Simultaneous Signature and Syndrome Compression

JOHN P. ROBINSON, SENIOR MEMBER, IEEE, AND NIRMAL R. SAXENA

Abstract—**An important organization for built-in self-test of VLSI circuits uses complete or psuedorandom test input generators followed by output data reduction. Two compression techniques which have been used are polynomial division (signature) and ones counting (syndrome). The simultaneous use of both of these approaches in parallel is investigated. Analytic and enumerative results indicate that the number of error patterns which are missed by both methods is nearly the theoretical minimum. The conclusion extends to signature compression combined with any other counter-based compression method such as Walsh spectral coefficients. Some suggestions for CAD implemented test design are given based on these results.**

I. INTRODUCTION

THE BUILT-IN self-test approach reduces the need for external test pattern generation or test input pattern storage. Exhaustive or extensive psuedorandom tests may be practical since parallel test activity is possible. Conceivably several chips on an undiced wafer could be under test in parallel. Electron beam scanning could check for valid test results. Compression methods can reduce the test data substantially.

Test data reduction may lead to information loss. Erroneous data sequences could be mapped into apparently correct behavior. With design aids, such as simulation, candidate compactors could be evaluated for effectiveness at the design stage. In this paper two compression methods, polynomial division and counting, are considered together. We show that their parallel use gives orthogonal behavior with respect to error patterns. Parallel compression is likely to be a viable test compaction structure.

Signature compression is a linear operation and thus partitions the space of binary m-tuples into regions, each having 2^{m-L} members, where L is the length of the signature register. This is represented by the equal-width bands in Fig. 1(a). All sequences in the same region have the same signature. For example if the compression polynomial is a primitive polynomial, then all those sequences in the zero signature zone are code words of a Hamming single error correcting code. All sequences with a specific signature form a coset of the Hamming code.

Syndrome compression is a nonlinear operation and partitions the space of binary m-tuples into unequally sized regions. The number of sequences having the same syndrome value w is just the number of ways w things can be chosen out of m or

$$\binom{m}{w}.$$

This is represented in Fig. 1(b) by the bands of varied width corresponding to different values for w. Each band represents those sequences which have the same syndrome.

When both signature and syndrome compression are used in parallel, certain misleading output data sequences can be confused with the correct output sequence. Such erroneous sequences or aliases would have the same signature, say S_j, as the correct test data and would also have the same number of ones, say w, as the valid sequence.

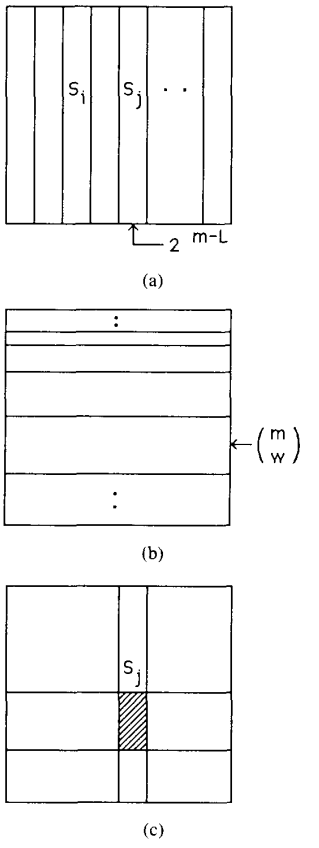

Fig. 1. Error region partitions. (a) Signature. (b) Syndrome. (c) Simultaneous.

Manuscript received December 13, 1985; revised June 11, 1987, and October 23, 1987. A preliminary version of this work was presented at the workshop "New Directions for IC Testing," Victoria, B.C., Canada, in March 1986. The review of this paper was arranged by Associate Editor J. A. Abraham.

J. P. Robinson is with the Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242.

N. R. Saxena is with Hewlett Packard, Cupertino, CA 95014.

IEEE Log Number 8718819.

0278-0070/88/0500-0584$01.00 © 1988 IEEE

A missed defective sequence would thus lie in the crosshatched area in Fig. 1(c). What we will show in this paper is that the syndrome region is almost equally divided up by the 2^L possible signatures. Hence, signature analysis will always reduce the error ambiguity of a counting-based compression method.

II. Maximum Length Case

Suppose that the length of the signature feedback shift register is L and that the feedback polynomial is primitive. If we assume that the test length is $m = 2^L - 1$, then there is enough structure to compute the exact overlap between a signature and a syndrome. Specifically we will determine the number of binary sequences of length $2^L - 1$ which have the same syndrome count and the same signature when the L stage signature unit uses a primitive polynomial for feedback.

A sequence with a syndrome count of w has w ones in a span of m; thus there are exactly

$$\binom{m}{w}$$

such sequences [6]. These sequences can be divided into rotational equivalence classes. Two sequences A and B are equivalent if A can be obtained from B by rotating B one or more steps. We next show that each member of an equivalence class has a distinct signature or all the signatures for the class are zero.

Notation: $A(x)$ is the polynomial corresponding to a sequence $A = a_0, a_1, \cdots, a_{m-1}$. The coefficient of x^i is a_i.

Example: If $A = 0100101$, then $A(x) = x + x^4 + x^6$.

Definition: Two sequences A and B are rotationally equivalent if $A(x) = x^t B(x)$, where all exponents are reduced modulo m.

All the members of a rotational equivalence class have the same number of ones or weight. Let w denote the numbers of ones in a sequence.

Theorem 1: Consider signature compression where:

i) The feedback corresponds to a primitive polynomial of degree L.
ii) The sequence length $m = 2^L - 1$.

All the members of a rotational equivalence class have signatures which are distinct and nonzero or all the signatures are zero.

Proof: Taking the last claim first, if a sequence has a zero signature then its polynomial is a multiple of the feedback polynomial [9]. Since the feedback polynomial is a primitive polynomial of degree L and $m = 2^L - 1$, all such zero signature sequences are code words from a Hamming code in cyclic form. All cyclic shifts of a code word yield a code word; thus if any sequence has a zero signature, then all members of the equivalence class have a zero signature.

Next consider the case where some signature is not zero. We assume that the claim is false and show that a contradiction results. Suppose that some sequence A has a nonzero signature and that a rotation of t steps has the same signature. Let S denote the nonzero signature. In all multiplications, exponents are reduced modulo L. By definition the signature $S(x)$ is the remainder found when $A(x)$ is divided by $P(x)$ or

$$A(x) = P(x)Q_1(x) + S(x)$$

where $P(x)$ is the feedback polynomial and the degree of $S(x)$ is less than L. Suppose some rotation, say t, of A has the same signature:

$$x^t A(x) = P(x)Q_2(x) + S(x).$$

Multiplying the first equation by x^t yields

$$x^t A(x) = x^t P(x)Q_1(x) + x^t S(x).$$

Equating these last two expressions and moving the terms involving $S(x)$ to the left,

$$(1 + x^t)S(x) = P(x)[Q_2(x) + x^t Q_1(x)]$$

$$\frac{(1 + x^t)S(x)}{P(x)} = [Q_2(x) + x^t Q_1(x)].$$

$S(x)$ is not zero by hypothesis and has degree less than L. Therefore $p(x)$ must divide $1 + x^t$ for the last equation to hold, where t is less than $2^L - 1$. By hypothesis $p(x)$ is primitive. Hence the smallest t where $p(x)$ divides $1 + x^t$ is $t = 2^L - 1$, which leads to a contradiction.

Q.E.D.

The number of zero signature sequences can be found from the weight generating function for a Hamming code (see [9, p. 90]). For binary codes,

$$G(x) = \frac{1}{n+1}\left[(1 + x)^n + n(1 + x)^a(1 - x)^{a+1}\right]$$

where $n = 2^L - 1$, $a = (n - 1)/2$, and the coefficient of x^w is the number of code words of weight w. Evaluating the first few terms,

$$g_0 = 1 \qquad g_1 = g_2 = 0$$

$$g_3 = \frac{n(n-1)}{3!}$$

$$g_4 = \frac{n(n-1)(n-3)}{4!}$$

$$g_5 = \frac{n(n-1)(n-3)(n-7)}{5!}$$

$$g_6 = \frac{n(n-1)(n-3)(n-5)(n-7)}{6!}.$$

For $w \geq 9$ the number of words of weight w is closely approximated by [10]

$$g_w = \frac{1}{(n+1)}\binom{n}{w}(1 + \epsilon)$$

where ϵ is small with respect to 1.

When $2^L - 1$ is a prime (e.g., $L = 3, 5, 7, 13, 17, 19, 31$), then the Hamming weight coefficients and Theo-

rem 1 can be combined to count exactly the number of sequences with the same signature and the same syndrome.

Theorem 2: Consider parallel signature and syndrome compression where:

i) The feedback corresponds to a primitive polynomial $p(x)$ of degree L.
ii) The sequence length $m = 2^L - 1$ is a prime.

Then:

1) The number of sequences of weight w and zero signature is g_w, the number of code words of weight w in the Hamming code generated by the primitive polynomial $p(x)$.
2) The number of weight w sequences with the same nonzero signature is $1/m(\binom{m}{w} - g_w)$.

Proof: If $m = 2^L - 1$ is a prime, then each rotational equivalence class has exactly m members. From Theorem 1, each member has a distinct signature value if any member has a nonzero signature. Thus each equivalence class adds one sequence to each nonzero signature or all the class members have zero signatures. Finally, a sequence has a zero signature if and only if it is a code word of a Hamming code. Q.E.D.

Example: Let $L = 5$ and $m = 2^L - 1 = 31$. Consider $w = 4$. There are 1085 weight 4 code words in the Hamming code, i.e., $g_4 = 1085$. From Theorem 2, there are 980 sequences with any particular nonzero signature. When $w = 6$ the corresponding numbers are $g_6 = 22\,568$ and for a nonzero signature 23 023. If all the weight 6 sequences of length 31 were to be partitioned as equally as possible into 32 classes, some classes would have 23 008 members and some 23 009. It can be seen that the signature partition is very close to the most uniform possible. As a measure of the uniformness we define a reduction factor R.

Definition: The reduction factor R of a signature unit with respect to a sequence length m and weight w is

$$R = \frac{\binom{m}{w}}{\text{maximum signature volume}}$$

where the signature volume is the number of weight w sequences of length m which have a particular signature.

Definition: The normalized reduction factor R^* is

$$R^* = \frac{R}{2^L}.$$

Note that $1 \le R \le 2^L$ and that $2^{-L} \le R^* \le 1$. The definitions of R and R^* are conservative and use the worst-case reduction. This guarantees that the actual reduction for any specific signature is at least R.

In the prime length case, using the expressions for the first few weight coefficients in Theorem 2, Table I results. It is conjectured that R approaches 2^L as w increases.

TABLE I
REDUCTION FACTOR R FOR PRIME LENGTH $m = 2^L - 1$

w, weight	1, 2	3, 4	5, 6
R, reduction factor	$2^L - 2$	$2^L - 3$	$2^L - \frac{15}{m^2 - 7m + 15}$

Using the approximation in [10]

$$R = \frac{2^L}{1 + \epsilon}$$

for $w \ge 9$, where $\epsilon \to 0$.

III. Enumerative Investigation

The cases examined in the prior section include only a small portion of those of interest. To obtain information for a wider range of parameter values, a program was written to simulate the action of a signature compressor. A constant weight sequence was compressed, the signature noted, and the process repeated. For a range of values of weight w and sequence length m, all possible sequences were simulated. The general conclusion was that the signature unit partitioned the constant weight sequences into nearly equal sized regions, as was found analytically for the prime length case. We next describe these experiments in detail.

The primitive feedback polynomial $p(x) = 1 + x + x^3 + x^4 + x^6$ was used in the cases summarized in Figs. 2 and 3. The reduction factor R would be at most $2^L = 64$. The vertical scale starts at 45 in Fig. 2 and the sequence length m ranges from 15 to 63 for three different weights $w = 3, 5, 7$. Note that R is not monotonic in m but is generally increasing. For $w = 3$, R approaches $2^L - 3$, which is the same value as that obtained analytically for prime lengths.

There were a few cases observed where R did not increase as w was made larger, for fixed m, and in these cases R remained the same for two consecutive values of w. The usual situation is illustrated in Fig. 3. R is seen to be monotonically increasing in w. Note that the vertical scale starts at 57 and that the two sample cases are for sequence lengths $m = 32$ and $m = 53$.

Next the characteristics for various polynomials were studied. All feedback polynomials of the form

$$1 + a_1 x + a_2 x^2 + \cdots a_{L-1} x^{L-1} + x^L$$

were considered. The packed representation was the number

$$PN = 2^{L-1} + \sum_{i=1}^{L-1} a_i 2^{i-1}.$$

For the polynomial used in Figs. 2 and 3, $PN = 45$. Fig. 4 is a plot of R for test length $m = 63$, signature compressor length $L = 6$, and $w = 3$, 4, and 5 as a function of PN, the packed polynomial coefficient. The lowest set of connected points corresponds to $w = 3$, the highest to $w = 5$. The encircled points are the six primitive poly-

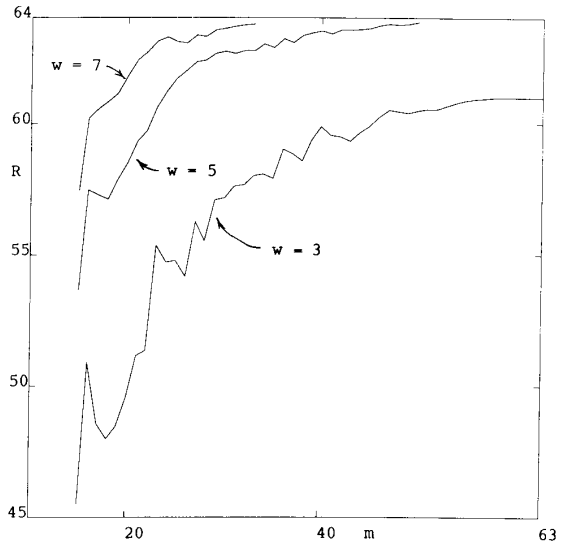

Fig. 2. Exact reduction factor R for $1 + x + x^3 + x^4 + x^6$.

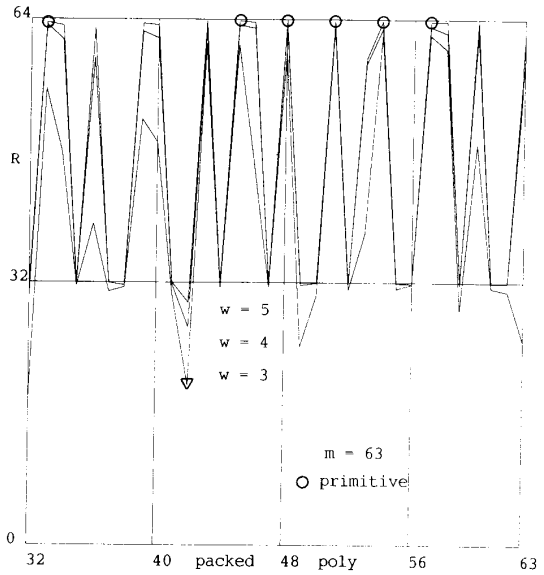

Fig. 4. R for feedback polynomials of length $L = 6$.

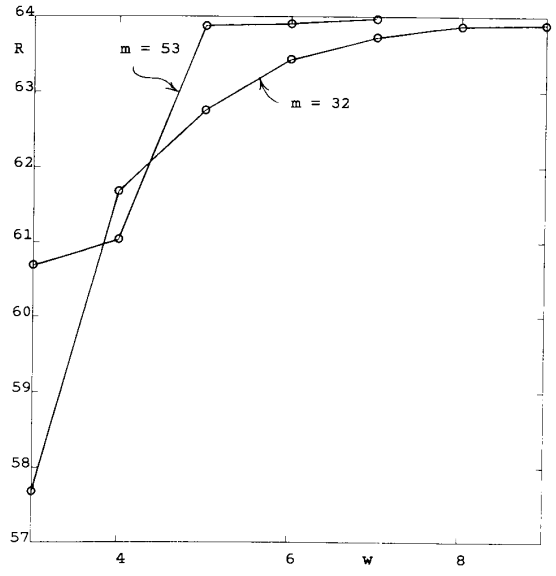

Fig. 3. Detail for $1 + x + x^3 + x^4 + x^6$ versus sequence weight w.

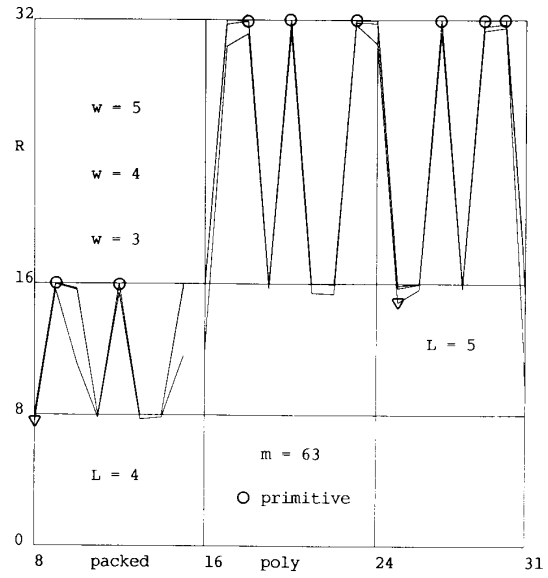

Fig. 5. R for polynomials of length 4 and 5.

nomials with packed values 33, 45, 48, 51, 54, 57. The poorest performance, as noted by the entriangled point, is for packed 42, which is $(1 + x)^6$.

It appears that any polynomial without 1 as a root offers comparable performance for w large and that a primitive polynomial performs well for w small.

The cases examined so far have considered test lengths m less than 2^L. It is known [5] that when m is 2^L or larger there will be some double error patterns that will be missed by signature compression. However there are few of these patterns and from a parallel compression view they do not significantly change the error behavior. Fig. 5 is a plot of R for test length $m = 63$ for all polynomials of degrees 4 and 5. Again, the lowest set of connected points correspond to $w = 3$, the highest to $w = 5$. The encircled points are primitive polynomials (best performance), while the entriangled points (poorest performance) are $(1 + x)^L$.

The overall characteristics found in this enumerative study are illustrated in Fig. 6. Note that the vertical scale is logarithmic. Each set of connected points corresponds to a primitive polynomial of degree L. Note, for sequence weight $w = 5$ or more, that the reduction factor R is essentially 2^L.

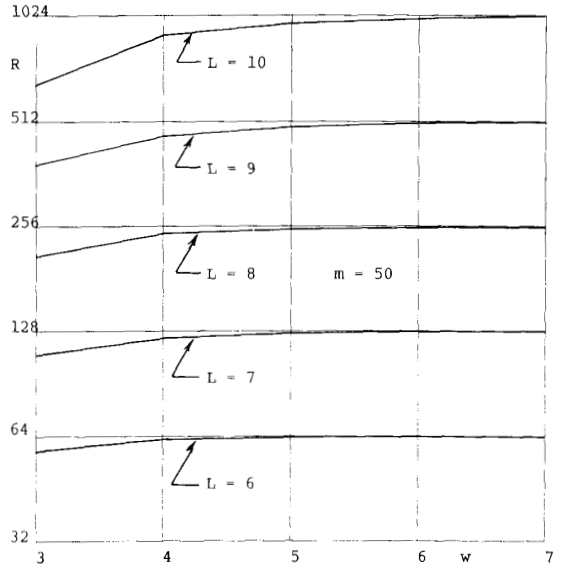

Fig. 6. R for sequence length 50 for primitive polynomials of degree L.

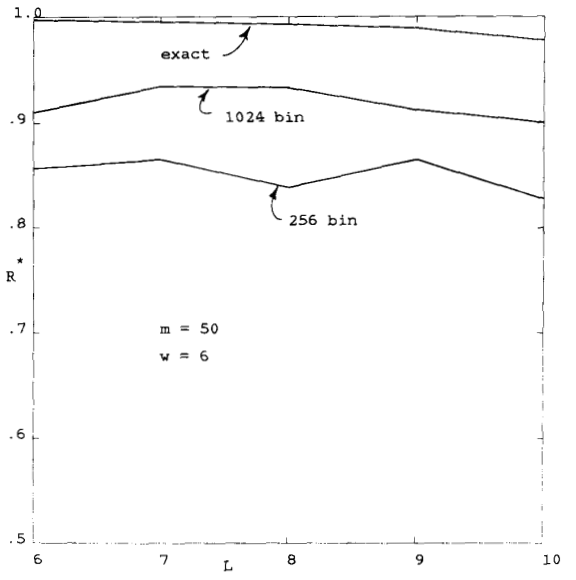

Fig. 7. Comparison between exact and Monte Carlo estimates for R.

IV. Monte Carlo Simulation

Above about 10^9 vectors, complete enumeration consumes a substantial amount of simulation time. A Monte Carlo approach was programmed to estimate the reduction factor for these larger cases. Fig. 7 is a validation check on this approach. The upper connected points are the exact values while the lower connected points are the Monte Carlo estimates. The reduction factor has been normalized by dividing by 2^L, the maximum possible value; thus $2^{-L} \leq R^* \leq 1$. Note that the estimates are uniformly less than the exact values. It is to be expected that the Monte Carlo estimates would be lower than the exact values since the largest sample bin count is used to define R^*.

Of course, it is not guaranteed that the Monte Carlo estimate will always be less than the exact value of R. It is conceivable that the random vectors might divide evenly into the 2^L bins, when R is less than 2^L. However, this event is extremely unlikely. The number of sequences simulated for each point was the bin number times 2^L. Note that the vertical axis starts at 0.5.

For the average bin size of 1024 (1024 bin) it is seen that the Monte Carlo estimate is at least 90 percent of the exact value. This average bin size, 1024, was used in subsequent estimates.

Figs. 8 and 9 summarize the characteristics found in this study. For both figures the cases examined are given in Table II, where the weight w is about $m/10$.

Fig. 8 reports cases where the sequence length m is mostly longer than 2^L, while Fig. 9 gives cases of both types. Again it is seen that the reduction factor R is close to 2^L. Additional cases were also examined and were found to give the same behavior.

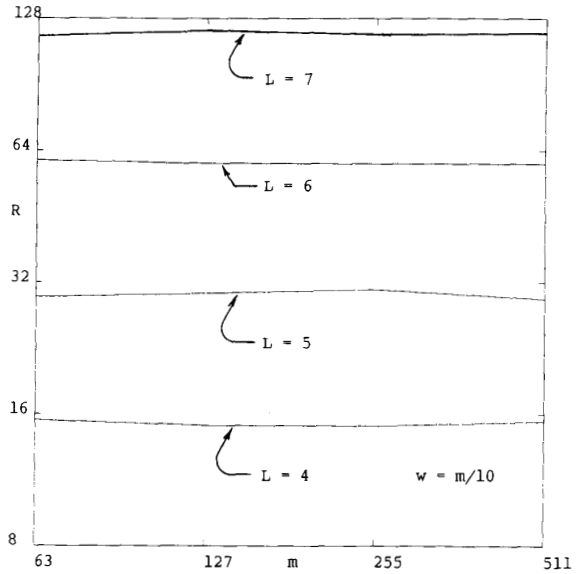

Fig. 8. Monte Carlo estimates of R for $L = 4$, 5, 6, and 7.

V. Conclusions

Simultaneous test data reduction using a length L signature compressor and a syndrome compressor will nearly always reduce the syndrome error ambiguity by a factor of almost 2^L, even if the test length is larger than 2^L. For a few special cases an exact expression of the reduction was obtained. There seems to be little difference in this reduction for various signature polynomials; however, the signature polynomial should not have an even number of nonzero terms. An even number of terms will approximately double the number of error aliases. Based on our

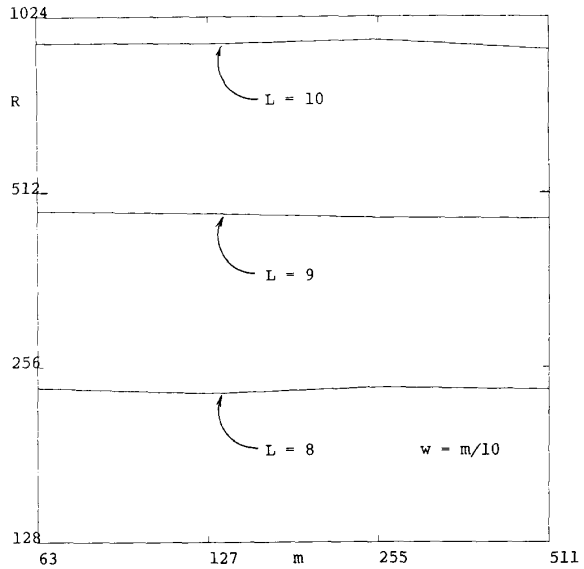

Fig. 9. Monte Carlo estimates of R for $L = 8$, 9, and 10.

TABLE II
CASES FOR FIGURES 8 AND 9

m, sequence length	63	127	256	511
w, weight	7	13	26	51

simulation studies, a primitive feedback polynomial (with about half nonzero coefficients) seems optimal. The differences between such a primitive polynomial and any polynomial without 1 as a root are only noticeable for test data with a very small number of ones (fewer than 6).

The reduction characteristics noted above also hold for transition counting, Walsh spectral coefficients, and any linear operation followed by counting when combined with signature compression. Hence any of these counter-based methods can have their error ambiguity reduced using simultaneous signature compression. From an error count view, the particular method having the smallest count total would result in the smallest overall error ambiguity.

REFERENCES

[1] T. W. Williams and K. P. Parker, "Design for testability—A survey," *Proc. IEEE*, pp. 98–112, Jan. 1982.
[2] T. W. Williams, "VLSI testing," *COMPUTER*, pp. 126–136, Oct. 1984.
[3] E. J. McCluskey, "Built-in self-test techniques," *IEEE Design and Test of Computers*, vol. 2, no. 2, pp. 21–28, Apr. 1985.
[4] N. Benowitz *et al.*, "An advanced fault isolation system for digital logic," *IEEE Trans. Comput.*, vol. C-24, pp. 489–497, May 1975.
[5] J. E. Smith, "Measures of the effectiveness of fault signature analysis," *IEEE Trans. Comput.*, vol. C-29, pp. 510–514, June 1980.
[6] J. Savir, "Syndrome testable design of combinational circuits," *IEEE Trans. Comput.*, vol. C-27, pp. 442–550, June 1980.
[7] A. K. Susskind, "Testing by verifying Walsh coefficients," *IEEE Trans. Comput.*, vol. C-32, pp. 198–201, Feb. 1983.
[8] D. M. Miller and J. C. Muzio, "Spectral fault signatures for single stuck-at faults in combinational networks," *IEEE Trans. Comput.*, vol. C-33, pp. 765–769, Aug. 1984.
[9] A. M. Michelson and A. H. Levesque, *Error-Control Techniques for Digital Communication*. New York: Wiley, 1985.
[10] T. Kasami, T. Fujinara, and S. Lin, "An approximation to the weight distribution of binary linear codes," *IEEE Trans. Inform. Theory*, vol. IT-31, Nov. 1985.

*

John P. Robinson (S'58–M'65–SM'77) received the B.S. degree in electrical engineering from Iowa State University, Ames, in 1960, and the M.S. and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, in 1962 and 1966, respectively.

From 1960 to 1962 he was at the RCA Laboratories, Princeton, NJ, and during the summers of 1963 and 1964, he was at the IBM Thomas J. Watson Research Center, Yorktown Heights, NY. Since 1965 he has been a member of the Department of Electrical and Computer Engineering, University of Iowa, Iowa City.

*

Nirmal R. Saxena received the B.E. degree in 1982 from Osmania University, Hyderabad, India, and the M.S.E.E. degree from the University of Iowa, Iowa City, in 1984.

He is currently a member of the technical staff in the Systems Architecture Laboratory of Hewlett Packard, Cupertino, CA. He is also working toward the Ph.D. degree at the Center for Reliable Computing, Standard University, Stanford, CA.