Noise Reduction for Optical OFDM Channel Estimation

Henning Paul[1], Abdulamir Ali[2], Karl-Dirk Kammeyer[1], Werner Rosenkranz[2]
[1]Department of Communications Engineering, University Bremen, Germany
[2]Chair for Communications, University of Kiel, Germany
[1]{paul, kammeyer}@ant.uni-bremen.de, [2]{aal, wr}@tf.uni-kiel.de

***Abstract*— In this paper, a noise reduction algorithm for channel estimation in optical Intensity Modulation and Direct Detection (IM/DD) OFDM systems is presented. Firstly, the properties of equivalent channels of IM/DD systems are discussed, then the noise reduction algorithm is developed. Simulation results for two different types of IM/DD systems conclude the paper.**

I. Introduction

In recent years, Orthogonal Frequency Division Multiplexing (OFDM) has been proposed for optical long-haul communications due to the ease of equalization of the dispersion-governed fiber optical channel [1]. The approaches pursued can be distinguished into two groups: First, there are coherent optical systems, which use quadrature modulation at the transmitter and zero-IF (Intermediate Frequency) complex valued downconversion using a local laser at the receiver [2]. Second, there exist incoherent approaches, which use envelope detection at the receiver and thus are restricted to real-valued modulation [3]. Since double sideband (DSB) modulation leads to undesired effects due to the phase properties of the fiber [4], single sideband (SSB) modulation is preferred, which can be accomplished by filtering in the bandpass domain [5] or by quadrature modulation [6].

However, the "compatible SSB" (CompSSB) approach presented in [6] is not a linear modulation scheme and therefore the end-to-end equivalent baseband channel, described in the electrical domain cannot be formulated [7]. This implies that there is no such thing as "perfect channel knowledge" in any CompSSB simulation. Usually, in these cases, the equivalent channel is modeled by an allpass with quadratic phase response, but in this paper we will show that by use of noise reduced [8] channel estimation, better results than with this approximation can be achieved.

The remainder of this paper is organized as follows: First, the system models for double sideband and compatible single sideband modulation are introduced and the equivalent channel is, if possible, is presented. The next section deals with channel estimation and the proposed noise reduction algorithm. After presentation of simulation results for two different system concepts, the discussion of results concludes this paper.

This work was supported by the German Research Foundation (DFG) under grants Ka814/19 and Ro2100/6.

II. System Model

In an Intensity Modulation/Direct Detection (IM/DD) system the real-valued, information-bearing signal $x(t)$ is modulated onto the carrier lightwave by means of a Mach-Zehnder modulator (MZM) and detected using a photo diode, which detects the instantaneous power. These two components expose severe nonlinearities, which can be compensated by means of digital signal processing [9]. This results in an overall system as depicted in Fig. 1 with a transmitter characteristic

$$y(t) = \beta g_0^1(m_{\mathrm{pre}} x(t) + b_{\mathrm{pre}}). \quad (1)$$

$g_0^1(\cdot)$ is a clipping function with lower and upper clipping thresholds 0 and 1, m_{pre} and b_{pre} are system parameters controlling the operation point, $\beta > 0$ is a scaling factor introduced to adapt to possible power constraints on the channel. In the "Back-2-Back" (B2B) case, i.e., with no channel present ($h_{\mathrm{c}}(t) = \delta_0(t)$) and with no additive noise $n(t)$, $z(t) = y(t)$ holds.

With a complex baseband channel $h_{\mathrm{c}}(t)$, for a zero-mean signal $x(t)$, from the input to the output an equivalent baseband channel

$$\tilde{h}(t) \propto \mathrm{Re}\left\{\mathrm{e}^{-j\varphi_0} h_{\mathrm{c}}(t)\right\} \quad (2)$$

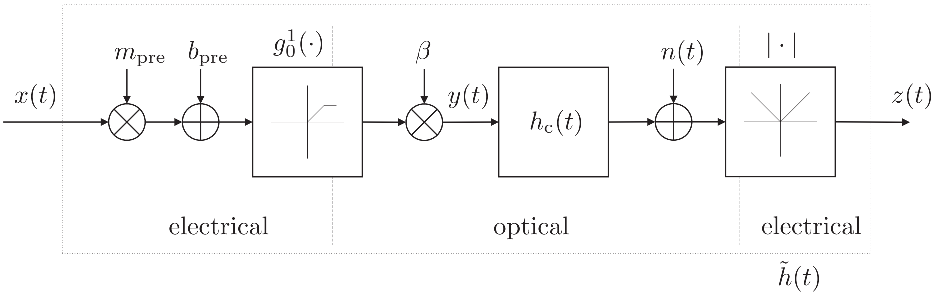

Fig. 1. System model for optical double sideband (DSB) transmission

with φ_0 being the channel phase at carrier frequency can be observed [7], in this case the detector output is given by

$$z(t) = \tilde{h}(t) * x(t) + d(t). \quad (3)$$

$d(t)$ is a term that contains all non-information-bearing terms resulting from distortion caused by the clipping characteristic and mixing products of the noise $n(t)$. In frequency domain, the equivalent channel can be described by

$$\tilde{H}(j\omega) \propto \left(\mathrm{e}^{-j\varphi_0} H_{\mathrm{c}}(j\omega) + \mathrm{e}^{j\varphi_0} H_{\mathrm{c}}^*(-j\omega)\right). \quad (4)$$

With the physical fiber optical channel having an allpass frequency response

$$H_{\mathrm{c}}(j\omega) = \mathrm{e}^{jb(\omega)}, \quad (5)$$

$b(\omega)$ being a (in good approximation) quadratic polynomial in ω, this results in an equivalent channel

$$\tilde{H}(j\omega) \propto \cos(b(\omega)), \quad (6)$$

which exposes frequency selective behavior. One common countermeasure against this effect is the application of sideband suppression by means of optical filtering: This causes $H_{\mathrm{c}}(j\omega)$ to vanish for $\omega < 0$ and thus results in an equivalent channel $\tilde{H}(j\omega) \propto \mathrm{e}^{-j\varphi_0} H_{\mathrm{c}}(j\omega)$ for $\omega > 0$. An alternative to sideband suppression is the direct generation of a SSB signal by means of complex valued I/Q or amplitude/phase modulation. A conventional SSB signal is generated by using the signal $s(t) = m_{\mathrm{pre}} x(t) + b_{\mathrm{pre}}$ for the I component and its Hilbert transform $\mathcal{H}\{s(t)\}$ for the Q component. A compatible SSB signal [6] is generated by amplitude modulation of the carrier with $s(t)$ and phase modulation with the Hilbert transform of its natural logarithm $\mathcal{H}\{\log(s(t))\}$. The latter has the advantage that in the B2B case, $s(t)$ and thus $x(t)$ can be recovered by magnitude detection. This type of system is depicted in Fig. 2.

Note that compatible SSB is not a linear modulation scheme due to the contained phase modulation. This leads to additional nonlinear distortion terms in $z(t)$ caused by the linear channel $h_{\mathrm{c}}(t)$. In this case, the equivalent channel $\tilde{h}(t)$ can not be formulated analytically [7]. This has the consequence that perfect channel knowledge can never be established, which is very disadvantageous for simulations. Usually, $h_{\mathrm{c}}(t)$ is chosen as a substitute for the unknown equivalent channel $\tilde{h}(t)$. Later in this paper, we will show that this assumption leads to worse system performance than channel estimation with the algorithm presented here.

In the following, $x(t)$ shall denote the analog representation of a discrete time, real-valued OFDM signal

$$x(k) = 4 \sum_{l=-\infty}^{\infty} \sum_{n=1}^{N_{\mathrm{FFT}}/2-1} \Bigg[\mathrm{Re}\{X_l(n)\} \cos(\frac{2\pi nk}{N_{\mathrm{FFT}}}) + \mathrm{Im}\{X_l(n)\} \sin(\frac{2\pi nk}{N_{\mathrm{FFT}}}) \Bigg] \cdot f(k - lN_{\mathrm{total}}), \quad (7)$$

with $f(k)$ being a rectangular window

$$f(k) = \begin{cases} 1 & -N_{\mathrm{g}} \leq k < N_{\mathrm{FFT}} \\ 0 & \text{else} \end{cases} \quad (8)$$

of length $N_{\mathrm{total}} = N_{\mathrm{g}} + N_{\mathrm{FFT}}$. $X_l(n)$, $n = 1 \dots N_{\mathrm{FFT}}/2 - 1$ represent the complex valued data symbols of the l-th OFDM symbol, $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ are real part and imaginary part operators. The DC and Nyquist frequency subcarriers are not used due to residual DC components in the receiver output signal $z(t)$.

In the following, the index l of the data symbols will be dropped since only one symbol will be considered; the data symbols $Z(n)$ at the receiver are defined analoguously to $X(n)$, resulting in an OFDM system model

$$Z(n) = \tilde{H}(n) X(n) + D(n). \quad (9)$$

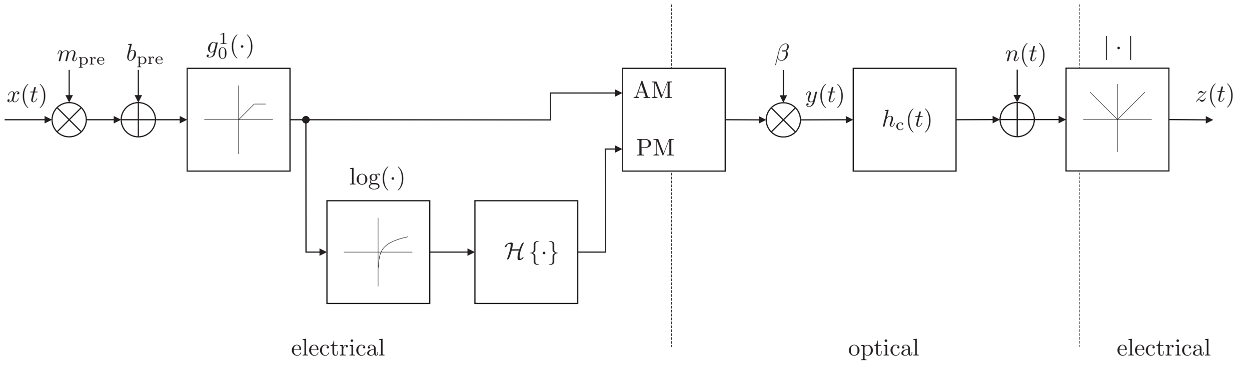

Fig. 2. System model for optical compatible single sideband (SSB) transmission

III. Channel Estimation

A. Least Squares Estimation

Using quaternary phase shift keying (QPSK) training symbols $X_{\rm tr}(n)$, the least squares OFDM channel estimation

$$\hat{H}(n) = \frac{Z(n)}{X_{\rm tr}(n)} \tag{10}$$

is yielded for subcarriers $1 \leq n \leq N_{\rm FFT}/2 - 1$. The estimation error is given by

$$\Delta H(n) = \frac{D(n)}{X_{\rm tr}(n)} \tag{11}$$

and might be correlated with $X_{\rm tr}(n)$, since the interference $D(n)$ might also contain intermodulation products of the subcarriers $X_{\rm tr}(n)$. However, it is modeled to have zero mean averaged over all possible training symbols. Due to this, channel knowledge can be improved arbitrarily by averaging least squares estimations over randomly varying training symbols, since the channel can be assumed to be static. In our case, random QPSK training symbols were chosen.

In $Z(n)$, a lower sideband (corresponding to negative indices $-1 \geq n \geq -N_{\rm FFT}/2 + 1$) can be observed, which stems from the magnitude operation and is a mirrored conjugate image of the upper sideband. Therefore, its subcarrier channel coefficients can be extended in conjugate complex fashion using the estimation of the upper sideband. But the coefficients for $n = 0$ and $n = N_{\rm FFT}/2$ are unknown, since these subcarriers were ommited. This implies that the estimated discrete time channel impulse response $\hat{h}(k)$ cannot be calculated directly, but an interpolation has to be performed.

B. Noise Reduction

The length of the impulse response of the physical channel (5) is infinite in theory, but becomes limited if a band limitation is applied. This inherently is the case for practical OFDM systems, thus $\tilde{h}(k)$ has a known maximum length, which is the basis for the dimensioning of the cyclic prefix, here defined to have length $N_{\rm g}$ samples. Its estimation $\hat{h}(k)$ instead has length $N_{\rm FFT}$, which implies that it contains only noise and interference after the first $N_{\rm g}$ samples. Based on this fact, a noise reduction algorithm was proposed in [8]. The philosophy of the proposed algorithm can be generalized to a minimum mean square error (MMSE) problem, which will be explained in the following. $\tilde{\mathbf{h}}$ shall denote the vector representation of $\tilde{h}(k)$ zero padded to length $N_{\rm FFT}$, $\hat{\mathbf{H}}'$ denotes the vector representation of the channel estimation $\hat{H}(n)$ where the unknown subcarrier coefficients have been omitted.

Now the equation

$$\hat{\mathbf{H}}' = \mathbf{W}_0 \tilde{\mathbf{h}} + \Delta\mathbf{H} \tag{12}$$

for the estimation can be formulated. $\mathbf{W}_0$ is a DFT matrix where the rows corresponding to the unknown subcarrier coefficients have been skipped, the vector $\Delta\mathbf{H}$ represents the estimation error $\Delta H(n)$. The objective is to design a filter matrix $\mathbf{W}_{\rm NR}$ such that the impulse response $\hat{\mathbf{h}}$ can be reconstructed:

$$\hat{\mathbf{h}} = \mathbf{W}_{\rm NR} \hat{\mathbf{H}}'. \tag{13}$$

This is performed by means of a MMSE approach: The mean square error

$$\mathrm{E}\left\{ ||\hat{\mathbf{h}} - \tilde{\mathbf{h}}||^2 \right\} \tag{14}$$

is supposed to be minimized for all possible channels, pilot symbols and noise realizations. Inserting (12) and (13) into this equation, we get

$$\begin{aligned} &\mathrm{E}\left\{ ||\mathbf{W}_{\rm NR}(\mathbf{W}_0 \tilde{\mathbf{h}} + \Delta\mathbf{H}) - \tilde{\mathbf{h}}||^2 \right\} \\ &= \mathrm{E}\left\{ ||(\mathbf{W}_{\rm NR}\mathbf{W}_0 - \mathbf{I})\tilde{\mathbf{h}} + \mathbf{W}_{\rm NR}\Delta\mathbf{H}||^2 \right\}, \end{aligned} \tag{15}$$

which, if $\tilde{\mathbf{h}}$ and $\Delta\mathbf{H}$ are assumed to have zero mean and to be uncorrelated, can be rewritten as

$$\begin{aligned}
&\mathrm{E}\left\{||(\mathbf{W}_{\mathrm{NR}}\mathbf{W}_0 - \mathbf{I})\tilde{\mathbf{h}}||^2\right\} + \mathrm{E}\left\{||\mathbf{W}_{\mathrm{NR}}\Delta\mathbf{H}||^2\right\} \\
&= (\mathbf{W}_{\mathrm{NR}}\mathbf{W}_0 - \mathbf{I})\mathrm{E}\left\{\tilde{\mathbf{h}}\tilde{\mathbf{h}}^{\mathrm{H}}\right\}(\mathbf{W}_0^{\mathrm{H}}\mathbf{W}_{\mathrm{NR}}^{\mathrm{H}} - \mathbf{I}) \\
&+ \mathbf{W}_{\mathrm{NR}}\mathrm{E}\left\{\Delta\mathbf{H}\Delta\mathbf{H}^{\mathrm{H}}\right\}\mathbf{W}_{\mathrm{NR}}^{\mathrm{H}} \qquad (16) \\
&= (\mathbf{W}_{\mathrm{NR}}\mathbf{W}_0 - \mathbf{I})\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}(\mathbf{W}_0^{\mathrm{H}}\mathbf{W}_{\mathrm{NR}}^{\mathrm{H}} - \mathbf{I}) \\
&+ \mathbf{W}_{\mathrm{NR}}\mathbf{R}_{\Delta\mathbf{H}\Delta\mathbf{H}}\mathbf{W}_{\mathrm{NR}}^{\mathrm{H}}.
\end{aligned}$$

Here, $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ and $\mathbf{R}_{\Delta\mathbf{H}\Delta\mathbf{H}}$are the correlation matrices of $\tilde{\mathbf{h}}$ and $\Delta\mathbf{H}$, respectively. Minimizing this expression by setting its derivative with respect to $\mathbf{W}_{\mathrm{NR}}^{\mathrm{H}}$ to zero, we get

$$(\mathbf{W}_{\mathrm{NR}}\mathbf{W}_0 - \mathbf{I})\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}\mathbf{W}_0^{\mathrm{H}} + \mathbf{W}_{\mathrm{NR}}\mathbf{R}_{\Delta\mathbf{H}\Delta\mathbf{H}} = \mathbf{0}, \qquad (17)$$

which can be rearranged into

$$\mathbf{W}_{\mathrm{NR}}(\mathbf{W}_0\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}\mathbf{W}_0^{\mathrm{H}} + \mathbf{R}_{\Delta\mathbf{H}\Delta\mathbf{H}}) = \mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}\mathbf{W}_0^{\mathrm{H}}. \qquad (18)$$

Solving this expression for $\mathbf{W}_{\mathrm{NR}}$ by inversion of $(\mathbf{W}_0\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}\mathbf{W}_0^{\mathrm{H}} + \mathbf{R}_{\Delta\mathbf{H}\Delta\mathbf{H}})$ leads to numerical problems, since in the nearly noise- and interference-free case the elements of $\mathbf{R}_{\Delta\mathbf{H}\Delta\mathbf{H}}$ become very small, making the overall matrix ill-conditioned. Due to the limited channel memory, $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ has block structure

$$\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} = \begin{bmatrix} \mathbf{R}'_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \qquad (19)$$

Obviously, the rank r of $\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ equals the channel memory in taps, which is smaller than N_{g} if the guard interval was designed properly. Inserting this knowledge (19) into (18), the problem can be reformulated using $\mathbf{W}'_0$, which is defined as containing only the r first columns of $\mathbf{W}_0$:

$$\mathbf{W}_{\mathrm{NR}}(\mathbf{W}'_0\mathbf{R}'_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}\mathbf{W}'^{\mathrm{H}}_0 + \mathbf{R}_{\Delta\mathbf{H}\Delta\mathbf{H}}) = \begin{bmatrix} \mathbf{R}'_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}\mathbf{W}'^{\mathrm{H}}_0 \\ \mathbf{0} \end{bmatrix}. \qquad (20)$$

This expression can generally be solved using the Moore-Penrose pseudo inverse. However, if a worst case scenario is assumed and $\mathbf{R}_{\Delta\mathbf{H}\Delta\mathbf{H}}$ and $\mathbf{R}'_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}$ both are assumed to be an identity matrix scaled by $\sigma^2_{\Delta H}$ and $\sigma^2_{\tilde{h}}$ respectively, in our special case a trivial solution can be found by

$$\mathbf{W}_{\mathrm{NR}} = \frac{\sigma^2_{\tilde{h}}}{\sigma^2_{\tilde{h}} + \sigma^2_{\Delta H}}(\mathbf{W}'^{\mathrm{H}}_0\mathbf{W}'_0)^{-1}\mathbf{W}'^{\mathrm{H}}_0. \qquad (21)$$

It can be seen that if PSK modulation is applied, the factor $\frac{\sigma^2_{\tilde{h}}}{\sigma^2_{\tilde{h}} + \sigma^2_{\Delta H}}$ can be ignored.

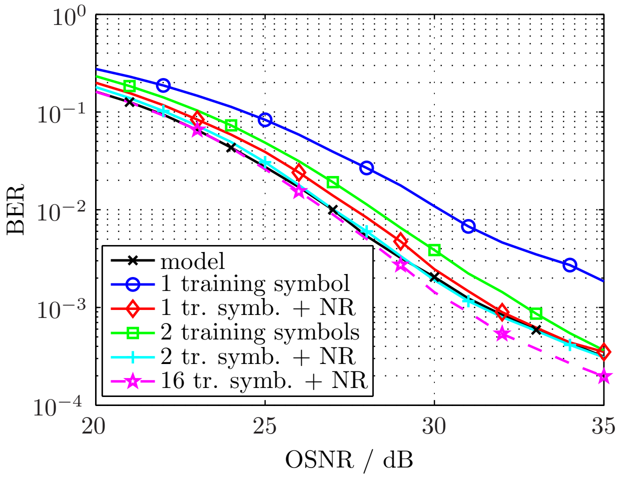

Fig. 3. Bit error rates with and without noise reduction for compatible SSB modulation

IV. Simulation Results

For the following simulations, the assumption $r = N_{\mathrm{g}}$, again representing the worst case scenario has been made. Fig. 3 shows the average bit error rate (BER) over the optical signal-to-noise ratio (OSNR) of an $N_{\mathrm{FFT}} = 1024$ optical IM/DD OFDM system using compatible SSB modulation, a cyclic prefix length of $N_{\mathrm{g}} = 1/4\,N_{\mathrm{FFT}}$, quaternary phase shift keying (QPSK) data symbols and training symbols with a bitrate of 42.8 Gb/s over 80 km standard single mode fiber and a carrier-to-sideband power ratio of approximately 12. A fiber bandwidth of 150 GHz was assumed and no optical filtering was performed. It can be seen that the usage of only one training symbol (plotted in blue) leads to a significant OSNR loss at the interesting target BER of 10^{-3} compared to the case where the model $\tilde{h}(k) = h_{\mathrm{c}}(k)$ was used for the equivalent baseband channel (plotted in black). The noise reduction of this single estimate improves system performance significantly (plotted in red), an additional gain can be observed when two different training symbols are used and the estimates are averaged before noise reduction (cyan curve). At a BER of 10^{-3}, this channel estimation already performs slightly better that the model-based equalization. Its inadequacy becomes apparent, when 16 independent channel estimations are averaged before noise reduction (magenta curve).

The scenario where no perfect channel knowledge is available and additionally the transmission is impaired by nonlinear effects is a case where the noise reduction performs especially well. It is, however, also advantageous for optical systems that are less impaired by nonlinear effects. Fig. 4

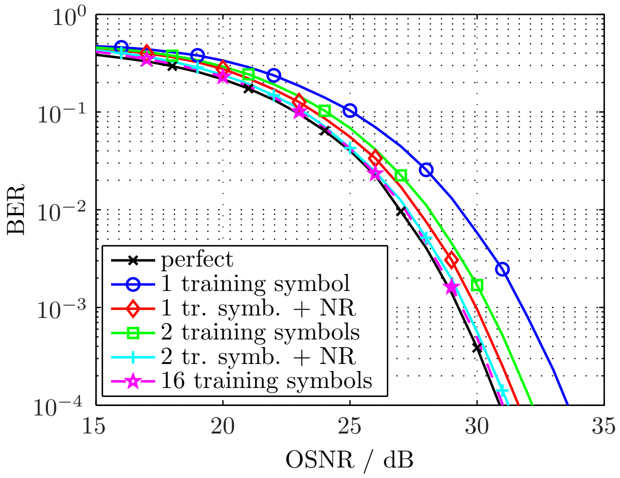

Fig. 4. Bit error rates with and without noise reduction for conventional SSB modulation

shows the simulated bit error rates of an optical IM/DD system using conventional, electrical SSB modulation by use of a I/Q modulator. Here the concept of a spectral gap was applied, doubling the required bandwidth and thus the FFT length to $N_{\mathrm{FFT}} = 2048$. The fiber length was kept fixed at 80 km, its bandwidth was doubled to 300 GHz. No linearization of modulators and detector was applied, the quadrature point was chosen for biasing, a standard deviation of $0.25\,V_\pi$ at the modulator input was established, resulting in a carrier-to-sideband power ratio of approximately 13. As in the previous example, $N_{\mathrm{g}} = 1/4\,N_{\mathrm{FFT}}$ and QPSK modulation were chosen. Also with this system, noise reduction improves the system performance, e.g., at a BER of 10^{-3} by approximately 2 dB if only one training symbol is used (comparing blue and red curve) and by approximately 1 dB if two symbols are used. A further increase in the number of training symbols only provides minor improvements; averaging over 16 estimations yields a bit error performance that is virtually indistinguishable from the case of equalization perfect channel knowledge which is available here, even without noise reduction. However, if only one training symbol can be provided for equalization, noise reduction allows a bit error performance, which is only approximately 1 dB away from the case of perfect channel knowledge.

V. Conclusion

In this paper, a noise reduction algorithm based on MMSE channel estimation was applied to optical communication systems. First, the system model of a compatible SSB system and its impairments on channel equalization were presented, then the MMSE criterion for the channel estimation was formulated. This criterion was approximated using several worst-case assumptions and the results were verified using computer simulations for two different system concepts. We have shown that the use of noise reduction on channel estimation can significantly improve the bit error performance and thus lower the required signal-to-noise ratio for a given target bit error rate.

References

[1] G. P. Agrawal, *Fiber-Optic Communication Systems, Third Edition*, John Wiley & Sons, Inc., 2002.

[2] W. Shieh and C. Athaudage, *Coherent Optical Orthogonal Frequency Division Multiplexing*, IET Electronics Letters, vol. 42, no. 10, pp. 587-588, 2006.

[3] H. Paul and K.-D. Kammeyer, *Modeling and Influences of Transmitter and Receiver Nonlinearities in Optical OFDM Transmission*, 13th International OFDM Workshop 2008 (InOWo '08), Hamburg, Deutschland, August 2008.

[4] J. Wang and K. Petermann, *Small Signal Analysis for Dispersive Optical Fiber Communication Systems*, IEEE Journal of Lightwave Technology, vol. 10, no. 1, pp. 96-100, 1992.

[5] A. Ali, J. Leibrich and W. Rosenkranz, *Spectrally Efficient OFDM-Transmission over Single-Mode Fiber Using Direct Detection*, 13th International OFDM Workshop 2008 (InOWo '08), Hamburg, Deutschland, August 2008.

[6] M. Schuster, B. Spinnler, C.-A. Bunge and K. Petermann, *Direct Detection OFDM for Optical Communication*, 13th International OFDM Workshop 2008 (InOWo '08), Hamburg, Deutschland, August 2008.

[7] H. Paul and K.-D. Kammeyer, *Equivalent Baseband Channels of Systems Using Envelope Detection*, International Journal of Electronics and Communications (AEÜ), Volume 63, Issue 7, July 2009, Pages 533-540, ISSN 1434-8411, DOI: 10.1016/j.aeue.2009.01.010.

[8] H. Schmidt, K.-D. Kammeyer, *Impulse Truncation for Wireless OFDM Systems*, 5th International OFDM-Workshop (InOWo 2000), Hamburg, Deutschland, September 2000.

[9] H. Paul, K.-D. Kammeyer, *Linearization of Transmitter and Receiver Nonlinearities in Optical OFDM Transmission*, 7th International Workshop on Multi-Carrier Systems & Solutions MC-SS 2009, May 2009.