Throughout this paper, we use the convention that $\sum_{\tau=0}^{-1} f(\tau) = 0$ and $\prod_{\tau=0}^{-1} f(\tau) = 1$ regardless of the function $f(\cdot)$, and also the convention $\log\left(\frac{0}{0}\right) = 0$.

We first consider a single sensor setting, under which the subscript $k = 1$ will be omitted, e.g., we use $\delta(t)$ and $\Sigma_{\hat{\mathbf{w}}}$ as shorthand for $\delta_1(t)$ and $\Sigma_{\hat{\mathbf{w}}_1}$.

Let $\{p_\delta : \delta \in [0, \delta_{\max}]\}$ denote the pmf of the random i.i.d. delay $\delta(\tau)$. Note that the above delay is defined from the sensor's perspective. Namely, the string $\mathbf{s}(\tau)$ will be received at time $t = \tau + \delta(\tau)$. We now consider *the current delay experienced by the controller*, which we denote it by θ. Namely, at time t, the *latest* string received by the controller is $\mathbf{s}(t-\theta)$. Given $\{p_\delta : \forall\delta\}$, we compute the pmf of θ by the following formula: Assume any arbitrary but fixed t value.

$$\overline{p}_\theta = \left(\sum_{\delta:\delta\leq\theta} p_\delta\right) \cdot \prod_{\tau=t-\theta+1}^{t} \left(\sum_{\delta:\delta>t-\tau} p_\delta\right), \ \forall\theta \in [0, \delta_{\max}] \tag{1}$$

where the first term of the right-hand side of (1) is the probability that string $\mathbf{s}(t-\theta)$ is received by time t and the rest of the right-hand side of (1) computes the probability that none of the strings $\mathbf{s}(\tau)$, $\forall\tau \in [t-\theta+1, t]$, has arrived by time θ. Even though the formula (1) involves t, the final value $\overline{p}_\theta$ does not since the delay experienced by the controller is a stationary random process.

A more compact expression of the above formula is as follows.

$$\overline{p}_\theta = p_{[0,\theta]} \cdot \prod_{i=1}^{\theta} p_{(\theta-i,\delta_{\max}]}, \ \forall\theta \in [0, \delta_{\max}], \tag{2}$$

which uses the subscript to indicate the range of interest when computing the probability and changes the index from τ to $i = \tau - t + \theta$ to remove the dependence of t.

Note that even though the sensor-side delay $\delta(\tau)$ is i.i.d., the current delay experienced by the controller usually has some memory. To dig deeper into the memory structure of θ, we also analyze the joint pmf $\overline{p}_{\theta_1,\theta_2}(m)$ of the delays experienced by the controller at two time instants, where θ_1 and θ_2 are the delays at time t and $t+m$, respectively, for some $m \geq 0$. The expression of $\overline{p}_{\theta_1,\theta_2}(m)$ contains three cases. Case 1: If $\theta_2 > \theta_1 + m$ then

$$\overline{p}_{\theta_1,\theta_2}(m) = 0 \tag{3}$$

Case 2: If $\theta_2 = \theta_1 + m$ then

$$\overline{p}_{\theta_1,\theta_2}(m) = p_{[0,\theta_1]} \cdot \prod_{\tau=t-\theta_1+1}^{t+m} p_{(t+m-\tau,\delta_{\max}]} \tag{4}$$

Case 1 is because even if there is no new arrival during time $(t, t+m]$, the delay θ_2 experienced at time $t+m$ is at most $\theta_1 + m$. Case 2 is because if $\theta_2 = \theta_1 + m$, then it means that string $\mathbf{s}(t-\theta_1)$ is received by time t and none of the strings $\mathbf{s}(\tau)$, $\forall\tau \in [t-\theta_1+1, t+m]$, has arrived by time $t+m$.

Case 3: If $\theta_2 < \theta_1 + m$ then

$$\begin{aligned}\overline{p}_{\theta_1,\theta_2}(m) =& p_{[0,\theta_1]} \cdot \left(\prod_{\tau=t-\theta_1+1}^{t+m-\theta_2-1} p_{(t-\tau,\delta_{\max}]}\right) \cdot p_{(\theta_2-m,\theta_2]} \\ &\cdot \left(\prod_{\tau=t+m-\theta_2+1}^{t+m} p_{(t+m-\tau,\delta_{\max}]}\right)\end{aligned} \tag{5}$$

Case 3 computes the probability that string $\mathbf{s}(t-\theta_1)$ is received by time t; none of the strings $\mathbf{s}(\tau)$, $\forall\tau \in [t-\theta_1+1, t+m-\theta_2)$, has arrived by time t; string $\mathbf{s}(t+m-\theta_2)$ has arrived between time $(t, t+m]$; and none of the strings $\mathbf{s}(\tau)$, $\forall\tau \in (t+m-\theta_2, t+m]$, has arrived by time $t+m$.

The marginal distribution $\overline{p}_\theta$ can be computed by (4) through the equation $\overline{p}_\theta = \overline{p}_{\theta,\theta}(0)$.

Proposition 1: For any given integers $m \in [1, \infty)$ and $\theta_1, \theta_2 \in [0, \delta_{\max}]$, define an $N \times N$ matrix $F_{\theta_1,\theta_2}(m)$ by

$$\begin{aligned}F_{\theta_1,\theta_2}(m) &\triangleq \sum_{i=\theta_2+1}^{\theta_1+m} A^i \cdot \Sigma_{\hat{\mathbf{w}}} \cdot (A^i)^{\mathrm{T}} \qquad (6)\\ &= V \cdot \mathsf{diag}\{\sigma_n^{(\theta_1,\theta_2,m)}\} \cdot V^{\mathrm{T}} \qquad (7)\end{aligned}$$

where (7) is the singular value decomposition of $F_{\theta_1,\theta_2}(m)$ with V being a unitary matrix and $\sigma_n^{(\theta_1,\theta_2,m)}$ being the n-th singular value of $F_{\theta_1,\theta_2}(m)$ for $n \in [1, N]$. Note that if $\theta_2 \geq \theta_1 + m$, then we simply have $F_{\theta_1,\theta_2}(m) = \mathbf{0}$ and $\sigma_n^{(\theta_1,\theta_2,m)} = 0$ per the convention adopted in this work.

Also define scalars $\mathsf{D}_{\min}^{(\theta)}$ and $\mathsf{D}_{\min}$ by

$$\mathsf{D}_{\min}^{(\theta)} \triangleq \mathrm{tr}\left(A^{\theta+1} P (A^{\theta+1})^{\mathrm{T}}\right) + \mathrm{tr}\left(\sum_{i=0}^{\theta} A^i \Sigma_{\mathbf{w}} (A^i)^{\mathrm{T}}\right) \tag{8}$$

$$\overline{\mathsf{D}}_{\min} \triangleq \sum_{\theta\in[0,\delta_{\max}]} \overline{p}_\theta \cdot \mathsf{D}_{\min}^{(\theta)}. \tag{9}$$

If $\mathsf{D} \leq \overline{\mathsf{D}}_{\min}$, then no scheme can achieve $\sup_t E(||\mathbf{x}(t)||^2) \leq \mathsf{D}$ regardless how large the communication rate R is. If $\mathsf{D} > \overline{\mathsf{D}}_{\min}$, then for any scheme that satisfies $\sup_t E(||\mathbf{x}(t)||^2) \leq \mathsf{D}$ with the expected bit length L, L must satisfy:

$$\begin{aligned}&L - \left(L \cdot \log_2\left(\frac{L}{L+1}\right) + \log_2\left(\frac{1}{L+1}\right)\right) \geq \\ &\frac{1}{m} \sum_{\forall\theta_1,\theta_2} \overline{p}_{\theta_1,\theta_2}(m) \cdot \left(\sum_{n=1}^{N} \frac{1}{2} \log_2\left(\frac{\sigma_n^{(\theta_1,\theta_2,m)}}{\mathsf{D}_n^{(\theta_1,\theta_2,m)}}\right)\right)\end{aligned} \tag{10}$$

where the value of $\mathsf{D}_n^{(\theta_1,\theta_2,m)}$ is the *water-filling* coefficients computed by

$$\mathsf{D}_n^{(\theta_1,\theta_2,m)} = \min(\eta^{(m)}, \sigma_n^{(\theta_1,\theta_2,m)}), \ \forall n, \theta_1, \theta_2 \tag{11}$$

and $\eta^{(m)}$ is chosen to be the largest possible value that still satisfies

$$\sum_{\forall\theta_1,\theta_2} \overline{p}_{\theta_1,\theta_2}(m) \cdot \left(\sum_{n=1}^{N} \mathsf{D}_n^{(\theta_1,\theta_2,m)}\right) \leq \mathsf{D} - \overline{\mathsf{D}}_{\min}. \tag{12}$$

(Recall that we set $\log\left(\frac{0}{0}\right) = 0$.)

I. Proof of Proposition 1

Proof: The proof consists of three parts and the value m is treated as a constant throughout the entire proof. Part I is a simple extension of the classic rate-distortion result.

Part I: Consider a time sharing random variable Q and an arbitrary deterministic 2-dimensional vector function $\vec{\Theta}(q) = (\Theta_1(q), \Theta_2(q))$. Also consider $(\delta_{\max}+1)^2$ random vectors Z_{θ_1,θ_2}, $\forall \theta_1, \theta_2 \in [0, \delta_{\max}]$. We assume that the joint distribution of $(Q, Z_{\vec{\Theta}(Q)})$ and the deterministic function $\vec{\Theta}(\cdot)$ satisfy the following two conditions: Assumption (i): The pmf of the distribution the 2-dimensional vector $\vec{\Theta}(Q) = (\Theta_1(Q), \Theta_2(Q))$ equals the pmf $\{\overline{p}_{\theta_1,\theta_2}(m) : \theta_1, \theta_2 \in [0, \delta_{\max}]\}$ defined in (3) to (5). Assumption (ii): Given $Q = q$, the conditional distribution of the vector $Z_{\vec{\Theta}(q)}$ is zero-mean Gaussian with covariance $F_{\Theta_1(q),\Theta_2(q)}(m)$ defined in (6). For notational simplicity, we often write Z_Q as shorthand for $Z_{\vec{\Theta}(Q)}$.

Our goal is to solve the following optimal quantizer problem:

$$\min \ I\left(Z_Q; \tilde{Z}_Q \middle| Q \right) \tag{13}$$

$$\text{subject to } E\left(\left\| Z_Q - \tilde{Z}_Q \right\|^2 \right) \leq \mathsf{D} - \overline{\mathsf{D}}_{\min} \tag{14}$$

where $\tilde{Z}_Q$ is the to-be-optimized quantizer, $I(\cdot;\cdot|\cdot)$ is the conditional mutual information, D is an arbitrarily given constant, and $\overline{\mathsf{D}}_{\min}$ is defined in (9). The above problem is a simple extension of the classic rate-distortion results and the minimum in (13) is thus

$$\sum_{\theta_1,\theta_2 \in [1,\delta_{\max}]} \overline{p}_{\theta_1,\theta_2}(m) \cdot \left(\sum_{n=1}^{N} \frac{1}{2} \log_2 \left(\frac{\sigma_n^{(\theta_1,\theta_2,m)}}{\mathsf{D}_n^{(\theta_1,\theta_2,m)}} \right) \right) \tag{15}$$

where $\mathsf{D}_n^{(\theta_1,\theta_2,m)}$ is the *water-filling* coefficient computed by (11). The proof of (15) is done by first fixing the values of $E\left(\left\| Z_Q - \tilde{Z}_Q \right\|^2 \middle| Q = q \right)$ for each $Q = q$ and minimizing $I(Z_Q; \tilde{Z}_Q | Q = q)$ separately for each $Q = q$ through the classic rate-distortion formula. We then further minimizing $I(Z_Q; \tilde{Z}_Q|Q)$ by reallocating the values of $E\left(\left\| Z_Q - \tilde{Z}_Q \right\|^2 \middle| Q = q \right)$ in an optimal way.

Part II describes a reduction-based construction that uses any arbitrarily given VLQ scheme of a linear control system to construct a quantizer $\tilde{Z}_Q$ for the composite Gaussian vector Z_Q described in Part I.

Part III shows that if the given VLQ scheme does not satisfy the inequality in (10), then the corresponding quantizer $\tilde{Z}_Q$ will have $I(Z_Q; \tilde{Z}_Q|Q)$ strictly smaller than (15), which contradicts the results in Part I. The proof is thus complete. In the following, we describe Parts II and III separately.

Part II: Consider a sufficiently large integer parameter T such that we can assume that the Kalman filter converges for all $t \geq T - \delta_{\max}$. That is, $P_t = P$, $\Phi_t = \Phi$, $\Gamma_t = \Gamma$ and $\Sigma_{\hat{\mathbf{w}}_{t-1}} = \Sigma_{\hat{\mathbf{w}}}$ for all $t \geq T - \delta_{\max}$. This assumption can be made rigorous by noting that all the statements in Proposition 1 are continuous with respect to the underlying matrices.

We first introduce the following *uncontrolled system*, which will be used extensively in the proof.

$$\mathbf{x}_{\mathrm{u}}(t+1) = A\mathbf{x}_{\mathrm{u}}(t) + \mathbf{w}(t) \tag{16}$$

$$\mathbf{y}_{\mathrm{u}}(t) = C\mathbf{x}_{\mathrm{u}}(t) + \mathbf{v}(t). \tag{17}$$

That is, the original discrete-time Gaussian linear system and this new uncontrolled system are driven by the same noises $\mathbf{w}(t)$ and $\mathbf{v}(t)$. The only difference between (16) and the original discrete-time Gaussian linear system is that the control action is set to $\mathbf{u}(t) = \mathbf{0}$ for all t in (16). One can easily verify that for any t, the *uncontrolled system* and the original linear system is related by the following equations.

$$\mathbf{x}(t) = \mathbf{x}_{\mathrm{u}}(t) + \sum_{\tau=0}^{t-1} A^{t-1-\tau}\mathbf{u}(\tau) \tag{18}$$

$$\mathbf{y}(t) = \mathbf{y}_{\mathrm{u}}(t) + C\left(\sum_{\tau=0}^{t-1} A^{t-1-\tau}\mathbf{u}(\tau) \right). \tag{19}$$

It is worth pointing out that the formulae of the P_t, Φ_t, Γ_t, $\Sigma_{\hat{\mathbf{w}}_{t-1}}$ matrices do not involve the control action $\mathbf{u}(\cdot)$. Therefore, both the original and the new uncontrolled systems share the same P_t, Φ_t, Γ_t, $\Sigma_{\hat{\mathbf{w}}_{t-1}}$ matrices. We also note that the delay $\{\delta(t) : t \in [0,\infty)\}$ only affects the delivery of the strings $\mathbf{s}(t)$ and thus the control actions $\mathbf{u}(t)$. Since the uncontrolled system has $\mathbf{u}(t) = 0$, the evolution of the uncontrolled system is independent of the delay $\{\delta(t) : t \in [0,\infty)\}$.

We now define the time sharing random variable Q. Consider the original discrete-time Gaussian linear control system. We first generate the random disturbance $\mathbf{w}(t)$ and random observation noise $\mathbf{v}(t)$ according to $\mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{w}})$ and $\mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{v}})$, respectively. We define the *past observation* of the uncontrolled system until time t by $\mathcal{Y}_{\mathrm{u}}(t) \triangleq \{\mathbf{y}_{\mathrm{u}}(\tau) : \tau \in [0,t]\}$. Namely, $\mathcal{Y}_{\mathrm{u}}(t)$ contains all the observation $\mathbf{y}_{\mathrm{u}}(\cdot)$ of the uncontrolled system until time t.

We also generate independently the sensor-side random delay $\delta(t)$ for all $t \in [0,\infty)$ according to the sensor-side distribution $\{p_\delta\}$. We use $\vec{\delta} \triangleq \{\delta(t) : t \in [0,\infty)\}$ to represent the entire random process of the sensor side delay from time 0 to ∞. We also define two random variables Θ_1 and Θ_2, where the former being the delay experienced by the controller at time T and the latter being the delay experienced by the controller at time $T+m$.

Finally, we define the time-sharing random variable Q as a pair

$$Q \triangleq \left\{ \vec{\delta}, \mathcal{Y}_{\mathrm{u}}(T - \Theta_1) \right\}. \tag{20}$$

That is, Q contains the entire sensor-side delay random process $\vec{\delta}$, and the overall observation of the uncontrolled system until time $T - \Theta_1$, where Θ_1 is the random delay experienced by the controller at time T.

We now define the deterministic 2-dimensional function $\vec{\Theta}(Q) \triangleq (\Theta_1, \Theta_2)$. We observe that the delays experienced by the controller depend exclusively on the realization of sensor-side delays $\vec{\delta}$, not on the underlying VLQ scheme nor on the values of $\mathbf{w}(t)$ and $\mathbf{v}(t)$. Since Q contains $\vec{\delta}$, the values of Θ_1 and Θ_2 are uniquely determined by Q and the function $\vec{\Theta}(Q)$ is indeed deterministic.

After defining the time-sharing random variable Q, we define the random variable $Z_{\vec{\Theta}(Q)}$. Again we use Z_Q as shorthand. To construct Z_Q, we assume there is no delay and no communication-rate constraint in the uncontrolled system in (16) and (17) so that a Kalman filter at time t can directly access $\mathbf{y}_{\mathrm{u}}(t)$ for all t. Under this *no-delay-in-the-uncontrolled-system* setting, we perform Kalman filtering on the uncontrolled system until time $T+m$ and denote the resulting series of estimates by $\hat{\mathbf{x}}_{\mathrm{u}}(t)$ for $t \in [0, T+m]$. Then for any given θ_1 and θ_2 values we compute two more estimates from this series of KF estimates $\{\hat{\mathbf{x}}_{\mathrm{u}}(t)\}$. That is

$$\hat{\mathbf{x}}_{\mathrm{u}}^{(\theta_1)}(T) = A^{\theta_1}\hat{\mathbf{x}}_{\mathrm{u}}(T-\theta_1) \tag{21}$$

$$\hat{\mathbf{x}}_{\mathrm{u}}^{(\theta_2)}(T+m) = A^{\theta_2}\hat{\mathbf{x}}_{\mathrm{u}}(T+m-\theta_2). \tag{22}$$

Intuitively speaking, $\hat{\mathbf{x}}_{\mathrm{u}}^{(\theta_1)}(T)$ is the KF estimate at time T if the delay experienced by the controller is θ_1 and all the observations before time $T-\theta_1$ are available. Similarly, $\hat{\mathbf{x}}_{\mathrm{u}}^{(\theta_2)}(T+m)$ is the KF estimate at time $T+m$ if the delay experienced by the controller is θ_2 and all the observations before time $T+m-\theta_2$ are available. Note that T and m are treated as constants throughout the proof.

We then define

$$Z_{\theta_1,\theta_2} \triangleq A \cdot \hat{\mathbf{x}}_{\mathrm{u}}^{(\theta_2)}(T+m) - A^{m+1} \cdot \hat{\mathbf{x}}_{\mathrm{u}}^{(\theta_1)}(T) \tag{23}$$

$$= A^{\theta_2+1} \cdot \left(\sum_{\tau=T-\theta_1}^{T+m-\theta_2-1} A^{T+m-\theta_2-1-\tau}\hat{\mathbf{w}}_{\mathrm{u}}(\tau) \right) \tag{24}$$

where (24) is derived by (21) and (22) and by iteratively using the Kalman filtering formula that computes the estimate $\hat{\mathbf{x}}_{\mathrm{u}}(t)$ for the uncontrolled system. The random variable Z_Q of our interest can then be defined as $Z_{\vec{\Theta}(Q)}$, i.e., we plug in the random variables Θ_1 and Θ_2 in (24).

We now verify that the above construction of Q, $\vec{\Theta}(\cdot)$, and Z_Q satisfies the two assumptions in Part I. Per our construction $\vec{\Theta}(Q)$ will have pmf being $\{\overline{p}_{\theta_1,\theta_2}(m) : \theta_1, \theta_2 \in [0, \delta_{\max}]\}$. Assumption (i) thus holds. By (24), one can see that Z_{θ_1,θ_2} depends only on $\hat{\mathbf{w}}_{\mathrm{u}}(\tau)$ for $\tau \in [T-\theta_1, T+m-\theta_2)$, which is independent of $\mathcal{Y}_{\mathrm{u}}(T-\theta_1)$, the past observations before and including time $T-\theta_1$. Therefore, conditioning on $Q = q = (\vec{\delta}, \mathcal{Y}_{\mathrm{u}}(T-\theta_1))$ with $\vec{\Theta}(\vec{\delta}) = (\theta_1, \theta_2)$, the distribution of $\{\hat{\mathbf{w}}_{\mathrm{u}}(\tau) : \tau \in [T-\theta_1, T+m-\theta_2)\}$ is i.i.d. Gaussian with covariance $\Sigma_{\hat{\mathbf{w}}}$. Then by (24), the conditional distribution of Z_{θ_1,θ_2} given $Q=q$ is Gaussian with mean zero and covariance $F_{\theta_1,\theta_2}(m)$. Assumption (ii) thus holds.

Remark 1: It is worth emphasizing that the above definition of Q, $\vec{\Theta}(\cdot)$, and Z_Q does not involve any VLQ scheme since they are based only on the random noises, random delays, and on performing KF on the uncontrolled system with zero-delay.

Remark 2: Thus far, we describe how to start from independently distributed $\mathbf{w}(t)$, $\mathbf{v}(t)$ and $\delta(t)$ to generate the joint distribution of (Q, Z_Q). Such a construction can be viewed as describing the joint distribution $P_{\{\mathbf{w}(\tau),\mathbf{v}(\tau),\delta(\tau):\tau\in[0,\infty)\},Q,Z_Q}$, which can later be used to derive the marginal distribution P_{Q,Z_Q} and the conditional distribution $P_{\{\mathbf{w}(\tau),\mathbf{v}(\tau),\delta(\tau):\tau\in[0,\infty)\}|Q,Z_Q}$. If we are given any random variable pair (Q, Z_Q) that has the same distribution as the derived marginal distribution P_{Q,Z_Q}, the pair (Q, Z_Q) can then be used to retrospectively generate the $\mathbf{w}(t)$, $\mathbf{v}(t)$ and $\delta(t)$ random variables. As a result, from the perspective quantizing Z_Q, it makes no difference whether the quantity Z_Q is given by an external source as a black box or whether it is generated through an elaborate procedure as described herein, provided the distribution of (Q, Z_Q) satisfies both Assumptions (i) and (ii).

In the following, we describe how any given VLQ scheme can be used to construct a quantizer $\tilde{Z}_Q$. Specifically, for any specific VLQ scheme, we can construct a quantizer of Z_Q by first running/simulating the original discrete-time Gaussian linear system using the same $\mathbf{w}(t)$ and $\mathbf{v}(t)$ as those used in the uncontrolled system, and using the independently generated delay $\vec{\delta}$. Note that while the random delay $\vec{\delta}$ has zero impact on the uncontrolled system, it will significantly alter the behavior of the original Gaussian system since the string $\mathbf{s}(t)$ now experience some random delay before arriving at the controller. We then use the given VLQ scheme to generate the control actions $\mathbf{u}(t)$ for all t. Finally we set

$$\tilde{Z}_Q \triangleq -\left(\sum_{\tau=0}^{T+m} A^{T+m-\tau}\mathbf{u}(\tau) \right) - A^{m+1}\hat{\mathbf{x}}_{\mathrm{u}}^{(\Theta_1)}(T). \tag{25}$$

Note that the above quantizer is coupled with the time-sharing random variable Q. Given different $Q = q$, the conditional distribution of $\tilde{Z}_Q$ may differ. The above concludes the description how to use any given VLC scheme to generate a quantizer $\tilde{Z}_Q$ for the random variable Z_Q.

Part III: Suppose the given VLQ scheme has expected length L and achieves expected distortion D for all time t. Then by choosing $t = T+m+1$, we have

$$E\left(\|\mathbf{x}(T+m+1)\|^2\right) \leq \mathsf{D} \tag{26}$$

We now characterize the performance of the quantizer $\tilde{Z}_Q$ constructed in (25). By (18), we have

$$\mathbf{x}(T+m+1) = \mathbf{x}_{\mathrm{u}}(T+m+1) + \sum_{\tau=0}^{T+m} A^{T+m-\tau}\mathbf{u}(\tau). \tag{27}$$

Since $\mathbf{x}_{\mathrm{u}}(T+m+1) = A\mathbf{x}_{\mathrm{u}}(T+m) + \mathbf{w}(T+m)$, we have

$$\begin{aligned}\mathbf{x}(T+m+1) =& A\left(\mathbf{x}_{\mathrm{u}}(T+m) - \hat{\mathbf{x}}_{\mathrm{u}}^{(\Theta_2)}(T+m)\right) \\ &+ A\hat{\mathbf{x}}_{\mathrm{u}}^{(\Theta_2)}(T+m) \\ &+ \sum_{\tau=0}^{T+m} A^{T+m-\tau}\mathbf{u}(\tau) + \mathbf{w}(T+m).\end{aligned} \tag{28}$$

By (23) and (25), we thus have

$$\begin{aligned}\mathbf{x}(T+m+1) =& A\left(\mathbf{x}_{\mathrm{u}}(T+m) - \hat{\mathbf{x}}_{\mathrm{u}}^{(\Theta_2)}(T+m)\right) \\ &+ (Z_Q - \tilde{Z}_Q) \\ &+ \mathbf{w}(T+m).\end{aligned} \tag{29}$$

There are three terms in (29), and we analyze the contribution of each term to the mean-square norm in (26).

We first notice that $\mathbf{w}(T+m)$ is zero mean and is independent of the other two terms. We thus have

$$E\left(\|\mathbf{x}(T+m+1)\|^2\right)$$
$$= E\left(\left\|A\left(\mathbf{x}_{\rm u}(T+m)-\hat{\mathbf{x}}_{\rm u}^{(\Theta_2)}(T+m)\right)+(Z_Q-\tilde{Z}_Q)\right\|^2\right)$$
$$+\,\mathrm{tr}(\Sigma_{\mathbf{w}}). \quad (30)$$

We now argue that both Z_Q and $\tilde{Z}_Q$ are deterministic functions of $(\vec{\delta},\mathcal{Y}_{\rm u}(T{+}m{-}\Theta_2))$, the entire delay realization and the past observation of the uncontrolled system until time $T{+}m{-}\Theta_2$.

Firstly, Z_Q in (23) is a deterministic function of the KF estimates $\hat{\mathbf{x}}_{\rm u}(T-\Theta_1)$ and $\hat{\mathbf{x}}_{\rm u}(T+m-\Theta_2)$, which in turn are deterministic functions of $(\vec{\delta},\mathcal{Y}_{\rm u}(T+m-\Theta_2))$. By applying the same argument, the $\hat{\mathbf{x}}_{\rm u}^{(\Theta_1)}(T)$ term in the definition of $\tilde{Z}_Q$ in (25) is also a deterministic function of $(\vec{\delta},\mathcal{Y}_{\rm u}(T{+}m{-}\Theta_2))$.

What remains to show is that $\{\mathbf{u}(\tau):\tau\in[0,T+m]\}$ is a deterministic function of $(\vec{\delta},\mathcal{Y}_{\rm u}(T{+}m{-}\Theta_2))$. The concept of the proof is quite straightforward, i.e., knowing $\mathbf{y}_{\rm u}(t)$ of the uncontrolled system allows one to know $\mathbf{y}(t)$ of the original system, which in turn allows one to know the transmitted string $\mathbf{s}(t)$. This then allows one to learn the computed control action $\mathbf{u}(t)$. In the following, we provide a rigorous proof.

The proof is by induction. Define $\mathcal{Y}(t)=\{\mathbf{y}(\tau):\tau\in[0,t]\}$ as all the past observations of the original system until time t. Note that while the observation $\mathcal{Y}_{\rm u}(t)$ of the uncontrolled system is not affected by the delay $\vec{\delta}$, the observation $\mathcal{Y}(t)$ depends heavily on the random delay $\vec{\delta}$ and the VLC scheme being used. The induction condition of the proof is: For any time t,

$$(\vec{\delta},\mathcal{Y}(t)) \text{ is a deterministic function of } (\vec{\delta},\mathcal{Y}_{\rm u}(t)). \quad (31)$$

Since $\{\mathbf{u}(\tau):\tau\in[0,T+m]\}$ is a deterministic function of $\{\mathbf{s}(\tau):\tau\in[0,T+m-\Theta_2]\}$ and since $\{\mathbf{s}(\tau):\tau\in[0,T+m-\Theta_2]\}$ is a deterministic function of $(\vec{\delta},\mathcal{Y}(T+m-\Theta_2))$, the above induction condition immediately implies $\{\mathbf{u}(\tau):\tau\in[0,T+m]\}$ being a deterministic function of $(\vec{\delta},\mathcal{Y}_{\rm u}(T+m-\Theta_2))$.

The induction condition holds naturally when $t=0$ since $\mathbf{y}(0)=\mathbf{y}_{\rm u}(0)$ according to (19). Suppose the induction condition holds for $t\leq t_0$. Then for any $t\leq t_0$ the string $\mathbf{s}(t)$, which, by definition, is a function of $\mathcal{Y}(t)$, is also a deterministic function of $(\vec{\delta},\mathcal{Y}_{\rm u}(t_0))$. Since the control action $\mathbf{u}(t)$, by definition, is a function of $\vec{\delta}$ and $\{\mathbf{s}(\tau):\tau\in[0,t-\theta(t)]\}$, it is also a deterministic function of $(\vec{\delta},\mathcal{Y}_{\rm u}(t_0))$. Since $\{\mathbf{u}(\tau):\tau\in[0,t_0]\}$ is a deterministic function of $(\vec{\delta},\mathcal{Y}_{\rm u}(t_0))$, eq. (19) implies $\mathbf{y}(t_0+1)$ is a deterministic function of $(\vec{\delta},\mathcal{Y}_{\rm u}(t_0{+}1))$. The induction condition thus holds for $t=t_0{+}1$ as well. The mini proof of (31) is thus complete.

Continue our analysis of (30). We notice that the KF estimate $\hat{\mathbf{x}}_{\rm u}^{(\Theta_2)}(T{+}m)$ is the conditional mean of $\mathbf{x}_{\rm u}(T{+}m)$ given $(\vec{\delta},\mathcal{Y}_{\rm u}(T{+}m{-}\Theta_2))$. Combining with the fact that both Z_Q and $\tilde{Z}_Q$ are deterministic functions of $(\vec{\delta},\mathcal{Y}_{\rm u}(T{+}m{-}\Theta_2))$, we have that $\mathbf{x}_{\rm u}(T+m)-\hat{\mathbf{x}}_{\rm u}^{(\Theta_2)}(T+m)$ has zero mean and is independent of $(Z_Q-\tilde{Z}_Q)$ when conditioning on $(\vec{\delta},\mathcal{Y}_{\rm u}(T+m-\Theta_2))$. We thus have

$$E\left(\|\mathbf{x}(T+m+1)\|^2\right)$$
$$= E\left(\left\|A\left(\mathbf{x}_{\rm u}(T+m)-\hat{\mathbf{x}}_{\rm u}^{(\Theta_2)}(T+m)\right)\right\|^2\right)$$
$$+\,E\left(\left\|Z_Q-\tilde{Z}_Q\right\|^2\right)+\mathrm{tr}(\Sigma_{\mathbf{w}}). \quad (32)$$

By analyzing (22) and assuming a fixed constant θ_2, the covariance matrix of the vector $\left(\mathbf{x}_{\rm u}(T+m)-\hat{\mathbf{x}}_{\rm u}^{(\theta_2)}(T+m)\right)$ is

$$\text{covariance}=A^{\theta_2}P\left(A^{\theta_2}\right)^{\rm T}+\sum_{i=0}^{\theta_2-1}A^i\Sigma_{\mathbf{w}}\left(A^i\right)^{\rm T}$$

where a sufficiently large T is used to ensure that $P_t=P$ for all $t\geq T-\delta_{\max}$. As a result,

$$E\left(\left\|A\left(\mathbf{x}_{\rm u}(T+m)-\hat{\mathbf{x}}_{\rm u}^{(\Theta_2)}(T+m)\right)\right\|^2\right)+\mathrm{tr}(\Sigma_{\mathbf{w}})$$
$$= E\left(\mathrm{tr}\left(A^{\Theta_2+1}P\left(A^{\Theta_2+1}\right)^{\rm T}\right)\right)$$
$$+\,E\left(\sum_{i=0}^{\Theta_2}\mathrm{tr}\left(A^i\Sigma_{\mathbf{w}}\left(A^i\right)^{\rm T}\right)\right) \quad (33)$$

where the expectation is with respect to the Θ_2 random variable. Eqs. (33), (32), and (26) jointly imply that

$$E\left(\left\|Z_Q-\tilde{Z}_Q\right\|^2\right)\leq \mathsf{D}-\overline{\mathsf{D}}_{\min} \quad (34)$$

for the $\overline{\mathsf{D}}_{\min}$ defined in (9).

We now analyze the mutual information $I(Z_Q;\tilde{Z}_Q|Q)$. We have

$$I(Z_Q;\tilde{Z}_Q|Q)\leq H(\tilde{Z}_Q|Q) \quad (35)$$

where $H(\cdot)$ is the entropy function. It is clear by definition (21) that $\hat{\mathbf{x}}_{\rm u}^{(\Theta_1)}(T)$ is a function of the past observation $\mathcal{Y}_{\rm u}(T{-}\Theta_1)$. By the induction condition (31) and the corresponding analysis, $\mathbf{u}(\tau)$ is a deterministic function of $Q=(\vec{\delta},\mathcal{Y}_{\rm u}(T-\Theta_1))$ for all $\tau\in[0,T]$. As a result, those $\mathbf{u}(\tau)$ do not contribute to the conditional entropy given Q. By (25), this implies

$$H(\tilde{Z}_Q|Q)\leq H(\{\mathbf{u}(\tau):\tau\in[T+1,T+m]\}|Q). \quad (36)$$

Conditioning on $Q=(\vec{\delta},\mathcal{Y}_{\rm u}(T{-}\Theta_1))$, the conditional entropy of the actions $\{\mathbf{u}(\tau):\tau\in[T+1,T+m]\}$ can only be resulted from the strings $\mathbf{s}(t)$ that have arrived during time $[T+1,T+m]$. The reason is that by definition, $\{\mathbf{u}(\tau):\tau\in[T+1,T+m]\}$ is a deterministic function of $\{\mathbf{s}(\tau):\tau\in[0,T+m-\Theta_2]\}$. By (31) and the accompanying analysis, $\{\mathbf{s}(\tau):\tau\in[0,T-\Theta_1]\}$ is a deterministic function of $Q=(\vec{\delta},\mathcal{Y}_{\rm u}(T-\Theta_1))$. If we define

$$\overline{\mathbf{S}}_{[t_1,t_2]}=\{\mathbf{s}(\tau):\tau+\delta(\tau)\in[t_1,t_2]\} \quad (37)$$

as the set of strings arriving between $[t_1,t_2]$. then this argument proves that

$$H(\{\mathbf{u}(\tau):\tau\in[T+1,T+m]\}|Q)\leq H(\overline{\mathbf{S}}_{[T+1,T+m]}|Q). \quad (38)$$

There are three major challenges when analyzing $H(\overline{\mathbf{S}}_{[T+1,T+m]}|Q)$. Challenge 1: $\overline{\mathbf{S}}_{[T+1,T+m]}$ contains a random number of messages, ranging from 0 (no message arriving during $[T+1,T+m]$) to $m+\delta_{\max}$ (all messages originated during time $[T+1-\delta_{\max},T+m]$ all arriving during $[T+1,T+m]$). So the analysis has to take into account a random number of incoming strings.

Challenge 2: The VLQ scheme used by the sensor to encode $\mathbf{s}(t)$ may be time-dependent. Therefore, no stationarity can be assumed during the analysis, which further complicates the analysis with a random number of incoming strings.

Challenge 3: Perhaps the most challenging obstacle is that for any time t, delays $\{\delta(\tau) : \tau+\delta(\tau) \leq t\}$ is known at the controller and may be fed from the controller back to the sensor. Such feedback can happen by embedding the information in the control action $\mathbf{u}(t)$. As a result, the entropy of $\mathbf{s}(t+1)$ may also depend on the delay realization $\vec{\delta}$. For example, it is possible that the sensor sends a string $\mathbf{s}(t)$ with expected length $> L$ when it notices that the delay experienced by the controller at time t is large and sends a shorter string with expected length $< L$ when the delay experienced by the controller at time t is small. Such a scheme can still have the overall length $E(|\mathbf{s}(t)|) = L$ since the two scenarios will be averaged probabilistically. On the other hand, the controller can use the longer string $\mathbf{s}(t)$ to better control the state $\mathbf{x}(t+1)$ in a relatively more adverse situation (facing longer delay), and use the shorter string $\mathbf{s}(t)$ to generate a more relaxed control $\mathbf{u}(t)$ when in a more lenient situation (facing shorter delay).

Recall that in Challenge 1, the number of strings in $\overline{\mathbf{S}}_{[T+1,T+m]}$ depends on the $\vec{\delta}$. Challenges 1 and 3 together show that both the number of strings and the content of each string may be simultaneously affected by Q, which significantly complicates the analysis.

Recall that thus far the only requirement for the starting point T is that the KF converges after $t \geq T-\delta_{\max}$. We now argue that for any $\epsilon > 0$ we can always choose a starting point T for which KF converges *and* the following inequality holds.

$$H(\overline{\mathbf{S}}_{[T+1,T+m]}|Q) \leq m\left(\epsilon + \sup_t H(\mathbf{s}(t))\right) \quad (39)$$

Consider a series of integer values $T_k \triangleq (k-1)m-1$, $k \in [1,\infty)$. For each T_k, define Θ_{T_k} as the delay experienced by the controller at time T_k. For any $K > 0$, we have

$$\begin{aligned}
&H(\{\mathbf{s}(\tau) : \tau \in [0, T_K+m]\}\,|\vec{\delta}) \\
&\geq H(\overline{\mathbf{S}}_{[0,T_k+m]}|\vec{\delta}) && (40)\\
&= \sum_{k=1}^{K} H\left(\overline{\mathbf{S}}_{[T_k+1,T_k+m]}\,\middle|\,\vec{\delta}, \overline{\mathbf{S}}_{[0,T_k]}\right) && (41)\\
&\geq \sum_{k=1}^{K} H\left(\overline{\mathbf{S}}_{[T_k+1,T_k+m]}\,\middle|\,\vec{\delta}, \mathcal{Y}_{\mathrm{u}}(T_k-\Theta_{T_k}), \overline{\mathbf{S}}_{[0,T_k]}\right) && (42)\\
&= \sum_{k=1}^{K} H\left(\overline{\mathbf{S}}_{[T_k+1,T_k+m]}\,\middle|\,\vec{\delta}, \mathcal{Y}_{\mathrm{u}}(T_k-\Theta_{T_k})\right) && (43)
\end{aligned}$$

where (40) follows from that the (deterministic) set $\{\mathbf{s}(\tau) : \tau \in [0, T_K+m]\}$, containing the strings sent during time $[0, T_k+m]$, is always a superset of the random set $\overline{\mathbf{S}}_{[0,T_k+m]}$, the strings arrived during the same time. (41) follows from the chain rule. (42) follows from that conditioning reduces entropy. (43) follows from that $\overline{\mathbf{S}}_{[0,T_k]}$, the *received* strings during time $[0,T_k]$, is completely determined by the entire delay random process $\vec{\delta}$ and the past observation $\mathcal{Y}(T_k-\Theta_{T_k})$ of the original system since Θ_{T_k} is the delay experienced by the controller at time T_k. This, together with (31), proves that we can drop the conditioning of $\overline{\mathbf{S}}_{[0,T_k]}$ without affecting the entropy, which leads to (43).

We also have

$$\begin{aligned}
&H(\{\mathbf{s}(\tau) : \tau \in [0, T_K+m]\}\,|\vec{\delta}) \\
&\leq H(\{\mathbf{s}(\tau) : \tau \in [0, T_K+m]\}) && (44)\\
&\leq \sum_{t=0}^{T_K+m} H(\mathbf{s}(t)) && (45)\\
&\leq K\cdot m\cdot\left(\sup_t H(\mathbf{s}(t))\right) && (46)
\end{aligned}$$

where (44) follows from removing a conditioning event, (45) follows from that the joint entropy is no larger than the sum of the marginal entropies, and (46) follows from replacing the sum by the product of the number of the terms and the maximum of the terms.

Jointly (43) and (46) show that at least one $k_0 \in [1,K]$ satisfies

$$H\left(\overline{\mathbf{S}}_{[T_{k_0}+1,T_{k_0}+m]}\,\middle|\,\vec{\delta}, \mathcal{Y}_{\mathrm{u}}(T_{k_0}-\Theta_{T_{k_0}})\right) \leq m\left(\sup_t H(\mathbf{s}(t))\right) \quad (47)$$

Note that if we set $T = T_{k_0}$ then this choice of T satisfies (39) with $\epsilon = 0$. However, the KF convergence condition may not hold for such a choice since T_{k_0} may be too small. However, since (43) and (46) hold for arbitrary K, we can show by basic limiting arguments that for any given T' and any given $\epsilon > 0$, we can always find a k_0 such that $T_{k_0} > T'$ and

$$\begin{aligned}
&H\left(\overline{\mathbf{S}}_{[T_{k_0}+1,T_{k_0}+m]}\,\middle|\,\vec{\delta}, \mathcal{Y}_{\mathrm{u}}(T_{k_0}-\Theta_{T_{k_0}})\right) \\
&\leq m\left(\epsilon + \sup_t H(\mathbf{s}(t))\right) && (48)
\end{aligned}$$

As a result, we can always choose a $T = T_{k_0}$ for some large k_0 such that KF converges after $T-\delta_{\max}$ and (39) holds.

Since the expected length of $\mathbf{s}(t) \leq L$, by Gibb's distribution, the entropy of $H(\mathbf{s}(t))$ is upper bounded by

$$\begin{aligned}
&\sup_t H(\mathbf{s}(t)) \\
&\leq L - \left(L\cdot\log_2\left(\frac{L}{L+1}\right) + \log_2\left(\frac{1}{L+1}\right)\right) && (49)
\end{aligned}$$

Combining (35), (36), (38), (39), and (49), we have

$$\begin{aligned}
&I(Z_Q; \tilde{Z}_Q|Q) \\
&\leq m\cdot\left(\epsilon + L - \left(L\log_2\left(\frac{L}{L+1}\right) + \log_2\left(\frac{1}{L+1}\right)\right)\right) && (50)
\end{aligned}$$

In Part I, we have already proven that $I(Z_Q; \tilde{Z}_Q|Q)$ is no less than (15). By letting ϵ in (50) being arbitrarily small, we have thus proven (10).

■

II. Proof of Proposition 2

Correction: In Proposition 2 we also need the assumption that **"$\Sigma_{\hat{\mathbf{w}}}$ is invertible"**, which holds naturally when the Gaussian perturbation $\mathbf{w}(t)$ is not degenerate, i.e., $\Sigma_{\mathbf{w}}$ being invertible.

Note: We assume that the system evolution matrix A is of real value. For complex-valued A, the proof still holds after changing the transpose of A to the conjugate transpose of A.

For any given $\overline{p}_{\theta_1,\theta_2}(m)$, m, $\Sigma_{\hat{\mathbf{w}}}$, and fixed difference $\Delta_{\mathsf{D}} \triangleq \mathsf{D} - \overline{\mathsf{D}}_{\min}$. Define

$$\begin{aligned}&\Phi(\overline{p}_{\theta_1,\theta_2}(m), m, \Sigma_{\hat{\mathbf{w}}}, \Delta_{\mathsf{D}})\\ &= \sum_{\forall\theta_1,\theta_2} \overline{p}_{\theta_1,\theta_2}(m) \cdot \left(\sum_{n=1}^{N} \frac{1}{2}\log_2\left(\frac{\sigma_n^{(\theta_1,\theta_2,m)}}{\mathsf{D}_n^{(\theta_1,\theta_2,m)}}\right)\right) \end{aligned} \tag{51}$$

where $\sigma_n^{(\theta_1,\theta_2,m)}$ is derived from $\Sigma_{\hat{\mathbf{w}}}$ and m using (6) and (7); and $\mathsf{D}_n^{(\theta_1,\theta_2,m)}$ is derived from the modified water-filling operations in (11) and (12) that take $\overline{p}_{\theta_1,\theta_2}(m)$, $\sigma_n^{(\theta_1,\theta_2,m)}$, and Δ_{D} as input.

Since we consider sufficiently large m, we can safely assume $m > \delta_{\max}$. In order to prove Proposition 2, we first claim that

$$\begin{aligned}&\Phi(1_{\{\theta_1=0,\theta_2=\delta_{\max}\}}, m, \Sigma_{\hat{\mathbf{w}}}, \Delta_{\mathsf{D}})\\ &\leq \Phi(\overline{p}_{\theta_1,\theta_2}(m), m, \Sigma_{\hat{\mathbf{w}}}, \Delta_{\mathsf{D}})\end{aligned} \tag{52}$$

where $1_{\{\theta_1=0,\theta_2=\delta_{\max}\}}$ is the probability distribution that $P(\theta_1 = 0, \theta_2 = \delta_{\max}) = 1$. Note that even though it is likely that one cannot find any p_δ that leads to $P(\theta_1 = 0, \theta_2 = \delta_{\max}) = 1$, all the computations of (6), (7), (11), and (12) can still be carried out verbatim as long as $m > \delta_{\max}$. It is just that the $\Phi(1_{\{\theta_1=0,\theta_2=\delta_{\max}\}}, m, \Sigma_{\hat{\mathbf{w}}}, \Delta_{\mathsf{D}})$ becomes a pure computation formula and no longer has direct connection to the original linear control problem.

The reason why (52) holds is because for any legitimate $\theta_1, \theta_2 \in [0, \delta_{\max}]$ values, $F_{\theta_1,\theta_2}(m)$ (6) is a sum of at least $m - \delta_{\max}$ terms and can be written as

$$F_{\theta_1,\theta_2}(m) = \left(\sum_{i=\delta_{\max}+1}^{m} A^i \cdot \Sigma_{\hat{\mathbf{w}}} \cdot (A^i)^{\mathrm{T}}\right) + F_{\text{rem}} \tag{53}$$

$$\triangleq F_{0,\delta_{\max}}(m) + F_{\text{rem}}. \tag{54}$$

where F_{rem} in (53) is a positive semi-definite matrix that represents the summands that are not counted during $i \in [\delta_{\max} + 1, m]$, i.e., the remainder term. In (54) we further use $F_{0,\delta_{\max}}(m)$ to denote the fixed sum from $i = \delta_{\max} + 1$ to m. Since the $F_{\theta_1,\theta_2}(m)$ of different (θ_1, θ_2) is always a sum of the common $F_{0,\delta_{\max}}(m)$ plus a positive semi-definite remainder term (the remainder term may vary for different (θ_1, θ_2)), by the *physically degrading channel* argument, the water-filling computation based on $\overline{p}_{\theta_1,\theta_2}(m)$ and $F_{\theta_1,\theta_2}(m)$ is no smaller than the water filling computation based on $1_{\{\theta_1=0,\theta_2=\delta_{\max}\}}$ and $F_{0,\delta_{\max}}(m)$. Ineq. (52) is thus proven.

Recall that we assume that A is non-defective, we can apply eigen-decomposition to A and obtain $A = U\mathsf{diag}(\lambda_n(A))U^{-1}$ where $\lambda_n(A)$ denotes the eigenvalues of A. We use U^{H} to denote the conjugate transpose of U. Since $U \cdot U^{\mathsf{H}}$ is positive semi-definite and $\Sigma_{\hat{\mathbf{w}}}$ is invertible, we can always find a sufficiently small $\beta_{\min}$ such that

$$\beta_{\min} \cdot \left(U \cdot U^{\mathsf{H}}\right) \prec \Sigma_{\hat{\mathbf{w}}} \tag{55}$$

We now claim that

$$\begin{aligned}&\Phi(1_{\{\theta_1=0,\theta_2=\delta_{\max}\}}, m, \beta_{\min}U(U^{\mathsf{H}}), \Delta_{\mathsf{D}})\\ &\leq \Phi(1_{\{\theta_1=0,\theta_2=\delta_{\max}\}}, m, \Sigma_{\hat{\mathbf{w}}}, \Delta_{\mathsf{D}}).\end{aligned} \tag{56}$$

This inequality is true since (55) implies

$$A^i \cdot \Sigma_{\hat{\mathbf{w}}} \cdot (A^i)^{\mathrm{T}} \succeq A^i \cdot (\beta_{\min} \cdot U(U^{\mathsf{H}})) \cdot (A^i)^{\mathrm{T}} \tag{57}$$

for all $i \in [\delta_{\max} + 1, m]$. Therefore the $F_{0,\delta_{\max}}(m)$ computed from $\Sigma_{\hat{\mathbf{w}}}$ is physically degraded from the new $F_{0,\delta_{\max}}(m)$ computed from a different covariance matrix $\beta_{\min}U(U^{\mathsf{H}})$. As a result the water filling rate decreases once we replace $\Sigma_{\hat{\mathbf{w}}}$ by $\beta_{\min} \cdot U(U^{\mathsf{H}})$. Ineq. (56) is thus proven. It is worth pointing out that we now expand the domain of interest from real to complex Gaussian vectors, and thus from real to complex covariance matrices.

For any Gaussian covariance matrix Σ and distortion target Δ, we define the corresponding water-filling rate by

$$\mathsf{WF}(\Sigma, \Delta) \tag{58}$$

We now claim that

$$\begin{aligned}&\mathsf{WF}(A^m(\beta_{\min}U(U^{\mathsf{H}}))(A^m)^{\mathrm{T}}, \Delta_{\mathsf{D}})\\ &\leq \Phi(1_{\{\theta_1=0,\theta_2=\delta_{\max}\}}, m, \beta_{\min}U(U^{\mathsf{H}}), \Delta_{\mathsf{D}}).\end{aligned} \tag{59}$$

This is true because even if the modified water-filling formula in (12) involves water-filling across different (θ_1, θ_2) values, we are currently focusing on the singleton probabiliity distribution $1_{\{\theta_1=0,\theta_2=\delta_{\max}\}}$. Therefore, the modified water-filling computation (especially (12)) collapses to the traditional water-filling formula. We thus have

$$\Phi(1_{\{\theta_1=0,\theta_2=\delta_{\max}\}}, m, \beta_{\min}U(U^{\mathsf{H}}), \Delta_{\mathsf{D}}) = \mathsf{WF}(\Sigma_a, \Delta_{\mathsf{D}}) \tag{60}$$

where

$$\Sigma_a \triangleq \sum_{i=\delta_{\max}+1}^{m} A^i(\beta_{\min}U(U^{\mathsf{H}}))(A^i)^{\mathrm{T}}. \tag{61}$$

Since $\Sigma_a \succeq A^m(\beta_{\min}U(U^{\mathsf{H}}))(A^m)^{\mathrm{T}}$, by the same physically degrading channel argument we have proven (59).

Define

$$\Lambda_m = \mathsf{diag}\left(|\lambda_n(A)|^{2m}\right) \tag{62}$$

as a diagonal matrix whose n-th diagonal element being the the $2m$-th order of the absolute value of $\lambda_n(A)$. We notice that

$$A^m(\beta_{\min} \cdot U(U^{\mathsf{H}}))(A^m)^{\mathrm{T}} = \beta_{\min} \cdot U\Lambda_m U^{\mathsf{H}}. \tag{63}$$

Also note that since $(U^{-1})^{\mathsf{H}}U^{-1}$ is positive semi-definite, we can always find a sufficiently large positive constant $\gamma_{\max}$ such that

$$(U^{-1})^{\mathsf{H}}U^{-1} \prec \gamma_{\max} \cdot I_n \tag{64}$$

where I_n is the n-dimensional identity matrix.

We now claim that

$$\mathsf{WF}(\beta_{\min}\Lambda_m, \gamma_{\max}\cdot\Delta_{\mathsf{D}}) \leq \mathsf{WF}(\beta_{\min}U\Lambda_m U^{\mathsf{H}}, \Delta_{\mathsf{D}}). \quad (65)$$

The proof is as follows. Given any n-dimensional Gaussian column vector W with covariance matrix $\beta_{\min}\Lambda_m$, we will construct a quantizer W_q such that $E(\|W-W_q\|^2) \leq \gamma_{\max}\Delta_{\mathsf{D}}$ while satisfying $I(W;W_q) = \mathsf{WF}(\beta_{\min}U\Lambda_m U^{\mathsf{H}}, \Delta_{\mathsf{D}})$. The quantizer construction is as follows.

We first compute a vector $\tilde{W} = UW$, which is of covariance $\beta_{\min}U\Lambda_m U^{\mathsf{H}}$. We then apply the optimal complex Gaussian quantizer on $\tilde{W}$ with the target distortion being Δ_{D}. We denote the output of the optimal quantizer by $\tilde{W}_q$. Clearly such a quantizer will satisfy

$$I(\tilde{W};\tilde{W}_q) = \mathsf{WF}(\beta_{\min}U\Lambda_m U^{\mathsf{H}}, \Delta_{\mathsf{D}}) \quad (66)$$

$$\text{and } E\left(\left\|\tilde{W}-\tilde{W}_q\right\|^2\right) \leq \Delta_{\mathsf{D}}. \quad (67)$$

We then compute our desired quantizer W_q by $W_q = U^{-1}\tilde{W}_q$. It is easy to verify that

$$I(W;W_q) = I(UW;UW_q) \quad (68)$$

$$= I(\tilde{W};\tilde{W}_q) = \mathsf{WF}(\beta_{\min}U\Lambda_m U^{\mathsf{H}}, \Delta_{\mathsf{D}}) \quad (69)$$

$$\text{and } E\left(\|W-W_q\|^2\right) = E\left(\|U^{-1}(\tilde{W}-\tilde{W}_q)\|^2\right) \quad (70)$$

$$\leq \gamma_{\max}E\left(\|\tilde{W}-\tilde{W}_q\|^2\right) \leq \gamma_{\max}\Delta_{\mathsf{D}} \quad (71)$$

where (68) follows from U being invertible; (69) follows from the construction of $\tilde{W} = UW$ and $W_q = U^{-1}\tilde{W}_q$ and from (66); (70) follows from the construction of $\tilde{W} = UW$ and $W_q = U^{-1}\tilde{W}_q$; and (71) follows from (64) and (67).

Note that the existence of such a quantizer W_q implies (65) since $\mathsf{WF}(\beta_{\min}\Lambda_m, \gamma_{\max}\Delta_{\mathsf{D}})$ lower bounds the mutual information (68)–(69) for any quantizer satisfying (71).

Finally, we notice that in the expression of $\mathsf{WF}(\beta_{\min}\Lambda_m, \gamma_{\max}\Delta_{\mathsf{D}})$, only Λ_m depends on m, the constants $\beta_{\min}$ and $\gamma_{\max}$ depend on $\Sigma_{\hat{\mathbf{w}}}$ and $A = U\mathsf{diag}(\lambda_n(A))U^{-1}$ but not on m, see (55) and (64). By analyzing the expression of $\mathsf{WF}(\beta_{\min}\Lambda_m, \gamma_{\max}\Delta_{\mathsf{D}})$, we can easily show that

$$\mathsf{WF}(\beta_{\min}\Lambda_m, \gamma_{\max}\Delta_{\mathsf{D}}) = m\cdot\left(\sum_{n=1}^{N}\max(\log_2(|\lambda_n(A)|),0)\right) + o(m) \quad (72)$$

where $o(m)$ is the standard small-o notation. By combining (52), (56), (59), (63), (65), and (72), we have proven

$$\liminf_{m\to\infty}\frac{1}{m}\cdot\Phi(\overline{p}_{\theta_1,\theta_2}(m), m, \Sigma_{\hat{\mathbf{w}}}, \Delta_{\mathsf{D}}) \geq \sum_{n=1}^{N}\max(\log_2(|\lambda_n(A)|),0). \quad (73)$$

Thus far we have proven half of Proposition 2. The other half can be proven in a similar way by showing

$$\Phi(\overline{p}_{\theta_1,\theta_2}(m), m, \Sigma_{\hat{\mathbf{w}}}, \Delta_{\mathsf{D}})$$

$$\preceq \Phi(1_{\{\theta_1=\delta_{\max},\theta_2=0\}}, m, \Sigma_{\hat{\mathbf{w}}}, \Delta_{\mathsf{D}}) \quad (74)$$

$$\preceq \Phi(1_{\{\theta_1=\delta_{\max},\theta_2=0\}}, m, \beta_{\max}U(U^{\mathsf{H}}), \Delta_{\mathsf{D}}) \quad (75)$$

$$= \mathsf{WF}(\beta_{\max}U\overline{\Lambda}_m U^{\mathsf{H}}, \Delta_{\mathsf{D}}) \quad (76)$$

$$\preceq \mathsf{WF}(\beta_{\max}\overline{\Lambda}_m, \Delta_{\mathsf{D}}/\overline{\gamma}_{\max}) \quad (77)$$

where (74) follows from the same physically degraded channel argument as in the proof of (52); $\beta_{\max}$ in (75) is a sufficiently large positive constant that satisfies

$$\Sigma_{\hat{\mathbf{w}}} \preceq \beta_{\max}U(U^{\mathsf{H}}) \quad (78)$$

which always exists since U is invertible; (75) then follows the same argument as in the proof of (56); (76) holds when we set $\overline{\Lambda}_m$ as a diagonal matrix whose n-th diagonal element is computed from $\lambda_n(A)$, the n-th eigenvalue of A, by

$$\frac{|\lambda_n(A)|^2}{|\lambda_n(A)|^2-1}\left(|\lambda_n(A)|^{2(m+\delta_{\max})}-1\right). \quad (79)$$

The reason that (76) holds for this special choice of $\overline{\Lambda}_m$ is because the singleton probability distribution $1_{\{\theta_1=\delta_{\max},\theta_2=0\}}$ being used.

The constant $\overline{\gamma}_{\max}$ in (77) is a sufficiently large positive constant that satisfies

$$U^{\mathsf{H}}U \prec \overline{\gamma}_{\max}I_n \quad (80)$$

where I_n is the n-dimensional identity matrix.

To prove (77), we follow similar steps as used in the proof of (65). That is, given any n-dimensional Gaussian column vector W with covariance matrix $\beta_{\max}U\overline{\Lambda}_m U^{\mathsf{H}}$, we will construct a quantizer W_q such that $E(\|W-W_q\|^2) \leq \Delta_{\mathsf{D}}$ while satisfying $I(W;W_q) = \mathsf{WF}(\beta_{\max}\overline{\Lambda}_m, \Delta_{\mathsf{D}}/\overline{\gamma}_{\max})$. The quantizer construction is as follows.

We first compute a vector $\tilde{W} = U^{-1}W$, which is of covariance $\beta_{\max}\overline{\Lambda}_m$. We then apply the optimal complex Gaussian quantizer on $\tilde{W}$ with the target distortion being $\Delta_{\mathsf{D}}/\overline{\gamma}_{\max}$. We denote the output of the optimal quantizer by $\tilde{W}_q$. Clearly such a quantizer will satisfy

$$I(\tilde{W};\tilde{W}_q) = \mathsf{WF}(\beta_{\max}\overline{\Lambda}_m, \Delta_{\mathsf{D}}/\overline{\gamma}_{\max}) \quad (81)$$

$$\text{and } E\left(\left\|\tilde{W}-\tilde{W}_q\right\|^2\right) \leq \left(\frac{\Delta_{\mathsf{D}}}{\overline{\gamma}_{\max}}\right). \quad (82)$$

We then compute our desired quantizer W_q by $W_q = U\tilde{W}_q$. It is easy to verify that

$$I(W;W_q) = I(U^{-1}W;U^{-1}W_q) \quad (83)$$

$$= I(\tilde{W};\tilde{W}_q) = \mathsf{WF}(\beta_{\max}\overline{\Lambda}_m, \Delta_{\mathsf{D}}/\overline{\gamma}_{\max}) \quad (84)$$

$$\text{and } E\left(\|W-W_q\|^2\right) = E\left(\|U(\tilde{W}-\tilde{W}_q)\|^2\right) \quad (85)$$

$$\leq \overline{\gamma}_{\max}E\left(\|\tilde{W}-\tilde{W}_q\|^2\right) \leq \Delta_{\mathsf{D}} \quad (86)$$

where (83) follows from U^{-1} being invertible; (84) follows from the construction of $\tilde{W} = U^{-1}W$ and $W_q = U\tilde{W}_q$ and from (81); (85) follows from the construction of $\tilde{W} = U^{-1}W$ and $W_q = U\tilde{W}_q$; and (86) follows from (80) and (82).

The existence of such a quantizer W_q implies (77) since $\mathsf{WF}(\beta_{\max}U\overline{\Lambda}_m U^{\mathsf{H}}, \Delta_{\mathsf{D}})$ lower bounds the mutual information (83)–(84) for any quantizer satisfying (86).

Finally, we notice that in the expression of $\mathsf{WF}(\beta_{\max}\overline{\Lambda}_m, \Delta_{\mathsf{D}}/\overline{\gamma}_{\max})$, only $\overline{\Lambda}_m$ depends on m, the constants $\beta_{\max}$ and $\overline{\gamma}_{\max}$ depend on $\Sigma_{\hat{\mathbf{w}}}$ and $A = U\mathsf{diag}(\lambda_n(A))U^{-1}$ but not on m, see (78) and (80). By analyzing the expression of $\mathsf{WF}(\beta_{\max}\overline{\Lambda}_m, \Delta_{\mathsf{D}}/\overline{\gamma}_{\max})$, we can easily show that

$$\begin{aligned}&\mathsf{WF}(\beta_{\max}\overline{\Lambda}_m, \Delta_{\mathsf{D}}/\overline{\gamma}_{\max})\\&= m \cdot \left(\sum_{n=1}^{N} \max(\log_2(|\lambda_n(A)|), 0)\right) + o(m)\end{aligned} \tag{87}$$

where $o(m)$ is the standard small-o notation. By combining (77) and (87), we have proven

$$\begin{aligned}&\limsup_{m\to\infty} \frac{1}{m} \cdot \Phi(\overline{p}_{\theta_1,\theta_2}(m), m, \Sigma_{\hat{\mathbf{w}}}, \Delta_{\mathsf{D}})\\&\leq \sum_{n=1}^{N} \max(\log_2(|\lambda_n(A)|), 0).\end{aligned} \tag{88}$$

Jointly (73) and (88) complete the proof of Proposition 2.