ABSTRACT

Soft-Input Soft-Output (SISO) equalizers based on linear filters have proven to be good, low complexity, alternatives to trellis-based SISO equalizers. In particular, the Soft Interference Canceller (SIC) has recently received great interest, especially for receivers performing Turbo Equalization. In this paper, we modify the way in which the SIC incorporates soft information. In existing literature the input to the cancellation filter is the expectation of the symbols based solely on the a-priori probabilities coming from the decoder, whereas here we propose to use the conditional expectation of those symbols, given both the a-priori probabilities and the received sequence. This modification results in performance gains at the expense of increased computational complexity. However, by introducing an approximation to the aforementioned algorithm a linear complexity SISO equalizer can be derived. Simulation results for an 8-PSK constellation and hostile radio channels have shown the effectiveness of the proposed algorithms in mitigating the Inter-Symbol Interference (ISI).

1. INTRODUCTION

Turbo Equalization [1] was motivated by the breakthrough of Turbo Codes [2], and has emerged as a promising technique for drastic reduction of the intersymbol interference in frequency selective wireless channels. Unfortunately, the trellis based SISO equalizer of [1] can be a heavy computational burden to wireless receivers with limited processing power. However, when the channel impulse response has long delay spread and the employed modulations are of high order, linear filter based SISO equalizers offer tremendous complexity reduction over trellis based ones, without sacrificing considerable performance.

In this context, it was proposed in [3] to replace the trellis based equalizer by an adaptive SIC of linear complexity. In [4] an MMSE-optimal equalizer based on linear filters was derived and it was proven that several other algorithms (such as the one in [3]) could be viewed as approximations of this one. In [5], a modified version of the sliding window algorithm of [4] was derived having similar performance to the original one while offering reduced computational complexity via the use of a Cholesky factorization technique. In [6], the authors modified the algorithm of [4] which involves complex valued matrices, into an algorithm that uses augmented real valued matrices yielding better performance at approximately the same complexity. More recently, the authors of [7] derived the theoretical (time invariant) transfer function of an MMSE optimal equalizer and showed that this equalizer reduces to a linear equalizer in the case of no a-priori information or to an MMSE SIC in the case of perfect a-priori information. Their algorithm, was shown to be identical to a low complexity algorithm derived in [8] in the case where the equalizer filters are restricted to finite length. In [9], channel output information was properly incorporated in the input of the SIC cancellation filter.

In the proposed turbo equalizer, we split the problem of a-priori probabilities based equalization into two distinct optimization problems. The first problem consists in the estimation of past and future symbols using a-priori probabilities and channel output information, while the second problem is the estimation of the current symbol based on past and future symbols. The solution to the first problem is to use an MMSE equalizer similar to the one developed in [8], but modified appropriately so as to provide all the required symbols instead of computing only the current symbol estimate. For the second problem an MMSE SIC is employed, which has been developed under the assumption that its input symbols are actually correct symbols (in practice they are provided by the aforementioned equalizer). As shown experimentally, the proposed approach, so-called Conditional Expectation - Soft Interference Canceller (CE-SIC), exhibits similar performance to the exact MMSE solution of [8], at almost the same computational cost. Furthermore, an approximate version, so-called Approximate Conditional Expectation - Soft Interference Canceller (ACE-SIC) has been derived, which has linear complexity and is experimentally shown to exhibit very good performance characteristics making it suitable for high data rate wireless communications.
The output of the channel is corrupted by complex-valued anti-causal parts, respectively, of the channel impulse response.

The output of the MAP decoder is extrinsic then non-negligible.

Equalization iterative detection algorithm can be found in [4].

We assume that the communication channel is frequency selective and constant during the packet transmission, so that the output of the channel (and input to the receiver) can be modelled as

\[ z_n = \sum_{i=-k+1}^{L_2} h_i x_{n-i} + w_n, \]

where \( L_1, L_2 + 1 \) denote the lengths of the anti-causal and causal parts, respectively, of the channel impulse response. The output of the multipath channel is corrupted by complex-valued Additive White Gaussian Noise (AWGN) \( w_n \).

At the receiver, we employ an equalizer to compute soft estimates of the transmitted symbols. As a part of the equalizer is also a scheme that transforms the soft estimates of the symbols into soft estimates of the bits corresponding to those symbols. The output of the equalizer is the log-likelihood \( L^{(E)}_c (c_m), m = 1, \ldots, S/R \), where the subscript stands for “extrinsic” and the superscript denotes that this log-likelihood ratio comes from the equalizer. Further details about the Turbo Equalization iterative detection algorithm can be found in [4].

It is interesting to note that, as it was also the case in [7], if the output of the MAP decoder is extrinsic then non-negligible performance degradation occurs for high order modulations. Thus, it is justifiable to use the entire a-posteriori probability information at the output of the decoder as input to the equalizer.

3. THE CE-SIC EQUALIZER

The CE-SIC, shown in Figure 2, consists of three distinct units, namely, an MMSE Soft Interference Canceller, a Conditional Expectation Computation unit that delivers symbol estimates to the cancellation filter of the SIC, and a Demapper. The Conditional Expectation Computation unit, provides estimates of the transmitted symbols given the a-priori information coming from the decoder and the output of the channel. Based on these estimates the SIC forms an estimate \( s_n \) of the current symbol. Finally, the Demapper, exploits the output of the SIC and the a-priori bit probabilities to compute the corresponding a-posteriori bit probabilities. In the following, we describe each of these units in more detail.

3.1. MMSE Soft Interference Cancellation

The SIC [3], [4] consists of two filters, i.e., the matched filter \( p = [p_{-k} \ldots p_0 \ldots p_l]^T \), \( M = k + l + 1 \) and the cancellation filter \( q = [q_{-K} \ldots q_{-1} 0 \ q_1 \ldots q_N]^T \). The input to the filter \( p \) is the sampled output of the channel at the symbol rate, whereas the input to the cancellation filter consists of past and future symbols. The output \( s_n \) of the SIC is the sum of the outputs of the two filters, i.e.,

\[ s_n = p^T z_n + q^T x_n, \]

where \( z_n = [z_{n+k} \ldots z_n \ldots z_{n-l}]^T \) and \( x_n = [\hat{x}_{n+K} \ldots \hat{x}_n \ldots \hat{x}_{n-N}]^T \), with \( N = l + L_2 \) and \( K = k + L_1 \). If we choose to minimize the mean squared error \( E[|s_n - x_n|^2] \) and assume that the cancellation filter contains correct symbols, then the involved filters are obtained via equations

\[ p = \frac{1}{\sigma_{w_n}^2 + E_h} H_d, \]

and

\[ q = -H_d p + dd^T H_d p \]

where \( E_h = d^T H_d H_d d \) is the square of the channel norm and \( H \) is the \( M \times (K + N + 1) \) channel convolution matrix. Vector \( d \) is defined as \( d = [0_{1 \times K} \ 1 \ 0_{1 \times N}]^T \). From the
above equations it is clear that, the output \( s_n \) of the canceller does not depend on the symbol estimate \( \hat{x}_n \) since the central tap \( q_0 \) of the cancellation filter has been set to zero. At this point, it is convenient to define a function \( \mathcal{T}(v, L, Q) \) which transforms the row vector \( v \) into an \( L \times Q \) Toeplitz matrix as

\[
\mathcal{T}(v_1 v_2 \ldots v_d), L, Q) = \begin{bmatrix}
v_1 & \cdots & v_d & 0 & \cdots & 0 \\
0 & \cdots & v_{d-1} & v_d & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & v_1 & \cdots & v_d \end{bmatrix}
\]

Thus, the convolution matrix \( H \) can be written as \( H = \mathcal{T}(h^T, M, K + N + 1) \) where \( h = [h_{-L} \cdots h_0 \cdots h_{L}]^T \).

### 3.2. Conditional Expectation Computation

After converting the log-likelihood ratios coming from the decoder to bit probabilities, symbol probabilities are computed as products of the corresponding bit probabilities (assumptions bits are independent). The mean and variance of the transmitted symbols are then given by:

\[
\overline{x}_n = E[x_n] = \sum_i \alpha_i P_r\{x_n = \alpha_i\}
\]

and \( \sigma^2_{x_n} = E[[x_n]^2] - E^2[x_n] = 1 - ||\overline{x}_n||^2 \) assuming unit average symbol power \( E [[x_n]^2] \).

The Conditional Expectation Computation unit, sets the input to the cancellation filter \( q \) equal to \( \hat{x}_n = E[x_n|z_n] \) instead of \( \overline{x}_n = E[x_n] \) as proposed in [3], where \( \overline{x}_n \) is computed using only a-priori probabilities. Vector \( z'_n \) is defined as \( z'_n = [z_{n+1} \cdots z_n \cdots z_{n-L}]^T \) and its length is selected so that all elements of \( \hat{x}_n \) use information from a window of at least \( M \) samples of the sequence \( \{z\} \). We may express vector \( z'_n \) in matrix form as

\[
z'_n = H'x'_n + w'_n
\]

where \( x'_n = [x_{n+2K} \cdots x_n \cdots x_{-2N}]^T \), vector \( w'_n \) contains the corresponding noise samples, and \( H' \) is the \((M' = K + M + N) \times (2(K + N) + 1) \) channel convolution matrix defined similarly to matrix \( H \). Thus, Theorem 10.3 of [10], concerning the Bayesian General Linear Model, may be applied assuming that the symbols \( x'_n \) have a prior p.d.f \( N(\overline{x}_n', C_{x_n'}) \).

Thus,

\[
\hat{x}'_n = \mathcal{E}[x'_n|z'_n] = \overline{x}'_n + C_{x_n'} H' (H'C_{x_n}' H' + C_{w'})^{-1} \cdot (z'_n - H' \overline{x}'_n)
\]

where \( \overline{x}'_n = E[x'_n] \), \( C_w = \sigma^2_w I_{M'} \) is the covariance matrix of the noise vector \( w' \) and \( C_{x_n'} = \text{diag}(\sigma^2_{x_{n+2K}} \cdots \sigma^2_{x_n} \cdots \sigma^2_{x_{-2N}}) \) is the diagonal covariance matrix of the symbols based solely on a-priori probabilities. Finally, the required vector \( \hat{x}_n \) is extracted from \( \hat{x}'_n \) by simply keeping only the \( K + N + 1 \) required elements.

Imposing the extrinsic-information constraint that \( \hat{x}'_n \) cannot be a function of a-priori knowledge about \( x_n \), and by keeping only the aforementioned \( K + N + 1 \) required elements we finally get:

\[
s_n = p^H z_n + q^H \hat{x}_n
\]
\[
= p^H z_n + q^H \hat{x}_n + q^H C_{x_n} F_n^e(z'_n - H' \overline{x}'_n + \overline{x}_n H'd')
\]

where

\[
\begin{align*}
F_n^e &= C(F_n^e, K + 1, 2K + 1 + N) \\
denotes a matrix consisting of the "central" \( K + 1 + N \) rows of \( F_n^e = H'H (H'C_{x_n}' H' + \sigma^2_w I)^{-1} \) (from line \( K + 1 \) to \( 2K + 1 + N \)) and the superscript \( (e) \) denotes that the corresponding statistical quantities have been computed setting \( \overline{x}_n = 0 \) and \( \sigma^2_{x_n} = 1 \).

From the above relation it is interesting to note that the suggested solution is, in fact, a classical Soft Interference Canceller (consisting of the first two terms of equation (7)) plus a term to compensate for the fact that, for low a-priori information, the symbol estimates \( \overline{x}_n \) are highly biased estimates of the actually transmitted symbols.

In order to transform the output of the CE-SIC into log-likelihood ratios, the mean and variance of \( s_n \), given that a particular symbol \( \alpha_i \) has been transmitted, must be computed. For these statistics, we get

\[
\mu_{i,n} = E[s_n|x_n = \alpha_i] = (p^H H d + q^H C_{x_n} F_n^e H'd') \cdot \alpha_i
\]

and

\[
\sigma^2_{i,n} = p^H H C_{x_n} C_{x_n} H^H + \sigma^2_w I_M \]
\[
+ 2 \text{Real} \{p^H H C_{x_n} C_{x_n} H^H + w' F_n^{(e)} C_{x_n} q\}
\]
\[
+ q^H C_{x_n} F_n^{(e)} H' (H'C_{x_n}' H' + \sigma^2_w I_M') F_n^{(e)} C_{x_n} q
\]

where

\[
W = [0_{M \times K} \sigma^2_w I_M] \quad 0_{M \times N}
\]

and \( C_{x_n,x_n}^{(e)} \) is the extrinsic covariance matrix between \( x_n \) and \( x_n' \). Also \( d' = [0_{1 \times 2K} \cdots 0_{1 \times 2N}]^T \). Finally, for the Demapping operation, we can use the exact demapper of [7].

### 4. THE ACE-SIC EQUALIZER

Assuming that low a-priori information is present, matrix

\[
F_n^e = H'H (H'C_{x_n}' H'H + \sigma^2_w I)^{-1}
\]

can be approximated by the matrix

\[
F_n^{(e)} = H'H (H'H + \sigma^2_w I)^{-1}
\]

Note that when no a-priori information is available then \( C_{x_n}^{(e)} \) equals the identity matrix. Furthermore, if we inspect the
rows of matrix $\hat{F}^{(e)}$, we can easily verify that each one corresponds to an MMSE linear equalizer of length $M'$, designed for a corresponding output delay. Thus, an approximation of matrix $F_n^{(e)}$ defined in (8) can be:

$$F = T \left\{ d^T \mathbf{H}^H (\mathbf{H} \mathbf{H}^H + \sigma_0^2 \mathbf{I}_M)^{-1}, K + 1 + N, K + M + N \right\}$$

(11)

that is, a single linear equalizer of length $M$ is used instead of $K + 1 + N$ different equalizers of length $M' > M$, for computing the symbol estimates. This approximation is valid when the linear equalizer length $M$ is adequately large, so that two linear equalizers of equal length $M' > M$, designed to give estimates of symbols $x_n$ and $x_{n-1}$ respectively, have equal taps but shifted by $i$ places, and both of them contain at least $M' - M$ zero elements. Of course, increasing the equalizer length $M$ and using a channel convolution matrix $\mathbf{H}$ of larger dimension in (11), makes this approximation more accurate. Thus, since we can design a sufficiently long linear equalizer quite easily, the most crucial approximation is the replacement of $C_{x_n}$ by $\mathbf{I}_{2K+1+2N}$.

The above suggested approximation of matrix $F_n^{(e)}$ by $\hat{F}$ is expected to affect the performance of the ACE-SIC algorithm compared to the performance of its exact counterpart, the CE-SIC. As a remedy to this performance degradation, we allow the (past and future) symbol estimates contained in the cancellation filter to depend on the a-priori information about the current symbol $x_n$. As these estimates are subsequently combined for the computation of the output of the ACE-SIC, it turns out that the extrinsic information restriction has been relaxed. On the other hand, using the a-priori information about $x_n$ improves the computed past and future symbol estimates. This modification yields the following filtering equation:

$$\hat{s}_n = \mathbf{p}^H \mathbf{z}_n + \mathbf{q}^H \mathbf{\bar{x}}_n + \mathbf{q}^H \mathbf{C}_{x_n} \hat{F} (\mathbf{z}_n - \mathbf{H} \mathbf{\bar{x}}_n)$$

(12)

in which the vector multiplying $\hat{F}$ does not include the term $\mathbf{z}_n \mathbf{H}^d \mathbf{d}'$, as opposed to equation (7). It is interesting to observe the interplay between $\mathbf{z}_n$ and the output of the linear equalizer in computing $\hat{x}_n$. For perfect a-priori information, $\sigma_2^2 = 0$ and $\hat{x}_n = \mathbf{z}_n$, whereas for no a-priori information $\sigma_2^2 = 1$ and $\hat{x}_n$ is equal to the output of the linear equalizer. For general a-priori information $\sigma_2^2$ serves as a "weight" between these two estimates.

Similarly to the CE-SIC, in order to transform the output of the algorithm into log-likelihood ratios the mean and variance of the output $\hat{s}_n$ must be estimated. For complexity reasons we assume that the required mean and variance remain fixed during each iteration, that is, they are computed once prior to each iteration. This can be achieved by keeping all symbol variances equal to a constant $\sigma^2$. We suggest using

$$\sigma^2 = \max \{ \sigma_{x_1}^2, \sigma_{x_2}^2, \ldots, \sigma_{x_{2l}/(lq)}^2 \}$$

which is valid whenever all symbol variances are equal, and does not amplify the reliability of initially more reliable symbols. Then, equations (9) and (10) can be used by substituting $F_n^{(e)}$ by $\hat{F}$ and all $\sigma_n^2$ by $\sigma^2$. It should be noted that after the initial equalization, some terms of (9) and (10) can be stored so that the mean and variance of the ACE-SIC output needed at later iterations are easily computed.

At this point it is interesting to note that the low complexity SISO equalizers proposed in [4] and [11], consist in fact of two equalizers, one designed for no a-priori information and the other designed for perfect a-priori information. In such a system, a suitable decision criterion must be used for deciding which of the two equalizers must be employed prior to each iteration. In contrast, the ACE-SIC equalizer is identical to its exact counterpart (CE-SIC) both for perfect a-priori information (i.e. $\sigma_n^2 \rightarrow 1$) and for no a-priori information (i.e. $\sigma_n^2 \rightarrow 0$), where in the latter case we assume approximation (11) is valid.

5. SIMULATION RESULTS

To test the performance of the proposed equalizers we performed some typical experiments. Information bits were generated in bursts of $S = 6144$ bits. Then an R.S.C. code with generator matrix $G(D) = [1 + D^2]$ of rate $R = 1/2$ was applied, and the resulting bits were interleaved using a K-Random interleaver ($K=23$) [12]. The interleaved bits, were mapped to an 8-PSK ($q = 3$) symbol alphabet using Gray code mapping. The 4096 symbols per burst were transmitted over a channel whose impulse response was set either $h_{-1} = 0.407, h_0 = 0.815, h_1 = 0.407$ (channel B of [13]) or $h_{-2} = 0.227, h_{-1} = 0.46, h_0 = 0.688, h_1 = 0.46, h_2 = 0.227$ (channel C of [13]). Figures 3 and 4 demonstrate the performance of various receivers performing turbo equalization for the aforementioned channels. For all simulations, the filter lengths were computed using $k = l = 10$.

In Figure 3, we notice that all equalizers exhibit similar performance. The MMSE equalizer of [4] has superior performance followed by the CE-SIC, the SWITCHED equalizer of [11] and the ACE-SIC. All algorithms after eight iterations attained the performance bound that corresponds to the AWGN channel. Thus, a high complexity algorithm for the channel B, does not seem very practical since the same performance can be obtained by the low complexity solutions.

In Figure 4, we notice that the ISI caused by the channel is quite severe so that none of the examined algorithms attains the performance bound after eight iterations. The MMSE equalizer of [4] and the CE-SIC have almost the same performance. The MMSE I equalizer of [8] attains better performance that the ACE-SIC, however, at a higher computational complexity. It is interesting to note that the ACE-SIC equalizer exhibits better performance than the SWITCHED equalizer of [11] (approximately 1dB less SNR is needed to achieve a BER of $10^{-3}$). Therefore, for hostile channels, switching between equalizers optimized for the two extreme cases (no a-priori and perfect a-priori information) can be a less efficient strategy than using an algorithm that can smoothly adapt to
the quality of the a-priori information (such as the ACE-SIC). Also, the ACE-SIC equalizer, at medium SNRs, achieves a performance close to the performance of its exact counterpart.

![Fig. 3. BER Performance after 8 turbo iterations, channel B](image1)

![Fig. 4. BER Performance after 8 turbo iterations, channel C](image2)

6. CONCLUSIONS

In this work, a novel SISO equalizer of linear complexity was presented. This algorithm was derived as an approximate implementation (ACE-SIC) of a new two step minimization algorithm (CE-SIC) which in turn was developed for the problem of equalization using a-priori probabilities. Simulation results have shown that (a) the exact implementation has almost identical performance to the MMSE equalizer of [8], and (b) the approximate implementation offers very good performance at linear complexity. Thus, the latter low complexity equalization algorithm is suitable for high data-rate wireless communication systems with limited processing power.

7. REFERENCES


