NEW APPROACH TO BIT LOADING AND POWER MINIMIZATION USING MERCURY/WATERFILLING

Miloš Jakovljević[1], *Santiago Zazo*[1], *José Luis Peña*[2]

[1]ETSIT-SSR, Signal Processing Application Group (GAPS),
Ciudad Universitaria s/n, 28040 Madrid, Spain
Emails:{milos, santiago}@gaps.ssr.upm.es

[2]Telefónica I+D ,
Emilio Vargas 6, 28043 Madrid, Spain
Email: sedano@tid.com

ABSTRACT

This paper investigates mercury/waterfilling (MWF) power distribution policy for multi-carrier systems and its potential benefits. Comparing the performance in terms of bit error rate with conventional bit loading algorithm for practical systems, like Levin Campello (LC), we demonstrated that its power distribution is not optimal as widely believed due to the gap approximation inaccuracy for low constellation sizes. This fact can be used to improve system stability by achieving lower bit error rates or decreasing noise margin. Furthermore, we proposed a new formulation for throughput optimization seen as a combinatorial problem due to discrete nature of modulations used in practical systems. Moreover, we developed a new bit loading algorithm based on MWF as solution to this problem that gives higher throughput than LC under the same set of constraints. We also developed an algorithm for power minimization that enables power savings compared to LC while keeping the same quality of service as before. These strategies can be beneficial to operators in reduction of their operational costs.

1. INTRODUCTION

In recent years, telecom operators showed a strong interest in improving the capacity utilization of their twisted-pair access networks. During the last two decades, several families of digital subscriber line (DSL) have been developed, standardized and installed. Some of them like the family of ADSL [1] and VDSL [2] technologies are based on multi-carrier discrete multi-tone modulation (DMT).

The allocation of power in order to maximize the throughput of conventional multi-carrier DSL systems was done by assuming that the inputs were Gaussian and then using the implementation gap in order to satisfy the capacity equation waterfilling policy was applied. The implementation gap introduced in capacity equation depends on modulation type and coding schemes and includes the Shannon gap approximation and a noise margin. It was demonstrated in [3] and [4] that Shannon gap can be approximately considered independent of constellation size for QAM modulation. Nevertheless, waterfilling policy gives real number of bits, not discrete, that has no meaning for practical systems. For practical system with constellation size constraints some of the algorithms developed can be found in [5] and [6]. Using the same gap approximation approach, in [7, 8, 9] the optimal power and bit allocation algorithm, named Levin Campello (LC), for practical multi-carrier systems is presented. This procedure certainly does not give the optimum power distribution because all procedures assume Gaussian inputs, but it was implemented in lack of explicit expression for the throughput function. Yet, recently in [10], fundamental relation between mutual information and minimum mean square error of conditional estimator was revealed. Based on this work Lozano in [11] formulated power allocation policy for arbitrary inputs of Gaussian parallel channel named mercury/waterfilling (MWF). It is using computable nonlinear MMSE of the inputs given their noisy outputs. In this way he went around the need for explicit mutual information expressions.However, this procedure needs to know in advance what inputs each channel has and it must be fixed. Thus, this is not a bit loading procedure.

In this paper we will first analyze the difference between mercury/waterfilling (MWF) policy from [11] and Levin Campello (LC) algorithm of [7, 8, 9] in terms of bit error rate (BER) for the same bit distribution. Thus, we will demonstrate that LC algorithm does not give the optimum power distribution since with MWF policy much lower BER can be obtained. Without going any further, this fact can be used to reduce the noise margin or improve the stability of the system. Taking the advantage of this fact, we formulated a new problem for throughput optimization. In order to solve this complex combinatorial optimization problem, we developed a novel bit loading algorithm based on MWF policy with the same set of restrictions as LC. We start with LC bit distribution and search for a better one that will give higher throughput. Therefore, we do not assume that we know the bit distribution in advance like in ordinary MWF, but rather improve the throuhout of the system keeping the same BER and aggregate power constraints as in LC algorithm. Energy consumption accounts for a big deal of operating expenses.Therefore, operators can choose power savings while keeping the same quality of service parameters. Hence, in this paper we also developed an algorithm based on MWF for power minimization while keeping the same throughput and BER that can be achieved by LC algorithm. Assuming the same discrete set of constellations inherited by involving MWF, the power minimization is also a complex combinatorial optimization problem. Although we used DSL systems for evaluation of our approaches, they can be applied on any multi-carrier system.

The rest of the paper is organized as follows: Section 2 gives a brief overview of the MF from [11]. In Section 3 we discuss the problem that we are trying to solve. Section 4 describes the novel algorithms and Section 5 gives simulation results and discussions that verify our approach. Section 6 concludes the paper with the major findings.

2. PRELIMINARIES

In order to optimize the system performance the most known and used criterion is the maximization of the input-output mutual information under a power constraint.

When the transmitted symbols are Gaussian, this criterion leads to the classical waterfilling policy, and when the mutual information is constrained due to the use of discrete constellations, the optimum policy is mercury/waterfilling (MWF) [11]. Moreover, this policy has no need for any gap approximation.

The independent N parallel Gaussian channels can be modeled as:

$$Y_k = h_k X_k + n_k \qquad (k = 1 \dots N), \tag{1}$$

where n_k is a zero - mean complex Gaussian random variable independent of the noise on the other channel and h_k is non zero channel gain. The complex valued inputs X_k are assumed to be independent and considering unit-power inputs S_k can be written as $X_k = \sqrt{p_k} S_k$, where $p_k \in [0, \infty)$ is the power allocated to sub-channel k. Any multi-carrier system can be represented in this way.

Under the aggregate power constraint P_{max} and the assumption that the inputs are normalized with respect to power constraint, the problem that was solved in [11] is the following:

$$[p_1^\star \dots p_N^\star] = arg \underbrace{max}_{\substack{p_1 \cdots p_N \\ \sum_i p_i = 1}} \sum_{i=1}^{N} I_i(SNR). \tag{2}$$

where p_i is the normalized power that should be assigned to the tone i, and I_i is the mutual information of the corresponding constellation.

The fundamental relation for any arbitrary input distribution revealed in [12] and used in further elaboration is given as:

$$\frac{d}{d(SNR)} I_k(SNR) = \mathbf{MMSE}_k(SNR), \tag{3}$$

where I_k is mutual information in Neper and $\mathbf{MMSE}_k$ is the mean square error of the Conditional Mean Estimate (CME) of the date in the k_{th} sub-channel with SNR signal to noise ratio.

To obtain the solution to this problem it is assumed that the receiver has the knowledge of magnitude and phase of channel gains. On the other hand the transmitter needs only the knowledge of the magnitudes. The power allocation $\{p_i^\star\}_1^N$ that solves (2) is given in [11] by:

$$p_i^\star = 0, \quad \gamma_i \leq \lambda, \tag{4}$$

$$\gamma_i \mathbf{MMSE}_i(p_i^\star \gamma_i) = \lambda, \quad \gamma_i > \lambda, \tag{5}$$

where λ is obtained from power constrained and γ_i is the normalized channel gain according to the power constraint.

The MMSE is a nonlinear function that is different for each constellation and if symbols are considered equally probable the expression given in [11] is:

$$\mathbf{MMSE}(SNR) = 1 - \frac{1}{M\pi} \int \frac{\left| \sum_{l=1}^{m} s_l e^{-\left|y - \sqrt{SNR} s_l\right|^2} \right|^2}{\sum_{l=1}^{m} e^{-\left|y - \sqrt{SNR} s_l\right|^2}} dy. \tag{6}$$

Taking into consideration that the integral in equation 6 does not have analytical solution it is very unpractical to implement MWF policy using this expression due to number of constraints. Therefore, we suggest to use two different approximations for MMSE in high power regime that can be assumed for vast number of applications. For example in DSL large SNR per tone is needed in order to achieve desired BER of 10^{-7}.

In [11] it is shown that the allocated power can be represented as reciprocally dependent to minimum distance that conform to input discrete constellations. Given the minimum distances $\{d_i\}_{i=1}^n$ optimum power in high power regime can be seen as:

$$p_i^\star = \frac{\alpha}{\gamma_i d_i^2} + O\left(\frac{\log P}{P}\right) \tag{7}$$

with

$$\frac{1}{\alpha} = \frac{1}{n} \sum_{l=1}^{n} \frac{1}{\gamma_i d_l^2} \tag{8}$$

As demonstrated in [13] MMSE for large SNR can be approximated as the π times the symbol error rate (SER) of the received constellation. Therefore the optimum power allocation for QAM constellations can be approximated as:

$$p_i^\star = 0, \quad \gamma_i \leq \lambda, \tag{9}$$

$$p_i^\star = \frac{2}{\gamma_i} \left(\frac{Q^{-1}\left(\frac{M\lambda}{K\pi\gamma_i}\right)}{d_{imin}} \right)^2, \quad \gamma_i > \lambda \tag{10}$$

where K is the number of pairs of points at minimum distance and M is the constellation size.

3. PROBLEM DEFINITION

For typical constellations, larger constellation sizes correspond to larger values of mutual information for any signal to noise ratio (SNR). For maximum mutual information, the richest available constellation should be used on each tone. Therefore, the definition of a bit loading policy under the maximization of the mutual information criterion is meaningless, and it is necessary to introduce some other practical constraints involving the achieved BER performance. Thus, given the arbitrary normalized channel inputs that can be chosen from a set of discrete constellations C with cardinality $\mathscr{C}$, the problem that we are trying to solve can be expressed as follows:

$$[p_1^\star \dots p_N^\star] = arg \underbrace{max}_{p_1 \cdots p_N, j} \sum_{i=1}^{N} I_i(SNR_i, \mathbf{c}_j(i)), \tag{11}$$

$$s.t. \sum_i p_i = P_{budget}, \tag{12}$$

$$P_e\left([p_1^\star \dots p_N^\star]\right) \leq BER_{target}, \tag{13}$$

$$j = 1, \dots (\mathscr{C})^N, \tag{14}$$

where

$$\mathbf{c}_j = \left[c_j(1), \dots c_j(N)\right], \tag{15}$$

$$c_j(k) \in \{0, \dots, \mathscr{C}\}, \tag{16}$$

$$k = 1, \dots, N, \tag{17}$$

where $p_i^\star$ is the optimum power for sub-carrier i, p_i is the power for sub-carrier i, I_i is the mutual information for sub-carrier i that depends on signal to noise ratio (SNR) and vector $\mathbf{c}_j$. Vector $\mathbf{c}_j$ has as elements the current bit distribution for each sub-carrier. These entries are denoted as $\mathbf{c}_j(i)$ for sub-carrier i and they can take values from the set 0 to $\mathscr{C}$. The index j defines all possible combinations of bit distributions and it goes from 1 to $\mathscr{C}^N$. Equations (16) and (17) denotes the element of vector $c_j(k)$ for kth sub-carrier.

Basically, we are trying to find the variation with repetition of constellation sizes on each tone or bit loading that will give the highest value for aggregate mutual information satisfying BER requirement. This problem having discrete set of input constellations presents complex combinatorial optimization problem. Certainly, in order to find the optimum solution brute force algorithm can be applied, but this is only efficient for very small number of tones and constellations. If we have N tones and $\mathscr{C}$ constellations there will be $\mathscr{C}^N$ possible solutions that should be investigated. For practical systems such as DSL, where number of tones is 512 for ADSL2+ [1] and 4096 for VDSL2 [2] and they use QAM constellations that can have sizes from 1 to 15 points, this approach is impossible. Therefore our goal is to find sub optimum solution that will satisfy the above constraints and we will demonstrate that it can achieve better throughput than the one obtained with LC algorithm.

The same problem can be seen as power minimization in the following manner:

$$min \quad \sum_{i=1}^{N} p_i(\mathbf{c}_j(i)), \tag{18}$$

$$s.t. \sum_i I_i(SNR, \mathbf{c}_j(i)) = R_{target}, \tag{19}$$

$$P_e([p_1 \dots p_N]) \leq BER_{target} \tag{20}$$

$$j = 1, \dots (\mathscr{C})^N, \tag{21}$$

where

$$\mathbf{c}_j = [c_j(1), \dots c_j(N)] \tag{22}$$

$$c_j(k) \in \{0, \dots, \mathscr{C}\} \tag{23}$$

$$k = 1, \dots, N \tag{24}$$

The notation has the same meaning as explained above. This is also a combinatorial optimization problem.

4. NOVEL ALGORITHMS

As we will demonstrate later, for the same constellation distribution among tones that is obtained with conventional LC algorithm, MWF has better performance in terms of BER. Therefore, this can be exploited to make the system more stable or to reduce the noise margin in order to keep the same performance. To take the advantage of this fact also two different strategies can be applied. First strategy that we propose is to find the bit loading or constellation distribution that will have better performance than LC solution while having the same aggregate power and BER. The other strategy is power minimization, where aggregate power can be decreased keeping the same throughput and BER of the system. This approach is similar to margin adaptive algorithms, but our solution due to combinatorial problem does not guarantee that the solution is optimum.

Algorithm 1 Optimization Bit Loading Algorithm

```
Calculate power and bit loading using LC algorithm b_LC
Calculate bit error rate for LC solution BER_LC
Set MWF bitloading b_MWF = b_LC
Calculate power distribution using MWF for bit loading b_MWF
Calculate bit error rate BERMWF for b_MWF
if BER_MWF ≥ BER_LC then
   End the algorithm. No improvement
end if
repeat
   for i = 1  to  N  {Do for every tone} do
      b_MWF(i,:) = b_MWF(i)+1{Increase for one bit on current tone }
      Calculate power distribution for b_MWF with MWF
      Calculate BER BER_MWF for b_MWF
   end for
   Find for which tone k BER_MWF is minimum
   Set b_MWF = b_MWF(k,:)
until BER_MWF < BER_LC
```

4.1 Bit Loading Algorithm

As already explained due to combinatorial nature of the problem we are trying to find suboptimal solution for the problem imposed in (11) that will satisfy constraints (12), (13) and (14).

As constraints we assume that the aggregate power has the same value for both LC and MWF and target BER should be less or equal to the BER obtained with LC. The pseudo code for the algorithm that we propose is labeled as 1.

Since we are trying to find a bit loading distribution that will give higher throughput than LC policy we take as a starting point the solution that is obtained with this procedure. Then, we calculate the power distribution using MWF as explained in Section 2. For this power distribution among tones we calculate the BER that can be achieved. If it is the same as for LC then this is the solution and no improvement in throughput is possible. If this is not the case then we search for better solution by forming new set of possible bit loading combinations where each member has one bit added on different tone. Now we search for new bit loading among the members of this new set and choose the one with lowest BER. The process is repeated iteratively until target BER is reached.

The complexity of the algorithm includes the complexity of LC and MWF. Besides, it must run thorough all bit loading cases of a newly formed set that has number of carriers members until the desired BER is achieved. Therefore, it has extra complexity of number of carriers times number of iterations needed to achieve the desired BER.

4.2 Power Minimization Algorithm

Since MWF gives optimal power distribution for a discrete input constellations some savings in power consumption compared to conventional LC algorithm can be achieved. In order to evaluate this gain we propose an algorithm labeled as 2. The algorithm first calculates the bit loading that is derived from LC, then calculates BER that can be achieved with such a power distribution. Afterwards, for the same bit

Algorithm 2 Optimization Bit Loading Algorithm

Calculate power and bit loading using LC algorithm b_{LC}
Calculate bit error rate for LC solution BER_{LC}
Set MWF bit loading $b_{MWF} = b_{LC}$
Set power constraint to P_{budget} used for LC
repeat
  Calculate power distribution with MWF
  Calculate BER BER_{MWF}
  Set $P_{budget} = P_{budget} - \Delta_p$
until $BER_{MWF} < BER_{LC}$

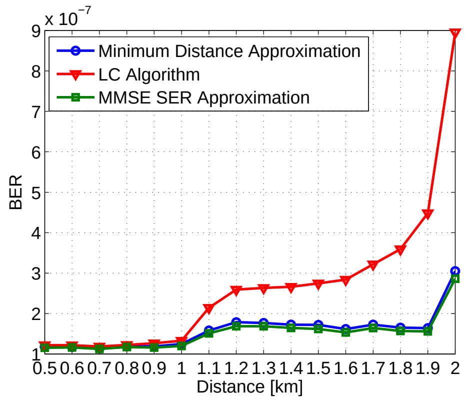

Figure 1: BER performance of LC algorithm, MWF with SER and minimum distance approximations dependent of cable lengths.

loading the algorithm calculates optimum power distribution using MWF policy and calculates BER that can be achieved. If this BER equals BER obtained with LC no power savings can be obtained. If not, the algorithm decreases the aggregate power for some step and recalculates the power distribution and BER by using MWF. The process iterates until the BER is smaller or equal to the one obtained with LC power distribution. The complexity of this algorithm includes the complexity of LC and MWF plus the number of iterations that are needed to achieve the desired BER.

5. SIMULATION RESULTS AND DISCUSSIONS

In order to evaluate the performance of the algorithms presented in Section 4 some simulations were performed. For this purpose we used DSL channel model from [14] and the rest of the parameters were according to ADSL2+ standard [1]. Thus, we assumed $\Gamma = 12\,\text{dB}$ as the implementation SNR gap for LC algorithm, the background noise level was set at $-140\,\text{dBm/Hz}$ and tone distance was $4.3125\,\text{kHz}$. For MWF no gap is needed. Moreover, to take into account the alien noise, in addition to the background noise, we have also added the ETSI ADSL Noise A [15]. We considered cable of $0.4\,\text{mm}$ (AWG24) for different lengths that were between $500\,\text{m}$ and 2000 m. As FEXT disturbers we considered 19 users that have the same length as a particular modem of interest. We did not consider NEXT since in today systems it can be avoided by echo cancellation or FDD techniques. We considered only downstream.

Figure 1 represents BER as a function of length for LC algorithm and MWF with two different approximations (equations (7),(8), (9), (10)) for MMSE using the same bit loading. As can be noted MWF has better performance over LC. This improvements is better for longer cables. Channel gain for longer cable is lower and therefore the constellations that have fewer number of points are chosen. Since MMSE for those constellations is further away from Gaussian inputs, MWF distributes power in an optimal way and achieves better performance. Also it is obvious that approximation of MMSE that is proportional to SER shows slightly better results because the bound is tighter. Figure 2 shows the relative gain in BER that can be achieved by MWF policy. The improvement is in the range of 10% on shorter lines up to 70% on longer lines.

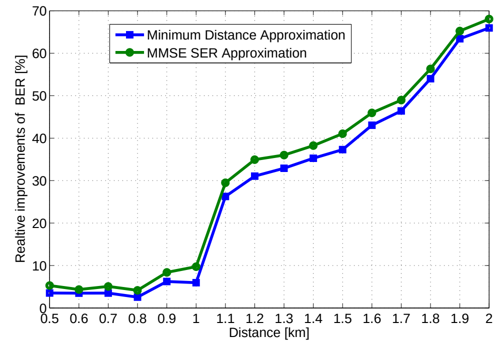

Figure 2: Relative gain in BER Performance of MWF with SER and minimum distance approximations compared to LC algorithm for different cable lengths.

Figure 3 presents relative gain in throughput that can be achieved by implementing algorithm 1. As can be seen the throughput that can be achieved is better from LC for $0,2\%$ on short lines up to 5% on longer lines that use smaller constellation sizes on more tones. Two MMSE approximations has almost the same performance. Minimum distance approximation is faster and easier to implement. This improvement can be beneficial for users that are further away. Clearly, increasing the range just a little can offer a dramatic savings to operators operational costs. Also the operators can deliver new services to the users that are further and that did not have them before.

Figure 4 shows excess power needed for LC algorithm compared with MWF policy assuming the same throughput and BER requirement. As it can be noted the excess power for short lines is small around 0.2% while for longer lines this extra power needed for LC increases to almost 10%. Taking this into cosideration operators can reduce substantialy their power consumption by implementing MWF policy with reduced aggregate power while achieving the same quality of service as in todays systems. Also electromagnetic compatibility issues can be relaxed since less power means less egress radiation. The difference between the two MMSE approximation is not significant.

6. CONCLUSION

In this paper we presented a new approach to bit loading and power minimization problem for DSL systems. We demonstrated that usual bit loading techniques such as LC algo-

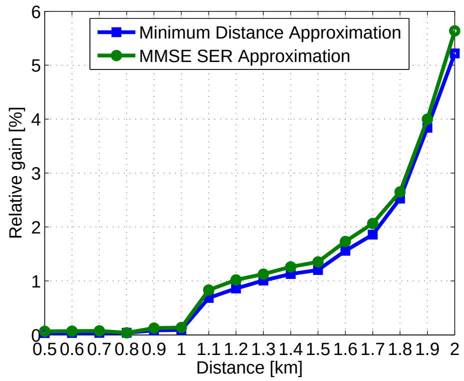

Figure 3: Relative gain in throughput of MWF with SER and minimum distance approximations compared to standard LC approach for different cable lengths.

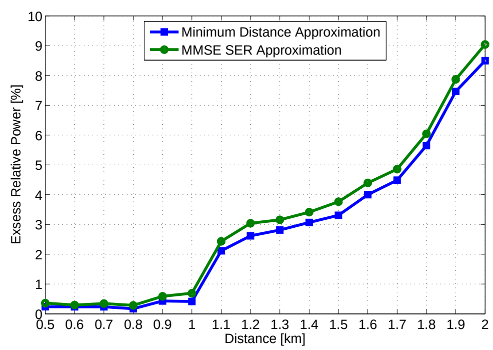

Figure 4: Relative excess power needed for LC power distribution compared with MWF policy with SER and minimum distance approximations for the same BER and throughput.

rithm does not give optimum power distribution for practical multi-carrier systems where discrete constellations are used. Applying MWF power distribution policy for the same constellations allocation as for LC, BER can significantly be improved. Taking the advantage of this fact we developed two sub optimal algorithms based on MWF policy that can improve overall system throughput or reduce power consumption while maintaining the same BER restriction as LC algorithm. These improvements are higher for the users that are using longer lines, because MWF policy treats the deviation of smaller constellations from Gaussian assumption of channel inputs better than the gap approximation. These algorithms can be beneficial for operators since they can reduce their operational costs.

7. ACKNOWLEDGMENT

We thank the Spanish Ministry of Education and Science for their support under Grant TEC2007-67520-C02-01/02/TCM.

REFERENCES

[1] ITU-T, "Asymmetric Digital Subscriber Line ADSL Transceivers - extended bandwidth ADSL2 (ADSL2plus)," Standard G.992.5, ITU, May 2003.

[2] ITU-T, "Very high speed digital subscriber," Standard G.993.1, ITU, Jun. 2004.

[3] R. Price, "Nonlinearly feedback equalized PAM versus capacity for noisy filter channels," in *Proc. of the IEEE International Conference on Communications, ICC*, Jun. 1972.

[4] G. D. Jr Forney and M. V. Eyuboglu, "Combined equalization and coding using precoding," *IEEE Communications Magazine*, vol. 29, no. 12, pp. 25–34, Dec. 1991.

[5] P.S. Chow, J.M. Cioffi, and J.A.C. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *Communications, IEEE Transactions on*, vol. 43, no. 234, pp. 773–775, 1995.

[6] R.F.H. Fischer and J.B. Huber, "A new loading algorithm for discrete multitone transmission," in *Global Telecommunications Conference, 1996. GLOBECOM '96. 'Communications: The Key to Global Prosperity*, 1996, vol. 1, pp. 724–728 vol.1.

[7] J. Campello, "Optimal discrete bit loading for multicarrier modulation systems," in *Proc. IEEE International Symposium on Information Theory*, 1998, pp. 193–.

[8] J. Campello, "Practical bit loading for dmt," in *Communications, 1999. ICC '99. 1999 IEEE International Conference on*, 6-10 June 1999, vol. 2, pp. 801–805vol.2.

[9] H.E. Levin, "A complete and optimal data allocation method for practical discrete multitone systems," in *Global Telecommunications Conference, 2001. GLOBECOM '01. IEEE*, 2001, vol. 1, pp. 369–374 vol.1.

[10] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.

[11] A. Lozano, A.M. Tulino, and S. Verdu, "Optimum power allocation for parallel gaussian channels with arbitrary input distributions," *IEEE Transactions on Information Theory*, vol. 52, no. 7, pp. 3033–3051, July 2006.

[12] D. Guo, S. Shamai, and S. Verdu, "Mutual information and minimum mean-square error in gaussian channels," *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, 2005.

[13] F. Perez-Cruz, M.R.D. Rodrigues, and S. Verdu, "Optimal precoding for digital subscriber lines," in *Communications, 2008. ICC '08. IEEE International Conference on*, 2008, pp. 1200–1204.

[14] R. F. M. van den Brink, "Cable reference models for simulating metallic access networks," *ETSI/STC TM6 contribution 970p02r3*, Permanent Document, Jun. 1998.

[15] ETSI, "Transmission and multiplexing (TM); Access transmission systems on metallic access cables; Asymmetric Digital Subscriber Line (ADSL) - European specific requirements [ITU-T Recomendation G.992.1 modified]," Standard TS 101 388, Version 1.3.1, ETSI, May 2002.