AN ABSTRACT OF THE THESIS OF

Howard C. Yang for the degree of Doctor of Philosophy in Electrical and
Computer Engineering presented on September 11, 1989.
Title: DESIGN CONSIDERATIONS FOR FAST-SETTLING OPERATIONAL AMPLIFIERS

Redacted for Privacy

Abstract approved:_____
                              David J. Allstot

The design considerations for fast-settling operational
amplifiers (opamps) differ significantly between sampled-data switched-
capacitor (SC) and conventional continuous-time applications. In SC
circuits, the shape of the output voltage waveform of an opamp is of no
consequence provided that the output settles to within a specified
tolerance of its steady-state value prior to the next sampling instant.
This feature allows for an optimum opamp frequency shaping to obtain a
minimum small-signal settling time. The theory applies to any opamp
that is well-approximated by a two-pole model, including the
conventional two-stage and single-stage folded-cascode topologies. As
the commonly-used equivalent-circuit Miller-effect model for frequency
compensation has generally been improperly applied to two-stage
transconductance amplifiers, it does not provide sufficient accuracy to
achieve the optimum phase margin condition. Therefore, the use of
equivalent-circuit models has been refined to provide greater accuracy
and to eliminate some previous misconceptions.

DESIGN CONSIDERATIONS FOR FAST-SETTLING

OPERATIONAL AMPLIFIERS

by

Howard C. Yang

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed September 11, 1989

Commencement June 1990

APPROVED:

Redacted for Privacy

Professor of Electrical & Computer Engineering
        in charge of major

Redacted for Privacy

Head of Department of Electrical & Computer Engineering

Redacted for Privacy

Dean of Graduate School

Date thesis is presented   Sept. 11, 1989

Typed by Howard C. Yang for   Howard C. Yang

To   My Parents:

Prof. Xun-Ren Yang
Prof. Shu-Zhuang Liang

You see things; and you say, "why?"
But I dream things that never were; and I say, "why not?"

———— G. Bernard Shaw

ACKNOWLEDGEMENTS

Finally, I want to dedicate this work to my parents. Their love, support and understanding have sustained me through all these years.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DESIGN CONSIDERATIONS FOR FAST-SETTLING OPERATIONAL AMPLIFIERS

I. INTRODUCTION

Since the first simple monolithic amplifier was fabricated in the early 1960s, analog integrated circuits (ICs) have expanded rapidly into one of the most important areas in modern electronics due to their high performance, low cost, and wide applications.

Prior to the mid-1970s, analog integrated circuits were implemented by using silicon bipolar transistor technology, while silicon metal-oxide-semiconductor field effect transistor (MOSFET) technology was utilized primarily for digital integrated circuits. In many applications (e.g., analog to digital converters), both analog and digital functions are required for the subsystem. It is undesirable to partition such subsystems into two separate technologies: bipolar analog and MOS digital ICs. In order to implement the subsystems containing both analog and digital functions on a single chip, analog MOSFET ICs have been developed. Although the first paper on analog MOS ICs was published in 1969 by Chalfan and Looney of Oregon State University [1], MOSFET technology did not make its profound impact to analog integrated circuits until a number of significant papers were published in the 1970s [2]. In the late-1970s, switched-capacitor (SC) circuits, which have been perhaps the most important application of MOS technology for analog ICs, were developed at AT&T Bell Labs and the University of California at Berkeley for

telecommunication systems [2]-[4]. Since then, analog MOS ICs have become one of the major members in the IC family.

MOS operational amplifiers (opamps) are core building cells for switched-capacitor circuits, and their performance, such as settling time and frequency response, directly affect the performance of SC circuits. Many aspects of the settling behavior and frequency response of operational amplifiers have been analyzed by various authors [5]-[10]. Generally, it has been shown that in order to obtain the desired settling/frequency characteristics, accurate frequency shaping must be employed. In conventional continuous-time applications, operational amplifiers are frequency compensated (either internally for two-stage opamps or via the loading capacitance for single-stage opamps) for a unity-gain phase margin of approximately 60 degrees to insure closed-loop stability, and to maximize flatness of the closed-loop amplitude response [6]. Flat gain characteristics are particularly important in minimizing waveform distortion in continuous-time pulse-amplifier applications. By contrast, the specific shapes of the opamp output waveforms are of no consequence in some applications such as in sampled-data switched-capacitor circuits where it is only necessary that the outputs settle to within a specified tolerance of their final values prior to the next sampling instant. By exploiting this unique feature of SC circuits, an optimum opamp phase margin criterion has been developed that gives the minimum settling time (MST) of an opamp for a given gain/bandwidth product.

Our analysis has shown that the small-signal settling time of a two-pole (one or two-stage) operational amplifier exhibits a well-defined minimum for a specific value of the unity-gain phase margin.

Since the settling time is also shown to be strongly dependent on phase margin, precise frequency shaping is required in order to achieve the minimum settling time. Unfortunately, the commonly-used equivalent-circuit model of the opamp based on the Miller approximation has in the past been improperly applied to two-stage transconductance amplifiers and does not provide sufficient accuracy with which to achieve the MST. Therefore, improved equivalent-circuit models have been developed. In addition to being significantly more accurate than the previous modeling wherein the dominant pole was always associated with the first stage, the improved modeling accounts for the possibility of a second-stage-dominant pole. In fact, for CMOS and GaAs operational transconductance amplifiers (OTA) that drive on-chip capacitive loads, the dominant pole is usually associated with the second stage. Our results show that whichever pole is dominant before compensation remains dominant after compensation. Hence, some considerable confusion that has existed previously regarding pole-splitting frequency compensation is eliminated.

In Chapter II, we develop a new optimum phase margin design criterion for OTA's used in SC applications. In Chapter III, we present some improved modeling techniques for two-stage opamps that provide sufficient accuracy with which to achieve the optimum phase margin condition. In Chapter IV, we extend the MST design technique to a one-stage folded-cascode CMOS opamp by using a newly developed two-pole small-signal model. Chapter V concludes this thesis with discussion on sensitivities to MOS process variations and suggestions for future work.

II. OPTIMUM PHASE MARGIN FOR SC APPLICATIONS

We begin this chapter with a brief review of the frequency and step-response equations of a second-order system. We then use this theory to derive an optimum unity-gain phase margin that allows for the maximum sampling frequency in SC circuits. Finally, design equations are developed to allow two-pole operational amplifiers to be compensated for the optimum phase margin condition.

2.1 Frequency/Time Response Equations for a Two-Pole System

In most switched-capacitor circuits, the maximum size of the output voltage step between sampling instants is small enough so that only small-signal analysis is necessary. Therefore, the operational amplifier is modeled using a linear two-pole model which is very useful in determining the pole-splitting compensation capacitance [11],[12]. The open-loop transfer function of the two-pole small-signal circuit prior to frequency compensation is given by

$$a(s) = \frac{a_0}{(1 - s/\omega_1)(1 - s/\omega_2)} \qquad (1)$$

where a_0 is the low-frequency gain of the operational amplifier and ω_1 and ω_2 are the radian frequencies of first and second left-half-plane (LHP) poles, respectively. The unity-gain closed-loop transfer function is given by

$$A(s) = \frac{a(s)}{a(s) + 1} = \frac{A_0}{(s/\omega_0)^2 + 2k(s/\omega_0) + 1} \qquad (2)$$

where $\qquad A_0 = \frac{a_0}{a_0 + 1} \qquad (3)$

$$\omega_0 = \sqrt{[\omega_1 \omega_2 (a_0 + 1)]} \qquad (4)$$

and
$$k = \frac{\omega_1 + \omega_2}{2\omega_0}. \tag{5}$$

Defining the pole separation factor as $\beta = \omega_2/\omega_1$, the damping factor is expressed in terms of β as

$$k = \frac{1 + \beta}{2\sqrt{[\beta(a_0 + 1)]}}. \tag{6}$$

The second-order system of Eqn. (2) has three possible responses to a voltage step input; namely, overdamped, critically damped and underdamped corresponding to $k > 1$, $k = 1$ and $k < 1$, respectively. The normalized time responses are obtained from Eqn. (2) using the inverse Laplace transform as:

Overdamped (exponentially approaching the steady-state value), $k > 1$:

$$v_0(t) = 1 - \frac{1}{2\sqrt{(k^2 - 1)}} \left[\frac{1}{k_1} exp(-k_1\omega_0 t) - \frac{1}{k_2} exp(-k_2\omega_0 t) \right] \tag{7}$$

where $k_1 = k - \sqrt{(k^2 - 1)}$ and $k_2 = k + \sqrt{(k^2 - 1)}$;

Critically damped (also exponentially approaching the final value), $k = 1$:

$$v_0(t) = 1 - (1 + \omega_0 t) \, exp(-\omega_0 t); \tag{8}$$

Underdamped (overshoot followed by exponentially damped sinusoid), $k < 1$:

$$v_0(t) = 1 - \left[\frac{k}{\sqrt{(1 - k^2)}} sin(\sqrt{1 - k^2} \, \omega_0 t) + cos(\sqrt{1 - k^2} \, \omega_0 t) \right] exp(-k\omega_0 t) \tag{9}$$

2.2 Small-Signal Settling Behavior of a Second-Order System

The small-signal settling time of an operational amplifier, t_s, is defined (Fig. 1) as the minimum time required for the output voltage of the amplifier to settle to within an error tolerance, D, of its final steady-state value. Using Eqns. (7)-(9), it can be shown that the settling behavior of a two-pole system is strongly dependent on the damping factor k. For a nominal value of the low-frequency open-loop voltage gain, a_0, the damping factor is only a function of the pole separation factor, β, according to Eqn. (6).

The time response of a second-order system is usually discussed in terms of the normalized variable $\omega_0 t$ [6]. Since ω_0 is a function of k from Eqn. (5), and therefore also of β from Eqn. (6), $\omega_0 t$ is a dependent variable. As a first step in finding the shortest response time of the second-order system, the time of the first peak of the underdamped response, t_p (Fig. 1), is determined by setting the first derivative of Eqn. (9) equal to zero to obtain

$$t_p = \frac{\pi}{\omega_0 \sqrt{(1 - k^2)}}.$$ (10)

Now, using Eqn. (4) in Eqn. (10), t_p as a function of β is

$$t_p = \frac{\pi}{\omega_1 \sqrt{[\beta(1 + a_0) - (1 + \beta)^2/4]}}.$$ (11)

The minimum of t_p with respect to β determines the shortest possible response time, and is determined by setting the first derivative of Eqn. (11) equal to zero and solving to obtain $\beta = 1 + 2a_0$. Substituting this value into Eqn. (6) gives the optimum damping factor as
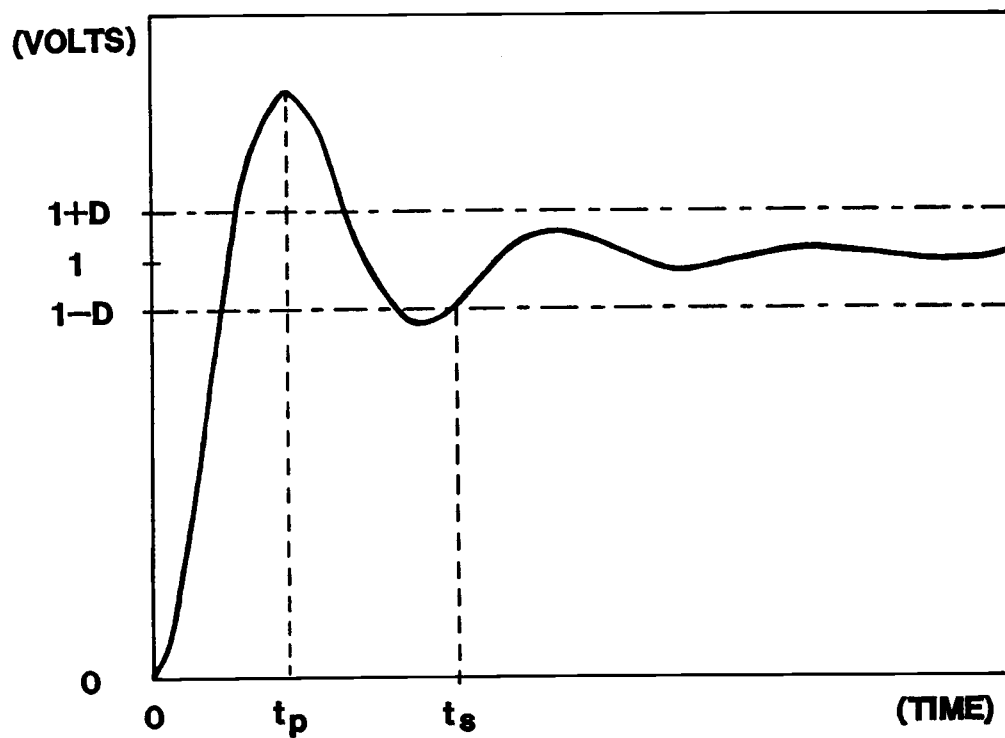
Fig. 1. Definition of small-signal settling time, t_s, with an error tolerance of D percent. The first positive peak of the underdamped response occurs at time t_p.

$$k = \frac{1 + a_o}{\sqrt{[(1 + 2a_o)(1 + a_o)]}}.$$ (12)

Since $a_o \gg 1$ is usually true in practical cases, Eqn. (12) simplifies to $k \simeq 0.707$ meaning that the absolute fastest response of the system corresponds to the Butterworth response. However, this does not indicate the minimum settling time in general since the error band, D, was not considered in the above analysis. In fact, it can be shown that the Butterworth response is underdamped exhibiting a transient overshoot of 4.3 percent, and therefore yields the minimum settling time only for the cases where $D \geq 0.043$. The Butterworth response provides a unity-gain phase margin of approximately 60 degrees.

For high precision SC circuits, an error tolerance smaller than 4.3 percent is required. The minimum settling time is obtained when the two-pole opamp is underdamped and the phase margin is chosen so that the first peak of the step response just touches the upper settling error limit as shown in Fig. 2. Referring to the figure, it is clear that this is the minimum since either an increase or decrease in the damping factor results in greater settling time*. Therefore, to realize the maximum sampling frequency, the phase margin condition "upper error bound = first peak" must be satisfied. The normalized voltage of the first peak determined from Eqn. (9) is equal to one plus the overshoot,

$$First\ peak = 1 + exp\ [-k\pi/\sqrt{(1 - k^2)}]$$ (13)

---

* Actually, the absolute minimum settling time occurs for the case wherein a doublet is deliberately introduced to give a three-pole one-zero response with the singularities placed so that the first peak just touches the upper bound, and the second peak just touches the lower error bound. It has been shown that this is probably not a practical solution because of high sensitivities to parameter variations [13].
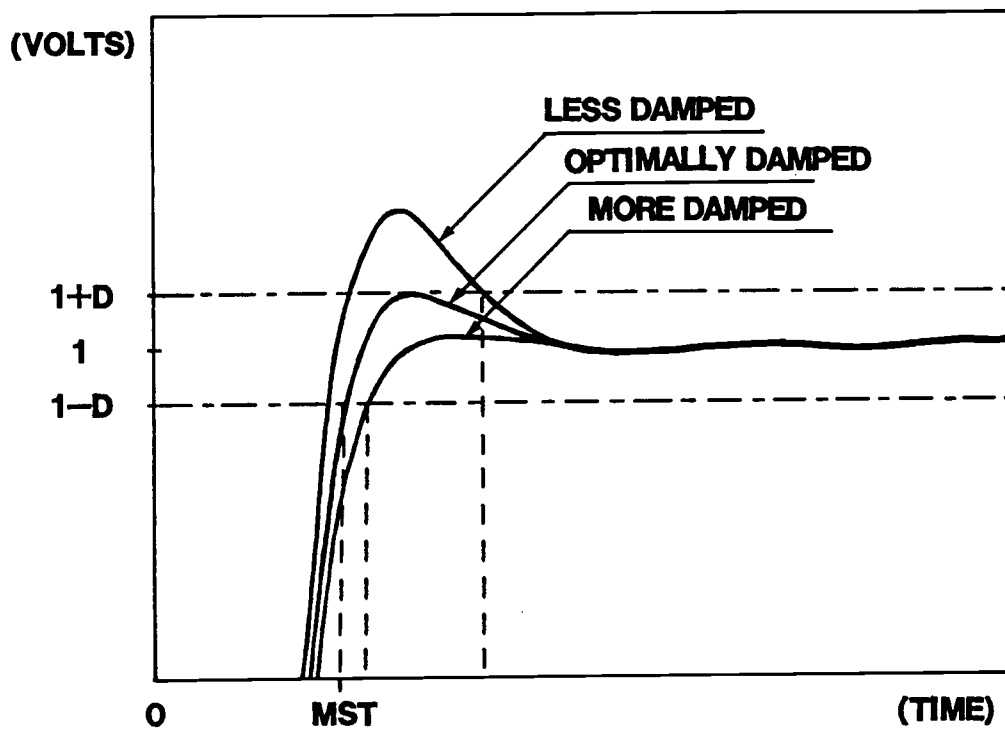
Fig. 2. Definition of the minimum small-signal settling time (MST) for
a two-pole system with an error tolerance of *D* percent.
Settling time is increased for either more or less damping.

and since the upper error bound voltage is $(1 + D)$, then

$$D = exp[-k\pi/\sqrt{(1 - k^2)}].$$ (14)

Using Eqn. (6) with $\beta \gg 1$, the optimum pole-separation factor, β_0, is determined from Eqn. (14) as [14],[15]

$$\beta_0 \simeq \frac{4(a_0 + 1)}{1 + (\pi/\ln D)^2}.$$ (15)

For a given error band, D, the minimum settling time of the SC circuit is achieved by designing the two-pole opamp so that the poles are separated in accordance with Eqn. (15). In opamp design, the settling time is usually considered as a function of phase margin, ϕ_m, rather than of the pole separation factor, β, or equivalently of the damping factor, k. Thus, a relationship between the settling time and the phase margin is desirable. Approximating the small-signal unity-gain frequency as $\omega_u = a_0\omega_1$, and the unity-gain time constant, as $\tau_u = 1/\omega_u$, the phase margin, ϕ_m, of a two-pole system (assuming widely separated poles) is expressed as

$$\phi_m \simeq 180^o - tan^{-1}(\omega_u/\omega_1) - tan^{-1}(\omega_u/\omega_2)$$
$$= 180^o - tan^{-1}(\omega_u/\omega_1) - tan^{-1}(\omega_u/\beta\omega_1).$$ (16)

For example, with $a_0 = 1000$ V/V and $f_1 = \omega_1/2\pi = 1$ KHz, the normalized small-signal settling time, t_s, of the two-pole system is plotted versus the unity-gain phase margin in Fig. 3. Clearly, t_s is a very strong function of the phase margin, and thus in order to obtain maximum operating speed, an accurate phase margin is required. The phase margin corresponding to the minimum settling time, $\phi_m(MST)$, is also a function of the error band tolerance as seen in Fig. 3, and is derived using Eqns. (15) and (16) (with the assumption that $tan^1 (\omega_u/\omega_1) \simeq 90^o$) as
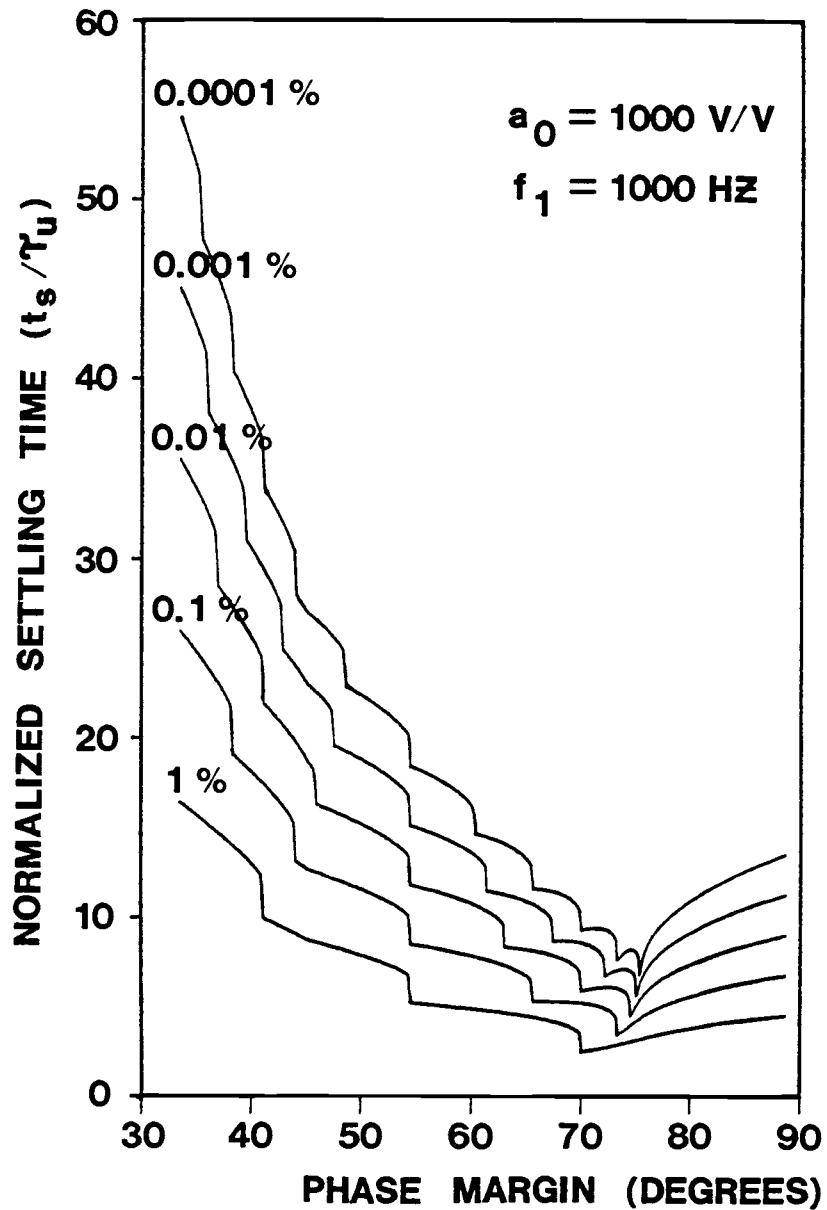
Fig. 3. Normalized small-signal settling time versus unity gain phase margin for common values of the error tolerance, D. The discontinuities in response time are associated with the natural frequency of the system, and occur when response peaks or valleys just enter the bounded region for increasing phase margin.

$$\phi_m(MST) \simeq 90^o - tan^{-1} \left[\frac{1 + (\pi/\ln D)^2}{4}\right].$$ (17)

For $D = 0.01$, Eqn. (17) gives $\phi_m(MST) \simeq 70^o$. Note that when $D \to 0$, $\phi_m(MST) \to 76^o$ which corresponds to the critically damped response of the two-pole system.

2.3 Frequency Compensation Considerations

In the previous section, it was shown that the Butterworth response provides the minimum settling time only for the rather impractical and imprecise cases wherein $D \geq 0.043$. For higher precision applications, D is much smaller, and therefore an improved compensation technique is required to achieve the minimum settling time.

Figure 4 shows a two-pole small-signal equivalent circuit of a two-stage operational amplifier which after adding compensation elements exhibits three LHP poles and one zero. Solving the network equations and using the dominant pole approximation technique [5],[16], the pole and zero frequencies after compensation are

$$\omega_1 \simeq - \frac{g_1 g_2}{(g_{m2} + g_1 + g_2)C_f + g_2 C_1 + g_1 C_2}$$ (18)

$$\omega_2 \simeq - \frac{(g_{m2} + g_1 + g_2)C_f + g_1 C_2 + g_2 C_1}{C_1 C_f + C_2 C_f + C_1 C_2}$$ (19)

$$\omega_3 \simeq - \frac{1}{R_f} \left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_f}\right)$$ (20)

and $$\omega_z = \frac{1}{C_f(1/g_{m2} - R_f)}$$ (21)

where $g_1 = 1/R_1$ and $g_2 = 1/R_2$. The values for the compensation capacitance, C_f, and the nulling resistance, R_f, are usually chosen so that the zero exactly cancels the first non-dominant pole [17]. Therefore, to obtain a simple two-pole system and to minimize the small-signal settling time for a given amplifier bandwidth, Eqns. (18)-(21) must satisfy

$$\omega_2 = \omega_z \tag{22}$$

and
$$\omega_3 = \beta_0 \omega_1 \tag{23}$$

where β_0 is determined from Eqn. (15). Solving Eqns. (22)-(23) gives the element values for optimum compensation as

$$R_f \simeq \frac{1 + (\pi/\ln D)^2}{4\ g_{m1}\ g_{m2}}\ \frac{1}{C_1}\ [(g_{m2}+g_1+g_2)C_f + g_1 C_2 + g_2 C_1] \tag{24}$$

and

$$C_f = \frac{C_1 + C_2}{(R_f - 1/g_{m2})(g_{m2} + g_1 + g_2)}. \tag{25a}$$

Substituting Eqn. (24) into (25a) and solving for C_f gives

$$C_f = [-b + \sqrt{(b^2 - 4ac)}]/2a \tag{25b}$$

where

$$a = \frac{[1 + (\pi/\ln D)^2]\ (g_{m2} + g_1 + g_2)^2}{4\ g_{m1}\ g_{m2}\ C_1} \tag{25c}$$

$$b = \frac{[1+(\pi/\ln D)^2](g_{m2}+g_1+g_2)(g_2 C_1+g_1 C_2)}{4\ g_{m1}\ g_{m2}\ C_1}\ \frac{(g_{m2}+g_1+g_2)}{g_{m2}} \tag{25d}$$

and

$$c = C_1 + C_2 \tag{25e}$$

The solutions of Eqns. (24) and (25) give the compensation element values which provide the MST response for the opamp.
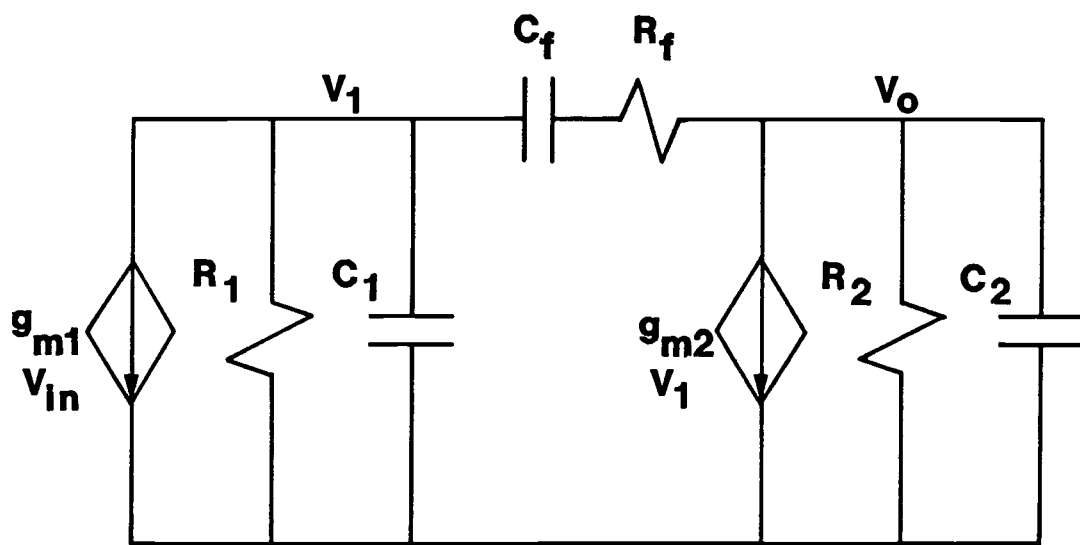
Fig. 4. Small-signal model of two-pole operational amplifier with pole-splitting compensation elements R_f and C_f.

III. IMPROVED EQUIVALENT CIRCUIT MODELING

In this chapter, we examine the use of small-signal equivalent circuits in frequency compensating two-pole operational amplifiers. Our results show that the previous techniques based on Miller-multiplied capacitance models are suitable only for those amplifiers in which the first-stage pole is dominant prior to compensation. In most SC circuits, the internal opamps are transconductance amplifiers which drive on-chip capacitive loads, and thus, for these opamps, the second-stage pole is dominant. Hence, in order to more accurately model the second-stage-dominant pole systems, and to subsequently realize the optimum phase margin condition, we have extended the small-signal model to include resistance effects in addition to Miller capacitance effects.

3.1 Problem Description

In order to test the accuracy of the compensation techniques, a CMOS two-stage transconductance amplifier (Fig. 5) was designed. As a first step in compensation, the parameter values (Table I) for the small-signal model of Fig. 6(a) were determined from Fig. 5. Unfortunately, this two-pole one-zero model is not convenient for calculating R_f and C_f because of the presence of the RHP zero introduced by C_{gd}**. Therefore, the next step usually involves using the Miller effect to develop the two-pole no-zero model of Fig. 6(b) [16],[18]. In this circuit, Miller capacitance C_{M1} (C_{M2}) is shunted

---

** With the series R_f-C_f compensation network connected in parallel with C_{gd}, the system actually exhibits three LHP poles and two zeros. The zero associated with the R_f-C_f network is the one that is moved into the left-half-plane to cancel the first nondominant pole. The zero associated with C_{gd} remains fixed in the right-half-plane at its original frequency.
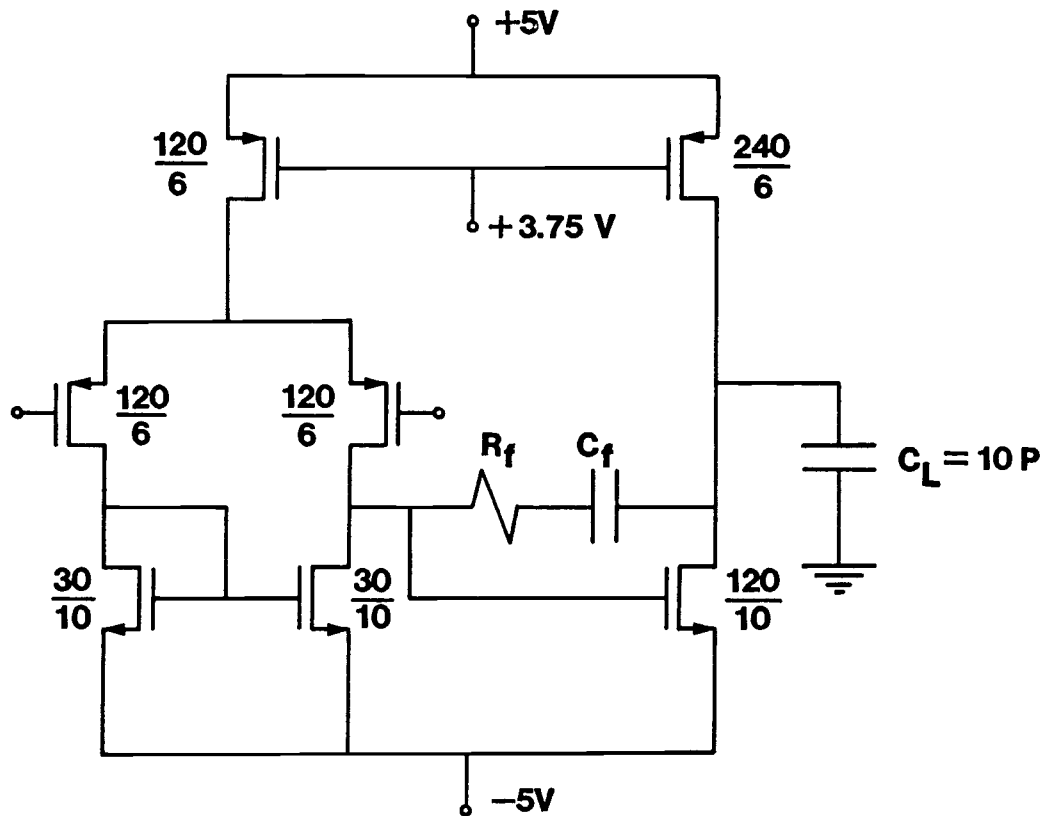
Fig.5. A typical two-stage CMOS operational amplifier used to exercise the compensation model. The fractions represent MOSFET (W/L) ratios in microns. R_f and C_f are the frequency compensation elements.

TABLE I

	1st Stage	2nd Stage
$g_m(\mu S)$	87.60	266.00
$C'(pF)$	0.59	10.28
$R'(k\Omega)$	866.55	197.63
$C_{gd}(fF)$	20.80	

Table I.   Small-signal parameter values for the CMOS two-stage opamp of
Fig. 5, and its small-signal model of Fig. 6(a).

across the first (second) stage to account for the effect of the gate-drain capacitance, C_{gd}, on the frequency of the first (second) stage pole. The RHP zero is assumed to be at a very high frequency relative to the unity-gain bandwidth, and is therefore ignored. In the two-pole model of Fig. 6(b), the total branch resistances are $R_1 = R_1' = 1/g_1'$ and $R_2 = R_2' = 1/g_2'$, and the total branch capacitances are given by

$$C_1 = C_1' + C_{M1} \tag{26}$$

and
$$C_2 = C_2' + C_{M2} \simeq C_2' \tag{27}$$

with
$$C_{M1} = C_{gd} \ (g_{m2}/g_2' + 1) \tag{28}$$

and
$$C_{M2} = C_{gd} \ \frac{(g_{m2}/g_2' + 1)}{g_{m2}/g_2'}. \tag{29}$$

We now show that when applied to a second-stage-dominant transconductance amplifier, this modeling approach does not allow frequency compensation to a sufficient degree of accuracy to obtain the MST response. Using the small-signal parameter values given in Table I, the two-pole one-zero model of Fig. 6(a), and the simplified two-pole model of Fig. 6(b) were simulated using SPICE2. It is clear from the responses shown in Fig. 7 that the simplified two-pole model (curve B) is a poor approximation to the two-pole one-zero model (curve A) especially at critical frequencies near the unity-gain frequency. This error in modeling before compensation significantly affects the accuracy of the unity-gain phase margin after compensation. For example, a desired error tolerance of $D = 0.01$ requires an optimum phase margin of 70 degrees. Using Eqns. (24)-(29) in conjunction with the model of Fig. 6(b), the compensation element values are calculated as $R_f = 12.3$ kΩ, and $C_f = 4.97$ pF. Using these values in compensating the CMOS two-stage opamp of Fig. 5 results in a simulated unity-gain

phase margin of 84 degrees. Because the resulting phase margin deviates significantly from the optimum value of 70 degrees, and because a doublet is also created due to inexact pole-zero cancellation, a very long settling time is observed. (Doublets may also result in reduced settling time in SC circuits [13] with a tradeoff in increased sensitivity to MOS process variations. In practice, an adaptive compensation technique [17] may be used to guarantee accurate pole-zero placement in the presence of typical variations.)

As previously stated, on-chip CMOS and GaAs transconductance amplifiers are usually second-stage dominant-pole systems with $R_1'C_1' < R_2'C_2'$ in the model of Fig. 6(a). Conventional application of the Miller-multiplied capacitance effect in developing the equivalent circuit of Fig. 6(b) predicts that the first-stage pole moves to a lower frequency for a nonzero value of C_{gd} and will become dominant if C_{gd} and/or g_{m2} is large enough as indicated in Fig. 8(a). This commonly-held view (c.f. [18], p. 377) is incorrect in terms of basic root locus theory wherein the poles of a two-pole system can never cross one another as pole-splitting negative feedback is applied [5]. The correct pole movements are as indicated in Fig. 8(b). With an increase of C_{gd} or g_{m2}, the second-stage pole moves to a lower frequency remaining dominant, while the first-stage pole moves to higher frequency remaining non-dominant. In general, the poles are always split apart with the application of pole-splitting compensation no matter which of the poles is dominant. It is now clear that the equivalent circuit model of Fig. 6(b) with Miller-multiplied capacitances is not an accurate representation of second-stage-dominant
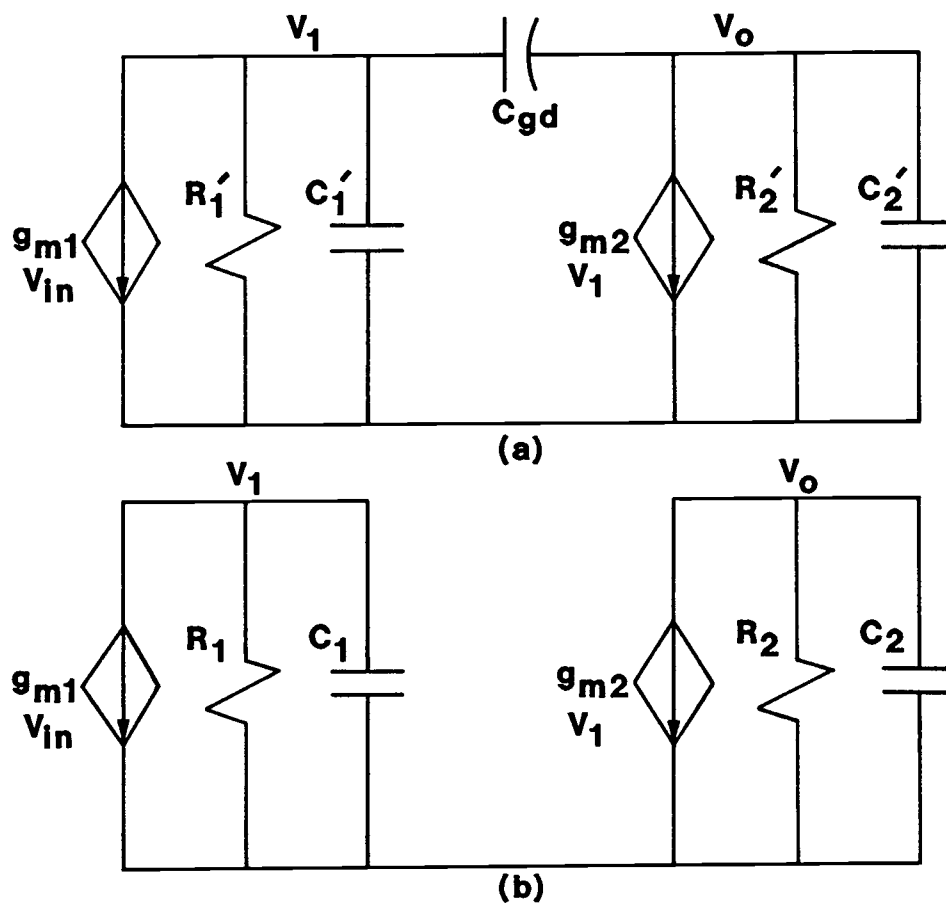
Fig. 6. Small-signal models of (a) the two-pole one-zero, and (b) the two-pole opamp before frequency compensation.
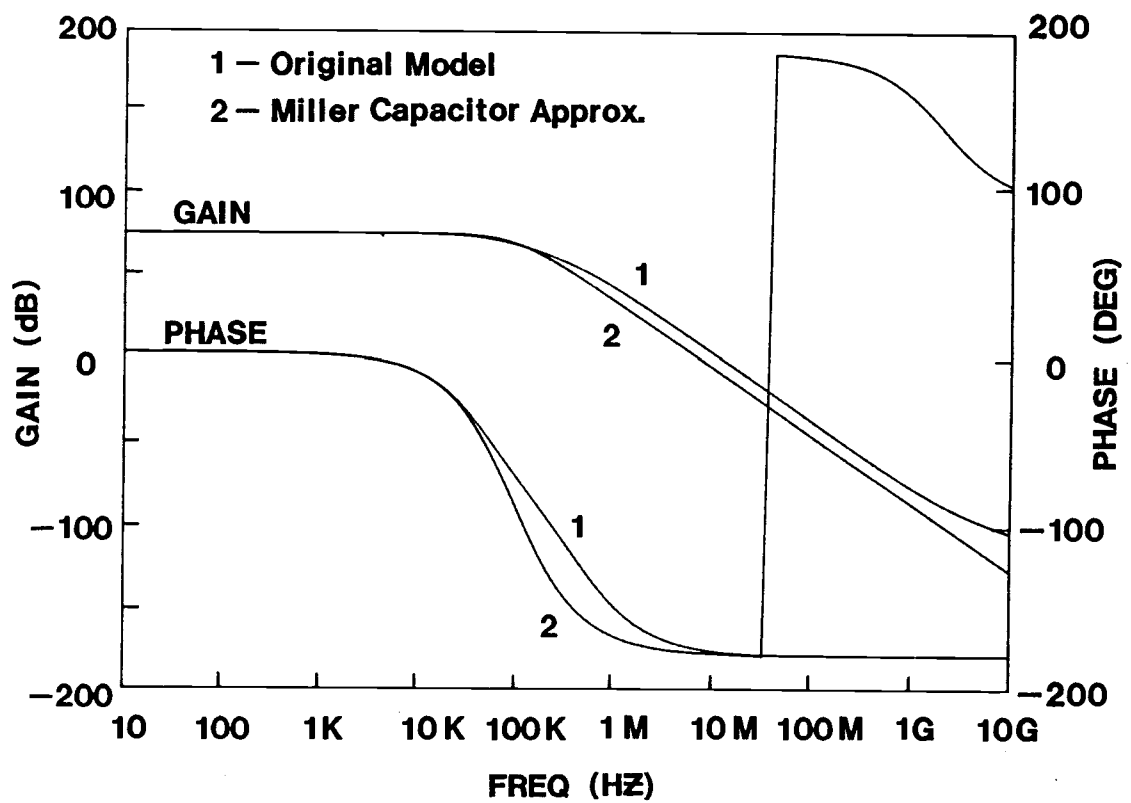
Fig. 7. Simulated frequency response for the original circuit of Fig. 6(a) and the Miller model of Fig. 6(b). Circuit parameters are given in Table I.
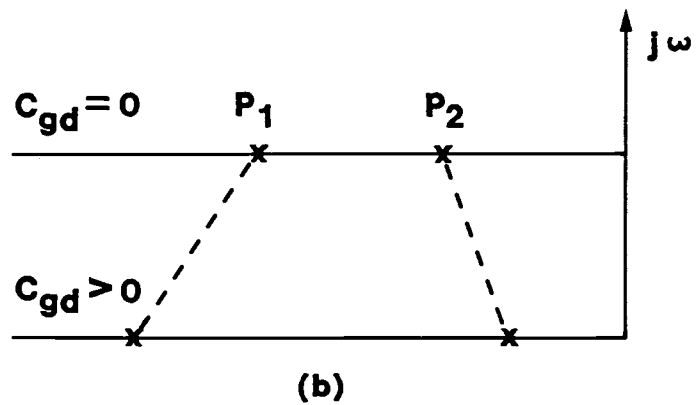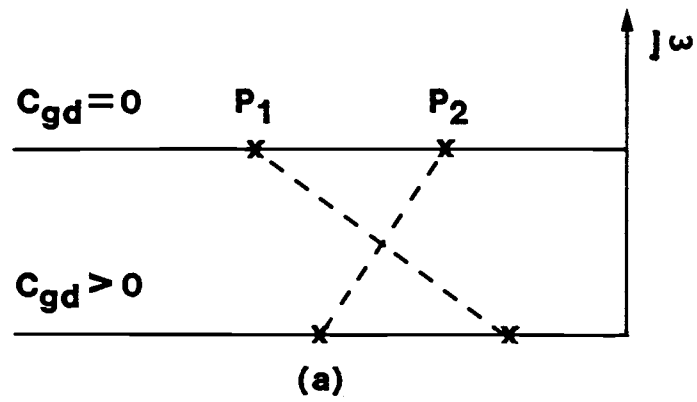
Fig. 8. (a) Unrealizable root-locus plot of pole splitting implied by conventional application of the Miller capacitance model to a second-stage dominant pole system, and (b) correct root locus plot predicted using the improved model.

transconductance amplifiers, and therefore cannot be used to achieve the MST response in this case.

3.2 Improved Equivalent Circuit Model

In order to achieve the MST response, modeling techniques have been developed that are more accurate at frequencies near the unity-gain frequency of the opamp. The improved equivalent circuit model includes the conventional first-stage-dominant one as a special case, and it is able to predict accurately the pole movement for optimum frequency compensation of second-stage-dominant transconductance amplifiers.

In the circuit of Fig. 6(a), the impedance looking into C_{gd} from the left, Z_1, is found from Fig. 9(a) as

$$Z_1 = \frac{1}{sC_{gd}(g_{m2}/g_2' + 1)} \frac{1 + s(C_2' + C_{gd})/g_2'}{1 + sC_2'/(g_{m2} + g_2')}. \tag{30}$$

When the operating frequency is low, Eqn. (30) becomes

$$Z_1 \simeq \frac{1}{sC_{M1}} = \frac{1}{sC_{gd}(g_{m2}/g_2' + 1)}. \tag{31}$$

Since $C_{gd}(g_{m2}/g_2' + 1)$ is simply the Miller capacitance, the conventional model clearly applies at low frequencies (Fig. 10). When the operating frequency is between $g_2'/(C_2' + C_{gd})$ and $(g_{m2} + g_2')/C_2'$,

$$Z_1 \simeq R_{M1} = \frac{1}{g_{m2} + g_2'} (1 + C_2'/C_{gd}). \tag{32}$$

In this frequency range, the impedance is not due to the Miller capacitance, but rather to a resistor, R_{M1}, which may be called a "Miller resistance". When the operating frequency is higher than $(g_{m2} + g_2')/C_2'$,

Fig. 9. Circuits for calculating impedance looking into C_{gd} (a) from the first stage, and (b) from the second stage.

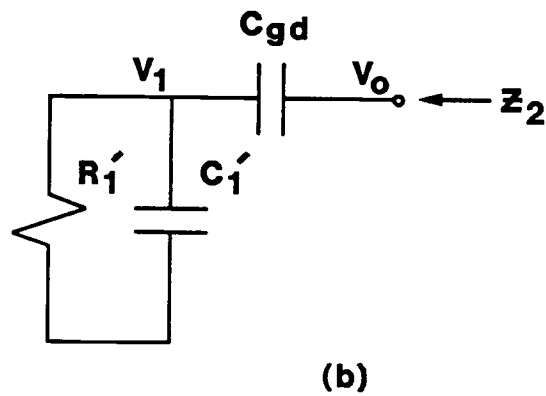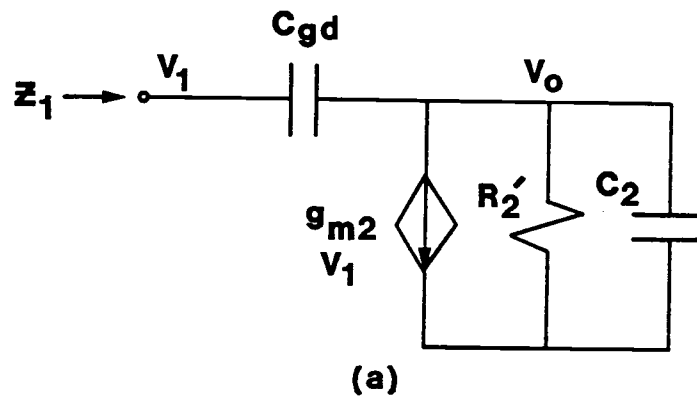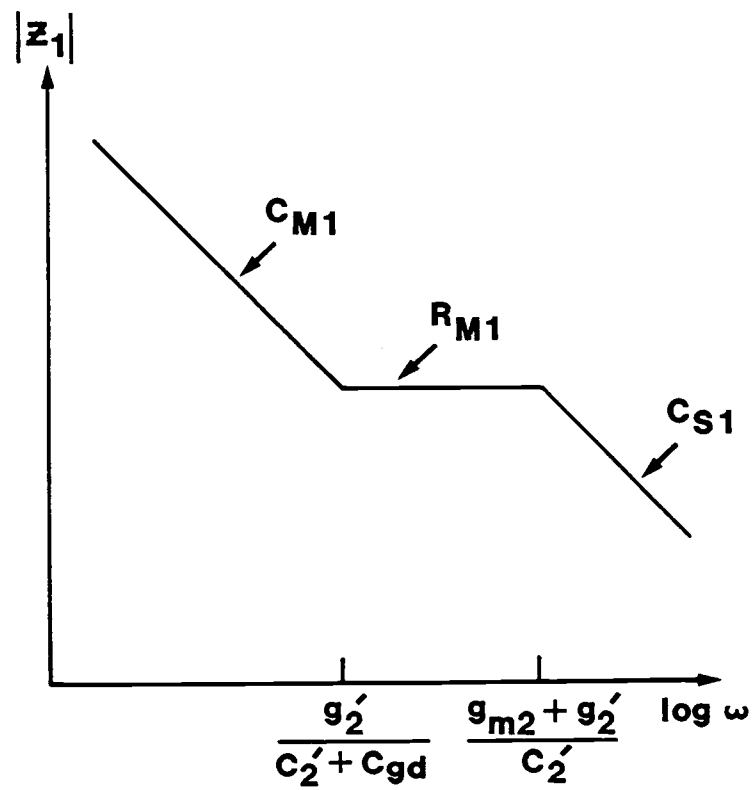Fig. 10. Impedance versus frequency looking into C_{gd} from the first stage.

$$Z_1 \simeq \frac{1}{sC_{s1}} = \frac{1}{s} \frac{C_{gd} + C_2'}{C_{gd} \, C_2'} \tag{33}$$

which is due to a capacitance, C_{s1}, equivalent to C_{gd} and C_2' in series. The impedance looking into C_{gd} from the right (Fig. 9(b)) is just C_{gd} in series with R_1' and C_1'. The complete equivalent circuit model shown in Fig. 11 is based on the impedances derived above. From this analysis, it is clear that the conventional equivalent circuit using the Miller-multiplied capacitors is an incomplete model accurate only for low frequencies.

SPICE simulations on the improved equivalent circuit model of Fig. 11 using the small-signal circuit parameters listed in Table I are compared with simulations of the original circuit model (Fig. 6(a)) in Fig. 12. Excellent agreement between the two circuits is obtained over the frequency range of interest. The high-frequency RHP zero due to C_{gd} is not included in the new equivalent circuit as the frequency of the zero is usually large relative to the unity-gain frequency.

Although the complete equivalent circuit is very accurate, it is complicated, and must be simplified for ease of use in frequency compensation calculations. In SC applications, the OTA is usually a second-stage dominant-pole system with the non-dominant pole at the output of the first stage. Therefore, in the high frequency range between $g_2'/(C_2' + C_{gd})$ and $(g_{m2} + g_2')/C_2'$, the "Miller resistor" (Eqn. (32)) dominates the effect of the impedance Z_1 on the first stage, while the capacitances C_{M1} and C_{s1} can be ignored. For the second stage, R_1' is removed for simplicity without a significant reduction in accuracy. The simplified equivalent circuit for the second-stage dominant-pole system is shown in Fig. 13(a). Perhaps

Fig. 11. A complete small-signal circuit model prior to frequency compensation.
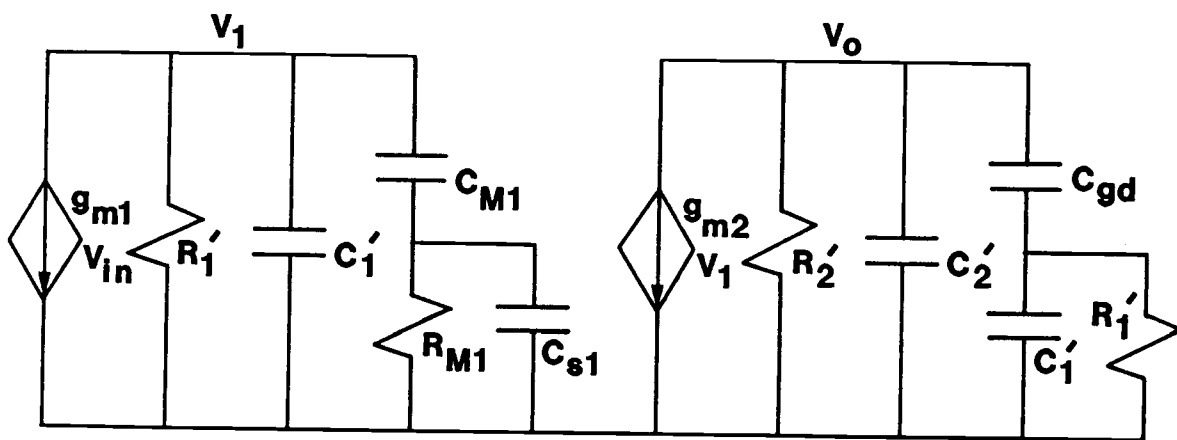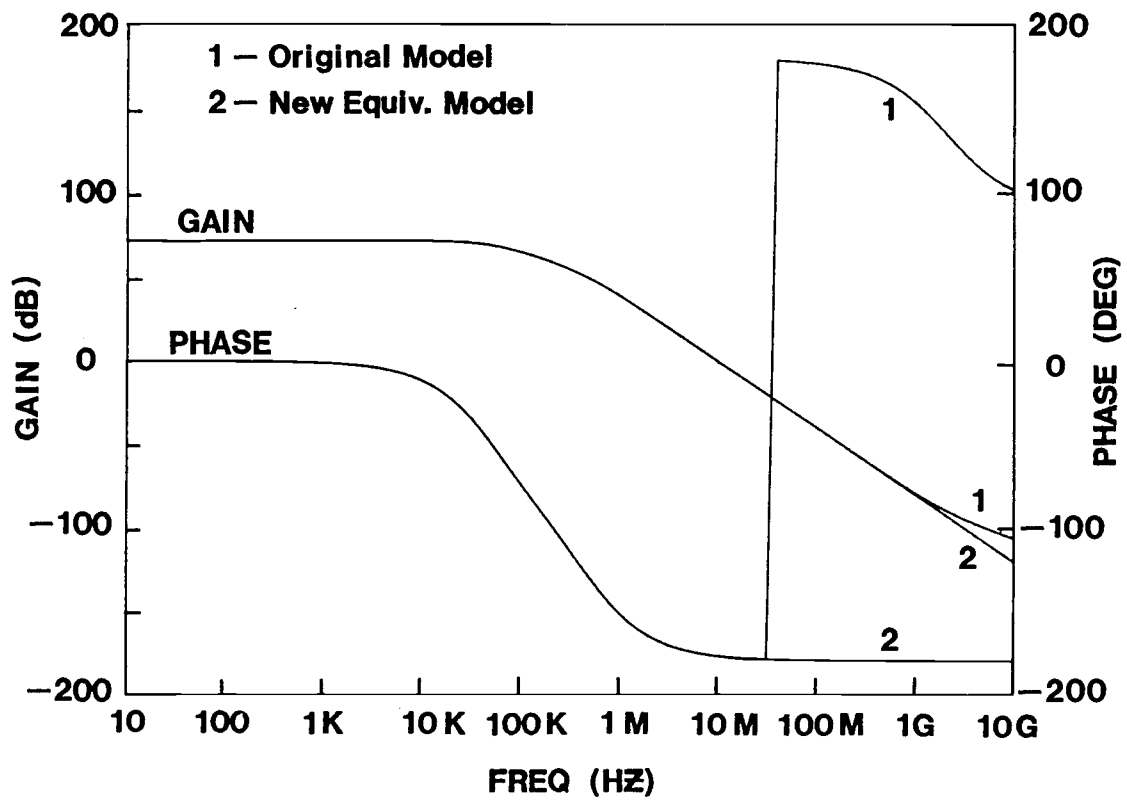
Fig. 12. Comparison of the simulated frequency responses of the complete small signal model of Fig. 11 to the original circuit of Fig. 6(a).
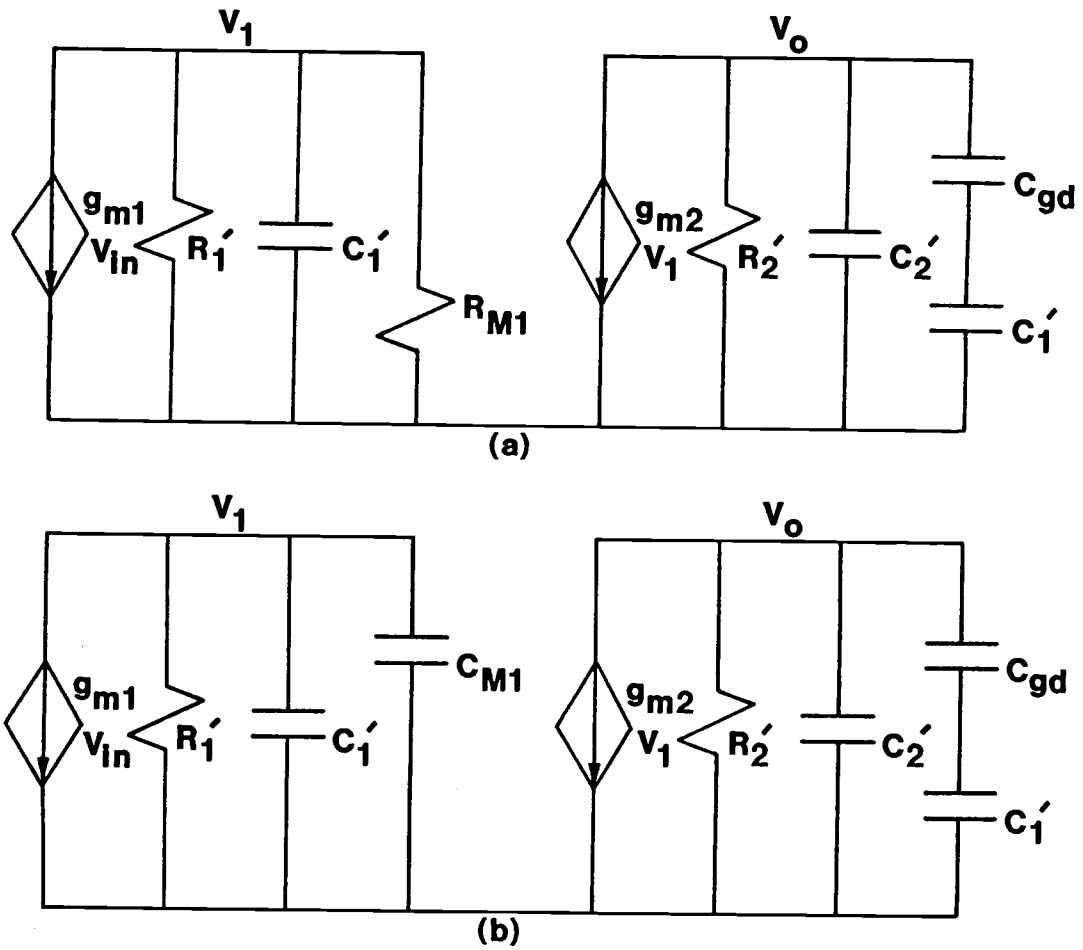
Fig. 13. Simplified small-signal models for (a) the second-stage dominant-pole system using the Miller resistance, and (b) the first-stage dominant-pole system using the Miller capacitance.

surprisingly, it has the same form as the previous simplified model of Fig. 6(b), with the only difference being in how the element values are determined. For the first stage of Fig. 6(b),

$$R_1 = R_1' R_{M1}/(R_1' + R_{M1}), \quad C_1 = C_1' \tag{34}$$

and for the second stage,

$$R_2 = R_2', \quad C_2 = C_2' + C_{gd} C_1'/(C_{gd} + C_1'). \tag{35}$$

This two-pole circuit with new element values is applied to frequency compensation in the same manner as before. In the case of the transconductance amplifier, note that R_1 is reduced significantly from R_1' due to the small Miller resistance, R_{M1} (Eqn. (32)), and C_1 is almost unchanged. Therefore, $R_1 C_1 < R_1' C_1'$, i.e., the first-stage pole moves to higher frequency because of the Miller *resistor* effect, while the second-stage pole moves to a lower frequency.

For most general purpose opamps with output stages, the condition $R_1' C_1' > R_2' C_2'$ is usually satisfied. Therefore, the opamp is a first-stage dominant-pole system prior to compensation. In this case, R_{M1} and C_{s1} in Fig. 11 may be ignored due to their very high impedances at low frequencies, while at the output of the second stage, removing R_1' is usually a valid approximation. With these simplifications, the equivalent circuit for the first-stage dominant-pole system reduces to the conventional Miller capacitor model of Fig. 13(b), [19].

3.3 Applications in Frequency Compensation

The simplified equivalent circuit for the second-stage dominant-pole system (Fig. 13(a)) was derived using Eqns. (34)-(35) and applied to frequency compensation for the MST response. The compensation element values were calculated using Table I, and Eqns. (24)-(25) with $D = 0.01$ as $R_f = 18.86$ kΩ and $C_f = 2.65$ pF. SPICE simulations with

these values for the simplified equivalent circuit and the original circuit of Fig. 6(a) are compared in Fig. 14(a). The agreement is excellent at frequencies near the unity gain crossover frequency where the phase margin, and thus the settling time, is determined. Although the agreement is not as good at low frequencies, this is of no concern in frequency compensation. The simulated unity-gain phase margins were 72.4 and 73.1 degrees for the original and the simplified equivalent models. The small difference was caused by neglecting the high-frequency zero introduced by C_{gd} in Fig. 6(a). The conventional first-stage dominant-pole equivalent circuit (Fig. 13(b)) using the Miller capacitance effect was also applied to the compensation of the transconductance amplifier with R_f = 12.63 kΩ and C_f = 4.97 pF, as calculated previously. The SPICE simulation results of Fig. 14(b) show poor agreement with the original circuit in both the low- and high-frequency ranges.

Finally, the more accurate compensation element values obtained using the improved model, R_f = 18.86 kΩ and C_f = 2.65 pF, were used in a simulation of the CMOS opamp of Fig. 5. The results indicated good cancellation of the pole-zero doublet, and a unity-gain phase margin of 67.6 degrees which is an error of less than four percent relative to the ideal 70 degree phase margin. For comparison, the previous technique yielded a phase margin of 84 degrees giving a 20 percent error. With some minor empirical adjustment, the 70 degree phase margin, and thus the MST response for D = 0.01 is achieved with R_f = 17 kΩ and C_f = 2.9 pF. The simulation results of the step responses for the 70 degree MST and the commonly-used 60 degree phase margin responses are shown in Fig. 15. It is seen that the MST

response, (the first peak of the response touches the upper error limit), reduces the settling time considerably as compared with the 60 degree phase margin case even though a slightly larger compensation capacitance is required to achieve the larger phase margin.
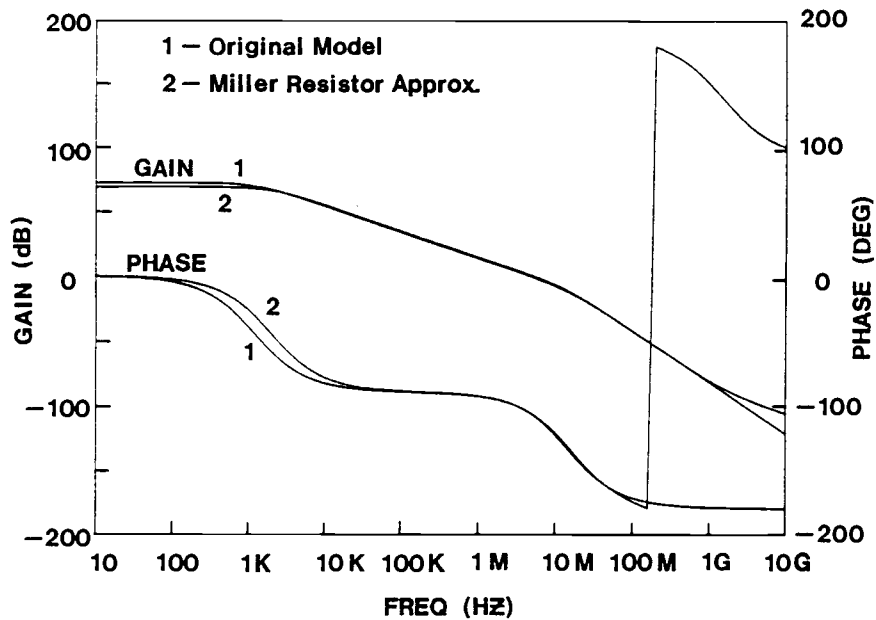
(a)
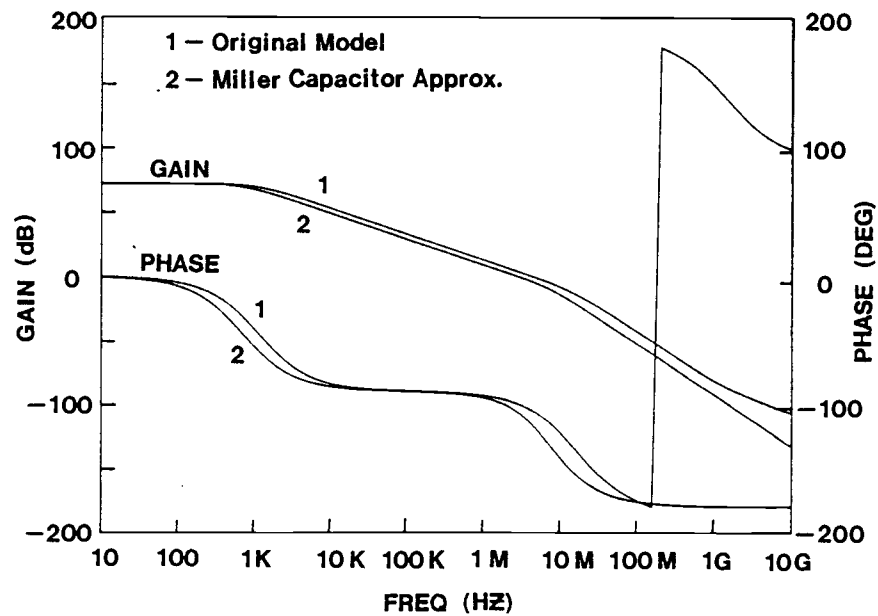
(b)

Fig. 14. Simulated frequency responses of the original circuit of Fig. 6(a) versus (a) the new model using the Miller resistance, and (b) the old model using the Miller capacitance.
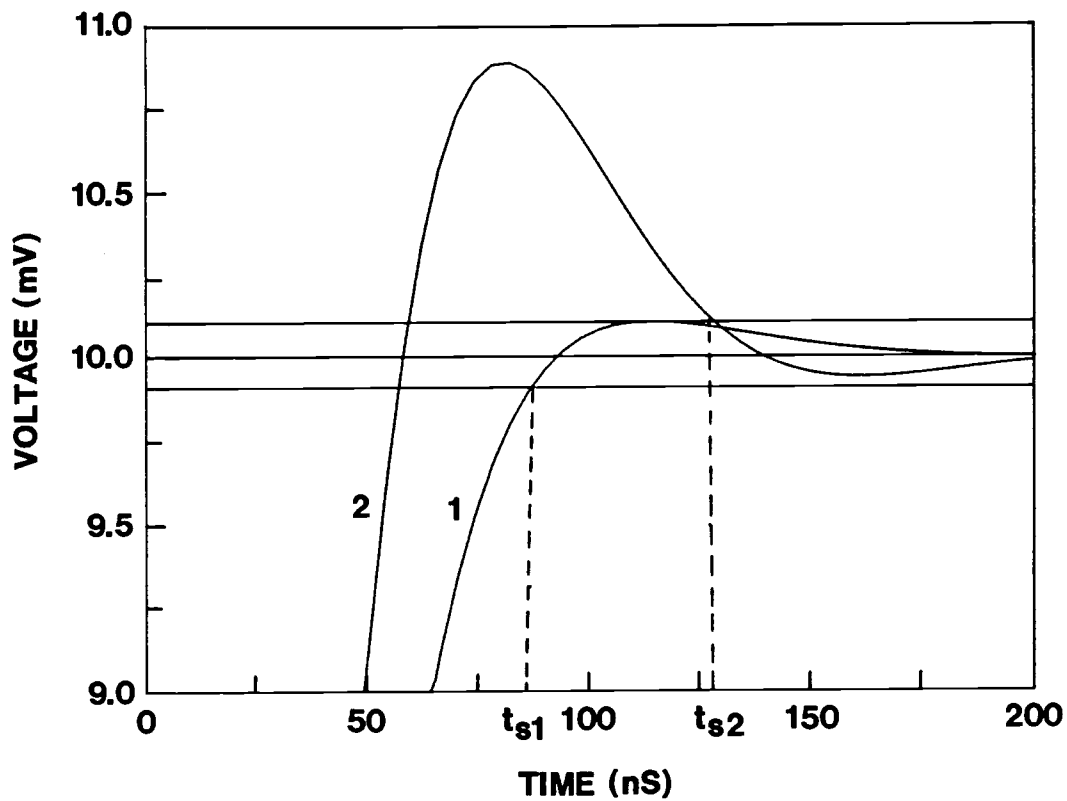
Fig. 15. Simulated step responses of the CMOS opamp for (1) the optimum (approximately 70 degree) cases, and (2) the conventional 60 degree. Note that the optimum design is faster even though the compensation capacitance required to achieve the larger phase margin is somewhat larger.

IV. MST DESIGN FOR ONE-STAGE FOLDED-CASCODE OPAMP

Another widely used CMOS opamp in SC circuits is the one-stage folded-cascode opamp. It provides higher frequency operation and a higher power supply rejection ratio than the conventional cascode and two-stage opamps. An early version of the folded-cascode opamp was analyzed by Ribner and Copeland in 1984 [20]. Figure 1 shows a newer version of the folded cascode opamp which has been used for several years [21], but was first published in 1987 [22]. For switched-capacitor applications, a major advantage of this version over the one analyzed by Ribner and Copeland is that this is a one stage opamp. Therefore, the loading capacitance is used to provide closed-loop stability and no additional compensation capacitance is required as in two-stage opamps.

For first-order analysis, the folded-cascode CMOS opamp is usually approximated as a one-pole opamp [18]. However, in order to design the opamp for the minimum settling time (MST) response, second-order effects must be considered. In this chapter, a two-pole model of the folded-cascode opamp of Fig. 16 is developed based on a complete small-signal analysis. The developed model represents the opamp very accurately, and is easily applied in the design for the MST response. Based on the optimization criterion proposed in Chapter II, the design equation for the MST response of the one-stage folded cascode opamp is derived and verified with SPICE simulations.
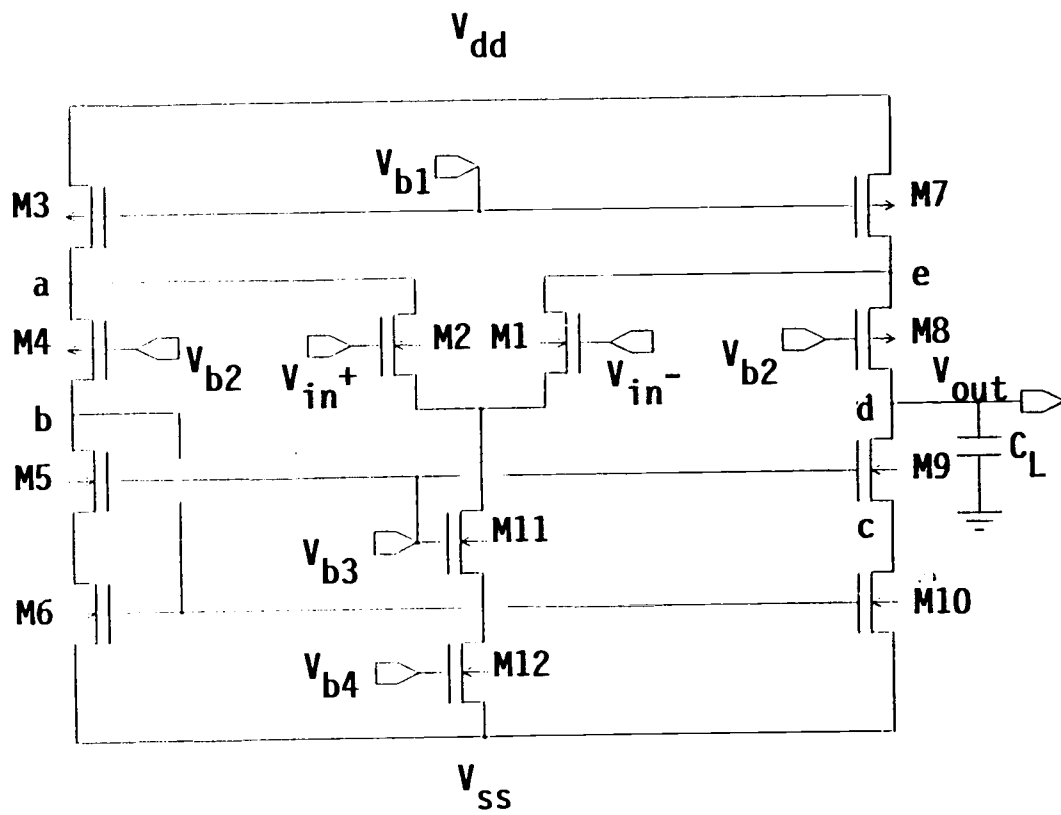
Fig. 16 A one-stage CMOS folded-cascode operational amplifier.

4.1 Small-Signal Modeling

A complete small-signal model for the folded-cascode opamp (Fig. 16) is shown in Fig. 17 [23], where the loading capacitance C_L is included in C_d and V_{out} is associated with node d in Fig. 16. The node equations for this circuit can be written as (neglecting C_{gd10})

$$g_{m1}(V_{in}/2) + (1/R_a + sC_a)V_a + g_{ds4}(V_a - V_b) + G_{m4}V_a = 0 \quad (36)$$

$$(1/R_b + sC_b)V_b + g_{ds4}(V_b - V_a) - G_{m4}V_a = 0 \quad (37)$$

$$g_{m10}V_b + (1/R_c + sC_c)V_c + g_{ds9}(V_c - V_{out}) + G_{m9}V_c = 0 \quad (38)$$

$$sC_dV_{out} + g_{ds9}(V_c - V_{out}) - G_{m9}V_c + g_{ds8}(V_d - V_e) - G_{m8}V_e = 0 \quad (39)$$

$$g_{m1}(-V_{in}/2) + (1/R_e + sC_e)V_e + g_{ds8}(V_e - V_d) + G_{m8}V_e = 0 \quad (40)$$

where $G_m = g_m + g_{mbs}$. In deriving the above equations, the very small C_{gd10} in Fig. 16 is neglected as a good approximation. Due to the fact that $R_bC_b \ll R_cC_c$, C_{gd10} does not act as a Miller capacitance, but rather as a Miller resistance, which is much larger than R_b, at node b. Using a dominant-pole approximation, we derive the transfer function of the small-signal circuit model of Fig. 17 as

$$a(s) = \frac{a_o(1 - s/s_{z1})(1 - s/s_{z2})}{(1 - s/\omega_d)(1 - s/\omega_e)(1 - s/\omega_b)(1 - s/\omega_c)} \quad (41)$$

with

$$a_o = \frac{g_{m1}G_{m8}}{(g_{ds1} + g_{ds7})g_{ds8} + g_{ds9}g_{ds10}(G_{m8}/G_{m9})}, \quad (42)$$

$$\omega_d = -\frac{(g_{ds1} + g_{ds7})g_{ds8}/G_{m8} + g_{ds9}g_{ds10}/G_{m9}}{C_d}, \quad (43)$$
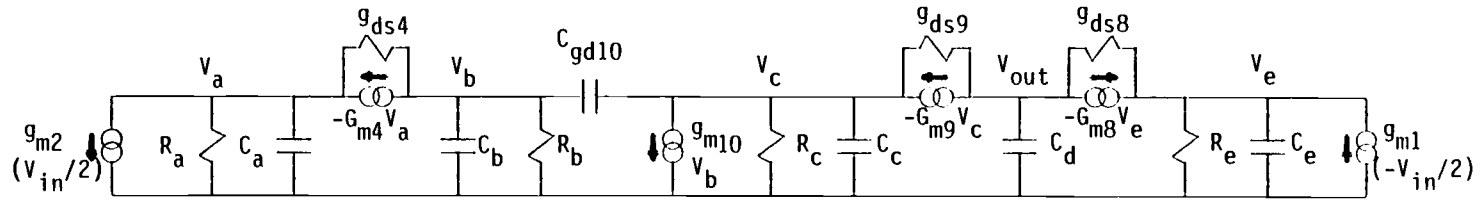
Fig. 17 A complete small-signal model for the opamp in Fig. 16. Nodal

voltages V_a-V_e are referred to the nodes a-e in Fig. 16, and

V_{out} is associated with node d.

$$\omega_e = -\frac{G_{m8}}{C_e} , \qquad (44)$$

$$\omega_b = -\frac{g_{m6}}{C_b} , \qquad (45)$$

$$\omega_c = -\frac{G_{m9}}{C_c} , \qquad (46)$$

and

$$s_{z1}, \ s_{z2} = (\omega_b + \omega_c)/2 \pm \sqrt{(\omega_b^2 + \omega_c^2)/4 - 3\omega_b\omega_c/2} \qquad (47)$$

where a_o is the DC gain of the opamp. Eqn. (40) shows that the transfer function of the folded-cascode opamp has four left-half-plane (LHP) poles and two LHP zeros. The dominant pole (Eqn. (43)) of the opamp is at the output node d due to the large values of loading capacitance and output impedance. The frequencies of the three non-dominant poles (Eqns. (44)-(46)) are of the same order. It can be shown that the two LHP zeros (Eqn. (12)) are usually a conjugate complex pair, where the real part of the frequency, $Re(s_{z1}, \ s_{z2}) = (\omega_b + \omega_c)/2$, is the average of ω_b and ω_c. For a first order approximation, $Re(s_{z1}, \ s_{z2})$ can be considered to be cancelled with the two poles ω_b and ω_c. Thus, the transfer function of Eqn. (40) is simplified to a two-pole model with

$$a(s) \simeq \frac{a_o}{(1 - s/\omega_d)(1 - s/\omega_e)} \qquad (48)$$

Figure 18 shows the two-pole small-signal circuit model of the folded-cascode opamp in which

$$g_{mI} = g_{m1}, \qquad (49a)$$

$$C_I = C_e, \qquad (49b)$$

$$R_I = 1/G_{m8}, \qquad (49c)$$
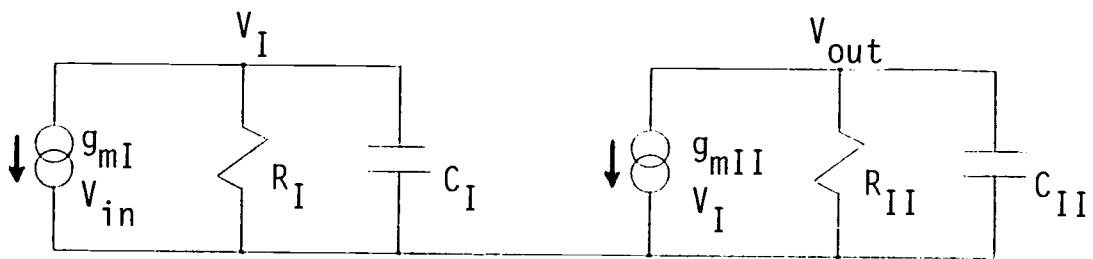
Fig. 18 A simplified two-pole small-signal model for the folded-cascode opamp.

and
$$g_{mII} = G_{m8}, \tag{50a}$$

$$C_{II} = C_d, \tag{50b}$$

$$R_{II} = \frac{1}{(g_{ds1} + g_{ds7})g_{ds8}/G_{m8} + g_{ds9}g_{ds10}/G_{m9}.} \tag{50c}$$

The dominant pole of this circuit is always associated with the second stage. Figure 18 shows a comparison of the SPICE simulated frequency responses of a folded-cascode CMOS opamp and its two pole model. The two-pole model is accurate up to the unity-gain frequency. Although the phase responses of the model and the opamp disagree slightly at the unity-gain frequency due to the inexact pole-zero cancellation, they are still in fairly good agreement and can be used to design the MST response for the opamp.

4.2 Opamp Design for MST Response

In Chapter II, we have shown that the minimum small-signal settling time response of a two-pole system is obtained when the response to a step input is underdamped so that the first peak just touches the upper error bound. Equation (15) shows that a well-defined pole separation factor provides the MST response for the two-pole opamp; This pole separation value for the folded-cascode opamp is:

$$\beta_0 \equiv \frac{\omega_e}{\omega_d} \simeq \frac{4 \, (a_0 + 1)}{1 + (\pi/\ln D)^2} \tag{51}$$

Using Eqn. (51) with Eqns. (7)-(9) and assuming $C_d \simeq C_L$, we obtain

$$C_L[1 + (\pi/\ln D)^2] - 4(g_{m1}/G_{m8})C_e = 0 \tag{52}$$

Equation (52) can be used either to determine the optimum loading capacitance for a given opamp, or as an opamp design equation for a given C_L.

A folded-cascode opamp was designed using Eqn. (52) with C_L = 5 pF and D = 0.01 with the device sizes listed in Table II. The frequency responses of the opamp and its two-pole model are shown in Fig. 19. Figure 20 shows the simulated MST response (C_L = 5 pF) compared with a more underdamped, yet wider bandwidth response (C_L = 4.5 pF) and a more overdamped, yet narrower bandwidth response (C_L = 5.5 pF). Interestingly, longer settling times are observed when C_L is either larger or smaller than 5 pF even though the unity-gain bandwidth increases for smaller C_L values. (Similar behavior has been reported for SPICE simulations of one-stage GaAs opamps [24].) The simulated phase margin for the MST response of this opamp is about 67 degrees which is slightly smaller than the MST phase margin predicted in Chapter II for the ideal two-pole opamps (around 70 degrees). This is because the third and fourth poles of the folded-cascode opamp are not exactly cancelled by the complex zeros, which results some additional phase shift. The small-signal settling time of the opamp versus the loading capacitance is plotted in Fig. 21. It is shown that the settling time is a very strong function of C_L, and the minimum settling time is obtained at 5 pF as designed. Similar to Fig. 3, the discontinuities in response time of Fig. 21 are associated with the natural frequency of the system, and occur when response peaks or valleys just enter the bounded region for increasing loading capacitance.

Table II

M1	120/6	M7	460/6
M2	120/6	M8	332/6
M3	460/6	M9	126/4
M4	332/6	M10	126/4
M5	126/4	M11	240/6
M6	126/4	M12	240/6

Table II. W/L ratios (in microns) of a one-stage folded-cascode opamp for MST response with C_L = 5 pF.
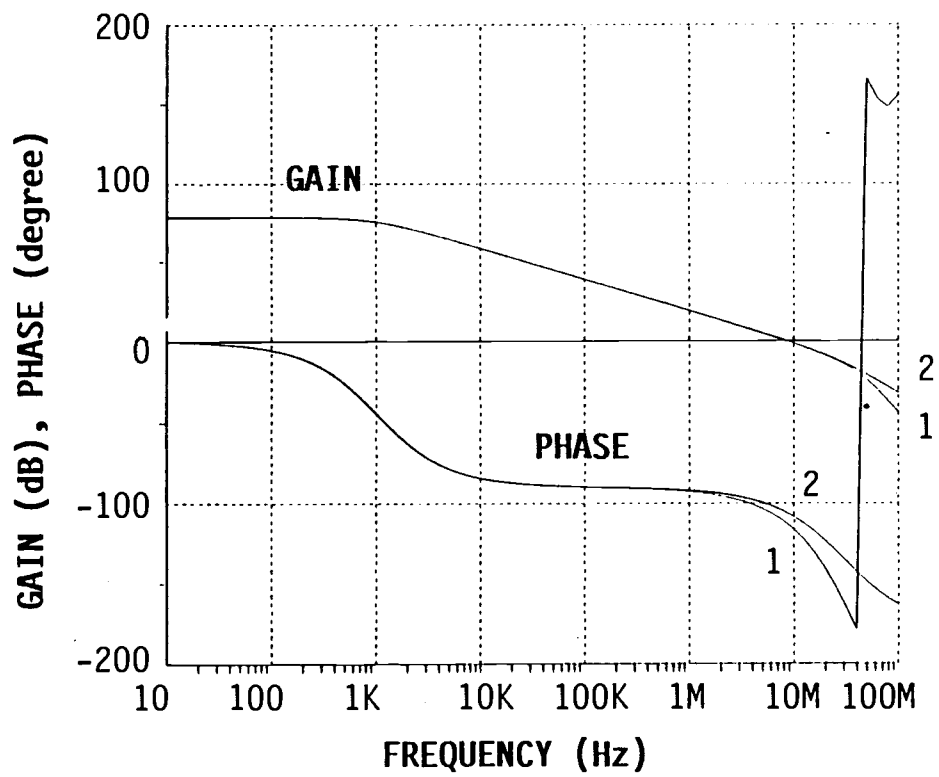
Fig. 19 SPICE simulated frequency responses for the original opamp of Fig. 16 (curve 1) and its two-pole model of Fig. 18 (curve 2).
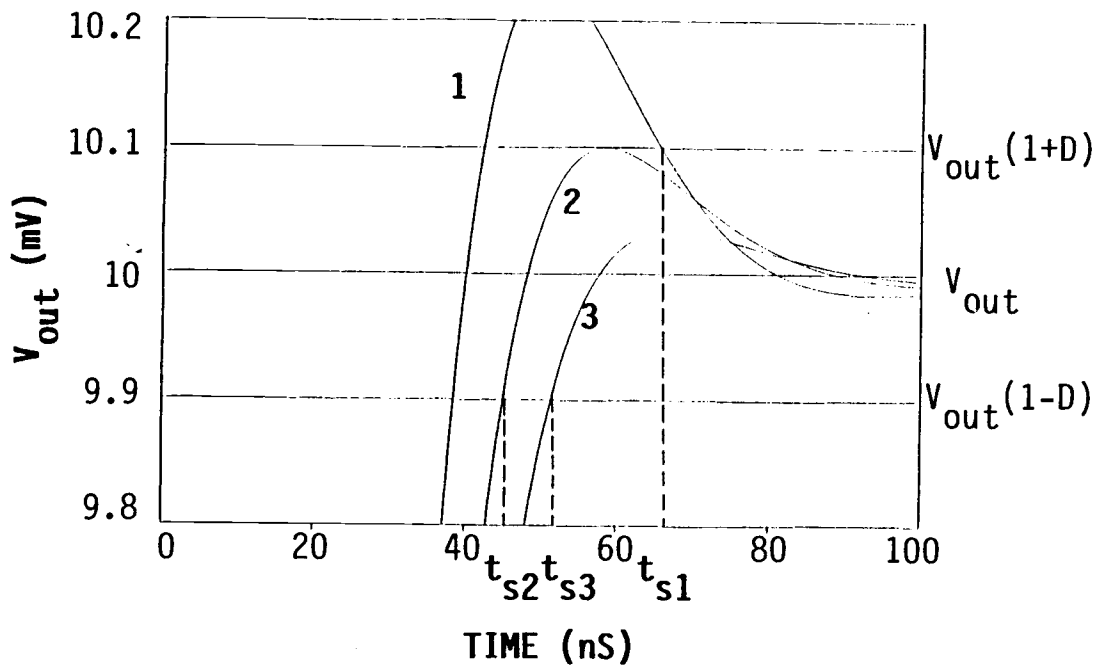
Fig. 20 Simulated MST response of the opamp with C_L = 5 pF (Curve 2) compared with a more underdamped response with C_L = 4.5 pF (Curve 1) and a more overdamped response with C_L = 5.5 pF. t_s indicates the settling time.
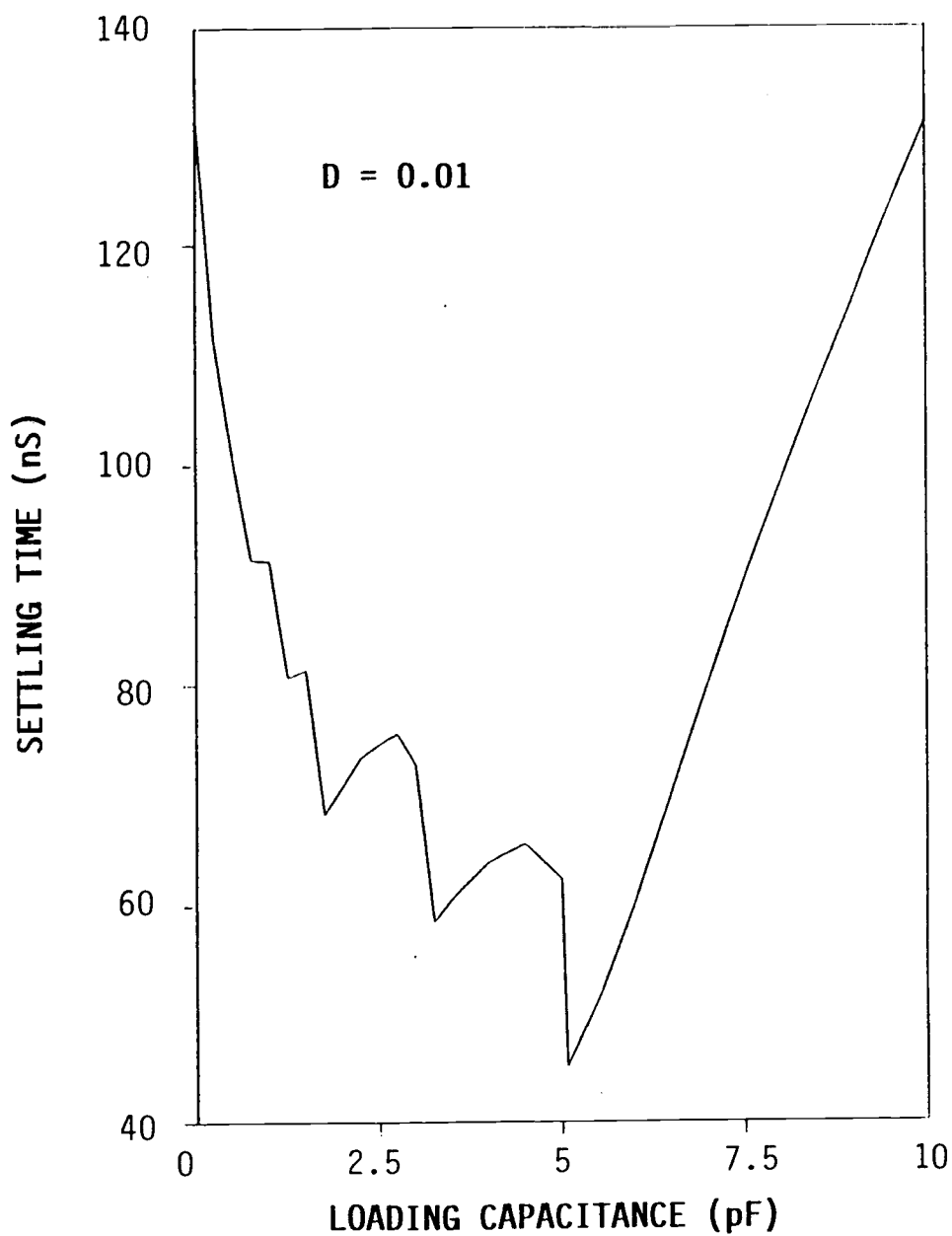
Fig. 21 Simulated small-signal settling time versus loading
capacitance for the one-stage folded-cascode opamp.

V. DISCUSSION AND CONCLUSIONS

A frequency compensation technique for fast-settling opamps has been presented, and an improved equivalent circuit model for a two-stage operational amplifier was also proposed. The results show that for a first-stage dominant-pole system, the Miller capacitance approximation is appropriate whereas for the second-stage dominant-pole system, the Miller resistance approximation is required. This improved modeling not only provides for much more accurate compensation, but also improves the understanding of pole-splitting compensation. The minimum small-signal settling time analysis is also applicable to the optimization of one-stage two-pole opamps as presented in Chapter IV.

One potential concern relates to the sensitivity of the MST response to process variations. Variations in phase margin were simulated with standard process variations about both the MST and nominal 60 degree cases for the two-stage opamp. The compensation capacitance was varied by ±10 percent, while the g_m's and the compensation resistance were varied by ±25 percent. As shown in Table III(a), the MST exhibited nearly identical sensitivity to the commonly-used 60 degree case. As shown in Table III(b), the sensitivities were also simulated for ±25 percent variations in R_1' and R_2', and ±10 percent variations in C_1' and C_2'. Again, the sensitivities were similar. Obviously, we cannot compensate the opamp exactly at the MST point over processing variations. The key point here is that if we select the MST as the nominal condition, then including process variations, the range of settling times is always less than the range obtained if we select 60 degrees as the nominal condition. The improvement is typically a factor of two. This argument

obviously applies for temperature variations as well. A similar sensitivity to process variation was also observed in the folded-cascode opamp.

We are somewhat justified in neglecting slew-rate effects in many applications. Using conventional two-stage or cascode opamps, slew-rate effects may be ignored for cases wherein the maximum output voltage change during any sampling period is less than about 500 mV. For high frequency CMOS SC filters, either class-AB or adaptively-biased [25]-[30] opamps may be used as they do not exhibit slew-rate limiting, and therefore small-signal analysis is applicable. The slew-rate limitation on the MST design is suggested for future study. Another suggestion for future work is the applications of the MST design techniques for SC filters.

TABLE III(a)

	$0.75C_f$	$1.0C_f$	$1.1C_f$
$0.75g_m$	65.7 (52.4)	68.6 (59.4)	70.8 (56.4)
$1.00g_m$	65.9 (56.2)	70.0 (60.0)	73.6 (62.7)
$1.25g_m$	61.7 (54.2)	65.7 (58.0)	69.7 (61.4)

TABLE III(b)

	$0.9 \, C'$	$1.1 \, C'$
$0.75 \, R'$	74.6 (64.0)	67.5 (57.4)
$1.25 \, R'$	73.7 (62.8)	66.9 (56.5)

Table III. Comparison of phase margins (in degrees) versus typical process variations for the MST compensation versus the 60-degree compensation (numbers in parentheses) for the circuit of Fig. 6(a) with R_f-C_f compensation elements.

VI. BIBLIOGRAPHY

1. J.L. Chalfan and J.C. Looney "Linear MOS integrated circuits," Solid-State Technology, pp. 31-39, May 1969.

2. P.R. Gray, D.A. Hodges and R.W. Brodersen (ed), *Analog MOS Integrated Circuits*, IEEE Press, 1980.

3. G.S. Moschytz (ed), *MOS Switched-Capacitor Filters: Analysis and Design*, IEEE Press, 1984

4. P.R. Gray, B.A. Wooley and R.W. Brodersen (ed), *Analog MOS Integrated Circuits (II)*, IEEE Press, 1989.

5. P.R. Gray and R.G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed., New York: Wiley, 1982.

6. S. Rosenstark, *Feedback Amplifier Principles*, New York: Macmillan, 1986.

7. B.Y. Kamath, R.G. Meyer and P.R. Gray, "Relationship between frequency response and settling time of operational amplifiers," *IEEE J. Solid-State Circuits*, vol. SC-9, pp. 347-352, Dec. 1974.

8. R.J. Apfel and P.R. Gray, "A fast-settling monolithic operational amplifier using doublet compression techniques," *IEEE J. Solid-State Circuits*, vol. SC-9, pp. 332-340, Dec. 1974.

9. C.T. Chuang, "Analysis of the settling behavior of an operational amplifier," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 74-80, Feb. 1982.

10. J.C. Lin and J.H. Nevin, "A modified time-domain model for nonlinear analysis of an operational amplifier," *IEEE J. Solid-State Circuits*, vol. SC-21, pp. 478-483, June 1986.

11. J.E. Solomon, "The monolithic op-amp: A tutorial study," *IEEE J. Solid-State Circuits*, vol. SC-9, pp. 314-332, Dec. 1974.

12. P.R. Gray and R.G. Meyer, "MOS operational amplifier design--A tutorial overview," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 969-982, Dec. 1982.

13. S.D. Jen, H.C. Yang and D.J. Allstot, "Operational amplifier compensation using doublet-decompression for switched-capacitor circuits," in *Proc. of the 31st IEEE Midwest Symposium on Circuits and Systems*, pp. 4-6, Aug. 1988.

14. H.C. Yang, J.R. Ireland and D.J. Allstot, "Improved operational amplifier compensation techniques for high-frequency switched-capacitor circuits," in *Proc. of the 30th IEEE Midwest Symposium on Circuits and Systems*, pp. 952-955, Aug. 1987.

15.   H.C. Yang and D.J. Allstot, "Considerations for fast-settling operational amplifiers," *IEEE Tran. Circuits and Systems,* March 1990 (to appear).

16.   R. Gregorian and G.C. Temes, *Analog MOS Integrated Circuits for Signal Processing,* New York: Wiley, 1986.

17.   W.C. Black, Jr., D.J. Allstot and R.A. Reed, "A high performance low power CMOS channel filter," *IEEE J. Solid-State Circuits,* vol. SC-15, pp. 929-938, Dec. 1980.

18.   P.E. Allen and D.R. Holberg, *CMOS Analog Circuit Design,* New York: Holt, Rinehart and Winston, 1987.

19.   H.C. Yang and D.J. Allstot, "An equivalent circuit model for two-stage operational amplifiers," in *Proc. of IEEE International Symposium on Circuits and Systems,* pp. 635-638, June 1988.

20.   D.B. Ribner and M.A. Copeland, "Design techniques for cascoded CMOS Op Amps with improved PSRR and common-mode input range," *IEEE J. Solid-State Circuits,* vol. SC-19, no. 6, pp. 919-925, Dec. 1984.

21.   D.J. Allstot, *Class Lecture Notes for EECS 240,* University of California at Berkeley, Fall Semester, 1985.

22.   J.N. Babanezhad and R. Gregorian, "A programmable gain/loss circuit," *IEEE J. Solid-State Circuits,* vol. SC-22, pp. 1082-1090, Dec. 1987.

23.   H.C. Yang, M.A. Abu-Dayeh and D.J. Allstot, "Analysis and Design of a fast-settling folded-cascode CMOS operational amplifier for switched-capacitor applications," in *Proc. of the 32nd IEEE Midwest Symp. Circuits and Systems,* PP. 4-6, Aug. 1989.

24.   C. Toumazou and D.G. Haigh, "Some designs and a characterization method for GaAs operational amplifiers for switched capacitor applications," *Electronics Letters,* vol. 24, no. 18, pp. 1170-1172, Sept. 1988.

25.   R. Castello and P.R. Gray, "A high-performance micropower switched-capacitor filter," *IEEE J. Solid-State Circuits,* vol. SC-20, pp. 1122-1132, Dec. 1985.

26.   S.H. Lewis and P.R. Gray, "Pipelined 5 ms/s 9-bit analog-to-digital converter," *IEEE J. Solid-State Circuits,* vol. SC-22, pp. 954-961, Dec. 1987.

27.   C. Yu, J.J. Yang and D.J. Allstot, "A high-swing class-AB CMOS operational amplifier," in *Proc. of the IEEE International Symposium on Circuits and Systems,* pp. 151-154, June 1988.

28. M.G. Degrauwe, J. Rijmenants, E.A. Vittoz and H.J. De Man, "Adaptive biasing CMOS amplifiers," *IEEE J. Solid State Circuits*, vol. SC-17, pp. 522-528, June 1982.

29. T.S. Fiez and D.J. Allstot, "A high-swing adaptively-biased CMOS amplifier," in *Proc of the 31st IEEE Midwest Symposium on Circuits and Systems*, pp. 852-854, Aug. 1988.

30. T.S. Fiez, H.C. Yang, J.J. Yang, C. Yu and D.J. Allstot, "A family of high swing CMOS operational amplifiers," *IEEE J. Solid-State Circuits*, Dec. 1989.