Optimal 8PSK mappings for BICM-ID over quasi-static fading channels

W. R. Carson and I. J. Wassell
Computer Laboratory
University of Cambridge
United Kingdom

M. R. D. Rodrigues
Instituto de Telecomunicações
Department of Computer Science
University of Porto, Portugal

Abstract—In this paper, we investigate the optimal mapping scheme for 8PSK for an iterative demapping and decoding system over quasi-static fading channels. We show that there are only 86 unique 8PSK mappings and conjecture that only 70 8PSK mappings will have unique error performance characteristics. We obtain an analytical expression for the demapper extrinsic information transfer (EXIT) functions using a Binary Erasure Channel (BEC) approximation to the AWGN channel model. In Section VI we also introduce a fixed a-priori criterion to classify all 8PSK mappings into 86 sub-sets [6].

Turbo or iterative techniques have been proposed for a large number of communication scenarios including multi-user processing [1], space-time processing [2] and the processing of bit interleaved coded modulation with iterative decoding (BICM-ID) [3]. They have been shown to exhibit excellent performance over the AWGN channel [4].

The quasi-static fading channel is extremely important because it models various practical scenarios characterised by extremely low time and frequency diversity, e.g., fixed wireless access (FWA) channels. Iterative techniques have been shown to improve performance in these channels however it has been observed that the choice of mapping scheme can have a dramatic effect on system performance [5]. It is therefore important to match the encoder and mapper appropriately. For our system model, two or more mapping schemes may result in identical system performance, this implies that we need only consider a subsection of the $8!=40,320$ mappings possible with 8PSK. Brännström and Rasmussen use a bit-wise distance criterion to classify all 8PSK mappings into 86 sub-sets [6].

This paper investigates the process of selecting an optimal 8PSK mapping scheme. Section II introduces the system model. Sections III, IV, V and VI demonstrate the existence of several equivalent mappings, leading to a reduction of the entire set of 8PSK mappings. In Section VI we also introduce the BEC approximation to the AWGN a-priori channel model for EXIT charts. Section VII provides the final BER simulation results and our optimal mapping. Finally, the main conclusions of this work are summarised in Section VIII.

II. SYSTEM MODEL

Fig.1 depicts the communications system model. We consider both single antenna systems ($N_T = N_R = 1$) often known as single-input single-output (SISO), which do not exploit space diversity, as well as multiple antenna systems ($N_T, N_R > 1$), which do exploit space diversity. The transmitter consists of four main stages: the encoder, the interleaver, the mapper and the space-time processor (see Fig.1). Initially, the information bits are convolutionally encoded at rate $R_c$, these coded bits are then pseudo-random interleaved. Finally, groups of 3 interleaved coded bits are mapped to a complex symbol from a unit power 8-ary phase shift keying (PSK) constellation.

In multiple transmit antenna systems ($N_T > 1$), the space-time processing block generates a space-time block code (STBC) according to the generator matrices $G_2, G_3$ or $G_4$ given by [7]. Essentially, a total of $K \times N_T$ symbols obtained from the original $K'$ modulation symbols are transmitted during $K$ time slots by $N_T$ transmit antennas.

The signal is distorted by a frequency-flat quasi-static fading channel as well as AWGN. Consequently, we express the relationship between the complex receive symbols and the complex transmit symbols associated with a specific STBC frame as:

$$ r = h s + n $$

(1)

Here, $r$ denotes the $N_R \times K$ matrix whose element $r_{j}(k)$ denotes the complex receive symbol at time slot $k$ and receive antenna $j$; $s$ denotes the $N_T \times K$ matrix whose element $s_i(k)$ denotes the complex transmit symbol at time slot $k$ and transmit antenna $i$; $h$ denotes the $N_R \times N_T$ matrix of channel gains whose element $h_{j,i}$ denotes the channel gain from transmit antenna $i$ to receive antenna $j$ (note that $h_{j,i}$ is independent of time slot $k$); and $n$ denotes the $N_R \times K$ matrix whose element $n_{j}(k)$ denotes the noise random variable at time slot $k$ and receive antenna $j$. The channel gains are uncorrelated circularly symmetric complex Gaussian with mean zero and variance $\frac{1}{2}$ per dimension. The noise random variables are uncorrelated circularly symmetric complex Gaussian with mean zero and variance $\frac{1}{2}SNR_{norm} = \frac{N_T}{2SNR}$ per dimension, where SNR denotes the average signal-to-noise ratio per receive antenna.
The receiver consists of two main parts: (i) the soft demapper and (ii) the soft-in soft-out decoder. These two stages are separated by pseudo-random interleavers and de-interleavers, and they exchange soft information in an iterative manner (see Fig.1). Specifically, the soft demapper takes as a priori information \( L_D^{\text{Dem}} \) on the code bits which is an interleaved version of the extrinsic information \( L_E^{\text{Dec}} \) on the code bits produced by the soft input-soft output decoder. Then, it computes the extrinsic information. The extrinsic information is then used as the a-priori input for the log-MAP algorithm [8] decoder, and is given by:

\[
\log_2 M \sum_{m' = 1}^{K'} \log_2 M' \sum_{m'' = 1}^{K''} p(r|m) \prod_{k'=1}^{K'} P_r(b_{m'}(k'))
\]

\[
\sum_{s \in \{s: b_{m}(k) = 0\}} \log_2 M \sum_{m' = 1}^{K'} P_r(b_{m'}(k'))
\]

where \( b_{m}(k) \) is the \( m \)-th bit conveyed by the \( k \)-th mapped symbol, \( s^+ = \{s: b_{m}(k) = 1\} \), and \( s^- = \{s: b_{m}(k) = 0\} \).

The eight positions in an 8PSK constellation are described by \( S(q) = \sqrt{E_b} e^{i \frac{2 \pi q}{7}} \) for \( q = 0, 1, ..., 7 \). The average symbol energy is given by \( E_s = 3R_c E_b \) where \( E_b \) is the average energy per information bit. The 3-bit symbols are mapped to the constellation positions using a mapping \( l \), such that the symbol at \( S(q) \) is given by \( l(q) \); the elements of \( l \) are expressed in octal notation. For example, the mapping in Fig. 2(a) is given by \( l_1 = [0, 1, 2, 3, 4, 5, 6, 7] \) and is known as natural mapping. A mapping is divided into three sub-mappings, each of which maps one of the bits in the 3-bit symbols to the constellation positions. We assign the following labels freely since this will define the other two sets. There are therefore \( 3! = 6 \) possible distinct labellings for each mapping.

Consider natural mapping, its three sub-mappings correspond to the three rings in Fig. 2(b). We choose the following labelling scheme for these sub-mappings:

\[
l_1 = BCD = \begin{bmatrix} 000 \\ 001 \\ 010 \\ 011 \\ 100 \\ 101 \\ 110 \\ 111 \end{bmatrix}^t
\]

where \( ^t \) denotes the transpose operator. The first sub-mapping label in \( l_1 \) is assigned to the inner most sub-mapping, the second label to the central sub-mapping and the final label to the outer most sub-mapping. Therefore for our choice of natural sub-mapping labels we get: \( B_s = 100, C_s = 010, D_s = 001 \) and \( T_s = 011, \bar{C}_s = 101, \bar{T}_s = 110 \).

III. MAPPING OPERATORS

When considering the mapping choice for our system model, we introduce four operators that take advantage of our assumptions on the channel noise and input bit distribution to drastically reduce the number of unique 8PSK mappings. The first two operators maintain the relative order of the symbols, using the assumption of circularly symmetric noise, to identify equivalent mappings. The next two operators maintain the Hamming distance between the constellation positions, using the assumption of equi-probable symbols at the channel input, to identify equivalent mappings. These final two operators have also been presented in [9] as the interchanging and complementation of rows (sub-mappings) respectively. If the mappings are viewed as a number expressed in octal, then we are able to index them in ascending order (See [6]), e.g. \( l_1 = [0, 1, 2, 3, 4, 5, 6, 7] \) and \( l_{40320} = [7, 6, 5, 4, 3, 2, 1, 0] \). We illustrate the effect of the operators on natural mapping (\( l_1 \)).

A. Constellation Rotation (Cr): \( l_{5914} = [1, 2, 3, 4, 5, 6, 7, 0] \)

Under the assumption of circularly symmetric noise, rotating the constellation positions of the symbols will not affect the complement of the bit sequence of symbol \( A_s \), for \( A \in \{B, C, D\} \). Therefore we are only able to assign one of these sets of labels freely since this will define the other two sets. There are therefore \( 3! = 6 \) possible distinct labellings for each mapping.

Consider natural mapping, its three sub-mappings correspond to the three rings in Fig. 2(b). We choose the following labelling scheme for these sub-mappings:

\[
l_1 = BCD = \begin{bmatrix} 000 \\ 001 \\ 010 \\ 011 \\ 100 \\ 101 \\ 110 \\ 111 \end{bmatrix}^t
\]
performance, i.e., all mappings \( l(q) = S(q + t) \) for \( t \in \mathbb{Z} \) are equivalent, e.g., \( l_1 \equiv l_{289} \). We refer to this operation of rotation as Constellation Rotation (Cr). The Cr operation gives a set expansion of eight, i.e., it gives a total of eight different mappings that are equivalent.

**B. Constellation Reflection (Cr):** \( l_{2890} = [0, 7, 6, 5, 4, 3, 2, 1] \)

Under the assumption of circularly symmetric noise, reflecting the constellation positions of the symbols will not affect performance, i.e., both mappings \( l(q) = S(tq) \) for \( t = \pm 1 \) are equivalent, e.g., \( l_1 \equiv l_{2890} \). We refer to this type of operation accordingly as constellation reflection (Cr). The Cr operation gives a set expansion of two.

**C. Sub-mapping Reordering (Sr):** \( l_{289} = \begin{bmatrix} 00110011 \\ 00001111 \\ 01010101 \end{bmatrix} \)

Under the assumption of equi-probable input symbols, reordering the sub-mappings will not affect performance, e.g., \( l_1 \equiv l_{289} \). We obtained \( l_{289} \) by exchanging sub-mappings \( B \) and \( C \). \( l_{289} = CBD \). We note that this reordering has affected the symbol labels, we now have \( B_s = 010, C_s = 100 \) and \( D_s = 001 \). We refer to this type of operation as sub-mapping reordering (Sr). The Sr operation gives a set expansion of \( 3! = 6 \).

**D. Bit Flipping (Bf):** \( l_{23617} = \begin{bmatrix} 11110000 \\ 00110011 \\ 01010101 \end{bmatrix} \)

Under the assumption of equi-probable symbols, complementation of one of the sub-mappings will not affect performance, e.g., \( l_1 \equiv l_{23617} \). This operation can also be represented by the exclusive OR operation with symbol \( X \), where \( X \in \{B_s, C_s, D_s, \overline{B_s}, \overline{C_s}, \overline{D_s}, 111\} \). We obtained \( l_{23617} \) by the complementation of sub-mapping \( B \) or \( X = B_s \). We note that the bit flipping operation does not affect the symbol values, i.e., \( B_s = 100, C_s = 010 \) and \( D_s = 001 \). The Bf operation gives a set expansion of \( 2^3 = 8 \).

**IV. SET REDUCING OPERATOR COMBINATIONS**

Using these four operators a mapping may generate a full set expansion of 768. Unfortunately, for some mappings combinations of the operators will generate the original mapping, they will therefore not have a full set expansion. Consider \( l_1 = [0, 1, 2, 3, 4, 5, 6, 7] \), if we Bf with \( X = B_s \) we generate \( l_{23617} = [4, 5, 6, 7, 0, 1, 2, 3] \) and we observe that with a simple Cr operation we can return to the original mapping, \( l_1 \). Therefore \( l_1 \) will not have a full set expansion.

We found that there are twelve combinations of operators that could render the original mapping, called \( T_1 \) to \( T_{12} \). These combinations were deduced by careful consideration of the effect that operators have on the phase relationships in a mapping. We introduce the following notation for Table I: the combination of mapping operators is denoted by the concatenation of their abbreviations, e.g., CrBf, and the phase of symbol \( S_2 \) anti-clockwise from \( S_1 \) is denoted by \( \theta(S_1, S_2) \). Some phase relationships are used repeatedly, so we abbreviate them further: \( \alpha = \theta(000, 111), \beta = \theta(B_s, \overline{B_s}), \gamma = \theta(C_s, \overline{C_s}) \) and \( \delta = \theta(D_s, \overline{D_s}) \). Each operator combination has particular phase requirements a mapping must satisfy, Table I.

The labelling scheme that we choose to apply to a mapping will not affect which combinations it satisfies, however in Table I we have chosen a labelling that focuses on symbols \( B_s \) or \( C_s \). Therefore when considering whether or not a mapping satisfies a particular combination, we should choose labellings that will also focus on symbols \( B_s \) and \( C_s \). If we are unable to find such a labelling scheme then the mapping will not satisfy the operator combination. For example, it is apparent that \( l_1 \) satisfies \( T_1 \) if we set \( B_s = 100, C_s = 010, D_s = 001 \). However because \( \theta(000, 100) = \pi \) for \( l_1 \) we will be unable to satisfy \( T_2 \) since phase requirement \( \theta(000, \overline{B_s}) = \pi \) is always violated regardless of our values of \( B_s, C_s, D_s \).

A single mapping generates an entire set expansion and we have calculated that there are a total of 86 non-equivalent non-overlapping set expansions in 8PSK. We deduced that there are 56 sets with reduced set expansions and calculated the size of these sets. Therefore we can calculate that there are 30 non-equivalent sets with full set expansions. We divide these 86 unique sub-sets of 8PSK into 15 groups according to the combinations they satisfy and label them accordingly, Table II. The table details the operator combinations a sub-set satisfies and its size (\( A_i \)).

A sub-set is generated from a mapping by using the four mapping operators, however the operators affect the phase

**TABLE I: Table of all set reducing operator combinations.**
relationships of a mapping and some Bf operations can generate a mapping with phase relationships that satisfy different operator combinations to the original mapping. For the first 12 sub-set groupings in Table II the mappings in a set expansion all satisfy the same combinations, but for the final three groups the mappings in the set expansions will satisfy one of two collections of operator combinations. We present both in the Multiple combinations column, and the number of mappings that satisfy each collection in a set expansion. The size of set expansion of a mapping and the number of operator combinations that are satisfied is halved for each non-parenthesised combination it satisfies. The combinations in parenthesis are satisfied coincidentally, they are the result of the non-parenthesised combinations and the phase restrictions these combinations place on a mapping, hence they do not contribute to the set reduction.

V. BIT-WISE DISTANCE SPECTRA

We are able to gain further insight into the unique 8PSK mappings by considering their bit-wise distance spectrum [6]. The bit-wise distance spectra with zero a-priori information for any mapping is defined below and is denoted by $W_0$; here and in the rest of the paper we use the sub-index $b$ to represent the number of known bits

$$ W_0 \triangleq \begin{bmatrix} w_0^1(1) & w_0^1(2) & w_0^1(3) & w_0^1(4) \\ w_0^2(1) & w_0^2(2) & w_0^2(3) & w_0^2(4) \\ w_0^3(1) & w_0^3(2) & w_0^3(3) & w_0^3(4) \end{bmatrix}, \tag{4} $$

where $w_0^i(j)$ denotes the average of the total Hamming distance for bit position $i = 1, 2, \ldots, m$ of a symbol to all other symbols at Euclidean distance $j$, for all symbols. We calculate this term below, where the Hamming distance between $i$th bits of the symbols at $S_n$ and $S_m$ is $d_H(S_n^i, S_m^i)$.

$$ w_0^i(j = 1, 2, 3) = \frac{1}{4} \sum_{n=0}^{7} d_H(S_n^i, S_{n+j}^i), \tag{5} $$

$$ w_0^i(j = 4) = \frac{1}{8} \sum_{n=0}^{7} d_H(S_n^i, S_{n+4}^i). \tag{6} $$

The elements of bit-wise distance spectrum for full a-priori information are defined in a similar fashion, except that symbols are considered if they only differ in the $i$th bit.

We introduce the notation $I_{E,b}$ to represent the average demapper Mutual information when $b$ bits are known. It was reasoned in [6] that the extrinsic Mutual information for zero a-priori information, $I_{E,0}$, is dependent on $W_0$, we would further conjecture that it is a function of the sum of the bit-wise distance spectra, $w_0(j) \triangleq \sum_{i=1}^{m} w_0^i(j)$. Similarly, it has been proved in [6] that the extrinsic Mutual Information for full a-priori information, $I_{E,2}$, is a function of the sum of the bit-wise distance spectra, $w_2(j)$:

$$ I_{E,2} = \frac{1}{m} \sum_{j=1}^{M/2} w_2(j) \left[ \sqrt{\gamma_3 \sin \left( \frac{\pi j}{M} \right)} \right]. \tag{7} $$

Using this classification, we can see from (5) and (6) that the mapping operators will not affect the values of $w_0$, or $w_2$, and hence equivalent mappings will have the same $[w_0, w_2]$ values. Furthermore, we have found that there are only 70 unique pairs of $[w_0, w_2]$, implying that some non-equivalent mappings have equivalent $[w_0, w_2]$ values. We therefore evaluated how these special non-equivalent mappings performed, through EXIT charts and BER plots.

Finally, we present all 86 unique sub-sets in Table III. We give the sub-set name, set expansion size $(A_e)$, the mapping with the lowest index in the sub-set ($l_e$), its mapping index $(c)$ and sub-set distance spectra values $[w_0, w_2]$. Some mappings are already well known and we also include these names: Gray, Natural (N), Set partitioning (SP) [10], modified set partitioning (MSP), semi-set partitioning (SSP) [11], and Maximum squared Euclidean weight (MSEW) [10].

VI. EXIT CHARTS

For a similar BICM-ID system over a quasi-static fading channel it was shown that the convergence threshold dictates error performance [12]. For an iteratively decoded system a standard tool for predicting the convergence threshold has been ten Brink’s EXIT charts [13]. Therefore although we are able to classify mappings for zero and full a-priori information, we are particularly interested in the form of the EXIT curve in the region between these two points.

There are three main methods to model the a-priori information in an iterative system [14]:

- Model it as an AWGN channel. This is ten Brink’s assumption for EXIT charts.
- Model it as a BEC with erasure probability $\epsilon$.
- Model it as a BSC with cross-over probability $p$.

We propose to approximate the AWGN a-priori channel with a BEC channel with erasure probability $\epsilon$. We are then able to derive the EXIT chart function (the extrinsic mutual information $I_E$) as a quadratic in terms of the a-priori mutual information $I_A$. We make use of the well-known relationship between the mutual information and erasure probability of the BEC, $I = 1 - \epsilon$

$$ I_E(I_A) = (1 - I_A)^2 I_{E,0} + I_A (1 - I_A) I_{E,1} + (I_A)^2 I_{E,2}. \tag{8} $$

It was further shown by ten Brink that for independent and uniformly distributed binary inputs the area under this EXIT
chart curve is equal to the capacity of the constellation, this is known as the area property. The capacity of a constellation is independent of the mappings scheme implemented, therefore we can rewrite the demapper EXIT function as a function of only two mapping dependent terms, \( I_{E,0} \) and \( I_{E,2} \), and the constellation constrained capacity, \( C_{8PSK} \):

\[
I_E(I_A) = \frac{1}{2} \left[ 3I_{E,2} + I_{E,0} - 2C_{8PSK} \right] I_A^2 + 2C_{8PSK} - 2I_{E,2} - 4I_{E,0} I_A + I_{E,0}.
\]

We have found that using the Bec as an approximation to the AWGN a-priori channel provides tight approximation for the EXIT chart curve of the demapper. We considered a wide range of \( E_b/N_0 \) and mapping schemes, Fig. 3. Using this model, we make a further conjecture that non-equivalent mappings with equivalent \([w_0, w_2]\) values will have the same error performance characteristics. Therefore we need only consider 70 non-equivalent 8PSK sub-sets.

Using the BEC a-priori approximation we are able to draw some guidelines for mapping selection for a particular convolutional code. Consider that at the convergence threshold of a turbo code, the decoding trajectory is able to just squeeze past the bottleneck between the two decoding EXIT functions and traverse the whole EXIT chart to the top right corner. This leads to a sudden sharp decrease in the BER curve. In order to achieve similar performance for our system we would also require the first intersection of the demapper and decoder EXIT curves at the convergence threshold to be as close to \( I_{Dem} = 1 \) as possible; this is more likely with concave, rather that convex, functions. For example, consider the SP mapping, Fig. 3; due to its convexity it intersects with the decoder EXIT curve earlier than would be desired. For this mapping there is a more gradual transition from a high to low BER at the convergence threshold.

The quadratic coefficient of (9) can be used to measure the concavity/convexity of EXIT functions. Furthermore due to the area property, for a fixed \( I_{E,0} \) the lower the value of \( I_{E,2} \) the greater the degree of convexity, but the higher the bit error rate. We define early convergence to a BER value of \( 10^{-5} \) as our optimisation criterion. There is therefore a trade-off between early convergence and low error floors.

VII. Simulation Results

This section investigates the performance of iterative demapping and decoding techniques over frequency-flat quasi-static fading channels both with and without antenna diversity. We considered the 86 unique 8PSK mapping schemes coupled with a rate 1/2 RSC code with octal generator polynomial \((1, 5, 7)\). We will emphasise differences in performance in the quasi-static fading and the AWGN regimes for BER performance curves. In our simulations, we consider a frame length of 2052 and a soft input-soft output decoder implementing the log-MAP algorithm.

We propose the mapping M8a (\( c = 48 \)) as optimal for our system model, the performance curves in Fig. 4 illustrate its performance both with and without antenna diversity. We compare its performance with three other mapping schemes: Gray mapping, which has been shown to have the best BER performance for non-iterative systems; semi-set partitioning (SSP) mapping, which has the lowest bit error floor for 8PSK; set partitioning (SP) mapping, which has a convex EXIT function and has the lowest convergence threshold. It is actually difficult to ascertain that SP mapping has the lowest convergence threshold from EXIT charts, and indeed there are a few mappings with very similar thresholds.

The EXIT functions for these four mappings are depicted in Fig. 3, as well as each Bec approximation curve and the decoder EXIT function. We observe that the M8a EXIT function has a high initial value and a slightly concave EXIT function, this combination enables it to have a much sharper BER plot decrease than SP mapping while having a much earlier convergence than SSP mapping.
We observe that iterating increases the performance in all scenarios with the exception of Gray mapping, which does not benefit. Both SP and M8a mappings have identical performance in systems with limited antenna diversity, and both outperform SSP mapping over all scenarios (antennas, iterations) over all $E_b/N_0$. For systems with significant antenna diversity for low $E_b/N_0$ SP mapping outperforms M8a mapping, however this corresponds to high BER values. We observe that M8a mapping outperforms the other three mappings for systems with significant antenna diversity at medium to high $E_b/N_0$ by approximately 1dB.

VIII. CONCLUSIONS

We observed that we can classify all 8PSK mappings into 86 sub-sets under comprehensive assumptions, with all mappings within a sub-set being equivalent. Furthermore we conjecture that only 70 of these sub-sets will have unique error performance characteristics. We show that using the BEC as an approximation to the AWGN for the a-priori channel in EXIT charts gives very good results and enables a quadric description of the EXIT curve. We use this analytical description to help aid our selection of an optimal mapping for our system model, M8a mapping, and compare its BER performance with mappings that are also optimal for different criteria. Results demonstrate that M8a mapping outperforms other mappings for our system with high antenna diversity and medium to high $E_b/N_0$. For systems that are non-iterative or have low antenna diversity, we observed that Gray mapping outperforms all other mappings for all $E_b/N_0$.

ACKNOWLEDGEMENT

The authors would especially like to thank Dr. Ioannis Chatzigeorgiou for valuable discussions and comments.

REFERENCES