MINIMUM SYMBOL ERROR RATE DETECTION IN SINGLE-INPUT MULTIPLE-OUTPUT CHANNELS WITH MARKOV NOISE

Lars P. B. Christensen

Technical University of Denmark, email: lc@imm.dtu.dk

ABSTRACT

Minimum symbol error rate detection in Single-Input Multiple-Output(SIMO) channels with Markov noise is presented. The special case of zero-mean Gauss-Markov noise is examined closer as it only requires knowledge of the second-order moments. In this special case, it is shown that optimal detection can be achieved by a Multiple-Input Multiple-Output(MIMO) whitening filter followed by a traditional BCJR algorithm. The Gauss-Markov noise model provides a reasonable approximation for co-channel interference, making it an interesting single-user detector for many multi-user communication systems where interference from other transmitters has a limiting effect.

1. INTRODUCTION

Interference from other users is a limiting factor in many real-life communication systems. The optimal solution is to jointly detect the desired and interfering users, but this has a complexity scaling exponentially with the number of interferers. Approaching the optimal performance with lower complexity is therefore of great interest and much work has been done within this field of research.

The idea followed in this work is to only detect the desired user and model the rest as noise. The solution to this problem of single-user minimum symbol error rate detection in channels with memory, today known simply as the BCJR algorithm after its inventors, is derived in [1] for memoryless noise. In [2] and [3], a Markov model of the noise is assumed and the optimal symbol-by-symbol and sequence detectors are derived, but only for Single-Input Single-Output(SISO) channels.

However, optimal representation of communication signals may require multiple observations per symbol. This is the case, if the signal has a bandwidth beyond the Nyquist frequency giving rise to a cyclostationary signal after digital sampling [4]. Another example is, if multiple antennas are available in the receiver. The resulting redundancy in the signal should therefore be exploited to better reject the interfering signal.

The presented framework is easily extended to include MIMO channels and/or multi-user detection by extending the discrete state-space, but this is outside the scope of this paper. Group detection of some streams/users and letting the remaining be approximated as Markov noise is also an option. Furthermore, the probabilistic nature of the detector makes it a good match for iterative decoding and parameter estimation schemes.

Section 2 presents the signal model and Section 3 derives the optimal symbol-by-symbol detector for SIMO channels with memoryless noise in order to establish the formalism that Section 4 extends to Markov noise. Next, the special case of SIMO channels with Gauss-Markov noise is looked into in Section 5. Section 6 contains simulation results of the presented detector using a GSM physical layer to highlight the benefit of the detector under the influence of co-channel interference.

2. SIGNAL MODEL

The received signal is assumed to have N_d receive dimensions and the received signal in the dth dimension is modeled as

$$r_d(t) = \sum_{n=-\infty}^{\infty} h_d(t - nT) s_n + \epsilon_d(t) \qquad (1)$$

with T being the symbol period and $\epsilon_d(t)$ being the noise process of the dth receive dimension. The nth complex symbol s_n belongs to the constellation set Ω and the number of constellation points is $|\Omega|$. Further, it is assumed that the constellation points have unit average power, i.e. $\Omega = \{+1, -1\}$ for binary modulation leading to $|\Omega| = 2$. The overall time-invariant impulse response $h_d(\tau)$ in the dth receive dimension includes pulse shaping in the transmitter, radio propagation channel and receive filtering. Furthermore, it is assumed to be zero for $\tau < 0$ and having finite temporal length LT.

In the detector, the received signal is sampled once every T for each dimension and the samples from all dimensions at the nth symbol is then put into a stacked column vector

notation as

$$\mathbf{r}\left(nT\right) = \left[r_0\left(nT\right), r_1\left(nT\right), ..., r_{N_d-1}\left(nT\right)\right]^T \quad (2)$$

where $(\cdot)^T$ indicates matrix transpose. Samples starting from the n_1'th symbol and ending at the n_2'th symbols is represented by

$$\mathbf{r}_{n_1}^{n_2} = \left[\mathbf{r}\left(n_1 T\right)^T, \mathbf{r}\left(\left(n_1+1\right)T\right)^T, \ldots, \mathbf{r}\left(n_2 T\right)^T\right]^T \quad (3)$$

Assuming $s_n = 0$ for $n < 0$ and that $N_s - L + 1$ symbols have been transmitted, the available received signal of length $N_s T$ is therefore $\mathbf{r}_0^{N_s-1}$.

3. OPTIMAL DETECTION IN SIMO CHANNELS WITH MEMORYLESS NOISE

The optimal symbol-by-symbol detector chooses the symbol having the highest posterior probability, that is

$$\hat{s}_n = \arg\max_{s_n} P\left(s_n \mid \mathbf{r}_0^{N_s-1}\right) \quad (4)$$

As a result of the finite constellation size $|\Omega|$ and the channel having a memory of $L - 1$ symbols, the task of finding the posterior distribution may be formulated as an inference problem on a Hidden Markov Model(HMM) with $|\Omega|^{L-1}$ states. The state vector of the jth state at the nth symbol determine the symbols interfering with the desired user in $\mathbf{r}_n^n$ assuming the jth combination of the interfering symbols. The state vector is therefore defined as

$$\boldsymbol{\sigma}_n^j = [\hat{s}_{n-L+1}^j, \hat{s}_{n-L+2}^j, ..., \hat{s}_{n-1}^j] \quad (5)$$

with $\hat{s}_{n-1}^j$ being the value of the $n-1$'th symbol in the jth state. Further define Ω_0 as the set of all states and $\Omega\left(s_{n-1}\right)$ as the set of states where $\hat{s}_{n-1}$ takes on a specific value from Ω. This makes it possible to rewrite the posterior probability as

$$P\left(s_{n-1} \mid \mathbf{r}_0^{N_s-1}\right) = \frac{P\left(s_{n-1}, \mathbf{r}_0^{N_s-1}\right)}{P\left(\mathbf{r}_0^{N_s-1}\right)}$$
$$= \frac{\sum_{\boldsymbol{\sigma}_n^j \in \Omega(s_{n-1})} P\left(\mathbf{r}_0^{N_s-1}, \boldsymbol{\sigma}_n^j\right)}{\sum_{\boldsymbol{\sigma}_n^j \in \Omega_0} P\left(\mathbf{r}_0^{N_s-1}, \boldsymbol{\sigma}_n^j\right)} \quad (6)$$

The result of Equation (6) is, that the desired posterior can be extracted from the joint probability $P\left(\mathbf{r}_0^{N_s-1}, \boldsymbol{\sigma}_n^j\right)$ by marginalization. Finding the joint probability is done by splitting it into a forward variable α_n^j and a backward variable β_n^j as

$$P\left(\mathbf{r}_0^{N_s-1}, \boldsymbol{\sigma}_n^j\right) = P\left(\mathbf{r}_0^{n-1}, \boldsymbol{\sigma}_n^j\right) P\left(\mathbf{r}_n^{N_s-1} \mid \mathbf{r}_0^{n-1}, \boldsymbol{\sigma}_n^j\right)$$
$$= \underbrace{P\left(\mathbf{r}_0^{n-1}, \boldsymbol{\sigma}_n^j\right)}_{\alpha_n^j} \underbrace{P\left(\mathbf{r}_n^{N_s-1} \mid \boldsymbol{\sigma}_n^j\right)}_{\beta_n^j} \quad (7)$$

The conditioning on $\mathbf{r}_0^{n-1}$ can be dropped as the state vector contains all information about the past. However, this is only true when the noise is memoryless as assumed here.

To find the forward and backward variables, let the apriori probability of a state transition be $P\left(\boldsymbol{\sigma}_n^j \mid \boldsymbol{\sigma}_{n-1}^{j'}\right)$ and the set of states where $P\left(\boldsymbol{\sigma}_n^j \mid \boldsymbol{\sigma}_{n-1}^{j'}\right) \neq 0$ as Ω_n. This makes it possible to find the forward variable recursively as

$$\alpha_n^j = P\left(\mathbf{r}_0^{n-1}, \boldsymbol{\sigma}_n^j\right) = \sum_{\boldsymbol{\sigma}_{n-1}^{j'} \in \Omega_n} P\left(\mathbf{r}_0^{n-1}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)$$
$$= \sum_{\boldsymbol{\sigma}_{n-1}^{j'} \in \Omega_n} \underbrace{P\left(\mathbf{r}_0^{n-2}, \boldsymbol{\sigma}_{n-1}^{j'}\right)}_{\alpha_{n-1}^{j'}} P\left(\boldsymbol{\sigma}_n^j \mid \boldsymbol{\sigma}_{n-1}^{j'}\right)$$
$$P\left(\mathbf{r}_{n-1}^{n-1} \mid \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right) \quad (8)$$

and the backward variable as

$$\beta_n^j = P\left(\mathbf{r}_n^{N_s-1} \mid \boldsymbol{\sigma}_n^j\right) = \sum_{\boldsymbol{\sigma}_{n+1}^{j'} \in \Omega_{n+1}} P\left(\mathbf{r}_n^{N_s-1}, \boldsymbol{\sigma}_{n+1}^{j'} \mid \boldsymbol{\sigma}_n^j\right)$$
$$= \sum_{\boldsymbol{\sigma}_{n+1}^{j'} \in \Omega_{n+1}} \underbrace{P\left(\mathbf{r}_{n+1}^{N_s-1} \mid \boldsymbol{\sigma}_{n+1}^{j'}\right)}_{\beta_{n+1}^{j'}} P\left(\boldsymbol{\sigma}_{n+1}^{j'} \mid \boldsymbol{\sigma}_n^j\right)$$
$$P\left(\mathbf{r}_n^n \mid \boldsymbol{\sigma}_{n+1}^{j'}, \boldsymbol{\sigma}_n^j\right) \quad (9)$$

In the special case where the noise $\epsilon_d\left(t\right)$ is zero-mean Additive White Gaussian Noise(AWGN) with variance σ^2 per dimension, the observation probability is conditionally Gaussian and is given by

$$-2ln\left(P\left(\mathbf{r}_{n-1}^{n-1} \mid \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)\right) + Z$$
$$= \frac{1}{\sigma^2}\left\|\mathbf{r}_{n-1}^{n-1} - \hat{\mathbf{r}}_{n-1}^{n-1}\left(\mathbf{h}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)\right\|^2 \quad (10)$$

with the normalization constant $Z = N_d ln\left(2\pi\sigma^2\right)$. The function $\hat{\mathbf{r}}_{n-1}^{n-1}\left(\mathbf{h}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)$ reconstructs the desired signal from the sampled impulse response $\mathbf{h}$ and the overall state. In the case of AWGN, the complexity of the algorithm is $\mathcal{O}\left(N_d |\Omega|^L\right)$ operations per symbol as $\mathcal{O}\left(N_d |\Omega|\right)$ operations must be performed in each of the $|\Omega|^{L-1}$ states.

4. OPTIMAL DETECTION IN SIMO CHANNELS WITH MARKOV NOISE

The noise process $\epsilon_d\left(t\right)$ is now assumed to be Markov with a finite temporal memory $N_m T$ and possibly coupled across

the receive dimensions. As before, it is desirable to model the desired signal by a HMM, but now the noise is independently modeled by a Markov model and the memory needed to correctly capture the state of the system is now $L - 1 + N_m$ as the memories add up. The number of states is now $|\Omega|^{L-1+N_m}$ and the new state vector is therefore defined as

$$\boldsymbol{\sigma}_n^j = [\hat{s}_{n-L-N_m+1}^j, \hat{s}_{n-L-N_m+2}^j, ..., \hat{s}_{n-1}^j] \quad (11)$$

As in Equation (7), the joint probability is found by splitting it into a forward variable and backward variable by

$$P\left(\mathbf{r}_0^{N_s-1}, \boldsymbol{\sigma}_n^j\right) = P\left(\mathbf{r}_0^{n-1}, \boldsymbol{\sigma}_n^j\right) P\left(\mathbf{r}_n^{N_s-1} \mid \mathbf{r}_0^{n-1}, \boldsymbol{\sigma}_n^j\right)$$
$$= \underbrace{P\left(\mathbf{r}_0^{n-1}, \boldsymbol{\sigma}_n^j\right)}_{\alpha_n^j} \underbrace{P\left(\mathbf{r}_n^{N_s-1} \mid \mathbf{r}_{n-N_m}^{n-1}, \boldsymbol{\sigma}_n^j\right)}_{\beta_n^j}$$
$$(12)$$

However, the conditioning on $\mathbf{r}_0^{n-1}$ can only be reduced to $\mathbf{r}_{n-N_m}^{n-1}$ as a result of the finite memory Markov noise and the fact that only the desired signal is described by $\boldsymbol{\sigma}_n^j$.

As before the forward variable α_n^j and the backward variable β_n^j may be found recursively. The forward variable is found by

$$\alpha_n^j = P\left(\mathbf{r}_0^{n-1}, \boldsymbol{\sigma}_n^j\right) = \sum_{\boldsymbol{\sigma}_{n-1}^{j'} \in \Omega_n} P\left(\mathbf{r}_0^{n-1}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)$$
$$= \sum_{\boldsymbol{\sigma}_{n-1}^{j'} \in \Omega_n} \underbrace{P\left(\mathbf{r}_0^{n-2}, \boldsymbol{\sigma}_{n-1}^{j'}\right)}_{\alpha_{n-1}^{j'}} P\left(\boldsymbol{\sigma}_n^j \mid \boldsymbol{\sigma}_{n-1}^{j'}\right)$$
$$P\left(\mathbf{r}_{n-1}^{n-1} \mid \mathbf{r}_{n-1-N_m}^{n-2}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)$$
$$(13)$$

and the backward variable by

$$\beta_n^j = P\left(\mathbf{r}_n^{N_s-1} \mid \mathbf{r}_{n-N_m}^{n-1}, \boldsymbol{\sigma}_n^j\right)$$
$$= \sum_{\boldsymbol{\sigma}_{n+1}^{j'} \in \Omega_{n+1}} P\left(\mathbf{r}_n^{N_s-1}, \boldsymbol{\sigma}_{n+1}^{j'} \mid \mathbf{r}_{n-N_m}^{n-1}, \boldsymbol{\sigma}_n^j\right)$$
$$= \sum_{\boldsymbol{\sigma}_{n+1}^{j'} \in \Omega_{n+1}} \underbrace{P\left(\mathbf{r}_{n+1}^{N_s-1} \mid \mathbf{r}_{n+1-N_m}^n, \boldsymbol{\sigma}_{n+1}^{j'}\right)}_{\beta_{n+1}^{j'}}$$
$$P\left(\boldsymbol{\sigma}_{n+1}^{j'} \mid \boldsymbol{\sigma}_n^j\right) P\left(\mathbf{r}_n^n \mid \mathbf{r}_{n-N_m}^{n-1}, \boldsymbol{\sigma}_{n+1}^{j'}, \boldsymbol{\sigma}_n^j\right)$$
$$(14)$$

The observation probability is now conditioned on previous received samples and a convenient way of avoiding

this is by rewriting using Bayes' rule as

$$P\left(\mathbf{r}_{n-1}^{n-1} \mid \mathbf{r}_{n-1-N_m}^{n-2}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)$$
$$= \frac{P\left(\mathbf{r}_{n-1-N_m}^{n-1} \mid \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)}{P\left(\mathbf{r}_{n-1-N_m}^{n-2} \mid \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)} \quad (15)$$

which can be evaluated directly. The complexity of the algorithm is $\mathcal{O}\left(|\Omega|^{L+N_m}\right)$ per symbol excluding the complexity involved in evaluating Equation (15).

5. OPTIMAL DETECTION IN SIMO CHANNELS WITH GAUSS-MARKOV NOISE

The special case where the noise is not only Markov, but also zero-mean Gaussian is described in this section. This model fits communication systems well and is interesting as only the covariance matrix of the noise must be known in order to evaluate Equation (15). Any Gauss-Markov process having non-zero mean is also described by this section, as it may be transformed into having zero-mean by subtracting the mean.

Let the noise samples be stacked in the same manner in $\boldsymbol{\epsilon}_{n_1}^{n_2}$ as in $\mathbf{r}_{n_1}^{n_2}$ for the received signal. To evaluate the top of Equation (15), the covariance

$$\boldsymbol{\Sigma} = E\left[\boldsymbol{\epsilon}_{n-N_m}^n \left(\boldsymbol{\epsilon}_{n-N_m}^n\right)^H\right] \quad (16)$$

is required where $(\cdot)^H$ indicates matrix transpose and complex conjugation. The lower part requires the covariance

$$\boldsymbol{\check{\Sigma}} = E\left[\boldsymbol{\epsilon}_{n-N_m}^{n-1} \left(\boldsymbol{\epsilon}_{n-N_m}^{n-1}\right)^H\right] \quad (17)$$

However, the last covariance is included in the first as can be seen by block partioning the first as

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\check{\Sigma}} & \mathbf{B} \\ \mathbf{B}^H & \mathbf{A} \end{bmatrix} \quad (18)$$

Defining

$$\mathbf{W} = \boldsymbol{\Sigma}^{-1} - \begin{bmatrix} \boldsymbol{\check{\Sigma}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (19)$$

makes it possible to rewrite Equation (15) as

$$-2ln\left(\frac{P\left(\mathbf{r}_{n-1-N_m}^{n-1} \mid \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)}{P\left(\mathbf{r}_{n-1-N_m}^{n-2} \mid \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)}\right) + Z =$$
$$\hat{\boldsymbol{\epsilon}}_{n-1-N_m}^{n-1}\left(\mathbf{h}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)^H \mathbf{W} \hat{\boldsymbol{\epsilon}}_{n-1-N_m}^{n-1}\left(\mathbf{h}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'}\right)$$
$$(20)$$

with the normalization constant being

$$Z = ln \left| \mathbf{A} - \mathbf{B}^H \check{\mathbf{\Sigma}}^{-1} \mathbf{B} \right| + N_d ln (2\pi) \qquad (21)$$

where $|\cdot|$ indicates the matrix determinant and the estimated noise is defined as

$$\hat{\boldsymbol{\epsilon}}_{n-1-N_m}^{n-1} \left(\mathbf{h}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'} \right)$$
$$= \mathbf{r}_{n-1-N_m}^{n-1} - \hat{\mathbf{r}}_{n-1-N_m}^{n-1} \left(\mathbf{h}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'} \right) \qquad (22)$$

However, the rank of $\mathbf{W}$ is not full. It is shown in Section 8 that $rank(\mathbf{W}) = N_d$, thus making it possible to express $\mathbf{W}$ as a sum of N_d vector outer products

$$\mathbf{W} = \sum_{d=0}^{N_d-1} \mathbf{f}_d \mathbf{f}_d^H$$
$$= \underbrace{[\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_{N_d-1}]}_{\mathbf{F}} \underbrace{[\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_{N_d-1}]^H}_{\mathbf{F}^H} \qquad (23)$$

Whitening of the noise process can therefore be achieved by the MIMO whitening filter $\mathbf{F}^H$. The whitened received signal and reconstructed signal are defined as

$$\tilde{\mathbf{r}}_{n-1}^{n-1} = \mathbf{F}^H \mathbf{r}_{n-1-N_m}^{n-1}$$
$$\tilde{\mathbf{r}}_{n-1}^{n-1} \left(\mathbf{h}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'} \right) = \mathbf{F}^H \hat{\mathbf{r}}_{n-1-N_m}^{n-1} \left(\mathbf{h}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'} \right) \qquad (24)$$

making it possible to rewrite Equation (20) as

$$-2ln \left(\frac{P \left(\mathbf{r}_{n-1-N_m}^{n-1} \mid \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'} \right)}{P \left(\mathbf{r}_{n-1-N_m}^{n-2} \mid \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'} \right)} \right) + Z$$
$$= \left\| \tilde{\mathbf{r}}_{n-1}^{n-1} - \tilde{\mathbf{r}}_{n-1}^{n-1} \left(\mathbf{h}, \boldsymbol{\sigma}_n^j, \boldsymbol{\sigma}_{n-1}^{j'} \right) \right\|^2 \qquad (25)$$

This is simply the AWGN version of the receiver as the signal has now been whitened by $\mathbf{F}^H$, but the channel length is now $L + N_m$ instead of L. The complexity of the receiver using the whitening filter is therefore $\mathcal{O}\left(N_d |\Omega|^{L+N_m} \right)$ operations per symbol excluding finding and applying the whitening filter. As expected, the results of [2] and [3] are recovered if $N_d = 1$ is inserted.

6. NUMERICAL SIMULATIONS

In this section, a number of simulations using a GSM physical layer is presented using the zero-mean Gauss-Markov model to approximate Co-Channel Interference(CCI). All interferers are fully synchonized in time with the desired user. The used modulation is Gaussian Minimum Shift Keying(GMSK) with $BT = 0.3$ in accordance with the GSM specifications. However, as this modulation is non-linear, it is linearized to fit the model of Equation (1) by the Laurent approximation [5]. The channel model used for all simulations is the Typical Urban(TU) model given by the GSM specifications using a mobile speed of 50 km/h at a carrier frequency of 945 MHz. All simulations are performed using perfect knowledge of the linearized impulse response of the desired user and covariance matrix of the linearized interference.

In the detector the received signal is oversampled by a factor of N_{sps} relative to the symbolrate. There are several reasons for having $N_{sps} > 1$ in a GSM receiver, one being that the bandwidth of the signal exceeds the Nyquist frequency. After sampling, the received signal is derotated to remove the rotation in the GMSK modulation [5]. Next, the received signal is split into a real and imaginary part as this improves the rejection of interferers with real constellation points, such as derotated GMSK, by exploiting redundancy in the interfering signal [6]. The number of real-valued receive dimensions for the receiver is therefore $N_d = 2N_{sps}$.

As a measure of the level of interference, the Carrier-to-Interference Ratio(CIR) is defined as the average received power for the desired user divided by the average received interference power. Further, AWGN is added to account for any thermal noise with the Signal-to-Noise Ratio(SNR) being defined as the average received signal power from the desired user divided by the noise variance.

For comparison, the performance of IQ-LMMSE[6] and BCJR detection without whitening is shown using perfect parameter estimates. Like the Gauss-Markov approximation, the IQ-LMMSE relies on second-order moments to suppress interference and therefore seems to be a reasonable comparison.

In Figure 1 the Bit Error Rate(BER) of a GSM user is plotted under the influence of one GMSK co-channel interferer. It can be seen that a gain can be achieved by increasing the oversampling from $N_{sps} = 1$ to $N_{sps} = 2$ and that the IQ-LMMSE is outperformed, especially at higher values of CIR. Higher values of N_{sps} are not included as only marginal gains compared to $N_{sps} = 2$ are achievable for this scenario. The value of $L = 7$ is selected in order to capture all energy spread out in time by the channel model and $N_m = 3$ was selected as higher values seem to provide little improvement.

In Figure 2 another interferer is added with the second interferer being 10dB weaker than the first. The performance is clearly impacted by the additional interferer, but a significant gain is still achieved over IQ-LMMSE and by using $N_{sps} = 2$ compared to $N_{sps} = 1$.

7. CONCLUSION

Minimum symbol error rate detection in SIMO channels with Markov noise has been derived as a generalization of
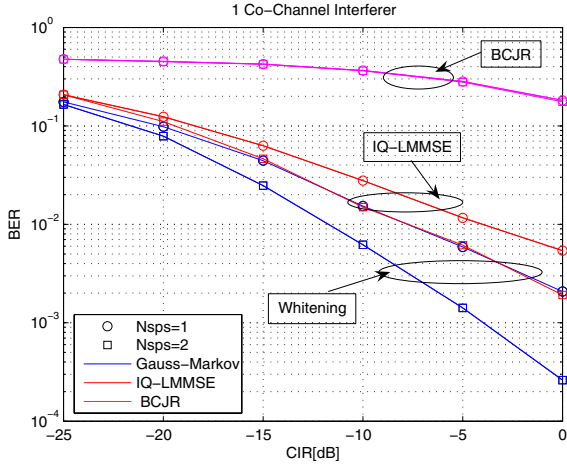
Fig. 1. BER plot of 1 CCI on TU50 channels with $L = 7$, $N_m = 3$, $SNR = 40dB$.

Fig. 2. BER plot of 2 CCIs on TU50 channels with $L = 7$, $N_m = 3$, $SNR = 40dB$.

the work presented in [2]. In the special case of Gauss-Markov noise, it has been shown that optimal detection can be achieved by a MIMO whitening filter followed by a traditional BCJR algorithm. Numerical simulations have shown that the presented detector can improve the performance compared to the IQ-LMMSE and the detector in [2] when multiple observations per symbol are available in interference-limited scenarios.

8. PROOF OF RANK

In this section it is shown that $rank\left(\mathbf{W}\right) = N_d$ with $\mathbf{W}$ defined by Equations (18)-(19). This is most easily done by using the inversion lemma for block partitioned matrices to yield

$$
\begin{aligned}
\mathbf{W} &= \mathbf{\Sigma}^{-1} - \begin{bmatrix} \mathbf{\check{\Sigma}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{\tilde{B}S\tilde{B}}^H & -\mathbf{\tilde{B}S} \\ -\mathbf{S\tilde{B}}^H & \mathbf{S} \end{bmatrix} = \begin{bmatrix} \mathbf{\tilde{B}} \\ -\mathbf{I} \end{bmatrix} \mathbf{S} \begin{bmatrix} \mathbf{\tilde{B}} \\ -\mathbf{I} \end{bmatrix}^H
\end{aligned} \quad (26)
$$

with

$$
\begin{aligned}
\mathbf{\tilde{B}} &= \mathbf{\check{\Sigma}}^{-1}\mathbf{B} \\
\mathbf{S} &= \left(\mathbf{A} - \mathbf{B}^H\mathbf{\check{\Sigma}}^{-1}\mathbf{B}\right)^{-1}
\end{aligned} \quad (27)
$$

where $\mathbf{I}$ indicates the identity matrix. As the Schur complement $\mathbf{S}$ is positive definite, it may be Cholesky factorized as $\mathbf{S} = \mathbf{CC}^H$ making it possible to rewrite Equation (26) as

$$
\mathbf{W} = \underbrace{\begin{bmatrix} \mathbf{\tilde{B}} \\ -\mathbf{I} \end{bmatrix} \mathbf{C}}_{\mathbf{F}} \underbrace{\mathbf{C}^H \begin{bmatrix} \mathbf{\tilde{B}} \\ -\mathbf{I} \end{bmatrix}^H}_{\mathbf{F}^H} \quad (28)
$$

The result in Equation (28) directly shows $rank\left(\mathbf{W}\right) = N_d$ and thereby completes the proof.

9. REFERENCES

[1] F. Jelinek L. R. Bahl, J. Cocke and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. on Info. Theory*, pp. 284–287, March 1974.

[2] A. Kavcic, "Soft-output detector for channels with intersymbol interference and markov noise memory," in *IEEE Global Telecommunications Conference*, 1999, vol. 1B, pp. 728–732.

[3] A. Kavcic and J. M. F. Moura, "The viterbi algorithm and markov noise memory," *IEEE Trans. on Info. Theory*, vol. 46, pp. 291–301, January 2000.

[4] J. G. Proakis, *Digital Communications*, McGraw-Hill, 3rd edition, 1995.

[5] P. Laurent, "Exact and approximate construction of digital phase modulations by superposition of amplitude modulated pulses (amp)," *IEEE Trans. on Communications*, vol. 34, no. 2, pp. 150–160, February 1986.

[6] R. Schober W. H. Gerstacker and A. Lampe, "Receivers with widely linear processing for frequency-selective channels," *IEEE Trans. on Communications*, vol. 51, no. 9, pp. 1512–1523, September 2003.