Abstract—This paper considers the stabilization of an unstable discrete-time linear system that is observed over a channel corrupted by continuous multiplicative noise. The main result is a converse bound that shows that if the system growth is large enough the system cannot be stabilized in a mean-squared sense. This is done by showing that the probability of the state magnitude remains bounded must go to zero with time.

It was known that a system with multiplicative observation noise can be stabilized using a simple linear strategy if the system growth is suitably bounded. However, it was not clear whether non-linear controllers could overcome arbitrarily large growth factors. One difficulty with using the standard approach for a data-rate theorem style converse is that the mutual information per round between the system state and the observation is potentially unbounded with a multiplicative noise observation channel. Our proof technique recursively bounds the conditional density of the system state (instead of focusing on the second moment) to bound the progress the controller can make.

I. INTRODUCTION

We consider the control of a system observed over a multiplicative noise channel, and provide a converse bound to define a region in which the system cannot be stabilized in the mean-square sense. Multiplicative observation noise can model the effects of fast-fading communication channel or sampling and quantization errors [1], [2]. A more detailed discussion of multiplicative noise models is available in [3].

Specifically, we analyze the following system (Fig. 1):

\[ X_{a,n+1} = a \cdot X_{a,n} - U_{a,n}, \]
\[ Y_{a,n} = Z_{a,n} \cdot X_{a,n}. \]

(1)

In the preceding formulation, the system state is represented by \( X_{a,n} \) at time \( n \), and the control \( U_{a,n} \) can be any function of the current and previous observations \( Y_{a,0} \) to \( Y_{a,n} \). The \( Z_{a,n} \)'s are i.i.d. random variables with a known continuous distribution. The realization of the noise \( Z_{a,n} \) is unknown to the controller, much like the fading coefficient of a channel might be unknown to the transmitter or receiver in non-coherent communication. The controller’s objective is to stabilize the system in a mean-square sense, i.e. ensure that \( \sup_n \mathbb{E}[|X_{a,n}|^2] < \infty \).

For a simple achievable strategy, suppose the \( Z_{a,n} \)'s are i.i.d. with mean 1 and variance \( \sigma^2 \). It is easy to verify that a simple memoryless linear strategy can mean-square stabilize the system (1) if and only if \( a^2 \leq (1 + \frac{1}{\sigma^2}) \).

A. Related work

This fundamental problem is inspired by many issues that have been long studied in control and we only mention some representative references here. It is connected to the vast body of work on data-rate theorems and control with communication constraints such as [4], [5]. Nair et al. provide a survey of related problems in [6]. For a system that is growing by a factor of \( a \) at each time, the data-rate theorems tell us that a noiseless observation data rate \( R > \log |a| \) is necessary and sufficient to stabilize the system [4].

A related problem is that of estimating a linear system over multiplicative noise. While early work on this had been limited to exploring linear estimation strategies [7], some more recent work used a genie-based approach to show a general converse result for the estimation problem over multiplicative noise for both linear and non-linear strategies [2]. The control problem in this paper can also be interpreted as an “active” estimation problem for \( X_0 \), and our converse (that also uses a side-information genie) applies to both linear and non-linear control strategies. However, techniques from the estimation converse or the data-rate theorems do not work for a converse here. Unlike the estimation problem, we cannot describe the distribution of \( X_n \) in our problem since the control \( U_n \) is arbitrary. For the same reason, we also cannot bound the range of \( X_n \), and we cannot bound the rate across the channel to use a data-rate theorem approach.

Another inspiration for our problem is the work on intermittent Kalman filtering [8], [9] and control over packet dropping networks [10] (i.e. estimation and control over Bernoulli multiplicative noise channels). The setup in this paper generalizes those setups to consider a general continuous multiplicative noise on the observation.
Finally, the result in this paper can also be thought of as extending classic uncertainty threshold principle [11] to also understand multiplicative observation uncertainty. The uncertainty threshold principle provides a threshold for the stabilizability of a system with Gaussian uncertainty on the system gain and the control gain.

We would eventually like to establish the control capacity of the system (1), as done for systems with multiplicative noise on the actuation channel (control gain) in [12], [3]. If the controller for a system with multiplicative observation noise is restricted to using linear control strategies, then its performance limit is the same as that of system with multiplicative actuation noise (as in [12]). However, the approach from [12] does not seem to work for the system considered here.

B. Proof approach

This paper introduces a non-standard converse approach and we believe these techniques are a primary contribution of the work. Instead of focusing on the second-moment, our proof bounds the density of the state and thus shows the instability of any moment of the state.

A key element of the proof is that a “genie” observes the state of the system and provides a quantized version of the logarithm of the state to the controller at each time as extra side-information in addition to the multiplicative noise observation. This side information bounds the state in intervals of size $2^{-k}$ around the initial state $X_0$ (with $k$ increasing as time increases). We know from results on non-coherent communication [13] and carry-free models [14] that only the order of magnitude of the message can be recovered from a transmission with multiplicative noise. As a result, this side-information does not effectively provide much extra information, but it allows us to quantify the rate at which the controller may make progress.

This paper focuses on a Gaussian distribution for the multiplicative noise, but the ideas can be extended to general distributions under mild assumptions.

II. PROBLEM STATEMENT AND SETUP

Consider the system $S_a$ in (1). For simplicity, let the initial state $X_{a,0}$ be distributed as $X_{a,0} \sim \mathcal{N}(0,1)$. Let $Z_n$ be i.i.d. Gaussian random variables such that $Z_n \sim \mathcal{N}(1, \sigma^2)$, $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-1)^2}{2\sigma^2}}$. Without loss of generality we can assume that the $Z_n$ have mean 1.

In this paper we show that there exists $a > 0$ such that $P(|X_{n,a}| < M) \to 0$ as $n \to \infty$ for all $M$. This implies that for the same $a$, the second moment $\mathbb{E}[X_{n,a}^2]$ is always unbounded, i.e. the system cannot be stabilized in the second-moment sense.

To obtain this threshold it suffices to consider the system, $S$ (i.e. set $a = 1$.) The distributions on the $Z_n$ variables and the initial state $X_0$ are the same as system $S_a$.

\begin{align}
X_{n+1} &= X_n - U_n, \\
Y_n &= Z_n \cdot X_n.
\end{align}

The system $S_a$ behaves as $S$ scaled by $a^n$ when the control $U_{a,n} = a^n U_n$ is applied. A similar result is given in [3, Ch. 5, Lemma 5.4.4]. Hence, we can write:

$$P(|X_{n,a}| < M) = P(|X_n| < a^{-n}M).$$

A. Notation and definitions

Let $S_n = \sum_{i=0}^{n-1} U_i$ for $n \geq 1$. Hence, $X_n = X_0 - S_n$.

The goal of the controller is to have $S_n$ be as close to $X_0$ as possible. We will track the progress of the controller through intervals $I_n$ that contain $X_0$ and are decreasing in length.

Let $d(I_n, S) := \inf_{x \in I_n} |S - x|$ denote the distance of a point $S$ from the interval $I_n$.

Definition 2.1: For all $n \geq 0$ and for any $k \in \mathbb{Z}$, there exists a unique integer $h(k)$ such that $X_0 \in [\frac{h(k)}{2^n}, \frac{h(k)+1}{2^n})$.

Let $J(k) := [\frac{h(k)}{2^n}, \frac{h(k)+1}{2^n})$. We now inductively define the intervals $I_n$ using the intervals $J(k)$. Define

$$K_0 := \min\{k \geq 0 \mid d(J(k), 0) \geq 2^{-k}\}, \quad K_n := \min\{k \mid k > K_{n-1}, d(J(k), S_n) \geq 2^{-k}\}.$$

Write $H_n := h(K_n)$ and $I_n := [\frac{H_n}{2^n}, \frac{H_n+1}{2^n})$.

Let $Y_0^+$ indicate the observations $Y_0$ to $Y_n$, and let $F_n := \{Y_0^+, X_0^+, H_n^+\}$, which is the total information available to the controller at time $n$. Let $f_{X_n}(x|F_n)$ be the conditional density of $X_n$ given $F_n$.

\[ I_n \]

\[ 0 \quad S_n \quad \geq 2^{-K_n} \quad X_0 \quad 2^{-K_n} \]

\[ I_n - S_n \]

\[ 0 \quad S_n \quad \geq 2^{-K_n} \quad X_0 \quad 2^{-K_n} \]

Fig. 2. A caricature illustrating the intervals $I_n$ and $I_n - S_n$.

B. Side-information lemma

This lemma uses $K_n$ to bound how fast $S_n$ approaches $X_0$.

The easy proof is defered to the Appendix.

Lemma 2.1: $2^{-K_n} \leq |X_0 - S_n|$, and if $K_n > K_{n-1} + 1$, then $|X_0 - S_n| \leq 2^{2-K_n}$.

III. CONVERSE RESULT

Theorem 3.1: There exists $a \in \mathbb{R}, 0 < a < \infty$ such that $P(|X_{a,n}| < M) \to 0$ for all $M < \infty$.

Strategy: The proof provides an exponential upper bound on the density $f_{X_n}(x|F_n)$, which implies that the controller cannot localize the state within an exponentially shrinking box. The bound on the density is obtained by estimating the change in the density from time $n$ to $n+1$ due to the observation $Y_n$. The interval $I_n$ (i.e. $H_n$ and $K_n$) is provided to the controller as side information at time $n$, which helps generate the bound.

The first step of the proof uses Lemma 3.2 to recursively bound the ratio $f_{X_{n+1}}(x|F_{n+1})$ by $f_{X_n}(x|F_n)$. This leads to an exponential bound
on the density $f_{X_n}(x \mid F_n)$ in terms of the side-information $K_n$ in (8).

This bound helps us generate bounds on the probability the event of interest $\{|X_n| < a^{-n}M\}$, which we show then must go to 0 as $n \to \infty$.

**Proof:** Consider

$$f_{X_n}(x \mid F_n) = f_{X_n}(x \mid Y_n, K_n, H_n, F_{n-1}) = \frac{f_{Y_n,K_n,H_n}(Y_n, K_n, H_n \mid X_n = x, F_{n-1}) \cdot f_{X_n}(x \mid F_{n-1})}{f_{Y_n,K_n,H_n}(Y_n, K_n, H_n \mid F_{n-1})}.$$

Formally, $f_{Y_n,K_n,H_n}$ is a density with respect to a product of the Lebesgue measure and two counting measures.

Since $X_0 \in I_n$, the controller knows that $X_n \in I_n - S_n$, where $I_n - S_n$ represents the interval $I_n$ shifted by $S_n$. We can calculate the ratio of the densities at $x, w \in I_n - S_n$ as:

$$\frac{f_{X_n}(x \mid F_n)}{f_{X_n}(w \mid F_n)} = \frac{f_{X_n}(x \mid F_{n-1}) f_{Y_n}(Y_n \mid X_n = x, F_{n-1})}{f_{X_n}(w \mid F_{n-1}) f_{Y_n}(Y_n \mid X_n = w, F_{n-1})}.$$  (3)

Since $K_n$ and $H_n$ are defined by $I_n$, the conditional distributions of $K_n$ and $H_n$ given $X_n = x$ and $X_n = w$ are equal for $x, w \in I_n - S_n$. So these terms cancel when we consider a ratio, giving (3).

Taking logarithms and using the triangle inequality gives the following recursive lemma, which we prove in the Appendix.

**Lemma 3.2:**

$$\left| \log \frac{f_{X_n}(x \mid F_n)}{f_{X_n}(w \mid F_n)} \right| \leq \frac{1}{\sigma^2} (\sigma^2 + 2Z_n(2Z_n - 1)) 2^{K_n} |x - w| + \left| \log \frac{f_{X_n}(x \mid F_{n-1})}{f_{X_n}(w \mid F_{n-1})} \right|.$$  (4)

Now, based on the control law we know that $f_{X_n}(x \mid F_{n-1}) = f_{X_n}(x(U_{n-1} - 1) \mid F_{n-1})$, since $U_{n-1}$ is $F_{n-1}$ measurable. Substituting this into (4) and unfolding recursively gives:

$$\left| \log \frac{f_{X_n}(x \mid F_n)}{f_{X_n}(w \mid F_n)} \right| \leq \frac{n}{\sigma^2} \left( \sigma^2 + 2Z_n(2Z_n - 1) \right) 2^{K_n} |x - w| + \left| \log \frac{f_{X_n}(x + S_n)}{f_{X_n}(w + S_n)} \right|.$$  (5)

The inequality (5) separates the effect of the uncertainty due to $X_0$ and the subsequent uncertainty due to the observations and control.

To bound the effect of the initial state let $\eta_n = \max_{x,w \in I_n - S_n} \left| \log \frac{f_{X_0}(x+S_n)}{f_{X_0}(w+S_n)} \right|$. Since $I_n$ is an interval of size at most $2^{-n}$ that contains $X_0$ we get that:

$$|\eta_n| \leq \frac{1}{2} ((X_0 + 2^{-n})^2 - (X_0 - 2^{-n})^2) \leq 2^{1-n} |X_0|.$$  (6)

Now, we define

$$\Psi_n = 2^{-K_n} \frac{1}{\sigma^2} \sum_{i=1}^{n} \left( \sigma^2 + 2Z_i(2Z_i - 1) \right) 2^{K_i},$$  (7)

and rewrite (5) as:

$$\left| \log \frac{f_{X_n}(x \mid F_n)}{f_{X_n}(w \mid F_n)} \right| \leq \Psi_n \cdot 2^{K_n} |x - w| + \eta_n.$$  (8)

Finally, we are in a position to bound $f_{X_n}(x \mid F_n)$:

$$f_{X_n}(x \mid F_n) \leq \exp\{\Psi_n 2^{K_n} |x - w| + \eta_n\} f_{X_n}(w \mid F_n).$$  (9)

We integrate (8) over an interval of length $\gamma = 2^{(-K_n+T)}$ with $x$ at one end point. So $|x - w| \leq 2^{(-K_n-T)}$. Here, $T \in \mathbb{R}$ is a constant that we will choose later. Such an interval can be fit into $I_n$ to the left or right of any $x$ depending on where $x$ is in the interval. Assuming without loss of generality that $x$ is the left endpoint of the integration interval we compute that

$$\int_x^{x+\gamma} f_{X_n}(x\mid F_n) \, dw \leq \int_x^{x+\gamma} \exp\{\Psi_n 2^{K_n} |x-w| + \eta_n\} f_{X_n}(w\mid F_n) \, dw$$

Bound $|x - w|$ on the RHS by $2^{(-K_n-T)}$. Integrating gives

$$\gamma \cdot f_{X_n}(x \mid F_n) \leq \int_x^{x+\gamma} \exp\{\Psi_n 2^{K_n} |x-w| + \eta_n\} f_{X_n}(w\mid F_n) \, dw$$

$$\leq \exp\{\Psi_n 2^{T-\gamma} + \eta_n\} \cdot 1,$$

since the density integrates out to 1. Hence,

$$f_{X_n}(x \mid F_n) \leq \exp\{\Psi_n 2^{T-\gamma} + \eta_n\} 2^{K_n+T}.$$  (10)

This gives us a bound on the density of $X_n$ in terms of $K_n$. Since the $K_n$’s are tracking the magnitude of $X_n$ what remains to be done is to bound the growth of the $K_n$’s. To do this we first state the following lemma about the crucial quantity $\Psi_n$.

**Lemma 3.3:** For $T > C_\sigma$, where $C_\sigma$ is a constant that depends only on $\sigma$, the expectation $\mathbb{E}[\Psi_n 2^{T}]$ is uniformly bounded for all $n$.

The next lemma uses Lemma 3.3 and (9) to show that the $K_n$ grow at most linearly.

**Lemma 3.4:** There exists a constant $C = C_{\sigma,T}$ depending only on $(\sigma, T)$ such that

$$\mathbb{P}(K_n - K_0 > C_{\sigma,T} \cdot n) \to 0.$$  (11)

The proofs of Lemmas 3.3, 3.4 are deferred to the Appendix.

Let $G_n$ denote the event that $K_n - K_0 > C_n$, and $G_n^c$ its complement. Then the event $\{|X_n| < a^{-n}M\}$ is covered by:

$$\mathbb{P}(|X_n| < a^{-n}M) \leq \mathbb{P}(G_n) + \mathbb{P}(K_0 > n) + \mathbb{P}(|X_n| \leq a^{-n}M, C_n, K_0 \leq n).$$  (12)

We evaluate the three terms one by one. For the first term in (12), we have $\mathbb{P}(G_n) = \mathbb{P}(K_n - K_0 > C_n) \to 0$ as $n \to \infty$ from Lemma 3.4.

The second term, $\mathbb{P}(K_0 > n)$, captures the case where the initial state $X_0$ might be very close to zero. However, eventually this advantage dies out for large enough $n$, since $\mathbb{P}(X_0 < 2^{-n}) \to 0$ as $n \to \infty$. 

The proofs of Lemmas 3.3, 3.4 are deferred to the Appendix.
The last term in (10) remains. By the law of iterated expectation:
\[
P(|X_n| < a^{-n}M, G_n, K_0 \leq n) = E[P(|X_n| < a^{-n}M, G_n, K_0 \leq n \mid F_n)].
\]
We focus on the term conditioned on \(F_n\):
\[
P(|X_n| < a^{-n}M, G_n, K_0 \leq n \mid F_n)
= E[I\{|X_n| < a^{-n}M\} \mid G_n, K_0 \leq n \mid F_n]
= P(|X_n| < a^{-n}M \mid F_n) \cdot I\{G_n\} \mid K_0 \leq n\}.
\]
(11)

Now, we can apply (9) to get
\[
P(|X_n| < a^{-n}M \mid F_n) = \int_{-a^{-n}M}^{a^{-n}M} f_{X_n}(x \mid F_n) \, dx
\leq 2M \cdot e^{-\frac{n}{a^{-n}M}} \cdot |F_n|^{-T} \cdot 2K_n + T.
\]
Then we can bound (11) as
\[
P(|X_n| < a^{-n}M \mid F_n) \cdot I\{G_n\} \mid K_0 \leq n
\leq 2M \cdot e^{-\frac{n}{a^{-n}M}} \cdot |G_n|^{-T} \cdot 2(C+1)n + T
\]
(12)
since \(K_n \leq Cn + K_0\) and \(K_0 \leq n\) implies \(K_n \leq (C+1)n\).

Taking expectations on both sides we get:
\[
P(|X_n| < a^{-n}M, G_n, K_0 \leq n)
\leq 2M \cdot e^{-\frac{n}{a^{-n}M}} \cdot |G_n|^{-T} \cdot E[I\{G_n\} \mid K_0 \leq n].
\]
(13)

By Lemma 3.3 and (6), the above expression (13) tends to 0 for \(a > 2^{C+1}\). Thus, all three probabilities in (10) converge to 0 as \(n \to \infty\). Hence, if \(a > 2^{C+1}\) then \(P(|X_n| < a^{-n}M) \to 0\) for all \(M\).

IV. CONCLUSION AND FUTURE WORK

This paper provides a first-proof-of-concept converse for a control system observed over continuous multiplicative noise. However, there is an exponential gap between the scaling behavior of the achievable strategy and the converse. It still remains to be seen if the achievable region can grow with non-linear controllers to significantly beat the \(a^2 < (1 + \frac{1}{2})\) bound, and computing the control capacity of the system remains open.

ACKNOWLEDGEMENTS

The authors would like to thank Anant Sahai for initial discussions regarding this problem.

REFERENCES


APPENDIX

A. Proof of Lemma 2.1

From the definition of \(I_n\), we know that \(d(I_n, S_n) \geq 2^{-K_n}\). Hence, \(I_n\) cannot contain \(S_n\). This gives \(2^{-K_n} \leq |X_0 - S_n|\), since \(X_0 \in I_n\).

To show the second half of the inequality, suppose that \(|X_0 - S_n| > 2^{-K_n}\). Then,
\[
2^{1-K_n} < |X_0 - S_n| - 2^{1-K_n}.
\]
Hence, there exists a smaller interval \(J(K_n - 1)\) that contains \(X_0\) such that
\[
2^{1-K_n} < d(J(K_n - 1), S_n),
\]
where \(J(K_n - 1)\) is an interval of length \(2^{1-K_n} < 2^{-K_n}\). Since we also assumed that \(K_n > K_{n-1} + 1\), this contradicts the assumption that \(K_n\) was the minimal \(k > K_{n-1}\) such that \(d(J(k), S_n) \geq 2^{-k}\).

B. Proof of Lemma 3.2

We take logarithms on both sides of (3) and apply the triangle inequality to get
\[
\log \left( \frac{f_{X_n}(x \mid F_n)}{f_{X_n}(w \mid F_n)} \right) \leq \log \left( \frac{f_{Y_n}(Y_n \mid X_n = x, F_{n-1})}{f_{Y_n}(Y_n \mid X_n = w, F_{n-1})} \right) + \log \left( \frac{f_{X_n}(x \mid F_{n-1})}{f_{X_n}(w \mid F_{n-1})} \right). \tag{14}
\]
Using $f_{Y_n}(Y_n|X_n = x, F_{n-1}) = \frac{1}{\sqrt{2\pi} \sigma_n} e^{-(\lambda_n - 1)^2 \frac{z^2}{2\sigma_n^2}}$, and the applying the triangle inequality again we get,

$$\log \left| f_{Y_n}(Y_n|X_n = x, F_{n-1}) \right| \leq \log \left| x \right| - \log \left| w \right| + \frac{1}{2\sigma^2} \left( \frac{Y_n}{x} - 1 \right)^2 - \left( \frac{Y_n}{w} - 1 \right)^2.$$  

(15)

We can use the derivatives of the functions to bound the two function differences above. Using $\frac{d}{dx} \log |x| = \frac{1}{x}$ and $\frac{d}{dx} \left( \frac{Y_n}{x} - 1 \right)^2 = 2 \left( \frac{Y_n}{x} - 1 \right) \left( \frac{1}{x^2} \right)$, we can bound (15) as:

$$\log \left| f_{Y_n}(Y_n|X_n = x, F_{n-1}) \right| \leq \left( \max_{t \in I_n - S_n} \frac{1}{|t|} \right) |x-w| + \left( \max_{t \in I_n - S_n} \frac{1}{2\sigma^2} \left( \frac{X_nZ_n}{t} - 1 \right) \left( \frac{X_nZ_n}{t^2} \right) \right) |x-w|.$$  

(16)

Since $X_n \in I_n - S_n$, the maximizations are over $t \in I_n - S_n$ in (16). First, note that $\max_{t \in I_n - S_n} \frac{1}{|t|} \leq 2K_n$. Second, for all $t \in I_n - S_n$, we have $\frac{1}{2} \leq \frac{1}{2|t|} \leq 2$. Hence, the second term on the RHS of (16) can upper bounded by:

$$\leq \left( \max_{t \in I_n - S_n} \frac{1}{2\sigma^2} \right)^2 |x-w|.$$  

Combining these two ideas we upper bound (16) as:

$$\leq 2K_n |x-w| + \frac{1}{2\sigma^2} (4Z_n(2Z_n - 1)2K_n)|x-w| = \left( 1 + \frac{2K_n (2Z_n - 1)}{\sigma^2} \right) 2K_n |x-w|.$$  

This implies the desired bound for (14).

C. Proof of Lemma 3.3

$$\mathbb{E}[e^{\psi_n 2^{-T}}] = \mathbb{E}[\exp\left\{ 2^{-T} - K_n \frac{1}{\sigma^2} \sum_{i=1}^{n} (\sigma^2 + 2Z_i(2Z_i - 1))2K_n \right\}].$$

Since the $K_i$’s increase by at least 1 each time, we can replace $2^{-K_n}2K_i$ by $2^{i-n}$. We take absolute values and use the triangle inequality to get the upper bound

$$\mathbb{E}[e^{\psi_n 2^{-T}}] \leq \mathbb{E}[\exp\left\{ 2^{-T} - \frac{1}{\sigma^2} \sum_{i=1}^{n} |\sigma^2 + 2Z_i(2Z_i - 1)|2^{i-n} \right\}].$$

Substituting $Z_i$ as $Z_i = \sigma \tilde{Z}_i + 1$, noting that the $Z_i$ are independent and some simple algebra gives the bound below. $\alpha, \beta$ are constants that depend on $\sigma$.

$$\mathbb{E}[e^{\psi_n 2^{-T}}] \leq \prod_{j=1}^{n} \mathbb{E} \left\{ \exp \left\{ 2^{-T} (\alpha \tilde{Z}_j^2 + \beta) 2^{-j} \right\} \right\}. \tag{17}$$

For each term in the product of (17):

$$\mathbb{E} \left\{ \exp \left\{ 2^{-T} (\alpha \tilde{Z}_j^2 + \beta) 2^{-j} \right\} \right\} = \exp \left\{ \beta 2^{-T-j} \right\} \int \frac{1}{\sqrt{2\pi}} e^{-(\alpha 2^{-j} - 1 - \beta) z^2} \, dz.$$

Choose $T > \log_2(2\alpha)$, so that $\xi := \alpha 2^{-T} < \frac{1}{2}$. Note $\xi < \infty$. The integral is equal to $\frac{1}{\sqrt{\xi(1-\xi)2^{-2j}}}$. Then,

$$\prod_{j=1}^{n} \frac{1}{\sqrt{1-\xi^2 2^{-4j}}} \leq \prod_{j=1}^{n} e^{2\xi e^{2\xi} 2^{-4j}} \leq e^{4\xi}.$$  

A similar bound can be obtained for the terms with $\beta$. The details are omitted and will be included in the full version.

D. Proof of Lemma 3.4

By construction, $K_{n+1} \geq K_n + 1$. In the case where $K_{n+1} > K_n + 1$, we can apply Lemma 2.1 and get that for $\ell \geq 2$

$$\mathbb{P}(K_{n+1} - \ell | F_n) \leq \mathbb{P}(K_{n+1} - \ell | F_n) \leq \mathbb{P}(X_{n+1} | \leq 2^{2-K_n-\ell} | F_n) \leq \mathbb{P}(X_n - U_n | \leq 2^{2-K_n-\ell} | F_n).$$

This is because the control $U_n$ must have been very close to $X_n$ for $K_n+1$ to be much larger than $K_n$. Then we calculate this probability by integrating out the density.

$$\mathbb{P}(X_n - U_n | \leq 2^{2-K_n-\ell} | F_n) = \int_{U_n - 2^{2-K_n-\ell}}^{U_n} f_{X_n}(t | F_n)\, dt \leq 2 \cdot 2^{-2-K_n-\ell} \mathbb{E}[f_{X_n}(t | F_n)].$$

Combined with (9), this gives us that

$$\mathbb{P}(K_{n+1} - \ell \geq \ell | F_n) \leq 2 \cdot 2^{-2-K_n-\ell} \mathbb{E}[f_{X_n}(t | F_n)] = 2^{3-\ell} + 2^{2-K_n+T}.$$

(18)

Let $D_n = K_{n+1} - K_n$ and

$$\bar{K}_n = \sum_{i=0}^{n-1} (D_i - \mathbb{E}[D_i | F_i]).$$

It is clear that $(\bar{K}_n)$ is a martingale with respect to $F_n$. In addition, (18) yields that the conditional distribution of $D_n$ given $F_n$ is stochastically dominated by the distribution of

$$G + (3 + T + \left| \frac{\eta_n + \Psi_n 2^{-T}}{\ln 2} \right|),$$

where $G$ is an independent geometric variable with mean 2. Combined with (6) and (7), it yields that

$$\mathbb{E}(D_i - \mathbb{E}[D_i | F_i])^2 \leq ED_i^2 \leq C_1,$$

where $C_1$ is a constant depending on $(T, \sigma)$. This implies

$$\mathbb{P}(\bar{K}_n \geq n) \to 0 \text{ as } n \to \infty. \tag{20}$$

Now, using (19) again, we get that

$$\mathbb{E}[D_i | F_i] \leq 2(5 + T + |\eta_n + \Psi_i 2^{-T}|). \tag{21}$$

A careful analysis of the quantity $\Delta_n = |\eta_n + \Psi_n 2^{-T}|$ can be used to show that

$$\mathbb{P}\left( \sum_{i=1}^{n} E[D_i | F_i] \geq 10^6(1 + T + \sigma^{-2})n \to 0, \right.$$  

which combined with (20) proves the lemma. Details will be included in the full version. ■