Université catholique de Louvain
Ecole Polythecnique de Louvain
LABORATOIRE DE
TÉLÉCOMMUNICATIONS
ET
TÉLÉDÉTECTION

B - 1348 Louvain-la-Neuve Belgique

DSL: Crosstalk Estimation Techniques, and Impact on DSM

Ali Kalakech

Thesis presented for the Ph. D. degree in
Engineering Sciences

Jury:

Luc Vandendorpe (UCL) - Supervisor
Jérôme Louveaux (UCL) - Co-Supervisor
Marc Moonen (KUL)
Jérôme Le Masson (ESM St-Cyr)
Koen Vanbleu (Broadcom)
Christophe De Vleeschouwer (UCL) - Chairman

September 2010

Remerciements

Effectuer une thèse de doctorat, c'est apprendre à avoir un esprit analytique, c'est apprendre à approcher les problèmes et savoir comment les résoudre. Par contre, effectuer une thèse c'est aussi apprendre à savoir s'arrêter parfois et à changer de piste. Effectuer une thèse de doctorat correspond à des moments de joie quand on trouve une solution originale mais ça peut aussi amener à des déceptions voir même de la dépression si on n'est pas entouré par les bonnes personnes. Après 5 longues années pleines de joie et de peu de déceptions, j'ai finalement réussi à entamer mon travail de doctorat. Cette réussite a été possible grâce à l'aide de plusieurs personnes que je tiens à remercier.

Tout d'abord je voudrais remercier mes parents qui m'ont encouragé à faire le choix de venir en Europe pour continuer mes études supérieures, malgré le fait que ça entrainait de longues périodes loin d'eux.

Je voudrais remercier mon promoteur Luc Vandendorpe, pour m'avoir accueilli au sein de l'équipe digicom, pour la confiance qu'il a portée en moi, et pour la liberté qu'il m'a accordé. Je remercie également mon co-promoteur Jérôme Louveaux pour tous les conseils et les astuces qu'il m'a donnés.

Je remercie les membres de mon jury pour leurs critiques constructives et leur contribution à l'amélioration de ce texte.

Je tiens aussi à remercier mes collègues de bureau Félix Brah, et Achraf Mallat, merci pour votre soutien et votre aide, je n'oublierai jamais nos discussions scientifiques et non scientifiques.

Je remercie également Samer Houri mon visiteur habituel, ainsi que tous les membres, anciens et actuels, du Labo TELE. Merci à Rony, Rémi, Annabelle, Sébastien, Bertrand, Xavier, David, Gemma, Vaibhav, Mohieddine,... et les autres. Vous avez réussi à faire de l'environnement académique du Labo TELE, un environnement sérieux mais aussi amical, gai, et plaisant. Je suis fier d'avoir fait partie de cet environnement.

Finalement je voudrais remercier Géraldine et sa maman, vous avez été une vraie famille pour moi, et votre soutien moral était essentiel durant cette période de ma vie.

Abstract

DSL systems "Digitial Subscriber Line", take advantage of the long time not used high frequency bands of twisted pairs in order to transmit signals that may go up to 30 MHz. This band is much higher than the traditional voice band of 4 kHz. Going up in frequency allows DSL systems to transmit at a very high Bit-Rates (24 Mbps in ADSL2+, 100 Mbps for VDSL), however it also poses various challenges. One of the challenges that may occur, is the increase of the interference between different lines, called crosstalk, at higher frequencies. Another drawback of using higher frequencies, is the fact that the direct channels of the twisted pairs are attenuated severly for long lines at higher frequencies. The attenuation of the direct channel gain for longer lines, coupled with crosstalk caused by other shorter lines in proximity would create a Near-Far problem in DSL systems when short and long lines are mixed together. Dynamic Spectrum Management or DSM was proposed to mitigate the effect of crosstalk on DSL systems. DSM exploits the information about the direct line channels, the crosstalk channels, and the ambient noise to improve the total system capacity. We may distinguish two DSM concepts: PSD coordination and Signal coordination. Both concepts necessitate the knowledge of crosstalk channels information, something that is not available under the current DSL systems.

In this thesis we consider the application of DSM algorithms under practical and realistic condition. In the first part we propose the Balanced Capacity concept as a possible solution for the Near-Far problem, then we proceed with the enhancement of several state of the art algorithms regarding DSM implementation. In the second part, we propose several practical estimation techniques for the crosstalk channel estimation. We start by estimating the crosstalk channel gain, which is sufficient for the application of PSD coordination algorithms, and then we proceed by the estimation of the total crosstalk channel which is required for the implementation of a full signal coordination. All of these estimation techniques are based on a passive or on a limited active observation of the DSL lines. Thus these estimators are practical, and can be used with the current DSL systems with minor modifications on the standards or on the users' modems.

Notations and Acronyms

Mathematical notations

j	imaginary unit: $j = \sqrt{-1}$
x	a scalar
x^*	the complex conjugate of x
$\mathrm{Re}(x)$	the real part of x
$\mathrm{Im}(x)$	the imaginary part of x
$\|x\|$	the absolute value or norm of x
$\mathbf{A}^T$	the transpose of matrix $\mathbf{A}$
$\mathbf{A}^H$	the conjugate transpose of matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	the inverse of matrix $\mathbf{A}$
$\mathrm{E}[x]$	the expected value (or mathematical expectation) of x
$\cong$	is approximately equal to
$[x]^+$	the positive part of x, i.e. $[x]^+ \triangleq \max(0, x)$

Acronyms

ADSL	Asymmetric Digital Subscriber Line
ANSI	American National Standards Institute
ASB	Autonomous Spectrum Balancing
ASP	Additional Starting Point
AWG	American Wire Gauge
AWGN	Additive White Gaussian Noise
BC	Balanced Capacity
BER	Bit Error Rate
BP	Band Preference
CIR	Carrier to Interference Impulse Ratio
CLS	Constrained Least Square
CO	Central Office
CP	Cyclic Prefix
DFT	Discrete Fourier Transform
DMT	Discrete Multi-Tone
DS	DownStream
DSL	Digital Subscriber line
DSM	Dynamic Spectrum Management
FDD	Frequency Division Duplexing
FDMA	Frequency Division Multiple Access
FEXT	Far End Crosstalk
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FTTB	Fiber To The Building
FTTH	Fiber To The Home
IDFT	Inverse Discrete Fourier Transform
IEEE	Institute of Electrical and Electronics Engineers
IFFT	Inverse Fast Fourier Transform
i.i.d.	independent and identically distributed
IPLS	Iterative Penalized Least Square
ISB	Iterative Spectrum Balancing
ISI	Inter Symbol Interference
IWI	Iterative Water Filling
LS	Least Square

MC	Multi-Carrier
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MSE	Mean Square Error
NEXT	Near End Crosstalk
NR	Newton Raphson
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
ONU	Optical Network Unit
OSB	Optimal Spectrum Balancing
PBO	Power Back-Off
pdf	probability density function
PLS	Penalized Least Square
PSD	Power Spectrum Density
SA	Steepest Ascent
SMC	Spectrum Management Center
SNR	Signal to Noise Ratio
SO	Successive Optimization
US	UpStream
VDSL	Very High Bitrate DSL
ZF	Zero Forcing

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 DSL Systems

Due to the high costs related to the fiber optics installation, it is still not possible to implement fiber to home technology at once and in many regions. DSL systems can be seen as the last mile solution to this problem, as it establishes the connection between remote users and the optical fibers present at a central office (CO) or at an optical network unit (ONU), using the available twisted pair lines of the telephone companies.

DSL systems allow a progressive implementation of the fiber to home technology. At the beginning, the fiber optics were installed at the central office, and the ADSL system was used to serve the remote users. ADSL for asymmetric digital subscriber line, exploited the existent local loops between the user terminal and the CO to make the connection possible. As fiber optics technology matured, the fibers were brought further more near the location of the users, by installing an optical network unit (ONU) near major urban areas, thus shortening the distance between the users and the fiber optics, which allowed the use of a very high frequency bands in VDSL system. In the future, fiber-to-the-building (FTTB) implementation is currently considered, and eventually fiber-to-the-home (FTTH). As the DSL systems are exploiting the local loops of an already existent telephony infrastructure, and since the fiber optics technology remains expensive for the time being, DSL systems are seen to be economically viable for the upcoming years.

1.1.1 Direct Lines and Crosstalk channels

1.1.1.1 Direct Channels

The direct line channel of a local loop can be derived from the electromagnetic characteristic of the twisted pair. If we consider a twisted pair line i, with no discontinuities nor mismatching, we can derive the direct channel of the line i as explained in the transmission line theory [1]:

$$H_{i,i}(f) = \exp\left(-\gamma_i(f) L_{tp,i}\right), \tag{1.1}$$

where $H_{i,i}(f)$ is the direct channel of the twisted pair line i at the frequency f, $L_{tp,i}$ is the length of line i, γ_i is the propagation constant. γ_i is given in function of the RLCG parameters of line i:

$$\gamma_i = \sqrt{(G_i'(f) + j2\pi f C_i'(f))(R_i'(f) + j2\pi f L_i'(f))} \tag{1.2}$$

where:

- R_i' is the resistance per unit length in Ohms/m
- L_i' is the inductance per unit length in Henry/m
- C_i' is the capacitance per unit length in Farad/m
- G_i' is the Electrical Conductivity in Siemens/m

R_i', L_i', C_i', G_i' vary with frequency, they represent the RLCG parameters that characterize a transmission line. In this thesis we have used the RLCG parameters of the twisted pair model used in [2, 3] where:

$$\begin{aligned}
R_i'(f) &= \frac{2d_i/a_i}{\sqrt{(d_i/a_i)^2 - 1}} \Re\{Z_{0i}(f)\} \\
L_i'(f) &= T_{ind}\frac{\mu_0}{\pi} cosh^{-1}(d_i/a_i) + \frac{d_i/a_i}{\pi f \sqrt{(d_i/a_i)^2 - 1} \Im\{Z_0(f)\}} \\
C_i'(f) &= \frac{\pi \varepsilon_r \varepsilon_0}{cosh^{-1}(d_i/a_i)} \\
G_i'(f) &= 0
\end{aligned}$$

where d_i is the center-spacing between the two wires of the twisted pairs, a_i is the diameter of the conductor, $\mu_0 = 4\pi * 10 - 7H/m$ is the permeability of air, T_{ind} is the twisting index, $\varepsilon_0 = 8.854{*}10{-}12F/m$ is the permittivity of air, ε_r is the relative permittivity of the insulation, and $Z_{0i}(f)$ is the internal impedance of the twisted pair and it is defined as:

$$Z_{0i}(f) = \frac{1}{d}\sqrt{\frac{2f\mu_0}{j\pi\sigma}\frac{J_0(d\sqrt{(\pi\mu_0\sigma/2j)f})}{J_1(d\sqrt{(\pi\mu_0\sigma/2j)f})}} \tag{1.3}$$

where σ la conductivity in Siemens/meter, and $J_n(x)$ is the Bessel function of order n. The values of a_i, d_i, σ, and T_{ind} are given for different types of twisted pairs in ANSI T1.601-1999 Standard [4]. For example, for a 24 AWG (American Wire Gauge) twisted pair we have the following characteristics at the room temperature (21 degree Celcius):

a	d	σ	T_{ind}	ε_r
0.67746 mm	0.85103 mm	$5.3256 * 10^7$ siemens/m	1.5758	1.3004

The channel model given by (1.1) represents a simple attenuation model. In this model, the direct channel gain decreases when either the frequency or the line length increases. Fig.1.1 shows the channel gain attenuation model in function of frequency for several lines having different lengths.

The channel gain attenuation model represents the direct channel gain when there are no impairments on the line, however numerous impairments can occur on the line. For example mismatching can happen at the source or at the end of a local loop when the impedance of the line is different then that of the source or of the load (look at Fig. 1.2). Another impairment that can occur, is the existence of a bridged tap on the line. Bridged tap represents a stub that is attached on the twisted pair line, it represents a "T" branch on the telephone cable as shown in Fig. 1.3. Mismatching and bridged taps cause reflections on the telephone cable, thus a part of the transmitted signal is reflected back, this would attenuate the signal severely on the line depending on the frequency. The attenuation caused by the impairment on the local loop are frequency selective.

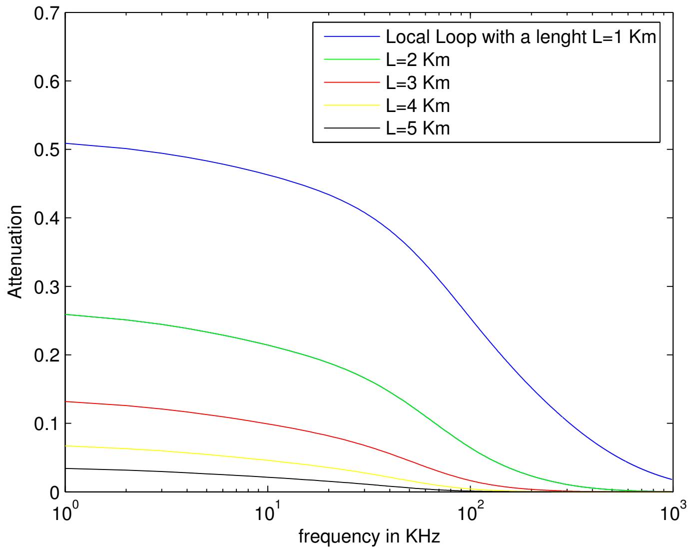


Figure 1.1: *Attenuation Channel Model*

One could include the effect of mismatching, or even the presence of a bridged tap, on the channel model of the line. Using an approach similar to [5, 6] we can derive the following models:

Mismatching model:

$$H_{ii}(f) == \frac{Z_{in}}{Z_s + Z_{in}}(1 + r_l)exp(-\gamma_{0i}L_i)(\frac{1}{1 - r_l r_s exp(-2\gamma_{0i}L_i)}) \quad (1.4)$$

where L_1 is the length of the line, γ_{01} is the propagation constant, Z_{in} is the imput impedance of the line, Z_{01} is the characteristic impedance, Z_s is the resistance at the source, Z_l is the resistance at the load, r_s and r_l are the reflection coefficient at the source and at the load respectively where $r_v = (Z_{01} - Z_v) / (Z_{01} + Z_v)$, where $r_v = r_s$ for $Z_v = Z_s$ and $r_v = r_l$ for $Z_v = Z_l$.

Bridged tap model: We consider a line with 3 sections as shown in Fig. 1.3, with each line section having different characteristics from

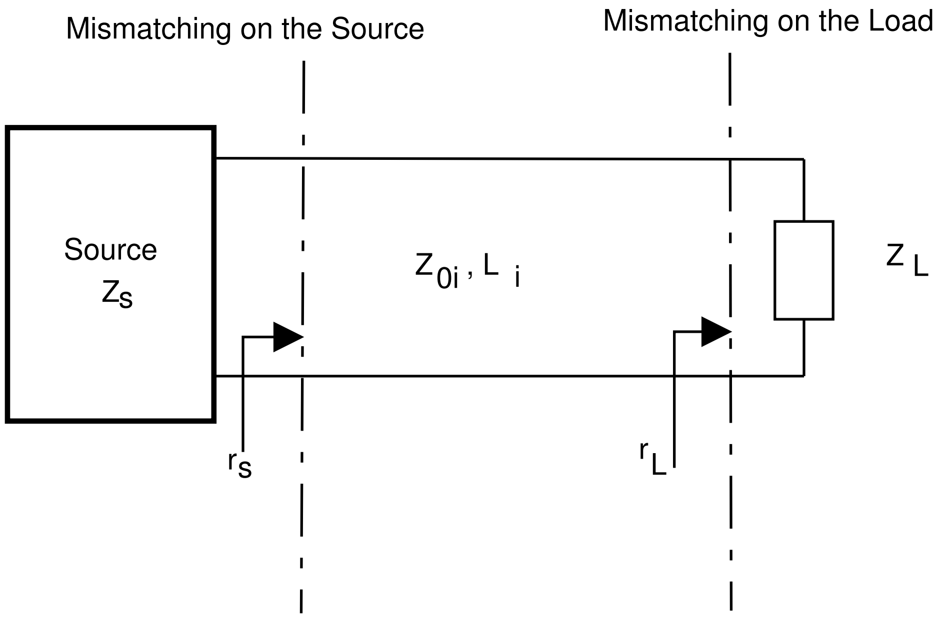


Figure 1.2: *Mismatching Effects*

the other 2 sections. Section 1 of the line corresponds to the part that start at the source and ends at the bridged tap, it has a length L_1 and a propagation constant γ_{01}. Section two corresponds to the part that goes from the bridged tap to the load, it has a length L_2 and a propagation constant equal to γ_{02}. Section 3 corresponds to the bridged tap itself, with a length L_3 and a propagation constant γ_{03}. The bridged tap model is given by:

$$H_{ii}(f) = (1+r_I)exp(-\gamma_{01}L_1)exp(-\gamma_{02}L_2)(1+(1+r_t)\frac{exp(-2\gamma_{03}L_3)}{1-r_t exp(-2\gamma_{03}L_{i3})}) \quad (1.5)$$

where r_I is the reflection coefficient of the echo caused by the bridged tap on section 1 of line, while r_t is the reflection coefficient on the bridged tap section (section 3 of the line).

We compare the attenuation model, the mismatching model, and the bridged tap model in Fig.1.4. From this figure, we can conclude that the direct channel gain of twisted pair is a frequency selective channel. Adding a bridged tap increases the frequency selectivity of the

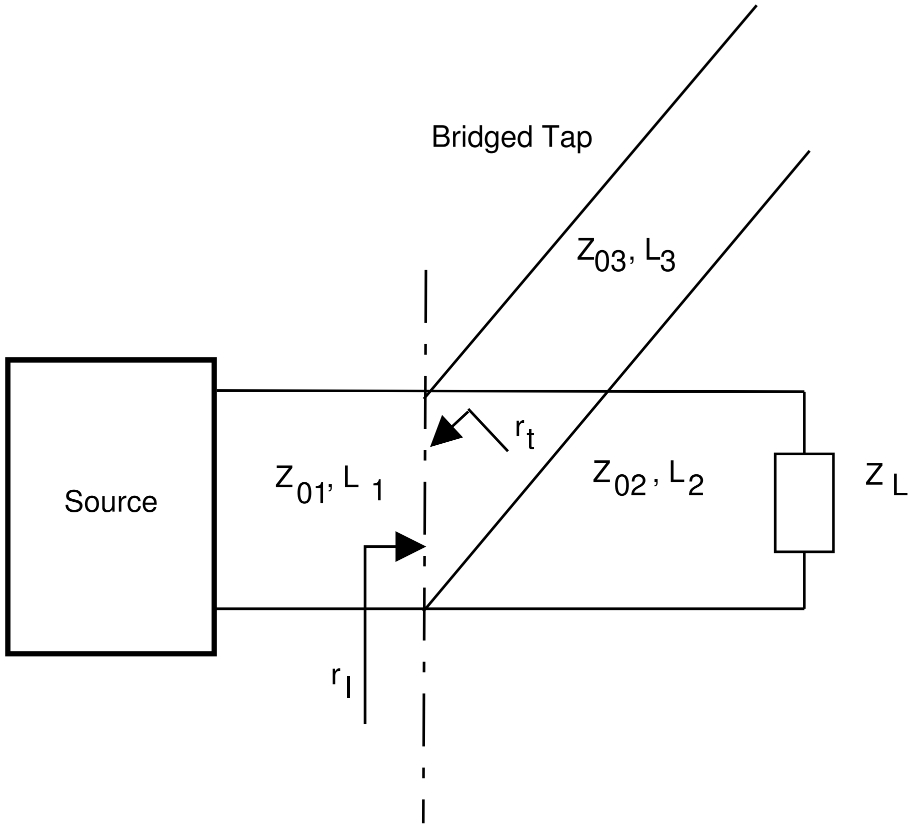


Figure 1.3: *Bridged Tap*

channel. The presence of Bridged taps and other impairments on the line, attenuates the direct channel gain when it is compared to the impairments-free model.

1.1.1.2 Crosstalk

Crosstalk is the interference created between different twisted pairs that belong to the same binder. This interference is caused by electromagnetic coupling between the different lines that lie in proximity of each others. Two types of crosstalk can be distinguished: Near end crosstalk (NEXT), and Far end crosstalk (FEXT).

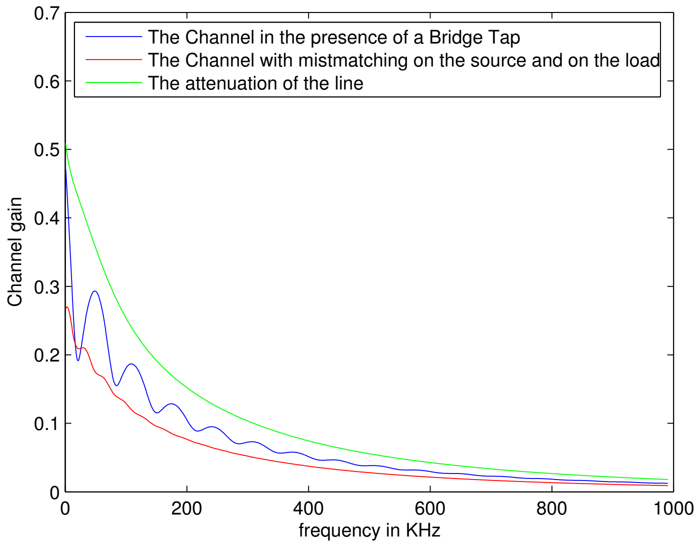


Figure 1.4: *Direct Channel Gain Models*

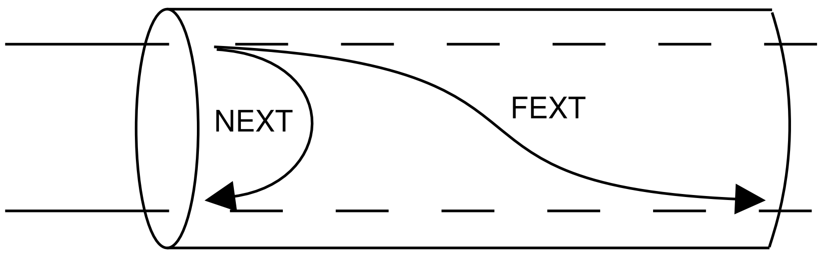


Figure 1.5: *Near and Far End Crosstalk*

1.1.1.2.1 NEXT Crosstalk Near end crosstalk is the interference that occurs between two twisted pairs when both the disturber line, and the victim line, exist on the same end of the binder. NEXT is caused by the transmitted signal on the transmitter side of the disturber line. These signals are usually strong signals as they are not attenuated by the channel. Thus NEXT effects can be extremely severe on the victim lines, especially on the received signals that got attenuated by the channel lines. Most of the DSL systems avoid NEXT by adopting an FDM strategy to separate the downstream signal (DS) from the upstream signal (US) by allocating different frequency bands to each of these two signals. NEXT channels may be expressed by the following model [7]:

$$H^2_{il,NEXT} = X_{49}(U/49)^{0.6} f^{3/2} \tag{1.6}$$

where $X_{49} = 1/1.13 \times 10^{13}$, U is the number of interference lines, and f is the frequency in Hz.

1.1.1.2.2 FEXT Crosstalk Far end crosstalk is the interference that occurs between two twisted pairs when both the disturber line, and the victim line, exist on two opposite ends of the binder. FEXT is less severe than NEXT, as the FEXT signals get attenuated by the direct line channel. FEXT channels can be modeled by the 1% worst case formula [7]:

$$H^2_{il,FEXT} = H^2_{ii}(U/49)^{0.6} d_{li} f^2 8 \times 10^{-20} \tag{1.7}$$

where d_{li} is the electromagnetic coupling distance between lines l and lines i.

In this thesis we consider DSL systems that adopt a frequency division duplexing (FDD) scheme to separate the upstream and the downstream bands, thus, in the rest of this work, we would only consider the case of the FEXT crosstalk.

1.1.2 DMT Modulation

The channel models of the twisted pairs lines used in the local loops are highly frequency selective. Due to the frequency selectivities of these channels, DSL systems suffer from inter-symbol interference (ISI). To combat the ISI, the majority of DSL systems have adopted the discrete multi tone (DMT) modulation. DMT modulation can be seen as decomposing the frequency selective direct channel into many flat fading sub-channels called tones [8]. At each tone, the DSL system may allocate different power, or even a different bitrate, depending on the ambient conditions on the tone such as the channel gain and the background noise. This procedure is called bitloading. Bitloading is implemented during the initialization period, however it can vary throughout the connection period. In fact, DMT based DSL systems, can keep tracking of the channel conditions (channel gain, and ambient noise) by implementing a bit swapping and a gain swapping techniques. In these techniques, the system can lower the number of bit allocation on one tone and increase it on another tone using bit swapping when the channel conditions change, same thing can be done by changing the allocated power on different tones using gain swapping.

Fig.1.8 and Fig.1.9 represent the DMT modulation block diagram. It is clear from these figures that the DMT modulation is very similar to OFDM used in wireless systems. DMT and OFDM differ in two main aspects: the first aspect is the variable bitloading adopted in DMT that allows the allocation of different power and bit at different tones, the second aspect is that the DMT modulation is a baseband modulation, thus it must use of the Mirror block in order to produce a real signal after the IFFT. The main DSL systems that have adopted the DMT modulation are: ADSL [9], ADSL2+ [10], and VDSL2 [11].

1.1.3 Spectrum Management

1.1.3.1 Near Far Problem

The attenuation of the direct channels in DSL systems is given in function of the length of the twisted pairs lines used in the local loops. The attenuation of the channel gain in longer line is more severe than that caused by shorter lines (Fig.1.1). The effects of cable lengths on the attenuation, coupled with crosstalk caused by strong signals from other lines may create a Near-Far problem in the DSL systems.

Signals that belong to long cables are subject to attenuation caused by the direct channel. These signals will cause a negligible crosstalk on their surroundings. Signals belonging to short lines suffer minimal attenuation, and they will cause important crosstalk on their neighboring lines. When long and short lines coexist on the same cable, the weak signals on the long line would suffer from the strong crosstalk caused by the strong signals transmitted on the short lines. Thus when a short line start transmitting, the long line would have to do a reinitialization and transmit data at a very low bitrate. The short line continue to transmit at nearly its maximal bitrate. Near-Far problem may occur in two cases: in the upstream and in the downstream.

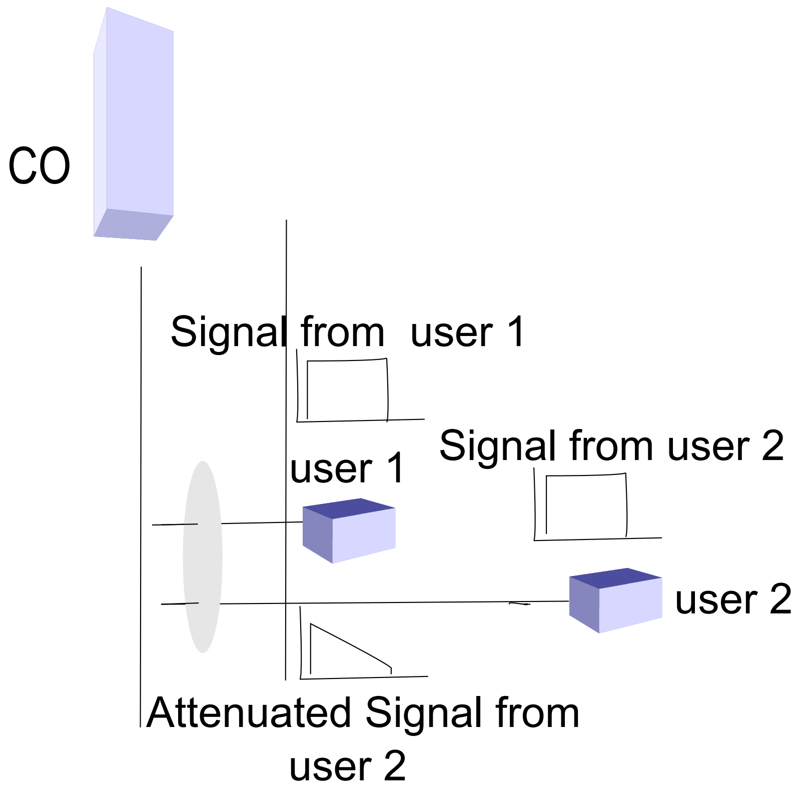


Figure 1.6: *Near-Far Problem in Upstream*

Near Far: Upstream Case In the upstream (US), lines that are served by the CO and share the same binder would suffer from a Near-Far problem if they have different lengths. Fig.1.6 gives a typical scenario of a Near-Far problem in the US. Signal transmitted by the modem of user 2 is attenuated by the time it reaches the same location of user 1. Thus the crosstalk caused by user 2 on user 1 is non significant. However, user 1 signal is not attenuated and it would cause a strong crosstalk on user 2.

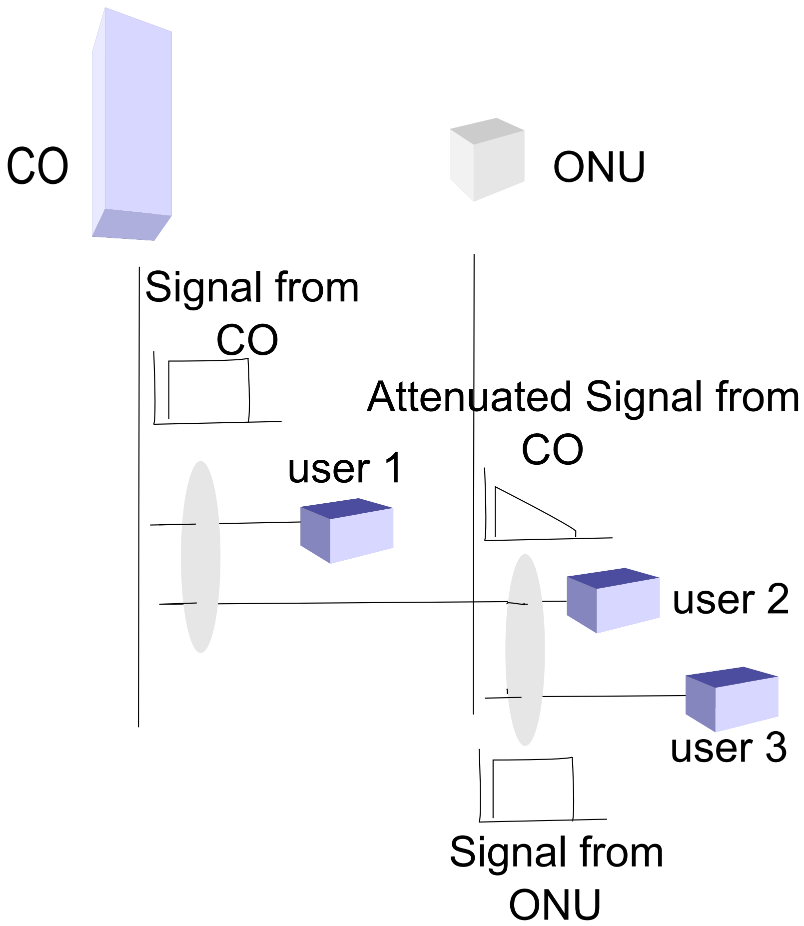


Figure 1.7: *Near-Far Problem in Downstream*

Near Far: Downstream Case In the downstream (DS) Near-Far problem can happen when lines served by the CO and by the ONU coexist in the same binder. In Fig.1.7 we show a DSL system that suffers from Near-Far in DS. In this system, the Near Far problem occurs between user 2 and user 3. As user 2 is served from the CO and user 3 is served from ONU. The same analysis used for the US can be applied on this case.

1.1.3.2 Spectrum Management

To limit the effect of crosstalk between different DSL systems and lines, telecoms companies and standardization bodies have imposed spectrum management techniques on the various DSL systems. These techniques include imposing fixed power spectrum density (PSD) mask and specific frequency and bandwidth allocation on the transmitted signal. Constraints on the total transmitted power may be imposed as well. Another way to insure the spectrum compatibility between the different lines and systems, is to measure the effect caused by a DSL line on other lines using the the 1% worst case formula.

Spectrum management insures the spectrum compatibility between signals belonging to different systems, and it reduces the impact of crosstalk between the different systems. However, it is far from being optimal, moreover techniques such as fixed PSD mask do not solve the Near-Far problem present in DSL systems. In the case of Near-Far situation, the power back-off methods are proposed as a possible solutions [12, 13, 14, 15, 16, 17, 18]. Power back-off methods (PBO) consist of lowering the PSD mask of the shorter loops in order to reduce their crosstalk on long loops.

Dynamic spectrum management (DSM) was proposed to further improve the power and frequency allocation in DSL systems [19, 7, 20]. DSM advocates the establishment of a spectral management center (SMC). The SMC uses the available information on each line such as the direct and the crosstalk channels, the ambient noise, the PSD levels in order to optimize the DSL lines performances in terms of stability and bitrate. 3 levels of DSM can be distinguished [19]:

1. DSM level 1 represents simple tasks that are done to improve the condition on a line based on the information provided by that line without acquiring information about other lines in the system. Adjusting the PSD of a line depending on its length (PBO methods) may be considered as a DSM level 1 operation.

2. DSM level 2: It represents a spectral coordination between the different DSL lines. DSM level 2 includes techniques for the optimization of discrete bitloading under individual power constraints [21, 22, 23, 24, 25, 26, 27, 28]. Maximization of DSL

system rate defined as the sum of aggregated continuous users rates is also considered [29, 30]. DSM level 2 techniques can reduce the effect of the crosstalk by optimizing the power allocation and the spectrum coordination among the different DSL lines that coexist in the same cable.

3. DSM level 3: Also called vectoring, it represents a coordination on the signals level between the different lines, in this case the DSL system is seen as a MIMO system. In this case and for each tone, the direct channels of the different lines, and the crosstalk channels are put together as a Matrix channel. Techniques such as Zero Forcing, and MMSE are used to equalize (to precode) the DSL system's channel in the US (in the DS).

1.2 Outline and Contributions

In this thesis we are interested in the application of DSM under practical and realistic conditions of the current DSL systems. In the first part of this dissertation we try to solve the Near-Far problem by incorporating the "Balanced Capacity" concept with the DSM level 2 techniques. Since it is only based on PSD coordination between the DSL lines, DSM level 2 seems to be the first candidate for implementation in DSL systems in the near future. DSM level 2 algorithms are complex and generally requires lot of computation times, thus our second objective in this thesis is the enhancement of state of the art DSM level 2 algorithms. This enhancement acts on two aspects: on one aspect we try to reduce the execution time required by the DSM level 2 optimization algorithms. This is done by adopting the successive optimization technique on one hand which insures a rapid convergence of the optimization, and by replacing the exhaustive and line search used in some state of the art algorithms, by less complex optimization algorithms such as Newton-Raphson and gradient algorithms on the other hand. The other aspect of enhancement is about the optimization outcome itself, in order to improve the optimization, and to reduce the chances of converging into a poor local optimum, we propose to use global optimization procedures such as multi start points, the optimal steepest ascent will be used as well to prevent the poor local optimum propagation in the case of successive optimization.

Most of the DSM level 2 algorithms suppose and need the total knowledge of the direct and of the crosstalk channel gain. While the direct channels are currently known by the CO, there are no information regarding the crosstalk channel gain. In the second part of this thesis we propose a technique for the estimation of the crosstalk channel gain without altering the modems and the standards of the current DSL systems. This estimation is based on the establishment of a monitoring system that observes the ambient conditions of each lines, information such as the SNR, the PSD of different signals, the time of connection and disconnection of the different users, and the background noise. These informations are supposed to be stored and analyzed. This passive observation of the DSL system would allow an effective estimation of the crosstalk channel gain when the SNR changes on a given line are correlated to the time of connection and disconnection of users transmitting on other lines in the systems. The monitoring system can be one and the same with the SMC.

For an effective implementation of DSM level 3 techniques, information on both the real part and the imaginary part of the crosstalk channels are required. In this case a passive observation of the DSL systems is not enough for the estimation of the crosstalk channel. We propose to make an active observation of the line by adding the effect of a virtual crosstalk channel on it. The virtual crosstalk will transmit the same signal as the interfering lines on the system. Thus the virtual crosstalk channel would change the gain of the actual crosstalk channel gain. This change can be detected by observing the rise or the drop of the SNR after inducing the virtual crosstalk. Changing the crosstalk channel gain twice is enough to estimate the real and imaginary part of the crosstalk channel using a triangulation techniques.

A time domain model for the crosstalk channel is also introduced, this model can improve the estimation of the crosstalk channel gain and of the real and imaginary parts of the crosstalk channels. This time domain model is usually limited in the number of taps, and it provides a relationship between the crosstalk channels of the different tones. This allows the estimation to be done over a limited number of pilot tones.

The rest of this thesis will be decomposed as follow:

Chapter 2
Chapter 2 is about the enhancement of the current state of the art algorithms, in this chapter we propose the "Balanced Capacity" concept as a possible solution to the Near-Far problem [31, 32]. We also use the correlation between the channel gain of the adjacent tones to propose the successive optimization concept, this concept allows the use of gradient type algorithms, such as steepest ascent, and Newton-Raphson [33, 34]. Multi start point technique is also suggested in order to improve the over all outcome of the optimization [35].

Chapter 3
Chapter 3 deals with estimation of crosstalk channel gain for DSM level 2 applications, it calls for the establishment of a monitoring system that observes the conditions on all the lines of the system. The estimation of the crosstalk channel gain is done by observing the SNR changes at the time of connection/disconnection of a user, or when a PBO is implemented. A time domain model can be used to improve the estimation and to compress the information of the various crosstalk channels.

Chapter 4
Chapter 4 considers the case of asynchronous crosstalk. In the first part of this chapter we derive an analytical model for the asynchronous crosstalk channel. In the second part we develop an estimation method for this type of channels. This method is largely based on the estimation technique used in chapter 3. A method for evaluating the non synchronization delay time is also provided.

Chapter 5
This chapter studies the effect of estimation error on DSM level 2. Taylor expansion was used to find a relationship between the optimal power allocation obtained when the different channels are fully known, and the optimal power allocation obtained under the presence of estimation error. A probabilistic tool is developed to calculate the expected loss in bitrate due to the estimation error.

Chapter 6
Chapter 6 considers the implementation of DSM level 3. DSM level 3 requires the full knowledge of the crosstalk channels. Thus the estimation of the Imaginary and Real part of the channel is necessary. The estimation is done by altering the actual crosstalk channel using an induced virtual crosstalk that transmit the same signal as the disturber line. Altering the real part first and then the imaginary part allow the estimation of the total crosstalk channel [36, 37]. The work presented in this chapter was patented by Alcatel-Lucent, and it eventually lead to many contributions in the standard and in the literature.

Chapter 7
In chapter 7 we optimize the crosstalk channel estimation by choosing the proper virtual crosstalk channel gain. A time domain model is also used. This model allows the estimation of several crosstalk channels at the same time.

Chapter 8
Chapter 8 provides an alternative estimation of the crosstalk channels. A least square time domain estimation is performed to estimate the time domain model of the crosstalk channel. Due to the sparsity of the time domain model, the estimation can be done over a limited number of pilot tones. Chapter 8 discusses the limitation of this method as it can result in an ill conditioned least square problem if the pilot tones are not chosen properly. Solutions for the ill conditioned LS problem are also provided [38].

Chapter 9
Chapter 9 presents our conclusions and provides few ideas for future work.

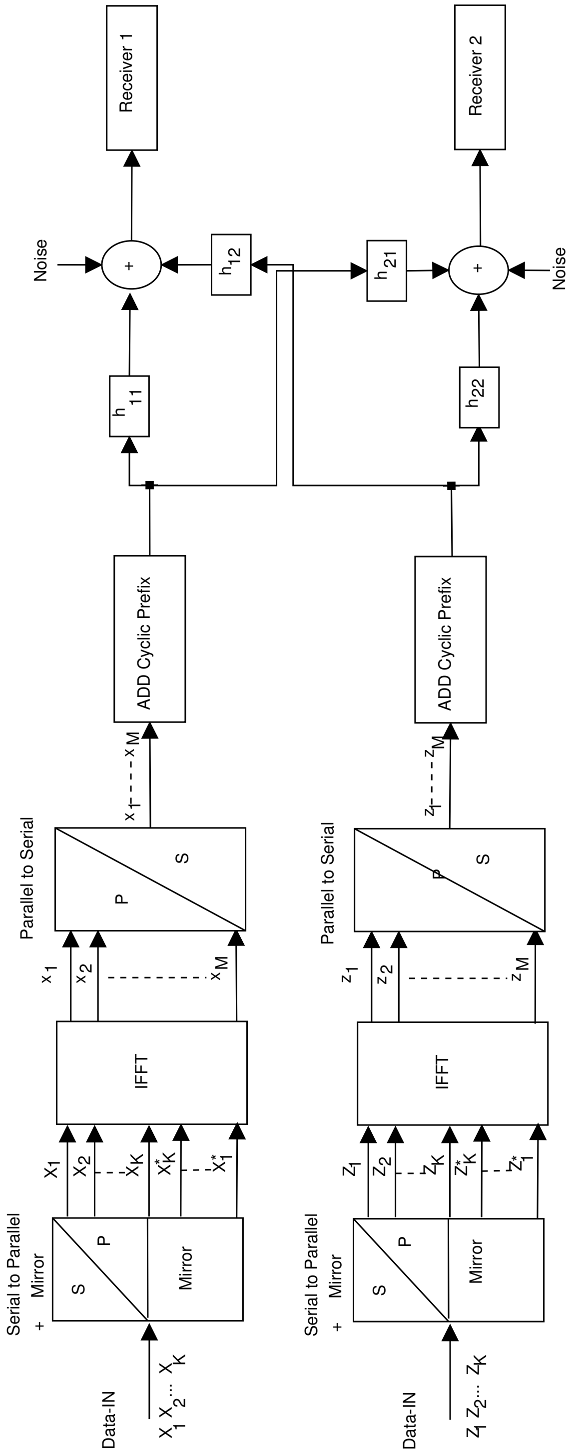


Figure 1.8: *DMT Transmitter*

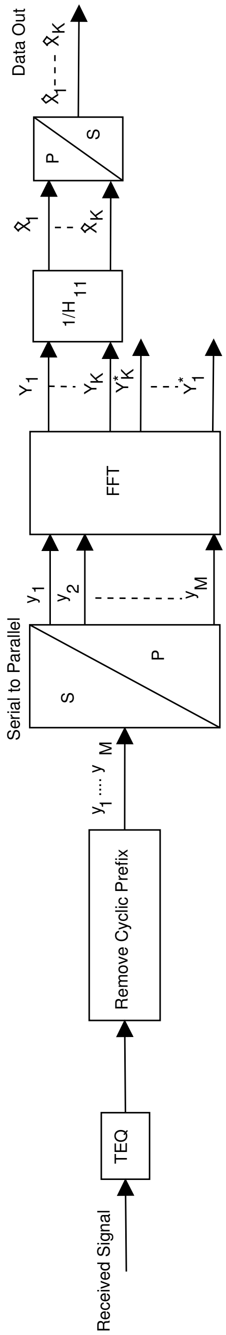


Figure 1.9: *DMT Receiver*

Chapter 2

Enhancement techniques for DSM algorithms

2.1 Introduction

One of the major limitations in DSL systems is crosstalk. Crosstalk is the electromagnetic coupling between twisted pairs that creates interference between the different lines in DSL systems. In VDSL systems it is the dominating noise factor. A user transmitting with an excessive power in a given binder can deteriorate all the lines of that binder. Operators and standardization bodies imposed limitations on the total power and on the PSD (power spectral density) mask of each system. PSD masks help to limit the level of crosstalk. However these PSD masks are based on the worst case scenario. Hence power allocations following these masks are far from optimal in practical situations.

To improve the power allocation in DSL systems, dynamic spectrum management (DSM) was introduced [7, 20, 39]. DSM tries to optimize users' bit rates by adapting their powers and spectral shapes to the channel loss and the ambient crosstalk across frequencies. DSM algorithms can be classified into two categories: distributed algorithms and centralized algorithms.

One of the first distributed algorithms that was proposed is Iterative Water Filling (IWF) [40]. In this algorithm each modem tries to optimize its own bit rate with respect to the ambient noise, including the current crosstalk. The optimization is done using the classical water filling algorithm. This procedure is repeated iteratively and independently on all modems until a constant level of powers and crosstalk is reached for all users. Autonomous Spectrum Balancing (ASB) [41] is another distributed algorithm where each modem tries to optimize the total sum of its own bit rate and the bit rate of a virtual line. The virtual line is supposed to represent a typical weak line, thus limiting the effect of each modem on its neighbors. Band preference algorithm (BP) [42] is a modified version of the IWF, BP utilizes power-scaling factors in the IWF algorithm to control the bit loading process, a large scaling factor given to a particular tone would correspond to a smaller number of bit being loaded to that particular tone, and vice versa (BP could be used to protect weak users).

Centralized algorithms propose the establishment of a system management center (SMC). In the SMC the total knowledge of the channel gains (both crosstalk channels and direct ones) is supposed to be given which allows the optimization of the entire system. Due to interference, the optimization of the DSL system total capacity is an NP hard non convex problem. OSB (optimal spectrum balancing) [43, 29] was proposed as the optimal centralized algorithm. OSB makes use of the Lagrange function to include the power constraints in the objective function. This enables the optimization to be decoupled over the different sub channels. In order to find the optimal solution at each sub channel, OSB resorts to an exhaustive search over all the different possible power allocations which makes it too complex for practical implementation. In order to reduce the complexity of OSB, the iterative spectrum balancing (ISB) algorithm was proposed [30, 44]. ISB replaces the exhaustive search over all the possible power allocations by a simple line search algorithm over individual users' powers, however this procedure is not globally optimum. Papers [45, 46] propose to solve the power allocation problem using an iterative convex approximation approach, in this approach the objective function is approximated by a convex function which allows the use of convex optimization.

In crosstalk limited DSL systems, a near far problem may occur when long lines coexist with short lines in the same binder. In this case, if the optimization problem is not formulated properly, the DSM algorithm may end up allocating small or no power to users with longer lines, while users with shorter line will be using almost the totality of their maximum allocated power. This means that some users would be sending at a rate that is nearly equal to their maximum bitrate, while other users would send at a very small rates. In this chapter, and in order to solve the near far problem in DSL systems, we propose to apply the concept of Balanced Capacity (BC) [47] to ensure fairness among different active users.

Another objective of this chapter is to propose practical techniques that can enhance the performance of existent state of the art DSM algorithms. For instance, the successive optimization (SO) principle is proposed, this principle may be used to speed up the convergence of several known DSM algorithms (ISB, SCALE), and it also allows the use of gradient type algorithms on DSL optimization which lowers the complexity of the DSM procedure. The use of multi start point technique is proposed to increase the overall outcome of the optimization.

2.2 System Model

We consider DSL systems using DMT (discrete multiple tone). Assuming a proper use of the cyclic prefix technique, the channel may be decomposed in K parallel subchannels. Let the total number of users on the system be U. For each particular tone the system can be viewed as an interference channel. Viewing the interference from other users due to crosstalk as Gaussian noise the bitrate of user i is given by the formula:

$$R_i = B \sum_{n=1}^{N} \log_2(1 + \frac{1}{\Gamma} \mathrm{SNR}_i(n)), \tag{2.1}$$

where Γ is the SNR gap, $\mathrm{SNR}_i(k)$ is given by

$$\mathrm{SNR}_i(k) = \frac{|H_{ii}(k)|^2 \, P_i(k)}{\sigma_i^2(k) + \sum_{l \neq i} |H_{il}(k)|^2 P_l(k)}. \tag{2.2}$$

Where $H_{ii}(k)$ is the direct channel gain of user i at tone k. $H_{il}(k)$ is the crosstalk gain from line l to line i at tone k, and $P_i(k)$ is the power transmitted by line i. The background noise variance of user i at tone k is denoted by $\sigma_i^2(k)$. It is assumed that a PSD mask and a total power limitation are imposed to each user, so the problem may be stated as:

$$\max_{\mathbf{P}} \quad \sum_{i=1}^{K} \omega_i R_i \tag{2.3}$$

$$\text{Subject to} \quad \begin{cases} \sum_{k=1}^{K} P_i(k) \leq P_t \\ P_i(k) \in [0, P_{max}(k)] \end{cases}$$

where $\mathbf{P}$ is a $K \times U$ matrix where each row correspond to the the power allocation at a tone k: $\mathbf{P}(k) = [P_1(k), P_2(k)...P_U(k)]$. The weighting coefficients ω_i are fixed parameters in the problem and supposed to be defined by external considerations on user's priorities.

2.3 Balanced Capacity

Due to the FEXT crosstalk, the near far problem may occur in several situation in DSL systems. In DownStream (DS), the near-far problem will occur only if in the same binder there is a mix between lines served by the central office (CO) and lines served by an optical network unit (ONU) located closer to the users. Typically, the signal transmitted on the ONU line will be strong in the DS case, while the signal of the CO lines will be attenuated due to the long distance from the CO. The strong signal of the ONU line will cause an important crosstalk on the CO lines thus limiting their bitrates, while the attenuated signals of the CO lines would cause a marginal crosstalk on the ONU line that continues to transmit at nearly its maximal bitrate. In the UpStream (US), the near far problem is caused by the coexistence of long lines and short lines in the same binder. Even if all the lines are served by the CO, in the US case the short lines would have the same effects that the ONU lines have in the DS.

To prevent the near far problem where some users transmit at nearly their maximal rates, while others transmit at a zero or a very low rates we propose to apply the "Balanced Capacity" concept on the

DSL systems. Balanced Capacity (BC) is defined as a situation where $R_i/R_{i,max}$ is a constant for all users i, where R_i is the bitrate of user i in the presence of crosstalk, and $R_{i,max}$ is the single user bitrate, i.e. it is the maximum achievable bitrate by user i, this happens when line i is free from interference. The concept of BC corresponds to a specific point of the capacity region boundary where the coexistence with the other users has the same relative cost for every user. Any point on the border of the capacity region can be achieved by maximizing $\sum_i^U \omega_i R_i$ with the appropriate set of ω_i, for $\sum_i^K \omega_i = 1$ and all ω_i being positive. In order to find the BC rates, one should find the appropriate values of ω_i that correspond to the BC point. The optimization problem now can be formulated as :

$$\begin{array}{ll} \max_{\mathbf{P}} & \sum_{i=1}^{K} \omega_i R_i \\ \text{Subject to} & \left\{ \begin{array}{l} R_1 = \gamma_i R_i \\ \sum_{k=1}^{K} P_i(k) \leq P_t \\ P_i(k) \in [0, P_{max}(k)] \end{array} \right. \end{array} \tag{2.4}$$

where for each user i we have $\gamma_i = R_{1,max}/R_{i,max}$. User 1 has been arbitrarily selected as reference user. We can see that the BC is included in the optimization problem by the constraints $R_1 = \gamma_i R_i$.

Solving problem (2.4) can be implemented by means of two loops as presented in Fig.2.1. The first loop has as target to search for the weighting factors corresponding to the BC point. The second loop aims at solving problem (2.3), using traditional or improved DSM algorithms, for fixed ω_i. The search for ω_i that correspond to BC can be done using the following procedure. Let $d_{bc,i} = R_1 - \gamma_i R_i$. At iteration t of the outer loop, the estimates of ω_i, denoted by $\omega_i^{(t)}$, are updated according to

$$\omega_i^{(t+1)} = \omega_i^{(t)} + \alpha_{bc}\, d_{bc,i}, \tag{2.5}$$

where α_{bc} is a small positive correction step. So if $\gamma_i R_i$ exceeds R_1, the weighting factor ω_i associated with R_i will decrease, and the optimization at iteration $t+1$ will result with a smaller R_i than that of

BC Finding Algorithm

1. *init* $\gamma_i = R_{1,max}/R_{i,max}$ *for* $i = 2 : U$
2. *init* $\omega_i^{(t=0)} = \frac{1}{U}$
3. *repeat until all* $|d_{bc,i}| < \epsilon$
   - *(inner loop):*
     - *Solve (2.3) for fixed* ω_i
   - $d_{bc,i} = R_1 - \gamma_i R_i$
   - $\omega_i^{(t+1)} = [\omega_i^{(t)} + \alpha_{bc}\, d_{bc,i}]^+$ *for* $i \neq 1$
   - $\omega_1^{(t+1)} = [1 - \sum_2^U \omega_i^{(t+1)}]^+$
   - $\omega_i^{(t+1)} = \omega_i^{(t+1)} / \sum_i \omega_i^{(t+1)}$ *for all* i

Figure 2.1: *BC algorithm with two loops*

iteration t. The opposite would happen R_1 exceeds $\gamma_i R_i$. As ω_i must be positive, negative ω_i are forced to zero. In all cases ω_1 is forced to $[1 - \sum_2^U \omega_i]^+$.

2.4 Optimization Algorithms Review

In this section we will review several optimization algorithms, some of these algorithms were proposed for DSM (OSB, ISB), others are general optimization algorithms (Newton-Raphson, gradient) that will be adopted for DSM application.

2.4.1 Review of OSB and ISB

Due to the presence of crosstalk, the objective function in (2.3) can be seen as a difference of two log functions which yields an NP hard non convex problem. Paper [43, 44] try to solve this problem globally by introducing the dual function to relax the power constraints. The

Lagrange function of problem (2.3) is given by:

$$g(\lambda, \mathbf{P}) = \sum_{i}^{U} \omega_i R_i - \lambda_i (\sum_{k=1}^{K} P_i(k) - P_t) \tag{2.6}$$

where $\lambda = [\lambda_1, ...\lambda_i, ...\lambda_U]$ and each λ_i represents a Lagrangian multiplier. For fixed λ, the function f is defined as:

$$f(\lambda) = \ \max_{\mathbf{P}} \quad g(\mathbf{P}). \tag{2.7}$$

The dual problem becomes:

$$\begin{array}{cl} \min_{\lambda} & f(\lambda) \\ \text{Subject to} & \lambda_i \geq 0. \end{array} \tag{2.8}$$

To solve problem (2.8) both OSB and ISB propose a double loop iterative procedure. An outer loop searches for the appropriate λ that minimizes $f(\lambda)$ to meet the power constraints. And for each set of fixed λ_i, an inner loop maximizes $g(\mathbf{P})$ with respect to $\mathbf{P}$. A simple sub-gradient algorithm was proposed in [30, 44] for finding λ in the outer loop. The search for λ was improved for better convergence in [48]. Examining $g(\mathbf{P})$ shows that for fixed λ there is no coupling between the tones k as $\sum_i^U \lambda_i P_t$ becomes a constant which no longer affects the optimization. Hence the optimizations can be carried out per tone. This makes the complexity linear in function of K. For each tone, ie for the inner loop optimization, the non-convexity property holds. So, to solve the tone wise optimization, OSB has been proposed where an exhaustive search over all the possible power allocations in a tone k is implemented. This renders the complexity exponential with U.

To reduce the complexity, ISB has been put forward. An exhaustive "line search" is performed over the power of individual users instead of a total exhaustive search. For each tone k, the power of $(U - 1)$ users is fixed and an exhaustive search is performed over the power $P_i(k)$ of the remaining users to maximize $g(\mathbf{P})$. This procedure is repeated iteratively over all users till a constant power allocation is reached.

2.4.2 Successive Optimization

In DSL systems, there is a strong correlation between adjacent sub-channels. This correlation holds for both crosstalk and direct channels where both empirical and practical channel models show that adjacent tones have a very similar channel gain. Due to this fact we can conclude that the optimization problem at a tone $k+1$ is very close to optimization at tone k. Thus the optimum value $\mathbf{P}(k) = [P_1(k), P_2(k)...P_U(k)]$ that represents the power allocation found at tone k (with $P_i(k)$ is the power allocated to user i) is very close to $\mathbf{P}(k+1)$ the optimal solution at $k+1$.

In this chapter, we propose the use of successive optimization (SO) as a method of enhancement of DSL algorithms. To use the SO: For each tone $k{+}1$ we start the search over the power with with the result found at the previous tone $\mathbf{P}(k)$. This simple procedure is able to speed up the convergence of typical DSM algorithms as it reduces significantly the number of iterations needed for the optimization. SO can also be used to propose gradient algorithm types for DSM applications.

2.4.3 Gradient Algorithm

In gradient algorithms, and in order to maximize the objective function around the point $\mathbf{P}(k)$, one should take steps proportional and of the same direction as the gradient $\nabla g[\mathbf{P}(k)]$:

$$\mathbf{P}^{(t)}(k) = \mathbf{P}^{(t-1)}(k) + \alpha \nabla g[\mathbf{P}^{(t-1)}(k)] \tag{2.9}$$

Where i is the iteration index, and α is a positive small step size. In this case ($\alpha > 0$), the gradient is called gradient ascent or steepest ascent. The steepest ascent procedure guaranties the convergence toward the nearest local optimum that lies in the proximity of the initial point $\mathbf{P}^0(k)$. In the case of convex optimization, the steepest ascent (SA) algorithm will almost certainly find the global maximum, however if the optimization is a non convex one, SA may converge into a poor local optimum. However, when coupled with SO, the gradient algorithm can be used for the optimization of DSL systems, even if the optimization is suffering from multiple local optima.

As it was previously shown, the optimum power allocations at two adjacent tones ($\mathbf{P}(k)$ and $\mathbf{P}(k+1)$) are at vicinity of each others. The vector $\mathbf{P}(k+1)$ is considered to be at least a local maximizer of the objective function $g(\mathbf{P})$ at the tone $k+1$. This means that $g(\mathbf{P})$ will be typically strictly concave in the neighborhood of $\mathbf{P}(k+1)$. Since $\mathbf{P}(k)$ lies in this neighborhood, initializing the optimization with $\mathbf{P}(k)$ put us in good situation to use gradient type algorithms.

2.4.4 Newton Raphson

Newton Raphson method approximate the objective function around $\mathbf{P}(k)$ by a quadratic function. So according to this method, the correction applied to vector $\mathbf{P}(k)$ at iteration t is given by

$$\mathbf{P}^{(t)}(k) = \mathbf{P}^{(t-1)}(k) - \left(\nabla^2 g[\mathbf{P}^{(t-1)}(k)]\right)^{-1} \nabla g[\mathbf{P}^{(t-1)}(k)] \tag{2.10}$$

where $\nabla^2 g[\mathbf{P}^{(t-1)}(k)]$ is the Hessian of the Lagrange function g, evaluated for $\mathbf{P}^{(t-1)}(k)$ and $\nabla g[\mathbf{P}^{(t-1)}(k)]$ is the gradient at the same value. NR method leads for a maximum when the Hessian matrix is negative definite. This means that NR works only in a strictly convex region, thus it is essential to couple NR method with SO. Even when SO is implemented with NR method, it may happen that the initial guess proposed is not close enough to the optimum value. This can be detected when the Hessian is non-negative definite. In such a case the NR method can no longer be used. Therefore we resort to a gradient method for a few iterations, until the Hessian becomes negative definite.

NR Complexity The complexity of the Newton-Raphson method is mainly given by the inversion of the Hessian matrix. The complexity of a matrix inversion using Gaussian elimination is in the order of $O(U^3)$ where U is the total number of users. The complexity of the line search method is in the order of $O(U^2 G_p)$, where for each user, at each possible power allocation for this user, we have to calculate the capacity for all the users (U times). G_p corresponds to the grid search, it is equal to the total number of possible power allocations for each user. In DSL systems the number of interfering users is usually small (typically under 10), so for small step size in the grid search (meaning good precision), the proposed method will exhibit a large

gain in complexity. For example for $K = 10$, and for $G_p = 10^2$ the NR complexity is $O(10^3)$ while that of LS is $O(10^4)$.
The Hessian inversion required for the NR method can be seen as draw back for this procedure. Thats why we propose next a quasi optimal steepest ascent algorithm that requires no matrix inversion.

2.5 Optimal Steepest Ascent Algorithm

In this section, we propose to implement an optimal steepest ascent algorithm for DSM optimization. Unlike the normal gradient algorithm that remains in the region of the nearest local optimal, as it simply make small correction steps in the direction of the gradient. Optimal SA actually finds the optimal value in the direction of gradient at each iteration. This property gives the optimal SA the ability to leave the region of a local optimal into the region of a better local or even global optimal.
For optimal steepest ascent the correction step α corresponds to α_m which is given by:

$$\alpha_m = max_\alpha \;\; g\left(\mathbf{P}(k) + \alpha \nabla g[\mathbf{P}(k)]\right) \tag{2.11}$$

Thus at each iteration, the optimal SA searches for the step size α_m that maximizes the objective function in the direction of the gradient. As in (2.3), problem (2.11) is a non convex problem. One way to find α_m is to find all the critical points and then testing them for optimality. To find all the critical points, one should find all the points α that satisfy the equation:

$$\frac{\partial}{\partial \alpha} g\left(\mathbf{P}(n) + \alpha \nabla g[\mathbf{P}(n)]\right) = 0 \tag{2.12}$$

Defining the elements of $\nabla g[\mathbf{P}(k)]$ by:

$$G_i(k) = \frac{\partial}{\partial P_i} g[\mathbf{P}(k)]$$

Equation (2.12) can be rewritten as:

$$\sum_i^U \frac{\omega_i'(k)}{(\eta_i(k) + \alpha)(\mu_i(k) + \alpha)} + \sum_i^U \lambda_i G_i(k) = 0 \tag{2.13}$$

with the terms:

$$\omega_i'(k) = \omega_i(k)\frac{b_i(k)c_i(k) - a_i(k)d_i(k)}{d_i(k)(c_i(k) + d_i(k))} \quad (2.14)$$

$$\eta_i(k) = \frac{a_i(k) + b_i(k)}{c_i(k) + d_i(k)} \quad (2.15)$$

$$\mu_i(k) = \frac{b_i(k)}{d_i(k)} \quad (2.16)$$

$$a_i(k) = |H|_{ii}^2(k)P_i(k) \quad (2.17)$$

$$b_i(k) = \sum_{l\neq i}^{U} |H|_{il}^2(k)P_l(k) + \sigma_i(k) \quad (2.18)$$

$$c_i(k) = |H|_{ii}^2(k)G_i(k) \quad (2.19)$$

$$d_i(k) = \sum_{l\neq i}^{U} |H|_{il}^2(k)G_l(k) \quad (2.20)$$

Fig.2.2 shows a typical plot of the LHS of equation (2.13) in function of α. The different asymptotes shown in Fig.2.2 corresponds to $\alpha = -\eta_i(k)$ or $\alpha = -\mu_i(k)$. Since α must be strictly positive, we are only concerned about negative η_i and μ_i.

Let A_1=min(-$\mu_i(k)$,-$\eta_i(k)$) for all negative $\mu_i(k)$ and $\eta_i(k)$. Since the channel gains are strictly positive, it can be easily shown that for $\alpha \geq A_1$ at least one element of the vector $\mathbf{P}^{(t)}(k) = \mathbf{P}^{(t-1)}(k) + \alpha\nabla g[\mathbf{P}^{(t-1)}(k)]$ must be negative. In practice these negative power components correspond to the lines being not active at the given tones. Thus if the optimum value requires transmitting powers on all lines, α_m should be typically between 0 and A_1. To find α_m between 0 and A_1 a hyperbola was used to approximate (2.13) between these two points.

The model $y = \frac{al}{x+A_1} + C_{st}$ is found to give good results. Parameters al and C_{st} can be calculated using two values of equation(2.13) (In the simulation equation (2.13) is evaluated arround 0 and $A_1/2$). The approximate value of α_m will be given by $\alpha_m' = -al/C_{st} - A_1$. Another hyperbolic approximation near α_m' may be done to improve the solution.

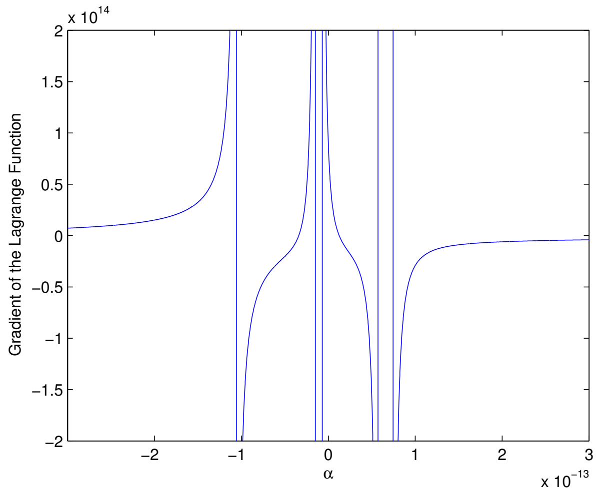


Figure 2.2: $\frac{\partial}{\partial \alpha} g\left(\mathbf{P}, \alpha\right)$ *in function of* α

An optimum value may exist for α_m larger than A_1. In this case the negative resultant power must be set to null which results in a new optimization with a lesser number of active lines. To check for optimality for each positive asymptotes replace the negative power with zeros and get the optimal value for the remaining active lines.

2.6 Enhanced ISB

In this section we will present several techniques to enhance the performance of ISB, some of these techniques may be applied to a variety of DSM optimization algorithms.

2.6.1 Complexity Reduction

In this section we describe an algorithm that maximizes the objective function $g(\lambda, \mathbf{P})$ at tone k with respect to $P_u(k)$. The power allocation at the other $U - 1$ users is considered constant. This algorithm tries to find directly the optimum by searching the critical points at which the gradient equal to zero. This algorithm replaces the exhaustive line search proposed in ISB for the inner loop. It reduces the complexity considerably.

2.6.1.1 Partial Derivative Approximation

A simple way to perform the optimization of $g(\lambda, \mathbf{P})$ at tone k with respect to $P_u(k)$ is to find all the roots of the derivatives.

$$\frac{\partial g(\lambda, \mathbf{P})}{\partial P_u(k)} = 0. \tag{2.21}$$

The partial derivative of $g(\lambda, \mathbf{P})$ with respect to $P_u(k)$ is given by

$$\begin{aligned} \frac{\partial g(\lambda, \mathbf{P})}{\partial P_u(k)} &= \frac{\omega_u / \ln(2)}{P_u(k) + C_u} \\ &+ \sum_{l \neq u} \omega_l / \ln(2) \Big(\frac{1}{P_u(k) + C_{l1}} - \frac{1}{P_u(k) + C_{l2}} \Big) \\ &- \lambda_u \end{aligned} \tag{2.22}$$

where C_u, C_{l1}, C_{l2} are defined as follows:

$$\begin{aligned} C_u &= \frac{\sum_{i \neq u} |H_{ui}(k)|^2 P_i(k) + \sigma^2}{|H_{uu}(k)|^2} \\ C_{l1} &= \frac{\sum_{i \neq u} |H_{li}(k)|^2 P_i(k) + \sigma^2}{|H_{lu}(k)|^2} \\ C_{l2} &= \frac{\sum_{i \neq u,l} |H_{li}(k)|^2 P_i(k) + \sigma^2}{|H_{lu}(k)|^2}. \end{aligned} \tag{2.23}$$

We notice that $\frac{\partial g(\lambda, \mathbf{P})}{\partial P_u(k)}$ is the sum of different hyperbolas of the form $y = \frac{a}{P_u(k)+C}$, where $-C_u$, $-C_{l1}$, $-C_{l2}$ represent the different asymptotes of these hyperbolas. These asymptotes are all negative. Since the direct channel gain $|H_{uu}(k)|^2$ is much bigger than the crosstalk channel gains $|H_{lu}(k)|^2$ we can conclude that C_u is much smaller than the other asymptotes. So for $0 < P_u(k) < P_{max}(k)$ and when P_{max} is low, the partial derivative $\frac{\partial g(\lambda, \mathbf{P})}{\partial P_u(k)}$ is mainly influenced by the hyperbola corresponding to C_u. Thus we may assume that it is monotonically decreasing over the search interval and $\frac{\partial g(\lambda,\mathbf{P})}{\partial P_u(k)}$ may be approximated by the hyperbolic model

$$\frac{\partial g(\lambda,\mathbf{P})}{\partial P_u(k)} \cong \frac{a}{P_u(k)+C} + C_t. \tag{2.24}$$

Fig. 2.3 illustrates the hyperbolic modeling of the partial derivative, where the solid blue line represents the real values of a partial derivative, while the red doted line represents a hyperbolic model around zero.
If we detect that the partial derivative is not monotonically decreasing we may use the model

$$\frac{\partial g(\lambda,\mathbf{P})}{\partial P_u(k)} \cong \frac{a}{P_u(k)+C} - \frac{d}{P_u(k)+D} + C_t \tag{2.25}$$

where a, C, d,D, C_t are constant parameters.

2.6.1.2 Root Finding Algorithm

This paragraph describes an algorithm that is able to find the root of (2.21) using only 4 realizations of the partial derivative. Thus the complexity of the ISB is reduced by a factor of $G/4$, where G is the total number of grids used in the exhaustive line search algorithm over the interval $[0, P_{max}(k)]$.

1. First compute $\frac{\partial g(\lambda,\mathbf{P})}{\partial P_u(k)}$ at 4 equidistant points p_1, p_2, p_3, p_4 in the interval $[0, 0.5P_{max}(k)]$ using the exact formula 2.22. The values of $\frac{\partial g(\lambda,\mathbf{P})}{\partial P_u(k)}$ at these points are denoted by A_1, A_2, A_3, A_4.

2. If the 4 realizations of the partial derivative prove to be monotonic, approximate $\frac{\partial g(\lambda,\mathbf{P})}{\partial P_u(k)}$ by model (2.24).

   - If A_1 and A_3 are both positive use the points p_1, p_3, p_4 to find a, C, and C_t.
   - If A_1 and A_3 are both negative or of different signs, use the points p_1, p_2, p_3 to find a, C,and C_t.

   - Find points $P_u^c(k)$ for which the model is equal to zero.

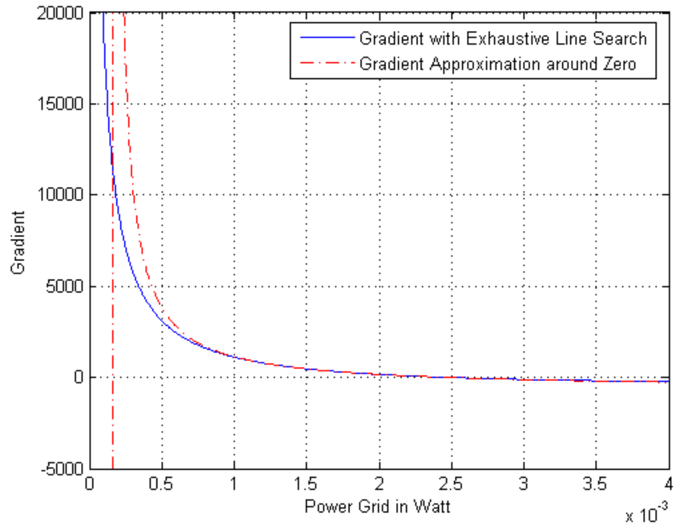


Figure 2.3: *Gradient given by exhaustive search vs its approximation around zero*

3. If the 4 realizations are not monotonic approximate $\frac{\partial g(\lambda,\mathbf{P})}{\partial P_u(k)}$ by model (2.25).
   - Use the points p_1, p_2, p_3, p_4 to find a, C, d, and C_t.
   - Find the roots $P_u^c(k)$ for which the model is equal to zero.
   - For model (2.25) 2 roots are found. Choose the root that maximizes the objective function.
4. If a is negative, the maximum is either 0 or $P_{max}(k)$.
5. If a is positive, the maximum is $P_u(k) = P_u^c(k)$
6. If a is positive, to improve the solution we may revisit the modeling around $P_u(k) = P_u^c(k)$.
   - Compute $\frac{\partial g(\lambda,\mathbf{P})}{\partial P_u(k)} = A_5$ at $P_u(k) = P_u^c(k)$.
   - If $P_u^c(k) > p_4$, $p_1 = P_u^c(k)$ and $A_1 = A_5$

- If $P_u^c(k) < p_4$, $p_4 = P_u^c(k)$ and $A_4 = A_5$
- Go to 2.

2.6.1.3 Complexity

The main advantage of this method with respect to existing ISB algorithm is its low complexity. The complexity of the exhaustive line search method used in conventional ISB is in the order of $O(G_pU^2)$. The root finding algorithm described above requires for each user, the computation of the gradient (U times operation), for at least 4 different power allocations, resulting in a complexity of order $O(4U^2)$. Thus the gain in complexity is of the order $G_p/4$.

2.6.2 Optimization Improvement

In the previous section we presented an algorithm to reduce the overall complexity of ISB. However, the described algorithm does not enhance the outcome of the optimization itself. This section proposes two techniques to improve the optimization: The additional starting point (ASP) and the successive optimization.

2.6.2.1 ASP

The ISB algorithm is only guaranteed to reach a local optimum. The objective function (2.6) is known to have many local optima. Due to its iterative structure (iteration across the users), the result of the algorithm depends on the initial point and there is a high chance that the ISB might fall on a local optimal and not on a global one. One technique to overcome this difficulty is to use a multi start point technique. In this technique the optimization algorithm is run many times. Every time, a different initial guess is used, thus increasing the chance to find the global optimum or at least to reach a solution closer to this global optimum.

As described in section 2.4.1, the ISB optimization is decoupled over tones. So instead of repeating the entire algorithm several times with different initial points, we propose to start with a different initial guess at each tone. Thus we can apply a multi start points procedure within a single realization of the algorithm. The use of different initial points at the different tones results in a perturbed power allocation over the tones. However combined techniques of successive optimization and multi start points can provide a smoothening effect and improve the optimization furthermore.

2.6.2.2 Double Successive Optimization

The reason behind the time sharing property described in [44] is the strong correlation between adjacent sub-channels. This correlation holds for both crosstalk and direct channels where both empirical and practical channel models show that adjacent tones have a very similar channel gain. In this chapter, we already used this observation to propose gradient type algorithms based on successive optimization: for each tone $k+1$ we start an optimization based on Newton-Raphson or on steepest ascent with the result found at the previous tone $\mathbf{P}(k)$. This simple procedure was proposed to speed up the convergence of existing algorithms as it reduces significantly the number of iterations needed for the inner loop. The same operation may be used with line search (ISB) instead of Newton-Raphson.

The drawback of successive optimization is that it may fall in the region of a local optimum for tone k. And due to the successive iteration it may get stuck in this region for several tones before reaching a better region. In this section we solve this problem by proposing a double successive optimization:

Forward Successive Optimization With these considerations in mind, successive optimization is used to smoothen the result of the ASP technique. At each tone $k+1$ the optimization is done twice. The first optimization is started with an initial guess $\mathbf{P}(k)$ (successive optimization). The second optimization is started with an initial guess $\mathbf{P}(k)+\mathbf{E}(k+1)$ (which explains the ASP name). $\mathbf{E}(k)$ is a random vector of the same dimension as $\mathbf{P}(k)$ and which varies with k. The objective of vector $\mathbf{E}(k)$ is to modify the initial guess $\mathbf{P}(k)$ so that the optimization start from another region. This avoids the optimization from falling into the regions of a poor local optimum over large period of tones. At the end the outcome of the two optimizations is compared and the best one is kept.

Reverse Successive Optimization The optimization proposed earlier is a forward successive optimization, where we start at tone 1 and we end at tone K. One way to improve the result further is to perform a reverse or backward successive optimization. This one is performed after the forward successive optimization is finished. The reverse one starts at tone $K-1$ and moves backward to tone 1. At each tone k the optimization is performed with an initial guess $\mathbf{P}(k+1)$. The outcome is compared to the previous result (obtained by the combined "ASP"/"Forward Successive Optimization" and the best power allocation is kept.

Final Procedure Overall, the ISB algorithm is implemented 3 times for each tone. One in the forward successive mode, initialized with $\mathbf{P}(k-1)$; one in the ASP technique initialized with $\mathbf{P}(k-1)+\mathbf{E}(k)$; and one in the reverse successive mode, initialized with $\mathbf{P}(k+1)$. The best solution is kept in the end.

2.7 Numerical Results

In this section we report the numerical results obtained for both BC algorithm and for the enhancement of DSM algorithms. The direct channel gain is considered to be the attenuation loss caused by the twisted pairs, while the crosstalk channels are modelled by the 1% worst case formula [7], the spacing between the tones is 4.3 kHz. We start by testing the BC algorithm using ISB and NR for the DSM optimization, then we compare the optimal gradient algorithm with ISB, and NR, for the sum rate criterion, and finally we report the results of the enhanced ISB algorithm.

2.7.1 Balanced Capacity

We use the BC algorithm in the case of FEXT crosstalk, with 3 users having different distances from the CO or ONU:

Users	user 1	user 2	user 3
Distances	7 Km	500 m	3 Km

The simulation is done for 128 tones, with tone zero corresponding to 258.75 Khz.

When applying the BC algorithm with a line search in the inner loop as suggested in [30], each search is performed on the interval [-50,-10] dBm/Hz with 0.5 dBm/Hz step. The algorithm usually converges after about 5 iterations to a 3% accuracy as shown by Fig.2.4

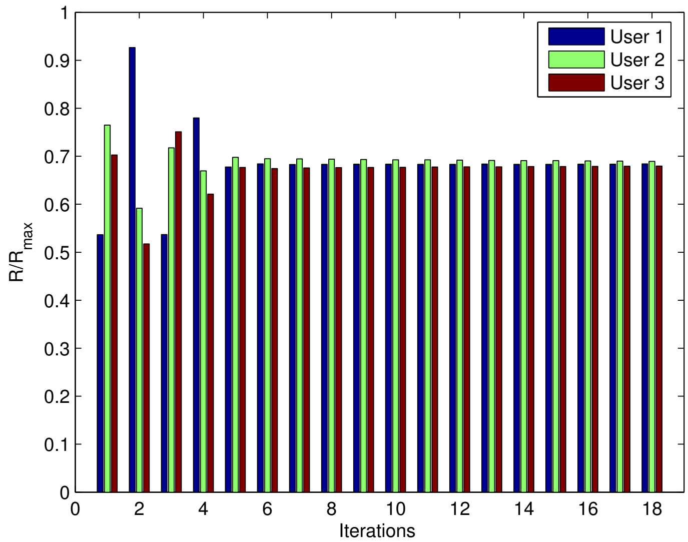


Figure 2.4: *BC algorithm using line search*

Now we run the BC algorithm using the NR as the inner loop. The step used for quantizing the possible power values is $10^{-4}P_{max}$ (which correspond to -50 dBm/Hz, i.e. the minimal allocation in the LS case). In this case, the quantization is done in linear scale, with a fixed step. The accuracy is thus higher than with the logarithmic step taken in the LS case. The evolution of the algorithm along the iterations is given in Fig. 2.5. The algorithm converges after 5 iterations to 3% accuracy.

The convergence of the outer loop (BC point search) is highly dependent on the choice of the correction step μ. If μ is small the convergence is slow, and the algorithm takes a lot of iterations before converging to an acceptable accuracy. Inversely when μ is large, the convergence is fast but the accuracy is lower. The solution used here is to use a variable correction step, starting with a high μ and reducing it as the number of iteration get bigger.

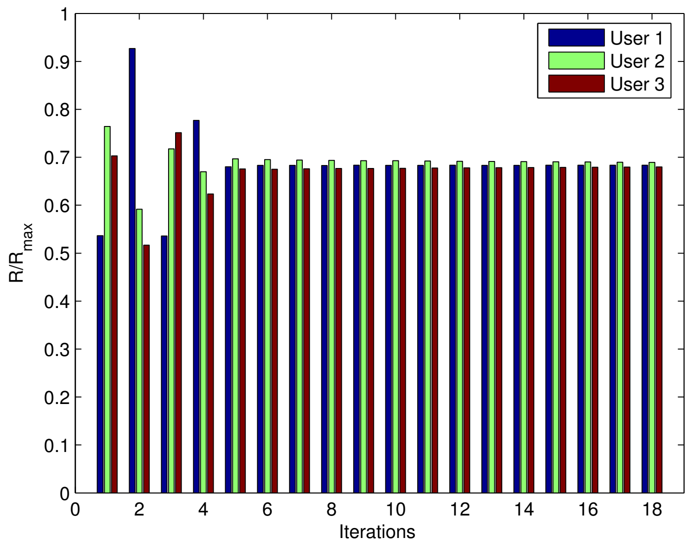


Figure 2.5: *BC algorithm using Newton-Raphson*

The comparison between the criterions of maximum sum capacity, BC, and transmitting at maximum PSD is done with respect to $R_i/R_{i,max}$ in Table 2.1, and with respect to the rates R_i, and sum rates $\sum_i R_i$ in Table 2.2. The interference from user 2 on user 1 and 3 is obviously seen in the maximum PSD strategy where user 1 and 3 can only transmit below 30 and 55 % of their maximum capacities respectively. On the contrary, user 2 does not suffer from much interference from 1 and 3. With the maximum sum rate strategy, user 2 has the highest percentage of its maximum capacity. This is due to its strong channel gain. User 1 is still strongly affected by crosstalk and can only transmit at about 50% of its maximum capacity. This issue is solved with the BC criterion which provides a balanced solution where all users send at almost 70% of their maximum capacities.

Inner Loop	Strategy	User 1	User 2	User 3
LS	max sum rate	0.5364	0.7648	0.7027
	BC concept	0.6840	0.6894	0.6795
	max PSD	0.2896	0.9659	0.5590
NR	max sum rate	0.5366	0.7643	0.7030
	BC concept	0.6835	0.6894	0.6798
	max PSD	0.2896	0.9659	0.5590

Table 2.1: *Comparison between R/R_{max} for sum rate maximization, BC, and maximum PSD results*

Algo	Strategy	User 1	User 2	User 3	$\sum_i R_i$
LS	max sum rate	4.777	12.393	10.0456	27.214
	BC concept	6.091	11.171	9.713	26.976
	max PSD	2.579	15.651	7.991	26.221
NR	max sum rate	4.779	12.385	10.050	27.214
	BC concept	6.087	11.172	9.718	26.976
	max PSD	2.579	15.651	7.991	26.221

Table 2.2: *Comparison between capacities in Mbps for sum rate maximization, BC, and maximum PSD results*

Table 2.1 & 2.2, show that the proposed NR algorithm provides results similar to the line search method in all cases. However the NR algorithm complexity is much smaller, as explained in the previous section. Figures 2.6 and 2.7 show the power allocation obtained by both the LS and the NR methods for the BC case. Once again, it is clear that the results of both methods are very close. Moreover, the figures confirm that the power allocation has minimal changes between two adjacent tones, so that we are in an appropriate situation to apply the Newton-Raphson and to speedup the convergence significantly.

In order to evaluate the complexity gain of the NR algorithm, several simulations have been done with different numbers of users. The line

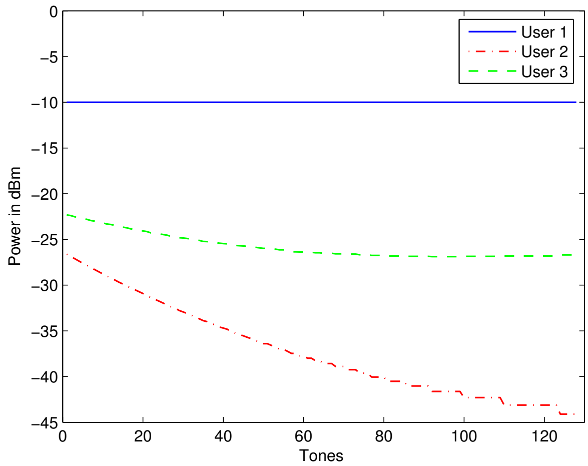


Figure 2.6: *Power allocation using Newton-Raphson step direction*

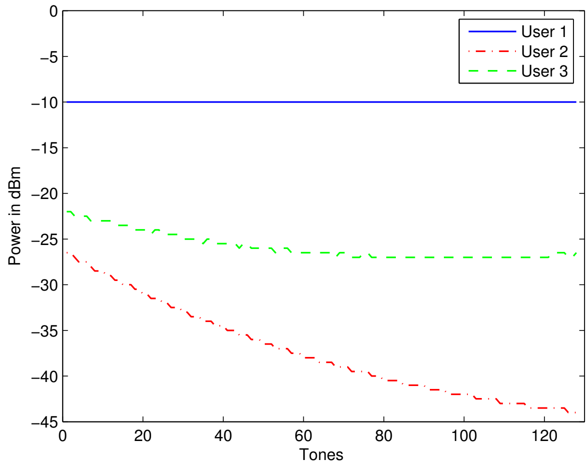


Figure 2.7: *Power allocation using Line Search*

lengths have been selected randomly in the interval [500 4500] m. In all cases, the BC algorithm has been applied with both methods (LS and NR). We only investigate the last inner loop corresponding to the BC

Inner Loop	Strategy	3 users	5 users	10 users	15 users
Line Search	$\sum_i \omega_i R_i$	1380.3	2294.1	2201.3	2155.3
	Execution Time	6.92	39.257	536.822	1067.595
	Iterations	495	1288	8645	4569
Newton Raphson	$\sum_i \omega_i R_i$	1380.3	2294.3	2202.6	2156.1
	Execution Time	0.221	0.671	3.695	8.112
	Iterations	194	829	1763	1461

Table 2.3: *Evolution of execution time, and iterations with the number of users for different inner loops*

point. Table 2.3 shows the normalized weighted sum rate ($\sum_i \omega_i R_i$) in bps/Hz obtained by the algorithm, the execution time of the last inner loop in seconds, and the corresponding number of iterations. Fig.2.8 shows the total complexity C for the different methods, given as the number of iteration times the order of complexity ($\frac{P_{max}}{\delta P}K^2$ for LS and K^3 for NR). The proposed NR method is significantly less complex.

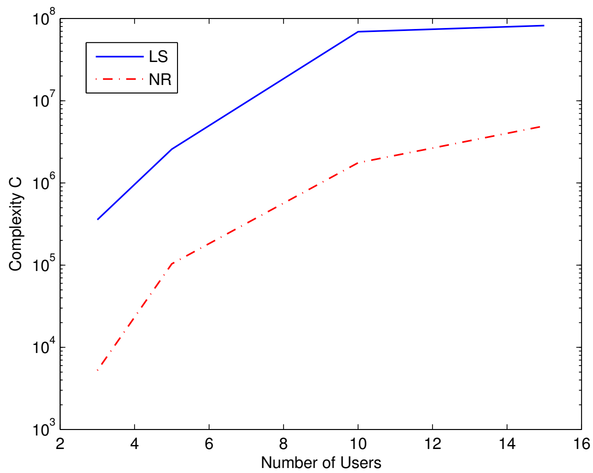


Figure 2.8: *Complexity vs users number*

2.7.2 Optimal Steepest Ascent

In this part we compare the numerical results obtained for the optimal SA, NR, and ISB algorithms. The comparison is made for the sum rate criterion, the number of tones is increased to 1024 and the algorithms are run for the following scenario: A system of 7 interfering users is considered, where the users have the following distances from the CO/ONU:

Users	user 1	user 2	user 3	user 4	user 5	user 6	user 7
Distances	4.8 km	400 m	3 km	2.4 km	4.5 m	3.8 km	2.3 km

The simulation is done for 1024 tones, with tone zero corresponding to 258.75 kHz. A PSD mask of -30dBm/Hz is imposed over all tones. The total power allowed per user is 20 dBm.

Number of users	3 users	5 users	7 users
SA	21542	24007	26944
NR	20562	23885	26896
ISB	21516	24029	26833

Table 2.4: *Sum capacity obtained for different algorithms in bit/DMT symbol*

The simulations are done in the presence of the first 3 users only, then for the first 5 users, and finally for all 7 users. Table 2.4 gives the sum capacities for the three simulations using different optimization algorithms. Table 2.5 gives the average execution of internal loops required for each one of these simulations.

Number of users	3 users	5 users	7 users
SA	1.4	3.7	7.6
NR	1.8	4.3	11
ISB	42	120	240

Table 2.5: *Execution time for different algorithms in seconds*

The ability to test for different values of α in the quazi-optimal steepest ascent guaranties that at each step the algorithm provides a near global optimal solution in the gradient direction even when it initializes at a point far from optimal or when we have multiple optimums. Thats why in the two scenarios explained above the SA algorithm provided near optimal results with relatively smaller complexity.

In the Newton-Raphson case the algorithm tends to follow the local optimum that lies at vicinity of the initial guess. The argument that "adjacent tones' optimums are close to each others and lie in the same region" works reasonably well, but over a large number of tones the global optimum may slip away from the initial optimal region while the NR algorithm is still stuck in it with local optimums.

2.7.3 Enhanced ISB Algorithm

Now we test the ISB algorithm for the same scenario set for the optimal SA. Table 2.6 compares the different versions of the ISB algorithm. ISB-LS refers to the original ISB algorithm with line search as the interior loop. ISB-LS with SO represents the original ISB algorithm with successive optimization. From the table we can see that the successive optimization speeds up the convergence at the expense of bit-rate. As explained in section 2.6.2.2, successive optimization increases the chance of staying in a local optimum region for a large range of tones. The enhanced ISB algorithm reduces considerably the execution time to less than a minute, and at the same time increases the bit-rate.

Algorithm	ISB-LS	ISB-LS with SO	Enhanced Fast ISB
Sum-rate	21877	21979	22340
Execution time	14.7 min	3.65 min	40 sec

Table 2.6: *Comparison of the different ISB algorithms*

Fig. 2.9 plots the sum capacity over the tones for the three stages of the enhanced ISB algorithm. As expected the ASP algorithm has several ups and downs for the different tones. The forward successive optimization (FSO) algorithm manages to smoothen the capacity and to improve the optimization. However we can see, in the example given, that the FSO get stuck in local optimums for the tones 136−284. It only reaches a better optimum at the tone $n = 284$. The advantage of using the reverse successive optimization is clear. When performing the RSO the optimization at tones $136 - 284$ can be improved from the use of the solution at $n = 284$ as starting point.

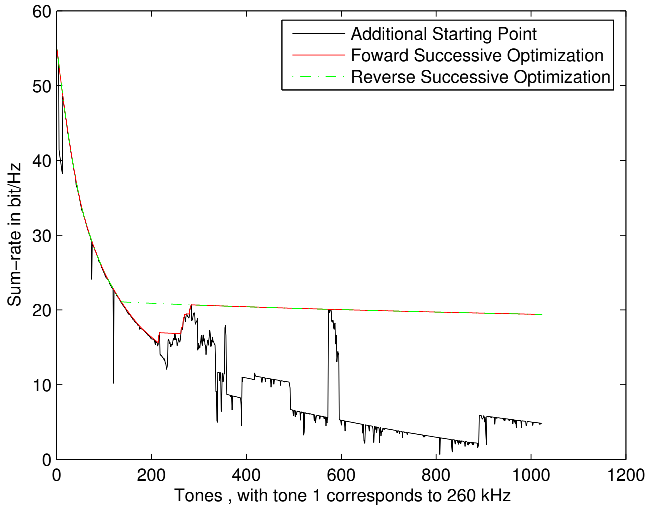


Figure 2.9: *Total sum capacity for the different stages of the enhanced ISB algorithm*

2.8 Conclusion

In this chapter we have presented an algorithm that is able to find the "Balanced Capacity" point in a DSL system environment. Furthermore, for fixed weighting (inner loop) coefficients we have shown how to take advantage of the similarity of the channel gains between two adjacent tones by using a simple technique based on Successive Optimization. This technique allows the implementation of fast gradient type algorithms for the DSL spectrum optimization. It may also be used to speed up the convergence of the existent state of the art algorithms such as ISB. We have also proposed several methods for the enhancement of the ISB algorithm. For example a method based on a gradient approximation was used to reduce the complexity of the ISB algorithm. We have also shown how to improve the performance of ISB using multi start points techniques. Finally, it was shown that a further performance gain can be obtained by coupling the multi start points with a double successive optimizations (both forward and reverse). These techniques allow the ISB algorithm to avoid local optimums on the different tones, which enhances the overall optimization.

Chapter 3

Crosstalk Channel Gain Estimation

3.1 Introduction

The main idea behind dynamic spectrum management (DSM) [7]-[39] is to establish a coordination between different DSL systems and lines. The objective of this coordination is to reduce the level of interference between the different lines and to increase the overall bit rate. This coordination between interfering lines, can be done on the signal level (level 3), or on the spectrum level (level 2). While level 3 DSM is complex and requires major changes to the current DSL systems, level 2 DSM is relatively easy to implement with the current modem designs, and it requires minimal changes to the standards.

In literature, level 2 DSM can be applied by 2 types of algorithms: distributed algorithms such as IWF [40] and ASB [41], and centralized algorithms such as OSB [43], ISB [30, 44]. In distributed algorithms, each modem tries to optimize it is own spectrum to achieve higher bit rates, no exchange of information between the different modems is assumed, and each modem relies on the knowledge of its direct channel gain, and the ambient SNR per tone to achieve the optimization. In centralized algorithms the coordination between the different lines is necessary, the optimization is done for the entire DSL system, where it takes into account the direct channel gains and the received and caused crosstalks of each modem, the optimization find the optimal power spectrum density (PSD) shape that improves the entire system capacity. Centralized DSM algorithms usually make 2 assumptions :

1. The existence of a spectrum management center where the algorithm will run and allocate a PSD for each modem.

2. The total knowledge of both direct and crosstalk channels of the DSL system.

From the two assumptions above, it is clear that in order to implement a centralized level 2 DSM algorithm, one should first proceed with an estimation of the crosstalk channel gains and then use any of the existing centralized algorithms to achieve the optimization.

This chapter deals with the estimation of the crosstalk channel gains for DSM level 2 applications. Crosstalk channels estimation techniques exist in the literature. In [49] a non-modem based technique is proposed, this technique correlates the noise PSD on a given line with a basis set of known DSL systems' PSD in order to determine the most dominant disturbers. In [50] the establishment of a new third party site was proposed, this site collects the transmitted and received signals on each DSL modem, a correlation process between these different signals is performed in order to remove the time difference between input and output signals, and then a least square estimator is applied to find the crosstalk coupling function. Paper [51] takes advantage of the initialization of new lines, where during initialization a pseudo random training sequence is used for the direct channel identification. In this method, the victim's receiver detects the training sequence, and then uses it to perform a least square estimation of the new crosstalk channels. In [52] a blind method for channel tracking is proposed, where the statistics of the different constellations is used to apply an EM algorithm, however the initial channel estimation is done using training sequences. In [53], a multi rate systems assumption is taken, thus different crosstalkers may have different sampling rates, blocking techniques are used to rend the crosstalk channels time invariant, then a least square estimation based on known pilot sequences is used. Most of these proposals requires modifications to the existent standards, and a modem based estimation.

In [54] the authors uses the SELT procedures (single-ended line testing) to measure the crosstalk PSD at the different lines of the DSL systems. From the measured PSD they deduce the crosstalk channel gain at the different frequencies. This procedure respect the DSL standards, however it requires a SELT testing over all DSL lines, during which only one line can transmit useful data.

In this chapter we assume the existence of a monitoring system that observes the DSL lines over a long period of time. In this monitoring system, the main events and the key indicators about the different DSL lines are stored. Thus information about the SNR, the time of connection/disconnection of each user, the noise level and the transmitted signal power level are supposed to be available. Using these information we propose a crosstalk gain estimator based on the SNR changes at the time of a connection/disconnection of a user. This estimation can be improved furthermore using a time domain model for the crosstalk channel between the different lines. This time domain model estimation plays the role of a compressing technique as well.

The chapter organization is as follow: Section 3.2 describes the system model. Section 3.3 describes the proposed channel gain estimator. Section 3.4 gives the confidence interval of the estimates. Section 3.5 deals with the compression and enhancement of the estimator. The simulation and numerical results are reported in section 3.6. Finally in section 3.7 we conclude.

3.2 DSL System Model

We consider DSL systems using DMT. The channel may be decomposed in K parallel sub-channels, where K represents the total number of tones. In a DSL system, and for lines served by the same DSLAM (Digital Subscriber Line Access Multiplexer), the assumption that all the lines are synchronous can be held. However this is not always true where crosstalk caused by lines served by different DSLAMs may be asynchronous. In this chapter we make the assumption that all the lines of the system are synchronized, the received signal Y_i of user i in sub-channel k is given by:

$$Y_i(k) = H_{ii}(k)X_i(k) + \sum_{l \neq i} H_{il}(k)X_l(k) + \omega(k) \tag{3.1}$$

where X_i is the useful signal of user i, ω is the additive noise, $H_{ii}(k)$ is the direct channel of user i for tone k, and $H_{il}(k)$ is the crosstalk channel from line l to line i for tone k. If the number of interfering users is large, the crosstalk signal maybe approximated by a Gaussian noise using the central limit theorem. From (3.1), one can deduce the SNR at tone k:

$$\mathrm{SNR}_i(k) = \frac{|H_{ii}|^2(k)P_i(k)}{\sum_{l \neq i} |H_{il}|^2(k)P_l(k) + \sigma_\omega^2(k)}, \tag{3.2}$$

where $P_i(k)$ and $P_l(k)$ are the power transmitted by lines i and l respectively. The variance associated with the background noise at tone k is denoted by $\sigma_\omega^2(k)$. The bitrate of user i can be evaluated by the formula:

$$R_i = \sum_k \log_2 \left(1 + \frac{1}{\Gamma}\mathrm{SNR}_i(k)\right), \tag{3.3}$$

where Γ is the SNR gap.

3.3 Crosstalk Channel Gain Estimator

In the current DSL systems, the CPEs (customer-premises equipments) periodically report the SNR per tone to the central office, this SNR is averaged over a large number of DMT symbols. SNR measurements can also be retrieved upon the request of CO (central office). The crosstalk gain estimator proposed in this chapter mainly depends on the reported SNR, it is thus essential to study the behavior of the SNR estimation performed at the CPE.

3.3.1 SNR Estimator

For the synchronous case, where all the disturbing lines are synchronized with the direct line, tone k is only affected by crosstalk caused by the same tone k of the other DSL lines, the analytical expression of the SNR is given by (3.2). If we suppose an error free decoding of the useful signals $X_i(k)$ at the receiver of user i, assuming that the direct channel H_{ii} is known perfectly for the receiver (realistic assumption under the current DSL standards [55]), the SNR estimator for line i at tone k averaged over N DMT symbols can be written as:

$$\hat{\mathrm{SNR}}_i(k) = \frac{|H_{ii}|^2(k)P_i(k)}{\frac{1}{N}\sum_{n=1}^{N}|\omega_{i,n}|^2(k)}, \tag{3.4}$$

where $\omega_{i,n}$ represents the crosstalk at line i plus the background noise for DMT symbol n: $\omega_{i,n}(k) = \sum_{l\neq i} H_{il}(k)X_{i,n}(k) + \omega_n(k)$, $\omega_n(k)$ is the background noise. The crosstalk power $\sigma^2_{\omega_i}$ can be written as:

$$\sigma^2_{\omega_i}(k) = \sum_{l\neq i}|H_{il}|^2(k)P_l(k) + \sigma^2_{\omega}(k), \tag{3.5}$$

where $\sigma^2_{\omega}(k)$ is the power of the background noise at the tone k. As $|H_{ii}|^2(k)P_i(k)$ is supposed to be known, the study of the SNR leads to the study of the noise variance estimator, which is a sample variance as shown in (3.4).

3.3.2 The Noise Variance Estimator

From (3.4), we can rewrite the noise variance estimator $\hat{\sigma}^2_{\omega_i}$ as:

$$\begin{aligned}\hat{\sigma}^2_{\omega_i}(k) &= \frac{1}{N}\sum_{n=1}^{N}|\omega_{i,n}|^2(k)\\ &= \frac{|H_{ii}|^2(k)P_i(k)}{\hat{\mathrm{SNR}}_i(k)}.\end{aligned} \tag{3.6}$$

When the number of interfering users is limited, the Gaussian approximation of ω_i can no longer be held, however the central limit theorem can still be applied on $\hat{\sigma}^2_{\omega_i}$ where the estimation is being done over a large number of DMT symbols N (typically $N > 100$) in this case we can assume that $\hat{\sigma}^2_{\omega_i} \sim \mathcal{N}(\sigma^2_{\omega_i}, \mathrm{Var}(\hat{\sigma}^2_{\omega_i}))$, where $\mathrm{Var}(\hat{\sigma}^2_{\omega_i})$ can be shown to be equal to:

$$\begin{aligned}\mathrm{Var}(\hat{\sigma}^2_{\omega_i}(k)) &= \frac{1}{N}\left(\sum_{l\neq i}|H_{il}|^4\left(E[|X_l|^4(k)] - P_l^2(k)\right)\right) \\ &+ \frac{1}{N}\sum_{l\neq i}\sum_{p\neq i,l}|H_{il}|^2(k)|H_{ip}|^2(k)P_l(k)P_p(k) \\ &+ \frac{2}{N}\sum_{l\neq i}|H_{il}|^2(k)P_l(k)\sigma^2_\omega(k) + \frac{2}{N}\sigma^4_\omega(k) \qquad (3.7)\end{aligned}$$

where E is the expectation operator. To use (3.7) we need to know a priori the crosstalk channel gains, this is not available for all crosstalks and for all the times. However, under the assumption that ω_i follows a Gaussian distribution (which is asymptotically true when the number of interfering users is very large), the variance of $\hat{\sigma}^2_{\omega_i}$ can be computed in function of $\sigma^2_{\omega_i}$:

$$\begin{aligned}\mathrm{Var}\left(\hat{\sigma}^2_{\omega_i}\right) &= \sum_{n=1}^{N}\mathrm{Var}\left(\frac{|\omega_{i,n}(k)|^2}{N}\right) \\ &\cong \frac{2\sigma^4_{\omega_i}}{N}. \qquad (3.8)\end{aligned}$$

Where $E\left[|\omega_{i,n}(k)|^4\right] = 3\sigma^4_{\omega_i}$ is used (fourth moment of the central Gaussian variable $\omega_{i,n}(k)$).

3.3.3 Crosstalk Lines Identification

The identification procedure described in this section assumes that the monitoring system was in function for a long period of time. This means that the system has acquired a database that includes events such as connection/disconnection of a user or a power cutback on a certain line, plus the corresponding SNR changes and the PSD of the users/lines. The simplest way to identify the crosstalk lines is to exploit the different events registered in the monitoring system to estimate the crosstalk channel gains between the different lines.

3.3.3.1 Connecting User

The straightforward approach to estimate the crosstalk channel gain caused by a given line u, is to observe the SNR changes on the other lines when user u is connecting/disconnecting: Let t_1 be the time at which user u is connecting to the network. The equivalent noise present on a line i before and after t_1 can be expressed as:

- When $t < t_1$ the equivalent noise ω_1 at tone k on line i is given by:

$$\omega_1(k) = \sum_{l \neq i} H_{il}(k) X_l(k) + \omega(k) \tag{3.9}$$

  while the variance is:

$$\sigma^2_{\omega_1}(k) = \sum_{l \neq i} |H_{il}|^2(k) P_l(k) + \sigma^2_{\omega}(k).$$

- When $t > t_1$, the additional interference caused by user u increases the variance of the equivalent noise present on line i, thus the SNR decreases. The new equivalent noise ω_2 is given by:

$$\omega_2(k) = \omega_1(k) + H_{iu} X_u(k) \tag{3.10}$$

  and the variance is:

$$\sigma^2_{\omega_2}(k) = \sigma^2_{\omega_1}(k) + |H_{iu}|^2(k) P_u(k).$$

  Now we may give an explicit expression of $|H_{iu}|^2$ in function of the noise variances:

$$|H_{iu}|^2(k) = \frac{\sigma^2_{\omega_2}(k) - \sigma^2_{\omega_1}(k)}{P_u(k)}. \tag{3.11}$$

To get an estimator of $|H_{iu}|^2$, we use the method of moment [56], which replaces the noise variance before and after t_1 ($\sigma^2_{\omega_1}$ and $\sigma^2_{\omega_2}$) by their respective samples variance:

$$
\begin{aligned}
|\hat{H}_{iu}|^2(k) =& \frac{\frac{1}{N}\sum_{n=1}^{N}|\omega_{2_n}(k)|^2 - \frac{1}{N}\sum_{n=1}^{N}|\omega_{1_n}(k)|^2}{P_u(k)} \\
=& \frac{|H_{ii}|^2(k)P_i(k)}{P_u(k)}\left(\frac{1}{S\hat{N}R_2(k)} - \frac{1}{S\hat{N}R_1(k)}\right).
\end{aligned} \tag{3.12}
$$

A similar procedure may be done if a user is disconnected or when the power changes.
To be noticed that the knowledge of the direct channel gain $|H_{ii}|^2$ is not necessary, as we only need the ratio $H_{iu}|^2/|H_{ii}|^2$ in order to perform DSM level 2, however the knowledge of $|H_{iu}|^2$ is needed for later work in this chapter.

3.3.3.2 Estimator Performance

By approximating the estimation error of the noise variance by a Gaussian noise, the estimator $|\hat{H}_{iu}|^2$ would be following a Gaussian distribution. The variance of $|\hat{H}_{iu}|^2$ can approximated by:

$$
\mathrm{Var}\left(|\hat{H}_{iu}|^2(k)\right) = \frac{2(\sigma^4_{\omega_2}(k) + \sigma^4_{\omega_1})(k)}{NP_u^2(k)}. \tag{3.13}
$$

From (3.13) we can conclude that the quality of the estimation depends on the ambient noise including the crosstalk of the different active lines. Different situations may occur, where for the same estimation we may get a good results on one tone and a bad results on another one depending on the background noise and the current active lines.

3.3.3.3 Maximum Likelihood Combiner

The monitoring system proposed in this chapter is supposed to record information for a long period of time. Different estimations can be performed for the same crosstalk channel and for different conditions. These estimations are considered to be independent and follow a Gaussian distribution. In this case, the available estimates may be combined using the following maximum likelihood combiner.

$$|\tilde{H}_{iu}|^2(k) = \frac{\sum_{t=1}^{T} |\hat{H}_{iu}^t|^2(k)/\mathrm{Var}\left(|\hat{H}_{iu}^t|^2(k)\right)}{\sum_{t=1}^{T} 1/\mathrm{Var}\left(|\hat{H}_{iu}^t|^2(k)\right)}. \tag{3.14}$$

Where $|\hat{H}_{iu}^t|^2(k)$ represents the result given by estimation t and T is the total number of estimations.

3.4 Confidence Interval for the Gain Estimator

In this section we will define a confidence interval for the estimator $|\hat{H}_{iu}|^2$ defined by (3.12). A confidence interval gives an estimated range of values, with a given probability that this range of values contains the real value of the estimated parameter. First we will start by defining the relative error of the gain estimator (3.12) from which we can deduce the confidence interval.

3.4.1 Relative Error of the Gain Estimator

The relative error of the noise estimator $\hat{\sigma}^2_{\omega_i}$ is defined as:

$$\rho_{\omega_i} = \frac{\hat{\sigma}^2_{\omega_i} - \sigma^2_{\omega_i}}{\sigma^2_{\omega_i}}. \tag{3.15}$$

While the relative error of the crosstalk gain estimator is given by:

$$\rho_{H_{iu}} = \frac{|\hat{H}_{iu}|^2 - |H_{iu}|^2}{|H_{iu}|^2}. \tag{3.16}$$

The relative error $\rho_{H_{iu}}$ can be rewritten as:

$$\rho_{H_{iu}} = \frac{\frac{\rho_{\omega_2}}{\sigma^2_{\omega_1}} - \frac{\rho_{\omega_1}}{\sigma^2_{\omega_2}}}{\frac{1}{\sigma^2_{\omega_1}} - \frac{1}{\sigma^2_{\omega_2}}}. \tag{3.17}$$

If the relative error of the noise estimator are the same ($\rho_{\omega_2} = \rho_{\omega_1} = \pm|\rho_{\omega_1}|$) we can find an upper bound on $\rho_{H_{iu}}$:

$$|\rho_{H_{iu}}| \leq \left|\rho_{\omega_1}\frac{1+r}{1-r}\right| \tag{3.18}$$

where $r = \sigma^2_{\omega_1}/\sigma^2_{\omega_2}$. From expression (3.18) it is clear that the upper bound on $|\rho_{H_{iu}}|$ is limited by $|\rho_{\omega_1}|$ as a lower level when r is very large or very small, and infinity as an upper level when the ratio $r = 1$. This means that when a crosstalk channel line is big enough to cause an important rise or drop in the SNR as the interfering user connect/disconnect, the relative error after the estimation of the crosstalk channel caused by the interfering line would be small. On the contrary, when the crosstalk channel influence is small, it causes insignificant changes to the SNR, and the relative error associated with the crosstalk channel estimation would be very large.

3.4.2 Confidence Interval

From the assumption $|\hat{H}_{iu}|^2 \sim \mathcal{N}\left(|H_{iu}|^2, \mathrm{Var}(|\hat{H}_{iu}|^2)\right)$ one can conclude that $\rho_{H_{iu}} \sim \mathcal{N}\left(0, \frac{\mathrm{Var}(|\hat{H}_{iu}|^2)}{|H_{iu}|^4}\right)$. The probability that $|\rho_{H_{iu}}| < a$ is given by:

$$\mathrm{Prb}(-a < \rho_{H_{iu}} < a) = \mathrm{erf}\left(a\frac{|H_{iu}|^2}{\sqrt{2\mathrm{Var}(|\hat{H}_{iu}|^2)}}\right) \tag{3.19}$$

where erf is the error function. From (3.19), and for $a < 1$, the confidence interval can be defined as:

$$\mathrm{Prb}\left(\frac{|\hat{H}_{iu}|^2}{1+a} < |H_{iu}|^2 < \frac{|\hat{H}_{iu}|^2}{1-a}\right) = \mathrm{erf}\left(a\frac{|H_{iu}|^2}{\sqrt{2\mathrm{Var}(|\hat{H}_{iu}|^2)}}\right).$$

3.5 Compression and Error Smoothening

In the current DSL systems, the total number of tones can be very large especially for VDSL systems, thus registering the values of all the crosstalk channels' gain and for all tones would require huge amounts of storage devices. In this section we propose a compression technique that would limit the information to be stored for each crosstalk channels. The main idea behind this technique is that the different tones of a crosstalk channels in the frequency domain represents a Fourier transform of a finite impulse response (FIR) in the time domain. Usually the FIR can be represented by a small number of taps that are limited to the length of the cyclic prefix (L_{CP}) of the DMT symbol. Thus the estimates of a crosstalk channel gain at all the different tones can be represented as the FFT of a time domain vector with limited size.

The crosstalk channel gain as defined in this chapter is equivalent to the power spectrum density of the crosstalk channels, thus the time domain vector used for the compression procedure represent the autocorrelation of the time domain crosstalk channel. This technique has the double advantage of limiting the amount of information to be stored on one hand, and of smoothening the estimation error on the other hand as this technique incorporates estimates of different tones together which has the effect of reducing the estimation variance.

Let $\mathbf{H_{iu}} = [H_{iu}(1)...H_{iu}(k)...H_{iu}(K)]^T$ be the crosstalk channel between line i and line u in the frequency domain. We have $\mathbf{H_{iu}} = \mathbf{W}^{K\times L_{CP}}\mathbf{h_{iu}}$ where $\mathbf{W}$ represents the FFT matrix and $\mathbf{h_{iu}}$ is the time domain representation of the crosstalk channel between the lines i and u, the length of $\mathbf{h_{iu}}$ is limited by the length of the cyclic prefix (L_{CP}) with $L_{CP} << K$. The gain of the crosstalk channel $\mathbf{H_{iu}}$ is given by the vector: $\mathbf{H^2_{iu}} = [|H_{iu}|^2(1)...|H_{iu}|^2(k)...|H_{iu}|^2(K)]^T$, let $\mathbf{R_{iu}}$ represents the autocorrelation of the vector $\mathbf{h_{iu}}$, we have that:

$$\mathbf{H^2_{iu}} = \mathbf{W}^{K\times 2L_{CP}-1}\mathbf{R_{iu}}. \tag{3.20}$$

Since $\mathbf{h_{iu}}$ represents a real time domain channel we can conclude that $\mathbf{R_{iu}}$ is real and symmetric around its center. The total length of $\mathbf{R_{iu}}$ is equal to $2L_{CP}-1$. The FFT matrix $\mathbf{W}^{K\times 2L_{CP}-1}$ can be seen as the combination of two matrices $\mathbf{W_1}$ and $\mathbf{W_2}$: $\mathbf{W}^{K\times 2L_{CP}-1} = [\mathbf{W_1}\ \ \mathbf{W_2}]$ where $\mathbf{W_1}$ represents the first L_{CP} columns of the FFT matrix while $\mathbf{W_2}$ represents the last $L_{CP}-1$ columns.

The autocorrelation $\mathbf{R_{iu}}$ can also be seen as the combination of two vectors $\mathbf{R_{iu,1}}$, and $\mathbf{R_{iu,2}}$: $\mathbf{R_{iu}} = \left[\mathbf{R_{iu,1}}^T \mathbf{R_{iu,2}}^T\right]^T$ where $\mathbf{R_{iu,1}}$ represents the first L_{CP} elements of $\mathbf{R_{iu}}$, while $\mathbf{R_{iu,2}}$ represents the last $L_{CP}-1$ elements. Since the autocorrelation $\mathbf{R_{iu}}$ is symmetrical on its center we have that $\mathbf{R_{iu,2}} = [\mathbf{R_{iu,1}}(L_{CP}-1)...\mathbf{R_{iu,1}}(1)]$. We define the matrix $\mathbf{W_{2,m}}$ as the following; for each row corresponding to tone k we have:

$$\mathbf{W_{2,m}}(k,:) = [\mathbf{W_2}(k, L_{CP}-1)...\mathbf{W_2}(k,1)\ \ 0].$$

Now we can rewrite expression (3.20) as:

$$\mathbf{H^2_{iu}} = \mathbf{W_t}\mathbf{R_{iu,1}}, \tag{3.21}$$

where $\mathbf{W_t} = \mathbf{W_1} + \mathbf{W_{2,m}}$.

A least square (LS) estimator can be used to compress the estimated crosstalk channel gain $\mathbf{\hat{H}^2_{iu}} = [|\hat{H}_{iu}|^2(1)...|\hat{H}_{iu}|^2(k)...|\hat{H}_{iu}|^2(K)]$ into a vector $\mathbf{R_{iu,e}}$ of length L_{CP}:

$$\mathbf{R_{iu,e}} = \left(\mathbf{W_t}^H \mathbf{W_t}\right)^{-1} \mathbf{W_t}^H \mathbf{\hat{H}^2_{iu}}. \tag{3.22}$$

To obtain the crosstalk channel gain from the compressed data we apply equation (3.21):

$$\mathbf{\hat{H}^2_{iu,c}} = \mathbf{W_t}\mathbf{R_{iu,e}}. \tag{3.23}$$

The estimation error variance of $\mathbf{\hat{H}^2_{iu,c}}$ is usually smaller than that of $\mathbf{\hat{H}^2_{iu}}$ since in (3.22) we combined the estimations of K tones to obtain the time domain model. We can further reduce the estimation error variance after compression by incorporating the crosstalk gain estimator variance at each tone in the expression (3.22), this can be done by using a weighted least square instead of a normal LS:

$$\mathbf{R_{iu,w}} = \left(\mathbf{W_t}^H \mathbf{D_{iu}} \mathbf{W_t}\right)^{-1} \mathbf{W_t}^H \mathbf{D_{iu}} \mathbf{\hat{H}^2_{iu}} \tag{3.24}$$

where $\mathbf{D_{iu}}$ is a diagonal matrix that represents the weight of the estimation at each tone. This weight is represented by the inverse of the estimation variance at k where: $\mathbf{D_{iu}}(k,k) = 1/\mathrm{Var}\left(|\hat{H}_{iu}|^2\right)(k)$.

3.6 Simulation

In this section, we consider a DSL system with 4 users: user 1 corresponds to the line over which the estimation is being done, all the observed SNR variation are being recorded at the end line of user 1, user 2 represents a new user during its connection on the DSL system thus the objective is to estimate the crosstalk channel from user 2 to user 1 by observing the SNR change on line 1 induced by the connection of user 2. User 3 and user 4 correspond to the remaining lines in this system, the crosstalk of these 2 users contributes to the background noise.

We suppose that the estimation is being done two times, one time under the presence of user 3 while user 4 being disconnected, and another time where user 3 is disconnected and user 4 is connected. This means that the estimation is being done under two different background noise represented by Fig.3.1. The 2 background noise represent the AWGN + crosstalk coming from the other connected user. The crosstalk channel gains used in this estimation are measured crosstalk channel between 400 m France Telecom (FT) cables. The DMT symbol is decomposed over 1024 tones where the spacing between tones is 4.3 kHz, and the first tone correspond to 258.75 kHz. The AWGN noise is supposed to have a PSD of -140 dBm/Hz over all tones and for all users.

From Fig.3.1 one can predict that under background noise 1 we will have a good estimation for the tones corresponding to lower frequencies and bad estimation for the tones corresponding to higher frequencies. This tendency would be inversed for the estimation done under background noise 2. Fig.3.2 shows typical estimations done under the two background noise. The graphics of these estimations confirm the performance tendencies explained earlier. Some of the estimated crosstalk channels' gain have negative values, this is due to the high estimation errors. The negative estimates are kept intact in order to maintain a non biased estimation.
Fig.3.3 and Fig.3.4 shows the variance of the crosstalk gain estimators done over background noise 1 and background noise 2 respectively. In both figures the red curve represents the crosstalk gain estimator variance computed using a Gaussian approximation of the crosstalk +

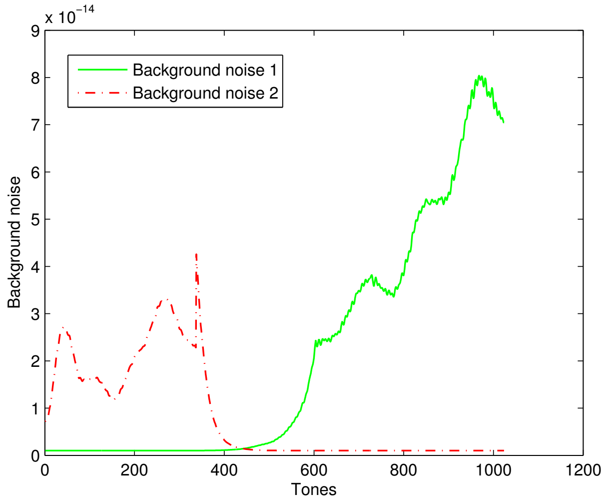


Figure 3.1: *Background Noise*

noise estimation variance (equation (3.8)), while the blue curve represents the crosstalk gain estimator variance computed using the theoretical formula of the crosstalk + noise estimation variance (equation (3.7)), the dashed green curve represents the estimated variance of the crosstalk gain estimator computed over 100 different estimations. We can see that there is a difference between the Gaussian approximation and the theoretical approach, this is due mainly to the fact that only 2 crosstalk are present in the current simulation which is not enough to apply the central limit theorem, however the Gaussian estimation can still be used for the ML combiner and for the weighted LS compressor as it gives a fair idea of the weight of each estimates.

To improve the overall estimation, we combine the two estimations using the ML combiner (equation (3.14)), where the Gaussian approximation was used for computing the estimator variances. The result of the ML combiner is then compressed into a 32 tap filter using the LS compressor (equation(3.22) & equation(3.23)), Fig.3.5 shows the results of the ML combiner and of the LS compressor. It is clear that the used compression technique has the double advantage of compressing the channel state information from 1024 to just 32 on one hand and of smoothening and reducing the estimation errors on the other hand.

Table 3.1 represents the mean square error (MSE) averaged over all

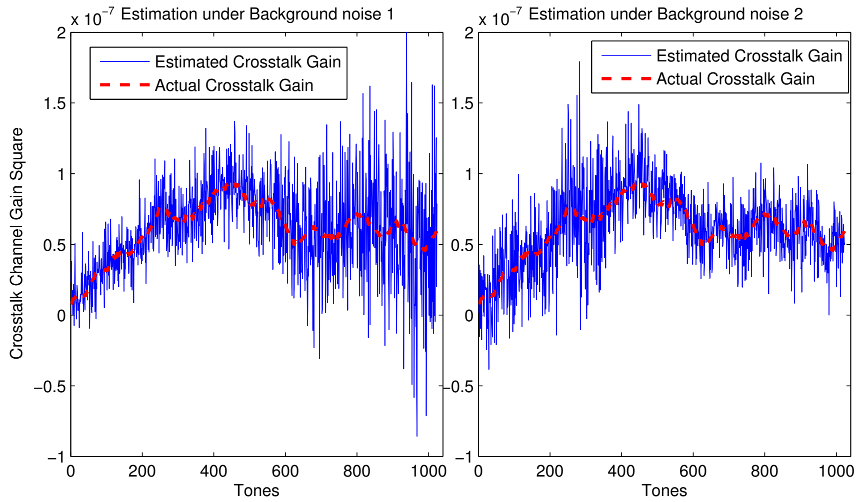


Figure 3.2: *Estimation done using SNR variation method*

tones for the different estimation techniques discussed earlier and for the two background noise. From this table we can conclude that the estimation can be improved just by using the compression technique: when the LS compressor is applied on the SNR variation estimator; the MSE is reduced by a ratio of 40 and 53 for background noise 1 and 2 respectively. The ML combiner reduced the MSE by a ratio of 4 for the estimation done over noise 1 and by a ratio of 2 for the estimation done over noise 2. Combining the ML combiner and the LS compressor gives the best results, where it reduced the MSE by a ratio of 143 for the estimation done over background noise 1, and by a ratio of 73 for the estimation done over background noise 2. The estimation procedure proposed in this chapter was tested with different crosstalk channels, and the results were consistent with the presented results in this section.

Estimation Technique	SNR Estimator	LS	WLS	ML	ML + LS
MSE (Noise 1)$\times 10^{16}$	8.186	0.202	0.192	1.998	0.057
MSE (Noise 2)$\times 10^{16}$	4.211	0.079	0.0705	1.998	0.057

Table 3.1: *Mean Square Error of Different Estimation Techniques*

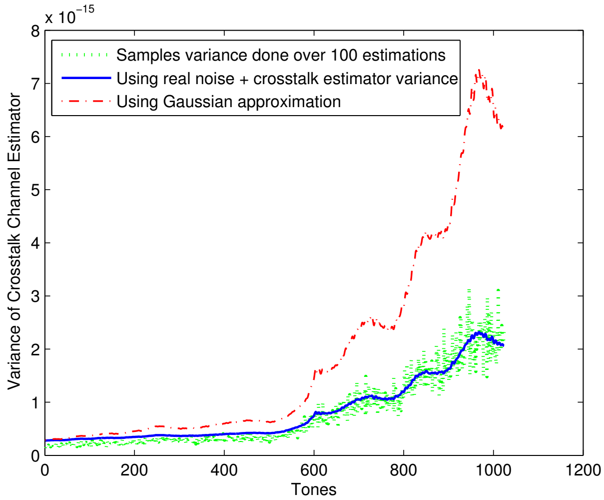


Figure 3.3: *Estimator Variance under Background Noise 1*

3.7 conclusion

In this chapter we showed that the estimation of the crosstalk channel gain is possible using a simple SNR observation, that relies on the informations available at the CO only without further techniques. In the case where we have a small background noise, that means users having the lowest crosstalk connecting first then those with stronger one connecting last, the estimator showed a good performance, in the inverse case we risk to not estimate the weak crosstalk channels as they will be submerged by the estimation noise caused by the strong crosstalk. A proposed remedy of this situation is to observe the lines for a long period of time, which increases the probability of having a good order of connections, then to combine the different estimations using a maximum likelihood combiner. Another option that can be studied is to lower the power of the user having a strong crosstalk, and to increase the power of the line having a weak crosstalk, but in this case the observation is being forced, and it is no longer a simple use of the data at the CO. We have also presented a compression technique that limits the crosstalk channel state information to be stored. The compression technique has been also shown to be beneficial in the smoothening and reducing of the estimation errors.

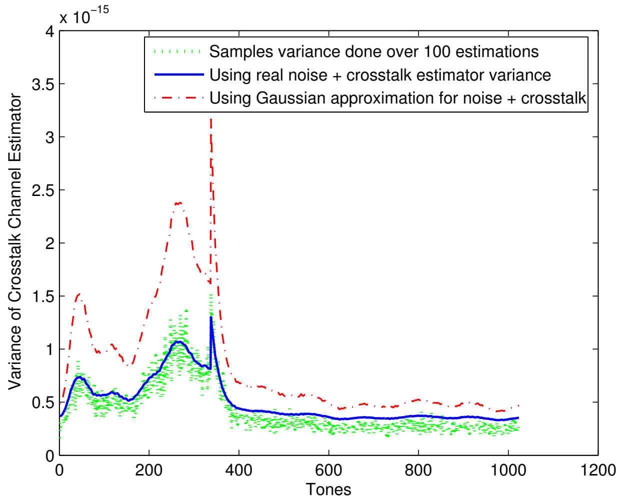


Figure 3.4: *Estimator Variance under Background Noise 2*

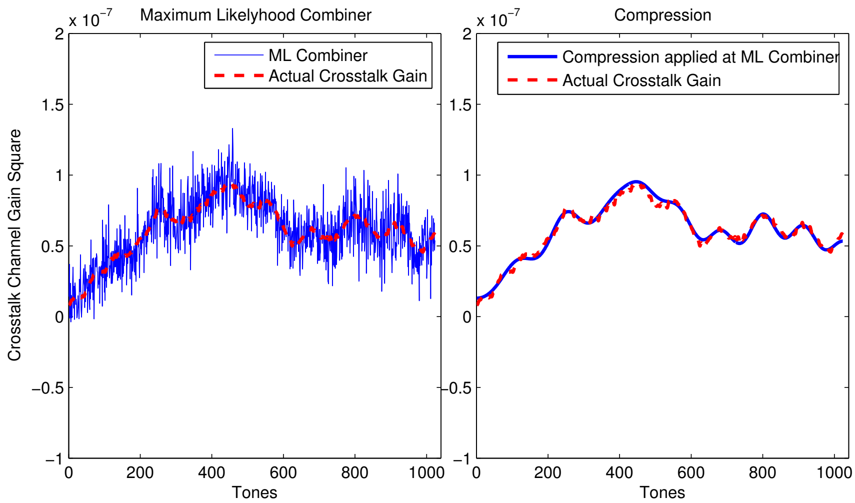


Figure 3.5: *Results of ML Combiner and LS Compressor*

Chapter 4

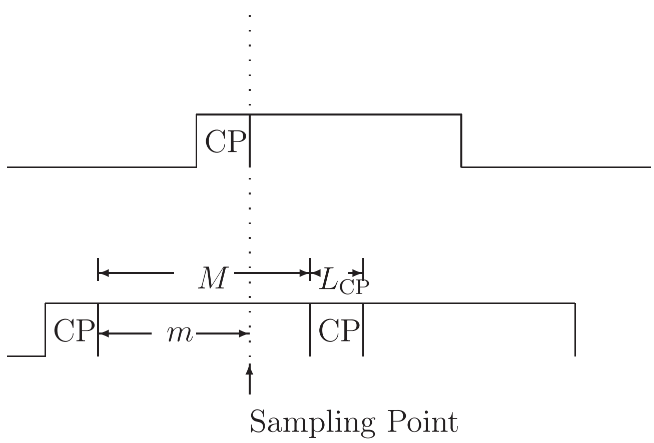
Asynchronous Crosstalk Estimation

4.1 Introduction

In literature, the assumption of a perfect synchronization between the different DSL users, is usually made, this synchronization is supposed to be done in both frequency and time. While synchronization is essential for applying crosstalk cancellation techniques, it is not considered a prerequisite for DSM level 1 applications, in fact distributed algorithms such as IWF and ASB do not make any particular assumption on the synchronization of the interfering users. In [59] the authors have proposed a centralized DSM algorithm for the asynchronous crosstalk case, however this algorithm is an heuristic algorithm and it does not guaranty the convergence toward a global optimum. DSM centralized algorithms that make the assumption of a perfect synchronization between the different DSL users are asymptotically optimal in the case of OSB, and they have a lower complexity of calculation, as the optimization in the case of a synchronous DSL system is linear in function of the total number of available tones, compared to an exponential complexity in the asynchronous case. Thus the synchronization of the different lines of a DSL system would have an important advantage for centralized DSM level 2 applications.

In this chapter, we will formulate an analytical model to the crosstalk caused by non synchronized users. We will show that when the synchronization is lost, the orthogonality between the different tones is lost as well, ie all tones that correspond to an interfering line, contribute to the crosstalk on each individual tone of a victim line. In fact, the crosstalk caused by an interfering line in the asynchronous case on a given victim's tone, can be seen as linear combination of the synchronous crosstalk caused by the same line on all the victim's tones. Thus, even in a non-synchronous case it is important to estimate the synchronous crosstalk channels of individual tones in order to implement centralized DSM algorithm, we will show that the crosstalk gain estimation procedure proposed in chapter 3 for the synchronous case, can be used in the case of non-synchronized crosstalk interference as well. Finally we will use the asynchronous crosstalk model in order to propose blind synchronization methods that can be applied on current non synchronized DSL systems without any major modification to the standards. This chapter is organized as the following, section 4.2 formulates an analytical model for the asynchronous crosstalk, section 4.3 proves that the crosstalk gain estimator proposed in chapter 3 can be used in the asynchronous case as well, section 4.4 presents a blind synchronization techniques, section 4.5 reports the simulation results, section 4.6 concludes this chapter.

4.2 Asynchronous Crosstalk



In the non synchronous case, the crosstalk signals are not synchronized to the main useful signal. In the case where the non synchronization delay between the desired signal and the crosstalk signal m_i is greater than the cyclic prefix length L_{CP}, the crosstalk signal will be formed of 2 consecutives DMT symbols and one cyclic prefix as illustrated in figure(4.2). The received time domain signal after sampling $y(n_s)$ at sample n_s for user i is given by:

$$y(n_s) = y_i(n_s) + y_c(n_s) + \nu(n_s), \tag{4.1}$$

where y_i is the desired signal, y_c is the noise due to the crosstalk signals, $y_c = \sum_{l \neq i} y_{c,l}$, where $y_{c,l}$ is the interference caused by user l on line i. As explained earlier, $y_{c,l}$ may be seen as the sum of two successive DMT symboles, plus the cyclic prefix. Let x_l and x_l^{next} be the two successive time domain DMT symbols at the transmitter side, and let h_{il} be the crosstalk channel impulse response between lines l and i

$$\begin{aligned} y_{c,l}(n_s) &= \Big(\sum_{v=0}^{L_{\mathrm{CP}}} h_{il}(v) x_l(n_s + m_l - v)\Big)\gamma_l(n_s + m_l) \\ &+ \Big(\sum_{v=0}^{L_{\mathrm{CP}}} h_{il}(v) x_l^{next}(n_s - v - (M - m_l + L_{\mathrm{CP}}))\Big) \\ &\times \gamma_l^{next}(n_s - (M - m_l + L_{\mathrm{CP}})) + s_l(n_s). \end{aligned} \tag{4.2}$$

Where:

- γ_l and γ_l^{next} are two time windows that limit the two DMT crosstalk symbols to the duration of the received signal y_i.
- M is the length of the DMT symbol
- L_{CP} is supposed to be equal to the crosstalk channel length
- m_l is the time delay between the user l under investigation and the reference user i
- $-m_l$ et $(M - m_l + L_{\mathrm{CP}})$ represent the respective delays of x_l and x_l^{next} with respect to the sampling point

- $s_l(n_s)$ is the cyclic prefix

γ_l and γ_l^{next} have the following shapes:

- $\gamma_l(n_s) = \begin{cases} 0 & for \quad 0 \leq n_s \leq m_l \\ 1 & for \;\; m_l + 1 \leq n_s \leq M \end{cases}$
- $\gamma_l^{next}(n_s) = \begin{cases} 1 & for \quad 0 \leq n_s \leq m_i - L_{\mathrm{CP}} \\ 0 & for \;\; m_i - L_{\mathrm{CP}} + 1 \leq n_s \leq M \end{cases}$

If we replace $n_1 = n_s + m_l$ and $n_2 = n_s - (M - m_l + L_{\mathrm{CP}})$ in 4.2, we may write an expression of the total noise in the asynchronous case:

$$\begin{aligned} \nu_{async}(n_s) &= \sum_{l \neq i} \left(\left(\sum_{v=0}^{L_{\mathrm{CP}}} h_{il}(v) x_l(n_1 - v) \right) \gamma_l(n_1) \right. \\ &+ \left. \left(\sum_{v=0}^{L_{\mathrm{CP}}} h_{il}(v) x_l^{next}(n_2 - v) \right) \gamma_l^{next}(n_2) + s_l(n_s) \right) \\ &+ \nu(n_s) \end{aligned} \tag{4.3}$$

We also have $[x_l(-L_{\mathrm{CP}}) \ ... \ x_l(-1)] = [x_l(M - L_{\mathrm{CP}}) \ ... \ x_l(M-1)]$, so we may consider that $\sum_{v=0}^{L_{\mathrm{CP}}} h_{il}(v) x_l(n_1 - v)$ is a circular convolution, then the fast fourier transform (FFT) of ν_{async} is given by ω_{async}:

$$\begin{aligned} \omega_{async}(k) &= \sum_{l \neq i} \left(\sum_{v=0}^{M} \Gamma_l(v) H_{il}(k - v) X_l(k - v) \exp(j2\pi k m_l / M) \right. \\ &+ \sum_{v=0}^{M} \Gamma_l^{next}(v) H_{il}(k - v) X_l^{next}(k - v) \\ &\times \left. \exp(-j2\pi k (M - m_l + L_{\mathrm{CP}})/M) + S_l(k) \right) + \omega(k) \end{aligned} \tag{4.4}$$

where H_{il} represents the crosstalk channel in the frequency domain, Γ_l, Γ_l^{next} are the FFT of γ_l, γ_l^{next} respectively, S_l is FFT of the s_l, and

ω is the FFT of ν.

The expression $\sum_{v=0}^{M} \Gamma_l(v) H_{il}(k-v) X_l(k-v)$ is a circular convolution where:
$[H_{il}(-M)X_i(-M) \; \ldots \; H_{il}(-1)X_i(-1)] = [H_{il}(0)X_i(0) \; \ldots \; H_{il}(M-1)X_i(M-1)]$.
Since ω_{async} has a zero mean, we may write the crosstalk expression in the asynchronous case as:

$$\sigma^2_{\omega,async}(k) = E\left[|\omega_{async}|^2(k)\right], \tag{4.5}$$

where $E[]$ is the expectation operator. In order to find the asynchronous crosstalk we need to find the expression $|\omega_{async}|^2$

$$\begin{aligned}
|\omega_{async}|^2(k) &= \left[\sum_{l\neq i}\left(\sum_{v=0}^{M} \Gamma_l(v) H_{il}(k-v) X_l(k-v)\exp(j2\pi k m_l/M)\right.\right. \\
&+ \sum_{v=0}^{M} \Gamma_l^{next}(v) H_{il}(k-v) X_l^{next}(k-v) \\
&\times \left.\left.\exp\left(\frac{-j2\pi k(M-m_l+L_{\mathrm{CP}})}{M}\right) + S_l(k)\right) + \omega(k)\right] \\
&\times \left[\sum_{l\neq i}\left(\sum_{v=0}^{M} \Gamma_l^*(v) H_{il}^*(k-v) X_l^*(k-v)\exp(-j2\pi k m_l/M)\right.\right. \\
&+ \sum_{v=0}^{M} \Gamma_l^{*next}(v) H_{il}^*(k-n) X_l^{*next}(k-v) \\
&\times \left.\left.\exp\left(\frac{j2\pi k(M-m_l+L_{\mathrm{CP}})}{M}\right) + S_l^*(k)\right) + \omega^*(k)\right].
\end{aligned} \tag{4.6}$$

Applying the expectation operator on (4.6) we get:

$$\begin{aligned}
E\left[|\omega_{async}|^2(k)\right] &= \sum_{l\neq i}\sum_{v=0}^{M}\left(|\Gamma_l|^2(v) + |\Gamma_l^{next}|^2(v)\right)|H_{il}|^2(k-v) P_l(k-v) \\
&+ \sigma^2_\omega(k) + C_s(k),
\end{aligned} \tag{4.7}$$

where P_l is the power transmitted on line l ($P_l = E[|X_l|^2]$), and C_s represents the reminder of equation (4.7), which is due to the cyclic prefixes effect S_l.
Writting $\Gamma_l(k)$ and $\Gamma_l^{next}(k)$ in function of m_l

$$\begin{aligned}
\Gamma_l(k) &= \sum_{v=0}^{M} \gamma_l \exp(-j2\pi vk/M) \\
&= \sum_{v=m_l+1}^{M} \gamma_l \exp(-j2\pi vk/M) \\
&= \sum_{v=0}^{M-m_l-1} \exp(-j2\pi(v+m_l+1)k/M) \\
&= \exp\Big(-j\frac{2\pi k}{M}(\frac{m_l+M}{2}+1)\Big)\frac{\sin\Big(\frac{\pi k(M-m_l-1)}{M}\Big)}{\sin\Big(\pi k/M\Big)} \qquad (4.8)
\end{aligned}$$

Using the same procedure we get

$$\begin{aligned}
\Gamma_l^{next}(k) &= \sum_{v=0}^{M} \gamma_l^{next} \exp(-j2\pi vk/M) \\
&= \sum_{v=0}^{m-L_{\mathrm{CP}}} \gamma_l^{next} \exp(-j2\pi vk/M) \\
&= \exp\Big(-j\frac{\pi k}{M}(m_l - L_{\mathrm{CP}} - 1)\Big)\frac{\sin\Big(\frac{\pi k(m_l-L_{\mathrm{CP}})}{M}\Big)}{\sin\Big(\pi k/M\Big)} \qquad (4.9)
\end{aligned}$$

4.3 Asynchronous Crosstalk Channel Gain Estimator

4.3.1 Correlated Power Allocation

To identify the crosstalk channel gain in the asynchronous case, we can imagine a similar scenario to the one proposed in chapter 3 where we can isolate the crosstalk caused by a particular user l during connection/disconnection of user l. The crosstalk caused by user l can be estimated by observing the SNR changes (equivalently the crosstalk changes), before and after user l get connected/disconnected as proposed in section 3.3.3.1. The estimate of the crosstalk caused by user l is given by:

$$\hat{\sigma}^2_{\omega_l,async}(k) = \sum_{v=0}^{M} A_{m,l}(v)|H_{il}|^2(k-v)P_l(k-v) + C_s(k) + e(k), \quad (4.10)$$

where $|H_{il}|^2(k)P_l(k)$ represents the crosstalk on tone k in the synchronous case, $A_{m,l}(v) = |\Gamma_l|^2(v) + |\Gamma_l^{next}|^2(v)$, $A_{m,l}(v)$ represents the inter-tone interference/leakage gain from the tone $(k-v)$ on the tone k due to non synchronization, $e(k)$ represents the estimation error at tone k. Fig.4.1 show $A(v)$ for a delay of non synchronization m_l equal to the half the length of a DMT symbol. Filter A will act as a low pass filter on the synchronous crosstalk, thus removing any sparks or rapid changes in the synchronous crosstalk vector: $\Big[|H_{il}|^2(0)P_l(0)$ ... $|H_{il}|^2(k)P_l(k)$... $|H_{il}|^2(K)P_l(K-1)\Big]$.
In fact from Fig.4.1 we can deduce that more than 92% of the asynchronous crosstalk at a tone k would be formed by a linear combination of the 5 adjacent synchronous crosstalk from tone $(k-2)$ to tone $(k+2)$.
Neglecting the effect of the cyclic prefix:

$$\hat{\sigma}^2_{\omega_l,async}(k) = \underbrace{\sum_{v=0}^{M} A(v)|H_{il}|^2(k-v)P_l(k-v)}_{circular\ convolution} + e(k), \quad (4.11)$$

where $e(k)$ is the estimation error on k. The crosstalk on tone k will be mainly caused by $|H_{il}|^2(k)P_l(k)$, and by $|H_{il}|^2(k_a)P_l(k_a)$ of the adjacent tones k_a. If we maintain the supposition that adjacent tones have a correlated crosstalk channel gain $|H_{il}|^2(k)$, and if we assume an equal or a highly correlated power allocation over the different tones, we would have:

$$\hat{\sigma}^2_{\omega_l,async}(k) \cong |H_{il}|^2(k)P_l(k) \quad (4.12)$$

The expression (4.12) is the same expression of the crosstalk caused by user l in the synchronized case, thus under these assumption we

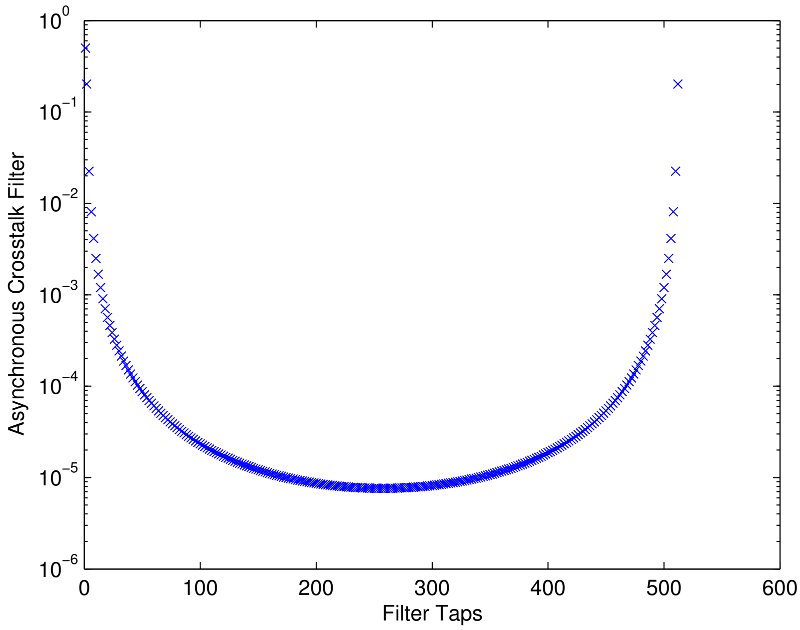


Figure 4.1: *Delay of 50% of the DMT symbole, logarithmic scale*

may use the same estimator proposed in chapter 3.

4.3.2 Uncorrelated Power Allocation

The assumption that the power allocation is constant or strongly correlated over all tones is not realistic, especially if DSM techniques are to be implemented. While allocated power can be strongly correlated among several numbers of tones, due to the correlation of the channel gains of these tones, the power allocation can suddenly change from one tone to another. This would happen when the global optimums of two adjacent tones belong to different regions. In this case, using the synchronized crosstalk method to estimate the crosstalk channels would increase the estimation error significantly.

Let $A_{m,l}$ represents the inter-tone interference gain filter for user l, $A_{m,l}$ is given as a function of the time delay m_l between user l and the user of interest (user i), let $\mathbf{M_{A_l}}$ be a circular matrix having $A_{m,l}$ as a first column, and let $\mathbf{D_{P,l}}$ be a diagonal matrix having the power allocation vector $P_l = [P_l(0) \ldots P_l(k) \ldots P_l(K)]$ as a diagonal. Equation (4.11) can be rewritten as:

$$\hat{\sigma}^2_{\omega_l,async} = \mathbf{M_{A_l}} \mathbf{D_{P,l}} |H_{il}|^2 + e, \tag{4.13}$$

where $\hat{\sigma}^2_{\omega_l,async}$, e, and $|H_{il}|^2$ represent vectors containing all the spectral elements of the crosstalk, the estimation error, and the crosstalk channel gain respectively, for example $|H_{il}|^2$ is the crosstalk channel gain vector: $|H_{il}|^2 = [|H_{il}|^2(1) \ldots |H_{il}|^2 k) \ldots |H_{il}|^2(K)]$.
The least square estimation of $|H_{il}|^2$ can be deduced from (4.13):

$$|\hat{H}_{il,m_l}|^2 = \left(\mathbf{D_{P,l}}^T \mathbf{M_{A_l}}^T \mathbf{M_{A_l}} \mathbf{D_{P,l}}\right)^{-1} \mathbf{D_{P,l}}^T \mathbf{M_{A_l}}^T \hat{\sigma}^2_{\omega_l,async}. \tag{4.14}$$

Both $\mathbf{D_{P_l}}$ and $\mathbf{M_{A_l}}$ are square matrices, equation (4.14) can be rewritten as:

$$|\hat{H}_{il,m_l}|^2 = \mathbf{D_{P_l}}^{-1} \mathbf{M_{A_l}}^{-1} \hat{\sigma}^2_{\omega_l,async}. \tag{4.15}$$

Since $\mathbf{M_{A_l}}$ is a circular matrix, $|\hat{H}_{il,m_l}|^2$ can be calculated in a fast way:

$$|\hat{H}_{il,m_l}|^2 = \mathbf{D_{P_l}}^{-1} \mathbf{W^H} diag\left(\mathbf{W} A_{m,l}\right)^{-1} \mathbf{W} \hat{\sigma}^2_{\omega_l,async}. \tag{4.16}$$

where $\mathbf{W}$ and $\mathbf{W^H}$ represent the FFT and IFFT matrices respectively, $diag\left(\mathbf{W} A_{m,l}\right)^{-1}$ is a diagonal matrix having $\mathbf{W} A_{m,l}$ as its diagonal. For $m_l = 0$ (for the synchronous case) the estimator given by equation (4.16) can be simplified to:

$$|\hat{H}_{il,m_l}|^2 = \mathbf{D_{P_l}}^{-1} \hat{\sigma}^2_{\omega_l,async}, \tag{4.17}$$

which has the same concept as the estimator proposed in chapter 3. From expression (4.15) we conclude that the mean square error between the asynchronous crosstalk model and the observed values $\hat{\sigma}^2_{\omega_l,async}$

is always zero and for all the possible values of m_l, as for all m_l we have:

$$\begin{aligned} &\left(\mathbf{M}_{\mathbf{A_l}}\mathbf{D}_{\mathbf{P_l}}|\hat{H}_{il,m_l}|^2 - \hat{\sigma}^2_{\omega_l,async}\right)^T \\ \times \quad &\left(\mathbf{M}_{\mathbf{A_l}}\mathbf{D}_{\mathbf{P_l}}|\hat{H}_{il,m_l}|^2 - \hat{\sigma}^2_{\omega_l,async}\right) = 0 \end{aligned} \tag{4.18}$$

From (4.18) follows that the LS crosstalk channel gain estimator (4.16) can be based on any possible delay value $\tilde{m}_l$ even if $\tilde{m}_l$ is far from the actual value m_l, as the mean square error will always be zero. We propose to estimate the crosstalk channel gain for the value $\tilde{m}_l = 0$, which means using the same estimator used for the synchronous case, for two reasons :

1. As it was shown in section 4.3.1, the synchronous crosstalk channel gain estimator works well over the tones were the power allocation is correlated, which is the case for most of the tones in DSL systems.
2. The synchronous crosstalk model is an exact model, thus we would not face an additional estimation error related to the modeling. This is not the case with the asynchronous crosstalk model, where the effect of the cyclic prefix were neglected.

Another crosstalk channel gain estimator can be proposed, if we consider that m_l is a random variable that can vary uniformly over $[0, M]$, the estimation of $|H_{il}|^2$ can be given simply by averaging over all possible values of m_l:

$$|\hat{H}_{il}|^2| = E\left[\hat{H}_{il,m_l}|^2\right] \tag{4.19}$$

Both estimators proposed in this section can be jointly put together, where the synchronous crosstalk channel gain estimator (4.17) would be given a large weight for tones that have a correlated power allocations on one hand, on the other hand and for tones having uncorrelated power allocation, the averaged asynchronous crosstalk channel gain estimator (4.19) would be privileged.

Since we know that the channel gain on the different tones are strongly correlated, especially among adjacent tones, we can reduce the estimation error further more using error smoothening techniques like the ones proposed in [60, 61] or simply by using the compression technique proposed in section 3.5.

4.4 Blind Synchronization

In this section we propose 2 methods to determine the value of the delay m_l between a user l and the line of interest i using blind techniques that do not require major changes on the current DSL systems and standards. These techniques exploit the fact that for an uncorrelated power allocation on adjacent tones, the non synchronization would actually smoothen the crosstalk observed over these different tones using the filter $A_{m,l}$. Thus the key points for a blind estimation of m_l is to observe the crosstalk variation among adjacent tones that have an uncorrelated power allocation. If the power allocation is correlated over all adjacent tones, we propose to chose a specific tone as pilot tone, and to vary the power allocation on this tone so that it would be noticeably different from its surrounding. This would put the pilot tone and its surrounding in an uncorrelated power allocation situation. Two $\hat{m}_l$ estimator may be proposed: A delay estimator that is based on a previous estimation of the crosstalk channel gain, and a joint delay and crosstalk channel gain estimator.

4.4.1 Delay Estimator based on the knowledge of Crosstalk Channel Gain

This method is a straight forward technique, let $|\tilde{H}_{il}|^2$ be the crosstalk channel gain estimator after smoothening and estimation error reduction, m_l is given as the value that would minimize the following least square expression:

$$\begin{aligned} \hat{m}_l \;=\; \operatorname{argmin}_{m_l} \quad & \left(\mathbf{M}_{\mathbf{A_l}}\mathbf{D}_{\mathbf{P_l}}|\tilde{H}_{il}|^2 - \hat{\sigma}^2_{\omega_l,async}\right)^T \\ \times \quad & \left(\mathbf{M}_{\mathbf{A_l}}\mathbf{D}_{\mathbf{P_l}}|\tilde{H}_{il}|^2 - \hat{\sigma}^2_{\omega_l,async}\right). \qquad (4.20) \end{aligned}$$

Where m_l affects the value of $\mathbf{M}_{\mathbf{A_l}}$ in equation (4.20). Testing different values of m_l over a discrete set $[0, M]$ will change $\mathbf{M}_{\mathbf{A_l}}$. In principle the

LS will be given for a value $\hat{m}_l$ which is close to the actual delay time m_l. In this case the LS is not zero, in contrast with (4.18), because we used an error reduction technique to smoothen the value of $|\hat{H}_{il}|^2$, thus equation (4.18) is not valid for the smoothened estimator $|\tilde{H}_{il}|^2$.

4.4.2 Joint Delay and Crosstalk Channel Gain Estimator

We can also imagine a joint estimation of the delay and of the crosstalk channel gain. In this case, and for all the possible values the non synchronization delay $\tilde{m}_l$, compute the crosstalk channel gain estimates using the estimator (4.16). For a value $\tilde{m}_l$ that is far from the actual delay value m_l, the estimation error around the pilot tone would be large, especially if the power on the pilot tone is much larger than its surroundings. Since the crosstalk channel gain are highly correlated among adjacent tones, the variance of the crosstalk channel gain estimates around the pilot tone would be extremely low in the absence of estimation error. A simple way to reject the non correct values $\tilde{m}_l$, is to observe the variance of the estimated crosstalk channel gains around the pilot tone. When $\tilde{m}_l = m_l$ this variance would be very low as the estimation error is small around the pilot tone in this case, and the adjacent pilot tones are almost equal, while the variance increases for $\tilde{m}_l$ getting further away from m_l. The estimator of m_l is now given by:

$$\hat{m}_l \;=\; \operatorname{argmin}_{m_l} \operatorname{Var}\left(\left[|\hat{H}_{il,m_l}|^2(k_p-2) \;...\; |\hat{H}_{il,m_l}|^2(k_p+2)\right]\right), \tag{4.21}$$

where Var represents the variance, k_p is the pilot tone, the variance of 5 tones was enough in practice.

4.5 Simulation Results

In this section, we observe the crosstalk variation on a DSL line i after the connection of a new user l. The PSD of the signal transmitted at line l is equal to -80 dBm/Hz over all tones with the exception of the pilot tone where a PSD of -72 dBm/Hz is transmitted. User l and line i are non synchronized, and the time delay between the useful signal and line i and the crosstalk caused by l is about 34% of the DMT symbol. The changes of the crosstalk before and after the connection of user l are shown in Fig.4.2. We can isolate the crosstalk part caused by the connecting user l by simply substracting the observed crosstalk before the connection of user l, and the observed crosstalk after the connection of l.

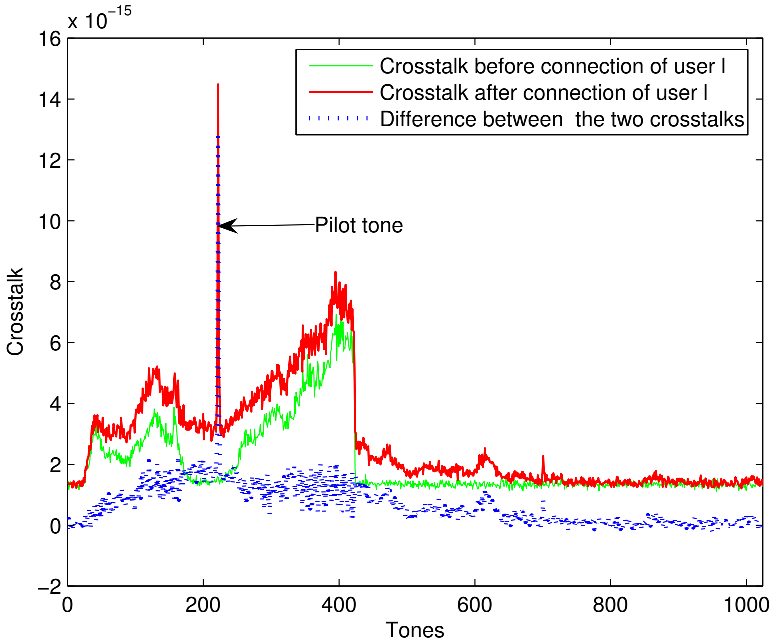


Figure 4.2: *Crosstalk before and after the connection of a new user*

Fig.4.3 shows the estimated crosstalk channel using the synchronized estimation and the non synchronized estimation for a known asynchronous delay time. We can see that the synchronized estimation works well for tones where the power allocation is correlated, however the synchronized estimation method did not give a good result for estimating the pilot tone's gain where the used power on this tone is uncorrelated to its surrounding. This is not the case for the non synchronized estimation, where the estimation is good around the pilot tone, and get worst for tones with correlated power allocation.

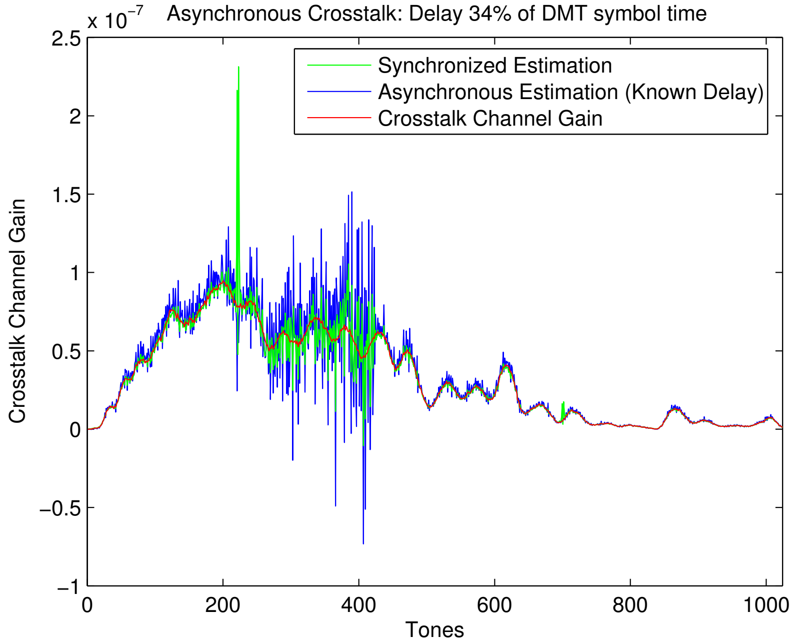


Figure 4.3: *Estimation of Asynchronous Crosstalk Channel Using Synchronized and Non Synchronized Models*

It is difficult to know in advance the delay time between the different DSL lines. In Fig.4.4 we show the estimation result obtained by averaging over all possible non synchronized models. Again, this method gives better results for the pilot tone and its neighbors when compared to the synchronized estimation method, while the synchronized method gives better results for tones having correlated power allocation.

To estimate the delay time, in the following we used the synchronized estimation method to obtain a first estimate of the crosstalk channel, then we used the smoothing method proposed in chapter 3 to reduce the estimation error around the pilot tone. The mean square error (MSE) between the crosstalk caused by the estimated channel and the observed crosstalk is checked for different delay time values. As

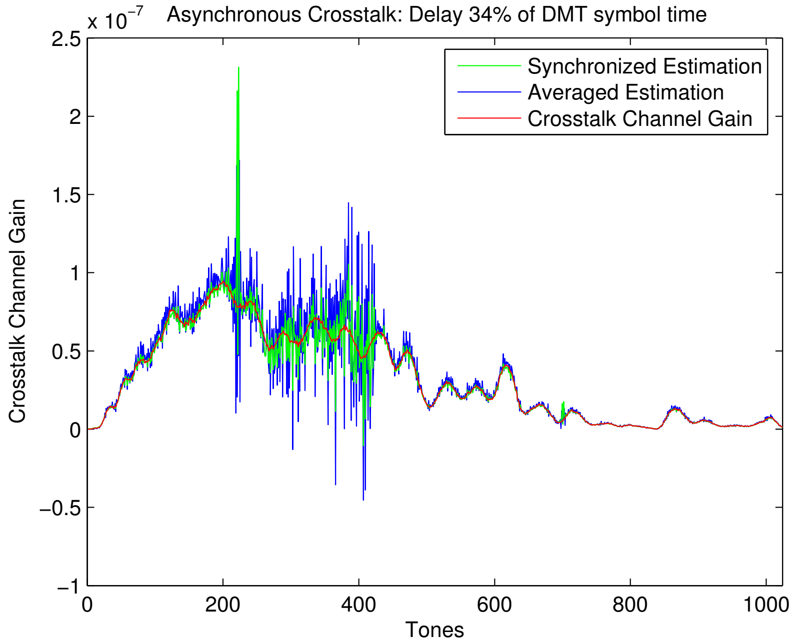


Figure 4.4: *Estimation of Asynchronous Crosstalk Channel by Averaging over all Non Synchronized Models*

shown in Fig.4.5, the minimum MSE corresponds to two values, the real delay time m_l which is $\hat{m}_{l1} = 0.34M$ of the DMT symbol time, and the second value that corresponds to $\hat{m}_{l2} = (1 - \hat{m}_{l1})M + L_{CP}$.
Fig.4.6 shows the variance of the estimated crosstalk channel of the pilot tone and its adjacent subchannels. The crosstalk channels are estimated using LS (4.16) for models corresponding to different delay times. Again we can see that the variance is minimal for the actual delay time $0.34M$ and for its symmetric opposite $0.66M + L_{CP}$.
Table 4.1 compare the absolute relative error $|\rho_{m_l}|$ obtained for the estimation of m_l using the minimal variance method and the minimal MSE method, from this table we can see that for this particular situation the minimal variance method outperformed the MSE method, furthermore the minimal variance method seemed to work better for small delay time, as $|\rho_{m_l}| = 0.004$ for $m_l = 0.1M$ using the minimal variance method, and it goes up to $|\rho_{m_l}| = 0.018$ for $m_l = 0.4M$. In

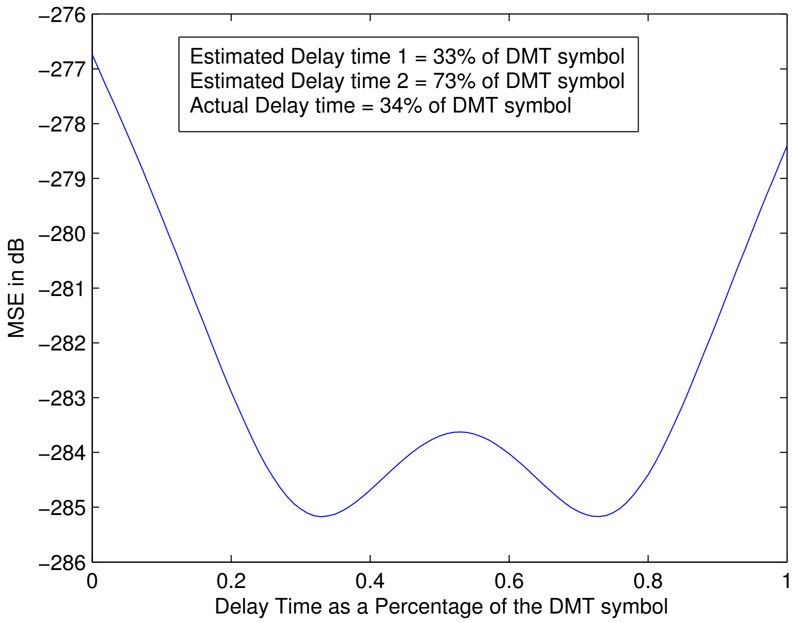


Figure 4.5: *MSE between Observed Crosstalk and Estimated Channel's Crosstalk for different Delay Time*

this simulation, $|\rho_{m_l}|$ was averaged over 10 estimations.

Delay Time	0.1M	0.2M	0.3M	0.4M
Minimum Variance ($\lvert\rho_{m_l}\rvert$)	0.004	0.0101	0.011	0.018
Minimum MSE ($\lvert\rho_{m_l}\rvert$)	0.0169	0.0128	0.0191	0.019

Table 4.1: *Comparison between Minimum Variance and Minimum MSE for Delay Time estimation*

4.6 Conclusion

In this chapter we adopted a more realistic approach for the current DSL systems by considering a non synchronized system. A crosstalk model for asynchronous lines was developed. This model was given

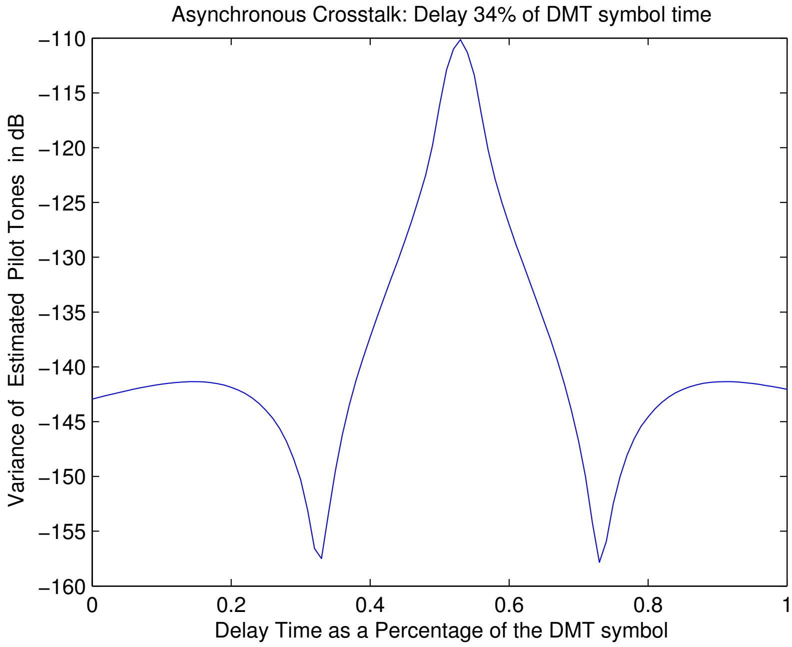


Figure 4.6: *Variance of the Estimated Channel around the Pilot Tone for different Delay Time*

in function of the delay time, and it was used to propose estimation procedures for the crosstalk channel under the assumption of non synchronization. Estimation techniques for the delay time were proposed as well. The effectiveness of the developed asynchronous crosstalk model and of the proposed estimation techniques is proved by simulations.

Chapter 5

Effect of Estimation Error on DSM

5.1 Introduction

There are few works in the literature dealing with the performance of DSM algorithms under the assumption of non perfect crosstalk channel estimation. Papers [62, 63] test the effect of imperfect knowledge of the crosstalk channels on the bitloading algorithms using simulation and actual implementation. In [64] the authors study the effect of crosstalk estimation on the total bitrate by deriving the probability density function of the bitrate in function of the estimation error, however in their analysis, they made the assumption that the power allocation is independent of the estimation error. In this chapter we will prove that the power allocation is actually dependent on the estimation error, where the error can affect the outcome of the optimization procedures. In fact, we will show that the main degradation in bitrate results when the estimation error leads the DSM algorithms to give a power allocation corresponding to a local optimum instead of the optimal power allocation that corresponds to the global optimum.

5.2 System Model

We consider a synchronized DSL system with imperfect knowledge of the crosstalk channel gains. The systems uses DMT modulation, where the different channels may be decomposed over K tones. Let $|H_{il}|^(2)(k)$ be the crosstalk channel between line l and line i for tone k, we define $\mathbf{G}(k)$ as a vector of variables representing all the elements of the different crosstalk channel gains at tone k, we consider two particular points of $\mathbf{G}(k)$:

- $\mathbf{G_t}(k) = [|H_{12}|^2(k)|H_{13}|^2(k)...|H_{il}|^2(k)...|H_{U(U-1)}|^2(k)]$ which represents the true values of the different crosstalk channels, U is the total number of users on the system.
- $\mathbf{G_e}(k) = [|\hat{H}_{12}|^2(k)|\hat{H}_{13}|^2(k)...|\hat{H}_{il}|^2(k)...|\hat{H}_{U(U-1)}|^2(k)]$ which contains all the estimates of the different crosstalk channel gains at tone k.

We have: $\mathbf{G_t}(k) = \mathbf{G_e}(k) + \mathbf{\Delta G}$, where $\mathbf{\Delta G}$ represents the estimation errors.

Let $\mathbf{P}(k)$ be a vector of variables that represents the power allocated for the different users at tone k where: $\mathbf{P}(k) = [P_1(k) \ ... \ P_i(k) \ ... \ P_U(k)]$, $P_i(k)$ is the power allocated for user i at tone k. We consider in particular the two optimal power allocations $\mathbf{P_t}(k)$ and $\mathbf{P_e}(k)$ associated with $\mathbf{G_t}(k)$ and $\mathbf{G_e}(k)$ respectively. $\mathbf{P_t}(k)$ and $\mathbf{P_e}(k)$ are related by $\mathbf{P_t}(k)$=$\mathbf{P_e}(k) + \mathbf{\Delta P_t}$.

We define the system's bitrate as the sum rate over all lines. The system's bitrate can be defined as a function of $\mathbf{G}(k)$ and $\mathbf{P}(k)$:

$$R\left(\mathbf{G}(k), \mathbf{P}(k)\right) = \sum_i R_i\left(\mathbf{G}(k), \mathbf{P}(k)\right) \tag{5.1}$$

where :

$$\sum_i R_i\left(\mathbf{G}(k), \mathbf{P}(k)\right) = log_2\left(1 + \frac{|H_{ii}|^2(k)P_i(k)}{\sum_{l \neq i} |H_{il}|^2(k)P_l(k) + \sigma_\omega^2(k)}\right) \tag{5.2}$$

where $\sigma_\omega(k)$ is the AWGN variance.

5.3 Effect of the Estimation Error on DSM

This section studies the effect of the estimation error on DSM centralized algorithms used for the DSL systems. We suppose that the optimization algorithm used for DSM is able to find the global optimum, this is particularly true for the OSB algorithm. In centralized DSM algorithms, the power constraints are incorporated within the objective function using Lagrange. This makes the optimization of the system's bitrate decoupled over tones for the inner loop, thus the optimization can be done over individual tones. Depending on the optimization problem, on each tone we may face one of the following cases: Dominant global optimum case, border points case, and on the presence of local optimums case.

5.3.1 Global Optimum

In the case of a dominant global optimum, due to the channel estimation error, the DSM algorithm would find a different power allocation then that found for error free estimations. However, since in this case there exists only one global optimum, the power allocated in the presence of estimation error $\mathbf{P_e}(k)$ stays in the same region of the error free power allocation $\mathbf{P_t}(k)$. We can find a relationship between $\mathbf{P_t}(k)$ and $\mathbf{P_e}(k)$ using Taylor expansion. If we expand the value of $R\left(\mathbf{G_t}(k), \mathbf{P}(k), k\right)$ around $\mathbf{G_e}(k)$ and $\mathbf{P_e}(k)$ we get:

$$\begin{aligned}
&R\left(\mathbf{G_t}(k), \mathbf{P}(k), k\right) = \\
&R\left(\mathbf{G_e}(k), \mathbf{P_e}(k), k\right) + \nabla^{\mathbf{T}} R_{\mathrm{G_e,P_e}} \left[\mathbf{\Delta G^T \Delta P^T}\right]^T + \\
&\frac{1}{2}\left[\mathbf{\Delta G^T \Delta P^T}\right] \nabla^{\mathbf{2}} R_{\mathrm{G_e,P_e}} \left[\mathbf{\Delta G^T \Delta P^T}\right]^T
\end{aligned} \tag{5.3}$$

where $\mathbf{P}(k){=}\mathbf{P_e}(k){+}\mathbf{\Delta P}$, $\nabla R_{\mathrm{G_e,P}}$ represents the gradient with respect to $\mathbf{G}(k)$ and $\mathbf{P}(k)$ calculated at $\mathbf{G_e}(k)$ and $\mathbf{P_e}(k)$:

$\nabla R_{\mathrm{G_e,P_e}} = [\nabla^{\mathbf{T}}_{\mathbf{G}} R_{\mathrm{G_e,P_e}} \;\; \nabla^{\mathbf{T}}_{\mathbf{P}} R_{\mathrm{G_e,P_e}}]^T$. where $\nabla_{\mathbf{G}} R_{\mathrm{G_e,P_e}}$ and $\nabla_{\mathbf{P}} R_{\mathrm{G_e,P_e}}$ are the gradient with respect to $\mathbf{G}(k)$ and $\mathbf{P}(k)$ respectively.
The expression $\nabla^{\mathbf{T}}$ represents the transpose of the gradient: $\nabla^{\mathbf{T}} =$

$(\nabla)^T$.

$\nabla^{\mathbf{2}} R_{\mathrm{G_e,P_e}}$ is the Hessian matrix calculated with respect to $\mathbf{G}(k)$ and $\mathbf{P}(k)$ at $\mathbf{G_e}(k)$ and $\mathbf{P_e}(k)$:

$$\nabla^{\mathbf{2}} R_{\mathrm{G_e,P_e}} = \begin{bmatrix} \nabla^{\mathbf{2}}_{\mathbf{G},\mathbf{G}} R_{\mathrm{G_e,P_e}} & \nabla^{\mathbf{2}}_{\mathbf{G},\mathbf{P}} R_{\mathrm{G_e,P_e}} \\ \nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{G}} R_{\mathrm{G_e,P_e}} & \nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{P}} R_{\mathrm{G_e,P_e}} \end{bmatrix}.$$

Deriving both sides of (5.3) with respect to $\mathbf{P}(k)$ we get:

$$\begin{aligned} &\nabla_{\mathbf{P}} R_{\mathrm{G_t,P}} = \nabla_{\mathbf{P}} R_{\mathrm{G_e,P_e}} \\ &+ \left[\nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{G}} R_{\mathrm{G_e,P_e}} \nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{P}} R_{\mathrm{G_e,P_e}}\right] \times \left[\mathbf{\Delta G^T \Delta P^T}\right]^T. \end{aligned} \tag{5.4}$$

Since $\mathbf{P_t}$ represents the optimal power allocation for an error free estimation, we have $\nabla_{\mathbf{P}} R_{\mathrm{G_t,P_t}} = 0$. Thus from (5.4) we get:

$$\mathbf{\Delta P_t} = -(\nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{P}} R_{\mathrm{G_e,P_e}})^{-1} \left(\nabla_{\mathbf{P}} R_{\mathrm{G_e,P_e}} + \nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{G}} R_{\mathrm{G_e,P_e}} \mathbf{\Delta G}\right) \tag{5.5}$$

where $\mathbf{P_t}(k) = \mathbf{P_e}(k) + \mathbf{\Delta P_t}$. If we replace $\mathbf{P}$ with the optimum power allocation found under the perfect knowledge of the crosstalk channel $\mathbf{P_t}$, equation (5.3) can be rewritten as:

$$\begin{aligned} &R\left(\mathbf{G_t}(k), \mathbf{P_t}(k)\right) = R\left(\mathbf{G_e}(k), \mathbf{P_e}(k), k\right) \\ &+\nabla^{\mathbf{T}}_{\mathbf{G}} R_{\mathrm{G_e,P_e}} \mathbf{\Delta G} + \frac{1}{2} \mathbf{\Delta G^T} \nabla^{\mathbf{2}}_{\mathbf{G},\mathbf{G}} R_{\mathrm{G_e,P_e}} \mathbf{\Delta G} \\ &+\frac{1}{2} \nabla^{\mathbf{T}}_{\mathbf{P}} R_{\mathrm{G_e,P_e}} \mathbf{\Delta P_t} + \frac{1}{2} \mathbf{\Delta P_t^T} \nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{G}} R_{\mathrm{G_e,P_e}} \mathbf{\Delta G}. \end{aligned} \tag{5.6}$$

The first 3 terms on the RHS of equation (5.6) represent a Taylor expansion of $R\left(\mathbf{G_t}(k), \mathbf{P_t}(k)\right)$ around $\mathrm{G_e}$, thus they can be replaced by $R\left(\mathbf{G_t}(k), \mathbf{P_e}(k)\right)$.
Let $\mathbf{\Delta R}(k) = R\left(\mathbf{G_t}(k), \mathbf{P_t}(k)\right) - R\left(\mathbf{G_t}(k), \mathbf{P_e}(k), k\right)$ be the degradation in bitrate due to the estimation error at tone k. The expected degradation around the optimal allocation $\mathbf{P_t}(k)$ can be shown to be

equal to:

$$
\begin{aligned}
&E\left[\mathbf{\Delta R}(k)\right] = \\
&-\frac{1}{2}(\nabla^{\mathbf{T}}_{\mathbf{P}} R_{\mathrm{G_e,P_e}})(\nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{P}} R_{\mathrm{G_e,P_e}})^{-1}\left(\nabla_{\mathbf{P}} R_{\mathrm{G_e,P_e}}\right) \\
&-\frac{1}{2}\mathbf{Tr}\Big((\nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{G}} R_{\mathrm{G_e,P_e}})^{T}(\nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{P}} R_{\mathrm{G_e,P_e}})^{-1}(\nabla^{\mathbf{2}}_{\mathbf{P},\mathbf{G}} R_{\mathrm{G_e,P_e}}) \\
&\times E\left[\mathbf{Diag}\left(\mathbf{\Delta G}\right)^{2}\right]\Big)
\end{aligned}
\tag{5.7}
$$

where $\mathbf{Diag}\left(\mathbf{\Delta G}\right)$ is a diagonal matrix having $\mathbf{\Delta G}$ as its diagonal. $E\left[\mathbf{Diag}\left(\mathbf{\Delta G}\right)^{2}\right] = \mathbf{Diag}\left(\mathbf{Var}\right)$, where $\mathbf{Var}$ is given by:

$$
\begin{aligned}
&\mathbf{Var}(k) = \\
&[\mathrm{Var}(|\hat{H}_{12}|^2)(k)...\mathrm{Var}(|\hat{H}_{il}|^2)(k)...\mathrm{Var}(|\hat{H}_{U(U-1)}|^2)(k)]^T
\end{aligned}
$$

5.3.2 Border Points

The above Taylor based approximation is valid only when the different elements of $\mathbf{P_t}(k)$ and $\mathbf{P_e}(k)$ lie within the linear region $[0\ P_{max}(k)]$, where $P_{max}(k)$ represent the power spectrum density (PSD) mask at tone k. When an element $P_{i,t}(k)$ of $\mathbf{P_t}(k)$ is equal to 0 or to $P_{max}(k)$, it generally means that $P_{i,t}(k)$ lies outside the linear region, thus the estimation error have no influence on the optimization procedure around $P_{i,t}(k)$, which implies that $P_{i,e}(k) = P_{i,t}(k)$. In this case, and in order to study the expected bitrate degradation using (5.7), we consider that all the elements that are equal to 0 or to $P_{max}(k)$ as constant, and we do the Taylor expansion around the elements that lies within the linear region. If $\mathbf{P_t}(k)$ and $\mathbf{P_e}(k)$ represent border points, ie all the elements are either equal to 0 or to $P_{max}(k)$, the expected bitrate degradation is null.

5.3.3 Local Optimums

The bitrate optimization of the DSL system is a non-convex optimization that present many local optimums. Expression (5.7) is a good indicator of the bitrate degradation in the presence of a dominant global optimum, where the modified optimum due to the estimation error is likely to be near the real global optimum. However, when the global optimum is not dominant, there is a chance that the estimation error would transform a local optimum into a global one, as it will be shown in this section.

Let $\mathbf{P_{1,t}}(k)$ and $\mathbf{P_{2,t}}(k)$ be the power allocations for a global and a local optimal, respectively. Let $\mathbf{P_{1,e}}(k)$ and $\mathbf{P_{2,e}}(k)$ be the modified power allocations due to the estimation error, where $\mathbf{P_{1,e}}(k)$ and $\mathbf{P_{2,e}}(k)$ are associated with $\mathbf{P_{1,t}}(k)$ and $\mathbf{P_{2,t}}(k)$, respectively because they lie within the same region. Now we can calculate the real optimums in function of the modified ones using expression (5.7):

$$R\left(\mathbf{G_t}(k), \mathbf{P_{1,t}}(k)\right) \approx R\left(\mathbf{G_e}(k), \mathbf{P_{1,e}}(k)\right) + \nabla_{\mathbf{G}}^{\mathbf{T}} R_{\mathrm{G_e,P_{1,e}}} \mathbf{\Delta G} \tag{5.8}$$

$$R\left(\mathbf{G_t}(k), \mathbf{P_{2,t}}(k)\right) \approx R\left(\mathbf{G_e}(k), \mathbf{P_{2,e}}(k)\right) + \nabla_{\mathbf{G}}^{\mathbf{T}} R_{\mathrm{G_e,P_{2,e}}} \mathbf{\Delta G} \tag{5.9}$$

where we limited expression (5.7) to the first two terms. This linear approximation is valid because the first order of Taylor expansion is obviously much greater than the second order expansion since it contains quadratic forms of the estimation error, and because $+\frac{1}{2}\nabla_{\mathbf{P}}^{\mathbf{T}} R_{\mathrm{G_e,P}} \mathbf{\Delta P} = 0$ for $\mathbf{P} = \mathbf{P_{1,e}}$ and $\mathbf{P} = \mathbf{P_{2,e}}$, since $\mathbf{P_{1,e}}$, $\mathbf{P_{2,e}}$ represent at least a local optimal power allocations for the system bitrate $R_{\mathrm{G_e,P}}$. Due to the estimation errors, sometimes the optimization algorithm would give $\mathbf{P_{2,e}}$ as an optimal power allocation instead of $\mathbf{P_{1,e}}$, this would happen when : $R\left(\mathbf{G_e}(k), \mathbf{P_{2,e}}(k)\right) > R\left(\mathbf{G_e}(k), \mathbf{P_{1,e}}(k)\right)$, thus the local optimum region of the original optimization problem would be modified into a global optimum region due to the error if the following inequation is maintained:

$$\left(\nabla_{\mathbf{G}}^{\mathbf{T}} R_{\mathrm{G_e,P_{1,e}}} - \nabla_{\mathbf{G}}^{\mathbf{T}} R_{\mathrm{G_e,P_{2,e}}}\right) \mathbf{\Delta G} \geq d(k) \tag{5.10}$$

where $d(k) = R\left(\mathbf{G}(k), \mathbf{P_{1,t}}(k)\right) - R\left(\mathbf{G}(k), \mathbf{P_{2,t}}(k)\right)$. The error $\left(\nabla_{\mathbf{G}}^{\mathbf{T}} R_{\mathrm{G_e,P_{1,e}}} - \nabla_{\mathbf{G}}^{\mathbf{T}} R_{\mathrm{G_e,P_{2,e}}}\right) \mathbf{\Delta G}$ follows a gaussian distribution $\mathcal{N}(0, S)$, where S is given by

$$S = \mathbf{Tr}\left(\mathbf{Diag}\left(\nabla_{\mathbf{G}}^{\mathbf{T}} R_{\mathrm{G_e,P_{1,e}}} - \nabla_{\mathbf{G}}^{\mathbf{T}} R_{\mathrm{G_e,P_{2,e}}}\right)^2 \mathbf{Diag}\left(\mathbf{Var}\right)\right)$$

The probability $P_{rb}(k)$ that $\mathbf{P_2}(k)$ is going to be chosen as the global optimum can be calculated now:

$$\begin{aligned} P_{rb}(k) &= \mathrm{Prb}\left(\left(\nabla^{\mathbf{T}}_{\mathbf{G}} R_{\mathrm{G_e,P_{1,e}}} - \nabla^{\mathbf{T}}_{\mathbf{G}} R_{\mathrm{G_e,P_{2,e}}}\right) \mathbf{\Delta G} \geq d(k) \right) \\ &= 0.5\left(1 - \mathrm{erf}\left(d(k)/\sqrt{2S}\right)\right) \end{aligned} \tag{5.11}$$

The expected bitrate degradation in the presence of a local optimun can be written as:

$$\begin{aligned} E\left[\mathbf{\Delta R_{lo}}(k)\right] &= P_{rb}(k)\left(d(k) + E\left[\mathbf{\Delta R_2}(k)\right]\right) \\ &+ (1 - P_{rb}(k))(E\left[\mathbf{\Delta R_1}(k)\right]) \end{aligned} \tag{5.12}$$

where $E\left[\mathbf{\Delta R_1}(k)\right]$, $E\left[\mathbf{\Delta R_2}(k)\right]$ represent the expected degradation in bitrate at the power allocations $\mathbf{P_{1,e}}$ and $\mathbf{P_{2,e}}$, respectively. They can be calculated using equation (5.7) if $\mathbf{P_{1,e}}$ and $\mathbf{P_{2,e}}$ lie in the linear region. For practical reasons we use:

$$d(k) = \left|R\left(\mathbf{G_e}(k), \mathbf{P_{1,e}}(k)\right) - R\left(\mathbf{G_e}(k), \mathbf{P_{2,e}}(k)\right)\right|$$

.

5.4 Numerical Results

In this part we study the effect of the estimation error on DSM algorithms, since our analysis is based on the assumption that the optimization algorithm is able to find the global optimum, we use OSB algorithm for the bitrate maximization. The objective function to be maximized is the DSL system bitrate defined as the sum of rates of all users. In the following we consider a DSL system based on two users. We use the theoretical twisted pair formula to model the direct channel gain of each line, and the worst case empirical formula to model the crosstalk channel gain. The lines considered in this DSL system have the lengths of 2.3 and 2.7 Km. The AWGN noise is supposed to have a PSD of -140 dBm/Hz over all tones and for all users. The total number of tones is 1024 tones. We consider a Gaussian estimation error that has a relative error of 10% ($\rho_{H_{iu}} \leq \pm 0.1$) for all tones with a probability $Prb(\rho_{H_{iu}} \leq \pm 0.1) = 85\%$. Fig.5.1 shows the actual bitrate degradation and the Bitrate degradation computed using Taylor expansion (equations (5.5),(5.6)) for a given estimation error. We can see that both curves are close enough. In this graph, the estimation error degradation lies in the vicinity of the global optimum, this is normal since we show only the first 100 tones of the DMT symbol, these tones correspond to a low frequency band in the used DMT symbol. For theoretical models used in this section for the direct and crosstalk channel gain, the SNR is high at low frequencies, thus we can apply the high SNR approximation at the first 100 tones and we can conclude the existance of a dominant global optimum [65].

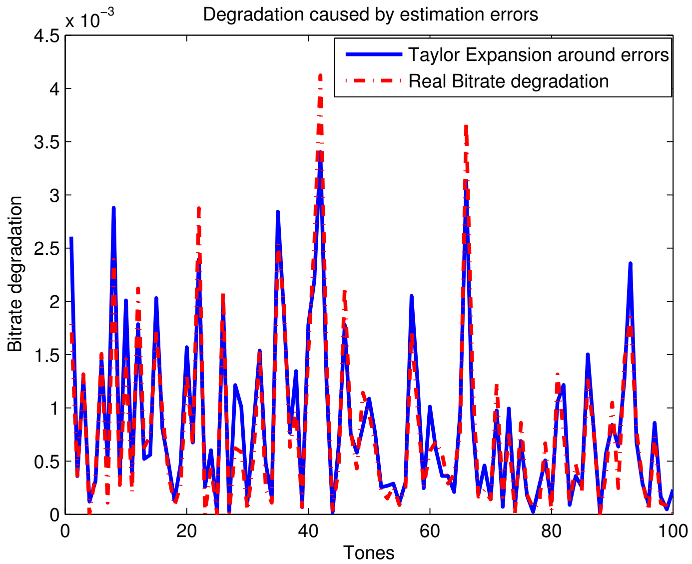


Figure 5.1: *Taylor Expansion vs Real Bitrate Degradation*

Fig.5.2 shows the expected bitrate degradation computed using equation (5.12), and the actual bitrate degradation averaged over 200 optimizations done using 200 different estimations with $(Prb(\rho_{H_{iu}} \leq \pm 0.1) = 85\%)$ for all tones. As explained previously, the expected bitrate degradation for the first 100 tones lies within the region of the dominant global optimal. For the frequency band that lies between tone 110 and tone 140, the optimization per tone is characterized by the presence of near-global optimal, so there is a possibility for the estimation error to modify the optimization problem so that the near-global optimal becomes a global one as explained in section 5.3.3. For tones that are higher than 140, the crosstalk channels become so strong that FDMA becomes the optimal strategy [66] thus power allocations for tones greater than 140 represent border points and the bitrate degradation is null (see section 5.3.2). Fig.5.3 shows the bitrate degra-

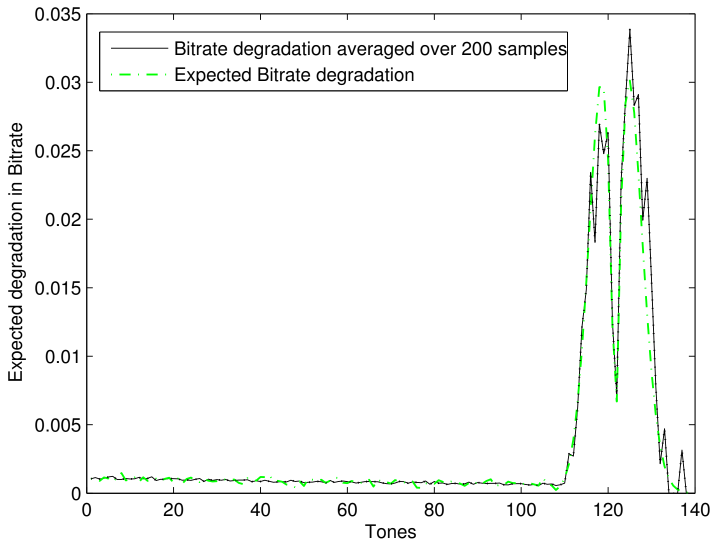


Figure 5.2: *Expected Bitrate Degradation vs Average Bitrate Degradation*

dation in function of relative error of crosstalk channel gain estimator ($\rho_{H_{iu}}$), as in the previous results, $\rho_{H_{iu}}$ has the same probability for all tones. From this figure we can conclude that the difference between the averaged bitrate degradation and the expected bitrate degradation over all tones is small for $\rho_{H_{iu}} < 0.1$, this difference increases as $\rho_{H_{iu}}$ increases, which is normal because Taylor approximation is not valid for large error. The overall bitrate degradation is small in this case (about 2 bits/DMT symbol for $Prb(\rho_{H_{iu}} \leq \pm 0.2) = 85\%$) this is due to the fact that the frequency band over which there exists a near-global optimum is limited by 30 tones (tones 110-140) out of 1024 available tones.

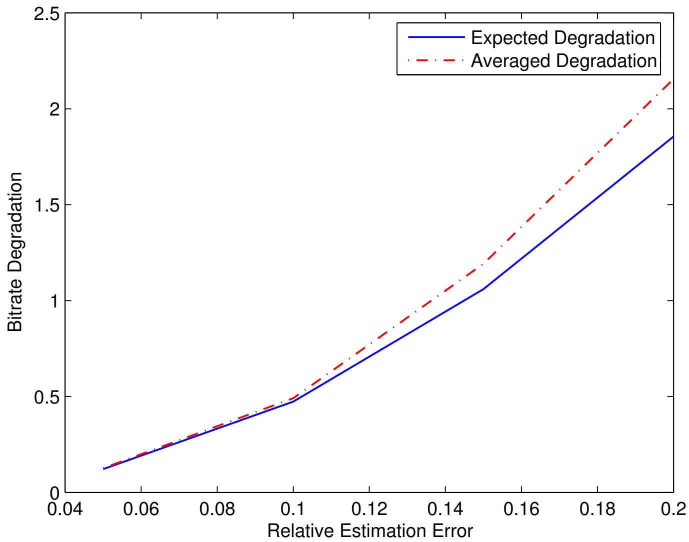


Figure 5.3: *Bitrate Degradation in Function of Relative Error of Crosstalk Estimator*

5.5 conclusion

In this chapter we studied the performance of DSM algorithms in the presence of estimation error, we have concluded that the estimation error mainly effect the optimization problems that have near-global optimums. The estimation error may change the optimization problem in a way that transforms a local optimum region into a global optimum one. This is particularly true when local optimum is near-global. We derived a probabilistic tool based on Taylor approximation, that can expect the degradation in bitrate due to the estimation error. The numerical results showed that this tool gives good results for estimations that have relative error that is less than 20%.

Chapter 6

Crosstalk Channel Estimation

6.1 Introduction

In order to satisfy the growing demand on higher bit-rates, the latest DSL systems like VDSL2 and VDSL2+ use higher bandwidths (up to 30 MHz) and shorter loops. The drawback of transmitting at high frequencies and using shorter loops is an increased crosstalk between the different DSL lines. Due to this increase in crosstalk, FEXT (far end crosstalk) is considered as the main degradation in these DSL systems. To maintain the promised high bit-rates, and to gain further in capacity, crosstalk cancellation or suppression is a must. For this reason, a number of precancellation techniques have been designed to decrease the effect of FEXT [67], [68], [69] in downstream, using the coordination at the CO (central office) and assuming no coordination at the receiver side, the CPE (customer premise equipment, i.e. the user's equipment).

All these precancellation schemes rely on a good estimation of the crosstalk channels between the various pairs of users (or equivalently pairs of lines). So the issue of crosstalk channel estimation has to be solved to be able to use those schemes. Downstream channel estimation appears to be a much more complicated task than the upstream channel estimation, and it has received some attention in the literature recently. In order to try to limit the quantity of overhead needed for feeding back the estimation, some methods have been proposed that only require to feedback the sign of the error samples (slicer errors) at the receiver [70]. The entire estimation processing is transferred at the central office. These sign-based methods exhibit slightly lower performance and slightly lower estimation speeds but allow a significant reduction of the overhead. In [71], it is proposed to simplify the precoder to its off-diagonal elements only, and an LMS tracking algorithm is proposed that converges to the optimal off-diagonal solution. This is essentially a pilot-based solution, that can also be used in a decision-directed mode, with the associated risk of error propagation.

Recently, it has been proposed in [72] to introduce a small perturbation on the transmitted signal, and to observe the effect on the apparent direct channels, seen from the receiver side. This method has the advantage of requiring a low overhead (only the direct channels need to be fed back), and requires minimal changes on the CPE. On another hand, the perturbation has an impact on the transmission of normal data, and has to be kept small enough to allow continuous transmission. Consequently, the changes on the direct channels may be very small and difficult to observe in some situations.

In this chapter we propose to induce a perturbation on the DSL line under investigation in order to change the crosstalk channel. Altering the real part of the crosstalk channel modifies its amplitude which is detectable by monitoring the SNR variations on the line. Applying the triangulation technique allows the estimation of the real crosstalk channel part normalized by the direct channel. The same procedure can be applied to extract the imaginary crosstalk channel part information. The estimated crosstalk channel will be normalized by the direct channel, however these informations are sufficient for a DSM level 3 signal precompensation.

6.2 System Model & available Information

Since no training sequences are to be used, and to stay in conformity with the different DSL standards, this article takes advantage of the available information under the current DSL systems to perform the crosstalk channels estimation. In this section we suppose the existence of a monitoring system. This system records information about the time of connection/disconnection of different users, the PSD level of the users' signals, the direct channels' gains, the SNR and the ambient noise for all the possible lines and at different tones. First we start with a brief model of the DSL system in section 6.2.1, then we formulate the estimator of the crosstalk channels in section 6.3. We analyse the performance of the estimator in section 6.4, and then in section 6.5 we provide a suitable precoding technique. Section 6.6 provides the simulation results of a typical 2 line DSL system, and in section 6.7 we conclude.

6.2.1 System Model

We consider DSL systems using DMT (discrete multiple tone). The channel maybe decomposed in K parallel sub-channels, where K represents the total number of tones. U is the total number of users. Under the assumption that all users are synchronized, the received signal Y of user i in sub-channel k is given by:

$$Y_i(k) = H_{ii}(k)X_i(k) + \sum_{l \neq i} H_{il}(k)X_l(k) + \omega(k). \tag{6.1}$$

Where X_i is the useful signal of user i, ω is the additive noise, $H_{ii}(k)$ is the direct channel gain of user i for tone k, and $H_{il}(k)$ is the crosstalk gain from line l to line i for tone k. If the number of interfering users is high, the additive crosstalk signal maybe considered as a Gaussian noise using the central limit theorem. We define the equivalent noise ω_e as:

$$\omega_e(k) = \sum_{l \neq i} H_{il}(k)X_l(k) + \omega(k). \tag{6.2}$$

The bitrate can be evaluated now by: $R_i(k) = B\log(1 + \mathrm{SNR}_i(k))$, the SNR is given by

$$\mathrm{SNR}_i(k) = \frac{|H_{ii}|^2(k)P_i(k)}{\sigma^2_{\omega_e}(k)} \tag{6.3}$$

where $\sigma^2_{\omega_e}(k) = \sum_{l \neq i} |H_{il}|^2(k)P_l(k) + \sigma^2_{\omega}(k)$, $P_i(k)$ is the power transmitted by line i. The AWGN variance at tone k is denoted by $\sigma^2_{\omega}(k)$, and the expression $\sum_{l \neq i} |H_{il}|^2(k)P_l(k)$ represents the crosstalk caused by all the interfering lines at the tone k.

6.2.2 The SNR Estimator

The main information that we have access to are the SNR, the PSD of the transmitted signal on each line, and the direct channels, since the crosstalk channel estimator depends mainly on these parameters, its interesting to study the behavior of the SNR estimator as performed at the receiver.
The direct channel is estimated during the training period, while the power is communicated between the modems prior to transmission.
The SNR estimator $S\hat{N}R(k)$ can be written as:

$$S\hat{N}R_i(k) = \frac{|H_{ii}|^2(k)P_i(k)}{\frac{1}{N}\sum_{n=1}^{N}|\omega_{e_n}(k)|^2} \tag{6.4}$$

where N is the total number of DMT symbol samples over which the SNR is estimated. As $|H_{ii}|^2(k)P_i(k)$ is supposed to be known, the study of the SNR leads to the study of the noise variance estimator, which is a sample variance as shown in (6.4).
The noise variance estimator can be derived from (6.4):

$$\begin{aligned}\hat{\sigma}^2_{\omega_e}(k) &= \frac{1}{N}\sum_{n=1}^{N}|\omega_{e_n}(k)|^2 \\ &= \frac{|H_{ii}|^2(k)P_i(k)}{S\hat{N}R(k)}.\end{aligned} \tag{6.5}$$

Under the assumption that ω_e follows a Gaussian distribution (which is asymptotically true when the number of interfering users is very large), the variance of $\hat{\sigma}^2_{\omega_e}$ can be computed in function of $\sigma^2_{\omega_e}$:

$$\begin{aligned}\mathrm{Var}\left(\hat{\sigma}^2_{\omega_e}\right) &= \sum_{n=1}^{N}\mathrm{Var}\left(\frac{|\omega_{e_n}(k)|^2}{N}\right) \\ &= \frac{2\sigma^4_{\omega_e}}{N}.\end{aligned} \tag{6.6}$$

Where we assume $E\left[|\omega_{e_n}(k)|^4\right] = 3\sigma^4_{\omega_e}$ (fourth moment of the central Gaussian variable $\omega_{e_n}(k)$).

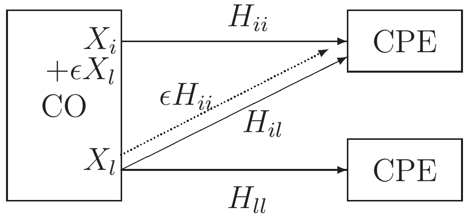


Figure 6.1: *DSL System during the Estimation of a Crosstalk Channel*

6.3 Tone-Wise Crosstalk Channel Estimation Based on SNR Perturbation

In this section we will propose a crosstalk channel estimation method based on an active observation of the DSL system. This method respects the current DSL standard, it imposes minimal changes on the CO side, and no changes at all on the CPE side. Unlike the passive observation method proposed in chapter 3, this new method is based on observing the SNR variation on a given line caused by inducing a perturbation on the crosstalk channel to be estimated.

The estimation procedure proposed in this section works on a tone by tone basis, it is able to extract both the real and the imaginary part of the crosstalk channel, not only the amplitude as in chapter 3.
Figure 6.1 represents the DSL system at a given tone k, during the estimation period. Line i is being the line under investigation, the system is concerned with estimating the crosstalk channel H_{il}.
Let the signal transmitted on line i be X_i, and the signal transmitted on line l be X_l. Thus the received signal at the CPE of line i at tone k would be:

$$Y_i(k) = H_{ii}(k)X_i(k) + H_{il}(k)X_l(k) + \omega_e(k).$$

To estimate the crosstalk channel, we propose to perturb the crosstalk channel H_{il} significantly enough to affect the resulting SNR. The perturbation may be done by adding an additional signal ϵX_l at the line i. This additional signal can be seen either as a sort of precoding, or as adding a virtual crosstalk channel ϵH_{ii} between line l and line i.

The new received signal at the CPE of line i is:

$$Y_{i,e}(k) = H_{ii}(k)X_i(k) + H_{ii}(k)\epsilon(k)X_l(k) + H_{il}(k)X_l(k) + \omega_e(k),$$

resulting in a modified crosstalk channel between l and i which equal to:

$$H_{e,il}(k) = H_{il}(k) + H_{ii}(k)\epsilon(k). \tag{6.7}$$

Where $\epsilon(k) = A(k)\exp(j\phi)$.

The tone wise crosstalk channel estimation can be described as procedure with 3 phases:

1. Phase 1:

   - $\epsilon = 0$
   - Use (6.5) to estimate $\sigma^2_{\omega_t}$ the total crosstalk power on line i, where:
     $$\sigma^2_{\omega_t}(k) = |H_{il}|^2(k)P_l(k) + \sigma^2_{\omega_e}(k)$$
     where P_l and $\sigma^2_{\omega_e}$ are the power associated with X_l and ω_e respectively.
     Let $S\hat{N}R_t$ be the measured SNR associated with this stage we have:
     $$\hat{\sigma}^2_{\omega_t}(k) = \frac{|H_{ii}|^2(k)P_i(k)}{S\hat{N}R_t(k)}.$$

2. Phase 2:

   - Set $\phi = 0$ thus $\epsilon(k) = A(k)$
   - Using (6.5), estimate the resultant noise variance $\sigma^2_{\omega_c}$ related to the modified crosstalk channel $H_{c,il}(k) = H_{il}(k) + H_{ii}(k)A(k)$ where:

     $$\sigma^2_{\omega_c}(k) = |H_{c,il}|^2(k)P_l(k) + \sigma^2_{\omega_e}(k).$$

So if $S\hat{N}R_c$ is the measured SNR during phase 2, we have:

$$\hat{\sigma}^2_{\omega_c}(k) = \frac{|H_{ii}|^2(k)P_i(k)}{S\hat{N}R_c(k)}.$$

- Estimate the expression:

$$\delta_{H_{c,il}}(k) = \left(|H_{c,il}|^2(k) - |H_{il}|^2(k)\right) / |H_{ii}|^2(k)$$

where:

$$\delta_{H_{c,il}}(k) = \frac{\sigma^2_{\omega_c}(k) - \sigma^2_{\omega_t}(k)}{|H_{ii}|^2(k)P_l(k)}$$

To the estimate $\hat{\delta}_{H_{c,il}}$, we replace $\sigma^2_{\omega_c}$ and $\sigma^2_{\omega_t}$ by their respective estimates:

$$\begin{aligned} \hat{\delta}_{H_{c,il}}(k) = & \frac{P_i(k)}{P_l(k)} \\ & \times \left(\frac{1}{S\hat{N}R_c(k)} - \frac{1}{S\hat{N}R_t(k)} \right). \end{aligned} \tag{6.8}$$

- Use the triangle formula to estimate $\Re\left(H_{il}(k)/H_{ii}(k)\right)$ (look at Fig.6.2):

$$\Re\left(\hat{H}_{il}(k)/H_{ii}(k)\right) = \frac{\hat{\delta}_{H_{c,il}}(k)}{2A(k)} - A(k)/2. \tag{6.9}$$

3. Phase 3:
   - Set $\phi = \frac{3\pi}{2}$ thus $\epsilon = -jA$
   - Using (6.5), estimate the resultant noise variance $\sigma^2_{\omega_s}$ related to the modified crosstalk channel $H_{s,il}(k) = H_{il}(k) - jH_{ii}(k)A(k)$ where:

$$\sigma^2_{\omega_s}(k) = |H_{s,il}|^2(k)P_l(k) + \sigma^2_{\omega_e}(k).$$

- Use equation (3.12) to estimate

  $$\delta_{H_{s,il}}(k) = \left(|H_{xs_{il}}|^2(k) - |H_{x_{il}}|^2(k)\right) / |H_{ii}|^2(k)$$

  where:

  $$\delta_{H_{s,il}}(k) = \frac{\sigma^2_{\omega_s}(k) - \sigma^2_{\omega}(k)}{|H_{ii}|^2(k) P_l(k)}.$$

  Again, if $S\hat{N}R_s$ is the measured SNR, the estimate of $\delta_{H_{s,il}}$ can be given as:

  $$\begin{aligned}\hat{\delta}_{H_{sil}}(k) = & \frac{P_i(k)}{P_l(k)} \\ & \times \left(\frac{1}{S\hat{N}R_s(k)} - \frac{1}{S\hat{N}R_t(k)} \right).\end{aligned} \tag{6.10}$$

- Use the triangle formula to estimate $\Im\left(H_{il}/H_{ii}\right)$ (see Fig.6.2)

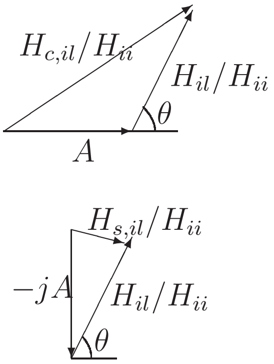
  $$\Im\left(\widehat{H}_{il}(k)/H_{ii}(k)\right) = A(k)/2 - \frac{\hat{\delta}_{H_{s,il}}(k)}{2A(k)}. \tag{6.11}$$

In the above procedure, ϵ is chosen to be either real or complex, thus it is changing one of H_{il} components at a time, which allows to estimate it, overall for each tone two perturbations are required to make the estimation. The value of A is chosen to be equal $|H_{il}|/2$, this value limits the SNR loss during the estimation.

6.4 Estimator Performance

Assuming that the estimation error of $1/S\hat{N}R_t$ follows a Gaussian distribution, we can use the equations (6.5) and (6.6) to estimate the variance $\mathrm{Var}\left(1/S\hat{N}R_t\right)$:

$$\mathrm{Var}\left(1/S\hat{N}R_t(k)\right) = \frac{2\sigma^4_{\omega_t}(k)}{N|H_{ii}|^4(k)P_i^2(k)}. \tag{6.12}$$



$$\Re(H_{il}/H_{ii}) = |H_{il}/H_{ii}|\cos\theta = \frac{(|H_{c,il}|^2 - |H_{il}|^2)/|H_{ii}|^2 - A^2}{2A}$$

$$\Im(H_{il}/H_{ii}) = |H_{il}/H_{ii}|\sin\theta = \frac{A^2 - (|H_{il}|^2 - |H_{s,il}|^2)/|H_{ii}|^2}{2A}$$

Figure 6.2: *Triangulation Technique to Compute the Imaginary and Real Part of the Crosstalk Channel*

Equation (6.12) can be used to estimate $\text{Var}\left(1/S\hat{N}R_c\right)$ and $\text{Var}\left(1/S\hat{N}R_s\right)$.

The difference of two Gaussian variables has a variance equal to the sum of variances of these two random variables, thus we have:

$$\text{Var}\left(1/S\hat{N}R_c - 1/S\hat{N}R_t\right) = \text{Var}\left(1/S\hat{N}R_c\right) + \text{Var}\left(1/S\hat{N}R_t\right) \tag{6.13}$$

and

$$\text{Var}\left(1/S\hat{N}R_s - 1/S\hat{N}R_t\right) = \text{Var}\left(1/S\hat{N}R_s\right) + \text{Var}\left(1/S\hat{N}R_t\right) \tag{6.14}$$

We can use equations (6.12),(6.13) and (6.14) to approximate the variances of $\hat{\delta}_{H_{c,il}}$ and $\hat{\delta}_{H_{s,il}}$ from which we can deduce the variances of the crosstalk channel's real and imaginary part estimators:

$$\text{Var}\left(\Re(\widehat{H}_{il}(k))/H_{ii}(k)\right) = \frac{\sigma^4_{\omega_c}(k) + \sigma^4_{\omega_t}(k)}{2NP_l^2(k)A^2(k)|H_{ii}|^4(k)}. \tag{6.15}$$

$$\text{Var}\left(\Im(\widehat{H}_{il}(k)/H_{ii}(k))\right) = \frac{\sigma^4_{\omega_s}(k) + \sigma^4_{\omega_t}(k)}{2NP_l^2(k)A^2(k)|H_{ii}|^4(k)}. \tag{6.16}$$

The estimation quality depends on the power of the interfering line l, on the number of DMT samples used in the estimation N, and on the amplitude A of the virtual crosstalk channel ϵ. If A is to be chosen too small, the quality of the estimation will be poor, however if A is too high, the line under investigation will suffer from high crosstalk.

6.5 Precoding

In DSM level 3, At each tone k the DSL system is considered as a MIMO system, thus the received signal may be given as:

$$\mathbf{Y}_k = \mathbf{H}_k \mathbf{X}_k, \tag{6.17}$$

where:

- $\mathbf{Y}_k$ is the received signals vector: $\mathbf{Y}_k = [Y_1(k) \ldots Y_i(k) \ldots Y_U(k)]$
- $\mathbf{X}_k$ is the transmitted signals vector: $\mathbf{X}_k = [X_1(k) \ldots X_i(k) \ldots X_U(k)]$
- $\mathbf{H}_k$ is the channel matrix:

$$\mathbf{H}_k = \begin{pmatrix} H_{11}(k) & H_{12}(k) & \ldots & H_{1U}(k) \\ H_{21}(k) & H_{22}(k) & \ldots & H_{2U}(k) \\ . & . & \ldots & . \\ . & . & \ldots & . \\ . & . & \ldots & . \\ H_{U1}(k) & H_{U2}(k) & \ldots & H_{UU}(k) \end{pmatrix}$$

We propose to use the following precoder at the transmitter:

$$\mathbf{X'}_k = \left(\mathbf{D}_{\mathbf{H}_k}^{-1} \mathbf{H}_k\right)^{-1} \mathbf{X}_k. \tag{6.18}$$

Where $\mathbf{D}_{\mathbf{H}_k}^{-1}$ is a diagonal matrix such that: $\mathbf{D}_{\mathbf{H}_k}^{-1}(i,i) = H_{ii}(k)$. Thus in order to make the precompensation of the crosstalk, we only need the off diagonal elements of $\mathbf{H}_k$ (crosstalk channels) normalized by the diagonal elements (direct channels), which is provided using our proposed estimation procedure.

6.6 Simulation

We test the estimation procedure proposed in this chapter on a 2×2 DSL system, the direct and crosstalk channels of this system are taken from measurement done on 400 m France Telecom (FT) cables. In this section we will present the estimation and precoding results obtained for one of these two lines. Fig.6.3 represents the estimation of the normalized crosstalk channels of the line under consideration. The estimation of the SNR was done over 100 DMT symbols, and for a Background noise level of -80 dBw/Hz. During the estimation a 16 QAM constellation was used over all tones.

Fig.6.4 represents the carrier to interference (CIR) ratio before and after precoding on the line. The average precoding gain is about 10 dB/tone in term of the CIR. Fig.6.5 compares the interference free bitloading, the bitloading under interference with no precompensation, and the bitloading under interference when the precompensation method was applied using our estimation results. Most of the tones had at least a bitloading gain of 1 bit, only few tones did not present a bitloading gain.

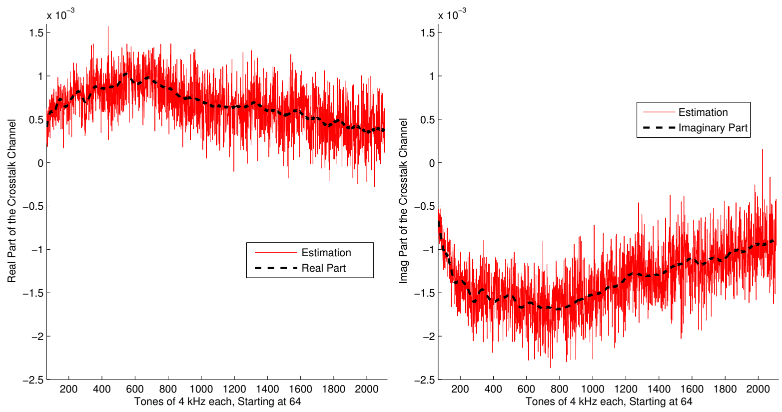


Figure 6.3: *Real and Imaginary Crosstalk Channel Estimation*

Table 6.1 compare the bitrates of the line when the precoding is done using different estimation results. We vary number of DMT symbols

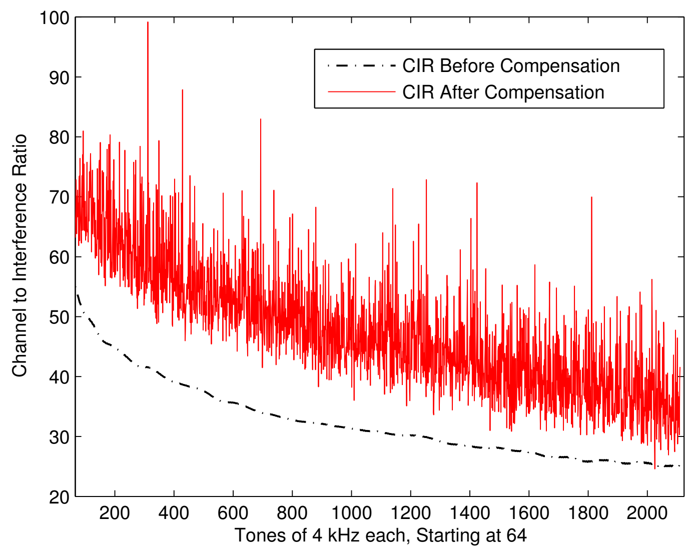


Figure 6.4: *Carrier to Interference Ratio*

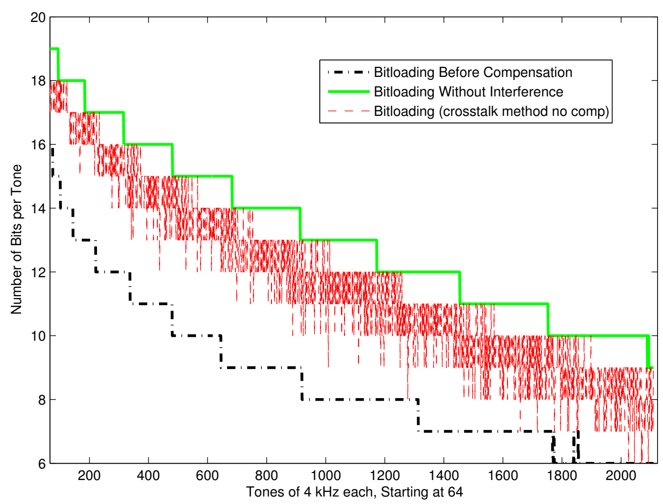


Figure 6.5: *Bitloading*

N required for the SNR estimation. From this table we can see that we have a precoding gain starting from $N = 10$.

Number of DMT symbols N	5	10	50	100	150	200
Bitloading (with Interference)	16743	16743	16743	16743	16743	16743
Bitloading (Interference Free)	25886	25886	25886	25886	25886	25886
Bitloading (Compensated)	16441	17978	21115	22095	22452	22695

Table 6.1: *Bitloading in Function of number of DMT Symbols used for Estimation*

6.7 conclusion

In this chapter we presented several optimized crosstalk channel estimation methods that do not require the use of a pilot sequence. These methods are based on precoding the transmitted signal and observing the SNR variation on different tones. The proposed estimation method requires minor changes in the current DSL systems, it is completely in conform with the current DSL user modems as the estimation procedure is implemented on the CO side only. A precoding method that only uses the off diagonal components normalized by the diagonal components was also presented. This precoding scheme is perfectly adapted to the proposed estimation method, as the estimated crosstalk channels are normalized by the direct channel of the victim line. The simulations showed the effectiveness of the proposed estimation and precoding methods.

Chapter 7

Optimized Crosstalk Channel Estimation

7.1 Introduction

The choice of the induced perturbation used for the crosstalk estimation remains an intriguing task, in chapter 6 the perturbation was chosen in a way to limit the SNR degradation during the estimation time by limiting the amplitude of the virtual crosstalk to the half of the crosstalk channel gain, in [73] an adaptive perturbation is proposed in order to limit the remaining interference level after crosstalk cancellation to a certain bound, however the same limitation on the perturbation amplitude is maintained.

In this chapter we start by deriving constraints on the perturbation power to limit the SNR degradation to 3 dB. This limitation is mainly concerned by limiting the degradation of the capacity, for the line under investigation, during the estimation time. However, practical implementation of the above method has shown [74] that the estimation time can be small compared to the time duration between estimations, especially as channels change very slowly in DSL lines. So another metric can be more appropriate for the choice of the perturbation power. Thus we proceed by relaxing the 3 dB SNR constraints, and we propose to choose the perturbation that optimizes an aggregated sum of the degraded bit-rates obtained during the estimation time and the improved bit-rates obtained after performing crosstalk cancellation (time duration estimations).

Another contribution of this chapter is to incorporate a time domain model for the crosstalk channel in the estimation procedure. In fact the above estimation techniques suppose that the sub-channels at different tones are independents, thus the estimation must be implemented two times over each tone, one time to estimate the real part of the crosstalk channel, and another time to estimate the imaginary part. However in DSL systems, tones (especially adjacent ones) are highly correlated. This correlation comes from the fact that the channel can be represented in the time domain by a finite impulse response with a limited number of taps. Estimating the time domain taps may reduce the estimation procedure by half, and limit the total number of tones to be estimated. It will be shown later on in this chapter that the time domain estimation necessitates only one amplitude perturbation, unlike the triangulation that requires two amplitude perturbations one on the real part and another on the imaginary part. The time domain model can be estimated by a limited number of tones (determined by the total number of taps).

This chapter will be divided as follow: Section 7.2 reviews the performance of the crosstalk estimation proposed in chapter 6 and then provides constraints to limit the SNR degradation to 3 dB. Section 7.3 compares the proposed crosstalk channel estimation method with other known estimation techniques. Section 7.4 will improve the crosstalk channel estimation based on the SNR changes by optimizing an aggregated sum of bitrates obtained during the estimation and after channel pre-compensation. Section 7.5 integrates the time domain model into the estimation technique, section 7.6 extends the estimation procedure for a multi line scheme, section 7.7 puts the theoretical approach under testing via simulations, and finally we conclude in section 7.8.

7.2 Constrained Perturbation

In this section we will limit the perturbation used in the active estimation of the crosstalk channels so that the predicted SNR changes would be limited to 3 dB loss. First we will review the estimator performance as derived in chapter 6, then we will propose a constraint on the choice of the perturbation A.

In chapter 6 we used a Gaussian approximation in order to estimate the variances of the crosstalk channel's real and imaginary part estimators:

$$\begin{aligned}\mathrm{Var}\left(\Re(\hat{H}_{il}(k)/H_{ii}(k))\right) &= \frac{\sigma^4_{\omega_c}(k)+\sigma^4_{\omega_t}(k)}{2NP_l^2(k)A^2(k)|H_{ii}|^4(k)}. &(7.1)\\ \mathrm{Var}\left(\Im(\hat{H}_{il}(k)/H_{ii}(k))\right) &= \frac{\sigma^4_{\omega_s}(k)+\sigma^4_{\omega_t}(k)}{2NP_l^2(k)A^2(k)|H_{ii}|^4(k)}. &(7.2)\end{aligned}$$

It is clear from these equation that A should be high in order to get good estimations, however if A is chosen too high, the line under investigation will suffer from high crosstalk, and the bitrate will drop severely during the estimation. One way of choosing A is to limit the SNR loss during estimation to less than 3 dB.

Let $\theta_{il}(k)$ be the phase related with the crosstalk channel H_{il} where we have $H_{il} = |H_{il}|\exp(j\theta_{il}(k))$. If we consider the phase θ_{il} as a random variable that varies uniformly on $[0, 2\pi]$. The crosstalk + noise variance with respect to θ_{il} will be:

$$\begin{aligned}E_{\theta,X,\omega}\Big[&\left(\left(H_{il}(k)+\epsilon(k)H_{ii}(k)\right)X_l(k)+\omega_e(k)\right)\\ &\times\left(\left(H_{il}(k)+\epsilon(k)H_{ii}(k)\right)X_l(k)+\omega_e(k)\right)^*\Big]\\ &=E_\theta\left[\sigma^2_{\omega_c}(k)\right]=E_\theta\left[\sigma^2_{\omega_s}(k)\right]\\ &=\sigma^2_{\omega_t}(k)+A^2(k)|H_{ii}|^2(k)P_l(k),\end{aligned} \tag{7.3}$$

where $\sigma^2_{\omega_t}(k) = |H_{il}|^2(k)P_l(k) + \sigma^2_{\omega_e}(k)$ is the crosstalk before we start the estimation procedure, and $\epsilon(k) = A(k)$ for the real part estimation and $\epsilon(k) = -jA(k)$ for the imaginary part.

The received SNR on the line under investigation during the estimation period may be predicted now as:

$$\mathrm{SNR}_e = \frac{|H_{ii}|^2(k)P_i(k)}{\sigma^2_{\omega_t}(k)+A^2(k)|H_{ii}|^2(k)P_l(k)}. \tag{7.4}$$

To limit the SNR degradation to 3 dB, A should fulfill the following constraint:

$$A(k) \leq \sqrt{\frac{\sigma^2_{\omega_t}(k)}{|H_{ii}|^2(k)P_l(k)}}. \tag{7.5}$$

Equation (7.3) can be used to predict the behavior of the crosstalk channel estimator prior to the estimation. Where we can simply replace the values of σ_{ω_c} and σ_{ω_s} in equations (7.1) and (7.2) by their expected value $\sigma^2_{\omega_t}(k) + A^2(k)|H_{ii}|^2(k)P_l(k)$, then we can evaluate the performance of the estimation in advance by predicting the crosstalk channel estimator variance using the following expression:

$$\begin{aligned}&\mathrm{Var}\Big(\hat{H}_{il}(k)/H_{ii}(k)\Big) = \\ &\quad \frac{\left(2\sigma^4_{\omega_t}(k) + A^4(k)|H_{ii}|^4(k)P_l^2(k) + 2A^2(k)|H_{ii}|^2(k)P_l^2(k)\sigma^2_{\omega_t}\right)}{NP_l^2(k)A^2(k)|H_{ii}|^4(k)}.\end{aligned} \tag{7.6}$$

7.3 Comparison to other Methods

In this section we will compare the performance of the crosstalk channel estimation based on the SNR method, with the classical method based on Pilot Symbols, and with the "channel abuse" method proposed in [72].

7.3.1 Pilot Symbols Method

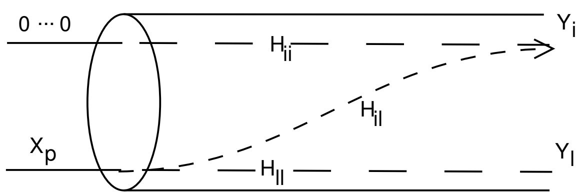


Figure 7.1: *DSL System in the Case of Pilot Symbols Estimation*

The "Pilot Symbols" method is resumed in Fig. 7.1. In order to estimate the crosstalk channel H_{il}, line i is put on quiet mode (no symbols are being transmitted there), while line l transmits known pilot symbols X_p to the receiver i. In this case the communication on both lines is halted. The received signal on line i, is given by:

$$Y_{ip,n}(k) = H_{il}(k)X_{p,n}(k) + \omega_e(k)$$

where $X_{p,n}$, and $Y_{ip,n}$, are the transmitted and the received signals for the DMT symbol n respectively. The estimation of crosstalk channel H_{il} can be given by:

$$\hat{H}_{il,p}(k) = \frac{1}{N_p}\sum_n \frac{Y_{ip,n}(k)}{X_{p,n}(k)} \tag{7.7}$$

where N_p is the total number of the DMT pilot symbols.

The variance of estimator $\hat{H}_{il,p}$ is given as:

$$\mathrm{Var}(\hat{H}_{il,p})(k) = \frac{\sigma^2_{\omega_e}(k)}{N_p P_l(k)} \tag{7.8}$$

where $X_{p,n}(k)$ are assumed to have the same power $P_l(k)$ for all DMT symbols.
In this case the bitrate loss during the estimation period is equal to the unused bitrate of lines i and l:

$$R_{L,p} = N_p(\log_2(1 + \frac{1}{\Gamma}\mathrm{SNR}_i) + \log_2(1 + \frac{1}{\Gamma}\mathrm{SNR}_l)) \tag{7.9}$$

7.3.2 Channel Abuse Method

Fig. 7.2 explains the principle of the crosstalk estimation using the "Channel Abuse" concept. In this case, the estimation of the channel H_{il} is done by inducing a small perturbation on the line l. The induced perturbation on line l is equal to ϵX_i. The perturbation is done in a similar way as proposed in this thesis, however instead of perturbing the crosstalk channel H_{il}, the term ϵX_i will affect the direct channel H_{ii}. An adaptive filter is assumed to track the modification on the

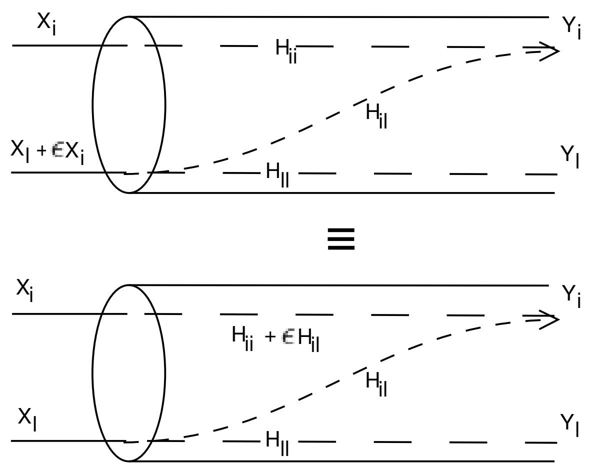


Figure 7.2: *DSL Systems in the Case of the Channel Abuse Estimation*

direct channel.

First, the direct channel H_{ii} estimator $\hat{H}_{ii,p}$ is obtained using a pilot symbol method similar to the estimation done in (7.7). The variance of the estimation is given by:

$$\mathrm{Var}(\hat{H}_{ii,p})(k) = \frac{\sigma^2_{\omega_t}(k)}{N_a P_i(k)} \tag{7.10}$$

where N_a in the number of the DMT pilot used for the channel estimation, $\sigma^2_{\omega_t}$ is crosstalk + background noise power, and P_i is the power on line i.

Second, perturb the signal X_2 in order to modify the direct channel H_{ii} as shown in Fig. 7.2, where the received signal is equal to:

$$Y_{ia,n}(k) = H_{ii,m}(k)X_{i,n} + \omega_t(k)$$

where $X_{i,n}$ and $Y_{ia,n}$ are the transmitted and the received signal of the DMT symbol n respectively, and the modified direct channel is given by: $H_{ii,m}(k) = H_{ii}(k) + \epsilon(k)H_{il}(k)$.

Third, use the hard decision $\hat{X}_{i,n}$ on Yia, n, in order to estimate $H_{ii,m}$ in a similar way to the pilot symbols method. For an estimation based on N_a DMT symbols, we have a probability of $(1-10^{-7})^{N_a}$ that $\hat{X}_{i,n} = X_{i,n}$ for all DMT symbols n (in DSL systems the probability of error is equal to 10^{-7}). In this case the variance of the modified direct channel is given by:

$$\mathrm{Var}(\hat{H}_{ii,m})(k) = \frac{\sigma^2_{\omega_t}(k)}{N_a P_i(k)} \tag{7.11}$$

The estimated crosstalk channel using the abuse method is equal to:

$$\hat{H}_{il,a}(k) = \frac{\hat{H}_{ii,m}(k) - \hat{H}_{ii,p}(k)}{\epsilon(k)} \tag{7.12}$$

The variance of $\hat{H}_{il,a}$ is given by:

$$\mathrm{Var}(\hat{H}_{il,a})(k) = \frac{2\sigma^2_{\omega_t}(k)}{N_a A^2(k) P_i(k)} \tag{7.13}$$

as before we have $A(k) = |\epsilon|(k)$.

In the case of the "channel abuse" method, the line l will suffer from an increased crosstalk due to the perturbation the same way as in the SNR method (same formulation), thus the epsilon should be small enough. Because ϵ is chosen to be small, and because $H_{il} << H_{ii}$, we can conclude that the SNR loss on line i is negligible.

The bitrate loss in this case is mainly due to the bitrate loss on line l caused by the SNR modification:

$$R_{L,a} = N_a(\log_2(1 + \frac{1}{\Gamma}\mathrm{SNR}_l) - \log_2(1 + \frac{1}{\Gamma}\mathrm{SNR}_{l,m})) \tag{7.14}$$

where $\mathrm{SNR}_{l,m}$ is the modified SNR on line l.

7.3.3 SNR Method

As explained earlier, in order to estimate H_{il} the SNR method induces perturbation on line i as shown in Fig. 7.3. This perturbation

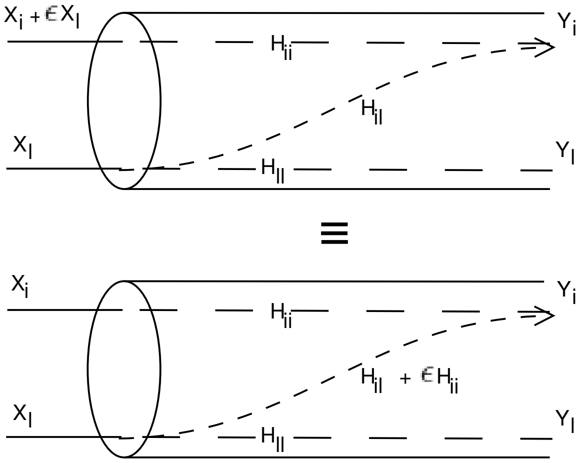


Figure 7.3: *DSL Systems in the Case of the SNR method Estimation*

causes the crosstalk channel to be modified. Unlike the channel abuse method, where the direct channel is modified.

When the modified SNR loss is limited to 3 dB, equations (7.1), (7.2) and (7.5) can be used to calculate the variance of the SNR method estimates:

$$\mathrm{Var}(\hat{H}_{il,\mathrm{SNR}})(k) = \frac{5\sigma^2_{\omega_t}(k)}{NP_l(k)} \tag{7.15}$$

where $\sigma^2_{\omega_s}(k) = \sigma^2_{\omega_c}(k) = A^2(k)P_l(k)|H_{ii}|^2(k) + \sigma^2_{\omega_t}(k) = 2\sigma^2_{\omega_t}(k)$.

The bitrate loss during SNR method is mainly due to the bitrate loss on line i caused by the SNR modification:

$$R_{L,\mathrm{SNR}} = N(\log_2(1 + \frac{1}{\Gamma}\mathrm{SNR}_i) - \log_2(1 + \frac{1}{\Gamma}\mathrm{SNR}_{i,m})) \tag{7.16}$$

where $\mathrm{SNR}_{i,m}$ is the modified SNR on line i. To be noticed that we only count the bitrate loss during the real part estimation. The bitrate loss during the imaginary part estimation can be compensated by using the estimated real part for a partial precoding of the crosstalk channel.

7.3.4 Comparision between the 3 methods

The pilot method and the SNR method would have the same performance if both estimators have the same variance. In this case we have:

$$\begin{aligned}\frac{N}{N_p} &= \frac{5\sigma^2_{\omega_t}(k)}{\sigma^2_{\omega_e}(k)} \\ &= 5 + \mathrm{SNR}_{i,p}(k) \end{aligned} \tag{7.17}$$

where $\mathrm{SNR}_{i,p}(k) = |H_{il}|^2(k)P_l(k)/\sigma^2_{\omega_e}(k)$ represent the SNR at line i using the "Pilot Method" estimation for H_{il}. In order for the SNR method and the "Pilot Method" to have the same performance, N should equal $(5 + \mathrm{SNR}_{i,p}(k))N_p$.

Fig. 7.4 shows the bitrate loss using the SNR method and the Pilot Symbols method, for different values of $\mathrm{SNR}_{i,p}$ and SNR_i. The values of N and N_p are chosen to have the same estimation performance using both methods. We can see that the SNR method can become advantageous in the case of low $\mathrm{SNR}_{i,p}$ and high SNR_i.

For the abuse method, if we assume that the modified SNR on line l is equal to 3 dB as well, we use the same procedure applied in the SNR method to get:

$$A^2(k)P_i(k)|H_{ll}|^2(k) = \sigma^2_{\omega_h}(k)$$

where $\sigma^2_{\omega_h}(k)$ represents the background noise and crosstalk power on line l.

The channel abuse estimation method and the SNR method would have the same performance when the following condition is achieved:

$$\begin{aligned}\frac{N}{N_a} &= \frac{5\sigma^2_{\omega_h}(k)}{|H_{ll}|^2(k)P_l(k)} \\ &= \frac{2.5}{\mathrm{SNR}_l}. \end{aligned} \tag{7.18}$$

We conclude from (7.18) that $N < N_a$, and that the bitrate loss in the SNR method is smaller than the bitrate loss in the abuse method, for $\mathrm{SNR}_l > 2.5$.

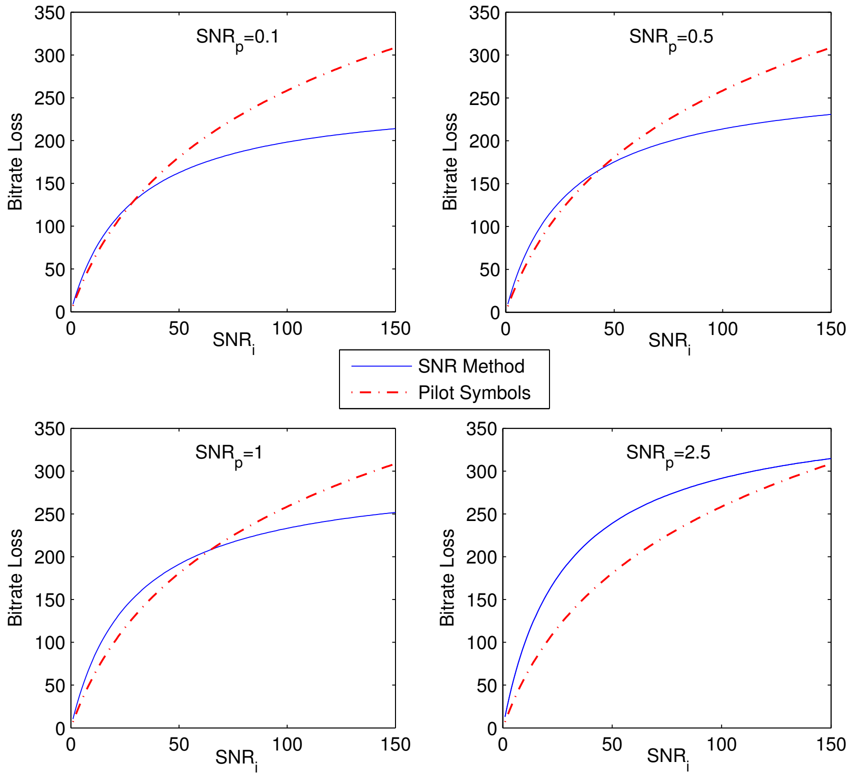


Figure 7.4: *Bitrate Loss Using SNR Method and Pilot Symbols Method*

7.4 Optimization of the Tone-Wise Crosstalk Channel Estimation

7.4.1 Optimizing the Estimation Technique

The value of A provided by equation (7.5) ensures the protection of the line under investigation during the estimation period, where it limits the SNR degradation to 3 dB maximum. However, this value may not be the optimum solution. Taking into the consideration that the estimation period is small, one may suggest that allowing the SNR degradation to be over 3 dB during the estimation period may increase the quality of the estimation. Thus the crosstalk cancellation after the estimation is done, may compensate for the SNR loss caused by a strong A. However a readaptation of the line for the time during estimation is suggested by adapting the line's bitrate to the predicted SNR loss, this readaptation of the bitrate prevent a line drop.
In the rest of this section we will present a crosstalk cancellation method that benefits from the estimation technique presented earlier, and then we will propose to choose A that optimizes an aggregated sum of user'rates during and after the estimation.

7.4.1.1 Successive Crosstalk Cancellation

For each line, the estimation must be repeated for all crosstalk channels caused by other interfering lines in the DSL system. Since each interfering line must be identified alone, one may propose a successive crosstalk cancellation procedure:

1. Choose a user i for crosstalk cancellation.
2. For the line under investigation i and for each tone of interest k:
    - Identify the line l that contribute the most on the crosstalk.
    - Apply the triangulation technique to estimate H_{il}.
3. Eliminate the identified crosstalk by precoding on line i.
4. Repeat steps (2,3) for all the remaining interfering lines.

5. Change i and go to step 2.

We can evaluate in advance the rate increase of user i after crosstalk cancellation. Let $\hat{H}_{il}(k)$ be the estimate of $H_{il}(k)$ at tone k, and let $e(k)$ be the estimation error $e_{il}(k) = H_{il}(k) - \hat{H}_{il}(k)$, the variance of e_{il} is the same as $\hat{H}_{il}(k)$ and it is given by equation (7.6). After precoding, the received signal at user i will be given as: $Y_{i,pc} = H_{ii}(k)X_i(k) + \omega_{il,r}(k)$. Where $\omega_{il,r}$ represents the remaining noise + crosstalk on the line after precoding: $\omega_{il,r}(k) = \underbrace{H_{il}(k)X_l(k) - \hat{H}_{il}(k)X_l(k)}_{e_{il}(k)X_l(k)} + \omega_e(k)$. The actual achievable bitrate after precoding is given in function of e_{il}: $R_{i,pc} = \sum_k R_{i,pc}(k)$, where at each tone k the actual bitrate after precoding is equal to:

$$R_{i,pc}(k) = \sum_k \log_2 \left(1 + \frac{1}{\Gamma} \frac{|H_{ii}|^2(k)P_i(k)}{(|e_{il}|^2(k)P_l(k) + \sigma^2_{\omega_e}(k))} \right). \tag{7.19}$$

The variance of e_{il} is known, thus we can predict the remaining noise + crosstalk power by replacing $|e_{il}|^2(k)$ by its expectancy :

$$\sigma^2_{\omega_{il,r}} = \text{Var}\left(\hat{H}_{il}(k)\right) P_l(k) + \sigma^2_{\omega_e}. \tag{7.20}$$

We can use equation (7.20) in order to predict user i's rate after precoding, $R_{i,pc}$ by: $\hat{R}_{i,pc} = \sum_k \hat{R}_{i,pc}(k)$, where $\hat{R}_{i,pc}(k)$ is the predicted rate at the tone k:

$$\hat{R}_{i,pc}(k) = \log_2 \left(1 + \frac{1}{\Gamma} \frac{|H_{ii}|^2(k)P_i(k)}{(\text{Var}\left(\hat{H}_{il}(k)\right) P_l(k) + \sigma^2_{\omega_e})} \right) \tag{7.21}$$

where Γ represents the SNR gap.

7.4.1.2 Joint Optimization of Users' Rates

As mentioned earlier the choice of $A(k)$ in equation (7.5) is based on limiting the SNR loss to 3 dB, in the following we propose to choose $A(k)$ that maximizes the user'rates during and after the estimation. After the crosstalk channel estimation, the crosstalk channel cancellation can be done, and the user rate after the estimation can be predicted by equation (7.21). We can apply the same approach to predict the line's rate during the estimation time. Let $R_{i,e}$ be the bitrate of the line during the estimation time, $R_{i,e}$ is given by $R_{i,e} = \sum_k R_{i,e}(k)$ where:

$$R_{i,e}(k) = \log_2\left(1 + \frac{1}{\Gamma}\frac{|H_{ii}|^2(k)P_i(k)}{(|H_{e,il}|^2(k)P_l(k) + \sigma^2_{\omega_e})}\right) \tag{7.22}$$

here $H_{e,il}$ is the modified crosstalk channel where $H_{e,il} = H_{c,il}$ or $H_{e,il} = H_{s,il}$.

The expectancy of noise and crosstalk variance during the estimation period $|H_{e,il}|^2(k)P_l(k) + \sigma^2_{\omega_e}$, is given by equation (7.3), thus we can predict the user rate $R_{i,e}(k)$ at tone k during the estimation as:

$$\hat{R}_{i,e}(k) = \log_2\left(1 + \frac{1}{\Gamma}\frac{|H_{ii}|^2(k)P_i(k)}{(A^2(k)|H_{ii}|^2(k)P_l(k) + |H_{il}|^2(k)P_l(k) + \sigma^2_{\omega_e})}\right), \tag{7.23}$$

The total predicted bitrate during the estimation is given by $\hat{R}_{i,e} = \sum_k \hat{R}_{i,e}(k)$.

Let A be a vector given by: $A = [A(1)...A(k)...A(K)]$. Now we can formulate the choice of A as an optimization problem of the following form:

$$\max_A \quad \alpha_1\hat{R}_{i,e} + \alpha_2\hat{R}_{i,pc}. \tag{7.24}$$

It is clear that optimization (7.24) is decoupled over tones, thus the optimization can be rewritten as:

$$A(k) \;=\; \max_{A(k)} \quad \alpha_1\hat{R}_{i,e}(k) + \alpha_2\hat{R}_{i,pc}(k). \tag{7.25}$$

The weighting factor α_1 represents the estimation time period, while α_2 is related to the period time between estimations. Normally the choice is $\alpha_1 << \alpha_2 < 1$ with $\alpha_2 = 1 - \alpha_1$. The optimization can be solved using a simple procedure such as searching for the optimum $A(k)$ over a given set of values ($[0, 2|H_{il}|]$ for example), this procedure is of low complexity since we are only searching over one parameter.

7.5 Time-model Crosstalk Channel Estimation Based on SNR Perturbation

In this section, we exploit the fact that in DSL systems a crosstalk channel may be represented by a time model with a limited number of taps. This fact will help reducing the number of the pilot tones and the time period over which the tone-wise estimation should be applied, further more it helps improving the quality of the estimation. In this section, first we will show how the time domain model may be incorporated with the tone-wise estimation, then we will optimize this estimation with respect to users' rate.

7.5.1 Time Domain Estimator

The estimation of the time domain model is based on the tone-wise estimation introduced earlier. Since the time domain model representing the crosstalk channel is usually formed by a limited number of taps L, that are usually less than the total number of tones ($L < K$), one can deduce a linear relationship between the different frequency estimates at different tones. This implies that the time domain model of the crosstalk channel may be estimated by using only one phase of the tone wise estimation over a limited number of pilot tones.
A small change is introduced to the tone wise estimation in this section, where $\phi(k)$ the phase of virtual crosstalk $\epsilon(k)$ at the tone k, will be allowed to have any values between 0 and 2π ($\phi(k) \in [0, 2\pi]$). In this case the modified crosstalk channel gain is given as:

$$\begin{aligned}|H_{e,il}|^2(k) &= |H_{il}(k) + \epsilon(k)H_{ii}(k)|^2 \\ &= (\Re(H_{il})(k) + A(k)|H_{ii}|(k)\cos(\phi(k) + \theta_{ii}(k)))^2 \\ &\quad + (\Im(H_{il})(k) + A(k)|H_{ii}|(k)\sin(\phi(k) + \theta_{ii}(k)))^2 ,\end{aligned} \tag{7.26}$$

where $\theta_{ii}(k)$ is the phase related to $H_{ii}(k)$. From equation (7.26) we can find a relation between the real and imaginary part of $H_{il}(k)$:

$$\begin{aligned}&\cos(\phi_m(k))\,\Re(H_{il})(k) + \sin(\phi_m(k))\,\Im(H_{il})(k) \\ &\qquad = \frac{|H_{e,il}|^2(k) - |H_{il}|^2(k) - A^2(k)|H_{ii}|^2(k)}{2A(k)|H_{ii}|(k)}\end{aligned} \tag{7.27}$$

where $\phi_m(k) = \phi(k) + \theta_{ii}(k)$. Let h_{il} be a vector of length L that represents the time domain model of the crosstalk channel from line l to line i. We can relate the frequency domain model to the time domain model using the following relationship:

$$H_{il} = \mathbf{W}h_{il}$$

where $\mathbf{W}$ is the FFT matrix.
The matrix $\mathbf{W}$ can be written as : $\mathbf{W} = \mathbf{W_r} + j\mathbf{W_i}$, where: $\mathbf{W_r} = \Re(\mathbf{W})$ and $\mathbf{W_i} = \Im(\mathbf{W})$.
Now we can rewrite the real and imaginary part of $H_{il}(k)$ as:

$$\Re(H_{il}) = \mathbf{W_r}h_{il}. \quad (7.28)$$
$$\Im(H_{il}) = \mathbf{W_i}h_{il}. \quad (7.29)$$

Replacing expressions (7.28), (7.29) in equation (7.27) and extending for all tones gives the following system:

$$D_A(\cos(D_\phi)\mathbf{W_r} + \sin(D_\phi)\mathbf{W_i})h_{il} = Z. \quad (7.30)$$

Where D_ϕ is a diagonal matrix having ϕ_m as a diagonal, D_A is a diagonal matrix having A as its diagonal, and Z is the observation vector:

$$Z(k) = \frac{|H_{e,il}|^2(k) - |H_{il}|^2(k) - A^2(k)|H_{ii}|^2(k)}{2|H_{ii}|(k)}.$$

Let matrix $\mathrm{M_W} = \cos(D_\phi)\mathbf{W_r} + \sin(D_\phi)\mathbf{W_i}$, using least square one may write an estimate for h_{il} as:

$$\hat{h}_{il} = (\mathrm{M_W}^H D_A^2 \mathrm{M_W})^{-1}\mathrm{M_W}^H D_A \hat{Z}, \quad (7.31)$$

where $\hat{Z}$ is the estimator of Z. Finally the frequency domain channel estimate $\hat{H}_{il}$ may be given as:

$$\hat{H}_{il} = \mathbf{W}\hat{h}_{il}. \quad (7.32)$$

7.5.1.0.1 Performance Evaluation Let $\sigma^2_{\omega_t}$ and $\sigma^2_{\omega_m}$ be the power of the noise + crosstalk before and after the crosstalk channel modification respectively. Let $\hat{\sigma}^2_{\omega_t}$ and $\hat{\sigma}^2_{\omega_m}$ be their respective estimates using (3.6), now $\hat{Z}$ components can be written as:

$$\hat{Z}(k) = \frac{1}{2}\frac{\hat{\sigma}^2_{\omega_m}(k) - \hat{\sigma}^2_{\omega_t}(k)}{P_l(k)|H_{ii}|(k)} - \frac{A^2(k)|H_{ii}|(k)}{2}. \tag{7.33}$$

We consider that the direct channel is perfectly known thus the variance of $\hat{Z}(k)$ is the same as the expression $\frac{\hat{\sigma}^2_{\omega_m}(k)-\hat{\sigma}^2_{\omega_t}(k)}{P_l(k)|H_{ii}|(k)}$ which may be given as:

$$\sigma^2_{\hat{Z}}(k) = \frac{\sigma^4_{\omega_t}(k) + \sigma^4_{\omega_m}(k)}{2NP_l^2(k)|H_{ii}|^2(k)}. \tag{7.34}$$

Again if we consider the phase θ_{il} as a random variable that varies uniformly on $[0, 2\pi]$. The variance with respect to θ_{il} can be evaluated the same way as in equation (7.3):

$$\sigma^2_{\hat{Z},\theta}(k) = \frac{2\sigma^4_{\omega_t}(k) + A^4(k)|H_{ii}|^4(k)P_l^2(k) + 2A^2(k)|H_{ii}|^2(k)P_l^2(k)\sigma^2_{\omega_t}(k)}{2NP_l^2(k)|H_{ii}|^2(k)}. \tag{7.35}$$

Let $\mathbf{C}_{\hat{Z}}$ be the covariance matrix of vector $\hat{Z}$, since each component is being estimated independently $\mathbf{C}_{\hat{Z}}$ is a diagonal matrix having $\sigma^2_{\hat{Z}}$ as its diagonal. Now we may write the covariance matrix of $\hat{H}_{il}$:

$$\begin{aligned}\mathbf{C}_{\hat{H}_{il}} &= \mathbf{W}(\mathrm{M_W}^H D_A^2 \mathrm{M_W})^{-1}\mathrm{M_W}^H D_A^2 \mathbf{C}_{\hat{Z}} \\ &\quad \times \mathrm{M_W}(\mathrm{M_W}^H D_A^2 \mathrm{M_W})^{-1}\mathbf{W}^H.\end{aligned} \tag{7.36}$$

We can evaluate the time domain estimation $\hat{H}_{il}$ in advance now using equation (7.36), as the variance of the time domain estimator is given by the diagonal of $\mathbf{C}_{\hat{H}_{il}}$.

7.5.2 Optimal Time Domain Estimation

This section searches for the optimal vector A that maximizes the aggregated sum of bit rates during the time domain estimation and after precoding. In the time-model case the tone-wise estimation can be done over a limited number of pilot tones that belongs to a finite set S. The tones that are not used in the estimation belong to $\bar{S}$. The actual rate during estimation is given in function of $H_{e,il}$:

$$\begin{aligned} R_{i,e_t} = &\sum_{k \in S} \log_2 \left(1 + \frac{1}{\Gamma} \frac{|H_{ii}|^2(k) P_i(k)}{(|H_{e,il}|^2(k) P_l(k) + \sigma^2_{\omega_e})}\right) \\ &+ \sum_{k \in \bar{S}} \log_2 \left(1 + \frac{1}{\Gamma} \frac{|H_{ii}|^2(k) P_i(k)}{(|H_{il}|^2(k) P_l(k) + \sigma^2_{\omega_e})}\right). \end{aligned} \tag{7.37}$$

The expression of the bit rate during estimation is very close to bit rate given in equation (7.23). The resemblance between the two expressions is normal since the two methods start by a tone-wise estimation.
We predict the bitrate during the estimation time by replacing $|H_{e,il}|^2(k)$ by its expectancy with respect to θ_{il}:

$$\begin{aligned} &\hat{R}_{i,e_t} = \\ &\sum_{k \in S} \log_2 \left(1 + \frac{1}{\Gamma} \frac{|H_{ii}|^2(k) P_i(k)}{(|H_{ii}|^2(k) A^2(k) P_l(k) + |H_{il}|^2(k) P_l(k) + \sigma^2_{\omega_e})}\right) \\ &+ \sum_{k \in \bar{S}} \log_2 \left(1 + \frac{1}{\Gamma} \frac{|H_{ii}|^2(k) P_i(k)}{(|H_{il}|^2(k) P_l(k) + \sigma^2_{\omega_e})}\right). \end{aligned} \tag{7.38}$$

After the estimation, assuming a crosstalk cancellation is applied like in section 7.4.1.1, the actual precoded bitrate has the same form as expression (7.19) and it can be written in function of $e_t = \hat{H}_{il} - H_{il}$:

$$R_{i,pc_t} = \sum_{k} \log_2 \left(1 + \frac{1}{\Gamma} \frac{|H_{ii}|^2(k) P_i(k)}{(|e_t|^2(k) P_l(k) + \sigma^2_{\omega_e})}\right). \tag{7.39}$$

one may predict the precoded bitrate as:

$$\hat{R}_{i,pc_t} = \sum_{k} \log_2 \left(1 + \frac{1}{\Gamma} \frac{|H_{ii}|^2(k) P_i(k)}{(\mathbf{C}_{\hat{H}_{il}}(k,k) P_l(k) + \sigma^2_{\omega_e})}\right). \tag{7.40}$$

Again A is chosen to optimize the following problem:

$$\max_A \quad \alpha_{t_1} \hat{R}_{i,e_t} + \alpha_{t_2} \hat{R}_{i,pc_t} \tag{7.41}$$

The weighting factor α_{t_1} and α_{t_2} represents respectively the estimation time period, and the time period between estimations. They are chosen such that $\alpha_{t_1} << \alpha_{t_2} < 1$ with $\alpha_{t_2} = 1 - \alpha_{t_1}$.

The optimization problem (7.41) is a non-convex optimization problem, and the different values of $A(k)$ are coupled between the different tones, so a simple line search per component is not a guaranteed solution as in problem (7.24).

To solve problem (7.41) a simple heuristic algorithm is proposed. For initial small values of $A = A^0$ a small increase of any A components will lead to increase $\hat{R}_{i,pc_t}$, this follows from the fact that at small A the expression $\mathbf{C}_{\hat{Z}}$ can be seen as independent from A ($A << \sigma_{\omega_t}$), thus $\mathbf{C}_{\hat{H}_{il}}$ is dominated by the matrix $(\mathrm{M_W}^H D_A^2 \mathrm{M_W})^{-1}$. The components of this matrix get smaller as A gets bigger.
On the other hand, increasing A will lead to the decrease of $\hat{R}_{i,e_t}$ as A represents the amplitude of the added virtual crosstalk. Due to the coupling between all the different tones in the $\hat{R}_{i,pc_t}$ case, increasing any of A components will have almost the same increasing effect, however in the $\hat{R}_{i,e_t}$ case the tones are decoupled, so the increase of the components of A will have different decreasing effect that varies among individual tones.

Based on the above observation, we propose an algorithm that takes into consideration only the decreasing effects on $\hat{R}_{i,e_t}$, this algorithm increases A in a way that minimizes the decrease of $\hat{R}_{i,e_t}$: Since in the case of $\hat{R}_{i,e_t}$ the tones are decoupled, one may work on minimizing the decrease of $\hat{R}_{i,e_t}$ by working on individual tones. If the gradient at the tone k $\partial\hat{R}_{i,e_t}/\partial A(k)$ is large, adding a large positive step on $A(k)$ will cause an important decrease of $\hat{R}_{i,e_t}$ while if $\partial\hat{R}_{i,e_t}/\partial A(k)$ is small, adding a large positive step on $A(k)$ will cause a limited decrease of $\hat{R}_{i,e_t}$. A simple way to do the optimization is to add a positive step to A that is inversely proportional to the gradient of $\hat{R}_{i,e_t}$. At the end of the iterations we may use a gradient algorithm for few steps to ensure that the calculated optimum is at least a local one. The optimization algorithm can be written as:

1. Initialize $A(k)$ by a small value $A^0(k)$
2. Compute $V^0 = \alpha_{t_1} R^0_{i_{c_t}} + \alpha_{t_2} R^0_{i_{c_t}}$
3. Go to the next iteration $t = t + 1$
4. Compute the gradient:
   $$\nabla R^t_{i_{e_t}} = [\partial R^t_{i_{e_t}} / \partial A^t(1) ... \partial R^t_{i_{e_t}} / \partial A^t(k) ... \partial R^t_{i_{e_t}} / \partial A^t(K)]^T$$
5. Update A using the formula
   $$A^{t+1}(k) = A^t(k) + s \frac{1}{\left| \partial R^t_{i_{e_t}} / \partial A^t(k) \right|}$$
   with s is a positive and small step
6. Compute $V^{t+1} = \alpha_{t_1} R^{t+1}_{i_{c_t}} + \alpha_{t_2} R^{t+1}_{i_{c_t}}$
7. if $V^{t+1} > V^t$ Go to 3, else decrease s and go to 5
8. Repeat while s is bigger than a threshold τ
9. for $s < \tau$ do some additional correction in the gradient direction

7.6 Multi-Line Identification

The single line crosstalk channel identification explained earlier may be extended to a multi-line identification, we may distinguish two cases: The case of one to many which corresponds to detecting crosstalk channels from one line to the other lines in the system, and the case of many to one where crosstalk channels from other lines in the system to one line are estimated.

7.6.1 One to Many

In this case we simultaneously precode all the lines of the system in order to estimate crosstalk channels coming from one line. The tone wise estimation or the time model technique may be used individually on each line of the DSL system and at the same time in order to estimate the crosstalk caused by a single line (Look at Fig.7.5).

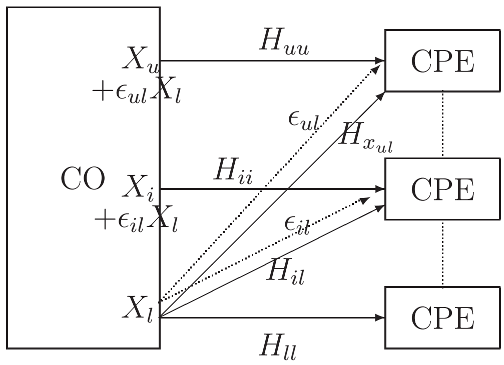


Figure 7.5: *Estimation of crosstalk channels caused by a single line on other DSL lines: The case of One-to-Many*

7.6.2 Many to One

In this case we estimate on a single line crosstalk channels coming from several lines in the DSL system (Fig.7.6). To be able to do the estimation of several crosstalk channels on a single lines and at the same time the time model estimation technique must be used: As already seen in section 7.5, when a time model is given for the crosstalk channel, the tone wise estimation may be done on a limited number of pilot tones, this can be seen as sampling in the frequency domain. Using this observation one can propose to estimate several crosstalk channels on different tones, and this by adding different precoders corresponding to different lines on different tones, the different tone wise estimation can be recombined using the time model technique. Fig.7.7 represents a precoded signal used for the tone wise estimation in the many to one case.

7.7 Simulation

In this section, the various algorithms introduced in this chapter are tested in the case of a single line identification. Fig.7.8 presents the direct channel of the line over which the estimation is taking place, the crosstalk channel to be estimated, and the remaining noise normalized by a constant PSD level of -80 dBw/Hz. This PSD is supposed to be the same for all users. The direct and crosstalk channels presented in

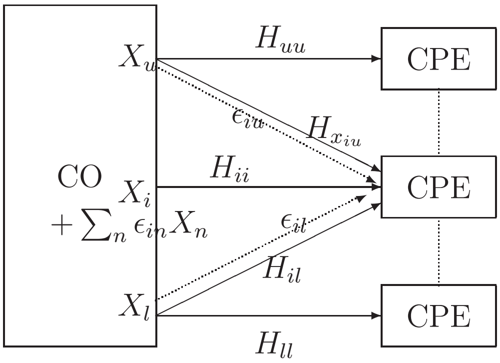


Figure 7.6: *Estimation of crosstalk channels caused by several DSL lines on a single line : The case of Many-to-One*

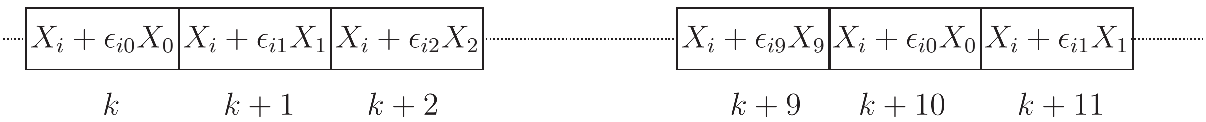


Figure 7.7: *Estimation in the case of Many-to-One: Different tones correspond to the estimation of different lines.*

Fig.7.8 are measured channels between 400 m France Telecom (FT) cables.

7.7.0.0.2 Time-domain method The evolution of the Time domain optimization with respect to iterations is presented in Fig.7.9. This plot is a numerical evaluation of the optimization algorithm presented in 7.5.2. Fig.7.9 represents the optimization procedure carried on prior to the estimation in order to choose the optimal value of A. This procedure is based on the predicted values of the bit-rates ($\hat{R}_{i,e_t}$ and $\hat{R}_{i,pc_t}$) and not on actual simulation. From the graph presented in Fig.7.9, one can see that the optimization started very slowly, this is due to choice of the initial values, where A^0 was chosen very small to insure that the algorithm could work (look at 7.5.2). After about 60 iterations, the algorithm finds a good startup A and the optimization accelerates with iterations. At the end, the loss of $\hat{R}_{i,e_t}$ is so great that it is no longer compensated by the gain in $\hat{R}_{i,pc_t}$, thus the optimizations slows down and stop.

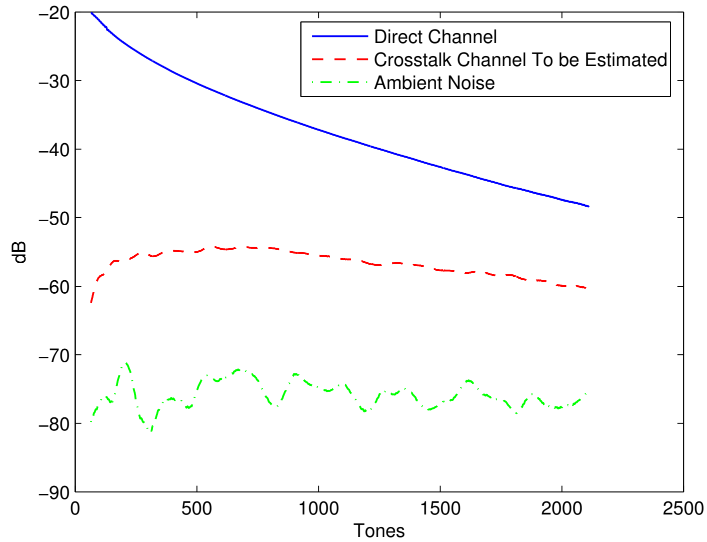


Figure 7.8: *The proposed Scenario*

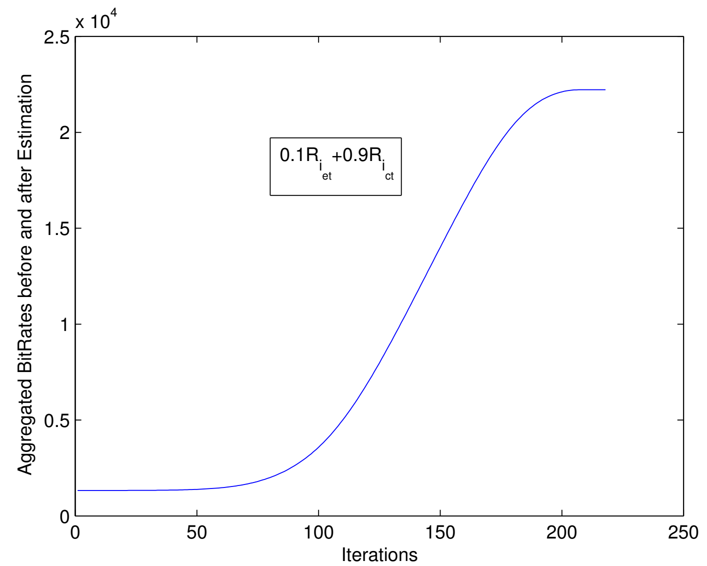


Figure 7.9: *Evolution of Time-domain Optimization with Iterations*

To test the concept of downsampling in which the estimations are taking place on a limited number of tones, Fig.7.10 shows the mean square error between the actual crosstalk channel and the estimated one for several downsampling ratios, the estimation is done over equidistant tones, with the distance between the pilot tones (used for estimation) is constant and is equal to the downsampling ratio. Fig.7.10 shows that we can go up to a downsampling ratio of 14 with a limited degradation in the estimation. The downsampling ratio is directly linked to the crosstalk channel to be estimated. If the crosstalk channel gain is smooth it can be represented by a time domain model with a limited number of taps thus the downsampling ratio can be high, if the crosstalk channel gain is highly uncorrelated, the time domain channel response is typically tall and the downsampling ratio is small.

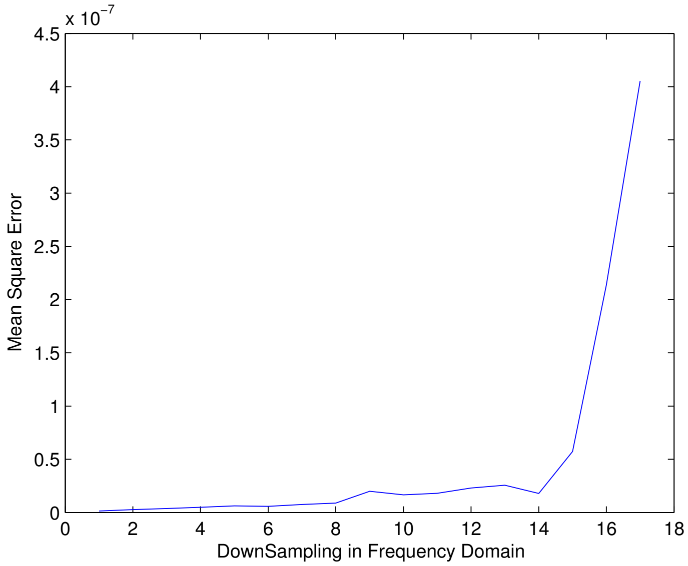


Figure 7.10: *Mean Square Error of the Estimated Channel in function of the Downsampling Ratio*

7.7.0.0.3 Comparison with the Tone-wise method This paragraph reports the comparison between the optimal time domain estimation, the optimal tone wise estimation, and the tone wise estimation with -3 dB maximum loss in SNR. The comparison is done using both the numerical results and the simulation results. Again, the numerical results represent a numerical evaluation of the optimization algorithms which are based on a predicted expression of the bit-rates. So after the optimization is done, the resultant optimal values A are fed back to the expressions (7.21), (7.23), (7.38), and (7.40) to get the numerical values the bit-rates. For the simulation results, the actual values of the crosstalk channel at the different tones are used to calculate the actual values of the bit-rates during estimation (7.22), and (7.37). For the bit-rates after the estimation, the optimal values A are used to simulate the estimation procedure, and then the estimated crosstalk channels are used to calculate the actual bitrates of the line after the crosstalk cancellation (7.19), and (7.39).

The time model estimation is used with an oversampling in frequency domain of (1/10) (only one tone out of 10 is used for the estimation, and the tones are chosen with an equal space). For the simulation a 16 QAM constellation is used on all tones, and an equal PSD of -80 dBw/Hz is given for all tones and for all users.

Fig.7.11 shows Bit-rate in function of the weighting factor α_e for the different estimation methods during the estimation period. The simulation results are shown to be better than the numerical ones, this is due to the pessimist approach used for the predicting of the different Bit-rates (the Gaussian distribution assumption).
Graphs shown in Fig.7.11 prove that limiting the loss of SNR to 3 dB maximum provides the ultimate protection for the line under investigation during the estimation period, as the line's bitrate during estimation time is the highest when 3 dB constraints are used. This is particularly true for small α_e (which is usually the case).

The numerical results predict that the time model estimation performs better than the tone-wise estimation in protecting the line capacity during estimation, which is intuitive because of the high downsampling ratio used in this particular test (1 out 10 tones is used). The simulation results did not prove this tendency, however both techniques, time model estimation and tone wise optimization, were very close from each others, and not far from the 3 dB SNR limitation.

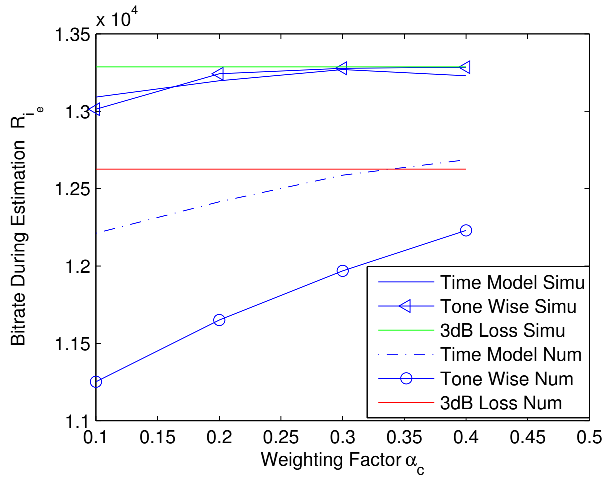


Figure 7.11: *Comparision between different methods for the estimation period*

Fig.7.12 shows Bit-rate in function of the weighting factor α_c for the different estimation methods during the estimation period. Once again, the numerical results predictions are too pessimists when compared to the simulation results. However this time both results give the same tendencies, where in both cases the tone wise optimization and the time model techniques outperformed the 3 dB limitation on the SNR. The simulation results show that the achievable bit-rate using crosstalk cancellation with the time model method did not change much with respect to α_c. The time model achieved bit-rate is around 25.9 kb/DMT symbol which means an 8.8 % increase over the 3 dB SNR limitation (achieved bitrate is 23.8 kb/DMT symbol). On the other hand, the tone wise optimization achieved bitrates between 24.2 and 24.5 kb/DMT symbol which is equivalent to an increase of 1.7 to 3 % over the 3 dB SNR limitation depending on the weighting factor. Based on the different results obtained so far, the time model technique has the edge over the tone wise techniques. The following list compares the advantages and inconvenient of the two techniques:

1. The time domain method outperformed the tone wise optimization, and the 3 dB SNR limitation, in crosstalk cancellation and achieved better rates (in both numerical and simulation results).

2. The time domain method achieved a comparable bitrate during the estimation period when compared to the 3 dB SNR limitation.

3. The time domain estimation needs only half the time required by the tone wise estimation: only one precoding is needed to achieve the time domain estimation unlike the tone wise estimation that requires two precoding (one for the imaginary part and one for the real part).

4. The time domain estimation can be done on a limited number of tones.

5. The time domain estimation allows the identification of several crosstalk channels at the same time.

6. The time domain estimation requires the full knowledge of the direct channel, this is not the case for the tone wise estimation.

7. The time domain optimization is complex when compared to the simple line search optimization used in the tone wise estimation.

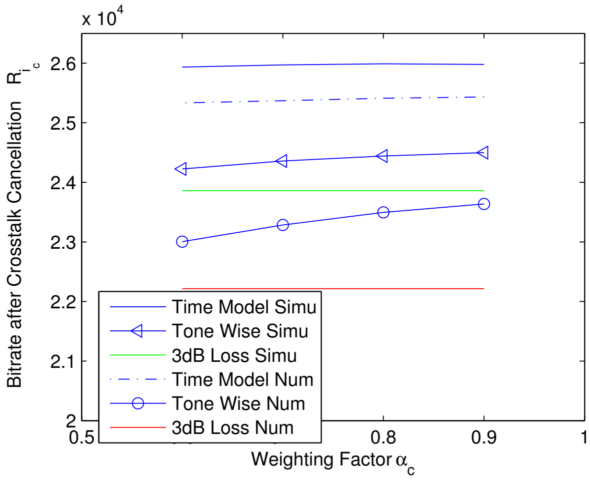


Figure 7.12: *Comparision between different methods after crosstalk cancellation*

7.8 conclusion

In this chapter we presented several optimized crosstalk channel estimation methods that do not require the use of a pilot sequence. The optimization was done to maximize the user rates during and after the estimation. A time domain model was also proposed. The time domain model estimation only requires few pilot tones. This allows the estimation of several crosstalk channels at the same time by allocating different pilot tones to different disturbers. The time domain model requires the full knowledge of the direct channel, while the current DSL system provides information on the direct channels of the lines, these information can have limited accuracy sometimes. The effect of direct channel inaccuracy on the time domain crosstalk estimation should be studied in future work.

Chapter 8

Channel Estimation Using few Pilot Tones

8.1 Introduction

The straight forward procedure to estimate the crosstalk channels DSL systems would be to use a set of training sequences, sent periodically, to perform the tracking of the downstream channels at the CPE (or upstream channels at the CO). Many solutions exist in the general framework of training sequences. An example of solution applicable to the VDSL system is analyzed in [75]. However, it requires to use part of the useful bit rate as pilot symbols. This causes important losses on the bit rate and some times it halts the communication on different lines.

In this chapter, we are interested in crosstalk channel estimation using training sequences. In order to limit the total bit rate loss caused by the pilot symbols, we propose to use a training signal that is formed in the frequency domain from a limited number of pilot tones. In this case, most of the tones are used for data transmission, and the communication between the end user and the CO is maintained, even during the estimation process. Unlike the procedure proposed in chapter 6, the pilot signal is sent on the line causing the disturbance and not on the victim lines, thus the SNR of the DSL system will be maintained the same during the estimation procedure (compared to the 3dB loss in chapter 6), and the only line that will be affected by the estimation is the line causing interference on the system. This can be advantageous when measuring the crosstalk channels from a newly connected line to an already established DSL system.

Using a limited number of adjacent pilot tones for estimation will result in an ill-conditioned problem, rendering the use of the classical LS algorithm non effective. To solve the conditioning problem, we add a regularization factor on the LS pseudo inverse matrix. In the case where the channel gain information are available, the regularization factor can be seen as the constraints on the channel gain values. When the channel gain information are not known, the regularization factor is a scaled identity matrix. The effect of the regularization can be reduced iteratively thus improving the overall estimation as the simulation results show.

This chapter will be organized as follow: Section 8.2 describes the DSL system. Section 8.3 explains the origin of the ill-conditioning problem in the time domain. Section 8.4 describes the algorithms used for the time domain estimation. Section 8.5 puts the theoretical approach under testing via simulations, and finally we conclude in section 8.6.

8.2 System Model

In this section, first we give the frequency domain model of a DSL line subject to crosstalk interference. Then we describe a time domain model of a pilot DMT signal with adjacent and limited tones. The resulting signal matrix is an ill-conditioned matrix that can not be used in least square estimation.

8.2.1 DSL System Model

We consider DSL systems using DMT. The channel may be decomposed in K parallel sub-channels, where K represents the total number of tones. Two main users are identified: the victim who communicates on the direct line d, and the disturber that causes the crosstalk u. Under the assumption that the two lines are synchronized, the received signal Y_d of user d in sub-channel k is given by:

$$Y_d(k) = H_{dd}(k)X_d(k) + H_{du}(k)X_u(k) + \omega(k) \tag{8.1}$$

where X_d is the useful signal of user d, X_u is the useful signal of user u, ω is the additive noise + the remaining crosstalks, $H_{dd}(k)$ is the direct channel gain of user d for tone k, and $H_{du}(k)$ is the crosstalk gain from line u to line d for tone k. If the number of interfering users is large, the additive crosstalk signal maybe considered as a gaussian noise using the central limit theorem. We define the equivalent noise ω_e as: $\omega_e(k) = H_{du}(k)X_u(k) + \omega(k)$. The capacity can be expressed now by the Shannon formula:

$$R_d = \sum_k \log_2 \left(1 + \frac{|H_{dd}|^2(k)P_d(k)}{\sigma^2_{\omega_e}(k)}\right) \tag{8.2}$$

where $\sigma^2_{\omega_e}(k) = |H_{du}|^2(k)P_u(k)+\sigma^2_{\omega}(k)$, $P_d(k)$ and, $P_u(k)$ are the power transmitted by line d and u respectively. The variance associated with the background noise and the remaining crosstalks at tone k is denoted by $\sigma^2_{\omega}(k)$.

8.2.2 Time Domain Model of the Training Sequence

During the estimation period of the crosstalk channel from a line u into a direct line d, the signal X_u transmitted on line u will be split in two: the training signal and the data signal.

Let N be the total number of DMT blocks in the training and the data signals. For each block n, the vector u_{I_n} represents the training symbols to be sent over the different tones of the training signal. Tones that correspond to the data signals will be replaced by zeros.

To obtain the time domain signal, an IFFT is applied over each DMT symbol, then a cyclic prefix is added to each DMT block. Finally the time domain signal can be written as:

$$v_u = [u^T_{P_1}u^T_{D_1}, u^T_{P_2}u^T_{D_2}, ...u^T_{P_n}u^T_{D_n}...u^T_{P_N}u^T_{D_N}]^T \tag{8.3}$$

where u_{D_n} is the vector obtained by applying an IFFT over u_{I_n}

$$u_{D_n} = \mathbf{W}^H u_{I_n} \tag{8.4}$$

and $\mathbf{W}$ represents the FFT matrix. Vector u_{P_n} represents the cyclic prefix obtained from u_{D_n}. The cyclic prefix length is given by L while the length of u_{D_n} is given by K. The total length of u is $N_t = N(K + L)$. We note by $u_{D_n}(i)$ the i^{th} element of the n^{th} DMT symbol. The cyclic prefix is given by $u_{P_n}(i) = u_{D_n}(K - L + i)$ for $i = 0, \ldots, L - 1$. Using a similar approach, the data signal v_d can be defined the same way as v_u. However in this case, the tones corresponding to the training signal bandwidth will carry zero information.The received time domain signal is represented by the vector y_d and it can be expressed as:

$$y_d = \mathbf{M_d} h_{dd} + (\mathbf{M_v} + \mathbf{M_u})\, h_{du} + \nu_0 \tag{8.5}$$

where h_{dd} and h_{du} are two vectors of length L_1 that represent the direct and the crosstalk channels impulse responses. Vector ν_0 represents the noise samples. Matrix $\mathbf{M_u}$ represents the matrix formed by the transmitted training signal on line u. The matrix $\mathbf{M_u}$ is the Toeplitz Matrix having v_u as the first column and having L_1 columns. Matrices $\mathbf{M_v}$, and $\mathbf{M_d}$ represent the matrices corresponding to the data signals on line u and on line d respectively, and they can be defined in a similar way as $\mathbf{M_u}$.

In DSL systems, the direct channel h_{dd} is supposed to be known perfectly. Thus if the transmitted signal on the direct line d is decoded with no errors one can remove the term $\mathbf{M_d} h_{dd}$ from equation (8.5), and the remaining signal is given by:

$$r_d = \mathbf{M_u} h_{du} + \nu. \tag{8.6}$$

The term $\mathbf{M_v} h_{du}$ is now included in the noise $\nu = \mathbf{M_v} h_{du} + \nu_0$.

The least square estimation of the channel h_{du} is given by:

$$\hat{h}_{ls} = \left(\mathbf{M_u}^T \mathbf{M_u}\right)^{-1} \mathbf{M_u}^T r_d. \tag{8.7}$$

However matrix $\mathbf{M} = \mathbf{M_u}^T\mathbf{M_u}$ is ill conditioned due to the presence of too many zeros in the spectral representation of u. These zeros come from the fact that no energy has been transmitted over the tone used for data.

8.3 Condition Number

For a matrix $\mathbf{A}$, the condition number $\kappa(\mathbf{A})$ indicates if a matrix is well conditioned or not. One way to measure the condition number is to use the formula $\kappa(\mathbf{A}) = \frac{|\lambda_{max}(\mathbf{A})|}{|\lambda_{min}(\mathbf{A})|}$, where $|\lambda_{max}(\mathbf{A})|$ and $|\lambda_{min}(\mathbf{A})|$ represent respectively the maximum and the minimum eigenvalues of $\mathbf{A}$. If $\kappa(\mathbf{A}) \approx 1$ the matrix is well conditioned, and if $\kappa(\mathbf{A}) >> 1$ the matrix is ill conditioned [76]. In the following we will describe the behavior of $\kappa(\mathbf{M})$ using a typical DSL training sequence v_u.

DSL systems use different types of training sequences for the direct channel estimation:

- Reverb1 where the same DMT symbol is repeated over different DMT blocks and without adding a cyclic prefix: $[u_D^T, u_D^T, ...u_D^T]^T$.
- Reverb2 where the same DMT symbol is repeated over different DMT blocks but with adding a cyclic prefix: $[u_P^T u_D^T, u_P^T u_D^T, ...u_P^T u_D^T]^T$.
- xMedley where different DMT symbols are sent over different blocks and the cyclic prefix is used: $[u_{P_1}^T u_{D_1}^T, ..., u_{P_n}^T u_{D_n}^T, ...u_{P_N}^T u_{D_N}^T]^T$.

It is important to study the condition number of the signal Matrix corresponding to this type of signals when the number of pilot tones to be used is limited. In the case of Reverb1, we assume the length of the channel impulse response L_1 to be equal to the length of an DMT block. Disregarding the first and the last DMT blocks transforms matrix $\mathbf{M}$ into a circular matrix. When disregarding the first and the last DMT block, matrix $\mathbf{M_u}$ can be seen as the concatenation of $N-2$ times a repetitive circular matrix $\mathbf{M_c}$. This circular matrix is formed by the repetitive DMT block used in Reverb1. We conclude that $\mathbf{M} = (N-2)\mathbf{M_c^T}\mathbf{M_c}$. Since $\mathbf{M_c}$ is a circular matrix we have $\mathbf{M_c} = \mathbf{W}^H\mathbf{\Lambda}\mathbf{W}$ where $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements represent the eigenvalues of $\mathbf{M_c}$. We define vector m_c to be the first column of matrix $\mathbf{M_c}$. The diagonal of matrix $\mathbf{\Lambda}$ is given as the FFT of m_c: $diag(\mathbf{\Lambda}) = \mathbf{W}m_c$. We can conclude:

$$\mathbf{M} = (N-2)\mathbf{W}^H\mathbf{\Lambda}^H\Lambda\mathbf{W}. \tag{8.8}$$

Equation (8.8) indicates that matrix $\mathbf{M}$ is a circulant matrix having as eigenvalues the diagonal of matrix $(N-2)\mathbf{\Lambda}^H\mathbf{\Lambda}$, where the eigenvalues of matrix $\mathbf{\Lambda}^H\mathbf{\Lambda}$ correspond to the amplitudes square of the eigenvalues of matrix $\mathbf{M_c}$.

In the Reverb1 case, and for L_1 equal to the length of the DMT block K, the matrix $\mathbf{M_c}$ is formed by a circularly shifted version of u_D^T. Thus when the FFT is applied on the first column of $\mathbf{M_c}$, the zeros corresponding to the Data signal band will appear, which makes the matrix $\mathbf{M}$ singular (since it has zero eigenvalues) with a rank equal to the number of pilot tones. In this case, time domain estimation can not be performed for a channel length L_1 bigger than the total number of pilot tones.
In the case of Reverb2, and if we start observing the received signal at the middle of a DMT block (inducing non synchronization), the first column of the matrix $\mathbf{M_c}$ would be formed by the parts of two consecutive DMT symbols separated by the cyclic prefix. The parts of the DMT symbols can be seen as truncated DMT symbol u_{T_n}: $u_{T_n} = \mathbf{T_n}u_{D_{sn}}$ where $\mathbf{T_n}$ represents a truncation matrix (diagonal matrix that has all the diagonal elements equal to one except for the truncated parts equal to zero), and $u_{D_{sn}}$ is a circularly shifted version of u_{D_n}.The cyclic prefix can also be seen as a truncated version of DMT symbol. Defining $u_{I_{sn}} = u_{I_n}\exp(j\theta)$ where θ corresponds to the circular shift between $u_{D_{sn}}$ and u_{D_n}, results in the relationship $u_{D_{sn}} = \mathbf{W}^H u_{I_{sn}}$. Applying FFT on u_{T_n} gives:

$$U_{T_n} = \mathbf{W}\mathbf{T_n}\mathbf{W}^H u_{I_{sn}} = \mathbf{M_{sinc_n}} u_{I_{sn}} \tag{8.9}$$

The matrix $\mathbf{M_{sinc_n}}$ is a circular matrix having as first column the FFT of a rectangle that corresponds to the diagonal of $\mathbf{T_n}$. According to (8.9), the FFT of a truncated DMT symbol is given as a circular convolution between the FFT of the DMT symbol and a truncated sinc that depends on the truncation matrix (which is by itself dependent on the non-synchronization). In this case spectral elements of m_c are non zero, however tones located far in the non used frequency band would have small energy compared to tones located on the used band and matrix $\mathbf{M_c}$ is ill conditioned.

The length of DMT block in Reverb2 case the is equal to $(K + L)$ (length of the DMT symbol + that of the cyclic prefix). Thus to obtain the eigenvalues of matrix $\mathbf{M_c}$, an FFT of length $(K+L)$ must be applied on its first column m_c. However the FFT of length $(K+L)$ is seen as an interpolation or zero padded version of the length K FFT. The spectral elements corresponding to the non used frequency band would always be small compared to the tones that correspond to the used band. Thus the condition number $\kappa(\mathbf{M_c}) = \frac{|\lambda_{max}(\mathbf{M_c})|}{|\lambda_{min}(\mathbf{M_c})|}$ is large, and matrix $\mathbf{M_c}$ remains ill-conditioned unless we limit the number of taps to be estimated.
In the case of xMedley signal, the matrix $\mathbf{M_u}$ is formed of $N-2$ different Toeplitz matrices $\mathbf{M_{t_n}}$, where each $\mathbf{M_{t_n}}$ corresponds to a DMT block n. The Toeplitz matrix $\mathbf{M_{t_n}}$ can be seen as the combination of a circular matrix $\mathbf{M_{c_n}}$, and a corrective matrix $\mathbf{D_n}$, so we have: $\mathbf{M_{t_n}} = \mathbf{M_{c_n}} + \mathbf{D_n}$. The signal matrix can now be expressed as:

$$\mathbf{M} = \sum_{n=2}^{N-1} (\mathbf{M_{c_n}} + \mathbf{D_n})^{\mathbf{T}} (\mathbf{M_{c_n}} + \mathbf{D_n}). \tag{8.10}$$

The sum of circular matrices is a circular matrix, and since all $\mathbf{M_{c_n}}$ have the same distribution of eigenvalues, which is the same as the distribution of the pilot tones, we can conclude that the circular matrix $\mathbf{M_{C_x}} = \sum_{n=2}^{N-1} \mathbf{M_{c_n}^T M_{c_n}}$ has the same eigenvalue distribution of any of its constituents $\mathbf{M_{c_n}}^T \mathbf{M_{c_n}}$, thus the same condition number. Expression (8.10) can be rewritten as: $\mathbf{M} = \mathbf{M_{C_x}} + \mathbf{D_x}$, where

$$\mathbf{D_x} = \sum_{n=2}^{N-1} \mathbf{D_n^T M_{c_n}} + \mathbf{M_{c_n}^T D_n} + \mathbf{D_n^T D_n}.$$

Matrix $\mathbf{M_{C_x}}$ is the dominant part of the signal matrix $\mathbf{M}$ and its inverse can be given by:

$$\mathbf{M^{-1}} = \mathbf{M_{C_x}^{-1}} \left[\mathbf{I} + \mathbf{D_x M_{C_x}^{-1}} + \left(\mathbf{D_x M_{C_x}^{-1}} \right)^2 + ... \right]. \tag{8.11}$$

As in the case of Reverb signals, equation (8.11) shows that the conditioning of the LS estimation is affected by the condition number of $\mathbf{M}_{\mathbf{C_x}}$ which is related to the distribution of the pilot tones.

8.4 Time Domain Estimation

Section 8.3 shows that in DMT systems, when using a training sequence with limited number of pilot tones, the LS estimation of the channel impulse response is an ill conditioned LS problem. In this section we will improve the LS estimation by simply adding a regularization factor to matrix $\mathbf{M}$. In the first part we suppose that the channel gain information are available, and the regularization factor would be simply seen as the constraints imposed to achieve the known gain values. In the second part we suppose no prior knowledge of the channel gain, in this case we will add a scaled identity matrix as the regularization factor, then we will propose a method to improve the estimation by iteratively minimizing the estimation error.

8.4.1 Constrained LS Estimation

In DSL systems, channels change slowly over long period of time, one may assume the existence of a monitoring system that is able to passively estimate the crosstalk channel gain ([54],[77]). In the following we suppose that we posses the crosstalk channel gain estimates. These estimates are incorporated as constraints in the time domain estimation to improve its conditioning. The least square problem becomes:

$$\begin{array}{ll} \min_{h_{du}} & (r_d - \mathbf{M_u} h_{du})^T (r_d - \mathbf{M_u} h_{du}) \\ \text{S.t} & \left| H_{du} \right| = \left| \hat{H}_{du_m} \right| \end{array} \tag{8.12}$$

where $\left|\hat{H}_{du_m}\right|$ are channel gain estimates given by the monitoring system. Problem (8.12) is a constrained optimization that can be solved iteratively using Lagrange formulation as the following:

$$\begin{aligned} F &= (r_d - \mathbf{M_u} h_{du})^T (r_d - \mathbf{M_u} h_{du}) \\ &+ \sum_i \mu_i \left(\left| H_{du}(i) \right| - \left| \hat{H}_{du_m}(i) \right| \right) \end{aligned} \tag{8.13}$$

where μ_i is the Lagrange multipliers for tone i, and F is the Lagrange function. Consider $\mathbf{\Sigma}$ to be a diagonal matrix with μ_i as diagonal elements, F can be written as:

$$\begin{aligned} F &= (r_d - \mathbf{M_u} h_{du})^T (r_d - \mathbf{M_u} h_{du}) + h_{du}^T \mathbf{W^H \Sigma W} h_{du} \\ &- \sum_i \mu_i \left| \hat{H}_{du_m}(i) \right| . \end{aligned} \tag{8.14}$$

To ensure that the expression $h_{du}^T \mathbf{W^H \Sigma W} h_{du}$ is always real, the Lagrange multipliers corresponding to the positive and negative frequencies should be the same. The minimization problem can be solved iteratively: At iteration t, and for fixed μ_i^t we can find $\hat{h}_{cls}^t$ that minimizes F in (8.14) and then using $\hat{h}_{cls}^t$ we can proceed to update μ_i^{t+1}. At fixed μ_i^t, minimum F with respect to h_{du} is given by:

$$\frac{\partial F}{\partial h_{du}} = 0 \tag{8.15}$$

Which is given by:

$$\hat{h}_{cls}^t = \left(\mathbf{M}^T \mathbf{M} + \mathbf{W^H \Sigma^t W} \right)^{-1} \mathbf{M_u}^T r_d \tag{8.16}$$

The search of the Lagrange multipliers is done in the gradient direction where at each iteration t, μ_i is updated according to:

$$\mu_i^t = \mu_i^{t-1} - \alpha \left(\left| \hat{H}_{cls}^t(i) \right| - \left| \hat{H}_{du_m}(i) \right| \right) \tag{8.17}$$

With α is a small step size, and $\hat{H}_{cls}^t$ is the FFT of $\hat{h}_{cls}^t$. The values μ_i^t are then used to update the matrix $\mathbf{\Sigma}^t$.

8.4.2 Penalized LS

One of the solutions used for ill-conditioned LS problems is to add a regularization factor to matrix $\mathbf{M}$ in order to make it invertible. Let the matrix $\mathbf{m_u}$ be the regularization factor; the channel estimator can now be expressed as:

$$\hat{h}_{pls} = \Big(\mathbf{M} + \mathbf{m_u}\Big)^{-1} \mathbf{M_u}^T r_d. \tag{8.18}$$

In fact, the constrained LS algorithm described in the previous section can be seen as a regularized LS solution, where the channel constraint matrix $\mathbf{W^H \Sigma W}$ is the regularization factor. In the lack of additional channel gain information, the simplest choice of the regularization factor is to use Tikhonov regularization [78], [79]. In this case, the matrix $\mathbf{m_u}$ is equal to $a\mathbf{I}$, where a is a positive scaling factor and $\mathbf{I}$ is the identity matrix.

In the case of Tikhonov regularization, $\hat{h}_{pls}$ can be seen as the result of the following minimization:

$$\min_{h_{du}} \quad (r_d - \mathbf{M_u} h_{du})^T (r_d - \mathbf{M_u} h_{du}) + a h_{du}^T h_{du}. \tag{8.19}$$

In (8.19) the LS criterion is modified to include a scaled square Euclidean norm of the channel impulse response. Hence an estimate with a large impulse response is penalized by the regularization term a, from where comes the name: "Penalized Least Square" or simply PLS.

8.4.3 Iterative Improvement of PLS

In this paragraph we try to improve the performance of the PLS algorithm by iteratively removing a part of the penalization factor. This procedure can be seen as improving the regularization factor used in the PLMS algorithm. Using the identity $(\mathbf{I} + \mathbf{AB})^{-\mathbf{1}} = \mathbf{I} - \mathbf{A}(\mathbf{I} + \mathbf{BA})^{-\mathbf{1}}\mathbf{B}$, expression (8.18) can be rewritten as:

$$h_{du} = \hat{h}_{pls} + (\mathbf{M} + \mathbf{m_u})^{-1} \mathbf{m_u} h_{du} + \epsilon \tag{8.20}$$

where $\epsilon = -\Big(\mathbf{M} + \mathbf{m_u}\Big)^{-1} \mathbf{M}^T \nu$. Expression (8.20) can be used to further improve the estimation of h_{du}. Let $\mathbf{M_e} = (\mathbf{M} + \mathbf{m_u})^{-1} \mathbf{m_u}$.

Minimizing the mean square of ϵ results in minimizing the following metric:

$$\begin{aligned} \|\epsilon\|_2^2 = & \Big(h_{du} - (\hat{h}_{pls} + \mathbf{M_e} h_{du})\Big)^T \\ \times & \Big(h_{du} - (\hat{h}_{pls} + \mathbf{M_e} h_{du})\Big). \end{aligned} \tag{8.21}$$

The gradient of (8.21) is given by:

$$\begin{aligned} \partial \|\epsilon\|_2^2 / \partial h_{du} &= 2\Big[h_{du} + \Big(\mathbf{M_e}^T\mathbf{M_e} - (\mathbf{M_e}^T + \mathbf{M_e})\Big) h_{du} \\ &+ \Big(\mathbf{M_e}^T - \mathbf{I}\Big)\hat{h}_{pls}\Big]. \end{aligned} \tag{8.22}$$

The minimization of $\|\epsilon\|_2^2$ is a least square minimization. Hence to find $\tilde{h}_{ipls}$ that minimizes (8.21), it is sufficient to set the gradient (8.22) to zero using the following iterative algorithm

$$\begin{aligned} \tilde{h}_{ipls}^{t+1} &= \Big(\mathbf{I} - \mathbf{M_e}^T\Big)\hat{h}_{pls} \\ &- \Big(\mathbf{M_e}^T\mathbf{M_e} - (\mathbf{M_e}^T + \mathbf{M_e})\Big)\tilde{h}_{ipls}^t \end{aligned} \tag{8.23}$$

where t represent the iterations. At the first iteration ($t = 1$), the initial estimation of the channel impulse response is given by: $\tilde{h}_{ipls}^1 = \hat{h}_{pls}$. The iterative algorithm (8.23) will be called iterative penalized least square (IPLS), and it can be justified using the identity: $(I + A)^{-1} = I + \sum_{n=1}^{\infty}(-1)^n A^n$, where if $\tilde{h}_{ipls}^1 = 0$ when t goes to infinity $\tilde{h}_{ipls}$ is given by:

$$\begin{aligned} \tilde{h}_{ipls} &= \Big(\mathbf{I} + \mathbf{M_e}^T\mathbf{M_e} - (\mathbf{M_e}^T + \mathbf{M_e})\Big)^{-1} \\ &\times \Big(\mathbf{I} - \mathbf{M_e}^T\Big)\hat{h}_{pls}. \end{aligned} \tag{8.24}$$

In the case of Tikhonov regularization, the matrix $\Big(\mathbf{M_e}^T\mathbf{M_e} - (\mathbf{M_e}^T + \mathbf{M_e})\Big)$ is reduced to $\mathbf{M_e}(\mathbf{M_e} - 2\mathbf{I})$, where $\mathbf{M_e} = (\mathbf{M} + a\mathbf{I})^{-1}a\mathbf{I}$. Considering that L_1 is equal to the DMT block length, and taking the Reverb1 case, using (8.8) gives:

$$\mathbf{M_e} = aW^H\left((N-2)\Lambda^H\Lambda + a\mathbf{I}\right)^{-1} W. \tag{8.25}$$

Hence $\mathbf{M_e}$ is a circular and symmetric matrix, with eigenvalues equal to $a((N-2)|\lambda(i)|+a)^{-1} \leq 1$, which makes $\mathbf{M_e}(\mathbf{M_e}-2\mathbf{I})$ negative definite. Most of the eigenvalues of matrix $\mathbf{M_e}(\mathbf{M_e}-2\mathbf{I})$ are equal to -1, the few eigenvalues that corresponds to the pilot tones are strictly bigger than -1. It can be shown that the additional error by iteration would have the same sign as the previous error thus they will add up. One way to mitigate this problem is to choose the number of iteration as a function of the noise power: if the noise is high limit the number of iterations, if the noise is small increase the number of iterations.

8.4.4 Performance Analysis

Since the proposed estimators in this chapter are biased, we propose to study the mean quadratic error as an indicator of performance. Let $\hat{h}$ be a biased estimator of h, and $m_{\hat{h}} = E\left[\hat{h}\right]$. The mean quadratic error $\eta_{\hat{h}}$ will be defined as the diagonal of the following matrix $\mathbf{C_E}$: $\mathbf{C_E} = E\left[(\hat{h}-h)(\hat{h}-h)^T\right]$.
It can be shown that: $\mathbf{C_E} = \mathbf{C_{\hat{h}}} + (h-m_{\hat{h}})(h-m_{\hat{h}})^T$, where $\mathbf{C_{\hat{h}}}$ is the covariance matrix of estimator $\hat{h}$. Thus we have:

$$\eta_{\hat{h}}(i) = \mathbf{C_{\hat{h}}}(i,i) + (h(i)-m_{\hat{h}}(i))^2 \tag{8.26}$$

ie the mean quadratic error is equal to the sum of the variance of the estimator and the square of the bias. In the frequency domain the mean square error is equal to the diagonal of $\mathbf{W}\mathbf{C_E}\mathbf{W^H}$.

8.4.4.0.4 Noise Covariance As it was previously seen, the noise of the received signal is given by $\nu = \mathbf{M_v}h_{du} + \nu_0$. Let v_d designs the vector corresponding to the data signal represented by matrix $\mathbf{M_v}$ and let $\mathbf{H}_{du}$ be a $(N_t \times N_t)$ Toeplitz matrix where the first L elements of the first column are equal to h_{du}^T while the rest are zeros. The noise can be rewritten as $\nu = \mathbf{H}_{du}v_d + \nu_0$, where v_d is a vector that represents the data signal on line d.
Supposing that the signal v_d has equal power $\sigma_{v_d}^2$ over all tones, and that the cyclic prefix and the pilot signal are small compared to the

data signal, we can write the noise covariance in the time domain as:

$$\mathbf{C}_\nu = \frac{1}{K}(\sigma^2_{v_d}\mathbf{H}_{du}\mathbf{H}^T_{du} + \sigma^2_\omega \mathbf{I}) \tag{8.27}$$

8.4.4.0.5 CLS Performance The covariance of CLS estimator is given by:

$$\begin{aligned} \mathbf{C}_{\hat{\mathbf{h}}_{\mathbf{cls}}} &= \Big(\mathbf{M}^T\mathbf{M} + \mathbf{W}^{\mathbf{H}}\mathbf{\Sigma}\mathbf{W}\Big)^{-1}\mathbf{M_u}^T\mathbf{C}_\nu \\ &\times \mathbf{M_u}\Big(\mathbf{M}^T\mathbf{M} + \mathbf{W}^{\mathbf{H}}\mathbf{\Sigma}\mathbf{W}\Big)^{-1}. \end{aligned} \tag{8.28}$$

And the bias is simply given by:

$$\mathtt{B}_{cls} = \Big(\Big(\mathbf{M}^T\mathbf{M} + \mathbf{W}^{\mathbf{H}}\mathbf{\Sigma}\mathbf{W}\Big)^{-1}\mathbf{M_u}^T - \mathbf{I}\Big)h_{du} \tag{8.29}$$

And the quadratic mean error is equal to:

$$\eta_{cls} = diag(\mathbf{C}_{\hat{\mathbf{h}}_{\mathbf{cls}}} + \mathtt{B}_{cls}\mathtt{B}^T_{cls}) \tag{8.30}$$

8.4.4.0.6 PLS & IPLS Performance The covariance of PLS estimator is equal to:

$$\mathbf{C}_{\hat{\mathbf{h}}_{\mathbf{pls}}} = \Big(\mathbf{M}^T\mathbf{M} + \mathbf{m_u}\Big)^{-1}\mathbf{M_u}^T\mathbf{C}_\nu\mathbf{M_u}\Big(\mathbf{M}^T\mathbf{M} + \mathbf{m_u}\Big)^{-1} \tag{8.31}$$

while the bias is given by:

$$\mathtt{B}_{pls} = \Big(\Big(\mathbf{M}^T\mathbf{M} + \mathbf{m_u}\Big)^{-1}\mathbf{M_u}^T - \mathbf{I}\Big)h_{du}.$$

To find the covariance of an IPLS estimator at iteration t first define the matrix:

$$\mathbf{S}_t = \sum_{c=1}^{t}\Big((\mathbf{M_e}^T + \mathbf{M_e}) - \mathbf{M_e}^T\mathbf{M_e}\Big)^{c-1}\Big(\mathbf{I} - \mathbf{M_e}^T\Big).$$

Now the covariance per iteration can be given by:

$$\mathbf{C}^{\mathbf{t}}_{\hat{\mathbf{h}}_{\mathbf{ipls}}} = \mathbf{S}_t\mathbf{C}_{\hat{\mathbf{h}}_{\mathbf{pls}}}\mathbf{S}^T_t \tag{8.32}$$

and the bias is defined by:

$$\mathtt{B}_{ipls}^{t} = \Big(\mathbf{S}_t\Big(\mathbf{M}^T\mathbf{M} + \mathbf{m_u}\Big)^{-1}\mathbf{M_u}^T - \mathbf{I}\Big)h_{du}. \tag{8.33}$$

The quadratic mean error in the case of PLS and IPLS estimators (η_{pls}, η_{ipls}) are defined the same way as equation (8.30). The quadratic mean error of these different estimators depends directly on the channel impulse response to be estimated, thus it will be impossible to evaluate the estimators' performance prior to the estimation. However the estimated values can be used to evaluate the performance after the estimation, or to propose a stop criterion to the IPLS estimation, where the iterations can be stopped when η_{ipls} starts increasing.

8.5 Simulation

In the following the different estimation algorithms presented in this chapter are tested using a 2 users DSL scenario. The victim's direct channel and the crosstalk channels are measured channels from a real case situation. The length of the DMT symbol is equal to 512 tones ($K = 512$). The training sequence used is an xMedley signal. Two types of pilot tones distribution are used: Regrouped pilot tones, and distributed pilot tones.

As shown in Fig.8.1, regrouped pilot tones refers to the case where pilot tones are adjacent to each others and occupy 2 separated frequency bands, while distributed pilot tones are uniformly distributed over the entire frequency band. In the regrouped distribution, the two frequency bands occupied by pilot tones are separated by a large frequency band. In this case, the side lobes caused by the non synchronization will fade away in the middle of this large band gap resulting in an ill conditioned problem. Regrouped tones can represent the US frequency bands in VDSL systems separated by the DS bands. If the DSL system suffer from non synchronization between different lines, it is important to estimate the side lobes caused by the US signal on the DS signals, and vice-versa.

In the uniform distribution, 2 pilot tones are separated by a small frequency band that can be easily overcome by side lobes if the number of pilot tones is high enough. In this case the estimation problem is well conditioned. This distribution should be used when estimating crosstalk channel between the same frequency bands (DS or US).
The length of the impulse response to be estimated is set to 45 taps.

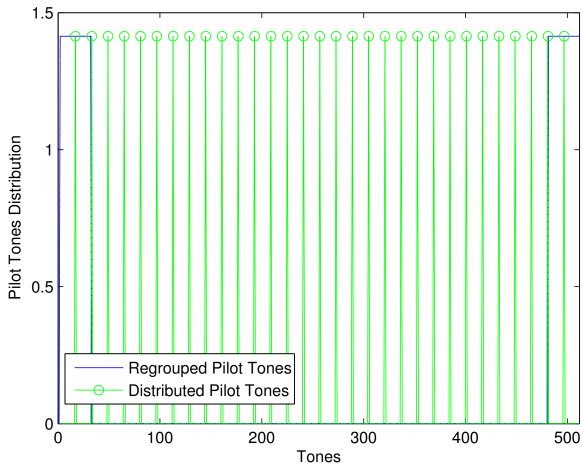


Figure 8.1: *Pilot tones distribution within the training signal*

The metric for testing the different algorithms is the victim's direct channel capacity before and after the crosstalk cancellation based on the estimated values. The channel capacity before the crosstalk cancellation R_d is given by expression (8.2) while the capacity after the cancellation is given by:

$$R_e = \sum_k \log_2 \left(1 + \frac{|H_{dd}|^2(k)P_d(k)}{|H_{du} - \hat{H}_{du}|^2(k)P_u(k) + \sigma_\omega^2(k)} \right). \tag{8.34}$$

$\hat{H}_{du}$ is the FFT of the h_{du} estimates obtained by means of one of the different estimation algorithms. Let R_{ls}, R_{pls}, R_{ipls}, and R_{cls} be the obtained post crosstalk cancellation capacities using LS, PLS, IPLS, and CLS respectively. The noise affecting the estimation process is mainly composed by the data signal transmitted on the other tones. In order to improve the estimation one may apply a power back off on the data symbol tones. We define F as the ratio between the power transmitted at the Data symbol tones during and before the estimation. When $F = 1$ it means that the power at the data symbol tones did not change during the estimation, thus the error is large. For $F = 0.5$ it means the power has been lowered by 3 dB over all the symbol tones.

8.5.1 Regrouped Distribution

The first simulation is done using a regrouped pilot tones training signal. A total of 64 pilot tones is used, with 32 adjacent tones occupying the lower part of the frequency band and another 32 tones occupying its higher part. The estimation is done over 160 DMT symbols, and using $F = 0.5$ for the data signal. Table 8.1, 8.2 and 8.3 report the capacity in bit/DMT symbols obtained using different estimators, and for different scaling factors a. For each scaling factor a, the estimators are tested over 3 different crosstalk channel impulses h_{du},$h_{du,1}$ and $h_{du,2}$. These channels are taken from measurements performed between 400 m France telecoms cables. The results show that the constrained LS estimator outperform the other estimators which is normal since it has access to the crosstalk channel gain information. The penalized LS estimator performance is highly dependent on the scaling factor, as if a is chosen too small or too large, it may deteriorates the results of PLS. However, if a is chosen too large the iterative PLS is able to remove the added bias, and improve the overall estimation as shown in Table 8.2 and 8.3.

Crosstalk channels	R_d	R_{ls}	R_{pls}	R_{ipls}	R_{cls}
h_{du}	5217.9	6568.2	7028.6	7217.7	7626.4
$h_{du,1}$	6488.5	7684	8179	8260.6	8636.1
$h_{du,2}$	6902	8221.1	8570.6	8723.7	9170.5

Table 8.1: *Results obtained for scaling factor $a = 0.0926$*

Crosstalk	R_d	R_{ls}	R_{pls}	R_{ipls}	R_{cls}
h_{du}	5217.9	6572.3	6702.5	7104.4	7562.3
$h_{du,1}$	6488.5	7658.1	7900.4	8200.5	8723.8
$h_{du,2}$	6902	8250.8	8189.2	8656.7	9125.5

Table 8.2: *Results obtained for scaling factor $a = 0.1852$*

Crosstalk	R_d	R_{ls}	R_{pls}	R_{ipls}	R_{cls}
h_{du}	5217.9	6568.2	6474.5	7010.3	7525
$h_{du,1}$	6488.5	7658.1	7644.3	8155	8818.6
$h_{du,2}$	6902	8071.8	7942.4	8478.6	8998.9

Table 8.3: *Results obtained for scaling factor $a = 0.3704$*

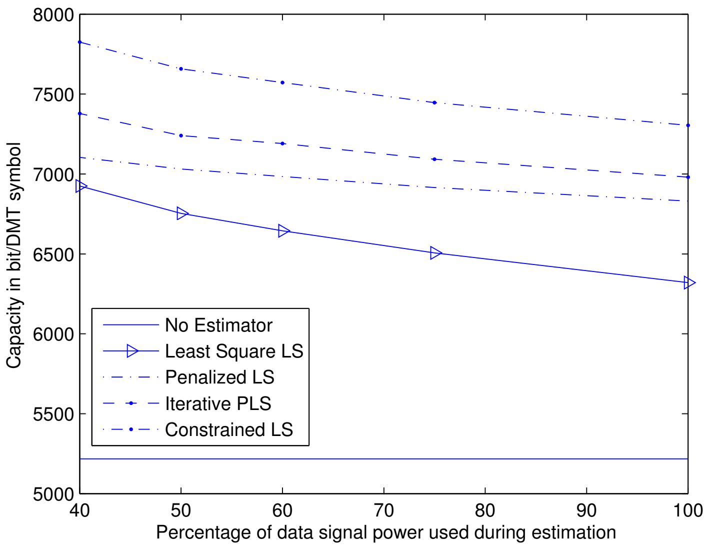


Figure 8.2: *The reached capacity using different estimation algorithms*

Fig.8.2 shows the capacity given by (8.34) when the crosstalk channel h_{du} is compensated using the different proposed estimators in function of $100F$. For the PLS algorithm the regularization factor is fixed to 0.0926, and for the IPLS we use the minimum η_{ipls} as stop criterion. This result shows that the CLS has outperformed the LS algorithm, where the capacity obtained using the CLS estimates for crosstalk cancellation is between 13% and 15.6% higher than the capacity obtained using the LS estimates. IPLS algorithm outperform LS as well, where it has a capacity gain over LS that varies between 6.56% and 10.46%. The capacity gain achieved by PLS algorithm over LS varies between 2.6% and 8%. It is clear that the LS algorithm is mainly deteriorated when the noise becomes large (F=1), this is due to the fact that the ill conditioning of the LS estimation amplifies the effect of noise. Fig.8.3 shows the evolution of the IPLS estimates error with respect to the iterations. The theoretical mean quadratic error (8.30) computed using the true crosstalk channel value h_{du}, and the estimated value $\hat{h}_{du}$ are also shown on the figure. The quadratic mean error calculated using $\hat{h}_{du}$ is about 3 to 4 dB higher than the actual estimation error between the estimated channel impulse response and the real one, however both curves have the same shape, and in this case they have the minimum values located at iteration 8. In general, using different simulations, the minimums of real and theoretical curves are found to be located at vicinity of each others (typically less than 2 to 3 iterations).

Finally, Fig.8.4 shows the residual crosstalk after cancellation of H_{du} using IPLS and LS algorithms. For the IPLS case, we can see that the crosstalk was reduced more than 50% for 300 tones. These tones are located at the higher and lower frequency bands were the pilot tones are used. In the middle band where the side lobes are not effective, the crosstalk reduction is limited however the resultant crosstalk does not exceed the original one, thus the capacity is either improved or stays the same over all tones. For the LS case, the crosstalk was reduced more than 50% for 150 tones. However, the resultant crosstalk in the middle band exceeds the original crosstalk by 50 % for some cases. Even though the capacity did not improve for the middle frequency band using the IPLS precoding, however the system SNR did not deteriorate. Thus the IPLS results can be used as an initialization of the method presented in chapter 6, where an accurate estimation can be done without the system 3 dB SNR loss suggested there.

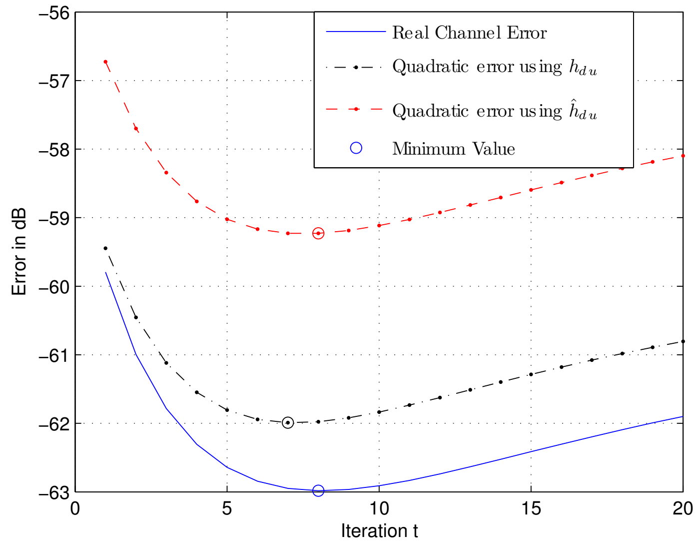


Figure 8.3: *The mean quadratic error in function of Iterations*

8.5.2 Uniform Distribution

Pilot Tones	Condition Number	R_{ls}	R_{pls}	R_{ipls}	R_{cs}
64	1.5990	9481.8	9108.6	9481.8	10204
32	2.8812	8380.7	8223.3	8397.6	10382
16	441.8778	6434.1	7353.4	7372.4	8686.3
8	2375.1	4857.6	5966.4	6101.1	6230.1

Table 8.4: *Results obtained using a uniformly distributed pilot tones signal*

The following simulation is done using only 16 DMT symbols, a 3 dB power back off is applied (F=0.5), and scaling factor $a = 0.3704$. Table 8.4 reports the capacity in bit/DMT symbol obtained after crosstalk cancellation for H_{du}. Using the same number of pilot tones as the regrouped distribution case (64 pilot tones) results in a well conditioned estimation problem where the condition number is equal to 1.6, in

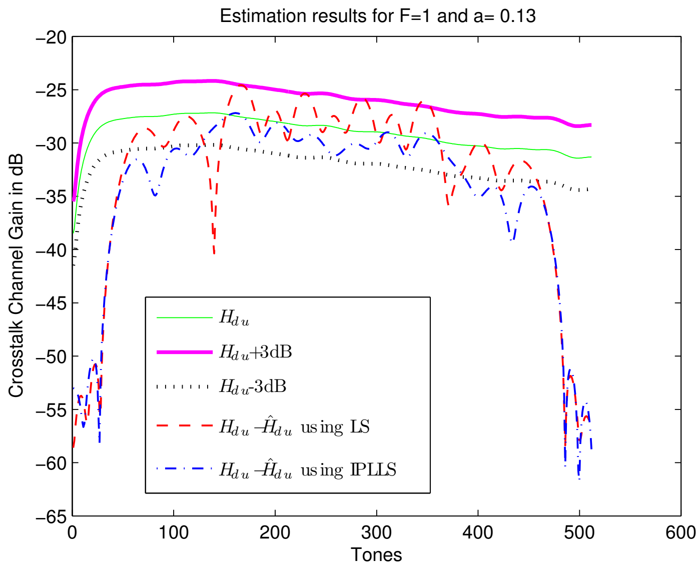


Figure 8.4: *Residual Crosstalk after Cancellation*

this case the LS estimator outperformed the PLS. However the IPLS finally converged to the same result obtained by the LS. As the number of pilot tones went down, the condition number increased, and the problem becomes ill conditioned. When the number of pilot tones is limited to 8, the LS estimator performance deteriorates, it even results in a capacity lower than the pre crosstalk cancellation capacity ($R_d = 5217.9$). The other estimators always conserve a capacity gain > 1 when compared to R_d. As seen in the regrouped pilot tones signal, the constrained estimator outperformed the other estimators.

8.6 Conclusion

In this chapter we have presented techniques to estimate crosstalk channels in DSL systems using a limited number of pilot tones. Using a limited number of pilot tones reduces the overhead caused by the crosstalk estimation, and insures a continues data communication on the DSL systems, however the crosstalk estimation may become an ill-conditioned problem. Due to the ill-conditioning theclassical LS algorithm becomes non-reliable. Several algorithms for solving an ill-conditioned estimation problem were introduced, Such as constrainted LS, penalized LS algorithms. An iterative reduction of the penalization factor was also proposed. These techniques improved the LS estimation of the crosstalk channels considerably. Even in the case of a well conditioned estimation problem, the iterative removal of the regularization facto in penalized LS algorithm was able to converge to the same results achieved by LS.

Chapter 9

Conclusions

In this thesis we addressed the problems of implementing DSM algorithms under practical and realistic conditions and constraints on the DSL systems. In the literature, proposed algorithms for DSM level 2 and DSM level 3 always assume the perfect knowledge of of both crosstalk channels and direct channels. While the direct channels is estimated and provided in the current DSL systems, the crosstalk channel is not known. Few estimation methods were proposed for the crosstalk channels' estimations. However, most of these methods includes coarse procedures that requires the halt of communication on the line under investigation, such as the use of pilot sequences or other DSL test methods (SELT, DELT). Furthermore, DSM level 2 centralized algorithms includes complex optimization techniques that may need long execution time. The main achievement of this thesis are focused on the enhancement of the state of the art DSM level 2 algorithms mainly ISB and OSB, and on proposing practical estimation techniques for the crosstalk channels that requires minimal changes on the current DSL standards and modems.

9.1 Achievements

In the first part of this thesis we proposed to solve the Near-Far problem in the DSL systems by adopting the "Balanced Capacity" concept in the DSM level 2 algorithms. By using the "Balanced capacity" concept we were able to ensure fairness between the different users in a

DSL system, where all users were able to transmit at the same percentage of their maximum achievable bitrate.

We then proceeded with the enhancement of some known state of the art DSM level 2 algorithms (mainly ISB and OSB) by incorporating the successive optimization (SO) concept within the optimization procedure. Successive optimization takes advantage of the correlation of the channels gain between adjacent tones to propose a proper initialization of the optimization at each tone based on the result found on the previous tone. The successive optimization techniques resulted in DSM algorithms with better convergence, where the execution time is decreased considerably.

Another advantage of the successive optimization is that it allows the use of low complex algorithms like Newton-Raphson and gradient algorithm (Steepest Ascent). We were able to adapt Newton-Raphson and steepest ascent to the DSL systems and the numerical results showed that the execution time was decreased further more when these two algorithms replaced the typical line search and exhaustive research used in ISB and OSB.

To improve the optimization results, and to prevent the successive optimization from falling in a poor local optimum region, we proposed to use a multi start points technique. At each tone, and after the implementation of the optimization starting by the results found on the previous tone, we run the optimization procedure a second time using an additional random starting point (ASP). The results of the SO and ASP are compared and the best value is kept. A reverse successive optimization was also implemented in order to take advantage the ASP results for the previous tones. These simple procedures prevented the optimization from being trapped in local optimum regions and provided better results than the traditional ISB algorithm. The enhancement techniques proposed in the first part of this thesis can be generalized and used to improve other DSM level 2 algorithms.

In the second part of this work we proposed practical algorithms for the estimation of the crosstalk channel. We started by suggesting the establishment of a monitoring center for the DSL systems. In this monitoring system, observations related to the SNR changes, the time of connection/disconnection of the users, the PSD, the noise level can be stored. These observations were used to estimate the channel gain of the crosstalk channels of a synchronized DSL system by correlating the SNR changes on a line to the connection/disconnection of other users. Techniques that involved a time domain model of the crosstalk channels were used to compress the crosstalk channel gain information, and to improve the overall estimation.

We extended the use of the passive estimation techniques for non synchronized DSL systems, where we developed a model for the asynchronous crosstalk channels' gain. Based on this model we proposed an estimation technique for the crosstalk channel gains based on the observation of the SNR changes. A blind estimation technique for the time delay between two different lines was also suggested. The proposed time delay estimation only required a non correlated power allocation at least between two adjacent tones.

We have also analyzed the effect of the estimation error on DSM level 2 algorithms. While in the literature some work is done to study the effect of the estimation error on DSM using simulation, or by studying the estimation error effect on the allocated bitrate. However, and to the best of our knowledge, there is no work that analyzes the effect of the estimation error on the power allocation itself (the result of the optimization). In this work, we used Taylor approximation in order to find a relationship between the estimation error and the allocated power and its effect on the system bitrate. We found that the estimation error could lead the optimization algorithm to a local optimal power allocation instead of the global one. This is more probable to happen when the estimation error is high, or when the local and global optimum are close to each others.

We have also proposed a total crosstalk channel estimation for DSM level 3 applications. This estimation procedure is based on an active observation of the SNR changes on a synchronous DSL lines. In this procedure, for each DSL line, and for each tone, we modified the crosstalk channels to be estimated by inducing perturbation on the line. By changing the crosstalk channel two times, one time by changing its real part, and another time by changing its imaginary part, we were able to estimate the crosstalk channel normalized by the direct channel. This estimation procedure was further improved by using a time domain model that relates the crosstalk channels over different tones. The use of a time domain model allowed the tone wise estimation to be done over a limited number of pilot tones.

Finally we proposed another time domain estimation technique for the crosstalk channels, this technique uses the LS to estimate the crosstalk channel time domain model based on a known training sequence. However, since the time domain model is usually formed of limited number of taps, and since adjacent tones have correlated crosstalk channels, we proposed to use a limited number of pilot tones in the training sequence. This allows a continuous communication in the DSL system even when using training sequences for the crosstalk channels' estimation. We found that this technique can give good results if the pilot tones are uniformly distributed over the DSL band. If not, the estimation problem becomes an ill-conditioned LS problem. However, even for ill-conditioned case we proposed several remedies that improve the estimation of the crosstalk channel.

9.2 Perspectives

The implementation of DSM level 2 and level 3 represents the next step for the DSL systems. This thesis presented several solutions that can help implementing DSM using the current DSL standards and modems. However more studies are needed before the full application of DSM. The present work can be expanded in several ways:

- Virtual binder identification: In this work we consider that the DSL lines that are at proximity of each others, thus influence each others, are known. This facilitates the estimation procedure. However, these informations are not always available, and a procedure to regroup the related DSL lines within one virtual binder must be proposed.

- Distributed estimation and optimization algorithms: While distributed DSM level 2 algorithm are already proposed in the literature, there is no work that combines a distributed bitrate optimization with a distributed crosstalk channels estimation.

- Proposal of DSM level 2 algorithm that takes into consideration the effect of the estimation errors.
- Estimation of the number of taps in the crosstalk channel time domain model based on the estimated crosstalk channel gain.
- Proposal of DSM level 3 iterative precoder that incorporate the crosstalk channel estimation based on the SNR method.
- Comparison between the time-domain crosstalk channels' estimation using SNR method and the time-domain estimation using limited number of pilot tones.
- The use of compressed sensing techniques to better estimated the crosstalk channels using training sequences with limited number of pilot tones.
- Proposal of an alternative training sequence with limited number of pilot tones: If the training sequence would apply a frequency hoping or a tone hoping on the pilot tones between the different DMT symbols, this may improve the the condition number of the channel matrix used for the LS time domain estimation.

Curriculum vitae

Ali Kalakech
Born in Beirut, Lebanon, on March, 10th 1981.

Education

2003-2004: Masters of Research, Université de Nice Sophia-Antipolis.
Orientation: Signal and Digital Communications.

1998 - 2003 : Electrical Engineer, Beirut Arab University.
Orientation : Telecommunications and Electronics.

Professional background

2005-2010 : Research assistant, Communications and Remote Sensing Laboratory, Université catholique de Louvain.

Research

2006-2010 : PhD in Engineering Sciences, Université catholique de Louvain.
Subject : "DSL: Crosstalk Estimation Techniques, and Impact on DSM".

Journal papers

1. Ali Kalakech, Jérôme Louveaux, Luc Vandendorpe, "Crosstalk Channel Gain Estimation Based on SNR Changes for DSM applications" in preparation for IEEE transactions on Signal Processing.

2. Ali Kalakech, Jérôme Louveaux, Luc Vandendorpe, "Optimized Crosstalk Channel Estimation based on SNR Variation" in preparation for IEEE transactions on Communications.

3. Ali Kalakech, Luc Vandendorpe, " Time Domain Estimation of Crosstalk Channels in DSL Systems Using limited number of Pilot Tones" in preparation for IEEE transactions on Signal Processing.

Conferences papers

1. Ali Kalakech, Jérôme Louveaux, Luc Vandendorpe, "A Balanced Capacity Implementation of DSM for DSL Systems", Symposium on Communication and Vehicular Technology in the Benelux, SCVT 2006, Liège, Belgium, November 2006. This paper received the best paper prize of the conference.

2. Ali Kalakech, Jérôme Louveaux, Luc Vandendorpe, "Applying the Balanced Capacity Concept on the DSL Systems" IEEE International Conference on Communications (ICC 2007). Glasgow, Scotland, United Kingdom, June 2007.

3. Ali Kalakech, Jérôme Louveaux, Luc Vandendorpe, "Enhancement of the Iterative Spectrum Balancing Algorithm for Power Allocation in DSL Systems", IEEE Global Communications Conference (Globecom 2008), New Orleans, LA, USA, November 2008.

4. Jérôme Louveaux, Ali Kalakech, Luc Vandendorpe, "An SNR-Assisted Crosstalk Channel Estimation Technique",Symposium on Communication and Vehicular Technology in the Benelux, SCVT 2008, Antwerp, Belgium.

5. J. Louveaux, A. Kalakech, M. Guenach, J. Maes, M. Peeters, L. Vandendorpe, "An SNR-Assisted Crosstalk Channel Estimation Technique", IEEE International Conference on Communications (ICC 2009), Dresden, Germany, June 2009.

6. Ali Kalakech, Jérôme Louveaux, Luc Vandendorpe, "Application of Gradient Algorithm for Optimizing Power Allocation in DSL Systems",Symposium on Communication and Vehicular Technology in the Benelux, SCVT 2009, Louvain-La-Neuve, Belgium, November 2009.

7. Ali Kalakech, Jérôme Louveaux, Luc Vandendorpe, "Application of Gradient Algorithm for Optimizing Power Allocation in DSL Systems", IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010, March 2010 Dallas, Texas, USA.

Bibliography

[1] S. J. Orfanidis, *Electromagnetic Waves and Antennas.* Rutgers University, 2004.

[2] A. A.(Louis) Beex, and R. Sedki , I. Takeshi , V. Krishnaraj , J. Yoho, and A. Maguibel, "Acterna project interim report 9: Presentation of causal models for twisted pair transmission lines and comparison with ansi t1e1.4 contribution models,," 2001.

[3] I. Takeshi , "Model-based identification of pots local loops for dsl connectivity prediction," Master dissertation, Blacksburg, Virginia, United States, 2001.

[4] A. N. S. for Telecommunications ñ Integrated Service Digital Network (ISDN), "Basic access interface for use on metallic loops for application on the network side of the nt (layer 1 specification)," New York: American National Standard Institutes, 1999.

[5] J.-J. Werner, "The hdsl environment [high bit rate digital subscriber line]," *Selected Areas in Communications, IEEE Journal on*, vol. 9, no. 6, pp. 785 –800, aug 1991.

[6] G. J. Semple, "The effect of bridged taps on digital reticulation systems," *Australian Communication Research*, vol. 19, no. 2, pp. 53 –69, 1985.

[7] T. Starr, P. Silverman, and J. M. Cioffi, *DSL Advances.* Pearson Education, 2002.

[8] J. Cioffi, *Digital Transmission Theory.* Stanford University, 2002.

[9] I.-T. R. G.992.1, "Asymmetric digital subscriber line (adsl) transceivers," July 1999.

[10] I.-T. R. G.992.2, "Asymmetric digital subscriber line (adsl) transceivers - extended bandwidth adsl2 (adsl2plus)," January 2003.

[11] ——, "Very high speed digital subscriber line transceivers 2 (vdsl2)," February 2006.

[12] K. S. Jacobsen, "Text proposal for upstream power back-off," 1999.

[13] ——, "Upstream power back-off for vdsl -practical consideration," 1999.

[14] K. Jacobsen, "Methods of upstream power backoff on very high speed digital subscriber lines," *Communications Magazine, IEEE*, vol. 39, no. 3, pp. 210 –216, mar 2001.

[15] D. F. T. R. TR-057, "Vdsl network element management," Feb 2003.

[16] D. Statovci, T. Nordstrom, and R. Nilsson, "Revised upstream power back-off for vdsl," in *in Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, pp. 633–636.

[17] M. J. et all, "Vdsl power back-off parameter optimization for a cable bundle," in *in the 15th European Signal Processing Conference (EUSIPCO 2007)*, Poznan, Poland, September 2007, pp. 550–554.

[18] D. Statovci and T. Nordstrom, "Performance evaluation of the cable bundle unique power back-off algorithm," in *Third International Conference on Access Networks, AccessNets*, Las Vegas, NV, USA, October 2008, pp. 365–380.

[19] A. committee NIPP pre-published document ATIS-PP-0600007, "Dynamic spectrum management technical report," May 2007.

[20] S. T. Chung, *Implementation and Applications of DSL Technology.* New York: Auerbach Publication, 2007.

[21] J. Lee, S. T. Chung, and J. Cioffi, "Optimal discrete bit loading for dmt-based dsl systems with equal-length loops," *Communications, IEEE Transactions on*, vol. 54, no. 10, pp. 1760 –1764, oct. 2006.

[22] J. Lee, R. Sonalkar, and J. Cioffi, "Multi-user discrete bit-loading for dmt-based dsl systems," vol. 2, nov. 2002, pp. 1259 – 1263 vol.2.

[23] N. Papandreou and T. Antonakopoulos, "Cooperative bit-loading and fairness bandwidth allocation in adsl systems," vol. 2, may 2003, pp. II–352 – II–355 vol.2.

[24] R. Sonalkar, "Bit- and power-allocation algorithm for symmetric operation of dmt-based dsl modems," *Communications, IEEE Transactions on*, vol. 50, no. 6, pp. 902 –906, jun 2002.

[25] E. Baccarelli, A. Fasano, and M. Biagi, "Novel efficient bit-loading algorithms for peak-energy-limited adsl-type multicarrier systems," *Signal Processing, IEEE Transactions on*, vol. 50, no. 5, pp. 1237 –1247, may 2002.

[26] N. Papandreou and T. Antonakopoulos, "A new computationally efficient discrete bit-loading algorithm for dmt applications," *Communications, IEEE Transactions on*, vol. 53, no. 5, pp. 785 – 789, may 2005.

[27] C. Akujuobi, J. Shen, and M. Sadiku, "A new parallel greedy bit-loading algorithm with fairness for multiple users in a dmt system," *Communications, IEEE Transactions on*, vol. 54, no. 8, pp. 1374 –1380, aug. 2006.

[28] D. Yu, K. Seong, and J. Cioffi, "Multiuser discrete bit-loading for digital subscriber lines," june 2007, pp. 2755 –2760.

[29] "Optimal multiuser spectrum balancing for digital subscriber lines," *Communications, IEEE Transactions on*, vol. 53, no. 12, pp. 2167 – 2167, dec. 2005.

[30] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, vol. 3, May 2005, pp. 1937–1941 Vol. 3.

[31] A. Kalakech, J. Louveaux, and L. Vandendorpe, "A balanced capacity implementation of dsm for dsl systems," in *Symposium on Communication and Vehicular Technology in the Benelux (SCVT 2006)*, November 2006.

[32] ——, "Applying the balanced capacity concept on the dsl systems," in *IEEE International Conference on Communications (ICC 2007)*, June 2007.

[33] ——, "Application of gradient algorithm for optimizing power allocation in dsl systems," in *Symposium on Communication and Vehicular Technology in the Benelux (SCVT 2009)*, November 2009.

[34] ——, "Application of gradient algorithm for optimizing power allocation in dsl systems," in *International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010*, March 2010.

[35] ——, "Enhancement of the iterative spectrum balancing algorithm for power allocation in dsl systems," in *IEEE Global Communications Conference (Globecom 2008)*, November 2008.

[36] J. Louveaux, A. Kalakech, and L. Vandendorpe, "An snr-assisted crosstalk channel estimation technique," in *Symposium on Communication and Vehicular Technology in the Benelux (SCVT 2008)*, November 2008.

[37] ——, "An snr-assisted crosstalk channel estimation technique," in *IEEE International Conference on Communications (ICC 2009)*, June 2009.

[38] A. Kalakech, J. Louveaux, and L. Vandendorpe, "Time domain estimation of crosstalk channels in dsl systems using limited number of pilot tones," in *submited to IEEE Global Communications Conference (Globecom 2010)*.

[39] K. B. Song, S. T. Chung, G. Ginis, and J. Cioffi, "Dynamic spectrum management for next-generation dsl systems," *Communications Magazine, IEEE*, vol. 40, no. 10, pp. 101–109, Oct 2002.

[40] W. Yu, G. Ginis, and J. Cioffi, "Distributed multiuser power control for digital subscriber lines," *Selected Areas in Communications, IEEE Journal on*, vol. 20, no. 5, pp. 1105–1115, Jun 2002.

[41] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *Signal Processing, IEEE Transactions on*, vol. 55, no. 8, pp. 4241 –4257, aug. 2007.

[42] W. Lee, Y. Kim, M. Brady, and J. Cioffi, "Spc09-6: Band-preference dynamic spectrum management in a dsl environment," in *Global Telecommunications Conference, 2006. GLOBECOM '06. IEEE*, 27 2006-Dec. 1 2006, pp. 1–5.

[43] R. Cendrillon, M. Moonen, J. Verlinden, T. Bostoen, and W. Yu, "Optimal multiuser spectrum management for digital subscriber lines," in *Communications, 2004 IEEE International Conference on*, vol. 1, 2004, pp. 1–5 Vol.1.

[44] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *Communications, IEEE Transactions on*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[45] J. Papandriopoulos and J. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user dsl networks," in *Communications, 2006. ICC '06. IEEE International Conference on*, vol. 7, June 2006, pp. 3270–3275.

[46] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "Convex relaxation based low-complexity optimal spectrum balancing for multi-user dsl," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3, April 2007, pp. III–349–III–352.

[47] T. Sartenaer, L. Vandendorpe, and J.Louveaux, "Balanced capacity of wireline multiple access channels with individual power

constraints," *Communications, IEEE Transactions on*, vol. 12, no. 53, pp. 2029–2042, December 2005.

[48] P. Tsiaflakis, I. Necoara, J. Suykens, and M. Moonen, "An improved dual decomposition approach to dsl dynamic spectrum management," in *The 17th European Signal Processing Conference (EUSIPCO). Glasgow, U.K.*, August 2009.

[49] S. Galli, C. Valenti, K. Kerpez, and S. Member, "A frequency-domain approach to crosstalk identification in xdsl systems," *IEEE J. Sel. Areas Commun*, vol. 19, pp. 1497–1506, 2001.

[50] C. Zeng, C. Aldana, A. A. Salvekar, S. Member, and J. M. Cioffi, "Crosstalk identification in xdsl systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 1488–1496, 2001.

[51] N. Papandreou and T. Antonakopoulos, "Far-end crosstalk identification method based on channel training sequences," *Instrumentation and Measurement, IEEE Transactions on*, vol. 54, no. 6, pp. 2204 – 2212, dec. 2005.

[52] C. Aldana and J. Cioffi, "Channel tracking for multiple input, single output systems using em algorithm," vol. 2, 2001, pp. 586 –590 vol.2.

[53] Y. Shi, F. Ding, and T. Chen, "Multirate crosstalk identification in xdsl systems," *Communications, IEEE Transactions on*, vol. 54, no. 10, pp. 1878 –1886, oct. 2006.

[54] F. Lindqvist, N. Lindqvist, B. Dortschy, P. Ödling, P. O. Börjesson, K. Ericson, and E. Pelaes, "Crosstalk channel estimation via standardized two-port measurements," *EURASIP J. Adv. Signal Process*, vol. 2008, pp. 1–14, 2008.

[55] K. J. Kerpez, "Automated loop identification on dsl lines," *International Journal of Communication Systems*, no. 22, pp. 1479–1493, 2009.

[56] W. W. Hines and D. C. Montgomery, *Probability and statistics in engineering and management science.* Wiley, 1980.

[57] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory.* Prentice Hall, 1993.

[58] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing.* Prentice Hall, 2000.

[59] V. M. K. Chan and W. Yu, "Multiuser spectrum optimization for discrete multitone systems with asynchronous crosstalk," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, 2007.

[60] "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259 – 268, 1992.

[61] S. Boyd and L. Vandenberghe, *Convex Optimization*, March 2004.

[62] A. Wyglinski, F. Labeau, and P. Kabal, "Effects of imperfect subcarrier snr information on adaptive bit loading algorithms for multicarrier systems," in *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, vol. 6, nov.-3 dec. 2004, pp. 3835 – 3839 Vol.6.

[63] E. Medeiros, N. Lindqvist, M. Monteiro, F. Lindqvist, H. Abrahim, B. Dortschy, and A. Klautau, "Dsm performance on practical dsl systems based on estimated crosstalk channel information," in *The 17th European Signal Processing Conference (EUSIPCO). Glasgow, U.K.*, August 2009.

[64] N. Lindqvist, F. Lindqvist, M. Monteiro, B. Dortschy, E. Pelaes, and A. Klautau, "Impact of crosstalk channel estimation on the dsm performance for dsl networks," *EURASIP J. Adv. Signal Process*, vol. 2009, 2009.

[65] M. Chiang, C. W. Tan, D. Palomar, D. O'Neill, and D. Julian, "Power control by geometric programming," *Wireless Communications, IEEE Transactions on*, vol. 6, no. 7, pp. 2640 –2651, july 2007.

[66] Y. Zhao and G. Pottie, "Optimal spectrum management in multiuser interference channels," 28 2009-july 3 2009, pp. 2266 –2270.

[67] R. Cendrillon, M. Moonen, J. Verlinden, T. Bostoen, and G. Ginis, "Improved linear crosstalk precompensation for dsl," vol. 4, may 2004, pp. iv – 1053–6 vol.4.

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

[69] A. Leshem and L. Youming, "A low complexity coordinated fext cancellation for vdsl," dec. 2004, pp. 338 – 341.

[70] J. Louveaux and A.-J. van der Veen, "Error sign feedback as an alternative to pilots for the tracking of fext transfer functions in downstream vdsl," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 91–91, uary.

[71] P. Duvaut, A. Mahadevan, M. Sorbara, E. Langberg, and P. Biyani, "Adaptive off-diagonal mimo pre-coder (odmp) for downstream dsl self fext cancellation," nov. 2007, pp. 2910 –2915.

[72] I. Ilani, "Point to multipoint channel estimation by "abuse" of receivers," nov. 2006, pp. 146 –150.

[73] P. Whiting, A. Ashikhmin, G. Kramer, C. Nuzman, A. van Wijngaarden, M. Zivkovic, M. Peeters, M. Guenach, J. Maes, and J. Verlinden, "Dsl crosstalk coefficient acquisition using snr feedback," 30 2008-dec. 4 2008, pp. 1 –5.

[74] M. Guenach, J. Louveaux, L. Vandendorpe, P. Whiting, J. Maes, and M. Peeters, "Performance analysis of the signal-to-noise ratio assisted crosstalk channel estimation for dsl systems," june 2009, pp. 1 –5.

[75] J. Le Masson, M. Ouzzif, and I. Wahibi, "Channel estimation using data and pilots for a coordinated dsl system," nov. 2007, pp. 2868 –2872.

[76] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2004.

[77] R. Cendrillon, F. Liming, J. Chou, G. Long, C. Hung, and D. Wei, "Dsm from theory to practice," 30 2008-dec. 4 2008, pp. 1 –4.

[78] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems.* New York: Wiley, 1977.

[79] A. Neumaier, *Solving ill-conditioned and singular linearsystems: a tutorial on regularization.* SIAM, 1998.

[80] J. Barthélemy, G. Cohen, and A. Lobstein, *Complexité algorithmique et problèmes de communications.* Masson, 1992.

[81] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing, year=2006, number=8, volume=24 pages=1426-1438,," *in IEEE Journal on Selected Areas in Communications, Special Issue on Nonlinear Optimization for Communication Systems,.*

[82] A. Garg and A. Chaturvedi, "A hybrid time divisioning scheme for power allocation in dmt-based dsl systems," *Communications Letters, IEEE*, vol. 10, no. 2, pp. 73 –75, feb. 2006.

[83] W. Yu, G. Ginis, and J. Cioffi, "Distributed multiuser power control for digital subscriber lines," *Selected Areas in Communications, IEEE Journal on*, vol. 20, no. 5, pp. 1105–1115, Jun 2002.

[84] J. Huang, R. Cendrillon, M. Chiang, and M. Moonen, "Autonomous spectrum balancing (asb) for frequency selective interference channels," in *Information Theory, 2006 IEEE International Symposium on*, July 2006, pp. 610–614.

[85] J. Louveaux and A.-J. van der Veen, "Adaptive dsl crosstalk precancellation design using low-rate feedback from end users," *Signal Processing Letters, IEEE*, vol. 13, no. 11, pp. 665 –668, nov. 2006.