



JOINT BLIND AND SEMI-BLIND DETECTION AND CHANNEL ESTIMATION FOR SPACE-TIME TRELLIS CODED SYSTEMS

Robert J. Piechocki, Christophe Andrieu, Andrew R. Nix and Joe P. McGeehan

University of Bristol, Centre for Communications Research
Woodland Road, MVB, BS8 1UB, Bristol, UK.

E-mail: r.j.piechocki@bristol.ac.uk, Tel. +44 117 9545203, Fax. +44 117 9545206

ABSTRACT

This paper considers a Multiple-Input Multiple-Output (MIMO) communication system, which uses Space-Time Trellis Coding (STTC). A novel method of decoding STTC without a need to transmit training sequences is developed. The technique uses only a single channel estimate to acquire a complete set of the channels' estimates while performing STTC detection. The method is akin to blind trellis search techniques (per-survivor processing - PSP) and adaptive Viterbi. Our solution consists of the deployment of a bank of Kalman Filters. The bank of Kalman Filters is coupled with Viterbi type decoders, which produce tentative decisions based on Kalman channel predictions. In return, the Kalman filters use the tentative decisions to update and track the MIMO channels corresponding to a number of tracked hypotheses. The proposed technique is particularly applicable to space-time systems operating in rapidly fading environments, where STTC can be decoded and MIMO channels efficiently tracked without relying on periodic pilot sequences.

1. INTRODUCTION

Until recently considerable effort was put into designing systems so as to mitigate the perceived detrimental effects of multipath propagation. However, recent work [1] has shown that by utilising multiple antenna architectures at both the transmitter and receiver, so-called multiple-input multiple-output (MIMO) architectures, vastly increased channel capacities are possible. The ideas behind space-time trellis coded modulation (STTCM) were first presented in [2]. The maximum likelihood detection of STTC requires provision of the channel state information (CSI). Typically the CSI is acquired via training sequences. The resulting CSI estimates are then fed to a space-frequency Viterbi decoder, which performs an MLSE search.

Kalman filter tracking of Space-time block coded system has been first reported in [3]. An attempt has also been made to jointly estimate and decode space-time trellis codes [4]. In [4] the authors propagate *posterior* distribution for the decoded symbols. This results in a complex nonlinear problem that can only be tackled by particle filters (sequential importance sampling). However, this also introduces a phase ambiguity in this application. Our approach is computationally less demanding and avoids the phase ambiguity problems even in blind version. The approach is based on per-survivor processing (PSP) (blind trellis search techniques) proposed in [5] and [6], to cope with the problem of unknown or fast changing channels. The PSP techniques are based on adaptive Viterbi and

The authors wish to thank Toshiba TREL Ltd. for sponsoring the work presented in this article. Christophe Andrieu is with the Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK.

LMS detection, and were developed in the context of blind MLSE equalisation. However, PSP is unsuitable in MIMO case.

2. SPACE-TIME CODING

To describe STTC we assume a MIMO system with N_T transmit and N_R receive antennas. The information symbols $\{\mathbf{s}_t; 1 \leq t \leq T\}$ are encoded to a codeword ($N_T \times T$ matrix): $[\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_T]$. Each element in \mathbf{c}_t is transmitted from a separate antenna simultaneously. Signals arriving at the receive antennas undergo independent fading. The received signal $\{\mathbf{y}_t; 1 \leq t \leq T\}$ is a linear mixture with the coefficients given by the channel \mathbf{H}_t , embedded in additive noise:

$$\mathbf{y}_t = \mathbf{H}_t \mathbf{c}_t + \mathbf{n}_t \quad (1)$$

In an AWGN channel, the maximum likelihood decoder can be realised using the Viterbi algorithm with the Euclidean metric:

$$\tilde{\mathbf{c}} = \arg \min_{\tilde{\mathbf{c}} = \tilde{\mathbf{c}}_1 \dots \tilde{\mathbf{c}}_T} \sum_{t=1}^T \left\| \mathbf{y}_t - \hat{\mathbf{H}}_t \tilde{\mathbf{c}}_t \right\|^2 \quad (2)$$

In Equation (2) it is assumed that some form of channel estimation resulting in $\{\hat{\mathbf{H}}_{1:T}\}$ has been performed prior to detection.

3. RECURSIVE CHANNEL ESTIMATION IN SPACE-FREQUENCY TRELLIS CODED SYSTEMS

In this section the novel technique is developed. It is assumed that only $\{\hat{\mathbf{H}}_0\}$ is available and that this knowledge suffices to estimate $\{\hat{\mathbf{H}}_{1:K}\}$ and decode the SFTC code. Since the concept rests on the developments from recursive Bayesian estimation theory, we start with a brief recap on recursive Bayesian estimation.

3.1. Recursive Bayesian estimation

The hidden states (unobserved signal) of interest $\{\mathbf{h}_t; t \in N\}$ are modelled as a Markov process:

$$f(\mathbf{h}_t | \mathbf{h}_0, \dots, \mathbf{h}_{t-1}, \mathbf{y}_1, \dots, \mathbf{y}_t) = f(\mathbf{h}_t | \mathbf{h}_{t-1}) \quad (3)$$

The observations $\{\mathbf{y}_{1:t}\}$ are independently and identically distributed (iid) conditioned on the current state:

$$f(\mathbf{y}_t | \mathbf{h}_0, \dots, \mathbf{h}_t, \mathbf{y}_1, \dots, \mathbf{y}_{t-1}) = f(\mathbf{y}_t | \mathbf{h}_t)$$

At time t the joint posterior distribution is given by Bayes' theorem:

$$f(\mathbf{h}_{0:t} | \mathbf{y}_{1:t}) = \frac{f(\mathbf{y}_{1:t} | \mathbf{h}_{0:t}) f(\mathbf{h}_{0:t})}{\int f(\mathbf{y}_{1:t} | \mathbf{h}_{0:t}) f(\mathbf{h}_{0:t}) d\mathbf{h}_{0:t}} \quad (4)$$

The problem amounts to finding a transform $f(\mathbf{h}_{0:t+1} | \mathbf{y}_{0:t+1}) = \Phi \{ f(\mathbf{h}_{0:t} | \mathbf{y}_{0:t}) \}$. To find out the required transform Φ we invoke Bayes' theorem:

$$\frac{f(\mathbf{h}_{0:t+1} | \mathbf{y}_{1:t+1})}{f(\mathbf{y}_{t+1} | \mathbf{h}_{0:t+1}, \mathbf{y}_{1:t})} = \frac{f(\mathbf{h}_{0:t+1} | \mathbf{y}_{t+1}, \mathbf{y}_{1:t})}{f(\mathbf{y}_{t+1} | \mathbf{y}_{1:t})} \quad (5)$$

However, the observations are conditionally independent: $f(\mathbf{y}_{t+1} | \mathbf{h}_{0:t+1}, \mathbf{y}_{1:t}) = f(\mathbf{y}_{t+1} | \mathbf{h}_{0:t+1})$. This leads to the required recursive formula:

$$f(\mathbf{h}_{0:t+1} | \mathbf{y}_{1:t+1}) = \frac{f(\mathbf{y}_{t+1} | \mathbf{h}_{0:t+1}) f(\mathbf{h}_{0:t+1} | \mathbf{y}_{1:t})}{f(\mathbf{y}_{t+1} | \mathbf{y}_{1:t})} \quad (6)$$

The above step updates the prior density $f(\mathbf{h}_{0:t+1} | \mathbf{y}_{1:t})$ once the measurements \mathbf{y}_{t+1} become available. To complete the recursions the prior density has to be specified. This is known as a prediction step:

$$f(\mathbf{h}_{0:t+1} | \mathbf{y}_{1:t}) = \int f(\mathbf{h}_{t+1} | \mathbf{h}_t) f(\mathbf{h}_t | \mathbf{y}_{0:t}) d\mathbf{h}_t \quad (7)$$

Equations (6) and (7) constitute a basis for Bayesian recursive estimation. Deceptively, the above recursions are straightforward to perform. However, the integrals involved are, in general, too difficult to compute. An exception is the case, where the states evolve according to some linear function and both the state and the observation noise are Gaussian.

3.2. Application to channel estimation in STTC

It is well known that the Kalman filter is an optimum Bayesian recursive estimator when both the state transitions and observation systems are linear and both the state and the observation noise are Gaussian. The Kalman filter performs the recursions from previous section, when the underlying problem has a form:

The estimated state \mathbf{h}_t evolves according to:

$$\mathbf{h}_{t+1} = \mathbf{A}_{t+1} \mathbf{h}_t + \mathbf{w}_{t+1} \quad (8)$$

And the observed signal is given by:

$$\mathbf{y}_{t+1} = \mathbf{C}_{t+1} \mathbf{h}_{t+1} + \mathbf{v}_{t+1} \quad (9)$$

The state noise \mathbf{w}_t and the observation noise \mathbf{v}_t are distributed according to:

$$f(\mathbf{w}_t) = N(\mathbf{0}, \mathbf{Q}) \quad (10)$$

$$f(\mathbf{v}_t) = N(\mathbf{0}, \mathbf{R}) \quad (11)$$

Equations (8,9,10,11) imply that the estimated process evolves sequentially and constitutes what is known as Gauss-Markov random process.

We can now cast our problem of channel estimation in STTC onto the above framework. In order to apply the Kalman filter framework we need to redefine equation (1) to the equivalent form as in (9), which is achieved by defining \mathbf{C}_t and \mathbf{h}_t as:

$$\mathbf{C}_t = \begin{bmatrix} \mathbf{c}_t^T & \mathbf{0}^T & \dots & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{c}_t^T & \dots & \mathbf{0}^T \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}^T & \mathbf{0}^T & \dots & \mathbf{c}_t^T \end{bmatrix} \quad (12)$$

$$\mathbf{h}_t = \text{vec} \{ \mathbf{H}_t^T \} \quad (13)$$

As mentioned in the previous section, sequential estimation amounts to repeated calculations of two alternating steps: prediction and update.

3.2.1. Prediction

Suppose that the random variable \mathbf{h}_t conditioned on the observations $\mathbf{y}_{1:t}$, is Gaussian:

$$f(\mathbf{h}_t | \mathbf{y}_{1:t}) = N(\mu_t, \mathbf{P}_t) \quad (14)$$

From equation (8) it can be deduced that $f(\mathbf{h}_{t+1} | \mathbf{h}_t) = N(\mathbf{A}\mathbf{h}_t, \mathbf{Q})$ (\mathbf{A} is a constant matrix here: $\mathbf{A}_t = \mathbf{A}$). Then from equation (7) the predictive marginal distribution is given by:

$$f(\mathbf{h}_{t+1} | \mathbf{y}_{1:t}) = \int N(\mathbf{A}\mathbf{h}_t, \mathbf{Q}) N(\mu_t, \mathbf{P}_t) d\mathbf{h}_t \quad (15)$$

After some tedious but straightforward manipulations¹ this becomes:

$$f(\mathbf{h}_{t+1} | \mathbf{y}_{1:t}) = N\left(\mathbf{A}\mu_t, \mathbf{Q} + \mathbf{A}\mathbf{P}_t\mathbf{A}^H\right) \quad (16)$$

Following definitions are made: $\mu_{t+1|t} = \mathbf{A}\mu_t$ and $\mathbf{P}_{t+1|t} = \mathbf{Q} + \mathbf{A}\mathbf{P}_t\mathbf{A}$. Then the predictive density is fully defined by:

$$f(\mathbf{h}_{t+1} | \mathbf{y}_{1:t}) = N(\mu_{t+1|t}, \mathbf{P}_{t+1|t}) \quad (17)$$

3.2.2. Up-date

Using (17) and (11), the update formula of (6) can be specified as:

$$f(\mathbf{h}_{t+1} | \mathbf{y}_{1:t+1}) = \frac{N(\mathbf{C}_{t+1} \mathbf{h}_{t+1}, \mathbf{R}) N(\mu_{t+1}, \mathbf{P}_{t+1|t})}{f(\mathbf{y}_{t+1} | \mathbf{y}_{1:t})} \quad (18)$$

After some tedious algebraic manipulations the posterior marginal density is expressed as:

$$f(\mathbf{h}_{t+1} | \mathbf{y}_{1:t+1}) = N(\mu_{t+1}, \mathbf{P}_{t+1}) \quad (19)$$

with the following notation:

$$\begin{aligned} \mathbf{P}_{t+1} &= [\mathbf{I} - \mathbf{K}_{t+1} \mathbf{C}_{t+1}] \mathbf{P}_{t+1|t} \\ \mu_{t+1} &= \mu_{t+1|t} + \mathbf{K}_{t+1} [\mathbf{y}_{t+1} - \mathbf{C}_{t+1} \mu_{t+1|t}] \\ \mathbf{K}_{t+1} &= \mathbf{P}_{t+1|t} \mathbf{C}_{t+1}^H [\mathbf{R} + \mathbf{C}_{t+1} \mathbf{P}_{t+1|t} \mathbf{C}_{t+1}^H]^{-1} \end{aligned} \quad (20)$$

Since both the predictive density $f(\mathbf{h}_{t+1} | \mathbf{y}_{1:t})$ and the up-dated posterior density $f(\mathbf{h}_{t+1} | \mathbf{y}_{1:t+1})$ are Gaussian, the mean and covariance describes them completely.

3.3. Description of the proposed algorithm

The initial training results in an estimate $\{\hat{\mathbf{H}}_0\}$. This estimate, together with the corresponding covariance matrix, is propagated to the next time instant using eq (16). The prior channel estimate ($\hat{\mathbf{H}}_{t+1|t}$ and thus $\hat{\mathbf{h}}_{t+1|t}$ via eq. (13)) is simply the mean ($\mu_{t+1|t}$) of the predictive density in (16). It is important for this method that the SFTC trellis always starts from a known state. This is depicted in figure 1, where it is assumed that the trellis starts from a zero state. For simplicity, of presentation figure 1 depicts a 4 state BPSK space-time code. In this case the two possible input symbols ('0' and '1') would result in two transitions (to state 0 and state 1 respectively). Corresponding to the two transitions there are two codewords: $\mathbf{c}^{(0,0)}$ and $\mathbf{c}^{(0,1)}$ (and thus $\mathbf{C}^{(0,0)}$ and $\mathbf{C}^{(0,1)}$ via eq. (12)). The superscript (i, j) denotes a transition from an i^{th} state to the j^{th} . Subsequently, the update of the channel estimates is performed. Using set of equations (20) the process (channels), the channels covariance matrices and the Kalman gain matrices all get updated. Since the $\mathbf{C}^{(i,j)}$ are in general different, the update

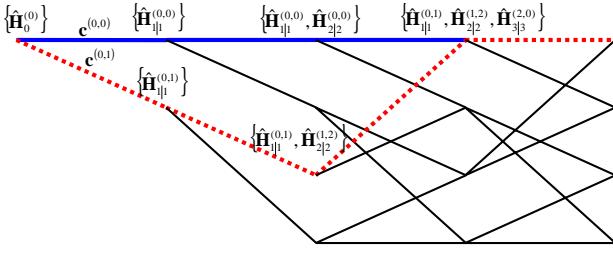


Fig. 1. Trellis description of the algorithm.

process results in obviously different posterior estimates for states 0 and 1.

This procedure continues until the state transitions in the trellis merge (the fourth segment of the trellis in figure 1) and all hypotheses till then are retained. At this stage a decision is made. The two paths merging at each stage correspond to two distinct hypotheses. Each with a set of codewords $\{c_{1:t}^{(i,j)}\}$ and a set of $\{\hat{H}_{1:t}^{(i,j)}\}$. Assuming that the Kalman filters track the channels with sufficient accuracy, a decision can be made to retain only one hypothesis using a Euclidean distance criterion eq (2). For example, in figure 1 the dashed path is retained and with it the history of channel posterior estimates $\{\hat{H}_{1|1}^{(0,1)}, \hat{H}_{2|2}^{(1,2)}, \hat{H}_{3|3}^{(2,0)}\}$. This is the last estimate in this set that will be propagated to obtain the prior estimate for all transitions originating from this state. This procedure is repeated for all states and all frequency tones. If the trellis is terminated (forced to the zero state) the last decision taken at the zero state will identify a path assumed to be correct and with it the whole space-frequency codeword $\{c_{1:T}^{(i,j)}\}$ and the channel estimates $\{\hat{H}_{1:T}^{(i,j)}\}$.

In the blind case there is no initial training and the initial estimate is set to zero $\{\hat{H}_0 = 0\}$.

Table 1 summarises the proposed algorithm.

Table 1. Algorithm summary

INITIALISE: $\hat{h}_0, \mathbf{A}, \mathbf{Q}, \mathbf{P}_0$
RECURSIONS: for $t=1:T$, for $j=1:J$
$\hat{h}_{t+1 t}^{(j)} = \mathbf{A}\hat{h}_{t t}^{(i)}$
$\mathbf{P}_{t+1 t}^{(j)} = \mathbf{Q} + \mathbf{A}\mathbf{P}_{t t}^{(i)}\mathbf{A}^H$
$[\mathbf{C}_{t+1}^{(j)}, \psi_{t+1}^{(j)}] = \arg \min_{\tilde{\mathbf{C}}^{(i,j)}} \left\{ \left\ \mathbf{y}_{t+1} - \tilde{\mathbf{C}}^{(i,j)}\hat{h}_{t+1 t}^{(j)} \right\ ^2 + \psi_t^{(i)} \right\}$
$\mathbf{K}_{t+1 t+1}^{(j)} = \mathbf{P}_{t+1 t}^{(j)} \mathbf{C}_{t+1}^{(j)H} \left[\mathbf{R} + \mathbf{C}_{t+1}^{(j)} \mathbf{P}_{t+1 t}^{(j)} \mathbf{C}_{t+1}^{(j)H} \right]^{-1}$
$\mathbf{P}_{t+1 t+1}^{(j)} = [\mathbf{I} - \mathbf{K}_{t+1 t+1}^{(j)} \mathbf{C}_{t+1}^{(j)}] \mathbf{P}_{t t}^{(i)}$
$\hat{h}_{t+1 t+1}^{(j)} = \hat{h}_{t+1 t}^{(j)} + \mathbf{K}_{t+1 t+1}^{(j)} [\mathbf{y}_{t+1} - \mathbf{C}_{t+1}^{(j)} \hat{h}_{t+1 t}^{(j)}]$
TERMINATE:
trace back $\{c_{T:1}^{(0)}, \hat{h}_{T:1}^{(0)}\}$

¹The exact algebraic manipulations involve expanding the two Gaussian densities, completing the square and integrating.

4. SIMULATION RESULTS

A basic solution for space-time code detection in rapid fading environments is a use of periodic pilot sequences and interpolation filters as originally proposed in [7]. This solution although simple, has some drawbacks. The pilot intervals are dictated by the amount of expected time variations in the channel, which has to be pre-determined before transmission. If the channel changes faster than expected, then this method fails. Conversely, if the channel fades slower than expected, then it means that the bandwidth (the resources) is wasted for too many pilots. In our method only a single initial pilot is needed at the beginning of the transmission. The performance of the proposed algorithm is investigated using 16 state 4 PSK space-time code of [8]. The channels are modelled as random walk: ($\mathbf{A} = \mathbf{I}$ and $\mathbf{Q} = .05\mathbf{I}$). Figures 2 and 3 show examples of MIMO channel tracking. Both the amplitude and phase are tracked very closely even at the Doppler spread of 500Hz. Figures 4 and 5 depict a frame error rate performance of the proposed techniques. For comparison, a training based system similar to that in [7] is investigated as well. A frame of the trained system is constructed using 3 orthogonal pilot sequences: at the beginning, the end, and in the middle of the frame. The frame comprises of 2×96 4-PSK symbols. Typically FER = 1% is used as a reference point. Only the proposed technique tolerates Doppler spreads experienced by a space-time coded system travelling at over 500 kph (carrier $f_c = 850$ MHz). Note that the performance of the trained technique (similar to [7]) could be improved by inserting more pilots, however this would degrade the bandwidth efficiency even further.

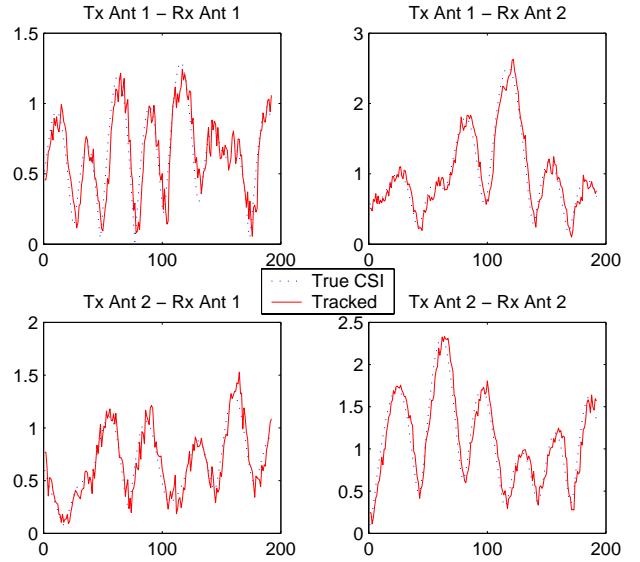


Fig. 2. Amplitude tracking versus consecutive data symbols (4-PSK) over 1 frame @ FDoppler = 500Hz (635kph) of our technique. The MIMO system uses 16 state 4-PSK code with 2 Tx and 2 Rx antennas.

5. CONCLUSIONS

A novel joint detection and channel estimation technique has been developed for Space-Time Trellis Coded systems. The technique uses only a single initial estimate in the semi-blind and none in the

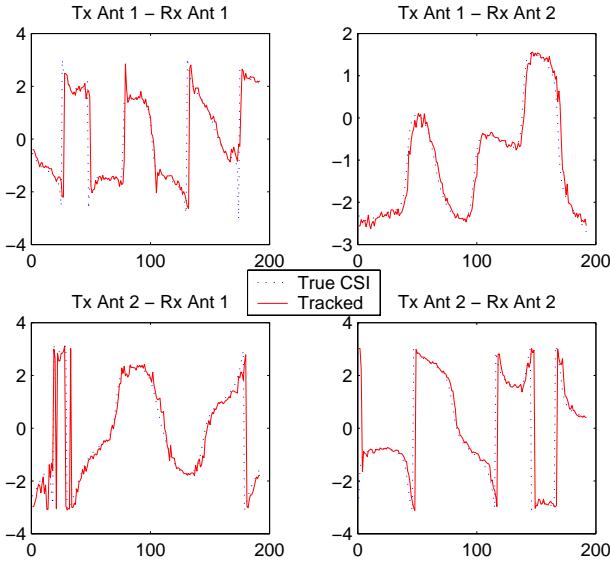


Fig. 3. Phase tracking versus consecutive data symbols (4-PSK) over 1 frame @ FDoppler = 500Hz (635kph) of our technique. The MIMO system uses 16 state 4-PSK code with 2 Tx and 2 Rx antennas.

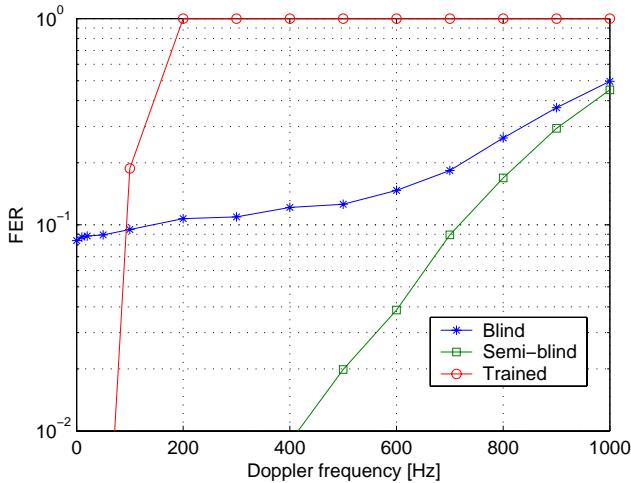


Fig. 4. FER of space-time coding enhanced D-AMPS (IS-136) versus Doppler frequency (carrier fc = 850MHz).

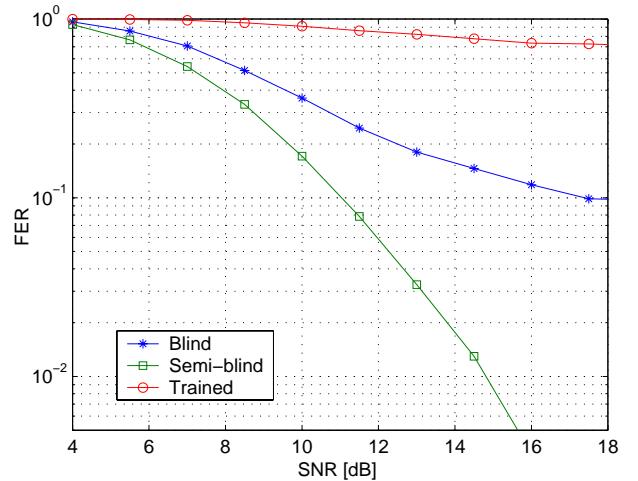


Fig. 5. FER of space-time coding enhanced D-AMPS (IS-136) versus SNR @ FDoppler = 120Hz (carrier fc = 850MHz).

blind mode. In both cases the entire channel estimate is fully recovered and SFTC is decoded even in fast faded channels. The technique allows space-time coded systems operate at much higher Doppler spreads. It does not require determination of the expected Doppler spread before transmission.

6. 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*, vol. 6, no. 3, pp. 311 – 335, 1998.
- [2] V. Tarokh, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication - performance criterion and code construction," *IEEE Trans on Information Theory*, vol. 44, no. 2, pp. 744 – 765, 1998.
- [3] Z. Liu, X. Ma, and G. Giannakis, "Space-time coding and kalman filtering for time selective fading channels," *IEEE Transactions on Communications*, vol. 50, no. 2, pp. 183 – 186, 2002.
- [4] J. Zhang and P. Djuric, "Joint estimation and decoding of space-time trellis codes," *EURASIP Journal on Applied Signal Processing*, vol. 2002, no. 3, pp. 305 – 315, 2002.
- [5] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Transactions on Communications*, vol. 42, no. 234, pp. 1000 – 1011, 1994.
- [6] R. Raheli, Polydoros, and C Tzou, "Per-survivor processing: A general approach to mlse in uncertain environments," *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 354 – 364, 1995.
- [7] A. Naguib, V. Tarokh, N Seshadri, and A. Calderbank, "A space-time coding based modem for high data rate wireless communications," *IEEE Journal on Selected areas in Communications*, vol. 16, pp. 1459 – 1478, 1998.
- [8] S. Baro, G. Bauch, and A. Hansmann, "Improved codes for space-time trellis coded modulation," *IEEE Communications Letters*, vol. 4, no. 1, 2000.