16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

ADAPTIVE OFF-DIAGONAL MIMO CANCELLER (ODMC) FOR VDSL UPSTREAM SELF FEXT MITIGATION

Patrick Duvaut, Pravesh Biyani, Amitkumar Mahadevan, Shailendra Singh, Satyavardhan Maheshwari

CONEXANT Systems Inc.
Red Bank, NJ 07701, USA

phone: + (01) 732-345-6119, fax: + (01) 732-345-7512, email: patrick.duvaut@conexant.com
web: www.conexant.com

ABSTRACT

This contribution presents a linear off-diagonal MIMO canceller (ODMC) that aims at mitigating the upstream self far-end cross-talk (FEXT) that severely impacts the performance of very high speed digital subscriber line (VDSL) services at the central office (CO). The ODMC operates per frequency and is thus well suited for DMT-based VDSL systems. We introduce a low complexity and low latency adaptive algorithm that converges towards the optimum ODMC that simultaneously maximizes the distributed Shannon capacities of the upstream links, under mild assumptions for nominal VDSL systems. The adaptive algorithm does not require any matrix inversion and is designed to learn and engage the ODMC seamlessly while operating in data mode. Further, we prove that the optimum ODMC approximately achieves FEXT-free capacity. Simulations show that the FEXT-free capacity can be achieved in roughly 200 iterations.

1. INTRODUCTION

Dynamic spectrum management level 3 (DSM-3) is a current initiative that the VDSL standard body [1] has undertaken to overcome one of the major limitations of VDSL deployment stemming from self far-end cross-talk (FEXT) impairments. The envisioned self-FEXT mitigation devices are CO-centric and require signal cooperation, also called signal vectoring [2], at the CO across the different users that are involved in a DSM-3 session. Pre-coding for mitigating downstream FEXT from the CO triggers most of the standardization activity [3], [4], because of the need for an inter-operable back channel to convey information about the downstream MIMO-DSL channel. Nevertheless, the mitigation of the upstream FEXT, which is more of a chip vendor proprietary technology, has already received a lot of attention [5]- [8]. Most of the above referred methods rely on MIMO channel matrix inversion and/or factorization, e.g., in [6], the author applies zero-forcing and in [2] a DFE based receiver is used at the CO. These methodologies, although capable of achieving near-optimum performance, especially in the low-to-mid upstream VDSL bands (see [9] for a comparative study on the performance of several self-FEXT cancellation solutions for upstream vectored VDSL), become computationally burdensome when pursued for hundred of tones in VDSL. Hence, the existing literature leaves a lot of room for the development of almost optimal and yet practical solutions for upstream FEXT cancellation, which motivates this work.

This paper proposes an off-diagonal MIMO canceller (ODMC) based on a low complexity and adaptive cross-talk mitigation algorithm for upstream VDSL which achieves near FEXT-free capacity. In [4], the authors have already introduced optimal off-diagonal MIMO pre-coders (ODMP) at the CO, to pre-compensate for the downstream FEXT. Contrary to intuitive expectations, because upstream self-FEXT mitigation at the CO relies neither on parameters conveyed via a back channel nor on pre-coding, the extension of the optimization approaches suggested in [4] to the derivation of optimum ODMC is not straightforward. Many new problems arise such as the need to jointly optimize the frequency domain equalizer (FEQ) at the CO and the ODMC coefficients. The derivations of the optimum ODMC and its associated maximum achievable distributed capacities turn out to be very different from the methods followed in [4]. Therefore, the current contribution unveils new theoretical results, and subsequently, new algorithms that reflect the specifics of the upstream self-FEXT mitigation problem.The paper solves the most challenging problem of learning and triggering the ODMC during the data mode, which eliminates the need to change the existing standard [1].

The remainder of the paper is organized as follows. Section 2 introduces the system model and describes the ODMC within the VDSL mode of operation. The problem formulation and derivation of the optimal ODMC are detailed in Section 3. In Section 4, equivalence between concurrent maximization of the distributed upstream capacities and the minimization of the upstream error variances, under mild assumptions is leveraged using the stochastic gradient paradigm [10]. Here, we derive a linear recursion that completes the adaptive learning of the ODMC, thus avoiding any matrix inversion. Simulation results are given in Section 5, followed by conclusions in Section 6. In the appendix, we show that the ODMC approximately achieves the FEXT-free capacity, which in turn can be proved to be very close to the full MIMO capacity.

2. SYSTEM MODEL AND ODMC

The motivation of the MIMO FEXT canceller described in the paper is best understood via the VDSL upstream system description given in Fig. 1. Let the number of vectored VDSL users participating in upstream self-FEXT cancellation be N. The $N\times 1$ received vectored signal in data mode at the CO on tone q at DMT symbol time instant t is denoted as $\mathbf{y}[q,t]$ (see Fig. 1) and can be written as:

$$\mathbf{y}[q,t] = (\mathbf{H}\mathbf{x}+\mathbf{v})[q,t], \qquad (1)$$

where $\mathbf{H}[q,t]$ is the upstream $N\times N$ MIMO-DSL channel matrix, $\mathbf{x}[q,t]$ is a $N\times 1$ column vector that represents the upstream transmit signals from the N different CPEs in data mode, $\mathbf{v}[q,t]$ is the $N\times 1$ column noise vector experienced at the CO receiver. By the central limit theorem, $\mathbf{v}[q,t]$ is Gaussian, complex circular with zero mean whose covariance matrix, Γ_v, is diagonal in the absence of alien cross-talks. The MIMO channel matrix can be written as

$$\mathbf{H}[q,t] = (\mathbf{H}_d(\mathbf{I}+\mathbf{C}))[q,t] \,, \qquad (2)$$

where $\mathbf{H}_d[q,t]$ is a diagonal matrix that represents the end-to-end direct line attenuation, and $\mathbf{C}[q,t]$ is an $N \times N$ off-diagonal matrix (all its diagonal entries are zero) whose element $C_{i,j}$ denotes the coupling that reflects the upstream self-FEXT from user j into user i. The $C_{i,j}$ are generally of the order of 30 to 40 dB below unity (0 dB). The off-diagonal self-FEXT channel model and its properties are detailed in [11] and will be leveraged in some key approximations of general derivations. Prior to entering data mode, in the presence of self-FEXT, and in the absence of an upstream self-FEXT canceller, a diagonal frequency domain equalization (FEQ) matrix $\mathbf{F}_{bc}[q,t]$ is learnt [1] (see Fig. 1). In data mode, the matrix $\mathbf{F}_{bc}[q,t]$ is updated at every DMT symbol time instant t and it compensates for the diagonal entries of the MIMO channel, i.e. $(\mathbf{F}_{bc}\mathbf{H}_d)[q,t] = \mathbf{I}$. As a consequence, the $N \times 1$ equalized received vectored data sample without self-FEXT canceller, $\mathbf{y}_{bc}[q,t]$, is given as

$$\mathbf{y}_{bc}[q,t] \triangleq (\mathbf{F}_{bc}\mathbf{y})[q,t] = ((\mathbf{I}+\mathbf{C})\mathbf{x}+\mathbf{w})[q,t], \qquad (3)$$

where $\mathbf{w} = \mathbf{F}_{bc}\mathbf{v} = \mathbf{H}_d^{-1}\mathbf{v}$. It is clear from (3) that the equivalent channel response post diagonal equalization is $\mathbf{I}+\mathbf{C}$. This serves as the motivation for introducing a self-FEXT canceller of the form $\mathbf{I}-\mathbf{R}$ where $\mathbf{R}$ is a $N \times N$ off-diagonal matrix. Furthermore, the strategy to split the equalization of the MIMO channel $\mathbf{H}$ into the diagonal-matrix component $\mathbf{F}_{bc}$ and the full-matrix component $\mathbf{I}-\mathbf{R}$ is tailored to allow for the engagement of the ODMC in data-mode; equalization of the diagonal component of the MIMO channel is indeed required before each vectored user can enter data-mode. Most importantly, as we will show later, enforcing this off-diagonal structure also leads to an adaptive algorithm without matrix inversion that converges to the optimal ODMC that simultaneously maximizes the distributed Shannon capacities of all the upstream receivers. The ODMC thus operates on the current equalized vectored signal $\mathbf{y}_{bc}[q,t]$ (see Fig. 1) to yield an enhanced (denoised) vectored signal $\mathbf{z}[q,t/t']$ given as

$$\mathbf{z}[q,t/t'] \triangleq (\mathbf{I}-\mathbf{R})[q,t']\mathbf{y}_{bc}[q,t]. \qquad (4)$$

In (4), the matrix $\mathbf{R}$ has been updated at time $t' < t$. Although matrices $\mathbf{R}$ and $\mathbf{C}$ both have zero diagonal entries, when multiplied together as suggested by (3) and (4), the result will, in general, contain diagonal components that need to be compensated by leveraging a post-canceller (pc) diagonal matrix equalizer $\mathbf{F}_{pc}$ (see Fig. 1). The updating and/or learning of the FEQs after the activation of the self-FEXT canceller is automatically done using seamless rate adaptation (SRA) [1] operation, as soon as the monitored noise margin crosses a pre-defined threshold for a sufficiently long duration. The SRA feature also involves an update of the bit loading of all the tones that have experienced a sufficiently long and large margin change, making sure that the BER is at most 10^{-7} under new SNR conditions. The post-canceller diagonal equalizer, $\mathbf{F}_{pc}[q,t'']$, can be actually updated at a

[1] in what follows, the subscript bc means before canceller

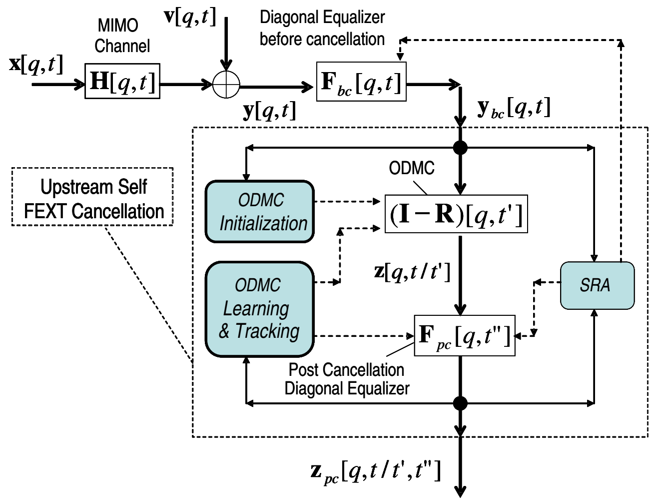

Figure 1: MIMO-VDSL upstream self-FEXT cancellation using ODMC on tone q at DMT symbol time instant t

different DMT time instant t'', than the ODMC ($t'' < t$). The final soft vectored information $\mathbf{z}_{pc}[q,t/t',t'']$ passed to the trellis demodulation block used in VDSL [1] is:

$$\begin{aligned}\mathbf{z}_{pc}[q,t/t',t''] &\triangleq \mathbf{F}_{pc}[q,t'']\mathbf{z}[q,t/t'] \\ &= \mathbf{F}_{pc}[q,t''](\mathbf{I}-\mathbf{R})[q,t']\mathbf{y}_{bc}[q,t]. \qquad (5)\end{aligned}$$

Having discussed the VDSL upstream system model, the main goal is to estimate the ODMC or the optimum matrix $\mathbf{R}$. In our approach, we concurrently maximize the distributed Shannon capacity of each upstream link. We also prove that maximizing the above capacity is equivalent to minimizing the mean-square error in the estimation of the transmitted symbols. This helps us to take advantage of the stochastic gradient paradigm and thus we can estimate $\mathbf{R}$ by linear recursion, thereby avoiding the need for matrix inversion, which is required by other previous approaches [5,6].

As shown in Fig. 1, this adaptive ODMC may be initialized with a non-zero value via the 'ODMC initialization' block; details of this initialization, however, are skipped for brevity. In addition, we can also show that the achieved capacity is very close to the MIMO capacity; again, the proof is omitted due to lack of space.

3. BLOCK OPTIMUM ODMC

3.1 Optimum ODMC based on the Distributed Upstream Shannon Capacity

In this section we will derive the explicit form of the optimal ODMC that simultaneously maximizes the distributed Shannon capacities of all the upstream receivers. To this end, we express the m^{th} component z_m, $1 \le m \le N$ of $\mathbf{z}$ (see (4)) as

$$\begin{aligned} z_m &= (\mathbf{c}_{-m}^T - \mathbf{r}_{-m}^T - row_{-m}\{(\mathbf{RC})_{nd}\})\mathbf{x}_{-m} \\ &\quad +(1-(\mathbf{RC})_{m,m})x_m + (w_m - \mathbf{r}_{-m}^T\mathbf{w}_{-m}) \,. \qquad (6)\end{aligned}$$

Here, $(.)_{nd}$ stands for the matrix operator that zeroes the diagonal entries while keeping the off-diagonal terms unchanged, $\mathbf{r}_{-m}^T$ is a $(N-1)$ length row vector that has all entries of $\mathbf{r}_m^T$ but the m^{th}(which is 0 as $\mathbf{R}$ is an off-diagonal

$$\omega_{z,m} = \log_2\left(1 + \frac{|1-(\mathbf{RC})_{m,m}|^2\,\sigma_x^2}{\left|\left|\mathbf{c}_{-m}^T - \mathbf{r}_{-m}^T - row_{-m}\{(\mathbf{RC})_{nd}\}\right|\right|^2 \sigma_x^2 + \sigma_{w,m}^2 (1-\mathbf{r}_{-m}^T)\Gamma_{w,m}\begin{pmatrix} 1 \\ -\mathbf{r}_{-m}^* \end{pmatrix}}\right) \quad (7)$$

matrix), and where $\mathbf{r}_m^T$ is the m^{th} row of matrix $\mathbf{R}$. Similar definitions holds for $\mathbf{c}_{-m}^T$, $row_{-m}\{\mathbf{RC}_{nd}\}$ and $\mathbf{w}_{-m}$. The Shannon capacity $\omega_{z,m}$ of the m^{th} mitigated channel is given in (7) and is derived directly from (6), where $\Gamma_{w,m} = \frac{1}{\sigma_{w,m}^2} E\left[\left[w_m|\mathbf{w}_{-m}^T\right]^H \left[w_m|\mathbf{w}_{-m}^T\right]\right]$. Defining

$$\zeta_{pc,m} \triangleq \frac{(\mathbf{r}_{-m}^T + row_{-m}\{(\mathbf{RC})_{nd}\} - \mathbf{c}_{-m}^T)\mathbf{x}_{-m} + (\mathbf{r}_{-m}^T \mathbf{w}_{-m} - w_m)}{(1-(\mathbf{RC})_{m,m})} \quad (8)$$

and combining (6), (7) and (8) yields the following simpler expression for $\omega_{z,m}$ in the presence of the ODMC:

$$\omega_{z,m} = \log_2\left(1 + \frac{\sigma_x^2}{VAR[\zeta_{pc,m}]}\right) . \quad (9)$$

Maximization of the distributed Shannon capacity may then be expressed as $\max_{\mathbf{r}_{-m}\in\mathbb{C}^{N-1}} (\omega_{z,m}),\ 1 \le m \le N$.

Notations: The notations used in the solution of the above optimization problem are described in what follows.

$$\mathbf{g}_m \triangleq col_m\{\mathbf{C}\}, \mathbf{r}_m^{'T} \triangleq [1| - \mathbf{r}_{-m}^T], \mathbf{g}_m^{'} \triangleq \left[\frac{1}{\mathbf{g}_{-m}}\right], \alpha_m^2 \triangleq \frac{\sigma_{w,m}^2}{\sigma_x^2} \quad (10)$$

$$\mathbf{1}_m^T \triangleq \left[\begin{array}{ccccc} 0 & 0 & 1(m^{th} entry) & 0 & 0 \end{array}\right] \quad (11)$$

Note that $\frac{1}{\alpha_m^2}$ is the FEXT-free SNR for the channel m. We define the $(N-1)\times(N-1)$ matrix $\mathbf{C}_{-m,-m}$ as that obtained by removing the m^{th} row and column from the $N \times N$ matrix $\mathbf{C}$. The following $N \times (N-1)$ matrix is also useful:

$$\mathbf{C}_m \triangleq \left[\frac{\mathbf{c}_{-m}^T}{\mathbf{I}_{N-1} + \mathbf{C}_{-m,-m}}\right] \quad (12)$$

Property: The components of the m^{th} row of the optimal ODMC that concurrently maximizes the distributed Shannon capacities are given by:

$$[1| - \mathbf{r}_{-m}^{o,T}] = \frac{[1|\mathbf{g}_{-m}^H](\mathbf{C}_m\mathbf{C}_m^H + \alpha_m^2\Gamma_{w,m})^{-1}}{[1|\mathbf{g}_{-m}^H](\mathbf{C}_m\mathbf{C}_m^H + \alpha_m^2\Gamma_{w,m})^{-1}\mathbf{1}_1} , \quad (13)$$

and the associated maximum achievable capacities $\omega_{z,m}^o$, $1 \le m \le N$ are

$$\omega_{z,m}^o = \log_2\left(1 + [1|\mathbf{g}_{-m}^H](\mathbf{C}_m\mathbf{C}_m^H + \alpha_m^2\Gamma_{w,m})^{-1}\left[\frac{1}{\mathbf{g}_{-m}}\right]\right) . \quad (14)$$

Proof: Using the above notations and the definition of the operator $row_{-m}\{\cdot\}$, we rewrite the terms in $\zeta_{pc,m}$ as:

$$(1-(\mathbf{RC})_{m,m}) = 1 - \mathbf{r}_{-m}^T\mathbf{g}_{-m} , \quad (15)$$

$$\begin{aligned} row_{-m}\{(\mathbf{RC})_{nd}\} &= \mathbf{r}_{-m}^T(\mathbf{C}_{-m,-m} - \mathbf{g}_{-m}\mathbf{1}_{-m}^T) \quad (16) \\ &= \mathbf{r}_{-m}^T\mathbf{C}_{-m,-m} , \quad (17) \end{aligned}$$

where the last equality follows since $\mathbf{1}_{-m} = \mathbf{0}$. Equations (8), (15) and (17) lead to the following scalar product representation of $\zeta_{pc,m}$:

$$\zeta_{pc,m} = -\frac{[1| - \mathbf{r}_{-m}^T]\left(\left[\frac{\mathbf{c}_{-m}^T}{\mathbf{I}_{N-1} + \mathbf{C}_{-m,-m}}\right]\mathbf{x}_{-m} + \left[\frac{w_m}{\mathbf{w}_{-m}}\right]\right)}{[1| - \mathbf{r}_{-m}^T]\left[\frac{1}{\mathbf{g}_{-m}}\right]} . \quad (18)$$

Equations (10), (12), and (18) enable us to write the SNR after applying the ODMC, $\sigma_x^2/VAR[\zeta_{pc,m}]$, as the quotient of two quadratic forms as follows:

$$\frac{\sigma_x^2}{VAR[\zeta_{pc,m}]} = \frac{\left|\mathbf{r}_m^{'T}\mathbf{g}_m^{'}\right|^2}{\mathbf{r}_m^{'T}(\mathbf{C}_m\mathbf{C}_m^H + \alpha_m^2\Gamma_{w,m})\mathbf{r}_m^{'*}} . \quad (19)$$

Defining the Hermitian scalar product $\langle\cdot,\cdot\rangle$ and considering the normalized vector $\tilde{\mathbf{r}}_m^{'T} \triangleq \mathbf{r}_m^{'T}\mathbf{Q}_m^{1/2}$ with $\mathbf{Q}_m \triangleq (\mathbf{C}_m\mathbf{C}_m^H + \alpha_m^2\Gamma_{w,m})$ give the following identity that considerably eases the maximization of $\sigma_x^2/VAR[\zeta_{pc,m}]$:

$$\frac{\sigma_x^2}{VAR[\zeta_{pc,m}]} = \frac{\left|\langle\mathbf{r}_m^{'}, (\mathbf{Q}_m^{-1/2}\mathbf{g}_m^{'})^*\rangle\right|^2}{\langle\tilde{\mathbf{r}}_m^{'}, \tilde{\mathbf{r}}_m^{'}\rangle} . \quad (20)$$

Indeed, (10) and the above definition of the matrix $\mathbf{Q}_m$ combined with the Schwartz inequality unveil the maximum value of the ratio $\sigma_x^2/VAR[\zeta_{pc,m}]$ as

$$\begin{aligned} \frac{\sigma_x^2}{VAR[\zeta_{pc,m}]} &\le \langle(\mathbf{Q}_m^{-1/2}\mathbf{g}_m^{'})^*), (\mathbf{Q}_m^{-1/2}\mathbf{g}_m^{'})^*)\rangle \\ &= [\mathbf{1}|\mathbf{g}_{-m}^H](\mathbf{C}_m\mathbf{C}_m^H + \alpha_m^2\Gamma_{w,m})^{-1}\left[\frac{1}{\mathbf{g}_{-m}}\right] \quad (21) \end{aligned}$$

In (21), the maximum value is reached when the vector $\tilde{\mathbf{r}}_m^{'o}$ is proportional to $(\mathbf{Q}_m^{-1/2}\mathbf{g}_m^{'})^*$, i.e.,

$$\mathbf{r}_m^{'o,T} = \eta\,\mathbf{g}_m^{'H}\mathbf{Q}_m^{-1} . \quad (22)$$

In keeping with (10), the constant η is chosen to ensure that the first element of the optimum vector $\tilde{\mathbf{r}}^{'o}$ equals unity. Thus,

$$\eta = \frac{1}{\mathbf{g}_m^{'H}\mathbf{Q}_m^{-1}\mathbf{1}_1} . \quad (23)$$

Using (9), (21)), (22), (23) and the definition of Matrix $\mathbf{Q}_m$, the optimal ODMC and their respective capacities are obtained as given in (13) and (14). □

3.2 Optimum ODMC, Error variance, MIMO-Capacity and FEXT-free performance

In the appendix, we prove that the optimum ODMC capacity in (14) (SNR in (21)) is close to the FEXT-free capacity (SNR). Though, distributed Shannon capacity and not MIMO Shannon capacity is used as the optimization criteria in this paper, due to the relatively small values of the FEXT coupling coefficients $C_{i,j}$, it can be easily shown that FEXT-free capacity Ω_{FF} (in nats), which the optimum ODMC achieves, is also close to the MIMO capacity Ω_{MIMO} given by $\Omega_{MIMO} \approx \Omega_{FF} + ||\mathbf{C}||_F^2$, where $||\mathbf{C}||_F^2$ is the Frobenius norm of the coupling matrix $\mathbf{C}$; $||\mathbf{C}||_F^2$ is generally of the order of 10^{-3} for frequencies around 10 MHz.

Starting from (6), assuming perfect detection and post-canceller diagonal equalization[2], the error in receiver m obtained by calculating the difference between the detected QAM symbol $x_{pc,m}[q,t]$ and denoised symbol $z_{pc,m}[q,t]$ (Fig. 1), is $e_{pc,m} = \frac{(\mathbf{r}^T_{-m}+row_{-m}\{(\mathbf{RC})_{nd}\}-\mathbf{c}^T_{-m})\mathbf{x}_{-m}+(\mathbf{r}^T_{-m}\mathbf{w}_{-m}-w_m)}{(1-(\mathbf{RC})_{m,m})} = \zeta_{pc,m}$. Hence, from (9) and the previous expression, it is clear that maximizing the distributed capacity is equivalent to minimizing the variance of the error. Therefore, we can use the LMS based adaptive algorithm and recursively estimate the ODMC and the post-canceller FEQ, as described next.

4. ADAPTIVE OPTIMUM ODMC

We now present an adaptive algorithm to learn the ODMC and post-canceller FEQ coefficients $\mathbf{F}_{pc}[q,t]$. In keeping with (5), we define the post-canceller vectored error as

$$\mathbf{e}_{pc}[q,t/t',t''] \triangleq \mathbf{x}_{pc}[q,t] - \mathbf{F}_{pc}[q,t''](\mathbf{I}-\mathbf{R})[q,t']\mathbf{y}_{bc}[q,t], \quad (24)$$

where $\mathbf{x}_{pc}[q,t]$ designates the demapped soft vectored symbol $\mathbf{z}_{pc}[q,t/t',t'']$, defined in (5). From (24), the component of the vectored error on channel m, $1 \le m \le N$ is as follows:

$$e_{pc,m}[q,t/t',t''] = x_{pc,m}[q,t] - f_{pc,m}[q,t''](y_{bc,m}[q,t] - \mathbf{r}^T_{-m}[q,t']\mathbf{y}_{bc,-m}[q,t]) \,. \quad (25)$$

In (25), $f_{pc,m}[q,t'']$ denotes the m^{th} diagonal entry of the $N \times N$ diagonal matrix $\mathbf{F}_{pc}[q,t'']$. Based on the explicit form of the error in (25), we can derive the adaptive LMS-based algorithm that converges [10] towards the ODMC that simultaneously minimizes the variance of the distributed errors of the upstream receivers at the CO side. As proved in Section 3, this adaptive algorithm also simultaneously maximizes the distributed Shannon capacities of the upstream links under mild assumptions. The stochastic gradient paradigm [10] leads to the following equations that define updates to the post-canceller equalizer every DMT symbol and the self-FEXT canceller coefficients every K DMT symbols: For $1 \le m \le N$, $n \ge 0$, and $1 \le i \le K$

$$f_{pc,m}[q,t_{nK+i}] = f_{pc,m}[q,t_{nK+i-1}] + \mu[t_{nK+i}](y_{bc,m}[q,t_{nK+i}] - \mathbf{r}^T_{-m}[q,t_{nK}]\mathbf{y}_{bc,-m}[q,t_{nK+i}])e^*_{pc,m}[q,t_{nK+i}/t_{nK},t_{nK+i-1}] \,, \quad (26)$$

$$\mathbf{r}_{-m}[q,t_{(n+1)K}] = \mathbf{r}_{-m}[q,t_{nK}] - \mu[t_{(n+1)K}]f_{pc,m}[q,t_{(n+1)K}] \cdot \mathbf{y}_{bc,-m}[q,t_{(n+1)K}]e^*_{pc,m}[q,t_{(n+1)K}/t_{nK},t_{(n+1)K}] \,. \quad (27)$$

The ODMC can also be intialized with a starting value $\mathbf{R}_{in}$ before engaging its adaptive learning to facilitate a faster convergence. One such intial estimate was discussed in [4], where $\mathbf{R}_{in}$ was set with an estimate of the coupling matrix $\mathbf{C}$. Additionally, once the ODMC is engaged, updates to the pre-canceller diagonal equalizer $\mathbf{F}_{bc}$ may be frozen and only the post-canceller equalizer $\mathbf{F}_{pc}$ may be updated, thereby minimizing the additional computing resources that are required.

[2]These two assumptions are easily met given that the mandatory BER value is 10^{-7} [1] and considering the long FEQ learning MEDLEY sequences recommended in VDSL standards [1].

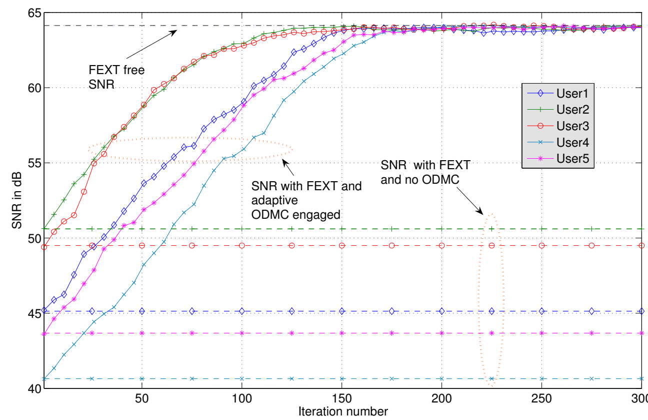

Figure 2: ODMC Learning for 5 users at 4MHz on a 0.3 km long loop.

In Section 5, we provide the performance achieved by the above algorithm based on Matlab simulations, where we see that the theoretical maximum capacity is reached by each upstream link after convergence of the adaptive ODMC. It is worth noting that the adaptive scheme described in (27) operates with a per-tone complexity of $\mathcal{O}\left(N^2\right)$ as compared to the $\mathcal{O}\left(N^3\right)$ complexity required by methods employing matrix inversion such as those based on zero-forcing [6], while achieving similar (and near-FEXT-free) performance. Additionally, the proposed adaptive scheme directly tracks the equalizer rather than the channel, thereby limiting numeric inaccuracies due to error propagation that may be more prevalent in the latter scenario. Furthermore, the proposed scheme can also address the seamless addition or removal of a vectored user, which is not straightforward in the case of the zero-forcing approach.

5. SIMULATIONS

Fig. 2 displays the SNR of five vectored users on the 4 MHz tone for a loop of length 0.3 km across the whole learning phase of the ODMC. Starting with an initial estimate of $\mathbf{R}_{in} = [0]$, the learning of the ODMC follows the recursive scheme described in (27) as explained in Section 4. It is seen in Fig. 2 that the adaptive algorithm completes its convergence within 200 iterations even for the worst affected user, i.e., the user that experiences the lowest SNR in the absence of any ODMC. As displayed in Fig. 2, after reaching steady state, the SNR achieved by all the users is very close to the FEXT-free SNR obtained if the only disturbance is the background noise and in the absence of any canceller. We see in Fig. 2 that the most impacted user (i.e. user # 4) enjoys an SNR improvement of $\sim 23dB$, from $\sim 41dB$ without ODMC up to $\sim 64dB$ after convergence of the adaptive ODMC.

6. CONCLUSIONS

An off-diagonal MIMO canceller (ODMC) that mitigates the upstream self-FEXT in VDSL systems at the CO has been introduced. A low complexity and low latency recursive algorithm that converges towards the ODMC that concurrently maximizes the distributed capacities of the upstream links has been derived and evaluated. As shown in the appendix, this algorithm can theoretically achieve close to FEXT-free SNR in steady state, while simulations show that convergence towards this FEXT-free SNR is possible within 200 iterations starting with zero as an initial value. The most FEXT-impaired users enjoy more than 20dB of SNR gain. Although the paper has only considered the learning of the ODMC with a fixed channel, the same adaptive scheme can be used for other crucial phases of upstream self-FEXT mitigation such as tracking the channel changes, and supporting the addition of new users that lead to the abrupt change of the channel dimensions [4].

7. APPENDIX A: PROOF THAT THE OPTIMAL ODMC ACHIEVES NEAR FEXT-FREE SNR

Using (10), the FEXT-free SNR is given by $\frac{1}{\alpha_m^2}$, while from (21), the optimum ODMC SNR in the absence of alien noise (i.e., $\Gamma_{w,m} = \mathbf{I}$) is given by

$$SNR_o = \left[\begin{array}{c|c} 1 & \mathbf{g}_{-m}^H \end{array}\right] (\mathbf{C}_m\mathbf{C}_m^H + \alpha_m^2\mathbf{I})^{-1} \left[\begin{array}{c} 1 \\ \hline \mathbf{g}_{-m}, \end{array}\right]. \quad (28)$$

Let $\mathbf{D} \triangleq (\mathbf{I}+\mathbf{C}_{-m,-m})(\mathbf{I}+\mathbf{C}_{-m,-m}^H)$. Using (12), we get

$$(\mathbf{C}_m\mathbf{C}_m^H + \alpha_m^2\mathbf{I}) = \left[\begin{array}{c|c} ||\mathbf{c}_{-m}||^2+\alpha_m^2 & -\mathbf{c}_{-m}^T+\delta_c^H \\ \hline \mathbf{c}_{-m}^*+\delta_c & \mathbf{D}+\alpha_m^2\mathbf{I}_{N-1} \end{array}\right], \quad (29)$$

where δ_c are second order terms of the coupling coefficients. Let $A = ||\mathbf{c}_{-m}||^2+\alpha_m^2$, $C = \mathbf{c}_{-m}^T+\delta_c^H$, $B = C^H$, $D = \mathbf{D} + \alpha_m^2\mathbf{I}$ and $\left[\begin{array}{c|c} X & Y \\ \hline Z & \Delta \end{array}\right] = \left[\begin{array}{c|c} ||\mathbf{c}_{-m}||^2+\alpha_m^2 & -\mathbf{c}_{-m}^T+\delta_c^H \\ \hline \mathbf{c}_{-m}^*+\delta_c & \mathbf{D}+\alpha_m^2\mathbf{I}_{N-1} \end{array}\right]^{-1}$. Noting that A is a positive scalar and using block-wise matrix inversion we have, $X = A^{-1}+A^{-1}C(D-BA^{-1}C)^{-1}BA^{-1}$, $Y = -A^{-1}C(D-BA^{-1}C)^{-1}$, $Z = -(D-BA^{-1}C)^{-1}BA^{-1}$, and $\Delta = (D-BA^{-1}C)^{-1}$. Therefore from (28) we have,

$$\begin{aligned} SNR_o &= \left[\begin{array}{c|c} 1 & \mathbf{g}_{-m}^H \end{array}\right] \left[\begin{array}{c|c} X & Y \\ \hline Z & \Delta \end{array}\right] \left[\begin{array}{c} 1 \\ \hline \mathbf{g}_{-m} \end{array}\right] \\ &= X + \mathbf{g}_{-m}^H Z + Y\mathbf{g}_{-m} + \mathbf{g}_{-m}^H \Delta \mathbf{g}_{-m} \quad (30) \end{aligned}$$

We can also write, $X = A^{-1}+(A^{-1})B^H(AD-BB^H)^{-1}B$. Now, using the matrix inversion lemma, we have

$$\begin{aligned} B^H(AD-BB^H)^{-1}B &= A^{-1}[B^HD^{-1}B - \frac{B^HD^{-1}BB^HD^{-1}BA^{-1}}{-1+B^HD^{-1}BA^{-1}}] \\ &= \frac{\alpha}{1-\alpha}, \text{ where } \alpha = A^{-1}B^HD^{-1}B. \end{aligned}$$

$$\therefore \quad X = A^{-1}+A^{-1}\frac{\alpha}{1-\alpha} = \frac{A^{-1}}{1-\alpha}. \quad (31)$$

Similarly, $Z = -(AD-BB^H)^{-1}B$, hence

$$\begin{aligned} \mathbf{g}_{-m}^H Z &= -\mathbf{g}_{-m}^H[A^{-1}D^{-1} - \frac{(A^{-1})^2D^{-1}BB^HD^{-1}}{-1+B^HD^{-1}BA^{-1}}]B \\ &= -\mathbf{g}_{-m}^H D^{-1}BA^{-1}\frac{1}{1-\alpha} \end{aligned}$$

Defining $\mathbf{g}_{-m}^H D^{-1}BA^{-1}$ as β, we have :

$$\mathbf{g}_{-m}^H Z = -\frac{\beta}{1-\alpha}, \; Y\mathbf{g}_{-m} = -\frac{\beta^H}{1-\alpha} \quad (32)$$

Now, $\mathbf{g}_{-m}^H\Delta\mathbf{g}_{-m} = \mathbf{g}_{-m}^H A(AD-BB^H)^{-1}\mathbf{g}_{-m}$

$$\begin{aligned} \mathbf{g}_{-m}^H\Delta\mathbf{g}_{-m} &= A\mathbf{g}_{-m}^H[A^{-1}D^{-1} - \frac{(A^{-1})^2D^{-1}BB^HD^{-1}}{-1+B^HD^{-1}BA^{-1}}]\mathbf{g}_{-m} \\ &= \mathbf{g}_{-m}^H D^{-1}\mathbf{g}_{-m} + \frac{A\beta\beta^H}{(1-\alpha)} \quad (33) \end{aligned}$$

Substituting (31)-(33) in (30), we get

$$\begin{aligned} SNR_o &= \frac{A^{-1}}{1-\alpha} - \frac{\beta}{1-\alpha} - \frac{\beta^H}{1-\alpha} + \frac{A\beta\beta^H}{1-\alpha} + \mathbf{g}_{-m}^H D^{-1}\mathbf{g}_{-m} \\ &= A^{-1}\frac{|1-A\beta|^2}{1-\alpha} + \mathbf{g}_{-m}^H D^{-1}\mathbf{g}_{-m} \quad (34) \end{aligned}$$

To proceed further we make following observations:

1. α_m^2 and $||\mathbf{c}_{-m}||^2$ are of the order 10^{-5}.
2. $\mathbf{D} \approx \mathbf{I}+\Omega$, where Ω is a non-diagonal matrix.
3. Using observation 1, $D = \mathbf{D}+\alpha_m^2\mathbf{I} \approx \mathbf{D}$.
4. We can neglect higher order terms in the presence of the lower orders.
5. $\mathbf{D}^{-1} \approx \mathbf{I}-\Omega$, $\mathbf{D}^{-1}$ is a positive definite matrix.
6. $B \approx \mathbf{c}_{-m}$ and $A\beta \approx \mathbf{g}_{-m}^H\mathbf{c}_{-m}$.
7. $\frac{1}{1-\alpha} \approx \frac{1}{1-\frac{||\mathbf{c}_{-m}||^2||-||\mathbf{c}_{-m}||^2||||\Omega||}{||\mathbf{c}_{-m}||^2+\alpha_m^2}} \approx \frac{\alpha_m^2+||\mathbf{c}_{-m}||^2||}{\alpha_m^2}$

Using the above observations and substituting A^{-1} in (34), we have the approximate ODMC SNR as :

$$SNR_o \approx SNR_{fextfree}|1-\mathbf{g}_{-m}^H\mathbf{c}_{-m}|^2 \approx SNR_{fextfree} \quad (35)$$

The above expression proves that the ODMC SNR is indeed very close to the FEXT-free SNR.

REFERENCES

[1] G.993.2, "Very High Speed Digital Subscriber Line Transceiver 2", ITU-T, VDSL2, 02/2006.

[2] G. Ginis and J. Cioffi, "Vectored Transmission for Digital Subscriber Line Systems", *IEEE J-SAC*, vol. 20, no. 5, pp. 1085-1104, June 2002.

[3] R. Cendrillon *et al.*, "Improved Linear Crosstalk Precompensation for DSL", *Proc. IEEE ICASSP 2004*, Montreal, 17-21 May 2004.

[4] P. Duvaut *et al.*, "Adaptive Off-Diagonal MIMO Pre-coder (ODMP) for Downstream DSL Self FEXT Cancellation", *Proc. IEEE GLOBECOM 2007*, Washington D.C., 26-30 Nov. 2007.

[5] R. Cendrillon *et al.*, "Partial Crosstalk Cancellation for Upstream VDSL", *EURASIP Journal on Applied Signal Processing*, 2004:10, pp. 1520-1535.

[6] R. Cendrillon *et al.*,"A Near-Optimal Linear Crosstalk Canceller for Upstream VDSL",*IEEE Trans. Signal Processing*, vol. 54, no. 8, pp. 3136-3146, Aug. 2006.

[7] Kok-Wui Cheong *et al.*, "Multiuser Soft Interference Canceller via Iterative Decoding for DSL Applications", *IEEE J-SAC*, vol. 20, no. 2, pp. 363-371, Feb. 2002.

[8] R. Hormis *et al.*, "Monte Carlo FEXT Cancellers for DSL Channels", *IEEE Trans. Circuits and Systems*, vol. 52, no. 9, pp. 1894-1908, Sep. 2005.

[9] S. Dolui *et al.*, "Evaluation of Algorithms for FEXT Cancellation in Upstream VDSL", *Proc. NCC 2008*, Mumbai, 1-3 Feb. 2008.

[10] A. Benveniste, *Adaptive Algorithms and Stochastic approximations*, Springer Verlag, New York, 1990.

[11] P. Duvaut *et al.*, "Construction of a DSL MIMO channel model for evaluation of FEXT cancellation systems in VDSL2", *Proc. 2007 IEEE Sarnoff Symposium*, Princeton, Apr. 30-May 2, 2007.