Robustness of Space-Time Turbo Codes

Wei Shi, Christos Komninakis, Richard D. Wesel, and Babak Daneshrad
University of California, Los Angeles
Los Angeles, CA 90095-1594

Abstract—**In this paper, we consider the performance of a turbo code with 2 transmit antennas and 2 receive antennas for flat fading channels. We consider robustness within and among families of channels with the same singular values (SV's). When the two SV's are similar, bit error rate performance is robust in terms of proximity to the channel capacity. When the 2 SV's differ significantly, robust performance can be achieved by a time-varying linear transformation (TVLT). Iterative decoding is the same as for standard parallel concatenated convolutional codes, except that the extrinsic information about encoder outputs are also exchanged.**

I. Introduction

Recently, the goal of providing high data rate wireless communications has motivated the use of multiple transmit and receive antennas. Foschini and Gans [1] showed the extraordinarily large capacity promised by using multiple transceiver antennas for the flat Rayleigh fading. This opened the new area of space-time codes, which take advantage of spatial diversity with multiple transceiver antennas and temporal diversity provided by channel coding and interleaving.

Tarokh *et al.* [2], [3] and Guey *et al.* [4] proposed the design criteria for space-time codes for flat Rayleigh fading in a variety of scenarios. To achieve a high level of diversity, turbo codes [5] can be used to get excellent performance in fading channels with multiple transceiver antennas. In [6], Liu *et al.* proposed a rank criterion of full space diversity QPSK space-time turbo codes. In [7], [8], Stefanov and Duman showed that a simple QPSK turbo code outperforms the space-time block and trellis codes [2], [3] significantly over (ergodic) block fading channels.

This paper shows the robustness of turbo codes in frequency-flat quasistatic fading with 2 transmit antennas and 2 receive antennas. This robustness is measured in terms of consistent bit error rate (BER) for a specified proximity to channel capacity. We consider robustness for families of channels with the same set of singular values (SV's). We have found that the robustness is determined by the difference of the 2 SV's. When the SV's are similar, a straightforward turbo decoding approach provides robust performance. When the SV's differ significantly, consistent performance may still be obtained, but requires a time-varying linear transformation (TVLT) in both the transmitter and the receiver. A similar concept to the TVLT technique presented in this paper is phase sweeping by A. Hiroike *et al.* [9].

Section II describes the channel and system models, the space-time turbo codes, and the decoding techniques. Section III presents the linear transformation in the receiver side and the time-varying linear transformation at both transmitter and receiver. Section IV shows the simulation results. Section V concludes the paper.

II. Channel and System Models

A. Space-time channels and theoretical limits

As in [1], we assume that the fading is fixed for a transmission block, but varies between blocks. Let the 2×2 complex matrix H represent the channel matrix for a given transmission block, where h_{ij} is the path gain from transmit antenna j to receive antenna i with $i, j \in \{1, 2\}$. The input-output equation for the multiple-input multiple-output (MIMO) channel is

$$\mathbf{y} = H \cdot \mathbf{x} + \mathbf{n},$$

or explicitly

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}, \qquad (1)$$

where x_j is the transmitted signal at the transmit antenna j, which is a two-dimensional (2-D) constellation point with energy $\bar{P}_x$ per dimension, $\bar{P}_x = P_x/2$. y_i is the received signal at receive antenna i. Noise n_i is i.i.d. additive white Gaussian noise (AWGN) with variance $N_0/2$ per dimension.

When the transmitter knows nothing about the channel and hence transmits x_1 and x_2 with the same energy, the maximum mutual information (MI) is

$$I(\mathbf{x}; \mathbf{y}) = \log_2\left(1 + \frac{\lambda_1^2 P_x}{N_0}\right) + \log_2\left(1 + \frac{\lambda_2^2 P_x}{N_0}\right), \qquad (2)$$

where λ_1 and λ_2 are the singular values of the channel matrix H. As defined in [1], the signal-to-noise ratio (SNR) is the average SNR at each receiver branch. In terms of the singular values, $SNR = \frac{(\lambda_1^2+\lambda_2^2)P_x}{2N_0}$.

Suppose the transmitter performs a linear transformation on the transmitted signal x_1 and x_2 by an 2×2 unitary matrix Q_t and the receiver performs linear transformation on the received signal y_1 and y_2 by an 2×2 unitary matrix Q_r. Then we have the following

$$Q_r \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = Q_r \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} Q_t \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + Q_r \begin{bmatrix} n_1 \\ n_2 \end{bmatrix},$$

where the effective channel $\tilde{H} = Q_r H Q_t$ has the same singular values as H, and the same MI by (2). Let the channel matrix H have the singular value decomposition (SVD) as

0-7803-7097-1/01/$10.00 ©2001 IEEE

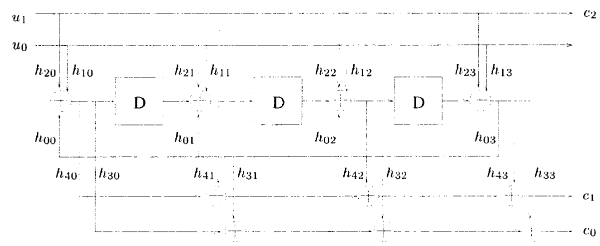

Fig. 1. Encoder structure with 3 memory elements

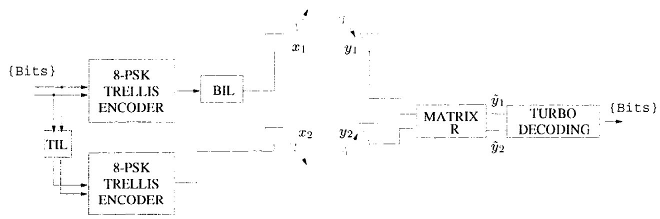

Fig. 2. Space-time model with turbo-coding.

$H = ASB^*$ with the singular values λ_1 and λ_2, then $\tilde{H}$ has an SVD as $\tilde{H} = \tilde{A}S\tilde{B}^*$ with the same singular values λ_1 and λ_2, since both $\tilde{A} = Q_rA$ and $\tilde{B} = Q_r^*B$ are unitary matrices. The noise $\begin{bmatrix} \tilde{n}_1 \\ \tilde{n}_2 \end{bmatrix} = Q_r \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}$ is still i.i.d. AWGN with variance $N_0/2$ per dimension. So if the original channel is in the family of channels with singular values λ_1 and λ_2, the channel $\tilde{H} = Q_rHQ_t$ is still in the family. The MI remains the same when such linear transformation techniques are applied later in this paper.

B. Space-time turbo codes

The constituent encoders we use in the symbol-interleaved turbo-trellis code of Figure 2 are depicted in Figure 1. They are characterized by the octal polynomial $(h_0, h_1, h_2, h_3, h_4) = (13, 1, 17, 13, 17)$, the meaning of which is described in Figure 1. These constituent encoders are optimal for an AWGN channel by the criteria proposed in [10], [11]. For the upper encoder the input bit u_0 is punctured. For the lower encoder the input bit u_1 is punctured.

The turbo interleaver (TIL) of Figure 2 is a 4098-symbol randomly generated interleaver with 2 input bits per symbol. The input bits are flipped in each interleaved symbol, as shown in Figure 2, which avoids repetition of systematic bit. A 4098-symbol block-interleaver (BIL) is used at the output of the upper encoder. x_1 is the transmitted signal in the upper transmit antenna which is the block-interleaved version of the output of the upper encoder. x_2 is the transmitted signal in the lower transmit antenna, which is the output of the lower encoder.

At the receiver, y_1 and y_2 are the linear combinations of x_1 and x_2 distorted by AWGN noise received by the two receive antennas. After a linear transformation R as shown in Figure 2 at the receiver, which will be discussed in the next section, the received signals y_1 and y_2 become $\tilde{y}_1$ and $\tilde{y}_2$. So the inputs to the turbo decoder are $\tilde{y}_1$ and $\tilde{y}_2$.

Figure 3 shows the decoding diagram. The only channel input to the upper decoder is the block-deinterleaved version of $\tilde{y}_1$. Similarly $\tilde{y}_2$ is the only channel input to the lower decoder. The channel input $\tilde{y}_i$ not considered by one decoder is accounted for through the extrinsic information on x_i provided by the other decoder.

Both upper and lower decoders use the forward-backward algorithm to get the reliabilities of input and the outputs of corresponding encoders. They exchange the reliabilities for the input in the same way as in standard parallel concatenated convolutional codes. Each decoder also sends the reliabilities about the outputs of the corresponding encoders to the other decoder. The upper and lower decoders are connected by interleavers, which is standard in soft-input soft-output (SISO) networks [12].

The crucial part of the forward-backward algorithm for the lower forward-backward iterative decoding is

$$\begin{aligned}
&\gamma_k(m', m) = Pr(\tilde{y}_{2k}, S_k = m | S_{k-1} = m') \\
&= Pr(S_k = m | S_{k-1} = m') \cdot Pr(\tilde{y}_{2k} | S_{k-1} = m', S_k = m) \\
&= Pr(u_k; I) \cdot Pr(\tilde{n}_{2k} = \tilde{y}_{2k} - \tilde{h}_{21}\, x_{1k} - \tilde{h}_{22}\, x_{2k} | x_{2k}) \\
&= Pr(u_k; I) \cdot \sum_{x_{1k}} Pr(x_{1k}; I) \\
&\cdot Pr(\tilde{n}_{2k} = \tilde{y}_{2k} - \tilde{h}_{21}\, x_{1k} - \tilde{h}_{22}\, x_{2k} \,|\, x_{1k},\, x_{2k}).
\end{aligned}$$

We can get a similar expression for the upper decoder. With the extrinsic information about x_1 or x_2, the upper and lower decoders can update its $\gamma_k(m', m)$ and calculate the reliabilities about the input and the output of the corresponding encoder.

Assume that the upper decoder is doing the decoding first in each iteration. In the first iteration, the extrinsic information $P(u; I)$ about the input will be zero (all u's are equally likely), when the logarithm is taken. The soft information $P(x_2; I)$ is needed by the upper encoder. It is possible to overcome

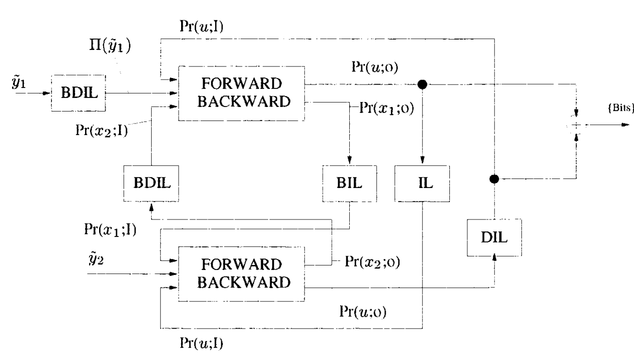

Fig. 3. Space-time iterative decoding.

the need for soft information $P(x_2; I)$ on the first iteration by using an upper/lower-triangular decomposition (ULTD).

Let $U = \begin{bmatrix} 1 & u_{12} \\ 0 & 1 \end{bmatrix}$ be the upper triangular matrix and $L = \begin{bmatrix} l_{11} & 0 \\ l_{21} & l_{22} \end{bmatrix}$ be the lower triangular matrix such that $\tilde{H} = UL$. U and L can be uniquely determined for a given $\tilde{H}$ and it can be shown that $l_{21} = \tilde{h}_{21}, l_{22} = \tilde{h}_{22}$. The inverse of matrix U is simply $U^{-1} = \begin{bmatrix} 1 & -u_{12} \\ 0 & 1 \end{bmatrix}$. Doing linear transformation on $\tilde{y}_1$ and $\tilde{y}_2$ by matrix U^{-1}, we get $\hat{y}_1$ and $\hat{y}_2$ as

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 \\ \tilde{h}_{21} & \tilde{h}_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} \hat{n}_1 \\ \hat{n}_2 \end{bmatrix} \quad (3)$$

where $\begin{bmatrix} \hat{n}_1 \\ \hat{n}_2 \end{bmatrix} = U^{-1} \begin{bmatrix} \tilde{n}_1 \\ \tilde{n}_2 \end{bmatrix}$. Clearly $\hat{y}_1 = l_{11}x_1 + \hat{n}_1$, so there is no dependency in x_2. $\hat{y}_2 = \tilde{y}_2$ by (3), so the channel input to the lower decoder is $\tilde{y}_2$ for all the iterations even with the ULTD performed in the first iteration. However the ULTD in the first iteration might introduce potential noise enhancement for $\hat{n}_1$.

III. Signal processing techniques

A. Linear transformation at the receiver side

From an MI perspective (2), all channels with the same SV's are equivalent. However diagonal channels perform best in our turbo decoding simulations because there is no interference from the other transmit antenna. Normally, the channel matrix is not diagonal in the space-time scenario; there is interference between the transmitted signals.

When the receiver has the knowledge of channel H, the receiver can perform linear operations on y_1 and y_2 to separate the transmitted signals. In this way, the receiver tries to create an effective channel $\tilde{H}$ that has the same mutual information but less interference. The linear transformation can be expressed as

$$\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \quad (4)$$

Assuming $H = ASB^*$ and $R = BA^*$, the effective channel matrix $\tilde{H} = RH = BSB^*$ is hermitian and positive-definite. If $\lambda_1 = \lambda_2$, $\tilde{H}$ is diagonal. This technique improves performance by mitigating variation in A in all cases and B when $\lambda_1 = \lambda_2$. For example,

Let

$$B_1 = \begin{bmatrix} \cos\theta e^{-j\theta_1} & \sin\theta \\ \sin\theta & -\cos\theta e^{j\theta_1} \end{bmatrix}, \quad (5)$$

then

$$B_1SB_1^* = \begin{bmatrix} \lambda_1\cos^2\theta + \lambda_2\sin^2\theta & (\lambda_1 - \lambda_2)\cos\theta\sin\theta e^{-j\theta_1} \\ (\lambda_1 - \lambda_2)\cos\theta\sin\theta e^{j\theta_1} & \lambda_1\sin^2\theta + \lambda_2\cos^2\theta \end{bmatrix}. \quad (6)$$

Another unitary matrix B_2 can be represented as

$$B_2 = \begin{bmatrix} \cos\theta e^{-j\theta_1} & -\sin\theta \\ \sin\theta & \cos\theta e^{j\theta_1} \end{bmatrix},$$

and $B_2SB_2^* = B_1SB_1^*$. Together, B_1 and B_2 represent all possible 2×2 unitary matrices.

We observe that smaller singular value difference $|\lambda_1 - \lambda_2|$ implies smaller off-diagonal (interference) elements in $\tilde{H} = B_1SB_1^*$. If $\lambda_1 = \lambda_2 = \lambda$, then $\tilde{H} = B_1SB_1^* = \lambda I_2$ regardless of θ. So, using the unitary transformation R produces the same interference-free (diagonal) effective channel $\tilde{H} = \lambda I_2$ for all channels with two identical singular values.

B. Time-varying linear transformations

When λ_1 and λ_2 are far apart, the channels with interference perform noticeably worse in our simulations than the diagonal channels despite the use of R. This can be observed from expression (6). For large singular value spread, the values in the off-diagonal elements depend on the specific value θ or the unitary matrix B. For channels with different θ, the performance can be noticeably different.

To achieve robustness (i.e. similar performance within a family of channels sharing the pair λ_1, λ_2), a time-varying but channel independent linear transformation at both the transmitter and the receiver can be introduced to make the performance on the family of channels consistent. This technique in a sense transforms the unitary matrix B in (5) to a time-varying unitary matrix T_kB, where T_k is a unitary matrix at the transmitter and is periodically varying for every channel transmission. Let

$$T_k = \begin{bmatrix} \cos\psi_i e^{-j\gamma_j} & \sin\psi_i \\ \sin\psi_i & -\cos\psi_i e^{j\gamma_j} \end{bmatrix}, \quad (7)$$

ψ_i is a phase-sweeping term taking values between 0 and π, i.e. $\psi_i = \frac{\pi}{p_1}i, i = 1\cdots, p_1$ and $\gamma_j = \frac{\pi}{p_2}j, j = 0\cdots, p_2 - 1$, where p_1 and p_2 can be any positive integer numbers. The index k is related with i and j as $k = (j-1)p_1 + i$.

Introducing a time-varying phase θ at both the transmitter and receiver (by T_k) is similar to the phase sweeping technique used in [9] to force fast fading at the receiver. To the extent that T_k cycles through a set of matrices that finely samples the space of possible B's, performance will be consistent throughout the class. We have found that the worst B's benefit substantially from this time variation. The best B's suffer a performance loss, but not as much as the gain for the worst B's.

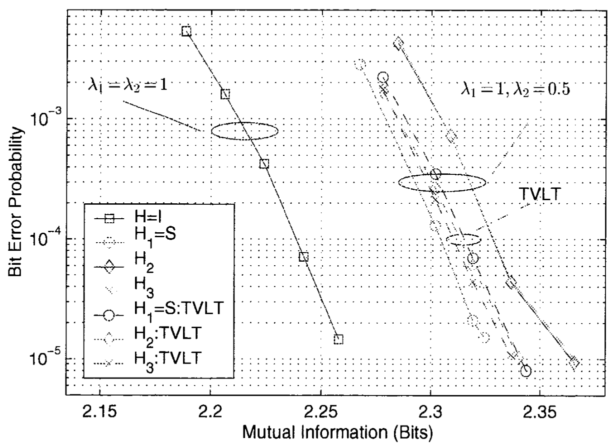

Fig. 4. Bit error rate vs. mutual information for two families of channels with different sets of singular values. The notation H_i:TVLT represents the time-varying linear transformation for channel H_i used at both the transmitter and receiver.

IV. Simulation

A. Proximity to Capacity

To have an idea about how closely the space-time turbo codes perform to the MI expressed in (2), we consider the plot of BER versus MI in plots of simulation results using Monte Carlo simulation with 15 iterations. All the simulations are terminated when 80 block errors occur. The standard plots of BER versus SNR make a fair performance comparison between different channels difficult so we use MI (2) instead of SNR, as in [13].

Figure 4 shows the BER versus MI for two families of channels. The $H = I$ curve represents all channels with singular values $\lambda_1 = \lambda_2 = 1$. The other six curves represent 3 channels all with $\lambda_1 = 1, \lambda_2 = 0.5$ without and with TVLT. With the assumption that the receiver has perfect knowledge of channel, all different channels with $\lambda_1 = \lambda_2 = 1$ have the same effective channel $\tilde{H} = I$, which is the identity matrix. At BER=10^{-5}, this class of channels needs 2.26 bits per 2 antennas. Thus, this family of channels requires an excess MI of 0.13 bits per antenna. At BER=10^{-5}, all 3 channels of $\lambda_1 = 1, \lambda_2 = 0.5$ without using TVLT need 2.33-2.37 bits per 2 antennas; they require an excess MI of 0.165-0.185 bits per antenna.

By using TVLT at both the transmitter and the receiver, the channels with singular values $\lambda_1 = \lambda_2 = 1$ still have the effective channel $\tilde{H} = I$ and perform the same as without using TVLT at both the transmitter and the receiver. However, for the channels with singular values with $\lambda_1 = 1, \lambda_2 = 0.5$, the TVLT technique makes the worst channels H_2 and H_3 perform much better and the best channel $H_1 = S$ perform a little bit worse. By using TVLT, the 3 channels perform much closer and need approximately 2.34 bits per 2 antennas for BER=10^{-5}; they require an excess MI of approximately 0.17 bits per antenna;

The equivalent AWGN SNR as described in [13] is used here to get a more familiar term. For a given MI, the equivalent AWGN SNR is the SNR which yields this MI for the AWGN channel capacity. For $H = I$, the 1.13 bits per antenna corresponds to equivalent AWGN SNR of 0.81 dB. For the other 3 channels, the 1.165-1.185 bits per antenna correspond to equivalent AWGN SNR of 0.94 dB-1.05 dB. Using TVLT, 1.17 bits per antenna corresponds to equivalent AWGN SNR of 0.97 dB. For comparison, the turbo-codes for transmission of 1 bit per channel use in the AWGN channel require an equivalent SNR of 0.7 dB-1.0 dB (or 0.121-0.176 bits of excess MI) for BER=10^{-5}. Thus, there might still be some room on the performance of space-time turbo-codes demonstrated here.

B. Examples

Here, we take a closer look at the 3 channels with singular values $\lambda_1 = 1, \lambda_2 = 0.5$ in Figure 4. Let $S = \begin{bmatrix} 1.0 & 0.0 \\ 0.0 & 0.5 \end{bmatrix}$. Channel matrix $H_1 = S$ is a diagonal matrix. $H_2 = A_2 S B_2^*$, where $B_2 = \begin{bmatrix} 0.6594 & -0.7518 \\ 0.7518 & 0.6594 \end{bmatrix}$ with $\theta = 49^o$, and A_2 can be any unitary matrix in real or complex domain. So the channel matrix H_2 represents a family of matrices which have the same effective channel matrix $\tilde{H}_2 = B_2 S B_2^*$ after the linear transformation at the receiver. Similarly $H_3 = A_3 S B_3^*$, where $B_3 = \begin{bmatrix} 0.7518 & -0.6594 \\ 0.6594 & 0.7518 \end{bmatrix}$ with $\theta = 41.25^o$. Similar to H_2, channel matrix H_3 represents a family of channel matrices which have the same effective channel matrix $\tilde{H}_3 = B_3 S B_3^*$.

As described in subsection IV-A, at BER=10^{-5}, the diagonal channel H_1 requires 2.33 bits of MI per 2 antennas using (2) while the other 2 channels H_2 and H_3 require approximately 2.37 bits per 2 antennas; the gap between bits the diagonal channel H_1 requires and the other 2 channels require is 0.04 bits per 2 antennas. The 3 channels require an excess MI of 0.165-0.185 bits per antenna. By using TVLT, the diagonal channel H_1 requires 2.335 bits per 2 antennas and the other 2 channels H_2 and H_3 require approximately 2.341 bits per 2 antennas; the gap between bits the diagonal channel H_1 requires and the other 2 channels require is 0.006 bits per 2 antennas. The 3 channels require an excess MI of 0.167-0.171 bits per antenna by using TVLT.

In Figure 5, we consider 3 channels, all with singular values $\lambda_1 = 1, \lambda_2 = 0.25$. Thus $S = \begin{bmatrix} 1.0 & 0.0 \\ 0.0 & 0.25 \end{bmatrix}$. The first is $H_1 = S = diag\{1.0, 0.25\}$. $H_2 = A_2 S B_2^*$, where $B_2 = \begin{bmatrix} 0.9239 & 0.3827 \\ 0.3827 & -0.9239 \end{bmatrix}$ with $\theta = 22.5^o$ from (6). $H_3 = A_3 S B_3^*$, where B_3 is the same as in IV-B, and the matrices A_2 and A_3 are irrelevant, as before.

At BER=2×10^{-4}, the diagonal matrix H_1 requires 2.46 bits per 2 antennas. Channels H_2 and H_3 require 2.52 bits and 2.65

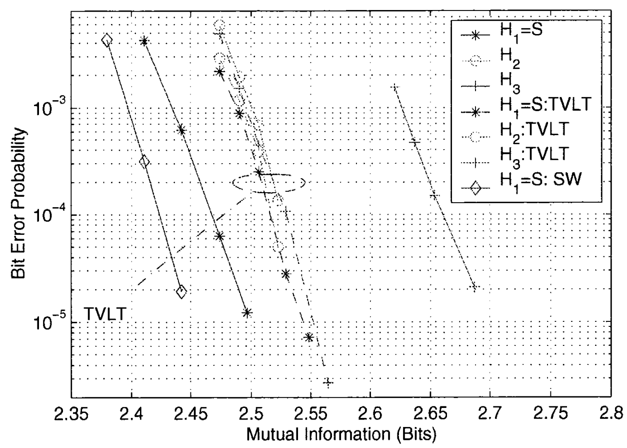

Fig. 5. Bit error rate vs. mutual information for the family of channels with λ_1 =1, λ_2=0.25. The notation H_i:TVLT represents the time-varying linear transformation for channel H_i used at both the transmitter and receiver. The notation $H_1 = S$:SW represents the transmitted signals x_1 and x_2 alternated for each antenna at the transmitter for the diagonal channel $H_1 = S$.

bits per 2 antennas respectively; the gap between bits the diagonal channel H_1 requires and the channel H_3 requires is 0.19 bits. The 3 channels need an excess MI of 0.23-0.325 bits per antenna. Also Figure 5 shows the performance improvement by alternating between the transmitted signals x_1 and x_2 for each antenna when channel matrix is $H_1 = S$.

In all cases studied, the time-varying linear transformation technique is used with parameters $p_1 = 12, p_2 = 1$. In the simulation, we only consider the linear transformation in the real domain, i.e. $\gamma_0 = 0$ as shown in (7). Using this technique, diagonal channel H_1 performs slightly worse. channel H_2 performs better, but the improvement is negligible. Only channel H_3 performs much better. The performances of the 3 channels are much more consistent after the TVLT technique. The performance loss suffered by the diagonal channel H_1 is much less than the performance gain by the channel H_3. However the performance gain by H_2 is negligible. At BER=2×10^{-4}, the diagonal matrix H_1 requires almost 2.5 bits per 2 antennas, while channels H_2 and H_3 require 2.51 bits per 2 antennas; the gap between the bits required by the 3 channels is 0.01 bits. The 3 channels consistently need an excess MI of 0.25-0.255 bits per antenna.

C. Rayleigh fading channels

Figures 6 shows the Block error rate vs. SNR in quasistatic Rayleigh fading. Each entry of the 2×2 channel matrix H is a zero mean, unit power complex Gaussian, i.i.d. from block to block but constant for each entire transmission block of 4098 symbols. Figure 6 plots the outage capacity along with four simulation curves for Block error rate, with or without linear transformation $R = BA^*$ at the receiver side and with or without time-varying linear transformation (TVLT) (7) at both the transmitter and receiver. Notice that the use of R improves performance significantly, by about 5 dB, while the TVLT has no visible effect on Block error rate. One possible explanation for that is that error performance in ergodic of Rayleigh fading is the average performance in good channels and bad channels. Figures 4 and 5 show that the TVLT improves performance in bad channels and deteriorates the Block error rate in good channels. If the quasistatic Rayleigh model produces a balanced mix of good and bad block channels, the limited effect of the TVLT on Rayleigh performance is justified.

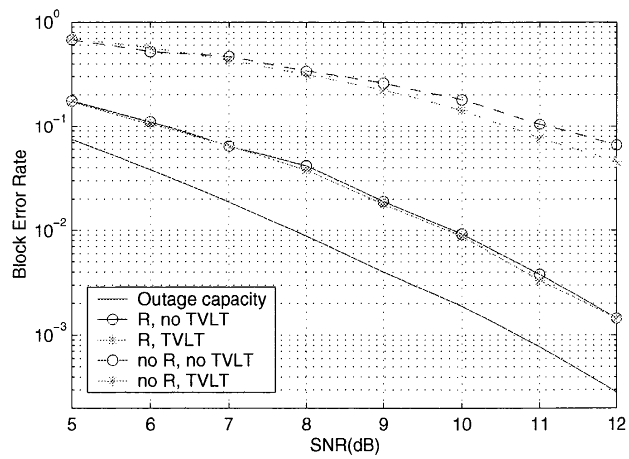

Fig. 6. Block error rate vs. SNR for quasistatic Rayleigh fading channels, with and without the linear transformation R at the receiver, and with or without the TVLT at both the transmitter and receiver.

We can see that the performance is 2.0 dB away from the capacity when the linear transformation R is used, which is equivalent to the 64-state space-time codes in [2].

D. Discussions

In the examples, we show that the performance is dictated by the set of singular values. The channel transformation and space-time iterative decoding techniques used in the receiver make the performance robust. When the values of λ_1 and λ_2 are close to each other, the performance in the family of channels is almost the same as the diagonal channel which can achieve turbo-coding performance similar to that in the AWGN channels.

When λ_1 and λ_2 are significantly different, there are two differences from the previous case. First the diagonal channels need more additional mutual information to get the desired BER. We saw in one example how alternating between the transmitted signals x_1 and x_2 for each antenna can improve performance in this case. Second there is some noticeable gap between the performance of the diagonal channels and the channels with the worst interference. In this case, a time-varying symbol-wise linear transformation technique provides consistent performance in the family of channels. We observed that the performance loss suffered by the good channels is much less than the performance gain obtained by the worst channels. However the performance gain by using TVLT for channels as H_2 in Figure 5 is negligible.

The performance of Rayleigh fading channels is 2.0 dB away from outage capacity. The time-varying linear transformation provides no advantage in this case, but the receiver transformation R does.

V. CONCLUSION

This paper shows one way to achieve robust performance in MIMO frequency-flat quasistatic fading channels with 2 transmit and 2 receive antennas, using turbo codes and linear receiver processing. The BER vs. mutual information (MI) performance is determined by the 2 singular values (SV's). When the 2 SV's are close to each other, the performance is consistent in the family of channels with the same SV's regardless of the unitary matrices A and B, where the channel matrix has singular value decomposition (SVD) in the form $H = ASB^*$. In this case all the channels in the family perform very close to each other and to channel capacity.

When the 2 SV's differ significantly, there is a noticeable gap between the diagonal channel $H = S$ and the worst channel with the most off-diagonal interference between transmit antennas. Also the channels in the family need more additional mutual information to get the desired BER and therefore perform farther from the channel capacity.

However, by using the time-varying linear transformation technique (TVLT), the channels in the family can achieve a kind of consistent performance in the sense that they perform much closer to each other, and the worst channels perform closer to the channel capacity. With this TVLT we observed a slight performance loss for the best channels, negligible performance gain for channels with moderate interference and much more performance gain for the bad channels. This TVLT has the same effect even for channels with singular values close to each other.

We also observed that the linear transformation R at the receiver can improve the performance significantly as shown in Rayleigh fading channels. The TVLT technique helps some worst channels but hurts other good channels. So the overall effect of this technique is negligible in Rayleigh fading channels.

VI. ACKNOWLEDGMENT

The authors would like to thank Christine Fragouli for contributing the symbol-interleaved turbo-trellis codes she found. She also gave invaluable comments and discussions regarding this work.

REFERENCES

[1] G. J. Foschini and M. J. Gans. On limits of Wireless Communications in a Fading Environment When Using Multiple Antennas. *Wireless Personal Communications*, 6:311–335, March 1998.

[2] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time Codes for High Data Rate Wireless Communications: Performance Criterion and Code Construction. *IEEE Trans. on Info. Theory*, 44(2):744–765, March 1998.

[3] V. Tarokh, H. Jafarkhani, and A. R. Calderbank. Space-time Block Codes for High data Rate Wireless Communications: Performance Results. *IEEE J. Select. Areas Commun.*, pages 451–460, March 1999.

[4] J. -C. Guey, M. P. Fitz, M. R. Bell, and W. -Y. Kuo. Signal Design for Transmitter Diversity Wireless Communication Systems Over Rayleigh Fading Channels. In *1996 IEEE Vehicular Technology Conference*, pages 136–140, Atlanta, GA, 1996.

[5] C. Berrou, A. Glavieux, and P. Thitimajshima. Near Shannon Limit Error-Correcting Coding and Decoding: Turbo Codes (1). In *ICC '93*, pages 1064–1070, 1993.

[6] Y. Liu, M. P. Fitz, and O. Y. Takeshita. QPSK Space-Time Turbo codes. In *ICC '00*, June 2000.

[7] A. Stefanov and T. M. Duman. Turbo Coded Modulation for Wireless Communications with Antenna Diversity. In *VTC-Fall*, 1999.

[8] A. Stefanov and T. M. Duman. Turbo Coded Modulation for Systems with Transmit and Receive Antenna Diversity. In *Globecom '99*, Dec. 1999.

[9] A. Hiroike, F. Adachi, and N. Nakajima. Combined Effects of Phase Sweeping Transmitter Diversity and Channel Coding. *IEEE Trans. on Vehicular Technology*, 41(2):170–176, May 1992.

[10] C. Fragouli and R. D. Wesel. Turbo Encoder Design for Symbol Interleaved Parallel Concatenated Trellis Coded Modulation. *IEEE Trans. on Comm.*, 2001.

[11] C. Fragouli and R. D. Wesel. Symbol Interleaved Parallel Concatenated Trellis Coded Modulation. In *ICC '99*, pages 42–46, 1999.

[12] S. Benedetto , D. Divsalar, G. Montorsi, and F. Pollara. Soft-Input Soft-Output Modules for the Construction and Distributed Iterative Decoding of Code Networks. *European Transactions on Telecommunications*, 9(2):155–172, March-April 1998.

[13] R. D. Wesel, X. Liu, and W. Shi. Trellis Codes for Periodic Erasures. *IEEE Trans. on Comm.*, 48(6):938–947, June 2000.