Sensitivity of Nonlinear Precoding to Imperfect Channel State Information in G.fast

Jochen Maes, Carl Nuzman, and Paschalis Tsiaflakis
Bell Labs, Nokia
Antwerp, Belgium and Murray Hill, New Jersey, USA
jochen.maes@nokia.com

***Abstract*—Nonlinear Tomlinson-Harashima precoding has been proposed as a near-optimal interference mitigation technique for downstream transmission in G.fast systems, particularly in the transmit spectrum above 106 MHz. We propose an alternative implementation of Tomlinson-Harashima precoding and examine performance of the common and alternative implementations with respect to quantization errors and imperfect channel state information. We show that Tomlinson-Harashima precoding is more sensitive than optimized linear precoding to varying channel state information due to fluctuations in ambient conditions and sudden changes in termination impedance. We also show that Tomlinson-Harashima precoding only outperforms optimized linear precoding when the channel state information is almost perfectly known.**

Keywords—G.fast; nonlinear precoding; Tomlinson-Harashima Precoding; vectoring; sudden termination change

I. INTRODUCTION

In the G.fast spectrum up to 212 MHz, the MIMO channel is not diagonally dominant and the linear precoder proposed in [1] is no longer near-optimal for far-end crosstalk cancellation. This is because the precompensation signal contributes significantly to the transmit power spectral density (PSD) of the victim line, affecting the aggregate transmit power and per-tone transmit PSD. Two methods can be employed to enforce power constraints. The first is the addition of per-user gain scaling to linear precoding, known as Transmitter Initiated Gain Adjustment (TIGA) [2].

The second approach is Nonlinear Precoding (NLP) which applies modulo arithmetic operations to shift a transmit constellation point with excessive power to a lower power equivalent. At the receiver, an independent modulo operation shifts the signal back to its intended position. The idea to employ modulo arithmetic to bound the value of the transmit signal was first introduced by Tomlinson and Harashima independently with application to single-user equalization [3,4]. Ginis and Cioffi applied the concept to multi-user system with precoding for crosstalk cancellation [5]. The near-optimal Tomlinson-Harashima Precoder (THP) is constructed using QR matrix decomposition. A sub-optimal approach is to decompose the channel matrix through lattice reduction methods such as the LLL algorithm by Lenstra, Lenstra and Lovats [6,7].

Nonlinear precoding promises significant performance enhancement if the channel matrix is sufficiently ill-conditioned [8], whereas optimized linear precoding schemes are shown to approach the performance of THP for moderate channels [9]. The increase of the modulation gap to capacity due to nonlinear precoding is quantified in [10], whereas [11] discusses the relation between encoding order and multi-user fairness. Above performance analyses assume perfect channel state information and ideal signal processing. However, since the gains of nonlinear precoding are highest precisely when channels are ill-conditioned, it is important to consider the impact of non-ideal factors. In this paper, we analyze the performance of NLP in the presence of quantization error, and of two sources of fluctuations in channel state information (CSI). We propose an alternative implementation for THP with enhanced robustness to quantization error, and propose a partial update algorithm for dealing with channel dynamics. We will compare the results with those obtained for an optimized linear precoder, such as the one presented in [8].

II. G.FAST DOWNSTREAM SYSTEM MODEL

We consider G.fast downstream transmission in which N transmitters located at the distribution point unit (DPU) transmit over a bundle of N twisted coppers pairs towards the corresponding N subscribers (or CPEs/users). More specifically, each user transmits discrete multi-tone (DMT) modulated data considering K tones over its twisted pair, while being synchronized with the transmission of all other users at the DPU. At the high frequencies employed in G.fast one obtains strong electromagnetic coupling between the twisted pairs, which is also referred to as far-end crosstalk. This results in the following per-tone MIMO system model

$$y_k = H_k x_k + z_k,\ k=\{1,\ldots,K\},$$

with y_k, x_k, z_k denoting the received vector, the transmit (precoded) vector and the received noise on tone k, respectively. H_k denotes the NxN matrix representing the MIMO channel on tone *k* with the diagonal elements $[H_k]_{n,n}$ denoting the complex direct channels and the off-diagonal elements $[H_k]_{n,m\ (n\neq m)}$ denoting the complex far-end crosstalk channels.

The transmit (precoded) vector represents the signal after a precoding operation on the data information symbol vector u_k.

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

The transmit vector has to satisfy power spectral density (PSD) constraints and aggregate transmit power constraints (ATP), as follows,

$$E\left(\left|[x_k]_n\right|^2\right) < PSD_k, \forall k, n,$$
$$\sum_k E\left(\left|[x_k]_n\right|^2\right) < ATP, \forall n.$$

Here, $E(x)$ denotes the expectation value of x, and $|x|$ denotes the absolute value of x. The objective is to design the precoding operation so as to maximize the achievable data rates R_n of each line *n*

$$R_n = \sum_k \eta f_s \log_2\left(1 + \frac{[SNR_k]_n}{\Gamma}\right),$$

with η denoting the transmission efficiency, f_s denoting the DMT symbol rate, SNR_k denoting the signal-to-noise ratio vector on tone *k* that depends on the precoding operation, and Γ denoting the SNR-gap with respect to capacity that depends on the modulation, noise margin and coding gain. In the remainder of this text, we continue with the per-tone case and drop the index k, for ease of representation.

III. Nonlinear precoder representations

A. General Representation

To represent a general zero-forcing nonlinear precoding scheme for N users, we suppose we wish to transmit random Nx1 vector u of complex constellation data, normalized to unit power. Denote by τ a value such that real and imaginary parts of the constellation are inside the range [-τ/2,τ/2], with some safety margin. We define a zero-forcing precoder $P = H^{-1}S$ where S is a diagonal scale matrix to be designed, and transmit the vector

$$x = P(u+\tau d(u))$$

Here d(u) is an Nx1 complex integer, chosen as a function of the data u to reduce the power of the transmitted vector x. . While we will sometimes use the notation d below to simplify notation, nonlinear dependence of d on u should be implicitly understood. The received vector is

$$y = HP(u+\tau d(u)) + z = Su + z + S\tau d(u)$$

such that individual receivers can independently remove the influence of the integer d by modulo arithmetic. The challenge is to implement the function d(u) with reasonable complexity.

B. Common QL implementation of THP

To obtain the common representation of THP, as laid out in [5], we use a variant of QR decomposition to decompose the channel inverse as $H^{-1} = QLS^{-1}$ where Q is unitary, L is lower-diagonal with unit diagonals, and S is diagonal. We define the scaled zero-forcing precoder $P = QL = H^{-1}S$. In the signal processing flow, the integer d and the elements of an intermediate vector $v = L(u+\tau d)$ are calculated sequentially as follows. At the n-th step, we know d_j and v_j for j =1,..,n-1 and wish to calculate d_n and v_n. If d_n were zero, then v_n would be

$$v_n' = \sum_{j=1}^{n-1} L_{nj}\left(u_j + \tau d_j\right) + u_n .$$

Hence the integer value of d_n that minimizes the magnitude of v_n is $d_n = \Phi_\tau(-v_n')$, where Φ_τ denotes rounding to nearest multiple of τ, and this minimal value is $v_n = v_n' + d_n$. Proceeding iteratively, all values of d and v are obtained. The output vector x is subsequently computed as x = Qv. As depicted in the left side of Figure 1, this can be implemented with a nonlinear unit that computes v from u, followed by a linear unit to compute x from v.

C. Proposed PL implementation of THP

The proposed alternate implementation is depicted on the right side of Figure 1. The first processing unit explicitly generates the shift vector d, using the recursion defined in the previous section, and then the linear processing unit computes x = P(u+d). We refer to this as the PL representation, because the stored coefficients in the second and first units are P and L, rather than Q and L. Mathematically, the two implementations are equivalent. In the next section, we examine some differences between these implementations when it comes to sensitivity to quantization. Following that, we examine the sensitivity of these schemes to changing channel gains.

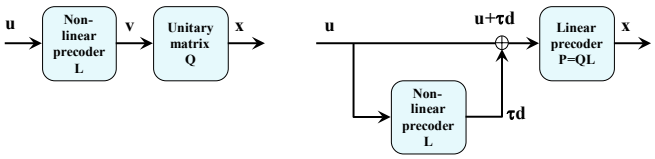

Fig. 1. QL (left) and PL (right) implementations of nonlinear precoding.

A qualitative advantage of the PL scheme is that it separates linear processing and shift-vector computation into separate modules. This can allow hardware and software designs from linear precoders to be re-used within the nonlinear structure.

IV. Coefficient Quantization

The level of quantization experienced by real-time signals in vectored systems plays an important role in determining performance. In this section we focus on quantization of precoder coefficients; these play a key role and are representative of other quantization effects, for example those due to finite precision arithmetic in the vector processor. In a THP implementation, we need to deal with quantization in the nonlinear unit (L matrix) and in the linear unit (Q or P matrices).

A. Linear processing unit

The linear processing unit simply calculates a matrix vector product x = Pw or x = Qv, where w = u+d or v = L(u+d) are the outputs of the nonlinear unit. In both cases, the main effect of coefficient quantization is residual interference due to imperfect zero forcing. For example, for a PL implementation, if we say the quantized precoder is P+Δ,

then the received value y' = H(P+Δ)w is the sum of the ideal y = HPw and the residual term r = HΔw. Modeling the quantization errors and data vector w as zero-mean and independent, the n-th residual term has variance

$$E\left[|r_n|^2\right] = \sum_{mj} |H_{nm}|^2 E\left[|\Delta_{mj}|^2\right] E\left[|w_j|^2\right]$$
$$= \beta^2 \sum_{mj} |H_{nm}|^2 |P_{mj}|^2 E\left[|w_j|^2\right]$$

Here we assume floating point representation of P, such that the variance of Δ_{mj} is approximately proportional to the squared magnitude of P_{mj}, by a factor β^2 that depends on the number of mantissa bits. For a QL representation, the corresponding residual interference would be

$$E\left[|r_n|^2\right] = \beta^2 \sum_{mj} |H_{nm}|^2 |Q_{mj}|^2 E\left[|v_j|^2\right].$$

For a given number of mantissa bits, the QL implementation has lower residual interference, since the unitary matrix Q is generally smaller than P, and the vector **v**, constrained to lie in the interval $[-\tau/2, \tau/2]$, is generally smaller than **w**.

B. *Nonlinear processing unit*

In the nonlinear unit, the PL implementation turns out to have advantages over the QL implementation. In the QL implementation, quantization errors in L are propagated forward into v and contribute to residual interference. In the PL implementation, errors in L can only lead to integer errors in d, which are removed at the receiver by modulo arithmetic. Hence L does not contribute to residual interference in this case. Errors in L, can however lead to sub-optimal choices for d_n, resulting in larger than ideal precoder output power. These errors would need to be handled in practice by leaving additional margin in the power constraints. Note that the computation of d_n is quite robust to errors in L, due to the nonlinear rounding function Φ_τ, and because the computation of element n takes into account any errors made in d_j, for j<n.

C. *Simulation results*

To further explore the effects of coefficient quantization in THP implementations, we simulate the impact of quantization on end-to-end data rate for a 212 MHz G.fast scenario. Simulations are performed using I51 channel data from [12]. An I51 cable is a bad quality untwisted single pair cable of 0.8 mm wire section. The data set from [12] provides far-end crosstalk measurements between ten such I51 cables arranged within a flexible conduit of 25 m length. Simulations assume a G.9700 power spectral density mask, an aggregate transmit power of 8 dBm, a maximum loading of 14 bits, a background noise of -140 dBm/Hz below 30 MHz and -150 dBm/Hz above 30 MHz. The gap to capacity is 10 dB, augmented for THP according to the table provided in [10] to take into account the effect of the receiver modulo on the trellis decoder.

For each implementation type (PL or QL), we simulated the impact of quantization in the linear unit with ideal nonlinear processing, and the impact of quantization in the nonlinear unit with ideal linear processing. Figure 2 shows the data rate obtained in each case, as a function of the number of mantissa bits used for each component (real and imaginary) in the floating point representation of the coefficient. The rate is normalized as a percentage of the rate obtained without quantization.

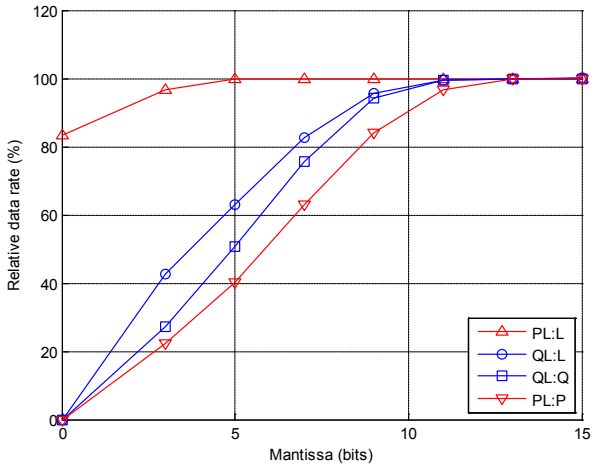

Fig. 2. Impact of quantization on relative data rate, for the QL and PL implementations of THP.

Quantization of the linear processing is labeled PL:P and QL:Q. As expected, the PL implementation requires more mantissa bits than the QL, although the difference is small in this case. Quantization of the nonlinear unit is labeled PL:L and QL:L. Here, ideal performance can be obtained with 5 mantissa bits with PL, compared with 11 for QL. If no bits are allocated to representing L in the PL scheme, performance reverts to that of a linear zero-forcing scheme. In this example, the THP data rate is almost 20% higher than the linear zero-forcing rate.

V. Channel Dynamics

A second consideration in implementing THP is the role of changing channel gains. Although channels gains in copper access systems are much more stable than say in wireless access, they do exhibit some dynamic behavior. Sources of channel dynamics include slow changes in ambient conditions [13], and abrupt termination impedance changes in coupled loops [14]. Even in linear systems, small changes of channel state can break the zero-forcing property of the precoder and lead to significant increases in residual interference. Fortunately such small changes can be estimated and compensated for with low complexity. In a THP implementation, it is similarly feasible to update the coefficients of the linear processing unit to maintain the zero-forcing property as channels change. However, performing the QL decomposition needed to ideally update the L matrix in the nonlinear unit is a non-trivial operation, which may not be feasible to do as often as the linear update. While using an out of date L matrix in THP implementation does not generate residual interference, it does in general lead to increased precoder output power, which must be handled by extra

margins, or by using the TIGA protocol to lower transmit gains. In this section, we will investigate the power penalty of out-of-date L matrices for THP. We will assume ideal processing (no quantization), in which case the QL and PL implementations are equivalent and interchangeable, though we will mostly use the PL notation for convenience. Furthermore, we will evaluate the use of inaccurate precoder matrix for both THP and optimized linear precoding.

A. Multiplicative perturbation model

A first model to analyze the implications for nonlinear precoding is through multiplicative perturbation in which the channel matrix changes from initial state H_0 to a perturbed state $H = (I+\Delta)H_0$. The zero-forcing precoder is initially $P_0 = H_0^{-1}S_0 = Q_0L_0$, but after the channel changes the linear processing matrix (P or equivalently Q) is updated to maintain the zero-forcing property, as $P = H^{-1}S_0 = QL_0$. To see how this effects the precoder output, we may further write

$$P = H_0^{-1}(I+\Delta)^{-1}S_0 = (I+H_0^{-1}\Delta H_0)^{-1}\,P_0.$$

For a given data vector u, the precoder output before the update would have been $x_0 = P_0(u+d)$. After the update, the precoder output is $x = P(u+d) = PP_0^{-1}x_0$, since the matrix L that determines d has not changed. The increase in the norm of x is bounded by

$$\|x\| \le \left\|\left(I + H_0^{-1}\Delta H_0\right)^{-1}\right\|\|x_0\| \le \frac{1}{1-c(H_0)\|\Delta\|}\|x_0\|$$

as long as $c(H_0)\ \|\Delta\| < 1$, where $c(A) = \|A\|\|A\|^{-1}$ is the condition number of matrix A [15]. Here, $\|A\|$ denotes the norm of A. Small channel perturbation generally lead to small increases in precoder output power, but sensitivity increases for channels with large condition number. When power increases are significant, a G.fast implementation can avoid violating power constraints by scaling the output down by a common factor, using the TIGA protocol.

The multiplicative perturbation model is suited to analyze sudden termination impedance changes, where the perturbation is chosen as $\Delta = H_N R$, where H_N is the near-end crosstalk channel matrix, and R is a diagonal matrix representing the reflection coefficient arising from an impedance mismatch between the transceiver and the line [14].

B. Element-wise perturbation model

A second model to analyze the implications for nonlinear precoding is through element-wise perturbation in which each element of the channel matrix changes from initial state $H_{nm,0}$ to a perturbed state $H_{nm} = (1+\delta_{nm})H_{nm,0}$. The perturbation size is proportional to the relative change in channel coefficient. The element-wise perturbation model is suited to analyze the effect of imperfect channel state information due to estimation errors or due to changes in ambient conditions [13].

C. Simulation Results

To quantify the impact of channel changes on the end-to-end data rate, we perform simulation experiments. In each case, we begin with a channel H_0 on every tone and corresponding nonlinear precoder defined by (P_0,L_0,S_0). We then perturb the channel, update the linear portion of the precoder as described above, and then empirically estimate the power increase on the worst case user. A common power back-off is applied to all users as needed to maintain power constraints, and the resulting bitloading and data rates are calculated. We use the same I51 channel data as in the coefficient quantization analysis.

In the first experiment, we simulate a channel change due to a change in termination impedance on a single alien line. We assume that initially the alien line is terminated by the line's characteristic impedance Z_0. A change in termination impedance Z will cause a reflection of magnitude $R = (Z-Z_0)/(Z+Z_0)$. The new data rate after power backoff is plotted as a function of the relative impedance change in percentage $100Z/Z_0$ in Figure 3. Near-end crosstalk coefficients are generated from the statistical channel model [16,17] applied to the empirical 99% worst case NEXT model [18]. We assume H_N is zero-diagonal.

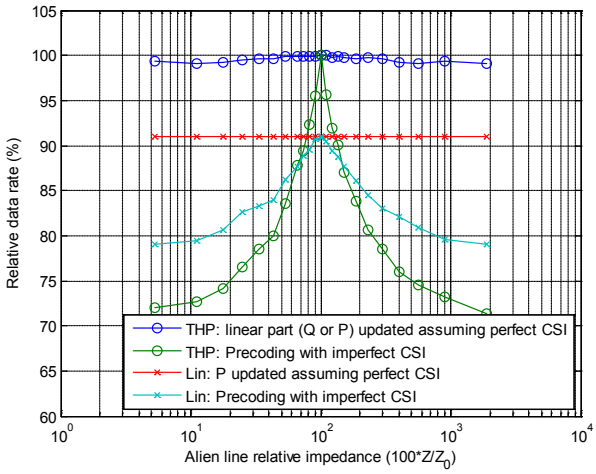

Fig. 3. Impact of alien line reflection on relative data rate.

In this example, there is 28% rate impact when the line is represented a high (open loop) or a low (short circuit) termination impedance. After the effects of a sudden termination change are observed, the transceivers will modify their frequency domain equalizer coefficients and apply fast rate adaptation to lower their bitloading to provide stable operation on the data rates presented by the green curve with open markers in Fig. 3. Updating precoder coefficients takes time beyond the error correction capabilities of the transceivers because of the required crosstalk channel estimation phase using probe sequences modulated on sync symbols, and because of the digital signal processing required to calculate new precoder coefficients from the measured residual channel [19]. The blue curve in Fig. 3 motivates the use of a fast partial update mechanism, in which the linear precoder is updated, but keeps the nonlinear L matrix unaltered. In this way, the computation of the QR

decomposition is avoided, and, if CSI can be acquired perfectly, the data rates are restored to 99% of their original values. In comparison, the data rate of the optimized linear precoder is reduced by 13% due to an alien impedance change without update (Fig. 3, cyan curve with 'x' markers), and not at all after update (red curve with 'x' markers).

In a second experiment, we use the element-wise model, with all channel coefficients changing by a random perturbation $H_{ij} = (1+\delta)H_{0,ij}$. This can model channel changes due to the ambient environment (e.g. temperature changes), as well as channel estimation errors. As shown in Figure 4, the resulting performance loss increases steadily with the magnitude of δ. The results show that under a wide range of imperfect CSI values, optimized linear precoding outperforms THP.

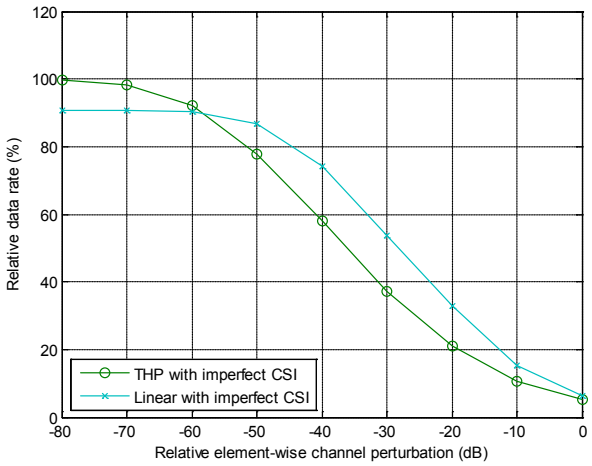

Fig. 4. Impact of imperfect channel state information on relative data rate.

VI. Conclusions

Nonlinear precoding via THP is a promising technique to achieve close to channel capacity on systems with severe far-end crosstalk. Quantization and channel dynamics are important factors that must be considered carefully in design of THP systems. We propose a PL-based THP implementation and show that it can be implemented with less precision in the nonlinear phase than is possible with a QL-based implementation. Channel dynamics and ability to accurately acquire channel state information are also important. Under several different models of channel dynamics, we show that fortunately the nonlinear unit coefficients do not need to be updated as frequently as the coefficients in the linear processing unit. We also show that under a wide range of CSI imperfections, optimized linear precoding outperforms THP. This can be understood by observing that nonlinear precoding exploits ill-conditioning to operate in an unstable optimum.

Acknowledgment

This work is supported by VLAIO through project IWT-5GBB.

References

[1] 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, no. 5, pp. 860–863, May 2007.

[2] M. Guenach, C. Nuzman, P. Tsiaflakis, and J. Maes, "Power optimization in vectored and non-vectored G.fast transmission", in IEEE Global Communications Conference (GLOBECOM), Dec. 2014, Austin, TX, USA, pp. 2229-2233

[3] M. Tomlinson, "New Automatic Equalizer Employing Modulo Arithmetic" Electronics Letters, 7(5-6), pp.138-139, Mar. 1971

[4] H. Harashima, and H. Miyakawa, "Matched-Transmission Technique for Channels with Inter Symbol Interference" IEEE Trans. on Communications, 20(4), pp. 774-780, Aug. 1972

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

[6] A.K. Lenstra, H.W. Lenstra, and L. Lovasz, 'Factoring polynomials with rational coefficients', *Math. Ann.*, 261:515–534, 1982.

[7] C. Windpassinger, and R.F.H. Fischer, 'Low-Complexity Near-Maximum-Likelihood Detection and Precoding for MIMO Systems using Lattice Reduction', proceedings of ITW2003, March, 2003.

[8] W. Lanneer, M. Moonen, P. Tsiaflakis, and J. Maes, "Linear and nonlinear precoding based dynamic spectrum management for downstream vectored G.fast transmission", in IEEE Global Communications Conference (GLOBECOM), Dec. 2015, San Diego, CA, USA

[9] R. Strobel, "Zero-forcing and MMSE precoding for G.fast", in IEEE Global Communications Conference (GLOBECOM), Dec. 2015, San Diego, CA, USA

[10] J. Neckebroek, M. Moeneclaey, W.Coomans, M. Guenach, P. Tsiaflakis, R.B. Moraes, and J. Maes, "Novel bitloading algorithms for coded G.fast DSL transmission with linear and nonlinear precoding", in IEEE International Conference on Communications, Jun 2015, London, UK.

[11] M. Hekrdla, A. Matera, W. Wang, D. Wei, and U. Spagnolini, "Ordered Tomlinson-Harashima Precoding in G.fast Downstream", in IEEE Global Communications Conference (GLOBECOM), Dec. 2015, San Diego, CA, USA

[12] A. Carrick, and T. Bongrd, 'G.fast: Release of full I51 cable measurements, ITU-T SG15 Q4 contribution C 13-02-07, Geneva, Feb. 2013.

[13] D. Statovci, T. Magesacher, M. Wolkerstorfer, and E. Medeiros, 'Analysis of Fast Initialization for Vectored Wireline Systems', in IEEE Global Communications Conference (GLOBECOM), Dec. 2013, Atlanta, GA, USA

[14] E. Medeiros, T. Magesacher, P. Ödling, D. Wei, X. Wang, Q. Li, P.-E. Eriksson, C. Lu, J. Boschma, and B. van den Heuvel, "Modeling Alien-Line Impedance Mismatch in Wideband Vectored Wireline Systems", IEEE Communications Letters, vol. 18, no. 9, pp 1527-1530, 2014.

[15] B. Noble and J. W. Daniel, Applied Linear Algebra, 3rd Edition, Prentice Hall.

[16] J. Maes, M. Guenach, and M. Peeters, 'Statistical MIMO channel model for gain quantification of DSL crosstalk mitigation techniques', proceedings of IEEE International Conference on Communications (ICC09), Dresden, May 2009.

[17] D. Van Bruyssel, and J. Maes, 'G.fast: Statistical model for Near End Crosstalk', ITU-T contribution 2016-06-Q4-040, June 2016.

[18] ITU-T, 'Test Procedures for Digital Subscriber Line (DSL) Transceivers', Recommendation G.996.1, February 2006.

[19] J . Maes, C. Nuzman, A. Van Wijngaarden, D. Van Bruyssel, 'Pilot-based crosstalk channel estimation for vector-enabled VDSL systems', proceedings of the 44th Annual Conference of Information Sciences and Systems (CISS 2010), Princeton, March 2010.