

DesignCon 2013

Understanding Apparent Increasing Random Jitter with Increasing PRBS Test Pattern Lengths

Dr. Martin Miller Ph.D., Teledyne LeCroy

[martin.miller@teledynelecroy.com]

Abstract

In the presence of an imperfect transmission "channel," there are physical and mathematical reasons that random jitter sometimes increases with increasing length PRBS test patterns. Essentially, the nature of the any inter-symbol interference (ISI) is such that the behavior of the total jitter (Tj) diverges significantly from the "dual-Dirac model". We rely upon that model for the basic deterministic and random jitter (Dj, Rj) breakdown. For multiple reasons, often the increase in observed Rj is not only apparent but can be verified by BERT measurements, and mathematically justified. The effect is larger than many engineers suspect. For that matter, one can ask whether long PRBS patterns serve as good test case at all.

Author's Biography

Employed at Teledyne-LeCroy (previously LeCroy Corporation), since August 1977 (35 years), I began as analog hardware engineer, working on charge sensitive analog-to-digital converters (ADCs) and time-to-digital converters (TDCs) for physics research. I was Engineering Manager from 1981 to 1985 for a division of LeCroy. I designed instruments using flash ADC's for fast frame capture (50MS/s) video digitizer. From 1985-86 devoted myself to writing the first waveform processing package for the 9400 digital oscilloscope, and specifying the next two packages for parametrics and Fourier Analysis. Led the development team for the 7200 modular digital oscilloscope as Director of R&D. I then spent the next 17 years at LeCroy in Geneva, Switzerland as Senior Scientist and then Chief Scientist. The most recent 15 years have been focused on Jitter and Noise Analysis for serial data and clocks.

I have participated in IEEE committees and JEDEC, contributing to test methods for digitizers and memories. I have actively participated in DesignCon since 2007 as both an "expert" panel member (tp_M1), as a paper presenter and tutorial presenter. I am also on the program committee for DesignCon.

The Trend and the Concern

There is a trend in the electronics industry to use longer and longer PRBS test sequences for Jitter Analysis. I have never been given a satisfying rationale for this trend apart from the fact that PRBS patterns are easy to generate and detect in test equipment, and longer PRBS sequences include the long run-lengths of 1's and 0's that some designs expect to "see" in situ.

Perhaps the existence of this paper will bring forward some intelligent justification for the trend; I hope so. Or perhaps we may stumble upon some better alternatives.

The basic measurement for timing analysis on data is the Time Interval Error (TIE) and its statistical distribution. The decomposition into jitter components requires (for oscilloscopes) analysis of many repetitions of the test data pattern. The concern I have as a counterweight to this trend, is that these longer PRBS patterns increase exponentially the amount of acquired and analyzed data, to the detriment of the quality of the results. To illustrate this point, to establish the systematic timing "edge-displacement" values (by averaging TIE for each bit position) for PRBS7 and PRBS23 by having 100 average TIE displacements for each case, which is to say 100 repetitions of the pattern recorded and analyzed, you need 2^{16} more data. That is NOT an insignificant penalty. Suppose it took 2 seconds to do the first analysis, the second would take more than 36 hours. That example is only for PRBS23.

Another concern is the confusion (in my opinion) about expectations for the invariance of random jitter (Rj) under varying test pattern PRBS length. I believe this paper will explain why this is not a reasonable expectation.

So while it is clear this trend exists, there are difficulties in managing the results from jitter testing with long PRBS sequences. So if you choose to go down this path, you should at least be aware of them.

The Observed effect

In my experience and using more than one methodology, I have observed that longer PRBS test patterns result in somewhat larger random jitter (Rj) values. By this, I mean a difference in the observed Rj for PRBS7 compared to PRBS23 that is significantly higher for the longer sequence.

This has been a nagging question for some time: "Why does the reported Rj seem to increase with increasing length Pseudo-Random-Bit-Sequence (PRBS) test patterns?" The question and the effect have become more problematic in the current culture of testing with longer and longer PRBS patterns.

I believe the reasons are not as difficult to understand as it might seem. I further claim this “apparent” growth is not merely “apparent”, but in most real-life scenarios it is real.

There is an underlying question which is probably essential to deciding whether this R_j “growth”, and accompanying “diminishing” D_j is real or significant to your application or not, and that is: “what is the purpose of obtaining R_j and D_j figures?” It should be clear to someone who is a student of statistics that the so called “dual-Dirac” model for jitter is a vast simplification, especially in the presence of an “impaired channel”. Yet the values for R_j and D_j , used to estimate a total-jitter is nonetheless a powerful tool ... so long as we remember the physical reality does not quite fit the model.

I claim there are two schools of thought on this subject. The purposes for determining a figure for R_j are to:

- 1) Characterize the random jitter in each edge of the test pattern.
- 2) Characterize the rate of growth of T_j over a region of interest of BER.

If your purpose for obtaining R_j and D_j values is to budget jitter, *and* if this includes the assumption that R_j describes the rate of growth of T_j at a BER of interest, then I believe the growing R_j is a reality you must take into account.

If your only purpose is to calculate a T_j at a BER (and only one BER), and if you intend to budget using only T_j to obtain a confidence interval at one target BER, then I would ask, why do you want R_j and D_j numbers at all? I think the answer to this is somewhat academic; it’s to quantify just that random jitter on each edge.

I don’t ask this last question lightly. It is clear to everyone that deterministic jitter increases and so total jitter increases. But if you want to estimate performance at lower and higher BER, the R_j value is important.

Assertion about Total Jitter (T_j)

I believe that all of the common methodologies in current use for oscilloscopes will predict similar values of T_j , but the methods are sometimes different for what contribution is associated with R_j , and consequently for D_j .

I also believe that the conclusions concerning the depth and shape of the bathtub curve of the eye diagram depend on a correct treatment of the statistics of the jitter decomposition, and that in many cases R_j can be mistakenly underestimated.

Review of the Methodologies

I believe it to be true that all the major oscilloscope suppliers have very similar basic processing algorithms to do basic jitter analysis. In a very brief summary this is the list of steps in the procedure:

1. A digitally recorded waveform data is analyzed for threshold crossing times, which are compared either to a regular sequence of intervals, or nearly regular intervals adapted by software PLL and “clock data extraction” (the mostly regular intervals representing the “expected” edge times of the recovered clock).
2. The observed crossing times (when there is one) are compared to the “expected” edge times and assembled into a sequence of “Time Interval Error” (TIE) values.
3. A determination of “systematic” edge displacement times for either the repeating data pattern case or for non-repeating patterns is formed by averaging TIE values, producing a set of Data Dependent Jitter (DDj) values.
4. A new set of timing error values is produced by subtracting the DDj values from TIE values at the respective positions in the sequence. The resulting set is often called the “Random + Bounded Uncorrelated Jitter” (RjBUj). This step is generally recognized as substantially simplifying the subsequent analysis.
5. The sequence of RjBUj values undergoes a spectral analysis whereby the peaks in the “jitter spectrum” are associated with additional “Deterministic” jitter, and the remaining background is assigned to “random” or Gaussian jitter.
6. A preliminary pair of Rj and Dj are determined, and from these an idealized PDF (not including the DDj distribution) is calculated.
7. The distribution of DDj values must then be re-combined with the idealized distribution to obtain an overall PDF and then a CDF that yields Tj as a function of BER.
8. To provide final reported Rj, Dj either:
 - a. Use the Rj from step 5 and work backwards to Dj using the standard dual-Dirac equation for Tj solve for Dj, or
 - b. Fit the Tj(BER) extrapolated curve to determine a new pair of Rj and Dj figures “best” representing the same standard dual-Dirac equation.

For many cases, where the DDj value distribution is either very simple or has few values (such as for short data patterns), this choice between 8a and 8b makes little difference. However the nature of long PRBS test patterns is to produce a structured and frequently “Gaussian-like” distribution of DDj values. Under this condition the difference in the final Rj and Dj can be dramatic.

The steps 8a and 8b reflect exactly the two different purposes for obtaining Rj and Dj. The result for the 8a case gives an estimate for any edge in the test pattern. The result from step 8b estimates the rate of growth of Tj(BER) at the specified BER.

Postulated Reasons for R_j Increasing with pattern length:

There are a number of reasons to believe that random jitter might increase with longer run-lengths present in long PRBS sequences.

1. One argument is that the edges following a long run of 0's or 1's can be "runt" edges, in some cases to the point that the slope (slew-rate) of the edge at the threshold crossing can be substantially different than for less challenged edges. The effect known as amplitude to phase (AM to PM) conversion, or vertical noise impact on timing jitter, affects the contribution to R_j for these "challenged" edges.

It is easy to see by looking at any lossy channel's eye diagram that such conditions exist. But, even though this is plausible, the methodology most commonly employed are relatively insensitive to this unless there is a lot of noise present, since the vast majority of edges are not impaired in this way.

2. As longer patterns are used, there is a natural compromise that is nearly always accepted which is to use a smaller set of overall pattern repetitions, thus reducing the statistical sample size and compromising the determination of systematic jitter for each edge present in the pattern. This reduction in the number of TIE values averaged to obtain DD_j values means that the uncertainty in these systematic values is increased. These statistical errors remain in the sequence of data analyzed to obtain an estimate of R_j, and can masquerade as additional random jitter. Please note, this contribution can *only* increase the calculated R_j.
3. Another issue, which is rarely mentioned, is the nature of the DD_j distribution and its impact on the final CDF. For short test patterns, this effect is nearly negligible, however for very long test patterns, whether they are PRBS or not, the consequence can be very significant. The effect is quite simply statistical and manifests as a difference between the results you obtain using step 8b instead of step 8a. That is, if you take into account the detail of the DD_j distribution, you may obtain significantly different R_j results for long test patterns as compared to shorter patterns.

The first of these can be ruled insignificant for cases where the vertical noise in the measurement is low. The second can be mitigated by collecting and analyzing more data. The third however is observed in very "unimpaired" measurement scenarios, like a lab generator connected via high quality cables directly to an oscilloscope. For this reason this 3rd statistical mechanism shall be the focus of the rest of this paper.

A Simple Monte-Carlo “Experiment”

If you observe the distribution of DDj for real or Monte-Carlo (simulated) serial data in the presence of “no channel”, the DDj distribution is a delta-function. It has no width because there is no mechanism to cause inter-symbol interference (ISI). If however you introduce a mild low pass filter as a source of ISI, not even enough to be called an “impairment”, since the eye is wide-open, you observe an DDj distribution that is bell-shaped. The width is small, but the shape is unmistakable.

Below, a Monte-Carlo simulation of a simulated PRBS23 30Gbps data stream that has a simple low pass filter at 22GHz, the eye diagram seems wide-open, nonetheless the DDj distribution exhibits a Gaussian-like shape, even though with a small sigma ($\sim 48\text{fs}$). The p-p is about 500fs or ± 5 sigma.

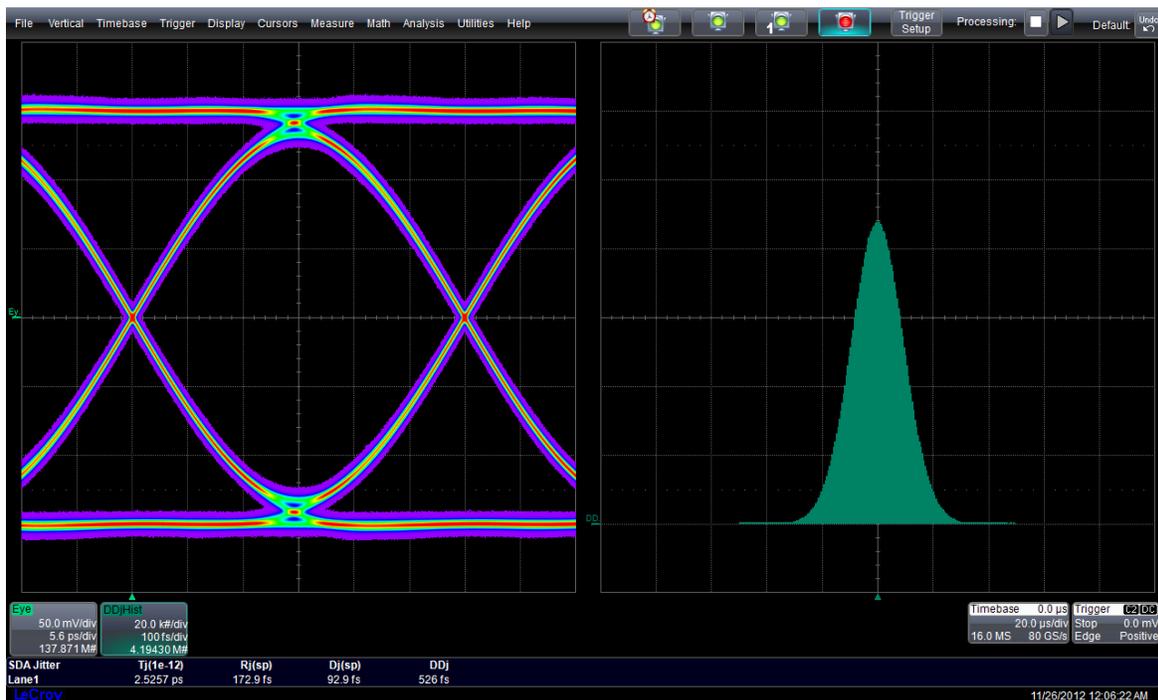


Figure 1: Monte-Carlo simulation of band limited serial data, using PRBS23, showing the DDj Histogram is Gaussian-like, as though it were truncated at roughly ± 5 sigma.

So you can see, even a small loss produces a DDj distribution for PRBS23 that is quite small, but has a characteristic shape. Be aware, nonetheless that as the ISI becomes stronger, more structure is introduced in the DDj distribution as multiple peaks appear. The fact that the tails are extremely rare does not change. This more complex case will not be treated in any detail.

Measurements on 4 pattern Lengths

Below is an image showing 4 measurements of jitter, and graphics of the DDj distribution for PRBS7, PRBS11, PRBS15 and PRBS23. You can easily identify which is which by the total populations of 64, 1024 (upper left and upper right), and 16384 and 4.19430 million (lower left and lower right). These 4 distributions tell an interesting story. The story is about how the distribution of DDj develops as the number of edges in the PRBS test pattern grows larger and larger, and the total breadth of the distribution also grows larger and larger.



Figure 2: A 30 Gbps data stream, at PRBS7, 11, 15 and 23 is analyzed for jitter and DDj Histogram is shown. Notice two things, the tails are more and more extreme, and the statistical significance of the contribution of the extremes is smaller and smaller. The PRBS23 DDj distribution has a p-p of about 8ps, and full-width-half max of about 2.7 ps.

This progression illustrates how the p-p width of the DDj distribution grows from 3.8 ps to 8.1 ps. And moreover, the contributors at each extreme (when convolved with the RjBUJ distribution) go from 1/64 for PRBS7, to 1/1024 for PRBS11, to 1/16384 for PRBS15 to 1/(4.2 million) for PRBS23.

To restate, for PRBS7 when the DDj distribution is re-convolved with the random jitter for every edge, the contribution of the displaced edges at the extreme left and right are only about 2 orders of magnitude less than unity. For PRBS11, about 3 orders of magnitude below unity and for PRBS15 4½ orders of magnitude less. But for the PRBS23 case, the extreme contributions are more than 6½ orders of magnitude less than unity. As an exercise, the reader can do the math for PRBS31.

The second part of the story is that the underlying shape of the DDj distribution is essentially the same. The PRBS15 and PRBS23 shapes are very similar, only the tails become more refined. This is much like a “truncated” Gaussian. In fact the standard deviations of all 4 distributions are approximately the same.

If for example you were required to make a histogram of an ideal Gaussian, but the total population of the histogram is limited to N, then the histogram would be cutoff at some maximum and minimum horizontal coordinate. The larger N is, the wider the histogram will be, even for the same sigma.

As a check for whether the observed DDj distribution is “real” the same signal source was measured using another oscilloscope. As you can see, the distribution is very similar in width and shape for the PRBS23 case. The point is that this effect, “the broadening of the DDj distribution” is indeed observable, and is likely not an artifact of either oscilloscope.

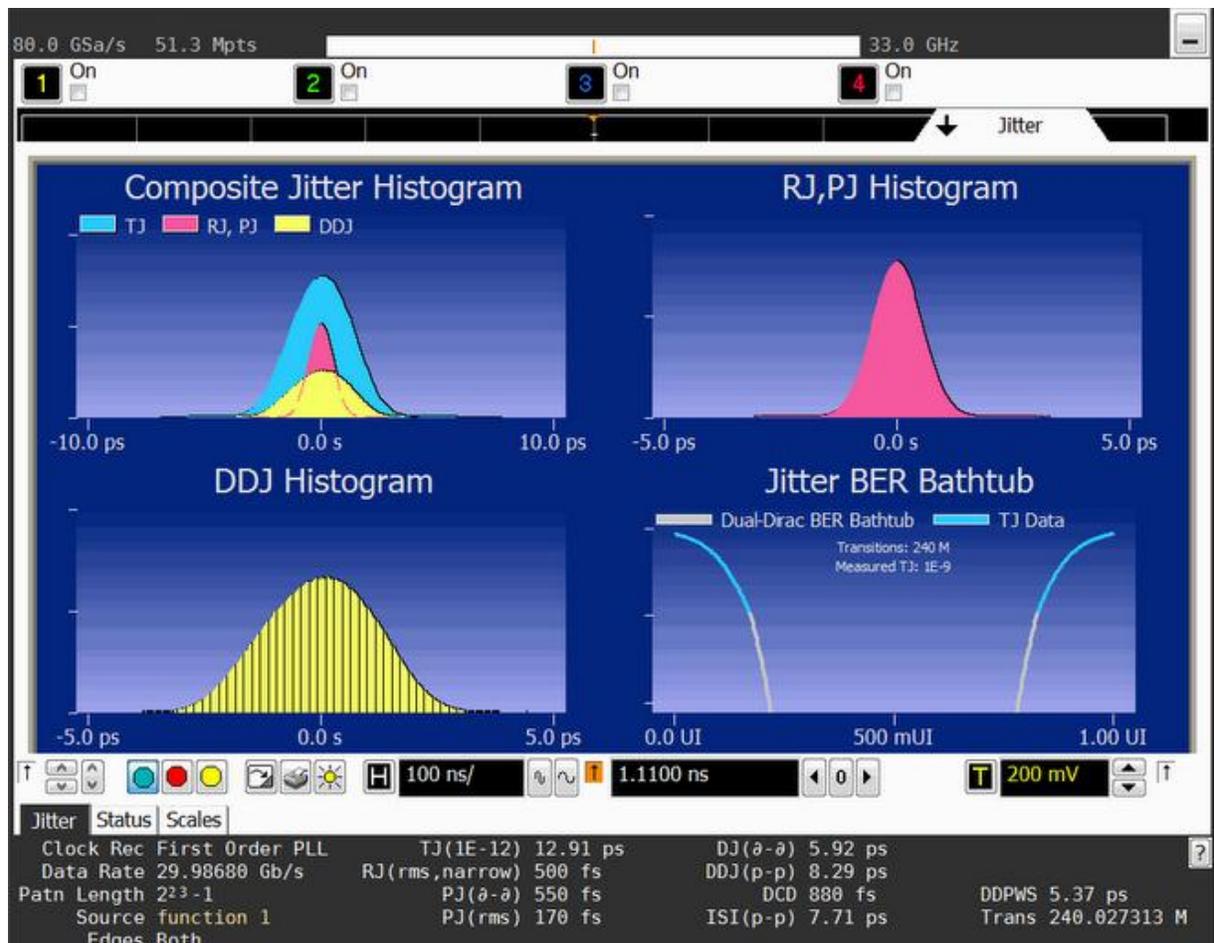


Figure 3: The same 30Gbps data stream, using high quality cable and another oscilloscope. Notice the DDj (DDJ) Histogram, and that it is peculiarly “Gaussian-like” with a p-p of about 8ps, and full-width-half-max of about 3 ps. This agrees well with the PRBS23 DDj distribution shown in Figure 2.

What is your purpose for estimating Rj?

If your purpose in determining a useful value for Rj is to budget jitter, you may be interested in how Tj is growing as BER decreases. For a perfect Gaussian, the growth of Tj as BER decreases is a constant. This constant is related to the sigma of the distribution. Of course this is unique to a pure Gaussian distribution. When the cumulative distribution function (CDF) of a Gaussian is shown on a Q-scale, the CDF appears as a straight line (and if symmetrized as two straight lines). The derivation is not given here, but the reciprocal of the slope of this line is the sigma of the Gaussian. This “sigma” for jitter measurements is associated with Rj.

When the distribution is NOT Gaussian, the slope of the CDF on the Q-scale is also NOT constant, but is rather a function of BER. We will not go into the derivation of this but the classic “dual-Dirac” equation can be expressed as:

$$T_j(BER) = D_j - 2Q\left(\frac{BER}{2\rho_x}\right) \cdot R_j$$

Where $Q(x)$ is the well-known function related to the inverse error function, and ρ_x is the transition density of the data pattern (approximately $\frac{1}{2}$ for PRBS patterns). It should be obvious the rate of growth of $T_j(BER)$ is proportional to R_j .

A common representation of the CDF on a Q-scale displays Q as the y-coordinate and the horizontal axis is time, $T_j(BER)$ being the width of the CDF at any given BER. The slope of the CDF on this display has the coordinates switched and so R_j is the reciprocal of the slope (i.e. the larger or steeper the slope, the less rapidly T_j is growing, and the smaller R_j is).

Since this slope for many distributions is not constant, perhaps we should not assume R_j is constant, but rather it is a function of BER or $Q(BER)$.

A Simple Calculation Concerning Truncated Gaussians in Combination with a Pure Gaussian:

Even if the DDj distributions are not exactly Gaussian, there is enough reason to investigate how truncated Gaussians combine with pure Gaussian distributions, and how the growth of the resulting $T_j(BER)$ behaves.

If we want to investigate the effect of a calculated truncated Gaussian and an ideal Gaussian, this can be easily accomplished by the following steps:

1. Calculate a truncated Gaussian PDF (i.e. histogram) with a fine horizontal resolution and a defined sigma.
2. Convolve this truncated histogram with an ideal Gaussian of defined sigma, by evaluating a sum for each bin at an offset equal to the bin position and a weight equal to its population in the truncated histogram with the ideal Gaussian, forming a new histogram, creating a new histogram with at least ± 22 sigma of range. (see pink histogram below)
3. Calculate the CDF of the composite histogram by summing the bins from the left and from the right, and display this versus BER on a log scale. (see lime-green trace below)
4. Calculate and display the CDF on a vertical axis of Q. (see red trace below)

5. Calculate the tangent (slope) of the CDF on this Q-scale, and determine a variable result for R_j as a function of BER (blue trace below)
6. And finally the resulting D_j as a function of BER (green trace)

For the graphs shown below, I have chosen an equal size R_j (truly random) and standard deviation of the truncated Gaussian. Both have sigma 0.25 ps. This particular case truncates the 1st Gaussian at $\pm 3\sigma$.

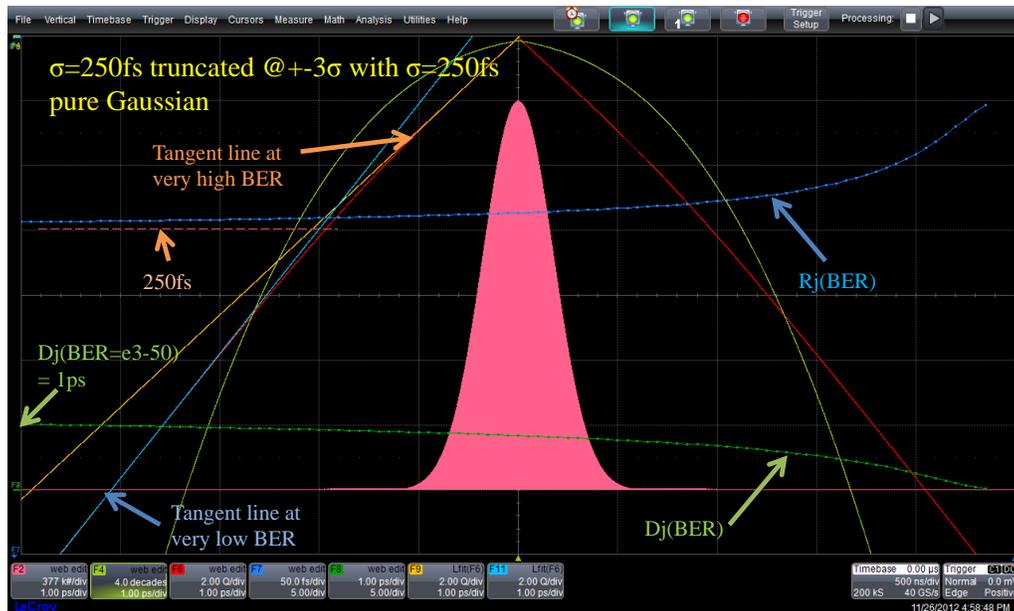


Figure 4: Calculations of truncated Gaussian ($\pm 3\sigma$) with ideal Gaussian, and derived traces. Note the blue trace is R_j from BER = $1e-50$ on the left to BER = $1e(-1.5)$ on the right (the far right of the grid is at BER = $1e-0$).

For contrast, below is a similar case, but with a 3 times larger standard deviation for the DD_j distribution (a more lossy channel), truncated at $\pm 6\sigma$. You see much more clearly the change in slope of the red CDF on Q-scale. For this case the quadrature sum for the .75ps and .25ps is .791ps which is the asymptotic value for R_j at very high BER. In this case it's noteworthy that the rate of growth of T_j has not reached its asymptote for R_j of 250 fs even at the BER as low as 10^{-50} . This should give pause to anyone who may think the effect is unimportant for budgeting timing error.

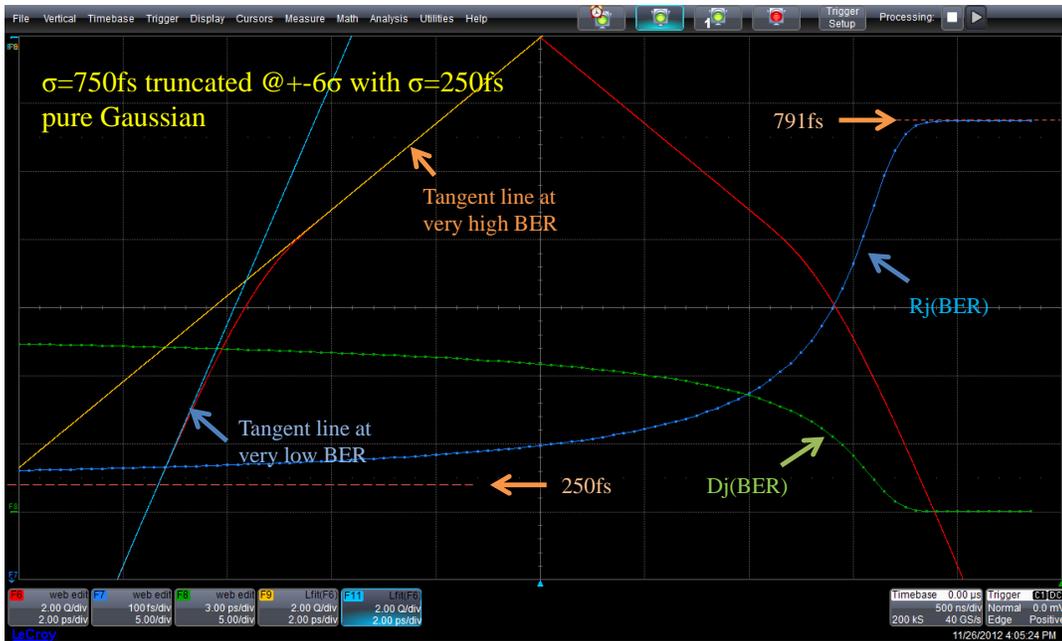


Figure 5: Calculations of truncated Gaussian ($\pm 6\sigma$) with ideal Gaussian, and derived traces. Note the blue traces Rj from BER = 10^{-50} on the left to BER = $10^{-1.5}$ on the right (the far right of the grid is at BER = 10^0).

Notice for figure 4 the Dj(BER) remains close to zero until the effect of the truncation sets-in at around BER= 10^{-7} . Notice also that at BER= 10^{-12} the apparent Rj(BER) is still 200fs (80% higher) above the asymptotic value of 250fs.

Below is a series of a sequence of models, with different levels of truncation. The composite graph is very informative, showing how the level of truncation pushes the transition from a quadrature combination of jitter to the asymptotic random jitter over a number of combinations.

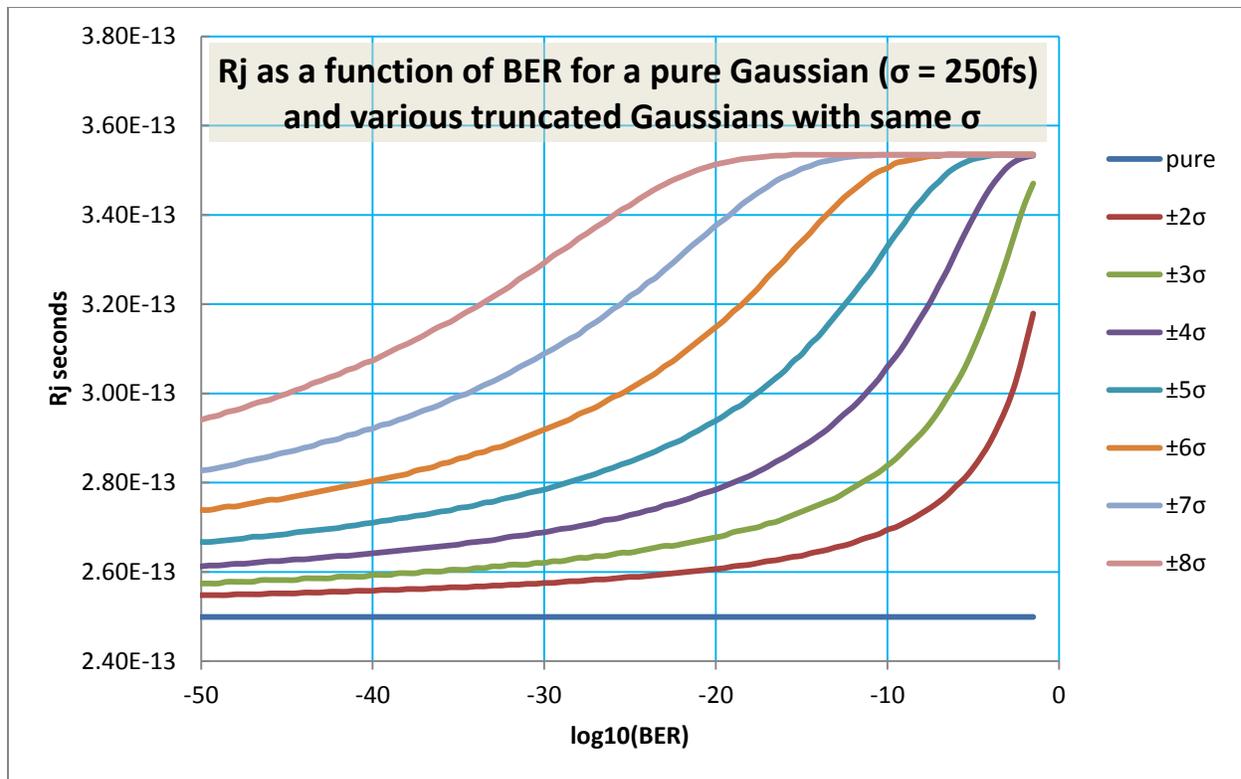


Figure 6: A perfect Gaussian with sigma = 250 fs convolved with various truncated Gaussians (sigma also 250fs). If the second distribution were not truncated, the composite sigma would be 354 fs. Here “Rj” values are the tangent of the CDF at each BER.

If two ideal Gaussians are convolved, the resulting sigma (assuming incoherence) is the quadrature sum of the two sigmas, so in this case 0.354 ps would be the quadrature sum. You will notice the cases with $\pm 8, 7, 6, 5, 4\sigma$, the very high BER points all converge at this value for $R_j(\text{BER} = 1e^{-1.5})$. Notice also, that even with a moderately small sigma for the (truncated) DDj distribution, the Rj values are not settled at $1e^{-12}$ and are still not settled to the expected asymptotic value of 250fs at $1e^{-50}$.

Comparison to a Bit Error Rate Tester’s decomposition of Rj and Dj

With regard to Rj, Dj decomposition, a BERT measures the edges of the bathtub curve looking precisely for the shape of Tj as a function of BER. So here is an example of how a minimum loss measurement of jitter also shows increasing Rj with PRBS pattern length. It is clear that whatever statistical effects may be at work, they cannot hide from this methodology. These measurements were made using a bit rate of 10.7 Gbps and short high quality cables. (I did not have BERT equipment at my disposal for a 30Gbps measurement).

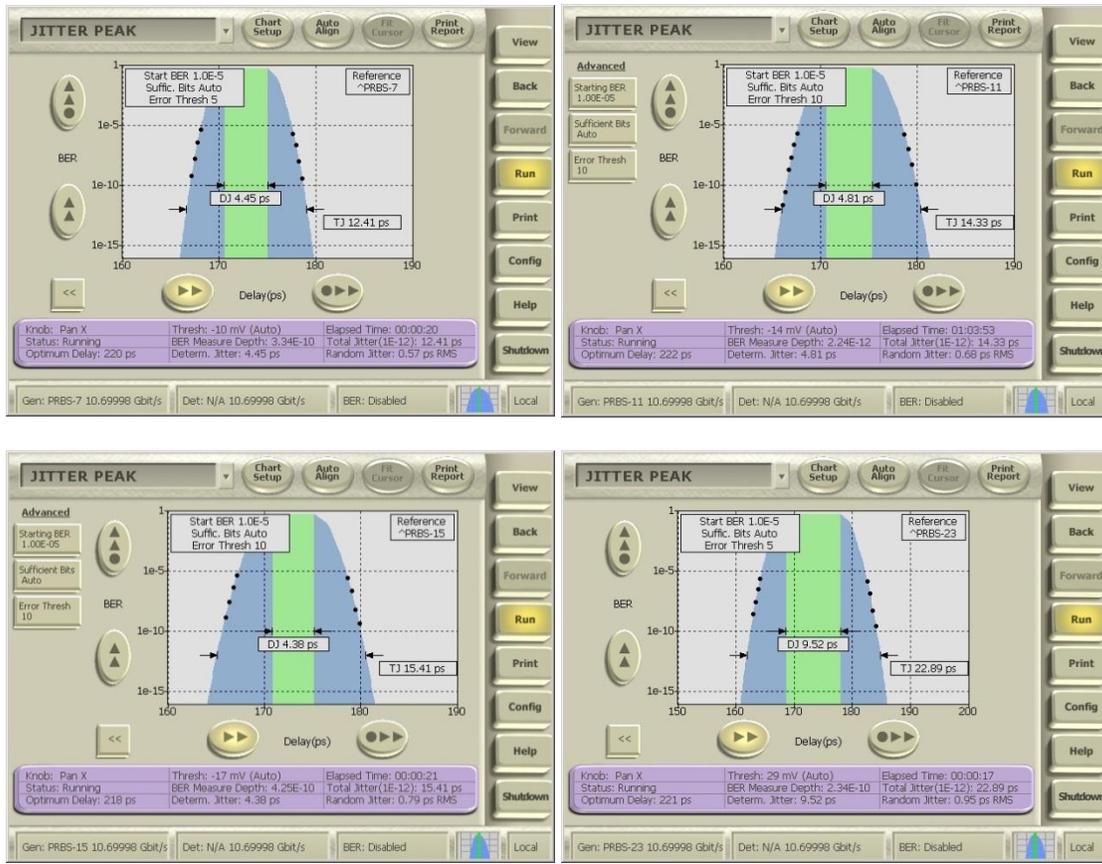


Figure 7: PRBS7, 11, 15 and 23 using the internal pattern generator of a BERT, shows increasing values for Rj even though the connection between generator and BERT input was made with short high-quality cables. Of course the BERT is oblivious to any notion of Rj for individual edges. A BERT can only see the impact on BER of the whole system.

This measurement illustrates how the “real” effect of increasing Rj cannot be ignored. It is clear that over the region of BER that the BERT has measured, the “effective” Rj measured by the BERT is *not* constant for different PRBS pattern lengths. A BERT cannot identify the nominal Rj for each individual edge, so it cannot report a value like the Rj from step 8a (see page 5). As such the BERT reflects the expected result for method 8b (see page 5) in oscilloscopes.

Bit Error Rate Tester	Tj(1e-12) (ps)	Rj (ps)	Dj (ps)
PRBS7	12.4	.57	4.45
PRBS11	14.3	.68	4.81
PRBS15	15.4	.79	4.38
PRBS23	22.9	1.09	9.52

Table 1: PRBS7, 11, 15 and 23 using the internal pattern generator of a BERT.

Compare the results for an oscilloscope measurement using steps 8a and 8b to the results from the BERT measurement. For the results from type 8a results, of course the Rj values do not systematically grow, and for method 8b they do.

Scope (8a)	Tj(1e-12) (ps)	Rj (ps)	Dj (ps)
PRBS7	12.9	.36	7.8
PRBS11	14.7	.39	9.1
PRBS15	16.1	.57	8.1
PRBS23	17.3	.56	9.3

Table 2: PRBS7, 11, 15 and 23 using the pattern generator of a BERT, but measured with an oscilloscope using method 8a. In this case we expect constant Rj.

Scope (8b)	Tj(1e-12) (ps)	Rj (ps)	Dj (ps)
PRBS7	12.9	.40	8.95
PRBS11	14.7	.45	4.81
PRBS15	16.1	.70	7.90
PRBS23	17.3	.82	7.28

Table 3: PRBS7, 11, 15 and 23 using the pattern generator of a BERT, but measured with an oscilloscope using method 8b. In this case, we do NOT expect a constant Rj

So it's clear that for the BERT measurements, even though this is a nearly ideal case , the reported Rj is not constant.

Conclusions:

When NRZ serial data is transmitted, there is always a medium or “channel”. However small, some loss is impossible to avoid. Measurements, models and Monte-Carlo results all confirm that Rj is not a constant over increasing PRBS length.

1. Whether you care about Rj as a measure of random jitter on each edge, or whether you are concerned with the rate of “growth” of Tj(BER) is a crucial factor in how to estimate Rj. Choose you method according to your need.
2. In the presence of an imperfect channel, the nature of the data dependent jitter distribution behaves roughly like a truncated Gaussian, for Monte-Carlo models and for observed cases using two different types of oscilloscopes. (Note: stronger ISI produces DDj distributions with more complex structure, but the tails are always there.)
3. Mathematically, a combination of a truncated Gaussian with a pure Gaussian predicts shapes for the Tj(BER) which show significant contributions to the slope of the jitter bathtub that are non-vanishing even for *very* low BER levels.

4. Measuring R_j with a BERT *does* exhibit increasing R_j with pattern length, which should not be a surprise since the effects of DDj cannot avoid being observed with this methodology. The shape of the edges of the bathtub curve necessarily contains this effect.

Recommendations and Questions:

There are a number of practical issues that make using long test patterns difficult. Compromise on the number of repetitions that can be analyzed is not the least of these difficulties. Time is valuable and limiting statistical uncertainty is important.

If your concern is the random jitter on each edge, then why bother with long patterns?

If on the other hand you are interested in the shape of the “bathtub”, then why not use a shorter pattern (say PRBS15) to eliminate pesky statistical uncertainty and time consumed in the lab, and get a solid estimate for individual edge R_j 's. Then use this and the standard deviation of the DDj distribution to predict the tails of an inevitable wider truncation for larger patterns to predict the overall impact on $T_j(\text{BER})$?

Ask yourself also: does the expected traffic resemble PRBS type patterns. Are the rare events in live traffic in any way similar to the rare extremes of PRBS test patterns? It should be clear that will affect how your real-world performance will relate to your test cases.