2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Signalling of Information via Coding in a Series Network of Unstable Stochastic Dynamical Control Systems

Charalambos D. Charalambous, Christos Kourtellaris and Ioannis Tzortzis

***Abstract*— We address the asymptotic problem of signalling information from one controller to another controller, in a series network, consisting of control system 1 (CS-1) and control system 2 (CS-2), as shown in Figure I.1, first analyzed in [1] for finite-horizon. Controller 2 of CS-2 has access to feedback information from its output, while controller 1 of CS-1 does not have access to feedback information from its output. Under suitable detectability and stabilizability conditions of matrix algebraic Riccati equations (AREs), it is shown that, if the rate of generating information by CS-1 is below the asymptotic control-coding (CC) capacity of CS-2, then we can synthesize, i) a controller-encoder for CS-2 that simultaneously controls the CS-2 and encodes the state of the CS-1, and operates at the CC capacity of CS-2,**

ii) a decoder for CS-2 that is optimal with respect to a mean-square error (MSE) criterion, and

iii) a controller for CS-1, which acts on the decoder output, and it is optimal with respect to the pay-off of CS-1.

Compared to [1], this paper includes bounds to MSE and Error probability of communicating digital messages.

I. Introduction

Shannon's coding capacity [2] is developed over the years with emphasis on communication system applications [3], [4]. It is demonstrated in [5] (see Theorem 4.1, Theorem 5.1) and [6], [7], that Shannon's coding capacity admits a natural generalization to decision models (DMs), such as, stochastic control systems, while in [8]–[10] is extended to unstable communication channels with memory. Further, in [11] it is demonstrated that optimal randomized strategies can be transformed into controller-encoders that simultaneously control outputs, encode Gaussian messages, and signal the messages to a decoder that reconstructs them with asymptotically arbitrary small MSE error.

The new paradigms of information signalling of control strategies constructed in [5], [11], illustrate that stochastic dynamical control systems with randomized control strategies, are also candidates of communication channels, capable of information transfer from their inputs to their outputs. The operational definition is a variant of Shannon's *coding rate*, called *control-coding (CC) rate*, with the encoder replaced by a controller-encoder [5], [11]. Consequently, a generalization of Shannon's direct coding theorem, states the following [5] (Theorem 4.1, Theorem 5.1): *"For any CC rate of R bits/second below the control-coding (CC) capacity (supremum of all rates R) of the DM, there exists a controller-encoder which controls the output process and encodes an information process, and a decoder or estimator, which operates asymptotically, with arbitrary small decoding error probability."*

However, real-time signalling of information from the input to the output of DMs, is still a challenging problem. Real-time signalling is often addressed via joint source channel coding (JSCC) design, using lossy compression of the information process. The JSCC is discussed in [12], using the nonanticipative rate distortion function (RDF). An example of signalling messages generated by a digital binary symmetric source over a binary symmetric channel with memory is given in [13].

The main objective of this paper is to analyze the per unit time limit of a series network of two unstable control systems, each assigned one controller, as shown in Fig.I.1, to signal information from one controller to another controller. The current analysis is focused on asymptotic analysis, and builds on the prior concepts and results of [1].

The Series Network of Control Systems of Figure I.1. The underlying hypotheses of the network are the following.

CS-2. The controller of CS-2 observes its output through feedback, and

CS-1. the controller of CS-1 does not observe its output.

Applications of such network include,

Scenario 1: CS-1 is controllable but its state process $\{X_i : i = 0, \ldots, n\}$ is not observable, possibly because CS-1 corresponds to the internal dynamics, and

Scenario 2: the state process $\{X_i : i = 0, \ldots, n\}$, which can be controlled or uncontrolled, is to be reproduced at the output of a decoder for the purpose of relaying it to another processor (not shown in the figure), etc.

To overcome the limitation of the controller of CS-1, the CC Capacity of CS-2 is determined and the randomized control strategy which achieves it, is found. Then the randomized control strategy of CS-2 is transformed into an controller-encoder, which simultaneously controls the output process $\{Y_i : i = 0, \ldots, n\}$ of CS-2, and encodes the state process $\{X_i : i = 0, \ldots, \}$ of CS-1, and a decoder or estimator is designed $\{\widehat{X}_i : i = 0, \ldots, \}$, with respect to a performance objective, while the decoder output $\{\widehat{X}_i : i = 0, \ldots, \}$ is made available to the controller of CS-1, to minimize the pay-off of CS-1.

For Gaussian systems we show under suitable detectability and stabilizability conditions that involve *three matrix AREs*, that the following hold.

(a) If the CC Capacity of CS-2 is above the rate at which the CS-1 generates its output $\{X_i : i = 0, \ldots, \}$, then the controlled process $\{X_i : i = 0, \ldots, \}$ can be encoded into the control strategy of CS-2 and decoded with arbitrary small

The authors are with the Department of Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus. E-mails: {chadcha,tzortzis.ioannis}@ucy.ac.cy

979-8-3503-0123-6/23/$31.00 ©2023 IEEE

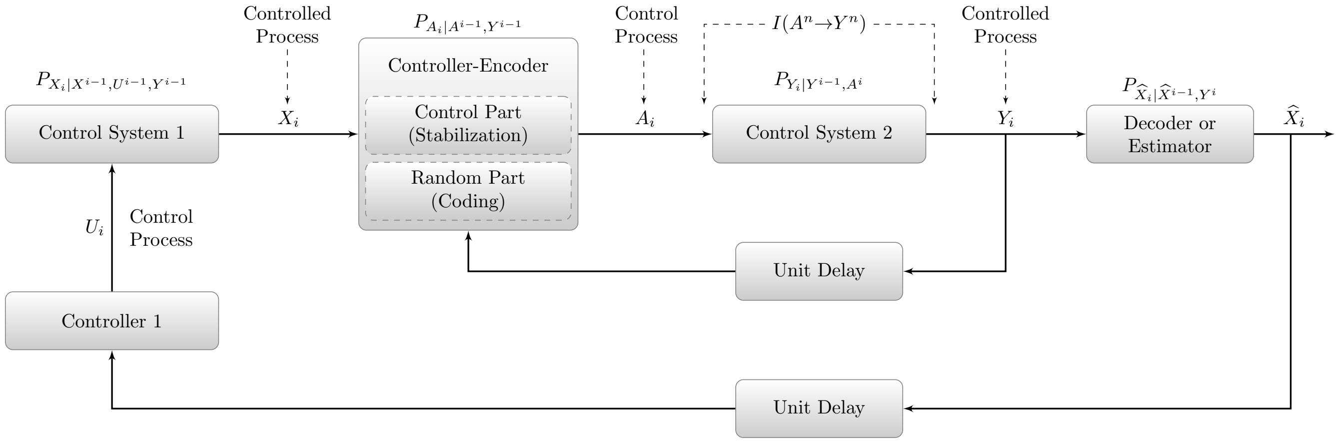

Fig. I.1. *Signalling of Information in a Series Network of Control Systems via Control-Coding.*

asymptotic error probability. For this to hold, it it necessary to allocate sufficient power to CS-2, which corresponds to the ergodic minimum cost of controlling CS-2. Any additional power is converted into information, which is signalled from the input to the output of CS-2.
(b) Under the condition of (a), the optimal controller of CS-1 is determined with respect to its ergodic control performance, and stability and optimality of the networked control system, and signalling of information is achieved, with small error probability, by proper control-coding-decoding.

Statements (a), (b) are one of the possible applications of Shannon's operational definition of capacity of noisy channels, extended to unstable dynamical systems. However, since this paper is a continuation of [1], then it is necessary to recall several results from [1], before we can address the per unit time infinite horizon problem.

II. Networked Control System with Signalling

Here, we introduce Fig. I.1 and definitions of optimality.

A. Networked Control System

Consider Fig. I.1. The control process of CS-2 is $A^n \triangleq \{A_i : i = 0, 1, \dots, n\}$ with values in $\mathbb{A}^n \triangleq \times_{i=0}^n \mathbb{A}_i$, its controlled processes is $Y^n \triangleq \{Y_i : i = 0, \dots, n\}$, with values in $\mathbb{Y}^n \triangleq \times_{i=0}^n \mathbb{Y}_i$, with initial state $S \triangleq Y_{-1}$, with values in $\mathbb{S} \triangleq \mathbb{Y}_{-1}$. Similarly, the control process of CS-1 is $U^n \triangleq \{U_i : i = 0, \dots, n\}$, with values in $\mathbb{U}^n \triangleq \times_{i=0}^n \mathbb{U}_i$, its controlled process is $X^n \triangleq \{X_i : i = 0, \dots, n\}$, with values in $\mathbb{X}^n \triangleq \times_{i=0}^n \mathbb{X}_i$.

CS-1. The conditional distribution of CS-1 is described by

$$\mathbf{P}_{X_i|X^{i-1},U^{i-1},A^{i-1},Y^{i-1},S} = \mathbf{P}_{X_i|X_{i-1},U_{i-1},Y_{i-1}} \equiv S_i(dx_i|x_{i-1}, u_{i-1}, y_{i-1}),\ i = 0, \dots, n. \quad \text{(II.1)}$$

For $i = 0$, $\mathbf{P}_{X_0|X_{-1},U_{-1},Y_{-1}} = S_0(dx_0)$.

The Control Strategies of CS-1 are measurable maps

$$\mathcal{U}_{[0,n-1]} \triangleq \Big\{ g_i : \mathbb{U}^{i-1} \times \mathbb{Y}^{i-1} \times \mathbb{S} \longmapsto \mathbb{U}_i, u_0 = g_0(s), \dots, u_i = g_i(u^{i-1}, y^{i-1}, s),\ i = 0, 1, \dots, n-2 \Big\}. \quad \text{(II.2)}$$

The Pay-off or Performance Criterion of CS-1 is

$$J_{0,n}(g^*) \triangleq \inf_{\{g_i\}_{i=0}^{n-1} \in \mathcal{U}_{[0,n-1]}} \mathbf{E}_s \Big\{ \ell_{0,n}(U^{n-1}, X^n) \Big\}, \quad \text{(II.3)}$$

$$\ell_{0,n}(u^{n-1}, x^n) \triangleq \sum_{i=0}^{n-1} \ell_i(u_i, x_i) + \varphi_n(x_n) \quad \text{(II.4)}$$

where $\ell_{0,n}(\cdot,\cdot)$ is a measurable function.
The controller of CS-1 does not have access to $\{X_i : i = 0, \dots, n\}$, but instead has access to the controlled process $\{Y_i : i = 0, \dots, n\}$ and initial state $S = s$ of CS-2.

CS-2. The conditional distribution of CS-2 is

$$\mathbf{P}_{Y_i|Y^{i-1},A^i,S,X^i,U^i} \equiv Q_i(dy_i|y_{i-1}, a_i),\ i = 0, \dots, n. \quad \text{(II.5)}$$

Definition 2.1: (Admissible controller-encoder-decoders)
(a) Controller-Encoder Strategies of CS-2. The controller-encoder strategies which control the controlled process $\{Y_i : i = 0, \dots, n\}$ and encode the controlled process $\{X_i : i = 0, \dots, n\}$ are measurable maps defined by

$$\begin{aligned} \mathcal{E}_{[0,n]}(\kappa) \triangleq \Big\{ & e_i : \mathbb{X}^i \times \mathbb{A}^{i-1} \times \mathbb{Y}^{i-1} \times \mathbb{S} \longmapsto \mathbb{A}_i, \\ & a_i = e_i(x^i, a^{i-1}, y^{i-1}, s),\ i = 0, \dots, n : \\ & \frac{1}{n+1} \mathbf{E}_s \Big(\gamma_{0,n}(A^n, Y^{n-1}) \Big) \le \kappa \Big\}, \\ & \gamma_{0,n}(A^n, Y^{n-1}) \triangleq \sum_{i=0}^n \gamma_i(a_i, y_{i-1}) \end{aligned}$$

where $\gamma_{0,n}(\cdot,\cdot) : \mathbb{A}^n \times \mathbb{Y}^{n-1} \longmapsto (-\infty, \infty]$ is a measurable function and $\kappa \in [0, \infty]$ is the total power.
(b) Decoder Strategies of CS-2. The decoder strategies which reconstruct or estimate the process $\{X_i : i = 0, \dots, n\}$ are square integrable sequences defined by

$$\mathcal{D}_{[0,n]} \triangleq \{ d_i : \mathbb{Y}^i \times \mathbb{S} \longmapsto \widehat{\mathbb{X}}_i,\ \widehat{x}_i = d_i(y^i, s), i = 0, \dots, n \}.$$

The CC-Capacity [5], [11] of CS-2 is defined using the randomized control strategies of Definition 2.2. These are transformed into an controller-encoder, which encodes $\{X_i : i = 0, \dots, n\}$ and operates at the CC Capacity of CS-2.

Definition 2.2: Admissible randomized control strategies of CS-2 belong to the constraint set

$$\mathcal{P}_{[0,n]}(\kappa) \triangleq \Big\{ P_i(da_i|a^{i-1}, y^{i-1}, s), i=0,\ldots,n : \frac{1}{n+1}\mathbf{E}_s\Big(\gamma_{0,n}(A^n, Y^{n-1})\Big) \leq \kappa \Big\}. \quad \text{(II.6)}$$

The Pay-off of CS-2 is the maximization over $\mathcal{P}_{[0,n]}(\kappa)$, of the directed information from $A^n \triangleq \{A_0,\ldots,A_n\}$ to $Y^n \triangleq \{Y_0,\ldots,Y_n\}$, for fixed initial data $S=s$ [14], [15],

$$I(A^n \rightarrow Y^n|s) \triangleq \mathbf{E}_s^P\Big\{ \sum_{i=0}^n \log\Big(\frac{dQ_i(\cdot|Y_{i-1}, A_i)}{d\mathbf{P}^P(\cdot|Y^{i-1}, S)}(Y_i)\Big)\Big\}$$

where for each i, $\mathbf{P}^P(dy_i|y^{i-1}, s) \equiv \mathbf{P}_{Y_i|Y^{i-1},S}$ is generated from $\{Q_i(\cdot|\cdot), P_i(\cdot|\cdot) : i=0,1,\ldots,n\}$.

B. Information Structures of Randomized Control Strategies

By [5], the information CC Capacity of CS-2 is a Markov Decision (MD) problem with randomized strategies:

$$\begin{aligned} C_{0,n}(\kappa) &\triangleq J_{A^n\rightarrow Y^n|s}(\pi^*, \kappa) \\ &= \sup_{\overset{\circ}{\mathcal{P}}_{[0,n]}(\kappa)} \mathbf{E}_s^\pi\Big\{ \sum_{i=0}^n \log\Big(\frac{Q_i(\cdot|Y_{i-1}, A_i)}{\Pi_i^\pi(\cdot|Y_{i-1})}(Y_i)\Big)\Big\} \\ &\equiv \sup_{\overset{\circ}{\mathcal{P}}_{[0,n]}(\kappa)} \sum_{i=0}^n I^\pi(A_i; Y_i|Y_{i-1}) \quad \text{(II.7)} \end{aligned}$$

$$\overset{\circ}{\mathcal{P}}_{[0,n]}(\kappa) \triangleq \Big\{ \pi_i(da_i|y_{i-1}),\ i=0,\ldots,n : \frac{1}{n+1}\mathbf{E}_s^\pi\Big(\sum_{i=0}^n \gamma_i(A_i, Y_{i-1})\Big) \leq \kappa \Big\} \subset \mathcal{P}_{[0,n]}(\kappa) \quad \text{(II.8)}$$

where the output process $\{Y_0,\ldots,Y_n\}$ is Markov, with corresponding transition probability distribution given by

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

By [5] the CC Capacity of CC-2 (under conditions) is

$$C(\kappa) \triangleq J_{A^\infty\rightarrow Y^\infty|s}(\pi^*, \kappa) = \lim_{n\rightarrow\infty} \frac{1}{n+1} J_{A^n\rightarrow Y^n|s}(\pi^*, \kappa).$$

By [11], the cost-rate denoted by $\kappa_{0,n}(C)$, is defined by

$$\begin{aligned} \kappa_{0,n}(C) &\triangleq \inf_{\pi_i(da_i|y_{i-1}), i=0,\ldots,n: \frac{1}{n+1}\sum_{i=0}^n I^\pi(A_i;Y_i|Y_{i-1})\geq C} \mathbf{E}_s^\pi\Big\{\gamma_{0,n}(A^n, Y^{n-1})\Big\}. \quad \text{(II.9)} \\ &\geq \inf_{\pi_i(da_i|y_{i-1}), i=0,\ldots,n} \mathbf{E}_s^\pi\Big\{\gamma_{0,n}(A^n, Y^{n-1})\Big\} \equiv \kappa_{0,n}(0). \end{aligned}$$

We should note that $\kappa_{0,n}(C) - \kappa_{0,n}(0)$ is the cost of signalling $\{X_t : t=0,\ldots,n\}$ to the output of CS-2.

Under certain conditions given in [5], $C(\kappa)$ is an upper bound on the the supremum of all achievable CC rates, and any CC rate below $C(\kappa)$ is achievable.

Algorithm for Synthesizing Strategies. The algorithm for signalling of information to CS-1 is the following.

(1) Compute the CC-Capacity $C(\kappa)$ of CS-2, and the optimal randomized control strategy $\{\pi_i^*(da_i|y_{i-1}) : i = 0,\ldots,n\} \in\overset{\circ}{\mathcal{P}}_{[0,n]}(\kappa)$, which achieves it.

(2) Transform $\{\pi_i^*(da_i|y_{i-1}) : i=0,\ldots,n\} \in\overset{\circ}{\mathcal{P}}_{[0,n]}(\kappa)$ into an controller-encoder, which controls $\{Y_i : i = 0,\ldots,n\}$, encodes $\{X_t : t=0,\ldots,n\}$, and reconstructs $\{X_t : t=0,\ldots,n\}$, by a decoder to produce $\{\widehat{X}_t : t = 0,\ldots,n\}$.

(3) Apply the estimated process $\{\widehat{X}_t : t=0,\ldots,n\}$ to the optimal controller of CS-1 that minimizes pay-off (II.3).

C. Multi-Objective Optimality of Strategies

By [1], since no controller-encoder strategy can operate at a higher information rate than $C_{0,n}(\kappa)$, and no decoder strategy can operate at an information rate higher than $C_{0,n}(\kappa)$, then we invoke the following definition of optimality.

Definition 2.3: A {controller-encoder, decoder, controller}, $(e^o(\cdot), d^o(\cdot), g^o(\cdot)) \in \mathcal{E}_{[0,n]}(\kappa)\times\mathcal{D}_{[0,n]}\times\mathcal{U}_{[0,n-1]}$ is optimal, if the following hold.

(i) For given $g(\cdot), d(\cdot)$ the strategy $e^o(\cdot, g(\cdot), d(\cdot)) \in \mathcal{E}_{[0,n]}(\kappa)$ operates at $J_{A^n\rightarrow Y^n|s}(\pi^*, \kappa)$, called information lossless.

(ii) For a given $g(\cdot)$ the decoder $d^o(\cdot) \in \mathcal{D}_{[0,n]}$ satisfies

$$\begin{aligned} &\widehat{J}_{0,n}(g, d^o(\cdot, g), e^o(\cdot, g, d^o)) \triangleq \mathbf{E}_s^{g,e^o,d^o}\Big\{\sum_{i=0}^n \rho_i(X_i, \widehat{X}_i)\Big\} \\ &\leq \widehat{J}_{0,n}(g, d(\cdot, g), e^o(\cdot, g, d)),\ \forall (d,g) \in \mathcal{D}_{[0,n]}\times\mathcal{U}_{[0,n-1]} \end{aligned}$$

where $\rho_i : \mathbb{X}_i \times \widehat{\mathbb{X}}_i \longmapsto [0,\infty), (x,\widehat{x}) \longmapsto \rho_i(x,\widehat{x}), i = 0,\ldots,n$ is the error fidelity. The mean square error (MSE) fidelity is defined by $\rho_i(x,\widehat{x}) \triangleq |x-\widehat{x}|^2, i=0,\ldots,n$.

(iii) The control strategy of the CS-1 $g^o(\cdot) \in \mathcal{U}_{[0,n-1]}$ satisfies

$$\begin{aligned} &J_{0,n}(g^o(\cdot), d^o(\cdot, g^o), e^o(\cdot, g^o, d^o)) \\ &\triangleq \mathbf{E}_s^{g^o,e^o,d^o}\Big\{\ell_{0,n}(U^{n-1}, X^n)\Big\} \\ &\leq J_{0,n}(g(\cdot), d^o(\cdot, g), e^o(\cdot, g, d^o)), \forall g(\cdot) \in \mathcal{U}_{[0,n-1]}. \quad \text{(II.10)} \end{aligned}$$

Next, we characterize the set of "information lossless controller-encoder strategies" using [1].

Theorem 2.1: (Information structures of information lossless strategies) [1]. For a given $g(\cdot), d(\cdot)$, let $e^o(\cdot, g(\cdot), d(\cdot)) \in \mathcal{E}_{[0,n]}(\kappa)$ be the optimal strategy for

$$\begin{aligned} &I_{X^n\rightarrow Y^n|s}(e^*, \kappa) \triangleq \sup_{\mathcal{E}_{[0,n]}(\kappa)} \mathbf{E}_s^e\Big\{\sum_{i=0}^n \\ &\log\Big(\frac{d\mathbf{P}(\cdot|Y_{i-1}, e_i(X^i, A^{i-1}, Y^{i-1}, S))}{\mathbf{P}^e(\cdot|Y^{i-1}, S)}(Y_i)\Big)\Big\}. \quad \text{(II.11)} \end{aligned}$$

Then the following hold.

(i) The optimal strategy in (II.11) occurs in the subset of Markov strategies in $\{X_i : i=0,\ldots,n\}$, defined by

$$\overset{\circ}{\mathcal{E}}_{[0,n]}(\kappa) \triangleq \Big\{\mu_i{:}\mathbb{X}_i\times\mathbb{Y}^{i-1}\times\mathbb{S}\longmapsto\mathbb{A}_i, a_i{=}\mu_i(x_i, y^{i-1}, s) \quad i=0,\ldots,n{:}\frac{1}{n+1}\mathbf{E}_s^\mu\Big(\gamma_{0,n}(A^n, Y^{n-1})\Big){\leq}\kappa\Big\} \subset \mathcal{E}_{[0,n]}(\kappa)$$

and moreover, $I_{X^n \to Y^n|s}(e^*, \kappa)$ reduces to the expression

$$I_{X^n \to Y^n|s}(\mu^*, \kappa) \triangleq \sup_{\mathring{\mathcal{E}}_{[0,n]}(\kappa)} \mathbf{E}_s^{\mu}\Big\{ \sum_{i=0}^{n} \log\Big(\frac{d\mathbf{P}(\cdot|Y_{i-1}, \mu_i(X_i, Y^{i-1}, S))}{\mathbf{P}^{\mu}(\cdot|Y^{i-1}, S)}(Y_i)\Big)\Big\} \tag{II.12}$$

$$\equiv \sup_{\mathring{\mathcal{E}}_{[0,n]}(\kappa)} \sum_{i=0}^{n} I^{\mu}(X_i; Y_i|Y^{i-1}, s) \tag{II.13}$$

$$\mathbf{P}^{\mu}(dy_i|y^{i-1}, s) = \int_{\mathbb{X}_i} \mathbf{P}(dy_i|y_{i-1}, \mu_i(x_i, y^{i-1}, s)) \mathbf{P}^{\mu}(dx_i|y^{i-1}, s). \tag{II.14}$$

That is, Markov coding of Markov process is optimal.
(ii) An optimal information lossless controller-encoder strategy $\{\mu_i^* : i = 0, \ldots, n\} \in \mathring{\mathcal{E}}_{[0,n]}(\kappa)$ satisfies the identity,

$$I_{X^n \to Y^n|s}(\mu^*, \kappa) = J_{A^n \to Y^n|s}(\pi^*, \kappa). \tag{II.15}$$

III. Signalling in Gaussian Networks

Now, we apply the concepts to the following CS-1, CS-2.

CS-2. A Gaussian Linear Decision Model (GL-DM-2) with quadratic cost function defined, for $i = 0, \ldots, n$:

$$Y_i = C_{i-1}\, Y_{i-1} + D_i\, A_i + V_i, \quad Y_{-1} = y_{-1} \equiv s, \tag{III.16}$$

$$\mathbf{P}_{V_i|V^{i-1}, A^i, Y_{-1}} = \mathbf{P}_{V_i}(dv_i), \quad V_i \sim N(0, K_{V_i}), \tag{III.17}$$

$$\gamma_i(a_i, y_{i-1}) \triangleq \langle a_i, R_i a_i\rangle + \langle y_{i-1}, Q_{i-1} y_{i-1}\rangle \tag{III.18}$$

where $(C_{i-1}, D_i) \in \mathbb{R}^{p\times p} \times \mathbb{R}^{p\times q}, (Q_{i-1}, R_i) \in \mathbb{S}_{+}^{p\times p} \times \mathbb{S}_{++}^{q\times q}$. The control system distribution is given by $Q_i(dy_i|y_{i-1}, a_i) \sim N(C_{i-1}y_{i-1} + D_i a_i, K_{V_i}), i = 0, \ldots, n$.

CS-1. A GL-DM-1 with quadratic cost function described recursive over the horizon $i = 0, 1, \ldots, n$:

$$X_{i+1} = H_i Y_i + F_i X_i + B_i U_i + G_i W_i, \;\; X_0 = x \in \mathbb{R}^q \tag{III.19}$$

$$\ell_i(x_i, u_i) \triangleq \langle u_i, \tilde{R}_i u_i\rangle + \langle x_i, \tilde{Q}_i x_i\rangle, \;\; \varphi_n(x) = \langle x_n, \tilde{M}_n x_n\rangle,$$

$$B_i \in \mathbb{R}^{q\times m}, \quad (\tilde{Q}_i, \tilde{R}_i) \in \mathbb{S}_{+}^{q\times q} \times \mathbb{S}_{++}^{m\times m}, \;\; \tilde{M}_n \in \mathbb{S}_{+}^{q\times q}$$

where $\{W_i \sim N\big(0, K_{W_i}\big) : i = 0, \ldots, n-1\}$ are $\mathbb{R}^k-$valued independent Gaussian processes, independent of the Gaussian RV X_0, (i.e., $\mathbf{P}_{X_0}(dx) \sim N\big(0, K_{X_0}\big)$). By (III.17), $\{W_i : i = 0, \ldots, n-1\}$ is independent of $\{V_i : i = 0, 1, \ldots, n\}$.

A. Information CC-Capacity: Hierarchical Optimality

Orthogonal Decomposition of Optimal Strategies. By [11], the optimal randomized control strategy $\{\pi_i^*(da_i|y_{i-1}) : i = 0, \ldots, n\}$ is induced by the Gaussian process $A_i = A_i^g$,

$$A_i^g = e_i^g(Y_{i-1}^g, Z_i^g) = U_i^g + Z_i^g = \Gamma_i Y_{i-1}^g + Z_i^g, \;\; U_i^g \triangleq \Gamma_i Y_{i-1}^g,$$

$$Y_i^g = \Big(C_{i-1} + D_i\Gamma_i\Big) Y_{i-1}^g + D_i Z_i^g + V_i, \quad Y_{-1}^g = y_{-1}$$

(i) Z_i^g independent of $(A^{g,i-1}, Y^{g,i-1})$, $i = 0, \ldots, n$,
(ii) $Z^{g,i}$ independent of V^i, for $i = 0, \ldots, n$,
(iii) $\{Z_i^g \sim N(0, K_{Z_i}) : i = 0, \ldots, n\}$ indep. Gaussian.

Hierarchical Decomposition and Separation Principle. The optimal strategy $\{\Gamma_i = \Gamma_i^* : i = 0, \ldots, n\}$, is given by

$$u_i^{g,*} = \overline{e}_i^{g,*}(y_{i-1}) = \Gamma_i^* y_{i-1}, \qquad i = 0, \ldots, n, \tag{III.20}$$

$$\Gamma_i^* = -\Big(D_i^T P(i{+}1) D_i {+} R_i\Big)^{-1} D_i^T P(i{+}1) C_{i-1} \tag{III.21}$$

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

$$P(i) = C_{i-1}^T P(i+1) C_{i-1} + Q_{i-1} - C_{i,i-1}^T P(i+1) D_i \Big(D_i^T P(i+1) D_i + R_i\Big)^{-1} \Big(C_{i-1}^T P(i+1) D_i\Big)^T, \quad P(n) = Q_{n-1}. \tag{III.22}$$

The optimal randomized part of the strategy $\big\{K_{Z_i} = K_{Z_i}^* : i = 0, \ldots, n\big\}$ is the solution of the water-filling problem:

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

$$\triangleq \sup_{K_{Z_i} \succeq 0, i=0,\ldots,n, \; \sum_{i=0}^{n} \kappa_i(K_{Z_i}) = \kappa(n+1)} \sum_{i=0}^{n} C_i^y(\kappa_i) \tag{III.23}$$

$$C_i^y(\kappa_i) \triangleq \frac{1}{2} \log \frac{|D_i K_{Z_i} D_i^T + K_{V_i}|}{|K_{V_i}|}, \quad i = 0, \ldots, n \tag{III.24}$$

$$\kappa_i \equiv \kappa_i(K_{Z_i}) \triangleq \begin{cases} tr\Big(R_n K_{Z_n}\Big), \; i = n \\ tr\Big(P(i+1)\big[D_i K_{Z_i} D_i^T + K_{V_i}\big] + R_i K_{Z_i}\Big), \; i = 1, \ldots, n-1 \\ tr\Big(P(1)\big[D_0 K_{Z_0} D_0^T + K_{V_0}\big] + R_0 K_{Z_0}\Big) + \langle y, P(0) y\rangle, \; i = 0. \end{cases} \tag{III.25}$$

Note that $K_{Z_i} = 0, \forall i$, implies $\kappa_{0,n}^y(0) \triangleq \sum_{i=0}^{n-1} tr\Big(P(i+1) K_{V_i}\Big) + \langle y, P(0) y\rangle$, which is the LQG optimal cost.

Example 3.1: (Scalar DM) For case $p = q = 1$, then

$$K_{Z_n}^* = \Big\{\frac{1}{2\lambda R_n} - \frac{K_{V_n}}{D_n^2}\Big\}^+, \;\; \{x\}^+ \triangleq \max\{0, x\} \tag{III.26}$$

$$K_{Z_i}^* = \Big\{\frac{1}{2\lambda\Big(P(i+1) D_i^2 + R_i\Big)} - \frac{K_{V_i}}{D_i^2}\Big\}^+ \tag{III.27}$$

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

$$\sum_{i=0}^{n-1} \Big\{\Big\{\frac{1}{2\lambda} - \frac{\Big(P(i+1) D_i^2 + R_i\Big) K_{V_i}}{D_i^2}\Big\}^+ + P(i{+}1) K_{V_i}\Big\} + \Big\{\frac{1}{2\lambda} - \frac{R_n K_{V_n}}{D_n^2}\Big\}^+ + y^2 P(0) = \kappa(n+1). \tag{III.28}$$

The information CC capacity is given by

$$C_{0,n}^y(\kappa) = \frac{1}{2} \sum_{i=0}^{n-1} \Big\{\log\Big(\frac{D_i^2}{2\lambda\Big(P(i+1) D_i^2 + R_i\Big) K_{V_i}}\Big)\Big\}^+ + \frac{1}{2}\Big\{\log\Big(\frac{D_i^2}{2\lambda R_n K_{V_n}}\Big)\Big\}^+ = \sum_{i=0}^{n} C_i^y(\kappa_i^*). \tag{III.29}$$

For $\kappa \in (\kappa^y(0), \infty)$, we obtain λ from (III.28). Clearly, in general, for each i, $C_i^y(\kappa_i^*) > 0$ provided $\kappa_i^* \in (\kappa_{min,i}, \infty)$.

B. Signalling of Information via Coding: Controller-Encoder-Decoder Strategies

In this section, we state a theorem from [1] that gives the optimal strategies {controller-encoder, decoder, controller}, according to Definition 2.3. This theorem is needed to address the infinite horizon version.

Theorem 3.1: (optimal quadruple of strategies {controller-encoder, decoder, controller}) [1]. Consider CS-1, i.e., $\{X_i : i = 0, 1, \ldots, n\}$, which is to be encoded and transmitted over the CS-2 defined by (III.16)-(III.18). Let $\Big\{\big(\Gamma_i^*, K_{Z_i}^*\big) : i = 0, \ldots, n\Big\}$ be the optimal strategy given by (III.21), (III.23) with optimal distribution $\{\pi_i^*(da_i|y_{i-1}) : i = 0, \ldots, \}$ and $\big\{(A_i^*, Y_i^*) : i = 0, \ldots, n\big\}$, which achieves $J_{A^n \rightarrow Y^n|s}(\pi^*, \kappa)$.
Define the filter estimates and conditional covariances by

$$\widehat{X}_{i|i-1} \triangleq \mathbf{E}_s\Big\{X_i\Big|Y^{*,i-1}\Big\}, \quad \widehat{X}_{i|i} \triangleq \mathbf{E}_s\Big\{X_i\Big|Y^{*,i}\Big\},$$
$$\Sigma_{i|i-1} \triangleq \mathbf{E}_s\Big\{\Big(X_i - \widehat{X}_{i|i-1}\Big)\Big(X_i - \widehat{X}_{i|i-1}\Big)^T\Big|Y^{*,i-1}\Big\},$$
$$\Sigma_{i|i} \triangleq \mathbf{E}_s\Big\{\Big(X_i - \widehat{X}_{i|i}\Big)\Big(X_i - \widehat{X}_{i|i}\Big)^T\Big|Y^{*,i}\Big\}, \ i = 0, \ldots, n.$$

Then the encoder strategy and corresponding controlled process, which operates at $J_{A^n \rightarrow Y^n|s}(\pi^{g,*}, \kappa)$, are given by

$$A_i^* = \mu_i^*(X_i, Y^{*,i-1}) = \Gamma_i^* Y_{i-1}^* + \Theta_i^* \Big\{X_i - \widehat{X}_{i|i-1}\Big\}, \quad \text{(III.30)}$$
$$\Theta_i^* = K_{Z_i}^{*,\frac{1}{2}} \Sigma_{i|i-1}^{-\frac{1}{2}}, \quad \Theta_i^* \succeq 0, \quad \text{(III.31)}$$
$$Y_i^* = \Big(C_{i-1} + D_i\Gamma_i^*\Big)Y_{i-1}^* + D_i\ \Theta_i^*\Big\{X_i - \widehat{X}_{i|i-1}\Big\} + V_i \quad \text{(III.32)}$$

for $i = 0, \ldots, n$. Moreover, the following hold.
(a) Filter Estimates. The innovations process defined by $\Big\{\nu_i^* \triangleq Y_i^* - \mathbf{E}\big\{Y_i^*\Big|Y^{*,i-1}\big\} : i = 0, \ldots, n\Big\}$ satisfies

$$\nu_i^* = Y_i^* - \Big(C_{i-1} + D_i\Gamma_i^*\Big)Y_{i-1}^* = D_i\ \Theta_i^*\Big\{X_i - \widehat{X}_{i|i-1}\Big\} + V_i,$$
$$\mathbf{E}_s\Big\{\nu_i^*\Big|Y^{*,i-1}\Big\} = \mathbf{E}_s\Big\{\nu_i^*\Big\} = 0,$$
$$\mathbf{E}_s\Big\{\nu_i^*(\nu_i^*)^T\Big|Y^{*,i-1}\Big\} = D_i K_{Z_i}^* D_i^T + K_{V_i} = \mathbf{E}_s\Big\{\nu_i^*(\nu_i^*)^T\Big\}$$

and the sequence of RVs, $\big\{\nu_i^* : i = 0, \ldots, n\big\}$ is uncorrelated. The optimal filter estimates satisfy the recursions:

$$\widehat{X}_{i+1|i} = H_i Y_i^* + F_i \widehat{X}_{i|i-1} + B_i g_i(Y^{*,i-1}) + \Psi_{i|i-1}\nu_i^*, \quad \widehat{X}_{0|-1} = \text{Given}, \quad \text{(III.33)}$$
$$\Sigma_{i+|i} = F_i\Sigma_{i|i-1}F_i^T + G_i K_{W_i} G_i^T - F_i\Sigma_{i|i-1}\Big(D_i\Theta_i^*\Big)^T . \Big[D_i K_{Z_i}^* D_i^T + K_{V_i}\Big]^{-1}\Big(D_i\Theta_i^*\Big)\Sigma_{i|i-1}F_i^T, \quad \text{(III.34)}$$
$$\Sigma_{0|-1} = \mathbf{E}_s\Big\{\Big(X_0 - \widehat{X}_{0|-1}\Big)\Big(X_0 - \widehat{X}_{0|-1}\Big)^T\Big\} \quad \text{(III.35)}$$
$$\Sigma_{i|i} = \Sigma_{i|i-1} - \overline{\Psi}_{i|i-1}\Big(D_i\Theta_i^*\Big)\Sigma_{i|i-1} \quad \text{(III.36)}$$

where the filter gains are defined by

$$\Psi_{i|i-1} \triangleq F_i\overline{\Psi}_{i|i-1},$$
$$\overline{\Psi}_{i|i-1} \triangleq \Sigma_{i|i-1}\Big(D_i\Theta_i^*\Big)^T\Big[D_i K_{Z_i}^* D_i^T + K_{V_i}\Big]^{-1} \quad \text{(III.37)}$$

and the controlled process $\big\{Y_i^* : i = 0, \ldots, n\big\}$ is given by

$$Y_i^* = \Big(C_{i,i-1} + D_i\Gamma_i^*\Big)Y_{i-1}^* + \nu_i^*, \quad i = 0, 1, \ldots. \quad \text{(III.38)}$$

(b) Information Lossless Controller-Encoder Operating at $J_{A^n \rightarrow Y^n|s}(\pi^*, \kappa)$. The controller-encoder $\{\mu_i^*(\cdot, \cdot) : i = 0, \ldots, \}$ is information lossless, that is,

$$I_{X^n \rightarrow Y^n|s}(\mu^*, \kappa) = \sum_{i=0}^{n}\Big\{H(\nu_i^*) - H(V_i)\Big\} = J_{A^n \rightarrow Y^n|s}(\pi^*, \kappa). \quad \text{(III.39)}$$

(c) The optimal decoder is $\widehat{X}_{i|i}, i = 0, \ldots, n$.
(d) The optimization problem of CS-1 is given by

$$J_{0,n}(g^*) = \inf_{\big\{g_i(\cdot):i=0,\ldots,n-1\big\}\in\mathcal{U}_{[0,n-1]}} \mathbf{E}_s\Big\{\sum_{i=0}^{n-1}\Big(\langle U_i, \tilde{R}_i U_i\rangle + \langle\widehat{X}_{i|i-1}, \tilde{Q}_i\widehat{X}_{i|i-1}\rangle\Big) + \langle\widehat{X}_{n|n-1}, \tilde{M}_n\widehat{X}_{n|n-1}\rangle\Big\} + \sum_{i=0}^{n-1} Tr\Big(\tilde{Q}_i\Sigma_{i|i-1}\Big) + Tr\Big(\tilde{M}\Sigma_{n|n-1}\Big) \quad \text{(III.40)}$$

subject to the constraint

$$\widehat{X}_{i+1|i} = H_i Y_i^* + F_i\widehat{X}_{i|i-1} + B_i g_i(Y^{*,i-1}) + \Psi_{i|i-1}\nu_i^*, \quad \widehat{X}_{0|-1} = \text{Given}, \quad \text{(III.41)}$$
$$Y_i^* = \Big(C_{i,i-1} + D_i\Gamma_i^*\Big)Y_{i-1}^* + \nu_i^*, \quad i = 0, 1, \ldots. \quad \text{(III.42)}$$

Moreover, the optimal strategy of CS-1 is linear in $\overline{X}_i \triangleq (Y_{i-1}^*, \widehat{X}_{i|i-1}), i = 0, \ldots, n$, i.e., $U_i^* = g_i^*(\overline{X}_i) = K_i\overline{X}_i, i = 0, \ldots, n$, where $K_i, i = 0, \ldots, n$ is determined from the solution of a control DRE.
Special Case. If $H_i = 0, i = 0, \ldots, n-1$ then the optimal strategy $g^*(\cdot) \in \mathcal{U}_{[0,n-1]}$ of CS-1 is given by

$$U_i^* = g_i^*(\widehat{X}_{i|i-1}) = K_i\widehat{X}_{i|i-1}, \quad i = 0, \ldots, n-1, \quad \text{(III.43)}$$
$$K_i \triangleq -\Big[\tilde{R}_i + B_i^T S(i+1)B_i\Big]^{-1} B_i^T S(i+1)F_i\widehat{X}_{i|i-1},$$

where $S(\cdot)$ satisfies the matrix DRE

$$S(i) = \tilde{Q}_i + F_i^T S(i+1)F_i - F_i^T S(i+1)B_i\Big[\tilde{R}_i + B_i^T S(t+1)B_i\Big]^{-1} B_i^T S(i+1)F_i, \ S(n) = \tilde{M}_n \quad \text{(III.44)}$$

and the optimal pay-off is given by

$$J_{0,n}(g^*, \mu^*, d^*, y) = \langle\widehat{X}_{0|-1}, S(0)\widehat{X}_{0|-1}\rangle + \sum_{i=0}^{n-1} Tr\Big(\tilde{Q}_i\Sigma_{i|i-1}\Big) + Tr\Big(\tilde{M}_n\Sigma_{n|n-1}\Big) + \sum_{i=0}^{n-1} Tr\Big(S(i+1)\overline{D}_i\overline{D}_i^T\Big), \quad \text{(III.45)}$$
$$\overline{D}_i \triangleq \Psi_{i|i-1}\Big(D_i K_{Z_i}^* D_i^T + K_{V_i}\Big)\Psi_{i|i-1}^T. \quad \text{(III.46)}$$

Example 3.2: By Theorem 3.1, with $p = q = 1$, then

$$\Sigma_{i|i} = F_{i-1}^2 e^{-2C_i^y(\kappa_i^*)} \Sigma_{i-1|i-1} + e^{-2C_i^y(\kappa_i^*)} G_{i-1}^2 K_{W_{i-1}}, \ \Sigma_{0|0} = e^{-2C_0^y(\kappa_0^*)} \Sigma_{0|-1}. \tag{III.47}$$

Special Case. Suppose $X_{i+1} = F_i X_i + B_i U_i, X_0 \sim N(0, \sigma_{X_0}^2), i = 0, \ldots, n$, that is, $G_i = 0$. Then

$$\Sigma_{n|n} = |F_0 F_1 \ldots F_{n-1}|^2 \, e^{-2\sum_{j=0}^{n} C_j^y(\kappa_j^*)} \Sigma_{0|-1}, \ n = 0, 1, \ldots$$

Moreover, the MSEs $\Sigma_{n|n}, n = 0, 1, \ldots$ satisfy:

$$\text{If } \sum_{i=0}^{n} C_i^y(\kappa_i^*) > \sum_{i \in \{0,\ldots,n-1\}: |F_i| > 1} \log |F_i|, \ \forall n = 0, \ldots \quad \text{then } \lim_{n \longrightarrow \infty} \Sigma_{n|n} = 0. \tag{III.48}$$

Remark 3.1: Condition (III.48) is fundamentally different from past literature [16]–[19], because it holds for time-varying systems and finite-time n.

C. Control-Coding Capacity of CS-2

We give conditions for $C(\kappa)$ to be the CC capacity of CS-2. Define the open unit disc of the space of complex number $\mathbb{C}$ by $\mathbb{D}_o \triangleq \{c \in \mathbb{C} : |c| < 1\}$, and the spectrum of a matrix $A \in \mathbb{R}^{q \times q}$ (the set of all its eigenvalues), by $spec(A) \subset \mathbb{C}$.

Theorem 3.2: (CC Capacity of CS-2). Consider the time-invariant (TI) version of CS-2, i.e., (III.16)-(III.18) with $(C_{i-1}, D_i, Q_{i-1}, R_i, K_{V_i}) = (C, D, Q, R, K_V)$. Assume
i) The pair (C, D) is stabilizable,
ii) the pair (G, C) is detectable, $Q = G^T G$, $G \in \mathbb{S}_+^{p \times p}$.
iii) $\pi_i(da_i|y_{i-1}) = \pi^\infty(da_i|y_{i-1}), \forall i$, i.e., TI.
Then $C(\kappa) \triangleq \lim_{n \longrightarrow \infty} \frac{1}{n+1} J_{A^n \rightarrow Y^n|Y_{-1}}(\pi^{\infty,*}, \kappa), \kappa \in (\kappa_{min}, \infty)$ exists and is finite. Moreover, the following hold.
(a) The optimal strategy is $\pi^{\infty,*}(\cdot|\cdot) \sim N(\overline{e}^{\infty,*}(y), K_Z^*)$, and the corresponding unique invariant distribution of $\{Y_i : i = 0, 1, \ldots\}$ is $\mathbf{P}^{\pi^{\infty,*}}(\cdot) \sim N(0, K_Y)$, where

$$\overline{e}^{\infty,*}(y) = \Gamma^* y, \quad \Gamma^* = -(D^T P D + R)^{-1} D^T P C, \tag{III.49}$$
$$P = C^T P C + Q - C^T P D (D^T P D + R)^{-1} D^T P C, \tag{III.50}$$
$$K_Y = (C + D\Gamma^*) K_Y (C + D\Gamma^*)^T + D K_Z^* D^T + K_V, \tag{III.51}$$
$$spec(C + D\Gamma^*) \subset \mathbb{D}_o. \tag{III.52}$$
$$A_i \triangleq \overline{e}^{\infty,*}(Y_{i-1}) + Z_i \equiv \overline{A}_i^* + Z_i, \ \overline{A}_i^* = \Gamma^* \, Y_{i-1}, \tag{III.53}$$
$$Y_i = C Y_{i-1} + D\overline{A}_i^* + D Z_i + V_i, \quad i = 0, 1, \ldots, \tag{III.54}$$

(b) $C(\kappa)$ is given by

$$C(\kappa) = \sup_{K_Z \in \mathbb{S}_+^{q \times q}} \Big\{ \frac{1}{2} \log \frac{|D K_Z D^T + K_V|}{|K_V|} + \lambda \kappa - \lambda \, tr\Big(R K_Z\Big) - \lambda \, tr\Big(P\Big[D K_Z D^T + K_V\Big]\Big) \Big\} \tag{III.55}$$

where $K_Z = K_Z^*$ is the optimal value and $\lambda \equiv \lambda(\kappa) \geq 0$ is the Lagrange multiplier found from the constraint

$$tr\Big(R K_Z\Big) + tr\Big(P\Big[D K_Z D^T + K_V\Big]\Big) \leq \kappa. \tag{III.56}$$

(c) $C(\kappa)$ is the CC capacity of the TI-GL-DM.

Remark 3.2: (Comments on Theorem 3.2)
Theorem 3.2, (c) is derived using the ergodicity of the directed information density and the optimal control cost. It implies that we can also encode digital messages into the randomized strategies of the CS-2, which may experience long coding delays. Applications include coding of actuator failures of CS-2. The technical problem of long coding delays is removed by joint source channel coding of digital sources over digital channels, as illustrated in [13].

D. Infinite Horizon Signalling of Information via Coding: Controller-Encoder-Decoder Strategies

Now, we are ready to state all assumptions that allow us to extend all statements of Theorem 3.1, to the per unit time limit, and thus ensuring asymptotic stationarity.

Theorem 3.3: (Infinite horizon optimal quadruple of strategies {controller-encoder, decoder, controller})
Consider the CS-1 and CS-2 of Theorem 3.1 with the following conditions.
(1) CS-2 is TI and satisfies the conditions of Thm 3.2, i)-iii).
(2) CS-1 is the TI version of (III.19) with $H_i = H = 0, i = 0, \ldots, n-1$ (for simplicity).
(3) i) The pair (D, F) is detectable, ii) the pair (F, W) is stabilizable, $W = G K_W G^T$, $G \in S_+^{q \times q}$, iii) the pair (F, B) is stabilizable, iv) the pair $(\tilde{G}, C)$ is detectable, where $\tilde{Q} = \tilde{G}^T \tilde{G}$, $\tilde{G} \in \mathbb{S}_+^{q \times q}$.
(4) The control strategies of CS-1 are TI, i.e., $\{g_i(\cdot) = g^\infty(\cdot) : i = 0, \ldots, \}$.
Then the following hold.
(a) The limits of Riccati difference equations of Theorem 3.2 $\Sigma \triangleq \lim_{n \longrightarrow \infty} \Sigma_{n|n-1}, \widehat{\Sigma} \triangleq \lim_{n \longrightarrow \infty} \Sigma_{n|n}, S \triangleq \lim_{n \longrightarrow \infty} S(n)$ exist, and satisfy algebraic Riccati equations.
(b) The controller-encoder strategy defined below is information lossless with respect to the CC capacity $C(\kappa)$ of Theorem 3.2.

$$A_i^* = \mu^{\infty,*}(X_i, Y^{i-1,*}) = \Gamma^* Y_{i-1}^* + \Theta^* \Big\{ X_i - \widehat{X}_{i|i-1} \Big\},$$
$$\Theta^* = K_Z^{*,\frac{1}{2}} \, \Sigma^{-\frac{1}{2}}, \quad \Theta^* \succeq 0,$$
$$Y_i^* = (C + D\Gamma^*) Y_{i-1}^* + B U_i^* + D \, \Theta^* \Big\{ X_i - \widehat{X}_{i|i-1} \Big\} + V_i.$$

where $Y_{-1}^* \sim N(0, K_Y)$. Specifically, Theorem 3.1, (a)-(c) holds with appropriate changes, such as, the following.
(a) The controller-encoder $\mu^{\infty,*}(\cdot, \cdot)$ satisfies

$$\lim_{n \to \infty} \frac{1}{n+1} I_{X^n \rightarrow Y^n|Y_{-1}}(\mu^{\infty,*}, \kappa) = J_{A^\infty \rightarrow Y^\infty|Y_{-1}}(\pi^{\infty,*} \kappa).$$

(b) $\lim_{n \longrightarrow \infty} \frac{1}{n+1} J_{0,n}(g^{\infty,*}, \mu^{\infty,*}, d^{\infty,*})$ exists, it is finite, and the optimal strategy $g^{\infty,*}(\cdot)$ is given by

$$U_i^{\infty,*} = g^{\infty,*}(\widehat{X}_{i|i-1}) = K^\infty \widehat{X}_{i|i-1}, \quad i = 0, \ldots, \tag{III.57}$$
$$K^\infty \triangleq -\Big[\tilde{R} + B^T S B\Big]^{-1} B^T S F \widehat{X}_{i|i-1}, \tag{III.58}$$
$$spec(F + B K^\infty) \subset \mathbb{D}_o. \tag{III.59}$$

where $S(\cdot)$ satisfies the matrix Riccati algebraic equation

$$S = \tilde{Q} + F^T SF - F^T SB\Big[\tilde{R} + B^T SB\Big]^{-1} B^T SF.$$

Proof: The derivation follows from Theorem 3.1 and the assumptions of detectability and stabilizability. ■

Theorem 3.3 illustrates the optimal signalling of information to stabilize CS-1, by encoding its unobserved state $\{X_i : i = 0, \ldots, \}$ using the randomized strategies of the CS-2, while ensuring optimality of control/communication objectives.

E. Digital Signalling of Information

Now, we discuss an application of CC capacity of CS-2, for digital messages, using generalizations of Schalkwijk-Kailath [20] coding scheme.

Problem 3.1: The objective is to reconstruct the state process X^n that satisfies the recursion $X_{i+1} = FX_i, X_0 \sim N(0, \sigma^2_{X_0}), i = 0, \ldots,$ at the output of a scalar CS-2, using digital coding and decoding.

The problem of signalling X^n via CS-2 is equivalent to the problem of coding-decoding of the initial value X_0, while operating at the CC capacity of CS-2.

Maximum Likelihood Error Probability. Consider a quantized representation of the initial state X_0 into equiprobable messages $X^{(n)} = x^{(n)} \in \mathcal{M}^{(n)} \triangleq \{0, 1, \ldots, M^{(n)}\}, n = 0, 1, \ldots$, where X_0 is an arbitrary RV, not necessarily Gaussian, with values in $\mathbb{R}$. By Example 3.2, and repeating [21], we can show that the probability of Maximum Likelihood (ML) decoding error at time n decreases doubly exponentially in $(n + 1)$, according to

$$\mathbf{P}^{ML}_{n,error} \leq 2Q\Big(\sqrt{\frac{3}{e}} \exp\Big\{(n+1)\Big(\frac{1}{n+1}C_{0,n}(\kappa) - R\Big)\Big\}\Big)$$

$$=2Q\Big(\sqrt{\frac{3}{e}} \exp\Big\{(n+1)\Big(C(\kappa) - R\Big)\Big\}\Big) \text{ for large } n$$

where $C(\kappa)$ is the CC capacity of the CS-2, and $M^{(n)} \triangleq \exp\{(n + 1)R\}$ is the rate. However, by Example 3.2,

$$\text{if } |F| > 1, \text{ then } C(\kappa) > \log|F|, \ \kappa \in (\kappa_{min}, \infty). \quad \text{(III.60)}$$

Hence, we have demonstrated that signalling digital messages through CS-2 is feasible, and this can be extended to the series network.

IV. CONCLUSIONS

An asymptotic hierarchical constructive procedure is developed to synthesize *optimal controllers-encoders-decoders*, to signal information from the controller of one control system to the controller of another control system, generalizing earlier work on CC Capacity of stochastic systems.

REFERENCES

[1] C. K. Charalambous, C. D. and I. Tzortzis, "Signalling of information in networked stochastic dynamical control systems," in *Proceedings of 56th IEEE Conference on Decision and Control*, Melbourne, Australia, December 12-15 2017, pp. 3138–3143.

[2] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech.*, vol. 27, pp. 379–423, July 1948.

[3] R. T. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, Inc., New York, 1968.

[4] S. Ihara, *Information theory for Continuous Systems*. World Scientific, 1993.

[5] C. Kourtellaris and C. D. Charalambous, "Information structures of capacity achieving distributions for feedback channels with memory and transmission cost: Stochastic optimal control & variational equalities," *IEEE Transactions on Information Theory*, vol. 64, no. 7, pp. 4962–4992, 2018.

[6] C. D. Charalambous, C. Kourtellaris, and S. Loyka, "Capacity achieving distributions & information lossless randomized strategies for feedback channels with memory: The LQG theory of directed information," *IEEE Transactions on Information Theory*, vol. 64, no. 9, pp. 6384–6418, 2018.

[7] C. D. Charalambous, C. Kourtellaris, I. Tzortzis, and Loyka, "The capacity of unstable dynamical systems-interaction of control and information transmission," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, 25-30, June 2017, pp. 2663–2667.

[8] C. Kourtellaris, C. D. Charalambous, and S. Loyka, "New formulas for ergodic feedback capacity of AGN channels driven by stable and unstable autoregressive noise," in *IEEE International Symposium on Information Theory (ISIT)*, July 21-26 2020, pp. 2091–2096.

[9] C. D. Charalambous, C. Kourtellaris, and S. Louka, "New formulas of feedback capacity for AGN a time-domain sufficient statistic approach," 13 October 2020. [Online]. Available: https://arxiv.org/abs/2010.06226

[10] S. Louka, C. Kourtellaris, and C. D. Charalambous, "Qualitative analysis of feedback capacity of agn channels driven by stable and unstable autoregressive moving average," in *in Proceedings of the 2021 IEEE Information Theory Workshop (ITW)*. IEEE, October 17-21 2021.

[11] Charalambous, C. Kourtellaris, and I. Tzortzis, "Information transfer of control strategies: Dualities of stochastic optimal control theory and feedback capacity of information theory," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 5010–5025, March 2017.

[12] C. Kourtellaris, C. D. Charalambous, and P. Stavrou, "Nonanticipative duality of sources and channels with memory and feedback," in *Coordination control of distributed systems*, ser. LNCIS, J. H. van Schuppen and T. Villa, Eds. Springer, Berlin, 2014, vol. 456.

[13] C. Kourtellaris, C. Charalambous, and J. Boutros, "Nonanticipative transmission of sources and channels with memory," in *IEEE International Symposium on Information Theory (ISIT)*, 2015.

[14] H. Marko, "The bidirectional communication theory–A generalization of information theory," *IEEE Transactions on Communications*, vol. 21, no. 12, pp. 1345–1351, Dec. 1973.

[15] J. L. Massey, "Causality, feedback and directed information," in *International Symposium on Information Theory and its Applications (ISITA '90)*, Nov. 27-30 1990, pp. 303–305.

[16] W. Wong and R. Brockett, "Systems with finite communication bandwidth constraints I: Stabilization with limited information feedback," *IEEE Transactions on Automatic Control*, vol. 44, pp. 1049–1053, May 1999.

[17] G. Nair, R. Evans, I. Mareels, and W. Moran, "Topological feedback entropy and nonlinear stabilization," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1585–1597, Sept. 2004.

[18] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056–1068, July 2004.

[19] N. Elia, "When Bode meets Shannon: control-oriented feedback communication schemes," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1477–1488, Sept. 2004.

[20] 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.

[21] R. G. Gallager and B. Nakiboglu, "Variations on a theme by schalkwijk and kailath," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 6–17, May 2010.