

WRITING ON DIRTY PAPER WITH FEEDBACK*

JIALING LIU[†] AND NICOLA ELIA[†]

Abstract. “Writing on dirty paper” refers to the communication problem over a channel with both noise and interference, where the interference is known to the encoder non-causally and unknown to the decoder. This problem is regarded as a basic building block in both the single-user and multiuser communications, and it has been extensively investigated by Costa and other researchers. However, little is known in the case that the encoder can have access to feedback from the decoder. In this paper, we study the dirty-paper coding problem for feedback Gaussian channels without or with memory. We provide the most power efficient coding schemes for this problem, i.e., the schemes achieve lossless interference cancelation. These schemes are based on the Kalman filtering algorithm, extend the Schalkwijk-Kailath feedback codes, have low complexity and a doubly exponential reliability function, and reveal the interconnections among information, control, and estimation over dirty-paper channels with feedback. This research may be found useful to, for example, power-constrained sensor network communication.

Key words: Feedback communication; Dirty-paper coding; Lossless interference cancelation; Capacity-achieving coding scheme; Interconnections among information, control, and estimation

1. Introduction. The study of lossless interference cancelation in a communication system has attracted considerable interest from researchers, since the publication of Costa’s celebrated article “Writing on Dirty Paper” [1]. Costa considered a power constrained discrete-time channel, in which there are two independent processes that corrupt the channel inputs. One process, a sequence $\{\xi_k\}$ of independently and identically distributed (i.i.d.) Gaussian random variables of mean zero and variance Q (i.e. $\xi_k \stackrel{\text{iid}}{\sim} \mathcal{N}(0, Q)$), is completely known to the encoder *non-causally* and is unknown to the decoder; this is referred to as the *interference* (or channel states). The other process, a sequence $\{N_k\}$ of i.i.d. $\mathcal{N}(0, 1)$ random variables (independent of $\{\xi_k\}$), is unknown to neither the the encoder nor the decoder; this is referred to as the noise. This problem was initially studied in [2] and later in [1] where Costa named this problem as the *writing on dirty paper* (WDP) problem. Fig. 1 illustrates this model, referred to as the *dirty-paper* model, in which W is the message, u is the channel input, ξ is the interference, N is the additive white Gaussian noise (AWGN), $y := u + \xi + N$ is the channel output, and \hat{W} is the decoded message. If ξ is zero (or is also known to the receiver), then the channel can achieve a rate \mathcal{R} with the channel input power being at least $\mathcal{P}(\mathcal{R}) := 2^{2\mathcal{R}} - 1$. If ξ is not zero and is not known to the

*This research was supported by NSF under Grant ECS-0093950. The material in this paper was presented in part at the 2006 IEEE International Conference on Networking, Sensing and Control (ICNSC), Ft. Lauderdale, Florida, U.S.A., April 2006.

[†]The authors are with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, USA. E-mail: {liujl,nelia}@iastate.edu

receiver, then obviously the minimum channel input power $\mathcal{P}_{WDP}(\mathcal{R})$ for achieving rate \mathcal{R} is bounded by

$$(1) \quad \mathcal{P}(\mathcal{R}) \leq \mathcal{P}_{WDP}(\mathcal{R}) \leq (1 + Q)\mathcal{P}(\mathcal{R}).$$

What is surprising is that the lower bound is achievable, namely there exists a strategy such that we can transmit across the channel *as if the interference did not exist*. In other words, we can achieve *lossless interference cancelation*, by which we mean that the interference is “canceled” without incurring any power increase or rate loss; such an optimal strategy was introduced by Costa in [1]. Note that the strategy of letting the encoder ignore the interference knowledge or letting the decoder try the brutal force way to cancel the interference incurs power increase or rate loss, and hence they are only suboptimal.

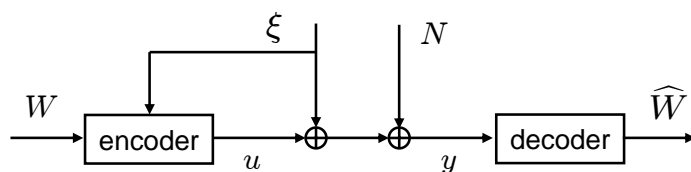


FIG. 1. AWGN dirty-paper channel model.

Costa’s results have been generalized to various situations; see [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13] and references therein. [6, 7] extended the results to the case of ergodic interference and colored Gaussian noise. [5, 8] showed that lossless interference cancelation is possible for arbitrarily varying interference, provided that the encoder and decoder share a common random dither signal. [9] showed that, as long as the interference and the noise are Gaussian (not necessarily memoryless, stationary, or ergodic), the channel has a capacity as if the interference did not exist. Various coding schemes were also provided; see e.g. [5, 8, 13]. On the other hand, if the interference is known to the transmitter only in a *causal* manner (in which case the problem is sometimes referred to as *writing on dirty tape* (WDT)), the problem is much more involved; in fact both the capacity computation problem and the capacity-achieving problem remain unsolved. See [3, 8, 4] for suboptimal coding strategies for the dirty-tape channels.

These results have found broad applications in information hiding [14], digital watermarking [7], precoding for intersymbol interference (ISI) channels [5], and precoding for broadcast channels [12]. To summarize its significance, the dirty-paper coding study has been considered to be a basic building block in both single-user and multiuser communication problems [13].

The above results are focused on the case where the encoder does not receive any feedback from the decoder. In many situations, however, it is possible for the

encoder to access the information at the decoder-side in a strictly causal way, namely the encoder has feedback from the decoder. This setup of feedback communication has drawn increasing attention in both the information theory community and control theory community; see [15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29] for a partial list. The availability of such feedback usually allows us to considerably simplify the coding scheme, to improve the performance, and to increase the capacity. We remark that little is done to extend the dirty-paper coding to feedback communication systems. The only notable exception is [30], where a feedback dirty-paper channel without memory was studied.

In this paper, we consider the dirty-paper channel where there is a *noiseless feedback* from the decoder to the encoder with one-step delay. We assume non-causal encoder-side information about the interference. In the case of arbitrarily varying interference and AWGN, we present a Kalman filter based coding scheme to achieve lossless interference cancelation. We then extend this result to the case of dirty-paper channels with both AWGN and inter-symbol interference, based on the techniques developed in the ISI-free case. Our coding schemes greatly simplify the encoder/decoder design and encoding/decoding processes, guarantee doubly exponentially decay of probability of error, and are optimal in the sense of achieving the stationary capacity (see Section 3 for details), namely, they are the *most power efficient* through these channels for any given communication rate over all possible stationary input distributions. The schemes are also extensions of the Schalkwijk-Kailath codes for AWGN non-dirty-paper channels to dirty-paper channels with both AWGN and ISI.

Our study reveals and utilizes the intimate connections among information, and control, and estimation over a dirty-paper channel. We show that the feedback communication over a dirty-paper channel is essentially equivalent to a Kalman filtering problem and a minimum-energy control problem (see [31, 29] for the study of minimum-energy control problem). This extends the perspective of integrating information, control, and estimation (proposed in [32, 33]) to dirty-paper channels with feedback. Such a unifying perspective may provide a systematic way to address more general feedback communication problems such as multiuser dirty-paper channels with feedback.

We remark that our assumption on the noiseless feedback link, though in general not realistic enough, is widely assumed in the feedback communication literature, has significant theoretic implications, and might be a useful step towards the study of the more realistic case of noisy feedback. On the other hand, a potential application area of the research under noiseless feedback assumption is the sensor networks, in which the forward communication from the sensors to the base station (or the cluster center, if any) may be very noisy due to the limited power of the sensors, whereas the feedback communication from the base station to the sensors may be viewed as

noiseless due to the high power of the base station. The interference to a sensor may be the signals sent by its neighboring sensors. Since the signals for neighboring sensors are usually correlated, the interference is partially known to this sensor. Our results say that we can significantly improve the forward transmission by taking advantage of both the feedback transmission and the knowledge about interference, which may be useful in sensor networks.

Organization: In Section 2, we introduce the optimal coding scheme for a dirty-paper channel with AWGN and arbitrarily varying interference. In Section 3, we study a dirty-paper channel with arbitrarily varying interference, AWGN, and ISI. In Section 4, we provide a numerical example. Finally we conclude this paper and discuss future research directions.

Notations: We represent time indices by subscripts, such as y_k . We denote by y^K the vector $\{y_0, y_1, \dots, y_K\}$, and $\{y_k\}$ the sequence $\{y_k\}_{k=0}^\infty$. Note that we assume the starting time of all processes is 0, consistent with the convention in dynamical systems but different from the information theory literature. We denote “defined to be” as “:=”.

2. The AWGN case. Consider the power constrained dirty-paper channel with AWGN shown in Fig. 1. Let the average power budget be $\mathcal{P} > 0$. Let $\{\xi_k\}$ be an interference sequence known non-causally to the encoder. This interference sequence can be deterministic or random; we assume that

$$(2) \quad \frac{1}{K+1} \left(\sum_{j=0}^K (1+\mathcal{P})^{-\frac{j}{2}} \mathbf{E}\xi_j \right)^2 \rightarrow 0$$

as K tends to infinity (noting that $\mathbf{E}\xi_j = \xi_j$ in the deterministic case). Let further $\{N_k\}$ be AWGN with $N_k \sim \mathcal{N}(0, 1)$. We first describe the proposed coding scheme and the coding process, then present the coding theorem, followed by the proof, and finally discuss the connections to a minimum-energy control problem and a Kalman filtering problem.

2.1. Coding scheme. Fig. 2 illustrates the designed coding system, in which we can identify the encoder, decoder, and dirty-paper channel. Let us fix the time horizon to be $\{0, 1, \dots, K\}$, namely the number of channel uses is $(K+1)$.

The encoder/decoder structures

In state-space, the encoder and decoder are described as

$$(3) \quad \text{encoder: } \begin{cases} x_k &= ax_{k-1} - L(y_{k-1} - \xi_{k-1}) \\ u_k &= cx_k \\ y_k &= u_k + \xi_k + N_k \end{cases}$$

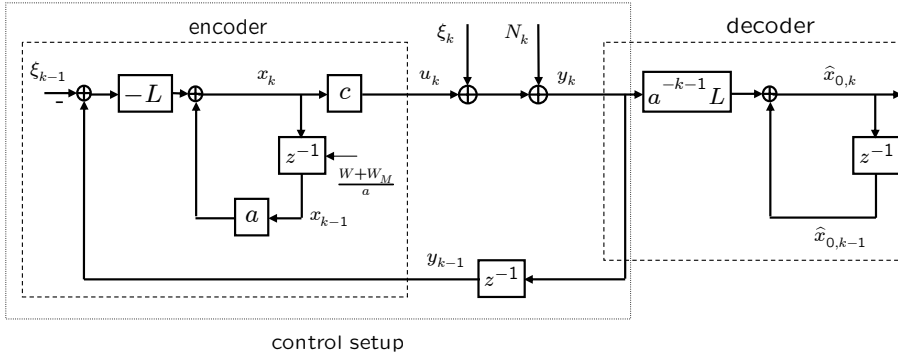


FIG. 2. The optimal coding scheme for a dirty-paper channel with AWGN. The dotted box indicates a control system, referred to as the control setup (see Section 2.3).

and

$$(4) \quad \text{decoder:} \quad \hat{x}_{0,k} = \hat{x}_{0,k-1} + a^{-k-1}Ly_k,$$

where

$$(5) \quad \begin{aligned} a &:= \sqrt{1 + \mathcal{P}} > 1 \\ c &:= 1 \\ L &:= a - \frac{1}{a}, \end{aligned}$$

$\xi_{-1} := 0$, $y_{-1} := 0$, $\hat{x}_{0,-1} := 0$, and x_0 will be determined shortly. We call x_k the *encoder state* and $\hat{x}_{0,k}$ the *decoder state*.

Transmission of analog source

The designed communication system can transmit either an analog source or a digital message. In the former case, we assume that the encoder wishes to convey a Gaussian random variable through the channel and the decoder wishes to learn the random variable, which is a rate-distortion problem or a successive refinement problem (see [18, 20, 34] and reference therein for study of successive refinement and its generalization, the sequential rate-distortion problem).

The coding process is as follows. Assume without loss of generality that the to-be-conveyed message W is distributed as $\mathcal{N}(0, \mathcal{P})$ (if the variance is not \mathcal{P} , we can scale W to have the desired variance). To encode, let

$$(6) \quad x_{-1} := \frac{W + W_M}{a},$$

i.e. $x_0 := W + W_M$, where ¹

$$(7) \quad W_M := -\frac{\sum_{j=0}^K a^{-j-1}L\xi_j}{1 - a^{-K-2}}.$$

¹Another choice which is asymptotically equivalent to (7) is $W_M := -\sum_{j=0}^K a^{-j-1}L\xi_j$.

Then run the system till instant of time K , generating $\hat{x}_{0,k}$ for $k = 0, 1, \dots, K$. To decode, let $\hat{W}_K := \hat{x}_{0,K}$. The distortion measure is

$$(8) \quad \text{MSE}(\hat{W}_K) := \mathbf{E}(W - \hat{W}_K)^2.$$

Transmission of digital message

To transmit digital messages over the communication system, let us fix $\epsilon > 0$ arbitrarily small. Suppose that we wish to transmit one of a set of

$$(9) \quad M_K := a^{(K+1)(1-\epsilon)}$$

messages. We equally partition the interval

$$(10) \quad \left[-\sqrt{\mathcal{P}} \left(1 + \frac{1}{M_K - 1} \right), \sqrt{\mathcal{P}} \left(1 + \frac{1}{M_K - 1} \right) \right]$$

into M_K sub-intervals, and map the sub-interval centers to a set of M_K equally likely messages; this is known to both the transmitter and receiver *a priori*.

Suppose now we wish to transmit the message represented by the center W . To encode, define x_{-1} according to (6). Then run the system till instant of time K . To decode, let the decoder estimate \bar{W}_K be

$$(11) \quad \bar{W}_K := \frac{\hat{x}_{0,K}}{1 - a^{-2K-2}}.$$

We then map \bar{W}_K into the closest sub-interval center and obtain the decoded message \hat{W}_K .² We declare an error if $\hat{W}_K \neq W$, and call a (an asymptotic) rate

$$(12) \quad \mathcal{R} := \lim_{K \rightarrow \infty} \frac{1}{K+1} \log M_K$$

achievable if the probability of error PE_K vanishes as K tends to infinity.

2.2. Coding theorem.

THEOREM 2.1. *Let $\{\xi_k\}$ be an arbitrarily varying interference sequence (deterministic or random) known to the encoder non-causally and satisfying (2), and $\{N_k\}$ be AWGN with $N_k \sim \mathcal{N}(0, 1)$. Then under the average power constraint $\mathbf{E}u^2 \leq \mathcal{P}$,*

i) The coding scheme constructed in Section 2.1 transmits an analog source $W \sim \mathcal{N}(0, \mathcal{P})$ from the encoder to the decoder at the capacity rate

$$(13) \quad C(\mathcal{P}) := \frac{1}{2} \log(1 + \mathcal{P}),$$

with MSE distortion $\text{MSE}(\hat{W}_K)$ satisfying the optimal rate-distortion tradeoff given by

$$(14) \quad C(\mathcal{P}) = \frac{1}{2(K+1)} \log \frac{\mathcal{P}}{\text{MSE}(\hat{W}_K)}$$

²Another decoding method is to map $\hat{x}_{0,K}$ directly into the closest sub-interval center and obtain the decoded message \hat{W}_K . This is asymptotically identical to the above decoding method. The scaling by $1/(1 - a^{-2K-2})$ is to remove the bias in the estimate of W ; see (24) and [35].

for each K .

ii) The coding scheme constructed in Section 2.1 can transmit digital messages from the encoder to the decoder at a rate arbitrarily close to $C(\mathcal{P})$, with PE_K decaying to zero doubly exponentially.

REMARK 1. Compared with the coding scheme for AWGN non-dirty-paper channels, the one for AWGN dirty-paper channels has two main differences: The presence of “process noise” (see [36] and Section 2.3) ξ_{k-1} at the encoder, and the presence of initial condition offset W_M . The presence of the process noise ξ_{k-1} at the encoder implies that, the interference ξ_k does not affect the encoder state x_k , since the process noise cancels the interference before the interference enters the encoder state. Thus, the channel input u_k is not affected by the interference since $u_k = cx_k$. In addition, by linearity of our coding scheme, the decoder state $\hat{x}_{0,K}$ depends affinely on x_0 , ξ , and N . The terms associated with x_0 and ξ are known to the encoder before the transmission because ξ^K is known to the encoder non-causally, and therefore the encoder can offset x_0 to cancel the term associated with ξ in $\hat{x}_{0,K}$, namely, the decoded message would not be affected by the interference.

Proof. To establish lossless interference cancelation, we need to show that, both the decoded message \hat{W}_K and the (asymptotic) average channel input power are not affected by the interference sequence. It is therefore sufficient to prove that both the decoder state $\hat{x}_{0,K}$ and the channel input u_k are decoupled (or asymptotically decoupled) from the interference sequence. Then the theorem would follow from standard results of feedback communication over an AWGN channel *without* interference. For completeness, we also include the proof for the AWGN non-dirty-paper channel case.

We first derive the expressions for u_k and $\hat{x}_{0,k}$. We can express the encoder state in terms of the initial condition, interference, and channel outputs as

$$(15) \quad x_k = a^k(W + W_M) + a^k \sum_{j=0}^{k-1} a^{-j-1} L(\xi_j - y_j).$$

Hence

$$(16) \quad \sum_{j=0}^{k-1} a^{-j-1} Ly_j = W + W_M - a^{-k}x_k + \sum_{j=0}^{k-1} a^{-j-1} L\xi_j.$$

On the other hand, noticing that $y_k = u_k + \xi_k + N_k$, the encoder state is also

$$(17) \quad \begin{aligned} x_k &= ax_{k-1} + L(\xi_{k-1} - u_{k-1} - \xi_{k-1} - N_{k-1}) \\ &= (a - Lc)x_{k-1} - LN_{k-1} \\ &= a^{-1}x_{k-1} - LN_{k-1} \\ &= a^{-k}(W + W_M) - \sum_{j=0}^{k-1} a^{-k+1+j} LN_j. \end{aligned}$$

Then the decoder state is

$$\begin{aligned}
\hat{x}_{0,k-1} &= \sum_{j=0}^{k-1} a^{-j-1} Ly_j \\
(18) \quad &= W + W_M - a^{-k} x_k + \sum_{j=0}^{k-1} a^{-j-1} L\xi_j \\
&= (1 - a^{-2k})(W + W_M) + a^{-2k} \sum_{j=0}^{k-1} a^{j+1} LN_j + \sum_{j=0}^{k-1} a^{-j-1} L\xi_j.
\end{aligned}$$

Therefore, by (7), we have

$$(19) \quad \hat{x}_{0,K} = (1 - a^{-2K-2})W + a^{-2K-2} \sum_{j=0}^K a^{j+1} LN_j,$$

and thus the interference does not affect the decoder state at time K and the decoded message \hat{W}_K . Now note that

$$(20) \quad u_k = cx_k = a^{-k}(W + W_M) - \sum_{j=0}^{k-1} a^{-k+1+j} LN_j.$$

That is, the presence of interference incurs an extra power overhead of $a^{-2k}W_M^2$, which vanishes exponentially and is negligible, since the power due to the term $\sum_{j=0}^{k-1} a^{-k+1+j} LN_j$ approaches a nonzero constant and the coding length $(K + 1)$ is sufficiently large. In summary, the interference does not affect either the decoded message or the asymptotic channel input power. This indeed leads to lossless interference canceling.

Then we compute the average channel input power, followed by the rate and distortion computation for i), and the rate and probability of error computation for ii). From (20), the input power is asymptotically determined only by the term $\sum_{j=0}^{k-1} a^{-k+1+j} LN_j$, which leads to that

$$(21) \quad \mathbf{E}u^2 = L^2 \lim_{k \rightarrow \infty} \sum_{j=0}^{k-1} a^{-2k+2+2j} = \frac{(a^2 - 1)^2}{a^2} \frac{a^2}{a^2 - 1} = \mathcal{P}.$$

i) The MSE distortion is

$$\begin{aligned}
\text{MSE}(\hat{W}_K) &= \mathbf{E} \left[a^{-4K-4}W^2 + a^{-4K-4} \sum_{j=0}^K a^{2j+2} L^2 (N_j)^2 \right] \\
(22) \quad &= a^{-4K-4} \mathcal{P} + a^{-4K-4} \mathcal{P}^2 \sum_{j=0}^K a^{2j} \\
&= a^{-4K-4} \mathcal{P} \left(1 + \mathcal{P} \frac{a^{2K+2} - 1}{a^2 - 1} \right) \\
&= a^{-4K-4} \mathcal{P} a^{2K+2} \\
&= \mathcal{P} a^{-2K-2}.
\end{aligned}$$

This distortion needs an information rate across the channel to be at least

$$(23) \quad R = \frac{1}{2K+2} \log \frac{\mathcal{P}}{\text{MSE}(\bar{W}_K)} = \log a = C(\mathcal{P}),$$

which is indeed the capacity of the AWGN channel and implies that the optimality is achieved.

ii) By (11) and (19), we have

$$(24) \quad \bar{W}_K = W + (a^{2K+2} - 1)^{-1} \sum_{j=0}^K a^{j+1} L N_j,$$

that is, \bar{W}_K is an unbiased estimate of W . For each given W , it holds that $\bar{W}_K \sim \mathcal{N}(W, (a^{2K+2} - 1)^{-2} \sum_{j=0}^K a^{2j+2} L^2)$.

The signalling rate is

$$(25) \quad \mathcal{R} := \lim_{K \rightarrow \infty} \frac{\log M_K}{K+1} = (1 - \epsilon) \log a.$$

This signalling rate is achievable if the probability of error vanishes. To compute the probability of error, note that for each message W , no error occurs if

$$(26) \quad |\bar{W}_K - W| \leq \frac{1}{M_K - 1} \sqrt{\bar{\mathcal{P}}}.$$

Then the probability of error satisfies

$$(27) \quad \begin{aligned} PE_K &\leq 2Q \left(\frac{\sqrt{\bar{\mathcal{P}}}/(M_K - 1)}{(a^{2K+2} - 1)^{-1} L \sqrt{\sum_{j=0}^K a^{2j+2}}} \right) \\ &= 2Q \left(\frac{\sqrt{\bar{\mathcal{P}}} a^{K+1} \sqrt{1 - a^{-2K-2}}}{(M_K - 1) \sqrt{a^2 - 1}} \right) \\ &= 2Q \left(a^{(K+1)\epsilon} \frac{\sqrt{1 - a^{-2K-2}}}{1 - a^{-(K+1)(1-\epsilon)}} \right) \\ &\stackrel{(a)}{\leq} 2Q \left(a^{(K+1)\epsilon} \right) \\ &\stackrel{(b)}{\leq} \frac{2}{\sqrt{2\pi} a^{(K+1)\epsilon}} \exp \left(-\frac{1}{2} a^{2(K+1)\epsilon} \right), \end{aligned}$$

where inequality (a) follows from

$$(28) \quad \frac{\sqrt{1 - a^{-2K-2}}}{1 - a^{-(K+1)(1-\epsilon)}} \geq 1$$

for any $\epsilon > 0$ and K , and (b) follows from the Chernoff bound

$$Q(t) \leq \frac{1}{\sqrt{2\pi} t} \exp\left(-\frac{1}{2} t^2\right).$$

Thus, PE_K decreases to zero doubly exponentially. \square

2.3. Connections to control problem and estimation problem. In this subsection, we briefly discuss the equivalent relations of our coding scheme to a minimum-energy control problem and a Kalman filtering problem (which can also be rewritten as a tracking-of-unstable-source problem). These relations are conceptually appealing since they provide another example that information, control, and estimation, three fundamental concepts, can be studied in a unified framework, and if any one of these problems is solved, other problems can be solved.

The dynamics of x_k in (3), repeated here as

$$(29) \quad \begin{cases} x_k &= ax_{k-1} + L(\xi_{k-1} - y_{k-1}) \\ u_k &= cx_k \\ y_k &= u_k + \xi_k + N_k, \end{cases}$$

is indeed a control system, as indicated in Fig. 2 as the control setup. The control setup is open-loop unstable with its pole at a , and is closed-loop stabilized with its pole at $1/a$, which is a *minimum-energy control* problem and the power of u is minimized over all possible choice of stabilizing L . Therefore, it can be established rigorously that the optimality in feedback communication coincides with the optimality in control, and reliable feedback communication using (3) and (4) is equivalent to feedback stabilization of (29); see [22] for relevant discussions.

Fig. 3 (a) illustrates a system closely related to the coding scheme shown in Fig. 2. The dynamics in Fig. 3 (a) is

$$(30) \quad \begin{cases} \tilde{x}_{k+1} &= a\tilde{x}_k + L\xi_k \\ r_k &= c\tilde{x}_k \\ u_k &= r_k - \hat{r}_k \\ y_k &= u_k + \xi_k + N_k \\ \hat{x}_{k+1} &= a\hat{x}_k + Ly_k \\ \hat{r}_k &= c\hat{x}_k \\ \hat{x}_{0,k} &= a^{-k-1}\hat{x}_{k+1}, \end{cases}$$

where $\tilde{x}_0 := (W + W_M)$ and $\hat{x}_0 := 0$. To see the relation to the coding scheme, letting $x_k := \tilde{x}_k - \hat{x}_k$ in (30), we indeed obtain the dynamics of (3) and (4) through straightforward manipulation. Therefore, (30) generates the same channel inputs, the same channel outputs, the same decoded message as (3) and (4) do. However, Fig. 3 (a) has an interpretation of tracking unstable source (i.e. $\{r_k\}$) over a communication channel by applying the internal mode principle, which has attracted considerable attention, see e.g. [20, 23, 22, 29]. It holds that asymptotic tracking of an unstable source over a (dirty-paper or non-dirty-paper) channel is equivalent to reliable feedback communication over the same channel.

We can furthermore arrange Fig. 3 (a) to obtain the block diagram shown in Fig. 3 (b), which is a Kalman filtering problem (cf. Fig. 1.1 in [36]). The dynamics of the

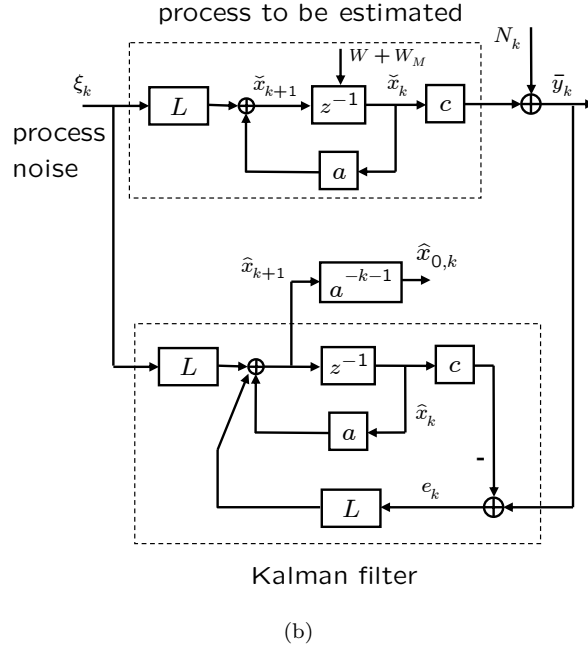
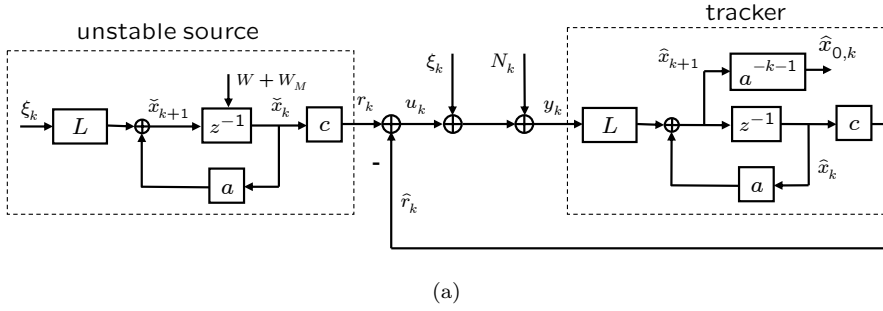


FIG. 3. (a) Tracking of unstable source. (b) The associated Kalman filtering problem. Note that the “process noise” ξ_k enters both the process to be estimated and the Kalman filter.

Kalman filtering shown in Fig. 3 (b) is

$$(31) \quad \begin{cases} \tilde{x}_{k+1} &= a\tilde{x}_k + L\xi_k \\ \bar{y}_k &= c\tilde{x}_k + N_k \\ \hat{x}_{k+1} &= a\hat{x}_k + L\xi_k + Le_k \\ e_k &= \bar{y}_k - c\hat{x}_k \\ \hat{x}_{0,k} &= a^{-k-1}\hat{x}_{k+1}, \end{cases}$$

where $\tilde{x}_0 := (W + W_M)$ and $\hat{x}_0 := 0$. Note that \tilde{x}_k , \hat{x}_k , and e_k in (31) are identical to those in (30), respectively. The feedback communication problem is then linked to an estimation problem, and the optimality in both problems coincides, see [32] for more detailed discussion. We also see from Fig. 3 (b) that the interference sequence becomes the process noise common to both the process to be estimated and the

Kalman filter. Then a well-known fact about Kalman filtering (cf. [36]) implies that the interference does not affect the estimation error ($\tilde{x}_k - \hat{x}_k$) and hence the channel input $u_k = c(\tilde{x}_k - \hat{x}_k)$. Therefore, our dirty-paper coding scheme is essentially a reformulation of Kalman filtering algorithm for the case of process noise known to both the process and the Kalman filter. It has been shown that the Kalman filter is optimal in information processing, storage, and dissipation [37]. The connection of the optimal dirty-paper coding scheme to the Kalman filter obtained here further confirms the optimality of the Kalman filter in the information theoretic sense in a more general setup. The role of the Kalman filter in more general communication setups is under current investigation.

3. The ISI Gaussian channel case. This section presents the optimal coding scheme for a dirty-paper channel with AWGN and ISI. We first describe the channel model. We then introduce the encoder/decoder structures and explain how to choose the parameters to ensure the optimality, and describe the encoding/decoding processes. Finally we present the coding theorem.

3.1. Channel model and the stationary capacity. The dirty-paper channel \mathcal{F} with ISI and AWGN is described in state-space as

$$(32) \quad \mathcal{F} : \begin{cases} s_{k+1} &= F s_k + G u_k \\ y_k &= H s_k + u_k + \xi_k + N_k, \end{cases}$$

where $s_0 := 0$, $F \in \mathbb{R}^m$ is stable, (F, G) is controllable, (F, H) is observable, the transfer function from u to y is minimum-phase, m is the *dimension* or *order* of \mathcal{F} , and ξ^K is an interference sequence known to the encoder non-causally and unknown to the decoder. See Fig. 4 (a) for the block diagram of \mathcal{F} . Through the equivalence shown in [18, 24, 29], the results developed for this channel hold for dirty-paper channels with colored Gaussian noise with rational power spectrum.

Our focus is to find a coding scheme to achieve the highest possible information rate when the channel input distributions are restricted to be stationary (noting that non-stationary input distributions are of little practical interest), in other words, we wish to achieve the stationary capacity.³ Precisely, our objective is to achieve

$$(33) \quad C^s := C^s(\mathcal{P}) := \sup_{\{u_k\}} \lim_{K \rightarrow \infty} \frac{1}{K+1} I(u^K \rightarrow y^K)$$

where the supremum is over all *stationary* input sequences $\{u_k\}$ satisfying the power

³It has been recently shown that the feedback capacity equals the stationary feedback capacity for a non-dirty-paper Gaussian channel, namely the input sequence achieving the feedback capacity can be chosen as (asymptotic) stationary (cf. [38, 39]). This follows that the proposed scheme in this paper achieves the feedback capacity for the dirty-paper channel.

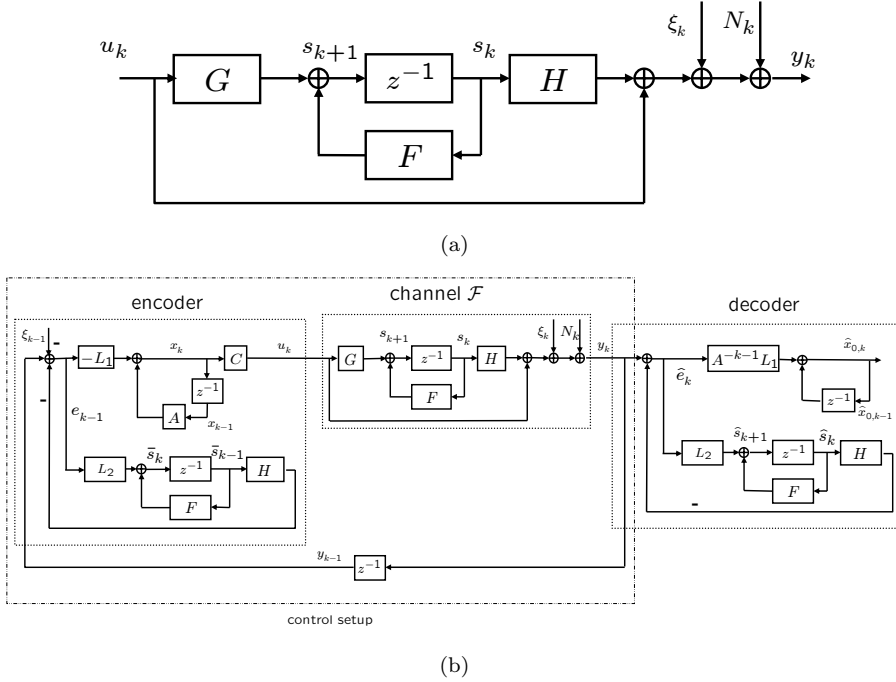


FIG. 4. (a) State-space representation of channel \mathcal{F} . (b) The encoder and decoder structures for \mathcal{F} .

constraint

$$(34) \quad \lim_{K \rightarrow \infty} \frac{1}{K+1} \mathbf{E} u^{K'} u^K \leq \mathcal{P}$$

and in the form of

$$(35) \quad u_k = \gamma_k u^{k-1} + \beta_k \xi^K + \eta_k y^{k-1} + \zeta_k$$

for any $\gamma_k \in \mathbb{R}^{1 \times k}$, $\beta_k \in \mathbb{R}^{1 \times (K+1)}$, $\eta_k \in \mathbb{R}^{1 \times k}$, and zero-mean Gaussian random variable $\zeta_k \in \mathbb{R}$. Here $\mathcal{P} > 0$ is the power budget and $I(u^K \rightarrow y^K)$ is the directed information from u^K to y^K (cf. [18]). The stationary capacity problem admits a finite-dimensional time-invariant solution, which has low complexity in system design and operation.

3.2. The coding scheme.

The encoder/decoder structures

In state-space, the encoder and decoder are described as

$$(36) \quad \text{Encoder: } \begin{cases} x_k & = Ax_{k-1} - L_1 e_{k-1} \\ u_k & = Cx_k \\ \bar{s}_k & = F\bar{s}_{k-1} + L_2 e_{k-1} \\ e_{k-1} & = y_{k-1} - \xi_{k-1} - H\bar{s}_{k-1} \end{cases}$$

and

$$(37) \quad \text{Decoder:} \begin{cases} \hat{s}_{k+1} &= F\hat{s}_k + L_2\hat{e}_k \\ \hat{e}_k &= y_k - H\hat{s}_k \\ \hat{x}_{0,k} &= \hat{x}_{0,k-1} + A^{-k-1}L_1\hat{e}_k, \end{cases}$$

where $s_0 := 0$, $\bar{s}_{-1} := 0$, $\xi_{-1} := 0$, $\hat{s}_0 := 0$, $\hat{x}_{0,-1} := 0$, $A \in \mathbb{R}^{(n+1) \times (n+1)}$, $C \in \mathbb{R}^{1 \times (n+1)}$, $L_1 \in \mathbb{R}^{n+1}$, and $L_2 \in \mathbb{R}^m$. We call $(n+1)$ the *encoder dimension*. See Fig. 4 (b) for the block diagram. Here A, C, u_k , etc. depend on n , but we do not specify the dependence explicitly to simplify notations.

Optimal choice of parameters

Fix a desired rate \mathcal{R} . Let $DI := 2^{\mathcal{R}}$ and $n := m - 1$ (recalling that m is the channel dimension), and solve the optimization problem

$$(38) \quad [\mathbf{a}_f^{opt}, \Sigma^{opt}] := \underset{\mathbf{a}_f \in \mathbb{R}^n}{\arg \inf} \mathbb{D}\Sigma\mathbb{D}',$$

s.t. $\Sigma = \mathbb{A}\Sigma\mathbb{A}' - \mathbb{A}\Sigma\mathbb{C}'\mathbb{C}\Sigma\mathbb{A}' / (\mathbb{C}\Sigma\mathbb{C}' + 1)$

where

$$(39) \quad \mathbb{A} := \left[\begin{array}{c|c} A & 0 \\ \hline GC & F \end{array} \right], \mathbb{C} := \begin{bmatrix} C & H \end{bmatrix}, \mathbb{D} := \begin{bmatrix} C & 0 \end{bmatrix},$$

$$A := \left[\begin{array}{c|c} 0_{n \times 1} & I_n \\ \hline \pm DI & \mathbf{a}_f \end{array} \right], C := \begin{bmatrix} 1 & 0_{1 \times n} \end{bmatrix}.$$

Note that we need to solve (38) twice (one for $+DI$ in A and one for $-DI$ in A), and choose the optimal solution as the one with the smaller objective function value. Then we form the optimal A^{opt} based on \mathbf{a}_f^{opt} , and let $(n^* + 1)$ be the number of unstable eigenvalues in A^{opt} , where $n^* \geq 0$.

Now let $n := n^*$, solve (38) again, and obtain a new \mathbf{a}_f^{opt} and Σ^{opt} . Then form A^{opt} , let $A^* = A^{opt}$, $\Sigma^* = \Sigma^{opt}$, $C^* := [1, 0_{1 \times n^*}]$, and form $\mathbb{A}^*, \mathbb{C}^*$, and \mathbb{D}^* . Let

$$(40) \quad L^* := \begin{bmatrix} L_1^* \\ L_2^* \end{bmatrix} := \frac{\mathbb{A}^*\Sigma^*\mathbb{C}^{*'}}{\mathbb{C}^*\Sigma^*\mathbb{C}^{*'} + 1}.$$

It holds that (A^*, C^*) is observable, and A^* has exactly $(n^* + 1)$ unstable eigenvalues.

We assign the encoder/decoder parameters to the scheme built in Fig. 4 (b) by letting

$$(41) \quad n := n^*, A := A^*, C := C^*, L_1 := L_1^*, L_2 := L_2^*.$$

We then drive the initial condition s_0 of channel \mathcal{F} to 0. Now we are ready to communicate at a rate \mathcal{R} using power $\mathbb{D}^*\Sigma^*\mathbb{D}^{*'}$, which is the minimum power needed to sustain the pre-specified rate \mathcal{R} .

The encoding/decoding processes

1) *Transmission of analog source*: Assume that the to-be-conveyed message W is distributed as $\mathcal{N}(0, I_{n^*+1})$ (noting that any non-degenerate (n^*+1) -variate Gaussian vector W can be transformed in this form). Assume that the coding length is $(K+1)$. To encode, let

$$(42) \quad x_0 := W + W_M,$$

where

$$(43) \quad W_M := -\sum_{j=0}^K A^{-j-1} L_1 \xi_j + \sum_{j=0}^K \sum_{i=0}^{j-1} A^{-j-1} L_1 H (F - L_2 H)^{k-1-j} L_2 \xi_j.$$

Then run the system till time epoch K . To decode, let $\hat{W}_K := \hat{x}_{0,K}$. The distortion is defined as

$$(44) \quad \text{MSE}(\hat{W}_K) := \mathbf{E}(W - \hat{W}_K)(W - \hat{W}_K)'$$

2) *Transmission of digital message*: To transmit digital messages over the communication system, let us first fix $\epsilon > 0$ small enough and the coding length $(K+1)$ large enough. Let

$$(45) \quad \Sigma_x^* := [I_{n^*+1}, 0] \Sigma^* [I_{n^*+1}, 0]'$$

Assume that the matrix $(A^{*'})^{-K-1} \Sigma_x^* (A^*)^{-K-1}$ has an eigenvalue decomposition as

$$(46) \quad (A^{*'})^{-K-1} \Sigma_x^* (A^*)^{-K-1} = E_K \Lambda_K E_K',$$

where $E_K = [e^{(1)}, \dots, e^{(n^*+1)}]$ is an orthonormal matrix and Λ_K is a positive diagonal matrix. Let $\sigma_{K,i}$ be the square root of the (i, i) th element of Λ_K . Let $\mathcal{B} \in \mathbb{R}^{n^*+1}$ be the hypercube spanned by columns of E_K , that is,

$$(47) \quad \mathcal{B} = \left\{ \sum_{i=0}^{n^*} \alpha^{(i)} e^{(i)} \mid \alpha^{(i)} \in \left[-\frac{1}{2}, \frac{1}{2}\right], i = 0, \dots, n^* \right\}.$$

Next we partition the i th side of \mathcal{B} into $(\sigma_{K,i})^{-(1-\epsilon)}$ segments. This induces a partition of \mathcal{B} into M_K sub-hypercubes, where

$$(48) \quad \begin{aligned} M_K &= \prod_{i=0}^{n^*} (\sigma_{K,i})^{-(1-\epsilon)} \\ &= [\det((A^{*'})^{-K-1} \Sigma_x^* (A^*)^{-K-1})]^{-\frac{1-\epsilon}{2}}. \end{aligned}$$

We then map the sub-hypercube centers to a set of M_K equally likely messages. The above procedure is known to both the transmitter and receiver *a priori*.

Suppose now we wish to transmit the message represented by the center W . To encode, define x_0 according to (42). Then run the system till time epoch K . To decode, we map $\hat{x}_{0,K}$ into the closest sub-hypercube center and obtain the decoded message \hat{W}_K . We declare an error if $\hat{W}_K \neq W$, and call a (an asymptotic) rate

$$(49) \quad \mathcal{R} := \lim_{K \rightarrow \infty} \frac{1}{K+1} \log M_K$$

achievable if the probability of error PE_K vanishes as K tends to infinity.

As we can see, the encoder/decoder design and the encoding/decoding processes are rather simple. The computation complexity involved in coding grows as $O(K+1)$.

3.3. Coding theorem.

THEOREM 3.1. *Construct the encoder/decoder shown in Fig. 4 (b) using n^* , A^* , C^* , L_1^* , L_2^* , and W_M . Then under the average power constraint $\mathbf{E}u^2 \leq \mathcal{P}$,*

i) The coding scheme transmits an analog source $W \sim \mathcal{N}(0, I_{n^+1})$ from the encoder to the decoder at rate $C^s(\mathcal{P})$, with MSE distortion $\text{MSE}(\hat{W}_T)$ achieving the optimal asymptotic rate-distortion tradeoff given by*

$$(50) \quad C^s(\mathcal{P}) = \lim_{T \rightarrow \infty} \frac{1}{2(T+1)} \log \frac{1}{\det \text{MSE}(\hat{W}_T)}.$$

ii) The coding scheme can transmit digital messages from the encoder to the decoder at a rate arbitrarily close to $C^s(\mathcal{P})$, with PE_K decaying to zero doubly exponentially.

REMARK 2. The main idea of this dirty-paper coding scheme is still the techniques used for the AWGN dirty-paper channel: The presence of process noise ξ_{k-1} at the encoder, and the presence of initial condition offset W_M . In addition, it still holds that the feedback communication problem is equivalent to a minimum-energy control problem and Kalman filtering problem; details are skipped for brevity.

Proof. Similar to the AWGN case, the interference ξ does not affect x_k and hence the channel input u_k asymptotically. In addition, the term in $\hat{x}_{0,K}$ associated with ξ is known to the encoder before the transmission, and therefore the encoder can offset x_0 to completely cancel the term associated with ξ . To facilitate our analysis, we rewrite the communication system as a control system and then establish the above claims.

We first define for the coding scheme shown in Fig. 4 (b) that

$$(51) \quad \begin{aligned} \tilde{s}_k &:= s_k - \bar{s}_k \\ \mathbb{X}_k &:= \begin{bmatrix} x_k \\ \tilde{s}_k \end{bmatrix}. \end{aligned}$$

Then the dynamics of the encoder becomes

$$(52) \quad \begin{cases} \mathbb{X}_k &= \mathbb{A}_{cl}\mathbb{X}_{k-1} - L N_{k-1} = \mathbb{A}\mathbb{X}_{k-1} - L e_{k-1} \\ e_{k-1} &= \mathbb{C}\mathbb{X}_{k-1} + N_{k-1} \\ u_k &= \mathbb{D}\mathbb{X}_k, \end{cases}$$

where $\mathbb{A}_{cl} := \mathbb{A} - LC$. Note that (52) is a control system (as indicated in Fig. 4 (b)), and is affected by the interference only through its initial condition \mathbb{X}_0 . We now have

$$(53) \quad \mathbb{X}_k = \mathbb{A}_{cl}^k \mathbb{X}_0 - \sum_{j=0}^{k-1} \mathbb{A}_{cl}^{k-j-1} L N_j.$$

By $u_k = \mathbb{C}\mathbb{X}_k$, any L stabilizing the control system ensures that, the channel input power is asymptotically equal to $\sum_{j=0}^{k-1} \mathbb{A}_{cl}^{k-j-1} L L' \mathbb{A}_{cl}^{k-j-1}$, independent of the interference. We can also verify that the term in $\hat{x}_{0,K}$ generated by x_0 is asymptotically x_0 .

In addition, the interference would generate an extra term in $\hat{x}_{0,K}$ given by

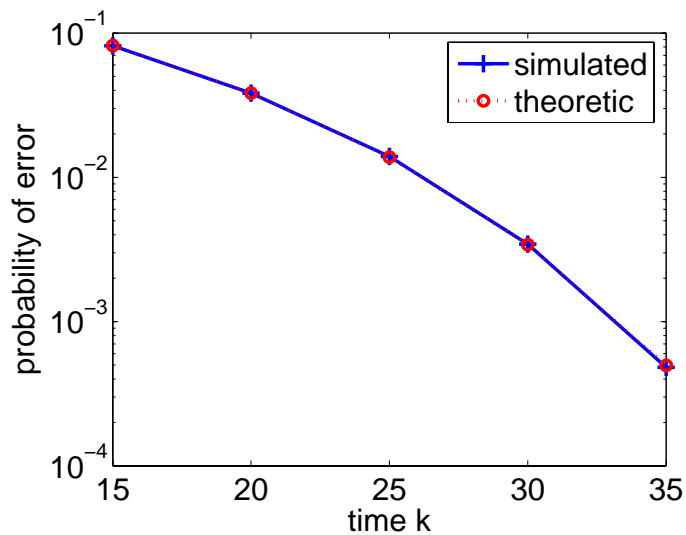
$$(54) \quad \sum_{j=0}^K A^{-j-1} L_1 \xi_j - \sum_{j=0}^K \sum_{i=0}^{j-1} A^{-j-1} L_1 H (F - L_2 H)^{k-1-j} L_2 \xi_j,$$

if no compensation to the initial condition x_0 is used. Then we offset x_0 by this amount and therefore we achieve (asymptotic) lossless cancelation of the interference. The proof for the coding scheme achieving $C^s(\mathcal{P})$ follows from the reasoning in [29].

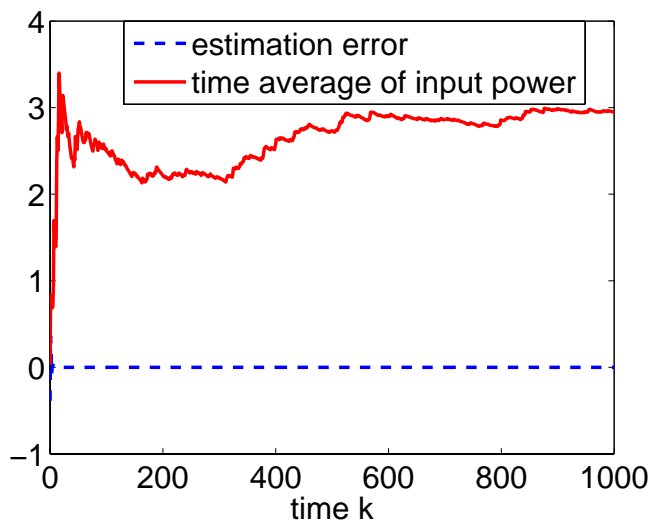
□

4. Numerical examples. In this section, we provide a numerical example of the proposed optimal coding scheme for AWGN dirty-paper channel. Assume that the power budget is $\mathcal{P} := 3$, that is, we may achieve any rate equal to $C(\mathcal{P}) - \epsilon = 1 - \epsilon$ bit per channel use for any $\epsilon > 0$. Let $\epsilon := 0.05$, i.e. the desired rate is $R = 0.95C(\mathcal{P}) = 0.95$ bit per channel use. Then we can construct the coding scheme according to Section 2.1.

Simulation shows that R is indeed achieved since the decoder can distinguish among $M_K := a^{(K+1)(1-\epsilon)}$ messages with the probability of error decaying to zero (in a doubly exponential fashion); see Fig. 5 (a). Due to the fast decay of probability of error, the coding length (and hence the coding delay) can be rather short (within 100) to attain a good performance. The shortened coding length also implies that the preview of the interference ξ^K can be short. Additionally, the average channel input power converges to \mathcal{P} , see Fig. 5 (b). For short coding length, however, the consumed power can be slightly different from the given power budget, since most of our analysis holds asymptotically. Fig. 5 (b) also shows the convergence of the decoder estimate \bar{W}_k to the message W .



(a)



(b)

FIG. 5. (a) Simulated probability of error and theoretic probability of error. (b) Convergence of the average channel input power, and vanishing of estimation error ($\bar{W}_K - W$).

5. Conclusions. In this paper, we have proposed capacity-achieving coding schemes for a dirty-paper channel with AWGN and for a dirty-paper channel with AWGN and ISI, both having noiseless output feedback. The interference is assumed to be known to the encoder non-causally before the transmission and unknown to the decoder, and can be arbitrarily varying, either deterministic or random. We achieved lossless interference cancelation, that is, we canceled the interference without extra

power overhead or rate loss. We exhibited the connections among feedback communication, feedback control, and estimation over a dirty-paper channel. In establishing lossless interference cancelation, we developed techniques which may be readily applied to more general dirty-paper channels. We also discussed the potential usefulness of our feedback communication schemes to sensor networks, in which we may wish to take advantage of the high-quality feedback link and the knowledge of interference to improve the forward communication. Future work includes the study of, first, multi-input multi-output communication systems with feedback, and second, channels with feedback and with causal interference information (i.e. dirty-tape channels with feedback), in which the perspective of investigating their control/estimation theoretic equivalences as well as the corresponding fundamental limitations (see [22, 29]) may play an important role.

REFERENCES

- [1] M. H. M. COSTA, *Writing on dirty paper*, IEEE Trans. Inform. Theory, 29:5(1983), pp. 439–441.
- [2] S. I. GEL'FAND AND M. S. PINSKER, *Coding for channel with random parameters*, Problems of Control and Information Theory, 9:1(1980), pp. 19–31.
- [3] F. M. J. WILLEMS, *On Gaussian channels with side information at the transmitter*, Proc. of the Ninth Symposium on Information Theory, 1988.
- [4] F. M. J. WILLEMS, *Signalling for the Gaussian channel with side information at the transmitter*, Proc. IEEE International Symposium on Information Theory (ISIT), page 348, 2000.
- [5] U. EREZ, S. SHAMAI (SHITZ), AND R. ZAMIR, *Capacity and lattice-strategies for cancelling known interference*, Proc. International Symposium on Information Theory and Applications (ISITA), Nov. 2000.
- [6] A. COHEN AND A. LAPIDOTH, *Generalized writing on dirty paper*, Proc. IEEE International Symposium on Information Theory (ISIT), page 227, 2002.
- [7] A. COHEN AND A. LAPIDOTH, *The Gaussian watermarking game*, IEEE Trans. Inform. Theory, 48:6(2002), pp. 1639–1667.
- [8] U. EREZ, S. SHAMAI (SHITZ), AND R. ZAMIR, *Capacity and lattice-strategies for cancelling known interference*, Submitted to IEEE Trans. Inform. Theory, 2002.
- [9] W. YU, A. SUTIVONG, D. JULIAN, T. COVER, AND M. CHIANG, *Writing on colored paper*, Proc. IEEE International Symposium on Information Theory (ISIT), page 302, 2002.
- [10] A. COHEN AND R. ZAMIR, *Writing on dirty paper in the presence of difference set noise*, Proc. 41st Annual Allerton Conference on Communication, Control and Computing, Oct. 2003.
- [11] R. ZAMIR AND A. COHEN, *The rate loss in writing on dirty paper*, DIMACS (Center for Discrete Mathematics and Theoretical Computer Science) Workshop on Network Information Theory, Mar. 2003.
- [12] G. CAIRE AND S. SHAMAI (SHITZ), *On the achievable throughput of a multi-antenna Gaussian broadcast channel*, IEEE Trans. Inform. Theory, 49:7(2003), pp. 1649–1706.
- [13] U. EREZ AND S. TEN BRINK, *A close-to-capacity dirty paper coding scheme*, IEEE Trans. Inform. Theory, 51:10(2005), pp. 3417–3432.
- [14] P. MOULIN AND J. O'SULLIVAN, *Information-theoretic analysis of information hiding*, Proc.

- IEEE International Symposium on Information Theory (ISIT), June 2000.
- [15] J. P. M. SCHALKWIJK AND T. KAILATH, *A coding scheme for additive noise channels with feedback, Part I: No bandwidth constraint*, IEEE Trans. Inform. Theory, IT-12:2(1966), pp. 172–182.
 - [16] J. P. M. SCHALKWIJK, *A coding scheme for additive noise channels with feedback, Part II: Bandlimited signals*, IEEE Trans. Inform. Theory, IT-12:2(1966), pp. 183–189.
 - [17] T. M. COVER AND S. POMBRA, *Gaussian feedback capacity*, IEEE Trans. Inform. Theory, IT-35(1989), pp. 37–43.
 - [18] S. TATIKONDA, *Control Under Communication Constraints*, PhD thesis, MIT, Cambridge, MA, Aug. 2000.
 - [19] V. S. BORKAR, S. K. MITTER, AND S. TATIKONDA, *Markov control problems under communication constraints*, Communications in Information and Systems, 1(2001), pp. 15–32.
 - [20] A. SAHAI, *Anytime Information Theory*, PhD thesis, MIT, Cambridge, MA, 2001.
 - [21] S. TATIKONDA AND S. MITTER, *Control under communication constraints*, IEEE Trans. Automat. Contr., 49:7(2004), pp. 1056–1068.
 - [22] N. ELIA, *When Bode meets Shannon: Control-oriented feedback communication schemes*, IEEE Trans. Automat. Contr., 49:9(2004), pp. 1477–1488.
 - [23] A. SAHAI AND S. MITTER, *The necessity and sufficiency of anytime capacity for control over a noisy communication link: Parts I and II*, Submitted to IEEE Trans. Inform. Theory, May 2005.
 - [24] S. YANG, A. KAVCIC, AND S. TATIKONDA, *Feedback capacity of power constrained Gaussian channels with memory*, Submitted to IEEE Trans. Inform. Theory, Nov. 2003. See also: *Linear Gaussian channels: Feedback capacity under power constraints*, Proc. IEEE International Symposium on Information Theory (ISIT), page 72, July 2004.
 - [25] A. SHAHAR-DORON AND M. FEDER, *On a capacity achieving scheme for the colored Gaussian channel with feedback*, Proc. 2004 IEEE International Symposium on Information Theory (ISIT), page 74, July 2004.
 - [26] YOUNG-HAN KIM, *The feedback capacity of the first-order moving average Gaussian channel*, Accepted by IEEE Trans. Inform. Theory, 2006. Available at <http://arxiv.org/abs/cs/0411036>.
 - [27] A. SAHAI, S. AVESTIMEHR, AND P. MINERO, *Anytime communication over the Gilbert-Eliot channel with noiseless feedback*, Proc. IEEE International Symposium on Information Theory (ISIT), 2005.
 - [28] J. LIU, N. ELIA, AND S. TATIKONDA, *Capacity-achieving feedback scheme for Markov channels with channel state information*, Submitted to IEEE Trans. Inform. Theory, Aug. 2005. See also: *Capacity-achieving Feedback Communication Scheme for Markov Channels with Channel State Information*, Proc. 2004 IEEE International Symposium on Information Theory (ISIT), page 71, June 2004.
 - [29] J. LIU AND N. ELIA, *Achieving the stationary feedback capacity for Gaussian channels with memory*, Submitted to IEEE Trans. Inform. Theory, Oct. 2005. See also: *Achieving the Stationary Feedback Capacity for Gaussian Channels*, Proc. 43rd Annual Allerton Conference on Communication, Control, and Computing, Sept. 2005. Available at <http://arxiv.org/abs/cs.IT/0512097>.
 - [30] N. MERHAV AND T. WEISSMAN, *Coding for the feedback Gel'fand-Pinsker channel and the feed-forward Wyner-Ziv source*, Proc. IEEE International Symposium on Information Theory (ISIT), 2005.
 - [31] H. KWAKERNAK AND R. SIVAN, *Linear Optimal Control Systems*, John Wiley & Sons, New York, 1972.
 - [32] J. LIU, *Fundamental Limits in Gaussian Channels with Feedback: Confluence of Communi-*

- tion, Estimation, and Control*, PhD thesis, Iowa State University, Ames, IA, Apr. 2006. Available at <http://archives.ece.iastate.edu/archive/00000219/>.
- [33] J. LIU AND N. ELIA, *Gaussian channels with feedback: fundamental limits and confluence of communication, estimation, and control*, Technical Report TR-2005-11-4, Iowa State University, Ames, IA, Dec. 2005. Available at <http://archives.ece.iastate.edu/archive/00000200/>.
- [34] M. GASTPAR, *On remote sources and channels with feedback*, Proc. 38th Annual Conference on Information Sciences and Systems (CISS), 2004.
- [35] R. G. GALLAGER, *Information Theory and Reliable Communication*, John Wiley and Sons, 1968.
- [36] T. KAILATH, A. SAYED, AND B. HASSIBI, *Linear Estimation*, Prentice Hall, 2000, Englewood Cliffs, NJ, 2000.
- [37] S. K. MITTER AND N. NEWTON, *Information and entropy flow in the Kalman-Bucy filter*, J. of Stat. Phys., 118(2005), pp. 145–176.
- [38] YOUNG-HAN KIM, *On the feedback capacity of stationary Gaussian channels*, Proc. 43rd Annual Allerton Conference on Communication, Control, and Computing, Sept. 2005. Available at <http://arxiv.org/abs/cs/0509078>.
- [39] YOUNG-HAN KIM, *Feedback capacity of stationary Gaussian channels*, Submitted to IEEE Trans. Inform. Theory, Feb. 2006. Available at <http://arxiv.org/abs/cs/0602091>.

