SWITCHED CAPACITOR SIGNAL PROCESSING CIRCUITS IN SCALED TECHNOLOGIES

by

Chorng-kuang Wang

Memorandum No. UCB/ERL M86/84

3 November 1986

SWITCHED CAPACITOR SIGNAL PROCESSING CIRCUITS IN SCALED TECHNOLOGIES

by

Chorng-kuang Wang

Memorandum No. UCB/ERL M86/84

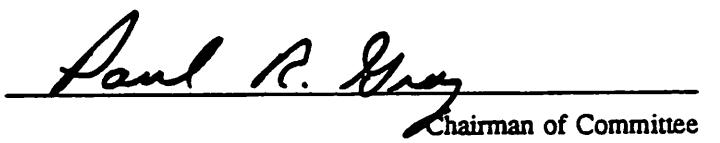
3 November 1986

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Switched Capacitor Signal Processing Circuits in Scaled Technologies

Ph. D. Chorng-kuang Wang Department of EECS



Chairman of Committee

Abstract

The objective of this project is to develop design techniques for switched capacitor signal processing circuits that take advantage of scaled technologies and closely approach their fundamental limits in area and power.

In the past, switched capacitor circuits for communications had to use both large power supplies and large devices to achieve sufficient gain and dynamic range. Today, scaled technologies are available. However, scaling generally increases flicker noise and reduces operational amplifier gain and output swing. This project is aimed at developing circuit solutions to these and other problems associated with technology scaling. To allow the use of minimum-sized transistors in the operational amplifier without degrading noise performance, chopper stabilization is used. Class AB operational amplifier architectures are used to allow large capacitive load charging currents without the large V_{DSAT} values and poor swing associated with large bias currents in minimum-sized devices. Fully differential architectures are used to preserve supply rejection and dynamic range.

An experimental fifth order elliptic PCM low-pass filter using these techniques has been fabricated in a 3-μm CMOS technology. This experimental chip achieves an area of 200 mil^2 per pole. Measured results for a ± 2.5 V power supply are: a dynamic range of 87 dB, and a power supply rejection ratio (PSRR) of 40 dB over 1MHz range.

ACKNOWLEDGEMENT

I would like to express my deep gratitude to Professor Paul R. Gray for his thoughtful guidance, constant support and encouragement throughout the course of my Ph. D. study, which has been the most enlightening and rewarding experience of my life.

I also wish to thank Dr. Rinaldo Castello for his many suggestions and technical assistance.

Friendship with my fellow graduate students has helped to make my life in Berkeley both enjoyable and memorable. And, ideas discussed with my colleagues have contributed quite a bit to this work.

Also, I appreciate the proof reading of my thesis done by Rinaldo Castello, Gordon Jacobs, Min-chi Jeng, Steve Lewis, Ting-ping Liu and Nan-shen Lin.

Finally, I thank my father Swei Wang for his encouragement and patience. I am indebted to those members of my family who have taken care of my father in my place for so many years. Last but not least, I would like to thank my wife Li-ming for her inspiring spirit and enduring effort through all these years.

This research was supported by the National Science Foundation grant ECS-8501488 and by the Semiconductor Research Corporation grant SRC-82-11-008.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

One important aspect of VLSI is the reduction of area. Both analog and digital integrated circuits use less area in scaled technologies than in non-scaled technologies. For the past ten years, digital signal processing circuits have taken advantage of scaled technologies that have resulted in dramatic area reduction and, hence, high level integration. Analog signal processing circuits, however, do not scale as easily as digital circuits due to certain fundamental as well as practical limiting factors. In particular, as the feature size is reduced, the KT/C noise increases and the power supply voltage is reduced, facts that pose some ultimate limitation to the achievable dynamic range of scaled analog circuits. Because digital circuits have been easier to scale than analog circuits, digital signal processing techniques are becoming more popular than in the past.

Digital techniques are robust to variations in temperature and minimum feature size because of the discrete nature of their amplitudes. Analog signals, on the other hand, are continuous in amplitude and are therefore more easily corrupted by those factors mentioned than digital signals. Because of this fundamental difference, there are some areas where analog and digital techniques rarely compete with each other. For example, digital techniques, are suitable for processing discrete variables in digital environments e.g. programmable algorithms, adaptation, frequency decimation. Analog techniques, on the other hand, are best suited for interfacing the analog world to digital environments (e.g. DSP and microprocessor). The following functions are often needed in such analog-to-digital conversion: low noise amplification, prefiltering and equalizing. The question as to whether to use analog or digital techniques is controversial only in the area of applications where both techniques can provide viable solutions. In this case, the choice between analog or digital processing techniques depends on many factors such as cost, reliability and function performance. For example, a two pole prefilter in a system can be easily implemented by switched

capacitor techniques, because just two op-amp's, several capacitors and several switches are required. A digital approach is not cost effective in this case, since analog to digital conversion, registers, ALU, etc are necessary to perform digital filtering. However, the above argument may not be true if, say, 15 poles of filtering is required, because CAD tools may make the digital design easy and turn-around time fast.

In principle, except for analog-to-digital conversions digital approaches can do what analog techniques can do. Then, there are, at least, two options to reduce the area. One choice is to use mostly digital techniques. The other alternative is to develop a way to scale analog circuits and continue using the analog techniques. The latter choice is the one discussed in this thesis. Specifically, the discussion concerns switched capacitor signal processing circuits. It will be shown that analog signal processing techniques are viable and valuable methods of taking advantage of scaled technologies, especially, in the low power and high frequency applications [1,2].

1.1. Motivations and Objectives

Switched capacitor techniques provide stability and accuracy to the MOS analog signal processing circuits. With the continuous evolution of scaled technologies, many complicated designs of telephony applications today provide entire functions on a chip, which particularly heavily depend on switched capacitor techniques [3-9].

Although there is improvement in the level of integration due to scaled technologies, switched capacitor processing circuits today are not scaling as fast as digital circuits; furthermore, switched capacitor circuits use power supplies and technologies that are not compatible with state of art digital technologies.

The main topics of this thesis are to identify the problems associated with the scaling of switched capacitor signal processing circuits, explore system and circuit solutions to minimize or remove the impact of the performance degradation due to scaling, and achieve the smallest possible area and power consumption for a prototype circuit. As a vehicle to test the feasibility of proposed

circuit approaches for scaling analog circuits, a PCM filter using switched capacitor techniques has been implemented. Such a filter was chosen because it has been widely used in telecommunication and has well defined frequency response and signal-to-noise ratio requirement. The prototype scalable PCM SCF achieves an area of only 200 mil^2/pole relative to 400 mil^2/pole of commercial product available today.

1.2. Outline

In chapter 2, MOS scaling concepts and objectives are discussed first. Then, the definition, advantages and problems of the three commonly used scaling laws, namely, CE (Constant field), CV (Constant Voltage) and QCV (Quasi Constant Field) are considered.

Chapter 3 deals with the circuit parameter performance of scaled MOS devices. First order equations are employed in order to be able to quickly perceive the parameter performance change by device scaling. Applying scaled device parameters, the performance of a basic two stage CMOS op-amp under scaling laws is discussed.

Chapter 4 is concerned with the performance degradation of scaled analog MOS circuits. These problems can be fundamental or practical. The most obvious ones are the increase of noise and the decrease of gain and voltage swing.

In chapter 5, design techniques for scalable MOS analog circuits are proposed. A system approach and op-amp design techniques are key solutions to the problems associated with scaling.

Chapter 6 describes the design and implementation of a prototype scalable filter using switched capacitor techniques with all the design approaches given in Chapter 5. Measured filter data and justification of circuit techniques employed will be discussed.

Chapter 7 discusses the fundamental/practical scaling limits for MOS devices and their circuit applications. Short channel effects that degrade circuit performance will be identified.

Finally, Chapter 8 draws conclusions from this project and gives possible future directions of switched capacitor signal processing circuits using scaled technologies.

CHAPTER 2

SCALING CONCEPTS AND STRATEGIES

The basic MOS CE (constant field) scaling law proposed by Dennard et. al. [10] in the early 1970's has been intensively used to maintain electrical characteristics of conventional long channel MOS devices and achieve improved performance of LSI circuits. Following the advent of the VLSI era however, CE scaling results in a degradation of the available dynamic range and lack of TTL compatibility. To avoid this problem non-constant field scaling laws, namely CV (Constant Voltage) and QCV (Quasi-Constant Voltage) [11] scaling laws are used instead. The QCV law provides optimal driving capability in digital circuits [11] and optimal overall performance in analog circuits [12].

Scaled MOS devices are most useful if the long channel characteristics can be maintained. A generalized scaling criterion [13] can be employed to monitor the usefulness of scaling laws.

In this chapter, scaling objectives, the scaling strategies, i.e. CE, CV, QCV, and a generalized criterion will be discussed. Device parameters after scaled by CE, CV and QCV laws and their application in analog building blocks will be discussed in the next chapter.

2.1. Objectives of Scaling

Improved performance of the MOS devices and a reduction of the RC time constants associated with the parasitic elements by scaling the device dimensions allow fast access time and a higher clocking rate in digital circuits. Practical data exhibiting these improvements have been reported numerously elsewhere.

In MOS analog applications, particularly in sampled data analog circuits such as switched capacitor filters, CCD's etc, one can take advantage of scaled MOS devices to extend the frequency range into the MHz range [14, 15] for applications such as clock recovery in data communication

systems, intermediate frequency in communication receivers, video processing in television receivers, and signal processing in LAN (Local Area Network). The extended capabilities are obtained not only from a reduction in RC time constants but also smaller interelectrode capacitances e.g. C_{gs}, which result in higher device unity gain bandwidth. Analog building blocks such as op-amps and comparators can run at faster speed with smaller area and lower power consumption. (Chapter 3).

One important advantage of the scalable analog circuits is the use of a power supply voltage which is compatible with digital technology. More analog functions can thus be integrated with digital circuits on the same chip. This higher level integration has been the major thrust driven by area reduction, yield increase and cost reduction using scaling technologies.

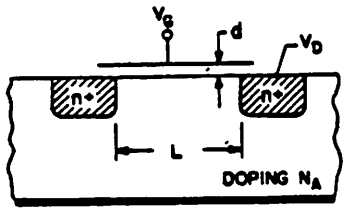
Several scaling laws that alleviate the short channel effects (Chapter 7) will be discussed in the following section.

2.2. Scaling Laws

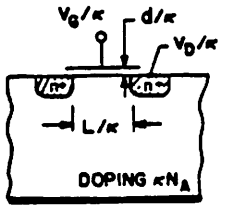
2.2.1. Constant Field Scaling [10]

Table 2.1 lists three scaling laws, namely, CE, CV and QCV. In CE or constant field scaling [10], horizon dimensions (channel width W and channel length L), vertical dimension (gate oxide thickness t_{ox}) and power supply are scaled by the same factor k. The substrate doping is increased by a factor of k. Fig. 2.1 shows the cross section of a CE scaled MOS device. Because all voltages including the threshold voltage, and dimensions including the depletion layer width are reduced by the same factor k, the internal electric field is unchanged after CE scaling.

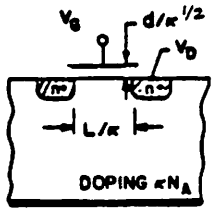
According to CE scaling, the maximum available output voltage becomes smaller, and it is not compatible with the existing TTL digital circuits. Because KT/C noise and 1/f noise are not scaled (Chapter 3), a smaller output swing results which severely degrades the signal to noise ratio in MOS analog applications. Another problem associated with CE scaling is the severe decrease of



(a)



(b)



(c)

Fig. 2.1 MOS device cross section
(a) before scaling
(b) after CE scaling
(c) after CV scaling

Device Parameters	Constant Field	Constant Voltage	Quasi-Constant Voltage
Voltage	1/k	1	$1/k^{1/2}$
Horizontal dimensions	1/k	1/k	1/k
Gate oxide thickness	1/k	$1/k^{1/2}$	1/k
Doping Concentration	k	k	k

Table 2.1 Definition of scaling laws

the drain current when the channel length is shorter than 1 μm due to impurity scattering effects [11]. The problems of TTL incompatibility, decrease of the drain current and degradation of the dynamic range by CE scaling can be solved by non-constant field scaling.

2.2.2. Constant Voltage Scaling

MOS devices by non-constant field scaling exhibit better current driving capability [11] and wider bandwidth than CE scaling. However non-constant field scaling can lead to high drain field, hot electron related problems and operation close to drain breakdown (Chapter 7).

In the CV or Constant Voltage scaling law, as its name implies, the circuit voltage is not scaled. Horizontal dimensions (channel length L and channel width W) and substrate doping N_B are scaled by the same factors as those of CE scaling (Table 2.1). The vertical dimension however is scaled by a smaller factor $\sqrt{k}$. The factor $\sqrt{k}$ has been chosen as most of the literatures [10 , 11 , 12 , 19] and does not have any physical significance. The increase of drain current I_D and device conductance g_m results in better driving capability in digital circuits and larger unity gain bandwidth in the MOS op-amp design.

For a short channel length (L < 2.5 μ), the reduction of drain current due to the saturation of the drift velocity becomes significant. It is especially a problem for submicron devices (0.5 μ) with CV scaling, because the combined effects of the velocity saturation and the degenerated source resistance reduce the drain current severely [11]. QCV is aimed to relieve both constraints especially when the channel length is shrunk to submicron region.

2.2.3. Quasi Constant Voltage Scaling [11]

Horizontal dimensions of QCV or Quasi Constant Voltage case are again scaled by the same factor as those of CE and CV cases (Table 2.1). However the voltage is reduced by a smaller factor $\sqrt{k}$ than that of CV law, and the oxide thickness is shrunk by the same factor k as that of the CE case (Table 2.1). In QCV scaling, the drain current is increasing down to submicron channel lengths, L= 0.5 μ and L= 0.3 μ for NMOS and PMOS transistors respectively [11]. QCV is therefore particularly important when narrow width (≃ 1μ) conductors are used because of the tendency of the conductor RC time constant to increase (chap 7).

2.2.4. Generalized Scaling Criterion [13]

In order to make the best compromise between the device electrical characteristics and process complexity when channel length is reduced, scaling parameter factors can be different from those for CE, CV and QCV in Table 2.1. In this case, a generalized scaling criterion [13] can be applied to monitor whether the resulting MOS device still maintains the long channel characteristics by any scaling laws (Chapter 7). The index of the generalized scaling law, M [12], is defined as

$$M = \frac{A[(x_j t_{ox})(W_S + W_D)^2]^{\frac{1}{3}}}{L} \tag{2.1}$$

where x_j is junction depth in μm, t_{ox} is the oxide thickness in Å, A is a constant; W_S and W_D are depletion widths of the source and the drain respectively, and are given by

$$W_S = [\frac{2\varepsilon_s}{qN_B}(2\phi_B - V_B]^{\frac{1}{2}} \tag{2.2}$$

$$W_D = [\frac{2\varepsilon_s}{qN_B}(2\phi_B - V_B + V_D]^{\frac{1}{2}} \tag{2.3}$$

where $2\phi_B$ is the built in voltage of the source-substrate junction, V_B is the substrate bias and V_D is the drain bias. For NMOS device, A is 0.41 $(\text{Å})^{\frac{1}{3}}$. If M ≤ 1, then the scaled device still maintains long channel behavior. On the other hand, if M > 1, then the MOS device has a tendency to have undesirable short channel effects (Chapter 7).

2.3. Summary

Major objectives for scaling MOS devices in VLSI technology are improvement in performance, increase in component density allowing higher circuit complexity, and reduction of cost.

CE scaling has been used to alleviate the problems associated with short channel effects (Chapter 7). However, considerations on output TTL compatibility, driving capability and dynamic range for shorter channel length MOS devices in the VLSI circuit applications lead to the use of the non-constant field, namely CV and QCV scalings. For the submicron range (e.g. 0.5 μ), QCV is favorable because of concerns about the decrease of drain current due to saturation of carrier velocity and degenerated source resistance.

CHAPTER 3

SCALING OF MOS ANALOG CIRCUITS

Most studies of MOS device scaling in circuits have been for digital applications [10 , 11 , 16 , 17 , 18]. Only a few discussions particularly focus on the impact of the scaled MOS devices on analog performance [12 , 19]. In this chapter, scaled MOS devices under the CE, CV and QCV laws will be discussed first. In order for a quick observation of the characteristics of scaled MOS devices, first order equations are employed similar to the approach by Dennard et. al. [10]. A two dimensional model [12 , 19] or even a three dimensional model [20] is necessary to accurately describe scaled MOS devices if field dependence, carrier drift velocity, short channel effects, etc. are considered. For channel lengths less than about 2 μm, two dimensional models can model the wide channel width scaled devices with reasonable accuracy [12 , 19] but not the narrow devices [20]. Narrow channel width geometry MOS devices need to be described by a three dimensional model for channel length less than about 2 μm [20 , 21].

Resistance, RC time constant, etc. of interconnections under scaling will be also discussed in this chapter. Scaling of interconnections is straightforward. However it places an ultimate limit on the overall scaling when the conductor width is reduced to less than 1 μm (Chapter 7).

Finally, applying the derived parameters of scaled MOS devices, the tendency of scaled op-amp performance such as gain and unity gain frequency can be easily observed.

3.1. Device Scaling

To a first order approximation, it is assumed that mobility degradation, hot electron effects, short and narrow channel effects, etc. (Chapter 7) are not considered in scaled MOS devices. This implies that scaled MOS devices still maintain long channel characteristics. For a channel length scaled down to 1.75 μ [12], first order equations can reasonably describe the trends of the

parameters under scaling except in the CV case.

Device parameters discussed in the following sections are corresponding to a MOS device operated in the saturation region. The notations, k_{WL}, k_v and k_{ox} are scale factors for decreasing the channel width and channel length, voltage, and oxide thickness respectively. k_b is the scale factor for increasing the substrate doping. All the parameters without ' are for scaled MOS devices.

Table 3.1 lists the device parameters scaled by CE, CV and QCV laws.

3.1.1. Threshold Voltage

The threshold voltage V_{TO} is given by

$$V_{TO} = \phi_{ms} + 2\phi_F - \frac{Q_{ss} + Q_B}{C_{ox}} \pm \frac{qD_i}{C_{ox}} \tag{3.1}$$

where ϕ_{ms} is the work function difference between gate and bulk materials, $2\phi_F = 2\frac{KT}{q}\ln\frac{N_B}{n_i}$ is surface potential, Q_{ss} is the fixed oxide charge, Q_B is the bulk depletion charge, C_{ox} is the gate oxide capacitance and D_i is the implant dose. The threshold voltage has to be scaled accordingly with other voltages. Although the term $\phi_{ms} + 2\phi_F$ can not be scaled proportionally, the scaling of the threshold voltage can still be achieved by adjusting the dose of implantation D_i such that the resulting threshold voltage V_{TO} equals to

$$V_{TO} = \frac{1}{k_v} V'_{TO} \tag{3.2}$$

3.1.2. Drain Current

The drain current I_{Dsat} in saturation after scaling is given by:

$$I_{Dsat} = \frac{\mu C_{ox}}{2}\left(\frac{W}{L}\right)(V_{GS} - V_T)^2 \tag{3.3a}$$

$$= \frac{k_{ox}}{k_v^2} I'_{Dsat} \tag{3.3b}$$

For CE scaling, drain saturation current is scaled down by a factor of k. For non-constant field

Device Parameters	Scaling Factor	CE	CV	QCV	Equation
V_{TO}	$\frac{1}{k_v}$	$1/k$	1	$1/k^{1/2}$	3.2
I_{Dsat}	$\frac{k_{ox}}{k_v^2}$	$1/k$	$k^{1/2}$	1	3.3b
g_m	$\frac{k_{ox}}{k_v}$	1	$k^{1/2}$	$k^{1/2}$	3.4b
r_d	$\frac{k_b^{1/2}k_v^{3/2}}{k_{WL}k_{ox}}$	1	$1/k$	$1/k^{3/4}$	3.5b
C_{gs}	$\frac{k_{ox}}{k_{WL}^2}$	$1/k$	$1/k^{3/2}$	$1/k$	3.6b
f_T	$\frac{k_{WL}^2}{k_v}$	k	k^2	$k^{3/2}$	3.7b
C_{gd}	$\frac{1}{k_{WL}}$	$1/k$	$1/k$	$1/k$	3.8b
$C_{db} \cong C_D$	$\frac{k_b^{1/2}k_v^{1/2}}{k_{WL}^2}$	$1/k$	$1/k^{3/2}$	$1/k^{5/4}$	3.9b
$C_{sb} \cong C_D$	$\frac{k_b^{1/2}k_v^{1/2}}{k_{WL}^2}$	$1/k$	$1/k^{3/2}$	$1/k^{5/4}$	3.9b
(RC)	$\frac{k_b^{1/2}k_v^{3/2}}{k_{ox}k_{WL}^2}$	$1/k$	$1/k^2$	$1/k^{7/4}$	3.10
P	$\frac{k_{ox}}{k_v^3}$	$1/k^2$	$k^{1/2}$	$1/k^2$	3.11
$\frac{P}{WL}$	$\frac{k_{ox}k_{WL}^2}{k_v^3}$	1	$k^{5/2}$	$k^{3/2}$	3.12
S	$\frac{1+\frac{k_b^{1/2}k_v^{1/2}}{k_{ox}}(\frac{C_{ch}}{C_{ox}})'}{1+(\frac{C_{ch}}{C_{ox}})'}$	1	1	≤ 1	3.13b
M	$\frac{k_{WL}}{(k_f k_{ox} k_t k_v)^{1/3}}$	$1/k^{1/3}$	$k^{1/3}$	$1/k^{1/6}$	3.14b
1/f noise	$\frac{k_{WL}^2}{k_{ox}^2}$	1	k	1	3.15b
Thermal noise	$\frac{k_v}{k_{ox}}$	1	$1/k^{1/2}$	$1/k^{1/2}$	3.16b

Table 3.1 Scaling results of device parameters

scaling, drain saturation current is unchanged for the QCV law and increases by a factor of $\sqrt{k}$ for the CV law.

3.1.3. Transconductance

Scaled transconductance g_m is given by

$$g_m = \frac{\partial I_D}{\partial V_{GS}} = \sqrt{2\mu C_{ox}(\frac{W}{L})I_D} \tag{3.4a}$$

$$= \frac{k_{ox}}{k_v} g_m' \tag{3.4b}$$

Transconductance is unchanged for CE law, but increases by a factor of $\sqrt{k}$ under both CV and QCV scalings.

3.1.4. Output Resistance

To a first order approximation, the device output resistance r_d is given by

$$r_d = \frac{1}{\lambda I_D} = \frac{1}{\frac{\partial I_D}{\frac{\partial V_{DS}}{L_{eff}} I_D}} = \frac{\sqrt{N_b(V_D - V_{Dsat})}\, L_{eff}}{\sqrt{\frac{\varepsilon_s}{2q}}\, I_D} \tag{3.5a}$$

$$= \frac{k_b^{\frac{1}{2}} k_v^{\frac{3}{2}}}{k_{WL} k_{ox}} r_d' \tag{3.5b}$$

where λ is the channel modulation coefficient, L_{eff} is the effective channel length. According to this simplified model, device output resistance is unchanged by CE scaling, however it decreases by k and $k^{\frac{3}{4}}$ for CV and QCV laws respectively. A decrease of the output resistance is undesirable for analog circuits because it results in op-amp gain degradation.

Output resistance r_d of a scaled MOS device is the most difficult parameter to model especially for narrow devices [20]. In (3.5a) only channel modulation is considered. No model has been developed thus far that can practically describe scaled device output resistance well enough if the DIBL (Drain Induced Barrier Lowering) [16] and hot electron effects are present.

3.1.5. Interelectrode Capacitance

Gate-source capacitance C_{gs} is given by

$$C_{gs} = \frac{2}{3}C_{ox}WL + C_{ox}WL_{ov} \approx \frac{2}{3}C_{ox}WL \tag{3.6a}$$

$$\doteq \frac{k_{ox}}{k_{WL}^2}C'_{gs} \tag{3.6b}$$

where L_{ov} is the gate-source overlap length in parallel to the direction of current flow as shown in Fig. 3.1 and assumed to be scaled by the same factor as gate oxide thickness and $C_{ox}WL_{ov}$ is the gate to source overlap capacitance, assumed much smaller than the contribution of the gate oxide capacitance $\frac{2}{3}C_{ox}WL$.

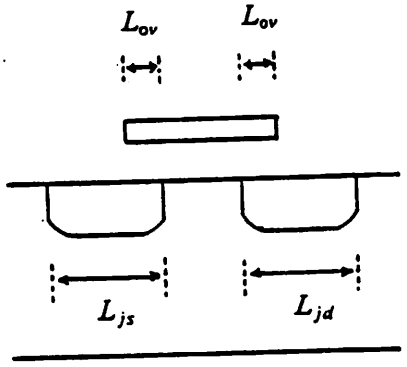


Fig. 3.1 Diffusion junction length L_j, gate to drain/source overlap length L_{ov} of a MOS device

C_{gs} reduces under CE, CV and QCV laws. Hence the unity gain frequency f_T of a MOS device increases because f_T is given by

$$f_T = \frac{g_m}{C_{gs}} = \frac{k_{ox}}{k_v} \cdot \frac{k_{WL}^2}{k_{ox}} f_T \tag{3.7a}$$

$$= \frac{k_{WL}^2}{k_v} f_T \tag{3.7b}$$

The fundamental limitation of fast analog circuits is the unity gain frequency of a MOS device f_T. The increase of the MOS device unity gain frequency f_T therefore allows MOS analog circuits to be faster.

The gate to drain/source overlap capacitance C_{gdo}/C_{gso} is given by

$$C_{gdo} = C_{ox} W L_{ov} \tag{3.8a}$$

$$= \frac{1}{k_{WL}} C'_{gdo} \tag{3.8b}$$

where L_{ov} is the gate to drain/source overlap length in parallel to the direction of current flow (Fig. 3.1). It is assumed to be scaled by the same factor as gate oxide thickness. C_{gdo}/C_{gso} reduces by all scaling laws, CE, CV and QCV and it is dependent only on the horizontal dimension scaling factor k_{WL}. The clock feedthrough will be smaller due to the decrease of the overlap capacitance.

Drain/source to substrate diffusion capacitance C_{db}/C_{sb} is generalized in notation as C_D in the following equation:

$$C_D = C_d W L_j = \frac{\varepsilon}{x_d} W L_j \tag{3.9a}$$

$$= \frac{k_b^{\frac{1}{2}} k_v^{\frac{1}{2}}}{k_{WL}^2} C'_D \tag{3.9b}$$

where C_d is the junction capacitance per unit area, L_j is the diffusion junction length in parallel to the direction of current flow (Fig. 3.1) assumed to be scaled by the same factor of channel width and channel length, and $x_d = \frac{1}{k_b^{\frac{1}{2}} k_v^{\frac{1}{2}}} x'_d$ is the depletion width between the diffusion and the substrate. From (3.9b), the junction capacitance becomes smaller after scaling. The device delay times

of the DC (large) signal and the AC (small) signal, $(RC)_{DC}$ and $(RC)_{AC}$, are given in the following equations respectively:

$$(RC)_{DC} \equiv (\frac{V_{DS}}{I_D} C_D) = \frac{k_v}{k_{ox}} \frac{k_b^{\frac{1}{2}} k_v^{\frac{1}{2}}}{k_{WL}^2} (\frac{V_{DS}}{I_D} C_D)' = \frac{k_b^{\frac{1}{2}} k_v^{\frac{3}{2}}}{k_{ox} k_{WL}^2} (RC)'_{DC} \tag{3.10a}$$

$$(RC)_{AC} \equiv (\frac{1}{g_m} C_D) = \frac{1}{\frac{k_{ox}}{k_v}} \frac{k_b^{\frac{1}{2}} k_v^{\frac{1}{2}}}{k_{WL}^2} (RC)'_{AC} = \frac{k_b^{\frac{1}{2}} k_v^{\frac{3}{2}}}{k_{ox} k_{WL}^2} (RC)'_{AC} \tag{3.10b}$$

According to (3.10a) and (3.10b), the device large and small signal delay times, $(RC)_{DC}$ and $(RC)_{AC}$, are scaled by the same factor and become smaller after scaling. The decrease of the delay times by the scaling is mainly because of the reduction of the diffusion capacitance. Notice that the device large and small signal resistances, $\frac{V_{DS}}{I_D}$ and $\frac{1}{g_m}$, are unchanged by CE law.

3.1.6. Device Power Dissipation and Density

Among all three scaling laws under study, only the CV scaling increases the power dissipation of both large (DC) and small (AC) signals. The DC and AC device power dissipations, P_{DC} and P_{AC}, can be expressed as follows:

$$P_{DC} \equiv (V_{DS} I_D) = \frac{1}{k_v} \frac{k_{ox}}{k_v^2} P'_{DC} = \frac{k_{ox}}{k_v^3} P'_{DC} \tag{3.11a}$$

$$P_{AC} \equiv \frac{1}{2} CV^2 f = \frac{k_{ox}}{k_{WL}^2} \frac{1}{k_v^2} \frac{k_{WL}^2}{k_v} P'_{AC} = \frac{k_{ox}}{k_v^3} P'_{AC} \tag{3.11b}$$

where f is the clock frequency and is assumed to be increased by the same factor as the device unity gain frequency f_T (Equation 3.7).

From (3.11a) and (3.11b), device power densities are expressed as

$$(\frac{P_{DC}}{WL}) = \frac{k_{ox}}{k_v^3} \frac{1}{\frac{1}{k_{WL}^2}} (\frac{P_{DC}}{WL})' = \frac{k_{WL}^2 k_{ox}}{k_v^3} (\frac{P_{DC}}{WL})' \tag{3.12a}$$

$$(\frac{P_{AC}}{WL}) = \frac{k_{ox}}{k_v^3}\frac{1}{\frac{1}{k_{WL}^2}}(\frac{P_{AC}}{WL})' = \frac{k_{WL}^2 k_{ox}}{k_v^3}(\frac{P_{AC}}{WL})' \qquad (3.12b)$$

Note the scaling factors of the device power dissipations/densities for both large and small signals are the same. Because device DC and AC power densities are unchanged by CE scaling, the cooling method for the circuit need not be changed even though more complicated circuits are integrated on the same chip. On the other hand, CV and QCV laws increase the power densities making thermal dissipation a concern in the chip complexity, reliability and packaging.

3.1.7. Subthreshold Current

The channel weak inversion current I_{subs} for $V_G < V_T$ and $V_D > \frac{3KT}{q}$ is important in circuit applications, because it determines how fast a MOS switch can turn off and what minimum power supply voltage is required to turn off the switch such that stored information will not leak away. This is especially important for charge storage circuits such as switched capacitor filters.

Because the subthreshold current depends on the gate voltage exponentially, subthreshold current is not scaled linearly. A commonly used figure of merit to describe the subthreshold behavior is the subthreshold swing S which is defined as the gate voltage below V_T needed to reduce the current by one decade (Chap 7).

$$S \equiv ln10 \cdot \frac{dV_G}{dlnI_D} \approx \frac{KT}{q} ln10 \cdot (1 + \frac{C_{ch}}{C_{ox}}) \qquad (3.13a)$$

$$\approx \frac{KT}{q} ln10 \cdot [1 + \frac{k_b^{\frac{1}{2}} k_v^{\frac{1}{2}}}{k_{ox}}(\frac{C_{ch}}{C_{ox}})'] \qquad (3.13b)$$

where C_{ch} is the channel depletion layer capacitance per unit area and C_{ox} is the gate oxide capacitance per unit area. From (3.13b), it can be seen that the subthreshold swing S either remains unchanged for CE and CV scalings or decreases slightly for QCV scaling. Therefore, the subthreshold current is either unchanged or reduced slightly after device is scaled. Hence, subthreshold current is a practical limiting factor in scaling MOS devices which will be discussed in Chapter 7.

Following the discussion of the generalized scaling criterion, the scaling index M, equation (2.1), was shown to be another way to describe how much a scaled device deviates from long channel subthreshold behavior.

$$M = \frac{A[(x_j t_{ox})\ (W_S + W_D)^2]^{\frac{1}{3}}}{L} \tag{3.14a}$$

$$= k_{WL}\left(\frac{1}{k_j k_{ox} k_b k_v}\right)^{\frac{1}{3}} M' \tag{3.14b}$$

Scaling index M of the CV law is the largest among all scaling laws discussed, therefore CV scaling approaches the subthreshold limitation more rapidly than CE and QCV scalings.

3.1.8. Noise

MOS devices exhibit a large noise component at low frequency due to 1/f noise. A commonly used 1/f noise power density is given by

$$\frac{\overline{v_f^2}}{\Delta f} = \frac{K_f}{WLC_{ox}^2 f} \tag{3.15a}$$

$$= \frac{k_{WL}^2}{k_{ox}^2}\left(\frac{\overline{v_f^2}}{\Delta f}\right)' \tag{3.15b}$$

1/f noise is not scalable under CE, CV and QCV laws, because the 1/f noise scaling factor $\frac{k_{WL}^2}{k_{ox}^2}$ is never less than 1. Thus 1/f noise is a practical limiting factor for scaled MOS devices in the design of analog circuits (Chapter 4).

MOS devices also exhibit thermal noise due to the resistive channel. The thermal noise is given by

$$\left(\frac{v_{th}^2}{\Delta f}\right) = 4KT\left(\frac{2}{3g_m}\right) \tag{3.16a}$$

$$= \frac{k_v}{k_{ox}}\left(\frac{v_{th}^2}{\Delta f}\right)' \tag{3.16b}$$

Device thermal noise decreases for the CV and QCV cases, and it remains unchanged for the CE

case. However the thermal noise contribution to a switched capacitor integrator is also associated with other circuit parameters (Chapter 4).

3.2. Capacitor Scaling

In MOS technology, a capacitor can be formed by two layers of heavily doped polysilicon, metals or a combination of the two. Thermally grown oxide or nitride can be used as the dielectric. Scaling of capacitors is straight forward just by changing the dielectric thickness and the capacitor area, i.e.

$$C = \frac{\varepsilon}{t_d} A \tag{3.17a}$$

$$= \frac{k_d}{k_A^2} C' \tag{3.17b}$$

where ε is the dielectric constant, t_d is the thickness of the dielectric material, A is the area of the capacitor, and k_d and k_A are scaling factors for dielectric thickness and capacitor area respectively. The minimum geometry of a capacitor is determined by matching and accuracy of the process technology available, and the KT/C noise of circuit performance required.

3.3. Resistor Scaling

Resistors are realized by either n^+/p^+ diffusion or polysilicon. The latter is favorable because it has less parasitic capacitance. Scaling of resistors is again straightforward. However accuracy and matching requirements limit resistor scaling.

3.4. Interconnection Scaling

Line delay, current density, and signal drop are major issues when interconnection lines are scaled. Other problems such as contact resistance, line fringe effects, line to line capacitance, etc. are fundamental limits to scaling (Chapter 7).

A line delay RC time constant is given by

$$(RC) = \rho\frac{l}{\omega t}\,\omega l C_{ox} \tag{3.18a}$$

$$= \rho\frac{l^2}{t}C_{ox} \tag{3.18b}$$

$$= \frac{k_{ll}k_{ox}}{k_{\omega l}^2}(RC)' \tag{3.18c}$$

where k_{ll}, k_{ox}, and $k_{\omega l}$ are the line thickness, line-substrate oxide thickness and line width/length scaling factors respectively. From (3.18b), the line delay RC time constant is proportional to the square of line length l, inversely proportional to the line thickness t and independent of the line width ω. If line thickness t remains unchanged ($k_{ll} = 1$), then line delay time constants are reduced by factors of k, $k^{\frac{3}{2}}$ and k for the CE, CV and QCV cases respectively. On the other hand, if thin film is required as the patterns become smaller, line delay time constants are unchanged for CE and QCV scalings, assuming line thickness is reduced by the same factor as oxide thickness. In this case ($k_{ll} \neq 1$), line delay constants are even larger if the interconnection lengths are not scaled, or even longer for more complex circuit. Then, other alternatives like refractory metal silicides [22] are necessary to be used.

Line current density $\frac{I}{A}$ is given by

$$\frac{I}{A} = \frac{k_{ox}k_{\omega l}k_{ll}}{k_v^2}(\frac{I}{A})' \tag{3.19}$$

where A is line area cross section. If line thickness t is reduced by the same factor as that of the oxide thickness, line current density are increased by factors of k, k^2 and k^2 under the CE, CV and QCV laws respectively. Even if the line thickness remains the same, line current densities of the CV and QCV cases increase by factors of $k^{\frac{3}{2}}$ and k respectively, the CE line density being unchanged. Increase of line current density by scaling is undesirable because it leads eventually to electromigration (Chapter 7).

Ratio of the signal drop on the line to the signal $\frac{IR}{V}$ becomes larger under CE, CV and QCV laws.

$$\frac{IR}{V} = \frac{k_{ox}k_{li}}{k_v} \tag{3.20}$$

However for moderate scaling, this is not a major problem. Table 3.2 lists the results of interconnection line scaling.

3.5. An Analog Building Block - The Operational Amplifier

Applying the derived scaled parameters in section 3.1, the effects of scaling on an analog basic building block, namely the op-amp, will be discussed in this section. A typical two-stage CMOS operational as shown in Fig. 3.2 is used as an example to demonstrate the performance variations due to CE, CV and QCV scaling laws. Such an op-amp is chosen because it is well discussed in the literature [23]. Table 3.3 lists scaling results on the performance of the two stage op-amp shown in Fig. 3.2

Line Thickness	Parameters	Scaling Factor	CE	CV	QCV
$k_{li} = k_{ox}$	RC	$\frac{k_{ox}^2}{k_\omega l^2}$	1	1/k	1
	$\frac{I}{A}$	$\frac{k_{ox}^2 k_\omega l}{k_v^2}$	k	k^2	k^2
	$\frac{IR}{V}$	$\frac{k_{ox}^2}{k_v}$	k	k	$k^{3/2}$
unchanged	RC	$\frac{k_{ox}}{k_\omega l^2}$	1/k	$1/k^{3/2}$	1/k
	$\frac{I}{A}$	$\frac{k_{ox} k_\omega l}{k_v^2}$	1	$k^{3/2}$	k
	$\frac{IR}{V}$	$\frac{k_{ox}}{k_v}$	1	$k^{1/2}$	$k^{1/2}$

Table 3.2 Scaling results of interconnection lines

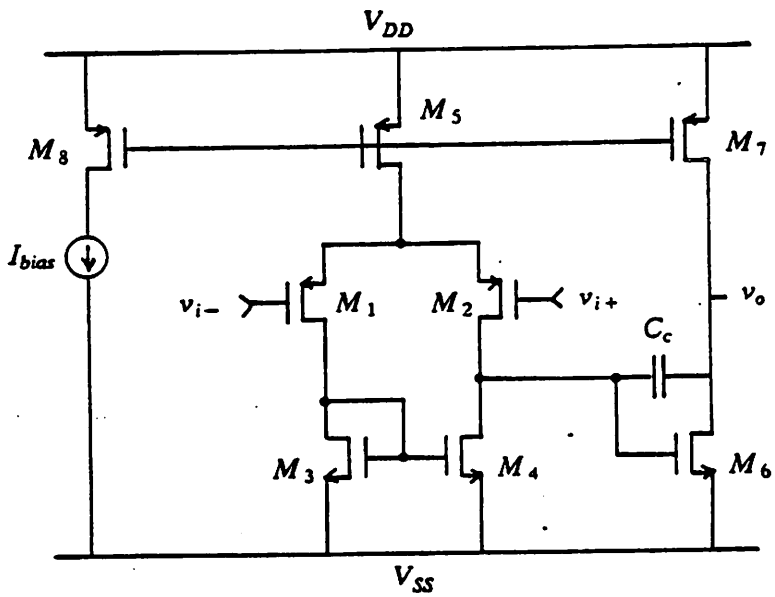


Fig. 3.2 Schematic of a basic two stage CMOS operational amplifier

Op-amp Performance	Scaling Factor	CE	CV	QCV	Equation
A_v	$\frac{k_b k_v}{k_{WL}^2}$	1	1/k	$1/k^{1/2}$	3.21b
ω_1	$\frac{k_{WL}^2}{k_v}$	k	k^2	$k^{3/2}$	3.22b
P_2	$\frac{k_{WL}^2}{k_v}$	k	k^2	$k^{3/2}$	3.23b
SR	$\frac{k_{WL}^2}{k_v^2}$	1	k^2	k	3.24
$\frac{SR}{V}$	$\frac{k_{WL}^2}{k_v}$	k	k^2	$k^{3/2}$	3.25
1/f Noise	$\frac{k_{WL}^2}{k_{ox}^2}$	1	k	1	3.27b
Thermal Noise	$\frac{k_v}{k_{ox}}$	1	$1/k^{1/2}$	$1/k^{1/2}$	3.28b

Table 3.3 Scaling results of performance of the op-amp as shown in Fig. 3.2

3.5.1. Open Circuit Voltage Gain

Overall open loop voltage gain A_v is the product of the gains of the first stage A_{v1} and the second stage A_{v2} i.e.

$$A_v = A_{v1}A_{v2} = \frac{g_{m1}}{g_{o2} + g_{o4}} \frac{g_{m6}}{g_{o6} + g_{o7}} = k_{g_m}^2 k_r^2 A_v' \tag{3.21a}$$

$$= \frac{k_b k_v}{k_{WL}^2} A_v' \tag{3.21b}$$

Op-amp open loop voltage gain is unchanged under CE scaling, but it decreases by k and $\sqrt{k}$ for the CV and QCV cases respectively. This is because in the CV and QCV cases, the decrease of the output resistance is larger than the increase of the device transconductance. If substrate current is present, the op-amp gain will be further degraded.

3.5.2. Unity Gain Frequency

Assume the internal parasitic capacitor C_1 is much smaller than the compensation capacitor C_c and the load capacitor C_2 and suppose the relative location of first non-dominant pole P_2 to

unity gain ω_1 i.e. $\frac{P_2}{\omega_1} = \frac{g_{m2}}{g_{m1}}\frac{C_c}{C_2}$ is unchanged after scaling. Then the compensation capacitance C_c is scaled by the same factor as that of output load capacitor C_2. The unity gain frequency after scaling is thus given by

$$\omega_1 = \frac{g_{m1}}{C_c} \tag{3.22a}$$

$$= \frac{k_{WL}^2}{k_v}\omega_1' \tag{3.22b}$$

Unity gain frequency increases by k, k^2 and $k^{\frac{3}{2}}$ respectively for the CE, CV and QCV cases respectively. Switched capacitor signal processing circuits, for instance, can run at a higher clocking frequency if op-amp has higher unity gain frequency.

The higher the first nondominant pole P_2, the higher the unity gain frequency. The first nondominant pole P_2 is given by

$$P_2 = \frac{-g_{m2}}{C_2} \tag{3.23a}$$

$$= \frac{k_{WL}^2}{k_v}P_2' \tag{3.23b}$$

Again the first nondominant pole P_2 increases by the same factor as the unity gain frequency when scaling. Notice that both the unity gain frequency and the first nondominant pole are proportional to the square of the horizontal scaling factor as expressed in (3.22b) and (3.23b).

3.5.3. Slew Rate

The slew rate of this op-amp is given by

$$SR = \frac{I_{D1}}{g_{m1}}\omega_1 \tag{3.24a}$$

$$= \frac{k_{WL}^2}{k_v^2}(SR)' \tag{3.24b}$$

The slew rate of the CE case remains unchanged. However it increases for the CV and QCV laws. Normalized slew rate is given by

$$\frac{SR}{V} = \frac{k_{WL}^2}{k_v}\left(\frac{SR}{V}\right)' \tag{3.25}$$

The normalized slew rate increases for all three scalings considered.

3.5.4. Noise Performance

The op-amp (Fig. 3.2) equivalent input noise power density can be calculated by the following equation:

$$v_{eqtot}^2 = v_{eq1}^2 + v_{eq2}^2 + \left(\frac{g_{m3}}{g_{m1}}\right)^2(v_{eq3}^2 + v_{eq4}^2) \tag{3.26}$$

where v_{eqi}^2 is the equivalent gate noise power density of the first stage MOS device M_i.

Using (3.15) for 1/f noise power density in (3.26), the op-amp input referred 1/f noise is given by

$$\left(\frac{v_{1/f}^2}{\Delta f}\right) = \frac{2K_p}{W_1 L_1 C_{ox}^2 f}\left(1 + \frac{K_n \mu_n L_1^2}{K_p \mu_p L_3^2}\right) \tag{3.27a}$$

$$= \frac{k_{WL}^2}{k_{ox}^2}\left(\frac{v_{1/f}^2}{\Delta f}\right)' \tag{3.27b}$$

where K_p and K_n are the flicker noise coefficients of PMOS and NMOS devices respectively. The scaling factor of op-amp 1/f noise is a strong function of device 1/f noise. In fact, it is dependent on the process. In (3.27b), the ratios of device mobility and flicker noise coefficients i.e. $\frac{\mu_n}{\mu_p}$ and $\frac{K_n}{K_p}$ are assumed unchanged after scaling. Nevertheless, 1/f noise has to be removed such that it is not process dependent.

Using (3.16) for device thermal noise in (3.26), the equivalent total op-amp thermal noise at input is given by

$$\frac{v_{ih}^2}{\Delta f} = 4KT\frac{4}{3g_m}\left(1 + \left[\frac{\mu_n\left(\frac{W}{L}\right)_3}{\mu_p\left(\frac{W}{L}\right)_1}\right]^{1/2}\right) \tag{3.28a}$$

$$= \frac{k_v}{k_{ox}}\left(\frac{v_{th}^2}{\Delta f}\right)' \qquad (3.28b)$$

Here a two stage op-amp is used as an example. For a single stage cascode op-amp [25] that will be discussed in Chapter 4, the op-amp performance after scaling has a similar tendency as the one discussed above.

3.6. Summary

Decrease of the interelectrode capacitance gives smaller RC time constants and higher device unity gain frequency resulting in faster digital and analog circuits. On the other hand, a decrease of the output resistance for CV and QCV scalings is undesirable because the op-amp gain will be smaller. The most undesirable scaled parameters are the subthreshold current and 1/f noise because they are not scalable in general. For non-constant field scalings, thermal power dissipation places a fundamental limit on circuit density and reliability.

Under CE scaling, other than the increase of the unity gain frequency due to decrease of the compensation capacitor, most of the op-amp performance parameters are unchanged except the severe degradation of the signal to noise ratio. Non-constant field laws give faster slew rate and higher unity gain frequency, however they result in lower op-amp gain and higher power density than the CE law. Both constant field and non-constant field scaling do not improve signal to noise ratio.

From the summary of the scaled op-amp performance (Table 3.3), the QCV law is the optimum for scaling MOS analog circuits because it offers faster speed and larger signal to noise ratio than the CE law, and it shows larger gain and and lower power dissipation than the CV law. The QCV law is also optimal for digital circuit scaling as proposed by Chatterjee [11]. Therefore, both digital and analog circuits can be integrated on the same chip by using the same scaling law, namely the QCV law.

Scaling of the MOS device and analog circuits has been discussed in this Chapter. In the next chapter, performance degradation caused by scaling MOS analog circuits will be discussed.

CHAPTER 4

LIMITATIONS TO SCALING IN MOS ANALOG CIRCUITS

The achievable density of digital circuits continues to increase in inverse proportion to technological feature size. Analog circuits on the other hand do not scale as easily due to certain fundamental as well as practical limiting factors [12]. The switch noise, commonly referred as KT/C noise, is a fundamental limit. Other practical limiting factors for example are the increase of 1/f noise, the difficulty of achieving a large operational amplifier voltage gain, and the difficulty of achieving a large voltage swing relative to the supply voltage.

In this chapter, some practical factors limiting the scaling of analog circuits will be discussed.

4.1. Voltage Swing

The maximum available output swing V_{max} is dictated by the power supply voltage V_{ps} used and the overdrive voltage $\Delta V = \left\{ \frac{I_D}{\frac{\mu C_{ox}}{2}(\frac{W}{L})} \right\}^{1/2}$ of the MOS transistors at the output stage. The maximum output swing V_{max} decreases if the power supply voltage V_{ps} is scaled. The available maximum output voltage V_{max} is given by

$$V_{max} = V_{ps} - \sum_i \Delta V_i \tag{4.1}$$

where $\sum \Delta V_i$ is the total overdrive voltage at the output stage.

For CE scaling, the maximum available output swing V_{max} is given by

$$V_{max} = \frac{1}{k}\left[V_{ps} - \sum_i \Delta V_i \right] \tag{4.2}$$

where k is the scaling factor.

For CV scaling,

$$V_{max} = V_{ps} - \sum_i \Delta V_i \tag{4.3}$$

As an example, in the case of a cascode output structure with the high swing biasing circuit of Fig. 4.1 [24], the maximum output voltage swing for CV scaling as the definition implies, remains unchanged $V_{max} = V_{ps} - 2\Delta V$. The maximum available output swing in CE scaling however is $V_{max} = \frac{1}{k}\left[V_{ps} - 2\Delta v\right]$ which is scaled by the same process scaling factor k.

4.2. Gain

Op-amp DC gain is determined by the product of the transconductance and the output resistance. For the cascode op-amp as shown in Fig. 4.1. The op-amp gain A_{vo} is given by

$$A_{vo} = \left[g_{mo}R_o\right]^2 \tag{4.4}$$

where g_{mo} and R_o are the device transconductance and output resistance respectively. Because the device transconductance and the output resistance are unchanged after CE scaling (cf Table 3.1), the op-amp gain is also unchanged after the scaling. In the CV case however, although the transconductance g_{mo} increases by a factor of $k^{1/2}$, the output resistance R_o decreases by a factor of k. Thus the gain under the CV scaling decreases. For a cascode op-amp in the CV case, the gain is given by

$$A_v = \left[g_m R\right]^2 = \left[k^{1/2} g_{mo} \frac{R_o}{k}\right]^2 = \frac{\left[g_{mo}R_o\right]^2}{k} = \frac{1}{k}(A_v) \tag{4.5}$$

The discussion here assumes the first order approximation of the transconductance g_{mo} and the output resistance R_o under scaling. The op-amp gain is worse if the short channel effects are considered. The limitations imposed by the short channel effects will be discussed more in detail in Chapter 7. One important practical limitation to the analog circuits due to the short channel effects is impact ionization that is particularly a problem in technologies involving thin oxide and high channel doping. Fig. 4.2 shows the output-input voltage characteristics of a depletion load

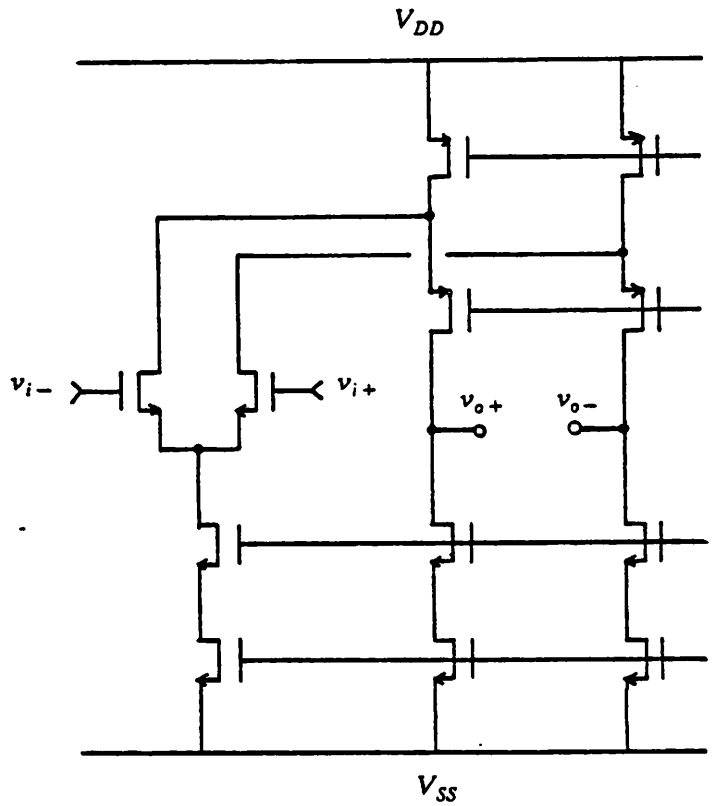


Fig. 4.1 Fully differential cascode op-amp

inverter with and without impact ionization[26]. From Fig. 4.2, it shows that impact ionization reduces noise margin and gain for digital circuits, but overall it has little effect on the digital circuit performance. On the other hand, for analog circuits as shown in Fig. 4.3a, impact ionization reduces the output impedance which hence reduces the op-amp gain (Fig. 4.3b) [26].

The effect of impact ionization on the gain of a simple CMOS inverter amplifier is shown in Fig. 4.4a. As the output voltage increases, the voltage drop between the end of the channel and the drain of device M_1 is also increases. This increases the energy of the carriers that are swept from the source to drain until these carriers have enough energy to ionize some of the lattice atoms on impact. This causes a current flow from the drain to the substrate as shown in Fig. 4.4b. Since this current depends on the output voltage, its effect can be represented as an equivalent resistance R_{eq}

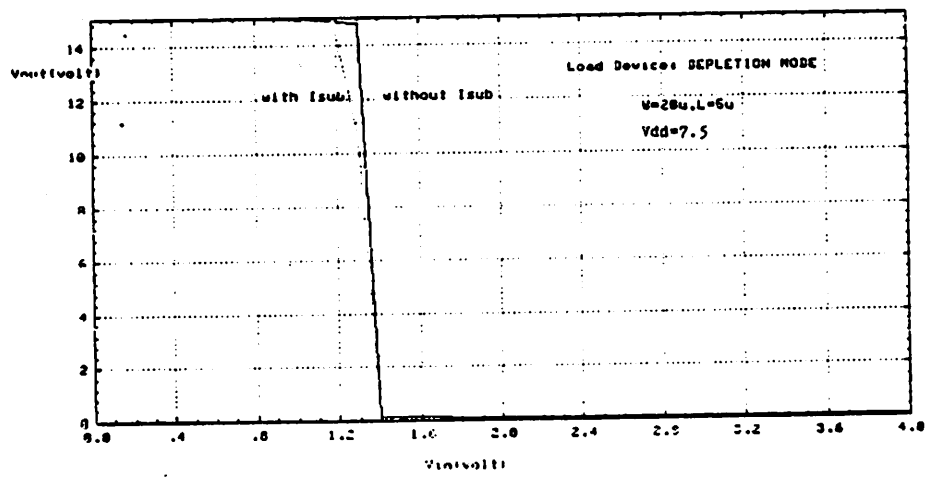
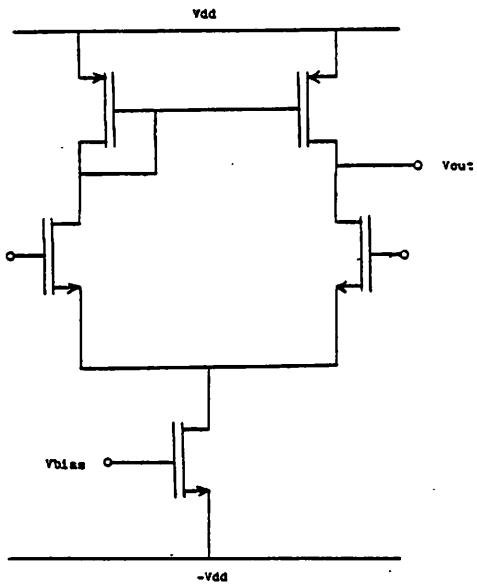
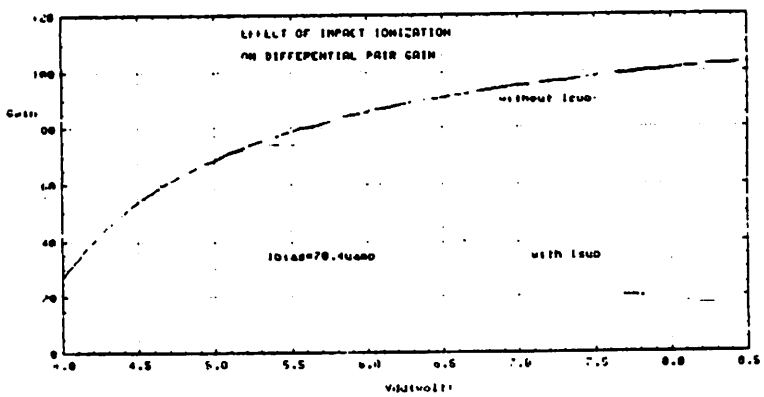


Fig. 4.2 Input-output characteristics of a depletion load inverter with and without impact ionization [26]

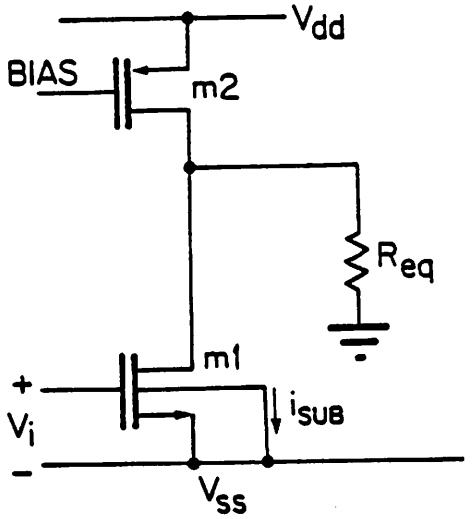


(a)

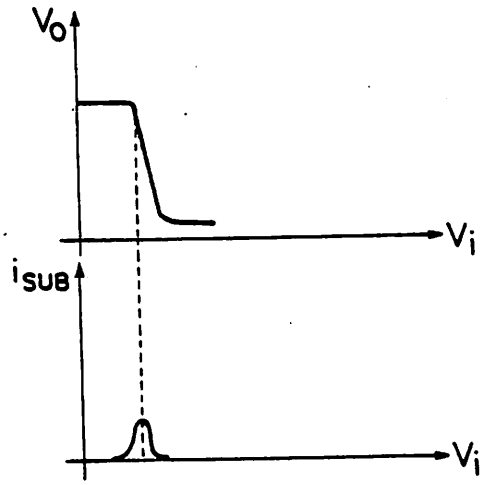


(b)

Fig. 4.3 (a) Simple differential amplifier
(b) Gain versus V_{dd} with and without impact ionization



(a)



(b)

Fig. 4.4 (a) simple inverter amplifier with R_{eq} representing the impact ionization
(b) the substrate current i_{subs} shown with the input output voltage characteristics

in parallel with the output of the op-amp shown in Fig. 4.4a. For a transconductance op-amp, this equivalent output resistance R_{eq} can dramatically degrade the op-amp output resistance and hence the voltage gain. Beside the degradation of the impedance, impact ionization also introduces excess noise and may induce latch up. In order to alleviate the problems caused by impact ionization, suitable circuit solutions have to be used (Chapter 5).

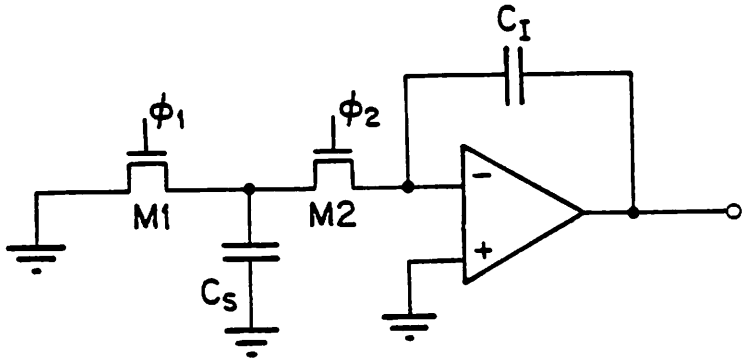
4.3. Noise

The dynamic range is defined as the available *rms* output voltage swing divided by the *rms* noise voltage. The maximum available output swing is governed by the scaling laws as in (4.2) and (4.3) for CE scaling and CV scaling respectively. Therefore, there is an upper bound for the noise given a certain dynamic range requirement when the circuit is scaled.

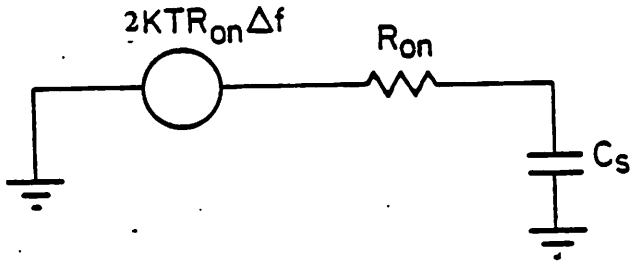
There are three major noise contributions. The first one is op-amp noise which is mainly composed of thermal noise and 1/f noise. The second noise source is the thermal noise of the switches which is commonly referred to as KT/C noise. The last one is the coupled spurious noise, e.g. digital noise which is injected to the circuit via power supply lines or coupling capacitors.

4.3.1. KT/C Noise

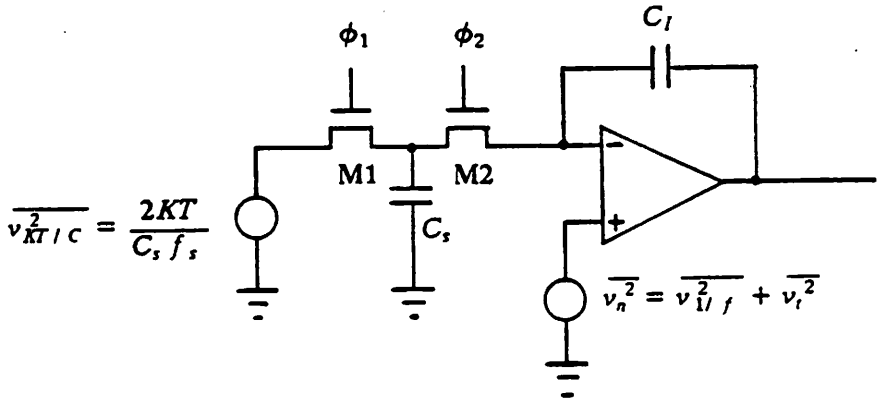
The source of the kT/C noise is shown in Fig. 4.5 for the case of a switched-capacitor integrator. The finite channel resistance of the MOS switch produces a transient thermal noise which will be sampled by the sampling capacitor. As shown in Fig. 4.5a during phase 1, the sampling capacitor C_s is connected to the input. During this time the voltage across C_s is equal to the input signal plus the noise due to the on-resistance of the MOS switch. At the end of phase 1, the input switch is open and C_s samples both the input signal and the noise. The variance of the noise samples $E(v_{cs}^2)$ that appear across the sampling capacitor C_s is equal to the integral over the entire frequency range of the white noise power density which is associated with the on-resistance of the switch bandlimited by the frequency response of the sampling circuit. This gives a noise variance



(a)



(b)



(c)

Fig. 4.5 (a) Switched capacitor integrator
(b) KT/C noise equivalent circuit when M_1 is on and M_2 is off
(c) Equivalent input KT/C noise power density of an integrator

equal to kT/C_s [27] in each sample appearing across sampling capacitor C_s independent of the switch resistance. This sampled noise depends only on the size of the capacitor C_s and therefore is usually referred to as KT/C noise. Table 4.1 gives the *rms* values of the kT/C noise samples for different capacitor sizes. Note the equivalent input KT/C noise power density due to switch M_1 (Fig. 4.5a) is $\frac{KT}{C_s f_s}$ for frequency $-\infty$ to $+\infty$ [27,29].

The right hand side switch M_2 (Fig. 4.5a) also contributes noise which is sampled by the sampling capacitor C_s. In this case, the resulting noise, strictly speaking, can not be considered as a first order low pass filtered noise, because the circuit which samples the switch noise of M_2 contains the op-amp. Assuming an ideal op-amp (with infinite bandwidth and gain) and, as in most of the practical cases, $\frac{C_s}{C_I} \ll 1$, the equivalent input KT/C noise power density contributed by switch M_2 is the same as that by switch M_1 i.e. $\frac{KT}{f_s C_s}$ [27]. Therefore, the total equivalent input KT/C noise density of an integrator for frequency range from $-\infty$ to $+\infty$ due to both switches M_1 and M_2 represented as a generator in Fig. 4.5c is

$$\frac{\overline{v^2_{KTC}}}{\Delta f} = \frac{2KT}{C_s f_s} \equiv 2KTR_s \tag{4.6}$$

where R_s is equivalent to $\frac{1}{C_s f_s}$. The corresponding in band output noise of the integrator [27,29] is

$$\overline{v^2_{KTC}} = \frac{KT}{C_I} \tag{4.7}$$

From the previous discussions, it seems that the equivalent integrator passband noise power contributed from switches M_1 and M_2 is $\frac{2KT}{C_s}$ at input or $\frac{KT}{C_I}$ at output respectively both of which only depend on the sampling capacitor C_s or the integrating capacitor C_I and environment temperature T. And the $\frac{KT}{C}$ noise is independent of the MOS resistance.

The equivalent KT/C noise density after the scaling is given by

KT/C Noise	
C(PF)	$\sqrt{(KT/C)}$ (μV)
100.0	6.4
10.0	20.2
1.0	64.0
0.1	202.0

Table 4.1 KT/C noise voltages for different values of capacitor values

$$\frac{\overline{v^2_{on\cdot KT/C}}}{\Delta f} = \frac{k_{WL}^2}{k_{ox}}\left(\frac{KT}{f_s C_s}\right)' \tag{4.8a}$$

$$= k_{KT/C}\left(\frac{KT}{f_s C_s}\right)' \tag{4.8b}$$

where k_{WL} is the scaling factor for the horizontal dimensions, k_{ox} is the oxide thickness scaling factor and $k_{KT/C} = \frac{k_{WL}^2}{k_{ox}}$ is the scaling factor of the KT/C noise. $k_{KT/C}$ is equal to k and $k^{\frac{3}{2}}$ for the CE and CV scaling respectively and both are always larger than one, hence the KT/C noise increases by scaling. The scaling factor of dynamic range due to KT/C noise is given by

$$k_{DR\cdot KT/C} \equiv \frac{1}{k_{ps}} \frac{1}{k_{KT/C}^{1/2}} \left(\frac{V_o}{v_{KT/C}}\right)' \tag{4.9a}$$

$$= \frac{1}{k_{ps}} \frac{k_{ox}^{1/2}}{k_{WL}} \left(\frac{V_o}{v_{KT/C}}\right)'$$

where V_o is the output available voltage and $v_{KT/C}$ is the output KT/C noise voltage. The factor $k_{DR\cdot KT/C}$'s of CE and CV are $k^{-\frac{3}{2}}$ and $k^{-\frac{3}{4}}$ respectively which are always less than one. Thus, increase of the KT/C noise by scaling reduces the dynamic range.

This KT/C noise is fundamental since it cannot be eliminated by either circuit or process improvement. Therefore, it determines the ultimate limit to the scaling of the circuit for a given performance level because the scaling decreases the capacitance value and hence increases the KT/C noise. In fact, for a given dynamic range, a minimum capacitor size is required . This in turn dictates a minimum chip area to integrate such a capacitor and a minimum power consumption to charge and discharge it at the signal frequency. Based on this argument, the absolute minimum capacitor area and power per pole for a switched-capacitor filter can then be computed and plotted versus the required dynamic range as shown on Fig. 4.6 and 4.7. From these plots it can be seen that for a fifth order filter with 3.4 kHz bandwidth and 95 dB dynamic range the theoretical minimum power consumption is 8 μW and the theoretical minimum area is 12 square mils, if a 10 volt supply is used and 23 square mils if a 5 volt supply is used. These limits are more than 2 orders of magnitude smaller than any PCM filter reported to date. These plots are based on many

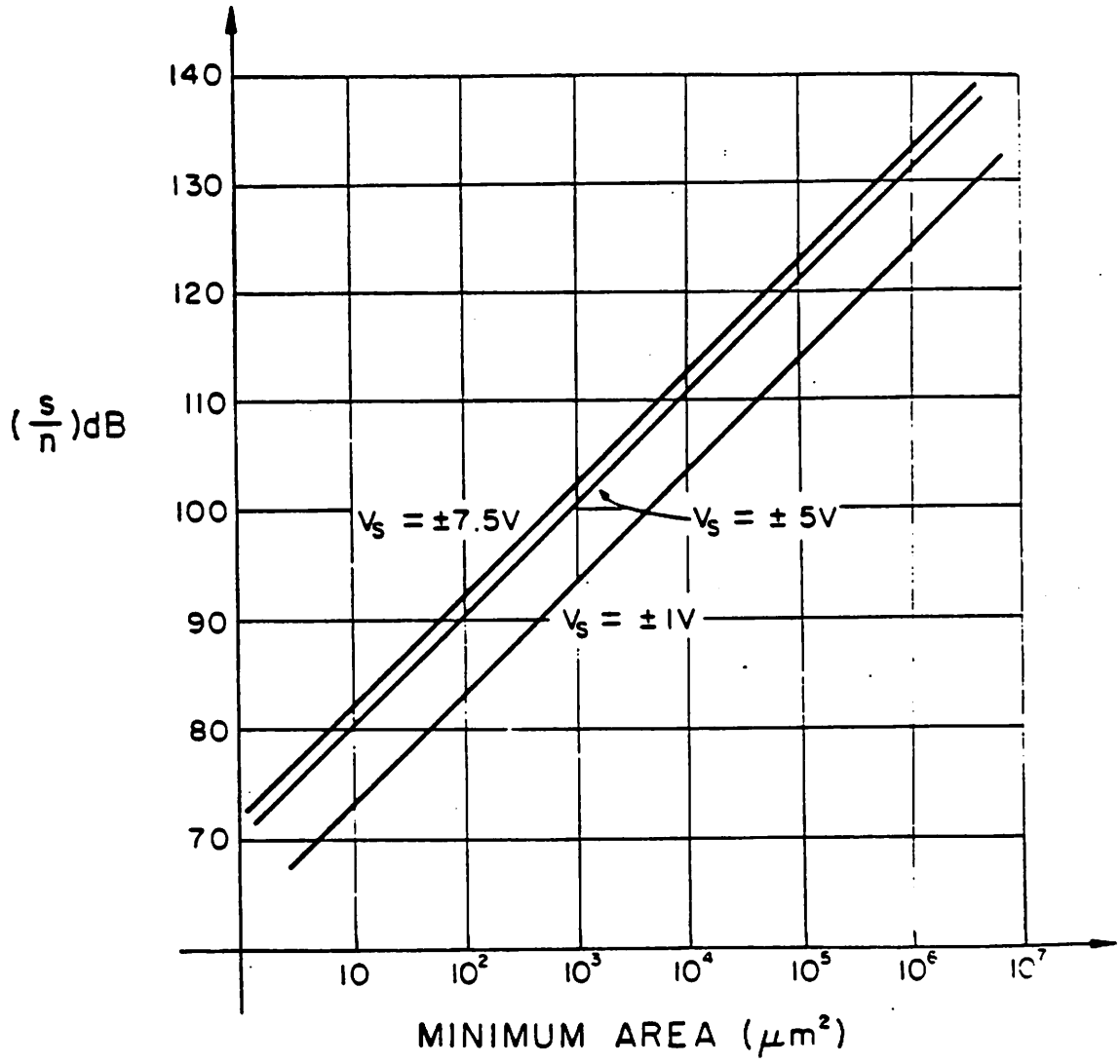


Fig. 4.6 KT/C limited dynamic range versus chip area per pole for different supply voltages

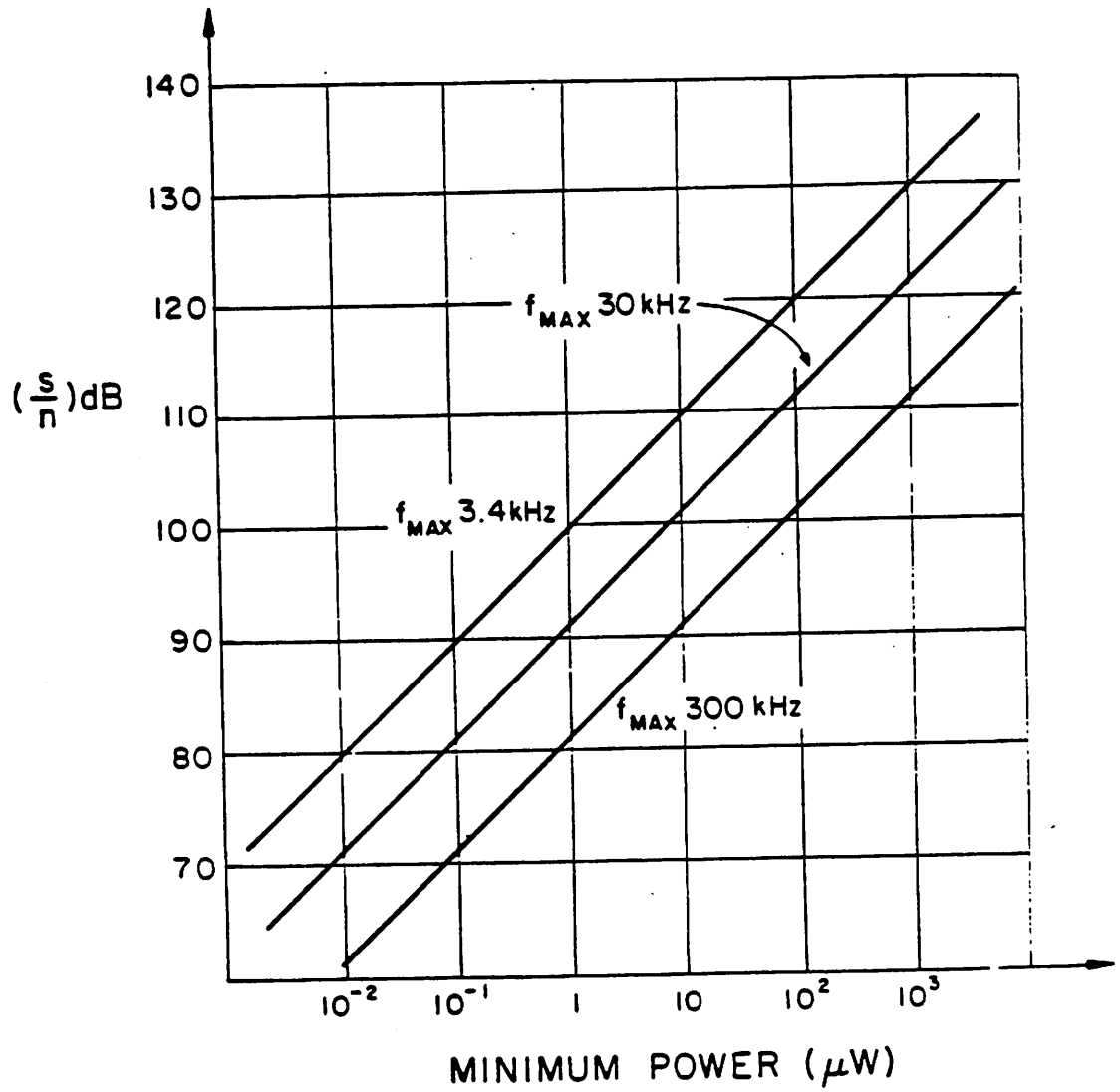


Fig. 4.7 KT/C limited dynamic range versus power dissipation per pole for different f_{max}s

idealized assumptions that are not valid in practical filters[29], but it is still true that if the practical limitations can be overcome or reduced, the area and power consumption of the filter can be dramatically reduced.

4.3.2. Op-amp Noise

Two kinds of op-amp noise are considered, thermal noise and 1/f noise. In this section, thermal noise property and its contribution to a switched capacitor circuits will be examined first. Then, practical consideration for the 1/f noise in scaled technologies will be discussed.

4.3.2.1. Thermal Noise

In a MOS transistor, the resistive channel is modulated by the gate source voltage so that the drain current is controlled by the gate source voltage. Since the channel is resistive, it exhibits thermal noise. For MOS devices, the equivalent thermal noise resistance R_{th} at gate in the saturation region is $\frac{2}{3}\frac{1}{g_m}$ (Chapter 3), where g_m is the device transconductance. For channel length less than 0.8 μ, it was reported that the equivalent gate thermal noise resistance R_{th} is about $\frac{1}{g_m}$ [30]. Thus the MOS thermal noise power density for frequency range from 0 to + ∞ can be written as

$$\frac{\overline{V^2}_{th}}{\Delta f} = 4KT\left[\frac{\alpha}{g_m}\right] \equiv 4KTR_{th} \tag{4.10}$$

where α is a constant and equals $\frac{2}{3}$ or 1 as discussed. Thermal noise in submicron devices, therefore, tends to be larger even if the transconductance is the same.

From (4.10), the thermal noise power density is inversely proportional to the transconductance g_m. From Table 3.1, transconductance is unchanged in the CE scaling and increases in the CV scaling. Therefore, to the first order approximation, the thermal noise is unchanged in the CE case and reduced in the CV case. However, the contribution of the op-amp thermal noise to a switched capacitor integrator is related to the op-amp unity gain frequency and other switch RC

time constants in the circuit. This will be discussed in the following.

4.3.2.1.1. Op-amp Thermal Noise Contribution to a SC Integrator

The op-amp thermal noise contribution to a switched capacitor integrator has been discussed by several authors [31,32,27]. One analysis is to model the op-amp as an ideal voltage source with a one pole roll off given by Fischer[31] Gobet[32] and Hsieh[27] as shown is Fig. 4.8. For a lot of applications however the op-amp finite resistance and capacitance at the output can not be neglected. In this case, the op-amp is modeled as in Fig. 4.9.

Based on the results derived by Gobet[32], Fischer[31] and Hsieh[27], if the op-amp is modeled as an ideal voltage source with one single pole roll off as in Fig. 4.8, the fold back thermal noise density at input of a switched capacitor integrator can be expressed as follows:

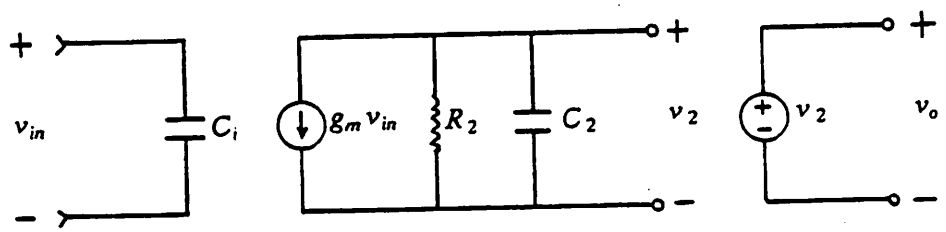


Fig. 4.8 Op-amp modeled as an ideal voltage source and having one pole roll off

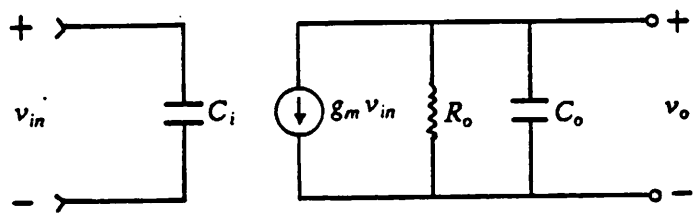


Fig. 4.9 Op-amp with finite output impedance

$$\frac{\overline{v_{th}^2}}{\Delta f} = Y \cdot S \cdot \frac{KT}{C f_s} \tag{4.11}$$

where Y is a function of $\frac{R_{th}}{R_{on}}$ and $\frac{\omega_u}{\omega_{on}}$, S is a $\frac{sin(x)}{x}$ function of $\frac{f}{f_s}$ which is due to the sample and hold effect, R_{th} is the equivalent op-amp thermal noise resistance, R_{on} is the on resistance of the switch, ω_u is the unity gain frequency of the op-amp and the ω_{on} is the unity gain frequency associated with the sampling switch and the capacitor. Combined with the KT/C noise of (4.6), the total equivalent input noise density of an integrator (Fig. 10e) with an ideal voltage source op-amp can be written as [27]

$$\frac{\overline{v^2_{total}}}{\Delta f} = \frac{\overline{v^2_{KT/C}}}{\Delta f} + \frac{\overline{v^2_{th}}}{\Delta f} = \frac{2KT}{f_s C_s} + \frac{KT}{f_s C_s}\left\{\frac{\frac{R_{th}}{R_{on1}}}{1+\frac{\omega_{on1}}{\omega_u}} + \frac{\frac{R_{th}}{R_{on2}}}{1+\frac{\omega_{on2}}{\omega_u}}\right\} \tag{4.12a}$$

For most applications, the on resistance of switch M_1 equals that of switch M_2. Therefore, $R_{on1} = R_{on2} \cong R_{on}$ and $\omega_{on1} = \omega_{on2} \cong \omega_{on}$. Then (4.12a) can be written as

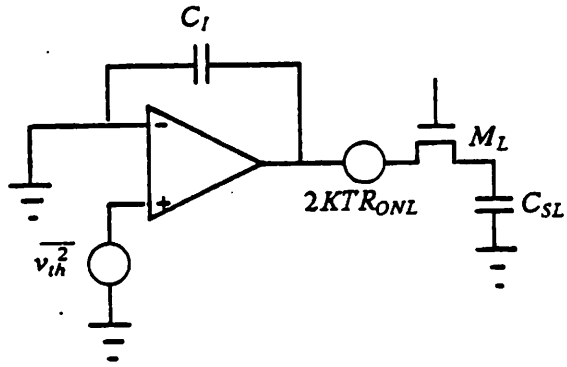
$$\frac{\overline{v^2_{total}}}{\Delta f} = \frac{2KT}{f_s C_s}\left\{\frac{\frac{1}{\omega_{th}} + \frac{1}{\omega_{on}} + \frac{1}{\omega_u}}{\frac{1}{\omega_{on}} + \frac{1}{\omega_u}}\right\} \tag{4.12b}$$

The second term of (4.12a) is the foldback thermal density $\overline{v^2_{th}}/\Delta f$ given by (4.11) if the sample and hold effect is not considered. From (4.12a)/(4.12b), the total equivalent input KT/C related noise density of an integrator is always larger than $\frac{2KT}{f_s C_s}$ for frequencies from $-\infty$ to $+\infty$ sampled by the input sampling capacitor if the op-amp is modeled as an ideal voltage source.

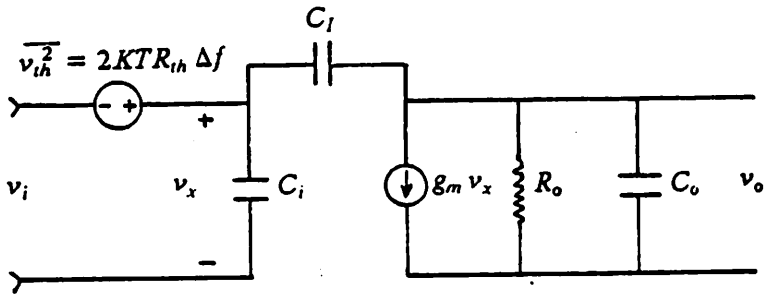
In the case of one pole op-amp with finite output impedance as shown in Fig. 4.9, the op-amp thermal noise contribution in a switched capacitor integrator is associated with the op-amp unity gain frequency, switch time constants, etc. Hsieh analyzed the noise of an integrator which is sampled by a switched capacitor at the op-amp output (Fig. 4.10a, 4.10b). The total noise power sampled by the switched capacitor C_{SL} (Fig. 4.10c) contributed from the op-amp thermal noise and the KT/C noise of the switch M_{SL} is given by [27]

$$\overline{v^2_{otoal}} = KT\left\{\frac{\frac{R_{th}}{b^2} + R_{onL}\left(1+\frac{C_{ot}}{bG_m R_{onL} C_{SL}}\right)}{R_{onL}C_{SL} + \frac{C_{ot}+C_{SL}}{bG_m}}\right\} \tag{4.13}$$

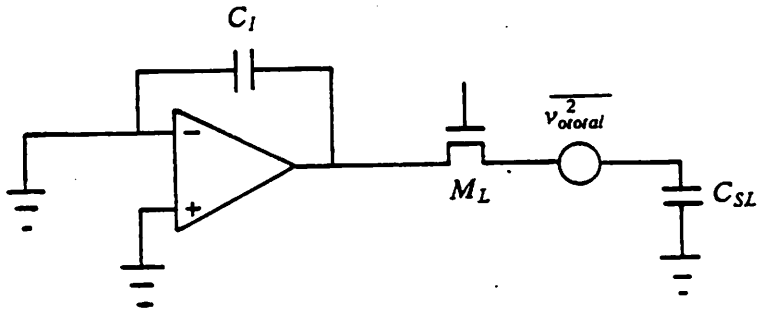
which can be rearranged and written in a noise power density form



(a)

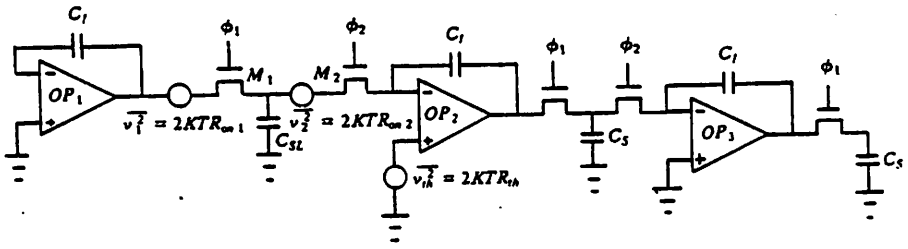


(b)

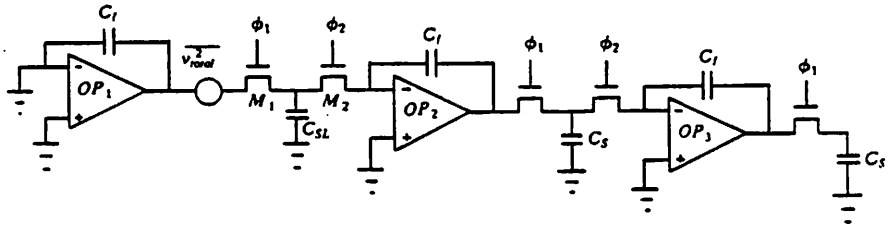


(c)

Fig. 4.10 (a) Integrator output sampled by a switched capacitor C_{SL}
(b) Equivalent circuit of the integrator composed of finite output impedance op-amp
(c) $\overline{v_{ototal}^2}$, equivalent noise due to op-amp thermal noise, sampled by C_{SL}



(d)



(e)

Fig. 4.10 (d) Integrator followed by another stage
(e) Equivalent thermal noise, $\overline{v_{total}^2}$, sampled by C_{SL}

$$\frac{\overline{v^2_{ototal}}}{\Delta f} = \frac{KT}{f_s C_{SL}} \left\{ \frac{\frac{R_{th}}{b^2} + R_{onL}\left(1 + \frac{C_{ot}}{bG_m R_{onL} C_{SL}}\right)}{R_{onL} + \frac{\frac{C_{ot}}{C_{SL}} + 1}{bG_m}} \right\} \tag{4.14}$$

where C_{SL} is the sampling capacitor at the integrator output as shown in Fig. 4.10, R_{th} is the op-amp equivalent input noise resistance, R_{onL} is the on resistance of the switch M_L (Fig. 4.10), G_m is the op-amp transconductance, $b \equiv \frac{C_I}{C_I + C_i}$, $C_{ot} \equiv bC_i + C_o$, C_i is the op-amp input parasitic capacitance, C_I is the integrating capacitor and C_o is the op-amp output capacitance as shown in Fig. 4.10b. The following relations are easily derived from (4.14).

$$\frac{\overline{v^2_{ototal}}}{\Delta f} > \frac{KT}{f_s C_{SL}} \qquad \text{if } R_{th} > \frac{b}{G_m} \tag{4.15a}$$

$$\frac{\overline{v^2_{ototal}}}{\Delta f} < \frac{KT}{f_s C_{SL}} \qquad \text{if } R_{th} < \frac{b}{G_m} \tag{4.15b}$$

$$\frac{\overline{v^2_{ototal}}}{\Delta f} = \frac{KT}{f_s C_{LS}} \qquad \text{if } R_{th} = \frac{b}{G_m} \tag{4.15c}$$

For most of the application, $b \simeq 1$, because the integrating capacitance C_I is much larger than the input parasitic capacitance. If $R_{th} < \frac{1}{G_m}$ then, from (4.15b) it turns out the op-amp finite output impedance not only reduces the op-amp thermal noise density sampled by sampling capacitor C_{SL} at the integrator output (Fig. 4.10a), but also makes the total noise density sampled by capacitor C_{SL} even less than the KT/C noise density sampled by the switched capacitor C_{SL} shown in (4.15b) i.e. $\frac{\overline{v^2_{ototal}}}{\Delta f} < \frac{KT}{f_s C_{SL}}$.

Considering also the noise contribution from the the following integrator composed of the same op-amp with finite output impedance as shown in Fig. 4.10d, Castello derived the integrator noise density sampled by the switched capacitor C_{SL} (Fig. 4.10e) [33]:

$$\frac{\overline{v^2_{total}}}{\Delta f} = \frac{KT}{C_{SL}f_s}\left[\frac{(G_mR_{th} + G_mR_{on1})C_{SL} + C_o}{(1 + G_mR_{on1})C_{SL} + C_o} + \frac{(G_mR_{th} + G_mR_{on2})C_{SL} + C_o}{(1 + G_mR_{on2})C_{SL} + C_o}\right] \tag{4.16}$$

where G_m is again the transconductance of the op-amp, C_o is the op-amp output capacitance, R_{th} is the op-amp equivalent thermal noise resistance, C_{SL} is the sampling capacitor at the output of the first integrator (Fig. 4.10d), and R_{on1} and R_{on2} are the on resistance of the sampling switches M_1 and M_2 respectively. The equivalent total noise density $\frac{\overline{v^2_{total}}}{\Delta f}$ shown as a generator in Fig. 4.10e can be easily manipulated in terms of time constants, $(\frac{1}{\omega})'s$ from (4.16).

$$\frac{\overline{v^2_{total}}}{\Delta f} = \frac{KT}{C_{SL}f_s}\left[\frac{\frac{1}{\omega_{th}}+\frac{1}{\omega_{on1}}+\frac{1}{\omega_u}}{\frac{1}{\omega_l}+\frac{1}{\omega_{on1}}+\frac{1}{\omega_u}} + \frac{\frac{1}{\omega_{th}}+\frac{1}{\omega_{on2}}+\frac{1}{\omega_u}}{\frac{1}{\omega_l}+\frac{1}{\omega_{on2}}+\frac{1}{\omega_u}}\right] \tag{4.17}$$

where $\omega_u = \frac{G_m}{C_L}$, $\omega_{th} = \frac{1}{R_{th}C_{SL}}$, $\omega_l = \frac{G_m}{C_{SL}}$, $\omega_{on1} = \frac{1}{R_{on1}C_{SL}}$, and $\omega_{on2} = \frac{1}{R_{on2}C_{SL}}$. For most applications, the on resistance of switch M_1 equals to that of switch M_2. Therefore, ω_{on1} equals to ω_{on2}. Then (4.17) can be simplied as

$$\frac{\overline{v^2_{total}}}{\Delta f} = \frac{2KT}{C_{SL}f_s}\left[\frac{\frac{1}{\omega_{th}}+\frac{1}{\omega_{on}}+\frac{1}{\omega_u}}{\frac{1}{\omega_l}+\frac{1}{\omega_{on}}+\frac{1}{\omega_u}}\right] \tag{4.18}$$

where $\omega_{on} \equiv \omega_{on1} = \omega_{on2}$.

With simple mathematical manipulation, the equivalent total noise density $\frac{\overline{v^2_{total}}}{\Delta f}$ of (4.16) can be rearranged as

$$\frac{\overline{v^2_{total}}}{\Delta f} = \frac{KT}{f_sC_{SL}}\left\{\frac{R_{th} + R_{on1}(1 + \frac{C_o}{G_mR_{on1}C_{SL}})}{R_{on1} + \frac{\frac{C_o}{C_{SL}}+1}{G_m}} + \frac{\frac{R_{th}}{b^2} + R_{on2}(1 + \frac{C_o}{G_mR_{on2}C_{SL}})}{R_{on2} + \frac{\frac{C_o}{C_{SL}}+1}{G_m}}\right\}$$

(4.19)

Again, because R_{on1} equals R_{on2} of switches M_1 and M_2 practically, i.e. $R_{on1} = R_{on2} \equiv R_{on}$, then (4.19) becomes

$$\frac{\overline{v^2_{total}}}{\Delta f} = \frac{2KT}{f_s C_{SL}} \left\{ \frac{R_{th} + R_{on}\left(1 + \frac{C_o}{G_m R_{on} C_{SL}}\right)}{R_{on} + \frac{\frac{C_o}{C_{SL}} + 1}{G_m}} \right\} \tag{4.20}$$

Comparing (4.14) and (4.20), other than the factor 2, these two equations are the same, if integrating capacitor C_I is much larger than the input parasitic capacitance C_i which is true for most of the practical cases. The factor of 2 difference is, as mentioned, because (4.20) considers also the noise contribution from the following stage. Similar to equations (4.15a, b, and c), the following relations can be easily derived from (4.20):

$$\frac{\overline{v^2_{total}}}{\Delta f} > \frac{2KT}{f_s C_{SL}} \qquad \text{if } R_{th} > \frac{1}{G_m} \text{ or equivalently } \omega_I > \omega_{eq} \tag{4.21a}$$

$$\frac{\overline{v^2_{total}}}{\Delta f} < \frac{2KT}{f_s C_{SL}} \qquad \text{if } R_{th} < \frac{1}{G_m} \text{ or equivalently } \omega_I < \omega_{eq} \tag{4.21b}$$

$$\frac{\overline{v^2_{total}}}{\Delta f} = \frac{2KT}{f_s C_{LS}} \qquad \text{if } R_{th} = \frac{1}{G_m} \text{ or equivalently } \omega_u = \omega_{th} \tag{4.21c}$$

From (4.21), if the equivalent noise resistance is smaller than the inverse of the op-amp transconductance i.e. $R_{th} < \frac{1}{G_m}$, then the total noise density is smaller than the noise density sampled by the sampling capacitor C_{SL} in Fig. 4.10d i.e. $\frac{\overline{v_{ntotal}^2}}{\Delta f} < 2\frac{KT}{C_I f_{SL}}$ (equation 4.21b). This result, equation (4.21b), is exactly the same as equation (4.15b) if the noise of the following stage is not considered. On the other hand, when the equivalent noise resistance R_{th} is larger, i.e. $R_{th} > \frac{1}{G_m}$, the total noise density will be larger than the amount $\frac{2KT}{C_{SL} f_s}$. In the design of a switched capacitor integrator with a finite output impedance op-amp, the op-amp equivalent noise resistance should be

smaller than the inverse of the op-amp transconductance in order to achieve less KT/C noise i.e. equation (4.21b) holds.

Comparing (4.12b) and (4.18), the term $\omega_l \equiv \frac{G_m}{C_{SL}}$ appears only in the noise equation of an integrator with a finite output impedance op-amp (4.18). In that case, there is an equivalent resistor of value $\frac{1}{G_m}$ at the input if the integrating capacitor C_l is large when the impedance of C_l is low. From the above discussions of (4.12) and (4.20) for an op-amp modeled either as an ideal voltage source or as a finite output impedance, the foldback equivalent input noise density due to the op-amp thermal noise is always $\frac{KT}{C_s f_s}$ times a constant. Therefore, in this sense, the op-amp thermal noise contribution to a switched capacitor integrator can be lumped as $\frac{KT}{C}$ noise. The op-amp 1/f noise has not been considered in the discussion of the integrator noise in (4.12) and (4.20) because it can be removed by the chopper stabilization technique (Chapter 5).

4.3.2.2. 1/f Noise

The phenomenon of the 1/f noise is well known. Its noise power density is inversely proportional to the frequency and gate capacitance. In the submicron channel, it was reported that the 1/f noise density is inversely proportional to the frequency to the order 0.5 [30]. It is almost independent of the saturation current. So far, there is no unified or conclusive analytical equation to explain all the parameter dependence of the 1/f noise phenomenon. However, the 1/f noise power density can be generalized as

$$\frac{\overline{v_f^2}}{\Delta f} = \frac{K_f}{WLC_{ox}^n f^m} \tag{4.22}$$

where n is an empirical parameters with values between 0 and 2 [12 ,34], m is also an empirical parameters and K_f is a parameter that is heavily dependent on the technology used but, to first order independent of the operating condition of the device (current level, gate voltage).

The 1/f noise after scaling is

$$\frac{\overline{v_f^2}}{\Delta f} = \frac{k_{WL}^2}{k_{ox}^2}\frac{K_f}{WLC_{ox}^n f^m}$$

$$= k_f \frac{K_f}{WLC_{ox}^n f^m} \tag{4.23}$$

where k_f is the 1/f noise scaling factor and equals $\frac{k_{WL}^2}{k_{ox}^2}$, k_{WL} is the scaling factor of the horizontal dimension, and k_{ox} is the oxide thickness scaling factor. To be consistent with most of the literature, here $n = 2$ and $m = 1$ are used for the study of the 1/f noise in scaled technologies. Then equation (4.23) will become

$$\frac{\overline{v_f^2}}{\Delta f} = k_f \frac{K_f}{WLC_{ox}^2 f} \tag{4.24}$$

As in Table 3.1, the 1/f noise by the CE scaling is unchanged, i.e. $k_f = 1$. On the other hand, the CV scaling increases the 1/f noise by a factor $k_f = k$. For the dynamic range, if only 1/f noise is considered,

$$k_{DRf} = \frac{1}{k_{ps}}\frac{1}{\sqrt{k_f}}\left(\frac{V_o}{V_f}\right)' \tag{4.25}$$

where V_o is the output voltage and v_f is the output 1/f noise voltage. In both the CE and CV scalings, $k_{DRf}'s$, $\frac{1}{k}$ and $\frac{1}{\sqrt{k}}$ respectively, are always less than one, therefore, the dynamic range is always worsened by scaling.

Combining the thermal noise given in (4.10) with $\alpha = \frac{2}{3}$, and the 1/f noise given in (4.24), the overall input equivalent noise of a MOS transistor can be represented by an equivalent gate noise power density

$$\frac{\overline{v_{ng}^2}}{\Delta f} = 4KT\left(\frac{2}{3}\frac{1}{g_m}\right) + \frac{K_f}{C_{ox}^2 WL}\frac{1}{f} \tag{4.26}$$

The corner frequency of a MOS transistor is given by

$$f_c = \frac{K_f}{C_{ox}^2 WL} \frac{\frac{3}{2} g_m}{4KT} \tag{4.27}$$

From (4.27), corner frequency is unchanged under the CE scaling and increases as factor of $k^{3/2}$ for CV scaling.

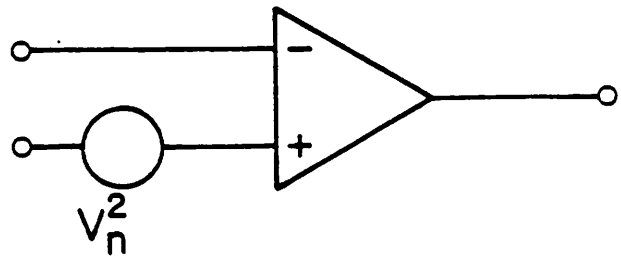
Because of the presence of the MOS 1/f noise, the input referred noise of MOS amplifiers shows a large component at low frequency as shown in Fig. 4.11. This 1/f noise is correlated color noise. It can be cancelled by circuit techniques which will be discussed in Chapter 5.

4.3.3. Coupled Digital Noise

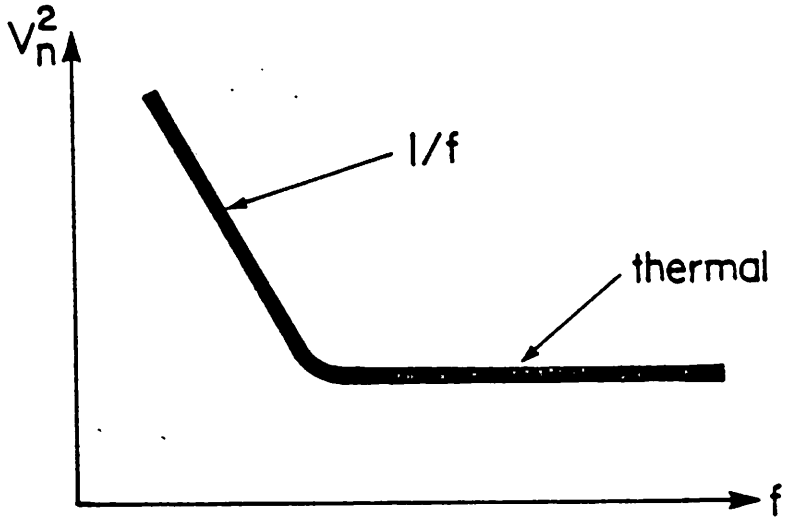
Another important practical limitation to the achievable dynamic range of a scaled analog circuit is the external noise pickup (particularly digital noise). This is not directly linked to the scaling itself but is due to the fact that scaled analog circuits will often be part of large digital/analog single chip systems. Since the complexity of the digital system that can be integrated on a single chip increases in inverse proportion to the square of the feature size, the amount of digital noise generated in the system also tends to increase in a scaled technology. Furthermore, the increased circuit density and reduced distance between devices tend to increase the noise coupling between the digital and analog parts of the chip. To reduce this form of noise pickup, the digital and analog portions of the system are usually decoupled by physically locating them far apart and using separate ground and power lines, however this must be supplemented with the use of analog circuit approaches that are inherently able to reject high frequency power supply noise which will be discussed in the next chapter.

4.4. Summary

In this chapter, the origins and the characteristics of the practical limiting factors for the scaling of the analog circuits are discussed. These limiting factors are the decrease of the op-amp gain and the output voltage swing, the increase of the op-amp noise (particularly the 1/f noise), KT/C



(a)



(b)

Fig. 4.11 (a) Op-amp with equivalent input noise voltage source
(b) Equivalent op-amp input noise spectrum

noise and the coupled digital noise. KT/C noise is the fundamental limit to the scaling of the analog circuits. Table 4.2 summarizes these limiting factors with respect to the scaling factors. Next chapter will discuss the circuit and system approaches to reduce or remove the impact of these practical limiting factors for the MOS analog circuits in scaled technology.

Performance Limiting Factors	CE	CV	Equation
Voltage swing	$1/k$	1	4.2, 4.3
Gain	1	$1/k$	4.5
1/f noise	1	k	4.23
KT/C noise	k	$k^{3/2}$	4.8
$\frac{\text{output swing}}{\text{1/f noise voltage}}$	k	$k^{1/2}$	4.25
$\frac{\text{output swing}}{\text{KT/C noise voltage}}$	$k^{3/2}$	$k^{3/4}$	4.9

Table 4.2 Summary of op-amp performance limiting factors in scaled technologies

CHAPTER 5

DESIGN TECHNIQUES FOR SCALABLE MOS ANALOG CIRCUITS

In the previous chapter, the practical limiting factors to the scaling of the analog circuits were discussed. They are the decrease of voltage swing and op-amp gain, and the increase of the noise.

System and circuit design techniques to solve the problems associated with the scaling of the analog circuits are the topics in this chapter. Two major subjects, the chopper stabilization technique and the op-amp architecture in scaled technologies, will be discussed in detail. Chopper stabilization technique is used to minimize the effect of the 1/f noise by shifting it to a higher frequency so that it is outside the band of interest. The class AB op-amp can provide high gain and high swing without sacrificing speed. A differential architecture, dynamic common moded feedback (DCMFB) and dynamic biasing are other features of the op-amp used in this project.

5.1. 1/f Noise Reduction

In the past, 1/f noise was reduced by simply increasing the input device size of the op-amp. This approach works well in the case where minimum feature size is large. In scaled technology, however, the fact that the low frequency 1/f noise does not scale with the feature size limits the ability to scale the input section of the op-amp. Area of op-amp is dominated by input transistors. Hence, the scaling of the physical dimensions of the op-amps is not possible if no increase of 1/f noise is allowed unless special measures are taken.

1/f noise is process dependent. The degradation of the dynamic range due to 1/f noise alone is worsened by a factor of k and $k^{1/4}$ in the CE (Constant Field) law and CV (Constant Voltage) law respectively (cf Table 4.2). 1/f noise can be reduced in several ways. First, a clean process with little contamination and few defects at the oxide silicon interface has less 1/f noise than a dirty

process. Second, buried channel MOS devices avoid the surface states that cause 1/f noise. Third, if an increase in input current can be tolerated, low 1/f noise can be achieved by using the lateral BJT. All the above methods to reduce 1/f noise increase either the process complexity or the chip area.

Another approach is to use circuit design techniques for reducing the 1/f noise through frequency translation methods such as the correlated double sampling (CDS) technique and the chopper stabilization technique [35,36] The former technique requires more constraints on the op-amp design, particularly on the settling time [27] and is not considered further here. The latter requires only additional MOS switches added to the op-amp. Thus the required op-amp performance is not constrained by the use of the chopper stabilization technique.

5.1.1. Chopper Stabilization Technique

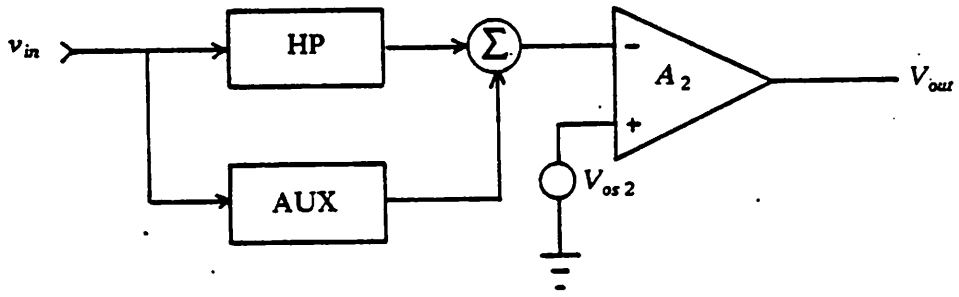
As stated, the 1/f noise is process dependent; it must be minimized to guarantee an adequate dynamic range over all possible process variations. This can be achieved by using chopper stabilization techniques [40] because 1/f noise is concentrated at low frequencies.

Chopper stabilization has been traditionally used in the design of precision DC amplifiers [36] whose offset voltage must be low and free from drift. Fig. 5.1 shows the basic block diagram of a chopper stabilized op-amp. The input signal is separated into both an AC path and a DC path [36]. The DC path contains an auxiliary amplifier, AUX which is composed of an input LP bandlimited filter, the DC chopped amplifier and the output LP filter (Fig. 5.1b) [36]. The offset of A_1 is shifted to the higher frequency and filtered out. The equivalent input offset voltage due to op-amp A_2 is

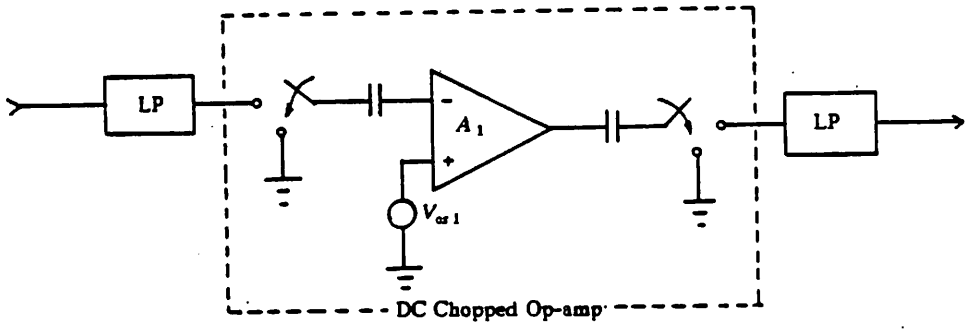
$$V_{os} = \frac{V_{os2}A_2}{A_2A_1} \tag{5.1a}$$

$$\approx 0 \tag{5.1b}$$

The chopping operation shifts all frequencies. Thus the equivalent low frequency noise can be



(a)



(b)

Fig. 5.1 (a) Block diagram of a chopper stabilized op-amp
(b) The AUX DC amplifier of the main amplifier A_2

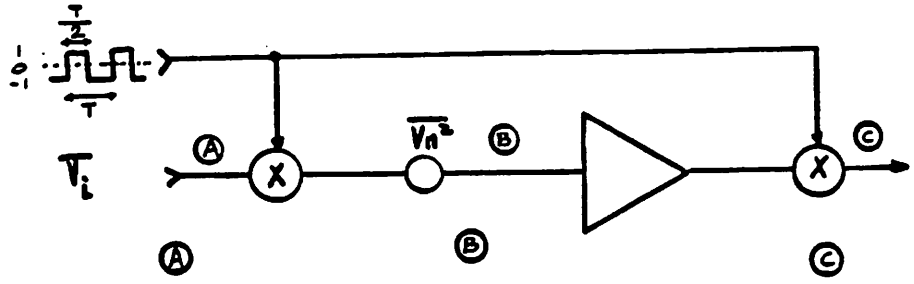
reduced as well by the chopper stabilization technique as long as the low frequency noise is nonwhite.

A chopper is a switched-mode pulse height modulator. It multiplies its main input by either +1 or -1. The chopper clock used is a clock with a duty cycle of 50%. The block diagram of a single stage operational amplifier with choppers at the input and the output is shown in Fig. 5.2. The magnitudes of the signal and the noise spectrum at each of the nodes are shown in Fig. 5.2b and Fig. 5.2c respectively. The first chopper modulates the input signal at the op-amp input up to the harmonics of the chopper clock frequency. The second chopper demodulates the signal at the op-amp output. Note that the modulating and demodulating clocks are synchronous with each other. Hence, the input signal is not affected by the choppers. The choppers are easily made to subsynchronize to the main filter clock. The generator $\overline{V_n^2}$ in Fig. 5.2a represents the op-amp equivalent input noise, which is mainly composed of 1/f noise and thermal noise. The noise due to the operational amplifier, however, is chopped by the output chopper only. Therefore, its spectrum is moved to the odd harmonics of the chopper clock frequency. The op-amp thermal noise n_t is white and thus is not affected by the chopper [27]. The magnitude of the 1/f noise S_{eq} after the output chopper is given by

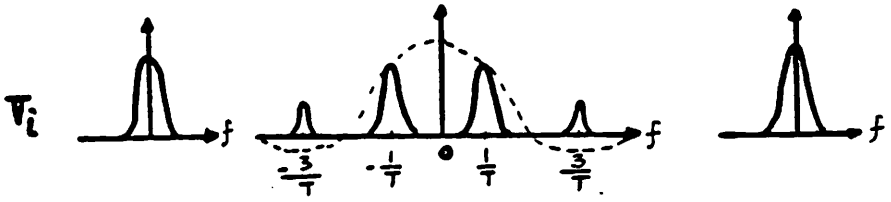
$$S_{eq} = \sum_{i=1}^{i=\infty} \left[\frac{2}{(2m-1)\pi} \right]^2 \left[S_f(f-(2m-1)f_s) + S_f(f+(2m-1)f_s) \right] \tag{5.2}$$

where $S_f(f) = \frac{A_f}{f}$ is the noise power density of the 1/f noise in the passband, and f_s is the sampling frequency. From (5.2), the magnitude of chopped 1/f noise decreases when the frequency increases. Its spectrum is mirrored at the odd harmonics of the chopper frequency as shown in Fig. 5.2c, because the chopper clock has symmetrical amplitude and 50% duty cycle. The passband 1/f noise S_{ineq} after the second chopper is the foldback of the 1/f noise, which is only a fraction of the original noise and is given by

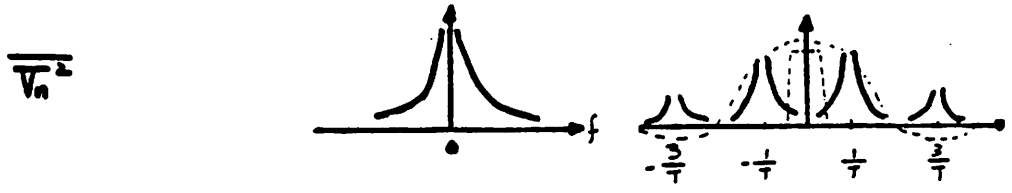
$$S_{ineq} = \frac{4}{\pi^2} \left[S_f(f-f_s) + S_f(f+f_s) \right] \tag{5.3a}$$



(a)



(b)



(c)

Fig. 5.2 (a) Block diagram of a chopper stabilized operational amplifier
(b) Magnitude of the frequency spectrum of the input signal at nodes indicated
(c) Magnitude of the frequency spectrum of the op-amp noise at nodes indicated

$$= \frac{4}{\pi^2}\left(\frac{A_f}{f - f_s} + \frac{A_f}{f + f_s}\right) \tag{5.3b}$$

Practical considerations concerning the use of nonideal chopper clocks will be discussed in Chapter 6.

5.1.2. MOS Realization of Chopper Stabilization Technique

A MOS realization of the chopper stabilization technique for a single stage op-amp is shown in Fig. 5.3a where v_n is the equivalent op-amp input noise voltage. A chopper is implemented by a set of 4 switches at the input and at the output that are controlled by two non-overlapping clocks, ϕ_c and $\overline{\phi_c}$. This simple solution is made possible by the availability of both signal polarities in a fully differential configuration. The sign of the input equivalent noise voltage, v_{ieq}, varies periodically, as shown in Fig. 5.3b and Fig. 5.3c. The equivalent input noise v_{ieq} can be written as $v_{ieq} = -v_n$ for clock phase at ϕ_c (Fig. 5.3b) and $v_{ieq} = +v_n$ for clock phase at $\overline{\phi_c}$ (Fig. 5.3c) respectively.

5.2. Operational Amplifier Architecture

Limitations to the scaling in analog circuits were discussed in chapter 4. These are the smaller voltage swing available and the increase of noise level. Except for the KT/C noise, which is fundamental limiting factor, these limitations can be considered as op-amp design problems. Another practical consideration is the chip area. In most conventional switched-capacitor filters, for example, the op-amps constitute the largest single component of the area. To reduce the overall size of the filter it is therefore important to be able to reduce the area of the op-amps. One concern in the reduction of the op-amp area is the increase of the 1/f noise, which can be removed by the chopper stabilization technique, as discussed in the previous section.

In this section, a choice of op-amp architecture that will remove or at least reduce the impact of the scaling on swing and gain will be discussed. Moreover, the operational amplifier can use all minimum channel length devices and is also optimized in order to use the smallest possible aspect

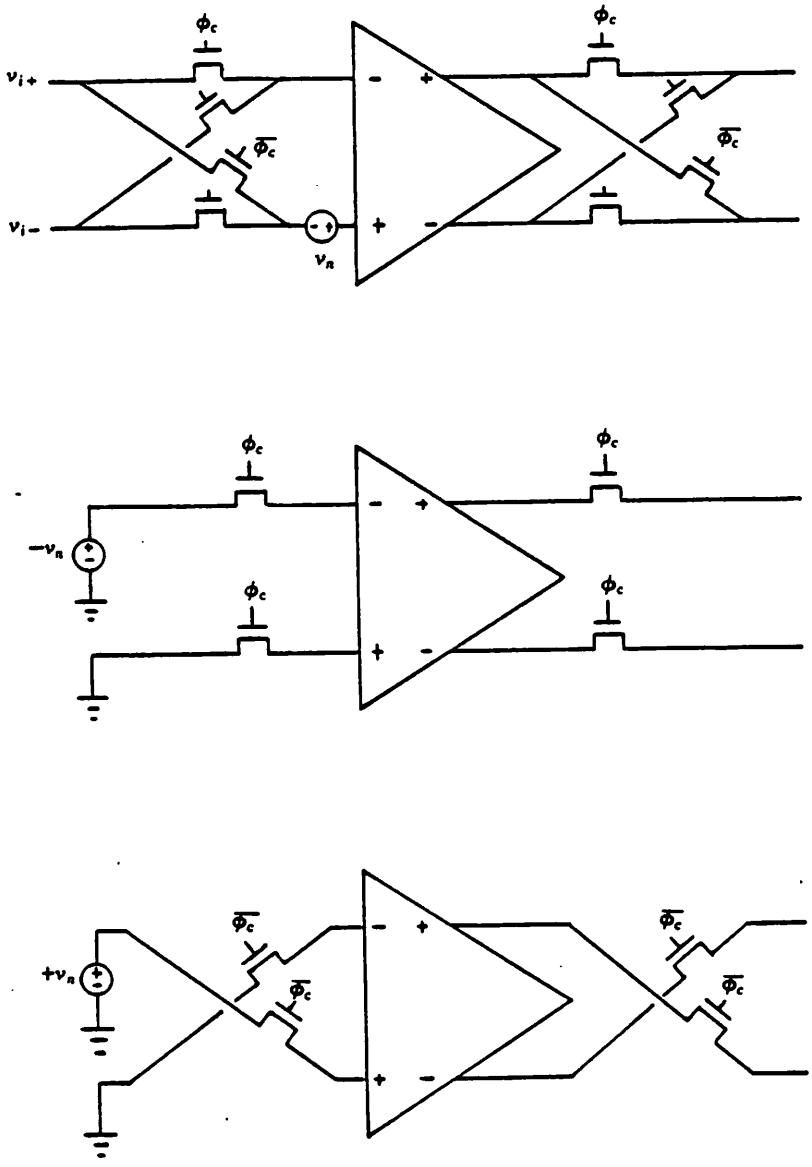


Fig. 5.3 (a) MOS implementation of a single stage chopper stabilized op-amp
(b) Equivalent input noise when chopper clock $\overline{\phi_c}$ is on
(c) Equivalent input noise when chopper clock ϕ_c is on

ratio in all of the devices.

5.2.1. Fully Differential Architecture

The distinct characteristics of a fully differential architecture are the balanced/symmetric circuit and the differential signal path. There are several well known advantages of the fully differential architectures [40,27]. The dynamic range is improved by 6 dB, because the signal is doubled by the fully differential configuration. It provides first order cancellation of the clock feedthrough and switch charge injection. PSRR (Power Supply Rejection Ratio) is improved because the spurious noises cancel each other out by the symmetric architecture. This is particularly important to guarantee a sufficient immunity from digital noise generated within the chip by using the differential architecture. Since both the signal polarities are available, a differential architecture eases the difficulty of desired signal manipulation. For example, a chopper in a fully differential architecture is just a set of 4 switches that either cross the signal to reverse its polarity or pass it unaffected (Fig. 5.3).

Unfortunately, there are several disadvantages of differential circuits. The main drawback is the need of a CMFB (Common Mode Feedback) circuit in order to set up the output common mode voltage. The area and power are therefore increased. However, the choice of the differential op-amp architecture and the layout strategy can still minimize these disadvantages.

5.2.2. Class AB

A class A op-amp usually has a large standby current when there is no AC input signal. This results in reduced gain and voltage swing (because of larger overdrive). Coupled with the short channel effects, these drawbacks are more of a problem in scaled technology. For a class A op-amp, the maximum available current to charge and discharge the output capacitors is less than or equal to the bias current. This limits the slewing speed of the circuit. The need for a large standby current often makes class A operation unsuitable for battery operated equipment. The large

power dissipated in the chip affects the chip size required. This is important in scaled technology. For the CV case, the power density (power/area) increases by a factor of $k^{5/2}$ (cf Table 3.1), where k is the scaling factor.

A class B op-amp consumes no power with zero input. Output and input characteristics of a class B op-amp are shown in Fig. 5.4. Although the class B op-amp alleviates the problems associated with the large standby current compared to a class A op-amp, there is a "dead zone" centered around zero input in its voltage characteristics that gives cross over distortion, which is severe when the input signal is small.

The cross over distortion of a class B op-amp can be eliminated by using a class AB op-amp, because the class AB op-amp is biased at a small current. However, it can sink and source a current that is larger than the standby current when there is a signal at the input. The original

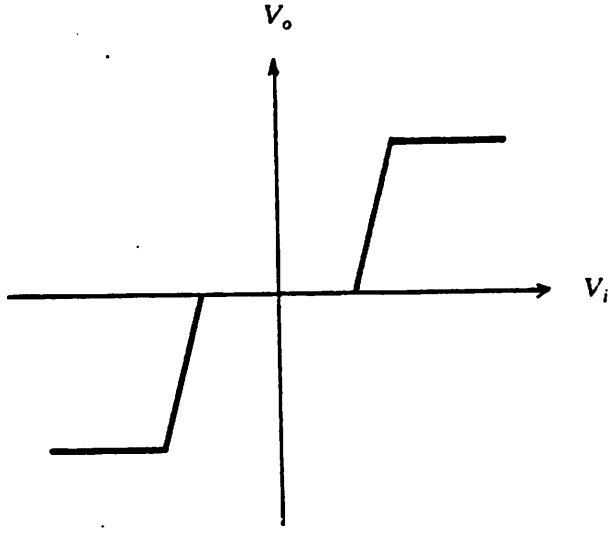


Fig. 5.4 Input output characteristics of a class B op-amp

motivation for choosing a class AB op-amp is the reduction of the power consumption. However, there are more important advantages in scaled technology than that of conserving power. Considerations on the choice of a class AB op-amp in scaled technology is the topic in this section.

The op-amp for the application of the switched capacitor circuits, for example a switched capacitor integrator as in Fig. 5.5, should be able to provide fast charging and no loss of charge transfer. A fast transfer of charge in switched capacitor circuits requires that the unity gain bandwidth of the op-amp be high enough. If the op-amp DC gain is large, the charge transfer from capacitor C_s to capacitor C_I can be complete (Fig. 5.5). These two basic properties, fast charging and no loss of charge, are required at the different times. Fast charging is required at the beginning of the clock (phase ϕ_T in Fig. 5.5) and a high DC gain is needed at the end of the transient. That is, an op-amp with a fast transient (large gain bandwidth product) and high standby DC gain is

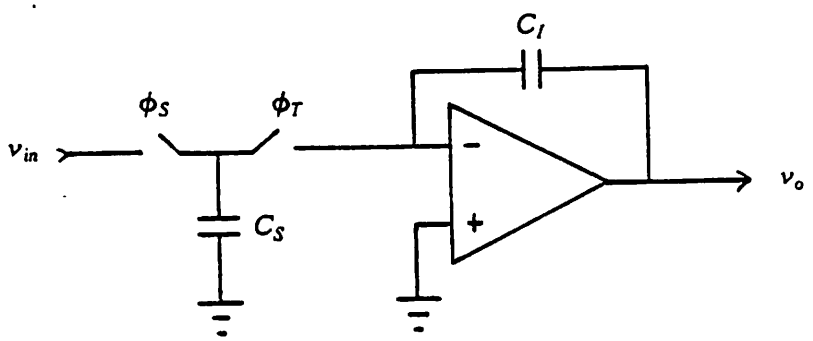


Fig. 5.5 Switched capacitor integrator

desirable for the application of charge transfer circuits such as switched capacitor integrators. One kind of op-amp that fulfills these two requirements is the class AB op-amp that will be discussed in this section. Dynamic biasing to further provide fast transient and high standby gain will be discussed in the next section.

Fig. 5.6a shows a simple CMOS inverter amplifier. Its low frequency gain is given by $A_{vo} = g_m r_{ds}$ And the unity gain bandwidth is given by $f_u = \frac{g_m}{C_s}$. The DC gain A_{vo} and the unity gain bandwidth f_u in the strong inversion region can be written in the following way:

$$A_{vo} = g_m r_{ds} \propto \sqrt{I_o} \cdot \frac{1}{I_o} = \frac{1}{\sqrt{I_o}} \tag{5.4a}$$

$$f_u = \frac{g_m}{C_s} \propto \sqrt{I_o} \tag{5.4b}$$

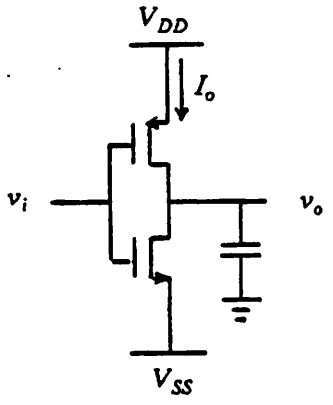
and in the weak inversion as:

$$A_{vo} = g_m r_{ds} \propto I_o \cdot \frac{1}{I_o} = constant \tag{5.5a}$$

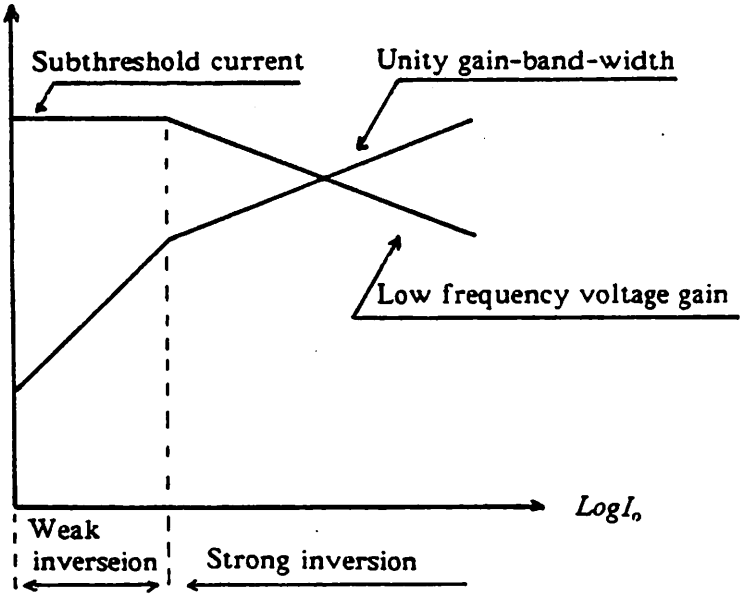
$$f_u = \frac{g_m}{C_s} \propto I_o \tag{5.5b}$$

The low frequency gain A_{vo} and the unity gain bandwidth f_u as function of the bias current I_o are shown in Fig. 5.6b. This figure shows that the inverter can be either at high unity gain bandwidth (fast transient) or at large DC gain. It can not have both properties at the same current level.

The use of class AB architecture also makes it possible to achieve the desired gain with small aspect ratio devices. The reason for this is that the gain per stage in a MOS amplifier is inversely proportional to the current level for a given device size. Thus the small signal voltage gain is high because of the low current level. And the fact that the small voltage gain falls off during the transient because of high current level is of little consequence in the application of the switched capacitor circuits. From this it follows that the desired gain can be maintained while reducing the device sizes provided that the current level is correspondingly reduced. This current reduction is achieved with the use of the class AB design. The combined effect of using class AB architecture together



(a)



(b)

Fig. 5.6 (a) CMOS inverter
(b) Unity gain bandwidth and low frequency gain as a function of bias current I_o

with a dynamic common mode feedback results also in a small filter power consumption.

The output swing tends to decrease when small W/L devices are used. This is because the output swing for a MOS amplifier is limited by the saturation voltage $\Delta V = \left\{ \frac{I}{\frac{\mu C_o}{2}\frac{W}{L}} \right\}^{1/2}$ of the devices in the output stage which, in turn, increases proportionally to the square root of their current level. As a consequence, the range of output voltages that correspond to the full voltage gain is increased as the current level is reduced. Because a class AB circuit can achieve the desired settling time with a smaller quiescent current level than a class A design, it follows that a class AB circuit gives a larger output swing for a given device size or alternatively can give the desired output swing with smaller device sizes than a class A op-amp. Although the above conclusion was derived without taking into account the effect of the cascode device, it is still possible to preserve a large output swing in a cascode amplifier combining a class AB design with dynamic biasing [37].

In summary, the advantages of a class AB op-amp are the following. The current level in a class AB op-amp is variable. At the beginning of a cycle, when a lot of current may be needed, a class AB op-amp can deliver a lot of current to charge or discharge the load quickly. A class AB op-amp does not have the slew rate problem, since the available slewing current is not limited by the biasing current in the usual sense. At the end of a cycle, however, the current level is much reduced. Thus it allows high swing, high gain and low power at the end of the cycle without sacrificing speed.

A class AB op-amp is a little complicated, however. It needs a structure to convert the input signal to internal current for the class AB operation. The current changes a lot from the standby value to the peak value during the transient. So the design of the biasing circuit is not trivial, since it should function in such a way that the op-amp can deliver as much current as possible to the output. One difficult problem in the design of the differential op-amp with a 5v power supply is being able to have a common mode feedback circuit that will not limit the output swing. This will be

discussed in this section.

Fig. 5.7 shows an example of a class AB op-amp block diagram [39]. This architecture is similar to the one used by Castello, etc. [37] but with dramatically reduced device sizes (Chapter 6). The input stage converts the input voltage to an internal current, which increases from its standby value with an increase in input voltage. This internal current is then mirrored at the top and at the bottom of the circuit to the output by the top and the bottom current mirrors. A cascode is used at the output to increase the output resistance and help provide high op-amp gain.

Realization and design of a class AB op-amp will be discussed in Chapter 6. In the following sections, dynamic common feedback, dynamic biasing and impact ionization protection that provide high gain, high swing and fast slewing for a class AB op-amp will be discussed.

5.2.3. Dynamic Common Mode Feedback

A fully differential op-amp architecture with common mode feedback is conceptually depicted in Fig. 5.8. The output voltage of the differential op-amp is determined by the difference of the output voltage $v_d = v_{o+} - v_{o-}$ (differential voltage) and the the average of the output voltage $v_{cm} = \frac{v_{o+} + v_{o-}}{2}$ (common mode voltage) with respect to a reference voltage V_{cm-} The former is controlled by the differential path op-amp as (DA) in Fig. 5.8. The latter is controlled by the common mode amplifier (CMFA) (Fig. 5.8). The differential path amplifier was discussed in the previous section. This section will discuss the common mode feedback (CMFB) circuit.

If the common mode voltage is continuously compared with a constant reference voltage e.g. ground by the common mode feedback circuit, then the common mode feedback circuit is referred to as a continuous common mode feedback. Fig. 5.9 shows an example of a differential op-amp with CCMFB (Continuous Common Mode Feed Back). On the other hand, if the output common mode voltage is periodically refreshed to the reference voltage [7] then the common mode feed back circuit is called the dynamic common mode feed back (DCMFB).

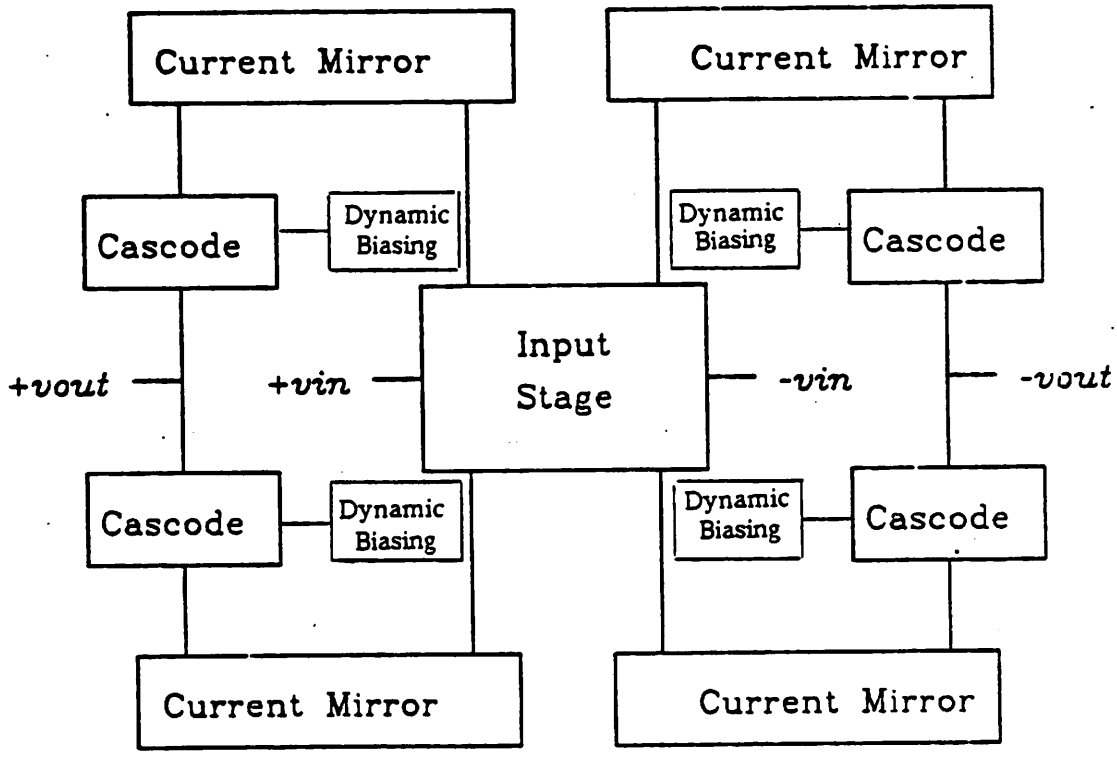


Fig. 5.7 The block diagram of a class AB op-amp example [39]

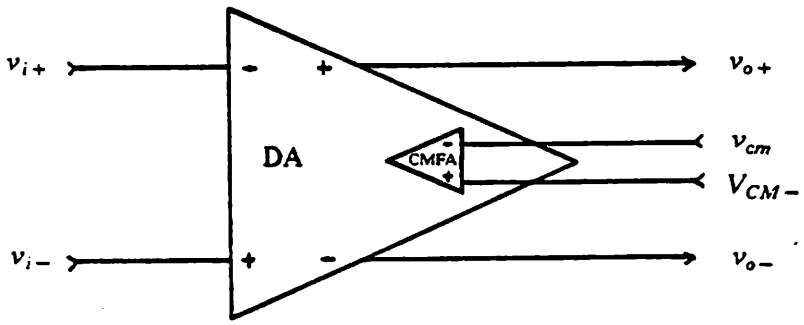


Fig. 5.8 Block diagram of a differential op-amp with CMFB

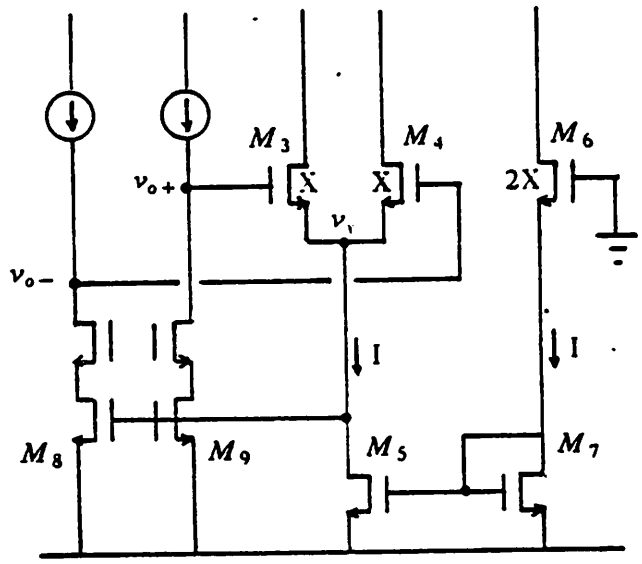


Fig. 5.9 Example of a continuous CMFB [38]

Fig. 5.10 shows an example of a switched capacitor dynamic common mode feedback circuit proposed by Senderowicz [7]. Together with a simple common mode feedback amplifier CMFA, the integrating capacitors, C_1 and C_2, and the switched capacitors, C_3 and C_4, provide the zero output common mode voltage while the output voltage swing is maximized. Capacitors C_1 and C_2 are chosen of the same value. These two capacitors divide the sum of the op-amp output voltages by two and thus produce a voltage v_{cm} common to both outputs. This can be easily seen from the following equation.

$$v_{cm} = \frac{v_{o+}C_1 + v_{o-}C_2}{C_1 + C_2 + C_{p1} + C_{p2}} \tag{5.6a}$$

$$= v_{o+} + \gamma v_{o-}\frac{C_1}{C_1(1+\gamma)+C_{p1}+C_{p2}} \tag{5.6b}$$

where v_{o+} and v_{o-} are the output voltages, C_{p1} and C_{p2} are the parasitic capacitors at the nodes of x (Fig. 5.10), and γ is the capacitor matching factor, which is close to 1 in practice.

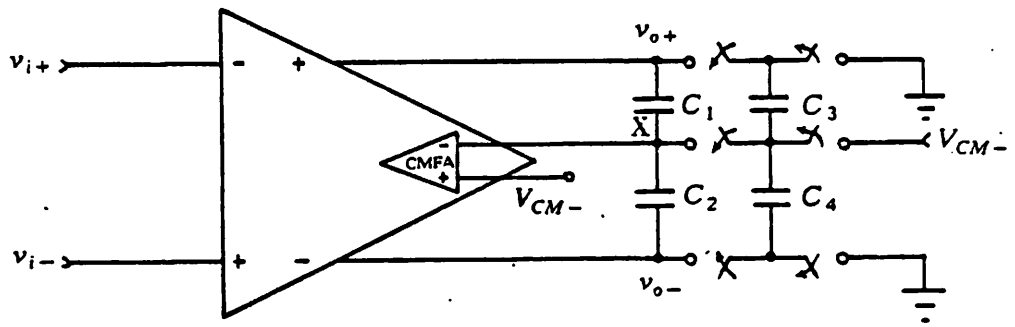


Fig. 5.10 Block diagram of a differential op-amp with DCMFB

If the matching of the capacitors C_1 and C_2 is good ($\gamma \approx 1$) and the parasitic capacitances C_{p1} and C_{p2} are small relative to the CMFB integrating capacitors C_1 and C_2, then the produced common voltage v_{cm} of (5.6) is approximately

$$v_{cm} \approx \frac{v_{o+} + \gamma v_{o-}}{1+\gamma} \approx \frac{v_{o+} + v_{o-}}{2} \tag{5.7}$$

where v_{cm} is the common voltage deviation from the bias voltage V_{cm-}. The total voltage at the node x is $v_{cm} + V_{cm-}$.

The operation of the switched capacitor CMFB shown in Fig. 5.10 involves mainly the capacitors. This operation of summation and division, performed as shown in (5.7), is not limited by the power supply. Hence the switched capacitor DCMFB (Fig. 5.10) will not limit the output swing. The continuous CMFB, on the other hand, will limit the output swing because it requires transistors

to sense the output swing. Fig. 5.9 shows an example of a continuous common mode feedback circuit. The common mode feedback circuit is composed of transistors M_3 to M_7. Gates of the input transistors M_3 and M_4 of the CMFB need a voltage of at least one threshold drop and one overdrive ($V_T + \Delta V$) in order to bias the transistors M_3 and M_4 in the saturation region. Because the op-amp outputs are the inputs of the CMFB transistors, the maximum output swing should be at least $V_T + \Delta V$ away from the power supply to guarantee that the input transistors M_3 and M_4 of the continuous common mode feedback are biased in the saturation region. For a $\pm 2.5V$ power supply, the maximum output swing is less than $\pm 1V$ for $V_T + \Delta V = 1.5V$.

Capacitors C_3 and C_4 in the dynamic common feedback (Fig. 5.10) periodically set up the proper voltage drop for the capacitors C_1 and C_2 to provide the zero output common mode voltage and bias the simple amplifier CMFA (Common Mode Feedback Amplifier) properly. Charge sharing between the capacitors C_1/C_2 and capacitors C_3/C_4 can also reduce the output common mode voltage. A circuit of only integrating capacitors $C_1 = C_2 \equiv C_i$ and switched capacitors $C_3 = C_4 \equiv C_s$ is shown in Fig. 5.11. V_{CM} is the bias voltage for the CMFA amplifier. Assume that there is a common mode voltage sensed at node x $v_x = v_{co}$ at time $t = 0$. The voltage v_x after i charge sharing cycles is v_{x_i} which is given by,

$$v_{x_i} = \left[v_{c(i-1)} \frac{C_i}{C_i + C_s} + V_{cm} \frac{C_i}{C_i + C_s} \right]$$

$$= v_{c(i-1)}\alpha + V_{cm}\beta$$

$$= v_{co}\alpha^i + V_{cm}\beta\left[1 + \alpha + \alpha^2 + \ldots\ldots + \alpha^{i-1}\right] \qquad (5.8)$$

where $\alpha = \dfrac{C_i}{C_i + C_s}$ and $\beta = \dfrac{C_s}{C_i + C_s}$. Practically, both α and β are < 1, and if the time is long enough, i.e. the i is large, then the v_{x_i} of (5.8) can be approximated as

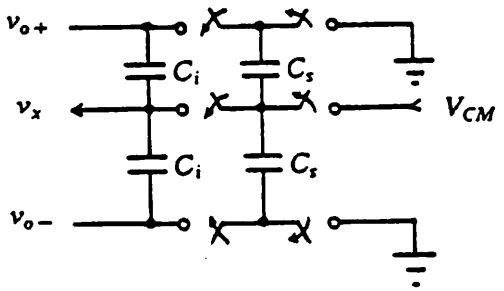


Fig. 5.11 DCMFB integrating and switched capacitors C_i and C_s respectively

$$v_{x_i} = v_{co}\alpha^i + V_{cm}\beta\left[\frac{1}{1-\alpha}\right] = 0 + V_{cm}\left[\frac{C_2}{1+C_2}\frac{1}{1-\frac{C_1}{C_1+C_2}}\right] = V_{cm} \tag{5.9}$$

From (5.9), it can bee seen that any of the perturbation of the output common mode voltage away from the V_{cm}, v_{co}, is eventually decreased to zero by the charge sharing process performed by the integrating capacitors C_i's and the switched capacitors C_s's.

Switched capacitor DCMFB is fast, because it requires mainly charging and discharging of the capacitors. On the other hand, continuous CMFB is slow compared with the switched capacitor DCMFB. Because the output swing is large, overdrives of the CMFB input transistors M_3 and M_4 have to be large in order to keep the input transistors M_3 and M_4 in saturation. This can be achieved by either increasing the current or decreasing the W/L of the input transistors M_3 and M_4. The previous approach will lead to more power consumption. The choice of the latter leads to

small g_m, which results in a slower circuit. That is, the continuous CMFB circuit can not be fast if the W/L ratio is reduced in order to have large output voltage swing.

The switched capacitor dynamic common feedback as shown in Fig. 5.10 improves the common mode nonlinearity problem. So long as capacitors C_1 and C_2 are equivalent and matched well, the sensed output common mode voltage v_{cm} is unchanged even for the large output swing as shown in (5.7). In contrast, for the continuous CMFB case, as in Fig. 5.9, the common mode voltage at node x is given by

$$v_x = \frac{v_{o+} + v_{o-}}{2} - V_T - \frac{\sqrt{I_o + \Delta I} + \sqrt{I_o - \Delta I}}{2} \tag{5.10}$$

where $I_o = \frac{I}{2}$ and ΔI is the current deviation from the standby current I_o. The sensed common mode voltage v_x relies on the square law of the device V-I characteristics and thus is not a linear function of the op-amp output voltage. The overall common mode voltage $v_x = v_{cm} + V_{cm}$ actually biases the transistor M_8 and M_9 (Fig. 5.9). Therefore, the output common mode voltage changes with the differential output voltages. This is more severe when the output swing is large, where V-I characteristics of the MOS transistor are nonlinear.

Node x in the switched capacitor DCMFB (Fig. 5.10) acts as a virtual ground for the signal. So capacitors C_1 and C_2 can be used also for compensation. The DCMFB consumes mostly transient power because the circuit is mainly composed of capacitors.

The DCMFB circuit uses the clock to periodically refresh a bias voltage on level shifting capacitors C_1 and C_2, therefore, some clocking noise and charge injection are expected, but these noises are canceled to the first order by the fully differential architecture.

Practical considerations on the noise contribution due to the DCMFB circuit and the nonideal clocks will be discussed in chapter 6.

5.2.4. Dynamic Biasing

The basic concept of dynamic biasing is that the biasing current of the op-amp is not constant as in the case of the statically biased op-amp, but changes during the op-amp operation [28]. In the switched capacitor application, the dynamic biasing can provide a lot of biasing current in the beginning of the the clock phase ϕ_T (Fig. 5.5). However, the bias current decreases toward high DC gain at the end of the ϕ_T. The dynamic bias circuit starts its operation in the strong inversion region with large unity gain bandwidth as can be seen from Fig. 5.6b, and then shifts into the weak inversion region with large DC gain at the end of ϕ_T, until the current goes almost to zero. Subthreshold operation of the MOS transistors has the advantages of low offset voltage, low temperature drift and good life drift stability [41]. There are several ways to accomplish dynamic biasing. One way is by pulsed tail current biasing [41] and the other way is by adaptive biasing [42,37].

Based on the structure of classical single ended op-amps, Hosticka proposed dynamic op-amps where the current sources were replaced by pulsed current sources [41]. The basic objective of pulsed biasing is the same as that of the MOS class AB op-amp discussed in the previous section. They both take advantage of the MOS electrical characteristics to achieve large DC gain and fast response.

The adaptive biasing scheme is a dynamic biasing method whose biasing current level is made signal-dependent. Benefiting from the advantages of the MOS characteristics, the op-amp is biased at low current level with high gain if no signal is applied, and the op-amp biasing current increases only when the signal applies. The adaptive biasing op-amp has the advantages mentioned, namely, high unity gain bandwidth f_u and large DC gain A_{vo} for fast and complete charge transfer, which is essential for the switched capacitor application. Several circuits have been realized. Vittoz's and Castello's dynamic biasing are signal dependent and not pulsed. Castello's adaptive biasing is used in conjunction with a class AB op-amp so that it can deliver as much current as possible to the output during transient [37].

The architecture of the dynamic biasing circuit used in this project shown in Fig. 5.7 is similar to the one used by Castello [37]. The cascode output transistor is dynamically biased in such a way that its bias voltage adapts to the changes of the input voltage during the transient. This adapted bias prevents the cascode output transistor from going to the linear region during the time when the output is delivering a lot of current to the load.

5.2.5. Impact Ionization Shielding

Effects of impact ionization on analog circuit performance have been discussed in the previous chapter. In this section, circuit solutions to the resistance degradation due to impact ionization will be shown. Normally, the impact ionization of the NMOS transistor starts for smaller drain-source voltages than those of the PMOS transistor [11] (Chapter 7). Therefore, the measures are taken mainly for impact ionization of the NMOS transistor.

A simple circuit solution to prevent impact ionization is to add an extra NMOS transistor in series with the output cascode as shown in Fig. 5.12. This extra transistor buffers the NMOS output cascode transistor, M_1, reducing its maximum drain to source voltage to a small fraction of the output voltage when the output goes in the positive direction. When the output goes in a negative direction, on the other hand, this shield transistor M_s does not appreciably reduce the output negative swing because the shield transistor, M_s, operates in the linear region. The shield transistor causes nonlinearity to the op-amp transfer characteristics because the shield transistor operates in different regions, i.e. saturation region and linear region. Fig. 5.13 shows another application of the shield transistor [(null)]. Notice that in this circuit, the gate of the shield transistor M_s is biased by a bias regulator so that the source of the shield transistor is set at the midpoint of the power supply [(null)]. In the case of ± 2.5 V power supply, the midpoint voltage is ground.

In a p-well operation, the impact ionization can be solved by simply tying the source of the NMOS transistor to its own independent well because the current created by the impact ionization will flow to the source instead of flowing to the substrate. The transistor output impedance there-

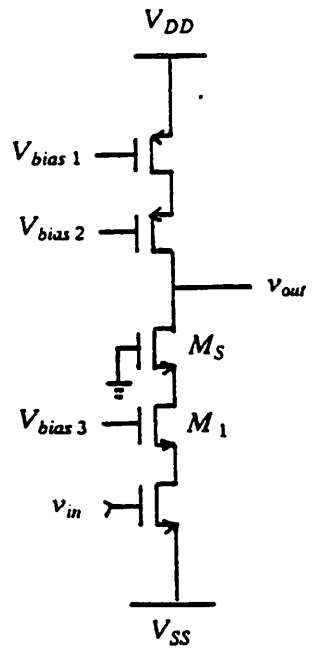


Fig. 5.12 NMOS transistor M_s reduces the impact of the impact ionization [43]

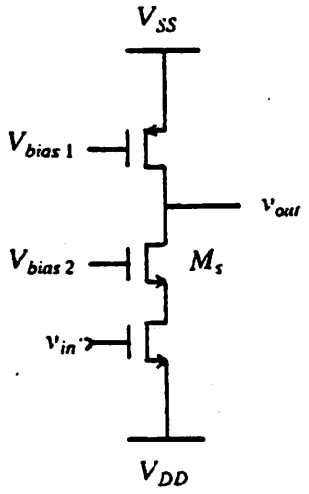


Fig. 5.13 Shield transistor M_s to minimize impact ionization

fore is not degraded by the equivalent shunt resistance R_{eq} (Fig. 4.4) due to impact ionization effect.

5.3. Summary

Circuit approaches to minimize the impact of the practical limitations on the MOS analog circuits in scaled technology were discussed. Chopper stabilization is used to remove the 1/f noise. A class AB op-amp increases the output swing, and the gain without sacrificing the speed. A fully differential architecture, dynamic common mode feedback and dynamic biasing also provide high output swing.

CHAPTER 6

AN EXPERIMENTAL SCALABLE SWITCHED-CAPACITOR FILTER

Fig. 6.1 shows the area and power consumption of some commercially available 5th. order switched capacitor PCM filters. Since the first chip was introduced in the market in 1977 and before the work described here, about a factor of 3 reduction in area has occurred. Even with this reduction, the smallest such filters required an area of about 400 mil^2 per pole. This is far greater than the theoretical minimum value, which has been shown in Fig. 4.6 and Fig. 4.7 to be 24 mil^2 of a fifth order PCM switched capacitor filter at a dynamic range of 95 dB if a power supply 5 V is used [29].

In this chapter, design techniques for a switched capacitor PCM transmitting filter suitable for fabrication in a scaled technology will be described. The primary objective of this design is to achieve the smallest possible area and power consumption. A fundamental requirement for this design is to be able to meet all the performance specifications while using a supply voltage that is compatible with that of the digital technology. The circuit design solutions adopted to overcome problems associated with the scaling of the analog circuits discussed in Chapter 5 will be explained. An additional result from this study may be the extension of frequency range over which the switched capacitor filters are useful because it is possible to design a fast op-amp by using a scaled technology.

6.1. Single Stage Chopper Stabilized Operational Amplifier

In the op-amp design, the 1/f noise problem introduced in the use of minimum device sizes is eliminated by employing a chopper stabilization technique. To achieve a voltage gain adequate for the switched capacitor application i.e. 60 dB or more, while using the minimum channel length devices, both two stage op-amps and single stage op-amps with cascode devices in series with the

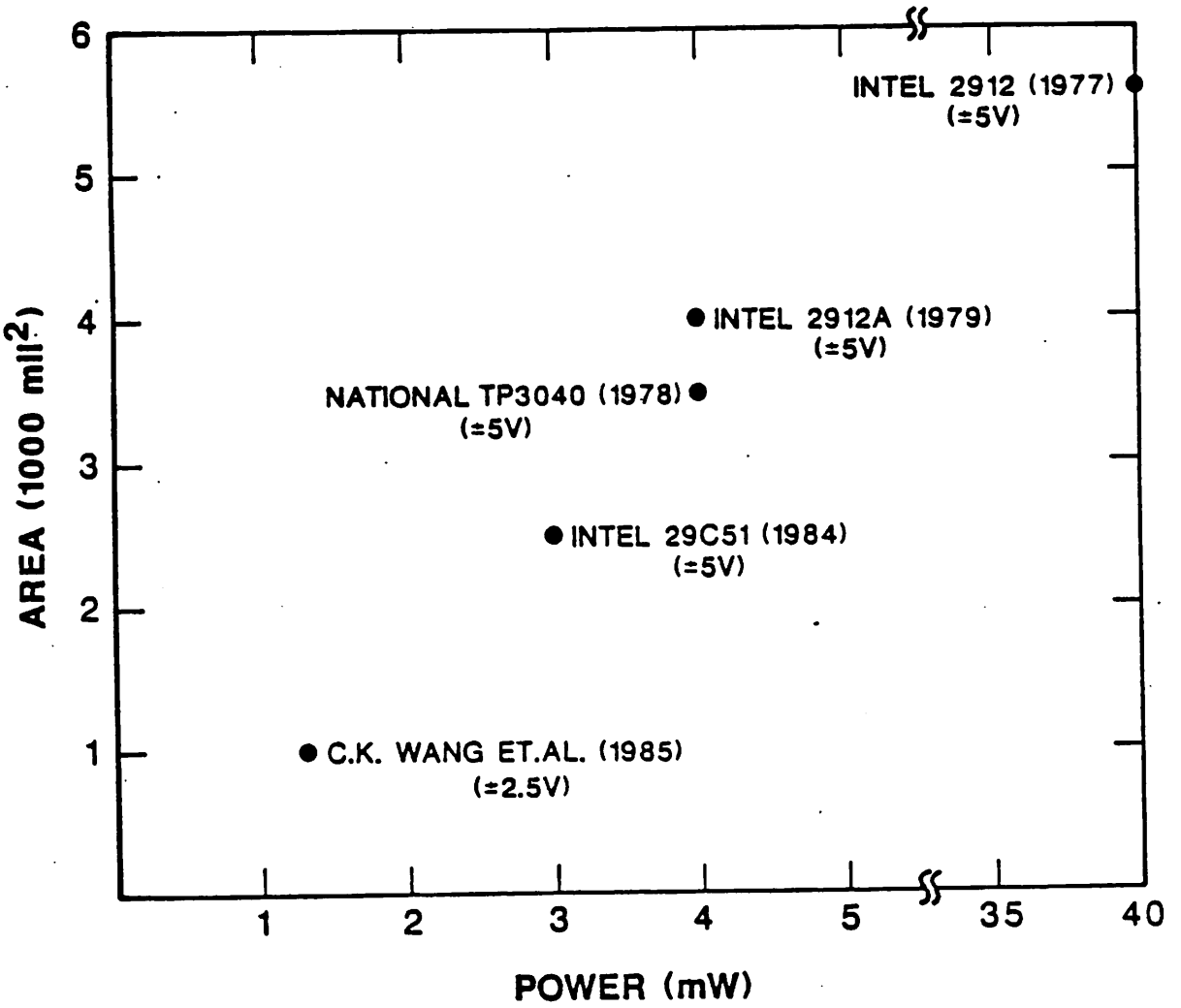


Fig. 6.1 Area and power dissipation of some commercially available 5th. order PCM switched capacitor filters together with the chip prototype of this project.

output transistors are considered. Because the single stage op-amp has only one high impedance node, does not need any compensation capacitors, and has good power supply rejection even at high frequency, it is selected for use in this project. Its major drawback is that the output swing is limited (especially when small W/L devices are used), however, this problem is virtually eliminated by using the class AB op-amp with dynamic common mode feedback discussed in Chapter 5. The circuits of a single stage class AB chopper stabilized op-amp with dynamic common mode feedback will be discussed in this section. The block diagram was shown in Fig. 5.7.

6.1.1. Overall Architecture

The circuit schematic of the operational amplifier is shown in Fig. 6.2. Other than the chopper stabilization technique employed, this circuit architecture is similar to the one described before [37], but with dramatically reduced device sizes (Table 6.1). The cross coupled transistors $M_1 - M_4$ convert the input voltage signal into two internal currents and provide class AB operation [43] and dynamic biasing. For maximum internal current given an input differential voltage, transistors $M_1 - M_4$ should be operated in the saturation region. The input common mode voltage that permits these transistors to operate in saturation for the largest range of input differential voltage is approximately equal to $\Delta V_{6,8} + V_T$, assuming $\Delta V_6 = \Delta V_8 \equiv \Delta V_{6,8}$ and $V_{TP} = V_{TN} \equiv V_T$. The maximum corresponding current is given by [39]

$$(I_{12})_{max} \approx \frac{\mu C_{ox}}{2}(\frac{W}{L})_{12}(\frac{V_{dd} - V_{ss}}{2} - V_T - \Delta V_{1,4})^2 \qquad (6.1)$$

The range of the input voltage swing is the smaller one of either the sum of the overdrives of the level shifter composed of transistors $M_5 - M_8$ or the maximum negative input swing available. Transistors $M_5 - M_8$ perform input level shifting and provide gate bias for the transistors M_3 and M_4. The optimum range of the input voltage swing is give by

$$\frac{1}{2}(\Delta V_{n5} + \Delta V_{p7}) = \frac{1}{3}[V_{ss} - (\Delta V_I + V_{Tn5} + V_{Tp7})] \qquad (6.2)$$

where ΔV_{n5}, ΔV_{p7} and ΔV_I are the transistor overdrives of the level shifter and the current source respectively. Four current mirrors $M_9 - M_{12}$ at the top and at the bottom mirror the input currents

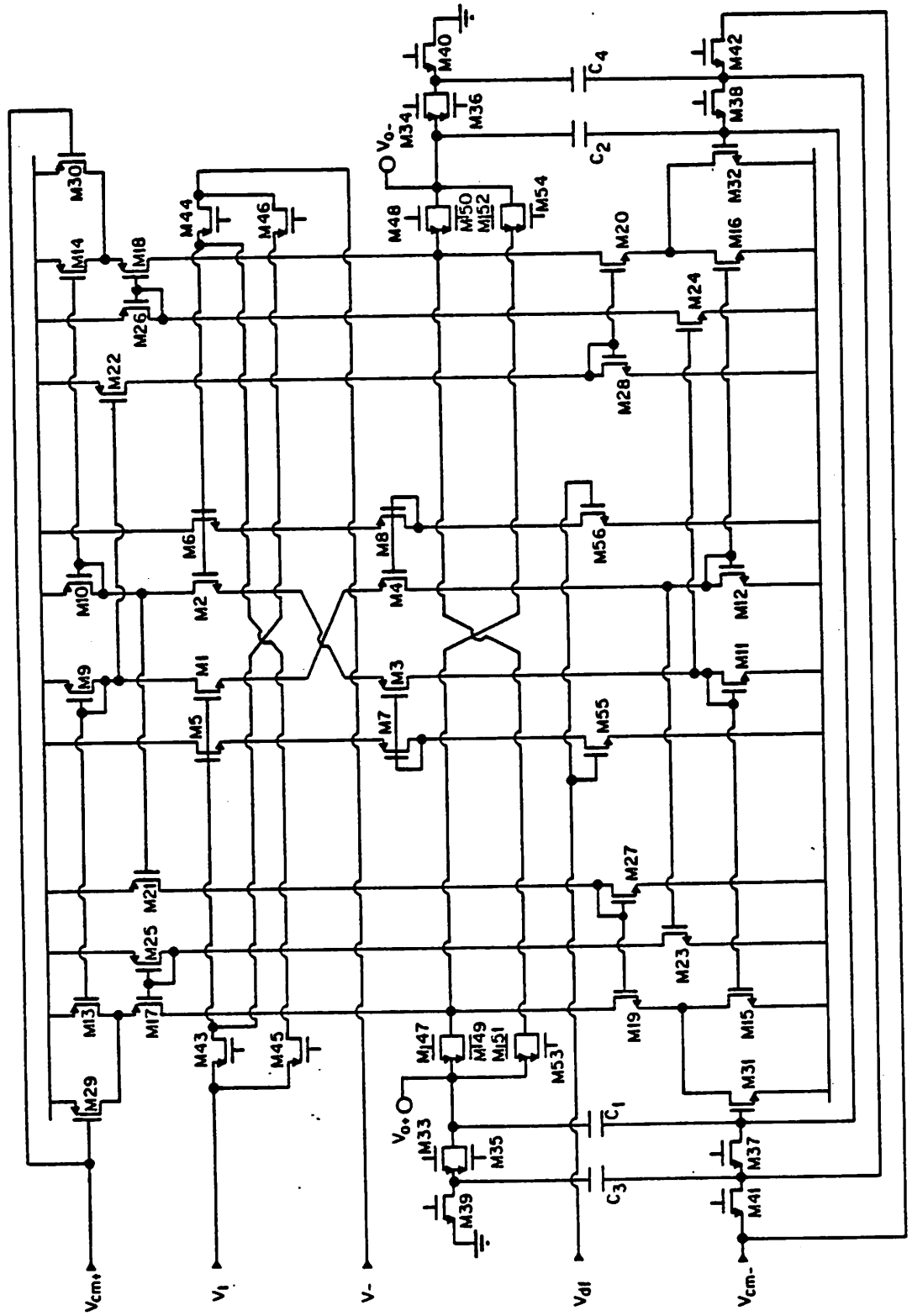


Fig. 6.2 Circuit schematic of a single stage class AB chopper stabilized operational amplifier with DCMFB

Device	W(μm)	L(μm)	Device	W(μm)	L(μm)
M1	25	3	M19	7	3.5
M2	25	3	M20	7	3.5
M3	25	3	M21	17.5	3.5
M4	25	3	M22	17.5	3.5
M5	25	3	M23	7	3.5
M6	25	3	M24	7	3.5
M7	25	3	M25	3.5	3.5
M8	25	3	M26	3.5	3.5
M9	10	3.5	M27	3	7
M10	10	3.5	M28	3	7
M11	4	3.5	M29	12	3.5
M12	4	3.5	M30	12	3.5
M13	17.5	3.5	M31	18	3.5
M14	17.5	3.5	M32	18	3.5
M15	7	3.5	M33-M34	7	3
M16	7	3.5	M55	7	3
M17	17.5	3.5	M56	7	3
M18	17.5	3.5			

Table 6.1 Transistor W/L dimensions of the chopper stabilized class AB operational amplifier as shown in Fig. 6.2

to the output for the current gain and dynamic biasing. Transistors $M_{17} - M_{20}$ are cascode transistors used to increase the output impedance and hence increase the gain.

A summary of equations describing the op-amp performance is given in table Table 6.2.

It is interesting to note that the cross coupled class AB input composed of transistors $M_1 - M_4$ can be used as a voltage-to-current transducer with high linearity and wide dynamic range (Appendix A).

The output cascode transistors $M_{17} - M_{20}$ are biased by a dynamic biasing circuit, which is composed of transistors $M_{21} - M_{28}$. The dynamic biasing circuit keeps the output transistors (M_{13} - M_{16}) in saturation for as wide a range of output currents as possible. A biasing circuit to achieve the maximum output swing [24] is employed. The ratio of the width of the mirror transistor of the dynamic biasing circuit to the width of the output cascode transistor, i.e. $\frac{W_{25}}{W_{17}}$, must be less than 1/4 so that the output cascode transistors are biased in saturation. To compensate for the body effect on the cascode transistors, this ratio is selected as 1/5.

The output common mode DC voltage (ground) is set by the dynamic common mode feedback circuit, which is composed of transistors $M_{29} - M_{42}$ and capacitors C1-C4 [7,37]. The overall operation of the DCMFB was discussed in the previous chapter. Basically, capacitors C_1 and C_2 provide AC path from the the op-amp outputs, V_{o+} and V_{o-}, to the gates of common mode feedback transistors, M_{31} and M_{32}. Capacitors C_1 and C_2 sum the output voltages and divide them by 2. Hence, the DCMFB circuit produces the common mode portion of the output voltage without affecting the differential voltage. The produced common mode voltage drives the common mode feedback amplifiers (CMFA) transistors M_{31} and M_{32}. Capacitors C_3 and C_4 provide proper DC values for the output common mode voltage and the bias of the common mode feedback amplifiers. That is, capacitors C_3 and C_4 periodically restore the voltage drop for DCMFB integrating capacitors C_1 and C_2 so that output common mode voltage is defined (as ground in this design) and common mode feedback transistors M_{32} and M_{32} are properly biased. So, transistors M_{31} and M_{31} take just enough current to make the output CM voltage equal to zero. Choice of the capacitance value

A. Overall Transconductance

$$\frac{4}{(\frac{1}{k_1^{1/2}}+\frac{1}{k_4^{1/2}})^2}(\Delta V_6+\Delta V_8)(\frac{W_{16}}{W_{12}})$$

B. DC Gain

$$\frac{4}{(\frac{1}{k_1^{1/2}}+\frac{1}{k_4^{1/2}})^2}(\Delta V_6+\Delta V_8)(\frac{W_{16}}{W_{12}})(g_{m20}r_{O20}r_{O16})$$

C. Unity Gain Frequency

$$\frac{4}{(\frac{1}{k_1^{1/2}}+\frac{1}{k_4^{1/2}})^2}(\Delta V_6+\Delta V_8)(\frac{W_{16}}{W_{12}})\frac{1}{C_L}$$

D. Non-dominant Poles at

D1. Gate of current mirrors M_{12}

$$\frac{g_{m12}}{C_{gs12}+C_{gs16}}$$

D2. Gate of cascode transistor M_{20}

$$\frac{g_{m20}}{C_{gs20}}$$

E. Input Doublets

$$pole=\frac{g_{m1}+g_{m4}}{C_{gs1}+C_{gs4}}$$

$$zero=\frac{g_{m1}}{C_{gs1}}$$

Table 6.2 Equations of the op-amp performance

of capacitors C_1 and C_2 is associated with the compensation of the op-amp stability and common mode settling time. Capacitors C_3 and C_4 cause voltage spikes each time when the switches $M_{33} - M_{38}$ close. The magnitude of the spike v_{sk} is given by

$$v_{sk} = v_{omax} \frac{C_3}{C_f + C_1 + C_3} \tag{6.3}$$

where C_f is the integrating capacitor and v_{omax} is the op-amp maximum output voltage available. If C_3 is large, it takes less number of cycles to reduce output common mode voltage (cf Equation (5.9)). However, if C_3 is large, the magnitude of the spike becomes so large that it takes a longer transient time for the op-amp to settle to its original steady state value. As a common practice, DCMFB sampling capacitors C_3 and C_4 are chosen such that their value is less than that of the DCMFB integrating capacitor C_1 and C_2.

Transistors $M_{43} - M_{46}$ and $M_{47} - M_{54}$ form the input chopper and the output chopper, respectively, which are controlled by two nonoverlapping clocks. Each set of chopper transistors steers the signal one way or the other. CMOS switches are used at the output chopper in order to accommodate the large output positive and negative swings. The input chopper, however, uses only the PMOS switches. This is because the input is biased at a common voltage larger than 1V and NMOS switches in p-type wells exhibit high threshold voltage due to body effect. Thus the NMOS switch is not efficient for the input chopper.

The transistor sizes are listed in Table 6.1. All minimum transistor sizes are used except when needed to optimize circuit performance. If the input NMOS transistors each has an independent well, the W/L of these devices is smaller than for the case where they share one well. This is because, with only one well, the body effect degrades the transconductance of the input devices and the W/L must be increased to compensate for this degradation. Although the W/L's are larger with only one well than with separate wells, the required area with only one well is less than with separate wells because of the well to well spacing requirements. The input transistors $M_1 - M_4$ and bias transistors $M_5 - M_8$, have a W/L ratio of 25/3, which is the largest device size among all op-amp transistors.

One practical consideration is the the op-amp noise. Noise contribution from dynamic biasing transistors $M_{25} - M_{28}$ and output cascode transistors $M_{17} - M_{20}$ are negligible because noise voltages at their gates are referred to the input through the transconductances of the cascode transistors (M_{17} - M_{20}), which are degraded by the output impedances of the output transistors $M_{13} - M_{16}$ and are therefore smaller than those due to output transistors $M_{13} - M_{16}$. The total equivalent input noise of the op-amp is given by the following equation:

$$\begin{aligned}\overline{v_{ieq}^2} = &(\sum_{i=1}^{i=8}\overline{v_i^2}) + 2\left\{\overline{v_{14}^2}(\frac{g_{m14}}{g_T})^2 + \overline{v_{16}^2}(\frac{g_{m16}}{g_T})^2\right\} \\ &+ 2\left\{\overline{v_{10}^2}(\frac{g_{m10}}{g_T})^2\left[(\frac{W_{14}}{W_{10}})^2 + (\frac{W_{15}}{W_{11}})^2\right]\right\} \\ &+ 2\left\{\overline{v_{12}^2}(\frac{g_{m12}}{g_T})^2\left[(\frac{W_{16}}{W_{12}})^2 + (\frac{W_{13}}{W_9})^2\right]\right\} \\ &+ 2\left[\overline{v_{mc30}^2}(\frac{g_{mc30}}{g_T})^2 + \overline{v_{mc32}^2}(\frac{g_{mc32}}{g_T})^2\right] \qquad (6.4)\end{aligned}$$

where g_T is the overall transconductance given in Table 6.2, and $\overline{v_j^2}$ is the equivalent gate noise power density of each MOS transistor for $j = 1 - 8$.

6.1.2. Simulation Results

The requirements of the op-amp low frequency gain A_{omin} and closed loop unity gain frequency $\omega_{c_{load}}$ in a SCF application can be calculated by the following equations [24];

$$A_{omin} = Q_{int}\left\{\frac{C_s + C_n + C_I}{C_s}\right\} \qquad (6.5)$$

where Q_{int} is the integrator Q factor, C_s is the sampling capacitor, C_n is the input parasitic capacitor and C_I is the integrating capacitor and

$$\omega_{c_{load}} = \frac{1}{T_s}\ln\left[Q_{int}\frac{C_s}{C_I}\right] \qquad (6.6)$$

where T_s is the total available integration time. For a nominal integrator $Q_{int} = 40$, and the $\frac{C_s}{C_I} \approx \frac{1}{20}$, the required DC gain is $A_{omin} \approx 800$ and the required closed loop unity gain frequency is

$\omega_{c_{load}} \approx 2\pi \cdot 0.1\ mHz.$

SPICE simulation shows that the open loop gain of the designed op-amp is 64 dB. With an output capacitance of 1.5pF, the unity gain frequency of the op-amp is 10 MHz with a phase margin of 55°. The duty cycle of the main filter clock of frequency 256kHz is 40%. Therefore the setting time should be smaller than 1562 ns. The op-amp settling time to 0.1 % of differential output 2 V step is 670 ns with an output capacitive load of 4.5pF. A summary of the op-amp SPICE simulated performance is shown in Table 6.3.

Fig. 6.3b shows the SPICE simulation of a chopper stabilized switched capacitor integrator. A top plate sampling is used for the case of SPICE simulation. Notice that only the polarity of the op-amp outputs (nodes 3 and 4 in Fig. 6.3) changes when the chop clock changes state. There is an obvious spike on the rising edge of the sampling clock ϕ_S (Fig. 6.3b) when at the same time the DCMFB circuit starts to refresh its reference voltage drop for the capacitors C_1 and C_2 so that the output common mode voltage and the bias of the common mode feedback transistors are refreshed. The spike (Equation (6.3)) has little effect on the filter performance (which will be explained more in detail in section 6.5) because the magnitude of the spike is small and the time duration is short. So, the output voltages will quickly restore their final values determined by the top plate of the integrating capacitor C_I.

Fig. 6.4 shows the feedthrough overshoot when the ϕ_T clock turns on. This occurs because it takes time for the op-amp to react as a negative feedback circuit. The magnitude of the overshoot (Fig. 6.4) is given by

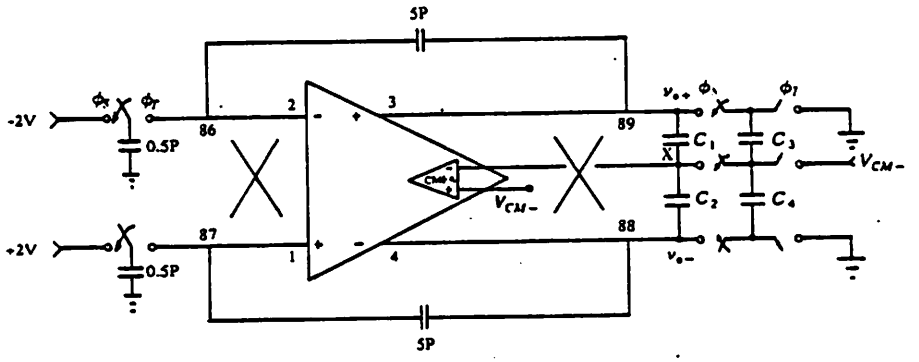
$$\Delta V_1 = \left[(\Delta V_2 \frac{C_I}{C_s}) \frac{C_s}{C_s + \frac{C_I C_L}{C_I + C_L}} \right] \frac{C_I}{C_I + C_L} \tag{6.7}$$

where ΔV_2 is the magnitude of the output step as shown in Fig. 6.4, capacitor C_L is the load capacitor. The larger the magnitude of the overshoot, the longer the slewing time is required.

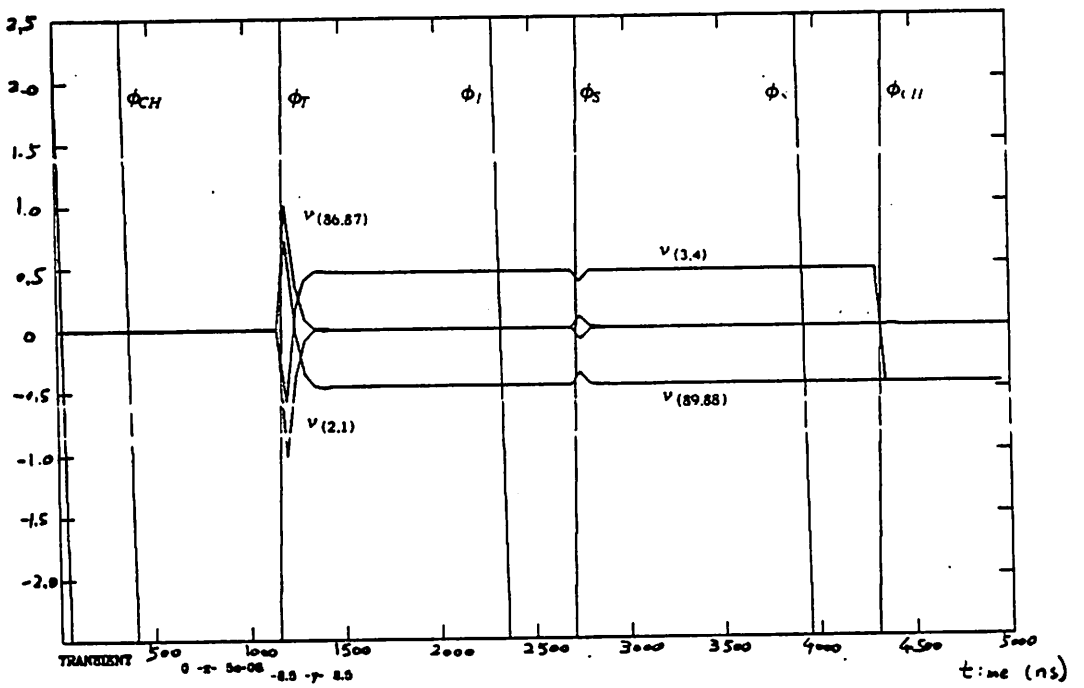
Supply voltage	±2.5	V
Active area*1	80	mil^2
Open loop voltage gain	64	dB
Power (stand by)	220	μ W
Ouput Swing	±2	V
Input-referred noise (@1KHz)	320	$nV/Hz^{1/2}$
Unity-gain bandwidth (@1.5PF load)	10	MHz
±1 V_{output} settle to within 1mV (capacitive load 4.5pF)	670	ns

Table 6.3 Op-amp designed performance

*1 Includes choppers and chopper bus lines

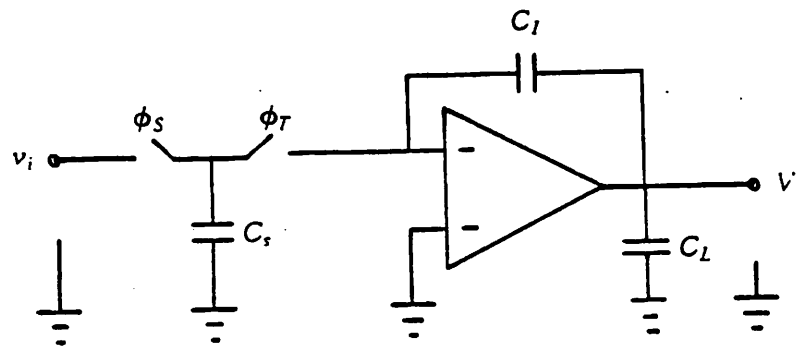


(a)

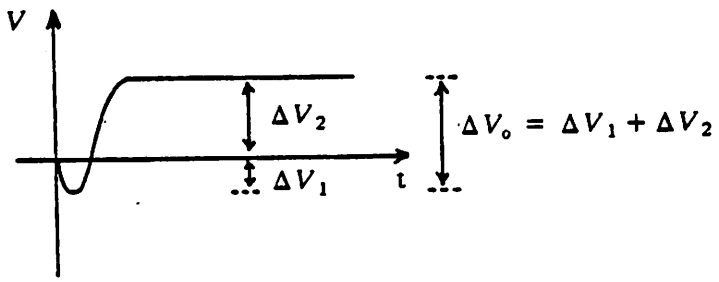


(b)

Fig. 6.3 SPICE simulation of a chopper stabilized integrator



(a)



(b)

Fig. 6.4 (a) Switched capacitor integrator
(b) Overshoot of an integrator at the beginning of clock ϕ_T

6.2. A Fifth Order LP PCM Chopper Stabilized Switched-Capacitor Filter

Applying the design techniques for scaling MOS analog circuits discussed in Chapter 5, an experimental high performance scalable switched capacitor filter has been implemented. The switched capacitor PCM filter realized is a fifth order elliptic low pass ladder filter with four transmission zeros. The filter requires five op-amps. Fig. 6.5 shows the filter in passive components from which the switched capacitor filter is derived. The pole/zero location and the frequency response of this filter are shown in Fig. 6.6 (not to scale).

Fig. 6.7 shows the microphotograph of the experimental chip prototype. The chip was fabricated in a 3μm p-well CMOS process with a gate oxide thickness of 500 Å and a capacitor oxide tickness of 600 Å. The process has one layer of metal, two layers of polysilicon. The metal pitch is 7μm. The total active area of the filter is 1000 square mils or 200 square mils per pole, where

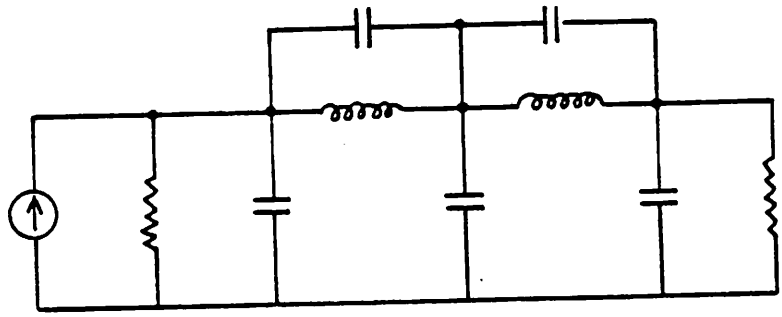
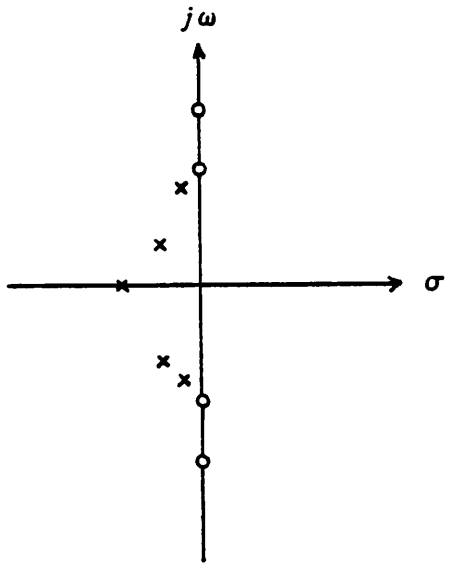
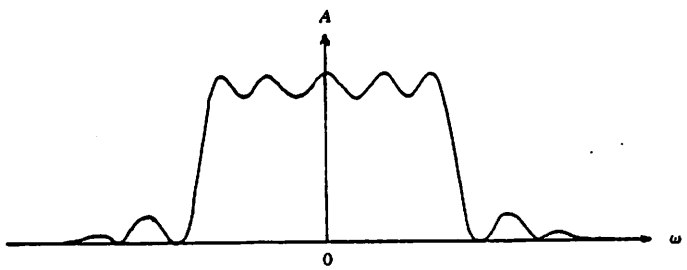


Fig. 6.5 A 5th. order elliptic low-pass LC ladder filter



(a)



(b)

Fig. 6.6 (a) Poles and zeros in an S place
(b) Magnitude frequency response of the LC ladder filter shown in Fig. 6.5 (not to scale)

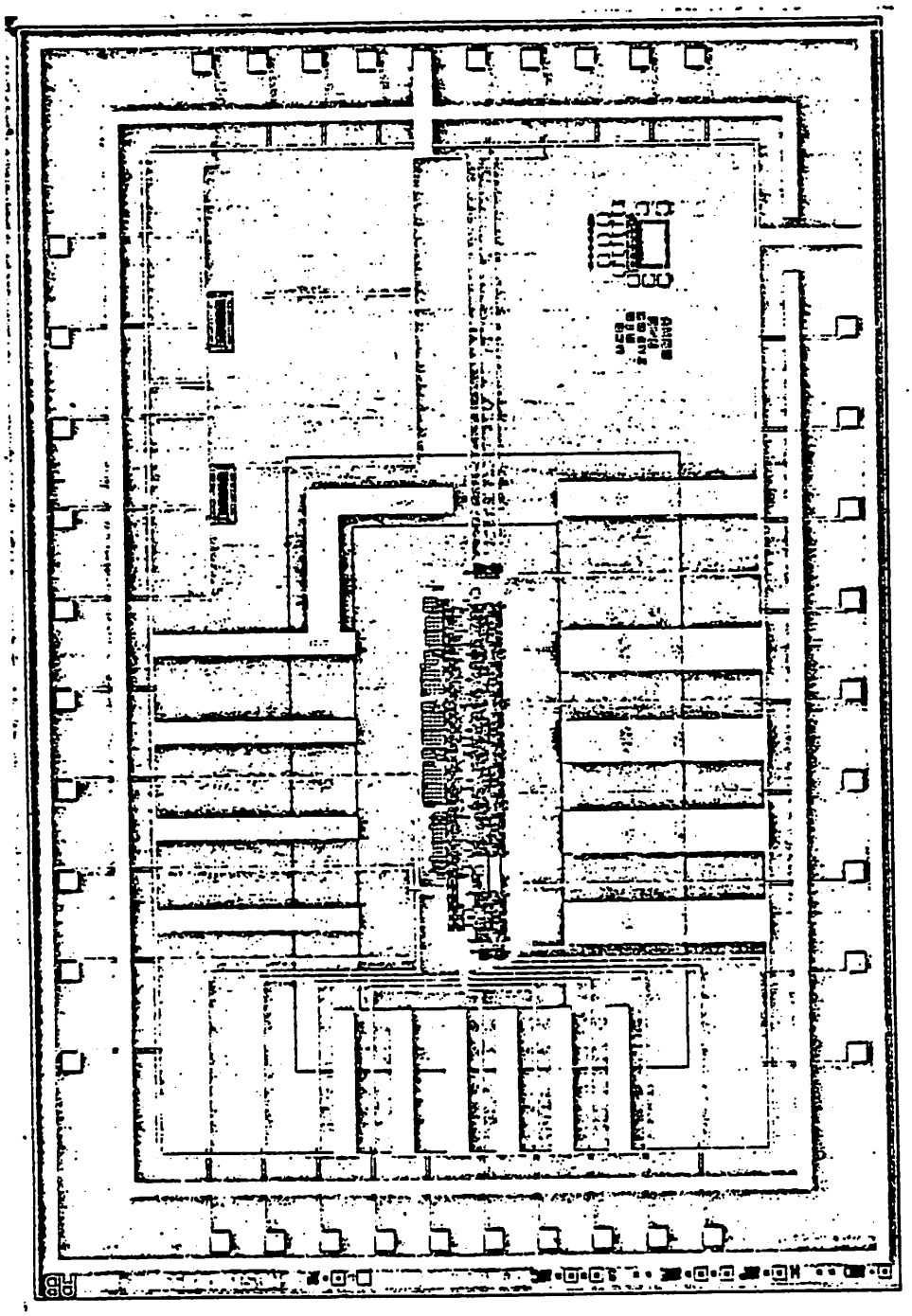


Fig. 6.7 Microphotograph of the chip prototype

the active area is interpreted as the nominal area of the minimum rectangular enclosing the whole filter, but not including clock generators.

The schematic of the switched capacitor PCM filter with 3 dB bandwidth 3.4 kHZ is shown in Fig. 6.8. SWAP [45] simulation of this chopper stabilized SCF is shown in Fig. 6.9. The capacitance of the unit capacitor is 0.26pF. One unit capacitor of 21μ by 21μ is used for all the sampling capacitors except the termination capacitors in the first and the last integrators.

In the following sections, practical design considerations on the scalable switched capacitor filter will be discussed. These are the fundamental limitation of KT/C noise, chopper stabilized integrator for the switched capacitor filters, clocking sequence of the chopper stabilized filter and layout strategies to minimize the area.

6.2.1. KT/C Noise Limited Dynamic Range

In order to achieve the required dynamic range of the filter (85dB or more) for a total supply voltage of 5 volts, the kT/C noise must be kept sufficiently small. This dictates a minimum value for the unit capacitors used in the integrators. A lower bound for the unit capacitor can be obtained by considering the kT/C limited dynamic range [29]. The actual value of the unit capacitor used has to be larger than this lower bound by an amount that depends on the particular filter structure and operational amplifier noise contribution. In this design, the final value of the unit capacitor 0.26pF was chosen based on the result of the filter noise simulation performed with the simulation program SWAP [45].

6.2.2. A Bottom Plate Chopper Stabilized Integrator

A basic building block for a chopper stabilized switched capacitor filter is shown in Fig. 6.10. The chopper stabilized integrator is composed of input bottom plate switched sampling capacitors and a chopper stabilized op-amp with dynamic common mode feedback circuits. PMOS switches only are used at the top plate of input sampling switches (cf Section 6.1.1). On the other hand,

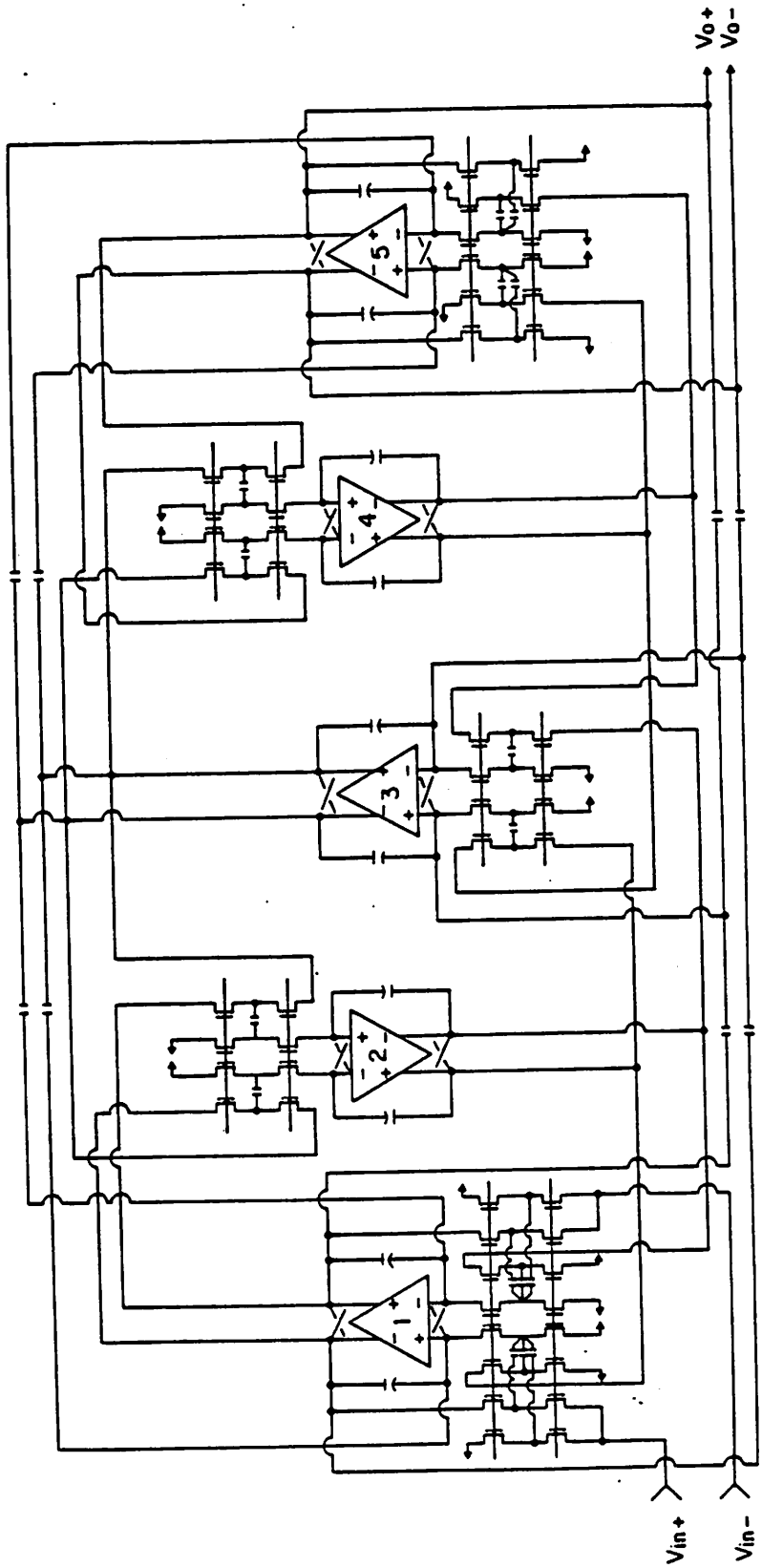


Fig. 6.8 Schematic diagram of the 5th. order PCM chopper stabilized low-pass switched capacitor filter

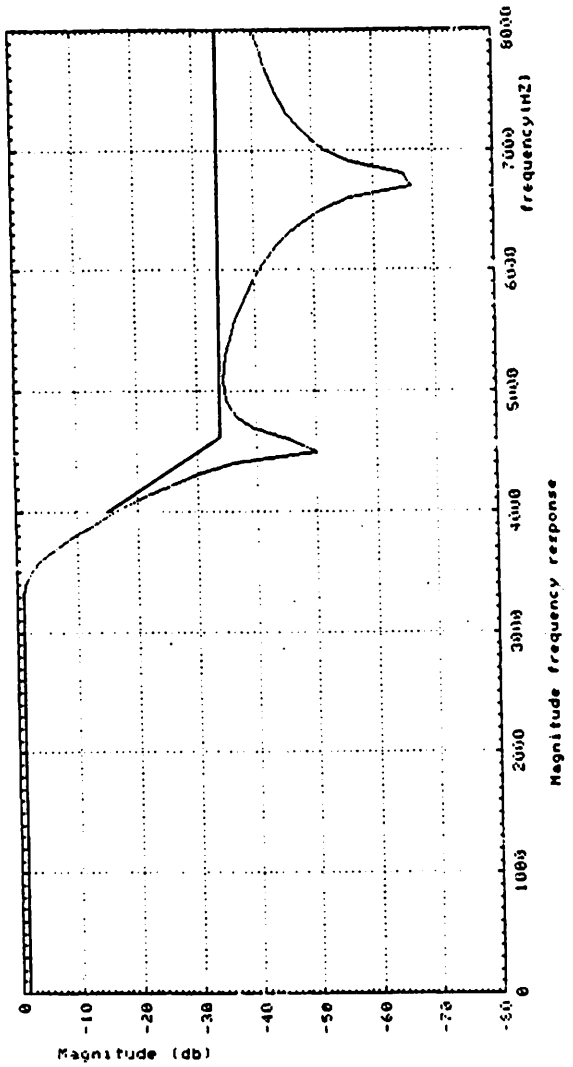


Fig. 6.9 SWAP simulation of the magnitude frequency response of the chopper stabilized SCF

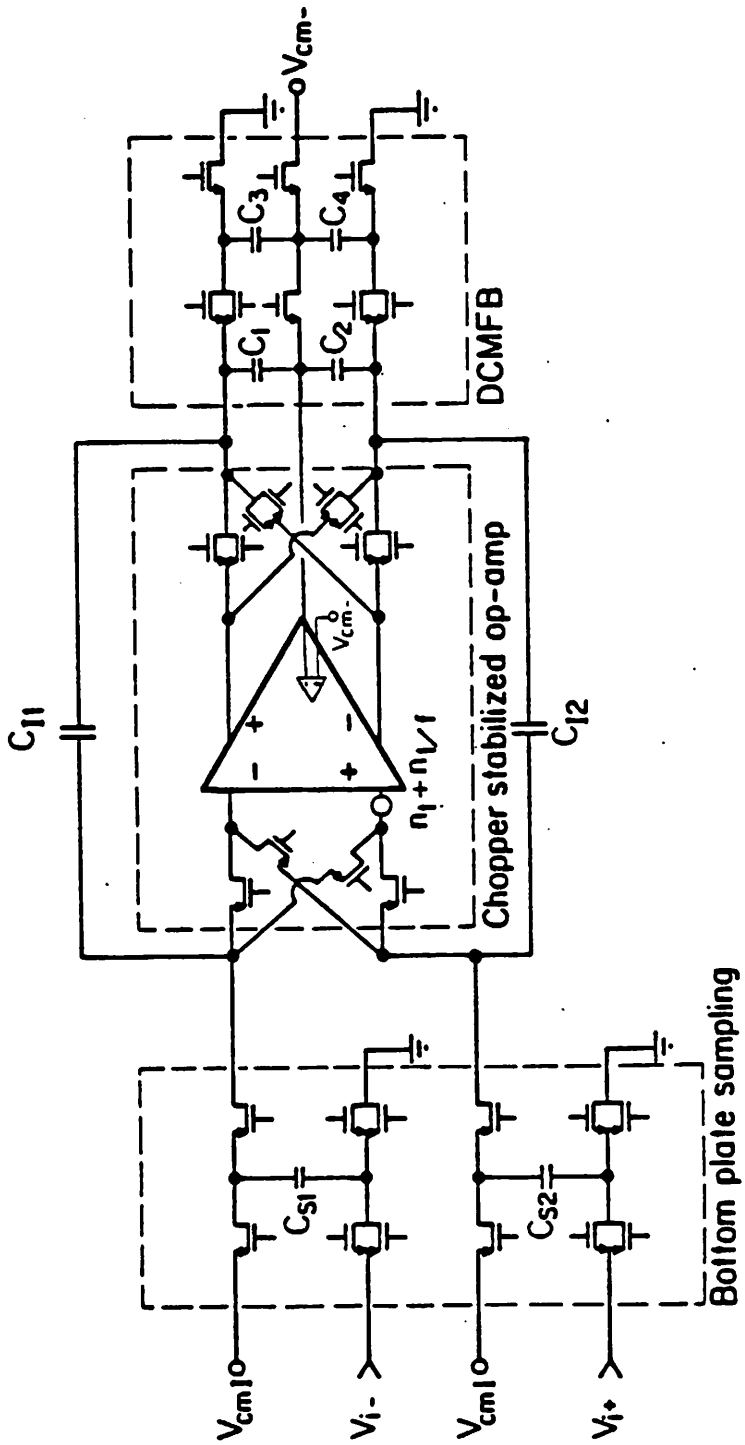


Fig. 6.10 Chopper stabilized integrator with DCMFB

because switches $M_{37} - M_{42}$ (Fig. 6.2) are associated with ground potential or negative bias, NMOS transistors are used. CMOS switches however, must be used at the op-amp output chopper in order to pass both positive and negative large signals.

The choppers are implemented by switches. This particularly simple implementation is possible because both signal polarities are available in a fully differential configuration. Note the signal polarity between the input and the output choppers is reversed twice during each chopper period. Signal polarity before the input chopper and after the output chopper, on the other hand, remains the same during any chopper period. Hence, the DCMFB circuit is placed after the output chopper instead of between the choppers in order to eliminate the need to charge and discharge the DCMFB capacitors $C_1 - C_4$ once per period.

6.2.3. Filter Clocking Sequence and Considerations

The entire circuit runs on a two phase nonoverlapping clock of frequency 256 kHz for the bottom plate sampling capacitors. The chopper runs at a frequency that is one half of the main filter clock frequency. The chopper clocks change state during the idle time between the sampling and the transferring of the integrator as shown in Fig. 6.11a. DCMFB uses the main filter clocks, which are also used for the bottom plate sampling and transferring. Only when the integrator samples the input, capacitors C_3 and C_4 set up reference voltages for capacitors C_1 and C_2. When the integrator transfers charge, capacitors C_3 and C_4 are refreshed to V_{cm-} instead (Fig. 6.10). The relative clocking sequence of sampling clock ϕ_S, transferring clock ϕ_T and chopper clock ϕ_{CH} is summarized in Fig. 6.11b.

Because the filter runs at a frequency of 256 kHz and the chopper clock runs at 128 kHz with a 50% duty cycle, 1/f noise components at the output of the chopper stabilized op-amp are moved to odd harmonics of the chopper frequency 128kHz, i.e. at $128\cdot(2m + 1)$kHz where m is an integer. From this argument, it can be easily seen that the op-amp output signal spectrum sampled by the following stage are located at the multiples of the main filter frequency 256 kHz, i.e. at

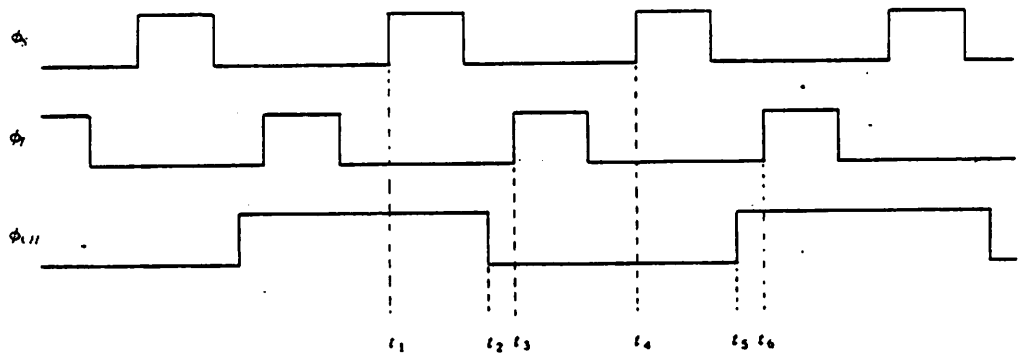


Fig. 6.11 (a) Clocking sequence of the main filter clock, chopper clock and the DCMFB clock

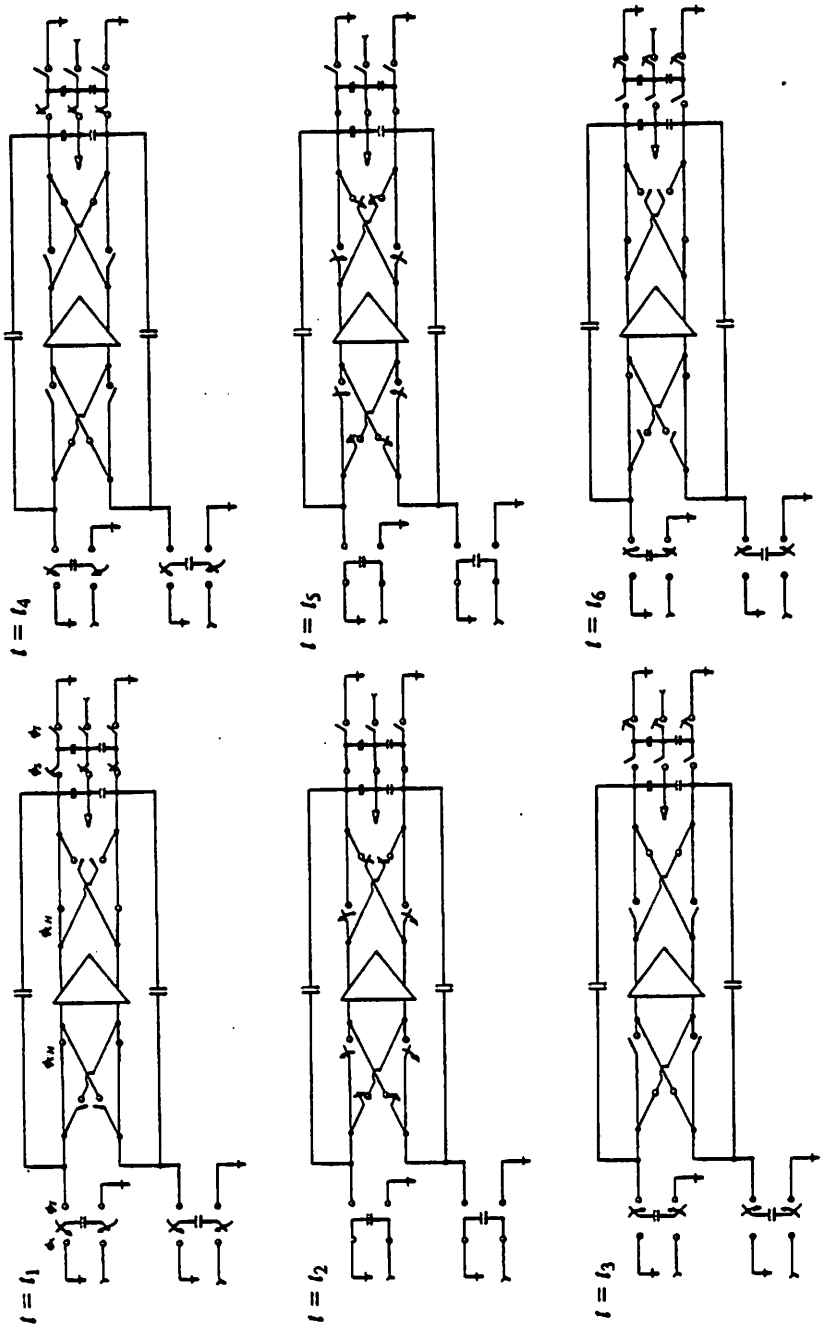


Fig. 6.11 (b) Switching sequence of the sampling switches, transfering switches and chopper switches

$128 \cdot (2m)$kHz. Signal and 1/f noise are alternate in the frequency domain. Fig. 6.12 shows the signal and 1/f noise spectrum of an open loop chopper stabilized op-amp sampled by a main filter clock of frequency 256 kHz. 1/f noise components are equally spaced in between the integer multiples of the filter clock frequency 256 kHz. This is optimum in terms of the aliasing minimum 1/f noise to the passband.

6.3. Layout Strategies

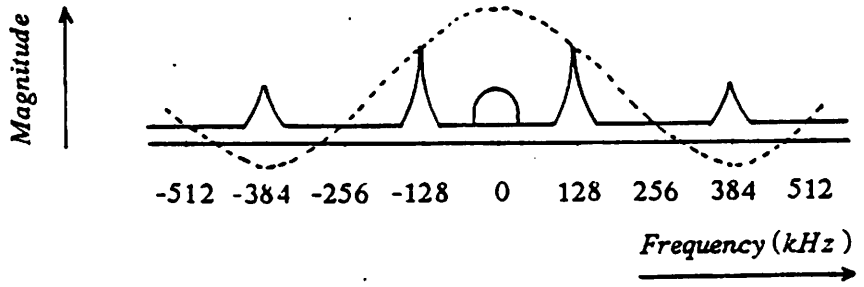
One of the research goals of this project is to achieve an area for the switched capacitor filter that approaches the theoretical minimum value. The metal pitch of this 3μ CMOS technology is 7μ and determines the minimum area possible. Therefore, in addition to using minimum transistor sizes whenever possible, some other layout strategies were used.

Symmetric layout is maintained throughout the filter as much as possible in order to achieve the advantages of a differential circuit.

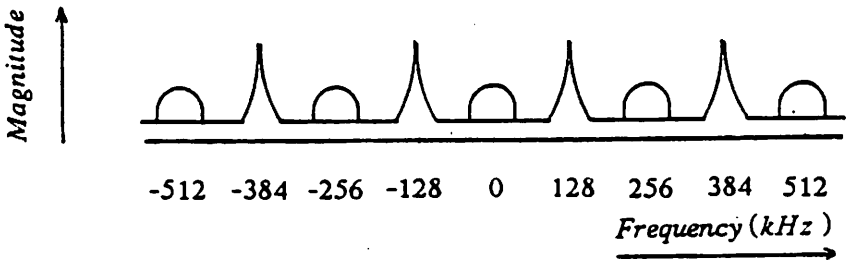
Chopper switches with chopper clock bus lines are laid out together with op-amp core transistors. To ease interconnection and save area, some of op-amp signal lines in metal pass over the gates of input transistors. Any coupling of spurious noises due to these overpassing metal lines are primarily common mode signals and hopefully do not affect the differential signals.

All clock bus lines run in polysilicon and also extend to form switches. Because of the small size of the filter, the RC time constant associated with the polysilicon clock lines is only several nano seconds, which is negligible for the clock rate used in the filter (256 kHz). In a scaled implementation, the RC time constant of a polysilicon line is reduced if the sheet resistance is still the same after the scaling (Table 3.2).

DCMFB capacitors $C_1 - C_4$ do not consume extra silicon area, since they are laid out underneath the interconnection metal bus lines. Any noise picked up by these capacitors is seen primarily as a common mode signal if the capacitors match well.



(a)



(b)

Fig. 6.12 (a) Signal and 1/f noise spectrum of an open loop chopper stabilized op-amp DCMFB clocked at 128 kHz
(b) sampled at output by a main filter clock of frequency 256 kHz

The resulting area of a chopper stabilized op-amp is 80 mil^2. The overall active area of the filter achieves 200 mil^2 per pole.

6.4. Experimental Results

In this section, measured data from the chip prototype will be discussed.

The filter supply voltage is ± 2.5V, and the total power consumption is 1.25 mW or 250 μW per pole. An overall filter response is shown in Fig. 6.13. The 3dB bandwidth is 3.4 kHz. The pass-band ripple is well within ± 0.125 dB. The stop band attenuation is better than 34 dB for frequencies higher than 4.6kHz. All of the above data agree well with simulated results obtained from the program SWAP. Detailed plots of the passband response of the filter for different values of the supply voltage are shown in Fig. 6.14. Curve (a) is for the nominal value of ± 2.5 V while curve (b) and (c) correspond to a 10 % increase and decrease of the supply voltage respectively. The magnitude variation is less than 0.005 dB. This demonstrates the small sensitivity of the filter to a change in the power supply voltage.

Power supply rejection ratio as a function of frequency in the range 0 to 10 kHz for both positive and negative supply is illustrated by Fig. 6.15. Better than 50 dB for positive and 40 dB for negative rejection at 1 kHz are achieved. The rejection of both negative and positive supplies for frequencies up to 1 MHz is shown in Fig. 6.16. This is an important specification for a scaled switched-capacitor filter because a large amount of digital noise energy easily falls in this frequency range.

Fig. 6.17 shows the measured equivalent input noise of the open loop op-amp with and without chopper. The op-amp noise is reduced by 20 dB at frequency 1 kHz if the DCMFB and the chopper are clocked at 256 kHz and 128 kHz respectively. The input equivalent op-amp noises with DCMFB clocked at 25 kHz (chopper clocked at 12.5 kHz) are shown in Fig. 6.18. It can be seen that the 1/f noise spectrum is, as expected, moved to 12.5 kHz and 37.5 kHz which are the odd harmonics of the chopper frequency 12.5 kHz. The noise at 25 kHz is corresponding to the

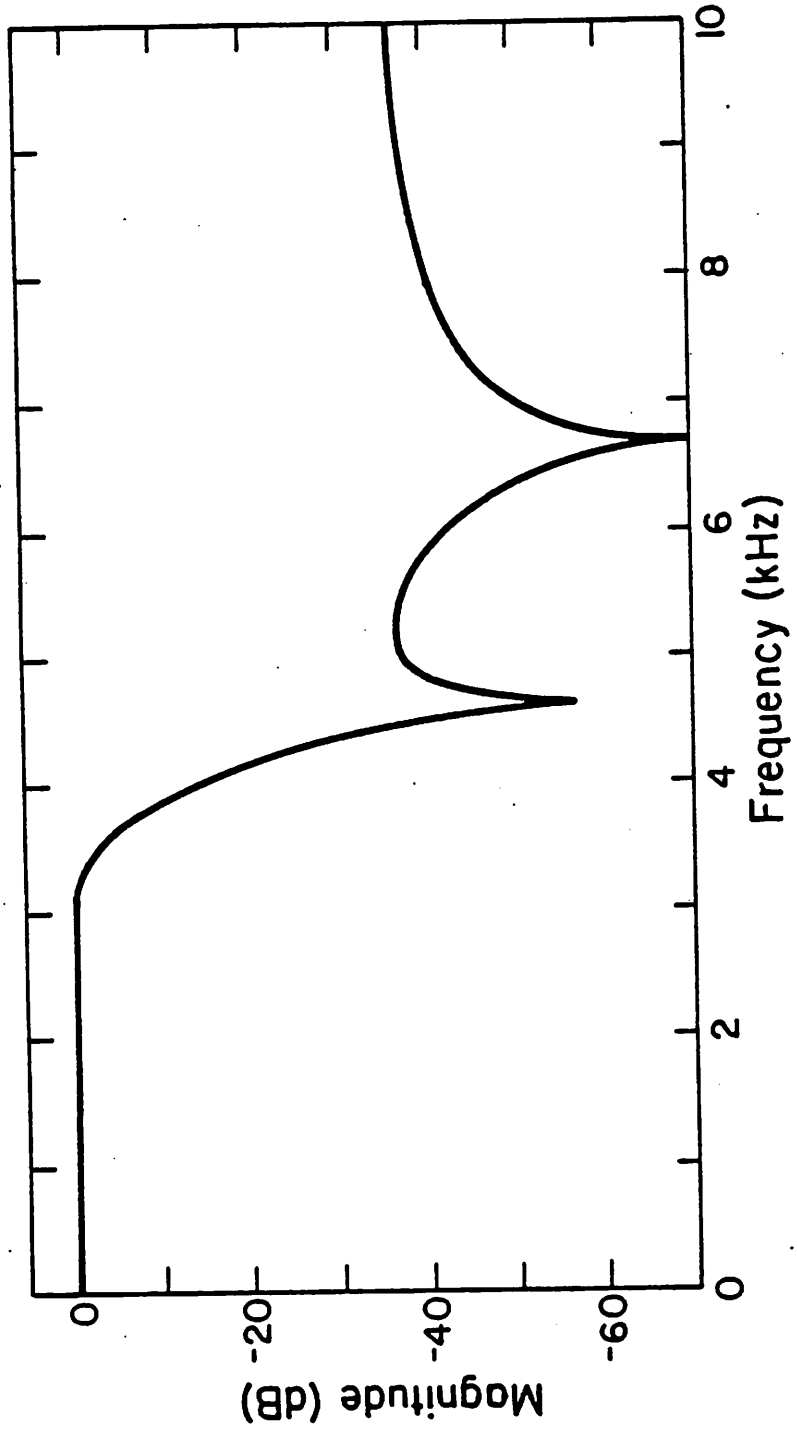


Fig. 6.13 Measured overall magnitude frequency response of the experimental filter

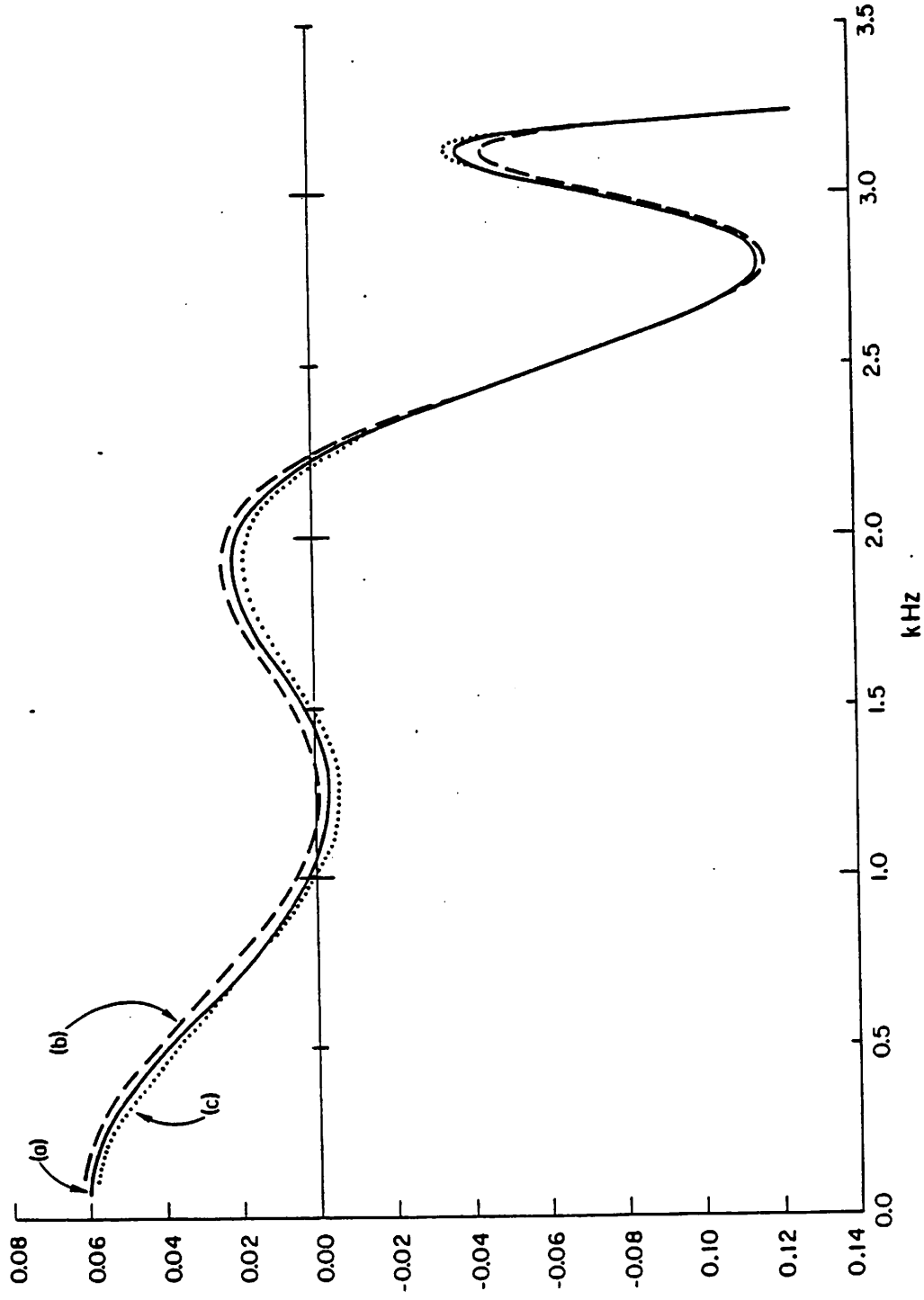


Fig. 6.14 Detailed passband ripples
(a) ± 2.5 V power supply
(b) ± 2.75 V power supply
(c) ± 2.25 V power supply

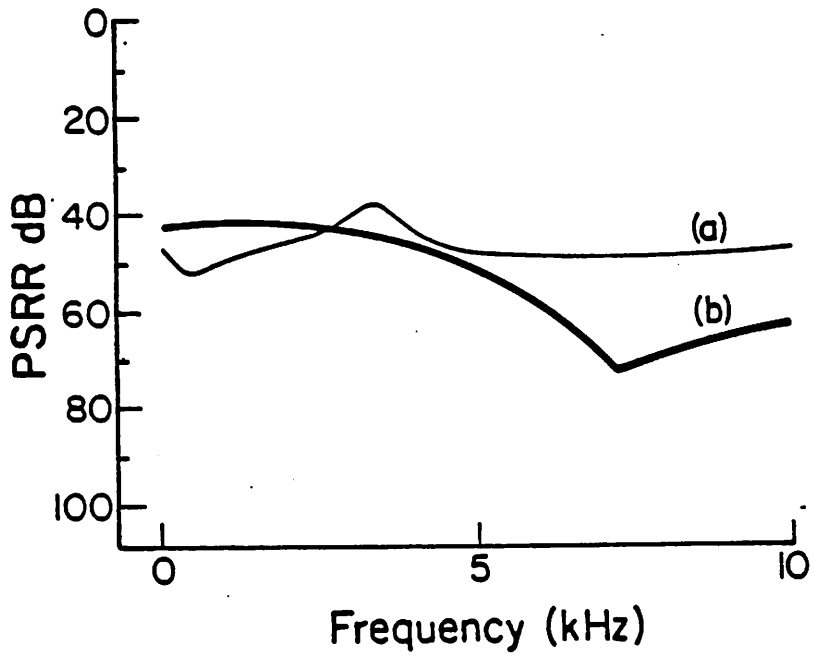


Fig. 6.15 Power supply rejection ratio (PSRR) for frequency range from 0 to 10 kHz
(a) positive PSRR
(b) negative PSRR

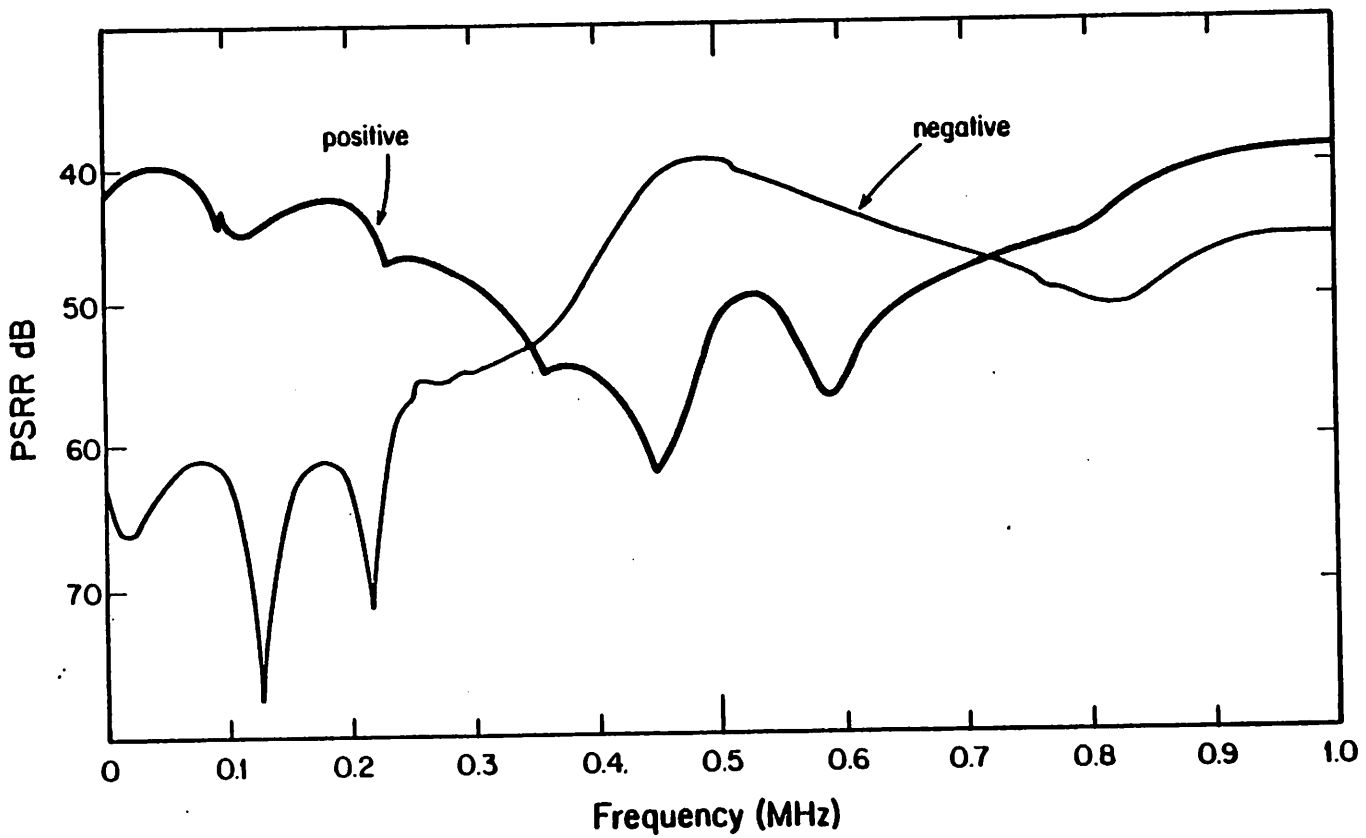


Fig. 6.16 Power supply rejection ratio for frequency up to 1 MHz

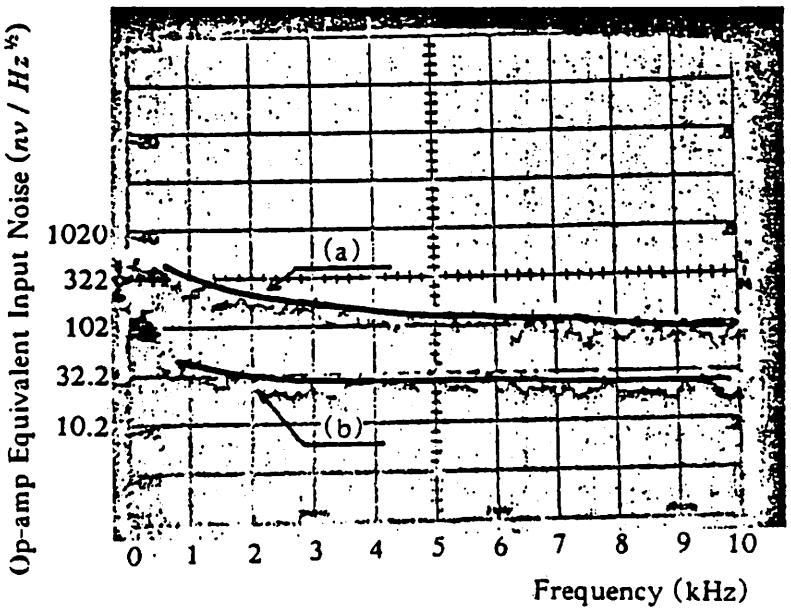


Fig. 6.17 Measured equivalent input noise spectrum of the open loop op-amp of frequency range from 0 to 10 kHz.
(a) without chopper
(b) with chopper of frequency 128kHz

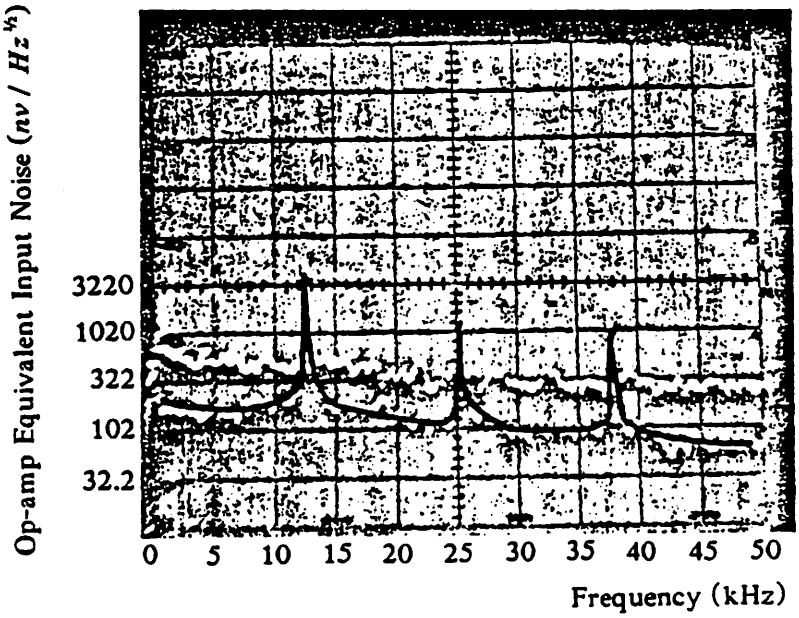


Fig. 6.18 Measured equivalent input noise spectrum of the open loop op-amp with chopper of frequency 12.5 kHz

spike generated by the DCFMB circuit which uses a clock rate of 25 kHz.

The filter output noise spectrum with and without chopper stabilization is shown in Fig. 6.19. The noise reduction at 1 kHz due to the choppers of frequency 256 kHz is 4 dB. Fig. 6.20 shows the measured filter output noise of main filter clock 25 kHz and chopper clock 12.5kHz. The 1/f noise harmonics appear at 12.5 kHz and 37.5 kHz which are the odd harmonics of the chopper frequency 12.5 kHz. The noise component at 25 kHz, as in the op-amp case, is the spike noise caused by the DCMFB circuit. In a switched capacitor application, DCMFB spike does not affect the passband signal, because the DCMFB circuit must be designed in such a way, as mentioned earlier, that the magnitude and the time duration of the spike are so small that the op-amp outputs will return to their original voltages quickly. Besides this, the DCMFB clocking noise is located at 256 kHz which is much higher than the 3 kHz of the passband frequency range of interest. There is nevertheless, a solution to totally remove this DCMFB spike noise aliased to the passband. This will be discussed in section 6.5.

Filter noise improvement due to the chopper will be larger in finer linewidth implementations. The noise reduction by the chopper is small due to the fact that the 1/f noise level for this particular run of the process was smaller than expected. Moreover, KT/C noise dominates and is not affected by the chopper stabilization technique. This will be discussed in section 6.5. The total c-message weighted filter output noise for the nominal clock (256 kHz) is 135 μV. The maximum differential output swing for a 1 kHz sinusoidal input that gives less than 1% total harmonic distortion (THD) is 3.11 V_{rms}. The resulting dynamic range is 87 dB.

The total harmonic distortion for a 4 V_{pp} output at 1 kHz is – 67 dB. A summary of the filter performance is shown in table Table 6.4.

6.5. Discussions

Two distinct features of the op-amp architecture employed are the dynamic common mode feedback and the chopper stabilization technique. The major concern is the extra noise and area

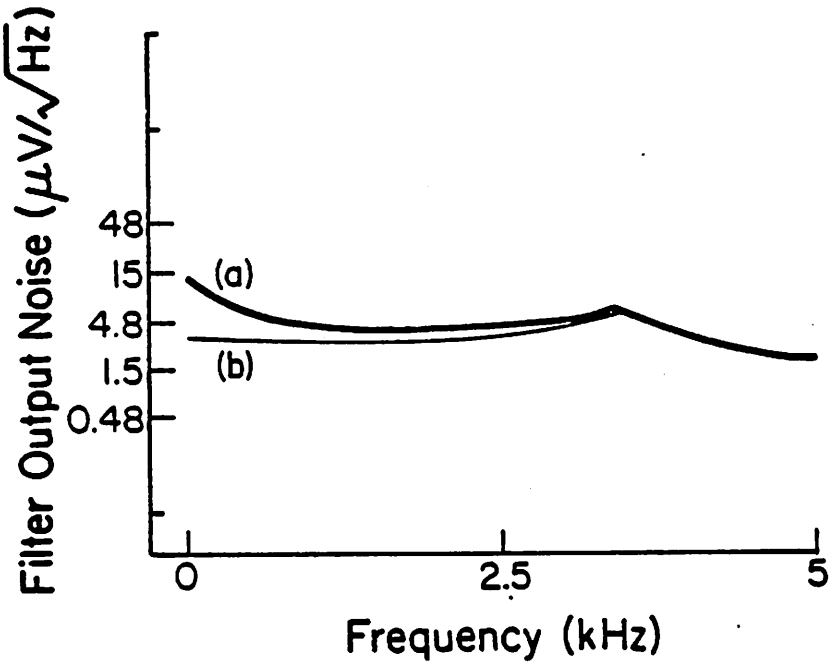


Fig. 6.19 Measured filter output noise spectrum of frequency from o to 5 kHz.
There is 4 db noise reduction at 1 kHz due to chopper.
(a) without chopper
(b) with chopper frequency of 128 kHz

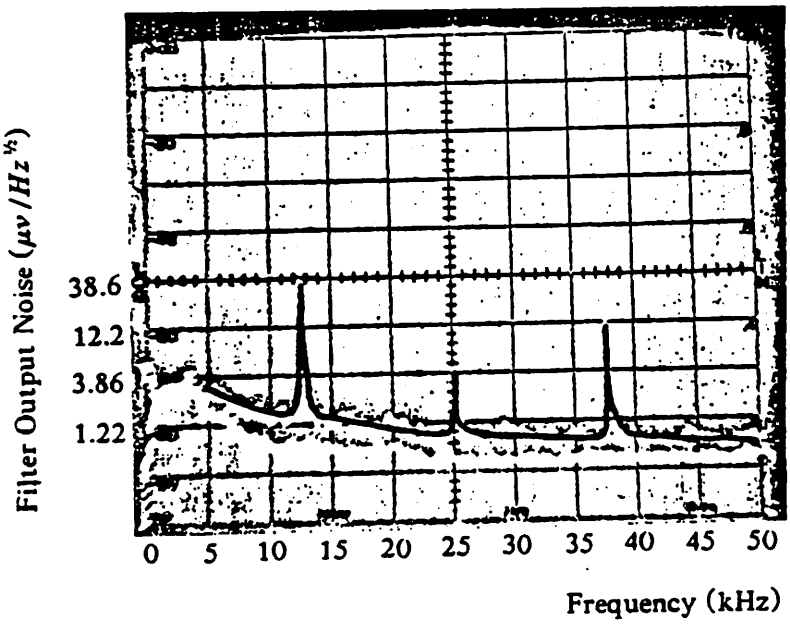


Fig. 6.20 Measured filter output noise spectrum of main filter clock frequency 25 kHz

Active area	1000	mil^2
Supply voltage	±2.5	V
Output swing (RMS) (<1% THD)	3.11	V
Power dissipation	1250	μ W
Clock frequency	256	KHz
Positive PSRR (@ 1 kHz)	50	dB
Negative PSRR (@ 1 kHz)	40	dB
Passband ripple	<±0.125	dB
Stop band rejection (>4.6 kHz)	34	dB
THD (2 V_{pp} output 1kHz)	-67	dB
Dynamic range (C weighted)	87	dB

Table 6.4 Filter measured performance at room temperature

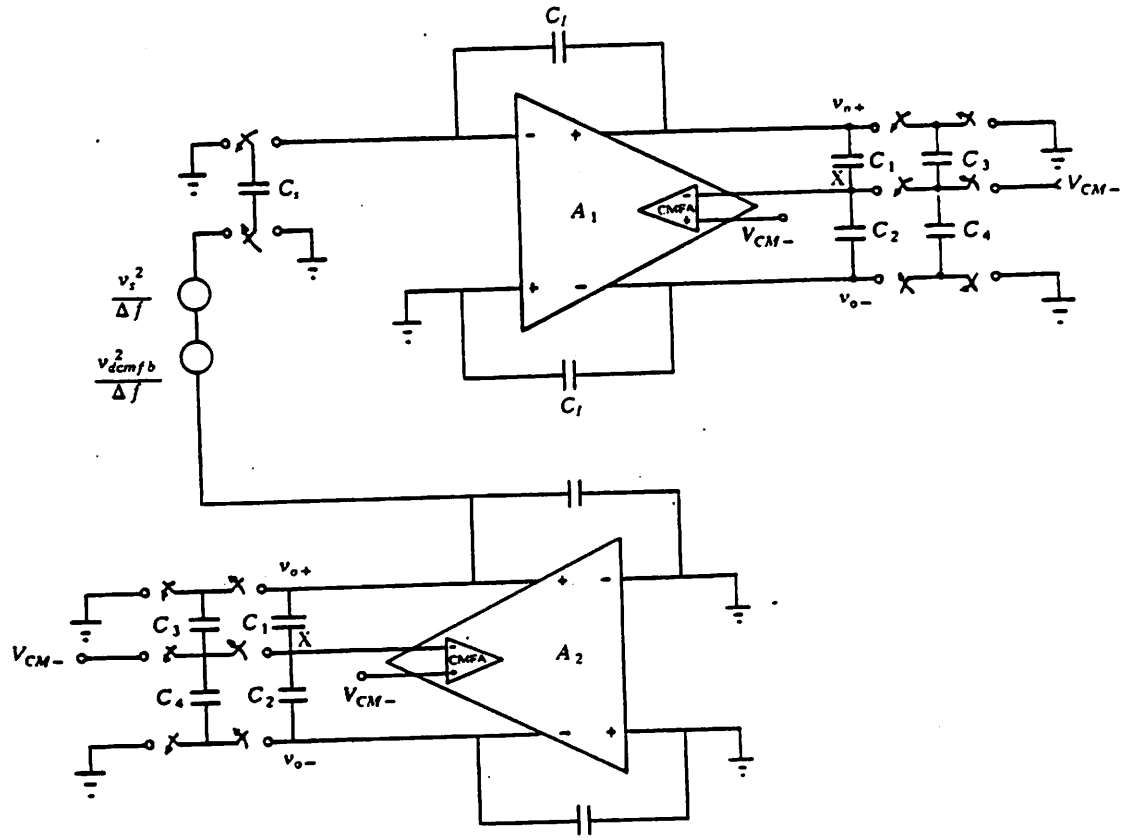
created by these two features, especially the op-amp area with additional chopper circuits. In this section, the noise contribution from the DCMFB and the nonideality of the chopper clocks will be discussed. Following these, area advantage of the chopper stabilization technique in scaled technology will be discussed.

6.5.1. Noise Contribution by the DCMFB Circuits

6.5.1.1. KT/C Noise

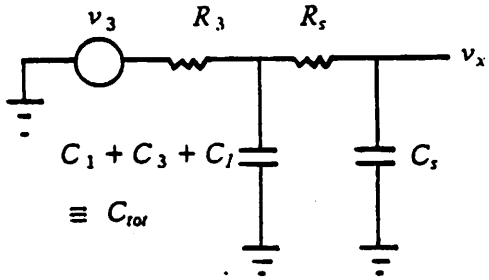
Fig. 6.21a shows the equivalent input KT/C noise power density voltage sources, $\frac{v^2_{dcmfb}}{\Delta f}$ and $\frac{v^2_s}{\Delta f}$, of integrator A with DCMFB op-amp. The former, $\frac{v^2_{dcmfb}}{\Delta f}$, is the equivalent input KT/C noise power density of integrator A due to DCMFB circuit of the previous stage. The latter $\frac{v^2_s}{\Delta f}$ is the equivalent input KT/C noise power density of integrator A which is given by $\frac{2KT}{f_s C_s}$ (Equation 4.6). For the time being, only the contribution of DCMFB KT/C noise from the previous stage is considered. The equivalent circuit to calculate the KT/C noise due to DCMFB circuit of the previous stage is shown in Fig. 6.21b where C_1 is the DCMFB integrating capacitor, C_3 is the DCMFB sampling capacitor, C_s is input sampling capacitor of the integrator A, $R_3 = \frac{1}{f_s C_3}$ is the equivalent resistance associated with the DCMFB switched capacitor C_3, $\frac{v^2_3}{\Delta f} = 2KTR_3$ is the noise power density due to R_3 and $R_s = \frac{1}{f_s C_s}$ is the equivalent resistance associated with the switched capacitor C_s of integrator A. Note that the noise for frequency range from $-\infty$ to $+\infty$ due to R_s is not considered at this moment. The transfer function from the noise voltage source v_3 to the voltage across capacitor C_s (Fig. 6.21b), v_x, is given by

$$H(s) = \frac{v_x}{v_3} \tag{6.8a}$$

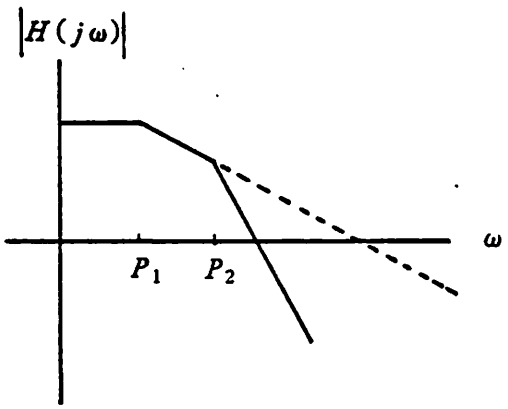


(a)

Fig. 6.21 (a) $\frac{v^2_{dcmfb}}{\Delta f}$ and $\frac{v^2_s}{\Delta f}$ are the equivalent input KT/C noise sources of integrator A composed of DCMFB op-amp.



(b)



(c)

Fig.6.21 (b) Equivalent circuit to calculate the KT/C noise due to DCMFB circuit of the previous stage
(c) Transfer function of equivalent circuit (b)

$$= \frac{1}{s^2 R_3 C_{tot} R_s C_s + s(R_3 C_{tot} + R_3 C_s + R_s C_s) + 1} \tag{6.8b}$$

where C_{tot} is the sum of capacitors C_1, C_3 and C_f. It can be shown that the system function H(s) has two negative real poles P_1 and P_2 (Fig. 6.21c).

For simplicity, suppose two sampling capacitors C_s and C_3 are of the same value, i.e. $R_3 = R_s$. The contributed noise is larger than the actual value if only the smaller pole P_1 is considered (Fig. 6.21c). $\frac{C_s}{C_{tot}} \ll 1$ for practical cases, so the smaller pole P_1 of the two poles is given by

$$P_1 = \frac{1}{R_3 C_{tot}}(1 - \frac{C_s}{C_{tot}}) \tag{6.9}$$

$$\approx \frac{1}{R_3 C_{tot}} \tag{6.9b}$$

The equivalent input KT/C noise power density contributed from DCMFB of the previous stage is therefore given by:

$$\frac{v^2_{dcmfb}}{\Delta f} = 2KTR_3 \frac{1}{R_3 C_{tot}} \frac{1}{f_s} = \frac{2KT}{f_s C_{tot}} \tag{6.10}$$

Combining (4.6) with (6.10), the total equivalent input KT/C noise power density v^2_{KTC} of integrator A is given by:

$$\frac{v^2_{KTC}}{\Delta f} = \frac{v^2_s}{\Delta f} + \frac{v^2_{dcmfb}}{\Delta f} \approx \frac{2KT}{f_s}(\frac{1}{C_s} + \frac{1}{C_{tot}}) \tag{6.11}$$

$$\equiv \frac{2KT}{f_s C_{eq}} \quad \text{for } -\infty < f < +\infty$$

where C_{eq} is the equivalent sampling capacitor of the integrator A (Fig. 6.21) to generate the amount of the total equivalent input KT/C noise power density. For most of the application, sampling capacitor C_s is much smaller than sum of the capacitors C_s, C_1 and C_f, i.e. $C_s \ll C_{tot}$. Thus, KT/C noise contribution due to DCMFB circuit is negligible from Equation 6.11.

6.5.1.2. Clocking Noise

A spike is generated each time when DCMFB integrating capacitors C_1 and C_2 are refreshed because of the charge redistribution of the DCMFB integrating capacitors (C_1/C_2) and DCMFB switched capacitors (C_3/C_4). This spike is a differential mode, however it mainly does not affect the passband signal. This is because spike magnitude is small practically (Equation 6.3) and appears only for a short time interval. The op-amp charges/discharges the DCMFB capacitors and restores the output voltages to their original values quickly. In a sample and hold system, only the quiescent state matters.

If the DCMFB circuit uses the same clock frequency as that of the chopper frequency which is half of the main filter clock frequency, then the spike spectrum will fall at the first harmonic of the chopper frequency. In this case, the DCMFB clocking noise (spike) sampled by the next stage will never be aliased into the passband. Instead, the DCMFB spike will be moved to the odd harmonics of the chopper frequency as those of the 1/f noise harmonics (Fig. 6.22). This method is not employed in this design because the concern of layout area.

6.5.2. Noise Contribution by the Chopper

The chopper stabilization technique mentioned earlier not only removes the colored noise at low frequency but also the DC noise. Therefore, any DC noise introduced by the the circuits between the input chopper and output chopper can be removed by the chopper operation. Besides, if any noise due to chopper appears as common mode voltages, they can cancel each other to the first order by the fully differential configuration. Possible noises generated by chopper process are charge injection, clock feedthrough, 1/f noise and thermal noise.

6.5.2.1. Chopper Noise

Charge injection noise due to the chopper at the summing node is thus negligible, because the summing nodes are at the virtual ground (same voltage). The fully differential configuration,

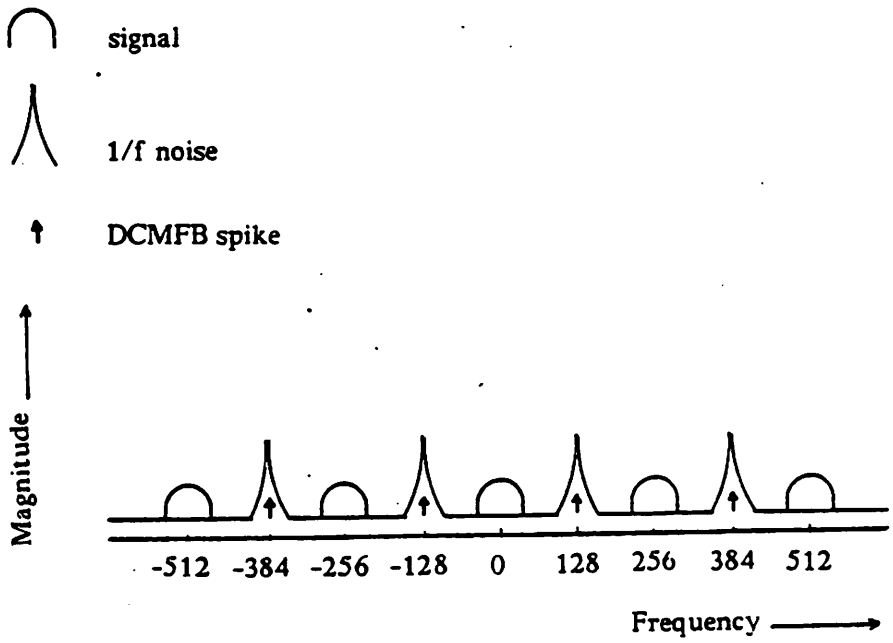


Fig. 6.22 Spectrum of the op-amp sampled by the main filter clock DCMFB is clocked at the same frequency of the chopper clock in order to remove the DCMFB spike

again, to the first order not only cancels the clock feedthrough but also the charge injection noise. The choppers change state during the idle time of the sampling clock ϕ_S and the transferring clock ϕ_T of the main filter (Fig. 6.11). 1/f noise contributed by chopper switches is only during a short transient time and negligible, because there is no current through the chopper switches at steady state.

Because the turn-on resistance of the chopper switch is small (only several $K\Omega$'s), the thermal noise caused by the switch is therefore small. Thermal noise caused by the chopper actually can be treated as part of the op-amp input thermal noise. Thus the chopper thermal noise is also seen as KT/C noise and is negligible for an integrator with finite op-amp output impedance if the equivalent thermal noise resistance due to the chopper and the op-amp is smaller than the inverse

of the op-amp overall transconductance (4.21b).

6.5.2.2. Nonideality of the Chopper Clocks

The ideal chopper clock is a periodic pulse with symmetric amplitude, 50% duty cycle and zero rising/falling time as shown in Fig. 6.23. The chopper frequency is chosen at half of the filter clock frequency as mentioned earlier in order to minimize the alias of the 1/f noise sampled by the next stage (Chapter 5). Therefore, it requires that chopper clock does not have DC component and other even harmonics of the chopper frequency.

Some examples of nonideal chopper clock which will be discussed in this section are (1) unsymmetric amplitude with duty cycle 50% as shown in Fig. 6.24, (2) symmetric amplitude but with duty cycle ≠50% as shown in Fig. 6.25, and (3) finite rising/falling time (Fig. 6.26).

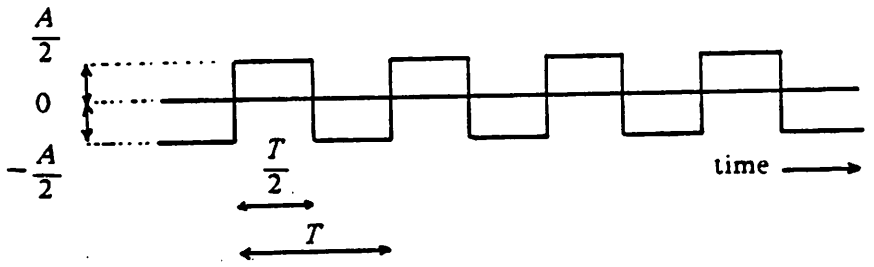
6.5.2.2.1. Case 1

There is a DC component (at $\omega = 0$) if the amplitude of the clock of duty cycle 50% is not symmetric. This will result in a 1/f noise component at the passband. The passband 1/f noise power density component due to the DC of the chopper clock is weighted by

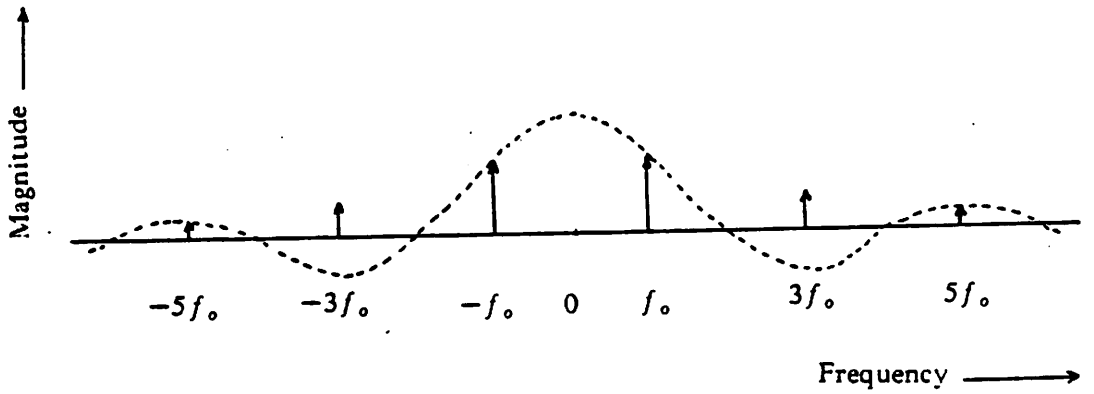
$$S_{DC} = (\delta\gamma A - \frac{A}{2})^2 \tag{6.12}$$

where A is the amplitude of the pulse (Fig. 6.24), δ is the duty cycle and γ is the factor to count for the unsymmetric magnitude. The magnitude of 1/f noise at DC can be neglected so long as the DC magnitude of the clock is small. For example, if $\delta = 50\%$ and $\gamma = 0.9$ (10% off the center), then the aliased 1/f noise in passband S_{DC} is only 1% of the original 1/f noise. Note that besides the passband 1/f noise, other 1/f noise components are at odd harmonics of the chopper frequency, because the chopper duty cycle is still 50%.

In this design, chopper operation is realized by steering the signal one way or the other using the switches which are controlled at gates of MOS devices by the chopper clocks. Therefore, the unsymmetric amplitude as shown in Fig. 6.24 is very unlikely.



(a)



(b)

Fig. 6.23 Ideal chopper clock with duty cycle 50% and symmetric amplitude
(a) in time domain
(b) in frequency domain

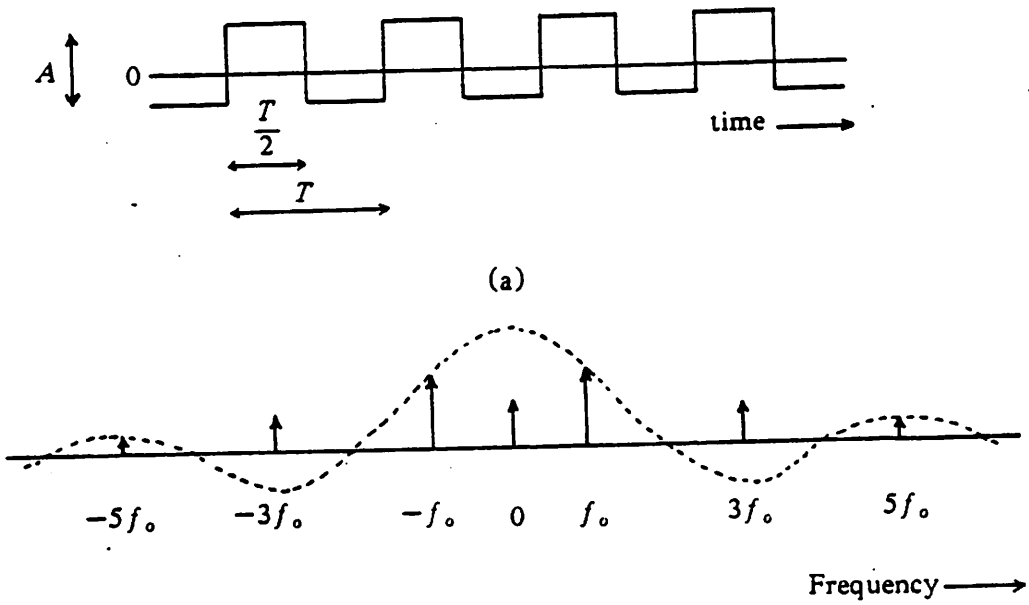
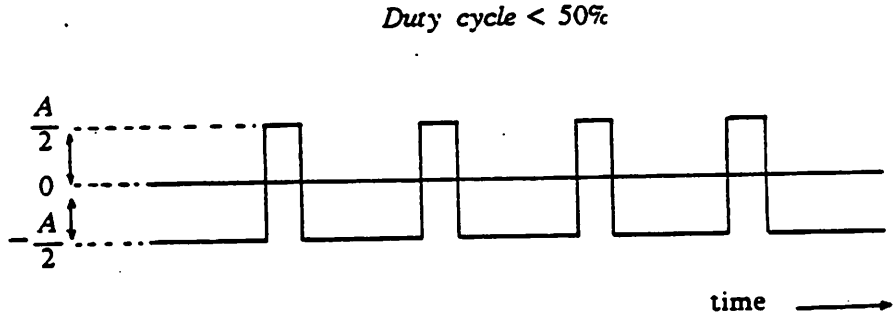
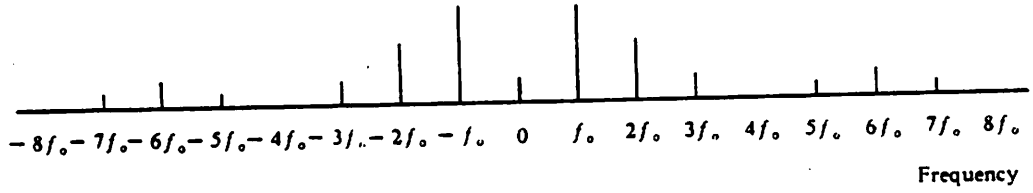


Fig. 6.24 Chopper clock of duty cycle 50% but with unsymmetric amplitude
(a) in time domain
(b) in frequency domain

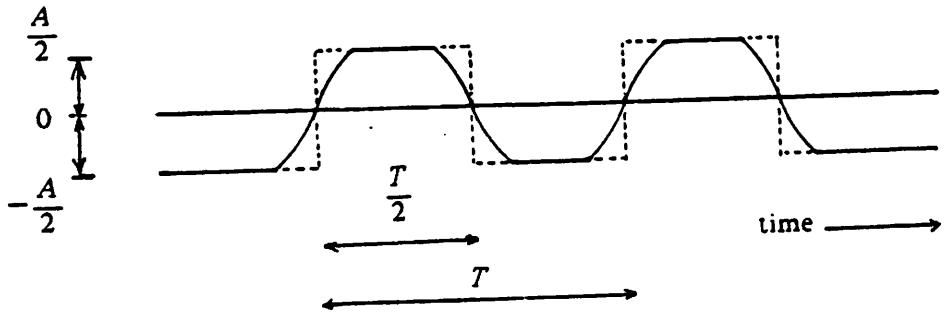


(a)

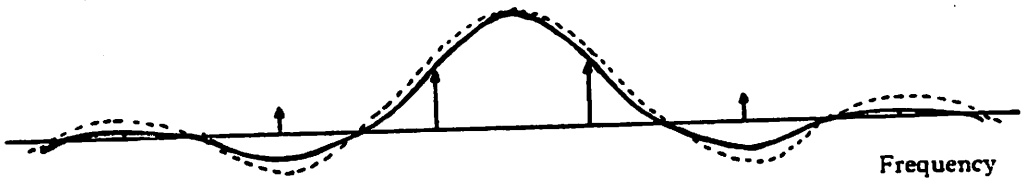


(b)

Fig. 6.25 Chopper clocks with symmetric amplitude but duty cycle ≠ 50%
(a) in time domain
(b) in frequency domain



(a)



---- Spectrum envelope of ideal chopper clock as shown in Fig. 6.23

——— Spectrum envelope of non-ideal chopper clock as shown in Fig. 6.26(a)

↑ Spectrum of non-ideal chopper clock as shown in Fig. 6.26(a)

(b)

Fig. 6.26 Clocks with symmetric rising and falling edges
(a) in time domain
(b) in frequency domain

6.5.2.2.2. Case 2

If the duty cycle of the chopper is not 50%, the chopper clock in the frequency domain has not only the DC component (6.12) but also has other even harmonic components of the chopper clock (Fig. 6.25). Thus the 1/f noise appears at the passband due to DC component of the clock is similar to case 1. And the 1/f noise at even harmonics of the chopper frequency will be aliased to the passband when sampled by the next stage.

6.5.2.2.3. Case 3

For simplicity, waveforms of the rising and falling edges are assumed to be symmetric as shown is Fig. 6.26. Then chopper clocks show smaller side lobes and, in the ultimate case, give more number of zeros in addition to the sampling zero crossings in frequency domain [46]. This results in a less number of noise harmonics and smaller 1/f noise components. However, there are still even harmonics which are undesirable if the duty cycle of the clock is not 50%.

The RC time constant of the chopper switches is only several nano seconds which is negligible compared with the chopper clock period of 7.8 micro-seconds (128 kHz). Thus the effect due to finite rising and falling time in this design is negligible.

If the signal is modulated by a sinusoidal wave at a frequency of half of the main filter clock, then, only one 1/f noise component appears at the chopper frequency. However the realization of the chopper is not so simple and besides, the chip needs an extra generator and additional clock bus lines.

6.5.3. Chopper Stabilized Op-amp in Scaled Technology

There are two major issues in the using of the chopper stabilization techniques. One is the degree of effectiveness of the filter noise reduced by chopper. The second issue is whether the area consumed by the chopper is justified.

6.5.3.1. Noise Improvement

As mentioned earlier, the reduction of the op-amp noise with chopper clock of 128 kHz turned on is 20 dB at 1kHz. However the improvement of the filter noise with the chopper of the same frequency is only 4 dB at 1kHz. The discrepancy is because the op-amp noise improvement due to chopper is the reduction of the noise density of a single stage op-amp. In the case of sampled data system, as the switched capacitor filter, the alias of the white noise due to the op-amp thermal noise and KT/C noise needs to be considered.

The calculated KT/C noise density of this particular filter caused by the 5 integrators is already 1.66 $\mu v/\sqrt{Hz}$ which accounts for 10 dB of the filter white noise that is not affected by the chopper operation.

6.5.3.2. Area Consideration

The goal of employing the chopper stabilization technique is to remove 1/f noise. Input device sizes can also be reduced because it is not necessary to use a large devices so long as the transconductance is large enough. The concern of using the chopper stabilization technique is the cost of the area for chopper switches and chopper clocks. The conventional method to minimize the 1/f noise is by increasing the area of input device sizes. For the same amount of reduction of the 1/f noise, a comparison in area of these two methods based on the chip prototype realized and the design rule used will be discussed in this section.

The op-amp noise reduction at 1 kHz by chopper is 20 dB or 10-fold improvement. If the chopper is not used, the area of the input device of the op-amp needs to be increased 10 times over the current device area used in order to achieve the same noise performance as the chopper stabilized op-amp does. The op-amp area with the chopper is only 80 mil^2 which is composed of the chopper area ($16mil^2$) and non-chopper area ($64mil^2$) including the chopper bus to be precise. Using the same design rules, the non-chopped op-amp requires an area of at least 126 mil^2 in order to achieve the same 20 dB 1/f noise improvement if the conventional method is used. For a fixed

dynamic range, because the chopper stabilized op-amp costs less area in scaled technologies, one can project that the chip prototype realized takes smaller area than the area of the non-chopped counterparts for channel length shorter than 4μ. The chip area of other analog circuits is efficient in general if a chopper stabilized op-amp is the major building block of the system.

Nevertheless, the 1/f noise has to be removed by the chopper stabilization such that the filter dynamic range is not process dependent.

6.6. Summary

Applying all the circuit techniques for scaling analog circuits discussed in Chapter 5, a scalable SCF with significantly reduced area per pole relative to commercially available devices of similar performance was built.

Noise contributions and the extra area introduced by the distinct circuit approaches, namely DCMFB, chopper stabilization are discussed and justified. These circuit techniques are more attractive particularly in scaled technology.

CHAPTER 7

SCALING LIMITS

With the advancement of lithography and pattern technology, scaling of MOS devices of the conventional structure to the micron or even submicron range is possible without showing severe short channel effects [20]. However, scaling solely for performance enhancement at IC levels becomes more difficult when other parasitic elements, such as interconnection line capacitance, line resistance [22] and contact resistance are considered because these parasitic elements are not scaled [47,20 , 16]. This limits the overall performance of the circuit even if the scaled transistor has better performance. Moreover, it is doubtful that the MOS device can be scaled beyond a certain limit e.g. 0.1 μm [20] and still gain any circuit performance advantages because of the fundamental scaling limit of the device physics.

Analog circuits suffer the MOS short channel effects more severely than digital circuits because analog circuits operate on continuous signals rather than binary digitized signals. Several short channel effects of the MOS devices such as DIBL (Drain Induced Barrier Lowering) and impact ionization mentioned in Chapter 4 not only degrade the circuit performance (e.g. gain and signal to noise ratio), but also, in the extreme case, eliminate the short channel MOS devices for analog circuit applications.

The impact of MOS device scaling on analog circuit performance and solutions associated with the (moderate) scaling were discussed in Chapter 4 and chapter 5, respectively. In this chapter, fundamental scaling limits from the device physics standpoint and practical considerations will be discussed first. Following these discussions, ultimate limiting factors of the short channel MOS devices for circuit designs will be discussed. It is neither the aim nor the scope of this project to suggest solutions to device scaling limits which could possibly be solved by new processes or device structure.

7.1. Fundamental Scaling Limits

Scaling of the MOS devices requires the maintenance of long channel behavior in order to achieve better circuit performance. However, the basic device physics of the MOS structure determines the fundamental limits of the device scaling. Before reaching the ultimate device limits, some of the device mechanisms due to short channel effects practically limit the MOS transistors for circuit applications.

Aside from the device limits, there are two other factors that need to be considered when the device is scaled down. One is the interconnections and the other is the thermal dissipation. Both limits affect the chip reliability, architecture and packaging.

7.1.1. Device Scaling Limits

7.1.1.1. Ultimate Limits

The ultimate dimension limit is 200 Å which is the minimum required for the electron interface of two materials (100 Å each side) such that no breakdown occurs. 100 Å is also the dopant spacing at the degeneracy limit. Statistical considerations of the device geometry and electron scattering prohibit the device to be scaled down beyond this limit [20].

7.1.1.2. Practical Limits

A useful MOS device should operate away from the undesirable short channel effects such as substrate current, threshold voltage shift and drain avalanche break down, etc., otherwise, the transistor performance will be degraded. These undesirable effects are mainly associated with the channel length, operating voltage and oxide thickness.

7.1.1.2.1. Channel Length Limitation

The feasible channel length for scaled MOS devices as a function of power supply was derived by Masuda [16] based on the constraints of the subthreshold leakage current, threshold voltage shift due to short channel length, drain avalanche breakdown, minimum required operating voltage and physical limits. Masuda showed that the minimum channel length is determined by the drain breakdown effect when power supply is large. Other than this large power supply range, the problem of the threshold voltage shift dominates in the determination of the minimum channel length.

Another consideration in the scaling limit associated with the channel length is the series source/drain resistance. Beyond some point of scaling the channel length, the series source/drain resistance is eventually comparable to the channel resistance when the MOS transistor is biased in the operating mode. Then the scaled MOS device can no longer preserve the long channel characteristics. The transconductance $g_m = \frac{g_{mo}}{1 + g_{mo}R_s}$ is degraded by the feedback source resistance.

7.1.1.2.2. Gate Oxide Thickness Limitation

For thin oxide devices, the oxide breakdown is an important limiting factor. Gate tunneling current occurs before the oxide breakdown is reached [16]. With the constraints of the subthreshold current and the tunneling current, Masuda showed the possible gate oxide thickness as a function of the power supply. The minimum oxide thickness derived is 94 Å [16].

7.1.2. Interconnection Scaling Limits

7.1.2.1. Electromigration

Under scaling, the current per unit width of a transistor increases. Coupled with the decrease of the conductor width the conductor current density becomes larger. For a particular kind of con-

ducting material, there is a certain limit of current density, e.g. $10^5 A/cm^2$ for aluminum, beyond which the electromigration occurs and results in the breaking of the conductors. Obviously, this can be prevented by just widening the metal with less area density.

7.1.2.2. Edge Capacitance

There are two parasitic capacitance components associated with the conductors. One is the parallel capacitance between the conductors and the substrate. The other one is the edge capacitance between the conductor edge and the substrate. The parallel component is proportional to the conductor width. The edge capacitance, however, is almost constant. After scaling, the former component becomes smaller. These two components are comparable to each other when the conductor width shrinks to around 1 μ [20]. The parallel capacitance component can of course be minimized by the increase of the field oxide thickness.

An additional capacitance is the line-to-line capacitance when the conductor spacing is small enough. This can be solved by routing a ground conductor line next to signal bus lines at the expense of extra area.

7.1.2.3. Wire Resistance

Fine patterns usually require thin films. This increases sheet resistance because the resistance is inversely proportionally to the line thickness. Together with the capacitance mentioned above, the increasing of the line resistance after scaling results in larger RC time constants which ultimately limit circuit speed and dynamic range.

Notice that in general the same circuit after scaling in the same technology can take advantage of the smaller RC time constants to improve circuit performance (Table 3.2) until the conductor is scaled beyond a limit such that the line edge capacitance dominates (<1 μ in aluminum case) or sheet resistance increases because thin film is used.

7.1.3. Thermal Considerations

The maximum power that a chip can dissipate is determined by either the number of the transistors or the clock rate of the chip. The ultimate limit of the power that silicon can tolerate is 10 W that is the band gap of silicon [20] and it is the fundamental limit of the material.

Standby power can be minimized by the use of CMOS devices in conjunction with the power off circuit in digital circuits. In analog applications, class AB op-amps consume small standby power.

7.2. Circuit Ultimate Limiting Factors

All the limiting factors in circuit applications are mainly caused by the high transverse electric field when the channel length is reduced. The substrate current, threshold voltage shift due to hot electrons injected into gate oxide (for NMOS devices) and drain avalanche breakdown are generally referred to as hot electron effects. Because the electron has longer mean free path to acquire more energy from the channel transverse field, the hot electron effect starts earlier in the scaled NMOS devices. The following discussion, therefore, is focused on the NMOS devices. The degradation of MOS devices due to impact ionization shows at least three symptoms [48]: (1) shift in the threshold ΔV_t [17] (2) shift in the subthreshold swing $\left(\frac{\partial V_g}{\partial \log I_d} \right) \Delta S$ [48], and (3) reduction in the transconductance in the linear or saturation regions Δg_m [50]. It is interesting to notice that these three symptoms, ΔV_t, ΔS and Δg_m are linearly related to one another [48]. One more symptom is the reduction of the device output resistance due to the substrate current.

There are other short channel hot electron effects such as light emission, hot electron emission and interface damage [48] which are far beyond the usable operating range for circuit application.

Beside problems of high field effects, one more limiting factor is the subthreshold current which occurs when the device is biased in the weak inversion region.

7.2.1. Subthreshold Leakage Currents

When the gate voltage is just below the threshold voltage, the device is biased in the weak inversion region, the corresponding current is called the subthreshold current which is a fundamental limit of the MOS circuits. In the subthreshold region, the drain current varies exponentially with the gate voltage. The subthreshold current is a criterion [51] for determining a long channel or a short channel device. In a long channel case, the subthreshold current is not drain dependent for $V_D > 3kT/q$ and $V_G < V_t$. For a short channel case, however, the subthreshold current is dependent of the drain voltage under the same conditions. A common figure of merit to monitor the subthreshold current is the subthreshold swing S which is defined as the gate voltage below threshold voltage needed to reduce the drain current by one decade (Equation (3.13a)), i.e.

$$S \equiv \ln 10 \cdot \frac{d V_g}{d \ln I_D} \tag{7.1a}$$

$$\simeq \frac{KT}{q} \ln 10 \cdot \left(1 + \frac{C_D}{C_{ox}} \right) \tag{7.1b}$$

where C_D is the surface depletion layer capacitance per unit area and C_{ox} is the gate oxide capacitance per unit area.

The subthreshold current determines how fast the MOS device is turned on and off, it is therefore important when MOS transistor is used as a switch. If the threshold voltage is designed too low, then when the gate voltage is turned off, there is still a significant amount of current that continues to flow in the channel. The impact of the subthreshold current on charge-storage circuits is the leakage of stored charge on a capacitor. In a switched capacitor circuit application, it means the signal stored will disappear.

Because the subthreshold swing S is not always scalable (Table 3.1), the weak inversion region is particularly important when the power supply is reduced in scaled technologies. In fact, the minimum power supply required for a MOS device used as a switch is determined by the subthreshold current which in turn is determined by the subthreshold swing S (Equation 7.1). The ideal subthreshold voltage swing at room temperature is $(S)_{min} = \ln(10) \cdot \frac{KT}{q} = 57\ mV$. If the

power supply $V_{ps} = V_{dd} - V_{ss}$ is, at least, n times larger than the threshold voltage V_t, and the current at V_{ss} is 10^m times smaller than the current at threshold voltage V_t, then the required power supply for the MOS as a switch is

$$V_{ps} \geq n\, V_t = n{\cdot}m\, S$$

For $n = 5$ and $m = 5$ [16] of a MOS switch at room temperature, the minimum power supply (V_{ps}) required is 1.43 V. In the analog application, the minimum power supply is associated with the voltage necessary to bias the transistors in the saturation region; and it is usually larger than 1.43 V.

7.2.2. Threshold Voltage Shift

Noticeable threshold voltage shift is observed in small geometry. In short channel devices, the threshold voltage decreases because of the source-drain charging effect. In narrow channel devices, the threshold voltage increases due to related depletion region spreading laterally in the substrate along the channel width. For a particular process, the channel width in a circuit should be large enough to avoid the narrow width effect in order not to degrade the transconductance.

Beside the voltage shift due to the shrinkage of the horizontal dimensions, for very short channel devices, DIBL (Drain Induced Barrier Lowering) also causes the threshold voltage decrease as the drain voltage increases [16]. Wong et al. derived an expression including the DIBL effect [12]:

$$r_o = (g_{ds})^{-1} = \left(\frac{\partial I_{Dsat}}{\partial V_{DS}}\right)^{-1} \simeq \frac{L^3}{g_m \eta} \tag{7.2}$$

where g_{ds} is the device output conductance, I_{Dsat} is the saturation drain current with the effect of threshold voltage shift considered, g_m is the device transconductance and η is a technology dependent parameter.

From (7.2), the device output resistance is proportional to the cubic power of the channel length L. This strong dependence of output resistance on the channel length is not tolerable for analog circuit application. Both DIBL and hot electron substrate current (next section) tend to

lower the output resistance and hence decrease the available gain of an amplifier.

One more effect which results in the positive shift of the threshold is the hot electrons injected into the gate oxide. This will be discussed separately later.

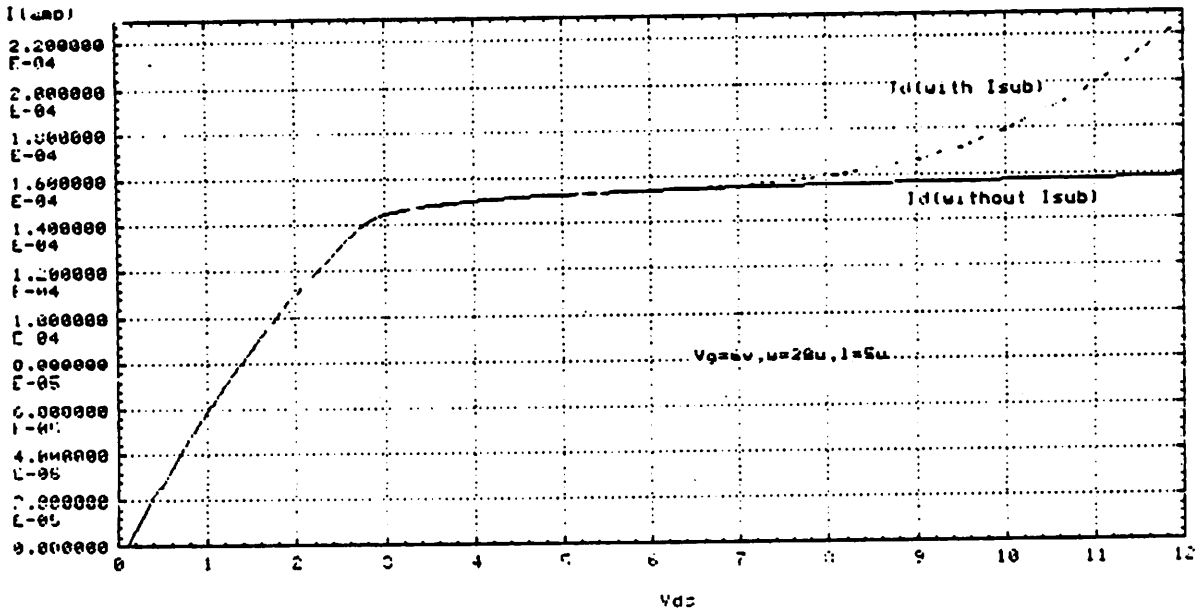
7.2.3. Substrate Currents

When the drain voltage is sufficiently high, substrate current will be generated because weak avalanche occurs within the drain punch through region [51]. The substrate current is in fact one of the impact ionization phenomena. The substrate current first increases, then decreases with the gate voltage at a given drain voltage. The rise of the current is due to the increase of the drain current with the gate voltage and eventually falls at larger gate voltages due to the decrease of the channel field at the drain punch through region. The peak occurs when these two factors are balanced. As the source to drain separation is reduced, the impact ionization occurs at a lower drain voltage as does the substrate current.

Using a unified equation for hot electron currents [52, 53], an analytical breakdown model [54, 55] and derived small signal parameters [26], the effect of the substrate current on drain current and output conductance are demonstrated in Fig. 7.1 [26]. The finite substrate resistance can be represented by a shunt resistance between the drain and the substrate as shown in Fig. 4.2. As the substrate current flows across this shunt resistance, the reverse junction potential between the source and the substrate decreases which causes the drain current to be increased and output resistance to be decreased. A circuit solution to this kind of performance degradation was discussed in Chapter 5.

7.2.4. Hot Electrons Injected to Gate Oxide

Another limiting factor related to high field impact ionization is the gate oxide charging [17] when a stress voltage is applied to the gate and drain. Electrons that overcome the Si–SiO_2 energy barrier and inject to the gate oxide are from the inversion layer or avalanche plasma formed near



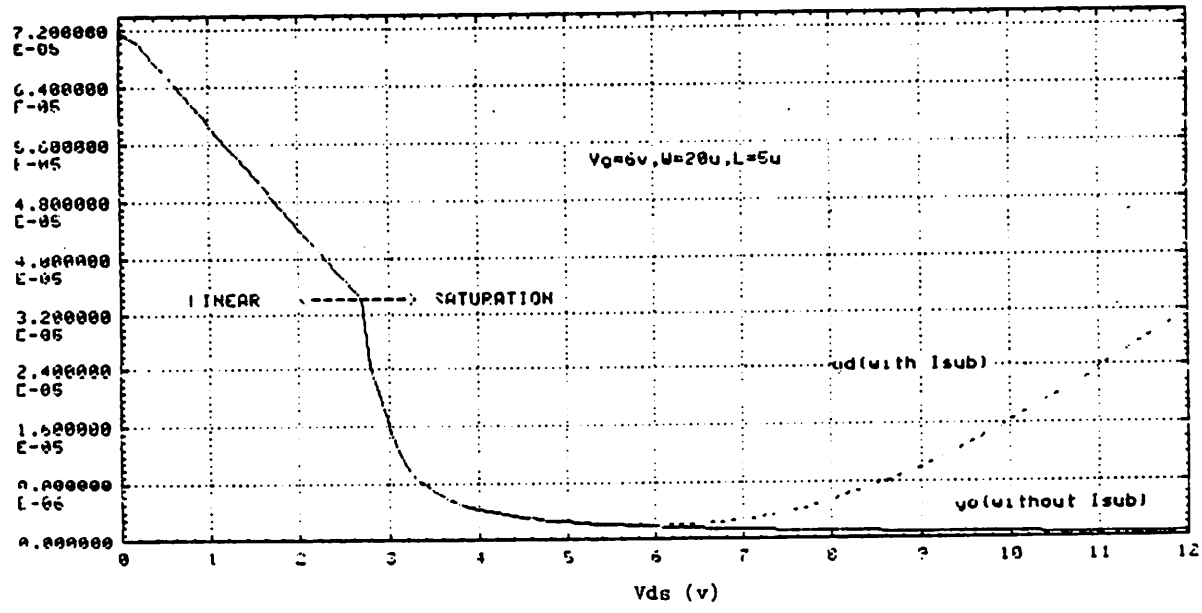


Fig. 7.1 Using unified model, the effect of substrate current on
(a) drain current (b) output transconductance [24]

the drain or the bulk. The most prominent effects of the injection of electrons into the gate oxide are the reduction of the transconductance due to reduced channel mobility, increase of substrate current due to the increase of the interface trap density, and the increase of the threshold voltage.

The degradation of the oxide charging is cumulative, therefore oxide charging effect limits the maximum voltage levels that can be applied or the minimum channel length that can be used for a given specific device lifetime.

7.2.5. Drain Breakdown

The drain breakdown of a MOS device is determined by avalanche breakdown of the PN drain junction caused by the high field between the gate and drain, and the lateral bipolar effect [16].

A MOS transistor with a shorter channel length has a higher channel transverse electric field that has its maximum at the drain end of the channel, hence the decrease of the drain breakdown voltage of a scaled MOS device is imminent [49]. For analog applications, the drain breakdown region should be completely avoided.

Fig. 7.2 [49] shows the maximum allowed drain voltage as a function of channel length with scaling limiting factors, i.e. drain breakdown, parasitic bipolar, oxide charging and punch through. As the channel length is reduced, different mechanisms limit the maximum drain voltage; and hence determine the power supply that can be used and, in turn decide the maximum output swing available in the circuit. These limiting factors are for a particular set of device parameters. For example, for a channel length less than 3 μ and power supply at 5 V, the oxide charging effect is more important than the parasitic bipolar effect. With other choices of parameters, the relative importance of the mechanisms changes. Nevertheless, the maximum voltage or equivalently, the minimum channel length which makes the device immune from these limiting factors imposes the power supply upper bound of the circuit. And the region beyond this upper bound must be avoided in the application of analog circuits because transconductance degradation, threshold shift, snapback

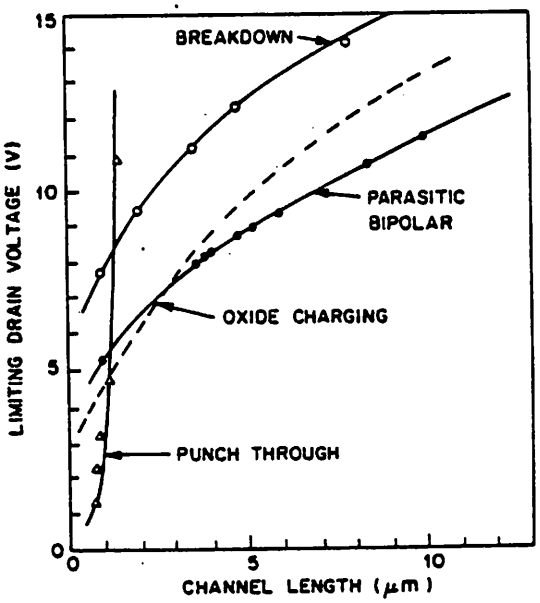


Fig. 7.2 Device limiting factors to the circuit performance [52].

current, etc. will occur.

7.3. Area Projection of the SCF

In this section, an area projection of the filter realized for a fixed dynamic range under CE and CV scaling laws will be discussed. Fig. 7.3 shows the area projection versus the channel length.

Under CE scaling, the projected filter area decreases for channel length scaled from 3 μ to 1.7 μ, but increases for channel length shorter than 1.7 μ. This is because the interconnection and the op-amp area dominate the total filter area for channel length larger than 1.7 μ. For channel length less than 1.7 μ, however, the capacitor area dominates, and the capacitor area must be increased in order to keep the dynamic range unchanged because the power supply is scaled.

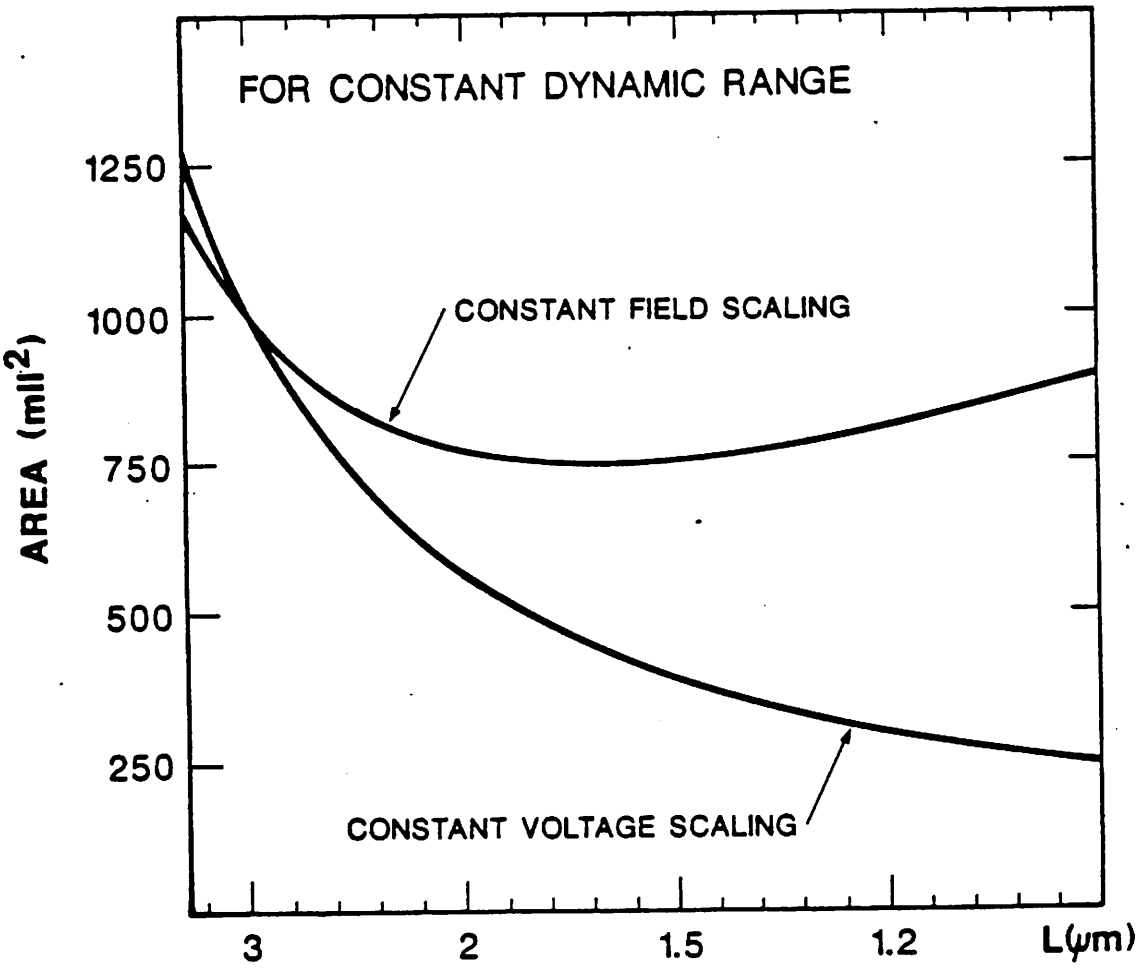


Fig. 7.3 Area projection of the 5th. order SCF in further scaled technology.

Overall area of the chip prototype decreases if CV scaling law is applied. The total area decreases monotonically even when capacitor area dominates. This is because the voltage is not scaled. So, it is not necessary to increase the capacitance in order to maintain constant dynamic range. Furthermore, because the capacitance per unit area increases with scaling, the required area to produce a fixed value of capacitance becomes smaller with scaling.

With a feature size of 1.5 μ, and still a 5 V power supply voltage, the expected area of the experimental filter is 82 mil^2 per pole instead of 200 mil^2.

7.4. Summary

The ultimate limit to scaling a MOS device is the minimum dimension required for the interface of different materials. Parasitic elements of interconnection and thermal power dissipation of the chip are basic constraints. Reduction of the channel length and the oxide thickness of a useful scaled MOS transistor is practically limited by the short channel effects, namely threshold voltage shift, hot electrons injected into gate oxide, and drain break down, which result in the variation of threshold voltage, the increase of subthreshold voltage swing, the decrease of transconductance and the output resistance. Another parameter that does not scale linearly (Equation 3.13b) is subthreshold current which shows the cut off characteristics of a MOS device. Table 7.1 shows the MOS scaling for LSI and VLSI [47].

A projection of the prototype switched capacitor filter realized in even more dramatically scaled technologies indicates that the area can still be further scaled.

Parameter	LSI (1970-1980)	VLSI 1980-
1. Feature	λ	• Limited by lithography, etching technologies
2. Density	$1/\lambda^2$	• Limited by design, registration, interconnect
3. Channel Length	λ	• Limited by short channel effect, system voltage, reliability requirements
4. Gate Delay	λ	• Intrinsic device delay limited by velocity saturation, gate dielectric, parasitic resistance
	λ	• In circuit delay limited by interconnect resistance, line capacitance, current drive capability
5. Supply Voltage	λ	• Lower system supply voltages mandatory from dielectric breakdown, hot electron injection reliability considerations
6. Power Dissipation	λ^2	• Sub-threshold leakage does not scale. Larger leakage due to barrier lowering, drain modulation
7. Power Delay Product	λ^3	• Speed power product enhancements limited by parasistics, interconnect limitations
8. Gate Oxide	λ	• Gate thickness 100Å required with significantly severe process control, defect constraints
9. Parasitic Elements		• Do not scale, contact resistance increases with area reduction with reliability hazard

Table 7.1 MOS scaling of LSI and VLSI [46]

CHAPTER 8

CONCLUSIONS

The overall objectives of this thesis were to explore new circuit and system techniques that will allow switched capacitor signal processing circuits realized in scaled technologies to more closely approach the theoretical minimum area calculated [29] than in the past.

Several circuit approaches to the implementation of switched capacitor signal processing techniques have been described. As a vehicle to test the feasibility of all these approaches, a scalable switched capacitor filter was realized with significantly reduced area relative to commercially available devices of similar performance. A projection of this result in even more dramatically scaled technologies than the one used indicates that the area can still be further reduced.

8.1. Future Directions

To reduce area, improve bandwidth, and improve matching, future work will be directed to take advantage of scaled technologies.

Applying the circuit approaches proposed for scaling analog circuits, the existing voice band applications of the switched-capacitor techniques, e.g. telecommunication [5,56], speech processing [57,58] and adaptive networks [59], can use scaled technology to achieve small area, low power and high performance. Because of the high device unity gain frequency, circuit approaches discussed are attractive for high performance and high system level integration in radio and video signal processings e.g. a single chip AM/FM receiver. Further development of MOS charge transfer techniques are likely for A/D and D/A conversions, especially in high precision application.

Device modeling of two dimensions or three dimensions for small feature size, particularly for channel length shorter than 1.5μ [20], is still an important task to work on in order to achieve simulation accuracy of analog circuit performance e.g. device output impedance.

8.2. Conclusions

Digital circuits have always been scalable. This work shows that analog circuits are scalable, too. Thus analog processing techniques will continue to provide important approaches for complex system function integrated with digital circuits in scaled technologies.

APPENDIX A

A CMOS Voltage Current Transducer

CCT (Current-to-Current Transducer), VVT (Voltage-to-Voltage Transducer), CVT (Current-to-Voltage Transducer), VCT (Voltage-to-Current Transducer) or their combination can be integrators for the active filters. Integrators with VCT is more applicable for high frequency continuous filters than that with VVT, since the integrating capacitor of VCT integrator is at its forward path instead of at the feedback path as that of VVT integrator [60]. This appendix devotes to a simple novel CMOS VCT.

A simple source coupled pair can be used as a VCT. High frequency continuous filters have been built with this kind of VCT elements [61,62]. The linearity and the dynamic range are inherently limited by the quadratic relation between I and V of the MOS devices. There are several measures to improve the linearity of the MOS VCT e.g. g_m degeneration [63], cancellation of second order error terms of two g_m's [64], quarter wave principle/algebraic cancellation [65,66], etc. All the circuits associated with these methods published so far are complicated circuit-wise and/or critical dimension-wise if linearity is required [65].

The linearity of the CMOS VCT propose is achieved by algebraic cancellation. That is a linear relation between output current i and input voltage υ of a VCT can be obtained by the following algebraic operation:

$$I_1 = (v + v_{m14})^2$$

$$I_2 = (v - v_{m23})^2$$

Then,

$$i = I_1 - I_2 = 4vv_m, \quad \text{if } v_{m14} = v_{m23} \equiv v_m \tag{A.1}$$

Then, current i is a linear function of v if v_m is constant in the region considered.

1. Derivation

The relation of output current i and the input voltage v of Fig. A.1 is derived in the following:

$$i = I_1 - I_2 = k[(v + v_{m14})^2 - (v - v_{m23})^2]$$

$$= k[(v_{m14}^2 - v_{m23}^2) + 2v(v_{m23} + v_{m14})]$$

$$= \frac{4}{(\frac{1}{\sqrt{k_p}} + \frac{1}{\sqrt{k_n}})^2}[V_o - (V_{TN} - V_{TP})]v \quad \text{if } v_{m14} = v_{m23} \text{ or } V_{TP3} = V_{TP3} \equiv V_{TP} \tag{A.2a}$$

$$= 4kv_m v \tag{A.2b}$$

where

$$v_m \equiv V_o - (V_{TN} + V_{TP})$$

$$v_{m14} = V_o - (V_{Tn1} + V_{Tp4})$$

$$v_{m23} = V_o - (V_{Tn2} + V_{Tp3})$$

$$v = v_{i+} - v_{i-}$$

$$\frac{1}{k} = \frac{1}{\frac{1}{\sqrt{k_p}} + \frac{1}{\sqrt{k_n}}}$$

$$k_p = \frac{\mu C_o}{2}(\frac{W}{L})_3 = \frac{\mu C_o}{2}(\frac{W}{L})_4$$

$$k_n = \frac{\mu C_o}{2}(\frac{W}{L})_1 = \frac{\mu C_o}{2}(\frac{W}{L})_2$$

If the voltage shift V_o is implemented as shown in Fig. A.2, then

$$V_o = \Delta V_{N5} + V_{TN5} + \Delta V_{P7} + V_{TP7}$$

and

$$v_m \equiv \Delta V_{N5} + V_{TN5} + \Delta V_{P7} + V_{TP7} - (V_{TN} + V_{TP}) \simeq \Delta V_{N5} + \Delta V_{P7} \tag{A.3}$$

where $V_{TN1} = V_{TN2} \equiv V_{TN}$, $V_{TP3} = V_{TP4} \equiv V_{TP}$

Incorporating (A.3), Equation (A.2b) becomes

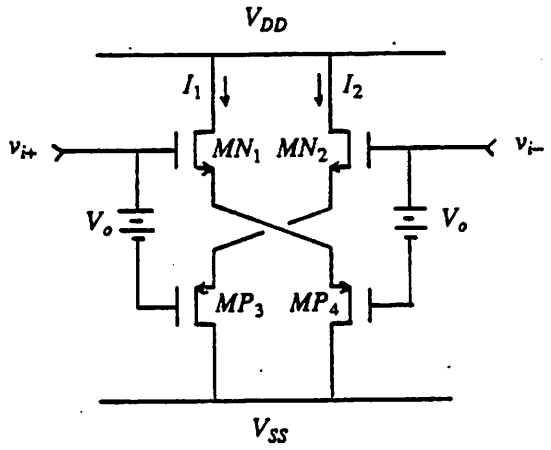


Fig. A.1 A class AB cross coupled pair as a Voltage-to-Current Transducer

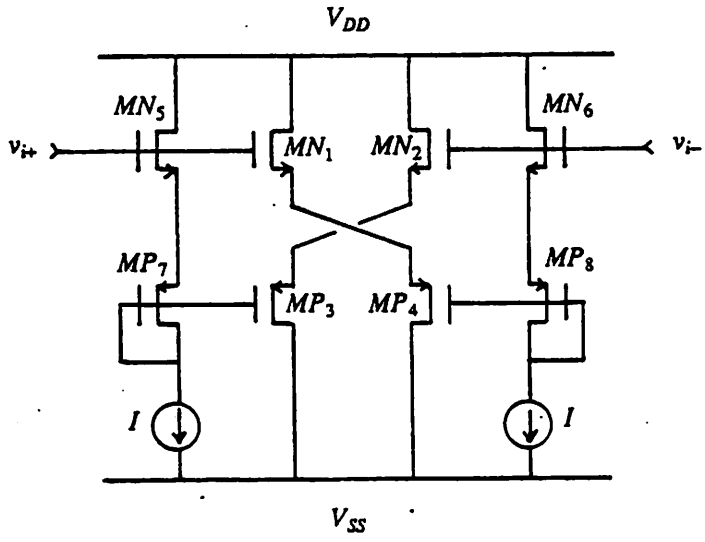


Fig. A.2 Voltage shift V_o implemented by transistors $M_5 - M_8$

$$i \simeq [\frac{4}{(\frac{1}{\sqrt{k_p}} + \frac{1}{\sqrt{k_n}})^2}(\Delta V_{N5} + \Delta V_{P7})]v \equiv g_m v$$

where g_m is the transconductance of Fig. A.2 and is defined as:

$$g_m \equiv \frac{4}{(\frac{1}{\sqrt{k_p}} + \frac{1}{\sqrt{k_n}})^2}(\Delta V_{N5} + \Delta V_{P7}) \tag{A.4}$$

2. Linearity

From (A.2a), the linearity is dependent on the matching of threshold voltages of NMOS transistors, MN_1 and MN_2; and PMOS transistors, MP_3 and MP_4, respectively. In a p-well technology, MN_1 and MN_2 do not have body effect, if the the source is tied to the well. Hence the nonlinearity is solely contributed by the difference of the threshold voltages between MP_3 and MP_4. If device sizes of MN_1/MN_2 and MP_3/MP_4 are so chosen such that $\mu C_o \frac{W}{L}$ of MN_1/MN_2 is equal to that of MP_3/MP_4, then source nodes of MP_3 and MP_4 are both at virtual ground. The threshold voltage of MP_3 is therefore equal to that of MP_4, since V_{sb} is the same.

3. Input Range

Input range is the smaller one of either the sum of the overdrives of the level shifter composed of MN_5 and MP_7 or maximum negative swing available. Optimal input range is achieved when the following relation maintained.

$$\frac{1}{2}(\Delta V_{n5} + \Delta V_{p7}) = \frac{1}{3}[V_{ss} - (\Delta V_I + V_{tn5} + V_{tp7})] \tag{A.5}$$

4. g_m Control

By (A.4), the effective g_m is varied by bias currents of level shifters whose gate voltages can be controlled by a PLL.

5. Circuits Connected to VCT

Circuits connected to a VCT can be class A or class AB of a single stage. There are additional nondominant poles due to current mirrors if class AB is used. Therefore class A of a single stage is more suitable for high frequency application because there is only one non-dominant pole.

6. High Frequency Application

This new CMOS VCT is attractive for high frequency application, since the overall input device size can be small. However, doublet at source nodes should be high enough.

7. SPICE Simulation Results

The following results are simulated by SPICE2.G and based on MOSIS process parameters. Device sizes $(\frac{W}{L})$'s of all NMOS's are 5μ/4μ and those of PMOS's are 10μ/4μ. Bias current for level shifters is 10μA. The power supply is ±5V.

Linearity Error	Differential Input Voltage (V) up to
1.00%	1.18
0.50%	1.02
0.05%	0.90

8. Conclusions

The circuit of this CMOS VCT is simple and its device size ratio required to achieve linearity is not critical. It demonstrates good linearity and large input dynamic range. It is attractive for high frequency application.

REFERENCES

[1] B. J. Hosticka, "Impact of Scaling on MOS Analog Performance," *IEEE Proceedings* 73 pp. 25-29 (Jan. 1985).

[2] E. A. Vittoz, "The Design of High-Performance Analog Circuits on Digital CMOS Chips," *IEEE J. Solid-State Circuits* SC-20 pp. 657-665 (June 1985).

[3] A. K. Takla and Y. A. Haque, "A Single-Chip 300 Baud FSK Modem," *IEEE J. Solid-State Circuits* SC-19 pp. 846-854 (Dec. 1984).

[4] K. Yamamoto, S. Fujii, and K. Matsuoka, "A Single Chip Modem," *Solid-State Circuits* SC-19 pp. 856-860 (Dec. 1984).

[5] C. A. Laber and P. P. Lemaitre, "A Monolithic 1200 Baud FSK CMOS Modem," *IEEE J. Solid-State Circuits* SC-19 pp. 861-868 (Dec. 1984).

[6] W. Black, D. J. Allstot, and R. A. Reed, "A High Performance Low Power CMOS Channel Filter," *IEEE J. Solid-State Circuits* SC-15 pp. 929-938 (Dec. 1980).

[7] D. Senderowicz, S. F. Dreyer, J. H. Huggins, C. F. Rahim, and C. A. Laber, "A Family of Differential NMOS Analog Circuits for a PCM Codec Filter Chip," *IEEE J. Solid-State Circuits* SC-17 pp. 1014-1023 (Dec. 1982).

[8] J. L. McCreary and P. R. Gray, "All-MOS Charge Redistribution Analog-to-Digital Conversion Techniques: Part I," *IEEE J. Solid-State Circuits* SC-109 pp. 371-379 (Dec. 1975).

[9] P. R. Gray, D. Senderowicz, H. O'Hara, and B. M. Warren, "A Single-Chip NMOS Dual-Channel Filter for PCM Telephony Applications," *IEEE J. Solid-State Circuits* SC-14 pp. 981-991 (Dec. 1979).

[10]

R. H. Dennard, F. H. Gaensslen, H. Yu, V. L. Rideout, E. Bassouts, and A. R. LeBlanc, "Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions," *IEEE J. Solid-State Circuits* SC-9 pp. 256-268 (Oct. 1974).

[11]

P. K. Chatterjee, W. R. Hunter, T. C. Holloway, and Y. T. Lin, "The Impact of Scaling Laws on the Choice of N-Channel or P-Channel for MOS VLSI," *IEEE Elect. Dev. Letters* EDL-1 pp. 220-223 (Oct. 1980).

[12]

S. Wong and C. A. T. Salama, "Impact of Scaling on MOS Analog Performance," *IEEE J. Solid-State Circuits* SC-18 pp. 106-114 (Feb. 1983).

[13]

J. R. Brews, W. Fichtner, E. H. Nicollian, and S. M. Sze, "Generalized Guide for MOSFET Miniaturization," *IEEE Elect. Dev. Letters* EDL-1 pp. 2-4 (Jan. 1980).

[14] K. Matsui, T. Matsuura, S. Fukasawa, Y. Izawa, Y. Toba, N. Miyake, and K. Nagasawa, "CMOS Video Filters Using Switched-Capacitor 14-MHZ Circuits," *IEEE Journal Solid-State Circuits* SC-20 pp. 1096-1102 (Dec. 1985).

[15] D. B. Ribner and M. A. Copeland, "Biquad Alternatives for High-Frequency Switched-Capacitor Filters," *IEEE Journal Solid-State Circuits* SC-20 pp. 1085-1095 (Dec. 1985).

[16]

H. Masuda, M. Nakai, and M. Kubo, "Characteristics and Limitation of Scaled-Down MOSFET's due to Two-Dimensional Field Effect," *IEEE Trans. Elect. Dev* ED-26 pp. 980-986 (June 1979).

[17]

T. H. Ning, P. W. Cook, R. H. Dennard, C. Osburn, S. Schuster, and H. N. Yu, "1 μm MOSFET VLSI Technology: Part IV -- Hot-Electron Design Constraints," *IEEE Trans. Elect. Dev* **ED-26** pp. 346-353 (April 1979).

[18] A. Reisman, "Device, Circuit, and Technology Scaling to Micron and Submicron Dimensions," *IEEE Proceedings* **71** pp. 550-565 (May 1983).

[19]

T. Enomoto, T. Ishihara, M. Yasumoto, and T. Aizawa, "Design, Fabrication, and Performance of Scaled Analog IC's," *IEEE J. Solid-State Circuits* **SC-18** pp. 395-402 (Aug. 1983).

[20]

M. P. Lepselter, D. S. Alles, H. J. Levinstein, G. E. Smith, and H. A. Watson, "A system Approach to 1-μm NMOS," *IEEE Proceedings* **71** pp. 640-656 (May 1983).

[21]

W. Fichtner, R. L. Johnson, and D. J. Rose, "Three-Dimensional Modeling of Small-Size MOSFET's," *IEEE Trans. Elect. Dev* **ED-28** p. 1215 (1981).

[22] P. L. Shah, "Refractory Metal Gate Process for VLSI Applications," *IEEE Trans. Elect. Dev* **ED-26** pp. 631-640 (April 1979).

[23]

P. R. Gray and R. G. Meyer, "MOS Operational Amplifier Design -- A Tutorial Overview," *IEEE J. Solid-State Circuits* **SC-17** pp. 969-982 (Dec. 1982).

[24] R. Kaneshiro, "Circuit and Technology Consideration for High Frequency Switched-Capacitor Filters," *Ph. D. Dissertation, Univ. of California, Berkeley* **M83/41**(July 1983).

[25]

T. C. Choi, R. T. Kaneshiro, R. W. Brodersen, and P. R. Gray, ", "High-Frequency CMOS

Switched-Capacitor Filters for Communications Application," *IEEE J. Solid-State Circuits* SC-18 pp. 652-664 (Dec. 1983).

[26] N. S. Lin and A. Oki, "Effect of Impact Ionization on Circuit Performance," *231 Term Paper Report*, (April 1984).

[27] K. C. Hsieh, "Noise Limitations in Switched-Capacitor Filters," *Ph. D. Dissertation, Univ. of California, Berkeley* UCB/ERL M82/39(May 1982).

[28] M. A. Copeland and J. M. Rabaey, "Dynamic Amplifier for MOS Technology," *Electron Lett.* 15 pp. 301-302 (May 1979).

[29] R. Castello and P. R. Gray, "Performance Limitations in Switched-Capacitor Filters," *IEEE Trans. Circuits Syst* CAS-32 pp. 865-876 (Sept. 1985).

[30]

K. Ogawa, B. Owen, and H. J. Boll, "A Long-Wave Optical Receiver Using a Short-Channel Si-MOSFET," *The Bell System Technical Journal* 62 pp. 1181-1187 (May-June 1983).

[31] J. H. Fischer, "Noise Sources and Calculation Techniques for Switched Capacitor Filters," *IEEE J. Solid-State Circuits* SC-17 pp. 742-752 (Aug. 1982).

[32] C. A. Gobet and A. Knob, "Noise Analysis of Switched Capacitor Networks," *IEEE Trans. Circuits Syst* CAS-30 pp. 37-43 (Jan. 1983).

[33] R. Castello, *Private Communication.*

[34] J. C. Bertails, "Low-Frequency Noise Considerations for MOS Amplifier Design," *IEEE J. Solid-State Circuits* SC-14 pp. 774-776 (Aug. 1979).

[35] E. M. Cherry and D. E. Hopper, "Amplifying Devices and Low-Pass Amplifier Design," *John Wiley and Sons, Inc.*, (1968).

[36] D. Jones and W. Webb, "Chopper-Stabilization Op-amp Combines MOS and Bipolar Elements on One Chip," *Electronics*, pp. 110-114 (Sept. 1973).

[37] R. Castello and P. R. Gray, "A High Performance Micropower Switched-Capacitor Filter," *IEEE J. Solid-State Circuits* SC-20 pp. 1122-1132 (Dec. 1985).

[38] P. R. Gray, "EE241," *class notes*, (Winter, 1983).

[39] S. Lewis, *Private Communication.*

[40] K. C. Hsieh, P. R. Gray, D. Senderowicz, and D. Messerschmitt, "A Low-Noise Chopper-Stabilized Differential Switched Capacitor Filtering technique," *IEEE J. Solid-State Circuits* SC-16 pp. 708-715 (Dec. 1981).

[41] B. J. Hosticka, "Dynamic CMOS Amplifiers," *IEEE J. Solid-State Circuits* SC-15 pp. 887-894 (Oct. 1980).

[42] M. Degrauwe, J Rijmenants, E. A. Vittoz, and H. J. De Man, "Adaptive biasing CMOS amplifiers," *IEEE J. Solid-State Circuits* SC-17 pp. 522-528 (June 1982).

[43] C. Laber, C. Rahim, G. Uehara, P. Kwok, and P. R. Gray, "A High-Performance 3μ CMOS Standard Cell Library," *IEEE CICC Proceedings*, p. 21 (1986).

[44] W. C. Black, "High Speed CMOS A/D Conversion Technique," *Ph. D. Dissertation, Univ. of California, Berkeley* UCB/ERL M80/54(Nov. 1980).

[45] R. Castello, "Low-Voltage Low-Power MOS Switched-Capacitor," *Ph. D. Dissertation, Univ. of California, Berkeley* UCB/ERL M84/67(Aug. 1980).

[46] Silvar-Lisco, Palo Alto, California, *The Switched Capacitor Network Simulator.*

[47] F. G. Stremler, "Introduction to Communication Systems," *Addison-Wesley Publishing Company*, (1982).

[48] G. Varnell, P. L. Shah, and R. H. Havemann, "MOS and Bipolar VLSI Technologies Using Electron-Beam Lithography," *IEEE Proceedings* 71 pp. 612-639 (May, 1983).

[49] C. Hu, S. Tam, F. Hsu, P. Ko, T. Chan, and K. Terrill, "Hot-Electron-Induced MOSFET Degradation -- Model, Monitor, and Improvement," *IEEE J. Solid-State Circuits* SC-20 pp. 295-305 (Feb. 1985).

[50] J. Matsunaga, M. Konaka, S. Kohyama, and H. Iiziku, *11th. Int. Conf. Solid State Devices*, p. 45 (Aug. 27, 1979).

[51] H. Gesch, J. P. Leburton, and G. E. Dorda, "Generation of Interface Stages by Hot-Hole Injection MOSFET's," *IEEE Trans. Electron Devices* ED-29 p. 913 (1982).

[52] S. M. Sze, "Physics of Semiconductor Physics," *Wiley Interscience*, (1981).

[53] P. K. Ko and R. S. Muller, "A Unified Model for Hot Electron Currents in MOSFET's," *IEDM Tech. Dig.*, p. 600 (1981).

[54] C. Hu, "Hot-Electron Effects in MOSFET's," *IEDM Tech. Dig.*, p. 176 (1983).

[55] F. C. Hsu, P. K. Ko, C. Hu, and R. S. Muller, "An Analytical Breakdown Model for Short-Channel MOSFET's," *IEEE Trans. Electron Devices* ED-29 pp. 1735-1740 (1982).

[56] F. C. Hsu, R. S. Muller, and C. Hu, "A Simplified Model of Short Channel MOSFET Characteristics in the Breakdown Mode," *IEEE Trans. Electron Devices* ED-30 pp. 571-576 (1983).

[57] K. Hanson, W. A. severin, E. R. Klinkovsky, D. C. Richardson, and J. Hochschild, "A 1200 Bit/s QPSK Full Duplex Modem," *IEEE J. Solid State Circuits* SC-19 pp. 878-887 (Dec. 1984).

[58] R. W. Brodersen, P. Hurst, and D. Allstat, "Switched-Capacitor Applications in Speech Processing," *ISCAS Proceeding*, pp. 732-737 (April 1980).

[59] L. McCready, A. Chowaniec, S. Watson, N. Methta, and D. Choii, "A Monolithic Formant-Based Parametric Speech Synthesizer," *ISCAS Proceeding*, pp. 986-988 (April 1981).

[60] O. Agazzi, D. Hodges, and D. Messerschmitt, "Echo Canceller for a 80 kbs Baseband Modem," *IEEE ISSCC Tech. Dig.*, pp. 155-145 (Feb. 1982).

[61] D. J. Vlach, T. R. Viswanathan, and K. Singhal, "Active Filters for Intermediate Frequencies Using Low-Gain Transducers," *IEEE Trans. Circuits and Sys.* CAS-21 pp. 79-74 (Jan. 1974).

[62] D. Senderowicz, "An NMOS Integrated Vector-Locked Loop," *Phd Dissertation, University of California, Berkeley* UCB/ERL M82/83, Nov. 1982(Dec. 1982).

[63] H. Khorramabadi and P. R. Gray, "High-Frequency CMOS Continuous Time Filters," *IEEE Journal of Solid-State Circuits* SC-19 pp. 939-948 (Dec. 1984).

[64] R. Torrance, T. R. Viswanatham, and J. V. Hanson, "CMOS Voltage to Current Transducers," *IEEE Trans. Circuits and Sys.* CAS().

[65] C. L. Stout, "The Design of a Continuous-Time Integrator for Use at 1MHZ," *MS Report, University of California, Berkeley*, (Oct. 1983).

[66] A. Nedungadi and T. R. Viswanathan, "Design of Linear CMOS Transconductance Elements," *IEEE Trans. Circuits and Sys.* CAS-31 pp. 891-894 (Oct. 1984).

[67] G. A. Korn and T. M. Korn, "Electronic Analog and Hybrid Computer," *McGraw-Hill*, (1972).

[68] IL S. Han and Song B. Park, "Voltage-Controlled Resistor by Two MOS Transistors and its Application to Active RC Filter MOS Integration," *IEEE Proceedings* 72 pp. 1655-1657 (Nov. 1984).