Equalization of multi-Gb/s chip-to-chip interconnects affected by manufacturing tolerances

Jelle Bailleul[a,*], Lennert Jacobs[a], Paolo Manfredi[b], Dries Vande Ginste[b], Marc Moeneclaey[a]

[a]*Department of Telecommunications and Information Processing*
Ghent University, Sint-Pietersnieuwsstraat 41, 9000 Gent, Belgium
[b]*Department of Information Technology*
Ghent University, Tech Lane Ghent Science Park – Campus A iGent, 9052 Gent, Belgium

Abstract

Electrical chip-to-chip interconnects suffer from considerable intersymbol interference at multi-Gb/s data rates, due to the frequency-dependent attenuation. Hence, reliable communication at high data rates requires equalization, to compensate for the channel response. As these interconnects are prone to manufacturing tolerances, the equalizer must be adjusted to each specific channel realization to perform optimally. We adopt a reduced-complexity equalization scheme where (part of) the equalizer is fixed, by involving the channel statistics into the equalizer derivation. For a 10 cm on-board microstrip interconnect with a 10% tolerance on its parameters, we point out that 2-PAM transmission using a fixed prefilter and an adjustable feedback filter, both with few taps, yields only a moderate bit error rate degradation, compared to the all-adjustable equalizer; at a bit error rate of 10^{-12}, these degradations are about 1.1 dB and 3 dB, when operating at 20 Gb/s and 80 Gb/s, respectively.

Keywords: MMSE Equalization, Chip-to-Chip Communication, Manufacturing Tolerances, Polynomial Chaos Theory.

*Corresponding author

Email addresses: `jelle.bailleul@ugent.be` (Jelle Bailleul), `lennert.jacobs@ugent.be` (Lennert Jacobs), `paolo.manfredi@ugent.be` (Paolo Manfredi), `dries.vande.ginste@ugent.be` (Dries Vande Ginste), `marc.moeneclaey@ugent.be` (Marc Moeneclaey)

1. Introduction

As predicted by Moore's law in 1965, the number of transistors on integrated circuits doubles roughly every two years. In order to exploit this drastic increase in on-chip processing power, the communication speed on the interconnection between chips should grow at the same pace, otherwise it becomes the system bottleneck. Chip-to-chip communication commonly makes use of electrical interconnects consisting of microstrips or striplines of several tens of cm of length, on a printed circuit board (PCB). Examples of chip-to-chip interconnects can be found in processor-to-memory interfaces [1], and in the multi-layer backplanes of server/router systems and multi-processor systems.

In order to achieve higher bitrates, higher signal bandwidths are required on the chip-to-chip interconnection. However, the electrical interconnects suffer from attenuation caused by skin effect and dielectric loss, which increase with the signal frequency. These frequency-dependent losses turn the communication channel into a low-pass channel with a high-frequency attenuation that increases with distance [2], causing considerable inter-symbol interference (ISI) at the input of the receiver (RX).

Currently, the main research focus in signal processing for electrical chip-to-chip interconnects is on reducing the ISI by compensating the frequency-dependent channel characteristic, using equalization. Data rates near 10 Gb/s per line are currently achieved using decision-feedback equalization (DFE) with a power consumption of about 1 mW/Gb/s [3–7]. Recent research (e.g., [7–10]) deals with increasing this speed up to 25 Gb/s per line, whereas speeds up to 100 Gb/s per line are targeted in the (near) future [11].

As the production of chip-to-chip interconnects is prone to manufacturing tolerances, the different realizations of the same interconnect are not identical. Ideally, the equalizer coefficients should be adjusted according to the specific realization. In principle, this could be achieved by measuring for each realization the corresponding frequency response of the interconnect, and adjusting the equalizer according to the measurement. A major disadvantages of this approach is the need for (time-consuming) measurements. These measurements could be avoided by using an adaptive equalizer, whose setting is auto-

matically adjusted according to the specific realization; however, this approach increases the implementation complexity because of the required adaptation circuit.

In this contribution we adopt a different approach, by keeping (part of) the equalizer fixed. The fixed (part of the) equalizer is set during the production process, and is determined based on the statistical properties of the channel. More specifically, the considered equalizer consists of a tapped delay-line prefilter (PRF) at the transmitter (TX) (which is feasible for multi-Gb/s data transmission [12]), and a feedback filter (FBF) operating on past symbol decisions at the RX. We investigate three equalization strategies: (i) the reference case, where both the PRF and the FBF are adjustable; (ii) the low-complexity approach, where both the PRF and the FBF are fixed; (iii) an intermediate approach, where the PRF is fixed but the FBF is adjustable. While strategy (i) constitutes the standard approach from literature when dealing with channel variability, the strategies (ii) and (iii) have, to the best of our knowledge, been investigated for the first time in [13]. The present paper extends the results from [13] in the following ways: a more comprehensive derivation of the optimum taps of the PRF and FBF is presented, the tap spacing of the PRF is not restricted to the symbol interval, the sampling delay at the RX is optimized, and the performance is evaluated for two levels of channel variability.

The organization of the paper is as follows. A generic baseband communication system, consisting of a PRF and a DFE, is described in section 2. In section 3 we determine the filter coefficients of the equalizer which minimize the mean-squared error (MSE) between the equalizer output and the symbol to be detected, under a constraint on the transmit energy per symbol. While the minimization of the MSE is standard for the reference case, the situation is different when (part of) the equalizer is fixed; in the latter case, we have to incorporate the statistical properties of the channel into the MSE to be minimized. The bit error rate (BER) performance of the equalized communication system is derived in section 4, for 2-PAM transmission. Numerical results regarding the minimal mean-squared error (MMSE) and BER performances are presented in section 5, for a chip-to-chip interconnect, with manufacturing tolerances on its parameters. The channel

frequency responses corresponding to the different realizations of these parameters are generated by making use of polynomial chaos theory, as outlined in [14–16]. Conclusions are drawn in section 6. One of the main conclusions is that the equalizer with fixed PRF and adjustable FBF has only a small degradation compared to the optimal fully adjustable equalizer, even for filters with few taps.
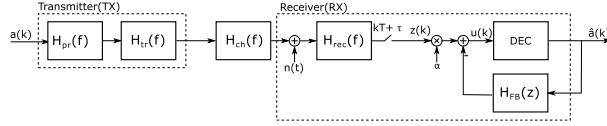
2. System Description

Figure 1: Communication system with equalization.

Let us consider the communication system displayed in Fig. 1 representing an equalized chip-to-chip interconnect. The input to the TX consists of a stream of i.i.d. equiprobable zero-mean data symbols $a(k)$ belonging to a real-valued constellation, with $\mathbb{E}[a^2(k)] = \sigma_a^2$. First the TX applies the data symbols at the symbol rate $1/T$ to a continuous-time tapped delay-line $H_{\mathrm{pr}}(f)$ that acts as PRF. The frequency response $H_{\mathrm{pr}}(f)$ is given by $H_{\mathrm{pr}}(f) = \sum_l h_{\mathrm{pr}}(l) \exp(-j2\pi f l T_d)$, where $h_{\mathrm{pr}}(l)$ is the lth PRF coefficient and T_d denotes the tap spacing of the PRF, not necessarily equal to T. When T/T_d is integer, the continuous-time PRF is equivalent to a fractionally spaced discrete-time PRF. Next, the output of the tapped delay-line is applied to the TX filter with frequency response $H_{\mathrm{tr}}(f)$. The resulting transmitted energy per symbol E_s is given by

$$E_s = \sigma_a^2 \mathbf{h}_{\mathrm{pr}}^T \mathbf{R}_{\mathrm{tr}} \mathbf{h}_{\mathrm{pr}} \tag{1}$$

where $(.)^T$ denotes transposition, $(\mathbf{R}_{\mathrm{tr}})_{m,n} = \int |H_{\mathrm{tr}}(f)|^2 e^{j2\pi f(mT_d - nT_d)} \mathrm{d}f$ and $(\mathbf{h}_{\mathrm{pr}})_l = h_{\mathrm{pr}}(l)$. When $H_{\mathrm{tr}}(f)$ represents a unit-energy square-root Nyquist filter with respect to the symbol rate $1/\mathrm{T}$ and $T_d = T$, we have $(\mathbf{R}_{\mathrm{tr}})_{m,n} = \delta_{m-n}$, yielding $E_s = \sigma_a^2 \mathbf{h}_{\mathrm{pr}}^T \mathbf{h}_{\mathrm{pr}}$.

4

The channel is characterized by a frequency response $H_{\mathrm{ch}}(f)$ and a zero-mean noise term $n(t)$ with power spectral density $S_n(f)$, representing RX noise and interference.

The received signal $r(t)$ is applied to the RX filter $H_{\mathrm{rec}}(f)$. The resulting signal is sampled at the symbol rate $1/T$. We denote by $z(k)$ the sample taken at instant $kT + \tau$, where the sampling delay τ compensates for the delay introduced by the cascade of the PRF, the TX filter, the channel and the RX filter. The decision $\hat{a}(k)$ about the symbol $a(k)$ is the constellation point which is closest to $u(k)$, with $u(k)$ given by

$$u(k) = \alpha z(k) - \sum_{l>0} \hat{a}(k-l)h_{\mathrm{fb}}(l) \qquad (2)$$

where α is a positive scaling factor, the second term in (2) represents a linear combination of past symbol decisions $\{\hat{a}(k-l), l > 0\}$, and $h_{\mathrm{fb}}(l)$ is the lth coefficient of the FBF. We stack the FBF coefficients into the vector $\mathbf{h}_{\mathrm{fb}}$, with $(\mathbf{h}_{\mathrm{fb}})_l = h_{\mathrm{fb}}(l)$.

The sample $z(k)$ at instant $kT + \tau$ can be represented as

$$z(k) = \sum_{l,m} a(k-m)h_{\mathrm{tot}}(mT - lT_d + \tau)h_{\mathrm{pr}}(l) + \nu(k) \qquad (3)$$

where $h_{\mathrm{tot}}(t) = \int H_{\mathrm{tr}}(f)H_{\mathrm{ch}}(f)H_{\mathrm{rec}}(f)e^{j2\pi ft}\mathrm{d}f$ is the impulse response of the cascade of TX filter, the channel, and the RX filter. The quantity $\nu(k)$ represents the contribution from the noise $n(t)$, and is characterized by its autocorrelation function $R_\nu(m) = E[\nu(k+m)\nu(k)] = \int S_n(f)|H_{\mathrm{rec}}(f)|^2 e^{j2\pi fmT}\mathrm{d}f$; the noise variance is then given by $\sigma_\nu^2 = R_\nu(0)$. In the case where $S_n(f) = N_0/2$ (white noise) and $H_{\mathrm{rec}}(f)$ is a unit-energy square-root Nyquist filter with respect to the symbol interval T, we obtain $R_\nu(m) = (N_0/2)\delta_m$.

Because of manufacturing tolerances, the channel frequency response $H_{\mathrm{ch}}(f)$ should be considered as a random process; the same holds for the channel impulse response $h_{\mathrm{ch}}(t)$ and the corresponding total impulse response $h_{\mathrm{tot}}(t)$. Taking into account that the pairs of (total impulse response, sampling delay) given by $(h_{\mathrm{tot}}(t), \tau)$ and $(h_{\mathrm{tot}}(t - t_0), \tau + t_0)$ yield equivalent samples $z(k)$, we will, without loss of generality, shift the time origin for

5

each realization of $h_{\text{tot}}(t)$ such that $h_{\text{tot}}(t)$ is maximum at $t = 0$. This means that the sample $z(k)$ occurs $kT + \tau$ seconds later than the peak of the specific realization $h_{\text{tot}}(t)$.

3. MMSE equalization

We restrict our attention to finite-length equalizers, i.e., $h_{\text{pr}}(l) = 0$ for $l \notin [-L_{\text{pr,min}}, L_{\text{pr,max}}]$ and $h_{\text{fb}}(l) = 0$ for $l \notin [1, L_{\text{fb}}]$. The equalizer coefficients are selected such that the normalized MSE, given by $\text{MSE}_{\text{avg}} = \mathbb{E}\big[(u(k) - a(k))^2\big]/\sigma_a^2$, is minimized under the restriction that the transmit energy per symbol equals a given value E_s; the expectation $\mathbb{E}[.]$ is with respect to the data symbols, the noise, and the realizations of the channel. The best strategy is to adjust both the PRF and the FBF according to the specific channel realization. A simpler but suboptimum strategy is to take fixed equalizers, irrespective of the channel realization; in this case the equalizers are determined from the statistical properties of $H_{\text{ch}}(f)$. In an intermediate strategy, the PRF is fixed while the FBF is adjusted according to the realization of $H_{\text{ch}}(f)$. Note that for each strategy we also have to determine the sampling delay τ which minimizes the MSE. In the minimization process we assume that the decisions on the past data symbols are correct.

3.1. Strategy S1: adjustable $(\boldsymbol{h}_{\text{pr}}, \boldsymbol{h}_{\text{fb}}, \alpha)$

All filter coefficients are adjusted according to the specific channel realization such that the conditional MSE, given by $\text{MSE}_{\text{cond}} = \mathbb{E}\big[(z(k) - a(k))^2 | H_{\text{ch}}(f)\big]/\sigma_a^2$ is minimized with respect to $(\mathbf{h}_{\text{pr}}, \mathbf{h}_{\text{fb}}, \alpha)$. MSE_{cond} for a given sampling delay τ can be expanded as

$$
\begin{aligned}
\text{MSE}_{\text{cond}} = {} & 1 - 2\alpha \mathbf{h}_{\text{tot},-}^T \mathbf{h}_{\text{pr}} + \alpha^2 \mathbf{h}_{\text{pr}}^T \mathbf{R} \mathbf{h}_{\text{pr}} + \alpha^2 \mu \\
& + \sum_{l=1}^{L_{\text{fb}}} \left(\alpha \sum_m h_{\text{tot}}(l, m, \tau) h_{\text{pr}}(m) - h_{\text{fb}}(l) \right)^2
\end{aligned}
\tag{4}
$$

where $\mu = \sigma_\nu^2/\sigma_a^2$, $(\mathbf{h}_{\text{tot},-})_l = h_{\text{tot}}(\tau - lT_d)$, $(\mathbf{R})_{m,n} = \sum\limits_{l \notin [1, L_{\text{fb}}]} h_{\text{tot}}(l, m, \tau) h_{\text{tot}}(l, n, \tau)$, and $h_{\text{tot}}(l, m, \tau) = h_{\text{tot}}(lT - mT_d + \tau)$.

6

In order to determine the optimal $(\mathbf{h}_{\mathrm{pr}}, \mathbf{h}_{\mathrm{fb}}, \alpha)$ we must minimize $\mathrm{MSE}_{\mathrm{cond}}$ from (4) under the restriction (1). In the Appendix we show that the solution is obtained as

$$\alpha^2 = \frac{\sigma_a^2}{E_s} \mathbf{h}_{\mathrm{tot},-}^T (\mathbf{R} + \mu' \mathbf{R}_{\mathrm{tr}})^{-1} \mathbf{R}_{\mathrm{tr}} (\mathbf{R} + \mu' \mathbf{R}_{\mathrm{tr}})^{-1} \mathbf{h}_{\mathrm{tot},-} \tag{5}$$

$$\mathbf{h}_{\mathrm{pr}} = \frac{1}{\alpha} (\mathbf{R} + \mu' \mathbf{R}_{\mathrm{tr}})^{-1} \mathbf{h}_{\mathrm{tot},-} \tag{6}$$

$$h_{\mathrm{fb}}(l) = \alpha \sum_m h_{\mathrm{tot}}(l, m, \tau) h_{\mathrm{pr}}(m) \qquad l \in [1, L_{\mathrm{fb}}] \tag{7}$$

where $\mu' = \sigma_\nu^2 / E_s$. The corresponding minimum value of $\mathrm{MSE}_{\mathrm{cond}}$ is given by

$$\mathrm{MSE}_{\mathrm{cond}} = 1 - \mathbf{h}_{\mathrm{tot},-}^T (\mathbf{R} + \mu' \mathbf{R}_{\mathrm{tr}})^{-1} \mathbf{h}_{\mathrm{tot},-} \tag{8}$$

It should be noted that, for a given realization $H_{\mathrm{ch}}(f)$, $\mathrm{MSE}_{\mathrm{cond}}$ from (8) still depends on the sampling delay τ. Hence, we have to further minimize $\mathrm{MSE}_{\mathrm{cond}}$ over τ, for each individual realization $H_{\mathrm{ch}}(f)$, to obtain the overall minimum $\mathrm{MSE}_{\mathrm{cond}}$.

3.2. Strategy S2: fixed $(\boldsymbol{h}_{\mathrm{pr}}, \boldsymbol{h}_{\mathrm{fb}}, \alpha)$

Here we derive a fixed $(\mathbf{h}_{\mathrm{pr}}, \mathbf{h}_{\mathrm{fb}}, \alpha)$, irrespective of the channel realization, which minimizes $\mathrm{MSE}_{\mathrm{avg}} = \mathbb{E}\big[\mathrm{MSE}_{\mathrm{cond}}\big]$ under the constraint (1), where the expectation $\mathbb{E}\big[.\big]$ is with respect to the channel realizations. In order to obtain manageable expressions, this optimization will be carried out under the assumption that the optimum sampling delay is the same for all channel realizations; although all channel impulse responses achieve their maximum value at $t = 0$, in practice the optimum sampling delay slightly depends on the shapes of the specific realizations. $\mathrm{MSE}_{\mathrm{avg}}$ for given τ is given by

$$\begin{aligned}
\mathrm{MSE}_{\mathrm{avg}} = {} & 1 - 2\alpha \mathbb{E}\big[\mathbf{h}_{\mathrm{tot},-}^T\big] \mathbf{h}_{\mathrm{pr}} + \alpha^2 \mathbf{h}_{\mathrm{pr}}^T \mathbf{R}_{\mathrm{avg}} \mathbf{h}_{\mathrm{pr}} + \alpha^2 \mu \\
& + \sum_{l=1}^{L_{\mathrm{fb}}} \Big(\alpha \sum_m \mathbb{E}\big[h_{\mathrm{tot}}(l, m, \tau)\big] h_{\mathrm{pr}}(m) - h_{\mathrm{fb}}(l) \Big)^2
\end{aligned} \tag{9}$$

7

where $(\mathbf{R}_{\text{avg}})_{m,n} = \sum_l \mathbb{E}\big[h_{\text{tot}}(l,m,\tau)h_{\text{tot}}(l,n,\tau)\big] - \sum_{l=1}^{L_{\text{fb}}} \mathbb{E}\big[h_{\text{tot}}(l,m,\tau)\big]\mathbb{E}\big[h_{\text{tot}}(l,n,\tau)\big]$.

Because of the similarity between (4) and (9), it follows that, for given τ, (5), (6) and (7) still hold, provided that $h_{\text{tot}}(t)$ and $\mathbf{R}$ are replaced by $\mathbb{E}[h_{\text{tot}}(t)]$ and $\mathbf{R}_{\text{avg}}$, respectively. The corresponding minimum value of MSE_{avg} is given by

$$\text{MSE}_{\text{avg}} = 1 - \mathbb{E}[\mathbf{h}_{\text{tot},-}^T](\mathbf{R}_{\text{avg}} + \mu'\mathbf{R}_{\text{tr}})^{-1}\mathbb{E}[\mathbf{h}_{\text{tot},-}] \tag{10}$$

The overall minimum MSE is obtained by minimizing MSE_{avg} from (10) over the sampling delay τ. Having found the optimum τ which minimizes (10), we compute the corresponding $(\mathbf{h}_{\text{pr}}, \mathbf{h}_{\text{fb}}, \alpha)$, which will be used as the fixed equalizer.

Note that the above fixed equalizer $(\mathbf{h}_{\text{pr}}, \mathbf{h}_{\text{fb}}, \alpha)$ depends both on the first-order and second-order moments of $h_{\text{ch}}(t)$. Hence, this equalizer is different from the equalizer which would result from the minimization of (4) with $h_{\text{ch}}(t)$ simply replaced by $\mathbb{E}[h_{\text{ch}}(t)]$.

3.3. Strategy S3: fixed $(\boldsymbol{h}_{\text{pr}}, \alpha)$, adjustable $\boldsymbol{h}_{\text{fb}}$

An intermediate strategy is investigated where the PRF $\mathbf{h}_{\text{pr}}$ and the scaling factor α are fixed for all channel realizations, while the FBF $\mathbf{h}_{\text{fb}}$ is adjusted according to the specific channel realization. As in S2, the optimization is carried out under the simplifying assumption that the optimum sampling delay τ is the same for all channel realizations. For given $\mathbf{h}_{\text{pr}}$, α and τ, $\mathbf{h}_{\text{fb}}$ minimizes MSE_{cond} from (4), and is selected according to (7). Averaging the resulting MSE_{cond} over the channel realizations yields

$$\text{MSE}_{\text{avg}} = 1 - 2\alpha\mathbb{E}\big[\mathbf{h}_{\text{tot},-}^T\big]\mathbf{h}_{\text{pr}} + \alpha^2\mathbf{h}_{\text{pr}}^T\mathbf{R}_{\text{avg}}\mathbf{h}_{\text{pr}} + \alpha^2\mu \tag{11}$$

where now $(\mathbf{R}_{\text{avg}})_{m,n} = \sum_{l \notin [1, L_{\text{fb}}]} \mathbb{E}\big[h_{\text{tot}}(l,m,\tau)h_{\text{tot}}(l,n,\tau)\big]$.

The fixed coefficients $(\mathbf{h}_{\text{pr}}, \alpha)$ minimize (11) under the constraint (1). The optimum $\mathbf{h}_{\text{pr}}$ and α^2 are again given by (6) and (5), but now with $h_{\text{tot}}(t)$ and $\mathbf{R}$ are replaced by $\mathbb{E}[h_{\text{tot}}(t)]$ and $\mathbf{R}_{\text{avg}}$, respectively. The corresponding minimum value of MSE_{avg} is given

by (10), with updated $\mathbf{R}_{\mathrm{avg}}$. Next, this minimum $\mathrm{MSE}_{\mathrm{avg}}$ is further minimized over τ.

3.4. Complexity considerations

When an interconnect is produced in large quantities, only a (much smaller) representative subset of the realizations must be measured to obtain reliable ensemble statistics. From these statistics, $(\mathbf{h}_{\mathrm{pr}}, \mathbf{h}_{\mathrm{fb}}, \alpha)$ for S2 (i.e., the fixed equalizer) can be computed and set to the corresponding value, which is the same for all realizations.

For S1 and S3, $(\mathbf{h}_{\mathrm{pr}}, \mathbf{h}_{\mathrm{fb}}, \alpha)$ or $\mathbf{h}_{\mathrm{fb}}$ must be adjusted according to the specific realization of the interconnect. When the adjustable part of the equalizer is to be determined from a measurement of the interconnect, *each* realization must be measured. Clearly, much more measurements are required than with S2.

Another approach with S1 and S3, is to implement an adaptive version of (part of) the equalizer. In this implementation, an adaptation circuit automatically sets the value of $(\mathbf{h}_{\mathrm{pr}}, \mathbf{h}_{\mathrm{fb}}, \alpha)$ or $\mathbf{h}_{\mathrm{fb}}$ during operation, depending on the specific channel. Knowledge about the impulse response $h_{\mathrm{tot}}(t)$ can be acquired at the RX from applying known pilot symbols to the TX filter $H_{\mathrm{tr}}(f)$ during a training period. However, for S1 this knowledge must be made available to the TX, in order that $\mathbf{h}_{\mathrm{pr}}$ can be set to the proper value; this requires the presence of a return channel from the RX to the TX. Since no return channel is needed for S3, where only $\mathbf{h}_{\mathrm{fb}}$ depends on the realization of the interconnect, its complexity is much less than for S1. Moreover, the computation of $\mathbf{h}_{\mathrm{fb}}$ for S3 is rather simple. Indeed, as $(\mathbf{h}_{\mathrm{pr}}, \alpha)$ is fixed, applying known pilot symbols to the input of the PRF allows the RX to directly estimate the sampled (at instants $lT + \tau$) impulse response related to the frequency response $\alpha H_{\mathrm{pr}}(f) H_{\mathrm{tr}}(f) H_{\mathrm{ch}}(f) H_{\mathrm{rec}}(f)$; according to (7), the FBF coefficients are equal to those samples at instants $T + \tau, ..., L_{\mathrm{fb}}T + \tau$.

4. BER performance

The BER is obtained by following the derivation outlined in [17]. For a given $H_{\text{ch}}(f)$, and assuming correct past decisions, the sample $u(k)$ from (2) can be decomposed as

$$u(k) = g(0)a(k) + \text{ISI}(k) + n_u(k) \tag{12}$$

where $n_u(k) = \alpha\nu(k)$ is the noise contribution, $\text{ISI}(k) = \sum_{l \neq 0} a(k-l)g(l)$ represents the ISI, and the pulse $g(l)$ is given by

$$g(l) = \begin{cases} \alpha \sum_m h_{\text{tot}}(l, m, \tau)h_{\text{pr}}(m) - h_{\text{fb}}(l) & l \in [1, L_{\text{fb}}] \\ \alpha \sum_m h_{\text{tot}}(l, m, \tau)h_{\text{pr}}(m) & l \notin [1, L_{\text{fb}}] \end{cases} \tag{13}$$

Note that for S1 and S3 we have $g(l) = 0 \; \forall l \in [1, L_{\text{fb}}]$ because of (7). We define L_g as the number of nonzero coefficients $g(l)$. For a Gaussian noise term $n(t)$ and a 2-PAM constellation $\mathcal{C} = \{-1, 1\}$, the BER for a given channel realization is given by

$$\text{BER} = \mathbb{E}\left[Q\left(\frac{g(0) + \text{ISI}(0)}{\sigma}\right)\right] \tag{14}$$

with $\sigma^2 = \alpha^2\sigma_\nu^2$; $Q(x) = 1 - \text{CDF}(x)$ where $\text{CDF}(x)$ denotes the cumulative distribution function of a zero-mean Gaussian random variable with unit variance. The expectation in (14) is with respect to the $L_g - 1$ data symbols $\breve{\mathbf{a}}_0$ contributing to $\text{ISI}(0)$; hence, this expectation involves $2^{L_g - 1}$ terms, which becomes computationally prohibitive for large L_g. This computational complexity is avoided by replacing BER from (14) by the estimate BER_{est} given by

$$\text{BER}_{\text{est}} = \frac{1}{N}\sum_{n=1}^{N} Q\left(\frac{g(0) + \text{ISI}^{(n)}(0)}{\sigma}\right) \tag{15}$$

where $\text{ISI}^{(n)}(0)$ involves the $L_g - 1$ data symbols $\breve{\mathbf{a}}_0^{(n)}$, and $\left(\breve{\mathbf{a}}_0^{(1)}, ..., \breve{\mathbf{a}}_0^{(N)}\right)$ denotes N independent realizations of $\breve{\mathbf{a}}_0$. The expectation of BER_{est} equals BER from (14), and

10

the estimation error variance $\sigma_{Q,\text{est}}^2 = \mathbb{E}[(\text{BER}_{\text{est}} - \text{BER})^2]$ is proportional to $1/N$. For large L_g, a value of N can be selected which achieves sufficient BER estimation accuracy while still satisfying $N \ll 2^{L_g-1}$, yielding a considerably smaller computation time than for the exact evaluation of the expectation in (14).

5. Numerical Results

We illustrate the above equalization strategies and their MMSE and BER performances, considering a chip-to-chip interconnect consisting of a 10-cm long on-board microstrip with a cross-section illustrated in Fig. 2. The nominal values of the geometrical and material parameters are as follows:

1. Width of signal conductor: $w = 100\,\mu\text{m}$

2. Thickness of the signal conductor: $t_k = 35\,\mu\text{m}$

3. Thickness of the dielectric substrate: $h = 500\,\mu\text{m}$

4. Conductivity of signal conductor (copper): $\sigma = 58\,MS/\text{m}$

5. Relative permittivity of dielectric substrate: $\epsilon_r = 4$

6. Loss tangent of dielectric substrate: $\tan\delta = 0.02$

The interconnect is driven by a source with a $50\,\Omega$ impedance, and is terminated by a $50\,\Omega$ load. The channel frequency response $H_{\text{ch}}(f)$ relates the voltage across the load to the voltage at the microstrip input, i.e., after the source impedance.
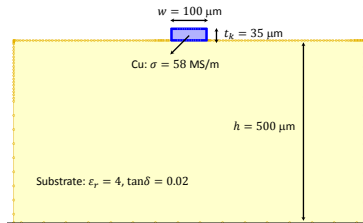

Figure 2: Cross section of microstrip.

11

Because of manufacturing tolerances, the produced interconnects are not identical. We consider the above parameters as independent Gaussian random variables, with a mean equal to their nominal value and a standard deviation of 10% of the mean (which is the order of magnitude for tolerances in consumer electronics). For the generation of the corresponding channel realizations we make use of the approach outlined in [14–16], which we briefly summarize. First, using polynomial chaos theory, the frequency-dependent RLGC parameters of the microstrip are represented as truncated multivariate polynomial expansions of the random parameters. Next, for each realization of the random parameters, the corresponding RLGC parameters are computed using these polynomial expansions. Finally, the corresponding realizations of the channel frequency response are obtained from the RLGC parameters using standard transmission line relations.

We have generated two data sets, each containing 1000 realizations of $H_{\mathrm{ch}}(f)$ with $|f| < 100$ GHz. In the first set (6RV), all parameters are treated as random variables; in the second set (3RV), only the width of the signal conductor w, the thickness of the dielectric substrate h, and the relative permittivity of dielectric substrate ϵ_r are considered as stochastic parameters. Comparing the two sets sets reveals that the higher variability of 6RV manifests itself mainly at frequencies above 20 GHz. Cascading each $H_{\mathrm{ch}}(f)$ with the TX and RX filters $H_{\mathrm{tr}}(f)$ and $H_{\mathrm{rec}}(f)$, which are 5th-order Butterworth filters with their 3 dB cutoff frequency at $1/(2T)$, the corresponding realizations of $h_{\mathrm{tot}}(t)$ have been obtained (and time-shifted so that the maximum of $h_{\mathrm{tot}}(t)$ occurs at $t = 0$). For $1/T = 20$ Gbaud (6RV) and 80 Gbaud (6RV and 3RV), Fig. 3 shows the expectation $\mathbb{E}\big[h_{\mathrm{tot}}(t)\big]$ together with $\mathbb{E}\big[h_{\mathrm{tot}}(t)\big] \pm \sigma(t)$, where $\sigma(t)$ denotes the standard deviation of $h_{\mathrm{tot}}(t)$. The figure shows only the dominant contributions to $h_{\mathrm{tot}}(t)$, consisting of the direct pulse and the first reflection, due to impedance mismatch; in between these contributions, $h_{\mathrm{tot}}(t)$ is nearly zero. Additional reflections are present, but their magnitudes are considerably lower than the first reflection. Note that the magnitude scales depend on the symbol rate, and that the time scales for the direct pulse and for the first refection are different. For $1/T = 20$ Gbaud, the set 3RV is not included and will not be considered: because

12

of the 10 GHz bandwidth of the TX and RX filters, the pulses $h_{\mathrm{tot}}(t)$ corresponding to the 3RV and 6RV sets are nearly identical. There is not much variability between the different realizations of $h_{\mathrm{tot}}(t)$ in the neighborhood of the direct pulse at $t = 0$. A higher variability is observed near the reflections, because the propagation speed is not constant over the different realizations. The differences between the responses for 20 Gbaud and 80 Gbaud come from the larger bandwidth of the TX and RX filter in the latter case, causing more attenuation of the direct and reflected pulses corresponding to 80 Gbaud. Note that the variability of the reflected pulses and their magnitude relative to the direct pulse are larger for 20 Gbaud than for 80 Gbaud; hence, the postcursor ISI at the RX filter output (caused by the falling edge of the direct pulse and by the dominant reflections of $h_{\mathrm{tot}}(t)$) is relatively more important at 20 Gbaud; on the other hand, the rising edge of the direct pulse is longer (when expressed in symbol intervals) at 80 Gbaud than at 20 Gbaud, indicating that the precursor ISI is more important at 80 Gbaud. The difference between the two data sets with $1/T = 80$ Gbaud consists of a slightly smaller variance for 3RV around the direct and reflected pulse.
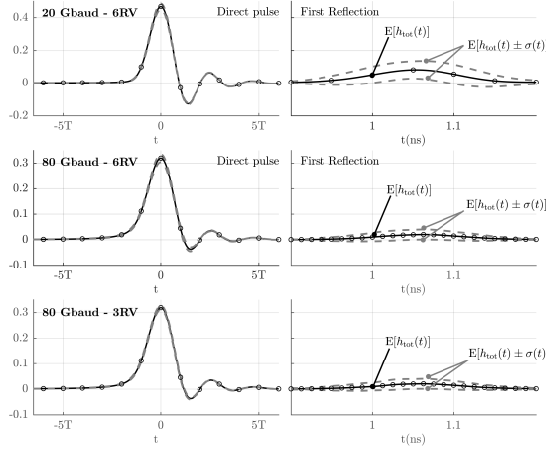

Figure 3: Plot of $\mathbb{E}[h_{\mathrm{tot}}(t)]$ and $\mathbb{E}[h_{\mathrm{tot}}(t)] \pm \sigma(t)$, for 20 Gbaud (6RV) and 80 Gbaud (6RV and 3RV).

Although longer (adjustable) equalizers typically yield less residual ISI, one must take into account that the proposed equalizers must be implementable in practice. We adopt

13

the following approach to keep the number of filter taps within reasonable limits.

- As far as the PRF is concerned, when $T_d = T$ we opt for a 5-tap PRF, which is of the order reported in literature for implemented equalizers [3, 5, 7, 12]. When the PRF operates at twice the symbol rate ($T_d = T/2$) then a 10-taps PRF is considered, such that both PRFs have the same time span, which includes the dominant samples of the direct pulse contained in $h_{\text{tot}}(t)$.

- We initially calculate the coefficients of the above short PRF combined with a long FBF, i.e. $h_{\text{fb}}(l) \; \forall l \in [1, ..., L_{\text{fb}}]$, where typically L_{fb} is chosen such that the FBF has a total span of $4\,\text{ns}$, which includes all relevant postcursor ISI. However this large amount of FBF taps involves a huge implementation complexity, so we drastically reduce the number of FBF taps by means of the following approach. As the dominant taps of the adjustable long FBF (S1 and S3) correspond to the locations of the direct pulse and the first reflection and the remaining taps are very small, we keep from the long FBF only those with the largest magnitude, and set the remaining taps to zero. For 20 and 80 Gbaud we keep only 5 and 10 FBF taps respectively. In S2, on the other hand, the fixed FBF is designed for coping with the average pulse $\mathbb{E}[h_{\text{tot}}(t)]$. When the actual pulse $h_{\text{tot}}(t)$ differs substantially from $\mathbb{E}[h_{\text{tot}}(t)]$, the fixed FBF will in some cases actually increase (rather than reduce) the postcursor ISI; this effect occurs especially for tap positions near the reflections, where the largest variance is observed. This increase in postcursor ISI is avoided by shortening the fixed FBF to a only few consecutive taps that deal with the postcursor ISI caused by the direct pulse (where the variability is small); we take 5 en 10 taps when operating at 20 Gbaud and 80 Gbaud, respectively. In the discussion of the MMSE and BER performance we will always consider both the long FBF and the FBF with the small number (5 or 10) of taps. The FBFs with reduced number of taps will be referred to as "sparse".

It should be noted that the short PRF and sparse FBF resulting from the above approach are not jointly optimum for the considered numbers of filter taps, because the PRF taps have been derived under the assumption of a long FBF. Obtaining the optimum filters with few taps would be prohibitive in terms of computation time, because for S1 and S3 the MMSE must be computed for all possible combinations of 5 FBF tap positions out of 80 positions (20 Gbaud transmission) or 10 FBF tap positions out of 320 positions (80 Gbaud transmission), after which the overall minimum MMSE must be selected.

5.1. MMSE performance

In this subsection we investigate the average minimum MSE for the equalization strategies considered in section 3, which is obtained by first minimizing over the sampling delay τ the conditional MSE for each individual channel realization, followed by averaging the resulting minimum conditional MSE over the channel realizations. The conditional MSE to be minimized over τ is given by (8) for S1, by (4) with $(\mathbf{h}_{\mathrm{pr}}, \mathbf{h}_{\mathrm{fb}}, \alpha)$ representing the fixed equalizer for S2, and by (4) with the last term removed and $(\mathbf{h}_{\mathrm{pr}}, \alpha)$ representing the fixed part of the equalizer for S3. The average MMSE will be displayed as a function of the signal-to-noise ratio (SNR), which we define as $\mathrm{SNR} = E_s \int |H_{\mathrm{rec}}(f)|^2 df / \sigma_\nu^2$ (which simplifies to $\mathrm{SNR} = 2E_s/N_0$ when $S_n(f) = N_0/2$).

Fig. 4a shows the MSE performance of the three equalization strategies when $1/T = 20$ Gbaud. As expected, the adjustable equalizer (S1) yields the lowest MSE. When only the FBF is adjustable (S3), the degradation compared to S1 is due to the fixed PRF, which cannot adjust to the (small) variability of the direct pulse of $h_{\mathrm{tot}}(t)$. A considerable MSE floor is observed when the entire equalizer is fixed (S2); this is attributed to the fixed FBF which cannot adjust to the (rather large) variability of the reflections in $h_{\mathrm{tot}}(t)$. Reducing the PRF tap spacing from T to $T/2$ slightly reduces the MSE for S2, whereas for S1 and S3 the MSE is essentially the same as with tap spacing T. For all three strategies, the MSE is increased when moving from the long FBF to the sparse FBF.
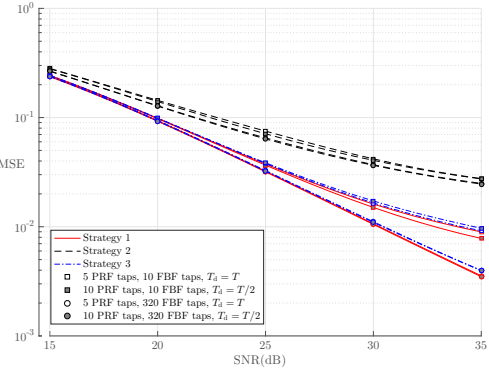
Fig. 4b shows the MSE performance at 80 Gbaud for the data set 6RV. The sparse
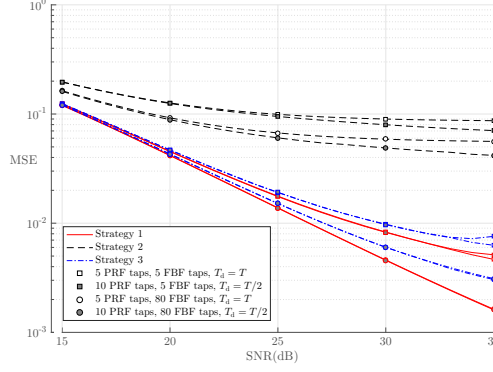
(a) $1/T$=20 Gbaud(6RV).
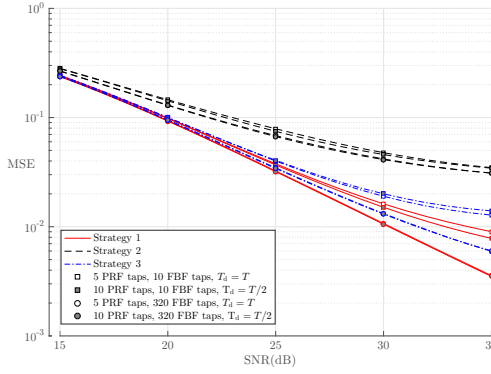
(b) $1/T$=80 Gbaud(6RV).

(c) $1/T$=80 Gbaud(3RV).

Figure 4: MSE$_{\text{avg}}$ versus SNR

FBF has 10 taps, since 5 taps give rise to a large MSE floor. Comparing the cases of 20 Gbaud and 80 Gbaud, we see that for S1 and S3 the MSE at high SNR (e.g., 35 dB) is larger for 80 Gbaud, whereas for S2 the MSE is larger for 20 Gbaud; this behavior for S2 is explained by the smaller magnitude and variability of the reflected pulses in $h_{\text{tot}}(t)$ at 80 Gbaud, compared to 20 Gbaud. For the sparse FBF, $T_d = T/2$ yields a smaller MSE compared to $T_d = T$ in the case of S1, whereas the opposite is observed for S3; this behavior for S3 is attributed to the fact that the fixed PRF yields more postcursor ISI, which cannot be canceled by the sparse FBF.

Fig. 4c displays the MSE performance for $1/T = 80$ Gbaud, but now for the data set 3RV. In S1 no noticeable difference in MSE performance is found compared to set 6RV.

For S2 and S3 the degradation with respect to S1 is clearly much less for 3RV than for 6RV, which can be explained by the smaller variability in the data set 3RV.
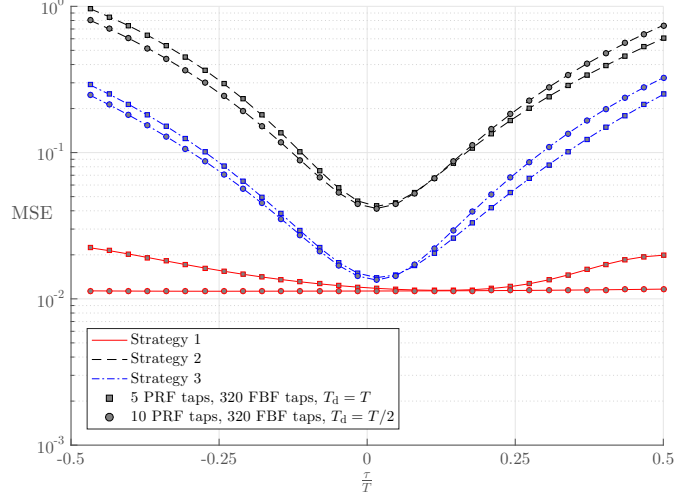

Figure 5: MSE$_{\text{avg}}$ versus τ for $1/T = 80$ Gbaud (6RV), SNR $= 30$ dB.

For SNR $= 30$ dB and 80 Gbaud (6RV), Fig. 5 illustrates the sensitivity of the MSE to the sampling delay τ. We display as a function of τ the averages (over the channel realizations) of the conditional MSE for the considered strategies (i.e., the expressions (8), (4), and (4) without its last term, which correspond to S1, S2 and S3, respectively), where the fixed parameters in S2 and S3 correspond to the optimum sampling delay. For S1 the adjustable PRF can to some extent compensate for a non-optimum sampling delay. This compensation works better with the fractionally-spaced ($T_d = T/2$) than with the symbol-spaced ($T_d = T$) PRF: as its frequency response is periodic with period $1/T_d$, the adjustable PRF is able to introduce an arbitrary additional delay over the entire useful signal frequency band $|f| \leq B$ (with B slightly exceeding $1/(2T)$ due to the TX and RX filter rolloff) only when $T_d \leq 1/(2B)$, which condition is satisfied for $T_d = T/2$ but not for $T_d = T$. For S2 and S3, the fixed PRF cannot compensate for a non-optimum sampling delay, yielding an increased sensitivity compared to S1; using $T_d = T/2$ does not reduce the sensitivity compared to $T_d = T$. The increased sensitivity of S2 and S3 is not problematic when a proper timing algorithm is used (see section 5.2).

17

5.2. BER performance

In this section we investigate the BER performance for 2-PAM transmission resulting from the three equalization strategies. We display as a function of SNR the average (over the set of channel realizations) of the BER corresponding to a given channel. The BER has been estimated as outlined in section 4, with $N = 10^7$, which provides for all channel realizations considered a sufficient accuracy of $\frac{\sigma^2_{Q,\text{est}}}{\text{BER}^2} \leq 10^{-3}$ down to BER=10^{-12}.

Fig. 6a shows the BER performance as a function of SNR at 20 Gb/s, with channel realizations from the set 6RV. Both a sparse (5 taps) and a long (80 taps) FBF are considered. The PRF tap spacing equals T (5 taps PRF) or $T/2$ (10 taps PRF). The following observations can be made.

- The best BER performance is obtained for the fully adjustable equalizer (S1). The PRF tap spacings T and $T/2$ yield virtually the same BER. A BER of 10^{-12} is achieved at SNR $\approx$ 23.7 dB for the long FBF, and at SNR $\approx$ 24.8 dB for the sparse FBF (i.e., only 1.1 dB degradation for the latter).

- Because of the small variability of the direct pulse in $h_{\text{tot}}(t)$, S3 yields nearly the same BER as S1.

- Due to the high variability of the reflected pulses in $h_{\text{tot}}(t)$, the fixed equalizer (S2) yields a poor BER performance, showing a BER floor near BER = 10^{-4} (long PRF) and BER = 10^{-6} (sparse PRF). The better high-SNR performance for the sparse FBF is attributed to the fact that (i) the fixed sparse FBF reduces the postcursor ISI from the direct pulse, which has small variability, and leaves the postcursor ISI from the reflections unaltered; (ii) the fixed long FBF cannot adapt to the strong and highly variable reflections, thereby increasing (rather than reducing) the peak ISI, compared to the sparse FBF. The BER is slightly smaller for $T_d = T/2$ compared to $T_d = T$.

18
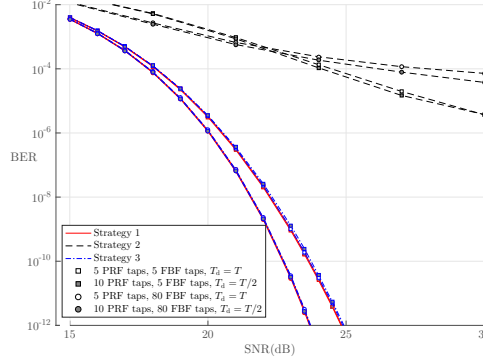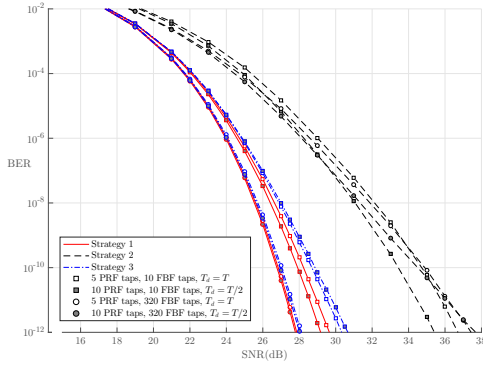
(a) 20 Gb/s (6RV).

(b) 80 Gb/s (6RV).
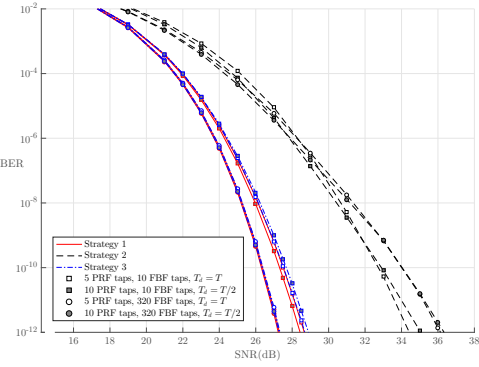
(c) 80 Gb/s (3RV).

Figure 6: BER versus SNR for 2-PAM.

Fig. 6b shows the BER performance at 80 Gb/s (6RV), for both the long FBF (320 taps) and the sparse FBF (10 taps). The following observations can be made.

- In the case of the long FBF, S1 with $T_d = T/2$ achieves a BER of 10^{-12} at SNR ≈ 27.8 dB, which is about 4 dB worse, compared to the 20 Gb/s bitrate; the BER performance for $T_d = T$ is only marginally worse. Only a small degradation (≈ 0.2 dB) is noted when keeping the PRF fixed (S3), with again $T_d = T/2$ performing slightly better than $T_d = T$.

- The BER performances of S1 and S3 are somewhat degraded when using the sparse instead of the long FBF. Taking the BER for S1 with the long FBF and $T_d = T/2$ as a reference, the degradation at a BER of 10^{-12} and $T_d = T/2$ amount to about

19

1.5 dB for S1 and 3 dB for S3. The PRF spacing $T_d = T/2$ performs about 0.5 dB better than $T_d = T$ for S1, whereas the opposite holds for S3; compared to the adjustable PRF, the fixed PRF gives rise to more postcursor ISI, which cannot be canceled by the sparse FBF.

- In contrast with the 20 Gb/s case, the BER for S2 does not show an error floor for the 80 Gb/s bitrate; this is because at 80 Gbaud the reflections in $h_{\text{tot}}(t)$ have a lower magnitude and smaller variability. However, compared to S1 with long FBF, we need to increase the SNR by at least 7.5 dB to achieve a BER of 10^{-12}, making this strategy less attractive to implement in practice.

Fig. 6c shows the BER curves at 80 Gb/s but now with data set 3RV. In general, a similar behavior as for the 6RV data set is observed. Because of the smaller variability of the 3RV data set, the SNR values needed to achieve a BER of 10^{-12} are lower than with the 6RV data set; considering for each strategy the curve yielding the best performance, these SNR values are reduced by about 0.6 dB, 1,0 dB and 0.8 dB for S1, S2 and S3, respectively.

In Table 1 the SNR (in dB) needed to obtain a BER of 10^{-12} is listed for various combinations of the number of filter taps, for 80 Gb/s and $T_d = T$. First of all, the table shows that increasing the number of PRF taps beyond 10 only yields a marginal change in performance[1] for S1 and S3. Next, we observe for S1 and S3 that the degradation rapidly drops when the number of FBF taps increases, making it possible to achieve a nearly optimal equalization with a limited amount of FBF taps. For S2 we observe as explained before that the degradation increases with the number of taps in the FBF. A similar dependence of the BER on the number of filter taps holds for 80 Gb/s with $T_d = T/2$, and for 20 Gb/s with $T_d = T$ and $T_d = T/2$.

In practice, the optimal sampling delay which minimizes the MSE for a given channel

---

[1]as the PRF and FBF are not jointly optimum (the former has been computed assuming the latter is very long), the performance does not necessarily improve with increasing PRF length

Table 1: SNR (dB) at BER=10^{-12} for different number of filter taps for 80 Gb/s (6RV) and $T_d = T$.

# PRF taps	# FBF taps	S1	S2	S3
5	10	29.7	36.7	30.3
5	20	28.4	36.7	28.8
5	40	27.9	36.7	28.3
5	320	27.9	37.4	28.1
15	10	29.8	35.2	30.4
15	320	27.8	38.4	28.0

Table 2: Degradation in SNR (dB) for $\tau = 0$ compared to optimal sampling (BER $= 10^{-12}$, $T_d = T$)

	# FBF taps	S1	S2	S3
$1/T = 20$ Gb/s, 6RV, $T_D = T$	5	0.7	-	0.7
	80	0.1	-	0.1
$1/T = 80$ Gb/s, 6RV, $T_D = T$	10	0.4	-0.3	0.2
	320	0.2	0.2	0.1
$1/T = 80$ Gb/s, 3RV, $T_D = T$	10	0.3	1.8	0.3
	320	0.2	1.8	0.2

realization might be difficult to implement at the RX, in which case we have to resort to a simpler approach. A possible solution is to take $\tau = 0$, i.e., for each channel realization the sampling instant $kT + \tau$ corresponds to the maximum of $h_{\text{tot}}(t - kT)$. This can be achieved by a timing algorithm which estimates $h_{\text{tot}}(t)$ (see section 3.4) and locates the position of its maximum. In Table 2 the degradation in SNR at BER $= 10^{-12}$ of this more practical sampling compared to the optimal sampling is presented for $T_d = T$, with the fixed equalizer parts (S2 and S3) optimized for $\tau = 0$. The negative entry (-0.3 dB) in the table indicates that for the considered configuration the BER performance for $\tau = 0$ is actually better than for optimum sampling; we have verified that although $\tau = 0$ yields a larger MSE, the peak value of the ISI for this particular case is smaller for $\tau = 0$ than with optimum sampling. From Table 2 we conclude that the degradation associated with sampling at $\tau = 0$ does not exceed 1 dB for most configurations.

Finally, we relate the above SNR values which yield BER $= 10^{-12}$ to representative SNRs encountered in practice. For a transmit signal with a root mean-square (rms) value of 0.5 V, RX noise characterized by a rms one-sided density of 10 nV$/\sqrt{\text{Hz}}$ (average over the Nyquist band $(0, 1/(2T))$) and a RX filter approximated by an ideal lowpass filter

with bandwidth $1/(2T)$, the RX noise gives rise to SNR values of 54 dB and 48 dB for bitrates of 20 Gb/s and 80 Gb/s, respectively. These SNR values are much larger than those required to achieve a BER of 10^{-12} with S3 (with sparse FBF), which according to Figs. 6a and 6b are about 25 dB (20 Gb/s bitrate) and 31 dB (80 Gb/s bitrate). The latter SNR values are met in the presence of interference levels (in addition to the RX noise) of about 28 mV rms (20 Gb/s bitrate) and 14 mV rms (80 Gb/s bitrate) within the corresponding Nyquist band. The PCB should be designed in order not to exceed these interference levels.

6. Conclusions

In this contribution we investigated the equalization for multi-Gb/s chip-to-chip interconnects affected by manufacturing tolerances. We considered a MMSE equalization scheme consisting of a PRF at the TX and a FBF operating on past decisions at the RX. Three strategies were investigated: in S1 both the PRF and the FBF are adjustable to the specific realization, S2 involves a fixed PRF and FBF based on the statistics of the interconnect, and in S3 the PRF is fixed but the FBF is adjustable. For a 10 cm microstrip with a 10% tolerance on its parameters, we have shown that S1 with a long FBF achieves a BER of 10^{-12} for SNR $\approx$ 23.7 dB at 20 Gb/s and SNR $\approx$ 27.8 dB at 80 Gb/s. The implementation complexity is considerably reduced when using S3 with a limited amount of FBF taps, at the expense of only a small BER degradation (about 1.1 dB at 20 Gb/s and 3 dB at 80 Gb/s). A considerably larger BER degradation occurs with S2, which suffers from a BER floor at 20 Gb/s, and shows a 7.5 dB degradation compared to S1 with long FBF at BER $= 10^{-12}$ when operating at 80 Gb/s.

We have shown that for some configurations a 10 tap PRF with half-symbol spacing brings a small BER performance advantage ($<$ 0.5 dB) compared to a 5-tap symbol-spaced PRF; however the cost of this advantage is a doubling of the implementation complexity at the TX. For S1, using half-symbol spacing at the PRF makes the perfor-

mance insensitive to a non-optimum sampling delay. For S2 and S3, the sensitivity to a non-optimum sampling delay is not reduced when moving to half-symbol spacing, so that taking the latter PRF spacing for these strategies is pointless.

7. Appendix

Obviously, for given α and $\mathbf{h}_{\mathrm{pr}}$, $\mathrm{MSE}_{\mathrm{cond}}$ from (4) is minimized by selecting $\mathbf{h}_{\mathrm{fb}}$ according to (7) which makes zero the last term in (4). The $\mathrm{MSE}_{\mathrm{cond}}$ is then given by

$$\mathrm{MSE}_{\mathrm{cond}} = 1 - 2\alpha \mathbf{h}_{\mathrm{tot},-}^T \mathbf{h}_{\mathrm{pr}} + \alpha^2 \mathbf{h}_{\mathrm{pr}}^T \mathbf{R} \mathbf{h}_{\mathrm{pr}} + \alpha^2 \mu \tag{16}$$

In order to take the constraint (1) into account, we consider the unconstrained minimization of the Lagrangian $\mathrm{MSE}_{\mathrm{cond}} + \lambda \mathbf{h}_{\mathrm{pr}}^T \mathbf{R}_{\mathrm{tr}} \mathbf{h}_{\mathrm{pr}}$ over $(\mathbf{h}_{\mathrm{pr}}, \alpha)$, where $\mathrm{MSE}_{\mathrm{cond}}$ is given by (16) and λ is the Lagrange multiplier. Equating to zero the derivatives of the Lagrangian with respect to $\mathbf{h}_{\mathrm{pr}}$ and α yields

$$-\alpha \mathbf{h}_{\mathrm{tot},-} + \alpha^2 \mathbf{R} \mathbf{h}_{\mathrm{pr}} + \lambda \mathbf{R}_{\mathrm{tr}} \mathbf{h}_{\mathrm{pr}} = 0 \tag{17}$$

$$-\mathbf{h}_{\mathrm{tot},-}^T \mathbf{h}_{\mathrm{pr}} + \alpha \mathbf{h}_{\mathrm{pr}}^T \mathbf{R} \mathbf{h}_{\mathrm{pr}} + \alpha \mu = 0 \tag{18}$$

Premultiplying (17) with $\mathbf{h}_{\mathrm{pr}}^T$, multiplying (18) with α and subtracting the resulting equations, we obtain $\alpha^2 \mu = \lambda E_s / \sigma_a^2$ or equivalently, $\lambda = \alpha^2 \mu'$ with $\mu' = \sigma_\nu^2 / E_s$. Substituting this value of λ into (17) and solving for $\mathbf{h}_{\mathrm{pr}}$ yields (6). The value of α results from the constraint (1), yielding (5). The resulting minimal $\mathrm{MSE}_{\mathrm{cond}}$ is given by (8).

8. Acknowledgments

References

[1] Y.-S. Kim, S.-K. Lee, S.-J. Bae, Y.-S. Sohn, J.-B. Lee, J.-S. Choi, H.-J. Park, J.-Y. Sim, An 8GB/s Quad-Skew-Cancelling Parallel Transceiver in 90nm CMOS for High-Speed DRAM Interface, in: 2012 IEEE Int. Solid-State Circuits Conf., 2012, pp. 136–138. doi:10.1109/ISSCC.2012.6176952.

[2] V. Stojanovic, M. Horowitz, Modeling and Analysis of High-Speed links, in: Proc. of the IEEE 2003 Custom Integrated Circuits Conf., 2003., 2003, pp. 589–594. doi:10.1109/CICC.2003.1249467.

[3] J. F. Bulzacchelli, M. Meghelli, S. V. Rylov, W. Rhee, A. V. Rylyakov, H. A. Ainspan, B. D. Parker, M. P. Beakes, A. Chung, T. J. Beukema, P. K. Pepeljugoski, L. Shan, Y. H. Kwark, S. Gowda, D. J. Friedman, A 10-Gb/s 5-Tap DFE/4-Tap FFE Transceiver in 90-nm CMOS Technology, IEEE J. Solid-State Circuits 41 (12) (2006) 2885–2900. doi:10.1109/JSSC.2006.884342.

[4] M. H. Nazari, A. Emami-Neyestanak, A 15-Gb/s 0.5-mW/Gbps Two-Tap DFE Receiver With Far-End Crosstalk Cancellation, IEEE J. Solid-State Circuits 47 (10) (2012) 2420–2432. doi: 10.1109/JSSC.2012.2203870.

[5] M. Kossel, T. Toifl, P. A. Francese, M. Brandli, C. Menolfi, P. Buchmann, L. Kull, T. M. Andersen, T. Morf, A 10 Gb/s 8-Tap 6b 2-PAM/4-PAM Tomlinson-Harashima Precoding Transmitter for Future Memory-Link Applications in 22-nm SOI CMOS, IEEE J. Solid-State Circuits 48 (12) (2013) 3268–3284. doi:10.1109/JSSC.2013.2279057.

[6] K. Fukuda, H. Yamashita, G. Ono, R. Nemoto, E. Suzuki, N. Masuda, T. Takemoto, F. Yuki, T. Saito, A 12.3-mW 12.5-Gb/s Complete Transceiver in 65-nm CMOS Process, IEEE J. Solid-State Circuits 45 (12) (2010) 2838–2849. doi:10.1109/JSSC.2010.2075410.

[7] J. F. Bulzacchelli, C. Menolfi, T. J. Beukema, D. W. Storaska, J. Hertle, D. R. Hanson, P.-H. Hsieh, S. V. Rylov, D. Furrer, D. Gardellini, A. Prati, T. Morf, V. Sharma, R. Kelkar, H. A. Ainspan, W. R. Kelly, L. R. Chieco, G. A. Ritter, J. A. Sorice, J. D. Garlett, R. Callan, M. Brandli, P. Buchmann, M. Kossel, T. Toifl, D. J. Friedman, A 28-Gb/s 4-Tap FFE/15-Tap DFE Serial Link Transceiver in 32-nm SOI CMOS Technology, IEEE J. Solid-State Circuits 47 (12) (2012) 3232–3248. doi:10.1109/JSSC.2012.2216414.

[8] D. G. Kam, M. B. Ritter, T. J. Beukema, J. F. Bulzacchelli, P. K. Pepeljugoski, Y. H. Kwark, L. Shan, X. Gu, C. W. Baks, R. A. John, G. Hougham, C. Schuster, R. Rimolo-Donadio, B. Wu, Is 25 Gb/s On-Board Signaling Viable?, IEEE Trans. Adv. Packag. 32 (2) (2009) 328–344. doi: 10.1109/TADVP.2008.2011138.

[9] J. Zhang, Q. B. Chen, K. Qiu, A. C. Scogna, M. Schauer, G. Romo, J. L. Drewniak, A. Orlandi, Design and Modeling for Chip-to-Chip Communication at 20 Gbps, in: 2010 IEEE Int. Symp. on Electromagnetic Compatibility, 2010, pp. 467–472. doi:10.1109/ISEMC.2010.5711320.

[10] C. Gao, J. Chen, X. Wu, P. Amleshi, The Generalized ICN for 25Gbps+ channel using NRZ, PAM-M or Duobinary coding scheme, in: 2012 IEEE Int. Symp. Electromagnetic Compatibility, 2012, pp. 22–27. doi:10.1109/ISEMC.2012.6351756.

[11] G. P. Fettweis, The Tactile Internet: Applications and Challenges, IEEE Trans. Veh. Technol. 9 (1) (2014) 64–70. doi:10.1109/MVT.2013.2295069.

[12] Y. Ban, T. De Keulenaer, Z. Li, J. Van Kerrebrouck, J. H. Sinsky, B. Kozicki, J. Bauwelinck, G. Torfs, A Wide-Band, 5-Tap Transversal Filter With Improved Testability for Equalization up to 84 Gb/s, IEEE Microwave Wireless Compon. Lett. 25 (11) (2015) 739–741. doi:10.1109/LMWC. 2015.2479844.

[13] J. Bailleul, L. Jacobs, P. Manfredi, D. Vande Ginste, M. Moeneclaey, MMSE equalization of multi-Gb/s chip-to-chip interconnects with M-PAM signaling affected by manufacturing tolerances, in: 2016 Symp. on Communications and Vehicular Technologies (SCVT), 2016, pp. 1–6. doi:10.1109/ SCVT.2016.779765.

[14] D. Vande Ginste, D. De Zutter, D. Deschrijver, T. Dhaene, P. Manfredi, F. Canavero, Stochastic Modeling-Based Variability Analysis of On-Chip Interconnects, IEEE Trans. Compon., Hybrids, Manuf. Technol. 2 (7) (2012) 1182–1192. doi:10.1109/TCPMT.2012.2192274.

[15] P. Manfredi, D. Vande Ginste, D. De Zutter, An Effective Modeling Framework for the Analysis of Interconnects Subject to Line-Edge Roughness, IEEE Microwave Wireless Compon. Lett. 25 (8) (2015) 502–504. doi:10.1109/LMWC.2015.2440776.

[16] P. Manfredi, D. Vande Ginste, D. De Zutter, F. G. Canavero, Generalized Decoupled Polynomial Chaos for Nonlinear Circuits With Many Random Parameters, IEEE Microwave Wireless Compon. Lett. 25 (8) (2015) 505–507. doi:10.1109/LMWC.2015.2440779.

[17] M. Guenach, L. Jacobs, B. Kozicki, M. Moeneclaey, Performance Analysis of Pre-Equalized Multilevel Partial Response Schemes, in: 2015 IEEE Int. Symp. on Signal Processing and Information Technology (ISSPIT), 2015, pp. 652–657. doi:10.1109/ISSPIT.2015.7394418.

Jelle Bailleul received the Master's degree in electrical engineering from Ghent University in 2015. He is now pursuing a Ph.D degree at the Telecommunications and Information Processing Department,Ghent University,Belgium. His research interests include signal processing, modulation and coding and statistical communication theory.

Lennert Jacobs received the Master's degree in electrical engineering and the Ph.D. degree in electrical engineering, both from Ghent University, Belgium, in 2006 and 2012, respectively. He is currently serving as a post-doctoral researcher at the Telecommunications and Information Processing Department, Ghent University, Belgium. His research interests are in fading channels, MIMO techniques, signal processing, and modulation and coding

Paolo Manfredi received the M.Sc. and Ph.D. degrees in electronic engineering from the Polytechnic University of Turin, Italy, in 2009 and 2013, respectively. He is now a Postdoctoral Research Fellow with the Found for Scientific Research - Flanders (FWO) at Ghent University, Belgium. His research interests include circuit and interconnect modeling, statistical and worst-case analysis, signal integrity, electromagnetic compatibility.

Dries Vande Ginste received the M.S. degree and the Ph.D. degree in electrical engineering from Ghent University, Gent, Belgium, in 2000 and 2005, respectively. He is an Associate Professor within the Electromagnetics Group, Dept. of Information Technology, Ghent University, Belgium, focusing on computational electromagnetics, electromagnetic compatibility, signal integrity, and antenna design.

Marc Moeneclaey (M.Sc. 1978, Ph.D. 1983) is Full Professor at the Telecommunications and Information Processing Department, Gent University, Belgium. His main research interests are in statistical communication theory. He authors more than 400 scientific papers, and co-authors the book *Digital communication receivers – Synchronization, channel estimation, and signal processing* (J. Wiley, 1998). He is co-recipient of the Mannesmann Innovations Prize 2000.