Information Preserving Quantization and Decoding for Satellite-Aided 5G Communications

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*—We consider the uplink of a non-terrestrial network where a relay node is forwarding digitized signals via a rate limited error-prone forward link to the serving satellite. The focus of our investigations are the design of suitable low-bit resolution quantization schemes to limit the rate on the forward link and the optimization of the decoder processing at the satellite. To this end, we investigate Information Bottleneck (IB) based quantizer design with mutual information as fidelity criterion. The IB approach can be extended to consider also the error statistic of the forward link within the quantizer design. By numerical investigations we demonstrate the performance gains of this forward-aware IB quantizer in case of erroneous forward links. Furthermore, we investigate a lookup-table based decoder which is optimized for the end-to-end statistic including the access link, the quantizer and the forward link. This decoder implementation processes only discrete values using lookup-tables of small size. The numerical results show that the performance of 3-bit forward-aware IB quantizer in combination with a 3-bit discrete decoder implementation is close to the double-precision floating-point belief propagation decoder even for strongly error-prone forward links.**

I. Introduction

The 3rd Generation Partnership Project (3GPP) has identified Non-Terrestrial Networks (NTNs) as a promising solution for the 5G service enablers Massive Machine-type Communication (mMTC) and Enhanced Mobile Broadband (eMBB) and several architecture options are currently under discussions [1], [2]. In particular, the User Equipment (UE) may use a direct access link to communicate with the satellite or, alternatively, an on-ground relay node (RN) provides the access link for the UE and forwards the information between the satellite and the UEs [3].

In this paper, we focus on the Uplink (UL) of this two-hop scenario where the RN provides the access link as depicted in Fig. 1. To this end, it is assumed that the signals received on the access link are processed by the RN and digitized messages are transmitted via the rate limited Forward Link (FL) to the satellite. Here, the application of low-bit quantizers is of significant interest in order to limit the required data rate on the FL. For the satellite, a regenerative payload configuration is assumed where the user messages are estimated by decoding the Forward Error Correction (FEC) [2] code. Thus, this scenario is similar to the UL of a Cloud Radio Access Network (Cloud-RAN) system where Small Cells (SCs) forward UE messages to the Central Processing Unit (CPU) for final processing [4], [5].

978-1-7281-3627-1/19/$31.00 © 2019 IEEE

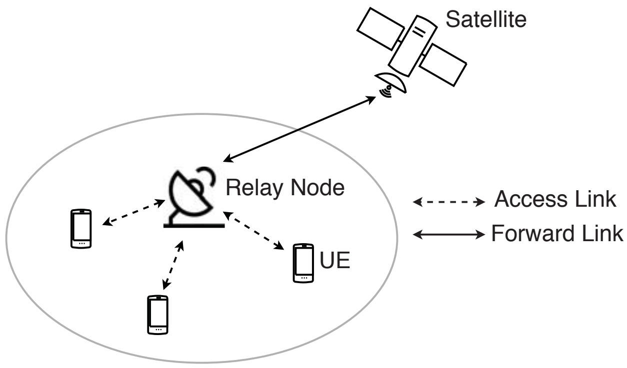

Fig. 1. Considered 5G architecture with relay node (RN) providing the access link for the User Equipments (UEs).

Considering modern FEC schemes like Turbo-Codes or Low Density Parity Check (LDPC) codes, the main computational complexity of the UL processing chain is required by the iterative decoder in the satellite [4]. In order to reduce the decoder complexity, we investigate discrete variants of the Message Passing (MP) decoder for LDPC codes with a low-bit representation for the internal messages [6]–[10]. As all internal operations of this Lookup-Table based Message Passing (LUT-MP) decoder are replaced by Lookup-Tables (LUTs) of small size being efficiently implementable in suitable hardware architectures, this approach promises high throughput [11].

Contributions: The considered communication system implements a classical two-hop transmission with a compress-and-forward relay consisting of an access- and a forward link. In order to transmit compressed signals by the RN, we apply the IB method [12]–[14] to design scalar quantizers at the RN based on the access channel statistic such that the mutual information between the source signal and the RN signal is maximized. In addition, we utilize the Forward-Aware Vector Information Bottleneck (FAVIB) algorithm [15] to consider also the statistic of the forward link in the quantizer design such that the end-to-end (e2e) mutual information between the source signal and the receive signal at the satellite is maximized. Compared to common quantization schemes, both approaches will provide improved e2e performance and increase the robustness against transmission errors on the FL. Furthermore, we extend the LUT-MP approach by optimizing the internal LUTs with respect to the e2e channel statistic given by the access channel, the quantizer, and the forward channel.

Outline: The remainder of this paper is organized as follows. Section II introduces the system model. In Section III we discuss the IB based quantization setup and the forward-aware quantization which considers the access- and forward channel statistics jointly. In Section IV the LUT-MP decoder, which considers the statistics of the access- and forward link jointly, is presented and evaluated by numerical investigations. The paper is summarized in Section V.

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 $\boldsymbol{x}$. Sets are denoted by calligraphic letters $\mathcal{X}$.

II. System Model

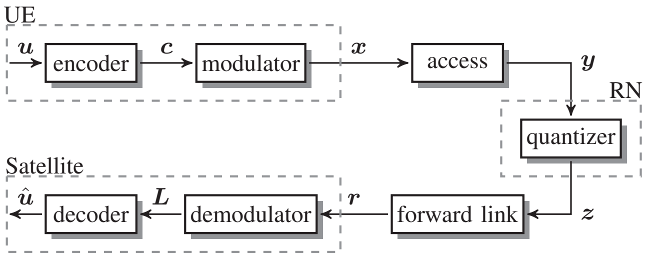

Fig. 2. Blockdiagramm for the uplink transmission of an UE to a satellite via a RN.

The uplink model for one UE is shown in Fig. 2. The UE encodes a binary information word $\boldsymbol{u} \in \mathbb{F}_2^K$ of length K into the code word $\boldsymbol{c} \in \mathbb{F}_2^N$ of a regular LDPC code with length N, where $\mathbb{F}_2$ denotes the binary Galois Field and $R = \frac{K}{N}$ is the rate of the code. The modulated sequence is denoted as $\boldsymbol{x} \in \mathcal{X}^J$, where $\mathcal{X}$ is the finite modulation alphabet and J is the length of the sequence. It is assumed that the source symbols are independent and identically distributed (iid) according to the distribution $p_{\boldsymbol{\mathsf{x}}}(\boldsymbol{x}) = \prod_{j=1}^{J} p_{\mathsf{x}}(x_j)$.

Depending on the used transmission scheme and the influence of the wireless transmission channel the RN processes the receive signals (e.g. down-conversion, equalization) before generating the discrete forward symbols. In order to allow for a general investigation, we describe the influence of the transmission channel and the preprocessing in the RN by a Discrete Memoryless Channel (DMC) with probability mass function (pmf) $p_{\boldsymbol{\mathsf{y}}|\boldsymbol{\mathsf{x}}}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{j=1}^{J} p_{\mathsf{y}|\mathsf{x}}(y_j|x_j)$. Furthermore, without loss of generality, for the numerical investigations this equivalent *access channel* is modelled as a finely quantized Additive White Gaussian Noise (AWGN) channel with output alphabet $\mathcal{Y}$ and noise variance σ_{n}^2 resulting in the transition probabilities $p_{\mathsf{y}|\mathsf{x}}(y|x)$.

The RN maps the access channel output $\boldsymbol{y} \in \mathcal{Y}^J$ by scalar quantization into discrete samples $\boldsymbol{z} \in \mathcal{Z}^J$, where $\mathcal{Z}$ is the set of quantizer labels and $|\mathcal{Z}| \ll |\mathcal{Y}|$ holds in general. The scalar quantizer is uniquely defined by its pmf $p_{\mathsf{z}|\mathsf{y}}(z|y)$ and a deterministic quantizer [16] is a special case where $p_{\mathsf{z}|\mathsf{y}}(z|y) \in \{0,1\}$ for all $(y,z) \in \mathcal{Y} \times \mathcal{Z}$. The corresponding vector-valued mapping is defined by $p_{\boldsymbol{\mathsf{z}}|\boldsymbol{\mathsf{y}}}(\boldsymbol{z}|\boldsymbol{y}) = \prod_{j=1}^{J} p_{\mathsf{z}|\mathsf{y}}(z_j|y_j)$. Due to the compression, the data rate on the FL is at most $\log_2(|\mathcal{Z}|)$ bit/channel use.

The transmission of discrete samples $\boldsymbol{z}$ to the satellite over the FL would also comprise encoding, modulation, channel impact, demodulation, decoding, etc. Again, in order to focus the investigations on the impact of the quantizer design and the transmission errors induced by the FL, all these processing steps are summarized in a second DMC $p_{\boldsymbol{\mathsf{r}}|\boldsymbol{\mathsf{z}}}(\boldsymbol{r}|\boldsymbol{z}) = \prod_{j=1}^{J} p_{\mathsf{r}|\mathsf{z}}(r_j|z_j)$, where $\boldsymbol{r} \in \mathcal{R}^J$ denotes the output sequence of the FL with output alphabet $\mathcal{R}$. For the subsequent numerical investigations a rather simple model to describe the transmission over the FL is applied [17], where each sample z is represented by a bit tuple with $\log_2(|\mathcal{Z}|)$ elements which are transmitted over a Binary Symmetric Channel (BSC) with bit flip probability P_e. The received bit tuple is mapped onto receive samples r with output cardinality $|\mathcal{R}| = |\mathcal{Z}|$.

In the satellite the receive samples $\boldsymbol{r}$ are fed to the decoder in order to estimate the transmitted information word $\boldsymbol{u}$. In case of common Floating-Point based Message Passing (FP-MP) decoder, Log-Likelihood Ratios (LLRs) $\boldsymbol{L} \in \mathcal{L}^J$ are calculated by the demodulator in order to represent a-posteriori information for the code bits as discussed subsequently in Subsection IV-A. In contrast, the LUT-MP decoder processes the samples $\boldsymbol{r}$ directly such that the calculation of LLRs is omitted, as explained in Subsection IV-B.

III. Quantizer Design at Relay Node

A. Quantizer Design based on Information Bottleneck

In contrast to common quantizers, the IB setup considers the statistic of the access channel $p_{\mathsf{y}|\mathsf{x}}(y|x)$ and the source $p_{\mathsf{x}}(x)$ for the quantizer design. Furthermore, it considers the mutual information[1] (MI) $I(\mathsf{x};\mathsf{z})$ between the source x and the quantizer output z as relevant information. The objective is to preserve a maximum amount of relevant information $I(\mathsf{x};\mathsf{z})$ under a constraint on the compression rate $I(\mathsf{y};\mathsf{z})$. The solution of the underlying trade-off is characterized by the IB functional [12]

$$p^{\star}_{\mathsf{z}|\mathsf{y}}(z|y) = \underset{p_{\mathsf{z}|\mathsf{y}}(z|y)}{\operatorname{argmin}} \frac{1}{\beta+1}\big(I(\mathsf{y};\mathsf{z}) - \beta I(\mathsf{x};\mathsf{z})\big) \ \text{ s.t. } |\mathcal{Z}| \leq N_{\mathsf{z}}\,, \tag{1}$$

where the parameter $\beta \geq 0$ controls the trade-off between compression rate $I(\mathsf{y};\mathsf{z})$ and relevant information $I(\mathsf{x};\mathsf{z})$ and the parameter N_{z} is an upper bound on the number of clusters $|\mathcal{Z}|$. The quantizer labels $\mathcal{Z}$ are immaterial since the objective function depends only on the statistic $p_{\mathsf{z}|\mathsf{y}}$. The optimization problem is neither convex nor concave in general and several heuristics exist to find local optimal solutions [14]. Subsequently, we assume $\beta \to \infty$, which leads to the maximization of the relevant information, i.e.

$$p^{\star}_{\mathsf{z}|\mathsf{y}}(z|y) = \underset{p_{\mathsf{z}|\mathsf{y}}(z|y)}{\operatorname{argmax}}\, I(\mathsf{x};\mathsf{z}) \ \text{ s.t. } |\mathcal{Z}| \leq N_{\mathsf{z}}\,. \tag{2}$$

[1] The mutual information [18] between two random variables a and b is given by $I(\mathsf{a};\mathsf{b}) = \sum_{a\in\mathcal{A}} \sum_{b\in\mathcal{B}} p_{\mathsf{a},\mathsf{b}}(a,b) \log \frac{p_{\mathsf{a},\mathsf{b}}(a,b)}{p_{\mathsf{a}}(a)p_{\mathsf{b}}(b)}$

In case of a deterministic quantizer mapping, the compression rate $I(\mathsf{y};\mathsf{z}) = H(\mathsf{z}) \leq \log_2(|\mathcal{Z}|)$ is determined by the Shannon entropy $H(\mathsf{z}) = -\sum_{z\in\mathcal{Z}} p_{\mathsf{z}}(z)\log_2(p_{\mathsf{z}}(z))$. The optimization problem belongs to the class of concave problems since it is a *convex maximization* problem. One can show that an optimal solution of deterministic type exists and for the special case of a binary input alphabet, an optimal solution can be found via dynamic programming [16].

B. Forward-Aware Quantization

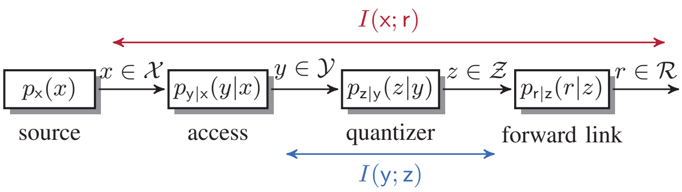

Fig. 3. Forward-aware quantizer design framework.

The IB framework can be extended to consider also the FL statistic $p_{\mathsf{r}|\mathsf{z}}(r|z)$ (see Fig. 3) in the quantizer design. In [15], the underlying variational problem was established and the derived Forward-Aware Vector Information Bottleneck (FAVIB) algorithm realizes a full extension of the IB method. In case of an error-free FL ($P_e = 0$), FAVIB and IB determine the same quantizer for given β. Subsequently, we consider again the special case $\beta \to \infty$ such that the e2e mutual information $I(\mathsf{x};\mathsf{r})$ between the transmitted signal x and the received signal r is maximized

$$p^{\star}_{\mathsf{z}|\mathsf{y}}(z|y) = \underset{p_{\mathsf{z}|\mathsf{y}}(z|y)}{\operatorname{argmax}} \; I(\mathsf{x};\mathsf{r}) \;\; \text{s.t.} \; |\mathcal{Z}| \leq N_{\mathsf{z}}\,, \tag{3}$$

leading to deterministic quatizers. As shown in [15], the optimization problem in (3) is equivalent to the Channel-Optimized Vector Quantization (COVQ) problem with MI as fidelity criterion [19]. Furthermore, a comparable problem setup has been analyzed in [17], where the authors developed an algorithm to maximize $I(\mathsf{x};\mathsf{r})$ by using a continuous model of the access channel.

IV. Satellite Processing

A. Demodulation

After the transmission of discrete values $\boldsymbol{z}$ over the FL, the satellite performs FEC decoding based on the discrete received signals $\boldsymbol{r}$ to estimate the user message $\boldsymbol{u}$. It is assumed, that the satellite has perfect knowledge about the e2e statistic $p_{\mathsf{r}|\mathsf{x}}(r|x)$ between the UE and the satellite given by

$$p_{\mathsf{r}|\mathsf{x}}(r|x) = \sum_{z\in\mathcal{Z}} p_{\mathsf{z}|\mathsf{x}}(z|x)\, p_{\mathsf{r}|\mathsf{z}}(r|z). \tag{4}$$

In case of BPSK modulation, the a-posteriori statistic for the codebits is given by $p_{\mathsf{r}|\mathsf{c}} = p_{\mathsf{r}|\mathsf{x}}$ and the LLRs are calculated by

$$L(r) := \log\left(\frac{p_{\mathsf{r}|\mathsf{c}}(r|0)}{p_{\mathsf{r}|\mathsf{c}}(r|1)}\right). \tag{5}$$

Please notice, that only discrete LLRs occur since r is a discrete random variable. Subsequently, the discrete LLRs are the input of the FP-MP decoder to estimate the transmitted information word $\boldsymbol{u}$.

B. LUT based MP Decoder

In order to optimize the decoder processing for discrete values, we utilize the discrete Density Evolution (DE) algorithm of [6] to generate LUT based decoding functions for a regular LDPC code with column weight d_v and row weight d_c. The LUT-MP decoder is optimized based on the e2e statistic $p_{\mathsf{r}|\mathsf{c}}(r|c)$ for every decoder iteration with design parameter σ_{n}, since we assume that the FL is fixed with parameter P_e. The LUTs are optimized for the noise threshold $\sigma^{\star}_{\mathsf{n}}$ of a (d_v,d_c)-regular LDPC code, i.e., where the mutual information $I(\mathsf{u};\hat{\mathsf{u}})$ between the transmitted UE message u and the estimated message at the satellite $\hat{\mathsf{u}}$ approaches one for the maximum number of iterations $i_{\max}$. The LUT-MP decoder has a three decomposed LUTs for every iteration [9], i.e. for variable nodes, check nodes and for the final decision. Since the LUT-MP decoder is designed based on the e2e statistic $p^{\star}_{\mathsf{r}|\mathsf{c}}$, it processes the discrete values of $\boldsymbol{r}$ directly, such that no LLR calculation step is required.

C. Performance Results

The numerical investigation shows the impact of quantization at the RN on the decoding performance at the satellite and the difference between the LUT-MP and the FP-MP decoder. We assume BPSK modulation such that the access channel is given by a finely quantized real valued AWGN channel, where the first and last boundary are set to $\mp(1 + 3\sigma_{\mathsf{n}})$. We use a regular LDPC code [20] with column weight $d_v = 3$ and row weight $d_c = 6$ of length $N = 2640$ and code rate $R = 0.5$ from [21].

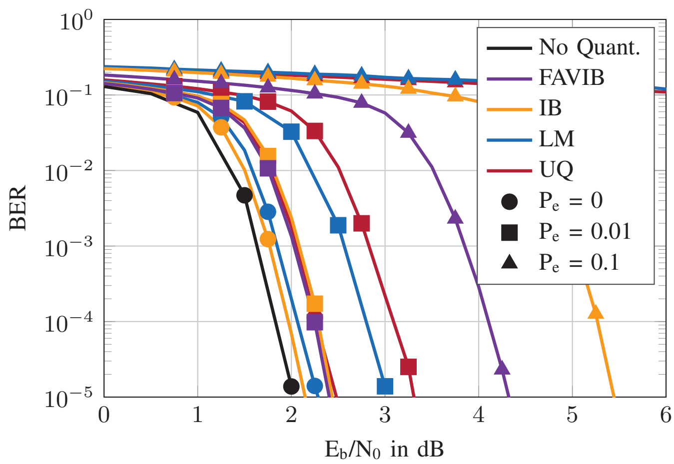

Fig. 4. BER simulations of the FP-MP decoder with LLR calculation (5) and 3-bit quantization at RN.

The performance of the FP-MP decoder with $i_{\max} = 50$ using 3-bit (i.e. $N_{\mathsf{z}} = 8$) quantization at the RN and LLR calculation as given in (5) is shown in Fig. 4. The FP-MP decoder for $P_e = 0$ without quantization serves as benchmark. The FAVIB approach outperforms Uniform (UQ), Loyd-Max (LM) and IB based quantization by considering also the FL statistic in the quantizer design. The reason for that is an inherent error protection in the quantizer mapping, since some of the quantizer labels are unused. This gives an additional gain of approximately 1.2 dB compared to the IB approach. Furthermore, we investigate the LUT-MP decoder and FP-MP decoder in case of an error free and erroneous FL. The BER performance is shown in Fig. 5. For $P_e = 0$, the performance loss for a BER of 10^{-5} of the discrete implementation with 3-bit FAVIB quantization and 3-bit LUT-MP decoder processing is approximately 0.4 dB compared to the FP-MP decoder without quantization at the RN. It is also notable that for an error free and erroneous FL, the LUT-MP decoder using only 3-bits per message is about 0.3 dB for a BER of 10^{-5} to the performance of the FP-MP decoder with 64 bit double precision per message.

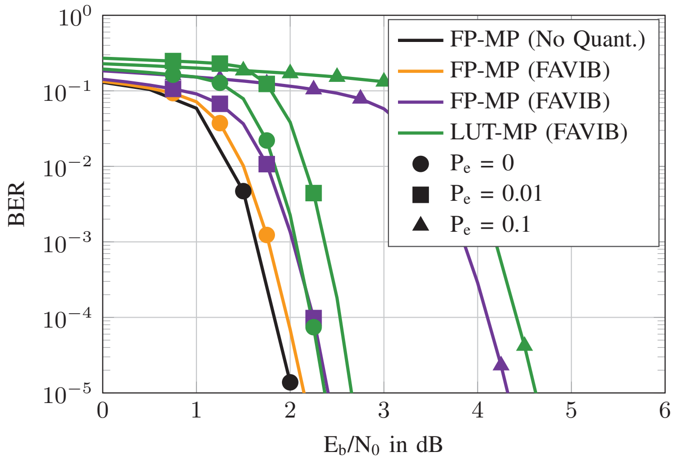

Fig. 5. BER performance of the 3-bit LUT- vs. FP-MP decoder.

V. Summary

We investigated the impact of quantization on the decoding performance in the uplink of a non-terrestrial network, where the relay node forwards digitized signals to the satellite. Our evaluation shows that Information Bottleneck based quantization is superior compared to common quantization approaches like uniform and Lloyd-Max. Furthermore, we also utilized the FAVIB algorithm to design quantizers which maximize the end-to-end performance between the user and the satellite. The optimized quantizer mapping contains an additional error protection leading to an improved decoding performance compared to Information Bottleneck based quantization. Finally, we investigated the performance of a lookup-table based message passing decoder which is optimized via a density evolution algorithm to process quantized samples. The lookup-table based decoder is designed for the threshold SNR of the code ensemble and used for all SNRs during the simulations. The BER performance of this decoder with 3-bit channel quantization and 3-bit message resolution is close to the floating-point implementation of the belief-propagation decoder without quantization at the relay node.

Acknowledgment

This work was partly funded by the European Regional Development Fund (ERDF) under grant LURAFO2012A.

References

[1] 3GPP TR 22.822, "Study on using Satellite Access in 5G, Stage 1 (Release 16)," Jun. 2018.

[2] 3GPP TR 38.811, "Study on New Radio (NR) to support Non-Terrestrial Networks; (Release 15)," Jun. 2018.

[3] A. Guidotti, A. Vanelli-Coralli, M. Conti, S. Andrenacci, S. Chatzinotas, N. Maturo, B. Evans, A. Awoseyila, A. Ugolini, T. Foggi, L. Gaudio, N. Alagha, and S. Cioni, "Architectures and Key Technical Challenges for 5G Systems Incorporating Satellites," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2624–2639, March 2019.

[4] D. Wübben, P. Rost, J. Bartelt, M. Lalam, V. Savin, M. Gorgoglione, A. Dekorsy, and G. Fettweis, "Benefits and Impact of Cloud Computing on 5G Signal Processing," *Special Issue The 5G Revolution of the IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 35–44, Nov 2014.

[5] 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.

[6] B. M. Kurkoski, K. Yamaguchi, and K. Kobayashi, "Noise Thresholds for Discrete LDPC Decoding Mappings," in *2008 IEEE Global Telecommunications Conference (GLOBECOM)*, Louisiana, USA, Dec. 2008.

[7] 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.

[8] J. Lewandowsky and G. Bauch, "Trellis based Node Operations for LDPC Decoders from the Information Bottleneck Method," in *2015 9th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Cairns, Australia, Dec. 2015.

[9] 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, Sept. 2016.

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

[11] R. Ghanaatian, A. Balatsoukas-Stimming, T. C. Mller, 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.

[12] 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, Sept. 1999, pp. 368–377.

[13] N. Slonim, "The Information Bottleneck: Theory and Applications," Ph.D. dissertation, Hebrew University of Jerusalem, Israel, 2002.

[14] 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.

[15] S. Hassanpour, T. Monsees, D. Wübben, and A. Dekorsy, "Forward-Aware Information Bottleneck-Based Vector Quantization for Noisy Channels," in *preparation*.

[16] B. M. Kurkoski and H. Yagi, "Quantization of Binary-Input Discrete Memoryless Channels," *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4544–4552, Aug. 2014.

[17] J. Bartelt, L. Landau, and G. Fettweis, "Improved Uplink I/Q-Signal Forwarding for Cloud-Based Radio Access Networks with Millimeter Wave Fronthaul," in *2015 International Symposium on Wireless Communication Systems (ISWCS)*, Brussels, Belgium, Aug. 2015.

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

[19] A. Winkelbauer, G. Matz, and A. Burg, "Channel-Optimized Vector Quantization with Mutual Information as Fidelity Criterion," in *2013 Asilomar Conference on Signals, Systems and Computers*, California, USA, Nov. 2013.

[20] R. G. Gallager, "Low-Density Parity-Check Codes," Ph.D. dissertation, 1963.

[21] D. J. C. MacKay, "Encyclopedia of Sparse Graph Codes," http://www.inference.org.uk/mackay/codes/data.html, 2018, [Online].