Concatenated fibre–wireless channel identification in a multiuser CDMA environment

Article *in* IET Communications · November 2007

DOI: 10.1049/iet-com:20045143 · Source: IEEE Xplore

2 authors, including:

Xavier N Fernando
Ryerson University
256 PUBLICATIONS **2,146** CITATIONS

SEE PROFILE

Concatenated fibre-wireless channel identification in a multiuser CDMA environment

S.Z. Pinter and X.N. Fernando

Abstract: Radio-over-fibre (ROF) has received increasing attention for its ability to enable broadband wireless access. This fibre-based wireless access scheme meets the demand for broadband service by integrating the high capacity of optical networks with the flexibility of radio networks (the optical and wireless channels are concatenated with one another). There are, however, impairments that come with this appealing technology. The nonlinear distortion of the optical link and the multipath dispersion of the wireless channel are two of the major factors. In order to limit the effects of these distortions, estimation, and subsequently equalisation, of the concatenated fibre-wireless channel needs to be done. An estimation algorithm for the fibre-wireless uplink in a multiuser code division multiple access (CDMA) environment is presented using pseudonoise training sequences. It has already been shown by Fernando *et al.* (2001) that identification of the fibre-wireless uplink is possible in a single user CDMA environment. However, the more difficult task of identification in a multiuser spread spectrum environment, which is more realistic, is shown. In the multiuser case, the cumulative effect of multiuser interference, multipath dispersion, nonlinear distortion and noise should all be handled together which makes it more challenging. Numerical evaluations of the developed algorithm show that a good estimation of both the linear and nonlinear systems is possible in the presence of 16 independent users and an signal-to-noise ratio (SNR) of 22 dB. The estimation accuracy increases with the length of the PN sequence.

1 Introduction

The increasing demand for high-capacity multimedia services challenges current personal communication systems (PCS) to retain wideband access. New wireless subscribers are signing up at an increasing rate demanding more capacity whereas the radio spectrum is limited. One scheme that has become increasingly popular to alleviate this demand is radio-over-fibre (ROF). ROF, where an optical signal is modulated at radio frequencies and transmitted via an optical fibre [1], provides for an excellent link allowing for high bandwidth communication of several channels.

The fibre-wireless architecture for cellular networks is shown in Fig. 1. In this scenario, there is an intermediate stage between the central base station and the mobile units. This intermediate stage is the optical fibre and the radio access point (RAP in Fig. 1). The RAPs provide wireless access instead of the conventional base station, and are connected to the central base station via the ROF links. Using the RAPs as an intermediate base station allows for a mirco/pico cell scenario. It is important to keep the RAPs complexity, cost and power at a minimum in order to allow for large-scale deployment. By doing so, a large cell can easily be split into smaller cells by dispersing RAPs throughout. This increases frequency reuse and enables broadband access.

doi:10.1049/iet-com:20045143

Paper first received 29th July 2004 and in revised form 22nd January 2007

The authors are with the Department of Electrical and Computer Engineering, Ryerson University, 350 Victoria Street, Toronto, Canada, Ontario M5B 2K3

E-mail: spinter@ieee.org

This technology does not come without its drawbacks. When the wireless link is in series with the ROF link – especially in a multiuser environment – nonlinear distortion of the ROF link, mainly because of the laser diode (and partly to the high-gain radio frequency (RF) amplifier at the optical receiver), becomes the primary concern [2]. Also, the dynamic range of the input signal is adversely affected in the fibre-wireless uplink where the received signal first travels through the wireless channel (resulting in path losses, fading, shadowing) before entering the optical (ROF) link. Because of the mutually coupled nature of these impairments, it is important to jointly estimate the two concatenated systems so that proper equalisation techniques can be applied. The objective is estimation of both the wireless and optical channels independently.

Estimation of the fibre-wireless uplink is considered and so the system of interest consists of a linear part (wireless channel) followed by a mildly nonlinear part (ROF link), in that particular order. (We do not assume an AWGN wireless channel (as in [2]), but instead consider a multipath wireless channel. This is more realistic with wideband access [3]. Clipping distortion is not considered, it is left for future work.) Fortunately, extensive work has been done with similar classes of systems that have shown linear/nonlinear systems modelled as a Wiener system [4–6]. In the aforementioned papers, the Wiener model was analysed in a single control signal (or single user) continuous-time baseband environment. Correlation analysis was used to decouple the identification of the linear and nonlinear component subsystems by using maximal-length pseudonoise (PN) sequence inputs (i.e. white noise-like inputs) [6]. Other estimation methods have also been established in the literature. Fang and Chow [7] used an orthogonal wavelet-based neural network (OWNN) to

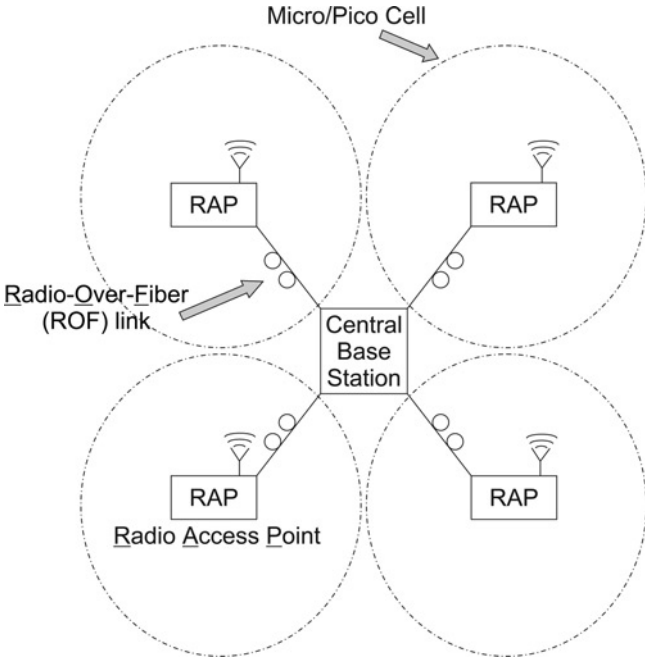

Fig. 1 *Fibre-wireless cell architecture*

identify a Wiener-type cascade model. Prakriya and Hatzinakos [8] showed blind identification of linear subsystems of Wiener-Hammerstein models with cyclostationary inputs. Tan and Godfrey [9] identified a Wiener-Hammerstein model in the frequency domain.

The work in [10] (where identification of a Wiener system is done using single PN sequences) is extended such that a Wiener system is identified in a discrete-time passband communication system with multiple maximal-length PN sequence inputs. The extension to multiple users is clear since real world systems do not operate with single users. In addition, the use of PN sequences in the estimation is attractive because these spreading codes are already widely used in spread spectrum communications [11]. In a multiple PN sequence environment, it is essential that the correlation properties of PN sequences be well understood. Fortunately, PN sequences have well-defined correlation properties [12–14]. One of the most important properties of a bipolar PN sequence is its excellent periodic autocorrelation. It is important to have this optimal autocorrelation otherwise there will be multiple identifications (further discussed in Section 3.3).

This paper is organised as follows. Section 2 discusses the issue of passband and baseband systems in a linear and nonlinear environment and how they relate to our paper. Section 3 provides estimation theory for the wireless channel. The correlation between the output and compound input and the correlation in the presence of multiple PN sequences are used to aid the discussion. The fibre channel estimation is performed in Section 4 using the least mean squares criteria. Simulation results and discussion as well as the performance of our identification in a hostile environment are presented in Section 5. In our simulations both the wireless channel and fibre link are identified in the presence of Gaussian noise (SNR of 22 dB). Section 6 is our conclusion.

2 Passband complex consideration

Communication signals and systems are passband. In order to use baseband signal processing, communication signals in the passband (i.e. real-valued signals [15]), must be appropriately translated from the passband to the baseband. Generally, this translation results in complex-valued baseband signals [15]. Therefore in a fibre-wireless passband system, the signals, as well as the channel impulse response and nonlinear component are complex-valued. We now show how these complex-valued quantities can be split into real-valued quadrature components.

When an RF signal undergoes a nonlinear transformation one of the major concerns is the amplitude-to-amplitude modulation (AM-AM) and amplitude-to-phase modulation (AM-PM) distortions. The complex-valued nonlinear fibre link in Fig. 2*a* introduces both of these distortions [2]. It has been shown in [16] and [17] that these nonlinear distortions can be expressed by inphase and quadrature phase components. Let the input signal in Fig. 2*a* be given as

$$q(t) = A(t)\cos[\omega_c t + \theta(t)] \tag{1}$$

Then the output $r(t)$ is

$$r(t) = R[A(t)]\cos\{\omega_c t + \theta(t) + \phi[A(t)]\} \tag{2}$$

where R is the AM-AM distortion and ϕ is the AM-PM distortion. The output $r(t)$ can also be expressed as

$$\begin{aligned} r(t) = {} & R[A(t)]\cos(\phi[A(t)])\cos(\omega_c t + \theta(t)) \\ & - R[A(t)]\sin(\phi[A(t)])\sin(\omega_c t + \theta(t)) \end{aligned} \tag{3}$$

using the trigonometric identity $\cos(A+B) = \cos(A)\cos(B) - \sin(A)\sin(B)$. Equation (3) can then be written as

$$r(t) = r_p[A(t)]\cos(\omega_c t + \theta(t)) - r_q[A(t)]\sin(\omega_c t + \theta(t)) \tag{4}$$

where

$$r_p[A(t)] = R[A(t)]\cos(\phi[A(t)]) \tag{5}$$

$$r_q[A(t)] = R[A(t)]\sin(\phi[A(t)]) \tag{6}$$

Equation (4) shows that the bandpass nonlinearity can be separated into an inphase component and a quadrature phase component with only AM-AM distortion. Therefore instead

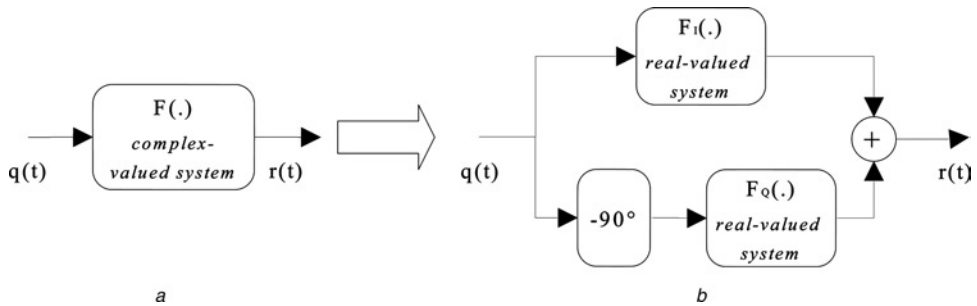

Fig. 2 *Quadrature model of a nonlinear system*

of dealing with the complex-valued system in Fig. 2*a* we can deal with two real-valued systems as shown by the quadrature model in Fig. 2*b*. Similarly, the bandpass impulse response $h(n)$ can also be separated into its inphase and quadrature phase components [15].

Therefore by using the quadrature model for both the linear and nonlinear systems, it can be stated that we estimate the real-valued inphase and quadrature phase components individually. Hence, all the variables introduced hereafter are real quantities unless otherwise specified.

3 Wireless channel estimation theory

A model of the fibre-wireless uplink is shown in Fig. 3. This block diagram represents one branch in the architecture of Fig. 1. This class of system consists of a wireless channel (linear system with impulse response $h(n)$) in cascade with an optical fibre link (nonlinear system with a function $F(\cdot)$).

The first step in estimating the wireless channel of the fibre-wireless uplink is to define the output of the system. According to the theorem of Weierstrass [5], any function which is continuous within an interval may be approximated to any required degree of accuracy by polynomials in this interval. Therefore the output of the nonlinear system plus the noise can be represented by a polynomial of the form

$$r(n) = A_1 q(n) + A_2 q^2(n) + \cdots + A_l q^l(n) + v(n) \quad (7)$$

where $v(n)$ is the total noise of both the optical and wireless channels. The system output $r(n)$ can then be expressed by the functional (Volterra) series [4]

$$\begin{aligned} r(n) = {} & A_1 \sum_{m_1=-\infty}^{\infty} h(m_1)u(n-m_1) \\ & + A_2 \sum_{m_1=-\infty}^{\infty}\sum_{m_2=-\infty}^{\infty} h(m_1)h(m_2)u(n-m_1)u(n-m_2) \\ & + A_l \sum_{m_1=-\infty}^{\infty} \cdots \sum_{m_l=-\infty}^{\infty} \prod_{i=1}^{l} h(m_i)u(n-m_i) + \cdots + v(n) \end{aligned} \quad (8)$$

where $u(n)$ is a compound input that can be written as

$$u(n) = x_1(n) + x_2(n) + \cdots + x_N(n) \quad (9)$$

where N is the number of PN sequences (or equivalently the number of users). It is at this point where we are considering multiple users. In [10], $u(n)$ was simply taken as a single PN sequence but $u(n)$ is taken as a summation of multiple independent PN sequences $x(n)$ of period N_c where

$$N_c = 2^n - 1 \qquad n \equiv \text{degree of the PN polynomial} \quad (10)$$

(details of PN properties will be discussed in Section 3.3). The output can also be written as a summation of the isolated lth order kernel as

$$r(n) = w_1(n) + w_2(n) + w_3(n) + \cdots + w_l(n) + v(n) \quad (11)$$

where

$$w_l(n) = A_l \sum_{m_1=-\infty}^{\infty} \cdots \sum_{m_l=-\infty}^{\infty} \prod_{i=1}^{l} h(m_i)u(n-m_i) \quad (12)$$

Therefore the output $r(n)$ consists of a compound input $u(n)$ that has been dispersed in time because of the impulse response and then raised to higher order powers because of the nonlinearity.

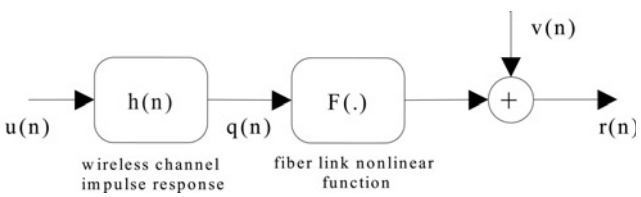

Fig. 3 *Fibre-wireless uplink modelled as a Wiener system*

3.1 Correlation relationship between the output and compound input

The next step in the estimation is to further process the output and input relations, as defined above, by utilising correlation relationships. The cross covariance between the output and the compound input is given as

$$\Re_{ur}(\sigma) = \overline{(r(n) - \overline{r(n)})(u(n-\sigma) - \overline{u(n-\sigma)})} \quad (13)$$

From this point onward let $r(n)$, $u(n)$, $q(n)$ and $v(n)$ refer to their respective signals with the mean removed. The cross covariance can then be written as

$$\Re_{ur}(\sigma) = \overline{r(n)u(n-\sigma)} \quad (14)$$

Substituting (8) into the above equation gives

$$\begin{aligned} \Re_{ur}(\sigma) = \Bigg[& \overline{A_1 \sum_{m_1=-\infty}^{\infty} h(m_1)u(n-m_1)} \\ & + \overline{A_2 \sum_{m_1=-\infty}^{\infty}\sum_{m_2=-\infty}^{\infty} h(m_1)h(m_2)u(n-m_1)u(n-m_2)} \\ & + \overline{A_l \sum_{m_1=-\infty}^{\infty} \cdots \sum_{m_l=-\infty}^{\infty} \prod_{i=1}^{l} h(m_i)u(n-m_i)} \\ & + \overline{\cdots + v(n)} \Bigg] \overline{[u(n-\sigma)]} \end{aligned} \quad (15)$$

which simplifies to

$$\Re_{ur}(\sigma) = \sum_{i=1}^{l} \Re_{uw_i}(\sigma) + \Re_{uv}(\sigma) \quad (16)$$

where the term $\Re_{uv}(\sigma)$ becomes zero assuming that the compound input and noise process are statistically independent, $\overline{v(n)u(n-\sigma)} = 0 \forall \sigma$. If $\Re_{ur}(\sigma)$ is evaluated directly as defined above, the terms $\sum_{i=2}^{l} \Re_{uw_i}(\sigma)$ give rise to anomalies associated with multidimensional autocovariances of PN sequences [4]. This problem can be overcome by isolating $\Re_{uw_1}(\sigma)$ using multilevel input testing. It should be noted that if the channel were linear there would be no need to isolate $\Re_{uw_1}(\sigma)$ because $\Re_{uw_1}(\sigma) = \Re_{ur}(\sigma)$.

3.2 Multilevel testing

Multilevel testing enables the extraction of $\Re_{uw_1}(\sigma)$ from $\Re_{ur}(\sigma)$. This step is crucial for successful estimation of the wireless channel. Multilevel testing is implemented at the RAP by using the signal $\alpha_i u(n)$ where $\alpha_i \neq \alpha_l \forall i \neq l$.

With multilevel inputs (16) becomes

$$\Re_{ur_{\alpha_i}}(\sigma) = \sum_{j=1}^{l} \alpha_i^j \Re_{uw_j}(\sigma) \quad i = 1, 2, \ldots, N_t \tag{17}$$

where N_t is the total number of multilevel inputs and hence the number of equations to solve. Representing (17) in matrix form gives

$$\begin{bmatrix} \Re_{ur_{\alpha_1}}(\sigma) \\ \Re_{ur_{\alpha_2}}(\sigma) \\ \cdot \\ \cdot \\ \Re_{ur_{\alpha_{N_t}}}(\sigma) \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_1^2 & \cdot & \cdot & \alpha_1^l \\ \alpha_2 & \alpha_2^2 & \cdot & \cdot & \alpha_2^l \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \alpha_{N_t} & \alpha_{N_t}^2 & \cdot & \cdot & \alpha_{N_t}^l \end{bmatrix} \times \begin{bmatrix} \Re_{uw_1}(\sigma) \\ \Re_{uw_2}(\sigma) \\ \cdot \\ \cdot \\ \Re_{uw_l}(\sigma) \end{bmatrix} \tag{18}$$

For every value of σ, (18) has a unique solution for $\Re_{uw_i}(\sigma)$, $i = 1, 2, \ldots, N_t$. (Refer to [4] for an explanation on this property of the matrix.) In order to successfully identify the linear system, the number of multilevel inputs N_t had to be equal to or greater than the order of the polynomial. This ensured that the correct linear system could be identified in the presence of any nonlinear function.

Now that $\Re_{uw_1}(\sigma)$ can be extracted, the final step in the identification process is to find how $\Re_{uw_1}(\sigma)$ relates to the channel impulse response.

3.3 Correlation relationship in the presence of multiple PN sequences

In order to accommodate for multiple PN sequences, the relationship in [4] had to be reworked. Initially, we derived the cross covariance relationship $\Re_{x_1x_2\cdots x_Nw_N}(\sigma)$; however, this result was undesirable because of the dependency on higher order kernels w_N. Another approach was to consider the cross covariance between the compound input $u(n)$ and $w_1(n)$, given as

$$\Re_{uw_1}(\sigma) = \overline{w_1(n)u(n-\sigma)} \tag{19}$$

where $w_1(n)$ is a zero mean process. Substituting (9) and $w_1(n)$ from (12) into (19) gives the following

$$\Re_{uw_1}(\sigma) = A_1 \sum_{m_1=-\infty}^{\infty} h(m_1) \times \overline{[x_1(n-m_1) + x_2(n-m_1) + \cdots + x_N(n-m_1)]} \; \overline{[x_1(n-\sigma) + x_2(n-\sigma) + \cdots + x_N(n-\sigma)]} \tag{20}$$

As an aside, another possibility for the above covariance was to consider the cross covariance between $w_1(n)$ and individual users inputs, instead of the compound input $u(n)$. However, it was found through simulations that using the compound input gave more accurate results. The reason being that a compound input gives N terms where two identical PN sequences are correlated, compared to only one term when individual inputs are used. Overall, having N terms results in better covariance properties as will be shown by the next few equations.

Expanding (20) gives two different types of terms: (1) where two identical PN sequences are multiplied

$$\overline{x_i(n-m_1)x_{j(j=i)}(n-\sigma)} \tag{21}$$

and (2) where two different PN sequences are multiplied

$$\overline{x_i(n-m_1)x_{j(j\neq i)}(n-\sigma)} \tag{22}$$

The outcome of the first case is the circular autocovariance which is given in [13] as

$$\Re_{xx}(\sigma) = \begin{cases} N_c & \text{if } \sigma = 0 \\ -1 & \text{if } \sigma \neq 0 \end{cases} \tag{23}$$

Clearly, the autocovariance properties of maximal-length PN sequences are very good. As mentioned above, using a compound input gives the summation of N autocovariance terms. This yields a large autocovariance value and hence improves the accuracy of identification. It is important to have this optimal autocovariance, otherwise there will be multiple identifications (one at each autocovariance peak).

The outcome of the second case is the cross covariance. The cross covariance properties of PN sequences are limited and it is known that m-sequences can have relatively large cross covariance peaks. However, numerical analysis showed that the peak cross covariances of the PN sequences in our simulations had typical peak values as shown in Table 1. In the simulations, a large number of PN sequences were used and peak values were recorded.

As mentioned above, the 'simulated peak cross covariance' column in Table 1 represents the peak cross covariance encountered for a specific PN sequence found through our simulations. Note that our simulation values in Table 1 are close to the cross correlation values of Table 1, column 8 of [12]. The row showing the peak cross covariance for $n = 10$ in Table 1 is most significant because it shows the maximum cross covariance value encountered during our simulations. For $n = 10$, the peak cross covariance can be presented as being 13.98% of N_c. Clearly, the cross covariance value is much smaller than the autocovariance value and below follows a discussion of why these cross covariance values can be neglected in the derivation.

There are three reasons why the cross covariance terms can be omitted in the derivation. The first is that the cross covariance values are small after removing the most offensive decimations. Since the actual number of users will always be less than the maximum possible, it will always be possible to remove the most offensive decimation. For example, we considered only 16 users with $N_c = 1023$. This removal of the most offensive PN sequences phenomenon significantly decreases the cross covariance value. Note that column 8 of [12] (after removal of the decimation

Table 1: Simulated peak cross covariance values of PN sequences

n	N_c	Number of PN sequences	Simulated peak cross covariance
5	31	6	7
6	63	6	3
7	127	18	15
8	255	16	63
9	511	48	79
10	1023	60	143

leading to the most significant bias) has much smaller cross correlation values compared to the values without the removal.

Secondly, because we have multiple users our autocovariance is actually in the order of $(N \times N_c)$; this autocovariance is relatively large compared to the cross covariance terms, thereby making it easy to differentiate between actual and erroneous impulse response peaks in the final cross covariance relationship. Thirdly, since cross covariance terms take on both positive and negative values, cancellations will be present and therefore the cross covariance values will not deviate much from the values listed in Table 1. Although the probability of every single cross covariance to be at its absolute worst case and all cross covariance values to have the same sign is very low, there is still the chance of this scenario occurring, but evaluating the probabilities of the occurrences of these rare cases and compensating for them is left for future work. As the number of users is increased from 1 to N, there will be varying effects on the cross covariance terms. For example, in the worst case scenario the maximum value of the cross covariance would be in the order of $(N) \times (N-1) \times$ (cross covariance value). When there are a low number of users, the worst case maximum cross covariance defined above is still much smaller than the autocovariance value, which is given by $(N \times N_c)$. However, as the number of users is significantly increased these cross covariance terms become a concern. Compensating for these effects shall be considered in a separate study.

Therefore identification of a multiuser fibre-wireless channel is possible using maximal-length PN sequences when typical cross covariance values occur and the most offensive decimations are removed, and as long as optimal autocovariance is maintained, as described by (23), the effect of the cross covariance terms will be minimal (although they will introduce some uncertainty into the final estimate). Hence, the cross covariance terms do not impair the identification significantly and it was assumed that the cross covariance between independent PN sequences is negligible when compared to the large autocovariance value. Applying the aforementioned assumption to (20) gives

$$\Re_{uw_1}(\sigma) = A_1 \sum_{m_1=-\infty}^{\infty} h(m_1)(\Re_{x_1x_1}(m_1-\sigma) + \Re_{x_2x_2}(m_1-\sigma) + \cdots + \Re_{x_Nx_N}(m_1-\sigma)) \quad (24)$$

Using the relationship $\Re_{x_ix_i}(\lambda) = N_c\delta_i(\lambda)$ and the convolution properties of the impulse function gives

$$\Re_{uw_1}(\sigma) = A_1N_c \sum_{m_1=-\infty}^{\infty} h(m_1)(\delta_1(m_1-\sigma) + \delta_2(m_1-\sigma) + \cdots + \delta_N(m_1-\sigma)) \quad (25)$$

$$= A_1N_c \sum_{i=1}^{N} h(\sigma) \quad (26)$$

Therefore the final cross covariance relationship can be written as

$$\Re_{uw_1}(\sigma) = A_1N_cNh(\sigma) \quad (27)$$

where N is the number of PN sequences (or users), N_c is the PN sequence length and A_1 is the linear gain of the nonlinear system. The estimated channel impulse response can be found by solving the above expression.

4 Fibre channel estimation theory

As mentioned in Section 1, the linear and nonlinear system identifications are independent in the presence of PN sequence inputs. Therefore a least squares polynomial fit is sufficient to identify the nonlinear system. The least squares polynomial fit is performed between the input and output of the nonlinear system. In the fibre-wireless channel, there is no access to the internal signal $q(n)$ (the input to the nonlinear system) and therefore it must be estimated. Referring to Fig. 3, this internal signal can be estimated by convolving the final impulse response estimate $\tilde{h}(n)$ (obtained from the previous section) with the compound input $u(n)$, giving an estimate of the internal signal, $\tilde{q}(n)$. The least squares polynomial fit is then applied to $\tilde{q}(n)$ and the measured output $r(n)$. So, let the estimated polynomial coefficients be given as

$$\hat{\boldsymbol{A}} = [\hat{A}_0 \quad \hat{A}_1 \quad \hat{A}_2 \quad . \quad . \quad . \quad \hat{A}_l]^{\mathrm{T}} \quad (28)$$

The estimated signal $\hat{r}(n)$ of the output of the nonlinear system is then given as

$$\hat{r}(n) = \hat{A}_0 + \hat{A}_1\tilde{q}(n) + \hat{A}_2\tilde{q}^2(n) + \cdots + \hat{A}_l\tilde{q}^l(n) + v(n) \quad (29)$$

Let us define vectors $\tilde{\boldsymbol{q}}$ and $\boldsymbol{r}$ of length N_{L}, which are made up of the signals $\tilde{q}(n)$ and $r(n)$, respectively. The Vandermonde matrix $\boldsymbol{V}$ can then be defined such that each row of $\boldsymbol{V}$ is a polynomial of the corresponding data point in $\tilde{\boldsymbol{q}}$

$$\boldsymbol{V} = \begin{bmatrix} 1 & \tilde{q}(0) & \tilde{q}^2(0) & \tilde{q}^3(0) & \cdots & \tilde{q}^l(0) \\ 1 & \tilde{q}(1) & \tilde{q}^2(1) & \tilde{q}^3(1) & \cdots & \tilde{q}^l(1) \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & \tilde{q}(N_{\mathrm{L}}-1) & \tilde{q}^2(N_{\mathrm{L}}-1) & \tilde{q}^3(N_{\mathrm{L}}-1) & \cdots & \tilde{q}^l(N_{\mathrm{L}}-1) \end{bmatrix} \quad (30)$$

In matrix notation, the equation for a polynomial fit is given by

$$\hat{\boldsymbol{r}} = \boldsymbol{V}\hat{\boldsymbol{A}} \quad (31)$$

Premultiplying by the matrix transpose $\boldsymbol{V}^{\mathrm{T}}$ and solving for $\hat{\boldsymbol{A}}$ gives

$$\hat{\boldsymbol{A}} = (\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V})^{-1}\boldsymbol{V}^{\mathrm{T}}\hat{\boldsymbol{r}} \quad (32)$$

The error between the actual data $\boldsymbol{r}$ and the estimated data $\hat{\boldsymbol{r}}$ is given by $e = \boldsymbol{r} - \boldsymbol{V}\hat{\boldsymbol{A}}$. Now, the order of the polynomial l must be selected to minimise the mean square error.

Once the impulse response has been estimated, the estimation of the nonlinear channel is straightforward. However, the accuracy of the nonlinear identification is highly dependent on the impulse response estimates, and so it is important that the wireless channel estimation algorithm work well.

5 Simulation results and discussion

The Simulink™ model shown in Fig. 4 was used to perform the multilevel testing on the fibre-wireless system.

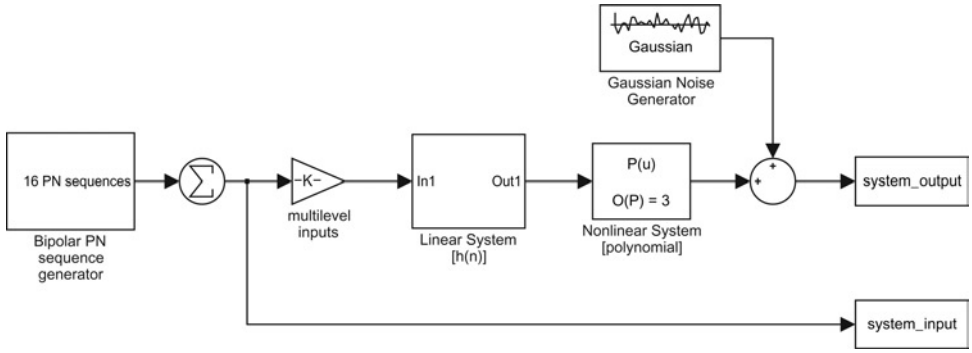

Fig. 4 *Simulink model*

The majority of the calculations were performed in MATLAB^{TM} (including all correlations) by sending the Simulink^{TM} outputs to the MATLAB^{TM} workspace. The next few sentences describe the individual blocks of the above Simulink^{TM} diagram. The 'multilevel inputs' block is a variable gain that changes depending on which multiple level is being simulated. The 'linear system' block simulates the impulse response of the wireless channel. The 'nonlinear system' block implements a polynomial expression on the input. The data recorded from the simulation is 'system_input' and 'system_output', from which we perform the identification.

5.1 Noise

Before presenting the simulation outcomes, we will briefly discuss our inclusion of Gaussian noise $v(n)$. It is important to consider the noise because in a practical environment noise from both the optical and wireless channels will be present in the output $r(n)$. Noise was not considered in [10]; however, we have included noise as well in this paper. The simulation was performed with varying SNRs to see how noise affected the accuracy of the polynomial estimation. Fig. 5 shows the mean square error between the actual and estimated polynomials (for a third-order odd polynomial) against the SNR. Using Fig. 5 it was found that an SNR of 22 dB (i.e. a Gaussian noise variance of 0.1262) still gave an acceptable identification and was therefore used in our simulations to show the effect of noise.

The plot of mean square error against SNR using a fifth-order (odd) polynomial was practically identical to the one above and therefore the same SNR was used throughout the whole simulation procedure. Although theoretically the noise is not correlated with the signal, the mean square error becomes very large for SNR $\leq$ 15 dB. This shows the practical limitation of the scheme.

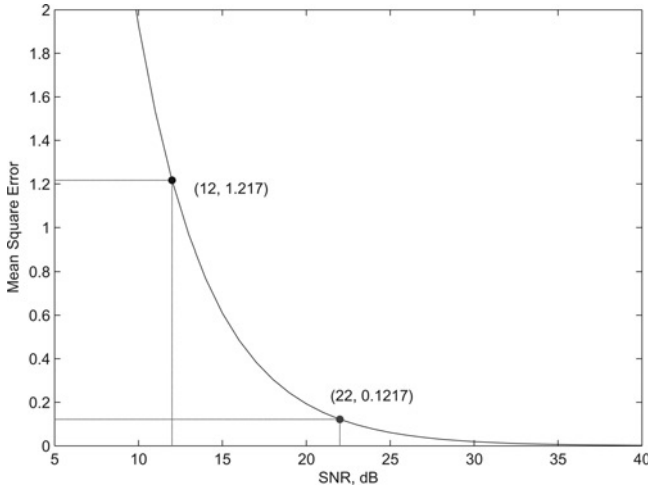

Fig. 5 *Mean square error against SNR using polynomial (35)*

5.2 Wireless channel identification

The simulations were performed with: (1) 16 PN sequences, each with a length of 1023 ($2^{10} - 1$), and (2) an SNR of 22 dB (selection of this SNR was discussed above). The sequence length was chosen based on various trial identifications. Inaccurate results were obtained at $N_c = 31$ and 63; however, better results were obtained starting at $N_c = 127$. Only at around $N_c = 511$ and 1023 was there a significant distinction between the actual impulse response and the erroneous impulse response peaks. From this observation, it can be stated that the longer the PN sequence, the better the covariance properties and hence the better the identification becomes. The ideal case is to have zero cross covariance and perfect autocovariance between all PN sequences. This would result in a perfect identification of the wireless channel.

The wireless channel was identified in the presence of a third-order polynomial; therefore the three input levels chosen were $\alpha = 1.0, 1.2, 1.4$. The wireless channel identification could have been performed just as accurately using any polynomial as long as the proper number of levels (α) was chosen. Figs. 6 and 7 show the estimated and actual impulse responses. Fig. 6 shows the identification using the impulse response

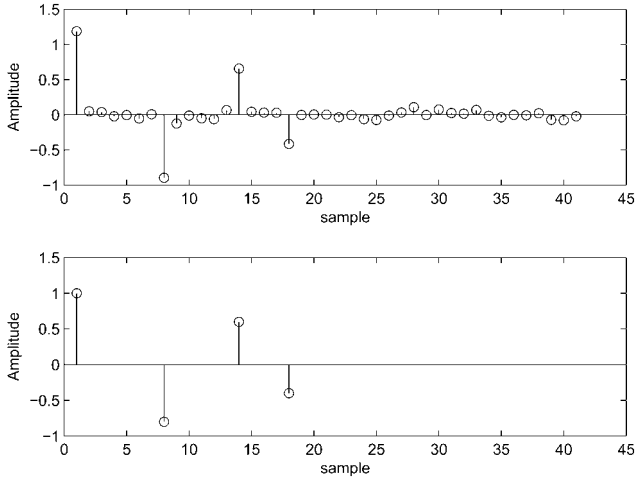

Fig. 6 *Estimated (top) and actual (bottom) impulse response ($h_1(n)$) of the wireless channel using 16 independent PN sequences, SNR = 22 dB, $N_c = 1023$*

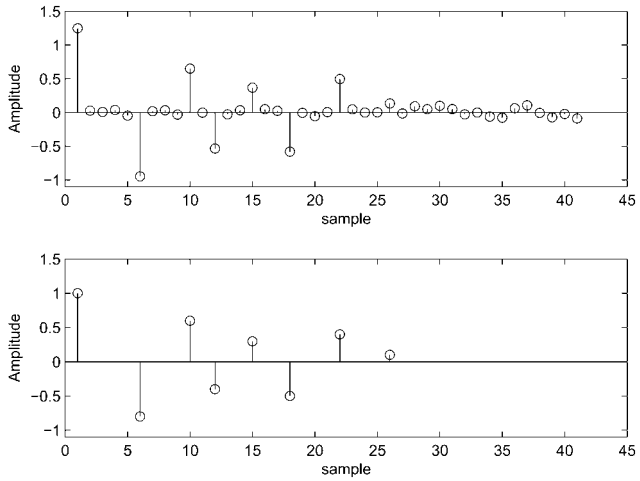

Fig. 7 *Estimated (top) and actual (bottom) impulse response ($h_2(n)$) of the wireless channel using 16 independent PN sequences, SNR = 22 dB, $N_c = 1023$*

$$h_1(n) = \delta(n) - 0.8\delta(n-7) + 0.6\delta(n-13) - 0.4\delta(n-17) \quad (33)$$

Fig. 7 shows the identification using the impulse response

$$h_2(n) = \delta(n) - 0.8\delta(n-5) + 0.6\delta(n-9) - 0.4\delta(n-11) + 0.3\delta(n-14) - 0.5\delta(n-17) + 0.4\delta(n-21) + 0.1\delta(n-25) \quad (34)$$

It is important to mention that each of the multiple PN sequences was generated using a separate maximal-length polynomial. This is in contrast to the common technique (used in current CDMA systems) of using delayed versions of a single maximal-length PN sequence to represent separate users. Using this technique resulted in multiple identifications at the locations of cross covariance peaks. This problem was solved by using separate polynomials to generate the multitude of PN sequences.

5.3 Fibre link identification

The results of the simulation for the fibre channel are shown in Figs. 8 and 9 for the nonlinear systems

$$r(n) = -0.35q^3(n) + q(n) \quad (35)$$

and

$$r(n) = 0.6q^5(n) + q(n) \quad (36)$$

Only odd order polynomials are relevant in our identification because we are performing the identification in the passband. Two important observations regarding the output of a nonlinear system in the passband can be made from [3, 18], they are

1. only odd order terms contribute to the system output after the bandpass filter, and
2. even order terms appear out of band and can easily be filtered out.

The accuracy of the nonlinear identification greatly depends on how much data are available. Specifically, we would like to have the input data cover a large dynamic range. In our case, we had the advantage of

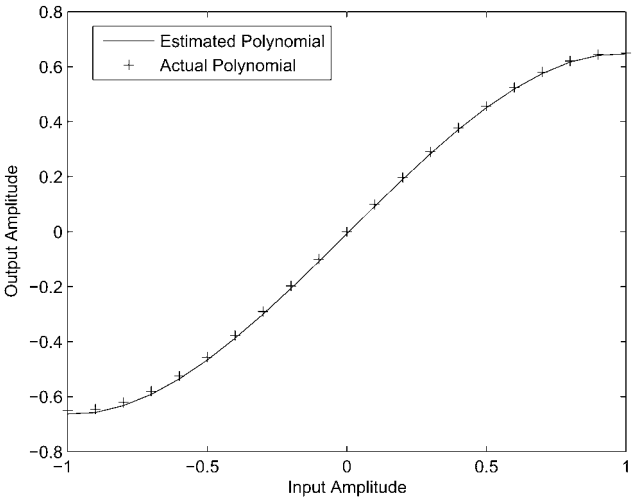

Fig. 8 *Nonlinear system identification using polynomial (35), SNR = 22 dB*

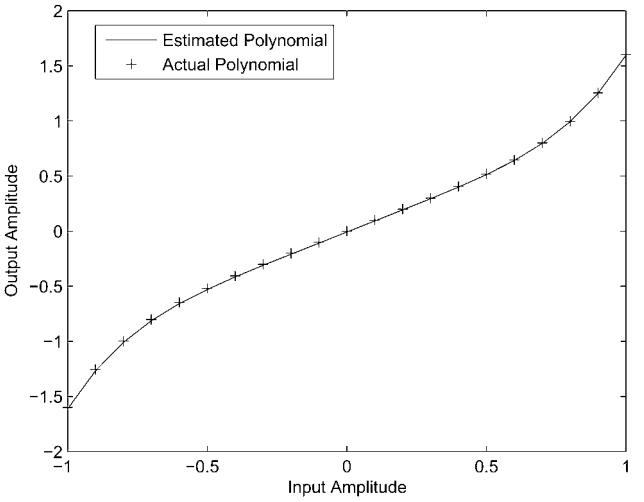

Fig. 9 *Nonlinear system identification using polynomial (36), SNR = 22 dB*

1. a PN sequence length of 1023, and
2. a linear channel with multipath conditions,

both of which contributed to a large dynamic range available for the least squares fit. (Also, once $h(n)$ is known we can send any signal to identify the nonlinear system, that is, one that covers a large dynamic range.) The estimated polynomials were

$$r(n) = -0.35q^3(n) + 1.005q(n) - 0.0079 \quad (37)$$

and

$$r(n) = 0.6q^5(n) + 1.005q(n) - 0.0079 \quad (38)$$

which are very close to the actual polynomials of (35) and (36). The mean square error was 0.1217 for both of the above cases. These estimates are obviously quite good; however, the accuracy of the identification deteriorates with decreasing SNR.

It is important to state that the nonlinear system must have a linear coefficient A_1. Otherwise, identification of neither the linear nor nonlinear channel is possible.

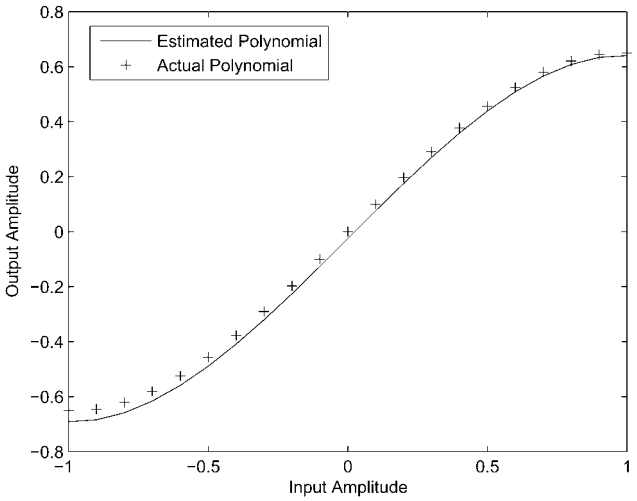

Fig. 10 *Nonlinear system identification using polynomial (35), SNR = 12 dB*

5.4 Hostile noise environment

In this section, we perform the identification using impulse response (33) and polynomial (35) in a hostile environment. We use an SNR of 12 dB to show this effect. Estimations of both the impulse response and nonlinear channel are adversely affected by the lower SNR; the nonlinear channel more so than the linear channel. In this special case, we make use of the fact that in (16) it is assumed that the compound input and noise process are statistically independent. Through simulation it was found that $\Re_{uv}(\sigma)$ was much smaller relative to $\Re_{uw_1}(\sigma)$ and consequently the impulse response estimation was not affected as much as the polynomial estimation. In case of the impulse response, the estimated values become slightly larger than those shown in Fig. 6, which is not terribly detrimental. However, in case of the polynomial, the estimated polynomial starts to diverge away from the actual polynomial as shown in Fig. 10. The mean square error for the polynomial in this case is 1.217 (from Fig. 5) which can be compared to a mean square error of 0.1217 using an SNR of 22 dB.

6 Conclusion

This paper presented a method for fibre-wireless system identification using the correlation properties of multiple PN sequences. We improved the single PN identification performed in [10] to accommodate a compound PN input and we showed the effect of noise in the output. Our approach is practical in the sense that multiple PN sequences are already widely used in existing spread spectrum communication systems.

Since we used multiple PN sequences (as opposed to a single PN sequence [10]), we encountered additional terms that introduced error into the estimation. These erroneous terms were the cross covariances in (20), where two different PN sequences are multiplied. Although we took these cross covariances to be negligible, they still introduced some error into the final impulse response estimate, and hence the polynomial estimate. One way our estimation algorithm can overcome this limitation is to increase the PN sequence length. However, this presents a practical limitation because this would increase the training time as well. This is an interesting area for future research.

As shown in the paper, the estimation of the fibre-wireless system was accurate. Our identification technique reveals the effects of distortion in an ROF system and hence provides a good basis for equalisation.

7 References

1 Morita, K., and Ohtsuka, H.: 'The new generation of wireless communications based on fiber-radio technologies', *IEICE Trans. Commun.*, 1993, **E76-B**, (9), pp. 1061–1068
2 Fernando, X.N., and Sesay, A.B.: 'Adaptive asymmetric linearization of radio over fiber links for wireless access', *IEEE Trans. Vehicular Technol.*, 2002, **51**, (6), pp. 1576–1586
3 Fernando, X.N., and Sesay, A.B.: 'A Hammerstein-type equalizer for concatenated fiber-wireless uplink', *IEEE Trans. Veh. Technol.*, 2005, **54**, (6), pp. 1980–1991
4 Billings, S.A., and Fakhouri, S.Y.: 'Identification of nonlinear systems using correlation analysis and pseudorandom inputs', *Int. J. Syst. Sci.*, 1980, **11**, (3), pp. 261–279
5 Billings, S.A., and Fakhouri, S.Y.: 'Identification of a class of nonlinear systems using correlation analysis', *Proc. IEE*, 1978, **125**, (7), pp. 691–697
6 Billings, S.A., and Fakhouri, S.Y.: 'Identification of nonlinear systems using the Wiener model', *Electron. Lett.*, **13**, (17), pp. 502–504
7 Fang, Y., and Chow, T.W.S.: 'Orthogonal wavelet neural networks applying to identification of Wiener model', *IEEE Trans. Circuits Syst. I: Fundam. Theory Appl.*, 2000, **47**, (4), pp. 591–593
8 Prakriya, S., and Hatzinakos, D.: 'Blind identification of linear subsystems of LTIZMNL- LTI models with cyclostationary inputs', *IEEE Trans. Signal Process.*, 1997, **45**, (8), pp. 2023–2036
9 Tan, A.H., and Godfrey, K.: 'Identification of Wiener-Hammerstein models using linear interpolation in the frequency domain (LIFRED)', *IEEE Trans. Instrum. Meas.*, **51**, (3), pp. 509–521
10 Fernando, X.N., and Sesay, A.B.: 'Fibre-wireless channel estimation using correlation properties of PN sequences', invited paper', *Can. J. Electr. Computer Eng.*, 2001, **26**, (2), pp. 43–47
11 Al-Raweshidy, H., and Komaki, S.: 'Radio over fiber technologies for mobile communications networks' (Artech House Publishers, Norwood, MA, 2002, 1st edn.)
12 Tirkel, A.Z.: 'Cross-correlation of M-sequences - some unusual coincidences', IEEE Proc., 4th Int. Symp. on Spread Spectrum Techniques and Applications, 1996, Vol. 3, pp. 969–973
13 Sarwate, D.V., and Pursley, M.B.: 'Crosscorrelation properties of pseudorandom and related sequences', *Proc. IEEE*, 1980, **68**, (5), pp. 593–619
14 Mutagi, R.N.: 'Pseudo noise sequences for engineers', *Electron. Commun. Eng. J.*, 1996, **8**, (2), pp. 79–87
15 Proakis, J.G.: 'Digital communications' (McGraw-Hill, New York, NY, 2001, 4th edn.)
16 Saleh, A.A.M.: 'Frequency independant and frequency dependant nonlinear models of TWT amplifiers', *IEEE Trans. Commun.*, 1981, **29**, pp. 1715–1720
17 Fernando, X.N.: 'Signal processing for optical fiber based wireless access', Ph.D. thesis, University of Calgary, 2001
18 Benedetto, S., and Biglieri, E.: 'Principles of digital transmission with wireless applications' (Kluwer Academic/Plenum, New York, NY, 1991, 1st edn.)