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

Nuno C Martins and Tsachy Weissman

Abstract

We present simple coding strategies, which are variants of the Schalkwijk-Kailath scheme, for communicating reliably over additive white noise channels in the presence of corrupted feedback. More specifically, we consider a framework comprising an additive white forward channel and a backward link which is used for feedback. We consider two types of corruption mechanisms in the backward 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 (that is, symbol-by-symbol scalar quantization), with bounded quantization error. The second corruption mechanism is an arbitrarily distributed additive bounded noise in the backward link. Here we allow symbol-by-symbol encoding at the input to the backward channel. We propose simple explicit schemes that guarantee positive information rate, in bits per channel use, with positive error exponent. If the forward channel is additive white Gaussian then our schemes achieve capacity, in the limit of diminishing amplitude of the noise components at the backward link, while guaranteeing that the probability of error converges to zero as a doubly exponential function of the block length. Furthermore, if the forward channel is additive white Gaussian and the backward link consists of an additive bounded noise channel, with signal-to-noise ratio (SNR) constrained symbol-by-symbol encoding, then our schemes are also capacity-achieving in the limit of high SNR.

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, is well known since Shannon's work [10]. 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. However, all the known simple schemes for reliable communication in the presence of feedback rely heavily on the assumption that the feedback is completely noise-free, and break down when noise is introduced into the backward link.

As a case in point, it was recently shown in [5] that *any* feedback scheme with linear encoding (of which the Schalkwijk-Kailath scheme and its variants are special cases) breaks down completely in the presence of additive white noise of arbitrarily small variance in the backward link: not only is it impossible to achieve capacity, but, with such schemes it is impossible to communicate reliably at any positive information rate.

It is therefore of primary importance, from both the theoretical and the practical viewpoints, to develop channel coding schemes that, by making use of *noisy* feedback, maintain the simplicity of noiseless feedback schemes while achieving a positive rate of reliable communication. It is the quest for such schemes that motivates this paper.

Our main contribution is the derivation of simple coding strategies, which are variants of the Schalkwijk-Kailath scheme, for communicating over additive white channels 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 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 [5]. In addition, the probability of error of our coding schemes converges to zero as a doubly exponential function of the block length, 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.

To our knowledge, the impact of noise in the feedback link on fundamental performance limits and on explicit schemes that attain them has hitherto received little attention. Exceptions are the papers [8], [2] which study the trade-off between reliability and delay in coding for discrete memoryless channels with noisy feedback, and suggest concrete coding schemes for this scenario. Another exception is the recent [6], which considers the capacity of discrete finite-state channels in the presence of non-invertible maps in the feedback link, such as quantization. Yet another paper is the aforementioned [5], which is primarily concerned with the impact of noise in the backward link on the error exponents.

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 large caps, such as Z.
- Stochastic processes are indexed by the discrete time variable t, like in X_t. We also use X^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^t is the empty set.
- A realization of a random variable Z is represented in small caps, such as z.
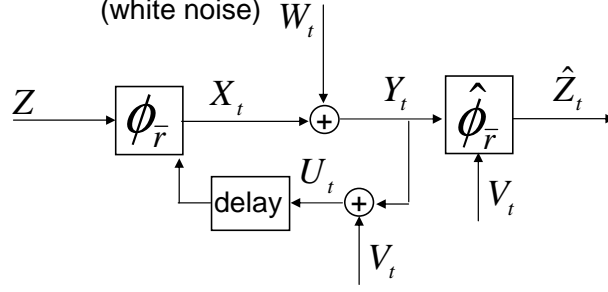
Fig. 1. Basic feedback scheme.

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 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^t are assumed independent for all t. The feedback noise V_t is a bounded real stochastic process whose amplitude has a least upper-bound given by:

$$\bar{\sigma}_V \overset{def}{=} \inf\{\alpha \in \mathbb{R}_{\geq 0} : Prob(|V_t| > \alpha) = 0, t \geq 0\}$$

meaning that the following holds:

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

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 an operator that maps Z and U^{t-1} into X_t for all t. Similarly, $\hat{\phi}_{\bar{r}}$ maps Y^t and V^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 $\phi_{\bar{r}} : (t, Z, U^{t-1}) \mapsto X_t$ and $\hat{\phi}_{\bar{r}} : (t, Y^t, V^t) \mapsto \hat{Z}_t$, represented in Fig 1, are defined as follows:

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

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

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:

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

which describes a system that is stable, in the bounded input implies bounded output sense. In the absence of backward link noise, i.e. $V_t = 0$, (1) and (2) are equivalent to the equations used in the original work by Schalkwijk-Kailath [9]. An alternative minimum variance control interpretation to (1) and (2), in the presence of *perfect* feedback, is given in [3]. In addition, the work by [3] extends Schalkwijk-Kailath's algorithm, with *perfect* feedback, to the multi-user case. A general control theoretic framework to feedback capacity is given in [11]. The following lemma states a few properties of (1) and (2) which motivate their use in the construction of coding schemes.

Lemma 2.1: 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)-(2) in conjunction with the following equations:

$$Y_t = X_t + W_t \tag{4}$$

$$U_t = X_t + V_t + W_t \tag{5}$$

The following holds:

$$X_t = 2^{\bar{r}t}(2^{\bar{r}} - 2^{-\bar{r}})(\hat{Z}_t - Z), \ t \geq 0 \tag{6}$$

$$E[X_t^2] \leq \left(\sigma_W \sqrt{2^{2\bar{r}} - 1} + \bar{\sigma}_V(2^{\bar{r}} + 1) + 2^{-\bar{r}t}(2^{\bar{r}} - 2^{-\bar{r}}) \right)^2, \ t \geq 0 \tag{7}$$

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

$$Prob\left(|X_t| \geq \alpha\right) \leq e^{-\frac{(\alpha-\gamma)^2}{2\beta^2}}, \ \alpha > 0, \ t \geq 0 \tag{8}$$

where γ and β are the following positive real constants:

$$\gamma \overset{def}{=} (2^{\bar{r}} + 1)\,\bar{\sigma}_V + 2^{\bar{r}} - 2^{-\bar{r}} \tag{9}$$

$$\beta^2 \overset{def}{=} \left(2^{2\bar{r}} - 1\right)\sigma_W^2 \tag{10}$$

Proof: In order to derive (6), we substitute $U_t = V_t + Y_t$ in (2). We now proceed to proving the validity of (7). Since the operators $\phi_{\bar{r}}$ and $\hat{\phi}_{\bar{r}}$ are linear, we can bound the variance of X_t

by separately quantifying the contribution of the external *inputs* Z, W_t and V_t. By making use of the triangular inequality, we arrive at the following bound:

$$\left(E[X_t^2]\right)^{1/2} \leq \left(\sigma_W^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|T\left(e^{j\omega}\right)\right|^2 d\omega\right)^{1/2} + \bar{\sigma}_V \max_{\omega \in (-\pi,\pi]} \left|T\left(e^{j\omega}\right)\right| + 2^{-\bar{r}t}(2^{\bar{r}} - 2^{-\bar{r}}) \quad (11)$$

where $T\left(e^{j\omega}\right)$ is the following transfer function:

$$T\left(e^{j\omega}\right) = \frac{2^{-\bar{r}} - 2^{\bar{r}}}{e^{j\omega} - 2^{-\bar{r}}} \quad (12)$$

The transfer function $T\left(e^{j\omega}\right)$ describes the input-output behavior of the feedback loop from V_t to X_t and from W_t to X_t. The first term in the right hand side of (11) quantifies the contribution from the white process W_t, while the second term is an upper-bound to the contribution of V_t and the last term comes from the *initial condition* determined by Z. Standard computations lead to the following results:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left|T\left(e^{j\omega}\right)\right|^2 d\omega = 2^{2\bar{r}} - 1 \quad (13)$$

$$\max_{\omega \in (-\pi,\pi]} \left|T\left(e^{j\omega}\right)\right| = \frac{2^{\bar{r}} - 2^{-\bar{r}}}{1 - 2^{-\bar{r}}} = 2^{\bar{r}} + 1 \quad (14)$$

After substituting (13) and (14) in (11), we arrive at (7). In order to prove (8)-(10), under the assumption that W_t is zero mean white Gaussian, we define the following auxiliary Gaussian process:

$$\tilde{X}_t = \begin{cases} 0 & \text{if } t = 0 \\ (2^{-\bar{r}} - 2^{\bar{r}}) \sum_{i=0}^{t-1} 2^{-\bar{r}(t-i-1)} W_i & \text{if } t \geq 1 \end{cases} \quad (15)$$

After simple manipulations, similar to the ones leading to (13)-(14), we get the following properties of $\tilde{X}_t$:

$$E[\tilde{X}_t^2] = \left(2^{2\bar{r}} - 1\right)\left(1 - 2^{-2\bar{r}t}\right)\sigma_W^2 \leq \beta^2 \quad (16)$$

$$|\tilde{X}_t - X_t| \leq \left(2^{\bar{r}} - 2^{-\bar{r}}\right)\left(\bar{\sigma}_V \frac{1 - 2^{-\bar{r}t}}{1 - 2^{-\bar{r}}} + 2^{-\bar{r}t}\right) \leq \gamma \quad (17)$$

where we used the definitions (9) and (10) along with (3). Consequently, we arrive at:

$$Prob\left(|X_t| \geq \alpha\right) \leq Prob\left(|\tilde{X}_t| \geq \alpha - \gamma\right) \leq \sqrt{\frac{2}{\pi\beta^2}} \int_{\alpha-\gamma}^{\infty} e^{-\frac{\mu^2}{2\beta^2}} d\mu, \ \alpha > 0 \quad (18)$$

where we used the facts that, by definition, $|\tilde{X}_t - X_t| \leq \gamma$, that $E[\tilde{X}_t^2] \leq \beta^2$ and that $\tilde{X}_t$ is normally distributed. The derivation of (8) is complete once we use the following upper-bound [7, page 220 eq. (5.1.8)]:

$$\sqrt{\frac{2}{\pi\beta^2}} \int_{\alpha-\gamma}^{\infty} e^{-\frac{\mu^2}{2\beta^2}} d\mu \leq e^{-\frac{(\alpha-\gamma)^2}{2\beta^2}} \quad (19)$$

□

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)-(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 which 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[1] transmission, in bits per channel 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 $\bar{\Theta}$ and Θ, respectively.

$$\bar{\Theta}(a) \stackrel{def}{=} \min\{n \in \mathbb{N} : a \leq n\}, \ a \in \mathbb{R} \tag{20}$$

$$\Theta(a) \stackrel{def}{=} \max\{n \in \mathbb{N} : a \geq n\}, \ a \in \mathbb{R} \tag{21}$$

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 $\{1, \ldots, 2^{\Theta(rn)}\}$ and a real stochastic process $\hat{Z}_t$, the following is the

---

[1]By reliable transmission we mean that the probability of error converges to zero with increasing block length n.
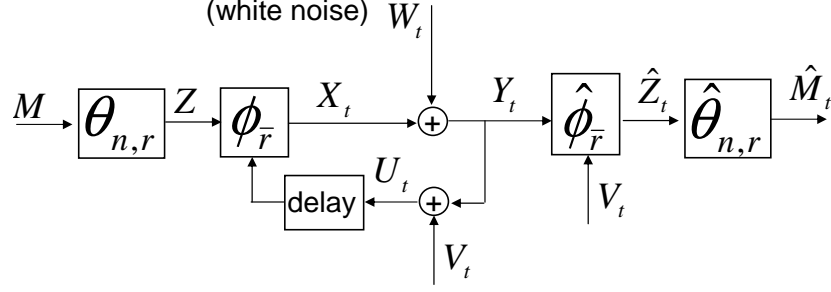
Fig. 2.   Basic feedback scheme with encoding and decoding.

definition of the maps $\theta_{n,r} : M \mapsto Z$ and $\hat{\theta}_{n,r} : \hat{Z}_t \mapsto \hat{M}_t$:

$$Z = \theta_{n,r}(M) \stackrel{def}{=} \left(M - \frac{1}{2} \right) 2^{-\Theta(rn)} \tag{22}$$

$$\hat{M}_t = \hat{\theta}_{n,r}(\hat{Z}_t) \stackrel{def}{=} \bar{\Theta} \left(2^{\Theta(rt)} \hat{Z}_t \right), \ t \in \{0, \dots, n\} \tag{23}$$

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.

$$\hat{M}_t = \hat{\theta}_{n,r} \left(\hat{\phi}_{\bar{r}}(t, Y^t, V^t) \right) \tag{24}$$

$$Y_t = W_t + \underbrace{\phi_{\bar{r}} \left(t, \theta_{n,r}(M), U^{t-1} \right)}_{X_t} \tag{25}$$

$$U_t = Y_t + V_t \tag{26}$$

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 $\{1, \dots, 2^{\Theta(rn)}\}$. If we adopt the scheme of Fig 2, alternatively described by (24)-(26), then the following holds:

$$Prob \left(M \neq \hat{M}_n \right) \leq \frac{2^{-2(\bar{r}-r)n} E[X_n^2]}{4(2^{\bar{r}} - 2^{-\bar{r}})^2} \tag{27}$$

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:

$$Prob\left(M \neq \hat{M}_n\right) \leq e^{-\frac{1}{2\beta^2}\left(2(2^{\bar{r}}-2^{-\bar{r}})2^{(\bar{r}-r)n}-\gamma\right)^2} \tag{28}$$

where γ and β are positive real constants given by (9) and (10), respectively.

Proof: We start by using (22)-(23) and the fact that $2^{\Theta(rn)}Z$ is in the set $\{\frac{1}{2}, \ldots, 2^{\Theta(rn)} - \frac{1}{2}\}$ to conclude the following:

$$\left|2^{\Theta(rn)}Z - 2^{\Theta(rn)}\hat{Z}_n\right| < \frac{1}{2} \implies M = \hat{M}_n \tag{29}$$

leading to:

$$Prob\left(M \neq \hat{M}_n\right) \leq Prob\left(\left|Z - \hat{Z}_n\right| \geq 2^{-(\Theta(rn)+1)}\right) \tag{30}$$

Using (6), (30) and the fact that $\Theta(rn) \leq rn$, we get:

$$Prob\left(M \neq \hat{M}_n\right) \leq Prob\left(|X_n| \geq 2(2^{\bar{r}} - 2^{-\bar{r}})2^{(\bar{r}-r)n}\right) \tag{31}$$

The inequality (27) follows from Markov's inequality applied to (31). Finally, the inequality (28) follows from (31) and (8). □

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 non-negative real solution ϱ of the following equation:

$$\sigma_W\sqrt{2^{2\varrho} - 1} = P_X - \bar{\sigma}_V\left(1 + 2^\varrho\right) \tag{32}$$

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 non-negative real solution of (32), 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.

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 (24)-(26), leads to:

$$E[X_t^2] \leq \left(P_X + \underbrace{2^{-\bar{r}t}(2^{\bar{r}} - 2^{-\bar{r}})}_{\text{vanishes with increasing } t} \right)^2, 0 \leq t \leq n \tag{33}$$

$$Prob\left(M \neq \hat{M}_n \right) \leq \frac{2^{-2(\bar{r}-r)n} E[X_n^2]}{4(2^{\bar{r}} - 2^{-\bar{r}})^2} \tag{34}$$

where M is a random variable arbitrarily distributed in the set $\{1, \ldots, 2^{\Theta(nr)}\}$. 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:

$$Prob\left(M \neq \hat{M}_n \right) \leq e^{-\frac{1}{2\beta^2}\left(2(2^{\bar{r}} - 2^{-\bar{r}})2^{(\bar{r}-r)n} - \gamma\right)^2}, \quad \alpha > 0 \tag{35}$$

where γ and β are positive real constants given by (9) and (10), 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[2] than P_X^2, allows for reliable transmission at any rate r strictly less than $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$. In addition, Theorem 3.2 shows that any rate of transmission r, if strictly less than $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$, leads to an achievable error exponent arbitrarily close to $2\left[r - \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)\right]$. 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 (35)).

Proof of Theorem 3.2: The inequalities (34) and (35) follow directly from Lemma 3.1. The derivation of (33) follows from (7) and from the fact that, from Definition 3.2, $\bar{r} < \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$ implies that $\sigma_W \sqrt{2^{2\bar{r}} - 1} + \bar{\sigma}_V(2^{\bar{r}} + 1) < P_X$. $\square$

It follows from its definition, as the solution to (32), that $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$ also satisfies the following 3 properties:

$$\lim_{\bar{\sigma}_V \to 0^+} \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V) = \frac{1}{2} \log_2\left(1 + \frac{P_X^2}{\sigma_W^2}\right), \quad \sigma_W^2 > 0, \ P_X^2 > 0 \tag{36}$$

$$\varrho\left(\sigma_W^2, P_X^2, \frac{P_X^2}{4}\right) = 0, \quad \sigma_W^2 > 0, \ P_X^2 > 0 \tag{37}$$

$$\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V) \simeq \log_2\left(\frac{P_X}{\sigma_W + \bar{\sigma}_V}\right), \quad P_X^2 >> \max\{\sigma_W^2, \bar{\sigma}_V^2\}, \tag{38}$$
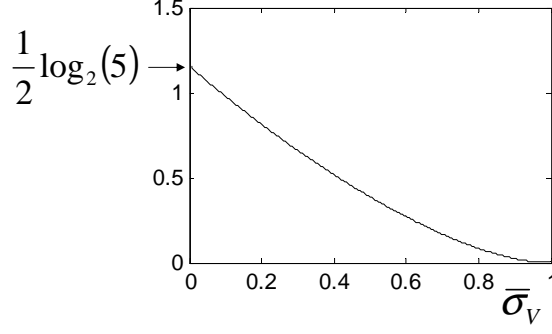
---

[2]See inequality (33).

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

where $\simeq$ indicates that the ratio between the left and right hand sides of (38) tends to 1 as $P_X \to \infty$. If W_t is white Gaussian then (36) indicates that in the limit, as the second moment of feedback noise goes to zero, the scheme of Fig 2 approaches capacity[3]. 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 (36) and (37), respectively.

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 re-constructs 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:

Definition 4.1: Given a positive real parameter b, a uniform quantizer with sensitivity b is a

---

[3]It is a standard fact [1] 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 $\frac{1}{2} \log_2 \left(1 + \frac{P_X^2}{\sigma_W^2}\right)$.
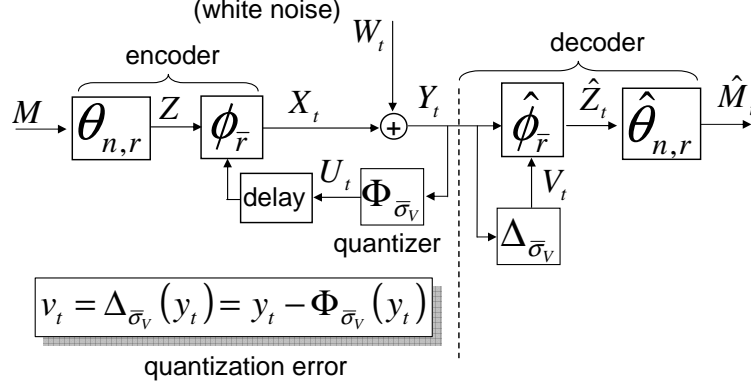
Fig. 4.   A coding and decoding scheme in the presence of uniformly quantized feedback.

function $\Phi_b : \mathbb{R} \to \mathbb{R}$ defined as:

$$\Phi_b(y) = 2b\Theta \left(\frac{y+b}{2b} \right) \tag{39}$$

where Θ is the *floor* function specified in (21). Similarly, the quantization error is given by the following function:

$$\Delta_b(y) = \Phi_b(y) - y, y \in \mathbb{R} \tag{40}$$

which satisfies the following bound:

$$|\Delta_b(y)| \le b, y \in \mathbb{R} \tag{41}$$

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

$$\hat{M}_t = \hat{\theta}_{n,r} \left(\hat{\phi}_{\bar{r}}(t, Y^t, V^t) \right) \tag{42}$$

$$Y_t = W_t + \underbrace{\phi_{\bar{r}} \left(t, \theta_{n,r}(M), U^{t-1} \right)}_{X_t} \tag{43}$$

$$U_t = \Phi_{\bar{\sigma}_V}(Y_t) = Y_t + V_t \tag{44}$$

$$V_t = \Delta_{\bar{\sigma}_V}(Y_t) \tag{45}$$

The Corollary below follows directly from Theorem 3.2 applied to the scheme of Fig 4, along with the upper-bound (41).

---

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

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 quantizer. In addition, select a positive transmission rate r and a positive real constant $\bar{r}$ satisfying $r < \bar{r} < \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V^2)$. For every positive integer block length n, the coding scheme specified by (42)-(45) (see Fig 4) leads to:

$$E[X_t^2] \leq \left(P_X + \underbrace{2^{-\bar{r}t}(2^{\bar{r}} - 2^{-\bar{r}})}_{\text{vanishes with increasing } t} \right)^2, 0 \leq t \leq n \tag{46}$$

$$Prob\left(M \neq \hat{M}_n \right) \leq \frac{2^{-2(\bar{r}-r)n}E[X_n^2]}{4(2^{\bar{r}} - 2^{-\bar{r}})^2} \tag{47}$$

where M is a random variable arbitrarily distributed in the set $\{1, \ldots, 2^{\Theta(nr)}\}$. 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:

$$Prob\left(M \neq \hat{M}_n \right) \leq e^{-\frac{1}{2\beta^2}\left(2(2^{\bar{r}} - 2^{-\bar{r}})2^{(\bar{r}-r)n} - \gamma \right)^2} \tag{48}$$

where γ and β are positive real constants given by (9) and (10), respectively.

Notice that Corollary 4.1 shows that, in the presence of uniformly quantized feedback with sensitivity $\bar{\sigma}_V$, any rate r strictly less than $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)$ allows for reliable transmission. This implies that the properties (36)-(37), 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 portrayed 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
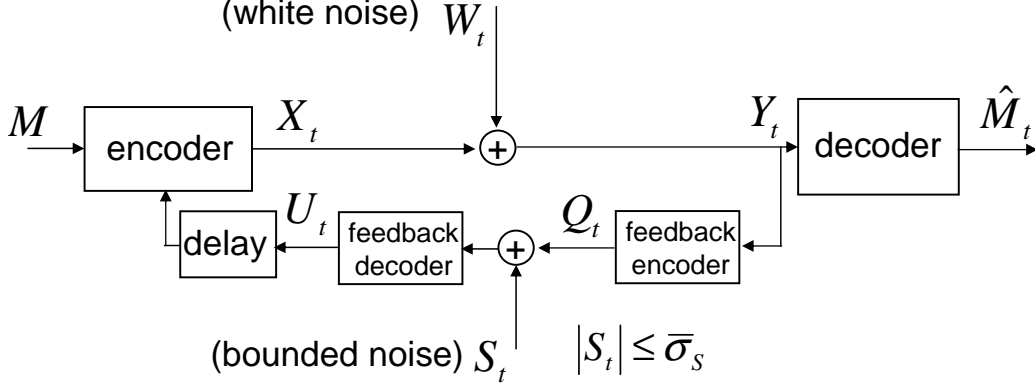
Fig. 5. Communication scheme in the presence of bounded feedback noise.

simplicity of the Schalkwijk-Kailath scheme [9], but, in contrast to the original scheme (which breaks down in the presence of noise in the backward link [9, Section 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. The main results of this Section are stated in Theorem 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 defined below:

$$\bar{\sigma}_S \overset{def}{=} \inf\{\alpha \in \mathbb{R}_{\geq 0} : Prob(|S_t| > \alpha) = 0, t \geq 0\}$$

meaning that the following holds:

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

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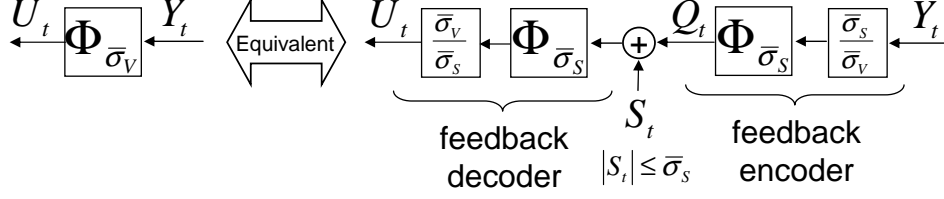
Fig. 6. Schematic representation of the equivalence expressed in Remark 5.1.

holds with probability one:

$$\frac{\bar{\sigma}_V}{\bar{\sigma}_S}\Phi_{\bar{\sigma}_S}\left(S_t + Q_t\right) = \Phi_{\bar{\sigma}_V}(Y_t) \tag{49}$$

where Q_t is given by:

$$Q_t = \Phi_{\bar{\sigma}_S}\left(\frac{\bar{\sigma}_S}{\bar{\sigma}_V}Y_t\right) \tag{50}$$

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.

Regarding the role of $\bar{\sigma}_V$, we have shown in (36) 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 (50) 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 σ_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
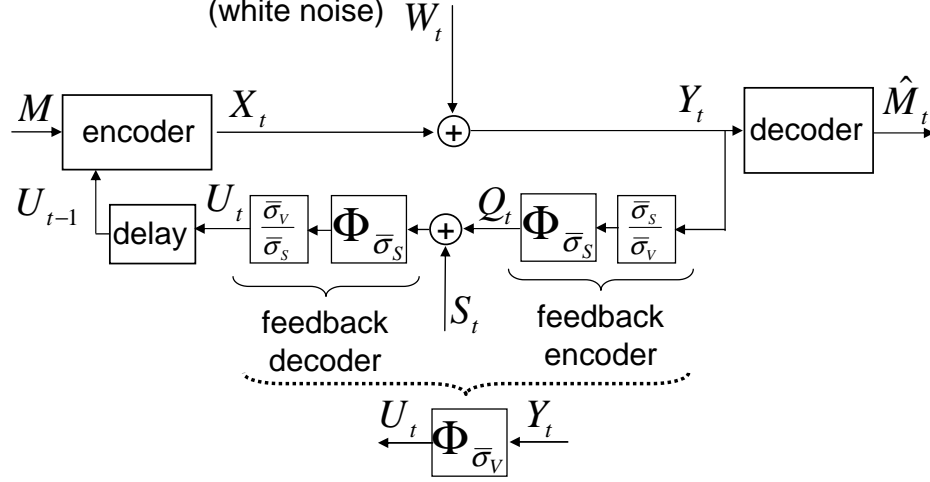
Fig. 7. Proposed coding scheme for dealing with the feedback corruption that results from additive noise S_t in the backward channel. The scheme is constructed by replacing the quantizer in the diagram of Fig 4 with the scheme of Fig 6. The encoder and decoder blocks are described in detail in Fig 4.

$\Gamma : \mathbb{R}^4_{\geq 0} \to \mathbb{R}_{\geq 0}$, which we will use as a selection for the design parameter $\bar{\sigma}_V$:

$$\Gamma\left(\sigma_W, \bar{\sigma}_S, P_X, P_Q\right) = \left(P_X + \sigma_W\right) \frac{\bar{\sigma}_S}{P_Q - \bar{\sigma}_S}, \ P_Q > \bar{\sigma}_S \tag{51}$$

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 $4\Gamma\left(\sigma_W, \bar{\sigma}_S, P_X, P_Q\right)^2 < P_X^2$ and $\bar{\sigma}_S < P_Q$. In addition, select a positive transmission rate r and a positive real constant $\bar{r}$ satisfying $r < \bar{r} < \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)|_{\bar{\sigma}_V = \Gamma\left(\sigma_W, \bar{\sigma}_S, P_X, P_Q\right)}$. For every positive integer block length n, the coding scheme of Fig 7, alternatively described by (42)-(45) and (50), leads to:

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

$$E[Q_t^2] \leq \left(P_Q + \underbrace{2^{-\bar{r}t}\frac{P_Q - \bar{\sigma}_S}{P_X + \sigma_W}\left(2^{\bar{r}} - 2^{-\bar{r}}\right)}_{\text{vanishes with increasing } t} \right)^2, \ 0 \leq t \leq n \tag{53}$$

$$Prob\left(M \neq \hat{M}_n \right) \leq \frac{2^{-2(\bar{r}-r)n}E[X_n^2]}{4(2^{\bar{r}} - 2^{-\bar{r}})^2} \tag{54}$$

where M is a random variable arbitrarily distributed in the set $\{1, \ldots, 2^{\Theta(nr)}\}$. 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:

$$Prob\left(M \neq \hat{M}_n\right) \leq e^{-\frac{1}{2\beta^2}\left(2(2^{\bar{r}}-2^{-\bar{r}})2^{(\bar{r}-r)n}-\gamma\right)^2} \tag{55}$$

where γ and β are positive real constants given by (9) and (10), respectively, where $\bar{\sigma}_V$ is given by the assumed selection $\bar{\sigma}_V = \Gamma\left(\sigma_W, \bar{\sigma}_S, P_X, P_Q\right)$.

Proof: The inequalities (52), (54) and (55) follow directly from Corollary 4.1. In order to arrive at (53), we start by noticing that we can use the triangular inequality to find the following inequalities:

$$\left(E[Y_t^2]\right)^{\frac{1}{2}} \leq \left(E[X_t^2]\right)^{\frac{1}{2}} + \sigma_W \tag{56}$$

$$\left(E[Q_t^2]\right)^{\frac{1}{2}} \leq \frac{\bar{\sigma}_S}{\bar{\sigma}_V}\left(E[Y_t^2]\right)^{\frac{1}{2}} + \bar{\sigma}_S \tag{57}$$

In addition, substitution of (56) in (57), leads to:

$$E[Q_t^2] \leq \left(\frac{\bar{\sigma}_S}{\bar{\sigma}_V}\left(E[X_t^2]^{\frac{1}{2}} + \sigma_W\right) + \bar{\sigma}_S\right)^2 \tag{58}$$

which, from (52), implies the following:

$$E[Q_t^2] \leq \left(\frac{\bar{\sigma}_S}{\bar{\sigma}_V}\left(P_X + 2^{-\bar{r}t}\left(2^{\bar{r}} - 2^{-\bar{r}}\right) + \sigma_W\right) + \bar{\sigma}_S\right)^2 \tag{59}$$

The proof is complete since (53) follows by substituting our choice $\bar{\sigma}_V = \Gamma\left(\sigma_W, \bar{\sigma}_S, P_X, P_Q\right)$ in (59). $\square$

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

$$\lim_{\bar{\sigma}_S \to 0^+} \varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)\big|_{\bar{\sigma}_V = \Gamma\left(\sigma_W, \bar{\sigma}_S, P_X, P_Q\right)} = \frac{1}{2}\log_2\left(1 + \frac{P_X^2}{\sigma_W^2}\right), \quad \sigma_W > 0, P_X > 0, P_Q > 0 \tag{60}$$

Notice that (60) 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 (60) 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
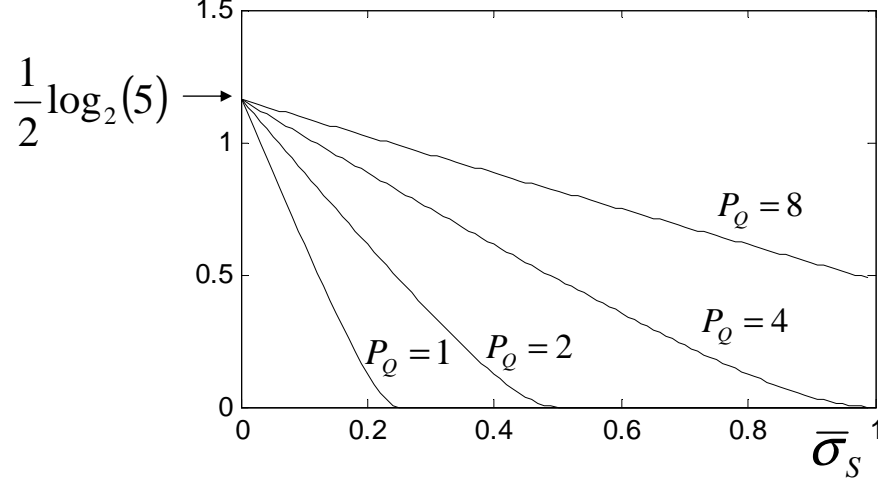
Fig. 8. Plot of $\varrho(\sigma_W^2, P_X^2, \bar{\sigma}_V)\big|_{\bar{\sigma}_V = \Gamma(\sigma_W, \bar{\sigma}_S, P_X, P_Q)}$ using $\sigma_W^2 = 1$, $P_X^2 = 4$ and $\bar{\sigma}_S \in [0, 1)$, for P_Q taking values 1, 2, 4 and 8.
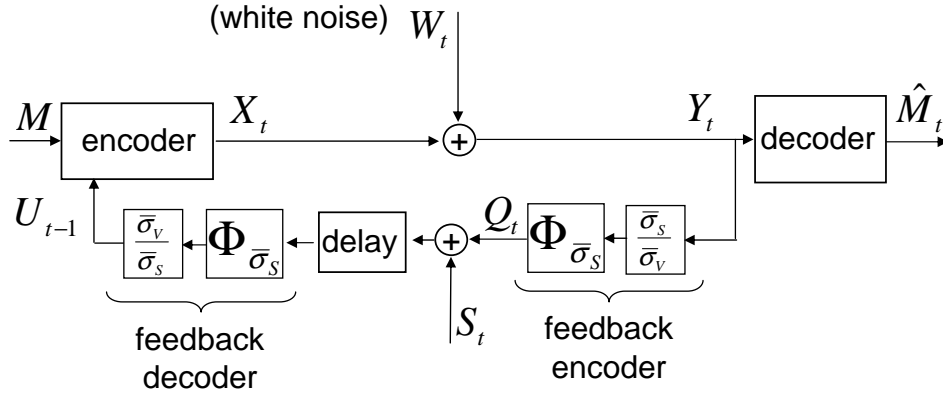

Fig. 9. A coding scheme equivalent to the one described by Fig 7.

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 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 Fig 7 and 9 are equivalent, implying that Theorem 5.1 holds also for the coding scheme of Fig 9.

VI. Conclusions

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.

We believe that our approach to the construction and analysis of coding schemes carries over naturally to the case where the noise in the forward channel is non-white. In this case, we expect to obtain variations on the schemes in [4] that are analogous to those in the present work and whose gap to capacity behaves similarly.

References

[1] T. M. Cover and J. A. Thomas; *"Elements of Information Theory,"* Wiley-Iterscience Publication, 1991

[2] S. C. Draper and A. Sahai, *"Noisy feedback improves communication reliability,"* Proceedings of the *International Symposium of Information Theory,* Seattle, Washington, July 2006

[3] N. Elia, *"When Bode Meets Shannon: Control-Oriented Feedback Communication Schemes,"* IEEE Transactions on Automatic Control, Vol. 49, No. 9, pp. 1477-1488, Sept. 2004

[4] Y. H. Kim, *"Feedback capacity of stationary Gaussian channels,"* submitted to *IEEE Transactions on Information Theory*. Available at " arxiv.org/abs/cs.IT/0602091"

[5] Y. H. Kim, A. Lapidoth and T. Weissman, *"Upper bounds on error exponents of channels with feedback,"* to appear in *IEEE 24th Convention of Electrical & Electronics Engineers in Israel*, Eilat, Israel, November 2006

[6] H. Permuter, T. Weissman and A. Goldsmith, *"Finite State Channels with Time-Invariant Deterministic Feedback,"* Available at "http://arxiv.org/abs/cs.IT/0608070"

[7] J. G. Proakis, M. Salehi, *"Fundamentals of communication systems,"* Prentice Hall, 2005

[8] A. Sahai and T. Şimşek, *"On the variable-delay reliability function of discrete memoryless channels with access to noisy feedback,"* Proceedings of the *IEEE Information Theory Workshop*, San Antonio, Texas, 2004.

[9] J. P. M. Schalkwijk and T. Kailath, *"A coding scheme for additive noise channels with feedback I: no bandwidth constraint,"* *IEEE Transactions on Information Theory*, vol. 12, pp. 172-182, April 1966

[10] C. E. Shannon, *"The zero-error capacity of a noisy channel,"* *IRE Transactions on Information Theory*, vol. IT-2, pp. 8–19, September 1956.

[11] S. Yang, A. Kavcic and S. Tatikonda, *"Feedback capacity of finite-state machine channels,"* *IEEE Transactions on Information Theory*, Volume 51, Issue 3, pp. 799 –810, March 2005