EV4-10/98-0xx

TO: EV/Chief, Avionic Systems Division

FROM: EV4/Chief, Avionics Test and Analysis Branch

SUBJECT: Theoretical Accuracy for ESTL Bit Error Rate Tests

Enclosed is the subject document for your information and retention. Bit Error Rate (BER) measurements are made at ESTL by evaluating a sample. A textbook approximation and an original exact technique is presented for determining a confidence interval containing the true BER given a trial outcome. This exact technique is important because the textbook approach diverges in the region where measurement error is the greatest: when a trial results in few or no errors.

For additional information on this analysis, please contact Chatwin Lansdowne at mail code EV41 or (281) 483-1265.

Sidney Novosad

Enclosure

cc:
EV/C. K. Land
EV/L. K. Bromley
EV41/C. A. Lansdowne
EV41/N. J. Robinson
EV41/B. G. Smith
EV41/M. S. Marston
EV41/S. S. Marston
EV41/C. G. Krpec
EV41/R. S. Wade
EV41/C. C. Sham
EV41/L. E. Hood

EV41/ATSC/J. A. Ross
EV41/ATSC/S. W. Phillips

EV41/LMSMSS/Library/C. Foster
ESTL RDC/K. Jolley
LMSMSS/C22/D. N. Cravey

LMSMSS/C22/C. D. Peña
LMSMSS/C22/J. E. Krafka
LMSMSS/C22/P. Allen
LMSMSS/C22/B. S. Collier
LMSMSS/C22/C. M. Delgado
LMSMSS/C22/W. D. Edwards
LMSMSS/C22/P. Hangsleben
LMSMSS/C22/T. B. Jenkins
LMSMSS/C22/G. L. McKinney
LMSMSS/C22/N. L. Mindiola
LMSMSS/C22/S. G. Morton
LMSMSS/C22/T. D. Pham
LMSMSS/C22/J. W. Raetzke
LMSMSS/C22/F. T. Shetz

NASA
National Aeronautics and
Space Administration

EV4-98-610

Lyndon B. Johnson Space Center
Houston, TX 77058

ELECTRONIC SYSTEMS TEST LABORATORY

THEORETICAL ACCURACY FOR ESTL BIT ERROR RATE TESTS

EV/AVIONIC SYSTEMS DIVISION
EV4/SYSTEMS ANALYSIS AND TEST BRANCH
SEPTEMBER 1998

THEORETICAL ACCURACY FOR
ESTL BIT ERROR RATE TESTS

Prepared by:

Chatwin Lansdowne

Reviewed by:

Ned J. Robinson, ESTL Facility Manager

Billy G. Smith, Test Manager

Approved by:

Sidney W. Novosad, Chief
Avionics Test and Analysis Branch

ACRONYMS AND ABBREVIATIONS

BER	Bit Error Rate
BPSK	Bi-phase Shift Key
ESTL	Electronic Systems Test Laboratory
NASA	National Aeronautics and Space Administration
PDF	Probability Density Function
PSD	Power Spectral Density
QPSK	Quadrature Phase Shift Key
RF	Radio Frequency
TRP	Total Received Power

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THEORETICAL ACCURACY FOR
ESTL BIT ERROR RATE TESTS

1. Introduction

"Bit error rate" [BER] for the purposes of this paper is the fraction of binary bits which are inverted by passage through a communication system. BER can be measured for a block of sample bits by comparing a received block with the transmitted block and counting the erroneous bits. Bit Error Rate [BER] tests are the most common type of test used by the ESTL for evaluating system-level performance.

The resolution of the test is obvious: the measurement cannot be resolved more finely than $1/N$, the number of bits tested. The tolerance is not. This paper examines the measurement accuracy of the bit error rate test. It is intended that this information will be useful in analyzing data taken in the ESTL.

This paper is divided into four sections and follows a logically ordered presentation, with results developed before they are evaluated. However, first-time readers will derive the greatest benefit from this paper by skipping the lengthy section devoted to analysis, and treating it as reference material. The analysis performed in this paper is based on a Probability Density Function [PDF] which is developed with greater detail in a past paper, *Theoretical Accuracy for ESTL Probability of Acquisition Tests,* EV4-98-609.

2. Assumptions

The entire test as well as this derivation hinge on one central assumption: it is valid to represent the transmitter/path/receiver system with a constant, P_e, the probability of a bit error, which is assumed to be uniformly constant for every bit. If errors occur in bursts or blocks, then P_e for a bit depends on whether the previous bit was an error and this analysis is invalid (errors must be independent). If symbols are not equally error-prone, then P_e is a function of the data stream. If the transmitter power varies, if the receiver noise temperature varies (*e.g.*, during warm-up), if the path loss varies, then P_e is not a constant but is a function of these variables.

3. Analysis

The exact confidence interval is somewhat complicated and is impractical to calculate for large sample sizes. A very good approximation will also be developed.

3.1 Exact Tolerance

An exact approach was developed for Bernoulli trials (repeated independent instances of an experiment that can have one of two possible outcomes) in *Theoretical Accuracy for ESTL Probability of Acquisition Tests*, EV4-94-618. The result was,

$$\frac{1-CL}{2}+CL=(N+1)\cdot C(N,n)\cdot\int_0^{P_{e,up}} P_e^n(1-P_e)^{(N-n)}\,dP_e$$

$$CL+\frac{1-CL}{2}=(N+1)\cdot C(N,n)\cdot\int_{P_{e,lo}}^{1} P_e^n(1-P_e)^{(N-n)}\,dP_e$$

where

P_e is the actual probability of bit error for the transmitter/attenuator/receiver configuration. This number is the unknown we seek to bound.

N is the number of samples (bits) per trial.

n is the number of bit errors in a trial (n will be measured).

CL is the confidence level desired for the interval (0.95 will be used here).

$P_{e,up}$ is the upper boundary of the interval that contains P_e with confidence CL.

$P_{e,lo}$ is the lower boundary of the interval that contains P_e with confidence CL.

and $C(N,n)$ is the combinations function, which gives the number of combinations of n objects that can be chosen from a population N.

$$C(N,n)=\frac{N!}{n!(N-n)!}$$

The integral must be solved by using the integration by parts technique recursively, giving a closed form of

$$\int x^n(1-x)^{N-n}\,dx=I=\frac{x^{n+1}(1-x)^{N-n}}{N+1}+\sum_{i=0}^{N-n-1}\left(\prod_{j=0}^{i}\frac{N-n-j}{N+1-j}\right)\frac{x^{n+1}(1-x)^{N-n-1-i}}{N-i}$$

This same integral reduces to a single-loop algorithm with good numerical properties in that it recursively adds two numbers of same sign and similar magnitude.

```
I=0.0E+0
for i=0 to N-n step 1
   I=(I*i + (1-x)^i)/(n+1+i)
next i
I=I*x^(n+1)
```

Practically, the equations can be solved easily by a computer using either symbolic, numeric, or direct methods. It is not obvious from these equations that the tolerance converges independently of N for $N>>n$. The general solution for any n is complicated, so convergence will be demonstrated by solving for the relatively simple case of n=1 and showing that for large N the tolerance is a constant. The rule of integration by parts, $\int udv=uv-\int vdu$, is used.

$$\frac{1-CL}{2}+CL=(N+1)\cdot C(N,1)\cdot\int_0^{P_{e,up}} P_e^1(1-P_e)^{(N-1)}\,dP_e$$

$$\frac{1+CL}{2}=(N+1)\cdot N\cdot\frac{-1}{N}\cdot\int_0^{P_{e,up}} P_e\left(-N(1-P_e)^{(N-1)}\right)dP_e$$

$$=-(N+1)\left(P_e(1-P_e)^N\Big|_0^{P_{e,up}}-\int_0^{P_{e,up}}(1-P_e)^N\,dP_e\right)$$

$$
\begin{aligned}
&= -(N+1)\left(P_{e,up}(1-P_{e,up})^N + \left.\frac{(1-P_e)^{N+1}}{N+1}\right|_0^{P_{e,up}} \right) \\
&= -(N+1)\cdot\left(P_{e,up}(1-P_{e,up})^N + \frac{(1-P_{e,up})^{(N+1)}}{(N+1)} - \frac{1}{(N+1)} \right) \\
&= \left(-(N+1)P_{e,up} - (1-P_{e,up})\right)(1-P_{e,up})^N + 1 \\
&= 1-(NP_{e,up}+1)(1-P_{e,up})^N
\end{aligned}
$$

Now substitute from the definition for positive tolerance, $t_+=(P_{e,up}-n/N)/(n/N)=NP_{e,up}-1$ for $n=1$, and take the limit for large N:

$$
\begin{aligned}
\frac{1+CL}{2} &= 1-(t_+ + 2)\left(1-\frac{t_+ + 1}{N}\right)^N \\
&= 1-(t_+ + 2)e^{-(t_+ + 1)}, \quad \text{for } \lim_{N\to\infty}
\end{aligned}
$$

This final equation can be solved (at least numerically) for its single unknown, t_+: if $CL=0.95$, then $t_+=4.572$. This illustrates the point that for $N>>n$, t is a function of n (=1 for this case) and of course the desired level of confidence.

Results for this section were found numerically and are included in the next section.

3.2 Exact Tolerance, Error-Free Case

For the relatively trivial case of no errors ($n=0$), the closed-form solution can be simplified. Here the equation has been slightly modified so that $P_{e,up}$ is the BER we are CL certain that the true BER is below.

$$
\begin{aligned}
CL &= (N+1)\cdot C(N,0)\cdot\int_0^{P_{e,up}} P_e^0(1-P_e)^{(N-0)}\,dP_e \\
&= (N+1)\cdot 1\cdot\left(\frac{-1}{(N+1)}\right)\cdot\int_0^{P_{e,up}} -(N+1)(1-P_e)^N\,dP_e \\
&= -(1-P_e)^{N+1}\Big|_0^{P_{e,up}} \\
&= \left(-(1-P_{e,up})^{N+1} - (-1)\right) \\
P_{e,up} &= 1-(1-CL)^{\frac{1}{N+1}}
\end{aligned}
$$

Results are given in the next section.

3.3 Gaussian Approximation of Tolerance

For large values of N where $n \approx N/2$, the shape of the binomial distribution approaches the Gaussian "bell-curve." For the binomial distribution the expected value and variance of n are

$$\mu = NP_e, \qquad \sigma^2 = NP_e(1 - P_e)$$

So the variance is a parabolic function of the system BER, it is zero for P_e=0 or P_e=1 and it peaks at P_e=0.5. Combining these facts, we can find a worst case confidence interval in the conventional fashion. First, define a half-width parameter e such that

$$P_{e,lo} = \frac{n}{N} - e \qquad P_{e,up} = \frac{n}{N} + e$$

then,

$$P(P_{e,lo} \le P_e \le P_{e,up}) \ge CL$$

$$P(\frac{n}{N} - e \le P_e \le \frac{n}{N} + e) \ge CL$$

The Standard Normal distribution ($F(z)=F((x-\mu)/\sigma)$ (tables available in most probability textbooks) can be applied to find a value for $P(-a<z<a)= P(-a<(x-\mu)/\sigma<a)$. Then,

$$P\left(\frac{-e\sqrt{N}}{\sqrt{P_e(1-P_e)}} \le \frac{n - NP_e}{\sqrt{NP_e(1-P_e)}} \le \frac{e\sqrt{N}}{\sqrt{P_e(1-P_e)}}\right) \ge CL$$

$$P\left(\frac{-e\sqrt{N}}{\sqrt{P_e(1-P_e)}} \le \frac{n - \mu}{\sigma} \le \frac{e\sqrt{N}}{\sqrt{P_e(1-P_e)}}\right) \ge CL$$

This is still a general application of the Gaussian distribution. It assumes that Pe is known. We will now substitute $P_e \approx n/N$ to approximate variance and average. Later it will be shown that a better substitution would be $P_e \approx (n+1)/N$; however, $n>>1$ for the region where the approximation of tolerance developed in this subsection is useful.

$$P\left(\frac{-e}{\sqrt{\frac{n}{N^2} - \frac{n^2}{N^3}}} \le \frac{n - \mu}{\sigma} \le \frac{e}{\sqrt{\frac{n}{N^2} - \frac{n^2}{N^3}}}\right) \ge CL$$

$$P\left(\frac{-e/(n/N)}{\sqrt{\frac{1}{n} - \frac{1}{N}}} \le \frac{n - \mu}{\sigma} \le \frac{e/(n/N)}{\sqrt{\frac{1}{n} - \frac{1}{N}}}\right) \ge CL$$

This last equation can be solved using the Standard Normal distribution. For a 95% confidence level (CL=0.95),

$$\frac{e/(n/N)}{\sqrt{\frac{1}{n}-\frac{1}{N}}} = 1.959961$$

$$\frac{e}{(n/N)} = 1.959961\ \sqrt{\frac{1}{n}-\frac{1}{N}}$$

This is an expression for tolerance as a fraction of the measured error rate, $\pm 100e/(n/N)\%$. The result can be simplified further for either end of the curve:

$$\frac{e}{(n/N)} = \frac{1.959961}{\sqrt{2n}}, \quad n = \frac{N}{2}$$

$$\frac{e}{(n/N)} = \frac{1.959961}{\sqrt{n}}, \quad N >> n$$

The general shape of the curve is dominated by n, with a $1/\sqrt{2}$ improvement in tolerance as the BER approaches 0.5. Although this approximation is good to a couple of digits for more than about 400 errors, it is not good to a single digit for fewer than 43 errors because the probability density function in this region is asymmetrically "bell" shaped. Results are illustrated in the next section.

3.4 Ideal BER function of Total Received Power

The probability of bit error for an "ideal" receiver using Binary Phase Shift Keying [BPSK] or Quadrature Phase Shift Keying [QPSK] is derived in most digital communications textbooks (see references). Many other modulation schemes have a similarly shaped ideal case. The general solution is,

$$P_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right) = Q(\sqrt{kP_{TRP}})$$

where E_b is the energy in a bit, $N_0/2$ is the Power Spectral Density [PSD] of the noise at the receiver input, and $Q(-x)$ is the area under the Standard Normal Probability Density Function between $-\infty$ and x. k is a constant (composed of noise temperature, noise figure, bit rate, and Boltzman's constant) and P_{TRP} is the Total Received Power. So,

$$\frac{P_{TRP1}}{P_{TRP2}} = \frac{Q^{-1}(P_{e1})^2}{Q^{-1}(P_{e2})^2}$$

This result is used in the next section to determine how much difference in power exists in the tolerance interval of the Bit Error Rate.

3.5 Expected Value of True BER

The expected (or "average") value of the true BER, based on a measurement, is ordinarily the measured BER. This is absolutely true when the probability density function for the true BER, given the measurement, is symmetrical about the measured value. The PDF is found in this paper to be approximately symmetrical (to a couple of digits) for measurements of 400 or more errors, and also for measured BERs near 0.5.

For small numbers of errors, the expected true BER is actually slightly higher than the measurement. That is, if more measurements were made, several measurements will more often than not average out to a higher BER than the initial measurement indicated. This effect should be accounted for in some situations.

The "average" outcome for any probability model may be determined by summing the possible outcomes, each weighted by its probability or frequency of occurrence (and divide by the sum of the weights, which is one-- one of the possible outcomes has to happen). For a continuous model, integrate over the range of possible outcomes, the product of the outcome and the PDF evaluated at that outcome:

$$E(X) \equiv \int_{-\infty}^{\infty} x \cdot f(x)dx$$

where

X is a random variable.

$E(X)$ is the Expected Value of the random variable X.

$f(x)$ is the Probability Density Function of X evaluated at x.

For the binomial distribution, the PDF is given by

$$f(P_e) = (N+1) \cdot C(N,n) \cdot P_e^n \cdot (1-P_e)^{(N-n)}$$

for P_e between 0 and 1. This was developed in *Theoretical Accuracy for ESTL Probability of Acquisition Tests*, EV4-94-618, and represents the ordinary binomial distribution scaled by the number of possible outcomes so that it obeys a fundamental rule for PDFs: the area beneath it is one. Applying these definitions to the problem at hand,

$$E(P_e) = (N+1) \cdot C(N,n) \cdot \int_0^1 x \cdot x^n \cdot (1-x)^{(N-n)}\, dx$$

$$= (N+1) \cdot C(N,n) \cdot \int_0^1 x^{(n+1)} (1-x)^{(N-n)}\, dx$$

where

P_e is the actual probability of bit error for the transmitter/attenuator/receiver configuration. This number is the unknown for which we seek an expectation.

N is the number of samples (bits) per trial.

n is the number of bit errors in a trial (n will be measured).

The integral must be solved by using the integration by parts technique recursively, giving a closed form of

$$\int x^{(n+1)} (1-x)^{N-n}\, dx = I = \frac{x^{n+2}(1-x)^{N-n}}{N+2} + \sum_{i=0}^{N-n-1} \left(\prod_{j=0}^{i} \frac{N-n-j}{N+1-j} \right) \frac{x^{n+2}(1-x)^{N-n-1-i}}{N+1-i}$$

This same integral reduces to a single-loop algorithm with good numerical properties in that it recursively adds two numbers of same sign and similar magnitude.

```
I=0.0E+0
for i=0 to N-n step 1
   I=(I*i + (1-x)^i)/(n+2+i)
next i
I=I*x^(n+2)
```

Further, a pseudo-code algorithm which gives I for the interval of 0 to 1 is,

```
I=1/(n+2)
```

```
for i=1 to N-n step 1
   I=I*i/(n+2+i)
next i
I=I*(N+1)*C(N,n)
```

Numerical evaluation of this algorithm indicates that for $N>>n$, the expression

$$E(P_e) \approx (n+1)/N$$

is at the least a very, very good approximation of expected value for the BER; however, the equations have proved too complicated to easily demonstrate that they truly reduce to this equivalence in the limit for large N.

4. Results

Results of applying the equations derived in the preceding section are illustrated in Figure 1 and tabulated in Table 1.

The approximation developed in the previous section can be expanded for other common choices of confidence level. The numerator is derived from the standard normal distribution, provided as a table in any probability textbook:

CL	$\pm e/(n/N) \approx$
0.99	$2.575835 \sqrt{\frac{1}{n}-\frac{1}{N}}$
0.95	$1.959961 \sqrt{\frac{1}{n}-\frac{1}{N}}$
0.9	$1.644853 \sqrt{\frac{1}{n}-\frac{1}{N}}$

Having calculated the 95% confidence case, then 31% more samples are needed for 99% confidence, and 16% fewer samples are needed for 90% confidence.

The error-free case may be summarized for most laboratory needs as follows: for $x>1$, if a $3.00 \cdot 10^x$ bit trial is error-free, then the BER is 95% likely to be below $1.00 \cdot 10^{-x}$. And if a $1.00 \cdot 10^x$ bit trial is error-free then it is 95% certain that the BER is below $3.00 \cdot 10^{-x}$.

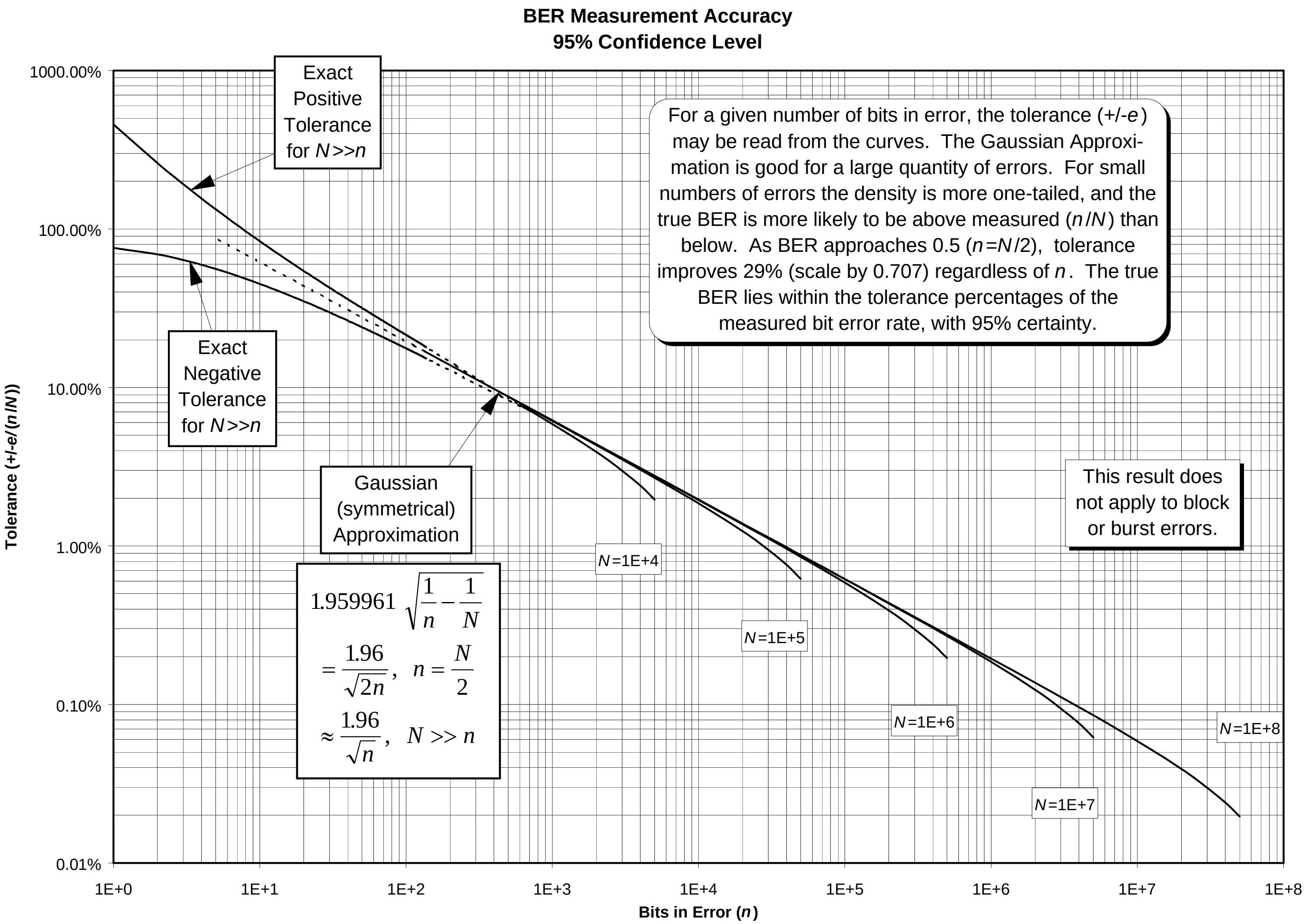

Figure 1. Tolerance of BER test for 95% Confidence Level

Table 1. Calculated Results for 95% Confidence Level

Errors n	True Positive Tolerance	True Negative Tolerance	Gaussian (Symmetrical) Approximation
1	+457.2%	-76.0%	±196.0%
2	+261.1%	-69.2%	±138.6%
3	+192.1%	-63.8%	±113.2%
4	+155.9%	-59.4%	±98.0%
5	+133.3%	-55.9%	±87.7%
6	+117.6%	-53.1%	±80.0%
7	+105.9%	-50.6%	±74.1%
8	+97.0%	-48.5%	±69.3%
9	+89.8%	-46.7%	±65.3%
10	+83.9%	-45.1%	±62.0%
11	+78.9%	-43.6%	±59.1%
12	+74.6%	-42.3%	±56.6%
13	+71.0%	-41.1%	±54.4%
14	+67.7%	-40.0%	±52.4%
15	+64.9%	-39.0%	±50.6%
16	+62.3%	-38.1%	±49.0%
17	+60.0%	-37.2%	±47.5%
18	+58.0%	-36.5%	±46.2%
19	+56.1%	-35.7%	±45.0%
20	+54.4%	-35.0%	±43.8%
21	+52.8%	-34.3%	±42.8%
22	+51.3%	-33.7%	±41.8%
23	+50.0%	-33.1%	±40.9%
24	+48.7%	-32.6%	±40.0%
25	+47.6%	-32.0%	±39.2%
26	+46.5%	-31.6%	±38.4%
27	+45.4%	-31.1%	±37.7%
28	+44.5%	-30.6%	±37.0%
29	+43.6%	-30.2%	±36.4%
30	+42.7%	-29.8%	±35.8%
31	+41.9%	-29.4%	±35.2%
32	+41.1%	-29.0%	±34.6%
33	+40.4%	-28.6%	±34.1%
34	+39.7%	-28.3%	±33.6%
35	+39.0%	-27.9%	±33.1%
36	+38.4%	-27.6%	±32.7%
37	+37.8%	-27.3%	±32.2%
38	+37.2%	-27.0%	±31.8%
39	+36.6%	-26.7%	±31.4%

Table 1. Calculated Results for 95% Confidence Level (continued)

Errors n	True Positive Tolerance	True Negative Tolerance	Gaussian (Symmetrical) Approximation
40	+36.1%	-26.4%	±31.0%
41	+35.6%	-26.1%	±30.6%
42	+35.1%	-25.9%	±30.2%
43	+34.6%	-25.6%	±29.9%
45	+33.7%	-25.1%	±29.2%
50	+31.8%	-24.0%	±27.7%
60	+28.6%	-22.2%	±25.3%
70	+26.3%	-20.8%	±23.4%
80	+24.4%	-19.6%	±21.9%
90	+22.8%	-18.6%	±20.7%
100	+21.5%	-17.7%	±19.6%
110	+20.4%	-16.9%	±18.7%

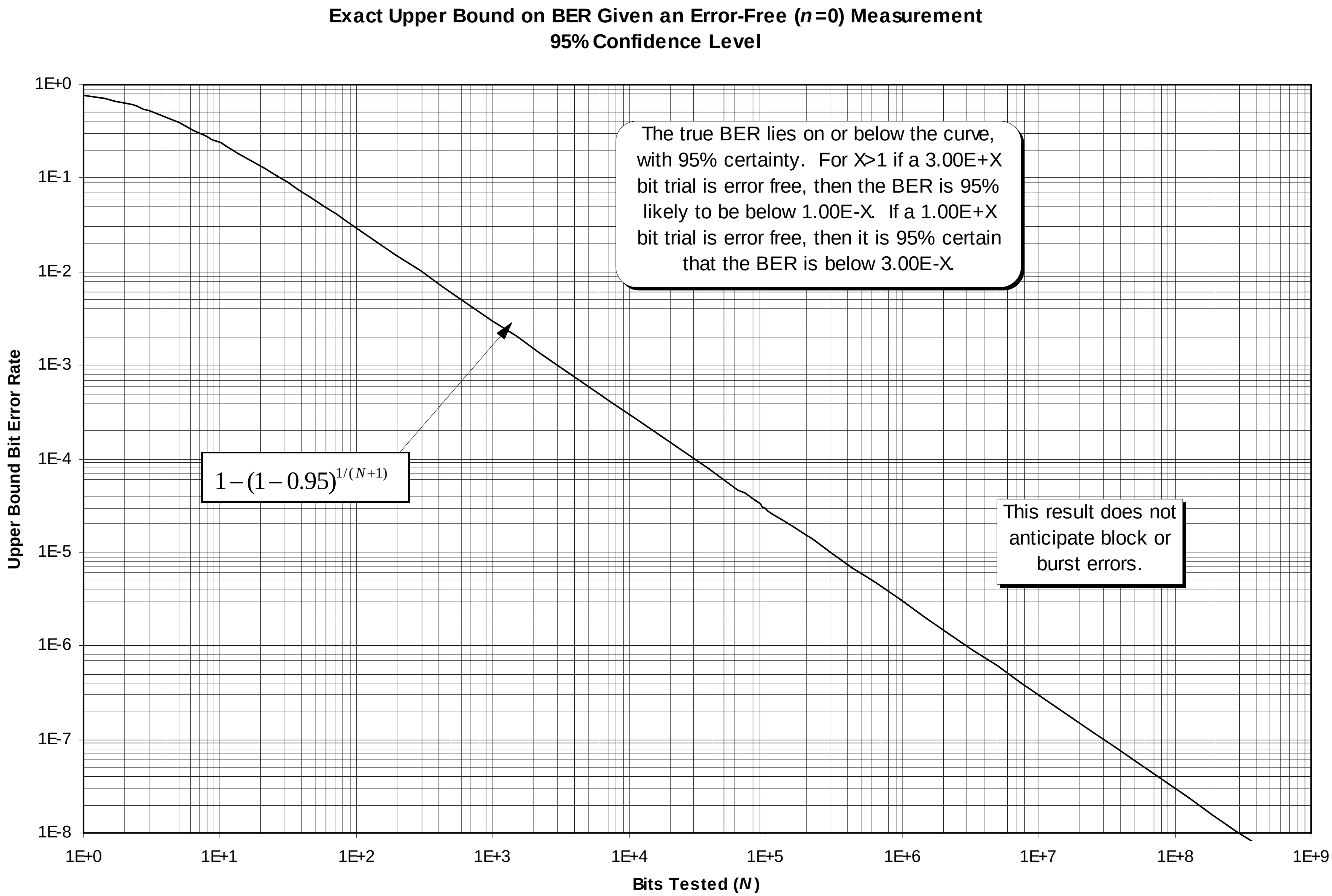

Figure 2. Upper Boundary on BER for an Error-Free Trial

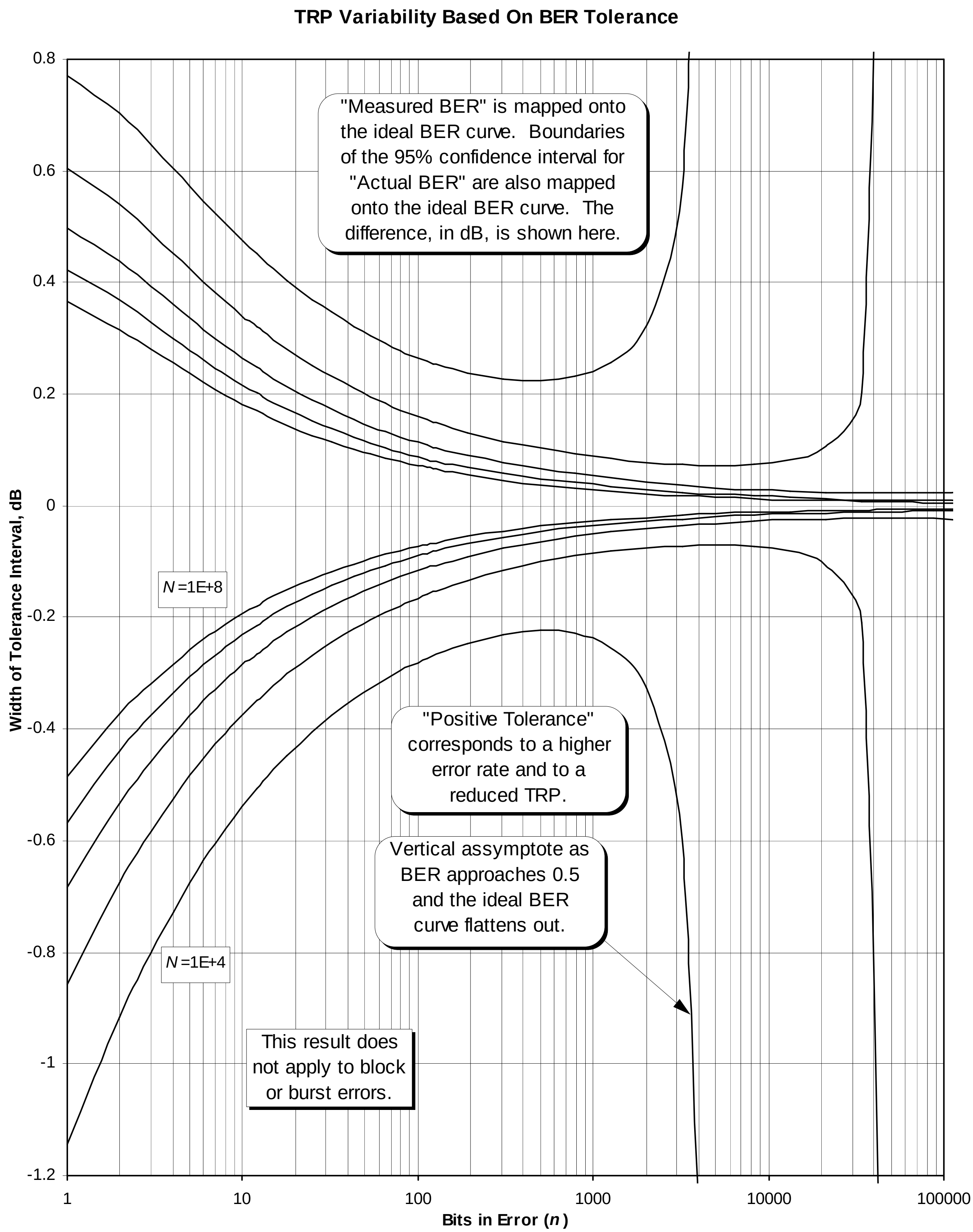

Figure 3. Power in the Tolerance Interval for the Ideal BER Curve

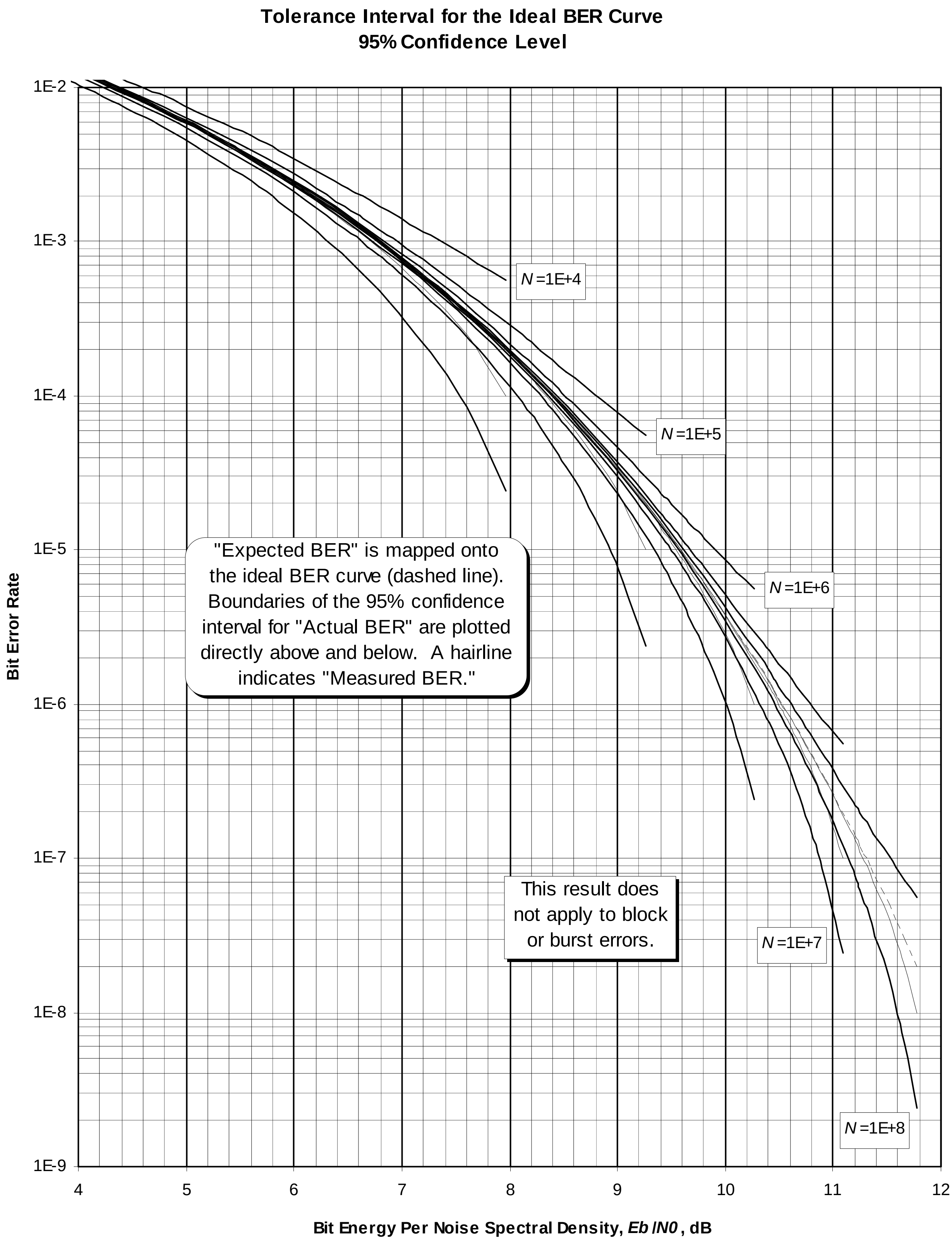

Figure 4. Tolerance Interval Boundaries on the Ideal BER Curve

5. Summary and Conclusion

BER measurement by sampling is a reasonable estimator. Tolerance of the measurement for small numbers of errors is an asymmetric function strongly influenced by the number of errors and weakly influenced by the number of samples. Asymmetry occurs because the distribution (PDF) for measurable BERs for small numbers of errors is single-tailed (no fewer than zero are possible) but for larger numbers of errors takes on a bell shape. At higher error rates there still may be significant variation in the measurements but the variation tends to be symmetrical.

If more accuracy is required, it can be obtained by decreasing confidence or by increasing the number of samples taken per trial.

A phenomenon called "kick-up" has been frequently observed in the ESTL for measurements of small numbers of errors. "Kick-up" describes the measurement at the lowest error rate, if it causes the rate of fall from left to right of the logarithmic plot of BER *vs*. RF power level to appear to diminish (the slope of this curve should monotonically become increasingly negative with increasing power). Kick-up causes concern because if it is accurate it suggests that the terminal BER at strong signal is something higher than error-free. Kick-up is shown to be irregular by the results presented here. Because the tolerance interval is "very" asymmetrical for fewer than about ten errors most measurements will be slightly below average, balanced by fewer but farther measurements on the high side. In fact, most measurements will kick *down*. Because the 95% confidence interval around a measurement will fail to contain the true BER for one in twenty measurements, a single kicked-up measurement does not justify drawing the conclusion that kick-up exists; rather, it suggests that the test should be repeated to gather more data and test the hypothesis that kick-up occurred.

6. References

1. *Theoretical Accuracy for ESTL Probability of Acquisition Tests*, EV4-94-618, October 1994.
2. *Elementary Statistics, Fifth Ed.* Mario F. Triola. Addison-Wesley Publishing Company, Inc., New York, ©1992. ISBN 0-8053-7631-3 (Sect. 5-5)
3. *Digital and Analog Communication Systems, Third Edition.* Leon W. Couch II. Macmillan, New York, ©1990. ISBN 0-02-325391-6