Joint channel estimation and decoding of root LDPC codes in block-fading channels

Iryna Andriyanova
ETIS Group
ENSEA/UCP/CNRS-UMR8051
F-95014 Cergy-Pontoise, France
iryna.andriyanova@ensea.fr

Ezio Biglieri
WISER S.r.l.
Via Fiume 23, I-57123 Livorno, Italy
e.biglieri@ieee.org

David Declercq
ETIS Group
ENSEA/UCP/CNRS-UMR8051
F-95014 Cergy-Pontoise, France
declercq@ensea.fr

***Abstract*—We study iterative receivers for joint decoding and channel-state estimation for transmission on block-fading channels of root-LDPC-coded signals. Root-LDPC codes are known to be most performant codes for block-fading channels, as their spacial "root" structure allows to get the full-diversity property. This property ensures a good error decoding performance of root LDPC codes, especially in contrast with the performance standard LDPC codes (having the maximum diversity equal to 1). However, as any channel code, root-LDPC codes also suffer from the diversity loss when the channel state information is not known at the receiver. In this work we propose a joint channel estimation- decoding scheme for root-LDPC codes that helps to overcome this problem and still to have the full-diversity.**

I. INTRODUCTION

Root-LDPC codes have been recently proposed for full-diversity transmission over block-fading channels [4]. 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 Fig. 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, $i = 0, 1, \dots, F-1$, and carrying coded symbols. The value of F can be interpreted as an indication of the delay constraint imposed on the overall transmission system [3]. 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 $R_0, \dots, R_{F-1}$). 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.

Previous work in this area includes [13], 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 [8], 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 [5], [7].

II. THE FACTOR-GRAPH APPROACH

The communication system we are examining includes a code $\mathcal{C}$ with block length n. Its symbols $x_0, \dots, x_{n-1}$ 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, $i = 0, \dots, c_{N-1}$, which we assume two-dimensional for simplicity (Fig. 2 refers to an example where the code is binary, $n/N = 3$, and $M = 8$). The symbols c_i are sent through a block-fading channel with parameter F, which causes them to be scaled by fading gains $R_0, \dots, R_{F-1}$ (the example illustrated in Fig. 2 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 $N_0/2$:

$$p(y_i \mid c_i, R_f) \propto \exp\{-|y_i - R_f c_i|^2/N_0\} \tag{1}$$

where $i = 0, \dots, N-1$, $f \triangleq \lfloor i/\nu \rfloor = 0, \dots, F-1$. The fading gains have common pdf $p(R)$. From now on, we shall write this density in the form $p(y_i \mid R_f c_i)$ to stress its dependence on the *product* $R_f c_i$. Maximum a posteriori decoding of symbols x_i requires the solution of

$$\max_{x_i} p(x_i \mid \mathbf{y}) = \max_{x_i} \sum_{\sim x_i} \sum_{\mathbf{R}} p(\mathbf{x}, \mathbf{y}, \mathbf{R}) \tag{2}$$

where $\mathbf{R} \triangleq (R_0, \dots, R_{F-1})$, $\mathbf{y}$ denotes the received vector, $\Sigma_{\mathbf{R}}$ marginalization with respect to $\mathbf{R}$, and $\sum_{\sim x_i}$ marginalization with respect to all components of $\mathbf{x}$ except x_i. Noting that

$$p(\mathbf{y}, \mathbf{x}, \mathbf{c}, \mathbf{R}) \propto P(\mathbf{x}) P(\mathbf{c} \mid \mathbf{x}) p(\mathbf{R}) p(\mathbf{y} \mid \mathbf{c}, \mathbf{R}) \tag{3}$$

under the assumptions that the one-to-one map $m : \mathbf{x} \mapsto \mathbf{c}$ 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 $\mathbf{R}$ are independent, we can write [1]

$$p(\mathbf{x} \mid \mathbf{y}, \mathbf{R}) \propto [\mathbf{x} \in \mathcal{C}][\mathbf{c} = m(\mathbf{x})] \prod_{i=0}^{N-1} p(y_i \mid R_{\lfloor i/\nu \rfloor} c_i) \prod_{k-0}^{F-1} p(R_k) \tag{4}$$

where $[\mathcal{P}]$ is the Iverson function, whose value is 1 if proposition $\mathcal{P}$ is true, and 0 otherwise. The corresponding factor graph is shown in Fig. 2, in the special case $F = 2$ and $n/N = 3$. The rectangular box with an $=$ sign represents the "repetition" function, while that with a $\times$ sign represents the product function (see Fig. 3).

A. Scheduling

A possible schedule for the generation of messages in the sum–product algorithm is shown in Fig. 2. Specifically, the figure describes the following steps

① Upward messages from nodes $p(y_i \mid R_{\lfloor i/\nu \rfloor} c_i)$.
② Messages from equality nodes.
③ Upward messages from product nodes.
④ Upward messages from mapper nodes.
⑤ One or more decoder iterations.
⑥ Downward messages from the decoder.
⑦ Downward messages from mapper nodes.
⑧ Messages towards equality blocks.
⑨ Messages towards nodes $p(y_i \mid R_{\lfloor i/\nu \rfloor} c_i)$.

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

B. Computing messages

1) Mapper node: Consider, for illustration's sake, the leftmost mapper of Fig. 2. The mapper is represented by the function $[c_0 = m(x,x_1,x_2)]$. The messages through the mapper (see Fig. 4) are given by [9]

$$\begin{aligned}\mu_{m\to c_0}(c_0) &= \sum_{x_0,x_1,x_2} [c_0 = m(x_0,x_1,x_2)]\mu_{x_0\to m}(x_0) \\ &\quad \mu_{x_1\to m}(x_1)\mu_{x_2\to m}(x_2); \qquad (5)\\ \mu_{m\to x_0}(x_0) &= \sum_{c_0,x_1,x_2} [c_0 = m(x_0,x_1,x_2)]\mu_{c_0\to m}(c_0) \\ &\quad \mu_{x_1\to m}(x_1)\mu_{x_2\to m}(x_2). \qquad (6)\end{aligned}$$

2) Repetition nodes: We have, focusing on a repetition function with three arguments and referring to Fig. 5:

$$\mu_{=\to R''}(R'') = \mu_{R\to =}(R'') \times \mu_{R'\to =}(R'') \qquad (7)$$

3) Product nodes: Observe from Fig. 6 that R and $c' \triangleq Rc$ are continuous variables, while c is discrete. We have

$$\mu_{\times\to R}(R) = \sum_c \mu c \to \times(c)\mu_{c'\to\times}(Rc) \qquad (8)$$

and

$$\mu_{\times\to c}(c) = \int \mu R \to \times(c)\mu_{c'\to\times}(Rc)\, dR \qquad (9)$$

4) Channel nodes: Consider the computation of the message of Fig. 7. We have

$$\mu_{p\to c_i}(c_i) = \int p(y_i \mid R_f c_i)\mu_{R_f\to p}(R_f)\, dR_f \qquad (10)$$

$$\mu_{p\to R_f}(R_f) = \sum_{c_i} p(y_i \mid R_f c_i)\mu_{c_i\to p}(c_i) \qquad (11)$$

C. Messages involving continuous variables

Some of the messages detailed above involve continuous variables. We can observe, for example, that closed-form calculation of (10) seems out of the question. Thus, approximations are needed, each of them yielding a different iterative algorithm. Following [7], we may consider the following approximations:

1) *Numerical integration.* This consists of replacing the integral in (10) 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 [7, Section III.C].
3) *Particle filtering.* See [7, Section III.D] and [6].

D. Trained CSI estimation

If CSI is estimated, before information transmission, by sending a known sequence of pilots, the factor graph of Fig. 2 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 $R_0, \ldots, R_{F-1}$ in a single step. Specifically, focusing (with no loss of generality) on the estimate of R_0 (see Fig. 8), we have the a posteriori density

$$p(R_0 \mid y_0, \ldots, y_{\nu-1}) \propto p(R_0) \prod_{i=0}^{\nu-1} p(y_i \mid R_0 c_i) \qquad (12)$$

whose maximization with respect to R_0 yields $\widehat{R}_0$. (Notice that the c_is in (12) are known constants.) As an example, assume normalized Rayleigh fading, so that, omitting the subscript 0 for notational simplicity,

$$p(R) = 2Re^{-R^2} \qquad (13)$$

and equal-energy signals $x_i = a_i\sqrt{\mathcal{E}}$, $|a_i|^2 = 1$ and $\mathcal{E}$ the common signal energy. Maximizing the a posteriori density of R is tantamount to finding

$$\begin{aligned}\widehat{R} &= \arg\max_R \left\{ \ln R - R^2 - \frac{1}{N_0}\sum_{i=0}^{\nu-1} |y_i - Rx_i|^2 \right\} \qquad (14)\\ &= \arg\max_R \left\{ \ln R - R^2 + 2R\frac{\mathcal{E}}{N_0}\sum_{i=0}^{\nu-1} \Re(y_i a_i^\star) - R^2\nu\frac{\mathcal{E}}{N_0} \right\} \qquad (15)\\ &= \arg\max_R \left\{ \ln R - \left(1 + \nu R^2\frac{\mathcal{E}}{N_0}\right) R^2 + 2\frac{\mathcal{E}}{N_0} wR \right\} \qquad (16)\end{aligned}$$

where

$$w \triangleq \frac{1}{\sqrt{\mathcal{E}}}\sum_{i=0}^{\nu-1} \Re(y_i a_i^\star) \qquad (17)$$

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

$$1 - 2(1 + \nu\eta)\widehat{R}^2 + 2\eta w\widehat{R} = 0 \qquad (18)$$

which is given by

$$\widehat{R} = \frac{\eta w + \sqrt{2 + 2\nu\eta + \eta^2 w^2}}{2(1 + \nu\eta)} \tag{19}$$

As a check, observe that, for high signal-to-noise ratios $\eta \to \infty$, (19) yields

$$\widehat{R} \approx \frac{2\eta w}{2\eta\nu} = \frac{w}{\nu} \approx R \tag{20}$$

as it should be.

This same procedure can be followed for iterative decision-directed estimation, whereby the sequence $\mathbf{c}$ is first decoded under the rough estimate $\widehat{R} = w/\nu$, and successively used to improve upon the estimate of $\mathbf{R}$. Notice that this corresponds to substituting, for the message going downward along branch c_i, the "hard decision" on c_i.

III. Simulation Results

Figure 9 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 (dotted curve), joint channel estimation and decoding (full curve) and no CSI at receiver (dashed curve).

The codelength of only 200 bits has been chosen for our simulations. However, note that the codelength value does not have a large impact on the decoding performance over block-fading channels, once F does not change [4].

It is seen how the joint estimator/decoder preserves the code diversity, and causes a very moderate loss of SNR. The mean-square error, achieved by the channel estimator after 5 channel estimation rounds, is only in the range from 0.04 to $3.5 \cdot 10^{-3}$ for the range of SNRs from 0 to 10 dB. Numerical results show that one can reduce the number of channel estimations to $2 - 3$, without degrading the performance too much. The reason for this comes from the fact that the channel estimation is performed jointly over *all* the received messages, that have experienced the same value R, and a very few number of channel estimation rounds sufficed to estimate the channel fading gain correctly.

Note that the decoding performance for standard LDPC codes has the diversity-1 slope even with the perfect CSI at receiver [4].

Acknowledgment

This work was supported by the European FP7 ICT-STREP DAVINCI project under the contract No. INFSO-ICT-216203.

References

[1] E. Biglieri, *Coding for Wireless Channels.* New York: Springer, 2005.

[2] E. Biglieri, A. Nordio, and G. Taricco, "Doubly-iterative decoding of spacetime turbo codes with a large number of antennas," *IEEE Trans. Commun.*, Vol. 53, No. 5, May 2005.

[3] E. Biglieri, J. Proakis, and S. Shamai (Shitz), "Fading channels: Information-thoretic and communications aspects, *IEEE Trans. Inform. Theory*, Vol. 44, No. 6, pp. 2619-2692, October 1998.

[4] J. J. Boutros, A. Guillén i Fàbregas, E. Biglieri, and G. Zémor, "Design and analysis of low-density parity-check codes for block fading channels," *2nd ITA Workshop*, San Diego, CA, January 28-February 2, 2007. The full paper "Low-density parity-check codes for nonergodic block-fading channels," will be published in *IEEE Trans. Inform. Theory* in the September, 2010, issue.

[5] J. H. G. Dauwels, *On Graphical Models for Communications and Machine Learning: Algorithms, Bounds, and Analog Implementations.* DS Dissertation, ETH Zürich, Switzerland, May 2006.

[6] J. Dauwels, S. Korl, and H.-A. Loeliger, "Particle methods as message passing," in *Proc. International Symp. Inform. Theory (ISIT 2006)*, Seattle, USA, pp. 2052-2056, July 9-14, 2006.

[7] J. Dauwels and H.-A. Loeliger, "Phase estimation by message passing," in *Proc. IEEE International Conference on Communications (ICC 2004)*, Vol. 1, pp. 523-527, June 20-24, 2004.

[8] Y. Liu, L. Brunel, and J. J. Boutros, "Belief propagation with Gaussian approximation for joint channel estimation and decoding," *IEEE PIMRC 2008*, Cannes, France, September 15-18, 2008.

[9] H.-A. Loeliger, "Some remarks on factor graphs," in *Proc. 3rd Int. Symp. on Turbo Codes and Related Topics*, Brest, France, pp. 111-115, September 1-5, 2003.

[10] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Proc. Mag.*, vol. 21, No. 1, pp. 28-41, January 2004.

[11] H. Niu, M. Shen, J. A. Ritcey, and H. Liu, "A factor graph approach to iterative channel estimation and LDPC decoding over fading channels," *IEEE Trans. Wireless Commun.*, Vol. 4, No. 4, pp. 1345-1350, July 2005.

[12] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge, U.K.: Cambridge University Press, 2008.

[13] A. P. Worthen and W. E. Stark, "Unified design of iterative receivers using factor graphs," *IEEE Trans. Inform. Theory*, Vol. 47, No. 2, pp. 843-849, February 2001.

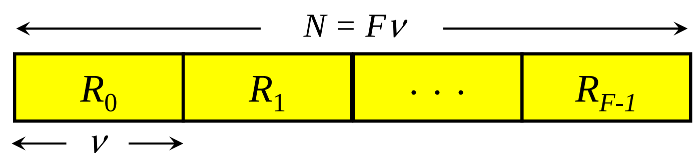

Fig. 1. Block-fading channel model.

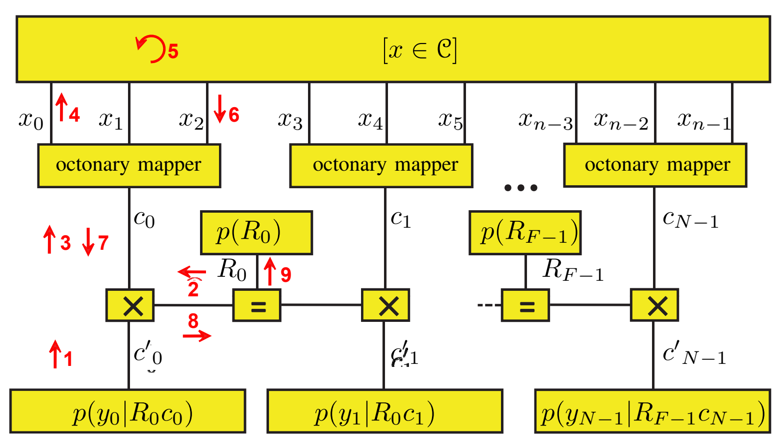

Fig. 2. Message update schedule.

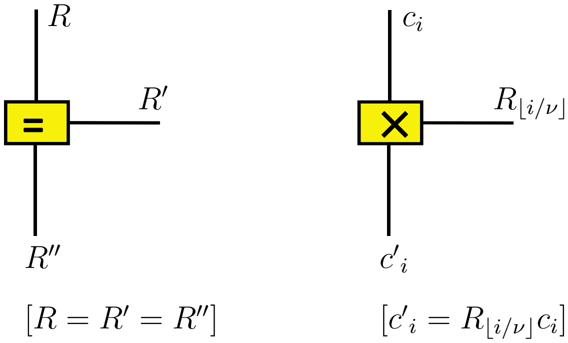

Fig. 3. Repetition and product functional blocks.

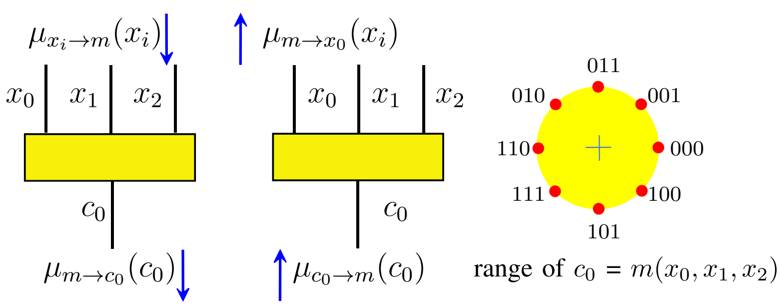

Fig. 4. Messages generated at the mapper.

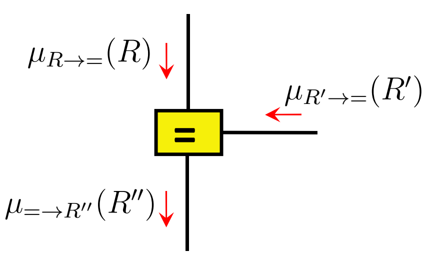

Fig. 5. Messages generated at the equality node.

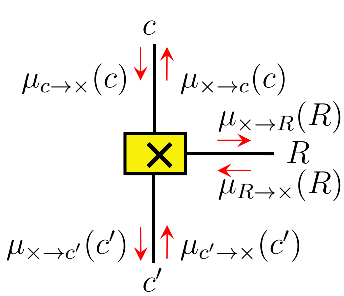

Fig. 6. Messages generated at the product node.

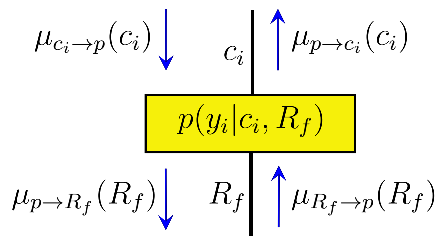

Fig. 7. A message computation.

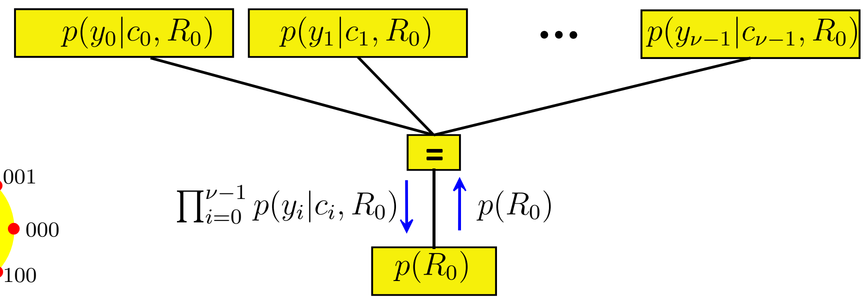

Fig. 8. Factor graph for trained estimation.

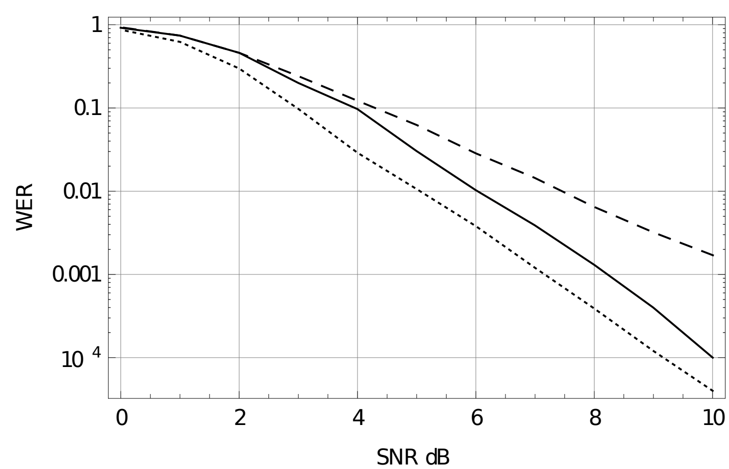

Fig. 9. Simulation results. Word error probability (WER) versus signal-to-noise ratio (SNR) achieved by a binary $(3, 6)$ root-LDPC code of length 200 used for transmission over a block-fading channel with $F = 2$ and fading coefficients 1.0 and 0.5 under three scenarios: Perfect CSI at receiver (dotted curve), joint channel estimation and decoding (full curve) and no CSI at receiver (dashed curve).