

## **The Meaning of Total Jitter**

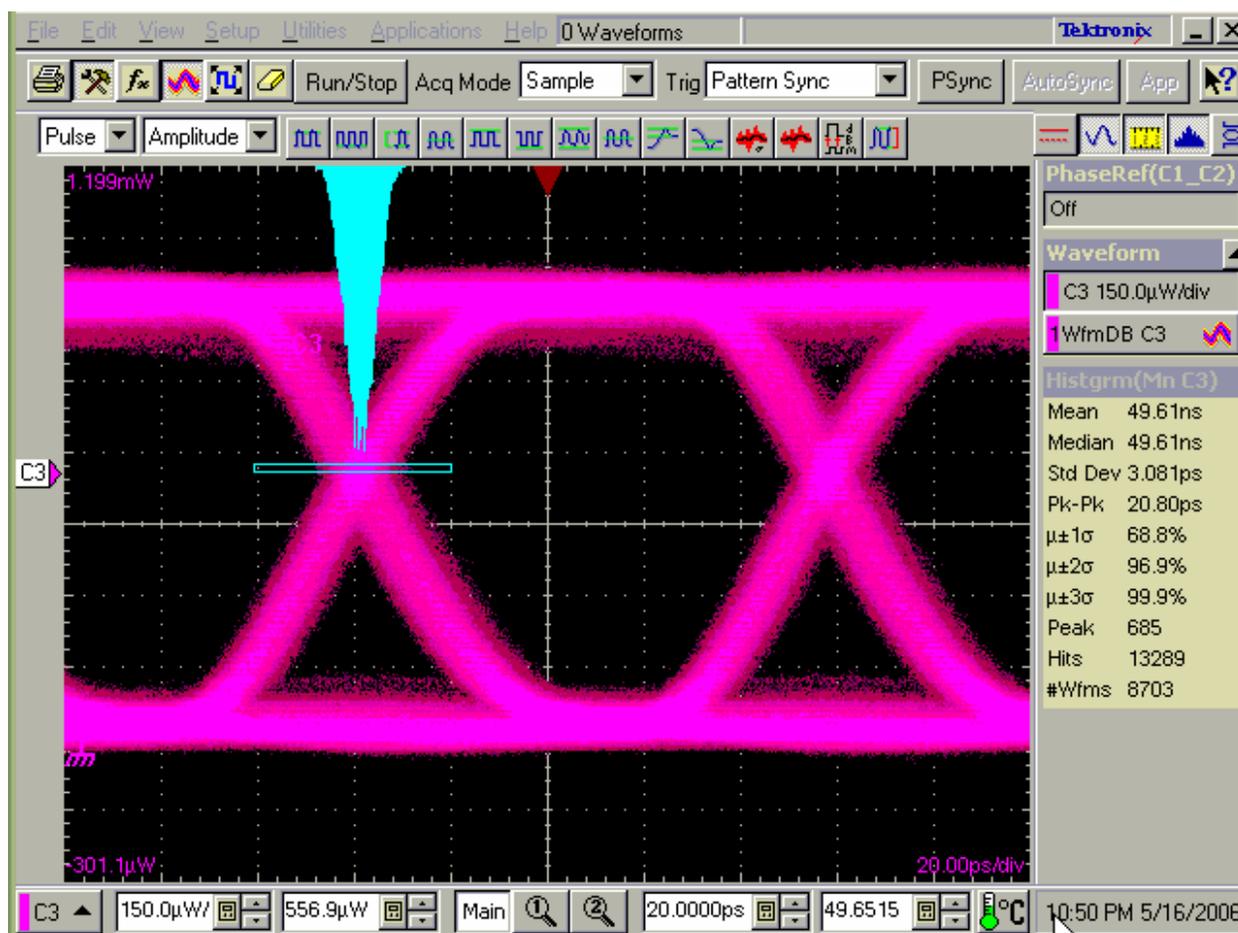
Ransom Stephens, Ph.D.

### **Abstract:**

Total Jitter is an increasingly important quantity in the development and specification of serial data links but, while it is well defined, it is not well understood. Total Jitter is like peak-to-peak jitter but referenced to a given Bit Error Ratio; it is the amount of eye-closure at a given BER. This paper clarifies the meaning of TJ and shows how measurements of jitter subcomponents, like Random Jitter (RJ) and Deterministic Jitter (DJ), can be used to provide fast accurate estimates of TJ(BER).

In analyzing jitter we have two separate but related concerns. First, we need a standard way of predicting how well two network elements, say a transmitter and a receiver, will work together and maintain an acceptable Bit Error Ratio (BER) – these are interoperability requirements. It is to this end that Total Jitter at a Bit Error Ratio was defined. We'll use the notation, TJ(BER), to keep things clear because the term "Total Jitter" is often used to describe quantities and distributions that are less well defined. Depending on the context, "Total Jitter" may refer to a measurement of the jitter Probability Density Function (PDF) which gives the probability of a logic transition occurring at a given time. A measurement of the jitter PDF can be performed on an oscilloscope, as shown in Figure 1, by making a histogram of the crossing point. "Total Jitter" is also used to distinguish the jitter of the whole system from jitter that is caused by its parts. In this paper, TJ(BER) is the total jitter defined at a bit error ratio.

Which brings us to our second concern in jitter analysis: diagnostics. If you have a problem with jitter – and that can only mean one thing, the BER is too high – just knowing the value of the BER, the peak-to-peak jitter, or even TJ(BER), doesn't help you solve the problem. For diagnostics, you need to break the jitter down into separate categories that, hopefully, point in the direction of the causes of jitter.



**Figure 1: An eye diagram showing a measurement of the jitter probability density function as a histogram projected from the crossing point.**

Assuring interoperability and performing diagnostics are not independent. Measurements of different types of jitter can be used to make accurate estimates of TJ(BER) – *estimates* that, in many cases, are more accurate than *measurements*! In this paper, we concentrate on TJ(BER) – what it is, why it is useful, and how it is measured and estimated. We'll cover diagnostics in another paper.

One point that you should keep in mind is that jitter analysis, by its very nature, is a simplification. Analyzing jitter independent of voltage noise can be useful, but no thorough signal integrity analysis neglects the relationship of the two.

## In Search of Peak-to-Peak Jitter

We need a quantity that manufacturers of different network elements can use to judge whether or not their parts can operate in a system to a specified BER. The naïve solution is simply to measure the peak-to-peak jitter. The problem with peak-to-peak jitter is that the value you measure depends on how long you make the measurement. It's even worse than that, if we all agree to measure peak-to-peak jitter for, say, 100 seconds, we'll still get different results. It is in this sense that peak-to-peak jitter is ill-defined.

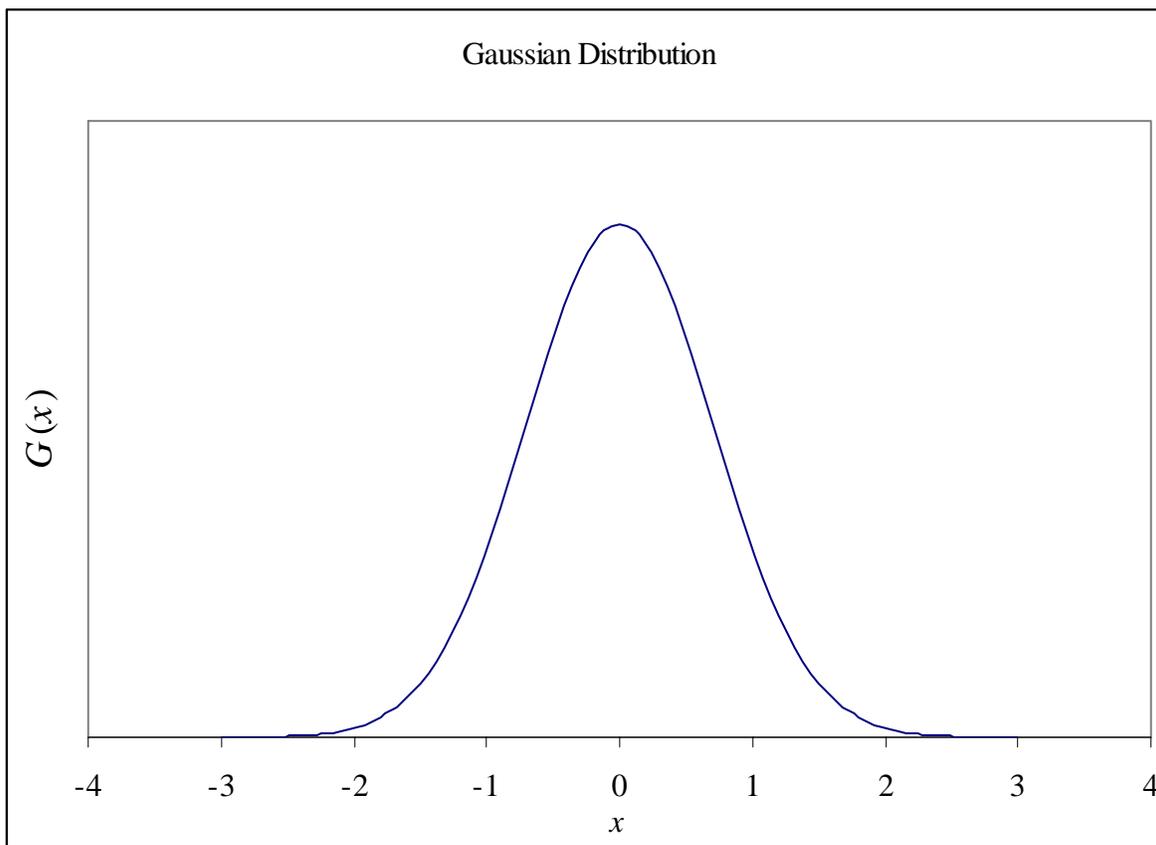
To find a quantity that both links jitter to the BER and is well defined, it's useful to distinguish random jitter from deterministic jitter. Deterministic Jitter (DJ, from here on) is caused by things like impedance mismatches and electromagnetic interference. If we knew everything about the circuit, the effect on the signal from such causes could be determined. DJ is also limited in its effect. That is, DJ can only shift the timing of a logic transition by some maximum amount. It is in this sense that DJ is bounded and, therefore, has a well defined peak-to-peak spread which we'll call DJ(p-p).

Random Jitter (RJ), as the name implies, is caused by things like thermal oscillations, flicker, and shot noise that result in levels of jitter that cannot be predicted on an edge-by-edge basis. The thing to understand about RJ is that it is caused by a huge number of very small effects. The variation in the width of the traces on a printed circuit board, fluctuations of conductivity in a conductor caused by impurities, variations in resistance caused by random fluctuations in the local voltage that individual electrons feel, and literally zillions of other tiny effects combine in a way that is dictated by the fundamental theorem of Statistical Physics: The Central Limit Theorem, which makes the whole thing very simple.

The Central Limit Theorem says that a large number of small independent processes combine in a way that follows a Gaussian distribution. This means that the variation in the timing of a specific logic transition from the ideal time is centered at zero, can occur arbitrarily early or late, but nearly always occurs within a few units of sigma from the ideal. Sigma, or  $\sigma$ , is the width of the Gaussian distribution. It's not the full-width-at-half-max, but it's the same idea and it's the only thing we need to know about the distribution;  $\sigma$  is also the rms value of the RJ distribution and gives a complete description of RJ. If we define  $x$  as the time-delay relative to the ideal time of a logic transition, then the RJ probability density function is given by

$$G(x) = N \exp\left(-\frac{x^2}{2\sigma^2}\right) \quad (1)$$

Don't worry about the constant,  $N$ , in front of the exponential in Eq. (1), it's just there to force the integral of the distribution to one. As you can see in Figure 2,  $G(x)$  is unbounded. The probability of RJ shifting a logic transition by an amount  $x$ , is proportional to  $G(x)$ , the height of the curve at  $x$ .



**Figure 2: A Gaussian, or normal, distribution. The probability of RJ causing a logic transition to be jittered in time by an amount  $x$  is proportional to the height of the curve at  $x$ ,  $G(x)$ .**

### **TJ(BER) – Total Jitter at a Bit Error Ratio**

Notice that, since  $G(x)$  never actually makes it to zero, it is possible for RJ to cause a logic transition to happen ten years from when it's supposed to. In the same sense, the laws of thermodynamics allow for the possibility that all of the air molecules in the room where you are sitting can spontaneously condense in a corner, leaving you in a vacuum where your eyes can pop out – it's possible, but so unlikely as to be irrelevant. But this possibility alludes to the reason that peak-to-peak jitter measurements get larger the longer that they are measured.

Because RJ is unbounded, peak-to-peak jitter is ill-defined.

But even if peak-to-peak jitter were well defined, it wouldn't do us much good. Remember, the only reason we care about jitter is if it causes errors. What we really need is something like a peak-to-peak jitter measurement that tells us about the amount of eye-closure. If we used an oscilloscope to measure peak-to-peak jitter and waited long enough, eventually the eye would close, that is, the peak-to-peak jitter would be the full bit period.

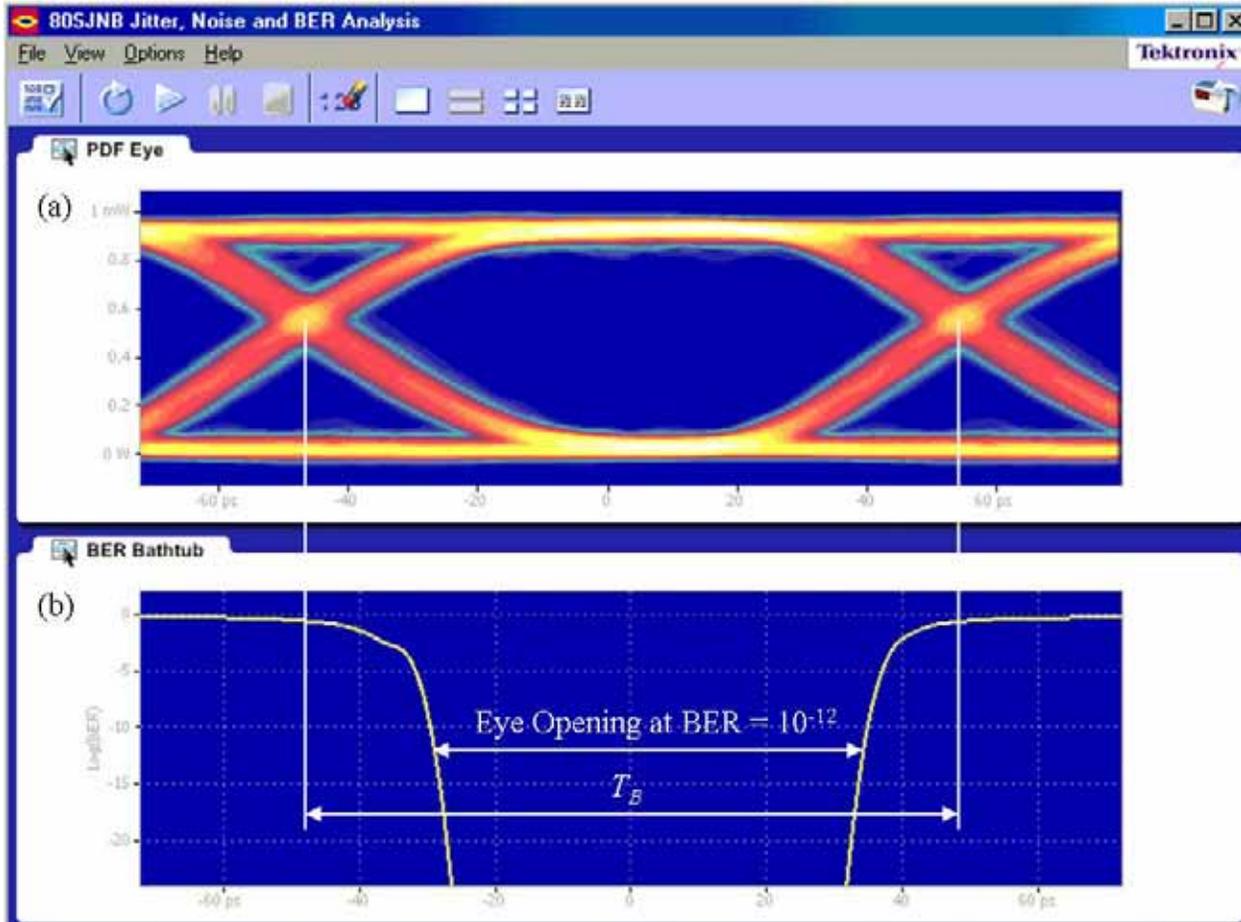
Think of a signal whose BER, when measured on an ideal receiver, is  $10^{-12}$ . The peak-to-peak jitter measurement would be the bit period, indicating total eye closure, after something like  $10^{12}$  bits were transmitted. It's not completely well defined because random fluctuations would still vary the amount of eye closure even after  $10^{12}$  bits, but now at least the peak-to-peak jitter is explicitly correlated to the BER. It is in this sense that Total Jitter defined at a Bit Error Ratio, TJ(BER), is like a peak-to-peak jitter measurement.

TJ(BER) relates the level of jitter to the bit error ratio of interest: it is the amount of eye closure at a given BER. In the example above, the simple peak-to-peak jitter measurement gave total eye closure after  $10^{12}$  bits. A measurement of TJ( $10^{-12}$ ) would reliably give TJ( $10^{-12}$ ) =  $T_B$ . But what if the actual BER were  $10^{-18}$  and the system were specified to  $10^{-12}$ ? In this case, after  $10^{12}$  bits were transmitted the eye would still be open, TJ( $10^{-12}$ ) would give us the amount of eye closure at a BER of  $10^{-12}$  for an eye that has a BER of  $10^{-18}$ .

Figure 3a shows an eye diagram and Figure 3b shows the "bathtub plot" that corresponds to it. A bathtub plot is a measurement of the BER as a function of the time-delay position of the sampling point – BER(x). It can be measured by stepping the sampling point horizontally across an eye diagram and measuring the BER at each point. At the edges, BER is very high and, as the sampling point is stepped to the eye-center, the BER drops precipitously until the sampling point approaches the other edge, when the BER increases again.

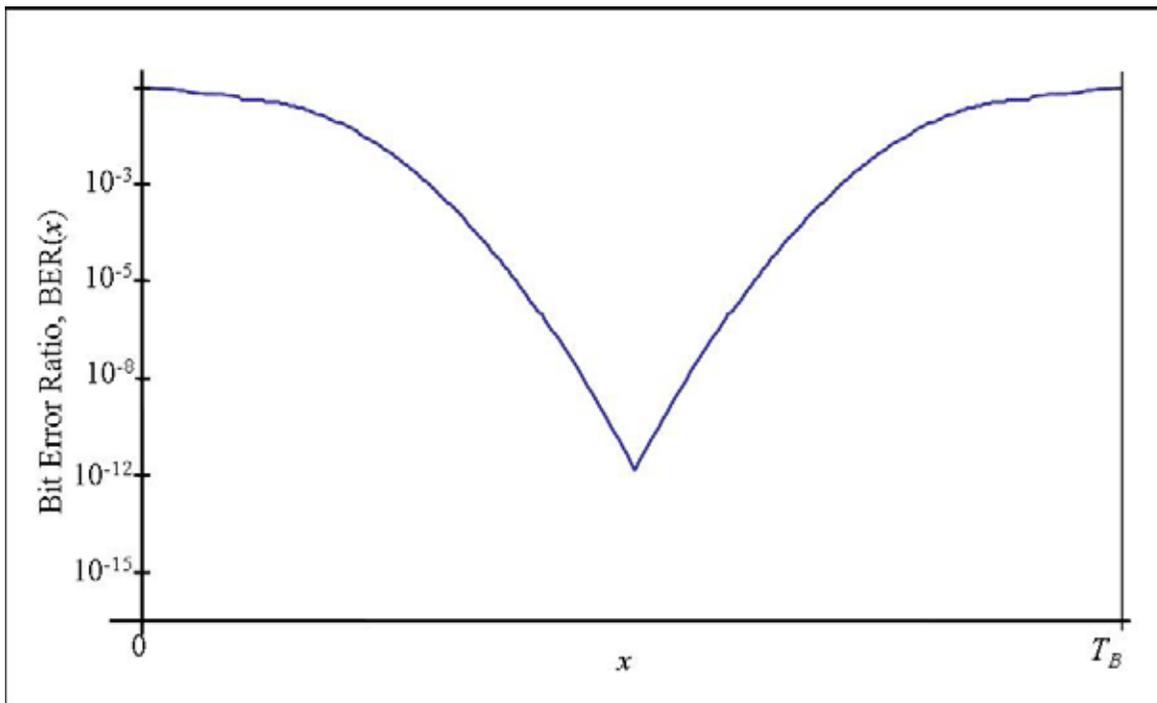
The *eye opening* at a BER is the distance between the two sloping curves in Figure 3b at the desired BER. TJ(BER) is the *eye closure* at the desired BER: the nominal bit period,  $T_B$ , less the eye opening at that BER.

$$TJ(BER) = T_B - \text{eye opening}(BER) \quad (2)$$



**Figure 3: (a) An eye diagram and, (b), the corresponding “bathtub-plot”: The bit error ratio as a function of the time-position of the sampling point with respect to the ideal logic transition time, BER(x).**

Now consider Figure 4. The BER of this system is exactly  $10^{-12}$ . You can tell because the eye is completely closed where the two slopes meet. This also means that  $TJ(10^{-12}) = T_B$ , and, for that matter,  $TJ(10^{-13}) = TJ(10^{-30}) = T_B$ , because the eye can only close completely.



**Figure 4: The bathtub plot of an eye that is closed at BER =  $10^{-12}$ .**

Strictly speaking the only practical way to *measure* TJ(BER) at BERs lower than about  $10^{-8}$  is with a Bit Error Ratio Tester. A BERT can measure TJ(BER) in two ways. It can do the brute force measurement, as in Figure 3 and Figure 4, of BER(x) to get TJ(BER), but measuring BER(x), to low BERs like  $10^{-12}$  takes a long time. It's the sort of measurement that you might start in the evening and expect to be finished the next morning. TJ(BER) can also be measured on a BERT by using a bracketing technique which cuts the measurement time down to about half an hour for a 5 Gb/s signal.

The reason that TJ(BER) measurements take so long is simple. In principle one cannot measure the rate of occurrence of an event without probing the system with a statistical sample sufficient to observe the event. Thus, measurements of TJ(BER) require the detection of at least  $6/\text{BER}$  bits ( $6 \times 10^{12}$  bits for BER =  $10^{-12}$  at 5 Gb/s would take 20 minutes) bits for the bracketing technique and about  $100/\text{BER}$  bits ( $10^{14}$  bits for BER =  $10^{-12}$  at 5 Gb/s would take five and a half hours) for an accurate brute force measurement of BER(x) to get sufficient statistics to reference the eye closure to the BER. In an industry where Moore's Law rules, it's easy to understand why such a great effort has gone into finding techniques that accurately *estimate* TJ(BER) from a much smaller data sample.

The largest uncertainty in a BERT measurement of TJ(BER) is introduced by the error detector sensitivity. Without going into a lot of BERT detail, “the error detector sensitivity” is the smallest difference between logic levels a BERT can distinguish, typically around 50 mV – much larger than the voltage resolution and accuracy of either a real-time or equivalent-time oscilloscope. The time-base of BERTs manufactured in the last few years are quite accurate and contribute less uncertainty than does the sensitivity; for BERTs manufactured more than five years ago, timing nonlinearities and time-base resolution make TJ(BER) measurements very difficult.

The point is that TJ(BER) can only be *measured* on a BERT, but if it could be measured on an oscilloscope, the measurement would be more accurate.

## RJ and DJ

There are two ways to get TJ(BER) without having to wait as long as is required to accumulate 6/BER bits. First, if each component of the jitter distribution can be separately measured, including RJ and all types of DJ, then they can be convolved into the jitter PDF from which BER(*x*) can be calculated for any *x*. Given BER(*x*), it’s easy to extract TJ(BER) as in Figure 3b and Eq. (2). A simple measurement of the Jitter PDF, like the histogram in Figure 1, doesn’t have enough data to build a PDF that is accurate enough to extract BER(*x*). Further, in complicated cases, even with the best jitter analysis techniques, it’s not possible to distinguish every DJ source and measure the correlations between the different DJ sources as is necessary to build the complete PDF.

The other way to get TJ(BER) is to use the dual-Dirac model. The dual-Dirac model provides an easy way to estimate TJ(BER) by measuring the RJ  $\sigma$  and a quantity much like the peak-to-peak spread of DJ, DJ(p-p).

$$TJ(BER) = 2Q_{BER} \times RJ + DJ(\delta\delta) \quad (3)$$

Here,  $Q_{BER}$  is a factor that relates the BER to the distance from the center of the RJ Gaussian, listed in Table 1 for different BERs, and DJ( $\delta\delta$ ) is a model dependent parameter that is easy to define and can be measured in many different ways.

BER	$Q_{BER}$
$10^{-10}$	6.35
$10^{-11}$	6.70

$10^{-12}$	7.05
$10^{-13}$	7.35
$10^{-14}$	7.65

**Table 1: Values of  $Q_{BER}$  for different BERs.**

We will cover the details of the dual-Dirac model in another paper. In a nutshell, the dual-Dirac model uses the fact that, for a signal with only RJ,  $BER(x)$  can be calculated from Eq. (1) and assumes that the tails of almost all jitter distributions follow the RJ Gaussian distribution, Figure 2, shifted toward the center of the eye.

Whether we calculate  $TJ(BER)$  by measuring RJ, all sources of DJ, convolving them together and integrating the resulting PDF, or by measuring RJ and  $DJ(\delta\delta)$  and using Eq. (3) we are making the same fundamental assumption: that there are no rare processes whose observation would require a huge statistical sample. Integral to this assumption is the idea that the RJ Gaussian accurately represents the tails of the jitter PDF. If it doesn't, then the estimates of  $TJ(BER)$  will be biased. At a deeper level, there is also some question about whether or not what we consider RJ is actually random. The Central Limit Theorem indicates that if there are enough small DJ effects, their distribution will mimic a Gaussian – a truncated or “bounded” Gaussian. In practice, these are not problems that test engineers need to worry about; they are problems that justifiably prevent the designers of technology standard specifications from sleeping at night. We can never be absolutely be certain that the tails of the jitter PDF follow the RJ Gaussian without acquiring that  $100/BER$  statistical sample, but, in the vast majority of cases, careful eye-diagram analysis, thorough study of jitter diagnostic information, and experience with your design provide clues of when your  $TJ(BER)$  estimates are being corrupted.

In most cases  $TJ(BER)$  can be estimated very quickly and accurately on high-fidelity, low noise oscilloscopes, such as the Tektronix CSA8200 equivalent-time sampling oscilloscope with 80SJNB jitter and noise analysis software and the Tektronix TDS6000 series real-time oscilloscope with TDSJIT3 jitter analysis software, that also provide a great deal of diagnostic information.

## Conclusion

The main points of this paper are:

1.  $TJ(BER)$  is an abstract form of peak-to-peak jitter that is referenced to a specific bit error ratio.
2.  $TJ(BER)$  is useful for assuring interoperability, but not so useful for diagnosing jitter problems.

3. While TJ(BER) can only be *measured* on a bit error ratio tester, TJ(BER) calculated from measurements of the jitter probability density function or the dual-Dirac model, Eq. (3), are frequently more accurate.

It is important that you also be aware of a few things that were *not* discussed in this paper. First, look at Eq. (2), the bit period,  $T_B$ , in a serial data system is much more complicated than just the reciprocal of the data rate. In jitter analysis all we care about is the jitter that can cause errors, so the clock that supplies “ $T_B$ ” in any of the analyses discussed here, can have a variety of different frequency characteristics.

We also didn’t go into any detail about the different sources of jitter, where they come from, how they are correlated with and can effect the signal-to-noise ratio, and how they need to be analyzed – we didn’t even mention that crosstalk can destroy any of these measurements. Nor did we discuss the details of the dual-Dirac model and why it’s so useful in compliance specifications. Each of these topics is the subject of another short paper.