Coding for Additive White Noise Channels With Feedback Corrupted by Quantization or Bounded Noise

Nuno C. Martins, *Member, IEEE*, and Tsachy Weissman, *Senior Member, IEEE*

***Abstract*—We present coding strategies, which are variants of the Schalkwijk–Kailath scheme, for communicating reliably over additive white noise channels in the presence of corrupted feedback. Our framework comprises an additive white forward channel and a feedback link. We consider two types of corruption mechanisms in the feedback link. The first is quantization noise, i.e., the encoder receives the quantized values of the past outputs of the forward channel. The quantization is uniform, memoryless and time invariant. The second corruption mechanism is an arbitrarily distributed additive bounded noise. Here we allow symbol-by-symbol encoding at the input to the feedback link. We propose explicit schemes featuring positive information rate and positive error exponent. If the forward channel is additive white Gaussian (AWGN) then, as the amplitude of the noise at the feedback link decreases to zero, the rate of our schemes converges to the capacity of the channel. Moreover, the probability of error is shown to converge to zero at a doubly exponential rate. If the forward channel is AWGN and the feedback link consists of an additive bounded noise channel, with signal-to-noise ratio (SNR) constrained symbol-by-symbol encoding, then our schemes achieve rates arbitrarily close to capacity, in the limit of high SNR (at the feedback link).**

***Index Terms*—AWGN channel, coding, error exponent, feedback, quantization.**

I. Introduction

THAT noiseless feedback does not increase the capacity of memoryless channels, but can dramatically enhance the reliability and simplicity of the schemes that achieve it, has been well known since Shannon's work [27]. The assumption of noiseless feedback is an idealization often meant to capture communication scenarios where the noise in the backward link is significantly smaller than in the forward channel. In his subsequent work, in the late 1950s, Shannon also investigated communication in the presence of noisy feedback [24], [25]. This was followed by vigorous research in the 1960s, where various coding strategies were derived for making use of noisy feedback. In particular, in the late 1960s, the success of the Schalkwijk–Kailath scheme in [26] also spurred the use of optimal linear quadratic estimation methods in the derivation of coding schemes for reliable transmission with noisy feedback [1], [5], [16], [14], [15], [12]. An attractive feature of these methods is that the analysis of their performance is tractable using second-order statistics, while the coding and decoding relies on simple iterative schemes. The use of feedback in communicating over white Gaussian interference networks has been addressed in [11], while the multiple-access and broadcast channels have been investigated in [19], [7], and [18]. The author of [4] uses optimal control principles to address the multiuser scenario, in the presence of *perfect* feedback. A general control-theoretic framework to feedback capacity is given in [28].

Another very important aspect of the scheme in [26] is that if the forward channel is white and Gaussian, then the probability of error decays as a double exponential function of the block length, as measured in channel uses. However, in contrast with [26], such a double exponential decay of the probability of error was never achieved in the presence of noisy feedback. In fact, it was recently shown in [10] that, in the presence of noisy feedback, no linear encoding (of which the Schalkwijk–Kailath scheme and its variants are special cases) can feature such a double exponential decay property. Notice that [10] does not contradict the result in [8], where it is shown that noisy feedback can be used to obtain a probability of error that decreases as a double exponential function of the coding block duration in seconds. The framework of [8] is specified in continuous time, and it also allows for an arbitrarily large number of channel uses per second, which is possible only with no bandwidth constraints on the feedback channel. In fact, as it is remarked in [8, p. 476], the strategy in [8] presupposes that the number of channel uses per second grows as an exponential function of the coding block duration in seconds. In contrast, the framework adopted in [10], which is also adopted in our paper, is in discrete time, meaning that not more than one channel use is allowed per time step. If we were to cast our formulation in continuous time then this would be equivalent to imposing a minimal duration (in seconds) for each channel use. A unified and comprehensive presentation of the main results that were available until 1998, on communication in the presence of feedback, is given in [17].

It is therefore of primary importance, from both the theoretical and the practical viewpoints, to develop channel coding

Manuscript received September 11, 2006; revised January 22, 2008. Published August 27, 2008 (projected). The work of N. C. Martins was supported by the National Science Foundation NSF Award EECS 0644764 CAREER and by AFOSR Award FA95500810120. The work of T. Weissman was supported by the National Science Foundation Awards 0512140 and 0546535.

N. C. Martins is with the Electrical and Computer Engineering Department and the Institute for Systems Research, University of Maryland, College Park MD20742 USA (e-mail: nmartins@umd.edu).

T. Weissman is with the Information Systems Laboratory in the Department of Electrical Engineering at Stanford University, Stanford, CA 94305-9510 USA, and with the Department of Electrical Engineering at the Technion–Israel Institute of Technology, Technion City, Haifa 3200, Israel (e-mail: tsachy@stanford.edu).

Communicated by G. Kramer, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2008.928289

schemes that, by making use of *noisy* feedback, maintain the simplicity of noiseless feedback schemes while achieving, for a white Gaussian forward channel, probability of error that decays as a double exponential function of the block length, as measured in channel uses. It is the quest for such schemes operating in discrete time, with one channel use per time step, that motivates this paper. Our main contribution is the derivation of simple coding strategies, which are variants of the Schalkwijk–Kailath scheme, for communicating in discrete-time over additive white channels[1] in the presence of corrupted feedback. More specifically, we consider two types of corruption mechanisms in the backward link.

- Quantization noise: the encoder receives the quantized values of the past outputs of the forward channel. The quantization is uniform, memoryless, and time invariant (that is, symbol-by-symbol scalar quantization), with bounded quantization error.
- Additive bounded noise: the noise in the backward link is additive, and has bounded components, but is otherwise arbitrarily distributed. Here we allow symbol-by-symbol encoding at the input to the backward channel.

The coding schemes that we present achieve positive information rate with positive error exponent. In addition, if the forward channel is additive white Gaussian then our schemes are capacity achieving, in the limit of diminishing amplitude of the noise components in the backward link. Furthermore, if the backward link consists of an additive bounded noise channel, with instantaneous encoding, then our schemes are also capacity achieving in the limit of high signal-to-noise ratio (SNR) (in the backward link). We note that the diminishing of the gap to capacity with vanishing noise in the backward link is a desired property, not to be taken for granted in light of the negative results in [10]. In addition, the probability of error of our coding schemes converges to zero as a doubly exponential function of the block length (channel uses), provided that the forward channel is additive, white, and Gaussian. As will be seen in subsequent sections, our analysis of the performance of the suggested schemes is based on elementary linear systems theory. Tools from linear systems theory were also used in [4], for addressing multiuser communication in the presence of noiseless feedback.

The impact of noise in the feedback link on fundamental performance limits and on explicit schemes that attain them has received attention recently. Examples are the papers [22], [3], where the authors study the tradeoff between reliability and delay in coding for discrete memoryless channels with noisy feedback, and suggest concrete coding schemes for this scenario. One of the ideas suggested in our paper is to address bounded noise in the feedback loop via quantization. A similar technique, in the control context, was proposed in [23]. Moreover, the authors of [20] consider the capacity of discrete finite-state channels in the presence of noninvertible maps in the feedback link, such as quantization, while [10] is primarily concerned with the impact of noise in the backward link on the error exponents. Further limits on the usefulness of noisy feedback are discussed in [6] and the Gaussian multiple-access channel (MAC) with imperfect feedback is studied in [13].

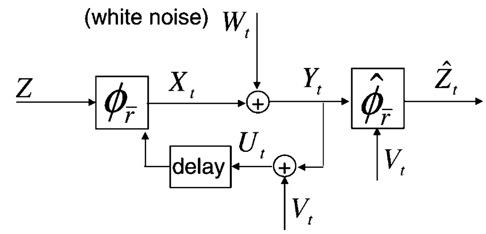

Fig. 1. Basic feedback scheme.

The remainder of this paper is structured as follows. Section II presents preliminary results and definitions, while Section III specifies and analyzes a coding scheme in the presence of feedback corrupted by bounded additive noise, under the assumption that the noise is observable at the decoder. The main results of the paper are presented in Sections IV and V, where we describe and analyze coding schemes for the cases where the backward link features uniform quantization or bounded additive noise, respectively. The paper ends with conclusions in Section VI.

Notation:

- Random variables are represented in upper case letters, such as Z.
- Stochastic processes are indexed by the discrete time variable t, like in X_t. We also use X_0^t to represent $(X_0, \ldots, X_t)$, provided that $t \geq 0$. If t is a negative integer then we adopt the convention that X_0^t is the empty set.
- A realization of a random variable Z is represented by lower case letters, such as z.

II. Preliminary Results and Definitions

In this section, we define and analyze a feedback system whose structure is described by the diagram of Fig. 1. The aforementioned system will be present in the coding schemes proposed in subsequent sections.

For the remainder of this paper, we consider that W_t is a zero mean and white[2] stochastic process of variance σ_W^2 and that Z is a real random variable taking values in $[0, 1]$. In addition, Z and W_0^t are assumed independent for all t. The feedback noise V_t is a real stochastic process with bounded support, whose amplitude has a least upper bound denoted as $\bar{\sigma}_V$, implying that the following holds:

$$\text{Prob}(|V_t| \leq \bar{\sigma}_V) = 1, \quad t \geq 0$$

Other than its boundedness, we impose no restrictions on V_t. In fact, all of our results hold under the sole assumption that V_t has bounded support. In addition, $\bar{\sigma}_V$ is the only *a priori* information about V_t that we use in our proofs. Hence, V_t can be any bounded arbitrarily varying stochastic process and it may be chosen so as it is dependent, or even a function of Z and/or W_0^t. In particular, from Section IV onwards, we focus on the

[1]Here we are not restricted to Gaussian channels. As such, white means that distinct time samples are uncorrelated, but not necessarily independent.

[2]Here we are not restricted to Gaussian channels. As such, white means that distinct time samples are uncorrelated, but not necessarily independent.

specific case where V_t is the quantization error associated with the uniform quantization of the output of the channel.

The remaining signals U_t, Y_t, and $\hat{Z}_t$ are also real stochastic processes. The block represented in Fig. 1 by $\phi_{\bar{r}}$ is a causal operator that maps Z and U_0^{t-1} into X_t for all t. Similarly, $\hat{\phi}_{\bar{r}}$ maps Y_0^t and V_0^t into $\hat{Z}_t$. The description of the maps $\phi_{\bar{r}}$ and $\hat{\phi}_{\bar{r}}$ is given in the following definition.

Definition 2.1: Given a positive real constant $\bar{r}$, the operators $\phi_{\bar{r}} : (t, Z, U_0^{t-1}) \mapsto X_t$ and $\hat{\phi}_{\bar{r}} : (t, Y_0^t, V_0^t) \mapsto \hat{Z}_t$, represented in Fig. 1, are defined as follows:

$$X_t = \phi_{\bar{r}}(t, Z, U_0^{t-1}) \overset{\text{def}}{=} \begin{cases} (2^{-\bar{r}} - 2^{\bar{r}}) \left(\sum_{i=0}^{t-1} 2^{\bar{r}(t-i-1)} U_i + 2^{\bar{r}t} Z \right), & \text{if } t \geq 1 \\ (2^{-\bar{r}} - 2^{\bar{r}}) Z, & \text{if } t = 0 \end{cases} \tag{1}$$

$$\hat{Z}_t = \hat{\phi}_{\bar{r}}\left(t, Y_0^t, V_0^t\right) \overset{\text{def}}{=} \begin{cases} -\sum_{i=0}^{t-1} 2^{-\bar{r}(i+1)} (V_i + Y_i), & \text{if } t \geq 1 \\ 0, & \text{if } t = 0. \end{cases} \tag{2}$$

Notice that (1) has a term, given by $2^{\bar{r}t} Z$, that grows exponentially. However, it should be observed that if the feedback loop is closed (see Fig. 1) by using $U_t = X_t + V_t + W_t$ then X_t is given by

$$X_t = (2^{-\bar{r}} - 2^{\bar{r}}) \left(\sum_{i=0}^{t-1} 2^{-\bar{r}(t-i-1)} (W_i + V_i) + 2^{-\bar{r}t} Z \right), \quad t \geq 1 \tag{3}$$

In order to arrive at (3), we observe that (1) can be equivalently expressed via the following recursion:

$$X_{t+1} = 2^{\bar{r}} X_t + (2^{-\bar{r}} - 2^{\bar{r}}) U_t, \quad t \geq 0 \tag{4}$$

By substituting $U_t = X_t + V_t + W_t$ into (4), we obtain the following:

$$X_{t+1} = 2^{-\bar{r}} X_t + (2^{-\bar{r}} - 2^{\bar{r}}) (V_t + W_t), \quad t \geq 0 \tag{5}$$

which, when expanded as a sum, leads to (3). Notice that (3) specifies a system that is stable, in the sense described in the following proposition.

Proposition 2.1: Let σ_W, $\bar{\sigma}_V$, and $\bar{r}$ be given positive constants and consider the equality expressed in (3). The following holds:

$$E[(X_t)^2] \leq (\sigma_W \sqrt{2^{2\bar{r}} - 1} + \bar{\sigma}_V (2^{\bar{r}} + 1) + 2^{-\bar{r}t} (2^{\bar{r}} - 2^{-\bar{r}}))^2 \tag{6}$$

If we view the feedback loop of Fig. 1, with input to state representation given by (3), as a system with three inputs $(\boldsymbol{V}_t, \boldsymbol{W}_t, \boldsymbol{Z})$, and state $\boldsymbol{X}_t$, then Proposition 2.1 leads to the conclusion that the second moment of the state will not grow unbounded.

Proof of Proposition 2.1: Let S be a random variable with finite second moment. We define the following semi-norm:

$$\|S\|_2 \overset{\text{def}}{=} (E[S^2])^{\frac{1}{2}} \tag{7}$$

Notice that $\|\cdot\|_2$ satisfies the triangle inequality. Hence, we can use (3) to write the following inequality:

$$\|X_t\|_2 \leq (2^{\bar{r}} - 2^{-\bar{r}}) \left(\|\sum_{i=0}^{t-1} 2^{-\bar{r}(t-i-1)} W_i\|_2 + \|\sum_{i=0}^{t-1} 2^{-\bar{r}(t-i-1)} V_i\|_2 + 2^{-\bar{r}t} \|Z\|_2 \right), \quad t \geq 1 \tag{8}$$

Now notice that we can use the fact that $E[(W_i)^2] \leq \sigma_W^2$ and $E[(V_i)^2] \leq \bar{\sigma}_V^2$ to obtain the following:

$$\begin{aligned} (2^{\bar{r}} - 2^{-\bar{r}}) \|\sum_{i=0}^{t-1} 2^{-\bar{r}(t-i-1)} W_i\|_2 &\leq (2^{\bar{r}} - 2^{-\bar{r}}) \sqrt{\sum_{i=0}^{t-1} 2^{-2\bar{r}(t-i-1)} \sigma_W^2} \\ &= \sqrt{(1 - 2^{-2t\bar{r}})(2^{2\bar{r}} - 1)} \sigma_W \end{aligned} \tag{9}$$

$$\begin{aligned} (2^{\bar{r}} - 2^{-\bar{r}}) \|\sum_{i=0}^{t-1} 2^{-\bar{r}(t-i-1)} V_i\|_2 &\leq (2^{\bar{r}} - 2^{-\bar{r}}) \sum_{i=0}^{t-1} 2^{-\bar{r}(t-i-1)} \|V_i\|_2 \\ &\leq (1 - 2^{-\bar{r}t})(2^{\bar{r}} + 1) \bar{\sigma}_V \end{aligned} \tag{10}$$

where the inequality in (9) follows by noticing that W_t is white and zero mean, while (10) is a direct application of the triangle inequality. The proof follows by substitution of (9) and (10) into (8). □

In the absence of noise at the backward link, i.e., $V_t = 0$, (1) and (2) are equivalent to the equations used in the original work by Schalkwijk–Kailath [26]. An alternative minimum variance control interpretation to (1) and (2), in the presence of *perfect* feedback, is given in [4]. The following lemma states a few properties of (1) and (2), motivating their use in the construction of coding schemes.

Lemma 2.2: Let σ_W^2, $\bar{\sigma}_V$ and $\bar{r}$ be given positive real constants. Consider the feedback system of Fig. 1, which is described by (1) and (2) in conjunction with the following equations:

$$Y_t = X_t + W_t \tag{11}$$

$$U_t = Y_t + V_t. \tag{12}$$

The following holds:

$$X_t = 2^{\bar{r}t} (2^{\bar{r}} - 2^{-\bar{r}}) (\hat{Z}_t - Z), t \geq 0 \tag{13}$$

$$E[(X_t)^2] \leq (\sigma_W \sqrt{2^{2\bar{r}} - 1} + \bar{\sigma}_V (2^{\bar{r}} + 1) + 2^{-\bar{r}t} (2^{\bar{r}} - 2^{-\bar{r}}))^2 \tag{14}$$

If W_t is zero-mean, white and Gaussian, with variance σ_W^2, then the following holds:

$$\text{Prob}(|X_t| \geq \alpha) \leq e^{-\frac{(\alpha - \gamma)^2}{2\beta^2}}, \alpha > \gamma, t \geq 0 \tag{15}$$

where γ and β are the following positive real constants:

$$\gamma \overset{\text{def}}{=} (2^{\bar{r}} + 1) \bar{\sigma}_V + 2^{\bar{r}} - 2^{-\bar{r}} \tag{16}$$

$$\beta^2 \overset{\text{def}}{=} (2^{2\bar{r}} - 1) \sigma_W^2. \tag{17}$$

Proof: In order to check the validity of (13), we only need to substitute (2) in (13) and verify that, indeed, we arrive at (3). The validity of (14) follows from Proposition 2.1. In order to prove (15), under the assumption that W_t is zero mean white Gaussian, we define the following auxiliary Gaussian process:

$$\tilde{X}_t = \begin{cases} 0, & \text{if } t = 0 \\ (2^{-\bar{r}} - 2^{\bar{r}}) \sum_{i=0}^{t-1} 2^{-\bar{r}(t-i-1)} W_i, & \text{if } t \geq 1. \end{cases} \quad (18)$$

The following inequality follows from (9):

$$E[(\tilde{X}_t)^2] = (2^{2\bar{r}} - 1)(1 - 2^{-2\bar{r}t})\sigma_W^2 \leq \beta^2 \quad (19)$$

From (10) and the fact that the equality below holds

$$\tilde{X}_t - X_t = (2^{\bar{r}} - 2^{-\bar{r}}) \left(\sum_{i=0}^{t-1} 2^{-\bar{r}(t-i-1)} V_i + 2^{-\bar{r}t} Z \right)$$

we obtain the following:

$$|\tilde{X}_t - X_t| \leq \bar{\sigma}_V (1 - 2^{-\bar{r}t})(2^{\bar{r}} + 1) + 2^{-\bar{r}t}(2^{\bar{r}} - 2^{-\bar{r}}) \leq \gamma \quad (20)$$

Consequently, we arrive at

$$\begin{aligned} \text{Prob}(|X_t| \geq \alpha) &\leq \text{Prob}(|\tilde{X}_t| \geq \alpha - \gamma) \\ &\leq \sqrt{\frac{2}{\pi\beta^2}} \int_{\alpha-\gamma}^{\infty} e^{-\frac{\mu^2}{2\beta^2}} d\mu, \alpha > \gamma \end{aligned} \quad (21)$$

where we used the facts that the inequalities $|\tilde{X}_t - X_t| \leq \gamma$ and $E[(\tilde{X}_t)^2] \leq \beta^2$ hold, as well as the fact that $\tilde{X}_t$ is normally distributed. The derivation of (15) is complete once we use the following upper bound [21, p. 220, eq. (5.1.8)]:

$$\sqrt{\frac{2}{\pi\beta^2}} \int_{\alpha-\gamma}^{\infty} e^{-\frac{\mu^2}{2\beta^2}} d\mu \leq e^{-\frac{(\alpha-\gamma)^2}{2\beta^2}}, \quad \alpha > \gamma \quad (22)$$

□

III. A Coding Scheme With Feedback

In this section, we describe a coding scheme in the presence of feedback according to the framework of Fig. 2, where $\phi_{\bar{r}}$ and $\hat{\phi}_{\bar{r}}$ are defined by (1) and (2), while the maps $\theta_{n,r}$ and $\hat{\theta}_{n,r}$ will be defined below. Notice that the scheme of Fig. 2 assumes that $\hat{\phi}_{\bar{r}}$ has direct access to the feedback noise V_t. Under such an assumption, in this section we construct an efficient and simple coding and decoding scheme that will be used as a basic building block in the rest of the paper. In Section IV, we use the fact that if the backward link is corrupted by uniform quantization then, in fact, V_t is the quantization error, which can be recovered from the output of the forward channel and used as an input to $\hat{\phi}_{\bar{r}}$. Finally, in Section V, we show that bounded noise in the feedback link can be dealt with by using a modification of the quantized feedback framework of Section IV. It should be noted that in the schemes presented in Sections IV and V, the decoder relies solely on the output of the forward channel.

The main result of this section is stated in Theorem 3.2, where we compute a rate of reliable[3] transmission, in bits per channel

[3]By reliable transmission we mean that the probability of error converges to zero with increasing block length n.

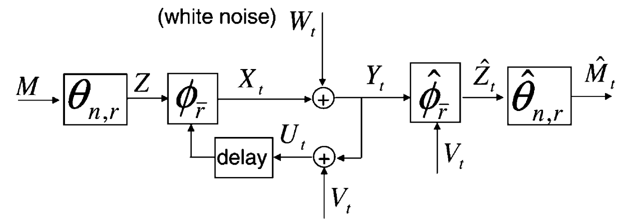

Fig. 2. Basic feedback scheme with encoding and decoding.

use, which is achievable by the scheme of Fig. 2, in the presence of a power constraint at the input of the forward channel. Such a transmission rate is a function of the parameters σ_W^2, $\bar{\sigma}_V$, and it also depends on the forward channel's input power constraint, which we denote as P_X^2. Theorem 3.2 also provides a lower bound on the error exponent of the resulting scheme. If the forward channel is additive, white, and Gaussian then Theorem 3.2 shows that the probability of error of the scheme of Fig. 2 decreases as a doubly exponential function of the block length.

We start with the following definitions of the *ceiling* and *floor* functions denoted by $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$, respectively.

$$\lceil a \rceil \overset{\text{def}}{=} \min\{n \in \mathbb{N} : a \leq n\}, a \in \mathbb{R} \quad (23)$$

$$\lfloor a \rfloor \overset{\text{def}}{=} \max\{n \in \mathbb{N} : a \geq n\}, a \in \mathbb{R}. \quad (24)$$

The following definition specifies the maps $\theta_{n,r}$ and $\hat{\theta}_{n,r}$ represented in Fig. 2.

Definition 3.1: Given a positive integer n, a positive real constant r, a random variable M taking values in the set $\{1, \ldots, 2^{\lfloor rn \rfloor}\}$, and a real stochastic process $\hat{Z}_t$, the following is the definition of the maps $\theta_{n,r} : M \mapsto Z$ and $\hat{\theta}_{n,r} : \hat{Z}_t \mapsto \hat{M}_t$:

$$Z = \theta_{n,r}(M) \overset{\text{def}}{=} \left(M - \frac{1}{2} \right) 2^{-\lfloor rn \rfloor} \quad (25)$$

$$\hat{M}_t = \hat{\theta}_{n,r}(\hat{Z}_t) \overset{\text{def}}{=} \lceil 2^{\lfloor rt \rfloor} \hat{Z}_t \rceil, t \in \{0, \ldots, n\} \quad (26)$$

For the remainder of this paper, n denotes the block length of the coding schemes and r represents a design parameter that quantifies the desired information rate, in bits per channel use. The following equations, describing the coding scheme of Fig. 2, will be used in the statement of Lemma 3.1 and Theorem 3.2:

$$\hat{M}_t = \hat{\theta}_{n,r} \left(\hat{\phi}_{\bar{r}} \left(t, Y_0^t, V_0^t \right) \right) \quad (27)$$

$$Y_t = W_t + \underbrace{\phi_{\bar{r}} \left(t, \theta_{n,r}(M), U_0^{t-1} \right)}_{X_t} \quad (28)$$

$$U_t = Y_t + V_t \quad (29)$$

Lemma 3.1: Let σ_W^2, $\bar{\sigma}_V$, and $\bar{r}$ be given positive real parameters. Consider that the block length is given by a positive integer n, that the desired transmission rate is a positive real number r strictly less than $\bar{r}$, and that M is a random variable arbitrarily distributed in the set $\{1, \ldots, 2^{\lfloor rn \rfloor}\}$. If we adopt the scheme of

Fig. 2, alternatively described by (27)–(29), then the following holds:

$$\text{Prob}(M \neq \hat{M}_n) \leq 4\frac{2^{-2(\bar{r}-r)n}E[(X_n)^2]}{(2^{\bar{r}}-2^{-\bar{r}})^2} \tag{30}$$

If W_t is zero mean, white, and Gaussian with variance σ_W^2 then the following doubly exponential decay, with increasing block size n, of the probability of error holds:

$$\text{Prob}(M \neq \hat{M}_n) \leq e^{-\frac{1}{2\beta^2}\left(2^{-1}(2^{\bar{r}}-2^{-\bar{r}})2^{(\bar{r}-r)n}-\gamma\right)^2}, \quad n > \frac{1}{\bar{r}-r}\log_2\left(\frac{2\gamma}{(2^{\bar{r}}-2^{-\bar{r}})}\right) \tag{31}$$

where γ and β are positive real constants given by (16) and (17), respectively.

Proof: We start by using (25) and (26) and the fact that $2^{\lfloor rn \rfloor}Z$ is in the set $\{\frac{1}{2},\ldots,2^{\lfloor rn \rfloor}-\frac{1}{2}\}$ to conclude the following:

$$\left|2^{\lfloor rn \rfloor}Z - 2^{\lfloor rn \rfloor}\hat{Z}_n\right| < \frac{1}{2} \Longrightarrow M = \hat{M}_n \tag{32}$$

leading to

$$\text{Prob}(M \neq \hat{M}_n) \leq \text{Prob}\left(|Z-\hat{Z}_n| \geq 2^{-(\lfloor rn \rfloor+1)}\right). \tag{33}$$

Using (13), (33), and the fact that $\lfloor rn \rfloor \leq rn$, we get

$$\text{Prob}(M \neq \hat{M}_n) \leq \text{Prob}\left(|X_n| \geq \frac{(2^{\bar{r}}-2^{-\bar{r}})2^{(\bar{r}-r)n}}{2}\right) \tag{34}$$

The inequality (30) follows from Markov's inequality applied to (34). Finally, the inequality (31) follows from (34) and (15). □

A. Lower Bounds on the Achievable Rate of Reliable Transmission in the Presence of a Power Constraint at the Input of the Forward Channel

Below, we define a function that quantifies an achievable rate of reliable transmission for the scheme of Fig. 2, in the presence of a power constraint at the input of the forward channel.

Definition 3.2: For every choice of positive real parameters σ_W^2, P_X^2, and $\bar{\sigma}_V$ satisfying $4\bar{\sigma}_V^2 \leq P_X^2$, define a function $\varrho : (\sigma_W^2, P_X^2, \bar{\sigma}_V) \mapsto \mathbb{R}_{\geq 0}$ as the nonnegative real solution ϱ of the following equation:

$$\sigma_W\sqrt{2^{2\varrho}-1} = P_X - \bar{\sigma}_V(1+2^{\varrho}) \tag{35}$$

If, instead, $4\bar{\sigma}_V^2 > P_X^2$ then $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V) \overset{\Delta}{=} 0$.

It is readily verifiable that a nonnegative real solution of (35), in terms of ϱ, exists and is unique, provided that σ_W^2 and P_X^2 are strictly positive and that $4\bar{\sigma}_V^2$ is less or equal than P_X^2.

Henceforward, we will show that, given P_X, $\bar{\sigma}_V$, and σ_W, the communication scheme proposed at the beginning of this section attains rates of reliable transmission that are arbitrarily close to $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$. Hence, from Definition 3.2, we can also infer that the power inequality $P_X > 2\bar{\sigma}_V$ is a sufficient condition for the existence of a positive rate of reliable transmission. This fact has physical significance since it establishes that, for any given $\bar{\sigma}_V$, there is a minimal critical value of P_X, above which our communication scheme is guaranteed to have positive rate of reliable communication.

Theorem 3.2: Let σ_W^2, P_X^2, and $\bar{\sigma}_V$ be given positive real parameters satisfying $4\bar{\sigma}_V^2 < P_X^2$. In addition, select a positive transmission rate r and a positive real constant $\bar{r}$ satisfying $r < \bar{r} < \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$. For every positive integer block length n, the coding scheme of Fig. 2, alternatively described by (27)–(29), leads to

$$E[(X_t)^2] \leq \left(P_X + \underbrace{2^{-\bar{r}t}(2^{\bar{r}}-2^{-\bar{r}})}_{\text{vanishes with increasing } t}\right)^2, \quad 0 \leq t \leq n \tag{36}$$

$$\text{Prob}(M \neq \hat{M}_n) \leq 4\frac{2^{-2(\bar{r}-r)n}E[(X_n)^2]}{(2^{\bar{r}}-2^{-\bar{r}})^2} \tag{37}$$

where M is a random variable arbitrarily distributed in the set $\{1,\ldots,2^{\lfloor nr \rfloor}\}$. If W_t is zero mean, white, and Gaussian with variance σ_W^2, then the following doubly exponential decay, with increasing block size n, of the probability of error holds:

$$\text{Prob}(M \neq \hat{M}_n) \leq e^{-\frac{1}{2\beta^2}\left(2^{-1}(2^{\bar{r}}-2^{-\bar{r}})2^{(\bar{r}-r)n}-\gamma\right)^2}, \quad n > \frac{1}{\bar{r}-r}\log_2\left(\frac{2\gamma}{(2^{\bar{r}}-2^{-\bar{r}})}\right) \tag{38}$$

where γ and β are positive real constants given by (16) and (17), respectively.

Theorem 3.2 shows that the scheme of Fig. 2, under the constraint that the time average of the second moment of X_t is less or equal[4] than P_X^2, allows for reliable transmission at any rate r strictly less than $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$. Theorem 3.2 also states that any rate of transmission r, if strictly less than $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$, is viable with error exponents that are arbitrarily close to $2[r - \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)]$. In addition, Theorem 3.2 shows that if the forward channel is additive, white ,and Gaussian then the probability of error decreases with the block length n at a doubly exponential rate (see (38)).

Proof of Theorem 3.2: The inequalities (37) and (38) follow directly from (30) and (31), as stated in Lemma 3.1. We now proceed to the derivation of (36), which we start by noticing that the following holds for every positive real $\bar{r}$ strictly less than $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$:

$$\sigma_W\sqrt{2^{2\bar{r}}-1} + \bar{\sigma}_V(2^{\bar{r}}+1) < \sigma_W\sqrt{2^{2\varrho(\sigma_W^2,P_X^2,\bar{\sigma}_V)}-1} + \bar{\sigma}_V(2^{\varrho(\sigma_W^2,P_X^2,\bar{\sigma}_V)}+1), \quad 0 < \bar{r} < \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V). \tag{39}$$

In the derivation of (39), we used the fact that, given positive σ_W, P_X, and $\bar{\sigma}_V$, the function $\eta(r) \overset{\text{def}}{=} \sigma_W\sqrt{2^{2r}-1} + \bar{\sigma}_V(2^r+1)$ is strictly increasing with respect to positive real r. Now, recall

[4]See inequality (36).

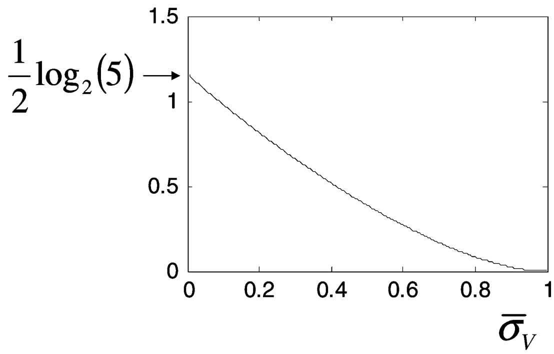

Fig. 3. Plot of $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$ using $\sigma_W^2 = 1$, $P_X^2 = 4$, and $\bar{\sigma}_V \in [0, 1]$.

that since we assume that $4\bar{\sigma}_V^2 < P_X^2$, from Definition 3.2 we conclude that the following equality holds:

$$\sigma_W \sqrt{2^{2\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)} - 1} + \bar{\sigma}_V (2^{\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)} + 1) = P_X \tag{40}$$

which, when substituted into (39), leads to the following inequality:

$$\sigma_W \sqrt{2^{2\bar{r}} - 1} + \bar{\sigma}_V (2^{\bar{r}} + 1) < P_X \tag{41}$$

In order to obtain (36), we only need to substitute (41) into (14), as stated in Lemma 2.2. □

It follows from its definition, as the solution to (35), that $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$ also satisfies the following three properties:

$$\lim_{\bar{\sigma}_V \to 0^+} \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V) = \frac{1}{2}\log_2\left(1 + \frac{P_X^2}{\sigma_W^2}\right), \ \sigma_W^2 > 0, P_X^2 > 0 \tag{42}$$

$$\varrho\left(\sigma_W^2, P_X^2, \frac{P_X}{2}\right) = 0, \quad \sigma_W^2 > 0, P_X^2 > 0 \tag{43}$$

$$\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V) \simeq \log_2\left(\frac{P_X}{\sigma_W + \bar{\sigma}_V}\right), \ P_X^2 \gg \max\{\sigma_W^2, \bar{\sigma}_V^2\} \tag{44}$$

where $\simeq$ indicates that the ratio between the left- and right-hand sides of (44) tends to 1 as $P_X \to \infty$. If W_t is white Gaussian then (42) indicates that in the limit, as the second moment of feedback noise goes to zero, the scheme of Fig. 2 approaches capacity.[5] We have computed $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$ for $\sigma_W^2 = 1$, $P_X^2 = 4$, and one thousand equally spaced values of $\bar{\sigma}_V$, ranging from zero to one and the results are plotted in Fig. 3. The plot illustrates a graceful (continuous) degradation of $\varrho(1, 4, \bar{\sigma}_V)$ as a function of $\bar{\sigma}_V$, going from the highest rate of $\frac{1}{2}\log_2 5$, achieving capacity when W_t is Gaussian, down to zero when $\bar{\sigma}_V = 1$, which is consistent with (42) and (43), respectively.

[5]It is a standard fact [2] that the capacity in bits per channel use of an additive Gaussian channel, with noise variance σ_W^2 and input power constraint P_X^2, is given by $\frac{1}{2}\log_2\left(1 + \frac{P_X^2}{\sigma_W^2}\right)$.

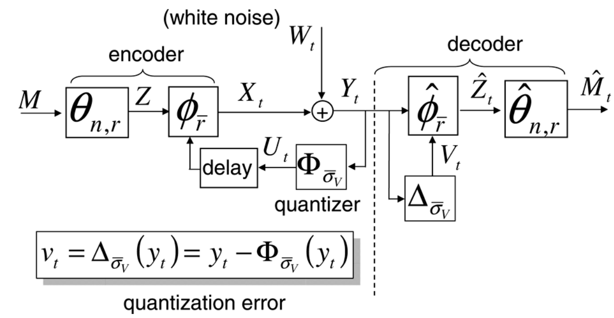

Fig. 4. A coding and decoding scheme in the presence of uniformly quantized feedback.

IV. Specification of a Coding Scheme Using Uniformly Quantized Feedback

In this section, we consider the scheme of Fig. 4, where $\Phi_{\bar{\sigma}_V}$ represents a memoryless uniform quantizer with sensitivity $\bar{\sigma}_V$ and $\Delta_{\bar{\sigma}_V}$ gives the associated quantization error. The main result of this section is Corollary 4.1, where we indicate that the results of Section III hold in the presence of uniformly quantized feedback. Notice that the diagram of Fig. 4 follows from Fig. 2 by adopting V_t as the quantization error, which the decoder reconstructs by making use of $\Delta_{\bar{\sigma}_V}$ applied to the output of the forward channel. The precise definitions of the uniform quantizer $\Phi_{\bar{\sigma}_V}$ and of the quantization error function $\Delta_{\bar{\sigma}_V}$ are given below as follows.

Definition 4.1: Given a positive real parameter b, a uniform quantizer with sensitivity b is a function $\Phi_b : \mathbb{R} \to \mathbb{R}$ defined as

$$\Phi_b(y) = 2b \left\lfloor \frac{y + b}{2b} \right\rfloor. \tag{45}$$

Similarly, the quantization error is given by the following function:

$$\Delta_b(y) = \Phi_b(y) - y, y \in \mathbb{R} \tag{46}$$

which satisfies the following bound:

$$|\Delta_b(y)| \le b, y \in \mathbb{R}. \tag{47}$$

The coding scheme of Fig. 4 can be equivalently expressed by the following equations:[6]

$$\hat{M}_t = \hat{\theta}_{n,r}\left(\hat{\phi}_{\bar{r}}(t, Y_0^t, V_0^t)\right) \tag{48}$$

$$Y_t = W_t + \underbrace{\phi_{\bar{r}}\left(t, \theta_{n,r}(M), U_0^{t-1}\right)}_{X_t} \tag{49}$$

$$U_t = \Phi_{\bar{\sigma}_V}(Y_t) = Y_t + V_t \tag{50}$$

$$V_t = \Delta_{\bar{\sigma}_V}(Y_t). \tag{51}$$

The next corollary follows directly from Theorem 3.2 applied to the scheme of Fig. 4, along with the upper bound (47).

Corollary 4.1: Let σ_W^2, P_X^2, and $\bar{\sigma}_V$ be positive real constants satisfying $4\bar{\sigma}_V^2 < P_X^2$, where $\bar{\sigma}_V$ represents the sensitivity of the

[6]Some of these equations have been used before, but we repeat them here for convenience.

quantizer. In addition, select a positive transmission rate r and a positive real constant $\bar{r}$ satisfying the sequence of inequalities $r < \bar{r} < \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V^2)$. For every positive integer block length n, the coding scheme specified by (48)–(51) (see Fig. 4) leads to

$$E[(X_t)^2] \leq \left(P_X + \underbrace{2^{-\bar{r}t}(2^{\bar{r}} - 2^{-\bar{r}})}_{\text{vanishes with increasing } t} \right)^2, \quad 0 \leq t \leq n \tag{52}$$

$$\text{Prob}(M \neq \hat{M}_n) \leq 4 \frac{2^{-2(\bar{r}-r)n} E[(X_n)^2]}{(2^{\bar{r}} - 2^{-\bar{r}})^2} \tag{53}$$

where M is a random variable arbitrarily distributed in the set $\{1, \ldots, 2^{\lfloor nr \rfloor}\}$. If W_t is zero mean, white, and Gaussian with variance σ_W^2 then the following doubly exponential decay, with increasing block size n, of the probability of error holds:

$$\text{Prob}(M \neq \hat{M}_n) \leq e^{-\frac{1}{2\beta^2}\left(2^{-1}(2^{\bar{r}}-2^{-\bar{r}})2^{(\bar{r}-r)n} - \gamma\right)^2}, \quad n > \frac{1}{\bar{r} - r} \log_2 \left(\frac{2\gamma}{(2^{\bar{r}} - 2^{-\bar{r}})} \right) \tag{54}$$

where γ and β are positive real constants given by (16) and (17), respectively.

Notice that Corollary 4.1 shows that, in the presence of uniformly quantized feedback with sensitivity $\bar{\sigma}_V$, any r strictly less than $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$ is a viable rate of reliable transmission. This implies that the properties (42) and (43), along with the conclusions derived in Section III, hold for uniformly quantized feedback. In particular, the achievable rate of reliable transmission of the coding scheme of Fig. 4 degrades gracefully as a continuous function of the quantizer sensitivity $\bar{\sigma}_V$ (see the numerical example shown in Fig. 3).

V. Coding and Decoding in the Presence of Feedback Corrupted by Bounded Noise

From Corollary 4.1, we conclude that there exist simple explicit coding strategies based on Schalkwijk–Kailath's framework that, even in the presence of *uniformly quantized* feedback, provide positive rates with positive error exponents. In this section, we aim at designing coding schemes in the presence of feedback corrupted by bounded noise. The main result of this section is discussed in Section V-A, where we describe a communication scheme whose structure is that of Fig. 5. In addition, we analyze the performance of such a scheme in the presence of power constraints at the input of the forward and backward channels. The proposed scheme retains the simplicity of the Schalkwijk–Kailath scheme [26], but, in contrast to the original scheme (which breaks down in the presence of noise in the backward link [26, Sec. III-D]), achieves a positive rate of reliable communication and is in fact capacity achieving in the limit of high SNR in the backward link (assuming white Gaussian noise in the forward channel). The scheme proposed in Section V-A also guarantees that, if the forward channel is additive, white, and Gaussian, then the probability of error converges to zero as a doubly exponential function of the block length, as measured in channel uses. The main results of this section are stated in Theorem 5.1.

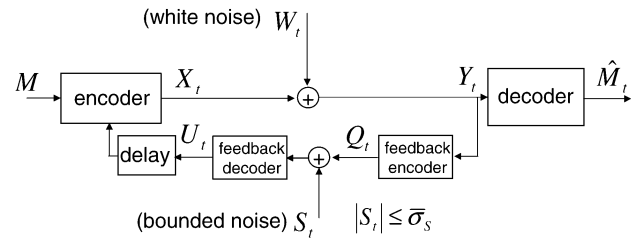

Fig. 5. Communication scheme in the presence of bounded feedback noise.

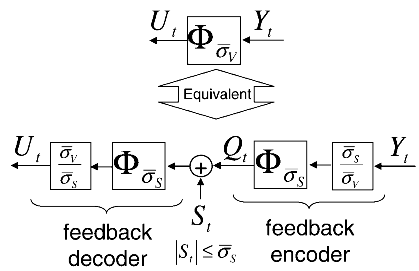

Fig. 6. Schematic representation of the equivalence expressed in Remark 5.1.

A. Performance in the Presence of a Power Constraint at the Input of the Backward Channel

For the remainder of this section, we will define a coding scheme whose structure is that of Fig. 5. The additive noise S_t in the feedback link is arbitrarily distributed, bounded and the tightest upper bound to its amplitude is denoted as $\bar{\sigma}_S$, meaning that the following holds:

$$\text{Prob}(|S_t| \leq \bar{\sigma}_S) = 1, t \geq 0.$$

The following remark will be used in the construction of a coding scheme with the structure of Fig. 5.

Remark 5.1: Let $\bar{\sigma}_S$ be a positive real constant and S_t be a real-valued stochastic process satisfying $|S_t| \leq \bar{\sigma}_S$ with probability one. Given a positive real parameter $\bar{\sigma}_V$, the following holds with probability one:

$$\frac{\bar{\sigma}_V}{\bar{\sigma}_S} \Phi_{\bar{\sigma}_S}(S_t + Q_t) = \Phi_{\bar{\sigma}_V}(Y_t) \tag{55}$$

where Q_t is given by

$$Q_t = \Phi_{\bar{\sigma}_S} \left(\frac{\bar{\sigma}_S}{\bar{\sigma}_V} Y_t \right). \tag{56}$$

The schematic representation of the equivalence expressed in Remark 5.1 is displayed in Fig. 6. In such a scheme, S_t is the bounded additive noise at the backward channel with input Q_t.

Aiming at constructing a coding scheme according to the structure of Fig. 5, we use Remark 5.1 to obtain a new coding strategy by substituting the feedback quantizer $\Phi_{\bar{\sigma}_V}$ of Fig. 4 with the equivalent additive noise channel diagram of Fig. 6. The resulting scheme, along with the encoding and decoding strategy of Section IV, provides a solution to the problem of designing encoders and decoders in the presence of an additive (bounded) noise backward channel (see Fig. 7). Under such a design strategy, $\bar{\sigma}_V$ becomes a design parameter. Notice that viewing $\bar{\sigma}_V$ as a design knob is in contrast with the framework of Section IV, where $\bar{\sigma}_V$ was a given constant.

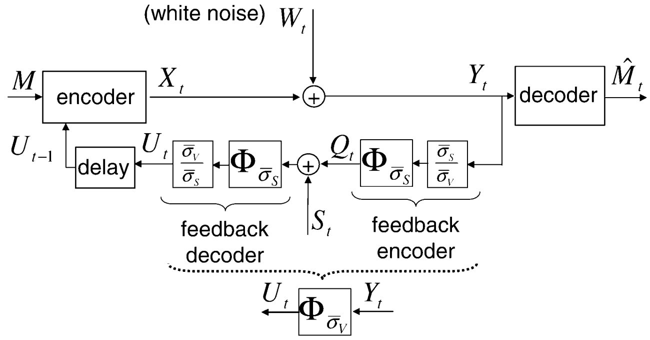

Fig. 7. Proposed coding scheme for dealing with the feedback corruption that results from additive noise S_t in the backward channel. The scheme is constructed by replacing the quantizer in the diagram of Fig. 4 with the scheme of Fig. 6. The encoder and decoder blocks are described in detail in Fig. 4.

Regarding the role of $\bar{\sigma}_V$, we have shown in (42) that as $\bar{\sigma}_V$ approaches zero, the achievable rate of reliable transmission converges to a positive value, which, in the case where W_t is white Gaussian, coincides with capacity. However, for any given positive real $\bar{\sigma}_S$, the smaller $\bar{\sigma}_V$ the larger the scaling constant $\frac{\bar{\sigma}_S}{\bar{\sigma}_V}$ in (56) and that may lead to Q_t having an arbitrarily large second moment. In Theorem 5.1, we show that the function defined below solves the aforementioned problem by providing a suitable choice for $\bar{\sigma}_V$, in the presence of power constraints at the input of the forward and backward channels.

Definition 5.1: Let σ_W, $\bar{\sigma}_S$, P_X, and P_Q be given positive real constants, where P_Q^2 symbolizes a power constraint at the input of the backward channel Q_t. Below, we define the function $\Gamma : \mathbb{R}^4_{\geq 0} \to \mathbb{R}_{\geq 0}$, which we will use as a selection for the design parameter $\bar{\sigma}_V$

$$\Gamma(\sigma_W, \bar{\sigma}_S, P_X, P_Q) = (P_X + \sigma_W)\frac{\bar{\sigma}_S}{P_Q - \bar{\sigma}_S}, P_Q > \bar{\sigma}_S \quad (57)$$

The following Theorem is one of the main results of this paper.

Theorem 5.1: Let σ_W^2, P_X^2, P_Q^2, and $\bar{\sigma}_S$ be positive constants satisfying $4\Gamma(\sigma_W, \bar{\sigma}_S, P_X, P_Q)^2 < P_X^2$ and $\bar{\sigma}_S < P_Q$. In addition, select a positive transmission rate r and a positive real constant $\bar{r}$ satisfying the sequence of inequalities

$$r < \bar{r} < \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)|_{\bar{\sigma}_V = \Gamma(\sigma_W, \bar{\sigma}_S, P_X, P_Q)}.$$

For every positive integer block length n, the coding scheme of Fig. 7, alternatively described by (48)–(51) and (56), leads to

$$E[(X_t)^2] \leq \left(P_X + \underbrace{2^{-\bar{r}t}(2^{\bar{r}} - 2^{-\bar{r}})}_{\text{vanishes with increasing } t} \right)^2, \quad 0 \leq t \leq n \quad (58)$$

$$E[(Q_t)^2] \leq \left(P_Q + \underbrace{2^{-\bar{r}t}\frac{P_Q - \bar{\sigma}_S}{P_X + \sigma_W}(2^{\bar{r}} - 2^{-\bar{r}})}_{\text{vanishes with increasing } t} \right)^2, \ 0 \leq t \leq n \quad (59)$$

$$\text{Prob}(M \neq \hat{M}_n) \leq 4\frac{2^{-2(\bar{r}-r)n}E[(X_n)^2]}{(2^{\bar{r}} - 2^{-\bar{r}})^2} \quad (60)$$

where M is a random variable arbitrarily distributed in the set $\{1, \ldots, 2^{\lfloor nr \rfloor}\}$. If W_t is zero mean, white, and Gaussian with variance σ_W^2 then the following doubly exponential decay, with increasing block size n, of the probability of error holds:

$$\text{Prob}(M \neq \hat{M}_n) \leq e^{-\frac{1}{2\beta^2}(2^{-1}(2^{\bar{r}} - 2^{-\bar{r}})2^{(\bar{r}-r)n} - \gamma)^2}, \quad n > \frac{1}{\bar{r} - r}\log_2\left(\frac{2\gamma}{(2^{\bar{r}} - 2^{-\bar{r}})}\right) \quad (61)$$

where γ and β are positive real constants given by (16) and (17), respectively, where $\bar{\sigma}_V$ is given by the assumed selection $\bar{\sigma}_V = \Gamma(\sigma_W, \bar{\sigma}_S, P_X, P_Q)$.

Proof: The inequalities (58), (60), and (61) follow directly from Corollary 4.1. In order to arrive at (59), we start by noticing that we can use the triangle inequality to find the following inequalities:

$$(E[(Y_t)^2])^{\frac{1}{2}} \leq (E[(X_t)^2])^{\frac{1}{2}} + \sigma_W \quad (62)$$

$$(E[(Q_t)^2])^{\frac{1}{2}} \leq \frac{\bar{\sigma}_S}{\bar{\sigma}_V}(E[(Y_t)^2])^{\frac{1}{2}} + \bar{\sigma}_S. \quad (63)$$

In addition, substitution of (62) in (63), leads to

$$E[(Q_t)^2] \leq \left(\frac{\bar{\sigma}_S}{\bar{\sigma}_V}\left(E[(X_t)^2]^{\frac{1}{2}} + \sigma_W \right) + \bar{\sigma}_S \right)^2 \quad (64)$$

which, from (58), implies the following:

$$E[(Q_t)^2] \leq \left(\frac{\bar{\sigma}_S}{\bar{\sigma}_V}\left(P_X + 2^{-\bar{r}t}\left(2^{\bar{r}} - 2^{-\bar{r}}\right) + \sigma_W \right) + \bar{\sigma}_S \right)^2 \quad (65)$$

The proof is complete since (59) follows by substituting our choice $\bar{\sigma}_V = \Gamma(\sigma_W, \bar{\sigma}_S, P_X, P_Q)$ in (65). □

Under the conditions of Theorem 5.1, including our choice of the design parameter $\bar{\sigma}_V$, the following limit holds:

$$\lim_{\bar{\sigma}_S \to 0^+} \varrho\left(\sigma_W^2, P_X^2, \bar{\sigma}_V\right)\big|_{\bar{\sigma}_V = \Gamma(\sigma_W, \bar{\sigma}_S, P_X, P_Q)} = \frac{1}{2}\log_2\left(1 + \frac{P_X^2}{\sigma_W^2}\right), \quad \sigma_W > 0, P_X > 0, P_Q > 0. \quad (66)$$

Notice that (66) leads to the conclusion that, under our choice of $\bar{\sigma}_V$, the performance of the scheme of Theorem 5.1 (see Fig. 7) degrades gracefully as a function of $\bar{\sigma}_S$, in terms of both the rate and the error exponent. If W_t is white Gaussian then (66) indicates that as $\bar{\sigma}_S$ tends to zero, the scheme of Theorem 5.1 can be used to reliably communicate at a rate arbitrarily close to capacity. Moreover, such a conclusion holds in the presence of an arbitrarily low power constraint at the backward channel. The plot of Fig. 8 displays how the achievable rate changes as a function of $\bar{\sigma}_S$, under the choice $\bar{\sigma}_V = \Gamma(\sigma_W, \bar{\sigma}_S, P_X, P_Q)$. Such a plot also illustrates that by increasing P_Q, we can reduce the sensitivity of the achievable rate of reliable transmission, relative to variations in $\bar{\sigma}_S$.

B. Further Comments on the Location of the One-Step Feedback Delay

In the framework of Fig. 7, the one-step delay block is located *after* the feedback decoder. However, we should stress

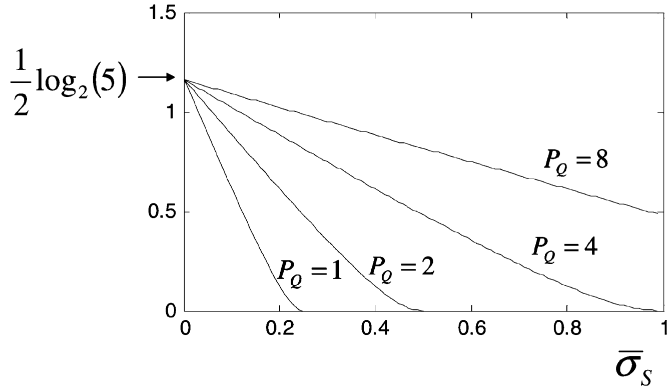

Fig. 8. Plot of $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)|_{\bar{\sigma}_V=\Gamma(\sigma_W,\bar{\sigma}_S,P_X,P_Q)}$ using $\sigma_W^2 = 1, P_X^2 = 4$, and $\bar{\sigma}_S \in [0, 1)$, for P_Q taking values $1, 2, 4$, and 8.

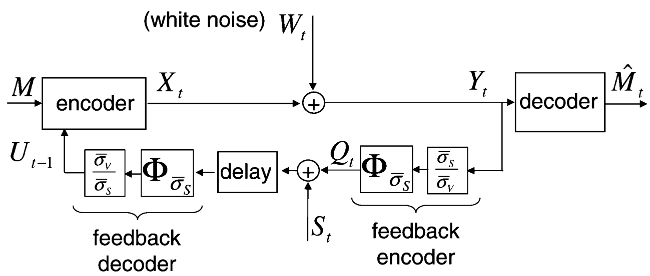

Fig. 9. A coding scheme equivalent to the one described by Fig. 7.

that, since the feedback decoder is time invariant, our coding scheme would be unaltered if we had placed the delay block *before* as indicated in Fig. 9. Indeed, the diagrams of Figs. 7 and 9 are equivalent, implying that Theorem 5.1 holds also for the coding scheme of Fig. 9.

VI. Conclusion

We derived simple schemes for reliable communication over a white noise forward channel, in the presence of corrupted feedback. Both the case of uniform quantization noise and the case of additive bounded noise in the backward link were considered, where, in the latter case, encoding at the input to the backward channel is allowed. The schemes were seen to achieve a positive rate of reliable communication, and in fact be capacity-achieving in the presence of an additive white Gaussian forward channel, in the limit of small noise (or high SNR when encoding is allowed) in the backward link. In addition, still under the assumption that the forward channel is additive white Gaussian, the proposed schemes guarantee that the probability of error converges to zero as a doubly exponential function of the block length.

Our approach to the construction and analysis of coding schemes carries over naturally to the case where the noise in the forward channel is nonwhite. In this case, we expect to obtain variations on the schemes in [9] that are analogous to those in the present work and whose gap to capacity behaves similarly.

References

[1] S. Butman, "A general formulation of linear feedback communication systems with solution," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 3, pp. 392–400, May 1969.

[2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley-Interscience, 1991.

[3] S. C. Draper and A. Sahai, "Noisy feedback improves communication reliability," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, Jul. 2006, pp. 69–73.

[4] N. Elia, "When Bode meets Shannon: Control-oriented feedback communication schemes," *IEEE Trans. Aut. Control*, vol. 49, no. 9, pp. 1477–1488, Sep. 2004.

[5] P. Elias, "Networks of gaussian channels with applications to feedback systems," *IEEE. Trans. Inf. Theory*, vol. IT-13, no. 3, pp. 493–501, Jul. 1967.

[6] M. Gastpar and G. Kramer, "On noisy feedback for interference channels," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Oct/Nov. 2006, pp. 216–220.

[7] M. S. Iacobucci and M. G. Di Benedetto, "A feedback for the multiple access channel (MAC): A case study," in *Proc. IEEE GLOBECOM*, Phoenix, AZ, Nov. 1997, pp. 128–132.

[8] R. L. Kashyap, "Feedback coding schemes for an additive noise channel with a noisy feedback link," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 3, pp. 471–480, May 1968.

[9] Y. H. Kim, "Feedback capacity of stationary Gaussian channels," *IEEE Trans. Inf. Theory* [Online]. Available: http://arxiv.org/abs/cs.IT/0602091, submitted for publication

[10] Y.-H. Kim, A. Lapidoth, and T. Weissman, "The Gaussian channel with noisy feedback," in *Proc. IEEE Int. Symp. Information Theory*, Nice, France, Jul. 2007, pp. 1416–1420.

[11] G. Kramer, "Feedback strategies for white gaussian interference networks," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1423–1438, Jun. 2002.

[12] A. J. Kramer, "Improving communication reliability by use of an intermittent feedback channel," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 1, pp. 52–60, Jan. 1969.

[13] A. Lapidoth and M. A. Wigger, "On the Gaussian MAC with imperfect feedback," in *Proc. 24th Convention of Electrical and Electronics Engineers*, Eilat, Israel, Nov. 2006, pp. 203–207.

[14] S. S. Lavenberg, "Feedback communication using orthogonal signals," *IEEE Trans. Inf. Theory*, vol. IT-15, no. 4, pp. 478–483, Jul. 1969.

[15] S. S. Lavenberg, "Repetitive signaling using a noisy feedback channel," *IEEE Trans. Inf. Theory*, vol. IT-17, no. 3, pp. 269–278, May 1971.

[16] J. K. Omura, "Optimum linear transmission of analog data for channels with feedback," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 1, pp. 38–43, Jan. 1968.

[17] J. Ooi, *Coding for Channels With Feedback*. Boston, MA: Kluwer, 1998.

[18] L. H. Ozarow and S. K. Leung-Yan-Cheong, "An achievable region and outer bound for the gaussian broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 4, pp. 667–671, Jul. 1984.

[19] L. H. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 4, pp. 623–629, Jul. 1984.

[20] H. Permuter, T. Weissman, and A. Goldsmith, Finite State Channels With Time-Invariant Deterministic Feedback. [Online]. Available: http://arxiv.org/abs/cs.IT/0608070

[21] J. G. Proakis and M. Salehi, *Fundamentals of Communication Systems*. Upeer Saddle River, NJ: Prentice-Hall, 2005.

[22] A. Sahai and T. Şimşek, "On the variable-delay reliability function of discrete memoryless channels with access to noisy feedback," in *Proc. IEEE Information Theory Workshop*, San Antonio, TX, 2004, pp. 336–341.

[23] A. Sahai, "Anytime Information Theory," Ph.D. dissertation, Dept. Elec. Eng. Comp. Sci., MIT, Cambridge, MA, Feb. 2001.

[24] C. E. Shannon, "Channels with side information at the transmitter," *IBM J. Res. Develop.*, vol. 2, pp. 289–293, Oct. 1958.

[25] C. E. Shannon, "Two-way communication channels," in *Proc. 4th Berkeley Symp. Mathematical Statistics and Probability*, Berkeley, CA, 1961, pp. 611–644.

[26] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback I: No bandwidth constraint," *IEEE Trans. Inf. Theory*, vol. IT-12, no. 2, pp. 172–182, Apr. 1966.

[27] C. E. Shannon, "The zero-error capacity of a noisy channel," *IRE Trans. Inf. Theory*, vol. 2, pp. 8–19, Sep. 1956.

[28] S. Yang, A. Kavčić, and S. Tatikonda, "Feedback capacity of finite-state machine channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 799–810, Mar. 2005.