Stability across a Gaussian Product Channel: Necessary and Sufficient Conditions

Utsaw Kumar, Vijay Gupta, and J. Nicholas Laneman

Abstract

We present necessary and sufficient conditions for stabilizing a discrete-time LTI plant in the mean squared sense when a sensor transmits the plant state information to a remotely placed controller across a Gaussian product channel. The Gaussian product channel models a continuous-time waveform Gaussian channel, where the encoder transmits information to the receiver across multiple noisy paths. The conditions presented are in terms of the power at the transmitter, noise variances and unstable eigenvalues. It is known that linear coding schemes may lead to overly restrictive stabilizability conditions in such scenarios. We present a non-linear coding scheme and present the resulting sufficient stablizability conditions. We then prove the necessity of the conditions using information theoretic tools.

I. Introduction

Networked control systems are now an active area of research. In this paper, we are interested in stabilizability of a scalar unstable linear time invariant (LTI) discrete time system across a Gaussian product channel (also known as parallel Gaussian channels). Stability conditions in the presence of one AWGN channel are available (e.g., [2]). Stabilizing the plant across a Gaussian relay [5], broadcast and multiple access channel [6] has also been considered. Interesting parallels of the problem with communication schemes achieving the capacity of a Gaussian channel with feedback through the Schalkwijk-Kailath (SK) scheme [9] are known [3].

The Gaussian product channel models a continuous-time waveform Gaussian channel in which the transmitter sends information to the receiver across multiple parallel channels, each parallel channel being individually modeled by an AWGN channel. The parallel channels may represent

Department of Electrical Engineering, University of Notre Dame. {ukumar, vgupta2, jnl}@nd.edu Work supported in part by NSF grant 0916716 for all the three authors, and by NSF grant 0846631 for the second author. Some results in the paper were presented in a preliminary form at CDC 2011.

different frequency bands, time instances, or in general different "degrees of freedom". Control across such a channel is inherently more difficult than control across a single channel. For instance, while it is known that for a single AWGN channel, the optimal encoding policies are linear [1], [2], for the parallel channel case, Yüksel and Taikonda [4] presented a counterexample which shows that in general linear encoding strategies may lead to overly restrictive stabilizability conditions.The inadequacy of linear controllers in achieving optimal stabilizability conditions or performance for a parallel Gaussian channel for discrete-time systems was also noted in [8]. In this paper, we consider non-linear encoder and decoder structures and present a tight characterization of the stabilizability conditions.

The remainder of the paper is organized as follows. We discuss the problem setup in Section II. In Section III the encoder, decoder and controller design are presented, which gives us the sufficiency results. In Section IV, a proof the necessary condition is presented.

Notation: The set of real numbers is denoted by $\mathbb{R}$, the l-dimensional Euclidean space by $\mathbb{R}^l$, and the set of positive integers by $\mathbb{Z}_+$. Denote the i-th basis vector by e_i, i.e. , $e_i \in \mathbb{R}^l$ has all elements 0, except for the i-th one which is unity. By $\log(x)$ we mean logarithm to base 2. Denote the eigenvalues of matrix M by $\lambda_j(M)$, $j = 1, \ldots, l$. M^T is the transpose of matrix M.

II. Problem Setup

Consider the setup in Fig. 1. An open loop unstable linear time-invariant process evolves as

$$S(k+1) = AS(k) + BU(k), \tag{1}$$

where $S(k) \in \mathbb{R}^l$ is the state and $U(k) \in \mathbb{R}$ is the control value. We assume that the pair (A, B) is controllable. We also assume that each component $S^i(0)$ ($i = 1, 2, \ldots, l$) of the initial condition $S(0)$ is a random variable uniformly distributed in the interval $[c, d]$ with a finite variance $\sigma^2_{S^i(0)} = \frac{(d-c)^2}{12}$.

For ease of exposition and without loss of generality, we assume that at every time step a sensor observes the state of the process $S(k)$ and transmits suitable information across the communication channel to the controller. The controller calculates a control input $U(k)$ and the actuator applies it to the process in (1). The communication channel from the plant to the controller is modeled as a Gaussian product channel, while the communication from the controller to the process is assumed to be noiseless. The Gaussian product channel consists of

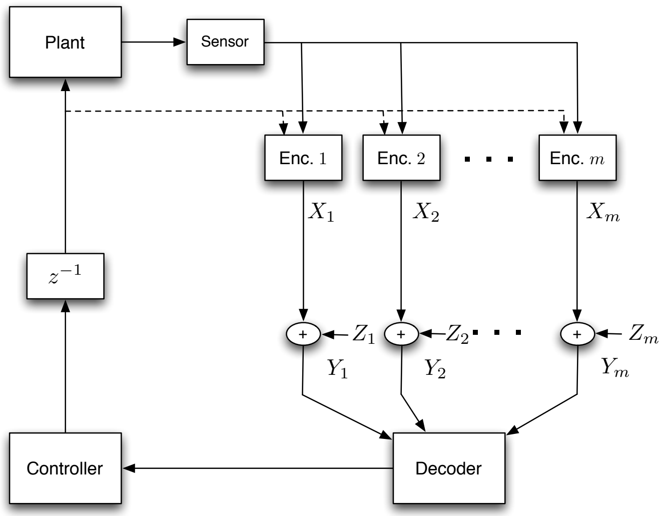

Fig. 1. Problem setup considered in the paper

m channels. The input and output of the i-th channel at time k is denoted by $X_i(k)$ and $Y_i(k)$ respectively, with $Y_i(k) = g_i X_i(k) + Z_i(k)$, where g_i is an attenuation factor due to path loss and $Z_i(k)$ is a noise term modeled by a zero-mean AWGN process with mean zero and variance σ_i^2. The noises on the various links are assumed to be mutually independent and white. We impose three constraints on the encoder and controller design:

- *Constraint C_1:* The control action must satisfy the cost constraint $\sum_{k=0}^{\infty} \mathbb{E}[U(k)^2] < \infty$.
- *Constraint C_2:* There is an average power constraint imposed on the signals transmitted by the encoders on the different channels and a common power constraint on the total power used. Thus, the encoding schemes must be such that the transmitted signals satisfy $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[X_i^2(k)] \leq P_i, \; i = 1, 2, \ldots m$, and that $\sum_{i=1}^{m} P_i \leq P$.
- *Constraint C_3:* The encoders are causal, but otherwise unconstrained in computation and memory. The information structure at the encoders is as follows. If h_i is the encoding policy at the encoder for the i-th input, then $X_i(k) = h_i(S(0), \ldots, S(k), U(0), \ldots, U(k-1))$.

Notice that we have assumed access to the control input at every encoder. Such an assumption makes sense since given the state values at time k and $k+1$, the sensor can calculate $U(k)$. In our construction, the controller will transmit additional information besides the control input. We assume that the actuator can isolate the correct component of the transmitted vector and apply it to the process.

Problem Statement: The problem we are interested in this paper is two fold. First, we want to design the maps h_i's and controller $U(k)$ so that the process (1) is mean square stabilized, while satisfying the design constraints C_1, C_2 and C_3, thus giving us sufficient conditions for stability. The design of the encoder map involves designing a scheme to divide the total power amongst the various inputs $X_1, X_2, \ldots X_m$ in an optimal way. Recall that a system is said to be stabilized in the mean squared sense if and only if, irrespective of the initial state $S(0)$, the following conditions are satisfied:

$$\mathbb{E}[S(k)] = 0, \quad \text{and} \quad \lim_{k\to\infty} \mathbb{E}[S(k)S^T(k)] = 0. \tag{2}$$

Second, we are interested in characterizing the necessary conditions of stability. We do this by using tools from information theory to lower bound the second moment of the state of the plant.

Main Result: The following is the main result of the paper:

Theorem 2.1: Consider the problem formulation presented in Section II. The necessary and sufficient condition for stabilizing the process (1) in mean squared sense over the Gaussian product channel is:

$$\sum_{i=1}^{l} \max\left\{0, \log|\lambda_i(A)|\right\} < \max \sum_{i=1}^{m} \frac{1}{2}\log\left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right), \tag{3}$$

where the maximization is over power allocations satisfying $\sum_{i=1}^{m} P_i = P$.
We prove the above theorem in two parts: section III presents a design that achieves stability if (3) is satisfied and the necessity of (3) for stability is proven in section IV.

Remark 2.1: The condition shown to be necessary for stability in Theorem 2.1 is less restrictive than the conditions in [8, Theorem 6]. This is because [8] concentrates on the class of linear encoders, decoders, and controllers. Since we allow for non-linear structures, we obtain less restrictive conditions. Further, the condition in Theorem 2.1 is also sufficient for stabilizability. In that sense, this is a tight characterization of the stabilizability region.

III. Sufficiency Results

In this section, we prove that if the condition (3) holds, then the system (1) can be stabilized. Our solution consists of two parts. First, we prove that if the controller has access to an estimate $\hat{S}(k)$ of the initial condition $S(0)$ of (1), such that the error $\epsilon(k) := \hat{S}(k) - S(0)$ has a covariance matrix that decreases geometrically in k, then the process can be stabilized by a suitable design of the controller. Then, we construct an encoding scheme that ensures that such an estimate is available to the controller. We begin with the first part.

A. Controller Design

Theorem 3.1: Consider the problem formulation stated in Section II. If

$$\mathbb{E}[\epsilon(k)] = 0, \tag{4}$$

$$\lim_{k\to\infty} A^k \mathbb{E}[\epsilon(k)\epsilon^T(k)](A^T)^k = 0, \tag{5}$$

the process (1) can be mean squared stabilized by a suitable choice of the controller.

Proof: Let K be such that the closed loop matrix $(A + BK)$ be Schur-stable. Such a K exists since (A, B) is controllable. Use the controller $U(k) = K\bar{S}(k)$, where

$$\bar{S}(k) = A^k \hat{S}(k) + \sum_{j=1}^{k} A^{k-j} BU(j-1). \tag{6}$$

With this controller, the process (1) evolves as $S(k+1) = (A + BK)S(k) + BK\delta(k)$, where $\delta(k) = S(k) - \bar{S}(k) = -A^k\epsilon(k)$. Since $S(0)$ is zero mean, (4) implies that $\mathbb{E}[\delta(k)] = 0$ and in turn $\mathbb{E}[S(k)] = 0$ at every k. Moreover, if (5) is satisfied, $\mathbb{E}[\delta(k)\delta^T(k)] \to 0$ as $k \to \infty$. Thus, since $A + BK$ is stable, we have $\lim_{k\to\infty} \mathbb{E}[S(k)S^T(k)] = 0$. ■

Note that even though we use a certainty equivalence controller, the controller does not generate its estimate using a Kalman filter. This result also provides the controller design.

We now provide a coding scheme that ensures that the relations (4) and (5) are satisfied. To set the notation, we begin by describing the scheme for the case of state dimension $l = 1$.

B. Scalar LTI Plant

In this section, we consider the process (1) evolving as

$$S(k+1) = aS(k) + U(k), \tag{7}$$

with $|a| > 1$. The encoders distribute the information about $S(0)$ amongst the various channel inputs $X_1(k), X_2(k), \ldots, X_m(k)$. Let $S_i(k)$ be the information about $S(0)$ transmitted through the i-th ($i \in \{1, 2, \ldots, m\}$) channel. $X_i(k)$ is calculated by scaling $S_i(k)$ to satisfy the power constraint C_2. The controller observes the outputs of the channels $Y_1(k), Y_2(k), \ldots, Y_m(k)$. and extracts relevant information $\hat{S}_1(k), \hat{S}_2(k), \ldots, \hat{S}_m(k)$. The estimate $\hat{S}(k)$ of the initial state $S(0)$ can be then calculated at the decoder as a function of the information collected from different links, i.e., $\hat{S}(k) = f(k, \hat{S}_1(k), \hat{S}_2(k), \ldots, \hat{S}_m(k))$. As before, define the overall estimation error as $\epsilon(k) := \hat{S}(k) - S(0)$. We let the estimation error for information sent through the i-th channel be denoted by $\epsilon_i(k) := \hat{S}_i(k) - S_i(0)$. Finally, let $\alpha(k)$ (resp. $\alpha_i(k)$) represent the variance of the estimation error $\epsilon(k)$ (resp. $\epsilon_i(k)$). We will now present our coding scheme and show that it satisfies the constraints in Theorem 3.1. We start with considering the special case when we have only one channel ($m = 1$), then extend it to the case when there are two channels ($m = 2$), and then generalize it for m channels.

1) Special Case m = 1: The coding scheme for the case when $m = 1$ works as follows [2]. Note that since there is only one channel, $S_1(0) = S(0), \hat{S}_1(k) = \hat{S}(k), \epsilon_1(k) = \epsilon(k)$ and $P_1 = P$.
Initialization: At time step $k = 0$, the encoder transmits

$$X_1(0) = \sqrt{\frac{P_1}{\sigma^2_{S_1(0)}}} S_1(0). \tag{8}$$

The decoder forms an estimate of $S_1(0)$ as $\hat{S}_1(0) = \frac{1}{g_1}\sqrt{\frac{\sigma^2_{S_1(0)}}{P_1}} Y_1(0)$. The estimation error $\epsilon_1(0)$ is zero-mean Gaussian with variance $\alpha_1(0)$, given by

$$\alpha(0) = \alpha_1(0) = \frac{\sigma^2_{S_1(0)} \sigma_1^2}{g_1^2 P_1}. \tag{9}$$

The controller calculates the control input according to (6) and transmits both the control $U(0) = K\bar{S}(0)$ and the estimate $\hat{S}(0)$ to the process.
Update: At each time step $k \geq 1$, the encoder transmits

$$X_1(k) = \sqrt{\frac{P_1}{\alpha_1(k-1)}} \epsilon_1(k-1). \tag{10}$$

The decoder updates its estimate as follows. At time $k \geq 1$, the decoder calculates the linear minimum mean squared error (MMSE) estimate of $S_1(0)$ given $Y_1(k)$ and $\hat{S}_1(k-1)$ as

$$\hat{S}_1(k) = \hat{S}_1(k-1) - \frac{\mathbb{E}[Y_1(k)\epsilon_1(k-1)]}{\mathbb{E}[{Y_1}^2(k)]} Y_1(k). \tag{11}$$

The controller calculates the control input according to (6) and transmits both the control $U(k) = K\bar{S}(k)$ and the estimate $\hat{S}(k)$ to the process. Note that the input $X_1(k)$ satisfies the power constraint C_2.

It can be seen that the estimation error $\epsilon_1(k)$ is Gaussian with zero mean and variance $\alpha_1(k)$. Since $\epsilon_1(k)$ is defined as $\hat{S}_1(k) - S_1(0)$, from (11) we obtain

$$\epsilon_1(k) = \epsilon_1(k-1) - \frac{\mathbb{E}[Y_1(k)\epsilon_1(k-1)]}{\mathbb{E}[{Y_1}^2(k)]} Y_1(k). \tag{12}$$

The variance of $\epsilon_1(k)$ can be obtained as

$$\alpha_1(k) = \mathbb{E}[\epsilon_1^2(k)] = \alpha_1(k-1) - \frac{\mathbb{E}^2[Y_1(k)\epsilon_1(k-1)]}{\mathbb{E}[{Y_1}^2(k)]}, \tag{13}$$

with the initial condition in (9). Now, since $\mathbb{E}[{Y_1}^2(k)] = g_1^2 P_1 + \sigma_1^2$ and $\mathbb{E}[Y_1(k)\epsilon_1(k-1)] = g_1\sqrt{P_1\alpha_1(k-1)}$, we have $\alpha_1(k) = \alpha_1(k-1)r_1$, where $r_1 = \left(\frac{\sigma_1^2}{g_1^2 P_1 + \sigma_1^2}\right)$. Note that since $\alpha(k) = \alpha_1(k)$ and $P_1 = P$,

$$\alpha(k) = \frac{\sigma_{S(0)}^2 \sigma_1^2}{g_1^2 P}\left(\frac{\sigma_1^2}{g_1^2 P + \sigma_1^2}\right)^k. \tag{14}$$

Proof for sufficiency of Theorem 2.1 *for scalar plants for* $m = 1$: It is easy to see that $\mathbb{E}[\epsilon(0)] = 0$. Further, since the linear minimum mean squared error estimator is unbiased, $\mathbb{E}[\epsilon(k)] = 0$ for all $k \geq 0$, which is the first condition in (5). The result follows by using the second condition in (5) and (14). $\square$

The right hand side of the condition in (3) is also the maximum rate at which information can be transmitted over a Gaussian point-to-point channel. Note that for a scalar system, if encoders have access to control law K and the matrices A and B, they can calculate $\hat{S}(k)$ from $U(k)$.

2) Special Case m=2: We consider next the case when $m = 2$. To develop a coding scheme for the case when more than one channel is present, we revisit a relevant result from information theory [10], and recognized in [4]. For a distributed source-channel coding to be optimal in the information-theoretic sense, two conditions need to be satisfied:

- The information transmitted on all the channels should be independent.
- Information is being transmitted through all the channels at respective channel capacities (the source and the channel need to be matched).

We develop a non-linear encoding scheme which will ensure that we transmit independent information over the two channels. Recall that $S(0)$ is uniformly distributed over $[c, d]$. Divide the interval $[c, d]$ into M_1 (we show how to choose M_1 later) disjoint, equal-length message intervals. Let c_j denote the partition points where $j \in \{1, 2, \ldots, M_1 - 1\}$. Also define $c_0 := c$ and $c_{M_1} := d$. A point x is said to be in the j-th interval I_j if $x \in [c_{j-1}, c_j]$. Consider a quantizer $T_1(\cdot)$ which maps each point of the j-th interval ($j \in \{1, 2\}$) to the midpoint of that particular interval. The message to be sent on the first channel is $T_1(S(0))$. The message to be sent on the second channel corresponds to the quantization error $S(0) - T_1(S(0))$. Thus, the messages are

$$\tilde{S}_1(0) = T_1(S(0)) \stackrel{(a)}{=} c + \left(j - \frac{1}{2}\right) \frac{1}{M_1} \text{ if } S(0) \in I_j,$$
$$\tilde{S}_2(0) = S(0) - \tilde{S}_1(0), \quad (15)$$

where (a) follows since the distance between adjacent quantized points is $\frac{1}{M_1}$.

Lemma 3.2: The random variables $\tilde{S}_1(0)$ and $\tilde{S}_2(0)$ defined in (15) are independent.

Proof: Consider the conditional probability $\Pr(\tilde{S}_2(0) = \gamma | \tilde{S}_1(0) \in I_j)$.

$$\Pr(\tilde{S}_2(0) = \gamma | \tilde{S}_1(0) \in I_j) = \frac{\Pr(\tilde{S}_2(0) = \gamma, \tilde{S}_1(0) \in I_j)}{\Pr(\tilde{S}_1(0) \in I_j)} = M_1 \Pr(\tilde{S}_2(0) = \gamma, \tilde{S}_1(0) \in I_j)$$
$$\stackrel{(a)}{=} \sum_{j=1}^{M_1} \Pr(\tilde{S}_2(0) = \gamma, \tilde{S}_1(0) \in I_j) = \Pr(\tilde{S}_2(0) = \gamma),$$

where (a) follows using the fact that $\Pr(\tilde{S}_2(0) = \gamma, \tilde{S}_1(0) \in I_j)$ is same for all j's. Thus, the random variables are independent. ■

Now, we define the information to be sent on the two parallel channels as $S_1(0) \triangleq \tilde{S}_1(0)$ and $S_2(0) \triangleq M_1 \tilde{S}_2(0)$. It can be seen that $S_1(0)$ takes values uniformly from a set with M_1 elements. The number of intervals M_1 is related the information rate R_1 and number of channel uses k over the first channel as $M_1 = 2^{kR_1}$. Note also that $S_2(0)$ is uniformly distributed in the interval $[-\frac{|d-c|}{2}, \frac{|d-c|}{2}]$ and thus has a variance $\sigma^2_{S_2(0)} = \frac{(d-c)^2}{12}$.

Now we transmit the messages $S_1(0)$ and $S_2(0)$ over the two channels recursively and independently of each other using the encoding scheme used for transmitting $S_1(0)$ for the case of $m = 1$ in Section III-B1 (see (8) and (10)). The decoder forms estimates $\hat{S}_i(k)$ of $S_i(0)$, $i = 1, 2$, the variances of which can be written using (14) as

$$\alpha_1(k) = \frac{\sigma^2_{S_1(0)} \sigma_1^2}{g_1^2 P_1} \left(\frac{\sigma_1^2}{g_1^2 P_1 + \sigma_1^2}\right)^k \triangleq \alpha_1(0) r_1^k, \quad (16)$$
$$\alpha_2(k) = \frac{\sigma^2_{S_2(0)} \sigma_2^2}{g_2^2 P_2} \left(\frac{\sigma_2^2}{g_2^2 P_2 + \sigma_2^2}\right)^k \triangleq \alpha_2(0) r_2^k. \quad (17)$$

Note that the estimation errors do not depend on the control inputs, and hence does not effect the controller design. The control input in this case is calculated as $U(k) = \left[\hat{S}_1(k), \hat{S}_2(k), K\bar{S}(k)\right]^T$, where the third component $K\bar{S}(k)$ is calculated using (6). The third component of $U(k)$ as defined above is extracted and applied to the process (7) by the actuator, whereas the i-th component ($i=1,2$) is used by the encoder i to update the i-th input, as given by equation (10). Note that because of the construction described above (15), the information sent on the two parallel channels is mutually independent. Also, except at time step $k=0$, the inputs to both the channels have a Gaussian distribution and are thus matched to the respective Gaussian channels. We now show that the condition (3) is sufficient for stability with this construction.

Proof for sufficiency of Theorem 2.1 *for scalar plants for* $m=2$: It is easy to see that $\mathbb{E}[\epsilon(0)]=\mathbb{E}[\epsilon_1(0)+\frac{\epsilon_2(0)}{M_1}]=0$. It is known that the linear minimum mean squared error is an unbiased estimator. Thus, $\mathbb{E}[\epsilon(k)]=0$ for all $k\geq 0$, which is the first condition in (5). To evaluate the estimation error variance $\alpha(k)$, we write

$$\mathbb{E}[\epsilon^2(k)]\overset{(a)}{=}\Pr(\hat{S}_1(k)\neq S_1(0))\mathbb{E}[\epsilon^2(k)|\hat{S}_1(k)\neq S_1(0)]+\Pr(\hat{S}_1(k)=S_1(0))\mathbb{E}[\epsilon^2(k)|\hat{S}_1(k)=S_1(0)].$$

The terms above can be written or bounded as follows.

$$\Pr(\hat{S}_1(k)\neq S_1(0))\leq \Pr\left[|\epsilon_1(n)|>\frac{1}{2M_1}\right]=2Q\left(\frac{1}{2M_1\sqrt{\alpha_1(n)}}\right)\overset{(a)}{=}2Q\left(\frac{2^{k\left(R_1-\frac{1}{2}\log(1+\frac{g_1^2P_1}{\sigma_1^2})\right)}}{2\sqrt{\frac{\sigma_{S_1(0)}^2\sigma_1^2}{g_1^2P_1}}}\right),$$

$$\mathbb{E}[\epsilon^2(k)|\hat{S}_1(k)\neq S_1(0)]\overset{(b)}{\leq}(d-c)^2,$$

$$\mathbb{E}[\epsilon^2(k)|\hat{S}_1(k)=S_1(0)]=\frac{\alpha_2(k)}{M_1^2}\overset{(c)}{=}\frac{\sigma_2^2}{12g_2^2P_2 2^{2kR_1}}\left(\frac{\sigma_2^2}{g_2^2P_2+\sigma_2^2}\right)^k,$$

where $Q(x)\triangleq\int_x^\infty\frac{1}{\sqrt{2\pi}}\exp(-\frac{y^2}{2})\,\mathrm{d}y$, (a) follows from the relation between M_1 and R_1, and (16), (b) follows using the fact the estimation error is upper bounded by the maximum distance between any two points on $[d,c]$ and (c) follows from (17). Further, since $\Pr(\hat{S}_1(k)=S_1(0))\leq 1$, we can upper bound $a^{2k}\mathbb{E}[\epsilon^2(k)]$ as

$$a^{2k}\mathbb{E}[\epsilon^2(k)]\leq a^{2k}Q\left(\frac{2^{k\left(R_1-\frac{1}{2}\log(1+\frac{g_1^2P_1}{\sigma_1^2})\right)}}{2\sqrt{\frac{\sigma_{S_1(0)}^2\sigma_1^2}{g_1^2P_1}}}\right)(d-c)+\frac{\sigma_2^2}{12g_2^2P_2}\frac{a^{2k}}{2^{2kR_1}}\left(\frac{\sigma_2^2}{g_2^2P_2+\sigma_2^2}\right)^k. \quad (18)$$

Since $Q(x)\sim\exp(\frac{-x^2}{2})$ for large x, the $Q(\cdot)$ term in (18) decreases doubly exponentially in k. On the other hand, the term a^{2k} increases exponentially. This implies that if $R_1<\frac{1}{2}\log\left(1+\frac{g_1^2P_1}{\sigma_1^2}\right)$, the first term in (18) goes to zero as $k\to\infty$ irrespective of the value of a. Moreover, if

$$\log a<R_1+\frac{1}{2}\log\left(1+\frac{g_2^2P_2}{\sigma_2^2}\right)<\frac{1}{2}\log\left(1+\frac{g_1^2P_1}{\sigma_1^2}\right)+\frac{1}{2}\log\left(1+\frac{g_2^2P_2}{\sigma_2^2}\right),$$

then the second term in (18) also approaches zero as $k \to \infty$. Now since we are allowed to choose P_1 and P_2 while satisfying the total power constraint in C_2, we can optimize the right hand side of the above equation to increase the stability region. Thus, if the condition in (3) is satisfied, then $a^{2k}\alpha(k) \to 0$ and mean square stability is obtained using the controller design. $\square$

3) Arbitrary value of m: The encoding scheme presented above can be generalized for transmission over the Gaussian product channel with $m > 2$ channels as follows. Divide the interval $[c, d]$ into M_1 disjoint, equal-length message intervals. Then divide each of these M_1 intervals into a further M_2 subintervals and so on till M_{m-1}. Define I_j^i to be the j-th interval ($j \in \{1, 2, \ldots, M_i\}$) for the i-th ($i \in \{1, 2, \ldots, m-1\}$) level quantizer T_i that maps the message in each point of the interval I_j^i to the midpoint of that interval. The message to be sent on the i-th channel ($i = 1, 2, \ldots, m-1$) corresponds to the output of the i-th quantizer $T_i(\cdot)$. The message to be sent on the m-th channel corresponds to the quantization error. Thus we design a set of $m-1$ quantizers as $\tilde{S}_1(0) = T_1(S(0))$, $\tilde{S}_2(0) = T_2(S(0) - \tilde{S}_1(0))$, $\tilde{S}_{m-1}(0) = T_{m-1}(S(0) - \sum_{i=1}^{m-2} \tilde{S}_i(0))$, and $\tilde{S}_m(0) = S(0) - \sum_{i=1}^{m-1} \tilde{S}_i(0)$. The following lemma is a generalization of Lemma 3.2 for arbitrary m presented without proof.

Lemma 3.3: The random variables $\tilde{S}_i$, $i = 1, \ldots, m$ defined above are mutually independent. Now, we define the messages to be sent on the i-th ($i = 1, \ldots, m$) parallel channel as $S_i(0) \triangleq \left(\prod_{j=1}^{i-1} M_j\right) \tilde{S}_i(0)$. Note that $S_i(0)$, $i = 1, 2, \ldots, m-1$ takes values uniformly from a set with M_i elements. As before, M_i is related to the information rate R_i and number of channel uses k over the i-th channel as $M_i = 2^{kR_i}$, $i = 1, 2, \ldots, m-1$. Also, $\tilde{S}_m(0)$ is uniformly distributed in the interval $[-\frac{|d-c|}{2}, \frac{|d-c|}{2}]$ and thus has a variance of $\frac{(d-c)^2}{12}$.

The encoding scheme is as follows. We transmit the messages $S_i(0)$, $i = 1, 2, \ldots, m$ over the m channels recursively in the same way as we transmitted $S_1(0)$ in for the case $m = 1$ (see (8) and (10)). The decoder forms estimates $\hat{S}_i(k)$ of $S_i(0)$, $i = 1, 2, \ldots, m$, the variances of which can written down using (14) as

$$\alpha_i(k) = \frac{\sigma_{S_i(0)}^2 \sigma_i^2}{g_i^2 P_i} \left(\frac{\sigma_i^2}{g_i^2 P_i + \sigma_i^2} \right)^k \triangleq \alpha_i(0) r_i^k. \tag{19}$$

For future reference, we denote the coding scheme by $\mathcal{S}(S(0), \hat{S}(k), m)$, where $S(0)$ is the initial condition, $\hat{S}(k)$ is the estimate of $S(0)$ at time k at the controller and m is the number of parallel channels.

The controller design is as follows. The controller calculates and transmits the input $U(k) = \left[\hat{S}_1(k), \ldots, \hat{S}_m(k), K\bar{S}(k)\right]^T$, where the last component is calculated using (6). The $m+1$-th component of $U(k)$ defined above is extracted and applied to the process (7), whereas the i-th component ($1 \le i \le m$) is used by the encoder i to update the i-th input. Note that because of the construction described above, the random variables transmitted on the parallel channels $i = 1, 2, \ldots, m$ are mutually independent. Also, except at time step $k = 0$, the inputs to both channels have a Gaussian distribution and are thus matched to the respective Gaussian channels. We can now prove the sufficiency part of Theorem 2.1.

Proof for sufficiency of Theorem 2.1 *for scalar plants*: We use the encoder, decoder and controller design outlined above. The first condition in (5) follows from an argument similar to the case when $m = 1, 2$. Define the event $E := (\hat{S}_1(k) = S_1(0), \hat{S}_2(k) = S_2(0), \ldots, \hat{S}_{m-1}(k) = S_{m-1}(0))$. We can write the estimation error variance $\alpha(k)$ as $\mathbb{E}[\epsilon^2(k)] = \Pr(\bar{E})\mathbb{E}[\epsilon^2(k)|\bar{E}] + \Pr(E)\mathbb{E}[\epsilon^2(k)|E]$ so that

$$a^{2k}\mathbb{E}[\epsilon^2(k)] = a^{2k}(\Pr(\bar{E})\mathbb{E}[\epsilon^2(k)|\bar{E}] + \Pr(E)\mathbb{E}[\epsilon^2(k)|E]) \tag{20}$$

Using arguments similar to the case when $m = 2$, we can prove that the first term in (20) goes to zero irrespective of the value of a if the following conditions are satisfied simultaneously.

$$R_i < \frac{1}{2}\log\left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right) \forall i = 1, 2, \ldots, m-1. \tag{21}$$

To obtain a condition for the second term to approach zero, rewrite

$$a^{2k}\Pr(E)\mathbb{E}[\epsilon^2(k)|E] \le a^{2k}\mathbb{E}[\epsilon^2(k)|E] = \frac{a^{2k}\alpha_m(k)}{\prod_{i=1}^{m-1} M_i} = \frac{\sigma_m^2}{12 g_m^2 P_m}\frac{a^{2k}}{\prod_{i=1}^{m-1} 2^{2kR_i}}\left(\frac{\sigma_m^2}{g_m^2 P_m + \sigma_m^2}\right)^k.$$

A sufficient condition for this term to approach zero is $\log a < \sum_{i=1}^{m-1} R_i + \frac{1}{2}\log\left(1 + \frac{g_m^2 P_m}{\sigma_m^2}\right) < \sum_{i=1}^{m} \frac{1}{2}\log\left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right)$, where the second inequality follows using (21). Now since we are allowed to choose P_i while satisfying the total power constraint in C_2, we can optimize the right hand side of the above equation to increase the stability region. Thus, if the condition in (3) is satisfied, then $a^{2k}\alpha(k) \to 0$ and mean square stability is obtained. □

C. Vector LTI plant

Without loss of generality, we assume that the matrice A is in the modal form

$$A = \begin{bmatrix} A_s & 0 \\ 0 & A_u \end{bmatrix}, \tag{22}$$

where $A_s \in \mathbb{R}^{(l-n)\times(l-n)}$ and $A_u^{-1} \in \mathbb{R}^{n\times n}$ are Schur stable. Note that $0 \le n \le l$, and we assume that an empty A_s (resp. A_u) corresponds to $n = l$ (resp. $n = 0$). Divide the state $S(k)$ into corresponding parts $S(k) = \begin{bmatrix} S_s^T(k) & S_u^T(k) \end{bmatrix}^T$.

Coding Scheme: The basic approach of the coding scheme for a vector process is to transmit the last n elements of the initial state $S(0)$ to the controller. To achieve this aim, n coding schemes proposed in Section III-B3 are used in parallel for each individual element. Thus, for each $j = 0, 1, \cdots, n-1$, at the sequence of times $kn+j$ ($k \in \mathbb{Z}_+$), the sensor, relay, and controller implement the coding scheme $\mathcal{S}(S^j(0), \hat{S}^j(k), m)$, with $S^j(k) = e_j^T S(0)$ and $\hat{S}^j(k) = e_j^T \hat{S}(k)$. The controller calculates the control input as follows. It maintains an estimate $\hat{S}(k)$ of the initial state $S(0)$. At each time k, such that $j = k \mod n$, it performs the following actions:

- Update $\hat{S}(k)$ as $\hat{S}(k) = \hat{S}(k-1) - e_j^T \hat{S}(k-1) e_j + \hat{S}^j(k) e_j$, with $\hat{S}_3(-1) = 0$.
- Calculate $\bar{S}(k)$ using the relation (6).
- Transmit the control input $U(k)$ given by. $U(k) = \left[\hat{S}_1^j(k), \ldots, \hat{S}_m^j(k), K\bar{S}(k)\right]^T$.

Stability Analysis: *Proof for sufficiency of Theorem* 2.1 *for vector plants*: We use the encoding scheme outlined above for each unstable state. Since A_s is stable, we do not need to update the first $l-n$ components of the error vector $\epsilon(k)$, which remain constant at 0. The other n components of $\epsilon(k)$ are updated every n time steps. Since $\epsilon(0) = 0$ and all updates in the coding scheme are linear, it is straightforward to see that equation (4) is satisfied. Since A_s is Schur stable, A_s^k approaches 0 as $k \to \infty$. Using the result from scalar case, we can see that $\lambda_j^{2k}\mathbb{E}\left[(\epsilon^j(k))^2\right] \to 0$, where $l - n + 1 \le j \le l$ and $\epsilon^j(k) = e_j^T \epsilon(k)$, if $\log|\lambda_j(A)| < \frac{1}{n} \max \sum_{i=1}^m \frac{1}{2} \log\left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right)$, where the maximization is over power allocations satisfying $\sum_{i=1}^m P_i = P$. Snce the diagonal elements in (5) approach 0 as $k \to \infty$, using the Cauchy-Schwarz inequality, it follows that the non-diagonal elements in (5) also approach 0. The theorem follows by noting that the above condition needs to be satisfied for $\forall 1 \le i \le m$. $\square$

We can rewrite the result in an alternative form.

Corollary 3.4: Consider the problem formulation presented in Section II with the coding scheme presented in Section III-B3 for each unstable state. The minimum power at the sensor that is needed for mean square stabilizing the process (1) of dimension l must be sufficient to guarantee that $\max \sum_{i=1}^m \frac{1}{2} \log\left(1 + \frac{g_i^2 P_i}{\sigma_i^2}\right) > \sum_{i=1}^l \max\{0, \log|\lambda_i(A)|\}$, where the maximization is over power allocations satisfying $\sum_{i=1}^m P_i = P$.

Remark 3.1: Water-filling Solution: The optimization over power allocations in (3) can be

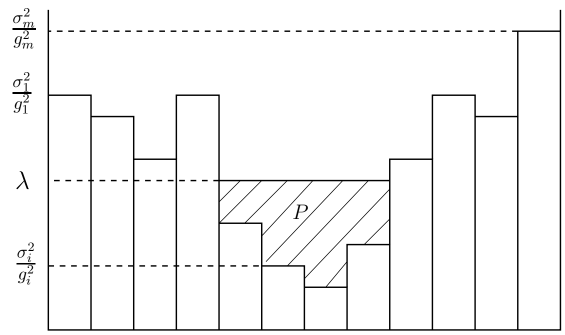

Fig. 2. Water-filling for the product channel. Area of the shaded region is equal to P.

solved using Lagrange multipliers. The solution is given by [11, Chapter 10] $P_i = \left(\lambda - \frac{\sigma_i^2}{g_i^2}\right)^+ = \max\left\{\lambda - \frac{\sigma_i^2}{g_i^2}, 0\right\}$, where the Lagrange multiplier λ is chosen to satisfy $\sum_{i=1}^m \left(\lambda - \frac{\sigma_i^2}{g_i^2}\right)^+ = P$. The optimal solution has the water-filling interpretation shown in Fig. 2. The vertical levels indicate the noise levels in the various channels. As the power P is increased, power is first alloted to the channel with lowest noise, then the next lowest and so on.

Remark 3.2: Constraint C_1: The constraints C_2 and C_3 are satisfied by construction of the coding scheme. We can also show that the constraint C_1 is satisfied by the proposed design.

Theorem 3.5: If $\hat{S}(k)$ is a linear MMSE estimate of $S(0)$, and the process (1) is mean squared stabilized, then the controller proposed in Theorem 3.1 satisfies $\sum_{k=0}^\infty \mathbb{E}[U^T(k)U(k)] < \infty$.

Proof: We have $\sum_{k=0}^\infty \mathbb{E}[U^T(k)U(k)] = \sum_{k=0}^\infty \mathbb{E}[\hat{S}^T(k)K^T K\hat{S}(k)]$, which is finite if the term $\sum_{k=0}^\infty \mathbb{E}[\text{tr}(\hat{S}(k)\hat{S}^T(k))]$ is finite. But $\sum_{k=0}^\infty \mathbb{E}[\text{tr}(\hat{S}(k)\hat{S}^T(k))] = \sum_{k=0}^\infty \mathbb{E}[\text{tr}(S(k)S^T(k))] + \sum_{k=0}^\infty \mathbb{E}[\text{tr}(\delta(k)\delta^T(k))]$. If the process (1) is mean squared stabilized, the first summation is finite. The second summation is $\sum_{k=0}^\infty \mathbb{E}[\text{tr}(\delta(k)\delta^T(k))] = \sum_{k=0}^\infty \mathbb{E}[\text{tr}(A^k\epsilon_3(k)\epsilon_3^T(k)(A^T)^k)] = \sum_{k=0}^\infty \sum_{j=1}^m \lambda_j^{2k}\mathbb{E}[(\epsilon_3^j(k))^2]$. Thus, if (3) is satisfied, the second term is finite as well. ■

IV. Necessity Results

To prove the necessary part of Theorem 2.1, we find a lower bound on the second moment of $S(k)$ and show that (3) is a necessary condition for this lower bound to converge to zero. A similar approach to proving necessity has been used for different settings [7], [12], [13]. Let Y_k represent a collection of all observations over the parallel channels at time k: $Y_k = \{Y_1(k), \dots, Y_m(k)\}$ and let Y^k represent a collection of $Y_j, 0 \le j \le k$: $Y^k = \{Y_0, \dots, Y_k\}$. Let $N(k)$ be the conditional entropy power of $S(k)$ conditioned on the event $\{Y^{k-1} = y^{k-1}\}$ averaged over all y^{k-1},

$$N(k) = \frac{1}{2\pi e}\mathbb{E}_{Y^{k-1}}\left[e^{\frac{2}{n}h(S_u(k)|Y^{k-1}=y^{k-1})}\right], \tag{23}$$

where S_u are unstable states of the plant under the decomposition (22).

Lemma 4.1: A necessary condition for (2) to hold is that $\lim_{k\to\infty} N(k) = 0$.

Proof: Conditional entropy power provides a lower bound on the mean square value of $S_u(k)$ [12]. Thus, $\mathbb{E}_{Y^{k-1}}\left[||S_u(k)||^2\right] \leq e^{(1-1/n)}\frac{1}{2\pi e}e^{\frac{2}{n}h(S_u(k)|Y^{k-1}=y^{k-1})}$. Taking expectation on both sides, we have $\mathbb{E}\left[||S_u(k)||^2\right] \leq e^{(1-1/n)}N(k)$, from which the result follows. ■

From Lemma 4.1, a necessary condition for mean square stability is provided by a necessary condition for $N(k)$ to be bounded. For this, we note that

$$\begin{aligned} N(k+1) &= \frac{1}{2\pi e}\mathbb{E}_{Y^k}\left[e^{\frac{2}{n}h(S_u(k+1)|Y^k=y^k)}\right] \\ &\stackrel{(a)}{=} \frac{1}{2\pi e}\mathbb{E}_{Y^k}\left[e^{\frac{2}{n}h(A_uS_u(k)|Y^k=y^k)}\right] \qquad (24) \\ &= \frac{|\det A_u|^{2/n}}{2\pi e}\mathbb{E}_{Y^{k-1}}\left[\mathbb{E}_{Y^k|Y^{k-1}}\left[e^{\frac{2}{n}h(S_u(k)|Y^k=y^k)}\right]\right] \\ &\stackrel{(b)}{\geq} \frac{|\det A_u|^{2/n}}{2\pi e}\mathbb{E}_{Y^{k-1}}\left[e^{\frac{2}{n}\mathbb{E}_{Y^k|Y^{k-1}}[h(S_u(k)|Y^k=y^k)]}\right] \\ &= \frac{|\det A_u|^{2/n}}{2\pi e}\mathbb{E}_{Y^{k-1}}\left[e^{\frac{2}{n}\mathbb{E}_{Y^k}[h(S_u(k)|Y^k=y^k)]}\right], \qquad (25) \end{aligned}$$

where (a) follows from the fact that the input $U(k)$ is a function of Y^k and (b) follows using Jensen's inequality [11, Chapter 2]. The expectation of the entropy term in (25) is given by $\mathbb{E}_{Y^k}[h(S_u(k)|Y^k = y^k)] = \mathbb{E}_{Y^k}[h(S_u(k)|Y^{k-1} = y^{k-1}) - I(S_u(k);Y_k|Y^{k-1} = y^{k-1})]$. Since $I(S_u(k);Y_k|Y^{k-1} = y^{k-1})] \leq C$, where C is the capacity of the Gaussian product channel, (25) and (23) yield $N(k+1) \geq |\det A_u|^{2/n}e^{(-2/m)C}N(k)$.Thus, using the expression for the capacity of a Gaussian product channel (right hand side of (3)), we prove the necessity of Theorem 2.1.

V. Conclusions

We derived sufficient and necessary conditions for mean square stabilizability of a linear time invariant open loop unstable plant over a Gaussian product channel. We present non-linear encoder and decoder designs to achieve mean square stability. When the sufficient conditions for stability are satisfied with equality, data about the initial condition is being transmitted at a rate equal to the capacity of a Gaussian product channel, which implies that our scheme is optimal.

REFERENCES

[1] R. Bansal and T. Basar, "Simultaneous design of measurement and control strategies for stochastic systems with feedback," *Automatica*, vol. 25, pp. 679-694, Sept. 1989.

[2] S. Tatikonda, and S. K. Mitter, "Control over noisy channels," *IEEE Trans. Auto. Control*, 49(7):1196 - 1201, July 2004.

[3] N. Elia, "When Bode meets Shannon: Control-oriented feedback communication schemes," *IEEE Trans. Auto. Control*, 49(9):1477-1488, Sep 2004.

[4] S. Yüksel, and S. Tatikonda "Distributed Sensing and Control over Communication Channels and a Counterexample," in *Proc. Annual Allerton Conference*, UIUC, Illinois, Sep. 2007.

[5] U. Kumar, V. Gupta, and J. N. Laneman, "Sufficient Conditions for Stabilizability over Gaussian Relay and Cascade Channels," in *Proc. Conference on Decision and Control*, Atlanta, Georgia, Dec. 2010.

[6] S. A. A. Zaidi, T. Oechtering, and M. Skoglund, "Sufficient Conditions for Closed-Loop Control Over Multiple-Access and Broadcast Channels," in *Proc. Conference on Decision and Control*, Atlanta, Georgia, Dec. 2010.

[7] G. N. Nair, F. Fagnani, S. Zampieri and R. J. Evans, "Feedback control under data rate constraints: an overview", *Proceedings of the IEEE*, 95(1):108-37, Jan. 2007.

[8] Z. Shu, and R. H. Middleton, "Stabilization over Power-Constrained Parallel Gaussian Channels," *IEEE Transactions on Automatic Control*, 56(7):1718–1724, July 2011.

[9] J. P. M. Schalkwijk and T. Kailath, "A Coding Scheme for Additive Noise Channels with Feedback - I: No bandwidth constraint," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 172– 182, April 1966.

[10] S. Shamai, S. Verdu, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inform. Theory*, 44(2):564–579, Mar. 1998.

[11] T. M. Cover, and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.

[12] J. S. Freudenberg, R. H. Middleton, and V. Solo, "The Minimal Signal-to-Noise Ratio Required to Stabilize over a Noisy Channel," *American Control Conference*, Minneapolis, Jun. 2006.

[13] P. Minero, M. Franceschetti, S. Dey, and G. N. Nair, "Data Rate Theorem for Stabilization Over Time-Varying Feedback Channels," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, Feb. 2009.