An Error Exponent for the AWGN Channel with Decision Feedback and Lattice Coding

Kenneth Palacio-Baus*†, Natasha Devroye*
*University of Illinois at Chicago, {kpalac2, devroye}@uic.edu
†University of Cuenca, Ecuador

***Abstract*—We present achievable error exponents for the AWGN channel with one bit noiseless feedback and an almost-sure power constraint. As in Forney's decision feedback approach for discrete memoryless channels, the backward channel is used to request retransmissions from the transmitter whenever an erasure is declared at the decoder. Our erasure/re-transmission-request scheme introduces a novel erasure decoding rule built on top of a lattice-based code inspired by De Buda. Numerically, our scheme is seen to outperform the sphere packing bound (valid for block codes) at lower rates and higher SNRs.**

***Index Terms*—Error exponents, AWGN, lattice, Forney, variable length coding**

I. Introduction

Forney [1] introduced *decision feedback* coding, where the sequences sent over the channel can be either decoded (correctly or incorrectly), or declared an erasure. The latter signals that the decoder is not confident enough to make a hard decision on the transmitted message. An erasure triggers a retransmission request, which is carried from the receiver to the transmitter over a noiseless feedback link. This approach nonetheless allows for considerable improvement of the error probability with respect to simple point-to-point feedback-free transmissions. Moreover, the use of feedback is very limited and the effective transmission rate is not affected. Forney's work focused on discrete memoryless channels, but his ideas can be extended to channels with continuous alphabets and different erasure criteria.

Wyner [2] applied a Forney-like scheme to the continuous time AWGN channel with average power P and noise variance σ^2. The capacity of this channel is well known to be unaffected by feedback and is given by $C = \frac{1}{2}\log_2\left(1+\frac{P}{\sigma^2}\right)$. Wyner's scheme is built on an orthogonal code, whose reliability function in the absence of feedback, denoted as $E_o(R)$, is known exactly and given by:

$$E_o(R) = \begin{cases} C/2 - R, & 0 \le R \le C/4 \\ \left(\sqrt{C}-\sqrt{R}\right)^2, & C/4 \le R \le C \end{cases}. \quad (1)$$

Wyner's approach is presented as a repeat-request coding strategy which includes the definition of an erasure decoding rule for an orthogonal code over the AWGN channel. Hence, Forney's principles can be directly applied and a higher reliability attained, noting that (1) is boosted to (2):

$$E_{\text{Wyn}}(R) = \left(\sqrt{C}-\sqrt{R}\right)^2 + C - R = 2\sqrt{C}\left(\sqrt{C}-\sqrt{R}\right). \quad (2)$$

Observe that (2) leads to larger gains for rates near C. Moreover, the transmission rate is not affected as the retransmission probability tends to zero as the block length tends to ∞.

If the feedback channel is able to convey more information without error, i.e. complete output feedback is available at the encoder, then much higher error exponents can be achieved. For example, Schalkwijk-Kailath (SK) [3] proposed a linear block coding scheme whose error probability has a double exponential decay as a function of the block length N. Later, Gallager and Nakiboğlu [4] demonstrated that a nested exponential decay can be achieved and that the number of nested terms grows linearly with the block length. A more restricted setting is addressed in [5], where achievable error exponents for the AWGN are studied under block coding and the presence of a noiseless but rate-limited feedback link. Forward codewords are subject to a different –more flexible– expected block power constraint, which allows for high amplitude transmissions. It was demonstrated that when the feedback rate is smaller than that of the forward channel, only an exponential decay in the error probability can be attained, in contrast to larger feedback rates where higher order exponential decays can be achieved. The benefits of noiseless feedback under block coding for AWGN channels at zero-rate are part of the line of work of Burnashev and Yamamoto, [6].

For variable length coding (VLC) settings, the availability of complete noiseless feedback was studied under a peak power constraint by Schalkwijk-Barron in [7] based on Viterbi's sequential decision feedback. The same reliability was later shown to be asymptotically attained using a simpler block-wise variable length coding by Yamamoto-Itoh [8]. Noisy feedback under VLC has been studied in [9], [10].

Error exponents of AWGN channels based on lattice codes have been previously analyzed under block coding in the absence of feedback in [11]. It was shown that Erez-Zamir's lattice encoding scheme is not only able to achieve capacity [12] but also, for rates sufficiently close to capacity, to attain error exponents similar to those under randomly generated codes (in contrast to the initial conjecture that only Poltyrev's error exponent could be achieved, see [13]). Their approach is based on transforming the AWGN channel model into a

This work was partially supported by NSF under award 1815428. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF.

Modulo-Lattice Additive Noise (MLAN) channel.

This work considers the case of a single bit of noiseless feedback, similar to that introduced by Forney and used by Wyner in [2] for AWGN channels. We focus on discrete time AWGN channels and introduce a novel erasure decoding technique that exploits the structure of a lattice-based code. We show that compared to block coding, it is possible to attain higher error exponents in the low-rate regime if this rule is used, especially at higher SNR.

In the following, we utilize capital letters to denote random variables and small letters for their realizations. Bold letters indicate N-dimensional vectors.

II. Problem statement

We consider an AWGN channel with a single bit of noiseless feedback. As in Forney's decision feedback scheme (or equivalently in Wyner's repeat-request coding strategy) the feedback channel is only used to tell the encoder that a retransmission of the last transmitted message is necessary. Such retransmission requests arise when the received sequence cannot be reliably decoded into a unique message but declared as an erasure (e.g. when a received sequence lies close to a decision boundary or is equidistant between different codewords). Hence, a message is transmitted repeatedly until it is received unerased before the transmitter moves to a new message. Next, we present a formal statement of the problem.

Consider a communication system in which a terminal (the transmitter) sends messages selected uniformly from the set $\mathcal{M} = \{1, 2, ..., M\}$ to another terminal (the receiver) over the forward direction. Let $\mathcal{X}, \mathcal{Y}$ be the set of all reals. Let $(\mathcal{X}, Q(y|x), \mathcal{Y})$ denote a one-way AWGN channel characterized by law $Q(y|x)$ in the forward direction with corresponding input and output alphabets $\mathcal{X}$ and $\mathcal{Y}$. Let Q^N denote N uses of channel with law $Q(y|x)$. A $(\mathcal{X}, Q(y|x), \mathcal{Y})$ channel is said to be memoryless if $Q^N\left(y^N|x^N\right) = \prod_{k=1}^{N} Q(y_k|x_k)$. At the k-th channel use, the forward direction is characterized by independent and identically distributed (iid) AWGN with zero mean and variance σ^2; which is described by the model:

$$Y_k = X_k + Z_k, \qquad Z_k \text{ iid } \sim \mathcal{N}(0, \sigma^2), \quad k = 1, 2, ...$$

Channel inputs are subject to an almost sure power constraint for a block of length N:

$$\Pr\left(\sum_{k=1}^{N} X_{ik}^2 \leq NP\right) = 1$$

where P corresponds to the average power of the transmitter. Moreover, we let the number of transmissions required to decode each message vary, so define a code for this channel in the variable length coding (VLC) setting as:

Definition 1. *A variable length code with a single bit of noiseless feedback, denoted by $\mathcal{C}_{vl}^{FB,AS}\left(M, P, \sigma^2\right)$, and used for the transmission of messages uniformly selected from the set $\mathcal{M} = \{1, 2, ..., M\}$ over a one-way AWGN channel $(\mathcal{X}, Q(y|x), \mathcal{Y})$, subject to an almost sure power constraint, comprises:*

1) *A set of forward encoding functions: defined for $k = 1, 2, ...$ as: $f_k : \mathcal{M} \to \mathcal{X}$, leading to channel inputs $X_k = f_k(W)$ and subject to an almost sure power constraint, where $W \in \mathcal{M}$ corresponds to the equiprobable message. We use the notation $\mathbf{x}(w)$ to denote the length N codeword linked to message $W = w$.*
2) *A set of decoding functions $\phi_k : \mathcal{Y}^k \to \mathcal{M}$ which yield the best estimate of the sent message W, denoted as $\widehat{W}$.*
3) *A non-negative random variable Δ, a stopping time of the filtration $\mathcal{G}_k = \sigma\{Y_1, ..., Y_k\}$, that indicates the time slot in which the receiver declares the estimate $\widehat{W}$. Here $\sigma\{\cdot\}$ denotes a σ-field.*

Moreover, we define the average communication rate as $\bar{R} = \frac{\log M}{\mathsf{E}[\Delta]}$, and denote the maximum probability of error attained by a $\mathcal{C}_{vl}^{\text{FB,AS}}\left(M, P, \sigma^2\right)$ code at rate $\bar{R}$ as:

$$\mathsf{P}_{\text{error}}^{\text{FB,AS}}\left(\bar{R}\right) := \max_{w \in \mathcal{M}} \Pr\left(\phi(y^{\Delta}) \neq w \mid W = w \text{ sent}\right)$$

Definition 2. *An achievable error exponent on the probability of error of a one-way AWGN channel with a single bit of noiseless feedback, and operating at rate $\bar{R}$ is defined as:*

$$E_{FB}(\bar{R}) := \liminf_{\mathsf{E}[\Delta] \to \infty} -\frac{1}{\mathsf{E}[\Delta]} \log \mathsf{P}_{error}^{FB,AS}\left(\bar{R}\right),$$

where the subscript FB indicates the availability of feedback.

Next, we present our main result statement followed by its corresponding proof.

III. Main result

We consider a repeat-request strategy like that proposed by Wyner [2] for the AWGN channel, using a single bit of noiseless feedback for retransmission request signaling only. Our scheme differs from Wyner's in that we use a lattice code instead of an orthogonal code. One contribution is thus to propose a novel erasure decoding rule tailored to the structure of lattices. Without loss of generality, assuming that $\sigma^2 = 1$, our main result is reflected in the following:

Theorem 1. *An achievable error exponent for the AWGN one-way channel, subject to an almost sure power constraint and using variable length coding is:*

$$E_{FB,Lat}(\bar{R}) \geq 4^{(C_\Lambda - R)} + 1 - 2 \cdot 2^{(C_\Lambda - R)}, \tag{3}$$

where $C_\Lambda = \log_2 \sqrt{P}$.

The term $C_\Lambda = \log_2 \sqrt{P}$, represents a rate slightly smaller than the channel capacity $C = \log_2 \sqrt{1 + P}$, and follows the notation introduced by De Buda's initial work [14] on the use of lattices for the AWGN channel. This scheme is thus valid only for a subset of the rates below capacity. Note that if $P \gg 1$, then $C_\Lambda \to C$.

A. Proof of Theorem 1

The result shown in (3) can be achieved by applying a repeat-request strategy over a lattice based code for the AWGN channel and the erasure decoding rule we describe in this

section. Let us first recall that the volume of a hypersphere of radius ρ is given by:

$$J_N(\rho) = \frac{\left(\pi\rho^2\right)^{N/2}}{\Gamma\left(\frac{N}{2}+1\right)} \tag{4}$$

where $\Gamma(n) = (n-1)!$ is the Gamma function. De Buda showed in his seminal work on lattices [14] that for any rate $R < \frac{1}{2}\log_2\left(\frac{P}{\sigma^2}\right)$, there exists a lattice code with arbitrarily small maximum error probability.

Following De Buda's notation in [14], let $\mathbf{r}$ be an N-dimensional radius vector (each of its components are denoted by x_i), and r indicate its magnitude, then $||\mathbf{r}^2|| = \sum_{i=1}^{N} x_i^2 = r^2$. We can express the noise distribution Z_N as:

$$Z_N(\mathbf{r}) = \frac{1}{(2\pi)^{N/2}} \exp\left(-\frac{r^2}{2}\right)$$

Note that the distribution depends on a scalar value r. Since for a normal distribution we have $Q(x) = \int_x^\infty Z_1(r)dr$, note that the probability that the sum of N squared noise samples exceeds χ^2 is given by the N-variate chi-square distribution $Q(\chi^2|N)$. Hence we have that:

$$Q(\chi^2|N) = \int_\chi^\infty Z_N(r)dJ_N(r)$$

Let Λ denote a lattice defined in $\mathbb{R}^N$ as the set $\Lambda = \{\lambda = G\mathbf{x} : \mathbf{x} \in \mathbb{Z}^N\}$, where G is an $N \times N$ real-valued generating matrix, so the lattice is generated by all integer linear combinations of the basis vectors. Then, consider all points $\lambda \in \Lambda$ contained within bounding region B_N, defined as the N-dimensional ball of radius $\sqrt{NP}$ corresponding to the volume that contains the codewords. Moreover, Λ has a fundamental Voronoi region of volume $\det\Lambda = \det\{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_N\} \neq 0$[1]. As in [14], the number of messages to be transmitted at a rate R (given by $M = 2^{NR}$) is determined by the ratio of the volumes of the hypersphere that defines the codebook with a radius of $\sqrt{NP}$ and a hypersphere of radius a. The latter has the same volume as that of the fundamental Voronoi region, hence $J_N(a) = \det\Lambda$, and thus a receives the name of *effective radius*, $r_{\text{eff}} = a$, see Figure 1. Then, from (4) we have that $2^{NR} = J_N(\sqrt{NP})/J_N(a)$. Hence, we can find the radius a that allows a communication rate R as:

$$a = \sqrt{NP}/2^R \tag{5}$$

Formally, let $\mathcal{C}$ be the codebook formed by the points of lattice Λ inside a hypersphere of radius $\sqrt{NP}$. Then, each $\mathbf{x} \in \mathcal{C}$ is associated with a decoding region for message $w \in \mathcal{M}$, defined as:

$$A_r(\mathbf{x}(w)) := \{\mathbf{y} : ||\mathbf{y} - \mathbf{x}|| < r\}, \tag{6}$$

where r is the radius of a hypersphere and chosen as $r = \alpha a$. Here, $\alpha \in (0,1)$ is a design parameter which will be selected later and relates to erasures, and a determined by the communication rate as in (5). The scheme operates as follows:

[1] $\{\mathbf{s}_1, \mathbf{s}_2, ..., \mathbf{s}_N\}$ are N linearly independent basis vectors that generate lattice Λ.

A message $W = w$ selected uniformly from $|\mathcal{M}|$ is encoded into codeword $\mathbf{x}(w)$ and sent over the forward channel. Upon receiving this sequence, the decoder utilizes the following:

Decoding rule: The decoder declares that the message $\widehat{W} = w_x$ associated to lattice point $\mathbf{x}$ was sent, if the length N received sequence $\mathbf{y}$ satisfies $\mathbf{y} \in A_r(\mathbf{x})$, otherwise, an erasure is declared.

Assume that message $W = w$ is sent, which leads to codeword $\mathbf{x}(w)$. Define the following two events as in [1]:

1) Event E_1: Corresponds to the case where the received signal falls outside the decoding region corresponding to the codeword sent (the circle of radius r denoted by $A_r(\mathbf{x}(w))$ in Figure 1).
2) Event $\mathcal{E}$: Corresponds to the case in which the received sequence lies within a decoding region associated to a message different to the one sent, i.e. lies in $A_r(\mathbf{x}(w'))$ for $w' \neq w$. This event is usually called the undetected error event.

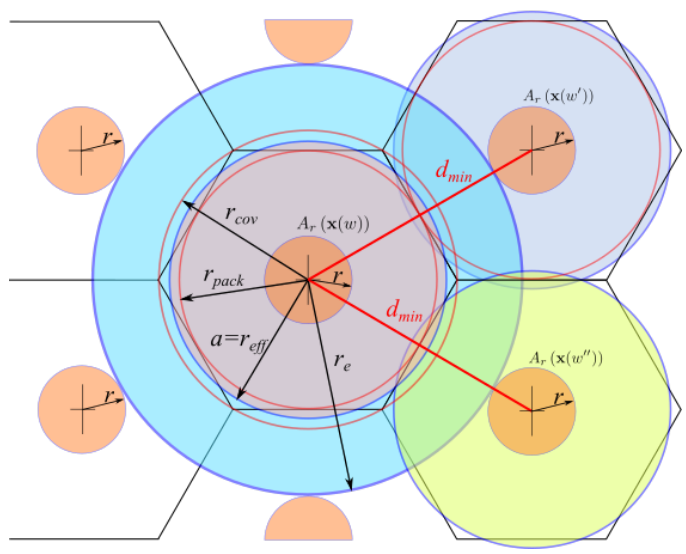

Fig. 1. Portion of a lattice based code $\mathcal{C}$ for $N = 2$. Decoding regions for messages w, w' and w'' are shown as the radius-r circles, whereas hexagons depict Voronoi regions. The effective radius r_{eff} is the radius of the hypersphere of the same volume as the lattice's fundamental Voronoi region.

Error probability: Next, assume that message $W = w$ is sent, then, following the decoding rule above, the error probality can be upper bounded as:

$$\begin{aligned}
\mathsf{P}_{\text{Error}} &= \Pr\left(\widehat{W} \neq w, \text{“Not Erasure”}|W = w\right) \\
&= \Pr\left(\widehat{W} \neq w|W = w, \text{“Not Erasure”}\right) \\
&\qquad \cdot \Pr\left(\text{“Not Erasure”}|W = w\right) \\
&= \Pr\left(\widehat{W} \neq w|W = w, \text{“Not Erasure”}\right) \\
&\qquad \cdot \underbrace{\left(1 - \Pr\left(\text{“Erasure”}|W = w\right)\right)}_{\leq 1} \\
&\leq \Pr\left(\widehat{W} \neq w|W = w, \text{“Not Erasure”}\right)
\end{aligned} \tag{7}$$

From (7), it is clear that the probability of error is upper bounded by $\Pr(\mathcal{E})$, since a decoding error occurs whenever the received sequence lies in the hyperspherical region $A_r(\mathbf{x}(w'))$ corresponding to a different codeword w' than the one sent w. This is the probability of undetected error, which can be upper bounded as that of having the received sequence outside the hypersphere of radius $(d_{\min} - r)$, where d_{min} is the minimum distance between two lattice points in $\mathcal{C}$, see Figure 1. Then:

$$\Pr(\mathcal{E}) \leq \int_{d_{\min}-r}^{\infty} Z_N(\theta) dJ_N(\theta) = Q\left((d_{\min} - r)^2 | N\right) \quad (8)$$

Next, since a sequence of lattices is said to be asymptotically good for packing if it achieves Minkowski-Hlawka's lower bound on the packing density $\rho^*(\text{pack})$ [15]: $\rho^*(\text{pack}) = \frac{r_{\text{pack}}}{r_{\text{eff}}} \geq \frac{1}{2}$, where r_{pack} is the *packing radius*, corresponding to the radius of the largest hypersphere contained in the Voronoi region of Λ as in Figure 1, Then, we have that $r_{\text{pack}} \geq \frac{1}{2} r_{\text{eff}}$. Therefore:

$$d_{\min} \geq 2 r_{\text{pack}} \geq r_{\text{eff}} = a$$

Next, we return to (8) and let $d_{\min} - r = r_e$. Note from the above bound and our choice of $r = \alpha a$, that:

$$r_e \geq a\,(1 - \alpha) \quad (9)$$

Hence, $\Pr(\mathcal{E}) \leq Q\left(r_e^2 | N\right)$. To evaluate this expression, we proceed as in [14], where the following approximation valid for $N > 100$ is used [16, Sec. 26.4.13, page 941]:

$$Q\left(\chi^2 | N\right) \approx Q(x_1), \;\; x_1 = \sqrt{2\chi^2} - \sqrt{2N - 1}.$$

Thus, we finally have:

$$\begin{aligned}
&\Pr(\mathcal{E}) \leq Q\left(\sqrt{2r_e^2} - \sqrt{2N-1}\right) \\
&\leq \frac{1}{2} \exp\left[-\frac{\left(\sqrt{2r_e^2} - \sqrt{2N-1}\right)^2}{2}\right] \\
&\overset{(a)}{\leq} \frac{1}{2} \exp\left[-\left([a(1-\alpha)] - \sqrt{N - 1/2}\right)^2\right] \\
&= \frac{1}{2} \exp\left[-\left(a^2(1-\alpha)^2 + N - \frac{1}{2} - 2a(1-\alpha)\sqrt{N - \frac{1}{2}}\right)\right]
\end{aligned}$$

where (a) results from (9). Next, using $a = \frac{\sqrt{NP}}{2^R}$ as in (5):

$$\begin{aligned}
&2\Pr(\mathcal{E}) \leq \\
&\exp\left[-\left(\frac{(1-\alpha)^2 NP}{2^{2R}} + N - \frac{1}{2} - \frac{2N(1-\alpha)\sqrt{P - \frac{P}{2N}}}{2^R}\right)\right] \\
&= \exp\left[-N\left((1-\alpha)^2 2^{2(C_\Lambda - R)} + 1 - \underbrace{\frac{1}{2N}}_{\to 0}\right.\right. \\
&\qquad\qquad \left.\left. -2\frac{(1-\alpha)}{2^R}\sqrt{P - \underbrace{\frac{P}{2N}}_{\to 0}}\right)\right] \\
&= \exp\left[-N\left((1-\alpha)^2 4^{(C_\Lambda - R)} + 1 - 2(1-\alpha)2^{\log_2 \frac{\sqrt{P}}{2^R}}\right)\right] \\
&= \exp\left[-N\left((1-\alpha)^2 4^{(C_\Lambda - R)} + 1 - 2(1-\alpha)2^{(C_\Lambda - R)}\right)\right],
\end{aligned}$$

$$= \exp\left[-N\left((1-\alpha)2^{(C_\Lambda - R)} - 1\right)^2\right], \quad (10)$$

where as in [14], $C_\Lambda = \log_2 \sqrt{P}$ represents a rate slightly smaller than the channel capacity, which is given by $C = \log_2 \sqrt{1 + P}$ (since we assumed $\sigma^2 = 1$).

Erasure probability: We evaluate this probability since it determines when retransmission requests occur. An "Erasure" event results from a sequence $\mathbf{y}$ being received in the space surrounding the decoding hyperspheres (of radius r) defined by (6) for each lattice point in $\mathcal{C}$. Let $B_r(\mathcal{C}) := \bigcup_{\mathbf{x} \in \mathcal{C}} A_r(\mathbf{x})$ describe the set formed by the union of hyperspherical decoding regions associated to each $\mathbf{x} \in \mathcal{C}$. Note that a retransmission must be triggered whenever an erasure is detected and reported via the feedback link. The decoder declares that an erasure occurred whenever the received sequence is an element of the set $B_r^c(\mathcal{C})$ (upperscript indicates the complement over $\mathbb{R}^N$):

$$B_r^c(\mathcal{C}) := \{\mathbb{R}^n \setminus \bigcup_{\mathbf{x} \in \mathcal{C}} A_r(\mathbf{x})\}$$

Let the erasure probability be denoted by P_s, then:

$$\mathsf{P}_s = \Pr\left(\mathbf{y} \in B_r^c(\mathcal{C})\right) = \Pr\left(E_1\right) - \Pr\left(\mathcal{E}\right) \leq \Pr\left(E_1\right). \quad (11)$$

Note that since the probability of event E_1 is much larger than that of event $\mathcal{E}$, the above inequality is a tight bound. We can further upper bound the term $\Pr(E_1)$ as follows:

$$\Pr\left(E_1\right) = \int_r^{\infty} Z_N(t) dJ_N(t) = Q\left(r^2 | N\right) \quad (12)$$

Thus, we have that (12) becomes:

$$\begin{aligned}
Q\left(r^2 | N\right) &\doteq Q\left(\sqrt{2r^2} - \sqrt{2N-1}\right) \\
&\leq \frac{1}{2} \exp\left[-\frac{\left(\sqrt{2r^2} - \sqrt{2N-1}\right)^2}{2}\right] \\
&= \frac{1}{2} \exp\left(-\left(r^2 + N - 1/2 - \sqrt{2r^2(2N-1)}\right)\right)
\end{aligned}$$

Next, from (5) and choosing $r = \alpha a$, we obtain an upper bound for the erasure probability $\mathsf{P}_s \leq \Pr(E_1)$:

$$\begin{aligned}
\Pr(E_1) &\leq \frac{1}{2} \exp\left(-(\alpha^2 a^2 + N - \frac{1}{2} - \sqrt{4N\alpha^2 a^2 - 2\alpha^2 a^2})\right) \\
&= \frac{1}{2} \exp\left[-\left(\alpha^2 \frac{NP}{4^R} + N - \frac{1}{2} - \sqrt{4N\alpha^2 \frac{NP}{4^R} - 2\alpha^2 \frac{NP}{4^R}}\right)\right] \\
&= \frac{1}{2} \exp\left[-N\left(\alpha^2 \frac{P}{4^R} + 1 - \underbrace{\frac{1}{2N}}_{\to 0} - 2\alpha\sqrt{\frac{P}{4^R} - \underbrace{\frac{P}{N4^R}}_{\to 0}}\right)\right] \\
&= \frac{1}{2} \exp\left[-N\left(\alpha^2 \frac{P}{4^R} + 1 - 2\alpha\frac{\sqrt{P}}{2^R}\right)\right] \\
&= \frac{1}{2} \exp\left[-N\left(\alpha^2 2^{(2\log_2\sqrt{P} - \log_2 2^{2R})}\right.\right. \\
&\qquad \left.\left. +1 - 2\alpha 2^{(\log_2\sqrt{P} - \log_2 2^R)}\right)\right]
\end{aligned}$$

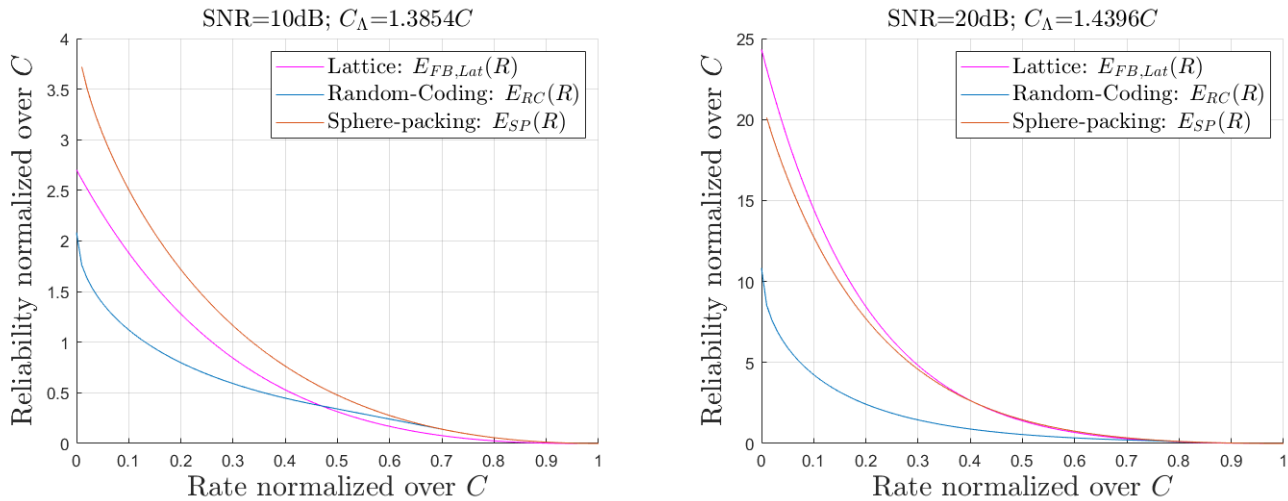

Fig. 2. Reliability of the AWGN channel as a function of rate for SNRs 10 and 20 dB. Our result, shown in magenta, is plotted only for $0 \leq R \leq C_\Lambda$.

$$= \frac{1}{2}\exp\left[-N\left(\alpha^2 4^{(C_\Lambda - R)} + 1 - 2\alpha 2^{(C_\Lambda - R)}\right)\right]$$
$$= (1/2)\exp\left[-N\left(\alpha 2^{(C_\Lambda - R)} - 1\right)^2\right]$$

Note that the bound above goes to zero as N increases.

Expected transmission time: The expected transmission time Δ is equal to the expected number of channel uses a message requires to be decoded (nor erased). Noting that the probability of retransmission is determined by the erasure probability, and since the occurrence of an erasure is independent for each block of length N, we have:

$$\begin{aligned}\mathsf{E}[\Delta] &= N + N\mathsf{P}_s + N\mathsf{P}_s^2 + \cdots \\ &= N\sum_{i=0}^{\infty}\mathsf{P}_s^i = N\frac{1}{1-\mathsf{P}_s} \approx N, \end{aligned} \tag{13}$$

since $\mathsf{P}_s \to 0$ as $N \to \infty$, see (11). This result indicates that on average, a single codeword of length N will be transmitted. Observe as well that the average rate $\bar{R} \to R$ as $N \to \infty$, and thus, the effective transmission rate is not affected by the repeat-request strategy. Next, we obtain the achievable error exponent based on this result which proves Theorem 1.

An achievable error exponent: An achievable error exponent results from utilizing (13) and the upper bound on the error probability (10):

$$\begin{aligned} E_{\text{FB,Lat}}(\bar{R}) &= \lim_{\mathsf{E}[\Delta]\to\infty}\frac{-1}{\mathsf{E}[\Delta]}\log \mathsf{P}_{\text{error}}\left(\bar{R}\right) \\ &\geq \lim_{\mathsf{E}[\Delta]\to\infty}\frac{-1}{\mathsf{E}[\Delta]}\log \Pr(\mathcal{E}) \\ &= (1-\alpha)^2 4^{(C_\Lambda - R)} + 1 - 2(1-\alpha)2^{(C_\Lambda - R)} \end{aligned} \tag{14}$$

Observe that the largest error exponent above can be obtained by picking $\alpha = \delta > 0$, where δ is a sufficiently small positive number:

$$E_{\text{FB,Lat}}(\bar{R}) \geq (1-\delta)^2 4^{(C_\Lambda - R)} + 1 - 2(1-\delta)2^{(C_\Lambda - R)}$$

The choice of $\alpha = \delta$ implies that the upper bound on the erasure probability is:

$$\mathsf{P}_s \leq \frac{1}{2}\exp\left[-N\left(\delta^2 4^{(C_\Lambda - R)} + 1 - 2\delta^2 2^{(C_\Lambda - R)}\right)\right]$$

The error exponent of Theorem (1) results from (14) and taking $\delta \to 0$.

IV. Numerical Simulation

Figure 2 shows a numerical simulation of our error exponent Theorem 1, along with the random coding lower bound $E_{\text{RC}}(R)$ and the sphere packing bound $E_{\text{SP}}(R)$ for block codes (recall that our schemes are for variable length codes). These are plotted using solid lines in magenta, blue and brown, respectively. The plots show two different SNR regimes, 10 and 20 dB, and both plots have rates and error exponents normalized by the channel capacity C.

In general, our scheme performs well for low rates and high SNR. Observe that it outperforms the random coding lower bound for lower rates for both SNRs we evaluated. In contrast, the sphere packing bound is not beaten at 10dB but only for a larger SNR. Note that at 20dB, the error exponent gains we obtain are significant at lower rates, where even the sphere packing bound is outperformed. However, recall that our scheme cannot operate at rates larger than $C_\Lambda < C$, and thus has an error exponent of zero for all rates in $C_\Lambda \leq R \leq C$. Nonetheless, note as well that as the SNR increases, C_Λ approaches C. Thus, one rule of thumb might be that for low rate, high SNR transmission (perhaps for low-rate control signaling, or whenever you have a strong channel but care more about reducing errors than rate) this type of single-bit retransmission strategy coupled with the new erasure-based lattice decoding strategy might be attractive.

V. Conclusions

We have proposed a coding scheme that builds on top of De Buda's original setting for rates below $C_\Lambda = \log_2\sqrt{P}$ using a lattice based code. Interestingly, we have conceived a way to incorporate a simple geometric argument to define an

erasure rule that exploits the natural structure of a lattice and that facilitates the retransmission of messages that the decoder finds ambiguous. We acknowledge the effect of missing the $+1$ term from the capacity expression of the Gaussian channel $C = \log_2 \sqrt{1+P}$, but emphasize that this loss is negligible at higher SNR (for which this scheme is attractive in the first place). The error exponents attained by our scheme in the high rate regime are not higher than those attained using other approaches, such as a simple open loop random code. However, at lower rates, our scheme exhibits much larger error exponents, especially at high SNR. Under VLC and decision feedback, the retransmissions carried over the channel in this scheme do not require high power – the AS power constraint is suffices, and there is no asymptotic rate penalty.

REFERENCES

[1] G. Forney, "Exponential error bounds for erasure, list, and decision feedback schemes," *IEEE Transactions on Information Theory*, vol. 14, no. 2, pp. 206–220, Mar. 1968.

[2] A. D. Wyner, "On the probability of error using a repeat-request strategy on the additive white gaussian noise channel," *The Bell System Technical Journal*, vol. 48, no. 1, pp. 71–86, 1969.

[3] J. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback–I: No bandwidth constraint," *IEEE Transactions on Information Theory*, vol. 12, no. 2, pp. 172–182, Apr 1966.

[4] R. G. Gallager and B. Nakiboğlu, "Variations on a Theme by Schalkwijk and Kailath," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 6–17, 2010.

[5] R. Mirghaderi, A. Goldsmith, and T. Weissman, "Achievable Error Exponents in the Gaussian Channel With Rate-Limited Feedback," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8144–8156, 2013.

[6] M. V. Burnashev and H. Yamamoto, "On optimal transmission strategies for channels with noiseless feedback," in *2016 IEEE International Symposium on Information Theory (ISIT)*, 2016, pp. 1282–1286.

[7] J. Schalkwijk and M. Barron, "Sequential signalling under a peak power constraint," *IEEE Transactions on Information Theory*, vol. 17, no. 3, pp. 278–282, May 1971.

[8] H. Yamamoto and K. Itoh, "Asymptotic performance of a modified Schalkwijk-Barron scheme for channels with noiseless feedback (Corresp.)," *IEEE Transactions on Information Theory*, vol. 25, no. 6, pp. 729–733, November 1979.

[9] A. Sato and H. Yamamoto, "Error exponents of discrete memoryless channels and awgn channels with noisy feedback," in *2010 International Symposium On Information Theory and Its Applications*, Oct 2010, pp. 452–457.

[10] K. Palacio-Baus and N. Devroye, "Variable-length coding error exponents for the awgn channel with noisy feedback at zero-rate," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 2254–2258.

[11] T. Liu, P. Moulin, and R. Koetter, "On error exponents of modulo lattice additive noise channels," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 454–471, 2006.

[12] U. Erez and R. Zamir, "Achieving 1/2 log (1+snr) on the awgn channel with lattice encoding and decoding," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2293–2314, 2004.

[13] ——, "Lattice decoded nested codes achieve the poltyrev exponent," in *Proceedings IEEE International Symposium on Information Theory,*, 2002, pp. 395–.

[14] R. de Buda, "The upper error bound of a new near-optimal code," *IEEE Transactions on Information Theory*, vol. 21, no. 4, pp. 441–445, 1975.

[15] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Transactions on Information Theory*, vol. 51, no. 10, pp. 3401–3416, 2005.

[16] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, 1972.