Iterative MAP detection of coded M-DPSK signals in fading channels with application to IS-136 TDMA

Andrea M. Tonello

Wireless Technology Laboratory - Bell Labs - Lucent Technologies
67 Whippany Rd - Whippany - NJ - 07981 - email: tonello@lucent.com - tonello@dei.unipd.it

Abstract – **Maximum a posteriori (MAP) detection of M-DPSK modulated signals is proposed for both time-non-dispersive and time-dispersive fading channels. When the communication system comprises an inner convolutional encoder with a bit-interleaver, it is shown that an iterative receiver obtained from the concatenation of a MAP M-DPSK detector and a MAP convolutional decoder is capable of improving the performance in both channel environments. Such a receiver is based on the turbo decoding concept of refining the reliability information provided at the output of each soft-in/soft-out receiver stage, by running multiple detection/decoding iterations in a receiver with feedback from the outer stage to the inner stage. Furthermore, the general case of multiple receiving antennas is considered. Combining of the received signals is addressed in both an environment characterized by thermal noise (i.e. AWGN) uncorrelated through the antenna elements, and an environment characterized by spatially correlated noise such that arising in the presence of co-channel interference in a TDMA system (e.g. IS-136).**

Introduction

In this paper we consider a communication system comprising a convolutional encoder, an interleaver, and an M-DPSK modulator. Differential detection is conventionally applied for detection of differential encoded M-PSK signals (M-DPSK) propagating through noisy time-non-dispersive fading channels (i.e. flat fading) [1],[2]. However, differential detection exhibits a penalty compared to coherent detection of M-DPSK [3] in terms of signal-to-noise ratio needed to achieve a given bit error rate performance. Furthermore, coherent detection of M-DPSK signals is required when dealing with problems such as equalization in time-dispersive fading channels (i.e. multi-path fading) [2], or mitigation of co-channel interference in TDMA wireless systems (e.g. IS-136) [3],[4].

We propose optimum coherent detection of coded M-DPSK signals through the application of the MAP algorithm [5],[6],[7]. With this approach the resulting receiver is obtained from the concatenation of a MAP detector, a de-interleaver, and a MAP convolutional decoder. The first stage (i.e. MAP detector) calculates optimum soft outputs (i.e. a posteriori log-likelihood ratios or L-values) for the coded bits by exploiting the Markovian property of the M-DPSK signal. Such a calculation is performed in a unified manner for both time non-dispersive and time-dispersive fading channels. In addition, the output L-values can be refined by taking into account the a priori L-values of the coded bits that can be approximated with the extrinsic L-values computed by the second stage (i.e. convolutional decoder). The resulting feedback receiver structure is similar to the one originally deployed in the decoding of turbo codes [8], then proposed for equalization [9], and recently applied to iterative demapping of ad hoc mapped 8-PSK [10] and QPSK [11] signals.

In this work we show that iterative detection/decoding of bit-interleaved coded M-DPSK signals achieves 'turbo' gains in both flat and multi-path fading. Furthermore, we extend the algorithm to a multiple receiver antenna scenario, where optimal antenna combining allows for diversity gains and co-channel interference rejection capability. The proposed unified receiver is capable of iteratively refining the L-values at both the detector and the decoder outputs in both flat fading and multi-path fading, with or without the presence of co-channel interference. Finally, the algorithm is applied to the IS-136 TDMA system. Performance results are reported for the case of perfect knowledge of the channel state information. Impressive gains are found over both conventional differential detection and hard decision receivers in all scenarios that we considered.

1 System model

We consider a communication system model as in Figure 1. The information bit sequence, belonging to a given user, is represented by b_i, $i=0,\dots,N_b-1$. The information bits are first channel encoded (bits c_i, $i=0,\dots,N_c-1$) with a rate k/n convolutional code, not necessarily recursive-systematic, interleaved (bits d_i, $i=0,\dots,N_d-1$, $N_d=N_c$), and then M-DPSK modulated to produce the symbol sequence x_i, $i=0,\dots,N_k-1$ with $N_k=N_d/\log_2 M$. It should be noted that for ease of notation the same time indices i are used in each variable. The complex symbols x_i, transmitted at rate $1/T$, after pulse shaping, are RF modulated. This signal then propagates through either a time-non-dispersive or a time-dispersive noisy fading channel. Thus, the channel may introduce intersymbol interference. Co-channel interference signals, if present, undergo the same operations.

A receiver with an antenna array of N_a elements captures such signals (Figure 2). After RF demodulation and matched filtering, the base-band signals $y_a(t)$ at each antenna a are sampled at rate $1/T$. We model the sequence $y_a(kT)=y_{a,k}$ of N_k complex samples at each antenna a, with a discrete time transversal filter:

$$y_{a,k} = \sum_{l=0}^{N_t-1} h_{a,k,l} x_{k-l} + y_{a,k}^{\text{int}} + n_{a,k} \tag{1}$$

where N_t is the number of taps in the FIR model.

The author is currently with *DEI - Dipartimento di Elettronica ed Informatica - Universita' di Padova - Padova - Italy*, on leave from *Bell Labs Lucent Technologies*.

0-7803-5435-4/99/$10.00 ©1999 IEEE

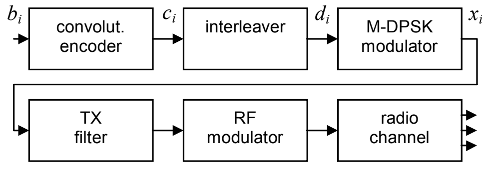

Figure 1: Communication system model.

The equivalent multiplicative channel state information (CSI) is represented by $h_{a,k,l}$. We model $h_{a,k,l}$ as statistically independent over index a complex Gaussian variables. Their amplitude is Rayleigh distributed, and their time correlation (i.e. over index k) is given by $\Omega_{a,l}J_0(2\pi f_d kT)$, where $f_d=vf_c/c$ ($J_0(.)$: zero-order Bessel function, v: mobile speed, f_c: carrier frequency, c: speed of light). The co-channel interference and AWG thermal noise are respectively given by $y_{a,k}^{\text{int}}$ and $n_{a,k}$ (overall by $w_{a,k}$). Equation (1) can be concisely rewritten in matrix form as:

$$\underline{\underline{y}} = \underline{\underline{h}}\,\underline{\underline{x}} + \underline{\underline{w}} \tag{2}$$

where $\underline{y}$ is a N_a x N_k matrix. In Figure 2, $\underline{y}_i$ represents a vector of N_a samples taken at time iT.

2 Iterative receiver structure

The receiver task is to determine the transmitted information bit sequence b_i. A two-stage receiver can accomplish such a goal. The first stage (detector) computes reliability information on the coded bits d_i from the channel observations $\underline{y}$. The reliability information is given by the log-likelihood ratios $L(d_i)=\ln[P(d_i=+1)/P(d_i=-1)]$ (referred to as L-values [6]). The second stage (decoder), after deinterleaving, performs decoding using the values calculated in the first stage.

We herein consider generation of optimum L-values in the first stage by application of the MAP algorithm [5]. The MAP algorithm can exploit the memory possessed by an M-DPSK signal that propagates through either a time-non-dispersive or a time-dispersive channel. Better L-values can be generated if the a priori probabilities of the coded bits are known. Such a priori probabilities can be estimated by the second stage (i.e. channel decoder) if a soft output decoder is deployed. The resulting receiver (Fig. 2) has a structure similar to the one originally proposed for decoding of turbo codes [8]. Thus, the detector computes a posteriori L-values for the coded bits $L(d_i|\underline{y})$ from only the channel observations, at the first iteration. These L-values are deinterleaved and fed to the decoder. The decoder (based on the MAP algorithm) computes L-values for both the information bits, $L(b_i)$, and for the coded bits, $L(c_i)$, by exploiting the memory of the convolutional code. The L-values $L(c_i)$ are used to derive an estimate of the a priori L-values of the coded bits $L_a(d_i)$ through the relation:

$$L_a(d_i) = \text{intlv}\left[L_e(c_i)\right] = \text{intlv}\left[L(c_i) - L_a(c_i)\right] \tag{3}$$

In equation (3) we attempt to minimize the correlation between the estimated a priori information and previous calculated information by feeding back only extrinsic information, $L_e(c_i)$ (i.e. incremental soft-decision information) [8]. The a priori L-values $L_a(d_i)$ can be used in a new detection iteration for refining the a posteriori L-values $L(d_i|\underline{y})$. Similarly, extrinsic information $L_e(d_i)$ is computed at the output of the detector and is fed to the decoder for a new decoding iteration. After the final iteration, hard decisions on $L(b_i)$ yield the decoded information sequence.

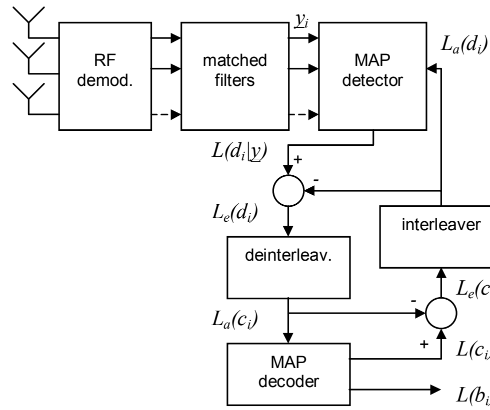

Figure 2: Iterative receiver with multiple antennas.

As it is shown in Section 5 multiple iterations can give significant performance gains. The steps involved in the MAP algorithm [5],[6],[7] are discussed in the next section. In particular, generation of a posteriori L-values $L(d_i|\underline{y})$ is addressed when the modulation is M-DPSK and the channel is either time-non-dispersive or time-dispersive in the presence of spatially correlated noise.

3 MAP detection of M-DPSK signals

An M-PSK differential encoder is a deterministic machine that given the previous state $S_{k-1} \in \{0,...,M-1\}$ and the input bits $d_{k,i}$, ($i=0,...,\log_2 M-1$) at signaling period kT, determines the output state S_k according to:

$$S_k = (S_{k-1} + \delta_k) \bmod M \tag{4}$$

where $\delta_k \in \{0,...,M-1\}$ is the phase transition index obtained by appropriate mapping (e.g. Gray) of a block of $N=\log_2 M$ input bits $d_{k,i}$. If an on-axis constellation is used the differential encoded symbol at time kT is given by $x_k = e^{j2\pi S_k/M}$. Extension to a π/M shifted M-DPSK modulation (e.g. π/4-DQPSK) is elementary.

Thus, an M-DPSK modulated signal is a discrete time finite state Markov process. The Markovian property is unaltered by propagation through either time-non-dispersive or time-dispersive channels. However, when the channel is time-dispersive (as in general expressed by the FIR model in (1)) the state at time kT is a vector $\underline{S}_k=[S_k\ S_{k-1}\ ...\ S_{k-N_t+2}]$ with $N_t>1$, whose time evolution is still derived from (4). With $N_t=1$ the state vector is a scalar $\underline{S}_k=[S_k]$.

At this point, from the sequence (block) of noisy channel observations $\underline{y}$ at the matched filters output, and given the channel state information $\underline{\underline{h}}$ (in practice estimated), we can use the MAP algorithm to determine the a posteriori log-likelihood ratios of the bits $d_{k,i}$. Following [5],[6],[7], the L-values of the bits $d_{k,i}$ are expressed by:

$$L(d_{k,i})=$$

$$=\ln\frac{P(d_{k,i}=+1|\underline{\underline{y}},\underline{\underline{h}})}{P(d_{k,i}=-1|\underline{\underline{y}},\underline{\underline{h}})}=\ln\frac{\sum_{(\underline{S}_k,\underline{S}_{k-1})\in D_i(+1)} p(\underline{S}_k,\underline{S}_{k-1},\underline{\underline{y}},\underline{\underline{h}})}{\sum_{(\underline{S}_k,\underline{S}_{k-1})\in D_i(-1)} p(\underline{S}_k,\underline{S}_{k-1},\underline{\underline{y}},\underline{\underline{h}})} \quad (5)$$

where $(\underline{S}_k, \underline{S}_{k-1})\in D_i(b)$ is the set of all possible state transitions determined by an input bit $d_{k,i}=b$ to the M-DPSK modulator. These two sets are constructed from relation (4). To proceed we make the assumption of a memoryless channel. Conditioned on the channel state information this assumption is true when $w_{a,k}$ is a sequence of independent (over index k) random variables. This condition is valid in the absence of co-channel interference (i.e. AWGN only), or in the presence of co-channel interference that is only spatially correlated (i.e. through the antenna elements). Then, the joint probabilities in equation (5) can be expressed as the product of three terms [6], a forward recursion, a transition probability, and a backward recursion:

$$p(\underline{S}_k,\underline{S}_{k-1},\underline{\underline{y}},\underline{\underline{h}})=\alpha_{k-1}(\underline{S}_{k-1})\gamma_k(\underline{S}_k,\underline{S}_{k-1})\beta_k(\underline{S}_k) \quad (6)$$

$$\alpha_k(\underline{S}_k)=\sum_{\underline{S}_{k-1}}\gamma_k(\underline{S}_k,\underline{S}_{k-1})\alpha_{k-1}(\underline{S}_{k-1}) \quad (7)$$

$$\beta_{k-1}(\underline{S}_{k-1})=\sum_{\underline{S}_k}\gamma_k(\underline{S}_k,\underline{S}_{k-1})\beta_k(\underline{S}_k) \quad (8)$$

The forward and backward recursions have to be initialized with initial and final known states. The transition probability is expressed as the product of a channel probability density function and the probability of transitioning between state $\underline{S}_{k-1}$ and state $\underline{S}_k$:

$$\gamma_k(\underline{S}_k,\underline{S}_{k-1})=p(\underline{y}_k,\underline{h}_k\,|\,\underline{S}_k,\underline{S}_{k-1})P(\underline{S}_k\,|\,\underline{S}_{k-1}) \quad (9)$$

where $\underline{y}_k$ and $\underline{h}_k$ are respectively a vector of N_a samples at time kT and the associated CSI. The transition probability, given the presence of the bit interleaver, is obtained from the L-values computed by the channel decoder (see also (3)) according to [6]:

$$P(\underline{S}_k\,|\,\underline{S}_{k-1})=A_k\, e^{\frac{1}{2}\sum_{i=0}^{N-1} d_{k,i}L_a(d_{k,i})} \quad (10)$$

where $d_{k,i}$ are the input bits to the M-DPSK modulator that are associated to the state transition $(\underline{S}_k, \underline{S}_{k-1})$, and A_k is a constant. The channel probability density function depends on the channel model assumption (flat fading, multi-path fading), and on the impairment $w_{a,k}$ statistics. The components of the N_a by 1 vector $\underline{w}_k$ are given by:

$$w_{a,k}=y_{a,k}-\sum_{l=0}^{N_t-1}h_{a,k,l}x_{k-l} \qquad ,a=0,...,N_a-1 \quad (11)$$

with x_{k-l} constellation symbols corresponding to the transition $(\underline{S}_k, \underline{S}_{k-1})$. Assuming $w_{a,k}$ to be a sequence of temporally uncorrelated, zero mean, Gaussian distributed random variables, with time varying spatial covariance matrix $\underline{\underline{R}}_{ww}(k)=E[\underline{w}_k\,\underline{w}_k^H]$, it follows that the channel probability density function is given by:

$$p(\underline{y}_k,\underline{h}_k\,|\,\underline{S}_k,\underline{S}_{k-1})=B_k\left|\underline{\underline{R}}_{ww}(k)\right|^{-1}e^{-\underline{w}_k^H\underline{\underline{R}}_{ww}^{-1}(k)\underline{w}_k} \quad (12)$$

where B_k is a constant and $|\underline{\underline{R}}|$ is the determinant of matrix $\underline{\underline{R}}$. In the absence of spatially correlated noise the covariance matrix is diagonal with elements $R_{a,a}=\sigma_a^2=E[|n_{a,k}|^2]$.

Typically, the MAP algorithm is applied in the log-domain by using the logarithm of expressions (6)-(12). Furthermore, a simplified implementation is achieved by applying the well-known dual maxima approximation leading to the Max-Log-MAP algorithm [7].

Finally, from the output L-values in equation (5) extrinsic L-values are derived. These L-values are de-interleaved and fed to the convolutional decoder where they are used in the transition metrics computations. The convolutional decoder applies the MAP algorithm following similar steps [6],[7]. The MAP decoder computes not only the L-values of the information bits, but also the L-values of the coded bits.

4 Application to IS-136 TDMA

In this section, we investigate the applicability of the suggested receiver to the IS-136 TDMA system [12]. The North American TDMA digital traffic channel (DTC) uses π/4-DQPSK modulation with Gray mapping, and a memory 5, rate ½ (with 8 bits puncturing) tail terminated convolutional code. The coding scheme with the ACELP vocoder is shown in Figure 3. We depict in Figure 4 the operations that take place every 20 ms.

A speech frame $\underline{b}$ of 148 information bits is partitioned into 3 sub-blocks or classes (see also Fig. 3): 52 C2 bits, 48 C1B bits, and 48 C1A bits. After adding 7 CRC bits to C1A, the C1A+C1B=C1 bits are convolutional encoded, while C2 bits are left uncoded. This generates a block $\underline{c}$ of 260 bits that are first reordered with a 26 by 10 matrix interleaver and then chain interleaved across 2 slots (20 ms apart). Thus, at a given transmission time interval, the 130 bits $\underline{d1}$ belong to the previous encoded speech frame, while the 130 bits $\underline{d2}$ belong to the current encoded speech frame. These bits together with other non-data bits, are slot formatted (Figure 5, considering the uplink) and modulated in order to generate blocks of symbols $\underline{x}$. The symbols are transmitted at rate 24.3 kbauds. Pulse shaping is obtained with a root raised cosine filter with roll-off=0.35.

As a result of chain interleaving, the input L-values $L_a(\underline{c})$ to the decoder are obtained from MAP detection of at least two slots. In turn, only partial amounts of soft information can be fed back from a given decoder to a given detector. Even though it is not the only solution, our receiver implementation (Figure 4) requires MAP detection over two consecutive received slots (20 ms apart), and MAP convolutional decoding over one encoded speech frame per iteration. Since only C1 bits are convolutionally encoded, only the a priori L-values of coded C1 bits are fed back from a given decoder. Thus, at the first iteration MAP detection over a slot takes place setting to zero the a priori L-values of coded C1 and C2 bits. In successive iterations, MAP detection over a slot is run using the observations and the a priori L-values of coded C1 bits generated from the decoder outputs (according to (3)).

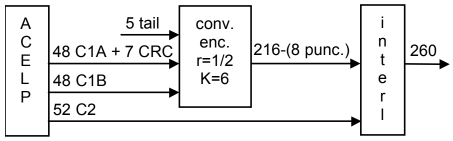

Figure 3: Coding scheme in IS-136 DTC.

It should be noted that the feedback receiver in Figure 4 doesn't introduce any extra processing delay, as otherwise would be the case if we increased the number of slots and/or encoded speech frames processed at each iteration. This would allow augmentation of the amount of feedback information to each processed slot.

Finally, simulations show that the proposed iterative receiver achieves significant BER gains in all classes, both C1 and C2, although C2 bits are not convolutionally encoded (see next section).

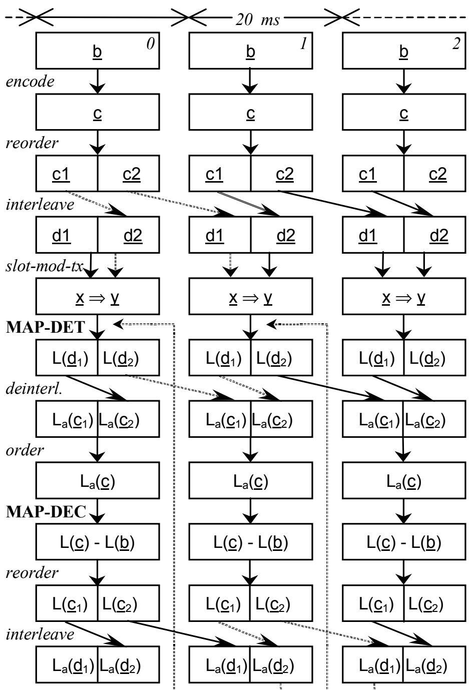

Figure 4: IS-136 processing with feedback.

gt	rt	data	sync	data	sacch	cdvcc	data
6	6	16	28	122	12	12	122

Figure 5: Slot format in IS-136 TDMA uplink.

5 Performance results in IS-136

Performance was evaluated through simulations in an IS-136 compliant system using the receiver structure in Figure 4. The Max-Log-MAP approximation was used in both the detector and the decoder implementation. We report error statistics collected over 10^4 slots, in terms of BER on decoded C1 bits and BER on C2 bits (uncoded). We consider both a flat fading environment and a multi-path fading environment assuming a 2-ray model. In this case, the rays are independently faded, with unity total power split equally, and with a relative time delay equal to 1 symbol (i.e. $T=41.2\ \mu s$). The receiver filters are root-raised cosine with roll-off 0.35. The sampling time and synchronization are assumed to be perfect.

First, we report in Figure 6 a comparison between the proposed MAP receiver and conventional differential detection with antenna combining (DD) [1] in flat fading without co-channel interference. The outputs from DD are used to generate soft inputs to the MAP decoder.

We assume a one tap parametric model in the MAP detector ($N_t=1$), and perfect knowledge of the CSI and of the noise variance. Both the single diversity ($N_a=1$) and the dual diversity ($N_a=2$) performances are reported. The fading is independent over antenna elements, with a maximum Doppler frequency equal to 100 Hz. The performance improvements over DD with Dopplers in the range 10-200 Hz do not differ significantly. The results are shown as a function of the signal-to-noise power ratio (C/N) observed at each antenna element (solid lines are for DD while dashed lines are for MAP). The performance of the proposed iterative MAP receiver is measured after the first pass through the decoder (i1), the second (i2), and the fourth (i4) pass.

As it can be seen in Figure 6, MAP gives a substantial gain over DD in both C2 BER and C1 BER. For C2 BER with one iteration (i1), there is up to 1.7 dB gain with $N_a=1$ and up to 2.2 dB gain with $N_a=2$, for the SNRs considered. Another iteration (i2) adds 2.2 dB gain with $N_a=1$ and 1.7 dB with $N_a=2$. More iterations do not yield improvement, and in fact the C2 BER curve with 4 iterations differs from the curve with 2 iterations only for low SNRs. Although C2 bits are not encoded, the associated iterative gain is due to the ability of the MAP detector of improving its outputs when a priori information for the coded C1 bits is fed back. With antenna diversity the gain over DD at the first iteration is larger than with a single antenna. The second iteration exploits the time diversity and yields a larger gain with a single antenna. Now, looking at C1 BER the gain over DD with $N_a=2$ is about 1.5 dB with 1 iteration and up to 2.2 dB with 4 iterations. With $N_a=1$ the overall gain over DD is in the range 1-1.9 dB for the SNRs considered.

Now, consider Figure 7 where performance in 2-ray fading is reported with single and dual diversity. For this case, a two taps parametric model ($N_t=2$), CSI, and noise variance knowledge are assumed. Given the high delay spread considered, DD performance is unacceptable. Equalization is required in this environment. The benefit from multiple iterations is clear. However, no advantage is found in C2 BER by running more than 2 iterations.

In order to show the benefit in C1 BER obtained by generating soft outputs (contrary to typical hard output maximum likelihood sequence estimators, MLSE), the proposed MAP receiver is compared to itself when the outputs at the first iteration are hard quantized (HMAP). Soft output MAP outperforms its hard output counterpart by more than 3 dB, with both single and dual diversity.

Finally, in Figure 8 we consider an environment with the presence of one TDMA co-channel interferer, which is flat faded and symbol synchronized with the desired user. The channel of the desired user is either flat or 2-ray. Double diversity with Doppler equal to 10 Hz is considered. Although we assumed knowledge of the desired user CSI, we estimated the impairment covariance matrix using an exponential forgetting

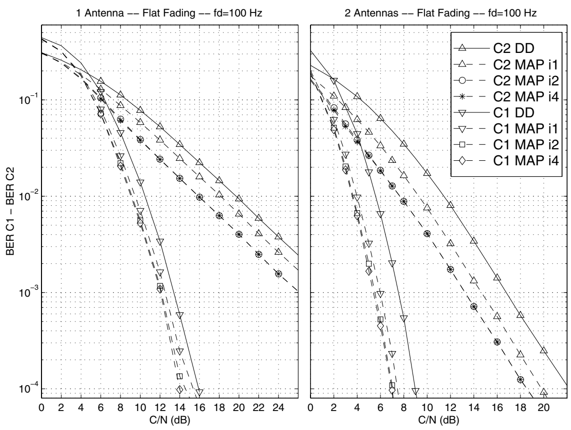

Figure 6: MAP versus DD in Flat Fading with $N_a=1,2$ ($C/I=\infty$).

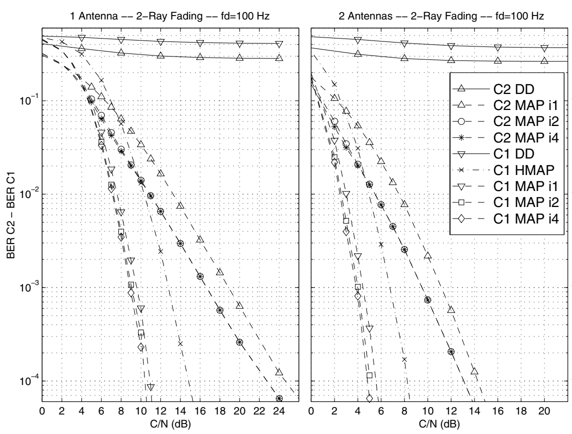

Figure 7: MAP versus DD and hard output MAP in 2-Ray Fading with $N_a=1,2$ ($C/I=\infty$).

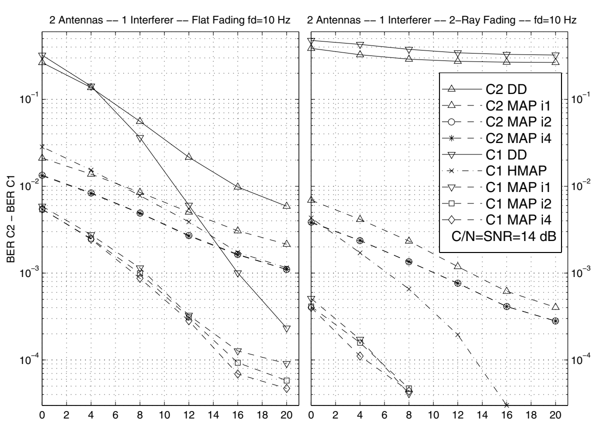

Figure 8: MAP versus DD and hard output MAP in Flat and 2-Ray Fading with $N_a=2$ and one TDMA co-channel interferer.

function. Training of the covariance matrix was first performed over known fields in the slot (*sync*, *cdvcc*). Then, tracking over the data fields was run. We report performance as a function of the signal-to-interference power ratio (C/I), for a fixed signal-to-noise ratio of 14 dB. Again, the proposed MAP receiver outperforms DD in both C2 and C1 BER. The gain in C2 BER, in flat fading, is more than 10 dB with 1 iteration, and more than 14 dB with 2 iterations. In C1 BER the gain is more than 6 dB with just 1 iteration. The performance in multi-path fading is greatly improved due to the diversity introduced by the second ray. The comparison between MAP and HMAP yields more than 8 dB gain in C1 BER in both flat and multipath fading. This final comparison clearly shows the superiority of the proposed soft output MAP receiver to interference cancellation receivers that provide hard outputs (e.g. based on hard output MLSE).

6 Conclusions

A unified approach for detection/decoding of bit-interleaved coded M-DPSK signals in time-non-dispersive and dispersive fading channels is proposed. The approach exploits the memory possessed by M-DPSK signals, and is based on iterative (turbo) MAP detection/decoding with optimum antenna combining in the presence of spatially uncorrelated or correlated noise (e.g. co-channel interference). Contrary to typical hard output MLSE based receivers, the proposed detector is capable of providing optimal soft outputs from which the outer decoder greatly benefits. Simulations with ideal CSI knowledge show that in IS-136 TDMA the proposed receiver outperforms both its hard output counterpart and conventional soft output differential detection in all the scenarios that we considered. The receiver achieves most of the gain with 2 iterations, while obtains diminishing returns with more iterations. The weak interleaving scheme in IS-136 probably limits the potential performance benefit of the proposed receiver. Although not reported, further simulations that incorporate practical channel estimation show performance advantages that are consistent with the ideal case.

Acknowledgments- The author would like to acknowledge James Seymour for the valuable discussions, and to thank Stephen Allpress, Quinn Li, Dirck Uptegrove for reviewing and helping refine this presentation.

References

[1] J. G. Proakis, ‘Digital communications’, NY: McGraw-Hill, 1995.

[2] A. Bahai, M.V. Clark, *et al.*, ’Link reliability for IS54/136 handsets with different receiver structures’, *IEEE Tr. Veh. Tech., January 1999, pp.213-223*.

[3] J.H. Winters, ‘Signal acquisition and tracking with adaptive arrays in the digital mobile radio system IS-54 with flat fading’, *IEEE Tr. Veh. Tech., November 1993, pp.377-384*.

[4] K.J. Molnar, G.E. Bottomley, ‘Adaptive array processing MLSE receivers for TDMA digital cellular/PCS communications’, *IEEE JSAC, October 1998, pp.1340-1351*.

[5] L.R. Bahl, J. Cocke, F. Jelinek, J. Raviv, ‘Optimal decoding of linear codes for minimizing symbol error rate*’. IEEE Tr. Info. Theory, March 1974*, pp. 284-287.

[6] J. Hagenauer, E. Offer, L. Papke, ‘Iterative decoding of binary block and convolutional codes’, *IEEE Tr. Info Theory, March 1996*, pp. 429-445.

[7] A.J. Viterbi, ‘An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes’, *IEEE JSAC February 1998*, pp. 260-264.

[8] C. Berrou, A. Glavieux, P. Thitimajshima, ‘Near Shannon limit error correcting coding and decoding: turbo codes’. *ICC 1993*, pp. 1064-1070.

[9] C. Douillard, A. Picart, *et al.* ‘Iterative correction of intersymbol interference: Turbo-equalizat.’, *ETT, Sep./Oct. 1995, pp. 507-511*.

[10] X. Li, J.A. Ritcey, ‘Bit-interleaved coded modulation with iterative decoding using soft feedback’, *IEE Elect. Letters, May 1998, pp. 942-943*.

[11] S. ten Brink, J. Speidel, R.H. Yan, ‘Iterative demapping for QPSK modulation’, *IEE Elect. Letters, July 1998, pp. 1459-1460* .

[12] TIA/EIA IS-136.2 standard, December 1994.