Charalambous, Charalambos D.; Kourtellaris, Christos K.; Charalambous, Themistoklis

A General Coding Scheme for Signaling Gaussian Processes over Gaussian Decision Models

Published in:
Proceedings of the IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2018

DOI:
10.1109/SPAWC.2018.8445982

Published: 24/08/2018

Document Version
Peer reviewed version

Please cite the original version:
Charalambous, C. D., Kourtellaris, C. K., & Charalambous, T. (2018). A General Coding Scheme for Signaling Gaussian Processes over Gaussian Decision Models. In *Proceedings of the IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications, SPAWC 2018* (Vol. 2018-June). [8445982] (IEEE International Workshop on Signal Processing Advances in Wireless Communications). IEEE. https://doi.org/10.1109/SPAWC.2018.8445982

A General Coding Scheme for Signaling Gaussian Processes over Gaussian Decision Models

Charalambos D. Charalambous, Christos K. Kourtellaris and Themistoklis Charalambous

***Abstract*— In this paper, we transform the $n-$finite transmission feedback information (FTFI) capacity of unstable Gaussian decision models with memory on past outputs, subject to an average cost constraint of quadratic form derived in [1], into controllers-encoders-decoders that control the output process, encode a Gaussian process, reconstruct the Gaussian process via a mean-square error (MSE) decoder, and achieve the $n-$FTFI capacity. For a Gaussian RV message $X \sim N(0, \sigma_X^2)$ it is shown that the MSE decays according to $\mathbf{E}|X - \widehat{X}_{n|n}|^2 = \exp\{-2C_{0,n}(\kappa)\}\sigma_X^2$, $\kappa \in (\kappa_{min}, \infty)$, where $C_{0,n}(\kappa)$ is the $n-$FTFI capacity, and κ_{min} is the threshold on the power to ensure convergence.**

***Index Terms*— Coding, $n-$finite transmission feedback information, Gaussian decision models, capacity.**

I. Introduction

It has been recently shown [2] that randomized strategies in decision systems are operational, in the sense that not only they stabilize the system but they also encode information, which can be decoded at the output of the control system with arbitrary small probability of decoding error. In other words, the control system is used to communicate information, with an operational meaning as defined by Shannon's capacity of communication channels.

Our main objective of this paper is to synthesize controller-encoder strategies for Gaussian recursive models (G-RMs), with input process $\{A_t : t = 0, 1, \ldots, n\}$ that simultaneously control the output process $\{Y_t : t = 0, 1, \ldots, n\}$, encode an information process $\{X_t : t = 0, 1, \ldots, n\}$, and synthesize a decoder at the output $\{Y_t : t = 0, 1, \ldots, n\}$, such that the process $\{X_t : t = 0, 1, \ldots, n\}$ is transmitted reliably to the decoder.

If the G-RM is a control system model, then the system, as depicted in Figure 1, gives rise to several application. One application is that of tracking the dynamics of $\{X_t : t = 0, 1, \ldots, n\}$ at the output of the decoder. If $\{X_t : t = 0, 1, \ldots, n\}$ is generated by a discrete deterministic recursion, then we show that it is possible for the decoder to track $\{X_t : t = 0, 1, \ldots, n\}$ with arbitrary small MSE. This is contrary to standard tracking design systems in which the noise of the G-RM, i.e., $\{V_t : t = 0, 1, \ldots, n\}$ imposes limitations on the minimum tracking error. Another application is that of signaling digital messages available to the controller, such as, values associated with actuating devices, for failure detection and monitoring applications.

C. D. Charalambous and C. K. Kourtellaris are with the Department of Electrical Engineering, University of Cyprus, Nicosia, Cyprus. E-mails: {chadcha,kourtellaris.christos}@ucy.ac.cy.

T. Charalambous is with the Department of Electrical Engineering and Automation, School of Electrical Engineering, Aalto University, Finland. E-mail: themistoklis.charalambous@aalto.fi

II. Problem Statement and Related Work

A. Notation

$\langle \cdot, \cdot \rangle$ denotes inner product of elements of linear spaces, $\mathbb{S}_+^{q\times q}$ denotes the set of symmetric positive semi-definite $q\times q$ matrices and $\mathbb{S}_{++}^{q\times q}$ the subset of positive definite matrices, with real entries.

B. Main Problem

We consider a multiple-input multiple output (MIMO) time-verying unstable G-RM given by

$$Y_i = C_{i-1}\ Y_{i-1} + D_i\ A_i + V_i,\ Y_{-1} = s, i = 0, \ldots, n, \quad (1)$$

subject to an average cost of quadratic form described by

$$\frac{1}{n+1}\mathbf{E}_s\Big\{\sum_{i=0}^n \gamma_i(A_i, Y_{i-1})\Big\} \leq \kappa, \quad (2)$$

$A^n \triangleq \{A_0, A_1, \ldots, A_n\}$ is the input process,
$Y^n \triangleq \{Y_0, A_1, \ldots, Y_n\}$ is the output process,
$V^n \triangleq \{V_0, \ldots, V_n\}$ is an independent zero mean Gaussian noise process, denoted by $V_i \sim N(0, K_{V_i}), K_{V_i} \succ 0, i = 0, \ldots, n$,
$\gamma_i(a_i, y_{i-1}) \triangleq \langle a_i, R_i a_i\rangle + \langle y_{i-1}, Q_{i-1}y_{i-1}\rangle$,
$(C_{i-1}, D_i) \in \mathbb{R}^{p\times p} \times \mathbb{R}^{p\times q},\ (Q_{i-1}, R_i) \in \mathbb{S}_+^{p\times p} \times \mathbb{S}_{++}^{q\times q}$.

The G-RM may correspond to a control system or a communication channel, as shown in Figure 1. The distribution of the G-RM is $\mathbf{P}_{Y_i|A^i,Y^{i-1},S} = Q_i(dy_i|y_{i-1}, a_i), i = 1, \ldots, n$, and for $i = 0$, the distribution is $Q_0(dy_0|s, a_0)$.

By [1], the characterization of the $n-$finite transmission feedback information (FTFI) capacity is

$$\begin{aligned} C_{0,n}(\kappa) &= J_{A^n\rightarrow Y^n|s}(\pi^*, \kappa) \\ &\triangleq \sup_{\overset{\circ}{\mathcal{P}}_{[0,n]}(\kappa)} \mathbf{E}_s^\pi\Big\{\sum_{i=0}^n \log\Big(\frac{Q_i(\cdot|Y_{i-1}, A_i)}{\mathbf{P}^\pi(\cdot|Y_{i-1})}(Y_i)\Big)\Big\} \end{aligned} \quad (3)$$

$$\mathbf{P}^\pi(dy_i|y_{i-1}) = \int_{\mathbb{A}_i} Q_i(dy_i|y_{i-1}, a_i) \otimes \pi_i(da_i|y_{i-1}), \quad (4)$$

for $i = 1, \ldots, n$, and

$$\mathbf{P}^\pi(dy_0|s) = \int_{\mathbb{A}_0} Q_0(dy_0|s, a_0) \otimes \pi_0(da_0|s),$$

for $i = 0$. The set of randomized strategies $\overset{\circ}{\mathcal{P}}_{[0,n]}\ (\kappa)$ are conditionally Gaussian and Markov defined by

$$\begin{aligned} \overset{\circ}{\mathcal{P}}_{[0,n]}\ (\kappa) \triangleq \Big\{&\pi_i(da_i|y_{i-1}),\ i = 0, \ldots, n : \\ &\frac{1}{n+1}\mathbf{E}_s^\pi\Big(\sum_{i=0}^n \gamma_i(A_i, Y_{i-1})\Big) \leq \kappa\Big\} \subset \mathcal{P}_{[0,n]}(\kappa). \end{aligned} \quad (5)$$

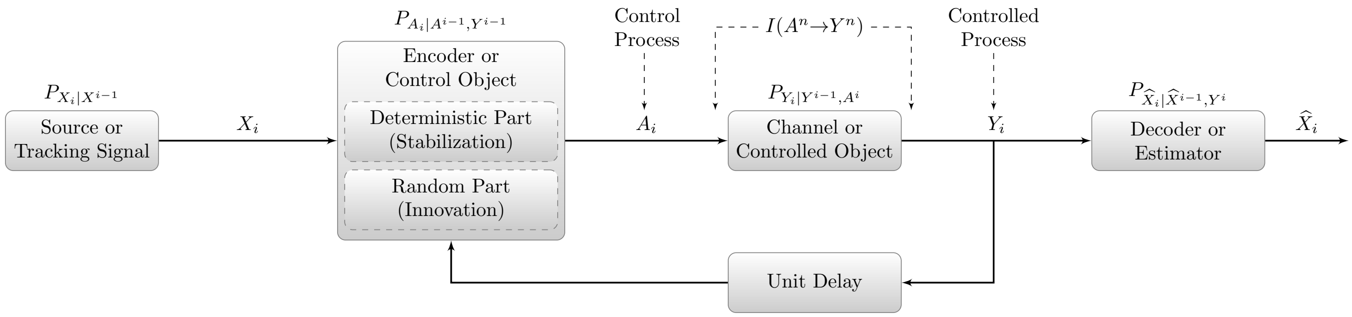

Fig. 1. *Depicts Shannon's communication block diagram and its analogy to stochastic control systems.*

The corresponding joint process $\{A_0, Y_0, \dots, A_n, Y_n\}$ and output process $\{Y_0, \dots, Y_n\}$, for fixed $S = s$, are Gaussian. Further, under the standard detectability and stabilizability conditions [1, Theorem 4.1] the feedback capacity is

$$C(\kappa) \stackrel{\triangle}{=} \lim_{n \longrightarrow \infty} \frac{1}{n+1} J_{A^n \to Y^n|s}(\pi^*, \kappa), \quad \kappa \in [\kappa_{min}, \infty).. \quad (6)$$

The *main objective* is to transform the distribution $\pi_i(da_i|y_{i-1}), i = 0, \dots, n$ that achieves the $n-$FTFI capacity and capacity $C_{0,n}(\kappa)$, into a controller-encoder and to construct a decoder, such that
(1) the controller-encoder operates at $C_{o,n}(\kappa)$,
(2) the decoder is MSE optimal,
for Gaussian, Markov process, described by the recursion

$$X_{i+1} = F_i X_i + G_i W_i, \quad X_0 = x \in \mathbb{X}_i \stackrel{\triangle}{=} \mathbb{R}^q \quad (7)$$

where $W_i \sim N(0, K_{W_i}), i = 0, \dots, n-1$ are $\mathbb{W}_i = \mathbb{R}^k-$valued zero mean Gaussian, independent of $V_i, i = 0, 1, \dots, n$, and $X_0 \sim N(0, K_{X_0})$.
It should be mentioned that controllers-encoders have contradictory goals. The controller aims at stabilization, while the encoder aims at communicating new information. A low-cost control strategy would want the state process to be kept near the origin, with as little randomness as feasible injected by the coding part, while a communication strategy requires informative deviations.

C. Related Work

The Shannon coding capacity For the memoryless additive Gaussian noise (AGN) channel

$$Y_i = A_i + V_i, \ i = 0, \dots, n, \frac{1}{n+1} \mathbf{E}\Big\{\sum_{i=0}^{n} |A_i|^2\Big\} \leq \kappa \quad (8)$$

with or without feedback, the capacity is given by [3]

$$C^{Sh}(\kappa) \stackrel{\triangle}{=} \frac{1}{2} \log\Big(1 + \frac{\kappa}{K_V}\Big). \quad (9)$$

The input process that achieves it is independent and identically distributed (IID) Gaussian $A_i^* \sim \mathbf{P}_A^*(da) \sim N(0, \kappa), i = 0, 1, \dots$.

Elias Coding Scheme of a Gaussian Message. Elias [4] introduced a coding scheme to communicate a Gaussian RV $X \sim N(0, \sigma_X^2)$ reliably over the memoryless AGN channel (8), that achieves $C^{Sh}(\kappa)$, given by

$$A_i = \sqrt{\frac{\kappa}{\mathbf{E}\Big(X - \mathbf{E}\Big\{X\Big|Y^{i-1}\Big\}\Big)^2}} \Big(X - \mathbf{E}\Big\{X\Big|Y^{i-1}\Big\}\Big). \quad (10)$$

$\widehat{X}_{i-1} \stackrel{\triangle}{=} \mathbf{E}\Big\{X\Big|Y^{i-1}\Big\}$ is linear in Y^{i-1}, $i = 0, \dots, n$, and it is computed recursively, using the Kalman-filter.

MSE Decoder of Gaussian Messages. When the Elias coding scheme is applied with a MSE decoder, the error is

$$\Sigma_n \stackrel{\triangle}{=} \mathbf{E}|X - \widehat{X}_n|^2 = \frac{\sigma_X^2}{\big(1 + \frac{\kappa}{K_V}\big)^{n+1}}, \quad n = 0, 1, \dots. \quad (11)$$

Maximum Likelihood (ML) Decoder of Digital Messages. Schalkwijk and Kailath [5] showed that, when the Elias coding scheme is applied to a set of equiprobable messages $\big\{0, 1, \dots, M_n \stackrel{\triangle}{=} \exp\{(n+1)R\} : n = 0, 1, \dots\big\}$, then the probability of ML decoding error at time n, decays doubly exponentially.

Butman Conjecture. For an AGN channel with stable and stationary noise $V^n \stackrel{\triangle}{=} \{V_0, \dots, V_n\}$ (with limited memory), *Butman [6] "conjectured"* that the Elias coding scheme of transmitting the error achieves capacity. Cover and Pombra [7] derived the characterization of feedback capacity for the AGN channel (8), when the noise V^n is nonstationary, nonergodic, with distribution $\mathbf{P}_{V^n}$. Kim [8] revisited the limited memory, stationary ergodic version of the Cover and Pombra [7] AGN channel, and applied frequency domain methods to conclude that *Butman's conjecture is true*. Variations of the Elias and Schalkwijk-Kailath schemes for network communication over memoryless AGN channels are extensive and given in [9]–[12].

III. The $n-$FTFI Capacity and Controller-Encoder-Decoder

By [1] the $n-$FTFI capacity of the G-RM is achieved by the input process

$$A_i = e_i(Y_{i-1}, Z_i) = U_i + Z_i = \Gamma_i Y_{i-1} + Z_i \quad (12)$$

(i) $U_i \stackrel{\triangle}{=} \Gamma_i Y_{i-1}, i = 0, \dots, n$ is the control process,
(ii) Z_i independent of (A^{i-1}, Y^{i-1}), $i = 0, \dots, n$,
(iii) Z^i independent of V^i, for $i = 0, \dots, n$,

(iv) $\{Z_i \sim N(0, K_{Z_i}) : i = 0, \ldots, n\}$ an independent Gaussian process.

The corresponding output process is

$$Y_i = \Big(C_{i-1} + D_i\Gamma_i\Big)Y_{i-1} + D_iZ_i + V_i, \quad Y_{-1} = s. \tag{13}$$

Furthermore, the optimal control and innovations parts of the strategy are found in [1, Section IV]. We include them below for completeness.

(a) The optimal control part of the strategy $\{U_i^* : i = 0, \ldots, n\}$, is given by

$$U_i^* = \Gamma_i^* Y_{i-1}^*, \qquad i = 0, \ldots, n, \tag{14}$$

$$\Gamma_i^* = -\Big(D_i^T P(i{+}1)D_i{+}R_i\Big)^{-1} D_i^T P(i{+}1)C_{i-1}, \tag{15}$$

where $\Gamma_n^* = 0$, $\big\{P(i) : i = 0, \ldots, n\big\}$ is a solution of the Riccati difference matrix equation

$$\begin{aligned} P(i) &= C_{i-1}^T P(i+1)C_{i-1} + Q_{i-1} \\ &\quad - C_{i-1}^T P(i+1)D_i\Big(D_i^T P(i+1)D_i + R_i\Big)^{-1} \\ &\quad \Big(C_{i-1}^T P(i+1)D_i\Big)^T, \quad P(n) = Q_{n-1} \end{aligned} \tag{16}$$

(b) The optimal innovations part of the strategy $\big\{K_{Z_i}^* \succeq 0 : i = 0, \ldots, n\big\}$ is the solution of the following problem.

$$J_{A^n \to Y^n|s}(\pi^*, \kappa) = C_{0,n}(\kappa_0^*, \ldots, \kappa_n^*) \equiv \sum_{i=0}^{n} C_i(\kappa_i^*) \tag{17}$$

$$\triangleq \sup_{K_{Z_i} \succeq 0, i=0,\ldots,n:\ \sum_{i=0}^n \kappa_i(K_{Z_i}) = \kappa(n+1)} \sum_{i=0}^{n} C_i(\kappa_i) \tag{18}$$

where

$$C_i(\kappa_i) \triangleq \frac{1}{2}\log\frac{|D_iK_{Z_i}D_i^T + K_{V_i}|}{|K_{V_i}|}, \quad i = 0, \ldots, n \tag{19}$$

$$\kappa_i \equiv \kappa_i(K_{Z_i}) \tag{20}$$

$$\triangleq \begin{cases} tr\Big(R_nK_{Z_n}\Big), \ i = n \\ tr\Big(P(i+1)\big[D_iK_{Z_i}D_i^T + K_{V_i}\big] \\ \quad + R_iK_{Z_i}\Big), \ i = 1, \ldots, n-1 \\ tr\Big(P(1)\big[D_0K_{Z_0}D_0^T + K_{V_0}\big] + R_0K_{Z_0}\Big) \\ \quad + \langle s, P(0)s\rangle, \ i = 0. \end{cases}$$

Example 3.1: Let $p = q = 1$. Then the solution of the Riccati difference matrix equation (16) does not depend on the covariance $K_{Z_i}, i = 0, \ldots, n$, and hence this simplifies the computation of the optimal $K_{Z_i}^*$ in the optimization problem (18). By the Kuhn-Tucker conditions we obtain

$$K_{Z_n}^* = \Big\{\frac{1}{2\lambda R_n} - \frac{K_{V_n}}{D_n^2}\Big\}^+, \quad \{x\}^+ \triangleq \max\{0, x\}, \tag{21}$$

$$K_{Z_i}^* = \Big\{\frac{1}{2\lambda\Big(P(i+1)D_i^2 + R_i\Big)} - \frac{K_{V_i}}{D_i^2}\Big\}^+ \tag{22}$$

for $i = n-1, n-2, \ldots, 0$, where $\lambda = \lambda_n(\kappa) \geq 0$ is chosen to satisfy the average constraint with equality given by

$$\begin{aligned} &\sum_{i=0}^{n-1}\Big\{\Big\{\frac{1}{2\lambda} - \frac{\Big(P(i+1)D_i^2 + R_i\Big)K_{V_i}}{D_i^2}\Big\}^+ + P(i{+}1)K_{V_i}\Big\} \\ &\quad + \Big\{\frac{1}{2\lambda} - \frac{R_nK_{V_n}}{D_n^2}\Big\}^+ + s^2P(0) = \kappa(n+1). \end{aligned} \tag{23}$$

Upon substituting into (19) then we obtain

$$C_{0,n}(\kappa) \triangleq J_{A^n \to Y^n|s}(\pi^*, \kappa) = \frac{1}{2}\sum_{i=0}^{n}\log\frac{|D_i^2K_{Z_i}^* + K_{V_i}|}{|K_{V_i}|} \tag{24}$$

Clearly, in general, for each i, $C_i(\kappa_i^*) > 0$ provided $\kappa_i^* \in (\kappa_{min,i}, \infty)$ and these critical values depend on whether the coefficients of the G-RM (13) lie outside or inside the unit circle, i.e., $|C_{i-1}| \geq 1$ or $|C_{i-1}| < 1$, for $i = 0, \ldots, n$ (see [1]). Expressions (21), (22) are known as water-filling in information theory. Clearly, the following hold.

$$\text{If } \kappa \triangleq \kappa_{min} \in [0, \infty) \text{ is such that} \tag{25}$$

$$\lambda_n(\kappa_{min}) > \frac{D_n^2}{2K_{V_n}R_n}, \tag{26}$$

$$\lambda_i(\kappa_{min}) > \frac{D_i^2}{2K_{V_i}\Big(P(i+1)D_i^2 + R_i\Big)}, i = 0, \ldots, n-1, \tag{27}$$

$$\text{then} \quad C_{0,n}(\kappa) = 0, \quad \text{and} \tag{28}$$

$$\kappa_{min} \triangleq \kappa_{0,n}(K_Z^*)\Big|_{K_Z* = 0} = \sum_{i=0}^{n-1}P(i+1)K_{V_i} + s^2P(0).$$

Hence, κ_{min} is the minimum power above which a non-negative information rate occurs, i.e., $C_{0,n}(\kappa) > 0, \forall \kappa \in (\kappa_{min}, \infty)$. κ_{min} is the solution of the Linear-Quadratic Gaussian stochastic optimal control problem of minimizing the average power, when $A_i = g_i(Y^{i-1}, s), i = 0, \ldots, n$, i.e., non-random; see the numerical example in Fig 2.

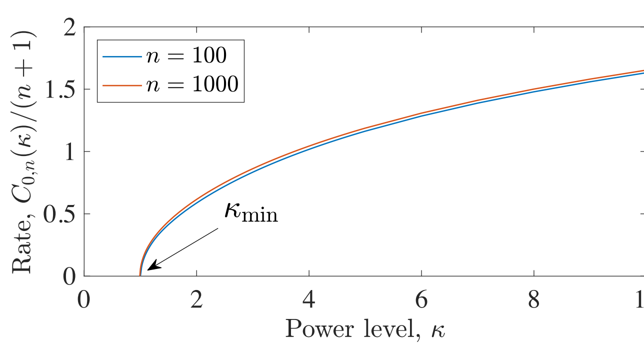

Fig. 2. *Example in which the rate of a scalar system is shown for different values of power level, κ, for $n = 100$ and $n = 1000$. Note that C_i is chosen such that $|C_i| > 1$.*

A. Optimal Controllers-Encoders

Now, we use the calculation of the $n-$FTFI capacity to show that linear controller-encoder strategies in (x_i, y^{i-1}), denoted by $\{\mu_i^L(x_i, y^{i-1}) : i = 0, \ldots, n\}$, achieve the $n-$FTFI, and that among such linear controller-encoders, then the conditional mean decoder minimizes the MSE.

Theorem 3.2: In the class of linear controller-encoders that encode the Gaussian Markov process X^n defined by (7) and operate at the $n-$FTFI capacity, the optimal controller-encoder exists, the conditional mean decoder minimizes the MSE, and these are given below.

Let $\Big\{\big(\Gamma_i^*, K_{Z_i}^*\big) : i = 0, \ldots, n\Big\}$ be the optimal strategy corresponding to (15) and (18), and joint process $\big\{(A_i^*, Y_i^*, Z_i^*) : i = 0, \ldots, n\big\}$.

Define the filter estimates[1] and conditional covariances

$$\widehat{X}_{i|i-1} \triangleq \mathbf{E}_s\Big\{X_i \Big| Y^{*,i-1}\Big\}, \quad \widehat{X}_{i|i} \triangleq \mathbf{E}_s\Big\{X_i \Big| Y^{*,i}\Big\},$$

$$\Sigma_{i|i-1} \triangleq \mathbf{E}_s\Big\{\Big(X_i - \widehat{X}_{i|i-1}\Big)\Big(X_i - \widehat{X}_{i|i-1}\Big)^T \Big| Y^{*,i-1}\Big\},$$

$$\Sigma_{i|i} \triangleq \mathbf{E}_s\Big\{\Big(X_i - \widehat{X}_{i|i}\Big)\Big(X_i - \widehat{X}_{i|i}\Big)^T \Big| Y^{*,i}\Big\}, \quad i = 0, \ldots, n.$$

(a) Controller-Encoder. The mutual information between X^n and Y^n for fixed $S = s$, denoted by $I^{\mu^{L,*}}(X^n; Y^n|s)$ of the controller-encoder strategy[2]

$$A_i^* = \mu_i^{L,*}(X_i, Y^{*,i-1}) = \Gamma_i^* Y_{i-1}^* + \Theta_i^* \Big\{X_i - \widehat{X}_{i|i-1}\Big\}, \quad (29)$$

$$\Theta_i^* = K_{Z_i}^{*,\frac{1}{2}} \Sigma_{i|i-1}^{-\frac{1}{2}}, \quad \Theta_i^* \succeq 0, \quad i = 0, \ldots, n, \quad (30)$$

$$Y_i^* = \Big(C_{i-1} + D_i \Gamma_i^*\Big) Y_{i-1}^* + D_i\, \Theta_i^* \Big\{X_i - \widehat{X}_{i|i-1}\Big\} + V_i, \quad (31)$$

operates at the $n-$FTFI capacity of the G-RM, i.e.,

$$I^{\mu^{L,*}}(X^n; Y^n|s) = J_{A^n \rightarrow Y^n|s}(\pi^*, \kappa). \quad (32)$$

Moreover, the following hold.

(b) Filter Estimates. Define the innovations process by $\big\{\nu_i^* \triangleq Y_i^* - \mathbf{E}_s\big\{Y_i^* \big| Y^{*,i-1}\big\} : i = 0, \ldots, n\big\}$. Then the optimal filter estimates satisfy the following recursions.

$$\begin{aligned} \widehat{X}_{i+1|i} &= F_i \widehat{X}_{i|i-1} + \Psi_{i|i-1}\Big\{Y_i^* - \Big(C_{i-1} + D_i \Gamma_i^*\Big) Y_{i-1}^*\Big\}, \\ &= F_i \widehat{X}_{i|i-1} + \Psi_{i|i-1} \nu_i^*, \quad \widehat{X}_{0|-1} = \text{Given}, \end{aligned} \quad (33)$$

$$\widehat{X}_{i+1|i+1} = F_i \widehat{X}_{i|i} + \overline{\Psi}_{i|i-1} \nu_i^*, \quad (34)$$

$$\begin{aligned} \Sigma_{i+|i} &= F_i \Sigma_{i|i-1} F_i^T + G_i K_{W_i} G_i^T - F_i \Sigma_{i|i-1}\Big(D_i \Theta_i^*\Big)^T \\ &\Big[D_i K_{Z_i}^* D_i^T + K_{V_i}\Big]^{-1}\Big(D_i \Theta_i^*\Big)\Sigma_{i|i-1} F_i^T, \end{aligned} \quad (35)$$

$$\Sigma_{0|-1} = \mathbf{E}_s\Big\{\Big(X_0 - \widehat{X}_{0|-1}\Big)\Big(X_0 - \widehat{X}_{0|-1}\Big)^T\Big\} \quad (36)$$

$$\Sigma_{i|i} = \Sigma_{i|i-1} - \overline{\Psi}_{i|i-1}\Big(D_i \Theta_i^*\Big)\Sigma_{i|i-1} \quad (37)$$

$$\Sigma_{i+1|i} = F_i \Sigma_{i|i} F_i^T + G_i K_{W_i} G_i^T \quad (38)$$

where the filter gains are difined by

$$\begin{aligned} \Psi_{i|i-1} &\triangleq F_i \overline{\Psi}_{i|i-1}, \\ \overline{\Psi}_{i|i-1} &\triangleq \Sigma_{i|i-1}\Big(D_i \Theta_i^*\Big)^T \Big[D_i K_{Z_i}^* D_i^T + K_{V_i}\Big]^{-1}. \end{aligned} \quad (39)$$

(c) Conditional Mean Decoder. For the controller-encoder $\mu_i^{L,*}(X_i, Y^{*,i-1}, s) = \widehat{\mu}_i^{L,*}(X_i, Y_{i-1}^*, \widehat{X}_{i|i-1}, s) \equiv \widehat{\mu}_i^{L,*}(X_i, Y_{i-1}^*, \widehat{X}_{i-1|i-1}, s)$, then the conditional mean decoder $\widehat{X}_{i|i} = \widehat{d}_i^*(Y^{*,i}, \mu^{L,*}(\cdot), s), i = 0, \ldots, n$ is optimal, in the sense of minimizing the MSE.

Proof: (a)-(c) are easily verified from [13]. ■

Next, we give an illustrative example of Theorem 3.2 to demonstrate the properties of the optimal controller-encoder, and its relation to the MSE, which is a generalization of the material discussed in Section II-C.

Example 3.3: Consider Theorem 3.2 with $p = q = 1$. By solving Riccati equation (36) we obtain

$$\begin{aligned} \Sigma_{i+1|i} &= F_i^2 \left(\frac{D_i^2 K_{Z_i}^* + K_{V_i}}{K_{V_i}}\right)^{-1} \Sigma_{i|i-1} + G_i^2 K_{W_i}, \quad \Sigma_{0|-1} \\ &= F_i^2 e^{-2C_i(\kappa_i^*)} \Sigma_{i|i-1} + G_i^2 K_{W_i}, \quad i = 0, \ldots, n \end{aligned} \quad (40)$$

$$\Sigma_{i|i} = F_{i-1}^2 e^{-2C_i(\kappa_i^*)} \Sigma_{i-1|i-1} + e^{-2C_i(\kappa_i^*)} G_{i-1}^2 K_{W_{i-1}}. \quad (41)$$

$$\Sigma_{0|0} = e^{-2C_0(\kappa_0^*)} \Sigma_{0|-1} \quad (42)$$

Clearly, by the above solutions there is a direct relation between the sequence of the MSEs at each time $\{\Sigma_{i|i-1}, \Sigma_{i|i-1}\}$, the information rates at each time instant $\{C_i(\kappa_i^*) : i = 0, 1, \ldots, n\}$, and the parameters of the information process $\{(F_i, G_i, K_{W_i}) : i = 0, \ldots, n-1\}$.

Case 1. Information Process $X_{i+1} = F_i X_i, X_0 \sim N(0, \sigma_{X_0}^2), i = 0, \ldots, n$, that is, we set $G_i = 0$. Then

$$\begin{aligned} \Sigma_{n|n} &= |F_0|^2 |F_1|^2 \ldots |F_{n-1}|^2\ e^{-2\sum_{j=0}^n C_j(\kappa_j^*)} \Sigma_{0|-1}, \\ & n = 0, 1, \ldots, \quad \Sigma_{0|0} = e^{-2C_0(\kappa_0^*)} \Sigma_{0|-1}, \end{aligned}$$

Then $\Sigma_{n|n}, n = 0, 1, \ldots$ converge monotonically to zero, i.e.,

$$\begin{aligned} &\text{If} \quad \sum_{i=0}^n C_i(\kappa_i^*) > \sum_{i \in \{0,\ldots,n-1\}: |F_i| > 1} \log|F_i|, \ \forall n = 0, \ldots \\ &\text{then} \quad \lim_{n \longrightarrow \infty} \Sigma_{n|n} = 0. \end{aligned} \quad (43)$$

Conditions (43) states that the larger the F_i's, i.e., more unstable, then the larger total power κ is needed to ensure the MSE of the decoder converges to zero.

Case 2. RV $X \sim N(0, \sigma_X^2)$. The MSE is obtained from Case 1, by setting $F_i = 1, i = 0, \ldots, n-1$, giving

$$\Sigma_{n|n} = e^{-2\sum_{i=0}^n C_i(\kappa_i)} \Sigma_{0|-1}, \quad n = 0, 1, \ldots. \quad (44)$$

This is the analog of Elias MSE decoding error (11); it is identical to (11), if the G-RM is memoryless, i.e., if $C_i = 0, i = 0, \ldots, n-1$ and $Q_i = 0, i = 0, \ldots, n-1, R_i = 1, i = 0, \ldots, n$.

Numerical example. In Fig. 4 we observe that as we increase the total power κ the estimation error $\Sigma_{n|n}$ given iteratively in (41) is reduced and eventually it converges exponentially to zero. However, below a certain value, then the constraint set is not feasible, i.e., it is empty.

[1] Without loss of generality we may assume $\sigma\{Y_{-1}^*\} = \sigma\{\Omega, \emptyset\}$, and omit the dependence of $\mathbf{E}_s$ on e.

[2] For any square matrix M with real entries $M^{\frac{1}{2}}$ is its square root.

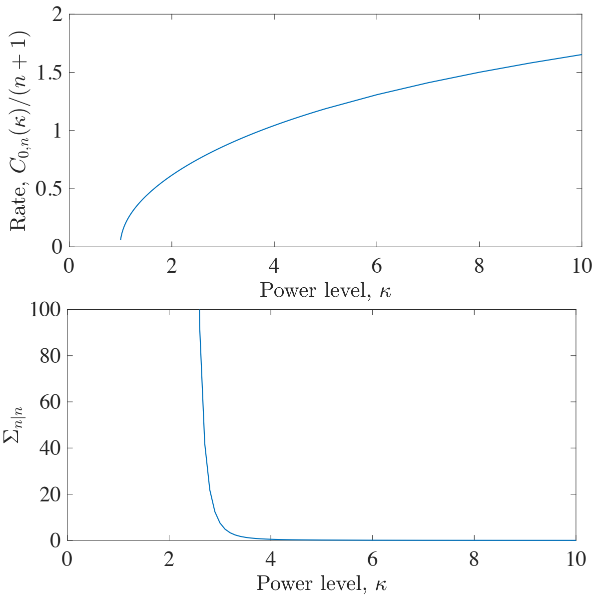

Fig. 3. *Example of the rate $C_{0,n}(\kappa)/(n+1)$ and exponential convergence of the* a posteriori *estimate $\Sigma_{n|n}$ to zero of a scalar system are shown for different values of power level, κ, for $n = 1000$. Note that C_i and F_i are chosen such that $|C_i| > 1$ and $|F_i| > 1$, respectively.*

B. Asymptotic Properties

The asymptotic properties of the controller-encoder-decoder are obtained by analyzing (6), under the following assumptions (see [14] on Linear Quadratic stochastic optimal control theory with complete information).

Assumption 3.4: (Time-invariant policies) Suppose the G-RM is time-invariant with parameters (C, D, K_V, R, Q) and

$i)$ the pair (C, D) is stabilizable;

$ii)$ the pair (G, C) is detectable, $Q \triangleq G^T G,\ G \in \mathbb{S}_+^{p\times p}$;

$iii)$ the policies are restricted to time-invariant policies $\pi_i(da_i|y_{i-1}) = \pi^\infty(da|y_{i-1}), i = 0, 1, \ldots$.

Under Assumptions 3.4, then $C(\kappa)$, defined by (6), is

$$C(\kappa) = \sup_{K_Z \in \mathbb{S}_+^{q\times q}} \frac{1}{2} \log \frac{|DK_Z D^T + K_V|}{|K_V|}, \tag{45}$$

$$tr\big(RK_Z\big) + tr\big(P\big[DK_Z D^T + K_V\big]\big) \leq \kappa \tag{46}$$

for $\kappa \in [\kappa_{min}, \infty)$, and the optimal policy $\Gamma^{\infty,*}$ is given by

$$\Gamma^* = -\Big(D^T P D + R\Big)^{-1} D^T P C, \tag{47}$$

$$P = -C^T P D\Big(D^T P D + R\Big)^{-1}\Big(C^T P D\Big)^T + C^T P C + Q, \quad spec\Big(C + D\Gamma^*\Big) \subset \mathbb{D}_o. \tag{48}$$

Here, $spec(A) \subset \mathbb{C}$ denotes the set of all eigenvalues of A, and $\mathbb{D}_o \triangleq \big\{c \in \mathbb{C} : |c| < 1\big\}$ denotes the open unit disc of the space of complex number $\mathbb{C}$. Similarly, we can also identify conditions for the asymptotic properties of optimal controllers-encoders-decoders of Section III-A.

Further, by [1, Theorem 4.1], then $C(\kappa)$ is the feedback capacity, irrespectively of whether matrix C of the G-RM is stable or unstable.

Numerical example. The asymptotic case can be easily solved using (45)-(48), for values satisfying Assumption 3.4.

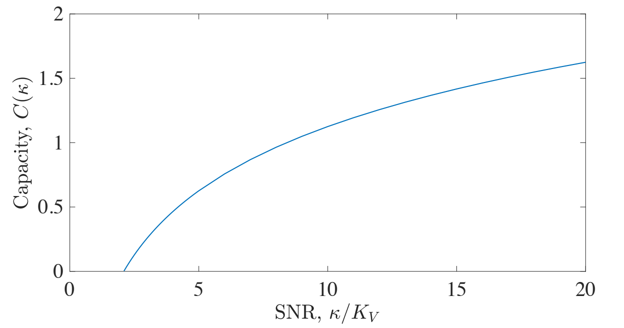

Fig. 4. *Example showing the evolution of the capacity over the SNR κ/K_V for the asymptotic case.*

IV. Conclusions

A constructive procedure is developed to synthesize {controller-encoder-decoder} strategies, that encode Gaussian Markov processes, communicate them over unstable Gaussian recursive models to the decoder. Examples illustrate the convergence of the MSE to zero, as the number of transmissions tends to infinity.

References

[1] C. K. Kourtellaris and C. D. Charalambous, "Information Structures for Feedback Capacity of Channels with Memory and Transmission Cost: Stochastic Optimal Control & Variational Equalities," *IEEE Transactions on Information Theory*, 2018, to appear.

[2] C. D. Charalambous, C. Kourtellaris, and I. Tzortzis, "Information transfer in stochastic optimal control with randomized strategies and directed information criterion," in *Proceedings of 55th IEEE Conference on Decision and Control*, Ddecember 12-16 2016, to appear.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.

[4] P. Elias, "Channel capacity without coding," MIT Research Laboratory of Electronics, Quarterly Report, October 1961, also in Lectures on Communication System Theory, E. Baghdady, Ed., New York: McGraw Hill, 1961.

[5] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback-I: no bandwidth constraints," *IEEE Transactions on Information Theory*, vol. 12, no. 2, pp. 172–182, April 1966.

[6] S. Butman, "Linear feedback rate bounds for regressive channels," *IEEE Transactions on Information Theory*, 1976.

[7] T. Cover and S. Pombra, "Gaussian feedback capacity," *IEEE Transactions on Information Theory*, vol. 35, no. 1, pp. 37–43, Jan. 1989.

[8] Y.-H. Kim, "Feedback capacity of stationary Gaussian channels," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 57–85, 2010.

[9] L. Ozarow, "The capacity of the white gaussian multiple access channel with feedback," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 623–629, July 1984.

[10] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the gaussian broadcast channel with feedback," *IEEE Transactions on Information Theory*, vol. 30, no. 4, pp. 667–671, July 1984.

[11] G. Kramer, "Feedback strategies for white gaussian interference networks," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1423–1438, June 2002.

[12] S. I. Bross and M. A. Wigger, "On the relay channel with receiver transmitter feedback," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 275–291, January 2009.

[13] P. E. Caines, *Linear Stochastic Systems*, ser. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., New York, 1988.

[14] P. R. Kumar and P. Varaiya, *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Prentice Hall, 986.