Research Article

Computationally Efficient Partial Crosstalk Cancellation in Fast Time-Varying DSL Crosstalk Environments

Amir R. Forouzan and Lee M. Garth

*Department of Electrical and Computer Engineering, University of Canterbury, Private Bag 4800,
Christchurch 8020, New Zealand*

Line selection (LS), tone selection (TS), and joint tone-line selection (JTLS) partial crosstalk cancellers have been proposed to reduce the online computational complexity of far-end crosstalk (FEXT) cancellers in digital subscriber lines (DSL). However, when the crosstalk profile changes rapidly over time, there is an additional requirement that the partial crosstalk cancellers, particularly the LS and JTLS schemes, should also provide a low preprocessing complexity. This is in contrast to the case for perfect crosstalk cancellers. In this paper, we propose two novel channel matrix inversion methods, the approximate inverse (AI) and reduced inverse (RI) schemes, which reduce the recurrent complexity of the LS and JTLS schemes. Moreover, we propose two new classes of JTLS algorithms, the subsort and Lagrange JTLS algorithms, with significantly lower computational complexity than the recently proposed optimal greedy JTLS scheme. The computational complexity analysis of our algorithms shows that they provide much lower recurrent complexities than the greedy JTLS algorithm, allowing them to work efficiently in very fast time-varying crosstalk environments. Moreover, the analytical and simulation results demonstrate that our techniques are close to the optimal solution from the crosstalk cancellation point of view. The results also reveal that partial crosstalk cancellation is more beneficial in upstream DSL, particularly for short loops.

1. INTRODUCTION

The main impairments in digital subscriber lines (DSL) are loop loss, crosstalk, background noise, impulse noise, and radio ingress. For the short loop lengths of very high-speed digital subscriber lines (VDSL), the dominant impairment is far-end crosstalk (FEXT). Recently, FEXT cancellation techniques in loops with coordination among the transceivers on one side have been proposed. Coordination results in effective FEXT cancellation with higher performance and complexity reduction [1]. However, the method and success of FEXT cancellation techniques strongly depend on the degree of coordination among the DSL transceivers and the available processing power. FEXT cancellation in downstream (DS) and upstream (US) discrete multitone (DMT) DSL can be done by coordinating the transmitter and the receiver modems, respectively.

In [2], a vector Tomlinson-Harashima precoder and in [3] a simpler technique called the diagonalizing precompensator have been proposed for crosstalk mitigation in DS DSL. For US transmission, a zero-forcing-generalized decision feedback equalizer (DFE) has been proposed for FEXT cancellation in [2]. In [4], it has been shown that the feedback portion of the DFE is not required, and a zero-forcing linear equalizer is near optimum for US VDSL.

These methods achieve the channel capacity for each tone very closely. However, new techniques requiring fewer computations are of crucial importance because of the huge complexity order of the system. In [5, 6], reduced complexity techniques for FEXT cancellation in DS and US have been studied. The proposed techniques decrease the computational complexity by ignoring crosstalk from nondominant crosstalkers (*line selection*) or by dedicating the processing power to the frequency bands where it is more beneficial (*tone selection*) or by combining line and tone selection techniques (*joint tone-line selection*).

Note that although the twisted-pair channel does not change quickly [7], the crosstalk profile can change very rapidly in DSL systems. These profile time variations can be due to a variety of causes. Most obviously, they can be

caused by quiescent modes in DSL transmitters. For example, a protocol, which reduces the transmitted power or switches the modem to an idle state when there is no information to be sent, would not only save money for the transmitter, but would also reduce the crosstalk power in the loop plant and allow other rate-adaptive modems to increase their rates [7]. Such a power-reduction scheme would have dramatic effect on the crosstalk profile as a significant share of traffic over DSL lines is due to Internet web browsing including variable-rate multimedia traffic. This bursty traffic yields a minimum transmission power for each DSL user which varies over time.

Such quiescent modes have been proposed in VDSL standards. For example, in short-term stationary VDSL systems, including burst transmission systems and systems that use quiescent modes, the transmitter is silent or generates only a pilot tone to reduce power consumption and crosstalk levels during idle IP packets [8]. Clause 5.4 of [9] describes the activation and power control procedure for a VDSL transmission unit (VTU). To reduce the crosstalk levels and radio frequency interference (RFI) of the VDSL system during a normal transmission session, the VTU dynamically switches between the steady-state transmission state and an idle state, a dynamic power-saving state, or a power-saving sleeping mode. The transition between these states is expected to take place in less than a hundred milliseconds. In more recent ADSL2 and ADSL2+ standards, however, the power control is activated within a time frame of seconds to minutes.

DSL systems can also be subject to time-varying crosstalk profiles from different coexisting DSL services with different symbol durations [10]. For example, in Annex F of [8] the time-varying and user data-dependent nature of T1 AMI and DDS systems has been studied, producing the conclusion that "the time duration of each PSD variant may vary from less than 1 millisecond to many hours." These variations can be greater than 20 dB and are caused by user data content.

As we will see in this paper, handling time-varying crosstalk is much easier for systems with crosstalk cancellation using a joint modem or a shared DSL access multiplexor (DSLAM). For in these cases, the DSLAM can easily control the power and bitrate of users in a joint fashion, avoiding the delay due to resynchronization of distributed modems.[1] In particular, we show how the DSLAM can avoid delay in partial crosstalk cancellation for fast time-varying crosstalk environments.

Considering the large number of independent DSL users in the cable, the crosstalk profile and therefore the set of dominant crosstalkers can change very rapidly over time for *short-term stationary DSL systems*. But the structures of partial crosstalk cancellers, particularly the line selection (LS) and joint tone-line selection (JTLS) schemes, depend substantially on the crosstalk profile and the set of dominant crosstalkers. Consequently, in contrast to perfect FEXT cancellation techniques, the initialization and recurrent complexity associated with partial crosstalk cancellers should be

reconsidered. In this paper, we propose two new channel matrix inversion (CMI) schemes and two novel classes of JTLS algorithms to reduce the recurrent preprocessing requirements of partial crosstalk cancellers for US and DS DSL.

Our first CMI method is based on a recently proposed power-series expansion technique for the inverse of the DSL channel transfer matrix [12]. Our second CMI method reduces the recurrent computational complexity by storing the inverse of the perfect channel matrix for each tone. When any change occurs in the crosstalk profile, the new structure for the partial crosstalk cancellers can be obtained from the stored information in a computationally efficient way. This method is a modified version of the scheme proposed in [5, 6], in which channel inversion is required every time the crosstalk profile changes over time. Since CMI is an essential part of the LS and JTLS schemes, our new CMI techniques result in a lower recurrent complexity for both the LS and JTLS schemes.

We also propose two new classes of algorithms for joint tone-line selection (JTLS). Our algorithms are much faster than a greedy algorithm recently proposed in [5, 6]. Our first JTLS scheme, the subsort JTLS algorithm, is a heuristic approach, which can nearly achieve the performance of the optimal JTLS algorithm. Our second JTLS scheme employs the Lagrange multiplier optimization technique to allocate the processing power efficiently. Our results show that the Lagrange JTLS algorithm is almost optimal for practical DSL channels.

The paper is organized as follows. In the next section, we describe the DSL channel. We review perfect crosstalk cancellation schemes in Section 3. We describe our partial crosstalk cancellation schemes in Sections 4 to 6. We evaluate the computational complexity of the new algorithms in Section 7. Finally, simulation results are presented in Section 8, and the conclusion is given in Section 9.

2.   DSL CHANNEL AND FEXT MODEL

Consider L VDSL users and the synchronized transmission of DMT symbols. In this case, the transmitted and received signals for each tone k can be arranged in the following matrix form [2]:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{n}_k, \quad 1 \le k \le N, \tag{1}$$

where N is the number of DMT tones, and $\mathbf{y}_k$, $\mathbf{x}_k$, and $\mathbf{n}_k$ are the L-dimensional vectors of received, transmitted, and noise samples for tone k, respectively. The ℓth elements of $\mathbf{y}_k$, $\mathbf{x}_k$, and $\mathbf{n}_k$ are denoted $y_k^{(\ell)}$, $x_k^{(\ell)}$, and $n_k^{(\ell)}$, respectively. Matrix $\mathbf{H}_k$ is an $L \times L$ channel transfer function for tone k, where the (i, j)th matrix component $h_k^{(i,j)} = [\mathbf{H}_k]_{ij}$ contains the single-tap complex channel from transmitter j to receiver i.

Throughout this work, we assume that crosstalk cancellation is performed by a joint modem or DSLAM located at the line termination side (central office or remote terminal). We assume that the modem has perfect knowledge of the crosstalk channel $\mathbf{H}_k$ on all tones in the DS and US directions and is aware of active and idle DSL users and their

---

[1] See [11] for an algorithm to jointly control the bitrates of the users.

transmitting power in each tone. We first concentrate on the downstream direction and then generalize our techniques to upstream DSL.

3. PERFECT CROSSTALK CANCELLATION IN DMT DSL

When the transmitter modems are colocated (e.g., the DS modems are colocated at the CO), the transmitted signals can be generated from users' data signals in a joint fashion, and it is possible to cancel crosstalk perfectly in a DMT DSL channel using vector coding schemes. In [2], a multiuser Tomlinson-Harashima precoder has been proposed, and it has been shown that the receiver nearly achieves the capacity of the twisted-pair channel as if there is no crosstalk.[2] In [3] a diagonalizing precompensator (DP) has been proposed, which nearly achieves the channel capacity on each line as well. In this paper, we consider DP for simplicity.

The diagonalizing precompensator consists of multiplying the vector $\mathbf{x}_k$ for each tone k by the following precoding matrix prior to transmission:

$$\mathbf{P}_{k,\mathrm{DP}} = \beta_k \mathbf{H}_k^{-1} \overbrace{\mathrm{diag}\{\mathbf{H}_k\}}^{\mathbf{\Lambda}}, \qquad (2)$$

where $\mathbf{H}_k^{-1}$ is the inverse of channel matrix and β_k is a normalizing factor, which ensures that the spectral mask is not exceeded on any line. Diagonal matrix $\mathbf{\Lambda} = \mathrm{diag}\{\mathbf{H}_k\}$ contains the diagonal elements of $\mathbf{H}_k$. Therefore, if we define $\overline{\mathbf{H}}_k \triangleq \mathbf{\Lambda}^{-1}\mathbf{H}_k$, we get

$$\mathbf{P}_{k,\mathrm{DP}} = \beta_k (\overline{\mathbf{H}}_k)^{-1}. \qquad (3)$$

It has been shown in [3] that $\beta_k \simeq 1$ for DSL loops. Thus, the DP is simply a ZF precompensator for the normalized channel matrix $\overline{\mathbf{H}}_k$. However, in contrast to a ZF precompensator with its constrained transmission power, the DP can nearly achieve the capacity of crosstalk-free loops [3].

4. COMPLEXITY REDUCTION BY PARTIAL CROSSTALK CANCELLATION

The perfect crosstalk cancellation schemes proposed in [2–4] require $\mathcal{O}(L^2N)$ operations per DMT symbol period. Since the number of twisted pairs in a binder group is up to 100 and N is 4096 in VDSL, the computational complexity of perfect crosstalk cancellers is too high for current processors [5, 6]. Therefore, we consider suboptimal partial crosstalk cancellation techniques.

It is widely accepted that the crosstalk to each loop is usually predominantly from a few crosstalkers, called the dominant crosstalkers [5, 6, 13, 14]. In the upstream direction, the dominant crosstalkers to a victim line are usually its neighbouring lines in the binder or shorter loops in the binder,

which overwhelm other lines due to the near-far effect [14]. In downstream DSL, where there is no near-far effect, the dominant crosstalkers are the ones closer to the victim line in the binder, assuming an equal transmission power in all loops. Moreover, crosstalk cancellation does not have the same benefit for all frequencies. Generally, at very low frequencies crosstalk has a negligible effect on performance, and at very high frequencies performance is bounded by the loop loss and the receiver noise rather than crosstalk.

Line selection (LS) schemes cancel crosstalk from dominant crosstalkers at all frequencies. Tone selection (TS) schemes, on the other hand, only cancel crosstalk for the tones which are most beneficial. LS and TS schemes improve the performance when the processing power is not enough to cancel all of the crosstalkers in all of the tones. However, superior performance can be achieved using joint tone-line selection (JTLS) schemes. In JTLS schemes, both the frequency tones and the lines are considered to determine how to expend the available processing power to get the highest possible bitrates.

The structures of LS and JTLS partial crosstalk cancellers depend substantially on the set of dominant crosstalkers. In the following sections, we propose new CMI schemes and novel JTLS algorithms to reduce the recurrent computational complexity of these techniques.

5. PARTIAL CROSSTALK CANCELLATION BY LINE SELECTION

A possible solution to the numerical complexity problem is to restrict the crosstalk cancellation to the crosstalk resulting from the dominant crosstalkers only (line selection). LS has been considered for downstream VDSL in [6]. In this method, the subset of users with the most crosstalk energy impinging on a victim line is selected, and their crosstalk is cancelled employing a CMI technique.

Here, the output of each CMI technique is an approximation of the inverse of the channel matrix for each tone with zero elements corresponding to the nondominant crosstalkers. The zero elements are essential to reduce the computational complexity. In [5, 6], a method is proposed to make a sparse partial crosstalk precompensator matrix $\mathbf{P}_k$ with off-diagonal nonzero elements only in the positions corresponding to the dominant crosstalkers. Assuming that the number of dominant crosstalkers per tone is p, the number of operations that the LS scheme requires per tone is $\mathcal{O}(pLN)$. In fast time-varying DSL channels, the computational complexity associated with updating $\mathbf{P}_k$ also has to be considered. Using the method in [5, 6], computation of $\mathbf{P}_k$ is an $\mathcal{O}(L(p+1)^3)$ operation, where $p < L$ is the number of dominant crosstalkers for each line.

When the state of any of the modems in the cable changes from the steady-state transmission state to an idle state, the set of dominant crosstalkers for the other users can change. As a result, these users are required to recompute their partial crosstalk cancellers $\mathbf{P}_k$ for $1 \le k \le N$. For N tones, this requires $\mathcal{O}(NL(p+1)^3)$ operations. On one hand, the users should switch between the idle and nonidle states as

---

[2] Note that the cyclic prefix in DMT modulation results in a loss in the capacity of the twisted pair channel regardless of the capacity loss due to crosstalk.

quickly as possible to reduce their crosstalk levels and RFI radiation. On the other hand, the computational constraints of the modems can prevent them from updating their partial crosstalk cancellers, which annihilates the potential gains of partial crosstalk cancellation. Therefore, it is of crucial importance for partial crosstalk cancellers to have a low recurrent computational complexity. In this section, we propose two low complexity CMI techniques for DSL.

In order to evaluate the performances of our CMI techniques, we compare the bitrates of the DSL loops using our methods with their bitrates when the dominant crosstalk entries are removed from the channel. For each tone k, we define a *dominant crosstalk-cancelled* (DCC) channel by

$$[\mathbf{H}_k^{\mathrm{DCC}}]_{(i,j)} \triangleq \begin{cases} 0 & \text{if } j \text{ is a dominant} \\ & \text{crosstalker for user } i, \\ h_k^{(i,j)} & \text{otherwise.} \end{cases} \quad (4)$$

We call a partial crosstalk canceller an *ideal partial crosstalk canceller* if it enables the VDSL users to achieve the same bitrates as they would achieve if they were communicating over the DCC channel.

5.1. Approximate inverse CMI

The elements of $\overline{\mathbf{H}}_k$ corresponding to the nondominant crosstalkers can be zeroed to get a matrix $\overline{\mathbf{H}}_k^0 \triangleq \overline{\mathbf{H}}_k - \mathbf{\Lambda}^{-1}\mathbf{H}_k^{\mathrm{DCC}} + \mathbf{I}_L$, where $\mathbf{I}_L$ is the $L \times L$ identity matrix. Assuming an equal transmission power for all of the modems, $\overline{\mathbf{H}}_k^0$ is simply formed by zeroing the smaller elements in each row of $\overline{\mathbf{H}}_k$. For the case when the transmission power of all of the modems is not the same, the elements with minimum $s_k^{(j)}|h_k^{(i,j)}|^2$ are nulled, where $s_k^{(j)}$ is the transmission power of modem j on tone k.

The approximate inverse (AI) CMI method uses $(\overline{\mathbf{H}}_k^0)^{-1}$ as an approximation to the ideal partial crosstalk canceller and then uses a power-series approximation for matrix inversion to efficiently compute $(\overline{\mathbf{H}}_k^0)^{-1}$ and to get a sparse precoding matrix. Precise calculation of $(\overline{\mathbf{H}}_k^0)^{-1}$ requires $\mathcal{O}(L^3)$ operations for each tone. Moreover, $(\overline{\mathbf{H}}_k^0)^{-1}$ is not generally a sparse matrix, which is essential in complexity reduction. To overcome these problems, we use the first-order terms of a power-series expansion. Use of a power-series expansion for the inverse of the DS channel-transfer matrix has been proposed in [12] to decrease the computational complexity of perfect crosstalk cancellation. The results reported in [12] predict a poor performance for the first-order power-series expansion of the exact inverse of $\mathbf{H}_k$ for short loops. Nevertheless, here we demonstrate that this method can be effectively used for partial crosstalk cancellation.

Using the first-order terms of a power-series expansion for the inverse of $(\overline{\mathbf{H}}_k^0)^{-1}$, we have

$$\beta_k(\overline{\mathbf{H}}_k^0)^{-1} \simeq \overline{\beta}_k(2\mathbf{I}_L - \overline{\mathbf{H}}_k^0) \triangleq \mathbf{P}_k^{\mathrm{AI}}, \quad (5)$$

or

$$[\mathbf{P}_k^{\mathrm{AI}}]_{ij} = \begin{cases} \overline{\beta}_k, & i = j, \\ -\overline{\beta}_k[\overline{\mathbf{H}}_k^0]_{ij}, & i \neq j, \end{cases} \quad (6)$$

where $\overline{\beta}_k \simeq 1$ is a normalizing factor and $\mathbf{I}_L$ is the $L \times L$ identity matrix. In Appendix A, we show that $(\overline{\mathbf{H}}_k^0)^{-1}$ reduces the power of the dominant crosstalkers to a level much lower than the crosstalk due to the nondominant crosstalkers. In Appendix B, we show that the condition number of $\overline{\mathbf{H}}_k^0$ is bounded by

$$\frac{|\lambda_{\overline{\mathbf{H}}_k^0}^{\max}|}{|\lambda_{\overline{\mathbf{H}}_k^0}^{\min}|} \leq \frac{1 + p\alpha}{1 - p\alpha}, \quad (7)$$

where $\lambda_{\overline{\mathbf{H}}_k^0}^{\max}$ and $\lambda_{\overline{\mathbf{H}}_k^0}^{\min}$ are the biggest and smallest eigenvalues of $\overline{\mathbf{H}}_k^0$, respectively and $\alpha < 0.01$ (see Appendix A for the definition of α). The right-hand side of (7) approaches one as $\alpha \to 0$. Since the number of dominant crosstalkers to each user p is typically around 3 to 4, we expect that the power-series expansion has a fast convergence to $(\overline{\mathbf{H}}_k^0)^{-1}$. As we show later in our simulation results, in contrast to perfect crosstalk cancellation, the performance of the AI scheme is very close to that of the ideal solution when employed in partial crosstalk cancellation.

5.2. Reduced inverse CMI

In the reduced inverse (RI) CMI scheme, we compute and store the structure of the perfect crosstalk cancellers, that is, $(\overline{\mathbf{H}}_k)^{-1}$ for $1 \leq k \leq N$, at network setup. Since the DSL crosstalk channel is essentially stationary, this data does not need to be recalculated for long periods of time.[3] The partial crosstalk cancellers are then simply calculated each time there is a change in the set of dominant crosstalkers by zeroing the elements of $(\overline{\mathbf{H}}_k)^{-1}$ corresponding to the nondominant crosstalkers. This is written as

$$[\mathbf{P}_k^{\mathrm{RI}}]_{ij} \triangleq \beta_k \begin{cases} 0 & \text{if } j \text{ is a nondominant,} \\ & \text{crosstalker for user } i, \\ [(\overline{\mathbf{H}}_k)^{-1}]_{ij} & \text{otherwise.} \end{cases} \quad (8)$$

Our simulation results show that the RI scheme almost achieves the performance of the ideal partial crosstalk canceller.

5.3. Generalization to upstream direction

As we discuss in Appendix A, the downstream DSL channel exhibits row-wise diagonal dominancy. The upstream DSL

---

[3] Although the DSL channel is essentially stationary, it may change over time because of several reasons such as change in customer wiring or temperature change. In practice, the DSL MIMO channel estimates must be updated and the matrix channel inverses must be recalculated if the channel has changed. In this paper, we ignore the computational complexity due to DSL channel changes.

channel, on the other hand, exhibits columnwise diagonal dominancy (CWDD) [2] (i.e., the diagonal elements of the US channel matrix are much larger than the off-diagonal elements in the same column). Recall that for the DS direction, we formed the normalized channel transfer matrix by pre-multiplying $\mathbf{H}_k$ by $\mathbf{\Lambda}^{-1}$. For the US direction, we form the normalized channel transfer matrix by post-multiplying $\mathbf{H}_k$ by $\mathbf{\Lambda}^{-1}$, that is,

$$\underline{\mathbf{H}}_k \overset{\Delta}{=} \mathbf{H}_k \mathbf{\Lambda}^{-1}. \tag{9}$$

Using the CWDD property of the DS DSL channel, it is simple to show that

$$\alpha \overset{\Delta}{=} \max_i \max_{j \neq i} \underline{h}_k^{(i,j)} \ll 1, \tag{10}$$

where $\underline{h}_k^{(i,j)} \overset{\Delta}{=} [\underline{\mathbf{H}}_k]_{ij}$.

It has been shown in [4] that a ZF equalizer is near optimal for US DSL. The ZF equalizer for US DSL consists of multiplying the received vector by the inverse of the channel transfer matrix followed by a slicer. Based on (1), at the receiver we can estimate the transmitted signal vector $\mathbf{x}_k$ using

$$\hat{\mathbf{x}}_k = \mathbf{H}_k^{-1} \mathbf{y}_k = \mathbf{\Lambda}^{-1} (\mathbf{\Lambda} \mathbf{H}_k^{-1}) \mathbf{y}_k = \mathbf{\Lambda}^{-1} \underline{\mathbf{H}}_k^{-1} \mathbf{y}_k. \tag{11}$$

Note that there is no need for a normalizing factor β_k in the US direction. The diagonal matrix $\mathbf{\Lambda}^{-1}$ has only a scaling effect on the slicer's thresholds. As $\underline{\mathbf{H}}_k$ has exactly the same properties as $\overline{\mathbf{H}}_k$ in (3), it is trivial to show that all of the results we have obtained in Section 5 for the downstream channel can be generalized to the upstream channel. More importantly, the computationally efficient channel matrix inversion schemes proposed in Sections 5.1 and 5.2 can be implemented in an analogous way. As we show by using simulations in Section 8, the proposed CMI schemes also essentially achieve the performance of the ideal partial crosstalk canceller in the US direction.

It is important to note that if a prewhitening filter $\mathbf{W}_k$ is used, we must replace $\mathbf{H}_k$ by the equivalent noise-whitened channel $\mathbf{W}_k \mathbf{H}_k$ in the corresponding formulas. Unfortunately, the CWDD property may not necessarily hold for this channel. In this paper, we assume that the elements of the received noise vector $\mathbf{n}_k$ are independent, and we ignore the prewhitening filter $\mathbf{W}_k$. The CMI techniques that we have introduced in this paper can still be applied to channels with correlated noise. However, their performance might be degraded with respect to the simulation results in this paper.

6. JOINT TONE-LINE SELECTION

In JTLS schemes, both the frequency tones and the lines are considered to determine how to expend the available processing power to get the highest possible bitrates. In these schemes, the number of dominant crosstalkers that are cancelled varies from tone to tone and line to line. We let $\rho_k^{(\ell)}$ denote the number of crosstalkers that are cancelled on tone k of line ℓ. A JTLS algorithm first determines the value of $\rho_k^{(\ell)}$ for all users and tones. It then forms sparse partial crosstalk

cancellation matrices using a CMI scheme such as the AI and RI schemes proposed in Section 5.

Given pN multiplications per user (if an average of p dominant crosstalkers are cancelled per tone),[4] the JTLS problem for user ℓ is written as [6]

$$\max_{\{\rho_k^{(\ell)}\}_{k=1,\dots,N}} \sum_k c_k^{(\ell)} \quad \text{s.t.} \sum_k \rho_k^{(\ell)} \leq pN, \tag{12}$$

where $c_k^{(\ell)}$ is the number of bits that can be loaded on the kth tone of user ℓ after cancelling $\rho_k^{(\ell)}$ dominant crosstalkers on this tone. We assume that the power of the users and the channel values are constant each time the algorithm is run.

Assuming $\rho_k^{(\ell)}$ dominant crosstalkers are cancelled, $c_k^{(\ell)}$ is calculated as

$$c_k^{(\ell)} = \log_2 \left(1 + \frac{1}{\Gamma} \frac{s_k^{(\ell)} |h_k^{(\ell,\ell)}|^2}{\sigma_{(k,\ell)}^2 + \sum_{j=1, j \neq \ell, j \notin \mathcal{D}_k^\ell(\rho_k^{(\ell)})}^{L} s_k^{(j)} |h_k^{(\ell,j)}|^2} \right), \tag{13}$$

where $s_k^{(\ell)} = \mathrm{E}\{|x_k^{(\ell)}|^2\}$, $\sigma_{(k,\ell)}^2 = \mathrm{E}\{|n_k^{(\ell)}|^2\}$, Γ is the signal-to-noise power ratio (SNR) gap [15], and $\mathcal{D}_k^\ell(\rho_k^{(\ell)})$ is the set of the $\rho_k^{(\ell)}$ largest dominant crosstalkers for user ℓ in tone k. It is clear that the larger $\rho_k^{(\ell)}$ is, the larger $c_k^{(\ell)}$ is. Therefore, in practice the optimal solution satisfies the equality condition $\sum_k \rho_k^{(\ell)} = pN$. Note too that the alien noise power is contained in $\sigma_{(k,\ell)}^2$. Therefore, the optimal JTLS partial crosstalk canceller should be recalculated from time to time in the presence of alien time-varying crosstalk, even if the DSL system does not have power control mode itself.

A greedy joint tone-line selection algorithm has been proposed in [5, 6]. In this step-by-step algorithm, the benefit of cancelling any number of crosstalkers is calculated for all of the tones, and in each step the crosstalkers with the most benefit from cancellation are added to the cancellation list. After adding them, the benefit of crosstalk cancellation for the remaining crosstalkers is updated, and the process is repeated until all of the processing power is consumed. The benefit of cancelling ρ crosstalkers on tone k of line ℓ is calculated using $v_k^{(\ell)}(\rho) = (c_k^{(\ell)}(\rho) - c_k^{(\ell)}(0))/\rho$. At the initialization, the benefit is calculated for all values of $\rho = 1, \dots, L - 1$ and all of the tones $k = 1, \dots, N$ for line ℓ. During each iteration, first the maximum benefit value for line ℓ is selected. If we denote the tone and number of crosstalkers of the largest benefit value to be k_s and ρ_s, then the number of crosstalkers to be cancelled in tone k_s is set to ρ_s. And finally, the update process is performed by zeroing $v_{k_s}^{(\ell)}(\rho)$ for $1 \leq \rho \leq \rho_s$ and setting $v_{k_s}^{(\ell)}(\rho) = (c_{k_s}^{(\ell)}(\rho) - c_{k_s}^{(\ell)}(\rho_s))/(\rho - \rho_s)$ for $\rho_s + 1 \leq \rho \leq L - 1$.

---

[4] In practice, some of the N tones can be neglected, depending on the transmission direction and the bandplan. Accordingly, N should be replaced by the actual number of tones that are used in the transmission for that particular direction.

By inspection, we realize that the algorithm is optimal, as it expends each bit of processing power for the most possible benefit in each step. The algorithm requires up to NL sort operations, which can have sizes as large as NL [6]. Therefore, using a fast-sort algorithm with computational complexity $\mathcal{O}(NL\log_2(NL))$, the computational complexities of the algorithm for one user and for the total L users are $\mathcal{O}(N^2L^2\log_2(NL))$ and $\mathcal{O}(N^2L^3\log_2(NL))$, respectively. Given the large number of tones in VDSL and twisted pairs in a typical cable, it is clear that much faster algorithms are required for fast time-varying crosstalk environments. A suboptimal JTLS algorithm for upstream DSL is proposed in [5] with a computational complexity for one user of $\mathcal{O}(NL\log_2(NL))$. We now propose two types of novel JTLS algorithms for both downstream and upstream DSL with much lower computational complexities than the optimal algorithm proposed in [5, 6].

6.1. Subsort JTLS algorithms

The family of subsort JTLS algorithms contains heuristic algorithms derived from the greedy JTLS algorithm in [5, 6]. Consider the benefit value selected at each step. It is easy to show that the benefit value is less than the benefit value selected at the previous step. As a result, on average we expect the aggregate benefit of the selected tone in each step (i.e., $\bar{v}_{k_s}^{(\ell)}(\rho_s) = (c_{k_s}^{(\ell)}(\rho_s) - c_{k_s}^{(\ell)}(0))/\rho_s)$ to be less than the aggregate benefit value of the tone selected at the previous step. The class of subsort algorithms that we propose here is based on this observation.

In these algorithms, we first calculate the benefit values $v_k^{(\ell)}(\rho)$ for all values of k and ρ at the initialization. If we denote $\bar{v}_{k_s}^{(\ell)}(\rho_s) \triangleq \theta^*$ at the final step of the greedy algorithm, to find θ^*, we consider an arbitrary threshold value θ (e.g., $\theta = 0.5$) and then perform one of the following algorithms.

Algorithm 1. For each tone k find the smallest $\rho_k^{(\ell)}$ with benefit $v_k^{(\ell)}(\rho_k^{(\ell)}) \geq \theta$. Set $\rho_k^{(\ell)} = L - 1$, if no $\rho_k^{(\ell)}$ is found with $v_k^{(\ell)}(\rho_k^{(\ell)}) \geq \theta$. Search for the largest threshold value θ that satisfies $\sum_{k=1}^N \rho_k^{(\ell)} \leq pN$.

Algorithm 2. For each tone k find the largest $\rho_k^{(\ell)}$ with benefit $v_k^{(\ell)}(\rho_k^{(\ell)}) \leq \theta$. Set $\rho_k^{(\ell)} = 0$, if no $\rho_k^{(\ell)}$ is found with $v_k^{(\ell)}(\rho_k^{(\ell)}) \leq \theta$. Search for the largest threshold value θ that satisfies $\sum_{k=1}^N \rho_k^{(\ell)} \leq pN$.

Algorithm 3. For each tone k find the smallest $\rho_k^{(\ell)}$ with benefit $v_k^{(\ell)}(\rho_k^{(\ell)}) \leq \theta$. Set $\rho_k^{(\ell)} = L - 1$, if no $\rho_k^{(\ell)}$ is found with $v_k^{(\ell)}(\rho_k^{(\ell)}) \leq \theta$. Search for the smallest threshold value θ that satisfies $\sum_{k=1}^N \rho_k^{(\ell)} \leq pN$.

Algorithm 4. For each tone k find the largest $\rho_k^{(\ell)}$ with benefit $v_k^{(\ell)}(\rho_k^{(\ell)}) \geq \theta$. Set $\rho_k^{(\ell)} = 0$, if no $\rho_k^{(\ell)}$ is found with $v_k^{(\ell)}(\rho_k^{(\ell)}) \geq \theta$. Search for the smallest threshold value θ that satisfies $\sum_{k=1}^N \rho_k^{(\ell)} \leq pN$.

For the above algorithms to work, we need to show that we can find an appropriate value of threshold θ that satisfies the processing power constraint $\sum \rho_k^{(\ell)} \approx pN$. In fact, for any of these algorithms we will show that the processing power $\sum \rho_k^{(\ell)}$ is an increasing or decreasing function of θ.

Theorem 1. *The processing power $\sum \rho_k^{(\ell)}$ is an increasing function of threshold value θ in Algorithms 1 and 2 and a decreasing function in Algorithms 3 and 4.*

Proof. Here we prove Theorem 1 for the first algorithm and leave the others to the reader. Assume that $\theta_1 \geq \theta_2$ for an arbitrary tone k. We denote the values of $\rho_k^{(\ell)}$ corresponding to θ_1 and θ_2 by $\rho_k^{(\ell)}(\theta_1)$ and $\rho_k^{(\ell)}(\theta_2)$, respectively. For $\rho_k^{(\ell)}(\theta_1) = L - 1$, clearly $\rho_k^{(\ell)}(\theta_1) \geq \rho_k^{(\ell)}(\theta_2)$. For $\rho_k^{(\ell)}(\theta_1) < L - 1$, we have $v_k^{(\ell)}(\rho_k^{(\ell)}(\theta_1)) \geq \theta_1$, and thus $v_k^{(\ell)}(\rho_k^{(\ell)}(\theta_1)) \geq \theta_2$. Since $\rho_k^{(\ell)}(\theta_2)$ is the smallest number that satisfies $v_k^{(\ell)}(\cdot) \geq \theta_2$, we must have $\rho_k^{(\ell)}(\theta_1) \geq \rho_k^{(\ell)}(\theta_2)$. Summing over all values of k, we get $\sum \rho_k^{(\ell)}(\theta_1) \geq \sum \rho_k^{(\ell)}(\theta_2)$. $\qquad \square$

Theorem 1 guarantees that the processing power is a monotonic function of θ. Therefore, we can search for the proper value of θ that satisfies the processing power constraint by simply using classic search schemes such as a bisection search. However, note that this value is not necessarily equal to θ^*, because, as we will see later, the subsort algorithms do not yield the same results as the greedy algorithm.

6.2. Lagrangian JTLS algorithm

The Lagrangian JTLS algorithm is based on the Lagrange multiplier method for constrained optimization, which is written here as [16]

$$\max_{\{\rho_k^{(\ell)}\}_k} \mathcal{L} = \sum_k c_k^{(\ell)} + \lambda\left(pN - \sum_k \rho_k^{(\ell)}\right), \qquad (14)$$

where $\lambda \geq 0$ is the Lagrangian multiplier. The dimension of the Lagrangian in (14) is extremely large. However, note that $c_k^{(\ell)}$ is independent of $c_{k'}^{(\ell)}$ and $\rho_{k'}^{(\ell)}$ for $k \neq k'$. Therefore, following the methodology as in [17], we can decouple the Lagrangian in (14) into N independent Lagrangians per tone, as follows:

$$\max_{\rho_k^{(\ell)}} \mathcal{L}_k = c_k^{(\ell)} - \lambda\rho_k^{(\ell)}, \quad k = 1,\ldots,N. \qquad (15)$$

Note that $\mathcal{L} = \lambda pN + \sum_k \mathcal{L}_k$.

For a particular value of λ, the optimal value of $\rho_k^{(\ell)}$ is obtained by examining all integer values of $\rho_k^{(\ell)}$ from 0 to $L - 1$ in (15). The optimal value of λ, λ^* is the one that satisfies the processing constraint $\sum \rho_k^{(\ell)} \approx pN$. To find λ^*, we first start with an arbitrary value of λ (e.g., $\lambda = 1$) and compute $\rho_k^{(\ell)}$ for $1 \leq k \leq N$ from (15). Then, we increase or decrease λ,

conditioned on $\sum \rho_k^{(\ell)}$ being greater or less than pN, respectively. We repeat this procedure until λ converges. At convergence, either the processing constraint is satisfied or λ^* is zero.

The optimality of the algorithm could be shown if the primal problem in (12) was convex [16]. Although this cannot be shown for DSL channels, it has been shown that when a time-sharing property is valid, the Lagrange multiplier method is optimal in multicarrier systems [18]. For the time-sharing property to occur in multicarrier systems, the number of subcarriers contributing to the signal at the receiver side should be infinite.[5] This is practically the case in high SNR loops, where hundreds to thousands of tones contribute to the signal power. On the other hand, for low SNR loops, where only a few tones contribute to the signal power, the processing power is almost always enough for perfect crosstalk cancellation on all of these tones. It is easy to show that the Lagrange JTLS algorithm converges to the optimal solution in this case. This justifies why the Lagrange JTLS algorithm is always optimal in practice. As we will show, our computer simulations verify this conclusion. This algorithm has recently been independently proposed by Tsiaflakis et al. [11].

7. COMPUTATIONAL COMPLEXITY

The total computational complexity of the partial crosstalk cancellers is the sum of the online and recurrent computational complexities. The online computational complexity is pN operations for each user per each DMT symbol for both the LS and JTLS schemes, when an average of p crosstalkers is cancelled for each tone. The DMT symbol period is $250\,\mu s$ in VDSL. In the following sections, we study the order of the recurrent operations needed by the partial crosstalk cancellers when the crosstalk profile varies over time. For a binder with tens of VDSL loops carrying variable rate traffic, it is expected that recomputation of the structure of the partial crosstalk cancellers is required every few milliseconds. Therefore, a practical partial crosstalk canceller should require as few recurrent operations as possible.

7.1. *Computational complexity of LS schemes*

The recurrent operations associated with the LS schemes in fast time-varying crosstalk environments consist of the following two phases: (1) sorting the crosstalkers to determine the dominant crosstalkers (tracking), (2) calculation of the sparse partial crosstalk cancellation matrices based on the order of the crosstalkers and the value of p (CMI).

Phase 1. Tracking requires N sorts of size $L - 1$ for each user, which is of order $\mathcal{O}(N(L - 1)\log_2(L - 1))$. If the users transmit only at the maximum power mask level when working and at zero power when idle, we can use a radix sort [19] to

reduce the computational complexity to $\mathcal{O}(N(L-1))$.[6] Moreover, if we assume that only one crosstalker has changed its power, even for a random channel and unlimited power levels, resorting the crosstalkers requires only $\mathcal{O}(N(L - 1))$ operations.

Phase 2. CMI does not require any further data processing when the dominant crosstalkers are determined using our proposed AI and RI schemes. There are only NL assignment operations per user associated with (6) and (8). In comparison, note that using the method proposed in [5, 6] to construct the sparse partial crosstalk cancellers requires $\mathcal{O}(N(p + 1)^3)$ calculations for each user and a total of $\mathcal{O}(NL(p + 1)^3)$ operations for all users. Moreover, there are $N(L + p)$ assignment operations for this method as well.

7.2. *Computational complexity of JTLS schemes*

The recurrent operations associated with the JTLS schemes in fast time-varying crosstalk environments consist of the following four phases: (1) tracking, (2) evaluating $c_k^{(\ell)}(\cdot)$ and $v_k^{(\ell)}(\cdot)$ for $1 \leq k \leq N$, (3) determining $\rho_k^{(\ell)}$ by means of a JTLS algorithm, (4) implementing CMI. We studied the computational complexity of tracking and CMI in the previous section. The computational complexity of tracking in JTLS is the same as in LS. It can be shown that the computational complexity of CMI in JTLS is greater than or equal to LS, when the parameter p is the same for the two schemes. We now study the computational complexity associated with Phases 2 and 2 for the ℓth user when a change occurs in its crosstalk profile. The total computational complexity is L times the computational complexity for a single user ℓ.

Phase 2. After sorting the crosstalkers, $c_k^{(\ell)}(\rho)$ can be calculated using (13). Calculation of $c_k^{(\ell)}(\rho)$ ($0 \leq \rho \leq L - 1$) for each tone k can be done in $\mathcal{O}(L)$ operations by evaluating $c_k^{(\ell)}(\rho)$ for the maximum value of ρ (i.e., $L - 1$) down to 0. This way, calculation of $c_k^{(\ell)}(\rho)$ in (13) for each value of ρ can be done in $\mathcal{O}(1)$ operations, given that we have stored the denominator of the fraction of the previous value of ρ. Therefore, the calculation of $c_k^{(\ell)}(\cdot)$ and $v_k^{(\ell)}(\cdot)$ for all tones $1 \leq k \leq N$ can be done in $\mathcal{O}(NL)$ operations for each user.

Phase 3. Determining $\rho_k^{(\ell)}$ is the core phase of the JTLS algorithms. The computational complexity of this phase is $\mathcal{O}(N^2L^2\log_2(NL))$ for each user for the greedy algorithm proposed in [5, 6]. For all of the subsort algorithms, the computational complexity is $\mathcal{O}(K_1NL)$ for each user, where K_1 is the number of iterations required to find the proper value

---

[5] For a detailed definition of the time-sharing property and the proof of the optimality of the Lagrange optimization technique in multicarrier systems when the number of subcarriers is large see [18].

[6] The radix sort is feasible when the numbers to be sorted are from a finite set of preknown values. A Boolean array is formed with size equal to the total number of possible values. The index of each element corresponds to a particular value, and the indices are presorted. First, all of the elements are preset to false. In the sort process, the elements corresponding to numbers in the list are set to true. Finally, the sorted list can be achieved by reading the values with corresponding boolean entries set to true.

of the threshold θ. Similarly, the computational complexity of the proposed Lagrangian JTLS algorithm is $\mathcal{O}(K_2NL)$ for each user, where K_2 is the number of iterations required to find the optimal Lagrange multiplier λ^*.

The required number of iterations depends on the desired precision. Threshold value θ and Lagrangian multiplier λ have values ranging from 0 to 15, with at most 15 bits loaded on a given tone. Therefore, for the error to be smaller than ϵ_x, the number of required iterations is $\log_2(15/\epsilon_x) - 1$. For example, for $\epsilon_x = 0.01$ the number of required iterations is 10. Similarly, the maximum processing power corresponding to perfect crosstalk cancellation is $N(L-1)$. Therefore, for the error in processing power to be smaller than ϵ_y, we need $\log_2[N(L-1)/\epsilon_y] - 1$ iterations on average. If we assume an average of p-dominant crosstalkers to be cancelled per tone and 1% error in processing power, we get $\epsilon_y = 0.01Np$, and the number of required iterations is $\log_2[100(L-1)/p] - 1$. When the available processing power is 20% of the required processing power, the number of iterations is approximately 8. Our simulations results show that $K_1 = 10$ to 14 iterations and $K_2 = 6$ to 11 iterations are usually enough to find the optimal values of θ and λ almost exactly. K_1 and K_2 can possibly be reduced using faster search methods like the sub-gradient search method as explained in [18]. In summary, we see that the largest portion of the processing complexity (see Phase 3) can be reduced significantly using the algorithms proposed in this article.

7.3. Calculation of computational complexity

The order of the total required number of floating point operations per second (flops) can be calculated using

$$\mathcal{O}_{\text{total}} = f_s \times \mathcal{O}_{\text{online}} + f_{\text{update}} \times \mathcal{O}_{\text{recurrent}}, \qquad (16)$$

where $f_s = 4000$ Hz is the DMT symbol rate and $\mathcal{O}_{\text{online}}$ is the order of online computational complexity. $\mathcal{O}_{\text{online}}$ is equal to LN and pN for perfect and partial crosstalk cancellers, respectively. f_{update} is the update rate (the rate of change in the overall crosstalk profile) and $\mathcal{O}_{\text{recurrent}}$ is the order of recurrent computational complexity. $\mathcal{O}_{\text{recurrent}}$ is zero for perfect crosstalk cancellers. It can be calculated for LS and JTLS schemes by summing the computational complexities corresponding to the phases discussed in this section.

As an example of order calculation using (16), we let the number of users be $L = 25$ and the average number of dominant crosstalkers to be cancelled per tone be $p = 5$. Parameter N is set to 1174, the number of tones in the US direction of the VDSL FDD 998 bandplan [20]. With $L = 25$ and $p = 5$, the online computational complexity of the partial crosstalk cancellers is $pLN/L^2N = p/L = 20\%$ of the perfect crosstalk canceller. However, the recurrent computational complexities of the partial crosstalk cancellers increase with the update rate.

Figure 1 illustrates the order of the total number of flops per user that are required for a perfect crosstalk canceller (perfect CC), for the LS and JTLS partial crosstalk cancellers proposed in this paper (new LS and new JTLS assuming

TABLE 1: Simulation parameters.

Tone width	4.3125 kHz
Symbol rate	4 kHz
Γ	12.9 dB
Transmission power	-60 dBm/Hz
Cable type	26 Gauge (0.4 mm) [8]
Load resistance	135 Ω
Noise model	ETSI Noise Model A [20]
Target error Prob.	10^{-7}
Band plan	998 FDD Bandplan [20]

$K_1 = K_2 = 11$), and for the LS and JTLS partial crosstalk cancellers proposed in [5, 6] (old LS and old JTLS). We have provided closeups of Figure 1(a) in Figures 1(b) and 1(c) to make the results more readable. We can see in Figures 1(a) and 1(b) that the increase of the computation complexity as a function of the update rate is very small using our new AI and RI LS schemes (2.5% increase in the total computational complexity for $f_{\text{update}} = 500$ Hz). The old LS scheme, proposed in [5, 6], also performs well compared to the old JTLS scheme. It, however, increases the total computational complexity by 24% for $f_{\text{update}} = 500$ Hz.

As it can be seen in Figures 1(a) and 1(c), the total computational complexity of the JTLS scheme in [5, 6] is very large in rapidly-varying crosstalk environments. For update rates greater than about 0.2 Hz, the total computational complexity of the scheme is even higher than the perfect crosstalk canceller. Our new JTLS schemes, however, provide a significantly lower computational complexity. We note that in Figure 1(b), even for the very high update rate of 500 Hz (update every 2 milliseconds), the increase in the computational complexity due to the recurrent complexity is about 30%. In comparison, to keep the increase in the computational complexity below 30% in the old JTLS scheme, the update rate should be less than 0.014 Hz (update every 71 seconds).

8. SIMULATION RESULTS

Having compared the relative computational complexities of the schemes, we now use worst-case channel simulations to compare the performances of the various techniques from a crosstalk-cancellation point of view. We have simulated the proposed algorithms for two typical scenarios for both the DS and US directions. Scenario 1 is a distributed scenario, and Scenario 2 is a near-far scenario. Scenario 1 consists of 10 VDSL users with lines varying in length from 300 m to 1200 m in 100 m increments. Scenario 2 consists of five VDSL users with 600 m line lengths and five with 300 m line lengths. The channel transfer matrix, $\mathbf{H}_k$, is simulated using the one percent worst-case coupling model in [20] and the line transfer function of [8]. The simulation parameters are listed in Table 1.

To see the benefits of a partial crosstalk canceller, we need to simulate a crosstalk channel which has a few dominant crosstalkers. To do this, we model the space selectivity of crosstalk [6] by taking the distance-squared law of
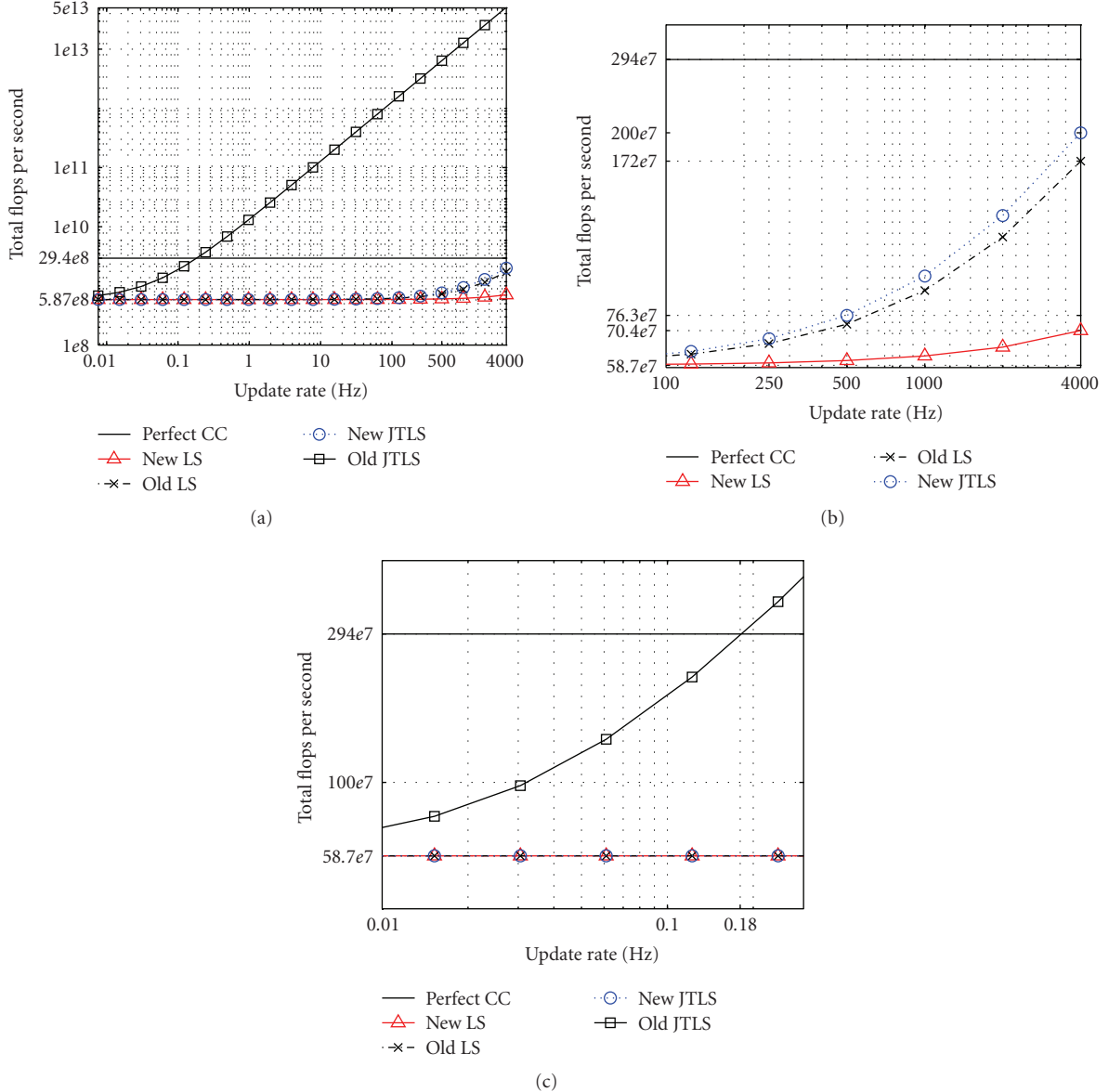
FIGURE 1: (a) The total number of flops (including online and recurrent complexities) per user for perfect crosstalk canceller (perfect CC), our new LS scheme (new LS), the LS scheme in [5, 6] (old LS), our new JTLS schemes (new JTLS) assuming $K_1 = K_2 = 11$, and the JTLS scheme in [5, 6] (old JTLS) for $L = 25$ and $p = 5$, (b) a closer look at the performance of our LS scheme, the LS scheme in [5, 6], and our new JTLS scheme and (c) a closer look at the performance of the JTLS scheme in [5, 6].

electromagnetic induction into account. **Figure 2** illustrates the cross-sections of the simulated 25-pair binder group for the two scenarios. Each circle represents a twisted pair. The length of each VDSL loop is written in the corresponding circle. The crosstalk couplings between pairs are considered to be inversely proportional to the square of the distance between the centers of the corresponding circles in **Figure 2**.[7] As a worst-case scenario, we select a tightly packed subset of

pairs at the center of the binder. The crosstalk couplings are normalized so that they are equal to the one percent worst-case model for tangent circles (e.g., pairs 1 and 2, 1 and 3, 2 and 7, etc.). If we order the crosstalkers by power, **Figure 3** shows the resultant cumulative average crosstalk power percentages for the 10 loops for the DS direction of Scenario 1, using the distance-squared law. This figure has a similar shape to the experimental measurements reported in **Figure 3** of [6].

Figures 4 and 5 show the performances of the LS schemes using the proposed CMI techniques for Scenarios 1 and 2, respectively. As it can be seen, both schemes nearly achieve the performance of the ideal LS partial crosstalk canceller. The

---

[7] The electromagnetic induction of twisted pairs into each other may not exactly follow the distance-squared law. However, our simulation results with a wide range of other powers for distance, ranging from $\sqrt{2}$ to 4, show that this does not affect the results reported in this article.
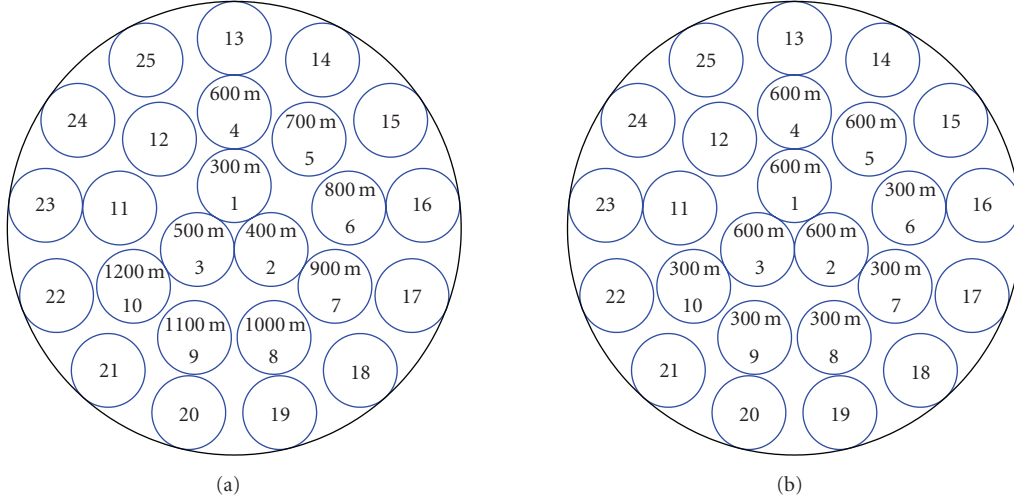
FIGURE 2: Cross-section of the binder and corresponding VDSL loop lengths: (a) distributed Scenario 1, (b) near-far Scenario 2.
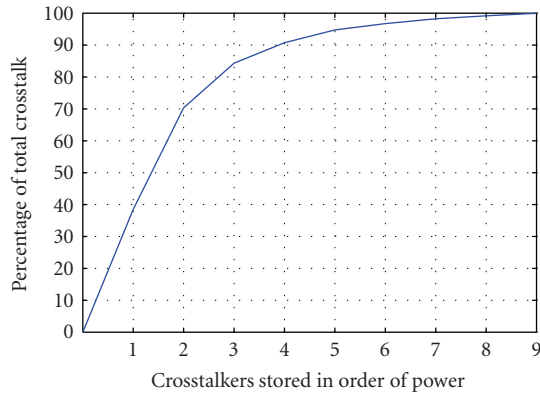

FIGURE 3: Cumulative average crosstalk percentages in DS direction for distributed Scenario 1 (crosstalkers are sorted by power).

RI scheme has a slightly superior performance to that of the AI scheme, especially for higher values of p. This is because of three phenomena. Firstly, as p increases a bigger fraction of the error is due to the residual crosstalk of the dominant crosstalkers for the AI scheme (compare (A.5) and (A.6)). Secondly, the condition number of $\overline{\mathbf{H}}_k^0$ increases as predicted by (7), and therefore, the error is bigger for the first-order terms of the power-series expansion of $\overline{\mathbf{H}}_k^0$. Thirdly, as p increases, the number of elements that should be eliminated from $\overline{\mathbf{H}}_k^{-1}$ decreases in the RI method, and therefore, the resultant matrix is a better approximation for the perfect crosstalk canceller. The cost we pay for using the RI scheme instead of the AI scheme is a higher complexity for matrix inversion and higher memory usage to store the channel inverse information.

Figures 6 and 7 illustrate the performance of the proposed JTLS algorithms compared to that of the optimal greedy algorithm. For each figure, the available processing power is governed by parameter p, the average number of dominant crosstalkers to be cancelled per tone. As it can be seen, the proposed algorithms can be sorted from the best to worst performance as follows: the Lagrange JTLS algorithm, subsort Algorithms 4, 3, 1, and 2. Among the subsort algorithms, the second one has the poorest performance and the fourth one has the best performance. As it can be seen, subsort Algorithm 4 has near optimal performance for most loop lengths in both scenarios. The Lagrange JTLS algorithm produces exactly the same performance as the optimal greedy algorithm for all loop lengths in both scenarios.

The fact that subsort Algorithm 4 has the best performance of the subsort algorithms can be explained by comparing it to the optimal greedy algorithm. Consider an arbitrary threshold value θ, and run the subsort Algorithm 4 at this threshold value. We denote the result of the algorithm by $\overline{\rho}_k^{(\ell)}$. Now consider the greedy algorithm being in the last step where the selected benefit is greater than θ (that is, the benefit value selected in the next step is less than θ), and denote the result of the greedy algorithm at this step by $\rho_k^{(\ell)}$. We can simply show that $\overline{v}_k^{(\ell)}(\rho_k^{(\ell)}) \geq \theta$. On the other hand, since $\overline{\rho}_k^{(\ell)}$ is the solution to subsort Algorithm 4 (i.e., $\overline{\rho}_k^{(\ell)}$ is the largest value that satisfies $\overline{v}_k^{(\ell)}(\overline{\rho}_k^{(\ell)}) \geq \theta$), we should have $\rho_k^{(\ell)} \leq \overline{\rho}_k^{(\ell)}$. When $\rho_k^{(\ell)} < \overline{\rho}_k^{(\ell)}$, using our assumption on the state of the greedy algorithm, we get $(c_k^{(\ell)}(\overline{\rho}_k^{(\ell)}) - c_k^{(\ell)}(\rho_k^{(\ell)}))/(\overline{\rho}_k^{(\ell)} - \rho_k^{(\ell)}) < \theta$. We know that $\overline{\rho}_k^{(\ell)}$ is greater than $\rho_k^{(\ell)}$ only for the tones that the aggregate benefit $\overline{v}_k^{(\ell)}(\rho_k^{(\ell)})$ is big enough to keep the aggregate benefit $\overline{v}_k^{(\ell)}(\overline{\rho}_k^{(\ell)})$ greater than θ. Since this phenomenon is unlikely to happen when the difference between $\overline{\rho}_k^{(\ell)}$ and $\rho_k^{(\ell)}$ is large, we expect that $\overline{\rho}_k^{(\ell)}$ and $\rho_k^{(\ell)}$ should have similar values, and consequently the fourth subsort algorithm should perform closely to the optimal greedy algorithm. Note that with the same threshold value θ, the greedy algorithm and the subsort Algorithm 4 do not necessarily require the same amount of processing power. However, we have just shown that for any value of θ the solution of the
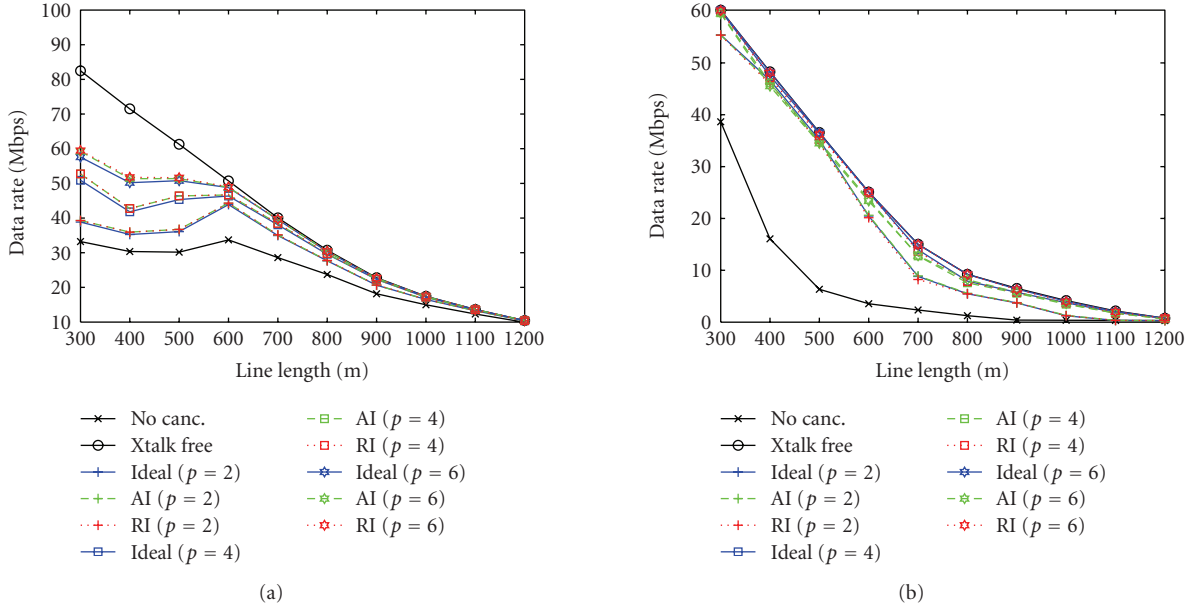
FIGURE 4: Data rates for LS scheme using the proposed AI and RI CMI methods compared to the ideal solution for distributed Scenario 1: (a) downstream and (b) upstream.
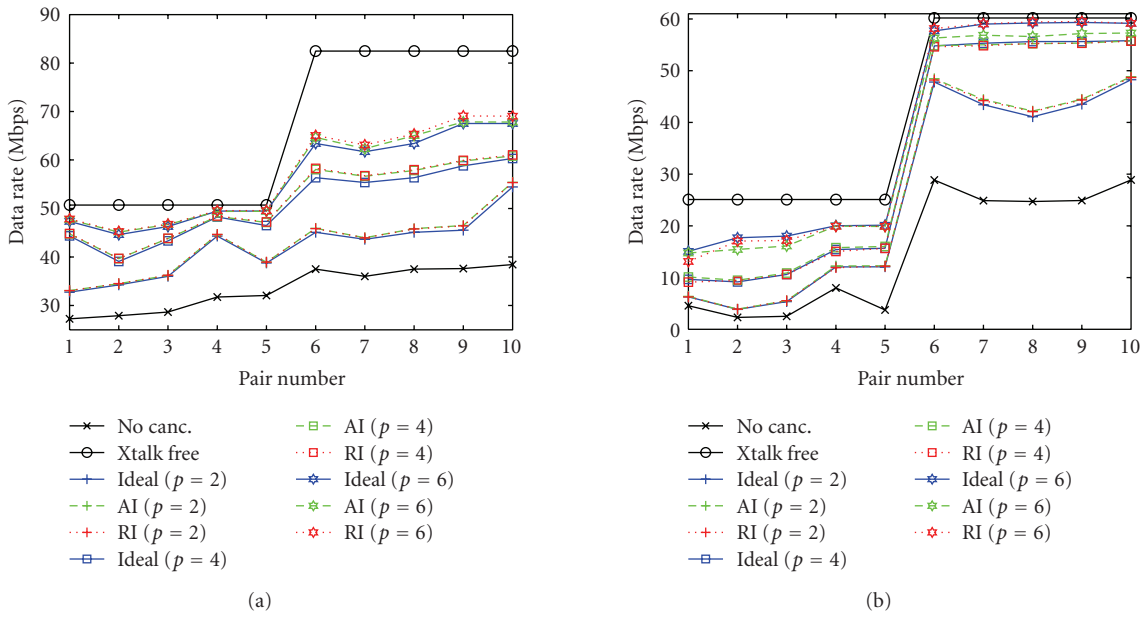

FIGURE 5: Data rates for LS scheme using the proposed AI and RI CMI methods compared to the ideal solution for near-far Scenario 2 (pairs 1 to 5: 600 m, pairs 6 to 10: 300 m): (a) downstream and (b) upstream.

subsort Algorithm 4 is close to the optimal solution. Since the required processing power is a monotonic function of the threshold value for both of the algorithms, we expect that the resultant processing power is also close to the processing power of the optimal solution.

Comparing the simulation results for DS and US, we see that the partial crosstalk cancellers are more beneficial for

short loops in the US direction. For example, as it can be seen in Figures 6(a) and 7(a), the data rates achieved by the optimal JTLS schemes are not as large as those achieved by perfect crosstalk cancellers (crosstalk free channel) for short loop lengths when the signal-to-noise ratio is high. We may justify this phenomenon by the following case study. The number of bits loaded on each tone using a partial crosstalk
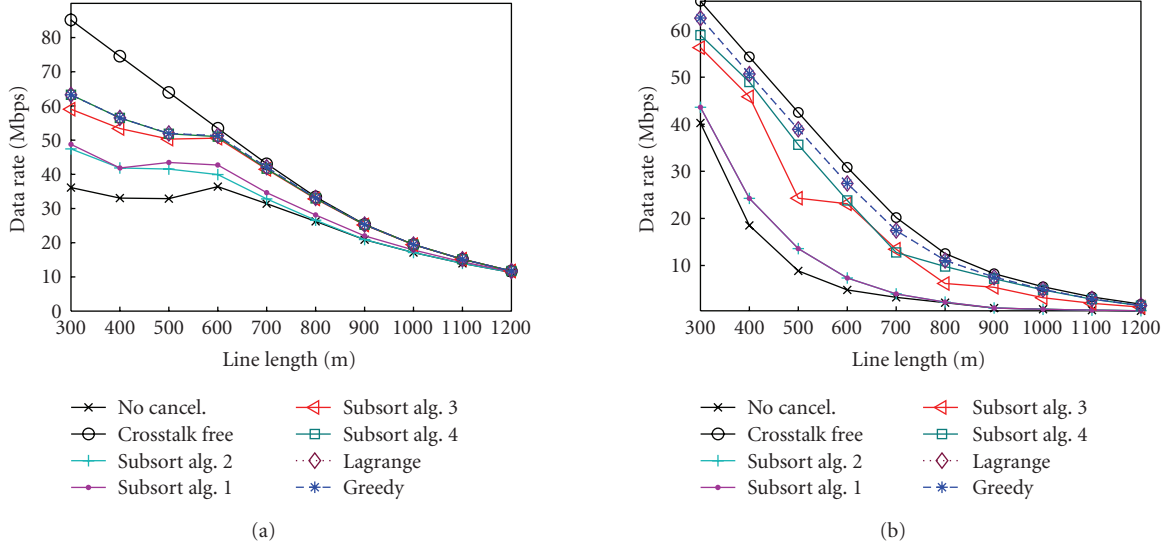
FIGURE 6: Data rates for proposed JTLS algorithms compared to the optimal solution for distributed Scenario 1: (a) downstream ($p = 4$) and (b) upstream ($p = 2$).

canceller that cancels 80% of the crosstalk can be written as

$$
\begin{aligned}
C_{80\% \text{ Canc.}} \\
&= \log_2 \left(1 + \frac{1}{\Gamma} \times \frac{\text{signal power}}{\text{noise power} + 20\% \text{ of crosstalk power}} \right) \\
&= \log_2 \left(\frac{\text{noise power} + 20\% \text{ of crosstalk power}}{0.2(5 \times \text{ noise power} + \text{ crosstalk power})} \right. \\
&\quad \left. + \frac{\Gamma^{-1} \text{signal power}}{0.2(5 \times \text{ noise power} + \text{ crosstalk power})} \right) \\
&< \log_2 \left(\frac{1}{0.2} \right) + \log_2 \left(1 + \frac{1}{\Gamma} \times \frac{\text{signal power}}{\text{noise power} + \text{crosstalk power}} \right) \\
&= 2.3 \text{ bits} + C_{\text{No Canc.}},
\end{aligned}
\tag{17}
$$

where $C_{\text{No Canc.}}$ is the capacity of the tone when there is no crosstalk cancellation. The average number of bits that can be loaded onto each tone of a 300 m loop with perfect crosstalk cancellation is about 15. Therefore, cancelling 80% of the crosstalk power results in an increase in data rate of less than $2.3/15 = 15\%$ of the perfect crosstalk cancellation rate. If we repeat the above steps for the cases where 95% and 99% of the crosstalk are cancelled, we get data rate increases of less than 29% and 44% of the crosstalk-free loop data rate, respectively.

On the other hand, it can be seen that the partial crosstalk canceller in US DSL operates much closer to the perfect crosstalk canceller than in DS DSL. This is because in the US channel the crosstalk couplings of the dominant crosstalkers are a few orders of magnitude greater than those of the nondominant crosstalkers due to the near-far effect. Therefore, in contrast to the DS case, cancelling these dominant crosstalkers blocks almost all of the crosstalk power.

The aforementioned case study together with the results obtained in Appendix A give us more insight as to why the AI scheme does not work well in high SNR loops, when used in perfect crosstalk cancellation, but works close to the ideal solution, when used in partial crosstalk cancellation. As shown in Appendix A, the AI technique reduces the residual crosstalk of the dominant crosstalkers to be small relative to the crosstalk due to the nondominant crosstalkers. Therefore, guided by our case study, we expect a small loss in performance compared to the ideal partial crosstalk canceller. On the other hand, when it is used for perfect crosstalk cancellation, the residual crosstalk of the crosstalkers can be much greater than the environment noise, resulting in a significant loss in performance compared to the performance of a perfect crosstalk canceller.

9. CONCLUSION AND FUTURE WORK

The DSL channel is essentially stationary. However, the crosstalk profile can change very rapidly in time for short-term stationary DSL systems. As the partial crosstalk cancellers involve recurring calculations for this type of traffic, it is crucial that they have low constructional complexity.

In this article, we have proposed new CMI and JTLS schemes for this purpose. We have studied the recurrent computational complexity of our schemes and showed that the proposed algorithms can effectively reduce the amount of required recurrent operations compared to previously proposed schemes.

Our AI and RI CMI schemes nearly achieve the performance of the ideal partial crosstalk canceller. The RI scheme requires more memory but provides a slightly superior performance to the AI scheme. The JTLS subsort Algorithm 4 produces a performance very close to that of the optimal greedy solution. In practice, the Lagrange JTLS algorithm almost achieves the performance of the optimal canceller, as shown by our extensive simulation results presented in this
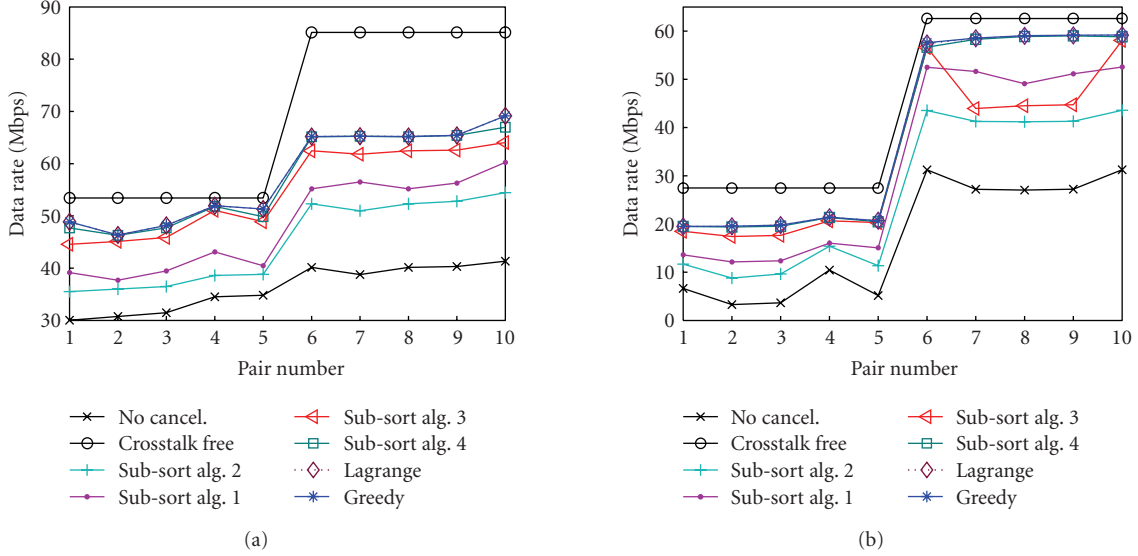
FIGURE 7: Data rates for proposed JTLS algorithms compared to the optimal solution for near-far Scenario 2 (pairs 1 to 5: 600 m, pairs 6 to 10: 300 m): (a) downstream ($p = 4$) and (b) upstream ($p = 4$).

paper and other simulation results not reported here. Our simulation results also show that the LS and JTLS schemes are particularly beneficial for crosstalk cancellation in short loops in the US direction.

To further refine these results, it would be of interest to simulate a VDSL network with variable-rate traffic, to find out how much processing power of the modems should be dedicated to the online portion of the partial crosstalk canceller and how often the structure of the partial crosstalk canceller should be updated, based on the characteristics of the traffic and the available processing power.

APPENDICES

A. CANCELLATION PROPERTIES OF $(\overline{\mathbf{H}}_k^0)^{-1}$

In this appendix, we consider the cancellation properties of $(\overline{\mathbf{H}}_k^0)^{-1}$. Without loss of generality, we assume that the normalizing factor β_k is equal to one. Let us define $\overline{\mathbf{H}}_k^{\text{ND}} \triangleq \overline{\mathbf{H}}_k - \overline{\mathbf{H}}_k^0$. Matrix $\overline{\mathbf{H}}_k^{\text{ND}}$ then contains the normalized crosstalk factors of the nondominant crosstalkers with its diagonal elements equal to zero. The net effect of the partial crosstalk precompensator $(\overline{\mathbf{H}}_k^0)^{-1}$ on the channel can be written as

$$\mathbf{H}_k(\overline{\mathbf{H}}_k^0)^{-1} = \mathbf{\Lambda}(\overline{\mathbf{H}}_k^0 + \overline{\mathbf{H}}_k^{\text{ND}})(\overline{\mathbf{H}}_k^0)^{-1} = \mathbf{\Lambda}\left[\mathbf{I} + \overline{\mathbf{H}}_k^{\text{ND}}(\overline{\mathbf{H}}_k^0)^{-1}\right].$$
(A.1)

Note that the DSL DS channel exhibits row-wise diagonal dominance (RWDD) [2]. That is, the diagonal elements of the DS channel transfer matrix $\mathbf{H}_k$ are much larger than the off-diagonal elements in the same row. In other words, for

any tone k and for any $j \neq i$, we have

$$|h_k^{(i,i)}| \gg |h_k^{(i,j)}|.$$
(A.2)

The magnitude of diagonal dominance is measured by parameter

$$\alpha \triangleq \max_i \max_{j \neq i} \frac{|h_k^{(i,j)}|}{|h_k^{(i,i)}|}.$$
(A.3)

Thus, we have $\alpha \ll 1$. Defining $\mathbf{G}^0 \triangleq \mathbf{I} - \overline{\mathbf{H}}_k^0$, from (6) we then have $\max_i \max_j |[\mathbf{G}^0]_{ij}| = \alpha$. Using a power-series expansion for $(\overline{\mathbf{H}}_k^0)^{-1}$, we can write

$$\mathbf{H}_k(\overline{\mathbf{H}}_k^0)^{-1} = \mathbf{\Lambda}\left\{\mathbf{I} + \overline{\mathbf{H}}_k^{\text{ND}}\left[\mathbf{I} + \sum_{\ell=1}^{\infty}(\mathbf{G}^0)^\ell\right]\right\}$$
$$= \mathbf{\Lambda} + \mathbf{\Lambda}\overline{\mathbf{H}}_k^{\text{ND}} + \mathbf{\Lambda}\overline{\mathbf{H}}_k^{\text{ND}}\sum_{\ell=1}^{\infty}(\mathbf{G}^0)^\ell.$$
(A.4)

The first and second terms of the last equality in (A.4) are the direct loop gains and the crosstalk factors of the nondominant crosstalkers, respectively. The third term, $\mathbf{\Lambda}\overline{\mathbf{H}}_k^{\text{ND}}\sum_{\ell=1}^{\infty}(\mathbf{G}^0)^\ell$, is the error term. The error term results in residual crosstalk from the dominant crosstalkers.

Consider an equal transmission power of $p_{\max}$ on all loops. Assuming the transmitted symbols from different users are independent and identically distributed, the average received crosstalk power on each loop due to the

nondominant crosstalkers is

$$
\frac{1}{L} \operatorname{Tr}\left[\mathbf{\Lambda} \overline{\mathbf{H}}_k^{\mathrm{ND}} \mathrm{E}\{\mathbf{x}_k \mathbf{x}_k^H\}\left(\overline{\mathbf{H}}_k^{\mathrm{ND}}\right)^H \mathbf{\Lambda}^H\right]
$$
$$
= \frac{p_{\max}}{L} \operatorname{Tr}\left[\mathbf{\Lambda} \overline{\mathbf{H}}_k^{\mathrm{ND}}\left(\overline{\mathbf{H}}_k^{\mathrm{ND}}\right)^H \mathbf{\Lambda}^H\right] \tag{A.5}
$$
$$
= \frac{p_{\max}}{L} \sum_{i,j=1}^L \left|\left[\mathbf{\Lambda} \overline{\mathbf{H}}_k^{\mathrm{ND}}\right]_{ij}\right|^2.
$$

Similarly, the average received power due to the error term is

$$
\frac{p_{\max}}{L} \operatorname{Tr}\left[\mathbf{\Lambda} \overline{\mathbf{H}}_k^{\mathrm{ND}} \mathbf{S}\mathbf{S}^H \left(\overline{\mathbf{H}}_k^{\mathrm{ND}}\right)^H \mathbf{\Lambda}^H\right]
$$
$$
= \frac{p_{\max}}{L} \sum_{i,j=1}^L \left|\left[\mathbf{\Lambda} \overline{\mathbf{H}}_k^{\mathrm{ND}} \mathbf{S}\right]_{ij}\right|^2
$$
$$
\leq \frac{p_{\max}}{L} \sum_{i,j=1}^L \left[\sum_{m=1}^L \left|\left[\mathbf{\Lambda} \overline{\mathbf{H}}_k^{\mathrm{ND}}\right]_{im}\right|^2 \sum_{n=1}^L \left|[\mathbf{S}]_{nj}\right|^2\right]
$$
$$
\leq \frac{pL\alpha^2}{(1-p\alpha)^2} \frac{p_{\max}}{L} \sum_{i,m=1}^L \left|\left[\mathbf{\Lambda} \overline{\mathbf{H}}_k^{\mathrm{ND}}\right]_{im}\right|^2, \tag{A.6}
$$

where $\mathbf{S} \triangleq \sum_{\ell=1}^\infty (\mathbf{G}^0)^\ell$. The first inequality follows from the relation $|\mathbf{w}^H \mathbf{z}|^2 \leq \|\mathbf{w}\|^2 \|\mathbf{y}\|^2$. The upper bound for the second inequality occurs in the worst-case scenario when the p dominant crosstalkers in $\mathbf{G}^0$ are the same for all L loops and their crosstalk couplings are at the maximum value α, that is,

$$
[\mathbf{G}^0]_{ij} = \begin{cases} \alpha, & 1 \leq j \leq p, \\ 0, & \text{otherwise.} \end{cases} \tag{A.7}
$$

Then we have

$$
\left[(\mathbf{G}^0)^\ell\right]_{ij} = \begin{cases} p^{\ell-1}\alpha^\ell, & 1 \leq j \leq p, \\ 0, & \text{otherwise,} \end{cases} \tag{A.8}
$$

and finally

$$
\sum_{n,j=1}^L \left|[\mathbf{S}]_{nj}\right|^2 \leq \frac{pL\alpha^2}{(1-p\alpha)^2}. \tag{A.9}
$$

Since $\alpha < 0.01$ [21], p is typically around 3 to 4, and $L = 25$, comparing (A.5) and (A.6), the average power due to the error will be a fraction of the average power due to the non-dominant crosstalkers.

B.   CONDITION NUMBER OF $\overline{\mathbf{H}}_k^0$

Define $\mathbf{G}^0 \triangleq \mathbf{I}_L - \overline{\mathbf{H}}_k^0$. We have $\max_i \max_j |[\mathbf{G}^0]_{ij}| = \alpha$ (see Appendix A). Now consider Perron's theorem [22, 23]: if $\{\mu_1, \ldots, \mu_L\}$ is an arbitrary set of positive numbers, then all eigenvalues $\{\lambda_G^\ell : \ell = 1, \ldots, L\}$ of matrix $\mathbf{G}$ lie on the disk $|z| \leq m_\mu$, where

$$
m_\mu = \max_{1 \leq i \leq L} \sum_{j=1}^L \frac{\mu_j}{\mu_i} \left|g^{(i,j)}\right|. \tag{B.1}
$$

By simply assuming $\mu_j = 1$ for all j, we get

$$
|\lambda_{G^0}^{\max}| \leq \max_{1 \leq i \leq L} \sum_{j=1}^L \left|[G^0]_{ij}\right| \leq p\alpha. \tag{B.2}
$$

To translate this result back to matrix $\overline{\mathbf{H}}_k^0$, note that from $\mathbf{G}^0 = \mathbf{I}_L - \overline{\mathbf{H}}_k^0$, we have $\lambda_{G^0} = 1 - \lambda_{\overline{H}_k^0}$. Therefore, we find that $|\lambda_{\overline{\mathbf{H}}_k^0}^{\max}| \leq 1 + p\alpha$, and as long as $p\alpha \leq 1$ we have $|\lambda_{\overline{\mathbf{H}}_k^0}^{\min}| \geq 1 - p\alpha$. Thus, the condition number of $\overline{\mathbf{H}}_k^0$ is bounded by

$$
\frac{|\lambda_{\overline{\mathbf{H}}_k^0}^{\max}|}{|\lambda_{\overline{\mathbf{H}}_k^0}^{\min}|} \leq \frac{1 + p\alpha}{1 - p\alpha}. \tag{B.3}
$$

REFERENCES

[1] T. M. Starr, M. Sorbara, J. M. Cioffi, and P. J. Silverman, *DSL Advances*, Prentice-Hall, Upper Saddle River, NJ, USA, 2003.

[2] G. Ginis and J. M. Cioffi, "Vectored transmission for digital subscriber line systems," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1085–1104, 2002.

[3] R. Cendrillon, M. Moonen, J. Verlinden, T. Bostoen, and G. Ginis, "Improved linear crosstalk precompensation for DSL," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '04)*, vol. 4, pp. 1053–1056, Montreal, Canada, May 2004.

[4] R. Cendrillon, M. Moonen, E. Van Den Bogaert, and G. Ginis, "The linear zero-forcing crosstalk canceler is near-optimal in DSL channels," in *Proccedings of IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 4, pp. 2334–2338, Dallas, Tex, USA, November-December 2004.

[5] R. Cendrillon, M. Moonen, G. Ginis, K. Van Acker, T. Bostoen, and P. Vandaele, "Partial crosstalk cancellation for upstream VDSL," *EURASIP Journal on Applied Signal Processing*, vol. 2004, no. 10, pp. 1520–1535, 2004.

[6] R. Cendrillon, G. Ginis, M. Moonen, and K. Van Acker, "Partial crosstalk precompensation in downstream VDSL," *Signal Processing*, vol. 84, no. 11, pp. 2005–2019, 2004.

[7] A. A. Salvekar, J. Louveaux, C. Aldana, J. L. Fang, E. de Carvalho, and J. M. Cioffi, "Profile detection in multiuser digital subscriber line systems," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, pp. 1116–1125, 2002.

[8] "Spectrum management for loop transmission systems," (Draft) ANSI Standard T1.417-2003, ANSI, Washington, DC, USA, February 2003.

[9] J. Cioffi, "Very-high-speed digital subscriber lines - system requirements," Tech. Rep. T1E1.4/98 Contribution 043R6, ANSI, Washington, DC, USA, November 1998.

[10] C. Zeng, C. Aldana, A. A. Salvekar, and J. M. Cioffi, "Crosstalk identification in xDSL systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1488–1496, 2001.

[11] P. Tsiaflakis, J. Vangorp, M. Moonen, J. Verlinden, and G. Ysebaert, "Partial crosstalk cancellation in a multi-user xDSL environment," in *Proceedings of IEEE International Conference on Communications (ICC '06)*, vol. 7, pp. 3264–3269, Istanbul, Turkey, June 2006.

[12] A. Leshem and L. Youming, "A low complexity coordinated FEXT cancellation for VDSL," in *Proceedings of the 11th IEEE International Conference on Electronics, Circuits and Systems (ICECS '04)*, pp. 338–341, Tel Aviv, Israel, December 2004.

[13] G. Ginis, "Multi-line coordinated communication for broadband access networks," Ph.D. dissertation, Stanford University, Stanford, Calif, USA, 2002.

[14] R. Cendrillon, M. Moonen, D. Gore, and A. Paulraj, "Low complexity crosstalk cancellation through line selection in upstream VDSL," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '03)*, vol. 4, pp. 692–695, Hong Kong, April 2003.

[15] T. M. Starr, J. M. Cioffi, and P. J. Silverman, *Understanding Digital Subscriber Line Technology*, Prentice-Hall, Upper Saddle River, NJ, USA, 1999.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[17] R. Cendrillon, M. Moonen, J. Verliden, T. Bostoen, and W. Yu, "Optimal multi-user spectrum management for digital subscriber lines," in *Proceedings of IEEE International Conference on Communications (ICC '04)*, vol. 1, pp. 1–5, Paris, France, June 2004.

[18] W. Yu, R. Lui, and R. Cendrillon, "Dual optimization methods for multiuser orthogonal frequency division multiplex systems," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '04)*, vol. 1, pp. 225–229, Dallas, Tex, USA, November-December 2004.

[19] R. Sedgewick, *Algorithms*, Addison-Wesley, Reading, Mass, USA, 2nd edition, 1988.

[20] ETSI, "Transmission and Multiplexing (TM); access transmission systems on metallic access cables; very high speed digital subscriber line (VDSL)—part I: functional requirements," ETSI Std. TS 101 270-1, Rev. V.1.3.1, 2003.

[21] R. Cendrillon, M. Moonen, R. Suciu, and G. Ginis, "Simplified Power Allocation and TX/RX Structure for MIMO-DSL," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '03)*, vol. 4, pp. 1842–1846, San Francisco, Calif, USA, December 2003.

[22] E. W. Weisstein, "Perron's theorem," http://mathworld.wolfram.com/PerronsTheorem.html.

[23] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, Academic Press, San Diego, Calif, USA, 2000.

Lee M. Garth received the B.S.E. degree (magna cum laude) from Princeton University, Princeton, NJ, in 1987, and the M.S. and Ph.D. degrees from the University of Illinois at Urbana-Champaign, in 1989 and 1996, respectively. He has had summer employment with Raytheon Company, GTE Corporation, and MITRE Corporation. From 1990 to 1996, he was a Senior Engineer at Techno-Sciences, Inc., Urbana, IL. From 1996 to 2000, he was a Member of the Advanced Data Communications Group of Bell Laboratories within Lucent Technologies, Holmdel, NJ. Since 2000, he has been a Faculty Member with the Department of Electrical and Electronic Engineering at the University of Canterbury, Christchurch, New Zealand. In 2006, he held a visiting appointment at the Samsung Advanced Institute of Technology in South Korea. His research interests include signal detection, array processing, adaptive equalization, and statistical signal processing with applications to communications systems. He is a Senior Member of the IEEE and a Member of Tau Beta Pi.

Amir R. Forouzan received his B.S. and M.S. degrees in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1998 and 2000, respectively, and his Ph.D. degree with highest distinction from University of Tehran in 2004. From August 1999 to May 2004, he was with the Iran Telecommunication Research Center as a Research Fellow. Since June 2004, he has been a Postdoctoral Fellow with the University of Canterbury (UC), Christchurch, New Zealand. During this period, he has been working on dynamic spectrum management and crosstalk cancellation schemes for digital subscriber lines (DSL). Since February 2006, he has been involved in establishing the New Zealand Broadband Research Facility at the UC, where he has led a software project team in developing an advanced DSL simulation tool. His research interests include dynamic spectrum management in DSL, optical CDMA, MIMO communication systems and information theory, and ultrawideband radio.