14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

A Novel Channel Estimator for a Superposition-based Cooperative System

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

Institute of Electronics, Communications and Information Technology
Queen's University of Belfast
Northern Ireland Science Park, Queen's Road, Belfast BT3 9DT, UK
email: y.gong,c.f.n.cowan@ee.qub.ac.uk, z.ding,T.Ratnarajah@ecit.qub.ac.uk

ABSTRACT

This paper investigates channel estimation and equalization for the relay node of a cooperative diversity system based on superposition modulation. Exploring the superposition structure of the transmission data, we propose a novel channel estimator for the relay node. Without any pilot sequence, the proposed estimator can converge to the ideal case as if the transmission data is known to the receiver. We also re-derive the soft-input-soft-output MMSE equalizer to match the superimposed data structure of the cooperative scheme. Finally computer simulation results are given to verify the proposed algorithm.

1. INTRODUCTION

Among various cooperative diversity schemes those that use non-orthogonal subspaces are of particular interest due to their high spectral efficiency [1, 2]. This paper focuses on a particular non-orthogonal decode-and-forward scheme based on superposition modulation described in [3], where a system with two source nodes, A and B, is considered, each transmitting data in turn to one destination node, and the transmission node transmits a superposition of its own data and the data received from the other source node during the previous slot.

A key point of 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 not the case in most scenarios. Channel estimation and equalization are thus necessary. Although pilot symbols are usually 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 [4] 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 as if all the transmission data are known. We also re-derive the soft-input-soft-output (SISO) MMSE equalizer [5] 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.

2. 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:

$$\mathbf{x}_{A,\ i} = \sqrt{1-\gamma^2}\cdot\mathbf{s}_{A,\ i} + \gamma\cdot\mathbf{s}_{B,\ i-1}, \tag{1}$$

where $\mathbf{s}_A$ and $\mathbf{s}_B$ 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 $\mathbf{s}_{B,\ i-1}$, 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

$$\mathbf{y}_{B,\ i} = \mathbf{H}_i\cdot\mathbf{x}_{A,\ i} + \mathbf{n}_i, \tag{2}$$

where $\mathbf{H}_i$ and $\mathbf{n}_i$ are the Sylvester channel matrix and noise vector from A to B respectively. For later use, we express $\mathbf{x}_{A,\ i} = [x_{A,\ i}(n),\ldots,x_{A,\ i}(n-M+1)]^{\mathrm{T}}$, and similarly for other vectors whenever necessary.

It is clear from (1) that $x_{A,\ i}(n)$ is an PAM signal with four constellation points. Because, for reliable detection, the constellation points should be separated as far as possible [6], the optimum γ must let the four constellation points be equally apart. Based on this observation, some simple derivation immediately leads to

$$\gamma^2_{\mathrm{opt}} = 0.2. \tag{3}$$

This result matches well with a statement in [3] that $0.075 \leqslant \gamma^2 \leqslant 0.2$. However such optimum γ only considers the general symbol detection. When the whole cooperative system is considered, the choice of γ becomes much more complicated since it also depends on the condition of the relay channels. The detail of this topic is beyond the scope of this paper.

3. 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 notation purposes, the channel estimation input is expressed as $\mathbf{c} = [c(n),\ldots,c(n-M+1)]^{\mathrm{T}}$. The LS channel estimator is then given by

$$\hat{\mathbf{h}} = (\mathbf{C}\mathbf{C}^{\mathrm{H}})^{-1}\cdot\mathbf{C}\,\mathbf{y}_B, \tag{4}$$

where $\mathbf{C} = [\mathbf{c}(n), \ldots, \mathbf{c}(n - N_{y_B} + 1)]^T$, $\mathbf{c}(n) = [c(n), \ldots, c(n - N_L + 1)]^T$, N_{y_B} and N_L are the lengths of $\mathbf{y}_B$ and the LS estimator respectively, and from (2) we have $N_{y_B} = M - N_h + 1$ and N_h is the channel length [1]. The simplest method to explore the superimposed data structure for the channel estimation is to regard $\mathbf{s}_B$ as the training sequence and $\mathbf{s}_A$ as the interference, or to let $\mathbf{c} = \gamma \cdot \mathbf{s}_B$ in (4). The performance is, obviously, severely limited by the so-called "co-packet interference" from $\sqrt{1-\gamma^2}\,\mathbf{s}_A$.

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, $\tilde{\mathbf{s}}_A$, to the channel estimator to suppress the co-packet interference so that

$$\mathbf{c} = \gamma \cdot \mathbf{s}_B + \sqrt{1-\gamma^2} \cdot \tilde{\mathbf{s}}_A. \tag{5}$$

The channel estimator and equalizer operate in an iterative way. Initially, $\tilde{\mathbf{s}}_A = \mathbf{0}$ and only $\mathbf{s}_B$ 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 an approach converges to the case as if both $\mathbf{s}_B$ and $\mathbf{s}_A$ 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 γ. An ideal input to such an iterative LS approach has the form of

$$\mathbf{c} = \gamma \cdot \mathbf{s}_B + \sqrt{1-\gamma^2} \cdot f(\hat{\mathbf{s}}_A), \tag{6}$$

where $\hat{\mathbf{s}}_A$ is an estimate of $\mathbf{s}_A$, $f(\hat{\mathbf{s}}_A) \to \mathbf{s}_A$ when $\hat{\mathbf{s}}_A$ is close to $\mathbf{s}_A$ and $f(\hat{\mathbf{s}}_A) \to 0$ 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

$$\mathbf{c} = \gamma \cdot \mathbf{s}_B + \sqrt{1-\gamma^2} \cdot \mathrm{E}[\mathbf{s}_A], \tag{7}$$

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

$$\mathrm{P}(s_A(n)) = \frac{1 + s_A(n) \cdot \tanh(\mathrm{LLR}(s_A(n))/2)}{2}, \tag{8}$$

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

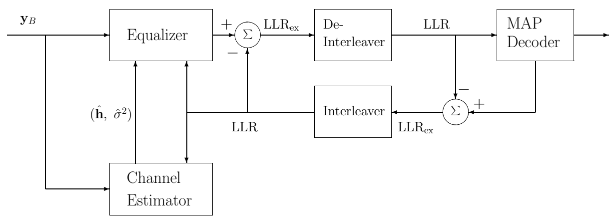

Figure 1: The turbo LS channel estimator.

Because only the extrinsic information LLR_{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 $\to \infty$, we have $\mathrm{LLR}_{ex}(s_A(n)) \to \infty$ and $\mathrm{E}[s_A(n)] = s_A(n)$. When SNR $\to -\infty$, on the other hand, we have $\mathrm{LLR}_{ex}(s_A(n)) \to 0$ and $\mathrm{E}[s_A(n)] = 0$. Therefore, (7) is a good realization of (6).

In most cases, the noise power is also unknown and can be estimated as

$$\hat{\sigma}^2 = \frac{|\hat{\mathbf{H}} \cdot \mathbf{c} - \mathbf{y}_B|^2}{N_{y_B}}, \tag{9}$$

where $\hat{\mathbf{H}}$ is the estimated channel matrix. It is obvious that (9) depends on not only $\hat{\mathbf{H}}$ but also $\mathbf{c}$. Thus if only $\mathbf{s}_B$ is used for the channel estimation, then even with $\hat{\mathbf{H}} = \mathbf{H}$, the noise power estimation is still limited to the co-packet interference from $\mathbf{s}_A$. The turbo channel estimator, however, can solve this problem well because it has not only a better estimation of $\hat{\mathbf{H}}$, but also less co-packet interference by including $\mathrm{E}[\mathbf{s}_A]$ in $\mathbf{c}$ as shown in (7).

4. 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 [5]. After the channel estimation, the known data $\mathbf{s}_B$ 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 $\mathbf{s}_B$ is removed before the equalization, the equalizer input is given by $\mathbf{y}' = \mathbf{y}_B - \gamma \mathbf{H} \cdot \mathbf{s}_B = \sqrt{1-\gamma^2}\mathbf{H} \cdot \mathbf{s}_A + \mathbf{n}$, and the equivalent channel SNR becomes

$$\mathrm{SNR} = \frac{1-\gamma^2}{\sigma^2}. \tag{10}$$

To the contrary, if the equalizer directly operates on $\mathbf{y}_B$ and removes $\mathbf{s}_B$ 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 γ. 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 holds. When the SNR is low, on the other hand, the BER performance deteriorates seriously, making little difference between the "pre- and post- cancellation" approaches. Therefore $\mathbf{s}_B$ 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 Figure 2, where $\mathbf{w}(n)$ is the equalizer vector, $b(n)$ is a DC term, Δ is the decision delay, $\mathbf{y}_{s_B}(n) = \gamma \hat{\mathbf{H}} \cdot \mathbf{s}_B$ which corresponds to the $\mathbf{s}_B$ part in $\mathbf{y}_B$ and $\hat{\mathbf{H}}$ 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 $\hat{x}_A(n-\Delta)$ by $\mathbf{w}^H(n)\mathbf{y}_{s_B}(n)$ gives $\hat{s}_A(n-\Delta)$, the estimation of $s_A(n-\Delta)$. Finally the LLR generator calculates the extrinsic information, $\mathrm{LLR}_{ex}(\mathbf{s}_A)$, based on the Gaussian assumption.

[1] Extra zeros need to be padded to $\mathbf{c}$ when $N_L > N_h$.

14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

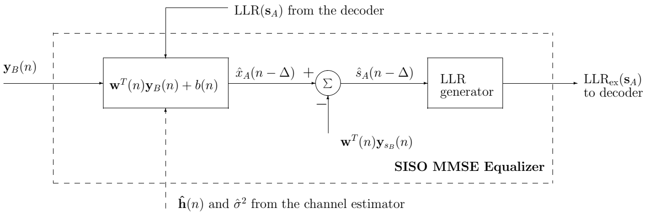

Figure 2: The SISO MMSE equalizer with superimposed data.

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

$$\begin{aligned} \bar{x}_A(n) &= \mathscr{X}_{A1} \cdot \mathrm{P}(s_A(n)=1) + \mathscr{X}_{A0} \cdot \mathrm{P}(s_A(n)=-1) \\ \mathrm{E}[x_A^2(n)] &= \mathscr{X}_{A1}^2 \cdot \mathrm{P}(s_A(n)=1) + \mathscr{X}_{A0}^2 \cdot \mathrm{P}(s_A(n)=-1), \end{aligned} \tag{11}$$

where $\mathrm{P}(s_A(n))$ is calculated according to (8) and $\bar{\mathbf{a}} = \mathrm{E}[\mathbf{a}]$ for any vector $\mathbf{a}$. Then using (11), setting $\mathrm{LLR}(s_A(n-\Delta)) = 0$, and with similar procedures as those in [5], we obtain the equalizer tap-vector and output as [2]

$$\begin{aligned} \mathbf{w}(n) =& (1-\gamma^2) \cdot \{\mathrm{Cov}(\mathbf{y}_B(n)) + \\ & [(1-\gamma^2) - \mathrm{Cov}(x_A(n-\Delta))]\hat{\mathbf{H}}_\Delta \hat{\mathbf{H}}_\Delta^{\mathrm{H}}\}^{-1}\hat{\mathbf{H}}_\Delta, \\ \hat{x}_A(n-\Delta) =& \gamma \cdot s_B(n-\Delta) + \mathbf{w}^{\mathrm{H}}(n)[\mathbf{y}_B(n) - \bar{\mathbf{y}}_B(n) + \\ & (\bar{x}_B(n-\Delta) - \gamma \cdot s_B(n-\Delta))\hat{\mathbf{H}}_\Delta], \end{aligned} \tag{12}$$

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

The mean and covariance of $\hat{s}_A(n-\Delta)$ for a given $s_A(n-\Delta) = \mathscr{S}_A$ are obtained as

$$\begin{aligned} \mu_{s_A,\,i}(n-\Delta) &= \mathrm{E}[\hat{s}_A(n-\Delta)|s_A(n-\Delta)] \\ &= \mathrm{E}[\hat{x}_A(n-\Delta)|s_A(n-\Delta)] - \mathrm{E}[\mathbf{w}^{\mathrm{H}}\mathbf{y}_{s_B}(n)] \\ &= \gamma s_B(n-\Delta) + \sqrt{1-\gamma^2}\mathscr{S}_A \mathbf{w}^{\mathrm{H}}(n)\hat{\mathbf{H}}_\Delta - \gamma \mathbf{w}^{\mathrm{H}}(n)\hat{\mathbf{H}}\mathbf{s}_B(n) \\ \sigma_{s_A}^2(n-\Delta) &= \mathrm{Cov}[\hat{s}_A(n-\Delta)|s_A(n-\Delta)] \\ &= (1-\gamma^2)\mathbf{w}^{\mathrm{H}}\hat{\mathbf{H}}_\Delta[1 - \hat{\mathbf{H}}_\Delta^{\mathrm{H}}\mathbf{w}(n)], \end{aligned} \tag{13}$$

where $\mu_{s_A,\,i}$ $(i=1,0)$ corresponds to $\mathscr{S}_A = \pm 1$ respectively. Note that the covariance of $\hat{x}_A(n-\Delta)$ and $\hat{s}_A(n-\Delta)$ are the same. Finally, with (13) and the Gaussian assumption, we obtain $\mathrm{LLR}_{\mathrm{ex}}(\mathbf{s}_A)$.

5. NUMERICAL SIMULATIONS

In this section, the channel vector is set as $\mathbf{h} = [0.1\ 0.3\ 1\ 0.3\ 0.1]^{\mathrm{T}}$, a half rate convolutional code with coding vectors of $[1\ 0\ 1]^{\mathrm{T}}$ and $[1\ 1\ 1]^{\mathrm{T}}$ 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 $\mathbf{s}_B$ is used for the channel estimation, the proposed turbo LS channel estimator is applied, both $\mathbf{s}_B$ and $\mathbf{s}_A$ are used for the channel estimation, and the channel is perfectly known, which are denoted as "LS-$\mathbf{s}_B$", "LS-turbo", "LS-both" and "Known-channel" respectively. For all of the cases, the turbo equalization is applied and the iteration number is set as 5. All the results below are obtained by averaging over 5,000 independent runs.

At first, we set $\gamma^2 = 0.2$ as was suggested in Section 2, and let $\mathbf{s}_B$ be removed after the equalizer. Figure 3 shows the mean-squared-error (MSE) of the channel tap estimation which is defined as $\mathrm{MSE}(\hat{\mathbf{h}}) = \mathrm{E}|\hat{\mathbf{h}} - \mathbf{h}|^2/|\mathbf{h}|^2$. It is clearly shown that in the working SNR range (e.g. SNR > 5dB), the proposed turbo LS estimator converges to the case as if both $\mathbf{s}_B$ and $\mathbf{s}_A$ 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 [3].

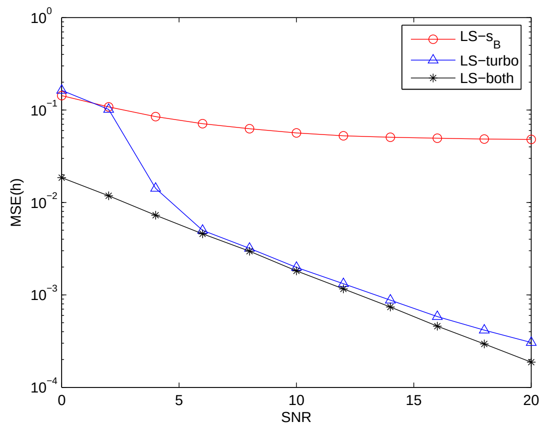

Figure 3: The Channel MSE for $\gamma^2 = 0.2$

It is clearly shown in Figure 4 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-$\mathbf{s}_B$' approach, where, for example, about 3dB improvement in SNR can be observed at BER $= 10^{-5}$.

The second example compares the performance between the approaches of "pre-cancellation" and "post-cancellation" for $\mathbf{s}_B$. For a better exposition, we particularly set $\gamma^2 = 0.45$, because, according to (10), the larger the γ^2 is, the bigger the difference between the two approaches appears. Figure 5 shows the equalizer output SNR which is obtained as $\mathrm{E}[\mu_{s_A,\,i}^2(n)]/\mathrm{E}[\sigma_{s_A}^2(n)]$, where $\mu_{s_A,\,i}(n)$ and $\sigma_{s_A}^2(n)$ are given by (13).The SNR advantage of the "post-cancellation" over the "pre-cancellation" approach can be easily observed.

Figure 6 shows the corresponding BER performances. It is clear 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-$\mathbf{s}_B$" with "pre-cancellation" gives the worst

[2]The detail of the derivation is omitted due to the space constraint of this paper.

[3]The MSE of the noise power estimation is similar to Figure 3, but is not shown here due to the space constraint.

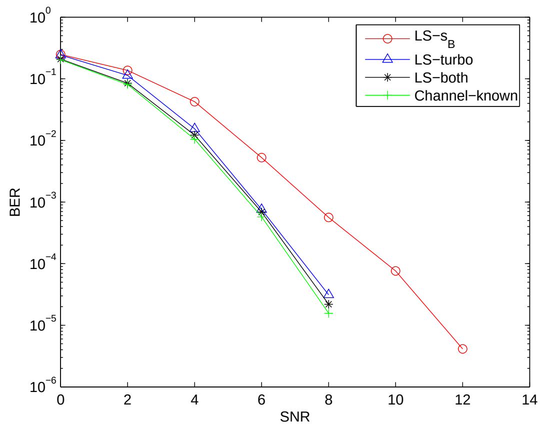

Figure 4: The BER performance for $\gamma^2 = 0.2$

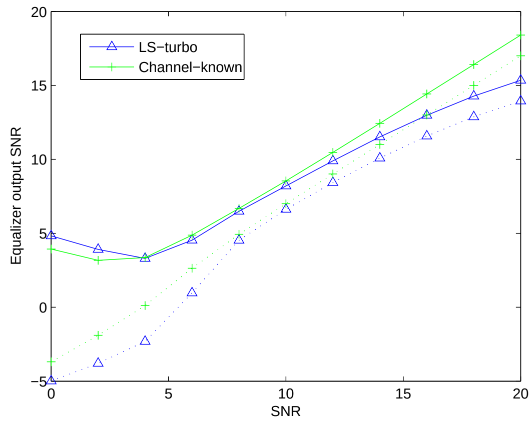

Figure 5: The equalizer output SNR for $\gamma^2 = 0.45$. Solid lines: "pre-cancellation"; Dash lines: "post-cancellation".

BER performance. It is interesting to observe that the performance for "LS-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 $\sqrt{1-\gamma^2}\mathbf{s}_A$ at the channel estimator and the neglect of $\gamma\mathbf{s}_B$ at the equalizer respectively. But with $\gamma^2 = 0.45$, the powers of $\sqrt{1-\gamma^2}\mathbf{s}_A$ and $\gamma\mathbf{s}_B$ are almost the same. This observation indicates that the information of $\mathbf{s}_B$ and $\mathbf{s}_A$ 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.

6. CONCLUSION

This paper proposed a novel turbo LS channel estimator for the relay node of a cooperative system based on superposition modulation. Without any pilot symbols, the proposed estimator converges to the ideal case as if both $\mathbf{s}_A$ and $\mathbf{s}_B$ are known to the receiver. Simulation results successfully verify the proposed algorithm. But we point out here that such a scheme is not totally "blind". At the very beginning of the transmission or at anytime that a node fails to decode the other node's data, the "superposition" structure is ruined and pilot symbols may be necessary to resume the process. Since such events rarely occur, over all, we improve the spectral efficiency by exploring the superimposed structure.

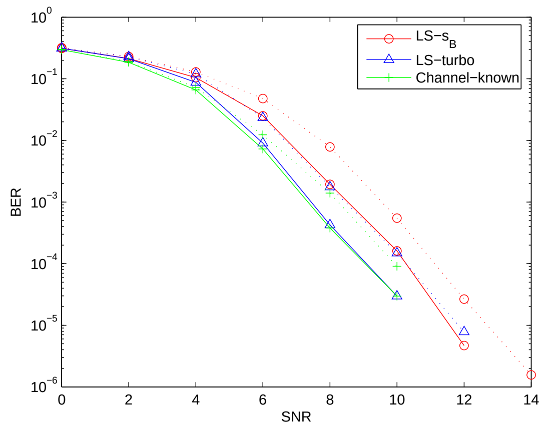

Figure 6: The BER performance for $\gamma^2 = 0.45$. Solid lines: "pre-cancellation"; Dash lines: "post-cancellation".

REFERENCES

[1] J. N. Laneman, D. N. C. Tse, and G. W. Wornell. Cooperative diversity in wireless networks: efficient protocals and outage behavior. *IEEE Trans. on Information Theory*, 50(12):3062 – 3080, Dec. 2004.

[2] K. Azarian, H. El Gamal, and P. Schniter. On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels. *IEEE Trans. on Information Theory*, 51(12):4152 – 4172, Dec. 2005.

[3] E. G. Larsson and B. R. Vojcic. Cooperative transmit diversity based on superposition modulation. *IEEE Communications Letters*, 9(9):778 – 780, Sept. 2005.

[4] M. Ghogho, D. McLernon, E. A.-Hernandez, and A. Swami. Channel estimation and symbol detection for block transmission using data-dependent superimposed training. *IEEE Signal Processing letters*, 12(3):226 – 229, March 2005.

[5] M. Tuchler, A. C. Singer, and R. Koetter. Minimum mean squared error equalization using a priori information. *IEEE Trans. on Signal Processing*, 50(3):673 – 683, March 2002.

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