Probabilistically Shaped Multi-Level Coding with Polar Codes for Fading Channels

Onurcan İşcan, Ronald Böhncke, Wen Xu
Huawei Technologies Düsseldorf GmbH, German Research Center
Email: {Onurcan.Iscan, Ronald.Boehnke, Wen.Dr.Xu}@Huawei.com

Abstract—A probabilistic shaping method for multi-level coding (MLC) is presented, where the transmitted symbols are forced to have a shaped non-uniform distribution. It is shown that shaping only a single bit-level suffices to compensate for most of the shaping loss on the fading channels. A polar code based implementation of the proposed scheme is presented, where shaping is performed by using a precoder at the transmitter without increasing the decoding complexity. Simulation results show that performance improvements can be obtained compared to BICM- and MLC-based polar coding without shaping on Rayleigh fading channels.

I. INTRODUCTION

In order to approach the theoretical capacity limits in a communication system, the transmitted symbols should in general have a non-uniform distribution. However, many existing systems use uniformly distributed transmit symbols, which can lead to a shaping loss. Probabilistic shaping (PS) is a promising approach that can reduce this loss for higher order modulation. For additive white Gaussian noise (AWGN) channels, it is shown in [1] that bit-interleaved coded modulation (BICM) with independent demapping combined with PS can compensate for the demapping and shaping losses, and approach the capacity. A scheme was also given in [1], where an additional distribution matcher/dematcher is used as a shaping encoder/decoder together with systematic error correction codes. In [2] this idea is extended to polar codes, where a precoder is used for systematic encoding, allowing the utilization of distribution matchers similar to [1]. Moreover, a multi-level demapping and successive decoding approach is utilized to avoid demapping loss.

Recently, another PS approach with polar coding for higher order modulation is introduced in [3, 4] (by extending the ideas from [5]), where a polar decoder is used as a precoder for signal shaping. The idea is to generate some shaping bits using the precoder and append them to the information bits prior to the polar encoding, such that the bits in the generated codewords have a certain probability distribution. Codewords with non-uniform distribution of bits are then mapped to channel input symbols, such that they have a desired distribution resulting in a shaping gain. Compared to PS approaches that use non-binary distribution matchers (such as [1] and [2]), this approach cannot compensate for the full shaping loss, since only an approximation of the optimal symbol distribution can be obtained. However, it is shown in [6] that still a large shaping gain can be obtained. Moreover, this approach does not require a distribution dematcher at the receiver, as the shaping bits are treated as regular information bits at the receiver that can be discarded after decoding.

The existing literature on PS mainly considers AWGN channel model (see [1], [6] and references therein). Here, we aim to develop a solution for higher order modulation with non-uniform distribution for transmission over fading channels. We first discuss different transmission strategies in Sec. II and show that BICM based PS (with independent demapping) cannot fully remove the shaping loss on fading channels (unlike on AWGN channels). We then propose in Sec. III a shaped multi-level coding (s-MLC) approach with single bit-level shaping (SBS) that can remove most of the loss. Motivated by the asymptotic results, we build polar codes by extending ideas from [5] and [3] (in Sec. IV), where shaping gain is obtained by modifying mainly the transmitter. Our simulation results indicate that the proposed scheme can significantly outperform conventional BICM and MLC (with uniform distribution) based on the 5G New Radio (NR) polar codes.

In this work, we use lowercase bold letters (e.g. x) for vectors and uppercase letters (e.g. X) for the random variables representing the elements of the associated vectors. $P_X(x)$, $E(X)$ and $E(X)$ denote the probability distribution, entropy and expected value of $X$. $P(X; Y)$ is the mutual information between $X$ and $Y$, and $BSC(p)$ represents a binary symmetric channel with crossover probability $p$. Calligraphic letters (e.g. $\mathcal{F}$) denote sets.

II. TRANSMISSION OVER FADING CHANNELS WITH HIGHER ORDER MODULATION

Consider the real-valued channel model $y = h \odot x + w$, where $\odot$ denotes element-wise multiplication. $h$ and $w$ contain the fading coefficients and the Gaussian noise samples, respectively, $x$ contains the 2$^m$ amplitude shift keying (ASK) channel input symbol taken from the alphabet $X_m = \{ \pm 1, \pm 3, \ldots, \pm (2^m - 1) \}$. The signal to noise ratio (SNR) is $\gamma = E(X^2)/E(W^2)$. A symbol mapper generates $x$ from $e_i, i = 1, \ldots, m$, containing the bits in the $i$th ASK bit-level. The extension to a complex-valued channel model is straightforward. The maximum achievable rate with perfect channel state information (CSI) at the receiver is given by $R = I(X; Y|H)$. In the rest of the paper, we consider

1This work is accepted for publication at Globecom 2018 Workshops. Copyright IEEE 2018.
Rayleigh distributed fading coefficients that are known to the receiver, i.e., each ASK symbol is subject to a different independent fading coefficient. Note that $R$ is maximized with a Gaussian distributed $X$ for continuous input alphabets, and with Maxwell-Boltzmann (MB) distribution with $P_X(x) \sim e^{-x^2}$ for non-continuous input alphabets. Below we consider different transmission strategies.

### A. Bit-Interleaved Coded Modulation

$k$ information bits are encoded by a channel code to obtain a binary codeword $c$ of length $n = mn_c$, which is interleaved and partitioned to $c_i$, $i = 1, \cdots, m$, i.e., each ASK bit-level contains bits from the same codeword. The transmission rate is $R = k/n_c$, and with independent demapping, non-negative rates up to

$$R_{BICM} = \mathbb{H}(X) - \sum_{i=1}^{m} \mathbb{H}(C_i|Y, H)$$

(1)

can be achieved [8]. Fig. 1 shows $R_{BICM}$ with uniform and MB symbol distribution for $m = 4$ with Gray labeling, which is typically used with BICM [9]. We optimized parameters of the MB distribution numerically to maximize the achievable rate at each SNR. We observe that BICM with MB distributed channel input symbols outperforms uniform BICM, but there is still a gap to the capacity. Recall that BICM with probabilistic shaping can approach the capacity of AWGN channels, if the symbol distribution is optimized for the target SNR. However, it suffers from the independent demapping loss due to the variation of the instantaneous SNR in fading channels, i.e., signal shaping does not completely compensate for the loss, because the dependence between bit-levels are not taken into account during demapping.

### B. Multi-Level Coding

$k$ information bits are first divided into $m$ parts with lengths $k_i$, $i = 1, \cdots, m$, which are encoded to codewords $c_i$ of length $n_c$ and transmitted over the $i$th ASK bit-level. The transmission rate is $R = \sum_{i=1}^{m} k_i/n_c$. The receiver performs multi-level demapping and successive decoding, i.e., after decoding each bit level, this information is used for demapping the next bit-levels. The achievable rate is [6]

$$R_{MLC} = \sum_{i=1}^{m} \mathbb{H}(C_i|Y, C_1, \cdots, C_{i-1}, H).$$

(2)

Observe that $R_{MLC}$ is the sum of bit-level capacities $I_c$, and the choice of the binary labeling does not affect the sum-rate, but only influences the rate of each individual bit-level. Therefore, asymptotically any binary labeling is optimal. Fig. 1 shows $R_{MLC}$ with uniform and MB channel input symbol distribution for $m = 4$, where the parameters of the MB distribution is optimized numerically for each SNR. We observe that MLC with MB distributed channel input symbols performs best, which motivates us to use signal shaping with MLC for fading channels.

Although the choice of the bit labeling does not influence the achievable rates asymptotically, it can have an effect on the performance in finite lengths. In [6] and [10] it is shown that a large variance of the bit-level capacities is beneficial in finite lengths, i.e., capacity of each bit-level should be preferably different. Therefore, we use set-partitioning labeling, which corresponds to natural binary code (NBC) for ASK, guaranteeing non-decreasing minimum Euclidean distance (and thus a non-decreasing capacity) for each bit-level if successive decoding is performed starting from the least significant bit.

### III. Signal Shaping for Multi-Level Coding

In the presented schemes, the channel input symbols are obtained from binary codewords. To obtain symbols with an MB distribution, the binary codewords transmitted on some bit-levels need to have non-uniform distribution of bits, i.e., $P_{C_i}(1) \neq 0.5$. Moreover, the bit-level distributions may need to be conditioned on each other, which implies joint encoding of bit-levels. By relaxing the condition on joint encoding, one can approximate the MB distribution by a product distribution (as in bit-level probabilistically shaped coded modulation, PSCM [11]), where each bit-level is encoded independently with $P_{C_i}(1) = p_i$. In the following, we show that having only a single bit-level with non-uniform distribution suffices to generate an approximation of the MB distribution that can compensate for most of the shaping loss in fading channels.

#### A. Single Bit-Level Shaping (SBS)

We consider the following piecewise constant approximation of the MB distribution for $2^m$-ASK symbols

$$P_X(x) = \begin{cases} 
  p/2^{m-1} & \text{if } |x| < 2^{m-1} \\
  (1 - p)/2^{m-1} & \text{otherwise}.
\end{cases}$$

(3)

This distribution requires only a single bit-level $t$ with $P_{C_t}(1) = p \geq 0.5$, where this bit-level determines whether $|x|$ is smaller than $2^{m-1}$ or not, i.e., $c_t = 1$ for $|x| < 2^{m-1}$. Accordingly, symbols with high energy (i.e., with $|x| > 2^{m-1}$) are transmitted less likely than the symbols with low energy. A labeling with such a property can be obtained from NBC by a circular shift of $2^{m-2}$, as shown in Fig. 2 for $m = 4$. 

![Fig. 1. Achievable rates on Rayleigh fading channels for $m = 4$ using MLC and BICM with uniform, MB and approximate MB (with single bit-level shaping) distributed ASK symbols. The parameters of the non-uniform distributions are numerically optimized at each SNR to maximize the achievable rate with the respective transmission scheme. Dotted line shows the channel capacity.](image-url)
and \( t = 4 \). We call this labeling \textit{shifted NBC}. The mapping from \( x \in X_m \) can be described by the binary representation of \( [(2^m - 1 - x)/2 + 2^m - 2 \mod 2^m] \) with most significant bit corresponding to the \( m \)th bit-level, and \( t = m \). Note that the circular shift does not change the Euclidean distances of the bit-levels. There exist also other equivalent labelings with the same distance properties. We leave further optimizations of the bit-labeling for non-uniform distribution and finite lengths as future work.

Observe that the \( m \)th bit-level determines the shape of the distribution and \( P_{C_{\text{rel}}} (1) \neq 0.5 \), i.e., the \( m \)th bit level requires special channel codes that we discuss in Sec. IV. Note that channel encoders usually generate codewords with \( P_{C} (1) = 0.5 \), thus there are no restrictions for the first \( m - 1 \) bit-levels.

\textbf{Achievable Rates for SBS:} For each operating SNR \( \gamma \), we find by numerical search a \( p^* \) that maximizes (2), shown in Fig. 3 for \( m = 4 \). Fig. 3 also depicts the resulting bit-level capacities with optimized \( p^* \) and with \( p = 0.5 \). Note that the optimal \( p^* \) leads to a decreased capacity of the last bit-level, but it also leads to increased capacities for other bit-levels, resulting in a larger sum-rate. Fig. 1 shows the achievable rates of SBS-MLC, which are close to optimal. This implies that \( R \) is not very sensitive with respect to the input distribution, and an approximation of the optimal distribution is nearly sufficient.

SBS can also be combined with Gray labeled BICM by using the method in [4] and shaping only a single bit-level. By maximizing (1) over \( p \) we obtain the achievable rates for SBS-BICM, which are also given in Fig. 1 as reference.

\textbf{IV. POLAR CODES WITH SINGLE BIT-LEVEL SHAPING}

Polar codes are known to work well with MLC [10], and they also allow generating codewords with \( P_C (1) \neq 0.5 \). In this work we use them in combination with s-MLC.

\textbf{A. Polar Codes}

Channel polarization [12] is an operation, where a physical channel is converted into virtual channels having either very high or very low reliabilities asymptotically. A polar encoder assigns information bits to reliable virtual channels, and (known) frozen bits to the unreliable ones. A polar decoder (e.g. successive cancellation list (SCL) decoder [13]) processes a noisy observation of a codeword together with the frozen bits to estimate the information bits. A polar codeword \( c \) of length \( n \) is obtained from a binary sequence \( u \) by

\[ c = uG, \] (4)

where \( G \) is the \( n \times n \) polar transform matrix. \( u \) contains information bits at the indices described by the set \( \mathcal{I} \), representing the reliable virtual channels, and frozen bits at the indices described by \( \mathcal{F} \). We will call \((n, \mathcal{I}, \mathcal{F})\) a polar code of length \( n \). The reliabilities of the virtual channels can be determined using density evolution (DE) [14].

\textbf{B. Polar Codewords with Non-Uniform Distribution of Bits}

Assuming independent and uniform choices of information and frozen bits, the numbers of ones and zeros in a codeword are roughly equal, i.e., \( P_{C} (1) = 0.5 \). In [5] and [15] a method was presented to obtain polar codewords with a target probability \( P_{C} (1) \neq 0.5 \) for transmission over asymmetric channels. The idea is to use some of the virtual channels (denoted by the set \( \mathcal{S} \)) to transmit \textit{shaping bits}, which are generated from the information and frozen bits. The shaping bits do not convey any new information, but they force the codeword to have a target probability \( P_{C} (1) \).
over a BSC(p), where e is the noise vector and the channel output is the all-zero vector. Accordingly, if we utilize a polar decoder by treating the all-zero vector as the noisy observation and the bits at $I \cup F$ as frozen bits, we can obtain the shaping bits at the output of the decoder if $(n, S, I \cup F)$ forms a polar code for a BSC(p), as described in [3].

**Example 1:** Consider transmitting $k$ information bits over a target channel with $n \to \infty$ and aim to have $P_C(1) = p$. Let the elements of $\epsilon$ represent the reliabilities of the virtual channels if a BSC(p) is polarized (obtained e.g. using DE). Similarly, let the elements of $\epsilon$ denote the reliabilities of the target channel after polarization. Note that asymptotically $S$ contains the most reliable $s = \lfloor n(1-h_2(p)) \rfloor$ virtual channels according to $\epsilon$ [3], where $h_2(p)$ is the binary entropy function. $I$ is constructed from the $k$ most reliable virtual channels of $\epsilon$ that are in the set $N \setminus S$. The remaining virtual channels form $F$. For simplicity, one can assume $\epsilon = e$, as for many binary symmetric channels the order of the reliabilities of the virtual channels is similar. As a result, if an ordered set of virtual channels is available (such as defined in [7]), one can use the most reliable $s$ channels for shaping bits, the least reliable channels for frozen bits and the rest for information bits.

Let $d$ and $f$ represent the vectors containing the information and frozen bits, respectively. We use a polar decoder as a precoder to obtain the shaping bits $s$, where $d$ and $f$ are treated as frozen bits, and $I_s = \log((1-p)/p)$ is used as the $L$-values representing the all-zero sequence. The decoder output would contain the shaping bits $s$. Note that $d$, $f$ and $s$ form $u$, which can be fed to the polar transform as in [4] to generate the codeword $e$ with $P_C(1) = p$. At the receiver, $s$ is treated as information bits (like $d$) that are discarded after decoding.

**Example 2:** For finite lengths, the assumption about the number of the shaping bits may be inaccurate. Therefore, we evaluate $P_C(1)$ numerically. For $n = 256$, we choose a target $p$, select an initial value for $s$ and randomly generate $n-s$ bits that are used as information and frozen bits. Then we use the most reliable $s$ virtual channels from the ordered set in [7] as $S$, and use an SCL decoder with list size 8 as a precoder to obtain $s$, where we treat all-zero vector as the output of a BSC(p). We use a polar transform to generate the codeword from the information, frozen and shaping bits, and evaluate $P_C(1)$ by Monte-Carlo simulations. We seek for $s$ that minimizes $|P_C(1) - p|$, which is shown in Fig. [4]. We observe that with increasing number of shaping bits, the distribution of ones and zeros within the codeword becomes more non-uniform.

### C. Shaped Multi-Level Coding with Polar Codes

We propose using s-MLC with polar codes, where we employ conventional polar codes for the first $m - 1$ bit levels and a polar code with $P_C(1) = p$ for the $m$th bit-level, all having codeword length $n_c$. To improve the finite length performance, we use $z_i$ cyclic redundancy check (CRC) bits at each bit-level and use CRC-aided decoding [13]. The shaping bits at the $m$th bit-level are generated after CRC bits are appended.

The design of polar codes requires obtaining the reliabilities of the polarized virtual channels, which depends on the operating SNR. Similarly, the set of virtual channels used for the transmission of the shaping bits depends on the target distribution. To simplify the design process, we use the ordered set of virtual channels from [7], which is specified for 5G NR control channels. Accordingly, for the bit-levels $i = 1, \ldots, m - 1$, the most reliable $k_i + z_i$ virtual channels are used for transmission of information and CRC bits and the rest is used for transmission of the frozen bits. For the $m$th bit-level, the most reliable $s$ virtual channels are used to transmit the shaping bits. From the remaining virtual channels, the most reliable ones are used for transmitting $k_m + z_m$ information and CRC bits, leaving the rest for frozen bits.

Note that s-MLC with polar codes requires only small modifications compared to a conventional MLC. The receiver takes $P_X(x)$ into account during demapping, and the decoded shaping bits are discarded at the receiver. Thus, the complexity increase at the receiver is negligible. The transmitter, however, needs to run an additional polar decoder to generate the shaping bits.

Below we discuss the choice of the parameters such that the number of transmitted information bits is maximized at a given operating SNR $\gamma$, a target block error rate (BLER) $P_e$ and $n_c$ channel uses.

**Choice of $p$:** The optimal $p^*$ can be obtained numerically by maximizing $R_{MLC}$ at $\gamma$, as given in Fig. [3] for $m = 4$.

**Choice of $s$:** After $p^*$ is determined, we use the approach in Example 2 to obtain the required number of shaping bits $s$.

**Choice of $k_i$:** One can approximate $k_i$ as $[I_n s_c]$ for very large $n_c$. However, for finite lengths and with a target $P_e$, the approximation will be inaccurate. Let $P_i^*$ denote the probability of incorrect decoding of the $i$th bit-level, which can be obtained by Monte-Carlo simulations. The BLER can be written as $P_e = 1 - \prod_{i=1}^m (1 - P_i)$. We seek for $k_i$ that produces a smaller error probability than $P_i$, assuming each bit-level has the same error probability $P_i = 1 - \sqrt{1 - P_e}$.

Let the elements of $e_i$ denote the reliability of the virtual channels after polarization at the $i$th ASK bit-level, assuming the previous bit-levels are error-free decoded and demapped at $\hat{\gamma}$. For given decoding parameters (e.g. list size) we find the maximum value of $k_i$ such that the block error probability (obtained by Monte-Carlo simulations) is less than $P_i$; if the most reliable $k_i + z_i$ virtual channels (excluding the ones used to transmit the shaping bits in the $m$th level) are used to transmit information and CRC bits.

### V. Numerical Evaluations and Discussions

We designed an s-MLC scheme uMLC for 2$^4$-ASK and shifted NBC labeling, as described in the previous section. We chose the operating SNR $\gamma$ such that a target error probability $P_e \leq 10^{-3}$ is obtained for $k = 512$ information bits with $n_c = 256$ channel uses. An SCL decoder with list size 8 was employed as the precoder and decoder, and pseudo-random sequences as frozen bits. We used $z_i$ CRC bits for the $i$th bit-level.

To compare the performance of the proposed scheme, we use three reference transmission schemes with similar distribution of $X$. As a first reference, an MLC scheme uMLC
with NBC labeling and uniform $P_X(x)$ was designed, where we used the same procedure as before except for the choice of $p^* = 0.5$ and $s = 0$. A second reference is the BICM scheme $uBICM$ with Gray labeling and uniform $P_X(x)$, where we used codewords of length $mn_c$ and a different pseudo-random interleave between encoder and symbol mapper for each codeword. As a further reference, we designed a BICM scheme $sBICM$ with Gray labeling and non-uniform $P_X(x)$ according to (3), where $s$ bits were used for shaping a single bit-level, resulting in a distribution given in (3). Note that all four schemes have $R = 2$ bits/use, and $z = 16$ additional CRC bits in total are employed for CRC-aided decoding. The parameters are given in Table I.

Fig. 5 shows the BLER performances. We observe that by using SBS instead of uniform signaling, the performance can be improved both for BICM and MLC. $sMLC$ performs 0.7dB better than $uMLC$ and 1dB better than $uBICM$, while the asymptotic gains (according to Fig 1) are 0.6dB and 1.18dB, respectively. Also note that MLC outperforms BICM, although shorter codewords (that cause a larger finite length loss) are employed. The proposed $sMLC$ with polar codes significantly outperforms the reference schemes.

Recall that for all designed codes we used the ordered set of virtual channels from [7] which is not necessarily optimal for any of the presented codes at the operating SNRs, and by evaluating the reliabilities of the virtual channels by using DE, one can improve the performance of each code. On the other hand, the performance curves with this suboptimal choice of virtual channels and the asymptotic results show similar trends.

In order to make a fair comparison, we only considered transmission strategies, where $X$ is distributed uniformly or according to (4). To obtain the full shaping gain on fading channels, one can use other transmission schemes resulting in an exact MB distribution with optimal parameters. For example, (16) proposes using non-binary channel codes together with non-binary distribution matchers and dematchers. This scheme would not suffer from a demapping loss on fading channels, but requires non-binary channel decoders, which are usually more complex than binary decoders. Another alternative is using the scheme presented in [2], where an additional precoder (for allowing systematic encoding of polar codes by dynamic freezing) and non-binary distribution matchers and dematchers are required. Compared to [2] and [16], the proposed scheme does not generate symbols with capacity achieving distribution, but as given in Fig. 1 the expected additional gain by using an exact MB distribution is small. Moreover, the proposed scheme has a simple receiver structure, as it only requires small modifications to a conventional polar MLC receiver, and does not require a distribution dematcher. Moreover, the precoder at the transmitter can be realized by a polar decoder, which is already included in the transmission chain of bidirectional communication systems.

**ACKNOWLEDGMENT**

This work has been partly performed in the framework of the H2020 project ONE5G (ICT-760809) receiving funds from the EU. The views expressed in this work are those of the authors and do not necessarily represent the project view.

**REFERENCES**

[1] G. Böcherer, F. Steinr, and P. Schulte, “Bandwidth efficient and rate-matched low-density parity-check coded modulation,” IEEE Trans. Commun., vol. 63, no. 12, pp. 4651–4665, 2015.

[2] T. Prinz, P. Yuan, G. Böcherer, F. Steinr, O. Ișcan, R. Böhnke, and W. Xu, “Polar coded probabilistic amplitude shaping for short packets,” in IEEE Int. Workshop on Sig. Proc. Adv. in Wirl. Comm., Jul. 2017, pp. 83–87.

[3] O. Ișcan, R. Böhnke, and W. Xu, “Shaped polar codes for higher order modulation,” IEEE Commun. Lett., vol. 22, no. 2, pp. 252–255, 2018.

[4] O. Ișcan and W. Xu, “Polar codes with integrated probabilistic shaping for 5G new radio,” preprint arXiv:1808.09360, 2018.

[5] J. Honda and H. Yamamoto, “Polar coding without alphabet extension for asymmetric models,” IEEE Trans. Inf. Theory, vol. 59, no. 12, pp. 7829–7838, 2013.

[6] U. Wachsmann, R. F. Fischer, and J. B. Huber, “Multilevel codes: Theoretical concepts and practical design rules,” IEEE Trans. Inf. Theory, vol. 45, no. 5, pp. 1361–1391, 1999.

[7] 3GPP TS 38.212, Technical Specification Group RAN, Radio Access Network, NR, Multiplexing and Channel Coding, 2017.

[8] G. Böcherer, “Achievable rates for probabilistic shaping,” preprint arXiv:1707.01134, 2017.

[9] G. Carre, G. Tanico, and E. Biglieri, “Bit-interleaved coded modulation,” IEEE Trans. Inf. Theory, vol. 44, no. 3, pp. 927–946, 1998.

[10] M. Stidd, A. Schenk, C. Stierstorfer, and J. B. Huber, “Polar-coded MC-MIMO,” IEEE Trans. Commun., vol. 61, no. 10, pp. 4108–4119, Oct. 2013.

[11] M. Pikus and W. Xu, “Bit-level probabilistically shaped coded modulation,” IEEE Commun. Lett., vol. 21, no. 9, pp. 1929–1932, 2017.

[12] E. Arikan, “Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels,” IEEE Trans. Inf. Theory, vol. 55, no. 7, pp. 3051–3073, Jul. 2009.

[13] L. Tal and A. Vardy, “List decoding of polar codes,” IEEE Trans. Inf. Theory, vol. 61, no. 5, pp. 2213–2226, May 2015.

[14] R. Mori and T. Tanaka, “Performance of polar codes with the construction using density evolution,” IEEE Commun. Lett., vol. 13, no. 7, 2009.

[15] M. Mondelli, R. Urbanke, and S. H. Hassani, “How to achieve the capacity of asymmetric channels,” in Proc. Allerton Conf. Commun., Contr. Comput., Sep. 2014, pp. 789–796.

[16] F. Steiner, G. Liva, and G. Boeckerh, “Ultra-sparse non-binary LDPC codes for probabilistic amplitude shaping,” in Proc. IEEE Global Commun. Conf. (GLOBECOM), Dec 2017, pp. 1–5.