantization levels.

### 2.3 Quantization Power

We will derive the quantization power for a general quantizer. We want to find the variance of the quantization error.
\[ \sigma_q^2 = E[Q^2] = \int_{-\infty}^{\infty} q^2(x) f_X(x) dx \] (2.4)

where \( X \) is the random variable representing the input, \( f_X(x) \) is the probability density function (pdf) of \( X \), \( Y \) is the random variable representing the output of the quantizer and \( Q \) is the random variable representing the quantization error. Thus,

\[ Q = X - Y \] (2.5)

Utilizing (2.5) and noting that the value from \( Y \) is picked from a finite set of values, we can put (2.4) in the following form,

\[ \sigma_q^2 = \frac{1}{K} \sum_{i=1}^{K} \int_{x_i}^{x_{i+1}} (x - y_i)^2 f_X(x) dx \] (2.6)

where \( y_i \) is the value assigned at the output when the input is between \( x_i \) and \( x_{i+1} \). Thus, we have an equation for the quantization power of a general quantizer. Looking at (2.6), we see that quantization power is highly dependent on the quantizer input and the quantizer's decision and quantization levels.

If we look at the case of having a uniform quantizer with \( x_1 = -\frac{V_{ref}}{2} \) and \( x_K + 1 = \frac{V_{ref}}{2} \) and a uniformly distributed input signal with the same limits and use (2.6), we will get the same result as we had before for the quantization power, \( \frac{V_{LSB}^2}{12} \). This should make sense, since a uniformly distributed input would produce quantization error that is uniformly distributed and the same result should follow. We should note that we will achieve the same noise power for other input signals, such as one that is rapidly changing so that all
possible quantization errors occur with approximately the same rate of occurrence as mentioned before.

2.4 Optimal Quantization

We are now interested to find the best possible quantizer for a particular input probability density function. That is, provided we know what $f_X(x)$ is, what is the best quantizer we can design that would produce the minimum amount of quantization noise for a given number of quantization levels? We will refer to an optimal quantizer as one in which the decision and output levels are chosen to obtain minimum quantization power.

This problem was originally addressed by Lloyd [Lloyd, 1957] and Max [Max, 1960]. They both considered the optimality of quantizers. To find the minimum quantization power in terms of $x_1, x_2, \ldots, x_{K+1}$ and $y_1, y_2, \ldots, y_K$, we need to set the derivatives of (2.6) with respect to each of these parameters to zero. Mathematically, we want

$$\frac{\partial \sigma^2}{\partial x_i} = 0 \quad i = 2, \ldots, K \quad (2.7)$$

$$\frac{\partial \sigma^2}{\partial y_i} = 0 \quad i = 1, \ldots, K \quad (2.8)$$

where it is assumed $x_1$ and $x_{K+1}$ are $-\infty$ and $+\infty$, respectively.

Evaluating (2.7) and (2.8) yields the following result [Lloyd, 1957][Max. 1960].

$$x_{i, \text{opt}} = \frac{(y_{i, \text{opt}} + y_{i-1, \text{opt}})}{2} \quad (2.9)$$

$$\int_{x_{i, \text{opt}}}^{x_{i+1, \text{opt}}} (x - y_{i, \text{opt}}) f_X(x) dx = 0 \quad (2.10)$$
That is, from (2.9), each decision level must be halfway between the output levels it connects or are adjacent to it. (2.10) implies that each output level should be the centroid of $f_X(x)$ in the appropriate region. In other words, the output level must be the mean value of $X$, when the input is in the appropriate region [Gersho, 1978]. We will call these two conditions the Lloyd-Max conditions and the quantizer it defines a Lloyd-Max quantizer. Solving (2.9) and (2.10) simultaneously is not easily done and the above conditions are used in the Lloyd-Max algorithm where an iterative procedure is used to find a set of parameters that satisfies the Lloyd-Max conditions.

While the Lloyd-Max conditions are necessary conditions for a minimum, they are not sufficient ones [Gersho, 1978]. There may be several sets of decision and output values that satisfy the Lloyd-Max conditions, producing several local minima, with only one being the desired absolute minimum (or perhaps several local minima are at the absolute minima). Thus, we can have a quantizer that satisfies these conditions and is not the optimum one. However, it was shown by Fleischer that there is a sufficient condition to obtain an optimal quantizer [Fleischer, 1964]. He showed that if we can satisfy the property that

$$\frac{d^2}{dx^2} \{ \ln f_X(x) \} < 0$$

(2.11)

for all $x$, then only one quantizer exists that satisfies the Lloyd-Max conditions and that this quantizer is optimum. We will call (2.11) the Fleischer condition.

In other words, (2.11) says that if $\ln(f_X(x))$ is concave downwards for all $x$, then we can guarantee that if we satisfy the Lloyd-Max conditions, we have parameters that define an optimal quantizer. It should be noted that the converse is not true, we can have an input signal which does not satisfy (2.11) and yet have a unique Lloyd-Max quantizer [Gersho, 1978].
2.5 Nonuniform Quantization in a 4-PAM System

We will examine quantization for a 4-PAM system using the architecture in Figure 1.3(c). In this case, we assume that the signal has been sampled and now needs to be quantized, so that a digital signal can be assigned to each sample.

2.5.1 Possible Optimal Quantization

First, let's see if the Fleischer condition is satisfied for a 4-PAM symbols with AWGN. Consider the following situation in Figure 2.3. $X$ is the random variable representing the 4-PAM symbols with equally likely symbols. $N$ is the random variable representing the AWGN with zero-mean and variance of $\sigma^2$ and $Y$ is the sum of the two. We want to know if the Fleischer condition is satisfied for $Y$. Note that we are essentially assuming there is no ISI on our signal. This could be a valid approximation if the analog equalization keeps the ISI small.

First, we need to find the pdf of the signal that will be input into the quantizer. This can be found by noting that

$$Y = X + N$$  \hspace{1cm} (2.12)

where the pdf of $X$ is

$$f_X(x) = \frac{1}{4} (\delta(x + 3) + \delta(x + 1) + \delta(x - 1) + \delta(x - 3))$$  \hspace{1cm} (2.13)

and the pdf of $N$ is
Chapter 2: Nonuniform Quantization

Nonuniform Quantization

implies that the pdf of \( Y \), \( f_Y(y) \), is given by [Leon-Garcia, 1994]

\[
f_Y(y) = f_X(x) \otimes f_N(n)
\]  

(2.15)

Thus, the pdf of \( Y \) is given by the convolution of the pdf's of \( X \) and \( N \). The convolution will produce the sum of four scaled versions of Gaussian random variables, each shifted by one of the PAM symbol values.

It is known that a Gaussian random variable satisfies the Fleischer condition [Fleischer, 1964]. However, we have the sum of four scaled and shifted Gaussian random variables. Substituting \( f_Y(y) \) into (2.11) and solving by hand is non-trivial and instead was calculated through simulations. It was found that the signal \( Y \) does not satisfy the Fleischer condition and there are numerous decision and output levels that satisfy the Lloyd-Max conditions, each set producing a different local minimum.

2.5.2 Fine-Coarse Quantization

Lets move away from optimal quantization and consider a different form of quantization. It would seem to make sense that for a fixed number of quantization levels, if we had smaller quantization steps where the probability of occurrence of the signal is higher at the expense of larger quantization steps where the probability of occurrence of the signal is lower, we would achieve smaller quantization noise [Jayant, 1984]. In other words, we could implement a form of nonuniform quantization where we have finer quantization where the pdf of the signal is high and coarser quantization where the pdf is low.

Figure 2.4 shows a simulation example of the distribution of possible samples.
presented to an ADC in the receiver of a digital communication system. For small ISI, the samples are centered around the PAM symbols. This means that most of the signal energy is concentrated around the PAM symbols with small deviations due to noise and ISI. The signal is rarely quantized to a level far away from the levels close to the PAM symbols. In other words, the signal is mostly quantized to values close to the PAM symbols. This means that certain levels are used extensively while others are rarely used. This supports the idea of using fine quantization at the expense of coarse quantization. However, as the ISI is increased, we see that we deviate more and more from the PAM symbols and the effect of nonuniform quantization would have less of an effect.

Figure 2.4 Possible samples introduced to the quantizer for varying ISI and fixed AWGN.
We are concerned with the receiver architecture in Figure 1.3(c). Some analog equalization will be performed on the incoming signal and the ISI will be reduced. This will keep the signal close the PAM values. If there was no equalization done prior to the ADC, like in the architecture in Figure 1.3(b), then there most likely would be a large amount of ISI on our signal. Nonuniform quantization, as discussed above, would have no benefit in such a case since the input to the ADC would be more like a uniformly distributed signal.

Now the question becomes how do we arrange the decision and output levels to achieve small quantization power? This needs to be done in a way that makes its physical implementation practical. An effective and simple way to do this would be to implement fine "uniform" quantization around the PAM symbols with coarser steps farther away from the PAM symbols. We will call this a fine-coarse (FICO) quantizer. This idea is shown in Figure 2.5 for a 32-level quantizer. Here, we see that we uniformly quantize around the PAM symbols with fine steps and then coarsely quantize farther away from the PAM symbols.

![Figure 2.5 The transfer curve of the FICO quantizer.](image)

We will define the nonuniform factor, $\Delta$, to be the fraction of the way between a PAM threshold and the PAM symbol that we begin fine quantization as shown in Figure 2.5. The
closer the value of $\Delta$ to one, the finer the quantization around each PAM symbol since there will be a fixed number of quantization levels assigned to each PAM symbol. Figure 2.6 shows three different 32-level quantizers each with different nonuniform factors. Note that when $\Delta=.25$, we have pure uniform quantization. Note that between a PAM threshold and PAM symbol, only the one decision level will be adjusted and then uniform quantization is performed between that decision level and the PAM symbol. This makes its implementation efficient and practical.

Figure 2.6 The FICO quantizer for various nonuniform factors.
2.6 Simulations Involving the Nonuniform ADC

2.6.1 Simulation Setup

Simulations were set up for the situation shown in Figure 2.7. Figure 2.7(a) shows the

situation introduced originally in Figure 1.3(c). After sampling the signal $r(t)$, we will have the
PAM symbols plus noise and ISI. The amount of noise and ISI present on the signal is highly
dependent on the channel and how the analog equalizer operates and can vary widely for
different systems and designs. Figure 2.7(a) can be conceptually thought of as the
simplification shown in Figure 2.7(b). $b(t)$ represents the impulse response of the transmit
filter, channel and analog equalizer. The ISI is introduced by $b(t)$ and the Gaussian noise
shown in Figure 2.7(b) has power $N (N=kN')$ since it has passed through the analog equalizer
and will experience noise enhancement.

Both noise and ISI will affect the variations on the transmitted PAM symbol. This will
affect how the nonuniform FICO ADC needs to be designed. Thus, we will consider certain

Figure 2.7 The system analyzed in the simulations, (a), and its simplification, (b).
situations for varying noise and ISI. The ISI is measured as explained here. Figure 2.8 shows a possible impulse response for a transmit filter, channel and analog equalizer. Here, every value after the first sample is post-cursor ISI. In this case, four taps would be required in the DFE to eliminate the ISI. The ISI is measured by summing all the ISI values at the sampling times and dividing by the symbol value to get the ISI as a fraction or percentage.

![Figure 2.8 Impulse response, b(t), and the illustration of ISI.](image)

The choice of the noise power values is based on the symbol error rate, SER vs. SNR curve shown in Figure 2.9 for 4-PAM. This curve is based on an equation from [Lee, 1994]. The SNR is the ratio of signal power to Gaussian noise power present on the received signal
right before the slicer. This curve assumes zero ISI. However, if we assume that the analog equalizer and DFE remove the ISI, then this curve may be a good assumption since all that would be present to the slicer would be the PAM symbols plus noise. Note that in our case, some of the noise at the slicer will include quantization noise as well as Gaussian noise. This will also alter the shape of this curve. We only use this curve to get a rough idea of what kind of SNR we are trying to achieve. Table 2.1 gives an idea of what kind of noise powers we are dealing with. Note that all simulations were done in MATLAB (with Simulink).

![SER vs. SNR for 4-PAM.](image)

**Figure 2.9** SER vs. SNR for 4-PAM.

<table>
<thead>
<tr>
<th>Noise Case</th>
<th>Signal Power (S)</th>
<th>Noise Power (N)</th>
<th>SNR</th>
<th>Corresponding SER</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7 dB</td>
<td>-10 dB</td>
<td>17 dB</td>
<td>$10^{-3}$</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>-13</td>
<td>20</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>-17</td>
<td>24</td>
<td>$10^{-12}$</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>-20</td>
<td>27</td>
<td>$10^{-23}$</td>
</tr>
</tbody>
</table>

**Table 2.1** The four different noise cases used in the simulations.
2.6.2 Simulation Results

Now that the simulation setup has been explained, the results of the 5-bit FICO quantizer will be shown. Referring to Figure 2.10 and Figure 2.11, we observe the results we have obtained. These graphs show plots of the quantization noise power vs. the nonuniform factor for the 5-bit FICO quantizer. There are four plots shown, each for a fixed amount of noise power. In each plot, there are a number of curves, each corresponding to a different amount of ISI. For instance, in Figure 2.10(a), the curves are plots of the quantization power vs. different nonuniform factors for SNR=27dB and various ISI values.

We have also shown in each plot the quantization power for a 5-bit and 6-bit uniform quantizer (calculated from (2.3)) to see how our 5-bit FICO quantizer is doing in comparison to the uniform case. If we were to use a uniform quantizer on our signal, we would achieve the quantization power expected from the uniform case presented in (2.3). Although we do not have a uniformly distributed signal, the signal to the quantizer is rapidly changing around the PAM symbols, producing a uniformly distributed quantization error.

Let us analyze the shape of the curves on one plot. If we look among the different curves, we see that they all exhibit a similar behavior. As we increase Δ, we get smaller quantization intervals around the PAM symbols and thus at first, we get an improvement in quantization noise due to having an improved resolution. However, as we keep reducing the quantization steps, we hit a point where the quantization noise begins to increase. This is because as Δ increases, the signal will eventually begin to fall more and more into the coarser quantization intervals. The point at which this happens depends on how close the signal is concentrated around the PAM symbols. For smaller ISI or noise, this point will happen at a larger Δ value since we can squeeze the levels tighter when we have less variation around the PAM symbols.
Figure 2.10  Quantization noise vs. the nonuniform factor for various noise and ISI values.
Figure 2.11  Quantization noise vs. the nonuniform factor for various noise and ISI values.
Looking at any one of the plots, we see that in general, we get less benefit from the quantizer as the ISI is increased. This should make sense, since if we increase ISI, there is more variations around the PAM signal and thus the nonuniform arrangement has less of an effect on quantization noise. On each figure, we have shown the case of no ISI as a reference. This situation would never really be achieved, since if there were no ISI, we would not need to do any digital equalization and thus would not need an ADC. However, it provides a good reference point to compare the other ISI curves to. Also, note that in all the figures, the case where we have over 40% ISI is another extreme. In this case, the signal is not highly concentrated anywhere, since the ISI is so high. We see that employing any nonuniform quantization in this case is not beneficial.

Also, as we look among the plots, we observe that as the SNR is decreased (or noise increased), the curves begin to shift up and to the left. This means larger quantization noise at a smaller nonuniform factor. This is because as we get larger noise power, the signal becomes less concentrated around the PAM symbols and again, nonuniform quantization has less of an effect on the signal.

Absorbing all of this information, we can say that the quantizer is highly dependent on what its input signal is. In particular, it depends on how much noise is added to the system, how this noise gets enhanced through the analog equalizer and how much ISI the analog equalizer leaves for the DFE. These factors all depend on such factors as the channel characteristics, the architecture of the analog equalizer and other specifications of the system.

Nevertheless, we do observe clear improvements of the 5-bit nonuniform quantizer over the 5-bit uniform quantizer. For higher SNR we can achieve up to the same dynamic range as a 6-bit quantizer. For other cases, we can achieve anywhere between that. Of course, there an infinite number of graphs that can be plotted and the design of the actual quantizer highly depends on the input signal presented to it.

Even achieving a fraction of a bit of improvement over the uniform case can be
advantageous. In the case where one can expect a certain range of ISI and noise, then the nonuniform factor can be chosen to achieve the lowest amount of quantization noise and achieve improvement in the performance of the quantizer.

2.7 Variable Nonuniform Quantization

As mentioned above, the nonuniform FICO quantizer is highly dependent on the input signal since we are picking the position of the levels based on the partially equalized signal and the value of the noise. In practical situations, we can not know precisely the power of the noise or the amount of ISI added onto the signal. If we design an ADC with a fixed $\Delta$, then variations on the expected signal will change the performance of the ADC from what is desired. We would have to know exactly know what the noise is and the amount of ISI added onto the signal to design an ADC with a fixed $\Delta$.

To improve the flexibility and robustness of the ADC, we will design a variable nonuniform FICO ADC. In other words, we will provide programmability to the possible spacing of the decision levels. Thus, we will be able to manually choose what $\Delta$ factor we want for our quantizer. In the ADC designed, we provide four possible $\Delta$ factors to choose from. The choice of the desired nonuniform factors is chosen digitally and will be discussed in Chapter 3.

2.8 Comparison to Lloyd-Max Quantizer

In this thesis, FICO quantization was implemented. However, it would be interesting to see how a Lloyd-Max quantizer performs in comparison to the FICO quantizer. Although we mentioned before that a 4-PAM system will not have a unique Lloyd-Max quantizer, we can still use the Lloyd-Max algorithm to find a quantizer with quantization power that is at a local minimum. According to the Lloyd-Max algorithm, you make an initial guess and then iteratively compute the values that satisfy the Lloyd-Max conditions. Since, there are
numerous local minima, depending on the initial guess you make, you may get a different set of decision and output levels.

What we will do is that for a fixed noise power and ISI on our signal, we will find the nonuniform factor that produces the minimum amount of quantization noise from the plots in Figure 2.10 and Figure 2.11. We will use this quantizer as our initial guess and let the Lloyd-Max algorithm find a quantizer that satisfies the Lloyd-Max conditions. We will do this for the specific case of SNR=24 dB to see what kind of results we get. The results are shown in Table 2.2. Since each bit represents 6.02 dB improvement in quantization noise, we represent the improvements of the quantizers in bits.

<table>
<thead>
<tr>
<th>ISI</th>
<th>improvement for FICO quantizer over 5-bit uniform case</th>
<th>improvement for Lloyd-Max quantizer over 5-bit uniform case</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2%</td>
<td>.8 bits @ Δ=.55</td>
<td>1.3 bits</td>
</tr>
<tr>
<td>6.7%</td>
<td>.6 bits @ Δ=.55</td>
<td>1.1 bits</td>
</tr>
<tr>
<td>11.5%</td>
<td>.25 bits @ Δ=.45</td>
<td>.7 bits</td>
</tr>
</tbody>
</table>

Table 2.2 Comparison of improvements in quantization noise for a FICO quantizer and Lloyd-Max quantizer for SNR=24 dB.

We see that we can achieve about half a bit of improvement using the Lloyd-Max quantizer. It was found in the simulations that the Lloyd-Max quantizer's transfer function, for these specific cases, employed finer quantization around the PAM symbols but added an extra coarse level further from the PAM symbol at the expense of a fine level. This is shown in Figure 2.12.

Due to the additional coarse quantization level, the implementation of the Lloyd-Max quantizer could be more complex than the FICO quantizer. Also, further evaluation is required to see how extensive the advantages of the Lloyd-Max quantizer are and how its transfer curve changes for different situations.
2.9 Digital Linearization

Consider the following nonuniform quantizer transfer curve shown in Figure 2.13. We know
that for an ADC, we will have to add digital signals to represent each output level. In our case, this is a 5-bit digital code. In a uniform ADC, each LSB change in the digital code represents a change of $V_{LSB}$ in the analog value it is representing. However, in the nonuniform case, each LSB change in the digital output can represent a different change in the analog value it is representing. For example, looking at Figure 2.13, the output value of -0.35 is represented by the digital code ‘01111’. If we move to the next output level and digital code, we have changed our output level by 0.7. However, moving to the next level after that, we will change our output level by 0.4. We see that when changing the digital code by 1 LSB, we will get different changes in the analog representation. This will cause a problem when performing digital signal processing due to having a nonlinear digital representation of our signals. We will need to perform some sort of *digital linearization*.

This can be done by adding additional bits on the digital signal after the ADC so that each time we change by 1 LSB, it represents a linear increase in the analog representation. This can be accomplished conceptually by the system in Figure 2.14. The output of the first

![Diagram](image)

**Figure 2.14** Digital linearization by using an 8-bit uniform quantizer.

nonuniform quantizer will be linearized by the second uniform quantizer since each change in the uniform quantizer’s LSB will be uniform. For instance, to illustrate this point, using the above example, the code ‘01111’ and ‘10000’ from the first quantizer might be changed to ‘1001110’ and ‘10001001’. Thus there are 21 LSBs between the two 8-bit codes to represent the 1 LSB change in the 5-bit code. There will be a mapping of each 5-bit code to an 8-bit code.

The addition of the second quantizer introduces additional quantization noise. The
higher the resolution of the second uniform quantizer, the less the effect of this additional quantization. We will pick the resolution of the second uniform quantizer depending on what kind of resolution we have around the PAM symbols. We have illustrated in Figure 2.14 an 8-bit digital linearizer. The choice of the resolution of the linearizer really depends on the specifications of the system. It was found for our system, an 8-bit linearizer sufficed and caused minimal additional quantization noise.

2.10 Summary

We have presented results for a 5-bit FICO nonuniform quantizer that achieves improved dynamic range over a 5-bit uniform quantizer under certain conditions. The amount of improvement is highly dependent on the signal presented to the FICO ADC. This will depend on the specific system and can vary widely for various situations. In particular, the noise and ISI added onto each symbol will affect the performance of the nonuniform ADC. The more concentrated the signal is around the PAM symbols, the better the effect of the nonuniform quantization.

We are not designing the ADC for a specific industrial application and do not have exact specifications. Instead, we are simply investigating the principle of nonuniform quantization and we find advantages for a nonuniform ADC in a PAM system. We have investigated nonuniform quantization for the 4-PAM case but the principles can be carried over for other multilevel PAM systems.
CHAPTER 3

Circuit Design

This chapter deals with the circuit design of the ADC. We have discussed the theory behind nonuniform quantization and now want to apply these principles into a working model. Recall that the ADC to be designed is a 5-bit, 150 MS/s converter with variable nonuniform FICO quantization as discussed in Chapter 2. The technology used for the design is the Hewlett Packard CMOS14TB Silicon Technology. It is a high-speed, high-density 0.5 μm CMOS process optimized for 3.3 V. Note that all circuit simulations were done in HSPICE with GSI used as its graphical counterpart.

3.1 High-Speed ADC Architectures

The flash architecture is the most straightforward way to perform analog-to-digital conversion for high-speed applications. The input signal is compared with $2^N-1$ comparators connected in
parallel as shown in Figure 3.1. Each comparator is connected to a reference voltage generated by a resistor string. The reference voltages determine the position of the decision levels of the quantizer. When the input signal is greater than the comparator's individual reference voltage, the comparator will toggle to the appropriate digital value. The output of the comparators produce a thermometer style code which can then be encoded to the appropriate binary code.

While the flash architecture is straightforward and fast, it has disadvantages. Due to the $2^N$ dependency, the size, power and input capacitance of these ADCs roughly doubles for every additional bit of resolution.

To try to improve on the disadvantages of the flash architecture, analog preprocessing techniques have been proposed. A block diagram of the analog preprocessing idea is shown in
Figure 3.2. The input signal is processed in the analog domain before being passed onto comparators to determine the digital value.

Analog preprocessing techniques include folding and interpolation. These two preprocessing architectures have been implemented together and separately [van de Grift, 1987][Steyaert, 1993][Nauta, 1995][Roovers, 1996][Flynn, 1996]. We will utilize interpolation in our ADC.

### 3.2 Interpolation

Interpolation is an analog preprocessing architecture that reduces the input capacitance placed onto the input signal. This is particularly significant at high-speeds since it can avoid or ease
the design of a power-hungry buffer to drive the input voltage.

**Figure 3.3 Comparison of the (a) flash and (b) interpolating architecture.**

Figure 3.3 shows a portion of the front-end of a flash ADC and interpolating ADC to help understand the difference between the two. Instead of directly feeding into three comparators as in the flash architecture, $V_{in}$ is first fed to two input stages for the interpolating architecture before being placed into three comparators. The values of the signals generated by the input stages depends on what the value of $V_{in}$ is with respect to the corresponding reference voltage. For instance, if $V_{in} > V_{ref1}$, then the output signal of input stage #1 will be above the threshold of the comparator it is being fed to. What is truly important is that the signal crosses the comparator threshold when $V_{in} = V_{ref1}$. All the comparator needs to do is latch high or low according to whether its input signal is above or below its threshold.

A plot of possible output signals of the input stages vs. $V_{in}$ for the interpolating
architecture is shown in Figure 3.4. We see that when $V_{in} > V_{ref1}$ the output signal from input stage #1 is above the comparator threshold and vice-versa when $V_{in} < V_{ref1}$. Thus, the output of comparator #1, $d_1$ (as shown in Figure 3.3), will be the same value as $d_1$ in the flash architecture since essentially, in both cases, we are comparing $V_{in}$ with $V_{ref1}$ and latching accordingly. This idea holds for all signals generated by the input stages.

However, notice that while we have generated appropriate signals for the comparison of $V_{in}$ with $V_{ref1}$ and $V_{ref3}$, we are missing the required signal between the two input stages to compare $V_{in}$ with $V_{ref2}$. In the flash architecture, $V_{in}$ is compared with $V_{ref2}$ (in the resistor string) right away but in the interpolating architecture, $V_{in}$ is not compared with $V_{ref2}$ at all. Thus, we are missing a signal that needs to have a threshold crossing at $V_{ref2}$ in the interpolating case. To generate this missing signal, we *interpolate* between the two input stages. This interpolated signal will have a threshold crossing at $V_{ref2}$ as shown in Figure 3.4.

In the interpolating case, $V_{in}$ is connected to two input stages while in the flash case, $V_{in}$ is connected to three comparators. Thus, the reduction of stages that $V_{in}$ is connected to results in a lower input capacitance. The interpolation factor is defined as the number of
interpolated signals between adjacent input stages plus one. In Figure 3.3, we have an interpolation factor of two. The higher the interpolation factor, the lower the input capacitance will be since there were be less input stages connected to $V_{in}$.

In bipolar technologies, interpolation has been obtained through a resistor ladder network and is based on the concept of voltage interpolation [van de Griff, 1987]. However, in CMOS, the output signals from the input stages cannot drive resistive loads easily. Therefore, in CMOS, current interpolation is performed with the scaling and adding of current signals [Steyaert, 1993]. This means that the outputs of the input stages are current signals, essentially making the input stages transconductors. Note that this means that the comparators now become current sensitive.

To understand how to interpolate, consider Figure 3.5. We have current signals $I_A$ and $I_B$ produced by the input stages and we want to obtain an arbitrary interpolated current signal, $I_C$, somewhere between $I_A$ and $I_B$. If we want $I_C$ to be a fraction, "$a$" (where $0 \leq a \leq 1$), of the way from $I_A$ to $I_B$, then

$$I_C = I_A - a(I_A - I_B) = (1 - a)I_A + aI_B$$

Note that if we go a fraction of the way in the vertical direction, this translation also produces a shift in the horizontal direction of the same fraction. This is shown in Figure 3.5.
For example, if we wanted an interpolated signal, $I_{CI}$, to lie $1/4$ of the way from $I_A$ to $I_B$ then using (3.1), we get

$$I_{C1} = \frac{3}{4}I_A + \frac{1}{4}I_B \tag{3.2}$$

and if we wanted $I_{C2}$ to lie in the middle of $I_A$ and $I_B$ ($a = 1/2$) then

$$I_{C2} = \frac{1}{2}I_A + \frac{1}{2}I_B \tag{3.3}$$

These types of equations can be implemented with weighted current mirrors summed at appropriate nodes. The circuit implementation of (3.2) and (3.3) is shown in Figure 3.6.

![Circuit implementation of (3.2) and (3.3).](image)

Note that for interpolation to work, the signals $I_A$ and $I_B$ must be linear over the interpolation range. Otherwise, the interpolated signal may not lie where desired. This would create a systematic offset, shifting the position of the threshold crossing of the interpolated signal.
3.3 Overall Circuit Architecture

The overall circuit architecture of the ADC is shown in Figure 3.7. The input stages (essentially transconductors) compare the input analog signal with reference voltages generated by a resistor ladder. Each one of the input stages has a reference voltage that
represents a PAM threshold or PAM symbol as shown in Figure 3.8. These voltages are

![Diagram of input stage reference voltages]

Figure 3.8 A nonuniform quantizer curve and the location of the reference voltages.

uniformly spread and fixed. We see that we need to interpolate by a factor of four to generate the three missing current signals (and accordingly missing reference voltages) between input stages. The nine input stages generate appropriate non-interpolated differential current signals. The current interpolation circuit then interpolates between adjacent input stages. Note that we need a total of 8 interpolating circuits, one between each input stage. This gives a total of 24 differential interpolated signals and 9 non-interpolated differential signals. This generates 33 differential current signals.

Variable quantization spacing is achieved within the interpolation circuit. A 2-bit digital signal becomes decoded and then determines the desired quantization spacing. There are four different quantization spacing possibilities based on the specified 2-bit digital code.

The differential current signals are fed to comparators. The latched comparators sense differences in the differential currents and amplify them to full digital values during latch mode.

At the output of the comparators, we have a digital thermometer style code. We need to convert this code into our desired 5-bit digital signal. Ideally, going from the bottom-up in
Figure 3.7, we should have a string of ‘1’s and then a string of ‘0’s at the outputs of the inverters. The transition from ‘1’ to ‘0’ depends on the value of $V_{in}$ at the sample time. A two-input NAND gates could be used to detect the transition point in the thermometer code. The NAND gate would only go low (high after inversion) if it had a ‘1’ and a ‘0’ above it. Then only one of the NAND gate outputs would be high (after inversion) and this allows for simple encoding. However, sometimes a lone ‘1’ or ‘0’ will occur in the wrong place due to comparator metastability, noise, cross talk, limited bandwidth, etc. [Johns, 1997.1]. Therefore, error correction has been added to improve performance using three-input NAND gates [Steyaert, 1993]. The three-input NAND gates attempt to correct the problem to a certain extent. There now must be two ‘0’s immediately above the ‘1’ in determining the transition point in the thermometer code.

After the error correction, only one of the remaining 31 digital signals will be high. That signal indicates what the 5-bit binary code should be. A 31-to-5 ROM encoder picks the appropriate code based on which line is high. This produces our 5-bit digital output. Note that since the ADC is a prototype, digital linearization as described in Section 2.9 is not implemented so as not to further complicate the design of the ADC.

Note that all transistor sizes labelled in the circuit diagrams of this chapter are in microns. For instance, $10/0.8$ implies $10\mu m/0.8\mu m$.

### 3.4 Bias Circuit

Although the bias circuit is not shown in Figure 3.3, it is required to generate the current sources needed in the input stages. The wide-swing constant-transconductance bias circuit
(Figure 3.9) was used [Johns, 1997.1]. This circuit allows us to have wide-swing cascode current mirrors. This provides us good high-impedance current sources with good output swing on them. The bias loop generates our desired bias current values. The wide-swing cascode bias, generates the wide-swing behavior of the circuit. It can be shown by performing KVL around the bottom loop, with $R_{bias}$, $M_3$ and $M_2$, it can be shown that

$$g_{M_2} = \frac{1}{R_{bias}} \quad (3.4)$$

if the size of $M_3$ is four times larger than $M_2$ [Johns, 1997.1]. This provides a stable transconductance, independent of power-supply voltages, process parameters or temperature.

A start-up circuit is also added. This is to ensure that at start-up, the circuit doesn't remain in a state in which all currents are zero. If all currents in the bias loop are zero, then $M_{17}$ will be turned off. $M_{18}$ will pull the gates of $M_{15}$ and $M_{16}$ high, in effect turning them on. This will inject current in the bias loop which in effect will turn $M_{17}$ on, pulling the gates of
M_{15} and M_{16} low. M_{18} is sized to be a high-impedance load so that its on-resistance is much larger than M_{17}'s on-resistance. This will ensure the voltage at the gates of M_{15} and M_{16} goes low enough.

### 3.5 Input Transconductance Stage

The input transconductance stage is shown in Figure 3.10. We see that we have a differential pair in which M_1 is connected to the input signal and M_2 is connected to a reference voltage from the resistor ladder. It is important for the differential pair to remain highly linear over the interpolation range as mentioned earlier. This implies picking a large $V_{eff1}$ to increase the linearity of the input stages. Any nonlinearity associated with the input stage over the interpolation range will produce an error that can be translated as a systematic input referred offset voltage. This will effect the overall linearity of the ADC. A high enough $V_{eff1}$ was picked so that the systematic input referred offset voltage is small and negligible.

The transconductance, $G_m$, of the input stage is also important. Any current offsets

![Figure 3.10 The input transconductance stage.](attachment:image_url)
generated in the interpolation circuit or comparator circuit will be divided \( G_m \) by to get equivalent input offset voltage. Therefore, a high \( G_m \) is desired.

Analyzing the above comments, and noting that

\[
V_{eff1} \propto \sqrt{I_{D1} \left( \frac{L}{W} \right)} \tag{3.5}
\]

\[
G_m \propto \sqrt{I_{D1} \left( \frac{W}{L} \right)} \tag{3.6}
\]

We see that changing the sizes affects one parameter positively and the other negatively. However, a high bias current has a positive impact on both parameters. This is set by \( M_3 \) and \( M_4 \), which are biased by the bias circuit. It should be noted while a high bias current will improve performance, it will also increase power dissipation. High output impedance is achieved on the current sources by using cascoded devices. It was observed in simulations that the current varied too much to maintain reasonable performance for non-cascoded current sources.

Another important factor is the analog bandwidth. This will limit the maximum input signal frequently that we can digitize effectively. A high analog bandwidth ensures better data conversion at high frequencies. The two nodes that limit the analog bandwidth are the input node and the current mirror node. The dominate node in the circuit is the current mirror node at the drain of \( M_5 \) and \( M_6 \). The capacitance is high on these nodes because of the gate-source capacitance of the all the current devices attached to this node for interpolation (not shown yet). The pole at this node is

\[
f_{p5} = \frac{1}{2\pi R_{eq} C_{eq}} \propto \frac{gM_5}{C_{eq}} \tag{3.7}
\]

The equivalent capacitance at this node depends on the number of transistors attached to this node (which depends on the interpolation factor) and their respective sizes. If we keep all the
lengths the same as $M_5$ for all the current devices in the interpolation circuit we get

$$C_{eq} \propto L_5 \sum_k W_k$$  \hspace{1cm} (3.8)

$$f_{p5} \propto \frac{\sqrt{I_{DS}(W/L)5}}{L_5 \sum_k W_k} - \frac{I_{DS}}{c W_5 L_5^3}$$  \hspace{1cm} (3.9)

where $c$ is set by the sum of the ratio of the widths of the scaled transistors to $W_5$.

So, we see that a high bias current increases the analog bandwidth. A small area for the devices is also needed to maximize the bandwidth. However, a trade-off will take place. The current mirror transistors are extremely important for matching. Matching among devices improves with larger areas (as will be discussed in Chapter 4).

### 3.6 Interpolation Stage

Figure 3.11 shows a uniform current interpolation circuit. $I_1$ and $I_5$ are the non-interpolated signals and $I_2$ to $I_4$ are the interpolated signals generated between the two input stages.

**Figure 3.11** A uniform current interpolating circuit.
[Steyaert, 1993]. They are generated as discussed before with the scaling and summing of current signals. Depending on the input voltage value, each input stage generates a non-interpolated differential current with a crossing at its respective reference voltage. These signals are used to interpolate for the “missing” reference voltages between the reference voltages of the input stages. Figure 3.12 shows the differential current signals generated with complementary signals and positive signals.

![Figure 3.12 Differential current interpolation signals.](image)

\( V_{ref1} = 0 \) and \( V_{ref2} = 100 \text{ mV} \). The differential implementation will improve the accuracy of the converter since the comparator will operate on the difference between signals and not on absolute values.

Now, consider a circuit that picks one nonuniform interpolated level and performs uniform quantization between that and an input stage. This is the kind of nonuniform quantization that was discussed in Chapter 2 and was called FICO quantization. The first decision level between a PAM threshold and PAM symbol is a coarse level and then fine uniform quantization is implemented between that decision level and the PAM symbol.
Chapter 3: Circuit Design

The circuit to implement this is shown in Figure 3.13.

Figure 3.13  A current interpolation circuit to implement FICO quantization.

By changing the sizes of \( M_7 \) and \( M_8 \) we can change the position the "missing" reference voltage. If we know where that decision level should be placed, then by using (3.1), we can pick the sizes of \( M_7 \) and \( M_8 \) accordingly. We pick one level \( (I_{A+}) \) depending on the sizes of \( M_7 \) and \( M_8 \). We then want to interpolate between \( I_{A+} \) and \( I_{5+} \). In order to re-interpolate, we need to buffer \( I_{A+} \). This is done by \( M_9 - M_{11} \). \( M_9 \) mirrors \( I_{A+} \) to \( M_{10} \) which passes the signal onto the p-channel diode connected transistor, \( M_{11} \). Then, "uniform" interpolation is performed between \( I_{2+} \) and \( I_{5+} \). With the size of \( 1/2 \) for \( M_7 \) and \( M_8 \), the signals generated from the circuit in Figure 3.13 is shown below in Figure 3.14. We can observe the
uniform spacing between $I_{2+}$ and $I_{5+}$.

Figure 3.14  Current signals corresponding to the circuit in Figure 3.13.

We will now take it one step further and implement variable nonuniform quantization as discussed in Chapter 2. Essentially, to vary the nonuniform factor, $\Delta$, we need to shift the position of $I_{A+}$ in Figure 3.13, in order to shift the position of that decision level. Figure 3.15 shows the circuit used to implement this. Only one of sw1, sw2, sw3 and sw4 will be high.

Figure 3.15  Current interpolation circuit with variable levels for FICO quantization.
This means that only one of the switches (M9-M12) will be on at any given time. Depending on which switch is on, the position of \( I_{A+} \) will be picked. Then the two remaining interpolated current signals (\( I_{3+} \) and \( I_{4+} \)) will be picked uniformly as discussed before. Only one of the pairs of transistors (\( M_1 \) and \( M_5 \), \( M_2 \) and \( M_6 \), \( M_3 \) and \( M_7 \), \( M_4 \) and \( M_8 \)) will actually be supplying current. Thus, the power is not increased over the circuit in Figure 3.13. We have four possible variations on the quantization levels and it is controlled by digital signals (sw1-sw4). The four possible nonuniform factors for the ADC were picked to be \( \Delta = .25 \) (5-bit), \( \Delta = .60 \) (5.9-bit), \( \Delta = .65 \) (6.1-bit) and \( \Delta = .75 \) (6.58-bit). The numbers shown in brackets are the resolution to which the fine quantization steps become for their respective \( \Delta \).

We are able to provide programmability and variability into the design without increasing the power consumption since only one of the switches will be on, allowing current to flow. Speed is degraded however, since unused transistors are adding gate capacitance to nodes \( n_1 \) and \( n_2 \) in Figure 3.15. Silicon area is also increased due to the programmability.

The input stage connected to node \( n_1 \) is the stage that the quantization levels are being pushed away from (representing the PAM threshold). The input stage connected to node \( n_1 \), is the stage that the quantization levels are squeezed in close to it (representing the PAM symbols). This is shown in Figure 3.16 (with five input stages and four interpolating regions) where we have indicated what node of their respective interpolating circuits each input stage is connected to.
Figure 3.16 Current signals generated by the circuit in Figure 3.15 for $\Delta = .60$.

### 3.7 2-to-4 Decoder

The 4-bit control signal (sw1-sw4) should be implemented in a simple and efficient way. To save on the number of pads used in the design and to provide a better design, we can utilize the fact that only one of these signals is supposed to be on at a given time. This can be designed by a 2-to-4 decoder. This way, a 2-bit digital signal is provided by the user and decoded onto 4-digital lines, only one of them being high depending on the 2-bit code. The circuit to implement this is shown in Figure 3.17 and is called a 2-to-4 NAND decoder [Rabaey, 1996].
3.8 Current Comparator

The differential current signals coming from the interpolation stage (interpolated signals) and input stage (non-interpolated signals) need to pass through comparators. Each comparator should be able to latch high or low depending on the polarity of the given differential current. All devices have dimensions of 2/0.6.

Figure 3.17 2-to-4 decoder circuit.
signal. The circuit used to implement the current comparator is shown in Figure 3.18

![Differential current comparator circuit](image)

Figure 3.18 Differential current comparator circuit.

[Steyaert, 1993][Roovers, 1996]. The latched comparator has a single phase clock. M₁ and M₂ are pull down devices that are always on. Transistors M₃, M₅ and M₄, M₆ form the positive feedback latch.

In the reset phase (clock is low), M₇ is turned on. The differential input currents get fed into low impedance nodes (due to being connected to the sources of M₃ and M₄). If I₈₊=I₈₋, then due to the symmetry of the design, V₀₊=V₀₋. If I₈₊>I₈₋ then a small voltage will be developed across M₇. This causes V₀₊ to become slightly greater than V₀₋. The opposite is true if I₈₊<I₈₋.

Once the clock goes high, M₇ turns off and we are in the latch mode. Any differences in V₀₊ and V₀₋ will get regenerated exponentially to full digital values by the positive feedback latch. The speed of the positive feedback latch in the latch mode is proportional to Gₘ/Cₜ [Roovers, 1996]. Cₜ is the capacitance at each of the output nodes and Gₘ is the transconductance of each inverter. Cₜ consists of parasitic capacitances at the output nodes due to devices in the current comparator circuit and the external load added onto the output.
3.9 SR Digital Latch

Since the output of the current comparators is passed onto digital logic, we need to be able to hold the digital values generated in the latch mode when the comparator goes into reset. This value should be held until the next latch cycle. Therefore, a digital latch is designed to hold the digital values once the clock goes low. This circuit is basically a clocked SR flip-flop and is shown in Figure 3.19 [Rabaey, 1996].

![Figure 3.19 Clocked SR latch.](image)

<table>
<thead>
<tr>
<th>clock</th>
<th>S</th>
<th>R</th>
<th>o+</th>
<th>o-</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>X</td>
<td>X</td>
<td>o+</td>
<td>o-</td>
</tr>
</tbody>
</table>

M₁, M₃ = 2/0.6  
M₂, M₄ = 8/0.6  
M₅, M₆, M₇, M₈ = 6/0.6

When the clock is high, the outputs take on the digital values given to it and retains those values once the clock goes low (reset phase in the comparator). It is important to size M₅-M₈ appropriately to allow the latch to switch states [Rabaey, 1996]. For instance, if we want to switch Q from high to low, then we must ensure that Q goes below the threshold of the inverter M₁, M₂.

The function of the SR latch can be observed in the simulation results shown in Figure 3.20. We see that when the comparator goes into the reset phase, the latch holds its digital value. When the comparator goes into the latch phase, the latch changes (or holds) its value according to the comparator's output.
Figure 3.20 Comparator output (above) and corresponding SR latch output (below).

3.10 31-to-5 ROM Encoder

After the error-correction stage (NAND gates and inverters), only one of the 31 digital signals will be high. This thermometer code needs to be translated into its 5-bit binary code. This is achieved by a 31-to-5 NOR ROM encoder [Rabaey, 1996]. Essentially, the input code selects the appropriate output code from a ROM table. A portion of this circuit demonstrating its function is shown in Figure 3.21.
This chapter has discussed the circuit design of the 5-bit nonuniform ADC. We have discussed the architecture of the ADC and the individual circuit blocks within the ADC. We also discussed how we implemented variable nonuniform FICO quantization as discussed in Chapter 2.
This chapter deals with the layout and practical implementation of the integrated circuit designed in this thesis. The circuits discussed in Chapter 3 were implemented in the Hewlett Packard 0.5 μm CMOS process. The layout of the ADC was done using the Cadence layout package.

4.1 Layout of the ADC

Certain layout techniques were used to improve the accuracy and effectiveness of the design. Special consideration was given to critical circuit blocks and layout wiring of the ADC. Also, since the ADC is a mixed-signal circuit (both analog and digital circuits on the same silicon substrate), noise issues had to be considered. It should be noted that many of the concepts
presented here (and implemented) are based on techniques and considerations presented in [Maloberti, 1994][Johns, 1997.1].

4.1.1 Matching Issues in Analog Portion

In the analog portion of the layout, matching is especially important among devices. As mentioned in Section 3.5, a finite $V_{eff}$ of the input stages produces the only systematic offset. However, in practical situations, there will be mismatches among transistors that are intended to be identical. Mismatch can be observed between parameters of equally designed devices as a result of several random processes which occur during fabrication [Pelgrom, 1989]. It has been shown that mismatches among the threshold voltage, $V_t$, of CMOS transistors can be expressed as a random variable with variance given as [Pelgrom, 1989]

$$\sigma^2(V_t) \propto \frac{1}{WL} \quad (4.1)$$

The same kind of relationship can be shown for the variation of $\beta = \mu_n C_{ox} \frac{W}{L}$. However, threshold voltage variations are typically dominant [Flynn, 1996]. Looking to (4.1), we see that larger transistors match better.

Mismatches among critical devices within the analog portions of an ADC can be very detrimental. Mismatch can be modeled as an input referred offset voltage and will adversely affect the linearity and accuracy of the converter. Therefore, an analysis of mismatch was important and necessary among the analog components. Monte-Carlo simulations were run in which the threshold voltage was varied among different transistors to see the effect on the input referred offset voltage. It was found that the p-channel devices in the cross-coupled inverters in the current comparator caused the largest offset errors. The input differential pair as well as the mirror transistors in the input stage caused the next largest error. The n-channel devices in the cross-coupled latch also caused significant offset errors. These are shown in
Figure 4.1. Since these devices were important for matching, their sizes were increased in an effort to improve their matching in the integrated circuit.

Also, special layout techniques were utilized to further improve the matching among critical devices. Critical devices for matching were laid out carefully to optimize their matching. To improve matching among devices, large objects were made out of several unit-sized transistors connected in parallel. This technique is very useful for matching in current mirrors. Also, outside or “dummy” transistors were included to better match boundary conditions. The gates of the dummy transistors were connected to the appropriate power supply to guarantee the dummy transistors were always turned off.

To achieve maximum accuracy in matching two transistors (that share a common node at their drain or source), a number of unit elements to realize each of the two transistors were used and the fingers of the two transistors were interdigitized. This approach is known as a common-centroid layout and helps to reduce mismatch errors caused by gradient effects across a microcircuit [Johns, 1997.1]. A common centroid layout for a differential pair is
shown in Figure 4.2. We see that the transistor is symmetric in both the horizontal and vertical
direction and any gradients across the microcircuit should affect $M_1$ and $M_2$ in the same way.

In the design of the ADC, a common-centroid layout was used for transistors in the input and interpolation stages. Also, the current comparator and SR latch were each laid out in a symmetric manner so that each single-ended signal went through identical paths.

By increasing the sizes of the relevant devices to improve matching, employing special layout techniques for matching and having differential operation (after the input stage), offsets will be kept to a minimum.

4.1.2 Noise Considerations

In mixed-signal integrated circuit design, special consideration must be given to digital noise coupling to the analog circuits through the power supplies or the substrate. The current in the analog portion of the ADC is fairly constant or slightly varying in time while the current in the digital portion of the ADC consists of sharp pulses. These pulses (which occur every time a digital gate or buffer changes state) cause a glitch on the digital power supply and surrounding

Figure 4.2 A common-centroid layout for a differential pair.
To avoid the digital noise from affecting the analog circuits, different power supplies and pins were used for the digital and analog portions of the design. Furthermore, different power supplies and pins were used for the digital output buffers since these inject very large current spikes as they drive large capacitive loads.

Also, the analog portion of the design was physically kept separate from the digital portion of the design. This separation will help minimize substrate noise from coupling to the analog circuits. In addition, the output digital buffers were kept as far away as possible from the analog circuits. The analog and digital circuitry were kept separate by guard rings and wells (Figure 4.3). The p⁺ connections to VSS provides a low impedance path between the substrate and VSS and minimizes the substrate noise as it propagates through the resistive substrate. The n well helps to further increase the resistance between the analog and digital portions. In the ADC, guard rings were placed around the digital portion of the design.

A shield was also used in the ADC to help isolate the substrate from the clock bus running along the integrated circuit. The use of the shield helps minimize capacitive coupling of noise between the clock bus and the substrate. The shield is realized as an n well between the substrate and the clock bus as shown in Figure 4.4.
4.1.3 Clock and Input Signal Skew

Small differences in the arrival of the clock and input analog signal to their respective circuitry can cause errors [Johns, 1997]. These signals can arrive at different times across the die due to delays over metal wires. These differences are intolerable for the speeds required for this design, since the input signal and clock are spread across 2 mm of die space.

To overcome this, a tree-like wiring structure was used for the clock and analog input signal in the layout [van de Plassche, 1994]. This is illustrated in Figure 4.5. In the conventional implementation, the signal travels different path lengths to get to its respective circuit. These different path lengths can cause large delay variations and consequently cause errors in the ADC. In the tree-like structure, we see that every circuit has identical path lengths to the signal source and therefore allows the signal to arrive to each circuit block at the same time.
Figure 4.5 Comparison of wiring layout for crucial high-speed signals.
4.1.4 Overall Layout

Figure 4.6 Layout of the ADC in 0.5 μm CMOS.
The layout of the entire circuit is shown in Figure 4.6. We see that the signal flow is from left to right. The physical layout of the blocks has a close correspondence with the electrical circuit topology (Figure 3.7). There are also test circuits for the input stage and current comparator for characterization in the upper left hand corner of the chip. Also, the lower right hand corner was used for an external, unrelated test circuit. The area of the chip is 2.02 mm x 2.02 mm. Approximately, 50% of the chip area was used for the analog portion and 50% was used for digital portion. The chip was packaged in a 44-pin Ceramic Quad Flat Package (CQFP). This package is suitable for speeds between 50 MHz to 1.5 GHz. 37 of the 44 pins were used for the ADC circuit and all related test points and circuits. We see that the layout is compact with limited whitespace.

Other important factors to account for in the layout included ensuring correct metal widths to meet current density requirements for all the different metal layers and having enough contacts and vias to handle the current going through certain connections [HP, 1995]. The wide metal lines running vertically on Figure 4.6, are power lines and needed to be wide enough to meet current density requirements. Also, metal layers, as opposed to polysilicon, were used for connections as much as possible to keep interconnect resistance and delay low. Crossing of metal lines were avoided as much as possible. However, when design constraints forced it, when possible, the lines were drawn in metal 1 and metal 3 to keep them far away from each other to decrease the coupling capacitance between the two lines. Also, due to the importance of matching in the analog portion of the design, no metal lines were drawn over transistors (except metal1, of course, to make connections to the junctions).

Resistors used in the ladder were realized with polysilicon with salicide block. Polysilicon is usually covered with salicide (conducting material) to decrease the polysilicon resistance in the integrated circuit. Blocking the polysilicon from having salicide cover it increases its resistance. A total resistance of 600 Ω was implemented with appropriate taps to generate the reference voltages. While the absolute value of the individual resistors is not
important but that they all hold the same value, matching among resistors in the ladder is
critical. If all the individual resistors change by the same amount, then the reference voltages
will remain the same, since we have voltage dividers at each tap, and only resistor ratios are
important. Therefore, careful consideration was given to keep each resistor matched to each
other. This meant maintaining the same contact configuration and layout orientation for each
resistor.

Electrostatic discharge (ESD) protection devices were not included on any I/O pads in
the design due to space limitations. These devices can occupy a large amount of silicon area
and can also degrade the performance of high-speed signal pads due to nonlinear capacitance
associated with ESD protection devices. However, for production chips, ESD protection
should be included to ensure reliable and robust protection from external sources.

4.2 Performance Results

Post extracted performance measurements of the ADC are shown in Table 4.1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Measured Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling Rate</td>
<td>150 MS/s</td>
</tr>
<tr>
<td>Input Capacitance</td>
<td>2.5 pF</td>
</tr>
<tr>
<td>Latency</td>
<td>4.5 ns (10 pF external load)</td>
</tr>
<tr>
<td>Input Range</td>
<td>1 V</td>
</tr>
</tbody>
</table>
| Total Power
Consumption     | 150 mW @ 3 V                                |
| Technology           | 0.5 µm CMOS (3M1P)                          |
| Die Size             | 2.02 x 2.02 mm$^2$                         |
| Package              | 44-pin CQFP                                 |

Table 4.1 Post extracted performance measurements of the ADC.
4.3 Test Board

Due to the frequency of operation, a two-sided printed-circuit board (PCB) (shown in Figure 4.7) was designed and built to test the ADC. The board contains all external components and connections required to test the chip. A schematic of the test board is shown in Figure 4.8.

Due to delays associated with the fabrication of the chip as well as time constraints, the chip was not characterized. Tests required for the ADC include DNL and INL measurements. These tests will give us some indication of the matching achieved in the design. Also, an indication of the dynamic performance of the ADC is required and would entail a plot of the dynamic range of the converter vs. the input analog frequency. Also, the

Figure 4.7 The layout of the two-sided PCB for the ADC.
variable nonuniform quantization should be tested and confirmed.

Figure 4.8 Schematic of the test board.

4.4 Summary

This chapter has discussed the important issues concerning the implementation of the ADC onto silicon. Special layout techniques for matching, noise issues and signal delays were employed. Post extracted measurements were presented. Also, the test board that was to be used for characterizing the ADC was presented.
5.1 Discussion

An ADC employing variable nonuniform quantization intended for a PAM digital communication system was designed and implemented in a standard digital CMOS process. The nonuniform quantization employed was referred to as FICO quantization where finer quantization was implemented around the PAM symbols and coarser quantization was implemented further away from the PAM symbols. FICO quantization showed improved dynamic range over uniform quantization for certain situations. Programmability of the nonuniform quantization was also added to the design with no additional power consumption but with increased area.

The interpolating architecture was used for the ADC. Implementation of the ADC
required careful consideration of the layout. This was due to the importance of matching among transistors in the analog portion of the chip, noise considerations of the mixed-signal circuit and other high-speed issues.

5.2 Recommendations for Future Work

As was shown in Chapter 2, the Lloyd-Max quantizer showed improvement in quantization noise over the FICO quantizer. Further exploration of the Lloyd-Max quantizer for a digital communication system is required. The ease of the implementation of the Lloyd-Max quantizer is an issue that needs to be addressed.

Also, the use of FICO quantization is highly dependent on what the input signal to the quantizer is. Having more variability in the design would improve the robustness and flexibility of the ADC within a digital communication system. In this design, there were four possible choices for the nonuniform factor. Increasing the choices in this design would not cause a power penalty but would increase the number of transistors required. Other circuit architectures to implement the nonuniform quantization would probably need to be explored to increase the number of choices of the nonuniform factor. The architectures would have to keep additional power and size to a minimum.

Since the ADC designed in this thesis is a prototype and separate from the remainder of the system it was intended for, digital linearization was not implemented. However, digital linearization would need to be implemented in a practical digital communication system.
REFERENCES


References


