

What the Dual-Dirac Model is and What it is Not

Ransom Stephens

October, 2006

Abstract:

The dual-Dirac model is a simple tool for estimating total jitter defined at a bit error ratio, TJ(BER), for serial data components and systems. By virtue of the fact that it can be used to combine the RJ and DJ of different elements to predict the TJ(BER) for a system, the dual-Dirac model is a key tool in specifications for serial data links. In this paper, we present the dual-Dirac model, how and why it is used in specifications, how it is used to estimate the Total Jitter of a system, the assumptions it makes and where they fail.

The dual-Dirac model was introduced several years ago as a tool for quickly estimating the total jitter defined at a bit error ratio, TJ(BER) [for the definition of TJ(BER), see Part 1 in this series, *The Meaning of Total Jitter*]. Since its introduction, the merits of the dual-Dirac model have been controversial. The controversy is almost always caused by misunderstandings in the meaning of the model parameters. The dual-Dirac model uses two parameters that, in most of the literature, are called RJ and DJ but, unless the model assumptions are strictly valid, these parameters are model dependent. To cut out that confusion, let's distinguish the model dependent parameters by calling them RJ($\delta\delta$) and DJ($\delta\delta$). As we'll see, if these model-dependent parameters are properly measured, then the relationship

$$TJ(BER) = 2Q_{BER} \times RJ(\delta\delta) + DJ(\delta\delta) \quad (1)$$

is model-*independent*. That is, TJ(BER) calculated from Eq. (1) is as accurate as the measurements of RJ($\delta\delta$) and DJ($\delta\delta$); any inaccuracy in TJ(BER) is due to the measurement, not the model. But I'm getting ahead of myself, let's go back and start at the beginning.

The dual-Dirac model

Since the following is a discussion of a mathematical model, the formalism is somewhat deep. The main reason we present it here is to show where Eq. (1) comes from and provide a rigorous foundation for the conclusions drawn in the subsequent section. If you're not intrigued by the formalism, then skim through this section, pay attention to the graphics, and you'll be well prepared for the following section where the important points of how to use the dual-Dirac model are made.

First, the Dirac-delta function, $\delta(x - x_0)$, is conveniently defined so that it is zero everywhere but at $x = x_0$, where it's infinite, but infinite in such a way that its integral is one:

$$\delta(x - x_0) \equiv \begin{cases} 0, & x \neq x_0 \\ \rightarrow \infty, & x = x_0 \end{cases} \quad \text{with} \quad \int \delta(x - x_0) dx = 1. \quad (2)$$

That is, $\delta(x - x_0)$ is a spike centered at $x = x_0$.

Second, as we know, the Probability Density Function (PDF) for RJ is a Gaussian,

$$\text{PDF}_{RJ}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{x^2}{2\sigma^2}\right]. \quad (3)$$

Third, the distributions of different components of jitter that are independent, like RJ and DJ, combine through convolution,

$$\begin{aligned} \text{PDF}(x) &= \text{PDF}_{DJ}(x) * \text{PDF}_{RJ}(x) \\ &= \int \text{PDF}_{DJ}(u) * \text{PDF}_{RJ}(x - u) du \end{aligned} \quad (4)$$

Figure 1 shows the pieces of the dual-Dirac model. The dual-Dirac DJ PDF is simply the sum of two Dirac delta functions, one centered at μ_L and one at μ_R ,

$$\text{PDF}_{\text{dual-Dirac DJ}}(x) = \delta(x - \mu_L) + \delta(x - \mu_R). \quad (5)$$

The peak-to-peak DJ of this distribution is simply the separation of the two Dirac deltas, $\mu_R - \mu_L$. Now introduce Gaussian RJ, Eq. (3) and combine it with the dual-Dirac DJ, Eq. (5), using the convolution, Eq. (4), to get the complete dual-Dirac PDF,

$$\begin{aligned} \text{PDF}_{\text{dual-Dirac}}(x) &= \text{RJ} * \text{DJ} = \frac{1}{\sqrt{2\pi}\sigma} \int [\delta(x' - \mu_L) + \delta(x' - \mu_R)] \exp\left(-\frac{(x - x')^2}{2\sigma^2}\right) dx' \\ &= \frac{1}{\sqrt{2\pi}\sigma} \left[\exp\left(-\frac{(x - \mu_L)^2}{2\sigma^2}\right) + \exp\left(-\frac{(x - \mu_R)^2}{2\sigma^2}\right) \right] \end{aligned} \quad (6)$$

two Gaussians, each displaced from the origin.

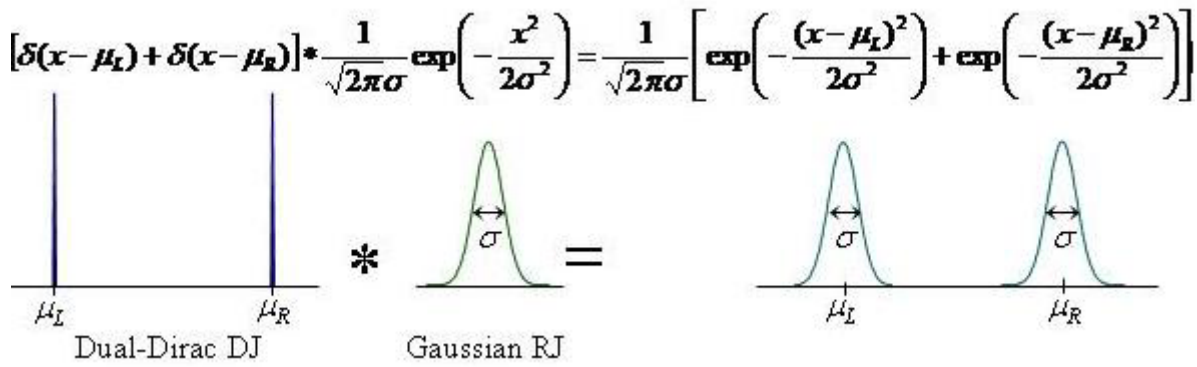


Figure 1: The dual-Dirac jitter distribution. In (a) the DJ and RJ distributions and, in (b), their convolution.

Another way to think of the dual-Dirac model that may be more intuitive is to see how it evolves in an eye diagram. Start with no jitter in Figure 2a and introduce only the dual-Dirac DJ in Figure 2b. Notice that there are two distinct logic-transition trajectories, one has its crossing point at μ_L and the other at μ_R . When RJ is introduced in Figure 2c, the two trajectories are smeared according to the Gaussian. The dual-Dirac jitter distribution is shown by the histogram in the upper left corner of Figure 2c. In practice, the dual-Dirac DJ distribution can be realized by square-wave phase modulation – hardly a realistic scenario.

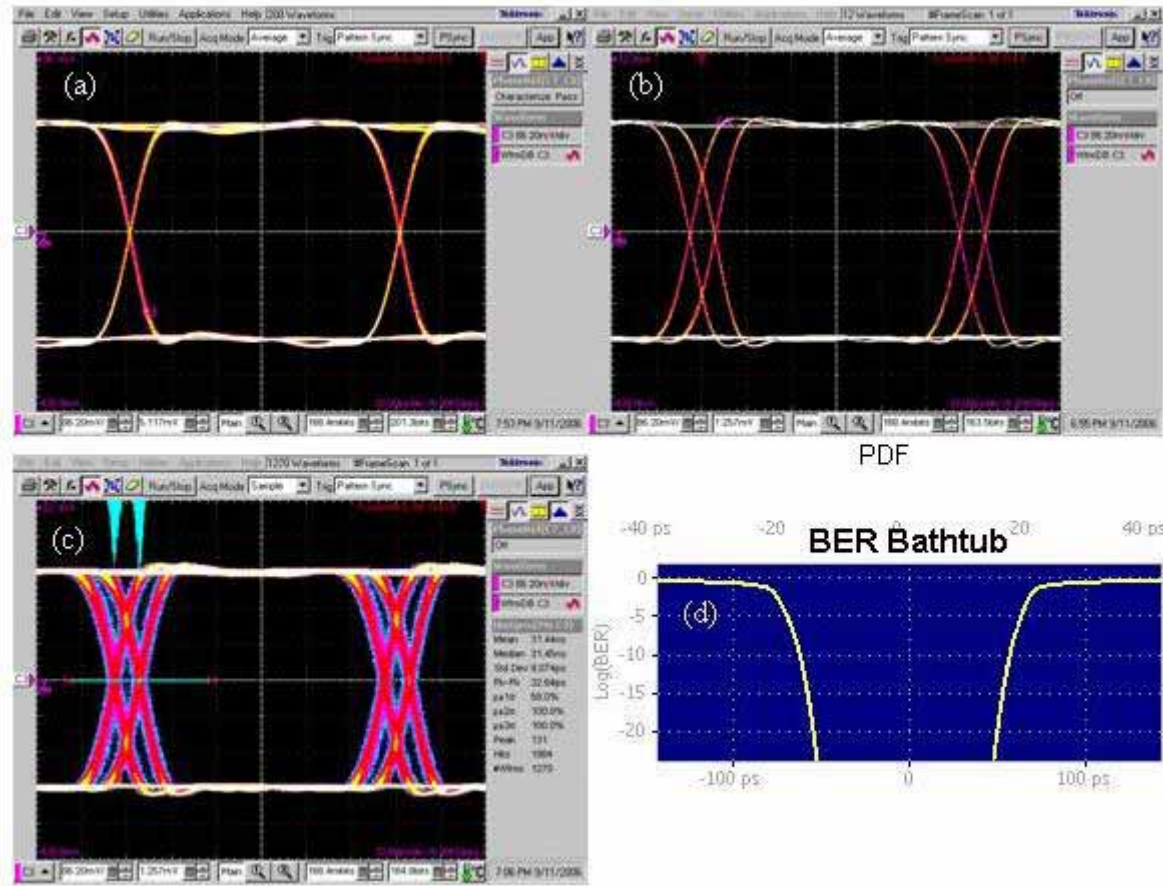


Figure 2: An eye diagram with, (a) no jitter, (b) dual-Dirac DJ, (c) RJ and dual-Dirac DJ, and (d) bathtub plot, BER(x).

Once we have the jitter PDF, the bathtub plot shown in Figure 2d can be calculated. BER(x) is the probability that an error will occur if the sampling point is positioned at the time-delay position, x . In general, BER(x) is given by

$$\text{BER}(x) = \rho_T \int_x^{\infty} \text{PDF}(x') dx' + \rho_T \int_{-\infty}^x \text{PDF}(x'-T) dx' \quad (7)$$

where ρ_T is the logic transition density (i.e., the ratio of the number of transitions to the number of bits). The first term on the right of Eq. (7) accounts for fluctuations across the sampling point from left to right and the second term accounts for fluctuations across the sampling point from right to left.

Plug the dual-Dirac model, Eq. (6) into Eq. (7), and get

$$\text{BER}_{\delta\delta}(x) = \rho_T \left[\text{erfc} \left(\frac{x - \mu_L}{\sqrt{2}\sigma} \right) + \text{erfc} \left(\frac{(x - T) - \mu_R}{\sqrt{2}\sigma} \right) \right] \quad (8)$$

where $\text{erfc}(x)$ is the “complementary error function.” Evaluate the complementary error functions, do a bunch of algebra, and get Eq. (1),

$$\text{TJ}(\text{BER}) = 2Q_{\text{BER}} \times \text{RJ}(\delta\delta) + \text{DJ}(\delta\delta) \quad (1)$$

where Q_{BER} is a constant that can be calculated from the complementary error function and is tabulated for a few values of BER in Table 1. RJ is given by σ and DJ is given by the separation of the two Dirac-delta functions, $\mu_R - \mu_L$. That is,

$$\text{RJ}(\delta\delta) = \sigma \text{ and } \text{DJ}(\delta\delta) = \mu_R - \mu_L. \quad (9)$$

Equation (1) with Eq. (9) is a nice result, but what good does it do with a real-life jitter distribution like the one in Figure 3?

BER	Q_{BER}
10^{-10}	6.3
10^{-11}	6.7
10^{-12}	7.0
10^{-13}	7.4
10^{-14}	7.7

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

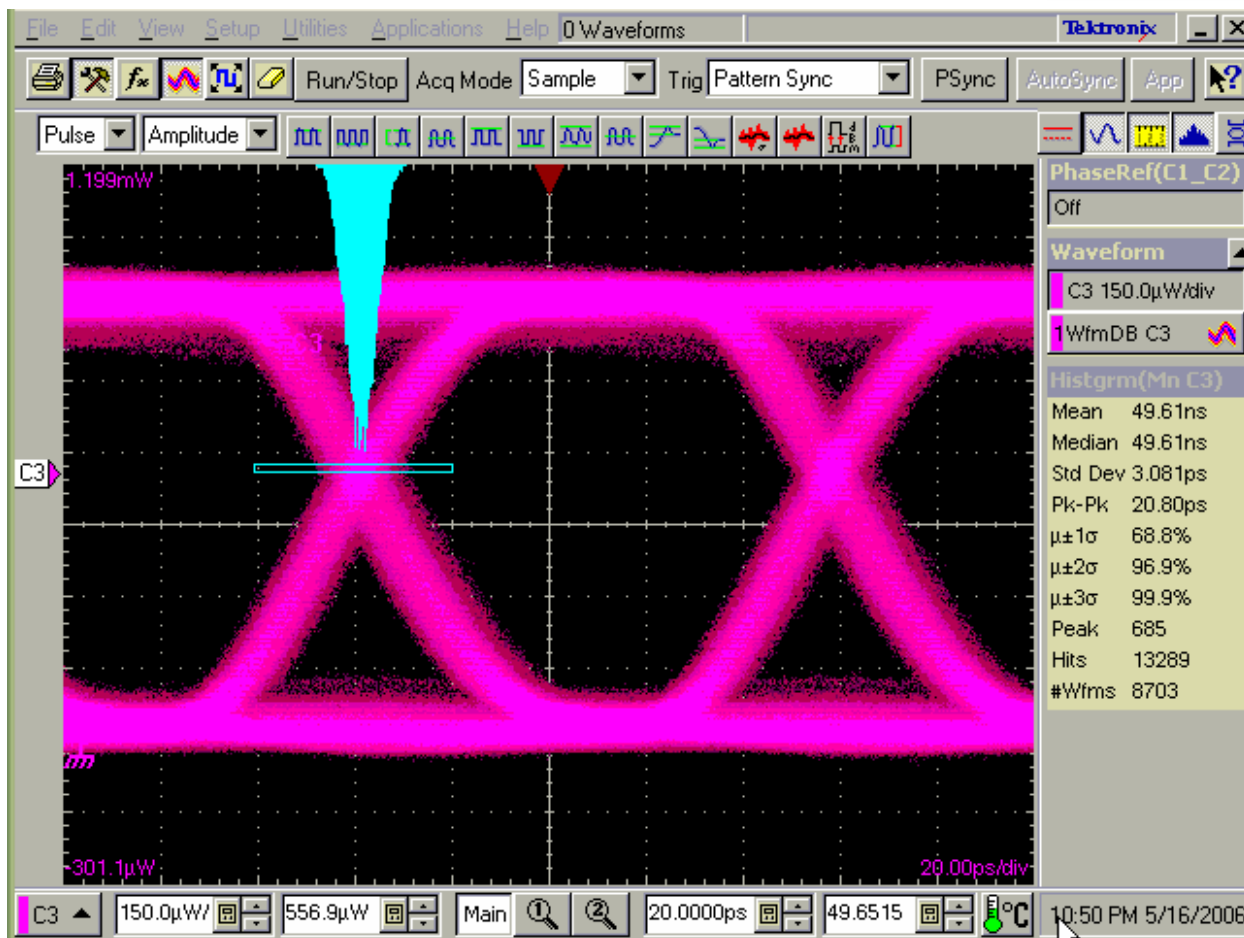


Figure 3: A more realistic eye-diagram.

What the dual-Dirac model is and what it is not

Since the calculation of TJ(BER) depends only on the tails of the jitter distribution, i.e., from “ x to ∞ ” and “ $-\infty$ to x ” as in Eq. (7), we can calculate TJ(BER) using Eq. (1) for *any* jitter distribution as long as:

- (1) we can neglect amplitude noise, and
- (2) the tails of the jitter distribution are dominated by RJ.

These are the two key assumptions of the dual-Dirac model.

Here’s another way of saying it:

As long as the tails of the distribution follow the RJ Gaussian at the BER we care about, TJ(BER) is given by Eq. (1).

The dual-Dirac model is just a convenient way of lining the Gaussian RJ distribution up with the data. That's all it is.

The trick to using the dual-Dirac model, is remembering that its two parameters, RJ($\delta\delta$) and DJ($\delta\delta$), are in general model-dependent and have to be measured from the data accordingly. Fortunately, RJ($\delta\delta$) and DJ($\delta\delta$) can be measured in many different ways for any jitter distribution for which the two assumptions above are valid.

Let's start with RJ($\delta\delta$) because it's much less model-dependent than DJ($\delta\delta$). As long as the rising and falling slew rates are the same, the dual-Dirac RJ is exactly the model-independent RJ given by the width, or rms, of the RJ Gaussian, σ . On the other hand, if the edges are not symmetric, then RJ($\delta\delta$) is defined as the average of the Gaussian tails of the left and right widths:

$$\text{Symmetric edges, RJ is model-independent: } RJ(\delta\delta) = RJ \equiv \sigma \quad (10)$$

$$\text{Asymmetric edges, RJ is model-dependent: } RJ(\delta\delta) \equiv \frac{1}{2} (\sigma_L + \sigma_R) \quad (11)$$

In the vast majority of systems the difference between Eq. (10) and (11) is smaller than the uncertainty in the measurements of σ_L and σ_R and we can safely assume $RJ(\delta\delta) = RJ = \sigma$.

By its nature the RJ of one element is statistically independent of the RJ of another element. That is, the RJ of one element doesn't interfere with the RJ of another. Independence means that the RJ distribution of two network elements is given by the convolution of the two RJ distributions. The convolution of two Gaussians is a Gaussian whose width is the square-root of the sum of the squares of the individual widths. Thus, the RJ for a system of N elements, is

$$\sigma_{system} = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_N^2} . \quad (12)$$

DJ($\delta\delta$) is more complicated. Where the model dependence of RJ is almost always negligible, the model dependence of DJ($\delta\delta$) is rarely negligible.

The only case where the actual peak-to-peak DJ, lets call it DJ(p-p), is the same as the dual-Dirac DJ, DJ($\delta\delta$), is when the DJ distribution really is given by the sum of two Dirac-delta functions, for example, when the phase is modulated by a square-wave. Figure 4 shows how DJ(p-p) differs from DJ($\delta\delta$). By their bounded nature, DJ distributions tend to have sharp edges and, by its Gaussian nature, RJ has long smooth tails. The process of convolving DJ with Gaussian RJ is a process of smoothing the edges of the DJ distribution. As the edges of the DJ distribution are sanded down, they appear to be pulled inward, closer to the nominal crossing point. The result is a general statement about the model-dependence of DJ,

$$DJ(\delta\delta) \leq DJ(p-p). \quad (13)$$

The result, Eq. (13), is the reason that the dual-Dirac model suffers so much unwarranted criticism. It's a model. It's okay for a model to have model-dependent parameters as long as the model-dependence is included when they're measured. In other words, to use Eq. (1), we need to measure DJ($\delta\delta$), not DJ(p-p) – which in some ways is good news. It's easier to measure DJ($\delta\delta$) than it is to measure DJ(p-p). An accurate measurement of DJ(p-p) requires a complete understanding of *all* sources of DJ, including their relative phases, which is difficult to measure accurately on most equipment. Accurate measurement of DJ($\delta\delta$) only requires that the data sample is large enough to have sampled all DJ processes.

The main point is that, from the compliance point of view, DJ($\delta\delta$) is more useful than DJ(p-p). The DJ($\delta\delta$) of different components can be combined to estimate the DJ($\delta\delta$) of a system, but the DJ(p-p) cannot.

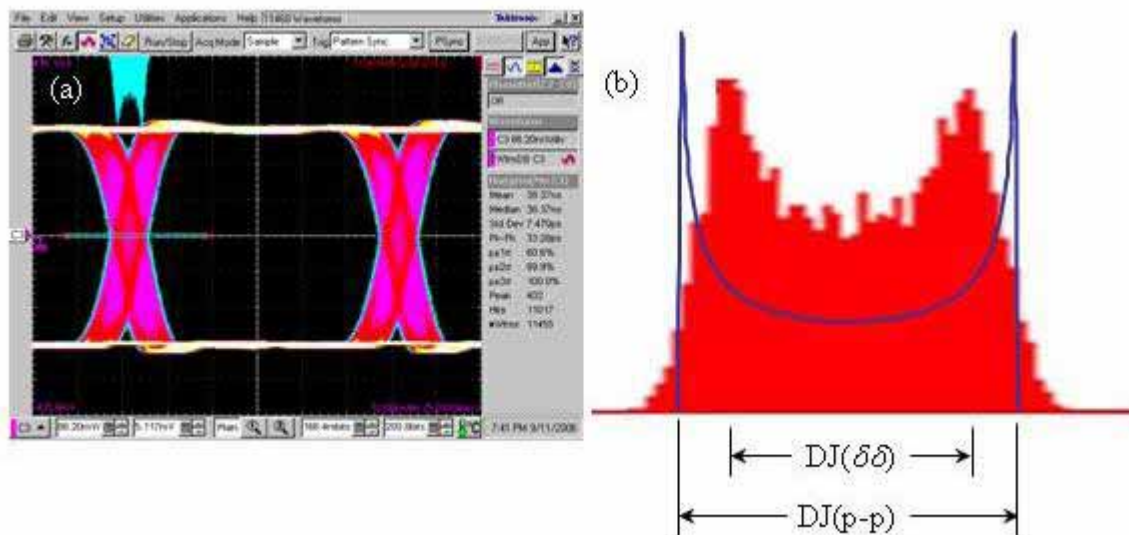


Figure 4: (a) An eye diagram and jitter histogram for a signal with sinusoidal DJ and RJ, (b), an exploded view of the jitter histogram with the underlying sinusoidal DJ PDF overlaid. Notice how the sharp edges of the DJ distribution are pulled inward when RJ is introduced. In both cases, the actual peak-to-peak DJ, $DJ(p-p)$, and the model-dependent dual-Dirac DJ, $DJ(\delta\delta)$, are shown, demonstrating Eq. (13).

If the PDF of one jitter source changes when the PDF of another source is changed, then those two sources are *dependent* or *correlated*. The peak-to-peak value of the convolution of two *independent* distributions is the sum of the peak-to-peak values of each distribution. On the other hand, the peak-to-peak value of the convolution of two *correlated* distributions is less than or equal to the sum of the peak-to-peak values of each distribution. Generalizing to a system of N components, we can say

$$DJ_{system}(p-p) \leq DJ_1(p-p) + DJ_2(p-p) + \dots + DJ_N(p-p). \quad (14)$$

The problem is that the DJ of different system elements is almost always correlated. For example, the ISI of a transmitted signal affects the ISI introduced by the transmission channel. If we used the sum of the component's $DJ(p-p)$ to estimate the DJ of the system we'd have two problems. First, we'd always overestimate the actual $DJ(p-p)$ and, second, $DJ(p-p)$ isn't the right parameter to use in Eq. (1) for estimating $TJ(BER)$, we need $DJ(\delta\delta)$ for that.

We need a simple expression for combining the DJ($\delta\delta$) of different components for use, along with the combination of component RJ of Eq. (12), in Eq. (1). The obvious candidate is

$$DJ_{system}(\delta\delta) \approx DJ_1(\delta\delta) + DJ_2(\delta\delta) + \dots + DJ_N(\delta\delta). \quad (15)$$

Equation (15) is an approximation, and should be treated with appropriate suspicion, though it's usually a pretty good approximation. Think of it like this, as more DJ sources are introduced, the resulting DJ distribution gets smoother around the edges. The smoother the distribution, the greater the discrepancy between DJ($\delta\delta$) and DJ(p-p). Now, since the maximum possible DJ(p-p) is given by the sum of the component DJ(p-p) values, Eq. (14), and since DJ($\delta\delta$) is generally smaller than DJ(p-p), it's reasonable to expect that $DJ_{system}(\delta\delta)$ should be smaller than that given by the sum, Eq. (15). We might then expect that Eq. (15) is a conservative estimate for the system DJ. That's the argument, anyway.

It's possible to get an idea of the accuracy of Eq. (15) by analyzing the different DJ components, a diagnostic technique that's the subject of another paper in this series, "All About the Acronyms: RJ, DJ, DDJ, ISI, DCD, PJ, SJ,..."

Equations (12) and (15) along with (1), rewritten here,

$$\sigma_{system} = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_N^2}. \quad (12)$$

$$DJ_{system}(\delta\delta) \approx DJ_1(\delta\delta) + DJ_2(\delta\delta) + \dots + DJ_N(\delta\delta). \quad (15)$$

$$TJ(BER) = 2Q_{BER} \times \sigma + DJ(\delta\delta) \quad (1)$$

provide the tools we need to estimate the TJ(BER) of a system from the RJ and DJ($\delta\delta$) of its components. It is for this reason that when DJ is quoted in the standards specification for a given technology like FibreChannel, PCI-Express, FBD, SATA, et cetera, it is DJ($\delta\delta$), not DJ(p-p) that is relevant.

Conclusion

The dual-Dirac model is useful both for estimating TJ(BER) and combining the RJ and DJ of separate components to estimate the TJ(BER) of a system – provided the model dependent parameters, DJ($\delta\delta$) and RJ($\delta\delta$) are used. The difference between the model-dependent RJ,

RJ($\delta\delta$), and the true RJ, σ , is almost always negligible and it's perfectly safe to assume RJ = σ . On the other hand, the model-dependent DJ, DJ($\delta\delta$), is different than the true peak-to-peak DJ, DJ(p-p). At least in the context of standards DJ($\delta\delta$) is more useful than DJ(p-p).

But it's still a model. It still rests on assumptions that can be debated. For example, most techniques for measuring RJ run into serious problems in environments that include crosstalk and other forms of "bounded uncorrelated jitter." Another potential problem is that if RJ doesn't follow a Gaussian distribution then the very foundation of the model collapses or if the DJ distribution mimics the RJ distribution then many measurement techniques will overestimate RJ. In fact, a large number of small deterministic effects can result in a distribution that is indistinguishable from a Gaussian down to a very low BER. These are called "truncated Gaussians." If the Gaussian is truncated at a BER of 10^{-8} and a data sample of 10^7 logic transitions is used to measure it, then the RJ is upwardly biased and Eq. (1) will yield a higher TJ(BER) than is the truth. Amplitude noise can also seriously complicate the situation; jitter analysis only considers one dimension of a two-dimensional problem.