A Multiple-Symbol Turbo Detector for Coded M-DPSK over Time-Varying Rayleigh Flat Fading Channels

Peter Klenner and Karl-Dirk Kammeyer
University of Bremen, Department of Communications Engineering
28359 Bremen, Germany
Email: {klenner,kammeyer}@ant.uni-bremen.de

***Abstract*— In the past few years, a number of publications concentrated on developing soft-in-soft-out algorithms for coded noncoherent detection, thereby enabling iterative processing and avoiding the need of channel estimation. This paper focuses on iterative noncoherent detection of convolutionally coded M-DPSK signals for time-variant Rayleigh flat fading channels without the receiver having channel state information (CSI). Specifically we show that a noncoherent soft demodulator for minimum-shift keying, which was previously reported in the literature, lends itself quite naturally for soft demodulation of DPSK. We present an extension of this receiver in terms of per-survivor processing and examine the influence of different symbol labelings.**

I. Introduction

DPSK is well-known as a robust means to transmit data because it renders the need of carrier phase tracking unnecessary and eludes the problem of phase ambiguities, thus allowing for a simple receiver design. To overcome the entailed penalty in shape of a loss in SNR (approx. 2-3dB for QDPSK against QPSK), multiple symbol differential detection (MSDD) schemes were introduced. Numerous publications center around the idea to extend the observation interval of two symbols for conventional detection of DPSK to larger intervals. Thus e.g. for the AWGN channel, it is possible to attain the performance of differentially encoded and coherently detected M-PSK. In [1] it is shown, how MSDD is implemented for AWGN as well as Rayleigh fading channels. There, a number of maximum-likelihood metrics are derived from the multivariate Gaussian distribution.

Several publications adopted the idea of MSDD to implement iterative strategies, in order to combine the benefits of coding and multiple symbol observations. In[2] a noncoherent BCJR-type algorithm is derived which introduces a MSDD-like transition metric by truncating the observations for a transition to a few previous symbols. Hence, the generated APPs are only approximate, which is made up for by feeding back extrinsic information from the channel decoder. Only AWGN is considered in [2]. A similar approach is taken in [3], where the focus lies on turbo-decoding. Unlike [2], the alphabet is not restricted to PSK, but also only AWGN is considered. [4] advances the scheme of [3] by incorporating block fading. Although minimum-shift keying is considered, the described transition metric is generally applicable.

Time-varying Rayleigh flat fading channels are examined in [5]. The applied BCJR is derived by using linear prediction, which basically enables the estimation of the channel coefficients. Finally, in [6] a noncoherent BCJR-type demodulator is derived quite naturally from the joint probability of received symbols and hypotheses. In [6] the respective receiver is designed for a minimum-shift keying modulator. We aim to show that the same approach yields straightforwardly an APP-demodulator for an M-DPSK modulated sequence. We will present an extension of this scheme by applying per-survivor processing (PSP) [7], which is motivated by the main findings of the authors of [5], i.e., their noncoherent BCJR-approach performs satisfying for slowly time-varying channels, whereas only small gains are achieved on fast varying channels. We will demonstrate that the use of PSP allows for a significant reduction of the error-floor. We also address the influence of different symbol labelings. Whereas coherent iterative receivers can only benefit from anti-Gray mapping [10], we found that our noncoherent iterative receiver benefits from anti-Gray mapping in the error-floor region, while it yields better performance in the waterfall region applying Gray mapping.

Before outlining the paper let us stipulate some notation. A vector is denoted as $\mathbf{x}_a^b = [x_b, x_{b-1}, \cdots, x_a]^{\mathrm{T}}$. The operator $\mathcal{D}\{\mathbf{x}\}$ places vector $\mathbf{x}$ on a diagonal matrix. A capital boldface letter denotes a matrix, e.g., $\mathbf{X}$. A matrix element in the μ-th row and the ν-th column is given by $[\mathbf{X}]_{\mu\nu}$. The 0-th order modified Bessel function is denoted as $\mathrm{J}_0(\cdot)$.

Section II describes the symbol-rate equivalent baseband system model. Section III considers the noncoherent turbo receiver following the standard BCJR procedure. Results are shown in Section IV. Conclusions are given in Section V.

II. System Model

The overall transceiving scheme is depicted in Fig. 1. I.i.d. information bits $b_i \in \{0, 1\}$ are fed to a convolutional encoder (CC) yielding encoded bits c_j. After interleaving (Π) the coded bits c'_j are mapped on the differential symbols $\Delta d_k \in \{\exp(\mathrm{j}2\pi\nu/M), \nu = 0, \cdots, M-1\}$. The differential encoding rule

$$d_k = \Delta d_k \cdot d_{k-1} \tag{1}$$

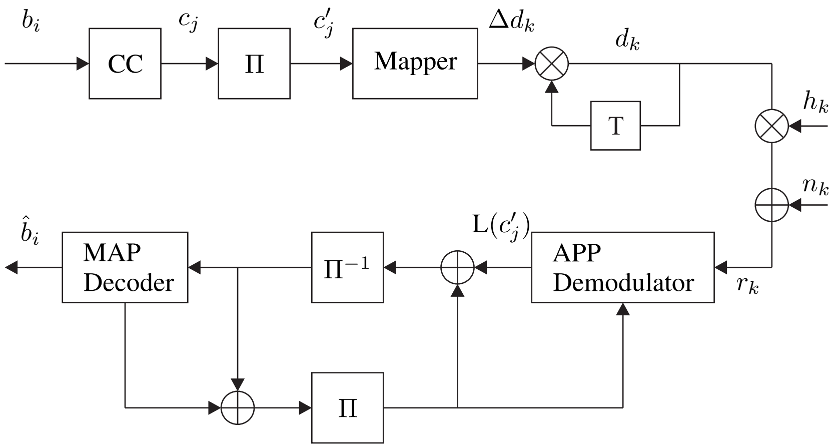

Fig. 1. Block diagram for convolutionally encoded bit-interleaved M-DPSK

yields the actually transmitted symbol d_k. Please note that d_k belongs to the same signal constellation as Δd_k. Let us also mention that the cascade of convolutional code, interleaver and differential modulation constitutes a serially concatenated system. With the received symbol r_k, the channel coefficient h_k and the noise w_k (AWGN with variance σ_w^2), the received sequence for a flat-fading channel at time instance $k > 0$ reads

$$\mathbf{r}_0^k = \mathcal{D}\left\{\mathbf{h}_0^k\right\} \cdot \mathbf{d}_0^k + \mathbf{w}_0^k, \tag{2}$$

The multivariate Gaussian distribution for $\mathbf{r}_0^k$, which is conditioned on the transmitted symbols $\mathbf{d}_0^k$ is given by

$$p\left(\mathbf{r}_0^k|\mathbf{d}_0^k\right) = \frac{\exp\left(-(\mathbf{r}_0^k)^{\mathrm{H}}\mathbf{C}_{rr}^{-1}[k]\mathbf{r}_0^k\right)}{\pi^{k+1}\det\mathbf{C}_{rr}[k]}. \tag{3}$$

where $\mathbf{C}_{rr}[k] \triangleq \mathrm{E}\{\mathbf{r}_0^k(\mathbf{r}_0^k)^{\mathrm{H}}|\mathbf{d}_0^k\}$. The determinant and the inverse of $\mathbf{C}_{rr}[k]$ are given by

$$\det\mathbf{C}_{rr}[k] = \det\mathbf{C}, \tag{4}$$

$$\mathbf{C}_{rr}^{-1}[k] = \mathcal{D}\{\mathbf{d}_0^k\}\mathbf{C}^{-1}\mathcal{D}\{\mathbf{d}_0^k\}^{\mathrm{H}}, \tag{5}$$

where the elements of matrix $\mathbf{C} = \mathrm{E}\{\mathbf{h}_0^k(\mathbf{h}_0^k)^{\mathrm{H}} + \mathbf{n}_0^k(\mathbf{n}_0^k)^{\mathrm{H}}\}$ for Rayleigh flat-fading with Jakes' Doppler spectrum ($f_{\mathrm{D,max}}T$: normalized maximum Doppler frequency) are given by

$$[\mathbf{C}]_{\mu\nu} = \mathrm{J}_0(2\pi f_{\mathrm{D,max}}T(\mu-\nu)) + \sigma_w^2\delta(\mu-\nu). \tag{6}$$

The received signal r_k enters an APP demodulator which computes soft-values $\mathrm{L}(c'_j)$ for the coded bits. Sec. III describes the details. Finally, extrinsic information is passed to an optimal maximum a posteriori (MAP) decoder [8], which performs the decoding of the convolutionally encoded bits to yield the decoded bits $\hat{b}_i$ and also delivers extrinsic information about the code bits. This can then serve as a priori information for a new demodulation/decoding cycle.

A. The role of symbol labeling

Soon after the discovery of turbo codes [9] the importance of different labeling strategies, i.e., the mapping from bits to symbols, became obvious, e.g. [10]. The main conclusion one can draw is that as soon as feedback information is available Gray-mapping may not be the preferred choice. With feedback information a higher-level signal constellation collapses to a binary signal constellation (assuming error-free feedback) in the demodulation process. Then it is important to have a mean Euclidean distance between all binary signal combinations which is as large as possible. Throughout the paper QPSK is considered, where the choice is simply between Gray mapping and natural mapping, which is the same as anti-Gray mapping for QPSK. In fact natural mapping of QPSK yields a larger mean Euclidean distance than Gray mapping. Sec. IV will demonstrate the influence of the different mapping strategies based on simulation results.

III. APP-Demodulation of M-DPSK

Having introduced the system model we are now equipped to describe the demodulator which computes APPs for the transmitted differential symbols and soft values for the interleaved code bits. We start with the definition of an L-value and basically follow the standard BCJR-procedure [8] to end in a computationally feasible algorithm. That same approach was chosen in [6]. We state the basic results to finally introduce PSP as an extension to BCJR based APP demodulation of M-DPSK.

Let us assume, that $L+1$ symbols have been received, i.e, L differential symbols Δd_k were transmitted. The respective L-value for the coded interleaved bits c'_j is then defined as

$$\mathrm{L}(c'_j) = \log\frac{\sum\limits_{\forall\Delta d_k(c'_j=0)} p(\Delta d_k, \mathbf{r}_0^L)}{\sum\limits_{\forall\Delta d_k(c'_j=1)} p(\Delta d_k, \mathbf{r}_0^L)}. \tag{7}$$

In the numerator and the denumerator $p(\Delta d_k, \mathbf{r}_0^L)$ describes the joint probability of the transmitted differential symbol Δd_k and the received sequence $\mathbf{r}_0^L$. The respective sums range over all those symbols Δd_k, which are labeled with either $c'_j = 0$ or $c'_j = 1$. One can now make use of the fact that the DPSK modulator can be interpreted as a rate one recursive convolutional code, which in turn can be described by a trellis with a certain number of states and transitions. This interpretation enables the expression of $p(\Delta d_k, \mathbf{r}_0^L)$ in terms of state transitions

$$p(\Delta d_k, \mathbf{r}_0^L) = \sum_{(s',s)\rightarrow \Delta d_k} p(s_{k-1} = s', s_k = s, \mathbf{r}_0^L). \quad (8)$$

In Eq. (8) the joint probability is expressed in terms of those state transitions (s', s) from state $s_{k-1} = s'$ to $s_k = s$ that belong to the associated transmitted symbol Δd_k.
Repeatedly applying Bayes' rule leads to a separation of the joint probability into recursively computable probabilities [8]

$$p(s', s, \mathbf{r}_0^L) = \alpha_{k-1}(s')\gamma_k(s', s)\beta_k(s), \quad (9)$$

where

$$\begin{aligned} \alpha_{k-1}(s') &= p(s', \mathbf{r}_0^{k-1}), & (10)\\ \gamma_k(s', s) &= p(r_k|s', s, \mathbf{r}_0^{k-1})\mathrm{Pr}(s|s'), & (11)\\ \beta_k(s) &= p(\mathbf{r}_{k+1}^L|s, \mathbf{r}_0^k). & (12) \end{aligned}$$

Note that in (11) the conditioning is based on the sequence $\mathbf{r}_0^{k-1}$ which led to the current receive symbol r_k. This is unlike MAP or Viterbi decoding of convolutional codes where the transitional probability only depends on the state transition and the currently received codebit, which only holds for non-correlated impairments of successively received code bits, e.g. one can savely assume that due to an interleaver code bits are received over a memoryless channel. The received DPSK sequence on the other hand inherently contains memory and moreover is strongly correlated due to the time-varying flat fading channel and that is taken into account by the respective conditioning.
Following the BCJR philosophy the recursive update rules are now given by

$$\begin{aligned} \alpha_k(s) &= \sum_{\forall s' \rightarrow \Delta d_k} \alpha_{k-1}(s')\gamma_k(s', s) & (13)\\ \beta_{k-1}(s') &= \sum_{\forall s \rightarrow \Delta d_k} \beta_k(s)\gamma_k(s', s) & (14) \end{aligned}$$

Finally the transitional probability in (11) is found by applying Bayes' rule

$$p(r_k|s', s, \mathbf{r}_0^{k-1}) = \frac{p(\mathbf{r}_0^k|s', s)}{p(\mathbf{r}_0^{k-1}|s', s)} \quad (15)$$

One could now be tempted to express this conditional probability in terms of (3) and, then, to apply it directly for the update of (13) and (14). Unfortunately, (15) faces a significant problem. The longer the sequences $\mathbf{r}_0^{k-1}$ becomes the more states have to be considered in order to cover all possible hypotheses, i.e., as is now (15) entails a time-variant trellis, probably with an exploding number of states. To avoid this problem a finite observation interval is introduced. This idea has been successfully applied e.g. in [2] and has its roots in multiple-symbol detection of M-DPSK [1]. This approach in fact leads to a receiver previously described by Hansson and Aulin [6] for minimum-shift keying and can basically be described by the following approximation

$$p(r_k|s', s, \mathbf{r}_0^{k-1}) \approx p(r_k|s', s, \mathbf{r}_{k-N+1}^{k-1}), \quad (16)$$

which consists of a new conditioning no longer on the total received sequence $\mathbf{r}_0^{k-1}$ but on the $N-1$ recently received symbols $\mathbf{r}_{k-N+1}^{k-1}$. Eq. (16) leads to a time-invariant trellis structure due to the constant observation interval and can be expressed by (3) and (15), (cf. [6, Eq.(12)])

$$p(r_k|s', s, \mathbf{r}_{k-N+1}^{k-1}) \propto \frac{\exp(-(\mathbf{r}_{k-N+1}^k)^{\mathrm{H}}\tilde{\mathbf{C}}_{rr}^{-1}[k]\mathbf{r}_{k-N+1}^k)}{\exp(-(\mathbf{r}_{k-N+1}^{k-1})^{\mathrm{H}}\tilde{\mathbf{C}}_{rr}^{-1}[k-1]\mathbf{r}_{k-N+1}^{k-1})}. \quad (17)$$

The correlation matrices are given by

$$\begin{aligned} \tilde{\mathbf{C}}_{rr}[k] &= \mathrm{E}\{\mathbf{r}_{k-N+1}^k(\mathbf{r}_{k-N+1}^k)^{\mathrm{H}}|\mathbf{d}_{k-N+1}^k\} & (18)\\ &= \mathcal{D}\{\mathbf{d}_{k-N+1}^k\}\tilde{\mathbf{C}}_{\mathrm{Num}}\mathcal{D}\{\mathbf{d}_{k-N+1}^k\}^{\mathrm{H}} & (19)\\ \tilde{\mathbf{C}}_{rr}[k-1] &= \mathrm{E}\{\mathbf{r}_{k-N+1}^{k-1}(\mathbf{r}_{k-N+1}^{k-1}|\mathbf{d}_{k-N+1}^{k-1})^{\mathrm{H}}\} & (20)\\ &= \mathcal{D}\{\mathbf{d}_{k-N+1}^{k-1}\}\tilde{\mathbf{C}}_{\mathrm{Den}}\mathcal{D}\{\mathbf{d}_{k-N+1}^{k-1}\}^{\mathrm{H}} & (21) \end{aligned}$$

The elements of the $(N \times N)$ - matrix $\tilde{\mathbf{C}}_{\mathrm{Num}} = \mathrm{E}\{\mathbf{h}_{k-N+1}^k(\mathbf{h}_{k-N+1}^k)^{\mathrm{H}}\} + \sigma_w^2\mathbf{I}_{N\times N}$ and the $(N-1 \times N-1)$ - matrix $\tilde{\mathbf{C}}_{\mathrm{Den}} = \mathrm{E}\{\mathbf{h}_{k-N+1}^{k-1}(\mathbf{h}_{k-N+1}^{k-1})^{\mathrm{H}}\} + \sigma_w^2\mathbf{I}_{N-1\times N-1}$ are given in (6). Note that both matrices are data- and time-independent. With the definition of

$$\mathbf{T} \triangleq -\tilde{\mathbf{C}}_{\mathrm{Num}}^{-1} + \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \tilde{\mathbf{C}}_{\mathrm{Den}}^{-1} & \\ 0 & & \end{pmatrix}, \quad (22)$$

where matrix $\tilde{\mathbf{C}}_{\mathrm{Den}}^{-1}$ in the bracketed matrix on the right hand side of (22) is meant to be enclosed by an additional upper row and an additional left hand column of zeros, Eq. (17) in the logarithmic domain can be simplified to

$$\begin{aligned} &\log p(r_k|s', s, \mathbf{r}_{k-N+1}^{k-1}) \propto \\ &(\mathbf{r}_{k-N+1}^k)^{\mathrm{H}}\mathcal{D}\{\mathbf{d}_{k-N+1}^k\}\mathbf{T}\mathcal{D}\{\mathbf{d}_{k-N+1}^k\}^{\mathrm{H}}\mathbf{r}_{k-N+1}^k. \quad (23) \end{aligned}$$

Please note the resemblance of (23) to [1, (10)] which provides the onset for the multiple-symbol detection metric for Rayleigh fading. With $t_{\mu\nu} \triangleq [\mathbf{T}]_{\mu\nu}$ Eq. (23) in scalar notation reads

$$\begin{aligned} &\log p(r_k|s', s, \mathbf{r}_{k-N+1}^{k-1}) \propto \\ &2 \cdot \mathrm{Re}\left(\sum_{\mu=0}^{N-1}\sum_{\nu=\mu+1}^{N-1} t_{\mu\nu} r_{k-\nu} r_{k-\mu}^* \prod_{\xi=\mu}^{\nu-1} \Delta\tilde{d}_{k-\xi}\right). \quad (24) \end{aligned}$$

The symbols $\Delta\tilde{d}_{k-\xi}, 0 \leq \xi \leq N-2$ represent the hypotheses for the state transition (s', s). Eq. (24) lies at the heart of an APP demodulator for M-DPSK and enables the computation of (11) which in turn allows for the recursive update of (13) in a forward sweep and of (14) in a backward sweep. Finally, (7) can be computed to yield the soft value of the code bits.

A. PSP-Extension

Eq. (24) is now extended to incorporate per-survivor processing. That is done by basically prolonging the observation interval from N symbols to $N+Z$ symbols. The transitional probability for this case reads

$$\log p(r_k|s', s, \mathbf{r}_{k-N-Z+1}^{k-1}) \propto \tag{25}$$
$$2 \cdot \mathrm{Re}\left(\sum_{\mu=0}^{N+Z-1} \sum_{\nu=\mu+1}^{N+Z-1} t_{\mu\nu} r_{k-\nu} r_{k-\mu}^{*} \prod_{\xi=\mu}^{\nu-1} \Delta\hat{\tilde{d}}_{k-\xi} \right)$$

Thereby we assumed that $\Delta\hat{\tilde{d}}_{k-\xi}$ is chosen from the set

$$\Delta\hat{\tilde{d}}_{k-\xi} \in \{\Delta\tilde{d}_k, \cdots, \Delta\tilde{d}_{k-N+2}, \Delta\hat{d}_{k-N+1}, \cdots, \Delta\hat{d}_{k-N-Z+1}\}, \tag{26}$$

which consists of the *true* hypotheses $\Delta\tilde{d}_k, \cdots, \Delta\tilde{d}_{k-N+2}$ which constitute the state transition (s', s) plus the survivors $\Delta\hat{d}_{k-N+1}, \cdots \Delta\hat{d}_{k-N-Z+1}$ which were associated with the state s'. Those survivors $s' = s_{\text{SUR}}$ are identified in the forward iteration by selecting that path which yields the largest incoming forward probability $\alpha_k(s)$ for a given state s.

$$s_{\text{SUR}} = \arg \max_{\forall s' \to s} \{\gamma_k(s', s)\alpha_{k-1}(s)\}, \tag{27}$$

i.e., instead of establishing new hypotheses for the Z additional symbols, the trial symbols are taken from the surviving path s_{SUR} which leads to the state under consideration s. An example is depicted in Fig. 2.

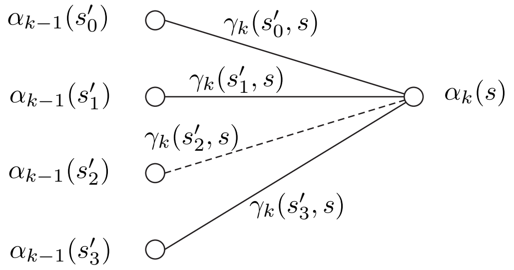

Fig. 2. Example for selecting a surviving path; the dashed line indicates the path with the largest path metric contribution, thus $s_{\text{SUR}} = s'_2$

IV. Simulation Results

A. Setup

Simulation results in terms of BER curves are presented for the time-varying Rayleigh flat-fading channel. Transmissions were simulated for a slowly time-varying channel with normalized maximum Doppler frequency $f_{\text{D,max}}T = 0.01$ and for a rapidly time-varying channel with $f_{\text{D,max}}T = 0.05$.

A block length of 10^4 information bits was assumed. For all results the $(133, 171)_8$ convolutional code with random interleaving was used. The signal alphabet was Q(D)PSK, where Gray and anti-Gray mapping were considered.

As a reference we simulated a DPSK transmission with perfect CSI, i.e., we interpreted the DPSK modulator as a recursive convolutional code of rate one followed by a PSK modulation with Gray mapping. The receiver with perfect CSI then performed an optimal APP demodulation of the coherent PSK upon which we employed a turbo detector which utilized the serially concatenation of convolutional code, interleaver and recursive code. 10 iterations were chosen. The performance of this receiver constitutes a lower bound for a receiver without CSI. The respective BER curves are labeled "perfect CSI".

Additionally we simulated the performance of conventional differentially detected QDPSK (Gray mapping) which is based on the correlation of two successive symbols. Decoding of the convolutional code is performed by the Viterbi algorithm. The respective curves are labeled "CDD".

Results for the turbo APP demodulator of Sec. III are then shown for a fixed observation interval of $N = 5$, which results for the 4-ary symbol alphabet into $M^{N-2} = 64$ states per trellis segment. For the receiver which utilizes PSP the observation interval was extended by three additional symbols, i.e., $Z = 3$. The results are depicted for three iterations. The curves for the noncoherent turbo receiver are labeled with "tMSDD, Gray" for the Gray mapping case and with "tMSDD, aGray" for the anti-Gray mapping case. Decoding is performed optimally by the BCJR-Log-Map algorithm.

B. Discussion

Let us first examine the performance of the receiver based on (24) without PSP. Fig. 3 shows results for a slowly time-varying channel.

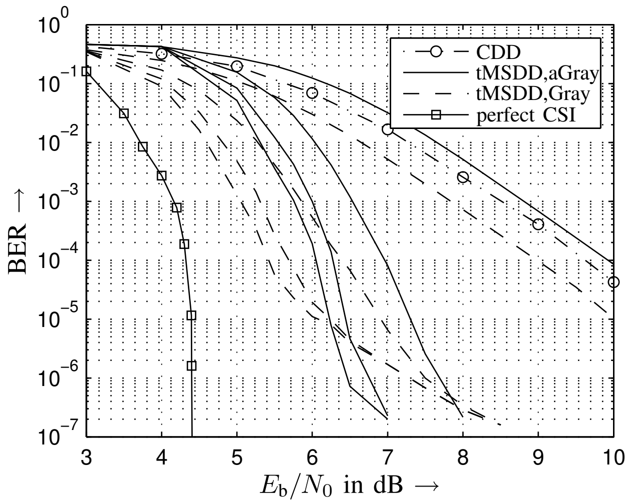

Fig. 3. Slowly time-varying channel with $f_{\text{D,max}}T = 0.01$, observation interval $N = 5$, no PSP

Generally, the curves exhibit the typical behavior of turbo codes, i.e., a water-fall region is followed by an ultimately ensuing error-floor. More specifically, the Gray mapping leads to a superior performance in the water-fall region. However, as soon as the error-floor becomes apparent the anti-Gray mapping leads to a better error performance.

For the rapidly time-varying channel in Fig. 4 the loss against the perfect CSI case becomes severe. The noncoherent turbo receiver with Gray mapping provides a small gain in the first iteration, whereas further iterations get basically stuck. Here, however, the anti-Gray mapping proves to be the more robust scheme at the more interesting higher SNRs. The gains for each iteration are larger and the error-floor is lower. Nevertheless, the overall bad performance calls for a remedy which is offered by the prolonged observation interval of the PSP approach.

In the following we will examine the performance of the receiver based on (25), i.e., per-survivor processing is now

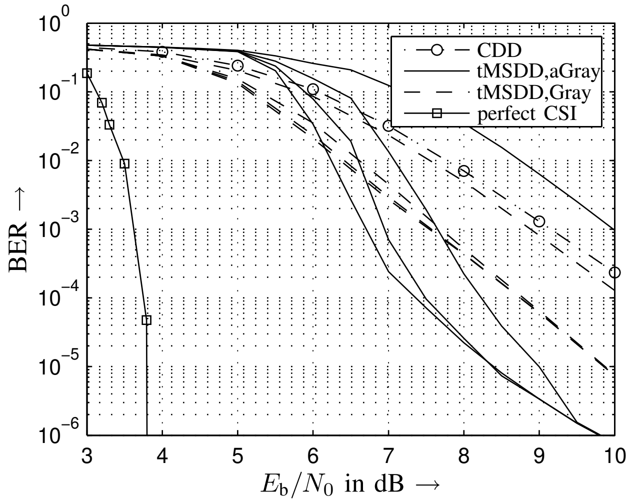

Fig. 4. Rapidly time-varying channel with $f_{\mathrm{D,max}}T = 0.05$, observation interval $N = 5$, no PSP

applied. Let us again start with the slowly time-varying channel in Fig. 5. The waterfall region compared to Fig. 3 is shifted closer to the perfect CSI and the error-floor is lowered. However, the improvements for the Gray mapping case as well as the anti-Gray mapping are fairly moderate against the non-PSP case.

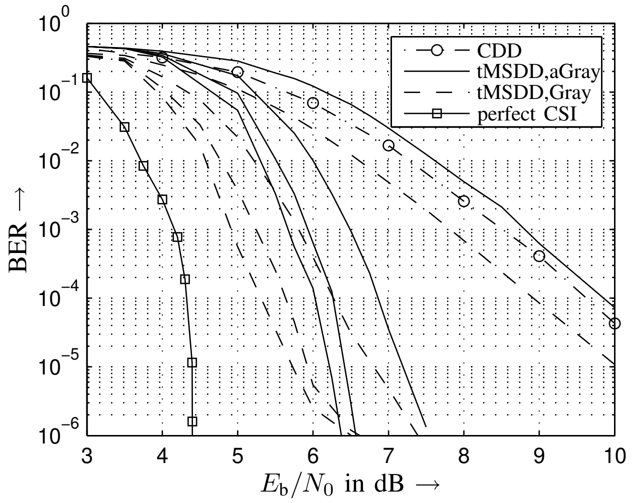

Fig. 5. Slowly time-varying channel with $f_{\mathrm{D,max}}T = 0.01$, observation interval $N = 5$, PSP with $Z = 3$

For the rapidly time-varying channel the improvements due to per-survivor processing become obvious. Both mappings benefit from the extended observation interval. But again the Gray mapping provides only small gains for more than one iteration and ends in a higher error-floor than the anti-Gray mapping, which is able to provide for much larger gains per iteration.

V. Conclusion

We have demonstrated that the APP-demodulator for MSK of Hansson and Aulin [6] can straightforwardly be applied to the noncoherent turbo reception of M-DPSK over Rayleigh fading channels. We have drawn the connection to multiple-symbol detection of Divsalar and Simon [1]. Owing to the fact that the overall detection performance improves with the length of the observation interval, which in turn increases the demodulator complexity in terms of number of states, we applied per-survivor processing to extend the observation interval while keeping the number of states constant. Thereby, we were able to improve the BER performance especially for rapidly fading channels. We additionally confirmed by simulations for QPSK that Gray mapping yields better performance in the waterfall region, whereas anti-Gray mapping proves to be superior in the error-floor region.

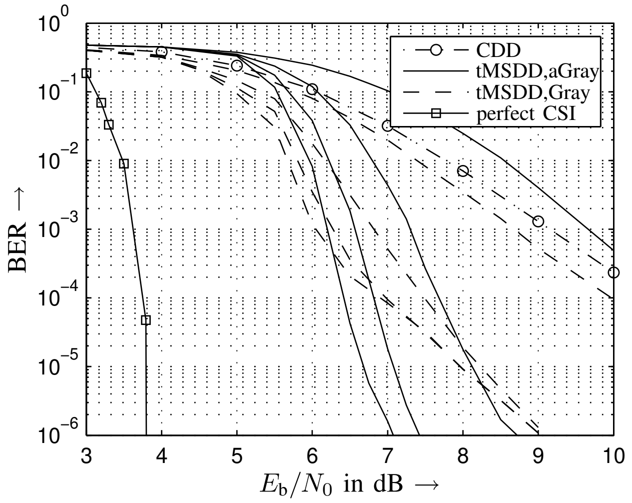

Fig. 6. Rapidly time-varying channel with $f_{\mathrm{D,max}}T = 0.05$, observation interval $N = 5$, PSP with $Z = 3$

References

[1] D. Divsalar, M. K. Simon, "Maximum-Likelihood Differential Detection of Uncoded and Trellis Coded Amplitude Phase Modulation Over AWGN and Fading Channels-Metrics and Performance", *IEEE Trans. Commun.*, vol. COM-42, pp. 76-89, January 1994

[2] I. D. Marsland, P. T. Mathiopoulos, "On the performance of iterative noncoherent detection of coded M-PSK signals", *IEEE Trans. Commun.*, vol. COM-48, pp. 588 - 596, April 2000.

[3] G. Colavolpe, G. Ferrari, R. Raheli, "Noncoherent iterative (turbo) decoding", *IEEE Trans. Commun.*, vol. COM-48, pp. 1488 - 1498, September 2000

[4] M. K. Howlader, X. Luo, "Noncoherent iterative demodulation and decoding of serially concatenated coded MSK", *Proc. IEEE GLOBECOM*, 2004, p. 2061, December 2004

[5] P. Hoeher, J. H. Lodge; "Turbo DPSK": Iterative Differential PSK Demodulation and Channel Decoding, *IEEE Trans. Commun.*, vol. COM-47, pp. 837-843, June 1999

[6] A. Hansson, T. Aulin; "Iterative diversity detection for correlated continuous-time Rayleigh fading channels", *IEEE Trans. Commun.*, vol. Com-51, pp. 240-246, Feb.2003

[7] R. Raheli, A. Polydoros, C.-K. Tzou, "Per-Survivor Processing: A General Approach to MLSE in Uncertain Environments"; *Trans. Comm*, vol. 43, pp. 354-364, Feb/Mar/April 1995

[8] L. Bahl, J. Cocke, F. Jelinek, J. Raviv; "Optimal decoding of linear codes for minimizing symbol error rate", *IEEE Trans. Inf.*, Mar 1974, pp. 284-287

[9] C. Berrou, A. Glavieux, P. Thitimajshima; "Near Shannon limit error-correcting codes and decoding", *Proc.ICC*, 1993, pp. 1064-1070

[10] S. ten Brink, J. Speidel, R.H. Yan; "Iterative demapping for QPSK modulation", *Electronic Letters*, July 1998, vol.34, pp. 1459-1460