Rodolpho Guedino de Siqueira

Analysis and Mitigation of the Effect of Repetitive Impulsive Noises on Digital Subscriber Lines

Recife - PE, Brasil

20 de dezembro de 2010

Rodolpho Guedino de Siqueira

Analysis and Mitigation of the Effect of Repetitive Impulsive Noises on Digital Subscriber Lines

Monografia apresentada para obtenção do grau de Engenheiro da Computação da Universidade Federal de Pernambuco.

Orientador:
Djamel Sadok

Co-orientador:
Ramide Dantas

Centro de Informática
Universidade Federal de Pernambuco

Recife - PE, Brasil

20 de dezembro de 2010

Monografia de Trabalho de Graduação sob o título *"Analysis and Mitigation of the Effect of Repetitive Impulsive Noises on Digital Subscriber Lines"*, defendida por Rodolpho Guedino de Siqueirae aprovada em 20 de dezembro de 2010, em Recife, Estado de Pernambuco, pela banca examinadora constituída pelos professores:

Prof. Djamel Sadok
Orientador

Prof. Judith Kelner

Abstract

This work deals with the effect of repetitive impulsive noises -REIN- on *Digital Subscriber Lines* - DSL. Based on simple assumptions about the time-domain characteristics of the noise, a model of the damage it causes on the symbols for each frequency is derived. With that model, a technique for obtaining a compensating *Signal to Noise Ratio Margin* for lines affected by that kind of noise is proposed, so that the noise effect can be mitigated. That value for SNR Margin is calculated in order to be just enough for the mitigation, so that there is no need to increase arbitrarily the protection and consequently decrease the data rate. The technique does not increase the INP value -which would increase delay- and that characteristic can benefit delay-sensible applications. Finally, for validating the research, an ADSL2+ testbed, with real standard-compliant modems communicating through a line emulator, is then used to test both the model and the mitigation technique based on it. The results show good accuracy of the predictions made and lead directly to related future work.

Dedication

I dedicate this work to those who have dedicated their life to me - my parents: Marco and Tânia; Antônio and Elydea.

Acknowledgments

This work has been written throughout several months, but it is indeed the ultimate result of many different things that happened during all my life as a person and as a student. Throughout those years, many people have helped me in different ways, and I thank them sincerely for the good influence they had on me.

First I would like to thank the MIT OpenCourse Ware, from which I learned so much, and the Free Software community, whose work has really helped me on the developing of the research and on the writing of this text.

Then I would like to thank my high school teachers Marinho and Eugenio. I haven't talked to them for a long while, but they were the first to show me what an impact a school teacher can have on somebody's life - and I still owe Eugenio some money from the Carnaval we spent together. Thanks for the inspiration, for the friendship and for the money!

I also thank Edna Natividade for her teaching and for the devoted coordination of our engineering course for a long time. I would also like to thank her and professor Milde Maria for the help they gave me when I mostly needed. I thank Marcilia Campos for her devotion to our understanding, for everything I learned with her as a Probability assistant student, and for her simple and lighthearted character. I also thank professor Hélio Magalhães de Oliveira for the passionate way of teaching and for all the inspiration he gave me. He made me more and more interested in the Communications research area, and it was during his classes that I first felt really proud to be studying Engineering.

I want to thank professor Djamel Sadok and professor Judith Kelner for all the help with this work and for all the opportunities they gave me. During the past two and a half years, I have been researching at GPRT -*Networks and Telecommunications Research Group*-, which is led by them, and that experience has really changed my life. It was the first time that I really started dealing with science and research, and I cannot express how much I have learned at that research group in so many different aspects. I also thank Ericsson for the investment given to our project, from which I acquired all the knowledge necessary to the research.

I thank Ramide Dantas for the help with this writing and for the coordination of our research project. I believe he got to know the people he was working with and he understood many of their different characteristics, so that it was possible to get the best from them. I feel really grateful for all the freedom of thought I had during the research. I thank Glauco Gonçalves for all the help and all the knowledge he shared with the just-arriving student that I was when I first met the group. I thank Andre Palhares for the inspiring and funny discussions we have had. I thank Petronio Gomes for the help of first getting into the group, for all the soccer matches and all the fun we had together as friends. I thank my long-time friend Rodrigo Perazzo for all the good work we have done together, which I am really proud of, and for all the patience and interest shown while listening to my ideas and thoughts - he definitely deserves greetings for that! I thank Thiago Cordeiro for all the discussions on Control Theory, for everything I learned from him about Music, and for all the fun we had together! I finally thank the people who have also shared the lab and their knowledge with me, and with whom I have had so many friendly talks for so much time: Marcio Neves, Douglas Damalio and Moisés Rodrigues. I also thank Rodrigo Melo for the friendship, for the company and for all the parties we went together!

I would also like to thank, from all my heart, all the friends from my class. The end of this year is a well-deserved celebration of so many years of study and effort together. And for you I leave the rest of the thanking to the days after the fun we are going to have in Argentina!

I sincerely thank my closest friends from university: Arthur Liraneto, André Braga, Daniel Britto, Felipe Cadena, Fernando Rodrigues, Guilherme Vasconcelos, Pedro Zuby, Petrônio Gomes, Renan Pires, Ricardo Albuquerque, Rodrigo Perazzo and Vinicius Kursancew. We have had fun in so many different aspects, and the kind of talks we have had makes me proud of being friends with you guys. There was so much inspiration and excitement in some of the talks that I cannot thank enough. I also thank all the guys from GVCin for all the partying time. There was so much fun in the trips and parties I have had with you that I already miss it!

I thank some friends who, in different ways, have played a important role in my life: Hugo Nobrega, Gabriel and Rafael Teixeira de Macedo, Felipe Salsano, Debora Baia, Maria Beatriz, Caio Moraes, Paula Martins, Gregório Bueno, Adah Alecrim, Rafael Querrer, Matheus Lustosa, Bruno Gross, Luisa, Laura Cançado, Sérgio Troyack, Marcella Perruci, Felipe Lobato, Nathalia Souza and Gustavo Polengo.

I want also to thank the friends that I consider to be my second family: Felipe Alecrim, Gilson Falcão, Hugo Lins and Thiago Girard. There is no way I can properly thank all that I have gone through with you guys. You made my life in Recife so much better and easier, and it was with you that I spent most of the best moments of my life! I am sure there is more to come, and I will be there for you wherever I am.

I finally thank my family for all the dedication and all the support given to me. They are the ones without whom I certainly could not have made it. I thank all my cousins, for they have been my friends all my life, and especially the older ones for being such role-models. I thank my cousin José Jr. for his close friendship during aIl our lives, as we have grown up together and have always supported each other.

I greatly thank my mom for her lighthearted spirit, which really inspires me, for the strength she has shown to me throughout life and for all the incredible food! She is one of the most talented persons I have known. I greatly thank my dad for all the lessons, for the impressive clearness of thought and all the braveness shown during some more difficult time. Everything I am is because of you two, and I am so proud of it!

I thank my stepmom, Ana, for being so good to me all this time and for giving me, along with my dad, one of the true joys of my life: my sister Lívia. I also thank another true joy of my life: my little brother, João. I am so proud of you, my brother. You have a brilliant future ahead of you!

I thank my grandma, Elydea, and my grandpa, Antônio, for all the effort and all the help. I also thank them for the good lessons on life and marriage and for providing me such good role-models on how to be a man: grandpa himself, my dad and my uncle.

At last, for this work of science I especially thank my dad, for all the answers he gave me, and my little brother, for all the questions he made me. They are indeed the best science teachers I have had. They are the ones who have never let the discovery spirit vanish in me.

Contents

List of Figures

List of Tables

1 Introduction p. 13

1.1 Goals . . . . p. 14

1.2 Structure . . . . p. 15

2 Literature Review p. 16

2.1 DSL Fundamentals . . . . p. 16

2.1.1 Coding . . . . p. 16

2.1.2 Interleaving . . . . p. 17

2.1.3 Modulation . . . . p. 18

2.1.4 Multiplexing . . . . p. 19

2.1.5 Synchronization . . . . p. 20

2.1.6 DMT Symbol and the Discrete Fourier Transform . . . . p. 21

2.2 Related work . . . . p. 22

3 Development p. 24

3.1 REIN's Effect Modeling . . . . p. 24

3.2 Background Noise Modeling . . . . p. 29

3.3 Mitigation Through Compensating Margin . . . . p. 31

4 Results and Validation p. 38

4.1 Simulation . . . . p. 38

4.2 Testbed and Experiment Description . . . . p. 42

4.3 Results . . . . p. 43

4.3.1 Experiment 1 . . . . p. 43

4.3.2 Experiment 2 . . . . p. 46

5 Conclusions and Future work p. 49

References p. 51

List of Figures

1 $I[n - \Phi]$ for $\Phi \in \{-10, -40, 250\}$ . . . . p. 26

2 $I[n - \Phi]$ for $\Phi \in \{450, 750, 1000\}$ . . . . p. 26

3 The waveform of the noise impulse . . . . p. 28

4 Behavior of the ratio of power measurements for $f = 100Hz$ . . . . . . p. 36

5 Behavior of the ratio of power measurements for $f = 50Hz$ . . . . p. 37

6 The waveform of the noise impulse . . . . p. 39

7 The Discrete Fourier Transform (magnitude) spectrum of the noise impulse p. 39

8 $I[n - \Phi]$ for $\Phi \in \{-10, -40, 250\}$ . . . . p. 40

9 $I[n - \Phi]$ for $\Phi \in \{450, 750, 1000\}$ . . . . p. 40

10 Impulse effect behavior for the 3rd carrier, as a function of the delay Φ p. 40

11 Impulse effect behavior for the 10th carrier, as a function of the delay Φ p. 41

12 Impulse effect behavior for the 50th carrier, as a function of the delay Φ p. 41

13 Impulse effect behavior for the 100th carrier, as a function of the delay Φ p. 41

14 The Discrete Fourier Transform (magnitude) spectrum of the noise impulse p. 44

15 Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 6dB$, and REIN with $f = 50Hz$ . . . . p. 44

16 Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 6dB$, and REIN with $f = 100Hz$ . . . . p. 45

17 Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 6dB$, and REIN with $f = 100$Hz and $f = 50$Hz . . . . p. 45

18 Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 6dB$, and AWGN . . . . p. 46

19 Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 6dB$, all three noises . . . . p. 47
20 Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 9dB$, all three noises . . . . p. 47

List of Tables

1 Loss before and after resynchronization, for different SNR Margin values p. 48

1 Introduction

The family of technologies entitled *Digital Subscriber Line* -DSL- emerged through innovations on the use of telephone line. The communication channel of those lines was designed so that its quality of transmission was enough to transmit the only signal the engineers had in mind by then, the voice signal - with a narrow bandwidth, estimated around 4KHz.

For a long time the, the engineers involved in Telecommunications found it impossible to transmit wide-band signals trough that telephone channel. Besides the distortion at high frequencies caused by the very copper cable, that channel had already been widely affected by Gaussian noise. However, the theory established by Claude Shannon provided the tools which made possible deeper studies about the workings of the communication channels, and those studies became the basis for more efficient design of new communication systems.

Nowadays the DSL lines work based on well established ideas, as *Orthogonal Frequency Division Multiplexing -OFDM-*, error-correcting codes called *Reed-Solomon* -RS- and Claude Shannon's *Channel Coding Theorem* -CCT. That set of fundamentals is formed by ideas expressed by an elegant and consistent mathematical theory, allowing good modeling of the workings of a communication channel. However, almost all the existing modeling of the behavior of those channels is based on the assumption that the noise present on the channel is the additive, white, gaussian type - called AWGN.

The modems are usually implemented based on the assumption of an AWGN channel [Golden e Jacobsen 2006]. So, when a noise which is indeed of that type starts affecting a line, it can just resynchronize in order to adjust to the new environment, and the transmission will be protected. When some other kind of noise affects the line, its effect is not well understood yet, and the modems do not usually implement anything against them [Golden e Jacobsen 2006]. For instance, if a given electric device induces a noise characterized by a repetitive generation of impulses at every 0.025s, the loss caused by that noise may be large and not easily fixable.

The reason why Gaussian noises are easier to protect against is because they have a stationary nature. Basically, their waveform characteristics do not vary greatly with time. On the other hand, impulsive kinds of noises are non-stationary- they alternate between peaks and valleys, having a behavior which may be harder to predict.

In order to illustrate the difference between the stationary noise and the non-stationary cases - and why this is a problem for DSL-, suppose a scenario in which a DSL line has been synchronized and is being affected by some Gaussian noise, but no significant loss of data is happening. Then the power of the Gaussian noise increases and the line starts losing the information it is sending. If it resynchronizes, there should be no loss of data anymore, as long as the noise on the lime remains the same, because the modem is implemented to measure the noise power at each frequency and protect the line against it. As the Gaussian noise nature is very similar through all the time, its power will remain about the same power against which the modem protected the line.

Now let us suppose that the line is synchronized, some impulsive noise appears on the line and starts damaging the transmission. If the line is resynchronized in order to fix it, the modem tries to sense the power of the noise present, but the power of impulsive noises varies greatly from time to time. When an impulse is present, the measured power will be much greater than the one measured during the time in between the impulses. Thus, if the modem decides the power value, against which it will protect the line, based on a measurement in between impulses, it is clearly not sensing the channel in a proper way. If, on the other hand, it tries to get better information about the power based on any kind of mean value of several measurements, it will conclude that there is some moderate powered noise on the line, and it will not protect the carriers from the actual power of the impulses.

That is exactly the kind of problem this work tries to address.

1.1 Goals

The general goals of the research developed and described throughout this text are:

1. Obtain a analytical model for the effect of Repetitive Impulsive Noise on the demodulation of the signals sent through DSL Lines, based on precise assumptions about its time domain characteristics.

2. Propose a resynchronization method which mitigates the effect of those kinds of noises and which does not increase the line protection arbitrarily - but yet removes all the error for at least 95% of the experiments.

1.2 Structure

The rest of the text is organized as follows: chapter 2 describes some of the most fundamental concepts of the DSL lines, giving more details to those most important to the understanding of the work done later, and gives an overview of the related work that has already been done in the same area of research.

Next, chapter 3 shows the derivation of the model proposed for the impact of REIN on the DSL lines and gives a description of a practical way to obtain a compensating margin for that kind of noise, so that its effect can be mitigated. Chapter 4 shows some of the results obtained during the research in order to validate both the model and the technique for mitigation.

Finally, chapter 5 gives the conclusions and outlines what is expected to be the most important contributions of the work done. Furthermore, it mentions future work directly related to what has already been achieved.

2 Literature Review

2.1 DSL Fundamentals

A Digital Subscriber Line works based on many concepts which, although being elegant, are also complex to explain in detail throughout a single dissertation. This section will cover most of them, but will only give details to the most important concepts for the understanding of the work developed in the rest of the text. The ones which are not essential to that goal will be discussed briefly, and the reader may find further information on the references mentioned.

There are several standards concerning what is here called by Digital Subscriber Line. They have been standardized by the *International Telecommunication Union*, ITU-T, and form the basis for the implementation of the DSL communication systems around the world. Some of them are: VDSL, ADSL, ADSL2, ADSL2+. They have some differences but many of the concepts are applicable to the standards just listed.

The ideas described and the claims made next are based on the ADSL2+ standard, which is the standard implemented by the testbed equipment used during the research. However, as the work done here is not really dependent on the small differences between the standards, the fundamentals listed throughout this chapter should apply to most of them. Moreover, when some specific claim is made -about the value of a parameter, for example-, an observation about which standard that value refers to is always made.

2.1.1 Coding

The information to be sent trough a DSL line is transformed into codewords of a error-correcting code called Reed-Solomon. This is done so that relatively small errors can occur during the transmission, at the physical layer, without making the whole data unusable.

Coding is basically a mapping from one message space to another. A simple example of coding would be the Hamming code which assigns to a message x a message of three equal symbols: $C(x) = xxx$. The resulting block of symbols is called *codeword*, and this is the coding part of the coding-decoding scheme. For the decoding, if we assume a binary code, the rule could be made base solely on a simple majority: if the receiver gets a majority of zeros, it decides that the information sent was a zero; otherwise, he decides it was a one. There can be no draw, as the number of bits is odd, and this code can correct a single bit of error.

Using that kind of error-correcting scheme, the communication system may have a final probability of error that is the same as another system which does not use coding, but instead uses more power to transmit the data and to overcome the noise. The power saved by the coding is usually called *coding gain* [Golden e Jacobsen 2006].

That code mentioned above could also be adjusted and extended so that a greater number of bits could be corrected, and normally all codes have different choices of parameters, such as the number of redundant bits. If we set $C(x) = xxxxx$ and leave the receiver to still decide based on majority, now there are 4 redundancy bits and the code can correct up to two errors. Sometimes, also, there can be no redundancy at all -and the codes do not do any correcting-, but they are there to keep the standards of implementation whether you need correction or not.

The Reed-Solomon codes are much more complex than the example given above. Nonetheless, the DSL systems use them for the same purposes, and the description given here should be enough for the understanding of the work, as it does not depend on the coding part of the standards. Any extra information about Reed-Solomon codes and error-correcting codes in general is found extensively in the literature, and for the specific details of the DSL implementation of those codes the reader is referred to [Golden e Jacobsen 2006].

2.1.2 Interleaving

Besides being coded, the information to be sent can be also interleaved among those codewords into which they were mapped [ITU-T 2009]. That means that the bits of the codewords are not necessarily sent in the usual order, but are exchanged among the codewords, sent, and then put back in order at the receiver [ITU-T 2009].

Consider the above code with 4 redundancy bits: if two messages x and y need to be sent, what is actually sent is their codewords $C(x) = xxxxx$ and $C(y) = yyyyy$, and a way to interleave the data could be to send $xyxyxyxyxy$. In that case, if up to four errors occur in sequence, at maximum two of them are of the same codeword, so the receiver can still correctly obtain the data - as the code corrects up to two errors in the same codeword.

As it is common to observe errors coming in bursts and several bits being corrupted in sequence [Nedev 2003], doing that kind of interleaving is usual and it can make the communication system more reliable. However, the DSL standards also allow the choice of not using any interleaving [ITU-T 2009].

For more information about interleaving, the reader is referred to [Gallager 1968].

2.1.3 Modulation

After the information has been coded and (optionally) interleaved, it is sent through the DSL line, made of twisted-pair of copper, as electric signals. That copper line, however, has already been used for the transmission of telephone signals, using the baseband from about 300Hz to 4kHz, much before the creation of DSL[Golden e Jacobsen 2006]. Therefore, the signals transmitted by the DSL system cannot be baseband, otherwise they would overlap the voice spectrum of the signal already on the line. They should, instead, be sent at higher frequencies of the spectrum. The standards then claim that the DSL signal should amplitude-modulate high-frequency sine-waves, called carriers, through a *Digital Quadrature Amplitude Modulation* scheme - digital QAM [ITU-T 2009].

Amplitude modulation is the changing of the amplitude of some carrier wave accordingly to some other signal. That digital QAM modulation uses two orthogonal waves at a certain frequency, a sine and a cosine. Two different information signals can then amplitude-modulate each one of the orthogonal waves separately, at the same frequency. Still, they can be properly recovered at the receiver through operations such as the dot product of the signal received and the carrier waves, separating the orthogonal components of the signal sent.

Furthermore, as mentioned, the modulation scheme is digital: not all amplitude values for the carrier waves are considered valid. The sender only sends waves with amplitudes from a discrete and finite set, and noise may or not change the amplitude of the wave to some other value. The receiver, at the other end, decides which valid amplitude was sent by choosing, among the valid amplitudes, the one closest to the amplitude received. As the valid amplitudes in that set are spaced by some value $delta$, the noise can add up to $\frac{delta}{2}$, and the decision will still be correct. That distance between valid amplitudes is another way of recovering from errors that the DSL systems -and all the digital communications- use. Moreover, it is a fundamental concept used for the mitigation technique proposed later.

2.1.4 Multiplexing

The DSL communication systems transmit the data through the modulation of many different carriers, each of them at a different frequency. By doing that, different numbers of bits can be assigned to each carrier, based on how good is the *Signal to Noise Ratio* at that frequency. If the channel is not good at some frequency, there is the choice of assigning few or none bits to it, so that the better frequencies can be used instead - and thus the information is better protected.

Those different carriers are all sent to through the same channel, and so there must be a way of separating their signals. That sending of many different signals through the same channel, separating them somehow, is called *multiplexing*. There are several ways to do that, and one of them is the separation done in the time domain, called *Time Division Multiplexing or TDM*, in which the signals alternate and are sent during different instants of time. Another one, called *Frequency Division Multiplexing or FDM*, is the separation of signals in the frequency domain, in which the signals modulate carriers of different frequencies, and can be separated though filtering etc. This is the kind of multiplexing done by the DSL systems.

The standards specify that, for DSL systems, all the carriers used should have frequencies that are multiples of a fundamental one - they are called harmonics of the fundamental frequency. For ADSL2+, the fundamental frequency for the orthogonal set is chosen to be at 4kHz, and consequently all the other carriers are spaced 4kHz in between in the frequency spectrum[ITU-T 2009]. As harmonic sinusoids are all orthogonal in their fundamental period, t_0, it makes possible the separation of the signals at the receiver and the independence of the all the carriers, if the separation is done every t_0 seconds. The system, then, uses a kind of multiplexing called *Orthogonal Frequency Division Multiplexing*. Moreover, being the carriers orthogonal and the system a discrete one, that scheme is also usually called *Discrete Multitone -DMT*.

2.1.5 Synchronization

A Digital Subscriber Line is composed by a pair of modems: one client modem at the *Customer Premises* and an equipment called *DSLAM* at the Central Office. The data traffic sent from the Central Office to the Customer Premises is called *downstream* traffic; the one from the Customer Premises to the Central office is called *upstream traffic*.

Before the DSLAM and the modem can actually communicate sending user data, the modems needs to perform what is called synchronizing. Basically, both sides need to decide how many and which are the valid amplitudes for the sine and cosine waves. There is usually a power constraint, and then, for each frequency, a voltage range is divided into pieces, depending on how many are the valid amplitudes. If a great number of amplitudes is chosen, that voltage range will be divided into really small pieces, and the spacing in between them, δ, will be small. This makes the communication system more susceptible to errors, as the noise can easier overcome the spacing between the valid amplitudes.

Thus, frequencies with low *Signal to Noise Ratio Margin* will have less amplitude values assigned to it, because they are more susceptible to error. That decision is the result of the calculation of how many bits should be allocated to each frequency, so that there is not significant loss, assuming an *Additive White Gaussian Noise* channel. The number or bits for each frequency is determined by a formula based on *Shannon's Coding Theorem*, and can be found in [Golden e Jacobsen 2006]:

$$b_k = \log_2 \left(\frac{SNR_k \cdot \gamma_c}{\gamma_m \cdot \Gamma(Pe)} + 1 \right) \tag{2.1}$$

where SNR_k is the ratio between the transmitted signal power and the noise power at the frequency of the kth carrier, γ_m is the value of the parameter *SNR Margin*, γ_c is the value of the *coding gain* given by the code used and $\Gamma(P_e)$ is the *Shannon Gap* of the desired probability of error P_e.

Besides deciding how many bits - or how many valid amplitudes - each carrier may have, the modems should also agree on a mapping from stream of bits to pairs of amplitudes - one for the sine and another for the cosine. For example, the sender and the transmitter can agree that if the data to be sent is *bbb*, then the sinusoids at some frequency k, with three bits allocated to it, will be sent with amplitudes equal to x for the cosine and y for the sine. If the receiver then sees the sinusoids arriving with those amplitudes, he knows the sender wanted to send *bbb* - and here we do not consider the effect of noise. Each pair (x, y) of valid amplitudes can be seen as a complex number $p = x + y \cdot i$

and is usually called a QAM *constellation point*.

2.1.6 DMT Symbol and the Discrete Fourier Transform

The information in ADSL2+ systems is sent at a synchronous rate of 4kHz, because the carriers are only orthogonal in the fundamental period as mentioned earlier[ITU-T 2009]. So, every 1/4 ms, a set of complex numbers which represent the amplitudes of all the carriers is chosen to be sent - this is called the *DMT Symbol*.

After the mapping from bits to complex numbers, many different analog carrier waves could be generated, modulated, and then their sum could the sent through the channel. However, the ADSL2+ standard established that the modulation and transmission scheme should be based on the *Discrete Fourier Transform* [ITU-T 2009].

Suppose there are T carriers and that all complex numbers $S[k]$ have been assigned to the T different carriers for a given DMT symbol. We can then see those complex number as the Fourier coefficients of a discrete-time signal, $s[n]$. That signal, in the discrete-time domain, can thus be obtained by the *Inverse Discrete Fourier Transform - IDFT*:

$$s[n] = \frac{1}{T} \cdot \sum_{n=0}^{T-1} S[k] \cdot \exp\left(\frac{j \cdot 2\pi \cdot k \cdot n}{T}\right) \tag{2.2}$$

where j is the imaginary unit - a notation which will be used throughout the whole text from now on.

Next, that discrete signal can be then transformed into a analog signal, with fundamental frequency at 4kHz, using an adequate discrete-to-analog converter.

However, there is not necessarily what is called *Hermitian symmetry* of those coefficients, and the resulting signal may not be real - thus it could not be transmitted through the analog medium which is the telephone line cable. For this reason, the coefficients $S[k]$ are repeated after the last one, doubling the size of coefficients, in such a way that symmetry exists, resulting in a necessarily real time-domain signal. Thus, if there are N carriers and N complex-numbers assigned to them, $T = 2N$. According to [ITU-T 2009], $N = 512$ and $T = 1024$ for ADSL2+.

For ADSL2+, that discrete signal, made of T samples, is what represents each DMT symbol to be sent through the lines. Every $\frac{1}{4}$ ms a set of 512 complex numbers is assigned to the carriers, the coefficients are doubled, and the 1024 samples are then sent in analog signal, after the conversion. That analog signal, when received by the other communicating side, is sampled again and its T-point DFT is calculated in order to obtain the T coefficients - but then only the first half is considered, excluding the artificially inserted ones.

The total number of bits sent by each symbol, L, is then given by the sum of the bits allocated to each carrier:

$$L = \sum_k b_k \tag{2.3}$$

2.2 Related work

The research area of this work, that involves the mitigation of the effect of impulsive noises on communication systems, has some worked published as scientific papers. Those papers relate to this work, but, in general, they are either more theoretical and limited to a statistical modeling of the effect or propose something for the actual mitigation, but not really feasible in practice with the standard-compliant modems of nowadays.

The works of [Henkel e Kessler 1994], [Levey e McLaughlin 1999] and many others are of the more theoretical type; they do not actually propose a method for the recuperation of the line when it is being affected by that kind of impulsive noises. The author of [Nedev 2003] also makes an analysis of the effect of impulse noise on DSL lines based on many different models from different authors, but no practical measure is proposed. Furthermore, their models are based in statistical data and have the disadvantage of not having the possibility for proving their correctness based on some assumption.

The author of [Sedarat et al. 2005], on the other hand, proposes a compensation through a change on the *Signal to Noise Ratio Margin* -which is a protection parameter-, in case an impulsive noise has been detected. However, the modeling he proposes does not seem clear nor based on an analytical deduction. Too much is assumed about the nature of the noise, but not much is justified. It is assumed, for example, that the "variation of the impulse amplitude is negligible comparing to the background noise", but there is no measured data or intuitive or mathematical reasoning to support that claim. Moreover, the method he proposes for the detection of noise parameters -as its variance and its mean- does not seem feasible on practice, without the changing of the modems' proprietary implementations.

This work, then, differentiates itself from the related works in two ways: it models the noises' effect based on precise assumptions about its time-domain characteristics and, furthermore, proposes a feasible method for the recovering of the lines affected by those noises, through the changing of a single standard-established parameter called SNR Margin.

3 *Development*

In order to develop the work and derive a model for the effect of REIN noises on the symbols, for each frequency, some considerations need to be made first. As the DSL communication systems are discrete, and the DMT symbols received come as the result of the sampling of analog signals, all the modeling done here is made in the discrete domain of the samples - called from now on discrete-time domain even though the independent variable is not time [Oppenheim e Schafer 1989]. Also, there is a need to first define precisely what is meant by "REIN". After a definition of what is the discrete-time signal representing REIN, a model is derived analytically based solely on assumptions of its discrete-time domain properties.

Throughout the rest of the text, *Repetitive Impulsive Noise -REIN- is defined as* the kind of *noise characterized by a train of impulses, each one equal to a signal $I[n]$ of duration $t > 0$, all created by some kind of fixed interference at a periodic rate. Besides, during the time between the impulses, the noise is zero.*

Defining mathematically, let $I[n]$ be the one of the REIN impulses considered in the discrete-time domain due to the sampling. Then,

$$R[n] = \sum_{m} I[n + m \left\lceil \frac{F}{f} \right\rceil]$$

where f is the impulse generation rate and F is the DSL sampling rate -both in Hz-, so that $\frac{F}{f}$ gives how many samples apart are the impulses.

3.1 REIN's Effect Modeling

Assume, in the discrete-time domain, that the transmitted signal is X, the background noise is B, the REIN noise is R and the received signal is $Y = X + N + R$, and let us suppose that a DSL symbol is made up of T samples if considered in the discrete-time domain.

The constellation point for the kth carrier, $P[k]$, demodulated from a certain symbol of the received signal, is thus obtained by a T-point DFT of the received samples, $Y[n]$:

$$
\begin{aligned}
&P[k] = \sum_{n=0}^{T-1} Y[n] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T}\right) \\
&= \sum_{n=0}^{T-1} (X[n] + B[n] + R[n]) \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T}\right) \\
&= \textstyle\sum_{n=0}^{T-1} X[n] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T}\right) + \sum_{n=0}^{T-1} B[n] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T}\right) + \sum_{n=0}^{T-1} R[n] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T}\right)
\end{aligned}
\tag{3.1}
$$

As we can see, the DFT is linear, and the error caused by the REIN, from now on called $E_R[n]$, is given by T-point DFT of its own samples:

$$
E_R[n] = \sum_{n=0}^{T-1} R[n] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T}\right) \tag{3.2}
$$

Furthermore, if the noise is of the REIN kind, it is made by impulses apart from each other, and thus sometimes it affects a symbol, and sometimes it does not. If the impulse generation is synchronized with the symbol sampling, the effect on the demodulation is always the same, deterministic, and not a random variable - at least approximately-, as the impulses always begin at the same time relatively to the beginning of the symbol sampling - i.e $n = 0$. Otherwise, if there is no synchronization, the REIN impulses affect the demodulation at different times of the symbol sampling, as illustrated by figures 1 and 2.

So let us assume that, when an impulse appears, it is the signal $I[n]$ delayed by some random variable $\Phi \in [-t, T]$ in respect to beginning of the symbol sampling - i.e $n = 0$. Assume also that the impulses occur far apart from each other, so that when we look at the effect of a single impulse on the demodulation of a certain symbol we can consider $I[n] = 0$, for $n < 0$. Under those assumptions, the effect of an impulse, $E_{imp}[k, \Phi]$, on the kth carrier's constellation point of a certain DMT symbol is then given by:

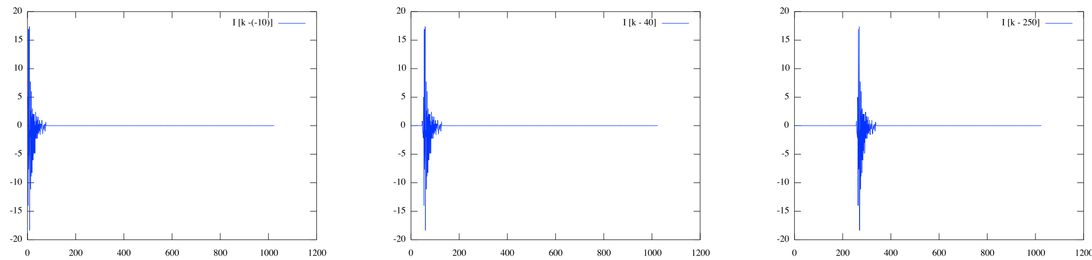

Figure 1: $I[n-\Phi]$ for $\Phi \in \{-10, -40, 250\}$

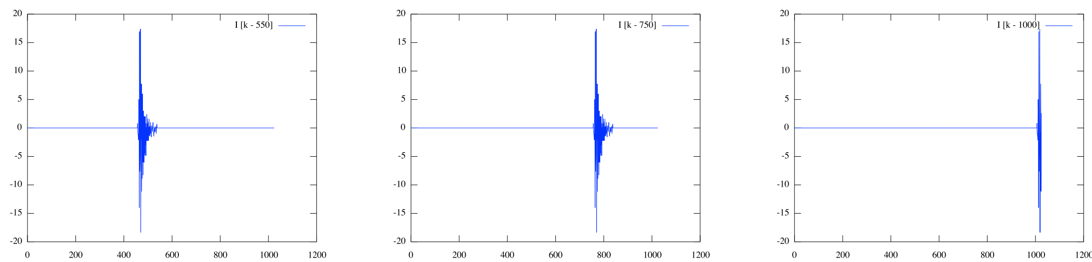

Figure 2: $I[n-\Phi]$ for $\Phi \in \{450, 750, 1000\}$

$$
\begin{aligned}
E_{imp}[k,\Phi] &= \sum_{n=0}^{T-1} I[n-\Phi] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T}\right) \\
&= \sum_{n=0}^{T-1} I[n-\Phi] \cdot \cos\left(\frac{2\pi \cdot n \cdot k}{T}\right) + j \cdot \sum_{n=0}^{T-1} I[n-\Phi] \cdot \sin\left(\frac{2\pi \cdot n \cdot k}{T}\right) \\
&= \sum_{x=-\Phi}^{T-\Phi-1} I[x] \cdot \cos\left(\frac{2\pi \cdot n \cdot (x+\Phi)}{T}\right) + j \cdot \sum_{x=-\Phi}^{T-\Phi-1} I[x] \cdot \sin\left(\frac{2\pi \cdot n \cdot (x+\Phi)}{T}\right) \\
&= \sum_{x=0}^{T-\Phi-1} I[x] \cdot \cos\left(\frac{2\pi \cdot n \cdot (x+\Phi)}{T}\right) + j \cdot \sum_{x=0}^{T-\Phi-1} I[x] \cdot \sin\left(\frac{2\pi \cdot n \cdot (x+\Phi)}{T}\right), \Phi \geq 0.
\end{aligned}
\tag{3.3}
$$

Thus, if $\Phi \geq 0$, the real and imaginary parts of the effect of an impulse on the kth carrier are equal to:

$$Re\{E_{imp}[k,\Phi]\} = \sum_{x=0}^{T-\Phi-1} I[x] \cdot \left\{\cos\left(\tfrac{2\pi \cdot n \cdot x}{T}\right) \cdot \cos\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) - \sin\left(\tfrac{2\pi \cdot n \cdot x}{T}\right) \cdot \sin\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right)\right\}$$
$$= \cos\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) \cdot \left\{\sum_{x=0}^{T-\Phi} I[x] \cdot \cos\left(\tfrac{2\pi \cdot n \cdot x}{T}\right)\right\} - \sin\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) \cdot \left\{\sum_{x=0}^{T-\Phi} I[x] \cdot \sin\left(\tfrac{2\pi \cdot n \cdot x}{T}\right)\right\}$$
$$= \cos\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) \cdot \left\{\sum_{x=0}^{t} I[x] \cdot \cos\left(\tfrac{2\pi \cdot n \cdot x}{T}\right)\right\} - \sin\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) \cdot \left\{\sum_{x=0}^{t} I[x] \cdot \sin\left(\tfrac{2\pi \cdot n \cdot x}{T}\right)\right\}, t \leq T - \Phi - 1. \tag{3.4}$$

$$Imag\{E_{imp}[k,\Phi]\} = \sum_{x=0}^{T-\Phi-1} I[x] \cdot \left\{\sin\left(\tfrac{2\pi \cdot n \cdot x}{T}\right) \cdot \cos\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) + \sin\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) \cdot \cos\left(\tfrac{2\pi \cdot n \cdot x}{T}\right)\right\}$$
$$= \cos\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) \cdot \left\{\sum_{x=0}^{T-\Phi} I[x] \cdot \sin\left(\tfrac{2\pi \cdot n \cdot x}{T}\right)\right\} - \sin\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) \cdot \left\{\sum_{x=0}^{T-\Phi} I[x] \cdot \cos\left(\tfrac{2\pi \cdot n \cdot x}{T}\right)\right\}$$
$$= \cos\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) \cdot \left\{\sum_{x=0}^{t} I[x] \cdot \sin\left(\tfrac{2\pi \cdot n \cdot x}{T}\right)\right\} - \sin\left(\tfrac{2\pi \cdot n \cdot \Phi}{T}\right) \cdot \left\{\sum_{x=0}^{t} I[x] \cdot \cos\left(\tfrac{2\pi \cdot n \cdot x}{T}\right)\right\}, t \leq T - \Phi - 1. \tag{3.5}$$

If we now define:

$$P[k] \doteq \sum_{x=0}^{t} I[x] \cdot \cos\left(\frac{2\pi \cdot n \cdot x}{T}\right) \tag{3.6}$$
$$Q[k] \doteq \sum_{x=0}^{t} I[x] \cdot \sin\left(\frac{2\pi \cdot n \cdot x}{T}\right) \tag{3.7}$$
$$M[k] \doteq \sqrt{P^2[k] + Q^2[k]} \tag{3.8}$$
$$\Theta[k] \doteq \arctan\frac{P}{Q}$$
$$\Psi[k,\Phi] \doteq \frac{2\pi \cdot n \cdot \Phi}{T} + \Theta \tag{3.9}$$

We finally have:

$$Re\{E_{imp}[k,\Phi]\} = M[k] \cdot \cos\left(\frac{2\pi \cdot n \cdot \Phi}{T} + \Theta\right), t \leq T - \Phi - 1, \Phi \geq 0.$$
$$Imag\{E_{imp}[k,\Phi]\} = M[k] \cdot \sin\left(\frac{2\pi \cdot n \cdot \Phi}{T} + \Theta\right), t \leq T - \Phi - 1, \Phi \geq 0.$$
$$E_{imp}[k,\Phi] = M[k] \cdot \{\cos(\Psi[k,\Phi]) + j \cdot \sin(\Psi[k,\Phi])\}, t \leq T - \Phi - 1, \Phi \geq 0. \tag{3.10}$$

That means that, if both the conditions a) "$\Phi \geq 0$" and b) "$t \leq T - \Phi - 1$" are satisfied, the effect of a REIN impulse on the kth carrier of a certain DMT symbol, as equation 3.10 shows, has constant magnitude in respect to the delay Φ and is equal to $M[k]$ as given by 3.8. Furthermore, its phase, $\Psi[k,\Phi]$ in equation 3.9, is an affine function

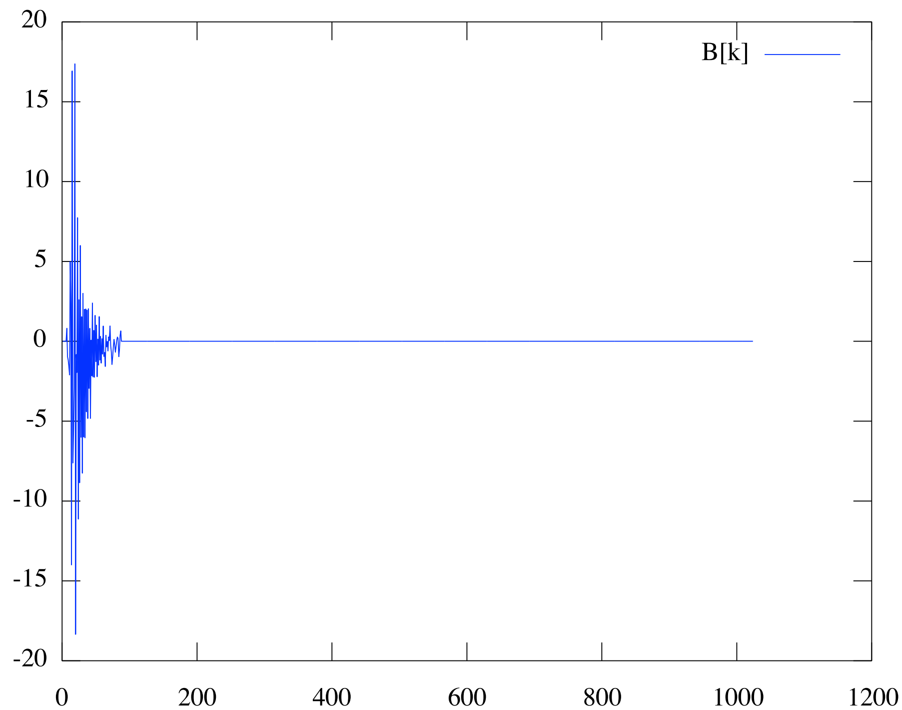

Figure 3: The waveform of the noise impulse

of the random variable Φ and has the same kind of distribution Φ does.

The condition "a" is not satisfied for $\Phi \in [-t, -1]$, and condition "b" is not satisfied for $\Phi \in [T-t, T-1]$. That gives at maximum a percentage of $\frac{2t}{T+t}$ of the values for Φ. If we assume Φ to be homogeneously distributed and T to be much greater than noise impulses' length in samples, t, both conditions are thus satisfied for most of the time. According to [ITU-T 2009], $T = 1024$ for ADSL2+, and the noise described on [ITU-T 2001] and illustrated by figure 3 has an impulse duration in number of samples, t, about a tenth of T. So the model described here should apply reasonably well for that scenario.

On the other hand, if the noise impulses' length are not much smaller than T, the variation of the magnitude may not be negligible. In that case, the condition "$\Phi \geq 0$" and "$t \leq T - \Phi - 1$" may be satisfied only for a few times and the magnitude of effect of the impulse may then depend on the delay and be a function of the random variable Φ for most of the time.

Finally, the error caused by the REIN noise on the demodulation of the DMT symbols will alternate, accordingly to the rate of impulse generation f and the frequency of the DSL sampling F, between none -in between the impulses- and $E_{imp}[k, \Phi]$ -when impulses appear:

$$E_R[k,\Phi] = (1 - \frac{f}{F}) \cdot 0 + \frac{f}{F} \cdot E_{imp}[k,\Phi] \tag{3.11}$$

On a scenario in which the background noise, B, is negligible in comparison to the REIN R, if for some k $M[k]$ is greater than the distance between the QAM points of that carrier, $delta_k$, we can consider approximately that there is an error independently of phase. This is is just an approximation, as the points are spaced over squares, and their diagonals, for example, are greater than the radius of the circle around the points; but this consideration simplifies the modeling and clearly works as an upper bound. We then consider only the magnitude, $M[k]$, of the effect of the impulses on the QAM-constellations.

Then, if $M[k] > delta_k$ for some k on that scenario, it is reasonable to assume that, every time an impulse is present during a symbol sampling, there is an error on that carrier k. In that case, the probability of symbol error thus depends basically on the ratio between the rate of impulse generation, f, and the symbol sampling frequency, F:

$$P_{sym} \sim \frac{f}{F} \tag{3.12}$$

As the xDSL implementations normally choose a bit loading algorithm so that all $delta_k$ are equal and chosen in order to achieve some wanted probability of error for the whole symbol, based on the assumption that the noise is AWGN[Golden e Jacobsen 2006], it is clear that, in order to fully mitigate that noise impact, the parameter *SNR Margin* should then be chosen so that $delta_i$ is greater than $max_{i \in [1,T]}^{\{M[k]\}}$ for all i. Section 3.3 describes it further.

3.2 Background Noise Modeling

Now, if we consider the scenario with both REIN and some other background noise, and as the magnitude of the effect of the noises is what is being considered, it is also in need to obtain a description of the modulus of the effect the background noise causes on the symbols.

So let us assume that, at every instant of the symbol sampling, the noise amplitudes

B[n] are independent random variables equally distributed, with mean μ_B and standard deviation σ_B. Thus, as the demodulation done through the DFT is linear, the effect $E[k]$ of the noise on the kth carrier of a certain symbol is a random variable function and can be calculated as:

$$E_B[k] = \sum_{n=0}^{T-1} B[n] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T}\right)$$
$$\left|\sum_{n=0}^{T-1} B[n] \cdot \cos\left(\frac{2\pi \cdot n \cdot k}{T}\right) + j \cdot \sum_{n=0}^{T-1} B[n] \cdot \sin\left(\frac{-2\pi \cdot n \cdot k}{T}\right)\right| \tag{3.13}$$

As we assumed $B[n]$ to be i.i.d, we have a sum of i.i.d random variables, multiplied by constants, at the real and the imaginary parts of the inner sum of (3.13). By the Central Limit Theorem, those sums approach a Gaussian distribution as T goes to infinity, independently of the equal distribution of the random variables at the single instants, $B[n]$. According to [ITU-T 2009], $T = 1024$ for ADSL2+ lines, which is relatively large, so that makes good a approximation. In that case, the noise in the time domain can be Gaussian or not, but its effect on the real and imaginary parts of the demodulated DMT symbols is still Gaussian.

As each one of those parts, real and imaginary, is a sum of random variables as mentioned before, their mean is the sum of the individual means. As we assumed $B[n]$ to be i.i.d, the means of the real and imaginary parts are given by:

$$\begin{aligned} &= \sum_{n=0}^{T-1} \mu_B \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T}\right) \\ &= \mu_B \sum_{n=0}^{T-1} \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T}\right) \\ &= \mu_B \cdot 0 = 0 \end{aligned}$$

Moreover, their variances are equal, called σ^2, given by what follows:

$$\begin{aligned}
\sigma^2 &= \sigma_B^2 \sum_{n=0}^{T-1} \cos^2\left(\frac{2\pi \cdot n \cdot k}{T}\right) \\
&= \sigma_B^2 \sum_{n=0}^{T-1} \sin^2\left(\frac{2\pi \cdot n \cdot k}{T}\right) \\
&= \sigma_B^2 \cdot T
\end{aligned}$$

And thus their standard deviation σ is equal to $\sigma_B \cdot \sqrt{T}$.

Being the real and imaginary parts zero-mean random variables with standard deviation σ, if we force the assumption of them being independent random variables, the resulting magnitude of the effect of the background noise on the kth carrier, $|E_B[k]|$, has a Rayleigh distribution with parameter σ. Its expectancy and variance, thus, are:

$$E\{|E_B[k]|\} = \sigma \cdot \sqrt{\frac{\pi}{2}} \qquad\qquad V\{|E_B[k]|\} = \sigma^2 \cdot \frac{(4-\pi)}{2} \tag{3.14}$$

Those results and the ones obtained on section 3.1 will be used next for proposing a way of mitigating the effect of the kinds of noises discussed here.

3.3 Mitigation Through Compensating Margin

This section proposes a way of mitigating the effect of some REIN that starts affecting a line after it has been synchronized. As exposed during the explanation of the problem in the Introduction, if we just resynchronize the line, it usually will not be enough to mitigate the loss in that case, as the modem assumes the noise is AWGN. Instead, we propose a way of mitigating that effect through resynchronization, too, but with a specific choice for the *SNR Margin* parameter that should be enough for the protection.

We know that, as mentioned by [Golden e Jacobsen 2006], the bit allocation algorithms implemented by the modems should consider the SNR at each frequency in order to choose the number of bits for each carrier. There is no clear definition, however, of how that value of SNR should be measured. On the other hand, it is clearly defined that the modulation and demodulation schema should be implemented using the DFT. So, if the SNR measurements are to help the bit allocation in any way, it seems reasonable to assume that the noise power measurements are done through a similar method - sampling the noise and projecting the resulting discrete-time signal into the discrete-time signal corresponding to each carrier. Furthermore, as the SNR is for each frequency, independent of phase, it makes sense to assume that the power measure is related to the magnitude of the DFT. Finally, the result of the measurement of the noise power should also be the calculated through some averaging, and not the result of a single calculation. Otherwise outliers could be common and the measurements could be inaccurate.

From now on we then assume that a single measurement, $U_Q[k]$, done by the modem, of the power of some signal Q at the frequency of the kth carrier is the square of of the magnitude of its T-point DFT, $D_Q[k]$:

$$\begin{aligned} U_Q[k] &= \{|D_Q[k]|\}^2 \\ U_Q[k] &= \left\{ \left| \sum_{n=0}^{T-1} Q[n] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T} \right) \right| \right\}^2 \end{aligned} \tag{3.15}$$

We assume also that, for the power of the signal Q, for each frequency k, the final value considered by the modem is an average of many measurements of the above taken over a long enough period so that we can approximate it by the expectancy of the random variable $U_Q[k]$:

$$P_Q[k] = E\{U_Q[k]\} \tag{3.16}$$

Now let us consider the power measurements done by the modem for the composite signal, S, made of the REIN - R, with impulse duration t, in number of samples, much smaller than T- and some background noise, B. First, $U_S[k]$ is given by:

$$
\begin{aligned}
U_S[k] &= \left\{ \left| \sum_{n=0}^{T-1} S[n] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T} \right) \right| \right\}^2 \\
&= \left\{ \left| \sum_{n=0}^{T-1} (B[n] + R[n]) \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T} \right) \right| \right\}^2 \\
&= \left\{ \left| \textstyle\sum_{n=0}^{T-1} B[n] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T} \right) + \sum_{n=0}^{T-1} R[n] \cdot \exp\left(\frac{-j \cdot 2\pi \cdot n \cdot k}{T} \right) \right| \right\}^2 \\
&= \{ |D_B[n] + D_R[n]| \}^2
\end{aligned}
\tag{3.17}
$$

As $D_B[n]$ and $D_R[n]$ are complex numbers, we have that the square of the magnitude of their sum is given by the cosine law, and that implies:

$$
\begin{aligned}
U_S[k] &= \{ |D_B[n] + D_R[n]| \}^2 \\
&= |D_B[n]|^2 + |D_R[n]|^2 + 2 \cdot |D_B[n]| \cdot |D_R[n]| \cdot \cos \Psi
\end{aligned}
\tag{3.18}
$$

where Ψ is the angle between the two complex numbers $D_B[n]$ and $D_R[n]$, which varies accordingly to the them, and thus is a random variable.

The expectancy of $U_S[k]$ is thus given by:

$$
\begin{aligned}
P_S[k] &= E\{U_S[k]\} \\
&= E\{ |D_B[n]|^2 + |D_R[n]|^2 + 2 \cdot |D_B[n]| \cdot |D_R[n]| \cdot \cos \Psi \} \\
&= E\{ |D_B[n]|^2 \} + E\{ |D_R[n]|^2 \} + 2E\{ \cdot |D_B[n]| \cdot |D_R[n]| \cdot \cos \Psi \}
\end{aligned}
\tag{3.19}
$$

On section 3.1, we concluded that the the T-point DFT of the kind of REIN mentioned here for R has magnitude approximately constant for each frequency, $M[k]$, when an impulse is present, and 0 in between the impulses. If the probability of a REIN impulse being present during the measurement is $p_i = \frac{f}{F}$, that leads to expectancy of $|D_R[n]|$ and $|D_R[n]|^2$:

$$\begin{aligned} E\{|D_R[n]|\} &= (1-p_i)\cdot 0 + p_i \cdot M[k] = p_i \cdot M[k] \\ E\{|D_R[n]|^2\} &= (1-p_i)\cdot 0^2 + p_i \cdot M[k]^2 = p_i \cdot M[k]^2 \end{aligned} \tag{3.20}$$

So we can calculate the value for the power measurement $P_S[k]$ made by the modem, if we force the assumption of the random variables $|D_B[n]|$, $|D_R[n]|$ and $\cos\Psi$ being independent:

$$\begin{aligned} P_S[k] &= E\{|D_B[n]|^2\} + E\{|D_R[n]|^2\} + 2E\{|D_B[n]|\}\cdot E\{|D_R[n]|\}\cdot E\{\cos\Psi\} \\ &= E\{|D_B[n]|^2\} + p_i \cdot M[k]^2 + 2\cdot p_i \cdot E\{|D_B[n]|\}\cdot M[k]\cdot E\{\cos\Psi\} \end{aligned} \tag{3.21}$$

On the other hand, if the impulses were present during all the measurements, we would have that the expectancy $E\{|D_R[n]|\}$ would be equal to $M[k]$, and the expectancy $E\{|D_R[n]|^2\}$ would be equal to $M[k]^2$. In that case the average power measurement would give a value P' equal to:

$$P'_S[k] = E\{|D_B[n]|^2\} + M[k]^2 + 2\cdot E\{|D_B[n]|\}\cdot M[k]\cdot E\{\cos\Psi\} \tag{3.22}$$

We concluded on section 3.2 that the magnitude of the effect of the background noise, satisfying the assumptions there made, has a Rayleigh distribution. If its parameter is x, then its expectancy is $E\{|D_B[n]|\} = x\sqrt{\frac{\pi}{2}}$ and its variance is $V\{|D_B[n]|\} = x\frac{(4-\pi)}{2}$. That leads to:

$$\begin{aligned} E\{|D_B[n]|^2\} &= V\{|D_B[n]|\} + E\{|D_B[n]|\}^2 = 2x^2 \\ P_S[k] &= 2x^2 + 2x\sqrt{\frac{\pi}{2}}\cdot M[k]\cdot E\{\cos\Psi\}\cdot p_i + p_i \cdot M[k]^2 \\ P'_S[k] &= 2x^2 + 2x\sqrt{\frac{\pi}{2}}\cdot M[k]\cdot E\{\cos\Psi\} + M[k]^2 \end{aligned} \tag{3.23}$$

Then, for facility of notation, we omit the details of the variables, like the indices for the frequencies, k, and we now define $y \doteq M[k]$ and $\beta \doteq \sqrt{\frac{\pi}{2}} \cdot E\{\cos\Psi\}$. That leads to:

$$P_S[k] = 2x^2 + 2xyp_i\beta + p_iy^2 \qquad\qquad P'_S[k] = 2x^2 + 2xy\beta + y^2 \tag{3.24}$$

By doing that, we can easier analyze the function

$$f(x, y, p_i, \beta) = \frac{P'_S[k]}{P_S[k]} = \frac{(2x^2 + 2xy\beta + y^2)}{(2x^2 + 2xyp_i\beta + p_iy^2)}$$

and see how greater than $P_S[k]$ could be $P'_S[k]$. In other words, we can calculate how greater would be the power measured by the modem, if the noise impulses were present during each single measure of the calculated average, in relation to the power it actually measures. By calculating a value $\gamma = \frac{P'_S[k]}{P_S[k]}$, we can set the parameter SNR Margin to exactly γ. By definition, the SNR margin is the factor by which the noise power can increase so that the line remains protected [Golden e Jacobsen 2006]. This should protect the line from the impulses.

We can rewrite f as:

$$\begin{aligned} f(x, y, p_i, \beta) &= \frac{(2 + 2\frac{y}{x}\beta + (\frac{y}{x})^2)}{(2 + 2\beta p_i\frac{y}{x} + p_i(\frac{y}{x})^2)} \\ a \doteq \frac{y}{x} &\rightarrow f(x, y, p_i, \beta) = g(a, p_i, \beta) = \frac{(2 + 2a\beta + a^2)}{(2 + 2a\beta p_i + p_ia^2)} \end{aligned} \tag{3.25}$$

Here, $a = \frac{y}{x}$ is the ratio of the magnitude of the effect caused by a REIN impulse, $I[n]$, and the parameter of the distribution of the effect caused by the background noise, $B[n]$.

The function $g(a, p_i, \beta)$ should finally give the factor by which the impulses' power could be greater than the one measured by the modem. If the values for a, p_i and β are known, the value for the parameter *Signal to Noise Ratio Margin* has thus been found. However, that solution is not of practical interest, as those values are usually not known. So it is worth mentioning that the function $g(a, p_i, \beta)$, for fixed pi and β, has a limit in the infinite, which can be calculated using the L'Hôpital Rule, of $\frac{1}{p_i} = \frac{F}{f}$ - where f is the rate of impulse generation and F is the frequency of the symbol sampling. Moreover, its derivative is always positive - this is illustrated by figure 4 for $f = 50Hz$ and by figure 5

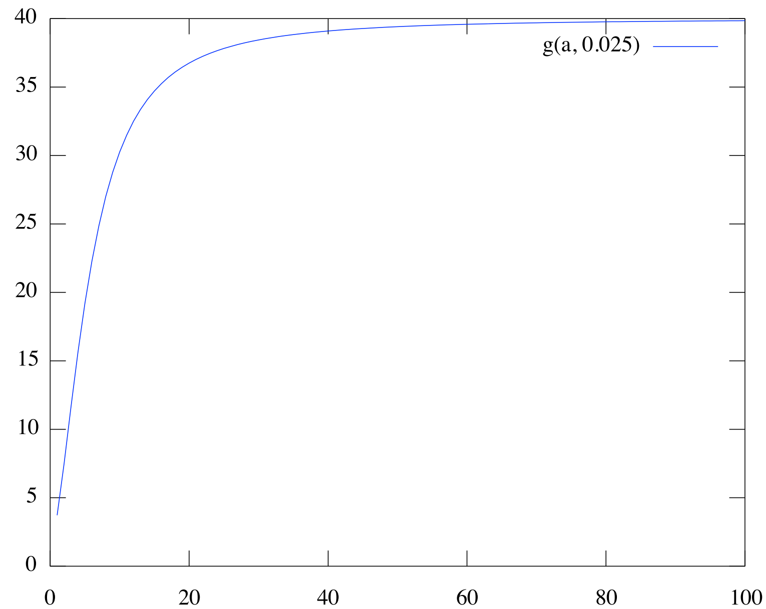

Figure 4: Behavior of the ratio of power measurements for $f = 100Hz$

for $f = 100Hz$.

This leads to the conclusion that the average power measured if the noise impulses were present during all the single measurements would be greater than the power measured by the modem by a factor of, at maximum, $\frac{F}{f}$. That factor only increases as $a = \frac{y}{x}$ gets bigger, but never exceeds $\frac{F}{f}$, independently of the power of the impulses. So, if we resynchronize the line and choose a SNR Margin equal to that factor, there should be no error at all, independently of how big is the ratio between the impulses' magnitude and the background noise's magnitude. Then the SNR Margin, in dB, can be chosen equal to:

$$SNR_{MGN_{db}} = 10 \cdot log_{10}\left(\frac{F}{f}\right) \tag{3.26}$$

Specifically, if a REIN noise, satisfying the definitions made here, starts affecting a line with an impulse generation rate of 100Hz, a resynchronization with the parameter SNR Margin set to about 16db should be able to mitigate the noise effect.

This hypothesis and other claims are validated in the following chapter.

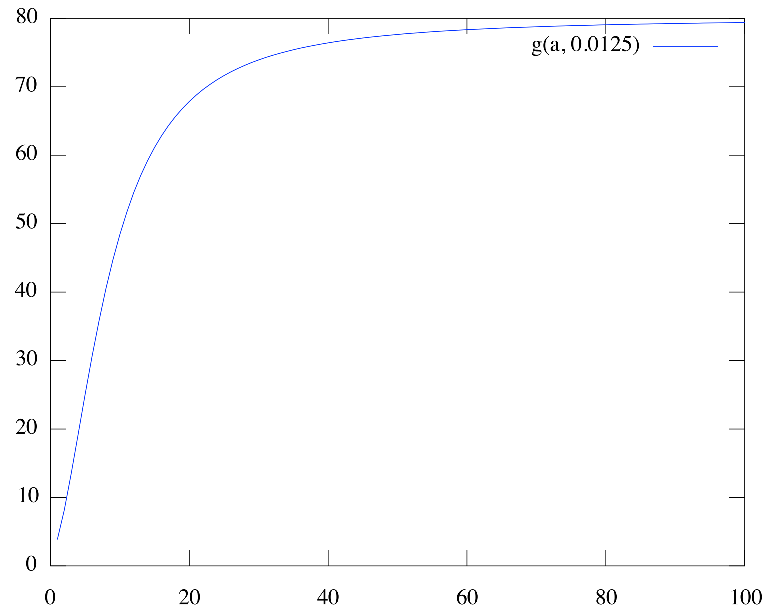

Figure 5: Behavior of the ratio of power measurements for $f = 50Hz$

4 Results and Validation

In order to validate the model and the technique for mitigation here proposed, experiments were realized on a ADSL2+ testbed, which will be described in section **??**. The noise used was the one mentioned in [ITU-T 2001], which has an impulse length of about a tenth of the DMT symbol sampling duration - $\approx 0.25 \cdot 10^{-4} s$. The waveform there listed was injected on the line at a impulse generation rate of 50Hz and 100Hz. Other noises were also used for a comparison between the results, but, from now on, when "the" noise is mentioned, it is the one mentioned by the standard, illustrated below by figure 6 and 7.

The magnitude spectrum illustrated by figure 7 belongs to waveform $I[n]$, not delayed by any sample. But the model proposed and the result in equation 3.10 say that the magnitude spectrum should be the same for $I[n-\Phi]$ for all values of Φ satisfying $t \leq T-\Phi$ and $\Phi \geq 0$. This is still validated below.

Besides the experiments realized on the testbed, the result in equation 3.10 was first validated through simulation, realized by calculating the T-point DFT of the samples of the noise. As the duration of that noise impulse is much smaller than the duration of the symbol sampling and as, with the impulse generation rates of 50Hz and 100Hz, the impulses are many symbols apart - about $\frac{4000}{50}$ symbols -, the rest of the samples was completed with zeros in order to fulfill the $T = 1024$ samples of ADSL2+.

The noise was then delayed in relation to the beginning of the sampling - i.e $n = 0$ -, and the result of its DFT was evaluated. The figures 8 and 9 illustrate the delaying, with the plots of $I[n-\Phi]$ for $\Phi \in \{-10, -40, 250, 750, 1000\}$.

4.1 Simulation

The result of 3.10 showed that, for values of Φ satisfying $t \leq T-\Phi$ and $\Phi \geq 0$, the real and imaginary parts of the error caused by the impulse $I[n-\Phi]$ on the carrier k of

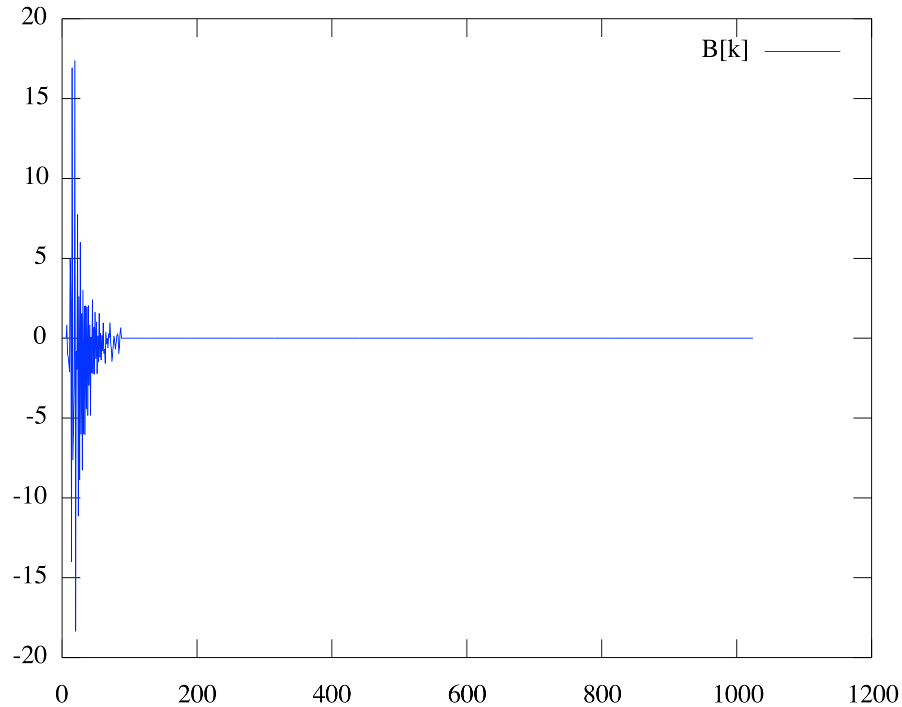

Figure 6: The waveform of the noise impulse

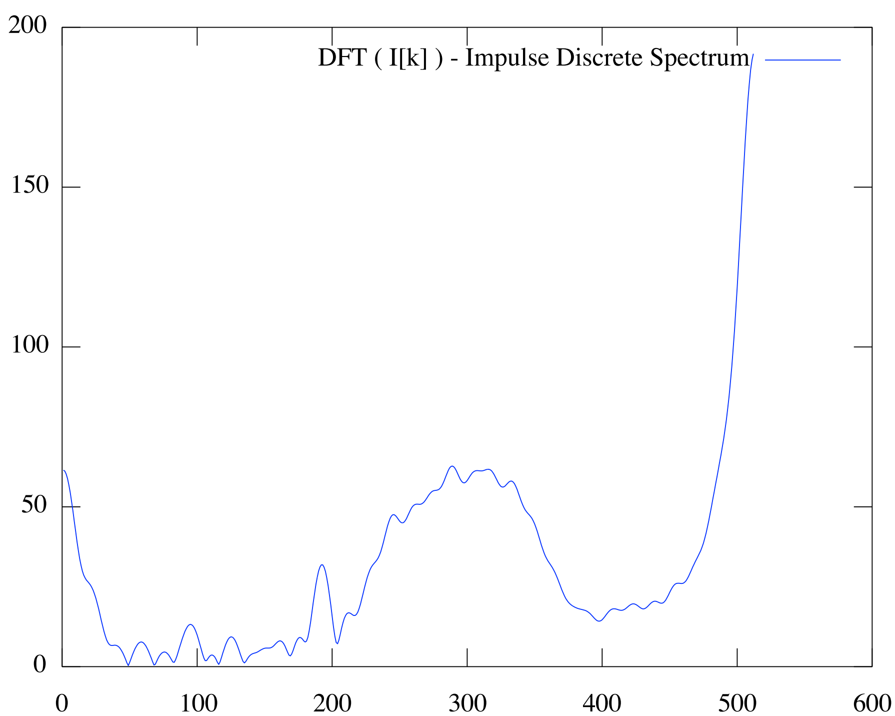

Figure 7: The Discrete Fourier Transform (magnitude) spectrum of the noise impulse

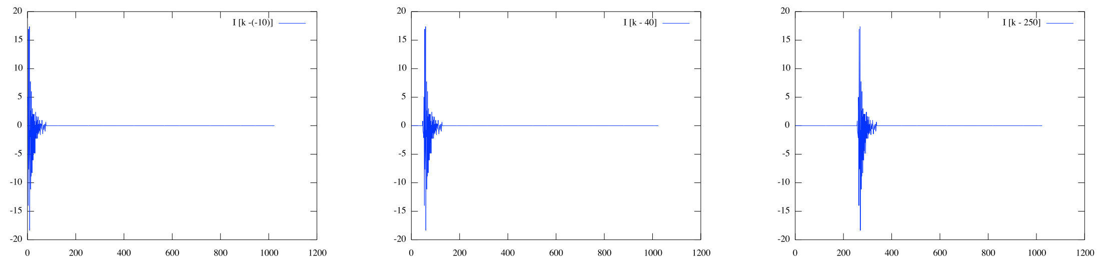

Figure 8: $I[n - \Phi]$ for $\Phi \in \{-10, -40, 250\}$

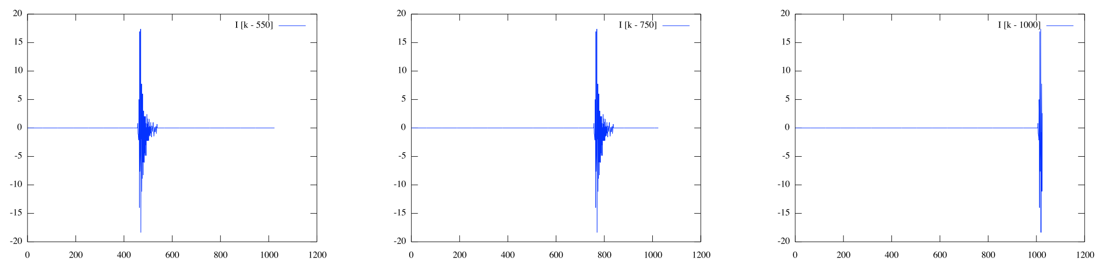

Figure 9: $I[n - \Phi]$ for $\Phi \in \{450, 750, 1000\}$

a certain symbol should vary as sinusoids, with frequencies depending on the carrier, and the error's magnitude should be constant. The results of calculating the T-point DFT of $I[n - \Phi]$ for $\Phi \in [-t, T]$ are shown below as a function of the delay, Φ, in number of samples, for several frequencies, by figures 10, 11, 12 and 13, with the real and imaginary parts on the left and the magnitude on the right. The y-axis is not normalized to any particular scale, but this should be of no importance to the discussion.

The point here is that, for values of Φ satisfying $t \leq T - \Phi$ and $\Phi \geq 0$, the result is

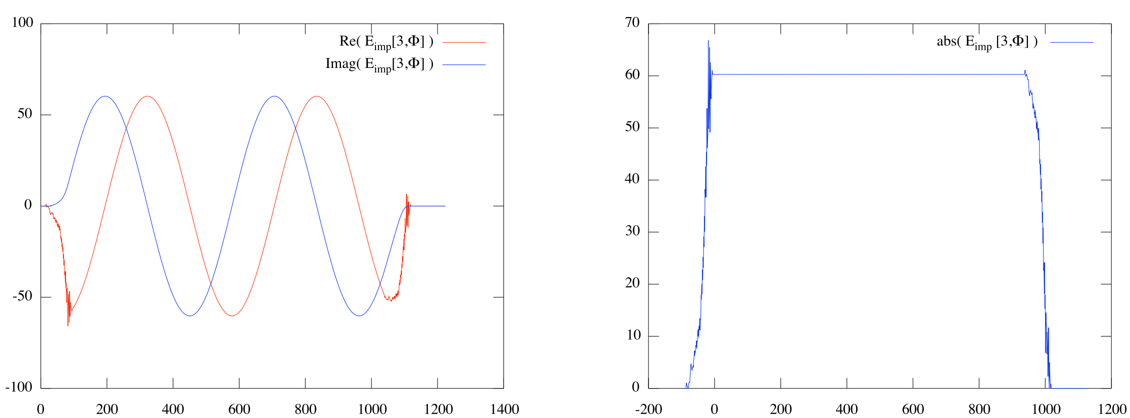

Figure 10: Impulse effect behavior for the 3rd carrier, as a function of the delay Φ

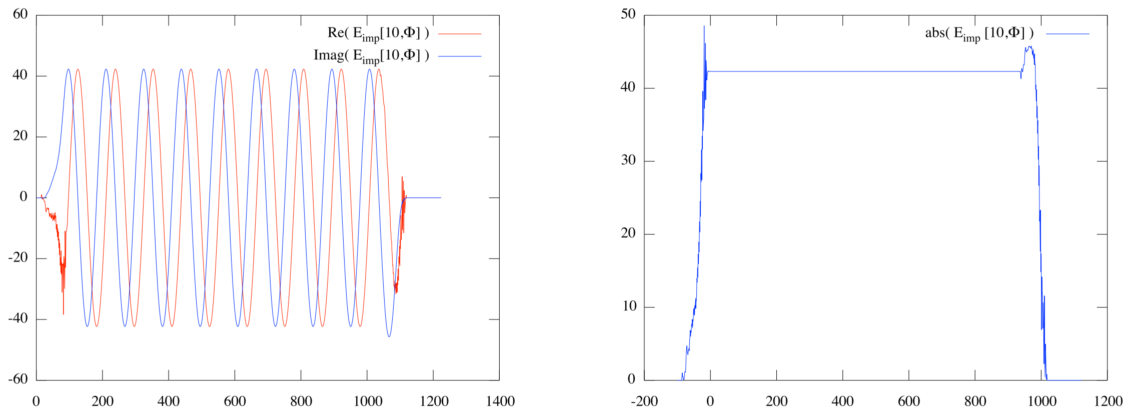

Figure 11: Impulse effect behavior for the 10th carrier, as a function of the delay Φ

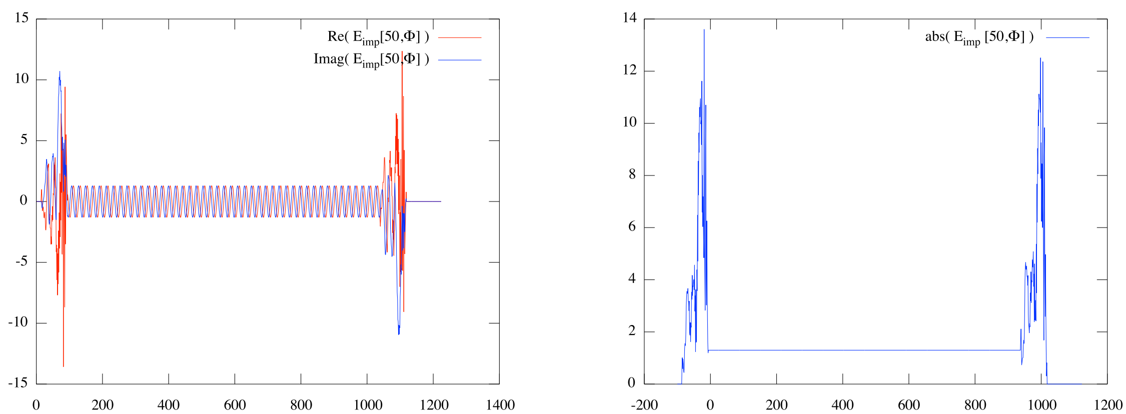

Figure 12: Impulse effect behavior for the 50th carrier, as a function of the delay Φ

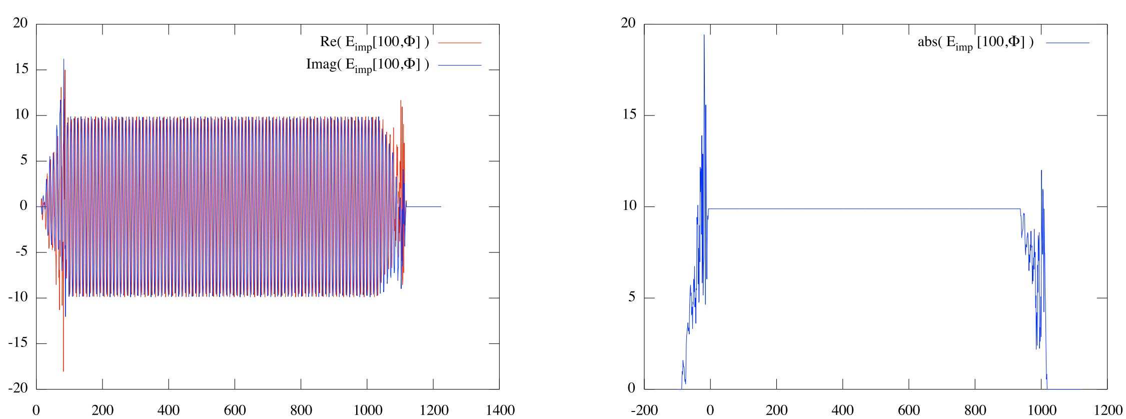

Figure 13: Impulse effect behavior for the 100th carrier, as a function of the delay Φ

really close to the predicted by the model, as shown by the figures. So, if the assumption of Φ being homogeneously distributed is good enough, the error caused by the noise on the experiments should behave approximately like predicted too.

In order to fully validate the research, the experiments briefly mentioned before were performed using the a testbed. The descriptions of the testbed and the experiments are given next.

4.2 Testbed and Experiment Description

To test the hypothesis, a final validation using an ADSL2+ testbed was performed. The testbed is composed by two Speedtouch modems, communicating through a line emulator capable of emulating lines with different loop lengths - which is the distance between the modems - subject to a waveform injector.

One of the modems played the role of the client, and the other one, the Central Office DSLAM. For each experiment, the modems were synchronized with some choice for the protection parameters *Signal to Noise Ratio Margin*, *LoopLength*, *Impulse Noise Protection*, and a video stream was sent from the DSLAM to the client.

For conclusions about loss, the metric utilized was the loss of IP datagrams, using traces of the dumping of packets on both sides of the line and measuring the ratio between ones successfully received and the total of packets sent.

As for the validation of the conclusion, given by equation 3.10, that the magnitude of the impact of the impulse on the demodulation of the symbols is approximately constant, we could just synchronize the modems, inject noise on the line, measure the SNR change for each carrier and see if it follows a pattern proportional to the noise's discrete spectrum shown by figure 14. However, as mentioned before, although the testbed used could measure the SNR for allocating the number of bits, there was no way to obtain the SNR values for different instants of time; the equipment used does not provide it. So, the experiment used for validation was the following, Experiment 1:

- synchronize the modems - synchronization 1;
- inject noise into the line, affecting the transmission between the modems;
- force re-synchronization of the modems, with noise still on the line - synchronization 2;

- measure the difference of the allocation of bits between synchronizations 1 and 2 and plot it as a function of frequency - curve 1;
- check similarity between curve 1 and the noise's discrete spectrum;

It is worth mentioning that, as the bits are allocated depending on the SNR at each frequency, the difference of the bits allocated to each synchronization gives indeed a sense of the noise power measured by the modem.

In order to test the hypothesis claimed on section 3.3, the following, Experiment 2, has been performed:

- synchronize the modems - synchronization 1;
- inject noise into the line, affecting the transmission between the modems;
- measure the loss of data caused by the noise;
- force re-resynchronization of the modems, with noise still on the line - synchronization 2;
- measure the loss of data caused by the noise after re-synchronization;

4.3 Results

4.3.1 Experiment 1

For the validation of the model, experiment 1 was performed for a number of times, and now some plots of the comparison between the difference of bits allocated to the two synchronizations and the noise's discrete spectrum are shown. The noise spectrum is first shown by figure 14. Once again, the y-axis is not normalized to any particular scale, but that should be of no importance to the discussion: we are only interested in comparing the shapes of the waveforms.

As a result of the experiment 1, now figure 15 illustrates, as a function of frequency, the difference of bits allocated to each synchronization.

On that figure, the high peaks are due to different pilot carriers chosen between the synchronizations. Even so, and despite the rounding caused by the integer constraint on the number of bits, we can see by those figures that the power measured by the modem

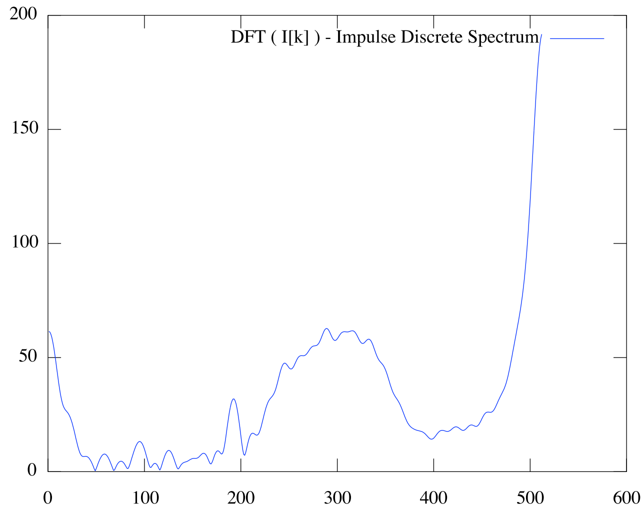

Figure 14: The Discrete Fourier Transform (magnitude) spectrum of the noise impulse

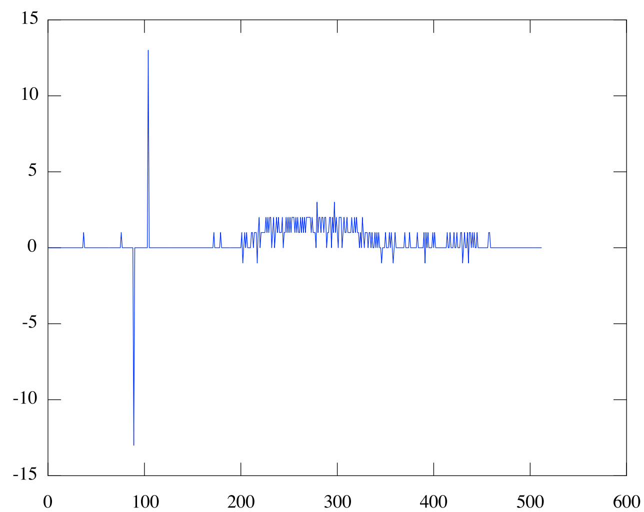

Figure 15: Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 6dB$, and REIN with $f = 50Hz$

when the noise is of the REIN type is indeed similar to the noise's discrete spectrum. Moreover, it should be clear now, after all the previous discussion, that this plot should become more and more similar to the noise's discrete spectrum as f increases if the all the assumptions made on section 3.3 hold. So we plot that again on figure 16, with the same line parameters, but with $f = 100$ Hz.

Figure 16 shows even more similarity with the noise's discrete spectrum. Figure 17 shows a comparison.

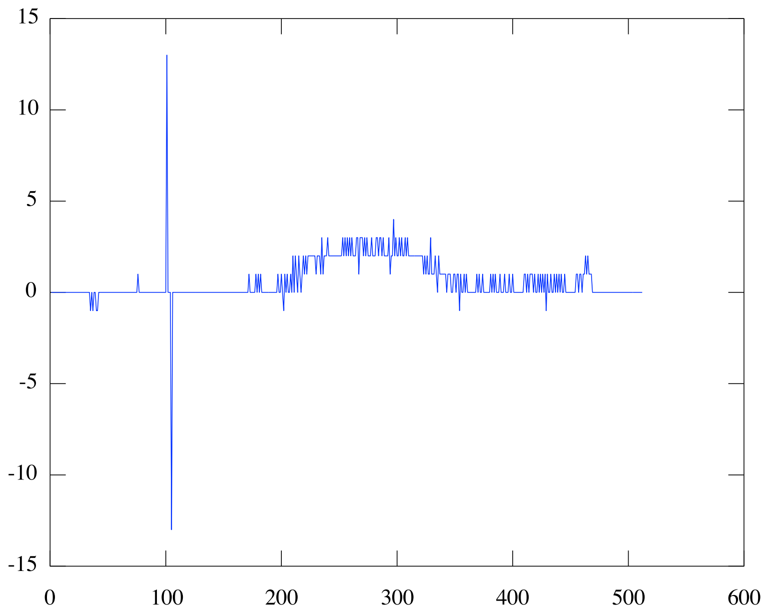

Figure 16: Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 6dB$, and REIN with $f = 100Hz$

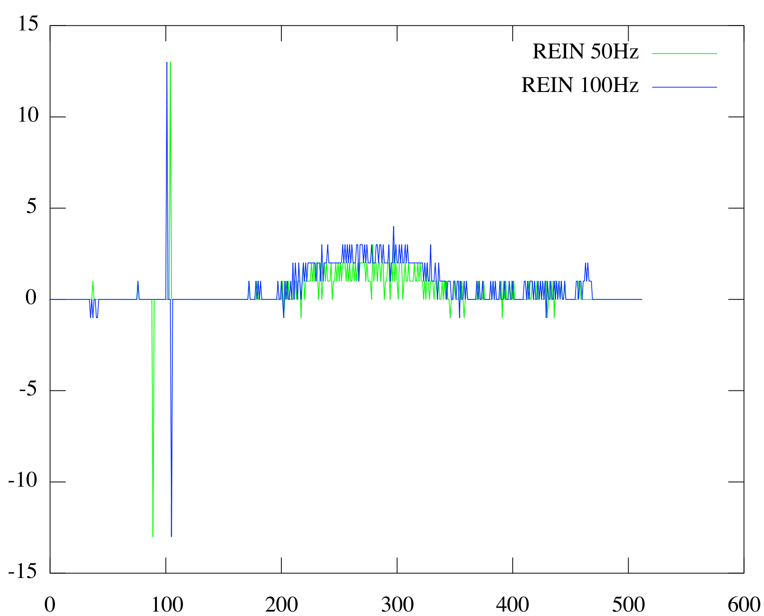

Figure 17: Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 6dB$, and REIN with $f = 100$Hz and $f = 50$Hz

On the other hand, for AWGN the plot of the difference of bits is much flatter. This is shown by figure 18.

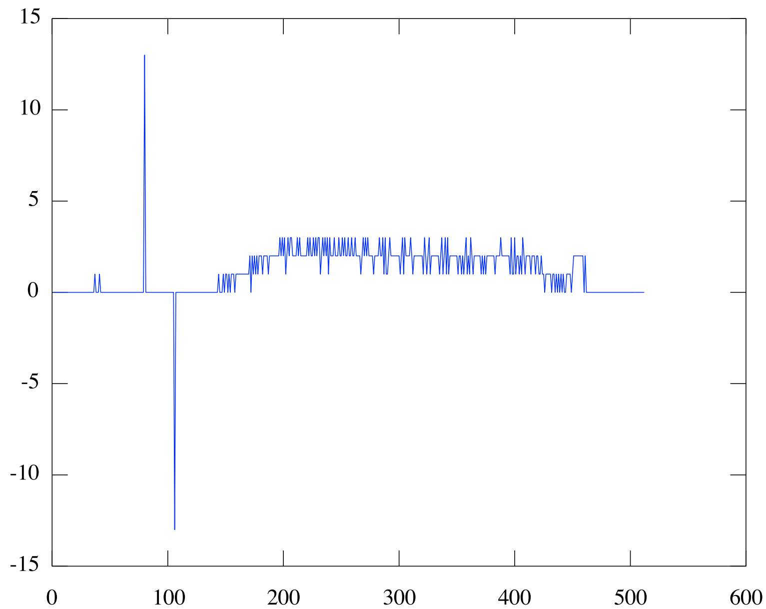

Figure 18: Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 6dB$, and AWGN

Just to make clear the distinction, figures 19 and 20 make the comparison of the three cases, for the parameter SNR Margin set to 6db and 9db.

All those graphs should be enough to validate the model on section 3.1 and the assumptions made on section 3.3 about the power measurements made by the modem.

We next show the results of Experiment 2, in order to validate the hypothesis made at the end of section 3.3.

4.3.2 Experiment 2

Here are shown the results of experiment 2. They were all performed on testbed described before, with $LoopLength = 2000$m and $INP = 0$. The choice of no *Impulse Noise Protection* is due to the factors mentioned before: we do not want any coding to interfere on the loss measurements; the intention is to protect the line only through changing parameters at the physical layer.

As mentioned on the description of Experiment 2, the line was synchronized, next noise was injected into it and the loss of IP datagrams was measured. Next, a resynchronization

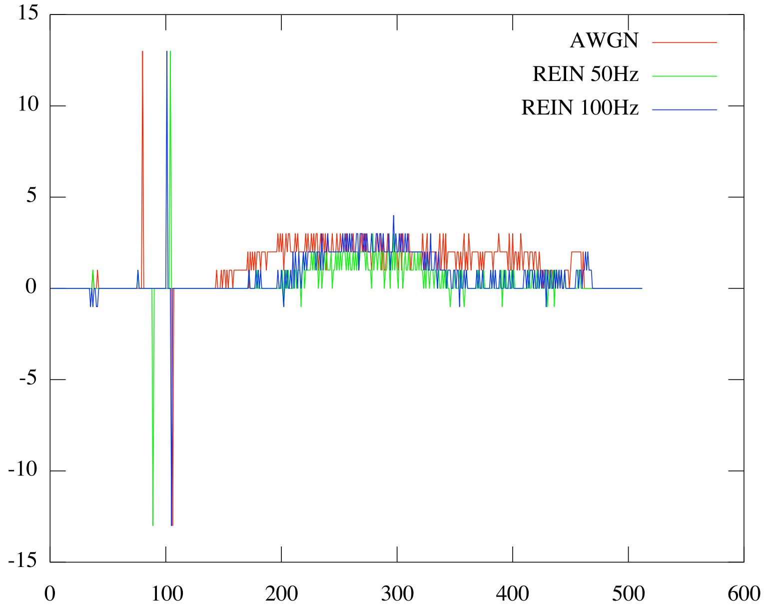

Figure 19: Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 6dB$, all three noises

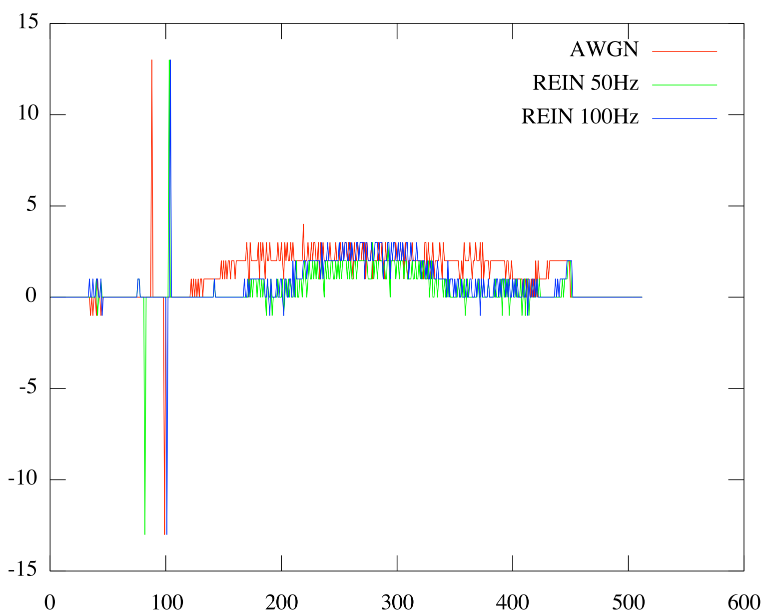

Figure 20: Difference of allocated bits for $LoopLenght = 2000$m, $INP = 0$, $SNR_{mgn} = 9dB$, all three noises

was forced, and the loss was measured again. Table 1 show the results.

From those results we can see that, with AWGN noise being injected, the line always

Table 1: Loss before and after resynchronization, for different SNR Margin values

Noise	SNR Margin	Before resync	After resync
AWGN	6dB	0.003150	0
	9dB	0	0
	12dB	0	0
	15dB	0	0
REIN 50Hz	6dB	0.033550	0.034170
	9dB	0.037100	0.033460
	12dB	0.034140	0.001090
	15dB	0.036040	0.000000
REIN 100Hz	6dB	0.066770	0.067510
	9dB	0.068630	0.066570
	12dB	0.069760	0.000000
	15dB	0.071970	0.000000

recovers itself after a synchronization, independent of the value for the parameter SNR Margin. However, when the noise injected is a REIN, the behavior is different - this is exactly what was explained on chapter 1, during the introduction to the problem.

The hypothesis of section 3.3 was that a resynchronization with a SNR Margin value of $10log_{10}\left(\frac{F}{f}\right)$ should be enough for REINs of any impulse power. That gives a value of 19.031 for f = 50Hz and 16.021 for f = 100Hz.

The table shows that, for f = 50Hz, a SNR Margin of 15dB was enough for the recovery of the line. Moreover, for f = 100Hz, 12dB was already enough. So, there was no need to perform experiments of higher values for SNR Margin, as they would only result in zero loss again. And those results are in accordance with the hypoteshis.

It is worth mentioning that the value for SNR Margin that was enough for the experiments with the noise we used could not be enough for another REIN with more impulse power. This work proposed a value for that parameter that should be enough for *any* impulse power, and that is why there can be differences between the value proposed and the one that is enough for a specific noise.

5 Conclusions and Future work

Throughout this work, an analytical model of the effect of REINs on the DMT symbols, for each frequency, was derived. The research has shown that, based on some simple assumptions about the time domain characteristics of the impulses, it becomes possible to better understand the behavior of a given repetitive interference on the demodulation of the symbols. It was shown that REINs with small duty cycle -with short impulses in comparison to the DTM symbol sampling duration- cause an error, on the demodulation of the symbol, whose magnitude is approximately constant for each frequency and whose phase is homogeneously distributed.

A contribution deriving from that model was a practical way of calculating and obtaining a non-arbitrary value for the *SNR Margin* parameter so that, when the line resynchronizes, there is no error anymore. With that parameter value, the noise effect has been compensated for all the experiments, with only one synchronization. The effect of the noise could be mitigated without increasing the *Impulse Noise Protection* parameter, and that characteristic can benefit specially those delay-sensible applications - since increasing the INP value increases the transmission delay.

Two main restrictions of the work here proposed could be pointed out: when you resynchronize the line with the parameter here proposed, you must assume the noise affecting it is of the REIN type; moreover, you must also assume a value for f.

The first restriction relates to the assumption that the noise is made of repetitive impulses which are much smaller in length than the DMT symbol sampling. That assumption is valid for the noise we used, but could not so for some other ones. Thus, the modeling of repetitive noises with impulses larger than that lead straight to future work.

As for the assumption of a value for f, we know that most of the electrical power around the world is of the alternate type, with frequencies of 50Hz and 60Hz mainly. Thus, those frequencies could be good default values for f, when no other is known.

In any case, the bit allocation algorithms of nowadays do assume the presence of

AWGN, so what we propose here is just another assumption, in case the first one does not seem to hold.

References

[Gallager 1968]GALLAGER, R. G. *Information Theory and Reliable Communication.* [S.l.]: Wiley, 1968.

[Golden e Jacobsen 2006]GOLDEN, H. D. P.; JACOBSEN, K. *DSL Fundamentals.* [S.l.]: Auerbach Publications, 2006.

[Henkel e Kessler 1994]HENKEL, W.; KESSLER, T. A wideband impulsive noise survey in the german telephone network, statistical description and modeling. *AEU*, v. 48, n. 6, p. 277,Äì288, 1994.

[ITU-T 2001]ITU-T. *Standard G.996.1, Test procedures for digital subscriber line (DSL) transceivers.* [S.l.], 2001.

[ITU-T 2009]ITU-T. *Standard G.992.5, Asymmetric digital subscriber line (ADSL) transceivers , Extended bandwidth ADSL2 (ADSL2plus).* [S.l.], 2009.

[Levey e McLaughlin 1999]LEVEY, D.; MCLAUGHLIN, S. Statistics of impulse noise: Interarrival times. *contribution to ETSI WG TM6*, 1999.

[Nedev 2003]NEDEV, N. H. *Analysis of the Impact of Impulse Noise in Digital Subscriber Line Systems.* Tese (Doutorado), 2003.

[Oppenheim e Schafer 1989]OPPENHEIM, A. V.; SCHAFER, R. W. *Discrete-time signal processing.* [S.l.]: Prentice-Hall, 1989.

[Sedarat et al. 2005]SEDARAT et al. Multicarrier bit-loading in presence of biased gaussian noise sources. *Consumer Communications and Networking Conference*, 2005.