SUCCESSIVE APPROXIMATION ANALOG-TO-DIGITAL CONVERSION TECHNIQUES IN MOS INTEGRATED CIRCUITS

by

James Leo McCreary

Memorandum No. ERL-554

9 October 1975

SUCCESSIVE APPROXIMATION ANALOG-TO-DIGITAL CONVERSION TECHNIQUES IN MOS INTEGRATED CIRCUITS

by

James Leo McCreary

Memorandum No. ERL-M554

9 October 1975

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

DEDICATION

This manuscript is dedicated to my wife Virginia for her patience and encouragement throughout our years of graduate study.

ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to Professor P.R. Gray for his patient supervision and ideas throughout the course of this work and to Professor D.A. Hodges for his special guidance and assistance. The recommendations of Professor L.F. Donaghey are also appreciated. In addition the author wishes to thank J. Albarran and G. Smarandoiu for their contributions and suggestions. The discussions with R. Suarez, R. Coen and R. Heald were also helpful. Special thanks go to D. McDaniel and D. Rogers for their technical support in the I.C. lab. The author also expresses his gratitude to Virginia for drafting the figures contained in this manuscript.

TABLE OF CONTENTS

Successive Approximation Analog-to-Digital Conversion Techniques in MOS Integrated Circuits

James Leo McCreary

SUCCESSIVE APPROXIMATION ANALOG-TO-DIGITAL CONVERSION

TECHNIQUES IN MOS INTEGRATED CIRCUITS

Ph.D. James Leo McCreary

Dept. of Electrical Engineering and Computer Sciences

Chairman of Committee

ABSTRACT

This research effort has been directed towards the exploration of new methods of analog-to-digital conversion which are suitable for monolithic construction. A new technique has been developed which utilizes charge-redistribution among binary weighted capacitors. An experimental integrated circuit, fabricated to test the new approach, performed analog-to-digital conversion in 23 μs with an accuracy of 10 bits ± 1/2 of the least significant bit.

CHAPTER I

INTRODUCTION

Of all the man-made methods of processing information, the fastest, most efficient and least expensive per operation is by electronic circuits. Such circuitry is usually classified as analog or digital depending upon the representation of the information it contains. An analog quantity may assume any level between two extremes and may be continuous in that region but a digital function is only permitted to have a set of discrete values which are usually separated by forbidden regions. The generalized electronic information processing machine, shown in Figure 1.1 must be capable of data flow into or out of the machine according to some control. It may also have processing units and memories. Although much information could be processed by either digital or by analog methods, some data processing is more suited to one particular method. Of course neither technique is best suited for all cases.

The rapid advancement of digital systems over the last 2 decades has been due to increasing applications that require digital methods. Some advantages of digital techniques are higher accuracies and greater immunity to noise. Large computers today are capable of maintaining precisions of the order of 1 part in 10^{18}. In contrast to analog circuits, this accuracy is not degraded or distorted by sequential operations (except for round-off error). Furthermore, digital memory systems have provided more practical means of nearly infinite storage capacity in printed, punched, or magnetic forms. Digital techniques would probably be preferred in applications requiring extensive computations, data manipulation or data management. However, both analog and digital circuits are fabricated in the same basic technologies and are therefore fundamentally capable of the same high operating speeds. Excluding analog circuits requiring coils or transformers both forms can

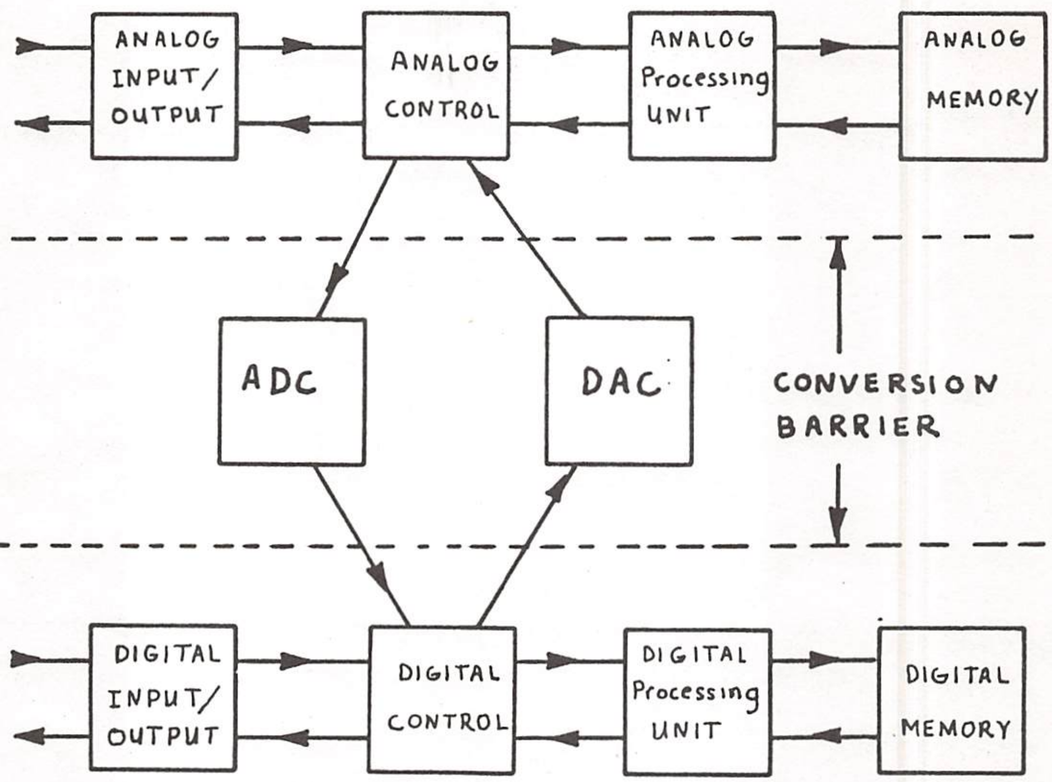


Figure 1.1: Generalized information processing machine.

generally be fabricated as integrated circuits (I.C.s), thereby realizing the advantages of small size, less power consumption, better reliability and lower costs. However, at present there is a greater availability of digital system I.C. components. A high-speed memory can be constructed to the user's specifications by wiring together several widely available I.C.s. Central processors are even available as single units.

We live in a world of analog functions. Data, usually originating as an analog quantity, is transcribed originally by people or by machines into a digital form by a process called analog-to-digital (A/D) conversion. The conversion is a prerequisite for digital processing of analog information. It should be pointed out that the accuracy of such processing cannot be more significant than the accuracy of conversion and that real-time processing cannot occur faster than the converter sampling rate. Two important qualities of the converter are, therefore, the accuracy and the conversion rate. Unfortunately both of these qualities are difficult to achieve without special fabrication techniques which are usually expensive. It is not surprising that the limitations and cost of conversion have inhibited full potential development of the generalized machine of Figure 1. That is, most information processing machines do not freely transit the conversion barrier as would be optimal conceptually. The converters are the only informational links between the two processing sections. Furthermore, the division between them is usually maintained on the premise that once the digital number is available it is more efficient to maintain it in that form until all processing is completed. This has been generally true. Although newer conversion techniques have improved both speed and accuracy, the cost of conversion has remained very high compared with the cost of other functions and therefore conversion has either been minimized or avoided.

It has been the focal point of this research to develop a substantially lower cost A/D converter (ADC) having both a high accuracy and a fast sampling rate. Such a new A/D conversion technique has been developed which can result in low cost circuits having high production yields. The new technique utilizes charge redistribution between binary weighted capacitors and a successive approximation method which rapidly converges an analog voltage into a binary number. In addition a standard N-channel MOS process can be used to implement this new algorithm and to fabricate on the same silicon chip the ADC as well as any additional logic which may comprise a larger system. The feasibility of the approach was investigated by fabricating an experimental integrated circuit which, when supported by a discrete logic system, simulated a complete ADC. The new type of ADC may be described as an "all-MOS, successive approximation, charge Redistribution ADC utilizing weighted CAPacitors" or RADCAP. The test circuit required only 23 μs to perform an accurate 10-bit binary conversion thereby verifying the new conversion algorithm [1].

An implied requirement for a lower cost circuit is that it be fabricated using only the standard processing steps required by all circuits. Any special procedures are usually expensive and therefore should be avoided. Hence only conventional photomasking techniques were used in the new design. Since the accuracy of the converter depends strongly upon the ability to construct precisely matched capacitors, the standard photolithographic process was identified as the ultimate limitation upon the conversion accuracy. This was established by gathering extensive data on capacitor ratio accuracy.

In Chapter II the principles and methods of A/D conversion will be discussed. This section is intended to be a survey of only the common conventional methods and to provide a theoretical background necessary for understanding the new technique.

In Chapter III the successive approximation method of conversion is discussed along with the basic framework for "a MOS charge-redistribution precision Voltage ATtenuator utilizing CAPacitors" or VATCAP. A conceptual model of the new RADCAP technique is then illustrated with VATCAP being a component.

The factors limiting accuracy in RADCAP techniques are examined in great detail in Chapter IV. The investigation of these sources of error represents a major effort of this research.

The factors which limit conversion rate for the RADCAP technique are discussed in Chapter V.

The experimental MOS I.C. and the digital logic system are discussed in Chapter VI and the circuit schematics are also given. In addition a precision MOS capacitor design is presented.

In Chapter VII the measurements from the first experimental I.C. and the subsequent design modifications leading to a second I.C. are examined. An experimental set-up is described in which the performance of the new A/D converter was evaluated.

The statement of conclusions from this research is made in Chapter VIII.

CHAPTER II

Principles and Methods of A/D Conversion

2.1 Introduction

The A/D conversion process consists of generalized D/A conversion and comparison operations which convert an analog input into a digitally encoded form utilizing some particular algorithm. In this chapter the most common algorithms and circuit methods of quantization will be examined.

2.2 Principles of A/D Conversion

2.2.1 Quantization Theory

The process of A/D conversion consists of two basic operations: generalized D/A conversion and comparison. This division is chosen because common functional blocks are not usually found in all types of converters. In general, however, N-bit ADCs require at least N comparison operations which may be performed in sequence or simultaneously by one or more comparators. Therefore the generalized DAC is the group of circuits which performs all functions other than those of comparison. The basic ADC is shown in a block diagram in Figure 2.1. A generalized DAC usually contains a DAC, digital logic circuits, and perhaps linear amplifiers. The function of this unit can best be understood by examining the digital output code. Assume that hereafter the analog input is a voltage level and that the digital output is binary. Although both of these assumptions are often true, they need not necessarily be true. In linear binary code, the N-bit number, $B(N) = b_{N-1}\, b_{N-2} \cdots\cdots b_1\, b_0$, where b_i equals 1 or 0, specifies 2^N binary numbers from zero through 2^N-1 inclusive and is bounded by 2^N. For a uniform ADC a linear correspondence or proportionality exists between a precise analog voltage V_i and a unique binary number B_i

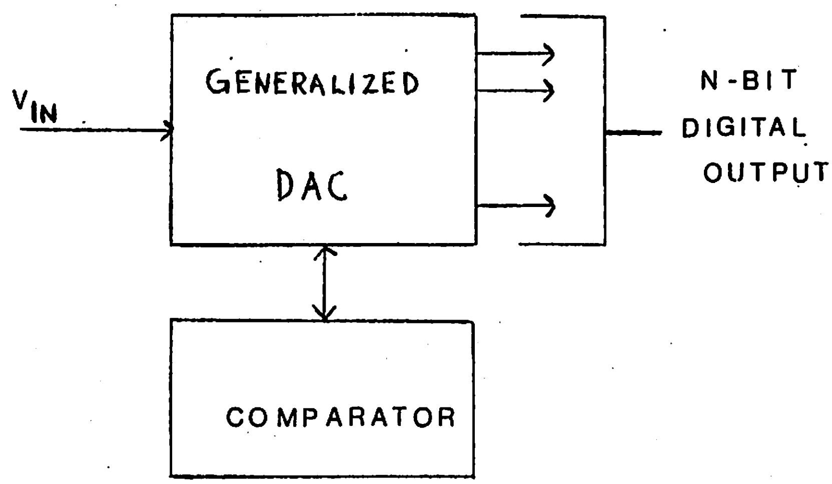


Figure 2.1: Block diagram of an ADC.

such that $0 \leq V_i < V_R$ and $0 \leq B_i < 2^N$ where $i = 0,1,\ldots.(2^N-1)$. All input voltages V_i must be less than V_R which corresponds to 2^N, the upper bound. Actually, V_R is a reference voltage which may be generated internally by the converter or else is supplied externally. External references are often chosen to be 5 or 10 volts.

At this point it has been established that for a linear converter a linear relationship must exist between a set of binary numbers B_i and a set of unique analog values V_i. Since the analog input V_{IN} is usually a continuous function, while the binary numbers are a set of discrete values, there must be an uncertainty in the quantization of the input. This uncertainty is a necessary result of correlating a continuous function and a discrete function. Exploring this further, if $V_R \rightarrow 2^N$, then $\frac{V_R}{2^N} \rightarrow 1$. The symbol "$\rightarrow$" implies linear correspondence or proportionality. Hence a voltage change of $\frac{V_R}{2^N}$ is required for a digital change of unity, or conversely,

$$(B_{i+1} - B_i = 1) \rightarrow (V_{i+1} - V_i = \frac{V_R}{2^N}).$$

The smallest change in analog input which can be resolved by a change in binary value is equal to $\frac{V_R}{2^N}$ and this is defined as the <u>resolution</u> of the converter. Again this is an inherent limitation associated with the process of quantizing a continuous-valued function into a discrete-valued function having a finite number of states. $\frac{V_R}{2^N}$ may also be called the <u>unit of quantization</u>. Therefore, assuming a uniform or linear encoder the set of values of V_i may now be computed:

$$V_i = \frac{V_R}{2^N} B_i \text{ for } B_i = 0,1,\ldots.(2^N-1).$$

A greater question still remains, however, and that concerns the relationship between the real continuous input V_{IN} and the discrete output B_i. The answer may be found by examining a segment of an ideal ADC transfer function, B_j vs V_{IN}, which is shown in Figure 2.2. The correspondence between V_j and B_j for $j = i-1, i, i+1$ is shown. A transition from B_{i-1} to B_i must occur somewhere between V_{i-1} and V_i and this is designated V_{Ti}. A similar transition occurs at $V_{T_{i+1}}$. Then $B_i \rightarrow V_{IN}$ such that $V_{T_i} \leq V_{IN} \leq V_{T_{i+1}}$. The <u>quantizing error</u> $\in$ for V_{IN} through one unit of quantization from $V_{i-1} \leq V_{IN} \leq V_i$ is:

$$\in_i = - V_{IN} + V_i \text{ for } B_i$$

and $$\in_i = - V_{IN} + V_{i-1} \text{ for } B_{i-1}.$$

For an ideal quantizer the worst case error $\in_j$, where $j = 1,2....(2^N-1)$, is minimized if all $\in_j$ are equal in magnitude to the same value $\in_q$. Evaluating the equations for $V_{IN} = V_{T_i}$ in the interval from V_{i-1} to V_i,

$$\in_q = - V_{T_i} + V_i$$

and $$- \in_q = - V_{T_i} + V_{i-1}, \text{ from which}$$

$$\in_q = \frac{V_i - V_{i-1}}{2} = \frac{1}{2} \frac{V_R}{2^N} .$$

$\in_q$ is defined as the <u>maximum quantization error</u> for an ideal, linear converter; and, as demonstrated, this value of error occurs at each transition i. Of course the particular magnitude of quantization error for any given value of V_{IN} could be less than $\in_q$ or even zero if $V_{IN} = V_i$. Although $\in_q$ is smaller for larger N, it can never be less than 1/2 of the resolution. Thus for an ideal converter, an output of B_i would imply an analog input

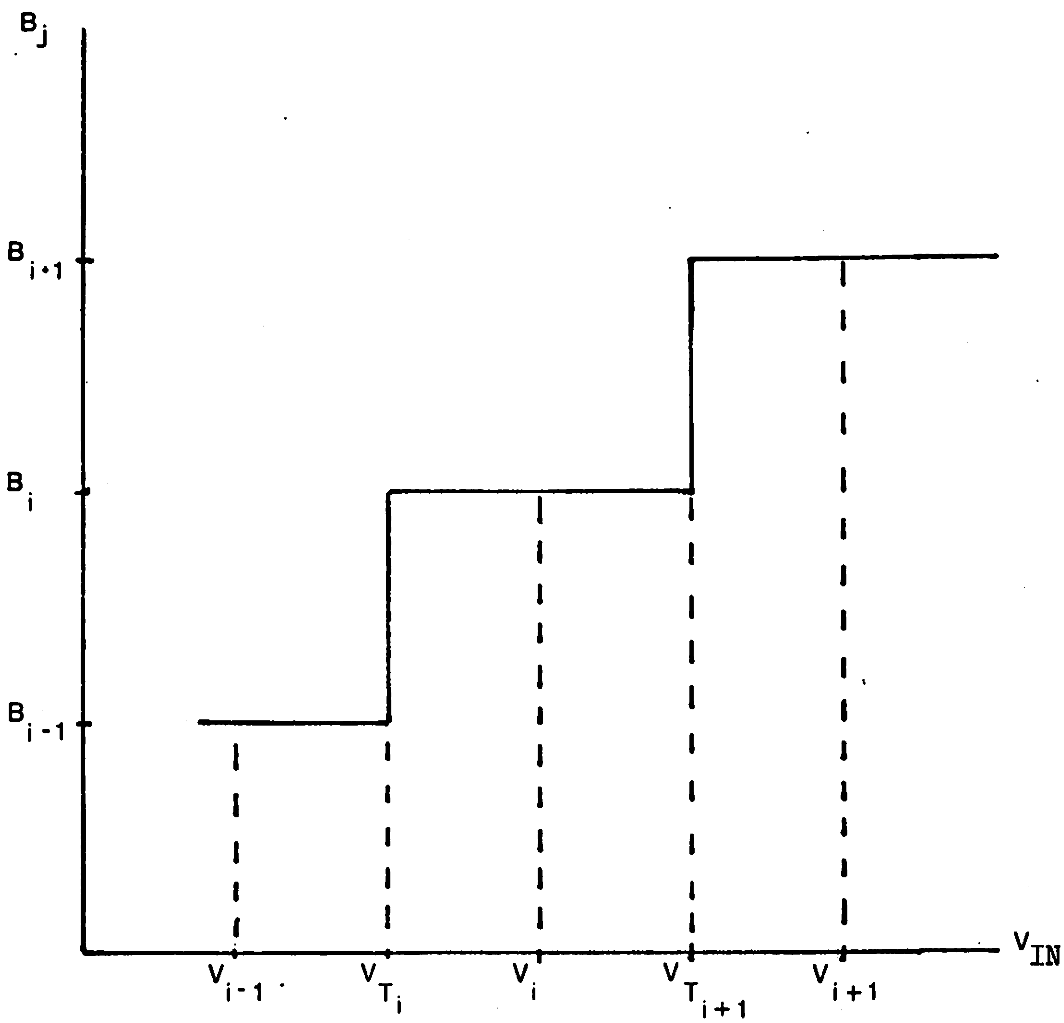


Figure 2.2: A segment of an ideal ADC transfer function illustrating the correspondence between analog input V_{IN}, the transition voltage V_j and the digital output B_j, for $j = i - 1$, i, and $i + 1$.

uncertainty of $\pm \frac{1}{2} \frac{V_R}{2^N}$:

$$B_i \rightarrow V_i \pm \frac{1}{2} \frac{V_R}{2^N} = V_{IN} = B_i \frac{V_R}{2^N} \pm \frac{1}{2} \frac{V_R}{2^N} = (B_i \pm \frac{1}{2}) \frac{V_R}{2^N} .$$

Moreover, the digital output of an ideal ADC can be no more accurate than $\pm \frac{1}{2}$ of the least significant bit (LSB). Hence the accuracy of the converter is no better than $\pm \frac{1}{2}$ LSB even in the ideal case. The transfer function and the quantization error of an ideal linear converter are plotted in Figure 2.3 for a 3-bit converter with a 10 V reference. From observation of states B_0 (000) and B_{2^N-1} (111), the ideal transfer function is characteristically asymmetric at its end points since

$$B_0 \rightarrow V_{IN} \text{ such that } 0 \leq V_{IN} \leq + \epsilon_q$$

but $\quad B_{2^N-1} \rightarrow V_{IN}$ such that

$$(V_{2^N-1} - \epsilon_q) \leq V_{IN} \leq (V_{2^N-1} + 2\,\epsilon_q)$$

Therefore the magnitude of the ideal quantization error of all digital states is no larger than ϵ_q except for B_{2^N-1} in which the error may be $2\,\epsilon_q$. This is a result of exceeding the linear input range of the converter which evidently terminates at $V_{IN} = V_{2^N-1} + \epsilon_q$.

In conclusion the generalized DAC and comparator must operate upon the analog input voltage thereby resulting in the transfer function illustrated in Figure 2.3.

2.2.2 Characterization of the Digital Output

The ideal uniform ADC was characterized by specifying the reference

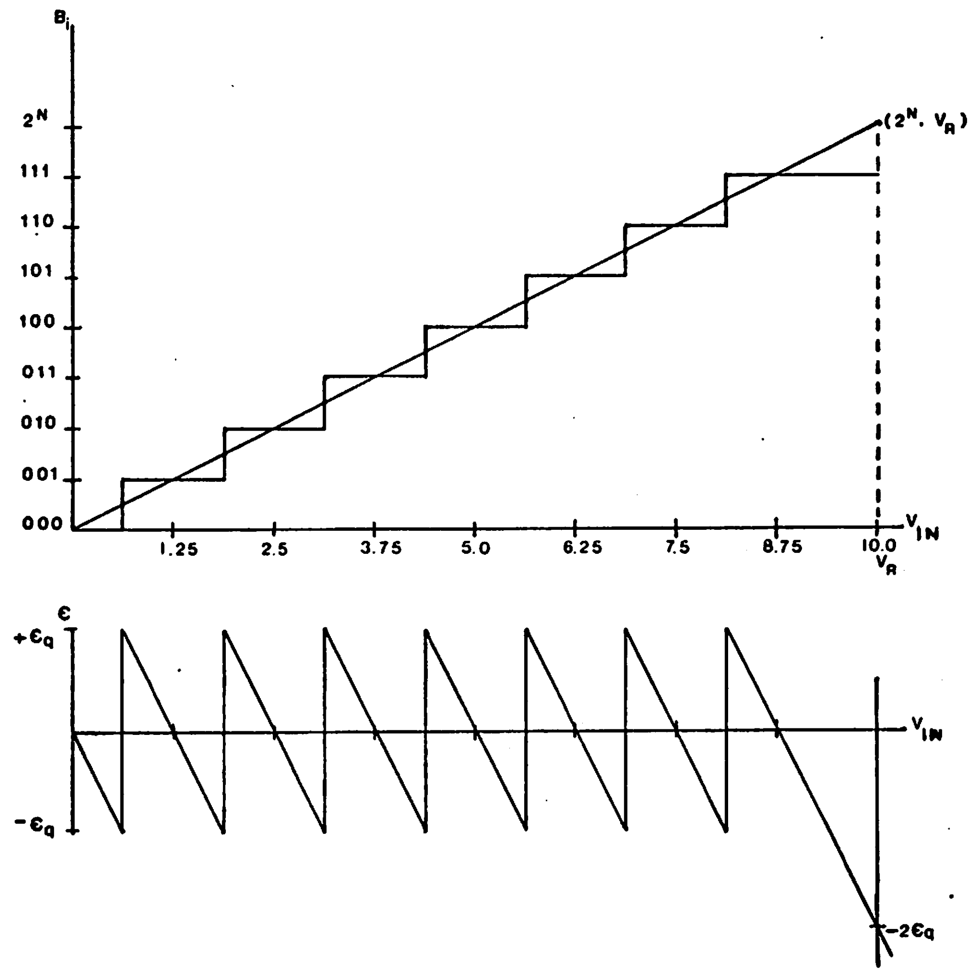


Figure 2.3: The transfer function and quantization error for an ideal ADC.

voltage and the number of bits which then determined the resolution, the accuracy and the maximum quantization error. Except for its endpoints, the transfer function was a regular staircase and each transition occurred at precisely designated points. As might be expected for a non ideal ADC the staircase is not regular and the transition points are not evenly spaced. Although the absolute transition voltage error referenced to V_R may remain unchanged the same error referenced to a fraction of 1 LSB usually becomes worse as efforts are made to increase the number of bits of resolution. One feature of the ideal converter transfer curve is its linearity. This is identified in Figure 2.3 from the fact that a straight line beginning at the origin will uniformly intersect the midpoint of each vertical transition as well as the boundary point $(2^N, V_R)$. This line is called the gain line. Furthermore, within the linear input range, a straight line will intersect all positive-going maxima on the quantization error curve and similarly for all negative-going minima also as illustrated in Figure 2.3. The set of 2^N-1 ideal transition midpoints may be listed:

$$(\text{transition midpoint})_i = B_{T_i} = [(B_i + \frac{1}{2}\text{ LSB}), \frac{V_R}{2^N}(i + \frac{1}{2})]$$

for $i = 0,1,\ldots.(2^N-2)$. The ideal transfer function has gain which is the slope of the gain line as computed from the coordinates of its endpoints:

$$\text{ideal gain} = \frac{2^N}{V_R} .$$

The transfer function for a real N-bit ADC is shown in Figure 2.4(a). The measured transition voltages V_{T_0}, V_{T_1} and $V_{T_{(2^N-2)}}$ are recorded. The difference between the real and ideal transition voltages for B_{T_0} is defined as the offset error and may be computed: offset $= V_{T_0} - \frac{1}{2}\frac{V_R}{2^N}$. The real

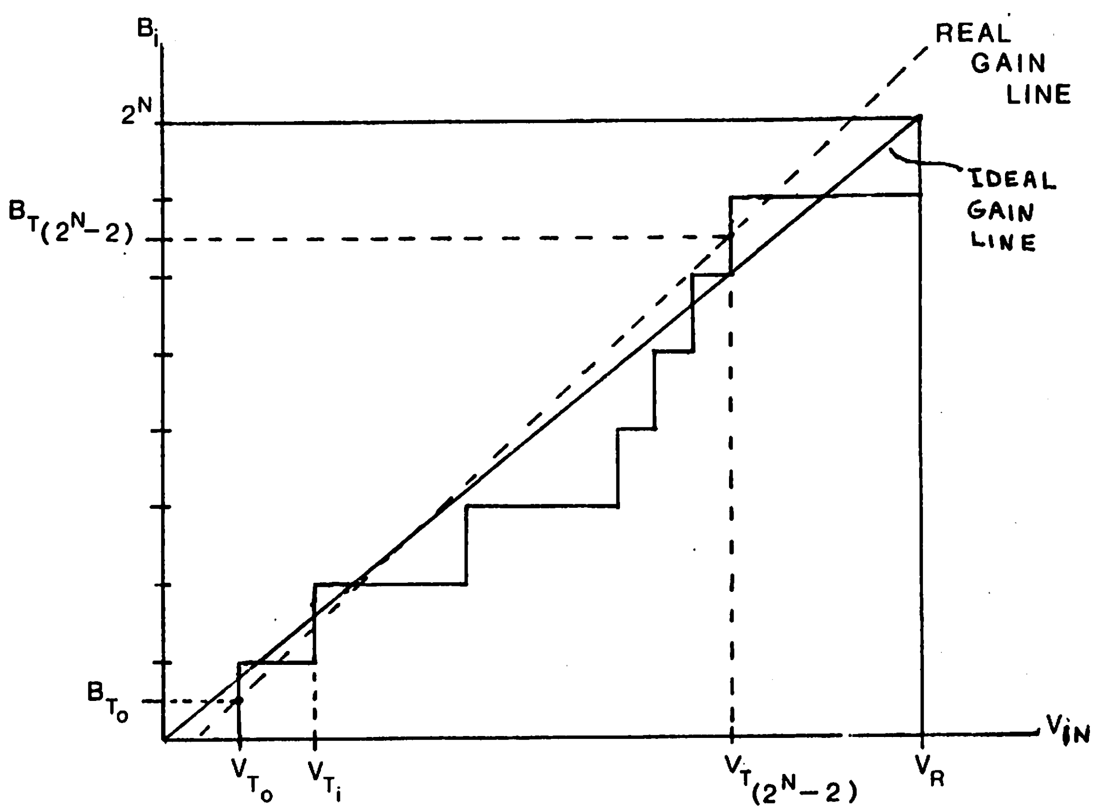


Figure 2.4(a): A real ADC transfer function having both gain and offset errors and nonlinearity.

gain is the slope of the real gain line that intersects the two points $(B_{T_{(2^N-2)}}, V_{T_{(2^N-2)}})$ and (B_{T_0}, V_{T_0}):

$$\text{real gain} = \frac{2^N - 2}{V_{T_{(2^N-2)}} - V_{T_0}} .$$

A gain error exists which may now be expressed as a percent:

$$\%\ \text{gain error} = \frac{\dfrac{2^N-2}{V_{T_{(2^N-2)}} - V_{T_0}} - \dfrac{2^N}{V_R}}{\dfrac{2^N}{V_R}} \times 100\%$$

If both gain and offset errors are adjusted to zero, as is possible with most ADCs, then the transition midpoints $B_{T_{(2^N-2)}}$ and B_{T_0} coincide with their ideal values and the real and ideal gain lines overlap. This is illustrated in Figures 2.4(b) and (c). Now the description of the real transfer function has been reduced to computing transition voltage errors relative to the ideal values. Since the curve which connects the transition midpoints is nonlinear, the parameter which describes it is the worst case deviation from the gain line or the <u>nonlinearity</u>. The nonlinearity may be defined as the worst case deviation of the transition voltage from its ideal value when gain and offset errors have been adjusted to zero. As shown in Figure 2.5 the nonlinearity may be referenced to the analog input axis and expressed as a percent of V_R or referenced to the digital axis as a fraction of the LSB:

$$\%\ \text{nonlinearity} = \frac{\Delta V_x}{V_R} \times 100\ \%$$

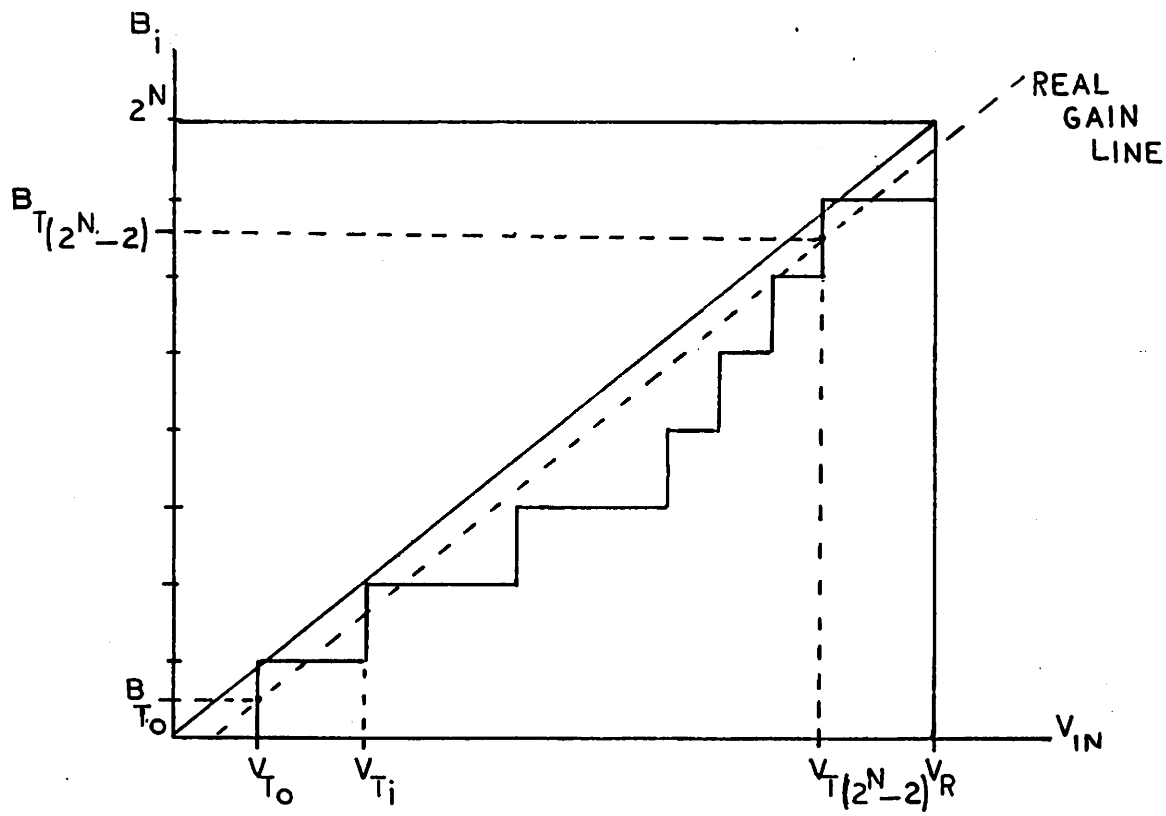


Figure 2.4(b): Gain error adjusted to zero.

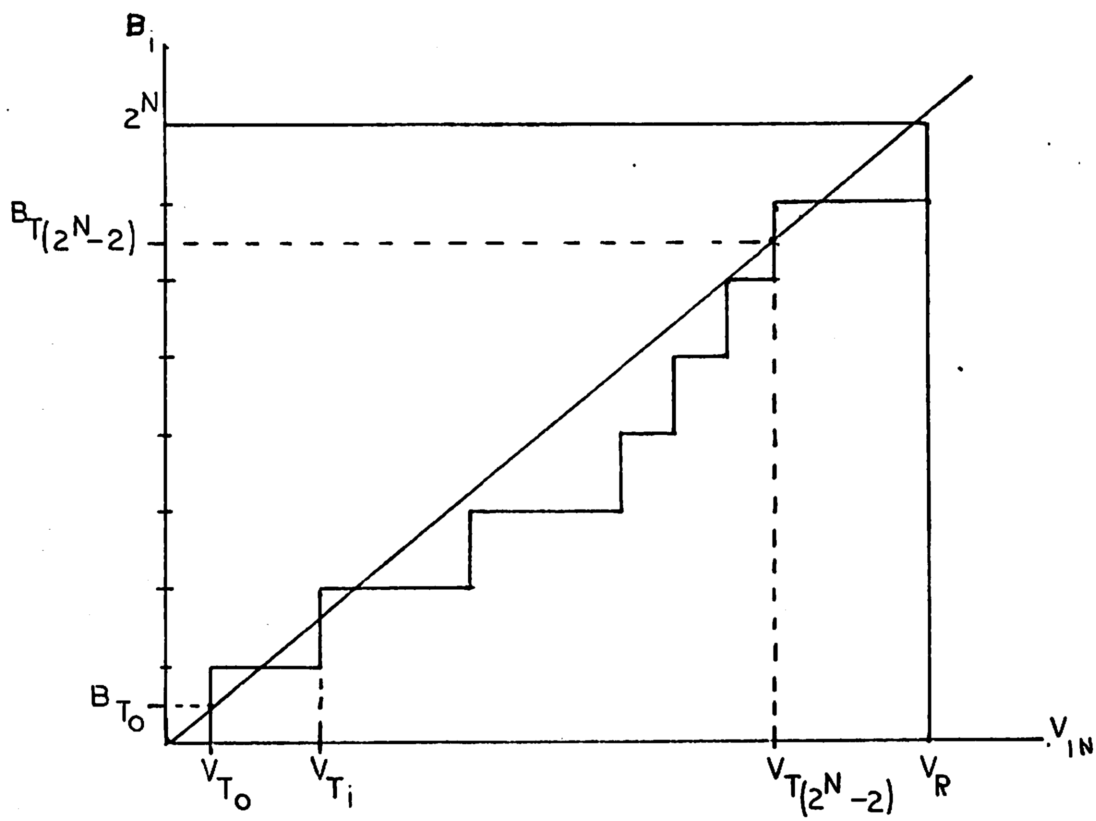


Figure 2.4(c): The ADC transfer function after offset and gain errors have both been adjusted to zero.

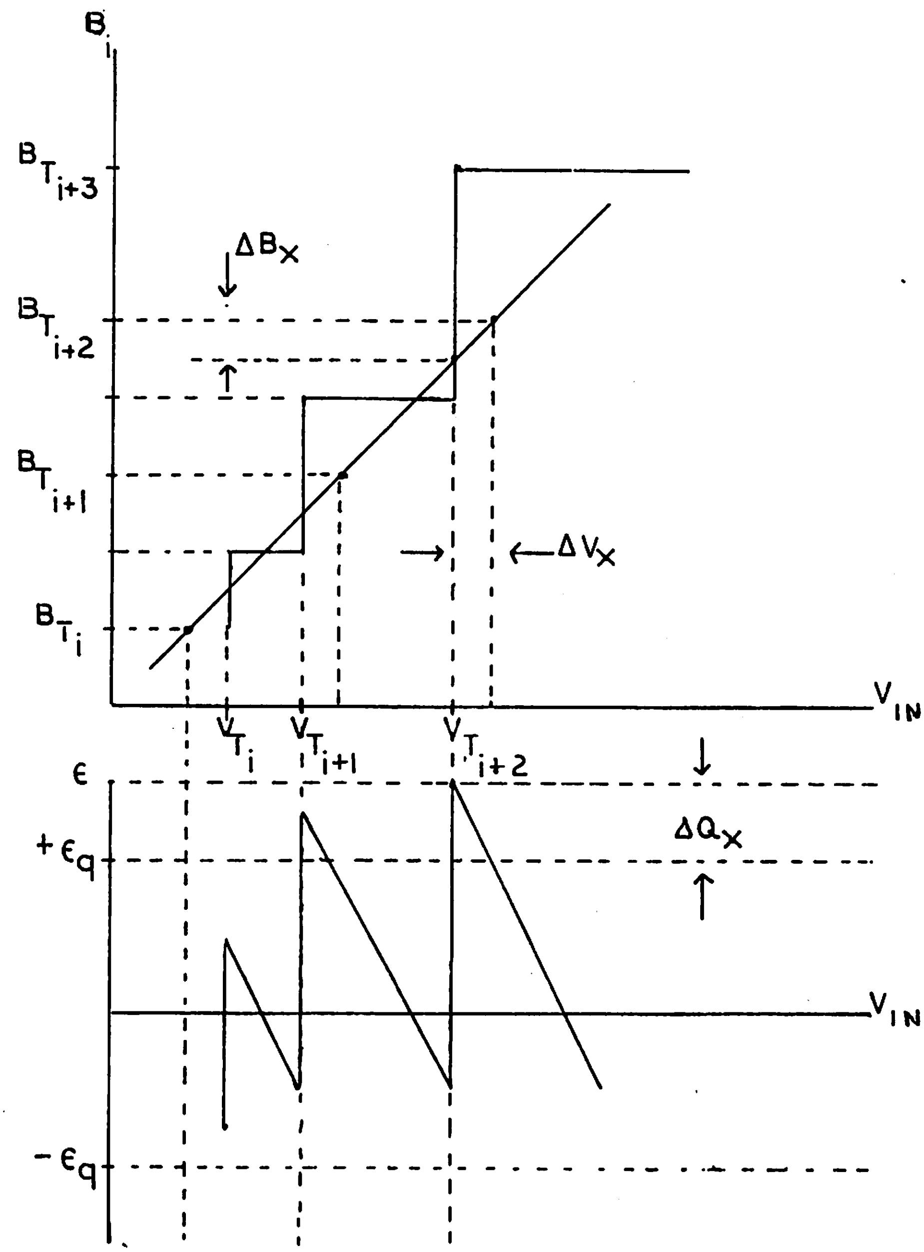


Figure 2.5: Transfer function and quantization error plots illustrating graphical determination of nonlinearity.

or (nonlinearity in bits) = ΔB_x LSB.

The corresponding quantization error plotted in Figure 2.5 indicates the nonlinearity directly as shown:

$$(\text{nonlinearity in bits}) = \Delta Q_x \frac{2^N}{V_R} \text{ LSB}.$$

The <u>accuracy</u> of an ADC is defined as the worst case deviation between V_{IN} and the digital output expressed as a binary fraction of V_R:

$$\text{accuracy} = \frac{B_i}{2^N} V_R - V_{IN}.$$

That is, an ideal converter has an accuracy of $\pm$ .5 LSB due to quantization error; but an encoder with a $\pm$ .6 LSB nonlinearity has an accuracy of only $\pm$ 1.1 LSB. The accuracy of any ADC may be expressed as the sum of quantization, nonlinearity, offset and gain errors.

Although the staircase transfer function for the real converter was shown to be <u>monotonic</u>, that is, uniformly increasing with V_{IN}, this might not be the case for some converters. This defect is called <u>nonmonotonicity</u>. Another defect, called <u>missing codes</u>, would exist if fewer than 2^N output states were present.

2.2.3 Conversion Dynamics

Since time is required for A/D conversion the dynamic performance of the converter will be degraded from its static characteristics especially for high frequency input voltages. One dynamic parameter, the <u>conversion time</u> T_c is the total time needed to perform one A/D conversion, and $\frac{1}{T_c}$ is the <u>conversion rate</u> or <u>sampling rate</u>. Two desirable features are a sampling accuracy of $\pm$ 1 LSB for any single sample, and a high sampling rate for a large input frequency range. The difficulty of achieving both of these will be investigated in the following example.

Let the input signal be a full-scale sine wave of amplitude $\frac{V_R}{2}$ and frequency f such that $V_{IN}(t) = \frac{V_R}{2} \sin 2\pi t$. If an N-bit ADC is to convert this signal within 1 LSB, then the maximum rate of change of the input is limited by the unit of quantization and the conversion rate:

$$f = \frac{1}{2^N \pi T_c} .$$

This assumes that the input voltage remains directly connected to the ADC during the entire conversion, hence it must remain stable during that period. From this equation a 10-bit converter with a 100 μs conversion period has a maximum input frequency of 3 Hz. However, the input frequency range of the converter may be extended by sampling the signal for a very short period of time and holding that value during the entire conversion. This function is commonly performed by a sample-and-hold circuit (S/H).

In conventional techniques a S/H is interfaced between the signal input and the converter input. Since this is usually realized with 1 or more operational amplifiers the cost is generally high. A simplified S/H circuit is shown in Figure 2.6. The two voltage followers act as unity gain buffers with very low output impedance. The equivalent time constant of the circuit is reduced by the low output resistance of the first follower; therefore, the only significant resistance is that of the closed switch. When the switch is opened the charge remains on the capacitor and its voltage is buffered into the converter. The time required for the S/H circuit to acquire the input voltage is called the <u>acquisition time</u> T_{aq} which is usually less than the conversion time because of the small time constant

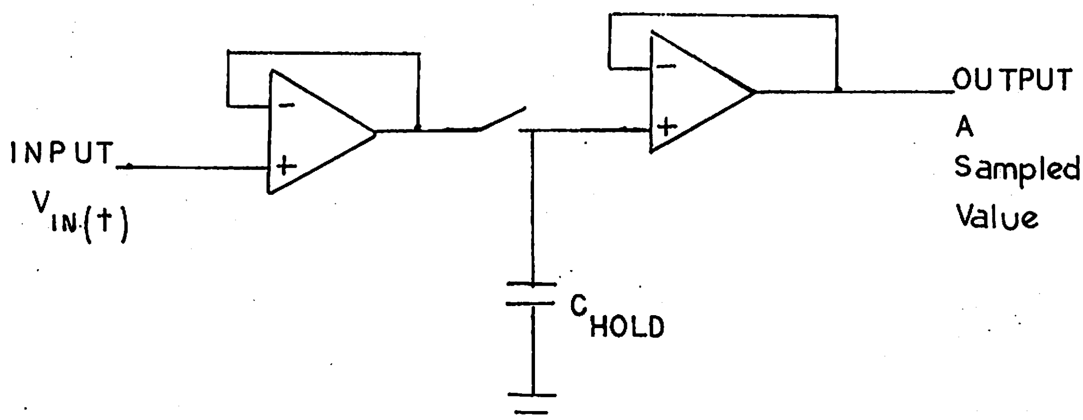


Figure 2.6: A simplified sample-and-hold circuit.

involved. At the end of the acquistion time the output is tracking the input voltage within the specified error range. The time required to open the switch is defined as the <u>aperture time</u> T_a and random variations in this time are manifested as an uncertainty in the sampled value for rapidly changing input signals. The composite ADC system shown in block diagram form in Figure 2.7 includes the S/H, the generalized DAC, and the comparator. In the timing diagram the acquistion time and the aperture time are contained in the total conversion time even though the operation of the converter does not begin until after $T_{aq} + T_a$.

For static inputs the aperture time causes no measurement uncertainty, and the acquistion time is not a limitation because the required sampling accuracy can be achieved if sufficient capacitor charging time is allowed. Therefore the only requirement is that $T_{aq} \geq M\tau$ where M is the number of $C_{HOLD} \times R_{SWITCH}$ time constants τ for the desired precision.

For dynamic measurements the maximum input frequency which may be sampled depends upon the criteria used to specify the sampled signal. For example let this criterion be a $\pm$ 1 LSB accuracy in the sampled value (in an idealistic case). In addition it will be assumed that the conversion time T_c only results in a time delayed output hence this will be ignored with respect to sampling accuracy. For this hypothetical case the maximum input signal frequency is determined by the aquisition time requirements. Figure 2.8 illustrates the equivalent S/H circuit during this time. A sine wave of frequency f and amplitude $\frac{V_R}{2}$ has a maximum rate of change equal to $V_R \pi f$. For convenience this will now be approximated by a ramp having the same slope:

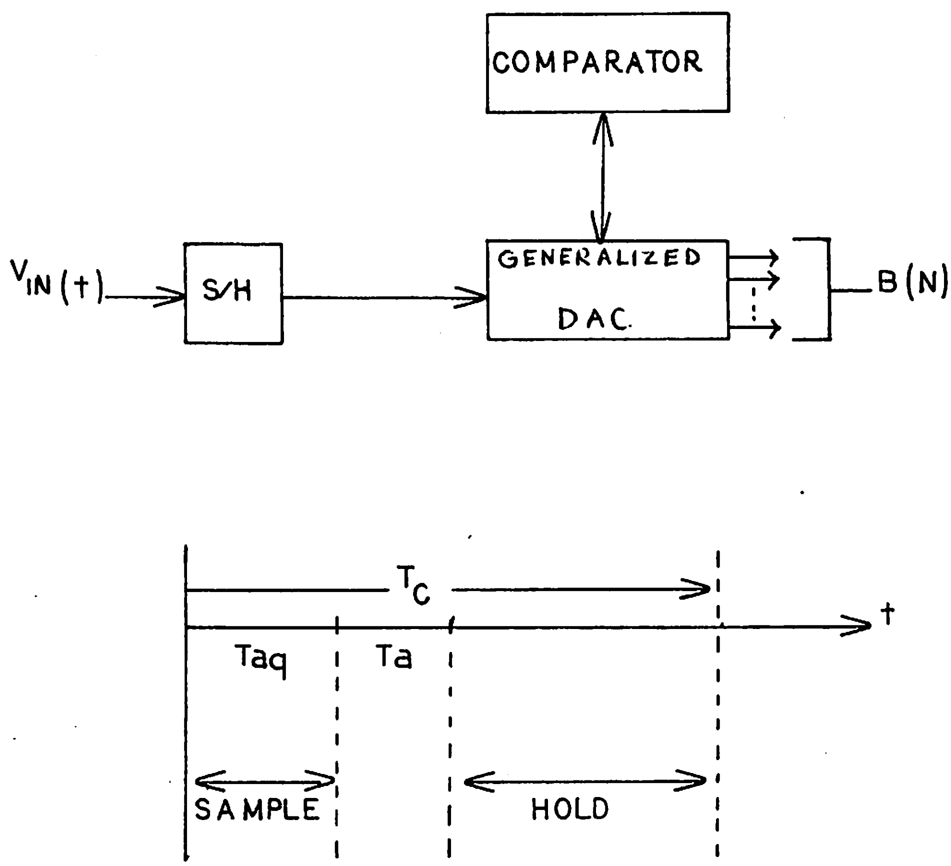


Figure 2.7: A block diagram and system timing diagram of an ADC which includes a sample-and-hold (S/H) circuit.

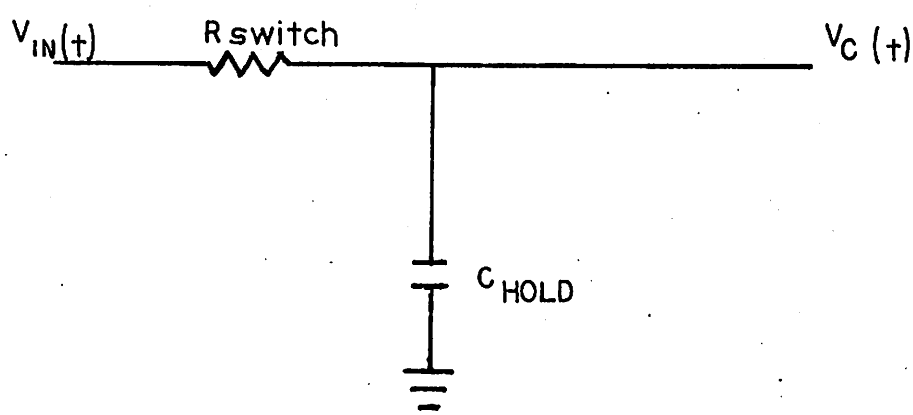


Figure 2.8: A simplified sample-and-hold circuit illustrating the important time constants.

$$V_{IN}(t) = (V_R \pi f)t.$$

Solving this circuit for capacitor voltage $v_c(t)$ and expressing the error as a fraction of an LSB:

$$\epsilon_{LSB} = 2^N \pi f \tau (1-e^{-\frac{t}{\tau}}).$$

From this equation the error converges to $2^N \tau \pi f$ for $t = \frac{1}{\pi f} >> \tau$. Evaluating this expression for 1 LSB error:

$$\epsilon_{LSB} = 1 = 2^N \tau \pi f$$

from which the maximum input frequency is

$$f_{max} = \frac{1}{2^N \pi \tau}.$$

Using nominal values of $C_{HOLD} = 1000$ pF and $R_{SWITCH} = 100\ \Omega$ for a 10-bit converter, then the maximum frequency which may be sampled to within $\pm$ 1 LSB accuracy by the S/H circuit is 3.1 kHz. The minimum acquisition time for this example is:

$$T_{aq} = \frac{1}{\pi f_{max}} = 2^N \tau = 100\ \mu s.$$

In practice, however, the acquisition time for this example would be substantially less than this value. This would be true for applications in which signals having much higher frequency components may be sampled if greater dynamic sampling error than $\pm$ 1 LSB can be tolerated (for these frequencies). For example, let the criteria for maximum input frequency be such that the sampled signal is allowed to be attenuated by $\frac{1}{\sqrt{2}}$ and phase shifted or time delayed by an arbitrary amount. However the sampled signal is not permitted to be distorted. There are many applications in communications having specifications such as these. It will be shown in Chapter V that the S/H circuit (an RC circuit) is actually a low-pass filter for the input signal. Hence from section 5.2.1 the minimum acquisition time required to meet these criteria is

$$T_{aq} = (N+1)\tau \ln 2$$

for which $f_{max} = \frac{1}{2\pi\tau}$.

In contrast with the previous example,

$$f_{max} = 1.6 \text{ mhz and}$$

$$T_{aq} = 0.8 \ \mu s \text{ for this case.}$$

It should be pointed out that according to the Nyquist principle a sampling system having a conversion rate of 10 kHz can only recover signals of less than 5 kHz. However there are some data acquisition methods which time-multiplex several 10 kHz converters to produce a sampling system capable of recovering higher frequency signals. An adaptation of this technique referred to as a pipeline converter may be used to perform parallel A/D conversion [2].

2.3 Techniques for D/A Conversion

All methods of A/D conversion require a generalized DAC as a component of the ADC. The process of D/A conversion will first be discussed as a necessary prelude to A/D conversion. Although current output converters are available, this discussion will be restricted to voltage output since only this concept is under study.

2.3.1 Summation of Binary Weighted Currents

A simplified form of weighted-current DAC is shown in Figure 2.9. Assuming that the operational amplifier is ideal, the binary ratioed currents are summed at the inverting input. Then

$$V_0 = -\frac{R_0}{R}\left(1 + \frac{1}{2} + \frac{1}{4} + \cdots\right) V_R = \left(-\frac{R_0}{R}\right)\frac{V_R}{2^N} B_i;\; B_i = 0, \cdots 2^N - 1$$

and V_0 becomes a quantized function of a binary number B_i. The difficulty of implementing this scheme is that the required range of resistor values becomes unmanageable. For example, a 10-bit converter would require that one resistor be of value (1024 $\pm$ .05%) R, and this precision is difficult to achieve with most fabrication processes due to the unfavorable aspect ratio of large resistors. In addition the large resistances and associated capacitances would result in slow speeds for some of the switches. A more practical circuit for achieving weighted currents is shown in Figure 2.10 [3]. In this schematic of a 4-bit D/A converter there are 4 current sources which comprise a QUAD. The temperature compensating diode D_T stabilizes the currents over a given temperature range. In order to maintain the same voltage drop across all base-emitter junctions the current density in each junction must be the same. This is achieved by having the number of emitters proportional to the total emitter current. The binary digit values are input through diodes. A '1' input will reverse bias the diode and the transistor will conduct; however, an '0' will divert all of the resistor current to ground. Although driving current into ground may be somewhat wasteful of power it is necessary in order to avoid thermal gradients when high precision is required. If the QUAD were extended for

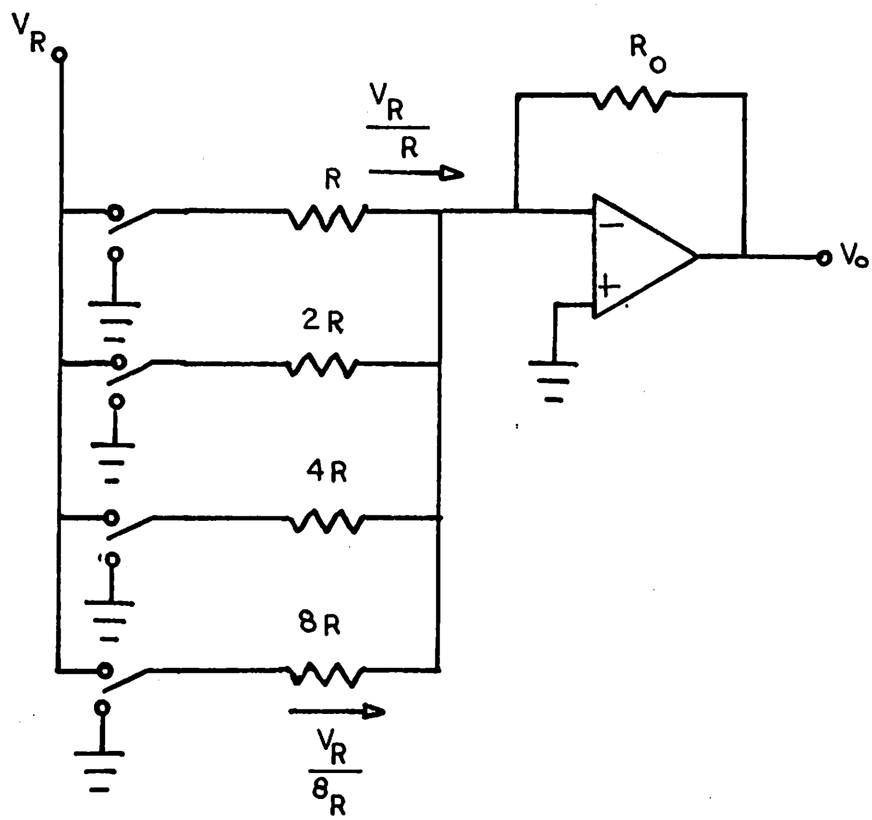


Figure 2.9: A simplified weighted-current DAC.

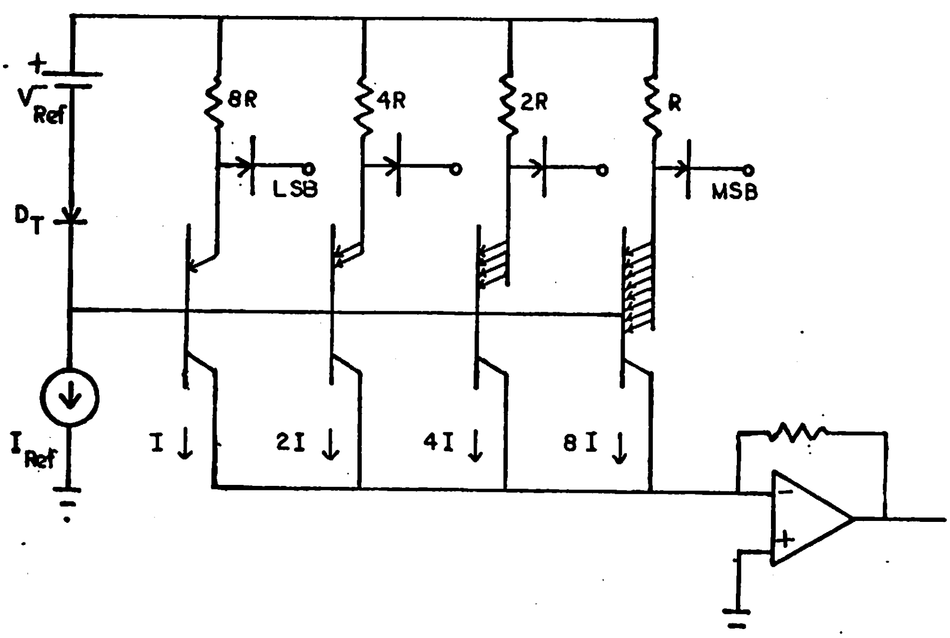


Figure 2.10: A circuit technique for achieving weighted currents.

a 10-bit converter, there would be one transistor requiring 1024 emitters and it would be difficult to retain matching precision for such a device. The actual circuit method groups the current sources in QUADs and interconnects these with current dividers in order to achieve the proper binary weighting. This is done in Figure 2.11. The QUADs are identical thereby avoiding awkward aspect ratios on both transistors and resistors. However, the resistive dividers now contribute to the total error, hence precision must be maintained in these elements too.

A second approach is the use of an R - 2R resistor ladder network to generate binary weighted currents. This is illustrated in Figure 2.12. Binary currents are established by the current division property of the ladder. This technique has the advantage of requiring matching of only two resistor sizes: R and 2R.

2.3.2 Binary Attenuation of Equal Currents

Another technique that also employs an R - 2R resistor ladder shown in Figure 2.13 performs binary attenuation of equal-valued currents. If $I_A = I_B = I_C = I_D = I$ then the output current of the ladder is given by $I_{out} = \frac{I_A}{24} + \frac{I_B}{12} + \frac{I_C}{6} + \frac{I_D}{3} = \frac{I}{24}(1+2+4+8)$ and the quantized output voltage is:

$$V_{out} = \left(-\frac{I}{24} R_{out}\right) B_i .$$

This method has the advantage of requiring identical current sources and resistors R and 2R which may be more easily matched. On the other hand there are more resistors required.

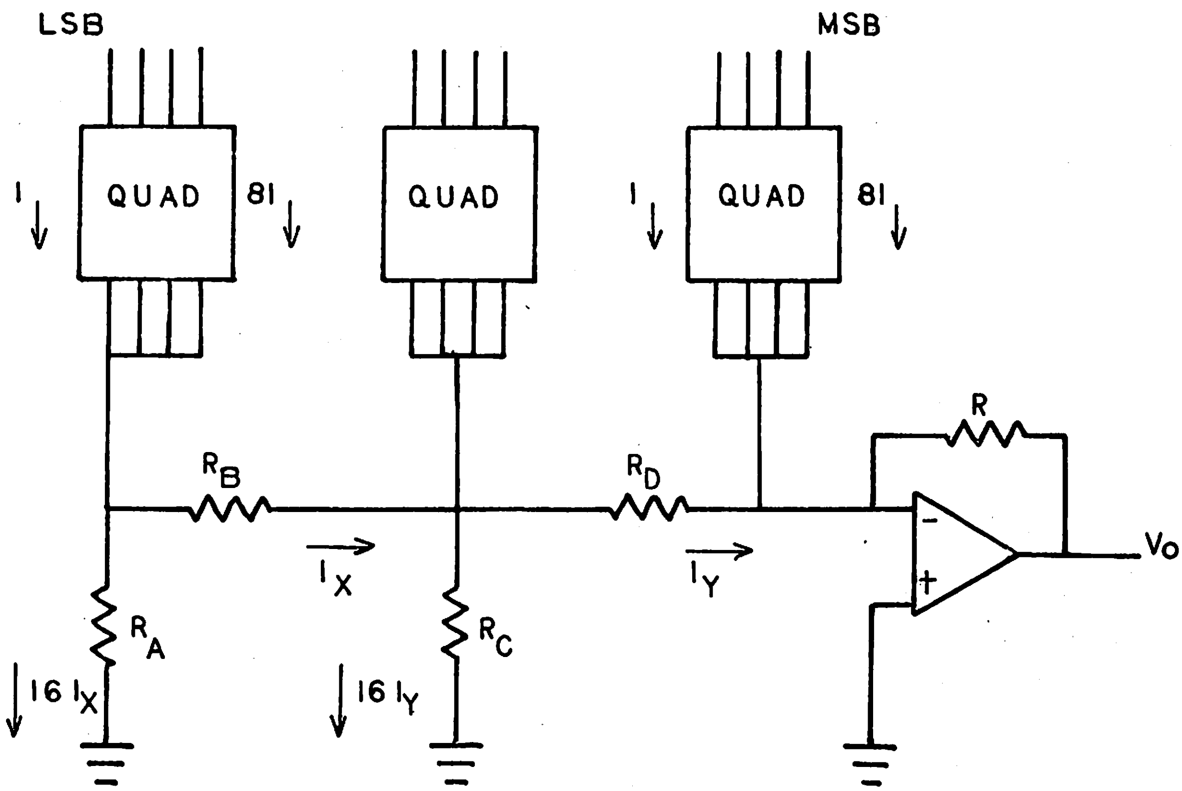


Figure 2.11: Quad current sources with two resistive current dividers.

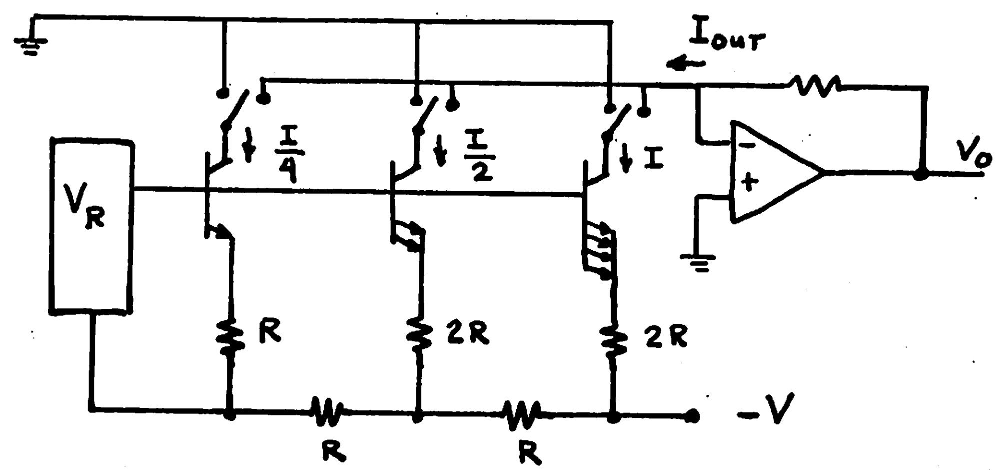


Figure 2.12: A DAC using summation of binary weighted currents generated by an R-2R resistor ladder network.

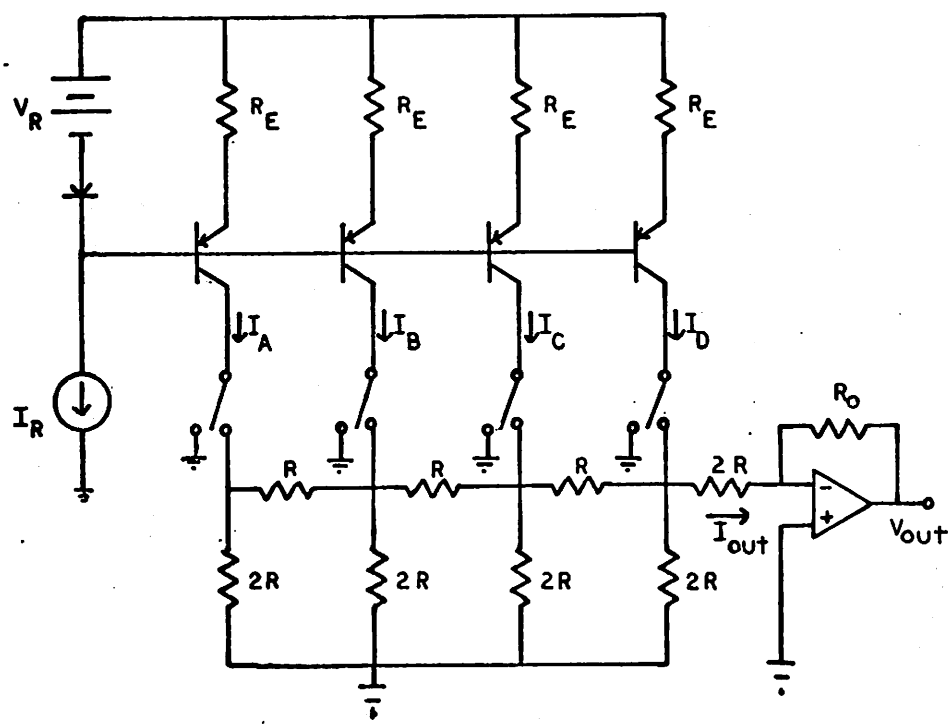


Figure 2.13: A DAC using binary attenuation of equal-valued currents.

2.3.3 Charge Redistribution

Two capacitors with different initial voltages may be connected together resulting in a final voltage that is dependent upon the total charge. Figure 2.14 illustrates a simplified 2-capacitor circuit. For equal capacitors

$$V_0 = \frac{V_A + V_B}{2}$$

since the initial voltage V_A equals zero or V_R and V_B is initially discharged by S3, then V_B must be a binary fraction of V_R:

$$V_B = \frac{V_R}{2^N} B_i = V_0.$$

After N redistributions the output voltage is quantized.

2.3.4 Integration Types

A DAC may be conceptually formed by integrating the charge on a capacitor. If a constant current source of value I charges the capacitor for a discrete number of clock periods, $T_{CLK}B_i$, then the final voltage is a quantized function of time:

$$V_0 = \frac{I}{C} T_{CLK}B_i.$$

This concept is illustrated in Figure 2.15 in which the counter is initially cleared and counts up to some desired number B_i. The term $\frac{I\ T_{CLK}}{C}$ must remain constant over the range of interest. This places severe tolerance limitations on the three variables involved. For this reason the structure illustrated in Figure 2.15 is not actually used to realize a DAC. However, integration methods which are fundamentally the same as that just discussed are often utilized for ADC circuits. Two particular forms of integration

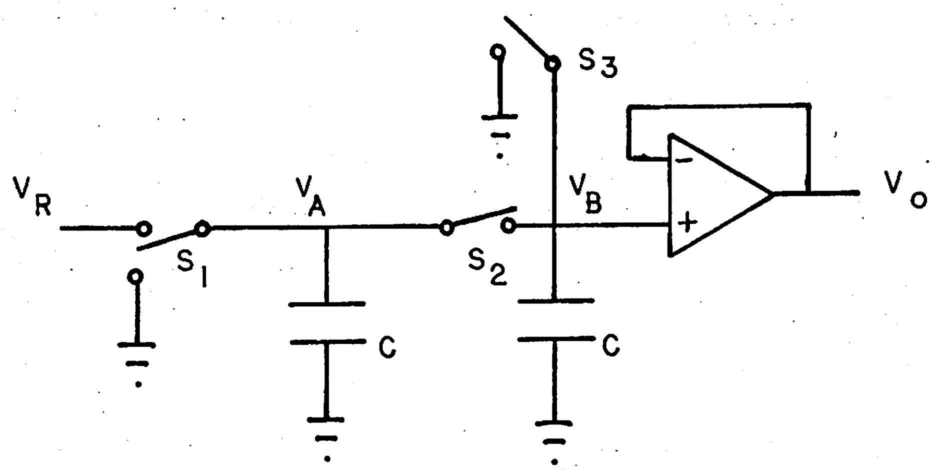


Figure 2.14: A 2-capacitor charge-redistribution DAC.

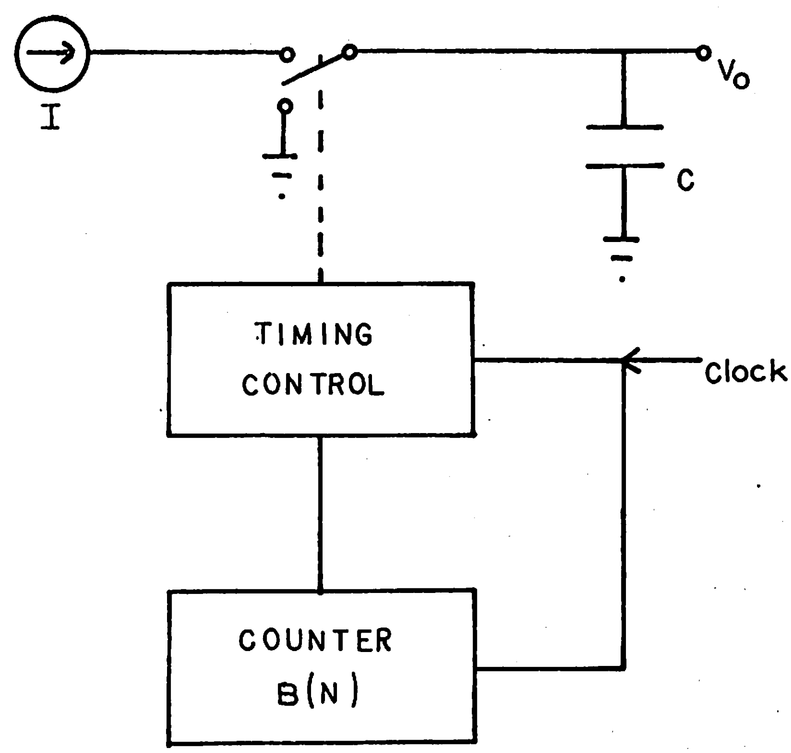


Figure 2.15: An Integration type of DAC.

type conversion circuits are single slope and dual slope converters which are discussed later.

2.4 Techniques for A/D Conversion

2.4.1 Serial Methods

The simplest method of A/D conversion is a serial method. This technique performs a linear sequential process of conversion such that 2^N comparisons are usually required for a full scale N-bit conversion. A conceptual serial ADC may require a few operational amplifiers, a capacitor and logic circuits as the primary components. A particular advantage is that none of these must be precision components. Although this method is characterized by circuit simplicity, its features include low cost and high accuracy. On the other hand it is very slow. For example, a 12-bit converter with a 5 mHz clock would require 820 μs to convert a full-scale input.

Serial ADCs are usually integration type conversion circuits. A particular form of serial ADC using integration methods is illustrated in Figure 2.16 [4]. This circuit may be classified as a single slope type because the slope $\frac{dv_c}{dt}$ has only one value of interest I/C. During the regular conversion cycle the capacitor is charged to V_{IN} and then discharged to the comparator threshold V_{TH} in time T_{IN} by the current source I. Then

$$T_{IN} = C(V_{IN} - V_{TH}) = T_{CLK}B_{IN}$$

where B_{IN} is the binary number of clock periods required to discharge C from V_{IN} to V_{TH}. During one phase of operation, the capacitor is charged to V_R and

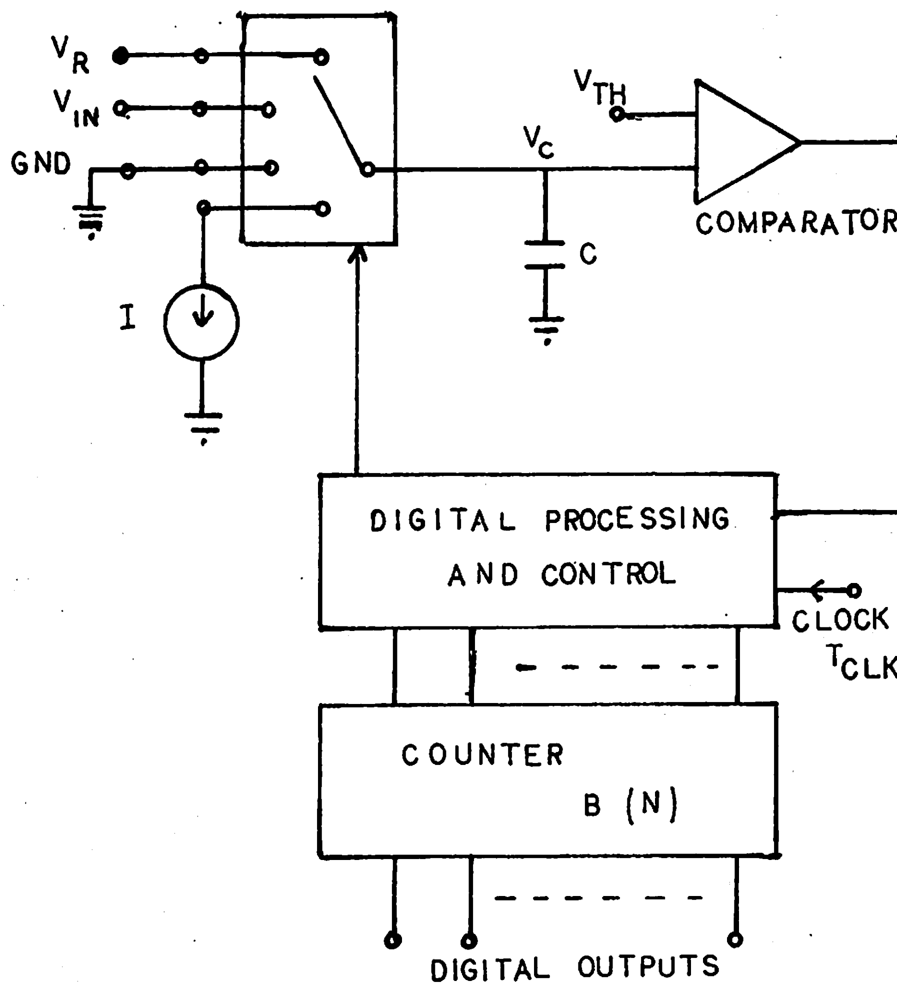


Figure 2.16: A single-slope type of ADC.

$$T_R = C(V_R - V_{TH}) = T_{CLK}B_R.$$

Since the comparator threshold V_{TH} is negative, ground may also be calibrated:

$$T_G = C(0 - V_{TH}) = T_{CLK}B_G.$$

The digital output is then expressed as a binary fraction of the reference:

$$B_i = \frac{V_{IN}}{V_R} \quad \frac{B_{IN} - B_G}{B_R - B_G}$$

and is independent of the value of T_{CLK} and C. The only precision requirements are that I and C be constant over the voltage range of interest during the time of each conversion. Disadvantages of this technique are the very slow conversion rate and the need for a division. Other characteristics of the single slope method are low cost, high accuracy and the inclusion of an intrinsic S/H function. Single slope ADCs have not been widely produced commercially since previous designs have required precision components or exact timing control.

Another serial method of A/D conversion which is more widely available than the single slope types is the <u>dual slope</u> ADC. It is shown in Figure 2.17. During a fixed time interval the switch contacts V_{IN} and the capacitor, which was initially discharged, now charges with a current $\frac{V_{IN}}{R}$. When V_O equals V_{TH}, the comparator threshold, the counter begins counting up to a fixed number:

$$B_X = \frac{RC\ (V_X - V_{TH})}{T_{CLK}\ V_{IN}} .$$

At this point the switch is connected to $- V_R$ and the counter is reset and begins counting again. The value of the count when V_O equals V_{TH} is:

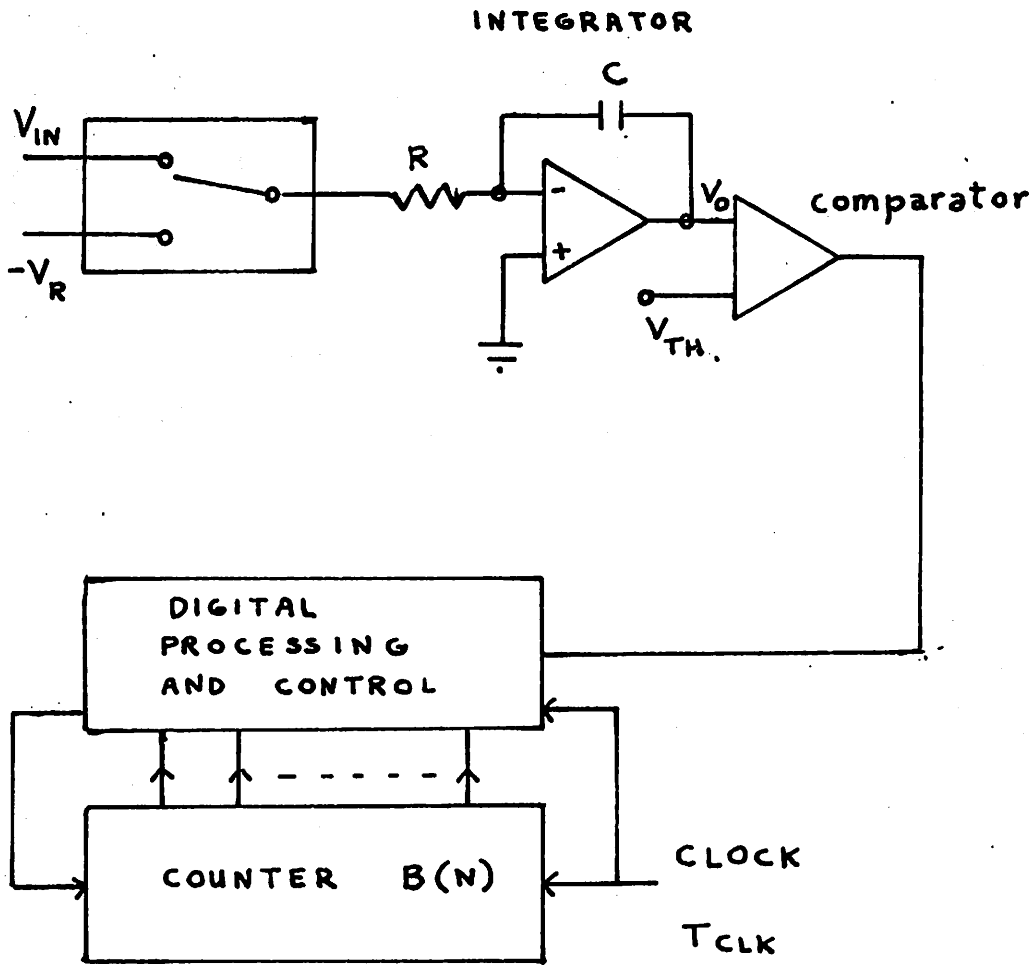


Figure 2.17: A dual slope ADC.

$$B_R = \frac{RC\ (V_X - V_{TH})}{T_{CLK}\ V_R} \ .$$

If B_X equals 2^N, a fixed count, then the following result is obtained:

$$B_R = 2^N \frac{V_{IN}}{V_R} \ .$$

The input has been quantized and converted. Furthermore, the accuracy is independent of the value of T_{CLK}, R, C, or V_{TH} and depends only upon the linearity of the integrator and the hysteresis of the comparator, both of which may be controlled. The dual slope converter has similar properties as other serial types: low conversion rates, low cost, and high accuracy. In contrast to the single slope method, the subtractions and divisions are not needed but also there is no S/H function and unfortunately on operational amplifier is required versus a constant current source. Also a negative reference is needed. The dual slope technique has gained wide acceptance in applications having only slowly varying inputs. Commercially available converters of this type are usually multi-chip realizations because the operational amplifier requires a bipolar technology while the digital circuitry is best realized in MOS technology.

2.4.2 Successive Approximation Methods

Successive approximation ADCs are probably more available commercially than any other types and usually have high resolution and fast conversion rates of 10 kHz to 10 mHz. This high speed is achieved by a quantizing algorithm which converges exponentially rather than linearly as in the previous methods. This is accomplished by successive comparisons of the input with binary fractions of the reference. The control logic operates upon the D/A converter by testing all N bits sequentially beginning with the MSB. This is done by assuming the bit value is '1' and then comparing the DAC output to V_{IN}. If this output is less than V_{IN} then the bit value is actually '1', otherwise, it is returned back to '0'. In this manner 2^N possible binary numbers could be tested in only N operations. The usual circuit configuration is illustrated in Figure 2.18. The DAC is usually the weighted current source or resistive ladder type. The digital control circuitry required for this scheme is usually quite simple. The accuracy of the output is dependent upon the comparator and the DAC. However, it is usually the resistor matching difficulty in the latter which creates the nonlinearity.

2.4.3 Parallel Conversion

A parallel ADC requires the simultaneous reference generation and comparison for all 2^N-1 transition voltages [5]. Thus 2^N-1 fractions of V_R must be formed by the converter and delivered to 2^N-1 separate comparators. This is shown conceptually in Figure 2.19 for a 2-bit parallel ADC. The parallel DAC consists of a string of 2^N resistors. Although this guarantees monotonicity at the DAC the offset variations of the comparators may result in missing codes. Also the linearity may become difficult to maintain if the number of resistors becomes very large. In addition the need for 2^N-1 comparators probably precludes monolithic high resolution parallel ADCs because the chip area required increases exponentially with N. Therefore this method is generally limited to very low resolutions. Parallel ADCs are very expensive however they offer conversion rates up to 25 mHz. Another feature is that they are inherently asynchronous and therefore do not require clock signals.

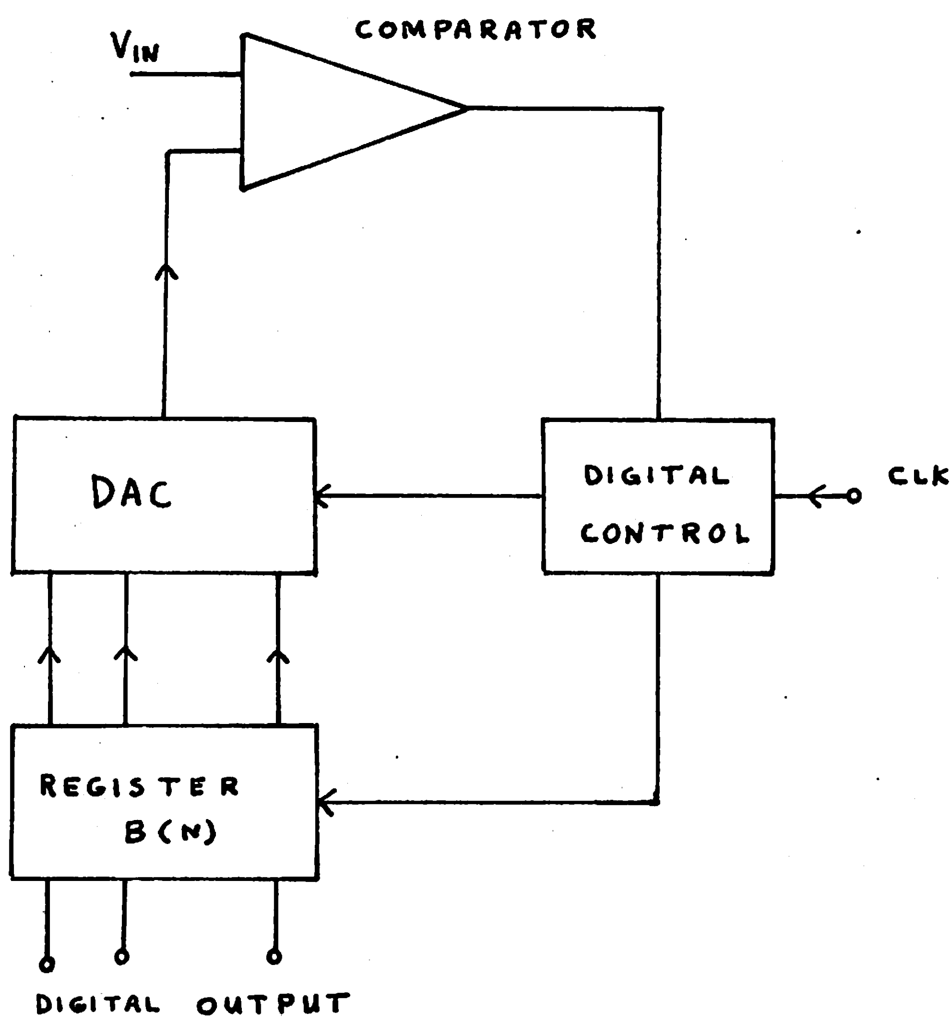


Figure 2.18: A successive approximation ADC.

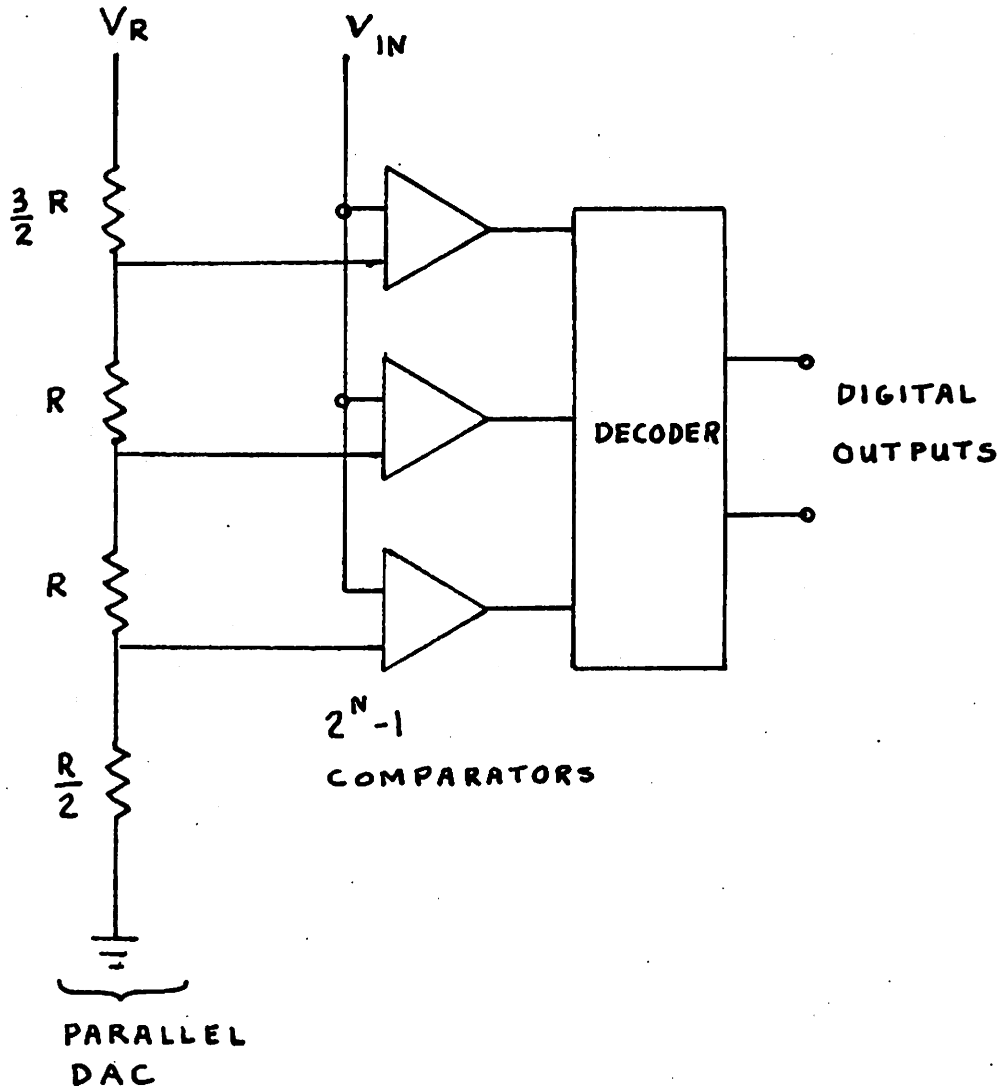


Figure 2.19: A parallel ADC.

2.5 Technologies for ADC Components

2.5.1 Introduction

Integrated circuit (I.C.) technologies, or fabrication techniques, have had a great impact upon electronic circuits. Linear circuits such as operational amplifiers, comparators, buffers, and current sources which were once made from discrete components are now commonly available as single-chip, bipolar I.C.s. An even greater development of digital I.C.s has occurred due to the low cost, high density advantages of MOS technology. In fact, MOS memories may be constructed from just a few chips while an entire CPU may be purchased as a single monolithic I.C. In contrast to either linear or digital circuits, converters must perform both analog and digital functions. While computational operations with digital circuits can be done with arbitrarily large precisions, such accuracies are not easily achieved in analog circuits, in fact, they are not always required. In order to minimize the error in analog operations, the standard bipolar technology often requires additional complexity such as component trimming and off-chip adjustments or external circuitry. In this section a survey is made of the conventional fabrication methods of ADC components. Methods of realizing comparators, control logic, and supporting functions will also be discussed.

2.5.2 ADC Component Technology for Charge Redistribution and Integration Methods

Charge-redistribution (C/R) methods have commonly been utilized in successive approximation ADC techniques, however these circuits have usually contained discrete components such as resistors and capacitors and bipolar op amps. On the other hand it has recently been demonstrated that a single chip, two capacitor C/R ADC may be realized with MOS technology [6].

Conventional integration methods, used in most serial type ADCs, are usually single slope or dual slope converters. These techniques commonly require op amps and often a discrete capacitor. In contrast a prototype serial ADC using the single slope method has been recently developed using MOS technology and having a potential for single chip realization [7].

2.5.3 ADC Component Technology for Precision Resistor Networks

Both the R-2R ladder and the weighted current source methods require precision resistors. Therefore, the following analysis will examine the three different methods of fabricating precision resistors for DAC circuits.

It would be convenient, of course, if diffused resistors could be made accurately since they are compatible with the standard bipolar technology and require no additional processing steps. However, sheet resistance ρ_s is usually limited to less than 200 Ω/square, creating difficulty in matching resistors over a wide range of binary values. Therefore, the binary weighted resistor method is not considered practical for high resolutions when diffused resistors are used. On the other hand, the R-2R ladder requires equal resistors of only two values. Therefore this approach is more suitable for diffused resistors. However there are still constraints which require proper design in order to achieve high accuracies. For example, the emitter resistors R_E of the equal current sources shown in Figure 2.12 must have an IR drop greater than 2^N times the V_{BE} mismatch. Small currents would be desirable for low power dissipation and small thermal gradients, therefore large resistors having high sheet resistance ρ_s would be required. On the contrary high resistivities in the R-2R ladder result in greater nonlinearities due to higher voltage coefficient of resistance (VCR). This arises because the resistors are actually formed with reverse-biased pn junctions having depletion region widths which are dependent upon junction voltage. The effective ρ_s is therefore somewhat dependent upon the IR drop. Since the depletion region width depends upon the doping gradient at the junction, this effect can be reduced by using low resistivity diffusions in the ladder. A tradeoff arises and the resistors must be fabricated in a manner which minimizes error due to the 2 effects just mentioned. At least one DAC is commercially available which utilizes a diffused R-2R network as part of a 10-bit monolithic bipolar I.C. [8].

In contrast to diffused resistor techniques, impurity ions may be accelerated by an electric field and driven into a silicon surface by a special process called ion implantation. The depth of penetration is generally shallow but the resultant impurity profile and concentration may be controlled. This technique is capable of providing very high resistivity values with dimensional tolerances determined primarily by the photomasking process. Present data indicates that better matching can be achieved than for diffused resistors but at the expense of one implantation. An experimental 10-bit DAC has been built by a commercial manufacturer [9]. It is an R-2R ladder type and is a complete monolithic I.C. utilizing ion implantation and bipolar technology.

A more complex method of resistor fabrication utilizes materials containing nichrome or tantalum or cermet (Cr-SiO) which are deposited as thin-films onto an insulating substrate. Patterns may then be etched to form thin-film resistors. Subsequent oxidation or annealing steps will increase or reduce ρ_s respectively. By carefully controlling these processes, sheet resistances from 10 to 2500 Ω/square can be achieved. In contrast to other types of resistors, thin-films may be trimmed by laser techniques thus enhancing the matching capabilities by one or two orders of magnitude [10]. It would appear that unlimited matching precision could be achieved in this manner; however, there are limitations. For example, the laser beam itself may oxidize or anneal parts of the resistor during the trim causing an uncertainty. Furthermore, materials evaporated by the beam may land on top of previously trimmed resistors thereby changing their values. In addition thin-film resistors have non-zero long-term drift and almost always require passivation layers. In spite of these difficulties, added cost, and complexity, thin-film resistors have generally been preferred over diffused or implanted resistors because they may be trimmed to 13 or 14-bit matching accuracies while the other 2 resistor types cannot be easily trimmed and have been generally limited to 10-bit precisions. Therefore, converters requiring high accuracies (exceeding 10 bits) almost always use discrete or trimmed thin-film or thick-film resistor networks. These may be either R-2R ladders or weighted current source types. The switches and other components of these DACs have been realized in bipolar technology. Some attempts to use CMOS technology and thin-film networks have needed external bipolar comparators and references [11].

2.5.4 Comparator

Most high-speed precision comparators available today are realized in bipolar technology. Although MOS and CMOS comparators do exist they have generally larger offset voltages, slower switching speed and different power supply requirements. Furthermore since most DACs already utilize bipolar circuits there is generally better compatibility with comparators of the same technology. For these reasons bipolar comparators prevail although MOS comparators have been realized as special purpose sub-units of MOS LSI circuits.

2.5.5 Digital Logic and Control

The most favorable technology for digital control and switching logic is MOS due to its high functional density, low cost, and process simplicity. CMOS technology has advantages of greater logic swing and lower d.c. power but at higher cost and complexity and lower functional density than MOS. Both of these methods require less power and area than bipolar logic and are usually much more desirable for digital circuits except in cases where speed is the ultimate concern.

2.5.6 Supporting Functions

Supporting functions such as amplification, buffering and voltage referencing have been realized almost exclusively in bipolar technology. Precision unity gain and high gain amplifiers are still difficult to achieve in MOS technology. In addition, stable, accurate voltage references have been developed only in bipolar technology.

2.5.7 Summary

In conclusion, ADC components which require high precision circuit elements such as resistor networks, voltage references, op amps and unity gain buffers strongly favor bipolar and thin-film technologies. However, complex digital circuits are more advantageously realized as MOS chips.

CHAPTER III

All-MOS, Successive Approximation, Weighted Capacitor, Analog-to-Digital Conversion Technique--RADCAP

3.1 Introduction

There are many circuit methods which perform A/D conversion as discussed in Chapter II. One particular method which is considered in more detail in this chapter is based on the successive approximation algorithm. This scheme has advantages of high speed operation for the amount of circuit complexity required. In section 3.3 MOS technology is considered as one fabrication method which may be used to realize analog and digital circuits. A particular technique for realizing a successive approximation ADC in MOS technology is by charge-redistribution on binary weighted capacitors, as discussed in section 3.4.

3.2 Successive Approximation A/D Conversion

3.2.1 A Comparison of Successive Approximation Method and Other Techniques

The conversion of an analog input voltage V_{IN} into N binary bits of resolution requires that the input be compared with a subset of (2^N-1) different fractions of the reference V_R. These fractions actually correspond to transition voltages as discussed in Chapter II. Hence it must be possible in the course of the A/D conversion that any fraction $(B_i - \frac{1}{2}) \frac{V_R}{2^N}$ (where $1 \leq B_i \leq 2^N-1$) be generated for comparison with V_{IN}. These fractions may be generated in the form of a continuous ramp (quantized by a clock or counter) in the case of a serial converter. In this scheme all values of $(B_i - \frac{1}{2}) \frac{V_R}{2^N}$ may be tested, but only in a sequential or linear manner. That is, after V_{IN} is tested against $(B_i - \frac{1}{2}) \frac{V_R}{2^N}$ the next comparison is with $(B_{i+1} - \frac{1}{2}) \frac{V_R}{2^N}$. The advantages of the serial method include greater circuit simplicity and fewer precision components than other methods. However this scheme is very slow since an average of $\frac{2^N}{2}$ tests may be required for conversion. In contrast, the parallel ADC compares V_{IN} with (2^N-1) fractions of V_R that are simultaneously generated and transmitted to (2^N-1) different comparators. The parallel converter is extremely fast since only 1 time interval is required for comparison; however, the need for such a large number of components results in a much greater circuit complexity than for other methods. The successive approximation ADC represents a compromise in both speed and complexity compared with serial and parallel methods. The converter must still be capable of generating (2^N-1) possible fractions of V_R but only N tests are required for N-bit resolution as contrasted with the serial method. This is because each subsequent test except for the first is actually a conditional test which depends upon the outcome of the last comparison. This method is characterized by high conversion rates (though not as high as for parallel converters) and high resolutions with intermediate circuit complexity. Although a particular application may be better suited for either serial or parallel conversion methods, there are many applications for successive approximation techniques and therefore this method has gained wide acceptance.

3.2.2 The Successive Approximation Algorithm

Attention is now focused upon the actual sequence of operations in the successive approximation algorithm. This analysis is aided with the flow chart characterizing the algorithm which is shown in Figure 3.1. The operation begins by assuming that the MSB is '1' since initially the "next most significant untested bit" is the MSB. The fraction $\frac{A}{B} V_R = (\frac{2^N}{2} - \frac{1}{2})$ is thereby generated and compared with V_{IN}. If V_{IN} is greater than this

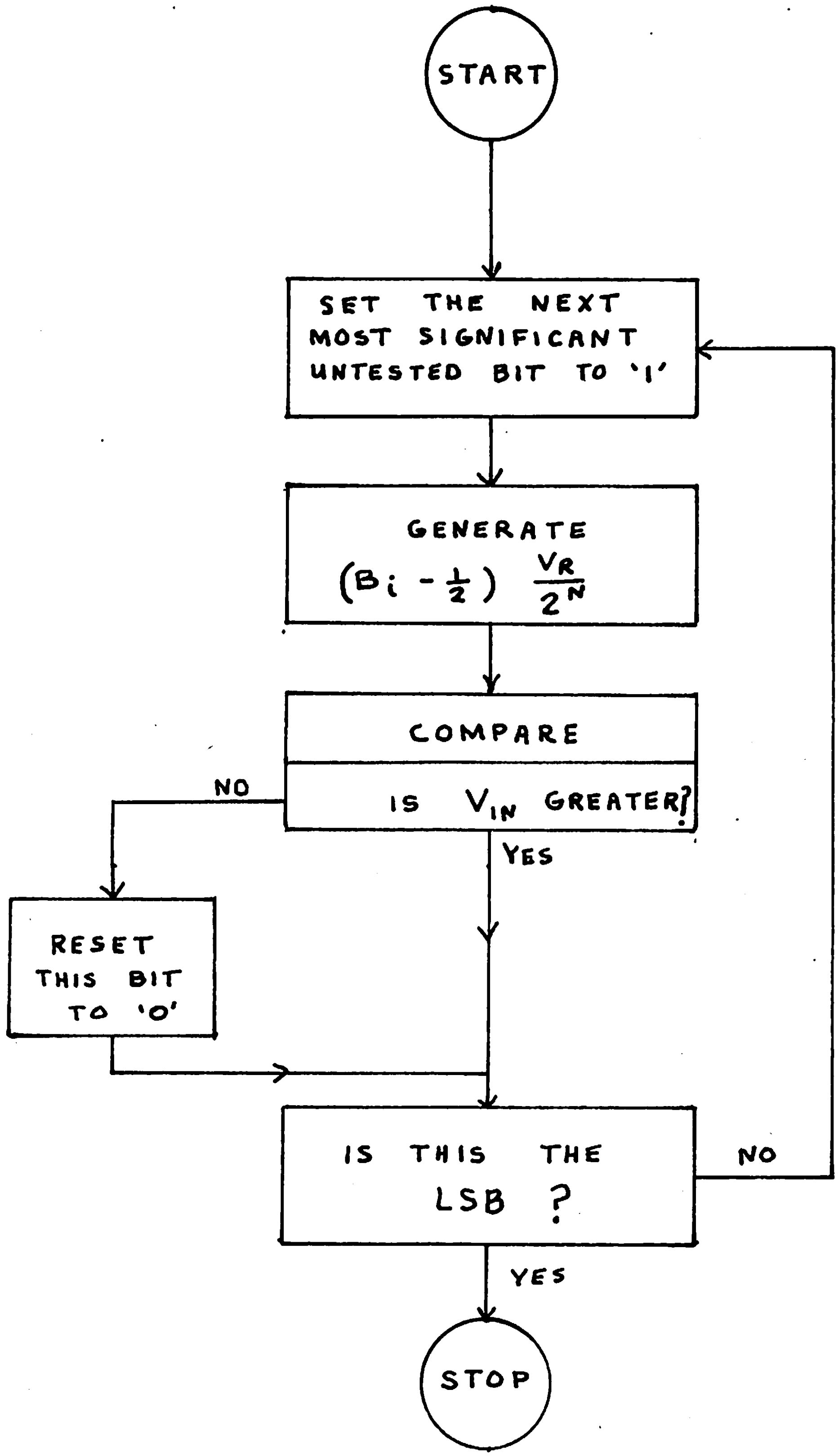


Figure 3.1: The successive approximation algorithm.

voltage then the value of this bit must be '1' as assumed. However if V_{IN} were less than the generated voltage then the value of the bit under test is reset to '0'. If the bit just determined were the LSB then the conversion is terminated, otherwise the next bit is tested just as before. The successive approximation converter therefore conducts a binary search for the best approximate digital value of V_{IN} by successively guessing V_{IN} and then determining whether or not the guess was correct. In this manner the conversion succeeds in N successive comparisons.

3.2.3 Precision Component Requirements for DACs

The most common circuit realization of the successive approximation ADC was previously shown in Figure 2.18. The register, digital control, and comparator may be realized as monolithic I.C.s without serious difficulty; however, the DAC must contain precision components of some kind. These have commonly been precision resistor networks or current sources. For high resolution converters the realization of these precision elements requires design considerations and often additional fabricational complexity. The need for precise components arises from the necessity of generating precise binary fractions of V_R which are accurate to within 1 part in 2^N. As the desired resolution increases, the allowed component tolerance diminishes.

3.3 Factors Influencing the Choice of Technology for Monolithic Realization of a Successive Approximation ADC

3.3.1 Advantages of MOS Realization

It was illustrated in section 2.4 that conventional techniques for high speed A/D conversion require high performance analog circuitry such as op amps, precision elements of some kind, and also digital circuits for counting sequencing and data storage. This has tended to result in multi-chip circuits consisting of one or more bipolar analog chips, possibly a thin-film resistor network, and also a MOS chip to economically perform the digital functions [12]. In converters utilizing precision resistor networks the high degree of resistor matching required has been incompatible with standard I.C. technology for resolutions greater than 8-10 bits. The high cost of precision networks has prevented the realization of high speed low cost converters. Since cost reduction is an industry objective as well as a goal of this research work, considerations must first be given to choice of technology. Lowest fabrication cost on a die area and gate count basis is achieved with MOS technology, having added advantages of high functional density, low power dissipation, and fabrication process simplicity. However, MOS technology has been primarily applied to digital logic. If the required analog processing for the A/D conversion can be performed in MOS technology, a low cost single-chip MOS realization is possible. In addition a MOS ADC has greater potential for future application in that it is process compatible with a MOS microprocessor circuit and therefore could possibly be placed on the same die [13]. Therefore MOS compatibility is desirable.

3.3.2 Realization of Precision Attenuator Networks Compatible with MOS Technology

A precision attenuator is necessary for successive approximation A/D conversion. This component performs quantized attenuation of a reference voltage and provides binary fractions of this reference for comparison. The attenuator performs a quantizing function similar to that of a DAC. In conventional successive approximation methods the attenuator usually contains a weighted resistor or R-2R network. In either case precisely matched current sources and resistors are required. The fabrication of matched-resistor networks and current sources in MOS technology does not appear to be practical for several reasons. First a weighted resistor method would require larger sheet resistivities than could be reasonably achieved with a standard MOS process. Second, even an R-2R ladder would need special design considerations in order to achieve the level of matching required for resolutions in the 10-bit range. Third, a MOS weighted current source or R-2R ladder network would require a MOS device to be used as a current switch. However, the "ON" resistance of the MOS device is much larger than for bipolar junction transistor switches and this resistance would have to be carefully scaled over such a wide range of values that high resolutions would not be easily obtained.

In contrast to its utilization as a current switch, the MOS device, used as a charge switch, has inherently zero offset voltage and as an amplifier has very high input resistance. In addition, capacitors are easily fabricated in metal gate technology. Therefore, one is led to use capacitors rather than resistors as the precision components, and to use charge rather than current as the working medium. This technique, referred to as charge-redistribution, has been used in some discrete component ADCs for many years [14]. However, these converters have required high-performance operational amplifiers which are difficult to realize in single channel MOS technology. Therefore, the design objective of this research centered upon the development of a precision charge-redistribution MOS attenuator which does not require an op amp.

3.3.3 VATCAP--A MOS DAC

In this section the utilization of weighted capacitors to perform precision binary attenuation is discussed. The basic component is the precision attenuator VATCAP. It has been shown in section 2.2.1 that for an ideal ADC

$$V_{IN} = B_i \frac{V_R}{2^N} - V_E \text{ where } B_i = 0, 1, \ldots (2^N-1).$$

The error voltage V_E represents the quantization uncertainty such that $-\frac{1}{2} \frac{V_R}{2^N} \leq V_E \leq \frac{1}{2} \frac{V_R}{2^N}$,

and $$V_E = \frac{B_i V_R}{2^N} - V_{IN}.$$

This equation illustrates the two basic operations which must be performed. First the fraction $\frac{B_i V_R}{2^N}$ must be generated by VATCAP, which is actually performing a D/A conversion function. Then this fraction must be compared with V_{IN}. The equation shows this operation as a subtraction which makes V_E either positive or negative.

The circuit realization of the first operation, the precision attenuation, can be achieved using charge-redistribution. More specifically the formation of the term $\frac{B_i V_R}{2^N}$ is desired. One way to accomplish this is with charge-redistribution between two capacitors. Assume that the initial voltages V_x and V_y in Figure 3.2 are equal to zero and that now V_y takes on the value V_R. The final voltage V_x, resulting from the charge-redistribution is:

$$V_x - V_R \frac{CB}{CA + CB}$$

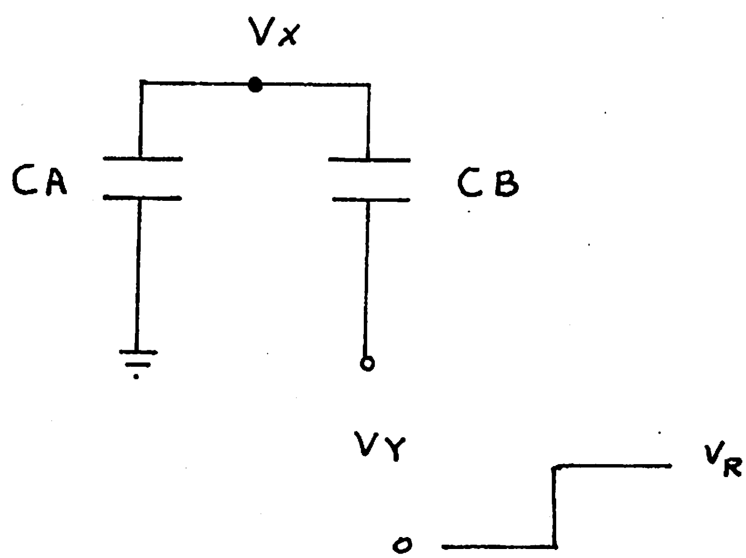


Figure 3.2: A charge-redistribution operation between 2 capacitors.

Furthermore it is desired that $V_x = B_i \frac{V_R}{2^N}$, therefore $\frac{B_i}{2^N} = \frac{CB}{CA + CB}$. The relationship between CA and CB for any value of $B_i = 0, 1, \ldots (2^N-1)$ may be determined by arbitrarily setting $CB = B_i C_1$ where C_1 is referred to as the capacitor of "unity weight" ($CB = C_1$ if $B_i = 1$). Now CA may be computed, $CA = (2^N - B_i)C_1$. Hence both CA and CB are integral multiples of the unity weight capacitor. Significant information can be extracted by expanding the equations for C_A and C_B. Since B_i is an arbitrary binary number from 0 through 2^N-1,

$$B_i = D_1 2^M + D_2 2^{M-1} + \cdots + D_M 2^1 + D_{M+1} 2^0$$

where $M < N$, and D is the binary bit value. Then $CB = (D_1 2^M + \cdots D_{M+1} 2^0)C_1$ and by expansion of 2^N,

$$CA = (2^{N-1} + 2^{N-2} + \cdots + 2^1 + 2^0 + 2^0$$

$$- D_1 2^M - \cdots - D_{M+1} 2^0)C_1.$$

From this expression it may be deduced that CA and CB can be configured from a string of binary weighted capacitors plus an additional capacitor of unity weight. Of the members of the string, those which are not used as components of C_B, must be contained in C_A. Any reference $\frac{B_i}{2^N} V_R$ may be generated from a single supply V_R and an array of binary weighted capacitors having two of unity weight. The error voltage V_E may now be formed by precharging V_x to an initial value $- V_{IN}$ rather than zero. The previous analysis remains valid by superposition. In conclusion the voltage V_x actually corresponds to the error voltage V_E which is the desired function.

3.4 A/D Conversion Using Charge-Redistribution on Weighted Capacitors--RADCAP

In section 3.3.3 the framework was developed for an MOS precision voltage attenuator, VATCAP, which performed the D/A conversion function. The attenuator was proposed to be an array of binary weighted capacitors with an additional capacitor of unity weight. If a DAC were desired the voltage output of the array after buffering would provide that function. However, an ADC is actually the design objective. In this case it is only necessary to add $- V_{IN}$ to the top plate of the array initially and then test the sign of V_E ($V_E = \frac{B_i V_R}{2^N} - V_{IN}$) after each redistribution. In this section a particular capacitor array structure is examined and the manner in which $- V_{IN}$ is stored and V_E subsequently compared with zero is illustrated in detail.

One realization using binary weighted capacitors to perform A/D conversion is illustrated with a conceptual 5-bit version of the converter shown in Figure 3.3 [15]. It consists of a comparator, an array of binary weighted capacitors plus one additional capacitor of weight corresponding to the LSB, and switches which connect the plates to certain voltages. A conversion is accomplished by a sequence of three operations. In the first, the "sample mode" (Figure 3.3), the top plate is connected to ground and the bottom plates to the input voltage. This results in a stored charge on the top plate which is proportional to the input voltage V_{IN}. In the "hold mode" of Figure 3.4 the top grounding switch is then opened, and the bottom plates are connected to ground. Since the charge on the top plate is conserved, its potential goes to $- V_{IN}$. The "redistribution mode," shown in Figure 3.5 begins by testing the value of the MSB. This is done by raising the bottom plate of the largest capacitor to the

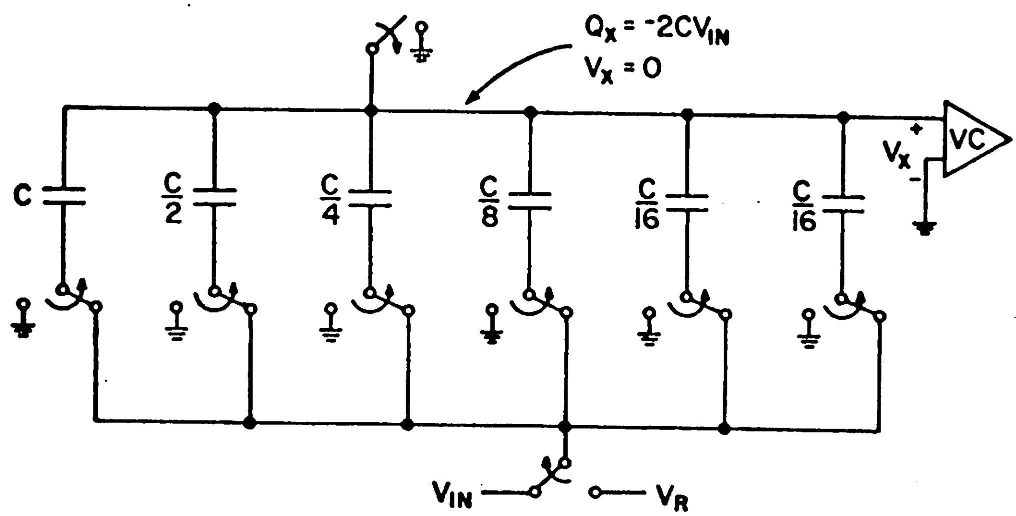


Figure 3.3: The sample mode.

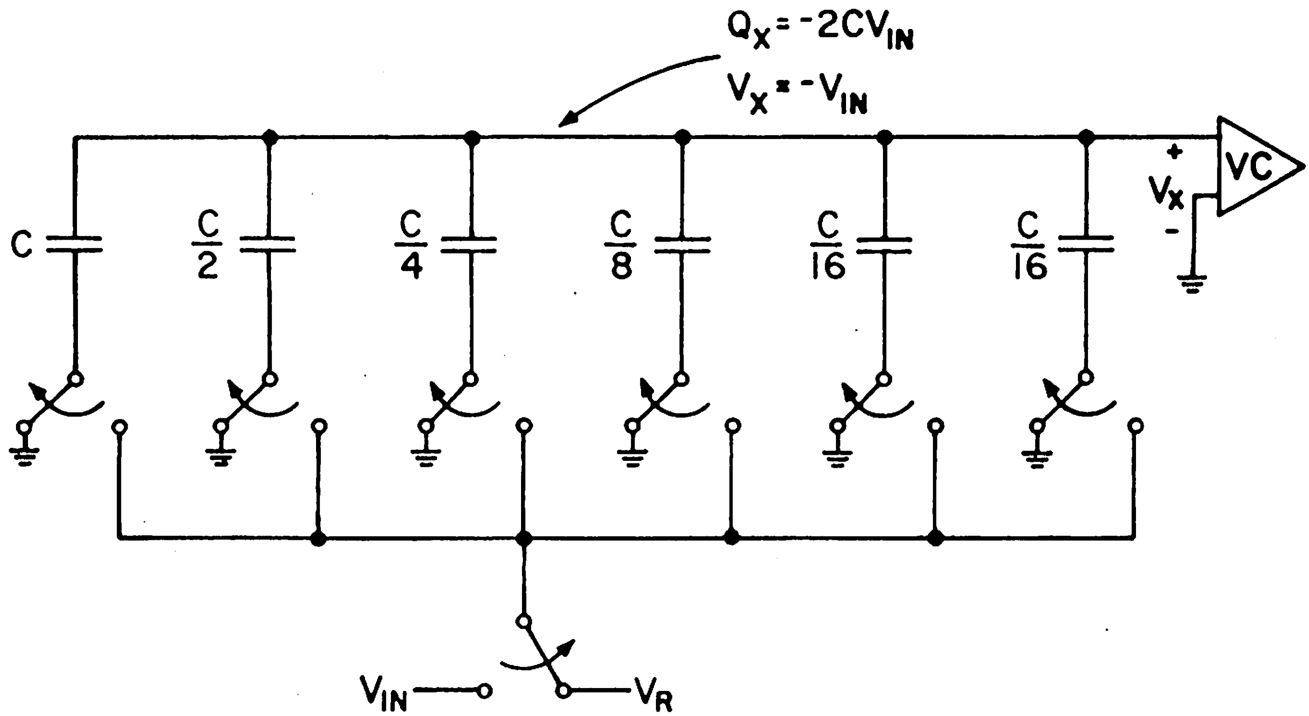
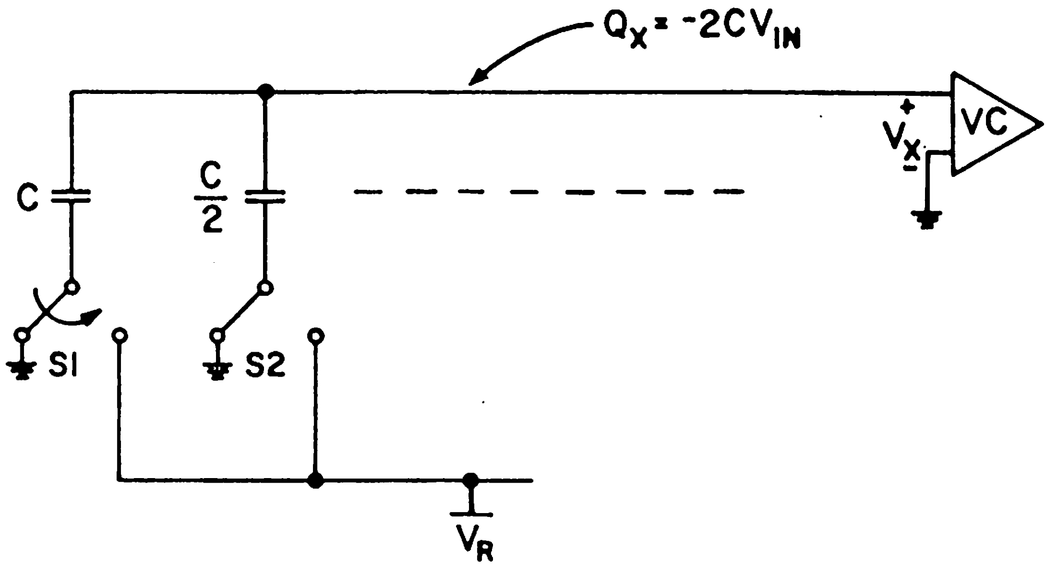


Figure 3.4: Pre-redistribution hold mode.



1. THROW S1 TO V_R $\rightarrow$ $V_X = -V_{IN} + \frac{V_R}{2}$

   DECIDE MSB, D_1

2. THROW S2 TO V_R $\qquad V_X = -V_{IN} + D_1 \frac{V_R}{2} + \frac{V_R}{4}$

   DECIDE D_2

   ⋮

Figure 3.5: Redistribution mode.

reference voltage V_R. The equivalent circuit is now actually a voltage divider between two equal capacitances. The voltage V_x which was equal to $- V_{IN}$ previously is now increased by $\frac{1}{2}$ the reference as a result of this operation.

$$V_x = - V_{IN} + \frac{V_R}{2}$$

The comparator senses the sign of V_x and its output is a logic '1' if $V_x < 0$ and is a '0' if $V_x > 0$. This is analogous to the interpretation that

$$\text{if } V_x < 0 \text{ then } V_{IN} > \frac{V_R}{2} \text{ hence the MSB} = 1$$

$$\text{but if } V_x > 0 \text{ then } V_{IN} < \frac{V_R}{2} \text{ therefore the MSB} = 0.$$

The output of the comparator is, therefore, the value of the binary bit being tested. Switch S1 is returned to ground only if the MSB, D_1 is a zero. In a similar manner, the next MSB is determined by raising the bottom plate of the next largest capacitor to V_R and checking the polarity of the resulting value of V_x. In this case however the voltage division property of the array causes $\frac{V_R}{4}$ to be added to V_x:

$$V_x = - V_{IN} + D_1 \frac{V_R}{2} + \frac{V_R}{4}$$

Conversion proceeds in this manner until all the bits have been determined. The final value of V_x is the error voltage V_E. A final configuration is illustrated in Figure 3.6 for the digital output 01001. Notice that all capacitors corresponding to a '0' bit are completely discharged. The total original charge on the top plates has been redistributed in a binary fashion and now resides only on the capacitors

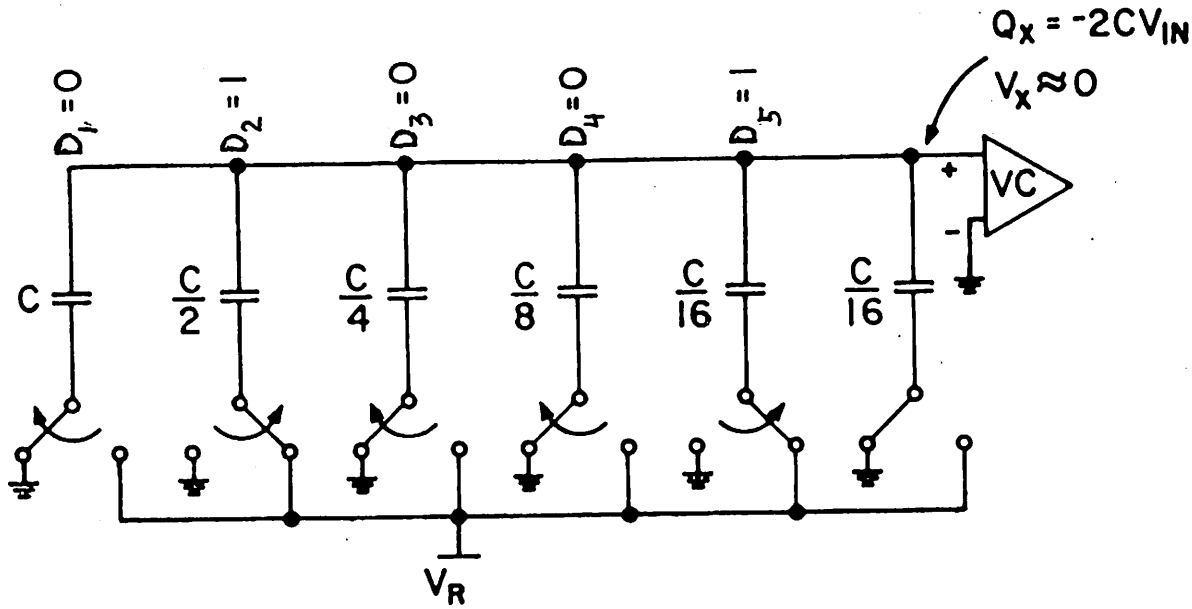


Figure 3.6: The final configuration (example).

corresponding to a '1' bit. N redistributions are required for a conversion resolution of N bits. In contrast to earlier charge-redistribution techniques the capacitance of the lower plate switch does not affect the accuracy of the conversion [16]. This fact is evident since the switch capacitance is either discharged to ground or is charged by V_R but never absorbs charge from the top plate. Therefore, the switch devices can be quite large permitting rapid redistributions. On the other hand the top plate of the array is connected to all the capacitors and to a switch and to the comparator resulting in a large parasitic capacitance from the top plate to ground. The nature of the conversion process, however, is such that V_x is converged back towards zero -- its initial value. Hence the charge on this parasitic is the same in the final configuration as it was in the sample mode. Therefore the error charge contributed by this parasitic is very near zero as will be further discussed in Chapter IV [17]. Because of this the smallest capacitor may be much smaller than the parasitic and consequently the largest capacitor may be reduced in value proportionally. Furthermore, the initial value of V_x need not necessarily be zero but can be the threshold voltage of the comparator. This fact allows cancellation of comparator offset by storing the offset in the array during the sample mode. The linearity then is primarily a function of the ratio accuracy of the capacitors in the array.

By only a slight modification of the array switching scheme bipolar voltage inputs can be encoded while still using only the single positive reference. This is achieved by connecting the bottom plate of the largest capacitor to V_R during the sample mode resulting in a stored charge:

$$Q_x = - C_{TOT} \left(\frac{V_{IN}}{2} + \frac{V_R}{2} \right)$$

Each bit is then tested in sequence just as before except that the largest capacitor is switched from V_R to ground during its test, while all the other capacitors are switched from ground to V_R. Also as before a bit value is true if V_x is negative after the test. The expression for V_x again converges back towards zero:

$$V_x = -\frac{V_{IN}}{2} + V_R \left(-\frac{D_1}{2^1} + \frac{D_2}{2^2} + \frac{D_3}{2^3} + \cdots + \frac{D_{10}}{2^{10}}\right) \approx 0$$

D_1 is '0' for $0 \leq V_{IN} \leq 10V$, but is '1' for $-10V \leq V_{IN} \leq 0$. Therefore D_1 represents the sign bit and its function is to level shift V_x in order to accommodate negative inputs. Hence a 10-bit conversion is achieved over the input range $\pm 10V$ with negative numbers expressed in 1's complement.

3.5 Summary

The fundamental basis for a successive approximation algorithm using charge-redistribution between binary weighted capacitors was examined in this chapter. The development of a circuit which implements this algorithm and the subsequent verification of the RADCAP method is the subject of later chapters.

CHAPTER IV

Factors Limiting Accuracy in RADCAP Type of Circuits

4.1 Introduction

In this chapter the factors limiting accuracy in RADCAP (MOS, successive approximation, charge-Redistribution, ADC utilizing weighted CAPacitors) type of circuits will be examined. The qualitative and quantitative compilation of these effects is essential in order to assess the advantages and disadvantages of various forms of capacitor structures and capacitor networks. The analysis of these effects led to a compatible MOS capacitor design geometry which is optimized to maintain ratio accuracy when conventional photolithography is used. In section 4.2 solutions will be proposed for the problem of input offset voltage cancellation for an MOS comparator. The effects of parasitic capacitance from the capacitor plates to ground will be investigated in section 4.3. Next the significance of temperature and voltage coefficients of capacitance and dielectric relaxation will be evaluated and an estimate will be made of how these affect the ability to design precision capacitors. Effects such as current leakage and parameter drift must also be considered. In addition an assessment is also made for the effects of several other factors which can cause ratio errors, some of which are capacitor oxide gradient and undercutting of the mask which defines the capacitors. Finally a description of the intrinsic offset voltage in RADCAP and its cancellation conclude this chapter.

4.2 MOS Comparator Input Offset Voltage Cancellation

The voltage comparison process is fundamental to A/D conversion. The offset voltage of the comparator is usually manifested as an offset error in the digital conversion. Because of the relatively large gate-source voltage mismatch in MOS differential amplifiers, the offset voltage of the all-MOS comparator needed for RADCAP must be eliminated as a source of error [18]. This can be accomplished either by digital means or by offset cancellation techniques. The problem is investigated with an ideal comparator shown in Figure 4.1. In this illustration the switching threshold voltage is zero and the gain at the threshold voltage is:

$$|A_c| = \frac{\Delta V_{out}}{\Delta V_{IN}} = \infty.$$

However, a real comparator (of any technology) may be modelled to a first order approximation by the transfer function shown in Figure 4.2. In this illustration the comparator has both input offset voltage VIOS and output d.c. voltage VOOS. VOOS is not actually an offset error but rather is a d.c. switching level intermediate between the two logic levels. Hysteresis has been neglected and finite gain is modeled by the fact that $|A_c| < \infty$. Both d.c. voltages may tend to be large for an MOS comparator; however a large value of VIOS causes a significant translation in d.c. voltage bias at the input. This tends to destroy the usefulness of the comparator for small signal inputs. Therefore the input offset voltage of an MOS comparator must be cancelled.

One particular method of performing offset voltage cancellation is illustrated in Figure 4.3. In this circuit the capacitor is precharged to the switching threshold voltage of the comparator. This is identical to the input offset if VIOS = VOOS as in the case of logic circuits. If the comparator gain is large in magnitude and negative then the momentary closure of S1 and S2 as shown during an initial precharge cycle forms a feedback path which stabilizes when $V_x \approx$ VIOS. When both switches are

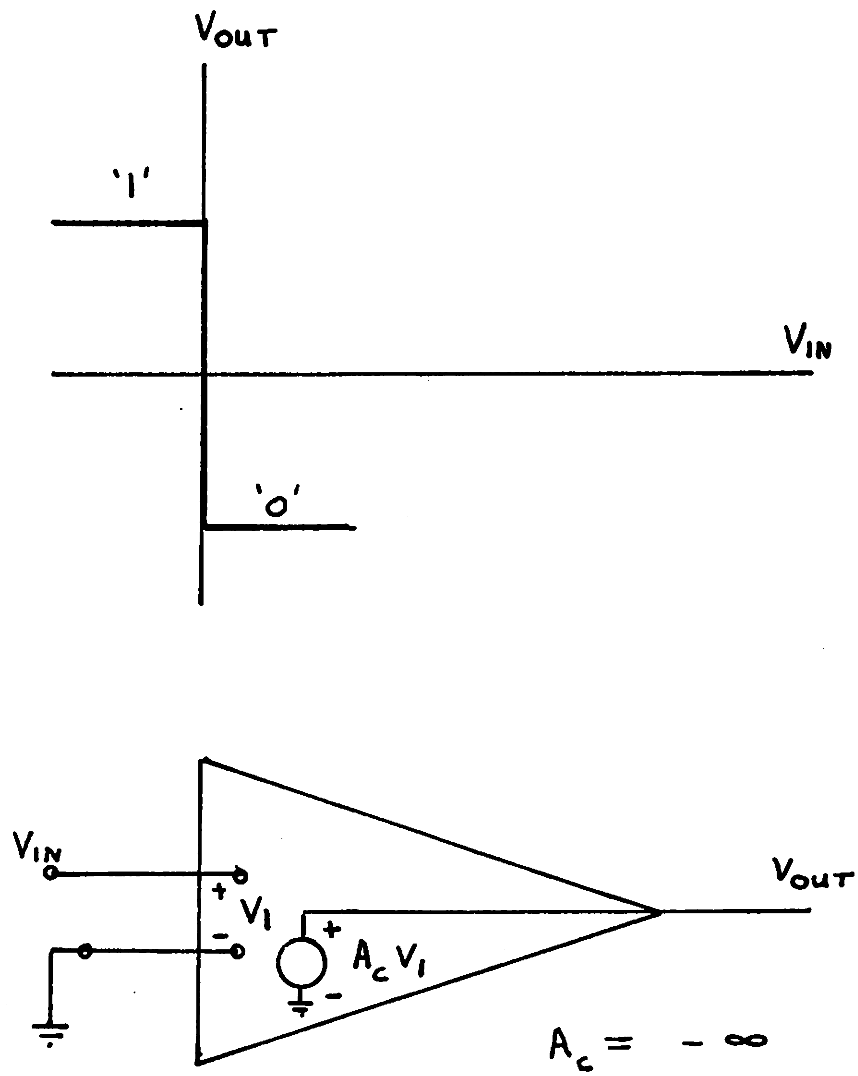


Figure 4.1: An ideal comparator.

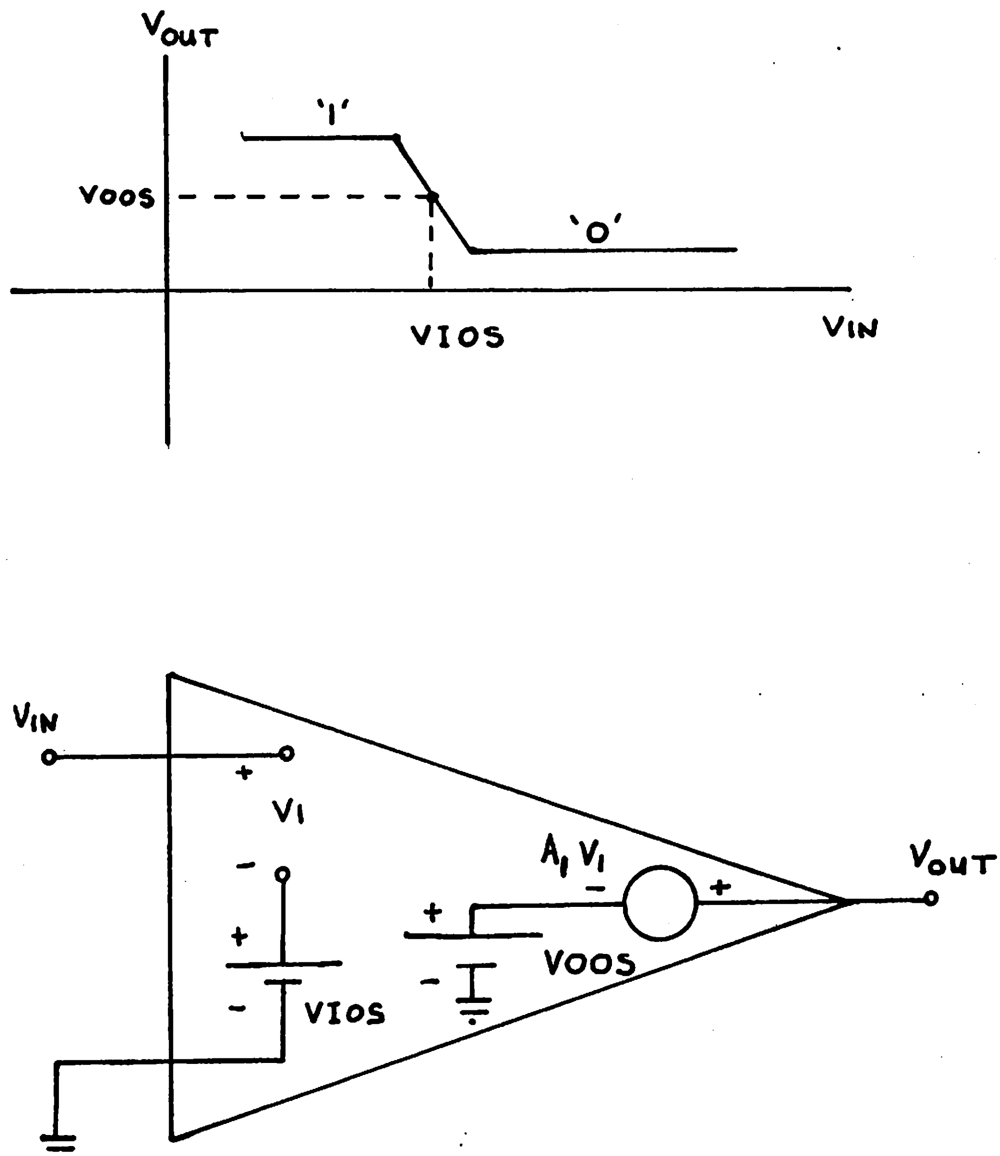


Figure 4.2: A real comparator.

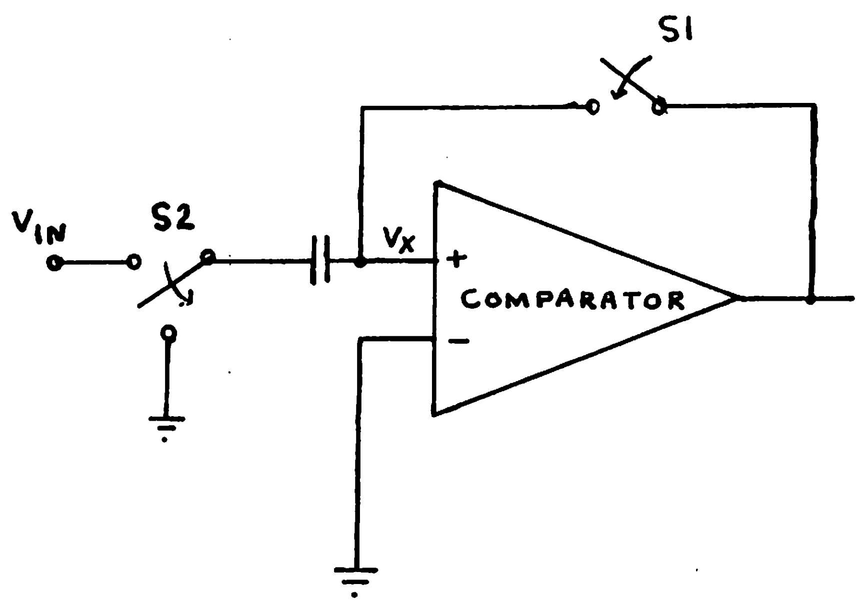


Figure 4.3: Offset cancellation by storing the offset in the capacitor.

opened the capacitor retains its d.c. charge and the effective offset referenced to V_{IN} is zero. If amplifier gain is low several stages are required in order to realize a high gain comparator. The example, shown in Figure 4.4 illustrates another technique to reduce the effects of input offset voltage. Two gain stages are used but offset cancellation by capacitor storage is performed only for stage A1. Hence there is no offset due to A1. However stage A2 has offset VIOS2, but when this is reflected back to the input V_{IN} through A1 the effective offset is:

$$\text{VINOS(EFFECTIVE)} = \frac{\text{VIOS2-VOOS1}}{\text{A1}}$$

If the gain magnitude A1 is large VINOS(EFFECTIVE) is small. Hence this effective value may be reduced to an acceptable value. Both of the offset cancellation methods just discussed are used for realization of the RADCAP technique.

4.3 Effects of Parasitic Capacitance from the Capacitor Plates to Ground

In RADCAP and VATCAP circuit techniques parasitic capacitance to ground exists at the top and bottom plates of the MOS capacitors. For either class of circuits the parasitic capacitance to ground from the bottom capacitor plates does not affect the accuracy of the charge-redistribution at the top plate. This is true because these parasitics are charged by V_R and discharged to ground by an MOS device which has no offset voltage when used as a charge switch. Therefore bottom plate parasitic capacitance does not participate in charge-redistribution at the top plate.

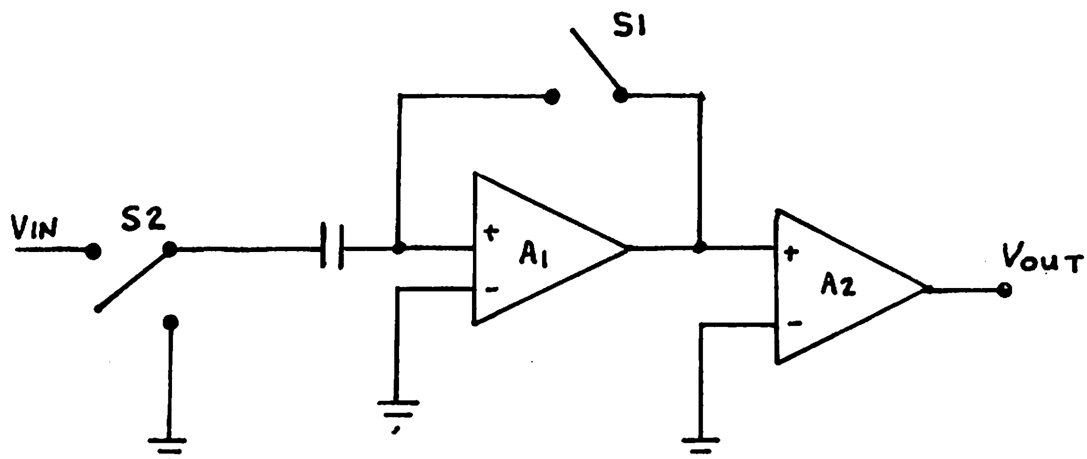


Figure 4.4: Reduction of input offset voltage by capacitive storage and by reflection through a high gain stage.

In VATCAP, however, the parasitic capacitance C_p to ground from the top plate of the capacitor network results in a charge-redistribution accuracy that is dependent upon the ratio of C_p to the total capacitance in the array C_T. The voltage signal V_x at the top plate in the VATCAP method is desired to be that of a precision attenuator or DAC:

$$V_x = B_i \frac{V_R}{2^N} .$$

However, if $C_p \neq 0$ as illustrated in Figure 4.5 for a simplified version of VATCAP then

$$V_x = B_i \frac{V_R}{2^N} \left(\frac{1}{1 + \frac{C_p}{C_T}} \right) \quad \text{, in which } C_p << C_T .$$

A gain error exists although the linearity and offset are unaffected provided that C_p is not heavily voltage dependent. One method of correcting this error is to increase the reference voltage to a new value V_R' such that:

$$V_R = \left(\frac{V_R'}{1 + \frac{C_p}{C_T}} \right) .$$

This procedure requires a stable external adjustable reference voltage.

In contrast, RADCAP incorporates VATCAP in such a way that C_p is not important as was asserted in section 3.4. This will now be demonstrated with the aid of Figure 4.6. As explained in Chapter III the successive approximation algorithm used in RADCAP will force V_x, the voltage at the top plate, to converge to zero plus or minus ϵ_q in the final configuration. For the ideal converter, having an ideal comparator as modelled in Figure 4.1, $V_x = B_i \frac{V_R}{2^N} - V_{IN}$ and $- \epsilon_q \leq V_x \leq \epsilon_q$. This was derived in Chapter II.

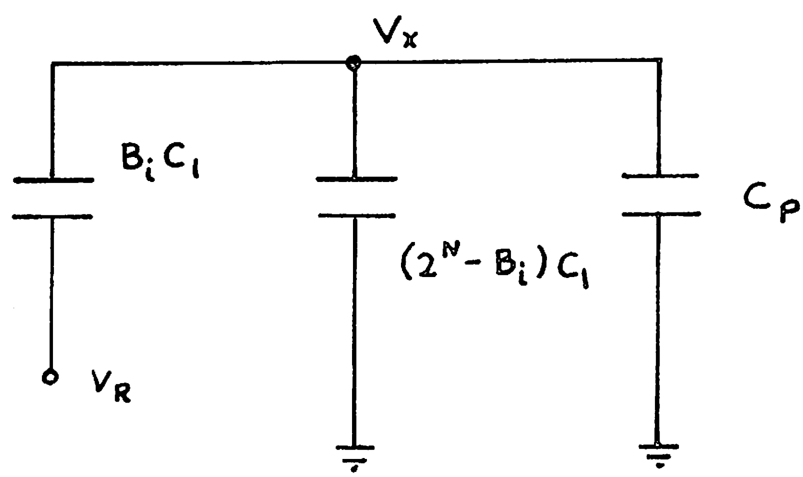


Figure 4.5: An illustration of top plate parasitic capacitance C_p at node X in the RADCAP circuit technique. $B_i C_1$ represents the equivalent parallel connection of all capacitors with bottom plates connected to V_R.

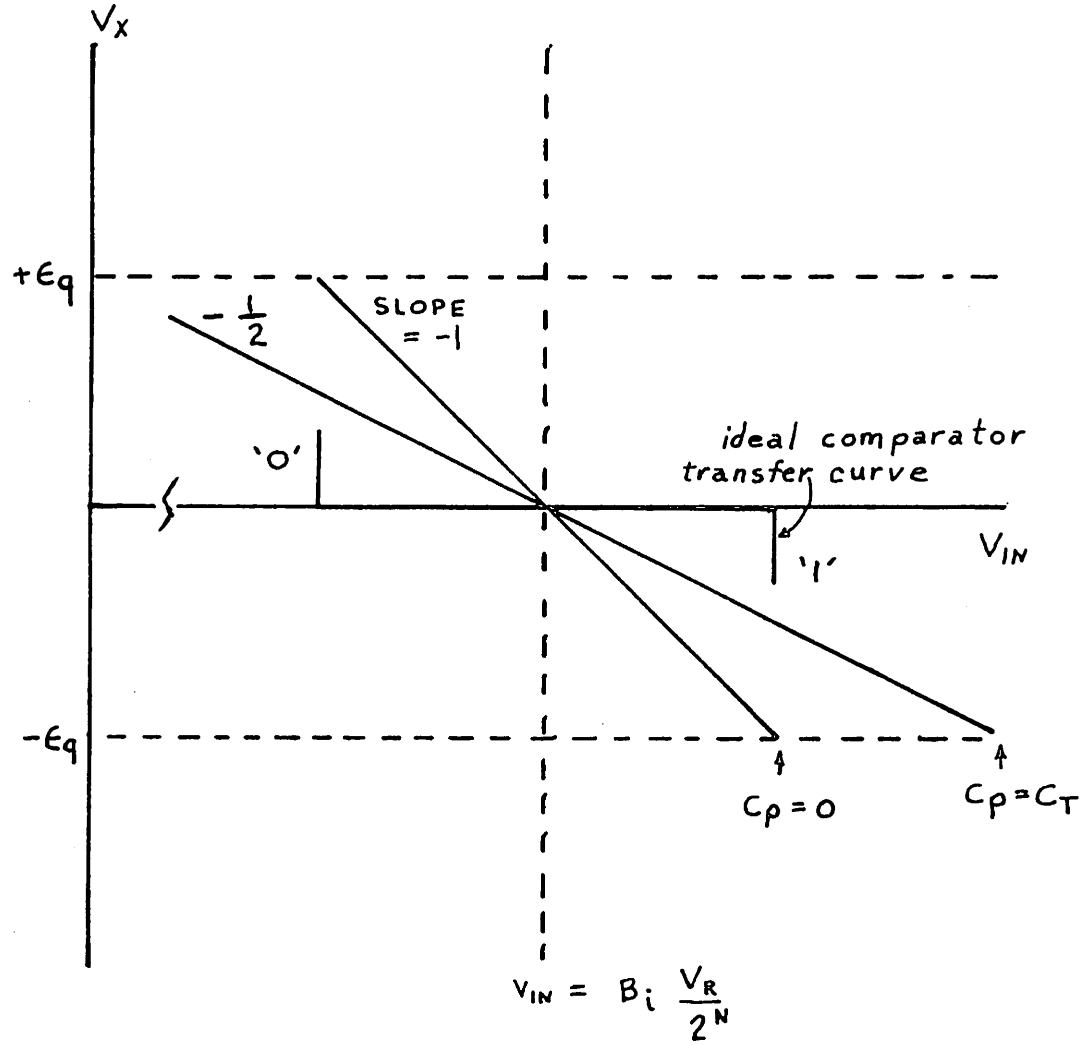


Figure 4.6: An illustration of the effect of C_p in RADCAP for an ideal comparator.

The expression for the total captured charge Q at x is:

$$Q = -\ 2^N C_1\ V_{IN} = V_x(2^N C_1 + C_p) - V_R\ B_i\ C_1 .$$

From this equation and $C_T = 2^N C_1$:

$$V_x = (B_i \frac{V_R}{2^N} - V_{IN}) \frac{1}{1 + \frac{C_p}{C_T}} .$$

Here V_x represents an effective error voltage. The influence of C_p may now be determined with the aid of Figure 4.6. This is the plot of V_x vs V_{IN} for a small region around $V_{IN} = B_i \frac{V_R}{2^N}$. Two lines are plotted which show the behavior of the transfer curve due to C_p. One line for which $C_p = 0$ has gain $\frac{\Delta V_x}{\Delta V_{IN}} = -\ 1$, and the other for the unrealisticly pessimistic case that $C_p = C_T$. In this latter case the gain has been reduced to $-\frac{1}{2}$. The ideal comparator transfer curve is also superimposed on the V_{IN} axis at the point $(B_i \frac{V_R}{2^N}, 0)$. The intersection of the comparator transfer curve and the transfer function for V_x is the point at which the ideal comparator switches. From observation this point is independent of C_p, hence C_p has no effect upon RADCAP if the comparator is ideal.

A different situation exists however for a real comparator having an uncancelled input offset voltage $-\ V_{OS}'$ and a finite gain representing an uncertainty Δu. This is illustrated in Figure 4.7 for $C_p = 0$ and $C_p = C_T$. For $C_p = 0$ the offset reflected to the V_{IN} axis is V_{OS}' and the input uncertainty range is Δu about that offset. However the comparator transfer curve is then translated to the intersection of the lines for which $C_p = C_T$ and $V_x = -\ V_{OS}'$. This models the unrealistic case in which parasitic capacitance is equal to the total capacitance in the array. From observa-

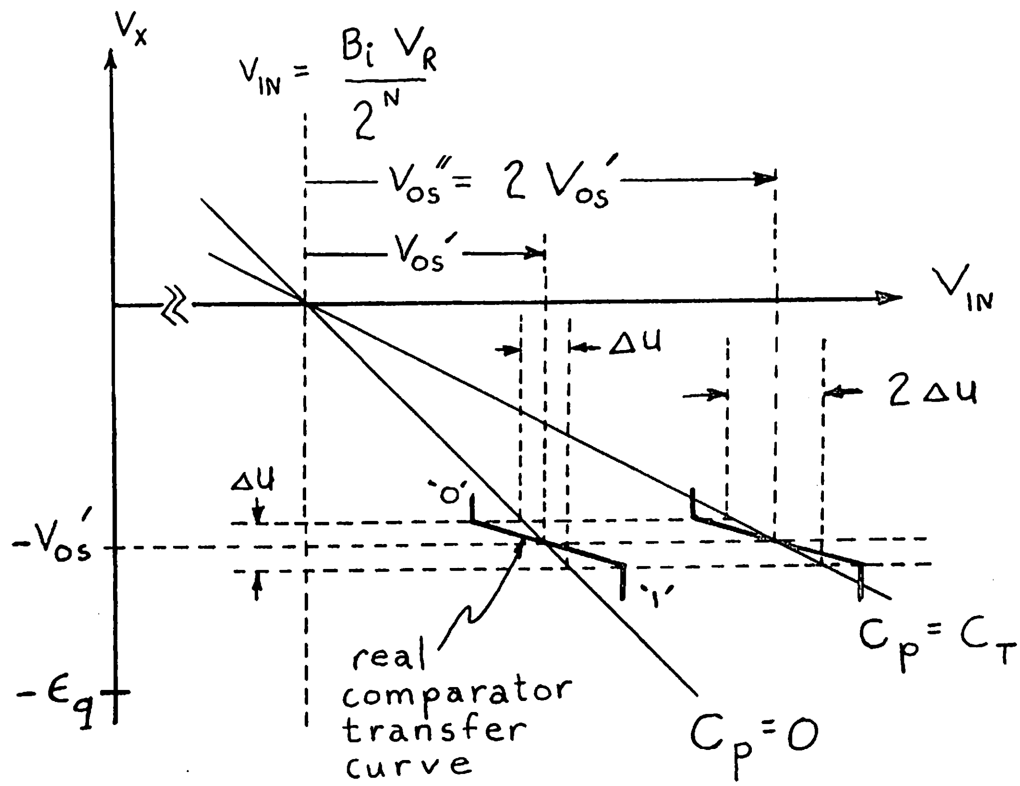


Figure 4.7: An illustration of the effect of C_p in RADCAP for a real comparator.

tion V_{OS}'' has increased to 2 V_{OS}' and the new uncertainty range is 2 Δu. Both changes represent a finite but negligible performance degredation for the RADCAP system. These same results could have been developed by considering VATCAP to have an effective gain of

$$A_{VAT} = - \left(\frac{1}{1+ \frac{C_p}{C_T}} \right)$$

which is the slope of the transfer function of VATCAP: $\frac{\Delta V_x}{\Delta V_{IN}}$. Now offset reflection techniques may be applied to the system as modelled in Figure 4.8. Hence, a comparator offset $- V_{OS}'$, when reflected to V_{IN} through VATCAP results in an effective input offset voltage:

$$\text{VINOS(effective)} = - \frac{V_{OS}'}{A_{VAT}} = V_{OS}' \left(1 + \frac{C_p}{C_T}\right).$$

A similar relationship exists for the uncertainty range when reflected to V_{IN}. The quantitative effect of C_p may now be determined. For RADCAP techniques it is estimated that $\frac{C_p}{C_T} \leq .05$ with normal layout guidelines, therefore it may be deduced that C_p has virtually no effect upon the RADCAP technique since V_{OS}' can be made small by offset cancellation methods. In addition the input uncertainty range can be reduced to a sufficiently low level as will be discussed later in Chapter VI.

4.4 Temperature Coefficient of Capacitance

A great advantage of a SiO_2 dielectric capacitor is its very low temperature coefficient of capacitance (TCC) of 24.8 ppm/°C [19]. Capacitance values over the range - 10°C to 140°C for both SiO_2 and $p^+ n^+$ junction capacitors are shown in Figures 4.9 and 4.10 respectively. The

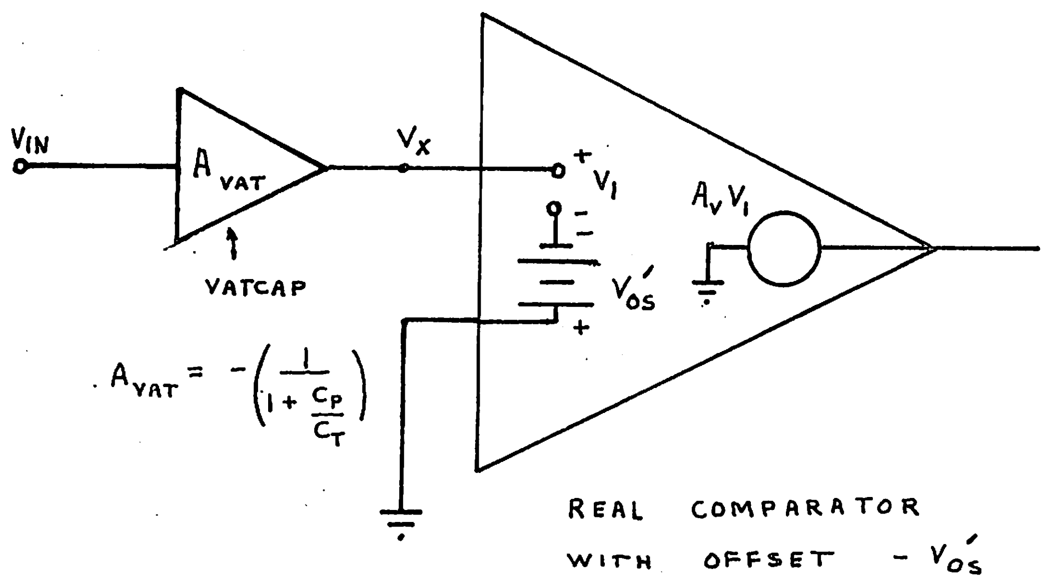


Figure 4.8: Analysis of the effect of C_p by reflection to the input, V_{IN}.

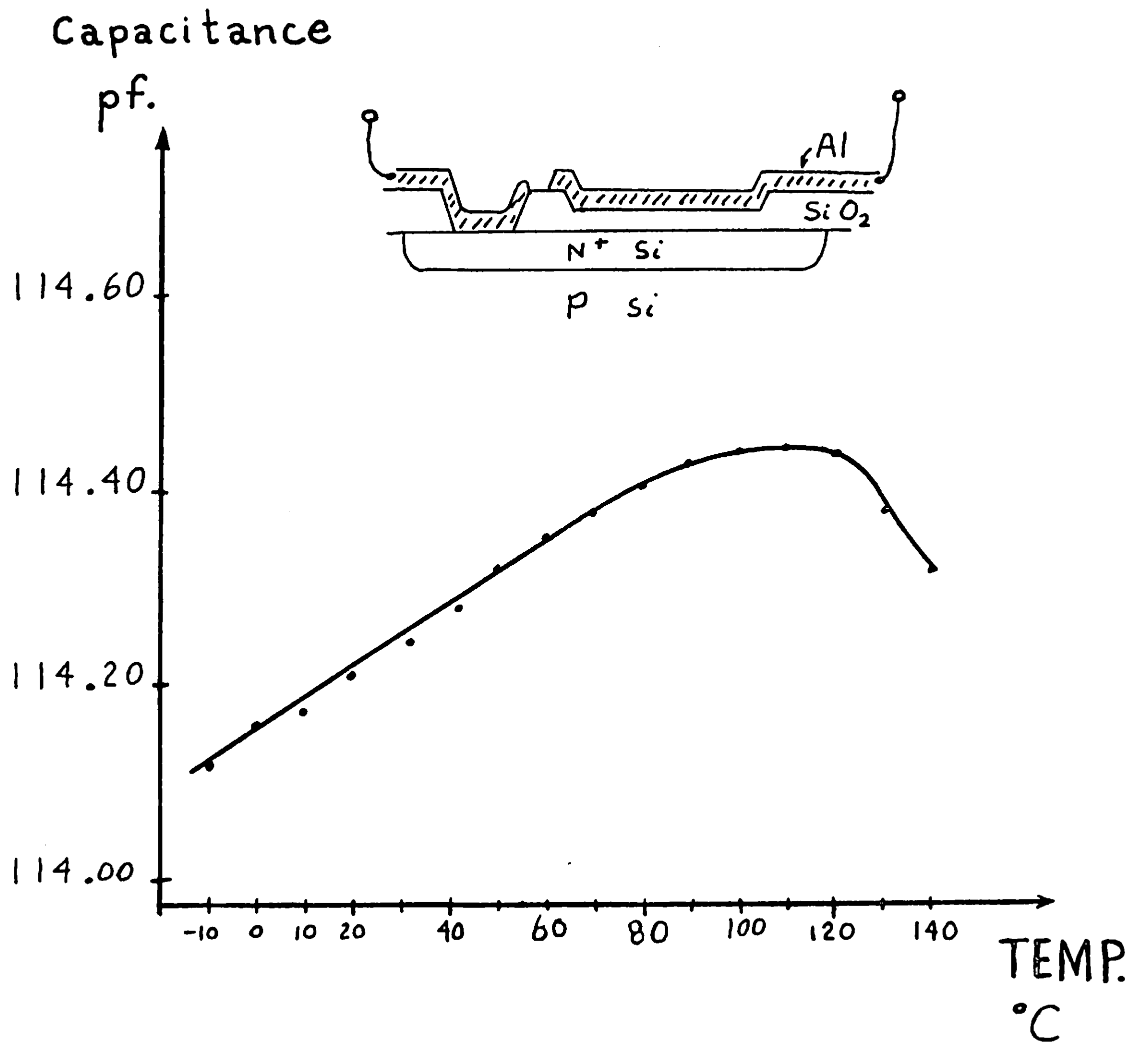


Figure 4.9: A plot of capacitance as a function of temperature for a SiO_2 capacitor.

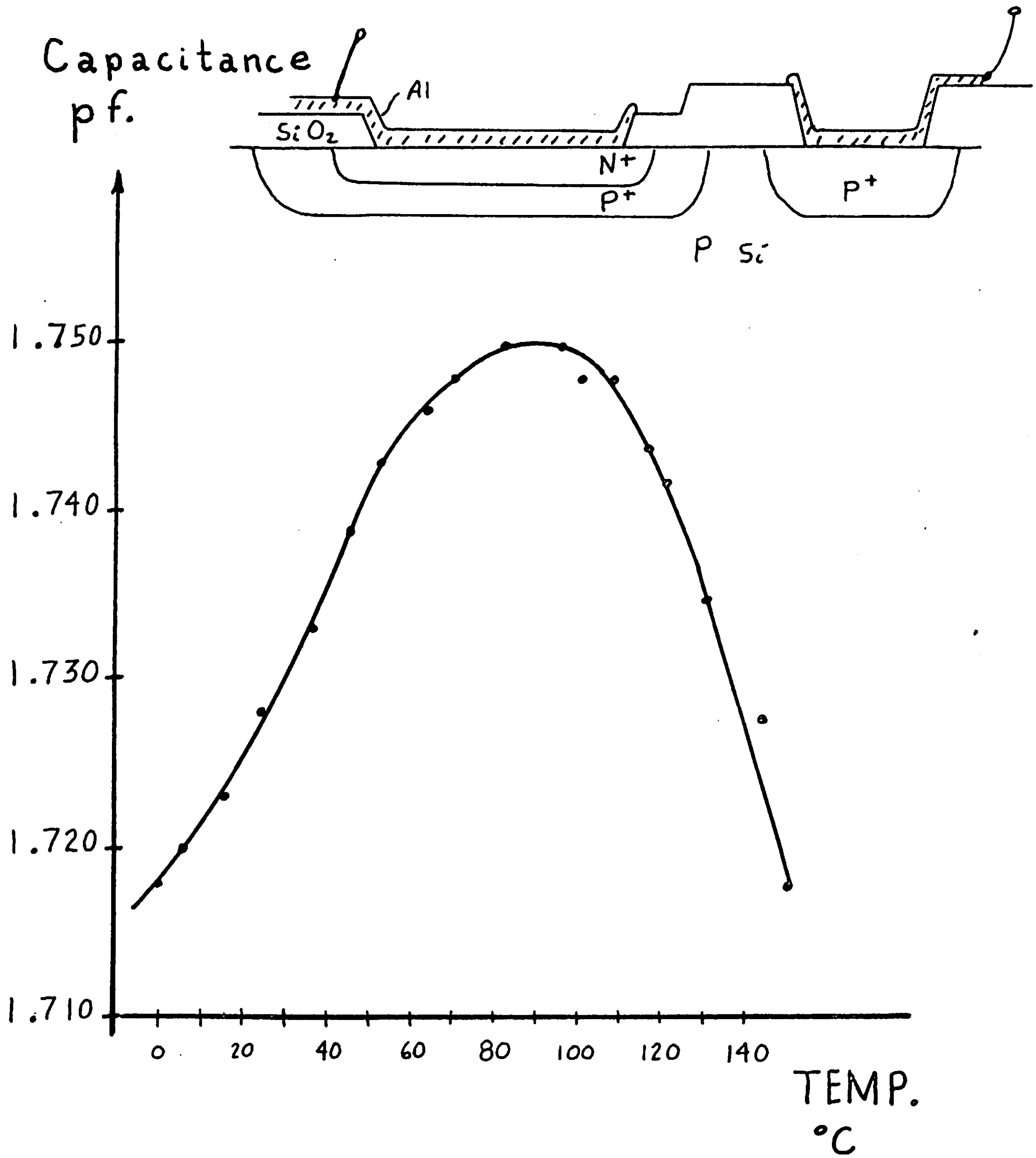


Figure 4.10: A plot of capacitance as a function of temperature for an N^+P^+ junction under reverse-bias conditions. Refer to Appendix C for details relating to diffusion times and temperatures.

TCC of the SiO_2 capacitor is contrasted with the temperature coefficient of resistance TCR in the range of 1000 to 2000 ppm/°C for diffused resistors and several hundred ppm/°C for thin-film and ion-implanted resistors [20]. This data is summarized in Table 4.1. An additional factor influencing circuit dependence upon temperature is that resistive networks used in high speed conversion techniques usually suffer from

Component	Typical Temperature Coefficient in ppm/°C
SiO_2 capacitor	+ 24.8
$p^+ n^+$ junction capacitor	+ 230
diffused resistor	+ 1500 [21]
thin-film resistor	- 200
ion-implanted resistor	+ 400

Table 4.1 Comparison of Component Temperature Coefficients over the military temperature range.

localized thermal gradients caused by the switching of large currents. Although some measures can be taken to reduce these effects, localized thermal gradients will still be present and will induce component mismatch errors. This situation is not present in RADCAP circuit methods since no large d.c. currents flow in the capacitors, hence no thermal gradients will exist due to the capacitor array. An additional benefit of charge-redistribution is therefore much lower power consumption. In addition the differential temperature coefficient (the variation in temperature coefficient between components) is expected to be less for SiO_2 capacitors than for resistors because TCC is an order of magnitude less than TCR. From these considerations it may be concluded that SiO_2 capacitor matching is less dependent upon local and environmental temperature variations than resistor matching [22]. Since component mismatch has a great effect upon linearity, the temperature coefficient of nonlinearity due to mismatching is low for the RADCAP technique in comparison with resistive approaches.

4.5 Voltage Coefficient of Capacitance

The value of a SiO_2 capacitor having a metal top plate and a heavily doped N^+ bottom plate is dependent upon d.c. terminal voltage [23]. This is due to the existence of accumulation, depletion, or inversion layers which may be formed at the N^+ surface. When the terminal voltage is positive as illustrated in Figure 4.11 the N type surface becomes accumulated with mobile electrons and acts as a low resistance. The capacitance in this case is almost entirely due to the SiO_2. However, as the terminal voltage V_c becomes negative the N type surface begins to become depleted and a very thin, high capacitance, space charge layer is added in series with the SiO_2 capacitor. The depletion capacitance becomes smaller as the reverse voltage increases and the depletion width widens. Therefore the total series capacitance decreases. At some value of reverse voltage V_c the surface becomes inverted and the depletion region width reaches its maximum value. This is illustrated by a minimum capacitance for some reverse bias in Figure 4.11. At a larger reverse bias voltage (more negative value of V_c) the surface becomes inverted and the total capacitance approaches the same value as that during accumulation. The effect of surface doping concentration of the N type material is also illustrated in Figure 4.11. It may be seen that increased doping reduces the fractional rate of change of capacitance with voltage. The capacitor voltage coefficient α is defined by the equation $\alpha = \frac{dC(v)}{dv}$.

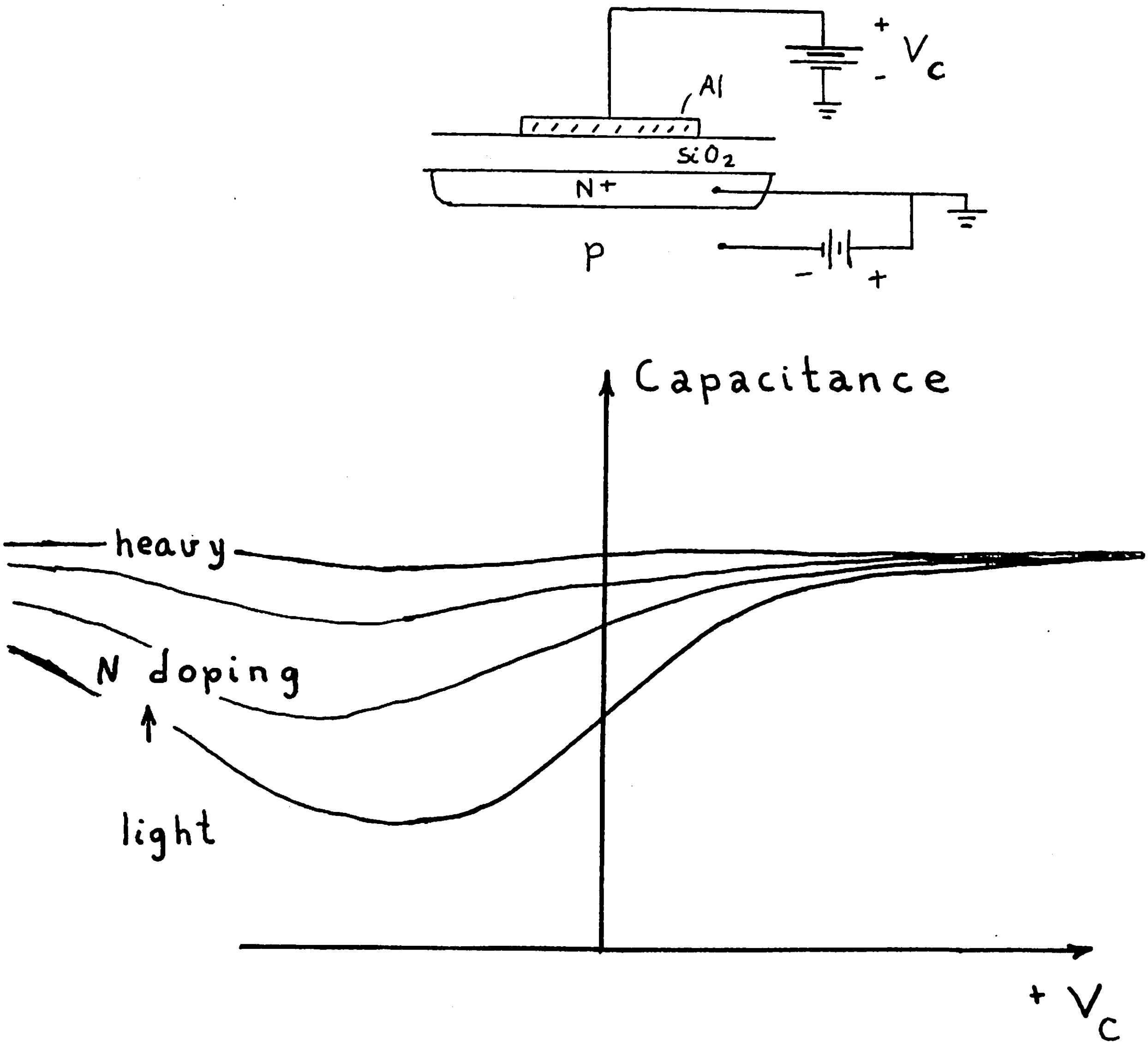


Figure 4.11: The voltage dependence of MOS capacitance.

A lower α can be obtained by a higher surface concentration. For this reason N^+ silicon is preferred as the capacitor back-plate since higher surface concentrations may be attained for N type impurities than P type impurities. The plot of capacitance versus voltage is shown in Figure 4.12. From this plot the linearly extrapolated α is 21.9 ppm/volt over the region - 10V to + 10V.

The effect of this value of α will now be illustrated with the aid of Figure 4.13. Any single charge-redistribution in the array may be modeled by a series combination of 2 capacitors C1 and C2. With the structures indicated:

$$C1(V1) = B_i C_0 (1 - \alpha\ V1)$$

and

$$C2(V2) = (2^N - B_i) C_0 (1 + \alpha\ V2).$$

These equations are solved simultaneously in Appendix A from which the worst case error in V2 is $\varepsilon = -\alpha \frac{V_R^2}{8}$ which occurs for $V2 = \frac{V_R}{2}$. The normalized error distribution as a function of V2 is plotted in Figure 4.14. In this graph, for the case that V_R = 10 volts, the error in mV is $-.3\alpha'$ where $\alpha' = \frac{\alpha}{22\ \text{ppm/V}}$ a normalization factor. Since the error is a function of V2 which is proportional to B_i and to V_{IN}, this error is manifested as a nonlinearity error. For the 10-bit experimental RADCAP circuit this corresponds to a worst case nonlinearity of +.03 LSB due to voltage coefficient, and this is an insignificant value. It is expected that a modified fabrication schedule that would increase the N^+ surface concentration would reduce α by a factor of 3 [24] [25].

4.6 Dielectric Relaxation

When a capacitor array is used as a precision voltage attenuator as in VATCAP, errors in voltage ratios can occur due to

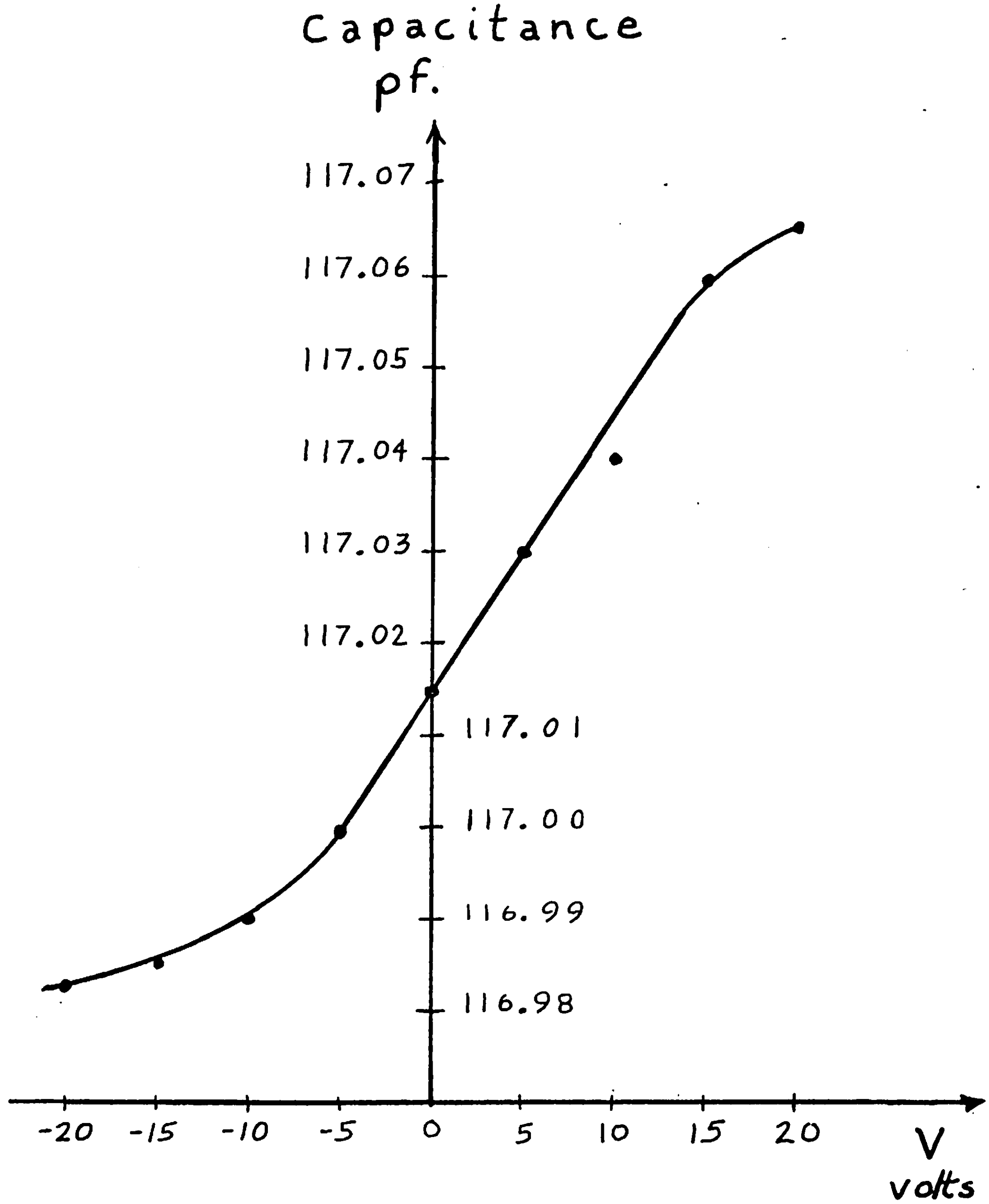


Figure 4.12: The measured voltage dependence of the largest capacitor in the array of an experimental I.C.

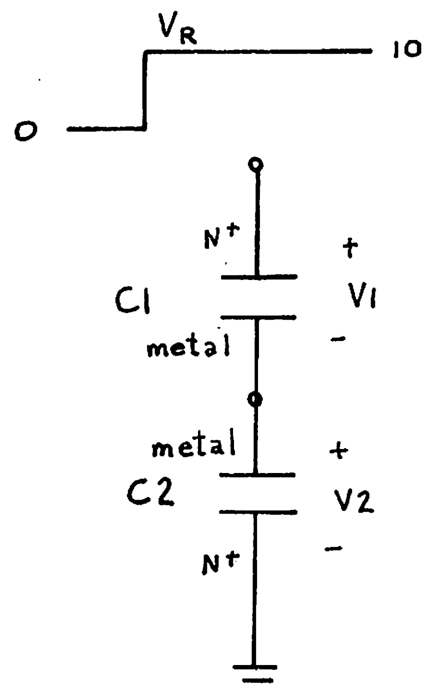


Figure 4.13: An equivalent circuit for the capacitor array.

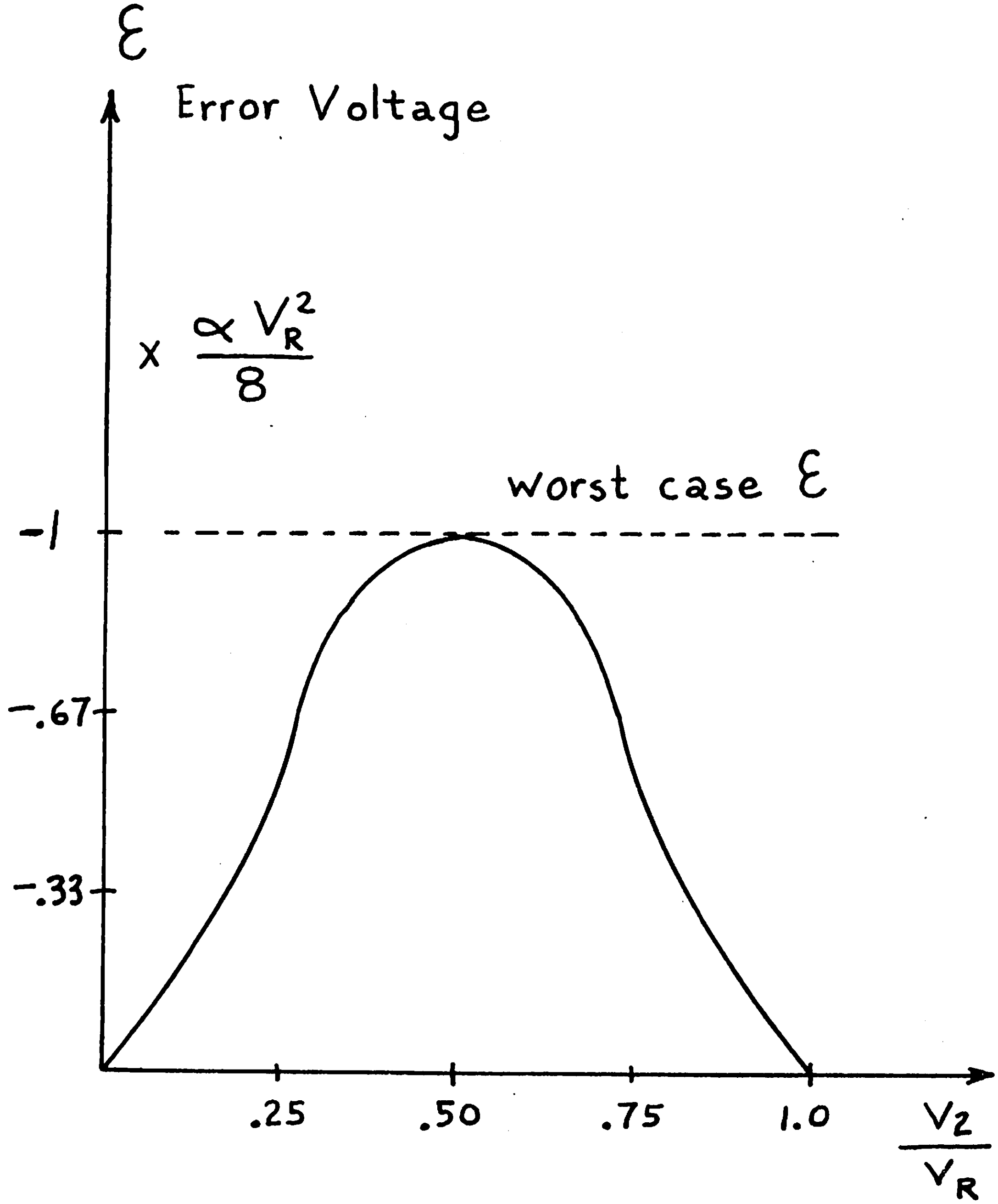


Figure 4.14: The normalized error due to capacitor voltage coefficient α. The vertical scale normalization factor α' is defined as $\frac{\alpha}{\alpha_0}$ where α_0 = 22 ppm/volt.

dielectric relaxation. Consider the two-capacitor circuit of Figure 4.15 having dielectric losses modeled by parallel resistances. In the ideal case R1 = R2 = ∞ and

$$V_x = \frac{C2}{C1 + C2} V_R$$

in steady state. But for a real insulator having finite resistivity, the final voltage would be $V_x = \frac{R1}{R1 + R2} V_R$ which could result in errors. In general, the maximum observation time (or conversion time) for a capacitor network is related to the dielectric relaxation time τ_R. For SiO_2 dielectric capacitors of area A and oxide thickness t:

$$\tau_R = RC = \frac{t}{\sigma SiO_2 \; A} \frac{\epsilon_{ox} A}{t} = \frac{\epsilon_{ox}}{\sigma SiO_2}$$

where the conductivity of SiO_2, $\sigma SiO_2 < 10^{-16}$ Ω cm and τ_R = 3000 sec. From these calculations the observation time must be much less than 3000 sec.

The parallel resistance actually models only one dielectric defect, the steady state generation, drift and recombination of mobile carriers through the oxide due to the high electric field. There are other effects, however, which may cause more serious problems [26]. Polarized molecules in the dielectric may become aligned with the electric field resulting in a residual polarization after rapid discharge. This phenomena is generally not considered significant in SiO_2 for two reasons; its time constant is much shorter than other circuit delays and the density of polar molecules may be kept to a small value by proper circuit fabrication techniques. Of a more serious consequence, however, is the density of mobile ions in the

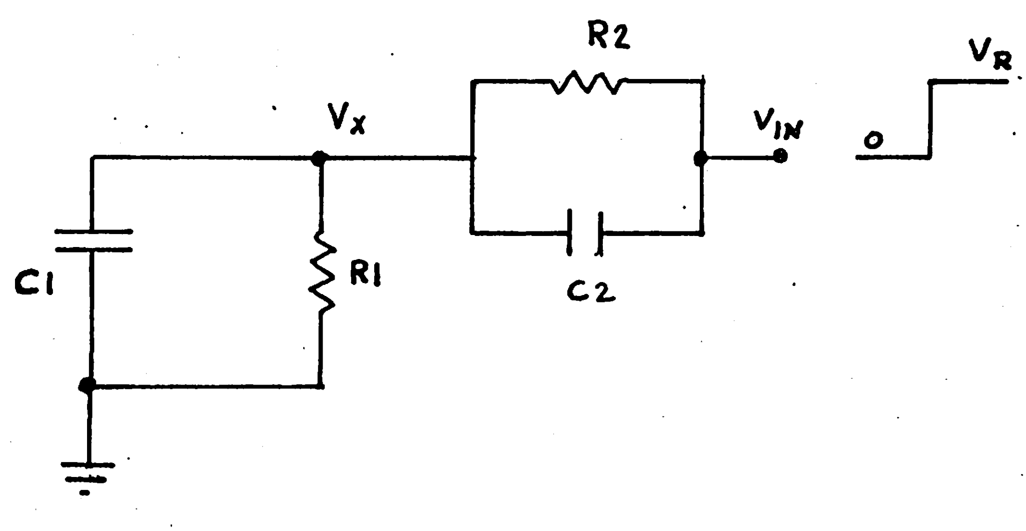


Figure 4.15: A 2-capacitor circuit modeling dielectric losses.

oxide or at its surface [27] [28]. These ions may drift in the electric field and accumulate at the oxide boundaries. This may result in a charge layer which dynamically augments the capacitance value. Furthermore if the concentration of mobile ions is very high, the conductivity of the oxide may appear to be large during the charging or discharging transient. Contamination of the oxide by alkali ions and water molecules enhances this dielectric relaxation phenomena by increasing the magnitude of the defects and reducing the relaxation time.

A typical dielectric relaxation effect is shown in Figure 4.16. These curves represent the voltage on the top plate $V_x(t)$ as the bottom plate of the largest capacitor is pulsed up to V_R for 50 μs then returned to ground. Figure 4.16(a) illustrates severe dielectric relaxation. This is manifested as an apparent residual voltage of opposite polarity remaining on the capacitor after it has been rapidly discharged. The magnitude of the defect was measured to be 10 mV, decaying to 5 mV after 50 μs. The 5 mV error had been previously detected though unexplained and subsequent investigation led to the discovery of the relaxation phenomena. This error alone created a nonlinearity of $\frac{1}{2}$ LSB. After some experimentation it was found that a mild heat treatment at 150°C to 200°C for 5 or 10 minutes reduced the magnitude of the effect beyond resolution capabilities as shown in Figure 4.16(b). In this case the relaxation effects had vanished within the surrounding system switching noise. These results tend to support the hypothesis that the main component of relaxation effect was associated with moisture trapped within the oxide or at its surface.

In conclusion, data from this study indicates that all contributions to dielectric relaxation combined do not cause significant error at the

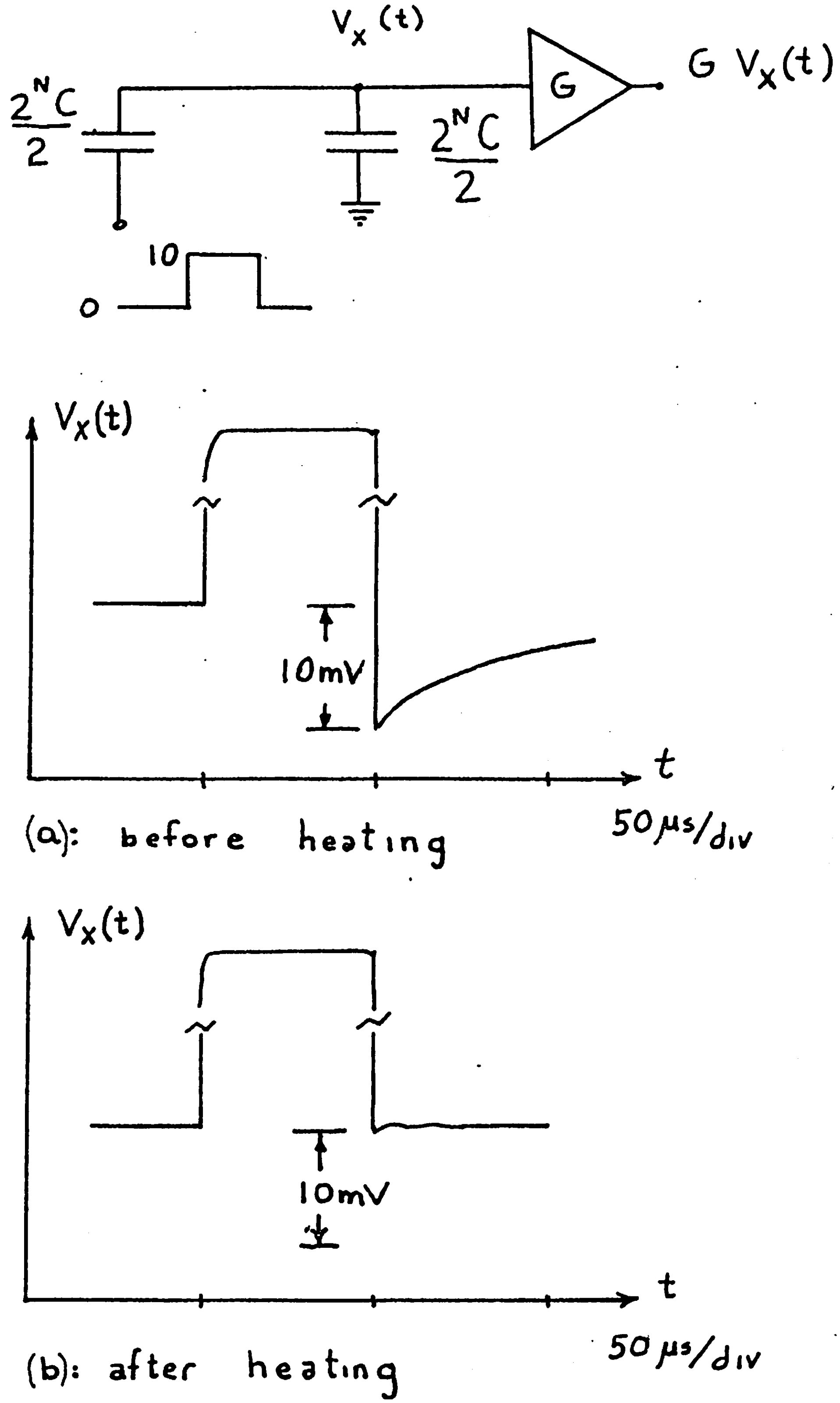


Figure 4.16: An example of a dielectric relaxation effect.

10-bit level.

4.7 Leakage Currents

A large leakage current from a reversed-biased pn junction connected to a capacitor plate can cause an error in capacitor voltage. This could result in an offset error in a circuit employing the VATCAP technique. This problem is especially acute in the case of high leakage junctions ($> 1nA/mil^2$) [29]. This is modeled in Figure 4.17. In the capacitor array technique the conversion time must be less than the time required for $\frac{1}{2}$ LSB error due to leakage. By estimating normal leakage to be $50(2)^{(\frac{T-25}{10})}$ pA/mil^2 as a function of Centigrade temperature T, then for the case of a 10 mil^2 junction at 75°C the maximum observation time for a voltage loss of less than $\frac{V_R}{2^{N+1}}$ is:

$$t = \frac{2^N C_1}{.5nA(2^5)} \frac{V_R}{2^{N+1}} = \frac{C_1}{32\ nA} V_R = 63\ \mu s,\ \text{for}$$

C_1 and V_R arbitrarily picked to be .2pF and 10 V. In this example the conversion must be completed within 63 µs or else greater than 1/2 LSB offset error will result. Since the nominal conversion time is about 20 µs, for RADCAP leakage effects will not be significant.

4.8 Parameter Drift

Parameter drifts in I.C. devices may result due to environmental temperature changes [30]. In most circuit designs this is not a serious problem if the circuit performance is dependent upon device ratios rather than upon absolute values. On the other hand large d.c. currents in an I.C. can cause localized thermal gradients. In conventional resistor network converters a symmetrical layout is usually required

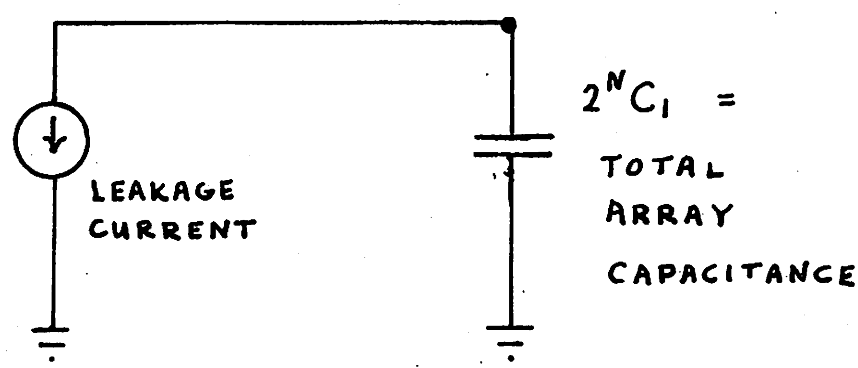


Figure 4.17: A circuit modeling leakage current from the top plate of the array.

to reduce the effects of these gradients. In contrast, there are no d.c. currents flowing in the capacitor network in RADCAP circuit methods hence no thermal gradients are caused by the capacitor array. However power dissipation in the logic circuitry may still cause thermal gradients unless proper design methods are used.

Long term parameter drifts are a characteristic problem with thin-film networks which are used in some converters [31]. Some improvement in stability is usually achieved at the cost of additional passivation layers during fabrication.

4.9 CAPACITOR RATIO ERRORS IN RADCAP

4.9.1 Capacitor Matching versus Resistor Matching

While monolithic circuit technology has had great impact on the cost of many analog circuit functions, such as operational amplifiers, the impact on the cost of A/D and D/A converters has not been as great. This is due to the complexity of a complete converter, and more importantly, to the problem of component matching. Because of the difference between the aspect ratios of diffused resistors of practical value versus those of capacitors, the attainable matching accuracy is higher for capacitors given the same overall die area [32]. The flexibility of capacitor geometry allows them to be made square or even circular so as to optimize matching accuracy. This is illustrated in Figure 4.18 by the fact that the resistors are usually long and thin. The resistor and capacitor are normalized for comparison by the constraint that they both have the same area 16 w^2 where w is a small dimension. The aspect ratio 16:1 would be the worst case

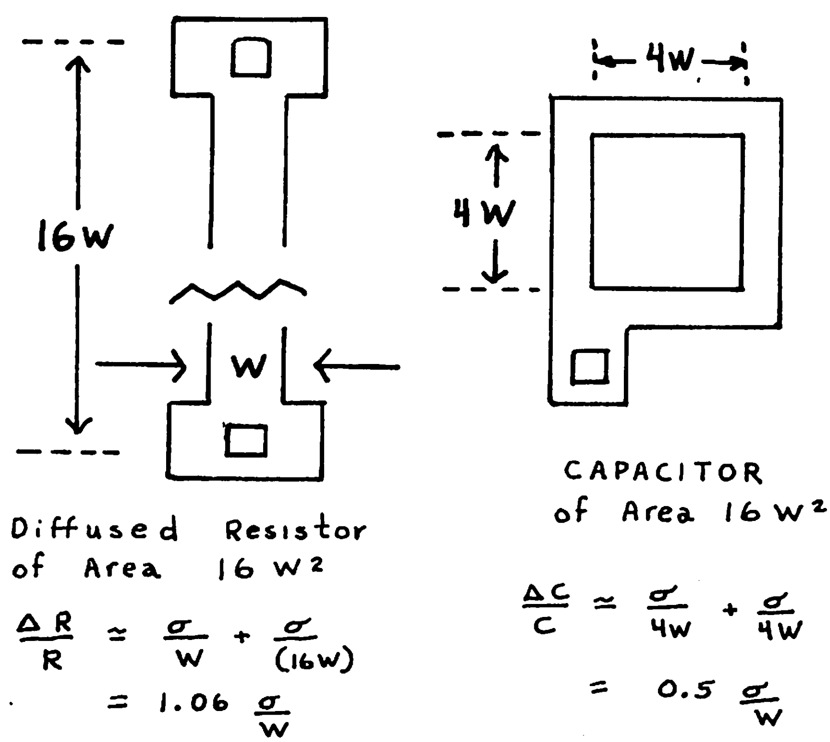


Figure 4.18: A comparison of capacitor matching and resistor matching.

value required for an 8-bit binary weighted resistor string. Then if σ is the uncertainty in line width due to the photolithography, the fractional variation in the resistor value is: $\frac{\Delta R}{R} = 1.06 \frac{\sigma}{w}$, but for the capacitor: $\frac{\Delta C}{C} = 0.5 \frac{\sigma}{w}$. Hence from a purely geometric argument, the capacitor matching is better by a factor of 2 for this example. Actually there is a third dimensional variable involved and this is the sheet resistance for the resistor and oxide thickness for the SiO_2 capacitor. Measurements indicate that excellent thin oxide uniformity can be obtained over an entire wafer.

4.9.2 Nonlinearity due to Ratio Errors

Consider the ideal case for a RADCAP type of circuit in which all the capacitors of the converter shown in Figure 4.19 have the precise binary weight values. For this case the digital output x is a regular staircase when plotted against V_{IN}, and every transition occurs at a precise value of V_{IN} designated V_T. On the other hand, changing one capacitor from its ideal value by a small amount ΔC_N causes all transition points to shift somewhat but there will be one worst case transition. The ratio $\frac{\Delta V_T}{V_R}$ is the normalized worst case fraction deviation in transition point from the ideal. This is also a measure of the nonlinearity. The ratio of this deviation to the fractional change in capacitor value $\frac{\Delta C_N}{C_N}$ represents the sensitivity of linearity to individual capacitor value. The plot of sensitivity also in Figure 4.19 shows that linearity is very sensitive to a fractional change in the large capacitors, but not very dependent upon similar fractional changes in the smaller capacitors. Therefore the smaller capacitors have greater allowable tolerances. It should be pointed out that actually all capacitors have simultaneous deviations

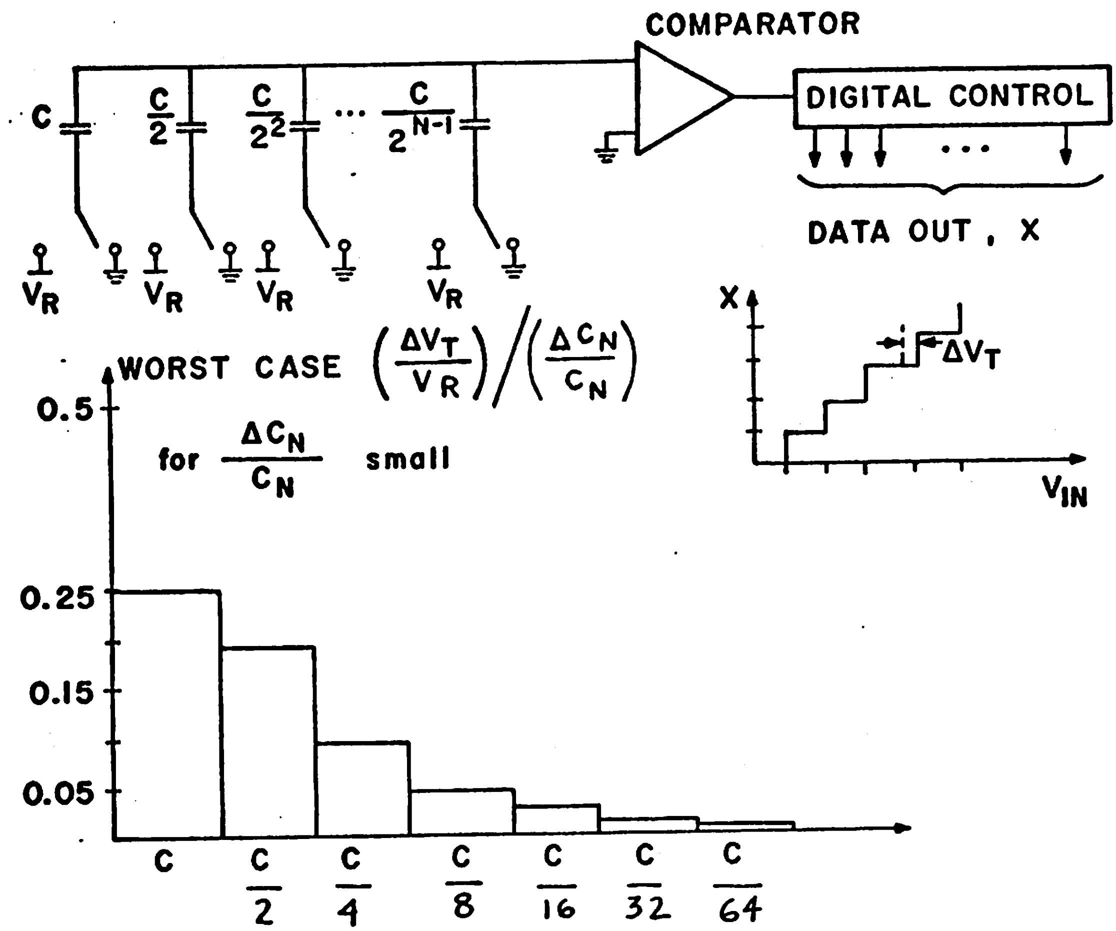


Figure 4.19: The sensitivity of A/D conversion linearity to deviation in individual capacitor values for RADCAP.

which cause ratio errors and the worst case combination of these must always be considered.

It will now be shown that even a simultaneous mismatch in the binary ratios of capacitors in the array causes only nonlinearity. It cannot cause a gain error because the end points of the transfer function, B_i vs V_{IN}, which is shown in Figure 2.3, are not dependent on capacitor matching. This results from the fact that no final charge-redistribution between capacitors occurs for either zero or full-scale inputs since all capacitors are either fully discharged or fully charged respectively. For the same reason no offset error can arise from capacitor mismatch since the mismatch cannot be manifested unless a charge-redistribution exists in the final configuration. This may be demonstrated analytically with the aid of Figure 4.20 for an N-bit RADCAP circuit. It is illustrated in the figure that for zero or full-scale inputs, the two boundary points in which all capacitors may be paralleled together, the resultant total mismatch error is zero. Another observation is that for $B_K = 2^N/2$ the total deviation is $\sum_{i=2^{N-1}}^{2^{N-1}} \Delta C_i$, but for $B_K - 1 = 2^{N-1} - 1$, the deviation is $\sum_{i=2^{N-2}}^{1B} \Delta C_i$.

Since $\sum_{i=2^{N-1}}^{1B} \Delta C_i$ equals zero then

$$\sum_{i=2^{N-1}}^{2^{N-1}} \Delta C_i = - \sum_{i=2^{N-2}}^{1B} \Delta C_i .$$

The significance of this is that the real transfer function must also pass through the midpoint of the ideal transfer function (considering only

Real Values		Ideal Values $C_T = 2^N C_1$		Capacitor Deviation
$C_{2^{N-1}}$	$=$	$C_T/2$	$+$	$\Delta C_{2^{N-1}}$
$C_{2^{N-2}}$	$=$	$C_T/4$	$+$	$\Delta C_{2^{N-2}}$
$\vdots$		$\vdots$		$\vdots$
C_{1A}	$=$	$C_T/2^N$	$+$	ΔC_{1A}
C_{1B}	$=$	$C_T/2^N$	$+$	ΔC_{1B}

$$C_T = \sum_{i=2^{N-1}}^{1B} C_i = C_T + \sum_{i=2^{N-1}}^{1B} \Delta C_i$$

$$\therefore \sum_{i=2^{N-1}}^{1B} \Delta C_i = 0$$

Figure 4.20. An illustration that the sum of capacitor deviations equals zero for RADCAP.

capacitor ratio errors). Moreover the generalized result:

$$\sum_{\substack{i = \text{all components} \\ \text{of } B_i}} \Delta C_i = - \sum_{\substack{j = \text{all components}, \\ \text{of } \overline{B}_i}} \Delta C_j$$

where $\overline{B}_i$ is the 1's complement of B_i, shows that the nonlinearity error at B_i is the negative of the nonlinearity error at $\overline{B}_i$. This may be summarized by stating that the capacitor ratio errors can cause only a nonlinearity, but that the behavior of this nonlinearity is such that the real transfer curve passes through the ideal midpoint and is an odd function about it. This is illustrated graphically in Figure 4.21 by the fact that both the ideal and real curves must intersect at the midpoint and the boundary points (considering only ratio errors which are independent of voltage).

4.9.3 Uniform Undercut

An important source of capacitor ratio error is uniform undercutting of the photoresist which defines the capacitor. Consider two capacitors C_4 and C_2 shown in Figure 4.22 which are nominally related by a factor of 2: $C_4 = 2C_2$. During the etching phase of the photomask process a poorly controlled lateral etch occurs called undercut. In most areas of the silicon wafer this effect will be evenly distributed and therefore will be characterized by a uniform reduction of each edge. Let Δx be the undercut length and L_4 be the side length of C_4, also as shown in Figure 4.22. Then a ratio error is produced between C_4 and C_2 which is proportional to the undercut length:

$$C_4 = 2C_2(1 + \epsilon); \quad \epsilon \simeq 4 \frac{\Delta x}{L_4}.$$

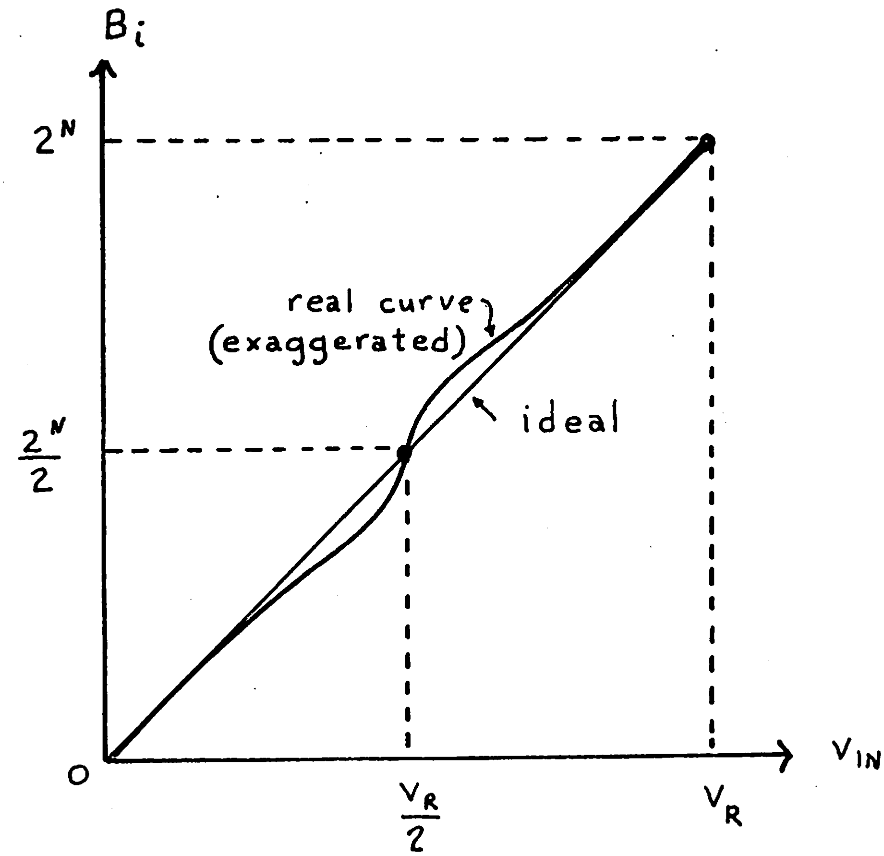
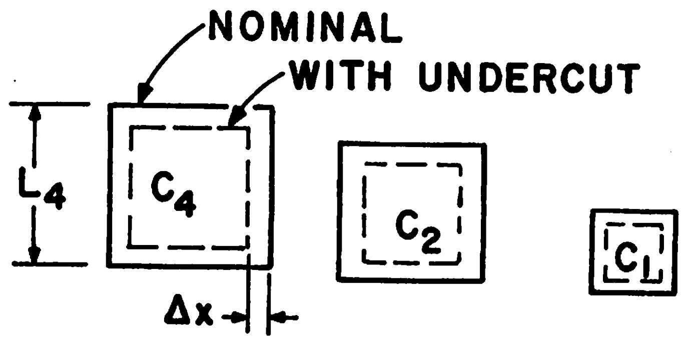


Figure 4.21: An illustration that capacitor ratio errors cause an odd-functioned nonlinearity about the input voltage midpoint.



NOMINAL : $C_4 = 2C_2$

WITH UNDERCUT : $C_4 = 2C_2(1+\epsilon)$; $\epsilon \simeq 4\dfrac{\Delta x}{L_4}$

UNDERCUT – INSENSITIVE GEOMETRY

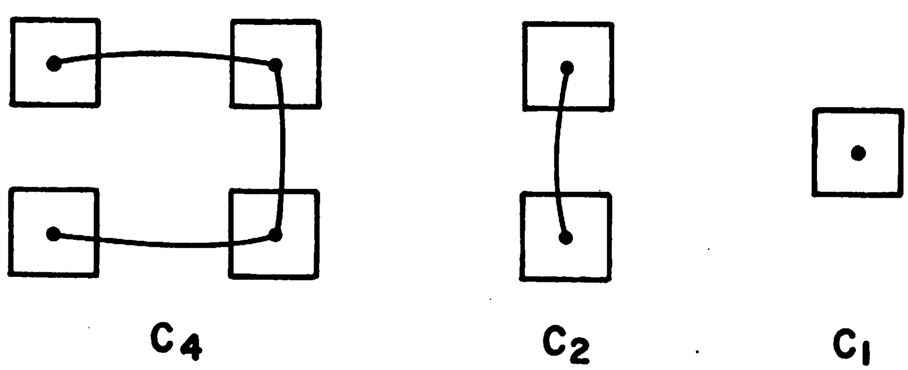


Figure 4.22: Capacitor ratio error due to photomask undercut.

This problem can be circumvented by a geometry such that the perimeter lengths as well as the areas are ratioed. This can be done as seen in Figure 4.22 by paralleling identical size plates to form the larger capacitors. Now the capacitor ratios are not affected by uniform undercut.

4.9.3 Oxide Gradient

Long range gradients in the thin capacitor oxide can also cause ratio errors. These gradients arise from non-uniform oxide growth conditions. If this variation in oxide thickness is approximated as first-order gradient as shown in Figure 4.23, then the resulting ratio error is proportional to the fractional variation in oxide thickness:

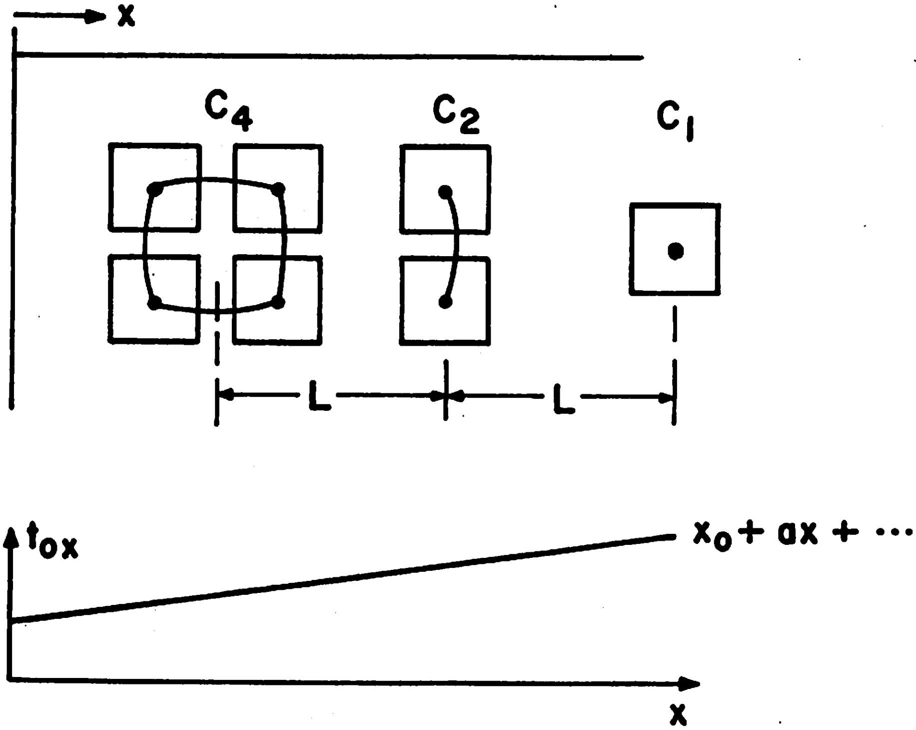
$$C_2 = 2C_1(1 + \in L)$$

$$C_4 = 4C_1(1 + 2\in L); \in = \frac{a}{x_0}.$$

Experimentally, values of 10 to 100 ppm/mil have been observed for the factor $\in$. Error from this source can be minimized by improved oxide growth techniques and by a common centroid geometry. This is done in Figure 4.23 by locating the elements of the capacitors in such a way that they are symmetrically spaced about a common center point. In this way the capacitor ratios may be maintained in spite of first-order gradients.

4.9.5 Non-Uniform Undercut

Non-uniform undercut usually appears in three forms as shown in Figure 4.24. Large-scale edge distortion is not clearly understood, however, one mechanism may involve chemical saturation. The etchant solution may become saturated with the etched material in some areas of the chip,



$$C_2 = 2C_1\left(1 + \frac{aL}{x_0}\right)$$

$$C_4 = 4C_1\left(1 + \frac{2aL}{x_0}\right) \quad ; \quad \frac{a}{x_0} \sim 10 \rightarrow 100 \text{ ppm/mil}$$

COMMON CENTROID GEOMETRY

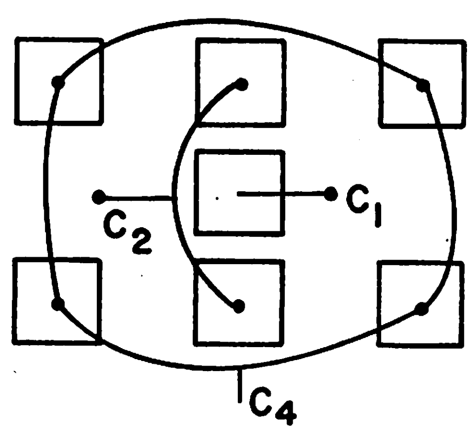


Figure 4.23: Capacitor ratio error due to oxide gradients.

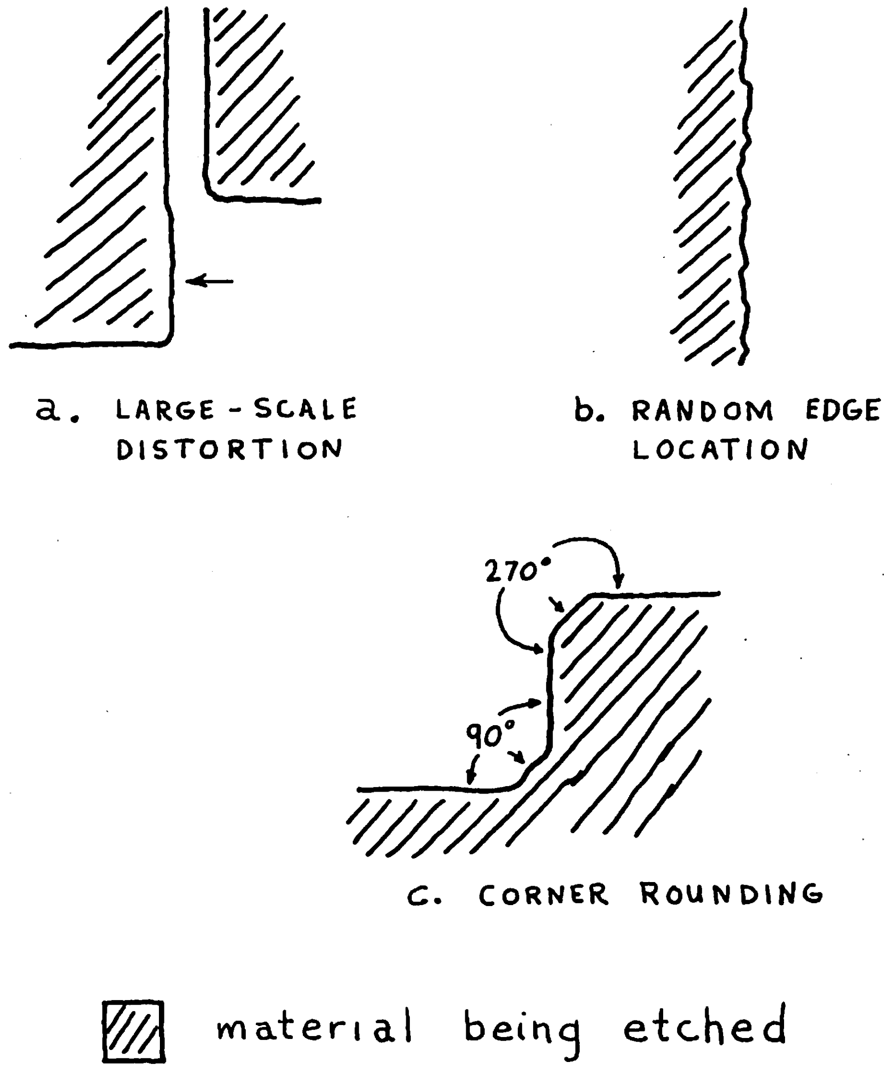


Figure 4.24: Non-uniform undercutting effects.

causing different etch rates. In addition regional temperature gradients caused by differences in amount of material being etched may cause regional etch rate variations.

The second type of non-uniform undercut is random edge location about which even less information appears in the literature. However, four processes may be involved. First the granular nature of the aluminum may cause local density variations and therefore localized etch rate variations. A "grainy" aluminum is usually caused by evaporation onto a heated wafer. This is usually done to promote adhesion of the metal to the dielectric. A cooled substrate would be a dubious improvement since the temperature of the metallic vapor would remain the same although the cooling-rate of the metal would increase. Another mechanism which might induce a random edge location is the error associated with the photolithography. For example, the glass plate emulsion mask may not have smooth edges. Or, the developed photoresist may have jagged edges due to light interference patterns. Random edge location may also result from localized temperature differences at the edge being etched since etch rate is a function of temperature and the chemical reaction releases energy. Finally, random edge location may result from the formation of stagnant gas bubbles about nucleation sites which retard the etch rate at localized points.

The last form of non-uniform undercut is corner-rounding. One difficulty in making precision capacitors with conventional photolithography is that corners cannot be made sharp. Figure 4.24 illustrates the distortion that results for both 270° and 90° corners. One apparent remedy for minimizing corners is to design circular capacitors. However, this is an unsatisfactory solution for a binary weighted capacitor array because at least 4 270° corners will be required due to the interconnect and mask alignment tab. A more realistic approach for a common centroid multiple plate capacitor array is one in which there are an equal number of 270° and 90° corners for each capacitor and also that the number of corners be binary ratioed between capacitors. As indicated in the figure this is an approximate solution since the excess area of the 270° corner is not exactly equal to that lost by the 90° corner.

Data indicates that large-scale edge variation may cause significant error at the 10-bit level, therefore, the design must involve control on this mechanism. A first-order control on etch rate is to locate every active edge the same distance from the opposite active edge. Furthermore, any reasonable fabrication changes which may reduce random edge should be considered.

4.9.6 Mask Alignment

The capacitor ratio must be independent of mask alignment. This can be done in conventional photolithography with the aid of an alignment tab which is parallel with the interconnect as shown in Figure 4.25.

4.9.7 Capacitor Ratio Error Due to Interconnect

Capacitors may differ from their design values due to the overlap of the interconnect on the thick oxide over the back electrode of the capacitor. The area of the interconnect which is effectively included in each capacitor value is not ratioed but is rather a constant offset added to each capacitor. Therefore this area must be included in design calculations in order to avoid ratio errors. Any absolute error in this overlap capacitance (as perhaps caused by thick oxide gradients) would cause a capacitor ratio error.

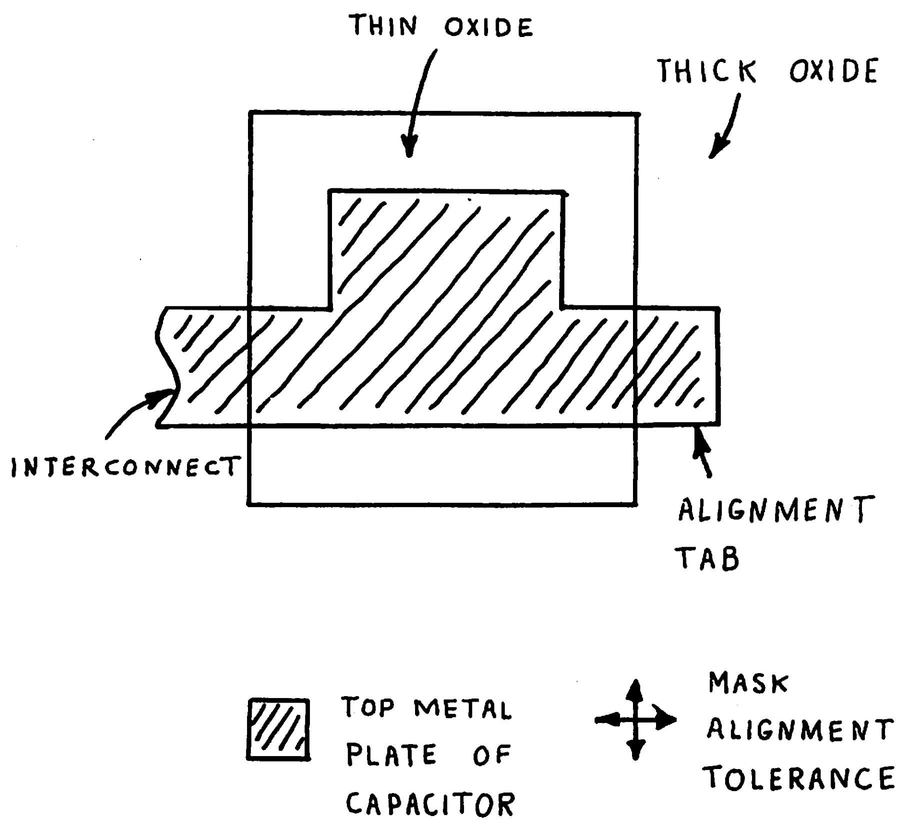


Figure 4.25: The insensitivity of capacitor area to mask alignment errors.

4.9.8 Fringing

A fringe field exists at the edges of the parallel plates as shown in Figure 4.26. The effect of fringing is to increase each dimension of the smallest plate by an amount equal to its thickness [33]. If the capacitor structure is such that this same plate suffers uniform from undercut, then there will be a first-order cancellation of undercut and fringing since undercut will tend to decrease the plate area by the same amount that it is increased by fringing. The two effects do not **exactly** cancel, however if the circuit is designed to be insensitive to undercut then it will automatically be unaffected by fringing.

4.11 Intrinsic Offset

It has been previously stated that a capacitor ratio error does not cause an offset and that the offset due to parasitic capacitance and comparator is small. However an intrinsic offset of $+\frac{1}{2}$ LSB still exists as indicated in Figure 4.27. The "intrinsic" transfer curve represents that for which the comparator is ideal and its offset has been cancelled as described in section 4.2. With the use of these techniques the comparator transfer function is defined by the equation

$$V_{out} = 1 \quad \text{if } V_x < V_T$$

but $$V_{out} = 0 \quad \text{if } V_x \geq V_T,$$

where V_T is the threshold voltage of the comparator. Then

$$V_x = B_i \frac{V_R}{2^N} - V_{IN} + V_T$$

due to comparator offset storage in the array. Hence,

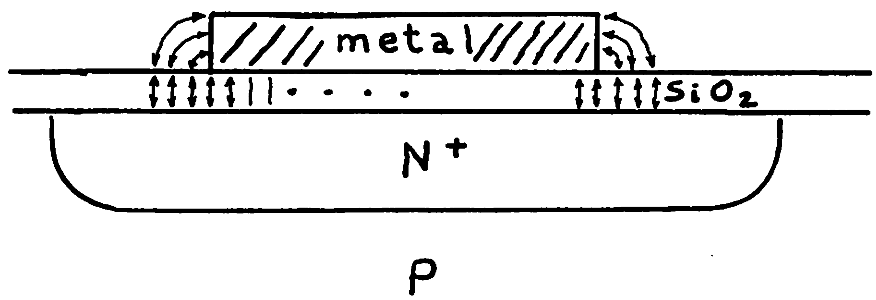


Figure 4.26: Electric field fringing effect.

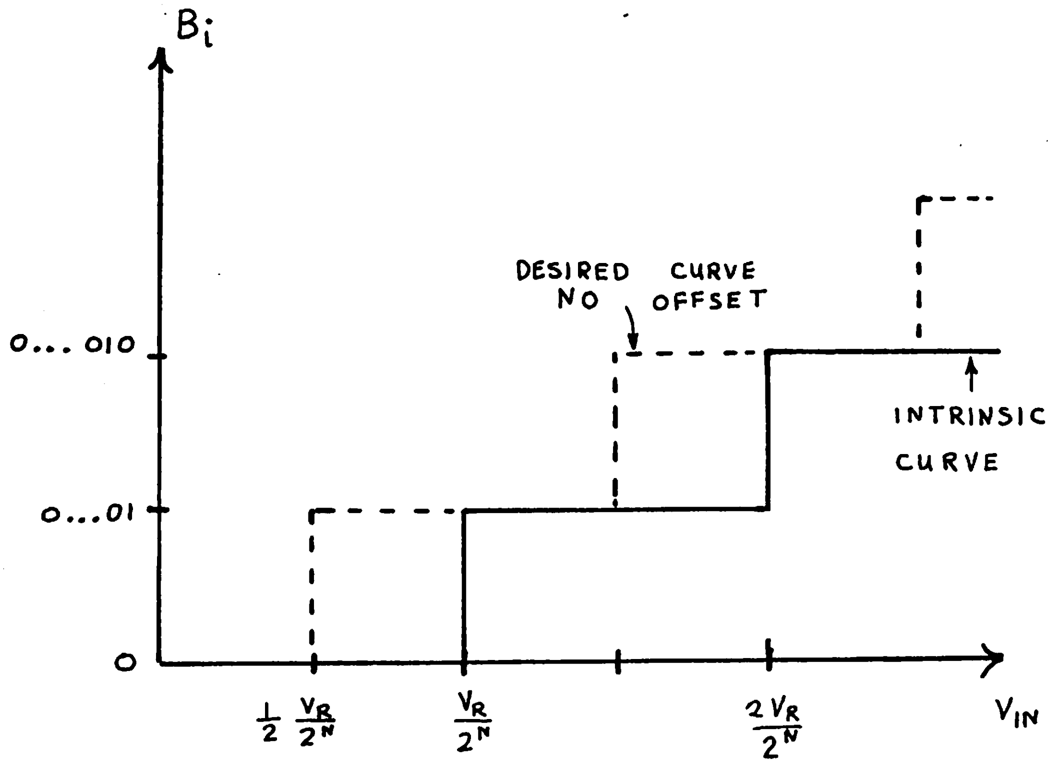


Figure 4.27: The $\frac{1}{2}$ LSB offset error of the intrinsic transfer function.

$$V_{out} = 1 \quad \text{if} \quad V_{IN} > B_i \frac{V_R}{2^N}$$

and therefore $B_i = 0$ until $V_{IN} > \frac{V_R}{2^N}$ at which point the first transition occurs. In view of this, a $-\frac{1}{2}$ LSB offset adjustment is required to give the "desired" curve which has its initial transition at $V_{IN} = \frac{1}{2}\frac{V_R}{N}$. This offset may be added to V_{IN} by level shifting the top plate by $-\frac{1}{2}\frac{V_R}{2^N}$ after the sample mode and comparator offset cancellation are complete, but before the redistribution mode. This may be accomplished on-chip with the extra unity weight capacitor C_{1A} which already exists in the array as shown in Figure 4.28. The voltage V_{1A} on the lower plate of C_{1A} pulses from V_{IN} down to $-\frac{V_R}{2}$ rather than to ground. The voltage $-\frac{V_R}{2}$ need only be accurate to $\pm$ 20% for an offset cancellation accuracy of $\pm$ .1 LSB. Also this voltage is readily available since the substrate bias supply is nominally - 5 volts. This technique provides the desired initial transition at $V_{IN} = \frac{1}{2}\frac{V_R}{2^N}$. An implication of offset cancellation is that now the top plate voltage during the final configuration is V_x such that:

$$V_T \geq V_x = \frac{B_i V_R}{2^N} - V_{IN} - \frac{1}{2}\frac{V_R}{2^N} + V_T \geq V_T - 2\epsilon_q .$$

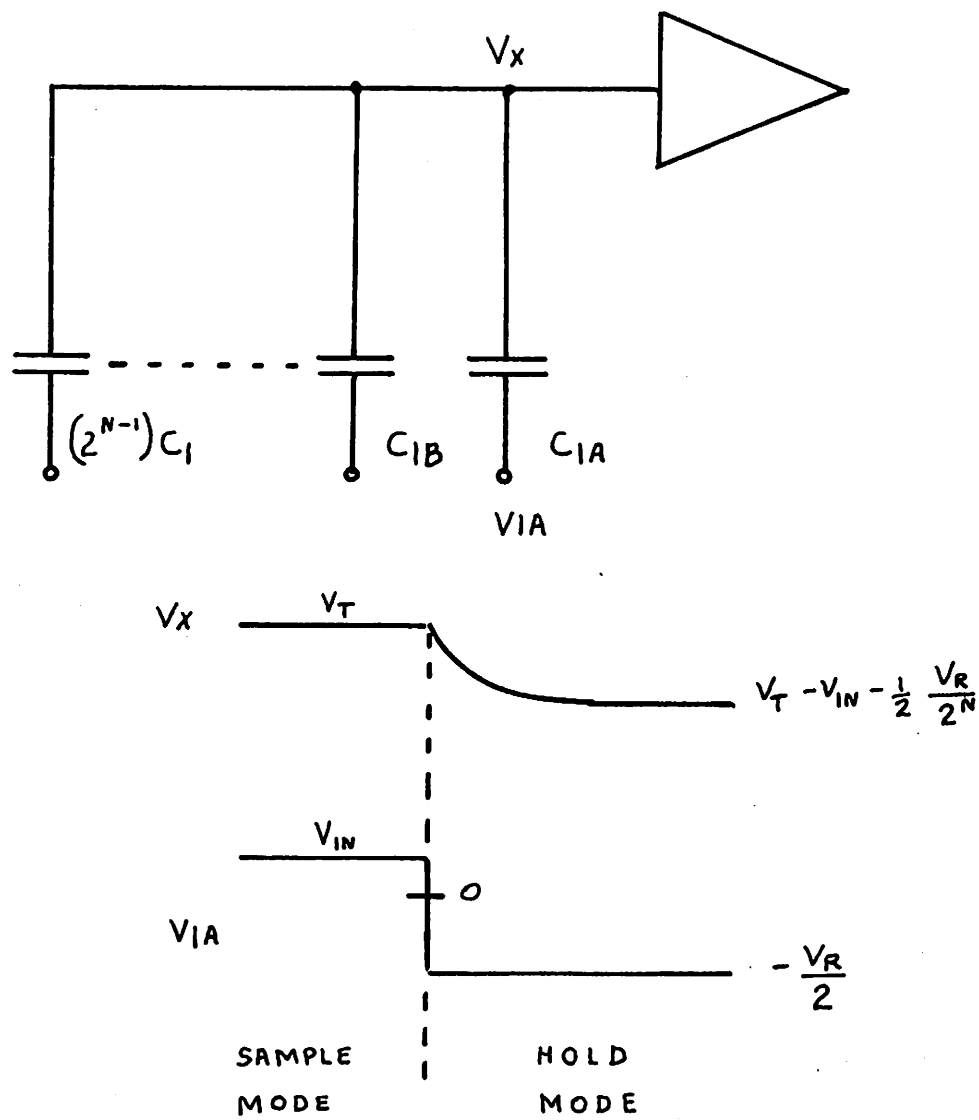


Figure 4.28: Intrinsic offset cancellation.

CHAPTER V

Factors Limiting Conversion Rate in RADCAP

5.1 Introduction

Compared with many conversion techniques the successive approximation method used in RADCAP is capable of rapid conversion. In this chapter the factors limiting the conversion rate are discussed from both theoretical and practical perspectives. In section 5.2 the acquisition time requirements for VATCAP when used as a sample/hold (S/H) circuit are examined. This analysis considers two criteria: the sampling accuracy specification and offset cancellation. This section also investigates the input bandwidth and distortion of high frequency signals. The minimum time required for one charge-redistribution cycle in VATCAP is discussed in section 5.3. An examination of the comparator delay time appears in section 5.4. The proper summation of all of these factors provides the maximum theoretical conversion rate as shown in section 5.5. However, there are practical limitations which dominate when conventional fabrication methods are used. These are discussed in section 5.6.

5.2 Factors Limiting the Minimum Acquistion Time when VATCAP is used as a S/H circuit

5.2.1 Relationship Between Acquistion Time and Sampling Accuracy

VATCAP serves two special functions in addition to D/A conversion. It provides offset cancellation for the comparator input stage as explained in Chapter IV and it also performs a S/H function for the input signal. This function results in a sample mode precharge delay referred to as the acquistion time T_{aq}. The nature of this delay is illustrated in Figure 5.1. The VATCAP attenuator is shown with scaled bottom plate switches. The

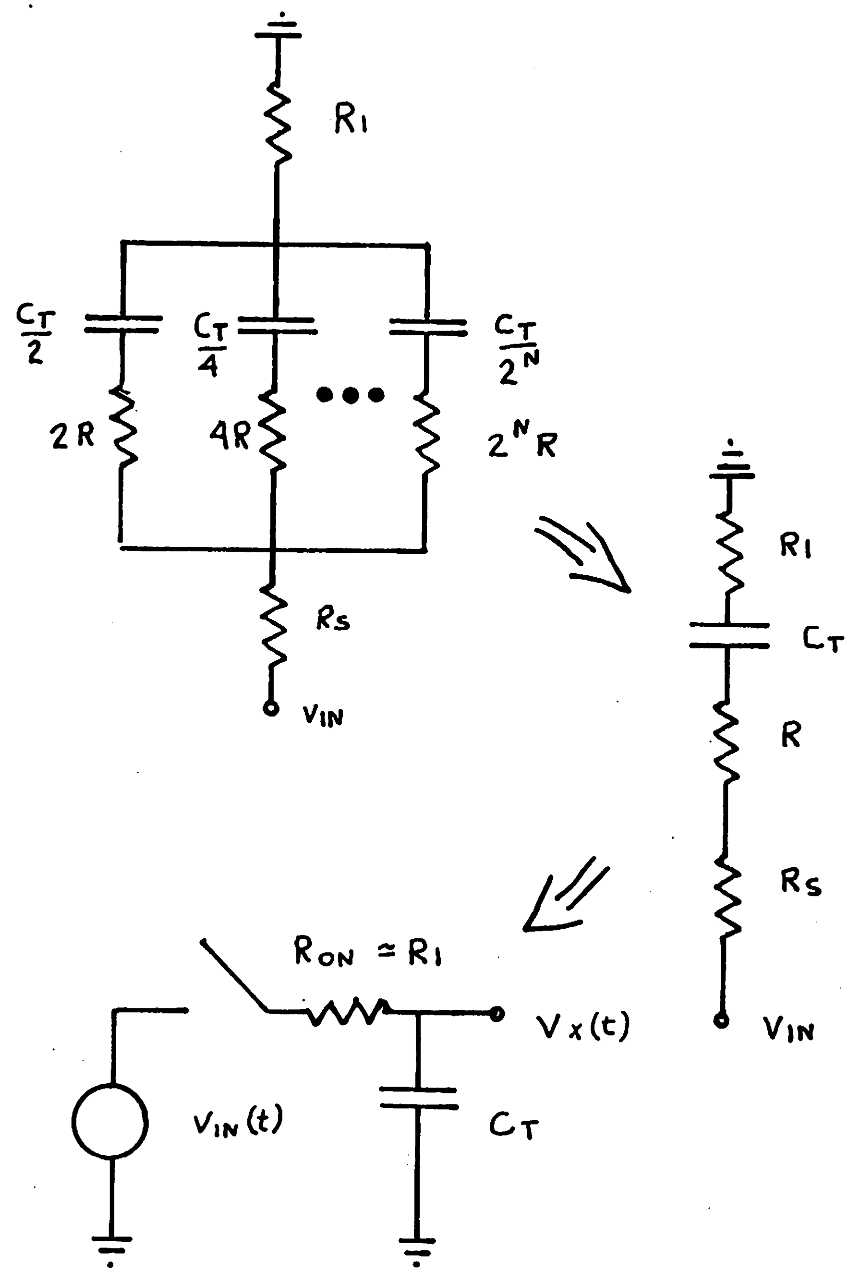


Figure 5.1: The equivalent circuit for VATCAP used as a sample-and-hold circuit during the sample mode precharge.

equivalent S/H circuit is reduced to a series RC circuit. R_{ON} represents the "ON" resistance of the MOS switches which will be approximately the same value as R_1 the "ON" resistance of the grounding switch because R and R_s can be made arbitrarily small compared with R_1. This assertion is based upon the fact that large devices may be used for lower plate switching without affecting the accuracy of the conversion but the same is not true for the grounding switch. This will be clarified in section 5.2.2. Therefore the time constant during precharge is

$$\tau = [R_s + R + R_1]C_T \approx R_1\ C_T.$$

The equation which describes $V_x(t)$ is

$$V_x(t) = \mathcal{L}^{-1}\left[\frac{V_{IN}(s)}{1 + \tau s}\right] \text{ where}$$

$\mathcal{L}^{-1}$ is the inverse Laplace transform and s is a complex number in frequency domain. For the case that $V_{IN}(t)$ is a step input of amplitude V_R then:

$$V_x(t) = V_R(1 - e^{-\frac{t}{\tau}}).$$

According to this solution the error voltage is $V_E = - V_R\ e^{-\frac{t}{\tau}})$. If the allowable error is $\frac{V_R}{2^{N+1}}$ or $\frac{1}{2}$ LSB then the minimum acquisition time is $T_{aq} = (N+1)\tau\ \ell n\ 2$. The solution may also be determined for the case in which the input is a ramp with a d.c. offset voltage V_o:

$$V_{IN} = \frac{V_R}{T_{IN}}\, t + V_o \quad \text{where } \frac{V_R}{T_{IN}}$$

represents the input slew rate. For this case

$$V_x(t) = V_{IN}(t) - \frac{V_R}{T_{IN}} \tau(1-e^{-\frac{t}{\tau}}) - V_o e^{-\frac{t}{\tau}}.$$

The error voltage is

$$V_E = - \frac{V_R}{T_{IN}} \tau(1-e^{-\frac{t}{\tau}}) - V_o e^{-\frac{t}{\tau}}.$$

The initial value of V_E is $- V_o$ at $t = 0$ but at $t = T_{aq} = (N+1)\tau \ln 2$ the error voltage converges to $- \frac{V_R \tau}{T_{IN}}$. Since this value is a constant, $V_x(t)$ may be described as following $V_{IN}(t)$ but attenuated in peak amplitude and delayed by a time τ.

If the input were a sine wave of amplitude V_A and offset V_o:

$$V_{IN}(t) = V_A \sin \omega t + V_o,$$

then

$$V_x(t) = \frac{V_A \sin \omega t}{1 + \omega^2\tau^2} + V_o - \frac{V_A\, \omega\tau(\cos \omega t - e^{-\frac{t}{\tau}})}{1 + \omega^2\tau^2} - V_o e^{-\frac{t}{\tau}}.$$

In this example V_o is used for convenience to represent an initial difference voltage between V_x and V_{IN} although this could also have been done with a phase shift. The error voltage at $t = 0$ is $- V_o$. At

$$t = T_{aq} = (N+1)\tau \ln 2$$

$$V_x(t) = \frac{V_A}{1 + \omega^2\tau^2} [\sin \omega t - \omega\tau \cos \omega t] + V_o.$$

After performing a trigonometric combination $V_x(t)$ becomes:

$$V_x(t) = \frac{V_A}{\sqrt{1 + \omega^2\tau^2}} \sin(\omega t - \theta) + V_o$$

where $\theta = \arc \tan \omega\tau$. It is apparent therefore that the S/H function results in amplitude reduction by a factor $\frac{1}{\sqrt{1 + \omega^2\tau^2}}$ and a phase shift of $-\theta$ for high frequency signals. The S/H behaves like a low pass filter having a bandwidth $\frac{1}{\tau}$. For example if $\omega = \frac{1}{\tau}$ then the signal amplitude of $V_x(t)$ is reduced to $\frac{V_A}{\sqrt{2}}$ and the phase shift is $-45°$.

In conclusion, for all three cases considered, the sampled signal $V_x(t)$ remains undistorted relative to quantization distortion provided that $T_{aq} \geq (N+1)\ \tau\ \ell n\ 2$. The minimum acquisition time is therefore proportional to the minimum value of τ.

5.2.2 Minimum Acquisition Time with Offset Cancellation Technique

It has just been determined that T_{aq} is proportional to τ. It is therefore desirable to determine the minimum value of τ. It will be shown that for the RADCAP technique this depends upon the extent to which feed-through effects can be cancelled. The offset cancellation scheme described in section 4.2 and Figure 4.3 cancels the additional offset error V_{FT} introduced by the capacitive feedthrough of switch S1. However there are limitations on the magnitude of V_{FT} since the stages A1 and A2 in Figure 4.4 must remain biased in a linear gain region. This technique will be described in detail in Chapter VI but at this point it is sufficient to model the circuit as shown in Figure 5.2. From this figure $\tau = R_1\ C_T$ as before but the "ON" resistance of switch S1 is given by

$$R_1 = \frac{1}{\frac{W}{L_c}\ \mu\ C_{ox}\ (\Delta V)}$$

where $\Delta V = V_{os}(ON) - V_T$ and $\frac{W}{L_c}$ is the channel width to length ratio [34].

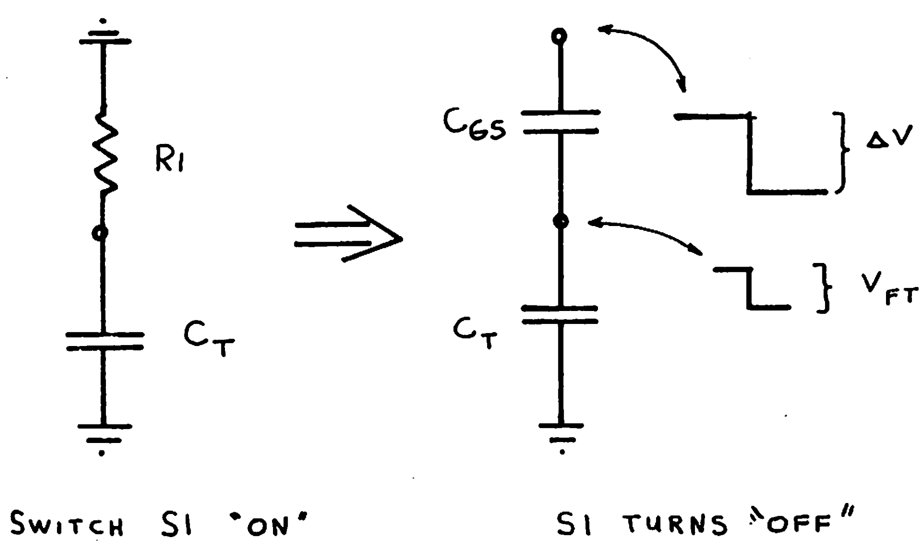


Figure 5.2: An illustration of the capacitive feedthrough of switch S1.

C_{ox} is channel capacitance per unit area of the thin oxide and μ is the effective electron mobility for N-MOS devices. Then the total array capacitance $C_T = A_T C_{ox}$ where A_T is the total top plate area. Therefore

$$\tau = \frac{A_T}{\frac{W}{L_c}\ \mu\ \Delta V} .$$

The effect of capacitive feedthrough when S1 turns off is a voltage drop at the comparator input of:

$$V_{FT} \cong \frac{C_{GS}\ \Delta V}{C_T} \quad \text{where}$$

C_{GS} is the gate-source capacitance of S1 and is assumed to be much less than C_T. Using C_{GS} approximately equal to $C_{ox}\ W\ L_c$ then the product:

$$V_{FT}\ \tau = \frac{L_c^2}{\mu} \quad \text{where}$$

L_c is the minimum channel length for the given supply voltage to avoid drain-source breakdown. Therefore a trade-off exists between V_{FT} and τ. The smallest value of τ is dependent upon the largest value of V_{FT} which may be cancelled by circuit techniques. Furthermore this trade-off prevents R_1 (the resistance of S1) from being as small as R or R_s since the capacitances of these switches do not cause errors. Hence the minimum value of τ is

$$\tau = \frac{L_c^2}{\mu\ V_{FT}}$$

and therefore the acquisition time is

$$T_{aq} = \frac{(N+1)}{V_{FT}} \quad \frac{L_c^2 \ \ln 2}{\mu} \quad .$$

The absolute minimum acquisition time will be estimated for the theoretical limit in which $V_{FT} = V_{DD} \simeq 15$ V and $L_C = 5\ \mu$, N = 10 bits, and $\mu = 500$ $\frac{cm^2}{V - s}$. Then for this example $T_{aq} = 0.25$ ns.

5.3 Factors Limiting the Minimum C/R Time for RADCAP Class of Circuits

The equivalent small signal model for VATCAP during C/R is illustrated in Figure 5.3 from which the circuit time constant is

$$\tau_{CR} = R(C_D + C_T).$$

R represents the total resistance of all bottom plate switches in parallel as before and C_D is the total drain to substrate capacitance of all switches in parallel. An equation for R is

$$R = \frac{1}{\mu \frac{W}{L_c} C_{ox} \Delta V}$$

as defined in section 5.2. The minimum value of drain capacitance may be expressed as $C_D = C_{pn}\ W\ L_D$ where C_{pn} is the capacitance per unit area of the drain pn junction and L_D is the minimum length of the drain diffusion. Therefore the resultant expression for τ_{CR} is

$$\tau_{CR} = \left(\frac{A_T}{A_c} + \frac{C_{pn}}{C_{ox}} \frac{L_D}{L_c}\right)\left(\frac{L_c^2}{\mu \Delta V}\right),$$

where A_c is the total channel area of all switches, $A_c = WL_c$. The first term dominates under normal conditions since $A_T \gg A_c$ but $C_{pn}L_D$ and $C_{ox}L_c$

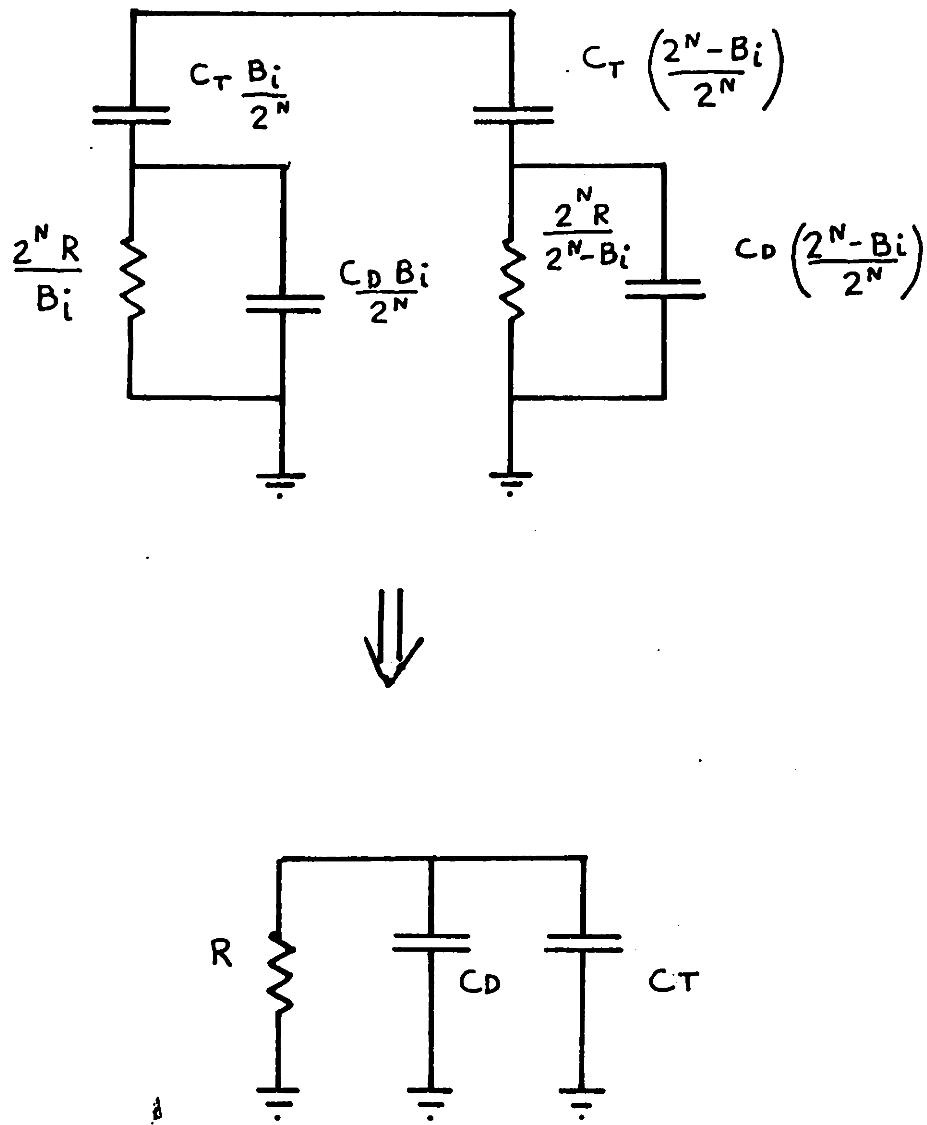


Figure 5.3: The equivalent circuit for VATCAP during the charge-redistribution mode.

may be of the same order of magnitude. In the theoretical limit that the capacitor array could be made small compared with the device size τ_{CR} approaches $\frac{C_{pn}}{C_{ox}} \frac{L_D L_C}{\mu \Delta V}$. In effect the switches must charge their own capacitances. The total time required for N C/R cycles, each requiring $(N+1)\ln 2$ time constants to go to completion, is $T_{CR} = N(N+1)\tau_{CR} \ln 2$.

5.4 Factors Causing Comparator Delay

For N bits of resolution, N comparisons must be performed by the comparator. It is of interest to examine the fundamental limitations resulting in comparator delay. The comparator must have sufficient gain to resolve the minimum input signal which is $\frac{V_R}{2^N}$. The output voltage swing must be approximately $V_{DD} - 2V_T$ since this must be compatible with the digital logic levels. Therefore the approximate minimum comparator gain required is

$$A = 2^N \frac{(V_{DD} - 2V_T)}{V_R} \quad .$$

Assume for a moment that this gain is realized by direct linear amplification. If slew rate limiting is not a problem then an optimistic transfer function describing this comparator is:

$$A(s) = \frac{A}{1 + s\,\tau_o}$$

in which $\tau_o = \frac{A}{2\pi f_T}$ and f_T is the unity-current-gain frequency of the devices. Hence the time required for N comparisons each settling to within $\frac{100}{2^{N+1}}$ % of final value is

$$T_{COMP} = \frac{N(N+1)\ 2^N \ln 2\ (V_{DD} - 2\ V_T)}{2\pi f_T\ V_R} \approx \frac{N^2\ 2^N \ln 2}{2\pi f_T}$$

for $\qquad V_{DD} - 2V_T \approx V_R.$

From this equation the gain-bandwidth product limitations involved in direct amplification cause a severe comparator delay.

Consider a different approach in which the comparator function is performed by a bistable latch having positive feedback as shown in Figure 5.4 [35]. The gain of the basic amplifier is expressed in terms of a single time constant as before:

$$A(s) = \frac{A}{1 + s\tau}$$

The total transfer function of the feedback amplifier is

$$\frac{V_2(s)}{V_1(s)} = \frac{-A(s)}{1 - (-A(s))^2}$$

for which the approximate time domain solution is:

$$V_2(t) = \frac{\Delta V_L}{2} e^{\frac{t(A-1)}{\tau}} .$$

In this expression ΔV_L is the initial voltage difference in the latch. Evaluating the delay time for the full output voltage swing, $V_2(T_{LATCH}) = V_{DD} - 2V_T$, then

$$T_{LATCH} = \frac{\tau}{A-1} \ln \left[\frac{2(V_{DD} - 2V_T)}{\Delta V_L} \right] .$$

From this expression the delay time is inversely proportional to (A-1) hence a large value of A is desirable. In fact, it would be desirable to realize the full gain of the comparator with the latch circuit due to the speed advantages of positive feedback over linear amplification. If this were done, the latch would have to switch properly for a minimum differential input signal of $\Delta V_L = \frac{V_R}{2^N}$ in order that 1 LSB be resolved.

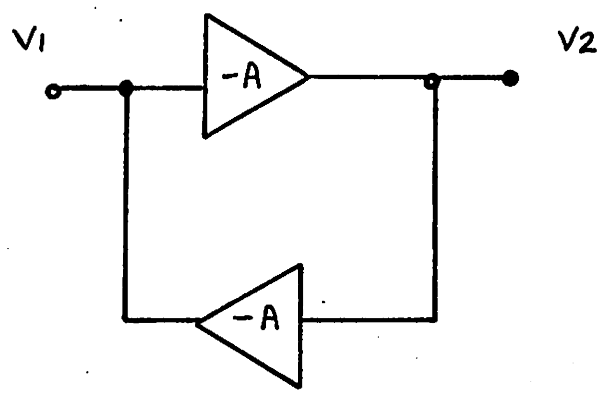


Figure 5.4: A bistable regenerative latch.

For simplicity $V_{DD} - 2V_T$ is taken to be approximately the value of V_R,

and $\frac{\tau}{A-1} \approx \frac{1}{2\pi f_T}$

$$T_{LATCH} \approx \frac{\tau}{A-1} (N+1) \ \ell n \ 2 \approx \frac{(N+1) \ \ell n \ 2}{2\pi f_T} .$$

Then $T_{COMP} = N \ T_{LATCH} \approx \frac{N^2 \ \ell n \ 2}{2\pi f_T}$. This quantity is compared with total comparator switching time on page 124. It is found that the regenerative latch switches approximately 2^N times faster than the amplifier. This result implies that ideally the comparator should consist of a latch and that linear amplification is to be avoided since it is a slower process than regenerative latching.

It is instructive to consider the theoretical switching time of the latch by expressing τ in terms of device and circuit parameters. By invoking an approach similar to that in section 5.3 the dominant time constant is modeled by a junction capacitance and thin oxide capacitance in parallel with the "ON" resistance of a large device. This analysis will assume for convenience that the amplifier having gain - A is a simple MOS active load inverter, hence, for a given $V_{GS}(ON)$ in the load:

$$\tau = C_L \times R_{ON} = \frac{W(L_D C_{pn} + L_C C_{ox}) A^2}{A^2 \ \mu \ \frac{W}{L_c} C_{ox} \ (V_{GS}(ON) - V_T)}$$

Taking an average value of $\frac{V_{DD} - 2V_T}{2}$ for $(V_{GS}(ON) - V_T)$ the following expression is obtained:

$$T_{COMP} = \frac{N(N+1)}{A-1} \left(\ell n \ 2\right) \frac{L_c^2}{\mu} \left(1 + \frac{L_D C_{pn}}{L_c C_{ox}}\right) \frac{2}{V_{DD} - 2V_T}$$

It is reiterated that this is a fundamental limitation rather than the nominal delay of a real circuit.

In reality the asymmetry and mismatches in the latch will reduce its ability to resolve small signals. Therefore the smallest practical input signal is determined by the effective input offset voltage of the latch, VINOS. Hence a linear amplification stage having a minimum gain of approximately $\frac{VINOS}{V_R} 2^{N+1}$ must preceed the latch in order to cancel the offset reflected to the input.

The results in this section support the conclusion that comparator delay will be largely determined by the amount of linear amplification required prior to the latch. Some reduction in delay can be realized by having this linear stage directly coupled to the comparator input during the C/R cycle. In this way the C/R delay and linear amplification delay may be nearly coincident. In retrospect the comparator having a regenerative latch may be modeled with infinite gain and an input offset voltage as asserted in Chapter IV.

5.5 The Theoretical Minimum Conversion Time

The theoretical minimum conversion time will be estimated on the basis that logic and aperture delays could conceivably be made small. In this case the total conversion time is given by:

$$T_c = T_{aq} + T_{CR} + T_{COMP}$$

$$= \frac{N+1}{V_{FT}} \frac{L_c^2 \ \ell n\ 2}{\mu} + \frac{N(N+1)}{\Delta V} \frac{L_c^2 \ \ell n\ 2}{\mu} \left(\frac{A_T}{A_c} + \frac{C_{pn} L_D}{C_{ox} L_c} \right)$$

$$+ \frac{N(N+1) L_c^2 \ \ell n\ 2}{(A-1)\mu} \left(1 + \frac{C_{pn} L_D}{C_{ox} L_c} \right) \left(\frac{2}{V_{DD} - 2V_T} \right)$$

The ultimate limit will now be computed with the following assumptions:

$$\Delta V = V_{FT} = V_{DD}; \ A = 2;$$

$$C_{pn} L_D = C_{ox} L_c;$$

and that $A_T < A_c$

Thus $$T_c \simeq \frac{(N+1)L_c^2 \ \ell n\ 2}{V_{DD}\ \mu}(5N + 1)$$

Using $L_c = 5\ \mu$, $N = 10$ bits, $V_{DD} = 15V$ and $\mu = 500\ \frac{cm^2}{V\text{-}s}$ the ultimate limit is approximately:

$$T_c = 22\text{ ns.}$$

This result indicates that theoretically the A/D conversion time may be quite small. Two important factors which influence this hypothetical limit are feedthrough cancellation and the charging of capacitance. The practical limits are considered in the next section.

5.6 Practical Limitations on Conversion Rate

Some practical considerations which limit conversion rate will now be discussed. The inspection of the time constants τ and τ_{CR} from sections 5.2 and 5.3 indicates that the time constants are directly proportional to the switch "ON" resistance and to the total capacitance C_T. Therefore increased speed may be achieved by larger switches (over some practical range); however, the switches may not be so large that the maximum instaneous current is excessive. Additional considerations are therefore power dissipation and chip area. A smaller total capacitance also reduces the time constants but the minimum total capacitance is limited by the resolution properties of the standard photomasking process rather than by the parasitic capacitance. The limitations of the photolithography are manifested as undercut error, corner-rounding, and random edge variations which combine statistically to produce a distribution of errors in total capacitor area with standard deviation σ_A. The minimum value of C_T is proportional to the minimum total area A_T such that $\frac{\sigma_A}{A_T} < \frac{1}{2^{N+1}}$ for N bits of resolution. This design constraint reflects the fact that higher yields of accurately ratioed arrays will result if the uncertainty in capacitor areas is small compared with the area of the unity weight capacitor.

CHAPTER VI

Description of an Experimental ADC

6.1 Introduction

The techniques described thus far should allow the fabrication of a single-chip MOS ADC. However only the critical portion of the circuit, the capacitor array and the comparator have been fabricated in order to demonstrate the feasibility of the RADCAP technique. The remaining element is the digital control circuitry, the MOS implementation of which is straightforward. The block diagram of Figure 6.1 shows a complete ADC and those components which were fabricated as part of the experimental I.C. The digital circuitry is composed of TTL gates. In this chapter the designs of the experimental chip and the logic system are discussed.

6.2 Optimization of MOS Capacitor Geometry

The realization of a high resolution ADC of the RADCAP type requires the fabrication of precisely matched capacitors. The implications of Chapter IV suggest that the key elements of this realization are multiple plates, common centroid geometry, corner compensation and etch rate control. Another requirement is mask misalignment tolerance. Finally the capacitor itself should depend upon as few variables as possible in order to reduce uncertainties.

Most of the design criteria can be achieved with conventional photolithography with the MOS capacitor structure shown in Figure 6.2. The capacitor top plate is aluminum and its area on top of the oxide covering the N^+ defines the capacitor. The heavily doped N^+ diffusion acts as the bottom plate. Most of the capacitance value is due to the thin oxide, however, a small fraction of the capacitor value is due to interconnect

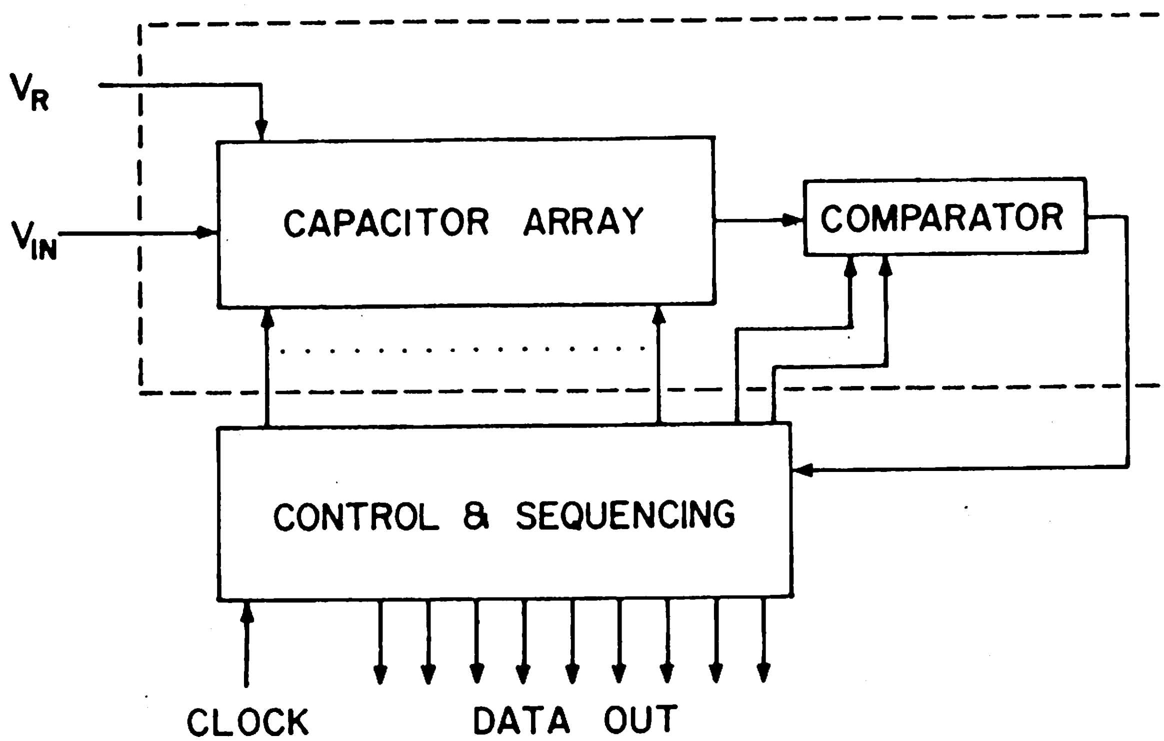


Figure 6.1: A complete ADC. The experimental I.C. is defined by the dashed lines.

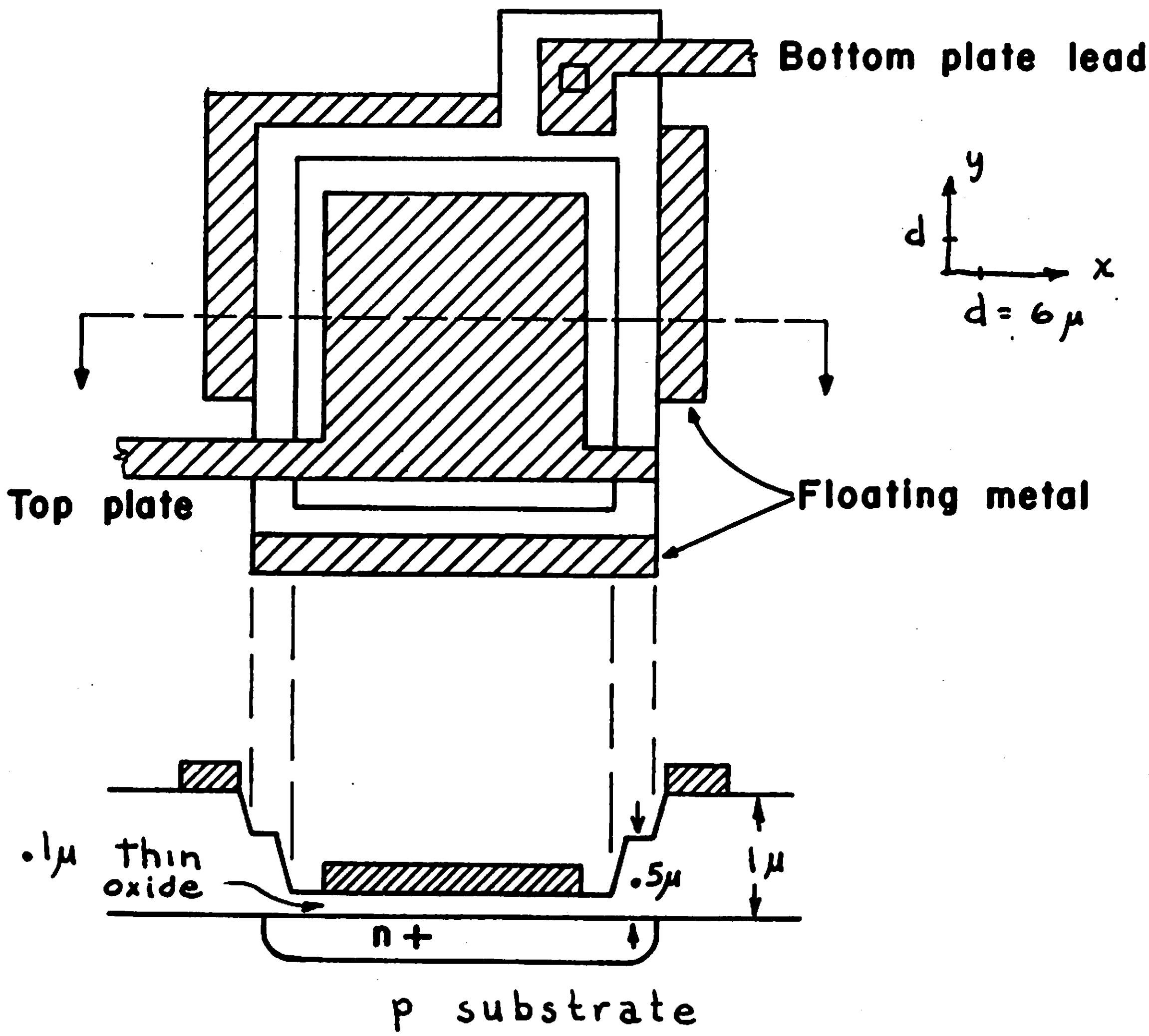


Figure 6.2: MOS precision capacitor structure.

passing over the thicker oxide above the N^+ region. The interconnect over the field oxide constitutes a parasitic capacitance to ground but does not add to the design value of the MOS capacitor. The thin oxide is nominally chosen as 1200 Å since this is the thickest gate oxide desirable. The field oxide is chosen to be 1 μ since this value results in a field device threshold voltage in excess of 20 volts which is desirable. The 0.5 μ oxide over the N^+ is arbitrarily picked to provide an intermediate step between oxides but should be thick enough to minimize overlap capacitance. On the other hand this oxide must not be so thick that the effective N^+ drive-in during oxide growth significantly reduces the N^+ surface concentration since this will increase the voltage coefficient of capacitance. The alignment tab shown as an extension of the capacitor plate makes the capacitor value independent of misalignment error d in the X-axis while no error results from a Y-axis misalignment of $\pm$ d. The floating metal strip helps to maintain a uniform undercut during the aluminum etch. The actual effect of the strip upon the circuit will be to add parasitic capacitance at the bottom plate which was discussed in Chapter IV. The circuit schematic for the capacitor array is shown in Figure 6.3.

6.3 MOS Comparator Realization

The MOS comparator is shown in Figure 6.3 along with the capacitor array. The device aspect ratio, $\frac{W}{L}$ (channel width divided by channel length), is given for each transistor. The basic operation of the comparator has been outlined in Chapter IV. It consists of one precharge cycle during the sample mode, a hold mode, and ten tests during the redistribution mode.

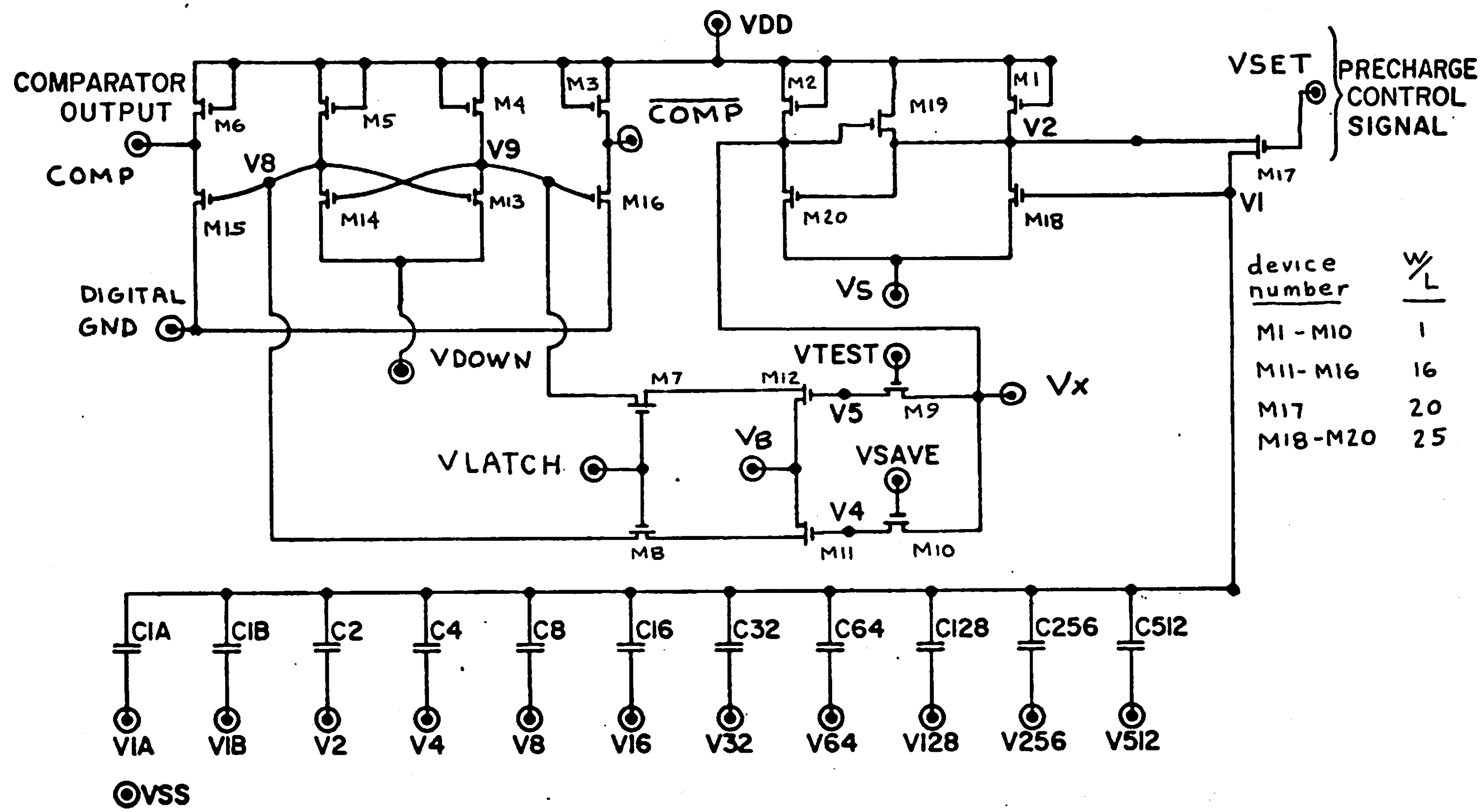


Figure 6.3: A circuit schematic of the experimental I.C.

The sample mode precharge cycle is initiated by a 15 volt VSET signal on the gate of M17. V_S is a supply voltage nominally chosen to be zero volts. In the worst-case precharge V_1 is negative and V_2 is also low because M17 is in the triode region of operation. This forces M18 off. Since M18-M1 and M20-M2 comprise two gain stages A1 and A2 respectively M20 must also be off because V2 is low. Therefore the output of A2, V_x, must be high which turns M19 heavily "ON". M19 is a feedback switch that shunts the output of stage A1 thereby reducing its effective output resistance. This causes a rapid precharge of the array since both M19 and M17 are large devices. When the array charges up positively and both V1 and V2 become larger, V_x will decrease as V2 increases above the nominal 2 volt threshold voltage of M20. M19 will eventually turn off when V2 is about 0.5 volt less than the "switching threshold" voltage of stage A2 (that value of voltage for a digital circuit for which $V_{IN} = V_{out}$ or $V2 = V_x$ in this case). However, when this occurs V1 is nearly equal to V2 since M17 is heavily "ON" in the "non-saturated" or "triode" region of operation. Therefore a d.c. steady state solution for A1 and A2 is:

$$V1 = V2 = V_x = V_{balance}.$$

M19 will remain off until V_x increases by about 2 volts above V2. Hence M19 does not destroy the small signal gain about the balance point, $V_{balance}$. In fact M19 actually performs an additional useful function by clamping the maximum value of V_x perhaps reducing the effects of capacitive coupling between the high gain stages and the rest of the circuit. After approximately 1 μs or 2 μs of precharge the steady state condition is reached and the sample mode precharge is complete.

Both VSET and VSAVE pulse down turning off M17 and M10 which were heavily "ON" at the completion of the precharge cycle. This initiates the hold mode and it is characterized by the storage of $V_{balance}$ at node V4. After the switching transients settle V_x will be significantly less than $V_{balance}$ due to the capacitive feedthrough effects of VSET upon V1: $V_x = \left(V_{balance} - V_{FT} A1\ A2\right)$ where A1 and A2 are the gains of stages A1 and A2. This is an apparent offset at V1 and may be cancelled simultaneously along with the intrinsic offset cancellation. The desired value of V1 is:

$$V1 = V_{balance} - \frac{1}{2}\frac{V_R}{2^N}$$

but after feedthrough of VSET:

$$V1 = V_{balance} - V_{FT} \quad .$$

hence V1 must be increased by a constant value: $\Delta V1 = V_{FT} - \frac{1}{2}\frac{V_R}{2^N}$. Since this value is small it may be added to V1 by an additional capacitor C_Z, in the array which is dedicated strictly to offset and feedthrough cancellation. The voltage applied to the bottom plate of C_Z may be a nominal 10 volt transition, hence the nominal value of C_Z for a 10 volt reference is:

$$C_Z = C_T\left(\frac{V_{FT} - \frac{1}{2}\frac{V_R}{2^N}}{10 - V_{FT} - \frac{1}{2}\frac{V_R}{2^N}}\right).$$

The offset and feedthrough cancellation scheme just described is illustrated in Figure 6.4. It is not necessary for V1 to go to $-V_{IN}$, as previously conceptualized in Chapter 3, unless the converter were operating to accept bipolar voltage inputs. However, for positive input voltages a test upon $-V_{IN}$ yields no information. Hence the hold mode ends with $V1 \simeq V_{balance}$.

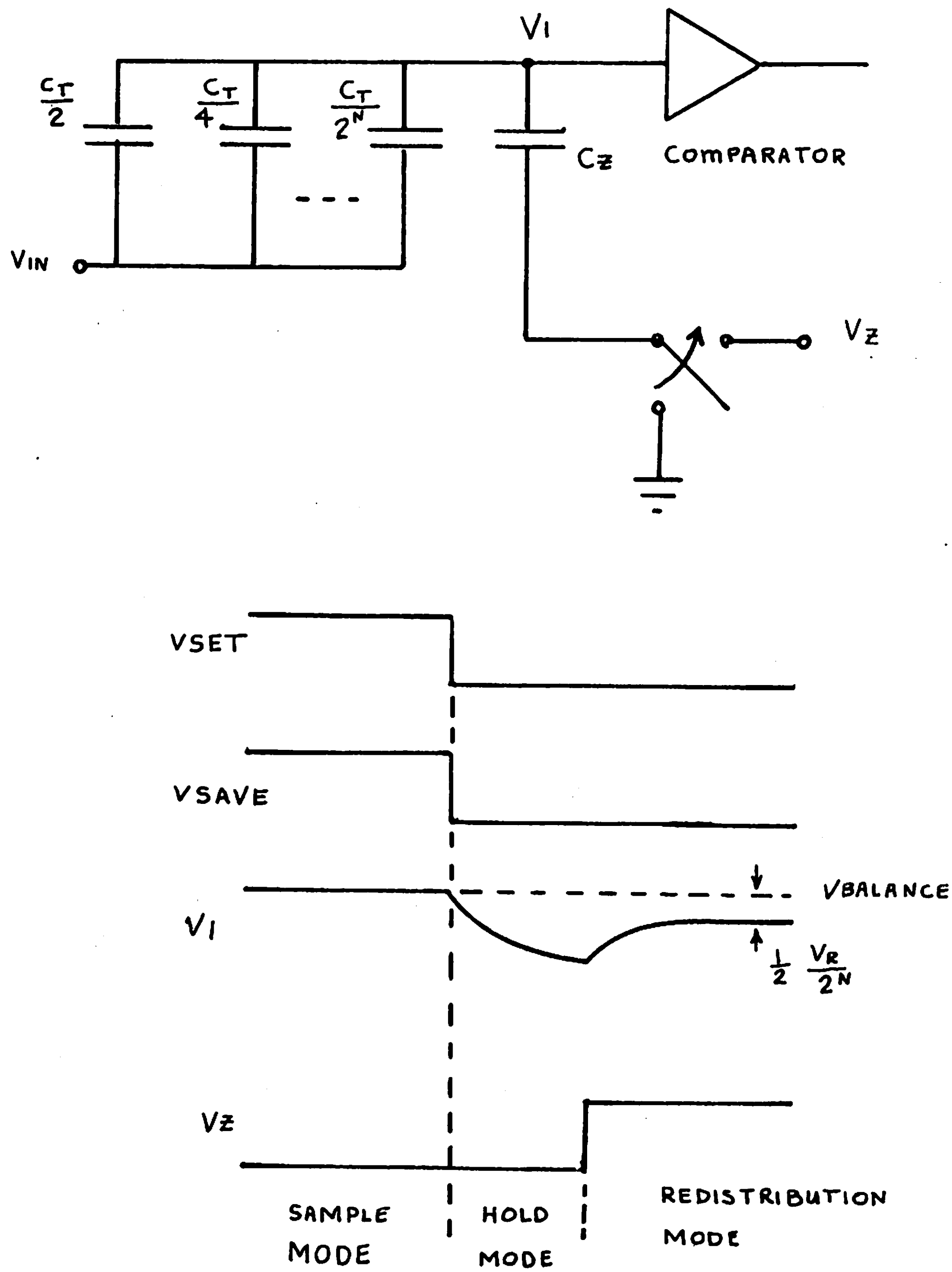


Figure 6.4: Cancellation of intrinsic offset and feedthrough.

The redistribution mode begins with the simultaneous occurrence of 3 events: intrinsic offset cancellation as just described; the switching of the bottom plate of the largest capacitor V512 to V_R; and the switching of all other capacitors to ground. In this way the charge-redistribution necessary to test the MSB is performed initially. After this V_x settles to V_x'' - some amplified value of V1. This value is stored at V5 by a pulse on the gate of M9. M9 turns off in order to simulate the same feed-through effects at V5 as M10 had upon V4. The difference signal (V5-V4) is proportional to $[V_x'' - V_x']$ where $V_x' = \left(V_{balance} - \frac{1}{2}\frac{V_R}{2^N} A1\ A2\right)$. During this phase of the redistribution cycle V_{DOWN} and V_{LATCH} are both high making M13 and M14 "off" but M7 and M8 "ON". V_B is a fixed d.c. voltage adjusted so that V4 biases M11 in the saturated region. During this time M5 and M4 act as saturated loads for M11 and M12 respectively. Hence these 4 devices comprise a third stage A3 having gain A3. However, this stage is actually a difference amplifier since (V8-V9) = A3(V5-V4). In contrast stages A1 and A2 form a single-input and single-output amplifier with gain A1 × A2. The inputs to A1 - A2 are time-multiplexed. This has the great advantage that offsets or mismatches in these two stages do not adversely affect the comparator since these affects are common to both V4 and V5. Only the polarity of $(V_x'' - V_x')$ is of importance. The signal at V8, V9 is:

$$(V8-V9) = G_{A1} \times G_{A2} \times G_{A3}(V_x'' - V_x')$$

$$\simeq 200\ (V_x'' - V_x').$$

Since the minimum signal is $\frac{V_R}{2^N} = 10$ mv, the minimum value of (V8-V9) is 2 volts. M4, M5, M13 and M14 form a bistable latch circuit. In the next clock cycle V_{DOWN} goes low and M14 and M13 turn "ON" and the latch partially regenerates. That is, (V8-V9) is increased in value by this operation. The regeneration becomes total on the next clock cycle when V_{LATCH} also goes low turning off the transmission gates M7 and M8. The latch ends in a d.c. state in which (V8-V9) is approximately equal to $(V_{DD}-V_{threshold})$ the full available logic swing. The gain of the entire comparator is therefore infinite. The outputs COMP and $\overline{\text{COMP}}$ are buffered out of the comparator and into the external digital system. The redistribution mode continues with similar tests until all 10 bits have been tested. The A/D conversion is then complete.

6.4 Logic Circuit Design

The logic circuit performs sequencing, control and data storage functions that are necessary to support the experimental chip as a complete ADC. A detailed system diagram illustrating the functional blocks of the logic circuit is shown in Figure B.1 of Appendix B. A state sequencer containing registers and a counter drives the "Capacitor Signal Generator" and the "Switch Signal Generator." The signals destined for the experimental I.C. require level shifting and buffering to convert them from TTL to MOS logic levels. This function is performed for the capacitor signals by the "CMOS Switches". MOS switches were not placed on chip because the number of bonding pads needed for gate signals would have been excessive. The "Switch Signal Generator" provides the timing signals for the comparator. The final configuration of capacitor signals is clocked into an output buffer and then may be channeled to a suitable display. The details of the logic system including the timing diagrams, state table, and circuit schematics are shown in Appendix B.

6.5 Summary

The design philosophy for both the experimental I.C. and the digital logic circuit was that a reasonable effort should be devoted to design flexibility. That is, the chip design should permit numerous methods of recovery in the case of partial circuit failure. This is evident by the number of d.c. levels that are externally adjustable off chip, and the optional outputs and internal bonding pads. In addition the logic system is designed with adjustable width timing signals. Although this philosophy adds some complexity it further enhances the capability for experimental evaluation of the new technique as well as aiding the investigation of errors.

CHAPTER 7

EXPERIMENTAL RESULTS

7.1 Introduction

The experimental results are discussed in 4 sections of this chapter. In section 7.2 the measured data taken for the first experimental I.C., IC1, is examined and the largest sources of error are identified. The subsequent design modifications in both the fabrication schedule and the layout that are required to correct the error are described in sections 7.3 and 7.4. The results of measurements taken from the second experimental IC, IC2, are analyzed in section 7.5. Both IC1 and IC2 have the same circuit schematic as illustrated in Figure 6.3 hence both are 10-bit RADCAP type of ADCs. However, they do not have the same circuit layout or topological geometric configuration. N-channel aluminum gate technology was chosen over p-channel technology due to the higher surface concentration of diffusant resulting in a lower voltage coefficient of capacitance and also due to the higher mobility of electrons over holes. The N-MOS metal gate fabrication schedule is given in Appendix C [36].

7.2 Experimental Results of IC1

7.2.1 Design of Circuit Layout for IC1

The layout of an integrated circuit is the plane geometrical configuration of the topology of various regions by which a circuit is realized. For IC1 the layout was designed so that 0.75 μ undercut could be tolerated. This requires the reproduction and parallel connection of smaller plates of identical geometry as discussed in section 4.9.3. In Appendix D an equation is derived for nonlinearity referenced to 10-bit resolution as a function of capacitor reproduction size and uniform undercut. This equation is plotted in Figure D.1. From this analysis for a VATCAP array it is concluded that a duplication size of 32 will retain 10-bit ratio accuracy for uniform undercut up to 0.75 μ. If, for example, capacitors in VATCAP are designated C512, C256, C128,...C2, C13 from largest to smallest, then a duplication of size 32 means that the largest single square capacitor plate has the same dimensions as C32 and that C64 is composed of 2 plates exactly identical in size to C32 but connected in parallel. Similarly, C512 is the parallel connection of 16 plates each identical to C32 in dimensions. This may be seen in the die photo of IC1 which is shown in Figure 7.1. Another layout design feature of IC1 includes common centroid geometry but only for the largest capacitor C512. This may also be seen in the die photo. The absence of available data on oxide growth uniformity of the particular furnace involved in capacitor oxide growth led to an initial assumption that over the region of 1 die the uniformity would be sufficiently good that common centroid geometry is needed only for the largest capacitor. Other sources of error that could not be accurately evaluated prior to wafer fabrication were not included in design considerations. The die photo shown in Figure 7.1 is about 70 mils square. The comparator is located below the capacitor array and next to the test devices. The layout for IC1 is realized on the silicon wafer with glass-emulsion photomasks containing a 400 times reduction of rubilith artwork. This large reduction size was chosen to minimize the effect of linear dimensional uncertainty involved in cutting the rubilith.

7.2.2 Threshold Voltage for the N-MOS Device in IC1

The N-MOS devices require a positive threshold voltage V_{TH} for proper

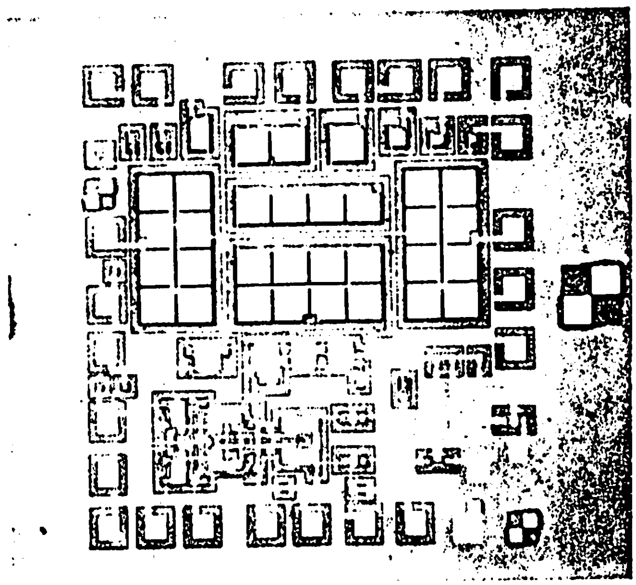


Figure 7.1: Die photo of IC1.

on/off switching action. A nominal threshold voltage of 2v is desirable because the voltage difference at the latch (V8-V9) can be no larger than V_{TH} from section 6.3, and 2V is a desired value for (V8-V9). There are no other particular device requirements although a large conduction coefficient or gain parameter is advantageous for faster switching. Before sintering the measured N-MOS threshold voltage V_{TH} for devices on IC1 was about 0 volts which could be increased to about 1 volt with - 10v substrate bias. After sintering V_{TH} became - .5V and + .5V for zero and - 10V substrate voltage respectively. The reduction in V_{TH} due to sintering is not clearly understood. However this may be caused by a negative ionic charge associated with moisture between the aluminum and the thin oxide which evaporates upon heating or to mobile positive charge contamination in the metal which migrates to the silicon-silicon dioxide surface during sintering. In any case the largest device threshold voltage is desirable hence the silicon wafer was not sintered. In contrast to sintered metal the unsintered aluminum may be easily removed which facilitates re-use of the wafer.

7.2.3 Sources of Error for IC1 due to Fabrication Procedures

Several sources of error were identified during the wafer fabrication process and subsequent wafer probe. The tri-chloroethylene (TCE) which was used during thin oxide growth resulted in visible pitting of the oxide when the ratio of TCE flow to dry O_2 flow was too large [37] [38]. It was also noted that the etch rate of the aluminum appeared to be faster in areas of the wafer for which a lesser amount of metal was being removed by the etchant per unit area. This was hypothesized to be an etchant saturation effect. Furthermore, the uniform undercut varied between 1 μ to 3 μ on the first few wafers produced. This was substantially larger than the expected undercut of 0.75 μ. Another source of error, observed at the probe station, was a high occurrence of low impedance capacitors. Upon close inspection the cause was determined to be pinholes in the thin oxide which created a low resistance between capacitor plates. In addition to this the capacitance value for a particular size of capacitor varied about 10% across the wafer. This may be modeled by a linear oxide gradient parameter of 100 ppm/mil. The subsequent investigation of the furnace associated with thin oxide growth revealed a defective furnace zone giving rise to a severe temperature gradient in the furnace tube. The sources of error just discussed were considerably larger in magnitude than expected and required fabrication schedule corrections before high precision matching accuracy could be achieved.

7.2.4 Data From Capacitance Bridge Measurements for IC1

IC1 was designed so that each capacitor in the array could be probed using a three-terminal measurement technique which nulled out the effects of stray and parasitic capacitance. Such data was taken for 21 arrays on one wafer and is plotted in Figure 7.2. In this plot the vertical axis represents the error in the binary weight of each capacitor from its ideal binary value. The horizontal axis is broken into regions corresponding to each capacitor. The horizontal segment which bisects each error band is the mean value of error for each capacitor. The arrow defines the standard deviation in the error distribution about the mean value. Several conclusions may be deduced from this data. First there exists a large systematic error in capacitor ratios since the mean values

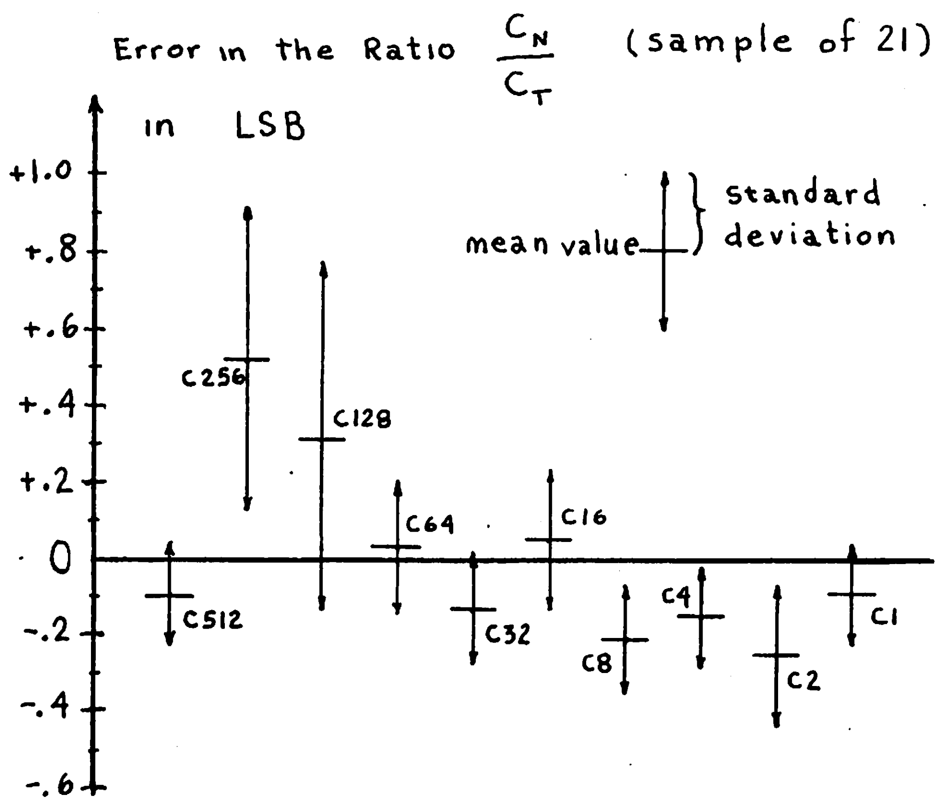


Figure 7.2: The mean value and standard deviations in capacitor ratios for IC1.

in some cases are significantly different from zero. Second the distribution of errors about each mean value is large and is probably due to a combination of factors which vary locally from die to die. There is no clear explanation for the small standard deviation of C512 compared to that for C256 and C128 although the common centroid for C512 could cause a reduced dependence of C512 upon variations in the direction of the oxide gradient.

An attempt is made to identify the sources of error leading to the particular distribution of mean value errors in Figure 7.3. The direction of decreasing oxide thickness is shown in Figure 7.3(a) superimposed upon the capacitor layout for IC1. The approximate direction of the first order gradient was determined by probing capacitors on each die and maping the change in their absolute values. The effect of this oxide gradient introduces a capacitor ratio error for each capacitor as shown qualitatively in Figure 7.3(b). Since C512 has common centroid geometry to some extent its mean error value is estimated to be less than that for C256 as shown. There are two additional sources of error which are computed analytically and also plotted. Figure 7.3(c) shows the approximate error distribution for 1 μ uniform undercut. An additional source of error for IC1 was the neglect of overlap capacitance to to interconnect and alignment stubs passing over the thicker oxide over the N^+ regions. It may be seen from the die photo in Figure 7.1 that neither half of C512 requires an alignment stub as do the other capacitors. The ratio error due to this additional capacitance is plotted in Figure 7.3(d). A composite summation of these 3 sources of error when weighted properly could give the observed mean value error distribution which is reproduced in Figure 7.3(e) for comparative analysis. One important conclusion from this study is

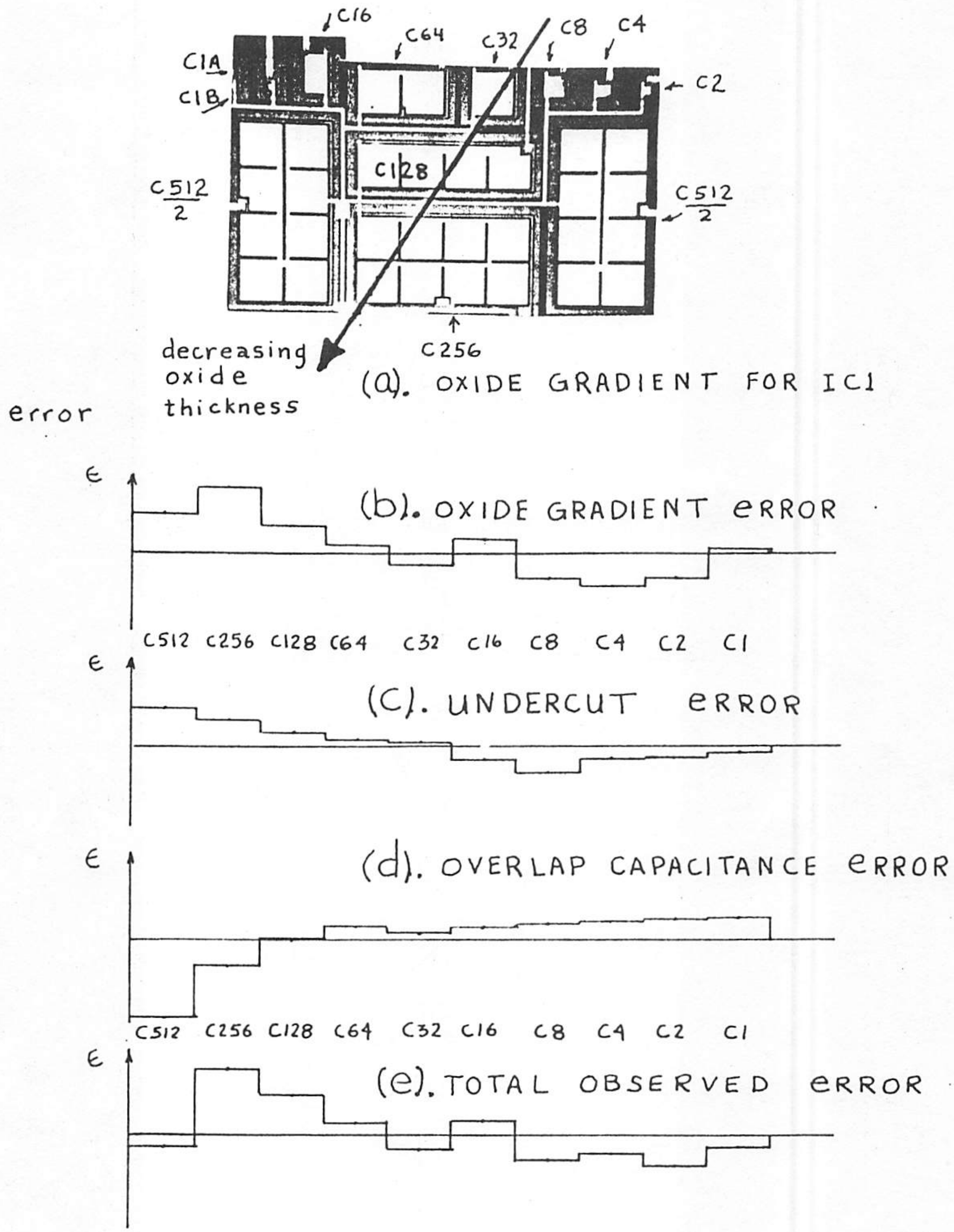


Figure 7.3: Sources of error for IC1.

that the oxide gradient effects must be considerably larger than the other 2 error sources in order to give the observed error distribution in Figure 7.3(e). In addition the error distribution including the systematic error is estimated from Figure 7.2 to be $\pm$ 1 LSB hence the nominal matching accuracy is sufficient for conversion linearity only for 8 or 9 bits of resolution.

7.2.5 Operation of IC1 in the ADC System

Several IC1 circuits were packaged and operated in the ADC system of Figure 6.1. It was verified by measurement that the system performed properly in a logical sense and the comparator offset cancellation functioned correctly. The entire system provided 10-bit resolution for input voltages from 0 to 10 volts.

The linearity of the RADCAP system was determined by measuring the transition voltage at which each individual bit turns on. The deviations in transition voltages were then computed and from these the worst case deviation (WCD) from linearity was determined.

The performance of 3 test circuits was evaluated and compared with their corresponding capacitance bridge measurements. These circuits were sintered to promote stronger contact pads for ultrasonic bonding. Figure 7.4 illustrates a comparison of the distribution of the mean error in capacitor ratios from bridge data to the distribution of mean errors derived by transition voltage measurement. This study indicates the relative reliability of the bridge data as a measure of the conversion nonlinearity. It further illustrates that capacitor ratio error accounts for nearly all of the observed deviation from linearity.

The performance of several circuits was evaluated and the average

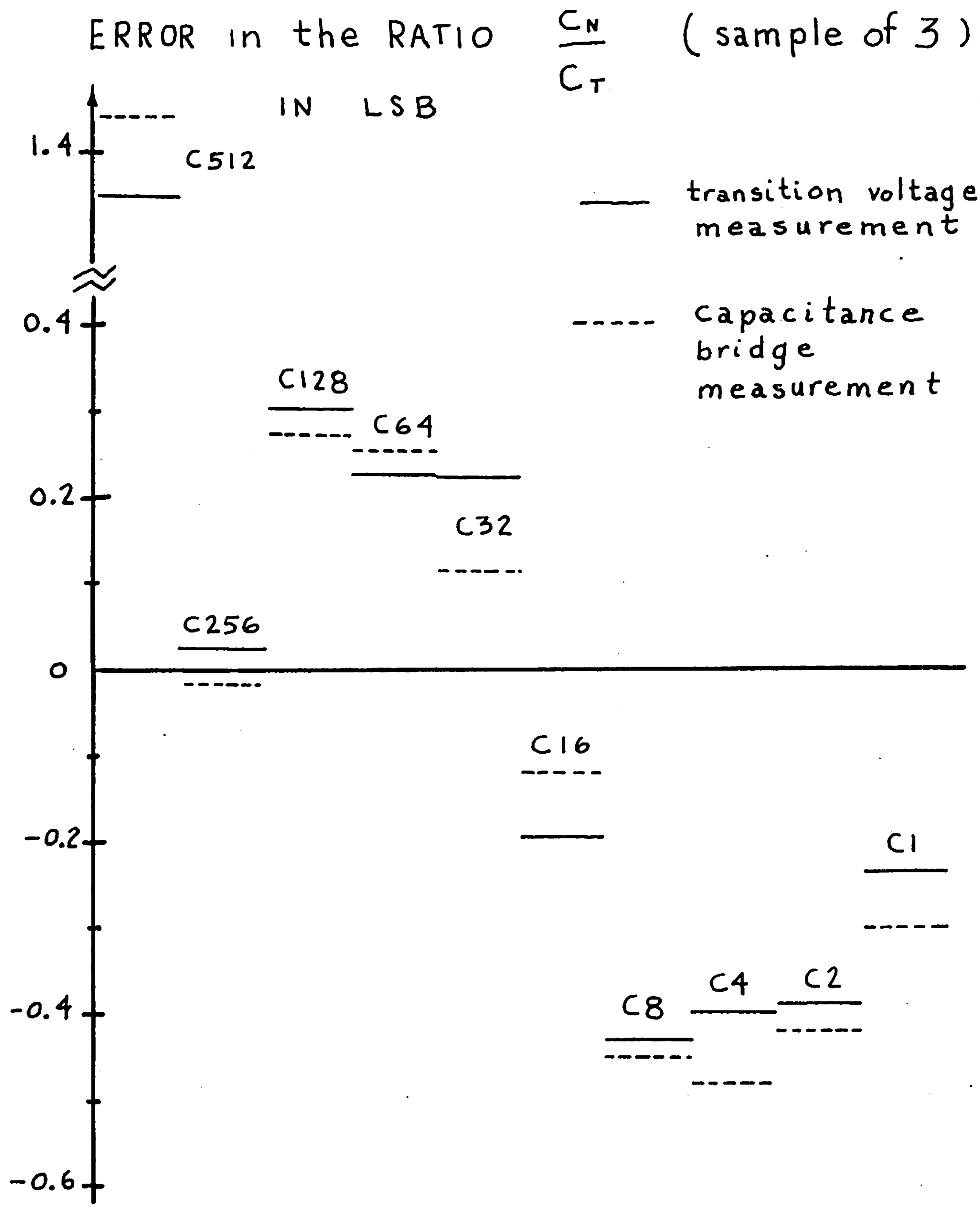


Figure 7.4: A comparison of mean error distributions from capacitance bridge and transition voltage measurements.

WCD from linearity was found to be $\pm$ 2 LSB for 10 bits of resolution. Conversely the maximum resolution for $\pm \frac{1}{2}$ LSB linearity was between 8 and 9 bits. It was apparent that the systematic errror (the mean error) must be eliminated from the system before further evaluation of performance data would be profitable. Analysis of the relative effects of oxide gradient and uniform undercut earlier in this chapter supports the conclusion that greater insensitivity to these two error sources by layout and fabrication modifications was required in order to achieve high yields at $\pm \frac{1}{2}$ LSB linearity for 10 bits of resolution.

7.3 Fabrication Modifications Required to Correct Defects in IC1

Several changes in the fabrication procedure were made in an effort to correct defects discovered in IC1. It was suspected that the N-MOS thresholds were low due to contamination of the vacuum chamber and associated equipment used during the filament evaporation of aluminum [39]. Therefore an electron beam evaporation system was developed and dedicated strictly to MOS in an attempt to reduce positive ion contamination and increase the device threshold voltage. In addition to this the TCE oxide growth with an initial TCE purge to avoid pitting observed with TCE oxides. Another modification intended to reduce damage to the thin oxide was a double exposure of the contact mask after shifting the working plate frame on the mask aligner. This technique insures that dust particles or spot defects will not create pinholes in the thin oxide thereby reducing the number of shorted or low impedance capacitors. In addition to this all contact windows were etched along with thin oxide regions. The advantage of this is that a much shorter etching time than before is needed to open the contact windows. This reduces the chances of photoresist lift and undercutting during this last oxide etch. A more radical modification in the fabrication schedule was required to offset the defective zone in the furnace used for thin oxide growth. The new oxide growth procedure required that the wafer lie flat on the boat rather than stand vertical. This was an effort to locate the entire wafer in approximately the same laminar flow streams and to utilize the boat itself as a constant temperature region. Furthermore the wafer was withdrawn from the furnace after each 25% of oxide growth and rotated 90° in an attempt to average the effects of temperature profile and flow stream variations upon oxide growth uniformity. The final fabrication changes involved improvements in photomasking associated with the aluminum. First the working plate for the metal mask was reproduced again with an improved focus for a sharper image. The photoresist development procedure was changed to a spray development for better resolution. Also the aluminum etching procedure was modified to include ultrasonic vibration during the etch in order to remove vapor bubbles from the wafer surface and to provide a better circulation of etchant and more uniform removal of aluminum.

7.4 Layout Modifications Required to Correct Defects in IC1

Evaluation of the measured capacitor matching data indicated that the two largest sources of error were uniform undercut and oxide gradient as discussed in section 7.2. Uniform undercut was estimated between 1 μ and 3 μ depending upon the care taken during the etch. From Appendix D the WCD due to this value of undercut is greater than .6 bit. However, if duplication size 8 were used to configure the array a maximum value of 2 μ undercut would be tolerable. Therefore this duplication size was chosen for the new layout. Another point of interest is that the array layout design for IC2 did not include in each capacitor that component due to interconnect passing over the diffused N^+ back plates. Actually this component was not neglected since from calculations it results in a ratio error distribution somewhat opposite to that caused by uniform undercut. Hence it was of interest to determine whether the undercut could be so well controlled that the ratio errors could be converged to zero by beginning with an "oversized" array. A theoretical mean error distribution including this interconnect capacitance is shown in Figure 7.5 for various values of undercut. The effect of uniform undercut upon capacitor ratio error is to increase the error in large capacitors and to decrease error for smaller capacitors. From this plot a nominal uniform undercut of 2 μ would result in minimal total error.

The viability of uniform undercut cancellation rests heavily upon the initial premise that undercut is indeed uniform at every metal edge. In Chapter 4 several mechanisms were proposed which could lead to regional non uniformities in etch rate. One mechanism, etchant saturation, was highly suspect since manifestations of this kind of effect were observed on several occasions. Therefore a layout modification was developed to increase the probability that the etch rate would tend to be uniform at the capacitor metal plate edges. This scheme involved floating metal strips or other metal lines placed the same distance from every capacitor plate edge so that the etchant concentration would tend to be equal at all capacitor plate edges. It can be shown that these floating strips have no adverse affect upon conversion accuracy since they merely constitute a stray capacitance to ground.

It was confirmed from measured data in section 7.2 that thin oxide

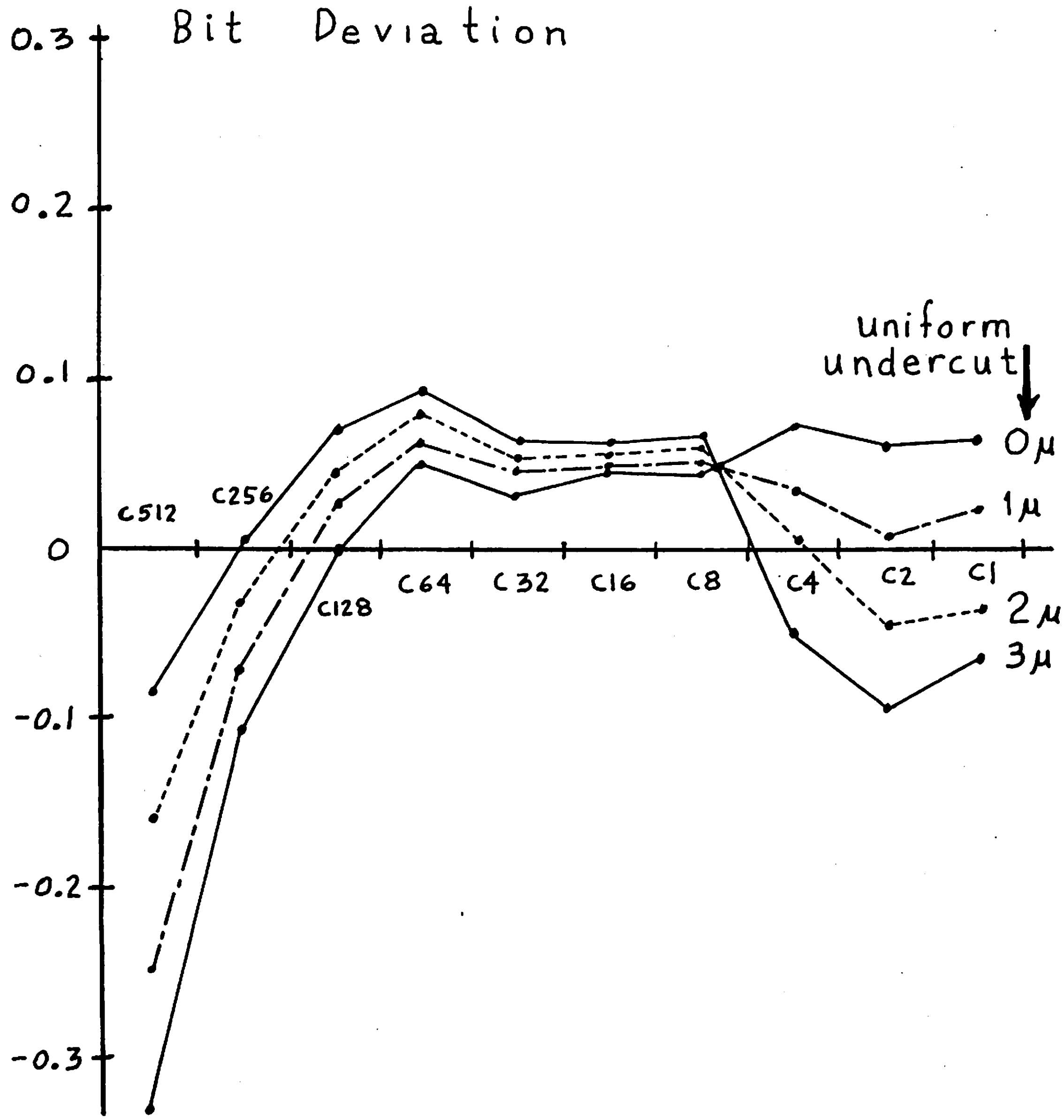


Figure 7.5: Capacitor ratio error distribution as a function of uniform undercut when interconnect capacitance is included.

gradients resulted in large ratio errors. In effect, a more extensive common centroid geometry was needed to reduce the effects of these gradients. A common centroid geometry was therefore chosen for the 5 largest capacitors since the errors in the smaller capacitors were small in absolute value.

Implementation of the layout modifications just discussed resulted in experimental integrated circuit #2 (IC2) shown in the die photo in Figure 7.6. The dimensions of the active area containing the capacitor array is 75 × 58 mils. The common centroid geometry and capacitor reproduction size can be clearly seen along with the floating metal strips. Also incorporated into this design was the existence for each capacitor of an equal number of 90° and 270° corners. This provided a first order cancellation of corner rounding effects.

An additional layout modification was a 200X reduction of the original artwork rather than 400X as used for IC1. While this increased the effect of rubylith cutting errors it had the added advantage of processing independence from commercial reduction facilities. This philosophy was consistent with the expectation that a systematic error would indeed be found but would be subsequently removed by a trim of the artwork and generation of a new working plate.

7.5 Experimental Results for IC2

This section contains the measured results for the second experimental integrated circuit. The N-MOS device parameters are examined in section 7.5.1 and the discovery and subsequent elimination of systematic capacitor ratio error is discussed in the next two sections. In sections 7.5.4 and 7.5.5 evaluations are made of the performance of the RADCAP ADC

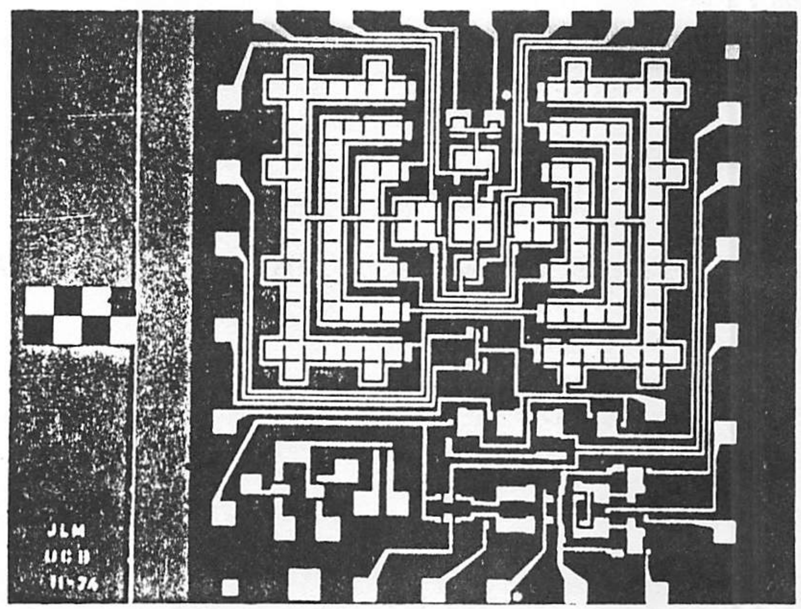


Figure 7.6: Die photo of IC2.

system. In the final section the analysis of remaining error is discussed.

7.5.1 N-MOS Device Parameters

The N-MOS device parameters were determined using test structures on IC2. These parameters apply to a p-type (100) substrate having a bulk doping of $N_B = 3 \times 10^{15}$ which was chosen for experimental work. The curves shown in Figure 7.7 illustrates the variations in drain current versus gate-source voltage, I_{DS} vs V_{GS}, for the case in which the drain and gate are connected, hence $V_{GS} = V_{DS}$. From left to right the curves are for bulk-source voltages (V_{BS}) of 0, - 2.5, - 5, - 10, and - 15 volts. Measured data points are plotted for convenience as shown in Figure 7.8. The vertical axis is scaled to vary as $\sqrt{I_{DS}}$. From this plot the extrapolated zero bias threshold voltages (for $V_{BS} = 0$) is obtained from the intersection of the V_{GS} axis and the best linear approximation for the set of data points: $V_{TO} = .17$ volt. The slope of this line also gives the conduction factor $K = \frac{W}{L}\frac{\mu}{2}\frac{\epsilon_{ox}}{t_{ox}} = \frac{I_D}{(V_{GS}-V_{TO})^2}$ and $K = 120\ \frac{\mu A}{V^2}$. Using $\frac{W}{L} = 9.4$ for the particular device and $\frac{\epsilon_{ox}}{t_{ox}} = 3.4 \times 10^{-8}\ \frac{F}{cm^2}$ the experimental electron mobility is $\mu = 750\ \frac{cm^2}{V\text{-}s}$. This value is somewhat lower, as expected, than the theoretical value of $1350\ \frac{cm^2}{V\text{-}s}$. In addition the best curve fit for V_T(effective) vs V_{BS} is for $N_B = 1.5 \times 10^{15}$. This is explained by a diffusion mechanism at the silicon surface called boron depletion which reduces the dopant concentration about 50% in this case [40]. A plot of V_T (effective) as a function of V_{BS} is shown in Figure 7.9. The slope of this plot is equal to γ, the bulk threshold parameter. From this data $\gamma = 0.657\ V^{\frac{1}{2}}$ which agrees with the theoretical value of $0.66\ V^{\frac{1}{2}}$ for a bulk surface concentration of 1.5×10^{15}. Figure 7.10 is

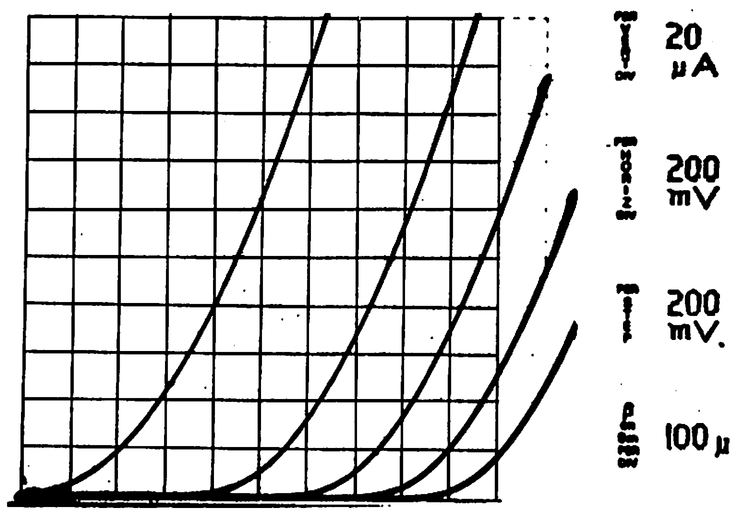


Figure 7.7: The N-MOS drain current characteristics for $V_{DS} = V_{GS}$ and $V_{BS} = 0$, -2.5, -5, -10, and -15 volts.

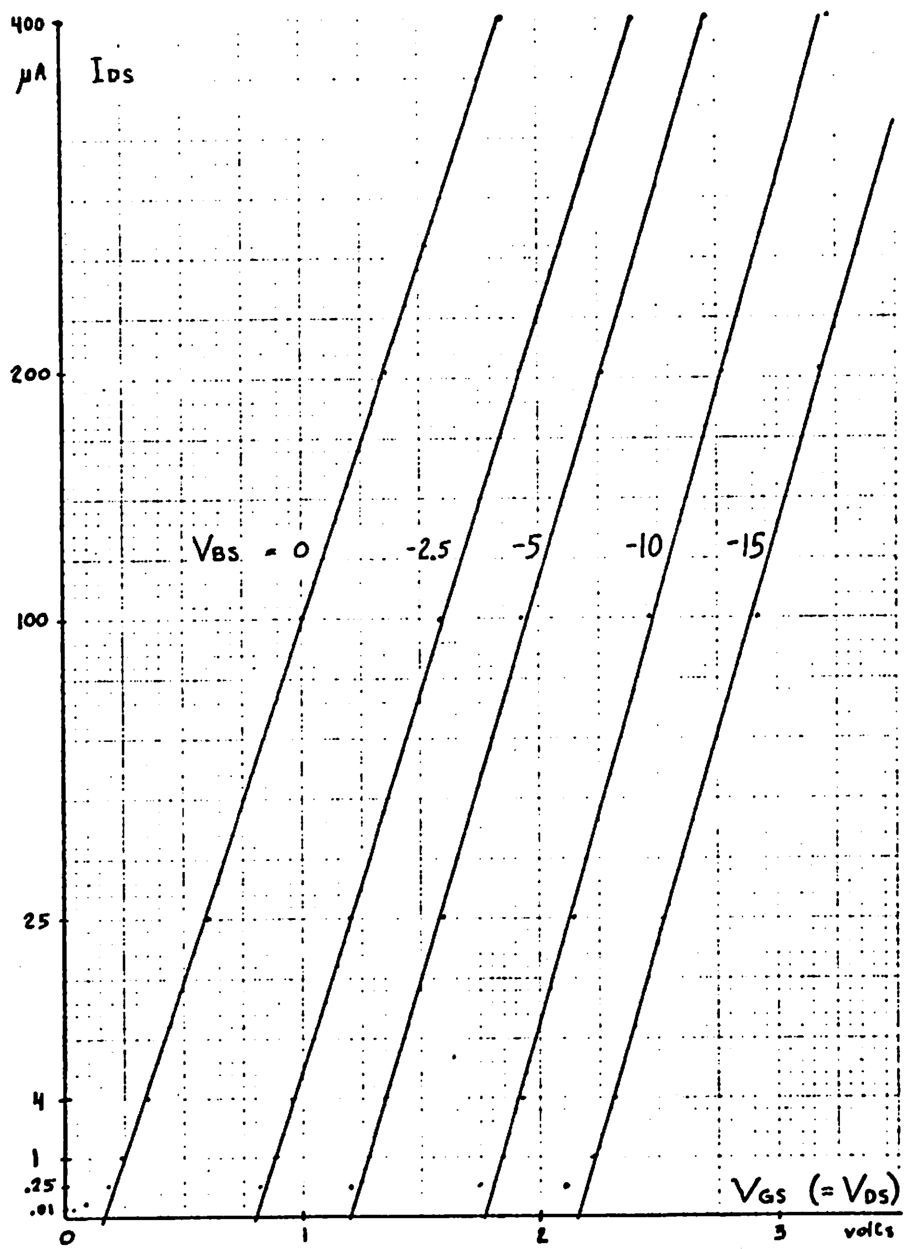


Figure 7.8: $\sqrt{I_{DS}}$ versus V_{GS} and the extrapolation of threshold voltage.

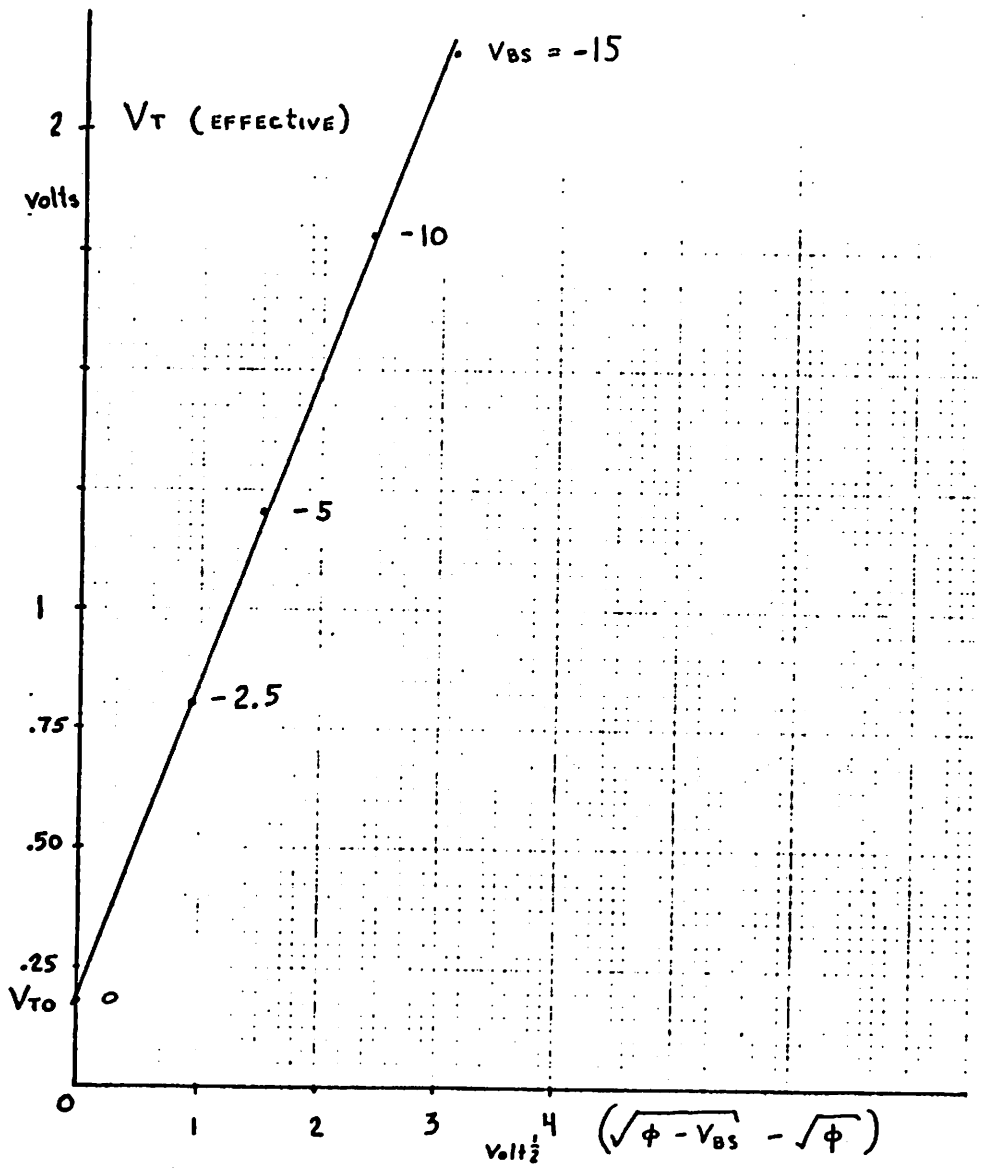


Figure 7.9: The determination of V_{TO} and the bulk parameter γ.

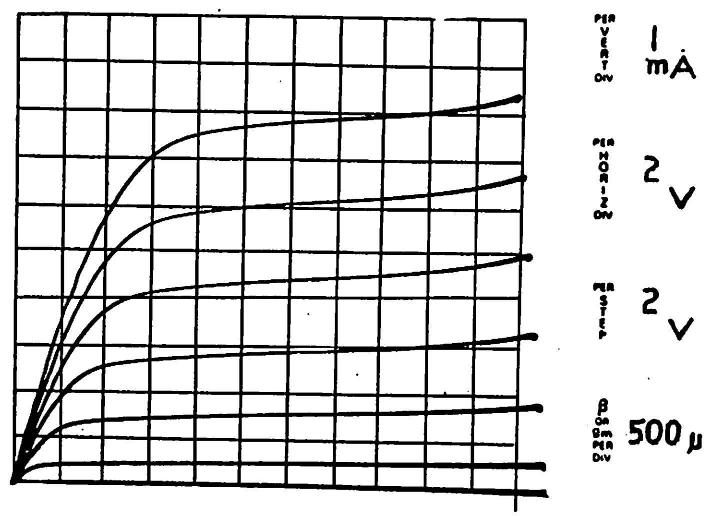


Figure 7.10: N-MOS drain current characteristics.

a photograph of the drain current characteristics I_{DS} vs V_{DS}. From this graph the channel length modulation parameter $\lambda = \frac{.01}{\text{volt}}$. This corresponds to an Early voltage of $\frac{1}{\lambda}$. In conclusion adequate device threshold voltages were realized by improved metal evaporation techniques. The N-MOS devices were enhancement type since the threshold voltage with $V_{BS} = 0$ is positive. Threshold voltages greater than 2 volts could be obtained with adequate body bias.

7.5.2 Measurement of Systematic Error in Capacitor Ratios

Nine capacitor arrays were probed and capacitor values measured. The capacitor ratios were computed and plotted in Figure 7.11. This plot illustrates the distribution in capacitor ratio errors. For each capacitor the mean error (m_e) is shown together with the standard deviation (σ_e) in errors about the mean. A large systematic error is indicated by the fact that m_e is significantly different from zero. The shape of the mean error distribution fits closely the expected error for oversized capacitors previously shown in Figure 7.11, however the value of the error requires that capacitors be oversized by an amount equal to at least twice the interconnect capacitance. This systematic error could have been caused by the photolithographic distortion in producing the working plates. This process involves a first reduction, a second reducion and a contact print. However, it is more likely that the thin oxide windows were actually larger due to uniform undercut during the oxide etch preceding thin oxide growth. In this case an effectively larger value of capacitance would be due to the interconnect but the error distribution would be precisely as shown in Figure 7.5 for zero undercut but with an increase in vertical scale dimensions. This interpretation

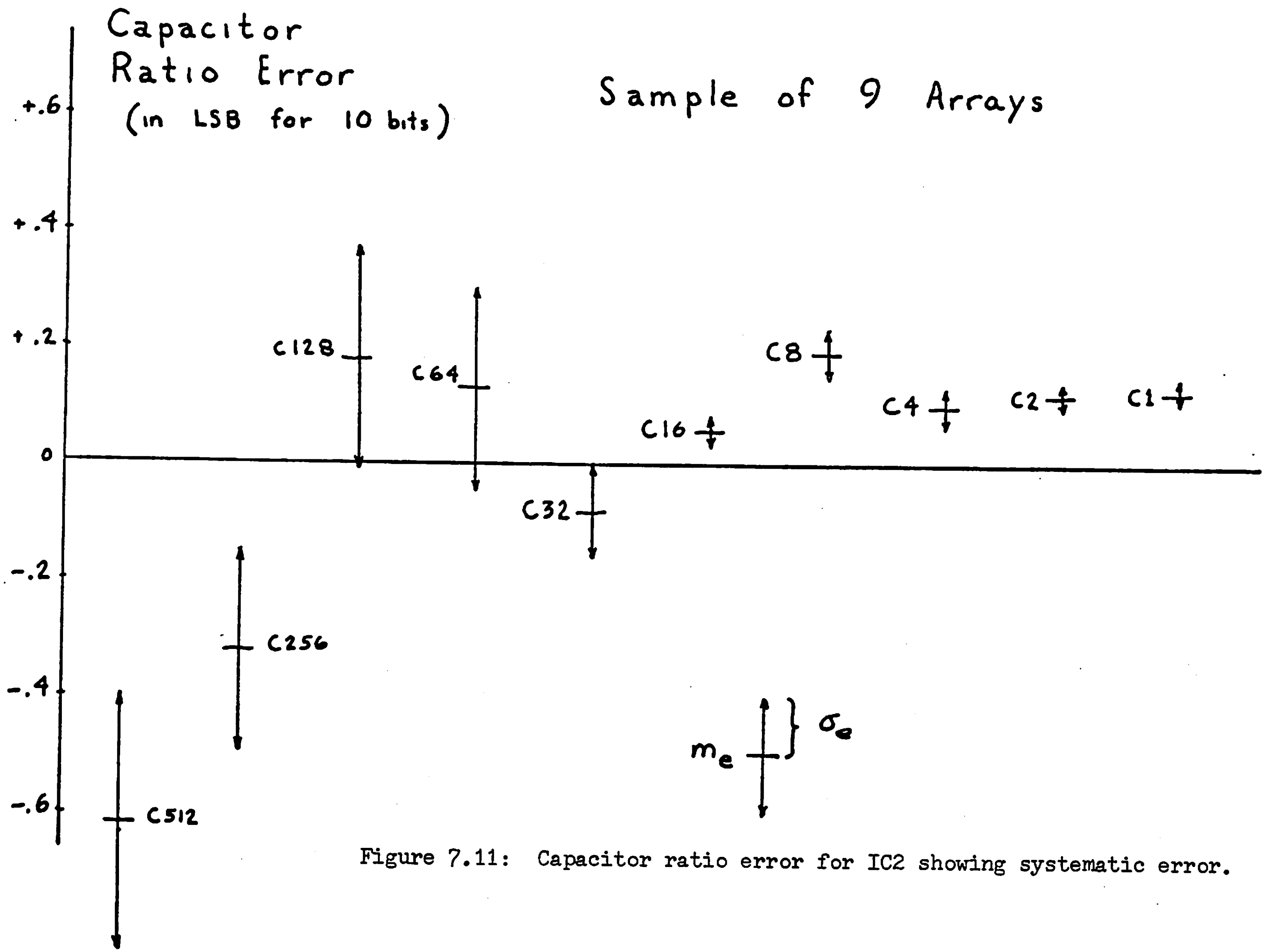


Figure 7.11: Capacitor ratio error for IC2 showing systematic error.

is consistent with the measured data shown in Figure 7.11.

7.5.3 Elimination of Systematic Error by Mask Trim

In the previous section the mean value of systematic error was determined for each capacitor by averaging a certain number of samples. The removal of this error was performed by the addition or subtraction of area to each capacitor on the rubylith defining the metal mask. Each area that was added or subtracted corresponded to the mean systematic error. A new working plate was produced and was used for the fabrication of 3 new wafers. From these a sample of 47 arrays were probed and the capacitor ratios were computed again as in section 7.5.2. The results of these measurements are shown in Figure 7.12. This plot represents the distribution of capacitor ratio errors. The mean error m_e and the standard deviation σ_e are indicated graphically for each capacitor. From this data it may be deduced that if capacitor ratio error were the only factor affecting yield the yield for $\pm \frac{1}{2}$ LSB linearity at 8, 9 and 10 bits of resolution would be 98%, 94% and 45% for this sample of 47 arrays. The ability to remove systematic error by a mask trim has been demonstrated.

The correct interpretation of the standard deviation in ratio error (σ_e) is not clear; however, it is helpful to plot these standard deviations as shown in Figure 7.13. The vertical scale is log σ_e while the horizontal scale is proportional to binary weight and consequently to the perimeter length ratio (and area ratio) for each capacitor. Assuming for the moment that this error distribution is due to a random mechanism which is uniformly distributed along the capacitor edges then the resultant random variation in capacitor ratio error could be expressed as $\sigma_e(i) = \sqrt{2^i}\, \sigma_e(0)$, for $i = 0,1,2,3,\ldots,(N-1)$, if the number of statistically averaged samples

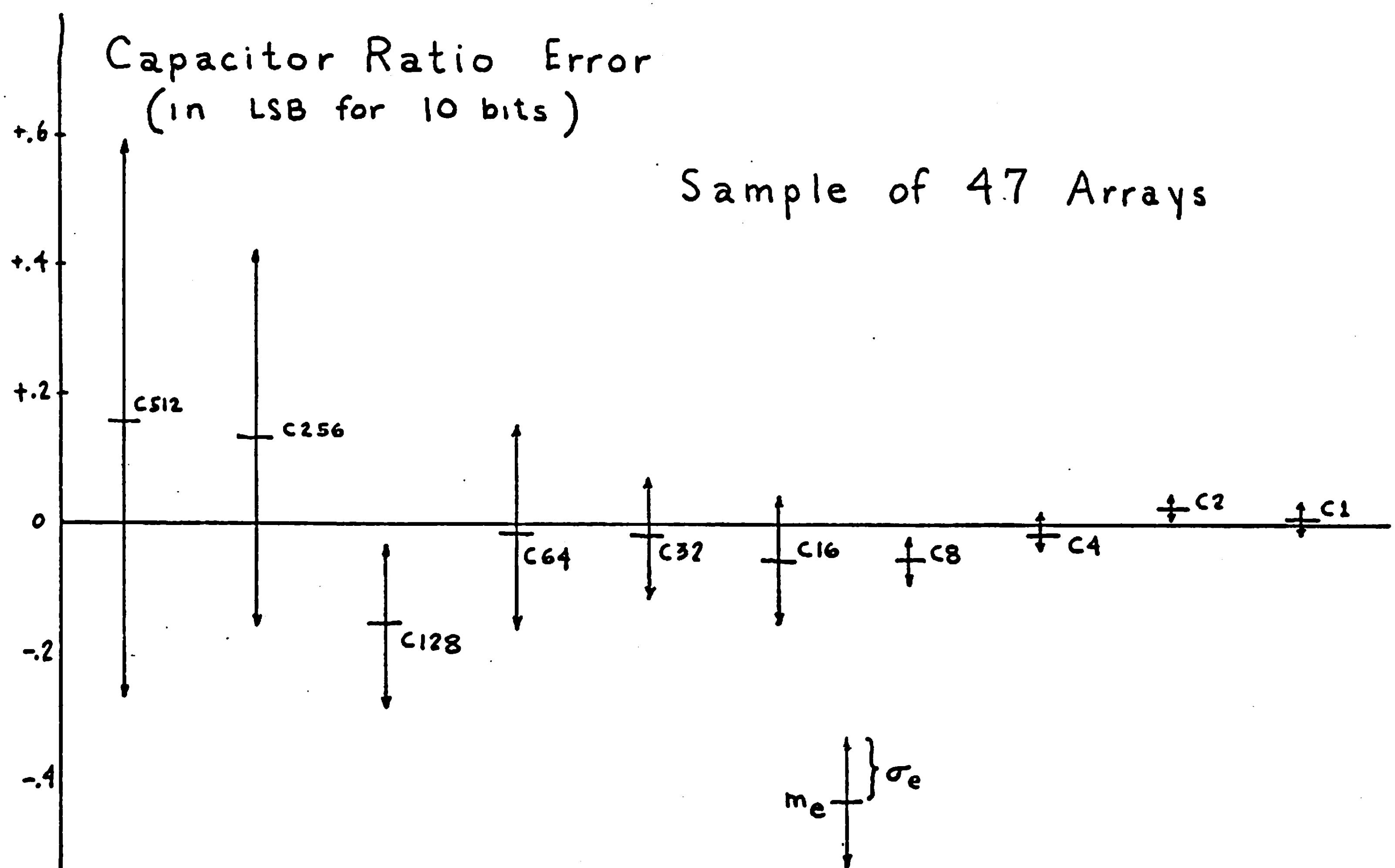


Figure 7.12: Capacitor ratio error after the first mask trim.

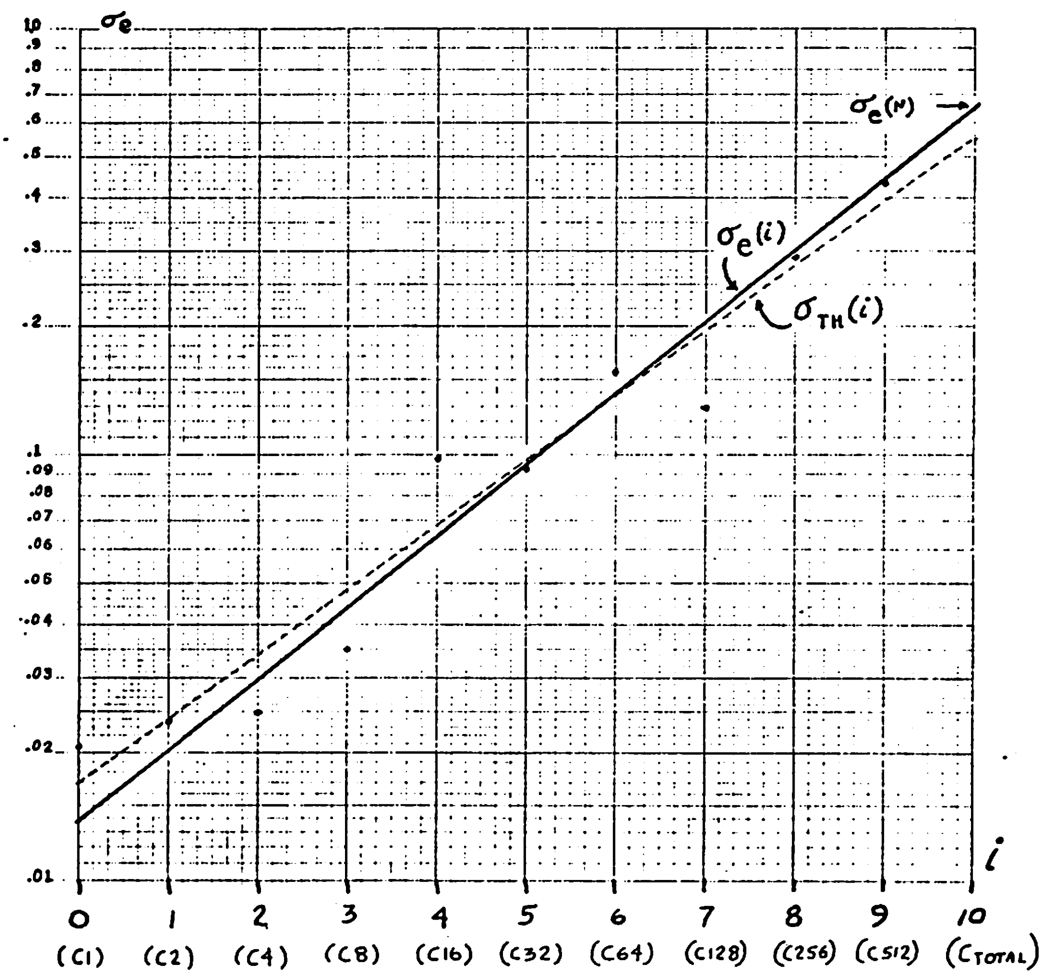


Figure 7.13: The measured standard deviation in capacitor ratio error $\sigma_e(i)$ compared with a theoretical distribution for a random variable $\sigma_{th}(i)$.

is large. This is due to the fact that the standard deviation of a random variable increases by $\sqrt{2}$ if the sample length (perimeter length in this case) is doubled. From this equation a plot of $\sigma_e(i)$ versus binary weight should theoretically have a slope of $\frac{1}{2}$. This theoretical line denoted as $\sigma_{TH}(i)$ is superimposed on the graph of $\sigma_e(i)$ versus binary weight i and a rough fit is observed. This tends to support the belief that the actual error distribution is due to the presence of a random mechanism which is uniformly distributed along the capacitor edge. It is instructive to pursue this analysis by constructing a line through the more significant data points. This line, denoted as $\sigma_e(i)$ in Figure 7.13, intersects the line corresponding to i = N (where N = 10 in this case) at $\sigma_e(N)$ = .66 LSB. This may be interpreted as the extrapolated standard deviation in the total error for all arrays which corresponds to the nonlinearity (assuming that capacitor ratio error is the most significant component of nonlinearity which is true for RADCAP). Hence it may be deduced that 68% (± one standard deviation) of all 47 arrays have a total worst case nonlinearity error $\sigma_e(N)$ no greater than ± .66 LSB (for N =10) or that 51% should have a standard deviation of ± .5 LSB. This can be correlated with the actual findings that 45% have an error less than ± .5 LSB.

In conclusion the existence of a random variable leading to ratio errors as described in this section requires proper control of geometry in order to minimize the relative effect of this error. That is, if high yield is desirable then large capacitors having long perimeters will tend to reduce the effect of the random variable if the random variable has an absolute effect upon the edges. However, small area capacitors are desirable to save chip area and reduce circuit time constants. Therefore an "optimal" geometry for the array may be defined as one for which the standard deviation in total array error is

$$\sigma_e(N) \leq \frac{1}{2^{N+1}} \text{ LSB} \quad \text{for}$$

68% yield at $\pm \frac{1}{2}$ LSB linearity. Conversely the optimal total area of the array A_{TOT} is such that

$$\frac{\sigma_{A_{TOT}}}{A_{TOT}} \leq \frac{1}{2^{N+1}}$$

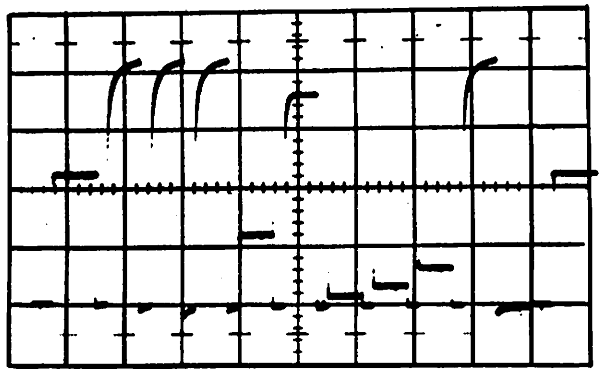
in which $\sigma_{A_{TOT}}$ is the standard deviation in the area A_{TOT} as determined by the particular fabrication technique.

7.5.4 Measurement of Transition Point Voltages

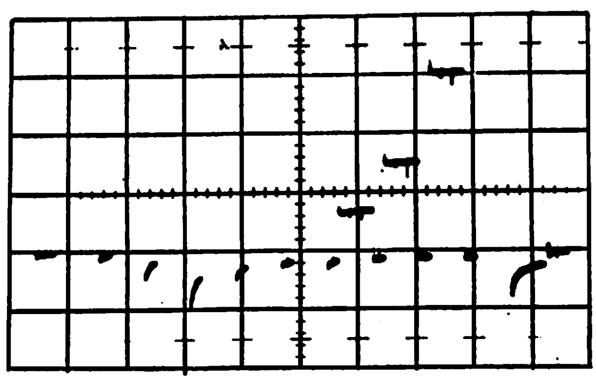
The transition point voltages were measured for IC2 as described in section 7.2.5 for IC1. With the systematic error removed high correlation was expected with the ideal transition points, and with the capacitance bridge data. However, an unexpected additional error was found that was proportionally larger for the larger capacitors. The magnitude of this error was - .5 LSB in the second largest capacitor. The presence of this error was investigated by switching the order in which capacitors were tested and by adding a redundant state between each test. For the following analysis the order of capacitor testing was:

C16 - C128 - C512 - C64 - C8 - C32 - C1 - C2 - C4 - C256

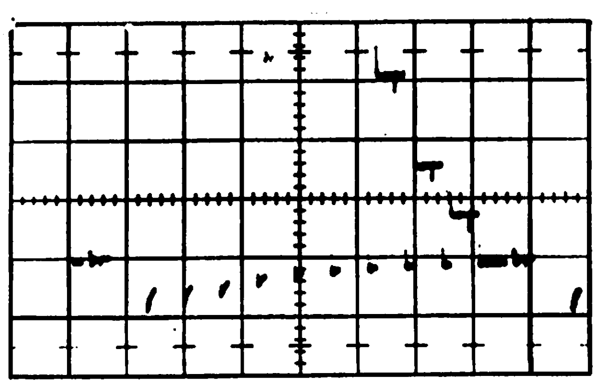
unless otherwise specified. The voltage variation at the top plate of the array is shown in Figure 7.14(a) and with an expanded vertical scale



(a). 80 mV / div ; 50 μs / div
UNORDERED ARRAY



(b). 15 mV / div ; 50 μs / div
UNORDERED ARRAY



(c). 15 mV / div ; 50 μs / div
ORDERED ARRAY

Figure 7.14: The top plate voltage waveform showing dielectric relaxation phenomenon before heat treatment.

in Figure 7.14(b). This photo shows the top plate voltage variation as each capacitor is charged to V_R and then rapidly discharged. The presence of an unexpected residual voltage after each negative going transition is evident. Figure 7.14(c) shows the voltage waveform when the correct capacitor order is restored. It was discovered that a mild heat treatment of 200°c for 5 minutes reduced the effect to a negligibly small value. The elimination of this effect for the same circuit as in Figure 7.14 is shown in Figure 7.15(a) and (b) for the unordered capacitor sequence and for the ordered capacitor sequence shown in Figure 7.15(c).

This effect was not detected in capacitance bridge data since these measurements determined only the small signal capacitance; however, the transition voltage measurements were performed with large signals. After heat treatments the transition voltage data correlated with capacitance bridge data as expected. The cause of this effect appears to have been related with moisture at the oxide surface which probably evaporated upon heating.

7.5.5 Experimental Measurement of Performance Parameters

When supported by a discrete logic system IC2 became a complete ADC simulating a RADCAP type of circuit. The performance of the complete system was evaluated by first observing the voltage waveforms to and from IC2. The signals VDOWN, VSET, COMP, and V_x are shown in Figure 7.16(a) as functions of time for an entire conversion cycle. These signals were defined in Chapter VI. The particular digital output shown here is the binary number 1010. Figure 7.16(b) shows the comparator switching waveform. From this photograph the average switching time of the latch is 125 ns. The voltage waveform V_x at the output of the high

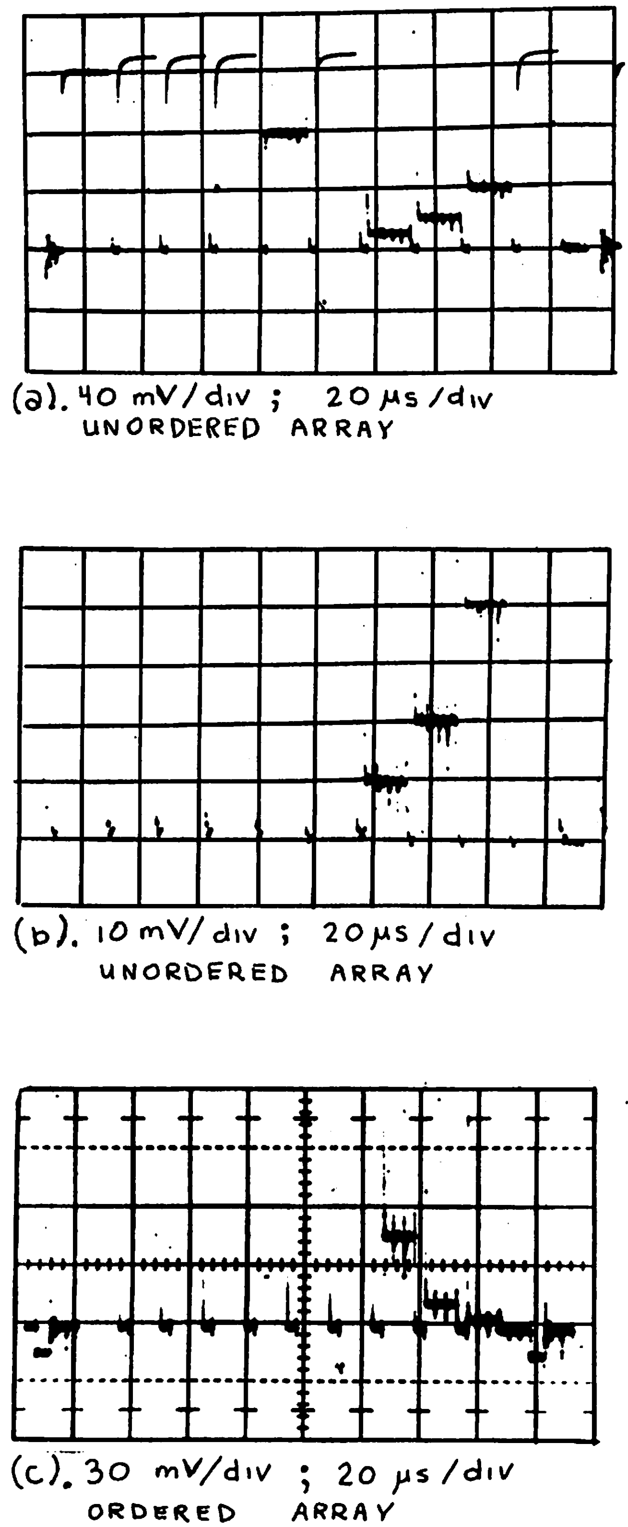


Figure 7.15: The top plate voltage waveform showing elimination of dielectric relaxation effect after heating.

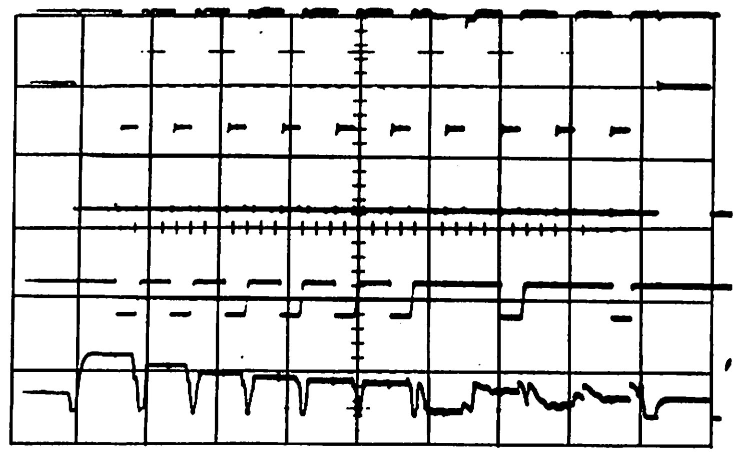


FIGURE 7.16(a): Waveforms : VDOWN (TOP), VSET, COMP, AND VX FOR 1 conversion cycle. vert. : 10 V/div ; horiz. : 5 μs/div.

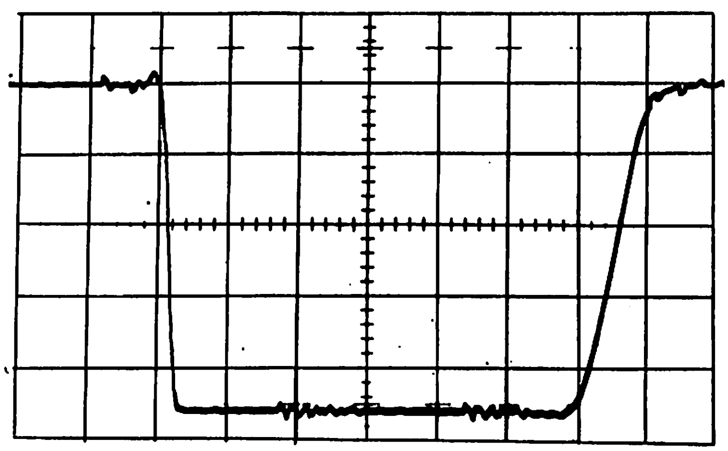


FIGURE 7.16(b) : COMPARATOR SWITCHING SIGNAL . vert. : 1 V/div ; horiz. : 250 ns/div

gain stage is illustrated in Figure 7.17(a) for a full scale 10 kHz triangle wave input. The segment of this waveform showing sample mode percharge is expanded for closer observation in Figure 7.17(b). From this photograph an adequate precharge cycle at this frequency is less than 2 μs.

The measurement of nonlinearity was performed using the experimental technique illustrated in Figure 7.18. The output of RADCAP was connected to a 12-bit DAC. Since V_{IN} was a ramp, V_{out} of the DAC was a staircase. Both of these waveforms were inputs to a differential amplifier having a sawtooth output corresponding to quantization error plus nonlinearity error. This waveform was recorded on an X-Y plotter and a typical recording is shown also in Figure 7.18. An expanded output is shown in Figure 7.19 and was used for the actual verification of RADCAP performance. This plot enabled a detailed examination of all 1024 states. Since all positive going peaks were between 0 and 10 mV and all negative going peaks were between 0 and - 10 mV the nonlinearity was less than $\pm$ .5 LSB for a resolution of 10 bits. Of 6 units tested three had $\pm \frac{1}{2}$ LSB linearity for 10 bits while the remaining units had 9 bit resolution. This roughly corresponded to the bridge data for these ICs.

The experimental IC was designed to cancel all offset error except the intrinsic error. In view of this the first transition point should be 1 LSB or 9.7 mV for V_R = 10 V. The typical value measured was about 9 mV with an average uncertainty less than $\pm$ 1 mV. Hence an offset error of less than 2 mV from the design value was observed for all 6 units tested.

An external 10 V reference was supplied to RADCAP hence the measurable gain error of less than 0.05% was due to this supply rather than to the experimental ADC.

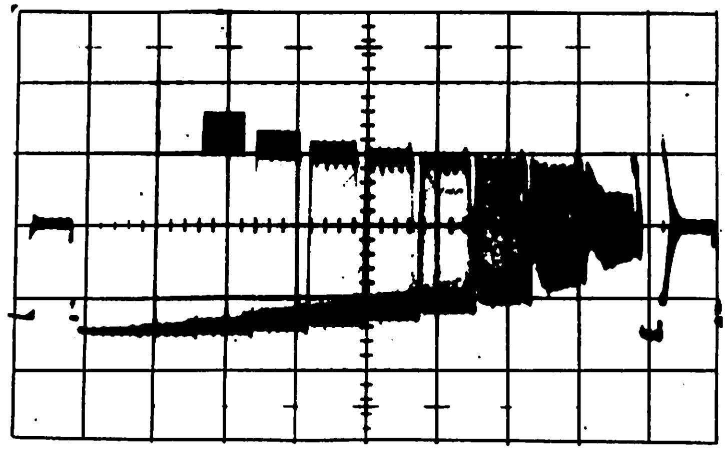


FIGURE 7.17(a): Vx vs time for V_{IN} = 10 KHz FULL SCALE RAMP ;
vert. : 2 V/div ; horiz. : 5 μs / div

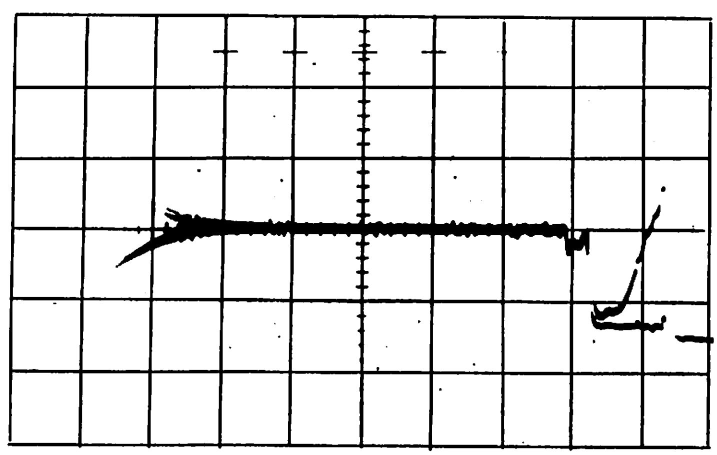


FIGURE 7.17(b): Vx during Precharge for 10 KHz full scale input.
vert. : 2 V/ div. ; horiz. : 0.5 μs/div.

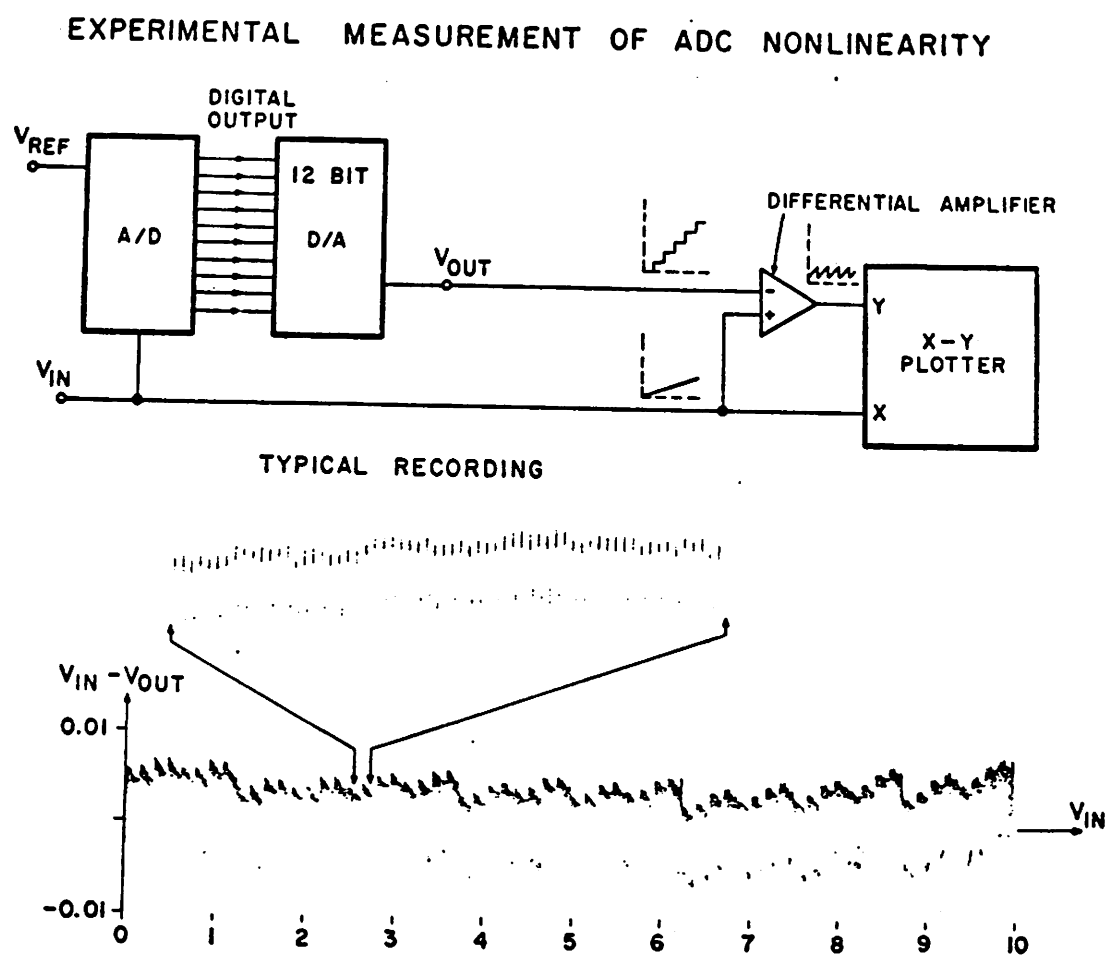


Figure 7.18: The circuit configuration and typical recording for measurement of nonlinearity.

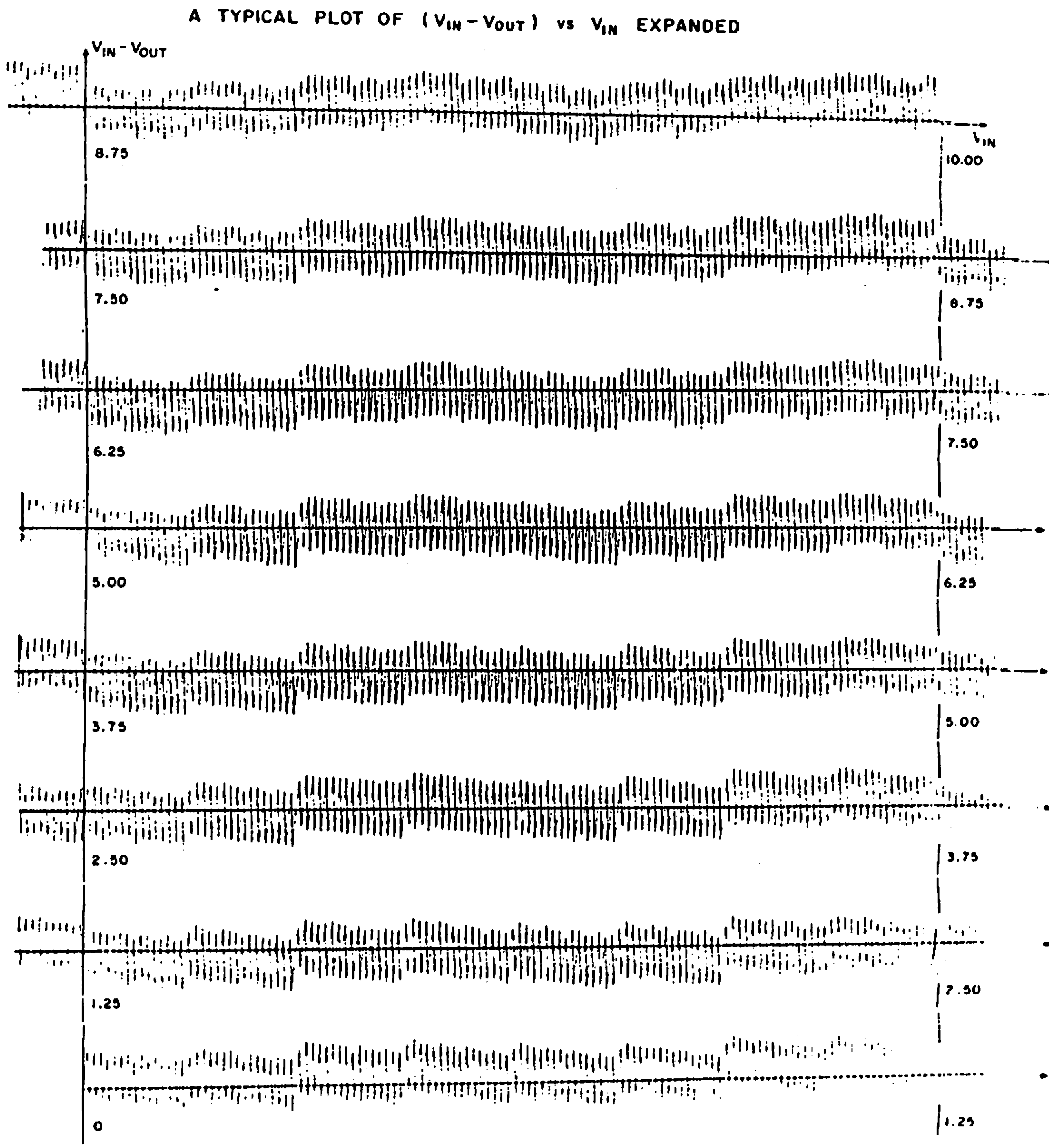


Figure 7.19: An expanded recording showing $\pm\frac{1}{2}$ LSB linearity for all 1024 digital states.

The sample mode acquisition time was the minimum precharge time required for an accurate conversion of a $\frac{1}{2}$ full-scale voltage step at the input. This was measured at 1.5 μs. The total conversion time was measured at 22.8 μs. A summary of performance specifications is given in Table 7.1.

Additional data was taken which illustrate in greater detail the performance characteristics of the RADCAP system when used to sample high frequency sine waves. Figure 7.20(a) is a photograph of 3 waveforms: V_{out} of the DAC (from Figure 7.18), V_{IN} (a 3 kHz sine wave), and V_x (the amplified waveform at the top plate of the array). This figure shows the sampling capability for sinusoids. Figures 7.20(b), (c), (d), (e), and (f) illustrate V_{out} (DAC) and V_{IN} for different frequencies. In these photos the lack of synchronization between sampling rate and V_{IN} results in an apparent continuous band for V_{out}. This is convenient since the envelope of V_{out} is visible. From these figures it is evident that V_{out} may be modeled as attenuated and phase-shifted with respect to V_{IN}. The actual recovery of the sinusoidal signal from the output V_{out}(DAC) is accomplished by a low pass filter. This was done using a 2-pole 5 kHz low pass filter. The filtered output V_{out} (5 kHz filter) was then connected to a distortion analyzer from which a THD (total harmonic distortion) of 0.35% for an input frequency range of 200 Hz to 3.5 kHz was measured. Two photographs taken in this frequency range are shown in Figure 7.21 (a) and (b).

7.5.6 Limitations on Matching Accuracy due to Random Edge Location

In section 7.5.3 a random mechanism operating uniformly along the capacitor perimeters during fabrication was postulated as the cause of

TABLE OF PERFORMANCE DATA

Resolution	10 bits
Linearity	$\pm\frac{1}{2}$ LSB
Input voltage range	0 – 10 V
Input offset voltage	2 mV
Gain error	< 0.05% (external reference)
Sample mode aquisition time	2.3 μs
Total conversion time	22.8 μs

Table 7.1: The measured performance data.

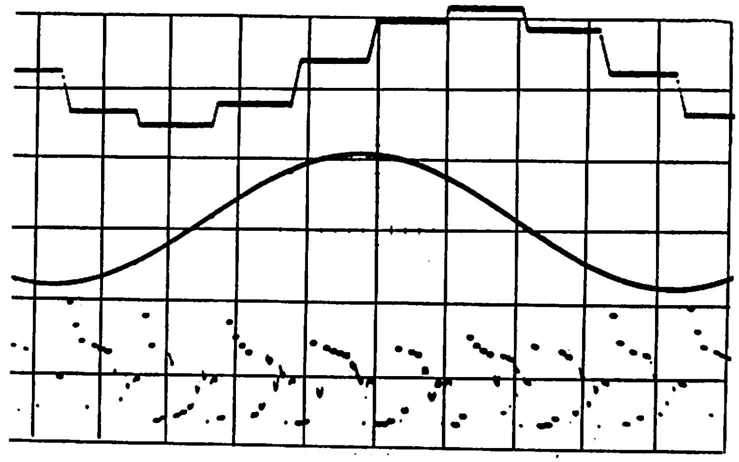


FIGURE 7.20(a): V_{OUT} (DAC) (TOP), V_{IN}, V_X.
for 3 KHz full scale sine wave input.
vert.: 5 V/div. ; horiz.: 40 μs/div.

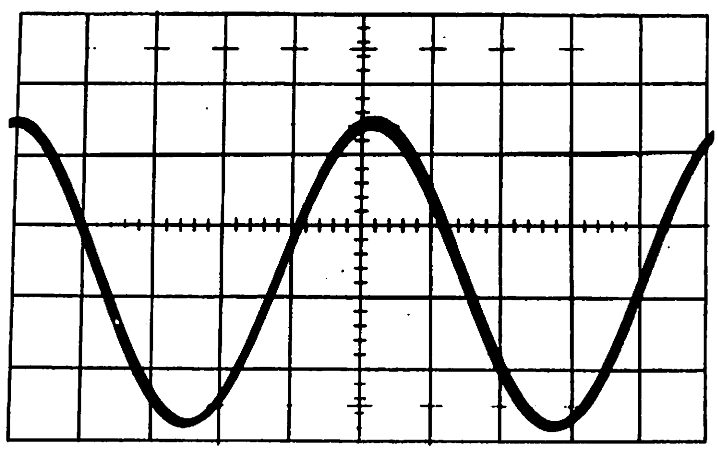


FIGURE 7.20(b): V_{OUT} (DAC) shown lagging V_{IN}
for 200 Hz 9V p-p sine wave input.
vert.: 2V/div ; horiz.: 1 ms/div.

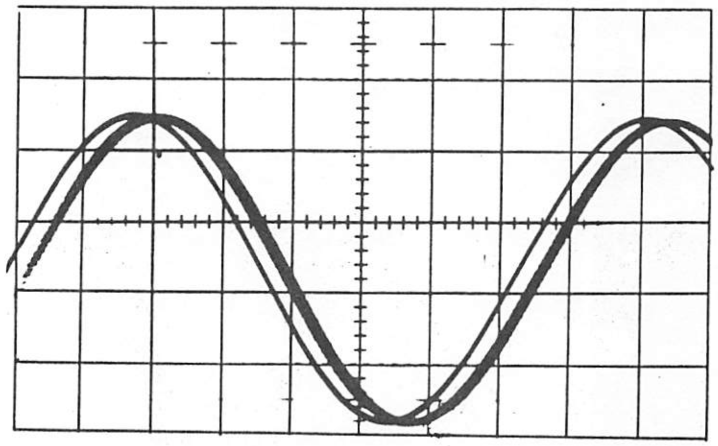


FIGURE 7.20 (c): V_{OUT} (DAC) shown lagging V_{IN} for 800 Hz, 9V p-p sine wave.
vert.: 2V/div. ; horiz.: 200 μs/div.

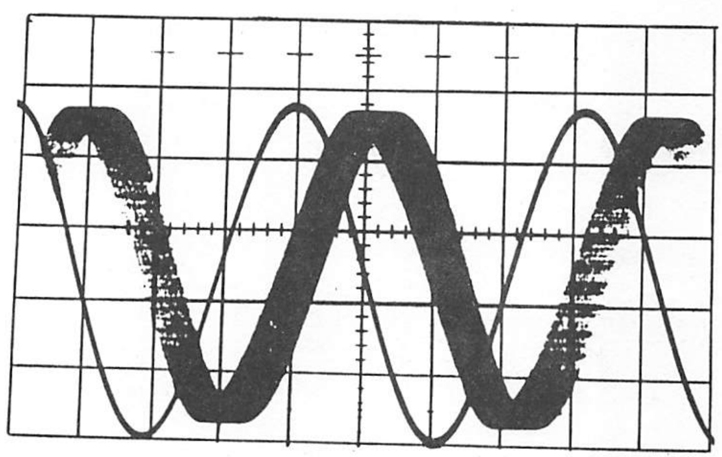


FIGURE 7.20 (d): V_{OUT} (DAC) shown lagging V_{IN} for 5 KHz, 9.5V p-p sine wave.
vert.: 2V/div. ; horiz.: 50 μs/div.

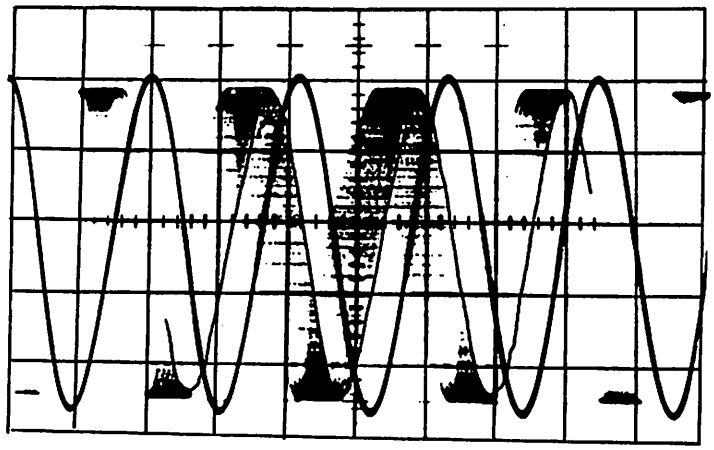


FIGURE 7.20(e): VOUT (DAC) shown lagging VIN for a 10 KHz, 9.5V p-p sine wave input. vert.: 2V/div ; horiz.: 50 μs/div. (image enhanced)

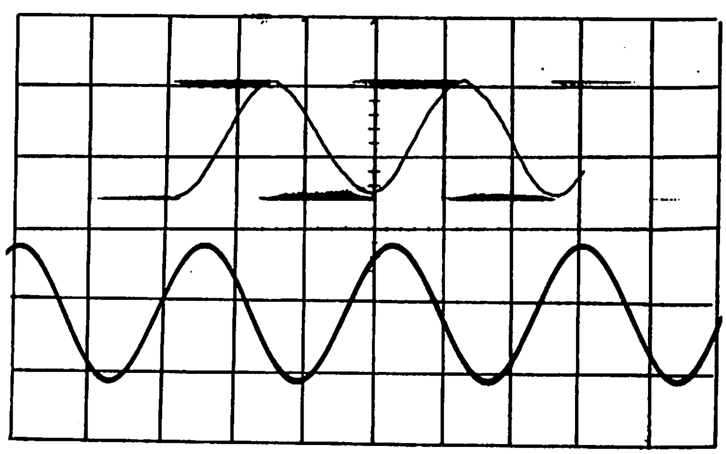


FIGURE 7.20(f): VOUT (DAC) shown above VIN for a 20 kHz, 9.5 V p-p sine wave input. vert.: 5V/div. ; horiz.: 20 μs /div. (image enhanced.)

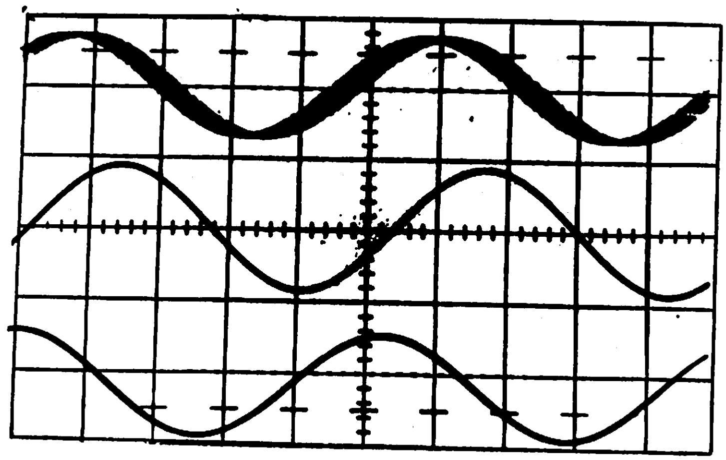


FIGURE 7.21(a): VOUT (DAC) (Top), VOUT (5 KHz filter), VIN (1 kHz, 9V p-p) (lower) for 0.35% THD.
vert.: 5V/div; horiz.: 200 μs/div.

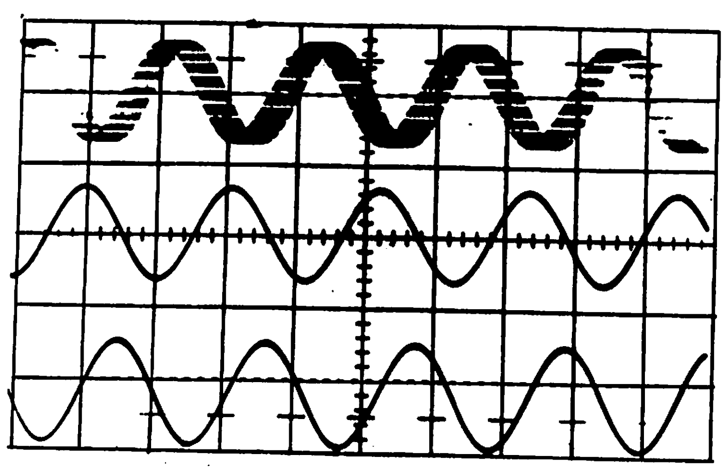


FIGURE 7.21(b): VOUT (DAC) (TOP), VOUT (5 KHz filter), VIN (1 KHz, 9V p-p) (lower) for 0.35 % THD.
vert.: 5V/div.; horiz.: 200 μs/div.

a random error distribution in capacitor ratios. In section 4.10 random edge location was discussed as a possible source of error. This will now be investigated with the aid of a scanning electron microscope (SEM).

Figure 7.22 shows two pictures of photoresist (PR) edges after development but before etching. In both cases the jagged edges are apparent. The PR was about 1 μ thick and the approximate random edge location variation is .1 μ to .2 μ. This defect may arise from the poor resolution capabilities of an emulsion working plate or from the resolution properties of the PR (AZ1350J). Figure 7.23 shows additional examples of the ripples in the PR edges.

In Figure 7.24 the aluminum has been etched but the PR has not yet been removed. A further more serious degradation of the PR is observable. It appears that the etchant also attacks the PR.

Figure 7.25 illustrates the jagged aluminum edges which remain after the metal has been etched and the PR removed.

In conclusion the most significant limitation to increased ratio accuracy is suspected to be the random edge location and this is probably caused by a combination of factors associated with the emulsion working plate, PR resolution, and etchant attack upon the PR as well as non-uniformities associated with etching the aluminum. The observed random error variation was not believed to be significantly dependent upon oxide gradient although this may be a minor factor. In spite of the fact that common centroid was not designed for the smaller capacitors the uniformity in oxide gradient was sufficiently good across the wafer that it would be negligible across one die. In addition the error distribution is not consistent with that which might be expected from an oxide gradient error.

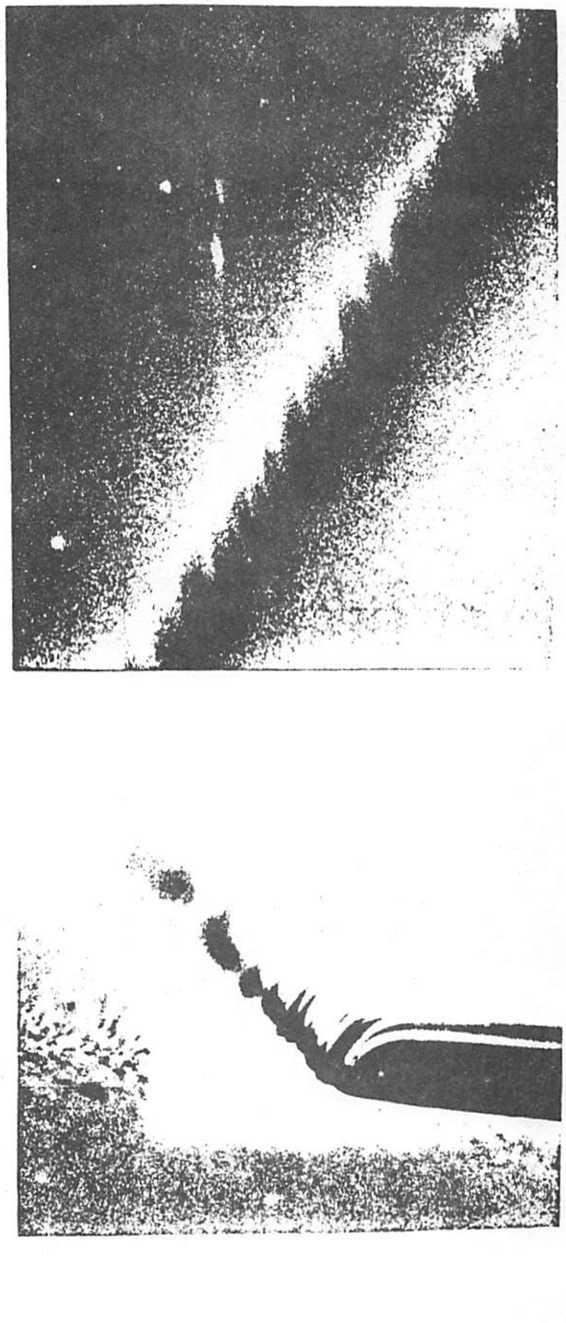


FIGURE 7.22 : PHOTORESIST edges magnified 10,000 x after development.

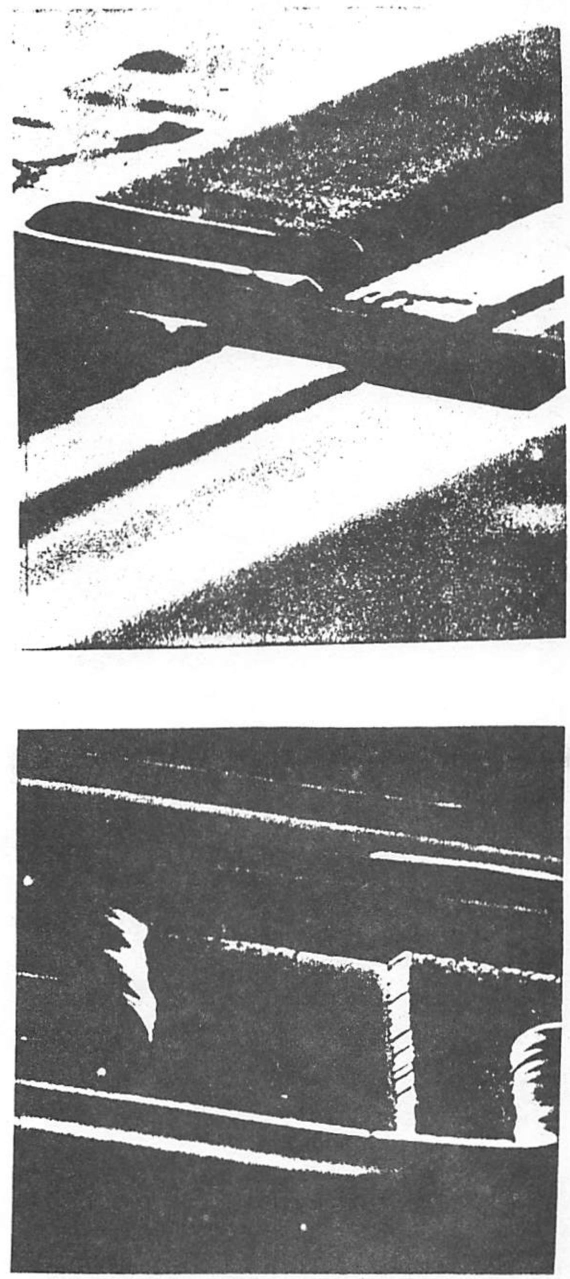
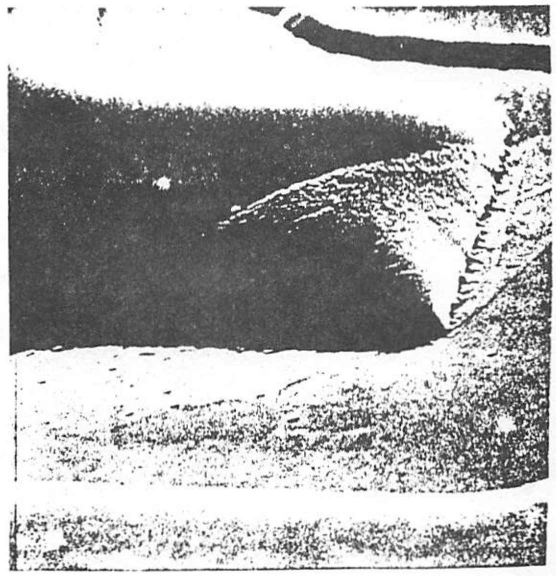
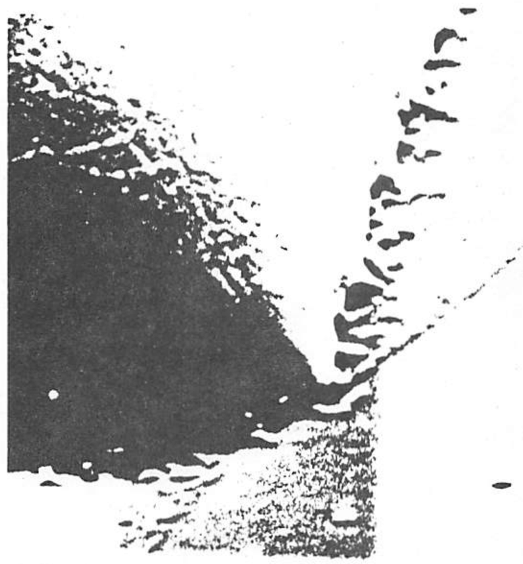


FIGURE 7.23: Developed Photoresist on Aluminum, before etching, showing ripples in resist edges. Magnified 1000x.

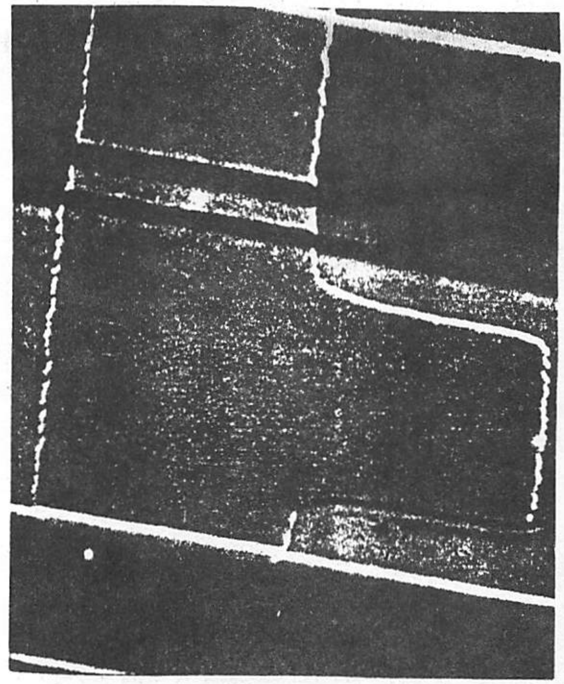


Magnification : 5000 x



Magnification : 12,000 X

FIGURE 7.24 : Photoresist on Aluminum after etching.



Magnification: 3400x

Magnification: 30,000x

FIGURE 7.25 : Aluminum pattern showing random edge location.

CHAPTER 8

CONCLUSION

The conclusions of this research effort are as follows:

1. The feasibility of single-chip realization of a high speed all-MOS ADC at low cost has been demonstrated. This was done by fabricating an experimental I.C. using N-channel aluminum gate technology. The experimental data indicates that conversion accuracies of 10 bits $\pm$ 1/2 LSB can be achieved at high yield.

2. The principal limitations on the accuracy of this technique are due to the conventional photomasking and chemical etching techniques used in the standard fabrication process. The resolution of the photolithography was identified as the practical limitation upon the ratio accuracy. Some improvement in matching could probably be achieved if special techniques were used to enhance photomasking resolutions such as electron beam exposure and ion beam etching of aluminum [41]. Improvement in matching accuracy could certainly be realized if on-chip trimming techniques were developed.

3. The major practical limitations on the conversion rate are due to the practical minimum value of array capacitance and the accuracy with which feedthrough cancellation can be achieved. Both of these are dependent upon the photolithographic resolution limits of conventional photomasking.

In conclusion, this investigation has demonstrated that, with the addition of an external reference voltage, a single-chip MOS ADC may be realized. It is estimated that this realization would require an active chip area of 90 × 90 mils square and consume 56 mW of power. Table 8.1

contains a summary of these requirements

Component	Area $(\text{mils})^2$	Power (mV)
Array	55 × 70	1
Comparator	30 × 25	30
Logic & interconnect	60 × 60	25
Total Chip	90 × 90	56

Power supplies: $\pm$ 15V; V_R = + 10V $\pm$ .05%

Input range: - 10V to + 10V for 20 mV resolution or 0 to + 10V for 10 mV resolution

10-bit conversion time: 23 μs

Table 8.1: Estimated single chip RADCAP Realization.

APPENDIX A

Calculation of Nonlinearity due to Capacitor Voltage Coefficient

As mentioned in section 4.5 the capacitor voltage coefficient α results in voltage dependent capacitors and this causes nonlinearity. This fact will now be supported by numerical calculations of the resultant error. For the capacitor structure illustrated in Figure 4.7 the equations of interest are:

$$C1(V1) = B_i CO\ (1 - \alpha\ V1) = CX(1 - \alpha\ V1)$$

and $$C2(V2) = (2^N - B_i)CO\ (1 + \alpha\ V2) = CY(1 + \alpha\ V2).$$

The change in charge caused by the transient must be equal for both C1 and C2:

$$\Delta Q1 = \Delta Q2.$$

Then $$\int_0^{V_R - V2} C1(V1)dV1 = \int_0^{V2} C2(V2)dV2$$

and $$CX\left(V1 - \frac{\alpha V1^2}{2}\right)\Bigg|_0^{V_R - V2} = CY\left(V2 + \frac{\alpha V2^2}{2}\right)\Bigg|_0^{V2}.$$

Solving for V2 after substituting the limits. The following result is obtained after simplification:

$$V2 = \frac{C_x}{C_x + C_y}\ V_R + \frac{1}{\alpha}\left[\sqrt{1 + {\frac{\alpha\ V_R\ B_i}{2^N}}^{2}\ \frac{1 - 2^N}{B_i}}\ - 1\right]$$

The first term in the equation above is the ideal value of V2 if α equals zero. The second term defined as ε represents the error in V2 due to voltage coefficient, and is expressed as a function of B_i. ε may also be simplified to the expression:

$$\frac{\alpha}{2}\left(\frac{V_R\, B_i}{2^N}\right)^2 \left(1 - \frac{2^N}{B_i}\right) .$$

Table A.1 lists the error as a function of the digital output for a 10-bit converter with α = 22 ppm/volt.

B_i	ε in millivolts	$\frac{\varepsilon\ (\alpha)}{\alpha}$
1024	0	0
992	-.036	$- 1.6 \times 10^3$
960	-.064	$- 2.9 \times 10^3$
896	-.122	$- 5.5 \times 10^3$
768	-.209	$- 9.5 \times 10^3$
640	-.259	$- 11.8 \times 10^3$
512	-.277	$- 12.6 \times 10^3$
384	-.259	$- 11.8 \times 10^3$
256	-.209	$- 9.5 \times 10^3$
128	-.122	$- 5.5 \times 10^3$
64	-.064	$- 2.9 \times 10^3$
32	-.036	$- 1.6 \times 10^3$
0	0	0

TABLE A.1

From the table, the worst case nonlinearity of $-.3$ mV occurs at $\frac{1}{2}$ full scale input. This may be generalized to the final result that the worst case error voltage always appears at an input of $\frac{V_R}{2}$ and its value is:

$$\varepsilon = -\alpha \frac{V_R^2}{8} .$$

APPENDIX B

Digital Logic Circuit

The digital logic circuit is shown in Figure B.1 along with connections to the experimental I.C. There are five logic blocks in addition to the chip: a sequencer, two signal generators, CMOS switches, and a buffer. All signal paths between the six circuits are labeled in the figure. The timing diagram for all signals to the chip are shown in Figure B.2. The timing diagram for capacitor signals is illustrated in greater detail in Figure B.3. The state table necessary to generate the desired timing is shown in Figure B.4. The implementation of the state table results in the circuit schematics for the Sequencer, the Switch Signal Generator, the Capacitor Signal Generator, the CMOS switches, and the Output Buffer which are respectively shown in Figures B.5 through B.9. Figure B.10 illustrates the minimal operating configuration for the experimental chip. The comparator outputs are translated from MOS levels to TTL levels by the CMOS NOR gates shown. Figure B.11 shows the bonding diagram for an experimental chip mounted in a 28 pin DIP.

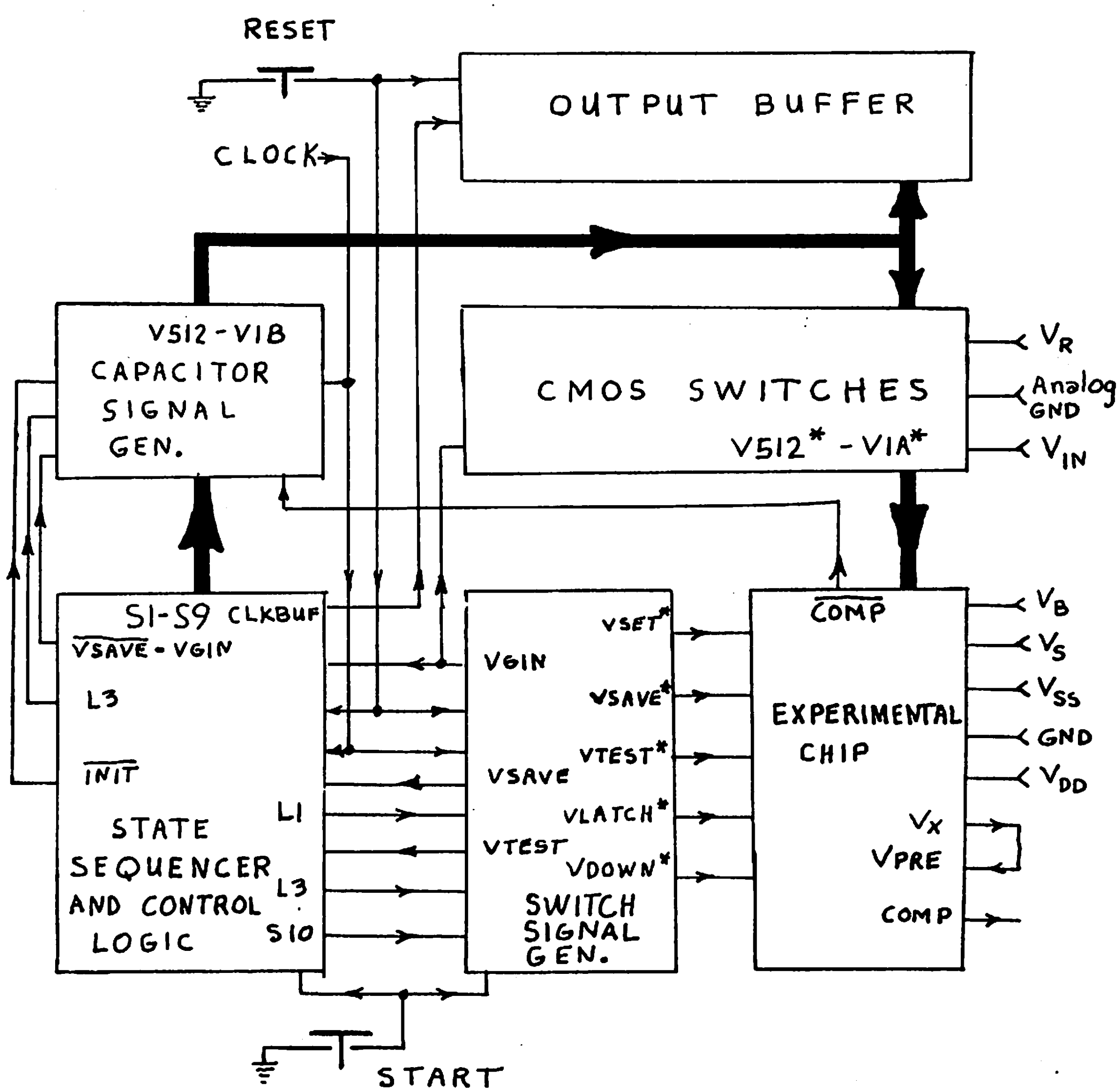


Figure B.1: The complete ADC.

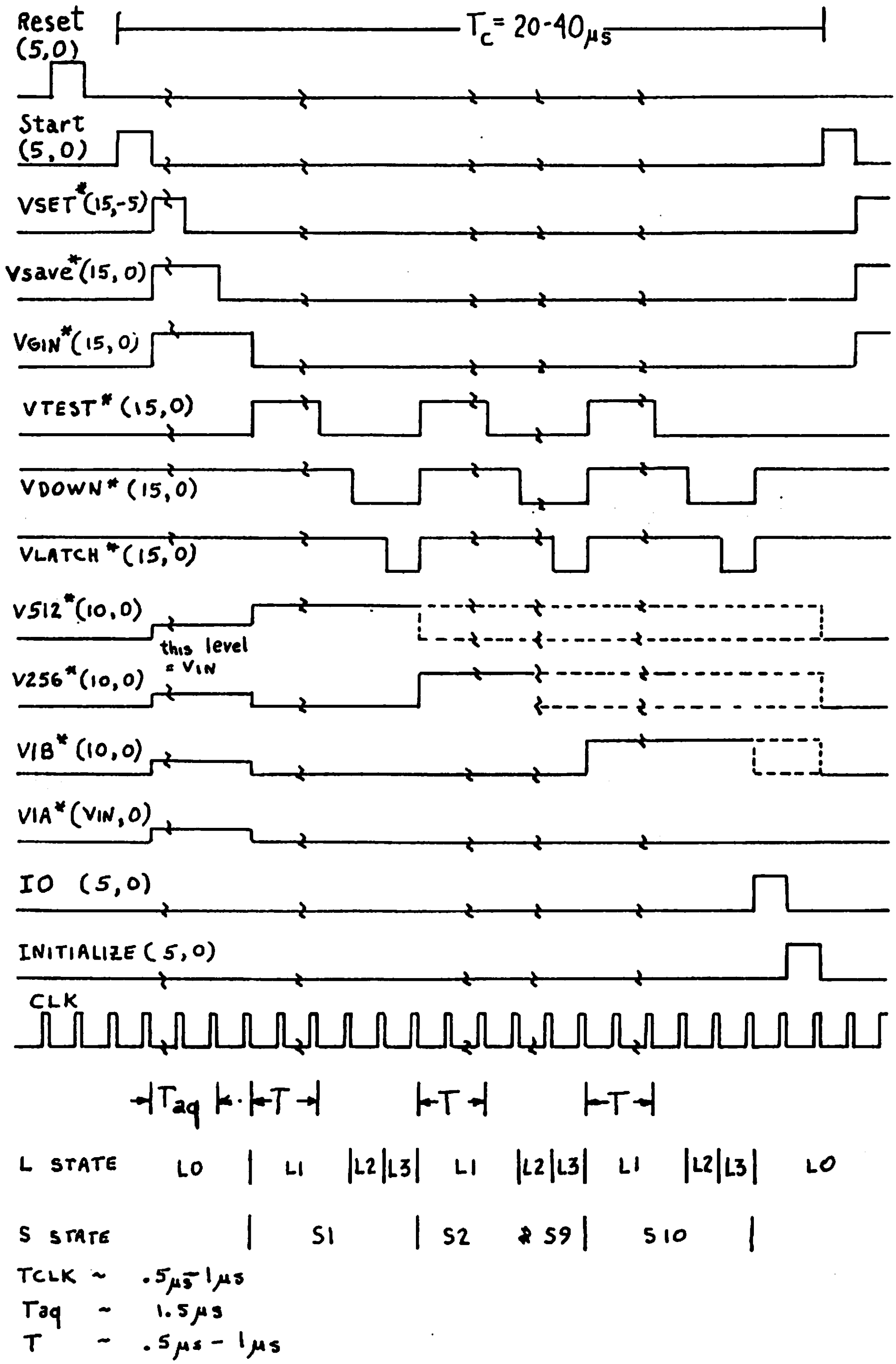


Figure B.2: Timing diagram for signals to experimental I.C.

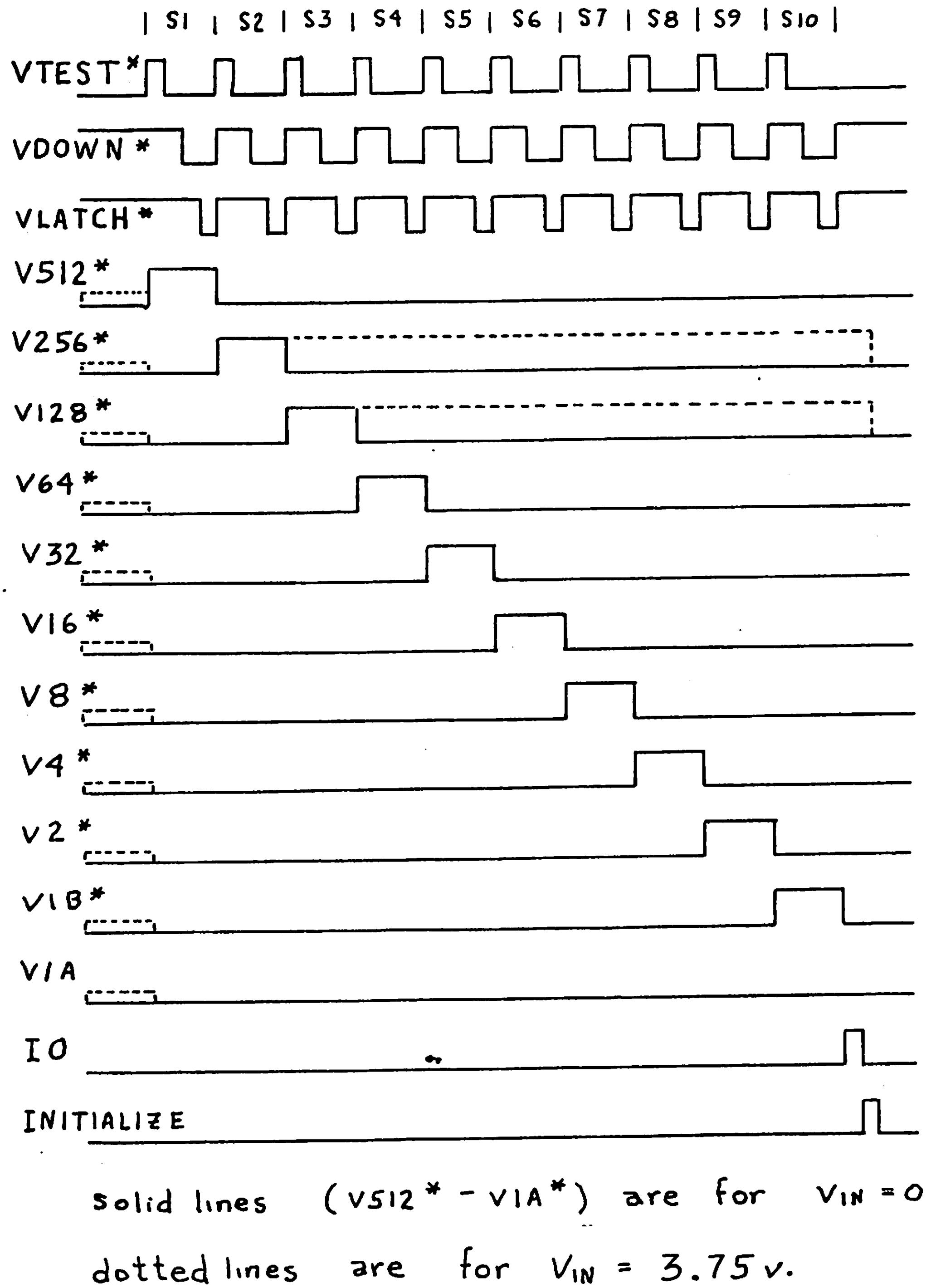


Figure B.3: An expanded timing diagram.

STATE TABLE

RESET ⇒ CLEAR: ST1, ST2, VSAVE, VGIN, VTEST, V512, V256 ... V1B, INITIALIZE, IO

PRESET: VDOWN, VLATCH

STATE:

	LO	RESET
	LO	
	LO	START ; JVSAVE ; JVGIN ; JVSET
	LO	RVSAVE = $\overline{\text{VSET}}$
	LO	RVGIN = $\overline{\text{VSAVE}}$; JV512 ; JVTEST ; $\overline{\text{VSAVE}}$·VGIN → L1, S1
	L1	RVTEST
	L1	RVDOWN ; $\overline{\text{VTEST}}$ → L2
S1	L2	RVLATCH
	L3	RV512 = $\overline{\text{COMP}}$; JV256 ; JVDOWN ; JVLATCH ; JVTEST
	L1	RVTEST
	L1	RVDOWN ; $\overline{\text{VTEST}}$ → L2
S2	L2	RVLATCH
	L3	RV256 = $\overline{\text{COMP}}$; JV128 ; JVDOWN ; JVLATCH ; JVTEST
	L1	RVTEST
	L1	RVDOWN ; $\overline{\text{VTEST}}$ → L2
S3	L2	RVLATCH
	L3	RV128 = $\overline{\text{COMP}}$; JV64 ; JVDOWN ; JVLATCH ; JVTEST
	⋮	⋮
	L3	RV2 = $\overline{\text{COMP}}$; JV1B ; JVDOWN ; JVLATCH ; JVTEST
	L1	RVTEST
	L1	RVDOWN ; $\overline{\text{VTEST}}$ → L2
S10	L2	RVLATCH
	L3	RV1B = $\overline{\text{COMP}}$; JVDOWN ; JVLATCH ; JIO ; L3·$\overline{\text{S10}}$ → L0
	LO	JINITIALIZE = IO ; CLKBUF = IO
	LO	(CLEAR V512 - V1B) = INITIALIZE ; JSTART = INITIALIZE
	LO	START ...
		⋮

Figure B.4: The state table for the digital logic circuit.

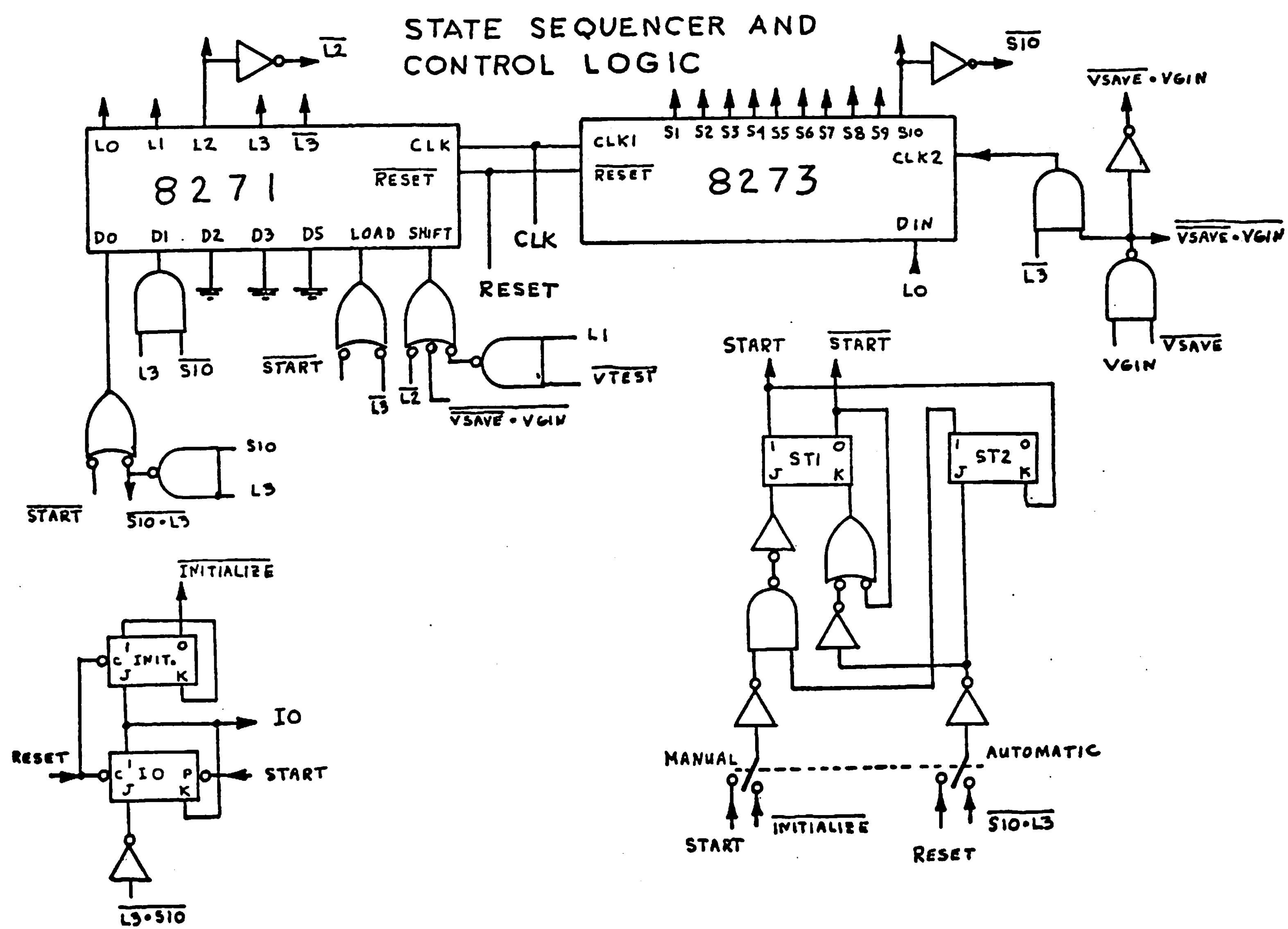


Figure B.5: The logic diagram of the state sequencer and control logic.

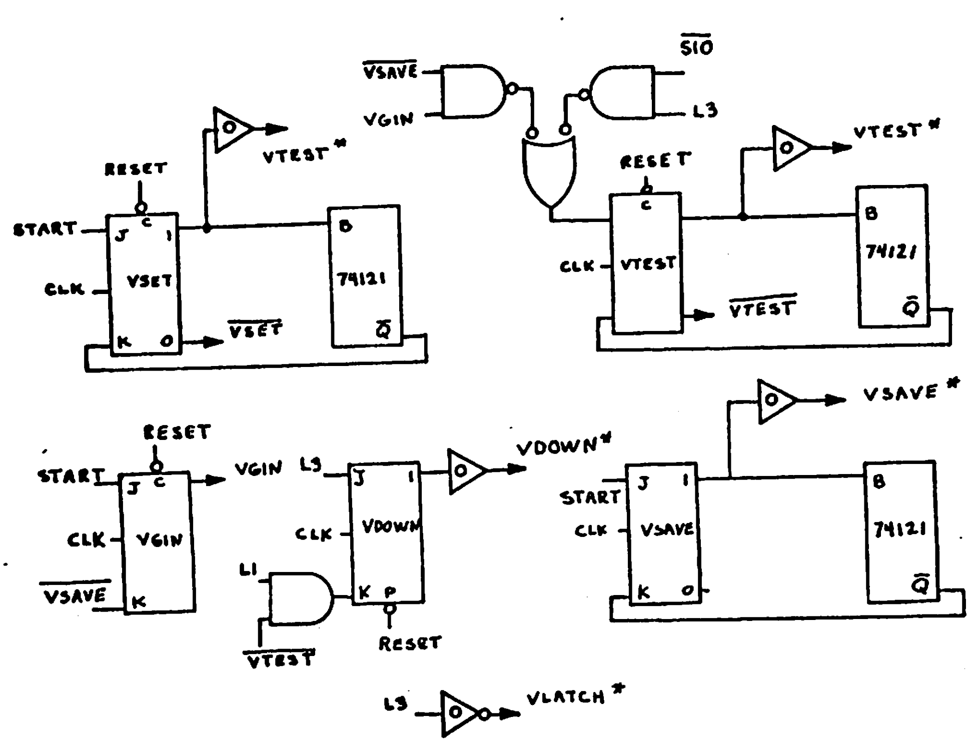


Figure B.6: The switch signal generator.

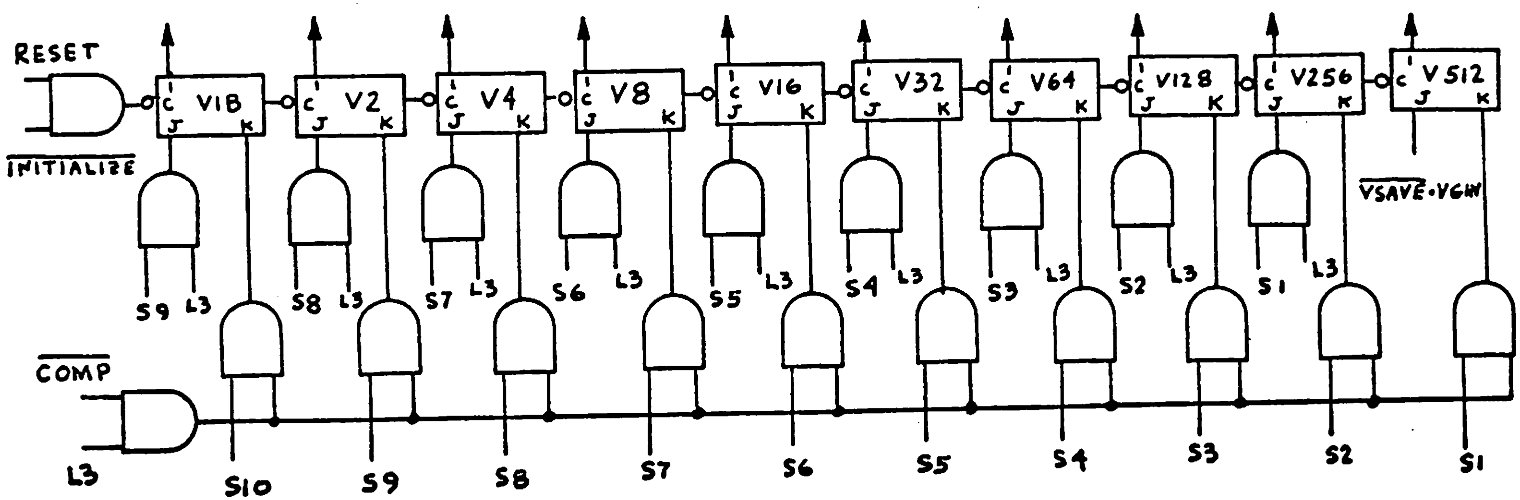


Figure B.7: The capacitor signal generator.

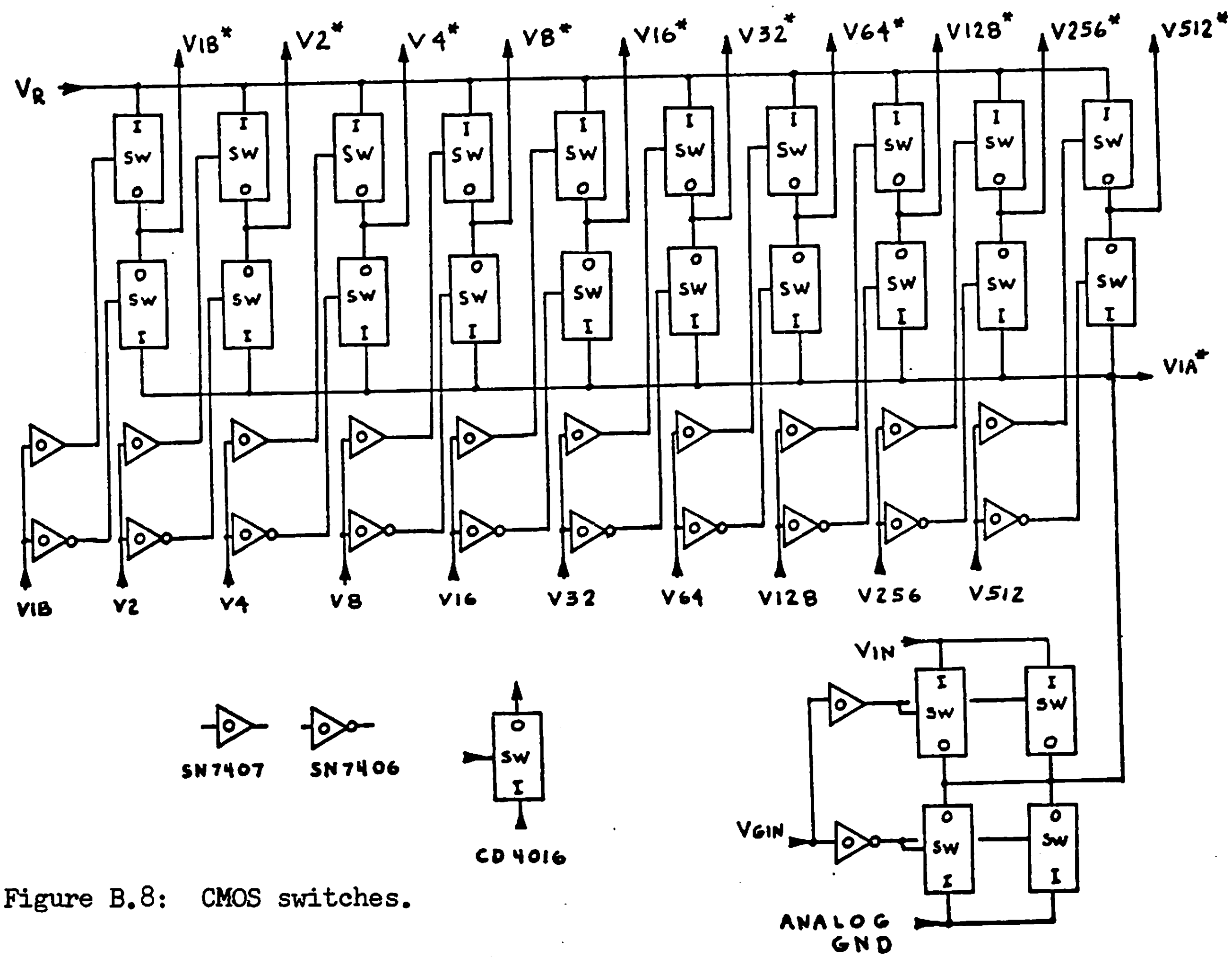


Figure B.8: CMOS switches.

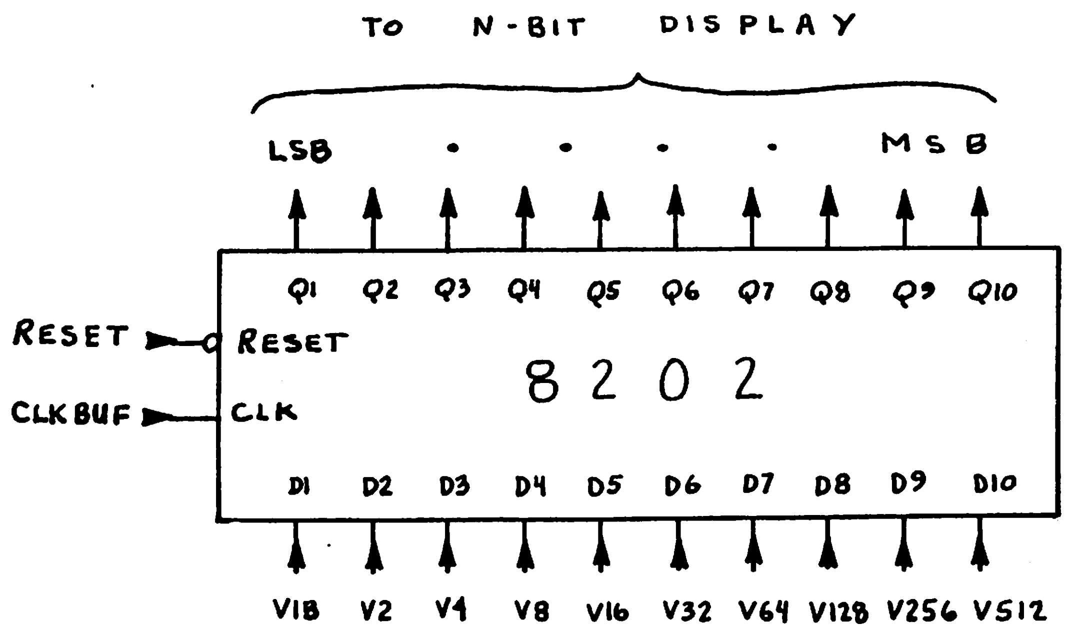


Figure B.9: The output buffer.

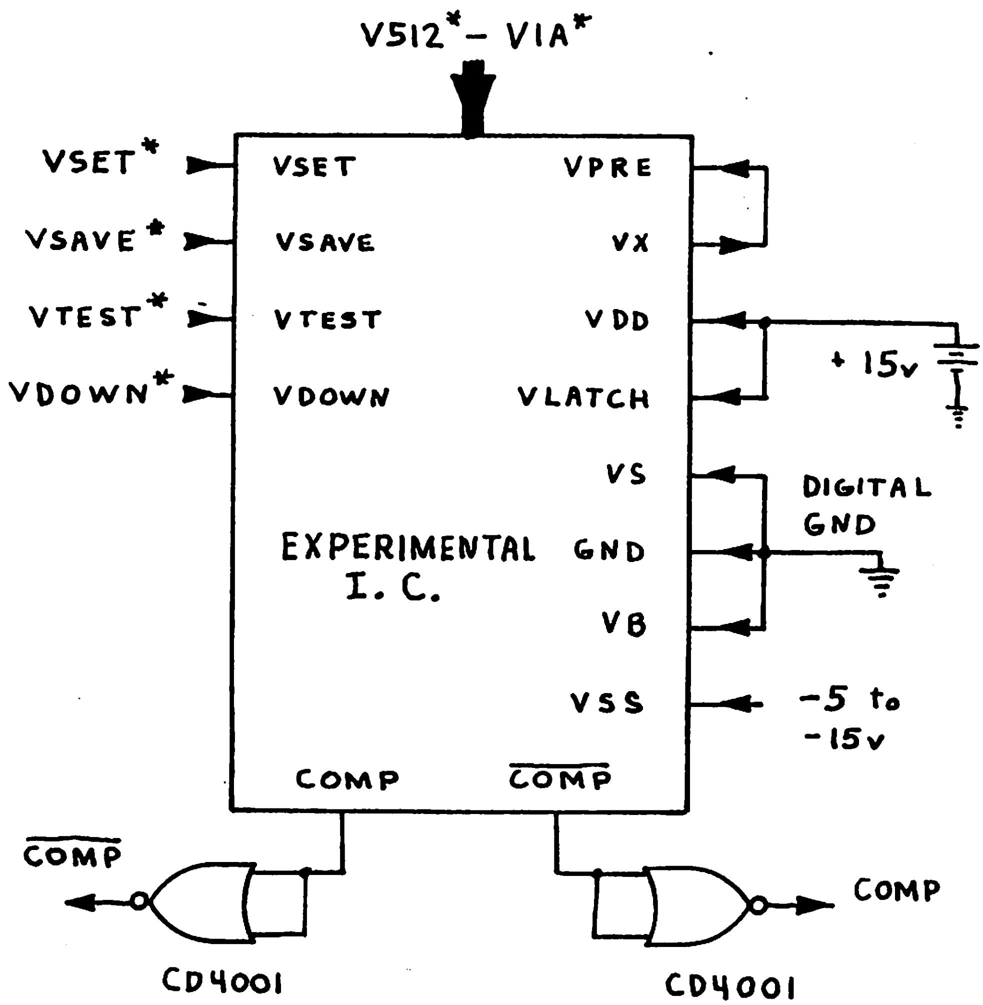


Figure B.10: The minimal operating configuration for the experimental chip.

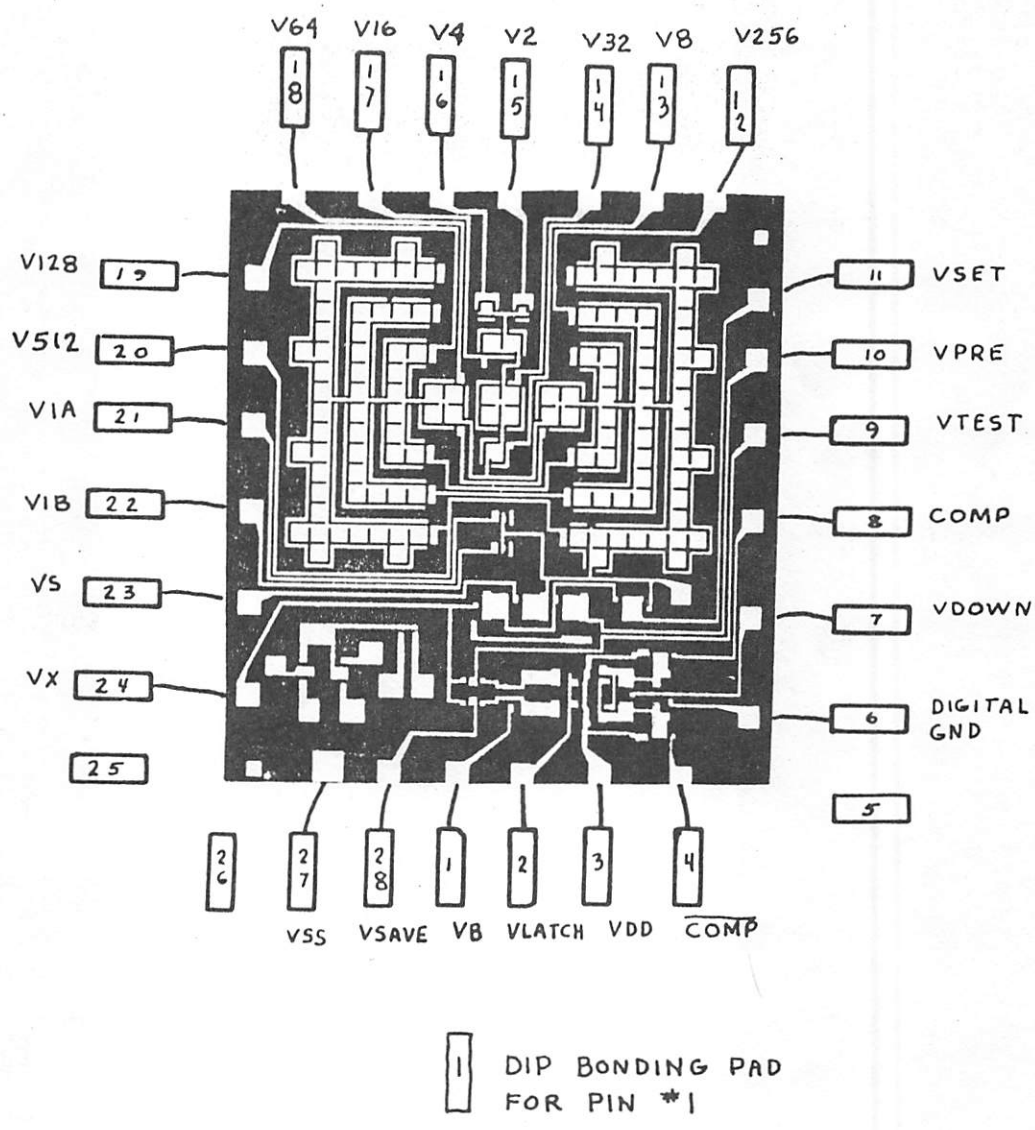


Figure B.11: 28-pin DIP bonding diagram.

APPENDIX C

N-MOS ALUMINUM GATE FABRICATION PROCESS

1. p-Type 100 substrate, 3-5 Ω-cm
2. Clean
   - DI: HF (9:1), dip
   - TCE, 60°C 10 min.
   - Acetone, 2 min.
   - DI rinse
   - RCA Clean:
   - RCAI: 1. NH_4OH: H_2O_2: DI (1:1:5)
     75°C 15 min.
     2. DI Rinse
   - RCAII: 1. HCl: H_2O_2: DI (1:1:6)
     75°C 15 min.
     2. DI Rinse
     3. N_2 Dry
3. Initial Oxidation; Initial Oxidation Furnace at 1150°C; 0.92 μ wet oxide;
   - wet O_2 0.5 ℓ/min. 90 min.
   - N_2 0.65 ℓ/min. 10 min.
4. PR (Photoresist) Step; p^+ isolation diffusion mask;
   - Kodak 747 (Micro neg) resist; 50 c.s.; 5000 rpm; 30 sec.
   - Air dry 10 min.
   - Prebake; 90°C, 30 min.
   - Expose 3 sec.
   - Spray develop; 30 sec.

- Spray rinse; 20 sec.
- Postbake; 125°C for 30 min.
- Oxide etch; NH_4F: HF (5:1);

  9.5 min for .92 μ at .1 μ/min
- Strip PR; H_2SO_4: H_2O_2 (4:1); 90°C; 5 min.

5. RCA (RCAI and RCAII as in Step 2)
6. p^+ Predeposit; Boron Predeposition Furnace @ 950 °C
   - B_2H_6 0.26 ℓ/min
   - O_2 13 cc/min 15 min.
   - N_2 1.3 ℓ/min
7. Cleaning
   - Etch boron glass 12% HF dip (HF:DI, 1:3)
   - RCAII
8. Oxide over p^+; p-type drive-in Furnace @ 1150°C;

   grow .4 μ wet O_2;
   - wet O_2 0.5 ℓ/min 16 min
   - N_2 1.0 ℓ/min 10 min
9. PR step; N^+ mask; .95 μ (9.5 min etch)
   - same as in Step 4
10. Cleaning
    - RCA
11. N^+ Predeposit; Phosphorous Predeposition furnace @ 1100°C

$$\left.\begin{array}{ll} -\ O_2 & 0.1\ \ell/\text{min} \\ -\ N_2 & 1.25\ \ell/\text{min} \end{array}\right\}\ 5\ \text{min}$$

$$\left.\begin{array}{ll} -\ O_2 & 0.1\ \ell/\text{min} \\ -\ N_2 & 1.25\ \ell/\text{min} \\ -\ N_2/POCl_3 & 96\ \text{cc/min} \end{array}\right\}\ 20\ \text{min}$$

- O_2 0.1 ℓ/min } 2 min
- N_2 1.25 ℓ/min }

12. Cleaning
    - Etch phosphorous glass; 12% HF dip
    - DI rinse
    - RCAII

13. Oxide over N^+; N-type Drive-in furnace @ 1100°C; wet O_2; 0.5 μ;
    - wet O_2 0.5 ℓ/min 34 min
    - N_2 1.0 ℓ/min 10 min

14. PR step; gate oxide mask;

    .99 μ doped oxide (6.5 min. etch)

15. Cleaning
    - RCA

16. Grow gate oxide; N-type drive-in furnace @ 1000°C; dry O_2; .1 μ; wafer horizontal on boat
    - O_2 1.5 ℓ/min 110 min total time

rotation	Δt
0°	10 min
180°	22 min
90°	33 min
270°	45 min

    - N_2 1.0 ℓ/min 5 min

17. PR Step; contact window mask;

    0.1 μ (1.5 min. etch)
    - same as step 4 except mask is double exposed for 2.5 sec. each after mask shifted 1 row. (This technique avoids pinholes.)

18. Clean

    - RCAII

19. Aluminum Evaporation; electron beam vacuum chamber;

    - IR Lamp drying 15 min.

    - Evaporation; .3 μ to .4 μ Aluminum

20. PR Step; AZ1350J resist; Metallization Mask;

    - IR Lamp heating 10 min

    - Spin coat AZ1350J at 8000 rpm; 30 sec;

    - 10 min air dry

    - Prebake 80°C for 45 min with air circulation

    - Expose 12 sec.

    - Spray develop MF312: DI (1:1); 45 sec.

    - DI Rinse

    - Postbake 20 min; 90°C

    - Aluminum etchant type A; 40°C; ultrasonic agitation; (~ 45 sec)

    - 1112 PR Stripper; 50°C; 2 min.

    - DI Rinse, N_2 dry

21. Clean

    - TCE dip

    - Acetone dip

22. Heat treatment; Sintering Oven; 200°C; 5 min

    - N_2: H_2 (9:1) - forming gas; 1 ℓ/min.

APPENDIX D

Calculation of Capacitor Plate Duplication Size Needed for Undercut Insensitivity

The effect of uniform undercut upon capacitor ratios will now be computed. This analysis will determine the duplication geometry needed for the larger capacitors. Let S be the design value of side length of a square plate representing the smallest capacitor C1. The uniform undercut (u) corresponds to a reduction in the edge length of each capacitor. The general approach used to determine the worst case deviation (WCD) in LSB for the entire array is outlined below:

1. Determine total area A_T.
2. Determine the resultant binary weight of the ith capacitor:

$$B_i = \frac{A_i}{A_T} 2^N.$$

3. Determine the deviation from ideal value:

$$Dev_i = B_i - 2^i$$

for $N-1 \geq i$.

4. Estimate WCD = Dev_N.

For an array having no duplication of smaller plates used in large capacitor construction, the total area is given by:

$$A_T(N,s,u) = (s-u)^2 + (s-u)^2 + (\sqrt{2}\ s-u)^2$$
$$+ (\sqrt{4}\ s-u)^2 + \cdots + (\sqrt{2^{N-1}}\ s-u)^2.$$

However, it has been shown in Chapter IV that it is necessary to design the larger capacitors by paralleling smaller plates of identical geometry in order to achieve insensitivity to uniform undercut. Let the kth capacitor corresponding to binary weight 2^k be used as the duplication unit. Hence the next largest capacitor would have binary weight $2(2^k)$ independent of undercut relative to 2^k. Using this approach the total capacitor area is:

$$A_T(N,k,s,u) = (s-u)^2 + (s-u)^2 + (\sqrt{2}\, s-u)^2$$

$$+ \cdots + (\sqrt{2^k}\, s-u)^2 + 2(\sqrt{2^k}\, s-u)^2$$

$$+ 4\,(\sqrt{2^k}\, s-u)^2 + \cdots + 2^{N-1-k}(\sqrt{2^k}\, s-u)^2,$$

$$= 2^N s^2 - 2us[2 + \sqrt{2} + \sqrt{4} + \cdots + \sqrt{2^k}(2^{N-k}-1)]$$

$$+ u^2[2^{N-k} + k].$$

The area of a large capacitor of weight i for $N-1 \leq i \leq k$ is:

$$A_i(N,k,s,u) = 2^{i-k}(\sqrt{2^k}\, s-u)^2.$$

The actual binary weight of this capacitor is:

$$B_i(N,k,s,u) = \frac{A_i\, 2^N}{A_T}$$

and neglecting the terms containing u^2 since $u \ll s$ then

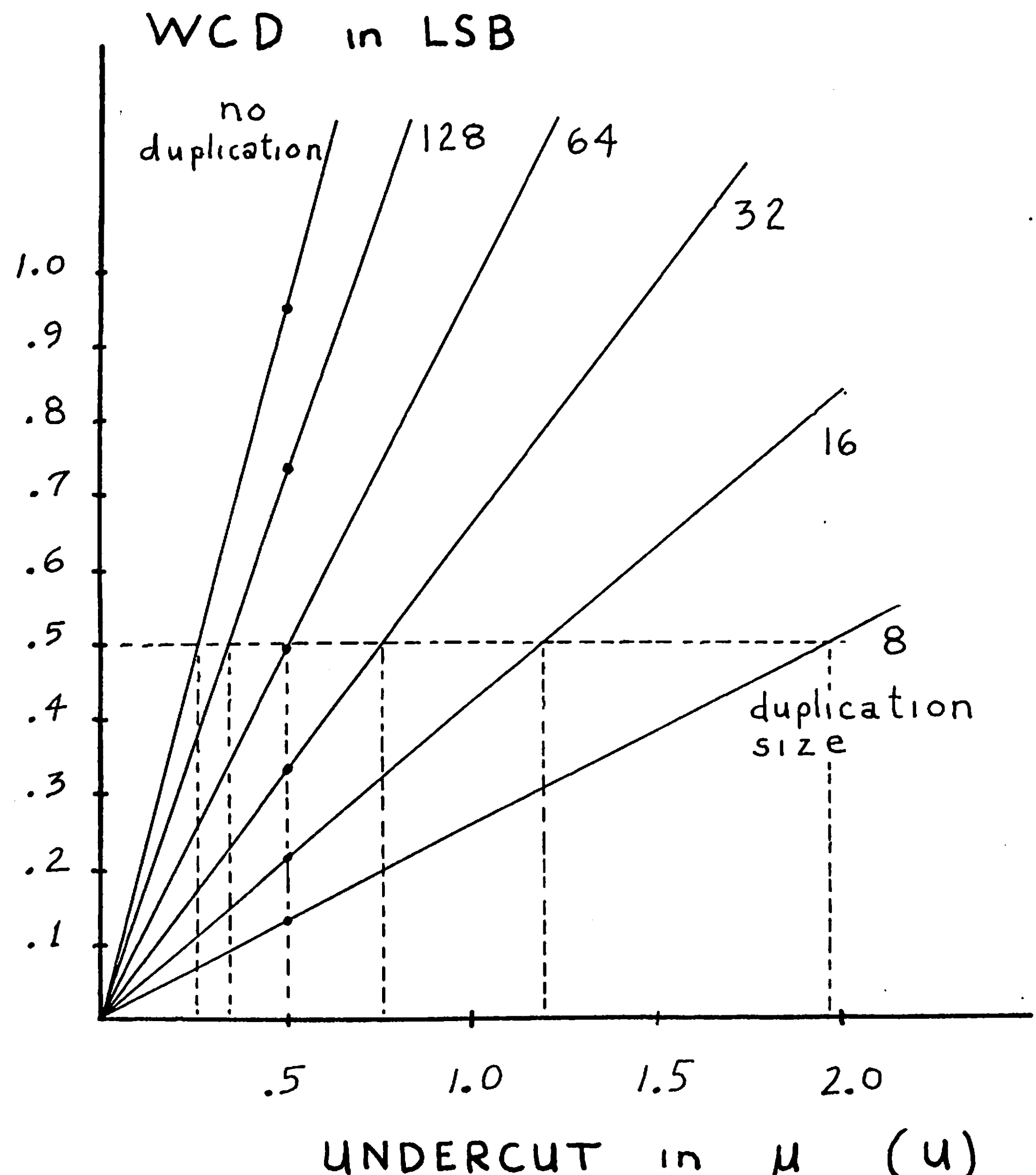
$$B_i = 2^i + \frac{u}{s} 2^i \left(\frac{2 + \sqrt{2} + \sqrt{4} + \cdots + 2^k(2^{N-k}-1)}{2^{N-1}} - 2^{1-\frac{k}{2}}\right).$$

The WCD for the entire array is approximately equal to the second term in this equation evaluated for i = N and 1 < k < N-1

$$WCD(N,k,s,u) = \frac{2u}{s}\left[2 + \sqrt{2} + \sqrt{4} + \cdots + \sqrt{2^{k-1}} - \sqrt{2^k}\right].$$

An interesting observation is that although the absolute error becomes smaller as N becomes larger the WCD in LSB is approximately independent of N provided that N is large.

The plot of WCD (k,u) versus u for nominal values of N = 10 bits and s = 20 μ is given in Figure D.1. From this graph, an array structure having duplication of capacitor size C8 (for which k = 3 and binary weight equals 8) will retain $\pm$ .5 LSB ratio accuracy in spite of 2 μ undercut.



Duplication Size	Allowable u for ± .5 LSB
none	.27 μ
128	.35 μ
64	.50 μ
32	.76 μ
16	1.20 μ
8	1.97 μ

Figure D.1: The worst case deviation from linearity (WCD) as a function of duplication size (k) and uniform undercut (u).

REFERENCES

1. J.L. McCreary and P.R. Gray, "A High Speed, All-MOS, Successive Approximation, Weighted Capacitor A/D Conversion Technique," ISSCC Digest of Technical Papers, pp. 38-39, Feb. 1975.
2. J. Albarran, Univ. of CA., Berkeley, CA., Private Communications.
3. "Engineering Product Handbook - A/D and D/A Converters," DATEL Systems, Inc., pp. 1-41, 1974.
4. K. Fukahori, "All MOS A/D Converter," M.S. Plan II Report, University of California, Berkeley, CA., 1974.
5. D.F. Hoeschele, Analog-to-Digital/Digital-to-Analog Conversion Techniques, Wiley & Sons, New York, 1968.
6. R.E. Suarez, P.R. Gray, and D.A. Hodges, "An All-MOS, Charge Redistribution, Successive Approximation A/D Conversion Technique," ISSCC Digest of Technical Papers, pp. 194-195, Feb. 1974.
7. G. Smarardoiu and D.A. Hodges, "An All-MOS Analog-to-Digital Converter Using a Constant Slope Approach," European Solid-State Circuits Conference, Canterbury, U.K., IEEE Conf. Pub. 130, pp. 60-61, Sep. 1975.
8. D.J. Dooley, "A Complete Monolithic 10-b D/A Converter," IEEE J. Solid-State Circuits, Vol. SC-8, pp. 404-408, Dec. 1973.
9. G. Kelson, H.H. Stellrecht, and D.S. Perloff, "A Monolithic 10-b Digital-to-Analog Converter Using Ion Implantation", IEEE J. Solid-Circuits, Vol. SC-8, pp. 396-403, Dec. 1973.
10. Analog-Digital Conversion Handbook, Engr. Staff of Analog Devices Inc., 6th Ed., 1972.
11. "Analogic Announces One-Chip Voltmeter," Electronic Engineering Times, pp. 2, Jan. 1, 1975.

12. J.A. Schoeff, "A Monolithic Analog Subsystem for High-Accuracy A/D Conversion," ISSCC Digest of Technical Papers, pp. 18-19, Feb. 1973.

13. "10-bit CMOS A/D Converter Chip Interfaces with LSI Microprocessors," Electronic Design 2, pp. 81-82, Jan 18, 1975.

14. H. Schmid, Electronic Analog/Digital Conversions, Van Nostrand-Reinhold, New York, 1970.

15. J.L. McCreary and P.R. Gray, "All-MOS A/D Conversion Techniques, Part I," IEEE J. Solid-State Circuits, Dec. 1975.

16. P.R. Gray, Univ. of CA., Berkeley, CA., Private Communication.

17. Ibid.

18. Ibid.

19. "Capacitance and Capacitors," Electronics, pp. 61-66, May 11, 1962.

20. G. Kelson, et al., op. cit.

21. P. Greiff, "Temperature Coefficient of Diffused Resistors," IEEE Proceedings (Corresp.), Vol. 53, pp. 215-216, Feb. 1965.

22. Integrated Silicon Device Technology--Capacitance, Research Triangle Institute, Technical Documentary Report No. ASD-TDR-63-316, Vol. 2, Oct. 1963.

23. A.S. Grove, Physics and Technology of Semiconductor Devices, Wiley and Sons, New York, 1967.

24. P.R. Gray, op. cit.

25. R.E. Suarez, "Analog-to-Digital Conversion in MOS Integrated Circuits," Ph.D. Thesis, University of California, Berkeley, 1975.

26. J.G. Simmons and G.W. Taylor, "Dielectric Relaxation and its Effect on the Isothermal Electrical Characteristics of Defect Insulators," Physical Review B, Vol. 6, No. 12, pp. 4793-4803, Dec. 15, 1972.

27. P.J. Burkhardt, "Dielectric Relaxation in Thermally Grown SiO_2 Films," IEEE Trans. on Electron Devices, Vol. 13, pp. 268-275, Feb. 1966.

28. R.J. Kriegler and R. Bartnikas, "Dielectric Relaxation in Si - SiO_2 - C_r Structures," IEEE Trans. on Electron Devices (Corresp.), pp. 1010-1011, Nov. 1970.

29. P. Richman, MOS Field-Effect Transistors and Integrated Circuits, Wiley & Sons, New York, 1973.

30. L. Vadasz and A.S. Grove, "Temperature Dependence of MOS Transistor Characteristics Below Saturation," IEEE Trans. on Electron Devices, Vol. 13, pp. 863-866, Dec. 1966.

31. W.M. Penney and L. Lau, Editors, MOS Integrated Circuits, Van Nostrand-Reinhold, New York, 1972.

32. R.E. Suarez, Univ. of CA., Berkeley, CA., private communication.

33. C.H. Sequin, "Fringe Field Corrections for Capacitors on Thin Dielectric Layers," Solid State Electronics, Vol. 14, pp. 417-420, 1971.

34. P. Richman, Characteristics and Operation of MOS Field-Effect Devices, McGraw-Hill, New York, 1967.

35. P.E. Gray and C.L. Searle, Electronics Principles, Physics, Models, and Circuits, Wiley and Sons, New York, pp. 898-904, 1969.

36. B.C. Young, "N-Channel Silicon-Gate MOS Transistor Fabrication Process," M.S. Thesis, University of California, Berkeley, 1974.

37. M. Chen and J.W. Hile, "Oxide Charge Reduction by Chemical Gettering with Trichloroethylene During Thermal Oxidation of Silicon," J. Electrochem. Soc., Vol. 119, pp. 223-226, Feb. 1972.

38. K. Hirabayashi and J. Iwamura, "Kinetics of Thermal Growth of HCl Oxides on Silicon," J. Electrochem. Soc., Vol. 120, No. 11, pp. 1597-1601, Nov. 1973.

39. D.A. Hodges, Univ. of CA., Berkeley, CA., private communication.

40. G. Schottky, "Decrease of FET Threshold Voltage Due to Boron Depletion During Thermal Oxidation," Solid State Electronics, Pergamon Press, Vol. 14, pp. 467-474, 1971.

41. D.A. Hodges, op. cit.