Stochasticity in Feedback Loops

GREAT EXPECTATIONS AND GUARANTEED RUIN

ROY S. SMITH and BASSAM BAMIEH

Stochastic components in a feedback loop introduce state behaviors that are fundamentally different from those observed in a deterministic system. The effect of injecting a stochastic signal additively in linear feedback systems can be viewed as the addition of filtered stochastic noise. If the stochastic signal enters the feedback loop in a multiplicative manner, a much richer set of state behaviors emerges. These phenomena are investigated for the simplest possible system: a multiplicative noise in a scalar, integrating feedback loop. The same dynamics arise when considering a first-order system in feedback with a stochastic gain. The dynamics of this form arise naturally in a number of domains, including compound investments in finance, chemical reaction dynamics, population dynamics, epidemiology, control over lossy communication channels, and adaptive control. Understanding the nature of such dynamics in a simple system is a precursor to recognizing them in more complex stochastic dynamical systems.

Summary

Stochastic feedback, or multiplicative noise, leads to heavy-tailed state distributions in which the median, mean, and variance of the state can diverge. A detailed study of this phenomenon in simple systems leads to precise control-theoretic conditions for its occurrence and provides insights into the underlying mechanism.

STOCHASTIC FEEDBACK LOOPS

Results on the evolution of the statistics of multiplicative systems have appeared in other research domains in the past but are not widely known within the control systems community. The presentation and the proof of the results given depend on only reasonably well-known statistical techniques. This article draws upon and augments such results to study the stability of a stochastic feedback loop (see "Summary"). One of the earliest observations of the unusual phenomena described in this article was reported by Rosenbloom [1], who examined the solution of the stochastic first-order differential equation

$$\dot{x} + a(t)x = S(t),$$

where $a(t)$ is a Gaussian stochastic function and $S(t)$ is a step function. The solution itself can approach one in probability, and yet, both the mean and variance become infinite as $t \to \infty$. The case when the variance goes to zero as $t \to \infty$ is referred to as mean-square stability, which was analyzed for continuous-time stochastic differential equations by Samuels [2].

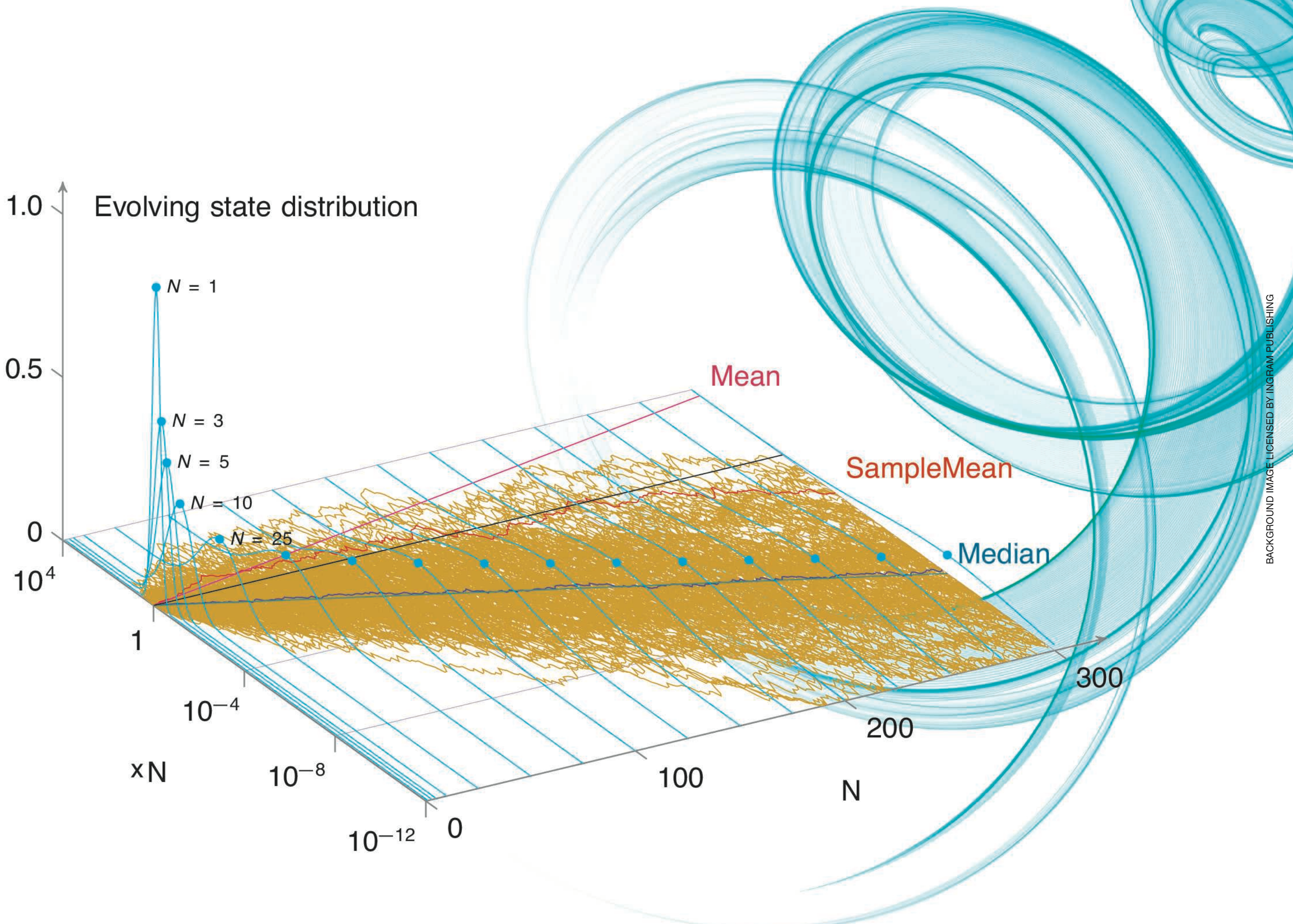
Another early formulation of the considered problem was given by Kalman [3] and reports on the results from his Ph.D. thesis. A discrete-time control design problem was posed, with the open-loop system being a vector-valued, discrete-time difference equation model of the form

$$x_{k+1} = A(w_{k+1})x_k + B(w_{k+1})u_k, \qquad (1)$$

where w_k is a sequence of independent random events. It is assumed that the probability distributions of $A(w_k)$ and $B(w_k)$ do not depend on k. The design criteria was, again, mean-square stability. This problem was also considered and characterized in the frequency domain for both continuous- and discrete-time systems by Willems and Blankenship in [4].

Mean-square stability conditions are appealing as they can be formulated in the multivariable case and relate directly to covariance matrices [5]. Furthermore, they lead to convex optimization problems for analysis and, in some

Digital Object Identifier 10.1109/MCS.2020.3048453
Date of current version: 24 March 2021



cases, controller design (see, for example, [6]). The disadvantage, which will become obvious in this article, is that mean-square stability is a very strong form of stability. In many applications, something weaker might be preferable. In the analysis of systems where mean-square stability is not satisfied, a weaker stability characterization would give a better understanding of the observed behavior.

Adaptive control research began in the 1960s and motivated the study of feedback systems with stochastically varying parameters. The work by Åström in [7] corresponds most closely to the approach discussed in this article, in that it characterizes the distributions that result in such systems. A continuous-time setting was used in [7] and much of the earlier works, which considers stochastic differential equations of the form

$$dx = x dw_1 + dw_2,$$

where w_1 and w_2 are Wiener processes. Some of the characteristics of the limiting distributions in [7] are also evident in the distributions arising in this article. In contrast, this work considers stochastic difference equations where the multiplicative term can be drawn from a wide range of possible distributions. A continuous-time setting was also used by Blankenship [8] with the system model

$$\dot{x}(t) = A(t)x(t), \quad x(t_0) = x_0 \in \mathbb{R}^n,$$

where the elements of $A(t)$ are stochastic stationary processes with certain continuity properties. The results use differential equation solution bounds to give sufficient conditions under which

$$\text{Prob}\left\{\lim_{t\to\infty} |x(t)| = 0\right\} = 1. \quad (2)$$

This article makes the case that, in many applications, the stability of the median is an important practical concept. Interestingly, a similar case has also been made in the domain of gambling strategies [9], where it was observed that proportional betting—a multiplicative strategy analogous to the stochastic feedback configuration—optimizes the median of the gambler's fortune. Gamblers care about the median as it characterizes their probability of making a profit. In contrast, the gambling house cares about the mean as it characterizes their risk.

From a probability theory perspective, the work presented here can be considered an application of renewal theory, which addresses the properties of the cumulative effect of a sequence of independent identically distributed (i.i.d.) random variables. The classic example is the study of the total number of "arrivals" in terms of the distribution of independent random interarrival times. In the context of (1), the cumulative effect of the stochastic $A(w_k)$ and $B(w_k)$ variables is reflected in the evolution of the state $x(k)$. In this framework, Kesten [10] studied the properties of the limiting distribution of the matrix evolution

$$x_{k+1} = A_k x_k + q_k, \quad x \in \mathbb{R}^n, \quad k \geq 0, \tag{3}$$

where A_k is a random matrix with positive entries and q_k a random vector. The conditions under which there exists a limiting distribution, $f_{x_\infty}(x)$, are given and essentially a generalization of the median stability results presented in this article. In the scalar case [10], Kesten showed that $f_{x_\infty}(x)$ can be heavy tailed, even if the distributions of A_k and q_k are relatively light tailed. Kesten's work was extended in work by Goldie [11], where a range of similar recursions was shown to also give power-law distribution tails. The limiting distribution, when it exists, was derived by Brandt [12]. The existence conditions are essentially equivalent to those for the stability of the median derived in this work. Work by de Saporta [13] describes an interesting variation on the recursion of (3) by considering the limiting distribution in the case where A_k comes from a Markov chain. Much more on the stochastic stability of Markov chains can be found in the comprehensive text of Meyn and Tweedie [14].

The discrete-time case in (1) and (3) can also be studied by considering the evolution of the state as the result of a product of random matrices. Random matrix products are known to lead to heavy-tailed distributions [15]. We show that this can lead to a situation where the mean does not characterize the typical behavior of the system. Stability can be considered in terms of the largest Lyaponov exponent

$$\lambda = \lim_{N \to \infty} \frac{1}{N} \mathcal{E}\left[\ln(\| A_N \|)\right], \quad \text{with } A_N = \prod_{k=1}^{N} A_k,$$

and where $\mathcal{E}[\ ]$ denotes the expectation. Calculating the upper and lower bounds on the largest Lyapunov exponent is currently an active area of research [16], [17].

One of the applications considered in this article is the stabilization of an unknown system via stochastic feedback. This has also been considered by Milisavljević and Verriest [18], who provide a stability condition that is an application of our results on median stability.

The growth in research interest and application of networked control systems has introduced another application of this theory. Sinopoli et al. [19] showed that Kalman filters with intermittent observations can lose mean-square stability once the probability of missing a packet reaches a threshold value. The focus on mean-square stability is natural with Kalman filtering as the construction of the time-varying Kalman filter requires a well-defined covariance matrix evolution. In the case of a static Kalman gain, the evolution of the estimation error is of the form given in (3). An analogous result on stabilization over fading channels was shown by Elia [20]. Elia also observed the emergence of heavy-tailed distributions in networked control systems in the case where mean-square stability is lost [21] and provided a mathematical characterization of this behavior in [22]. Work by Mo and Sinopoli [23] extended the packet loss model and provided bounds on the tail of the error distribution. Dey and Schenato [24] studied the distinction between the instability of the second moment and the conditions required for the existence of a limiting power-law distribution. As also noted in this work, this is the distinction between median stability and variance stability.

The adaptive control application that provided motivation for the analysis of these systems in the 1960s has recently received renewed attention. Rantzer [25] considers a single-parameter case and examines the stability of various moments. Concentration bounds on the distribution of the parameter error are derived.

This article focuses on the scalar discrete-time case given in (3), without the random exogenous input q_k, and shows that even though the distribution of $f_{x_\infty}(x)$ might be heavy tailed, it is still possible that (2) holds; the state x_k decays to zero with probability one. By focusing on the scalar case, the distributions are derived and calculated as a function of the time index k and give explicit conditions for the stability of the median, mean, and variance of the state. The results given are not unexpected in light of the prior work outlined earlier in this section. However, the explicit characterization of stability conditions and the calculation of the distributions involved at each time step provide insight into the manner in which the instabilities manifest themselves.

Notation

The notation $a \sim f_a(a)$ denotes that the random variable a is drawn from a distribution with probability density function $f_a(a)$. The cumulative distribution function is denoted by $F_a(a)$ and the complementary cumulative distribution by $\bar{F}_a(a)$ $(=1-F_a(a))$. The expected value of a is denoted by $\mathcal{E}[a] = \mu_a$. The normal distribution of mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$, and the log-normal distribution is denoted by $\mathcal{LN}$. The set of (nonnegative) integers is denoted by $(\mathbb{Z}_+)\,\mathbb{Z}$ and the reals by $(\mathbb{R}_+)\,\mathbb{R}$.

PROBLEM DESCRIPTION

The plant is a first-order system with the scalar state x_k evolving with dynamics given by

$$x_{k+1} = a_k x_k, \quad k = 0, 1, \ldots, \tag{4}$$

where a_k are independent random variables drawn from a distribution $f_a(a)$ with mean μ_a and variance σ_a^2. The distribution $f_a(a)$ is assumed to have support on only $a > 0$, and x_0 is assumed to be strictly positive.

The dynamics in (4) can be viewed as multiplicative noise a_k entering a feedback loop. An alternative view is that of a stochastic feedback gain. Both interpretations are illustrated in Figure 1, and both involve a feedback loop around a delay.

This is the simplest case of the type of processes described previously. As only a scalar state and uncorrelated a_k are considered, it is too simplistic for many real processes of this type. However, it is a prototypical case and illustrates some of the phenomena that may arise in more complex systems. Understanding the stability characteristics of this system is a precursor to understanding those for more complex systems.

Stability

At the time index N, the state x_N is given by

$$x_N = \prod_{k=0}^{N-1} a_k x_0. \tag{5}$$

As $a_k \sim f_a(a)$, state x_N is also a stochastic variable with a probability density function denoted by $f_{x_N}(x)$. Of interest are the properties of this distribution as $N \to \infty$. More specifically, conditions will be derived for the following three notions of stability.

At every time step N, the median of x_N [referred to as $\text{median}(x_N)$ in the sequel] is defined as any value m_N, such that

$$\int_0^{m_N} f_{X_N}(x)dx \geq \frac{1}{2} \quad \text{and} \quad \int_{m_N}^{\infty} f_{X_N}(x)dx \geq \frac{1}{2}.$$

Next, the stability conditions of interest are defined. The system in (5) is *median stable* if the density of the state x_N satisfies

$$\lim_{N\to\infty} \text{median}(x_N) = 0.$$

The system in (5) is *mean stable* if the density of the state x_N satisfies

$$\lim_{N\to\infty} \mathcal{E}[x_N] = 0,$$

and the system in (5) is *variance stable* if the density of the state x_N satisfies

$$\lim_{N\to\infty} \mathcal{E}[(x_N - E\{x_N\})^2] = 0.$$

The approach taken will involve analyzing the system in terms of the probability density functions of the logarithmic variables

$$\zeta_k := g(x_k) = \ln(x_k) \tag{6}$$

and

$$\alpha_k := g(a_k) = \ln(a_k). \tag{7}$$

The function g is defined here for convenience in subsequent derivations.

Assume that $x_0 > 0$ is given; thus, $x_N \geq 0$ for all $N \in \mathbb{Z}_+$. For simplicity in the following, assume without loss of generality that $x_0 = 1$. The evolution of the dynamics in (4) now becomes

$$\zeta_{k+1} = \zeta_k + \alpha_k. \tag{8}$$

This allows for expression of x_N as

$$x_N = e^{\zeta_N}.$$

This evolution for N timesteps can be examined to illustrate the way in which stability results will be derived. Under this assumption that $x_0 = 1$,

$$\mathcal{E}[\zeta_N] = \mathcal{E}\left[\sum_{k=0}^{N-1} \alpha_k\right] = N\mathcal{E}[\alpha] = N\mu_\alpha$$

if the distributions $f_{\alpha_k}(\alpha)$ are i.i.d.

It is tempting to say that if $\mu_\alpha < 0$, then $\mathcal{E}[x_N] < 1$. This is not true as the results in the following sections will make clear. What will be true is that, for distributions where the distribution of $\ln(a)$ satisfies certain moment assumptions,

$$\mathcal{E}[\alpha] < 0 \quad \Leftrightarrow \quad \lim_{N\to\infty} \text{median}(x_N) = 0.$$

Commutative Variable Relationships

The system will be studied in terms of the probability distributions of both the x_k, a_k variables and their logarithmic versions: ζ_k, α_k. The logarithmic/exponential relationship between these variables means that one can map the distributions from one set of variables to the other. Figure 2 provides a commutative diagram of these relationships as they evolve over the time index.

The mappings to the logarithmic variables in (6) and (7) map the corresponding distributions. This is described for the α variable, but it also applies to the x_k variables. Suppose that a has a probability distribution $f_a(a)$ defined on a support $S_a \subseteq \mathbb{R}_+$,

$$S_a = \{a \mid f_a(a) > 0\}.$$

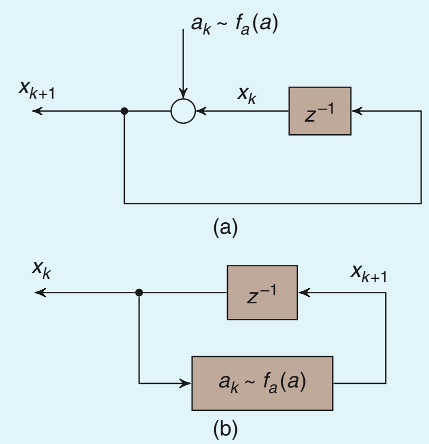


FIGURE 1 The process dynamics. (a) The system generating the stochastic dynamics is shown as a multiplicative noise signal in a feedback loop. (b) The equivalent system is a feedback loop with a stochastic gain.

If $g(a)$ is monotonically increasing and invertible on S_a, then the probability density of α is given by

$$f_\alpha(\alpha) = \begin{cases} f_a\left(g^{-1}(\alpha)\right)\left|\frac{dg^{-1}(\alpha)}{d\alpha}\right|, & \text{if } \alpha \in S_\alpha, \\ 0, & \text{if } \alpha \notin S_\alpha \end{cases} \quad (9)$$

and has support

$$S_\alpha = \{\alpha = g(a) \mid a \in S_a\}.$$

In this case,

$$g(a) = \ln(a), \quad g^{-1}(\alpha) = e^\alpha$$

and

$$\left|\frac{dg^{-1}(\alpha)}{d\alpha}\right| = |e^\alpha| = e^\alpha.$$

Of interest is the distribution of x_N as N increases, and, as shown from Figure 2, there are several ways of calculating this distribution. One can directly consider the evolution of the variable

$$x_{k+1} = a_k x_k,$$

where x_k has density $f_{x_k}(x)$ and a_k has density $f_a(a)$. This can be calculated as

$$f_{x_{k+1}}(x) = \int_{-\infty}^{\infty} f_{x_k}(\xi) f_a(x/\xi) \frac{1}{|\xi|} d\xi.$$

The $f_{x_{k+1}}(x)$ distribution can also be obtained by first transforming x_k to ζ_k using the mapping in (9):

$$f_{\zeta_k}(\zeta) = f_{x_k}(e^\zeta) e^\zeta.$$

The ζ dynamics are simply additive [see (8)]; thus, the ζ_{k+1} distribution is given by the convolution

$$f_{\zeta_{k+1}}(\zeta) = \int_{-\infty}^{\infty} f_{\zeta_k}(\xi) f_\alpha(\zeta - \xi) d\xi.$$

The $f_{x_{k+1}}(x)$ distribution is then given by the inverse of the mapping in (9):

$$f_{X_{k+1}}(x) = f_{\zeta_{k+1}}(\ln(x)) \frac{1}{x}, \quad x > 0.$$

Log-Normal Distributions

The case where $f_a(a)$ is a log-normal distribution is, in some sense, generic. If $f_a(a)$ is a log-normal distribution, then $f_\alpha(\alpha)$ is normal. Then ζ_N is a sum of independent, normal random variables and is, consequently, also normally distributed. Equivalently, the distribution of x_N is log-normal for all N. In other words, log-normal distributions are closed under the multiplication of random variables. Log-normal distributions, in conjunction with multiplicative noise, have been studied extensively in communication theory. See [26] for a widely applicable example.

The central-limit theorem implies that even if the $f_\alpha(\alpha)$ distribution is not normal, the scaled distribution of ζ_N tends to a normal distribution. One therefore also expects that the scaled x_N distribution tends to a log-normal distribution. Note that this argument requires that the random variable α have a finite second moment. This is a very weak assumption, and there are examples where the random variable a may not even have a finite mean, yet all of the moments of the corresponding α random variable are finite. An example of such a case is presented later.

Several properties of log-normal distributions are presented here for later use. All of these results can be found in [27]. The log-normal distribution can be defined by considering a as given by

$$a = e^\alpha, \quad \text{where} \quad \alpha \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2).$$

The parameters μ_α and σ_α are known as the location and scale parameters, respectively; however, they are simply referred to as the mean and standard deviation in the α-space. This distribution definition ensures that $S_a = \mathbb{R}_+$ and so fits the assumptions of this problem.

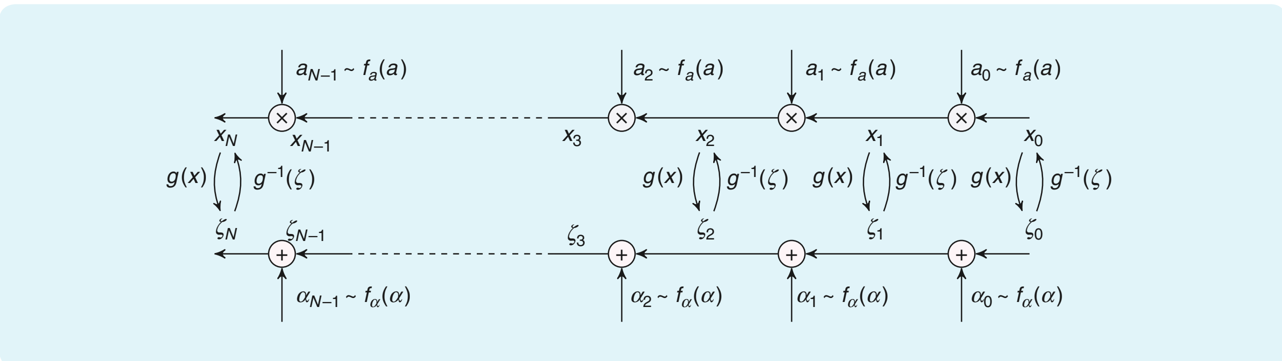


FIGURE 2 The state variable evolution. The mapping between the x and ζ variables is given by $\zeta_i = g(x_i) = \ln(x_i)$, and the inverse mapping is $x_i = g^{-1}(\zeta_i) = e^{\zeta_i}$.

It is clear from (8) that ζ_N is the sum of N random variables α_k with each $\alpha_k \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2)$. The sum of independent, normally distributed variables is also normally distributed and

$$\zeta_N \sim \mathcal{N}(N\mu_\alpha, N\sigma_\alpha^2).$$

This gives a closed-form expression for the distribution of ζ_N:

$$f_{\zeta_N}(\zeta) = f_{\mathcal{N}}(\alpha, N\mu_\alpha, N\sigma_\alpha^2), \tag{10}$$

where

$$f_{\mathcal{N}}(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}.$$

Closed-form expressions relate the mean and variance of a_k to the mean and variance of α_k [28], [29].

$$\mu_\alpha = \ln\left(\frac{\mu_a}{\sqrt{1 + \frac{\sigma_a^2}{\mu_a^2}}}\right) \tag{11}$$

$$\sigma_\alpha^2 = \ln\left(1 + \frac{\sigma_a^2}{\mu_a^2}\right). \tag{12}$$

The inverse mapping is given by

$$\mu_a = \mathrm{e}^{\mu_\alpha + \sigma_\alpha^2/2} \tag{13}$$

$$\sigma_a^2 = (\mathrm{e}^{\sigma_\alpha^2} - 1)(\mathrm{e}^{2\mu_\alpha + \sigma_\alpha^2}). \tag{14}$$

The fact that the mean and variance of a are not simply the exponentiation of the corresponding α domain values leads to interesting characterizations of stability in the x domain. The mode and median of a are also given by simple expressions:

$$\text{mode}(a) = \mathrm{e}^{\mu_\alpha - \sigma_\alpha^2} \tag{15}$$

$$\text{median}(a) = \mathrm{e}^{\mu_\alpha}. \tag{16}$$

The median condition—and any other quantile value—is transformed via exponentiation, making it a simple matter to characterize properties of the median or quantile value.

STABILITY CONDITIONS

The following sections derive the conditions under which the median, mean, and variance of $f_{X_N}(x)$ converge to zero as $N \to \infty$.

Mean Stability

The condition for the stability of the mean of x_N is a simple consequence of the fact that for two independent, random variables, the product of the expectations is equal to the expectation of the product.

Theorem 1: Mean Stability

$$\lim_{N\to\infty} \text{mean}(x_N) = 0 \;\Leftrightarrow\; \mu_a < 1.$$

For log-normal distributions, (13) shows that the mean stability condition can also be stated by the mean and variance of the $f_\alpha(\alpha)$ distribution

$$\lim_{N\to\infty} \text{mean}(x_N) = 0 \;\Leftrightarrow\; \mu_\alpha + \sigma_\alpha^2/2 < 0. \tag{17}$$

Variance Stability

Goodman [30] derives the variance of a product of arbitrary random variables, which directly leads to the following variance stability result.

Theorem 2: Variance Stability

$$\lim_{N\to\infty} \text{variance}(x_N) = 0 \;\Leftrightarrow\; \mu_a^2 + \sigma_a^2 < 1.$$

Proof of Theorem 2

From [30]

$$\text{variance}(x_N) = \text{variance}\left(\prod_{k=0}^{N-1} a_k\right) = \prod_{k=0}^{N-1}(\sigma_a^2 + \mu_a^2) - \prod_{k=0}^{N-1}\mu_a^2,$$

which, in this i.i.d. case, gives

$$\text{variance}(x_N) = (\sigma_a^2 + \mu_a^2)^N - \mu_a^{2N} \tag{18}$$

$$= \sigma_a^2(\sigma_a^2 + \mu_a^2)^{N-1} + \mu_a^{2N}\left(\left(1 + \frac{\sigma_a^2}{\mu_a^2}\right)^{N-1} - 1\right). \tag{19}$$

If $\sigma_a^2 + \mu_a^2 < 1$, then both of the terms in (19) go to zero as $N \to \infty$. In the case where $\sigma_a^2 + \mu_a^2 = 1$ and $\sigma_a^2 > 0$, (19) also shows that the variance of x_N goes to one as $N \to \infty$. If $\sigma_a^2 + \mu_a^2 > 1$, then both of the terms in (19) grow without bound as $N \to \infty$. ■

In the log-normal distribution case, substituting (13) and (14) into the condition of Theorem 2 gives an equivalent condition in terms of the normal $f_\alpha(\alpha)$ distribution,

$$\lim_{N\to\infty} \text{variance}(x_N) = 0 \;\Leftrightarrow\; \mu_\alpha + \sigma_\alpha^2 < 0. \tag{20}$$

The $f_\alpha(\alpha)$ condition for variance stability in (20) is clearly stronger than the mean stability condition given in (17). We also note that the variance stability condition implies that $x_N \to 0$ in $\mathcal{L}_2$ as $N \to \infty$.

Median Stability: Log-Normal Case

The least-restrictive stability condition to be considered is that for the median of x_N. This result is easy to obtain for a log-normal distribution and is thus presented first.

Theorem 3: Median Stability—Log-Normal Distribution

If $f_a(a)$ is a log-normal distribution,

$$\lim_{N\to\infty} \text{median}(x_N) = 0 \;\Leftrightarrow\; \mu_\alpha < 0.$$

TABLE 1 Stochastic feedback gain stability conditions for log-normal distributions.

Stability Property	$f_a(a)$ Distribution	$f_\alpha(\alpha)$ Distribution
median(x_N)	$\mu_a^2 - \frac{\sigma_a^2}{\mu_a^2} < 1$	$\mu_\alpha < 0$
mean(x_N)	$\mu_a < 1$	$\mu_\alpha + \frac{\sigma_\alpha^2}{2} < 0$
variance(x_N)	$\mu_a^2 + \sigma_a^2 < 1$	$\mu_\alpha + \sigma_\alpha^2 < 0$

TABLE 2 Stochastic feedback gain stability conditions for more general distributions.

Stability Property	$f_a(a)$ Distribution	$f_\alpha(\alpha)$ Distribution
median(x_N)	—	$\mu_\alpha < 0$
mean(x_N)	$\mu_a < 1$	—
variance(x_N)	$\mu_a^2 + \sigma_a^2 < 1$	—

Proof of Theorem 3

As $f_{\zeta_N}(x)$ is a normal distribution, its median is equal to its mean:

$$\text{median}(\zeta_k) = \text{mean}(\zeta_N) = N\mu_\alpha.$$

This immediately gives $\lim_{N\to\infty} \text{median}(\zeta_N) = -\infty$ if and only if (iff) $\mu_\alpha < 0$. As $\text{median}(x_k) = e^{\text{median}(\zeta_N)}$, the result follows. ■

Equation (11) demonstrates that the condition of Theorem 3 can also be expressed in terms of the $f_a(a)$ distribution

$$\lim_{N\to\infty} \text{median}(x_N) = 0 \;\Leftrightarrow\; \mu_a^2 - \frac{\sigma_a^2}{\mu_a^2} < 1.$$

Note that, depending on the variance σ_a^2, systems with a mean of μ_a greater than one might still be median stable. This point is discussed in greater detail later in this article. The stability results for log-normal distributions are summarized in Table 1. The conditions can be expressed in terms of either the $f_a(a)$ or $f_\alpha(\alpha)$ distribution because the mapping between the distributions involves only the means and variances.

Median Stability: General Distributions

Now consider median stability in the case where $f_a(a)$ distributions are other than log-normal. For the purposes of this section, assume that $f_\alpha(\alpha)$ is a nonlattice distribution with a bounded third moment. Note that a lattice distribution is one where there exist parameters $b \in \mathbb{R}$ and $h > 0$ such that $\text{Prob}\{\alpha \in b + h\mathbb{Z}\} = 1$. The mean and variance relationships between the $f_a(a)$ and $f_\alpha(\alpha)$ distributions given in (11)–(14) no longer hold. Unfortunately, this is also true for all the values of N as well as in the limit as $N \to \infty$.

The situation is more complex for more general $f_a(a)$ distributions as $f_\alpha(\alpha)$ is not normal. In the context of the central-limit theorem, it is perhaps surprising that, although the distribution $f_{\zeta_N}(\zeta)$ is the N-fold convolution of $f_\alpha(\alpha)$ distributions, the median of $f_{\zeta_N}(\zeta)$ does not necessarily converge to the mean of $f_{\zeta_N}(\zeta)$. The difference can be quantified.

Lemma 1

Assume that $f_\alpha(\alpha)$ is a nonlattice distribution with a bounded third moment. Then,

$$\lim_{N\to\infty} \text{median}(\zeta_N) - \mathcal{E}[\zeta_N] = -\frac{\mathcal{E}[(\alpha-\mu_\alpha)^3]}{6\sigma_\alpha^2}.$$

Proof of Lemma 1

$$\begin{aligned}\text{median}(\zeta_N) - \mathcal{E}[\zeta_N] &= \text{median}\left(\sum_{k=0}^{N-1}\alpha\right) - N\mu_\alpha \\ &= \sigma_\alpha \text{median}\left(\sum_{k=0}^{N-1}\frac{\alpha-\mu_\alpha}{\sigma_\alpha}\right).\end{aligned}$$

Define a new stochastic variable y with probability density function $f_y(y)$ by

$$y \;:= h(\alpha) = \frac{\alpha - \mu_\alpha}{\sigma_\alpha},$$

and note that $\mathcal{E}[y] = \mu_y = 0$ and $\mathcal{E}[y^2] = \sigma_y^2 = 1$. Define τ as the third moment of $f_y(y)$,

$$\tau \;:= \mathcal{E}[y^3],$$

and, by assumption, $\tau < \infty$. The following result from Hall [31] provides the key step. If $f_y(y)$ is a nonlattice distribution with $\mu_y = 0, \sigma_y = 1$ and $\mathcal{E}[y^3] = \tau < \infty$, then

$$\lim_{N\to\infty} \text{median}\left(\sum_{k=0}^{N-1} y\right) = \frac{-\tau}{6}.$$

As

$$\begin{aligned}\lim_{N\to\infty} \text{median}(\zeta_N) - \mathcal{E}[\zeta_N] &= \sigma_\alpha \text{median}\left(\sum_{k=0}^{N-1}\frac{\alpha-\mu_\alpha}{\sigma_\alpha}\right) \\ &= \sigma_\alpha \text{median}\left(\sum_{k=0}^{N-1} y\right) \\ &= \frac{-\sigma_\alpha \tau}{6}.\end{aligned} \tag{21}$$

The only thing remaining is to determine the value of τ. As $h^{-1}(y) = \sigma_\alpha y + \mu_\alpha$,

$$\begin{aligned}\tau &= \int_{-\infty}^{\infty} y^3 f_y(y)dy = \int_{-\infty}^{\infty} y^3 f_\alpha(h^{-1}(y)) \left|\frac{dh^{-1}(y)}{dy}\right| dy \\ &= \int_{-\infty}^{\infty} y^3 f_\alpha(\sigma_\alpha y + \mu_\alpha)\,|\sigma_\alpha|\,dy \\ &= \int_{-\infty}^{\infty} \frac{(\alpha-\mu_\alpha)^3}{\sigma_\alpha^3} f_\alpha(\alpha)|\sigma_\alpha| \frac{d\alpha}{|\sigma_\alpha|} \\ &= \frac{1}{\sigma_\alpha^3}\int_{-\infty}^{\infty}(\alpha-\mu_\alpha)^3 f_\alpha(\alpha)d\alpha \\ &= \frac{\mathcal{E}[(\alpha-\mu_\alpha)^3]}{\sigma_\alpha^3}.\end{aligned}$$

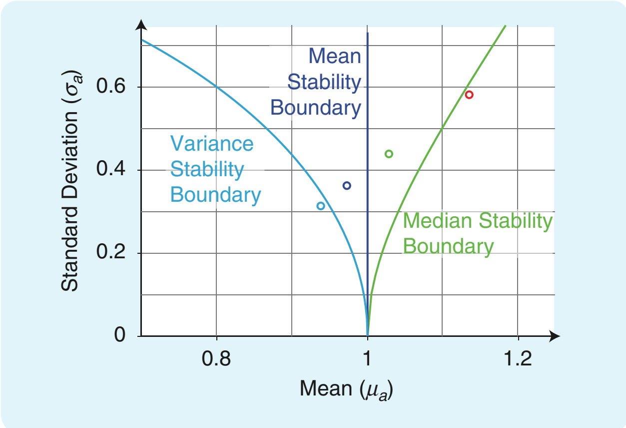


FIGURE 3 The stability regions in the a domain. The region where the condition is satisfied is to the left of the correspondingly colored boundary. The mean and variance stability regions are applicable to general distributions. The median stability region shown here applies to log-normal $f_a(a)$ distributions. Other distributions will have slightly different median stability regions but identical mean and variance stability regions. The four circles indicate the values of μ_α and σ_α of the distributions shown in Figure 4.

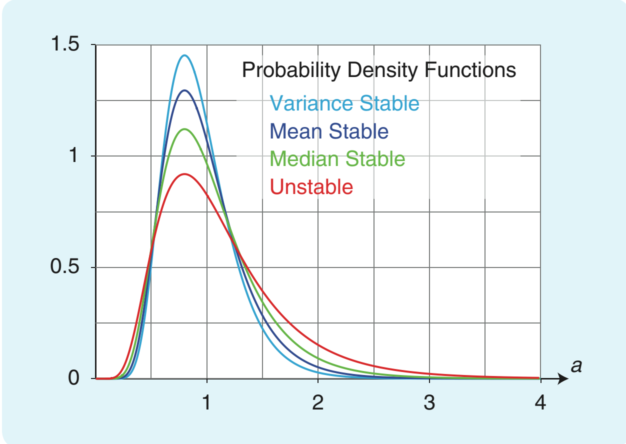


FIGURE 4 The probability density functions of a for four stability cases: variance stable, mean stable, median stable, and unstable. All four cases have the same mode; their means and standard deviations are presented in Figure 3.

Substituting this expression into (21) gives the desired result. ■

The key point in determining the median stability is that the limit in Lemma 1 is independent of N.

Theorem 4: Median Stability

Assume that $f_\alpha(\alpha)$ is a nonlattice distribution with a bounded third moment. Then,

$$\lim_{N\to\infty} \text{median}(x_N) = 0 \quad \Leftrightarrow \quad \mu_\alpha < 0.$$

Proof of Theorem 4

If $\mu_\alpha = 0$, then $\mathcal{E}[\zeta_N] = N\mu_\alpha = 0$, and from Lemma 1, $\lim_{N\to\infty} \text{median}(\zeta_N)$ is a finite constant. Therefore,

$$\lim_{N\to\infty} \text{median}(x_N) = \lim_{N\to\infty} e^{\text{median}(\zeta_N)} \neq 0.$$

By Lemma 1, for every $\epsilon > 0$, there exists an integer $\bar{N}$ such that for all $N > \bar{N}$

$$\left|\text{median}(\zeta_N) - N\mu_\alpha + \frac{\mathcal{E}[(\alpha-\mu_\alpha)^3]}{6\sigma_\alpha^2}\right| < \epsilon.$$

This implies that

$$|\text{median}(\zeta_N) - N\mu_\alpha| < \left|\frac{\mathcal{E}[(\alpha-\mu_\alpha)^3]}{6\sigma_\alpha^2}\right| + \epsilon.$$

If it is assumed that $\mu_\alpha < 0$, then $N\mu_\alpha < 0$ and

$$\text{median}(\zeta_N) < N\mu_\alpha + \left|\frac{\mathcal{E}[(\alpha-\mu_\alpha)^3]}{6\sigma_\alpha^2}\right| + \epsilon.$$

The right-hand side clearly goes to $-\infty$ as $N \to \infty$. An analogous argument for $\mu_\alpha > 0$ gives a lower bound on $\text{median}(\zeta_N)$ that goes to ∞ as $N \to \infty$. Exponentiating $\text{median}(\zeta_N)$ provides the required result. ■

This result also follows from Cantelli's inequality (see Lemma 3) without the requirement of a bounded third moment. However, the method of proof stated previously in this section illustrates the manner in which a sum of nonnormal distributions does not converge to a normal distribution. It also gives the following interesting boundary condition.

Corollary 1: Median Limit—Zero-Mean Log Distribution

If $f_\alpha(\alpha)$ is a nonlattice distribution with $\mu_\alpha = 0$ and $|\mathcal{E}[(\alpha-\mu_\alpha)^3]| < \infty$, then

$$\lim_{N\to\infty} \text{median}(x_N) = e^{-\frac{\mathcal{E}[(\alpha-\mu_\alpha)^3]}{6\sigma_\alpha^2}}.$$

Ethier [9] also uses the result of Hall to prove a similar result on the median of a gambler's fortune. The results in [9] suggest that it might be possible to remove the assumption of the nonlattice distribution in Lemma 1. Lattice distributions are important for gambling applications but may be of less interest in many control applications. The stability conditions for more general distributions (that is, nonlattice and with a bounded third moment) are summarized in Table 2. In the table, "—" denotes the fact that no mean and variance condition exists for every distribution. It is possible to find distributions with nonzero third moments that violate the corresponding normal-log-normal conditions of Table 1.

The median stability conditions given here for a stochastic gain $a_k \sim f_a(a)$ have a similar form to the stability conditions for a time-varying gain (see "Stabilization By Time-Varying Gains and the Geometric Mean").

The implications of median stability can be stated in terms of well-known convergence characterizations for

Stabilization By Time-Varying Gains and the Geometric Mean

Some of the initially nonintuitive phenomena observed for stochastic feedback may be better understood by considering systems with certain types of deterministic, but time-varying, feedback gains. For the case of a scalar state, a complete analysis is easy to accomplish (see [S1] for a more complete analysis of the periodic multivariable case). Consider the single-state, discrete-time system and its solution

$$x_{k+1} = a_k x_k \Rightarrow x_N = \left(\prod_{k=0}^{N-1} a_k\right) x_0. \tag{S1}$$

If a is a periodic signal with period N, then the growth of x can be characterized by observing the behavior every N time steps. Define the *subsampled state*

$$\hat{x}_k := x_{kN}.$$

Note that x decays iff $\hat{x}$ decays because the growth of x between the subsamples is bounded. The recursion for $\hat{x}$ is time invariant:

$$\hat{x}_{k+1} = x_{(k+1)N} = \left(\prod_{k=0}^{N-1} a(k)\right) x_{kN} =: \hat{a}\,\hat{x}_k,$$

where $\hat{a} := \prod_{k=0}^{N-1} a_k$ is the so-called monodromy gain. Thus, the sequence $\hat{x}$ decays iff

$$|\hat{a}| < 1 \;\Leftrightarrow |\hat{a}|^{\frac{1}{N}} < 1 \;\Leftrightarrow\; \left(\prod_{k=0}^{N-1} |a_k|\right)^{\frac{1}{N}} < 1. \tag{S2}$$

The last quantity in (S2) is the *geometric mean* of the absolute value of the signal a, which is the right quantity that characterizes stability in this system. The geometric mean can also be expressed using the arithmetic mean of the logarithm

$$\left(\prod_{k=0}^{N-1} |a_k|\right)^{\frac{1}{N}} < 1 \;\Leftrightarrow\; \ln\left(\prod_{k=0}^{N-1} |a_k|\right)^{1/N} < 0$$
$$\Leftrightarrow\; \frac{1}{N}\sum_{k=0}^{N-1} \ln|a_k| < 0.$$

Thus, the system is asymptotically stable iff the *arithmetic mean of* $\{\ln|a_k|\}$ *is negative*. Note how this is analogous to the condition $E\{\ln|a|\} < 0$ when a is a stochastic process. The relation between the geometric and arithmetic means through the logarithm function is illustrated in Figure S1. The figure depicts a periodic gain a that is symmetrically distributed around one. The $\ln(a_k)$ mapping tends to boost those values of $\{a_k\}$ that are lower than one more heavily toward large negative numbers while tempering the values of $\{a_k\}$ that are larger than one by mapping them to smaller positive numbers. The result is that, even though a may be symmetrically distributed around one, the product $\prod_{k=0}^{N-1}|a_k|$ will be strictly smaller than one.

Next, examine (S1) in the case where the sequence $\{a_k\}$ is a general time-varying gain. The asymptotic behavior of the solution is completely determined by the limit of the product of the gains $\{a_k\}$, which can be studied as a series limit by taking the logarithm

$$\ln\left|\prod_{k=0}^{N-1} a_k\right| = \sum_{k=0}^{N-1} \ln|a_k|.$$

Next, explore the limit

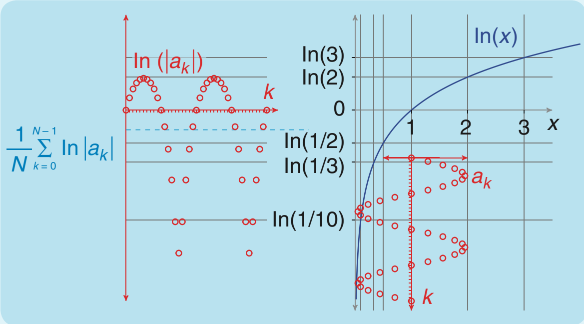


FIGURE S1 An illustration of the mapping from a_k to $\ln|a_k|$, showing that when a is symmetrically distributed around one, $\ln|a|$ is distributed more heavily toward negative numbers due to distortion by the ln mapping. The cyan dashed line on the graph of $\ln|a|$ indicates the arithmetic mean of that signal (this is the logarithm of the geometric mean of a), showing how it is negative while the ln of the mean of a is zero.

$$\begin{aligned}\lim_{N\to\infty}\sum_{k=0}^{N-1}\ln|a_k| &= \lim_{N\to\infty} N\left(\frac{1}{N}\sum_{k=0}^{N-1}\ln|a_k|\right)\\ &= N\left(\lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}\ln|a_k|\right)\\ &=: N\,\mathcal{E}[\ln|a|],\end{aligned} \tag{S3}$$

where the last limit is expressed in terms of the *asymptotic average*, which, for any signal u, is defined by

$$\mathcal{E}[u] := \lim_{N\to\infty}\frac{1}{N}\sum_{k=0}^{N-1} u_k.$$

This asymptotic average can be thought of as a "deterministic expectation" of u, which is equivalently the time average of the realization of a stochastic process.

It can thus be concluded that if the asymptotic average of $\ln|a|$ exists and is negative, then the state will asymptotically converge to zero, that is,

$$\mathcal{E}[\ln|a|] < 0 \;\Rightarrow\; \lim_{N\to\infty} x_N = 0.$$

Something slightly stronger can be concluded:

$$\mathcal{E}[\ln|a|] =: \ln(\gamma) < 0 \;\Rightarrow\; |x_N| \le \alpha\,\gamma^N, \tag{S4}$$

that is, the convergence is geometric, with a decay rate $\gamma < 1$. Finally, note that condition (S4) is necessary only for exponential convergence. Slower convergence can still occur, even when this condition does not hold. For convergence, it is necessary only for the sequence on the right-hand side of (S3) to go to $-\infty$. This can occur even when the asymptotic average is converging to zero (from below) as long as it converges at a rate slower than $1/N$. More precisely, it can be stated that

$$\lim_{N\to\infty} x_N = 0 \;\Leftrightarrow\; \lim_{N\to\infty} N\left(\frac{1}{N}\sum_{k=0}^{N-1}\ln|a_k|\right) = -\infty.$$

REFERENCE

[S1] S. Bittanti and P. Colaneri, *Periodic Systems: Filtering and Control.* New York: Springer-Verlag, 2009.

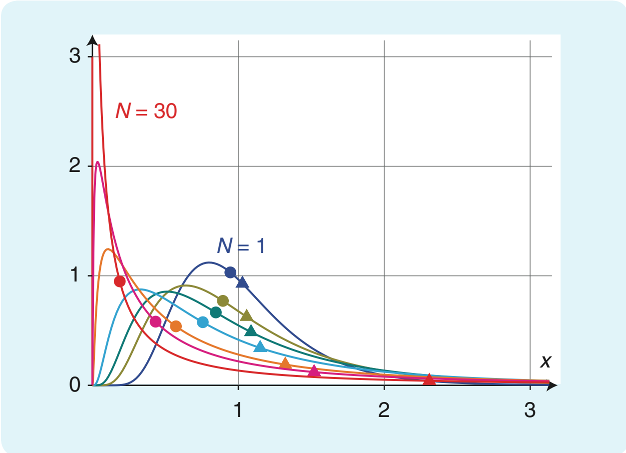


FIGURE 5 An evolution of the probability density function of x_N for $N = 1, 2, 3, 5, 10, 15$, and 30. The mean values of each x_N distribution are indicated by triangles, and the median values are indicated by solid dots. As $N \to \infty$, the mean of x_N increases, and the median decreases.

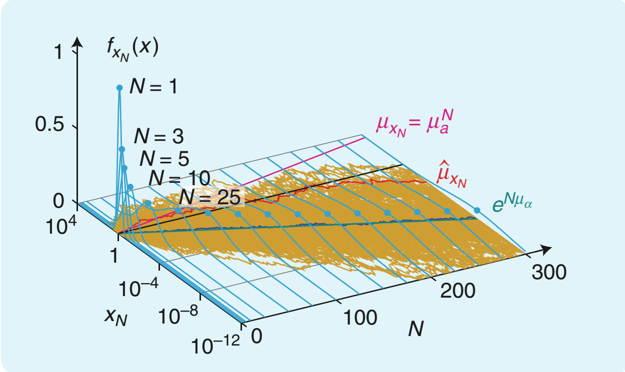


FIGURE 6 A simulation of the median stable/mean unstable case. The vertical axis shows the evolution of the probability density function $f_{x_N}(x)$ for a range of values of N. The horizontal plane shows 200 simulation sample paths of x_N as a function of N. Also shown are the sample median (blue line) and sample mean (red line: $\hat{\mu}_{x_N}$), along with the theoretical median (cyan line: $e^{N\mu_\alpha}$) and theoretical mean (magenta line: μ_{x_N}). The theoretical values are derived by mapping the α-space distribution in (10) through (15) and (16).

stochastic sequences. The sequence of medians of x_N arises from the stochastic dynamics in (5). The next section will show that, under the assumption that the variance of $f_\alpha(\alpha)$ is finite, the variance of x_N that arises from these same dynamics leads to exponentially decreasing bounds on the integral of the tail of the $f_{X_N}(x)$ distribution. These two facts imply that if the system in (5) is median stable, then x_N converges in probability to the degenerate random variable $x = 0$:

$$\lim_{N \to \infty} \text{Prob}\{\,|\,x_N\,| > \epsilon\} = 0, \quad \text{for all } \epsilon > 0.$$

Stability Regions

The variance, mean, and median stability regions for the $f_a(a)$ distribution are displayed in Figure 3. The most interesting observation is that there exists a region in which median (x_N) is stable and the mean (x_N) is unstable. It will subsequently be shown that, in this case, the sample paths x_N go to zero in probability, but the mean of x_N goes to ∞. This analysis can also be applied to other feedback loops; see "Stochastic Gain Stabilization," wherein the stability regions for a first-order system are derived with an unknown gain and pole position. The Matlab code that generated all the simulation data and figures for this article is publicly available at https://doi.org/10.3929/ethz-b-000457726.

Figure 4 gives the probability distribution functions for four stability cases: unstable, median stable, mean stable, and variance stable. In all the cases, the mode of the distribution is lower than one. The remarkable feature of these distributions is that they are not particularly different and yet give very different stability characteristics in the evolution of the state.

The most intriguing case is where the median of x_N is stable, but the mean is unstable. Figure 5 shows the evolution of the log-normal probability density function of x_N for a range of values of N. The evolution of the median toward zero and the mean toward infinity are clear in the distributions.

The median stable/mean unstable case is illustrated by simulating 200 sample paths. The $f_a(a)$ distribution is the log-normal distribution with a probability density function presented as case 3 of Figure 4 $(\mu_a = 1.0283, \sigma_a = 0.4389)$. Figure 6 illustrates the sample paths and the evolution of

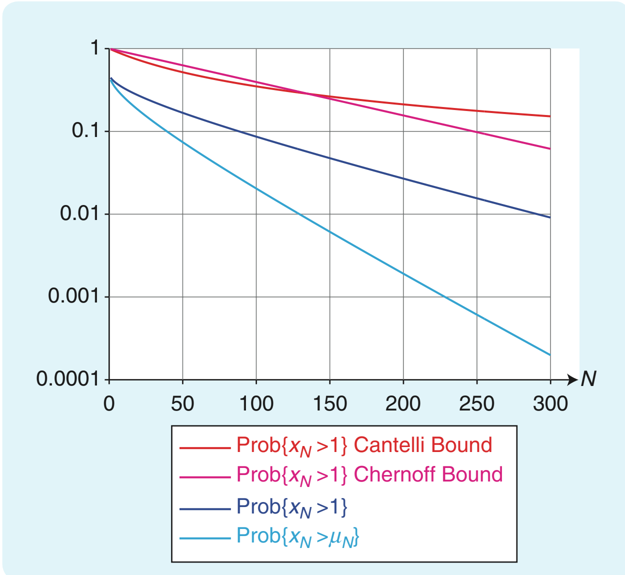


FIGURE 7 The concentration inequalities and tail distributions for x_N. The probability that $x_N > 1$ is calculated for the log-normal distribution case studied in Figure 6 (blue line). Also shown is the probability that $x_N > \mu_{x_N}$ (cyan line). For comparison, several concentration inequalities are also illustrated: the Cantelli inequality (red) and Chernoff inequality (magenta) for the same distribution. For the simulation shown in Figure 6, the probability that a sample trajectory will exceed the mean $(\mu_{x_N} = \mu_a^N)$ at the end of the interval $(N = 300)$ is 0.0002.

Stochastic Gain Stabilization

The stability results for stochastic feedback can be easily applied to the slightly more difficult problem of stabilizing a general first-order system via stochastic feedback. Figure S2 illustrates the configuration for this problem. This is a simple case of a more general stochastic stabilization problem, referred to as stabilization by noise. This problem has been studied in stochastic vibration control context (see the review in [S2]). In vibration control, the assumption of an oscillatory nominal response is usually exploited. The more general case is studied in [S3] and is based on earlier work in [S4]. This work focuses on the continuous-time equivalent to the mean stability case considered in this article. The application example has also been studied in [18], where a result, which is essentially equivalent to the median stability boundary below, is presented.

The plant is given and has the transfer function $G(z)$,

$$G(z) = \frac{\gamma}{z-\tau}.$$

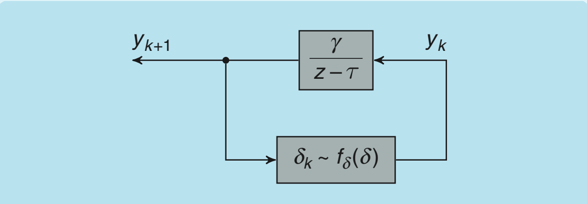


FIGURE S2 A stochastic stabilization problem. A first-order plant is connected in feedback with a stochastic gain $\delta_k \sim f_\delta(\delta)$.

Denote the plant output by y_k. The closed-loop dynamics of the feedback system illustrated in Figure S2 are given by

$$y_{k+1} = (\tau + \gamma\delta_k) y_k,$$

where $\delta_k \sim f_\delta(\delta)$ is the stochastic feedback drawn from a known distribution at each time instant. Now define $x_k = |y_k|$, and note that

$$x_{k+1} = |\tau + \gamma\delta_k| x_k.$$

As $x_k \geq 0$ for all k, the results summarized in Table 2 are directly applicable by defining

$$a_k = |\tau + \gamma\delta_k|.$$

The mean of the $f_a(a)$ distribution is

$$\mu_a = |\tau + \gamma\mu_\delta|.$$

The variance may be more difficult to evaluate precisely, but it can be easily estimated numerically. If the distribution $f_\delta(\delta)$ were such that $\delta_k > 0$, then

$$\sigma_a^2 = \gamma\sigma_\delta^2.$$

However, the absolute value in the definition of a complicates this somewhat, particularly in the case of interest, where $\tau \neq 0$. As expected, the conditions for median, mean, and variance stability differ, and for a given distribution, $f_\delta(\delta)$, a stability boundary diagram (analogous to that which is in Figure 3) can be drawn. Figure S3 depicts the stability regions for the case where δ_k is drawn from a normal distribution $\delta_k \sim \mathcal{N}(\mu_\delta, \sigma_\delta^2)$.

the distribution $f_\alpha(\alpha)$. As N increases, the median stability condition ensures that the probability of a sample path not going toward zero is zero; however, $\mu_a > 1$. Thus, the $\mathrm{mean}(x_N)$ is unstable and goes to ∞. For very large N, this results in an x_N distribution with a very high peak close to $x_N = 0$, with still-enough weight in the positive tail so that the mean of x_N is very large (and growing with N). The sample estimate of $\mathrm{mean}(x_N)$ (denoted by $\hat{\mu}_{x_N}$) drops below the theoretical mean μ_{x_N} as increasingly fewer sample paths are near or above the mean. This phenomenon is investigated in more detail in the next section.

CUMULATIVE DISTRIBUTIONS AND CONCENTRATION RESULTS

In this section we examine the observation made in the previous section (that is, in the median stable case, the mass of the distribution falls below the mean) in more detail. Specifically, the goal is to calculate (or at least provide an upper bound for) the probability that x_K exceeds a certain value. Denote the complementary cumulative distribution function by

$$\bar{F}_{a_N}(x_{\mathrm{bnd}}) = \mathrm{Prob}\{x_N > x_{\mathrm{bnd}}\} = \mathrm{Prob}\left\{\prod_{k=1}^{N} a_k > x_{\mathrm{bnd}}\right\},$$

where it is assumed that $x_0 = 1$. Furthermore, we are interested in the properties of $\bar{F}_{a_N}(x_{\mathrm{bnd}})$ as $N \to \infty$ because this gives information about the mass of the distribution of x_N as N increases.

Results of this nature are referred to as concentration inequalities in the statistics literature and have a long history. See [32] for a much more extensive treatment of concentration inequalities in stochastic processes similar to the ones considered in this article. Assume that $\mu_\alpha < 0$ (median stable case) and observe that two choices of x_{bnd} are of potential interest.

1) $x_{\mathrm{bnd}} = 1$: This gives the probability that $x_N > x_0$, addressing the question of the probability that a realization of the x_k trajectory grows over the interval $[0, N]$.
2) $x_{\mathrm{bnd}} = \mathrm{mean}(x_N) = \mu_a^N$: This provides insight into the ability (or lack thereof) to estimate the mean of x_N from a finite number of sample path realizations.

For simplicity, this article focuses on the first case. The results are easily extended to the second at the expense of more complex formulas in some cases.

An $f_\delta(\delta)$ distribution with a nonzero mean can be viewed as a constant feedback gain of μ_δ in parallel with a zero-mean stochastic gain. The static feedback effect of μ_δ is accounted for in the stability boundary figure by plotting the nominal case as $|\tau+\gamma\mu_\delta|$. Similarly, the standard deviation of the stochastic feedback is scaled by $1/\sqrt{|\gamma|}$ to normalize for the gain-scaling effect of γ. The condition for the nominal stability of the plant is that $|\tau+\gamma\mu_\delta|<1$. The median stability boundary shows that, for a range of variance, the median of $|y_k|$ is stable; however, if the nominal plant is not stable, then neither the mean nor the variance of $|y_k|$ can be stabilized by stochastic feedback. It is also interesting to note that, for any given nominal stability margin, there are increasingly large values of the variance of the stochastic feedback that will destabilize the variance, mean, and median (in that order).

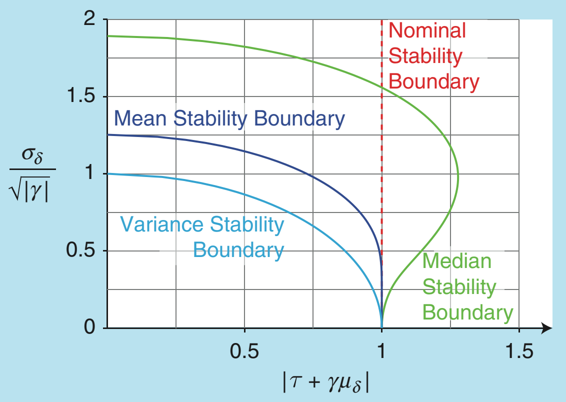


FIGURE S3 The stability boundaries for the state-magnitude evolution for the plant, $G(z)=\gamma/(z-\tau)$, in feedback with a stochastic gain, $\delta_k\sim\mathcal{N}(\mu_\delta,\sigma_\delta^2)$. Nominally unstable open-loop plants may be median stabilized by a stochastic gain with the appropriate variance.

Another observation is that the stability boundaries involve the absolute values of functions of the plant parameters τ and γ. This has an interesting robustness interpretation and implies that in the $\mu_\delta=0$ case, the plant can be median stabilized for a range of τ and γ, irrespective of their signs. For example, for a plant with τ in the range $-1.05\le\tau\le 1.05$, there exists a zero-mean, normally distributed stochastic feedback of a certain variance that will median stabilize the plant. This exceptional robustness should not be interpreted as an indication that the stochastic controller is practical. The mean and variance of the realizations of the trajectories are still growing without bound, and the random excursions could be extremely large. The stochasticity in the feedback loop leads to distributions of y_k that are heavy tailed. The potential value of these results is in avoiding the case where stochasticity in a feedback loop inadvertently leads to destabilization.

REFERENCES

[S2] J. Roberts and P. Spanos, "Stochastic averaging: An approximate method of solving random vibration problems," *Int. J. Non-Linear Mech.*, vol. 21, no. 2, pp. 111–134, 1986. doi: 10.1016/0020-7462(86)90025-9.

[S3] L. Arnold, H. Crauel, and V. Wihstutz, "Stabilization of linear systems by noise," *SIAM J. Control Optimiz.*, vol. 21, no. 3, pp. 451–461, 1983. doi: 10.1137/0321027.

[S4] V. Oseledec, "A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems," *Trans. Moscow Math. Soc.*, vol. 19, pp. 197–231, 1968.

The analysis is, of course, easier in the α-space. Thus,

$$\text{Prob}\{x_N>1\}=\text{Prob}\{\zeta_N>0\}.$$

The objective is to provide bounds on this probability as a function of N.

Log-Normal Distribution Case

First consider the log-normal $f_a(a)$ case as exact formulas are easily derived. In this case,

$$f_{\alpha_N}(\alpha)=f_{\mathcal{N}}(\alpha,N\mu_\alpha,N\sigma_\alpha^2),$$
$$F_{\alpha_N}(\alpha)=\int_{-\infty}^{\alpha} f_{\mathcal{N}}(y,N\mu_\alpha,N\sigma_\alpha^2)dy,$$

and

$$\bar{F}_\alpha(\alpha)=\frac{1}{2}\left(1-\text{erf}\left(\frac{\alpha-N\mu_\alpha}{\sqrt{2N\sigma_\alpha^2}}\right)\right),$$

where $\text{erf}(x)$ is the error function. The tail probability is then

$$\text{Prob}\{x_N>1\}=\frac{1}{2}\left(1-\text{erf}\left(\frac{-\sqrt{N}\mu_\alpha}{\sqrt{2\sigma_\alpha^2}}\right)\right). \quad (22)$$

In the median stable case, $\mu_\alpha<0$; thus, the argument of the error function is positive.

Figure 7 displays the application of the bound in (22) to the example simulated in Figure 6. When the distribution is known, $\text{Prob}\{x_N>1\}$ can be calculated numerically, which is shown as a function of N for the $f_a(a)\sim\mathcal{LN}$ case. Also shown is

$$\text{Prob}\{x_N>\mu_{x_N}\}=\text{Prob}\{x_N>\mu_a^N\},$$

and the exponential decrease of this probability illustrates why the sample-based estimate of μ_{x_N} rapidly deteriorates with increasing N.

The exponential decay in Figure 7 might seem counterintuitive as the complementary cumulative distribution of a standard normal distribution satisfies

$$\bar{F}_\alpha(\alpha)<\frac{1}{\sqrt{2\pi}\,\alpha}e^{\frac{-\alpha^2}{2}},$$

for all $\alpha>0$, which appears to be a significantly faster decay. However, (22) and Figure 7 consider the decay with respect to N, and the effect of the mean $(\mu_{\alpha_N}=N\mu_\alpha)$

A Heavy-Tailed Example

It is natural to ask how heavy tailed the distribution of $f_a(a)$ can be and still lead to median stability

$$\lim_{N\to\infty} \text{median}(x_k) = 0.$$

To illustrate an extreme case, consider the $f_a(a)$ probability distribution to be given by

$$f_a(a) = \begin{cases} \dfrac{2}{\pi\gamma}\dfrac{1}{1+(a/\gamma)^2} & \text{for } a \geq 0, \\ 0 & \text{for } a < 0, \end{cases} \qquad \text{(S5)}$$

where $\gamma > 0$ is a real-valued parameter. Figure S4(a) illustrates this distribution on a log-log scale for three choices of the parameter γ. The distribution is equal to the magnitude of a Cauchy distribution, and all of the moments of this distribution (including the mean) are infinite. This is also clear from the a^{-2} power law decay in the tail shown in Figure S4(a). The calculation of the $f_\alpha(\alpha)$ distribution is given by (9), and in this case is

$$\begin{aligned} f_\alpha(\alpha) &= \frac{2}{\pi}\frac{e^\alpha/\gamma}{1+(e^\alpha/\gamma)^2} = \frac{2}{\pi}\frac{e^{(\alpha-\ln(\gamma))}}{1+e^{2(\alpha-\ln(\gamma))}} \\ &= \frac{1}{\pi}\frac{1}{\cosh(\alpha-\ln(\gamma))} = \frac{1}{\pi}\text{sech}(\alpha-\ln(\gamma)). \end{aligned} \qquad \text{(S6)}$$

This distribution, without the γ parameter, is known in statistics literature as a hyperbolic secant distribution and has been studied for nearly 100 years [S5]–[S8]. Most of the main properties of the distribution can be found in [S9]. The applications of the hyperbolic secant distribution are not all that common [S9], [S10]. There is a range of generalizations to the distribution, with application to specific domains in finance and actuarial statistics (see [S8], [S9], and [S11]).

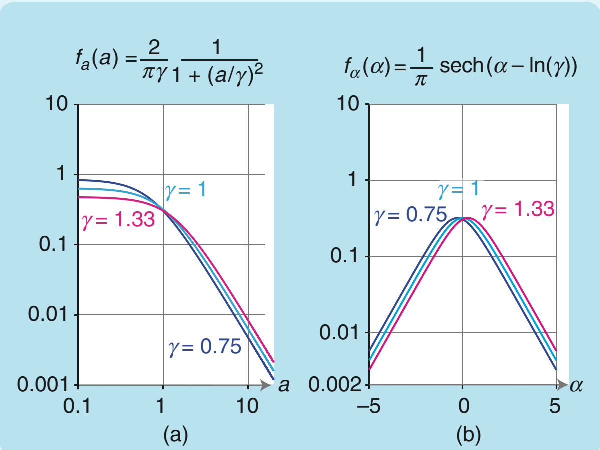


FIGURE S4 Heavy-tailed example probability distributions. (a) The $f_a(a)$ distribution on a log-log scale. The linear decay of the tail on the log-log plot shows a power law characteristic with decay a^{-2}. All moments of the $f_a(a)$ distribution are infinite. (b) The $f_\alpha(\alpha)$ distribution on a log-linear scale. All moments of the $f_\alpha(\alpha)$ distribution are finite.

Figure S4(b) shows the probability density function of the $f_\alpha(\alpha)$ distribution for three choices of γ. The exponential decay of the probability density function is clear from the log-linear plot. All the moments of this distribution are finite, and the moment-generating function (for $\gamma = 1$) is

$$\phi_\alpha(\lambda) = \frac{1}{\cos(\pi\lambda/2)}, \quad |\lambda| < 1. \qquad \text{(S7)}$$

The symmetry of (S6) about $\alpha = \ln(\gamma)$ shows that

$$\mu_\alpha = \ln(\gamma).$$

By applying Theorem 4,

$$\lim_{N\to\infty} \text{median}(x_N) = 0 \iff \gamma < 1.$$

Note, however, that for all $\gamma > 0$,

$$\mathcal{E}[x_N] = \infty$$

for all N. This is an extreme example of an unstable mean. The symmetry of the $f_\alpha(\alpha)$ distribution implies that the median and mean are equal and the median of the state x_N can therefore be given analytically:

$$\text{median}(x_N) = e^{\text{median}(\zeta_N)} = e^{\text{mean}(\zeta_N)} = e^{N\ln(\gamma)} = \gamma^N.$$

For illustration (and in comparison to Figure 6), 500 random trajectories of $\zeta_N = \ln(x_N)$, for $N = 1, \ldots, 100$ are shown in Figure S5. The predicted evolution of the median of ζ_N is compared to a sample-based estimate and found to be accurate. As a result of the very heavy-tailed nature of the $f_a(a)$ distribution, the range of the ζ_N trajectories in Figure S5 is much greater than in the normal/log-normal case shown in Figure 6.

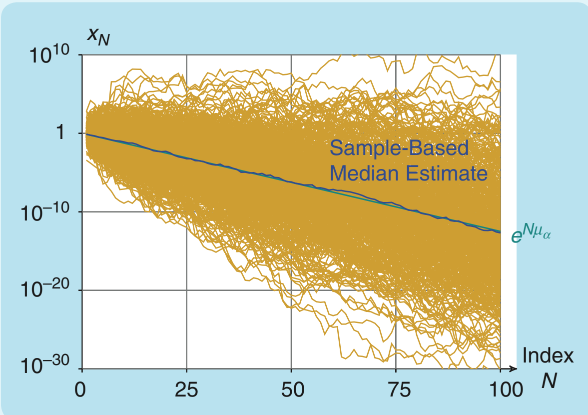


FIGURE S5 A total of 500 sample trajectory simulations for $\gamma = 0.75$. The median of the ζ_N distribution is compared with a sample estimate of the median.

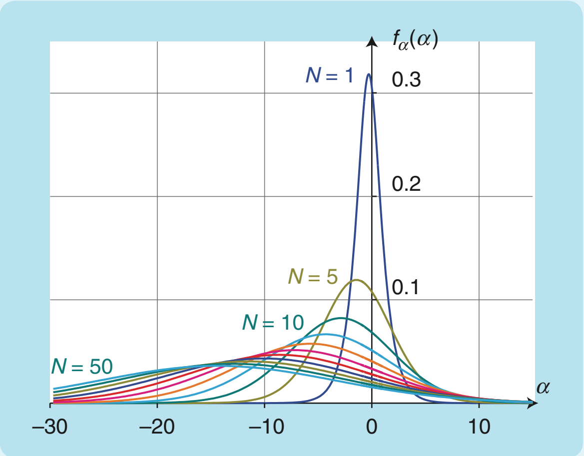


FIGURE S6 An evolution of the probability distribution function of ζ_N for $N = 1, 5, \ldots, 50$. The initial $(N = 1)$ distribution is that which is shown in the right plot in Figure S4 for $\gamma = 0.75$.

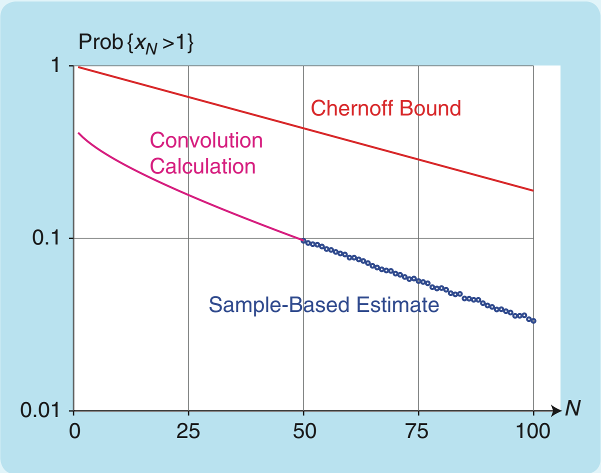


FIGURE S7 A comparison between the Chernoff bound (red) given in (S8)–(S10) and $\text{Prob}\{x_N > 1\}$ as a function of N for $\gamma = 0.75$. For N up to 50, the numerical evaluation (via Chebfun [S12]) of tail of the N-fold convolution of $f_\alpha(\alpha)$ gives the actual probability and is shown (magenta). For $N > 50$ (where numerical limitations prevent the numerical probability calculation), the probability is estimated from 100,000 sample trajectories (blue circles).

The behavior of the distribution of ζ_N as $N \longrightarrow \infty$ is given by the distribution of the N-fold sum of random variables, α_k, with each drawn from the $f_\alpha(\alpha)$ distribution. The distribution of α_N is the N-fold convolution of $f_\alpha(\alpha)$, which was numerically calculated (using the Chebfun Matlab Toolbox [S12]) in Figure S6. The characteristics of a sum of hyperbolic secant random variables were first studied in [S13].

As the moment-generating function for $f_\alpha(\alpha)$ in (S7) is finite in a range around zero, the probability that $x_N > 1$ decays to zero exponentially as $N \to \infty$. The bound can be calculated from the moment-generating function in (S7) and Lemma 5

$$\text{Prob}\left\{\prod_{k=1}^{N} x_k > 1\right\} \leq e^{-cN}, \tag{S8}$$

where

$$c = -\lambda^* \ln(\gamma) + \ln(\cos(\pi\lambda^*/2)) \tag{S9}$$

and

$$\lambda^* = \frac{2}{\pi}\arctan\left(\frac{-2\ln(\gamma)}{\pi}\right). \tag{S10}$$

Figure S7 shows this bound. The actual probably can be calculated numerically from the distributions in Figure S6 and estimated from samples in the simulation in Figure S5. Both of these comparisons are made and indicate that the exponent in the Chernoff bound is tight; however, the bound itself could be divided by a factor of at least two.

This is an extreme example, and it is interesting to put it into the context of a simple investment finance problem. Consider the accumulated return on an investment with an independent identically distributed random rate of return at every time step. In this context, x_0 is the initial investment, $f_a(a)$ is the probability distribution of the rate of return at each time step, and x_N is the investment value after N time steps. Then this example is a case where the expected rate of return is infinite for each of the time steps, and yet, the probability of making a profit after N time steps decays exponentially to zero as N increases.

REFERENCES

[S5] R. A. Fischer, "On the 'probable error' of a coefficient of a correlation deduced from a small sample," *Metron*, vol. 1, no. 4, pp. 3–32, 1921.

[S6] E. L. Dodd, "The frequency law of a function of variables with given frequency laws," *Ann. Math.*, vol. 27, no. 1, pp. 12–20, 1925. doi: 10.2307/1967829.

[S7] E. Roa, "A number of new generating functions with applications to statistics," Ph.D. dissertation, Univ. of Michigan, Ann Arbor, 1924.

[S8] W. Perks, "On some experiments in the graduation of mortality statistics," *J. Inst. Actuaries*, vol. 63, no. 1, pp. 12–57, 1932. doi: 10.1017/S0020268100046680.

[S9] M. J. Fischer, *Generalized Hyperbolic Secant Distributions With Applications to Finance*. New York: Springer-Verlag, 2014.

[S10] P. Ding, "Three occurrences of the hyperbolic-secant distribution," *Amer. Statist.*, vol. 68, no. 1, pp. 32–35, 2014. doi: 10.1080/00031305.2013.867902.

[S11] W. Harkness and M. Harkness, "Generalized hyperbolic secant distributions," *Amer. Statist. Assoc. J.*, vol. 63, no. 321, pp. 329–337, Mar. 1968. doi: 10.2307/2283852.

[S12] T. Driscoll, N. Hale, and L. Trefethen, Eds., *Chebfun Guide*. Oxford: Pafnuty Publications, 2014.

[S13] W. Baten, "The probability law for the sum of n independent variables, each subject to the law $(1/(2h))\text{sech}(\pi x/(2h))$," *Bull. Amer. Math. Soc.*, vol. 40, no. 4, pp. 284–290, 1934.

becoming more negative as N increases is countered to some extent by the convolution with $f_\alpha(\alpha)$ broadening the distribution (as it evolves with increasing N).

This approach also shows that as $N \to \infty$, the probability that x_N exceeds any arbitrarily small number goes to zero. The following lemma states this more formally.

Lemma 2: Log-Normal Convergence to Zero

Assume that $a \sim \mathcal{LN}$, $\mu_\alpha < 0$, and for simplicity, $x_0 = 1$. For any $\epsilon > 0$,

$$\lim_{N \to \infty} \text{Prob}\{x_N > \epsilon\} = 0.$$

Proof of Lemma 2

Again write

$$\begin{aligned}\text{Prob}\{x_N > \epsilon\} &= \text{Prob}\{\alpha_N > \ln(\epsilon)\}\\ &= \left(1 - \text{erf}\left(\frac{\ln(\epsilon) - N\mu_\alpha}{\sqrt{2N\sigma_\alpha^2}}\right)\right).\end{aligned}$$

For all $N > \ln(\epsilon)/\mu_\alpha$, the argument of the erf function is positive and increasing without bound as a function of N. As

$$\lim_{x \to \infty} \text{erf}(x) = 1,$$

the result follows. ■

It is interesting that this result holds, even though $\text{mean}(x_N)$ may be growing to $+\infty$.

More General Distributions

The tail-probability results in the last section can be generalized to a wider range of distributions, and the strength of the bound depends upon the assumptions placed on the underlying distribution. Exponential bounds are still possible for a wide range of distributions, and some examples are provided in the following lemmas. A decaying bound is available under the assumptions that $\mu_\alpha < 0$ and the distribution $f_\alpha(\alpha)$ has a finite variance, σ_α^2. These conditions are weaker than those considered for median stability in Theorem 4. Under the assumption of the pairwise independence of the α_k variables, which is satisfied here by assumption, Cantelli's inequality [33] leads to the following bound.

Lemma 3

Assume that $\alpha \sim f_\alpha(\alpha)$ has a finite mean, $\mu_\alpha < 0$, and a finite variance, σ_α^2. Then,

$$\text{Prob}\{x_N > 1\} \leq \frac{1}{1 + N\dfrac{\mu_\alpha^2}{\sigma_\alpha^2}}. \quad (23)$$

This bound is also illustrated in Figure 7. In this general case, the distribution converges to zero with a $1/N$ rate. A more general version of Lemma 2 is immediate.

Lemma 4

Assume that $\alpha \sim f_\alpha(\alpha)$ has a finite mean $\mu_\alpha < 0$ and a finite variance σ_α^2. Then, for any $\epsilon > 0$,

$$\lim_{N \to \infty} \text{Prob}\{x_N > \epsilon\} = 0.$$

As $x_N > 0$ for all N, Lemma 4 states that the random variable x_N converges in probability to the degenerate random variable $x = 0$. The assumption that the variance of $f_\alpha(\alpha)$ is finite is very weak and is satisfied in many cases, even when the corresponding $f_a(a)$ random variable does not have finite moments. Given this assumption, it is clear that, for systems given by (5), median stability implies convergence of x_N in probability to $x = 0$.

The Cantelli inequality (Lemma 3) requires the fewest assumptions on the distribution $f_\alpha(\alpha)$ and has only a decay rate approximating $1/N$ for large N. For smaller values of N, this bound is actually more accurate than some of the other bounds. There exist distributions for which the Cantelli bound is tight, and in some cases, it is not possible to find a better bound.

Tighter bounds are possible if higher moments are known, and the next most significant assumption is that $\alpha \sim f_\alpha(\alpha)$ comes from a distribution that has a finite, moment-generating function within an open interval around zero. This implies that all moments of the distribution are bounded. Distributions not satisfying this assumption can be defined as being *heavy tailed*. Note that this assumption is on the α random variable—the $a \sim f_a(a)$ may be heavy tailed, and "A Heavy-Tailed Example" gives a rather extreme example. The moment-generating function is defined as

$$\phi_\alpha(\lambda) := \mathcal{E}[e^{\lambda\alpha}],$$

which is assumed to be finite within a region of the origin

$$\phi_\alpha(\alpha) < \infty, \quad \text{for all } |\lambda| \leq \beta,\ \beta > 0. \quad (24)$$

The moment-generating function is used in the calculation of the Chernoff bound on the tail on the distribution. In this case,

$$\text{Prob}\{\alpha - \mu_\alpha > t\} \leq e^{-\left(\sup\limits_{\lambda \in [0,\beta]} (\lambda(t - \mu_\alpha) - \ln(\phi_\alpha(\lambda)))\right)}.$$

This is not the most general form of the Chernoff bound, and other forms give tighter bounds for low values of t. However, the behavior of the $f_{x_N}(x)$ distribution for large values of N is the primary concern of this article and is addressed by the simpler bound given above. This can be applied directly to the evolution of the $f_\alpha(\alpha)$ distribution in the following way.

Lemma 5

Assume that $\alpha \sim f_\alpha(\alpha)$ has a moment-generating function that is finite over an open interval, including zero (24). Also assume that $\mu_\alpha < 0$. Then,

The conditions for median, mean, and variance stability are different, and it is natural to consider which is more appropriate for use in any particular problem.

$$\mathrm{Prob}\{x_N > 1\} = \mathrm{Prob}\left\{\prod_{k=1}^{N} a_k > 1\right\} \leq \mathrm{e}^{-cN},$$

where

$$c = \sup_{\lambda \in [0,\beta]} -\lambda\mu_\alpha - \ln(\phi_\alpha(\lambda)).$$

The proof of Lemma 5 follows immediately from substituting

$$\phi_{\alpha_N}(\lambda) = \phi_\alpha(\lambda)^N$$

and $t = 0$ into the Chernoff bound. Thus, the existence of a finite, moment-generating function around zero implies an exponential decay of the tail bound of the distribution of x_N as $N \to \infty$. However, calculating the constant for the exponent requires knowledge of the moment-generating function.

The Chernoff bound (Lemma 5) is also displayed in Figure 7. The exponent on this bound is the closest single exponent bound for the actual tail distribution. Tighter exponential bounds require sums of exponentials. This bound can also be tightened by scaling by 0.5 (see the discussion in [34] and the references therein for further details). However, having $\phi_\alpha(\alpha)$ finite in an open interval of the origin uniquely determines the probability density function and the corresponding cumulative probability density function. This can then be integrated numerically to calculate the required probability.

DISCUSSION

The goal of this article was to precisely specify and illustrate the conditions for the stability of the median, mean, and variance in discrete-time stochastic feedback settings. The discrete-time setting enables a far wider range of distributions to be considered than is possible in the continuous-time case, and it is, at the same time, relevant to a wide range of problems. The focus on the scalar variable case is, of course, much more restrictive and has allowed precise statements to be made about the probability of the distributions of solutions to the difference equations. This is particularly true for median statistics.

The differences between the stability conditions arise because of the heavy-tailed nature of the resulting distributions. This allows the phenomenon of the mean growing exponentially while the distribution converges exponentially to zero to arise. Note that the stochastic component of the system need not be heavy tailed for this to be observed; it suffices that the effect of the stochastic component is integrated via a feedback interconnection with a dynamical system.

The variance stability condition is a simple case of the more widely known mean-square stability criterion from the 1970s [5]. This condition has the advantage that it is also exact for the multivariable case. However, it is acknowledged that mean-square stability is a strong form of stability [6, p. 136]. The results in this article emphasize this point, particularly in comparison to median stability.

The conditions for median, mean, and variance stability are different, and it is natural to consider which is more appropriate for use in any particular problem. Very different answers can arise from the exact statement of the problem and can easily lead to interpretations which are—at least from a cursory point of view—contradictory. For example, in an investment problem, there is a relatively wide range of circumstances in which a return on investment will have an expected value greater than one (and consequently the expected profit grows with time) and yet in which the probability of making any profit at all decays to zero. The equivalent conditions in a population dynamics or epidemiological context would indicate that an expected survival rate may be greater than one, and yet the probability of extinction is also going to one.

These apparent paradoxes illustrate that, in stochastic feedback situations, seemingly similar questions may have widely divergent answers. Thus, it is important to pose the correct measure of stability in the problem formulation and its analysis. The increasing use of interconnected feedback networks (and particularly those where online data-based updating leads to stochasticity in the feedback components) requires the careful selection of analysis criteria and design methods.

ACKNOWLEDGMENTS

The authors would like to thank Tryphon Georgiou, Sean Meyn, Vikram Krishnamurthy, and Subhrakanti Dey for useful discussions on this topic. The work of Bassam Bamieh is partially supported by National Science Foundation Awards CMMI-1763064 and ECCS-1932777.

AUTHOR INFORMATION

Roy S. Smith (rsmith@control.ee.ethz.ch) is a professor of electrical engineering at the Swiss Federal Institute of Technology, Zürich, 8092, Switzerland, which he joined in 2011. From 1990 to 2010, he was on the faculty of the University of California. He received the Ph.D. degree from the California Institute of Technology in 1990 and the M.Eng. and B.Eng. degrees from the University of Canterbury, New Zealand,

in 1980. He has been a long-time consultant to the NASA Jet Propulsion Laboratory and has industrial experience in automotive control and power system design. His research interests are broadly focused on the modeling, identification, and control of uncertain systems. Control application domains of interest include chemical processes, flexible structure vibration, spacecraft and vehicle formations, semiconductor fabrication facilities, automotive engines, Mars aeromaneuvering entry design, energy management in buildings, and thermoacoustic machines. He is an associate fellow of the American Institute for Aeronautics and Astronautics and a Fellow of IEEE and the International Federation of Automatic Control.

Bassam Bamieh is a professor of mechanical engineering at the University of California at Santa Barbara (UCSB), Santa Barbara, 93106, California, USA. He received his B.Sc. degree in electrical engineering and physics from Valparaiso University, Indiana, in 1983 and the M.Sc. and Ph.D. degrees in electrical and computer engineering from Rice University, Houston, Texas, in 1986 and 1992, respectively. Prior to joining UCSB in 1998, he was an assistant professor in the Department of Electrical and Computer Engineering and the Coordinated Science Laboratory at the University of Illinois at Urbana-Champaign from 1991 to 1998. His research interests include robust and optimal control, distributed and networked control and dynamical systems, shear flow transition and turbulence, quantum control, and the use of feedback in thermoacoustic energy-conversion devices. He is a past recipient of the IEEE Control Systems Society G.S. Axelby Outstanding Paper Award (twice), an AACC Hugo Schuck Best Paper Award, and a National Science Foundation CAREER Award. He is a Distinguished Lecturer of the IEEE Control Systems Society and a Fellow of IEEE and the International Federation of Automatic Control.

REFERENCES

[1] A. Rosenbloom, "Analysis of randomly time varying linear systems," Ph.D. dissertation, Univ. of California, Los Angeles, 1954.

[2] J. C. Samuels, "On the mean square stability of random linear systems," *IRE Trans. Circuit Theory*, vol. 6, no. 5, pp. 248–259, 1959. doi: 10.1109/TCT.1959.1086608.

[3] R. Kalman, "Control of randomly varying linear dynamical systems," in *Proc. Symp. Appl. Math.*, 1962, vol. 13, pp. 287–298.

[4] J. C. Willems and G. L. Blankenship, "Frequency domain stability criteria for stochastic systems," *IEEE Trans. Autom. Control*, vol. 16, no. 4, pp. 292–299, 1971. doi: 10.1109/TAC.1971.1099733.

[5] J. Willems, "Mean square stability criteria for stochastic feedback systems," *Int. J. Syst. Sci.*, vol. 4, no. 4, pp. 545–564, 1973. doi: 10.1080/00207727308920036.

[6] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, vol. 15. Philadelphia: SIAM, 1994.

[7] K. Åström, "On a first-order stochastic differential equation," *Int. J. Control*, vol. 1, no. 4, pp. 301–326, 1965. doi: 10.1080/00207176508905484.

[8] G. Blankenship, "Stability of linear differential equations with random coefficients," *IEEE Trans. Autom. Control*, vol. 22, no. 5, pp. 834–838, Oct. 1977. doi: 10.1109/TAC.1977.1101612.

[9] S. Ethier, "The Kelly system maximizes median fortune," *J. Appl. Probab.*, vol. 41, no. 4, pp. 1230–1236, 2004. doi: 10.1017/S002190020002101X.

[10] H. Kesten, "Random difference equations and renewal theory for products of random matrices," *Acta Math.*, vol. 131, pp. 207–248, Dec. 1973. doi: 10.1007/BF02392040.

[11] C. M. Goldie, "Implicit renewal theory and tails of solutions of random equations," *Ann. Appl. Probab.*, vol. 1, no. 1, pp. 126–166, 1991. doi: 10.1214/aoap/1177005985.

[12] A. Brandt, "The stochastic equation $y_{n+1} = a_n y_n + b_n$ with stationary coefficients," *Adv. Appl. Probab.*, vol. 18, no. 1, pp. 211–220, 1986.

[13] B. de Saporta, "Tail of the stationary solution of the stochastic equation $y_{n+1} = a_n y_n + b_n$ with Markovian coefficients," *Stochastic Processes Appl.*, vol. 15, no. 12, pp. 1954–1978, 2005.

[14] S. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[15] D. Buraczewski, E. Damek, and T. Mikosch, *Stochastic Models with Power-Law Tails: The Equation $X = AX + B$*. New York: Springer-Verlag, 2016.

[16] V. Protasov and R. Jungers, "Lower and upper bounds for the largest Lyapunov exponent of matrices," *Linear Algebra Appl.*, vol. 438, no. 11, pp. 4448–4468, 2013. doi: 10.1016/j.laa.2013.01.027.

[17] D. Sutter, O. Fawzi, and R. Renner, "Bounds on Lyapunov exponents via entropy accumulation," 2019, arXiv:1905.0327v3.

[18] M. Milisavljević and E. I. Verriest, "Stability and stabilization of discrete systems with multiplicative noise," in *Proc. Eur. Control Conf.*, 1997, pp. 3503–3508. doi: 10.23919/ECC.1997.7082656.

[19] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, 2004. doi: 10.1109/TAC.2004.834121.

[20] N. Elia, "Remote stabilization over fading channels," *Syst. Control Lett.*, vol. 54, no. 3, pp. 237–249, 2005. doi: 10.1016/j.sysconle.2004.08.009.

[21] N. Elia, "Emergence of power laws in networked control systems," in *Proc. IEEE Conf. Decision Control*, 2006, pp. 490–495. doi: 10.1109/CDC.2006.377020.

[22] J. Wang and N. Elia, "Distributed averaging under constraints on information exchange: Emergence of Lévy flights," *IEEE Trans. Autom. Control*, vol. 57, no. 10, pp. 2435–2449, 2012. doi: 10.1109/TAC.2012.2186093.

[23] Y. Mo and B. Sinopoli, "Kalman filtering with intermittent observations: Tail distribution and critical value," *IEEE Trans. Autom. Control*, vol. 57, no. 12, pp. 677–689, 2012. doi: 10.1109/TAC.2011.2166309.

[24] S. Dey and L. Schenato, "Heavy-tails in Kalman filtering with packet losses: Confidence bounds vs second moment stability," in *Proc. Eur. Control Conf.*, 2018, pp. 1480–1486. doi: 10.23919/ECC.2018.8550235.

[25] A. Rantzer, "Concentration bounds for single parameter adaptive control," in *Proc. Amer. Control Conf.*, 2018, pp. 1862–1866. doi: 10.23919/ACC.2018.8431891.

[26] E. Hoversten, R. Harger, and S. Halme, "Communication theory for the turbulent atmosphere," *Proc. IEEE*, vol. 58, no. 10, pp. 1626–1650, 1970. doi: 10.1109/PROC.1970.7986.

[27] J. Aitchison and J. Brown, *The Lognormal Distribution With Special Reference to its use in Economics*. Cambridge, U.K.: Cambridge Univ. Press, 1957.

[28] D. Finney, "On the distribution of a variate whose logarithm is normally distributed," *Suppl. J. Roy. Statist. Soc.*, vol. 7, no. 2, pp. 155–161, 1941. doi: 10.2307/2983663.

[29] G. Shellard, "Estimating the product of several random variables," *J. Amer. Statist. Assoc.*, vol. 47, no. 258, pp. 216–221, 1952. doi: 10.1080/01621459.1952.10501165.

[30] L. A. Goodman, "The variance of the product of K random variables," *J. Amer. Statist. Assoc.*, vol. 57, no. 297, pp. 54–60, Mar. 1962. doi: 10.2307/2282440.

[31] P. Hall, "On the limiting behavior of the mode and median of a sum of independent random variables," *Ann. Probab.*, vol. 8, no. 3, pp. 419–430, 1980. doi: 10.1214/aop/1176994717.

[32] B. Bercu, B. Delyon, and E. Rio, *Concentration Inequalities for Sums and Martingales*. New York: Springer-Verlag, 2015.

[33] F. Cantelli, "Sui confini della probabilità," *Atti del Congr.o Internazionale dei Matematici*, vol. 6, pp. 47–59, Sept. 1928.

[34] S.-H. Chang, P. C. Cosman, and L. B. Milstein, "Chernoff-type bounds for the Gaussian error function," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 2939–2944, 2011. doi: 10.1109/TCOMM.2011.072011.100049.