Analytical evaluation of the role of estimation and quantization errors in downstream vectored VDSL systems

Marco Baldi, Franco Chiaraluce

Dip. Ing. Biomedica, Elettronica e Telecomunicazioni
Università Politecnica delle Marche
Ancona, Italy
{m.baldi, f.chiaraluce}@univpm.it

Roberto Garello

Dipartimento di Elettronica
Politecnico di Torino
Torino, Italy
garello@polito.it

Marco Polano

Telecom Italia
Torino, Italy
marco.polano@telecomitalia.it

Marcello Valentini

Torino, Italy
marcello.valentini@ieee.org

Abstract— **We investigate the effect of non idealities in the diagonalizing precoder vectoring technique used for cancellation of the far-end crosstalk in downstream VDSL networks. By using analytical formulas, we estimate the average bit rate achievable as a function of both relative and absolute estimation errors. Several numerical examples are provided, in different scenarios and operation conditions. Results are presented covering an important aspect for practical applications: the transmitted power required to achieve a target bit rate, as a function of the line length. Finally, we provide some results on the impact of a smart quantization law, which limits the performance loss.**

Keywords-VDSL; vectoring; channel estimation errors; quantization errors

I.   INTRODUCTION

A great amount of work has been done to improve the features of digital subscriber line (DSL) systems. Several papers appeared in past and recent literature for performance evaluation of such systems (see [1], [2], and the references therein). In this paper, our goal is to study and quantify the impact of some practical impairment parameters on the achievable bit rate and the other performance figures, under realistic scenarios. We mainly focus on the solutions adopted for overcoming the limitations imposed by far-end crosstalk (FEXT), which is a major problem in very high speed networks [3], [4]. Most of these solutions adopt user coordination techniques [5]; in particular, in the downstream direction, that will be considered in this paper, coordination is possible as the transmitting modems are co-located at the central office. So, FEXT can be completely canceled, at least in principle, through a pre-distortion to apply, tone by tone, at each modem's signal before its transmission. Using these discrete multi-tone (DMT) *vectored* transmission techniques, the achievable bit rates can be significantly increased with respect to non-vectored solutions. As a matter of fact, the recently issued ITU-T G.vector

(G.993.5) [6] allows expanded use of 100 Mbps DSL. Additionally, a distinctive feature of the vectored VDSL systems is that its steady-state performance is more predictable than non-vectored systems [7]. In essence, the vectored systems become more stable, and this should encourage the adoption of analytical models for performance analysis and evaluation. Actually, such models are already available, and allow to forecasting easily the best performance achievable.

However, some problems exist, that can make difficult to reach this optimal behavior. The first problem is the effect of estimation errors. Pre-distortion requires the knowledge of the channel transfer function; this is rather easy to obtain for the direct channels, but becomes more difficult for the crosstalk channels. The latter are not yet completely described and, although suitable models have been proposed, they require in depth verification. As regards the coordination technique, in this paper we consider the adoption of the diagonalizing precoder (DP) proposed in [8], that has the advantage of an easy implementation and does not require modifications of the customer premise equipment. In [9] some analytical and simulation results were provided to study this technique in a low complexity implementation. In [1], [2] we developed a theoretical approach that permits to analytically determine the average bit rate achievable, as a function of the relative or the absolute estimation error. In this paper, we extend that theoretical analysis and we further support it by a number of simulations, that allow a more complete characterization of the random variables involved, for example by evaluation of the probability density function (p.d.f.). Moreover, we present new results characterizing the target bit rate and the transmitted power necessary to achieve it, that are two aspects of key importance for practical applications.

Another significant aspect concerns the need to take into account the effect of finite word length in the representation of the precoder variables, i.e., to measure the impact of quantization errors. This issue is extremely

important, due to its influence on the performance complexity trade-off: on one hand, coarse quantization can imply intolerable rate loss but, on the other hand, a large number of quantization bits can yield high hardware complexity and a great amount of memory needed for the precoding process. In [10], it has been shown that to obtain a limited capacity loss, due to quantization errors, a 14 bits representation of the precoder entries may be necessary. On the other hand, in [2] we showed that, by adopting a suitable quantization law, the same loss can be ensured by using only 10 bits. A further study on such quantization technique is presented in this paper, focusing on different practical scenarios and confirming the performance gain.

In essence, the main contributions of this paper consist in facing the issues described above both in (simple) analytical terms and through an extensive numerical evaluation. Various realistic examples are considered. The statistical character of the variables involved is taken into account, which is not very common in previous literature. The results of the analysis are also used to discuss relevant aspects, like evaluation of the power required for achieving a target bit rate, which are of great importance in practical applications.

The organization of the paper is as follows. In Section II we introduce the system and its relevant performance parameters in ideal conditions. In Section III we describe the channel model adopted and the statistical issues it involves. In Section IV we discuss the effect of the estimation errors, through theoretical arguments and simulations. Section V describes the simulation environments and gives several numerical examples. In Section VI we present the impact of the quantization errors, and we discuss the quantization law that allows to reduce the number of quantization bits. Finally, Section VII concludes the paper.

II. IDEAL BEHAVIOR

We consider the VDSL 998 17 standard [11], characterized by 4096 tones with frequency separation $\Delta = 4312.5$ Hz. For downstream, the Power Spectral Density (PSD) cannot exceed the mask described in Table I and depicted in Fig. 1 [11].

Denoting by s_k^{mask} the value of the PSD at the kth tone, the power transmitted on line n at tone k must satisfy the constraint $P_k^n \leq s_k^{mask} \cdot \Delta$. On each line we consider a total power $P_T^n = \sum_{k=1}^M P_k^n$ equal to 14.5 dBm (a typical value for cabinet transmission), distributed by the water-filling algorithm (see [12], for example) on the $M = 2454$ tones allocated for downstream.

Let us consider Fig. 2, that refers to the kth downstream tone f_k. In the figure, $\mathbf{X}_k = [X_k^1, X_k^2, \ldots, X_k^L]^T$ is an L-component vector collecting the symbols transmitted by L users on as many lines. $\mathbf{H}_k$ is the $L \times L$ channel matrix; its (i, j)th element, H_k^{ij}, represents the channel from transmitter j to receiver i.

TABLE I. PSD MASK FOR VDSL 998 17 DOWNSTREAM

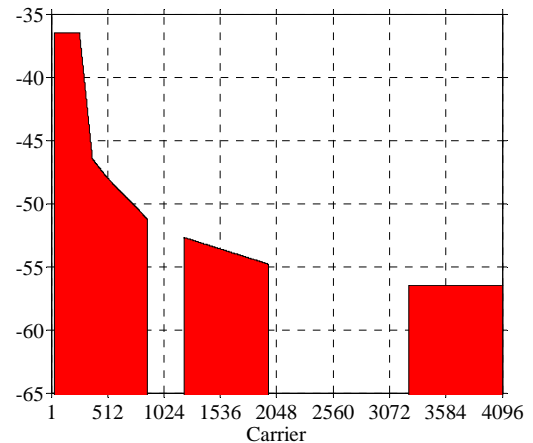| tone		PSD [dBm/Hz]	
from	to	from	to
32	256	−36.5	−36.5
256	376	−36.5	−46.5
376	512	−46.5	−48.0
512	870	−48.0	−51.2
1026	1971	−52.7	−54.8
3246	4096	−56.5	−56.5

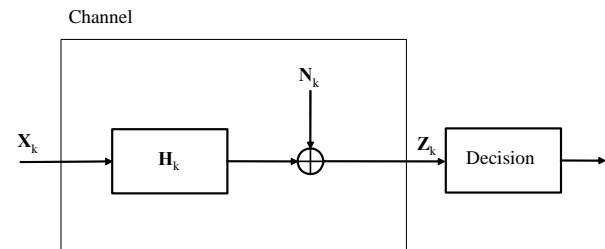Figure 1. Graphic representation of the PSD mask for VDSL 998 17 downstream.

Figure 2. Schematic representation of the VDSL system.

The matrix $\mathbf{H}_k$ is row-wise diagonal dominant (RWDD); this means that, on each row of $\mathbf{H}_k$, the diagonal element has typically much larger magnitude than the off-diagonal elements (i.e., $\left| H_k^{ii} \right| \gg \left| H_k^{ij} \right|$, $\forall j \neq i$). Also shown in Fig. 2, $\mathbf{N}_k$ is the L-component vector describing the additive thermal noise.

If all the L lines of the binder are controlled by the same operator, and the line drivers are co-located (in the same cabinet or central office) then the vector of symbols $\mathbf{X}_k$ can be made available to an apparatus able to coordinate the L lines. Ideally, this knowledge can be used to completely eliminate the FEXT interference by applying a proper precoder. This idea was first presented in [5], where vectored DMT was introduced.

In the simplest form of precoding (from a conceptual viewpoint), the vector $\mathbf{X}_k$ is pre-multiplied by a matrix $\mathbf{P}_k = \left\{ p_k^{ij} \right\}$, such that $\mathbf{P}_k = \alpha_k \mathbf{H}_k^{-1}$, which is able to completely remove the FEXT interference. This method requires:
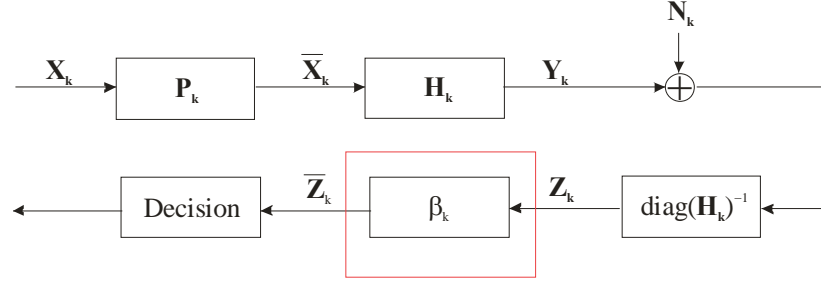
Figure 3.   Schematic representation of the vectored system based on DP.

- The exact knowledge of the channel matrix $\mathbf{H}_k$ at the transmitter side.
- The use of a scaling factor α_k for not exceeding the PSD mask. The Tomlinson precoder was proposed to solve this problem in [5], but has a quite large complexity..

To overcome the complexity of the Tomlinson precoder, some alternative schemes have been considered. Among them, one of the most promising is the Diagonalizing Precoder (DP) [8]. The DP system at tone k is shown in Fig. 3; the precoding matrix $\mathbf{P}_k$ is now defined as:

$$\mathbf{P}_k = \beta_k^{-1} \mathbf{H}_k^{-1} \cdot \text{diag}(\mathbf{H}_k) , \qquad (1)$$

with $\beta_k \triangleq \max_i \left\| \left[\mathbf{H}_k^{-1} \cdot \text{diag}(\mathbf{H}_k) \right]_{\text{row } i} \right\|$. In (1), $\text{diag}(\mathbf{H}_k)$ is the diagonal matrix having elements $H_k^{11}, \ldots , H_k^{LL}$.

In ideal conditions (i.e., perfectly compensated FEXT), the signal-to-noise ratio at the nth receiver and the kth tone is:

$$SNR_k^n = \frac{P_k^n \left| H_k^{nn} \right|^2}{\sigma_N^2} \qquad (2)$$

where P_k^n is the transmitted power and σ_N^2 is the variance of the thermal noise, that is independent of k and n.

By using the well-known gap approximation, the number of bit/symbol of user n at tone k is given by:

$$c_k^n = \log_2 \left\{ 1 + \frac{SNR_k^n}{\Gamma} \right\} , \qquad (3)$$

where Γ is the transmission gap, and the achievable bit rate is:

$$C^n = R_S \sum_{k=1}^{M} c_k^n , \qquad (4)$$

where $R_S = 4000$ symbol/s is the net symbol rate (which differs from Δ because of the cyclic prefix). Actually, in (3) and (4), the integer part of c_k^n must be assumed. To

simplify the notation, however, this will be not explicitly indicated afterward (but will be used in the numerical evaluations).

III. CHANNEL MODEL

We describe the diagonal elements of matrix $\mathbf{H}_k$ (direct channel) through the well-known Marconi (MAR1) model [13]. As regards the FEXT, such model can be extended by using the following expression, that takes into account the statistical coupling dispersion with respect to the 1% worst case [14].

$$H_{FEXT}(f,d) = \left| H(f,d) \right| f \sqrt{d} \chi 10^{-X/20} e^{i\Phi} . \qquad (5)$$

In (5), $H(f, d)$ is the direct channel at frequency f (in MHz) and distance d (in km), $\chi = 10^{-2.25}$ is a coupling coefficient, X is a Gaussian random variable with mean value (in dB) μ_X and standard deviation (in dB) σ_X, Φ is a uniform random variable $\in [0, 2\pi)$. For the subsequent analysis, it will be useful to compute the square modulus of (5). Because of the assumption on X, $10^{-X/10}$ is a log-normal variable whose mean value and variance are well known and can be expressed as functions of μ_X and σ_X. These expressions will be used in Section V, where several numerical examples will be discussed.

The values of μ_X and σ_X depend on the type of cable adopted. In the next numerical examples, we will set $\sigma_X = 7.8$ dB and $\mu_X = 2.33 \times \sigma_X = 18.174$ dB.

Eq. (5) is self-consistent in the case of lines with equal length and permits us to obtain the off-diagonal elements H_k^{nj}, with $n \neq j$. In a more practical scenario, where the lines within the cable have different lengths, say d_n and d_j, H_k^{nj} can be obtained through the following expression:

$$H_k^{nj} = \left| H_k^{nn} \right| f_k \sqrt{\min\left(d_j, d_n\right)} \chi 10^{-X/20} e^{i\Phi} , \qquad (6)$$

as the FEXT contribution is limited to the common length, while it is further attenuated in the longest loop.

In previous literature, (5) has been often used neglecting the statistical issues and considering the maximum $\left| H_k^{nj} \right|$ value, that is obtained by setting $X = 0$. This represents a worst case situation, that rarely occurs in practice. Instead, by following the approach presented in

[1] and [2], we can face the problem in statistical terms, so deriving more significant values for the quantities involved. Examples will be given in Section IV.

Even assuming the worst case situation, the crosstalk channel H_k^{nj} from a disturber j into a victim n is always much weaker than H_k^{nn}. As mentioned above, this yields the RWDD character of the downstream VDSL channel matrix, that is known to be a key issue for the efficiency of the vectored system.

IV. EFFECT OF ESTIMATION ERRORS

An analytical approach to quantify the impact of estimation errors was developed in [2] and is reported here for paper completeness. In Section V, this approach will be used to produce a set of results, matched to practical applications, much more extended with respect to the scenarios considered in [2].

Let $\hat{\mathbf{H}}_k$ be the estimated channel matrix. If an estimation error is present, it is modeled through a matrix $\mathbf{E}_k$ such that:

$$\hat{\mathbf{H}}_k = \mathbf{H}_k + \mathbf{E}_k . \qquad (7)$$

It is reasonable to assume that the direct channels are estimated correctly, so that $\mathbf{E}_k$ has zero diagonal elements. Moreover, for a given error percentage e, assumed constant for all the off-diagonal elements, we have $\hat{H}_k^{ij} = H_k^{ij}(1+e)$.

Matrix $\hat{\mathbf{H}}_k$ must replace, in (1), the actual matrix $\mathbf{H}_k$ and, looking at Fig. 3, through simple algebra, we find [2]:

$$\begin{aligned}
\bar{\mathbf{Z}}_k &= \left\{ \mathbf{I} - \left[\mathrm{diag}\left(\hat{\mathbf{H}}_k\right) \right]^{-1} \cdot \mathrm{diag}\left(\mathbf{E}_k \cdot \hat{\mathbf{H}}_k^{-1}\right) \cdot \mathrm{diag}\left(\hat{\mathbf{H}}_k\right) \right\} \cdot \mathbf{X}_k \\
&- \left[\mathrm{diag}\left(\hat{\mathbf{H}}_k\right) \right]^{-1} \cdot \left[\mathbf{E}_k \cdot \hat{\mathbf{H}}_k^{-1} - \mathrm{diag}\left(\mathbf{E}_k \cdot \hat{\mathbf{H}}_k^{-1}\right) \right] \cdot \mathrm{diag}\left(\hat{\mathbf{H}}_k\right) \cdot \mathbf{X}_k \\
&+ \beta_k \left[\mathrm{diag}\left(\hat{\mathbf{H}}_k\right) \right]^{-1} \cdot \mathbf{N}_k .
\end{aligned} \qquad (8)$$

Taking into account the RWDD character of the channel matrix, it is possible to verify that $\beta_k \approx 1$, $\hat{\mathbf{H}}_k^{-1} \cdot \mathrm{diag}\left(\hat{\mathbf{H}}_k\right) \approx \mathbf{I}$ and $\mathrm{diag}\left(\mathbf{E}_k \cdot \hat{\mathbf{H}}_k^{-1}\right) \approx 0$. Then, (8) becomes:

$$\bar{\mathbf{Z}}_k \approx \mathbf{X}_k - \left[\mathrm{diag}\left(\hat{\mathbf{H}}_k\right) \right]^{-1} \cdot \mathbf{E}_k \cdot \mathbf{X}_k + \left[\mathrm{diag}\left(\hat{\mathbf{H}}_k\right) \right]^{-1} \cdot \mathbf{N}_k . \qquad (9)$$

So, by elaborating (9) to obtain the signal-to-noise ratio, the number of bit/symbol (3) results in:

$$c_k^n = \log_2 \left\{ 1 + \frac{P_k^n \left| H_k^{nn} \right|^2}{\left| e \right|^2 \displaystyle\sum_{\substack{j=1 \\ j \neq n}}^{L} \left| H_k^{nj} \right|^2 P_k^j + \sigma_N^2} \cdot \frac{1}{\Gamma} \right\} . \qquad (10)$$

Because of the presence of the random variables H_k^{nj}, c_k^n is a random variable as well. In this paper, we consider two different approaches for estimating its mean $\langle c_k^n \rangle$; they are described next. From the knowledge of $\langle c_k^n \rangle$, the mean of C^n can be derived as well, by using (4).

Approximation 1: A first coarse approximation consists in replacing, in (10), the mean of $\left| H_k^{nj} \right|^2$, so that:

$$\langle c_k^n \rangle_1 = \log_2 \left\{ 1 + \frac{P_k^n \left| H_k^{nn} \right|^2}{\left| e \right|^2 \displaystyle\sum_{\substack{j=1 \\ j \neq n}}^{L} \left\langle \left| H_k^{nj} \right|^2 \right\rangle P_k^j + \sigma_N^2} \cdot \frac{1}{\Gamma} \right\} . \qquad (11)$$

Based on the channel model described in Section III, it is possible to verify that:

$$\left\langle \left| H_k^{nj} \right|^2 \right\rangle = \left| H_k^{nn} \right|^2 f_k^2 \min\left(d_j, d_n \right) \chi^2 \mathrm{e}^{-\ln(10)/10 \mu_X + [\ln(10)/10]^2 \sigma_X^2 / 2} . \qquad (12)$$

Approximation 2: A more accurate analysis should consider c_k^n as a function of the random variable:

$$Y = \sum_{\substack{j=1 \\ j \neq n}}^{L} Y_j = \sum_{\substack{j=1 \\ j \neq n}}^{L} P_k^j \min\left(d_j, d_n \right) 10^{-X_j/10} , \qquad (13)$$

and then obtain $\langle c_k^n \rangle$ accordingly. Once having fixed the scenario, P_k^j and $\min\left(d_j, d_n \right)$ are known, so that Y is the sum of properly scaled log-normal variables. With simple algebra we have (details are given in the Appendix):

$$\begin{aligned}
\langle c_k^n \rangle_2 = \langle c_k^n \rangle_1 + \log_2 &\left[\sqrt{\frac{\left(b \mu_Y + \sigma_N^2 \right)^2 + b^2 \sigma_Y^2}{\left(b \mu_Y + a + \sigma_N^2 \right)^2 + b^2 \sigma_Y^2}} \right. \\
&\left. \cdot \left(1 + \frac{a}{b \mu_Y + \sigma_N^2} \right) \right] ,
\end{aligned} \qquad (14)$$

where μ_Y and σ_Y^2 are the mean value and variance of Y, and

$$a = \frac{P_k^n \left| H_k^{nn} \right|^2}{\Gamma} \quad , \quad b = \left| e \right|^2 \left| H_k^{nn} \right|^2 f_k^2 \chi^2 . \quad (15)$$

Some analytical methods are known for computing the statistical averages of Y, as required in (14). Among them: 1) Wilkinson's method is particularly simple and provides an explicit solution [15]; 2) Schwartz & Yeh's method is more accurate but requires a recursive approach [16].

By applying Wilkinson's method, we find the following expressions:

$$\mu_Y = e^{-\ln 10/10 \mu_X + (\ln 10/10)^2 \sigma_X^2 /2} \sum_{\substack{j=1 \\ j \neq n}}^{L} P_k^j \min\left(d_j, d_n\right),$$

$$\sigma_Y^2 = e^{-2\ln 10/10 \mu_X + (\ln 10/10)^2 \sigma_X^2} \left\{ e^{(\ln 10/10)^2 \sigma_X^2} \sum_{\substack{j=1 \\ j \neq n}}^{L} \left[P_k^j \min\left(d_j, d_n\right) \right]^2 \right.$$

$$\left. + \sum_{\substack{j=1 \\ j \neq n}}^{L} \sum_{\substack{m=1 \\ m \neq j,n}}^{L} \left[P_k^j \min\left(d_j, d_n\right) \right] \left[P_k^m \min\left(d_m, d_n\right) \right] \right\} - \mu_Y^2 .$$

$$(16)$$

Channel estimation is typically realized through Least Squares (LS) or Recursive Least Squares (RLS) algorithms [17] and it is based on the transmission of S training symbols with constant power level. The goodness of the estimate depends on the value of S and on the ratio P_k^n / σ_N^2 [10]. Following an analytical procedure similar to that reported in [18], it is easy to verify that the average number of bit/symbol for user n at tone k, in the case of using the LS algorithm, can be written as [2]:

$$\left\langle c_k^n \right\rangle = \log_2 \left\{ 1 + \frac{P_k^n \left| H_k^{nn} \right|^2}{\left(\dfrac{L-1}{S} + 1 \right) \sigma_N^2} \cdot \frac{1}{\Gamma} \right\} . \quad (17)$$

Whilst (10), (11) and (14) refer to a fixed relative error percentage, (17) takes into account the variance of the estimation error, on the basis of the actual algorithm used for channel estimation. It also suggests the main method to limit the impact of the error, that consists, as obvious, in increasing the value of S. Numerical examples will be given in the following section.

V. NUMERICAL EXAMPLES AND SIMULATIONS

In our numerical examples we have used the PSD mask of the VDSL2 998 bandplan, up to 17 MHz, described in Section II. σ_N^2 has been computed on the basis of a background spectral density of −140 dBm/Hz.

Moreover, we have assumed $\Gamma = 9.75$ dB and a maximum number of bit/carrier (bit clipping) equal to 15.

A. Considered scenarios

By varying the number and length of the interfering lines, a huge number of different scenarios can be analyzed. Among them, just for explicative purposes, we have focused on the following three scenarios, that extend a first series of cases already presented in [2].

Scenario 1: In the first scenario, we assume that the lines have four different lengths, all multiple of a minimum value Δd. The situation is schematically shown in Fig. 4 (where DSLAM means Digital Subscriber Line Access Multiplexer and CPE denotes the Customer-Premises Equipment). For each length we have $L/4$ lines, with $L = 8$ (Case 1.1) or $L = 16$ (Case 1.2). In the simulations that will be discussed next, we have considered $\Delta d = 0.3$ km.

Scenario 2: In the second scenario, the $L/4$ lines in the first group have length d, while those in the ith group, with $i = 2, 3, 4$, have length $d + (i - 1) \Delta d$. This scenario is schematically shown in Fig. 5. For each length we have $L = 8$ (Case 2.1) or $L = 16$ (Case 2.2). In the simulations that will be discussed next, in particular, we have assumed $d = 0.9$ km and $\Delta d = 0.1$ km. A case that this scenario is well suited to model is the one plotted in Fig. 6, in which a first long span arrives up to the cellar of a building and, from that, at regular intervals, $L/4$ lines are separated to cover different building floors.

Scenario 3: In the third scenario, we have $L = 8$ (Case 3.1) or $L = 16$ (Case 3.2) lines with equal length. This can be seen as a special case of Scenario 2, when $\Delta d = 0$ is assumed.
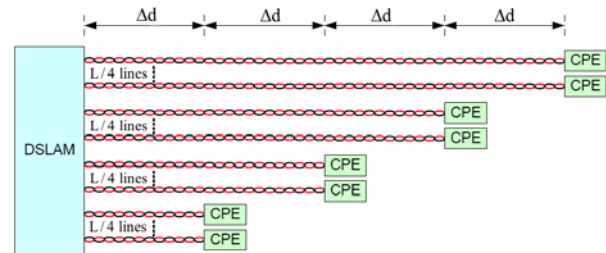

Figure 4. Schematic description of Scenario 1.
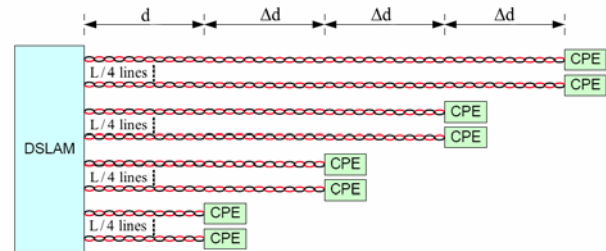

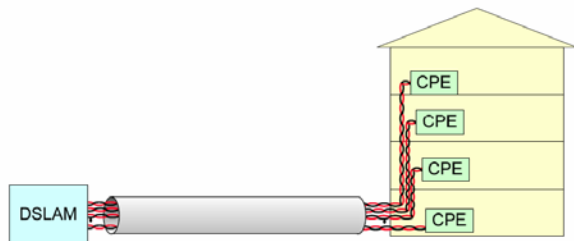Figure 5. Schematic description of Scenario 2.

Figure 6.   Example of application of the Scenario 2.

B.   Examples of average bit rates

Tables II-V report some examples of computation of the average bit rates, for the Scenarios 1 and 2 introduced in Section V.A, by assuming the presence of an error percentage e in the evaluation of the channel matrix. Both the analytical approximations discussed in Section III have been considered.

TABLE II.   EXAMPLES OF AVERAGE BIT RATES FOR CASE 1.1

Line length (km)	Simulation based on (9) (Mbps)	Simulation based on (8) (Mbps)	$\langle C^n \rangle_1$ (Mbps)	$\langle C^n \rangle_2$ (Mbps)
		$e = 0.1$		
0.3	140.56	140.53	136.29	140.44
0.6	99.93	99.86	100.05	100.41
0.9	63.25	63.24	63.25	63.28
1.2	40.91	40.91	40.90	40.90
		$e = 0.5$		
0.3	115.00	114.95	105.29	114.03
0.6	89.05	88.94	87.21	92.38
0.9	60.02	60.01	59.16	61.21
1.2	40.11	40.11	39.58	40.26

TABLE III.   EXAMPLES OF AVERAGE BIT RATES FOR CASE 1.2

Line length (km)	Simulation based on (9) (Mbps)	Simulation based on (8) (Mbps)	$\langle C^n \rangle_1$ (Mbps)	$\langle C^n \rangle_2$ (Mbps)
		$e = 0.1$		
0.3	135.99	135.92	130.22	135.02
0.6	98.70	98.52	98.82	99.38
0.9	63.00	62.98	62.99	63.06
1.2	40.86	40.85	40.82	40.83
		$e = 0.5$		
0.3	102.28	102.19	95.76	101.19
0.6	82.25	82.01	80.83	84.68
0.9	57.20	57.16	56.16	58.10
1.2	39.13	39.12	38.27	39.08

TABLE IV.   EXAMPLES OF AVERAGE BIT RATES FOR CASE 2.1

Line length (km)	Simulation based on (9) (Mbps)	Simulation based on (8) (Mbps)	$\langle C^n \rangle_1$ (Mbps)	$\langle C^n \rangle_2$ (Mbps)
		$e = 0.1$		
0.9	63.13	63.12	63.17	63.22
1.0	54.33	54.32	54.36	54.38
1.1	46.49	46.48	46.49	46.50
1.2	40.89	40.88	40.86	40.87
		$e = 0.5$		
0.9	58.65	58.62	58.06	60.28
1.0	51.22	51.20	50.90	52.52
1.1	44.44	44.42	43.99	45.19
1.2	39.59	39.58	39.01	39.89

TABLE V.   EXAMPLES OF AVERAGE BIT RATES FOR CASE 2.2

Line length (km)	Simulation based on (9) (Mbps)	Simulation based on (8) (Mbps)	$\langle C^n \rangle_1$ (Mbps)	$\langle C^n \rangle_2$ (Mbps)
		$e = 0.1$		
0.9	62.77	62.72	62.83	62.93
1.0	54.10	54.06	54.14	54.19
1.1	46.35	46.33	46.34	46.37
1.2	40.80	40.79	40.76	40.78
		$e = 0.5$		
0.9	55.33	55.27	54.80	56.69
1.0	48.73	48.67	48.41	49.92
1.1	42.59	42.55	41.98	43.16
1.2	38.28	38.26	37.44	38.36

The reliability of the approximated average bit rates has been tested also through a comparison with the simulation results. For this purpose, the samples of $\left| H_k^{nj} \right|^2$ have been generated, according to their statistics, and used in (10). The confidence of the estimation can be made high by increasing the number of random extractions. The numerical elaboration has been managed through simple programs written in Matlab© and C++.

Simulations have been performed by considering either the exact expression (8) or the approximate expression (9). The results confirm that the use of (9) is quite acceptable, being related to the RWDD character of matrix $\mathbf{H}_k$.

We also observe that Approximation 2 is generally better for the case of short lengths, while, for longer lengths, the gap between the two approximations becomes less pronounced. In all the considered cases, the agreement between the simulated and the analytical results is good, thus proving the effectiveness of the proposed model.

On the other hand, simulations permit us to derive any statistical description of the random variable c_k^n. In Fig. 7, for example, we have plotted the estimated p.d.f. for $d = 0.6$ km and $e = 0.5$, in Case 1.1. Coherent with Table II, the calculated mean value of the p.d.f. is 88.94 Mbps; it is also noticeable the fact that the curve is very narrow around the mean, thus demonstrating a small variance of the achievable bit rates, that is another recognized property of the vectored systems [7], even in the presence of estimation errors.
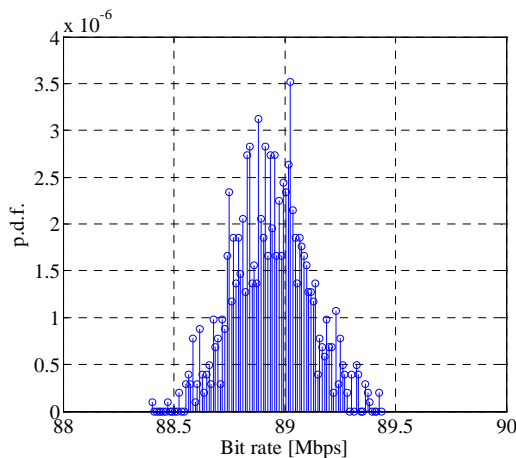
Figure 7.   Estimated p.d.f. for $d = 0.6$ km and $e = 0.5$ in Case 1.1.

As stressed above, (17) also permits us to evaluate the impact of the estimation error induced by a limited number of training symbols. The achievable bit rate for different values of S, by considering the Scenarios 1 and 2 described in Section V.A, are reported in Tables VI-IX. For the non-vectored system, the results have been obtained by adopting the Approximation 2.

TABLE VI.   AVERAGE BIT RATES AS A FUNCTION OF THE NUMBER OF TRAINING SYMBOLS FOR CASE 1.1

Line length (km)	Non-vector. (Mbps)	Vector. $S = 1$ (Mbps)	Vector. $S = 10$ (Mbps)	Vector. $.S = 100$ (Mbps)	Vector. $S = 1000$ (Mbps)	Vector. Ideal (Mbps)
0.3	96.79	133.88	141.97	144.41	144.66	144.69
0.6	82.14	79.32	95.86	100.23	100.88	100.96
0.9	57.11	48.08	59.71	63.01	63.41	63.46
1.2	38.80	30.95	38.31	40.60	40.92	40.96

TABLE VII.   AVERAGE BIT RATES AS A FUNCTION OF THE NUMBER OF TRAINING SYMBOLS FOR CASE 1.2

Line length (km)	Non-vector. (Mbps)	Vector. $S = 1$ (Mbps)	Vector. $S = 10$ (Mbps)	Vector. $S = 100$ (Mbps)	Vector. $S = 1000$ (Mbps)	Vector. Ideal (Mbps)
0.3	83.10	127.90	139.91	143.96	144.63	144.69
0.6	71.77	71.92	91.96	99.58	100.80	100.96
0.9	51.82	42.82	56.84	62.49	63.36	63.46
1.2	36.07	27.78	36.42	40.24	40.88	40.96

TABLE VIII.   AVERAGE BIT RATES AS A FUNCTION OF THE NUMBER OF TRAINING SYMBOLS FOR CASE 2.1

Line length (km)	Non-vector. (Mbps)	Vector. $S = 1$ (Mbps)	Vector. $S = 10$ (Mbps)	Vector. $S = 100$ (Mbps)	Vector. $S = 1000$ (Mbps)	Vector. Ideal (Mbps)
0.9	55.32	48.08	59.71	63.01	63.41	63.46
1.0	48.86	39.87	50.82	54.08	54.50	54.54
1.1	42.41	34.58	43.43	46.20	46.57	46.62
1.2	37.80	30.95	38.31	40.60	40.92	40.96

TABLE IX.   AVERAGE BIT RATES AS A FUNCTION OF THE NUMBER OF TRAINING SYMBOLS FOR CASE 2.2

Line length (km)	Non-vector. (Mbps)	Vector. $S = 1$ (Mbps)	Vector. $S = 10$ (Mbps)	Vector. $S = 100$ (Mbps)	Vector. $S = 1000$ (Mbps)	Vector. Ideal (Mbps)
0.9	49.68	42.82	56.84	62.49	63.36	63.46
1.0	44.39	35.37	48.12	53.58	54.44	54.54
1.1	38.76	30.91	41.15	45.78	46.53	46.62
1.2	34.74	27.78	36.42	40.24	40.88	40.96

From the tables we see that, just by using $S = 100$ training symbols, the average bit rate is very close to the ideal result, thus providing the expected high gain with respect to the non-vectored system (also shown in the table for the sake of reference). We have developed a number of similar comparisons for different scenarios; when longer loops are considered, the requirement on S can become more stringent. Moreover, as clearly shown by (17), for an increasing number of lines, the value of S has to be increased as well. However, assuming $S = 1000$ symbols should guarantee a limited impact of the estimation errors for any VDSL2 scenario of practical interest.

From a different viewpoint, the designer can be interested to fix a value of the bit rate and to determine the maximum line length that is compatible with it, for a suitable total power. Obviously, this length depends on the number and length of the interfering lines. Some examples are shown in Table X, for "target" bit rates of 50, 75 and 100 Mbps. The Case 3.1 has been analyzed, and therefore the system consists of $L = 8$ lines of equal length; the value of S has been assumed high enough to make the impact of the absolute estimation errors negligible. The table shows the total power required for reaching the specified bit rates, as a function of the line length. Where the label "n.r." appears, this means that the target bit rate cannot be reached within a total power of 14.5 dBm that, in Section II, has been assumed as a typical value. Focusing attention on the bit rate of 100 Mbps, that is the expected value in the G.993.5 Standard, we see that the non-vectored system is unable to reach it even for the shortest lengths. On the contrary, by introducing vectored transmission, the 100 Mbps target can be reached up to distances longer than 600 m. Obviously, this conclusion holds for the specific scenario here considered (note, in particular, that the tagged line and all the interfering ones have the same length) but it can be easily extended, updating the limits, to other scenarios.

Another relevant aspect that results from the numerical analysis is that, depending on the line length, the desired bit rate can often be reached by using a total power P_T^n that is significantly smaller than the typical value of 14.5 dBm. This power saving is another merit of the vectored system, and it is seen with particular interest by the vendors, for the technological advantages it provides.

TABLE X.   TOTAL POWER (IN dBm) REQUIRED FOR ACHIEVING THE TARGET BIT RATE IN CASE OF L = 8 LINES OF EQUAL LENGTH

Bit rate (Mbps)	Technique	Line length d (km)										
		0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2
50	Vectored	−35.2	−28.8	−22.8	−17.9	−13.2	−8.3	−3.3	1.9	7	12.2	n.r.
	Non-vectored	−30.2	−23.5	−17.1	−11.4	−6.3	−0.4	5.9	12.6	n.r.	n.r.	n.r.
75	Vectored	−27.3	−20.9	−14.4	−7.8	−1	5.6	11.6	n.r.	n.r.	n.r.	n.r.
	Non-vectored	−20.5	−13.4	−5.3	4.9	n.r.	n.r.	n.r.	n.r.	n.r.	n.r	n.r.
100	Vectored	−19.7	−13.1	−6.2	1.1	9	n.r.	n.r.	n.r.	n.r.	n.r	n.r.
	Non-vectored	n.r.	n.r.	n.r.	n.r.	n.r	n.r	n.r	n.r	n.r	n.r.	n.r.

VI. EFFECT OF QUANTIZATION ERRORS

The effect of quantization errors can be modeled in a way similar to that discussed in Section IV for the estimation errors. This analysis has been reported in [2], and is here repeated for the sake of completeness.

Let us suppose that matrix $\mathbf{P}_k$ is represented, in finite precision, as a matrix $\hat{\mathbf{P}}_k$, such that:

$$\hat{\mathbf{P}}_k = \mathbf{P}_k + \mathbf{E}_k , \qquad (18)$$

where matrix $\mathbf{E}_k$ now expresses the quantization errors. The latter, in turn, can be related to a matrix $\mathbf{\Delta}_k$ as follows:

$$\mathbf{\Delta}_k = \mathbf{P}_k^{-1} \cdot \mathbf{E}_k . \qquad (19)$$

In ideal conditions, we have $\mathbf{\Delta}_k = \mathbf{E}_k = \mathbf{0}$. Through simple algebra, the signal-to-noise ratio at the nth receiver and the kth tone, in the presence of quantization errors, is [10]:

$$S\bar{N}R_k^n = \frac{\left|H_k^{nn}\right|^2 \left|1+\Delta_k^{nn}\right|^2 P_k^n}{\left|H_k^{nn}\right|^2 \sum_{\substack{j=1 \\ j \neq n}}^{L} \left|\Delta_k^{nj}\right|^2 P_k^j + \sigma_N^2} , \qquad (20)$$

being Δ_k^{nj} the (n, j)th element of $\mathbf{\Delta}_k$. Eq. (20) can be used to replace SNR_k^n in (3), thus reducing the achievable bit rate with respect to the ideal condition. By investigating the statistical properties of c_k^n in presence of quantization errors, it is possible to find the number of quantization bits needed to have a penalty smaller than a prefixed percentage. In this view, an important analytical work was done in [10], where a number of bounds were determined, and their reliability tested through simulations.

In that paper, however, the elements of $\mathbf{E}_k$ were modeled as random variables uniformly distributed in the range $[-2^{-v}, 2^{-v}]$, where v is the number of quantization bits. No specific quantization law was considered.

Noting by $\bar{c}_k^n$ the number of bit/symbol for user n at tone k in presence of quantization errors, and using (3) and (4), the effect on the bit rate is measured by the following parameter:

$$\frac{\langle L^n \rangle}{C^n} \cdot 100 = \frac{\sum_{k=1}^{M} \langle L_k^n \rangle}{C^n} \cdot 100 , \qquad (21)$$

where $L_k^n = c_k^n - \bar{c}_k^n = \log_2 \left\{ \dfrac{1+\Gamma^{-1} SNR_k^n}{1+\Gamma^{-1} S\bar{N}R_k^n} \right\}$ has the meaning of transmission rate loss for the nth receiver at the kth tone [10]. Taking into account that the modulus of the diagonal elements of matrix $\mathbf{P}_k$ is about 1, a first choice consists in assuming a midtread quantization law between -1 and 1. Because of the RWDD property of matrix $\mathbf{H}_k$, however, the off-diagonal elements are very small. So, following this quantization law, most of the off-diagonal elements become zero after quantization, particularly in the case of small v and low frequencies.

Explicitly, this means that the vectoring procedure is made ineffective by quantization. In spite of this, for small values of v, the error due to the quantization is, on average, smaller than that resulting from the assumption of a uniform error. Tables XI and XII show the values of $\langle L^n \rangle / C^n \cdot 100$, in Case 3.2 (i.e., with $L = 16$ lines of equal length) as obtained by the model in [10] and by the midtread quantization law here considered, for various line lengths, namely: $d = 0.3$ km, $d = 0.6$ km, $d = 0.9$ km and $d = 1.2$ km. The difference between the two groups of results is evident for small v, while it almost disappears for large v. Both tables confirm that, wishing to have a rate loss below 4% for line lengths ≥ 0.3 km, $v = 14$ bits is almost always required. Though this value could be implemented on the basis of the current technology, it seems exaggeratedly high, and, in fact, it can be reduced by using a smarter quantization law.

The key point is the need to distinguish between the dynamics of the diagonal elements of $\mathbf{P}_k$, that are close to 1, and that of the off-diagonal elements, that are much smaller than 1. So, in [2] we proposed to adapt the midtread quantization law to such dynamics, by assuming different quantization thresholds for the two classes of data. This approach is further developed and assessed in the following. In practice, the 2^v quantization levels are distributed between $-T_{h1}$ and T_{h1} for the diagonal elements, and between $-T_{h2}$ and T_{h2} for the off-diagonal elements. The assumption of T_{h1} equal to 1 seems a natural choice. On the contrary, the choice of T_{h2} should take into account the dynamics of the off-diagonal elements. Fig. 8 shows an example of max and average values of $\left|P_k^{ij}\right|$, with $i \neq j$, for $L = 16$ and line length $d = 0.3$ km or $d = 1.2$ km.

While the maximum of $\left|P_k^{ij}\right|$ can be locally rather high, the average value is very small, and the assumption of $T_{h2} = 0.05$ is a reasonable choice, particularly for the short line lengths. So, our quantization law assumes a uniform distribution in the range $[-1, +1]$, for the diagonal elements, and in the range $[-0.05, +0.05]$, for the off-diagonal elements. It should be noted that the implementation of this quantization scheme does not require any additional processing, but only a selective management of the elements of the precoding matrix.

The results obtained by using the midtread quantization law with different thresholds are shown in Table XIII. In comparison with Tables XI and XII, we see a significant improvement for any value of v. In particular, the target of capacity loss below 4%, for $d \geq 0.3$ km, can now be achieved by using only $v = 10$ bits (or even $v = 8$ bits, for the longest lines), with a significant saving with respect to the case of equal thresholds.

TABLE XI. $\langle L^n \rangle / C^n \cdot 100$ WITH UNIFORM GENERATION OF THE QUANTIZATION ERRORS FOR CASE 3.2

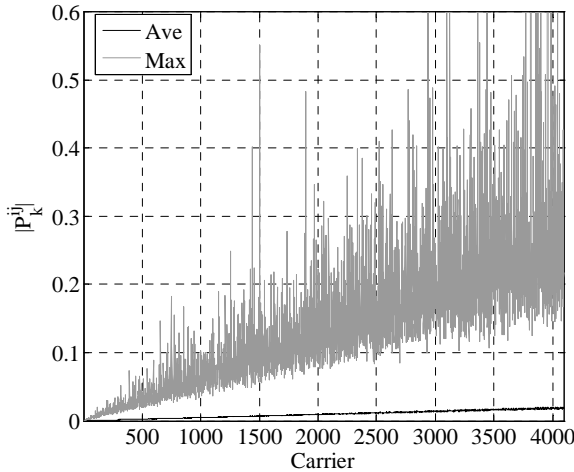| d (km)	$v = 6$	$v = 8$	$v = 10$	$v = 12$	$v = 14$
0.3	86.28	59.33	33.75	13.79	3.34
0.6	64.99	37.95	17.73	5.88	1.08
0.9	60.15	33.40	14.90	4.53	0.71
1.2	61.23	33.13	14.02	4.11	0.60

TABLE XII. $\langle L^n \rangle / C^n \cdot 100$ WITH MIDTREAD QUANTIZATION FOR CASE 3.2

d (km)	$v = 6$	$v = 8$	$v = 10$	$v = 12$	$v = 14$
0.3	61.69	50.84	31.85	13.59	3.37
0.6	38.56	29.24	16.00	5.77	1.13
0.9	28.76	22.85	12.75	4.36	0.72
1.2	24.27	20.03	11.37	3.92	0.61

Similar conclusions were drawn in [2], by considering different scenarios and parameter values, thus proving the effectiveness of the proposed quantization law under rather general conditions.
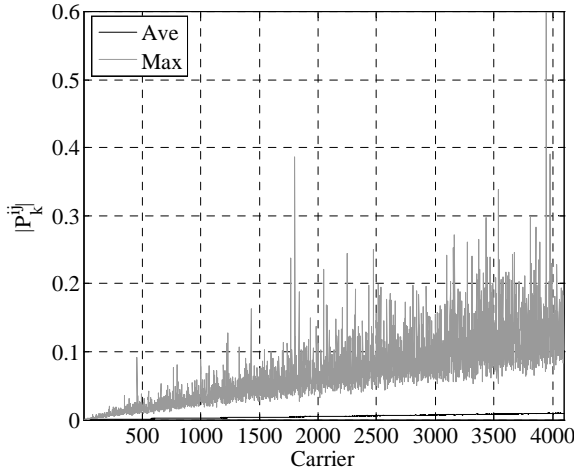
Figure 8. Simulated dynamics for the off-diagonal elements of the precoding matrix: (a) $d = 0.3$ km, (b) $d = 1.2$ km, in Case 3.2.

TABLE XIII. $\langle L^n \rangle / C^n \cdot 100$ WITH MIDTREAD QUANTIZATION ADOPTING DIFFERENT THRESHOLDS FOR CASE 3.2

d (km)	$v = 6$	$v = 8$	$v = 10$	$v = 12$	$v = 14$
0.3	29.79	12.66	3.85	1.65	1.48
0.6	15.03	5.17	1.21	0.54	0.49
0.9	11.41	3.58	0.59	0.11	0.07
1.2	10.17	3.14	0.45	0.05	0.03

VII. CONCLUSION

Estimation errors and quantization errors can severely limit the performance of vectored VDSL systems. The analysis of their effects must be performed by taking into account the statistical nature of the FEXT. In this paper, we have considered the case of a downstream VDSL link, where crosstalk is nominally canceled by using a diagonalizing precoder. Starting from analytical formulas for quantifying the effect of such impairments, we have performed a study on practical VDSL scenarios. We have verified that the impact of the estimation errors can be made negligible by using a number of training symbols in the order of $S = 1000$. Additionally, only 10 bits or fewer are required to maintain the capacity loss below 4% in the presence of quantization errors in the precoding matrix representation. The presented approach also allows to analytically determine the transmitted power that is necessary to achieve a target bit rate at a given distance, that is a key parameter for the design of VDSL systems.

The analysis has been focused on the VDSL2 17 MHz profiles, but it can be extended, for example, to the VDSL2 30 MHz profiles, whose adoption is planned for the near future.

APPENDIX

Derivation of Approximation 2

Let us define:

$$SNR_k^n = \frac{P_k^n \left| H_k^{nn} \right|^2}{\left| e \right|^2 \sum_{\substack{j=1 \\ j \neq n}}^{L} \left| H_k^{nj} \right|^2 P_k^j + \sigma_N^2}, \tag{22}$$

the signal-to-noise ratio in the presence of the error percentage e. With the positions (13) and (15), we have:

$$\log_2 \left(1 + \frac{SNR_k^n}{\Gamma} \right) = \log_2 \left(1 + \frac{a}{bY + \sigma_N^2} \right)$$
$$= \log_2 \left(Y + \frac{a + \sigma_N^2}{b} \right) - \log_2 \left(Y + \frac{\sigma_N^2}{b} \right). \tag{23}$$

Let us define:

$$W_1 = Y + \frac{a + \sigma_N^2}{b}, \quad W_2 = Y + \frac{\sigma_N^2}{b}. \qquad (24)$$

Both W_1 and W_2 are log-normal variables, with mean value and variance:

$$\mu_{W_1} = \mu_Y + \frac{a + \sigma_N^2}{b}, \quad \mu_{W_2} = \mu_Y + \frac{\sigma_N^2}{b}, \qquad (25)$$

$$\sigma_{W_1}^2 = \sigma_{W_2}^2 = \sigma_Y^2. \qquad (26)$$

Consequently,

$$\Pi_1 = \log_2 W_1, \quad \Pi_2 = \log_2 W_2 \qquad (27)$$

are Gaussian variables, with mean value

$$\langle \Pi_{1/2} \rangle = \log_2 \left(\mu_{W_{1/2}} \right) - \frac{1}{2} \log_2 \left(1 + \frac{\sigma_{W_{1/2}}^2}{\mu_{W_{1/2}}^2} \right). \qquad (28)$$

So, through simple algebra, we obtain:

$$\langle \Pi_1 \rangle - \langle \Pi_2 \rangle = \log_2 \left(\frac{\mu_{W_1}}{\mu_{W_2}} \right) + \frac{1}{2} \log_2 \left(\frac{\langle W_2^2 \rangle}{\langle W_1^2 \rangle} + \frac{\mu_{W_1}^2}{\mu_{W_2}^2} \right) \quad (29)$$

where $\langle W_i^2 \rangle$, $i = 1, 2$, is the average square value of W_i. Finally, by using the expressions above, (14) is derived.

REFERENCES

[1] M. Baldi, F. Chiaraluce, R. Garello, M. Polano, and M. Valentini, "On the Effect of Estimation and Quantization Errors in Downstream VDSL Systems", Proc. Third International Conference on Communication Theory, Reliability and Quality of Service, CTRQ 2010, Athens/Glyfada, Greece, 13 - 19 Jun. 2010, pp. 49-54.

[2] M. Baldi, F. Chiaraluce, R. Garello, M. Polano, and M. Valentini, "Simple Statistical Analysis of the Impact of Some Nonidealities in Downstream VDSL with Linear Precoding", Eurasip Journal on Advances in Signal Processing, Volume 2010, Article ID 454871, 14 pages, doi: 10.1155/2010/454871.

[3] J. Louveaux and A.-J. van der Veen, "Adaptive Precoding for Downstream Crosstalk Precancelation in DSL Systems Using Sign-Error Feedback", IEEE Transactions on Signal Processing, vol. 58, Jun. 2010, pp. 3173-3179.

[4] R. Y. Gilad, E. Domanovitz, A. Priebatch, I. Sharfer, A. Matza, and E. Tsur, "Vectored VDSL from a Practical Perspective,'" Proc. 2010 IEEE Sensor Array and Multichannel Signal Processing Workshop, Kibbutz Ma'ale Hahamisha, Israel, Oct. 2010, pp. 225-228.

[5] G. Ginis and J. M. Cioffi, "Vectored Transmission for Digital Subscriber Line Systems", IEEE Journal on Selected Areas in Communications, vol. 20, Jun. 2002, pp. 1085-1104.

[6] ITU-T Recommendation G.993.5, "Self-FEXT Cancellation (Vectoring) for Use with VDSL2 Transceivers," Apr. 2010.

[7] V. Oksman, et al., "The ITU-T's New G.vector Standard Proliferates 100 Mb/s DSL", IEEE Communications Magazine, vol. 48, Oct. 2010, pp. 140-148.

[8] R. Cendrillon, G. Ginis, E. Van den Bogaert, and M. Moonen, "A Near-Optimal Linear Crosstalk Precoder for Downstream VDSL", IEEE Transactions on Communications, vol. 55, May 2007, pp. 860-863.

[9] A. Leshem and L. Youming, "A Low Complexity Linear Precoding Technique for Next Generation VDSL Downstream over Copper", IEEE Transactions on Signal Processing, vol. 55, Nov. 2007, pp. 5527-5534.

[10] E. Sayag, A. Leshem, and N. D. Sidiropoulos, "Finite Word Length Effects on Transmission Rate in Zero Forcing Linear Precoding for Multichannel DSL," IEEE Transactions on Signal Processing, vol. 57, Apr. 2009, pp. 1469-1482.

[11] ITU-T Recommendation G.993.2, Amendment 1, ``VTU-R Limit PSD masks for band plan 998 (and its extensions)," Table B.6, Profile 998E17-M2x-NUS0, Apr. 2007.

[12] W. Yu, "Multiuser water-filling in the presence of crosstalk,'" Proc. IEEE Inf. Theory Appls. Workshop, San Diego, CA, 29 Jan.-2 Feb. 2007, pp. 414-420.

[13] J. Musson, "Maximum Likelihood Estimation of the Primary Parameters of Twisted Pair Cables," ETSI/STC TM6 981t08a1, Madrid, Spain, Jan. 1998.

[14] M. Sorbara, P. Duvaut, F. Shmulyian, S. Singh, and A. Mahadevan, "Construction of a DSL-MIMO Channel for Evaluation of FEXT Cancellation Systems in VDSL2," Proc. 2007 IEEE Sarnoff Symposium, Princeton, New Jersey, 30 Apr.-2 May 2007, pp. 1-6.

[15] P. Cardieri and T. S. Rappaport, "Statistics of the Sum of Log-normal Variables in Wireless Communications,'" Proc. 51st IEEE Vehicular Technology Conference (VTC'00), Tokyo, Japan, May 2000, pp. 1823-1827.

[16] S. C. Schwartz and Y. S. Yeh, "On the distribution function and moments of power sums with log-normal components", The Bell System Technical Journal, vol. 61, Sep. 1982, pp. 1441-1462.

[17] P. Turcza and T. Twardowski, "RLS Based MIMO Channel Identification for FEXT Compensation in Vectored xDSL System,'" Proc. European Conference on Circuit Theory and Design, Cork, Ireland, 29 Aug.-2 Sep. 2005, pp. 251-254.

[18] J. Le Masson, M. Ouzzif, and I. Wahibi, "Channel Estimation Using Data and Pilots for a Coordinated DSL System", Proc. IEEE Globecom 2007, Washington, DC, USA, 26 -30 Nov. 2007, pp. 2868-2872.