A Riccati-Lyapunov Approach to Nonfeedback Capacity of MIMO Gaussian Channels Driven by Stable and Unstable Noise

Charalambos D. Charalambous[1], Stelios Louka[1] and Sergey Loyka[2]
University of Cyprus[1] *and University of Ottawa*[2]
chadcha@ucy.ac.cy,louka.stelios@ucy.ac.cy,sergey.loyka@ieee.org

2022 IEEE Information Theory Workshop (ITW) | 978-1-6654-8341-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/ITW54588.2022.9965821

***Abstract*—We show that the nonfeedback capacity of multiple-input multiple-output (MIMO) additive Gaussian noise (AGN) channels, when the noise is nonstationary and unstable, is characterized by an asymptotic optimization problem–the per unit time limit of the characterization of a finite block or transmission without feedback information (FTwFI) capacity, that involves two generalized matrix difference Riccati equations (DREs) of filtering theory, and a matrix difference Lyapunov equation of stability theory, of Gaussian systems. Further, we identify conditions and prove, that the characterization of nonfeedback capacity is the uniform asymptotic per unit time limit, over all initial distributions. The asymptotic characterization of capacity involves two generalized matrix algebraic Riccati equations (AREs) and a matrix algebraic Lyapunov equation. We also present an example to illustrate that our characterization of capacity produces a known closed-form expression of the water-filling solution of capacity (for power levels above a minimum power).**

I. Introduction, Problem, and Main Results

The nonfeedback capacity of time-invariant Gaussian channels, with stable impulse response, when the noise is stationary or asymptotically stationary, is often characterized in frequency-domain, by the so-called water-filling solution. It can be found in several books [1]–[4] and research papers, such as Tsybakov [5]. The analysis of channel capacity for asymptotically equivalent matrices for multiple-input multiple-output (MIMO) Gaussian channels, is found in [6]–[8], while for Gaussian channels with intersymbol interference in [9]. A characterization of nonfeedback capacity for single-input single-output (SISO) AGN channels with nonstationary noise is given in Cover and Pombra [10], while bounds on nonfeedback capacity and comparisons to feedback capacity, are given in [10]–[13].

The Channel Model. In this paper, we analyze the nonfeedback capacity of MIMO AGN channels, driven by general unstable, nonstationary, nonergodic noise,

$$Y_t = H_t X_t + V_t, \quad t=1,\ldots,n, \quad \frac{1}{n}\mathbf{E}\Big\{\sum_{t=1}^{n} ||X_t||^2_{\mathbb{R}^{n_x}}\Big\} \leq \kappa \qquad \text{(I.1)}$$

where $\kappa \in [0,\infty)$, $X_t : \Omega \rightarrow \mathbb{X} \stackrel{\triangle}{=} \mathbb{R}^{n_x}$, $Y_t : \Omega \rightarrow \mathbb{Y} \stackrel{\triangle}{=} \mathbb{R}^{n_y}$, and $V_t : \Omega \rightarrow \mathbb{V} \stackrel{\triangle}{=} \mathbb{R}^{n_y}$, are the channel input, channel output and noise random variables (RVs), respectively, (n_x, n_y) are finite positive integers, $H_t \in \mathbb{R}^{n_y \times n_x}$ is nonrandom and the distribution of the sequence $V^n = \{V_1,\ldots,V_n\}$, i.e., $\mathbf{P}_{V^n} \stackrel{\triangle}{=} \mathbb{P}\{V_1 \leq v_1,\ldots,V_n \leq v_n\}$, is jointly Gaussian, and $\mathbf{P}_{V_1}$ is the distribution of the RV V_1.

Operational Nonfeedback Code. The code consists of (a) a set of uniformly distributed messages $M : \Omega \rightarrow \mathscr{M}^{(n)} \stackrel{\triangle}{=} \{1,\ldots,M^{(n)}\}$, known to the encoder and decoder, (b) a set of encoder strategies mapping messages $M=m$ and past channel inputs into current inputs, defined by[1]

$$\mathscr{E}_n(\kappa) \triangleq \Big\{g_i : \mathscr{M}^{(n)} \times \mathbb{X}^{i-1} \rightarrow \mathbb{X}_i,\, x_1 = g_1(m),\, x_2 = g_2(m,x_1), \ldots, x_n = g_n(m,x^{n-1}) \;\Big|\; \frac{1}{n}\mathbf{E}^g\Big\{\sum_{t=1}^{n}||X_t||^2_{\mathbb{R}^{n_x}}\Big\} \leq \kappa\Big\} \qquad \text{(I.2)}$$

where $g_i(\cdot)$ are measurable maps and, (c) a decoder $d_n(\cdot) : \mathbb{Y}^n \rightarrow \mathscr{M}^{(n)}$, with average probability of decoding error

$$\mathbf{P}^{(n)}_{error} \triangleq \frac{1}{M^{(n)}} \sum_{m \in \mathscr{M}^{(n)}} \mathbf{P}^g\big\{d_n(Y^n) \neq m \big| M = m\big\}. \qquad \text{(I.3)}$$

The messages $M : \Omega \rightarrow \mathscr{M}^{(n)}$ are independent of V^n, i.e., $\mathbf{P}_{V^n|M} = \mathbf{P}_{V^n}$. We emphasize that, in general $\mathbf{P}^{(n)}_{error}$ depends on the distribution $\mathbf{P}_{V^n}$ and g, and this is different for different choices of the distribution of $\mathbf{P}_{V_1}$ of the initial RV V_1. *The code rate* is $r_n \triangleq \frac{1}{n}\log M^{(n)}$. A rate R is called an *achievable rate*, if there exists an encoder and decoder sequence satisfying $\lim_{n\longrightarrow\infty} \mathbf{P}^{(n)}_{error} = 0$ and $\liminf_{n\longrightarrow\infty} \frac{1}{n}\log M^{(n)} \geq R$. The operational *nonfeedback capacity* is defined by $C^{op}(\kappa) \triangleq \sup\{R \big| R \text{ is achievable}\}, \forall \mathbf{P}_{V_1}$, i.e., *it is required to be independent of the choice of the initial RV distribution* $\mathbf{P}_{V_1}$.

Two main fundamental differences from [1]–[9], [11]–[13], are i) the consideration of nonstationary and unstable noise, which rules out the characterization of capacity using a frequency-domain approach, and
ii) the requirement that $C^{op}(\kappa)$ does not depend on the distribution of the initial RV V_1, i.e., $\mathbf{P}_{V_1}$, which induces $\mathbf{P}_{Y_1}$ (which is challenging to show due the generality of the noise in i)).
To deal with i), we invoke a time-domain approach. Our starting point is an information theoretic characterization of the $n-$finite transmission, or block length without feedback (n-FTwFI) capacity, denoted by $C_n(\kappa, \mathbf{P}_{Y_1})$, which is analogous to the one derived by Cover and Pombra [10], generalized to the MIMO AGN channel (I.1). To deal with ii), we derive an equivalent *sequential characterization* of $C_n(\kappa, \mathbf{P}_{Y_1})$, and we identify conditions such that $\lim_{n\longrightarrow\infty} \frac{1}{n} C_n(\kappa, \mathbf{P}_{Y_1}) = C(\kappa)$, i.e., for all initial distributions of $\mathbf{P}_{Y_1}$ and hence $\mathbf{P}_{V_1}$. Our approach is motivated by the feedback capacity characterization of MIMO AGN channels with memory presented in [14], and the analysis of single-input single-output (SISO) of [15]–[19]. However, it will become obvious that the treatment of

[1] Notation '$\mathbf{E}^g$ indicates that the corresponding distribution $\mathbf{P}$ depends on the encoding strategy g.

nonfeedback capacity in time-domain (for unstable noise) is much more difficult compare to that of feedback capacity.

II. Asymptotic Characterization of Capacity

Notation. $\mathbb{Z}_+ \triangleq \{1,2,\ldots\}$, $\mathbb{R} \triangleq (-\infty,\infty)$, $\mathbb{R}^m$ is the finite-dimensional Euclidean space, and $\mathbb{R}^{n\times m}$ is the set of n by m matrices. $I_n \in \mathbb{R}^{n\times n}$ denotes the identity matrix, $tr(A)$ denotes the trace of $A \in \mathbb{R}^{n\times n}, n \in \mathbb{Z}_+$. $\mathbb{C} \triangleq \{a+jb : (a,b) \in \mathbb{R}\times\mathbb{R}\}$ is the space of complex numbers, and $\mathbb{D}_o \triangleq \{c \in \mathbb{C} : |c| < 1\}$. $spec(A) \subset \mathbb{C}$ is the spectrum of a matrix $A \in \mathbb{R}^{q\times q}, q \in \mathbb{Z}_+$ (the set of all its eigenvalues). A matrix $A \in \mathbb{R}^{q\times q}$ is called exponentially stable if all its eigenvalues are within the open unit disc, that is, $spec(A) \subset \mathbb{D}_o$. $X \in G(\mu_X, K_X), K_X \succeq 0$ denotes a Gaussian distributed RV X, with $\mu_X = \mathbf{E}\{X\}$ and $K_X = cov(X,X) \triangleq \mathbf{E}\big\{\big(X-\mathbf{E}\{X\}\big)\big(X-\mathbf{E}\{X\}\big)^T\big\} \succeq 0$. Given another Gaussian RV $Y : \Omega \to \mathbb{R}^{n_y}$, which is jointly Gaussian distributed with X, i.e., with joint distribution $\mathbf{P}_{X,Y}$, the conditional covariance of X given Y is $K_{X|Y} = cov(X,X\big|Y) \triangleq \mathbf{E}\big\{\big(X-\mathbf{E}\{X\big|Y\}\big)\big(X-\mathbf{E}\{X\big|Y\}\big)^T\big|Y\big\}$.
Throughout this paper, we consider the noise of Definition II.1.

Definition II.1. *A time-varying partially observable state space (PO-SS) realization of the Gaussian noise V^n, is*

$$S_{t+1} = A_tS_t + B_tW_t, \quad t=1,\ldots,n-1 \tag{II.4}$$
$$V_t = C_tS_t + N_tW_t, \quad t=1,\ldots,n, \tag{II.5}$$
$$S_1 \in G(\mu_{S_1}, K_{S_1}), \quad K_{S_1} \succeq 0, \tag{II.6}$$
$$W_t \in G(0, K_{W_t}), \quad K_{W_t} \succ 0, \quad t=1\ldots,n, \tag{II.7}$$
$$S_t : \Omega \to \mathbb{R}^{n_s},\ W_t : \Omega \to \mathbb{R}^{n_w},\ V_t : \Omega \to \mathbb{R}^{n_y}, \tag{II.8}$$
$$R_t \triangleq N_tK_{W_t}N_t^T \succ 0, \quad t=1,\ldots,n \tag{II.9}$$

where $W_t, t=1\ldots,n$ is an independent Gaussian process, independent of S_1. n_y, n_s, n_w are arbitrary positive integers. Note, $\mathbf{P}_{V^n}$ depends on $\mathbf{P}_{V_1}$ and hence on the choice of $\mathbf{P}_{S_1}$.

Converse Coding Theorem. Suppose there exists a sequence of achievable nonfeedback codes with error probability $\mathbf{P}^{(n)}_{error} \longrightarrow 0$, as $n \longrightarrow \infty$, then $R \leq \lim_{n\longrightarrow\infty}\frac{1}{n}C_n(\kappa, \mathbf{P}_{Y_1})$, where $C_n(\kappa,\mathbf{P}_{Y_1})$ is the sequential characterization of the n-FTwFI capacity formula [14, Section I, III] (note that (II.11) follows from [10]),

$$C_n(\kappa,\mathbf{P}_{Y_1}) = \sup_{\frac{1}{n}\mathbf{E}\left\{\sum_{t=1}^n ||X_t||^2_{\mathbb{R}^{n_x}}\right\}\leq\kappa} \sum_{t=1}^n I(X_t,V^{t-1};Y_t|Y^{t-1}) \tag{II.10}$$
$$= \sup_{\frac{1}{n}\mathbf{E}\left\{\sum_{t=1}^n ||X_t||^2_{\mathbb{R}^{n_x}}\right\}\leq\kappa} H(Y^n) - H(V^n) \in [0,\infty] \tag{II.11}$$

where (II.11) follows from the channel definition (I.1) (if the probability density functions exist) and the supremum is over $\mathbf{P}_{X_t|X^{t-1}}, t=1,\ldots,n$ induced by jointly Gaussian inputs X^n,

$$X_t = \sum_{j=1}^{t-1}\Lambda_{t,j}X_j + Z^o_t = \Lambda_t\mathbf{X}^{t-1} + Z^o_t,\ X_1 = Z^o_1, \tag{II.12}$$
$$Z^o_t \in G(0, K_{Z^o_t}),\ K_{Z^o_t} \succeq 0,\ t=1,\ldots,n, \text{ indep. Gaus.,} \tag{II.13}$$
$$Z^o_t \text{ independent of } (V^{t-1}, X^{t-1}, Y^{t-1}, Z^{o,t-1}), \forall t, \tag{II.14}$$
$$\Lambda_t \in \mathbb{R}^{n_x\times(t-1)n_x}, \forall t \text{ is nonrandom.} \tag{II.15}$$

The consideration of unstable noise V^n implies Y^n is unstable, therefore for the asymptotic analysis, we need to use the two innovations processes of V^n and Y^n, as in [15]–[18], giving rise the characterization of $C_n(\kappa,\mathbf{P}_{Y_1}) \in [0,\infty]$,

$$C_n(\kappa,\mathbf{P}_{Y_1}) = \sup_{(\Lambda_t, K_{Z_t}), t=1,\ldots,n, \frac{1}{n}\mathbf{E}\left\{\sum_{t=1}^n ||X_t||^2_{\mathbb{R}^{n_x}}\right\}\leq\kappa} \Bigg\{ \sum_{t=1}^n \Big(H(I_t) - H(\hat{I}_t)\Big)\Bigg\}, \tag{II.16}$$
$$I_t \triangleq Y_t - \mathbf{E}\Big\{Y_t\Big|Y^{t-1}\Big\}, \quad \hat{I}_t \triangleq V_t - \mathbf{E}\Big\{V_t\Big|V^{t-1}\Big\}. \tag{II.17}$$

where $I_t, \hat{I}_t$ are the innovations processes of Y^n, V^n. Clearly, the analysis of the convergence properties of $\lim_{n\longrightarrow\infty}\frac{1}{n}C_n(\kappa,\mathbf{P}_{Y_1})$, is directly related to the convergence properties of $(I_t, \hat{I}_t, X_t), t = 1,2,\ldots,n, \frac{1}{n}\mathbf{E}\big\{\sum_{t=1}^n ||X_t||^2_{\mathbb{R}^{n_x}}\big\}$, as $n \longrightarrow \infty$.

State Space Realization of Channel Input. By (II.12) the input process X^n is causal, and not finite-memory. Hence, it can be generated by the infinite-dimensional state space realization,

$$\Xi_{t+1} = F_t\Xi_t + G_tZ_t, \quad t=1,\ldots,n-1, \tag{II.18}$$
$$X_t = \Gamma_t\Xi_t + D_tZ_t, \quad t=1,\ldots,n, \tag{II.19}$$
$$\Xi_1 \in G(\mu_{\Xi_1}, K_{\Xi_1}),\ K_{\Xi_1} \succeq 0, \tag{II.20}$$
$$Z_t \in G(0, K_{Z_t}), \quad K_{Z_t} \succeq 0, \quad t=1\ldots,n, \tag{II.21}$$
$$Z^n \text{ indep. seq.,} \quad (\Xi_1, Z^n, W^n) \text{ mutually indep.} \tag{II.22}$$
$$\Xi_t : \Omega \to \mathbb{R}^{n_\xi},\ Z_t : \Omega \to \mathbb{R}^{n_z},\ X_t : \Omega \to \mathbb{R}^{n_x} \tag{II.23}$$

where n_ξ is the dimension of Ξ_t which is nondecreasing with t, and n_z is an arbitrary finite positive integer, and $(F_t, G_t, \Gamma_t, D_t, K_{\Xi_1}, K_{Z_t})$ are nonrandom matrices $\forall t$. In this paper, we restrict our analysis to asymptotically time-invariant matrices, $\lim_{n\longrightarrow\infty}(A_n, B_n, C_n, N_n, K_{W_n}) = (A,B,C,N,K_W), K_W \succ 0$, $\lim_{n\longrightarrow\infty}(F_n, G_n, \Gamma_n, D_n, K_{Z_n}) = (F,G,\Gamma,D,K_Z), K_Z \succeq 0$. For such restriction, follows directly that a necessary condition for convergence of the average power, $\lim_{n\longrightarrow\infty}\frac{1}{n}\mathbf{E}\big\{\sum_{t=1}^n ||X_t||^2_{\mathbb{R}^{n_x}}\big\} \in [0,\infty)$, is the stability of F, i.e., the eigenvalues of F lie inside the unit disc in the space of complex numbers. The stability of F further implies that $(\Xi_n, X_n), n = 1,2,\ldots$ is asymptotically stationary. Consequently, whether the asymptotic dimension of the state space realization is asymptotically finite, $n_\xi < \infty$, is determined from the Hankel matrix, of the covariance matrix $R_X(t) \triangleq \mathbf{E}\big\{X(t+s)X^T(s)\big\}, t = 1,2,\ldots$ as follows. Define the finite Hankel matrix, for $(k,m) \in \mathbb{Z}_+\times\mathbb{Z}_+, \mathbb{Z}_+ = \{1,2,\ldots\}$:

$$H_X(k,m) = \begin{pmatrix} R_X(1) & R_X(2) & \ldots & R_X(m) \\ R_X(2) & R_X(3) & \ldots & R_X(m+1) \\ \vdots & \vdots & \vdots & \vdots \\ R_X(k) & R_X(k+1) & \ldots & R_X(k+m-1) \end{pmatrix}$$

Define the rank of the infinite Hankel matrix by, $rank(H_X) \triangleq \sup_{k,m\in\in\mathbb{Z}_+\times\mathbb{Z}_+} rank(H_X(k,m)) \in \mathbb{Z}_+\cup\{+\infty\}$. The state space realization is asymptotically finite-dimensional if and only if $rank(H_X) < \infty$. For our analysis we assume $rank(H_X) < \infty$, and hence there exists a finite integer, $n_\xi \leq rank(H_X)$.

Main Problems and Assumptions. Now, we state our main problems and accompanied assumptions.

Problem #1. Identify conditions, such that asymptotic limit exists and does not depend on $\mathbf{P}_{Y_1}$ (and hence on $\mathbf{P}_{V_1}$),

$$C^o(\kappa,\mathbf{P}_{Y_1}) \stackrel{\triangle}{=} \lim_{n\longrightarrow\infty}\frac{1}{n}C_n(\kappa,\mathbf{P}_{Y_1}) = C(\kappa)\in[0,\infty), \forall \mathbf{P}_{Y_1}. \quad \text{(II.24)}$$

Assumptions II.1. *Considered for the asymptotic analysis are*

Case 1: *Time-invariant,*

$$(A_n,B_n,C_n,N_n,K_{W_n}) = (A,B,C,N,K_W),\ K_W\succ 0,\quad \forall n \quad \text{(II.25)}$$
$$(F_n,G_n,\Gamma_n,D_n,K_{Z_n}) = (F,G,\Gamma,D,K_Z), K_Z\succeq 0,\quad \forall n. \quad \text{(II.26)}$$

Case 2: *Asymptotically time-invariant,*

$$\lim_{n\longrightarrow\infty}(A_n,B_n,C_n,N_n,K_{W_n}) = (A,B,C,N,K_W),\ K_W\succ 0, \quad \text{(II.27)}$$
$$\lim_{n\longrightarrow\infty}(F_n,G_n,\Gamma_n,D_n,K_{Z_n}) = (F,G,\Gamma,D,K_Z),\ K_Z\succeq 0 \quad \text{(II.28)}$$

where the limits are element wise. We also assume $rank(H_X)<\infty$, hence $n_\xi \leq rank(H_X)$, and the realizations are finite-dimensional and of minimal dimensions.

Problem #2. Determine conditions for the limit and the supremum to be interchanged, so that

$$C^\infty(\kappa,\mathbf{P}_{Y_1}) \stackrel{\triangle}{=} \sup_{\lim_{n\longrightarrow\infty}\frac{1}{n}\mathbf{E}\left\{\sum_{t=1}^n ||X_t||^2_{\mathbb{R}^{n_x}}\right\}\leq\kappa} \lim_{n\longrightarrow\infty}\frac{1}{n}\Big\{\sum_{t=1}^n\Big(H(I_t) - H(\hat{I}_t)\Big)\Big\} = C^o(\kappa,\mathbf{P}_{Y_1}) = C(\kappa)\in[0,\infty),\quad \forall\mathbf{P}_{Y_1}. \quad \text{(II.29)}$$

Main Results. For Cases 1 and 2, we identify conditions on 1) channel matrices (H,A,B,C,N,K_W) and 2) channel input matrices $(F,G,K_Z,\Gamma,D), K_Z\succeq 0$, such that the limit in (II.24) exists, is independent of $\mathbf{P}_{Y_1}$ (hence of $\mathbf{P}_{V_1}$), and capacity is

$$C(\kappa) \stackrel{\triangle}{=} \sup \frac{1}{2}\ln\Big\{\frac{\det\big(\mathbf{C}\Pi\mathbf{C}^T + \mathbf{D}K_{\overline{W}}\mathbf{D}^T\big)}{\det\big(C\Sigma C^T + NK_WN^T\big)}\Big\}^+ \quad \text{(II.30)}$$

where the supremum is over (F,G,Γ,D,K_Z) and

$$K_Z\succeq 0,\ tr(\Gamma P\Gamma^T + DK_ZD^T)\leq\kappa, \quad \text{(II.31)}$$

$\Pi\succeq 0,\ \Sigma\succeq 0$ satisfy matrix Algebraic Riccati Eqns, (II.32)

$P\succeq 0$ satisfies a matrix Lyapunov Equation (II.33)

and where $\{\cdot\}^+ \stackrel{\triangle}{=} \max\{1,\cdot\}$, and $(\mathbf{C},\mathbf{D},K_{\overline{W}})$ are specific matrices given in Theorem II.1. Further, since the convergence in (II.30) and (II.29) are uniform $\forall\mathbf{P}_{Y_1}$, then asymptotic equipartition (AEP) and information stability hold, which imply $C(\kappa)$ is the nonfeedback capacity, even for unstable channels, similar to feedback capacity in [20], [21].

A. Sequential Characterizations of $n-$FTwFI Capacity

Next, we determine the characterization of $C_n(\kappa,\mathbf{P}_{Y_1})$ of (II.16), and its dependence on two DREs and a Lyapunov equation.

Theorem II.1. *Sequential characterization of $C_n(\kappa,\mathbf{P}_{Y_1})$. Consider the MIMO channel (I.1), the noise of Definition II.1 and input (II.12)-(II.15), and assume n_ξ is finite. Define*

$$\Theta_t \stackrel{\triangle}{=} \begin{pmatrix} \Xi_t^T & S_t^T\end{pmatrix}^T,\quad \overline{W}_t \stackrel{\triangle}{=} \begin{pmatrix} Z_t^T & W_t^T\end{pmatrix}^T,$$
$$\Pi_t \stackrel{\triangle}{=} cov\big(\Theta_t,\Theta_t\big|Y^{t-1}\big) = \mathbf{E}\Big\{\Big(\Theta-\widehat{\Theta}_t\Big)\Big(\Theta_t-\widehat{\Theta}_t\Big)^T\Big\},$$
$$\widehat{\Theta}_t \stackrel{\triangle}{=} \mathbf{E}\Big\{\Theta_t\Big|Y^{t-1}\Big\},\ t=2,\ldots,n,\ \widehat{\Theta}_1 \stackrel{\triangle}{=} \mu_{\Theta_1},\ \Pi_1 \stackrel{\triangle}{=} K_{\Theta_1},$$
$$P_t \stackrel{\triangle}{=} cov\big(\Xi_t,\Xi_t\big) = \mathbf{E}\Big\{\Big(\Xi_t - \mathbf{E}\Big\{\Xi_t\Big\}\Big)\Big(\Xi_t-\mathbf{E}\Big\{\Xi_t\Big\}\Big)^T\Big\}.$$

(i) The joint Gaussian process (X^n,Y^n,V^n) is represented by

$$\Theta_{t+1} = \mathbf{A}_t\Theta_t + \mathbf{B}_t\overline{W}_t,\quad t=1,\ldots,n-1, \quad \text{(II.34)}$$
$$Y_t = \mathbf{C}_t\Theta_t + \mathbf{D}_t\overline{W}_t,\quad t=1,\ldots,n \quad \text{(II.35)}$$
$$\mathbf{A}_t \stackrel{\triangle}{=} \begin{pmatrix} F_t & 0\\ 0 & A_t\end{pmatrix},\quad \mathbf{B}_t \stackrel{\triangle}{=} \begin{pmatrix} G_t & 0\\ 0 & B_t\end{pmatrix}, \quad \text{(II.36)}$$
$$\mathbf{C}_t \stackrel{\triangle}{=} \begin{pmatrix} H_t\Gamma_t & C_t\end{pmatrix},\quad \mathbf{D}_t \stackrel{\triangle}{=} \begin{pmatrix} H_tD_t & N_t\end{pmatrix} \quad \text{(II.37)}$$

where $\mathbf{A}_t,\mathbf{B}_t,\mathbf{C}_t,\mathbf{D}_t$ are appropriate matrices.

(ii) The error $\widehat{E}_t = \Theta_t - \widehat{\Theta}_t$, and covariance $\Pi_t \stackrel{\triangle}{=} \mathbf{E}\{\widehat{E}_t\widehat{E}_t^T\}$, satisfy the recursion and generalized matrix DRE,

$$\widehat{E}_{t+1} = \mathbf{F}_t^{CL}(\Pi_t)\widehat{E}_t + \big(\mathbf{B}_t - \mathbf{F}_t(\Pi_t)\mathbf{D}_t\big)\overline{W}_t,\ t=1,\ldots,n, \quad \text{(II.38)}$$
$$\mathbf{F}_t^{CL}(\Pi_t) = \mathbf{A}_t - \mathbf{F}_t(\Pi_t)\mathbf{C}_t, \quad \text{(II.39)}$$
$$\mathbf{F}_t(\Pi_t) = \big(\mathbf{A}_t\Pi_t\mathbf{C}_t^T + \mathbf{B}_tK_{\overline{W}_t}\mathbf{D}_t^T\big)\big(\mathbf{D}_tK_{\overline{W}_t}\mathbf{D}_t^T + \mathbf{C}_t\Pi_t\mathbf{C}_t^T\big)^{-1}.$$
$$\Pi_{t+1} = \mathbf{A}_t\Pi_t\mathbf{A}_t^T + \mathbf{B}_tK_{\overline{W}_t}\mathbf{B}_t^T - \Big(\mathbf{A}_t\Pi_t\mathbf{C}_t^T + \mathbf{B}_tK_{\overline{W}_t}\mathbf{D}_t^T\Big) \cdot\Big(\mathbf{D}_tK_{\overline{W}_t}\mathbf{D}_t^T + \mathbf{C}_t\Pi_t\mathbf{C}_t^T\Big)^{-1}\Big(\mathbf{A}_t\Pi_t\mathbf{C}_t^T + \mathbf{B}_tK_{\overline{W}_t}\mathbf{D}_t^T\Big)^T,$$
$$\Pi_t\succeq 0,\quad \Pi_1 = K_{\Theta_1}\succeq 0,\quad t=1,\ldots,n. \quad \text{(II.40)}$$

(iii) The innovations process I_t of Y^n for $t=1,\ldots,n$, is

$$I_t \stackrel{\triangle}{=} Y_t - \mathbf{E}\big\{Y_t\big|Y^{t-1}\big\} = \mathbf{C}_t\big(\Theta_t - \widehat{\Theta}_t\big) + \mathbf{D}_t\overline{W}_t, \quad \text{(II.41)}$$
$$I_t\in G(0,K_{I_t}),\quad K_{I_t} = \mathbf{C}_t\Pi_t\mathbf{C}_t^T + \mathbf{D}_tK_{\overline{W}_t}\mathbf{D}_t^T. \quad \text{(II.42)}$$

(iv) The matrix $P_t = cov\big(\Xi_t,\Xi_t\big)$ satisfies Lyapunov recursion,

$$P_{t+1} = F_tP_tF_t^T + G_tK_{Z_t}G_t^T,\quad P_t\succeq 0,\quad P_1 = K_{\Xi_1}. \quad \text{(II.43)}$$

(v) The average power constraint is

$$\frac{1}{n}\mathbf{E}\Big\{\sum_{t=1}^n ||X_t||^2_{\mathbb{R}^{n_x}}\Big\} = \frac{1}{n}\sum_{t=1}^n tr(\Gamma_tP_t\Gamma_t^T + D_tK_{Z_t}D_t^T). \quad \text{(II.44)}$$

(vii) The entropy of Y^n is $H(Y^n) = \sum_{t=1}^n H(I_t)$, is given by

$$H(Y^n) = \frac{1}{2}\sum_{t=1}^n\ln\big((2\pi e)^{n_y}\det\big(\mathbf{C}_t\Pi_t\mathbf{C}_t^T + \mathbf{D}_tK_{\overline{W}_t}\mathbf{D}_t^T\big)\big) \quad \text{(II.45)}$$

and the entropy of V^n is $H(V^n) = \sum_{t=1}^n H(\hat{I}_t)$, is given by

$$H(V^n) = \frac{1}{2}\sum_{t=1}^n\ln\big((2\pi e)^{n_y}\det\big(C_t\Sigma_tC_t^T + N_tK_{W_t}N_t^T\big)\big) \quad \text{(II.46)}$$
$$\hat{I}_t\in G(0,K_{\hat{I}_t})\ \text{ an orth. innov. proc. indep. of } V^{t-1}, \quad \text{(II.47)}$$
$$K_{\hat{I}_t} \stackrel{\triangle}{=} cov(\hat{I}_t,\hat{I}_t) = C_t\Sigma_tC_t^T + N_tK_{W_t}N_t^T = K_{V_t|V^{t-1}}. \quad \text{(II.48)}$$

where Σ_t satisfies ($M^{CL}(\Sigma_t)$ is similar to $\mathbf{F}_t^{CL}(\Pi_t)$) the DRE

$$\Sigma_{t+1} = A_t\Sigma_tA_t^T + B_tK_{W_t}B_t^T - \Big(A_t\Sigma_tC_t^T + B_tK_{W_t}N^T\Big) \cdot\Big(N_tK_{W_t}N_t^T + C_t\Sigma_tC_t^T\Big)^{-1}\Big(A_t\Sigma_tC_t^T + B_tK_{W_t}N_t^T\Big)^T,$$
$$\Sigma_t\succeq 0,\quad t=1,\ldots,n,\quad \Sigma_1 = K_{S_1}\succeq 0. \quad \text{(II.49)}$$

(v) An equivalent characterization of $C_n(\kappa,\mathbf{P}_{Y_1})$ is

$$C_n(\kappa,\mathbf{P}_{Y_1}) = \sup_{\frac{1}{n}\mathbf{E}\left\{\sum_{t=1}^n ||X_t||^2_{\mathbb{R}^{n_x}}\right\}\leq\kappa}\frac{1}{2}\sum_{t=1}^n\ln\Big\{\frac{\det(K_{I_t})}{\det(K_{\hat{I}_t})}\Big\}^+ \quad \text{(II.50)}$$

where the supremum is over $(F_t,G_t,\Gamma_t,D_t,K_{Z_t})$, $t=1,\ldots,n$.

Proof. Since n_ξ is finite, there always exists a realization (II.18)-(II.23). We compute (II.16) using the two innovations processes of Y^n, V^n, given in (II.17). Upon substituting (II.18)-(II.23) into the channel (I.1) we obtain representation (i). Then by applying the general equations of Kalman filtering [22], and evaluating the entropies we obtain statements (ii)-(v).

B. Convergence of Generalized DRE and Lyapunov Equations

To address the limit $\lim_{n\longrightarrow\infty}\frac{1}{n}C_n(\kappa,\mathbf{P}_{V_1})$, under Cases 1, 2 we need to investigate the convergence properties of generalized matrix DREs (II.40), (II.49) and Lyapunov matrix difference equation (II.43), using the properties in [21, Theorem A.1] and [15, Theorem III.2]. We present sufficient conditions for convergence in Corollary II.1, Theorem II.2, Theorem II.3.

Corollary II.1. *Consider Case 1 or Case 2*
Let $\Sigma_t, t=1,2,\ldots$ denote the solution of the matrix DRE (II.49). Let $\Sigma=\Sigma^T\succeq 0$ be a solution of the corresponding ARE

$$\Sigma = A\Sigma A^T + BK_WB^T - \Big(A\Sigma C^T + BK_WN^T\Big) .\Big(NK_WN^T + C\Sigma C^T\Big)^{-1}\Big(A\Sigma C^T + BK_WN^T\Big)^T. \quad \text{(II.51)}$$

Define the matrices

$$A^* \stackrel{\triangle}{=} A - BK_WN^T\big(NK_WN^T\big)^{-1}C, \quad G \stackrel{\triangle}{=} B,$$
$$B^* \stackrel{\triangle}{=} K_W - K_WN^T\Big(NK_WN^T\Big)^{-1}\Big(K_WN^T\Big)^T, \quad \text{(II.52)}$$

and suppose (see [22]–[24] for definitions)

$$\{A, C\} \text{ is detectable, and } \{A^*, GB^{*,\frac{1}{2}}\} \text{ is stabilizable.} \quad \text{(II.53)}$$

Any solution $\Sigma_t, t=1,2,\ldots,n$ to the generalized matrix DRE (II.49) with arbitrary initial condition $\Sigma_1\succeq 0$, is such that $\lim_{n\longrightarrow\infty}\Sigma_n=\Sigma$, where $\Sigma\succeq 0$ is the unique solution of matrix ARE (II.51) such that $M^{CL}(\Sigma)\in\mathbb{D}_o$.

Proof. For Case 1, the convergence of $\Sigma_n, n=1,2,\ldots$, follows from the detectability and stabilizability conditions [22]–[24]. For Case 2, the statements of convergence of $\Sigma_n, n=1,2,\ldots$ hold, due to continuity property of solutions of generalized difference Riccati equations, with respect to its coefficients.

Theorem II.2. *Consider Case 1 or Case 2.*
Let $\Pi_t, t=1,\ldots,$ denote the solution of the DRE (II.40). Let $\Pi=\Pi^T\succeq 0$ be a solution of the corresponding ARE

$$\Pi = \mathbf{A}\Pi\mathbf{A}^T + \mathbf{B}K_{\overline{W}}\mathbf{B}^T - \Big(\mathbf{A}\Pi\mathbf{C}^T + \mathbf{B}K_{\overline{W}}\mathbf{D}^T\Big) .\Big(\mathbf{D}K_{\overline{W}}\mathbf{D}^T + \mathbf{C}\Pi\mathbf{C}^T\Big)^{-1}\Big(\mathbf{A}\Pi\mathbf{C}^T + \mathbf{B}K_{\overline{W}}\mathbf{D}^T\Big)^T. \quad \text{(II.54)}$$

Define the matrices [22]–[24]

$$\mathbf{A}^* \stackrel{\triangle}{=} \mathbf{A} - \mathbf{B}K_{\overline{W}}\mathbf{D}^T\big(\mathbf{D}K_{\overline{W}}\mathbf{D}^T\big)^{-1}\mathbf{C}, \quad \mathbf{G} \stackrel{\triangle}{=} \mathbf{B},$$
$$\mathbf{B}^* \stackrel{\triangle}{=} K_{\overline{W}} - K_{\overline{W}}\mathbf{D}^T\Big(\mathbf{D}K_{\overline{W}}\mathbf{D}^T\Big)^{-1}\Big(K_{\overline{W}}\mathbf{D}^T\Big)^T. \quad \text{(II.55)}$$

Suppose [22]–[24]

$$\{\mathbf{A}, \mathbf{C}\} \text{ is detectable and } \{\mathbf{A}^*, \mathbf{G}\mathbf{B}^{*,\frac{1}{2}}\} \text{ is stabilizable.} \quad \text{(II.56)}$$

Any solution $\Pi_t, t=1,2,\ldots,n$ to the generalized matrix DRE (II.40) with arbitrary initial condition $\Pi_1\succeq 0$, is such that $\lim_{n\longrightarrow\infty}\Pi_n=\Pi$, where $\Pi\succeq 0$ is the unique solution of the generalized matrix ARE (II.54), such that $spec\big(\mathbf{F}^{CL}(\Pi)\big)\in\mathbb{D}_o$.

Proof. Similar to Corollary II.1.

Theorem II.3 identifies conditions for the average power (II.44) to converge, using $P_t = cov\big(\Xi_t,\Xi_t\big)$, which satisfies (II.43).

Theorem II.3. *Convergence of average power*
Consider the average power of Thm II.1, for Cases 1 or 2. Let $P_t, t=1\ldots,n$ be a solution of Lyapunov recursion (II.43). Let $P\succeq 0$ be a solution of

$$P = FPF^T + GK_ZG^T. \quad \text{(II.57)}$$

Suppose F is an exponentially stable matrix. Any solution $P_t, t=1,2,\ldots,n$ to the Lyapunov recursion DRE (II.43), with arbitrary initial condition $P_1\succeq 0$, is such that $\lim_{n\longrightarrow\infty}P_n=P$, where $P\succeq 0$ is the unique solution of (II.57). Moreover,

$$\lim_{n\longrightarrow\infty}\frac{1}{n}\mathbf{E}\Big\{\sum_{t=1}^{n}||X_t||^2_{\mathbb{R}^{n_x}}\Big\} = \lim_{n\longrightarrow\infty}\frac{1}{n}\sum_{t=1}^{n}tr\Big(\Gamma P_t\Gamma^T + DK_ZD^T\Big)$$
$$= tr\Big(\Gamma P\Gamma^T + DK_ZD^T\Big),\ \forall P_1\succeq 0. \quad \text{(II.58)}$$

Proof. The conditions for Case 1 are known [22]–[24], and imply (II.58). For Case 2, we use the continuity of solutions of Lyapunov equations, with respect to their coefficients.

C. Asymptotic Characterizations of Nonfeedback Capacity

First, we address Problem #2 and then Problem #1.

Theorem II.4. *Characterization of $C^\infty(\kappa,\mathbf{P}_{Y_1})$ for Case 1*
Consider the time-invariant noise and channel input strategies of Case 1, i.e., (II.25) and (II.26) hold.
Define the per unit time limit and supremum by[2]

$$C^\infty(\kappa,\mathbf{P}_{Y_1}) \stackrel{\triangle}{=} \sup_{\mathscr{P}_\infty(\kappa)}\lim_{n\longrightarrow\infty}\frac{1}{2n}\sum_{t=1}^{n}\ln\Big\{\frac{\det\big(\mathbf{C}\Pi_t\mathbf{C}^T + \mathbf{D}K_{\overline{W}}\mathbf{D}^T\big)}{\det\big(C\Sigma_tC^T + NK_WN^T\big)}\Big\}^+ \quad \text{(II.59)}$$

$$\mathscr{P}_\infty(\kappa) \stackrel{\triangle}{=} \Big\{(F,G,\Gamma,D,K_Z)\in\mathscr{P}^\infty\Big| \lim_{n\longrightarrow\infty}\frac{1}{n}\sum_{t=1}^{n}tr\big(\Gamma P_t\Gamma^T + DK_ZD^T\big)\le\kappa\Big\}, \quad \text{(II.60)}$$

$\mathscr{P}^\infty \stackrel{\triangle}{=} \Big\{(F,G,\Gamma,D,K_Z), K_Z\succeq 0,$ *such that the following hold,*

(i) the detectability and stabilizability of (II.53), (II.61)

(ii) the detectability and stabilizability of (II.56), (II.62)

(iii) F is exponentially stable$\Big\}$. (II.63)

Then, $C^\infty(\kappa,\mathbf{P}_{Y_1})$ is given by

$$C^\infty(\kappa,\mathbf{P}_{Y_1}) = C^\infty(\kappa), \quad \forall\mathbf{P}_{Y_1}$$
$$C^\infty(\kappa) \stackrel{\triangle}{=} \sup_{\mathscr{P}^\infty(\kappa)}\frac{1}{2}\ln\Big\{\frac{\det\big(\mathbf{C}\Pi\mathbf{C}^T + \mathbf{D}K_{\overline{W}}\mathbf{D}^T\big)}{\det\big(C\Sigma C^T + NK_WN^T\big)}\Big\}^+ \quad \text{(II.64)}$$
$$\mathscr{P}^\infty(\kappa) \stackrel{\triangle}{=} \Big\{(F,G,\Gamma,D,K_Z)\in\mathscr{P}^\infty\Big| tr\big(\Gamma P\Gamma^T + DK_ZD^T\big)\le\kappa\Big\}$$

and $\Sigma\succeq 0$ and $\Pi\succeq 0$ are the unique and stabilizable solutions, i.e., $spec\big(M^{CL}(\Sigma)\big)\in\mathbb{D}_o$ and $spec\big(\mathbf{F}^{CL}(\Pi)\big)\in\mathbb{D}_o$ of the generalized matrix AREs (II.51) and (II.54) respectively, $P\succeq 0$ is the unique solution of the matrix Lyapunov equation (II.57),

[2]By [25], if at any time t, the information $\max\{0, H(Y_t|Y^{t-1}) - H(V_t|V^{t-1})\}\in[0,\infty]$ is $+\infty$, then at this time no transmission is allowed.

provided there exists $\kappa \in [0,\infty)$, such that $\mathscr{P}^\infty(\kappa)$ is non-empty. Moreover, the optimal $(F,G,\Gamma,D,K_Z) \in \mathscr{P}^\infty(\kappa)$, is such that,
(i) if the noise is stable, then the input and the output processes $(X_t,Y_t), t=1,\ldots$ are asymptotic stationary and
(ii) if the noise is unstable, then the input and the innovations processes $(X_t,I_t), t=1,\ldots$ are asymptotic stationary.

Proof. This follows from the definition of the set $\mathscr{P}^\infty$, which imply Corollary II.1, Theorem II.2 and Theorem II.3 hold. The complete steps are given in [26] (due to space limitation).

Next, we show that Theorem II.4 remains valid for Case 2.

Corollary II.2. *Characterization of $C^\infty(\kappa,\mathbf{P}_{Y_1})$ for Case 2 Consider the asymptotically time-invariant noise and channel input strategies of Case 2, i.e., (II.27) and (II.28) hold. Define the per unit time limit and supremum by*

$$C^{\infty,+}(\kappa,\mathbf{P}_{Y_1}) \stackrel{\triangle}{=} \sup_{\mathscr{P}^+_\infty(\kappa)} \lim_{n\longrightarrow\infty} \frac{1}{2n}\Big\{ \sum_{t=1}^n \ln\Big\{\frac{\det\big(\mathbf{C}_t\Pi_t\mathbf{C}_t^T+\mathbf{D}_tK_{\overline{W}_t}\mathbf{D}_t^T\big)}{\det\big(C_t\Sigma_tC_t^T+N_tK_{W_t}N_t^T\big)}\Big\}^+\Big\} \tag{II.65}$$

$$\mathscr{P}^+_\infty(\kappa) \stackrel{\triangle}{=} \Big\{\{(F_n,G_n,\Gamma_n,D_nK_{Z_n})|n=1,2,\ldots\} \in \mathscr{P}^+_\infty \Big| \lim_{n\longrightarrow\infty}\frac{1}{n}\sum_{t=1}^n tr\big(\Gamma_tP_t\Gamma_t^T+D_tK_{Z_t}D_t^T\big) \leq \kappa\Big\}, \tag{II.66}$$

$$\mathscr{P}^+_\infty \stackrel{\triangle}{=} \Big\{\{(F_n,G_n,\Gamma_n,D_n,K_{Z_n})|n=1,2,\ldots\}\Big| \lim_{n\longrightarrow\infty}(F_n,G_n,\Gamma_n,D_nK_{Z_n}) = (F,G,\Gamma,D,K_Z) \in \mathscr{P}^\infty\Big\}. \tag{II.67}$$

Then, $C^{\infty,+}(\kappa,\mathbf{P}_{Y_1}) = C^\infty(\kappa,\mathbf{P}_{Y_1}) = C^\infty(\kappa) =$ *(II.64)*, $\forall \mathbf{P}_{Y_1}$ (II.68)

and the statements of Theorem II.4.(i), (ii), remain valid.

Proof. The solutions of the DREs and the Lyapunov equation are, $\Sigma_{n+1} = \Sigma_{n+1}(\Sigma_n,A_n,B_n,C_n,N_n,K_{W_n})$, $\Pi_{n+1} = \Pi_{n+1}(\Pi_n,\Sigma_n,\mathbf{A}_n,\mathbf{B}_n,\mathbf{C}_n,\mathbf{D}_n,K_{\overline{W}_n})$,
$P_{n+1} = P_{n+1}(P_n,F_n,G_n,\Gamma_n,D_n,K_{Z_n})$, $n=1,2,\ldots$ and these are continuous with respect to their coefficients. Moreover, for all elements of the set $\mathscr{P}^\infty$, by (II.27), then

$$\lim_{n\longrightarrow\infty}\frac{1}{n}\sum_{t=1}^n tr(\Gamma_tP_t{\Gamma_t}^T+D_tK_{Z_t}{D_t}^T) = tr(\Gamma P\Gamma^T+DK_ZD^T), \forall \mathbf{P}_{Y_1},$$
$$\lim_{n\longrightarrow\infty}\frac{1}{2n}\sum_{t=1}^n \ln\Big\{\frac{\det\big(\mathbf{C}_t\Pi_t\mathbf{C}_t^T+\mathbf{D}_tK_{\overline{W}_t}\mathbf{D}_t^T\big)}{\det\big(C_t\Sigma_tC_t^T+N_tK_{W_t}N_t^T\big)}\Big\}^+ = \frac{1}{2}\ln\Big(\frac{\det\big(\mathbf{C}\Pi\mathbf{C}^T+\mathbf{D}K_{\overline{W}}\mathbf{D}^T\big)}{\det\big(C\Sigma C^T+NK_WN^T\big)}\Big), \quad \forall \mathbf{P}_{Y_1}.$$

The statement follows from the proof of Thm II.4, see [26].

Identity $C^o(\kappa,\mathbf{P}_{Y_1}) = C^\infty(\kappa,\mathbf{P}_{Y_1}) = C^\infty(\kappa), \forall \mathbf{P}_{Y_1}$ for Case 2, is due to the uniform convergence of Theorem II.4, Corollary II.2.

Theorem II.5. *Characterization of $C^o(\kappa,\mathbf{P}_{Y_1})$ for Case 2 Consider the asymptotically time-invariant noise and channel input strategies of Case 2, i.e., (II.27) and (II.28) hold. Define the per unit time limit and supremum by*

$$C^o(\kappa,\mathbf{P}_{Y_1}) \stackrel{\triangle}{=} \lim_{n\longrightarrow\infty} \sup_{\mathscr{P}^{o,+}_n(\kappa)} \frac{1}{2n}\Big\{ \sum_{t=1}^n \ln\Big\{\frac{\det\big(\mathbf{C}_t\Pi_t\mathbf{C}_t^T+\mathbf{D}_tK_{\overline{W}_t}\mathbf{D}_t^T\big)}{\det\big(C_t\Sigma_tC_t^T+N_tK_{W_t}N_t^T\big)}\Big\}^+\Big\} \tag{II.69}$$

$$\mathscr{P}^+_\infty(\kappa) \stackrel{\triangle}{=} \Big\{\{(F_n,G_n,\Gamma_n,D_nK_{Z_n})|n=1,2,\ldots\} \in \mathscr{P}^+_\infty \Big| \frac{1}{n}\sum_{t=1}^n tr(\Gamma_tP_t\Gamma_t^T+D_tK_{Z_t}D_t^T)\big\} \leq \kappa\Big\}. \tag{II.70}$$

Then, $C^o(\kappa,\mathbf{P}_{Y_1}) = C^\infty(\kappa,\mathbf{P}_{Y_1}) = C^\infty(\kappa) =$ *(II.64)*, $\forall \mathbf{P}_{Y_1}$ (II.71)

and the statements of Theorem II.4.(1), (ii), remain valid.

Proof. The derivation uses the uniform limits in the proof of Theorem II.4 and Corollary II.2.

Example II.1. *Consider the scalar channel, (I.1), with $n_y = n_x = 1, H_n = 1$, and $V_n, n=1,\ldots$ an autoregressive noise, AR(c), with $c \in (-\infty,\infty)$, i.e., $V_n = cV_{n-1}+W_n, K_{W_n}=1$.*
(a) Stable $c \in (0,1)$. The power spectral density (PSD) is, $S_V(e^{j\theta}) = \frac{1}{|e^{j\theta}-c|^2}$. By [2, Example 5.5.1], for $\kappa \geq \kappa_{min} \stackrel{\triangle}{=} \frac{2c}{(1-c)^2(1+c)}$, the water-filling capacity is $C^{WF}(\kappa) = \frac{1}{2}\ln\big(\kappa+\frac{1}{1-c^2}\big)$. It can be verified that, for $\kappa \geq \kappa_{min}$, the time-domain capacity using $C(\kappa)$ of (II.30) is achieved by optimal input, $\Xi_{n+1} = c\Xi_n+Z_n, X_n = (c-a)\Xi_n+Z_n$, with parameters, $a = \frac{Ac}{K_Z}$, $A = \kappa - \frac{2c}{(1+c)(1-c)^2}+\frac{1}{(1-c)^2}$, $K_Z = \frac{A(1+c^2)-1+\sqrt{\Delta}}{2}$, $\Delta = \big(1-A(1+c^2)\big)^2-4(Ac)^2$, and $a \in (-1,1)$ holds. Then by calculations follows, $C(\kappa) = (II.30) = C^{WF}(\kappa), \forall \kappa \geq \kappa_{min}$, verifying [2, Example 5.5.1].
(b) Unstable $|c| \geq 1$. The above input X_n is not optimal, because the variance of Ξ_n and hence P_n, as $n \longrightarrow \infty$, grows unbounded. We do not provide the optimal input X_n of $C(\kappa)$ given by (II.30), for $|c| \geq 1$. However, for $|c| \geq 1$ we can easily derive achievable lower bounds on $C(\kappa)$ given by (II.30), by considering sub-optimal finite memory inputs, i.e., IID inputs $X_n = Z_n$ or Markov inputs $X_n = \Lambda X_{n-1}+Z_n$, as in [18, Theorem 3.2, Section IV].

III. Conclusion

New asymptotic characterizations of nonfeedback capacity of MIMO additive Gaussian noise (AGN) channels are presented, when the noise is nonstationary and unstable. The asymptotic characterizations of nonfeedback capacity, involve two generalized matrix algebraic Riccati equations (AREs) of filtering theory and a Lyapunov matrix equation of stability theory of Gaussian systems. Identified, are conditions for uniform convergence of the asymptotic limits, which imply that the nonfeedback capacity is independent of the initial states.

IV. Acknowledgements

This work was supported in parts by the European Regional Development Fund and the Republic of Cyprus through the Research Promotion Foundation Projects EXCELLENCE/1216/0296.

REFERENCES

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

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

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

[4] R. W. Yeung, *Information Theory and Network Coding*, 1st ed. Springer Publishing Company, Incorporated, 2008.

[5] B. S. Tsybakov, "A different approach to finding the capacity of a Gaussian vector channel," *Problems of Information Transmission*, vol. 42, no. 3, pp. 186–196, 2006.

[6] L. H. Brandenburg and A. D. Wyner, "Capacity of the Gaussian channel with memory: The multivariate case," *The Bell System Technical Journal*, vol. 53, no. 5, pp. 745–778, May 1974.

[7] J. Gutirrez-Gutierrez and P. M. Crespo, "Asymptotically equivalent sequences of matrices and Hermitian block Toeplitz matrices with continuous symbols: Applications to MIMO systems," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 5671–5680, December 2008.

[8] J. Gutirrez-Gutierrez, P. M. Crespo, M. Zarraga-Rodriguez, and B. O. Hogstad, "Asymptotically equivalent sequences of matrices and capacity of a discrete-time Gaussian MIMO channel with memory," *IEEE Transactions on Information Theory*, vol. 9, no. 63, pp. 6000–6003, September 2017.

[9] W. Hirt and J. L. Massey, "Capacity of the discrete-time Gaussian channel with intersymbol interference," *IEEE Transactions on Information Theory*, vol. 56, no. 34, pp. 380–388, May 1998.

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

[11] K. Yanaki, "Necessary and sufficient conditions for the capacity of the discrete-time Gaussian channel to be increased by feedback," *IEEE Transactions on Information Theory*, vol. 38, pp. 1788–1791, May 1992.

[12] ——, "An upper bound on the discrete-time Gaussian channel with feedback-II," *IEEE Transactions on Information Theory*, vol. 40, pp. 588–593, May 1994.

[13] H. W. Chen and K. Yanaki, "Refiniements of the half-bit and factor-of-two bounds for capacity in Gaussian channels with feedback," *IEEE Transactions on Information Theory*, vol. 45, pp. 316–325, January 1999.

[14] C. D. Charalambous, C. Kourtellaris, and S. Louka, "Sequential characterizations of Cover and Pombra Gaussian feedback capacity: Generalizations to MIMO channels via sufficient statistic," in *2021 IEEE Information Theory Workshop (ITW), Kanazawa, Japan*, October 17-21 2021.

[15] ——, "New formulas of feedback capacity for AGN channels with memory: A time-domain sufficient statistic approach," submitted for publication, August 2020. [Online]. Available: https://arxiv.org/abs/2010.06226

[16] S. Louka, C. Kourtellaris, and C. D. Charalambous, "Qualitative analysis of feedback capacity of AGN channels driven by unstable versus stable autoregressive moving average noise," in *2021 IEEE Information Theory Workshop (ITW), Kanazawa, Japan*, October 17-21 2021.

[17] C. Kourtellaris, C. D. Charalambous, and S. Loyka, "New formulas of 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. 2073–2078.

[18] ——, "From feedback capacity to tight achievable bounds without feedback for AGN channels with stable and unstable autoregressive noise," in *IEEE International Symposium on Information Theory (ISIT)*, July 21-26 2020, pp. 2091–2096.

[19] C. D. Charalambous, C. Kourtellaris, and S. Loyka, "Time-invariant feedback strategies do not increase capacity of AGN channels driven by stable and certain unstable autoregressive noise," 13 August 2021. [Online]. Available: https://arxiv.org/abs/1907.10991v3

[20] 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, July 2018.

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

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

[23] T. Kailath, A. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000.

[24] J. H. van Schuppen, *Control and System Theory of Discrete-Time Stochastic Systems*. Springer, 2021.

[25] M. I. Gelfand and M. Yaglom, "Calculation of the amount of information about a random function contained in another such function," *Amer. Math. Soc. Transl.*, vol. 2, pp. 199–246, 1959.

[26] C. D. Charalambous and S. Louka, "A Riccati-Lyapunov approach to nonfeedback capacity of MIMO Gaussian channels driven by stable and unstable noise," *arXiv preprint arXiv:2205.01430*, 2022. [Online]. Available: https://arxiv.org/abs/2205.01430