A Balanced 0.9-μm CMOS Transconductance-C Filter Tunable Over the VHF Range

W. Martin Snelgrove, *Member, IEEE,* and Ayal Shoval, *Student Member, IEEE*

***Abstract*—A balanced transconductance-C biquad implemented in the digital subset of a 0.9-μm CMOS process operates at frequencies up to 450 MHz and Q factors from a nominal value near 1 to approximately 100 with 30–40-dB dynamic range. By switching in capacitors and adjusting control voltages it can be tuned to below 30 MHz, demonstrating the capability of operating over the entire VHF range. Active area is 0.029 mm^2 and power consumption is 8–12 mW with a 5-V power supply.**

I. Introduction

SEVERAL competing circuit technologies are available for implementing monolithic filters: active-RC [1], MOSFET-C [2], switched-C [3], [4], transconductance-C [5], [6], and digital [7]. The different characteristics of each defines its applications niches, and in new technologies it is generally appropriate to test the capabilities of the circuit styles in their areas of natural strength. We have developed and tested a chip to examine the high-frequency limits of transconductance-C (G_m–C) filters in submicrometer CMOS.

Because transconductance-C filters use an integrator built from an open-loop transconductance amplifier driving a capacitive load they can be very fast (no compensation capacitors are needed) but should not be expected to be very linear. Linearizing circuits exist, but provide limited improvement in linearity and often compromise speed. At high speeds they tend to store state variables on small capacitors and therefore suffer large kT/C noise: this combines with their nonlinearity to limit their useful dynamic range.

The most natural applications for these filters are therefore those where speed is vital and low or moderate SNR acceptable, for example, pulse-shaping and equalization for very high-speed data communications. For these reasons we chose to investigate the capabilities of G_m–C circuits in modern CMOS by emphasizing speed. Linearity, power consumption, and area were considered only insofar as they did not seriously compromise high-frequency performance, but were all quite reasonable in the final design.

Placing the highest priority on speed implies use of minimum-length devices run at high gate-source voltages, and selects for simple circuits to eliminate parasitic poles. The single low supply voltage (5 V, with a strong interest in going down to 3.3 V) further discourages complex circuits, since even cascodes quickly consume all available bias headroom. In this sense the requirement of high operating frequency tends automatically to produce small areas.

In the following section, the versatility of G_m–C technology in implementing general-purpose filters is reviewed. In Section III, the transconductor that was implemented is discussed and its limitations are highlighted. In Section IV the biquad loop, which was chosen as a structure to demonstrate the capabilities of the technology, is described. Experimental results are presented in Section V.

II. State-Space Structures for Transconductance-C Filters

The basic building block we use is a differential transconductor loaded by a grounded capacitor. Summing the output currents of several transconductors $G_{m,i}$ with different inputs v_i onto a single capacitor produces an output described by

$$C\dot{v}_c = G_{m,1}v_1 + G_{m,2}v_2 + \cdots .$$

This is all that is needed to produce an arbitrary transfer function, except high-pass functions where there is finite gain in the limit as frequency goes to infinity. Interconnecting N capacitors with transconductors allows direct implementation of any mathematical system of the form

$$\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}u \tag{1}$$

where the states x_i are represented by capacitor voltages, matrix element A_{ij} is implemented by a transconductor with input x_j and output x_i, and b_i is implemented by a transconductor from the overall input u to state i. Arbitrary matrix values are obtained by scaling transconductors ($G_{m,ij} = A_{ij}/C_i$) and the freedom to exchange inputs of a differential circuit allows arbitrary signs. An important special case is the two-integrator loop: the circuit implemented in this experiment, for instance, was a frequency-scaled version of

$$\boldsymbol{A} = \begin{pmatrix} -0.5 & 1 \\ -1 & -0.5 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2}$$

Manuscript received August 6, 1991; revised November 11, 1991. This work was supported by the Canadian federal government through NSERC operating grants and the Micronet program.

The authors are with the Department of Electrical Engineering, University of Toronto, Toronto, Ont., Canada M5S 1A4.

IEEE Log Number 9105501.

for which the transfer functions from the input to the two outputs (the states) can be computed from the general formula

$$\boldsymbol{X}/U = (s\boldsymbol{I} - \boldsymbol{A})^{-1}\boldsymbol{b} = \frac{1}{(s + 0.5)^2 + 1}\begin{pmatrix} s + 0.5 & 1 \\ -1 & s + 0.5 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{s^2 + s + 1.25}\begin{pmatrix} s + 0.5 \\ -1 \end{pmatrix}. \quad (3)$$

One state implements a low-pass filter and the other a function with a real-axis zero. By adding a second input, so that $\boldsymbol{b} = (1 \quad -0.5)^T$, a true bandpass can be obtained.

Other important structures [8]–[10] simulate various types of *LC* ladders, so as to obtain desirable sensitivity and dynamic range properties. In fact, the system described by (2) simulates a doubly terminated low-pass ladder filter with $N = 2$. This is the design style we see as the natural extension of this work to higher order. Many of these circuits can also be seen as gyrator-*C* circuits [6], because a back-to-back pair of transconductors is a gyrator, but the state-space formulation is more general.

At low frequencies, the preferred implementation of a high-order filter, for flexibility and simplicity of tuning, is often a cascade of second-order sections (biquads). Unfortunately, a second-order notch circuit has

$$T(s) = \frac{s^2 + \omega_z^2}{s^2 + s\frac{\omega_0}{Q} + \omega_0^2}$$

which cannot be realized in the form of (1) because it has a high-pass component. Notch biquads solve this problem by using floating capacitors [11], [12], which we would like to avoid (see Section IV-A). They also typically have low and reactive input impedance and so may need interstage voltage buffers, which are relatively slow.

Transconductance-*C* filters have to be tuned during operation because their time constants vary widely with process, bias, and environment. Bias dependencies are generally used to adjust transconductance, and at the system level, tuning is done either by a master–slave scheme [6] or using adaptive filtering [13], [14]. The fact that this tuning hardware is needed anyway suggests that a natural niche for the technology is one where adaptation is required, again pointing to data communications.

Adaptive technology is very robust, because the filter is inside a control loop. A Q-factor error from finite g_o, for instance, will simply be tuned out. This means that the signal-path filter can be designed more for speed than for accuracy.

Often it is possible to tune elements of $\boldsymbol{A}$ and $\boldsymbol{b}$ independently, but that gives $N^2 + N$ degrees of freedom—far more than necessary to define a transfer function of order N. The circuit we implemented is more constrained, and only two degrees of freedom were used, to set frequency and Q. The main additional constraints come from internal signal levels.

III. Linearized Transconductors

A. The Single-Ended Transconductor

Fig. 1 shows the use of a simple differential pair as a transconductor. Well-known circuits for large-signal linear transconductors exist [15]–[19]. Fig. 2 shows the circuit we use, which is simply a differential pair without the current source and is large-signal linear under the same general conditions as the others. A similar block was recently reported [20]. The linearized circuits operate by canceling the quadratic components of the drain currents of two devices, and are exactly linear when devices are exactly described by the long-channel pinch-off model

$$i_D = \frac{K'}{2}\frac{W}{L}(v_{GS} - V_T)^2 \quad (4)$$

where $K' = \mu_0 C_{ox}$. Thus, if the inputs consist of a differential signal superimposed on a common-mode bias voltage, $V_{CM} \pm v_d/2$, and $M3$ and $M4$ implement an ideal current mirror, then

$$i_0 = \frac{K_1'}{2}\frac{W_1}{L_1}\left(V_{CM} + \frac{v_d}{2} - V_T\right)^2 - \frac{K_2'}{2}\frac{W_2}{L_2}\left(V_{CM} - \frac{v_d}{2} - V_T\right)^2 \quad (5)$$

and when the two transistors are identical ($K_1' = K_2'$, $W_1 = W_2$, $L_1 = L_2$)

$$i_0 = K'\frac{W}{L}(V_{CM} - V_T)v_d \quad (6)$$

which is linear in the signal voltage v_d.

B. Short-Channel Effects

Since we are optimizing for speed, and hence using short-channel devices (0.9 μm in this technology), (4) is not a good model. A simple correction term for transverse-field degradation of mobility [21] gives the equation

$$i_D = \frac{K'}{2}\frac{W}{L}\frac{(v_{GS} - V_T)^2}{1 + \theta(v_{GS} - V_T)}. \quad (7)$$

A stronger effect for minimum L comes from velocity saturation for which pinch-off models are quite complex [22]. A simple first-order approximation uses the θ term in (7) to model this too. The form is plausible because it correctly predicts that short-channel devices are nearly linear at high v_{GS}.

When transistors are modeled by (7), for $\theta(v_{GS} - V_T) \ll 1$, (6) still applies, while for $\theta(v_{GS} - V_T) \gg 1$, the transistors are essentially linear transconductors already, and there are no quadratic components to cancel: the result is still large-signal linear, with

$$i_0 = \frac{K'}{2}\frac{W}{L\theta}v_d. \quad (8)$$

Between these two extremes, however, the devices will produce output currents with odd-order terms not canceled by the differential configuration and will not be ex-

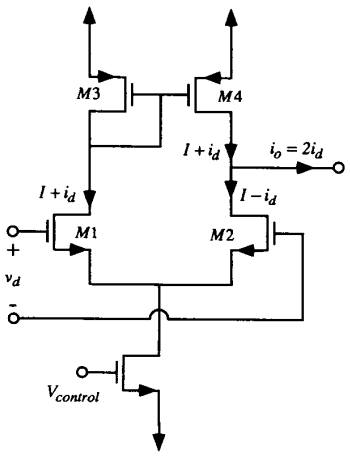


Fig. 1. A differential pair as a transconductor.

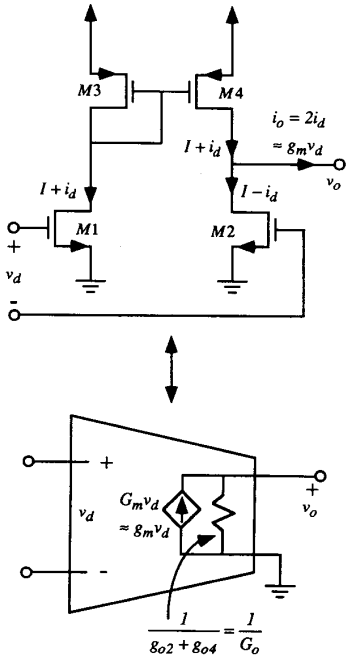


Fig. 2. The single-ended transconductor.

actly linear. For this reason there is an inherent trade-off between linearity, best at long channels, and speed, best with minimum channels. Our test chip is an experiment with one extreme of the resulting range.

Note that even using long channels will not make these circuits perfectly linear in a modern process: the transverse-field term is set by oxide thickness and depends on process, not L.

One way to estimate the cubic distortion in the output current is to represent (7) in the form of a Taylor series about the bias point. For the input transistor $M1$ of the differential pair in Fig. 2, (7) gives

$$i_1 = \frac{K_1'}{2}\frac{W_1}{L_1}\frac{(V_{ON}^2 + V_{ON}v_d + (v_d/2)^2)}{1 + \theta(V_{ON} + v_d/2)} \tag{9}$$

where $V_{ON} = V_{CM} - V_T$. Rewriting (9) in the form

$$i_1 = \frac{K_1'}{2}\frac{W_1}{L_1}\frac{1}{1+\theta V_{ON}} \cdot \frac{(V_{ON}^2 + V_{ON}v_d + (v_d/2)^2)}{1 + \frac{\theta v_d/2}{1+\theta V_{ON}}}$$

allows a Taylor expansion to be made and, with $\Theta = \theta/(1+\theta V_{ON})$, results in

$$i_1 = \frac{K_1'}{2}\frac{W_1/L_1}{1+\theta V_{ON}}\left(V_{ON}^2 + V_{ON}v_d + \left(\frac{v_d}{2}\right)^2\right) \cdot \left[1 - \Theta\frac{v_d}{2} + \left(\Theta\frac{v_d}{2}\right)^2 - \left(\Theta\frac{v_d}{2}\right)^3 + \cdots\right].$$

For the second transistor a similar expression is obtained but with v_d replaced with $-v_d$. Assuming matching, even-order Taylor terms from $M1$ and $M2$ cancel. To estimate third harmonic distortion, let $v_d = V_{pd}\cos\omega t$ and solve for the output current

$$i_0 = \frac{V_{ON}K'W/L}{1+\theta V_{ON}}\left[\left(1 - \frac{\Theta}{2}V_{ON}\right)V_{pd}\cos\omega t + \frac{\Theta V_{pd}^3}{32} \cdot \left(2\Theta - \Theta^2 V_{ON} - \frac{1}{V_{ON}}\right)(3\cos\omega t + \cos 3\omega t)\right].$$

Third harmonic distortion is therefore

$$\mathrm{HD}_3 \approx \left|\frac{\Theta V_{pd}^2\left(2\Theta - \Theta^2 V_{ON} - \frac{1}{V_{ON}}\right)}{32\left(1 - \frac{\Theta}{2}V_{ON}\right)}\right|. \tag{10}$$

For a short-channel device with $\theta = 0.6\ \mathrm{V}^{-1}$, biased at a high on-voltage V_{ON} of 1.8 V, this predicts about 0.16% distortion for a 1-V signal level. For comparison, a long device in the same technology might have $\theta = 0.1\ \mathrm{V}^{-1}$ and would still show 0.11% distortion at the same signal level. This is a very simplified model of short-channel distortion, and useful primarily for a qualitative understanding at low signal levels.

C. Common-Mode Gain and Signal Swing

In the presence of finite output conductance G_o, the circuit of Fig. 2 has nonzero common-mode gain, resulting from the asymmetry between connections of $M3$ and $M4$. A matching argument can be used to show that this produces a *voltage* gain of approximately -1. Transistors $M1$ and $M3$ form a ratioed inverter which was designed to have a voltage gain of about -1; for a purely common-mode signal in, and with like transistors matched and having finite output conductance, the $M2$, $M4$ pair will operate at exactly the same gate–source and drain–source voltages as the $M1$, $M3$ pair when the drain of $M2$ (which is the output) tracks the drain of $M1$ (the ratioed inverter).

This common-mode gain is expressed as a voltage gain: the common-mode rejection ratio should be expressed as a ratio of transconductances, and is G_m/G_o. Because of the use of short-channel devices, this is low: about 20 dB.

The transconductor is designed to operate with all devices in pinch-off. $M1$ or $M2$ will enter triode if their drain voltages drop more than one threshold voltage below their gates, which sets an upper limit of about 1 V on differential signal levels. $M1$ and $M2$ can also be driven into cutoff by signals larger than $2(V_{CM} - V_T)$, and this mechanism will dominate for low V_{CM} (which will be needed when tuning for low transconductances, which in turn correspond to low frequencies, low input gains, or high Q factors). $M4$ can enter triode for high common-mode input and output.

The complex relationships between output swing and control voltage severely limit the useful tuning range of this transconductor. Many of these problems are common to all CMOS transconductors, but in the short-channel case the transverse-field and velocity saturation effects further reduce the range of transconductance available from a given range of tuning voltage $(V_{CM} - V_T)$. Note that in the limiting case, (8), no adjustment of transconductance is possible. Wide tuning range can only be practically obtained by augmenting the individual transconductor's poor tuning range with some other mechanism, such as by switching-in capacitors or by partially canceling the outputs of two transconductors.

D. The Balanced Transconductor

A balanced-output transconductor can be obtained from a pair of single-ended circuits each producing one of the two outputs as shown in Fig. 3. This kind of arrangement has to be analyzed for the effects of mismatch between the two half-circuits.

If the two transconductors do not match, and have outputs

$$i_1 = G_{m,1}v_{id} + G_{m,CM,1}v_{iCM}$$
$$i_2 = -G_{m,2}v_{id} + G_{m,CM,2}v_{iCM}$$

and these currents drive capacitors C_1 and C_2 that are nominally $2C$ but slightly different, then the overall integrator has outputs with differential and common-mode components

$$\begin{pmatrix} v_{o,d} \\ v_{o,cm} \end{pmatrix} = \begin{pmatrix} v_{o1} - v_{o2} \\ \dfrac{v_{o1} + v_{o2}}{2} \end{pmatrix}$$
$$= \begin{pmatrix} \dfrac{G_{m,1}}{sC_1} + \dfrac{G_{m,2}}{sC_2} & \dfrac{G_{m,CM,1}}{sC_1} - \dfrac{G_{m,CM,2}}{sC_2} \\ \dfrac{1}{2}\left(\dfrac{G_{m,1}}{sC_1} - \dfrac{G_{m,2}}{sC_2}\right) & \dfrac{1}{2}\left(\dfrac{G_{m,CM,1}}{sC_1} + \dfrac{G_{m,CM,2}}{sC_2}\right) \end{pmatrix} \begin{pmatrix} v_{id} \\ v_{icm} \end{pmatrix}$$
$$\approx \begin{pmatrix} \dfrac{G_m}{sC} & \Delta\left(\dfrac{G_{m,CM}}{sC}\right) \\ \dfrac{1}{2}\left(\Delta\left(\dfrac{G_m}{sC}\right)\right) & \dfrac{1}{2}\left(\dfrac{G_{m,CM}}{sC}\right) \end{pmatrix} \begin{pmatrix} v_{id} \\ v_{icm} \end{pmatrix}$$

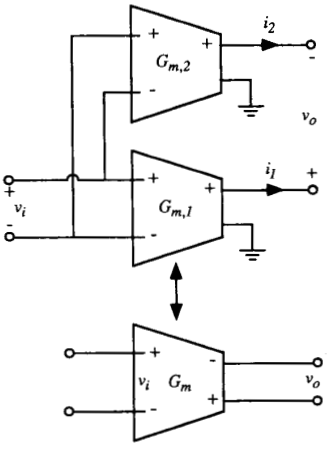


Fig. 3. The balanced transconductor.

where G_m and $G_{m,CM}$ refer to average values and Δ terms to mismatches.

The undesirable terms here are off-diagonal, and cause conversion gains between common-mode and differential signals. In particular, the $\Delta(G_m/sC)$ term converts differential signals to common-mode signals, which will then cause intermodulation distortion in transconductors and damping resistors connected to the output since differential gain and damping resistance (Section IV-B) are set by common-mode voltage.

The $\Delta(G_{m,CM}/sC)$ term is likely to be bigger for this circuit than for a transconductor based on a classical differential pair. This can cause gain errors, but an adaptive system would not be sensitive to them. The term's most significant effect is that it may limit rejection of power-supply noise and input common mode.

Our use of extremely small devices means that mismatches will be large.

E. Effects of G_o and f_t on Q

An ideal transconductor, when capacitively loaded, would produce a frequency response that rolls off at exactly 20 dB/decade with exactly 90° of phase shift at all frequencies. Finite output conductance reduces gain and phase at low frequencies; parasitic poles in G_m cause ex-

cess phase and reduced gain; and finite gain–drain capacitance causes a right-half-plane zero for excess phase and increased gain [19], [23]. Practical transconductors work well at frequencies (geometrically) halfway between the low-frequency pole $G_o/2C$ and the high-frequency pole (near f_t) or zero (G_m/C_{gd}); at frequencies below this point the output conductance provides excess damping (hence reducing Q) and at higher frequencies the excess phase causes Q enhancement or even oscillation. Our chip was overdamped up to frequencies near 300 MHz, indicating that G_o damping dominates over the VHF range.

IV. The Biquad Loop

A second-order loop was implemented because it is a simple structure that demonstrates the capabilities of the technology. This structure consists of a pair of balanced transconductors connected in a loop as shown in Fig. 4, together with capacitors to define frequency ω_o, damping devices R_D, and voltage sources V_{CM} to set common-mode levels and thereby tune the transconductors. All components were implemented with MOS devices to obtain simplicity, reduced process sensitivity, and tunability, at some cost in linearity.

When all devices are operated at equal bias levels so that $R_D = 2/G_m$ the biquad operates as described by (3) in Section II. It has a Q of $\sqrt{5}/2 \approx 1.1$ and a center frequency $\omega_0 \approx 1.1(G_m/C)$.

A. Capacitors

We chose to use grounded, rather than differential, capacitors in order to stabilize the common-mode loop caused by the finite common-mode gain of the transconductors. This loop would ordinarily be highly unstable, since a loop of two integrators with gains having the same signs has real poles at $\pm(G_{m,CM}/C_{CM})$. Choosing to use grounded capacitors (and minimizing common-mode gain) brings these roots as close to the origin as possible for a given differential-loop ω_o. Grounded damping devices can then force both poles into the left half-plane to obtain common-mode stability.

Looking at the loop from a voltage point of view, the transconductors have approximately unity common-mode gain; the common-mode loop has a loop gain of 1 and is just on the edge of instability. Grounded damping devices then guarantee stability. Mismatches may increase common-mode gain, and this may be a problem at high Q where damping is small.

Because transconductance can only be controlled over a limited range, adjustable capacitors are used to obtain a broad tuning range. We implemented switchable capacitors by using the channel capacitances of a string of NMOS devices as shown in Fig. 5(a). These were laid out in a ring pattern [24] (Fig. 5(b)) to minimize parasitic $C_{\min}$, which defines the maximum operating frequency. As transistors M_{C1} to M_{C4} are turned on by external control inputs, progressively lower frequency ranges are obtained. The overall balanced circuit structure cancels even-order distortion products, and also cancels signal currents in the control inputs. The signal-swing constraint resulting from the need to maintain channels in these devices is identical to that required to bias the transconductors, and hence not a limitation.

As well as allowing coarse frequency tuning and avoiding the need for a double-poly capacitor, using MOS capacitors reduces the effect of process variation; oxide thickness and channel width appear identically in expressions for transconductance and capacitance, leaving G_m/C insensitive to these highly variable parameters. Channel length continues to contribute variability.

A second version of the filter, using fixed double-poly capacitors, was also fabricated and tested for comparison. Its measured overall distortion levels were comparable.

At low-frequency settings the series string of ON devices has the structure of an RC line, which complicates the frequency response. This effect is tolerable since the lower passband frequency at these settings makes high-frequency poles less critical.

B. Damping and Common-Mode Control

Although a self-connected transconductor can be used for damping, it is not the best choice here: it implements a differential resistor, and so does not stabilize the common-mode loop. We chose instead to use the source impedance of an NMOS device biased by a current source as shown in Fig. 6(a). With only two transistors per side, this gives us common-mode control as well as damping by approximating the Thévenin equivalent of Fig. 6(b). Even-order distortions are, of course, canceled by the balanced structure. Using the same circuit for common-mode control and damping requires that both integrators be damped, dictating the filter structure chosen.

The ratio between device sizes in the transconductor and the damping circuit sets the nominal circuit Q (1.1, (3)), which can be adjusted at the cost of dynamic range.

At high Q (relative to the nominal level) these devices operate at low bias levels and appear to dominate distortion performance. Back-bias terms contribute nonlinearity in addition to the short-channel terms discussed in Section III-B. We biased the two integrators identically (all the $V1$ inputs were tied together, as were the $V2$ lines).

C. Probes

At the high-frequency extreme, node capacitances are well under 1 pF. It is therefore difficult to probe the nodes without affecting the circuit.

While a conventional solution to this problem would be to add source followers as buffers, we preferred to use common-source devices M_{T1}–M_{T4} as shown in Fig. 4. These are linear under the same assumptions as the transconductor, and can directly drive a 50-Ω load for RF measurements. A source follower using these small devices, driving a 50-Ω load, where $50\ \Omega \ll 1/g_{mT}$, would in any case have a low voltage gain ($\approx g_{mT}R_L$)—the same as the common-source device. The dependence of the output on

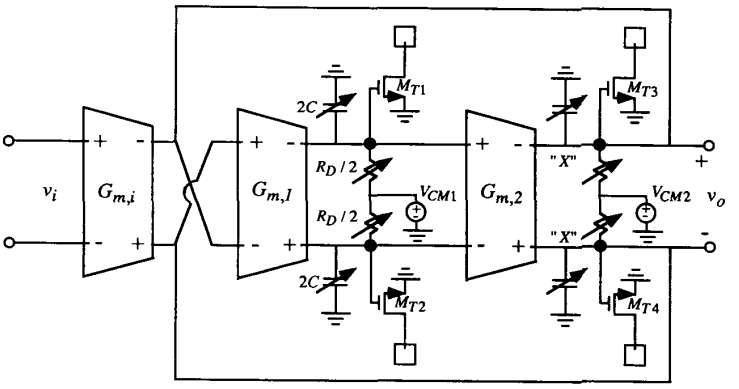


Fig. 4. The second-order biquad loop. Capacitors and damping devices are detailed in Figs. 5 and 6(a).

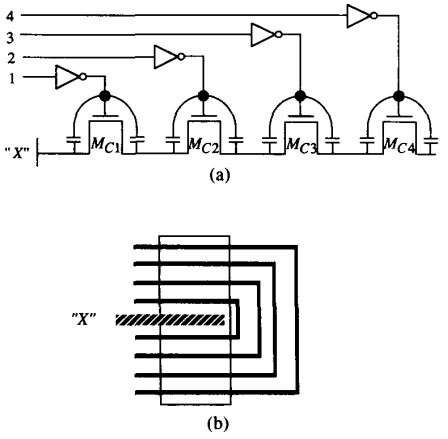


Fig. 5. (a) Switchable capacitor chain. (b) Layout pattern. Node "X" is from Fig. 4.

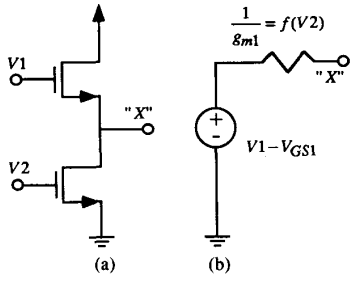


Fig. 6. (a) Common-mode control and damping details. (b) Thévenin equivalent of Fig. 6(a).

g_{mT} makes it difficult to measure internal signal levels, which must be estimated using simulated g_{mT} values. A test transistor on-chip would have been useful. The low gain and output levels from these test devices also make noise measurement impossible, so again simulated values are needed.

In a practical system we expect the signals from these filters would never come off-chip, so that these probe devices should be regarded as test equipment.

V. Experimental Results

A photomicrograph of the circuit appears in Fig. 7. The active area is approximately 255 μm × 112 μm. The transconductors may be recognized by the three center sections while the capacitors are the four distinct squares surrounding the transconductors. The entire biquad occupies roughly twice the area of a single bonding pad without having been very tightly laid out.

A. Frequency Response

The test setup consisted of the filter and two 50-Ω 180° power splitters. These were connected at the filter input to provide single-ended-to-differential conversion and at the filter output to achieve differential-to-single-ended conversion. Termination resistors (50 Ω) were placed at the filter input and output ports.

A typical measured frequency response of the filter is shown in Fig. 8. Measured f_o and Q were 274 MHz and 14, respectively.

Simulated and experimental values for I_{MT} (probe bias current) at given bias settings agreed to within 3%, suggesting that the simulated model was quite good. Hence, probe transconductance values derived from simulations were used to estimate internal signal levels. The peak internal output differential level for the response shown in Fig. 8 is about 20 mV_{rms} for a differential input level of 3.2 mV_{rms}.

For this response (f_o = 274 MHz), the core filter consumed 2.4 mA (and the probes consumed 2.2 mA) for a filter power dissipation of 12 mW. For comparison, at a center frequency of 220 MHz and a Q of 2.5 the core filter dissipated 8 mW. As expected, higher speeds cost power.

B. Tuning Range

Depending on tuning voltages and the number of capacitors in use, a variety of f_o and Q values can be obtained. Fig. 9 is a scatter plot showing (f_o, Q) values obtained experimentally.

The lines shown connect (f_o, Q) pairs at constant $V2$ levels and different $V1$. These two bias levels set the com-

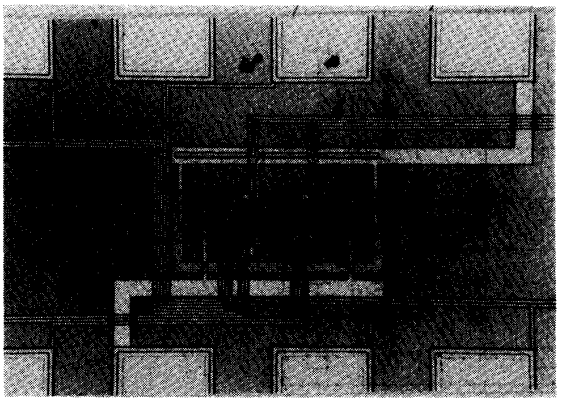


Fig. 7. A photomicrograph of the 0.9-μm CMOS biquad.

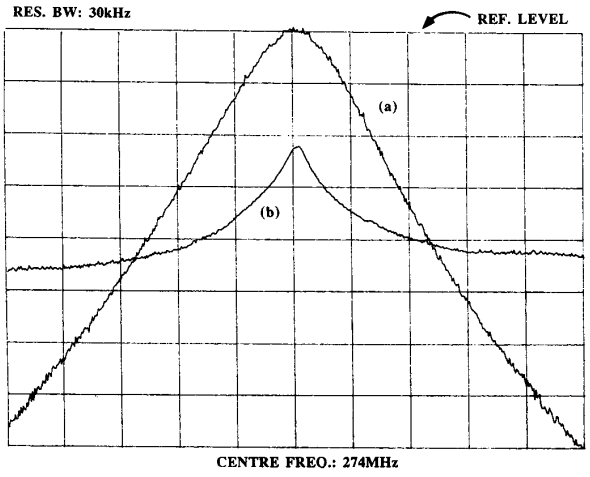


Fig. 8. Filter frequency response: curve (*a*) 5 MHz/div, 1 dB/div, −62-dBm reference level; curve (*b*) 50 MHz/div, 10 dB/div, −40-dBm reference level.

mon-mode voltage, which in turn defines the bias current in the transconductors and the test devices. The latter bias current (I_{MT}) can be measured and is indicated in the figure. The lines are roughly constant Q (for low Q), so $V2$ primarily sets Q. Obviously the two controls interact. At low (near nominal) Q the interaction does not appear severe enough to stop an adaptive filter from working, but above $Q \approx 10$ adaptation will be difficult.

The plot shows, as expected, that higher Q's are obtained at higher frequencies, as f_t-related effects begin to cancel the G_o-induced damping that dominates at lower frequencies. These effects appear to balance at about 300 MHz, below which the circuit is stable even with the explicit damping devices turned off. At the highest frequency range, operation is dominated by parasitics and oscillation is common.

The tuning regions overlap in frequency at low (near nominal) Q, so the entire VHF range is covered. A more generous overlap would be needed in a manufacturable product, and can be obtained by changing capacitor ratios.

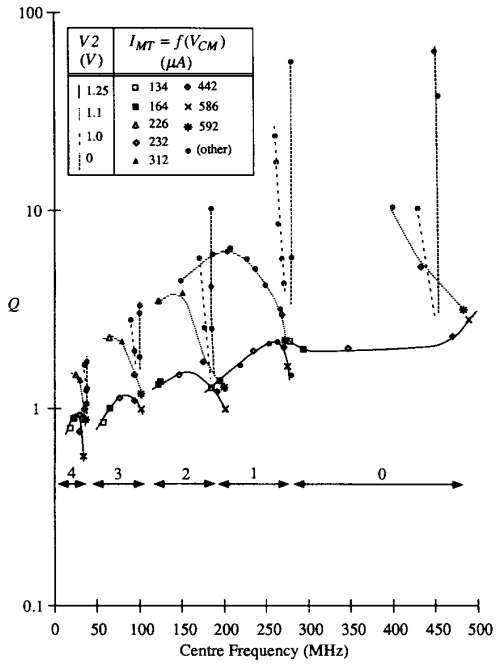


Fig. 9. A scatter plot illustrating overall tuning range. Lines indicate different damping resistances (Q settings). Points indicate different common-mode levels (f_o settings). The number of capacitors switched ON is indicated above the arrows (five distinct frequency regions of operation).

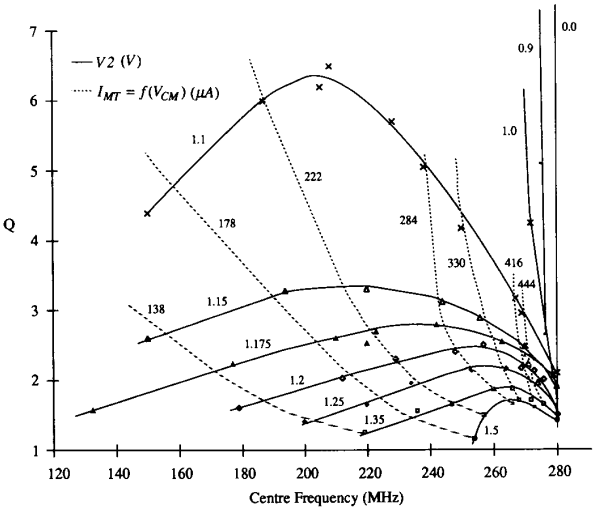


Fig. 10. A detailed scatter plot for region 1 of Fig. 9.

Fig. 10 is a more detailed view of Fig. 9 and concentrates on region 1 near nominal Q. The solid lines connect (f_o, Q) pairs at constant $V2$ levels while the dotted lines represent constant V_{CM} (I_{MT}) levels.

The figure shows that as V_{CM} is increased the dotted lines close up indicating reduced frequency tuning range. This effect is a direct result of the $\theta(V_{CM} - V_T)$ term discussed in Section III-B. In the limiting case, the center frequency settles to 280 MHz ($V_{CM} \doteq 2.9$ V or

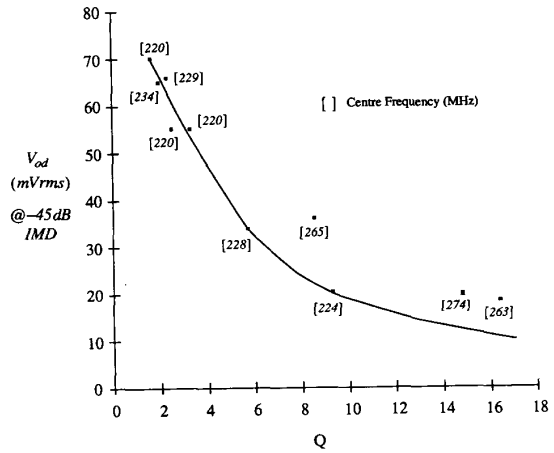


Fig. 11. A plot of internal signal level at constant (−45 dB) IM distortion as function of Q.

I_{MT} = 650 μA), implying constant transconductance, as predicted by (8).

At low bias levels for the damping devices ($V2 < 0.9$ V) the common-mode level is no longer set by $V1$ and $V2$ but by the transconductors (about midsupply) and again no tuning of the center frequency is possible. The Q, however, can still be adjusted from about 1 to 56 by varying $V1$.

C. Dynamic Range

An intermodulation-distortion (IM) test was performed to estimate the cubic distortion of the whole filter by measuring the output component at the upper 3-dB frequency when inputs are at the band center and the lower 3-dB point. This test was performed over a range of Q values with f_o held near 220 MHz. Fig. 11 depicts the results.

The signal level plotted in the figure was the internal differential output level (f_o component, which is about 2 dB below the total signal) for which the output level at the upper 3-dB frequency was 40 dB below that at the lower 3-dB frequency. This corresponds to 0.58% distortion. The curve shows how the usable signal level drops with increasing Q. We presume that at higher Q's linearity suffers from the small ON voltage of the damping devices. A better method for tuning Q is thus needed for a practical device.

To measure the cubic distortion of the input transconductor alone, a tone at $f_o/3$ was supplied to the filter. A reasonably high Q (14) was established to ensure that internal out-of-band signals will not contribute to the output while the input transconductor's cubic distortion component will be amplified and appear at the center frequency. By comparing the gain for the third harmonic to that obtained for an in-band tone, S/D for the input transconductor can be deduced.

For this test, we obtained 1% distortion for a differential filter input of 355 mV_{rms}. This signal level is 14 dB lower than the value (10), derived from the simplified analysis in Section III-B, would predict for $\theta = 0.6\ V^{-1}$ and $V_{ON} = 1.8$ V.

An unexpected feature of the actual transconductor nonlinearity is that a further 6-dB increase in input level only raises the third harmonic to 1.1%, rather than the 4% that would be expected for a pure cubic. This can be interpreted to mean that there is a zero-crossing distortion component. Zero-crossing distortion *decreases* as a percentage of signal as signal levels increase, while saturation effects produce distortion that *increases* with signal. A system with both distortion mechanisms has a composite relation between signal level and distortion. At high Q settings the damping devices are almost off, so zero-crossing distortion can be expected.

Output noise was dominated by the test setup, because of the low gain of the probe devices. A lower bound on the noise power can be obtained by estimating kT/C. For the response in Fig. 8 the value for C is 131 fF, which predicts a minimum noise power of 0.18 mV_{rms} on a differential output.

Using a noise current of $\overline{i_n^2} = (8/3)kTg_m$ for a transistor, a value of 1.3-mV_{rms} noise power on a differential output is obtained. This estimate is optimistic as short-channel devices are noisier. The simulated value was 2.43 mV_{rms}, about 23 dB higher than the kT/C bound. The transconductor is noisy because it has eight devices contributing equally to output noise current. One way to reduce noise by up to 3 dB, while maintaining the basic circuit, would be to use smaller devices for the current mirrors in the transconductors. However, the smaller devices would degrade common-mode stability and signal swing. An even simpler transconductor based on a single differential pair would be better. We are working on this type of transconductor using a BiCMOS process and investigating an alternative method for common-mode control.

From Fig. 11, for Q near the nominal design value,

0.58% IM distortion occurs at differential signal levels of about 86 mV_{rms} (total signal: 70 mV_{rms} at the center frequency and 50-mV_{rms} at the lower 3-dB frequency). Using the simulation figure for the noise we obtain an SNR of 31 dB at 0.58% IM distortion (−45 dB). $S/(N + D)$ reaches a maximum of 33 dB when signal level is increased to 158 mV_{rms}.

SNR performance at the high-frequency setting is worst case because the higher node capacitance, when more capacitors are switched, in filters noise better (the same noise density appears in a smaller bandwidth). Distortion levels also appear to be lower at lower frequencies, perhaps because nonlinear parasitic capacitances are less important.

D. PSRR and CMRR

A swept sinusoidal signal was injected to each supply line with the filter inputs grounded. The in-band gain was compared to the in-band gain that a swept sinusoidal signal produced when applied to the filter input. Measured results (at a center frequency of 120 MHz and a Q of 9) indicate a power-supply rejection ratio of 30 dB from the positive supply and 23 dB from the negative supply. The lower rejection for inputs injected onto the negative supply is probably due to the output probes, which sense only the negative supply, and from back-gating of the top transistor of each damping circuit. This test was only performed on the version of the filter with fixed double-poly capacitors so we do not yet know the PSRR effects of the switchable capacitors.

The measured common-mode rejection ratio is 32 dB under the same test conditions. Assuming that this performance is dominated by the input transconductor (because the common-mode circuit is low pass and attenuates signals at f_o so that later stages have less common-mode signal power to convert) gives $\Delta(G_{m,CM}/sC) \approx 0.025(G_m/sC)$. About 20 dB of this is attributed to G_m/G_o and the remaining 12 dB to mismatching.

VI. Conclusions

We have demonstrated that a 0.9-μm CMOS process is capable of performing filtering in the VHF band with good Q factors and moderate SNR. Very simple circuits that were sparing of silicon area and power were used. The most likely areas of application of this capability are in data communications, where high speeds, moderate SNR, and tunability are appropriate.

Dynamic range performance is limited by poor linearity, particularly when the filter is operated far from its nominal design frequency and Q. "Linearized" transconductors are of limited effectiveness when channels are short, and perhaps even for long devices in a process with thin gate oxide. A BiCMOS transconductor is being designed to separate the bias and tuning functions so that maximum linearity can be maintained over the full tuning range. There is room for improvement in the transconductor noise performance. Linearity and tunability of the damping devices will also have to be improved.

Acknowledgment

The design was largely done during a sabbatical at AT&T Bell Laboratories in Reading, PA, in the CMOS group led by T. R. Viswanathan. Discussions with him, and frequent discussions with S. Lewis, J. Fernandez, J. Sonntag, and S. Fetterman, all contributed to the work. Test equipment was provided by the VLSI research group, University of Toronto.

References

[1] J. J. Friend, C. A. Harris, and D. Hilberman, "STAR: An active biquadratic filter section," *IEEE Trans. Circuits Syst.*, vol. CAS-32, no. 2, pp. 115-121, Feb. 1975.

[2] M. Banu and Y. Tsividis, "Fully integrated active *RC* filters in MOS technology," *IEEE J. Solid-State Circuits*, vol. SC-18, no. 6, pp. 644-651, Dec. 1983.

[3] J. T. Caves, M. A. Copeland, C. F. Rahim, and S. O. Rosenbaum, "Sampled analog filtering using switched capacitors as resistor equivalents," *IEEE J. Solid-State Circuits*, vol. SC-12, no.6, pp. 592-600, Dec. 1977.

[4] K. Martin, "Improved circuits for the realization of switched-capacitor filters," *IEEE Trans. Circuits Syst.*, vol. CAS-27, no. 4. pp. 237-244, Apr. 1980.

[5] J. J. Culbert and W. M. Snelgrove, "A continuous time 5 MHz biquad filter in BNR CMOS-1B," in *Proc. 1986 Canadian Conf. Very Large Scale Integration*, Oct. 1986, pp. 247-252.

[6] K. W. Moulding, J. R. Quartly, P. J. Rankin, R. S. Thompson, and G. A. Wilson, "Gyrator video filter IC with automatic tuning," *IEEE J. Solid-State Circuits*, vol. SC-15, no. 6, pp. 936-968, Dec. 1980.

[7] A. V. Oppenheim and R. W. Schafer, *Discrete Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[8] H. Khorramabadi and P. R. Gray, "High-frequency CMOS continuous time filters," in *Proc. IEEE Int. Symp. Circ. Syst.* (Montreal, Canada), May 7-10, 1984, pp. 1498-1501.

[9] W. M. Snelgrove and A. S. Sedra, "Synthesis and analysis of state-space active filters using intermediate transfer functions," *IEEE Trans. Circuits Syst.*, vol. CAS-33, no. 3, pp. 287-301, Mar. 1986.

[10] D. A. Johns, W. M. Snelgrove, and A. S. Sedra, "Orthonormal ladder filters," *IEEE Trans. Circuits Syst.*, vol. 36, no. 3, pp. 337-343, Mar. 1989.

[11] C. Plett, M. A. Copeland, and R. A. Hadaway, "Continuous time filters using open loop tunable transconductance amplifiers," in *Proc. IEEE Int. Symp. Circ. Systems* (San Jose, CA), May 5-7, 1986, pp. 1173-1176.

[12] R. L. Geiger and E. Sanchez-Sinencio, "Active filter design using operational transconductance amplifiers: A tutorial," *IEEE Circuits Devices Mag.*, vol. 2, no. 1, pp. 20-32, Mar. 1985.

[13] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[14] D. A. Johns, W. M. Snelgrove, and A. S. Sedra, "Continuous-time LMS adaptive recursive filters," *IEEE Trans. Circuits Syst.*, vol. 38, no. 7, pp. 769-778, July 1991.

[15] F. J. Fernandez, "Linear CMOS transconductance element," U.S. Patent 4 734 654, Mar. 29, 1988.

[16] C. S. Park and R. Schaumann, "Fully integrated analog filters in CMOS technology," in *Proc. IEEE Int. Symp. Circ. Systems* (San Jose, CA), May 5-7, 1986, pp. 1161-1164.

[17] R. Castello and P. R. Gray, "A high-performance micropower switched-capacitor filter," *IEEE J. Solid-State Circuits*, vol. SC-20, no. 6, pp. 1122-1132, Dec. 1985.

[18] A. Nedungadi and T. R. Viswanathan, "Design of linear CMOS transconductance elements," *IEEE Trans. Circuits Syst.*, vol. CAS-31, no. 10, pp. 891-894, Oct. 1984.

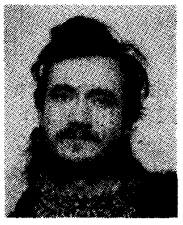
[19] J. J. Culbert, "On transconductance-C filters," M.A.Sc. thesis, Univ. of Toronto, Toronto, Canada, Jan. 1988.

[20] S. T. Dupuie and M. Ismail, "High frequency CMOS transconductors," in *Analogue IC Design: The Current-Mode Approach*, C. Toumazou, F. J. Lidgey, and D. G. Haigh, Eds. London: Peter Peregrinus, 1990, pp. 181-238.

[21] C. Duvvury, "A guide to short-channel effects in MOSFETs," *IEEE Circuits Devices Mag.*, vol. 2, no. 6, pp. 6-10, Nov. 1986.

[22] K.-Y. Toh, P.-K. Ko, and R. G. Meyer, "An engineering model for

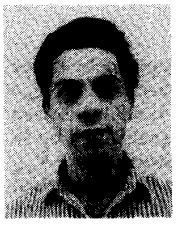
short-channel MOS devices," *IEEE J. Solid-State Circuits*, vol. 23, no. 4, pp. 950–958, Aug. 1988.
[23] K. D. Peterson, A. P. Nedungadi, and R. L. Geiger, "Amplifier design considerations for high frequency monolithic filters," in *Proc. ECCTD 87 (Circuit Theory and Design)*, R. Gerber, Ed. Amsterdam: Elsevier Science, 1987, pp. 321–326.
[24] J. Sonntag, AT&T Bell Laboratories, private communications.



W. Martin Snelgrove (S'75–M'81) received the B.A.Sc. degree in chemical engineering and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, Ont., Canada, in 1975, 1977, and 1982, respectively.

In 1982 he worked at INAOE in Mexico as Visiting Researcher in CAD. Since then he has been at the University of Toronto, where he is currently an Associate Professor. In 1989 and 1990 he spent six-month sabbatical leaves as a Resident Visitor at AT&T Bell Laboratories in Reading, PA, working in CMOS analog design. In the spring of 1991 he taught an analog VLSI course at Oulu University, Finland, from an early draft of a new textbook. He is interested in architectures for high-performance integrated circuits for signal processing applications. This work includes RF signal conditioning, high-speed data conversion, real-time DSP, and CAD for signal processing.

Dr. Snelgrove was the winner of the 1986 Circuits and Systems Society Guillemin-Cauer Award for a 1986 paper coauthored with A. Sedra.



Ayal Shoval received the B.A.Sc. and M.A.Sc. degrees in electrical engineering from the University of Toronto, Canada, in 1989 and 1991, respectively. His M.A.Sc. thesis focused on hybrid analog/digital trimming techniques involving the use of median estimation to obtain offset-free update circuitry for analog adaptive filters. He is currently working towards the Ph.D. degree in the Department of Electrical Engineering at the University of Toronto. His present research interest is in the area of high-frequency adaptive analog filters for data communications.