MMSE TURBO EQUALIZER FOR CHANNELS WITH COCHANNEL INTERFERENCE

Y. Gong, Z. Ding and C. F. N. Cowan*

Institute of Electronics, Communications and Information Technology
The Queen's University of Belfast
Queen's Road, Queen's Island Belfast BT3 9DT, UK

ABSTRACT

The MMSE turbo equalizer has excellent performance with moderate complexity. Current realizations of an MMSE turbo equalizer usually assume the equalizer output is Gaussian, which, however, may not be true for channels with co-channel interference (CCI). In this paper, we first show that, in presence of CCI, a linear pre-whitener is required to decorrelate the CCI-plus-noise. Then we apply density estimation for the residual noise at the equalizer output to further improve the performance. Numerical simulations are also shown to verify the analysis.

1. INTRODUCTION

The turbo MMSE equalizer can achieve good performance by iteratively exchanging soft information, or the loglikelihood (LLR), between the equalizer and decoder [1, 2]. A common problem of previous approaches is that they assume the residual noise at the equalizer output is Gaussian, which, however, may not hold in real applications. This is because in many systems, besides ISI and the Gaussian noise, the channel also suffers from co-channel interference (CCI) which, for example, may come from a nearby system operating in the same frequency band or a unknown user in the adjacent cell in mobile communications. The presence of CCI-plus-noise brings two problems to the current SISO MMSE equalizer design: First, since usually the CCI signals are also transmitted through multipaths, the CCI-plus-noise is not white and the corresponding correlation matrix is not diagonal. Thus in general, the full knowledge of the CCI channel, which is unfortunately not available in most cases, is required to obtain the MMSE equalizer. Second, with CCI, the channel noise is normally a Gaussian mixture, as is the residual noise at the equalizer output, resulting in performance loss if the equalizer still retains the Gaussian assumption.

The main task of this paper is, therefore, to propose methods to overcome these problems. We will show that the first problem can be solved by processing the received signal with a linear pre-whitener which, as shown in [3], can perfectly decorrelate the CCI-plus-noise. For the second problem, we may resort to the kernel density function [4] to on-line estimate the probability density of the residual noise. One of the main difficulties is that the residual noise is not directly observable, but, as will be shown later in this paper, can only be constructed with training sequences. Although sometimes training sequences used for channel estimation may also be used for the density estimation, in many cases, the training sequences are not available. A possible alterative is to use the hard decision of the equalizer output which is, however, not reliable as it is highly dependent on the accuracy of the currently detected symbol. In the turbo equalization system, fortunately, the a priori information fed back from the decoder can be used to form *self-generated* training symbols. The use of the a priori information is much more robust than directly handling the equalizer output not only because the decoder always gives lower bit-error-rate (BER) output than the equalizer, but also because the a priori information feedback from the decoder, due to the deinter-/inter-leaver, can be regarded as "independent" from the current detected symbols, preventing *error propagation*.

2. MMSE EQUALIZER WITH PRE-WHITENER

First, we introduce frequently used notation in this paper: $\mathbf{C_{xy}} = \mathrm{E}[\mathbf{x}\mathbf{y}^{\mathrm{H}}] - \mathrm{E}[\mathbf{x}]\mathrm{E}[\mathbf{y}^{\mathrm{H}}]$, $\mathbf{R_{xy}} = \mathrm{E}[\mathbf{x}\mathbf{y}^{\mathrm{H}}]$ and $\bar{\mathbf{x}} = \mathrm{E}[\mathbf{x}]$.

Assuming the length of the equalizer is N_f, the received vector can be expressed as

$$\mathbf{y}(n) = \mathbf{H} \cdot \mathbf{x}(n) + \mathbf{w}(n), \tag{1}$$

where $\mathbf{H}$ is the $N_f \times (N_f + N_h - 1)$ channel matrix, N_h is the channel length, $\mathbf{x}(n)$ is the channel input and $\mathbf{w}(n)$ is the channel CCI-plus noise vector which is again given by $\mathbf{w}(n) = \mathbf{w}_g(n) + \mathbf{w}_c(n)$, where the last two terms represent the Gaussian channel noise and CCI vectors respectively.

Minimizing the cost function of $\mathrm{E}|x(n-\Delta) - (\mathbf{f}^{\mathrm{H}}(n)\mathbf{y}(n) + d(n))|^2$, and removing the influence of the a priori information for the current symbol, gives the SISO MMSE equalizer as [1, 5]:

$$\mathbf{f}(n) = K(n)\mathbf{C}_{\mathbf{yy}}^{-1}(n) \cdot \mathbf{p}(n), \tag{2}$$

where $K(n) = [1 + (1 - \sigma_x^2(n-\Delta)]^{-1}$, Δ is the decision delay, $\mathbf{p}(n) = \sigma_x^2(n-\Delta)\mathbf{h}_\Delta$, $\mathbf{h}_\Delta$ is the $(\Delta+1)$th column of $\mathbf{H}$ and $\sigma_x^2(n) = \mathrm{E}[x^2(n)] - \bar{x}^2(n)$.

Assuming $\mathbf{w}(n)$ has zero mean, we have $\bar{\mathbf{y}}(n) = \mathbf{H}\bar{\mathbf{x}}(n)$, substituting into the definition of $\mathbf{C}_{\mathbf{yy}}(n)$ gives $\mathbf{C}_{\mathbf{yy}}(n) = \mathbf{R}_{\mathbf{yy}}(n) - \mathbf{H}\bar{\mathbf{x}}(n)\bar{\mathbf{x}}^{\mathrm{H}}(n)\mathbf{H}$, where $\mathbf{R}_{\mathbf{yy}}(n) = \mathbf{H}\mathbf{R}_{xx}(n)\mathbf{H}^{\mathrm{H}} + \mathbf{R}_{\mathbf{ww}}(n)$ and

$$\mathbf{R}_{\mathbf{ww}}(n) = \sigma_g^2\mathbf{I} + \sum_{k=1}^{K}\mathbf{H}_{c_k}\mathbf{H}_{c_k}^{\mathrm{H}} \tag{3}$$

where σ_g^2 is the variance of the Gaussian noise and $\mathbf{H}_{c_k}$ is the channel matrix for the kth CCI user. It is clear from (3) that, if at least one of the CCI channels' memory is larger than one, $\mathbf{R}_{\mathbf{ww}}(n)$ is not diagonal, and the assumption of $\mathbf{R}_{\mathbf{ww}}(n) = \sigma_w^2\mathbf{I}$, as was used in previous approaches (e.g. [1]), leads to performance degradation. Thus, the calculation of $\mathbf{C}_{\mathbf{yy}}(n)$ requires full knowledge of not only the transmission but also the CCI channels which, unfortunately, are not available in most cases. Moreover, in the SISO MMSE equalizer, $\mathbf{R}_{xx}(n)$ is a function of the a priori information which normally varies with time, making $\mathbf{R}_{\mathbf{yy}}(n)$ vary with time as well. This implies that the traditional time average can not be used to estimate $\mathbf{R}_{\mathbf{yy}}(n)$.

To solve this problem, a linear pre-whitener can be applied on the channel output before equalization. The equalizer input vector then becomes:

$$\mathbf{y}'(n) = \mathbf{H}' \cdot \mathbf{x}(n) + \mathbf{w}'(n), \tag{4}$$

where $\mathbf{H}' = \mathbf{P}_w\mathbf{H}$, $\mathbf{w}'(n) = \mathbf{P}_w\mathbf{w}(n)$ and $\mathbf{P}_w$ is the pre-whitener matrix. Since it is known [3] that a pre-whitener can decorrelate the CCI-plus-noise, we have

$$\mathbf{R}_{\mathbf{w}'\mathbf{w}'}(n) = \sigma_{w'}^2\mathbf{I}, \tag{5}$$

where $\sigma_{w'}^2 = (\sigma_g^2 + \sigma_c^2)\mathbf{p}_w^{\mathrm{H}}\mathbf{p}_w$, σ_c^2 is the interference power and $\mathbf{p}_w$ the pre-whitener vector. Then the MMSE equalizer can be obtained by regarding $\mathbf{y}'(n)$, $\mathbf{H}'$ and $\mathbf{w}'(n)$ as the equivalent equalizer input, channel matrix and channel noise respectively. Supposing the length of the pre-whitener is N_p, the equivalent channel memory becomes $N_h + N_p - 1$, which means a longer equalizer should be applied.

3. TURBO EQUALIZER WITH DENSITY ESTIMATION

3.1. Density of the Residual Noise

For clarity of exposition, only the BPSK system is considered, but the results can be readily extended to more advanced systems.

When $x(n-\Delta) = \alpha_i$ ($\alpha_i = \pm 1$), the output from the MMSE equalizer can be expressed as

$$z(n) = \mathrm{E}[z(n)|\alpha_i] + z_w(n), \tag{6}$$

where $\mathrm{E}[z(n)|\alpha_i]$ is the mean of $z(n)$ on the condition of $x(n-\Delta) = \alpha_i$, and $z_w(n)$ is the residual noise, which can be expressed as

$$z_w(n) = z_x(n) + z_g(n) + z_c(n), \tag{7}$$

where $z_x(n)$ is the residual ISI, $z_g(n)$ and $z_c(n)$ are the noise and CCI after passing through the pre-whitener and equalizer respectively. Assuming there are K CCI users, we have $z_c(n) = z_{c,1}(n) + \cdots + z_{c,K}(n)$, where $z_{c,k}(n)$ corresponds to the kth CCI signal which is given by

$$z_{c,k}(n) = f(n) \otimes p_w(n) \otimes h_{c_k}(n) \otimes x_{c_k}(n), \tag{8}$$

where $f(n)$, $p_w(n)$, $h_{c_k}(n)$ are the impulse responses for the equalizer, pre-whitener and kth CCI channel respectively, and $x_{c_k}(n)$ is the transmitted signal from the kth CCI user. We assume each CCI user transmits either $+1$ or -1 with binary distribution of $f_{x_{c_k}}(x) = 1/2 \cdot [\delta(x-1) + \delta(x+1)]$.

It is clear from (8) that, $z_{c,k}(n)$ has $2^{N_{z,k}}$ possible values, where $N_{z,k} = N_f + N_p + N_{c_k} - 2$ and N_{c_k} is the kth CCI channel length. Then $z_c(n)$ has 2^{N_z} possible values, where $N_z = N_{z,1} + \cdots + N_{z,K}$. Thus the density of $z_c(n)$ is given by

$$f_{z_c}(x) = \frac{1}{2^{N_z}}\sum_{i=1}^{2^{N_z}}\delta(x - z_c^{(i)}), \tag{9}$$

where $z_c^{(i)}$ is the ith realization of $z_c(n)$. Similarly, for a given $x(n-\Delta)$, the density of $z_x(n)$ can be obtained as

$$f_{z_x}(x) = \frac{1}{2^{N_x-1}}\sum_{i=1}^{2^{N_x-1}}\delta(x - z_x^{(i)}), \tag{10}$$

where $N_x = N_f + N_p + N_h - 2$ and $z_x^{(i)}$ is the ith realization of $z_x(n)$.

We assume $w_g(n)$ is Gaussian, so is $z_g(n)$. Finally, from (7), the convolution of (9), (10) and the Gaussian distribution of $z_g(n)$ gives the density of $z_w(n)$ which is given by

$$\begin{aligned} f_{z_w}(x) &= \frac{1}{2^{N_w}}\sum_{i=1}^{2^{N_z}}\sum_{j=1}^{2^{N_x-1}} g_{z_g}(x - z_c^{(i)} - z_x^{(j)}) \\ &= \frac{1}{2^{N_w}}\sum_{i=1}^{2^{N_w}} g_{z_g}(x - z_w^{(i)}), \end{aligned} \tag{11}$$

where $N_w = N_z + N_x - 1$, $z_w^{(i)}$ is the ith realization of $z_w(n)$ and $g_{z_g}(x)$ is the Gaussian density of $z_g(n)$.

It is clear from (11) that, in general, the residual noise is a Gaussian mixture. But if there is no CCI, the density of the residual noise can be well approximated by a Gaussian distribution because the residual ISI $z_x(n)$ is normally small. With CCI, on the other hand, especially when some of the CCI are particularly strong, the residual noise is obviously a Gaussian mixture.

3.2. Density Estimation

Assuming the CCI-plus-noise has zero mean, we have $\mathrm{E}[z(n)|\alpha_i] = \alpha_i \cdot \mu_z(n)$, where $\mu_z(n) = \mathbf{f}'(n)\mathbf{h}_\Delta$, substituting into (6) gives

$$z_w(n) = z(n) - \alpha_i \cdot \mu_z(n). \tag{12}$$

The density of $z_w(n)$ can be estimated by applying the kernel density function:

$$f_{z_w}(x) = \frac{1}{N}\sum_{i=1}^{N} \Phi(x - z_w(n)), \tag{13}$$

where N is the number of samples. In this paper, the Gaussian kernel is used so that

$$\Phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

where σ can be chosen as (see [4]) $\sigma = \sigma_{z_w} \cdot N^{-1/5}$, and σ_{z_w} is the standard deviation of $z_w(n)$. With (13), the LLR of $z(n)$ can be obtained as

$$\mathrm{LLR}(z(n)) = \ln\frac{\mathrm{P}(z(n)|\alpha_i = +1)}{\mathrm{P}(z(n)|\alpha_i = -1)} = \ln\frac{f_{z_w}(x - \mu_z(n))}{f_{z_w}(x + \mu_z(n))}$$

Without channel coding, and assuming the MMSE equalizer cancels most of the ISI, near optimum decision can be achieved by a one-symbol ML detector, i.e. the decision is $+1$ if $\mathrm{LLR}(z(n)) > 0$ and vice versa. In general, the one-symbol ML decision gives better performance than the hard decision. But if $f_{z_w}(x)$ is symmetric (i.e. $f_{z_w}(-x) = f_{z_w}(x)$), and satisfies

$$f_{z_w}(x - \mu_z(n)) > f_{z_w}(-x - \mu_z(n)), \tag{14}$$

for any $x > 0$, it can be easily verified that the one-symbol ML detector is equivalent to the hard decision with threshold at 0. It is clear from (11) that the symmetric condition is always satisfied, and the condition of (14) is equivalent to

$$\mu_z^2(n) > \max_i{}^2(z_w^{(i)}), \tag{15}$$

where $\mu_z^2(n)$ is in fact the signal power at the equalizer output and $\max^2(z_w^{(i)})$ reflects the output CCI power. Thus (15) holds when the channel SIR $> \beta$dB (and vice versa), where, roughly speaking, β is about equal to 0.

We note that a one symbol ML decision based on the Gaussian assumption is equivalent to the hard decision [6], which means that, when SIR $> \beta$dB, the Gaussian assumption is good enough and estimating the density in a uncoded system is not necessary.

3.3. The Modified Turbo Structure

In turbo equalization, the LLR is iteratively exchanged between the equalizer and decoder, and the decision is not made until satisfactory performance is reached. Under such a scenario, the density estimation of the residual noise becomes significant because it gives better accuracy of the LLR at the equalizer output, which again improves the decoding performance.

It is clear from (12) that training symbols are required to obtain $z_w(n)$, and then $f_{z_w}(x)$. If the training symbol is not available, the hard decision of the equalizer output may be used as an alternative. However, the hard-decision is unliable to apply as it is too strong a feedback which highly depends on the accuracy of the current decision, causing error propagation. In fact, as will shown in an example later, density estimation using the direct hard decision usually gives little performance improvement. In the turbo MMSE equalizer, fortunately, the training symbols can be replaced by the hard decision of the feedback a priori information from the decoder which, as has been explained in the introduction, is much more reliable than directly handling the equalizer output.

With the above analysis, the structure of the modified turbo MMSE equalizer is shown in Fig. 1, where the channel output is first pre-whitened before processed by the equalizer, and the kernel density function then estimates the density of the residual noise, based on which the LLR is generated and, after being interleaved, forwarded to the decoder. We highlight that, at the first iteration when no a priori information is available, the hard decision of the equalizer output is used for the density estimation.

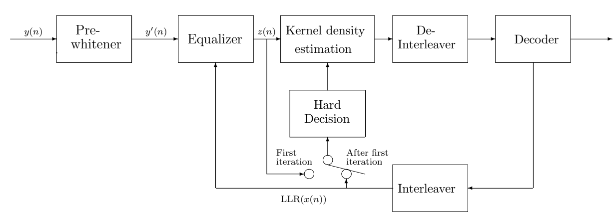

Fig. 1. The turbo MMSE equalizer with density estimation.

4. SIMULATION RESULTS

In this section, numerical simulations are shown to verify the above analysis. First, we consider a system with no channel coding, where the channel vector is $[0.9325\ 0.2798\ 0.1865\ 0.0933\ 0.0933]^{\mathrm{T}}$, there is one CCI user in the system with channel vector of $[0.1925\ 0.9623\ 0.1925]^{\mathrm{T}}$, the channel SNR $= 14$dB, 1024 symbols are transmitted, the lengths of the pre-whitener and equalizer are 2 and 10 respectively, and 100 symbols are used for the density estimation.

Fig. 2 shows the BER curves with respect to the SIR for the hard decision and the one-symbol ML decision based on the kernel density estimation using training-symbols and direct hard-decision respectively. It is clear that, when SIR$<$ 0.5dB, the ML decision using training symbols has significantly better performance than the hard decision, but, as expected, has no advantage over the latter when SIR$>$ 0.5dB since then (14) is satisfied. It is also shown that the ML decision using the hard decision of the equalizer output gives no performance improvement at all.

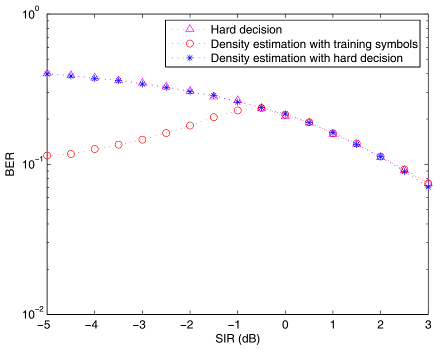

Fig. 2. BER vs SIR for a uncoded system.

For the following experiments, a half rate convolutional code with coding vectors of $[1\ 0\ 1]^{\mathrm{T}}$ and $[1\ 1\ 1]^{\mathrm{T}}$ is added to the system. At the receiver, the modified turbo equalizer shown in Fig. 1 is used to recover the original data. The iteration number is set to be 5.

As an example, Fig. 3 shows, when SIR$= -2$dB, the density of the residual noise obtained by the kernel density function using the training symbols, a priori information and direct hard decision respectively. It is clear that the density obtained by training symbols and a priori information are close, all demonstrate Gaussian mixture. But the density estimation based on the direct hard decision deviates significantly from the other two.

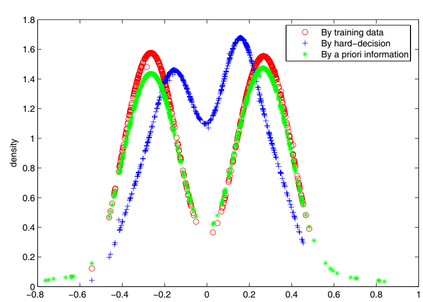

Fig. 3. The density of the residual noise (SIR$= -3$dB)

Fig. 4 shows the BER with respect to the SIR for the decoder output. It is obvious that the performance of the receiver based on the kernel density estimation using the training symbols and a priori information are almost same, and both are better than that based on the white Gaussian assumption.

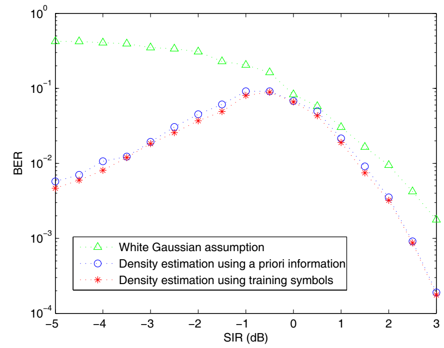

Fig. 4. BER vs. SIR for turbo MMSE equalizer

5. CONCLUSIONS

This paper investigated the turbo MMSE equalizer for channels with CCI. First we showed that a pre-whitener is necessary to decorrelate the CCI-plus-noise, then, after studying the density of the residual noise, we used the kernel density function to estimate its density. Finally numerical simulations were given to verify the analysis.

6. REFERENCES

[1] M. Tuchler, A. C. Singer, and R. Koetter, "Minimum mean squared error equalization using a priori information," *IEEE Trans. on Signal Processing*, vol. 50, no. 3, pp. 673 – 683, March 2002.

[2] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. on communications*, vol. 47, pp. 1046 – 1061, 1999.

[3] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, NJ, 1996.

[4] B. W. Silverman, *Density Estimation for Statistics and Dada Analysis*, London: Chapman Hall, 1986.

[5] H. V. Poor, *An introduction to signal, detection and estimation*, New York: Springer-Verlag, 1994.

[6] J. G. Proakis, *Digital Communications*, McGraw-Hill, Inc, 1995.