Probabilistic Theory for Semi-Blind Oversampling Burst-Mode Clock and Data Recovery Circuits

Bhavin J. Shastri, and David V. Plant
Photonic Systems Group, Dept. of Electrical and Computer Eng., McGill University, Montreal, QC H3A 2A7, Canada
shastri@ieee.org

***Abstract*—We develop a unified probabilistic theory for phase-locked clock and data recovery circuits (CDRs), CDRs based on $N\times$-oversampling techniques in either the time- or space domain, and burst-mode CDRs built from oversampling CDRs. This theory quantitatively explains the performance of these circuits in terms of the bit error rate.**

I. Introduction

Passive optical networks (PONs) are recognized as an economic solution to alleviate the bandwidth bottleneck in access networks by deploying fiber-to-the-home [1]. The challenge in the design of a chip set for PONs arises from the upstream data path as the network is point-to-multipoint. Using time division multiple access, multiple optical network units transmit data to the optical line terminal (OLT) in the central office. Due to optical path differences, packets can vary in amplitude and phase—bursty data. Consequently, this necessitates burst-mode receivers (BMRx) at the OLT. The BMRx front-end is responsible for amplitude recovery, whereas clock and data recovery (CDR) together with fast phase acquisition is performed by a burst-mode CDR (BM-CDR) circuit.

Random noise which is always present at the BMRx front-end affects the determination of the decision threshold and introduces sensitivity penalty. A sensitivity penalty using Gaussian noise statistics for BMRx was first addressed in [2]. A more accurate model is provided in [3], while a unified theory which includes the interaction of Gaussian noise with the finite charging/discharging time of the adaptive threshold detection circuitry is derived in [4]. In [5], the influence of random direct current offsets on the sensitivity of BMRx is analyzed. While there has been an appreciable amount of research on the theory of BMRx front-end circuits in literature, virtually no attention has been paid to the mathematical modeling of BM-CDRs. In this paper, we develop a probabilistic theory for BM-CDRs.

II. Burst-Mode Clock and Data Recovery

Problem: A CDR circuit that is in phase-lock samples (with recovered lock) the incoming data in the center of the data eye. Fig. 1 depicts the nature of bursty traffic in a PON upstream with asynchronous phase steps $|\Delta\varphi| \leq 2\pi$ rad, between the consecutive k^{th} and $(k+1)^{\text{th}}$ packet. This phase step will result in the sampling clock t_{inst}, in-phase with the last bit of the k^{th} packet, to be out-of-phase by $\Delta\varphi$ with the first bit of the $(k+1)^{\text{th}}$ packet. This inevitable presence of phase steps can cause conventional CDRs to lose pattern synchronization. Preamble bits can be inserted at the beginning of each packet

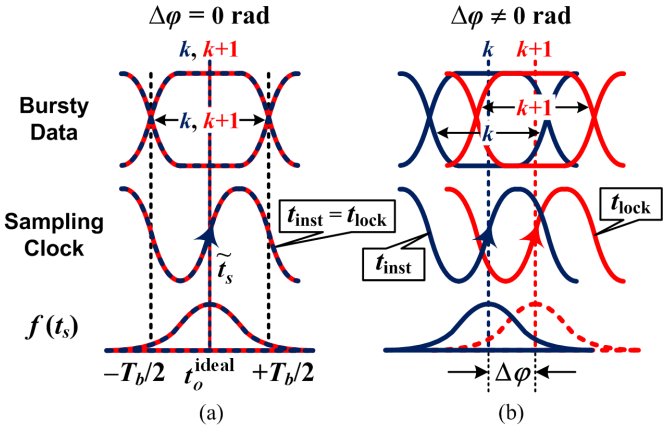


Fig. 1. Graphical depiction of the actual clock sampling point $\widetilde{t}_s$, in the presence of random jitter, and the associated probability density function $f(t_s)$, when the phase difference: (a) $\Delta\varphi = 0$ rad; and (b) $\Delta\varphi \neq 0$ rad.

to allow the CDR feedback loop enough time to settle down and thus acquire lock; that is, align the instantaneous clock t_{inst}, to the lock state t_{lock}, so as to sample in the middle of the data bit. However, the use of a preamble introduces overhead, reducing the effective throughput and increasing delay.

Solution: The most important characteristic of a BM-CDR is its phase acquisition time which must be as short as possible. We define the lock acquisition time as the number of preamble bits l, needed to achieve error-free operation. Fig. 2 shows a block diagram of a BM-CDR architecture that has recently become popular [1], [6], [7]. The BM-CDR is based on a phase-tracking $N\times$-oversampling CDR and a clock phase aligner/picker (CPA) that makes use of a phase picking algorithm. The hybrid combination of phase-tracking and a blind oversampling CDR is referred to as a semi-blind oversampling CDR [10]. One can either oversample in the time domain using a clock frequency $N\times$ the bit rate, or oversample in the space domain using N multi-phase clocks with a frequency equal to the bit rate. Oversampling in time results in N data samples, whereas oversampling in space results in N clock samples. The end result is that the N (data or clock) samples are forwarded to the CPA for picking the correct phase. The phase picking algorithm for time oversampling is based on selecting the best data sample by relying on a simple comparison with a known pattern. This technique requires faster electronics. On the other hand, the phase picking algorithm for space oversampling is based on a more complex algorithm which tries to determine the clock sample closest to the center

978-1-4244-7772-2/10/$26.00 ©2010 IEEE

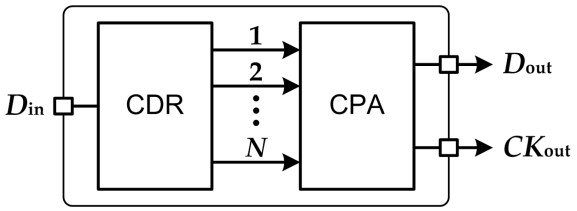


Fig. 2. BM-CDR architecture based on an oversampling CDR and CPA.

of the data eye. This technique requires low skew between multiple phases of the clock. In either case, the phase picking algorithm guarantees at least one data sample or clock edge that will yield uncorrupted data regardless of any phase step $|\Delta\varphi| \leq 2\pi$ rad between the consecutive packets. The phase picker then uses a feedback mechanism to select the correct sample from the N possibilities. It has been experimentally demonstrated in [1], [6], [7] that this BM-CDR architecture achieves instantaneous phase acquisition ($l = 0$ preamble bits) with error-free operation for any phase step $|\Delta\varphi| \leq 2\pi$ rad. In this paper, we develop a unified probabilistic theory for BM-CDRs built from $N\times$ oversampling CDRs in either the time- or space domain.

III. Theoretical Modeling

The probabilistic theory developed here is for data transmitted in the non-return-to-zero format, and it is independent of the bit rate and pulse shape, as long as the intersymbol interference (ISI) at the sampling point is negligible. This will remain valid at high bit rates, as long as the channel remains limited by Gaussian noise [2].

Jitter can be interpreted as the perturbations of the threshold-crossing time of data transitions from their ideal position in time. A part of the jitter of the data is inherited as phase uncertainty of the recovered sampling clock in the clock recovery circuit. As a result, the regenerated (retimed) data sequence by the CDR may be erroneous, degrading the bit error rate (BER) performance. Jitter is in general classified as being either random or deterministic. Random jitter (RJ) is unpredictable, unbounded, and results from physical noise sources based on random processes. RJ is attributed to thermal noise, shot noise, and flicker noise. The generation of RJ is approximated to a Gaussian probability distribution. This follows from the central limit theorem which states that the composite effect of many uncorrelated noise sources, regardless of the distributions, approaches a Gaussian distribution. The Gaussian approximation is sufficiently accurate for design purposes and is far easier to evaluate than the more exact probability distribution within the receiver [8]. RJ is characterized by the root-mean-square (RMS) value of the Gaussian probability distribution. Deterministic jitter (DJ) is predictable, bounded, and is attributed to duty cycle distortions. DJ is classified as ISI and data-dependent jitter, pulse-width-distortion jitter, sinusoidal jitter, and uncorrelated bounded jitter. The effect of DJ is to shrink the data eye by a finite amount and will only further deteriorate a device under test's performance. Thus, in order to simplify the mathematical modeling, DJ is ignored.

In deriving the theoretical probabilistic model, we make use of continuous random variables $\widetilde{x}$, that follow a Gaussian distribution denoted as $\widetilde{x} \sim N(\mu, \sigma^2)$, where μ is the mean, $\sigma > 0$ is the standard deviation, and the probability density function (PDF) of $\widetilde{x}$, is given by $f(x) = \left(1/\sqrt{2\pi}\cdot\sigma\right) \cdot \exp\left[-(x-\mu)^2/2\sigma^2\right]$, $x \in \mathbb{R}$, with the following characteristics: $f(x) > 0$, for all x and $\int_{-\infty}^{+\infty} f(x)\cdot dx = 1$. In the context of clock and data recovery, we define the following continuous random variables with a Gaussian distribution:

- $\widetilde{\xi} \sim N(0, \sigma_{t_s}^2)$, with PDF $f(\xi)$, is the jitter on the edges of the data bits with a zero mean, where σ_{t_s} corresponds to the RMS jitter on the sampling clock signal;
- $\widetilde{t_s} \sim N(\widetilde{t_o}, \sigma_{t_s}^2)$, with PDF $f(t_s)$, is the *actual* clock sampling point in the presence of random jitter; and
- $\widetilde{t_o} \sim N(t_o^{\text{ideal}}, \sigma_{t_o}^2)$, with PDF $f(t_o)$, is the clock sampling point *determined* by the CDR, where t_o^{ideal} is the *ideal* clock sampling point in the middle of the data bit, and $\sigma_{t_o}^2 = \kappa\cdot\sigma_{t_s}^2$, with κ being a constant of proportionality.

For convenience, the left and right edges of the data eye are located at $-T_b/2$ and $+T_b/2$, respectively [see Fig. 1(a)]. Thus, the expectation of the clock sampling point is given by

$$E\left[\widetilde{t_o}\right] \triangleq \int_{-\infty}^{+\infty} t_o \cdot f(t_o)\cdot dt_o = t_o^{\text{ideal}} = 0, \tag{1}$$

as the ideal clock sampling point is in the center of the data bit. Let $\widetilde{\xi}_j^{\text{left}}$ and $\widetilde{\xi}_j^{\text{right}}$ be the jitter on the left edge and right edge of the j^{th} bit of an l-bit preamble. We assume that $\widetilde{\xi}_j^{\text{left}}$ and $\widetilde{\xi}_j^{\text{right}}$ are independent with common RMS jitter σ_{t_s}. Then the mid-point of the j^{th} bit $\widetilde{\tau}_j$, is expressed as

$$\widetilde{\tau}_j = \frac{\widetilde{\xi}_j^{\text{left}} + \widetilde{\xi}_j^{\text{right}}}{2}. \tag{2}$$

After the l-bit preamble, the clock sampling point determined by the CDR $\widetilde{t_o}$, at the first bit where the decision circuit will start sampling the data bits, is given by the average of the individual mid-points $\widetilde{\tau}_j$, in (2) as

$$\widetilde{t_o} = \frac{1}{(l+1)}\cdot\sum_{j=1}^{l+1}\widetilde{\tau}_j. \tag{3}$$

Thus, σ_{t_o} can be related to sampling clock RMS jitter σ_{t_s}, as

$$\sigma_{t_o}^2 \triangleq E\left[\left(\widetilde{t_o} - \underbrace{E\left[\widetilde{t_o}\right]}_{=0}\right)^2\right] = \underbrace{\frac{1}{2(l+1)}}_{=\kappa}\cdot\sigma_{t_s}^2. \tag{4}$$

Hence, the PDFs $f(t_s)$ (actual sampling point) and $f(t_o)$ (sampling point determined by CDR), can be expressed as:

$$f(t_s) = \frac{1}{\sqrt{2\pi}\cdot\sigma_{t_s}}\cdot\exp\left(-\frac{\left(t_s-\widetilde{t_o}\right)^2}{2\sigma_{t_s}^2}\right); \tag{5}$$

$$f(t_o) = \frac{1}{\sigma_{t_s}}\cdot\sqrt{\frac{(l+1)}{\pi}}\cdot\exp\left(-\frac{(l+1)\cdot t_o^2}{\sigma_{t_s}^2}\right). \tag{6}$$

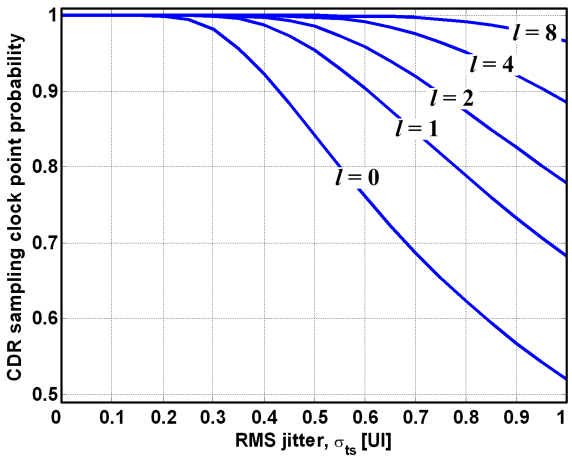


Fig. 3. Probability of the clock sampling point determined by the CDR $\widetilde{t_o}$, to be within the data bit after an l-bit preamble.

The probability that the sampling point determined by CDR $\widetilde{t_o}$, will be within the data bit after l preamble bits is given by

$$\Pr\left(|\widetilde{t_o}| < \frac{T_b}{2}\right) = \int_{-T_b/2}^{+T_b/2} f(t_o)\cdot dt_o = 1 - 2Q\left(\frac{1}{\sigma_{t_s}[\text{UI}]}\cdot\sqrt{\frac{(l+1)}{2}}\right) \quad (7)$$

where $Q(x) \triangleq \left(1/\sqrt{2\pi}\right)\int_x^\infty \exp\left(-\lambda^2/2\right)d\lambda$ is the normalized Gaussian tail probability. Note that (7) has been made independent of the data rate; thus, the RMS jitter σ_{t_s}, is expressed in terms of the unit interval (UI). In Fig. 3 we plot (7) as a function of σ_{t_s} for different l. The probability $\Pr\left(|\widetilde{t_o}| < T_b/2\right)$, decreases with increasing jitter but can be compensated by increasing the preamble length. Also, for $\sigma_{t_s} \leq 0.25$ UI, $\Pr\left(|\widetilde{t_o}| < T_b/2\right) \sim 1$ with no preamble bits.

When there is no phase difference $\Delta\varphi = 0$ rad, between two consecutive packets in a PON uplink [see Fig. 1(a)], the CDR's sampling error probability is equivalent to the probability that the clock transition occurs either before the leading data transition or after the trailing data transition, $\Pr\left(|\widetilde{t_s}| > T_b/2\right)$, given that the sampling point determined by the CDR $\widetilde{t_o}$, is within the data eye. Assuming uncorrelated data with equiprobable ONEs and ZEROs, the sampling error probability P_s, of the CDR can be expressed as

$$P_s = \frac{1}{2}\cdot\Pr\left(|\widetilde{t_o}| < \frac{T_b}{2}\right)\cdot\Pr\left(|\widetilde{t_s}| \geq \frac{T_b}{2}\right), \text{ and} \quad (8)$$

$$\Pr\left(|\widetilde{t_s}| \geq \frac{T_b}{2}\right) = \int_{-\infty}^{-T_b/2} f(t_s)\cdot dt_s + \int_{+T_b/2}^{+\infty} f(t_s)\cdot dt_s. \quad (9)$$

Ideally, the sampling clock must bear a well-defined phase relationship with respect to the received data so that the decision circuit samples each bit at the mid-point of the data eye. Thus, it is desirable that the CDR sampling point be as close as possible to the ideal sampling point, $\widetilde{t_o} \sim t_o^{\text{ideal}} = 0$. Also, since the PDF $f(t_s)$, is even-symmetric, then $\Pr\left(\widetilde{t_s} < -T_b/2\right) = \Pr\left(\widetilde{t_s} > +T_b/2\right)$, and the sampling error probability is given as

$$P_s = Q\left(\frac{T_b}{2\sigma_{t_s}}\right). \quad (10)$$

With a finite phase difference $\Delta\varphi \neq 0$ rad, between the consecutive packets [see Fig. 1(b)], the phase step has the effect of displacing the instantaneous CDR sampling clock t_{inst}, by $|\Delta\varphi|\cdot(T_b/2\pi)$. By inserting preamble bits, the CDR feedback loop will have time to settle down. Specifically, after an l-bit preamble, the sampling point determined by the CDR $\widetilde{t_o}$, will be displaced by $t_{|\Delta\varphi|} = |\Delta\varphi|\cdot\left(1-\eta(l)\right)\cdot(T_b/2\pi)$, where $\eta(l)$, is the CDR feedback loop function analytically derived for a second-order PLL to be [9]

$$\eta(l) = 1 - \exp\left(-l\cdot\zeta\cdot\omega_n\cdot T_b\right)\times\left\{\cosh\left(l\cdot\omega_n\cdot T_b\cdot\sqrt{\zeta^2-1}\right) - \frac{\zeta}{\sqrt{\zeta^2-1}}\cdot\sinh\left(l\cdot\omega_n\cdot T_b\cdot\sqrt{\zeta^2-1}\right)\right\}, \zeta > 0 \quad (11)$$

where ζ is the "damping ratio" and ω_n in [rad/s] is the "natural frequency", both dependent on CDR circuit parameters. Note that the expression for $t_{|\Delta\varphi|}$ is only valid for phase steps $|\Delta\varphi| \leq \pi$ rad, and does not account for $\pi < |\Delta\varphi| \leq 2\pi$ rad. Thus, a correcting factor ψ, must be introduced to account for the symmetrical performance about the edges of the data bit such that $(\psi, \Delta\varphi) \in \left\{\left(0, [-\pi, +\pi]\right), \left(2\pi, [\pm\pi, \pm 2\pi]\right)\right\}$; hence,

$$t_{|\Delta\varphi|} = \left[\left(|\Delta\varphi| - \psi\right)\cdot\left(1-\eta(l)\right)\right]\cdot\frac{T_b}{2\pi}. \quad (12)$$

It follows from (12), that the PDF $f(t_s)$ in (5), can therefore be modified to account for this phase step as

$$f(t_s) = \frac{1}{\sqrt{2\pi}\cdot\sigma_s}\cdot\exp\left(-\frac{\left(t_s - \widetilde{t_o} \mp t_{|\Delta\varphi|}\right)^2}{2\sigma_{t_s}^2}\right) \quad (13)$$

Subsequently, the probability that the clock transition occurs either before the leading data transition or after the trailing data transition can then be expressed as

$$\Pr\left(|\widetilde{t_s}| \geq \frac{T_b}{2}\right) = \frac{1}{2}\cdot\left\{Q\left(\frac{\widetilde{t_x} - t_{|\Delta\varphi|}}{\sigma_{t_s}}\right) + Q\left(\frac{\widetilde{t_x} + t_{|\Delta\varphi|}}{\sigma_{t_s}}\right)\right\} \quad (14)$$

where $\widetilde{t_x} = T_b/2 - \widetilde{t_o}$. Before we proceed, we make two assumptions: (1) the sampling point determined by the CDR is ideally located at the center of the data eye ($\widetilde{t_o} = 0$) *before* a phase step $|\Delta\varphi|$; and (2) the RMS jitter on the clock signal $\sigma_{t_s} \leq 0.25$ UI, implying the probability that the CDR clock sampling point is within the data eye *after* the phase step is $\Pr\left(|\widetilde{t_o}| < T_b/2\right) \sim 1$, for any number of preamble bits l. Consequently, for a given phase step $|\Delta\varphi| \leq 2\pi$ rad, the sampling error probability P_s, in (8) can be expressed as

$$P_s\left(|\Delta\varphi|\right) = \frac{1}{2}\cdot\left\{Q\left(\frac{\pi - \left(|\Delta\varphi| - \psi\right)\cdot\left(1-\eta(l)\right)}{2\pi\cdot\sigma_{t_s}[\text{UI}]}\right) + Q\left(\frac{\pi + \left(|\Delta\varphi| - \psi\right)\cdot\left(1-\eta(l)\right)}{2\pi\cdot\sigma_{t_s}[\text{UI}]}\right)\right\}. \quad (15)$$

For a CDR that is based on an $N\times$-oversampling architecture in either time or space, the absolute value of the maximum phase difference between the ideal sampling point and the sampling point determined by the CDR, is $\max\big(|t_o^{\text{ideal}}-\widetilde{t_o}|\big) = T_b/2N \equiv \pi/N$ [rad]. For $t_o^{\text{ideal}} = 0$, the N-clock sampling points determined by the CDR $t_o^n|_N$, are located at:

$$\widetilde{t_o} \in \Big\{t_o^n|_N\Big\} = \Big\{\frac{\pi}{N}(2n+1-N)\Big\}, n = 0, 1, \ldots, N-1. \quad (16)$$

For each of the N data samples, the sampling error probabilities $P_s^n|_N$, can be calculated by convolving $P_s\big(|\Delta\varphi|\big)$ in (15), with the N-sampling points $t_o^n|_N$ in (16), as

$$P_s^n|_N = P_s\big(|\Delta\varphi|\big) \otimes \delta\big(|\Delta\varphi| - t_o^n|_N\big), \text{ and} \quad (17)$$

$$\delta\big(|\Delta\varphi| - t_o^n|_N\big) \triangleq \begin{cases} 1 & \text{if } |\Delta\varphi| = t_o^n|_N, \\ 0 & \text{if } |\Delta\varphi| \neq t_o^n|_N. \end{cases} \quad (18)$$

is the Dirac-delta function. It follows from the sifting property

$$\begin{aligned} P_s^n|_N &= \int_{-\infty}^{+\infty} P_s\big(|\Delta\varphi| - \lambda\big) \cdot \delta\big(\lambda - t_o^n|_N\big) \cdot d\lambda \\ &= P_s\big(|\Delta\varphi| - t_o^n|_N\big). \end{aligned} \quad (19)$$

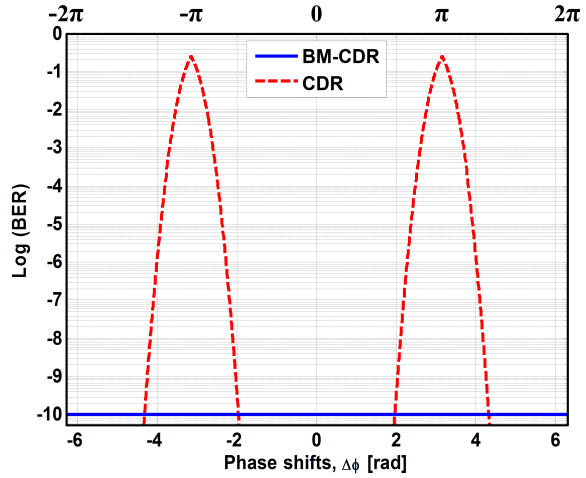
For a BM-CDR based on the $N\times$-oversampling CDR and a CPA which selects the correct set of samples with the aid of a phase picking algorithm, the sampling error probability $P_s^{\text{BM-CDR}}$, is expressed as

$$P_s^{\text{BM-CDR}} = \min\Big\{P_s\big(|\Delta\varphi| - t_o^n|_N\big)\Big\} \quad (20)$$

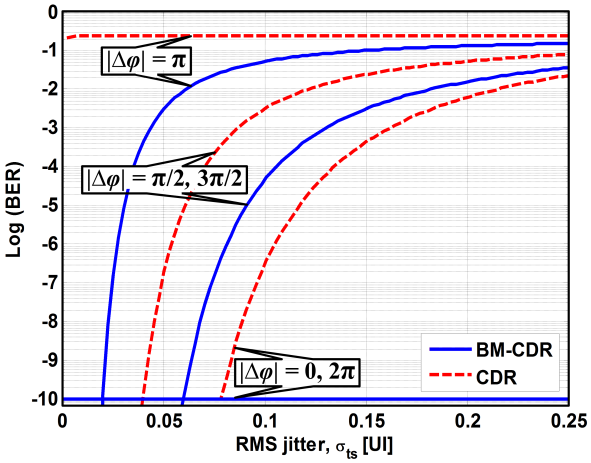
We define the BER, denoted as P_e, of the CDR, $N\times$-oversampling CDR, and BM-CDR, from the sampling error probabilities in (15), (19), and (20), as follows:

$$\text{BER} \equiv P_e \triangleq \begin{cases} P_s\big(|\Delta\varphi|\big) & \text{for CDR,} \\ \Big\{P_s\big(|\Delta\varphi| - t_o^n|_N\big)\Big\} & \text{for } N\times\text{-CDR,} \\ \min\Big\{P_s\big(|\Delta\varphi| - t_o^n|_N\big)\Big\} & \text{for BM-CDR.} \end{cases} \quad (21)$$

Fig. 4(a) shows the BER performance of the CDR and BM-CDR as a function of phase step for a zero preamble length ($l = 0$). As expected the worst-case phase steps for the CDR are $\pm\pi$ rad because these represent the half-bit periods, and therefore the CDR is sampling exactly at the edge of the data eye, resulting in a BER ~ 0.5. At phase shifts near 0 or 2π rad, we can easily achieve error-free operation, BER $< 10^{-10}$, because the CDR is almost sampling at the middle of each data bit. For the BM-CDR we achieve error-free operation, for any phase step $|\Delta\varphi| \leq 2\pi$ rad. Similar results have been obtained experimentally in [1], [6], [7], clearly validating our probabilistic theoretical model. In Fig. 4(b) we plot the BER performance of the CDR and BM-CDR as a function of the RMS jitter for different phase steps and zero preamble bits. As anticipated, for a given BER and phase step, the allowable RMS jitter on the sampling clock is higher with the BM-CDR than the CDR in each case. More importantly, it can be perceived that the BM-CDR achieves far superior BERs for any given phase step and RMS jitter.



(a)



(b)

Fig. 4. BER performance of the CDR and BM-CDR (for zero preamble length) versus: (a) phase step; and (b) sampling clock RMS jitter.

IV. Conclusion

We have developed a unified probabilistic theory for conventional CDRs, $N\times$-oversampling CDRs (in time or space), and BM-CDRs built from oversampling CDRs. The theoretical model quantitatively explains the performance of these circuits in terms of the BER by taking into account the phase steps between successive packets, preamble length, and jitter on the sampling clock. This model will help refine theoretical models of PONs and provide input for establishing realistic power budgets.

References

[1] B. J. Shastri, et. al., *Analog Integr. Circuit and Signal Process.*, vol. 60, no. 1-2, 2009.
[2] C. A. Eldering, *J. Lightw. Technol.*, vol. 11, no. 12, 1993.
[3] P. Menendez-Valdes, *J. Lightw. Technol.*, vol. 13, no. 11, 1995.
[4] C. Su, et. al., *J. Lightw. Technol.*, vol. 15, no. 4, 1997.
[5] P. Ossieur, et. al., *IEEE J. Lightw. Technol.*, vol. 24, no. 3, 2006.
[6] B. J. Shastri, et. al., *J. Opt. Commun. and Netw.*, vol. 2, no. 1, 2010.
[7] B. J. Shastri, et. al., *IEEE J. Sel. Topic Quantum Electron.*, vol. 16, no. 5, 2010, to appear.
[8] P. P. Webb, et. al., *RCA Review*, vol. 35, no. 2, 1974.
[9] F. M. Gardner, *Phaselock Techniques*, 2nd ed. New York: Wiley, 1979.
[10] M. van Ierssel, et. al., *IEEE J. Solid-State Circuits*, vol. 42, no. 10, 2007.