Relative Entropy based Message Combining for Exploiting Diversity in Information Optimized Processing

Tobias Monsees, Dirk Wübben and Armin Dekorsy
Department of Communications Engineering
University of Bremen, 28359 Bremen, Germany
Email: {tmonsees, wuebben, dekorsy}@ant.uni-bremen.de

***Abstract*—Maximizing information has become a powerful technique for the design of efficient receiver components with very low bit resolution. In the established information processing approach, the algorithmic tasks are executed on discrete messages and the processing steps are designed to optimize the mutual information. In this paper, we extend the concept of information optimized processing for exploiting diversity. To that end, we propose the Relative Entropy based Message Combining (REMC) approach in order to merge discrete messages with different underlying distributions, e.g., stemming from different diversity branches. We exemplary evaluate the proposed REMC for a single-user uplink-model with multiple Radio Access Points (RAPs) and higher order modulation schemes. The numerical results show that message combining is required to leverages diversity gains in an information optimized receiver.**

I. Introduction

Clustering to maximize Mutual Information (MI) via Information Bottleneck Method (IBM) [1] has become a powerful technique for the design of information optimized channel quantizers [2]. The information optimized design approach has also been successfully utilized to design discrete receiver concepts optimized for coarsely quantized signals [3], [4]. Since a large part of the processing complexity lies in the decoding part, especially efficient decoder designs are of great interest. Information optimized decoder implementations offer excellent performance with a very low bit resolution [5]–[8], suitable for efficient implementations [9]. It has been demonstrated, that such designed discrete decoders with only 3 or 4 bits per variable can perform very close to floating-point decoder implementations. The low bit resolution greatly reduces the implementation complexity in terms of storage and interconnections of internal functions [10].

In an information optimized decoder implementation, the representatives of variables are interpreted as abstract messages with a specific probabilistic interpretation, i.e. they can be substituted by any desired representative without changing its probabilistic relations and therefore the value of information they preserve. To generalize the information optimized design procedure for higher-order modulations schemes and irregular Low Density Parity Check (LDPC) codes, the concept of message alignment was proposed in [11], [12]. The information-optimized decoder is commonly designed for one specific design channel [13] that shows good performance even if the actual channel is mismatched to the design channel [7]. Distributed RAP processing offers a flexible assignment of functionalities between spatially distributed RAPs and centralized joint processing enables to leverage spatial diversity and improve reliability [14]. However, if receive diversity is available, the reliability of different RAPs has to be taken into account in order to exploit diversity. The proposed centralized message combining transforms messages from different RAPs into a combined message that is matched to an information optimized receiver design.

The main contribution of this paper[1] is the Relative Entropy based Message Combining (REMC) approach that leverages joint processing gains for an information-optimized decoder and also allows to reuse the same implementation for diversity schemes and higher-order modulation schemes.

The remainder of this paper is structured as follows. The system model is introduced in Section II. The information-optimized processing is discussed in Section III. Section IV reviews state-of-the-art concepts and introduces the proposed REMCs. In the numerical evaluation in Section V, we focus on the decoding of regular LDPC codes. We show that the performance of this approach is close to the performance of floating-point sum-product decoding with double precision. The final conclusion is given in Section VI.

II. Distributed RAP Processing

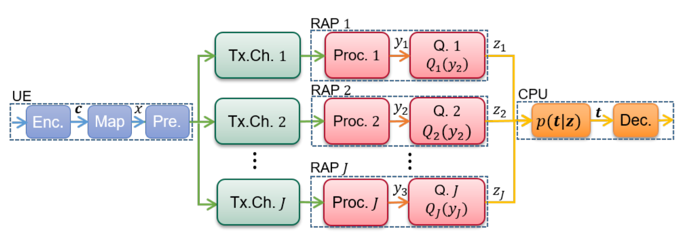

Fig. 1. Single-user uplink-model with J distributed RAPs and final data estimation in the CPU

[1] *Notation:* Random variables are denoted by sans-serif letters x, random vectors by bold sans-serif letters $\boldsymbol{\mathsf{x}}$, realizations by serif letters x and vector valued realizations by bold serif letters $\mathbf{x}$. Sets are denoted by calligraphic letters $\mathcal{X}$. The distribution $p_{\mathsf{x}}(x)$ of a random variable x is abbreviated as $p(x)$.

In order to demonstrate the general concept of message combining for diversity exploitation we consider a single-user uplink system where the user equipment (UE) signal is observed by J distributed RAPs as visualized in Fig. 1. Each RAP forwards a compressed representation of its local observation to the CPU for estimating the source message based on all observations [15]–[17].

In the UE, the source message is encoded by a linear channel code, e.g., an LDPC code and mapped onto symbols from the modulation alphabet $\mathcal{X}$ of size M. By omitting the time index for simplicity, the mapper assigns to any binary code bit vector $\mathbf{c} \in \mathbb{F}_2^m$ of length $m = \log_2 M$ one transmit symbol $x \in \mathcal{X}$. Before transmission, further processing (e.g., transformation to time domain for Orthogonal Frequency Division Multiplexing (OFDM) and insertion of pilot symbols) is performed. In the RAPs, the noisy observations are again processed (e.g., transformation to frequency domain and sub-carrier wise equalization for OFDM) yielding the local observation $y_j \in \mathcal{Y}_j$ for the transmitted symbol x. Subsequently, these observations y_j are quantized/compressed in order to reduce the fronthaul (FH) data rate for forwarding the local observations to the CPU for final data estimation (e.g. decoding the LDPC code).

For appropriate processing in the UE and in the RAPs the local observation y_j depends only on the single transmit symbol x and the dependency can be modelled by an equivalent Single-Input Single-Output (SISO) channel [18], [19] with fading-coefficient $\mathsf{h}_j \sim \mathcal{N}_{\mathrm{C}}(0,1)$. Thus, the influence of the signal pre-processing at the UE, the physical transmission channel and the signal processing at the RAPs for each pair of transmit and receive signals is described by

$$y_j = h_j x + n_j \tag{1}$$

with the *effective* Signal-to-Noise Ratio (SNR) $\mathrm{SNR}_j = |h_j|^2 \sigma_{\mathsf{x}}^2 / \sigma_{\mathsf{n}_j}^2$. Here, σ_{x}^2 and $\sigma_{\mathsf{n}_j}^2$ denote the variance of the transmit signal and the effective noise realization, respectively. Thus, the influence of varying SNRs per RAPs or subcarriers are also incorporated in the corresponding effective SNR. Subsequently, the effective transmission model (1) will be denoted as *access channel* and the relation between x and y_j is described by the distribution $p(y_j|x)$. Prior to forwarding local observations to the CPU, the observation y_j is quantized into a message $z_j \in \mathcal{Z}_j$ from the discrete alphabet $\mathcal{Z}_j$ by the quantizer function $z_j = Q_j(y_j)$ (details for the design of the quantizer will be provided in Section III-A). Subsequently, each RAP forwards its message z_j via rate-limited FH links to the CPU. The CPU utilizes the proposed REMC unit that yields a combined message $t_\nu \in \mathcal{T}$ for each code bit c_ν with index $\nu = 1, ..., m$ of one modulated symbol x based on all received messages z_j, $1 \le j \le J$.

III. Information-Optimized Processing

A. Mutual Information Optimized Quantizer Design

In order to forward compressed messages for the local observations y_j to the CPU, a joint design of the local quantizers would be desirable [20]. However, here we restrict ourself to an independent design of the local quantizers Q_j per branch j such that the MI[2] $I(\mathsf{x};\mathsf{z}_j)$ between the source symbol x and the quantizer output z_j per RAP is maximized. We assume that the source distribution $p(x)$ is fixed, implying that the achievable rate of the channel between the symbol x and the quantizer output z_j is given by $I(\mathsf{x};\mathsf{z}_j)$ [22]. The objective function of the quantizer design problem Q_j is formalized as

$$Q_j^\star = \underset{Q \in \mathcal{Q}}{\operatorname{argmax}}\, I(\mathsf{x};\mathsf{z}_j) \quad \text{s.t. } |\mathcal{Z}_j| \le N_j\ , \tag{2}$$

for $j = 1, ..., J$, where $\mathcal{Q}$ is the set of all possible quantizer mappings and N_j denotes the upper bound on the cardinality of the set $\mathcal{Z}_j$ in branch j. We consider rate limitations of individual FH links by limiting the cardinality of $\mathcal{Z}_j$ to N_j such that at most $\log_2 N_j$ bits are forwarded by RAP j allowing for unequal cardinalities $|\mathcal{Z}_j|$ for $j = 1, \ldots, J$. The objective in (2) is a special case of the IBM [1] with trade-off parameter $\beta \to \infty$. For (2) we can restrict our design to deterministic quantizers, i.e., $p(z_j|y_j) \in \{0,1\}$, $\forall \{z_j, y_j\} \in \mathcal{Z}_j \times \mathcal{Y}_j$, such that every realization in $\mathcal{Y}_j$ is mapped to only one specific cluster in $\mathcal{Z}_j$ [23]. The quantization $Q_j^\star$ can be formalized as probability mass function $p(z_j|y_j)$ such that under the presumed Markovian relation $\mathsf{x} - \mathsf{y}_j - \mathsf{z}_j$, the *meaning* of z_j w.r.t. x is given by

$$p(x|z_j) = \frac{\sum_{y_j \in \mathcal{Y}_j} p(z_j|y_j) p(y_j|x) p(x)}{\sum_{x \in \mathcal{X}} \sum_{y_j \in \mathcal{Y}_j} p(z_j|y_j) p(y_j|x) p(x)}\ . \tag{3}$$

The meaning $p(x|z_j)$ of message z_j per RAP j depends on the individual access channel $p(y_j|x)$ (i.e., on the individual SNR_j) and the quantizer mapping $p(z_j|y_j)$. Furthermore, in case of higher-order modulation schemes, the individual message meaning w.r.t. bit c_ν of one modulated symbol x is given by the distribution $p(c_\nu|z_j)$ with

$$p(c_\nu = \xi | z_j) = \sum_{x \in \mathcal{X}_\nu^\xi} p(x|z_j) \ \text{ for } \ \xi \in \{0,1\} \tag{4}$$

and $\mathcal{X}_\nu^\xi = \{x(\mathbf{c}) \in \mathcal{X} \,|\, c_\nu = \xi\}$. In general, the meanings $p(c_\nu|z_j)$ are different for unequal bit levels c_i and c_j with $i \neq j$.

Fig. 2 visualizes the different meanings of z_j by means of Log-Likelihood Ratios (LLRs) $L(c_\nu|z_j) = \ln \frac{p(c_\nu=0|z_j)}{p(c_\nu=1|z_j)}$ for two RAPs with $\mathrm{SNR}_1 = 5$ dB and $\mathrm{SNR}_2 = 15$ dB in case of 4-ASK mapping and SNR optimal quantizers $Q_1^\star$ and $Q_2^\star$. Obviously, the same cluster index has quite different meanings. For example, in case of $z_1 = 3$ the LLRs are given by $L(c_1|z_1 = 3) \approx 1.5$ and $L(c_2|z_1 = 3) \approx -0.5$, whereas for $z_2 = 3$ the LLRs are $L(c_1|z_2 = 3) \approx 5.1$ and $L(c_2|z_2 = 3) \approx -4.2$ for c_1 and c_2, respectively. Consequently, the question arises how to process the different discrete messages jointly in the centralized estimation step in the CPU. The resulting message combining problem will be discussed in detail in Section IV.

[2] The mutual information between two random variables a and b is given by $I(\mathsf{a};\mathsf{b}) = \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} p(a,b) \log_2 \frac{p(a,b)}{p(a)p(b)}$ [21].

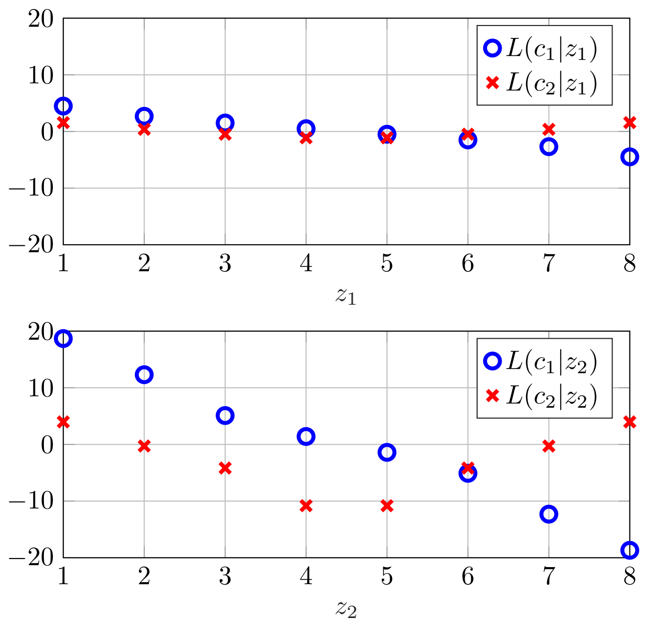

Fig. 2. Gray-coded 4-ASK with equiprobable $p(x)$ (i.e. $\sigma_x^2 = 5$) and AWGN with variance $\text{SNR}_1 = 5$ dB (top) and $\text{SNR}_2 = 15$ dB (bottom), both quantized into 8 clusters

B. Information-Optimized Message Passing Decoder

As an alternative to the Floating-Point Sum-Product Algorithm (FP-SPA) [24] for decoding LDPC codes, decoder realizations with low bit resolution using mutual information maximizing message mappings have been proposed [13]. The designed discrete decoder is optimized for the processing of discrete signals and all internal operations can be implemented by simple Lookup Tables (LUTs). For binary codes the so called Lookup-Table based Message Passing (LUT-MP) decoder performs with a message resolution of only 3 or 4 bit is close to the FP-SPA with 64-bit double precision [5]. The key idea for the decoder design is the concept of discrete density evolution [13] that yields discrete decoder functions that maximize mutual information for a fixed design distribution $p^\star(c,t)$ between the binary input c and the compression variable $t \in \mathcal{T}$ of cardinality $|\mathcal{T}|$. Usually, the design distribution $p^\star(c,t)$ at the threshold SNR of the code is used for the design yielding good performance even if the transmission channel is mismatched to the design channel [7].

IV. Relative Entropy based Message Combining (REMC)

A. Message Alignment

In Section III-A the problem, that discrete values with varying meaning for the signals of interest (i.e., transmit symbols x or code bits c_ν) need to be handled accordingly has been illustrated. To this end, the Message Alignment (MA) approach for the alignment of clusters with similar probabilistic meaning that enables the design of discrete LDPC decoders for higher-order modulation schemes has been proposed [11]. For designing discrete decoders for irregular LDPC codes the iterative MA has been presented [12] to align messages w.r.t. an averaged degree distribution. Furthermore, iterative MA has also been applied for distributed wireless sensor networks where local sensors forward an aligned message to the fusion center to enable MAP-detection on an averaged distribution [25] without transmitting additional sensor information to the fusion center. However, the MA within each local sensor node requires the knowledge of an averaged distribution that depends on the access channel of all other sensor nodes and might be problematic in case of highly dynamic networks. In contrast to that, we propose a combining approach that enables an optimized central message combining and considers also different rate limitations, i.e., $|\mathcal{Z}_j|$ might be different per branch j.

B. Message Combining

The general objective of message combining is to generate a meaning $\mathbf{t}$ for the message vector $\mathbf{z} = [z_1, ..., z_J] \in \mathcal{Z}_1 \times ... \times \mathcal{Z}_J$. To that end, it is necessary to specify a target distribution $p(\cdot|\mathbf{t})$. As shown in (3), the meaning $p(x|z_j)$ of the message $z_j \in \mathcal{Z}_j$ of RAP j depends on the quantizer mapping $p(z_j|y_j)$ and the access channel $p(y_j|x)$. For our setup, we select the index t_ν where the meaning $p(c_\nu|\mathbf{z})$ of the message vector $\mathbf{z}$ is similar to the design distribution of the discrete decoder $p^\star(c|t = t_\nu)$, i.e. $p(c_\nu = \xi|\mathbf{z}) \approx p^\star(c = \xi|t = t_\nu)$ for $\xi \in \{0, 1\}$. The meaning $p(c_\nu|\mathbf{z})$ of the message vector $\mathbf{z}$ is given under the presumed Markovian assumption $\mathsf{x} - \mathbf{y} - \mathbf{z}$ by

$$p(c_\nu = \xi|\mathbf{z}) = \sum_{x \in \mathcal{X}_\nu^\xi} \frac{\sum_{\mathbf{y} \in \mathcal{Y}'} p(\mathbf{z}|\mathbf{y}) p(\mathbf{y}|x) p(x)}{\sum_{x' \in \mathcal{X}} \sum_{\mathbf{y} \in \mathcal{Y}'} p(\mathbf{z}|\mathbf{y}) p(\mathbf{y}|x') p(x')}, \tag{5}$$

with $\mathcal{Y}' = \mathcal{Y}_1 \times ... \times \mathcal{Y}_J$, $p(\mathbf{z}|\mathbf{y}) = \prod_j p(z_j|y_j)$ and $p(\mathbf{y}|x) = \prod_j p(y_j|x)$ for $\xi \in \{0, 1\}$. The meaning in (5) depends on all quantizer mappings $p(z_j|y_j)$ and all access channels $p(y_j|x)$.

Similar to MA, we utilize the relative entropy[3] $D_{\text{KL}}(\cdot\|\cdot)$ or Kullback-Leiber divergence to find an aligned message related to a target posterior distribution. The relative entropy is a natural measure that occurs in case of mutual information optimized receiver processing [26].

In our proposed REMC, we translate the meaning $p(c_\nu|\mathbf{z})$ of the combined message vector $\mathbf{z}$ into a meaning $p^\star(c|t)$ of t, that is determined by the design distribution of the discrete decoder, i.e. the translated message $t_\nu(\mathbf{z})$ is given by

$$t_\nu(\mathbf{z}) = Q_{C,\nu}(\mathbf{z}) = \operatorname*{argmin}_{t \in \mathcal{T}} D_{\text{KL}}(p(c_\nu|\mathbf{z}) \| p^\star(c|t)) \tag{6}$$

for every bit c_ν, $\nu = 1, ..., m$ in one modulated symbol x. Consequently, message combining in (6) finds the combinded messages $t_\nu(\mathbf{z})$ for every bit c_ν that minimizes the missmatch between the meaning of $p(c_\nu|\mathbf{z})$ and $p^\star(c|t = t_\nu)$, measured by the relative entropy. The resulting deterministic combining mapping $\mathcal{Z}_1 \times ... \times \mathcal{Z}_J \to \mathcal{T}$ is implemented as a REMC LUT that generally also depends on the individual access channels

[3]The relative entropy or Kullback-Leibler divergence between two distributions $p(x_1)$ and $p(x_2)$ with the same set of possible outcomes $\mathcal{X}$ is defined as $D_{\text{KL}}(p(x_1)\|p(x_2)) = \sum_{x \in \mathcal{X}} p(x_1 = x) \log_2 \frac{p(x_1=x)}{p(x_2=x)}$

$p(y_j|x)$ and quantizer mappings $p(z_j|y_j)$ for $j = 1, ..., J$. However, as we will discuss in the next subsection, the output of the REMC LUT is highly informative even if we use only a small number REMC LUTs and select the one that best fits the actual channel conditions. The combined message t_ν is forwarded to the information-optimized decoder (c.f. Section III-B). Note that the message combining in (6) also provides a general framework to combine messages of a source related to a target distribution, i.e. $p^\star(c|t)$ can be substituted by any desired (possibly vector valued) target distribution $p(\cdot|\mathbf{t})$ of further information optimized processing steps. We also note that the message combining is not restricted to individual quantizer design per RAP, i.e. it can also be applied in case of joint quantizer design [27] and also for generalized quantizer design measures [28].

C. Preserved Information after Message Combining

In order to analyze the impact of centralized message combining with REMC we consider a system with $J = 2$ RAPs and transmission of BPSK symbols x. The effective SNR of the first branch has been fixed to $\mathrm{SNR}_1 = 1.55\,\mathrm{dB}$ and SNR_2 of the second branch is varied. Each RAP quantizes the observations y_j with a fixed 3-bit quantizer which has been optimized for a design SNR of $\mathrm{SNR}_{\mathrm{Design}} = 1.55\,\mathrm{dB}$. Figure 3 shows the resulting Mutual Information (MI) terms.

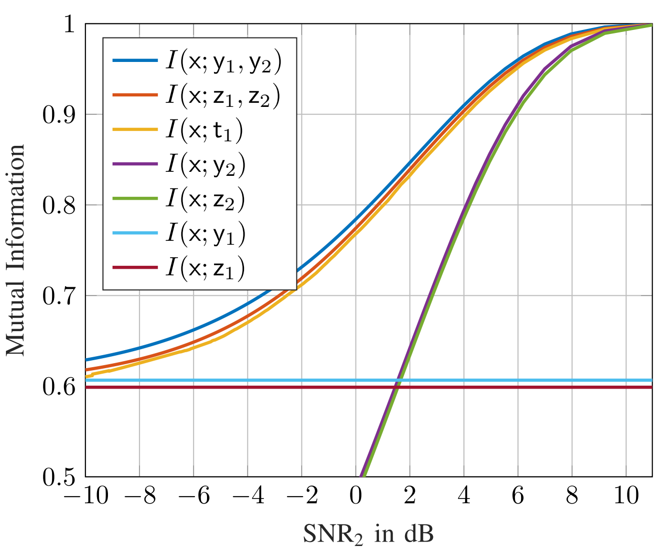

Fig. 3. Information rates for two diversity branches with fixed quantizers at RAP and message combining at CPU

$I(\mathsf{x};\mathsf{y}_1)$ and $I(\mathsf{x};\mathsf{y}_2)$ indicate the MI between the source and the individual observations in the RAPs, whereas $I(\mathsf{x};\mathsf{y}_1,\mathsf{y}_2)$ denotes the *total available* MI if both observations are considered jointly. Clearly, the gain of a joint processing is demonstrated and $I(\mathsf{x};\mathsf{y}_1,\mathsf{y}_2)$ serves as an upper bound for the total MI *after quantization*, $I(\mathsf{x};\mathsf{z}_1,\mathsf{z}_2)$. Although, per RAP the same quantizers optimized for a fixed $\mathrm{SNR}_{\mathrm{Design}} = 1.55\,\mathrm{dB}$ are applied, the loss induced by 3-bit quantization is quite small w.r.t. the upper bound. The REMC approach (6) yields in the CPU for z_1 and z_2 a combined message t_1 which is aligned to the given meaning $p^\star(c|t)$. Here, $p^\star(c|t)$ has been selected such that it corresponds to the threshold $\mathrm{SNR}_{\mathrm{Threshold}} = 1.55\,\mathrm{dB}$ for ($d_V = 3$, $d_C = 6$)-regular LDPC code [18]. By investigating the MI $I(\mathsf{x};\mathsf{t}_1)$ between the source signal x and the REMC output t_1 only a small loss compared to the upper bound $I(\mathsf{x};\mathsf{z}_1,\mathsf{z}_2)$ is visible. Thus, the varying meaning of the compressed signals z_1 and z_2 is successfully considered by the REMC if individual access distributions and the applied quantizers (or, alternatively the meaning $p(x|z_j)$) are known at the CPU.

V. Performance Evaluations

In this section we investigate the Bit Error Rate (BER) performance for a coded system applying a very short (3, 6)-regular LDPC code of length $N_C = 96$ and code rate $R_C = 0.5$. The channel between the UE and the varying number of RAPs are modelled as independent block Rayleigh fading channels. As a benchmark we consider the direct forwarding of observations to the CPU, combining of individual LLR and decoding by the FP-SPA. For the quantized systems each RAP forwards 3-bit messages per observation to the CPU. If, once again, FP-SPA is used for decoding, the CPU calculates LLRs based on the compressed signals z_j and performs LLR combining prior to decoding. Alternatively, the REMC approach is applied for combining the disrete messages and the LUT-MP decoder optimized for the threshold $\mathrm{SNR}_{\mathrm{Threshold}} = 1.55\,\mathrm{dB}$ is used. Subsequently, we will first analyze a mismatched design for the quantized system indicating a severe performance loss as the diversity is not exploited. Then, the approach with REMC is considered for BPSK and 16-QAM, respectively.

A. Missmatched Message Combining

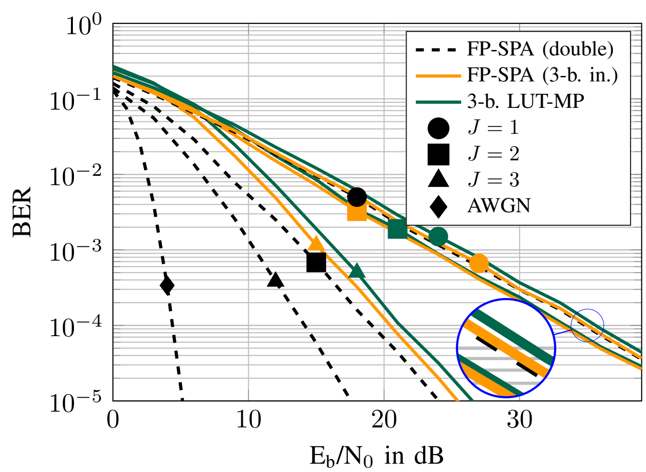

Fig. 4. BER performance with J RAPs applying per RAP fixed 3-bit quantizer and fixed message combining

In this setup it is assumed, that fixed quantizers Q designed for the $\mathrm{SNR}_{\mathrm{Threshold}} = 1.55\,\mathrm{dB}$ are utilized in every RAP, but the CPUs is not aware of the effective access SNRs. Thus, for LLR combining the fixed e2e distribution $p(c_1|\mathbf{z})$ given by the design SNR and the quantizer mappings $p(\mathbf{z}|\mathbf{y})$ is assumed. Similarly, for REMC only the fixed e2e distribution $p(c_1|\mathbf{z})$ is applied. Correspondingly, the instantaneous realization of the access channel and, thus, the instantaneous effective SNR, is not considered in the processing.

Fig. 4 shows the BER performance of the FP-SPA without quantization and optimal LLR combining (benchmark), FP-SPA with 3-bit fixed IB quantization at the RAPs as well as 3-bit LUT-MP decoding for a varying number of RAPs $J = 1, 2, 3$. The case of FP-SPA without quantization and without fading (i.e. additive white Gaussian noise (AWGN) channel) is added as an additional benchmark.

As expected, for $J = 1$ a significant performance loss due to fading compared to the AWGN case becomes obvious. For two RAPs ($J = 2$) the system without quantization shows the expected diversity gain in the slope of the BER curve leading to substantial performance improvements. However, both systems with quantization schemes do not realize any diversity gain and show only a small improvement compared to $J = 1$. Both approaches do not consider the reliability of the instantaneous realizations leading to a this severe loss in performance. Although, for $J = 3$ the schemes with quantized FH signals realize some diversity gain, the loss compared to the non-quantized FH system is roughly ≈ 7 dB at a BER of 10^{-4}. Obviously, by using only one fixed 3-bit quantizer and ignoring the current influence of the access channel is not sufficient to exploit the available spatial diversity. As the instantaneous e2e distributions of the different RAP signals is not considered, the different reliability of the RAP signals z_j is not exploited in the CPU.

B. Optimized Message Combining

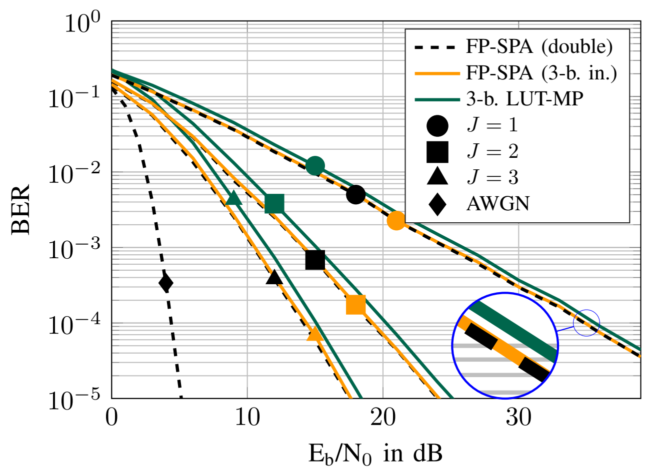

Fig. 5. BER performance with RAPs applying SNR-adapted 3-bit quantizer per RAP and message combining in CPU for $J > 1$

Fig. 5 shows the BER performance if the RAPs forward also their instantaneous SNR_j to the CPU such that the LLR combining or the REMC can be performed based correct e2e distribution $p(c_1|\mathbf{z})$. For $J = 1$ almost no difference compared to Fig 4 is visible. In case of $J = 2$, FP-SPA with quantizer selection at the individual RAPs and LLR calculation/combining based on the e2e distributions achieves the same diversity gain and only a marginal SNR loss compared to the benchmark without any quantization. The LUT-MP decoder with REMC at the CPU considers also the reliability of the two different RAP signals, leading to the same diversity gain and only a small performance loss of ≈ 1.1 dB at a BER of 10^{-4} compared to the benchmark. For $J = 3$, the LUT-MP decoder with message combining and the FP-SPA with LLR combining also achieves the same diversity gain compared to the benchmark. Furthermore, relative performance of the LUT-MP with message combining and the FP-SPA using LLR combining is quite similar to the case of $J = 2$.

C. Optimized Message Combining for 16-QAM Modulation

Subsequently, we investigate the performance in case of 16-QAM modulation. We analyse the forwarding of to 6-bit per complex-valued rx signal over the FH, while the FP-SPA or the 3-bit LUT-MP decoder is used.

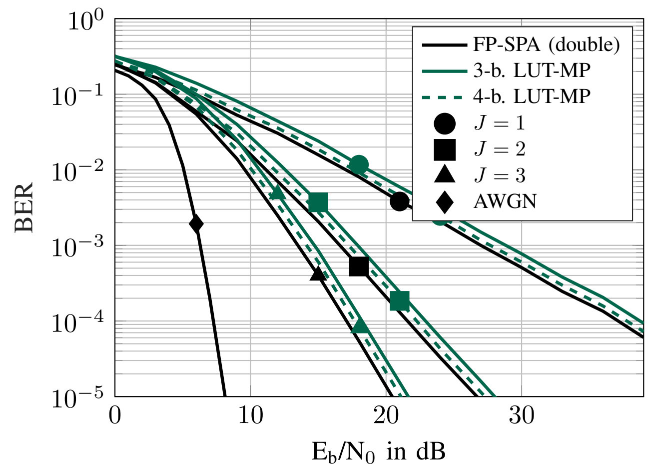

Fig. 6. BER performance with RAPs applying SNR-adapted 6-bit quantizer per RAP and message combining in CPU for $J > 1$ for 16-QAM modulation

A comparison between the 3-bit LUT-MP and 4-bit LUT-MP for 6-bit channel quantization is shown in Fig. 6. The 4-bit LUT-MP achieves at a BER of 10^{-3} a performance gain of ≈ 1 dB for $J = 1$ and ≈ 0.6 dB for $J = 2, 3$. The performance improvement can be further increased by increasing the number of bits of the LUT-MP. Hence, the e2e performance by using a low bit resolution for the forwarding of I/Q data via the FH and the joint processing at the CPU (REMC and LUT-MP decoding) is very close to the benchmark without quantization and FP-SPA decoding.

VI. Conclusion

We introduced a relative entropy based message combining approach that generalizes an information-optimized receiver to leverage joint processing gains by exploiting receive diversity. The proposed message combining enables a flexible usage an information-optimized receiver implementation for higher-order modulation schemes that considers also the possibly different rate limitations of the corresponding forward links.

Acknowledgment

This work was partly funded by the German ministry of education and research (BMBF) under grant 16KIS0720 (TACNET 4.0) and 16KIS1180K (FunKI).

References

[1] N. Tishby, F. C. Pereira, and W. Bialek, “The Information Bottleneck Method,” in *37th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, Sep. 1999, p. 368–377.

[2] S. Hassanpour, D. Wübben, and A. Dekorsy, “Overview and Investigation of Algorithms for the Information Bottleneck Method,” in *11th Int. Conference on Systems, Communications and Coding (SCC)*, Hamburg, Germany, Feb. 2017.

[3] J. Lewandowsky, M. Stark, and G. Bauch, “Information Bottleneck Graphs for Receiver Design,” in *2016 IEEE International Symposium on Information Theory (ISIT)*, Barcelona, Spain, Jul. 2016, pp. 2888–2892.

[4] J. Lewandowsky, M. Stark, R. Mendrzik, and G. Bauch, “Discrete Channel Estimation by Integer Passing in Information Bottleneck Graphs,” in *11th International ITG Conference on Systems, Communications and Coding (SCC)*, Hamburg, Germany, Feb. 2017.

[5] F. J. C. Romero and B. M. Kurkoski, “LDPC Decoding Mappings That Maximize Mutual Information,” *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 9, pp. 2391–2401, Sep. 2016.

[6] A. Balatsoukas-Stimming, M. Meidlinger, R. Ghanaatian, G. Matz, and A. Burg, “A Fully-Unrolled LDPC Decoder based on Quantized Message Passing,” in *2015 IEEE Workshop on Signal Processing Systems (SiPS)*, Hangzhou, China, Oct. 2015.

[7] J. Lewandowsky and G. Bauch, “Information-Optimum LDPC Decoders Based on the Information Bottleneck Method,” *IEEE Access*, vol. 6, pp. 4054–4071, Jan. 2018.

[8] S. A. A. Shah, M. Stark, and G. Bauch, “Coarsely Quantized Decoding and Construction of Polar Codes Using the Information Bottleneck Method,” *Algorithms*, vol. 12, no. 9, Sep. 2019.

[9] R. Ghanaatian, A. Balatsoukas-Stimming, T. C. Müller, M. Meidlinger, G. Matz, A. Teman, and A. Burg, “A 588-Gb/s LDPC Decoder Based on Finite-Alphabet Message Passing,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 2, pp. 329–340, Feb. 2018.

[10] J.-S. Lee and J. Thorpe, “Memory-Efficient Decoding of LDPC Codes,” in *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.*, Adelaide, Australia, Sep. 2005, pp. 459–463.

[11] J. Lewandowsky, M. Stark, and G. Bauch, “Message Alignment for Discrete LDPC Decoders with Quadrature Amplitude Modulation,” in *2017 IEEE International Symposium on Information Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 2925–2929.

[12] M. Stark, J. Lewandowsky, and G. Bauch, “Information-Optimum LDPC Decoders with Message Alignment for Irregular Codes,” in *2018 IEEE Global Communications Conference (GLOBECOM 2018)*, Abu Dhabi, UAE, Dec. 2018.

[13] B. M. Kurkoski, K. Yamaguchi, and K. Kobayashi, “Noise Thresholds for Discrete LDPC Decoding Mappings,” in *2008 IEEE Global Telecommunications Conference (GLOBECOM 2008)*, New Orleans, LA, USA, Dec. 2008.

[14] D. Wübben, P. Rost, J. S. Bartelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, and G. Fettweis, “Benefits and Impact of Cloud Computing on 5G Signal Processing: Flexible Centralization through Cloud-RAN,” *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 35–44, Nov. 2014.

[15] I. Estella Aguerri, A. Zaidi, G. Caire, and S. Shamai Shitz, “On the Capacity of Cloud Radio Access Networks With Oblivious Relaying,” *IEEE Transactions on Information Theory*, vol. 65, no. 7, pp. 4575–4596, Feb. 2019.

[16] D. Chen and V. Kuehn, “Alternating Information Bottleneck Optimization for the Compression in the Uplink of C-RAN,” in *2016 IEEE International Conference on Communications (ICC)*, Kuala Lumpur, Malaysia, Jul. 2016.

[17] S. Hassanpour, T. Monsees, D. Wübben, and A. Dekorsy, “Forward-Aware Information Bottleneck-Based Vector Quantization for Noisy Channels,” *IEEE Transactions on Communications*, vol. 68, no. 12, pp. 7911–7926, Dec. 2020.

[18] T. Monsees, D. Wübben, and A. Dekorsy, “Channel-Optimized Information Bottleneck Design for Signal Forwarding and Discrete Decoding in Cloud-RAN,” in *12th International ITG Conference on Systems, Communications and Coding (SCC)*, Rostock, Germany, Feb. 2019.

[19] J. Demel, T. Monsees, C. Bockelmann, D. Wübben, and A. Dekorsy, “Cloud-RAN Fronthaul Rate Reduction via IBM-based Quantization for Multicarrier Systems,” in *24th International ITG Workshop on Smart Antennas (WSA)*, Hamburg, Germany, Feb. 2020.

[20] S. Hassanpour, D. Wübben, and A. Dekorsy, “Generalized Distributed Information Bottleneck for Fronthaul Rate Reduction at the Cloud-RANs Uplink,” in *2020 IEEE Global Communications Conference (GLOBECOM 2020)*, Taipei, Taiwan, Dec. 2020.

[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.

[22] B. M. Kurkoski and H. Yagi, “Finding the Capacity of a Quantized Binary-Input DMC,” in *2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, Cambridge, USA, Jul. 2012, pp. 686–690.

[23] ——, “Quantization of Binary-Input Discrete Memoryless Channels,” *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4544–4552, Aug. 2014.

[24] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, “Factor Graphs and the Sum-product Algorithm,” *IEEE Trans. Inf. Theor.*, vol. 47, no. 2, pp. 498–519, Sep. 2006.

[25] M. Stark, J. Lewandowsky, and G. Bauch, “Iterative Message Alignment for Quantized Message Passing between Distributed Sensor Nodes,” in *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*, Porto, Portugal, Jun. 2018.

[26] J. A. Zhang and B. M. Kurkoski, “Low-Complexity Quantization of Discrete Memoryless Channels,” in *2016 International Symposium on Information Theory and Its Applications (ISITA)*, Monterey, California, USA, Oct. 2016, pp. 448–452.

[27] S. Hassanpour, D. Wübben, and A. Dekorsy, “Generalized Distributed Information Bottleneck for Fronthaul Rate Reduction at the Cloud-RANs Uplink,” in *2020 IEEE Global Communications Conference (GLOBECOM 2020)*, Taipei, Taiwan, Dec. 2020.

[28] X. He, K. Cai, W. Song, and Z. Mei, “Dynamic Programming for Sequential Deterministic Quantization of Discrete Memoryless Channels,” *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 3638–3651, Jun. 2021.