result is

$$S'_{B_H}(\omega; a) = a|G(a\omega)|^2 \cdot \frac{1}{|\omega|^{2H+1}}, \tag{27}$$

which corresponds to the FBM spectrum, as seen through the filter of nominal scale a.

In fact, since (22) is related to $x(t)$ by a linear filter with impulse response

$$h_a(t) = \frac{1}{\sqrt{a}} g\left(-\frac{t}{a}\right), \tag{28}$$

it would have been possible to obtain the result of (27) by using the fact that, in this case,

$$S'_{B_H}(\omega; a) = \int_{-\infty}^{+\infty} W_{h_a}(t-s, \omega) W_{B_H}(s, \omega)\, ds. \tag{29}$$

Equation (29) results from the compatibility of the WVS with linear filtering and its reduction to power spectral density in the case of stationary processes [12]. It can be checked by direct computation that substitution of (6) and (28) into (29) leads to (27).

Since the spectral content of FBM is characterized relative to each scale by (27), it is at last possible to obtain a global spectrum by adding up the contributions pertaining to each scale. Taking into account the natural measure associated with the energy distribution of the wavelet transform in the scale direction [6], we end up with the result

$$S'_{B_H}(\omega) = \int_0^{+\infty} S'_{B_H}(\omega; a) \cdot \frac{da}{a^2} = \frac{1}{|\omega|^{2H+1}}, \tag{30}$$

which is again the desired behavior for a spectrum of FBM.

V. Conclusion

Two approaches have been proposed for the analysis of FBM's. The first one focused on their nonstationary character; by means of a time–frequency analysis, it was shown that their spectral behavior can be described in terms of average spectra. The second approach focused on self-similarity properties of FBM's; by means of a time–scale analysis, a decomposition in terms of self-similar stationary processes was effected. In both cases, the description of a global spectral behavior of FBM's relative to a measurement process (either a time average or a scale filter) was achieved, providing new insight into investigations of such processes.

References

[1] R. J. Barton and H. V. Poor, "Signal detection in fractional Gaussian noise," *IEEE Trans. Inform. Theory*, vol. 34, pp. 943–959, Sept. 1988, Part I.

[2] I. Daubechies, "The wavelet transform, time–frequency localization and signal analysis," preprint, AT&T Bell Labs., 1987.

[3] P. Flandrin, "On the positivity of the Wigner–Ville spectrum," *Signal Proc.*, vol. 11, no. 2, pp. 187–189, 1986.

[4] A. Grossmann and J. Morlet, "Decomposition of Hardy functions into square integrable wavelets of constant shape," *SIAM J. Math. Anal.*, vol. 15, no. 4, pp. 723–736, 1984.

[5] J. M. Keller, R. M. Crownover, and R. Y. Chen, "Characteristics of natural scenes related to the fractal dimension," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, no. 5, pp. 621–627, 1987.

[6] R. Kronland-Martinet, J. Morlet, and A. Grossmann, "Analysis of sound patterns through wavelet transforms," *Int. J. Pattern Recogn. Artif. Intell.*, vol. 1, no. 2, pp. 97–126, 1987.

[7] R. M. Loynes, "On the concept of the spectrum for nonstationary processes," *J. Roy. Stat. Soc. B*, vol. 30, no. 1, pp. 1–20, 1968.

[8] T. Lundhal, W. J. Ohley, S. M. Kay, and R. Siffert, "Fractional Brownian motion: A maximum likelihood estimator and its application to image texture," *IEEE Trans. Med. Imaging*, vol. MI-5, no. 3, pp. 152–161, 1986.

[9] B. Mandelbrot, "Some noises with $1/f$ spectrum, a bridge between direct current and white noise," *IEEE Trans. Inform. Theory*, vol. IT-13, no. 2, pp. 289–298, 1967.

[10] ——, *The Fractal Geometry of Nature*. San Francisco, CA: Freeman, 1983.

[11] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM Rev.*, vol. 10, no. 4, pp. 422–437, 1968.

[12] W. Martin and P. Flandrin, "Wigner-Ville spectral analysis of nonstationary processes," *IEEE Trans. Acoust., Speech, Sig. Proc.*, vol. ASSP-33, no. 6, pp. 1461–1470, 1985.

[13] E. W. Montroll and M. F. Shlesinger, "On $1/f$ noise and other distributions with long tails," *Proc. Nat. Acad. Sci. USA*, vol. 79, pp. 3380–3383, 1982.

[14] H. Takayasu, "$f^{-\beta}$ power spectrum and stable distribution," *J. Phys. Soc. Jap.*, vol. 56, no. 4, pp. 1257–1260, 1987.

Optimum Linear Causal Coding Schemes for Gaussian Stochastic Processes in the Presence of Correlated Jamming

TANGÜL Ü. BAŞAR, MEMBER, IEEE, AND TAMER BAŞAR, FELLOW, IEEE

***Abstract*—The complete solution is obtained to the following problem. A Gaussian stochastic process $\{\theta_t, t \in [0, t_f]\}$ satisfying a certain stochastic differential equation is to be transmitted through a stochastic channel to a receiver under minimum mean-squared error distortion measure. The channel is to be used for exactly t_f seconds and, in addition to white Gaussian noise with a given energy level, the channel is corrupted by another source whose output may be correlated with the input to the channel and which satisfies a given power constraint. There is an input power constraint to the channel, and noiseless feedback is allowed between the receiver (decoder) and the transmitter (encoder). We determine the linear causal encoder and decoder structures that function optimally under the worst admissible noise inputs to the channel. The least favorable probability distribution for this unknown noise is found to be Gaussian and is correlated with the transmitted signal. Also included is a comparative study of these results with earlier ones that addressed a similar problem without a causality restriction imposed on the transmitter.**

I. Introduction and Problem Description

Consider the communication system depicted in Fig. 1. A stochastic process $\{\theta_t\}$, $t \in [0, t_f]$, which satisfies the equation

$$d\theta_t = a(t)\theta_t\, dt + b(t)\, dV_t; \qquad \theta_0 \sim N(0, \gamma), \tag{1}$$

where V_t is a standard Wiener process and is independent of θ_0, and $a(t)$ and $b(t)$ are uniformly bounded (i.e., for some $k > 0$, $|a(t)| \le k$, $|b(t)| \le k$), is to be transmitted through a continuous-time stochastic channel which is corrupted by additive white Gaussian noise and is also tapped by an intelligent jammer. The jammer sends a stochastic process (possibly correlated with the encoded message process) at a certain power level so as to maximize the interference to the channel. The transmission channel is to be used for exactly t_f seconds and has an instantaneous

Manuscript received July 10, 1987; revised August 2, 1987. This work was supported in part by the Joint Services Electronics Program under Contract N00014-84-C-0149 through the University of Illinois. This paper was presented at the 1986 Conference on Information Sciences and Systems, Princeton University, Princeton, NJ, March 19–21.

T. Ü. Başar is with the Electrical and Computer Engineering Department, Illinois Institute of Technology, Chicago, IL 60616.

T. Başar is with the Coordinated Science Laboratory and the Department of Electrical and Computer Engineering, University of Illinois, 1101 West Springfield Avenue, Urbana, IL 61801.

IEEE Log Number 8825353.

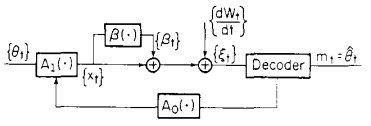

Fig. 1. Continuous-time communication system.

input power constraint P, i.e.,

$$E\left[x_t^2\right] \le P. \tag{2}$$

Its output process ξ_t, $t \in [0, t_f]$, is described by

$$d\xi_t = x_t\,dt + \beta_t\,dt + dW_t, \qquad \xi_0 = 0. \tag{3}$$

Here $\{W_t,\ t \in [0, t_f]\}$ is a standard Wiener process, $\{\beta_t,\ t \in [0, t_f]\}$ is a second-order stochastic process determined by the jammer and is subject to the power constraint

$$E\left[\beta_t^2\right] \le K \tag{4}$$

where $K < P$, and x_t (the input to the channel) is determined by two (nonanticipatory) coding functions $A_0(t; \xi_s,\ s \le t)$ and $A_1(t; \xi_s,\ s \le t)$ such that

$$x_t = A_0(t; \xi_s,\ s \le t) + A_1(t; \xi_s,\ s \le t)\,\theta_t. \tag{5}$$

Note that this imposes a memoryless linear structure on the encoder, in terms of its dependence on $\{\theta_t\}$; however, the dependence on $\{\xi_t\}$ through the feedback channel could be arbitrary, provided that it is nonanticipatory. Furthermore, the dependence of $\{\beta_t\}$ on the message process $\{\theta_t\}$ is only through the output of the encoder $\{x_t\}$, again under a nonanticipatory transformation.

Finally, the distortion measure adopted for this communications problem is

$$D(A_0, A_1, \hat{\theta}; \beta) = E\left\{\int_0^{t_f}\left[\theta_t - \hat{\theta}_t(\xi_s,\ s \le t)\right]^2 dt\right\} \tag{6}$$

which is to be minimized by the choice of $A_0, A_1, \hat{\theta} := \{\hat{\theta}_t\}$ and maximized by the choice of $\beta := \{\beta_t\}$. Note that $\hat{\theta}$ is the output of the decoder, which depends on $\{\xi_t\}$ in a nonanticipatory manner. Hence a natural criterion here would be the upper value of the corresponding game:

$$\min_{A_0, A_1, \hat{\theta}} \max_{\beta} D(A_0, A_1, \hat{\theta}; \beta). \tag{7a}$$

In this correspondence we seek and prove the stronger result that a quadruple $A_0^*, A_1^*, \hat{\theta}^* := \{\hat{\theta}_t^*,\ t \in [0, t_f]\}, \beta^* = \{\beta_t^*,\ t \in [0, t_f]\}$ exists satisfying the pair of saddle-point inequalities

$$\begin{aligned} D(A_0^*, A_1^*, \hat{\theta}^*; \beta) &\le D(A_0^*, A_1^*, \hat{\theta}^*; \beta^*) \\ &\le D(A_0, A_1, \hat{\theta}; \beta^*) \end{aligned} \tag{7b}$$

A number of historical remarks are now in order to place the contributions of this correspondence in proper perspective. First, a different version of this problem was studied earlier [1] where no causality restriction was imposed on the encoder, which also premitted the consideration of a more general structure for the encoder, basically that of all measurable functions of the message process $\{\theta_t\}$. The saddle-point solution obtained there was indeed strictly noncausal for nontrivial stochastic processes, and it involved a Karhunen–Loeve expansion of the message process, storage of these data, and transmitting only a few leading terms in this expansion. Hence, under a causality restriction, it is natural to expect the solution to be different than the one given in [1], which indeed is the case as to be shown in the sequel. We should note, however, that since the optimum design for the jammer-free version of the problem (i.e., which involves an extension of Ihara's result [2] from random variable input to stochastic process message processes, under the causality restriction) is known to be linear under some regularity conditions [5], it seems possible to relax the structural restriction (5) on the encoder and still obtain the result presented in this paper.

Second, one of our basic assumptions above has been that the jammer can choose its policy as a function of the output of the encoder but not directly as a function of its input; in other words, the jammer is allowed to tap the channel after the encoder (and not before). This assumption provides us some convenience in proving the basic result presented in Theorem 1; however, it brings in no loss of generality in the following sense. As will be shown in Corollary 1 just following Theorem 1, the basic result remains intact if the jammer *also* has access to the past values of $\{\theta_t\}$. In other words, this additional information has no value, since it passes through a causal transformation the output of which is the relevant quantity already known to the jammer. On the other hand, if the jammer has access *only* to $\{\theta_t\}$ and *not* $\{x_t\}$, the result would no longer hold, because then a random choice for the encoder transformation A_1 (as hinted by the results of [3], derived primarily for static channels) would nullify the useful information content of $\{\theta_t\}$. A detailed study of this problem will be presented in a future paper.

II. Optimum Coding–Decoding Schemes

The solution to the problem in Section I is provided in the following theorem and is depicted in Fig. 2.

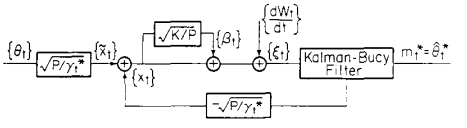

Fig. 2. Optimum structure for communication system.

Theorem 1: The communication system described in Section I and depicted in Fig. 1 admits a solution which is given as follows.

1) The optimum coding functions are

$$A_1^*(t) = \sqrt{P/\gamma_t^*} \tag{8}$$

$$A_0^*(t) = -A_1^*(t)\,m_t^* \tag{9}$$

where $\{m_t^*\}$ is the estimator process satisfying the stochastic differential equation

$$dm_t^* = a(t)\,m_t^*\,dt + \sqrt{P\gamma_t^*}\,d\xi_t^*, \qquad m_0^* = 0, \tag{10a}$$

with $\{\gamma_t^*\}$ the unique solution of

$$\dot{\gamma}_t^* = \left[2a(t) - \tilde{P}\right]\gamma_t^* + b^2(t), \qquad \gamma_0^* = \gamma \tag{10b}$$

and

$$\tilde{P} := \left(1 - \sqrt{K/P}\right)^2 P. \tag{10c}$$

2) The optimum decoder output is

$$\hat{\theta}_t^* = m_t^*, \qquad t \in [0, t_f]. \tag{11}$$

3) The maximizing jammer policy is

$$\beta^*(t, x_s,\ s \le t) = -\sqrt{K/P}\,x_t^* =: \beta_t^* \tag{12}$$

where

$$x_t^* = A_0^* + A_1^*\theta_t. \tag{13}$$

4) The worst-case distortion level is

$$D(A_0^*, A_1^*, \hat{\theta}^*; \beta^*) = \int_0^{t_f}\gamma_t^*\,dt =: D^*. \tag{14}$$

Proof: The proof proceeds in two steps. We first establish the validity of the right-hand side (RHS) inequality of (7b) when β^* is given by (12), and then prove the left-hand side (LHS)

inequality of (7b) when A_0^*, A_1^* and $\hat{\theta}_t^*$ are given by (8), (9), and (11) respectively.

a) The RHS Inequality: With $\beta_t = -(\sqrt{K/P})x_t$, the channel output (3) can be written as

$$\begin{aligned} d\xi_t &= x_t\,dt - \sqrt{K/P}\,x_t\,dt + dW_t \\ &= \left(1-\sqrt{K/P}\right)x_t\,dt + dW_t \end{aligned} \tag{15}$$

where

$$x_t = A_0(t) + A_1(t)\theta_t.$$

For each fixed A_0, A_1, let

$$\gamma_t = \min_{\hat{\theta}_t} E\left\{\left[\theta_t - \hat{\theta}_t\right]\right\}^2.$$

Then the Kalman–Bucy theory [4, p. 246] is directly applicable, and the minimizing $\hat{\theta}_t$ is m_t, where m_t satisfies the following stochastic differential equation:

$$\begin{aligned} dm_t &= a(t)m_t\,dt + \gamma_t\left(1-\sqrt{K/P}\right) \\ &\quad \cdot A_1\left[d\xi_t - \left(1-\sqrt{K/P}\right)\left(A_0(t) + A_1(t)m_t\right)dt\right] \end{aligned} \tag{16}$$

$$\dot{\gamma}_t = 2a(t)\gamma_t - \gamma_t^2\left(1-\sqrt{K/P}\right)^2 A_1^2 + b^2 \tag{17}$$

with $m_0 = 0$, $\gamma_0 = \gamma$. From the channel power constraint (2), we can find

$$\begin{aligned} &E\left[A_0 + A_1\theta_t\right]^2 = E\left[\left(A_0 + m_t A_1\right) + \left(\theta_t - m_t\right)A_1\right]^2 \le P \\ &\quad \Leftrightarrow E\left\{\left(A_0 + A_1 m_t\right)^2\right\} + E\left\{\gamma_t A_1^2\right\} \le P. \end{aligned} \tag{18}$$

Note that (17) is equivalent to the integral equation

$$\begin{aligned} \gamma_t &= \gamma\exp\left\{2\int_0^t a(s)\,ds - \left(1-\sqrt{K/P}\right)^2\int_0^t \gamma_s A_1^2(s,\xi)\,ds\right\} \\ &\quad + \int_0^t b^2(s)\exp\left\{2\int_s^t a(u)\,du - \left(1-\sqrt{K/P}\right)^2\right. \\ &\quad \left.\cdot\int_s^t \gamma_u A_1^2(u,\xi)\,du\right\}ds. \end{aligned} \tag{19}$$

By Jensen's inequality ($E[e^{-\eta}] \ge e^{-E[\eta]}$), we have

$$\begin{aligned} \gamma_t &\ge \gamma\exp\left\{2\int_0^t a(s)\,ds - \left(1-\sqrt{K/P}\right)^2\int_0^t E\left[\gamma_s A_1^2\right]ds\right\} \\ &\quad + \int_0^t b^2(s)\exp\left\{2\int_s^t a(u)\,du - \left(1-\sqrt{K/P}\right)^2\right. \\ &\quad \left.\cdot\int_s^t E\left[\gamma_u A_1^2(u,\xi)\,du\right]\right\}ds \end{aligned} \tag{20}$$

which, together with the inequality $E[\gamma_t A_1^2(t,\xi)] \le P$ (following from (18)) yields the following lower bound for $E[\gamma_t]$:

$$\begin{aligned} E[\gamma_t] &\ge \gamma\exp\left\{2\int_0^t\left(a(s) - \frac{\tilde{P}}{2}\right)ds\right\} \\ &\quad + \int_0^t b^2\exp\left\{2\int_s^t\left(a(u) - \frac{\tilde{P}}{2}\right)du\right\}ds. \end{aligned} \tag{21}$$

Following [5], we choose

$$A_0^*(t,\xi_t) = -A_1^*(t,\xi_t)m_t^* \tag{22}$$

$$A_1^*(t,\xi_t) = \sqrt{P/\gamma_t^*} \tag{23}$$

where

$$\gamma_t^* = E[\theta_t - m_t^*]^2, \qquad 0 \le t \le t_f.$$

ξ_t^* is the solution of

$$d\xi_t^* = \sqrt{P/\gamma_t^*}\,(\theta_t - m_t^*)\,dt + dW_t, \qquad \xi_0 = 0, \tag{24}$$

and m_t^*, γ_t^* are given as

$$dm_t^* = a(t)m_t^*\,dt + \sqrt{P\gamma_t^*}\,d\xi_t^*, \qquad m_0^* = 0 \tag{25}$$

$$\dot{\gamma}_t^* = \left(2a(t) - \tilde{P}\right)\gamma_t^* + b^2(t), \qquad \gamma_0^* = \gamma \tag{26}$$

$$\tilde{P} = \left(1-\sqrt{K/P}\right)^2 P. \tag{27}$$

The linear equation (26) has the unique solution

$$\begin{aligned} \gamma_t^* &= \gamma\exp\left\{2\int_0^t\left(a(s) - \frac{\tilde{P}}{2}\right)ds\right\} \\ &\quad + \int_0^t b^2\exp\left\{\left(a(u) - \frac{\tilde{P}}{2}\right)du\right\}ds \end{aligned} \tag{28}$$

which coincides with the bound given in (21), and thus its integral over $(0, t_f)$ gives us the realizable minimum error. This completes the proof of validity of the RHS inequality of (7).

b) The LHS Inequality: Let $A_0^*(t,\xi)$ and $A_1^*(t,\xi)$ be defined by (8), (9). Then,

$$x_t = A_1^*(t,\xi)[\theta_t - m_t] = \sqrt{P/\gamma_t^*}\,[\theta_t - m_t]$$

where m_t satisfies the stochastic differential equation

$$\begin{aligned} dm_t &= a(t)m_t\,dt + \gamma_t^*\tilde{A}_1^*(t,\xi)\,d\xi_t \\ m_0 &= 0 \end{aligned} \tag{29}$$

and

$$\tilde{A}_1^*(t,\xi_1) \triangleq \left(1-\sqrt{K/P}\right)A_1^*(t,\xi_t).$$

Since

$$d\xi_t = A_1^*(t,\xi)[\theta_t - m_t]\,dt + \beta_t\,dt + dW_t, \tag{30}$$

we obtain

$$\begin{aligned} dm_t &= a(t)m_t\,dt + \gamma_t^*\tilde{A}_1^*A_1^*[\theta_t - m_t]\,dt \\ &\quad + \gamma_t^*\tilde{A}_1^*\beta_t\,dt + \gamma_t^*\tilde{A}_1^*\,dW_t. \end{aligned} \tag{31}$$

Let $p_t = m_t - \theta_t$; then p_t satisfies the stochastic differential equation

$$\begin{aligned} dp_t &= a_t p_t\,dt - \gamma_t^*\tilde{A}_1^*A_1^*p_t\,dt + \gamma_t^*\tilde{A}_1^*\beta_t\,dt + \gamma_t^*\tilde{A}_1^*\,dW_t - b_t\,dV_t \\ p_0 &= -\theta_0. \end{aligned} \tag{32}$$

Note the the distortion measure (6) can be rewritten as

$$D = E\left[\int_0^{t_f} p_t^2\,dt\right].$$

We now let $M_t = p_t^2$ and obtain the following stochastic differential equation for M_t, by direct application of Ito's formula

$$\begin{aligned} dM_t &= 2a_t M_t\,dt - 2p_t b_t\,dV_t - 2\gamma_t^*\tilde{A}_1^*A_1^*M_t\,dt \\ &\quad + 2\gamma_t^*\tilde{A}_1^*\beta_t p_t\,dt + 2\gamma_t^*\tilde{A}_1^*p_t\,dW_t + b_t^2\,dt + \gamma_t^{*2}\tilde{A}_1^{*2}\,dt \\ M_0 &= \theta_0^2. \end{aligned} \tag{33}$$

Finally, letting $\epsilon(t) = E[M_t]$, we arrive at the following differential equation describing the evolution of $E[M_t]$:

$$\dot{\epsilon} = 2a_t\epsilon - 2\gamma_t^*\tilde{A}_1^*A_1^*\epsilon + 2\gamma_t^*\tilde{A}_1^*E[\beta_t p_t] + b^2 + \gamma_t^{*2}\tilde{A}_1^{*2}. \tag{34}$$

Now, $\{\beta_t\}$ should be chosen so as to maximize $\int_0^{t_f}\epsilon(t)\,dt$ [which is $D(A_1^*, A_0^*, \hat{\theta}^*; \beta)$] under the power constraint (4). Since $\gamma_t^*\tilde{A}_1^*$ is positive and the power constraint is instantaneous, maximizing $\epsilon(t)$ for each t is equivalent to maximizing the quantity $E[\beta_t p_t]$.

Noting the bound (from Cauchy–Schwartz inequality)

$$\{E[\beta_t p_t]\}^2 \le E[\beta_t^2]\,E[p_t^2] \le K\epsilon(t), \tag{35}$$

the unique maximizing solution is

$$\beta_t^* = \sqrt{K/\epsilon_t}\,p_t \tag{36}$$

and hence

$$\max_{E[\beta_t^2]\le K} E[\beta_t p_t] = \sqrt{K\epsilon_t}\,.$$

Substituting this in (34) and also noting the equality $\epsilon(t) = \gamma_t^*$ leads to the differential equation

$$\dot{\gamma}_t^* = (2a_t - \tilde{P})\gamma_t^* + b^2(t); \qquad \gamma_0^* = \gamma$$

which is identical to (26). In view of this, (36) can be rewritten as

$$\beta_t^* = \sqrt{K/P}\,A_1^* p_t,$$

and since

$$x_t = A_1^*[\theta_t - m_t],$$

we finally obtain

$$\beta_t^*(x_t) = -\sqrt{K/P}\,x_t, \qquad t \in [0, t_f]$$

as the unique jammer policy that maximizes $D(A_1^*, A_0^*, \hat{\theta}^*; \beta)$. This then validates the LHS inequality of (7).

Corollary 1: The result of the theorem remains intact if the stochastic process $\{\beta_t\}$ is taken to be adapted to the larger sigma field, generated by $\{\theta_t\}$ and $\{x_t\}$, i.e.,

$$\beta_t = \beta(t; x_s, \theta_s, s \le t).$$

Proof: All we need to show is the validity of the LHS inequality, since for the RHS inequality the choice $\beta_t = -\sqrt{K/P}\,x_t$ generates a stochastic process that is adapted to the new information field. Now, for the LHS inequality, we proceed as in the proof of the theorem until we reach (34). Here, maximization of $\int_0^{t_f}\epsilon(t)\,dt$ under the power constraint (4) is gain equivalent to maximizing $E[\beta_t p_t]$ for each t, whose unique solution under the new (enlarged) functional dependence is still (36); in other words, the additional knowledge of $\{\theta_t\}$ is redundant information as far as the jammer's maximization problem is concerned.

Remark 1: When the message to be transmitted is a Gaussian random variable (i.e., $a = b = 0$ in the formulation of Section I), the saddle-point distortion level $\gamma_{t_f}^*$ becomes (from (10)):

$$\gamma_{t_f}^* = \gamma \exp\{-\tilde{P}t_f\}$$

which is a result already known from [1]. Hence the distortion levels obtained in [1] and in this correspondence coincide for this special case (as expected). However, if either $a \neq 0$ or $b \neq 0$, D^* of this correspondence should be strictly larger than its counterpart in [1], because of the causality restriction. This is clearly illustrated in the example of the next section.

III. An Example

Let the message process θ_t, $t \ge 0$ be the standard Wiener process V_t, $t \ge 0$; that is

$$\theta_t = V_t \qquad \theta_0 \equiv 0$$

(in the formulation of Section I, $a(t) \equiv 0$, $b(t) \equiv 1$, $\gamma = 0$). Furthermore, suppose that

$$t_f = 1 \qquad K = 4 \qquad P = (2+\sqrt{5})^2 \cong 17.94.$$

Then, the integral (28) can be evaluated to give

$$\gamma_t^* = \int_0^t \exp\{(s-t)\tilde{P}\}\,ds = \frac{1}{4.998}[1 - e^{-4.998t}]$$

leading to

$$D^* = \int_0^1 \gamma_t^*\,dt = \frac{1}{4.998}\left[1 - \frac{1}{4.998}(1 - e^{-4.998})\right] \cong 0.160, \quad \text{(i)}$$

which is the minimum guaranteed distortion level. If, on the other hand, we use the solution given in [1], we readily obtain a guaranteed distortion level of

$$D^* \cong 0.074,$$

which is considerably smaller than (i), the difference being the loss due to causality. In both cases, the least favorable jammer policy is

$$\beta_t^*(x_t) \cong -0.472 x_t.$$

The input to the channel, on the other hand, is

$$\tilde{x}_t^* = 9.469(1 - e^{-4.998t})^{1/2}\theta_t$$

for the causal system of this correspondence, whereas for the noncausal system of [1],

$$\tilde{x}_t^* = \sum_{i=1}^{3} a_i(t)\,b_i$$

where

$$b_i = \left(\left(i - \frac{1}{2}\right)^2 \Pi^2\right)^{-1} \int_{-1}^{0} \sin\left[\left(i - \frac{1}{2}\right)\Pi t\right]\theta_t\,dt\sqrt{2}$$

$$a_i(t) \cong \begin{cases} 6.653\exp[5t/2], & 0 \le t \le 0.694 \\ 19.961\exp[5(t-0.694)/2], & 0.694 < t \le 0.949 \\ 33.270\exp[5(t-0.949)/2], & 0.949 < t \le 1. \end{cases}$$

Note that in this latter case we assume that the entire sample path of the process is available at the encoder prior to $t = 0$. Furthermore, following a transmission of two time units, the signal estimate is reconstructed at the output of the decoder, again for a period of one time unit. Hence the processing interval is, in fact, at least three times the length of that in the former problem.

References

[1] T. Ü. Başar and T. Başar, "Optimum coding and decoding schemes for the transmission of a stochastic process over a continuous-time stochastic channel with partially unknown statistics," *Stochastics*, vol. 8, pp. 2213–237, 1982.

[2] S. Ihara, "Optimal coding in white Gaussian channel with feedback," in *Proc. 2nd Japan–USSR Symp. Probability Theory* (published in *Lecture Notes in Math Series*, no. 330. Berlin: Springer-Verlag, 1973, pp. 120–123).

[3] T. Başar and Y.-W. Wu, "A complete characterization of minimax and maximin encoder–decoder policies for communication channels with incomplete statistical description," *IEEE Trans. Inform. Theory*, vol. IT-31, no. 4, pp. 482–489, 1985.

[4] K. Astrom, *Stochastic Control*. New York: Academic, 1970.

[5] R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes. II Applications*. New York: Springer-Verlag, 1978.