### Extinction ratio induced crosstalk system penalty in WDM networks

J. Song, C.K. Chan, F.K. Tong and I.K. Chen

**Introduction:** A wavelength router is the key element in scalable wavelength division multiple access (WDM) networks using wavelength reuse [1]. Both static and dynamic wavelength routers suffer a small but nonzero crosstalk from neigbouring inputs causing severe degradation in system performance. This type of crosstalk, which originates from other input ports carrying identical or similar wavelength with that of the signal, is difficult to eliminate by filtering. Planar SiO$_2$ waveguide devices demonstrate a crosstalk level of $\sim 25$dB [2]. These unwanted signals will beat with the selected optical signal and introduce a new kind of noise in the receiver [3 – 7]. Previous work on system penalty evaluation either ignored the extinction ratio in their calculations [3, 5, 7], or did not isolate the effect of extinction ratio in their simulations [4, 5]. Here, we have derived a closed form analytical expression for the system penalty, incorporated with extinction ratio $\nu$ (the power ratio of signal ‘mark’ to ‘space’) as a key factor. Both theoretical and experimental results suggest that the extinction ratio should be $\geq 15$dB when several routers are in cascade, assuming a typical crosstalk level of $\sim 25$dB in SiO$_2$-based routers. We further deduce that if the maximum number of nodes is limited to $K$ for an extinction ratio of 15dB, reducing the extinction ratio to 7dB decreases this number to $K/2$.

![Analytical model of our WDMA network and homo-wavelength crosstalk at each wavelength router](image)

**Fig. 1** Analytical model of our WDMA network and homo-wavelength crosstalk at each wavelength router

- a Analytical model
- b Homo-wavelength crosstalk at each wavelength router

**System configuration and analysis:** Fig. 1a shows a partial all-optical, wavelength reuse WDMA network. For transmitting $\lambda_i$ from node 1 to node $K$, the wavelength router in each node will relate $\lambda_i$ from node 1 towards node $K$. Each wavelength router has $N$ input and $N$ output ports and each port carries $M$ wavelengths, $\lambda_0 - \lambda_M$ (Fig. 1b). All $M$ wavelengths at $N$ input ports will be demultiplexed and routed to different outputs according to either the pre-determined (static) or re-configurable (dynamic) pathways. The small but nonzero leakage at the same wavelength, with that of the signal from the other $N - 1$ neighbouring input ports, will interfere with the signal.

We adopt a worst-case scenario where the interferring channels have the same polarisation with that of the signal. Without loss of generality, we assume an identical leakage among all ports at the router, that is, the crosstalk to signal ratio $C = C_0 = C$ for all $i$ and $j$ ports. This ratio can be expressed as $C = P_i/P_s$, where $P_i$ is the power leaking from any other port to the signal port and $P_s$ is the transmitted signal power. We only consider the beat noise originating from the signal and crosstalk because the crosstalk-crosstalk beatings are insignificant [6]. With randomised phases arriving at the detector, the total crosstalk power $C_0$ after $K$ nodes can be simplified [4 – 6] as $C_0 = CK(N - 1)$. We assume that in our network, the signal attenuation in the fibre and wavelength router is fully compensated for by the in-line gain-flattened optical amplifiers, and the noise from the amplifiers is neglected.

The Q-factor for the optimised decision threshold [6] can be written as

$$Q = \frac{2\mu(t-1)(t+1)}{\sqrt{1 + T_s I_{\text{shot},0} + T_s I_{\text{shot},1} + I_{\text{shot},0} + I_{\text{shot},1}}}$$

where $\mu$ is the average photocurrent in the presence of crosstalk channels, $I_{\text{shot}}$ is the circuit thermal noise; $T_s$ and $T_s$ are

**References**


© IEEE 1996
Electronics Letters Online No: 19961443
24 September 1996

P.D. Karaminas and A. Manikas (Digital Communications Section, Electrical and Electronic Engineering Department, Imperial College of Science, Technology and Medicine, London SW7 2BT, United Kingdom)

The authors derive an analytical solution for crosstalk-induced penalty in wavelength routers with consideration of the extinction ratio. Both theoretical and experimental results suggest that the extinction ratio can induce a system penalty and should be kept to at least $\geq 15$dB to maintain the penalty below 1dB for a total crosstalk level of $\sim 25$dB.

**Fig. 1** Outage probabilities for 1000 trials

$C = 1, N = 9, R = 1500$ m, $L = 100$ snapshots

$P_s - 5$ dBm, $P_l - 10$ dBm

$\text{Extinction ratio} = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm

$\nu = 2\nu = 7.5$ cm
the signal-crosstalk beat noises for ‘mark’ and ‘space’, respectively. It can be shown that \( P_{\text{crosstalk}} = 4n_i^2C_{aw}(r + 1) \), and \( P_{\text{crosstalk}} = 4n_i^2C_{aw}(r + 1) \).

When there is no crosstalk, let \( i_0 \) be the average photocurrent in this case; we have

\[
Q = \frac{2i_0(r - 1)(r + 1)}{\sqrt{P_{\text{th}}^2 + P_{\text{th}}^2}} \tag{2}
\]

We need to compare \( i_0 \) with \( i_0 \) to evaluate the crosstalk-induced system penalty \( X \) (in dB). An analytical expression for \( X \) is given by

\[
X = 10\log_{10} \left( \frac{P_{\text{crosstalk}}}{P_{\text{crosstalk}}(r - 1)(r + 1)} \right)
\]

\[
= 10\log_{10} \left( \frac{\sqrt{4n_i^2C_{aw}(r + 1)} + \sqrt{4n_i^2C_{aw}(r + 1)} + P_{\text{th}}}{\sqrt{P_{\text{th}}^2 + P_{\text{th}}^2}} \right)
\]

\[
= -10\log_{10} \left( \frac{\sqrt{4n_i^2C_{aw}} + \sqrt{4n_i^2C_{aw}} + \sqrt{4n_i^2C_{aw}}}{\sqrt{b}} \right) \tag{3}
\]

where \( b \equiv (r - 1)(r + 1) \). For an infinite \( r \), i.e. \( r \to \infty \), \( b \to 1 \). Eqn. 3 converges to eqn. 10 in [6].

![Fig. 2 Calculated system penalty \( P \) against total crosstalk ratio \( C_{aw} \) under different extinction ratio \( r \)](image)

<table>
<thead>
<tr>
<th>BER</th>
<th>Theoretical result</th>
<th>Experimental data</th>
</tr>
</thead>
<tbody>
<tr>
<td>10^-2</td>
<td>(i) ( r ) = 7dB</td>
<td>▼ 10dB</td>
</tr>
<tr>
<td></td>
<td>(ii) ( r ) = 10dB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(iii) ( r ) = 15dB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(iv) ( r ) = \infty</td>
<td></td>
</tr>
</tbody>
</table>

Results and discussion: For Fig. 2 shows \( X \) against \( C_{aw} \) at various \( r \), all calculated under a fixed BER at 10^-2. We can see that when \( C_{aw} \) is small, say < -35dB, the circuit thermal noise \( F_{th} \) is dominant, and the difference in \( X \) for various \( r \) is very small. However, when \( C_{aw} \) increases, signal-crosstalk beat noise becomes dominant, the difference in \( X \) for various \( r \) becomes substantial and the effect of \( r \) should be taken into consideration. Note that \( X \) increases drastically as \( C_{aw} \) approaches ~24dB, a typical value for the wavelength routers using SOI technologies.

![Fig. 3 Experimental setup to measure homo-wavelength crosstalk-induced system penalty \( X \)](image)

We used the experimental setup in Fig. 3 to verify the calculated results. To simulate the homo-wavelength crosstalk, the signal is split into two paths, with one path 7km longer than the other to avoid coherent interference. The polarisation of the crosstalk channel is carefully controlled to maximise the beat noise at the receiver. To generate different extinction ratios, different biases and modulation currents are applied to the DFB laser. Using a 2^1-1 NRZ PRBS at 622Mbits/s, we confirm that \( r \) greatly influences \( X \), especially when signal-crosstalk beat noise becomes dominant (see Fig. 2). The discrepancy ~2dB when \( X = 1dB \) between the theoretical values and the experimental data, arises from a single crosstalk channel performed in the experiment, resulting in an over-estimate of beat noise due to Gaussian noise assumption [3, 6, 7]. This assumption becomes valid when the number of crosstalk channels increases to 16 and beyond, a practical case for WDM networks. From theoretical prediction, \( r \) should be > -15dB for a typical \( C_{aw} \) of ~23dB in wavelength routers.

![Fig. 4 Relationship between total number of ports \( K \) and channel crosstalk ratio \( C \) under different extinction ratio \( r \)](image)

Calculations are based on BER = 10^-2 and using system penalty \( X \) of 1dB as a criterion.

(i) \( r \) = infinite
(ii) \( r \) = 15dB
(iii) \( r \) = 10dB
(iv) \( r \) = 7dB

We can further deduce the network size using \( X = 1dB \) as a criterion. The relationship between the total number of ports \( K = N - 1 \), i.e. the total number of the crosstalk channels, and the crosstalk ratio \( C \) per channel with different \( r \) is shown in Fig. 4, again using BER = 10^-2. Poor \( r \) of ~7dB will reduce the maximum number of nodes in a WDM network to about half of that with \( r \) of ~15dB. For example, when \( C = -40dB \) and \( N = 4 \), the maximum \( K \) for \( r = 15dB \) is 16, and \( K \) reduces to 8 when \( r \) decreases to ~7dB. This strongly illustrates how the extinction ratio heavily influences the maximum number of nodes in WDM networks.

Summary: The influence of the extinction ratio on a homo-wavelength crosstalk-induced system penalty in WDM networks is theoretically and experimentally investigated. An analytical expression is derived incorporated with the extinction ratio, showing that the extinction ratio has significant impact on crosstalk-induced system penalty.

© IEE 1996

Electronics Letters Online No: 19961437
J. Song, C.K. Chan, F.K. Tong and L.K. Chen (Department of Information Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong)

References
Fast-convergence filtered regressor algorithms for blind equalisation

S.C. Douglas, A. Cichocki and S. Amari

Indexing terms: Equalisers, Mobile communication systems, Adaptive equalisers

The authors present a simple extension of the standard Bussgang blind equalisation algorithms that significantly improves their convergence properties. The technique uses the inverse channel estimate to filter the regressor signal. The modified algorithms provide quasi-Newton convergence in the vicinity of a local minimum of the chosen cost function with only a modest increase in the overall computational complexity of the system. An example of the technique as applied to the constant-modulus algorithm indicates its superior convergence behaviour.

Introduction: Blind equalisation and deconvolution have numerous applications in communications, geophysical exploration, image processing, and array processing [1]. In blind equalisation, a complex-valued signal $x(k)$ is assumed to be generated from a linear, time-invariant filter as

$$x(k) = \sum_{i=-\infty}^{\infty} h_i s(k-i)$$

where $s(k)$ is the unknown source signal and $h_i, i = -\infty < i < \infty$ is the impulse response of the unknown filter. The task is to process the signal $x(k)$ with a second filter such that $c(k - \Delta)$ can be recovered, where $c$ and $\Delta$ are complex-valued scaling and real-valued time-delay factors, respectively. It is assumed that the unknown source signal is i.i.d., such that $E_r(x(k)c^*(k)) = E_r(c(k)x^*(k))$ if $i \neq k$, where $*$ denotes complex conjugate.

In the so-called Bussgang methods [2], the coefficient vector $W(k) = [w_1(k), \ldots, w_m(k)]$ is updated as

$$W(k+1) = W(k) + p(k)f(y(k))X(k)$$

where $Y = \text{diag}(W(k))$ is the linear output of the equaliser, $X = \text{diag}(x^2(k))$. The input signal vector is the step size at time $k$, $X$ denotes the transpose of the vector $X$, and $f(y(k))$ is a nonlinearity that depends on the statistics of $y(k)$. The Bussgang methods include the Sato and Godard or constant modulus algorithms as specific cases. While efforts have focused on characterising the stationary points of these algorithms [3, 4], an important practical issue is their relatively slow convergence speeds, as Bussgang techniques may fail to obtain proper equalisation in a reasonable number of iterations for channels with even modest variations in their amplitude spectra. Quasi-Newton techniques in this letter are used to increase an equaliser's convergence speed, but the overall system is computationally complex and difficult to implement in real-time hardware.

In this Letter, we propose a simple modification of the algorithm in eqn. 2 that dramatically improves the convergence properties of the system for ill-conditioned channels. The technique uses the estimate $\hat{W}(k)$ of the inverse of the channel to equalise the modes of convergence of the system, such that the system's behaviour is similar to those of quasi-Newton schemes near the convergence point of the equaliser. Simulations using the constant-modulus algorithm (CMA) cost function suggest that the modified CMA equaliser can acquire the approximate inverses of arbitrary non-minimum phase channels in many fewer iterations as compared to the standard CMA equaliser.

Algorithm: The proposed algorithm is inspired by techniques to improve the convergence properties of multichannel blind separation algorithms for instantaneous signal mixtures [6]. The modified algorithm is described by the update equation

$$W(k+1) = W(k) + \frac{p}{\sigma} \frac{\lambda}{\|Y(k)\|} f(y(k-L)Y(k))(3)$$

where $Y(k) = [y(k), \ldots, y(k-L+1)]^T$ contains samples from the equaliser output, $\sigma$ is a small constant, and $\lambda = \|a(k) - \ldots, a(k-L)^T\|$ contains samples computed as

$$\lambda(u(i)) = \sum_{i=0}^{L-1} w_{L-1}(k)y^*(k-i)$$

Note that this algorithm is computationally-simple, requiring an additional $(n+1)$ complex multiplies and a single divide over the $2L+3$ complex multiplies used in the original Bussgang algorithm if a recursive FIR filter is used to estimate $\hat{Y}(k)$.

To understand the nature of the vector $U(k)$, consider a case where $\mu$ is small such that the vector $\hat{W}(k)$ is nearly constant with time. In this case, we see that

$$\lambda(u(k)) = \sum_{i=0}^{L-1} w_{L-1}(k)y^*(k-i)$$

where $z^*$ is the delay operator, $W(z) = \sum_{i=0}^{L-1} w_i z^{-i}$, and $W(z) = \sum_{i=0}^{L-1} w_i^* z^i$. Near convergence, we find

$$\frac{1}{\lambda}(W(z)z^{-1})f(z) = \sum_{i=0}^{L-1} w_i z^{-i} e^{-\mu x_i}$$

Let $S_{\mu}(\omega)$ be the discrete-time Fourier transform of the stationary input autocorrelation sequence $r(k) = E[x(k)x^*(k-n-k)]$. Therefore, the vector $U(k)$ can be viewed as a rough approximation to the modified regressor vector $R(\hat{X}(k))$, where $R(\hat{X}(k)) = E[\hat{X}(k)X(k)^*]/\lambda$. Such an argument can be made stronger by noting that as $\lambda \to \infty$, multiplication of $U(k)$ by $R^{-1}(\mu)$ is equivalent to filtering the signal $x(k)$ by a filter whose frequency response is $S_{\mu}(\omega)$. We introduce an L-sample delay in the signal $R(k)$ in eqn. 3 because the time-varying FIR filter relating $x(k)$ and $y(k)$ is linear phase with a constant group delay of L samples, irrespective of the value of $\hat{W}(k)$. Therefore, delaying the signal $r(k)$ by L samples in the algorithm maintains proper phase coherence of this signal with the filtered regressor vector. Filtered-regressor algorithms are robust and are used in numerous feedback adaptive control schemes [8]. Although we have not proven the asymptotic stability of our modified blind equalisation scheme, the formulation of the algorithm is technically sound, and our modified algorithm has never diverged in any of our simulation tests for suitably small step size values. Since the inverse channel estimate may not be accurate initially, we use normalisation of the step size using the $L$-norm of the output vector. Such a choice is motivated from quasi-Newton techniques as well, as $\|y(k)^2\| = \lambda E[|X(k)|^2]$. Near equaliser convergence for large $L$, the use of blind equalisation and deconvolution schemes, the choice of the initial coefficient vector $W(0)$ affects the convergence behaviour of the system. For our tests, choosing

$$\lambda(0) = \delta_1$$

where $j$ is a suitable value in the range $0 \leq j \leq L$, led to adequate convergence of the system.

Simulations: We now compare through simulation the performance of the modified algorithms in eqn. 3 with both the algorithm in eqn. 2 and a quasi-Newton version of eqn. 2, in which the regressor vector in the update is replaced by $R(\hat{X}(k))$, where

$$R(\hat{X}(k)) = \frac{Y(k)X(k)^*}{\|Y(k)X(k)^*\|}$$

and $\lambda = 0.995$. For these simulations, the chosen length of each of the equalisers is $L+1 = 11$ coefficients, and each equaliser uses the constant-modulus cost function nonlinearity given by $f(y) = (1-j\lambda N(y))$. In these tests, $\delta(0)$ is chosen as a quadratic-saturation-modulated (QAM) signal with unity modulus, in which case $A = 1$ for the algorithms, and $\lambda(0)$ is generated using a non-minimum phase channel with transfer function

$$H(z) = \frac{e^\mu}{1-0.7e^{-z^*-1}}$$

such that exact equalisation with a finite-length equaliser is not possible, a realistic situation [4]. For the standard CMA algorithm, we have chosen $\mu(0) = \frac{\hat{\lambda}}{1 + \|\hat{X}(k)\|^2}$ where $\hat{\lambda} = 0.01$, and

2114 ELECTRONICS LETTERS 7th November 1996 Vol. 32 No. 23