The Truth about 2-level Transition Elimination in Bang-Bang PAM-4 CDRs

Marijn Verbeke, Guy Torfs, and Pieter Rombouts

***Abstract*—Reception of 4-level pulse amplitude modulation (PAM-4) requires a clock and data recovery (CDR) circuit, typically implemented by a PLL-like structure. An essential block in such a CDR is the phase detector which should detect whether the recovered clock leads or lags the incoming data edges. In typical implementations an incoming data edge is detected by sensing whether the incoming waveform crosses a data threshold level. However, there is some ambiguity in detecting the incoming data edge because PAM-4 modulation has 3 thresholds. If the waveform crosses multiple threshold levels, the level crossings will occur at different time instants due to the finite rise/fall time of the incoming waveform.**

In this work, we first analyze qualitatively and quantitatively CDR systems that use one threshold for phase adjustment. Here, *eliminating the 2-level transitions* decreases the amount of jitter injected by the phase detector. However, the available transitions for phase adjustment are also reduced, which lowers the CDR's robustness. Secondly, for CDR systems using three thresholds, a combination of two techniques: *majority voting* and *elimination of 2-level transitions* is investigated. We prove that in this case, the *elimination of 2-level transitions* is not needed and even gives a worse performance when implemented.

***Index Terms*—4-level pulse amplitude modulation (PAM-4), clock and data recovery (CDR), Alexander Bang Bang phase detector, majority voting, transition elimination, transition selection, minor, middle, intermediate, asymmetric, major transition.**

I. Introduction

TO support ever increasing data rates, the traditional on-off-keying (OOK) non-return-to-zero (NRZ) signaling scheme employed in wireline and optical interconnects is progressively being replaced by 4-level pulse amplitude modulation (PAM-4) [1], [2]. Thanks to the higher spectral efficiency, the data rate can be doubled within the same bandwidth budget. However, this comes at the cost of increased complexity both at the transmitter as well as at the receiver side. This manuscript focuses on a critical block in the receiver: i.e. the clock-and-data recovery (CDR), which converts the incoming waveform into a synchronized data stream with its clock. A typical block diagram of such a CDR is shown in Fig. 1. One of its core blocks is the phase detector (PD), which compares the position of the edges of the recovered clock relative to the edges of the incoming data waveform D_{in}. Depending on this comparison, either an *early* or a *late* signal can be produced, indicating that the recovered clock leads or lags. Since it is the goal of the CDR to align the edges of the recovered clock and the incoming waveform, the *early* and *late* signal can be interpreted as an error signal. This error signal is then filtered by a loop filter and used to adjust a voltage controlled oscillator (VCO).

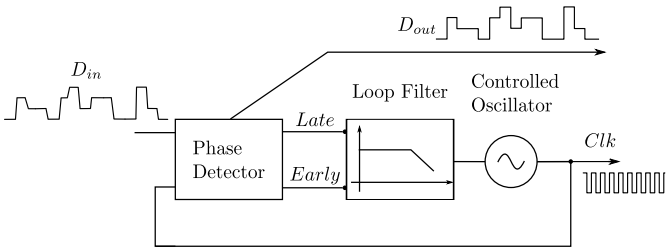


Fig. 1. Conceptual block diagram of a PLL-based CDR circuit.

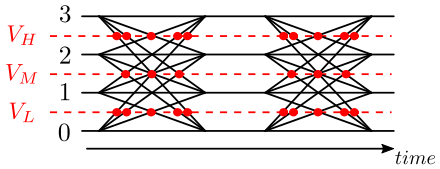


Fig. 2. Illustration of the possible transitions in a PAM-4 waveform. The dots indicate the threshold crossings.

Compared to OOK, especially the PAM-4 phase detector block becomes significantly more complex, because multiple transition scenarios need to be considered to detect the edges of the incoming waveform. In practice, this edge detection is implemented by detecting whether the incoming waveform crosses a certain threshold level. Since there are 4 valid signal levels, three threshold levels can be considered: V_L, V_M and V_H. This is illustrated in Fig. 2 together with all the possible data transitions. An important observation is that the data transitions always occur with a finite rise/fall time due to bandwidth limitations in the channel between the transmitter and receiver. This creates the key problem for PAM-4 CDRs: i.e. the location of the threshold crossing is data dependent. This is also evident from the dots in Fig. 2.

This problem was already recognized by numerous authors [3]–[28] which all report on the difficulty of multi-level phase detection. A first solution is to use only 1 threshold level to detect the data transitions. This reduces the intrinsic performance of the phase detection, but it is a quite common

M. Verbeke and G. Torfs are with imec - IDLab, Dept. Information technology of Ghent University, 9000 Gent, Belgium (marijn.verbeke@imec.be and guy.torfs@ugent.be). P. Rombouts is with the Electronics and Information Systems Department, Ghent University, 9000 Gent, Belgium (pieter.rombouts@ugent.be).

This paper is a postprint of a paper submitted to and accepted for publication in IEEE Transactions on Circuits and Systems I: Regular Papers. The copy of record is available at IEEE Xplore: M. Verbeke et al., "The Truth about 2-level Transition Elimination in Bang-Bang PAM-4 CDRs," in IEEE Transactions on Circuits and Systems I: Regular Papers. doi: 10.1109/TCSI.2020.3033202

approach followed in e.g. [4]–[13]. It is well known that in this simplified implementation, transitions that cross exactly 2 threshold levels will lead to incorrect edge detection. Such transitions are called *2-level crossings* (or *middle* or *intermediate* or *asymmetric* transitions). Transitions that cross either only 1 or 3 threshold levels result in correct edge detection. The mechanism behind this, is reviewed in Section II and a new quantitative analysis is performed in Section III. Some authors [11]–[13] decided to tolerate the incorrect edge detection and we will show that in some cases this yields a better performance than omitting the timing information from the 2-level crossings. We will also analyze another solution provided by [10]. Here, the 2-level crossings are only used to correct the sampling position if the sampling position occurs "very late" or "very early". We will show that this *partial 2-level elimination* outperforms the case where all the 2-level crossings are omitted. An alternative strategy [8], [9] is to use line coding techniques to avoid 2-level crossings. Unfortunately, this creates an overhead due to the bit/symbols that need to be added to eliminate the 2-level crossings.

Section IV proceeds with the case where three thresholds are used for timing adjustment. Also here quite some implementations have been published [13]–[25]. Surprisingly, all these publications still incorporate techniques to eliminate 2-level crossings. This is what this paper is about: we will demonstrate in Section V with a full quantitative analysis of a CDR system with three thresholds that it is completely unnecessary to remove these transitions. This means that the additional overhead to deal with the 2-level transitions can be eliminated, leading to a more efficient implementation. Additionally this also means that more edge information is available to the CDR circuit such that a higher performance can be achieved. It should be noted that there is already one publication [26] which uses all three thresholds without 2-level transition elimination. Below we will also discuss the solution of [26] and compare it to other edge detection schemes. The analysis will also show that a phase detector which uses all three thresholds has a better performance compared to its one threshold alternative.

Phase detection techniques that use even more than 3 thresholds are discussed in [27], [28]. But in these cases, more samplers are required. This further increases the power consumption. Therefore these techniques are considered less favorable and are not discussed in this work.

Finally, simulation results are presented in Section VI and Section VII concludes the paper.

II. Phase Detection Based on One Threshold

In high-speed application, the PD is typically implemented as a binary phase detector that only indicates whether the recovered clock leads ($Early$) or lags ($Late$) the input data. This binary phase detection operation can be achieved by using one comparator (with threshold V_M) and an Alexander phase detector (Fig. 3) [29]. Alexander phase detection is based on three successive data samples (S_0, S_1 and S_2), which are sampled at twice the data clock frequency. In a typical CDR, this is done by sampling the data both on the rising as well as the falling edges of the full-rate recovered clock Clk.

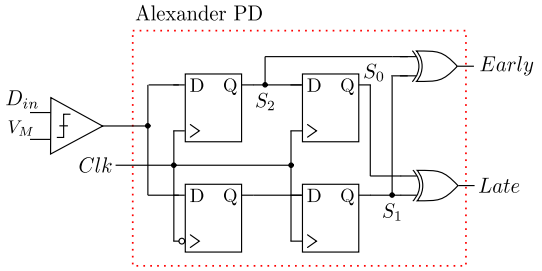


Fig. 3. Implementation of a PD with one threshold for timing adjustment.

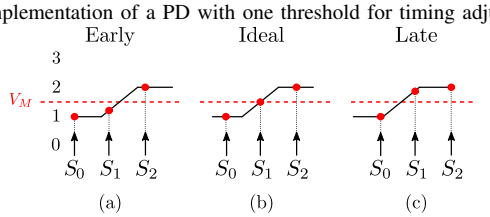


Fig. 4. Waveforms for the locking behavior of a 1-level transition using one threshold for timing adjustment. (a) the clock leads the data, (b) the clock is in the ideal locked position and (c) the clock lags the data.

Fig. 4 shows the three samples taken by an Alexander PD in the possible three cases: the clock is leading ($Early$), aligned ($Ideal$) or lagging ($Late$) with respect to the data edge. The output of the PD is calculated using the XOR operation and summarized as [29], [30]:

$$\begin{aligned} \text{Early}: S_0 \oplus S_1 = 0,\ \ S_1 \oplus S_2 = 1 \quad &\rightarrow \quad \text{Clk frequency} \downarrow \\ \text{Late}: S_0 \oplus S_1 = 1,\ \ S_1 \oplus S_2 = 0 \quad &\rightarrow \quad \text{Clk frequency} \uparrow \\ \text{Others}: S_0 \oplus S_1 = S_1 \oplus S_2 \quad &\rightarrow \quad \text{Do not adjust clk} \end{aligned}$$

The phase estimation based on a single threshold is correct when receiving OOK-modulated data. However, for PAM-4 modulation, the transition is data dependent. The decomposition of the PAM-4 eye diagram in Fig. 5 shows four distinct transition categories: *zero-level transitions, 1-level transitions, 2-level transitions* and *3-level transitions*. In literature, the three later categories are sometimes also referred to as *minor, middle* and *major transitions*, respectively.

For a zero-level transition (Fig. 5(b)), there is no edge information and hence these transitions are not relevant for this work. The 1-level and 3-level transitions are shown in Fig. 5(c) and (f), respectively. Here, the crossing point of these transitions is located at the midpoint of the data symbols. Please note that due to the use of one threshold, only the 3-level transitions and two out of the six 1-level transitions can be detected. This is indicated by the red dot in Fig. 5.

The last category comprises the 2-level transitions, which are shown by Fig. 5(d) and (e). Depending on the transitions, the crossing of the threshold V_M occurs at two different moments in time (indicated by the red dots). We will note the offset between these time instants and the ideal crossing by the offset d_2.

The probability that the data crosses the threshold V_M forms the threshold crossing distribution. As seen in the second row of Fig. 5, this threshold crossing distribution has three peaks: the center peak originates from the 1-level and 3-level transitions, while the two side peaks are the result of the 2-level transitions.

The bottom three traces analyze the response of three different PD topologies. Clearly, in the case of a 1-level and 3-

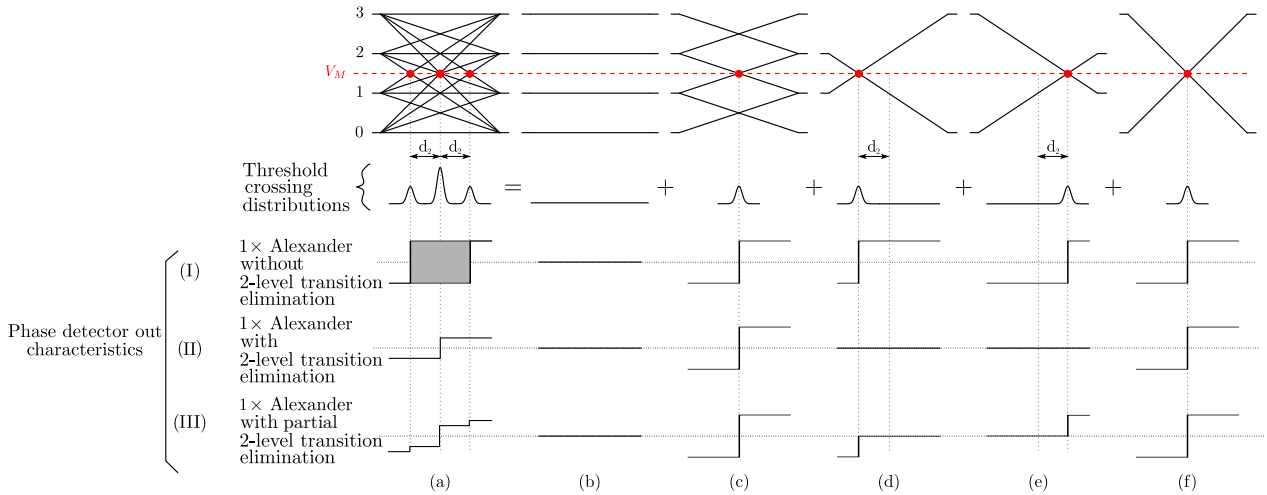


Fig. 5. The decomposition of different transitions of a PAM-4 signal when only considering the middle transition level V_M. (a) Combination of all possible transitions, (b) zero-level transitions, (c) 1-level transitions, (d) 2-level transitions leading to a left-shifted edge detection, (e) 2-level transitions leading to a right-shifted edge detection and (f) 3-level transitions. For each category of transitions, the threshold crossing points are indicated by the red dots and the threshold crossing distributions are given. (a) forms the combination of all possible transitions, corresponding threshold crossing points and threshold crossing distributions. The bottom three traces indicate the phase detector output characteristics for different scenarios.

level transition, the PD operation will detect the correct edge location (in the center of gravity of all possible transitions) corresponding to the ideal crossing. This results in the ideal bang-bang phase detector characteristic as illustrated in the bottom three traces of Fig. 5, column (c) and (f) respectively. For the case of a 2-level transition, the crossing of V_M will differ from the ideal crossing (either shifted to the left or to the right, see Fig. 5(d) and (e)). If these 2-level transitions are used (trace (I) of Fig. 5), the phase detector will produce a non-zero and contradictory control signal for the different transitions at certain sampling phases. Since the data transitions are highly uncorrelated, this occurs in a random way which manifests itself as a significant injection of additional jitter in the CDR. This is highlighted by the gray area in the overall combined phase detector characteristic for trace (I) in Fig. 5 column (a).

In order to avoid this excess jitter injection mechanism, the 2-level transitions of the incoming data should not be taken into account. This is also called *2-level transition elimination* or *middle transition elimination* and is represented by trace (II) in Fig. 5. An undesired side effect of this is that only four transitions (two 1-level and two 3-level) out of all 16 possible transitions are used for the clock recovery operation. This limited amount of usable transitions decreases the robustness of the CDR operation, because the CDR may have to wait a long time to receive an update signal to tune the recovered clock. Another approach to avoid the excess jitter injection mechanism is presented in [10] and uses *partial 2-level transition elimination*. In trace (III) of Fig. 5, the system only generates a correction signal if the recovered clock is "very early" or "very late". If the sampling point falls outside the very early/late regions, the PD ignores the extracted phase information. Note that in this scenario, the phase detector will also never provide conflicting information and therefore does not introduce additional jitter. However, some 2-level transitions are also not considered, which reduces the robustness of the CDR operation.

A quantitative analysis will provide a clear trade-off between the jitter injection due to the 2-level transitions in one case and the lower update rate in the other case.

Note that in reality the rise and fall times could be different or vary in time. Also the threshold V_M can deviate from its ideal value. These are additional parasitic effects that can also affect the effectiveness of the phase detector. A detailed elaborate study of such additional effects is out of the scope of this manuscript, but we performed a preliminary study which suggests that our conclusions remain unchanged in the presence of these effects.

III. Pseudo-Linearized Phase Detector Gain Using One Threshold

In order to get a quantitative measure whether the partial or complete 2-level transition elimination is beneficial for a PD using 1 threshold, we will study both the injected jitter and the loop dynamics of the CDR operation using a pseudo-linear analysis technique called describing functions. The use of describing functions [31] to model the non-linear PD behavior, has been proven to be a straightforward and accurate analysis method [32], [33].

First, the Alexander phase detector (Fig. 3) can be represented by a slicer in the phase domain, e.g. if the phase difference between input data and recovered clock ϕ_e is positive (the clock is $Early$) the phase detector will output +1. Vice versa if the phase difference ϕ_e is negative, the clock is $Late$ and the phase detector will output -1.

This non-linear element can now be converted into its describing functions. Assuming that the phase difference ϕ_e has two components: a random noise component $\phi_{e,n}$[1] and a bias $\phi_{e,B}$, we obtain the describing function model depicted by Fig. 6. Each input component is multiplied by its corresponding gain factor (resp. K_n and K_B). Similar to the phase difference, the output of the slicer ϕ_u comprises a random noise component $\phi_{u,n}$ and a bias $\phi_{u,B}$. The instantaneous difference ϕ_q between the output of the

[1]A random noise component in the phase domain is equivalent to jitter.

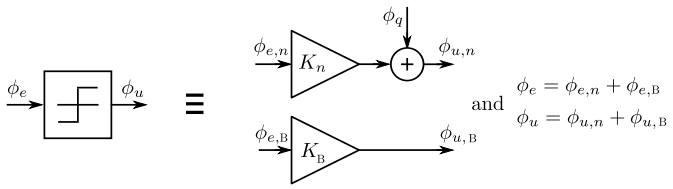


Fig. 6. The Alexander Bang-Bang phase detector can be decomposed in its pseudo-linear gains K_n and K_B.

Alexander phase detector ϕ_u and the output of the pseudo-linearized model $(K_n\phi_{e,n} + K_B\phi_{e,B})$ can be approximated as random distributed quantization noise[2]. This noise term is added to obtain a more accurate model [32], [33]. The linearized gains K_n and K_B should be calculated such that the power of the quantization noise is minimized [31], [32]. We can see from Fig. 5, that the input-output characteristics of the phase detector will differ for each transition type. Therefore, we decompose these gain factors using Bayes' rule for the different transitions types:

$$K_n = \frac{E\{\phi_{e,n}\phi_u\}}{E\left\{\phi_{e,n}^2\right\}} = \sum_X \text{Prob}[X]\frac{E\{\phi_{e,n}\phi_u|X\}}{E\left\{\phi_{e,n}^2\right\}} \equiv \sum_X \text{Prob}[X]K_{n,X} \quad (1)$$

$$K_B = \frac{E\{\phi_{e,B}\phi_u\}}{E\left\{\phi_{e,B}^2\right\}} = \sum_X \text{Prob}[X]\frac{E\{\phi_{e,B}\phi_u|X\}}{E\left\{\phi_{e,B}^2\right\}} \equiv \sum_X \text{Prob}[X]K_{B,X} \quad (2)$$

where X represents the transition type, i.e.: 1-level, 2-level (left), 2-level (right) and 3-level transitions indicated by Fig. 5(c), (d), (e) and (f) respectively. The "Prob$[X]$" operator represents the probability that one of the above transitions X occurs[3]. Furthermore, we define the sub-describing functions $K_{n,X}$ as $\frac{E\{\phi_{e,n}\phi_u|X\}}{E\left\{\phi_{e,n}^2\right\}}$ and $K_{B,X}$ as $\frac{E\{\phi_{e,B}\phi_u|X\}}{E\left\{\phi_{e,B}^2\right\}}$, because it can be seen as the pseudo-linearized gain of the Alexander PD for the case that only transition X occurs. Each of these transition dependent gain factors $K_{n,X}$ and $K_{B,X}$ can be calculated and are summarized in the following sections for the different scenarios.

A. Scenario (I): without 2-level transition elimination

For the first scenario, the 1-level, 2-level and 3-level transitions should be considered. For two of the 1-level transitions (Fig. 5(c)) and the 3-level transitions (Fig. 5(f)), the input-output characteristic of the phase detector for these transitions is symmetrical around zero, and thus there will be no DC offset in the phase error in steady state regime. Therefore the describing functions with two components (random noise and bias) reduce to a describing function with only one component, i.e. the random noise [31]. This is however not the case for the describing functions for the 2-level transitions. As shown in Fig. 5(d) and (e), the crossing occurs with an offset d_2 from its ideal locking position. For these transitions, this offset can be modeled as a bias component $\phi_{e,B}$ which is equal to $+d_2$ for Fig. 5(d) and equal to $-d_2$ for Fig. 5(e).

The sub-describing functions become:

$$K_{n,\text{1-level}} = K_{n,\text{3-level}} = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma} \quad (3)$$

$$K_{n,\text{2-level (right)}} = K_{n,\text{2-level (left)}} = 2\mathcal{N}\left(\frac{d_2}{\sigma}\right)\frac{1}{\sigma} \quad (4)$$

$$K_{B,\text{1-level}} = K_{B,\text{3-level}} = 0 \quad (5)$$

$$K_{B,\text{2-level (right)}} = K_{B,\text{2-level (left)}} = \frac{1}{d_2}\left[2\Phi\left(\frac{d_2}{\sigma}\right) - 1\right] \quad (6)$$

with d_2 the offset between the crossing of the 2-level transitions and the ideal crossing, σ the standard deviation of the phase difference $\phi_{e,n}$. The functions $\mathcal{N}()$ and $\Phi()$ are, respectively, the probability density function and the cumulative distribution function of the Standard Normal distribution and are defined by:

$$\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right) \quad (7)$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{v^2}{2}\right)dv \quad (8)$$

Using (1), we can write the noise gain K_n as:

$$K_{n,(I)} = \left(\frac{1}{4}\sqrt{\frac{2}{\pi}} + \frac{1}{2}\mathcal{N}\left(\frac{d_2}{\sigma}\right)\right)\frac{1}{\sigma} \quad (9)$$

Its upper and lower limit are defined by:

$$K_{n,(I),\text{upper}} = K_{n,(I)}\Big|_{d_2 \ll \sigma} = \frac{1}{2}\sqrt{\frac{2}{\pi}}\frac{1}{\sigma} \quad (10)$$

$$K_{n,(I),\text{lower}} = K_{n,(I)}\Big|_{d_2 \gg \sigma} = \frac{1}{4}\sqrt{\frac{2}{\pi}}\frac{1}{\sigma} \quad (11)$$

As shown by (6), the bias gain factors for the 2-level transitions, $K_{B,\text{2-level (left)}}$ and $K_{B,\text{2-level (right)}}$ are equal and not zero. However, as the transitions are uncorrelated, the expected value of the bias term of the output of the phase detector $\phi_{u,B}$ will be zero.

Any variation of ϕ_u due to a bias components (originating from the asymmetric 2-level transitions) can be perceived as random variation due to the uncorrelated nature of the transitions and will be included in the quantization error ϕ_q, for which its variance σ_q^2 is determined by (12) and (13).

[2]The random noise component of the phase error $\phi_{e,n}$ and the quantization noise ϕ_q can be described as a discrete-time white noise process which is sampled at the symbol rate. Therefore they can be approximated as white noise with a bandwidth of half the symbol rate.

[3]There are 2 possible *1-level transitions*, 4 possible *2-level transitions* and 2 possible *3-level transitions* which will generate an output different from zero.

$$E\left[\phi_u^2\right] = E\left[\left(K_n\phi_{e,n} + \phi_q\right)^2\right]$$
$$= K_n^2 E\left[\phi_{e,n}^2\right] + E\left[\phi_q^2\right]$$
$$= \left(\frac{1}{4}\sqrt{\frac{2}{\pi}} + \frac{1}{2}\mathcal{N}\left(\frac{d_2}{\sigma}\right)\right)^2 + \sigma_q^2 \quad (12)$$

$$E\left[\phi_u^2\right] = \sum_{\phi_u} \text{Prob}\left[\phi_u\right]\phi_u^2 = \frac{1}{2} \quad (13)$$

The variance of the quantization noise σ_q^2 and its lower and upper limit are given by:

$$\sigma_{q,(I)}^2 = \frac{1}{2} - \left(\frac{1}{4}\sqrt{\frac{2}{\pi}} + \frac{1}{2}\mathcal{N}\left(\frac{d_2}{\sigma}\right)\right)^2 \quad (14)$$

$$\sigma_{q,(I),\text{lower}}^2 = \sigma_q^2\Big|_{d_2 \ll \sigma} = \frac{1}{2} - \frac{1}{2\pi} \quad (15)$$

$$\sigma_{q,(I),\text{upper}}^2 = \sigma_q^2\Big|_{d_2 \gg \sigma} = \frac{1}{2} - \frac{1}{8\pi} \quad (16)$$

This gives a measure of the amount of jitter injected into the CDR loop due to the phase detector. The equations above also show that the quantization noise introduced in the system increases if the ratio of the offset between the crossing of the 2-level transitions and the ideal crossing d_2, and the standard deviation of the phase difference σ increases.

B. Scenario (II): with 2-level transition elimination

The sub-describing functions for the Alexander phase detector with 2-level transition elimination are the same as for the scenario without 2-level transition elimination, except for the 2-level transitions. These are omitted and the corresponding sub-describing functions are equal to zero. The sub-describing functions and combined noise gain are:

$$K_{n,\text{1-level}} = K_{n,\text{3-level}} = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma} \quad (17)$$

$$K_{n,\text{2-level (left)}} = K_{n,\text{2-level (right)}} = 0 \quad (18)$$

$$K_{n,(II)} = \frac{2}{16}\left(K_{n,\text{1-level}} + K_{n,\text{3-level}}\right) = \frac{1}{4}\sqrt{\frac{2}{\pi}}\frac{1}{\sigma} \quad (19)$$

where σ represents the standard deviation of the phase error $\phi_{e,n}$. If we compare (19) with (10) and (11), we can see that the noise gain K_n for the 2-level transition elimination case is equal to the lower limit of the noise gain $K_{n,\text{lower}}$ without 2-level transition elimination.

The variance of the quantization noise term σ_q^2 can be derived from (21) and (22):

$$E\left[\phi_u^2\right] = K_n^2 E\left[\phi_{e,n}^2\right] + E\left[\phi_q^2\right] \quad (20)$$
$$= \left(\frac{1}{4}\sqrt{\frac{2}{\pi}}\right)^2 + \sigma_q^2 \quad (21)$$

$$E\left[\phi_u^2\right] = \sum_{\phi_u} \text{Prob}\left[\phi_u\right]\phi_u^2 = \frac{1}{4} \quad (22)$$

The variance of the quantization noise σ_q^2 is thus given by:

$$\sigma_{q,(II)}^2 = \frac{1}{4} - \frac{1}{8\pi} \quad (23)$$

If we compare (15), (16) and (23), we can observe that the introduced quantization noise is the lowest when the 2-level transition elimination is used. This is in-line with the conclusions of prior-art work, and with our expectations which have been discussed in the previous section. But, the analysis also shows that the gain of the phase detectors is different when employing 2-level transition elimination. In order to make an accurate comparison between both phase detectors, a comparison should be made on the system level to incorporate both the effect of the quantization noise and phase detector gain. As we will see later this leads to a more nuanced conclusion.

C. Scenario (III): with partial 2-level transition elimination

The last scenario using 1 threshold for phase alignment applies partial 2-level transition elimination. The sub-describing functions are now given by:

$$K_{n,\text{1-level}} = K_{n,\text{3-level}} = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma} \quad (24)$$

$$K_{n,\text{2-level (right)}} = \mathcal{N}\left(\frac{-d_2}{\sigma}\right)\frac{1}{\sigma} \quad (25)$$

$$K_{n,\text{2-level (left)}} = \mathcal{N}\left(\frac{d_2}{\sigma}\right)\frac{1}{\sigma} \quad (26)$$

with d_2 the offset between the crossing of the 2-level transitions and the ideal crossing, σ the standard deviation of the phase difference $\phi_{e,n}$. The function $\mathcal{N}()$ is defined by (7).

The overall describing function gain is :

$$K_{n,(III)} = \frac{1}{4}\left(\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_2}{\sigma}\right)\right)\frac{1}{\sigma} \quad (27)$$

As expected, this gain lies in between the values of the scenarios without and with 2-level transition elimination given by (9) and (19), respectively.

The variance for the quantization noise σ_q^2 is given by:

$$\sigma_{q,(III)}^2 = E\left[\phi_u^2\right] - K_n^2 E\left[\phi_{e,n}^2\right] \quad (28)$$
$$= \frac{1}{2} - \frac{1}{4}\Phi\left(\frac{d_2}{\sigma}\right) - \left(\frac{1}{4}\left(\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_2}{\sigma}\right)\right)\right)^2 \quad (29)$$

for which the variance of ϕ_u is determined by:

$$E\left[\phi_u^2\right] = \sum_{\phi_u} \text{Prob}\left[\phi_u\right]\phi_u^2 = \text{Prob}\left[|\phi_u| = 1\right] \quad (30)$$
$$= \text{Prob}\left[\text{1-level}\right] + \text{Prob}\left[\text{3-level}\right]$$
$$+ \text{Prob}\left[\text{2-level (right)}\right] \cdot \text{Prob}\left[\phi_{e,n} < -d_2 | \text{2-level (right)}\right]$$
$$+ \text{Prob}\left[\text{2-level (left)}\right] \cdot \text{Prob}\left[\phi_{e,n} > d_2 | \text{2-level (left)}\right] \quad (31)$$

$$= \frac{1}{2} - \frac{1}{4}\Phi\left(\frac{d_2}{\sigma}\right) \quad (32)$$

and where the function $\Phi()$ is given by (8).

In Section III-D, this scenario will also be compared on a system level together with the two other scenarios.

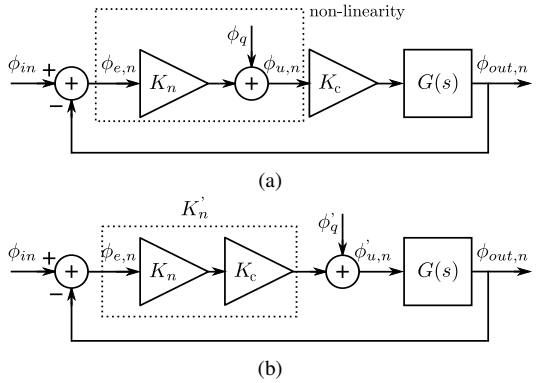


Fig. 7. General phase domain model of a CDR with describing functions and with a gain compensation block K_c. (a) Original model, (b) The compensation gain K_c is combined with the noise gain K_n to create the overall equivalent gain K_n'. Note that the quantization noise ϕ_q' is now also dependent on K_c.

D. Comparison on system level

The conceptual block diagram of a CDR in Fig. 1 is converted to a general phase domain model shown in Fig. 7(a). The combination of the loop filter and the VCO is equivalent to the linear block $G(s)$ [33]:

$$G(s) = \frac{\omega_0}{s}\frac{1+\frac{\omega_z}{s}}{1+\frac{s}{\omega_p}} \tag{33}$$

where ω_z represents the frequency of the zero, ω_p the frequency of the pole and ω_0 the overall amplification factor of the linear block. Note that if ω_0 has a value between ω_z and ω_p, ω_0 also represents the unity gain frequency. Furthermore the non-linear behavior of the phase detector is modeled by the describing function gain K_n and the variance of the quantization noise σ_q^2 that were calculated in the previous section. Due to the different noise gain K_n of the phase detector in the different scenarios, the bandwidth of each CDR system will differ. In order to perform a fair comparison, we should compare CDR systems with the same bandwidth, for this we introduce a compensation gain K_c, which is such that every considered system has the same bandwidth.

By shifting the gain compensation block K_c, we transform Fig. 7(a) to Fig. 7(b). The new equivalent noise gain K_n' and the variance of the equivalent quantization noise term $\sigma_{q'}^2$ of Fig. 7(b) become:

$$K_n' = K_n \cdot K_c \tag{34}$$
$$\sigma_{q'}^2 = \sigma_q^2 \cdot K_c^2 \tag{35}$$

The overall equivalent noise gain K_n' for the CDR systems is set to $\frac{1}{2\sigma}\sqrt{\frac{2}{\pi}}$, which will allow an easy comparison to the CDR systems using 3 thresholds later in the manuscript. The compensation gains K_c are therefore:

$$K_{c,(I)} = \frac{2\sqrt{\frac{2}{\pi}}}{\sqrt{\frac{2}{\pi}} + 2\mathcal{N}\left(\frac{d_2}{\sigma}\right)} \tag{36}$$
$$K_{c,(II)} = 2 \tag{37}$$
$$K_{c,(III)} = \frac{2\sqrt{\frac{2}{\pi}}}{\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_2}{\sigma}\right)} \tag{38}$$

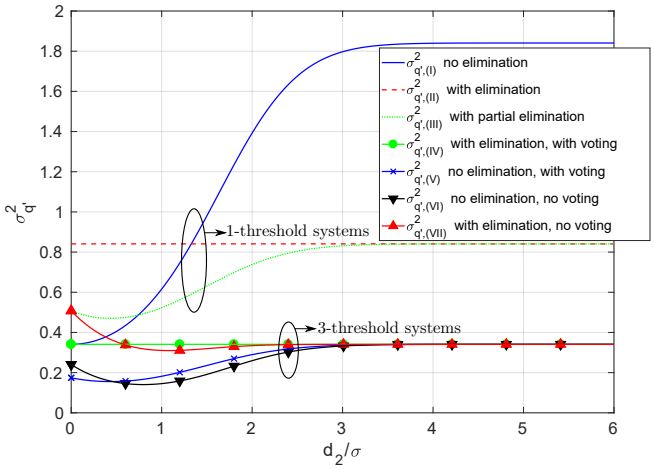


Fig. 8. Comparison of the equivalent quantization noise $\sigma_{q'}^2$ for 1 and 3 thresholds with the overall equivalent gain K_n' set to $\frac{1}{2\sigma}\sqrt{\frac{2}{\pi}}$.

By combining the noise gain K_n and the compensation gain K_c to a fixed equivalent noise gain K_n', we can obtain an equivalent quantization noise term with a variance $\sigma_q'^2$ which is only dependent on one variable, i.e.: $\frac{d_2}{\sigma}$.

Note that, σ, the variance of the jitter at the input of the phase detector $\phi_{e,n}$ comprises two components: a component coming from the jitter in the overall CDR input ϕ_{in}, and a component due to injected quantization noise ϕ_q (which is fed back in the loop). This can be understood by the inspection of Fig. 7(b). With iterative techniques similar to [33], for a given loop filter and input jitter variance, the value of σ can be calculated, but this is not needed for the rest of the discussion and hence is not done in this paper.

Fig. 8 shows the variances of the equivalent injected quantization noise $\sigma_{q'}^2$ for the compensated CDR systems without, with and with partial 2-level transition elimination. It is shown that if the distance between the 2-level crossings and the ideal crossing is significant with respect to the random phase error $(d_2 >> \sigma)$, the injection of jitter due to these 2-level transitions becomes dominant.

Clearly, more quantization noise will lead to a larger variance at the input of the phase detector, reducing the performance of the CDR. A system which uses 2-level transition elimination or partial 2-level transition elimination is thus beneficiary in this case.

However, if d_2/σ is sufficiently small, the penalty of a reduced update-rate due to 2-level transition elimination becomes more severe. That is, a CDR with 2-level transition elimination (and with an equal transfer function) introduces more quantization noise in the system than a CDR system without 2-level transition elimination. In this case, it is better to tolerate the incorrect edge detection but use all transitions to perform a phase update.

Therefore, in contrast to common believe, there is a trade-off between including the excess jitter injection mechanism due to the 2-level transitions and decreasing the robustness of the CDR operation by eliminating the 2-level transitions. This trade-off depends on the ratio of the distance between 2-level crossings and the ideal crossing d_2, and the standard deviation

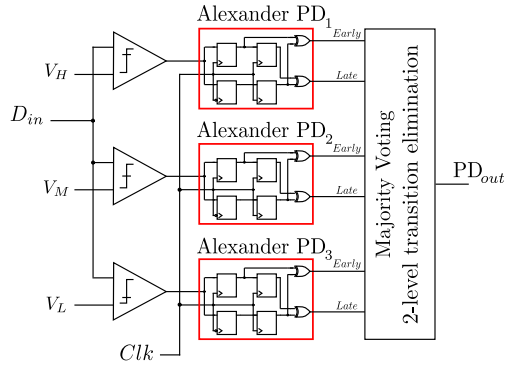


Fig. 9. A conceptual diagram of a typical implementation of a PD with three thresholds for timing adjustment.

of the random phase error σ.

Note that such a trade-off is made by the scenario with partial 2-level transition elimination. It removes only the 2-level transitions which would introduce jitter in the system. As a result it always performs better compared to the scenario with 2-level elimination and for most values of d_2/σ it is also better than the scenario without 2-level transition elimination (only for very small values of d_2/σ, the scenario without 2-level transition elimination injects the lowest amount of jitter).

IV. Phase detection with three thresholds

To increase the robustness of the PAM-4 CDR, three threshold levels can be used in the phase detection. A conceptual diagram of a typical implementation is shown in Fig. 9 [13]. The implementation comprises three comparators each with a corresponding threshold level (V_L, V_M, V_H), and three Alexander phase detectors. The outputs of the Alexander phase detectors are combined by a *majority voting and 2-level transition elimination* block.

The majority voting block selects whether more *early* or more *late* signals are generated by the three Alexander phase detectors. Its output can be *majority early*, *majority late* or *zero*. The zero output occurs if there are no transitions or if the number of incoming *early* signals is equal to the number of *late* signals. Typical implementations eliminate the 2-level transitions of the incoming PAM-4 modulated waveform.

We have shown for the case where timing adjustment is based on one threshold, that if the distance between the 2-level crossings and the ideal crossing is significant with respect to the random phase error ($d_2 >> \sigma$), then this 2-level transition elimination is beneficiary. However, it is a misconception that it is also advantageous to eliminate the 2-level transitions for timing adjustment with three thresholds. This will be discussed in detail below.

Similar to the case of timing adjustment with one threshold, the transitions of PAM-4 data can be decomposed into 3 categories that contain phase information for timing adjustment: *1-level, 2-level* and *3-level transitions*. This is illustrated in Fig. 10 together with all three threshold levels. Additionally, the threshold crossings of the PAM-4 data are depicted by the red dots in Fig. 10. For each 1-level transition, shown in column (c) of Fig. 10, there is one threshold crossing which resides at the ideal location. Secondly, for each 3-level transition, Fig. 10(e), there are three threshold crossings: one at the ideal location and one before and one after the ideal location, which results in three peaks in the threshold crossing distribution. The difference between the ideal crossing and the one before, respectively the one after the ideal crossing are denoted by d_3. Finally, every 2-level transition (Fig. 10(d)) crosses two threshold levels (one before and one after the ideal location), which results in two peaks in the threshold crossing distribution. The difference between the ideal crossing and the one before, respectively the one after the ideal crossing for the 2-level transitions are denoted by d_2.

Summarizing, in total there are five time instances for which a threshold crossing is observed. The ideal threshold crossing instant is at the center between two successive data symbols and occurs for the 3-level and 1-level transitions. Furthermore, there are four side peaks[4] in the threshold crossing distribution shown in the second row of Fig. 10 which originate from the 3-level and 2-level transitions. Fig. 10 also depicts in the third and fourth row, the amount of *Late* and *Early* signals that are generated by the three Alexander PDs. The bottom four traces analyze the response of four different PD topologies. Clearly, the case of a 1-level transition always results in the ideal threshold crossing instant. This results in the ideal bang-bang phase detector characteristic as illustrated in the bottom four traces of Fig. 10, column (c). The 2-level and 3-level transitions for the different scenarios are discussed below.

A. Scenario (IV): Majority voting with 2-level transition elimination

In this scenario, shown in trace (IV) of Fig. 10, the 2-level transitions are not used. For the 3-level transitions, all three Alexander PD blocks (see Fig. 9) will generate an *early/late* signal. Their outputs are combined with majority voting into either *majority early* or *majority late*. This leads to an ideal bang bang characteristic as shown in column (e). The combined effect is the ideal bang bang phase detector characteristic shown in column (a). This scenario is typically used [13]–[22].

B. Scenario (V): Only majority voting

This scenario, shown in trace (V) of Fig. 10, differs from the previous scenario in the sense that the 2-level transitions are also used. In this case, 2 of the Alexander PD blocks (see Fig. 9) will generate an *early/late* signal. There are 3 possible cases: either *two times early* which after voting becomes *majority early*, or *1 early and 1 late* which after voting becomes *zero*, or finally *two times late* which after voting becomes *majority late*. If we look at the resulting PD characteristic for the case of 2-level transitions, see Fig. 10 column (d), we see that it is similar to the case of 2-level transition elimination, except for the case where the phase error is very large (very *late* or very *early*). In this situation the correct majority *early/late* signal is still generated and no error or jitter is introduced. Note that this is in contrast to the

[4]Due to jitter, the threshold crossing distribution peaks broaden and cause the four side peaks to flow together into two wider peaks.

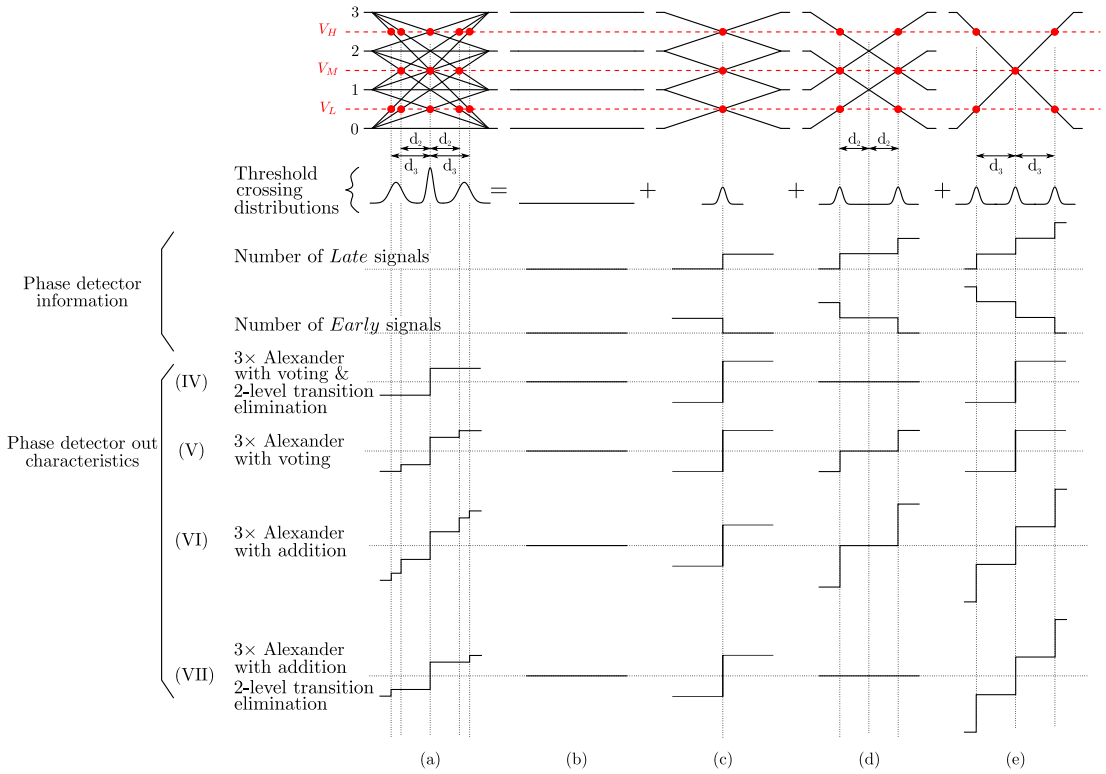


Fig. 10. The decomposition of different transitions of PAM-4 data when considering three transition levels V_L, V_M and V_H. (a) Combination of all possible transitions, (b) zero-level transitions, (c) 1-level transitions, (d) 2-level transitions and (e) 3-level transitions. For each category of transitions, the threshold crossing points are indicated by the red dots and the threshold crossing distributions, number of generated *late* and *early* signals are given. (a) forms the combination of all possible transitions, corresponding threshold crossing points and threshold crossing distributions. The bottom four traces indicate the phase detector output characteristics for different scenarios.

case when only one threshold was used for phase detection (Section II).

If we take the probability[5] of the different transitions into account, this leads to the overall phase detector characteristic shown in trace (V), column (a). By comparing this to the case where the 2-level transition is eliminated, we see that the resulting overall PD characteristic is very similar but slightly improved (it gives a stronger response for large phase errors). Besides, the implementation is simpler because the hardware to eliminate the 2-level transition elimination is no longer required. Therefore this implementation is preferred. However, no reports on this scenario were found in literature.

C. Scenario (VI): No 2-level transition elimination and addition instead of majority voting

A third scenario was implemented in [26]. Here, the outputs of the 3 Alexander PDs (see Fig. 9) are combined by addition instead of by majority voting. This case is represented in trace (VI) in Fig. 10. For the case of 1-level crossings, this leads to the same result as in the previous cases, see column (c). For the case of 2-level crossings, the 3 possible outputs are now: two times *early*, one *early* and one *late* which cancel each other out (=zero) and two times *late*. The resulting characteristic is shown in column (d). And finally for the case of 3-level crossings, the 4 possible net outputs are now: three times *early*, one time *early*, one time *late* and three times *late*. The resulting characteristic is shown in column (e). If we again take the probability of the different transitions into account, this leads to the overall phase detector characteristic shown in column (a). By comparing this to the previous cases, we see that the resulting overall PD characteristic is again very similar.

[5]There are 6 possible *1-level transitions*, 4 possible *2-level transitions* and 2 possible *3-level transitions*

D. Scenario (VII): 2-level transition elimination and addition instead of majority voting

In the last scenario the outputs of the 3 Alexander PDs are again combined by addition instead of by majority voting but the 2-level transitions are not used. This case is represented in trace (VII) in Fig. 10. This case is almost similar to the previous case (with addition instead of majority voting). The only difference is that the input-output characteristic for the 2-level transitions is now equal to zero. The combined input-output characteristic is shown in column (a).

As with the single-threshold scenarios, the rise and fall times can differ from each other or vary in time. Also the threshold levels can deviate from their ideal value. Again a detailed elaborate study of such additional effects is out of the scope of this manuscript, but also here, a preliminary study suggests that our conclusions remain unchanged: i.e. for the case with three thresholds, it is better not to eliminate the 2-level transitions.

V. Pseudo-linearized phase detector gain using three thresholds

To perform a quantitative analysis for the comparison of the phase detectors using three thresholds, we construct a phase domain model that includes the describing functions of the phase detector using three thresholds. For this, we use the same model as in Section III (Fig. 7(a)). The combination of the loop filter and the VCO is again equivalent to the linear block $G(s)$, provided by (33). The non-linear behavior of the phase detector (with 3 thresholds) can be modeled by only using a describing function gain K_n and the variance of the quantization noise σ_q^2. This is because the input-output characteristic for every transition of the phase detector is symmetrical around zero (Fig. 10), which reduces the pseudo-linear analysis to the random-input describing functions without a bias term. The compensation gain K_c that will be used in the comparison on system level (later in the manuscript) in order to compensate the noise gain K_n of the different phase detectors, is also already included in the model.

The noise gain K_n of the different phase detectors with three thresholds can again be decomposed as in (1). Below all describing noise gains K_n and quantization noise sources σ_q^2 for the phase detector topologies using three thresholds for phase alignment are summarized. The full calculations are added in the respective appendix sections.

In the equations below, σ represents the standard deviation of the phase error $\phi_{e,n}$ and, d_2 and d_3 represent the difference between a 2-level transition crossing, respectively a 3-level transition crossing, and the ideal crossing (Fig. 10). The functions $\mathcal{N}()$ and $\Phi()$ are respectively defined by (7) and (8).

A. Scenario (IV): Majority voting with 2-level transition elimination

For the scenario shown in trace (IV) of Fig. 10, the total noise gain K_n is described by:

$$K_{n,(IV)} = \frac{1}{2\sigma}\sqrt{\frac{2}{\pi}} \tag{39}$$

The variance for the quantization noise σ_q^2 is given by:

$$\sigma_{q,(IV)}^2 = \frac{1}{2} - \frac{1}{2\pi} \tag{40}$$

B. Scenario (V): Only majority voting

For the scenario shown in trace (V) of Fig. 10, the total noise gain K_n is described by:

$$K_{n,(V)} = \frac{1}{2\sigma}\left(\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_2}{\sigma}\right)\right) \tag{41}$$

The variance for the quantization noise σ_q^2 is given by:

$$\sigma_{q,(V)}^2 = 1 - \frac{1}{2}\Phi\left(\frac{d_2}{\sigma}\right) - \left(\frac{1}{2}\left(\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_2}{\sigma}\right)\right)\right)^2 \tag{42}$$

C. Scenario (VI): No 2-level transition elimination and addition instead of majority voting

For the scenario shown in trace (VI) of Fig. 10, the total noise gain K_n is described by:

$$K_{n,(VI)} = \frac{1}{2\sigma}\left(\sqrt{\frac{2}{\pi}} + 2\mathcal{N}\left(\frac{d_2}{\sigma}\right) + \mathcal{N}\left(\frac{d_3}{\sigma}\right)\right) \tag{43}$$

The variance for the quantization noise σ_q^2 is given by:

$$\sigma_{q,(VI)}^2 = \frac{9}{2} - 2\Phi\left(\frac{d_2}{\sigma}\right) - 2\Phi\left(\frac{d_3}{\sigma}\right) - \left(\frac{1}{2}\left(\sqrt{\frac{2}{\pi}} + 2\mathcal{N}\left(\frac{d_2}{\sigma}\right) + \mathcal{N}\left(\frac{d_3}{\sigma}\right)\right)\right)^2 \tag{44}$$

D. Scenario (VII): 2-level transition elimination and addition instead of majority voting

For the scenario shown in trace (VII) of Fig. 10, the total noise gain K_n is described by:

$$K_{n,(VII)} = \frac{1}{2\sigma}\left(\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_3}{\sigma}\right)\right) \tag{45}$$

The variance for the quantization noise σ_q^2 is given by:

$$\sigma_{q,(VII)}^2 = \frac{5}{2} - 2\Phi\left(\frac{d_3}{\sigma}\right) - \left(\frac{1}{2}\left(\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_3}{\sigma}\right)\right)\right)^2 \tag{46}$$

E. Comparison on system level

By shifting the gain compensation block K_c, we transform Fig. 7(a) to Fig. 7(b). We obtain an equivalent noise gain K_n' and an equivalent quantization noise term ϕ_q'.

The overall equivalent noise gain K_n' for all CDR systems is set to the gain in scenario (IV), i.e. $\frac{1}{2\sigma}\sqrt{\frac{2}{\pi}}$ for easy comparison.

As we plug in all the above noise gains for the different scenarios given by (39), (41), (43) and (45) in (34), we obtain the following compensation gains K_c:

$$K_{c,(IV)} = 1 \tag{47}$$

$$K_{c,(V)} = \frac{K_{n,(IV)}}{K_{n,(V)}} = \frac{\sqrt{\frac{2}{\pi}}}{\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_2}{\sigma}\right)} \tag{48}$$

$$K_{c,(VI)} = \frac{K_{n,(IV)}}{K_{n,(VI)}} = \frac{\sqrt{\frac{2}{\pi}}}{\sqrt{\frac{2}{\pi}} + 2\mathcal{N}\left(\frac{d_2}{\sigma}\right) + \mathcal{N}\left(\frac{d_3}{\sigma}\right)} \tag{49}$$

$$K_{c,(VII)} = \frac{K_{n,(IV)}}{K_{n,(V)}} = \frac{\sqrt{\frac{2}{\pi}}}{\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_3}{\sigma}\right)} \tag{50}$$

The corresponding equivalent quantization noise sources have a variance $\sigma_q'^2$ that are found by multiplying the original variance with the square of the respective compensation gain.

If we use a linear approximation of the transitions and basic geometry, it can be shown that the relation between d_2 and

d_3 is given by: $\frac{d_2}{d_3} = \frac{3}{4}$. This is used to give a graphical representation of the variance $\sigma_q'^2$ of the corresponding equivalent quantization noise sources as a function of $\frac{d_2}{\sigma}$. This representation was added to Fig. 8.

We can distinguish two different regions for the 3 threshold scenarios in Fig. 8, i.e.: $\frac{d_2}{\sigma} \geq 3$ and $\frac{d_2}{\sigma} < 3$. If $\frac{d_2}{\sigma}$ is larger than 3, the resulting quantization noise become equal, next to the already equal equivalent noise gain K_n'. Therefore if $\frac{d_2}{\sigma}$ is larger than 3, there is absolutely no difference between the systems which have implemented 2-level transition elimination or majority voting and the systems without 2-level transition elimination or majority voting.

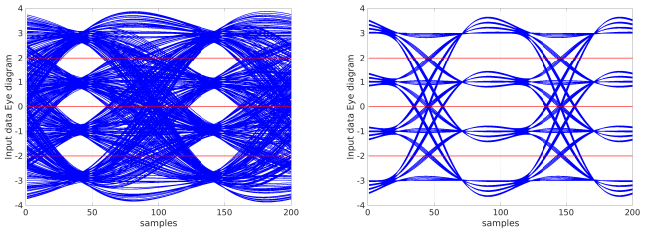
For the other region ($\frac{d_2}{\sigma} < 3$), we can clearly see that the systems with 2-level transition elimination have more jitter injection than the systems without 2-level transition elimination. The use of addition instead of majority voting only provides less jitter injection if $\frac{d_2}{\sigma} > \approx 0.6$. However, the difference is less pronounced than the usage of 2-level transition elimination.

Comparing the results for three thresholds with the results for one threshold (Fig. 8) reveals that the injected jitter $\sigma_q'^2$ for systems with 1 threshold is always bigger than the injected jitter $\sigma_q'^2$ for scenario (IV) with 3 thresholds, i.e. the scenario with elimination and with voting. As a result, a CDR system which uses 3 thresholds is more robust than a system that uses only 1 threshold.

VI. Simulations

We choose to compare the CDR systems in simulations by their jitter tolerance (JTOL) performance, which is the preferred standardized performance metric. For such jitter tolerance simulations, sinusoidal jitter is applied at the input, and for each jitter frequency, the jitter amplitude is increased until we achieve a specified symbol error rate. Although, random Gaussian jitter was used in the previous analyses instead of sinusoidal jitter, several links with the previous analyses can be made. First, at lower jitter tolerance frequencies (in-band frequencies) the sinusoidal component at the input of the CDR will be dominant compared to the injected random quantization term. The describing functions of the phase detector will therefore reduce to the sinusoidal input describing functions. Although the absolute value of the gain will be different, it can be shown that the relation of the upper and lower limits of the noise gain K_n for the different scenarios remains the same. Therefore the maximum and minimum values for the compensation gains K_c for the random Gaussian jitter given by (36)–(38) and (47)–(50) are the same for the sinusoidal input describing functions. Based on the lower JTOL frequencies, we will be able to verify if the systems have the same equivalent gain K_n'. Second, for the higher JTOL frequencies (out-band frequencies), the quantization term becomes dominant and the above analyses with random Gaussian jitter holds true. Here, we can see that an increased injected quantization noise σ_q', will increase the variance of the phase error σ, making it more susceptible to symbol errors.

The previous analyses showed that the equivalent quantization noise σ_q' can be divided into 2 regions, i.e.: a small input channel bandwidth region ($\frac{d_2}{\sigma} \gg 1$) and a large input channel bandwidth region ($\frac{d_2}{\sigma} \ll 1$). Depending on these regions, some topologies introduce more or less quantization noise. Therefore the performance of the all PD implementations (i.e. three phase detectors with 1 threshold and four phase detectors using 3 thresholds) will be verified with two channels. A 50 Gbaud/s PAM-4 modulated signal (with sinusoidal jitter) is filtered by a fourth order Butterworth filter with either a corner frequency of 25 GHz or 50 GHz before it is applied to the input of the CDR instances, each with a different PD implementation. Please note that the analysis, the simulations and conclusions are independent of the data rate, i.e.: the simulations can be executed with a different frequency but by scaling the variables, we would obtain the same results.



(a) (b)

Fig. 11. An eye diagram of a jitterless PAM-4 input signal of the CDR, which is passed through a bandwidth limited channel of (a) 25 GHz and (b) 50 GHz.

Two eye diagrams of a jitterless input signal which is passed through a respectively bandwidth limited channel of 25 GHz and 50 GHz, are shown in Fig. 11(a) and Fig. 11(b). The red horizontal lines indicate the different threshold level, and we can clearly see the multiple phase locations where the input waveform crosses the thresholds due to the different transitions. As illustrated by the figures, a bandwidth limited channel of 25 GHz will results in a case where the distance between the non-centered crossings and the ideal crossing is large w.r.t. the noise variance ($\frac{d_2}{\sigma} \gg 1$), while the other case (with a bandwidth limited channel of 50 GHz) will result in large noise variance compared to the difference between non-centered and ideal crossings ($\frac{d_2}{\sigma} \ll 1$). These two cases will cover both regions in Fig. 8.

An illustration of the simulation model of a CDR system with a phase detector with 3 thresholds, 2-level transition elimination and majority voting is presented in Fig. 12.

The "DataGeneration" block generates a jittery PAM-4 waveform. This signal is passed to three comparators, which slice the input wave at the desired thresholds. Three $Early$-$Late$ signals are generated by three Alexander phase detectors (identical to Fig. 3). These $Early$-$Late$ signals are then processed to provide the desired control signal. For this scenario, the $Early$-$Late$ are first added and then a majority voting operation is performed by a saturation block. To perform the 2-level transition elimination operation, the resulting signal is multiplied by 0 if two out of the three $Early$-$Late$ signals are different from zero. The resulting phase correction signal is then amplified by the compensation gain K_c before it is sent to the loop filter and VCO.

Note that from this testbench all other CDR systems can be derived. To disable the 2-level transition elimination, the output of the "Majority voting" block should be directly

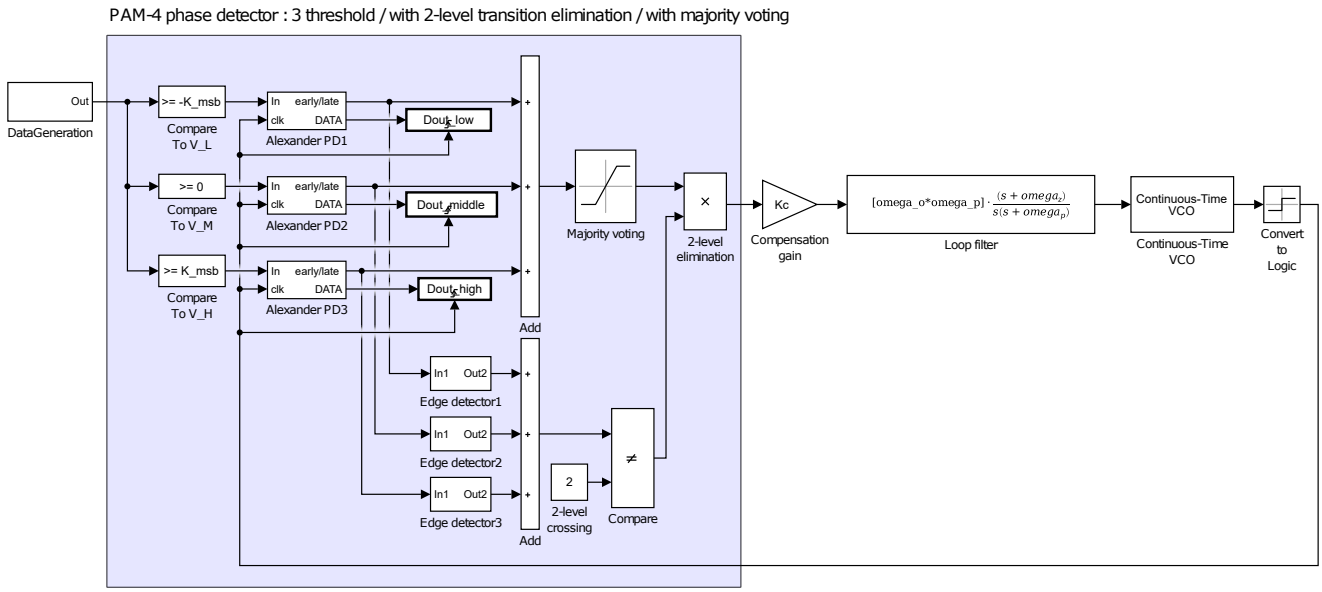


Fig. 12. Testbench of a CDR system with a phase detector with 3 thresholds, 2-level transition elimination and majority voting.

connected to the compensation gain K_c. If addition is used instead of majority voting, the "Majority voting" block should be omitted. Finally, if only 1 threshold is used instead of 3 thresholds, only the middle Alexander phase detector should be connected to the "Majority voting" block instead of the sum of all three Alexander phase detectors. For partial 2-level transition elimination, additional logic needs to be added to determine if a 2-level transition is leading to a left-shifted or right-shifted edge detection (not present on the figure).

To make a fair comparison between all the CDRs, they have the same loop filter and VCO. The loop filter is given by (33) with the following values for ω_z, ω_0 and ω_p respectively : $2\pi \cdot 100$ KHz, $2\pi \cdot 100$ MHz and $2\pi \cdot 10$ GHz. These values are chosen such that the CDRs have a large bandwidth and have sufficient margin to have a stable operation. Multiple batches of simulation runs where executed to test the jitter tolerance of the CDRs, which are discussed in detail in the following sections.

A. 1-threshold CDRs results - High channel bandwidth

First, we perform the two batches of jitter tolerance simulations where we set the channel bandwidth to 50 GHz. For both batches of simulations, we apply PAM-4 data with sinusoidal jitter to each CDR which uses 1 threshold and capture the symbol error rate. For each frequency of the sinusoidal jitter, we increase the jitter amplitude until we achieve a symbol error rate of 10^{-5}.

In the first batch of simulations, we apply the same compensation gain K_c, i.e.: $K_c = 2$ to all CDR systems, which corresponds to the maximum gain provided by (36)–(38). In this case the systems are identical except for the PD.

In the second batch of simulations, we set the compensation gains $K_{c,(I)} = 1$, $K_{c,(II)} = 2$ and $K_{c,(III)} = \frac{4}{3}$. This is because we expect that the ratio $\frac{d_2}{\sigma}$ will be much smaller than 1 due to the large channel bandwidth. Therefore the compensation gain K_c of scenario (I) without 2-level transition elimination given by (36) drops to 1 and the compensation gain K_c of scenario (III) with partial 2-level transition elimination given by (38) drops to 4/3. We expect that all CDR systems now have the same transfer function.

The corresponding jitter tolerance curves when the compensation gains K_c are equal, are shown by Fig. 13(a). It shows at the lower sinusoidal jitter frequencies that the overall gain of the phase detectors is indeed different (as indicated by Section III). At higher sinusoidal jitter frequencies, we can see that all systems coincide, which implies that the systems inject about the same amount of quantization noise. If we reduce the compensation gains K_c of the CDR without elimination and with partial elimination, we can see in Fig. 13(b) that the overall gains of all systems become equal, i.e. the curves coincide at lower sinusoidal jitter frequencies. Additionally, the injected quantization noise of the CDR with partial elimination and without elimination reduces, which results in a higher JTOL at higher sinusoidal jitter frequencies.

Overall, we can conclude that a CDR without 2-level transition elimination (with or without compensation gain) has a better JTOL performance compared to a CDR system with 2-level transition elimination for the case of a broad bandwidth input channel. The difference between the scenarios without 2-level transition elimination and partial 2-level transition elimination is however very small when the corresponding CDR systems are compensated. This matches with our calculations, depicted in Fig. 8.

B. 1-threshold CDRs results - Low channel bandwidth

Secondly, we will test the JTOL performance when we set the channel bandwidth to 25 GHz. We expect now that the ratio $\frac{d_2}{\sigma}$ will be much larger than 1. Therefore, the corresponding compensation gains K_c given by (36)–(38) will all be equal.

A first observation from the simulations with a low input channel bandwidth (and K_c equal to 2) showed that the bandwidth of all the CDR systems increased compared to the cases with a channel bandwidth of 50 GHz. An explanation

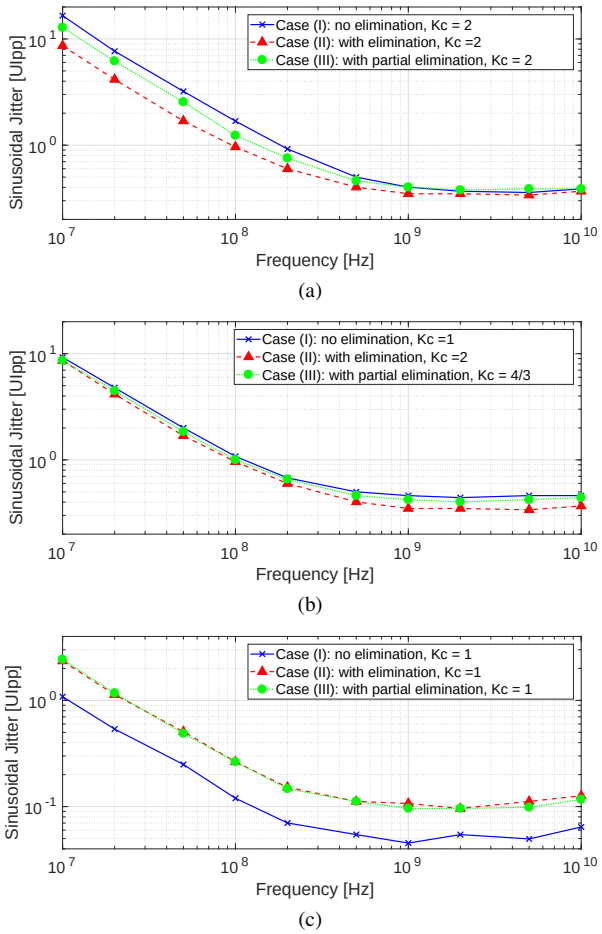


Fig. 13. Simulation results of the JTOL for the different phase detectors with 1 threshold for different values for the channel bandwidth and the compensation gain K_c: (a) has a channel bandwidth of 50 GHz, and $K_c = 2$ for all scenarios, (b) has a channel bandwidth 50 GHz, and K_c is given by the minimum values of (36)–(38) for the different scenarios, (c) has a channel bandwidth of 25 GHz, and $K_c = 1$ for all scenarios. The input data rate is 50 Gbaud/s.

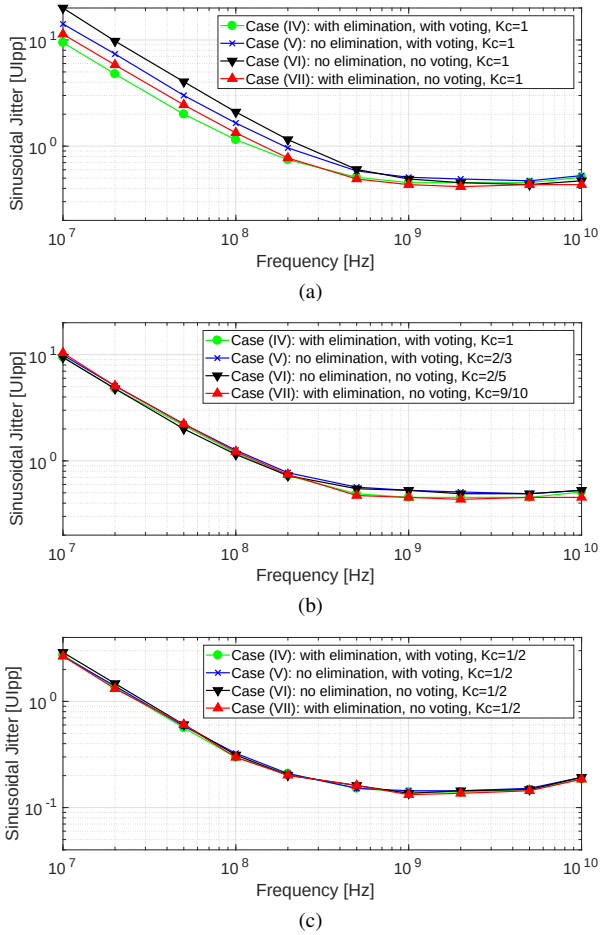


Fig. 14. Simulation results of the JTOL for the different phase detectors with 3 thresholds for different values for the channel bandwidth and the compensation gain K_c: (a) has a channel bandwidth of 50 GHz and $K_c = 1$ for all scenarios, (b) has a channel bandwidth of 50 GHz and K_c is given by (51)–(54) for the different scenarios, (c) has a channel bandwidth of 25 GHz and $K_c = 1/2$ for all scenarios. The input data rate is 50 Gbaud/s.

for this change in CDR bandwidth for the different channels is given by observing the eye diagram in Fig. 11. For a 25 GHz input channel bandwidth, the eye width of the input signal is decreased, making the data recovery much more susceptible to out-band jitter in the CDR (which is always present). Therefore, smaller jitter levels will result in the same bit error rate (e.g.: 1e-5 in our JTOL figures). As shown by (9), (19) and (27), a reduction of the standard deviation of the phase error σ will result in a higher pseudo-linear gain, which in turn increases the CDR bandwidth.

In order to have a similar CDR bandwidth as for the simulation results with a channel bandwidth of 50 GHz, the compensation gain is set to 1 for all scenarios. The resulting JTOL simulation curves are shown in Fig. 13(c). Here, we can see that the scenario without 2-level elimination has a far worse jitter tolerance compared to the other scenarios. This can be explained by observing the large increase in injected quantization noise in Fig. 8. Due to the large injected quantization noise, the performance at low jitter frequencies is degraded and the curves for the different scenarios do not coincide. Note however that for all scenarios the corner frequency of the jitter tolerance curves are equal indicating that the CDR transfer functions are equal. Furthermore, the scenarios with 2-level elimination and with partial 2-level elimination coincide, as expected from Fig. 8.

C. 3-threshold CDRs results - High bandwidth channels

The same simulation are performed for the CDR implementations which use phase detectors with three thresholds. For the first batch of simulations we set the bandwidth of the channel to 50 GHz and we do not include any compensation gain (K_c = 1 for all scenarios). This also corresponds to the maximum value of the compensation gain K_c given by (47)–(50). The systems are identical except for the phase detector.

The results from the jitter tolerance simulation are shown by Fig. 14(a). The CDR systems with a phase detector that use 2-level transition elimination have lower jitter tolerance at the lower sinusoidal jitter frequencies, indicating that the intrinsic gain of the phase detector is lower. Additionally, the use of majority voting also reduces the gain of the phase detector compared to systems which use addition in the phase detector.

From these simulation results, we can fit the corresponding compensation gains K_c, which need to be introduced to make

the bandwidth of the CDR systems equal:

$$K_{c,(IV)} = 1 \tag{51}$$

$$K_{c,(V)} = \frac{K_{n,(IV)}}{K_{n,(V)}} = \frac{2}{3} \tag{52}$$

$$K_{c,(VI)} = \frac{K_{n,(IV)}}{K_{n,(VI)}} = \frac{2}{5} \tag{53}$$

$$K_{c,(VII)} = \frac{K_{n,(IV)}}{K_{n,(V)}} = 0.9 \tag{54}$$

Note that these compensation gains K_c for a sinusoidal input jitter are almost all equal to the lower analytical limits of (47)–(50) derived for Gaussian input jitter.

The JTOL simulations for the compensated 3-threshold phase detectors are shown in Fig. 14(b). As desired, the curves coincide for the lower sinusoidal jitter frequencies, indicating that the transfer functions of all the systems are equal. We can however notice that the systems using 2-level transition elimination have a lower jitter tolerance at higher sinusoidal jitter frequencies due to a higher quantization noise injection (see Fig. 8 for $\frac{d_2}{\sigma} < 3$). Furthermore, the use of majority voting in the phase detector has no impact on the jitter tolerance when the systems have the same transfer function.

The compensated 3-threshold phase detectors of Fig. 14(b), can be compared to the compensated 1-threshold phase detectors shown in Fig. 13(b), because the overall equivalent phase detector gains K_n' are all approximately equal to $\frac{1}{2\sigma}\sqrt{\frac{2}{\pi}}$.

D. 3-threshold CDRs results - Low bandwidth channels

Finally, we perform a batch of simulations to compare the JTOL for CDR systems with 3 thresholds and with a low bandwidth channel.

Identical to the low bandwidth channel case for CDR systems with 1 threshold, we expect that the ratio $\frac{d_2}{\sigma}$ will be much larger than 1. Therefore, the corresponding compensation gains K_c given by (47)–(50) will all be equal. Furthermore, the injected quantization noises (shown in Fig. 8) are also all equal. Therefore we should not see any difference in the JTOL and the simulation results should coincide. Fig. 14(c) confirms our calculations and expectations. Note that similar to the CDR systems with 1 threshold, the compensation gains for all scenarios are now all set to $1/2$ to make the CDR bandwidths for the low bandwidth channels equal to the CDR bandwidths for the high bandwidth channels.

E. Threshold level variations

To further study the validity of our observations, similar simulations as the ones shown in Fig. 14 were performed with up to 10% deviations on the comparator threshold levels. Also in the presence of this non-ideal effect, it was found that all conclusions remain valid: i.e. for the case with three thresholds, it is better not to eliminate the 2-level transitions.

VII. Conclusion

In this work, the phase detection for clock and data recovery for PAM-4 is investigated. For the case where the phase detection is based on 1 threshold levels, we have proven that it is always better to perform *partial 2-level transition elimination* compared to the traditional *2-level transition elimination*. Additionally, it is beneficial to use *(partial) 2-level transition elimination* when the bandwidth of the input channel is small. However, if the bandwidth of the input channel is large, we can achieve a better jitter tolerance if we do not perform *2-level transition elimination*. For the case where the phase detection is based on 3 threshold levels, contrary to the established common practice, we have demonstrated that *2-level transition elimination* is not needed for correct operation of the overall phase detector. If the bandwidth of the input channel is large, it is even better to avoid *2-level transition elimination* to achieve less quantization noise injection and obtain a higher jitter tolerance. Finally, the use of majority voting does not have a significant impact on the jitter tolerance. These results are backed by analytical calculations and computer simulations.

Appendix

A. Calculations: Scenario (IV)

For the scenario shown in trace (IV) of Fig. 10, the sub-describing functions gains are given by:

$$K_{n,\text{1-level}} = K_{n,\text{3-level}} = \frac{1}{\sigma}\sqrt{\frac{2}{\pi}} \tag{55}$$

$$K_{n,\text{2-level}} = 0 \tag{56}$$

The variance for the quantization noise σ_q^2 is given by:

$$\sigma_q^2 = E\left[\phi_u^2\right] - K_n^2 E\left[\phi_{e,n}^2\right] = \frac{1}{2} - \frac{1}{2\pi} \tag{57}$$

B. Calculations: Scenario (V)

For the scenario shown in trace (V) of Fig. 10, the sub-describing functions gains are given by:

$$K_{n,\text{1-level}} = K_{n,\text{3-level}} = \frac{1}{\sigma}\sqrt{\frac{2}{\pi}} \tag{58}$$

$$K_{n,\text{2-level}} = \frac{2}{\sigma}\mathcal{N}\left(\frac{d_2}{\sigma}\right) \tag{59}$$

The variance for the quantization noise σ_q^2 is given by:

$$\sigma_q^2 = E\left[\phi_u^2\right] - K_n^2 E\left[\phi_{e,n}^2\right] \tag{60}$$

$$= E\left[\phi_u^2\right] - \left(\frac{1}{2}\left(\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_2}{\sigma}\right)\right)\right)^2 \tag{61}$$

for which the variance of ϕ_u is determined by:

$$E\left[\phi_u^2\right] = \sum_{\phi_u} \text{Prob}\left[\phi_u\right]\phi_u^2 \tag{62}$$

$$= \text{Prob}\left[\phi_u = 1\right]1^2 + \text{Prob}\left[\phi_u = -1\right](-1)^2 \tag{63}$$

$$= \text{Prob}\left[\text{1-level}\right] + \text{Prob}\left[\text{3-level}\right] + \text{Prob}\left[\text{2-level}\right] \cdot \text{Prob}\left[|\phi_{e,n}| > d_2 | \text{2-level}\right] \tag{64}$$

$$= 1 - \frac{1}{2}\Phi\left(\frac{d_2}{\sigma}\right) \tag{65}$$

C. Calculations: Scenario (VI)

For the scenario shown in trace (VI) of Fig. 10, the sub-describing functions gains are given by:

$$K_{n,\text{1-level}} = \frac{1}{\sigma}\sqrt{\frac{2}{\pi}} \tag{66}$$

$$K_{n,\text{2-level}} = \frac{4}{\sigma}\mathcal{N}\left(\frac{d_2}{\sigma}\right) \tag{67}$$

$$K_{n,\text{3-level}} = \frac{1}{\sigma}\sqrt{\frac{2}{\pi}} + \frac{4}{\sigma}\mathcal{N}\left(\frac{d_3}{\sigma}\right) \tag{68}$$

The variance for the quantization noise σ_q^2 is given by:

$$\sigma_q^2 = E\left[\phi_u^2\right] - K_n^2 E\left[\phi_{e,n}^2\right] \tag{69}$$

$$= E\left[\phi_u^2\right] - \left(\frac{1}{2}\left(\sqrt{\frac{2}{\pi}} + 2\mathcal{N}\left(\frac{d_2}{\sigma}\right) + \mathcal{N}\left(\frac{d_3}{\sigma}\right)\right)\right)^2 \tag{70}$$

for which the variance of ϕ_u is determined by:

$$E\left[\phi_u^2\right] = \sum_{\phi_u} \text{Prob}\left[\phi_u\right]\phi_u^2 \tag{71}$$

$$\begin{aligned} &= \text{Prob}\left[\text{1-level}\right] \\ &\quad + \text{Prob}\left[\text{3-level}\right] \cdot \text{Prob}\left[|\phi_{e,n}| < d_3 | \text{3-level}\right] \\ &\quad + 4 \cdot \text{Prob}\left[\text{2-level}\right] \cdot \text{Prob}\left[|\phi_{e,n}| > d_2 | \text{2-level}\right] \\ &\quad + 9 \cdot \text{Prob}\left[\text{3-level}\right] \cdot \text{Prob}\left[|\phi_{e,n}| > d_3 | \text{3-level}\right] \end{aligned} \tag{72}$$

$$= \frac{9}{2} - 2\Phi\left(\frac{d_2}{\sigma}\right) - 2\Phi\left(\frac{d_3}{\sigma}\right) \tag{73}$$

D. Calculations: Scenario (VII)

For the scenario shown in trace (VII) of Fig. 10, the sub-describing functions gains are given by:

$$K_{n,\text{1-level}} = \frac{1}{\sigma}\sqrt{\frac{2}{\pi}} \tag{74}$$

$$K_{n,\text{2-level}} = 0 \tag{75}$$

$$K_{n,\text{3-level}} = \frac{1}{\sigma}\sqrt{\frac{2}{\pi}} + \frac{4}{\sigma}\mathcal{N}\left(\frac{d_3}{\sigma}\right) \tag{76}$$

The variance for the quantization noise σ_q^2 is given by:

$$\sigma_q^2 = E\left[\phi_u^2\right] - K_n^2 E\left[\phi_{e,n}^2\right] \tag{77}$$

$$= E\left[\phi_u^2\right] - \left(\frac{1}{2}\left(\sqrt{\frac{2}{\pi}} + \mathcal{N}\left(\frac{d_3}{\sigma}\right)\right)\right)^2 \tag{78}$$

for which the variance of ϕ_u is determined by:

$$E\left[\phi_u^2\right] = \frac{5}{2} - 2\Phi\left(\frac{d_3}{\sigma}\right) \tag{79}$$

REFERENCES

[1] J. Lee *et al.*, "Design of 56 Gb/s NRZ and PAM4 SerDes Transceivers in CMOS Technologies," *IEEE J. Solid-State Circuits*, vol. 50, no. 9, pp. 2061–2073, sep 2015.

[2] K. D. Sadeghipour *et al.*, "Design of a sample-and-hold analog front end for a 56Gb/s PAM-4 receiver using 65nm CMOS," in *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, may 2015, pp. 1606–1609.

[3] A. Amirkhany, "Basics of Clock and Data Recovery Circuits: Exploring High-Speed Serial Links," *IEEE Solid-State Circuits Mag.*, vol. 12, no. 1, pp. 25–38, dec 2020.

[4] C. Wang *et al.*, "A 52-Gb/s Sub-1pJ/bit PAM4 Receiver in 40-nm CMOS for Low-Power Interconnects," in *IEEE Symposium on VLSI Circuits, Digest of Technical Papers*, vol. 2019-June, jun 2019, pp. C274–C275.

[5] R. Ma *et al.*, "A 5/10 Gb/s dual-mode NRZ/PAM4 CDR in 65-nm CMOS," in *2019 IEEE International Conference on Electron Devices and Solid-State Circuits, EDSSC 2019*, jun 2019.

[6] A. Roshan-Zamir *et al.*, "A 56-Gb/s PAM4 receiver with low-overhead techniques for threshold and edge-based DFE Fir- and IIR-TAP adaptation in 65-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 54, no. 3, pp. 672–684, mar 2019.

[7] Q. Liao *et al.*, "The Design Techniques for High-Speed PAM4 Clock and Data Recovery," in *Proceedings of 2018 IEEE International Conference on Integrated Circuits, Technologies and Applications, ICTA 2018*, jul 2018, pp. 142–143.

[8] H.-Y. Chen *et al.*, "DC-Balance Low-Jitter Transmission Code for 4-PAM Signaling," *IEEE Trans. Circuits Syst. II*, vol. 53, no. 9, pp. 827–831, sep 2006.

[9] R. Farjad-Rad *et al.*, "A 0.3-μm CMOS 8-Gb/s 4-PAM serial link transceiver," *IEEE J. Solid-State Circuits*, vol. 35, no. 5, pp. 757–764, may 2000.

[10] B. Dehlaghi *et al.*, "A 1.41-pJ/b 56-Gb/s PAM-4 Receiver Using Enhanced Transition Utilization CDR and Genetic Adaptation Algorithms in 7-nm CMOS," *IEEE Solid-State Circuits Letters*, vol. 2, no. 11, pp. 248–251, aug 2019.

[11] F. Lv *et al.*, "Design of 80-Gb/s PAM4 wireline receiver in 65-nm CMOS technology," in *2017 International Conference on Electron Devices and Solid-State Circuits (EDSSC)*. IEEE, oct 2017, pp. 1–2.

[12] Ming-ta Hsieh *et al.*, "SiGe BiCMOS PAM-4 clock and data recovery circuit for high-speed serial communications," in *IEEE International [Systems-on-Chip] SOC Conference, 2003. Proceedings.* IEEE, 2003, pp. 305–308.

[13] J. L. Zerbe *et al.*, "Equalization and clock recovery for a 2.5-10-Gb/s 2-PAM/4-PAM backplane transceiver cell," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2121–2130, dec 2003.

[14] F. Lv *et al.*, "Design of 56 Gb/s PAM4 wire-line receiver with ring VCO based CDR in a 65 nm CMOS technology," in *2017 IEEE 12th International Conference on ASIC (ASICON)*. IEEE, oct 2017, pp. 537–540.

[15] F. A. Musa *et al.*, "Clock recovery in high-speed multilevel serial links," in *Proceedings of the 2003 International Symposium on Circuits and Systems, 2003. ISCAS '03.*, vol. 5. IEEE, 2003, pp. V–V.

[16] F. A. Musa *et al.*, "Modeling and Design of Multilevel Bang–Bang CDRs in the Presence of ISI and Noise," *IEEE Trans. Circuits Syst. I*, vol. 54, no. 10, pp. 2137–2147, oct 2007.

[17] P.-J. Peng *et al.*, "6.1 A 56Gb/s PAM-4/NRZ transceiver in 40nm CMOS," in *2017 IEEE International Solid-State Circuits Conference (ISSCC)*. IEEE, feb 2017, pp. 110–111.

[18] Q. Liao *et al.*, "A Dual-28Gb/s Digital-Assisted Distributed Driver with CDR for Optical-DAC PAM4 Modulation in 40nm CMOS," in *Digest of Papers - IEEE Radio Frequency Integrated Circuits Symposium*, vol. 2019-June, jun 2019, pp. 219–222.

[19] L. Chang *et al.*, "A 50Gb/s-PAM4 CDR with on-chip eye opening monitor for reference-level and clock-sampling adaptation," in *2018 Optical Fiber Communications Conference and Exposition, OFC 2018 - Proceedings*, 2018, pp. 1–3.

[20] Q. Liao *et al.*, "A 50-Gb/s PAM4 Si-Photonic Transmitter With Digital-Assisted Distributed Driver and Integrated CDR in 40-nm CMOS," *IEEE J. Solid-State Circuits*, pp. 1–15, jan 2020.

[21] T. Toifl *et al.*, "A 22-Gb/s PAM-4 Receiver in 90-nm CMOS SOI Technology," *IEEE J. Solid-State Circuits*, vol. 41, no. 4, pp. 954–965, apr 2006.

[22] D.-H. Kwon *et al.*, "A 32-Gb/s PAM-4 Quarter-Rate Clock and Data Recovery Circuit with an Input Slew-rate Tolerant Selective Transition Detector," *IEEE Trans. Circuits Syst. II*, p. 1, 2018.

[23] E. Depaoli *et al.*, "A 64 Gb/s low-power transceiver for short-reach PAM-4 electrical links in 28-nm FDSOI CMOS," *IEEE J. Solid-State Circuits*, vol. 54, no. 1, pp. 6–17, jan 2019.

[24] A. G. Bessios *et al.*, "Transition-limiting codes for 4-PAM signaling in high speed serial links," in *GLOBECOM '03. IEEE Global Telecommunications Conference (IEEE Cat. No.03CH37489)*. IEEE, 2003, pp. 3747–3751.

[25] V. Stojanovic *et al.*, "Autonomous dual-mode (PAM2/4) serial link transceiver with adaptive equalization and data recovery," *IEEE J. Solid-State Circuits*, vol. 40, no. 4, pp. 1012–1026, apr 2005.

[26] L. Tang *et al.*, "A 40 Gb/s 74.9 mW PAM4 receiver with novel clock and data recovery," in *Proceedings - IEEE International Symposium on Circuits and Systems*. IEEE, may 2017, pp. 1–4.
[27] Chih-kong Ken Yang *et al.*, "Analysis of timing recovery for multi-Gbps PAM transceivers," in *Proceedings of the IEEE 2003 Custom Integrated Circuits Conference, 2003.* IEEE, 2003, pp. 67–72.
[28] J. T. Stonick *et al.*, "An adaptive pam-4 5-Gb/s backplane transceiver in 0.25-um CMOS," *IEEE J. Solid-State Circuits*, vol. 38, no. 3, pp. 436–443, mar 2003.
[29] J. Alexander, "Clock recovery from random binary signals," *Electronics Letters*, vol. 11, no. 22, pp. 541–542, 1975.
[30] M. Verbeke *et al.*, "A 1.8-pJ/b, 12.5–25-Gb/s Wide Range All-Digital Clock and Data Recovery Circuit," *IEEE J. Solid-State Circuits*, vol. 53, no. 2, pp. 470–483, feb 2018.
[31] A. Gelb *et al.*, *Multiple-input describing functions and nonlinear system design.* New York: New York : McGraw-Hill, 1968.
[32] M. J. Park *et al.*, "Pseudo-linear analysis of bang-bang controlled timing circuits," *IEEE Trans. Circuits Syst. I*, vol. 60, no. 6, pp. 1381–1394, 2013.
[33] M. Verbeke *et al.*, "Influence of jitter on limit cycles in bang-bang clock and data recovery circuits," *IEEE Trans. Circuits Syst. I*, vol. 62, no. 6, pp. 1463–1471, jun 2015.