Improved Implementation of the Silicon Cochlea

Lloyd Watts, Douglas A. Kerns, Richard F. Lyon, Member, IEEE, and Carver A. Mead

Abstract—The original "analog electronic cochlea" of Lyon and Mead (1988) used a cascade of second-order filter sections in subthreshold analog VLSI to implement a low-power, real-time model of early auditory processing. Experience with many silicon-cochlea chips has allowed the identification of a number of important design issues, namely dynamic range, stability, device mismatch, and compactness. In this paper, the original design is discussed in light of these issues, and circuit and layout techniques are described which significantly improve its performance, robustness, and efficiency. Measurements from test chips verify the improved performance.

I. INTRODUCTION

The "analog electronic cochlea" of Lyon and Mead [1] used subthreshold analog VLSI to implement a low-power, real-time, biologically motivated model of early auditory processing. The model was based on a serial cascade of second-order filter sections whose cutoff frequencies decrease exponentially with distance into the cascade, to capture the essence of normal unidirectional wave propagation in real biological cochleas. By varying the degree of resonance in the second-order sections, it was possible to model both passive and active cochlear function.

Since the original work, the silicon cochlea has been successfully used in several analog VLSI models of higher auditory function, including spatial localization [2], pitch detection [3], and a variety of correlator-based signal representations [4], [5].

Other analog VLSI models have recently been proposed in order to capture more of the fine detail of the biological cochlea function. These models include a cascade of third-order filter sections [6], a bidirectional transmission line based directly on cochlear fluid mechanics [7], and a first-order delay line which couples each stage into a gyrator-based resonant circuit [8]. While each of these approaches has strengths and weaknesses, the original cascade of second-order sections is desirable for its simplicity, compactness, and good performance.

A number of major design issues have been identified to improve the performance of the original silicon cochlea design, namely dynamic range, stability, device mismatch, and compactness. In this paper, the original design is discussed in light of these issues, and circuit and layout techniques are described which significantly improve its performance, robustness, and efficiency. Test results are presented from a number of working chips.

II. THE ORIGINAL CIRCUIT

In this section we review the important aspects of the original cochlea design, as relevant to the present work.

A. The Transconductance Amplifier

Fig. 1 shows the basic transconductance amplifier used in the original cochlea circuit. A transconductance amplifier biased in subthreshold has a hyperbolic tangent transfer characteristic given by

$$I_{out} = I_{bias} \tanh \left( \frac{\kappa (V_+ - V_-)}{2U_T} \right)$$

where $\kappa$ is the back-gate coefficient (typical value 0.7) and the thermal voltage $U_T = kT/q = 25.6$ mV at room temperature.

For small inputs, the amplifier behavior can be approximated as a linear transconductance:

$$I_{out} = g_m (V_+ - V_-), \quad |V_+ - V_-| < 60 \text{ mV}$$

where the transconductance $g_m$ is given by

$$g_m = \frac{I_{bias} \kappa}{2U_T}$$

and $I_{bias}$ is exponentially controlled by the bias voltage

$$I_{bias} \propto \exp \left( \frac{\kappa V_{bias}}{U_T} \right).$$

For large inputs, the output current $I_{out}$ saturates at $\pm I_{bias}$.

B. The Second-Order Section

The original second-order section used in the filter cascade is shown in Fig. 2, where the amplifiers are transconductance amplifiers as described above.

The transfer function for the second-order section is given by

$$H(s) = \frac{V_{out}(s)}{V_{in}(s)} = \frac{1}{1 + \frac{s \tau}{Q} + \tau^2 s^2}$$

Manuscript received September 30, 1991; revised January 17, 1992. This work was supported by the Office of Naval Research and the System Development Foundation.

L. Watts and D. A. Kerns are with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125.

R. F. Lyon is with Apple Computer, Inc., Cupertino, CA 95014, and the Department of Computer Science, California Institute of Technology, Pasadena, CA 91125.

C. A. Mead is with the Department of Computer Science, California Institute of Technology, Pasadena, CA 91125.

IEEE Log Number 9107225.
where the time constant \( \tau = C/g \), and the filter quality factor

\[
Q = \frac{1}{2(1 - \alpha)}
\]

where \( \alpha = g_0/(2g_r) \). Typical values for the parameters are \( C = 1 \text{ pF} \), and \( 10^{-10} \Omega^{-1} < g_r < 10^{-7} \Omega^{-1} \), resulting in audio-frequency time constants. Sections with \( Q \) less than 0.707 have a purely low-pass character, while sections with \( Q \) greater than 0.707 have a resonant peak, whose height increases and width decreases with increasing \( Q \).

C. The Cochlea

The silicon cochlea consists of a serial cascade of the second-order filter sections, as shown in Fig. 3. Each filter section has a signal input \( V_{in} \), a signal output \( V_{out} \), and two bias voltages \( V_i \) and \( V_Q \) which control its time constant \( \tau \) and filter quality factor \( Q \), respectively. Since the time constant of the filter decreases exponentially with the bias voltage \( V_i \), it is possible to configure the cascade to have an exponentially increasing time constant (exponentially decreasing cutoff frequency) with distance into the cascade by applying a decreasing linear "tilt" on the \( V_i \) bias voltage seen by each filter section. The tilted line is easily achieved by using a thin polysilicon wire as a long resistor, applying appropriate voltages at the two ends (\( V_i \) and \( V_{out} \)), and taking the bias voltages for the filter stages at regularly spaced intervals. A similarly tilted bias line is used to apply the \( V_Q \) bias inputs, usually such that all stages have the same \( Q \).

Since each stage of the cascade has unity gain at dc, the propagating voltage signal is interpreted as a fluid pressure wave in the cochlea. The composite effect of many stages results in very steep high-frequency cutoff, often 100 dB/octave or more, as measured in real cochleas [9]. In order to qualitatively match the measurements of basilar membrane velocity, a differentiator circuit is often appended to each output tap to tilt the low side of the response by 6 dB/octave, giving an overall asymmetric bandpass character to the outputs [1].

In order to model active processes in the cochlea associated with outer hair cells, it is desirable to tune each second-order section in the cascade to have a small resonant peak (\( Q > 0.707 \)) [1]; the resulting cascade will combine the effects of the small individual resonances to produce a large composite "pseudoresonance," which is more broadly tuned than a single resonance of the same gain.

III. MAJOR ISSUES

A. Dynamic Range

Typically, the linear range of the transconductance amplifiers limits the size of input signals to about 60 mVpp. With on-chip noise levels of about 0.1 mVpp, the input dynamic range has an upper limit of about 55 dB for a single section, and significantly less for a cascade. Increasing the linear range of the amplifiers would increase the dynamic range of the silicon cochlea.

B. Stability Limits

It can easily be seen that the second-order section has a small-signal stability limit at \( \alpha = g_0/(2g_r) = 1 \), since this point \( Q \to \infty \). Not so obvious is the existence of a large-signal stability limit, derived using a piecewise linear approach [10], at the point \( \alpha = 0.809 \), corresponding
Fig. 4. Chip data illustrating small-signal and large-signal stability limits for a single second-order section. The input is a sine wave at the resonant frequency. When \( \alpha \) is increased from 0, the second-order section becomes more and more resonant and the amplitude of the output increases until the small-signal stability limit is reached at \( \alpha = 1 \). At this point, large-signal limit-cycle oscillations occur, and the amplitude of the output is large and constant. In order to restore small-signal behavior, it is necessary to reduce \( \alpha \) below the large-signal stability limit of \( \alpha = 0.809 \).

to \( Q = 2.63 \). The stability of the second-order section may be summarized:

- \( 0 < \alpha < 0.809 \) unconditionally stable
- \( 0.809 < \alpha < 1 \) small-signal stable, large-signal unstable
- \( \alpha > 1 \) unstable

and is illustrated with chip data in Fig. 4.

Since in practice we cannot guarantee that the input to each section will be small, we must operate each section in the unconditionally stable regime, which requires that \( \alpha < 0.809 \) and \( Q < 2.63 \) for all sections. The elimination of the large-signal instability is therefore an important design objective.

C. Device Matching

It is well known [13] that matching in subthreshold MOS devices is poor; that is, the current flowing in each of two transistors with the same drawn geometry and bias conditions may differ by as much as a factor of 2 if the devices are small. This poor control over the parameters of each section can have a number of undesirable effects.

Since the tilted polysilicon resistive wires that are used to bias the \( V_r \) and \( V_Q \) of each section are passive devices, we can guarantee that the bias voltages to each stage will decrease monotonically with distance into the cascade, although local variations in the width, thickness, and resistivity of the polysilicon wire may cause the voltage drops from stage to stage to be slightly nonuniform. However, the large random variations in the transconductance of each amplifier may easily destroy the monotonicity of the resonant frequency as a function of position.

An even more serious problem is the effect of device matching on the stability of each section. In order to ensure that the worst-case \( \alpha < 0.809 \), we should make the typical \( \alpha < 0.404 \) to allow for a factor of 2 error and still have all stages unconditionally stable. Thus, the typical second-order section in the cascade will have a \( Q < 0.84 \) and will contribute only a very small amount of gain; most of the gain will be contributed by a few worst-case sections.

D. Compactness

It is possible to improve the matching by increasing the size of the devices—the matching error decreases roughly as the square root of device area. However, in order to model real cochleas, we will need a large number of stages in the filter cascade, which means that the sections should be made as compact as possible. The desired compactness can be achieved by eliminating redundant circuit elements, and by using small devices everywhere except where matching is important to the system-level behavior.

IV. CIRCUIT AND LAYOUT TECHNIQUES

In this section, we address the major issues listed above with circuit and layout techniques to improve the performance of the original cochlea design.

A. Increasing Dynamic Range

There are a number of techniques for increasing the linear range of subthreshold MOS transconductance amplifiers, including capacitive division and source degeneration. We have used source degeneration successfully in our cochlea designs, as described below.

Fig. 5 shows a transconductance amplifier with source degeneration via diode-connected transistors, one on each side of the differential pair. It can easily be shown that

\[
I_{\text{out}} = I_{\text{bias}} \tanh \left( \frac{k_r (V_+ - V_-)}{2U_T} \right)
\]

where

\[
k_r = \begin{cases} 
\kappa & \text{(no diodes)} \\
\kappa^2 / (k + 1) & \text{(one diode per side)} \\
\kappa^2 / (k^2 + k + 1) & \text{(two diodes per side)}
\end{cases}
\]

For a typical value of \( \kappa = 0.7 \), we expect degeneration with one diode per side to widen the linear range by a factor of 2.4, and expect degeneration with two diodes per side to widen the range by a factor of 4.4. Data from subthreshold transconductance amplifiers are shown in Fig. 6.

The price for the increased linear range is a decrease in the common-mode operating range and increased thermal noise injection. For a normal transconductance amplifier (no diodes) in subthreshold operation in a follower configuration, the inputs must be constrained between \( V_{\text{sat}} + V_{\text{dd}} / \kappa \) and \( V_{\text{dd}} - V_{\text{sat}} \) in order to keep the bias transis-
Fig. 5. Wide-range-input transconductance amplifier. (a) Schematic, with one diode per side. (b) Symbol.

Fig. 6. Data from transconductance amplifiers with zero, one, and two source-degeneration diodes per side. Increasing the number of diodes per side increases the width of the response.

We find that the amplifier with one diode per side gives a reasonable increase in linear range without undue restrictions on the common-mode operating range, for a modest improvement in the input dynamic range of about 7.6 dB.

A proposed symbol for the wide-range-input transconductance amplifier is shown in Fig. 5; this symbol is usually used only when necessary to distinguish a wide-range-input transconductance amplifier from other transconductance amplifiers in a circuit.

B. Eliminating the Large-Signal Instability

In order to achieve a high $Q$ in a second-order section, $g_Q$ must be nearly twice as large as $g_\text{f}$. However, increasing the transconductance $g_Q$ results in an increase in the saturation current $I_{\text{bias},Q}$ (equation (3)), causing the second-order section to become large-signal unstable when the saturation current $I_{\text{bias},Q}$ is less than $1.62 I_{\text{bias}}$.

The large-signal instability may be eliminated by increasing the linear range (and hence the saturation current) of the amplifiers in the feedforward direction, while keeping the feedback amplifier range narrow, as shown in Fig. 7. The transconductance of the feedback amplifier $g_\text{f}$ can then be increased to be as large as the transconductance of the feedforward amplifiers $g_\text{f}$, while the saturation current $I_{\text{bias},Q}$ is less than $1.62$ times as large as $I_{\text{bias}}$.

It is apparent that the ratio of feedforward linear range to feedback linear range must be at least $2/1.62 = 1.23$ in order to eliminate the large-signal instability. This condition is easily achieved by using a conventional amplifier in the feedback direction, and wide-range amplifiers with one diode per side in the feedforward direction, as shown in Fig. 7, for a width ratio of 2.4; another good variation is to use two diodes per side in the feedforward direction and one diode per side in the feedback direction, for a width ratio of $4.4/2.4 = 1.83$.

C. Improving Matching

The traditional methods for improving matching [11] include making devices larger to reduce the effect of process variations, placing match-sensitive devices close to each other to reduce the effect of gradients in process parameters and temperature, and common-centroid techniques, in which devices to be matched are duplicated and connected in parallel around a central point to eliminate the effect of gradients to first order. These techniques are applicable to the present work. However, since we want to put many stages on a single chip, we must devote the silicon area within the cell to those devices whose matching has the greatest effect on the system performance.

In the case of the silicon cochlea, the bias transistors are the most critical devices, since they have the greatest effect on the transconductance of the amplifiers, and thus the resonant frequency and quality factor of each section. Mirror-transistor mismatch also affects the transconductance value, but only half as much as the corresponding inaccuracy in the bias transistor. Mismatches in the differential-pair transistors and mirror transistors will also cause a random voltage offset in the output of each stage, but this offset is not a serious problem in many applications. It can be shown that the optimum control of transconductance is achieved for a given total area when the bias transistor is twice as large as the individual mirror transistors. In our test chips, other layout considerations make it convenient and useful to make the mirror and dif-
ferential-pair transistors small (6 μm × 6 μm) and to devote four to eight times as much area to bias transistors.

Controlling the Q of a second-order section implies controlling the value of \( \alpha = g_Q/(2g_a) \). Since the Q is such a sensitive function of \( \alpha \), it is reasonable to consider devoting silicon area to a common-centroid structure. A brute force approach would require duplicating each of the three bias transistors, laying the six transistors out in a hexagonal arrangement around a central point, and connecting the pairs together in parallel. However, a detailed analysis will reveal that, for equal capacitances in the second-order sections, it is important only for the transconductance \( g_Q \) to match the sum of the two feedforward transconductances, not the individual transconductances. So a more efficient approach which achieves the same objective is to duplicate only the feedback bias transistor, and juxtapose this pair of transistors with the feedforward bias transistors in a “pseudoquad” formation, as shown in Fig. 8.

A final improvement is related to the fact that in the original design, we are relying on an identical tilt on the \( V_a \) and \( V_Q \) lines to achieve a uniform \( Q \) at each stage. It is considerably easier to use the same tilted bias line \( V_a \), and to vary the transconductance of the feedback amplifier by varying the source of the bias transistor instead of the gate, as shown in Fig. 9; this scheme requires only one tilted polysilicon line \( (V_a) \) and a global nontilted \( Q \) control line which should be made of metal since it must supply the bias current from all the feedback amplifiers. This “\( Q \)-source control” would seem at first to be a trivial modification but it has a number of important benefits:

1) the tuning problem has been reduced from four degrees of freedom to three;
2) two input parameters \((V_{Q_0} \text{ and } V_{Q_1})\) which had to be tuned relative to other parameters \((V_a \text{ and } V_Q)\) have been replaced by a single absolute parameter \( Q_{\text{cont}} \), which is tuned with respect to ground, which simplifies chip testing considerably;
3) another source of error in \( \alpha \) has been eliminated, namely the fluctuations in the resistivity and dimensions of the \( V_Q \) polysilicon wire, which were uncorrelated with the fluctuations of the \( V_a \) polysilicon wire.

D. Compactness

It is possible to eliminate two redundant transistors in the original circuit by observing that the first feedforward amplifier and the feedback amplifier have a common output node. Whenever this situation occurs, a single current mirror can be shared between the two differential pairs, as shown in Fig. 10.

Another important observation relates to the conversion from fluid pressure to membrane velocity, which is usually done with a high-pass or bandpass circuit. It is apparent in Fig. 2 that the output current of the rightmost
Fig. 10. Sharing a current mirror between two amplifiers with a common output node. (a) The differential output currents from two transconductance amplifiers are summed at the common output node via Kirchoff's current law. (b) The branch currents in the differential pairs are summed before computing the difference via the current mirror. The two schemes are equivalent.

Fig. 11. Current copy techniques. (a) Direct current copy. (b) Scaled unidirectional current copy. The roles of $V_{\text{ref}}$ and $V_{\text{out}}$ are interchangeable, as shown, the circuit is designed to provide an amplified current copy for $V_{\text{ref}} < V_{\text{out}}$.

Fig. 12. Schematic of improved second-order section.

Feedforward amplifier is related to the output voltage of the section by $I_{\text{out}} = sCV_{\text{out}}$, i.e., the output current is proportional to the time derivative of the output voltage. There is no need to devote extra circuitry to the time-derivative operation, since it is already being done by the output amplifier, provided that we are willing to accept the result in the form of a current instead of a voltage. In order to use this current as the input to a subsequent computation, or to observe the current from off-chip, it is necessary to make a copy of it, as shown in Fig. 11(a). If a unidirectional current is acceptable, then a single transistor can be used to make a scaled copy of the current, as shown in Fig. 11(b). We often tilt the $V_{\text{scale}}$ line in order to give approximately equal peak current at each tap.

E. The Improved Design

Fig. 12 shows a schematic of an improved second-order section which incorporates all of the above circuit techniques, namely source-degenerated wide-range-input amplifiers in the forward direction only, $Q$-source control, large bias transistors in a pseudoquad configuration (as evidenced by the duplicate bias transistor in the feedback amplifier), shared current mirror, and a scaled unidirectional current copy.

V. Test Results

A typical silicon cochlea test chip is built in the MOSIS TinyChip frame (2.22 mm x 2.25 mm) in a standard double-poly double-metal 2-µm CMOS technology. Typical projects contain between 43 and 51 cochlea stages, in two rows of about 25 stages each.

The outputs are observed using a scanner (serial analog multiplexer), which serially switches each of the cochlea outputs onto a single pin [12]. The scanner allows the demonstration of real-time traveling waves, facilitates automated data acquisition, and allows larger designs without the need for more pins.

Most of our chips scan out both the voltage corresponding to fluid pressure and the current corresponding to membrane velocity. The output voltage of each second-order section is buffered by a fast voltage follower to pre-
vent excessive capacitive loading and switching noise which would disturb the cochlea operation. The currents are usually scanned out into an off-chip current-sense amplifier, which converts the current into a voltage signal.

In order to show the effectiveness of the circuit techniques in improving the performance of the silicon cochlea, results are presented for two chips—an early chip which used an unsophisticated layout, and a later chip with the improvements as described above. In the early layout, all transistors were 6 μm × 6 μm. In the improved layout, all transistors were 6 μm × 6 μm except for the bias transistors which were 12 μm × 12 μm.

In Fig. 13, the frequency response curves for all voltage signal taps are shown for both early layout and the
improved layout. Each curve represents the composite transfer function from cochlea input to the selected tap, including the effects of all taps in between.

The contribution from an individual tap is manifested in the difference in the log-scale plot of adjacent composite curves. Fig. 14 shows the derived individual transfer functions for the two designs. The severe effect of mismatching in the quality factor parameter is very evident in the data from the early layout; a few worst-case taps are clearly dominating the behavior of the entire cascade.

It is a simple matter to fit the canonical transfer function of (5) to each of the individual transfer functions. Fig. 15 shows a plot of natural frequency versus tap number for the two chips. The downward tilt in the plot corresponds to the decreasing natural frequency with distance into the cascade.

Fig. 16 shows a plot of quality factor $Q$ versus tap number for the two chips. The worst-case taps are clearly evident in the data from the early layout. In the later layout, the improved matching has made it possible to tune the typical $Q$ to a slightly higher value, with much better worst-case behavior than that of the early layout. The standard deviation about the mean has been reduced by about 40%.

Finally, Fig. 17 shows a plot of the current output signal corresponding to membrane velocity for each tap of the improved layout. Each curve has the desired asymmetric bandpass shape.

The power consumption of the chips is 7.5 mW with the scanner running at 1.2 kHz and the follower buffers biased above threshold. The power consumption of the 51-stage cochlea itself is 11 μW. 99.9% of the power consumption in normal operation is devoted to making the results of the computation externally observable.

VI. CONCLUSIONS

The silicon cochlea based on a cascade of second-order filter sections is capable of excellent performance and is more compact than any known alternative. A thorough investigation of the important design issues has led to layout and circuit design techniques that improve the performance of this subthreshold analog VLSI system to the point where high-level system behavior can be well controlled, even though the precision of the individual devices is limited.

ACKNOWLEDGMENT

The authors are pleased to acknowledge many helpful discussions with X. Arreguit and J. Lazzaro. We thank M. Sivilotti for suggesting the symbol for the wide-range-input amplifier. Chip fabrication was provided by the Defense Advanced Research Projects Agency and the MOSIS Service.

REFERENCES


Lloyd Watts received the B.Sc. degree in engineering physics from Queen's University, Kingston, Ont., Canada, in 1984. He received the M.A.Sc. degree from Simon Fraser University, Burnaby, B.C., Canada, in 1989, in the area of digital speech coding. Since 1989 he has been a doctoral student in electrical engineering at the California Institute of Technology, where he conducts research in the area of analog VLSI models of hearing.

From 1984 to 1987 he was with Microtel Pacific Research of Burnaby, B.C., Canada, where he worked on digital satellite communications, error control coding, and VLSI design.

Douglas A. Kerns received the B.S.E.E. degree in 1987 from Northwestern University in Evanston, IL, where he was also associated with the Biomedical Engineering Department. He received the M.S.E.E. degree in 1988 from the California Institute of Technology, Pasadena, and is currently working on the Ph.D. degree there.

As an undergraduate, he worked as a summer student technician at the Fermi National Accelerator Laboratory in Batavia, IL. During the summers of 1989 and 1990 he worked at the Jet Propulsion Laboratory in Pasadena, CA, designing analog microelectronics. He currently works part-time for Tanner Research in Pasadena, designing autocorrective microelectronics.

Richard F. Lyon (M'78) received the B.S. degree in engineering and applied science from California Institute of Technology, Pasadena, in 1974, and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 1975.

He has worked on a variety of projects involving communication and information theory, digital system design, analog and digital signal processing, VLSI design and methodologies, and sensory perception at Caltech, Bell Labs, Jet Propulsion Laboratory, Stanford Telecommunications Inc., Xerox PARC, and Schlumberger Palo Alto Research. In his current position at Apple Computer, he leads the Perception Systems group in applying models of hearing and in investigating brain-like computing approaches. He is also currently on the Computer Science faculty at Caltech, where he works on sensory modeling research and analog VLSI techniques.

Carver A. Mead, Gordon and Betty Moore Professor of Computer Science, has taught at the California Institute of Technology, Pasadena, for more than 30 years. He has contributed in the fields of solid-state electronics and the management of complexity in the design of very large-scale integrated circuits, and has been active in the development of innovative design methodologies for VLSI. He wrote with Lynn Conway the standard text for VLSI design, Introduction to VLSI Systems. His recent work is concerned with modeling neuronal structures, such as the retina and the cochlea using analog VLSI systems. His new book on this topic, Analog VLSI and Neural Systems, has recently been published (Addison-Wesley).

Prof. Mead is a member of the National Academy of Sciences, the National Academy of Engineering, a foreign member of the Royal Swedish Academy of Engineering Sciences, a Fellow of the American Physical Society, and a Life Fellow of the Franklin Institute. He is also the recipient of a number of awards including the centennial medal of the IEEE.