

Jitter Analysis: The dual-Dirac Model, RJ/DJ, and Q-Scale

31-December-2004

The dual-Dirac model is a tool for quickly estimating total jitter defined at a low bit error ratio, TJ(BER). The deterministic and random subcomponents of the jitter signal are separated within the context of the model to yield two quantities, root-mean-square random jitter (RJ) and a model-dependent form of the peak-to-peak deterministic jitter, DJ($\delta\delta$). The total jitter of a system is then estimated from RJ and DJ($\delta\delta$).

This paper provides a complete description of the dual-Dirac model, how it is used in technology standards and a summary of how it is applied on different types of test equipment. The Q-scale formulation is described in detail and is used to provide a simple visual description of the model's features and to show how different implementations of the model can lead to different results. For an introduction to jitter analysis please see reference¹.

Section one is a short self-contained summary. The succeeding sections provide the complete reference material for understanding the summary. This format provides both the story-in-a-nutshell and the details-in-full so that you can, hopefully, access the information you need as quickly as possible – if you have comments or suggestions, accolades or criticisms, please send me a note:

ransom_stephens@keysight.com.

¹ To learn about the basics of jitter analysis, see Ransom Stephens, "The rules of jitter analysis," <http://www.analogzone.com/nett0927.pdf>, 2004; Ransom Stephens, "Analyzing Jitter at High Data Rates," IEEE Communications Magazine, vol. 42, no. 2, February 2004, pp. 6-10; Keysight Technologies' Application Note, "Jitter Fundamentals: Keysight 81250 ParBERT Jitter Injection and Analysis Capabilities," Literature Number AN-5988-9756EN, available from www.Keysight.com, 2003.

1. What The dual-Dirac Approximation is and What it is Not

The dual-Dirac model is universally accepted for its utility in quickly estimating total jitter defined at a bit error ratio, TJ(BER), and for providing a mechanism for combining TJ(BER) from different network elements. It relies on the five assumptions given in Table 1.

Table 1. The dual-Dirac model assumptions.

1	Jitter can be separated into two categories, random jitter (RJ) and deterministic jitter (DJ). Jitter can be separated into two categories, random jitter (RJ) and deterministic jitter (DJ).
2	RJ follows a Gaussian distribution and can be fully described in terms of a single relevant parameter, the rms value of the RJ distribution or, equivalently, the width of the Gaussian distribution, σ .
3	DJ follows a finite, bounded distribution.
4	DJ follows a distribution formed by two Dirac-delta functions. The time-delay separation of the two delta functions gives the dual-Dirac model-dependent DJ, as shown in Figure 1.
5	Jitter is a stationary phenomenon. That is, a measurement of the jitter on a given system taken over an appropriate time interval will give the same result regardless of when that time interval is initiated.

A measurement of TJ(BER) at low BER can only be performed on a Bit Error Ratio Tester (BERT) and takes as long as is necessary to transmit about $10/\text{BER}$ bits¹ (e.g., to measure TJ(10^{-12}) would take about 70 minutes). The first two assumptions in Table 1, that jitter is caused by a combination of random and deterministic processes whose distributions can be separated and that RJ follows a Gaussian, are standard industry-wide assumptions; they are the key to how the dual-Dirac model is used to estimate TJ(BER) from measurements of comparatively low statistics.

The components of jitter combine through convolution. We can think of the jitter distribution, e.g., the histogram of the crossing-point of an eye diagram, as having three regions: at the crossing-point the distribution is dominated by DJ, at time-delays farther from the crossing-point the distribution is increasingly dominated by RJ until, far from the crossing point, in the asymptotic limit, the tails follow the Gaussian RJ distribution. The asymptotic tails of the distribution usually cause errors at the level of $\text{BER} < 10^{-8}$. The ideal way to determine the behavior of the tails, and hence, TJ(BER) would be to deconvolve RJ and DJ. But without knowing the DJ distribution beforehand, there is no practical way to deconvolve the distribution.

¹ Marcus Müller, Ransom Stephens, and Russ McHugh, "Total Jitter Measurement at Low Probability Levels, Using Optimized BERT Scan Method," DesignCon 2005, <http://www.designcon.com/conference/7-ta4.html>.

The dual-Dirac model, Figure 1, provides the simplest possible distribution: the crossing-point is separated into two Dirac-delta functions positioned at μ_L and μ_R , the DJ dominated region, followed by an artificially abrupt transition to the RJ dominated tails. There are many ways to implement the dual-Dirac model, in all of them estimating TJ(BER) is a matter of describing the tails of the jitter distribution with the tails of two Gaussians of width σ separated by a fixed amount $DJ(\delta\delta) \equiv |\mu_L - \mu_R|$:

The dual-Dirac model is a Gaussian approximation to the outer edges of the jitter distribution displaced by $DJ(\delta\delta)$.

Conceptually, think of DJ closing the eye a fixed amount, $DJ(\delta\delta)$, and the Gaussian RJ tails closing the eye an amount that depends on the bit error ratio of interest. Once σ and $DJ(\delta\delta)$ are measured the eye closure at any BER can be estimated with:

$$TJ(BER) \equiv 2QBER\sigma + DJ(\delta\delta) \quad (1)$$

where Q_{BER} is calculated from the complementary error function, as given in Table 2. Since σ is multiplied by $2 Q_{BER}$, the accuracy of TJ(BER) depends first on the accuracy of the RJ measurement, σ , and second on the accuracy of the $DJ(\delta\delta)$ measurement.

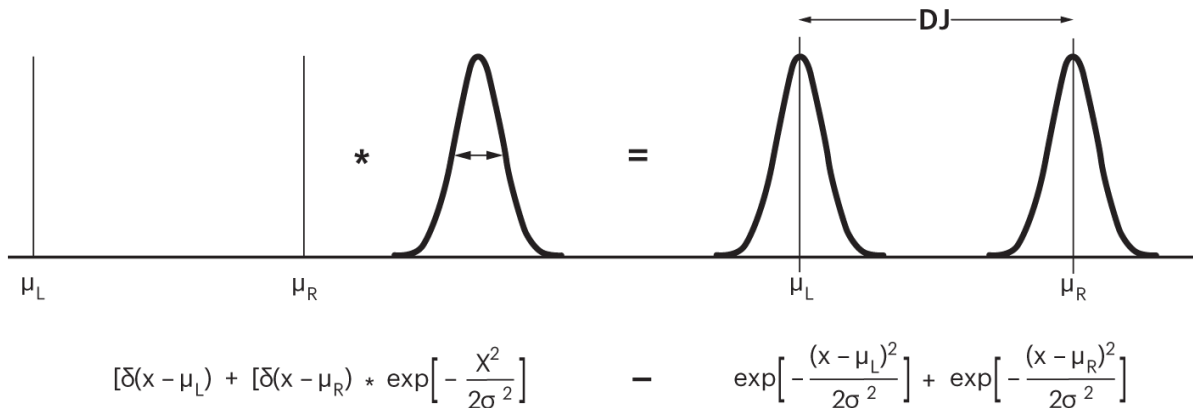


Figure 1. The convolution of the sum of two delta functions separated by DJ and a Gaussian RJ distribution of width σ . The underlying assumption of the dual-Dirac approximation is that any jitter distribution can be modeled this way.

Q _{BER}	BER
6.4	10 ⁻¹⁰
6.7	10 ⁻¹¹
7.0	10 ⁻¹²
7.3	10 ⁻¹³
7.6	10 ⁻¹⁴

Table 2. Values of Q_{BER}, the multiplicative constant for determining eye closure due to RJ, for different BER values.

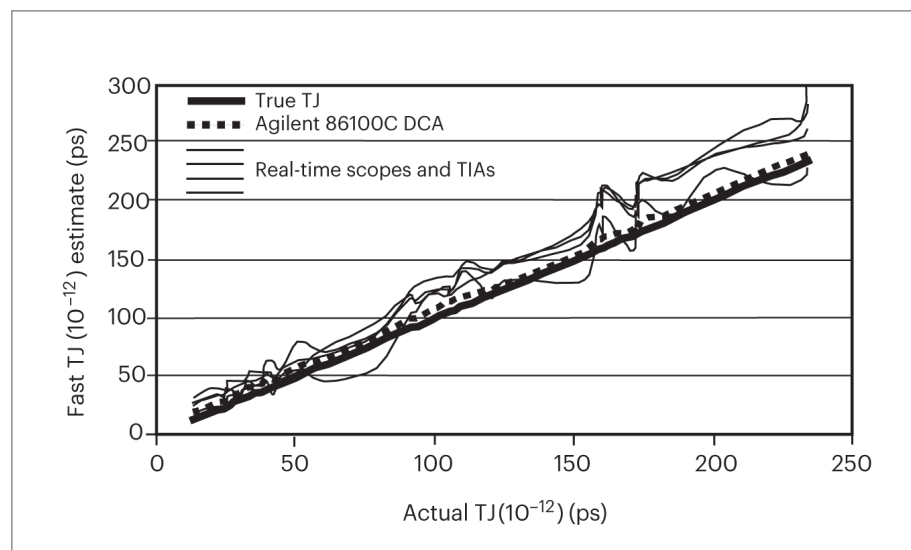


Figure 2. Comparison of different jitter analysis techniques in estimating TJ(10⁻¹²) for a variety of different combinations of random, periodic, and data-dependent jitter. The heavy solid line indicates the actual TJ(10⁻¹²) and the other lines give results from different test sets. The DCA-J (heavy dotted line) gives the most accurate results. For the details of the analysis, see ref.¹.

Different implementations of the dual-Dirac model can give widely varying results in σ and DJ($\delta\delta$) and hence TJ(BER). There are two different categories types of dual-Dirac implementations, those that determine σ by fitting some type of Gaussian expression (e.g., the complementary error function) to a jitter distribution and those that measure the timing noise spectrum and equate it to σ . In both cases problems can occur when DJ is mistaken for RJ. Keysight Technologies, Inc. solved this problem when we introduced the 86100C DCA-J. The DCA-J measures the timing noise that is independent of the bounded deterministic jitter to obtain an accurate value for σ . As shown in Figure 2 [details in Ref.¹], extensive tests of all the available techniques on a precision jitter transmitter [Ref.²] show that the DCA-J technique (described in Ref.³) is the most accurate, sensitive, and repeatable jitter analysis tool available.

¹ Ransom Stephens et al., "Comparison of jitter analysis techniques with a precision jitter transmitter," Keysight Technologies Whitepaper, 2005. Available from www.Keysight.com.

² Jim Stimple and Ransom Stephens, "Precision Jitter Transmitter," DesignCon 2005, <http://www.designcon.com/conference/7-wa1.html>.

³ Keysight Technologies Product Note, 86100C-1, "Precision jitter analysis using the Keysight 86100C DCA-J," Keysight Literature Number 5989-1146EN, <http://cp.literature.keysight.com/litweb/pdf/5989-1146EN.pdf>.

Most of the confusion surrounding the application of the dual-Dirac model is caused by the fact that the dual-Dirac DJ is a completely different quantity than the peak-to-peak DJ. To distinguish the two, we'll use the notation DJ(p-p) and DJ($\delta\delta$). The distinction is easy to understand: real DJ never follows the simple dual-Dirac distribution and so it is unreasonable to expect the DJ extracted from the dual-Dirac model to approximate the actual peak-to-peak DJ. DJ($\delta\delta$) is a model dependent quantity that must be derived under the assumption that DJ follows a distribution formed by two Dirac-delta functions, as shown in Figure 1. Generally,

$$DJ(\delta\delta) < DJ(p-p) \quad (2)$$

While DJ($\delta\delta$) can always be measured, DJ(p-p) can only be measured in special cases. For example, when DJ is composed exclusively of data-dependent jitter (DDJ) it can be measured by comparing the average transition times of a repeating data pattern [Ref. ¹]. The inequality, Eq. (2), can be confusing when DJ($\delta\delta$) is mistaken as an estimate of DJ(p-p); for example DJ($\delta\delta$) may be smaller than a subcomponent of DJ, but DJ(p-p) cannot be smaller than one of its subcomponents.

The inequality, Eq. (2), is a consequence of the dual-Dirac model but does not detract from the model's utility. On the contrary, DJ($\delta\delta$) meets the two necessary requirements of a quantity of interest for a standards body:

DJ($\delta\delta$) is both well defined and observable

DJ($\delta\delta$) can also be measured on a variety of different types of test equipment. The true peak-to-peak DJ, DJ(p-p), on the other hand can only be measured in special circumstances, is not useful for estimating TJ(BER), and doesn't provide any benefit in diagnosing problems.

When well-considered standards refer to DJ, it is DJ($\delta\delta$) to which they refer, either explicitly or implicitly.

Once values for DJ($\delta\delta$) and σ are obtained for a given network element, the values for combinations of elements can be estimated with these rules:

$$\sigma_{\text{Total}} = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2} \quad (3)$$

$$DJ_{\text{Total}}(\delta\delta) \approx DJ_1(\delta\delta) + DJ_2(\delta\delta) + \dots + DJ_N(\delta\delta)$$

The simple rules, Eqs. (1) and (3), provide standards committees an easy way to divide a BER budget among a systems' components. Notice in Eq. (3) that σ_{Total} is equal to the sum of the squares but DJ($\delta\delta$)_{Total} is an approximation to the sum of DJ($\delta\delta$)_i, the distinction is explained in Section 3.3.

The full story is given in the rest of this note. We start in section 2, with some background material defining bit error ratio and TJ(BER). Section 3 gives a precise description of the dual-Dirac model, including the Q-scale representation, and how the model provides a simple technique for combining the TJ(BER) of different network elements to estimate TJ(BER) of a system, Eq. (3). In Section 4 we discuss different implementations of the model, identify common pitfalls and show, for a specific system, the relationship between DJ($\delta\delta$) and DJ(p-p) and how accurately σ and TJ(BER) are estimated in a naïve dual-Dirac implementation. In Section 5, we conclude with a discussion of the veracity of the standard industry wide assumptions and a bulleted summary.

¹ Keysight Technologies Product Note, 86100C-1, "Precision jitter analysis using the Keysight 86100C DCA-J," Keysight Literature Number 5989-1146EN, <http://cp.literature.keysight.com/litweb/pdf/5989-1146EN.pdf>.

2. Total Jitter Defined at a Bit Error Ratio

Uncorrelated jitter distributions combine through convolution. The industry-wide assumption is that random jitter (RJ) follows a Gaussian distribution that is independent of all sources of DJ. The probability density function, or jitter distribution, can be measured by making a histogram of an eye-diagram crossing point, Figure 3a. The jitter distribution, $J(x)$ Figure 3b, can then be described as $RJ(x)_{DJ}(x)$ where the functional form of $DJ(x)$ is not generally known or even observable, but $RJ(x)$ is given by a Gaussian distribution whose width is σ and x is the time-delay, or horizontal axis of the eye diagram. If we write the jitter distribution as

$$J(x) = RJ(x) * DJ(x) \quad (4)$$

With

$$RJ(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{x^2}{2\sigma^2}\right] \quad (5)$$

$$\text{then } J(x) = \frac{1}{\sqrt{2\pi}\sigma} \int DJ(x') \exp\left[-\frac{(x-x')^2}{2\sigma^2}\right] dx' \quad (6)$$

It's not hard to show that for bounded $DJ(x)$ the asymptotic behavior of $J(x)$ is the same as that of a Gaussian,

$$\lim_{x \rightarrow -\infty} J(x) \rightarrow A \exp\left[-\frac{(x-\xi_L)^2}{2\sigma^2}\right] \quad (7)$$

$$\lim_{x \rightarrow \infty} J(x) \rightarrow A \exp\left[-\frac{(x-\xi_R)^2}{2\sigma^2}\right]$$

where ξ_L and ξ_R are constants that depend on $DJ(x)$. In the dual-Dirac model $\xi_L = \mu_L$ and $\xi_R = \mu_R$ as in Figure 1.

The asymptotic behavior of the jitter distribution, Eq. (7), is the reason that the dual Dirac model can be used to represent a jitter distribution like that shown in Figure 3b.

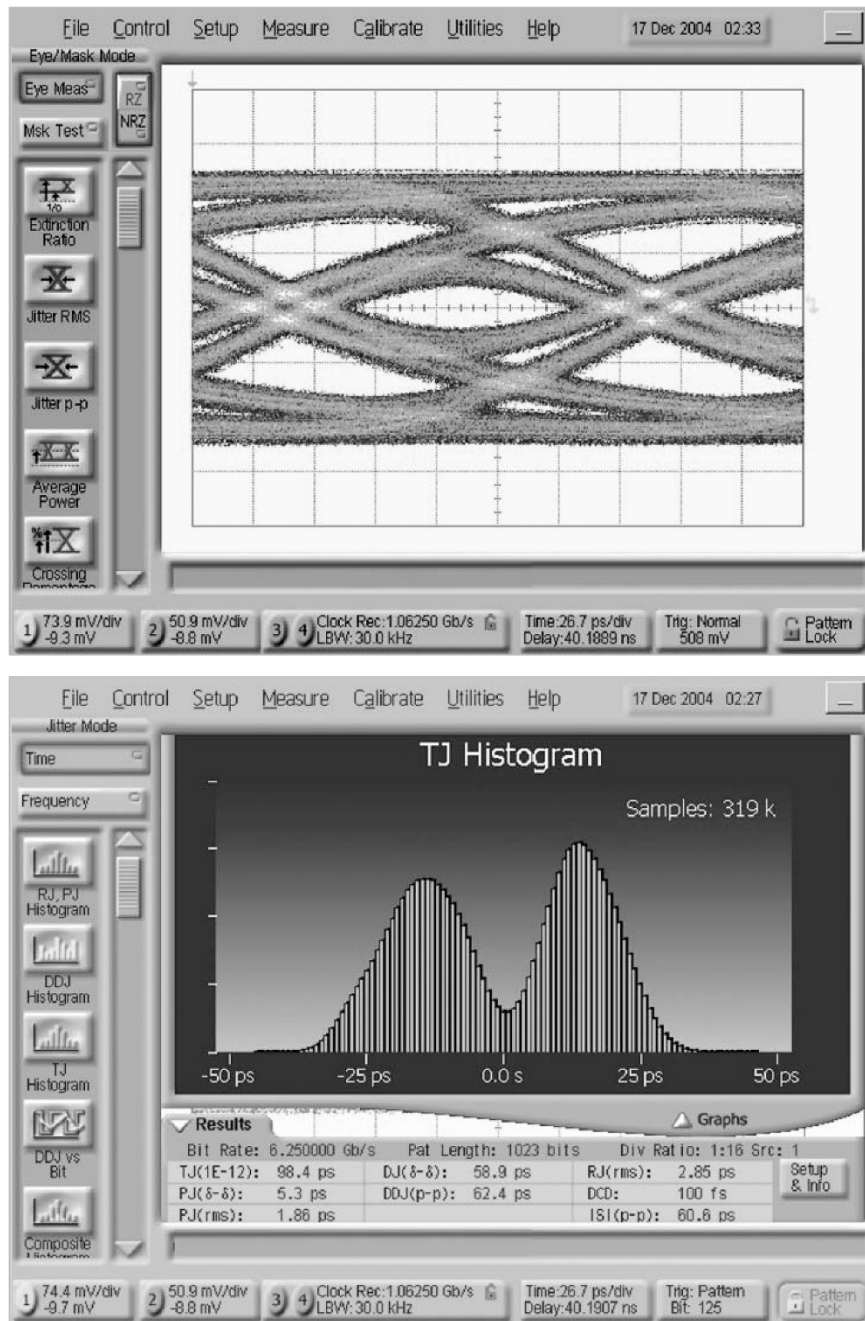


Figure 3. (a) An eye diagram measured on a Keysight 86100C DCA-J and, (b), the jitter distribution, $J(x)$.

The value of the dual-Dirac model is in its ability to provide a technique to quickly estimate total jitter at a given bit error ratio, TJ(BER). Since TJ(BER) is defined in terms of the bit error we need to back up a bit and define a few other things to make sense of it. The BER is defined as

$$\text{BER}(x, V) \equiv \lim_{N \rightarrow \infty} \frac{N_{\text{err}}(x, V)}{N} \quad (8)$$

where (x, V) is the position of the sampling point and $N_{\text{err}}(x, V)$ is the number of errors that would be detected from a total of N transmitted bits. An error is detected, for example, on a logic '0' bit if the observed potential (or, for differential signals, the difference of the two potentials on the differential lines) is greater than V at the sampling time-delay x . In this document, we assume that all errors result from timing errors alone – that is, from jitter – which means that we can measure the dependence of the BER on the time-delay, x , without being concerned about amplitude noise.

The dependence of BER on x is the heart of how TJ(BER) is defined. By scanning the time-delay position of the sampling point across the eye a bathtub plot, $\text{BER}(x)$, as shown in Figure 4, can be measured on a bit error ratio tester. $\text{BER}(x)$, can also be derived from the jitter distribution, $J(x)$. Since $\text{BER}(x)$ is given by the probability for a logic transition fluctuating across the sampling point time position, x , if we consider the left edge of Figure 3a then the probability of a transition fluctuating across the point x is given by

$$\text{BER}_L(x) = \rho_T \int_x^\infty J(x') dx' \quad (9)$$

where ρ_T is the ratio of the number of logic transitions to the total number of bits, the transition density. Similarly, on the right side of the eye diagram, near $x = T_B$,

$$\text{BER}_R(x) = \rho_T \int_{-\infty}^x J(x') dx' \quad (10)$$

so that

$$\text{BER}(x) = \text{BER}_L(x) + \text{BER}_R(x)$$

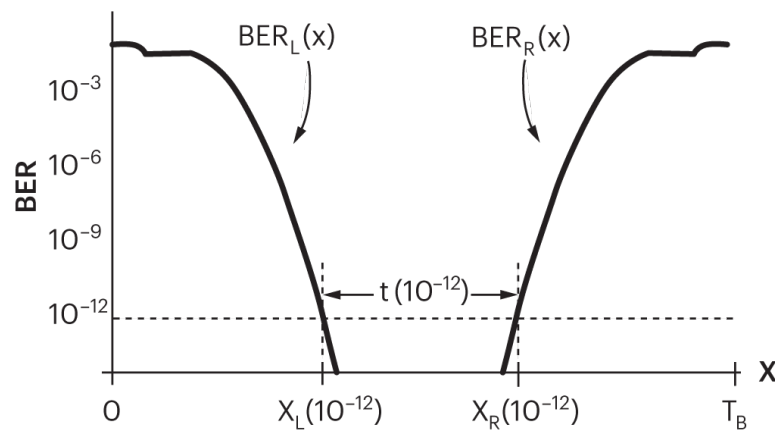


Figure 4. A bathtub plot or BERTscan; the bit error ratio as a function of sampling point delay, x .

The eye opening at a given BER, $t(\text{BER})$, is given by the separation of the left and right BER curves at a given BER. For example, in Figure 4, the eye opening at $\text{BER} = 10^{-12}$, is given by the difference of x_L and x_R – those points where $\text{BER} = 10^{-12}$. Formally, if we invert $\text{BER}(x)$ to get $x(\text{BER})$, then

$$t(\text{BER}) = x_R(\text{BER}) - x_L(\text{BER}) \quad (11)$$

$TJ(\text{BER})$ is defined as the amount of eye closure due to jitter at a given BER; that is, $TJ(\text{BER})$ is the difference in the bit period and the eye opening:

$$TJ(\text{BER}) \equiv T_B - t(\text{BER}) \quad (12)$$

3. The dual-Dirac Model in Action

The fundamental assumption behind the dual-Dirac approximation is that any deterministic jitter distribution can be approximated by two delta functions separated by DJ, as shown in Figure 1. In this special case, $DJ(\delta\delta) = DJ(p-p)$. A jitter distribution that closely follows a dual-Dirac could result from pure duty cycle distortion (DCD) or square wave phase modulation but in practical applications it is much more complicated.

Separate jitter components combine through convolution – the mathematical process of folding one distribution over another, Eq. (6) – it is important to understand the process of convolution. To illustrate how RJ and DJ convolve, think of the smooth Gaussian RJ as a smearing function imposed on the bounded DJ distribution. In Figure 5a the dashed DJ distribution is given by a single frequency of sinusoidal jitter and, in Figure 5b, by a flat, bounded DJ distribution. The DJ distributions are convolved with a Gaussian resulting in the smooth solid curves. The effect of the convolution is to smooth the sharp edges of the DJ distribution and $J(x)$ obtains its Gaussian tails. Notice how the smoothing effect of the convolution brings the sharp DJ edges inward. The vertical lines in Figure 5 are set at μ_R and μ_L demonstrating the inequality, $DJ(\delta\delta) = |\mu_R - \mu_L| < DJ(p-p)$. The two Gaussian curves (dash-dot) give the dual-Dirac approximation to the solid curve. It doesn't matter that the central part of the dual-Dirac distribution doesn't match the actual distribution; the important feature is that the Gaussian tails match the tails of the true jitter distribution as in Eq. (7) so that $TJ(\text{BER})$ can be estimated using Eqs. (9) through (12).

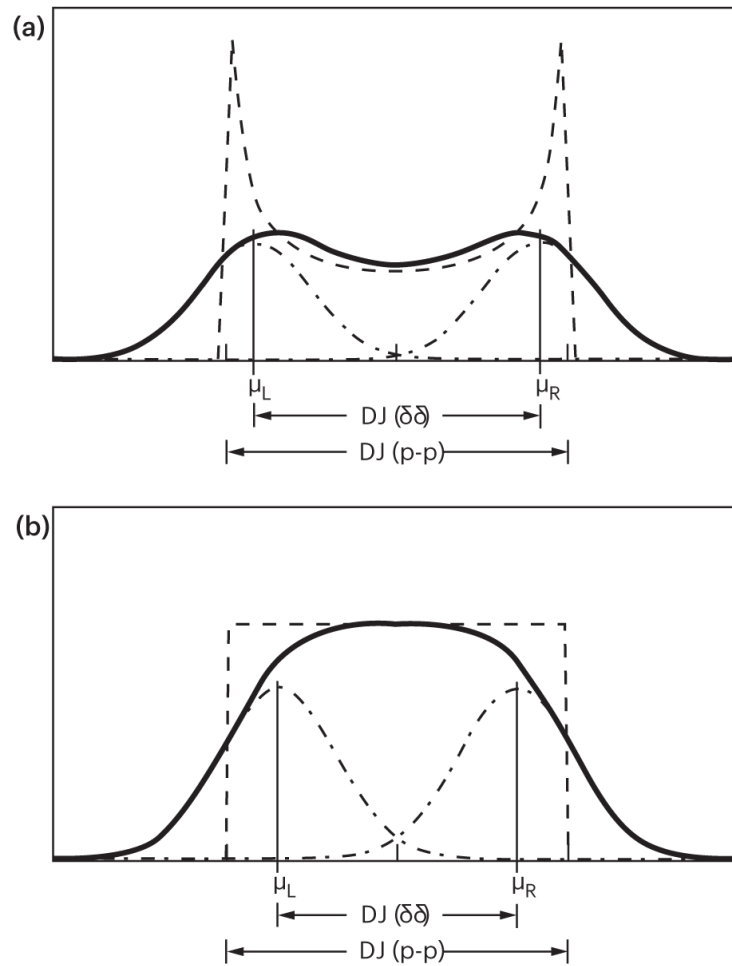


Figure 5. Application of the dual-Dirac model to two ideal cases. In (a) the dashed DJ distribution is caused by a single frequency of sinusoidal jitter and in (b) the dashed bounded, constant (square wave) DJ distribution could be caused by triangle-wave phase modulation. In these examples, $\sigma = 0.15 \times DJ(p-p)$, the solid curve is the convolved jitter distribution and the dash-dot lines are the dual-Dirac approximation. The vertical lines indicate where the dual-Dirac model sets the means of Gaussians, μ_R and μ_L .

3.1 Q-scale

To fully understand the dual-Dirac model it is useful to study a linearized version of the bathtub plot, Figure 4. We introduce a variable Q, instead of BER, for the vertical axis. The advantage of the “Q-scale” is that a Gaussian jitter distribution is a straight line in Q(x) which makes it easier to see how the dual-Dirac model really works in estimating TJ(BER) and why $DJ(\delta\delta) < DJ(p-p)$.

If we set the time-delay sufficiently far from the DJ distribution so that J(x) can be described by a Gaussian, then BER_L(x), using Eqs. (5) and (9), is given by

$$BER_L(x) = \rho_T \frac{1}{\sqrt{2\pi}\sigma} \int_x^\infty \exp\left[-\frac{(\mu_L - x')^2}{2\sigma^2}\right] dx' \quad (13)$$

Now let

$$Q = \frac{\mu_L - x}{\sigma} \quad (\text{if } J(x) \text{ is Gaussian}) \quad (14)$$

so that Eq. (13) becomes

$$BER_L(Q) = \rho_T \int_Q^\infty \exp\left(-\frac{Q'^2}{2}\right) dQ'$$

The complementary error function is given by

$$\text{erfc}(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_x^\infty \exp\left[-\frac{x'^2}{2\sigma^2}\right] dx' \quad (15)$$

so that Eq. (14) can be written

$$\begin{aligned} BER_L(Q) &= \rho_T \text{erf}\left(\frac{Q}{\sqrt{2}}\right) \\ &= \rho_T \left(1 - \text{erf}\left(\frac{Q}{\sqrt{2}}\right)\right) \end{aligned} \quad (16)$$

In the last step, I substituted the error function, erf(x), for the complementary error function, erfc(x); the two are related by $\text{erfc}(x) = 1 - \text{erf}(x)$.

The reason we're doing this is to map BER(x) as shown in the bathtub plot of Figure 5 to Q(x) where Gaussian distributions are straight lines. Next, we invert Eq. (16) from BER(Q) to Q(BER),

$$Q(BER) = \sqrt{2} \text{erf}^{-1} \left[1 - \frac{1}{\rho_T} BER \right] \quad (17)$$

where erf^{-1} indicates the inverse of the error function (there are several good approximations to erf^{-1} , but no closed form solution). To this point we have assumed a Gaussian distribution. But now that we have $Q(\text{BER})$ we can generalize to any case. Remember, $\text{BER}(x)$ describes the dependence of the BER on the time-delay position of the sampling point. Similarly, by replacing BER with $\text{BER}(x)$ in Eq. (17) we get a definition of Q that does not rely on the form of the jitter distribution.

$$Q(x) = \sqrt{2} \text{erf}^{-1} \left[1 - \frac{1}{\rho_T} \text{BER}(x) \right] \quad (\text{general definition}) \quad (18)$$

The important distinction is that Eq. (14) gives Q in the special case where $J(x)$ follows a Gaussian distribution, but Eq. (18) is true for any jitter distribution.

Finally, turning to Figure 6 we can decipher how the dual-Dirac distribution models the actual distribution. In Figure 6a the solid line gives $Q(x)$ and the dashed line gives the dual-Dirac approximation to $Q(x)$.

The statement – the dual-Dirac model is a Gaussian approximation to the outer edges of the jitter distribution displaced by a fixed amount, $DJ(\delta\delta)$ – is demonstrated in Figure 6. In the absence of DJ, $J(x)$ would be purely Gaussian, the dual-Dirac model would give $\mu_L = 0$ and $\mu_R = T_B$ giving $DJ(p-p) = DJ(\delta\delta) = 0$. In Figure 6a the convolution of a bounded DJ distribution with a Gaussian RJ distribution results in $Q(x)$ whose tails asymptotically approach the Gaussian tails as described by Eq. (7) with slope $1/\sigma$ as required by Eq. (14). Figure 6b shows all of the parameters of the dual-Dirac model explicitly, it presents $Q(x)$ centered about a crossing point rather than about the center of the eye; the same symmetry as the jitter distributions shown in Figure 3a and Figure 5. Figure 6b is made by moving the right slope of $Q(x)$ – due to eye closure from the right crossing point – across the left slope and joining it at $x = 0$. The advantage of Figure 6b is that $TJ(10^{-12})$, $DJ(p-p)$, and $DJ(\delta\delta)$ can all be shown explicitly.

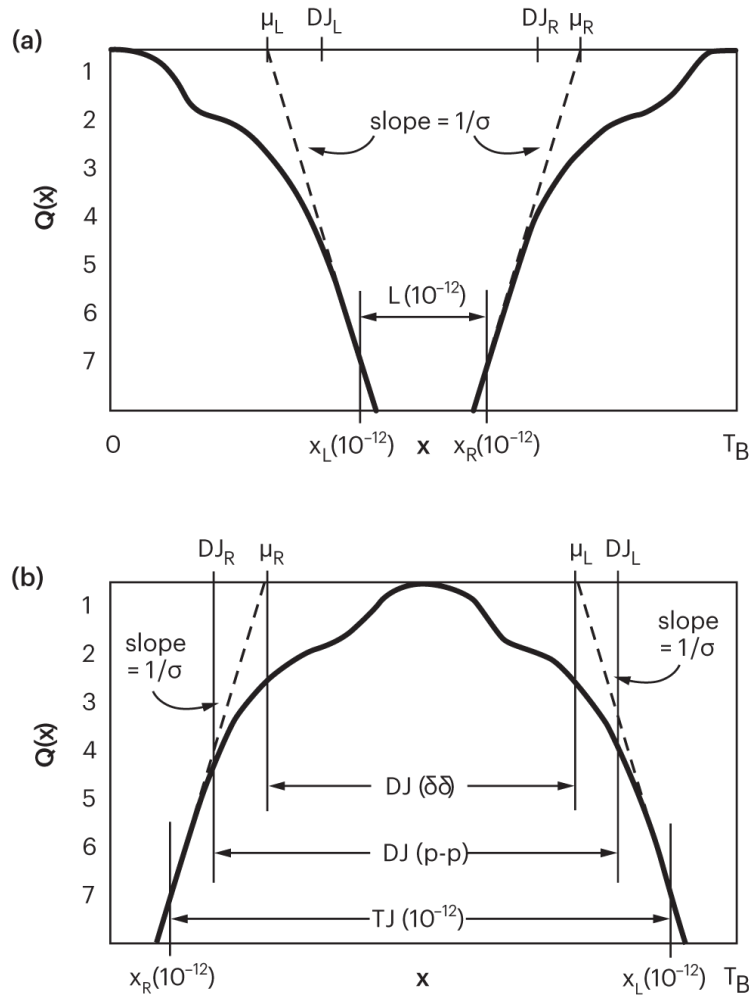


Figure 6. (a) The Q-scale version of a bathtub plot – $Q(x)$ rather than $BER(x)$ – where Gaussian effects are straight lines of slope $1/\sigma$. The dashed line gives the dual-Dirac approximation to $Q(x)$. (b) The same plot, but with the right side of the bathtub plot shifted over to the left to explicitly show $TJ(10^{-12})$ and emphasize the difference between $DJ(p-p)$ and $DJ(\delta\delta)$.

Notice the difference between $DJ(p-p)$ and $DJ(\delta\delta)$ in Figure 6: they are completely separate quantities. While the accuracy of the dual-Dirac model is improved by inserting a value for σ of greater accuracy, the accuracy is degraded by replacing $DJ(\delta\delta)$ with $DJ(p-p)$; the dashed lines in Figure 6a would be displaced inward from μ_L to DJ_L and μ_R to DJ_R resulting in a considerable overestimate of $TJ(BER)$. Once again, $DJ(\delta\delta)$ does not estimate of $DJ(p-p)$.

Given Eqs. (13) through (18) we can see where the values for Q_{BER} in Table 2 come from. The BER due solely to the Gaussian tails is given by Eq. (16) with Q given by Eq. (14). For transition density, $p_r = 1/2$, $BER = 10^{-11}$ corresponds to $Q = 6.7$, for $BER = 10^{-12}$, $Q = 7.0$, and so on. Equation (1) gives $TJ(BER)$ due to the Gaussian tails, $2Q_{BER}\sigma$, plus a fixed displacement given by $DJ(\delta\delta)$: $TJ(BER) = 2Q_{BER}\sigma + DJ(\delta\delta)$.

3.2 Application of the dual-Dirac model

Once μ_R , μ_L , and σ are obtained, TJ(BER) can be estimated by using Eqs. (9) and (10) with the dual-Dirac jitter distribution in place of $J(x)$ to calculate $x_L(\text{BER})$ and $x_R(\text{BER})$ and then TJ(BER) as in Eq. (12), or just use Eq. (1), TJ(BER) = $2Q_{\text{BER}}\sigma + DJ(\delta\delta)$. Ultimately this process is the extrapolation of the tails of the observed distribution to the BER of interest. It can be a huge extrapolation. In most cases $J(x)$ is measured from the transitions of about 10^6 bits, corresponding to a BER of more than 10^{-6} ; extending the tails down to 10^{-12} is an extrapolation of at least six orders of magnitude.

The dual-Dirac model is applied in many different ways on different types of test equipment. Real-time oscilloscopes from different vendors use vastly different techniques that frequently yield widely varying results. A concise analysis of the major differences of the techniques is provided below in Section 4.

3.3 Combining the jitter of different elements in a system

The jitter for a system of network elements can be combined according to Eq. (3), repeated below, under the assumption that the jitter of different network elements is uncorrelated.

$$\sigma_{\text{Total}} = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2} \quad (3)$$
$$DJ_{\text{Total}}(\delta\delta) \approx DJ_1(\delta\delta) + DJ_2(\delta\delta) + \dots + DJ_N(\delta\delta)$$

Remember that uncorrelated jitter distributions are combined by convolution, Eqs. (4) and (6). Since the convolution of two uncorrelated Gaussian distributions gives a Gaussian distribution whose width is the RSS of the individual widths, the Gaussian RJ of each component can be combined through RSS as in Eq. (3).

The peak-to-peak DJ of a convolution of uncorrelated distributions is the sum of the peak-to-peak DJs of the components. But, unlike for RJ, we have no assurance that the DJ sources are uncorrelated. For example, if a transmitter sends a signal with ISI into a channel, the response of the channel – its filtering effects and so forth – will be different than it would were there no ISI from the transmitter. In this case DJ(p-p) emitted from the channel is at most the sum of the DJ(p-p) of the transmitter and the channel,

$$DJ_{\text{Total}}(\text{p-p}) \leq DJ_1(\text{p-p}) + DJ_2(\text{p-p}) + \dots + DJ_N(\text{p-p})$$

Of course what we really need is $DJ_{\text{Total}}(\delta\delta)$, not $DJ_{\text{Total}}(\text{p-p})$. It is not necessarily the case that the sum of $DJ_i(\delta\delta)$ gives $DJ_{\text{Total}}(\delta\delta)$; hence the approximate relationship denoted by “ $\approx$ ” in Eq. (3). Still the approximation should be reasonably accurate. As more DJ sources are convolved the resulting DJ distribution is smoother about the edges. The difference between DJ(p-p) and DJ($\delta\delta$) is larger for smoother distributions – for example, compare DJ($\delta\delta$) between Figure 5a and b. It is therefore reasonable to expect that $DJ_{\text{Total}}(\delta\delta)$ resulting from the convolution of the DJ distributions of each network element should be smaller than that given by the naïve sum in Eq. (3), rendering Eq. (3) a conservative estimate of the combined RJ and DJ($\delta\delta$) of the system. In some cases the estimate for DJ($\delta\delta$) given in Eq. (3) can give an appreciable overestimate that could lead you to believe that a system would fail its BER requirement when it would actually pass. The best way to determine how the effects combine is to study the data-dependent jitter of the system. For example, a measurement of DDJ vs bit can be performed quite accurately and provide concise information of how the DJ of different sources interferes.

From the perspective of interoperability – whose assurance is the primary role of a technology specification – Eq. (3) is a conservative way to combine σ and $DJ(\delta\delta)$ for use in Eq. (1) to estimate $TJ(BER)$.

4. Different Implementations Of The dual-Dirac Model

The dual-Dirac model is applied in two fundamentally different ways:

- The fitting techniques analyze $J(x)$ or $BER(x)$ to determine both σ and $DJ(\delta\delta)$.
- The *independent- σ* techniques determine σ with a direct measurement of timing noise and then derive $DJ(\delta\delta)$ with a variety of techniques¹.

The fitting techniques are used to estimate $TJ(BER)$ quickly on BERTs, oscilloscopes, and TIAs. There are two distinct approaches to the fitting techniques that operate in nearly equivalent ways. They both fit separate Gaussian distributions to the left and right edges of the distribution to yield four parameters σ_R , σ_L , μ_R , and μ_L ; they both average σ_R and σ_L and assign $\sigma = 1/2(\sigma_R + \sigma_L)$. In one approach a simple Gaussian Eq. (5) is used to fit a fraction of the tails of the jitter distribution, $J(x)$. In the other approach the Gaussian inspired complementary error function, Eq. (15), is fit to the tails of a bathtub plot, $BER(x)$ or $Q(x)$. In ideal conditions the two techniques are equivalent. In ordinary conditions, the fit to the bathtub plot tends to be more repeatable because it is less affected by random fluctuations.

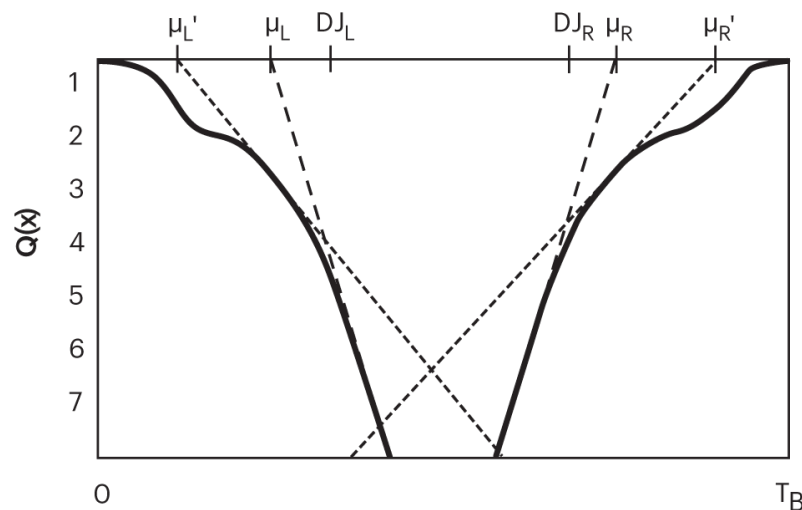


Figure 7. A plot of $Q(x)$ demonstrating how fitting techniques can overestimate RJ if the fit is applied to regions of Q above the asymptotes. The dashed (green) lines indicate the properly applied dual-Dirac model, the dot-dash (red) lines indicate a fit that over-estimates RJ and underestimates DJ resulting in the mistaken impression that the eye is completely closed at $BER=10^{-12}$ (i.e., $Q \sim 7$).

¹ Ransom Stephens et al., "Comparison of jitter analysis techniques with a precision jitter transmitter," Keysight Technologies Whitepaper, 2005. Available from www.Keysight.com.

The accuracy of the fitting techniques is extremely sensitive to the region of the distribution to which the fit is applied. The fit should be performed on only the tails of the distribution that truly follows the underlying RJ Gaussian given by Eq. (7), but it is difficult to determine where this asymptotic behavior begins.

Consider $Q(x)$ in Figure 7. When Gaussian tails are fit to the region $Q(x) > 5$, where $Q(x)$ is linear, the fit matches the asymptotes (i.e., the green dashed lines) and the slope gives an accurate estimate of $1/\sigma$. But if the fit is applied to the region $2 < Q(x) < 4$ then a line with a smaller slope (larger σ) results (i.e., the red dot-dashed lines) and yields an overestimate of $TJ(BER)$. In this example, the fitting technique would indicate that the eye is completely closed at $BER = 10^{-12}$ (which corresponds to $Q = 7.0$). Whether or not the fit is applied low enough on the slopes depends on the statistical significance of the measurement. That is, the more data acquired, the farther down the slopes one can fit. The problem is that the effect of convolving the RJ Gaussian and the bounded DJ distribution yields a distribution whose tails are well parameterized by a Gaussian with a larger value of σ than that of the actual underlying RJ Gaussian even very close to the DJ dominated part of the curve. As a result, goodness-of-fit estimators (e.g., fit confidence levels or correlation values) are rarely useful indicators of where the distribution genuinely follows the underlying RJ Gaussian. The only way to be certain that the fitting technique is accurate is to acquire a generous statistical sample and appreciable increases in the sample size require exponentially longer test times.

While there is no way to be certain that the fit is performed far enough down the tails, you can at least assure that the fit doesn't contradict itself with some simple convergence criteria. By repeating the fit for successively larger statistical samples to get to lower values of Q , $DJ(\delta\delta)$ should converge to a constant value. The slopes on the left and right edges should also converge to a common value (for electrical systems [7]). Depending on the test device, it may not be possible to acquire enough data to meet these convergence requirements – remember, only a BERT can realistically provide data well down the slopes to BERs of 10^{-9} and lower. To make it worse, convergence doesn't guarantee that your fit is far enough down the asymptotes; convergence is necessary but not sufficient.

To illustrate the biases introduced by a fitting technique consider Figure 8. Figure 8a, the ratio of the RJ fit results, σ_{fit} , to the actual RJ, σ_{true} , as a function of $\sigma_{true}/DJ(p-p)$, shows how RJ tends to be overestimated. Notice that in the limit of low and high RJ, σ_{fit} converges to σ_{true} . Figure 8b, the ratio of the $DJ(\delta\delta)$ results from a fitting technique to the actual $DJ(p-p)$ as a function of $\sigma_{true}/DJ(p-p)$, shows that $DJ(\delta\delta)$ and $DJ(p-p)$ are very different quantities; only in the limit of zero RJ does $DJ(\delta\delta)$ approach $DJ(p-p)$. Figure 8b is an excellent demonstration of the relationship $DJ(\delta\delta) < DJ(p-p)$. Figure 8c, the ratio of $TJ(10^{-12})$ from a fitting technique to the true $TJ(10^{-12})$ as a function of $\sigma_{true}/DJ(p-p)$, shows how fitting techniques tend to overestimate TJ. Since the data in this example is from a simulation, there are no statistical fluctuations; hence, the results are more accurate than one should reasonably expect from actual data.

When σ is measured independently, for example by measuring the rms timing noise in the frequency domain, an accurate value can be obtained from a much smaller data set than by fitting. The slopes of $BER(x)$, or $Q(x)$, are then fixed and the sum of two Gaussians can be matched to the tails with only the centers of the Gaussians, μ_R and μ_L , allowed to vary.

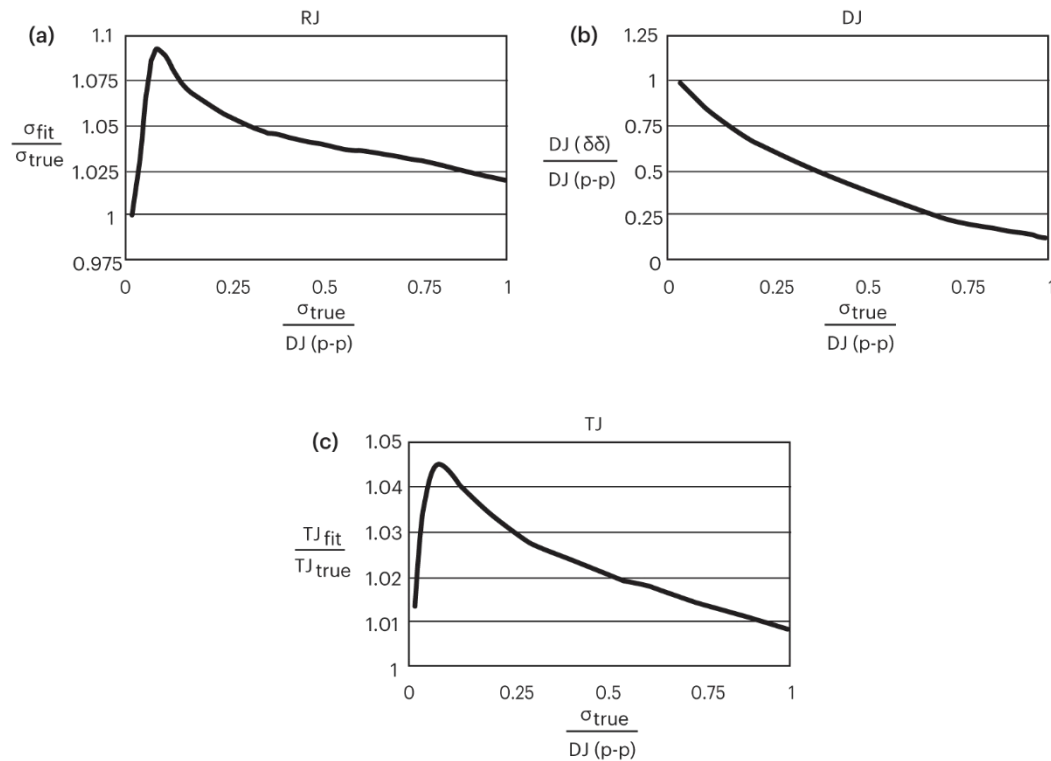


Figure 8. Implementation of a fitting technique to the dual-Dirac model for different simulated RJ/DJ ratios. The fit is applied to simulated data from $10^{-6} < \text{BER} < 10^{-3}$. In (a) the dual-Dirac estimate of RJ is typically overestimated by 5% but varies with the ratio RJ/DJ; (b) demonstrates that DJ($\delta\delta$) and DJ(p-p) are different quantities; and in (c), the resulting over-estimate of TJ(10^{-12}) is shown. DJ in this example is dominated by ISI from a backplane model and the data pattern is a pseudo-random binary sequence of length 27^{-1} at 2.5 Gb/s.

It is interesting to contrast the fitting and independent- σ techniques. In the fitting techniques one assumes that RJ follows a Gaussian distribution before the measurement and in the independent- σ techniques the assumption is made after it is measured. In practice, the fitting techniques break down when a set of DJ components convolve to form a DJ distribution with smoothly decaying tails. The central limit theorem of statistics says that the convolution of an infinite number of small uncorrelated processes, regardless of their individual distributions, follows a Gaussian distribution. This means that the more DJ components are included, the more that the resulting total DJ distribution resembles a Gaussian and the larger the statistical sample required of the fitting techniques. The effect is easy to observe by introducing data dependent jitter with a backplane or by increasing the length and complexity of the test pattern; the fitting techniques give larger values for σ unless the region included in the fit is appropriately adjusted. The fitting techniques cannot discriminate the bounded DJ from the unbounded RJ and end up overestimating σ . Since σ has a greater effect on TJ(BER) than does DJ($\delta\delta$), an appreciable overestimate of σ can give a substantial overestimate of TJ(BER).

To see how different implementations give different results, consider Figure 9. The precision jitter transmitter described in Ref.¹ was used to generate jitter signals on a 2.5 Gb/s, 2^7-1 pseudo-random binary sequence. Measurements from real-time oscilloscopes, time interval analyzers, the Keysight N4901B SerialBERT, and the Keysight 86100C DCA-J equivalent-time sampling oscilloscope are shown – only the measurements performed by Keysight equipment are labeled – the details of the study are given in Ref.². The data in Figure 9 is separated by column. Each column includes data from a given jitter condition described by the bars at the bottom. You can see that all of the test sets perform adequately when only periodic jitter is applied and, to a lesser extent when only RJ is applied. But as data dependent jitter is introduced – by introducing different lengths of backplane – the test sets give wildly disparate results. Figure 9b shows that every test-set except the DCA-J mistake DDJ for RJ. The situation deteriorates for most test-sets as the number and magnitude of different jitter sources are introduced.

Figure 9 demonstrates that the Keysight 86100C DCA-J is the only jitter analysis system that provides consistent, accurate measurements of $TJ(10^{-12})$ and its subcomponents. The Keysight N4901B SerialBERT is the only other test-set that gives consistently accurate estimates of $TJ(10^{-12})$. The BERT performs the simplest form of the fitting technique implementation of the dual-Dirac model described in Section 4. The other fitting techniques include algorithmic variations beyond the simple dual-Dirac model that make their results more difficult to understand. It is ironic that the real-time oscilloscopes and TIAs use more complicated algorithms for estimating $TJ(BER)$ and RJ, but yield less accurate results.

The BERT measurements in Figure 9 were performed by transmitting 3×10^9 bits in 1 ps steps of time-delay; the complementary error function, Eq. (15), was fit to the bathtub plots in the region $BER < 10^{-4}$ in all cases. You can see how the fitting technique over-estimate σ and $TJ(10^{-12})$ as the complexity of the DJ distribution increases.

The advantages of the BERT fitting technique compared to fitting techniques implemented on oscilloscopes and time interval analyzers are:

1. It's easy to understand the simple implementation of the dual-Dirac model.
2. There is a lot of freedom in setting the region of the curve that is included in the fit because the statistical sample can be increased at the full data rate.
3. And, most importantly, fast estimates of $TJ(BER)$ can be confirmed by performing a measurement of $TJ(BER)$ when in doubt³.

1 Jim Stimple and Ransom Stephens, "Precision Jitter Transmitter," DesignCon 2005, <http://www.designcon.com/conference/7-wa1.html>.

2 Ransom Stephens et al., "Comparison of jitter analysis techniques with a precision jitter transmitter," Keysight Technologies Whitepaper, 2005. Available from www.Keysight.com.

3 Marcus Müller, Ransom Stephens, and Russ McHugh, "Total Jitter Measurement at Low Probability Levels, Using Optimized BERT Scan Method," DesignCon 2005, <http://www.designcon.com/conference/7-ta4.html>.

The reasons that the DCA-J gives the most accurate results are simple. The DCA-J's first step is to separate jitter that is correlated to the test pattern (i.e., data dependent jitter) by using its built in pattern trigger and trigger divider to analyze each edge of the pattern independently. The hardware-based DDJ separation technique immediately prevents DDJ from being confused for RJ. RJ is then measured from just a data set whose DDJ has been removed with an independent- σ technique.

One of the real-time oscilloscope techniques also uses an independent- σ technique but consistently mistakes DDJ for RJ anyway. The DCA-J technique is described in detail in Ref.¹. For several years the real time oscilloscopes and TIAs have delivered inconsistent measurements of RJ, PJ, DDJ, and estimates of TJ(10^{-12}) but it is only with the introduction of the DCA-J that reliable measurements in this field have been realized.

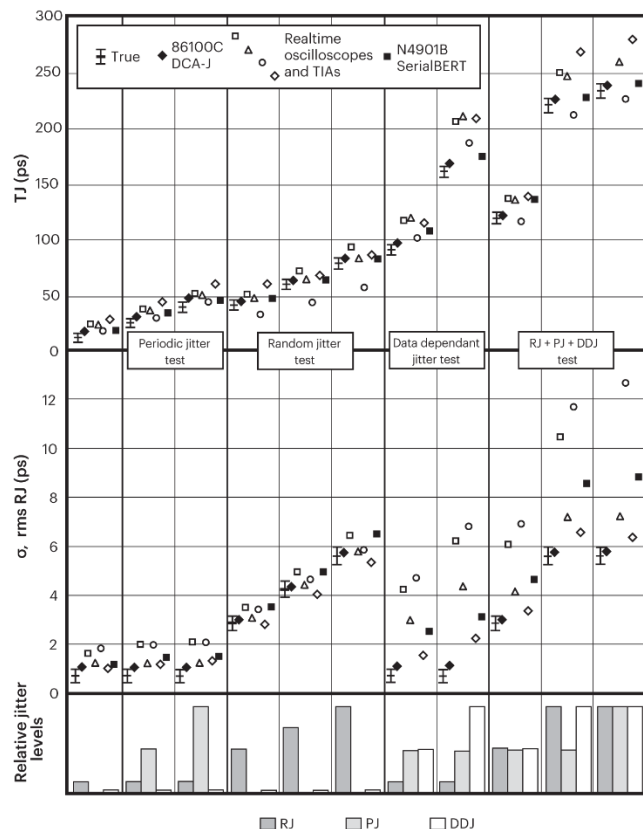


Figure 9. Comparison of jitter test-set results for TJ(10^{-12}) and σ (i.e., RJ). Each column of the graphs corresponds to a jitter condition indicated by the bars at the bottom of the column. The left most data-point in each column is the calibrated level of jitter from the precision jitter transmitter described in Ref.². The error bars on the calibrated values indicate the systematic uncertainty of the calibration. The unlabeled measurements were performed on common jitter test-sets such as real-time oscilloscopes and time interval analyzers; one of these measurements is off the scale in the rightmost column. The data from the Keysight equipment is labeled.

1 Keysight Technologies Product Note, 86100C-1, "Precision jitter analysis using the Keysight 86100C DCA-J," Keysight Literature Number 5989-1146EN, <http://cp.literature.keysight.com/litweb/pdf/5989-1146EN.pdf>.

2 Jim Stimpale and Ransom Stephens, "Precision Jitter Transmitter," DesignCon 2005, <http://www.designcon.com/conference/7-wa1.html>.

5. Conclusion

5.1 The reality of the standard, industry-wide assumptions

The assumptions of the dual-Dirac model, Table 1, can be debated. The Gaussian RJ assumption is based on the idea that RJ is caused by the thermal behavior of a Fermi gas of electrons in a conductor that induces timing fluctuations that follow a Gaussian distribution¹. The debate over the Gaussian RJ assumption centers on whether the observed RJ levels are consistent with the levels one should expect from purely thermal processes. The argument is that thermal processes, in most cases, can only account for one or two picoseconds of rms RJ, but we frequently see three or four times that. The central limit theorem begs the possibility that many small DJ sources could convolve into a distribution that follows a Gaussian that is truncated and could easily be mistaken for RJ.

The last assumption, that jitter is a stationary phenomena, is nearly intractable. All jitter test-sets that estimate TJ(BER) sample the data. None of them analyze a continuous data stream. Real-time oscilloscopes capture finite lengths of data and post-process them; even when separate captures are combined there is a long trigger re-arm time between data captures. BERTs analyze every bit as a digital stream with a given sampling point position, and equivalent time sampling oscilloscopes build statistical distributions by repetitively sampling the time-delay and voltage of a signal. None of these techniques analyze a continuous data stream; none of them can reliably detect nonstationarity of a jitter signal. The only jitter analysis technique that analyzes a continuous data stream is phase-noise analysis of clock signals commonly used in SONET/SDH jitter analysis. The SONET/SDH jitter analysis is band limited and does not relate timing-noise to BER; Ref.² provides a comprehensive description of SONET/SDH jitter analysis.

The discrimination of timing and amplitude noise is another difficulty. Crosstalk between neighboring data paths causes errors by changing the crossing point position in time and is frequently referred to as nonperiodic bounded uncorrelated jitter (BUJ). But cross-talk is amplitude noise, not timing noise - it's not jitter, though it causes errors the same way that jitter does. None of the jitter test-sets on the market are capable of reliably analyzing BUJ simply because it can take on too many forms. But without a more comprehensive technique for analyzing amplitude and timing noise BUJ interferes with the implementation of the dual-Dirac model on all test equipment.

¹ L.D.Landau and E.M.Lifshitz, Statistical Physics, Part 2, 3rd ed., Pergamon Press, 1980.

² Keysight Tutorial, "Understanding Jitter and Wander Measurements and Standards, Second Edition," Keysight Literature number 5988-6254EN, 2003. Available from www.Keysight.com.

The finite bandwidth of a receiver can cause amplitude noise to timing noise conversion. The jitter properties of a receiver are usually tested with some type of jitter tolerance measurement – a test of the level of jitter that a receiver can tolerate. A jitter test-set with a receiver of limited bandwidth can experience the same amplitude noise to jitter conversion and overestimate the jitter on a signal. However, a receiver's intrinsic noise increases with bandwidth. The conflict is another case where the DCA-J is the best solution. Receivers with very high bandwidth are available and the noise floor of the DCA-J is significantly lower than that of real-time oscilloscopes or TIAs.

In most cases noise is noise. The separation of jitter and amplitude noise is not always transparent or performed independently. For example, inter-symbol interference contributes to both. ISI affects the slew rate and amplitudes of different bits in a signal as well as the transition time. Generally, in diagnosing the problems of a system, it is wise to keep in mind the complexity of the signal.

5.2 Summary

The dual-Dirac model:

- Provides a useful way to quickly estimate TJ(BER) at low BER.
- Is built on the assumption that RJ follows a Gaussian distribution that can be described by a single parameter, σ
- Defines an observable DJ($\delta\delta$) that can be measured in a variety of different ways but should not be confused with the true peak-to-peak DJ, DJ(p-p). Generally, DJ($\delta\delta$) < DJ(p-p).
- Provides a transparent way to combine σ and DJ($\delta\delta$) of different network elements to estimate TJ(BER) for a system.
- Provides observable, well defined quantities for use in technology standards.
- Can be applied in different ways that have different systematic uncertainties that can lead to different errors that make comparison of results difficult.
- Is most accurately implemented through the technique introduced on Keysight's 86100C DCA-J that combines a hardware separation of jitter that is correlated to the test pattern from that which is uncorrelated with a direct measurement of timing noise, σ .

Keysight enables innovators to push the boundaries of engineering by quickly solving design, emulation, and test challenges to create the best product experiences. Start your innovation journey at www.keysight.com.



This information is subject to change without notice. © Keysight Technologies, 2017 – 2023, Published in USA, April 5, 2023, 5989-3206EN