Abstract: In this deliverable, we study iterative receivers for joint decoding and channel-state estimation for transmission on block-fading channels of root-LDPC-coded signals.

Keyword list: DAVINCI

Disclaimer: N/A
Executive Summary

This report summarizes some of our most recent findings about LDPC codes designed for use over the block-fading channels. They are mainly meant to supplement the results that were documented in previous reports. Sections 2 and 3 of this document review the main principles underlying the design of full-diversity LDPC codes for the block-fading channel, and in particular focus on nonbinary codes and on codes designed for channels affected by more than 2 fading values. All these results, as well as the results previously documented, focus on a situation in which the channel state information, viz., the values of the realizations of the fading values in the different blocks, are known at the receiver. Section 4 shows how the two steps of reception, i.e., channel estimation and decoding, can be unified in a single iterative structure based on factor graphs.
## Authors

<table>
<thead>
<tr>
<th>Partner</th>
<th>Name</th>
<th>Phone / Fax / e-mail</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WISER</strong></td>
<td>Ezio Biglieri</td>
<td>Phone: +39 0586 888487</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fax: +39 0586 888487</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e-mail: <a href="mailto:e.biglieri@ieee.org">e.biglieri@ieee.org</a></td>
</tr>
<tr>
<td><strong>ETIS Group (ENSEA/UCP/CNRS-UMR8051)</strong></td>
<td>Iryna Andriyanova</td>
<td>e-mail: <a href="mailto:iryna.andriyanova@ensea.fr">iryna.andriyanova@ensea.fr</a></td>
</tr>
<tr>
<td></td>
<td>David Declercq</td>
<td>e-mail: <a href="mailto:declercq@ensea.fr">declercq@ensea.fr</a></td>
</tr>
</tbody>
</table>
Table of Contents

1. Introduction ...................................................................................................................... 6

2. Notes on root LGPC codes .............................................................................................. 7
   2.1 Channel model............................................................................................................... 7
   2.1.1 Coding for the block-erasure channel ...................................................................... 7
   2.1.2 Coding for the block-fading channel ...................................................................... 8
   2.2 Code design for the BLEC .......................................................................................... 9
       2.2.1 A maximum-diversity, rate-1/3 binary code for the BLEC with $F = 3$ .................. 10
       2.2.2 A diversity-3, rate-1/2 binary code for the BLEC with $F = 4$ ......................... 11
   2.3 $q$-ary LDPC codes .................................................................................................... 12

3. The factor-graph approach ............................................................................................... 13
   3.1 Scheduling .................................................................................................................. 14
   3.2 Computing messages .................................................................................................. 14
       3.2.1 Mapper node ....................................................................................................... 14
       3.2.2 Repetition nodes ............................................................................................... 15
       3.2.3 Product nodes ................................................................................................... 15
       3.2.4 Channel nodes ................................................................................................... 15
   3.3 Messages involving continuous variables ..................................................................... 16
   3.4 Trained CSI estimation ............................................................................................... 16

4. About the performance of the joint estimator—decoder .............................................. 18

References .............................................................................................................................. 19
List of Acronyms and Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BFC</td>
<td>Block-Fading Channel</td>
</tr>
<tr>
<td>BLEC</td>
<td>Block-Erasure Channel</td>
</tr>
<tr>
<td>CSI</td>
<td>Channel State Information</td>
</tr>
<tr>
<td>FEC</td>
<td>Forward Error Correction</td>
</tr>
<tr>
<td>GF</td>
<td>Galois Field</td>
</tr>
<tr>
<td>LDPC</td>
<td>Low-Density Parity Check</td>
</tr>
<tr>
<td>OFDM</td>
<td>Orthogonal Frequency-Division Multiplexing</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal-to-Noise Ratio</td>
</tr>
<tr>
<td>WER</td>
<td>Word Error Probability</td>
</tr>
</tbody>
</table>
1. Introduction

Current challenges in wireless communications focus on providing a wide variety of users, including mobile ones, with reliable high-speed data services. The candidate technologies to meet this goal must then exploit the most promising state-of-the-art techniques to achieve a high level of spectral and energy efficiency, including forward error correction (FEC) coding.

Root-LDPC codes have been recently proposed for full-diversity transmission over block-fading channels [BOU10]. In this paper we study iterative receivers for joint decoding and channel-state estimation for the transmission of root-LDPC-coded signals. The block-fading channel model we are using is shown in Figure 1.1, which also defines our notations. A code word of length $N$ is transmitted over $F$ independently faded channels, each affected by a fading gain $R_i$, $i = 0, 1, \ldots, F-1$, and carrying coded symbols. The value of $F$ can be interpreted as an indication of the delay constraint imposed on the overall transmission system [BIG98]. It is known that smaller values of $F$, corresponding to tight delay constraints, limit the achievable diversity and hence impair the performance of a coded system under the assumption of known channel state information (CSI) at the receiver (in our context, CSI corresponds to the value taken on by the random variables $R_0, \ldots, R_{F-1}$). Our goal is to examine joint decoding of transmitted symbols and estimation of CSI, whose statistics are known. We are interested, in particular, in the tradeoff between decoding and CSI-estimation accuracy caused by the choice of the parameter $F$. In fact, while a smaller value of $F$ yields a worse error performance when the CSI is perfectly known at the receiver, for a fixed block length $N$ more symbols are affected by the same fading gain, and consequently the CSI can be estimated with more accuracy. We examine how all this affects the ultimate performance of the transmission system, as described by error probability.

![Figure 1.1 - Block-fading channel model](image)

Previous work in this area includes [WOR01], where iterative joint decoding and estimation of CSI for block-fading channels was examined, under the assumption of binary codes and a specific simplifying approximation for the messages generated at the output of the graph nodes associated with the fading gains. In [LIU08], the implementation of an iterative joint channel estimator and decoder is discussed by comparing different approximations for continuous messages. The graphical structure of our CSI estimator owes to the analysis presented in [DAU06a, DAU04].
2. Notes on root LGPC codes

Recent work [BOU07] has introduced a new family of binary LDPC codes, designed to approach the outage limit over the non-ergodic block-fading channel (BFC) when decoded using an iterative algorithm. Among the findings of [BOU07] most relevant here, we recall that

1. Standard random LDPC codes fail to achieve maximum diversity on the BFC, and new design criteria are called for.
2. Design criteria must be matched to the BFC and to iterative decoding, as the decoding performance of maximum likelihood designed codes is not satisfactory.

The main goal of these notes is to supplement that work with some additional observations.

2.1 Channel model

The block-fading channel (BFC) model, a special case of the block-interference channel defined in [MCE84], proves to be exceedingly useful in several situations in which transmission takes place over a channel with a low Doppler spread, and at the same time delay constraints force the interleaver/deinterleaver to be shorter than it would be required in order to make the channel memoryless. For example, it is well known that channel coding, in combination with interleaving, provides time diversity, especially in fast-fading environments with low coherence time. However, it may be insufficiently effective when fading is very slow, and at the same time large interleaving delays cannot be tolerated. In this case, the BFC model allows one to quantify the effect of finite depth interleaving, and hence of delay constraints, on the performance of a coded system [BIG98].

OFDM transmission offers another example where the BFC model may be a realistic, and useful, one. We consider the BFC model with \( F \) fading blocks, where each block has length \( N \) symbols, chosen from a \( q \)-ary alphabet. The alphabet may be chosen to form a Galois field \( \text{GF}(q) \). Codes over rings can also be envisaged. See, e.g., [RIC01a, Section V] and [SRI05]. See also [BEN06, p. 549] and [SRI05, Section I] for discussions of the signals to be used in the implementation of nonbinary codes.

Fading is flat, constant on each block, and i.i.d. on different blocks. The baseband equivalent channel model is given by

\[ y_f = \alpha_f x_f + n_f, \quad f = 1, \ldots, F \]  

where \( \alpha_f \) denotes the \( f \)th block fading coefficient and the noise \( n_f \) is i.i.d. complex circularly symmetric Gaussian, with components \( \sim \mathcal{N}(0,\sigma^2) \). The decoder is assumed to have perfect side information, it knows the values of all fading coefficients. The block-erasure channel (BLEC) can be viewed as a special case of BFC, whereby the fading coefficients can take on only two values, viz., 0, corresponding to an erasure, or \( \infty \), corresponding to a correct reception of the block. The side information of the decoder consists of knowing which code symbols were erased (if any), and its inputs are limited to either correct symbols or erasures (see [RIC01a, Chap. 3] and [GUI06b, LAP94] for discussions of erasure channels).

As observed in [BEN06], two differences distinguish codes on \( \text{GF}(q) \) from binary codes. First, each variable node is assigned an element of \( \text{GF}(q) \). Second, each edge of the Tanner graph carries a label in \( \text{GF}(q) \setminus \{0\} \). A word \( x \) is a code word if, at every check node \( j \), we have

\[ \sum_{i \in \mathcal{N}(j)} g_{i,j} x_i = 0 \]  

where \( g_{i,j} \) is the label of the edge joining symbol node \( i \) to check node \( j \), and \( \mathcal{N}(j) \) is the set of symbols checked by node \( j \).

2.1.1 Coding for the block-erasure channel

In the framework of the ensuing discussion, BLECs are relevant because the diversity of a code transmitted over them remains unchanged when a BFC is used. However, while diversity is the only relevant parameter to describe the performance of a code over the BLEC, two codes with the same
diversity may behave very differently over the BFC. Consideration of the BLEC provides considerable insight into the behavior of a code over the BFC, and code design for the BLEC yields a class of codes, characterized by the same diversity, among which one should pick the best code for the BFC. Specifically, we say that a code has diversity $d$ when its information content is lost over the BLEC if and only if $d$ or more of its blocks are erased. A sufficient condition for correct decoding over the BLEC is that the number of blocks erased, denoted $\epsilon$, satisfies $\epsilon < d$. A code has maximum diversity if and only if all its codewords differ in each block. For linear codes, this is tantamount to asking that no codeword have an all-zero block. We define the minimum blockwise Hamming weight as

$$w^b = \min \{w_1, \ldots, w_f\}$$

where $w_f$ is the Hamming weight of block $f$, and the minimum is taken with respect to all nonzero codewords. From (3), we see that $w^b > 0$ implies $d = F$. We can easily observe how the introduction of a nonbinary alphabet, which increases the number of codewords, makes it easier to generate a code combining high rate with maximum diversity.

**Example:** Block-repetition codes. Consider a code $B$ with rate $R$ and $K$ codewords. A new code $C$ is generated whose words are obtained by repeating $F$ times each codeword of $B$. The code $C$ has maximum diversity, and rate $R/F$.

### 2.1.2 Coding for the block-fading channel

On the BFC with Rayleigh fading, the word error probability is, asymptotically for large SNR, inversely proportional to a negative power of signal-to-noise ratio, whose exponent is the diversity $d$ of the code, that is, the minimum Hamming distance of a code of length $F$ constructed over an alphabet of size $q^r$. The Singleton bound yields an upper bound on $d$ [BIG05a, p. 104]:

$$d \leq 1 + F \left(1 - \frac{R}{\log_2 q}\right)$$

where $R$ is the code rate, expressed in bits per $q$-ary symbol. This bound has been proved to be optimal in an information-theoretical sense, as the outage probability of a coded scheme has an exponent dictated by (4), and this exponent achieves the RHS of (4) in the ensemble of random codes [10]. The worst possible bound occurs for binary codes: $q = 2$, yields $d \leq 1 + F(1 - R)$, a case examined in depth in [BOU07] for the case $F = 2$. To achieve maximum diversity, we need $\rho \leq 1/F$, where $\rho = R/\log_2 q$, i.e., for maximum diversity the code must have a number of information bits per binary digit not exceeding the inverse of $F$. The challenge, of course, is to derive a set of maximum-diversity codes, and among them minimize the error probability.

We say that a code has full diversity if it has a minimum Hamming block-distance equal to $F$. For full-diversity codes, the outage probability, the word-error, and the bit-error probability have the same asymptotic slope (for long blocks and large SNR) equal to $F$. In addition, on the BLEC full-diversity codes approach the outage-probability bound, while on BFCs having full diversity is only a necessary condition for this behavior.

The discussion above provides a rationale for the procedure we advocate to design nonbinary codes for the BFC. First, we choose a rate, and design a maximum-diversity binary code focusing on the BLEC. We thus find a set of constraints that the code must satisfy. Under these constraints, we look for the best code for the BF channel. In particular, the original set of constraints, translated into a graphical (Tanner) model, will include an equal degree for each symbol node and for each check node. This constraint will lead in a simple way to the design of (regular) LDPC codes. Next, we transform the binary code into a $q$-ary code by replacing the 1s in its parity-check matrix with elements of $\text{GF}(q) \setminus \{0\}$. In other words, the code designed for the BLEC will provide us with the topology of the Tanner graph of the nonbinary code, while the final structure of the latter is obtained by labeling its branches with elements of $\text{GF}(q) \setminus \{0\}$.
2.2 Code design for the BLEC

Following [DI02] and [RIC08, Chap. 3], we define a stopping set $S$ of a Tanner graph as the subset of symbol nodes such that all neighbors of $S$ are connected to $S$ more than once. The relevance of stopping sets in iterative decoding for the BLEC is explained by simple modification of an argument in [RIC08, Chap. 3]. This proves that iterative decoding over a BLEC which erases all symbols in a set $E$ cannot remove the erasures contained in the maximal stopping set included in $E$. In other words, the iterative decoder can remove all the erasures in $E$ if and only if the latter does not contain a stopping set. In fact, a symbol can be recovered from erasures if and only if at least one of its neighboring check nodes is connected to correctly received symbols. To see this in a simple way, observe that on the erasure channel a symbol can be recovered without error if and only if at most one symbol is affected by an erasure (Figure 2.1). In general we know that, in iterative decoding over the erasure channel, the following holds: (a) At a symbol node, the outgoing message is an erasure if and only if all incoming messages are erasures. Otherwise, the outgoing message is equal to the common value of the nonerasure messages. (b) At a check node, the outgoing message is an erasure if any of the incoming messages is an erasure. Otherwise, the outgoing message is equal to the opposite of mod-$q$ sum of the incoming messages.

To describe the relevance of stopping sets for codes over the BLEC, consider the simple example of the $(12, 6)$ binary code whose Tanner graph is shown in Figure 2.2, and whose parity-check matrix is

$$H = \begin{bmatrix} I_k & 0_k & S & S^- \\ S^T & S^- & I_k & 0_k \end{bmatrix}$$  \hspace{1cm} (5)

where $I_k$ and $0_k$ denote the $k \times k$ identity and null matrices, respectively, and

$$S^T \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad S^- \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$  \hspace{1cm} (6)

Figure 2.2 - A $(12,6)$ binary block code for the BLEC. Branch labels denote symbol degrees.
It is easily verified that the sets of symbols \{4, 5, 6\} and \{10, 11, 12\} are the maximal stopping sets contained in an erasure of the first and of the second block, respectively. In fact, we can see directly from (5) that symbol nodes \{1, 2, 3\} have neighbors 1c and 2c, but they are connected only once to 1c. Similarly, symbol nodes \{4, 5, 6\} have neighbors 2c, to which they are connected more than once, as reflected by the structure of submatrix \(S'\), whose columns have weight 3. The same arguments hold for the remaining symbol nodes.

Thus, iterative decoding of a codeword whose symbols 1\(\ldots\)6 are erased will cause the loss of symbols in positions 4, 5, and 6, while recovering correctly the symbols in positions 1, 2, and 3. Similarly, if symbols 7\(\ldots\)12 are erased, the decoder will recover correctly symbols 7, 8, and 9. Consequently, if this code is used for transmission over a BLEC with \(F = 2\), and its first block is erased, then symbols \{1, 2, 3\} can be recovered. Similarly, if the second block is erased, then only symbols \{7, 8, 9\} can be recovered. In conclusion, this code, as used over the BLEC with \(F = 2\), yields maximum diversity if its parity symbols are associated with the two stopping sets, and its information symbols with the remaining symbol nodes.

This example illustrates the fact that, for diversity \(d\), the code graph should be so designed in such a way that the maximal stopping set in each block excludes the information bits. Equivalently, no information symbol can be in a stopping set (notice also that the union of stopping sets is itself a stopping set). This concept was expressed in terms of “root checks” in [BOU07]: the root checks of [BOU07] are those prohibiting stopping sets on information symbols.

This code can be easily transformed into a (3, 6) LDPC code. It suffices to consider its parity-check matrix, and replace matrices \(S\) and \(S'\) with two sparse matrices of appropriate dimensions having column and row weights equal to 2 and 3, respectively. Some additional designs are described in the subsections that follow.

### 2.2.1 A maximum-diversity, rate-1/3 binary code for the BLEC with \(F = 3\)

Figure 2.3 illustrates a (18, 6) binary linear code for the BFC with \(F = 3\), having \(d = 3\) and rate \(\rho = 1/3\).

![Figure 2.3](image)

**Figure 2.3 - A rate-1/3, (18,6) binary block code with maximum diversity over the BLEC.**

Its parity-check matrix is
Figure 2.4 - A rate-1/3, (4,6) binary block code for the BLEC. Branch labels denote symbol degrees.

where

$$S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad S' = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$ (8)

2.2.2 A diversity-3, rate-1/2 binary code for the BLEC with $F = 4$

The Singleton bound suggests the possibility of a code for the channel with $F = 4$ and having diversity $d = 3$, $\rho = 1/2$. Figure 2.5 illustrates such a code. Its parity-check matrix is
where $S'$ has row and column weights both equal to 1, and $S''$ has row and column weights both equal to 3. The smallest length of the regular LDPC code is 24, obtained by choosing for example $S' = I$, and $S'' = J_k$, where $J_k$ is the $k \times k$ matrix all of whose entries are 1.

Observe that, if the channel degrades from $F = 4$ to $F = 2$ (which could occur when the fading becomes slower, for example because of a reduced mobility of the user), then this LDPC code can be transformed into the $(3, 6)$ LDPC code examined at the onset of this section. It suffices to merge symbol nodes 1 and 3, symbol nodes 2 and 4, check nodes 1c2 and 3c4, and check nodes 2c3 and 4c1. Thus, this code has maximum diversity for both channels having $F = 2$ and $F = 4$.

\[
H = \begin{bmatrix}
1 & 0 & S & 0 & 0 & S & S'' \\
S' & S & I & 0 & S & 0 & 0 \\
0 & 0 & S & S'' & I & 0 & S \\
S' & 0 & 0 & 0 & S & S'' & 1 & 0
\end{bmatrix}
\] (9)

2.3 $q$-ary LDPC codes

As discussed before, the generation of $q$-ary LPDPC codes be performed as follows. An LDPC code designed for the BLEC is described in terms of a certain number of sparse matrices. Choosing these at random in the binary field will generate a Tanner graph corresponding to a binary LDPC code; this can finally be transformed into the Tanner graph of a nonbinary code by attaching to the graph branches (as suggested in [DAV98]) random values in $GF(q) \setminus \{0\}$. Then, designing a full-diversity code for the BLEC yields full diversity over the BFC. However, nonbinary labels on the graph branches should be selected in a way that guarantees a high coding gain on the BFC.
3. The factor-graph approach

The communication system we are examining includes a code $\mathcal{C}$ with block length $n$. Its symbols $x_0, \ldots, x_{n-1}$ are grouped into $N$ blocks, each of them being sent to a mapper which accepts $n/N$ coded symbols and outputs an $M$-ary symbol $c_i$, $i = 0, \ldots, c_{N-1}$, which we assume two-dimensional for simplicity (Figure 3.1 refers to an example where the code is binary, $n/N = 3$, and $M = 8$).

![Figure 3.1 - Message update schedule](image)

The symbols $c_i$ are sent through a block-fading channel with parameter $F$, which causes them to be scaled by fading gains $R_0, \ldots, R_{F-1}$ (the example illustrated in Figure 3.1 refers to the case $F = 2$). The channel is described by the conditional probability density function (pdf) corresponding to additive white Gaussian noise with variance $N_0/2$:

$$p(y_i | c_i, R_f) \propto \exp\left[-\frac{1}{2} y_i - R_f c_i \right] / N_0$$

(10)

where $i = 0, \ldots, N-1$, $f \in \{i/n\} = 0, \ldots, F - 1$. The fading gains have common pdf $p(R)$. From now on, we shall write this density in the form $p(y_i | R_f, c_i)$ to stress its dependence on the product $R_f c_i$.

Maximum a posteriori decoding of symbols $x_i$ requires the solution of

$$\max_x p(x_i | y) = \max_x \sum_{y} \sum_{R} p(x, y, R)$$

(11)

where $R \in \{R_0, \ldots, R_{F-1}\}$, $y$ denotes the received vector, $\sum_R$ marginalization with respect to $R$, and $\sum_{x_i}$ marginalization with respect to all components of $x$ except $x_i$. Noting that

$$p(y, x, c, R) \propto P(x) P(c | x) p(R) p(y | c, R)$$

(12)

under the assumptions that the one-to-one map $m: x \rightarrow c$ is deterministic, that the code words of $\mathcal{C}$ are transmitted with equal probabilities, that the channel is stationary and invariant, and that the components of $R$ are independent, we can write [BIG05a]

$$p(x | y, R) \propto [x \in \mathcal{C}] [c = m(x)] \prod_{i=0}^{N-1} p(y_i | R_{i/n}, c_i) \prod_{f=0}^{F-1} p(R_f)$$

(13)

where $[P]$ is the Iverson function, whose value is 1 if proposition $P$ is true, and 0 otherwise. The corresponding factor graph is shown in Figure 3.1, in the special case $F = 2$ and $n/N = 3$. The
rectangular box with an = sign represents the “repetition” function, while that with a × sign represents the product function (see Figure 3.2).

![Figure 3.2 – Repetition and product functional blocks](image)

### 3.1 Scheduling

A possible schedule for the generation of messages in the sum–product algorithm is shown in Figure 3.1. Specifically, the figure describes the following steps:

3. Upward messages from nodes \( p(y_i | R_{ij}, c_i) \).
4. Messages from equality nodes.
5. Upward messages from product nodes.
6. Upward messages from mapper nodes.
7. One or more decoder iterations.
8. Downward messages from the decoder.
9. Downward messages from mapper nodes.
10. Messages towards equality blocks.
11. Messages towards nodes \( p(y_i | R_{ij}, c_i) \).

Steps 1 to 9 are iterated until a preassigned stopping condition is met.

### 3.2 Computing messages

#### 3.2.1 Mapper node

Consider, for illustration’s sake, the leftmost mapper of Figure 3.1. The mapper is represented by the function \( [c_0 = m(x_0, x_1, x_2)] \). The messages through the mapper (see Figure 3.3) are given by [LOE03]

\[
\mu_{m \rightarrow c_0}(c_0) = \sum_{d_0, d_1, d_2} [c_0 = m(x_0, x_1, x_2)] \mu_{d_0 \rightarrow m}(x_0) \mu_{d_1 \rightarrow m}(x_1) \mu_{d_2 \rightarrow m}(x_2) \tag{14}
\]

\[
\mu_{x_0 \rightarrow m}(x_0) = \sum_{d_0, d_1, d_2} [c_0 = m(x_0, x_1, x_2)] \mu_{d_0 \rightarrow m}(c_0) \mu_{d_1 \rightarrow m}(x_1) \mu_{d_2 \rightarrow m}(x_2) \tag{15}
\]

![Figure 3.3 - Message generated at the mapper](image)
3.2.2 Repetition nodes
We have, focusing on a repetition function with three arguments and referring to Figure 3.4:

\[ \mu_{\rightarrow R}^{\mu} \left( R' \right) = \mu_{\rightarrow R}^{\mu} \left( R \right) \times \mu_{\rightarrow R}^{\mu} \left( R' \right) \]  

(16)

![Figure 3.4 - Message generated at the equality node](image)

3.2.3 Product nodes
Observe from Figure 3.5 that \( R \) and \( c' \) are continuous variables, while \( c \) is discrete.

![Figure 3.5 - Message generated at the equality node](image)

We have

\[ \mu_{\rightarrow R(c)}(R) = \sum_{c} \mu_{\rightarrow R(c)}(R) \]  

(17)

and

\[ \mu_{\rightarrow R(c)}(c) = \int \mu_{R \rightarrow R(c)}(R) dR \]  

(18)

3.2.4 Channel nodes
Consider the computation of the message of Figure 3.6.

![Figure 3.6 – E message computation](image)
We have

\[
\mu_{p\rightarrow c_i}(c_i) = \int p(y_i | R_i, c_i) \mu_{R_i\rightarrow p}(R_i) dR_i
\]

(19)

\[
\mu_{p\rightarrow R_j}(R_j) = \sum_{c_i} p(y_i | R_j, c_i) \mu_{c_i\rightarrow p}(c_i)
\]

(20)

### 3.3 Messages involving continuous variables

Some of the messages detailed above involve continuous variables. We can observe, for example, that closed-form calculation of (19) seems out of the question. Thus, approximations are needed, each of them yielding a different iterative algorithm. Following [DAU04], we may consider the following approximations:

1. **Numerical integration.** This consists of replacing the integral in (19) by a finite sum, which is tantamount to using a quantized fading model. This method may become unfeasible if \( F \) is large.
2. **Gradient methods.** These have the lowest complexity [DAU04, Section III.C].
3. **Particle filtering.** See [DAU04, Section III.D] and [DAU06b].

### 3.4 Trained CSI estimation

If CSI is estimated, before information transmission, by sending a known sequence of pilots, the factor graph of Figure 3.1 should be modified by removing the whole part above the branches labeled \( c_0, \ldots, c_{N-1} \). Observe that the modified graph has no loops, and hence the sum-product algorithm computes the MAP estimates of \( R_0, \ldots, R_{N-1} \) in a single step.

\[
\begin{align*}
\mathcal{Z} & = \prod_{i=0}^{\nu-1} p(y_i | c_i, R_0) \\
p(R_0) & = \int p(y_0, \ldots, y_{\nu-1} | R_0) \prod_{i=0}^{\nu-1} p(y_i | R_i, c_i) dR_0
\end{align*}
\]

Figure 3.7 – Factor graph for trained estimation

Specifically, focusing (with no loss of generality) on the estimate of \( R_0 \) (see Figure 3.7), we have the a posteriori density

\[
p(R_0 | y_0, \ldots, y_{\nu-1}) \propto p(R_0) \prod_{i=0}^{\nu-1} p(y_i | R_i, c_i)
\]

(21)

whose maximization with respect to \( R_0 \) yields \( \hat{R}_0 \) (Notice that the \( c_i \) in (21) are known constants). As an example, assume normalized Rayleigh fading, so that, omitting the subscript 0 for notational simplicity,

\[
p(R) = 2Re^{-\|} \]

(22)

and equal-energy signals \( x_i = a_i \sqrt{E} \), \( |a_i|^2 = 1 \) and \( E \) the common signal energy. Maximizing the a posteriori density of \( R \) is tantamount to finding
\[
\hat{R} = \text{arg max}_R \left\{ \ln R - R^2 - \frac{1}{N_0} \sum_{i=0}^{N-1} y_r - R x_i \right\}^{1/2}
\]

(23)

\[
= \text{arg max}_R \left\{ \ln R - R^2 + 2R \frac{E}{N_0} \sum_{i=0}^{N-1} g(y, a_i^c) - R^2 \nu \frac{E}{N_0} \right\}
\]

(24)

\[
= \text{arg max}_R \left\{ \ln R - \left(1 + \nu R \frac{E}{N_0} \right) R^2 + 2 \frac{E}{N_0} wR \right\}
\]

(25)

where

\[
w = \frac{1}{\sqrt{E}} \sum_{i=0}^{N-1} g(y, a_i^c)
\]

(26)

turns out to be a Gaussian random variable with mean \( \nu R \) and variance \((E/N_0)^{-1} \nu / 2\). Taking the derivative, and defining \( \eta = E / N_0 \), we obtain the maximum a posteriori estimate of \( R \) as the positive solution of the equation

\[
1 - 2(1 + \nu \eta) \hat{R}^2 + 2\nu \eta \hat{R} = 0
\]

(27)

which is given by

\[
\hat{R} = \frac{\eta w \pm \sqrt{2 + 2
\nu \eta + \eta^2 w^2}}{2(1 + \nu \eta)}
\]

(28)

As a check, observe that, for high signal-to-noise ratios \( \eta \to \infty \), (28) yields

\[
\hat{R} \approx \frac{2\eta w}{2\eta \nu} = \frac{w}{\nu} = \frac{R}{\nu}
\]

(29)

as it should be.

This same procedure can be followed for iterative decision-directed estimation, whereby the sequence \( c \) is first decoded under the rough estimate \( \hat{R} = w / \nu \), and successively used to improve upon the estimate of \( R \). Notice that this corresponds to substituting, for the message going downward along branch \( c_i \), the “hard decision” on \( c_i \).
4. About the performance of the joint estimator—decoder

The performance of the iterative joint estimator—decoder does not seem to be amenable to analysis, as it would depend strongly on the specific channel model, and in particular on the values taken on by the $F$ fading values. What can be said in general is that the performance should lie in the middle between that of a, “uninformed” receiver in which no channel state information is known (i.e., a designed for channel assumed to be affected by AWGN only) and a receiver in which the channel state information is perfectly known, as if provided by an external genie. In addition, we conjecture that our iterative receiver preserves the code diversity, contrary to the uninformed receiver. The example below illustrates the statement above, and substantiates our conjecture.

Figure 4.1 shows the word error probability (WER) versus signal-to-noise ratio (SNR) achieved by a binary (3, 6) root-LDPC code of length 200 used for transmission over a block-fading channel with $F = 2$ and fading coefficients 1.0 and 0.5 under three scenarios: Perfect CSI at receiver (lower curve), joint channel estimation and decoding (middle curve) and no CSI at receiver (upper curve). It is seen how the joint estimator/decoder preserves the code diversity, and causes a moderate loss of SNR.

![Figure 4.1](image_url)

**Figure 4.1** – Simulation results. Word error probability (WER) versus signal-to-noise ratio (SNR) achieved by a binary (3, 6) root-LDPC code of length 200 used for transmission over a block-fading channel with $F = 2$ and fading coefficients 1.0 and 0.5 under three scenarios: Perfect CSI at receiver (lower curve), joint channel estimation and decoding (middle curve) and no CSI at receiver (upper curve).
References


