Turbo channel estimation and equalization for a superposition-based cooperative system

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: GONG, Y. ... et al, 2009. Turbo channel estimation and equalization for a superposition-based cooperative system. IET Communications, 3 (11), pp. 1790-1799.

Additional Information:

- This paper is a postprint of a paper submitted to and accepted for publication in IET Communications and is subject to Institution of Engineering and Technology Copyright. The copy of record is available at IET Digital Library.

Metadata Record: https://dspace.lboro.ac.uk/2134/25670

Version: Accepted for publication

Publisher: © The Institution of Engineering and Technology

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

Please cite the published version.
Turbo Channel Estimation and Equalization for a Superposition-based Cooperative System

Yu Gong, Zhiguo Ding, T. Ratnarajah and Colin F. N. Cowan

Abstract

Cooperative diversity based on superposition modulation demonstrates high spectral efficiency. A key point of this scheme is for the relay node to reliably estimate the channel and detect the transmission data. Exploring the superposition structure of the transmission data, this paper proposes a novel turbo least-square channel estimator which, without any pilot sequence, can converge to the ideal case as if all the transmission are known to the receiver. The soft-in-soft-out MMSE equalizer is also carefully re-derived to match the superimposed data structure. Finally computer simulation results are shown to verify the proposed algorithm.

Index Terms - cooperative diversity, channel estimation, turbo equalization, superimposed training.

I. INTRODUCTION

Cooperative diversity becomes an attractive alternative to the multiple antenna approach to achieve space diversity in a mobile system [1]. Of particular interests are the schemes using non-orthogonal subspace due to their high spectral efficiency [2]. In this paper, we focus on a particular non-orthogonal decode-and-forward scheme based on superposition modulation described in [3], [4], in which the authors consider a system with two source nodes, A and B, transmitting data in turn to one destination node. At a time slot, the transmission node transmits a superposition of its own data and the data received from the other source node during the previous slot. This scheme can be extended to a general multiple source scenario by grouping the source nodes with two each.

A key point in this scheme is for the relay node to reliably detect the transmission data. In the original protocol [3], the channel is assumed to be flat fading and known, which is, unfortunately, not the case in most scenarios. Channel estimation and equalization are thus necessary. Although pilot symbols are usually

Y. Gong is with the School of Systems Engineering of Reading University, Reading, RG6 6AY, UK. Tel.:+44 0118 378 8581. Email: y.gong@reading.co.uk. Z. Ding is with the Department of Communication Systems of Lancaster University, Lancaster, LA1 4WA, UK. Tel., +44 015245 10399, Email: z.ding@lancaster.ac.uk. T. Ratnarajah and C. Cowan are with ECIT, Queen’s University of Belfast, Belfast, BT3 9DT, UK. Tel: +44 028 9097 1078. Email: T.Ratnarajah@ecit.qub.ac.uk.
required for the channel estimation, they can be saved at the relay node by exploring the superposition structure of the transmission data where part of the data is known to the receiving node. This falls into the general area of the superimposed training (see [5] and the references therein). Many related algorithms have been proposed, most of which attempt to minimize the influence of the information data on the training sequence by exploring some periodic properties of the training data. Such approaches, unfortunately, can not be applied for the superposition-based cooperative system, since we now have little control of the “training sequence” which is in fact the information data of a source node and generally not periodic.

In this paper, we propose a novel turbo least-square (LS) channel estimator by using the a priori information fed back from the decoder to iteratively improve the channel estimation. In the working SNR range, the proposed estimator converges to the ideal estimator that all the transmission data are known. We also re-derive the soft-in-soft-out (SISO) MMSE equalizer described in [6] for this particular superimposed data structure. For clarity of exposition, we assume BPSK modulation in this paper, but the results can be readily extended to other modulation methods as well.

II. SYSTEM MODEL

Without losing generality, we always assume node A transmits and B receives at a time slot $i$. If A successfully decodes the data from B during the time slot $i - 1$, it transmits a packet of $M$ superimposed symbols as:

$$x_{A,i} = \sqrt{1 - \gamma^2} \cdot s_A, i + \gamma \cdot s_B, i - 1,$$

(1)

where $s_A$ and $s_B$ are the data vector for A and B respectively, the subscript $i$ represents the time index, and $0 < \gamma^2 < 1$. Otherwise if A fails to decode $s_{B,i - 1}$, it only transmits its own data packet.

Since this paper mainly considers the relay node, we only show the received signal vector at node B which is given by

$$y_{B,i} = H_i \cdot x_{A,i} + n_i,$$

(2)

where $H_i$ and $n_i$ are the Sylvester channel matrix and noise vector from A to B respectively. For later use, we express $x_{A,i} = [x_{A,i}(n), \cdots, x_{A,i}(n - M + 1)]^T$, and similarly for other vectors whenever necessary.
The diversity gain achieved by such a cooperative protocol can be illustrated as following. According to [4], the density of the sum rate $I$ for the superposition cooperation can be easily shown as

$$q_{SP}(I) = \frac{2e^{2I}}{\rho^2 \lambda_1,0 \lambda_2,0} \int_1^{e^{2I}} f \left( \frac{e^{2I} - \psi_1}{\rho \lambda_1,0 \psi_1} \right) f \left( \frac{\psi_1 - 1}{\rho \lambda_2,0} \right) \frac{1}{\psi_1} d\psi_1,$$

(3)

where $\rho$ denotes SNR and $\lambda_i$ is the constant determined by the coefficient matrix. On the other hand, we have the mutual information for the direct transmission $I_D = \ln \left( 1 + \rho |h_1|^2 \right)$. And it is straightforward to show the density function of $I_D$ as

$$q_1(I_D) = \int_{0}^{\infty} \delta \left( I_D - \ln(1 + \rho x) \right) f_1(x) dx = \frac{e^{I_D}}{\rho} e^{-\left( \frac{e^{I_D} - 1}{\rho} \right)},$$

(4)

where $\delta(\cdot)$ denotes the Delta function. The second equation follows from the property of the Delta function $\delta(f(x)) = \sum_i \frac{\delta(x - x_i)}{|df/dx|_{x_i}}$ where $x_i$ is the $i$th root of $f(x)$. Following similar steps, the density of the mutual information for the selection relay scheme in [1] can be shown as $q_2(I_{SR}) = \frac{2e^{2I_{SR}}}{\rho} \left( \frac{e^{2I_{SR}} - 1}{\rho} \right) e^{-\left( \frac{e^{2I_{SR}} - 1}{\rho} \right)}$.

From the density of the mutual information, we can express the outage probability as as $P_{out} = \int_{0}^{I_{out}} q(I) dI$. Fig. 1 shows the density of the mutual information for the proposed, selection and direct transmission schemes. As can be seen from these figures, the value of the density function for the superimposed protocol is smaller than that for the two comparable schemes for small $I$, which ensures the superimposed scheme to achieve the best reception robustness among the three transmission schemes.
III. Turbo LS Channel Estimator

In this paper, we assume the channel is quasi-static (slow fading), i.e. remains unchanged within one packet. We also assume that, with guarded interval, no inter-packet interference occurs. So the time index \( i \) is dropped whenever no confusion is caused.

For denotation purpose, the input to the channel estimator is expressed as \( c = [c(n), \cdots, c(n - M + 1)]^\top \). The LS channel estimator is then given by

\[
\hat{h} = (CC^H)^{-1} \cdot C y_B,
\]

where

\[
C = [c(n), \cdots, c(n - N_{y_B} + 1)]^\top, c(n) = [c(n), \cdots, c(n - N_L + 1)]^\top, N_{y_B} \text{ and } N_L \text{ are the lengths of } y_B \text{ and the LS estimator respectively, and from (2) we have } N_{y_B} = M - N_h + 1 \text{ and } N_h \text{ is the channel length}.
\]

The simplest method to explore the superimposed data structure for the channel estimation is to regard \( s_B \) as the training sequence and \( s_A \) as the interference, or to let \( c = \gamma \cdot s_B \) in (5). The performance is, obviously, severely limited to the so-called “co-packet interference” from \( \sqrt{1 - \gamma^2} \cdot s_A \).

Since an equalizer is usually required for a selective fading channel, similar to the decision feedback equalizer (DFE), we may feed back the hard decision of the equalizer output, \( \tilde{s}_A \), to the channel estimator to suppress the co-packet interference so that

\[
c = \gamma \cdot s_B + \sqrt{1 - \gamma^2} \cdot \tilde{s}_A.
\]

The channel estimator and equalizer operate in an iterative way. Initially, \( \tilde{s}_A = 0 \) and only \( s_B \) is applied for the channel estimation. The estimated channel is then used by the equalizer whose output, after the hard decision device, is fed back for the next channel estimation. Although ideally such approach converges to the case as if both \( s_B \) and \( s_A \) are known to the LS estimator, it may suffer from serious error propagation when, for example, the channel SNR is low or the co-packet interference is large due to a small \( \gamma \). An ideal

\[1\text{Extra zeros need to be padded to } c \text{ when } N_L > N_h.\]
input to such iterative LS approach has the form of

\[ c = \gamma \cdot s_B + \sqrt{1 - \gamma^2} \cdot f(\hat{s}_A), \]  

(7)

where \( \hat{s}_A \) is an estimate of \( s_A \), \( f(\hat{s}_A) \rightarrow s_A \) when \( \hat{s}_A \) is close to \( s_A \) and \( f(\hat{s}_A) \rightarrow 0 \) as otherwise.

Since the original superposition-based cooperative diversity belongs to the general *decode-and-forward* scheme, a decoder is usually followed after the equalizer. Inspired by the excellent performance of the turbo equalizer, we propose a so-called turbo LS estimator so that

\[ c = \gamma \cdot s_B + \sqrt{1 - \gamma^2} \cdot E[s_A], \]  

(8)

where \( E[s_A(n)] = 1 \cdot P(s_A(n) = 1) + (-1) \cdot P(s_A(n) = -1), \)

\[ P(s_A(n)) = \frac{1 + s_A(n) \cdot \tanh(\text{LLR}(s_A(n))/2)}{2}, \]  

(9)

and \( \text{LLR}(s_A(n)) = \ln[P(s_A(n) = 1)/P(s_A(n) = -1)] \) which is the log-likelihood fed back from the decoder. The overall structure of the turbo channel estimator is illustrated in Fig. 2, where initially \( \text{LLR}_{\text{ex}}(s_A(n)) = 0 \) for all \( n \).

Because only the extrinsic information \( \text{LLR}_{\text{ex}} \) is fed back from the decoder, and also due to the deinter-/inter-leaver, the error propagation can be effectively suppressed. To be specific, when SNR \( \rightarrow \infty \), we have \( \text{LLR}_{\text{ex}}(s_A(n)) \rightarrow \infty \) and \( E[s_A(n)] = s_A(n) \). When SNR \( \rightarrow -\infty \), on the other hand, we have \( \text{LLR}_{\text{ex}}(s_A(n)) \rightarrow 0 \) and \( E[s_A(n)] = 0 \). Therefore, (8) is a good realization of (7).
In most cases, the noise power is also unknown and can be estimated as
\[ \hat{\sigma}^2 = \frac{|\hat{H} \cdot c - y_B|^2}{N_{y_B}}, \] (10)
where \( \hat{H} \) is the estimated channel matrix. It is obvious that (10) depends on not only \( \hat{H} \) but also \( c \). Thus if only \( s_B \) is used for the channel estimation, then even with \( \hat{H} = H \), the noise power estimation is still limited to the co-packet interference from \( s_A \). The turbo channel estimator, on the contrary, can solve this problem well because it has not only a better estimation of \( \hat{H} \), but also a less co-packet interference by including \( E[s_A] \) in \( c \) as shown in (8).

IV. SISO MMSE Equalizer with Superimposed Data

In this paper, we are particularly interested in the linear SISO MMSE equalizer due to its simplicity and nature connection to the turbo structure [6]. After the channel estimation, the known data \( s_B \) must be removed either before or after the equalization, which are, for clarity of exposition, denoted as “pre-cancellation” and “post-cancellation” respectively. Although it looks straightforward, the “pre-cancellation” approach suffers performance loss in SNR. To illustrate this phenomena, we first assume the channel is perfectly known. Then if \( s_B \) is removed before the equalization, the equalizer input is given by
\[ y'_{B} = y_{B} - \gamma H \cdot s_{B} = \sqrt{1 - \gamma^2} H \cdot s_{A} + n, \]
and the equivalent channel SNR becomes
\[ \text{SNR} = \frac{1 - \gamma^2}{\sigma^2}. \] (11)
On the contrary, if the equalizer directly operates on \( y_{B} \) and removes \( s_{B} \) after the equalization, the channel SNR is \( 1/\sigma^2 \). This clearly reveals the SNR loss from the “pre-cancellation” approach, where the exact value of loss depends on \( \gamma \). When the channel is not perfectly known, the analysis is more complicated since the channel estimation error becomes another source of “noise”. However, in a working SNR range, the proposed turbo channel estimator gives very small error and the above conclusion still approximately hold. When the SNR is low, on the other hand, the BER performance deteriorates seriously, making it little different between the “pre-” and “post-cancellation” approaches. Therefore \( s_B \) should always be removed after the equalizer, and it is then necessary to re-derive the SISO MMSE equalizer to match the superimposed data
structure.

The detail of the equalizer is shown in Fig. 3, where \( w(n) \) is the equalizer vector, \( b(n) \) is a DC term, \( \Delta \) is the decision delay, \( y_{sB}(n) = \gamma \hat{H} \cdot s_B \) which corresponds to the \( s_B \) part in \( y_B \) and \( \hat{H} \) is the estimated channel matrix. In particular, \( \hat{x}_A(n-\Delta) \) is the equalizer output, or the estimation of \( x_A(n-\Delta) \), subtracting which by \( w^H(n)y_{sB}(n) \) gives \( \hat{s}_A(n-\Delta) \), the estimation of \( s_A(n-\Delta) \). Finally the LLR generator calculates the extrinsic information, \( \text{LLR}_{ex}(s_A) \), based on the Gaussian assumption.

It is clear from (1) that, for a known \( s_B(n) \), \( x_A(n) \) can only take two values: \( \lambda'_{A1} = \sqrt{1 - \gamma^2 + \gamma \cdot s_B(n)} \) and \( \lambda'_{A0} = -\sqrt{1 - \gamma^2 + \gamma \cdot s_B(n)} \), corresponding to \( s_A(n) = \pm 1 \) respectively. Then we have

\[
\begin{align*}
\bar{x}_A(n) &= \lambda'_{A1} \cdot P(s_A(n) = 1) + \lambda'_{A0} \cdot P(s_A(n) = -1), \\
E[x_A^2(n)] &= \lambda''_{A1} \cdot P(s_A(n) = 1) + \lambda''_{A0} \cdot P(s_A(n) = -1),
\end{align*}
\]

(12)

where \( P(s_A(n)) \) is calculated according to (9) and \( \bar{a} = E[a] \) for any vector \( a \). Then using (12), setting \( \text{LLR}(s_A(n - \Delta)) = 0 \), and with similar procedures as those in [6], we obtain the equalizer tap-vector and output as

\[
w(n) = (1 - \gamma^2) \cdot \{\text{Cov}(y_B(n)) + [(1 - \gamma^2) - \text{Cov}(x_A(n - \Delta))]\hat{H}_\Delta \hat{H}_\Delta^H\}^{-1}\hat{H}_\Delta,
\]

(13)

\[
\hat{x}_A(n - \Delta) = \gamma \cdot s_B(n - \Delta) + w^H(n)[y_B(n) - \bar{y}_B(n) + (\bar{x}_B(n - \Delta) - \gamma \cdot s_B(n - \Delta))\hat{H}_\Delta],
\]

respectively, where \( \hat{H}_\Delta \) is the \( \Delta + 1 \)th column of \( \hat{H} \) and \( \text{Cov}(a) = E[aa^H] + E^2[a] \) for any vector \( a \). Note that \( \text{Cov}(y_B(n)) \) and \( \bar{y}_B(n) \) can be easily further decomposed in term of channel parameters and \( \text{LLR}(s_A) \).

The detail of the derivation is omitted due to the space constraint of this paper.
The mean and covariance of $\hat{s}_A(n - \Delta)$ for a given $s_A(n - \Delta) = S_A$ are obtained as

$$
\mu_{s_A, i}(n - \Delta) = \mathbb{E}[\hat{s}_A(n - \Delta)|s_A(n - \Delta)] = \mathbb{E}[\hat{x}_A(n - \Delta)|s_A(n - \Delta)] - \mathbb{E}[w^Hy_{s_B}(n)] \\
= \gamma s_B(n - \Delta) + \sqrt{1 - \gamma^2}S_Aw^H(\hat{H}_\Delta - \gamma w^H(n)\hat{H}s_B(n)}
$$

(14)

$$
\sigma^2_{s_A}(n - \Delta) = \text{Cov}[\hat{s}_A(n - \Delta)|s_A(n - \Delta)] = (1 - \gamma^2)w^H\hat{H}_\Delta[1 - \hat{H}_\Delta^Hw(n)],
$$

where $\mu_{s_A, i}$ corresponds to $S_A = \pm 1$ for $i = 1, 0$ respectively. Note that the covariance of $\hat{x}_A(n - \Delta)$ and $\hat{s}_A(n - \Delta)$ are the same. Finally, with (14) and the Gaussian assumption, we obtain $\text{LLR}_{ex}(s_A)$.

V. Numerical Simulations

In this section, the channel vector is set as $h = [0.1 0.3 1 0.3 0.1]^T$, a half rate convolutional code with coding vectors of $[1 0 1]^T$ and $[1 1 1]^T$ is used to encode the information data of $s_A(n)$, each packet contains 128 symbols, the length of the channel estimator and equalizer are given by 5 and 10 respectively. We consider four cases, i.e. only $s_B$ is used for the channel estimation, the proposed turbo LS channel estimator is applied, both $s_B$ and $s_A$ are used for the channel estimation, and the channel is perfectly known, which are denoted as “LS-$s_B$”, “LS-turbo”, “LS-both” and “Known-channel” respectively. For fair comparison, the turbo equalization is applied for all cases and the iteration number is set as 5. All the results below are obtained by averaging over 5,000 independent runs.

In the first example, we set $\gamma^2 = 0.2$, and let $s_B$ be removed after the equalizer (i.e. “post-cancellation”). Fig. 4(a) and (b) shows the mean-squared-error (MSE) of the channel tap estimation which is defined as $\text{MSE}(\hat{h}) = \mathbb{E}[\|\hat{h} - h\|^2]/\|h\|^2$, and the BER performance respectively. It is clearly shown in Fig. 4(a) that, in the working SNR range (e.g. SNR > 5dB), the proposed turbo LS estimator converges to the case as if both $s_B$ and $s_A$ are known to the receiver. On the other hand, when SNR is low (e.g. SNR < 2dB), the error propagation can be effectively suppressed, since then the turbo LS estimator works like a traditional LS estimator. Fig. 4(b) clearly shows that the BER performance with the turbo LS estimator is almost identical to that of the ideal case where the channel is perfectly known, and is significantly better than that with the “LS-$s_B$” approach, where, for example, about 3dB improvement in SNR can be observed at

3The MSE of the noise power estimation is similar to Fig. 4(a), but is not shown here due to the space constraint.
BER = $10^{-5}$.

In the second example, we compare the performance between the approaches of “pre-cancellation” and “post-cancellation” for $s_B$. For a better exposition, we particularly set $\gamma^2 = 0.45$, because, according to (11), the larger the $\gamma^2$ is, the bigger the difference between the two approaches appears. Fig. 5(a) and (b) show the equalizer output SNR and the BER performances for both approaches respectively. The equalizer output SNR is obtained as $E[\mu_{s_A, i}(n)]/E[\sigma_{s_A}^2(n)]$, where $\mu_{s_A, i}(n)$ and $\sigma_{s_A}^2(n)$ are given by (14).

Fig. 5(a) clearly reveals the SNR advantage of the “post-cancellation” over the “pre-cancellation” approach. It is clear from Fig. 5(b) that, with “post-cancellation”, the best BER performance, which is achieved when the channel is known, is almost identical to that by applying the turbo LS estimator. On the other
hand, though it looks straightforward, the approach of “LS-\textit{s}_B” with “pre-cancellation” gives the worst BER performance. There are about 3dB difference in SNR at BER = $10^{-5}$ between the best and worst cases. It is interesting to observe that the performance for “LS-\textit{s}_B” with “post-cancellation” is close to that for “LS-turbo” with “pre-cancellation”, because the performance loss suffered by the two cases are due to the neglect of $\sqrt{1-\gamma^2 s_A}$ at the channel estimator and the neglect of $\gamma s_B$ at the equalizer respectively. But with $\gamma^2 = 0.45$, the powers of $\sqrt{1-\gamma^2 s_A}$ and $\gamma s_B$ are almost the same. This observation indicates that the information of $s_B$ and $s_A$ should be used as much as possible by the channel estimator and equalizer, which is in fact the philosophy behind the proposed approach of this paper.

VI. CONCLUSION

In this paper, we have studied the problem how to reliably achieve channel estimation and symbol detection at cooperative relays. Exploring the superposition structure of the transmission data, this paper proposes a novel turbo least-square channel estimator which, without any pilot sequence, can converge to the ideal case as if all the transmission are known to the receiver. The soft-in-soft-out MMSE equalizer is also carefully re-derived to match the superimposed data structure. Finally computer simulation results are shown to verify the proposed algorithm.

REFERENCES


