



TITLE:

Stable Process Approach to Analysis of Systems Under Heavy-Tailed Noise: Modeling and Stochastic Linearization

AUTHOR(S):

Kashima, Kenji; Aoyama, Hiroki; Ohta, Yoshito

CITATION:

Kashima, Kenji ...[et al]. Stable Process Approach to Analysis of Systems Under Heavy-Tailed Noise: Modeling and Stochastic Linearization. IEEE Transactions on Automatic Control 2019, 64(4): 1344-1357

ISSUE DATE:

2019-04

URL:

<http://hdl.handle.net/2433/263835>

RIGHT:

This is an open access article.

Stable Process Approach to Analysis of Systems Under Heavy-Tailed Noise: Modeling and Stochastic Linearization

Kenji Kashima¹, Member, IEEE, Hiroki Aoyama², and Yoshito Ohta³, Senior Member, IEEE

Abstract—The Wiener process has provided a lot of practically useful mathematical tools to model stochastic noise in many applications. However, this framework is not enough for modeling extremal events, since many statistical properties of dynamical systems driven by the Wiener process are inevitably Gaussian. The goal of this work is to develop a framework that can represent a heavy-tailed distribution without losing the advantages of the Wiener process. To this end, we investigate models based on *stable processes* (this term “stable” has nothing to do with “dynamical stability”) and clarify their fundamental properties. In addition, we propose a method for stochastic linearization, which enables us to approximately linearize static nonlinearities in feedback systems under heavy-tailed noise, and analyze the resulting error theoretically. The proposed method is applied to assessing wind power fluctuation to show the practical usefulness.

Index Terms—Extremal events, linearization, renewable energy, stochastic systems.

I. INTRODUCTION

IN MANY engineering applications, it is important to evaluate the effect of probabilistic uncertainty. There are many mathematical representations for dynamical noise signals. Among them, the Wiener process is the most classical and powerful one [1]. From a modeling point of view, the Wiener increment can be viewed as a continuous-time counterpart of the discrete-time independent and identically distributed random variables having finite variance, which is a standard expression

Manuscript received March 13, 2017; revised December 3, 2017 and April 4, 2018; accepted April 21, 2018. Date of publication May 30, 2018; date of current version March 27, 2019. This work was supported in part by the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (KAKENHI) under Grant 26289130 and in part by the Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, under Grant JPMJCR15K1. Recommended by Associate Editor S. Yuksel. (Corresponding author: Kenji Kashima.)

K. Kashima and Y. Ohta are with the Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan, and also with the Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Kawaguchi 332-0012 Japan (e-mail: kk@i.kyoto-u.ac.jp; yoshito_ohta@i.kyoto-u.ac.jp).

H. Aoyama is with the Research and Engineering Division Aero-Engine and Space Operations, IHI Corporation, Tokyo 135-8710, Japan (e-mail: hiroki_aoyama@ihi.co.jp).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TAC.2018.2842145

to represent probabilistic noise. From a technical point of view, many statistical properties of linear dynamical systems driven by the Wiener process conform to a *normal distribution*. This enables one to develop beautiful theories for analysis and synthesis, e.g., H^2 -control design and Kalman filtering [2].

On the other hand, in recent years, it has been pointed out that extremal events often cause severe damage in many situations. Thus, it is important to suitably handle such events to develop resilient systems. When we attempt to investigate this problem in the Wiener process setting, the aforementioned affinity to the normal distribution results in conclusions that are no longer meaningful, and even misleading. Actually, the Gaussian function, which is the probability density function of the normal distribution, has quickly decaying tails as seen in the popular 3σ rule (99.7% of samples lie within three standard deviations of the mean). For example, the statistical property of wind power fluctuation, which does not fit to the normal distribution because of its high probability of the extremal outlier [3], is known as an “umbrella curve” and regarded as a source of severe damage to power systems. In addition, the scale-free property attracts much attention in various scientific disciplines [4]. Mathematically, this is a power law of the probability distribution, which represents significantly frequent outliers. These facts imply that the Wiener process framework is not capable of capturing such *non-Gaussian* phenomena adequately. There are several interesting works to represent sudden shifts in control systems, e.g., Markovian jump [5] and Laplacian distributions [6], [7]. However, these treatments are rather different from the Wiener process case. As a result, the technical advantage arising from the Gaussianity would easily be lost.

A. Contribution and Literature Review

The goal of this work is to develop a framework that can represent heavy-tailed distributions arising from extremal events, without losing the aforementioned advantages of the Wiener process. To this end, we alternatively employ *stable processes* (this term “stable” has nothing to do with “dynamical stability” in control theory) [8]–[10]. Stable processes are a natural extension of the Wiener process. As a result, the technical advantages of the Wiener process setting can be retained considerably in comparison to other non-Gaussian frameworks. Nevertheless, the behavior of linear systems driven by a stable process is expressed in terms of the so-called stable distribution with a

non-Gaussianity parameter $\alpha \in (0, 2]$. This distribution is just a normal distribution when $\alpha = 2$, but obeys a power law otherwise. This enables us to model heavy-tailed distributions. In addition, beyond the linear systems analysis, we analyze systems equipped with static nonlinearity, such as saturation or a dead zone. In this paper, we investigate statistical properties of the output signal of such feedback systems driven by stable processes. In the main results, we propose a stochastic linearization method [11] to provide user-friendly modeling and analysis methods.

Some related works are listed here. In [12], linear–quadratic–Gaussian (LQG) theory was extended to another class of jump diffusions. However, stable processes can model completely different statistics, which require careful mathematical treatment. In addition, we also study the effect of nonlinearity. Next, in [11] and [13], stochastic linearization of systems driven by the Wiener process is thoroughly investigated. It should be emphasized that the extension to our setting is not straightforward. This is because the stable distribution with $\alpha \neq 2$ does not have an analytic representation for the probability distribution and finite second-order moment. Consequently, the error variance minimization, which was a crucial building block in the conventional methods, becomes nontrivial, or even ill-posed. We circumvent this issue by employing an alternative error metric. A preliminary version of this paper was presented in [14]. The original contribution of the present paper is summarized as follows: the results are presented in a more general (by means of multiple-input multiple-output generalized plant and various nonlinearities) and mathematically rigorous (in terms of the convergence in law) framework; the power spectrum density is characterized in a closed form in Theorem 1; the key parameter that determines the effect of non-Gaussianity is given explicitly in Theorem 3; a novel theoretical error bound is derived in Theorem 4; and, finally, in the application to wind power assessment where a more realistic situation is considered, a practically meaningful sensitivity analysis is conducted employing the presented techniques in Section V.

B. Organization

The rest of this paper is organized as follows. In Section II, we introduce the stable process and formulate our problem. In Section III, we characterize the asymptotic behavior of outliers and briefly explain the frequency domain modeling. In Section IV, we propose a stochastic linearization method for this class of feedback systems. In Section V, we verify the effectiveness of the proposed method via its application to renewable energy assessment. Some concluding remarks are given in Section VI. Proof of the Theorems are provided in the Appendixes.

Notation: The set of real numbers is \mathbb{R} and $j = \sqrt{-1}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space equipped with a natural filtration $\{\mathcal{F}_t\}_{t \geq 0}$. The expectation is denoted by \mathbb{E} . Only when we need to specify the underlying probability measure, or conditioning, will this be shown in the subscript. The indicator function $\mathbb{1}_S(x)$ is 1 if $x \in S$, and 0 otherwise. For stochastic process x_t , the convergence in law¹ to a random variable x_∞

¹ $\lim_{t \rightarrow \infty} \mathbb{P}[x_t \in B] = \mathbb{P}[x_\infty \in B]$ for any Borel set B with boundary ∂B satisfying $\mathbb{P}[x_\infty \in \partial B] = 0$; see [15] and [16].

is denoted by $x_t \xrightarrow{d} x_\infty$. The probability distribution of x_∞ is referred to as the stationary distribution of x_t . For positive real constants s, x , the gamma function $\Gamma(s)$ and scaled lower incomplete Gamma function $\Gamma_\ell(s, x)$ are defined by

$$\Gamma(s) := \int_0^\infty t^{s-1} e^{-t} dt$$

$$\Gamma_\ell(s, x) := \frac{1}{\Gamma(s)} \int_0^x t^{s-1} e^{-t} dt.$$

II. GENERALIZED WHITE NOISE AND PROBLEM FORMULATION

In this paper, we utilize a more general framework for random signals than the Wiener process. The definitions are introduced in Section II-A. Then, the main problem in this paper is formulated in Section II-B.

A. Stable Process

We begin with an extension of the normal distribution. There are several equivalent definitions for the normal distribution. Among them, we rely on the representation based on its characteristic function² (see [8]).

Definition 1: A real-valued random variable X is said to have a symmetric stable distribution with parameter $\alpha \in (0, 2]$ and $\sigma > 0$, or simply $X \sim \mathbf{S}\alpha\mathbf{S}(\alpha, \sigma)$, if its characteristic function satisfies

$$\mathbb{E}[\exp(j\nu X)] = \exp(-\sigma^\alpha |\nu|^\alpha), \quad \nu \in \mathbb{R}. \quad (1)$$

□

The parameter α represents the degree of non-Gaussianity. In particular, $\mathbf{S}\alpha\mathbf{S}(2, \sigma)$ obeys the normal distribution with variance $2\sigma^2$. As α becomes small, the resulting distribution has a heavier tail. Similarly to the standard deviation of the normal distribution, σ is a scaling parameter. The parameter is omitted when $\sigma = 1$ such as $X \sim \mathbf{S}\alpha\mathbf{S}(\alpha)$. For $X \sim \mathbf{S}\alpha\mathbf{S}(\alpha)$, we have

$$\kappa X \sim \mathbf{S}\alpha\mathbf{S}(\alpha, |\kappa|), \quad \kappa \in \mathbb{R} \quad (2)$$

by definition. We denote $X_t \xrightarrow{d} \mathbf{S}\alpha\mathbf{S}(\alpha, \sigma)$ when $X_t \xrightarrow{d} X_\infty$ such that $X_\infty \sim \mathbf{S}\alpha\mathbf{S}(\alpha, \sigma)$.

Next, we extend the notion to stochastic processes. The Wiener process can be generalized in a similar way to Definition 1.

Definition 2: A stochastic process L_t is said to be a (normalized) α -stable process with parameter α if

$$L_t \sim \mathbf{S}\alpha\mathbf{S}\left(\alpha, t^{1/\alpha}\right). \quad (3)$$

□

Here, $\sqrt{2}L_t$ with $\alpha = 2$ is the Wiener process. Finally, similarly to the Itô integral for the Wiener process, we can introduce a stochastic differential equation associated with stable processes [8, Ch. 3].

²There is another definition for which (1) is replaced by $\mathbb{E}[\exp(j\nu X)] = \exp(-\sigma^\alpha |\nu|^\alpha / \alpha)$.

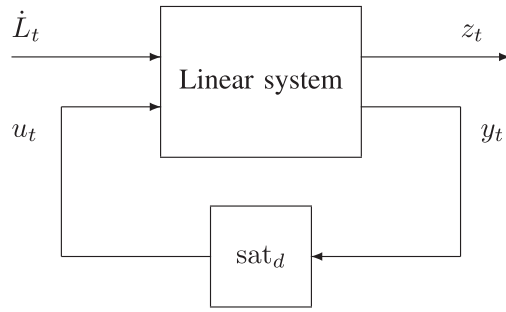


Fig. 1. Block diagram for Problem 1.

B. Problem Formulation

The goal of this paper is to derive a simple method to find statistical properties of the output signal of feedback systems driven by stable processes. In particular, we focus on systems consisting of linear systems and saturation nonlinearity in Fig. 1

$$dx_t = Ax_t dt + Bu_t dt + b dL_t \quad (4)$$

$$z_t = c_z x_t \quad (5)$$

$$y_t = C_y x_t \quad (6)$$

$$u_t = \text{sat}_d(y_t) \quad (7)$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^r$, $z_t \in \mathbb{R}$, and $y_t \in \mathbb{R}^r$ are the state, control input, evaluation output, observation variables, respectively, $b, c_z^\top \in \mathbb{R}^n$, and A, B , and C_y are matrices of compatible dimension. The stochastic input L_t is an α -stable process with parameter α and

$$\text{sat}_d(y) := \begin{bmatrix} \text{sat}_{d_1}(y_1) \\ \vdots \\ \text{sat}_{d_r}(y_r) \end{bmatrix} \quad (8)$$

with the given threshold vector $d \in [0, \infty)^r$ and

$$\text{sat}_{d_i}(y_i) := \begin{cases} -d_i, & y_i < -d_i \\ y_i, & |y_i| \leq d_i \\ d_i, & y_i > d_i \end{cases} \quad (9)$$

where y_j (respectively, d_j) is the j th element of y (respectively, d); see Fig. 2(a). Then, we are ready to state our problem.

Problem 1: Consider the feedback system with saturation given above. Suppose that $(A + BKC_y)$ is Hurwitz for any $K \in \{\text{diag}(k_1, \dots, k_r) : k_i \in [0, 1]\}$ and that x_t converges in law to a random variable x_∞ defined on \mathbb{R}^n as $t \rightarrow \infty$. Then, find the stationary probability distribution of z_t . \square

For the case with $\alpha = 2$ investigated in [11], it is not difficult to compute the desired probability distribution via the Monte Carlo method. However, it is challenging to capture the effect of rare events based on the Monte Carlo simulation [17]. Therefore, it should be emphasized that the Monte Carlo method is not enough for our purpose to quantify the effect of heavy-tailed noise represented by the stable processes with $\alpha < 2$. In addition, it is difficult to obtain insight into the parameter dependence from the obtained simulation result (see Section V). Another approach, such as [18], is not necessarily computation-

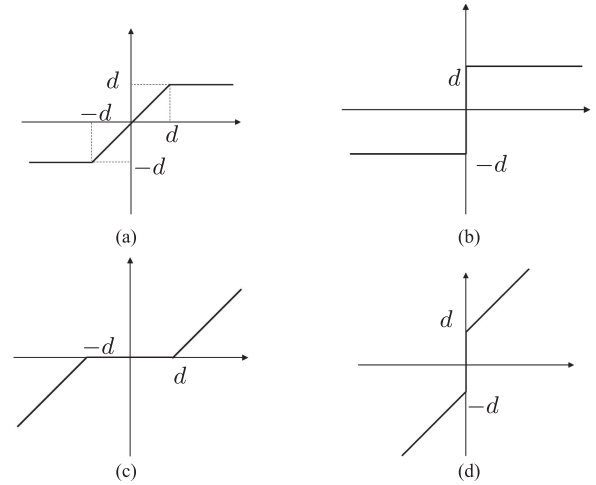


Fig. 2. Static nonlinearities. (a) Saturation. (b) Relay. (c) Deadzone. (d) Friction.

ally tractable for multidimensional systems. Obviously, simply ignoring the saturation ($y = u$) is problematic. In view of this, we construct a stochastic linearization method [19], [20], where the saturation is approximated by a suitable linear function.

Derivation of existence criteria of the invariant measure is out of the scope of this paper; see [21] and [22] for the Wiener process. The stability assumption with $K = I$ implies that the feedback system is stable when the saturation is not active. Though we confine ourselves to saturation in Problem 1, the method proposed in this paper is applicable to other nonlinearity blocks (see Section IV-E).

III. MODELING SYSTEMS UNDER HEAVY-TAILED NOISE

A. Sample Path Properties

In principle, the probability density function can be uniquely determined by the inverse Fourier transform of the characteristic function. For $\alpha = 2$, the probability density function is again a Gaussian function, which can easily be sampled. Though the probability density function of the stable distribution with $\alpha < 2$ has no analytic representation, it is possible to generate samples by the following property [8, Proposition 1.7.1]:

$$X = \frac{\sin(\alpha U)}{(\cos(U))^{\frac{1}{\alpha}}} \left(\frac{\cos((1-\alpha)U)}{E} \right)^{\frac{1-\alpha}{\alpha}} \sim S\alpha S(\alpha) \quad (10)$$

where U and E are mutually independent and, respectively, a uniform random variable on $(-\frac{\pi}{2}, \frac{\pi}{2})$ and an exponential random variable with intensity 1. However, it is still difficult to precisely evaluate the statistical properties resulting from outliers via the Monte Carlo simulation. This is because the statistical property of outliers is not negligible, explained as follows (see [8, Property 1.2.15]):

Proposition 1: For $X \sim S\alpha S(\alpha)$ with $\alpha \in (1, 2)$, we have

$$\lim_{\lambda \rightarrow \infty} \lambda^\alpha \mathbb{E}[X > \lambda] = C_\alpha := \frac{1-\alpha}{2\Gamma(2-\alpha)\cos(\frac{\pi\alpha}{2})}. \quad (11)$$

\square

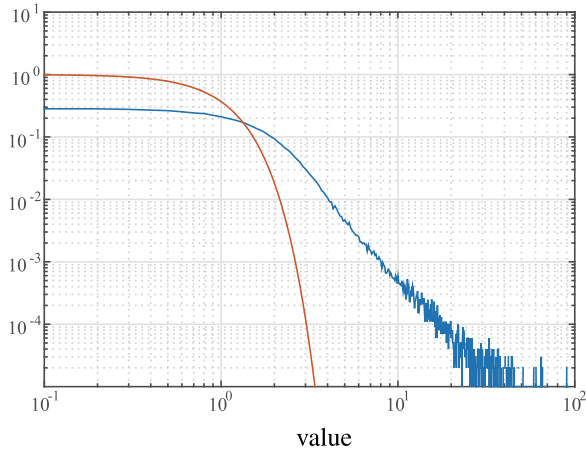


Fig. 3. log-log plot of the histogram of the stable distribution $S_{\alpha}S(1.6)$ and Gaussian density function of $S_{\alpha}S(2)$.

This is a power law for stable random variables. In comparison to the normal distribution for $\alpha = 2$, extremal values occur with a significantly larger probability. Fig. 3 shows the log–log plot of the tail distribution. Such a linear decay is often called scale free. As a direct consequence of this power law, the moment $\mathbb{E}[|X|^p]$ is not finite for $p \geq \alpha$ (see [9, Example 25.10]).

Proposition 2: For $X \sim S_{\alpha}S(\alpha)$ with $\alpha \in (0, 2)$ and $p \in (-1, \alpha)$, we have

$$\mathbb{E}[|X|^p] = \frac{2^p \Gamma\left(\frac{1+p}{2}\right) \Gamma\left(1 - \frac{p}{\alpha}\right)}{\sqrt{\pi} \Gamma\left(1 - \frac{p}{2}\right)}. \quad (12)$$

□

Concerning the continuity of sample paths, we have the following property [9].

Proposition 3: Every α -stable process is a Lévy process, that is, a càdlàg process (right continuous with left limits) having stationary and independent increment. □

Lévy processes allow discontinuity of sample paths although the Wiener process is known to have almost sure continuous sample paths. This causes a significant difference in terms of the trajectories. This proposition, combined with Definition 1, means

$$L_{t+dt} - L_t \sim (dt)^{1/\alpha} S_{\alpha}S(\alpha) \quad (13)$$

for any α , $dt > 0$, which allows one to employ the Euler–Maruyama method [23] for simulating dynamical systems driven by stable processes [8].

Example 1: Let us consider

$$dn_t = -0.001n_t dt + 1.365dL_t \quad (14)$$

where L_t is an α -stable processes, which is used as a wind power fluctuation model (with time t [s]) in Section V. Fig. 4 shows sample paths for $\alpha = 2$, 1.6, which are scaled such that they have identical stationary 1st order moment (see Proposition 2 and Theorem 2). Even a simulation over such a short time interval suggests that we can use stable processes with $\alpha < 2$ to represent stochastic noises that exhibit frequent and large sudden shifts.

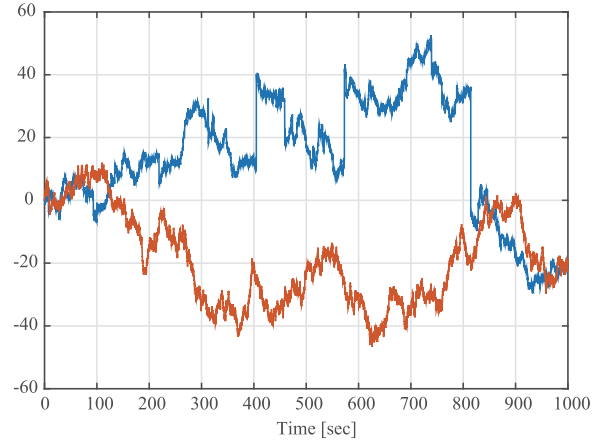


Fig. 4. Scaled sample path for $\alpha = 2$ (red) and 1.6 (blue).

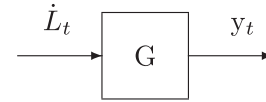


Fig. 5. Block diagram of the Ornstein–Uhlenbeck process.

B. Frequency-Domain Modeling

Before we proceed, let us examine the advantage of the use of stable processes from a modeling point of view.

In order to model noise signals from time-series data, it is standard to use the *Ornstein–Uhlenbeck processes* driven by Wiener processes as follows. Let us assume that the noise one wants to model is represented as y_t defined by³

$$dx_t = Ax_t dt + b dL_t, \quad y_t = cx_t \quad (15)$$

where $A \in \mathbb{R}^{n \times n}$ is Hurwitz, $b, c^T \in \mathbb{R}^n$, and L_t is an α -stable process with $\alpha \in (0, 2]$. We need to choose suitable

$$G(s) := c(sI - A)^{-1}b \quad (16)$$

to capture the statistical property of the noise (see Fig. 5). Suppose that sufficiently many long sample paths of y_t are available so that we can approximately compute its power spectrum density. When $\alpha = 2$, one of the easiest ways of choosing $G(s)$ is to make $|G(j\omega)|$ close to the power spectrum density because the Wiener process has a “flat” frequency distribution. Such a gain diagram fitting provides a good approximation to capture the *size* and *smoothness* of the given noise signals.

The following result shows that a similar modeling approach is also possible for stable process cases.

Theorem 1: For (15) and any $\alpha \in (0, 2)$, $p \in (-1, \alpha)$, and $\omega > 0$,

$$\lim_{T \rightarrow \infty} \mathbb{E} \left[\left| \frac{1}{T^{1/\alpha}} \int_0^T e^{-j\omega t} y_t dt \right|^p \right] = \kappa_{(\alpha,p)} |G(j\omega)|^p \quad (17)$$

³Throughout this paper, Italic fonts (e.g., x_t , A) are used to represent symbols related to the nonlinear dynamics in Problem 1, while Roman fonts (e.g., x_t , A) are used for the general linear dynamics in (15).

holds with

$$\kappa_{(\alpha,p)} = \frac{(2\kappa_\alpha)^p \Gamma(1 + \frac{p}{2}) \Gamma(1 - \frac{p}{\alpha})}{\Gamma(1 - \frac{p}{2})} \quad (18)$$

$$\kappa_\alpha := \left(\frac{1}{\pi} \int_0^\pi |\cos t|^\alpha dt \right)^{1/\alpha}. \quad (19)$$

□

This equality also holds for $\alpha = p = 2$, where the left- and right-hand sides of (17) are equal (up to scaling) to the power spectrum density of y_t and $|G(j\omega)|^2$, respectively. This fact justifies the aforementioned gain diagram fitting. This theorem implies that, for any fixed α , we should approximate the given power spectrum density by $|G(j\omega)|$ in the sense of (17). Therefore, we can use the new parameter α to capture other statistical characteristics without sacrificing the capability of representing the given frequency-domain information about the smoothness; see also Remark 2 in Section IV-B for a reasonable way to determine α .

Remark 1: The Student's t distribution is widely used to model heavy-tailed *random variables*. However, to the best of our knowledge, its associated *stochastic process* does not have similar properties to the Wiener process [24]. For example, the stable distribution cannot be replaced by the Student's t distribution in any of Definition 2 and Theorems 1 and 2.

IV. STOCHASTIC LINEARIZATION

In this section, we derive some theoretical results related to stochastic linearization and introduce our proposed method to Problem 1. In what follows, we confine ourselves to the case with $\alpha \in (1, 2]$ because even the mean absolute value is not finite for $\alpha \leq 1$.

A. Wiener Process Case

In this section, we briefly review the stochastic linearization for $\alpha = 2$. The following are two important properties for performing stochastic linearization of feedback systems consisting of linear systems and saturation.

- The probability distribution of y_t in (15) converges to the normal distribution $\mathcal{S}\alpha\mathcal{S}(2, \|\text{ce}^{\mathbf{A}t}\mathbf{b}\|_2)$. This well-known result plays a crucial role in LQG theory. By using this, once the nonlinearity is approximated by linear systems in Problem 1, it is straightforward to know the stationary distribution of y_t and z_t based on the associated transfer function.
- Given real scalar constants $\sigma > 0$ and $d > 0$, the gain $k > 0$ that minimizes

$$\mathbb{E}_{Y \sim \mathcal{S}\alpha\mathcal{S}(2,\sigma)} [|\text{sat}_d(Y) - kY|^2] \quad (20)$$

is given by

$$k = \text{erf} \left(\frac{d}{2\sigma} \right) \quad (21)$$

where the error function is given by

$$\text{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

This fact shows that, assuming that the underlying random variable is Gaussian, the best approximation gain, in the sense of the error variance, is given analytically. This is a consequence of the fact that the probability density function has an explicit representation. An important feature of this characterization is that the optimal gain depends not only on the threshold d , but also σ .

Let us go back to the feedback systems in Problem 1. If the stationary distribution of v is close to $\mathcal{S}\alpha\mathcal{S}(2, \sigma_v)$, then it is reasonable to approximate the saturation by $k = \text{erf}(d/(2\sigma_v))$. Conversely, once the saturation is approximated by a linear gain k , the stationary distribution of v should be close to the normal distribution characterized in terms of the system from \dot{L}_t to v . By combining these two observations, we obtain an equation with respect to k , which can be a reasonable linearized gain for the saturation; see the next section for detail. Note that the resulting gain depends not only on d , but also on linear dynamics.

B. Stable Process Case

The next step is to provide a stable process counterpart for the properties in the previous section.

1) Stationary Distribution of the Ornstein-Uhlenbeck Process: We begin with the output of linear systems driven by stable processes.

Theorem 2: For (15) with $\alpha \in (1, 2]$,

$$y_t \xrightarrow{d} \mathcal{S}\alpha\mathcal{S}(\alpha, \|\text{ce}^{\mathbf{A}t}\mathbf{b}\|_\alpha) \quad (22)$$

holds, where the L^α -norm of a real-valued function $f(t)$ defined on $t \geq 0$ is defined by

$$\|f\|_\alpha := \left(\int_0^\infty |f(t)|^\alpha dt \right)^{1/\alpha} \quad (23)$$

provided that it is finite. □

This theorem indicates that the output of a linear system driven by stable processes has a stable stationary distribution with the same α . The *gain* is the L^α -norm of the impulse response. Note that this value is identical to the H^2 -norm of the transfer function when $\alpha = 2$. In the proposed method, Theorem 2 for general linear dynamics is used to find a suitable linear approximation of the nonlinear dynamics in Problem 1.

Remark 2: From a modeling viewpoint, it seems easy and reasonable to determine α based on (11) and the line fitting of the log–log plot (as in Fig. 3) of the histogram of the long history of the noise data (an alternative of the stationary distribution).

2) Optimal Linear Approximation: Theorem 2 motivates us to investigate the optimal approximation gain for the saturation function, where the underlying random variable conforms to a stable distribution. However, by Proposition 1, the error variance as in (20) cannot be finite when $\alpha < 2$. In view of this, we approximately linearize the saturation based on the mean absolute error instead of the error variance.

Theorem 3: Given $d > 0$, $\sigma > 0$, and $\alpha \in (1, 2)$, the gain $k > 0$ that minimizes

$$\mathbb{E}_{Y \sim \mathcal{S}\alpha\mathcal{S}(\alpha,\sigma)} [|\text{sat}_d(Y) - kY|] \quad (24)$$

TABLE I
 γ_α VERSUS α

α	γ_α
1.1	864.6556
1.3	7.0057
1.5	2.7854
1.9	1.7504
2.0	1.6651

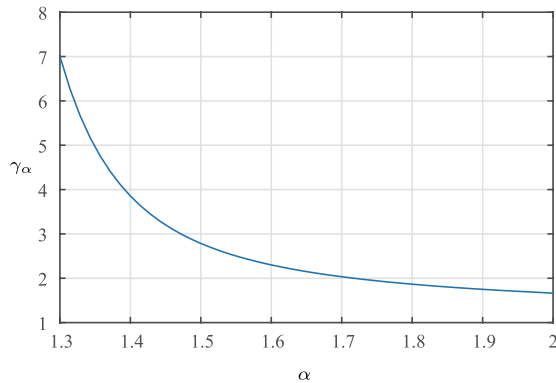


Fig. 6. γ_α for $\alpha \in [1.3, 2]$.

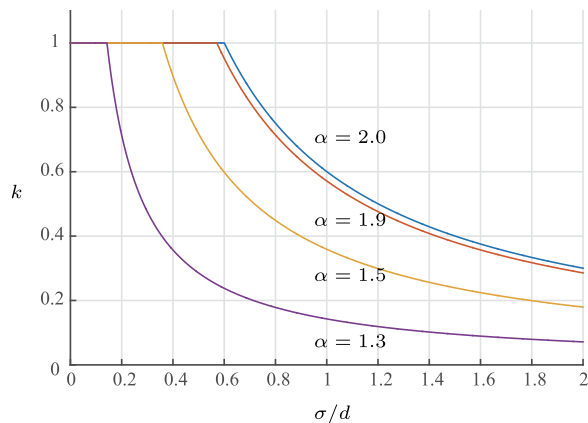


Fig. 7. Optimal gain k versus $\sigma/d \in [0, 2]$ for $\alpha = 1.3, 1.5, 1.9, 2.0$.

is given by

$$k_{\text{sat}} := \min \left(1, \frac{d}{\gamma_\alpha \sigma} \right) \quad (25)$$

where γ_α is a unique positive constant satisfying

$$\int_0^{\frac{\pi}{2}} \mathcal{F}_\alpha(u) \Gamma_\ell \left(2 - \frac{1}{\alpha}, \left(\frac{\gamma}{\mathcal{F}_\alpha(u)} \right)^{\frac{\alpha}{\alpha-1}} \right) du = \frac{1}{2} \frac{\alpha}{\alpha-1} \quad (26)$$

and $\mathcal{F}_\alpha(u)$ is defined by

$$\mathcal{F}_\alpha(u) := \frac{\sin(\alpha u)}{(\cos u)^{\frac{1}{\alpha}}} (\cos((1-\alpha)u))^{\frac{1-\alpha}{\alpha}}. \quad (27)$$

□

The value of γ_α for some α is shown in Table I and Fig. 6. The optimal gain in (25) is a monotonically decreasing function of σ/d for any fixed α , as is expected from (2); see Fig. 7,

which shows σ/d and the corresponding optimal gain k for $\alpha = 1.3, 1.5, 1.9, 2.0$. It should be noted that it is optimal to ignore the saturation (i.e., $k = 1$) if $\sigma/d < \gamma_\alpha$ despite the heavy-tailed distribution of the underlying random variable. When $\sigma/d > \gamma_\alpha$, the optimal gain is inversely proportional to σ/d .

Remark 3: As stated in the proof (see the Appendixes), (26) is equivalent to

$$\mathbb{E}_{Y \sim \mathcal{S}\alpha\mathcal{S}(\alpha)} [Y \cdot \mathbb{1}_{[0, \gamma_\alpha]}(Y)] = \frac{1}{2\pi} \Gamma \left(1 - \frac{1}{\alpha} \right). \quad (28)$$

In [14], the left-hand side of (28) was computed via the Monte Carlo simulation, by which 100% confidence intervals cannot be obtained. This fact can lead to invalid estimates particularly when α is close to 1. In contrast, the integral in (26) is much easier to evaluate. Note that the normalized incomplete Gamma function Γ_ℓ is implemented as `gammainc` in MATLAB [25].

C. Proposed Method

First, let us examine the boundedness of the moment of y in the original dynamics:

Corollary 1: In Problem 1, the stationary distribution of y_j has a finite p th order moment for any $p \in [1, \alpha)$ and $j = 1, \dots, r$. □

In view of this, similarly to the Gaussian case, the stationary distribution of y_j is approximated by

$$\mathcal{S}\alpha\mathcal{S}(\alpha, \sigma_{y_j}). \quad (29)$$

Then, it is reasonable to approximate $\text{sat}_{d_j}(y_j)$ by linear gain

$$k_j = \min \left(1, \frac{d_j}{\gamma_\alpha \sigma_{y_j}} \right) \quad (30)$$

by Theorem 3. Conversely, once each saturation $\text{sat}_{d_j}(y_j)$ is approximated by a linear gain k_j , the stationary distribution of y_j is given in the form of (29) with

$$\sigma_{y_j} = \left\| C_{y_j} e^{(A+BKC_y)t} b \right\|_\alpha \quad (31)$$

by Theorem 2, where C_{y_j} is the j th row of C_y and $K := \text{diag}(k_1, \dots, k_r)$. Here, we utilized the fact that the transfer function of the linearized system from \dot{L}_t to the y_j is $C_{y_j}(sI - (A+BKC_y)^{-1}b)$. Combining (29)–(31), we are ready to state our proposed method for Problem 1:

Proposed method

Solve the coupled equations

$$k_j = \min \left(1, \frac{d_j}{\gamma_\alpha \|C_{y_j} e^{(A+BKC_y)t} b\|_\alpha} \right), \quad j = 1, \dots, r \quad (32)$$

with respect to $K := \text{diag}(k_1, \dots, k_r)$ in order to determine the approximate gain $k_j \in [0, 1]$ for the j -th saturation. Then, regard

$$\mathcal{S}\alpha\mathcal{S}(\alpha, \|C_z e^{(A+BKC_y)t} b\|_\alpha)$$

as an approximate stationary distribution of z_t .

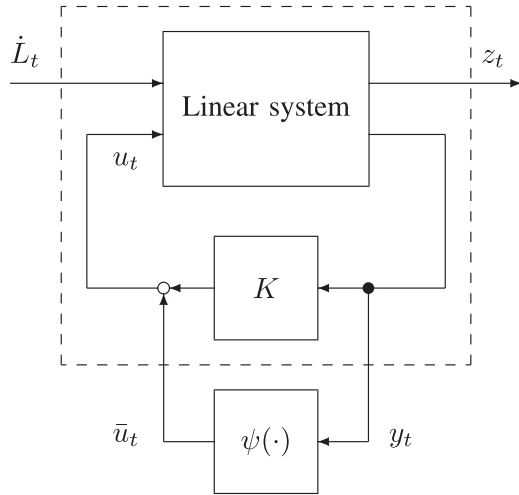


Fig. 8. Block diagram for the loop shifting.

The formula in Proposition 2 can be used to compute the moments with respect to this distribution. We solve the r coupled equations (32) numerically. This implies that it is important that the equation is given in a simple form so that no sample path generation is needed at any iteration.

D. Error Analysis

To the best of the authors' knowledge, existing Gaussian approximation error analysis methods (e.g., [26], [27]) are not applicable to our case. Actually, even in the conventional Gaussian case, theoretical error bounds for stochastic linearization of feedback systems have not been obtained [11]. In this section, we provide a novel theoretical bound and numerical experiments to examine the practical usefulness.

1) Theoretical Bound: First, the loop shifting in Fig. 8 is applied to obtain

$$dx_t = (A + BKC_y)x_t dt + B\bar{u}_t dt + bdL_t \quad (33)$$

$$\bar{u}_t = \psi(y_t) \quad (34)$$

with (5) and (6), where

$$\psi(y) := \text{sat}_d(y) - Ky. \quad (35)$$

Furthermore, thanks to the linearity of the dynamics, the solutions to

$$d\tilde{x}_t = (A + BKC_y)\tilde{x}_t dt + bdL_t \quad (36)$$

$$d\delta_t = (A + BKC_y)\delta_t dt + B\bar{u}_t dt \quad (37)$$

with $x_0 = \tilde{x}_0 + \delta_0$ satisfy

$$x_t = \tilde{x}_t + \delta_t. \quad (38)$$

See the block diagram in Fig. 9. The proposed method approximates x_∞ by \tilde{x}_∞ , which satisfies $c_z \tilde{x}_\infty \sim \mathbf{S}\alpha\mathbf{S}(\alpha, \|c_z e^{(A+BKC_y)t} b\|_\alpha)$. It is, therefore, crucial for error analysis to evaluate δ_t . The main difficulty lies in the fact that δ_t is a signal generated in the nonlinear feedback loop. Recall

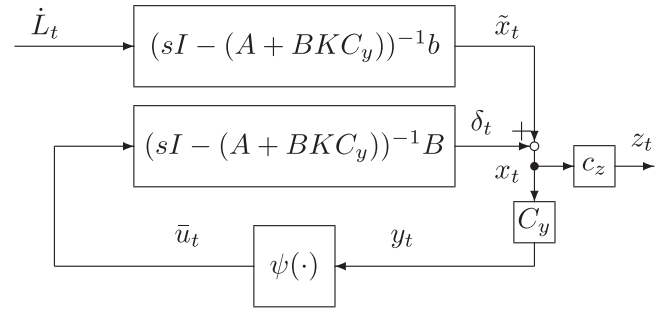


Fig. 9. Block diagram for Theorem 4.

that our proposed method determines K aiming at minimizing

$$\eta(K) := \mathbb{E}_{Y \sim \mathbf{S}\alpha\mathbf{S}(\alpha, \|C_y e^{(A+BKC_y)t} b\|_\alpha)} [|\text{sat}_d(Y) - KY|]. \quad (39)$$

The following theorem gives an error bound in terms of $\eta(K)$.

Theorem 4: In Problem 1 with $r = 1$, given $K \in [0, 1]$, define

$$\zeta := \max(K, 1 - K) \in [1/2, 1]$$

$$\eta_y := \|C_y e^{(A+BKC_y)t} B\|_1$$

$$\eta_z := \|c_z e^{(A+BKC_y)t} B\|_1.$$

Then, if $\zeta\eta_y < 1$, then

$$\mathcal{E} := \limsup_{t \rightarrow \infty} \mathbb{E}[|c_z(x_t - \tilde{x}_t)|] \leq \frac{\eta_z}{1 - \zeta\eta_y} \eta(K). \quad (40)$$

□

When the mean absolute value of z_t is of interest in Problem 1, one can use

$$\mathbb{E}[|c_z \tilde{x}_\infty|] - \mathcal{E} \leq \mathbb{E}[|c_z x_\infty|] \leq \mathbb{E}[|c_z \tilde{x}_\infty|] + \mathcal{E} \quad (41)$$

which follows from (40) and the triangle inequality.

The obtained *a posteriori* error bound can be understood intuitively as follows: Concerning the static nonlinearity, the only property we utilized in the proof is

$$|\psi(y + \Delta)| \leq |\psi(y)| + \zeta|\Delta|, \quad y, \Delta \in \mathbb{R}. \quad (42)$$

Note that the H^∞ -norm of $C_y(sI - (A + BKC_y))^{-1}B$ is less than or equal to η_y (see, e.g., [2]). Therefore, ζ and η_y can be viewed as (incremental) gain upper bounds of the static nonlinearity $\psi(\cdot)$ and the linear system from \bar{u}_t to y_t , respectively. Consequently, the assumption $\zeta\eta_y < 1$ plays a role of a *small gain* condition to ensure the stability of the nonlinear feedback loop in Fig. 9. Under this condition, $1/(1 - \zeta\eta_y)$ can be seen as the *sensitivity function* that characterizes how much the additive signal to \bar{u}_t , which is related to $\eta(K)$, affects \bar{u}_t . Similarly, η_z is the gain of the linear system from \bar{u}_t to z_t . The upper bound in (40) is the product of these three terms.

Remark 4: Although the case of $r = 1$ and the saturation is considered here for simplicity, similar bounds can be obtained for general cases with other nonlinearities shown in Fig. 3. It is also expected that there are several ways to tighten the bounds.

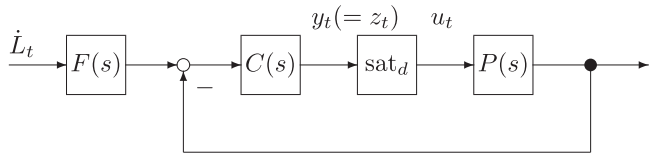


Fig. 10. Unity feedback system.

TABLE II
MONTE CARLO METHODS (MC) AND THE PROPOSED METHOD (SL)

σ	p	d	MC	SL	\mathcal{E}
1	0.1	1	0.6123	0.5840	-
1	0.1	3	0.5864	0.5840	-
5	0.1	1	3.3268	3.4689	-
5	0.1	3	3.1699	3.1136	-
1	2	1	0.5904	0.5765	0.0599
1	2	3	0.6259	0.5765	0.0089
5	2	1	3.4630	3.5825	0.6630
5	2	3	3.1238	3.1381	0.7500

For example, it seems better to replace (42) by

$$|\psi(y + \Delta)| \leq |\psi(y)| + K|\Delta| + d(1 - 2K) \quad (43)$$

when K is close to 0.

2) Numerical Experiments by Randomly Generated Systems: Similarly to [11, Sec. 3], we apply the proposed method to the unity feedback system in Fig. 10, where $F(s)$, $C(s)$, $P(s)$, $d > 0$, and $1 < \alpha \leq 2$ (the parameter in the stable process L_t) are randomly generated and examine the resulting accuracy.

For the proposed method, the approximate gain for the saturation, i.e., the solution to (32), was computed within the accuracy 10^{-3} . For the Monte Carlo method, we calculated long time average of $|y_t|$ (i.e., $z_t = y_t$) of sample paths, which are generated by time discretization with width $dt = 0.01$ until these average values *seem to converge*. For each value, the computation time is more than 10 min for the Monte Carlo method and less than 1 s for the proposed method with MacBook Air with an Intel Core i7 processor. Table II shows the obtained data of the mean absolute value of the stationary distribution of u_t for some cases with

$$F(s) = \frac{\sigma}{s+1}, \quad C(s) = \frac{1}{s+1}, \quad P(s) = \frac{1}{s+p}, \quad \alpha = 1.7.$$

We can observe that the proposed method provides reasonable estimates. It should be emphasized that the computation time required for the Monte Carlo method quickly increases as $\alpha \searrow 1$, while it does not affect that for the proposed method significantly.

Next, Theorem 4 is applied to obtain a hard bound. In this feedback system, $\eta_y = \eta_z$ is the L^1 -norm of the impulse response of the system

$$\frac{C(s)P(s)}{1 + KC(s)P(s)}.$$

Unfortunately, Theorem 4 is not applicable to the case $p < 0.5$ because $\zeta\eta_y < 1$ does not hold. For other cases, \mathcal{E} is shown in the rightmost column of Table II. These bounds for $\sigma = 1$ provide

useful information about the accuracy, while it is desirable to tighten them for $\sigma = 5$. It is worth noting that the result for $(\sigma, p, d) = (1, 2, 3)$ again indicates that the Monte Carlo method is not trivial.

E. Other Nonlinearities and Impact of Non-Gaussianity α

We have so far investigated the saturation only. The proposed scheme is also applicable to other piecewise affine nonlinearities.

Theorem 5: Let d , $\sigma > 0$, $\alpha \in (1, 2)$ and \mathbb{E} denotes $\mathbb{E}_{Y \sim \mathcal{S}\alpha\mathcal{S}(\alpha, \sigma)}$.

1) For the *relay* in Fig. 3(b)

$$\text{rel}_d(y) := \begin{cases} -d, & y < 0 \\ 0, & y = 0 \\ d, & y > 0 \end{cases} \quad (44)$$

$\mathbb{E}[|\text{rel}_d(Y) - kY|]$ is minimized by

$$k_{\text{rel}} := \frac{d}{\gamma_\alpha \sigma}. \quad (45)$$

2) For the *deadzone* in Fig. 3(c)

$$\text{dz}_d(y) := \begin{cases} d + y, & y < -d \\ 0, & |y| \leq d \\ -d + y, & y > d \end{cases} \quad (46)$$

$\mathbb{E}[|\text{dz}_d(Y) - kY|]$ is minimized by

$$k_{\text{dz}} := 1 - k_{\text{sat}}. \quad (47)$$

3) For the *friction* in Fig. 3(d)

$$\text{fr}_d(y) := \begin{cases} -d + y, & y < 0 \\ 0, & y = 0 \\ d + y, & y > 0 \end{cases} \quad (48)$$

$\mathbb{E}[|\text{fr}_d(Y) - kY|]$ is minimized by

$$k_{\text{fr}} := 1 + k_{\text{rel}}. \quad (49)$$

□

The relationships $k_{\text{sat}} = 1 - k_{\text{dz}}$ and $k_{\text{fr}} = 1 + k_{\text{rel}}$ are the same as in the conventional Gaussian case (see [11]).

Remark 5: For the *saturation with deadzone* $\text{sat}_d(\text{dz}_\delta(y))$ with $d, \delta > 0$, the gain k that minimizes $\mathbb{E}[|\text{sat}_d(\text{dz}_\delta(Y)) - kY|]$ is given by the unique $k \in [0, d/(d + \delta)]$ satisfying

$$2\mathbb{E}[Y \cdot \mathbb{1}_{(\delta/(1-k), d/k)}(Y)] = \mathbb{E}[Y \cdot \mathbb{1}_{(0, \infty)}(Y)]. \quad (50)$$

The result for the saturation and deadzone can be viewed as a special case of (50) with $\delta = 0$ and $d = +\infty$, respectively; see (70) in the proof of Theorem 3.

Next, we make some comments on the relationship between parameters. All the optimal gains in Theorems 3 and 5 are determined by the ratio between the saturation threshold d and noise scale σ and γ_α . The contribution of the non-Gaussianity parameter α is characterized as follows.

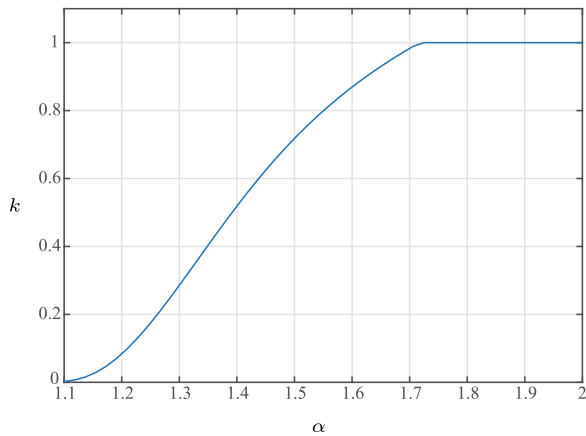


Fig. 11. Optimal gain k for $\alpha \in [1.1, 2)$ when $\sigma/d = 2$.

Theorem 6: For $\alpha \in (1, 2)$, γ_α defined by (28) is monotonically decreasing with respect to α , and

$$\lim_{\alpha \searrow 1} \gamma_\alpha = +\infty. \quad (51)$$

□

Consequently, when $\sigma/d > 0$ is fixed, the optimal gain k_{sat} in (26) is monotonically increasing with respect to α , and

$$\lim_{\alpha \searrow 1} k_{\text{sat}} = 0. \quad (52)$$

This fact can be explained as follows. First, a smaller gain achieves a smaller error for the extremal outliers (i.e., $|d - kx|$ for $|x| \gg d$) that frequently occur for stable processes with α close to 1. While a smaller gain leads to a larger error on the unsaturated interval (i.e., $|x - kx|$ for $|x| < d$), this is less important because the unsaturated input is close to 0 with a high probability. See Fig. 11 for α -dependence when $\sigma/d = 2$.

V. APPLICATION: UNCERTAINTY ASSESSMENT OF WIND POWER GENERATION

A. Model Description

In this section, we apply the proposed scheme to evaluate the frequency fluctuation of power systems interconnected to uncertain renewable energy [28], [29]. We utilize the load frequency control (LFC) model in Fig. 12; see [29, Sec. 1] for a strong motivation to compute interconnectable capacity of wind power in Japan. The physical meaning of each signal and block (the role of the path with q_t and P_3 is explained later) is

- t [s] time;
- n_t [MW] wind power fluctuation;
- e_t [MW] power deviation;
- f_t [Hz] frequency deviation;
- v_t [MW] power adjustment by LFC;
- W frequency domain model of the wind power fluctuation;
- P_1 physical inertia, e.g., load characteristics and system inertia of the power system;
- P_2 power generator for LFC.

The leeway available by the controller to adjust the power output of the power generator is bounded by $d_1 > 0$. Thus, the saturation function sat_{d_1} is appropriate to model this. The value of d_1 significantly affects the running cost and the resulting frequency deviation. The feedback loop with $1/s$ and sat_{d_2} ($d_2 > 0$) represents the rate limiter for the command input to the power generator. In what follows, we analyze the frequency deviation for the case, where L_t is an α -stable process with $\alpha \in (1, 2]$. Actually, several real data in [3] had a tendency similar to Fig. 3, which did not fit well to the normal distribution.

Note that the model with $q_t = 0$ mainly focuses on the high- or middle-frequency behavior controlled by inertia of the power system and the adjustment of the thermal generation v_t . In order to consider the low-frequency fluctuation (e.g., longer than 20 min) caused by the frequent outliers, we also include the slow adjustment q_t by means of *combined cycle* or *economic dispatch control* (EDC), whose responsiveness is characterized by P_3 . This signal plays a role of adjusting the saturation range of control input (see Fig. 13).

We take $d_1 = 25$ MW, $d_2 = 0.166$ MW/s, $k_{\text{sys}} = 1/250$ Hz/MW,

$$P_1 = \frac{1}{3s + 1}, \quad P_2 = \frac{1}{\frac{1}{0.15}s + 1}, \quad W(s) = \frac{0.0021 \times 650}{s + 0.001} \quad (53)$$

as in [29, Sec. 4], and

$$P_3(s) = \frac{1}{T_c s + 1}. \quad (54)$$

Note that (14) in Example 1 corresponds to n_t in Fig. 12 with $W(s)$ in (53). In summary, the dynamics in Fig. 12 with $y_t = [y_1 \ y_2]^T$ and $z_t = e_t$ are given by (4)–(7) with

$$A = \begin{bmatrix} -1/3 & 0 & 0 & -1/3 & 1/3 \\ 0 & -0.15 & 0.15 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1/T_c & 0 & 0 & -1/T_c & 0 \\ 0 & 0 & 0 & 0 & -0.001 \end{bmatrix}$$

$$B = \begin{bmatrix} -1/3 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$C_y = \begin{bmatrix} C_{y1} \\ C_{y2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 \end{bmatrix}$$

$$c_z = [1 \ 0 \ 0 \ 0 \ 0]$$

$$b = [0 \ 0 \ 0 \ 0 \ -0.0021 \times 650]^T.$$

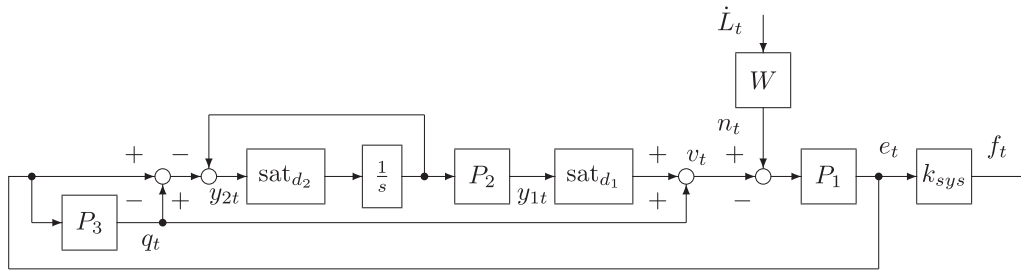


Fig. 12. Block diagram of the LFC system.

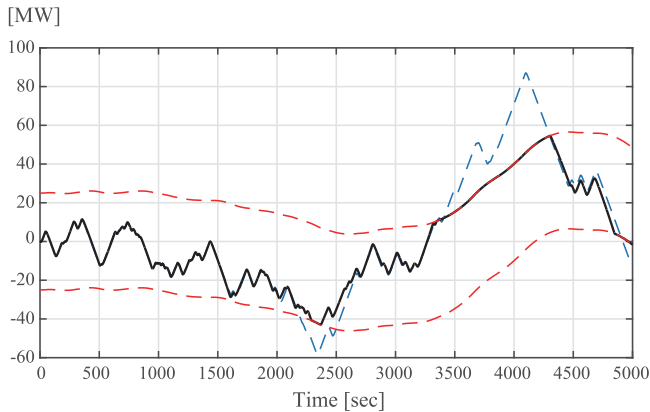


Fig. 13. Sample path of v_t (black), $y_{1t} + q_t$ (blue), and $q_t \pm d_1$ (red).

Finally, we set

$$\alpha = 1.9 \quad (55)$$

unless otherwise stated.

To confirm the effect of q_t , a sample path of v_t is shown in Fig. 13, where $T_c = 10^3$. While $v_t \in [-d_1, d_1]$ if $q_t = 0$, we can generate a wider range of v_t by slowly adjusting the saturation range.

B. Simulation Result

In order to assess the resulting average frequency fluctuation, we approximately compute $\mathbb{E}[e_\infty]$. For the Monte Carlo method, we generate a sample path by time discretization with width $dt = 0.01$ and calculated long time average $\int_0^{T_{MC}} |e_t| dt / T_{MC}$ with large T_{MC} . First, let us consider the case without feedback, i.e., $v_t = 0$. In this case, e_t is an Ornstein–Uhlenbeck process as in Fig. 5 with $G = P_1 W$. Thus, its exact stationary distribution can be obtained by Theorem 2; $\mathbb{E}[e_\infty] \approx 43.8695$. On the other hand, the result of the Monte Carlo method is shown in Fig. 14. We can see that even this simple Ornstein–Uhlenbeck process case is not trivial to evaluate by sample path generation, due to its slow convergence. This result implies that the suitable sample number is difficult to determine. Actually, even for such linear processes, their effective Monte Carlo simulation requires sophisticated techniques when $n > 1$ [30].

Next, we include feedback control. Fig. 15 shows the results for the system with/without adjustment q_t of the satura-

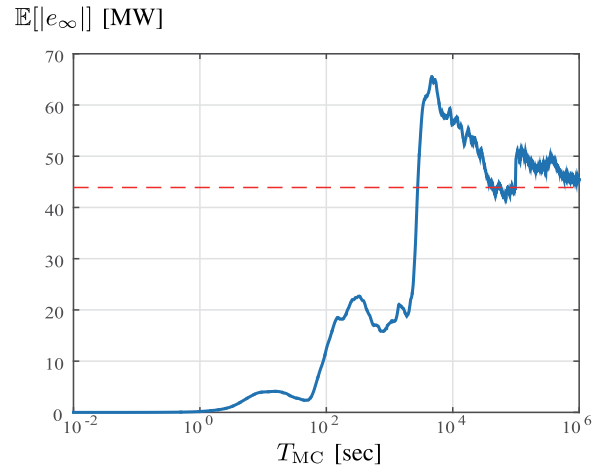


Fig. 14. Slow convergence of the Monte Carlo method for the stable process case with $\alpha = 1.9$: Monte Carlo result versus averaging time length (blue) and theoretical value (red).

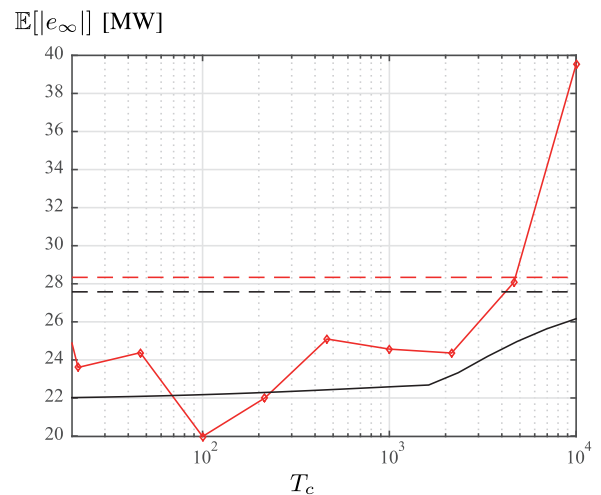


Fig. 15. Proposed method (black) and the Monte Carlo method (red): $q_t = 0$ case (dashed) and $T_c \in [10^1, 10^4]$ (solid).

tion range of v_t , where P_3 is (54) with $T_c \in [10^1, 10^4]$. The Monte Carlo method, for which we took $T_{MC} = 10^6$ in view of Fig. 14, was performed only for $q_t = 0$ and 10 different values for T_c . Recall that a smaller T_c implies better responsiveness of the complementary power generation or EDC. The result by the proposed method suggests that $T_c = 10^3$ is enough because its fluctuation suppression effect saturates. On the other hand,

the Monte Carlo method does not provide such an insight into the parameter dependence. In most trials, we encountered an unsuitable result, as in that for $T_c = 10^4$ in Fig. 15, which shows the difficulty of the Monte Carlo simulation.

VI. CONCLUSION

In this paper, we proposed a framework to deal with dynamical systems under heavy-tailed stochastic noise. We utilized stable processes to formulate non-Gaussianity caused by nonnegligible extremal events. From an identification point of view, this model enables us to easily incorporate noise statistics available from time-series data. In Section IV, we gave two key results for stochastic linearization of feedback systems driven by stable processes. The representation in Theorems 2 and 3 is simple enough to calculate iteratively to solve (32). It should be emphasized that the Monte Carlo simulation for systems with a lot of rare events is highly nontrivial or requires an infeasibly high number of samples. Therefore, the advantage of our proposed method over the Monte Carlo method is more significant as the non-Gaussianity increases (smaller α). The effectiveness of the proposed method was examined through the uncertainty assessment of wind power generation. Finally, the most important property that is lost for $\alpha < 2$ is that *the associated signal and system spaces are Hilbert (i.e., the Lebesgue space L^2 and Hardy space H^2)*, based on which LQG theory and Kalman filtering have been developed for $\alpha = 2$. We expect that our approximation scheme is useful to circumvent this issue when practical controller design and state estimation methods are investigated.

APPENDIX A

PROOF OF THEOREMS 1 AND 2 AND COROLLARY 1

A. Theorem 2

First, we prove Theorem 2 based on the following proposition (see [8], Proposition 3.4.1):

Proposition 4: For any real-valued measurable function f defined on $[0, T]$,

$$\mathbb{E} \left[\exp \left[j\nu \int_0^T f(s) dL_s \right] \right] = \exp \left[- \int_0^T |\nu f(s)|^\alpha ds \right] \quad (56)$$

for $\nu \in \mathbb{R}$ holds provided that $\int_0^T |f(s)|^\alpha ds$ is finite. \square

Equation (15) admits the explicit strong solution

$$x_t = e^{At} x_0 + \int_0^t e^{A(s-t)} b dL_s. \quad (57)$$

In what follows, we take $x_0 = 0$ as the first term vanishes in the large t limit. Then, Proposition 4 with $f(s) = ce^{A(t-s)} b$ gives

$$\mathbb{E}[e^{j\nu y_t}] = \exp \left[-|\nu|^\alpha \int_0^t |ce^{As} b|^\alpha ds \right], \quad \nu \in \mathbb{R}. \quad (58)$$

Therefore, by taking the large t limit, the desired result follows.

B. Theorem 1

Combining (15) with the Itô product formula

$$d(e^{-j\omega t} x_t) = -j\omega e^{-j\omega t} x_t dt + e^{-j\omega t} dx_t$$

we obtain

$$e^{-j\omega t} x_t dt = (A - j\omega I)^{-1} (d(e^{-j\omega t} x_t) - e^{-j\omega t} b dL_t).$$

By integrating this from 0 to T and multiplying by $c/T^{1/\alpha}$, we have

$$\begin{aligned} & \frac{1}{T^{1/\alpha}} \int_0^T e^{-j\omega t} y_t dt \\ &= \frac{1}{T^{1/\alpha}} \left(c(A - j\omega I)^{-1} e^{-j\omega T} x_T + G(j\omega) \int_0^T e^{-j\omega t} dL_t \right). \end{aligned}$$

It follows from Proposition 4 that the first term converges in probability to zero. Consequently, by focusing on the second term, we have

$$\left| \frac{1}{T^{1/\alpha}} \int_0^T e^{-j\omega t} y_t dt \right| = |G(j\omega)| \|v_T\| \quad (59)$$

where we introduced the two-dimensional random vector

$$\begin{aligned} v_T &:= \frac{1}{T^{1/\alpha}} \left[\operatorname{Re} \int_0^T e^{-j\omega t} dL_t, \operatorname{Im} \int_0^T e^{-j\omega t} dL_t \right]^\top \\ &= \frac{1}{T^{1/\alpha}} \int_0^T [\cos \omega t, -\sin \omega t]^\top dL_t. \end{aligned}$$

Next, take an arbitrary $\nu \in \mathbb{R}^2$ and denote its argument by $\psi_\nu \in [0, 2\pi)$, i.e., $\nu = \|\nu\| [\cos \psi_\nu, \sin \psi_\nu]^\top$. Then, by the multidimensional version of Proposition 4 (see [8, Ch. 3]), we have

$$\begin{aligned} \mathbb{E}[\exp[j\nu^\top v_T]] &= \exp \left[-\frac{1}{T} \int_0^T |\nu^\top [\cos \omega t, -\sin \omega t]^\top|^\alpha dt \right] \\ &= \exp \left[-\|\nu\|^\alpha \frac{1}{T} \int_0^T |\cos(\omega t + \psi_\nu)|^\alpha dt \right] \\ &\rightarrow \exp[-\|\nu\|^\alpha (\kappa_\alpha)^\alpha], \text{ as } T \rightarrow \infty \end{aligned}$$

independent of ω, ψ_ν . Hence, $v_T \xrightarrow{d} v_\infty$ as $T \rightarrow \infty$, where v_∞ is an *isotropic* stable distribution on \mathbb{R}^2 defined by

$$\mathbb{E}[\exp[j\nu^\top v_\infty]] = \exp[-\|\nu\|^\alpha (\kappa_\alpha)^\alpha].$$

Finally, Theorem 1 follows from (59) and a closed expression for $\mathbb{E}[\|v_\infty\|^p]$ (see [31]).

C. Corollary 1

Corollary 1 is a direct consequence of Theorem 2. The solution to (4) is given by

$$x_t = e^{At} x_0 + \int_0^t e^{A(t-s)} b dL_s + \int_0^t e^{A(t-s)} B u_s ds. \quad (60)$$

Here, the first and second terms converge in law to a symmetric stable distribution with α by Theorem 2. Recall that $u_t \in [-d_1, d_1] \times \cdots \times [-d_r, d_r]$ by definition. Therefore, the support of the third term is uniformly bounded for all t . This completes the proof.

APPENDIX B PROOF OF THEOREM 4

In what follows, the following explicit representation is employed:

$$\delta_t = e^{(A+BKC_y)(t-t')} \delta_{t'} + \int_{t'}^t e^{(A+BKC_y)(t+t'-s)} B\psi(y_s) ds \quad (61)$$

for any $t > t' > 0$. We ignore the first term, since it vanishes as $t \rightarrow \infty$ independent of the choice of t' . This implies that

$$\mathbb{E}[|C_y \delta_t|] \leq \eta_y \cdot \sup_{s \geq t'} \mathbb{E}[|\psi(y_s)|], \quad t > t'. \quad (62)$$

Furthermore, due to the assumption $x_t \xrightarrow{d} x_\infty$,

$$\lim_{t' \rightarrow \infty} \sup_{s \geq t'} \mathbb{E}[|\psi(y_s)|] = \lim_{t \rightarrow \infty} \mathbb{E}[|\psi(y_t)|]. \quad (63)$$

In summary,

$$\mathcal{E}_y := \limsup_{t \rightarrow \infty} \mathbb{E}[|C_y \delta_t|] \leq \eta_y \cdot \lim_{t \rightarrow \infty} \mathbb{E}[|\psi(y_t)|]. \quad (64)$$

By the same argument,

$$\mathcal{E} \leq \eta_z \cdot \lim_{t \rightarrow \infty} \mathbb{E}[|\psi(y_t)|] \quad (65)$$

also holds.

Next, inequality (42) yields

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{E}[|\psi(y_t)|] &\leq \limsup_{t \rightarrow \infty} \mathbb{E}[|\psi(C_y \tilde{x}_t)| + \zeta |C_y \delta_t|] \\ &\leq \lim_{t \rightarrow \infty} \mathbb{E}[|\psi(C_y \tilde{x}_t)|] + \zeta \limsup_{t \rightarrow \infty} \mathbb{E}[|C_y \delta_t|] \\ &= \eta(K) + \zeta \mathcal{E}_y \\ &\leq \eta(K) + \zeta \eta_y \cdot \lim_{t \rightarrow \infty} \mathbb{E}[|\psi(y_t)|] \end{aligned} \quad (66)$$

where $C_y \tilde{x}_t \xrightarrow{d} \mathcal{S}\alpha\mathcal{S}(\alpha, \|C_y e^{(A+BKC_y)t} b\|_\alpha)$ and (64) are applied. Therefore, under the assumption $\zeta \eta_y < 1$, we have

$$\lim_{t \rightarrow \infty} \mathbb{E}[|\psi(y_t)|] \leq \frac{1}{1 - \zeta \eta_y} \eta(K). \quad (67)$$

Finally, this inequality with (65) completes the proof.

APPENDIX C PROOF OF THEOREMS 3, 5, AND 6

In the remaining of this paper, \mathbb{E} represents $\mathbb{E}_{Y \sim \mathcal{S}\alpha\mathcal{S}(\alpha, \sigma)}$ unless otherwise stated.

A. Theorem 3

Let us regard (24) as a function of k . Note that this function is continuous on \mathbb{R} because

$$\begin{aligned} &|\mathbb{E}[|\text{sat}_d(Y) - kY|] - \mathbb{E}[|\text{sat}_d(Y) - lY|]| \\ &= |\mathbb{E}[|\text{sat}_d(Y) - kY| - |\text{sat}_d(Y) - lY|]| \\ &\leq |k - l| \mathbb{E}[|Y|] \end{aligned}$$

where $\mathbb{E}[|Y|]$ is finite.

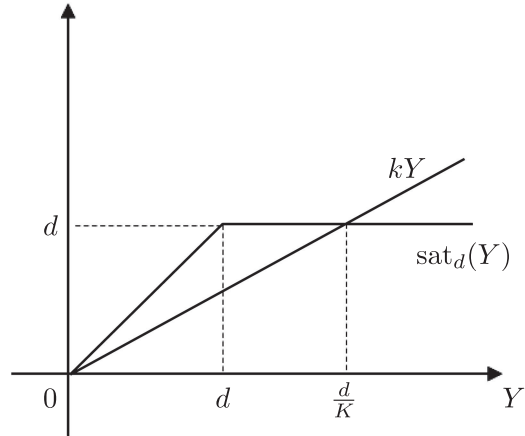


Fig. 16. Saturation function and its linear approximation.

Trivially, the optimal k that minimizes this expectation satisfies $0 \leq k \leq 1$. For $0 < k \leq 1$, splitting the integration interval as in Fig. 16, we obtain

$$\begin{aligned} &\mathbb{E}[|\text{sat}_d(Y) - kY|] \\ &= 2\mathbb{E}[(Y - kY) \cdot \mathbb{1}_{(0,d)}(Y)] + 2\mathbb{E}[(d - kY) \cdot \mathbb{1}_{(d,d/k)}(Y)] \\ &\quad + 2\mathbb{E}[(kY - d) \cdot \mathbb{1}_{(d/k,+\infty)}(Y)] \\ &= 2k\mathbb{E}[Y \cdot \mathbb{1}_{(0,+\infty)}(Y)] - 4k\mathbb{E}[Y \cdot \mathbb{1}_{(0,d/k)}(Y)] \\ &\quad + 4d\mathbb{E}[\mathbb{1}_{(0,d/k)}(Y)] + \text{constant} \end{aligned}$$

where the last term is a constant independent of k . Thus, by the linearity of the expectation, the derivative is

$$\begin{aligned} \frac{d}{dk} \mathbb{E}[|\text{sat}_d(Y) - kY|] &= 2\mathbb{E}[Y \cdot \mathbb{1}_{(0,+\infty)}(Y)] \\ &\quad - 4\mathbb{E}[Y \cdot \mathbb{1}_{(0,d/k)}(Y)] \\ &\quad - 4k \frac{d}{dk} \{ \mathbb{E}[Y \cdot \mathbb{1}_{(0,d/k)}(Y)] \} \\ &\quad + 4d \frac{d}{dk} \{ \mathbb{E}[\mathbb{1}_{(0,d/k)}(Y)] \} \\ &= 2\mathbb{E}[Y \cdot \mathbb{1}_{(0,+\infty)}(Y)] \\ &\quad - 4\mathbb{E}[Y \cdot \mathbb{1}_{(0,d/k)}(Y)] \end{aligned} \quad (68)$$

where we used the equality

$$k \frac{d}{dk} \{ \mathbb{E}[Y \cdot \mathbb{1}_{(0,d/k)}(Y)] \} = d \frac{d}{dk} \{ \mathbb{E}[\mathbb{1}_{(0,d/k)}(Y)] \}. \quad (69)$$

This means that $\frac{d}{dk} \mathbb{E}[|\text{sat}_d(Y) - kY|]$ is a strictly increasing function with respect to k and is negative for sufficiently small k . Therefore, the desired minimizer is given as $\min(1, k^*)$, where k^* is the unique solution to

$$\mathbb{E}[Y \cdot \mathbb{1}_{(0,+\infty)}(Y)] = 2\mathbb{E}[Y \cdot \mathbb{1}_{(0,d/k)}(Y)]. \quad (70)$$

By (12), the left-hand side is equal to

$$\frac{\sigma}{\pi} \Gamma \left(1 - \frac{1}{\alpha} \right).$$

On the other hand, the right-hand side is equal to

$$2\sigma \mathbb{E}_{Y \sim \mathcal{S}_{\alpha} \mathcal{S}(\alpha)} [Y \cdot \mathbb{1}_{(0, d/(\sigma k))}(Y)].$$

In summary,

$$\gamma_{\alpha} := \frac{d}{\sigma k^*} \quad (71)$$

satisfies (28).

Based on (10), it can be verified that the left-hand side of (28) can be rewritten as

$$\begin{aligned} & \mathbb{E}_{Y \sim \mathcal{S}_{\alpha} \mathcal{S}(\alpha)} [Y \cdot \mathbb{1}_{(0, \gamma)}(Y)] \\ &= \frac{\Gamma(2 - \frac{1}{\alpha})}{\pi} \int_0^{\frac{\pi}{2}} \mathcal{F}_{\alpha}(u) \Gamma_{\ell} \left(2 - \frac{1}{\alpha}, \left(\frac{\gamma}{\mathcal{F}_{\alpha}(u)} \right)^{\frac{\alpha}{\alpha-1}} \right) du. \end{aligned} \quad (72)$$

Finally, $\Gamma(2 - \frac{1}{\alpha}) = (1 - \frac{1}{\alpha})\Gamma(1 - \frac{1}{\alpha})$ completes the proof of Theorem 3.

B. Theorem 5

Similarly to the case of the saturation, we have

$$\begin{aligned} \mathbb{E} [|\text{rel}_d(Y) - kY|] &= 2k \mathbb{E}[Y \cdot \mathbb{1}_{(0, +\infty)}(Y)] \\ &\quad - 4k \mathbb{E}[Y \cdot \mathbb{1}_{(0, d/k)}(Y)] \\ &\quad + 4d \mathbb{E} [\mathbb{1}_{(0, d/k)}(Y)] + \text{constant} \end{aligned}$$

where the last term is a constant independent of k . The only difference is that this equality holds for all $k \geq 0$. That is, the optimal gain is not bounded by 1. Therefore, we obtain k_{rel} instead of k_{sat} .

Finally, $\text{dz}_d(y) = y - \text{sat}_d(y)$ (respectively, $\text{fr}_d(y) = y + \text{rel}_d(y)$) readily yields $k_{\text{dz}} = 1 - k_{\text{sat}}$ (respectively, $k_{\text{fr}} = 1 + k_{\text{rel}}$).

C. Theorem 6

Note that the left-hand side of (28) is strictly increasing with respect to γ_{α} , and that the Gamma function $\Gamma(\cdot)$ is strictly decreasing function on the interval $(0, 1/2]$. Therefore, γ_{α} is strictly decreasing with respect to α . In addition, since $\Gamma(1 - 1/\alpha) \rightarrow +\infty$ as $\alpha \rightarrow 1$, we have (51).

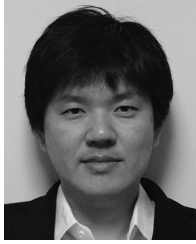
ACKNOWLEDGMENT

The authors would like to thank Dr. R. Kawai of the University of Sydney for his proof reading of Theorems 1 and 2. The authors would also like to thank Prof. M. Kato of Tokyo Denki University for his comment on the electricity network modeling in Section V.

REFERENCES

- [1] B. K. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, 5th ed. New York, NY, USA: Springer, 1998.
- [2] K. Zhou, J. C. Doyle, and K. K. Glover, *Robust and Optimal Control*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.
- [3] T. Yoshida, M. Kato, and K. Kashima, "Probabilistic evaluation method of interconnectable capacity for wind power generation using real data," in *Proc. Int. Conf. Renew. Energy Power Qual. J.*, vol. 1, no. 13, 2015, pp. 66–70.

- [4] A. Barabási, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.
- [5] O. L. Costa, M. D. Fragoso, and M. G. Todorov, *Continuous-Time Markov Jump Linear Systems*, 1st ed. Berlin, Germany: Springer-Verlag, 2013.
- [6] H. Ohlsson, F. Gustafsson, L. Ljung, and S. Boyd, "Smoothed state estimates under abrupt changes using sum-of-norms regularization," *Automatica*, vol. 48, no. 4, pp. 595–605, Apr. 2012.
- [7] L. R. G. Carrillo, W. J. Russell, J. P. Hespanha, and G. E. Collins, "State estimation of multiagent systems under impulsive noise and disturbances," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 1, pp. 13–26, Jan. 2015.
- [8] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*. London, U.K.: Chapman & Hall, 1994.
- [9] K.-i. Sato, *Lévy Processes and Infinitely Divisible Distributions*. Cambridge, U.K.: Cambridge Univ. Press, 1999.
- [10] D. Applebaum, *Lévy Processes and Stochastic Calculus*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2009.
- [11] S. Ching, Y. Eun, C. Gokcek, P. T. Kabamba, and S. M. Meerkov, *Quasilinear Control: Performance Analysis and Design of Feedback Systems with Nonlinear Sensors and Actuators*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [12] J. Westman and F. Hanson, "The LQGP problem: A manufacturing application," in *Proc. Amer. Control Conf.*, 1997, vol. 1, pp. 566–570.
- [13] D. Liberzon and R. Brockett, "Nonlinear feedback systems perturbed by noise: Steady-state probability distributions and optimal control," *IEEE Trans. Autom. Control*, vol. 45, no. 6, pp. 1116–1130, Jun. 2000.
- [14] K. Kashima, H. Aoyama, and Y. Ohta, "Modeling and linearization of systems under heavy-tailed stochastic noise with application to renewable energy assessment," in *Proc. 54th IEEE Conf. Decis. Control*, Dec. 2015, pp. 1852–1857.
- [15] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus* (ser. Graduate Texts in Mathematics), 2nd ed. New York, NY, USA: Springer-Verlag, 1998.
- [16] P. Billingsley, *Probability and Measure* (Wiley Series in Probability and Mathematical Statistics). New York, NY, USA: Wiley, 2012.
- [17] J. A. Bucklew, *Introduction to Rare Event Simulation* (Springer Series in Statistics). New York, NY, USA: Springer, 2010.
- [18] K. Kashima and R. Kawai, "On weak approximation of stochastic differential equations through hard bounds by mathematical programming," *SIAM J. Sci. Comput.*, vol. 35, no. 1, pp. A1–A21, Jan. 2013.
- [19] P. Bernard, "Stochastic linearization: What is available and what is not," *Comput. Struct.*, vol. 67, pp. 9–18, 1998.
- [20] S. H. Crandall, "A half-century of stochastic equivalent linearization," *Struct. Control Health Monit.*, vol. 13, no. 1, pp. 27–40, Jan. 2006.
- [21] R. N. Bhattacharya, "Criteria for recurrence and existence of invariant measures for multidimensional diffusions," *Ann. Probab.*, vol. 6, no. 4, pp. 541–553, Aug. 1978.
- [22] S. Albeverio, V. Bogachev, and M. Röckner, "On uniqueness of invariant measures for finite- and infinite-dimensional diffusions," *Commun. Pure Appl. Math.*, vol. 52, no. 3, pp. 325–362, Mar. 1999.
- [23] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, corrected ed. New York, NY, USA: Springer, 2010.
- [24] B. Grigelionis, *Student's t-Distribution and Related Stochastic Processes* (Springer Briefs in Statistics). New York, NY, USA: Springer, 2013.
- [25] *MATLAB, R2015b*, The MathWorks Inc., Natick, MA, USA, 2015.
- [26] J. C. Spall, "The Kantorovich inequality for error analysis of the Kalman filter with unknown noise distributions," *Automatica*, vol. 31, no. 10, pp. 1513–1517, 1995.
- [27] J. L. Maryak, J. C. Spall, and B. D. Heydon, "Use of the Kalman filter for inference in state-space models with unknown noise distributions," *IEEE Trans. Autom. Control*, vol. 49, no. 1, pp. 87–90, Jan. 2004.
- [28] P. T. Kabamba, S. M. Meerkov, and H. R. Ossareh, "Quasilinear control approach to wind farm power control," in *Proc. 52nd IEEE Conf. Decis. Control*, Dec. 2013, pp. 1307–1312.
- [29] K. Kashima, M. Kato, J.-i. Imura, and K. Aihara, "Probabilistic evaluation of interconnectable capacity for wind power generation," *Eur. Phys. J. Special Topics*, vol. 223, no. 12, pp. 2493–2501, Sep. 2014.
- [30] R. Kawai, "Sample path generation of Lévy-driven continuous-time autoregressive moving average processes," *Methodol. Comput. Appl. Probab.*, vol. 19, no. 1, pp. 175–211, Mar. 2017.
- [31] J. P. Nolan, "Multivariate elliptically contoured stable distributions: Theory and estimation," *Comput. Statist.*, vol. 28, no. 5, pp. 2067–2089, Oct. 2013.



Kenji Kashima (S'01–M'06) was born in Oita, Japan, in 1977. He received the B.Sc. degree in engineering in 2000 and the M.Sc. and Ph.D. degrees in informatics in 2002 and 2005, respectively, all from Kyoto University, Kyoto, Japan.

He was an Assistant Professor with the Graduate School of Information Science and Engineering, Tokyo Institute of Technology, from 2005 to 2011. From April 2010 to March 2011, he was with Universität Stuttgart. He was an Associate Professor with the Graduate School of

Engineering Science, Osaka University, from 2011 to 2013. Since 2013, he has been with the Graduate School of Informatics, Kyoto University, where he is currently an Associate Professor. His research interests include system and control theory for distributed and stochastic phenomena in large-scale dynamical systems, as well as their applications.

Dr. Kashima was a semiplenary speaker at the 19th International Symposium on the Mathematical Theory of Networks and Systems. He was a recipient of the Humboldt Research Fellowship for Experienced Researchers from the Alexander von Humboldt Foundation, Germany, in 2010, and the Pioneer Award of SICE Control Division in 2012. He has been an Associate Editor for the IEEE TRANSACTIONS OF AUTOMATIC CONTROL since 2017, the IEEE CSS Conference Editorial Board since 2011, and the *Asian Journal of Control* since 2014.



Yoshito Ohta (S'83–A'84–M'86–SM'07) received the Dr.Eng. degree in electronic engineering from Osaka University, Suita, Japan, in 1986.

From 1986 to 1988, he was a Visiting Scientist with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology. Since 2006, he has been a Professor with the Department of Applied Mathematics and Physics, Kyoto University, Kyoto, Japan. His research interests include modeling of control

systems, networked control systems, and robust control.

Dr. Ohta was an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL from 2001 to 2005 and a member of the Board of Governors of the IEEE Control Systems Society from 2008 to 2010. He was the General Chair of the 54th IEEE Conference on Decision and Control, Osaka, Japan, in 2015. He is an Associate Editor for the *European Journal of Control* and the Editor-in-Chief of the *SICE Journal of Control, Measurement, and System Integration*.



Hiroki Aoyama was born in Kumamoto, Japan, in 1990. He received the B.Sc. degree in engineering and the M.Sc. degree in informatics from Kyoto University, Kyoto, Japan, in 2014 and 2016, respectively.

He then joined IHI Corporation, Tokyo, Japan, where he is currently with the Control Systems Engineering Department, where research on the accessories (components that configure engine control system, e.g., sensors, pumps, actuators, and engine control unit) of aircraft engines is

conducted. He is involved in designing the control rule and analyzing effects that noise gives in an engine control system.