Behzad Razavi

The Decision-Feedback Equalizer

The decision-feedback equalizer (DFE) dates back to the 1960s [1] and began to appear in high-speed wireline communication systems in the early 2000s. In this article, we study the properties of this circuit and describe its "analog" implementations.

The Need for Equalization

As high-speed random data propagates through a medium with a limited bandwidth (also called a "lossy" medium), it is dispersed. That is, the data edges become slower, possibly disallowing full transition if a 010 or 101 sequence occurs [Figure 1(a)]. This sluggishness of the channel also makes the zero crossing times of the data a function of the bit amplitudes, causing significant jitter. Both degradations increase the bit-detection error rate. In the frequency domain, the channel attenuates the high-frequency content of the data [Figure 1(b)].

The frequency-dependent channel loss depicted in Figure 1(b) can be undone by means of a circuit having the inverse response, i.e., a high-pass filter (HPF). As illustrated in Figure 1(c), if subjected to such a response, the received data assumes its original, undispersed shape and more easily lends itself to detection. This HPF exemplifies a "linear" equalizer as it can be approximated by a finite impulse response filter incorporating only linear stages (delay units and scaling coefficients). We say the equalizer provides a high-frequency "boost" to compensate for the channel loss.

Digital Object Identifier 10.1109/MSSC.2017.2745939
Date of publication: 16 November 2017

The Need for Decision-Feedback Equalization

While intuitively appealing, linear equalization faces three issues. First, since it requires a large amount of boost for very lossy channels, it significantly amplifies high-frequency noise, corrupting the data. Second, a high boost demands multiple stages, each one inevitably limiting the bandwidth and consuming considerable power. Third, the inverse response provided by a linear equalizer does not suffice in most practical cases. Specifically, a typical channel introduces, in the signal path, impedance *discontinuities* (mismatches) resulting from connectors and other physical interfaces between boards, cables, etc. Such discontinuities manifest themselves as deep notches in the channel's frequency response (Figure 2) that would be difficult to compensate by a linear equalizer.

To appreciate the beauty of DFEs, we first return to the time domain and view the data waveform as the superposition of random steps shifted in time by integer multiples of T_b [Figure 3(a)].

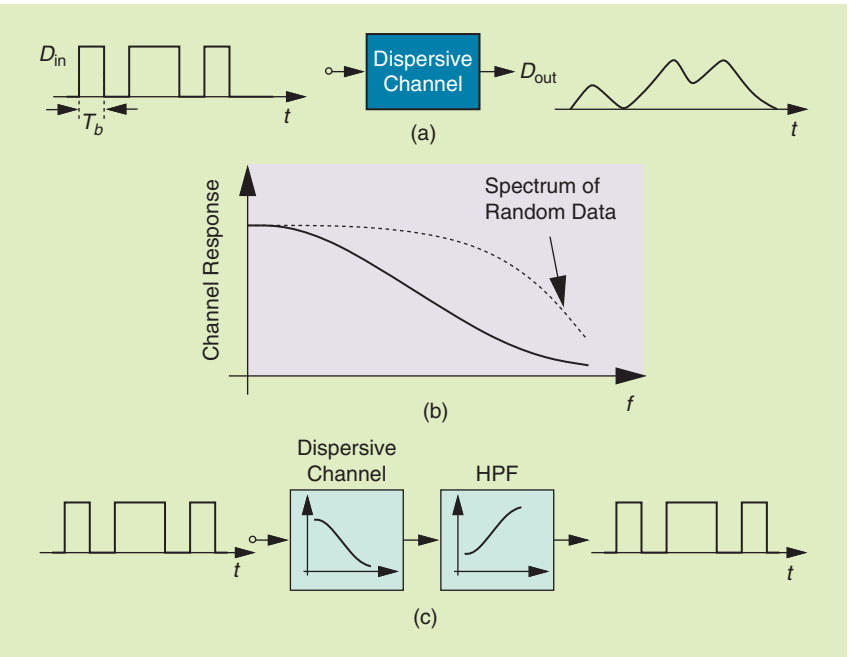

FIGURE 1: (a) The dispersion of random data in a lossy channel, (b) the channel frequency response showing attenuation of high-frequency components, and (c) the use of an HPF to equalize the channel.

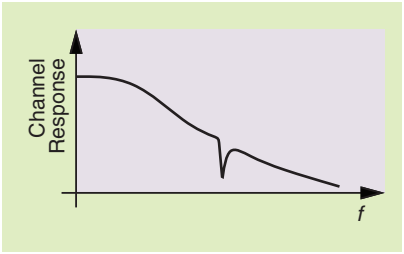

FIGURE 2: Notch in frequency response due to impedance discontinuity.

Due to the channel imperfections, each output bit is broadened, exhibiting a tail that interferes with the next bit(s). Called intersymbol interference (ISI), this phenomenon is more clearly seen in the impulse response of the channel [Figure 3(b)]. We observe that the tail values at T_b, $2T_b$, etc., (called the "postcursors") respectively represent the ISI introduced in the next bit, the bit after it, and so on. In general, energy removal from a signal's spectrum causes ISI, whether it occurs as wideband loss [Figure 1(b)] or as narrowband rejection (Figure 2). We surmise that ISI can be suppressed if we reconstruct the tail values and subtract them from the next bit(s).

Let us implement this idea for canceling the first postcursor in Figure 3(b). We must delay the present bit by one bit period, scale this result by a factor equal to h_1, and subtract it from the next bit. Figure 4(a) shows such an arrangement.

If we consider D_{in} as the present bit, D_{out} holds the previous bit and D_F a scaled copy thereof. Thus, $D_{in}-D_F$ is free from the first postcursor, whether it is created by wideband loss or impedance discontinuities. With a linear delay element, this loop is still a "linear" equalizer. A side effect is that the amplitude noise at the output is scaled by a factor of h_1 and added to the input data, degrading the signal-to-noise ratio of each bit.

This issue can be remedied if the delay element is followed by a limiter, also known as a "slicer," so as to remove the amplitude noise [Figure 4(b)]. The loop thus stops the noise from circulating and acts as a nonlinear equalizer. For robust operation, we replace the delay stage and the slicer with a flipflop (FF) [Figure 4(c)], recognizing that typical FFs provide both a one-period delay and limiting action on the amplitude. We can say the loop feeds the FF's decision back to the input, hence the term DFE. As illustrated in Figure 4(d), the summer output, D_{sum}, is free from the first postcursor, making greater voltage excursions with less jitter and allowing better detection. Sensed by the flipflop, this output is the most critical node in the circuit.

In the DFE loop of Figure 4(c), two parameters must be adjusted to reach optimum performance. First, the clock sampling edges must occur at the peaks of D_{sum}, necessitating a clock recovery circuit. Second, the first "tap" value, h_1, must be chosen according to the actual channel response. This is typically accomplished by monitoring the eye diagram at the summer output and adjusting h_1 to maximize its height.

Higher-order postcursors can also be removed by a DFE. Depicted in Figure 5, a two-tap realization returns scaled copies of the last two bits to the input.

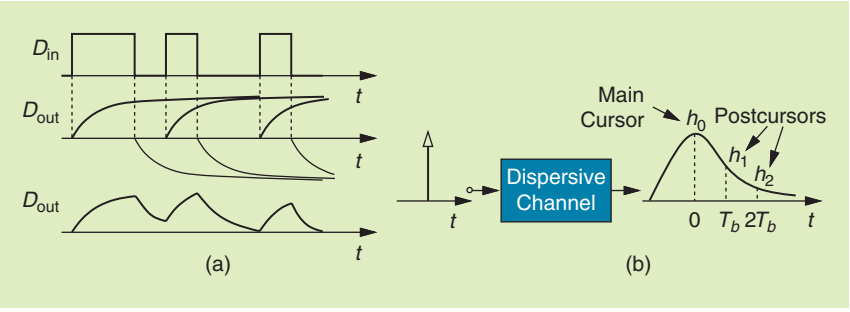

FIGURE 3: (a) Random data viewed as superposition of steps and (b) the impulse response of a lossy channel.

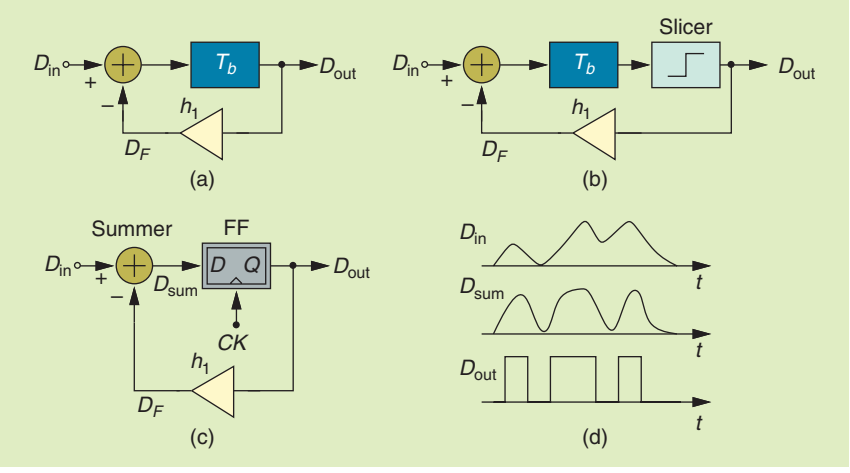

FIGURE 4: (a) A feedback loop canceling the first postcursor, (b) the addition of a slicer to suppress amplitude noise, (c) the use of an FF as a delay element and a slicer, and (d) the resulting waveforms.

Design Issues

In addition to clock phase alignment and proper setting of the feedback tap, the DFE shown in Figure 4(c) must also deal with the total loop delay. We predict that, at a sufficiently high data and clock rate, the circuit begins to incur errors.

To study the DFE speed limitations, consider the differential topology in Figure 6 and suppose the slave latch in the FF enters the sense mode on the falling edge of the clock, at $t = t_1$. The slave output requires a certain amount of time to change state, called the "clock-to-Q" delay, $T_{CK-Q} = t_2 - t_1$. This transition propagates through

the scaling stage and causes a change at the summing node. This node has a finite time constant, introducing its own delay, $T_{FB} = t_3 - t_2$. When *CK* goes high, the master latch enters the sense mode and must change its output according to the new value of D_{sum} before *CK* goes low again. The necessary time for this change is the setup time of the FF, $T_{setup} = t_5 - t_4$. Thus, $T_{CK-Q} + T_{FB} + T_{setup}$ must not exceed one clock cycle and hence one bit period:

$$T_{CK-Q} + T_{FB} + T_{setup} \leq T_b. \qquad (1)$$

Similar speed limitations exist in other variants of this architecture as well (see the "DFE Variants" section).

Three other nonidealities affect the performance of the DFE. First, the data input port of the summer in Figure 6 cannot be arbitrarily nonlinear because the dispersed data's amplitude carries information about the channel and must not experience significant limiting. We can see intuitively that, if the D_{in} waveform in Figure 4(d) is greatly amplified and sliced, then all of the bits exhibit a full swing but the jitter introduced by the channel remains. As a guideline, we choose the 1-dB compression point of this port to be greater than the main cursor amplitude [2] so that the nonlinearity negligibly increases the ISI.

Second, the input offset of the FF, V_{OS}, shifts the net voltage sensed at the summing node, D_{sum}, equivalently degrading the voltage margin for the negative or positive data values sampled by the FF. Third, the total noise in D_{sum}, V_n, yields a finite bit error rate. This noise includes that produced by the summer and the stages preceding the DFE and the input-referred noise of the FF. As a rule of thumb, we ensure that $8\,(4V_{OS} + V_{n,rms})$ remains less than the peak data swing D_{sum}. The factor of eight is chosen to ensure error rates on the order of 10^{-12} and the factor of four represents the four-sigma variance of the offset.

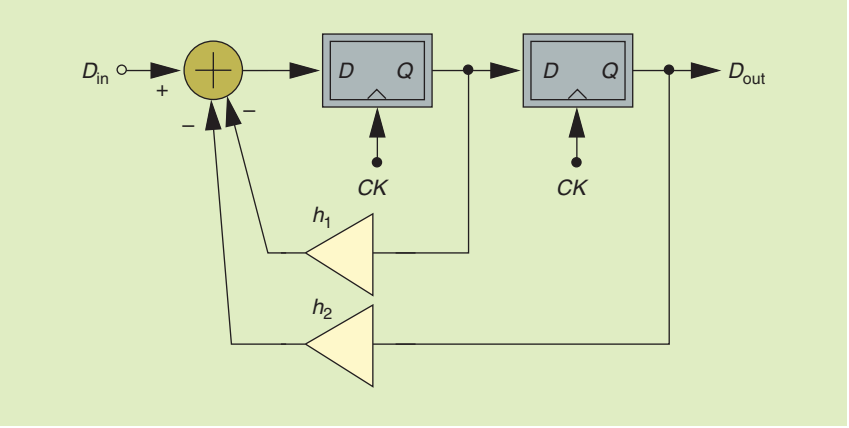

FIGURE 5: A two-tap DFE architecture.

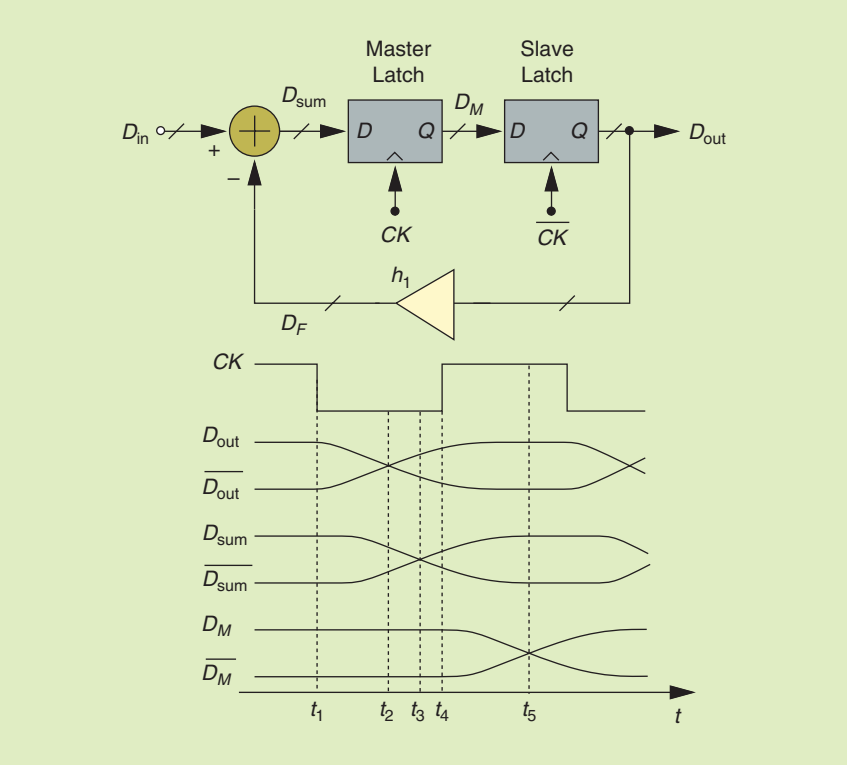

FIGURE 6: A DFE with differential signal paths.

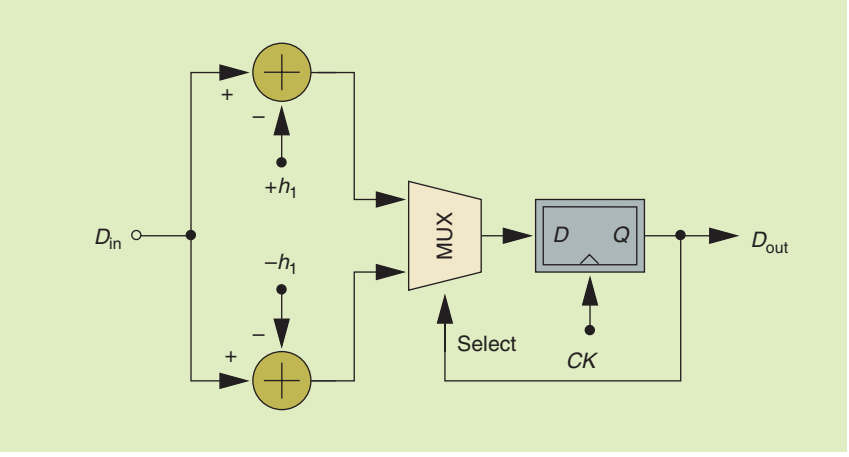

FIGURE 7: An unrolled DFE architecture.

DFE Variants

A multitude of DFE architectures have been proposed to ease the design tradeoffs. We study some here.

It is possible to transform the feedback loop of Figure 4(c) to a predictive or "unrolled" topology. Suppose D_{in} and D_{out} swing between –1 and +1. Since we wish to compute $D_{in} - h_1 D_{out}$, we can equivalently consider $D_{in} - h_1$ and $D_{in} + h_1$ as the only two possible levels that must reach the FF. The selection between these two values can be made by the previous bit. Figure 7 shows the resulting "unrolled DFE" [3]. Here, the previous bit available at D_{out} decides whether $D_{in} - h_1$ or $D_{in} + h_1$ must travel through the multiplexer and be sliced by the FF. We note that the summing nodes lie outside the feedback loop, which is the principal advantage of this arrangement. The timing budget is now given by $T_{CK-Q} + T_{setup} + T_{MUX} < T_b$, where T_{MUX} denotes the delay from the select input of the multiplexer to its output. In some cases, T_{MUX} is less than T_{FB} in (1). However, the D_{out} signal must be level shifted and/or amplified to properly switch the multiplexer, leading to additional delay.

At very high speeds, it is desirable to drive the DFE with a half-rate clock, $CK_{1/2}$, which is simpler to generate and distribute. Figure 8(a) shows a half-rate DFE [4], where the FFs are clocked by $CK_{1/2}$ and $\overline{CK_{1/2}}$, thereby demultiplexing the data by a factor of two. Each output bit lasts for $2T_b$ seconds and, after subtraction from D_{in}, is fed to the FF in the other branch. This topology nonetheless does not relax the loop timing budget given by (1). It also consumes about twice as much power as the full-rate DFE of Figure 4(c).

Another half-rate DFE architecture is depicted in Figure 8(b) [5]. Here, the half-rate outputs are multiplexed so as to reconstruct the full-rate data, with the result serving as the feedback signal. While using only one summer, this method adds the multiplexer delay to $T_{CK-Q} + T_{FB} + T_{setup}$, degrading the speed.

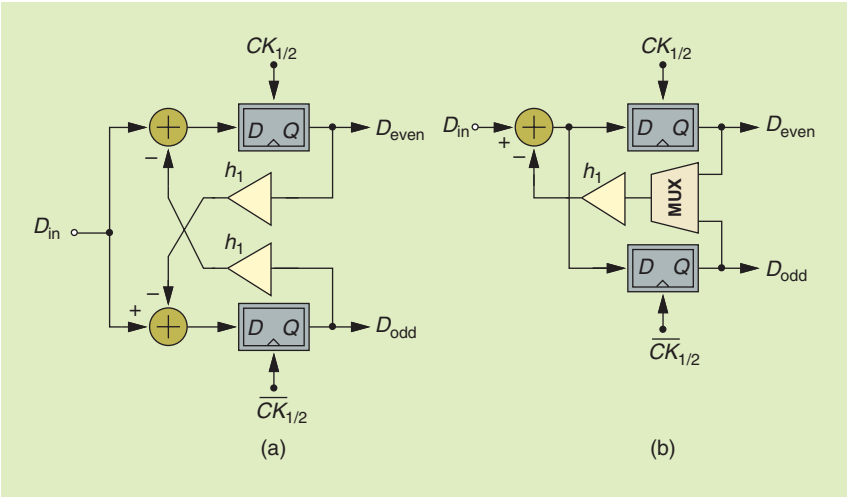

FIGURE 8: Half-rate DFE architectures with (a) two summers and (b) one summer and one multiplexer.

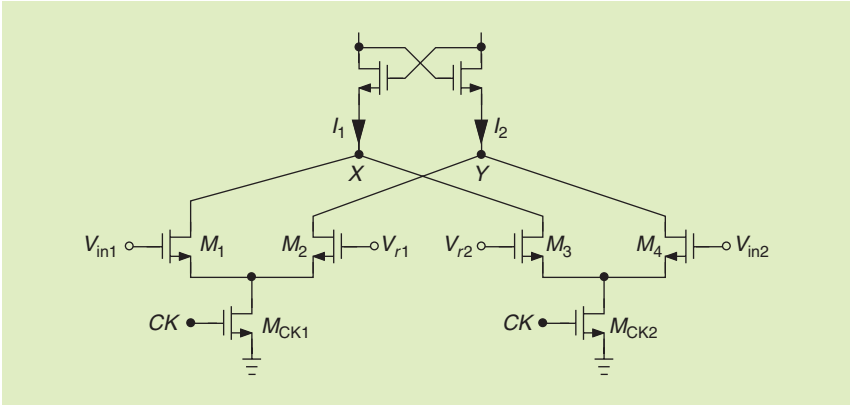

FIGURE 9: A comparator input stage based on two differential pairs.

At very high speeds, the summing node and the FFs can incorporate inductive peaking for a greater bandwidth and a smaller loop delay. This improvement comes at the cost of a more complex layout and signal distribution difficulties.

Questions for the Reader

1) Can the delay stage and the slicer in Figure 4(b) be realized as a single limiting differential pair?
2) Can the unrolled DFE of Figure 7 accommodate a second tap?

Answers to Last Issue's Questions

1) In Figure 9, why can we not apply V_{in1} and V_{in2} to M_1 and M_2 and V_{r1} and V_{r2} to M_3 and M_4?

   In such a case, each differential pair can experience a large input difference even when the comparator is making a critical decision. As a result, the transconductance of the two pairs falls considerably, making the offsets of the subsequent stages significant.
2) How does the characteristic shown in Figure 10(b) change if the front-end comparator has an offset equal to 1.5 least-significant bits (LSBs)?

   In the ideal case, we have $V_F^+ - V_F^- = V_{in}^+ - V_{in}^-$ if $V_{in}^+ - V_{in}^- > 0$ and $V_F^+ - V_F^- = -(V_{in}^+ - V_{in}^-)$ if $V_{in}^+ - V_{in}^- < 0$. With a comparator offset of 1.5 LSBs, the former holds if $V_{in}^+ - V_{in}^- > 1.5$ LSBs and the latter, if $V_{in}^+ - V_{in}^- < 1.5$ LSBs. That is, the circuit negates the differential input even for values reaching $= 1.5$ LSBs. The resulting characteristic is shown in Figure 10(c).

(continued on p. 132)

SSCS Vice President Bram Nauta Inducted to the Royal Dutch Academy of Arts and Sciences

IEEE Solid-State Circuits Society (SSCS) Vice President Bram Nauta was inducted into the Royal Dutch Academy of Arts and Sciences in June. Nauta is a professor at the University of Twente, heading the Integrated Circuits Design group. His current research interests are high-speed analog complementary metal-oxide-semiconductor circuits, software-defined radio, cognitive radio, and beamforming.

Academy membership is a great honor in The Netherlands. The academy appoints a maximum of 16 new members every year. Membership is awarded based on an individual's scientific and scholarly achievements.

Digital Object Identifier 10.1109/MSSC.2017.2746193
Date of publication: 16 November 2017

Bram Nauta

Once appointed, individuals are members for life. Members meet and discuss issues of interest to science, scholarship, and society. Academy members represent a wide spectrum of scientific and scholarly disciplines, giving all members the opportunity to embrace new fields in science and scholarship.

Nauta was inducted as a result of the work he performed throughout his career and it was a great honor. "It was a surprise for me," Nauta said, "especially because I'm an electrical engineer working on the application side of science."

He hopes his induction will open new doors for him, especially outside his own scientific field.

For more information about the Royal Dutch Academy of Arts and Sciences, visit https://www.knaw.nl/nl.

—Abira Sengupta

A CIRCUIT FOR ALL SEASONS

(continued from p. 16)

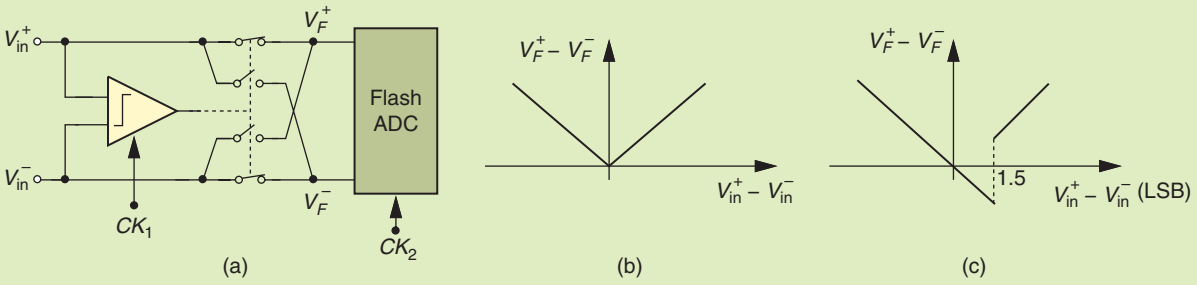

FIGURE 10: (a) The flash stage preceded by a polarity detector and (b) the resulting characteristic and the characteristic in the presence of comparator offset.

References

[1] M. E. Austin, "Decision-feedback equalization for digital communication over dispersive channels," Tech. Rep. 437, Lincoln Laboratory, Aug. 1967.

[2] J. Jung and B. Razavi, "A 25 Gb/s 5.8 mW CMOS equalizer," *IEEE J. Solid-State Circuits*, vol. 50, pp. 515–526, Feb. 2015.

[3] S. Kasturia and J. H. Winters, "Techniques for high-speed implementation of nonlinear cancellation," *IEEE J. Selected Areas Commun.*, vol. 9, pp. 711–717, June 1991.

[4] J. F. Bulzacchelli, M. Meghelli, S. V. Rylov, W. Rhee, A. Rylyakov, H. A. Ainspan, B. D. Parker, M. P. Beakes, A. Chung, T. J. Beukema, P. K. Pepeljugoski, L. Shan, Y. H. Kwark, S. Gowda, and D. J. Friedman, "A 10-Gb/s 5-tap DFE/4-tap FFE transceiver in 90-nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2885–2900, Dec. 2006.

[5] R. Payne, P. Landman, B. Bhakta, S. Ramaswamy, S. Wu, J. D. Powers, M. U. Erdogan, A. Yee, R. G. L. Wu, Y. Xie, B. Parthasarathy, K. Brouse, W. Mohammed, K. Heragu, V. Gupta, L. Dyson, and W. Lee, "A 6.25-Gb/s binary transceiver in 0.13-um CMOS for serial data transmission across high loss legacy backplane channels," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2646–2657, Dec. 2005.