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

Control of systems with multiplicative observation noise

Moses Won and Gireeja Ranade.

***Abstract*— We consider the control of a linear system observed over multiplicative-noise. Specifically, the controller must stabilize the system using a control action based on observations of the system state that have been multiplied by i.i.d. random variables. While there is a long history of work on this fundamental problem, much of it has focused on understanding the performance of linear controllers, and the optimal control strategy for such a system remains unknown. In this paper, we consider the case of uniform multiplicative observation noise, and provide a non-linear control strategy based on the maximum a-posteriori (MAP) estimator of the state. We explicitly compute the convergence rates of different moments of the system under this control strategy, and find that the MAP-based strategy outperforms the best memoryless linear strategy when the "signal-to-noise" ratio (SNR) of the multiplicative noise, i.e. the ratio of the mean to the standard deviation, is low. In the high SNR regime we see that the MAP strategy is also a linear memoryless strategy, however, it is suboptimal and is outperformed by the optimal linear controller.**

I. INTRODUCTION

Systems with state-dependent or multiplicative noise can arise in the context of voltage-controlled resistors [1], [2], pulse-width and pulse-amplitude modulation systems [3]–[5], timing jitter, communications applications, and phase-lock loops [6], and linearization [7]. They also arise in physical settings where certain model parameters are dependent on changing quantities (e.g. vehicle velocity) [8], and in the modeling of biomechanical control problems [9]–[11].

In this paper, we consider the stabilization of the system:

$$X_{n+1} = a \cdot X_n + U_n \tag{1}$$
$$Y_n = C_n X_n. \tag{2}$$

Here, the initial state X_0 has density $f_X(\cdot)$. The multiplicative noise is captured by i.i.d. continuous random variables C_n with density $f_C(\cdot)$. The goal of the controller is to stabilize the system in an η-moment sense, i.e. $\sup \mathbb{E}[|X_n|^\eta] < \infty$ causally using information from the observations Y_i.

A. Background

Systems with multiplicative noise have a long-history in the field starting in the 70s and 80s. The robust control community has been driving the discussion around issues of modeling errors in control [12]–[22]. However, there are two main challenges with the robust control perspective. First, it can be overly conservative (e.g. [23]). Second, we are often interested in state-dependent noise in the time-domain (e.g. from linearization as in [7]), but this uncertainty does not translate easily to the frequency domain as often considered in the robust control literature. The H_∞ and H_2 robust control approaches with stochastic uncertainty in [24]–[27] only look at systems that are open-loop stable. The body of work on linear parameter varying (LPV) control [28], [29] assumes bounds on the parameter variations, which may not always be available. The related family of work on linear time varying systems (LTV) often assumes non-causal access to the varying parameters, which may be unrealistic [28].

A substantial body of work has focused on linear control strategies for systems with parameter uncertainty. For instance, [30]–[34] study linear control strategies for general multiplicative noise systems. For the case where there is multiplicative noise on the state and control only the uncertainty threshold principle shows that linear strategies can be optimal in a second-moment sense [35], [36].

The case where the multiplicative noise (on control or observation) is a discrete Bernoulli random variable is well understood from the perspective of second-moment stability by Sinopoli et al. [37] and related works [34], [38]–[40]. Here, we see that linear strategies are optimal in the second-moment sense for both multiplicative observation noise [37] as well as for multiplicative noise on the control action [39] (i.e. U_n is multiplied by an i.i.d. random variable.) [23] shows that linear strategies are optimal for the case on continuous multiplicative noise on the control action in all moment senses.

In contrast to this, we know from [41] that for continuous multiplicative observation noise, linear strategies are suboptimal for the problem, and one can get unbounded performance gains through using non-linear strategies. Recent work used neural-networks to identify non-linear control strategies for multiplicative observation noise systems [42], and provided examples of non-linear time-varying strategies that empirically perform well. However, no provable guarantees were given. More recent work has also taken other novel approaches to multiplicative noise systems using robust control techniques as well as techniques such as policy gradient [43]–[48].

Multiplicative observation noise can also be considered to be an informational constraint on the control of the system. Non-linear strategies have been shown to be required in other scenarios where explicit or implicit communication must occur in a linear control system (e.g. Witsenhausen's counterexample [49]). A large body of work has explored the intersection of information-theoretic ideas with control theory [50], [51] and the impact of explicit communication/information

M. Won and G. Ranade are with the department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA, USA moseswon@berkeley.edu, ranade@eecs.berkeley.edu

constraints on control. Data rate theorems explore the impact of noiseless communication channels [52]–[62]. A related family of work considers control with noisy communication channels [63]–[66].

Our problem setup corresponds to control over a non-coherent communication channel [67], i.e. a communication channel where the multiplicative fading coefficient is unknown and time-varying, which is common in fast-fading scenarios for rapidly changing communication channels. In this case, we know that the capacity of such a channel is bounded by $\log\log \text{SNR}$, where SNR is the signal-to-noise ratio of the channel. For the channel in (2), the channel SNR may be unbounded, since the magnitude of X_n is unbounded, effectively, there is no power-constraint on the channel. Hence, the traditional perspective of rate-limited control does not easily apply to analyze this system. Nonetheless, we take an informational perspective by explicitly computing the maximum a-posteriori (MAP) estimator for the system to compute a control action in this paper.

B. Main Contributions

In this paper we characterize the convergence regime and rate of the MAP-based control strategy for a system with multiplicative observation noise as in (3) for different values of SNR of the multiplicative noise and different moments of stability. This is given in Theorem 4.1. Further, we compare the performance of the MAP-strategy to the optimal linear memoryless strategy (see Fig. 1). As the SNR of the system goes to infinity, we see that the MAP strategy is outperformed by the simple linear memoryless strategy (i.e. the best linear controller for η-moment stability.) We will see that this is because as SNR grows the MAP-strategy converges to a linear memoryless strategy, but since it is agnostic to the moment being stabilized, ends up being suboptimal. We discuss this comparison in detail in Section. V. The proof of the main result is in Section. VI.

II. Problem setup

In this paper, we consider the control of a system in the presence of uniform multiplicative observation noise. In particular we consider:

$$X_{n+1} = X_n + U_n. \qquad X_0 \sim \text{Unif}[-1,1]. \tag{3}$$

$$Y_n = C_n X_n. \qquad C_n \sim \text{Unif}[\mu - w, \mu + w]. \tag{4}$$

X_n is the scalar state of the system, and the goal of the controller is to choose U_n such that $\lim_{n\to\infty} \mathbb{E}[|X_n|^\eta] = 0$. The control U_n is chosen as any function of the observations $Y_0^n := (Y_0, Y_1, \dots, Y_n)$. The distributions of X_0 and C_n are known to the controller. Let $\mathcal{H}_n$ be the σ-algebra generated by all observations Y_0^n.

Note in (3) we set the growth of the system to be $a = 1$. We are interested in the rate of decay i.e. $r = \liminf_{n\to\infty} -\frac{1}{n\eta} \log \mathbb{E}[|X_n|^\eta / |X_0|^\eta]$, since the largest a for which (1) can be stabilized in an η-moment sense is given by 2^r.

Without loss of generality, we will consider $C_n \sim \text{Unif}[\mu - 1, \mu + 1]$, i.e. we will set the uniform random variable to be drawn from an interval of length 2, centered at mean μ. Thus, the variance of C_n, $\text{Var}(C_n) = \sigma^2 = \frac{1}{3}$. As the parameter μ varies from 0 to ∞, the "signal-to-noise" ratio (SNR), defined as $\frac{\mu}{\sigma}$ varies between 0 to ∞ as well. We notice that the SNR is a sufficient parameter to capture the informational impact of the multiplicative noise, since the mutual information $\mathcal{I}(X; CX) = \mathcal{I}(X; \frac{CX}{b})$, for any constant $b \neq 0, b \in \mathbb{R}$. We define $s = \frac{\mu}{w}$, and note that $\frac{\mu}{\sigma} = \sqrt{3}s$.

We are interested in understanding the η-th moment stability of the system, i.e.

Definition 2.1 (η-moment stability): We say the system in (3) is stable in an η-moment sense using control strategy U_0^n if $\sup_n \mathbb{E}[|X_n|^\eta] < \infty$.

As $\eta \to 0$, this converges to the notion of stability in probability [50].

III. Control Strategies

We define the two strategies we will compare: the linear control strategy and the MAP control strategy.

Definition 3.1 (Linear Control): A control strategy of the from $U_n = dY_n$ for all n for some fixed $d \in \mathbb{R}$ is called a linear control strategy.

To define the MAP control strategy, first we define the MAP estimator for the state at time k, using observations up to time n as:

$$\widehat{X}_{k|n}^{\text{MAP}} := \arg\max_x f_X(X_k = x | \mathcal{H}_n). \tag{5}$$

We will be most concerned with the MAP estimation of the initial state X_0, given by $\widehat{X}_{0|n}^{\text{MAP}}$.

With this, we define the MAP control strategy as the difference between the previous and current estimate of X_0.

Definition 3.2 (MAP control): For the system in (3) define the sequence of MAP controls as

$$U_0 = -\widehat{X}_{0|0}^{\text{MAP}} \tag{6}$$

$$U_n = \widehat{X}_{0|n-1}^{\text{MAP}} - \widehat{X}_{0|n}^{\text{MAP}}, n \geq 1. \tag{7}$$

With this we have $\sum_{i=0}^n U_i = -\widehat{X}_{0|n}^{\text{MAP}}$. We can also show that $U_n = -\widehat{X}_{n|n}^{\text{MAP}}$.

Finally, we also know that the system state X_n is always equal to the difference between X_0 and the current MAP estimate.

Lemma 3.3: For the system in (3) with MAP Control, the state for all $n \geq 1$ is:

$$X_n = X_0 - \widehat{X}_{0|n-1}^{\text{MAP}}. \tag{8}$$

The proof is straightforward and is omitted.

We now consider the evolution of the conditional density which allows evaluation of $\widehat{X}_{0|n}^{\text{MAP}}$.

Lemma 3.4 (Conditional Density of $X_0|\mathcal{H}_n$): For the system in (3), with controls u_0^n that are functions of the observations y_0^n, the conditional density of X_0 given $\mathcal{H}_n$ for $n \geq 1$ is

$$f_X(X_0 = x | \mathcal{H}_n) = \frac{f_X(x)}{f_{Y_0^n}(y_0^n)} \prod_{j=0}^{n} \frac{f_C\left(\frac{y_j}{x + \sum_{i=0}^{j-1} u_i}\right)}{|x + \sum_{i=0}^{j-1} u_i|}, \tag{9}$$

where

$$f_Y(Y_n = y_n \mid X_0 = x, \mathcal{H}_{n-1}) = \frac{f_C\left(\frac{y_n}{x+\sum_{i=0}^{n-1} u_i}\right)}{|x + \sum_{i=0}^{n-1} u_i|}.$$

By the definition of our system, the densities of X_0 and C_n, $f_X(\cdot)$ and $f_C(\cdot)$ can be written as $f_X(x) = \frac{1}{2}\mathbb{1}\{x : |x| \leq 1\}$ and $f_C(c) = \frac{1}{2}\mathbb{1}\{c : |c - s| \leq 1\}$, where $\mathbb{1}\{S\}$ is the indicator function for the set S.
The proof will be included in the full version of the paper.

We now investigate the effect of s on $f_X(X_0 = x|\mathcal{H}_n)$. The influence of s enters through the $f_C(\cdot)$ expressions in $f_Y(Y_n = y_n|X_0 = x, \mathcal{H}_{n-1})$ where under the map strategy with $w = 1$ we have that:

$$f_Y(Y_n = y_n|X_0 = x, \mathcal{H}_{n-1}) = \frac{f_C\left(\frac{y_n}{x-\widehat{X}^{\text{MAP}}_{0|n-1}}\right)}{\left|x - \widehat{X}^{\text{MAP}}_{0|n-1}\right|}.$$

Since $f_C\left(\frac{y_n}{x-\widehat{X}^{\text{MAP}}_{0|n-1}}\right) = \frac{1}{2}\mathbb{1}\left\{x : \left|\frac{y_n}{x-\widehat{X}^{\text{MAP}}_{0|n-1}} - s\right| \leq 1\right\}$, the intervals where $f_X(X_0 = x|\mathcal{H}_n)$ is not zero, is the intersection of the intervals of x satisfying $\left|\frac{y_n}{x-\widehat{X}^{\text{MAP}}_{0|n-1}} - s\right| \leq 1$. This is formalized later in Lemma 6.1.

Impact of increasing s: Consider a fixed σ. As the value of μ (and therefore s increases), the support of the multiplicative noise C_n shifts further and further away from zero. If $s = 0$ the observation Y_n contains no information about the sign of X_n, however for $s \geq 1$ the sign of X_n is perfectly disambiguated from Y_n. When $s > 1$, not only do we know $\text{sgn}(X_n)$, we now have an upper bound and lower bound on the value of $|X_n|$.

IV. Main Result

Our main result gives an explicit convergence rate for the system (3) using the MAP controller.

Theorem 4.1: For the system in (3) with MAP control, $X_n \to 0$ a.s. and $\widehat{X}^{\text{MAP}}_{0|n} \to X_0$ a.s. Further, the convergence rate of the η-th moment of the state to zero is given by

$$\liminf_{n\to\infty} -\frac{1}{n\eta}\log \mathbb{E}\left[\left|\frac{X_n}{X_0}\right|^\eta\right] = \begin{cases} \frac{1}{\eta}\log(1+\eta) & 0 \leq s < 1 \\ \frac{1}{\eta}\log\left((1+\eta)\left(\frac{s+1}{2}\right)^\eta\right) & 1 \leq s. \end{cases}$$

We will prove the two cases $0 \leq s < 1$ and $s \geq 1$ separately. We start with the case $s \geq 1$ which is more straightforward, as in this case the observation Y_n contains full information about the sign of the state X_n. The case when $0 \leq s < 1$ has a similar flavor, however must follow a more nuanced analysis to account for the uncertainty on the sign of the state. Surprisingly, if there is even a small amount of sign uncertainty, we see that increased SNR does not help — the convergence rate has no dependence on s. As soon as the sign issue is resolved, i.e. $s \geq 1$, increased SNR increases the rate of convergence.

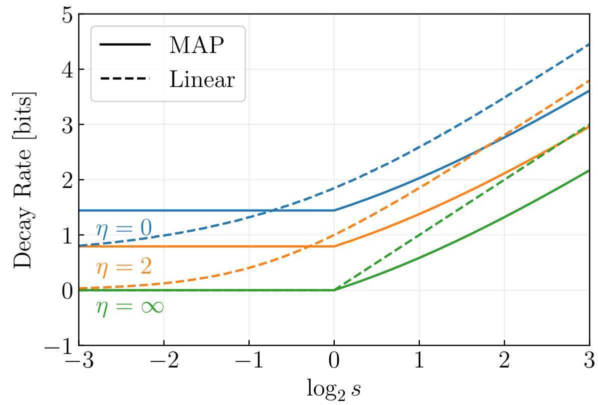

Fig. 1. This plot shows decay rate for system in an η-moment sense for different values of η. The dashed lines show the performance of the best memoryless linear strategy for each of the moment-senses, while the solid lines show the performance of the MAP-strategy.

V. Comparison with Linear Controller

The memoryless linear strategy that maximizes the η-moment decay rate for the system (3) is given by $U_n = d^\star_\eta \cdot Y_n$, where $d^\star_\eta = -\arg\min_d \mathbb{E}[|1 + d \cdot C|^\eta]$ and $C \sim \text{Unif}[\mu - w, \mu + w]$, i.e. has the distribution of the multiplicative noise. In the case of $\eta = 2$, a straightforward computation gives $d^\star_\eta = -\frac{\mu}{\mu^2+\sigma^2} = \frac{-s}{s^2+1/3}$. In the case where $\eta = 0$, we can compute $d^\star_0$ numerically as $\arg\min_d \mathbb{E}[\log|1 + d \cdot C|]$ as in [23, Thm. 3.4]. For $\eta = \infty$, $d^\star_\infty = -\frac{w}{\mu} = -\frac{1}{s}$ if $s > 1$ and $d^\star_\infty = -1$ otherwise as in [23]. The maximum decay rate when $\eta = \infty$ is given as $-\log(s)$ if $s > 1$ and 0 otherwise.

This performance of the memoryless linear control strategy in comparison with the MAP-strategy is illustrated in Fig. 1. We clearly see that at low values of s (i.e. Low SNR), the MAP strategy outperforms the linear strategy for small moments, but as $\eta \to \infty$, both strategies perform identically and decay with rate 0 in the limit. This is because the controller does not know the sign of the state and thus neither the MAP nor the linear strategy can keep the state bounded with probability one (as is required in the $\eta = \infty$, i.e. robust control scenario).

As $s \to \infty$ we see that the linear controller outperforms the MAP strategy. As we will see in Sec. VI, when $s > 1$ the MAP strategy ends up being a linear memoryless strategy that chooses $U_n = -\frac{1}{s+1}Y_n$. This strategy does not depend on η, but ends up being too conservative across all the moment senses (e.g. the optimal linear strategy for the ∞-moment is $d^\star_\infty = -\frac{1}{s}$.

These findings support the conjecture that follows from the consideration of bit-level-models of the system, as in [68], [69], that in the high SNR regime, linear control strategies are optimal for this problem.

VI. Proofs

We start by providing a few lemmas that will be helpful in the main proofs. This first lemma identifies the support of the terms in the conditional density that will be useful for computing the MAP estimator.

Lemma 6.1: Let $\mathcal{X}$ be the set of values of x that satisfy the inequality

$$\left|\frac{y}{x-u} - s\right| \leq 1,$$

for $s \geq 0$. Then

$$\mathcal{X} = \begin{cases} \frac{|y|}{\text{sgn}(y)s+1} \leq x-u \leq \frac{|y|}{\text{sgn}(y)s-1} & s>1, \\ \frac{|y|}{2} \leq (x-u)\text{sgn}(y) & s=1, \\ x-u \leq \frac{|y|}{\text{sgn}(y)s-1} \text{ or } \frac{|y|}{\text{sgn}(y)s+1} \leq x-u & s<1. \end{cases}$$

The proof of this is provided in the full version of this paper.

A. $s \geq 1$

Without loss of generality, let us assume in this case $\mu \geq 1$ and $X_0 > 0$. The proofs for the cases $\mu \leq -1$ and $X_0 < 0$ will proceed identically with some sign flips. Since we assume $\mu \geq 1$, we note that $\text{sgn}(C_n) = 1$.

First, we prove a lemma that identifies the MAP-control in this case.

Lemma 6.2: For all $n \geq 0$ we have that

$$\widehat{X}^{\text{MAP}}_{0|n} = \frac{1}{s+1}\sum_{i=0}^{n} y_i, \quad \text{and}$$

$$X_{n+1} = X_0 \cdot \prod_{i=0}^{n}\left(1 - \frac{C_i}{s+1}\right).$$

Proof: We proceed by induction on n. For $n=0$, we must show that $\widehat{X}^{\text{MAP}}_{0|0} = \frac{y_0}{s+1}$, from which it follows that $X_1 = X_0 - \widehat{X}^{\text{MAP}}_{0|0} = X_0\left(1 - \frac{C_0}{s+1}\right)$.

Note that $\text{sgn}(Y_0) = 1$, since we assume $X_0 > 0$. To evaluate $\widehat{X}^{\text{MAP}}_{0|0}$, by lemmas 3.4 and 6.1 we know the conditional density only has nonzero values $f_X(X_0 = x|\mathcal{H}_0) = \frac{1}{f_Y(y_0)} \cdot \frac{1}{4|x|}$ for $x \in \mathcal{X}_0 \cap [-1,1]$ where $\mathcal{X}_0$:

$$\mathcal{X}_0 := \begin{cases} \left\{x : \frac{y_0}{2} \leq x\right\} & s=1 \\ \left\{x : \frac{y_0}{s+1} \leq x \leq \frac{y_0}{s-1}\right\} & s>1. \end{cases}$$

Since $\widehat{X}^{\text{MAP}}_{0|0} = \arg\max_{x\in\mathcal{X}_0\cap[-1,1]} \frac{1}{|x|}$, $\widehat{X}^{\text{MAP}}_{0|0}$ will be the smallest magnitude value of x in $\mathcal{X}_0$. We observe that in both cases where $s=1$ and $s>1$ that this smallest magnitude value is $\frac{y_0}{s+1}$, and thus $\widehat{X}^{\text{MAP}}_{0|0} = \frac{y_0}{s+1}$. This establishes the base case.

Now, assume for all $n \leq k$ that $\widehat{X}^{\text{MAP}}_{0|n} = \frac{1}{s+1}\sum_{i=0}^{n} y_i$ and $X_{n+1} = X_0 \cdot \prod_{i=0}^{n}\left(1 - \frac{C_i}{s+1}\right)$. This implies that $X_n \geq 0$ for all $n \leq k$, since $C_i < s+1$. Hence, $Y_n \geq 0$ for all $n \leq k$. Thus $\widehat{X}^{\text{MAP}}_{0|n}$ must be increasing in n.

Now to evaluate $\widehat{X}^{\text{MAP}}_{0|k+1}$, consider $f_X(X_0 = x|\mathcal{H}_{k+1})$. By Lemma 3.4 we have that

$$f_X(X_0 = x|\mathcal{H}_{k+1}) = \frac{1}{2f_Y\left(y_0^{k+1}\right)}\prod_{i=0}^{k+1}\frac{1}{2\left|x - \widehat{X}^{\text{MAP}}_{0|i-1}\right|}$$

on the interval $\mathcal{X}_{k+1} \cap [-1,1]$. This function is decreasing in $|x|$ for $|x| \geq |\widehat{X}^{\text{MAP}}_{0|k}|$. We have from Lemma 6.1:

$$\mathcal{X}_{k+1} = \bigcap_{i=0}^{k+1}\begin{cases} \left\{x : \frac{y_i}{2} \leq x - \widehat{X}^{\text{MAP}}_{0|i-1}\right\} & s=1 \\ \left\{x : \frac{y_i}{s+1} \leq x - \widehat{X}^{\text{MAP}}_{0|i-1} \leq \frac{y_i}{s-1}\right\} & s>1. \end{cases}$$

As the $|\widehat{X}^{\text{MAP}}_{0|i}|$ are increasing, $\mathcal{X}_{k+1}$ becomes

$$\mathcal{X}_{k+1} = \begin{cases} \left\{x : \frac{y_{k+1}}{2} \leq x - \widehat{X}^{\text{MAP}}_{0|k}\right\} & s=1 \\ \left\{x : \frac{y_{k+1}}{s+1} \leq x - \widehat{X}^{\text{MAP}}_{0|k} \leq \frac{y_{k+1}}{s-1}\right\} & s>1. \end{cases}$$

By similar reasoning as in the base case, $\widehat{X}^{\text{MAP}}_{0|k+1} = \widehat{X}^{\text{MAP}}_{0|k} + \frac{y_{k+1}}{s+1}$, and the rest follows similarly as well. ■

Lemma 6.3: For the system in (3) with MAP control and $s \geq 1$, $X_n \to 0$ a.s., $\widehat{X}^{\text{MAP}}_{0|n} \to X_0$ a.s., and

$$-\frac{1}{n}\log\mathbb{E}\left[\left|\frac{X_n}{X_0}\right|^{\eta}\right] = \log\left((1+\eta)\left(\frac{s+1}{2}\right)^{\eta}\right).$$

Proof: From lemma 6.2, $X_n = X_0\prod_{i=0}^{n-1}\left(1 - \frac{C_i}{s+1}\right)$. Let $V_i := 1 - \frac{C_i}{s+1}$. Then $V_i \sim \text{Unif}[0, \frac{2}{s+1}]$. Since $s \geq 1$, $0 \leq V_i \leq 1$ w.p. 1. By the second Borel-Cantelli lemma, since the V_i are independent and $\sum_{i=0}^{\infty}\mathbb{P}(0 \leq V_i < 1) = \infty$, $\mathbb{P}(0 \leq V_i < 1 \text{ i.o.}) = \mathbb{P}(\lim_{n\to\infty} X_n = 0) = 1$. So $X_n \to 0$ a.s. Since $X_{n+1} = X_0 - \widehat{X}^{\text{MAP}}_{0|n}$, then $\widehat{X}^{\text{MAP}}_{0|n} \to X_0$ a.s.

Now we compute the η-moment decay rate. Since all the $V_i \sim \text{Unif}[0, \frac{2}{s+1}]$ are i.i.d., we have that:

$$\mathbb{E}\left[\left|\frac{X_n}{X_0}\right|^{\eta}\right] = \mathbb{E}[V_0^{\eta}]^n = (1+\eta)\left(\frac{s+1}{2}\right)^{\eta},$$

which gives the result. ■

B. $s \in [0,1)$

Once again, we will assume that $X_0 > 0$, the other case will proceed similarly. Unlike in the case $s \geq 1$, in this case the sign of X_n cannot be resolved from the observation Y_n. Hence, the system must spend controls learning the sign of X_0. We will first show that $\widehat{X}^{\text{MAP}}_{0|n} \to X_0$ a.s., and in doing so, there must be a finite number of steps with probability 1 until $\text{sgn}\left(\widehat{X}^{\text{MAP}}_{0|n}\right) = \text{sgn}(X_0)$.

We see from Lemma 6.1 when $s \in [0,1)$, the support of $f_X(X_0 = x|\mathcal{H}_n)$ has a single growing interval about 0 where X_0 cannot be. The endpoints of this interval are the key points of interest, since these are the values that $\widehat{X}^{\text{MAP}}_{0|n}$ can take on, as we will see. For this we first define:

Definition 6.4 (Positive and Negative $\widehat{X}^{MAP}_{0|n}$ Candidates): We define the positive candidate for the MAP estimator as

$$\widehat{X}^{+}_{0|n} = \arg\max_{x\geq 0} f_Y(Y_0^n = y_0^n \mid X_0 = x) \tag{10}$$

i.e. the value $x \geq 0$ that maximizes the conditional density of Y_0^n given X_0. Similarly, we define the negative candidate:

$$\widehat{X}^{-}_{0|n} = \arg\max_{x\leq 0} f_Y(Y_0^n = y_0^n \mid X_0 = x). \tag{11}$$

These candidates update according to the following lemma, which is omitted for space reasons.

Lemma 6.5 (Update rule for $\widehat{X}^+_{0|n}, \widehat{X}^-_{0|n}$): At each time n, $\widehat{X}^{\text{MAP}}_{0|n}$ must be one of the two values $\widehat{X}^+_{0|n}$ or $\widehat{X}^-_{0|n}$. Furthermore,

$$\widehat{X}^+_{0|n+1} = \max\left(\widehat{X}^{\text{MAP}}_{0|n} + \frac{|Y_{n+1}|}{1+\text{sgn}(Y_{n+1})s}, \widehat{X}^+_{0|n}\right). \quad (12)$$

If $\widehat{X}^+_{0|K} > 1$, then $\widehat{X}^{\text{MAP}}_{0|n} = \widehat{X}^-_{0|n} \forall n \geq K$. Similarly,

$$\widehat{X}^-_{0|n+1} = \min\left(\widehat{X}^{\text{MAP}}_{0|n} - \frac{|Y_{n+1}|}{1-\text{sgn}(Y_{n+1})s}, \widehat{X}^-_{0|n}\right). \quad (13)$$

If $\widehat{X}^-_{0|K} < -1$ then $\widehat{X}^{\text{MAP}}_{0|n} = \widehat{X}^+_{0|n} \forall n \geq K$.

Our proof approach will be to show that $\widehat{X}^{\text{MAP}}_{0|n}$ must converge to one of $\widehat{X}^+_{0|n}$ or $\widehat{X}^-_{0|n}$, which in turn converges to X_0. For this we first establish the following properties.

Lemma 6.6: Under MAP control, $\widehat{X}^+_{0|n}$ is increasing and $\widehat{X}^-_{0|n}$ is decreasing in n. Furthermore, if $X_0 > 0$, then $\forall n, \widehat{X}^+_{0|n} \leq X_0$ and $\widehat{X}^-_{0|n} \leq X_0$.
The proof for this and subsequent lemmas are in the full version of the paper.

Now, recall that we are assuming without loss of generality that $X_0 > 0$. Through the next lemma we characterize the impact on the system of making the "right" choice, i.e., choosing $\widehat{X}^{\text{MAP}}_{0|n} = \widehat{X}^+_{0|n}$ two times in a row, as well as the impact of making the "wrong" choice, i.e. choosing $\widehat{X}^{\text{MAP}}_{0|n} = \widehat{X}^-_{0|n}$ two times in a row. We see that in one case the state magnitude decays and in the other it increases.

Lemma 6.7: If $\widehat{X}^{\text{MAP}}_{0|n-1} = \widehat{X}^+_{0|n-1}$ and $\widehat{X}^{\text{MAP}}_{0|n} = \widehat{X}^+_{0|n}$, then $X_{n+1} = Q_n X_n$ where $Q_n \sim \text{Unif}[0,1]$. If instead $\widehat{X}^{\text{MAP}}_{0|n-1} = \widehat{X}^-_{0|n-1}$ and $\widehat{X}^{\text{MAP}}_{0|n} = \widehat{X}^-_{0|n}$, then $X_{n+1} = Q_n X_n$ where $Q_n \in [1, \frac{2}{1-s}]$.

This lemma helps us show that $\widehat{X}^{\text{MAP}}_{0|n} \to \widehat{X}^+_{0|n}$ a.s. and then that $\widehat{X}^+_{0|n} \to X_0$ a.s..

Lemma 6.8: $\widehat{X}^{\text{MAP}}_{0|n} \to \widehat{X}^+_{0|n}$ a.s. as $n \to \infty$.
The proof of Lemma 6.8 proceeds by showing that $\widehat{X}^{\text{MAP}}_{0|n}$ cannot equal $\widehat{X}^-_{0|n}$ infinitely many times, since every time this happens we will strictly decrease the negative candidate $\widehat{X}^-_{0|n}$, and thus it must eventually cross -1. Hence $\widehat{X}^{\text{MAP}}_{0|n}$ must converge to $\widehat{X}^+_{0|n}$.

Lemma 6.9: $\widehat{X}^+_{0|n} \to X_0$ a.s. as $n \to \infty$.
The core idea behind Lemma 6.9 is the monotone convergence theorem. We show that X_0 must be the essential supremum of $\widehat{X}^+_{0|n}$. Thus the sequence must converge to X_0.

Combining Lemmas 6.8 and 6.9, we obtain that $\widehat{X}^{\text{MAP}}_{0|n} \to X_0$ a.s. as $n \to \infty$ for the case $X_0 > 0$. Finally, we compute the η-rate for $0 \leq s < 1$.

Lemma 6.10: For (3) with MAP control and $0 \leq s < 1$,

$$\lim_{n\to\infty} -\frac{1}{n\eta} \log \mathbb{E}\left[\left|\frac{X_n}{X_0}\right|^\eta\right] = \frac{1}{\eta}\log(1+\eta).$$

The proof of Lemma 6.10 upper and lower bounds the rate by considering the worst case and best case performance during the X_0 sign estimation period, and showing that as $n \to \infty$ both bounds converge to the same value.

VII. Conclusion and Future Work

To the best of our knowledge, we provide the first known strategy with a provable convergence rate for control of the system in (3). We conjecture that the linear strategy is optimal in the case of high SNR.

VIII. Acknowledgments

Thanks to the NSF for CAREER Grant ECCS-2240031.

References

[1] R. W. Newcomb, "Active integrated circuit synthesis," Stanford University, Department of Electrical Engineering, Tech. Rep., 1969.

[2] L. Strauss, "Wave generation and shaping." *Wave generation and shaping*, 1970.

[3] G. Murphy and S. Wu, "A stability criterion for pulse-width-modulated feedback control systems," *IEEE Transactions on Automatic Control*, vol. 9, no. 4, pp. 434–441, 1964.

[4] M. Barr, "Pulse width modulation," *Embedded Systems Programming*, vol. 14, no. 10, pp. 103–104, 2001.

[5] J. Holtz, "Pulsewidth modulation — a survey," *IEEE Transactions on Industrial Electronics*, vol. 39, no. 5, pp. 410–420, 1992.

[6] A. S. Willsky and S. I. Marcus, "Analysis of bilinear noise models in circuits and devices," *Journal of the Franklin Institute*, vol. 301, no. 1-2, pp. 103–122, 1976.

[7] D. J. Webb, K. L. Crandall, and J. van den Berg, "Online parameter estimation via real-time replanning of continuous Gaussian POMDPs," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 5998–6005. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/6907743/

[8] P. Hingwe, H.-S. Tan, A. K. Packard, and M. Tomizuka, "Linear parameter varying controller for automated lane guidance: experimental study on tractor-trailers," *Control Systems Technology, IEEE Transactions on.*, vol. 10, no. 6, pp. 793–806, 2002.

[9] E. Todorov and W. Li, "A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems," in *Proceedings of the 2005, American Control Conference, 2005.* IEEE, 2005, pp. 300–306.

[10] E. Todorov and M. Jordan, "A minimal intervention principle for coordinated movement," *Advances in neural information processing systems*, vol. 15, 2002.

[11] E. Todorov and M. I. Jordan, "Optimal feedback control as a theory of motor coordination," *Nature neuroscience*, vol. 5, no. 11, pp. 1226–1235, 2002.

[12] K. Zhou and J. C. Doyle, *Essentials of robust control*. Prentice hall Upper Saddle River, NJ, 1998, vol. 104.

[13] W. De Koning, "Optimal estimation of linear discrete-time systems with stochastic parameters," *Automatica.*, vol. 20, no. 1, pp. 113–115, 1984.

[14] T. Hayakawa, H. Ishii, and K. Tsumura, "Adaptive quantized control for linear uncertain discrete-time systems," *Automatica*, vol. 45, no. 3, pp. 692–700, 2009.

[15] V. N. Phat, J. Jiang, A. V. Savkin, and I. R. Petersen, "Robust stabilization of linear uncertain discrete-time systems via a limited capacity communication channel," *Systems & control letters.*, vol. 53, no. 5, pp. 347–360, 2004.

[16] L. Xie, Y. C. Soh, and C. E. de Souza, "Robust Kalman filtering for uncertain discrete-time systems," *Automatic Control, IEEE Transactions on.*, vol. 39, no. 6, pp. 1310–1314, 1994.

[17] A. B. Kurzhanski and P. Varaiya, "On reachability under uncertainty," *SIAM Journal on Control and Optimization.*, vol. 41, no. 1, pp. 181–216, 2002.

[18] M. Fu and L. Xie, "Quantized feedback control for linear uncertain systems," *International Journal of Robust and Nonlinear Control.*, vol. 20, no. 8, pp. 843–857, 2010.

[19] C.-J. Seo and B. K. Kim, "Robust and reliable H-infinity control for linear systems with parameter uncertainty and actuator failure," *Automatica*, vol. 32, no. 3, pp. 465–467, 1996. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0005109895001492

[20] P. Gahinet, P. Apkarian, and M. Chilali, "Affine parameter-dependent Lyapunov functions and real parametric uncertainty," *IEEE Transactions on Automatic control*, vol. 41, no. 3, pp. 436–442, 1996. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/486646/

[21] L. Xie, “Output feedback H-infinity control of systems with parameter uncertainty,” *International Journal of control*, vol. 63, no. 4, pp. 741–750, 1996. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/00207179608921866

[22] L. Xie, M. Fu, and C. E. de Souza, “H-infinity control and quadratic stabilization of systems with parameter uncertainty via output feedback,” *IEEE Transactions on Automatic Control*, vol. 37, no. 8, pp. 1253–1256, 1992. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/151120/

[23] G. Ranade and A. Sahai, “Control capacity,” *IEEE Transactions on Information Theory*, vol. 65, no. 1, pp. 235–254, 2018.

[24] E. Gershon, U. Shaked, and I. Yaesh, “H-infinity control and filtering of discrete-time stochastic systems with multiplicative noise,” *Automatica.*, vol. 37, no. 3, pp. 409–417, 2001.

[25] E. Gershon, D. Limebeer, U. Shaked, and I. Yaesh, “Robust H-∞ filtering of stationary continuous-time linear systems with stochastic uncertainties,” *IEEE Transactions on Automatic Control*, vol. 46, no. 11, pp. 1788–1793, 2001.

[26] E. Gershon, U. Shaked, and I. Yaesh, “Robust H-∞ estimation of stationary discrete-time linear processes with stochastic uncertainties,” *Systems & control letters*, vol. 45, no. 4, pp. 257–269, 2002.

[27] E. Gershon and U. Shaked, “Static H-2 and H-infinity output-feedback of discrete-time LTI systems with state multiplicative noise,” *Systems & Control Letters.*, vol. 55, no. 3, pp. 232–239, 2006.

[28] J. S. Shamma, “An overview of lpv systems,” *Control of linear parameter varying systems with applications*, pp. 3–26, 2012.

[29] F. Wu, *Control of linear parameter varying systems*. University of California, Berkeley, 1995.

[30] W. M. Wonham, “Optimal stationary control of a linear system with state-dependent noise,” *SIAM Journal on Control*, vol. 5, no. 3, pp. 486–500, 1967.

[31] J. L. Willems and J. C. Willems, “Feedback stabilizability for stochastic systems with state and control dependent noise,” *Automatica*, vol. 12, no. 3, pp. 277–283, 1976.

[32] U. Haussmann, “Optimal stationary control with state control dependent noise,” *SIAM Journal on Control*, vol. 9, no. 2, pp. 184–198, 1971.

[33] L. El Ghaoui, “State-feedback control of systems with multiplicative noise via linear matrix inequalities,” *Systems & Control Letters*, vol. 24, no. 3, pp. 223–228, 1995.

[34] N. Elia, “Remote stabilization over fading channels,” *Systems & Control Letters*, vol. 54, no. 3, pp. 237–249, 2005.

[35] M. Athans, R. Ku, and S. Gershwin, “The uncertainty threshold principle: Some fundamental limitations of optimal decision making under dynamic uncertainty,” *IEEE Transactions on Automatic Control*, vol. 22, no. 3, pp. 491–495, 1977.

[36] R. Ku, M. Athans, and P. Varaiya, “The Effects of Discounted Cost on the Uncertainty Threshold Principle,” in *Annals of Economic and Social Measurement, Volume 6, number 5*, 1977, pp. 11–18.

[37] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, “Kalman filtering with intermittent observations,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[38] S. Park and A. Sahai, “Intermittent Kalman filtering: Eigenvalue cycles and nonuniform sampling,” in *American Control Conference (ACC).*, 2011, pp. 3692–3697.

[39] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, “Foundations of control and estimation over lossy networks,” *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, 2007.

[40] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, “A survey of recent results in networked control systems,” *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138–162, 2007.

[41] J. Ding, Y. Peres, G. Ranade, and A. Zhai, “When multiplicative noise stymies control,” *Annals of Applied Probability*, 2018.

[42] V. Subramanian, M. Won, and G. Ranade, “Learning a neural-network controller for a multiplicative observation noise system,” in *2020 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2020, pp. 2849–2854.

[43] B. Gravell and T. Summers, “Robust learning-based control via bootstrapped multiplicative noise,” in *Learning for Dynamics and Control*. PMLR, 2020, pp. 599–607.

[44] B. Gravell, P. M. Esfahani, and T. Summers, “Learning optimal controllers for linear systems with multiplicative noise via policy gradient,” *IEEE Transactions on Automatic Control*, vol. 66, no. 11, pp. 5283–5298, 2020.

[45] Y. Xing, B. Gravell, X. He, K. H. Johansson, and T. Summers, “Linear system identification under multiplicative noise from multiple trajectory data,” in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 5157–5261.

[46] B. Gravell, M. Gargiani, J. Lygeros, and T. H. Summers, “Policy iteration for multiplicative noise output feedback control,” *arXiv preprint arXiv:2203.17165*, 2022.

[47] P. Coppens and P. Patrinos, “Safe learning LQR of linear dynamics with multiplicative noise,” *arXiv e-prints*, pp. arXiv–2207, 2022.

[48] P. Coppens, M. Schuurmans, and P. Patrinos, “Data-driven distributionally robust LQR with multiplicative noise,” in *Learning for Dynamics and Control*. PMLR, 2020, pp. 521–530.

[49] H. Witsenhausen, “A counterexample in stochastic optimum control,” *SIAM Journal on Control.*, vol. 6, p. 131, 1968.

[50] S. Yuksel and T. Basar, “Stochastic networked control systems,” 2013.

[51] S. Fang, J. Chen, and H. Ishii, *Towards Integrating Control and Information Theories*. Springer, 2016.

[52] W. S. Wong and R. W. Brockett, “Systems with finite communication bandwidth constraints I: State estimation problems,” *IEEE Transactions on Automatic Control*, vol. 42, no. 9, pp. 1294–1299, 1997.

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

[54] G. Nair and R. Evans, “Stabilizability of stochastic linear systems with finite feedback data rates,” *SIAM Journal on Control and Optimization.*, vol. 43, no. 2, pp. 413–436, 2004.

[55] G. N. Nair, F. Fagnani, S. Zampieri, and R. J. Evans, “Feedback control under data rate constraints: An overview,” *Proceedings of the IEEE*, vol. 95, no. 1, pp. 108–137, 2007.

[56] 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, pp. 243–255, 2009.

[57] K. You and L. Xie, “Minimum data rate for mean square stabilizability of linear systems with Markovian packet losses,” *IEEE Transactions on Automatic Control*, vol. 56, no. 4, pp. 772–785, 2011.

[58] P. Minero, L. Coviello, and M. Franceschetti, “Stabilization over Markov feedback channels: the general case,” *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 349–362, 2013.

[59] V. Kostina, Y. Peres, G. Ranade, and M. Sellke, “Exact minimum number of bits to stabilize a linear system,” *IEEE Transactions on Automatic Control*, 2021.

[60] J. S. Freudenberg, R. H. Middleton, and J. H. Braslavsky, “Minimum variance control over a Gaussian communication channel,” *IEEE Transactions on Automatic Control*, vol. 56, no. 8, pp. 1751–1765, 2011.

[61] L. Bao, M. Skoglund, and K. H. Johansson, “Iterative encoder-controller design for feedback control over noisy channels,” *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 265–278, 2011.

[62] A. S. Matveev and A. V. Savkin, *Estimation and control over communication networks*. Springer Science & Business Media, 2009.

[63] A. Sahai and S. Mitter, “The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link part I: scalar systems,” *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3369–3395, 2006.

[64] J. H. Braslavsky, R. H. Middleton, and J. S. Freudenberg, “Feedback stabilization over signal-to-noise ratio constrained channels,” *IEEE Transactions on Automatic Control*, vol. 52, no. 8, pp. 1391–1403, 2007.

[65] E. I. Silva, G. C. Goodwin, and D. E. Quevedo, “Control system design subject to SNR constraints,” *Automatica*, vol. 46, no. 2, pp. 428–436, 2010.

[66] E. I. Silva and S. A. Pulgar, “Performance limitations for single-input LTI plants controlled over SNR constrained channels with feedback,” *Automatica*, vol. 49, no. 2, pp. 540–547, 2013.

[67] A. Lapidoth and S. Moser, “Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels,” *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2426–2467, 2003.

[68] G. Ranade and A. Sahai, “Non-coherence in estimation and control,” in *Communication, Control, and Computing, 51st Annual Allerton Conference on*, 2013.

[69] ——, “Control Capacity,” in *IEEE International Symposium on Information Theory*. IEEE, 2015.