Reduced-Complexity Noncoherent Communication with Differential QAM and Iterative Receiver Processing

N. Jacobsen and U. Madhow[1]
ECE Department
UC Santa Barbara
Santa Barbara, CA 93106
{jacobsen,madhow}@ece.ucsb.edu

Abstract —

This work investigates a noncoherent communication system employing Differential Quadrature Amplitude Modulation (QAM), serially concatenated with a binary convolutional code. Reduced-complexity methods for iterative noncoherent demodulation and decoding are developed and evaluated for two blockwise constant channel models, the phase-noisy AWGN and the Rayleigh fading channels. We present a simple averaging estimate of the unknown channel amplitude based on conventional differential detection statistics, which is key to efficient block noncoherent demodulation for QAM alphabets. The unknown channel phase is quantized over several phase bins. Maximum a posteriori probability (MAP) block demodulation is performed based on the estimated amplitude and quantized phase. A Bayesian combination of the outputs of the parallel phase branches is passed up to the decoder. Significant complexity reduction is obtained by pruning the number of parallel phase branches down to two based on a log-likelihood ratio metric that incorporates feedback from the decoder. The reduced-complexity receiver comes within 2.4 dB of capacity for the block Rayleigh fading channel for signaling at rate .675 bits/symbol using 8-QAM.

I. Introduction

The standard approach to estimation and tracking of time-varying wireless channels is to employ pilots. There are two main drawbacks of this approach: the pilot overhead required to accurately track a rapidly varying channel may be excessive, and a channel estimate based purely on the pilot is suboptimal, since it does not exploit the bulk of the transmitted energy, which is in the data. A number of recent papers [1, 2, 3, 4] consider an alternative approach to wireless transceiver design, in which the channel and data are estimated jointly without necessarily using pilots. Such *noncoherent* techniques are particularly efficient when the channel can be approximated as constant over several symbols. In this case, block differential demodulation is known to approach coherent (i.e., channel known at the receiver) performance for uncoded systems [5, 6]. For coded systems, iterative information exchange between a decoder and a block noncoherent demodulator is shown [4] to approach the capacity, computed in [7], of the block fading channel. Earlier work on iterative demodulation and decoding includes [1, 2, 3].

Most practical communication strategies over fading channels, including the constructive encoding and decoding strategies referenced above, employ PSK alphabets. However, it is known [4] that, for a given constellation size, amplitude/phase modulation is more power-efficient at moderate to high SNR for noncoherent communication, just as it is for coherent communication over the classical AWGN channel. Our current work therefore investigates noncoherent systems based on QAM-like alphabets. Serial concatenation of an outer binary code with differential QAM (using an obvious generalization of differential PSK) is employed, together with iterative demodulation and decoding as in prior work. However, a key new technical issue that must be solved is that of obtaining, at reasonable complexity, amplitude estimates that are sufficiently accurate for QAM demodulation. We show that this can be achieved by a bootstrap stage, using decisions from conventional differential demodulation based on two consecutive symbols. This stage also yields initial soft decisions to be fed up to the decoder. The unknown phase is handled by quantizing it into a number of bins as in [4, 6, 8], running Maximum A Posteriori Probability (MAP) demodulators in parallel for each phase bin, and combining the outputs in Bayesian fashion. However, in contrast to earlier work, a phase arbitration mechanism based on a log-likelihood ratio (LLR) criterion is used to reduce the number of phase bins per block to two after the first iteration. With this simplification, the complexity of noncoherent block demodulation begins approaching the benchmark of (idealized) coherent MAP demodulation.

We denote scalar random variables with capitals (e.g. X), using lower case (e.g. x) for their realization. Likewise, random vectors, such as $\mathbf{W}$, will appear boldface, and a lower case bold $\mathbf{w}$ denotes a vector of deterministic values. The notation $\mathbf{W} \sim \mathcal{CN}(0, K)$ is used for a vector of circularly symmetric complex Gaussian random variables of mean zero and covariance matrix $E[\mathbf{W}\mathbf{W}^H] = K$, in particular,

$$f_{\mathbf{W}}(\mathbf{w}) = \frac{1}{\det(\pi K)} \exp(-\mathbf{w}^H K^{-1} \mathbf{w}) \tag{1}$$

where ^{H} is the conjugate transpose operator.

II. System model

Figure 1 depicts the baseband transmitter and channel model. The information symbol sequence $\mathbf{u}$ is coded and interleaved, producing code word $\mathbf{c}$. Consider the two ring 8-QAM constellation and differential encoding of Figure 2. The most significant bit of a 3-bit code symbol $\mathbf{c}_t = \{c_0^t c_1^t c_2^t\}$ affects amplitude transition between the $(t-1)^{th}$ and t^{th} transmitted

[1]This work was supported by Motorola, with matching funds from the University of California Cooperative Research Program under a Communications Research (CoRe) Grant

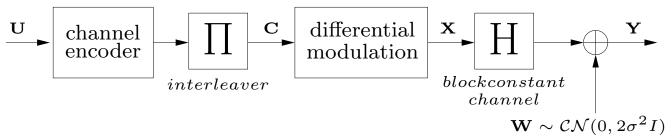

Figure 1: Transmitter and channel

8-QAM symbol. Phase transitions are given by a Gray encoding of the remaining two bits. Letting g denote the differential mapping, the t^{th} transmitted symbol is $x_t = g(\mathbf{c}_t, x_{t-1})$. This bit-to-symbol mapping is characterized by modulo-$\frac{\pi}{2}$ rotational invariance, with respect to differential detection. As such, we need only to recover the unknown channel phase in $[0, \frac{\pi}{2})$.

The block constant channel model is $\mathbf{Y} = H\mathbf{x} + \mathbf{W}$, where $\mathbf{Y}$ and $\mathbf{x}$ are one block of block of transmitted and received symbols, $H = Ae^{j\theta}$ is the unknown channel and $\mathbf{W}$ is complex additive white Gaussian noise, variance $2\sigma^2 I$. For the phase-noisy AWGN channel, we set $A = 1$, with θ uniformly distributed on $[0, 2\pi]$. For the block Rayleigh fading model, $H \sim \mathcal{CN}(0, 1)$; equivalently, A is Rayleigh, θ is $\mathcal{U}[0, 2\pi]$, and A, θ, are independent.

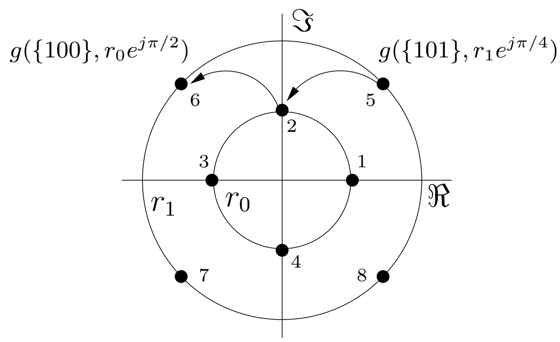

Figure 2: Differential 8-QAM signaling

III. MAP Demodulation

Our methods rely on block-wise MAP demodulation of the differentially encoded QAM data. Viewing g as a unit rate/memory recursive convolutional code, differential encoding can be represented graphically with the trellis of Figure 3. Associated with each trellis edge e are an initial and final state, input code symbol $\mathbf{c}_t(e)$, and output channel symbol $x_t(e)$. Coherent demodulation is performed with the BCJR algorithm. The *a posteriori* log-likelihood of the t^{th} code symbol is given by (2)

$$\lambda_t(\mathbf{c}|h) = \max_{e:\mathbf{c}_t(e)=\mathbf{c}}^{\star} \left\{ \alpha_{t-1}(s^I(e)) + \gamma_t(e|h) + \beta_t(s^F(e)) \right\} \quad (2)$$

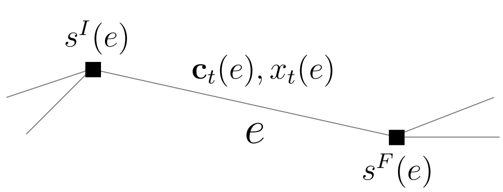

Figure 3: Trellis section of block demodulator

where $\alpha_t(s)$ and $\beta_t(s)$ are computed via the standard forwards/backwards recursions [9, 10], and $\max^\star(x, y) = \log(e^x + e^y)$. The branch metric $\gamma_t(e|h)$ of edge e is as in (3).

$$\gamma_t(e|h) = \lambda_t^I(\mathbf{c}_t(e)) + \frac{1}{\sigma^2} \Re < y_t, hx_t(e) > \quad (3)$$

Prior differential symbol probabilities $\lambda_t^I(\mathbf{c})$ are initially zero, and then set by the outer decoder in an iterative receiver.

IV. Channel amplitude estimate

We present an amplitude estimate that is well matched to ring type QAM constellations. Presented here for the case of 2-ring, 8-ary signaling, the estimate generalizes easily to larger constellations. The complexity is linear in the product of the number of rings and the alphabet size.

We employ conventional two-symbol differential detection to obtain likelihoods for the transmitted amplitude levels, which are then used to obtain an averaged amplitude estimator for a block of symbols. Note that this bootstrap stage can be skipped when estimating the channel amplitude for constant-amplitude PSK signaling [4]. The vector of two received symbols is (4)

$$\mathbf{y}_t \equiv \begin{bmatrix} y_{t-1} \\ y_t \end{bmatrix} = h \begin{bmatrix} R_{t-1} \\ R_t e^{j\Phi_t} \end{bmatrix} + \begin{bmatrix} w_{t-1} \\ w_t \end{bmatrix} = h\mathbf{x}_t + \mathbf{w}_t \quad (4)$$

where $R_t \in \{r_0, r_1\}$ is the modulus of the t^{th} transmitted symbol and Φ_t the differential phase. The phase of the $(t-1)^{th}$ symbol has been factored into the channel, without changing its circularly symmetric density. The noncoherent *a priori* density (less constant terms) is (5).

$$\log f_{\mathbf{x}}(\mathbf{y}_t) = \frac{1}{2\sigma^2} \frac{|\mathbf{y}_t^H \mathbf{x}|^2}{\|\mathbf{x}\|^2 + 2\sigma^2} - \log(\|\mathbf{x}\|^2 + 2\sigma^2) \quad (5)$$

$$\pi_r^t \equiv Pr(R_t = r|\mathbf{y}_t) \propto \sum_{\mathbf{x}:R_t=r} f_{\mathbf{x}}(\mathbf{y}_t) \quad (6)$$

Denoting the probability of the t^{th} symbol amplitude R_t as in (6), the fading gain estimate $\hat{A}$ is computed with (7) where T denotes the block length, or coherence interval of the channel.

$$\hat{A}^2 = \max \left\{ 0, \frac{\mathbf{y}^H \mathbf{y} - 2T\sigma^2}{\sum_{t=0}^{T-1} \sum_{r=r_0,r_1} \pi_r^t r^2} \right\} \quad (7)$$

V. Phase quantization and iterative reception

Direct implementation of MAP block noncoherent demodulation is computationally infeasible, because the noncoherent *a posteriori* density does not factor into a product of T i.i.d. terms, so that the BCJR algorithm cannot be directly applied. However, we can approximate MAP noncoherent demodulation using the coherent BCJR algorithm as a building block, by plugging in the amplitude estimate of (7), and Q-level quantization of the unknown channel phase θ in the range $[0, \frac{\pi}{2})$. For each quantization branch $q \in \mathcal{Q} = \{0, 1, \ldots, Q-1\}$ we can use (2) and (3) with $h = \hat{A}\exp(\frac{j\pi q}{2Q})$. The resulting

$$\lambda_t^O(\mathbf{c}) = \max_{q\in\mathcal{Q}}^{\star} \left\{ \lambda_t(\mathbf{c}|\hat{A}\exp\left(\frac{j\pi q}{2Q}\right)) \right\} - \lambda_t^I(\mathbf{c}) \quad (8)$$

code symbol likelihoods are then averaged (8) to yield extrinsic likelihoods of an approximation to the true noncoherent density. See [4] for the details of this approximation. Figure 4

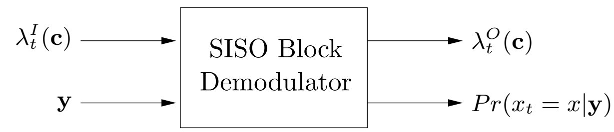

Figure 4: Module for iterative demodulation/decoding

depicts the soft-input soft-output noncoherent demodulation module that represents (8) and enables turbo-like processing between demodulator and channel decoder.

Joint noncoherent demodulation and decoding is illustrated in Figure 5. First, a conventional two symbol detector is used to noncoherently demodulate the received data sequence. We refer to this step as bootstrapping the receiver, as it provides the initial estimate of code symbol likelihoods. These likelihoods are (i) used for channel amplitude estimation and (ii) passed directly to the outer SISO convolutional decoder. The channel decoder computes code symbol extrinsic likelihoods which are subsequently fed back to the demodulator. The SISO demodulation module performs block demodulation of the received differential data sequence using parallel coherent BCJR algorithms, as described previously. Demodulation and decoding are performed iteratively until the termination criterion is satisfied.

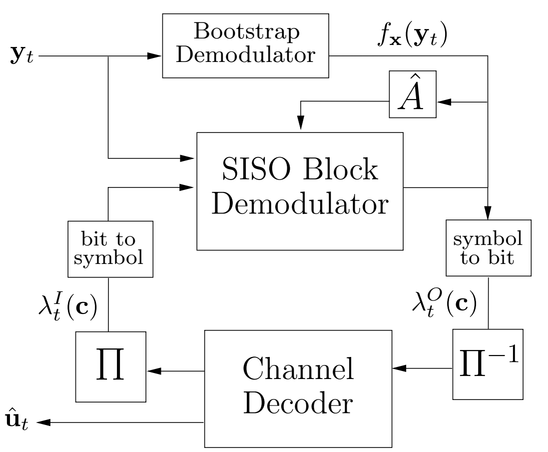

Figure 5: Iterative receiver structure

VI. Parallel phase branch arbitration

Although the method of phase quantization demonstrates near capacity performance on the block fading channel, each phase branch requires its own BCJR computation per iteration. Such a receiver requires Q times as many computations as a coherent communication system employing the same number of iterations. However, experiments show that a genie-based system which uses only the branch with phase closest to the true channel phase provides excellent performance. This motivates the development of a criterion for ranking and pruning parallel phase branches as iterative demodulation and decoding are performed.

For each block, we associate a metric for each phase branch which measures the reliability of the soft decisions output by that branch. The idea is to keep the phase branches which are the most reliable. In particular, we employ the average (or equivalently, the sum) of the code bit LLR magnitudes as a

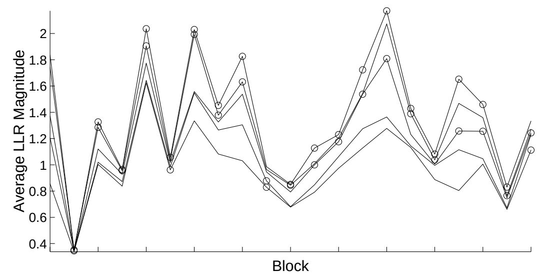

Figure 6: Average LLR magnitude for all phase branches, with two closest noted

reliability measure. Other reliability measures such as the sum of the binary entropies of the code bits could also be used, but the LLR-based reliability metric leverages the computations already performed by the BCJR algorithm. Figure 6 plots typical values of the average LLR magnitude for a sequence of blocks. The two "best" phase branches, i.e., the ones closest to the true phase, are marked by circles. Note that, for most blocks, at least one of these two branches is the one with the largest metric. Furthermore, for blocks with larger metrics, the two best branches are also the ones with the largest metric values. This leads to a pruning scheme which chooses the two branches with the largest average LLR values, as described below.

Our proposed scheme is as follows. Bootstrap the receiver with conventional two symbol detection as before, to obtain amplitude estimates and soft decisions to be passed up to the decoder. The decoder performs one iteration, and provides extrinsic information back to the demodulator. Now, perform parallel block demodulation with *all* phase branches, and compute the extrinsic code bit likelihoods for each phase branch. Use the average code bit LLR magnitude criterion for eliminating all but two phase branches on each block. Since $\lambda_t(\mathbf{c}|\hat{A}\exp(\frac{j\pi q}{2Q}))$ are already known, the receiver will still compute (8). All subsequent demodulations will only consider the selected subset of two phase branches. As such, after the first iteration of full phase quantization, the arbitrated scheme requires only twice as many demodulation computations as its coherent equivalent.

VII. Results

We first present simulation results for the phase-noisy AWGN channel. The information bits are encoded with the $G = [20\ 25\ 27\ 33]$ rate-$\frac{1}{4}$ non-recursive convolutional code. Codeword length is 64,000 bits. We have chosen a moderate coherence interval, $T = 10$, for the blockwise constant channel. The information rate of the coded 8-QAM system is then $\frac{3(T-1)}{4T}$ since the first symbol of each block serves as reference. Simulations have shown that $Q = 5$ quantization levels in $[0, \frac{\pi}{2})$ are sufficient to achieve averaging gains in the noncoherent system. Figure 7 demonstrates the performance of the iterative receiver on a phase-noisy AWGN channel, after 20 iterations. The reduced-complexity noncoherent receiver is less than .9 dB from the idealized benchmark of coherent reception (i.e., known amplitude and phase, unrealizable for our channel model) for the same code. This gap can be roughly decomposed as follows: about .5 dB for unknown phase (i.e., for using parallel phase branches instead of the true phase), less than .2 dB for unknown amplitude (i.e., for using the amplitude estimate instead of the true amplitude), and about .2 dB for complexity reduction by pruning the number of phase branches to two.

Figure 8 provides results for the Rayleigh block fading channel with same system parameters as above. Again, the reduced-complexity noncoherent receiver is about .9 dB from (unrealizable) coherent reception of the same code, with about .3 dB for unknown phase, .5 dB for unknown amplitude, and about .1 dB for phase branch pruning. Indeed, the reduced complexity receiver is only about .3 dB from a genie-based receiver which employs the estimated amplitude and the quantized phase closest to the true phase. Note that the loss due to amplitude estimation is larger for the Rayleigh fading channel, possibly because the contribution of too many blocks to the likelihood computations is being zeroed out (erased) when the amplitude estimator is set to zero. Addressing this issue is an important topic for future work, especially when we consider larger constellations. Note that the noncoherent block Rayleigh fading capacity benchmark for 8-QAM at this information rate is 2.7 dB [4], so that our reduced complexity noncoherent receiver is about 2.4 dB away from capacity, with 10^{-4} taken as the reliable communication BER threshold. Since even the unrealizable coherent benchmark for the given code is 1.6 dB away from this, we can attribute at least 1.6 dB of this gap to code construction. Part of the remaining .8 dB could potentially be closed by improving the noncoherent demodulation strategy.

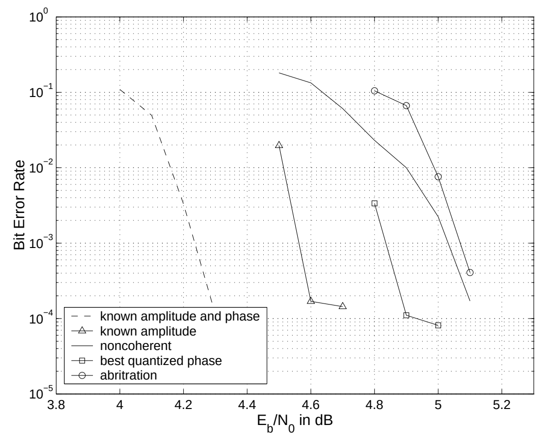

Figure 8: Rayleigh block fading results

VIII. Conclusions

The results in this paper suggest the feasibility of efficient noncoherent communication at moderate to large SNR using QAM alphabets. For an M-ary constellation consisting of n PSK rings, the complexity of the simple amplitude estimator proposed here scales as nM, so that the use of larger constellations is computationally tractable. A natural application of these techniques is to Orthogonal Frequency Division Multiplexed (OFDM) systems, for which the channel seen by a contiguous time-frequency block is often well modeled as a constant complex scalar. Our experience suggests the following design strategy for approaching capacity in noncoherent systems, by addressing the issues of encoding and decoding separately. First, we can obtain a lower bound on the gap due to *encoding* by comparing the unrealizable coherent performance of the code with noncoherent capacity. This quickly indicates how much effort must be expended in refining code constructions, in a manner that is decoupled from the complexity reduction strategies required for practical noncoherent demodulation. Second, we can quantify how much of the gap between noncoherent and coherent demodulation is due to not knowing the amplitude, not knowing the phase, and due to complexity-reducing techniques. This indicates the effort required to refine demodulation and decoding strategies. The ultimate goal is to simplify joint noncoherent demodulation and decoding strategies for large QAM alphabets, so as to have complexity comparable to conventional pilot-based reception, while approaching noncoherent capacity.

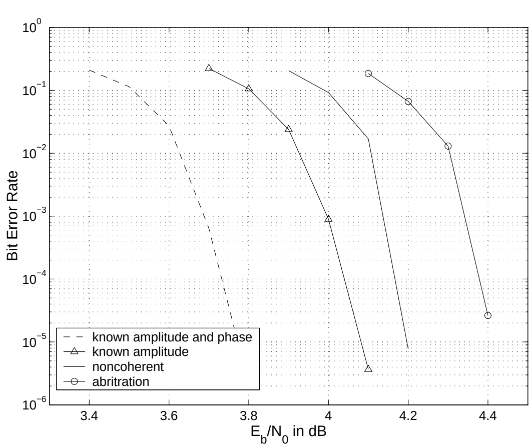

Figure 7: Phase-noisy AWGN results

The operating regime where the framework considered here is of most interest is that of low to moderate mobility, and moderate to large SNR: the lower the mobility, the more feasible it is to operate at higher SNR and higher spectral efficiency using a large constellation. However, as the mobility (and hence the rate of channel time variations) increases, the block-wise constant channel model used here starts breaking down, and it becomes extremely power-inefficient to operate at high SNR [11]. Indeed, even the structure of the noncoherent demodulator must be reconsidered in this setting [12].

References

[1] R.-R. Chen, D. Agrawal and U. Madhow, "Noncoherent detection of factor-graph codes over fading channels," *Proc. 2000 Conference on Informations Sciences and Systems (CISS)*, March, 2000.

[2] M. Peleg, S. Shamai, and S. Galan, "Iterative decoding for coded noncoherent MPSK communications over phase-noisy channels," *IEE Proc. Comm.*, vol. 147, no. 2, pp. 87–95, April 2000.

[3] P. Hoeher and J. Lodge, "Turbo DPSK: iterative differential PSK demodulation and channel decoding", *IEEE Trans. Comm.*, vol. 47, no, 6, pp. 837–842, June 1999.

[4] R.-R. Chen, R. Koetter, D. Agrawal and U. Madhow, "Joint demodulation and decoding for the noncoherent block fading channel: a practical framework for approaching channel capacity", To appear *IEEE Trans. Comm.*

[5] D. Divsalar and M.K. Simon, "Multiple-symbol differential detection of MPSK," *IEEE Trans. Comm.*, vol. 38, no. 3, pp. 300-308, March 1990.

[6] D. Warrier and U. Madhow, "Spectrally efficient noncoherent communication," *IEEE Trans. Inform. Theory*, vol. 48, no. 3, pp. 651–668, March 2002.

[7] T.L. Marzetta and B.M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Inform. Theory*, vol. 45, no. 1, pp. 139–157, January 1999.

[8] O. Macchi and L.L. Scharf, "A dynamic programming algorithm for phase estimation and data decoding on random phase channels," *IEEE Trans. Inform. Theory*, September 1981.

[9] L.R. Bahl, J. Cocke, F. Jelinek and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–287, March 1974.

[10] S. Benedetto, D. Divsalar, G. Montorsi and F. Pollara, "A soft-input soft-output maximum a posteriori (MAP) module to decode parallel and serial concatenated codes, Jet Propulsion Laboratory, Pasadena, CA, JPL TDA Prog. Rep. 42-127, November 1996.

[11] A. Lapidoth and S. Moser, "Capacity bounds via duality with applications to multi-antenna systems on flat fading channels," Preprint, 2002.

[12] R.-R. Chen, R. Koetter and U. Madhow, "Joint noncoherent demodulation and decoding for fast rayleigh fading channels," To appear *Proceedings CISS 2003*, March 2003.