This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory

Achievable Error Exponents of One-Way and Two-Way AWGN Channels

Kenneth Palacio-Baus, Natasha Devroye

***Abstract*—Achievable error exponents for the one-way with noisy feedback and two-way AWGN channels are derived for the transmission of a finite number of messages M under almost sure (AS) and expected block (EXP) transmit power constraints. In the one-way setting under noisy AWGN feedback, under an AS power constraint, known linear and non-linear passive schemes are modified to incorporate AS constraints in the feedback link as well. In addition, a new active feedback scheme is presented in which the receiver feeds back the most likely pair of codewords, and the transmitter re-transmits which of these two was originally sent. This active feedback scheme outperforms one of the passive feedback schemes for all channel parameters; the linear scheme outperforms the others for low feedback noise variance. Under the EXP constraint, a known achievable error exponent for the transmission of two messages is generalized to any arbitrary but finite number of messages M through the use of simplex codes and erasure decoding. In the two-way AWGN setting, each user has its own message to send in addition to (possibly) aiding in the transmission of feedback for the opposite direction. Two-way error exponent regions are defined and achievable error exponent regions are derived for the first time under both AS and EXP power constraints. For the presented achievability schemes, feedback or interaction leads to error exponent gains in one direction, possibly at the expense of a decrease in the error exponents attained in the other direction. The relationship between M and n supported by our achievable strategies is explored.**

***Index Terms*—AWGN channel, two-way AWGN channel, error exponents, noisy feedback, zero-rate, error exponents region.**

I. INTRODUCTION

The reliability function [1]–[3], or error exponent, of a one-way channel characterizes the rate of decay of the probability of error when communicating one of 2^{nR} messages as

$$E(R) = \lim_{n\to\infty} -\frac{1}{n}\log \mathsf{P}_e^{(n)},$$

where $\mathsf{P}_e^{(n)}$ is the smallest probability of error that can be achieved by a code of rate R with block length n. Error exponents have been the subject of intense interest in both the absence [3], [4] and presence of feedback. If feedback is available, the transmitter is given access to a (possibly noisy, possibly encoded) function of the received output, that may dramatically increase the error exponents of one-way channels relative to when feedback is absent. Error exponents with ideal, noiseless, feedback have been considered for discrete memoryless channels (DMC) by Berlekamp [5] and for AWGN channels by Pinsker [6]. While even perfect feedback cannot increase the capacity of non-anticipatory DMC channels [7], it may greatly improve the error exponents achieved. This was first demonstrated for the AWGN channel under the EXP power constraint in [8], in which the probability of error was made to decay double-exponentially in the block length n; [9] later demonstrated a decay rate equal to any number of exponential levels. In recent years, it has been shown that even noisy feedback is useful in improving error exponents (though less dramatically), with a limited number of available results. In this article we continue this line of work and study error exponents of one-way additive white Gaussian noise (AWGN) channels with noisy AWGN feedback for the transmission of a finite number of messages (zero-rate).

We also extend the study of error exponents to the two-way channel, where two terminals exchange independent messages. Each transmitter's encoding function is *adaptive* or *interactive*, meaning that at time k it is a function of the message and the past available channel outputs. The capacity region of the two-way AWGN channel (with independent noise across terminals) is a rectangular region where both users may simultaneously attain their interference-free AWGN capacity, reflecting that adaptation at the transmitters is useless from a capacity perspective. Here we show that the same may not be said about error exponent regions: that is, adaptation *does* improve the error exponents of this two-way AWGN channel for the transmission of a finite number of messages. This is the first study of two-way error exponents to the best of our knowledge. This finite message regime is of initial interest as the exact error exponent for the one-way channel with noisy feedback is still open in this regime. Error exponents for positive rates are left for future work, and may extend existing work such as [8], [10]–[14]. We also initially study error exponents for block-coding-based codes rather than variable-length codes, as studied in [15]–[17] and the references therein for DMCs. This resolves synchronization issues that manifest when feedback is noisy and transmission is variable-length.

Error exponents of AWGN channels are sensitive to the type of power constraint imposed on the channel inputs and feedback. In this work, we consider two constraints, defined as the

Manuscript received December 3, 2018; revised February 16, 2020; accepted February 16, 2021;

This work was partially supported by NSF under awards 1645381 and 1815428. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NSF. Part of this work was presented at the IEEE International Symposium on Information Theory 2018.

Kenneth Palacio-Baus and Natasha Devroye are with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL 60607 USA (e-mail: kpalac2@uic.edu; devroye@uic.edu). K. Palacio-Baus is also with the University of Cuenca, Ecuador.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory

- Almost sure (AS) power constraint:

$$\mathsf{P}\left(\sum_{k=1}^{n} X_{i,k}^2 \leq nP_i\right) = 1 \tag{1}$$

- Expected block (EXP) power constraint:

$$E\left[\sum_{k=1}^{n} X_{i,k}^2\right] \leq nP_i. \tag{2}$$

In (1) and (2), $X_{i,k}$ corresponds to the k-th channel input of user i with power constraint P_i. The EXP constraint is less stringent than the AS constraint, and allows very high amplitude transmissions to occur with exponentially small probability. These rare events may correspond to decoding errors, and transmissions may be used to correct such errors, increasing achievable error exponents. The authors of [13] comment that this "trick" can not be used for general DMCs, but is useful for continuous-valued channels characterized by additive noise.

A. *Summary of Contributions*

We present results on the one-way and two-way AWGN channel error exponents under both the AS and EXP power constraints. Our contributions are as follows.

For the *one-way AWGN channel with noisy AWGN feedback*:

1) Theorems 1 and 2 each present an achievable error exponent for the one-way channel with AS constraints in both the forward and feedback links for a finite number of messages $M \geq 3$. Having power constraints in both directions is motivated by practical constraints, and to be suitable for the two-way channel. Theorem 1 is based on a modification of the non-linear coding scheme with an AS power constraint in the forward direction and *passive* (not power constrained) noisy feedback of [18, Theorem 1]. This modification scales the feedback signal to ensure that the feedback AS power constraint is met, and is shown in Appendix A. In Theorem 2, we propose a new *active* feedback scheme, proven in Section IV and extended for larger M in Appendix B. The resulting error exponent is shown to always outperform the modified non-linear *passive* feedback scheme of Theorem 1, and to lead to gains over the feedback-free exponent for a strong feedback link. This gains are due to feeding back the most likely pair of received codewords in an *encoded* fashion, using a simplex code.
2) Theorem 3 presents another achievable error exponent for the one-way channel with AS constraints in both the forward and feedback links for a finite number of messages $M \geq 2$. This a modification of the *linear passive* feedback coding scheme (not power constrained in the feedback link) presented in [18, Theorem 2], in which the AS constraint is now imposed in the feedback direction as well. Its proof is summarized in Appendix C. This linear passive feedback scheme outperforms the exponents achieved in Theorems 1 and 2 for very high signal to noise ratios (SNR) in the feedback link.
3) Theorem 4 presents an achievable error exponent for the one-way channel with EXP constraints in both the forward and feedback links for a finite number of messages M. This is a generalization of the Kim-Lapidoth-Weissman [19] achievable error exponent for the transmission of two messages to the transmission of any finite number of messages $M \geq 2$. This extension, proven in Section V, is based on the use of simplex codes, active feedback, and a geometric approach based on erasure decoding and "NACK bands / volumes".

For the *two-way AWGN channel*:

5) Theorems 5, 6 and 7 demonstrate – for the first time – achievable error exponent regions for the two-way AWGN channel under the AS power constraint, provided that one channel's SNR is better than the other. This non-symmetric SNR scenario is of interest since for the one-way AWGN channel under the AS constraint, only a feedback link significantly stronger than the forward has been shown to lead to error exponent gains over feedback-free transmissions. These results follow as a direct application of Theorem 2 for active feedback and Theorems 1 and 3 for passive feedback.
6) Theorem 8 demonstrates – again for the first time – an achievable error exponent region for the two-way channel under EXP power constraints in both directions. This result follows from direct application of time / power sharing of the scheme introduced in Theorem 4 in each direction. In contrast to the results under AS constraints, the regions here are valid for all SNRs in each direction.

The achievable error exponents for the transmission of a finite number of messages M are based on the use of a simplex code for the non-feedback transmissions. The use of simplex codes allows for a geometry-based upper bound on the probability of error that can be visualized for a small number of messages ($M = 3$) and extended to any finite $M \geq 3$. Under the EXP constraint, some feedback and retransmission signals occurring with exponentially small probability employ very high amplitude signals to ensure an exceedingly small probability of error.

B. *Article outline*

The remainder of this article is organized as follows. Section II presents a summary of known error exponents for the One-Way and Two-Way AWGN channels. Section III presents our findings: In Subsection III-A we present new achievable error exponents for the one-way AWGN channel in Theorems 1, 2, 3, and 4. In Subsection III-B, we present Theorems 5, 6, 7 and 8 as the main contributions, which follow from the one-way results. A short discussion on outer bounds is presented in Subsection III-C. The proofs of Theorems 2 and 4 for the one-way channel and the transmission of three messages are shown respectively in Sections IV and V, and later extended to general $M \geq 3$ in Appendices B and D. Section VI addresses the relation between the number of messages M and the block length n for our proposed schemes. Numerical simulations are presented in Section VII. Section VIII concludes the paper.

Notation. We use $\mathsf{P}_w(x)$ to indicate the probability of any event x conditioned on the transmission of message $W = w$, i.e. the probability of error given that $W = 1$ has been sent is $\mathsf{P}(\text{error}|w = 1) = \mathsf{P}_1(\text{error})$. The length of a sequence is denoted using a superscript, i.e. $x_i^{\lambda n}$ denotes that sequence x lasts for λn channel uses. Subscript i indicates the terminal that generated the sequence. Error exponents are denoted by E, accompanied by a superscript to denote the forward direction power constraint, and a subscript to denote the feedback link power constraint, if applicable. We use $\mathsf{E}[\cdot]$ to denote the expectation operator. Random variables are upper case letters, taking on instances in lower cases from alphabets in calligraphic font (random variable X takes on $x \in \mathcal{X}$). We use $a_n \doteq b_n$ to indicate that $\frac{1}{n}\ln(\frac{a_n}{b_n}) \to 0$ as $n \to \infty$. We use the terms *terminal* and *user* interchangeably. For one-way channels, our definitions refer to systems with feedback; definitions in the absence of feedback should be clear from context and omission of the corresponding feedback-related terms.

Achievable transmission schemes are based on simplex codes denoted by $\mathcal{C}(\Omega_M, \mathcal{E})$ for the feedback-free transmission of M messages $\omega \in \Omega_M := \{1, \cdots, M\}$. This code consists of a constellation of M data symbols $X^n(\omega)$, each of energy $\mathcal{E} = \sum_{i=1}^n X_i(\omega)^2$, which are the vertices of an M-simplex tetrahedron. As indicated in [18, Appendix], the codewords of the simplex code $\mathcal{C}(\Omega_M, \mathcal{E})$ of length n are given by (3). Note that the first $M-1$ transmissions correspond to the coordinates of the data symbols of the M-simplex, and the remaining entries are zero. $M-1$ dimensions are thus required to represent each symbol.

$$X^n(\omega) = A\left(e_\omega - \frac{1}{M}\sum_{\omega=1}^{M} e_\omega\right) \quad \text{for } \omega \in \{1, \cdots, M\} \tag{3}$$

where $A = \sqrt{\mathcal{E}\frac{M}{M-1}}$, with $e_\omega = (\underbrace{0, \cdots, 0}_{\omega-1}, 1, 0, \cdots, 0)$.

II. One-Way and Two-Way AWGN channels: error exponents at zero-rate background

A. The one-way AWGN channel

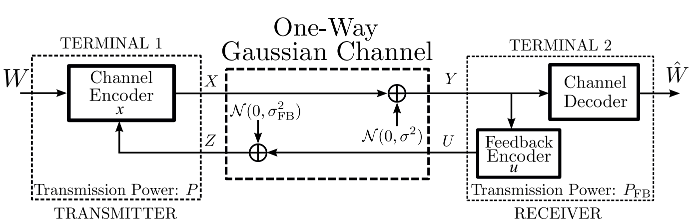


Fig. 1. One-way AWGN channel with active feedback.

Figure 1 shows the one-way AWGN channel with active feedback, where terminal 1 (T1 / transmitter) sends message W, selected uniformly from the set $\mathcal{W} = \{1, \cdots, M\}$, to terminal 2 (T2 / receiver) using a code of block length n. Both communication directions are independent channels characterized by AWGN of zero mean and respective variances σ^2 and σ^2_{FB}, identically distributed and independent across users and channel uses. The channel model for the k-th time slot is described by (4) and (5), with inputs subject to a $\Psi \in \{\text{AS}, \text{EXP}\}$ power constraint according to (1) and (2). Noiseless feedback is captured by $\sigma^2_{\text{FB}} \to 0$. No feedback is captured by $\sigma^2_{\text{FB}} \to \infty$.

$$Y_k = X_k + N_k, \qquad N_k \sim \mathcal{N}(0, \sigma^2), \;\; k = 1, ..., n \tag{4}$$

$$Z_k = U_k + N_{\text{FB}_k}, \quad N_{\text{FB}_k} \sim \mathcal{N}(0, \sigma^2_{\text{FB}}), \;\; k = 1, ..., n. \tag{5}$$

Let $\mathcal{X}, \mathcal{Y}, \mathcal{U}, \mathcal{Z}$ all be the set of reals and $\left(M, P, \sigma^2, P_{\text{FB}}, \sigma^2_{\text{FB}}, n\right)$ be a code for the transmission of M messages consisting of n forward and feedback encoding functions[1], $x_k : \mathcal{W} \times \mathcal{Z}^{k-1} \to \mathcal{X}$ and $u_k : \mathcal{Y}^k \to \mathcal{U}$ (for $k = 1, ..., n$), leading to channel inputs $X_k = x_k\left(W, Z^{k-1}\right)$ and $U_k = u_k\left(Y^k\right)$, and a decoding rule $\phi : \mathcal{Y}^n \to \{1, \cdots, M\}$ that determines the best estimate of the transmitted message W, denoted by $\hat{W}$. Let $\mathsf{P}_e^\Psi\left(M, P, \sigma^2, P_{\text{FB}}, \sigma^2_{\text{FB}}, n\right) := \frac{1}{M}\sum_{w=1}^{M} \Pr(\phi(y^n) \neq w \mid W = w \text{ sent})$ denote the probability of error attained by a particular $\left(M, P, \sigma^2, P_{\text{FB}}, \sigma^2_{\text{FB}}, n\right)$ code under a $\Psi \in \{\text{AS}, \text{EXP}\}$ power constraint. We define the achievable error exponent for the one-way AWGN channel with feedback under the Ψ power constraint as:

$$E^\Psi_{\text{FB}}(M, P, \sigma^2, P_{\text{FB}}, \sigma^2_{\text{FB}}) := \liminf_{n\to\infty} -\frac{1}{n}\log \mathsf{P}_e^\Psi\left(M, P, \sigma^2, P_{\text{FB}}, \sigma^2_{\text{FB}}, n\right),$$

where the subscript FB indicates the presence of feedback and is omitted for the non-feedback case, following the notation of [19]. Literature involving rate-limited noiseless feedback is beyond the scope of this work but important contributions for the one-way AWGN channel can be found in [20], [21]. Contributions for multi-user settings not considered here, such as the broadcast and interference Gaussian channels with feedback can be found in [22]–[26]. Further results for the positive rate regime for point-to-point channels can be found in [11], [27], [28], while noisy feedback is studied in [29]. Error exponents under the variable length coding (VLC) setting for the AWGN channel with perfect feedback have been addressed in [16], [30]–[32]; noisy feedback is considered in [17], [33].

Next, we present a review of error exponents results for the one-way AWGN channel with noisy feedback and the transmission of a finite number of messages:

1) ***Error exponents for the one-way AWGN channel without feedback and with perfect output feedback:*** Error exponents for the transmission of a finite number messages without feedback and with perfect output feedback serve as lower and upper bounds on those achievable with noisy feedback. The feedback-free setting ($\sigma^2_{\text{FB}} \to \infty$) was studied by Shannon [4], who showed that for the transmission of M messages under the AS power constraint, the best achievable error exponent is attainable using a simplex code and described by (6):

$$E^{\text{AS}}\left(M, P, \sigma^2\right) = \frac{P}{\sigma^2}\frac{M}{4(M-1)}. \tag{6}$$

Note that $E^{\text{AS}}\left(2, P, \sigma^2\right) = \frac{P}{2\sigma^2}$, and $E^{\text{AS}}\left(3, P, \sigma^2\right) = \frac{3}{8}\frac{P}{\sigma^2}$, while for large M, $E^{\text{AS}}\left(M, P, \sigma^2\right) \approx \frac{P}{4\sigma^2}$. For the perfect-

[1]Note that the forward and backward encoding functions shown here as x_k and u_k, are denoted in [19] as f_k and g_k respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory

feedback setting ($\sigma^2_{\text{FB}} \to 0$), Pinsker [6] showed that for $M \geq 3$ messages under the AS power constraint, the error exponent can be improved up to:

$$E^{\text{AS}}_{\text{FB}}\left(M, P, \sigma^2, P_{\text{FB}}, \sigma^2_{\text{FB}} = 0\right) = \frac{P}{2\sigma^2}. \tag{7}$$

Evaluating (6) for $M = 2$ leads to (7), hence no improvements over the feedback-free error exponent are possible subject to the AS power constraint for $M = 2$, even using perfect feedback [34], [35]. Bound (7) also upper bounds the error exponent under the AS constraint for the one-way AWGN channel with noisy feedback.

*2) **Error exponents for the One-Way AWGN channel with noisy feedback**:* Feedback over a noisy channel is called *passive* when the forward channel output observed at the receiver is directly sent to the source uncoded (perhaps scaled to meet a power constraint) and *active* when the k-th feedback channel input is a function u_k (see Figure 1) of the received outputs up until time k, i.e. $U_k = u_k(Y^k)$. To mimic passive feedback, u_k can be modeled as a one-to-one scaling of signal Y_k. Thus feedback provides a noisy version of the observed forward channel outputs $Z_k = \tilde{Y}_k$.

Error exponents under passive feedback: Yamamoto and Burnashev obtained error exponents for the AWGN channel and passive noisy feedback under the AS power constraint for the transmission of a non-exponentially growing number of messages M (zero-rate) using block coding (rather than variable-length coding) [36]–[38]. Their work extends previous techniques used in the Binary Symmetric Channel (BSC) with noisy feedback [39], [40] to the AWGN channel. Under the AS power constraint and for a large number of messages M, such that $\ln M = o(n)$, $n \to \infty$ and $\sigma^2_{\text{FB}} \to 0$, the following error exponent is achievable [36, Theorem 1 (a)]:

$$E^{\text{AS}}_{\text{FB}}(M, P, \sigma, \sigma^2_{\text{FB}}) \geq \frac{PM}{4\sigma^2(M-1)}\left[1 + \frac{1}{2+\sqrt{5}} - \frac{1}{2M} + o(1)\right]. \tag{8}$$

As $\sigma^2_{\text{FB}} \to \infty$, (9) is achievable and strictly larger than the feedback-free error exponent [36, Theorem 1 (b)]:

$$\begin{aligned} E^{\text{AS}}_{\text{FB}}(M, P, \sigma, \sigma^2_{\text{FB}}) &\geq \frac{PM}{4\sigma^2(M-1)}\left[1 + \frac{1}{56\sigma^2_{\text{FB}}} + O(\sigma^{-4}_{\text{FB}})\right] \\ &> E(M, P, \sigma, \sigma^2_{\text{FB}} = \infty) = \frac{PM}{4\sigma^2(M-1)}. \end{aligned} \tag{9}$$

For very small feedback noise (8) was improved as follows for $M \to \infty$ with $\ln M = o(n)$ as $n \to \infty$ in [37], yielding a larger improvement over feedback-free than that of (8):

$$E^{\text{AS}}_{\text{FB}}\left(M \to \infty, P, \sigma^2, \sigma^2_{\text{FB}}\right) \geq \frac{P}{3\sigma^2}\left(1 - \sigma^2_{\text{FB}}\right).$$

Kim, Lapidoth, and Weissman derived bounds on the reliability function of the AWGN channel with AWGN passive noisy feedback of very small and non-zero noise variance $\sigma^2_{\text{FB}} = \epsilon^2 : 0 < \epsilon^2 \ll 1$ in [41]. They presented bounds on the error exponent for $M = 2$ under the expected average power constraint $\frac{1}{n}\sum_{i=1}^{n} \mathsf{E}\left[X_i^2\right] \leq P$, as:

$$\frac{P}{2\sigma^2_{\text{FB}}} \leq E^{\text{EXP}}_{\text{FB}}\left(M = 2, P, \sigma^2 = 1, \sigma^2_{\text{FB}}\right) \leq \frac{P(1+\sigma^2_{\text{FB}})}{2\sigma^2_{\text{FB}}}.$$

Passive feedback is also exploited by Xiang and Kim in [18], where achievable error exponents for the transmission of $M \geq 3$ messages were presented under the constraint $\mathsf{P}[\sum_{i=1}^{n} x_k^2(w, \tilde{Y}^{k-1}) \leq nP] = 1, \forall w \in \{1, \cdots, M\}$ on the forward channel and noisy AWGN feedback. They proposed passive (simply send back the output signal over the noisy feedback channel without any power constraints) *non-linear* and *linear* feedback schemes presented below, and our altered versions – in which a symbol-by-symbol (hence passive) scaled version of the output symbols is fed back and shown to meet the AS power constraint in the feedback direction – in Section III, as a part of our contributions. The error exponents are re-computed for these altered schemes.

The non-linear passive scheme with an AS power constraint in the forward direction, and no power constraint in the (passive, noisy) feedback direction of [18] achieves a lower bound on the error exponent for general $M \geq 3$ as (10), outperforming the feedback-free error exponent of (6) when $\sigma^2_{\text{FB}} < \frac{\sigma^2}{4}$.

$$\begin{aligned} &E^{\text{AS-NL}}_{\text{FB}}\left(M, P, \sigma^2, s\right) \geq \\ &\quad \min\left\{M\frac{P}{2\sigma^2}\frac{s^2-2s+4}{M(s^2-2s+4)+3(M-2)}, \right. \\ &\quad \left. \underbrace{\frac{P}{\sigma^2_{\text{FB}}}\frac{3M}{8}\frac{s^2}{M(s^2-2s+4)+3(M-2)}}_{\text{passive feedback}}\right\} \end{aligned} \tag{10}$$

where, $s \in [0,1]$ is chosen to equate both terms in (10), resulting in the optimal choice: $s = \frac{1-\sqrt{3\frac{\sigma^2}{\sigma^2_{\text{FB}}}-3}}{1-\frac{3}{4}\frac{\sigma^2}{\sigma^2_{\text{FB}}}}$. Note that (10) is presented in a form that explicitly shows the contribution to the error exponent of the forward and feedback transmissions respectively. In Theorem 1 we show how this non-linear passive scheme may be altered to accommodate an AS power constraint in the feedback link as well.

The other coding scheme proposed in [18], which we refer to as the *linear coding scheme*, provides an error exponent that outperforms the non-linear scheme when the noise variance in the feedback link is very low. For the transmission of $M \geq 2$ messages, [18, Theorem 2] shows:

$$\begin{aligned} &E^{\text{AS-LIN}}_{\text{FB}}(M, P, \sigma^2 = 1, \sigma^2_{\text{FB}}) \\ &\geq \frac{P}{2}\frac{1}{1+\sigma^2_{\text{FB}}+4\left(\lfloor M/2\rfloor\right)^2\sigma^2_{\text{FB}}+4\left(\lfloor M/2\rfloor\right)\sqrt{\sigma^2_{\text{FB}}(1+\sigma^2_{\text{FB}})}} \\ &\geq \frac{P}{2}\frac{1}{\left(\sqrt{\sigma^2_{\text{FB}}}M+\sqrt{1+\sigma^2_{\text{FB}}}\right)^2}, \end{aligned}$$

where the super-script AS-LIN refers to linear coding scheme under AS. In Theorem 3 we show how this scheme can be extended to include the AS constraint in the feedback link as well.

Error exponents under active feedback: Error exponents under active noisy feedback (see Figure 1) were presented in [19] for the EXP power constraint for the forward channel and both AS and EXP power constraints for the feedback link and $M = 2$ messages. The corresponding lower and upper bound results are:

- **AS feedback power constraint**: $\sum_{k=1}^{n} u_k^2(Y^k) \leq nP_{\text{FB}}$

$$E^{\text{EXP}}_{\text{FB}^{\text{AS}}}\left(2, P, \sigma^2, P_{\text{FB}}, \sigma^2_{\text{FB}}\right) \geq \frac{P}{2\sigma^2} + \frac{2P_{\text{FB}}}{\sigma^2_{\text{FB}}},$$

$$E^{\text{EXP}}_{\text{FB}^{\text{AS}}}\left(2, P, \sigma^2, P_{\text{FB}}, \sigma^2_{\text{FB}}\right) \leq \frac{P}{2\sigma^2} + \frac{2\sqrt{\left(P_{\text{FB}} + \sigma^2_{\text{FB}}\right) P_{\text{FB}}}}{\sigma^2_{\text{FB}}}.$$

- **EXP feedback power constraint**: $\mathsf{E}\left[\sum_{k=1}^{n} u_k^2(Y^k)\right] \leq nP_{\text{FB}}$

$$E^{\text{EXP}}_{\text{FB}^{\text{EXP}}}\left(2, P, \sigma^2, P_{\text{FB}}, \sigma^2_{\text{FB}}\right) \geq 2\left(\frac{P}{\sigma^2} + \frac{P_{\text{FB}}}{\sigma^2_{\text{FB}}}\right), \quad (11)$$

$$E^{\text{EXP}}_{\text{FB}^{\text{EXP}}}\left(2, P, \sigma^2, P_{\text{FB}}, \sigma^2_{\text{FB}}\right) \leq \frac{\left(\sqrt{P + \sigma^2} + \sqrt{P}\right)^2}{\sigma^2} + \frac{\left(\sqrt{P_{\text{FB}} + \sigma^2_{\text{FB}}} + \sqrt{P_{\text{FB}}}\right)^2}{\sigma^2_{\text{FB}}}.$$

3) A discussion on active and passive feedback: The various achievable error exponents for the one-way AWGN channel with noisy feedback strongly depend on the type of power constraint imposed for both directions. For example, the more flexible EXP constraint yields an error exponent at least four times larger than the feedback-free exponent for any M when noisy (active) feedback is used. This may be seen from Theorem 4 by comparing (20) with (6). This differs remarkably from conclusions under the AS constraint, where even with perfect (passive) feedback, the feedback-free error exponent may only be improved up to $\frac{P}{2\sigma^2}$, see (7), which does not depend on the number of messages being sent and is attainable only for $M = 2$ without the use of feedback. If perfect feedback is available, Pinsker's [6] error exponent $\frac{P}{2\sigma^2}$ is achievable for any $M \geq 2$, such that $\lim_{n\to\infty} \frac{\log M}{n} = 0$. The *linear-scheme* of [18] approaches Pinsker's error exponent as $\sigma^2_{\text{FB}} \to 0$ using passive noisy feedback.

B. The two-way AWGN channel

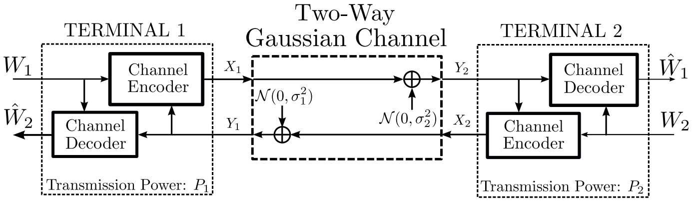


Fig. 2. Two-way AWGN channel.

The two-way AWGN channel was introduced by Shannon [42] and further studied by Han [43]. Its capacity region (with independent noise across terminals) is known [43], [44], and is a rectangular region where both users may simultaneously attain their interference-free AWGN capacity (denoted by $R_{12} \leq C_{12}$ and $R_{21} \leq C_{21}$ for each direction respectively). The capacity region cannot be increased by interaction or adaptation between the two terminals. The question pursued here is whether the same is true of error exponents. Figure 2 shows the two-way AWGN channel, comprising two users denoted by terminal i, for $i \in \{1, 2\}$, each transmitting message W_i, uniformly selected from $\mathcal{W}_i := \{1, \cdots, M\}$ to terminal $(3 - i)$ using a block code of block length n. The general two-way AWGN channel model is given by (12), characterizing the channel output observed at the i-th terminal

$$Y_i = X_i + a_i X_{3-i} + N_i, \quad (12)$$

where, a_i is a constant, $X_i \in \mathbb{R}$ corresponds to channel inputs satisfying an input block power constraint, $Y_i \in \mathbb{R}$ corresponds to channel outputs and $N_i \sim \mathcal{N}(0, \sigma_i^2)$ to zero-mean Gaussian noise processes, each independent and identically distributed across channel uses. The model of (12) can be simplified by noting that each terminal can subtract its own signal X_i. Hence, the two-way AWGN channel may equivalently be represented as

$$Y_i = X_{3-i} + N_i; \quad N_i \sim \mathcal{N}(0, \sigma_i^2).$$

We characterize error exponents for the zero-rate operational rate pair, i.e. for $(R_{12}, R_{21}) = (0, 0)$, and for different ratios between the SNRs in the two directions. Let $\frac{P_i}{\sigma^2_{(3-i)}}$ be the signal-to-noise ratio for link $i \to (3 - i)$. We consider both symmetric $\left(\frac{P_1}{\sigma_2^2} = \frac{P_2}{\sigma_1^2}\right)$ and non-symmetric $\left(\frac{P_1}{\sigma_2^2} \neq \frac{P_2}{\sigma_1^2}\right)$ channels. The latter case is of particular importance under the AS power constraint, since only a feedback channel stronger that the forward direction leads to gains over feedback-free error exponents [18], [36], [38]. We have been unable to achieve error exponents greater than without feedback for symmetric SNR channels under the AS constraint. We formalize the problem next.

Let $\left(M, P_1, \sigma_2^2, P_2, \sigma_1^2, n\right)$ be a code of block length n which consists of two encoding and two decoding rules as in Figure 2. The i-th terminal's encoding rule consists of n functions, defined for the k-th channel use as:

$$x_{i,k} : \{1, \cdots, M\} \times Y_i^{k-1} \to \mathcal{X}_i, \quad \text{for } k = 1, ..., n,$$

leading to the k-th channel inputs for terminal i, $X_{i,k} = x_{i,k}\left(W_i, Y_i^{k-1}\right)$. Decoding rules are denoted by ϕ_i, and estimate the received message based on the sequence Y_i^n as:

$$\phi_i : \mathcal{Y}_i^n \times \mathcal{W}_i \to \mathcal{W}_{3-i}, \quad \text{for } i = 1, 2.$$

The probability of error simultaneously achieved by a particular $(M, P_1, \sigma_2^2, P_2, \sigma_1^2, n)$ code in the forward and backward directions under a $\Psi \in \{\text{AS}, \text{EXP}\}$ power constraint is denoted by $\mathsf{P}^{\Psi}_{e_{12}}\left(M, P_1, \sigma_2^2, P_2, \sigma_1^2, n\right) = \frac{1}{M^2}\sum_{w_1,w_2} \mathsf{P}(\phi_2(y_2^n, w_2) \neq w_1 | w_1, w_2 \text{ sent})$ and equivalently $\mathsf{P}^{\Psi}_{e_{21}}\left(M, P_1, \sigma_2^2, P_2, \sigma_1^2, n\right)$.

Definition 1. *A pair of error exponents $(E_{12}, E_{21})^{\Psi}$ is achievable for the transmission of a finite number of messages M under Ψ power constraint for the two-way AWGN channel, if there exists an $\left(M, P_1, \sigma_2^2, P_2, \sigma_1^2, n\right)$ code such that for large*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory

n, simultaneously

$$-\frac{1}{n}\log \mathsf{P}^{\Psi}_{e_{12}}\left(M,P_1,\sigma_2^2,P_2,\sigma_1^2,n\right) \geq E^{\Psi}_{12}\left(M,P_1,\sigma_2^2,P_2,\sigma_1^2\right),$$

$$-\frac{1}{n}\log \mathsf{P}^{\Psi}_{e_{21}}\left(M,P_1,\sigma_2^2,P_2,\sigma_1^2,n\right) \geq E^{\Psi}_{21}\left(M,P_1,\sigma_2^2,P_2,\sigma_1^2\right).$$

Definition 2. *The error exponent region (EER) for the two-way AWGN channel and the transmission of M messages corresponds to the union over all achievable error exponent pairs $(E_{12},E_{21})^{\Psi}$, where we will often drop the arguments of E_{ij} for simplicity and sometimes we may refer to $(E_{12},E_{21})^{AS}$ as $(E_{12}^{AS},E_{21}^{AS})$.*

Two-way error exponents have not been studied to the best of our knowledge. An initial characterization for the two-way AWGN was presented by the authors in [45] which we extend here. In this channel, each terminal intuitively performs two tasks: 1) transmission of their own message, and 2) transmission of feedback information for the other terminal, which may consume part of its available power. Our achievable error exponent regions suggest that gains resulting from interaction may come at the price of a reduction of the error exponent of the terminal providing feedback.[2]

III. Contributions

We outline our contributions in a series of Theorems in this section; later sections provide the proofs, remarks on extensions to larger M, and numerical evaluations.

A. Achievable error exponents for the one-way AWGN channel with noisy feedback

We present new achievable error exponents for the one-way AWGN channel for the transmission of a finite number of messages M under the AS and EXP power constraints and both passive and active AWGN feedback.

Theorem 1. *Under the AS power constraint in both directions of a one-way AWGN channel with non-linear passive noisy feedback and the transmission of finite $M \geq 3$ messages, an achievable error exponent for $s \in [0,1]$ is determined by* (13) *To maximize this expression, take $s \in [0,1]$ that equates both terms in* (13) *as $s = \frac{1-\sqrt{3\frac{\sigma^2}{\sigma_{FB}^2}\frac{P_{FB}}{P+5\sigma^2+\sigma\sqrt{P}}-3}}{1-\frac{3}{4}\frac{\sigma^2}{\sigma_{FB}^2}\frac{P_{FB}}{P+5\sigma^2+\sigma\sqrt{P}}}$.*

Remark 1. *The left term of the* min *operation in* (13) *leads to an error exponent strictly larger than the feedback-free error exponent presented in* (6) *for any choice of $s<1$, and attains its largest value at $s=0$, since the factor in parentheses is monotonically decreasing with s for all M. The right term attains its largest value when $s=1$, since the term in parentheses is monotonically increasing with s for all M. Hence, a critical point (denoted as $\sigma^2_{FB^2}$ in Figure 8) exists for which, for $\sigma_{FB}^2 < \frac{\sigma^2}{4}\frac{P_{FB}}{P+5\sigma^2+\sigma\sqrt{P}}$ an error exponent strictly larger than the feedback-free error exponent is achievable.*

[2]Since we operate at zero-rate, the tradeoff with rate is not captured. However, the EER does generally depend on the number of messages M being sent in each direction. Extensions to different numbers of messages in the two directions is left for future work.

This Theorem is shown in Appendix A and follows by scaling the feedback signal used in the non-linear passive scheme of [18] achieving the exponent in (10) by a factor which ensures the AS power constraint of P_{FB} is met in the feedback direction. This scaling alters the achieved error exponent, resulting in (13). We note that this error exponent shows a dependence on the feedback power constraint P_{FB}, whereas the original non-linear passive scheme of [18, Theorem 1] in (10) does not impose a power constraint (nor do they scale the received signal) in the feedback direction and hence does not. The relation between both results is presented in the following remark.

Remark 2. *The achievable error exponent in* (13) *(non-linear passive feedback, AS feedback power constraint) yields that of* (10) *from [18, Theorem 1] (passive feedback, no feedback power constraint) when $P_{FB} = P+5\sigma^2+\sigma\sqrt{P}$. It is easy to show that if $P_{FB} > P+5\sigma^2+\sigma\sqrt{P}$, then the scaled passive scheme's error exponent in* (13) *outperforms that of* (10)*; the reverse is true otherwise, though the schemes should perhaps not be compared as one has a power constraint in the feedback direction and the other does not.*

The use of active feedback leads to our next contribution:

Theorem 2. *Under the AS power constraint in both directions of a one-way AWGN channel with non-linear active noisy feedback and the transmission of finite $M \geq 3$ messages, an achievable error exponent for $s \in [0,1]$ is determined by* (14).

Remark 3. *The largest achievable error exponent for the active feedback coding scheme expression in* (14)*, occurs when $s=0$, thus:*

$$E^{AS}_{FB^{AS}}\left(M,P,\sigma^2,P_{FB},\sigma^2_{FB}\right) \geq \min\left\{\frac{P}{\sigma^2}\frac{2M}{7M-6)},\frac{P_{FB}}{\sigma_{FB}^2}\frac{M(M-1)}{4\left(M(M-1)-2\right)}\right\}. \quad (16)$$

This expression results from the use of active feedback, as opposed to passive as done in the linear and non-linear noisy coding schemes of [18]. We use three non-overlapping stages: first an initial transmission of message W in $\lambda(n-1)$ channel uses using a simplex code of size M. As in [18], if the received sequence is in a protection region, then all remaining transmissions are ignored until the next message is sent. Otherwise, an active feedback stage of duration $(1-\lambda)(n-1)$ using a simplex code of size $\binom{M}{2}$ is used to feed back the most likely pair of messages candidates the receiver has determined (denoted by $q=\{\hat{w}_1,\hat{w}_2\}$, where $\hat{w}_1,\hat{w}_2 \in \{1,\cdots,M\}$ and represents the two closest (minimum distance) messages to the sequence received in the first stage). Once q is decoded by the transmitter as $\hat{q}=\{\tilde{w}_1,\tilde{w}_2\}$, a single channel use retransmission signal is sent to indicate whether the true message W is the first or second element in $\hat{q}$. If W is not in $\hat{q}$, this counts as an error, and nothing is sent to the receiver.

Remark 4. *The resulting expressions of Theorems 1 and 2 only differ in the term of the right of the* min *operation, which*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory

$$E^{\text{AS}}_{\text{FB}^{\text{AS}}}\left(M,P,\sigma^2,P_{\text{FB}},\sigma^2_{\text{FB}},s\right) \geq \min\Bigg\{ M\frac{P}{2\sigma^2}\left(\frac{s^2-2s+4}{M(s^2-2s+4)+3(M-2)}\right), \underbrace{\frac{P}{\sigma^2_{\text{FB}}}\frac{3M}{8}\left(\frac{s^2}{M(s^2-2s+4)+3(M-2)}\right)\frac{P_{\text{FB}}}{P+5\sigma^2+\sigma\sqrt{P}}}_{\text{passive feedback}}\Bigg\}. \tag{13}$$

$$E^{\text{AS}}_{\text{FB}^{\text{AS}}}\left(M,P,\sigma^2,P_{\text{FB}},\sigma^2_{\text{FB}},s\right) \geq \min\left\{ M\frac{P}{2\sigma^2}\left(\frac{s^2-2s+4}{M(s^2-2s+4)+3(M-2)}\right), \underbrace{\frac{P_{\text{FB}}}{\sigma^2_{\text{FB}}}\frac{M(M-1)}{4\left(M(M-1)-2\right)}}_{\text{active feedback}}\right\}. \tag{14}$$

$$E^{\text{AS-LIN}}_{\text{FB}^{\text{AS}}}\left(M,P,\sigma^2,\sigma^2_{\text{FB}},\alpha\right) \geq \frac{P}{2}\frac{1}{\sigma^2+\frac{\sigma^2_{\text{FB}}}{\alpha^2}+4\left(\lfloor M/2\rfloor\right)^2\sigma^2_{\text{FB}}/\alpha^2+4\left(\lfloor M/2\rfloor\right)\sqrt{\frac{\sigma^2_{\text{FB}}}{\alpha^2}\left(\sigma^2+\frac{\sigma^2_{\text{FB}}}{\alpha^2}\right)}}. \tag{15}$$

depends on how feedback is handled. In both schemes, P_{FB} can be chosen to be large enough such that (13) *and* (14)*, are dominated by the term on the left, which is the same for both. This term is monotonically decreasing in s, maximized as $s \to 0$, over the valid choices of $s \in [0,1]$. The active feedback scheme is able to achieve this maximal error exponent at a lower P_{FB} than that needed by the non-linear passive scheme.*

We can also directly show that the non-linear passive scheme's error exponent is always upper bounded by the active scheme. To do so, we show that the right term in the min function of the passive scheme -(13)- is no larger than the right hand side of the min function of the active scheme -(14)-,

$$\begin{aligned}
&\frac{P}{\sigma^2_{\text{FB}}}\frac{3M}{8}\frac{s^2}{M(s^2-2s+4)+3(M-2)}\frac{P_{\text{FB}}}{P+5\sigma^2+\sigma\sqrt{P}} \\
&= \frac{P_{\text{FB}}}{\sigma^2_{\text{FB}}}\frac{3M}{8}\frac{s^2}{M(s^2-2s+4)+3(M-2)}\underbrace{\frac{P}{P+5\sigma^2+\sigma\sqrt{P}}}_{\leq 1} \\
&\leq \frac{P_{\text{FB}}}{\sigma^2_{\text{FB}}}\frac{3M}{8}\frac{s^2}{M(s^2-2s+4)+3(M-2)} \\
&\overset{(a)}{\leq} \frac{P_{\text{FB}}}{\sigma^2_{\text{FB}}}\frac{3M}{8}\frac{1}{6(M-1)} \\
&= \frac{P_{\text{FB}}}{\sigma^2_{\text{FB}}}\frac{1}{4}\frac{M}{4(M-1)} \qquad (17)\\
&\overset{(b)}{\leq} \frac{P_{\text{FB}}}{\sigma^2_{\text{FB}}}\frac{1}{4}\frac{M(M-1)}{M(M-1)-2} \qquad (18)
\end{aligned}$$

where (a) follows as $\frac{s^2}{M(s^2-2s+4)+3(M-2)}$ is monotonically increasing in $s \in [0,1]$, for any M and is maximized with $s = 1$; (b) follows as the term in parenthesis of (17): $\frac{M}{4(M-1)} < 1$, whereas that in (18): $\frac{M(M-1)}{M(M-1)-2} > 1$.

Intuitively, the active coding scheme renders the feedback transmission resilient to noise. In the passive coding scheme, the transmitter obtains a noisy version of the received sequence $y^{\lambda(n-1)}$ and estimates the most likely codeword pair determined by the receiver. In the active feedback scheme, this pair is determined at the receiver directly, encoded using a simplex code and sent over the noisy feedback link to the transmitter. Section IV presents the proof of this result for $M = 3$, and the generalization for any finite $M \geq 3$ is presented in Appendix B.

Next we present an achievable error exponent resulting from scaling the feedback direction of the linear, passive coding scheme of [18, Theorem 2] to meet an AS power constraint in the feedback direction:

Theorem 3. *Under the AS power constraint in both directions of a one-way AWGN channel with linear passive noisy feedback and the transmission of finite $M \geq 2$ messages, an achievable error exponent is determined by* (15)*, where α is the scaling factor that guarantees the AS constraint is met, and determined by the solution of:*

$$\frac{\left(\frac{2L\left(\sigma_{FB}/\alpha\right)}{2L\left(\sigma_{FB}/\alpha\right)+\sqrt{\sigma^2+\frac{\sigma^2_{FB}}{\alpha^2}}}\right)P+}{2\sigma\sqrt{\left(\frac{2L\left(\sigma_{FB}/\alpha\right)}{2L\left(\sigma_{FB}/\alpha\right)+\sqrt{\sigma^2+\frac{\sigma^2_{FB}}{\alpha^2}}}\right)P}} = \frac{P_{FB}}{\alpha^2}, \tag{19}$$

given $P, P_{FB}, \sigma^2, \sigma^2_{FB}$ and $L = \lfloor M/2 \rfloor$.

This result is a direct extension of [18, Theorem 2] which is recovered by setting $\alpha = 1$. We present its derivation in Appendix C.

Remark 5. *This coding scheme leads to an error exponent that is higher than those attained in Theorems 1 and 2 when the noise in the feedback link is very low, as was reported in [18, Theorem 2]. This continues to be true even the feedback input channels are also subject to an AS power constraint. Furthermore, as $\sigma^2_{FB} \to 0$ it recovers Pinsker's result in (7), the error exponent under noiseless feedback.*

Next, we present a generalization of the Kim-Lapidoth-Weissman [19] achievable error exponent for $M = 2$ under the EXP constraint for both directions to any arbitrary but finite number $M \geq 2$:

Theorem 4. *Under the EXP power constraint in both directions of a one-way AWGN channel with active noisy feedback* *and the transmission of finite $M \geq 2$ messages, an achievable error exponent is:*

$$E_{FB^{EXP}}^{EXP}\left(M, P, \sigma^2, P_{FB}, \sigma_{FB}^2\right) \geq \frac{M}{M-1}\left(\frac{P}{\sigma^2}+\frac{P_{FB}}{\sigma_{FB}^2}\right). \quad (20)$$

Observe that (20) yields the result of [19, Equation (15)] when $M = 2$, shown in (11). Hence, the achievable error exponent under the EXP constraint and noisy feedback at least quadruples the one attained under AS in the absence of feedback, provided in (6), not only for $M = 2$ (first shown in [19]) but for any finite number of messages being sent. The error exponent of (20) is achieved by a three-stages coding scheme based on a building block (BB) which is used to send message W using $\lambda(n-M)$ channel uses. This is decoded as W' and feed back using the same building block in $(1-\lambda)(n-M)$ channel uses and decoded as W''. Finally, the transmitter compares the true W and W'' and generates a length M retransmission codeword used to correct transmission errors.

B. Error exponents for the Two-Way AWGN channel

We derive achievable error exponent regions for the two-way AWGN channel, both for interactive and non-interactive terminals.

1) Non-interactive terminals: When terminals ignore (for the purpose of generating their channel inputs) the received outputs, the following error exponent region is achievable:

Proposition 1. *An achievable EER for non-interactive terminals in a two-way AWGN channel and the transmission of a finite number of messages M is formed by the union over all simultaneously achieved error exponents pairs under the AS and EXP power constraints and given by:*

$$E_{12} \geq \frac{P_1}{\sigma_2^2}\frac{M}{4(M-1)}, \quad (21)$$

$$E_{21} \geq \frac{P_2}{\sigma_1^2}\frac{M}{4(M-1)}. \quad (22)$$

Equations (21) and (22) follow directly in the form of (6) from [4], and are based on a simplex code. Each terminal focuses on its own transmission only without allocating any resources to help the other direction.

2) Interactive terminals: Two-way error exponents may benefit from interaction in certain scenarios, but not others. For example, under the AS constraint, interaction cannot improve the two-way error exponents for $M = 2$ since the error exponent attained without feedback matches that with perfect feedback, and is given by Pinsker's bound in (7). Pinsker's upper bound suggests that for $M > 2$, error exponent gains over non-feedback are possible in both directions when interaction or feedback is employed.

Achievable EER under the AS power constraint: We study the transmission of $M \geq 3$ messages over an AWGN two-way channel with $\frac{P_1}{\sigma_2^2} < \frac{P_2}{\sigma_1^2}$. As in the one-way channel with noisy feedback under AS constraints, [18], [36], [37], we have not been able to obtain gains over Proposition 1 for symmetric SNRs. Assuming the $1 \to 2$ direction is noisier than the $1 \leftarrow 2$ direction, we consider a two-way coding scheme in which the stronger $1 \leftarrow 2$ link is used during a fraction of the block length n to transmit W_2 without feedback, and for another fraction of time, to improve the error exponent of the weaker $1 \to 2$ direction using either passive or active feedback for the transmission of W_1. This results in the following two theorems:

Theorem 5. *For the transmission of $M \geq 3$ messages and passive feedback, any error exponent pair $(E_{12}^{AS}, E_{21}^{AS})$ satisfying* (23) *and* (24) *is achievable, for some $\lambda \in [0,1]$, and $s \in (0,1)$ and $\frac{P_1}{\sigma_2^2} < \frac{P_2}{\sigma_1^2}$.*

Equation (23) follows from (6) and the use of a non-feedback, simplex-code-based transmission for message W_2 in the first λn channel uses. Equation (24) follows from direct application of the non-linear noisy feedback coding scheme in Theorem 1, used to transmit message W_1 in the remaining $(1-\lambda)n$ channel uses. The factor α^2 is used to ensure the AS power constraint in the $1 \leftarrow 2$ direction is satisfied, and is derived in a similar way as in the one-way channel presented in Theorem 1, shown in Appendix A. The power available at the receiver to provide passive feedback is $(1-\lambda)P_2$, since λP_2 is allocated for the transmission of the messages in the $1 \leftarrow 2$ direction. Equivalently to Theorem 1, (24) is maximized for the choice of $s \in [0,1]$ of $s = \frac{1-\sqrt{3\frac{\sigma_2^2}{\sigma_1^2}\frac{P_2(1-\lambda)}{P_1+5\sigma_2^2+\sigma_2\sqrt{P_1}}-3}}{1-\frac{3}{4}\frac{\sigma_2^2}{\sigma_1^2}\frac{P_2(1-\lambda)}{P_1+5\sigma_2^2+\sigma_2\sqrt{P_1}}}$, and the equivalent to Remark 1 applies. Next, active feedback is considered.

Theorem 6. *For the transmission of $M \geq 3$ messages and active feedback, any error exponent pair $(E_{12}^{AS}, E_{21}^{AS})$ satisfying the following is achievable, for some $\lambda \in [0,1]$, and $\frac{P_1}{\sigma_2^2} < \frac{P_2}{\sigma_1^2}$:*

$$E_{21}^{AS} \geq \frac{M}{4(M-1)}\lambda\frac{P_2}{\sigma_1^2} \quad (27)$$

$$E_{12}^{AS} \geq \min\left\{\frac{P_1}{\sigma_2^2}\frac{2M}{7M-6}, \underbrace{\frac{P_2}{\sigma_1^2}\frac{(1-\lambda)M(M-1)}{4\left(M(M-1)-2\right)}}_{\textit{active feedback}}\right\}. \quad (28)$$

As in Theorem 5, (27) follows from using (6) for λn channel uses and a simplex code of M symbols for the transmission of W_2 in the $1 \leftarrow 2$ direction. Equation (28) results from the use of encoded feedback for the transmission of message W_1 in the $1 \to 2$ direction, using the scheme for the one-way AWGN channel presented in Section III-A, Theorem 2, and the error exponent given in (16).

Figure 3 shows block diagrams of the schemes used to obtain Theorems 5 and 6. Message W_2 is sent over the stronger link without feedback, and W_1 is transmitted in the opposite direction respectively with (a) passive or (b) active feedback. In (a), the weaker channel remains idle while W_2 is being sent during λn channel uses. Terminal 1 initiates its own transmission (helped by Terminal 2) once the $1 \leftarrow 2$ message transmission concludes. In the remaining $(1-\lambda)n$ channel uses, W_1 is transmitted employing the non-linear passive feedback scheme presented in Theorem 1. Since Terminal 2 used part of its available power P_2 to send W_2, only the

$$E_{21}^{\mathrm{AS}} \geq \frac{M}{4(M-1)}\lambda\frac{P_2}{\sigma_1^2}, \tag{23}$$

$$E_{12}^{\mathrm{AS}} \geq \min\left\{ M\frac{P_1}{2\sigma_2^2}\frac{s^2-2s+4}{M(s^2-2s+4)+3(M-2)}, \underbrace{\frac{P_1}{\sigma_1^2}\frac{3M}{8}\frac{s^2}{M(s^2-2s+4)+3(M-2)}\overbrace{\frac{P_2(1-\lambda)}{P_1+5\sigma_2^2+\sigma_2\sqrt{P_1}}}^{\alpha^2}}_{\text{passive feedback}} \right\}. \tag{24}$$

$$E_{21}^{AS} \geq \frac{M}{4(M-1)}\lambda\frac{P_2}{\sigma_1^2} \tag{25}$$

$$E_{21}^{\mathrm{AS}} \geq \frac{P_1}{2}\frac{1}{\sigma_2^2+\frac{\sigma_1^2}{\alpha^2}+4\left(\lfloor M/2\rfloor\right)^2\sigma_1^2/\alpha^2+4\left(\lfloor M/2\rfloor\right)\sqrt{\frac{\sigma_1^2}{\alpha^2}\left(\sigma_2^2+\frac{\sigma_1^2}{\alpha^2}\right)}}, \tag{26}$$

remainder can be used to serve the other direction through feedback. Note that during passive feedback, the i-th output received at Terminal 2 is immediately fed back to the Terminal 1 without delay and both directions are busy at the same time. In (b), the transmission of W_2, occurs in $\lambda(n-1)$ channel uses, simultaneously to the first stage of the active feedback assisted transmission of W_1 which occurs over the weaker channel. This is possible since both directions are independent, and because in our scheme the active feedback stage starts once the first stage transmission of W_1 is concluded.

Apparently, error exponent gains over feedback-free transmission are possible only for the weaker direction. This is not surprising given the results for the one-way channel with noisy feedback under an AS power constraint. The linear passive feedback scheme of Theorem 3 may be used for two-way communication as well, leading to the following theorem:

Theorem 7. *For the transmission of M messages and passive linear feedback, any error exponent pair $(E_{12}^{AS}, E_{21}^{AS})$ satisfying* (25) *and* (26) *is achievable, for some $\lambda \in [0,1]$ and $\frac{P_1}{\sigma_2^2} < \frac{P_2}{\sigma_1^2}$, where α is the scaling factor that guarantees the AS constraint is met, and determined by the solution of:*

$$\left(\frac{2L\left(\sigma_1/\alpha\right)}{2L\left(\sigma_1/\alpha\right)+\sqrt{\sigma_2^2+\frac{\sigma_1^2}{\alpha^2}}}\right)P+ 2\sigma_2\sqrt{\left(\frac{2L\left(\sigma_1/\alpha\right)}{2L\left(\sigma_1/\alpha\right)+\sqrt{\sigma_2^2+\frac{\sigma_1^2}{\alpha^2}}}\right)P}=\frac{P_2(1-\lambda)}{\alpha^2}, \tag{29}$$

given $P_1, P_2, \sigma_1^2, \sigma_2^2$ and $L = \lfloor M/2 \rfloor$.

This theorem results from a time-sharing argument, in which during λn channel uses, message W_2 is sent in the stronger $1 \leftarrow 2$ without feedback, thus (25) results from the feedback-free error exponent derived by Shannon [4] and shown in (6). In the remaining $(1-\lambda)n$ channel uses, the linear coding scheme with noisy feedback from [18, Theorem 2] and modified in Theorem 3 is utilized to transmit message W_1 in the weaker $1 \to 2$ direction, taking into account that terminal 2 has already used part of its available energy for the transmission of W_2. This results in (26) and the associated expression for the choice of α in (29). This scheme leads to error exponent gains over feedback-free transmissions in the weaker direction at the expense of a decrease of the error exponent of the stronger direction. Such gains are possible when the $1 \leftarrow 2$ direction operates in the SNR regime for which the linear passive feedback scheme outperforms the feedback-free error exponent, i.e. see Figure 8 (right).

Next, we consider the EXP power constraint, which permits very rarely occurring high amplitude transmissions. We show how this feature is used in the two-way AWGN channel as well, which holds for all relative SNR conditions.

Achievable error exponents region under the EXP power constraint: We propose a coding scheme based on the one-way building block introduced by Kim-Lapidoth-Weissman [19], which we modified to operate for a general number of messages (Theorem 4), and further customized here to support two-way communications. Our scheme involves three stages in which each terminal first sends their message, then, feeds back an estimate of the previously received message, and finally (if necessary) triggers a high amplitude retransmission to correct erroneous decoding decisions. This scheme results in the EER described by the following theorem:

Theorem 8. *For the transmission of $M \geq 2$ messages and **active feedback**, any error exponent pair $(E_{12}^{EXP}, E_{21}^{EXP})$ satisfying the following is achievable, for some $\lambda \in [0,1]$*

$$E_{12}^{EXP} \geq \frac{M}{M-1}\left(\lambda K_1\frac{P_1}{\sigma_2^2}+(1-\lambda)J_2\frac{P_2}{\sigma_1^2}\right) \tag{30}$$

$$E_{21}^{EXP} \geq \frac{M}{M-1}\left(\lambda K_2\frac{P_2}{\sigma_1^2}+(1-\lambda)J_1\frac{P_1}{\sigma_2^2}\right), \tag{31}$$

where $K_1, K_2 \in [0, \frac{1}{\lambda}]$ and $J_1, J_2 \in [0, \frac{1}{1-\lambda}]$ are power allocation parameters such that

$$\lambda K_i + (1-\lambda)J_i \leq 1, \quad \textit{for } i = 1, 2. \tag{32}$$

An achievable error exponent sum-rate is given by adding Equations (30) and (31).

$$E_{12}^{\mathrm{EXP}} + E_{21}^{\mathrm{EXP}} \geq \frac{M}{M-1}\left(\lambda K_1\frac{P_1}{\sigma_2^2}+(1-\lambda)J_2\frac{P_2}{\sigma_1^2}\right)+$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory

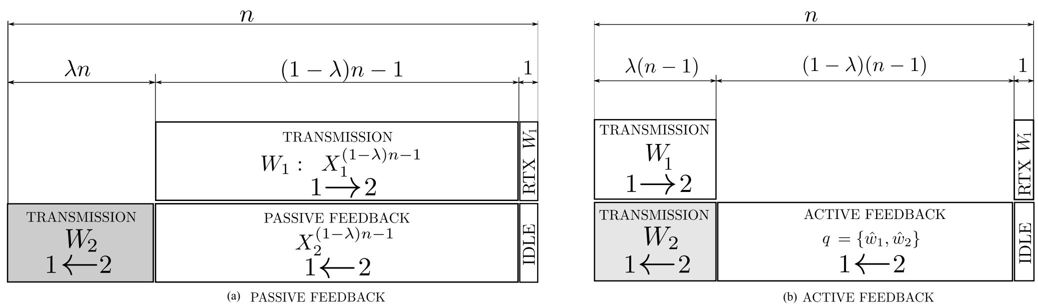


Fig. 3. Two-Way AWGN channel coding schemes block diagrams under AS power constraint, where the $1 \leftarrow 2$ direction is stronger (has higher SNR) than the $1 \rightarrow 2$ direction.

$$
\begin{aligned}
&\frac{M}{M-1}\left(\lambda K_2 \frac{P_2}{\sigma_1^2} + (1-\lambda) J_1 \frac{P_1}{\sigma_2^2}\right) \\
&= \frac{M}{M-1}\left(\frac{P_1}{\sigma_2^2} \underbrace{(\lambda K_1 + (1-\lambda) J_1)}_{=1 \text{ by } (32)} + \right. \\
&\qquad \left. \frac{P_2}{\sigma_1^2} \underbrace{(\lambda K_2 + (1-\lambda) J_2)}_{=1 \text{ by } (32)} \right) \\
&= \frac{M}{M-1}\left(\frac{P_1}{\sigma_2^2} + \frac{P_2}{\sigma_1^2}\right),
\end{aligned}
$$

where the last equality follows by choosing the parameters such that (32) is satisfied with equality. Figure 4 shows this EER. Theorem 8 results from the simultaneous use of the three-stage coding scheme presented for the one-way channel for both directions. Thus the BB is used for $\lambda(n-M)$ channel uses to transmit messages W_1 and W_2, which are respectively decoded as W_1' and W_2'. Then, the BB is used during $(1-\lambda)(n-M)$ channel uses to transmit these estimates back to the terminal where each message originated. These are respectively decoded as W_1'' and W_2''. The last stage involves M channel uses in each direction based on the index location code similar to the one-way scheme. Thus, each direction operates for a fraction of time aided by the other terminal. Below, the axis-crossing points result from dedicating the whole block length n and all power of one terminal to provide feedback for the transmission of the other direction only.

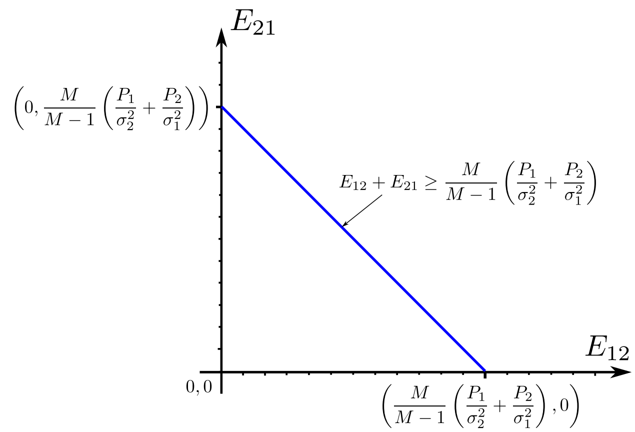


Fig. 4. Achievable error exponent region for the transmission of M messages under the EXP power constraint.

C. A discussion on outer bounds

Outer bounds on the error exponent region for the two-way AWGN channel may be derived from existing results for the one-way AWGN channel. Under the EXP power constraint and $M = 2$, outer bounds on the error exponent region for the two-way AWGN channel follow directly from using the one-way channel outer bound under noisy feedback presented in [19] for each direction:

Proposition 2. *Any achievable error exponent pair $(E_{12}^{EXP}, E_{21}^{EXP})$ under the EXP power constraint for $M = 2$ must satisfy*

$$
E_{12}^{EXP}\left(M=2, P_1, \sigma_2^2, P_2, \sigma_1^2\right) \leq \frac{\left(\sqrt{P_1 + \sigma_2^2} + \sqrt{P_1}\right)^2}{\sigma_2^2} + \frac{\left(\sqrt{P_2 + \sigma_1^2} + \sqrt{P_2}\right)^2}{\sigma_1^2}
$$

$$
E_{21}^{EXP}\left(M=2, P_1, \sigma_2^2, P_2, \sigma_1^2\right) \leq \frac{\left(\sqrt{P_2 + \sigma_1^2} + \sqrt{P_2}\right)^2}{\sigma_1^2} + \frac{\left(\sqrt{P_1 + \sigma_2^2} + \sqrt{P_1}\right)^2}{\sigma_2^2}.
$$

An upper bound for the two-way AWGN channel error exponent region under the AS power constraint is given directly by providing perfect noiseless output feedback to both transmitters (like two one-way channels with perfect feedback, not using those perfect links to convey messages). Then, from [6]:

Proposition 3. *Any achievable error exponent pair $(E_{12}^{AS}, E_{21}^{AS})$ under the AS power constraint for any M must satisfy*

$$
\begin{aligned}
E_{12}^{AS}\left(M, \frac{P_1}{\sigma_2^2}, \frac{P_2}{\sigma_1^2}\right) &\leq \frac{P_1}{2\sigma_2^2}, \\
E_{21}^{AS}\left(M, \frac{P_1}{\sigma_2^2}, \frac{P_2}{\sigma_1^2}\right) &\leq \frac{P_2}{2\sigma_1^2}.
\end{aligned}
$$

This concludes the statement of our main results. The remainder of the paper consists of the proofs and numerical evaluations of these regions, so that their performance can be visually compared.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory

IV. Proof of Theorem 2: An active noisy feedback coding scheme for the One-Way AWGN channel for three messages under the AS power constraint

Equation (14), follows from altering the non-linear passive feedback technique from Xiang-Kim [18, Section II] for $M \geq 3$, to allow active feedback. We present the proof next in some detail for $M = 3$, and generalize this to arbitrary but finite $M \geq 3$ in Appendix B. For the transmission of three messages from $\Omega_3 = \mathcal{W} = \{1, 2, 3\}$, we use the simplex code defined in (3), which results in $\mathcal{C}(\Omega_3, \mathcal{E})$ and length n codewords $X^n(\omega)$:

$$X^n(\omega) = \begin{cases} \sqrt{\mathcal{E}} \cdot (0, 1, 0, ..., 0), & \text{if } \omega = 1 \\ \sqrt{\mathcal{E}} \cdot \left(\frac{-\sqrt{3}}{2}, -\frac{1}{2}, 0, ..., 0\right), & \text{if } \omega = 2 \\ \sqrt{\mathcal{E}} \cdot \underbrace{\left(\frac{+\sqrt{3}}{2}, -\frac{1}{2}, 0, ..., 0\right)}_{n \text{ channel uses}}, & \text{if } \omega = 3 \end{cases}. \quad (33)$$

Consider transmitting a message W uniformly chosen from $\mathcal{W}$ over the channel of Figure 1 assuming the feedback link is strictly better than the forward, $\frac{P_{\text{FB}}}{\sigma^2_{\text{FB}}} > \frac{P}{\sigma^2}$. The block diagram of the non-linear noisy feedback scheme from Xiang-Kim using passive feedback (thus $U_k = Y_k$) is shown in Figure 5 (a), whereas (b) shows the scheme we propose for active feedback, where the transmission of W occurs in n channel uses through three non-overlapping stages: transmission, active feedback, and retransmission (RTX). Thus, a stage starts only once the previous one is complete. Moreover, we have sketched consecutive transmissions of messages to show how the transmitter initiates the transmission of the next message while it is receiving the feedback message from the previous one.

1. Transmission: This first stage lasts for $\lambda(n-1)$ channel uses, where message $W = w$ is transmitted without feedback as codeword $X^{\lambda(n-1)}(w)$ from (33), using the simplex code $\mathcal{C}(\mathcal{W}, \beta nP)$. At the receiver, sequence $y^{\lambda(n-1)}$ is decoded using protection regions B_w (as in [18]), one for each transmitted codeword as shown in Figure 6 and defined in (34). Sequences received inside region B_w are immediately decoded as the codeword w (and the other stages are ignored). These regions are parameterized by $s \in [0, 1]$ following [18, Equation (6)] for a parameter $t \in [0, \frac{\sqrt{3}-1}{2}]$ which is geometrically coupled to s by the relation: $\left|\frac{1+s/2}{\sqrt{3}} - \frac{\sqrt{s^2-2s+4}}{2\sqrt{3}}\right| = t$. This guarantees that distance $d_5 = \min_{y^{\lambda(n-1)} \in A'_{23} \cup B_2 \cup B_3} ||x^{\lambda(n-1)}(1) - y^{\lambda(n-1)}||$.

$$\begin{aligned} B_w = \Big\{ y^{\lambda(n-1)} : & ||x^{\lambda(n-1)}(w) - y^{\lambda(n-1)}|| \\ & \leq ||x^{\lambda(n-1)}(w') - y^{\lambda(n-1)}|| \text{ for } w' \neq w, \\ \text{and } & \left| ||x^{\lambda(n-1)}(w') - y^{\lambda(n-1)}|| - ||x^{\lambda(n-1)}(w'') - y^{\lambda(n-1)}|| \right| \\ & \leq td', \text{ for } w' \neq w, \ w'' \neq w \Big\}. \end{aligned} \quad (34)$$

Figure 6 also illustrates regions $A'_{12}, A'_{13}, A'_{23}$, that represent regions in which the receiver is undecided between two codewords. The definition of these regions follows as in [18, Section II-B]:

$$A'_{ww'} = Aww' \setminus \left(\bigcup_{w''} B_{w''} \right), \quad w, w' \in [1:3] \quad (35)$$

which can be generalized to any M by taking $w, w' \in [1 : M]$, and where:

$$\begin{aligned} A_{ww'} = \Big\{ & y^{\lambda(n-1)} : \\ & \max\left\{ ||y^{\lambda(n-1)} - x^{\lambda(n-1)}(w)||, ||y^{\lambda(n-1)} - x^{\lambda(n-1)}(w')|| \right\} \\ & \leq ||y^{\lambda(n-1)} - x^{\lambda(n-1)}(w'')||, w'' \neq w, w'' \neq w' \Big\} \end{aligned}$$

From now on, without loss of generality (WLOG) assume $W = 1$ is sent. The probability of error of this stage corresponds to the occurrence of event $\mathcal{E}_{\text{T}}$, defined in (36) which denotes sequence $y^{\lambda(n-1)}$ being received in the wrong protection region $B_2 \cup B_3$, or in the ambiguous region excluding $W = 1$, i.e. $y^{\lambda(n-1)} \in A'_{23}$:

$$\mathcal{E}_{\text{T}} = \{ y^{\lambda(n-1)} \in B_2 \cup B_3 \cup A'_{23} \}. \quad (36)$$

The achievable error exponent of this stage is upper bounded by [18, Equation (7)] as[3] $\mathsf{P}(\mathcal{E}_{\text{T}}) \leq 2Q\left(\frac{d_5}{\sigma}\right) \leq \exp\left(-n\frac{\beta P}{8\sigma^2}(s^2 - 2s + 4) \right)$ (see Figure 6), and is hence given in (37) as

$$\lim_{n \to \infty} -\frac{1}{n} \log\left(\mathsf{P}(\mathcal{E}_{\text{T}})\right) := E_{\mathcal{E}_{\text{T}}} \geq \frac{\beta P}{8\sigma^2}(s^2 - 2s + 4). \quad (37)$$

2. Active feedback: When $y^{\lambda(n-1)}$ is not in a protection region, the receiver determines which two codewords form the most likely codeword pair $q = \{\hat{w}_1, \hat{w}_2\}$, given by the two closest codewords to the received sequence as:

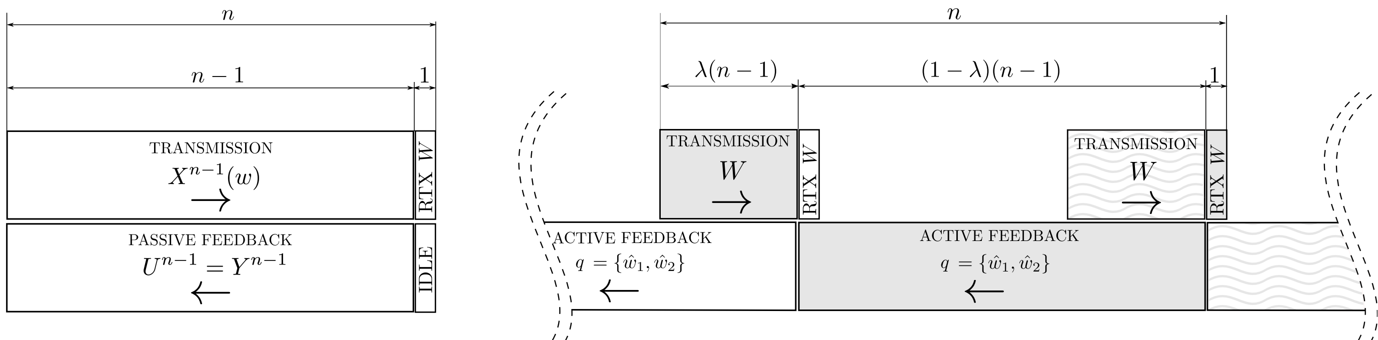
$$q = \begin{cases} \{1, 2\}, & \text{if } y^{\lambda(n-1)} \in A'_{12} \\ \{2, 3\}, & \text{if } y^{\lambda(n-1)} \in A'_{23} \\ \{1, 3\}, & \text{if } y^{\lambda(n-1)} \in A'_{13} \end{cases}.$$

Let $\mathcal{Q}$ be the set of $q = \{\hat{w}_1, \hat{w}_2\}$ values defined above. The most likely unordered pair (labeled lexicographically) is encoded and sent to the transmitter as $U^{(1-\lambda)(n-1)}(q)$ using the simplex code $\mathcal{C}(\mathcal{Q}, nP_{\text{FB}})$ over the feedback link in $(1-\lambda)(n-1)$ channel uses (active feedback). The transmitter estimates the sent message q as $\hat{q} = \{\tilde{w}_1, \tilde{w}_2\}$. An error occurs if q, is incorrectly decoded at the transmitter $\hat{q}$. We denote this event as $\mathcal{E}_{\text{AFB}} = \{\hat{q} \neq q\}$, and its probability of occurrence as $\mathsf{P}(\mathcal{E}_{\text{AFB}})$.

Note that even if $\hat{q}$ contains the true message, when $\hat{q}$ does not match q exactly, this is counted as an error. We do this since in the last stage, whether the true message is the first or second element in the pair is encoded in the polarity of an antipodal signaling. As such, while in some cases an erroneous $\hat{q}$ could still lead to the correct re-transmission, we choose to simply count them all as errors and require $\hat{q} = q$. the achievable error exponent for this stage is given by (38), which results from (6) since we use a simplex code [4] of three data

[3]Where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{y^2}{2}\right) dy$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory



(a) PASSIVE FEEDBACK (b) ACTIVE FEEDBACK

Fig. 5. Block diagrams for the non-linear noisy coding schemes under the AS power constraint.

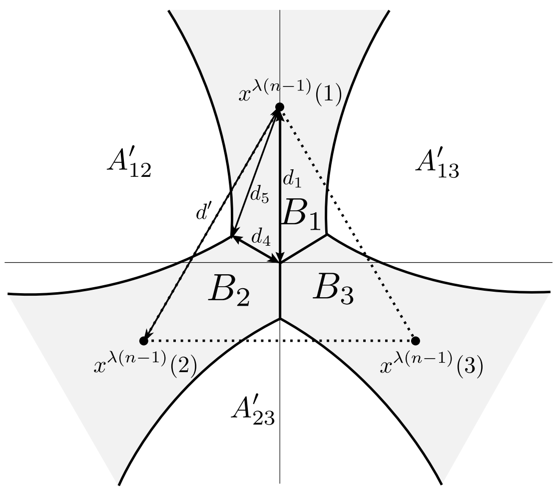


Fig. 6. Protection regions decoding introduced by Xiang-Kim [18] and used for decoding message W for the transmission stage. Here, $d_1 = \sqrt{\beta n P}$, $d_4 = \frac{s}{2} d_1$, $d' = \sqrt{3\beta n P}$, and $d_5 = \sqrt{\frac{\beta n P}{4}(s^2 - 2s + 4)}$.

symbols.

$$\lim_{n\to\infty} -\frac{1}{n}\log\left(\mathsf{P}(\mathcal{E}_{\text{AFB}})\right) := E_{\mathcal{E}_{\text{AFB}}} \geq \frac{3}{8}\frac{P_{\text{FB}}}{\sigma^2_{\text{FB}}}. \tag{38}$$

Note that since only $(1-\lambda)(n-1)$ channel uses are available for this stage, the receiver may scale up its transmission by $\frac{1}{1-\lambda}$. Note as well that during the active feedback stage, the transmitter remains idle for a portion of the channel uses, as the simplex code for M data points concentrates the energy in $M-1$ dimensions.

3. Retransmission: The last stage occurs in the last (single) channel use of the block of length n. At this point, the transmitter and receiver have estimates of the two most likely codewords, $q = \{\hat{w}_1, \hat{w}_2\}$, and $\hat{q} = \{\tilde{w}_1, \tilde{w}_2\}$, which are the same otherwise this is counted as an error in the feedback stage. Then, based on $\hat{q}$, the transmitter uses antipodal signaling to send the true codeword among these two candidates using its remaining power:

$$X_n = \begin{cases} +\sqrt{(1-\beta)nP}, & \text{if } \; w = \min\{\tilde{w}_1, \tilde{w}_2\} \\ -\sqrt{(1-\beta)nP}, & \text{if } \; w = \max\{\tilde{w}_1, \tilde{w}_2\} \\ 0, & \text{otherwise.} \end{cases} \tag{39}$$

Note in the above, that parameter β is used to allocate the power available at the transmitter between the Transmission and Retransmission stages, whereas λ is used to characterize the duration of the transmission and the active feedback stages. Parameter λ requires attention as M becomes large, since the number of channel uses required in the transmission and feedback stages is respectively determined by the number of dimensions imposed by the simplex code used; for example in the transmission stage, $M-1$ dimensions are required, whereas in the active feedback stage, this number is determined by the number of codeword pairs possible out of the M messages that could be transmitted. We address this in Section VI. At the end of this stage, the message decoding rule follows that of [18], in which $\hat{W} = \hat{w}$:

$$\hat{w} = \arg\min_{w\in\{\hat{w}_1,\hat{w}_2\}} ||x^{\lambda(n-1)}(w) - y^{\lambda(n-1)}||^2 + ||x_n(w) - y_n||^2.$$

The decoder errs if given that $y^{\lambda(n-1)}$ is outside the protection regions and the feedback stage led to $q = \hat{q}$, the following event occurs:

$$\mathcal{E}_{\text{RT}} = \left\{ (W \in q = \hat{q}) \cap \left(\hat{W} \neq W\right) \mid W = 1 \right\}.$$

The error exponent of this stage is derived in a similar way as in [18, $\mathsf{P}(\mathcal{E}_2)$ in Section II-A], thus we have that:

$$\mathsf{P}(\mathcal{E}_{\text{RT}}) = Q\left(\sqrt{\left(1 - \frac{\beta}{4}\right)\frac{P}{\sigma^2}n}\right). \tag{40}$$

Equation (40) results as for this stage, the event $\hat{q} = q$ renders the subsequent analysis equivalent to the noiseless passive feedback case analyzed in [18], which thus derives the error exponent by considering the distance between the codewords of the simplex code of the transmission stage (given by $d' = \sqrt{3\beta n P}$), and the antipodal signaling of the retransmission stage (given by $d_a = 2\sqrt{(1-\beta)nP}$). These are used to characterize the distance between two codewords as $\sqrt{d'^2 + d_a^2} = \sqrt{3\beta nP + 4(1-\beta)nP}$ and utilized to obtain (40), which leads to the error exponent:

$$\lim_{n\to\infty} -\frac{1}{n}\log\left(\mathsf{P}(\mathcal{E}_{\text{RT}})\right) := E_{\mathcal{E}_{\text{RT}}} \geq \left(1 - \frac{\beta}{4}\right)\frac{P}{2\sigma^2}. \tag{41}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory

To obtain the overall error probability upper bound, note that by symmetry and WLOG $W = 1$ is sent. Denote this conditional probability of error as $\mathsf{P}(\text{error} \mid W = 1) = \mathsf{P}_1(W \neq \hat{W})$. Note that $\mathsf{P}_1(W \neq \hat{W}) = \mathsf{P}_2(W \neq \hat{W}) = \mathsf{P}_3(W \neq \hat{W})$, and hence, since the messages are uniformly distributed $\mathsf{P}(\text{error}) = \frac{1}{3}\sum_{i=1}^{3} \mathsf{P}_i(W \neq \hat{W}) = \mathsf{P}_1(W \neq \hat{W})$. Then, we have

$$\mathsf{P}(\text{error}) = P_1(\hat{W} \neq W) \leq \mathsf{P}(\mathcal{E}_{\text{T}}) + \mathsf{P}\left(\mathcal{E}_{\text{AFB}}\right) + \mathsf{P}\left(\mathcal{E}_{\text{RT}}\right)$$

From (37), (38) and (41), the overall error exponent is thus given by the minimum of the three stages:

$$E_{12}\left(M = 3, \frac{P}{\sigma^2}, \frac{P_{\text{FB}}}{\sigma_{\text{FB}}^2}, s\right) \geq \min\left\{\frac{\beta P}{8\sigma^2}(s^2 - 2s + 4), \frac{3}{8}\frac{P_{\text{FB}}}{\sigma_{\text{FB}}^2}, \left(1 - \frac{\beta}{4}\right)\frac{P}{2\sigma^2}\right\}. \quad (42)$$

Equating the first and third terms, we obtain $\beta = \frac{4}{s^2 - 2s + 5}$. This reduces the number of arguments of (42) to two, which leads to (14) evaluated at $M = 3$. The generalization to $M \geq 3$ is presented in Appendix B. □

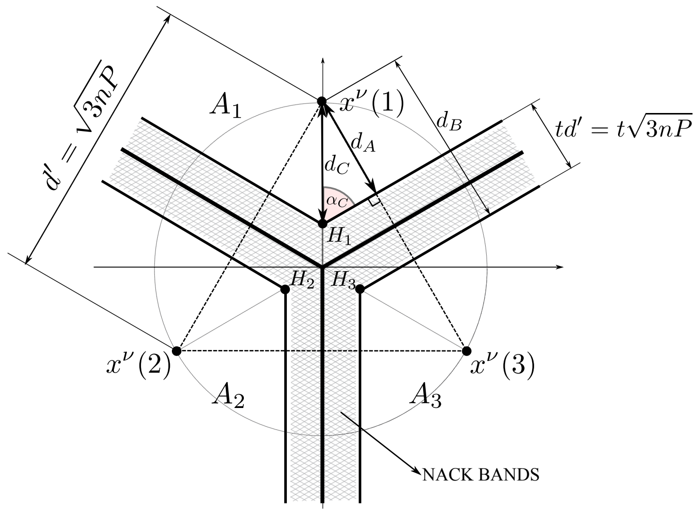


Fig. 7. Decoding regions based on "NACK-Bands". The receiver performs immediate decoding for sequences received in the A_i regions, otherwise, a retransmission is necessary. Note that Regions A_i shown here are not related to regions $A_{ww'}$ shown in Figure 6.

V. Proof of Theorem 4: An active noisy feedback coding scheme for the One-Way AWGN channel for $M \geq 3$ messages under the EXP power constraint

Theorem 4 results from a generalization of the Kim-Lapidoth-Weissman coding scheme [19] (for the transmission of $M = 2$ messages over the channel of Figure 4 under the EXP power constraint) to support any finite $M \geq 3$. We next illustrate results for $M = 3$ using a geometric approach, which is easily extended to larger finite M.

A. A communication building block for $M = 3$ messages

Consider the AWGN channel with active feedback of Figure 1 and described by (4) and (5), where both channel inputs are subject to EXP power constraints: $\mathsf{E}\left[\sum_{k=1}^{n} X_k^2\right] \leq nP$ and $\mathsf{E}\left[\sum_{k=1}^{n} U_k^2\right] \leq nP_{\text{FB}}$.

1) Building block operation: The scheme for $M = 2$ [19] is based on the usage of a "building block" (BB), which is a three-stage (transmission, active feedback and retransmission) scheme achieving an error exponent of $\frac{2P}{\sigma^2}$ if used in the forward direction ($\frac{2P_{\text{FB}}}{\sigma_{\text{FB}}^2}$ if used in the feedback direction). We will show that our modification for $M = 3$ results in a BB of error exponent $\frac{3P}{2\sigma^2}$.

In the first stage, W, taken uniformly from $\mathcal{W} = \{1, 2, 3\}$, is sent in $\nu := n - 4$ channel uses, using the simplex code $\mathcal{C}\left(\Omega_3 = \mathcal{W}, nP\right)$, while the receiver remains silent. Figure 7 shows the constellation resulting from encoded messages $x^\nu(W)$ and their corresponding decoding regions A_i for $i \in \{1, 2, 3\}$. Once the first stage is complete, the receiver determines whether the received sequence y^ν is valid (ACK, meaning $y^\nu \in \bigcup_{i=1}^{3} A_i$), or non-valid (NACK, meaning $y^\nu \notin \bigcup_{i=1}^{3} A_i$). The latter occurs for signals received in the gray "NACK bands" shown below. The width of these bands is proportional to the distance between any two codewords and parameterized by $t \in (0, 1)$. Each ACK region A_i corresponds to a space determined by $M - 1$ hyperplanes defined in $M - 1$ dimensions (in Figure 7, this corresponds to two lines). Specifically, an ACK region is bounded by the hyperplanes separating codeword $x^\nu(i)$ from each of the other $M - 1 = 2$ codewords $x^\nu(j)$, for $j \neq i \in \{1, 2, 3\}$. Note that each codeword is in ν dimensions, but only the first 2 are non-zero for $M = 3$, and hence we visually represent these in 2-D. Each separating hyperplane (line) is perpendicular to the line connecting codewords i, j, represented by the vector $\mathbf{w}_{ij} \in \mathbb{R}^\nu$ and given by $\mathbf{w}_{ij} = x^\nu(i) - x^\nu(j)$. These lines are located at a distance d_A from codeword i. In Figure 7 for example, A_1 is determined by two hyperplanes (two lines), one between codewords 1-2 and another between 1-3. Let the point $H_i \in \mathbb{R}^\nu$ denote the intersection of the lines and $\mathbf{u}_i$ denote a unitary vector in the direction of codeword $x^\nu(i)$. Then, $H_i = \left(\sqrt{nP} - d_C\right)\mathbf{u}_i$, where $d_C = \frac{1}{2}(1 - t)\sqrt{3nP}\sin\alpha_C$, and the angle α_C is given by the geometry of the constellation. Next, we use point H_i and vector $\mathbf{w}_{ij}$ to characterize each separating line defining region A_i as: $\{\mathbf{x} \in \mathbb{R}^\nu \mid \langle \mathbf{w}_{ij}, \mathbf{x}\rangle = b_i\}$, where $b_i = \langle \mathbf{w}_{ij}, H_i\rangle$. The region A_1 is defined as

$$A_1 = \bigcap_{j=2,3} \left\{y^\nu := \mathbf{y} : \langle \mathbf{w}_{1j}, \mathbf{y}\rangle \geq \langle \mathbf{w}_{1j}, H_1\rangle\right\}, \quad (43)$$

and the other two regions follow in a similar manner.

In the following, we assume that $W = 1$ is sent. Then, in the second phase of the BB, the receiver's decoding decision (ACK/NACK) is encoded using the single channel use feedback codeword $U_{\nu+1}$ in (44): NACK is encoded as a very high amplitude signal since $\mathsf{P}_1(\text{NACK})$ is exponentially small; for ACKs nothing is sent:

$$U_{\nu+1} = \begin{cases} 0, & \text{if ACK} \\ \sqrt{\frac{\Delta}{\mathsf{P}_1(\text{NACK})}}, & \text{if NACK} \end{cases}, \quad (44)$$

for some $\Delta > 0$. By symmetry, $\mathsf{P}_1(\text{NACK}) = \mathsf{P}_2(\text{NACK}) = \mathsf{P}_3(\text{NACK})$. To verify that NACK events occur very rarely,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory

observe first that $\mathsf{P}_1(\text{NACK}) = 1 - \mathsf{P}_1(\text{ACK})$ and that $\mathsf{P}_1(\text{ACK}) \geq \mathsf{P}_1(y^\nu \in A_1) \geq 1 - 2Q\left(\frac{d_A}{\sigma}\right)$, and hence

$$\mathsf{P}_1(\text{NACK}) \leq 2Q\left(\frac{d_A}{\sigma}\right) \leq \exp\left(-\frac{d_A^2}{2\sigma^2}\right).$$

Plugging in $d_A = \frac{1}{2}(d' - td')$ and since $d' = \sqrt{3nP}$, $\mathsf{P}_1(\text{NACK}) \leq \exp\left(-n\frac{3P}{8\sigma^2}(1-t)^2\right)$.

Feedback transmission $U_{\nu+1}$ is decoded by the transmitter by comparing the received $z_{\nu+1}$ with a large threshold Υ. A "NACK" is declared if $z_{\nu+1} > \Upsilon$, otherwise, an "ACK" is declared. As in [19], Υ is chosen to be n, and one can verify that with this choice, $\mathsf{P}_1(\text{"NACK"} \mid \text{ACK})$ and $\mathsf{P}_1(\text{"ACK"} \mid \text{NACK})$ decay to zero in n faster than any exponential, for all $W \in \{1,2,3\}$, i.e.

$$-\lim_{n\to\infty} \frac{1}{n} \ln \mathsf{P}_1(\text{"NACK"} \mid \text{ACK}) = +\infty \tag{45}$$

$$-\lim_{n\to\infty} \frac{1}{n} \ln \mathsf{P}_1(\text{"ACK"} \mid \text{NACK}) = +\infty. \tag{46}$$

During the retransmission stage (the third and final stage of the BB), the transmitter sends a length three codeword that depends on the true message W and whether a "NACK" or "ACK" is declared after the feedback stage. This codeword is generated using (47), for the transmission of $|\mathcal{W}|+1=4$ messages from $\{\mathcal{W}\cup\{0\}\}$.

This encoding consists of an all zeros codeword associated to an "ACK" event, and an index location based signaling associated to a "NACK", in which the codeword is non-zero only in the location indexed by the true message $W=w$, i.e. transmits a very high amplitude $\sqrt{\frac{\Delta'}{\mathsf{P}_w(\text{"NACK"})}}$ in position w (since $\mathsf{P}_w(\text{"NACK"})$ is exponentially small) and zero in all the remaining $M-1=2$ positions, for some $\Delta' > 0$,

$$(X_{n-2}, X_{n-1}, X_n) = \begin{cases} \sqrt{\frac{\Delta'}{\mathsf{P}_w(\text{"NACK"})}} \cdot (0,0,0), & \text{if "ACK"} \\ \sqrt{\frac{\Delta'}{\mathsf{P}_w(\text{"NACK"})}} \cdot (1,0,0), & \text{if "NACK" \& } W=1 \\ \sqrt{\frac{\Delta'}{\mathsf{P}_w(\text{"NACK"})}} \cdot (0,1,0), & \text{if "NACK" \& } W=2 \\ \sqrt{\frac{\Delta'}{\mathsf{P}_w(\text{"NACK"})}} \cdot (0,0,1), & \text{if "NACK" \& } W=3. \end{cases} \tag{47}$$

At the end of the retransmission stage, the receiver compares each of the signals received in the last $M=3$ time slots, i.e. (y_{n-2}, y_{n-1}, y_n) with a very large threshold $\Upsilon = n$, such that for $i \in \{n-2, n-1, n\}$:

$$-\lim_{n\to\infty} \frac{1}{n} \ln \mathsf{P}_1\left(Y_i < \Upsilon \mid X_i = \sqrt{\frac{\Delta'}{\mathsf{P}_1(\text{"NACK"})}}\right) = +\infty \tag{48}$$

$$-\lim_{n\to\infty} \frac{1}{n} \ln \mathsf{P}_1\left(Y_i > \Upsilon \mid X_i = 0\right) = +\infty. \tag{49}$$

Note from (47) and the assumption that $W=1$ is sent, that the received signal at the i-th position during the last $M=3$ channel uses, Y_i, is a random variable distributed in the case of "ACK" as: $Y_i \sim \mathcal{N}\left(0,\sigma^2\right)$, for $i \in \{n-2, n-1, n\}$; and in the case of "NACK" as: $Y_i \sim \mathcal{N}\left(\sqrt{\frac{\Delta'}{\mathsf{P}_1(\text{"NACK"})}}, \sigma^2\right)$, if $i = n-2$, and $Y_i \sim \mathcal{N}\left(0,\sigma^2\right)$, if $i \in \{n-1, n\}$. At the end of the three stages, (48) and (49) guarantee that once the retransmission sequence of length $M=3$ is received, the decoder can use the following rule to determine $\hat{W}$ as:

$$\hat{W} = \begin{cases} \tilde{W}, & \text{if ACK} \\ 1, & \text{if NACK and } Y_{n-2} > \Upsilon\ ,\ Y_{n-1} < \Upsilon\ ,\ Y_n < \Upsilon \\ 2, & \text{if NACK and } Y_{n-2} < \Upsilon\ ,\ Y_{n-1} > \Upsilon\ ,\ Y_n < \Upsilon \\ 3, & \text{if NACK and } Y_{n-2} < \Upsilon\ ,\ Y_{n-1} < \Upsilon\ ,\ Y_n > \Upsilon, \end{cases} \tag{50}$$

where $\tilde{W}$ corresponds to the codeword decoded at the receiver based on the sequence received in the first $\nu = n-4$ (more generally $n-(M+1)$) channel uses using minimum distance decoding: $\tilde{W} = \arg\min_{w\in\{1,2,3\}} ||x^\nu(w) - y^\nu||$.

2) Probability of error analysis: Assuming $W=1$ was sent, from the geometry of the problem in Figure 7, it suffices to upper bound $\mathsf{P}(\text{error}) = \mathsf{P}_1(W \neq \hat{W})$ as

$$\begin{aligned} &\mathsf{P}_1(\text{error}) \\ &= \overbrace{\mathsf{P}_1(\text{NACK})\mathsf{P}_1(\text{"NACK"}|\text{NACK})}^{\leq 1}\overbrace{\mathsf{P}_1(\text{error}|\text{NACK}\ ,\ \text{"NACK"})}^{\text{exceedingly small by (48) and (49)}} \\ &+ \overbrace{\mathsf{P}_1(\text{NACK})}^{\leq 1}\overbrace{\mathsf{P}_1(\text{"ACK"}|\text{NACK})}^{\text{exceedingly small by (46)}} \cdot \overbrace{\mathsf{P}_1(\text{error}|\text{NACK}\ ,\ \text{"ACK"})}^{\leq 1} \\ &+ \overbrace{\mathsf{P}_1(\text{ACK})}^{\leq 1}\overbrace{\mathsf{P}_1(\text{"NACK"}|\text{ACK})}^{\text{exceedingly small by (45)}} \cdot \overbrace{\mathsf{P}_1(\text{error}|\text{ACK}\ ,\ \text{"NACK"})}^{\leq 1} \\ &+ \mathsf{P}_1(\text{ACK})\mathsf{P}_1(\text{"ACK"} \mid \text{ACK}) \cdot \mathsf{P}_1(\text{error} \mid \text{ACK}\ ,\ \text{"ACK"}) \\ &\doteq \mathsf{P}_1(\text{ACK})\mathsf{P}_1(\text{"ACK"} \mid \text{ACK}) \cdot \mathsf{P}_1(\text{error} \mid \text{ACK}\ ,\ \text{"ACK"}) \\ &\leq \mathsf{P}_1(\text{error} \mid \text{ACK}\ ,\ \text{"ACK"}) = \mathsf{P}_1(\text{error} \mid \text{ACK}). \end{aligned} \tag{51}$$

Equation (51) is similar to [19, Equation (132)]. To upper bound $\mathsf{P}_1(\text{error} \mid \text{ACK})$ note first that an ACK implies the received signal is within one of the three codeword regions A_w (shown in Figure 7). Since $W=1$, an error occurs only if the received signal is in one of the two codeword regions A_2 or A_3, and so

$$\begin{aligned} \mathsf{P}_1(\text{error} \mid \text{ACK}) &= \frac{\mathsf{P}_1(\text{error}, \text{ACK})}{\mathsf{P}_1(\text{ACK})} \\ &= \frac{\mathsf{P}_1\left(y^\nu \in A_2 \cup A_3\right)}{\mathsf{P}_1\left(y^\nu \in A_1 \cup A_2 \cup A_3\right)}. \end{aligned} \tag{52}$$

By symmetry, $\mathsf{P}_1(y^\nu \in A_2) = \mathsf{P}_1(y^\nu \in A_3)$, and hence

$$\begin{aligned} \mathsf{P}_1\left(y^\nu \in A_2 \cup A_3\right) &\leq \mathsf{P}_1\left(y^\nu \in A_2\right) + \mathsf{P}_1\left(y^\nu \in A_3\right) \\ &= 2\mathsf{P}_1\left(y^\nu \in A_2\right) \leq 2Q\left(\frac{d_B}{\sigma}\right) \\ &\leq \exp\left(-n\frac{3P}{8\sigma^2}(1+t)^2\right) \end{aligned}$$

where, $d_B = \frac{1}{2}(d' + td')$ is as shown in Figure 7. We can further upper bound (52) by lower bounding the denominator as $\mathsf{P}_1\left(y^\nu \in A_1 \cup A_2 \cup A_3\right) \geq \mathsf{P}_1\left(y^\nu \in A_1\right) \geq 1 - 2Q\left(\frac{d_A}{\sigma}\right) \geq 1 - \exp\left(-n\frac{3P}{8\sigma^2}(1-t)^2\right)$.

Therefore, $\mathsf{P}_1(\text{error} \mid \text{ACK}) \doteq \exp\left(-n\frac{3P}{8\sigma^2}(1+t)^2\right)$.

Finally, the achievable error exponent for the BB, since parameter t can be chosen sufficiently close to 1, is:

$$E_{\text{BB}}^{\text{EXP}} \geq \frac{3P}{8\sigma^2}(1+t)^2 = \frac{3P}{2\sigma^2}. \tag{53}$$

Equation (53) shows that the BB leads to a four-fold gain over non-feedback transmission under the AS power constraint in (6), $E^{\text{AS}}(P,\sigma^2)=\frac{3P}{8\sigma^2}$ as a result of a more flexible power constraint and noisy feedback. Next, following [19], we use the BB for the transmission of $M=3$ messages in a three-stage communication scheme.

B. A transmission scheme based on the building block for the transmission of $M=3$ messages

This scheme comprises three stages: transmission, active feedback and retransmission, lasting $\lambda(n-3)$, $(1-\lambda)(n-3)$ and 3 channel uses respectively for $\lambda\in(0,1)$. We describe each of them as follows:

1. Transmission: Message W is transmitted using the BB, using power $\frac{1}{\lambda}(P-\eta)$ on the forward link and $\frac{\eta}{\lambda}$ in the feedback link. The transmitter reserves power $0<\eta<\min\{P,P_{\text{FB}}\}$ to provide feedback to the receiver's transmission in the next stage. Denoting the estimation of the true message at the end of this stage by W', from (53), the probability of error of this stage is:

$$\begin{aligned}\mathsf{P}_1(W'\neq W)&\doteq\exp\left(-\lambda(n-3)\frac{3\frac{1}{\lambda}(P-\eta)}{2\sigma^2}\right)\\&=\exp\left(-n\frac{3(P-\eta)}{2\sigma^2}\right).\end{aligned}\tag{54}$$

2. Active feedback: W' is sent to the transmitter using the BB. The receiver uses power $\frac{1}{1-\lambda}(P_{\text{FB}}-\eta)$ and the transmitter uses $\frac{\eta}{1-\lambda}$. The estimate of W' is W'' leads to the probability of error, given by (53):

$$\begin{aligned}&\mathsf{P}_1(W''\neq W'\mid W')\doteq\\&\exp\left(-(1-\lambda)(n-3)\frac{3\frac{1}{(1-\lambda)}(P_{\text{FB}}-\eta)}{2\sigma_{\text{FB}}^2}\right)\\&=\exp\left(-n\frac{3(P_{\text{FB}}-\eta)}{2\sigma_{\text{FB}}^2}\right).\end{aligned}\tag{55}$$

3. Retransmission: In the retransmission stage, the transmitter first compares W'' with W, and generates a length M retransmission codeword as in the third stage of the BB, with the single difference that the amplitude of the transmissions is $\sqrt{\frac{\Delta'}{\mathsf{P}_1(W''\neq W)}}$. Finally, the receiver uses the decoding rule of (56) (setting $\hat{W}=W'$) to estimate message W, based on the length $M=3$ codeword received during the retransmission stage:

$$\hat{W}=\begin{cases}W', & \text{if } Y_{n-2}<\Upsilon\ ,\ Y_{n-1}<\Upsilon\ ,\ Y_n<\Upsilon\\1, & \text{if } Y_{n-2}>\Upsilon\ ,\ Y_{n-1}<\Upsilon\ ,\ Y_n<\Upsilon\\2, & \text{if } Y_{n-2}<\Upsilon\ ,\ Y_{n-1}>\Upsilon\ ,\ Y_n<\Upsilon\\3, & \text{if } Y_{n-2}<\Upsilon\ ,\ Y_{n-1}<\Upsilon\ ,\ Y_n>\Upsilon\end{cases}.\tag{56}$$

As in the BB decoding rule, each of the three signals Y_i for $i=n-2,n-1,n$, is compared with a threshold $\Upsilon=n$. Since $\mathsf{P}_w(W''\neq W)$ can be shown to be exponentially small, it follows that

$$-\lim_{n\to\infty}\frac{1}{n}\ln\mathsf{P}_1\left(Y_i<\Upsilon\mid X_i^M=\sqrt{\frac{\Delta'}{\mathsf{P}_1(W''\neq W)}}\right)=+\infty\tag{57}$$

$$-\lim_{n\to\infty}\frac{1}{n}\ln\mathsf{P}_1\left(Y_i>\Upsilon\mid X_i^M=0\right)=+\infty.\tag{58}$$

The probability of error of the scheme, considering decoding rule (56), along with (57) and (58) is:

$$\begin{aligned}&\mathsf{P}_1(\text{error})\\&=\overbrace{\mathsf{P}_1(\text{error}\mid W''=1,W'=1)}^{\text{negligible by (58)}}\cdot\overbrace{\mathsf{P}_1(W''=1,W'=1)}^{\leq 1}\\&+\overbrace{\mathsf{P}_1(\text{error}\mid W''=2,W'=1)}^{\text{negligible by (57) and (58)}}\cdot\overbrace{\mathsf{P}_1(W''=2,W'=1)}^{\leq 1}\\&+\mathsf{P}_1(\text{error}\mid W''=1,W'=3)\cdot\mathsf{P}_1(W''=1,W'=3)\\&+\mathsf{P}_1(\text{error}\mid W''=1,W'=2)\cdot\mathsf{P}_1(W''=1,W'=2)\\&+\overbrace{\mathsf{P}_1(\text{error}\mid W''=2,W'=2)}^{\text{negligible by (57) and (58)}}\cdot\overbrace{\mathsf{P}_1(W''=2,W'=2)}^{\leq 1}\\&+\overbrace{\mathsf{P}_1(\text{error}\mid W''=2,W'=3)}^{\text{negligible by (57) and (58)}}\cdot\overbrace{\mathsf{P}_1(W''=2,W'=3)}^{\leq 1}\\&+\overbrace{\mathsf{P}_1(\text{error}\mid W''=3,W'=1)}^{\text{negligible by (57) and (58)}}\cdot\overbrace{\mathsf{P}_1(W''=3,W'=1)}^{\leq 1}\\&+\overbrace{\mathsf{P}_1(\text{error}\mid W''=3,W'=2)}^{\text{negligible by (57) and (58)}}\cdot\overbrace{\mathsf{P}_1(W''=3,W'=2)}^{\leq 1}\\&+\overbrace{\mathsf{P}_1(\text{error}\mid W''=3,W'=3)}^{\text{negligible by (57) and (58)}}\cdot\overbrace{\mathsf{P}_1(W''=3,W'=3)}^{\leq 1}\\&\doteq\mathsf{P}_1(\text{error}\mid W''=1,W'=2)\cdot\mathsf{P}_1(W''=1,W'=2)\\&+\mathsf{P}_1(\text{error}\mid W''=1,W'=3)\cdot\mathsf{P}_1(W''=1,W'=3)\\&\leq\mathsf{P}_1(W''=1,W'=2)+\mathsf{P}_1(W''=1,W'=3).\end{aligned}$$

Since $\mathsf{P}_1(W''=1,W'=2)=\mathsf{P}_1(W''=1,W'=3)$,

$$\begin{aligned}\mathsf{P}_1(\text{error})&\leq 2\mathsf{P}_1(W''=1,W'=2).\\&=2\mathsf{P}(W''=1|W'=2,W=1)\cdot\mathsf{P}(W'=2|W=1)\\&\doteq\exp\left(-n\frac{3(P_{\text{FB}}-\eta)}{2\sigma_{\text{FB}}^2}\right)\exp\left(-n\frac{3(P-\eta)}{2\sigma^2}\right)\quad(59)\\&=\exp\left(-n\left(\frac{3(P-\eta)}{2\sigma^2}+\frac{3(P_{\text{FB}}-\eta)}{2\sigma_{\text{FB}}^2}\right)\right),\quad(60)\end{aligned}$$

where (59) results from using (54) and (55). Then, (60) implies that $E^{\text{EXP}}\geq\frac{3}{2}\left(\frac{P}{\sigma^2}+\frac{P_{\text{FB}}}{\sigma_{\text{FB}}^2}\right)$ is achievable since η can be chosen sufficiently small. $\square$

VI. On the largest number of transmitted messages M

The main restriction on M comes from the use of simplex codes for both the one-way and two-way AWGN channels. These codes require all symbols to have the same energy and pairwise distance. It is known [46], [47] [48, pp. 65–67, Proposition 4.1] that the unique solution to placing $M\leq n+1$ equally likely points on the surface of a unitary sphere in $\mathbb{R}^n$ such that the distance between any two points is maximized corresponds to a regular simplex. Thus, codewords of length n may be used to transmit a point from a regular simplex code with M messages as long as $M\leq n+1$. The regular simplex code is not necessarily optimal in the sense of minimizing the overall probability of error for the AWGN under the AS constraint, but it has been shown to be globally optimal in the sense of maximizing the union bound on the error probability for all SNR [49], as it maximizes the minimum distance between codewords.

A. Bounds on M for the AS constraint

In Theorem 2 the active feedback transmits the set of codeword pairs that is most likely to have occurred in the first stage. This feedback stage becomes a bottleneck, as this number is $L = \binom{M}{2} = \frac{M(M-1)}{2}$, and hence requires at least $L-1$ channel uses when using a simplex code for transmission. In the first stage, the number of dimension of the simplex code is $\lambda(n-1) = M-1$. The total block length n is thus $\lambda(n-1)+(1-\lambda)(n-1)+1 = n$, which becomes $(M-1)+\left(\frac{M^2-M}{2}-1\right)+1 = n$. The above result can be represented as the quadratic equation $M^2+M-(2+2n) = 0$ whose positive root is $M = \frac{\sqrt{8n+9}-1}{2} = \sqrt{2n+\frac{9}{4}}-\frac{1}{2}$. As n becomes large, $M \approx \sqrt{2n}$ may thus be supported for $\lambda \approx \frac{2}{\sqrt{2n}}$.

In the case of non-linear passive feedback of Theorem 1, M can be as large as n by setting the duration of the retransmission stage to one channel use and leaving $n-1$ channel uses for the simplex code of the transmission/passive-feedback stage, thus leaving room for exactly $M = n$ symbols. Notice that error exponent of the passive linear scheme of Theorem 3 deteriorates for large M, and hence also in Theorem 7.

For the two-way channel, recall Theorems 5 and 6, and the block diagrams depicted in Figure 3 of Section II-B where both terminals transmit the same number of messages M. We first consider passive feedback in Theorem 5. This scheme allocates the first λn slots to the transmission of message W_2 over the $1 \leftarrow 2$ channel without feedback. The remaining channel uses are used for the transmission of W_1 in the opposite (weaker) direction with the help of passive feedback using two stages: transmission/passive-feedback in $(1-\lambda)n-1$ channel uses, and retransmission in a single channel use (see [18, Section II-A]). The total block length n is $\lambda n+[(1-\lambda)n-1]+1 = n$. Since a simplex code for M symbols in each direction requires $M-1$ dimensions, it follows that the number of messages to be transmitted can be as large as $M \approx \frac{n}{2}$, with $\lambda \approx n/2$.

For active feedback, consider the block diagram of Figure 3 (b). Note that the transmission of message W_2 without feedback over the stronger $1 \leftarrow 2$ direction, and the feedback-free transmission of the first stage of the active feedback scheme employed in the weaker $1 \rightarrow 2$ direction (for message W_1) occur simultaneously over $\lambda(n-1)$ channel uses. These transmissions are based on a simplex code of M symbols and therefore $\lambda(n-1) = M-1$. The active feedback stage occurs in the next $(1-\lambda)(n-1)$ channel uses, followed by a single channel use for the retransmission, see (39). As in the one-way channel, the allowable number of messages is mainly determined by the active feedback stage, which uses the largest fraction of the block length n. Thus, again, $M \approx \sqrt{2n}$.

B. Bounds on M for the EXP constraint

For the scheme presented in Theorem 4 the largest value that M can take is determined by the duration of the retransmission stage of the BB, which is used in the (overall) three-stage-scheme. For the transmission of M messages, the BB requires $M-1$ channel uses to support the simplex code of the transmission stage, a single channel use for the feedback signaling, and finally, a length M codeword for retransmission. Thus, as a function of M, the block length of the BB is $2M$. In the three-stage-scheme, the BB is used two times: to transmit W in the forward direction; and to feed back the estimate W' of W to the transmitter. Then, a length M codeword is used in the third stage. Hence, the number of channel uses required for this scheme is $2M+2M+M \leq n$, from which we conclude that $M \leq \frac{1}{5}n$. This also holds for the two-way communication scheme in Theorem 8.

The results summarizing how large M may be (as $n \rightarrow \infty$) are in Table I.

TABLE I
SUMMARY OF RESULTS FOR THE LARGEST M IN EACH ACHIEVABILITY SCHEME

Power Contraint	FB-type	One-Way	Two-Way
AS	NL-Passive	n	$\frac{n}{2}$
	NL-Active	$\sqrt{2n}$	$\sqrt{2n}$
EXP	Active	$\frac{n}{5}$	$\frac{n}{5}$

VII. Numerical simulations

We numerically evaluate the newly derived inner bounds for different SNR conditions and number of messages. Simulations include the feedback-free and perfect feedback achievable error exponents (or regions in case of the two-way channel) as references.

A. Simulations for the one-way AWGN channel

Figure 8 depicts achievable error exponents under the AS constraint as a function of the noise variance in the feedback link for $\sigma_{\mathrm{FB}}^2 \in [10^{-7}, 1]$, $M = 3$, $P = P_{\mathrm{FB}} = 1$, and $\sigma^2 = 1$. The figure on the left utilizes a linear scale for σ_{FB}^2 on the horizontal axis; the right utilizes a logarithmic one. The plots include the upper bound under perfect feedback, given in (7) and corresponding to $\frac{P}{2\sigma^2} = 0.5$; and the lower bound in the absence of feedback, given in (6), corresponding to $\frac{3P}{8\sigma^2} = 0.375$, which are shown with a dashed and a dotted black lines respectively. The linear and non-linear *passive* feedback coding schemes from [18] (modified in this work to satisfy an AS power constraint in both directions as in Theorems 3 and 1) are plotted in continuous red and dash-dotted blue lines respectively, whereas the *active* feedback coding scheme proposed in Theorem 2 is plotted in gray.

B. Simulations for the two-way AWGN Channel

1) Numerical simulation for $M = 3$ under AS power constraint: This simulation illustrates how our schemes lead to an achievable EER that shows an improvement over the feedback-free EER of the weaker direction, at the cost of a reduction of the one in the stronger direction. The forward and feedback SNRs used in Figure 9 are $\frac{P_1}{\sigma_2^2} = 1$ and $\frac{P_2}{\sigma_1^2} = 1000$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3066855, IEEE Transactions on Information Theory

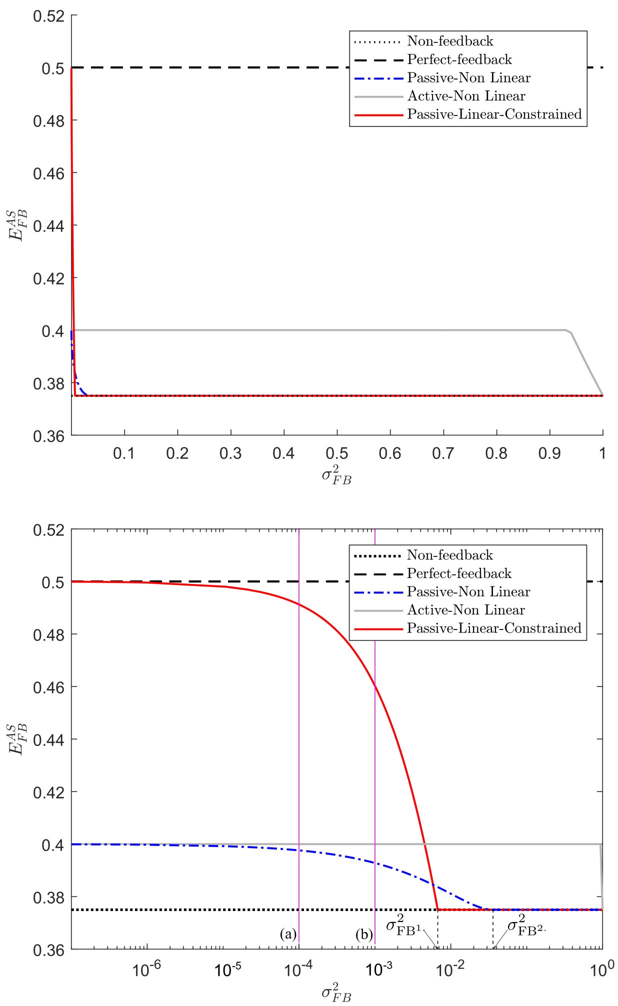


Fig. 8. Achievable error exponent under AS: coding schemes comparison for $M = 3$, where $\sigma^2_{\mathrm{FB}^1} \approx 6.73 \times 10^{-3}$ and $\sigma^2_{\mathrm{FB}^2} \approx 35.71 \times 10^{-3}$. Points (a) and (b) correspond to the pair of SNR values we evaluated in Figure 9 for two-way communication.

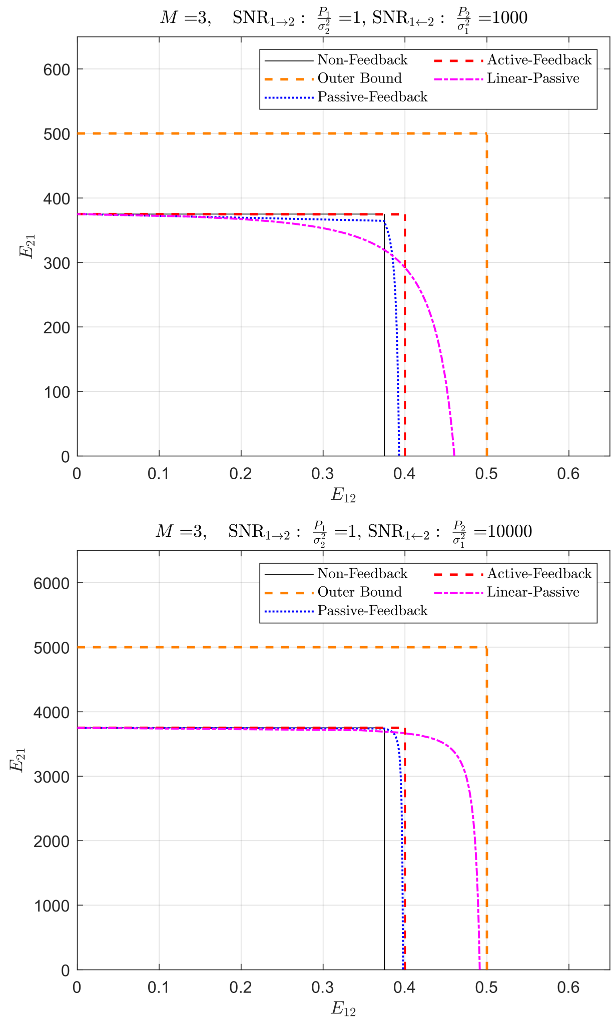


Fig. 9. Error exponent region for the two-way AWGN channel under the AS constraint and the transmission of three messages.

for the left, and $\frac{P_2}{\sigma_1^2} = 10000$ for the right plot (these reflect the values (a) and (b) for σ^2_{FB} in Figure 8). Both figures include the non-feedback inner bound of Proposition 1 in solid black; the outer bound of Proposition 3 in dashed orange; the non-linear passive feedback (Theorem 5) in dotted blue; the non-linear active feedback (Theorem 6) is dashed red; the linear feedback scheme of Theorem 7 in dash-dotted magenta. Consider the $1 \to 2$ direction and observe that the achievable error exponent of the non-linear active feedback scheme is higher than the one achieved by the non-linear passive feedback scheme. At these SNRs, the linear passive feedback scheme achieves a larger error exponent the $1 \to 2$ direction at the cost of a larger decrease of the error exponent in the $1 \leftarrow 2$ direction. Such high error exponents are possible only when one direction is much stronger than the other. See Figure 8 for the one-way channel and observe that the linear scheme produces error exponent gains over feedback-free transmissions as the feedback noise decreases, i.e. when $\sigma^2_{\mathrm{FB}} \le \sigma^2_{\mathrm{FB}^2}$.

Figure 10 shows the achievable EER obtained by Theorems 5, 6 and 7 for different values of M. Note that as M increases, the achievable feedback-free error exponent in both directions decreases until $1/4$ of the corresponding channel SNR. This numerical evaluation shows the active feedback scheme attains error exponents above those attained with the non-linear passive feedback even as M increases. Moreover, this figure visualizes how the error exponent achieved with the linear passive feedback scheme is reduced dramatically as M increases.

2) Numerical simulation for M under EXP power constraint: Figure 11 shows the achievable EER for the two-way AWGN channel with active noisy feedback under the EXP power constraint for increasing M. The largest region is attained for the transmission of $M = 2$. Note that the red line corresponds to a large value of M and that even for this case, the EXP constraints yield an achievable region that completely contains what is achievable under the AS, which is shown with a solid black line at the bottom left square region.

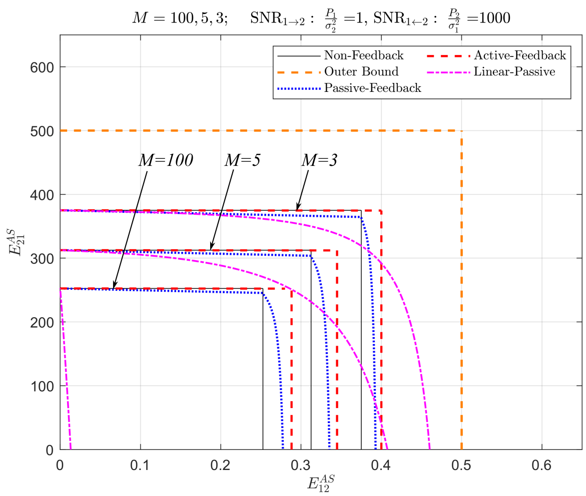


Fig. 10. Achievable error exponent regions provided by Theorems 5, 6 and 7 for different values of M.

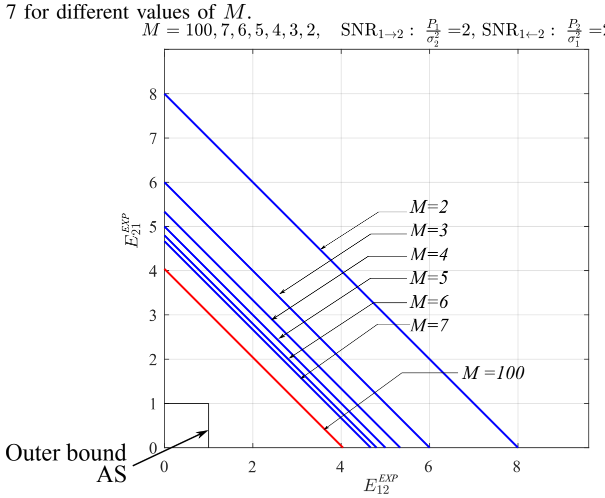


Fig. 11. Achievable error exponent region for the two-way AWGN channel with active noisy feedback under the EXP power constraint and symmetric SNR, Theorem 8.

VIII. Conclusions

Achievable error exponents for the one-way with noisy, power constrained feedback, and two-way AWGN channel under the AS and EXP power constraints were derived for the transmission of a finite number of messages. Under the AS power constraint, noisy feedback only *appears* (based on achievability results) to provide error exponent gains over feedback-free transmission when the feedback channel SNR is (considerably) larger than the forward channel SNR. Proving this is an open problem. In the two-way setting, we demonstrated achievable error exponent pairs that illustrate a trade off between allocating resources towards one direction's transmission versus sending feedback for the other. For the EXP constraint in particular, finding an achievable error exponent region that outperforms time sharing is an open problem. Deriving tighter outer bounds also remains to be done.

Appendix A
Proof of Theorem 1 for passive non-linear feedback under the AS constraint.

We adjust the non-linear passive feedback scheme of [18] by feeding back a symbol-by-symbol scaled output signal that ensures that the AS constraints are met for both forward ($\mathsf{P}\left(\sum_{k=1}^n X_k^2 \le nP\right) = 1$) and feedback ($\mathsf{P}\left(\sum_{k=1}^n U_k^2 \le nP_{\text{FB}}\right) = 1$) directions. Passive feedback is meant here to denote feedback of the form $U_k = \alpha Y_k$, where α is an optional scaling factor. The scheme of [18, Theorem 1] uses $\alpha = 1$. First, we show that

$$\mathsf{P}\left(\sum_{k=1}^n Y_k^2 > nP + 5\sigma^2 n + \sigma n\sqrt{P}\right) \to 0, \quad \text{as } n \to \infty,$$

which follows as

$$\begin{aligned}
&\mathsf{P}\left(\sum_{k=1}^n Y_k^2 > nP + 5\sigma^2 n + \sigma n\sqrt{P}\right)\\
&= \mathsf{P}\left(\sum_{k=1}^n \left(X_k^2 + N_k^2 + 2X_kN_k\right) > nP + 5\sigma^2 n + \sigma n\sqrt{P}\right)\\
&\le \mathsf{P}\left(\sum_{k=1}^n \left(X_k^2 + N_k^2\right) > nP + 5\sigma^2 n\right)\\
&\quad + \mathsf{P}\left(\sum_{k=1}^n 2X_kN_k > \sigma n\sqrt{P}\right)\\
&\le \mathsf{P}\left(\sum_{k=1}^n X_k^2 > nP\right) + \mathsf{P}\left(\sum_{k=1}^n N_k^2 > 5\sigma^2 n\right)\\
&\quad + \mathsf{P}\left(\sum_{k=1}^n 2X_kN_k > \sigma n\sqrt{P}\right)
\end{aligned}$$

The first term vanishes given that the AS constraint imposed in the forward channel inputs $\mathsf{P}\left(\sum_{k=1}^n X_k^2 \le nP\right) = 1$ is satisfied. The second vanishes as

$$\begin{aligned}
&\mathsf{P}\left(\sum_{k=1}^n N_k^2 > \sigma^2(n + 2\sqrt{n^2} + 2n)\right)\\
&= \mathsf{P}\left(\sum_{k=1}^n \left(\frac{N_k}{\sigma}\right)^2 \ge 5n\right) \le \exp\left(-n\right) \qquad (61)
\end{aligned}$$

as $\sum_{k=1}^n \left(\frac{N_k}{\sigma^2}\right)^2$ corresponds to a chi-square random variable with n degrees of freedom χ_n^2. Its tail may be upper bounded using the Laurent-Massart upper bound given in [50, Lemma 1], that gives an exponential upper bound for any random variable $U \sim \chi_D^2$ with D degrees of freedom, for any positive x as:

$$\mathsf{P}\left(U - D \ge 2\sqrt{Dx} + 2x\right) \le \exp\left(-x\right). \qquad (62)$$

Thus, (61) results from (62), by taking $U = \sum_{k=1}^n \left(\frac{N_k}{\sigma}\right)^2$, $D = n$, and $x = n$. The third term vanishes as

$$\mathsf{P}\left(\sum_{k=1}^n 2X_kN_k \ge \sigma n\sqrt{P}\right) = Q\left(\frac{\frac{\sigma n\sqrt{P}}{2}}{\sqrt{\sigma^2 \sum_{k=1}^n X_k^2}}\right)$$

$$\stackrel{(a)}{\leq} Q\left(\frac{n\sqrt{P}}{2\sqrt{nP}}\right) = Q\left(\frac{\sqrt{n}}{2}\right)$$
$$\stackrel{(b)}{\leq} \frac{1}{2}\exp\left(-n/2\right)$$

where (a) follows by noticing that $\sum_{k=1}^{n} X_k N_k \sim \mathcal{N}\left(0, \sigma^2 \sum_{k=1}^{n} X_k^2\right)$ since the N_k's are i.i.d. for all k, and by the forward channel input constraint $\sum_{k=1}^{n} X_k^2 \leq nP$ and (b) from the Chernoff bound for the Q-function: $Q(x) \leq \frac{1}{2}\exp\left(-x^2/2\right)$.

We alter the scheme of [18] by sending $U_k = \alpha Y_k$, for

$$\alpha = \sqrt{\frac{nP_{\text{FB}}}{nP + 5\sigma^2 n + \sigma n\sqrt{P}}} = \sqrt{\frac{P_{\text{FB}}}{P + 5\sigma^2 + \sigma\sqrt{P}}}$$

which will guarantee the AS power constraint of the feedback direction $\mathbb{P}\left(\sum_{k=1}^{n} U_k^2 \leq nP_{\text{FB}}\right) = 1$ is met. We next consider the effect this has on the error exponent.

The error exponent achieved by the passive non-linear scheme in [18] may be expressed for general M as in (10), which was proven in detail in [18] for $M = 3$ as:

$$E_{\text{FB}}^{\text{AS-NL}}\left(M, P, \sigma^2, s\right) \geq \min\left\{\frac{P}{2\sigma^2}\left(\frac{s^2 - 2s + 4}{s^2 - 2s + 5}\right), \underbrace{\frac{P}{\sigma_{\text{FB}}^2}\frac{3}{8}\left(\frac{s^2}{s^2 - 2s + 5}\right)}_{\text{passive feedback}}\right\} \tag{63}$$

The term on the right of (63) corresponds to the error exponent of the passive feedback stage, and results from an upper bound on the probability of event $\tilde{\mathcal{E}}_{12}$ in [18, Eq. (8)]:

$$\mathsf{P}(\tilde{\mathcal{E}}_{12}) \leq 2Q\left(\frac{d_6}{\sigma_{\text{FB}}}\right), \tag{64}$$

where $d_6 = s\sqrt{\frac{3\lambda nP}{16}}$ is indicated in [18, Figure 6(b)] (and where eventually $\lambda = \frac{4}{s^2 - 2s + 5}$ is chosen. The term on the right of (63) follows by further upper bounding (64) using the Chernoff bound.

When the feedback signals $U_k = \alpha Y_k$ are scaled, so is the distance d_6 used in (64), which is now denoted as d_6' and given by $d_6' = \alpha s\sqrt{\frac{3nP}{4(s^2-2s+5)}} = s\sqrt{\frac{3nP}{4(s^2-2s+5)}\left(\frac{P_{\text{FB}}}{P+5\sigma^2+\sigma\sqrt{P}}\right)}$. Hence, the error exponent of the passive feedback stage is given by:

$$\begin{aligned}
\mathsf{P}(\tilde{\mathcal{E}}_{12}) &\leq 2Q\left(\frac{d_6'}{\sigma_{\text{FB}}}\right) \\
&= 2Q\left(\frac{s}{\sigma_{\text{FB}}}\sqrt{\frac{3nP}{4(s^2 - 2s + 5)}\left(\frac{P_{\text{FB}}}{P + 5\sigma^2 + \sigma\sqrt{P}}\right)}\right) \\
&\leq \exp\left[-n\frac{P}{\sigma_{\text{FB}}^2}\frac{3}{8}\left(\frac{s^2}{s^2 - 2s + 5}\right)\frac{P_{\text{FB}}}{P + 5\sigma^2 + \sigma\sqrt{P}}\right].
\end{aligned}$$

Since the forward error exponent (the left term) remains the same, the overall error exponent expression obtained is (63). Applying the scaling factor for general M we obtain (13). □

Appendix B
Proof of Theorem 2 for $M \geq 3$ messages

We provide an abbreviated generalization to $M > 3$ that mimics the proof for three messages. The distances utilized in the proof of Theorem 2 are generalized to any M using the mapping (65) introduced in [18]:

$$d_j^{(3)} = d_j^{(M)}\sqrt{3(M-1)/(2M)}. \tag{65}$$

Consider the transmission of M messages chosen equally likely from the set $\mathcal{W} = \{1, \cdots, M\}$, and assume message $W = 1$ is sent. For the transmission stage, codewords of length $\lambda(n-1)$ are generated using the simplex code $\mathcal{C}(\mathcal{W}, \beta nP)$ defined in (3) and decoded using protection regions defined in (34) with $w, w', w'' \in [1:M]$. The probability of the error event $\mathcal{E}_{\text{T}} = \left\{Y^{\lambda(n-1)} \in \bigcup_{w\neq 1} B_w \cup \left(\bigcup_{w,w'\neq 1} A'_{ww'}\right)\right\}$ is upper bounded by $\mathsf{P}(\mathcal{E}_{\text{T}}) \leq M^2 Q\left(d_5^{(M)}\right) \leq \frac{M^2}{2}\exp\left(-n\frac{M\beta}{12(M-1)}\frac{P}{\sigma^2}(s^2 - 2s + 4)\right)$, where $d_5^{(M)}$ is obtained from d_5 as shown in Figure 6, and $A'_{ww'}$ is defined in (35), following [18, Section II-B] with $w, w', w'' \in [1:M]$.

In the active feedback stage, the most likely codeword pair (in lexicographical order) q is returned to the transmitter. As before, the antipodal signaling of the retransmission stage sends a positive signal to indicate the true message is the first of the pair, and a negative signal to indicate the true message is the second. Since there are $\binom{M}{2}$ lexicographically ordered pairs, the number of data / constellation points of the simplex code used in the active feedback stage is $\binom{M}{2}$, and the probability of error can be upper bounded for the event $\mathcal{E}_{\text{AFB}} = \{q \neq q\}$, using (6) as in (66). Note that since this stage lasts for $(1-\lambda)(n-1)$ channel uses, the receiver may scale up its transmissions by $1/(1-\lambda)$ to use all the available power and satisfy the AS power constraint. Hence, we obtain

$$\begin{aligned}
\mathsf{P}(\mathcal{E}_{\text{AFB}}) &\leq (M-1)\cdot Q\left(\sqrt{n\frac{P_{\text{FB}}}{\sigma_{\text{FB}}^2}\frac{\binom{M}{2}}{2\left(\binom{M}{2}-1\right)}}\right) \\
&\leq \frac{M-1}{2}\exp\left(-n\frac{P_{\text{FB}}}{\sigma_{\text{FB}}^2}\frac{\binom{M}{2}}{4\left(\binom{M}{2}-1\right)}\right).
\end{aligned} \tag{66}$$

Finally, for the retransmission stage, the probability of occurrence of the event $\mathcal{E}_{\text{RT}} = \{(W \in q = q) \cap (\hat{W} \neq W)\}$ can be upper bounded as $\mathsf{P}(\mathcal{E}_{\text{RT}}) = Q\left(-\sqrt{\left(1 - \beta\frac{M-2}{2(M-1)}\right)\frac{P_{\text{FB}}}{\sigma_{\text{FB}}^2}}\right)\frac{1}{2}\exp\left[-n\frac{P}{2\sigma^2}\left(1 - \beta\frac{M-2}{2(M-1)}\right)\right]$. Hence, the overall error exponent is:

$$\begin{aligned}
E_{12}^{\text{AS}} &= \limsup_{n\to\infty} -\frac{1}{n}\ln\mathsf{P}(\text{error}) \\
&\geq \limsup_{n\to\infty} -\frac{1}{n}\max\left\{\ln\mathsf{P}(\mathcal{E}_{\text{T}}), \ln\mathsf{P}(\mathcal{E}_{\text{AFB}}), \ln\mathsf{P}(\mathcal{E}_{\text{RT}})\right\} \\
&\geq \min\left\{\beta\frac{MP}{12\sigma^2(M-1)}(s^2 - 2s + 4), \frac{P_{\text{FB}}}{\sigma_{\text{FB}}^2}\frac{\binom{M}{2}}{4\left(\binom{M}{2}-1\right)},\right. \\
&\qquad \left.\frac{P}{2\sigma^2}\left(1 - \beta\frac{M-2}{2(M-1)}\right)\right\}.
\end{aligned}$$

Since $\binom{M}{2} = \frac{M(M-1)}{2}$, by equating the first and third arguments, $\beta = \frac{6(M-1)}{M(s^2-2s+4)+3(M-2)}$, thus yielding (14). $\square$

APPENDIX C
PROOF OF THEOREM 3 FOR PASSIVE LINEAR FEEDBACK UNDER THE AS CONSTRAINT

The linear passive feedback coding scheme of [18, Theorem 2] is slightly modified in order to ensure that the AS constraint in the feedback link is met. To remain a passive scheme, the feedback signal is assumed to be the forward signal scaled by the factor α as $U_i = \alpha Y_i$. The forward encoding function is thus

$$X_1(w) = \begin{cases} \frac{L+1-w}{L}\sqrt{\lambda n P} & \text{if } M = 2L+1, \\ \frac{L+1/2-w}{L}\sqrt{\lambda n P} & \text{if } M = 2L, \end{cases}$$

and $X_i = (1+\delta)(N_{i-1} + \frac{1}{\alpha}N_{\text{FB}_{i-1}})$ for $2 \le i \le \eta$, where $\eta = \eta(w, N^n, N^n_{\text{FB}})$ is the largest $k \le \bar{n} = \sqrt{n}$ such that: $\sum_{i=1}^k X_i^2 \le nP$. Each X_i is obtained from the feedback direction channel output Z_{i-1}, first by a scaling of $1/\alpha$, followed by a subtraction of the previously sent X_{i-1}. Hence: $X_{i+1} = \frac{1}{\alpha}Z_i - X_i$. Transmissions in the feedback direction are given by:

$$U_i = \begin{cases} \alpha\left(X_1 + N_1\right), & \text{if } i = 1 \\ \alpha(1+\delta)(N_{i-1} + \frac{1}{\alpha}N_{\text{FB}_{i-1}}) + \alpha N_i, & \text{if } 2 \le i \le k \\ \alpha N_i, & \text{if } k+1 \le i \le \bar{n}. \end{cases} \tag{67}$$

Equation (19) is obtained following similar steps as those in [18], but considering the channels' inputs defined above, and choosing $(1+\delta)^2 = 1 + \sqrt{\frac{4L^2\sigma^2_{\text{FB}}/\alpha^2}{\sigma^2 + \frac{\sigma^2_{\text{FB}}}{\alpha^2}}}$.

The scaling factor α is obtained such that, for the feedback channel inputs $U_i = \alpha Y_i$:

$$\mathsf{P}\left(\sum_{i=1}^n U_i^2 > nP_{\text{FB}}\right) \to 0, \quad \text{as } n \to \infty. \tag{68}$$

To prove (68), we follow a similar approach as in Appendix A to first quantify the energy used by the receiver to provide feedback. Using (67) with $k = \bar{n}$, we have that:

$$\begin{aligned}
&\sum_{i=1}^n U_i^2 \\
&= U_1^2 + \sum_{i=2}^{\bar{n}} U_i^2 \\
&= \alpha^2\left[X_1^2 + N_1^2 + 2X_1N_1\right] \\
&\quad + \alpha^2 \sum_{i=2}^{\bar{n}} \Bigg\{ \left[(1+\delta)(N_{i-1} + \frac{1}{\alpha}N_{\text{FB}_{i-1}})\right]^2 \\
&\quad + N_i^2 + 2\left[(1+\delta)(N_{i-1} + \frac{1}{\alpha}N_{\text{FB}_{i-1}})\right]N_i \Bigg\} \\
&= \alpha^2X_1^2 + 2\alpha^2X_1N_1 + \alpha^2\sum_{i=1}^{\bar{n}} N_i^2 \\
&\quad + 2\alpha^2(1+\delta)\left(\sum_{i=1}^{\bar{n}-1} N_iN_{i+1} + \frac{1}{\alpha}\sum_{i=1}^{\bar{n}-1} N_{\text{FB}_i}N_{i+1}\right) \\
&\quad + \alpha^2(1+\delta)^2\left(\sum_{i=1}^{\bar{n}-1} N_i^2 + \frac{1}{\alpha^2}\sum_{i=1}^{\bar{n}-1} N^2_{\text{FB}_i} + 2\frac{1}{\alpha}\sum_{i=1}^{\bar{n}-1} N_iN_{\text{FB}_i}\right).
\end{aligned}$$

We upper bound $\mathsf{P}\left(\sum_{i=1}^n U_i^2 > nP_{\text{FB}}\right)$ by choosing parameters α and λ such that $nP_{\text{FB}} = \sum_{k=1}^8 C_k$, where:

$$\begin{aligned}
C_1 &= n\lambda P\alpha^2 \\
C_2 &= n2\alpha^2\sqrt{\lambda P}\sigma \\
C_3 &= \bar{n}\alpha^2\sigma^2\left[1 + 2\sqrt{\sigma^2} + 2\sigma^2\right] \\
C_4 &= \alpha^2(1+\delta)\sigma^2\left[(\bar{n}-1) + 2\sqrt{(\bar{n}-1)\bar{n}} + 2\bar{n}\right] \\
C_5 &= \frac{\alpha}{2}(1+\delta)\left(\sigma^2 + \sigma^2_{\text{FB}}\right)\left[(\bar{n}-1) + 2\sqrt{(\bar{n}-1)\bar{n}} + 2\bar{n}\right] \\
C_6 &= \alpha^2(1+\delta)^2\sigma^2\left[(\bar{n}-1) + 2\sqrt{(\bar{n}-1)\bar{n}} + 2\bar{n}\right] \\
C_7 &= (1+\delta)^2\sigma^2_{\text{FB}}\left[(\bar{n}-1) + 2\sqrt{(\bar{n}-1)\bar{n}} + 2\bar{n}\right] \\
C_8 &= \frac{\alpha}{2}(1+\delta)^2\left(\sigma^2 + \sigma^2_{\text{FB}}\right)\left[(\bar{n}-1) + 2\sqrt{(\bar{n}-1)\bar{n}} + 2\bar{n}\right].
\end{aligned}$$

Then,

$$\begin{aligned}
&\mathsf{P}\left(\sum_{i=1}^n U_i^2 > nP_{\text{FB}}\right) \\
&= \mathsf{P}\Bigg(\alpha^2X_1^2 + 2\alpha^2X_1N_1 + \alpha^2\sum_{i=1}^{\bar{n}} N_i^2 + \\
&\quad 2\alpha^2(1+\delta)\sum_{i=1}^{\bar{n}-1} N_iN_{i+1} + 2\alpha(1+\delta)\sum_{i=1}^{\bar{n}-1} N_{\text{FB}_i}N_{i+1} + \\
&\quad \alpha^2(1+\delta)^2\sum_{i=1}^{\bar{n}-1} N_i^2 + (1+\delta)^2\sum_{i=1}^{\bar{n}-1} N^2_{\text{FB}_i} + \\
&\quad 2\alpha(1+\delta)^2\sum_{i=1}^{\bar{n}-1} N_iN_{\text{FB}_i} > \sum_{k=1}^8 C_k\Bigg) \\
&\le \mathsf{P}\left(\alpha^2X_1^2 > C_1\right) + \mathsf{P}\left(2\alpha^2X_1N_1 > C_2\right) \\
&\quad + \mathsf{P}\left(\alpha^2\sum_{i=1}^{\bar{n}} N_i^2 > C_3\right) \\
&\quad + \mathsf{P}\left(2\alpha^2(1+\delta)\sum_{i=1}^{\bar{n}-1} N_iN_{i+1} > C_4\right) \\
&\quad + \mathsf{P}\left(2\alpha(1+\delta)\sum_{i=1}^{\bar{n}-1} N_{\text{FB}_i}N_{i+1} > C_5\right) \\
&\quad + \mathsf{P}\left(\alpha^2(1+\delta)^2\sum_{i=1}^{\bar{n}-1} N_i^2 > C_6\right) \\
&\quad + \mathsf{P}\left((1+\delta)^2\sum_{i=1}^{\bar{n}-1} N^2_{\text{FB}_i} > C_7\right) \\
&\quad + \mathsf{P}\left(2\alpha(1+\delta)^2\sum_{i=1}^{\bar{n}-1} N_iN_{\text{FB}_i} > C_8\right) \\
&\to 0, \quad \text{as } n \to \infty,
\end{aligned}$$

where the last step follows since terms C_k's are appropriately computed / chosen (as in the previous Appendix, making use of the Laurent-Massart bound from [50] for the terms resulting in a χ^2 distribution, and given that $X_1^2 \leq \lambda n P$) to ensure that each term in the sum above vanishes as $n \to \infty$. The parameter α is chosen such that $nP_{\text{FB}} = \sum_{k=1}^{8} C_k = n\lambda P\alpha^2 + n2\alpha^2\sqrt{\lambda P}\sigma + \bar{n}\alpha^2\sigma^2\left[1+2\sqrt{\sigma^2+2\sigma^2}\right] + \left[(\bar{n}-1)+2\sqrt{(\bar{n}-1)\bar{n}}+2\bar{n}\right]K(\alpha)$, where, $\bar{n} = \sqrt{n}$ and $K(\alpha) = (1+\delta)\left[\alpha^2\sigma^2 + \frac{\alpha}{2}\left(\sigma^2+\sigma_{\text{FB}}^2\right)\right] + (1+\delta)^2\left[\alpha^2\sigma^2+\sigma_{\text{FB}}^2+\frac{\alpha}{2}\left(\sigma^2+\sigma_{\text{FB}}^2\right)\right]$, with the choice of $\lambda = \left(1+\sqrt{\frac{\sigma^2+\frac{\sigma_{\text{FB}}^2}{\alpha^2}}{4L^2\sigma_{\text{FB}}^2/\alpha^2}}\right)^{-1}$, and $L = \lfloor M/2 \rfloor$, which leads to (19). $\square$

Appendix D
Proof of Theorem 4 for $M \geq 3$ messages

We generalize this result to any finite number of messages M. First, we show the generalization of the building block operation, followed by its use in the three-stage communication scheme.

A. Building block for general finite $M \geq 3$

This generalization is based on the use of a simplex code of M messages for the transmission and feedback stages, with symbols having energy nP and nP_{FB} respectively. For $M = 3$ the simplex code lives in 2-D space; for general M the constellation lives in $M-1$ dimensional space. For notational convenience, let distances shown in Figure 7, i.e. d' for $M=3$, be denoted as $d'^{(3)}$. Then, the equivalent distance for general M is denoted as $d'^{(M)}$ and obtained from (65). The building block operation for $M > 3$ remains unchanged. Thus, it suffices to prove that the probability of seeing a NACK is still exponentially small.

1) NACK $(M-1)$-volume and probability of NACK event: For M symbols, a simplex code lives in $(M-1)$-dimensional space and the NACK bands become $(M-1)$-dimensional volumes defined as the complement of the union of the ACK volumes: $\bigcup_{i=1}^{M} A_i$, where each A_i is defined by extension of (43) as follows:

Recall Figure 7 used for $M = 3$. To characterize each of the $M-1$ hyperplanes that describe the ACK region A_i around codeword $x^\nu(i)$ of length $\nu = n-(M+1)$ (each hyperplane separating codewords $x^\nu(i)$ and $x^\nu(j)$ for $j \neq i \in \{1,\cdots,M\}$), a point $H_i := (\sqrt{nP}-d_C)\mathbf{u}_i$ ($\mathbf{u}_i$ is the unitary vector in the direction of codeword $x^\nu(i)$, distance d_C is derived geometrically from distance d_A and characterized by parameter t and angle α_C which depends on the geometry of the codewords constellation) and a vector normal to each hyperplane, $\mathbf{w}_{ij} := x^\nu(i) - x^\nu(j)$ are needed.

Extending (43), each hyperplane bounding ACK region A_i is given by: $\{\mathbf{x} \in \mathbb{R}^\nu \mid \langle \mathbf{w}_{ij}\mathbf{x}\rangle = b_i\}$ for $j \neq i \in \{1,\cdots,M\}$, where similarly, $b_i = \langle \mathbf{w}_{ij}, H_i\rangle$. The ACK region A_1 for general M is hence defined as $A_1 = \bigcap_{j=2}^{M}\{y^\nu := \mathbf{y} : \langle \mathbf{w}_{1j}, \mathbf{y}\rangle \geq \langle \mathbf{w}_{1j}, H_1\rangle\}$, and the definitions for the remaining $M-1$ regions follow similarly.

Now assume that $W = 1$ is transmitted, we first prove the NACK event occurs with exponentially small probability. Since $\mathsf{P}_1(\text{NACK}) = 1 - \mathsf{P}_1(\text{ACK})$, we obtain a lower bound on $\mathsf{P}_1(\text{ACK})$ to upper bound $\mathsf{P}_1(\text{NACK})$. An ACK event is declared whenever the received signal y^ν lies within one of the A_i regions:

$$\begin{aligned}\mathsf{P}_1(\text{ACK}) &= \mathsf{P}_1(y^\nu \in \bigcup_{i=1}^{M} A_i)\\ &\geq \mathsf{P}_1(y^\nu \in A_1) \geq 1 - \left[(M-1)Q\left(\frac{d_A^{(M)}}{\sigma}\right)\right],\end{aligned}$$

where $d_A^{(M)}$ denotes the distance between symbol $W=1$ and the NACK volume closest boundary, i.e. $d_A^{(M)} = (1-t)\frac{d'^{(M)}}{2}$. Then,

$$\mathsf{P}_1(y^\nu \in A_1) \geq 1 - \frac{(M-1)}{2}\exp\left\{-n\frac{P}{4\sigma^2}\frac{M}{M-1}(1-t)^2\right\}. \tag{69}$$

Therefore, the upper bound on $\mathsf{P}_1(\text{NACK}) \leq 1 - \mathsf{P}_1(\text{ACK})$ can be written as:

$$\begin{aligned}\mathsf{P}_1(\text{NACK}) &\leq (M-1)Q\left(\frac{d_A^{(M)}}{\sigma}\right)\\ &\leq \frac{M-1}{2}\exp\left(-\frac{(d_A^{(M)})^2}{2\sigma^2}\right)\\ &\leq \frac{M-1}{2}\exp\left(-n\frac{3P}{2\sigma^2}\left(\frac{M}{6(M-1)}\right)(1-t)^2\right),\end{aligned}$$

where the final equality follows from the mapping in (65), and the equivalent distance $d_A^{(M)} = (1-t)\sqrt{3nP}\sqrt{\frac{M}{6(M-1)}}$.

2) Probability of error analysis for the building block for M messages: Once the transmitter has sent codeword $x^\nu(w)$, the receiver determines whether y^ν lies inside the NACK volume or in one of the ACK regions A_i. Then, the signaling of (44) is used to report this result to the transmitter. Depending on the decoding result, "NACK" or an "ACK", the length M retransmission codeword follows exactly as for $M = 3$. Then, a decoding error may occur only when the receiver declares an ACK, and the decoding decision is solely based on y^ν.

$$\begin{aligned}\mathsf{P}_1(\text{error}|\text{ACK}) &= \frac{\mathsf{P}_1(\text{error, ACK})}{\mathsf{P}_1(\text{ACK})}\\ &= \frac{\mathsf{P}_1\left(y^\nu \in \bigcup_{i=2}^{M} A_i\right)}{\mathsf{P}_1\left(y^\nu \in \bigcup_{i=1}^{M} A_i\right)} \leq \frac{\sum_{i=2}^{M}\mathsf{P}_1\left(y^\nu \in A_i\right)}{\mathsf{P}_1\left(y^\nu \in A_1\right)},\end{aligned}$$

where the last inequality comes from the lower bound $\mathsf{P}_1\left(y^\nu \in \bigcup_{i=1}^{M} A_i\right) \geq \mathsf{P}_1\left(y^\nu \in A_1\right)$ for the denominator and the union bound on the numerator. Next, noting that for all $i \geq 2$, (see Figure 7):

$$\mathsf{P}_1(y^\nu \in A_i) \leq Q\left(\frac{d_B^{(M)}}{\sigma}\right), \tag{70}$$

where $d_B^{(M)}$ is the distance between symbol $W=1$ and the furthest NACK region boundary: $d_B^{(M)} = (1+t)\frac{d'^{(M)}}{2} = (1+t)\sqrt{3nP}\sqrt{\frac{M}{6(M-1)}}$. Then, by (69) and (70), and the Chernoff bound for the Q-function:

$$\mathsf{P}_1(\text{error}|\text{ACK}) \leq \frac{\frac{M-1}{2}\exp\left\{-n\frac{P}{4\sigma^2}\frac{M}{M-1}(1+t)^2\right\}}{1-\frac{M-1}{2}\exp\left\{-n\frac{P}{4\sigma^2}\frac{M}{M-1}(1-t)^2\right\}}. \quad (71)$$

After some algebraic manipulation and for large n, (71) leads to:

$$\mathsf{P}_1(\text{error}|\text{ACK}) \doteq \frac{M-1}{2}\exp\left\{-n\frac{P}{4\sigma^2}\frac{M(1+t)^2}{M-1}\right\}. \quad (72)$$

Since parameter t can be chosen to be very close to 1, the achievable error exponent of the BB may be bounded as

$$E_{\text{BB}}^{\text{EXP}} \geq \frac{P}{\sigma^2}\frac{M}{M-1}. \quad (73)$$

Note that for $M=2$, (73) leads to the result of [19]: $E^{\text{EXP}} \geq \frac{2P}{\sigma^2}$; and for $M=3$, it leads to (53): $E^{\text{EXP}} \geq \frac{3P}{2\sigma^2}$.

B. A three stage transmission scheme based on the building block for M messages

In general, the three stage operation remains unchanged. In the transmission stage, W is transmitted using the BB, where the transmitter uses power $\frac{P-\eta}{\lambda}$ and receiver $\frac{\eta}{\lambda}$. Note that $\mathsf{P}_1(W' \neq W) = \mathsf{P}_2(W' \neq W) = \cdots = \mathsf{P}_w(W' \neq W)$. The receiver estimates W' with a probability of error given by (72):

$$\begin{aligned}&\mathsf{P}_1(W' \neq W)\\ &\quad\leq \frac{M-1}{2}\exp\left\{-\lambda(n-M)\frac{\frac{1}{\lambda}(P-\eta)}{4\sigma^2}\frac{M(1+t)^2}{M-1}\right\}\\ &\quad\doteq \frac{M-1}{2}\exp\left\{-n\frac{(P-\eta)}{4\sigma^2}\frac{M(1+t)^2}{M-1}\right\}.\end{aligned}$$

For the feedback stage, W' is sent to the transmitter using the BB in $(1-\lambda)(n-M)$ channel uses. Here, the receiver uses power $\frac{P_{\text{FB}}-\eta}{1-\lambda}$ and transmitter $\frac{\eta}{1-\lambda}$. By the simplex code geometry, $\mathsf{P}_1(W'' \neq W' \mid W') = \mathsf{P}_2(W'' \neq W' \mid W') = \cdots = \mathsf{P}_w(W'' \neq W' \mid W')$. The feedback transmission yields a probability of error given by (72):

$$\begin{aligned}&\mathsf{P}_1(W'' \neq W' \mid W')\\ &\quad\leq \frac{M-1}{2}\cdot\\ &\qquad\exp\left\{-(1-\lambda)(n-M)\frac{\frac{1}{1-\lambda}(P_{\text{FB}}-\eta)}{4\sigma_{\text{FB}}^2}\frac{M(1+t)^2}{M-1}\right\}\\ &\quad\doteq \frac{M-1}{2}\exp\left\{-n\frac{(P_{\text{FB}}-\eta)}{4\sigma_{\text{FB}}^2}\frac{M(1+t)^2}{M-1}\right\}.\end{aligned}$$

In the retransmission stage the transmitter compares W'' with W, and generates a retransmission codeword of length M based on the index location code directly extended from (47). The final decoding rule follows directly from extending (50) to $M>3$, noting that (57) and (58) still hold.

The probability of error of this scheme is dominated by the events where an incorrect symbol is decoded after the transmission stage, and the feedback stage leads to incorrectly decoding W'' as the true message. It can be shown that: $\mathsf{P}_1(\text{error}) \leq \sum_{k=2}^{M}\mathsf{P}_1(W''=1, W'=k)$. Since $\mathsf{P}_1(W''=1, W'=2) = \mathsf{P}_1(W''=1, W'=3) = \ldots = \mathsf{P}_1(W''=1, W'=M)$, we have

$$\begin{aligned}&\mathsf{P}_1(\text{error})\\ &\leq (M-1)\mathsf{P}(W'=2|W=1)\mathsf{P}(W''=1|W'=2, W=1)\\ &= (M-1)\frac{M-1}{2}\exp\left\{-n\frac{(P-\eta)}{4\sigma^2}\frac{M}{M-1}(1+t)^2\right\}\\ &\quad\cdot\frac{1}{2}\exp\left\{-n\frac{(P_{\text{FB}}-\eta)}{4\sigma_{\text{FB}}^2}\frac{M}{M-1}(1+t)^2\right\}\\ &\leq \frac{(M-1)^2}{4}\cdot\exp\left[-n\frac{M(1+t)^2}{M-1}\left(\frac{P-\eta}{4\sigma^2}+\frac{P_{\text{FB}}-\eta}{4\sigma_{\text{FB}}^2}\right)\right].\end{aligned}$$

To conclude, the following error exponent is achievable for η sufficiently small and t very close to 1:

$$E_{\text{FB}^{\text{EXP}}}^{\text{EXP}}\left(M, P, \sigma^2, P_{\text{FB}}, \sigma_{\text{FB}}^2\right) \geq \frac{M}{M-1}\left(\frac{P}{\sigma^2}+\frac{P_2}{\sigma_{\text{FB}}^2}\right). \quad \square$$

Acknowledgements

The authors wish to thank the anonymous reviewers for their observations and fruitful comments which allowed us to improve the presentation of this article and some of our proofs.

References

[1] R. M. Fano, *Transmission of Information: A Statistical Theory of Communications*. Cambridge, Massachusetts: The M.I.T. Press, 1961.
[2] R. Gallager, *Information Theory and Reliable Communication*. New York, NY: Wiley, 1968.
[3] ——, "A simple derivation of the coding theorem and some applications," *IEEE Transactions on Information Theory*, vol. 11, no. 1, pp. 3–18, January 1965.
[4] C. E. Shannon, "Probability of error for optimal codes in a Gaussian channel," *The Bell System Technical Journal*, vol. 38, no. 3, pp. 611–656, May 1959.
[5] E. R. Berlekamp, "Block coding with noiseless feedback," Thesis, Massachusetts Institute of Technology, 1964.
[6] M. S. Pinsker, "The probability of error in block transmission in a memoryless Gaussian channel with feedback," *Problemy Peredachi Informatsii*, vol. 4, no. 4, pp. 3–19, 1968.
[7] C. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, pp. 8–19, Sept. 1956.
[8] 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.
[9] K. S. Zigangirov, "Upper bounds for the error probability for channels with feedback," *Problemy Peredachi Informatsii*, vol. 6, no. 2, pp. 87–92, 1970.
[10] A. Sahai, S. C. Draper, and M. Gastpar, "Boosting reliability over AWGN networks with average power constraints and noiseless feedback," in *IEEE Transactions on Information Theory*, 2005, pp. 402–406.
[11] M. Agarwal, D. Guo, and M. L. Honig, "Error exponent for Gaussian channels with partial sequential feedback," *IEEE Transactions on Information Theory*, vol. 59, no. 8, pp. 4757–4766, Aug 2013.
[12] A. Ben-Yishai and O. Shayevitz, "Interactive Schemes for the AWGN Channel with Noisy Feedback," *IEEE Transactions on Information Theory*, vol. 63, no. 4, pp. 2409–2427, 2017.
[13] M. V. Burnashev and H. Yamamoto, "Noisy feedback improves the BSC reliability function," in *Proc. IEEE Int. Symp. Inf. Theory*, June 2009, pp. 1501–1505.
[14] ——, "On the reliability function for a BSC with noisy feedback," *Problems of Information Transmission*, vol. 46, no. 2, pp. 103–121, 2010.
[15] M. V. Burnashev, "Data transmission over a discrete channel with feedback: random transmission time," *Problemy Peredachi Informatsii*, vol. 12, no. 4, pp. 10–30, 1976.

[16] B. Nakiboglu and R. G. Gallager, "Error exponents for variable-length block codes with feedback and cost constraints," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 945–963, March 2008.

[17] 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*, 2010, pp. 452–457.

[18] Y. Xiang and Y. H. Kim, "Gaussian channel with noisy feedback and peak energy constraint," *IEEE Transactions on Information Theory*, vol. 59, no. 8, pp. 4746–4756, Aug. 2013.

[19] Y. H. Kim, A. Lapidoth, and T. Weissman, "Error exponents for the Gaussian channel with active noisy feedback," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1223–1236, March 2011.

[20] R. Mirghaderi and A. Goldsmith, "Communication over the Gaussian channel with rate-limited feedback," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sep. 2010, pp. 451–457.

[21] 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, Dec 2013.

[22] Y. Wu, P. Minero, and M. Wigger, "Reliability of the Gaussian broadcast channel with common message and feedback," in *2013 IEEE 14th Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, June 2013, pp. 210–214.

[23] M. Gastpar, A. Lapidoth, Y. Steinberg, and M. Wigger, "Coding schemes and asymptotic capacity for the Gaussian broadcast and interference channels with feedback," *IEEE Transactions on Information Theory*, vol. 60, no. 1, pp. 54–71, Jan 2014.

[24] S. Belhadj Amor, Y. Steinberg, and M. Wigger, "MIMO MAC-BC duality with linear-feedback coding schemes," *IEEE Transactions on Information Theory*, vol. 61, no. 11, pp. 5976–5998, Nov 2015.

[25] A. Bracher and M. Wigger, "Feedback and partial message side-information on the semideterministic broadcast channel," *IEEE Transactions on Information Theory*, vol. 63, no. 8, pp. 5052–5073, Aug 2017.

[26] Y. Wu and M. Wigger, "Coding schemes with rate-limited feedback that improve over the no feedback capacity for a large class of broadcast channels," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 2009–2033, April 2016.

[27] J. Schalkwijk, "A coding scheme for additive noise channels with feedback–ii: Band-limited signals," *IEEE Transactions on Information Theory*, vol. 12, no. 2, pp. 183–189, April 1966.

[28] A. Kramer, "Improving communication reliability by use of an intermittent feedback channel," *IEEE Transactions on Information Theory*, vol. 15, no. 1, pp. 52–60, January 1969.

[29] A. Ben-Yishai and O. Shayevitz, "Interactive Schemes for the AWGN Channel with Noisy Feedback," *IEEE Transactions on Information Theory*, vol. 63, no. 4, pp. 2409–2427, April 2017.

[30] 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.

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

[32] A. Sahai, S. C. Draper, and M. Gastpar, "Boosting reliability over AWGN networks with average power constraints and noiseless feedback," in *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.*, Sep. 2005, pp. 402–406.

[33] K. Palacio-Baus and N. Devroye, "Variable-length coding error exponents for the AWGN channel with noisy feedback at zero-rate," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2019, pp. 2254–2258.

[34] L. Shepp, J. Wolf, A. Wyner, and J. Ziv, "Binary communication over the Gaussian channel using feedback with a peak energy constraint," *IEEE Transactions on Information Theory*, vol. 15, no. 4, pp. 476–478, July 1969.

[35] M. V. Burnashev and H. Yamamoto, "On optimal transmission strategies for channels with noiseless feedback," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2016, pp. 1282–1286.

[36] ——, "On the reliability function for a noisy feedback Gaussian channel: Zero rate," *Problems of Information Transmission*, vol. 48, no. 3, pp. 199–216, 2012.

[37] ——, "On using noisy feedback in a Gaussian channel," *Problems of Information Transmission*, vol. 50, no. 3, pp. 217–231, 2014.

[38] ——, "Noisy feedback improves the Gaussian channel reliability function," in *Proc. IEEE Int. Symp. Inf. Theory*, June 2014, pp. 2554–2558.

[39] ——, "On BSC, noisy feedback and three messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 886–889.

[40] ——, "On the zero-rate error exponent for a BSC with noisy feedback," *Prob. of Inf. Transmission*, vol. 44, no. 3, pp. 198–213, Sep. 2008.

[41] Y. H. Kim, A. Lapidoth, and T. Weissman, "The Gaussian channel with noisy feedback," in *Proc. IEEE Int. Symp. Inf. Theory*, June 2007, pp. 1416–1420.

[42] C. E. Shannon, "Two-way communications channels," in *4th Berkeley Symp. Math. Stat. Prob.*, Chicago, IL, Jun. 1961, pp. 611–644.

[43] Te Han, "A general coding scheme for the two-way channel," *IEEE Transactions on Information Theory*, vol. 30, no. 1, pp. 35–44, 1984.

[44] H. Sato, "Two-user communication channels," *IEEE Trans. Inf. Theory*, vol. 23, 1977.

[45] K. Palacio-Baus and N. Devroye, "Two-Way AWGN Channel Error Exponents at Zero Rate," in *2018 IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 1685–1689.

[46] A. Balakrishnan, "A contribution to the sphere-packing problem of communication theory," *Journal of Mathematical Analysis and Applications*, vol. 3, no. 3, pp. 485–506, Oct 1961.

[47] H. J. Landau and D. Slepian, "On the optimality of the regular simplex code," *The Bell System Technical Journal*, vol. 45, no. 8, pp. 1247–1272, Oct 1966.

[48] H. Cohn, "Packing, coding and ground states," in *Mathematics and Materials*, M. J. Bowick, D. Kinderlehrer, G. Menon, and C. Radin, Eds. American Mathematical Society, Institute for Advanced Study, Society for Industrial and Applied Mathematics, 2017, pp. 59–68.

[49] M. Steiner, "The strong simplex conjecture is false," *IEEE Transactions on Information Theory*, vol. 40, no. 3, pp. 721–731, May 1994.

[50] B. Laurent and P. Massart, "Adaptive estimation of a quadratic functional by model selection," *The Annals of Statistics*, vol. 28, no. 5, pp. 1302–1338, 10 2000.

Kenneth S. Palacio-Baus received the M.Sc. degree from the Department of Electrical and Computer Engineering (ECE) at the University of Illinois at Chicago (UIC) in 2011, supported by a Fulbright scholarship. From 2013 to 2016 he joined the University of Cuenca, Ecuador, as a lecturer and researcher in the areas of Signal Processing and Artificial Intelligence. In 2016 he rejoined the ECE department at UIC and received the UIC College of Engineering Graduate Student Award. He is currently a Ph.D. candidate working under the supervision of Dr. Natasha Devroye. His current research topic focuses on error exponents of two-way channels.

Natasha Devroye is a Professor in the Department of Electrical and Computer Engineering at the University of Illinois at Chicago (UIC), which she joined in January 2009. From July 2007 until July 2008 she was a Lecturer at Harvard University. Dr. Devroye obtained her Ph.D in Engineering Sciences from the School of Engineering and Applied Sciences at Harvard University in 2007, and a Honors B. Eng in Electrical Engineering from McGill University in 2001. Dr. Devroye was a recipient of an NSF CAREER award in 2011 and was named UIC's Researcher of the Year in the "Rising Star" category in 2012. She has been an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS, the IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING and the IEEE TRANSACTIONS ON INFORMATION THEORY. Her research focuses on multi-user information theory and applications to cognitive and software-defined radio, radar, relay, spectrum sharing, zero-error and two-way communication networks.