Making Power Line Communications More Reliable by Using Optimal Soft Information

Stefano Galli

Telcordia Technologies, Inc.

445 South Street, Morristown

New Jersey 07960, USA

Abstract

The Power Line Channel (PLC) is a very harsh transmission medium because of its frequency selective and time-varying transfer function that exhibits several notches in the band of interest. Moreover, a strong time varying noise is normally present and poses even more challenges to the modem designer. The time-varying and frequency selective nature of the PLC requires that adaptive and, possibly, nonlinear equalization be performed at the receiver. However, adaptive receivers suffer from severe performance degradation on time varying and Inter-Symbol Interference (ISI)-impaired links because of the unreliability of the hard decisions used for channel estimation and tracking. In particular, unreliable hard-decisions are a major issue especially in harsh and time-varying environments. In the present paper, we propose to alleviate this performance loss by the use of optimal soft decisions: the Nonlinear Minimum Mean Square Error (NL-MMSE) filtered and fixed-lag smoothed estimates of the transmitted symbols.

I. Introduction

The PLC is a very harsh transmission medium because of its frequency selective and time-varying transfer function that exhibits several notches in the band of interest; moreover, a strong time varying noise (colored background noise, narrowband noise, impulsive noises) is present and poses even more challenges to the modem designer [1], [2], [3], [4]. The time-varying and frequency selective nature of the PLC requires that adaptive and nonlinear equalization be performed at the receiver. However, as will be shown in Sect. III.A, the optimal NL-MMSE channel estimator is *intrinsically non-recursive*. This result suggests that pursuing the truly optimal NL-MMSE channel estimator would lead to an impractical non-recursive channel estimator. For this reason, another approach is followed here. Instead of pursuing the optimal channel estimator, we will try to enhance the performance of linear receivers by introducing some sort of nonlinearity without giving up the property of recursivity.

We have ascertained that most of the performance degradation suffered by adaptive receivers is due to the unreliability of hard-decisions [9]. In the present paper, we propose to alleviate this performance loss by the use of optimal soft decisions. In particular, it is shown here that it is possible to compute recursively the NL-MMSE filtered and fixed-lag smoothed estimates of the transmitted symbols and use those estimates in place of the usual hard decisions for channel estimation and tracking. It is worth pointing out that the use of these soft decisions is not the result of mere intuition, but the consequence of the correct statement of the problem of NL-MMSE estimation of the channel.

The paper is organized as follows. The considered system model is described in Sect. II. The role of the NL-MMSE filtered and fixed-lag smoothed estimates in nonlinear estimation is pointed out in Sect. III. In Sect. IV, a method for the recursive computation of the NL-MMSE estimates of the transmitted symbols is described. Finally, conclusive remarks are drawn in Sect. V.

II. System Model

Referring to a digital link impaired by time-varying multipath phenomena and noise, the baud-rate sampled (T_S-sampled) sequence received at the output of the equivalent low-pass randomly time varying ISI channel can be modeled as:

$$\begin{aligned} y(i) &= \sum_{m=0}^{L-1} g(i;m)s(i-m)+v(i) = \\ &= \underline{G}^T(i)\underline{x}(\mathbf{i})+v(i),\ i\geq 1 \end{aligned}, \qquad (1)$$

where the transmitted data sequence $\{s(k)\in A\equiv\{s_1,\dots,s_M\}\subset C\}$ consists of M-ary generally complex independent identically distributed symbols taken from an assigned modulation constellation A; L is the delay spread of the channel expressed in terms of multiples of the signaling time; $\{g(i;m)\equiv g(t=iT_S;\ \tau=mT_S)\in C,\ 0\leq m\leq L-1,\ i\geq 1\}$ is the T_S-sampled complex time-variant input-delay spread function $g(t;\tau)$ of the overall link including the transmitting filter, the transmission channel and the receiving filter; $\underline{G}^T(i)\equiv[g(i;0)\dots g(i;L-1)]\in C^L$ is the resulting L-variate complex impulse-response vector of the channel at the i-th epoch; $\underline{x}(\mathbf{i})\equiv[s(i)\dots s(i-L+1)]^T$ is the vector containing the last L transmitted symbols; and $\{v(i)\in C^1\}$ is a complex zero-mean colored random sequence which accounts for the noise.

The L-variate random sequence $\underline{x}(i) \in A^L \equiv \{\xi_1, ..., \xi_N\} \subset C^L$ is a first-order Markov chain known as "state sequence" of the channel [7]. The random variable $\{\underline{X}(i)\}$ may assume $N \equiv M^L$ distinct and equally likely values $\{\xi_i, 1 \le i \le N\}$, corresponding to the L-long ordered sub-sequences of constellation symbols, that is:

$$\underline{\xi}_j \equiv \left[\sigma_1^{(j)} \ldots \sigma_L^{(j)}\right]^T, 1 \le j \le N, \tag{2}$$

where $\sigma_i^{(j)} \in A$ ($1 \le i \le L$, $1 \le j \le N$) is the i-th component of the j-th determination ξ_j of the channel-state.

III. NL-MMSE Estimates of the Transmitted Symbols: a *Necessary Consequence* of the NL-MMSE Channel Estimation Problem

The basic novelty of the proposed approach is to use the NL-MMSE estimates of the transmitted symbols in place of the usual hard decision data for channel estimation and tracking. The intuitive idea behind this approach is that the use of "estimates" in place of "decisions" is advantageous whenever the reliability of the hard decisions is low. As an example, this concept was also developed in a seminal work by D. P. Taylor [5], where a sub-optimum MMSE estimate of the transmitted symbols was used in the feedback section in place of hard decisions. The usefulness of NL-MMSE estimates for mitigating the effects of erroneous hard decisions has also been recently pointed out in [6], although no method for computing them was suggested. However, the use of NL-MMSE estimates in place of hard decisions is not only a mere intuition. In fact, as will be shown in the following, it can be proved that this is a straightforward consequence of the MMSE channel estimation problem.

On the basis of the innovations approach proposed by Kailath in [10], it is possible to express the MMSE estimate of the channel sequence $\underline{G}(i)$ as a linear combination of the innovations process of the observations $y(i)$:

$$\hat{\underline{G}}(i/i) \equiv E\left\{\underline{G}(i) \mid \underline{y}_1^i\right\} = \sum_{k=1}^{i} \underline{c}_k \eta(k), \tag{3}$$

where $\underline{y}_1^i$ denotes the sequence of the available observations from step 1 to step i, the sequence $\eta(1)$, $\eta(2)$, …, $\eta(i)$ is the innovations sequence, and the coefficients $\underline{c}_k$ are to be determined so as to minimize the mean square value of the estimation error $\underline{\varepsilon}(i) = \underline{G}(i) - \hat{\underline{G}}(i/i)$. Although only the case of linear innovations is specifically addressed in [10], eq.(3) still holds for the more general case of NL-MMSE estimation and nonlinear innovations as long as it is taken into account that the nonlinear innovations are Martingale Difference sequences (and not simply uncorrelated sequences as are the linear innovations) and are not Gaussian (see, for example, the considerations made in Sect. III of [11]). In fact, the estimator (3) is indeed a nonlinear estimator since the general innovations $\eta(k)$ are not linearly dependent on the observation sequence $\underline{y}_1^i$.

The innovations of the observations can be easily expressed by resorting to the Doob-Meyer representation [12], that decomposes the random variable $y(i)$ into the sum of two terms, a predictable term and a non-predictable term, the former being the MMSE one-step prediction of the observations and the latter being the innovation of the observations. In so doing, we obtain:

$$y(i) = \hat{y}(i/i-1) + \eta(i) = E\left\{y(i) \mid \underline{y}_1^{i-1}\right\} + \eta(i). \tag{4}$$

Eq.(3) can now be rewritten as the following:

$$\begin{aligned}\hat{\underline{G}}(i/i) &\equiv E\left\{\underline{G}(i) \mid \underline{y}_1^i\right\} = \sum_{k=1}^{i-1} \underline{c}_k \eta(k) + \underline{c}_i \eta(i) = \\ &= \sum_{k=1}^{i-1} \underline{c}_k \eta(k) + \underline{c}_i \left[y(i) - \hat{y}(i/i-1)\right]\end{aligned}. \tag{5}$$

By definition, the summation on the right-hand side of eq.(4) represents the estimate of the channel vector at the (i-1)-th epoch given the observations up to step (i-1), i.e. $\hat{\underline{G}}(i-1/i-1) \equiv E\left\{\underline{G}(i-1) \mid \underline{y}_1^{i-1}\right\}$, so that we can finally write:

$$\hat{\underline{G}}(i/i) = \hat{\underline{G}}(i-1/i-1) + \underline{c}_i \left[y(i) - \hat{y}(i/i-1)\right]. \tag{6}$$

If $\{\underline{G}(i)\}$ were a Gauss-Markov process, the estimate $\hat{\underline{G}}(i-1/i-1)$ would be replaced by the one-step MMSE prediction $\hat{\underline{G}}(i/i-1)$ [12], so that we would have:

$$\hat{\underline{G}}(i/i) = \hat{\underline{G}}(i/i-1) + \underline{c}_i \left[y(i) - \hat{y}(i/i-1)\right]. \tag{7}$$

As far as the one-step prediction of the observations $\hat{y}(i/i-1)$ is concerned, we can write the following:

$$\hat{y}(i/i-1) \equiv E\left\{y(i) \mid \underline{y}_1^{i-1}\right\} = E\left\{\underline{G}(i)^T \underline{x}(i) \mid \underline{y}_1^{i-1}\right\}. \tag{8}$$

Now, depending on the approximations made, the computation of the expected value in (8) leads to the two cases discussed in Sects. III.A and III.B. Considerations on the optimal nonlinear estimate of the channel are made in Section III.C.

A. *Computation of the one-step MMSE prediction of the observations: approximation of correct hard decisions*

Under the assumption that the transmitted symbols are available to the receiver (e.g., for the case of training sequences) or, equivalently, that the detector is delivering correct hard-decisions, we can rewrite eq.(8) as the following:

$$\begin{aligned}\hat{y}(i/i-1) &= E\left\{y(i) \mid \underline{y}_1^{i-1}, \underline{x}(i)\right\} = \\ &= E\left\{\underline{G}(i)^T \mid \underline{y}_1^{i-1}, \underline{x}(i)\right\} \underline{x}(i)\end{aligned} \quad (9)$$

The direct estimation of the channel $\underline{G}(i)$ given $\underline{y}_1^{i-1}$ is a non-Gaussian problem since the probability density function of $\underline{G}(i)$ conditioned to the available observations $\underline{y}_1^{i-1}$ is non-Gaussian. However, under the assumption of correct decisions, conditioning $\underline{G}(i)$ to both $\underline{y}_1^{i-1}$ and $\underline{x}(i)$ turns the channel estimation problem into a Gaussian problem.

The approximation of correct hard decisions is usually used to give rise to conventional decision-driven channel estimators. However, this method turns an intrinsically non-Gaussian problem into a Gaussian one and, therefore, conventional RLS-like decision-driven channel estimators represent the suboptimal linear MMSE solution to the problem of the estimation of the channel vector $\underline{G}(i)$.

B. Computation of the one-step MMSE prediction of the observations: approximation of the conditional independence between {X(i)} and {G(i)}, given the observations.

The general case where no assumption is made about the reliability of the decisions is addressed in this subsection. We start by pointing out that although the sequences $\{\underline{G}(i)\}$ and $\{\underline{x}(i)\}$ are statistically independent, they are not independent anymore when they are both conditioned to $\underline{y}_1^{i-1}$. This implies that the correct computation of the expected value in (8) is not straightforward since it requires the computation of all the cumulants. However, it is here conjectured that the approximation of considering $\{\underline{G}(i)\}$ and $\{\underline{x}(i)\}$ at least uncorrelated, even when conditioned to $\underline{y}_1^{i-1}$, is not a harsh approximation. Moreover, it is also conjectured that the approximation of correct hard decisions is a much harsher approximation than the one made here. The validity of this approximation is more thoroughly discussed in [9]. In so doing, we may rewrite eq.(8) in the following way:

$$\begin{aligned}\hat{y}(i/i-1) &= E\left\{\underline{G}(i)^T \underline{x}(i) \mid \underline{y}_1^{i-1}\right\} \approx \\ &\approx E\left\{\underline{G}(i)^T \mid \underline{y}_1^{i-1}\right\} E\left\{\underline{x}(i) \mid \underline{y}_1^{i-1}\right\} = \\ &= \hat{\underline{G}}(i/i-1)^T E\left\{\underline{x}(i) \mid \underline{y}_1^{i-1}\right\}.\end{aligned} \quad (10)$$

It is now clear that the term $E\left\{\underline{x}(i) \mid y_1^{i-1}\right\}$, i.e. the MMSE one-step prediction of the state sequence of the channel, appears in the problem of MMSE estimation of the channel impulse response $\underline{G}(i)$ and that the computation of the expected value $E\left\{\underline{x}(i) \mid y_1^{i-1}\right\}$, if simple, would allow us to avoid the overly-optimistic approximation of correct hard decisions.

Since the MMSE estimate of a random vector is the vector consisting of the MMSE estimates of the single elements, we can rewrite eq.(10):

$$E\left\{\underline{x}(i) \mid y_1^{i-1}\right\} \equiv \begin{bmatrix} E\left\{s(i) \mid \underline{y}_1^{i-1}\right\} \\ E\left\{s(i-1) \mid \underline{y}_1^{i-1}\right\} \\ \vdots \\ E\left\{s(i-L+1) \mid \underline{y}_1^{i-1}\right\} \end{bmatrix} \equiv \equiv \hat{\underline{s}}_{NL-MMSE}(i/i-1) \quad (11)$$

Now, we can finally rewrite eq.(7) as:

$$\begin{aligned}&\hat{\underline{G}}(i/i) \approx \\ &\approx \hat{\underline{G}}(i/i-1) + \underline{c}_i \left[y(i) - \hat{\underline{G}}(i/i-1)^T \hat{\underline{s}}_{NL-MMSE}(i/i-1)\right]\end{aligned} \quad (12)$$

The problem of estimating the sequence $\underline{x}(i)$ on the basis of the available observations $\underline{y}_1^{i-1}$ is a non-Gaussian problem since the probability density function of $\underline{x}(i)$ conditioned to the available observations $\underline{y}_1^{i-1}$ is non-Gaussian. Therefore, the exact computation of the expected value $E\left\{\underline{x}(i) \mid \underline{y}_1^{i-1}\right\}$ is a problem of NL-MMSE estimation. As eqs.(11), (12) clearly show, the NL-MMSE estimate of the channel $\underline{G}(i)$ requires the NL-MMSE estimate of the channel state vector $\underline{x}(i)$ which, in turn, requires the computation of the NL-MMSE predicted, filtered and fixed-lag smoothed estimates of the transmitted symbols. This confirms that these estimates are *necessary* to NL-MMSE channel estimation.

C. Optimal NL-MMSE estimation

In order to address properly the problem of optimal NL-MMSE estimation of $\{G(i)\}$, let us define the following sequences

$$\begin{aligned}\underline{\zeta}(i) &\equiv \hat{\underline{G}}(i/i) - \hat{\underline{G}}(i/i-1) = \\ &= E\left\{\underline{G}(i) \mid \underline{y}_1^{i}\right\} - E\left\{\underline{G}(i) \mid \underline{y}_1^{i-1}\right\},\end{aligned} \quad (13.a)$$

$$\begin{aligned}\eta(i) &\equiv y(i) - \hat{y}(i/i-1) = \\ &= y(i) - E\left\{y(i) \mid \underline{y}_1^{i-1}\right\},\end{aligned} \quad (13.b)$$

as the *estimation* and *observation* innovation sequences, respectively. It is possible to show that the innovations sequences $\{\zeta(i)\}$ and $\{\eta(i)\}$ have the property of being MD sequences. Roughly speaking, the MD property is something intermediate between statistical independence and uncorrelation, so that independence implies the MD property that, in turn, implies uncorrelation. If a suboptimal linear MMSE solution for the estimation of $\underline{G}(i)$ is desired, it is sufficient to consider (13.a) and (13.b) as uncorrelated sequences. However, if the optimal NL-MMSE solution is desired, the additional MD structure embedded in the innovations sequences cannot be neglected and has to be taken into account. This may be done by exploiting the MD Representation Theorem [11].

MD Representation Theorem. *The estimation innovations sequence (13.a) is an MD sequence with respect to the σ-algebra built on* $\{\underline{y}_1^{i-1}\}$ *and, therefore, may be represented as a transformation of the observations innovation (13.b) sequence as in the following:*

$$\underline{\zeta}(i) = \underline{C}_G(i/i)\eta(i) \tag{14}$$

where $\underline{C}_G(i)$ *is an* *adapted* *sequence with respect to the σ-algebra built on* $\{\underline{y}_1^{i-1}\}$ *and can be computed as follows:*

$$\underline{C}_G(i/i) = \frac{E\left\{\underline{\zeta}(i)\eta^*(i) \mid \underline{y}_1^i\right\}}{E\left\{\eta(i)\eta^*(i) \mid \underline{y}_1^i\right\}} \tag{15}$$

The theorem states that the gain sequence is an adapted sequence (see Definition 3.2 in [11]), thus the optimal NL-MMSE estimator of the sequence $\{\underline{G}(i)\}$ does not appear to be recursive, since recursivity is only ensured by the property of predictability (see Definition 4.1 in [11]). However, it has been shown that in some special cases the gain sequence is predictable with respect to the σ-algebra built on $\{\underline{y}_1^{i-1}\}$. In those cases it is possible to rewrite (15) as the following:

$$\begin{aligned}\underline{C}_G(i/i) &= \frac{E\left\{\underline{\zeta}(i)\eta^*(i) \mid \underline{y}_1^i\right\}}{E\left\{\eta(i)\eta^*(i) \mid \underline{y}_1^i\right\}} = \\ &= \frac{E\left\{\underline{\zeta}(i)\eta^*(i) \mid \underline{y}_1^{i-1}\right\}}{E\left\{\eta(i)\eta^*(i) \mid \underline{y}_1^{i-1}\right\}} = \underline{C}_G(i/i-1)\end{aligned} \tag{16}$$

The predictable form of the theorem has been proven only in the following three cases:

1. Continuous-time case for signals observed in Gaussian white noise [13];
2. Continuous-time case for signals observed through point processes [14];
3. Discrete-time case for signals observed through binary point processes [14].

Unfortunately, the predictability of the gain sequence has been proven to be *false* in the important case of discrete-time observations in white Gaussian noise, the most recurrent case in communications problems. Therefore, in principle, the optimal NL-MMSE estimator of any sequence observed in white Gaussian noise is *intrinsically non-recursive* (see [9] for more details).

IV. Computation of the NL-MMSE Estimates

It has been recently proposed [8] to use the posterior probabilities of the message symbols for the computation of the NL-MMSE estimates of the transmitted symbols. However, this approach is optimal only for the case of systems with no memory since it can only yield the filtered NL-MMSE estimates of the transmitted symbols, i.e. $\hat{s}_{NL-MMSE}(i/i) \equiv E\left\{\underline{s}(i) \mid y_1^i\right\}$. As will be shown in the following, if the posterior probabilities of the states of the channel are used in place of those of the message symbols, the fixed-lag smoothed estimates can also be obtained, i.e. $\hat{s}_{NL-MMSE}(i-d/i) \equiv E\left\{s(i-d) \mid \underline{y}_1^i\right\}$. In the case of systems with memory, the availability of fixed-lag smoothed estimates ensures better performances.

The MMSE estimate of the symbol $s(i)$ on the basis of the observations from step 1 to step i is given by the following definitory relationship:

$$\begin{aligned}\hat{s}_{MMSE}(i) &\equiv E\left\{s(i) \mid \underline{y}_1^i\right\} \equiv \\ &\equiv \sum_{k=1}^{M} s_k \, Pr\left(s(i) = s_k \mid \underline{y}_1^i\right),\end{aligned} \tag{3}$$

As is well known, in Gaussian environments a solution to the estimation of the transmitted symbol $s(i)$ can be directly obtained via a linear MMSE Wiener-like sliding-window filter for the desired data-stream $\{s(i)\}$. However, the problem is intrinsically non-Gaussian since the probability density function of $s(i)$ conditioned to the available observations $\underline{y}_1^i$ is non-Gaussian, so that the linear Wiener-like approach is characterized by modest performances. Therefore, the exact computation of the expected value $E\left\{s(i) \mid \underline{y}_1^i\right\}$ is a problem of NL-MMSE estimation. It is possible to show that the NL-MMSE estimates of the transmitted symbols can be expressed as follows [9]:

$$\begin{aligned}\underline{\hat{s}}_{NL-MMSE}(i/i) &\equiv \begin{bmatrix} \hat{s}_{NL-MMSE}(i/i) \\ \hat{s}_{NL-MMSE}(i-1/i) \\ \vdots \\ \hat{s}_{NL-MMSE}(i-L+1/i) \end{bmatrix} = \\ &= \begin{bmatrix} \sigma_1^{(1)} & \sigma_1^{(2)} & \cdots & \sigma_1^{(N)} \\ \sigma_2^{(1)} & \sigma_2^{(2)} & \cdots & \sigma_2^{(N)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_L^{(1)} & \sigma_L^{(2)} & \cdots & \sigma_L^{(N)} \end{bmatrix} \begin{bmatrix} P\left(\underline{x}(i) = \underline{\xi}_1 \mid \underline{y}_1^i\right) \\ P\left(\underline{x}(i) = \underline{\xi}_2 \mid \underline{y}_1^i\right) \\ \vdots \\ P\left(\underline{x}(i) = \underline{\xi}_N \mid \underline{y}_1^i\right) \end{bmatrix} = \\ &= \Xi \, \underline{\pi}(i/i)\end{aligned} \tag{4}$$

where we have introduced the LxN matrix Ξ whose columns consists of the vectors $\{\xi_i\}$ of (2) and the N-variate vector of the APPs of the state sequence of the channel defined as

$$\underline{\pi}(i/i) \equiv \left[P\left(\underline{x}(i) = \underline{\xi}_1 \mid \underline{y}_1^i\right) \ldots P\left(\underline{x}(i) = \underline{\xi}_N \mid \underline{y}_1^i\right)\right]^T. \tag{5}$$

The relationship in eq.(4) shows that the NL-MMSE estimation of the last L transmitted symbol $\underline{\hat{s}}_{NL-MMSE}(i/i)$ can be expressed as a linear transformation of the APPs of the state of the channel. The estimate $\underline{\hat{s}}_{NL-MMSE}(i/i)$ can be viewed as a linear combination of all the possible realizations of the state of the channel with probabilistic weights equal to the APPs of the states of the channel. Moreover, the proposed approach allows us to obtain the filtered NL-MMSE estimate of $s(i)$, i.e. $\hat{\underline{s}}_{NL-MMSE}(i/i) \equiv E\left\{s(i) \mid \underline{y}_1^i\right\}$, and the fixed-lag smoothed NL-MMSE estimates of $s(i)$, i.e. $\hat{\underline{s}}_{NL-MMSE}(i-d/i) \equiv E\left\{s(i-d) \mid \underline{y}_1^i\right\}$ for $1 \leq d \leq L$-1.

V. Conclusions

A new and efficient method for generating NL-MMSE estimates on the basis of the APPs of the states of the channel has been developed. The proposed approach to compute NL-MMSE estimates holds in general and can be applied to all those problems that yield to a state-space representation. Several applications can be foreseen in a wide variety of detection and estimation problems, such as in multiuser detection, iterative decoding and in all those situations where hard decisions are heavily employed despite the low reliability of the detection process. Moreover, this method makes the use of probabilistic symbol detectors, as the ones described in [7] and references therein, very appealing because they are able to generate three kinds of information: "hard" information (the hard decisions), "soft" information (the APPs) and a "quasi-soft" intermediate case represented by the NL-MMSE filtered and fixed lag smoothed estimates of the transmitted symbols.

References

[1] T. Banwell, S. Galli, "A new approach to the modeling of the transfer function of the power line channel", *IEEE International Symposium on Power Line Communications and Its Applications*, ISPLC 2001, Malmö, Sweden, April 4-6, 2001.

[2] M. Zimmermann, K. Dostert, "An Analysis of the Broadband Noise Scenario in Power-Line Networks", *IEEE International Symposium on Power Line Communications and Its Applications*, ISPLC 2000, Limerick, Ireland, April 5-7, 2000.

[3] H. Philipps, "Development of a Statistical Model for Power Line Communications Channels", *IEEE International Symposium on Power Line Communications and Its Applications*, ISPLC 2000, Limerick, Ireland, April 5-7, 2000.

[4] C. Assimakopoulos, N. Pavlidou, "Measurements and Modeling of In-House Power Lines Installation for Broadband Communications", *IEEE International Symposium on Power Line Communications and Its Applications*, ISPLC 2001, Malmö, Sweden, April 4-6, 2001.

[5] D.P. Taylor, "The Estimate Feedback Equalizer: A Suboptimum Nonlinear Receiver", *IEEE Trans. Commun.*, vol. 21, pp.979-990. Sept. 1973.

[6] F. Tarköy, "MMSE-Optimal Feedback and its Applications", *Proceedings of the IEEE Intern. Symp. on Infor. Theory*, ISIT'95, Whistler, Canada, 17-22 Sep. 1995.

[7] Y. Li, B. Vucetic, Y. Sato, "Optimum Soft-Output Detection for Channels with Intersymbol Interference", *IEEE Trans. Inform. Theory*, vol.41, no.3, pp.704-713, May 1995.

[8] X. Wang, H.V. Poor, "Iterative (Turbo) Soft Interference Cancellation and Decoding for Coded CDMA", *IEEE Trans. Commun*., vol.47, no.7, July 1999.

[9] S. Galli, "A New Family of Soft-Output Adaptive Receivers Exploiting Non-Linear MMSE Estimates of the Transmitted Symbols for TDMA-Based Wireless Links", submitted to *IEEE Transactions on Communications*, June 2000.

[10] T. Kailath, "The Innovations Approach to Detection and Estimation Theory", *Proceedings of the IEEE*, vol. 58, pp. 680-695, 1970.

[11] A. Segall, "Stochastic Processes in Estimation Theory", *IEEE Trans. Inform. Theory*, vol.22, no.3, pp.275-286, May 1976.

[12] G. Kallianpur, *Stochastic Filtering Theory*, Springer, New York, 1980.

[13] M. Fujisaki, F. Kallianpur, H. Kunita, "Stochastic Differential Equations for the Nonlinear Filtering Problem", *Osaka J. Math.*, vol.9, pp.19-40, 1972.

[14] P.A. Meyer, "Martingales and Stochastic Integrals", *Lecture Notes in Mathematics*, vol.284, New York, Springer, 1972.