

BiCMOS Circuits for Analog Viterbi Decoders

Mohammad Hossein Shakiba, David A. Johns, *Senior Member, IEEE*, and Kenneth W. Martin, *Fellow, IEEE*

Abstract—Analog Viterbi decoders are finding widespread use in class-IV partial-response disk-drive applications. These analog realizations are often used because they are smaller and consume less power than their digital counterparts. However, class-IV signaling allows simplifications during Viterbi detection and thus existing analog decoders have limited applications. The purpose of this paper is to develop efficient analog circuits that can be used for general Viterbi detection. To demonstrate the feasibility of the proposed approach, the analog portions of two analog Viterbi decoders were fabricated in a 0.8- μm BiCMOS process. With an off-chip digital path memory, operation up to 50 Mb/s is demonstrated. However, simulations indicate that with on-chip digital path memory, speeds on the order of 300 Mb/s can be achieved. The power consumption of the proposed approach is estimated to be 15 mW/state drawn from a single 5-V power supply.

Index Terms—Analog, BiCMOS, communications, Viterbi.

I. INTRODUCTION

IN ESTIMATING a digital sequence, it is well known that if one received symbol contains information about other symbols, symbol-by-symbol detection will no longer lead to optimum performance [1]. This degradation arises when the symbol-by-symbol detector removes the effects of the other symbols in the detection process, hence, ignoring some useful information. For example, consider transmitting a sequence over a channel causing a time dispersion of the signal energy and hence, intersymbol interference (ISI). Since the interference contains information about the transmitted symbols, for optimum performance, the whole received sequence should be used to detect any symbol or group of symbols. Partial-response signaling (PRS) [2] falls in this category. Sequence detection of partial-response signals receives most of our attention in this paper.

The Viterbi algorithm (VA) [3] provides a practical means for realizing a maximum-likelihood sequence detection (MLSD) scheme. This technique was first proposed for decoding convolutional codes [4] and was later extended to the problem of optimum detection of digital sequences experiencing linear ISI [5], [6]. The basic idea is to consider the received sequence as a finite-state discrete-time Markov process contaminated by memoryless noise. A trellis diagram

is conceptually constructed by unwrapping the state diagram in time. The detector assigns a metric to each branch of the trellis, proportional to the error signal between the received value and the ideal signal resulting from that transition. The maximum-likelihood sequence is the one which results in the minimum accumulated error through the trellis. This approach is algorithmic in the sense that at each time step, for each one of the states of the trellis, the state metric (defined to be the accumulated error signal) is calculated using the previous state metrics and the branch metrics at that time step. In addition to state metrics, enough knowledge regarding the paths along which these optimum metrics have been obtained should also be saved. A block of memory, used with different management methods [7], can be utilized to save the required information. This information, stored in the form of digital sequences, enables the decoder to trace back the optimum paths ending on each state. Following the literature, we shall refer to this memory as path memory and its contents as survivor sequences.

Although the VA has been traditionally implemented in the digital domain, high-speed, small size, and low-power constraints have motivated researchers to look at analog realizations of the algorithm as well. Analog Viterbi decoders have demonstrated size and power advantages over their digital counterparts [8]–[14] and many of today's state-of-the-art partial-response maximum-likelihood read channels employ an analog Viterbi decoder in their processor core. In present partial-response class-IV systems, the power and size savings are mainly due to the elimination of the A/D converter which is commonly realized as a flash or interpolating architecture due to its high speed requirements. Similar tradeoffs occur in the choice of analog versus digital equalization where many of today's industry choose to perform analog equalization for power and size savings benefits.

The integrated analog decoders reported so far have been implemented based on some simplifications (but no approximations) made to the VA and are limited to certain applications. To extend the idea of analog realizations to general Viterbi decoders, a new approach should be taken. The purpose of this paper is to show that such an extension, with a simple circuit realization, is indeed possible. The proposed approach presented here is particularly applicable to the disk-drive industry, where an ever-increasing demand in the storage density calls for more sophisticated signal processing and, consequently, sequence detectors with increased numbers of states. Extended PRS Viterbi decoders are one example of such detectors.

Generalizing other analog solutions is not feasible in practice. The technique proposed in [15] is limited to a hard

Manuscript received October 31, 1996; revised January 8, 1998. This work was supported in part by the Canadian Network of Centres of Excellence in Microelectronics (MICRONET). This paper was recommended by Associate Editor B. Leung.

M. H. Shakiba was with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 3G4 Canada. He is now with Gennum Corp., Burlington, ON, Canada (e-mail: shakiba@eecg.toronto.edu).

D. A. Johns and K. W. Martin are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 3G4 Canada.

Publisher Item Identifier S 1057-7130(98)09976-5.

decision detection. An extension to soft detection results in a large number of diodes connected in series, which requires an unreasonably high-voltage supply and a need for an A/D converter, which eliminates the motivation for an analog realization altogether. Furthermore, the technique is nonalgorithmic and should not be used in decoders with deep path memories. It is also likely to be slow, since it requires a diode-configured circuit with many diodes in cascade to settle at each time step. The technique introduced in [16] can be generalized, but, with an unnecessary growth in the number of analog building blocks. This results in a size and power inefficient circuit realization.

The approach described here exploits the ability of simple analog circuitry to perform the required functions in the Viterbi decoder. Simplicity is considered a basic requirement since speed, size, and power consumption are the major concerns. The original idea was published in [17], however, with few circuit-level descriptions and analysis and no experimental results. This paper begins with a general overview of the approach, followed by details of a circuit-level implementation. An IC implementation of two decoders is described and experimental results for one decoder are given. The decoders were implemented on a common silicon core and the chip was fabricated in a 0.8- μm BiCMOS process. To save design time, digital path memories were not included on chip. As a result, experiments were carried up to 50 Mb/s. However, simulations indicate that with an on-chip path memory, speeds on the order of a few hundred-megahertz could be achieved. The power consumption of the decoder is estimated to be about 15 mW/state drawn from a 5-V single power supply.

It should be pointed out here that the implementation approach is general and can be used in error-correction coding (including convolutional codes), M -ary communications, and irregular trellises. For example, application of the technique to a quaternary dicode partial-response Viterbi decoder appears to be a likely candidate. Quaternary PRS has recently been proposed for high-rate data transmission over unshielded twisted-pair cables [18]. Also, due to its regularity, the implementation approach is well suited for automated design of analog Viterbi decoders. The analog-automation tool, which can extend down to layout [19], enables the analog design to be easily transferred from one technology to another.

II. GENERAL OVERVIEW

In an algorithmic realization of an MLSD, the metric assigned to each state of the trellis is updated using the previous state metrics and the branch metrics. Each branch in the trellis corresponds to a transition between two states with the branch metric representing the distance of the received symbol from the noiseless signal resulting from that transition. In the case of additive white Gaussian noise, the error (distance) criterion is based on minimizing the squared Euclidean distance. The update mechanism takes place such that the accumulated error of the estimated sequence for each state is minimized. In a trellis with $L = M^N$ states, there are M transitions initiating from or ending to each state, where M is the size of the

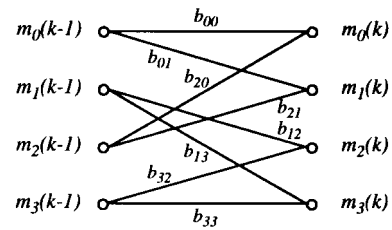


Fig. 1. A typical four-state trellis diagram.

alphabet and N is the memory size of the encoder.¹ As an example, Fig. 1 shows the trellis diagram for $N = 2$ and $M = 2$.

The arithmetic functions performed by the signal processor of the Viterbi decoder are of the form of add-compare-select (ACS). These operations can be described by the general expression²

$$m_i(k) = \max_j \{m_j(k-1) - b_{ji}(k)\}, \quad i = 0, 1, \dots, L-1 \\ j = 0, 1, \dots, L-1 \quad (1)$$

where $m_i(k)$ is the i th state metric at time step k and $b_{ji}(k)$ is the metric of the branch connecting state j at time $k-1$ to state i at time k . Note that for those values of i and j for which there does not exist a transition between states, branch metrics equal to infinity should be considered. This does not increase the implementation complexity and has been assumed for the generalization of the above mathematical expression.

The inputs to the algorithm are the branch metrics, which should be calculated based on the error criterion. For a square-Euclidean criterion, these metrics are of the quadratic form, but, can be reduced to linear combinations of the received sample and some constant values [20]. The idea is to expand the quadratic terms and cancel out the common terms, which are not required during the ACS operations. Also, a fixed gain might be considered for all of the branches in the trellis.

The above mathematical operations can be realized by the simple circuit depicted in Fig. 2 where the diodes have electrically adjustable turn-on or threshold voltages. In this circuit, there is a one-to-one correspondence between diode branches and the branches of the trellis diagram. In other words, each branch of the trellis is represented by a diode branch. The threshold voltage of each diode is proportional to the metric of its corresponding branch. Note that for those values of i and j for which there exists no transition between the states, the diode branches should simply be omitted. The previous state metrics are translated to voltage sources, driving the bus lines b_0 to b_{L-1} . The new values of state metrics appear as voltages at nodes 0 to $L-1$.

Assuming sharp I - V characteristics for the diodes, only one diode turns on in each set and conducts current. For junction diodes and the accuracy needed in a Viterbi decoder (typically 4–6 bits), this assumption seems reasonable. We shall return to this point when circuit implementation issues are addressed.

¹ Although only regular trellises fall in this category, the idea can easily be extended to irregular trellises as well.

² Note that here, equivalently, the \max function can be converted to a \min function with an inversion in the sign of the branch metrics.

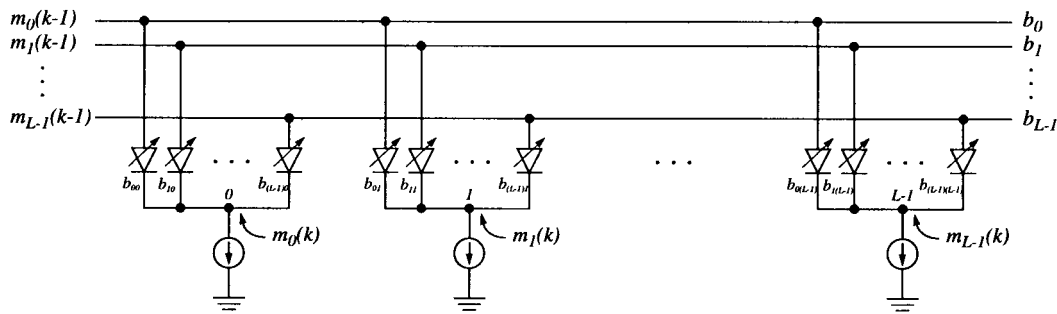


Fig. 2. The diode network used to realize the ACS function.

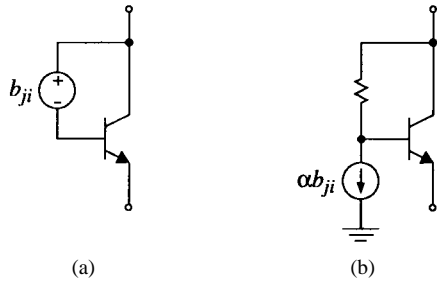


Fig. 3. A threshold-programmable diode. The floating voltage source in (a) is replaced by a resistor and a current source in (b).

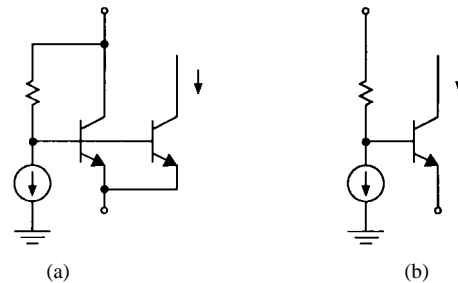


Fig. 5. Sensing the current of a programmable diode. (a) Employing a mirror transistor. (b) Directing the collector current.

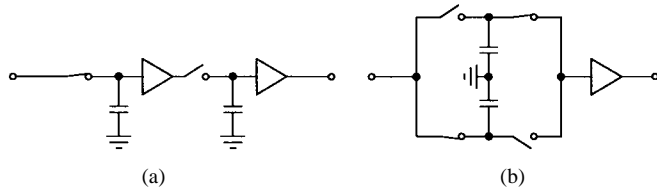


Fig. 4. A two-stage S/H: (a) master-slave and (b) ping-pong.

The threshold-programmable diode is shown in Fig. 3(a). This diode is composed of a diode-connected bipolar junction transistor (BJT) and a voltage source placed in the loop. Note that the negative swing of the voltage source should be limited to prevent saturation. Also, note that with the sharp $I-V$ characteristic assumption, the v_{BE} drops of the transistors will have negligible effect on the operation of the circuit. The floating voltage sources can be implemented by resistors fed by current sources as illustrated in Fig. 3(b). The current sources are proportional to the branch metrics and can be generated by combining the input sample with some dc values.

To avoid any destructive effects on the stored state metrics, two-stage sample-and-holds (S/H's) should be employed when feeding the new metrics back. While one stage is in track mode, sampling the new value of its corresponding state metric, the other is in the hold mode, holding the previous value across the bus line. Although master-slave S/H's can also be used, ping-pong S/H's are preferred, as they provide the potential of doubling the speed. This is particularly important as the speed of operation is mainly limited by S/H's. Fig. 4 illustrates the idea. Note that the output buffer should provide a very small output impedance as it will be used to drive one of the bus lines in Fig. 2. Also, note that one such circuit is required per state.

As seen from (1), unbounded growth of the state metrics is an inherent problem of the VA. To reduce the required dynamic

range in internal calculations, several methods for normalizing the state metrics have been proposed in digital decoders [21]. Taking an average over the state metrics and setting it in each iteration to a desired value minimizes the required dynamic range. This optimum solution, costly for a digital realization, works well in this analog approach by employing fast common-mode feedback (CMFB) circuitry. The CMFB operates on the state metrics by continuously monitoring these metrics and maintains a constant common-mode voltage for them. In other words, a CMFB circuitry is employed similar to those used in fully differential analog circuits but rather than setting the common-mode voltage of only two voltages, the common-mode voltage of all the state-metrics is maintained at a constant value. It should be mentioned here that in the case of a Viterbi decoder with a large number of states, there may still be a difficulty in accommodating the dynamic range of the state metrics. In other words, although the CMFB sets the average of the state metrics to a known value, the circuit must be able to handle the smallest path metric while not saturating the largest path metric.

The comparison results are the currents through the diodes. It is necessary to find all the L conducting diodes, each from individual groups in Fig. 2, which carry the branch currents. To sense the branch currents, either mirror transistors can be used or the collector currents of the diodes can be directed rather than being drawn from the buses. Fig. 5 depicts a programmable diode with current sensing.

In the mirror-transistor approach, the collector-to-emitter voltage variations of the different diodes are typically not as large as those in the other approach. This results in $V_{BE} - i_C$ characteristics less affected by the output circuits. However, this effect is usually negligible and the latter approach may be taken to reduce transistor count.

The sensed currents should be used to update the contents of the path memory. Memory management is beyond the scope of this paper; however, the register-exchanged method used in other analog implementations is straightforward and is often used in decoders with low number of states. To guarantee that digital levels will be developed from the sensed currents, the currents of the branches ending on each state should be compared and only the maximum current should be mapped to "one." In a binary scheme, the voltages at the bases of the diode transistors can be compared to generate the required digital data. This approach is advantageous since the comparison can be done by a simple comparator, eliminating the need for current sensing and comparing altogether. It should be noted here that since both examples presented later in this paper correspond to binary schemes, this voltage sensing approach was adopted.

The decoded data is obtained by tracing back that portion of the path memory which corresponds to the optimum survivor sequence. The optimum sequence reflects the state transitions ending on the state with the minimum metric value (minimum accumulated error). As a result, a multi-input comparator is also needed. Alternatively, one can choose any arbitrary state for a trace back. In general, this local-optimum trace back results in a detected sequence different from that of the global-optimum trace back. Toward the start, both sequences are most likely the same because of merging which takes place in the histories of different state transitions. However, toward the end of the sequences, the depth of the path memory decreases and the probability of diverging increases. As a result, with a deep path memory, both methods yield the same detected bits. In applications where the increase in the decoding delay is not critical, the local optimum trace back is often preferred since increasing the length of the path memory is straightforward.

Finally, it should be mentioned that each set of diodes in the diode network described here can be viewed as a generalized differential cell which compares the base voltages and directs the maximum input (with a v_{BE} drop) to the output taken from the common-emitter terminal. This equivalent approach was taken in [17].

III. CIRCUIT REALIZATION

Having introduced a general overview of the analog implementation approach, more details of the functional blocks will be presented and circuit-level issues will be addressed in this section.

A. Ping-Pong S/H

The ping-pong S/H, shown in Fig. 4(b), consists of two basic S/H's operating on a common input signal. A commutator directs the previous sample value, held by one of the S/H's, to the output, while the other S/H samples the new value of the input. The output buffer should have very high input impedance and very low output impedance. In a BiCMOS process, a combination of a MOS input-stage and a BJT output-stage is the best choice. Source follower and emitter follower stages were employed in our design. To keep the voltage swings within reasonable values, a PMOS source

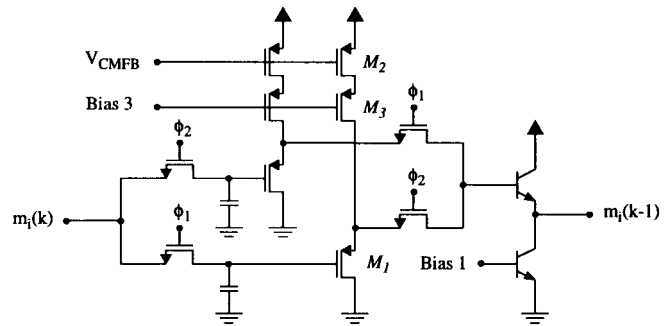


Fig. 6. The ping-pong S/H circuit. Input and output are the new and previous state metrics of state i .

follower was chosen to alternate the level shifts introduced by the two stages. As described shortly, the CMFB mechanism adjusts the level shifts of the source-followers by adjusting their biasing currents. It should be pointed out here that instead of one source follower at the output, two identical stages were used at the inputs of the commutator. This prevents charge transfer from the holding capacitors to the input capacitance of the MOS transistor. Also, note that due to the existence of the CMFB circuit, only the signal-dependent charge transfer between different state metrics in the S/H's is of concern when calculating new state metrics. With a holding capacitor of 0.5 pF used in our implementation, and the required accuracy, the effect of this signal dependent charge transfer is negligible and there is no need for using a charge cancellation technique. Fig. 6 depicts the S/H circuit. In this figure, ϕ_1 and ϕ_2 are two nonoverlapping phases obtained from a divide-by-two version of the master clock. The clock-generator circuit is discussed later in this section.

B. CMFB Circuit

The need for a CMFB control and its efficiency in minimizing the required dynamic range of the circuits was explained earlier. In our implementation, shown in Fig. 7, this mechanism is applied by continuously monitoring the common-mode (CM) signal of all of the state metrics and adjusting a value added to all of them to keep the CM signal equal to a reference voltage. This is accomplished by changing the level shifts of the source-follower buffers in the S/H circuits through adjusting their biasing currents in a continuous-time feedback loop.

C. Branch-Metric Generators

It was mentioned that the branch metrics (error signals) are usually expressed as linear combinations of the input samples and some dc values. In constructing these combinations, both polarities of their generating components are usually required and the combinations are needed in the form of currents. In an efficient realization, differential transconductors might be used to convert the input voltages to differential currents. The resulting signals, with appropriate polarities, can be combined by simply interconnecting the outputs of the transconductors. Since high linearity is usually not a

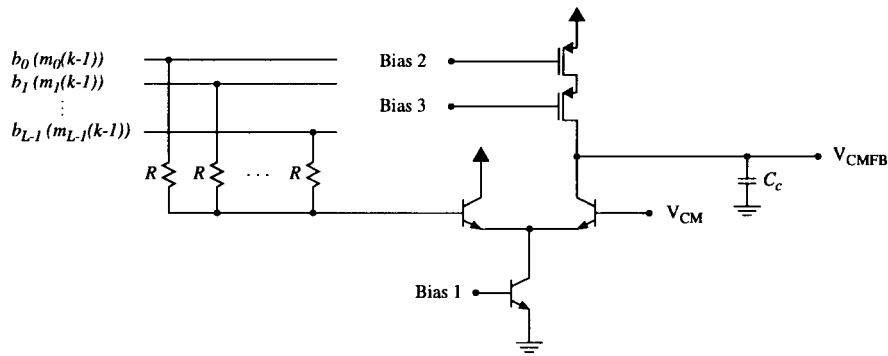


Fig. 7. The CMFB circuit.

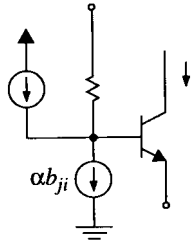


Fig. 8. Reducing the dc threshold of the diodes by partially (entirely) bypassing the dc components of the error signals.

requirement, differential pairs with degeneration are suggested because of their simplicity.

The error currents program the thresholds of the diodes shown in Fig. 2 by developing proportional voltages across the resistors depicted in Figs. 3 and 5. While the necessary signal voltages are developed across the various resistors, one should be careful to match dc voltage drops across the resistors of the various branch metric generators. Fortunately, this matching requirement is not severe as it only needs to satisfy the accuracy requirement of the Viterbi decoder (around 6 bit accuracy). In addition, to reduce the unnecessary dc voltage drops across these resistors (and even eliminate them altogether), constant current sources can be added to the combinations. These sources partially (entirely) bypass the biasing currents of the differential cells from the resistors. The technique reduces the operating-voltage requirement of the ACS circuitry and is suggested in low-voltage applications. Fig. 8 illustrates the idea. Note that since equal constant currents will be added to all of the branch metrics, there is no matching requirement between the sinking and sourcing current sources shown in this figure. Any mismatch between this sinking and sourcing current sources will only result in a residual dc voltage across the resistor which should be the same for all branch metric generators. However, the output impedances of these current sources should be considerably higher than the resistor value. This is usually the case in practice.

D. Clock Generator

A symbol-rate synchronous clock is needed to update the path memory at the end of each iteration. Such a clock is also required if comparators are employed after the ACS circuitry.

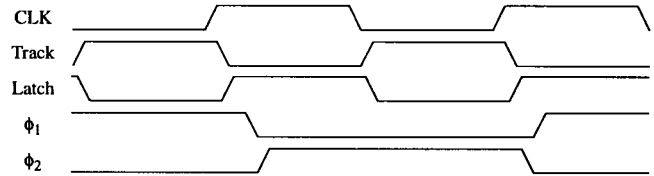


Fig. 9. A typical timing diagram for the clock generator.

Also, a divide-by-two version of the clock should be used to control the ping-pong S/H's. Two phases of the latter clock are required as shown in Fig. 6. Nonoverlapping clocks are essential to guarantee prevention of the destructive feedback explained earlier.

If comparators are used, their outputs should change only after the path memory is updated. Fig. 9 shows a typical timing diagram. With these timings, the path memory should be updated at the falling edges of the clock. Track and latch are used by the latched comparators. The nonoverlapping phases ϕ_1 and ϕ_2 are also shown. These signals toggle the S/H's after outputs of the comparators are latched.

The circuit depicted in Fig. 10 can be used to generate the required clocks from a single-phase clock. In this circuit, a T flip-flop, formed by feeding the complementary output of a D flip-flop back to its input, divides the frequency of the clock by two. The divided-by-two output is then converted to phases ϕ_1 and ϕ_2 by a nonoverlapping clock generator. To drive the D-flip-flop, two symbol-rate nonoverlapping signals and their complementary overlapping signals are required. These signals are generated by another nonoverlapping clock generator which works at the symbol rate. The symbol-rate clock signals required by the comparators are obtained from those outputs which satisfy the above timing diagram.

IV. DESIGN EXAMPLES

To investigate the feasibility of the proposed implementation approach, two PRS Viterbi decoders were designed and fabricated on a common silicon core. Due to availability of the experimental setup, a dicode decoder was chosen. Although this approach is not efficient in this special case,³ it was fabricated to prove the concept. As well, an extended PRS decoder was designed to show that the approach can easily

³For a preferred analog implementation of a dicode Viterbi decoder, the reader is referred to [14].

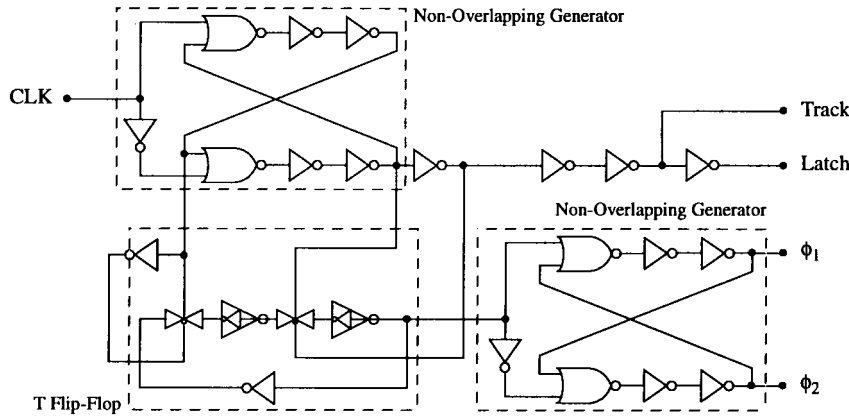


Fig. 10. The clock generator circuit.

be extended to decoders with considerably higher number of states. The extended partial-response (EPR4) scheme was chosen due to its future application in the read channels of magnetic recording systems. However, experimental results are not yet available due to limitations in the test equipment.

A. Binary Dicode

A dicode PRS system is a communication system in which the input data undergoes a $1 - D$ operation before being transmitted. Here, D denotes a unit delay and equals one symbol time. From this definition it can easily be verified that the trellis diagram of a dicode PRS scheme with binary inputs is a simple butterfly with the following branch metrics

$$b_{ji}(k) = (y(k) - (i - j))^2, \quad \begin{matrix} j = 0, 1 \\ i = 0, 1 \end{matrix} \quad (2)$$

where $y(k)$ denotes the received sample. Note that here the squared-Euclidean error criterion has been adopted.

Cancelling out the terms that do not depend on i or j in the above expressions and applying a gain of 0.5 to all of the branches yield the following equivalent branch metrics

$$b_{ji}(k) = (j - i) \left(y(k) + \frac{j - i}{2} \right), \quad \begin{matrix} j = 0, 1 \\ i = 0, 1 \end{matrix} \quad (3)$$

The trellis diagram and its diode network representation are shown in Fig. 11. The threshold voltage of each diode is proportional to the metric of its corresponding branch. Constant values are added to threshold voltages to prevent them from becoming negative.

Applying the implementation method to the above decoder, the circuit depicted in Fig. 12 results. In this decoder, the diode shown in Fig. 3(b) is used (i.e., without current sensing). Error signals are obtained by first converting the input signal $y(t)$ and the dc signals to currents and then combining the currents with appropriate polarities. The biasing currents of the transconductance cells are absorbed by bypass current sources to prevent flowing excess dc currents in the resistors. As a result, the threshold voltages of two cross-coupled diodes fluctuate around the levels generated by the dc components of the error signals plus the v_{BE} drops of the transistors. Two outer diodes have fixed threshold voltages of v_{BE} .

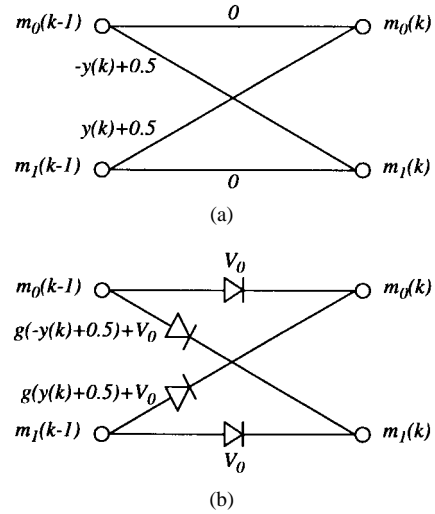


Fig. 11. (a) The binary dicode trellis diagram. (b) The corresponding diode network.

The transconductance cells and the bypass current sources connected to these diodes result in a net current of zero and are included to maintain matching.

An intuitive explanation was given earlier that the v_{BE} drops introduced by the “on” diodes will not have a significant effect on the performance of the decoder. To investigate the effect, consider the generic circuit shown in Fig. 13. Assuming the exponential characteristic $i_E = I_S \exp(v_{BE}/V_T)$ for the transistors, straightforward analysis of the circuit yields the following expression for v_e :

$$v_e = v_1 - e_1 - V_T \ln \left(\frac{I}{I_S} \right) + V_T \ln \left(1 + e^{(v_2 - v_1 + e_1 - e_2)/V_T} \right). \quad (4)$$

Equation (4) can be applied to two metric generators in the dicode decoder⁴ and the results can be adopted by simulations to illustrate the effects of $I-V$ dependency of the diodes on the performance of the decoder. Fig. 14 shows behavioral simulation results in which the bit-error rate (BER) performance is plotted for two branch gain values. The branch gain g , shown in Fig. 11, is the proportionality coefficient in translating the

⁴Note that I_S does not have any effect since the third term is a common term for all state-metric generators and can be discarded in calculations.

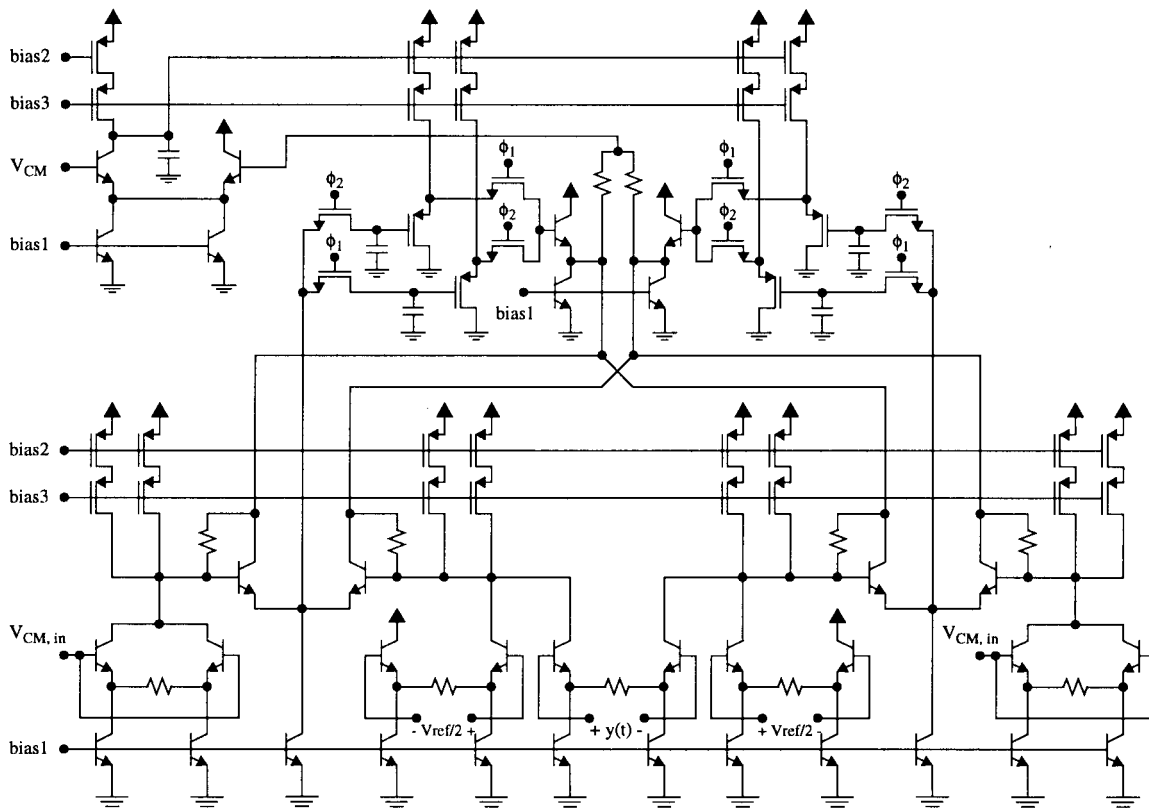


Fig. 12. The binary dicode decoder.

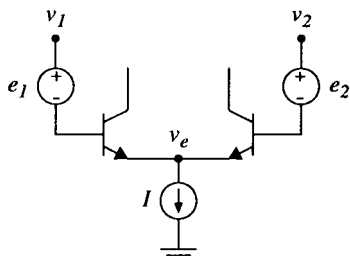


Fig. 13. The generic circuit for generating the new state metrics.

branch metrics to threshold voltages of the diodes and is set by the transconductances of the V/I converters and the resistors in the threshold-programmable diodes. In Fig. 14, the Viterbi bound and the performance of the symbol-by-symbol detector are also included for comparison.

Fig. 14 shows that performance degradation due to the differences in the v_{BE} of different transistors (which result if the “on” diodes are partially on and do not carry all of the tail currents) is indeed negligible. However, this degradation becomes considerable as the branch gain is reduced. This is expected, since decreasing the gain results in having v_{BE} differences comparable to the threshold voltages. This fact is further illustrated in Fig. 15, where the BER of the decoder at $SNR = 12$ dB is plotted as a function of the branch gain. As seen from this figure, for very small gains the performance suddenly drops, however, even for practically small gains the decoder nearly achieves its optimum performance. A gain of 0.2 develops threshold voltages in the range of -0.1 to 0.3 V, whereas much higher swings can be handled even with the transistors in diode-connected configuration. Had the collector

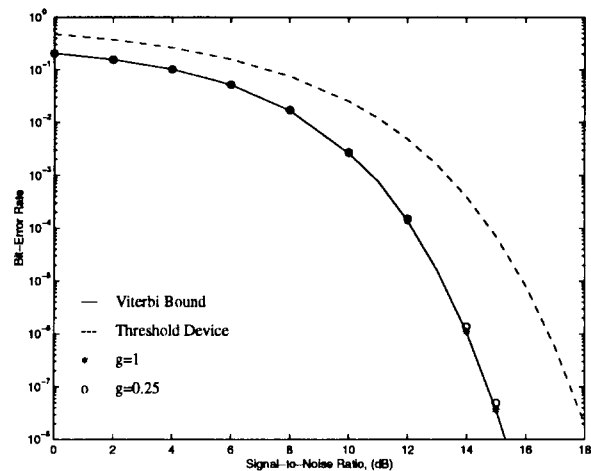


Fig. 14. Effect of the nonideal $I-V$ characteristics of the diodes on the performance of the dicode decoder for two branch gains.

of the transistors been connected to higher potential levels, higher-swing threshold voltages could have been established.

In the two-state dicode decoder, two S/H 's are used to feed the new state metrics back to the bus lines. The CMFB circuit keeps the CM signal of these lines equal to a CM reference voltage. In our implementation, a 3-PF compensation capacitor was used, resulting in a Q factor of 1.15 and a pole frequency of 300 MHz. Figs. 16 and 17 show SPICE simulations for the open- and closed-loop responses, respectively. Note that the fast roll-off in the frequency response, resulting from high frequency poles, can indeed be neglected in the frequency range of our interest (up to at least 300 MHz).

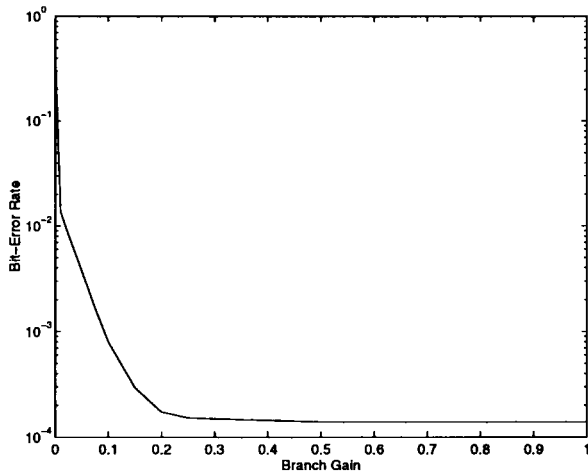


Fig. 15. BER performance of the dicode decoder as a function of the branch gain at SNR = 12 dB.

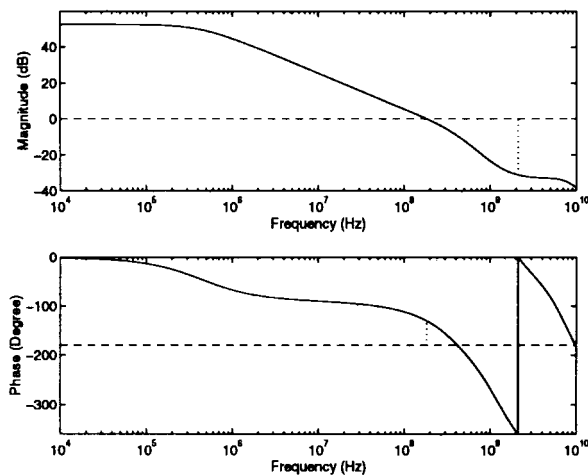


Fig. 16. Bode plot of the CMFB circuit in the binary dicode decoder obtained by SPICE.

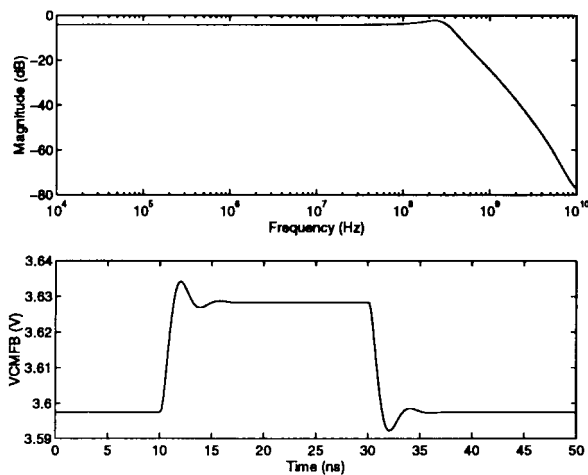


Fig. 17. Frequency and time domain SPICE simulation results of the closed-loop CMFB circuit.

The comparison results can be taken from the bases of the diode transistors. Two comparators were employed to compare the signals and output the digital information to update the

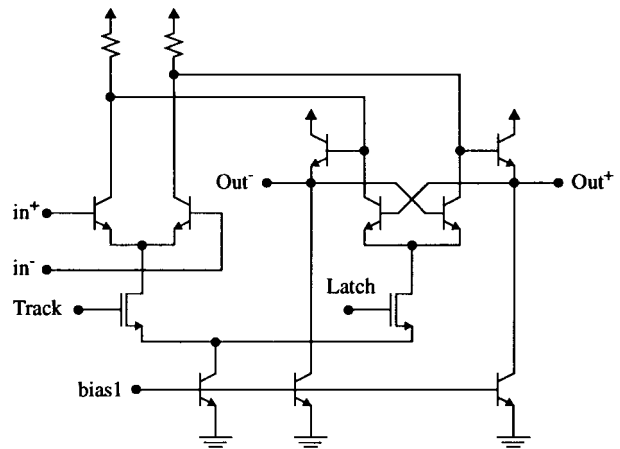


Fig. 18. The latched-comparator circuit.

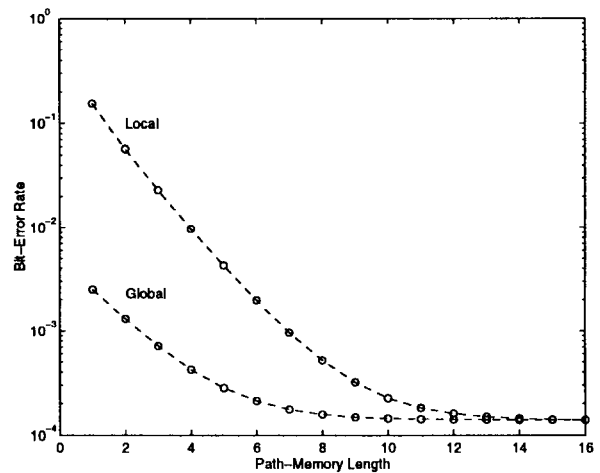


Fig. 19. BER performance of the dicode decoder at SNR = 12 dB as a function of the path-memory depth, when local and global optimum sequences are traced back in detection.

path memory. Fig. 18 depicts the comparator circuit. The circuit consists of a preamplifier stage, which amplifies the input in the track mode, and a latch which develops two complementary outputs, when the comparator is switched to the latch mode.

The decoded data would be available once the contents of the path memory are traced back. The local-optimum trace-back method was chosen in our implementation. As was mentioned earlier, the performance degradation associated with this trace-back method, compared to a global-optimum trace back, can be compensated by an increase in the length of the path memory. Fig. 19 depicts a typical simulated BER performance of the dicode Viterbi decoder in both cases as a function of the path-memory length.

B. Binary EPR4

It has been shown that at high densities a class of EPR schemes can be utilized to resemble the spectrum of the read signal in a magnetic recording system [23]. Consequently, much effort has been directed toward an efficient implementation of their sequence detectors. EPR4, an EPR scheme with

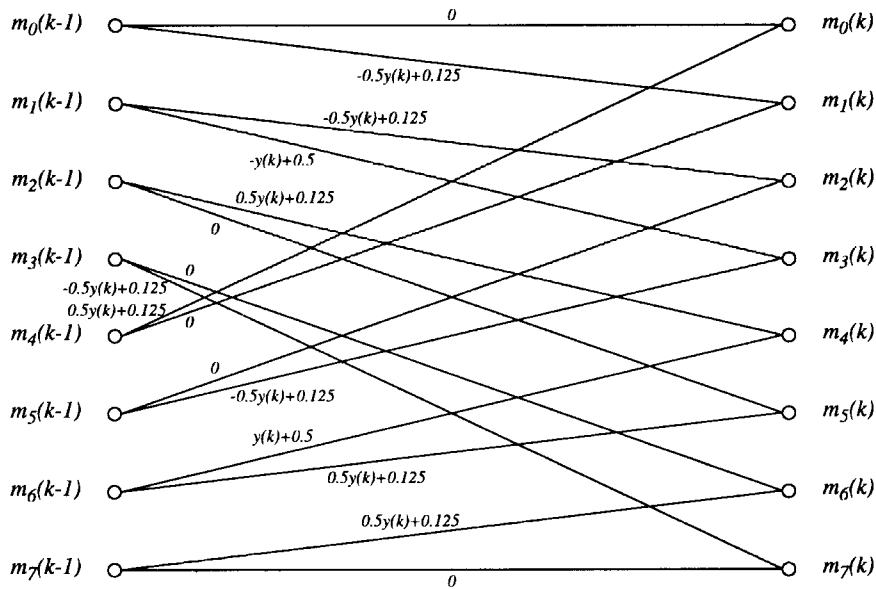


Fig. 20. The trellis diagram of the binary EPR4 signaling scheme.

four ISI terms, is one of the most promising schemes for this application. This system is expressed by the following coding polynomial:

$$F(D) = (1 - D)(1 + D)^2. \quad (5)$$

With binary input (as is the case for saturated magnetic storage systems such as disk drives), the above coding polynomial results in the eight-state trellis diagram shown in Fig. 20, where the branch metrics were derived in a manner similar to what was done for the dicode trellis.

Application of the implementation technique to the above trellis is an extension to the dicode decoder design. The procedure is straightforward and will not be repeated here.⁵ Eight binary outputs update the contents of the path memory which, in a register-exchange configuration, consists of eight interconnected shift registers. The decoded signal can be obtained by tracing the contents of any one of the shift registers back to its very first stage. Here, the local-optimum trace back is chosen in favor of reducing the complexity by eliminating the need for an eight-input comparator. Again, a deeper path memory compensates the degradation.

V. INTEGRATED-CIRCUIT IMPLEMENTATION

An IC containing both the dicode and the EPR4 decoders was designed and fabricated in a 0.8- μm BiCMOS process.⁶ To save design time, the digital path memories were not included on the chip. Rather, comparators were employed to provide the serial and parallel loading controls for updating the off-chip path memories. To be able to drive the pads and the external circuitry, the comparator outputs were driven off-chip through open-collector differential pairs. The idea was to maintain flexibility in adjusting the output impedances, the

⁵For the same compensation capacitor, however, a slightly more damped CMFB loop results.

⁶The Northern Telecom BATMOS process, available through Canadian Microelectronics Corporation.

driving capabilities, and the signal swings at the outputs of the chip. Fig. 21 illustrates the layout.

The inputs to the decoders are the signals needed to generate the branch metrics. These are $y(t)$ and $V_{\text{ref}}/2$, for the dicode and $y(t)$, $y(t)/2$, $V_{\text{ref}}/2$, and $V_{\text{ref}}/8$, for the EPR4 decoder. Obviously, $y(t)/2$ and $V_{\text{ref}}/8$ can be obtained from $y(t)$ and $V_{\text{ref}}/2$ in a real situation. The common-mode reference voltage was also input to the decoders.

VI. EXPERIMENTAL RESULTS

Since the digital path-memory was left off-chip, the experiments were conducted at speeds much lower than what could have been achieved with a fully integrated decoder. The path memory was implemented by wire-wrapping ECL shift registers. Fig. 22 illustrates the experimental results at a speed of 50 Mb/s. From this figure, it can be seen that the decoder performance indeed follows that of a Viterbi decoder. The sudden drop in the performance at high SNR (low BER) is caused by the truncated path memory and the local-optimum trace back. This deviation is not related to our implementation approach and is also observed in system-level simulations.

VII. CONCLUSIONS

Analog integrated Viterbi decoders have already demonstrated many advantages over the conventional digital realizations. The benefits can mainly be summarized in achieving faster decoders with significant savings in the size and power consumption. However, the reported implementations are limited to certain applications. In this paper it was shown that these advantages could be extended to other Viterbi decoders. A general approach to implementing Viterbi decoders in the analog domain, which is not based on any simplification to the decoding algorithm, was proposed. After describing the approach, its basic building blocks were considered in detail and some circuit-level issues were discussed. Simplicity of the

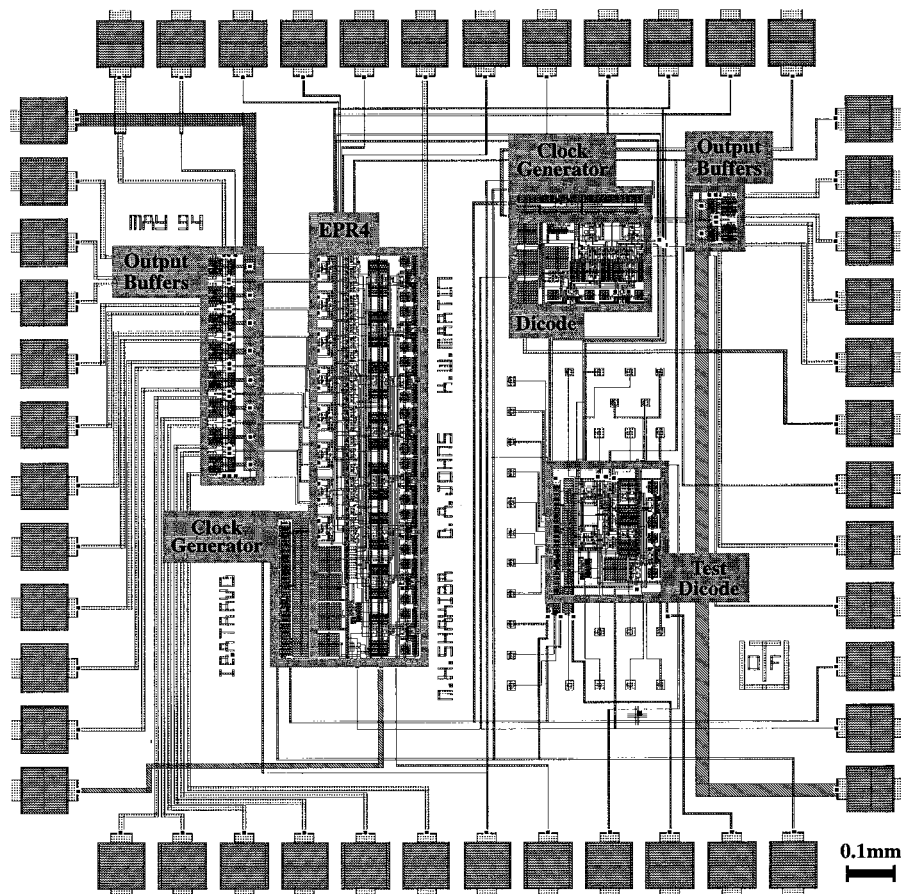


Fig. 21. The layout of the dicode and EPR4 analog Viterbi decoders, fabricated in a $0.8\text{-}\mu\text{m}$ BiCMOS process.

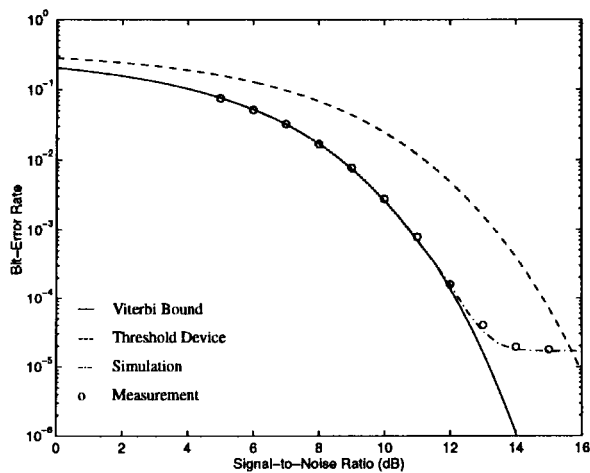


Fig. 22. Measured BER performance of the dicode decoder at 50 Mb/s. Note that a relatively short path memory (12 bits) with the local-optimum trace-back method has been used.

implementation was a basic requirement since speed, size, and power consumption were the major concerns.

A proof-of-concept dicode Viterbi decoder was designed and fabricated in a $0.80\text{-}\mu\text{m}$ BiCMOS process. To save design time, the digital path memory was not included. With an off-chip path memory, speeds in the order of tens of megabits-per-second were achieved. With an on-chip path memory, however, higher speeds should be attainable. This

was confirmed by simulations, where speeds in the order of few hundreds of megabits-per-second were observed. The power consumption of the decoder was measured to be about 15 mW/state drawn from a 5-V single power supply.

To illustrate the ease of extending the implementation approach to more complicated decoders, an EPR4 Viterbi decoder was also fabricated on the same chip. This particular extended partial-response scheme was chosen since it is presently finding its first application in the disk-drive industry.

Finally, it should be mentioned that the implementation approach is general and can be used in error-correction coding systems, M -ary communications, and irregular trellises. The quaternary dicode scheme is a good example where application of the present approach in realizing its Viterbi decoder seems promising. This scheme is recently receiving some attention in high-rate data transmission over twisted-pair cables.

REFERENCES

- [1] E. A. Lee and D. G. Messerschmitt, *Digital Communication*. Boston, MA: Kluwer, 1994.
- [2] P. Kabal and S. Pasupathy, "Partial-response signaling," *IEEE Trans. Commun.*, vol. COM-23, pp. 921-934, Sept. 1975.
- [3] G. D. Forney, Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, pp. 268-278, Mar. 1973.
- [4] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 260-269, Apr. 1967.
- [5] H. Kobayashi, "Correlative level coding and maximum-likelihood decoding," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 586-594, Sept. 1971.

- [6] G. D. Forney, Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 363–378, May 1972.
- [7] C. M. Rader, "Memory management in a Viterbi algorithm," *IEEE Trans. Commun.*, vol. COM-29, pp. 1399–1401, Sept. 1981.
- [8] A. S. Acampora and R. P. Gilmore, "Analog Viterbi decoding for high speed digital satellite channels," *IEEE Trans. Commun.*, vol. COM-26, pp. 1463–1470, Oct. 1978.
- [9] T. Suzuki and Y. Saitoh, "A 100Mbps optical transmission experiment employing a Viterbi decoder composed of analog circuits," in *Proc. URSI Int. Symp. Signals, Systems, and Electronics*, 1989.
- [10] T. W. Matthews and R. R. Spencer, "An analog CMOS Viterbi detector for digital magnetic recording," in *Proc. IEEE Int. Solid-State Circuits Conf.*, 1993, pp. 214–215.
- [11] ———, "An integrated analog CMOS Viterbi detector for digital magnetic recording," *IEEE J. Solid-State Circuits*, vol. 28, pp. 1294–1302, Dec. 1993.
- [12] R. G. Yamasaki, *et al.*, "A 72Mb/S PRML disk-drive channel chip with an analog sampled-data signal processor," in *Proc. IEEE Int. Solid-State Circuits Conf.*, 1994, pp. 278–279.
- [13] M. H. Shakiba, D. A. Johns, and K. W. Martin, "Analog implementation of class-IV partial-response Viterbi detector," in *Proc. IEEE Int. Symp. Circuits and Systems*, 1994, vol. 4, pp. 91–94.
- [14] ———, "A 200 MHz 3.3 V BiCMOS class-IV partial-response analog Viterbi decoder," in *Proc. IEEE Custom Integrated-Circuits Conf.*, 1995, pp. 567–570.
- [15] R. C. Davis, "Diode-configured Viterbi algorithm error correcting decoder for convolutional codes," U.S. patent 4545 054, Oct. 1, 1985.
- [16] A. Acampora, "Decoder for implementing an approximation of the Viterbi algorithm using analog processing techniques," U.S. patent 4087 787, May 2, 1978.
- [17] M. H. Shakiba, D. A. Johns, and K. W. Martin, "General approach to implementing analogue Viterbi decoders," *Electron. Lett.*, vol. 30, no. 22, pp. 1823–1824, Oct. 27, 1994.
- [18] G. Cherubini, S. Olcer, and G. Ungerboeck, "A quaternary partial-response class-IV system for 125 Mbit/s data transmission over unshielded twisted-pair cables," in *Proc. IEEE Int. Conf. Communications*, 1993, vol. 3, pp. 1814–1819.
- [19] B. R. Owen *et al.*, "BALLISTIC: An analog layout language," in *Proc. IEEE Custom Integrated-Circuits Conf.*, 1995, pp. 41–44.
- [20] J. F. Hayes, "The Viterbi algorithm applied to digital data transmission," *IEEE Commun. Mag.*, vol. 13, no. 2, pp. 15–20, Mar. 1975.
- [21] C. Shung *et al.*, "VLSI architectures for metric normalization in the Viterbi algorithm," in *Proc. IEEE Int. Conf. Communications*, 1990, vol. 4, pp. 1723–1728.
- [22] D. A. Johns and K. W. Martin, *Analog Integrated Circuit Design*. NY: Wiley, 1997.
- [23] H. K. Thapar and A. M. Patel, "A class of partial response systems for increasing storage density in magnetic recording," *IEEE Trans. Magn.*, vol. MAG-23, pp. 3666–3668, Sept. 1987.



Mohammad Hossein Shakiba was born in Hamedan, Iran, in 1960. He received the B.Sc. and M.Sc. degrees from the Isfahan University of Technology, Iran, in 1985 and 1988 and the Ph.D. degree from the University of Toronto, Canada, in 1997, all in electrical engineering.

From 1986 to 1991, he worked as a Research and Teaching Member in the Department of Electrical and Computer Engineering, Isfahan University of Technology. While working on the Ph.D. degree, he held the position of Research Fellow in the

Department of Electrical and Computer Engineering, University of Toronto. During this time, his research was focused on the area of high-speed data communication systems. His doctoral work was related to analog implementations of Viterbi detectors with emphasis on the partial-response systems. He is currently interested in analog integrated circuit design for wired and wireless data communication systems. He is currently with Gennum Corporation, Burlington, ON, Canada.



David A. Johns (S'81–M'89–SM'94) received the B.A.Sc., M.A.Sc., and Ph.D. degrees from the University of Toronto, Canada, in 1980, 1983, and 1989, respectively.

From 1980 to 1981, he worked at Mitel Corp., Ottawa, ON, Canada, while from 1983 to 1985 he was an Analog Microchip Designer at Pacific Microcircuits Ltd., Vancouver, BC, Canada. His doctoral work focused on analog and digital adaptive filters including the development of an orthonormal structure for analog filters. In 1988, he was hired

at the University of Toronto, where he is currently a Full Professor. He has ongoing research programs in the general area of analog integrated circuits including filters, oversampling, and data converters. His more recent work is directed toward circuits and systems for digital communications over wired, magnetic, and infrared channels. His research work has resulted in more than 40 publications as well as one textbook entitled *Analog Integrated Circuit Design* (New York: Wiley, 1997), which was co-authored by K. Martin. He has been involved in numerous industrial short courses as well as consulting for various companies such as Brooktree, Lucent, IBM, and others. He served as an Associate Editor for *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, PART II* from 1993 to 1995 and for *Part I* from 1995 to 1997.



Kenneth W. Martin (S'75–M'80–SM'89–F'91) received the B.A.Sc., M.A.Sc., and Ph.D. degrees from the University of Toronto, Canada, in 1975, 1977, and 1980, respectively.

From 1977 to 1978, he was a member of the Scientific Research Staff at Bell Northern Research, Ottawa, Canada, where he did some of the early research in integrated, switched-capacitor networks. Between 1980 and 1992, he was consecutively an Assistant, Associate, and Full Professor at the University of California at Los Angeles. In 1992,

he accepted the endowed "Stanley Ho Professorship in Microelectronics" at the University of Toronto. He has also been a consultant to many high-technology companies including Xerox Corp., Hughes Aircraft Co., Intel Corp., and Brooktree Corp. in the areas of high-speed analog and digital integrated circuit design. He has ongoing research programs in the areas of integrated circuits and systems. Recently, most of his research has focused on data-communication systems, both wired and wireless, as well as CAD for analog IC design. Recently, he completed, along with co-author D. Johns, a textbook entitled *Analog Integrated Circuit Design* (New York: Wiley, 1997). In addition, he has co-authored three other books in cooperation with former Ph.D. students.

Dr. Martin was appointed as the Circuits and Systems IEEE Press Representative (1985–1986). He was selected by the Circuits and Systems Society for the Outstanding Young Engineer Award that was presented at the IEEE Centennial Keys to the Future Program in 1984. He was elected by the Circuits and Systems Society to their Administrative Committee (ADCOM 1985–1987), and as a member of the Circuits and Systems BOG (1995–1997). He served as an Associate Editor of the *IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS* from 1985–1987, as an Associate Editor of the *IEEE PROCEEDINGS* (1995–1997), and has served on the Technical Committee for many International Symposia on Circuits and Systems. He was awarded a National Science Foundation Presidential Young Investigator Award (1985–1990). He was a co-recipient of the Beatrice Winner Award at the 1993 ISSCC.