Linear Gaussian Channels: Feedback Capacity under Power Constraints

Shaohua Yang[1], Aleksandar Kavčić[1]
DEAS, Harvard University, USA
e-mail: [yangsh, kavcic]@deas.harvard.edu

Sekhar Tatikonda
Dept. of EE, Yale University, USA
e-mail: sekhar.tatikonda@yale.edu

Abstract — **For discrete-time power-constrained linear Gaussian channels with z-domain rational power spectral density functions: 1) Gauss-Markov sources achieve the feedback capacity; 2) a Kalman filter is optimal for processing the feedback information; 3) dynamic programming optimizes the Gauss-Markov source and computes the feedback capacity. Further, if the optimal Kalman filter reaches a steady state as the time $k \to \infty$, then the asymptotic feedback capacity exists and its formula is given.**

I. Summary of Results

We consider a discrete-time communication channel with input X_k, output $R_k = X_k + N_k$, and additive stationary Gaussian noise N_k whose power-spectral-density function is rational, i.e., $S_N(z) = H(z)H(z^{-1})$, where $H(z) = \left(1 - \sum_{l=1}^{L} a_l z^{-l}\right) \Big/ \left(1 + \sum_{l=1}^{L} c_l z^{-l}\right)$ is a stable minimum phase filter and is determined by coefficient vectors $\underline{a} = [a_1, a_2, \ldots, a_L]^{\mathrm{T}}$ and $\underline{c} = [c_1, c_2, \ldots, c_L]^{\mathrm{T}}$. It is assumed that there is a noiseless feedback link from the receiver to the transmitter. An equivalent state-space channel model is derived as

$$\underline{S}_k = \mathbf{A}\underline{S}_{k-1} + \underline{b}X_k, \tag{1}$$

$$Y_k = (\underline{a} + \underline{c})^{\mathrm{T}} \underline{S}_{k-1} + X_k + W_k. \tag{2}$$

Here, $\underline{S}_k = [S_k(1), \ldots, S_k(L)]^{\mathrm{T}}$ is the channel state whose initial value $\underline{S}_0 = [0, 0, \ldots, 0]^{\mathrm{T}}$ is known to both the transmitter and receiver, Y_k is the filtered channel output, i.e., $Y(z) = H^{-1}(z)R(z)$, W_k is a unit-variance white Gaussian random process, and coefficients $\mathbf{A}$ and $\underline{b}$ are

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & \ldots & a_{L-1} & a_L \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{3}$$

For any block length $n \in \mathbb{N}$, the feedback capacity [1, 2], or the maximal information rate, is

$$C^{\mathrm{fb(n)}} = \max_{\mathcal{P}} h(Y_1^n)/n - \log\left(\sqrt{2\pi \mathrm{e}}\right), \tag{4}$$

where the maximization is under the average input power constraint $\mathrm{E}\left[\frac{1}{n}\sum_{k=1}^{n}(X_k)^2\right] \le P$, and over the source

$$\mathcal{P} = \{P(X_k|X_1^{k-1}, Y_1^{k-1}, \underline{s}_0), \text{for } k = 1, 2, \ldots, n\}. \tag{5}$$

We solve this source optimization and feedback capacity computation problem (4) in Theorem 1. Then, in Theorem 2, the formula for the asymptotic feedback capacity, i.e., for $n \to \infty$, is derived under a steady-state assumption, which has been numerically verified by utilizing Theorem 1 for all example channels that we tested.

[1]This work was supported by NSF Grant CCR-9904458 and by the National Storage Industry Consortium.

Theorem 1: The feedback capacity $C^{\mathrm{fb(n)}}$ is achieved by a Gauss-Markov source, that is, for $k = 1, 2, \ldots, n$,

$$X_k = \underline{d}_k^{\mathrm{T}} \left(\underline{S}_{k-1} - \underline{m}_{k-1}\right) + e_k V_k, \tag{6}$$

where $\underline{m}_{k-1} = \mathrm{E}\left[\underline{S}_{k-1} \middle| \underline{s}_0, y_1^{k-1}\right]$, variables $V_k \sim \mathcal{N}(0,1)$ form a white random process, and coefficients $\underline{d}_k$ and e_k depend on $\mathbf{K}_{k-1} = \mathrm{E}\left[\left(\underline{S}_{k-1} - \underline{m}_{k-1}\right)\left(\underline{S}_{k-1} - \underline{m}_{k-1}\right)^{\mathrm{T}}\right]$, i.e., $\underline{d}_k = \underline{d}_k(\mathbf{K}_{k-1})$ and $e_k = e_k(\mathbf{K}_{k-1})$. Thus, a Kalman filter that computes $\underline{m}_{k-1}$ and $\mathbf{K}_{k-1}$ is optimal for processing the feedback information. Further, the capacity-achieving coefficients $\underline{d}_k$ and e_k are solution to

$$\max_{\{\underline{d}_k, e_k\}} \sum_{t=1}^{n} \left[\frac{1}{2}\log\left(1 + (\underline{a} + \underline{c} + \underline{d}_k)^{\mathrm{T}} \mathbf{K}_{k-1} (\underline{a} + \underline{c} + \underline{d}_k) + (e_k)^2\right) - \gamma\left(\underline{d}_k^{\mathrm{T}} \mathbf{K}_{k-1} \underline{d}_k + (e_k)^2\right)\right], \tag{7}$$

where $\gamma > 0$ is an auxiliary shadow-price variable which monotonically decreases with the power budget P. The optimization (7) is solved by dynamic programming. $\square$

Theorem 2: For $n \to \infty$, if the Kalman filter reaches a steady state as $k \to \infty$, then $\underline{d}_k \to \underline{d}$, $e_k \to e$, and the asymptotic feedback capacity exists and equals to

$$C^{\mathrm{fb}} \triangleq C^{\mathrm{fb}(\infty)} = \max_{\underline{d}, e} \frac{1}{2}\log\left(1 + (\underline{a} + \underline{c} + \underline{d})^{\mathrm{T}} \mathbf{K} (\underline{a} + \underline{c} + \underline{d}) + e^2\right), \tag{8}$$

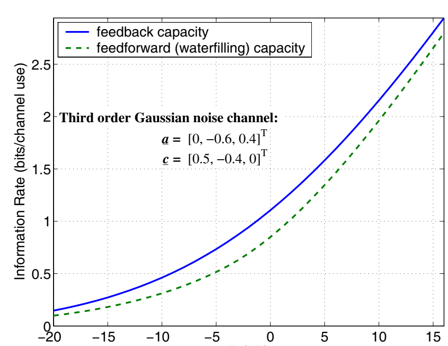
where the maximization is under the following constraints

$$\underline{d}^{\mathrm{T}} \mathbf{K} \underline{d} + e^2 = P, \tag{9}$$

$$\mathbf{K} = \mathbf{Q}\mathbf{K}\mathbf{Q}^{\mathrm{T}} + \underline{b}\,\underline{b}^{\mathrm{T}} e^2 - \frac{\left(\mathbf{Q}\mathbf{K}(\underline{a} + \underline{c} + \underline{d}) + \underline{b}\, e^2\right)\left(\mathbf{Q}\mathbf{K}(\underline{a} + \underline{c} + \underline{d}) + \underline{b}\, e^2\right)^{\mathrm{T}}}{(\underline{a} + \underline{c} + \underline{d})^{\mathrm{T}} \mathbf{K} (\underline{a} + \underline{c} + \underline{d}) + e^2 + \sigma_W^2}. \tag{10}$$

Here, $\mathbf{Q} = \mathbf{A} + \underline{b}\,\underline{d}^{\mathrm{T}}$, and $\mathbf{K}$ is non-negative definite. $\square$

With no rigorous proof of the steady state assumption in Theorem 2, formula (8) is a lower bound on the asymptotic feedback capacity. Numerical evaluations, however, reveal that the feedback capacity computed by Theorem 1 for large n equals the one computed by Theorem 2.



References

[1] Cover et al., *IEEE Trans. Inform. Theory*, p. 37, vol. 35, 1989.

[2] Tatikonda, MIT PhD Thesis, Sept. 2000.