On the Nonasymptotic Performance of Variable-Length Codes with Noisy Stop Feedback (Extended Version with Proofs)

Johan Östman, Rahul Devassy, Giuseppe Durisi, and Erik G. Ström
Chalmers University of Technology, Gothenburg, Sweden

***Abstract*—We present an upper bound on the error probability achievable using variable-length stop-feedback codes, for a fixed size of the information payload and a given constraint on both the average and the maximum latency. Differently from the bound proposed in Polyanskiy *et al.* (2011), which pertains to the scenario in which the stop signal is sent over a noiseless feedback channel, our bound applies to the practically relevant scenario in which the feedback link is noisy. Through numerical results, we illustrate that, in scenarios in which the desired average latency is small, noise in the feedback link can deteriorate the performance of variable-length stop-feedback codes to the extent that it becomes comparable to that of fixed-length codes without feedback.**

I. Introduction

Variable-length stop-feedback (VLSF) coding schemes, i.e., schemes such as simple automatic repeat request (ARQ) and hybrid ARQ, in which information is transmitted until the reception of a positive acknowledgment, are ubiquitous in modern wirelss communication systems. This is because they offer a simple yet effective way to adapt the transmission rate to the channel conditions.

The question investigated in this paper is whether such schemes are suitable for ultra-reliable low-latency communications (URLLC)—one of the new use cases in next-generation wireless systems (5G). VLSF coding schemes are attractive because, as shown in [1], they approach capacity much faster in the average blocklength than fixed-length codes without feedback. Mathematically, one can show that the *dispersion* of suitably designed VLSF coding schemes is zero (see [1, Thm. 2 and Thm. 3]). However, the result obtained in [1] pertains to the setup in which the feedback latency is ignored, and the feedback link is assumed noiseless. As argued in, e.g., [2], [3] these two assumptions are not suitable for URLLC.

The impact of the feedback-link latency on performance in a URLLC setup has been recently investigated in [4], [5] for the case of noiseless feedback. However, most of the available analyses dealing with the noisy-feedback case are asymptotic in the average blocklength (see, e.g., [6]) and cannot be used to infer design guidelines for URLLC. One exception is [7] where the simulated frame-error rate of an actual VLSF code is used within a theoretical framework that captures the impact of errors in the feedback link.

This work was supported by the Swedish Research Council under grants 2014-6066 and 2016-03293.

As discussed in [8] in the context of error-exponent analyses of variable-length coding schemes with *full feedback* (rather than stop feedback, which is the focus of this paper), one challenge that arises when the feedback is noisy is that the transmitter and the receiver may fall out of synchronization and start operating on different messages. The analysis in [8] shows that one can construct a variable-length coding scheme for the full-feedback case that is robust against noise and synchronization errors. More precisely, the corresponding error exponent, although smaller than the error exponent for the variable-length noiseless full-feedback case [9], is still larger than the sphere-packing bound, which governs the performance of fixed-length coding schemes without feedback.

Contributions: We present a non-asymptotic random-coding upper bound on the error probability achievable using VLSF coding schemes, for a given size of the information payload and for a given constraint on both the average and the maximum latency. Differently from the bound in [1, Thm. 3], our bound pertains to the setup in which both the forward and the feedback link are noisy. Furthermore, our analysis accounts for the feedback latency, and accommodates for the presence of a constraint on the maximum number of transmission rounds.

We use the bound derived in this paper to study the performance of VLSF coding schemes operating over a binary-input AWGN channel, focusing on the URLLC-relevant regime in which both the information payload and the average latency are small. Our results indicate that the presence of noise on the feedback link causes a significant degradation in the performance of VLSF codes. For example, when the size of the information payload is 30 bits and both forward and feedback channels operate at $0\,\mathrm{dB}$ of SNR, the error probability achievable for an average latency of 110 channel uses is 10^{-5} if the feedback link is assumed ideal but only 10^{-2} if errors on the feedback link are accounted for. For such a scenario, the performance of the VLSF code becomes comparable to that of a fixed-blocklength no-feedback coding scheme.

Notation: Upper case letters are used to denote random vectors, e.g., $\boldsymbol{X}$ and their realizations are written in lower case, e.g., $\mathbf{x}$. We use superscripts to denote the concatenation of vectors of equal size, e.g., $\boldsymbol{X}^\nu = [\boldsymbol{X}_1, \dots, \boldsymbol{X}_\nu]$. The distribution of a real Gaussian random variable with mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$. Finally, the expectation operator is denoted by $\mathbb{E}[\cdot]$, $\mathbb{P}[\cdot]$ is used for probabilities, $\mathbb{1}\{\cdot\}$ denotes the

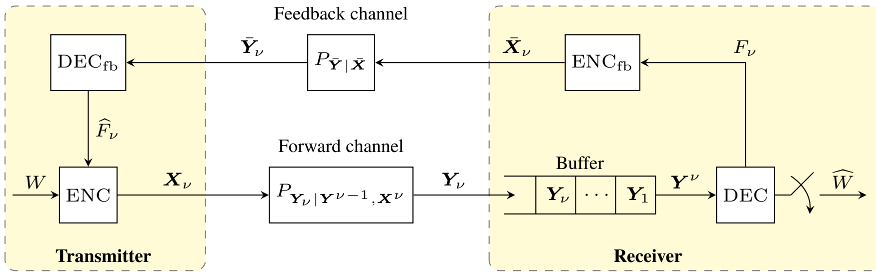


Fig. 1: Round ν of the variable-length noisy stop-feedback transmission scheme. Here W denotes the information message, $\widehat{W}$ is its estimate at the receiver, F_ν is the feedback bit computed at the receiver, and $\widehat{F}_\nu$ its estimate at the transmitter.

indicator function, and $Q(\cdot)$ stands for the Gaussian Q-function.

II. System Model

We consider a point-to-point communication system in which information is transmitted using a VLSF coding scheme. Specifically, transmission occurs over a variable number of rounds, where each round is divided into a data phase and a feedback phase, not necessarily of equal duration. Throughout, we assume that the number of transmission rounds does not exceed the integer $\ell_{\mathrm{m}} < \infty$.

In the data phase, a segment (spanning n channel uses) of the codeword associated to the current information message is sent to the receiver over the forward channel. The forward channel is modeled as a sequence of conditional probability kernels $\{P_{\boldsymbol{Y}_\nu|\boldsymbol{Y}^{\nu-1},\boldsymbol{X}^\nu}\}_{\nu=1}^{\ell_{\mathrm{m}}}$, where the random vectors $\boldsymbol{Y}_\nu$ and $\boldsymbol{X}_\nu$, $\nu = 1,\dots,\ell_{\mathrm{m}}$, take values on the sets $\mathcal{Y}^n$ and $\mathcal{X}^n$, respectively.

At the end of each data phase, the receiver decides whether to perform decoding based on the channel outputs received that far, or to request an additional transmission. The outcome of this decision—a single bit of information conveying the message "stop", which we denote by s or "continue", which we denote by c is transmitted in the feedback phase over the feedback channel using n_{f} channel uses. We model the feedback channel as a sequence of conditional probability kernels $\{P_{\bar{\boldsymbol{Y}}_\nu|\bar{\boldsymbol{Y}}^{\nu-1},\bar{\boldsymbol{X}}^\nu}\}_{\nu=1}^{\ell_{\mathrm{m}}}$, where the random vectors $\bar{\boldsymbol{Y}}_\nu$ and $\bar{\boldsymbol{X}}_\nu$, $\nu = 1,\dots,\ell_{\mathrm{m}}$, take values on the sets $\bar{\mathcal{Y}}^{n_{\mathrm{f}}}$ and $\bar{\mathcal{X}}^{n_{\mathrm{f}}}$, respectively. We assume for simplicity that the feedback channel is block-wise stationary and memoryless, i.e.,

$$P_{\bar{\boldsymbol{Y}}_\nu|\bar{\boldsymbol{Y}}^{\nu-1},\bar{\boldsymbol{X}}^\nu}(\bar{\mathbf{y}}_\nu|\bar{\mathbf{y}}^{\nu-1},\bar{\mathbf{x}}^\nu) = P_{\bar{\boldsymbol{Y}}|\bar{\boldsymbol{X}}}(\bar{\mathbf{y}}_\nu|\bar{\mathbf{x}}_\nu). \tag{1}$$

Upon observing the output of the feedback channel, the transmitter decides whether s or c was sent. This implies that the feedback channel can be viewed as a binary asymmetric channel, with crossover probabilities $p_{\mathrm{sc}} = \mathbb{P}[\mathrm{s} \to \mathrm{c}]$ and $p_{\mathrm{cs}} = \mathbb{P}[\mathrm{c} \to \mathrm{s}]$, which depend both on n_{f} and on the encoder-decoder pair used to transmit the binary message over the feedback channel.

To emphasize that our setup allows for errors in the feedback link, we denote the VLSF coding scheme just described, and depicted in Fig. 1, as variable-length noisy stop-feedback (VLNSF) coding scheme.

Some remarks on our setup are in order. We allow for $p_{\mathrm{sc}} \neq p_{\mathrm{cs}}$ since the $\mathrm{s} \to \mathrm{c}$ and the $\mathrm{c} \to \mathrm{s}$ events have a different impact on performance. Indeed, the $\mathrm{c} \to \mathrm{s}$ event causes the premature stop of the transmission of the current message. We assume, perhaps pessimistically, that this always results in an error at the decoder. This error needs to be handled by higher layers, hence causing a significant latency increase. On the contrary, the $\mathrm{s} \to \mathrm{c}$ event triggers an unnecessary additional transmission round, which causes only a moderate increase in latency. Treating these two error events on the feedback channel differently is in agreement with current wireless standards, where one typically sets $p_{\mathrm{cs}} \ll p_{\mathrm{sc}}$. For example, in long term evolution (LTE), $p_{\mathrm{sc}} = 10^{-2}$ whereas $p_{\mathrm{cs}} \in [10^{-4}, 10^{-3}]$ [10].

Note that an error on the feedback channel may result in the transmitter and the receiver falling out of synchronization, i.e., operating on different messages. To prevent this, we assume that the receiver has also access to an error-free control channel that is used by the transmitter to indicate whether a codeword segment corresponding to the current message or to a new message is to arrive at the receiver. Note that this is again in line with current wireless communication standards such as LTE [10, p. 249]. From a modeling perspective, this is equivalent to assuming that the noisy estimate of the feedback bit produced at the transmitter is known to the receiver.

In our setup, a transmission error occurs if

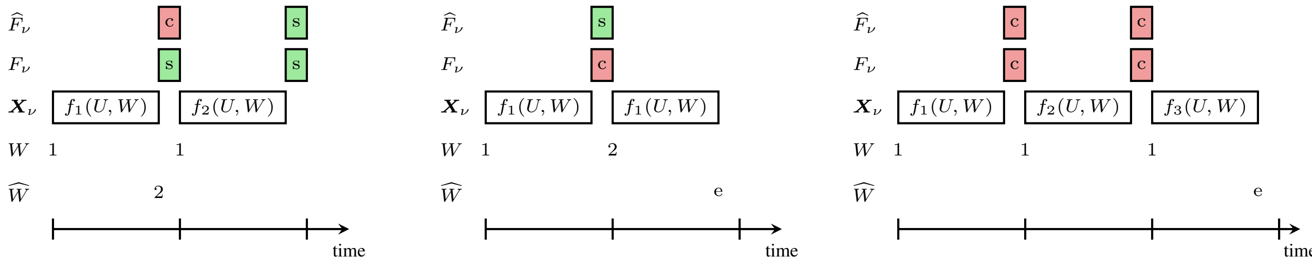
- The receiver decides to perform decoding but produces the wrong codeword estimate—an event typically referred to as *undetected error*. This event is shown in Fig. 2a along with an $\mathrm{s} \to \mathrm{c}$ event, which does not cause an error, but increases latency.
- A $\mathrm{c} \to \mathrm{s}$ event occurs on the feedback channel, see Fig. 2b.
- The receiver is not able to perform decoding within the available ℓ_{m} rounds, see Fig. 2c.

In the last two cases, the decoder declares an erasure, which we denote by the symbol e.

A. Definition of a VLNSF Code

Before providing a formal definition of a VLNSF code, we introduce some additional notation. We let $F_\nu \in \{\mathrm{s}, \mathrm{c}\}$ be the feedback bit generated by the receiver in round $\nu = 1,\dots,\ell_{\mathrm{m}}$ and $\widehat{F}_\nu \in \{\mathrm{s}, \mathrm{c}\}$ its estimate at the transmitter.

We are interested in the trade-off between error probability and average latency, which we define as the average number of rounds needed by the transmitter to complete the processing of the current message. Note that in the presence of errors on the feedback link, the number of rounds after which the receiver produces an estimate of the transmitted message (or declares



(a) The message $W = 1$ is incorrectly decoded as $\widehat{W} = 2$ after the first round; A $\mathrm{s} \to \mathrm{c}$ event causes the retransmission of message $W = 1$.

(b) A $\mathrm{c} \to \mathrm{s}$ event makes the transmitter move to message $W = 2$ after the first round. The receiver observes a packet out of sequence and declares an erasure.

(c) The receiver declares an erasure after $\ell_\mathrm{m} = 3$ unsuccessful transmission rounds.

Fig. 2: Example of the three types of error events for $\ell_\mathrm{m} = 3$.

an erasure) does not necessarily coincide with the number of transmission rounds (see Fig. 2a). Our definition of latency is relevant in a full-buffer situation, where the long-term throughput is defined as the ratio between the number of bits necessary to represent a message, and the average latency defined as above, as a consequence of the renewal-reward theorem [11, Th. 10.5.1].

The definition of a VLNSF coding scheme provided below is an adaptation to the noisy feedback case of the definition of a VLSF coding scheme given in [1].

Definition 1: An $(\ell_\mathrm{a}, M, \epsilon, \ell_\mathrm{m})$-VLNSF coding scheme where M, ℓ_m are positive integers, ℓ_a is a nonnegative real number, and $\epsilon \in (0, 1)$, consists of

- A random variable U defined on a set $\mathcal{U}$ of cardinality $|\mathcal{U}| \leq 2$ that is revealed to both the transmitter and the receiver before the start of the transmission. This random variable acts as common randomness and allows for the use of randomized encoding and decoding strategies.
- A sequence of ℓ_m encoders for the forward channel $f_\nu : \mathcal{U} \times \{1, \dots, M\} \to \mathcal{X}^n$, $\nu = 1, \dots, \ell_\mathrm{m}$, defining the input
$$\boldsymbol{X}_\nu = f_\nu(U, W) \tag{2}$$
for a given message W, which we assume to be uniformly distributed over $\{1, \dots, M\}$.
- A sequence of ℓ_m decoders for the forward channel $g_\nu : \mathcal{U} \times \mathcal{Y}^{n\nu} \to \{1, \dots, M\}$, $\nu = 1, \dots, \ell_\mathrm{m}$, providing an estimate $g_\nu(U, \boldsymbol{Y}^\nu)$ of the message W.
- A sequence of binary random variables $F_\nu \in \{\mathrm{s}, \mathrm{c}\}$, $\nu = 1, \dots, \ell_\mathrm{m}$, which are the output of a stopping rule defined on the filtration $\sigma(U, \boldsymbol{Y}_1, \dots \boldsymbol{Y}_\nu)$. These random variables are the binary messages transmitted by the receiver on the feedback channel.
- An encoder for the feedback channel $\bar{f} : \{\mathrm{s}, \mathrm{c}\} \to \bar{\mathcal{X}}^{n_\mathrm{f}}$ defining the input $\bar{\boldsymbol{X}}_\nu = \bar{f}(F_\nu)$ at round $\nu = 1, \dots, \ell_\mathrm{m}$.
- A decoder for the feedback channel $\bar{g} : \bar{\mathcal{Y}}^{n_\mathrm{f}} \to \{\mathrm{s}, \mathrm{c}\}$ that produces the estimate $\widehat{F}_\nu = \bar{g}(\bar{\boldsymbol{Y}}_\nu)$ at round ν.
- A stopping time at the transmitter τ_tx and a message estimate $\widehat{W} \in \{1, \dots, M\} \cup \mathrm{e}$ defined through the procedure detailed in Algorithm 1, which satisfy the average latency constraint
$$\mathbb{E}[\tau_\mathrm{tx}] \leq \ell_\mathrm{a} \tag{3}$$
and the error probability constraint
$$\mathbb{P}\left[\widehat{W} \neq W\right] \leq \epsilon. \tag{4}$$

Some remarks are in order. Compared to the definition of VLSF codes provided in [1], our definition involves two stopping times (see Algorithm 1), one at the transmitter and one at the receiver. This is needed to account for errors on the feedback link. Also, the decoder employs an erasure option, which is used if a $\mathrm{c} \to \mathrm{s}$ event occurs, or if the stopping rule is not triggered after ℓ_m rounds. Note that we measure the latency in transmission rounds. Each transmission round involves n channel uses on the forward channel and n_f channel uses on the feedback channel.

III. Main Result

We provide an achievability bound, i.e., an upper bound on the error probability achievable using a VLNSF code (see Definition 1), for a fixed number of messages M, a fixed average latency ℓ_a, and a fixed maximum latency ℓ_m.

Before presenting our bound, we characterize the pairs $(p_\mathrm{sc} = \mathbb{P}[\mathrm{s} \to \mathrm{c}], p_\mathrm{cs} = \mathbb{P}[\mathrm{c} \to \mathrm{s}])$ that are achievable for a given choice of the encoder for the feedback channel.

Lemma 1: For a given n_f and for a given encoder $\bar{f} : \{\mathrm{s}, \mathrm{c}\} \to \bar{\mathcal{X}}^{n_\mathrm{f}}$ for the feedback channel, all pairs $(p_\mathrm{sc}, p_\mathrm{cs})$ in the convex hull of the union on the following two sets are achievable
$$\bigcup_{\gamma_\mathrm{f} \in \mathbb{R} \cup \{\pm\infty\}} \left(P^{(\mathrm{s})}\left[\frac{\mathrm{d}P^{(\mathrm{c})}}{\mathrm{d}P^{(\mathrm{s})}} > \gamma_\mathrm{f}\right], P^{(\mathrm{c})}\left[\frac{\mathrm{d}P^{(\mathrm{c})}}{\mathrm{d}P^{(\mathrm{s})}} \leq \gamma_\mathrm{f}\right] \right) \tag{5}$$
and
$$\bigcup_{\gamma_\mathrm{f} \in \mathbb{R} \cup \{\pm\infty\}} \left(P^{(\mathrm{s})}\left[\frac{\mathrm{d}P^{(\mathrm{c})}}{\mathrm{d}P^{(\mathrm{s})}} \geq \gamma_\mathrm{f}\right], P^{(\mathrm{c})}\left[\frac{\mathrm{d}P^{(\mathrm{c})}}{\mathrm{d}P^{(\mathrm{s})}} < \gamma_\mathrm{f}\right] \right) \tag{6}$$
where $P^{(\mathrm{c})} = P_{\bar{\boldsymbol{Y}} | \bar{\boldsymbol{X}} = \bar{f}(\mathrm{c})}$ and $P^{(\mathrm{s})} = P_{\bar{\boldsymbol{Y}} | \bar{\boldsymbol{X}} = \bar{f}(\mathrm{s})}$.

Proof: The result follows from a direct application of the Neyman-Pearson lemma [12]. ■

We present next our achievability bound, which generalizes [1, Thm. 3] to the case of noisy feedback and finite maximum number of transmission rounds.

Theorem 1: Let $(p_\mathrm{sc}, p_\mathrm{cs})$ be an achievable pair according to Lemma 1 for a given choice of n_f and encoder for the feedback channel. Assume that $0 \leq p_\mathrm{sc} + p_\mathrm{cs} \leq 1$. Fix ℓ_m, n and a scalar $\gamma_\mathrm{dec} > 0$. Let $(\boldsymbol{X}_1, \boldsymbol{X}_2, \dots)$ be an arbitrary stochastic

Algorithm 1 Procedure at the transmitter and the receiver to compute the message estimate $\widehat{W}$, the transmitter stopping time τ_{tx} and the receiver stopping time τ_{rx}.

```
Initialize:
    τ_tx = τ_rx = ∞;  F_0 = F̂_0 = c;
for ν = 1 → ℓ_m do
    Transmitter:
    if ν > 1 then
        compute F̂_{ν−1} = ḡ(Ȳ_{ν−1})
    end if
    if F̂_{ν−1} = c then
        transmit f_ν(U, M) over the forward channel
        if ν = ℓ_m then
            set τ_tx = ν
        end if
    else
        set τ_tx = ν − 1
    end if
    Receiver:
    if F_{ν−1} = s then
        if F̂_{ν−1} = c then
            set F_ν = s
        else
            STOP
        end if
    else
        if F̂_{ν−1} = s then
            set Ŵ = e, τ_rx = ν
            STOP
        else
            use stopping rule to compute F_ν
            if F_ν = s then
                set Ŵ = g_ν(Y^ν, U), τ_rx = ν
            end if
        end if
    end if
    if ν < ℓ_m then
        send f̄(F_ν) on the feedback channel
    else
        if τ_rx = ∞ then
            set Ŵ = e, τ_tx = ℓ_m
        end if
    end if
end for
```

process where $\boldsymbol{X}_\nu \in \mathcal{X}^n$ for every nonnegative integer ν. Define a probability space with distribution

$$P_{\boldsymbol{X}^\nu,\boldsymbol{Y}^\nu,\widetilde{\boldsymbol{X}}^\nu}(\mathbf{x}^\nu,\mathbf{y}^\nu,\widetilde{\mathbf{x}}^\nu) = P_{\boldsymbol{X}^\nu}(\mathbf{x}^\nu)\, P_{\boldsymbol{X}^\nu}(\widetilde{\mathbf{x}}^\nu) \prod_{i=1}^{\nu} P_{\boldsymbol{Y}_i|\boldsymbol{X}^i,\boldsymbol{Y}^{i-1}}\left(\mathbf{y}_i|\mathbf{x}^i,\mathbf{y}^{i-1}\right), \quad (7)$$

a sequence of information density functions $\mathcal{X}^{\nu n}\times\mathcal{Y}^{\nu n}\to\mathbb{R}$

$$\imath_\nu(\mathbf{x}^\nu,\mathbf{y}^\nu) = \log\frac{\mathrm{d}P_{\boldsymbol{Y}^\nu|\boldsymbol{X}^\nu}(\mathbf{y}^\nu|\mathbf{x}^\nu)}{\mathrm{d}P_{\boldsymbol{Y}^\nu}(\mathbf{y}^\nu)}, \quad \nu = 1,2,\dots \quad (8)$$

and two stopping times

$$\tau = \inf\{\nu \ge 1 : \imath_\nu(\boldsymbol{X}^\nu,\boldsymbol{Y}^\nu) \ge \gamma_{\mathrm{dec}}\}, \quad (9)$$

$$\widetilde{\tau} = \inf\{\nu \ge 1 : \imath_\nu\left(\widetilde{\boldsymbol{X}}^\nu,\boldsymbol{Y}^\nu\right) \ge \gamma_{\mathrm{dec}}\}. \quad (10)$$

Then, there exists an $(\ell_{\mathrm{a}}, M, \epsilon, \ell_{\mathrm{m}})$-VLNSF code with

$$\ell_{\mathrm{a}} \le \sum_{\nu=0}^{\ell_{\mathrm{m}}-1} (H_{\nu+1} - H_\nu)\,\mathbb{P}[\tau > \nu], \quad (11)$$

$$\epsilon \le \sum_{\nu=1}^{\ell_{\mathrm{m}}} \xi_\nu \Big(\alpha_\nu \mathbb{P}[\tau > \nu] + (M-1)\,\mathbb{P}[\tau \ge \nu, \widetilde{\tau} = \nu]\Big) \quad (12)$$

where $\alpha_\nu = p_{\mathrm{cs}}$ for $\nu = 1,\dots,\ell_{\mathrm{m}}-1$ and $\alpha_{\ell_{\mathrm{m}}} = 1$; furthermore, $\xi_\nu = (1-p_{\mathrm{cs}})^{\nu-1}$ and

$$H_\nu = \sum_{k=1}^{\nu-1} k\xi_k p_{\mathrm{cs}} + \xi_\nu \left[\sum_{k=\nu}^{\ell_{\mathrm{m}}-1} k p_{\mathrm{sc}}^{k-\nu}(1-p_{\mathrm{sc}}) + \ell_{\mathrm{m}} p_{\mathrm{sc}}^{\ell_{\mathrm{m}}-\nu}\right] \quad (13)$$

for $\nu = 1,\dots,\ell_{\mathrm{m}}$, whereas $H_0 = 0$.

Proof: See Appendix A. ■

Some remarks about our achievability bound are in order. As discussed in Appendix A, our bound is based on a decoder that tracks the accumulated information density between each codeword and the received signal, and triggers the stopping rule whenever the accumulated information density exceeds the threshold γ_{dec}. The random variable τ in (9) denotes the first round in which the information density corresponding to the desired codeword exceeds the threshold, whereas $\widetilde{\tau}$ in (10) denotes the first round in which a codeword different from the transmitted one exceeds the threshold. Clearly, the event $\tau > \widetilde{\tau}$ will correspond to an undetected error, provided that $\widetilde{\tau} \le \ell_{\mathrm{m}}$ and no $\mathrm{c}\to\mathrm{s}$ error has occurred in the previous rounds. This is captured by the second term in the error-probability bound (12). The first term in (12) captures instead the error resulting from a $\mathrm{c}\to\mathrm{s}$ event. Note that one can recover [1, Thm. 3] from Theorem 1 by setting $p_{\mathrm{sc}} = p_{\mathrm{cs}} = 0$ and letting $\ell_{\mathrm{m}} \to \infty$.

IV. Numerical Results

We show in this section how one can use the achievability bound in Theorem 1 to design a short-packet transmission system operating over a wireless channel. Although our framework is general, we shall consider for simplicity the case in which both the forward and the feedback channel are real-valued binary-input AWGN channels with same SNR.

We assume that the additive noise has unit variance and that each transmit symbol belongs to the alphabet $\{-\sqrt{\rho}, \sqrt{\rho}\}$, where ρ denotes the SNR. We also assume that the feedback channel assigns the n_{f} dimensional vector $[\sqrt{\rho},\dots,\sqrt{\rho}]$ to the message s and $[-\sqrt{\rho},\dots,-\sqrt{\rho}]$ to c. Under these assumptions,

$$p_{\mathrm{sc}} = Q(\sqrt{n_{\mathrm{f}}\rho} + \gamma_{\mathrm{f}}) \quad (14)$$

$$p_{\mathrm{cs}} = Q(\sqrt{n_{\mathrm{f}}\rho} - \gamma_{\mathrm{f}}). \quad (15)$$

The bound in Theorem 1 is evaluated for a stationary memoryless input process with marginal distribution uniform over $\{-\sqrt{\rho}, \sqrt{\rho}\}$. For such a distribution, (8) reduces to

$$\imath_\nu(\boldsymbol{X}^\nu, \boldsymbol{Y}^\nu) \sim \sum_{i=1}^{\nu n} \log 2 - \log(1 + \exp(-2Z_i)) \quad (16)$$

where $Z_i \sim \mathcal{N}(\rho, \rho)$. Since evaluating (12) directly is challenging, we use the following upper bound

$$\begin{aligned} \mathbb{P}[\tau \geq \nu, \widetilde{\tau} = \nu] &\leq \mathbb{P}[\widetilde{\tau} = \nu] && (17)\\ &= \mathbb{E}[\exp(-\imath_\nu(\boldsymbol{X}^\nu, \boldsymbol{Y}^\nu))\, \mathbb{1}\{\tau = \nu\}]. && (18) \end{aligned}$$

The last step follows from a change of measure (see [1, Eq. (110)]).

In our numerical simulations, we fix the number of channel uses per transmission round $n_{\mathrm{tot}} = n + n_{\mathrm{f}}$ and the maximum number of transmission rounds ℓ_{m}. For a given number of information bits $\log_2 M$, we use Theorem 1 to obtain an upper bound on the error probability ϵ achievable for a given *average* number of transmission rounds ℓ_{a}, and, hence, for a given average latency $\ell_{\mathrm{a}} n_{\mathrm{tot}}$. The error probability is minimized over the Neyman-Pearson threshold γ_{f} in Lemma 1, and over the number of feedback symbols n_{f}, under the constraint that $n + n_{\mathrm{f}} = n_{\mathrm{tot}}$.

In Fig. 3, we assume a maximum latency of $\ell_{\mathrm{m}} n_{\mathrm{tot}} = 400$ channel uses. We consider both the case $n_{\mathrm{tot}} = 25$, for which the maximum number ℓ_{m} of transmission rounds is 16 and the case $n_{\mathrm{tot}} = 50$, for which $\ell_{\mathrm{m}} = 8$. For comparison, we plot the error probability achievable for the case of noiseless feedback, for which the choice $n_{\mathrm{f}} = 1$ is optimal. We also illustrate an upper bound on the error probability obtained using a fixed-length code (no feedback) of blocklength $\ell_{\mathrm{m}} n_{\mathrm{tot}}$. Specifically, we use the bound provided in [13, Eq. (95)].

Our results illustrate the deleterious effect of noise on the feedback channel on the performance of VLSF coding schemes. Consider for example a target error probability $\epsilon = 10^{-5}$. For the case of a noiseless feedback link, the minimum average latency obtainable with a VLSF scheme is 111 channel uses when $n_{\mathrm{tot}} = 25$ and 117 channel uses when $n_{\mathrm{tot}} = 50$. When noise in the feedback link is accounted for, though, the average latency increases to 141 and 153 channel uses, respectively. These values are similar to the blocklength required by a fixed-length coding scheme to operate at $\epsilon = 10^{-5}$, which is 149 channel uses. The performance degradation of the VLSF coding scheme is caused by the resources that need to be allocated to the feedback link to decrease the frequency of $\mathrm{c} \to \mathrm{s}$ and $\mathrm{s} \to \mathrm{c}$ errors.

Specifically, to achieve $\epsilon = 10^{-5}$ with minimum average latency for the case $n_{\mathrm{tot}} = 25$, one needs to set $n_{\mathrm{f}} = 8$ and $\gamma_{\mathrm{f}} = -2.13$. This results in $p_{\mathrm{sc}} = 0.24$ and $p_{\mathrm{cs}} = 3.55 \times 10^{-7}$. For the case $n_{\mathrm{tot}} = 50$, one needs to set $n_{\mathrm{f}} = 10$ and $\gamma_{\mathrm{f}} = -1.65$, which results in $p_{\mathrm{sc}} = 0.065$ and $p_{\mathrm{cs}} = 7.46 \times 10^{-7}$. Note that in both cases, the threshold γ_{f} is chosen so that the $\mathrm{c} \to \mathrm{s}$ event occurs with much smaller probability than the $\mathrm{s} \to \mathrm{c}$ event. Indeed, since in our setup the $\mathrm{c} \to \mathrm{s}$ event results in an error, its probability must always be smaller than the target error probability 10^{-5}.

Observe that noise in the feedback link is actually helpful for

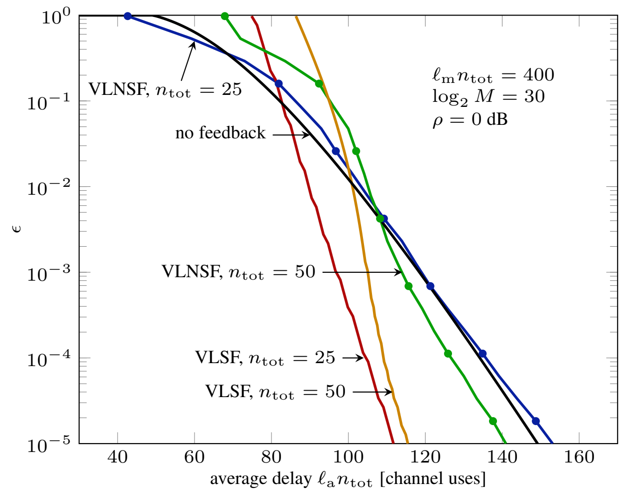


Fig. 3: VLNSF performance over the real–input binary AWGN channel for different resource allocations. The VLNSF curves are optimized over n_{f} and γ_{f} and $n_{\mathrm{f}} = 1$ for the VLSF curves.

the type of decoder used in the achievability bound when the average latency is small and the error probability is high. Indeed, in this regime the occurrence of $\mathrm{c} \to \mathrm{s}$ errors allows one to set a higher decoding threshold γ_{dec} compared to the noiseless feedback case, which reduces undetected errors.

V. Conclusion

We have generalized the achievability bound for VLSF coding schemes presented in [1, Thm. 3] to the case in which the feedback channel is noisy. Our numerical results for the bi-AWGN channel, illustrate that, in some regimes, the performance gain of VLSF coding schemes over fixed length transmission is reduced significantly once the assumption of a noiseless feedback link is dropped.

Appendix A
Proof of Theorem 1

Similar to [1, Thm. 3], we define a random variable U on the set[1]

$$\mathcal{U} = \underbrace{\mathcal{X}^\infty \times \cdots \times \mathcal{X}^\infty}_{M \text{times}} \quad (19)$$

with probability mass function

$$P_U = \underbrace{P_{\boldsymbol{X}^\infty} \times \cdots \times P_{\boldsymbol{X}^\infty}}_{M \text{times}} \quad (20)$$

where $P_{\boldsymbol{X}^\infty}$ denotes the distribution of the process $\{\boldsymbol{X}_1, \boldsymbol{X}_2, \dots\}$. Each realization of U produces M infinite-dimensional codewords $[\boldsymbol{C}_1(w), \boldsymbol{C}_2(w), \dots]$, $w = 1, \dots, M$ where each codeword segment $\boldsymbol{C}_\nu(w)$ belongs to $\mathcal{X}^n$, $n = 1, 2, \dots$ The encoder f_ν maps the message w to the codeword segment $\boldsymbol{C}_\nu(w)$. As detailed in Algorithm 1, the

[1]Similar to [14, Section II] (see also [1, Thm. 19]) one can reduce the cardinality of this random variable to 2.

transmitter is also equipped with a stopping rule, which defines a stopping time τ_{tx} as follows:

$$\tau_{\mathrm{tx}} = \min\{\ell_{\mathrm{m}}, \min\{\nu : \hat{F}_\nu = \mathrm{s}\}\}. \tag{21}$$

Here, we use the convention that the minimum of an empty set is ∞.

At the decoding side, we consider the following stopping rule: stop at round ν if $\imath_\nu(\boldsymbol{C}^\nu(w), \boldsymbol{Y}^\nu) \geq \gamma_{\mathrm{dec}}$ for some w. Let now

$$\tau_w = \inf\{\nu : \imath_\nu(\boldsymbol{C}^\nu(w), \boldsymbol{Y}^\nu) \geq \gamma_{\mathrm{dec}}\} \tag{22}$$

and let

$$\tau_{\mathrm{dec}} = \min\{\tau_1, \dots, \tau_M\}\,. \tag{23}$$

Finally, let[2]

$$\tau_{\mathrm{rx}} = \min\{\tau_{\mathrm{dec}}, \tau_{\mathrm{tx}}\} \tag{24}$$

be the stopping time at the decoder. If $\tau_{\mathrm{rx}} = \tau_{\mathrm{dec}}$, the decoder sets $\widehat{W} = \max\{w : \tau_w = \tau_{\mathrm{dec}}\}$. Otherwise it sets $\widehat{W} = \mathrm{e}$. In words, if no codeword results in a threshold crossing or a $\mathrm{c} \to \mathrm{s}$ error occurs, an erasure is declared. Otherwise, the index of the codeword that resulted in a threshold crossing is taken as message estimate. If more than one codeword yields a threshold crossing, the codeword with the largest index is chosen.

We next prove that $\mathbb{E}[\tau_{\mathrm{tx}}]$ can be upper-bounded as in (11). Set $H_0 = 0$ and $H_\nu = \mathbb{E}[\tau_{\mathrm{tx}} | \tau_{\mathrm{dec}} = \nu]$. One can show that for $\nu = 1, \dots, \ell_{\mathrm{m}} - 1$, the conditional expectation H_ν takes the form given in (13), whereas for $\nu \geq \ell_{\mathrm{m}}$

$$H_\nu = \sum_{k=1}^{\ell_{\mathrm{m}}} k(1-p_{\mathrm{cs}})^{k-1} p_{\mathrm{cs}} + \ell_{\mathrm{m}}(1-p_{\mathrm{sc}})^{\ell_{\mathrm{m}}-1}. \tag{25}$$

Note that this quantity does not depend on ν. We next evaluate $\mathbb{E}[\tau_{\mathrm{tx}}]$ as follows

$$\begin{aligned}
\mathbb{E}[\tau_{\mathrm{tx}}] &= \sum_{\nu=1}^{\infty} H_\nu \mathbb{P}[\tau_{\mathrm{dec}} = \nu] && (26)\\
&= \sum_{\nu=1}^{\infty} H_\nu \left(\mathbb{P}[\tau_{\mathrm{dec}} > \nu - 1] - \mathbb{P}[\tau_{\mathrm{dec}} > \nu]\right) && (27)\\
&= \sum_{\nu=0}^{\infty} (H_{\nu+1} - H_\nu) \mathbb{P}[\tau_{\mathrm{dec}} > \nu] && (28)\\
&= \sum_{\nu=1}^{\ell_{\mathrm{m}}-1} (H_{\nu+1} - H_\nu) \mathbb{P}[\tau_{\mathrm{dec}} > \nu]\,. && (29)
\end{aligned}$$

In the last step we used that $H_{\nu+1} = H_\nu$ for all $\nu \geq \ell_{\mathrm{m}}$ as a consequence of (25). Note now that $H_1 > H_0$ by definition and that, for $\nu = 1, \dots, \ell_{\mathrm{m}} - 1$,

$$\begin{aligned}
H_{\nu+1} - H_\nu &= \frac{(1-p_{\mathrm{sc}}-p_{\mathrm{cs}})(1-p_{\mathrm{cs}})^{\nu-1}}{p_{\mathrm{sc}}} \\
&\quad \times \left[\sum_{k=\nu}^{\ell_{\mathrm{m}}-1} k p_{\mathrm{sc}}^{k-\nu}(1-p_{\mathrm{sc}}) + \ell_{\mathrm{m}} p_{\mathrm{sc}}^{\ell_{\mathrm{m}}-\nu} - \nu\right] && (30)\\
&= \frac{(1-p_{\mathrm{sc}}-p_{\mathrm{cs}})(1-p_{\mathrm{cs}})^{\nu-1}\left(1-p_{\mathrm{sc}}^{\ell_{\mathrm{m}}-\nu}\right)}{1-p_{\mathrm{sc}}}. && (31)
\end{aligned}$$

This implies that $H_{\nu+1} - H_\nu \geq 0$ whenever $p_{\mathrm{sc}} + p_{\mathrm{cs}} \leq 1$. To obtain the desired result, we notice that

$$\mathbb{P}[\tau_{\mathrm{dec}} > \nu] \leq \frac{1}{M}\sum_{w=1}^{M} \mathbb{P}[\tau_w > \nu | W = w] = \mathbb{P}[\tau > \nu] \tag{32}$$

where τ is defined in (9).

We now prove (12). First note that

$$\begin{aligned}
\epsilon &= \frac{1}{M}\sum_{w=1}^{M} \mathbb{P}\left[\widehat{W} \neq w | W = w\right] && (33)\\
&\leq \mathbb{P}\left[\widehat{W} \neq 1 | W = 1\right] && (34)\\
&= \sum_{\nu=1}^{\ell_{\mathrm{m}}} \mathbb{P}\left[\tau_{\mathrm{rx}} = \nu, \widehat{W} \neq 1 | W = 1\right]. && (35)
\end{aligned}$$

Next, we decompose each term on the right-hand-side of (35) as

$$\begin{aligned}
&\mathbb{P}\left[\tau_{\mathrm{rx}} = \nu, \widehat{W} \neq 1 | W = 1\right] \\
&= \mathbb{P}[\tau_{\mathrm{tx}} = \nu, \tau_{\mathrm{dec}} > \nu, | W = 1] \\
&\quad + \mathbb{P}\left[\tau_{\mathrm{tx}} \geq \nu, \tau_{\mathrm{dec}} = \nu, \widehat{W} \neq 1 | W = 1\right] && (36)\\
&= \mathbb{P}[\tau_{\mathrm{tx}} = \nu | \tau_{\mathrm{dec}} > \nu, W = 1]\, \mathbb{P}[\tau_{\mathrm{dec}} > \nu | W = 1] \\
&\quad + \mathbb{P}\left[\tau_{\mathrm{tx}} \geq \nu | \tau_{\mathrm{dec}} = \nu, \widehat{W} \neq 1, W = 1\right] \\
&\quad \times \mathbb{P}\left[\tau_{\mathrm{dec}} = \nu, \widehat{W} \neq 1 | W = 1\right]. && (37)
\end{aligned}$$

The first term on the right-hand side of (37) is the probability that an erasure is declared at step ν because of a $\mathrm{c} \to \mathrm{s}$ event or because the maximum number of transmission round is exceeded. The second term on the right-hand side of (37) corresponds to the probability of an undetected error. Observe now that

$$\mathbb{P}[\tau_{\mathrm{dec}} > \nu | W = 1] \leq \mathbb{P}[\tau > \nu]\,. \tag{38}$$

Furthermore,

$$\begin{aligned}
&\mathbb{P}\left[\tau_{\mathrm{dec}} = \nu, \widehat{W} \neq 1 | W = 1\right] \\
&= \mathbb{P}\left[\cup_{m=2}^{M} \{\tau_1 \geq \nu, \tau_m = \nu\} | W = 1\right] && (39)\\
&\leq (M-1)\, \mathbb{P}[\tau_1 \geq \nu, \tau_2 = \nu | W = 1] && (40)\\
&= (M-1)\, \mathbb{P}[\tau \geq \nu, \widetilde{\tau} = \nu] && (41)
\end{aligned}$$

where $\widetilde{\tau}$ is defined in (10). Finally, we have that

$$\mathbb{P}\left[\tau_{\mathrm{tx}} \geq \nu | \tau_{\mathrm{dec}} = \nu, \widehat{W} \neq 1, W = 1\right] = (1-p_{\mathrm{cs}})^{\nu-1} \tag{42}$$

and that

$$\mathbb{P}[\tau_{\mathrm{tx}} = \nu | \tau_{\mathrm{dec}} > \nu, W = 1] = (1-p_{\mathrm{cs}})^{\nu-1} p_{\mathrm{cs}} \tag{43}$$

for $\nu = 1, \dots, \ell_{\mathrm{m}} - 1$, whereas

$$\mathbb{P}[\tau_{\mathrm{tx}} = \nu | \tau_{\mathrm{dec}} > \nu, W = 1] = (1-p_{\mathrm{cs}})^{\nu-1} \tag{44}$$

for $\nu = \ell_{\mathrm{m}}$. We obtain the desired bound by substituting (38), (41), (42), (43), and (44) into (37) and then (37) into (35).

[2] Recall that the decoder is assumed to know the estimate at the transmitter of the feedback bit through a control channel that announces the presence of new packets. This allows the decoder to learn the stopping time τ_{tx}.

References

[1] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Feedback in the non-asymptotic regime," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4903–4925, Aug. 2011.

[2] M. Bennis, M. Debbah, and H. V. Poor, "Ultrareliable and low-latency wireless communication: Tail, risk, and scale," *Proc. IEEE*, vol. 106, no. 10, pp. 1834–1853, Oct. 2018.

[3] H. Shariatmadari, S. Iraji, R. Jantti, P. Popovski, Z. Li, and M. A. Uusitalo, "Fifth-generation control channel design: Achieving ultrareliable low-latency communications," *IEEE Veh. Technol. Mag.*, vol. 13, no. 2, pp. 84–93, Jun. 2018.

[4] J. Östman, R. Devassy, G. C. Ferrante, and G. Durisi, "Low-latency short-packet transmissions: Fixed length or HARQ?" in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018.

[5] A. Avranas, M. Kountouris, and P. Ciblat, "Energy-latency tradeoff in ultra-reliable low-latency communication with retransmissions," *IEEE J. Sel. Areas Commun.*, vol. 11, no. 36, pp. 2475–2485, Nov. 2018.

[6] P. Wu and N. Jindal, "Coding versus ARQ in fading channels: How reliable should the PHY be?" *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3363 – 3374, Dec. 2011.

[7] T. Breddermann, B. Eschbach, and P. Vary, "On the design of hybrid automatic repeat request schemes with unreliable feedback," *IEEE Trans. Commun.*, vol. 62, no. 2, pp. 758–768, 2014.

[8] S. C. Draper and A. Sahai, "Variable-length channel coding with noisy feedback," *Eur. Trans. Telecommun.*, vol. 19, pp. 355–370, Apr. 2008.

[9] M. V. Burnashev, "Data transmission over a discrete channel with feedback. Random transmission time," *Probl. Inf. Transm.*, vol. 12, no. 4, pp. 10–30, Dec. 1976.

[10] E. Dahlman, S. Parkvall, and J. Sköld, *4G LTE/LTE-Advanced for Mobile Broadband*, Burlington, MA, U.S.A., 2011.

[11] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, 3rd ed., Oxford, U.K., 2001.

[12] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Phil. Trans. Roy. Soc. A*, vol. 231, pp. 289–337, Jan. 1933.

[13] T. Erseghe, "Coding in the finite-blocklength regime: Bounds based on Laplace integrals and their asymptotic approximations," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 6854–6883, Dec. 2016.

[14] K. F. Trillingsgaard, W. Yang, G. Durisi, and P. Popovski, "Common-message broadcast channels with feedback in the nonasymptotic regime: Stop feedback," *IEEE Trans. Inf. Theory*, vol. 64, no. 12, pp. 7686–7718, Dec. 2018.