SEPARATE TEMPORAL AND SPATIAL PARAMETRIC CHANNEL ESTIMATION

Jonas Strandell and Erik Lindskog

Signals and Systems, University of Uppsala, Box 528, SE-751 20 Uppsala, Sweden

Phone: +46-18-4713071, Fax: +46-18-555096, Email: jost@radiodesign.se, el@signal.uu.se

ABSTRACT

A temporally parameterized channel estimate can be improved by projection onto a spatially parameterized subset. The projection is performed in a spectrum norm sense, and is investigated by means of simulations. By utilizing the channel estimation method in a multidimensional MLSE detector, significant improvements in the bit-error rates can be obtained, compared to using the initial estimate of the channel.

1. INTRODUCTION

Due to the increasing demand of capacity in wireless telecommunications, it is of great interest to find ways to increase the spectral efficiency. Since the total available bandwidth is limited, the focus is set on the spatial dimension. The spatial dimension can be exploited by means of an array antenna at the base station, in combination with a spatio-temporal equalizer. An indirect equalizer requires an estimate of the spatio-temporal channel from the mobile to the base station.

In [1], a maximum likelihood method for estimation of direction of arrivals (DOAs) and relative gains in a path model, for several possibly coherent signals, is derived. In [2] this method is utilized in order to exploit the spatial structure when identifying channels to an array of antennas. In this paper we propose an improved method. First, an initial estimate of the channels to the individual antenna elements is formed, exploiting the *a priori* known temporal filtering in the transmitter and the receiver, as described in [3]. The spatial structure of the channel is then exploited by projecting the channel onto a subset of channels parameterized in DOAs and relative gains.

2. CHANNEL IDENTIFICATION UTILIZING PULSE SHAPING INFORMATION

Consider the uplink transmission. The array is assumed to have M elements. The symbols are pulse shaped at the transmitter, upconverted to the air interface frequency and transmitted through the air interface. Filtering and downconversion is then performed at the receiver. The resulting system can be modeled by a continuous time filter with impulse response $p(t_c)$ (t_c denotes continuous time) containing all *a priori* known filtering at the transmitter and the receiver, a physical channel with impulse response $h_i(t_c)$ from the mobile to the i:th antenna element, and additive noise $n_i(t_c)$. Figure 1 shows the continuous time model, where the derivative operator p is used. Setting the sampling period T to 1, we get the output sequence $y_i(t)$ for the i:th antenna, where t represents discrete time. The transfer function from the symbols $d(t)$ to the discrete time output $y_i(t)$ will therefore be modeled by the FIR-filter

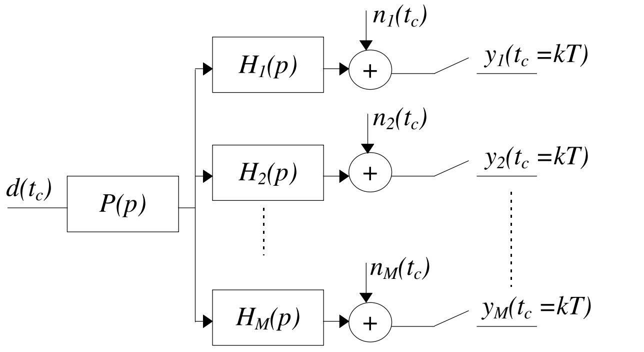


Fig. 1. Continuous time transmission model.

$B_i(q^{-1})=b_{i0}+b_{i1}q^{-1}+...+b_{i,nb}q^{-nb}$, where q^{-1} represents the backward shift operator, i.e. $q^{-1}x(t)=x(t-1)$. The discrete time model is depicted in Figure 2.

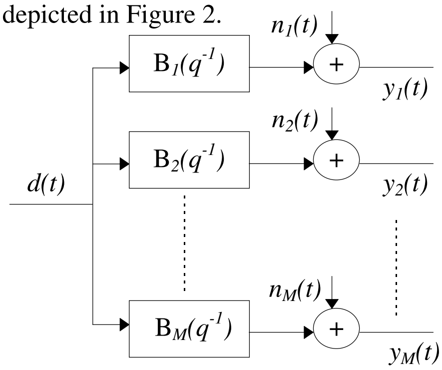


Fig. 2. Simple discrete time model.

The channels $B_i(q^{-1})$ can now be identified by formulating a least-squares problem, but since the filtering performed at both the receiver and transmitter can be considered known, it is wiser to use this information and only estimate the unknown part of the channel [3]. In Figure 3 this approach is applied to an array of antennas. Two samples per symbol and two pulse shaping branches per channel are used. A least-squares problem can now be formulated using the known signals $x_{ik}(t)$ and $y_{jk}(t)$ to identify the polynomials $H_{ji}(q^{-1})$, giving the overall Mx2 polynomial channel matrix

$$\hat{B}_0(q^{-1}) = H(q^{-1})P(q^{-1}), \tag{1}$$

with $H(q^{-1})$ and $P(q^{-1})$ defined in equations (11) and (13).

3 PROJECTION ONTO A SPATIALLY PARAMETERIZED SUBSET

In this section, the *overall* channel estimate obtained according to the method described above, here denoted $\hat{B}_0(q^{-1})$, is projected onto a spatially parameterized subset $B_{\theta,\gamma}(q^{-1})$. This will in general improve the channel estimate since the *spatial structure* of the channel is exploited. The Mx2 polynomial matrix $\hat{B}_0(q^{-1})$ is used as an initial estimate of the channel from the mobile to the array antenna. If we assume that the channel is constituted of multiple reflections of the transmitted signal,

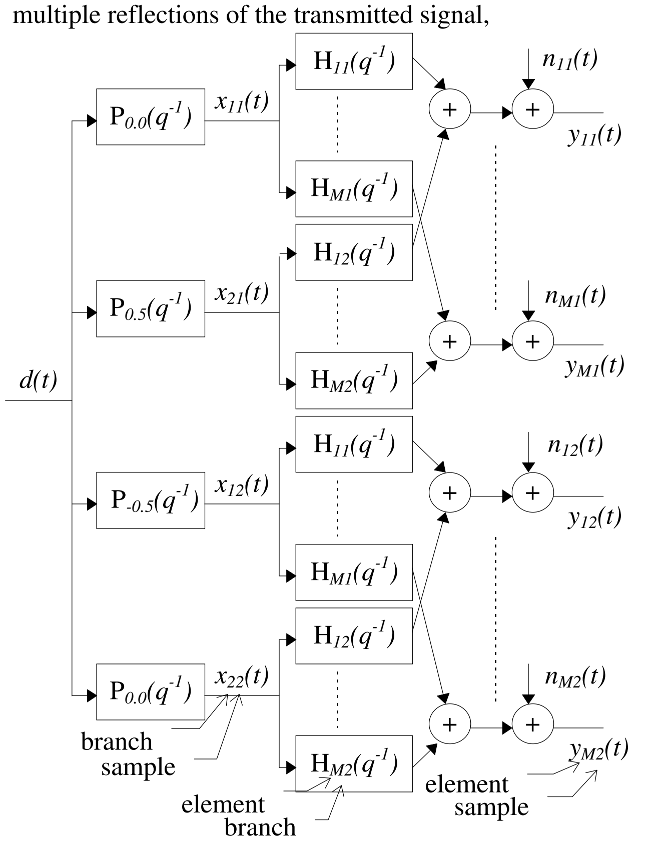


Fig. 3. Channel model utilizing multiple pulse shaping branches and oversampling.

the true channel will belong to a subset, parameterized by angles, θ, and relative gains (signal strength), γ, of the multipath components. The initial estimate $\hat{B}_0(q^{-1})$ will in general not belong to the subset $\left\{B_{\theta,\gamma}(q^{-1})\right\}$, and the idea is therefore to project $\hat{B}_0(q^{-1})$ onto $\left\{B_{\theta,\gamma}(q^{-1})\right\}$. The projection can be illustrated as in Figure 4.

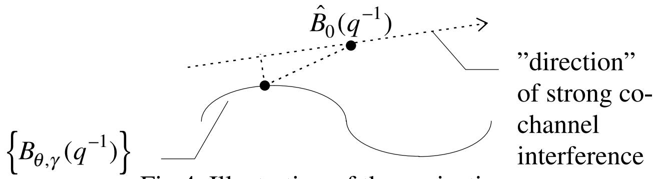


Fig 4. Illustration of the projection.

Errors in the initial estimate in "directions" where we have strong co-channel interference are paid little attention. A more detailed description of the projection is given below, but first the involved parameters and variables are defined.

The parameterized subset $\left\{B_{\theta,\gamma}(q^{-1})\right\}$ has the elements

$$
\begin{aligned}
B_{\theta,\gamma}(q^{-1}) &= \sum_{l=0}^{nb} b_l(\theta_l,\gamma_l)q^{-l} = \\
&= \sum_{l=0}^{nb} \left[b_{1l}(\theta_{1l},\gamma_{1l}) \quad b_{2l}(\theta_{2l},\gamma_{2l})\right]q^{-l}
\end{aligned}, \tag{2}
$$

where $\theta_{ij} \in [0,2\pi]$ and $\gamma_{ij} \in \Re$. The *l*:th *vector* tap corresponding to the *k*:th fractional sample can be written as

$$
b_{kl}(\theta_{kl},\gamma_{kl}) = \sum_{m=1}^{K_{kl}} \gamma_{klm} a(\theta_{klm}) \tag{3}
$$

where $a(\theta_{klm})$ is the Mx1 array response vector parameterized by the angle θ_{klm}. The number of impinging plane waves for the *l*:th vector tap and the *k*:th fractional sample is given by K_{kl}. The angles and relative gains corresponding to each of the plane waves are collected in the parameter *vectors* θ_{kl} and γ_{kl}, respectively, given by

$$
\begin{aligned}
\theta_{kl} &= \left[\theta_{kl1} \quad \dots \quad \theta_{klK_{kl}}\right]^T \\
\gamma_{kl} &= \left[\gamma_{kl1} \quad \dots \quad \gamma_{klK_{kl}}\right]^T
\end{aligned}. \tag{4}
$$

The parameter *matrices* θ and γ are defined as

$$
\begin{aligned}
\theta &= \left[\theta_{10} \quad \theta_{20} \quad \theta_{11} \quad \theta_{21} \quad \cdots \quad \theta_{1nb} \quad \theta_{2nb}\right] \\
\gamma &= \left[\gamma_{10} \quad \gamma_{20} \quad \gamma_{11} \quad \gamma_{21} \quad \cdots \quad \gamma_{1nb} \quad \gamma_{2nb}\right]
\end{aligned}. \tag{5}
$$

In a realistic situation, one would have to consider the effects of over- and underestimation of the number of paths K_{kl} per sample and tap. In this study, we also make the assumption that all angles are distinct, i.e. no angular spread.

The initial channel estimate $\hat{B}_0(q^{-1})$ is projected onto the subset $\left\{B_{\theta,\gamma}(q^{-1})\right\}$ in a *spectrum norm* sense, i.e.

$$
\hat{\theta},\hat{\gamma} = \arg\min_{\theta,\gamma} \left\| \hat{B}_0(q^{-1}) - B_{\theta,\gamma}(q^{-1}) \right\|^2_{R^{-1}_{\Delta\hat{B}_0,\Delta\hat{B}_0}(q,q^{-1})} \tag{6}
$$

giving $B_{\hat{\theta},\hat{\gamma}}(q^{-1})$. The norm $\left\|A(q^{-1})\right\|^2_{W(q,q^{-1})}$ is defined as the *trace of the constant term* of the matrix polynomial

$$
A^H(q)W(q,q^{-1})A(q^{-1})\ . \tag{7}
$$

The *error* of the initial estimate $\hat{B}_0(q^{-1})$ is defined as

$$
\Delta\hat{B}_0(q^{-1}) = \hat{B}_0(q^{-1}) - B_{true}(q^{-1}) \tag{8}
$$

where $B_{true}(q^{-1})$ is the true channel (in the subset $\left\{B_{\theta,\gamma}(q^{-1})\right\}$). The double-sided matrix polynomial $R_{\Delta\hat{B}_0,\Delta\hat{B}_0}(q,q^{-1})$ is the expectation of the spectrum of $\Delta\hat{B}_0(q^{-1})$, i.e.

$$
R_{\Delta\hat{B}_0,\Delta\hat{B}_0}(q,q^{-1}) = E\left[\Delta\hat{B}_0(q^{-1})\Delta\hat{B}_0{}^H(q)\right]. \tag{9}
$$

The weighting $R^{-1}_{\Delta\hat{B}_0,\Delta\hat{B}_0}(q,q^{-1})$ puts high weights in the areas of the spatio-temporal spectrum $\hat{B}_0(q^{-1})\hat{B}_0^H(q)$ that have low uncertainties, and low weights in the areas of $\hat{B}_0(q^{-1})\hat{B}_0^H(q)$ that have high uncertainties.

We now derive an expression for $R_{\Delta\hat{B}_0,\Delta\hat{B}_0}(q,q^{-1})$. The output matrix $Y(t)$ can be expressed as

$$
Y(t) = \mathrm{H}(q^{-1})X(t) \tag{10}
$$

where $H(q^{-1})$ and $X(t)$ are defined as

$$\mathrm{H}(q^{-1}) = \begin{bmatrix} \mathrm{H}_{11}(q^{-1}) & \mathrm{H}_{21}(q^{-1}) & \ldots & \mathrm{H}_{M1}(q^{-1}) \\ \mathrm{H}_{12}(q^{-1}) & \mathrm{H}_{22}(q^{-1}) & \ldots & \mathrm{H}_{M2}(q^{-1}) \end{bmatrix}^T$$

and (11)

$$X(t) = \begin{bmatrix} x_{11}(t) & x_{12}(t) \\ x_{21}(t) & x_{22}(t) \end{bmatrix} = P(q^{-1})d(t) \tag{12}$$

with

$$P(q^{-1}) = \begin{bmatrix} P_{0.0}(q^{-1}) & P_{-0.5}(q^{-1}) \\ P_{0.5}(q^{-1}) & P_{0.0}(q^{-1}) \end{bmatrix} \tag{13}$$

and

$$P_{0.0}(q^{-1}) = p_{0.0,0} + p_{0.0,1}q^{-1} + \ldots + p_{0.0,np}q^{-np}. \tag{14}$$

The polynomial matrix $\mathrm{H}(q^{-1})$ can be estimated with the least-squares method as

$$\hat{H}(q^{-1}) = \hat{R}_{YX}(q,q^{-1})\hat{R}_{XX}^{-1}(q,q^{-1}) \tag{15}$$

where

$$\hat{R}_{YX}(q,q^{-1}) = \frac{1}{N_u}\sum_t \sum_{m=-np}^{np+nh} Y(t+m)X^H(t)q^{-m} \tag{16}$$

$$\hat{R}_{XX}(q,q^{-1}) = \frac{1}{N_u}\sum_t \sum_{m=-np}^{np} X(t+m)X^H(t)q^{-m} \tag{17}$$

and N_u is the number of terms used in the approximation. Above, np+1 and nh+1 are the number of taps in the pulse shaping filter and in the physical channel, respectively.

The output matrix $Y(t)$ can also be written as

$$Y(t) = Y_{true}(t) + N(t) \tag{18}$$

where Y_{true} is the "noise-free" output given by

$$Y_{true}(t) = \mathrm{B}_{true}(q^{-1})d(t) \tag{19}$$

and $N(t)$ is an Mx2 noise matrix. It is important to note that we have neglected errors in the channel estimate due to an incorrect model structure, i.e. only errors due to the noise is considered. Define the residues $\varepsilon(t)$ as

$$\varepsilon(t) = Y(t) - \hat{B}_0(q^{-1})d(t). \tag{20}$$

The channel error is then given by

$$\begin{aligned} &\Delta\hat{B}_0(q^{-1}) = \hat{B}_0(q^{-1}) - B_{true}(q^{-1}) = \\ &= (\hat{H}(q^{-1}) - H_{true}(q^{-1}))P(q^{-1}) = \Delta H(q^{-1})P(q^{-1}) = \\ &= \hat{R}_{\varepsilon X}(q,q^{-1})\hat{R}_{XX}^{-1}(q,q^{-1})P(q^{-1}) \end{aligned} \tag{21}$$

where $\hat{R}_{\varepsilon X}(q,q^{-1})$ is defined similarly to (16). For a large number of training symbols N, the following holds

$$\hat{R}_{XX}(q,q^{-1}) \approx P(q^{-1})P^H(q) \tag{22}$$

and

$$\begin{aligned} &R_{\Delta\hat{B}_0\Delta\hat{B}_0}(q,q^{-1}) = E\left[\Delta\hat{B}_0(q^{-1})\Delta\hat{B}_0^{\,H}(q)\right] \approx \\ &\approx E\{\sum_{t_1,m_1,t_2,m_2} \varepsilon(t_1+m_1)d^H(t_1)q^{m_2-m_1} \times \\ &\times P^H(q)\left[P(q^{-1})P^H(q)\right]^{-1}P(q^{-1}) \times \\ &\times P^H(q)\left[P(q^{-1})P^H(q)\right]^{-H}P(q^{-1})d^H(t_2)\varepsilon(t_2+m_2)\} \approx \\ &\approx E\{\varepsilon(t)P^H(q)\left(P(q^{-1})P^H(q)\right)^{-H}P(q^{-1})\varepsilon^H(t)\}. \end{aligned} \tag{23}$$

In the approximations above we have assumed the symbols $d(t)$ and the residues $\varepsilon(t)$ to be temporally white. In practice, with co-channel interference, the noise will be temporally colored, but we still make the assumption above, since this leads to a radically simplified decoupled algorithm, as will be explained below.

For the special case of two branches and two samples per symbol, $\mathrm{P}(q^{-1})$ will be a square matrix, and the expression above simplifies to

$$\begin{aligned} &R_{\Delta\hat{B}_0\Delta\hat{B}_0}(q,q^{-1}) = E\left[\Delta\hat{B}_0(q^{-1})\Delta\hat{B}_0^H(q)\right] \approx \\ &\approx E\left[\varepsilon(t)\varepsilon^H(t)\right] = \hat{R}_{\Delta\hat{B}_0\Delta\hat{B}_0} \end{aligned}, \tag{24}$$

i.e. a constant matrix (no q-dependence), which means that $\hat{R}_{\Delta\hat{B}_0,\Delta\hat{B}_0}(q,q^{-1}) \equiv \hat{R}_{\Delta\hat{B}_0,\Delta\hat{B}_0}$ is purely spatial. This means that the projection in this case is purely spatial, and therefore we call the projection here a *parametric spatial projection*. The minimization of the norm in (6) now decouples as shown below, i.e.

$$\begin{aligned} &\left\|\hat{B}_0(q^{-1}) - B_{\theta,\gamma}(q^{-1})\right\|^2_{\hat{R}^{-1}_{\Delta B,\Delta B}} = \\ &= tr\left\{\Delta b_0^H\hat{R}^{-1}_{\Delta B_0,\Delta B_0}\Delta b_0 + \ldots + \Delta b_{nb}^H\hat{R}^{-1}_{\Delta B_0,\Delta B_0}\Delta b_{nb}\right\} \\ &= tr\left\{\Delta b_0^H\hat{R}^{-1}_{\Delta B_0,\Delta B_0}\Delta b_0\right\} + \ldots + tr\left\{\Delta b_{nb}^H\hat{R}^{-1}_{\Delta B_0,\Delta B_0}\Delta b_{nb}\right\}. \end{aligned} \tag{25}$$

where

$$\Delta b_l = \hat{b}_l - b_l(\theta_l,\gamma_l). \tag{26}$$

The terms of (25) can therefore be minimized independently. The l:th term can be expressed as

$$\begin{aligned} &tr\left\{\Delta b_l^H\hat{R}^{-1}_{\Delta B_0,\Delta B_0}\Delta b_l\right\} = tr\left\{\begin{bmatrix}\Delta b_{1l}^H \\ \Delta b_{2l}^H\end{bmatrix}\hat{R}^{-1}_{\Delta B_0,\Delta B_0}\begin{bmatrix}\Delta b_{1l} & \Delta b_{2l}\end{bmatrix}\right\} \\ &= \Delta b_{l1}^H\hat{R}^{-1}_{\Delta B_0,\Delta B_0}\Delta b_{l1} + \Delta b_{l2}^H\hat{R}^{-1}_{\Delta B_0,\Delta B_0}\Delta b_{l2}. \end{aligned} \tag{27}$$

We thus get two independent minimizations per vector tap, and the minimizing angles and gains are therefore given by

$$\hat{\theta}_{kl},\hat{\gamma}_{kl} = \arg\min_{\theta_i,\gamma_i}\left\{\left\|\hat{b}_{kl} - b_{kl}(\theta_{kl},\gamma_{kl})\right\|_{\hat{R}^{-1}_{\Delta\hat{B}_0,\Delta\hat{B}_0}}\right\}. \tag{28}$$

This can be solved as (see [1])

$$\begin{cases} \hat{\gamma}_{kl} = \left\{\hat{R}^{-1/2}_{\Delta B,\Delta B}A(\hat{\theta}_{kl})\right\}^+\hat{R}^{-1/2}_{\Delta B,\Delta B}\hat{b}_{kl} \\ \hat{\theta}_{kl} = \arg\min_{\theta_{kl}}\Big\{\hat{b}_{kl}^H\Big[\hat{R}^{-1}_{\Delta B,\Delta B} - \hat{R}^{-1}_{\Delta B,\Delta B}A(\theta_{kl}) \\ \times\left[A^H(\theta_{kl})\hat{R}^{-1}_{\Delta B,\Delta B}A(\theta_{kl})\right]^{-1}A^H(\theta_{kl})\hat{R}^{-1}_{\Delta B,\Delta B}\Big]\hat{b}_{kl}\Big\} \end{cases} \tag{29}$$

where

$$A(\theta_{kl}) = \left[a(\theta_{kl1}) \quad a(\theta_{kl2}) \quad \cdots \quad a(\theta_{klK_k})\right] \tag{30}$$

and $^+$ denotes pseudo inverse. The above non-quadratic minimization with respect to θ_{kl} can be solved by some gradient method, and the resulting angles and relative gains of each path in each tap give the final channel estimate.

4. EXAMPLES

In this section the channel identification methods discussed in Sections 2 and 3 are applied to a multipath channel with intersymbol interference. Independent Rayleigh fading is assumed. The example physical channel have taps of equal mean power at delays 0, 0.33T, 0.67T and T. A raised cosine pulse with roll-off factor 0.35 is used. For simplicity a ULA (uniform linear array) is used. Distinct angles are assumed (no distributed transmitters).

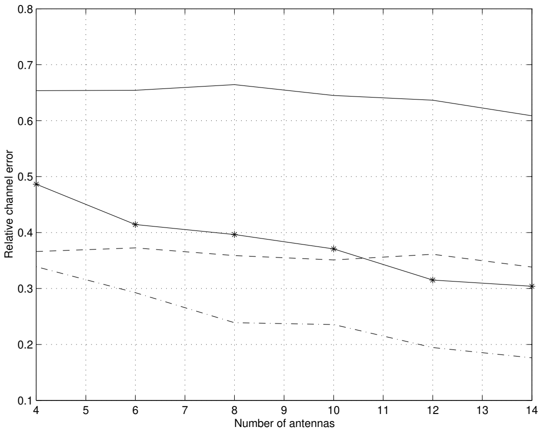
Four identification methods have been investigated and compared; unparameterized least-squares identification (LS), parametric spatial projection initialized with the unparameterized LS-estimate of the channel (LS-PSP), the method of [3] utilizing the pulse shaping information with two branches and two times oversampling (PS), and finally the here derived parametric spatial projection initialized by PS (PS-PSP). The gradient method for the non-linear minimization in the PSP-methods is initialized by the true angles. The true angles would of course not be known in a real situation, but if the angles vary slowly, estimated angles from the previous frame could be used to initiate the angle calculation of the current frame. We only need initial estimates good enough to avoid false local minima.

The methods can be compared by studying the relative channel error, defined as

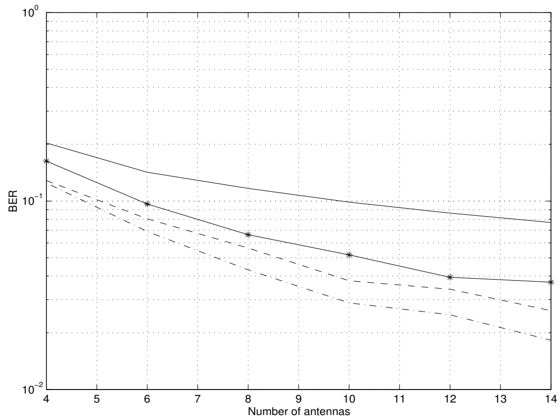
$$\left\|\hat{B}_0 - B_{true}\right\|_{Frobenius} \Big/ \left\|B_{true}\right\|_{Frobenius} \tag{31}$$

where $\hat{B}_0$ is a matrix with columns formed by the Mx2 matrix taps of $\hat{B}_0(q^{-1})$. A more useful performance criterion is the bit-error-rate (BER) achieved by using the channel estimates in a multi-channel MLSE [4].

Figure 5 shows the relative channel error and the BER as a function of the number of antennas. The signal-to-noise ratio (SNR) was chosen to -5 dB. A co-channel-interferer with the same delay profile as the signal of interest but with different DOAs was simulated, giving a signal-to-interference ratio (SIR) of -3 dB. The length of the BPSK-modulated training sequence was chosen to 26 symbols. Since LS and PS are purely temporal methods, it is not surprising that the channel errors remain constant as we increase the number of antennas. LS-PSP and PS-PSP utilizes the spatial structure of the channel, and therefore the channel errors decrease as the number of antennas increases.



(a)



(b)

Fig. 5. Relative channel error (a) and BER (b) as a function of the number of antennas. Solid is LS, asterisk LS-PSP, dashed PS and dash-dot PS-PSP.

We can see that the BER is significantly improved for PS-PSP compared to LS-PSP. If 4 % is an acceptable BER, the number of antennas can be reduced from 12 to 9. Note that the BER for all methods decreases as the number of antennas increases because of the improved antenna gain.

5. CONCLUSIONS

The channel estimate obtained by the method of [3] was improved by projection in a spectrum norm sense onto a subset parameterized in angle of arrivals and relative gains. The resulting channel estimate was compared to other estimates in terms of relative channel error and BER. The spatio-temporal methods outperform the purely temporal ones as the number of antennas is increased. A notable improvement can also be seen when comparing the proposed method to the method where an unparameterized initial LS estimate is used, see [1].

REFERENCES

[1] M. Cedervall and R. Moses, ”Efficient Maximum Likelihood DOA Estimation for Signals with Known Waveforms in the Presence of Multipath”, *IEEE Transactions on Signal Processing*, vol. 45, pp. 808-811, March 1997.

[2] E. Lindskog, ”Array Channel Identification Using Direction of Arrival Parameterization”, in *Proceedings of the IEEE International Conference on Universal Personal Communication, Cambridge, Massachusetts, U.S.A*, Sept. 29 - Oct. 2, 1996, vol. 2, pp. 999-1003.

[3] E. Lindskog and J. Strandell, “Multi-User Channel Estimation Exploiting Pulse Shaping Information”, in *Proceedings of the 9th European Signal Processing Conference,* Island of Rhodes, Greece, 8-11 Sept., 1998.

[4] E. Lindskog, ”Multi-channel maximum likelihood sequence estimation”, in *Proceedings of the 47th IEEE Vehicular Technology Conference, vol. 2, Phoenix, Arizona, USA*, May 5-7 1997, pp. 715-719.