

MITSUBISHI ELECTRIC RESEARCH LABORATORIES
<http://www.merl.com>

On Probabilistic Data Association for Achieving Near-Exponential Diversity over Fading Channels

Yellepeddi, A.; Kim, K.J.; Duan, C.; Orlik, P.

TR2013-062 June 2013

Abstract

Machine-to-Machine (M2M) wireless communication requires the transmission of short blocks of data with high reliability over fading channels. We discuss the use of the probabilistic data association (PDA) detector in conjunction with precoding to design high-performance systems for these links. First, the performance of the traditional PDA algorithm with precoding over ideal Rayleigh fading links is analyzed, which provides insight into its performance, and evidence of an error floor at high SNRs. Then, a novel ordering mechanism is proposed that takes advantage of the precoder characteristics. It is shown by simulation that the proposed modified algorithm can achieve near-ML performance for block sizes as small as 32 symbols

IEEE International Conference on Communications (ICC)

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

Copyright © Mitsubishi Electric Research Laboratories, Inc., 2013
201 Broadway, Cambridge, Massachusetts 02139

On Probabilistic Data Association for Achieving Near-Exponential Diversity over Fading Channels

Atulya Yellepeddi

Electrical Engineering-Applied Ocean Science and Engineering
MIT-WHOI Joint Program
Cambridge, MA 02139

Kyeong Jin Kim, Chunjie Duan and Philip Orlik

Mitsubishi Electric Research Laboratories (MERL)
Cambridge, MA 02139
{kkim,porlik,duan}@merl.com

Abstract—Machine-to-Machine (M2M) wireless communication requires the transmission of short blocks of data with high reliability over fading channels. We discuss the use of the probabilistic data association (PDA) detector in conjunction with precoding to design high-performance systems for these links. First, the performance of the traditional PDA algorithm with precoding over ideal Rayleigh fading links is analyzed, which provides insight into its performance, and evidence of an error floor at high SNRs. Then, a novel ordering mechanism is proposed that takes advantage of the precoder characteristics. It is shown by simulation that the proposed modified algorithm can achieve near-ML performance for block sizes as small as 32 symbols

Index Terms—Machine-to-Machine (M2M) Communication, single-input single-output (SISO) channel, precoding, probabilistic data association (PDA) algorithm, ordered-PDA.

I. INTRODUCTION

Machine-to-Machine (M2M) Communication is an important part of various modern settings, and has applications in factory automation, wireless telemetry and advertising, to name a few. M2M applications require the transmission of short blocks of information wirelessly, with extremely high reliability and low latency.

The M2M channel is usually a fading channel, which makes achieving these requirements difficult. Single-input single-output (SISO) channels with frequency-flat fading have poor reliability when the channel is not known at the transmitter [1]. The average uncoded error probability for such links falls off as a polynomial function of the received signal-to-noise ratio (SNR), in contrast to the performance of the optimal detector over an additive white Gaussian noise (AWGN) link [2], for which the probability of error decays exponentially with SNR, a limit termed infinite or exponential diversity [3].

One common way of improving the reliability over flat-fading channels is to use spatial diversity [4] (multiple-input multiple-output (MIMO) systems). A variety of near-maximum likelihood (ML) algorithms have been developed [5] for large MIMO systems that achieve close to exponential diversity at reasonable complexity. The likelihood ascent search (LAS) [6], the QRD-M algorithm [7], graphical model based approaches [8] and the probabilistic data association (PDA) algorithm [9] are some examples.

A. Yellepeddi was with Mitsubishi Electric Research Laboratories during this work.

However, devices in M2M networks are usually small and simple devices [10], and are unlikely to be equipped with a large number of antennas. Thus, the interest is in SISO links, or possibly small MIMO links. In [11], the authors discussed a method to turn SISO links into virtual-MIMO links, through precoding with a random matrix, a technique called pseudo-random phase precoding (PRPP); followed by detection with the LAS algorithm. This technique performed extremely well in practice, and the concept of precoding seems promising for SISO fading links. However, with the PRPP precoder and LAS detector, a block size of 400 symbols or more was required to achieve near-exponential diversity. Many practical M2M systems require still smaller block sizes, in which regime the performance of LAS falls off.

In this paper, we explore the PDA algorithm for SISO systems with precoding. First, we analyze the PDA algorithm to demonstrate that for large block sizes, it achieves AWGN performance over SISO links with precoding, but that it is subject to a finite block size error-propagation effect that limits performance. We then propose a novel ordering algorithm that improves performance by reducing the error propagation. This algorithm is shown to be particularly advantageous when the precoder is a Hadamard matrix (referred to as the *Hadamard Precoder*) or a discrete Fourier transform (DFT) matrix (*DFT precoder*).

Notation: Boldface lowercase math symbols represent vector quantities, boldface uppercase symbols represent matrices, as in \mathbf{A} , and a_{jk} is the (j, k) th element of \mathbf{A} . The symbol $(\cdot)^\dagger$ indicates the conjugate transpose of a matrix, $(\cdot)^*$ denotes complex conjugation, and the subscripts r and i , when applied to a scalar, denote its real and imaginary parts, respectively. $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ will indicate a complex normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. \mathbf{I}_N represents the $N \times N$ identity matrix, and \mathbf{e}_n represents its n th column.

II. SYSTEM MODEL

A simplified block diagram of the system under consideration is shown in Fig.1. A block of information bits \mathbf{b} is mapped into a vector $\mathbf{x} \triangleq [x_1, \dots, x_N]^T$ of symbols to be transmitted. The length of the vector \mathbf{x} is the *block size* of transmission, denoted N . We shall assume the binary phase shift keying (BPSK) for mapping, so that $\mathbf{x} \in \{-1, 1\}^N$. Also, assume that \mathbf{x} is uniformly distributed, so that any sequence is equally

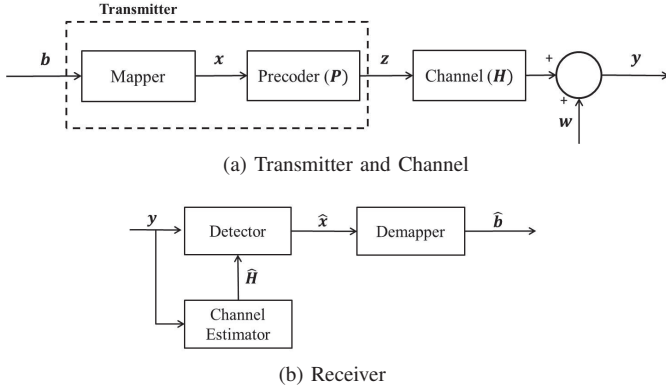


Fig. 1: Block diagram of the considered system.

likely. The block \mathbf{x} is mapped by the $N \times N$ precoding matrix \mathbf{P} into a set of precoded symbols $\mathbf{z} = \mathbf{P}\mathbf{x}$, where in general, $\mathbf{z} \in \mathbb{C}^N$. \mathbf{P} is assumed to be unitary, and its every element is assumed to have a magnitude $1/\sqrt{N}$. This implies that every element of the precoder may be written:

$$p_{jk} = \frac{e^{j\theta_{jk}}}{\sqrt{N}}, \quad \theta_{jk} \in [-\pi, \pi) \quad (1)$$

Precoded symbol block \mathbf{z} is then transmitted across the channel. The received signal vector \mathbf{y} is given by:

$$\begin{aligned} \mathbf{y} &= \sqrt{\gamma} \mathbf{H} \mathbf{P} \mathbf{x} + \mathbf{w} \\ &= \mathbf{G} \mathbf{x} + \mathbf{w} \\ &= \sqrt{\gamma} \tilde{\mathbf{G}} \mathbf{x} + \mathbf{w} \end{aligned} \quad (2)$$

where the channel is assumed to be an ideal Rayleigh fading channel, so that $\mathbf{H} \triangleq \text{diag}(h_1, \dots, h_N)$, where $h_k \sim \mathcal{CN}(0, 1)$ and $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$. The average received SNR is denoted γ . We also define $\mathbf{G} \triangleq \sqrt{\gamma} \mathbf{H} \mathbf{P}$ and $\tilde{\mathbf{G}} \triangleq \mathbf{H} \mathbf{P}$ as the equivalent channels with and without signal gain being applied. We assume perfect channel state information at the receiver, so that $\hat{\mathbf{H}} = \sqrt{\gamma} \mathbf{H}$. The received signal is then passed to the detector which forms an estimate $\hat{\mathbf{x}}$ of \mathbf{x} , and then a demapper, which maps this back into an estimate $\hat{\mathbf{b}}$ of the bits \mathbf{b} .

III. THE PROBABILISTIC DATA ASSOCIATION ALGORITHM

The principle of the PDA algorithm in [9] and [12] is to compute the probability distribution of each of the transmitted symbols, assuming the rest of the symbols and the channel noise together form an equivalent noise, and making a Gaussian assumption on the distribution of the equivalent noise. We rewrite (2), for any n , as:

$$\tilde{\mathbf{y}} = \mathbf{G}^{-1} \mathbf{y} = x_n \mathbf{e}_n + \sum_{k \neq n} x_k \mathbf{e}_k + \mathbf{G}^{-1} \mathbf{w} \quad (3)$$

PDA updates the distribution of x_n by treating the entire term $\sum_{k \neq n} x_k \mathbf{e}_k + \mathbf{G}^{-1} \mathbf{w}$ as an equivalent noise for the update of x_n . For the BPSK modulation, let $p_n \triangleq \mathbb{P}(x_n = 1)$. Then, PDA performs the update for p_n by computing the probability

that $x_n = 1$, conditioned on the knowledge of $\tilde{\mathbf{y}}$ and on all the probabilities computed so far, $p_k, k \neq n$.

$$\begin{aligned} p_n &= \mathbb{P}(x_n = 1) \\ &= \mathbb{P}(x_n = 1 | \tilde{\mathbf{y}}, \{p_k, k \neq n\}) \\ &= \frac{\mathbb{P}(\tilde{\mathbf{y}} | x_n = 1, \{p_k, k \neq n\})}{\mathbb{P}(\tilde{\mathbf{y}} | x_n = 1, \{p_k, k \neq n\}) + \mathbb{P}(\tilde{\mathbf{y}} | x_n = -1, \{p_k, k \neq n\})} \end{aligned} \quad (4)$$

To compute the probabilities of (4), PDA makes the assumption that, conditioned on x_n , $\tilde{\mathbf{y}}$ is a complex Gaussian. Thus, the mean and covariance of $\tilde{\mathbf{y}}$ evaluated as follows:

$$\begin{aligned} \boldsymbol{\mu}_n(x_n) &= \mathbb{E}[\tilde{\mathbf{y}} | x_n = 1, \{p_k, k \neq n\}] \\ &= x_n \mathbf{e}_n + \sum_{k \neq n} (2p_k - 1) \mathbf{e}_k \end{aligned} \quad (5)$$

and:

$$\begin{aligned} \boldsymbol{\Sigma}_n &= \text{cov}(\tilde{\mathbf{y}} | x_n = 1, \{p_k, k \neq n\}) \\ &= (\mathbf{G}^\dagger \mathbf{G})^{-1} + \sum_{k \neq n} 4p_k(1 - p_k) \mathbf{e}_k \mathbf{e}_k^\dagger \end{aligned} \quad (6)$$

Note that the covariance does not depend on the value of x_n . During the update of the n th symbol, using the assumption that, conditioned on x_n , $\tilde{\mathbf{y}} \sim \mathcal{CN}(\boldsymbol{\mu}_n(x_n), \boldsymbol{\Sigma}_n)$, along with (4), the update equation for p_n is given as:

$$p_n = \frac{1}{1 + \exp(-4\Re(\boldsymbol{\mu}_n^\dagger(1) \boldsymbol{\Sigma}_n^{-1} \mathbf{e}_k))} \quad (7)$$

The PDA algorithm starts by initializing $p_n = 0.5, n = 1, 2, \dots, N$. Then for each n , it runs the update step (7). After all the probabilities have been updated, the algorithm goes back and runs all the updates again. The process continues until all the probabilities have converged or each probability is updated at least a certain number of times (a fixed number of iterations are run).

IV. PERFORMANCE ANALYSIS OF THE PDA DETECTOR

The performance is analyzed in terms of symbol error rate (SER) for the BPSK signal. Assume that the probability of the n th symbol is being updated, and that all the other probabilities have converged, i.e., that for $k \neq n, p_k(1 - p_k) \approx 0$. Then, from (6), it is clear that $\boldsymbol{\Sigma}_n \approx (\mathbf{G}^\dagger \mathbf{G})^{-1}$. Also define $\bar{x}_k = 2p_k - 1$ for all k . Then, from (7), after a little algebra, we have:

$$p_n \approx \frac{1}{1 + \exp(-4\Re([\mathbf{y} - \sum_{k \neq n} \bar{x}_k \mathbf{G} \mathbf{e}_k]^\dagger \mathbf{G} \mathbf{e}_k))} \quad (8)$$

An error is made in the n th symbol if $p_n > 0.5$ when $x_n = -1$ or $p_n < 0.5, x_n = 1$. By conditioning on the input sequence \mathbf{x} , the other estimated symbols, $\{\bar{x}_k, k \neq n\}$, and the matrix \mathbf{G} , the error probability of $x_n = 1$ is given by:

$$\begin{aligned} P_{\text{err},n}(1) &= \mathbb{E}_{\{x_k, k \neq n\}, \mathbf{G}, \{\bar{x}_k, k \neq n\}} [\mathbb{P}[p_n < 0.5 | x_n = 1, \dots \\ &\quad \dots \{x_k, k \neq n\}, \mathbf{G}, \{\bar{x}_k, k \neq n\}]] \end{aligned} \quad (9)$$

From (8), we note that $p_n < 0.5$ if, and only if:

$$\Re\left[y - \sum_{k \neq n} \bar{x}_k \mathbf{G} e_k\right]^\dagger \mathbf{G} e_n < 0 \quad (10)$$

Using (2) and the definition of \mathbf{G} , (10) becomes, for $x_n = 1$:

$$\Re(\mathbf{w}^H \mathbf{G} e_n) < -\Re(e_n^\dagger \mathbf{G}^\dagger \mathbf{G} e_n + \sum_{k \neq n} \Delta x_k e_k^\dagger \mathbf{G}^\dagger \mathbf{G} e_n) \quad (11)$$

where $\Delta x_k \triangleq x_k - \bar{x}_k$.

Let us define $\tilde{w}_n \triangleq \Re(\mathbf{w}^H \mathbf{G} e_n)$, which can be written as follows:

$$\tilde{w}_n = \sqrt{\gamma} \sum_{l=1}^N (w_{l_r} \tilde{g}_{l_{n_r}} + w_{l_i} \tilde{g}_{l_{n_i}}) \quad (12)$$

where w_{l_r}, w_{l_i} being all independent real Gaussians with 0 mean and variance $1/2$. It is easy to see that $\mathbb{E}[\tilde{w}_n] = 0$, and:

$$\text{var}(\tilde{w}_n) = \frac{1}{2} \sum_{l=1}^N (\tilde{g}_{l_{n_r}}^2 + \tilde{g}_{l_{n_i}}^2) = \frac{\gamma}{2N} \sum_{l=1}^N |h_l|^2 \quad (13)$$

Based on (9) and (13), $p_{\text{err},n}(1)$ is evaluated as follows:

$$p_{\text{err},n}(1) = \mathbb{E} \left[\mathbb{Q} \left(\frac{\sqrt{2\gamma} \Re(e_n^\dagger \tilde{\mathbf{G}}^\dagger \tilde{\mathbf{G}} e_n + \sum_{k \neq n} (\Delta x_k) e_k^\dagger \mathbf{G}^\dagger \mathbf{G} e_n)}{\sqrt{\frac{1}{N} \sum_{l=1}^N |h_l|^2}} \right) \right] \quad (14)$$

where $\mathbb{Q}(\cdot)$ denotes the Q-function. The expectation is taken over the same quantities as in (9). Also, we can readily derive:

$$\Re(e_k^\dagger \tilde{\mathbf{G}}^\dagger \tilde{\mathbf{G}} e_n) = \frac{1}{N} \sum_{l=1}^N |h_l|^2 \cos(\theta_{l_n} - \theta_{l_k}) \quad (15)$$

where the θ_{j_k} s were defined in (1).

Substituting (15) into (14), the probability of error when $x_n = 1$ is given as follows:

$$p_{\text{err},n}(1) = \mathbb{E} \left[\mathbb{Q} \left(\frac{\sqrt{2\gamma}}{\sqrt{\frac{1}{N} \sum_{l=1}^N |h_l|^2}} \left(\frac{1}{N} \sum_{l=1}^N |h_l|^2 + \sum_{k \neq n} (\Delta x_k) \frac{1}{N} \sum_{l=1}^N |h_l|^2 \cos(\theta_{l_n} - \theta_{l_k}) \right) \right) \right] \quad (16)$$

A similar procedure can be applied for $x_n = -1$, and this yields:

$$p_{\text{err},n}(-1) = \mathbb{E} \left[\mathbb{Q} \left(\frac{\sqrt{2\gamma}}{\sqrt{\frac{1}{N} \sum_{l=1}^N |h_l|^2}} \left(\frac{1}{N} \sum_{l=1}^N |h_l|^2 - \sum_{k \neq n} (\Delta x_k) \frac{1}{N} \sum_{l=1}^N |h_l|^2 \cos(\theta_{l_n} - \theta_{l_k}) \right) \right) \right] \quad (17)$$

Thus, the overall probability of error is given by as:

$$p_{\text{err},n} = \frac{p_{\text{err},n}(1) + p_{\text{err},n}(-1)}{2} \quad (18)$$

A. Large Block Size Effects

As $N \rightarrow \infty$, we note that, by the weak law of large numbers, we have:

$$\frac{1}{N} \sum_{l=1}^N |h_l|^2 \rightarrow 1 \quad (19a)$$

$$\frac{1}{N} \sum_{l=1}^N |h_l|^2 \cos(\theta_{l_n} - \theta_{l_k}) \rightarrow 0 \quad (19b)$$

For PRPP, it is immediately obvious why the expected value for the θ s uniformly distributed in $[-\pi, \pi)$ is zero. For other precoders, Eqs. (19a) and (19b) are not true. However, for DFT precoders, it is approximately valid for large enough N . Further, for Hadamard precoders, the $\cos(\theta_{l_n} - \theta_{l_k})$ can be thought of as being equiprobably drawn from 1 and -1 for large enough N , so that the argument holds. We hypothesize that the condition holds for most reasonable unitary precoders satisfying the magnitude constraints on the elements.

Using (19a) and (19b) in (16), it becomes clear that:

$$p_{\text{err},n}(1) = p_{\text{err},n}(-1) \stackrel{N \rightarrow \infty}{\approx} \mathbb{Q}(\sqrt{2\gamma}) \quad (20)$$

Therefore, at large N , the performance of the PDA algorithm for precoded SISO channels is near the AWGN performance.

B. Finite Block Size Effects

For finite N , where the law of large numbers does not apply, the coefficient of Δx_k in (16) and (17) does not go to zero. Thus, any error made in one symbol raises the error probability of all the other symbols. In practice, the multiplication of the $\sqrt{2\gamma}$ term with the error terms leads to the performance breaking away from the AWGN performance at high SNRs. Errors propagate irrespective of SNR- i.e., once the first error is made, the propagation does not depend on the SNR, which can be observed from (16). Thus, the error propagation terms become more and more of a factor at high SNRs. While the probability error still decreases with SNR, the falloff becomes slower and is not as fast as the AWGN falloff. This effect is more pronounced at smaller block sizes, as, intuitively, there is less chance of recovering from an error once made because there is less information to combine.

V. ORDERED PDA (O-PDA)

One typical method of order in which PDA updates the symbol probabilities is the the optimal decision feedback ordering of [13]. However, for the ideal Rayleigh fading model under consideration, every permutation has the same cost for the proposed scheme of [13]. In other words, the decision feedback ordering is unable to differentiate between orderings for this channel. The proof is omitted here for lack of space.

However, it is clear that if the first few symbols to be decoded are correct, there is a definite overall advantage to the performance of the algorithm since there is less chance of error propagation. Hence, we would expect that the performance follows the ML curve for longer. Therefore we investigate orderings based on likelihood metrics.

A. Proposed Ordering Scheme

To design a mechanism for ordering, we consider the space of precoded symbols $z = P\mathbf{x}$. The n th element of z is given by:

$$z_n = \frac{1}{\sqrt{N}} \sum_{j=1}^N e^{j\theta_{nj}} x_j. \quad (21)$$

Let us assume that each z_n belong to the set $\mathcal{Z} \subset \mathbb{C}$, and let \mathcal{Z} be a finite set. Note that this may not be true (as in the case of the PRPP precoder) or, even if it is, $|\mathcal{Z}|$ may be very large (as for the DFT precoder).

We can rewrite (2) as $y_n = h_n z_n + w_n$, $n = 1, \dots, N$. Thus, knowing the set \mathcal{Z} , we can compute the likelihood, $p(y_n | z_n = z, h_n)$, for each $z \in \mathcal{Z}$, by:

$$p(y_n | z_n = z, h_n) \propto \exp\left(-\frac{|y_n - h_n z|^2}{2\sigma_w^2}\right). \quad (22)$$

If we choose a prior on the z_n s, $p(z_n = z)$, we can map these likelihoods into a posterior via Bayes' theorem:

$$p(z_n = z | y_n, h_n) = \frac{p(y_n | z_n = z, h_n) p(z_n = z)}{\sum_{z_0 \in \mathcal{Z}} p(y_n | z_n = z_0, h_n) p(z_n = z_0)}. \quad (23)$$

Using this posterior probability, we compute the expectation on the z_n 's as:

$$\mathbb{E}[z_n | h_n, y_n] = \sum_{z \in \mathcal{Z}} z p(z_n = z | y_n, h_n). \quad (24)$$

Now using the linearity of expectations, we can map this into an expected value of \mathbf{x} , as follows:

$$\boldsymbol{\mu}_x = P^{-1} [\mathbb{E}[z_1 | h_1, y_1], \dots, \mathbb{E}[z_N | h_N, y_N]]^T. \quad (25)$$

Once we have an expected value of \mathbf{x} , we can map that into probabilities of the x_n s, which can serve as the initial probabilities for the PDA algorithm. Specifically, for the BPSK case, it is given by:

$$[p_1, \dots, p_N]^T = \frac{1 + \boldsymbol{\mu}_x}{2}. \quad (26)$$

While (26) provides an initial set of probabilities for the algorithm, it still does not give the ordering itself. However, it is evident that the quantity $v_n \triangleq p_n(1 - p_n)$ measures the variance of the random variable x_n . The smaller v_n , indicates the higher confidence. Thus, we simply order the symbols in the ascending order of the value v_n , and update probabilities in that order.

Since the employed precoder is unitary, it is clear that none of the z_n s can have a magnitude larger than \sqrt{N} , as this is the root power of the sequence \mathbf{x} . Also, the typical sequence of \mathbf{x} has an equal number of 1s and -1 s, so linearity implies that the typical values of z_n , which is a complex linear combination of the values in \mathbf{x} , are all near the 0 in the complex plane. Thus, we choose a set of points in the complex plane, restricted to be within the circle of radius \sqrt{N} , and with the bulk of the points close to the origin, we should expect that to be sufficient for ordering. The black crosses in Fig.2 show a possible choice

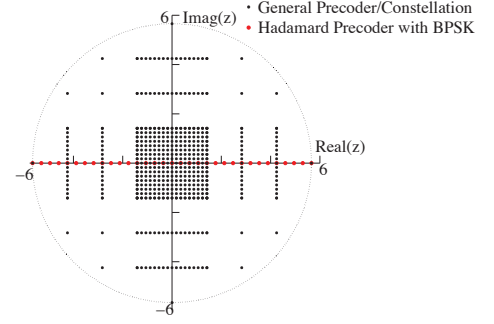


Fig. 2: Possible sets \mathcal{Z} .

of the set \mathcal{Z} for general complex precoders, with $N = 32$. A special case occurs when the precoder is the Hadamard matrix. As each entry is either $\pm 1/\sqrt{N}$, it can be seen that the set \mathcal{Z} is finite and real, and in fact can be found exactly as:

$$\mathcal{Z} = \left\{ -\sqrt{N} + \frac{2k}{\sqrt{N}}, k = 0, 1, \dots, N \right\}. \quad (27)$$

The red dots of Fig.2 shows this set for Hadamard precoders with $N = 32$.

B. Choice of Priors

To choose an effective prior, we assume that z_n is uniform in the set \mathcal{Z} . In particular, for the Hadamard precoder, it is easy to show that if \mathbf{x} is uniform, then every z_n is distributed binomially with parameters $(N, 0.5)$. Specifically:

$$\mathbb{P}\left(z_n = \frac{N - 2k}{\sqrt{N}}\right) = \frac{1}{2^N} \binom{N}{k}. \quad (28)$$

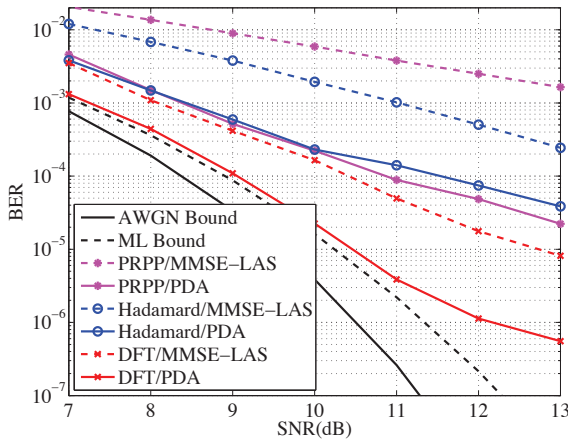
For the more general case, from (21), we note that z_n is essentially a weighted sum of a number of i.i.d. random variables, x_j . Evidently, therefore, we can make a central limit argument on the z_n , and assume that $z_n \sim \mathcal{CN}(0, 1)$. This is the prior that is used. We now summarize the complete O-PDA algorithm.

C. Summary of O-PDA Algorithm

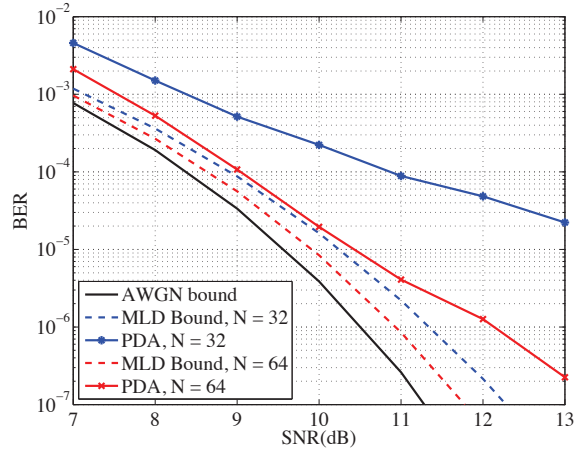
- 1) Choose a set \mathcal{Z} and a prior $p(z)$ on \mathcal{Z} , depending on the precoder and N .
- 2) For every $n = 1, \dots, N$, compute the likelihood, $p(y_n | z_n = z, h_n)$, for every $z \in \mathcal{Z}$, using (22).
- 3) Map that to a posterior, using (23).
- 4) Map that into an expectation on \mathbf{x} , using (24) and (25).
- 5) Use that to form probabilities p_n with (26).
- 6) Determine the ascending order of v_n .
- 7) Run PDA, using p_n as the initial probabilities and in the ordering determined above- the first symbol to be run should be the one with the smallest value of v_n .

VI. SIMULATION RESULTS

The channel for simulation is the ideal Rayleigh fading channel with uncorrelated channel taps, wherein the noise is white Gaussian. Performances are depicted at block sizes of



(a) PDA and MMSE-LAS [11] with various precoders, $N = 32$



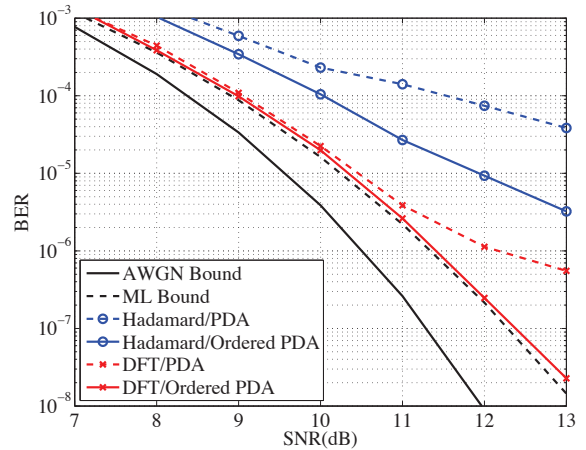
(b) PDA with PRPP precoding, at block sizes 32 and 64

Fig. 3: Simulated Performance of the PDA Detector

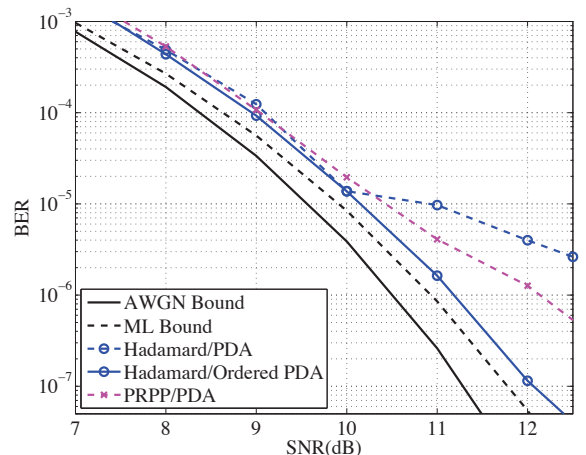
$N = 32$ and 64 , which are significantly smaller than in [11]. Also depicted are the AWGN and ML bounds. Note that the ML bound is different for different block sizes. As $N \rightarrow \infty$, the ML bound matches the AWGN bound.

Fig.3a demonstrates the performance of the PDA detector versus the MMSE-LAS detector of [11], with various precoders such as the PRPP precoder, the DFT matrix precoder and the Hadamard precoder. Fig. 3a shows that PDA shows a large improvement over the MMSE-LAS detector for all the different precoders. Evidently, the PDA algorithm in general provides a much more reliable estimate than the LAS detector. The predicted error propagation is also observed in the PDA algorithm. The performance follows the ML bound (which in turn follows the AWGN bound) until the high SNR regime. At high SNRs, the error rate falloff slows down and the curves break away from the bound. Another important feature of Fig.3a is that the DFT precoder is significantly better than either the PRPP or Hadamard precoder, for both PDA and LAS.

Comparing the PRPP and Hadamard precoded system decoded with LAS and PDA, in Fig.3a, we note that for LAS, the Hadamard precoder performs better whereas for PDA, the performance of the two precoders is comparable. This effect arises as a result of the fact that while LAS is a local search, for which the Hadamard precoder has some advantage in terms of codeword distances, PDA is a global search algorithm. An even more interesting effect is observed- at high SNRs, the PRPP precoder is better when coupled with PDA than the Hadamard precoder. Due to the structure of the Hadamard precoder, it is more strongly susceptible to error propagation than with PRPP. In Fig.3b, we consider the effect of increasing N with the PDA algorithm and the PRPP precoder. Evidently, the performance is improved when the block size is raised from 32 to 64- the SNR at which the performance breaks away from the corresponding ML bound is increased. This is in consonance with the analysis that we carried out- the longer



(a) $N = 32$



(b) $N = 64$

Fig. 4: Performance of PDA and O-PDA

the block size, the smaller the error propagation effect.

Fig.4 demonstrates the performance of the O-PDA algorithm, vis-a-vis the conventional PDA algorithm with arbitrary ordering. In Fig.4a, we compare the performance for a block size $N = 32$, for the DFT and Hadamard precoders. Note that the constellations chosen for ordering in the Hadamard and DFT cases are the respective ones of Fig.2 when $N = 32$, and that the priors chosen are, respectively for the Hadamard and DFT precoders, the binomial distribution of (28) and the complex standard normal distribution.

It is obvious that the ordering does in fact, as hypothesized, mitigate the error propagation, to the extent that the DFT performance is near ML, even at the relatively small block size of 32. The Hadamard precoder, evidently has poorer performance than the DFT precoder, but even here, the ordering provides an improvement. In this case, the performance improvement provided raises the SNR at which the performance characteristic becomes significantly non-exponential diversity. When $N = 64$, a similar effect is observed in Fig.4b). Given several observations, we would expect the DFT precoder to perform even better.

Note that an effect at high SNRs between the PRPP and Hadamard precoded systems, similar to that observed in Fig.3a, is demonstrated here also. The Hadamard and PRPP precoders perform about the same with PDA when it is run without regard to the ordering, but at high SNRs, the Hadamard precoder is slightly worse. However, the Hadamard precoded system with O-PDA is close to the ML bound. This clearly demonstrates the advantages of the O-PDA. For both $N = 32$ and 64, O-PDA is able, with a suitable precoder, to achieve within 0.5dB of the ML bound at an error rate of 10^{-8} . We also note that we have not applied ordering to the PRPP precoded system. The performance of PRPP/O-PDA is in fact rather poor, and the benefit is rather marginal.

VII. CONCLUSIONS AND FUTURE WORK

In this work, we considered the use of the PDA algorithm for the detection of signals transmitted over wireless fading links with precoding. We first analyzed the performance of the algorithm for this channel, and demonstrated the presence of error propagation, which would limit performance at high SNRs. We then provided an ordering mechanism that allowed us to initialize the algorithm with a set of probabilities and an order of updating that mitigated the error propagation, and called it the O-PDA algorithm. Simulation demonstrated the

use of the algorithms proposed, and provided us with insight into the functioning of various precoders across these channels and how to optimize the decoder for these various cases.

However, this work considered idealized Rayleigh channel models. The performance of the proposed scheme for channels other than independent Rayleigh fading remains to be investigated. Future work may also investigate the performance under correlated fading. A comparison against other schemes, such as Sphere Decoding is also being considered. Finally, the effect of unknown system parameters (such as the covariance matrix of the input) remains to be studied.

REFERENCES

- [1] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Information-theoretic and communications aspects," *Information Theory, IEEE Transactions on*, vol. 44, no. 6, pp. 2619–2692, 1998.
- [2] J. Proakis, *Digital communications*. McGraw-hill, 1987, vol. 1221.
- [3] L. Li and H. Jafarkhani, "Short-term performance limits of mimo systems with side information at the transmitter," *Arxiv preprint arXiv:1106.0733*, 2011.
- [4] M. Simon and M. Alouini, *Digital communication over fading channels*. Wiley-IEEE Press, 2005, vol. 86.
- [5] A. Chockalingam, "Low-complexity algorithms for large-mimo detection," in *Communications, Control and Signal Processing (ISCCSP), 2010 4th International Symposium on*. IEEE, 2010, pp. 1–6.
- [6] S. Mohammed, A. Chockalingam, and B. Rajan, "Low-complexity detection and performance in multi-gigabit high spectral efficiency wireless systems," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*. IEEE, 2008, pp. 1–5.
- [7] K. Kim, J. Yue, R. Iltis, and J. Gibson, "A qrd-m/kalman filter-based detection and channel estimation algorithm for mimo-ofdm systems," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 2, pp. 710–721, 2005.
- [8] P. Som, T. Datta, N. Srinidhi, A. Chockalingam, and B. Rajan, "Low-complexity detection in large-dimension mimo-isi channels using graphical models," *Selected Topics in Signal Processing, IEEE Journal of*, no. 99, pp. 1–1, 2011.
- [9] D. Pham, K. Pattipati, P. Willett, and J. Luo, "A generalized probabilistic data association detector for multiple antenna systems," *Communications Letters, IEEE*, vol. 8, no. 4, pp. 205–207, 2004.
- [10] G. Wu, S. Talwar, K. Johnsson, N. Himayat, and K. Johnson, "M2m: From mobile to embedded internet," *Communications Magazine, IEEE*, vol. 49, no. 4, pp. 36–43, 2011.
- [11] R. Annavejla and P. Orlik, "Achieving near-exponential diversity on uncoded low-dimensional mimo, multi-user and multi-carrier systems without transmitter csi," in *Information Theory and Applications Workshop (ITA), 2011*. IEEE, 2011, pp. 1–11.
- [12] Y. Jia, C. Vithanage, C. Andrieu, and R. Piechocki, "Probabilistic data association for symbol detection in mimo systems," *Electronics Letters*, vol. 42, no. 1, pp. 38–40, 2006.
- [13] M. Varanasi, "Decision feedback multiuser detection: A systematic approach," *Information Theory, IEEE Transactions on*, vol. 45, no. 1, pp. 219–240, 1999.