Incremental Nonlinear Stability Analysis for Stochastic Systems Perturbed by Lévy Noise

SooJean Han and Soon-Jo Chung, *Senior Member, IEEE*

***Abstract*— The notion of a stochastic system is often thought of as a deterministic system perturbed by white noise which arises from factors such as measurement error. However, there are many instances where the system suffers from large perturbations that are unable to be represented by pure white noise. Motivated by this modeling flaw, this paper builds a theoretical framework for stability analysis and controller design of systems that are affected by large, purely discontinuous noise generated according to a Poisson process. We refer to such noise as "shot noise". The concept of stability is defined for systems with shot noise, and conditions to achieve global exponential stability towards a bounded error ball about the nominal system behavior are derived with the use of incremental stability analysis. This will help easily segway into a treatment of stochastic systems that are perturbed by the general Lévy noise, which is the white and shot noises combined.**

***Index Terms*— Nonlinear control systems, Poisson processes, Random processes, Robust stability, Stability analysis, Stability criteria, Stochastic processes, Stochastic systems, Uncertain systems**

I. Introduction

WHEN the extension of deterministic concepts to the stochastic case are made, only white noise is typically accounted for. Additive white Gaussian noise (AWGN) is a popular noise model that is used in the field of information theory and extensive theory has been developed around it [1]. In applications that are more relevant to robotics, such as spacecraft navigation [2], vision-based localization/mapping [3], and motion-planning [4], white noise is often used to model measurement noise. In fact, in many engineering scenarios, white noise appears to be the initial go-to model when anything stochastic is considered. This is not surprising because white noise has many desirable properties that make it easy to work with. For example, a linear combination of Gaussian random vectors is still Gaussian-distributed, and Gaussian distributions arise quite commonly in practice, according to the Central Limit Theorem (CLT).

However, a major flaw with pure white noise models is their lack of generality. White noise is small in magnitude, continuous in the sense that large drifts occur steadily over time, and affects a system persistently for a measurable duration of time. It does not account for large instantaneous disturbances. Consider a robotic flight system experiencing a sudden gust of wind, a collision against a tree branch, or a single rough push by a human. The magnitude of the perturbation is too large for it to be considered white noise and occurs as a single impulse in time. Such types of noise are referred to as *shot noise* [5] or *Poisson noise*; we will be using the two terms interchangeably throughout this work. An intuitive visualization of a pure white noise process, and a white noise process superimposed atop shot noise impulses are shown in Fig. 1. This decomposition of a general *Lévy noise process* into a continuous part Brownian motion and a purely discontinuous part Poisson random measure (to be formally defined in the present paper) has been shown to be useful in financial economics [6] and in signal processing for neuroscience applications [7]. However, to the authors' knowledge, the inclusion of shot noise effects in practical robots and physical systems is still limited.

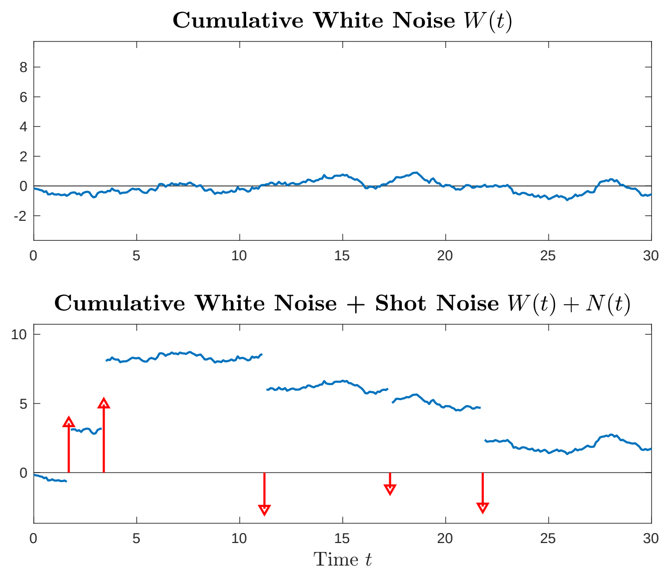

Fig. 1: In the top, a pure white noise process. In the bottom, a white noise process superimposed atop shot noise impulses. As a result of the shot noise, the effect over time is more severe (larger drift) than pure white noise.

With this motivation, there are two questions that are of interest to us in this work. First, the *stability problem*: for some given system dynamics, what class of shot noise perturbations can it handle? Second, the *control/estimation problem*: for

SooJean Han is with the Department of Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125, USA. Email: soojean@caltech.edu.

Soon-Jo Chung is with the Graduate Aerospace Laboratory, California Institute of Technology, Pasadena, CA 91125, USA. Email: sjchung@caltech.edu.

given system dynamics, how should controllers be designed such that it is robust to a class of shot noise perturbations broader than the baseline class determined in the answer to the first question?

A. Related Work

Some Lyapunov function characterizations of stability for nonlinear systems are discussed in [8], and the well-known direct and indirect methods of Lyapunov can be found in any standard control theory textbook (e.g. [9], [10]). Extensions to Lyapunov stability for nonlinear stochastic systems are discussed in [11]–[13]. Stabilization for stochastic nonlinear systems perturbed by noise of unknown covariance is discussed in [14], but it is still Gaussian white noise type. Asymptotic stability of systems driven by combined white noise and jump processes is developed in [15] and an exponential stability result is derived in [16]. In contrast, we choose to utilize contraction theoretic characterizations of stability in this paper rather than a Lyapunov one because it allows us to generalize by considering *incremental stability* of solution trajectories with respect to any desired time-varying trajectory, with guaranteed global exponential convergence towards this desired trajectory. Incremental stability of multiple trajectories arises in numerous applications such as cooperative control over multi-agent swarm systems [17] and directed networks [18], [19]. Further connections between the Lyapunov and incremental stability can be found in [20].

There has been an extensive amount of work done on deciding incremental stability for deterministic nonlinear systems [21], [22]. The work of [23] takes the first step in extending incremental stability to stochastic systems driven by white Gaussian noise. Stability is considered with respect to both constant and time-dependent metrics. Informally, a metric is a coordinate transformation of the original system dynamics which allows for easier stability analysis of said system. Finally, [24] extends the theory in [23] to more general state-dependent metrics, which is useful in their state-dependent construction of nonlinear observers or controllers.

We emphasize that by introducing jump-discontinuous shot noise, we are able to handle noise more generally than what discretizing our system dynamics would allow. Namely, the time of arrival of each jump can be allowed to take continuous values (in the Poisson case, interarrival times between jumps are exponentially distributed), but using a fixed-sample-time discrete system would restrict our ability to model these times accurately. This distinction is important to make because it illustrates how our noise modeling approach differs from the modeling of hybrid systems as in [25]–[27]. An analysis of contraction theory for hybrid systems is done in [28].

Extensive work on Hamilton-Jacobi-Bellman (HJB) based control methods for deterministic systems with deterministic disturbances have been studied in the past [29], [30] and such HJB methods have been modified for the stochastic control of systems with white and shot noise in [31], with applications to finance. A fuzzy interpolation approach is developed in [32] to stabilizing nonlinear systems with shot noise in this way while avoiding the need to solve the Hamilton-Jacobi inequality directly. Machine learning approaches to the rejection of general non-white disturbances have been considered in [33], [34]. But to account for scenarios with insufficient training data or training time, and potentially expensive failures that may occur during the learning process (e.g, a quadrotor part breaks when performing a certain maneuver), we are focused on developing an analytical framework that can be applied to minimize the use of machine learning in the controller design as much as possible.

We organize this paper as follows. The fundamental contraction theorem and the robust contraction theorem for bounded, deterministic disturbances are briefly reviewed in Section II. In Section III, we provide a uniqueness and existence result for SDEs perturbed by Poisson noise, which we then use in Section IV. In Section II, we extend contraction-theoretic results to systems with Poisson noise and in particular, we address the stability problem by proving globally exponential convergence towards a bounded error ball via incremental stability analysis. In Section V, we numerically demonstrate that trajectories can still remain bounded under shot noise and illustrate that a simple LQR scheme can still reasonably control linear systems affected by shot noise. We conclude the paper and provide directions for future work in Section VI.

II. Contraction Analysis for Incremental Stability

We begin with a brief review of established deterministic contraction results that we will extend to the setting of systems perturbed by both white noise and jump noise. A more comprehensive study, including any omitted proofs to the theorems presented, can be found in [17], [23], [35]–[37].

Definition 1. The deterministic, noiseless system $\dot{\mathbf{x}} = f(t, \mathbf{x})$ is said to be *incrementally (globally exponentially) stable* if there exist constants $\kappa, \beta > 0$ such that

$$\|\mathbf{x}(t) - \mathbf{y}(t)\| \leq \kappa \|\mathbf{x}_0 - \mathbf{y}_0\| e^{-\beta t} \tag{1}$$

for all $t \geq 0$ and for all solution trajectories $\mathbf{x}(t)$ and $\mathbf{y}(t)$ of the system with respective initial conditions $\mathbf{x}_0$ and $\mathbf{y}_0$. Assume $\mathbf{x}_0 \neq \mathbf{y}_0$, otherwise the two trajectories are exactly the same for all t and (1) is trivially satisfied with equality.

Before we formally define what a contracting system is, we will derive an intuitive analogy. For $\mathbf{x} \in \mathbb{R}^n$, we denote $\delta\mathbf{x} := (\delta x_1, \cdots, \delta x_n)$ to be the infinitesimal displacement length over a fixed interval of time Δt. We construct a virtual volume out of the dynamics such that it is proportional to $\delta\mathbf{x}^T\delta\mathbf{x} = \|\delta\mathbf{x}\|^2$. A contracting volume is indicative of a contracting system.

Furthermore, the rate of change of $\delta\mathbf{x}$ with time can be approximated by multiplying the slope of the tangent line by the displacement in time. This is exactly how we would approximate the derivative in basic calculus.

$$\delta\mathbf{x}(t_k) = \delta\mathbf{x}(t_{k-1}) + \left(\frac{\partial f}{\partial \mathbf{x}}\delta\mathbf{x}(t_{k-1})\right)\Delta t \Longrightarrow \delta\dot{\mathbf{x}} = \frac{\partial f}{\partial \mathbf{x}}\delta\mathbf{x}$$

A main tool that will be used throughout is the *Comparison Lemma* [9], which can be simply stated as follows. Suppose we have an initial-value problem of the form $\{\dot{u} = g(u,t), u(0) = u_0\}$ and corresponding solution $u(t)$. Then if we were to consider an analogous problem $\{\dot{v} \le g(v,t), v(0) \le u_0\}$, the solution $v(t)$ satisfies $v(t) \le u(t)$ for all $t \ge 0$.

Theorem 1 (Basic Contraction Theorem)**.** If there exists a uniformly positive definite matrix $S(t,\mathbf{x}) := \Theta(t,\mathbf{x})^T\Theta(t,\mathbf{x})$ for Θ being some smooth invertible square matrix that represents the differential coordinate transformation of $\delta\mathbf{z} := \Theta\delta\mathbf{x}$, and if the generalized Jacobian F is uniformly negative definite,

$$F := \left(\dot{\Theta}(t,\mathbf{x}) + \Theta(t,\mathbf{x})\frac{\partial f}{\partial \mathbf{x}}\right)\Theta(t,\mathbf{x})^{-1} \le -\alpha I \tag{2}$$

for some $\alpha > 0$ and $\forall t \ge 0$, then all system trajectories converge globally exponentially to a single trajectory with a convergence rate equal to α. Such a system is called *contracting* with respect to the Riemannian metric associated with $S(t,\mathbf{x})$.

The introduction of a metric $S(t,\mathbf{x})$ is made because one may find it hard to determine from the form of the original dynamics whether a system is contracting or not. It may be easier to analyze the system within a different coordinate frame. But oftentimes, constructing a metric $S(t,\mathbf{x})$ is just as hard as finding a Lyapunov function to identify stability.

One can also extend Theorem 1 to systems with an additive bounded deterministic perturbation: $\dot{\mathbf{x}} = f(t,\mathbf{x})+w(t,\mathbf{x})$ such that $\sup_{t,\mathbf{x}} \|w(t,\mathbf{x})\| \le c$ for some $c > 0$. Note that because of this perturbation, the dynamics of the infinitesimal length $\delta\mathbf{z}$ become $\delta\dot{\mathbf{z}} = \dot{\Theta}\delta\mathbf{x} + \Theta\delta\dot{\mathbf{x}} = F\delta\mathbf{z} + \Theta\delta\mathbf{w}$, This is usually referred to as the *virtual dynamics*.

Theorem 2 (Robust Contraction Theorem)**.** Let $\mathbf{x}_1(t)$ be the solution of the unperturbed system (U) : $\dot{\mathbf{x}} = f(t,\mathbf{x})$, and assume (U) is contracting with a rate of α. Let $\mathbf{x}_2(t)$ be the solution of the perturbed system (P) : $\dot{\mathbf{x}} = f(t,\mathbf{x})+w(t,\mathbf{x})$.

Define the path integral to be the shortest path/"distance" between the two trajectories:

$$R(t) = \int_{\mathbf{x}_1}^{\mathbf{x}_2} \|\delta\mathbf{z}(t)\| \quad \forall t \ge 0 \tag{3}$$

Then we have:

$$R(t) \le R(0)e^{-\alpha t} + \frac{1}{\alpha}\sup_{\tau,\mathbf{x}} \|\Theta(\tau,\mathbf{x})w(\tau,\mathbf{x})\| \left(1 - e^{-\alpha t}\right) \tag{4}$$

from which we can conclude that the distance exponentially converges to the following error ball:

$$\lim_{t\to\infty} R(t) \le \sup_{\tau,\mathbf{x}} \frac{1}{\alpha} \|\Theta(\tau,\mathbf{x})w(\tau,\mathbf{x})\| \tag{5}$$

III. Background for the Poisson SDE

Theorem 2 is insufficient for stochastic systems because this supremum value $\sup\|w\|$ is kept constant and the w is deterministic and a bounded-variation process. Informally speaking, a function is of bounded variation if it is "well-behaved" and does not vary wildly within a small increment of time. For instance, the Brownian motion process has *unbounded variation*, yet the variance is kept proportional to the approximate incremental timestep, which decreases to zero. So there is a very small probability that consecutive random variables will have a large discrepancy between them [38].

We consider processes that are perturbed by random noise not only of a Gaussian type, but also of sudden discrete jumps whose distributions are better modeled as a Poisson process (see Fig. 1). From here on, we use the term "jumps" to describe individual arrivals of a shot noise process.

The analysis of such systems requires a modification of the standard Itô integral and Itô formula for SDEs which only involve Gaussian white noise (i.e. Brownian motion). We will establish a few definitions and properties before delving into existence and uniqueness theorems for SDEs perturbed by shot noise (Section III-B), and the contraction theorems that arise (Section IV).

A. Preliminaries

Definition 2. Let E be a Borel-measurable set such that $0 \in \overline{E}$, where $\overline{E}$ denotes the closure of E. We will refer to E as the *jump space*. Consider a random measure $\overline{N}([0,T]\times E)$ on set E until time $T > 0$ with *intensity measure* Leb $\times\, \nu$, where Leb denotes the standard Lebesgue measure. We will denote the corresponding *intensity (parameter)* as $\lambda := \nu(E)$. It is important to notice the difference between Leb (the measure in time) and ν (the probability measure describing the distribution of the jump "size"). Then $\overline{N}$ is called a *Poisson random measure* if the following are satisfied:

1) if $E_1, \cdots, E_n$ are pairwise disjoint elements of E, then $\overline{N}([0,T]\times E_1), \cdots, \overline{N}([0,T]\times E_n)$ are independent.
2) for each $E_i \subseteq E$, $\overline{N}([0,T]\times E_i)$ is distributed as a Poisson random variable with intensity parameter $\lambda_i := \nu(E_i)$.

Every Poisson random measure $\overline{N}([0,T]\times E)$ has an associated *standard Poisson process* (counting process) $N(t)$, which counts the number of jumps in E that have occurred in the time interval $[0,t]$ for $t \le T$. We use the subscript 1 because it is a special case of the Poisson random measure with jump space $E = \{1\}$. Both the random measure and its counting process have the same intensity λ. Denote the time of the i-th arrival with random variable T_i. The mean is given by $\mathbb{E}[N(t)] = \lambda t$ and the variance is $\mathrm{Var}(N(t)) = \lambda t$. The *compensated Poisson process* $N(t) - \lambda t$ is a martingale with respect to $\mathcal{F}_t := \sigma(N(s) : s \in [0,t))$. This results from the fact that $\mathbb{E}[N(t) - \lambda t] = 0$, i.e., has centered increments.

Example 1. A Poisson random measure with heights taking more than one value in E can also be expressed as a *compound Poisson process* $Y(t)$; it is essentially a standard Poisson process where the size of the ith jump is a random variable ξ_i not necessarily equal to 1. We write it as the cumulative sum $Y(t) = \sum_{i=1}^{N(t)} \xi_i$. Again, we denote arrival times by T_i and its intensity by λ.

To make the illustration more concrete, consider a scalar process which models the arrivals of vehicles in a particular parking space and the variables ξ_k which represent the type of the vehicle: $\xi_k = i$ if it is type i, where $i \in E := \{1, \cdots, K\}$. Then we can partition the space into K disjoint sets $E_i =$

$\{i\}, i = 1, \cdots, K$. For each i, we associate a counting process $N_{E_i}(t)$ with intensity λ_i. It counts the number of arrivals of type i there are (i.e. number of jumps in set E_i) in the time interval $[0, t]$. Now suppose an entering vehicle is type i with probability p_i, where $\sum_{i=1}^{K} p_i = 1$. Then the intensity of each split process is related to the intensity λ of the original arrival process via $\lambda_i = p_i \lambda$.

Let $g : [0, T] \times E \to \mathbb{R}$ be a Borel-measurable function and $\overline{N}$ be a Poisson random measure on $[0, T] \times E$ with intensity measure Leb $\times\, \nu$. We define the *Poisson integral* of g as follows:

$$I_g := \int_{[0,T]\times E} g(t, y)\overline{N}(dt, dy) \tag{6}$$

In contrast to the previous Example 1, we use lowercase y to denote the jump sizes instead of ξ.

A well-known interpretation of this integral (see [39] for a formula with time-invariant g) is the following summation form:

$$I_g = \sum_{0<t\le T} g(t, \Delta\overline{N}(t))\mathbb{1}_E\{\Delta\overline{N}(t)\} \tag{7}$$

where $\Delta\overline{N}(t) = \overline{N}(t) - \overline{N}(t-)$, and the indicator $\mathbb{1}_E\left\{\Delta\overline{N}(t)\right\}$ is equal to 1 if a jump occurred at time t and 0 otherwise. The subscript E in the indicator is to show that a jump is counted only if it satisfies a membership property of E. For example, if $E = (-\infty, -1] \cup [1, \infty)$, then any $\Delta\overline{N}(t)$ such that $|\Delta\overline{N}(t)| < 1$ is not considered a jump.

Overall, the definition of the integral has quite an intuitive interpretation: if we think of the noise process as a sequence of impulses where the ith impulse arrives at time T_i, then integrating a function g with respect to it over an interval of time $[0, T]$ would only pick out the values of g at $T_i \in [0, T]$.

Example 2. We continue from Example 1. Let $g(t, y) = g(y)$ be the time-independent cost of parking in the parking site, which is a function of the vehicle type in the following way: $g(y) = 10y$. Partition the space according to Example 1. Then evaluating I_g over a fixed time period $[0, T]$ is equivalent to determining the total cost of all vehicles which have entered and parked in this time. Suppose that for a specific instance, the arrival process of vehicle types (i.e. the jump heights of the process $Y(t)$) during this time interval is $\{2, 2, 3, 2, 3, 1, 2\}$ so that there have been $N(T) = 7$ total arrivals. We get

$$\begin{aligned} I_g &= \sum_{i=1}^{7}\sum_{k=1}^{3} g(T_i, \Delta Y(T_i))\mathbb{1}_{E_k}\{\Delta Y(T_i)\} \\ &= g(2) + g(2) + g(3) + g(2) + g(3) + g(1) + g(2) = 150 \end{aligned}$$

Another needed result for our analysis is the following: the simplest version of a well-known formula called *Campbell's Formula*. The version in [40] presents the formula for time-invariant functions g; here, we extend the result more generally to functions g that also depend on time.

Lemma 1 (Basic Campbell's Formula)**.** Let $g : [0, \infty) \times E \to \mathbb{R}$ be a Borel-measurable function and $\overline{N}([0, T] \times E)$ denote the Poisson random measure with intensity $\lambda := \nu(E)$ over the jump space E. Then

$$\mathbb{E}[I_g] = \int_{[0,T]\times E} \mathbb{E}[g(t, y)]dt\nu(dy) \tag{8}$$

In order for the integral (6) to be well-defined, we need:

$$\int_{[0,T]\times E} |g(t, y)|dt\nu(dy) < \infty \text{ a.s.} \tag{9}$$

where "a.s." denotes shorthand for "almost-surely". Campbell's formula and other uses of this integral are stated while assuming this property is satisfied.

Proof: The compensated integral is given by replacing $\overline{N}(dt, dy)$ with a compensated Poisson process:

$$\int_{[0,T]\times E} g(t, y)(\overline{N}(dt, dy) - dt\nu(dy))$$

Because the compensated Poisson is a martingale, this integral has mean zero. Splitting this apart yields our desired result:

$$\mathbb{E}\left[\int_{[0,T]\times E} g(t, y)(\overline{N}(dt, dy) - dt\nu(dy))\right] = 0$$

which implies

$$\begin{aligned} \mathbb{E}\left[\int_{[0,T]\times E} g(t, y)\overline{N}(dt, dy)\right] &= \mathbb{E}\left[\int_{[0,T]\times E} g(t, y)dt\nu(dy)\right] \\ &= \int_{[0,T]\times E} \mathbb{E}[g(t, y)]dt\nu(dy) \end{aligned}$$

where the last equality results from Fubini's theorem. □

Remark 1. A special case of the formula in Lemma 1 is when the function g is deterministic and independent of time.

$$\mathbb{E}\left[\int_{[0,T]\times E} g(y)N(dt, dy)\right] = T\int_E g(y)\nu(dy) \tag{10}$$

Essentially, because the function is independent of time, the integral over t becomes just a constant multiplication.

Finally, we discuss one additional property of the Poisson integral [39], [41] which will be used in the proof of the existence and uniqueness theorem to be presented in the following subsection.

Lemma 2. If the following properties hold:

$$\int_{[0,T]\times E} |g(t, y)|dt\nu(dy) < \infty \tag{11a}$$

$$\int_{[0,T]\times E} g^2(t, y)dt\nu(dy) < \infty \tag{11b}$$

then

$$\begin{aligned} \mathbb{E}[I_g^2] = &\int_{[0,T]\times E} \mathbb{E}[g(t, y)^2]dt\nu(dy) \\ &+ \left(\int_{[0,T]\times E} \mathbb{E}[g(t, y)]dt\nu(dy)\right)^2 \end{aligned} \tag{12}$$

We defer the proof of this property to Appendix I.

Definition 3. A process $L(t)$ is said to be a *Lévy process* if all paths of L are cádlág, $\mathbb{P}(L(0) = 0) = 1$, and L has stationary and independent increments.

Note that both white noise and the Poisson process are Lévy processes. The compound Poisson process is also a Lévy process. A linear combination of white and shot noise, such as in Fig. 1 is also a Lévy process. Thus, modeling noise as a Lévy process enables us to generalize it. The *Lévy-Khintchine Decomposition Formula*, which is presented in Theorem 6, states that any Lévy noise process can be expressed as a combination of white and shot noise. As mentioned, white noise has been studied extensively in the past literature, so we will focus primarily on Poisson noise throughout this paper.

B. Existence and Uniqueness of Solutions

We consider systems that can be expressed as SDEs of the following form:

$$d\mathbf{x}(t) = f(t, \mathbf{x})dt + \sigma(t, \mathbf{x})dW(t) + \xi(t, \mathbf{x})dN(t) \tag{13}$$

where $f : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}^n$ is the unperturbed part of the system, $\sigma : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$ is the variation of the white noise, $W : \mathbb{R}^+ \to \mathbb{R}^d$ is a d-dimensional Brownian motion process, $\xi : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^{n\times \ell}$ is the size of the jumps that occur, and $N(t)$ is the ℓ-dimensional standard Poisson random measure with intensity $\lambda := \nu(\{1\}^\ell)$. Hence, $N(t) := N([0, t] \times \{1\}^\ell)$. Starting from here, we will take $N(t)$ to mean $N_1(t)$ (as defined in the previous subsection with $\ell = 1$), and we write it without the 1 subscript for simplicity.

Note that when $\xi(t, \mathbf{x}) \equiv 0$, we recover the standard white or Gaussian noise SDE:

$$d\mathbf{x}(t) = f(t, \mathbf{x})dt + \sigma(t, \mathbf{x})dW(t) \tag{14}$$

The conditions for existence and uniqueness of solutions are the standard ones (see [42]): the functions must be Lipschitz with respect to the time argument and have bounded growth with respect to the state argument.

A similar result can be derived in the case of shot noise perturbations. For the moment, we consider SDEs with $\sigma(t, \mathbf{x}) \equiv 0$.

$$d\mathbf{x}(t) = f(t, \mathbf{x})dt + \xi(t, \mathbf{x})dN(t) \tag{15}$$

First, we present the well-known Gronwall inequality, a standard result of which can be found in any classical control-theoretic textbook (e.g. [9], [10]).

Lemma 3 (Gronwall inequality)**.** Let $I = [t_1, t_2] \subset \mathbb{R}$ and $\phi, \psi, \rho : I \to \mathbb{R}^+$ be continuous, nonnegative functions.

If the following inequality holds true:

$$\phi(t) \leq \psi(t) + \int_{t_1}^{t} \rho(s)\phi(s)ds \quad \forall\, t \in [t_1, t_2] \tag{16}$$

Then it follows that

$$\phi(t) \leq \psi(t) + \int_{t_1}^{t} \psi(s)\rho(s)e^{\int_s^t \rho(\tau)d\tau}ds \tag{17}$$

Theorem 3 (Existence and Uniqueness for SDE (15))**.** For fixed $T > 0$, let $f : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\xi : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n\times\ell}$ be measurable functions satisfying the following conditions

1) Lipschitz: $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, t \in [0, T]$,

$$\begin{aligned}\|f(t, \mathbf{x}) - f(t, \mathbf{y})\| + \|\xi(t, \mathbf{x}) - \xi(t, \mathbf{y})\|_F \\ \leq K\|\mathbf{x} - \mathbf{y}\|\end{aligned} \tag{18}$$

2) Bounded growth: $\forall \mathbf{x} \in \mathbb{R}^n, t \in [0, T]$

$$\|f(t, \mathbf{x})\|^2 + \|\xi(t, \mathbf{x})\|_F^2 \leq C(1 + \|\mathbf{x}\|^2) \tag{19}$$

for positive constants C and K where the norm on ξ is the Frobenius norm and the norms on the vector-valued functions are any vector norm. Further, let $\mathbf{x}_0 \in \mathbb{R}^n$ have $\mathbb{E}[\|\mathbf{x}_0\|] < \infty$ and be independent of the noise processes. Then the SDE (15) with initial condition $\mathbf{x}(0) = \mathbf{x}_0$ has a unique solution $\mathbf{x}(t)$ adapted to the filtration $\mathcal{F}_t$ generated by $\mathbf{x}_0$ and $N(s)$, where $s \leq t$ and

$$\mathbb{E}\left[\int_0^T \|\mathbf{x}(t)\|^2 dt\right] < \infty$$

Remark 2. Following the construction of (6), we take the definition of the Poisson integral used in the Lipschitz and bounded growth conditions to be

$$\int_0^t \xi(s, \mathbf{x})dN(s) = \sum_{i=1}^{N_1(t)} \xi(T_i, \mathbf{x}(T_i)) \tag{20}$$

since the size of every jump of the standard Poisson process is 1. We have essentially isolated the size of each jump as a multiplicative factor $\xi(t, \mathbf{x})$. From here on out in our analysis, we will be using this multiplicative form instead of the general Poisson random measure $N(dt, dy)$. Thus, we can express the Poisson integral as:

$$\begin{aligned}\mathbb{E}\left[\int_0^t \xi(s, \mathbf{x})dN(s)\right] &= \mathbb{E}\left[\int_{[0,t]\times\{1\}^\ell} \xi(s, \mathbf{x})N(ds, dy)\right] \\ &= \int_{[0,t]\times\{1\}^\ell} \mathbb{E}\left[\xi(s, \mathbf{x})\right] ds\nu(dy) \\ &= \lambda \int_0^t \mathbb{E}\left[\xi(s, \mathbf{x})\right] ds\end{aligned} \tag{21}$$

Remark 3. Theorem 3 was also presented in [31], but without a proof. We present a complete proof here. There has been previous work done on describing such conditions for solutions to SDEs of the form (15) while imposing different, non-Lipschitz conditions on f and ξ. For instance, [43] relaxes the Lipschitz conditions by instead assuming that f and ξ are bounded above by any concave function of the normed difference in trajectories $\|\mathbf{x} - \mathbf{y}\|$. Alternatively, [44] presents a result for conditions where f is upper-bounded in norm by a constant and the bound on ξ depends on the size of the jump (which is not easily applicable to our case because we are only considering standard Poisson process noise, i.e. the jump size is always one). We choose to work with simple Lipschitz conditions because it is easier to relate to the well-known white noise version.

Proof: The proof is similar to that of pure white noise, e.g. [42]. First we construct an approximate sequence using Picard iterations, recursively defined as

$$\mathbf{z}^{(n)}(t) = \mathbf{z}_0^{(n)} + \int_0^t f(s, \mathbf{z}^{(n-1)}(s))ds + \int_0^t \xi(s, \mathbf{z}^{(n-1)}(s))dN(s) \tag{22}$$

where $n \in \mathbb{N}$.

Taking the difference between two trajectories $\mathbf{z}^{(n)}(t)$ and $\mathbf{z}^{(m)}(t)$ results in

$$\mathbf{z}^{(n,m)}(t) = \mathbf{z}_0^{(n,m)} + \int_0^t f^{(n-1,m-1)}(s)ds + \int_0^t \xi^{(n-1,m-1)}(s)dN(s) \tag{23}$$

with $n, m \in \mathbb{N}$ and the notation

$$\begin{aligned}
&\mathbf{z}^{(n,m)}(t) := \mathbf{z}^{(n)}(t) - \mathbf{z}^{(m)}(t),\ \mathbf{z}^{(n,m)}(0) := \mathbf{z}^{(n,m)}(0)\\
&f^{(n,m)}(t) := f(t, \mathbf{z}^{(n)}(t)) - f(t, \mathbf{z}^{(m)}(t))\\
&\xi^{(n,m)}(t) := \xi(t, \mathbf{z}^{(n)}(t)) - \xi(t, \mathbf{z}^{(m)}(t))
\end{aligned}$$

Taking the mean-squared difference, and applying the triangle and Cauchy-Schwarz inequalities leads to

$$\begin{aligned}
\mathbb{E}\left[\left\|\mathbf{z}^{(n,m)}(t)\right\|^2\right] &\leq \mathbb{E}\left[\left(\left\|\mathbf{z}_0^{(n,m)}\right\| + \int_0^t \left\|f^{(n-1,m-1)}(s)\right\| ds + \int_0^t \left\|\xi^{(n-1,m-1)}(s)\right\|_F dN(s)\right)^2\right]\\
&\leq 3\mathbb{E}\left[\left\|\mathbf{z}_0^{(n,m)}\right\|^2\right] + 3\mathbb{E}\left[\int_0^t ds \int_0^t \left\|f^{(n-1,m-1)}(s)\right\|^2 ds\right]\\
&\quad + 3\mathbb{E}\left[\left(\int_0^t \left\|\xi^{(n-1,m-1)}(s)\right\|_F dN(s)\right)^2\right]
\end{aligned} \tag{24}$$

Note that the Lipschitz bound (18) can be squared on both sides:

$$\begin{aligned}
&\|f(t,\mathbf{x}) - f(t,\mathbf{y})\|^2 + \|\xi(s,\mathbf{x}) - \xi(s,\mathbf{y})\|_F^2 +\\
&2\|f(t,\mathbf{x}) - f(t,\mathbf{y})\| \|\xi(s,\mathbf{x}) - \xi(s,\mathbf{y})\|_F \leq K^2 \|\mathbf{x} - \mathbf{y}\|^2
\end{aligned} \tag{25}$$

Because norms are nonnegative and the integral of nonnegative functions (whether it is standard ds or Poisson $dN(s)$) is also nonnegative, the bound also holds for each individual term in the left-hand side sum. Using (25) on the second expectation term yields:

$$\mathbb{E}\left[\int_0^t ds \int_0^t \left\|f^{(n-1,m-1)}(s)\right\|^2 ds\right] \leq K^2 t \int_0^t \mathbb{E}\left[\left\|\mathbf{z}^{(n-1,m-1)}(s)\right\|^2\right] ds \tag{26}$$

and for the final term, we can apply Lemma 2.

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t \left\|\xi^{(n-1,m-1)}(s)\right\|_F dN(s)\right)^2\right]\\
&\quad = \lambda \int_0^t \mathbb{E}\left[\left\|\xi^{(n-1,m-1)}(s)\right\|_F^2\right] ds\\
&\quad\quad + \mathbb{E}\left[\left(\lambda \int_0^t \left\|\xi^{(n-1,m-1)}(s)\right\|_F ds\right)^2\right]
\end{aligned} \tag{27}$$

using the fact that $\lambda := \int_{\{1\}^\ell} \nu(dy)$, as in Remark 2. By Cauchy-Schwarz inequality and the squared Lipschitz bound (25):

$$\begin{aligned}
(27) &\leq K^2\lambda \int_0^t \mathbb{E}\left[\left\|\mathbf{z}^{(n-1,m-1)}(s)\right\|^2\right] ds\\
&\quad + \lambda^2 \mathbb{E}\left[\int_0^t ds \int_0^t \left\|\xi^{(n-1,m-1)}(s)\right\|_F^2 ds\right]\\
&\leq K^2\lambda \int_0^t \mathbb{E}\left[\left\|\mathbf{z}^{(n-1,m-1)}(s)\right\|^2\right] ds\\
&\quad + K^2\lambda^2 t \int_0^t \mathbb{E}\left[\left\|\mathbf{z}^{(n-1,m-1)}(s)\right\|^2\right] ds
\end{aligned} \tag{28}$$

Finally, note that $\mathbf{z}_0^{(n,m)} = 0$ because both trajectories $\mathbf{z}^{(n)}$ and $\mathbf{z}^{(m)}$ begin with the same initial conditions. In combination, we get:

$$\begin{aligned}
\mathbb{E}\left[\left\|\mathbf{z}^{(n)}(t) - \mathbf{z}^{(m)}(t)\right\|^2\right] &\leq 3K^2(t + \lambda + \lambda^2 t)\\
&\quad \times \int_0^t \mathbb{E}\left[\left\|\mathbf{z}^{(n-1)}(s) - \mathbf{z}^{(m-1)}(s)\right\|^2\right] ds
\end{aligned} \tag{29}$$

Choose $n = k+1, m = k$ for $k > 0$. By induction, we get:

$$\mathbb{E}\left[\left\|\mathbf{z}^{(k+1)}(t) - \mathbf{z}^{(k)}(t)\right\|^2\right] \leq \frac{c^k t^{k+1}}{(k+1)!} \quad \forall k \geq 0, t \in [0,T] \tag{30}$$

where $c := 3K^2(T+\lambda+\lambda^2 T)$. From there, it is straightforward to show that $\{\mathbf{z}^{(k)}(t)\}$ is a Cauchy sequence which converges to a limit since $\mathbf{z} \in \mathbb{R}^n$.

To show that the solution is unique, consider two solution trajectories $\mathbf{x}(t,\omega)$ and $\mathbf{y}(t,\omega)$ of (15) with respective initial conditions $\mathbf{x}_0$ and $\mathbf{y}_0$ where ω is a specific sample path of the noise process N. We can apply the same calculations as before on the mean-squared error difference between $\mathbf{x}$ and $\mathbf{y}$ to get

$$\begin{aligned}
\mathbb{E}\left[\|\mathbf{x}(t) - \mathbf{y}(t)\|^2\right] &\leq 3\mathbb{E}\left[\|\mathbf{x}_0 - \mathbf{y}_0\|^2\right]\\
&\quad + c\mathbb{E}\left[\int_0^t \|\mathbf{x}(s) - \mathbf{y}(s)\|^2 ds\right]
\end{aligned} \tag{31}$$

By Gronwall's inequality Lemma 3, (31) becomes

$$\mathbb{E}\left[\|\mathbf{x}(t) - \mathbf{y}(t)\|^2\right] \leq 3\mathbb{E}\left[\|\mathbf{x}_0 - \mathbf{y}_0\|^2\right] e^{ct} \tag{32}$$

Now we set the two initial conditions $\mathbf{x}_0$ and $\mathbf{y}_0$ equal to each other. This implies that $c_1 = 0$ and so $h(t) = 0$ for all $t \geq 0$. Thus,

$$\mathbb{P}\left(\|\mathbf{x} - \mathbf{y}\| = 0\right) = 1 \quad \text{for all } t \geq 0$$

This holds for all sample paths of N. Thus, the solution is indeed unique for all $t \in [0, T]$. The proof is complete. □

Remark 4. We have shown convergence of two trajectories in the mean-squared sense: in expectation, the trajectories will converge toward each other. It is weaker than the almost-sure sense of convergence, meaning we do not guarantee trajectory convergence for every noise process sample paths ω. For a more comprehensive treatment of this topic, see Chapter 2 of [45], [46], and [47], which additionally develops a CLT-like theorem for semimartingales (informally defined, SDEs which involve both a continuous martingale part such as white noise, and a purely discontinuous part such as shot noise).

IV. Stochastic Contraction with Poisson Noise

A. White Noise Review

To prove contraction of the white noise SDE (14), we make the following assumption.

Assumption 1 (Bounded Metric)**.** We will assume that the metric $S(t, \mathbf{x})$ is bounded in both arguments $\mathbf{x}$ and t from above and below, and that its first and second derivatives with respect to the $\mathbf{x}$ argument are also bounded from above. We thus define the following constants

$$\underline{s} = \inf_{t,\mathbf{x}} \lambda_{\text{min}}(S(t, \mathbf{x})), \quad \overline{s} = \sup_{t,\mathbf{x}} \lambda_{\text{max}}(S(t, \mathbf{x})) \tag{33}$$
$$s' = \sup_{t,\mathbf{x},i,j} \|(\partial_x S(t, \mathbf{x}))_{i,j}\|, \quad s'' = \sup_{t,\mathbf{x},i,j} \left\|(\partial_x^2 S(t, \mathbf{x}))_{i,j}\right\|$$

We have previously established contraction as a concept of convergence between different solution trajectories of the system starting from different initial conditions. In the stochastic setting, the difference between trajectories will also stem from using different noise processes. Specifically, we define two trajectories $\mathbf{x}(t)$ and $\mathbf{y}(t)$ as solutions to (14) driven by completely different noise processes with different variances:

$$d\mathbf{x} = f(t, \mathbf{x})dt + \sigma_1(t, \mathbf{x})dW_1(t) \tag{34a}$$
$$d\mathbf{y} = f(t, \mathbf{y})dt + \sigma_2(t, \mathbf{y})dW_2(t) \tag{34b}$$

and for both of these systems stochastically contracting, we will denote η_1 and η_2 such that $\|\sigma_1(t, \mathbf{x})\|_F \le \eta_1$ and $\|\sigma_2(t, \mathbf{y})\|_F \le \eta_2$. Then, the following holds:

$$\text{tr}(\sigma_1^T S \sigma_1(t, \mathbf{x})) \le \overline{s}\eta_1^2, \quad \text{tr}(\sigma_2^T S \sigma_2(t, \mathbf{y})) \le \overline{s}\eta_2^2 \tag{35}$$

We construct a *virtual system* in terms of $\mathbf{z}(t) \in \mathbb{R}^n$ such that its particular solutions are $\mathbf{x}(t)$ and $\mathbf{y}(t)$. If this virtual system is contracting in $S(t, \mathbf{z})$, then $\mathbf{x}$ and $\mathbf{y}$ converge towards each other globally and exponentially fast.

Definition 4 (Virtual System for (14))**.** We can represent the infinitesimal differential length $\delta\mathbf{z}$ as a path integral and reparameterize using measure $\mu \in [0, 1]$

$$\mathbf{y}(t) - \mathbf{x}(t) = \int_{\mathbf{x}}^{\mathbf{y}} \delta\mathbf{z} = \int_0^1 \left(\frac{\partial \mathbf{z}}{\partial \mu}\right) d\mu \tag{36}$$

where $\mu \in [0, 1]$ is a measure parameter such that:

$$\mathbf{z}(\mu = 0, t) = \mathbf{x}(t), \, \mathbf{z}(\mu = 1, t) = \mathbf{y}(t) \tag{37}$$
$$\sigma_{\mu=0}(t, \mathbf{z}) = \sigma_1(t, \mathbf{x}), \, \sigma_{\mu=1}(t, \mathbf{z}) = \sigma_2(t, \mathbf{y})$$
$$W_{\mu=0}(t, \mathbf{z}) = W_1(t), \, W_{\mu=1}(t, \mathbf{z}) = W_2(t)$$

e.g., $\mathbf{z}(\mu, t) := \mu\mathbf{x}(t) + (1-\mu)\mathbf{y}(t)$. Using this, we can rewrite (14) as the virtual system:

$$d\mathbf{z}(\mu, t) = f(t, \mathbf{z}(\mu, t))dt + \sigma_\mu(t, \mathbf{z}(\mu, t))dW_\mu(t) \tag{38}$$

This further enables us to describe the virtual dynamics as follows:

$$d\delta\mathbf{z} = F\delta\mathbf{z}dt + \delta\sigma_\mu dW_\mu \tag{39}$$

where we denote

$$\delta\sigma_\mu = \begin{bmatrix} \frac{\partial \sigma_{\mu,1}}{\partial \mathbf{z}}\delta\mathbf{z} & \cdots & \frac{\partial \sigma_{\mu,d}}{\partial \mathbf{z}}\delta\mathbf{z} \end{bmatrix}$$

for $\sigma_\mu := \begin{bmatrix} \sigma_{\mu,1} & \cdots & \sigma_{\mu,d} \end{bmatrix}$, where $\sigma_{\mu,i}$ is the ith column of σ_μ.

Note that by Cauchy-Schwarz and the triangle inequality for integrals, we can bound the path integral as follows:

$$\|\mathbf{y} - \mathbf{x}\|^2 = \left\|\int_{\mathbf{x}}^{\mathbf{y}} \delta\mathbf{z}\right\|^2 \le \int_{\mathbf{x}}^{\mathbf{y}} \|\delta\mathbf{z}\|^2 = \int_0^1 \left\|\frac{\partial \mathbf{z}}{\partial \mu}\right\|^2 d\mu \tag{40}$$

and multiplying across by $\underline{s}$ yields

$$\begin{aligned} \underline{s}\,\|\mathbf{y} - \mathbf{x}\|^2 &\le \underline{s} \int_0^1 \left\|\frac{\partial \mathbf{z}}{\partial \mu}\right\|^2 d\mu \\ &\le \int_0^1 \left(\frac{\partial \mathbf{z}}{\partial \mu}\right)^T S(t, \mathbf{z}(\mu, t)) \left(\frac{\partial \mathbf{z}}{\partial \mu}\right) d\mu \end{aligned} \tag{41}$$

We will use the right-side expression above as our Lyapunov function:

$$V(t, \mathbf{z}, \delta\mathbf{z}) = \int_0^1 \left(\frac{\partial \mathbf{z}}{\partial \mu}\right)^T S(t, \mathbf{z}(\mu, t)) \left(\frac{\partial \mathbf{z}}{\partial \mu}\right) d\mu \tag{42}$$

Note that there is a dependence of V on $\delta\mathbf{z}$ because we can alternatively (and informally) express $V(\mathbf{z}, \delta\mathbf{z}, t) = \int_{\mathbf{x}}^{\mathbf{y}} \delta\mathbf{z}^T S(t, \mathbf{z})\delta\mathbf{z}$.

As in the proof for existence and uniqueness, stochastic contraction results for both white and shot noise will be examined in the mean-square sense.

Theorem 4 (Stochastic Contraction Theorem for White Noise [24])**.** For the SDE system described in (14), suppose the following two conditions are satisfied:

1) the unperturbed system $\dot{\mathbf{x}} = f(t, \mathbf{x})$ is contracting in $S(t, \mathbf{x})$ at a rate α:

$$\left(\frac{\partial f}{\partial \mathbf{x}}\right)^T S(t, \mathbf{x}) + S(t, \mathbf{x}) \left(\frac{\partial f}{\partial \mathbf{x}}\right) + \dot{S}(t, \mathbf{x}) \le -2\alpha S \tag{43}$$

2) there exists $\eta > 0$ such that $\|\sigma(t, \mathbf{x})\|_F \le \eta$ for all $\mathbf{x} \in \mathbb{R}^n$ and $t \ge 0$. This implies that $\text{tr}(\sigma^T S\sigma(t, \mathbf{x})) \le \overline{s}\eta^2$, and justifies the assumption of (35).

Further assume that the initial conditions adhere to some probability distribution $p(\mathbf{z}_0) = p(\mathbf{x}_0, \mathbf{y}_0)$. Then (14) is stochastically contracting if

$$\mathbb{E}_{\mathbf{z}_0}[V(t, \mathbf{z}, \delta\mathbf{z})] \le V(0, \mathbf{z}_0, \delta\mathbf{z}_0)e^{-\beta t} + \frac{\kappa}{\beta}(1 - e^{-\beta t}) \tag{44}$$

where

$$\beta = \alpha - \frac{1}{2\underline{s}}(\eta_1^2 + \eta_2^2)\left(\frac{s''}{2} + s'\right) \tag{45a}$$

$$\kappa = \frac{1}{2}(\overline{s} + s')(\eta_1^2 + \eta_2^2) \tag{45b}$$

Because the Lyapunov-like function V can be arbitrarily defined for stability analysis, we simplify the inequality further to express it in terms of the given trajectories $\mathbf{x}$ and $\mathbf{y}$. Dividing across by $\underline{s}$ yields:

$$\frac{1}{\underline{s}}\mathbb{E}_{\mathbf{z}_0}[V(t,\mathbf{z},\delta\mathbf{z})] \leq \frac{1}{\underline{s}}\mathbb{E}_{\mathbf{z}_0}[V(0,\mathbf{z}_0,\delta\mathbf{z}_0)]e^{-\beta t} + \frac{\kappa}{\underline{s}\beta}(1-e^{-\beta t})$$

Finally, we lower-bound the left-hand side of the inequality according to the construction in (42). This gives us:

$$\mathbb{E}_{(\mathbf{x}_0,\mathbf{y}_0)}[\|\mathbf{y}-\mathbf{x}\|^2] \leq \frac{1}{\underline{s}}\|\mathbf{y}_0-\mathbf{x}_0\|^2 e^{-\beta t} + \frac{\kappa}{\underline{s}\beta}\left(1-e^{-\beta t}\right) \tag{46}$$

Remark 5. The inequality in Theorem 4 is analogous to the Robust Contraction form (4). In the stochastic form, the path integral (squared) now has an expectation over it because of randomness introduced by the Brownian motion disturbance. Furthermore:

$$\lim_{t\to\infty}\mathbb{E}_{(\mathbf{x}_0,\mathbf{y}_0)}\left[\|\mathbf{y}-\mathbf{x}\|^2\right] \leq \frac{\kappa}{\underline{s}\beta} \tag{47}$$

the right-hand-side of which behaves as $\mathrm{tr}(\sigma^T\sigma(\tau,\mathbf{x}))/\alpha$, similar to the bound (5) for the deterministic case. Finally, the initial conditions in both cases are exponentially forgotten because α and β are both positive constants.

B. Shot Noise

In this subsection, we will present the version of Theorem 4 for Poisson noise perturbed systems of the form (15). We have the same setup and notation as in the white noise case and the same metric assumptions as in Assumption 1. Compare two trajectories: $\mathbf{x}(t)$ as the solution to (15) and $\mathbf{y}(t)$ as the solution to the unperturbed system:

$$d\mathbf{x} = f(t,\mathbf{x})dt + \xi(t,\mathbf{x})dN(t) \tag{48a}$$

$$d\mathbf{y} = f(t,\mathbf{y})dt \tag{48b}$$

where f and ξ satisfy the conditions in Theorem 3 required for the existence and uniqueness of solutions. Take $\mathrm{tr}(\xi^T S\xi(t,\mathbf{x})) \leq \overline{s}\eta^2$ for $\eta > 0$.

Definition 5 (Virtual System for (15))**.** As in Definition 4, represent the infinitesimal differential length $\delta\mathbf{z}$ as a path integral and reparameterize using measure parameter $\mu \in [0,1]$ such that:

$$\begin{aligned} &\mathbf{z}(\mu=0,t) = \mathbf{x}(t),\ \mathbf{z}(\mu=1,t) = \mathbf{y}(t) \\ &\xi_{\mu=0}(t,\mathbf{z}) = \xi(t,\mathbf{x}),\ \xi_{\mu=1}(t,\mathbf{z}) = 0 \\ &N_{\mu=0}(t,\mathbf{z}) = N(t),\ N_{\mu=1}(t,\mathbf{z}) = 0 \end{aligned} \tag{49}$$

Using this, we can rewrite (15) as the virtual system:

$$d\mathbf{z}(\mu,t) = f(t,\mathbf{z}(\mu,t))dt + \xi_\mu(t,\mathbf{z}(\mu,t))dN_\mu(t) \tag{50}$$

This further enables us to describe the virtual dynamics as follows:

$$d\delta\mathbf{z} = F\delta\mathbf{z}dt + \delta\xi_\mu dN_\mu \tag{51}$$

where $\delta\xi_\mu$ is defined in the same way as in (39).

A fundamental difference for stochastic contraction with respect to shot noise arises in contrast to the white noise case because of the large, discontinuously-occurring nature of the noise process. Thus, we provide guarantees which are not constant: within a fixed interval of time, the difference between the two trajectories will remain bounded by a time-varying function. The following useful fact allows us to do so.

Remark 6. For any fixed $t > 0$, $\mathbb{P}(N(t) = \infty) = 0$. That is, it is with probability zero that the number of jumps that have occurred by time t is infinite. We make the further assumption that the sum of all the sizes of the jumps that occurred until time t is finite.

We further introduce the Itô formula with jumps, which we will use in our proof of contraction. The version of the formula stated is also used in [15], [48], and a comprehensive treatment of its derivation can be found in [40].

Lemma 4. For functions $F \in \mathcal{C}^{(1,2)}$,

$$\begin{aligned} F(t,\mathbf{x}(t)) &= F(0,\mathbf{x}_0) + \int_0^t \partial_t F(s,\mathbf{x}(s))ds \\ &+ \sum_{i=1}^n \int_0^t \partial_{x_i} F(s,\mathbf{x}(s))d\mathbf{x}_i(s) \\ &+ \frac{1}{2}\sum_{i,j=1}^n \int_0^t \partial_{x_i}\partial_{x_j} F(s,\mathbf{x}(s))d[\mathbf{x}_i,\mathbf{x}_j]^c(s) \\ &+ \int_0^t \left(F(s,\mathbf{x}(s)) - F(s,\mathbf{x}(s-))\right)dN(s) \end{aligned} \tag{52}$$

where $\mathbf{x} \in \mathbb{R}^n$.

Remark 7 (Quadratic Variation)**.** Consider two generic scalar SDEs of the form (13):

$$dx_1 = f(t,x_1)dt + \sigma(t,x_1)dW(t) + \xi(t,x_1)dN(t) \tag{53a}$$

$$dx_2 = f(t,x_2)dt + \sigma(t,x_2)dW(t) + \xi(t,x_2)dN(t) \tag{53b}$$

The *quadratic variation* term $d\langle x_1,x_2\rangle(t)$ is computed to be $\sigma(t,x_1)\sigma(t,x_2)dt + \xi(t,x_1)\xi(t,x_2)dN(t)$ since $dW(t)\cdot dW(t) = dt$ and $dN(t)\cdot dN(t) = dN(t)$ while the dot products between all other terms vanish ($dt\cdot dt = dt\cdot dW(t) = dt\cdot dN(t) = dW(t)\cdot dN(t) = 0$). It is comprised of two parts: the continuous part $d[x_1,x_2]^c(t) = dt$ and the purely discontinuous part $d[x_1,x_2]^d(t) = dN(t)$. For further information about this notation, one may refer to [42], [48] and references therein.

Theorem 5 (Stochastic Contraction Theorem for Shot Noise)**.** For the SDE system described in (15), suppose the following two conditions are satisfied:

1) the unperturbed system satisfies (43).
2) there exists $\eta > 0$ such that $\|\xi(t,\mathbf{x})\|_F \leq \eta$ for all $\mathbf{x} \in \mathbb{R}^n$ and $t \geq 0$.

Further assume that the initial conditions adhere to some probability distribution $p(\mathbf{z}_0) = p(\mathbf{x}_0, \mathbf{y}_0)$. Then (15) is stochastically contracting if

$$\mathbb{E}_{\mathbf{z}_0}[V(t, \mathbf{z}, \delta\mathbf{z})] \leq V(0, \mathbf{z}_0, \delta\mathbf{z}_0)e^{-\beta t} + \kappa(t) \tag{54}$$

where V is the same Lyapunov-like function (42) as in the white noise analysis, and

$$\beta := 2\alpha - \lambda\eta\left(\frac{s'}{\underline{s}} + \frac{\overline{s}}{\underline{s}}\right) \tag{55a}$$

$$\kappa(t) := \lambda \int_0^t (\overline{s}\eta + c(s))\, e^{-\beta(s-t)} ds \tag{55b}$$

and λ is the intensity of the Poisson process $N(t)$ and $c(t)$ is the continuous and bounded function defined in (62).

As in Theorem 4, we can express the bound (54) in terms of the mean-squared difference between the two trajectories $\mathbf{x}$ and $\mathbf{y}$ using the construction in (42):

$$\mathbb{E}_{(\mathbf{x}_0,\mathbf{y}_0)}[\|\mathbf{y} - \mathbf{x}\|^2] \leq \frac{1}{\underline{s}}\|\mathbf{y}_0 - \mathbf{x}_0\|^2 e^{-\beta t} + \frac{\kappa(t)}{\underline{s}} \tag{56}$$

Remark 8. By Remark 6, the number of jumps by time t is finite with probability 1, so we are taking a maximum over a finite number of values. Hence, the supremum in the bound of Theorem 5 is justified.

Proof: By applying Itô's formula (52) to the Lyapunov-like function (42), we obtain

$$\begin{aligned}
&V(t, \mathbf{z}, \delta\mathbf{z}) = V(0, \mathbf{z}_0, \delta\mathbf{z}_0) + \int_0^t \partial_t V(s, \mathbf{z}, \delta\mathbf{z}) ds + \\
&\int_0^t \sum_{i=1}^n \left[\partial_{z_i} V(s, \mathbf{z}, \delta\mathbf{z}) f_i(s, \mathbf{z}) + \partial_{\delta z_i} V(s, \mathbf{z}, \delta\mathbf{z}) \left(\frac{\partial f}{\partial \mathbf{z}}\delta\mathbf{z}\right)_i\right] \\
&\quad + \int_0^t \sum_{i=1}^n [\partial_{z_i} V(s, \mathbf{z}, \delta\mathbf{z})\xi_{\mu,i}(s, \mathbf{z}) \\
&\qquad\qquad + \partial_{\delta z_i} V(s, \mathbf{z}, \delta\mathbf{z})\delta\xi_{\mu,i}]\, dN_\mu(s) \\
&+ \int_0^t (V(s, \mathbf{z}, \delta\mathbf{z}) - V(s, \mathbf{z}(s-), \delta\mathbf{z}(s-))) dN_\mu(s)
\end{aligned} \tag{57}$$

where the subscript of i in $\xi_{\mu,i}$ (and other similar notation) denotes the ith component of the respective vector. Note that these are dimension $1 \times d$, and the Poisson dN_μ is dimension $d \times 1$, so the overall product is a scalar, as expected.

The first three terms are derived directly from deterministic contraction of an unperturbed system $\dot{\mathbf{x}} = f(t, \mathbf{x})$. For the fourth and fifth terms, we will derive bounds with the expected value included so that we may directly apply Campbell's formula Lemma 1.

By Campbell's formula, the bound on the fourth sum term is written as

$$\begin{aligned}
&\mathbb{E}_{\mathbf{z}_0}\left[\int_0^t \sum_{i=1}^n (\partial_{z_i} V(s, \mathbf{z}, \delta\mathbf{z})\xi_{\mu,i}(\mathbf{z}, s)\right. \\
&\qquad\qquad \left. + \partial_{\delta z_i} V(s, \mathbf{z}, \delta\mathbf{z})\delta\xi_{\mu,i}(\mathbf{z}, s))\, dN_\mu(s)\right] \\
&= \lambda \sum_{i=1}^n \left(\int_0^t \mathbb{E}_{\mathbf{z}_0}[\partial_{z_i} V(s, \mathbf{z}, \delta\mathbf{z})\xi_{\mu,i}(\mathbf{z}, s)]\, ds\right. \\
&\qquad \left. + \int_0^t \mathbb{E}_{\mathbf{z}_0}[\partial_{\delta z_i} V(s, \mathbf{z}, \delta\mathbf{z})\delta\xi_{\mu,i}(\mathbf{z}, s)]\, ds\right)
\end{aligned} \tag{58}$$

where λ can be taken outside of the integral by Remark 2.

Applying submultiplicativity, we get:

$$\begin{aligned}
(58) \leq \lambda \sum_{i=1}^n &\left(\int_0^t \mathbb{E}_{\mathbf{z}_0}[\|\partial_{z_i} V(s, \mathbf{z}, \delta\mathbf{z})\|\, \|\xi(s, z)_i\|]\, ds\right. \\
&\left. + \int_0^t \mathbb{E}_{\mathbf{z}_0}[\|\partial_{\delta z_i} V(s, \mathbf{z}, \delta\mathbf{z})\|\, \|\delta\xi(s, z)_i\|]\, ds\right)
\end{aligned} \tag{59}$$

Note that

$$\|\partial_{z_i} V(s, \mathbf{z}, \delta\mathbf{z})\| \leq \sup_{t,\mathbf{z}} \|\partial_{\mathbf{z}_i} S(t, \mathbf{z})\|\, \|\delta\mathbf{z}^T\delta\mathbf{z}\| \leq \frac{s'}{\underline{s}} V(s, \mathbf{z}, \delta\mathbf{z})$$

$$\|\partial_{\delta z_i} V(s, \mathbf{z}, \delta\mathbf{z})\| \leq 2\sup_{t,\mathbf{z}} \|S(t, \mathbf{z})\|\, |\delta\mathbf{z}_i| = 2\overline{s}|\delta\mathbf{z}_i|$$

Hence, (59) is bounded above by:

$$\begin{aligned}
(59) &\leq \lambda\left(\frac{s'}{\underline{s}}\right)\int_0^t V(s, \mathbf{z}, \delta\mathbf{z}) \sum_{i=1}^n \|\xi(s, \mathbf{z})_i\|\, ds \\
&\qquad + 2\lambda\overline{s}\int_0^t \sum_{i=1}^n |\delta\mathbf{z}_i|\, \|\delta\xi(s, \mathbf{z})_i\|\, ds \\
&\leq \lambda\eta\left(\frac{s'}{\underline{s}}\right)\int_0^t V(s, \mathbf{z}, \delta\mathbf{z}) ds \\
&\qquad + 2\lambda\overline{s}\eta\int_0^t \|\delta\mathbf{z}\| \text{ by Cauchy-Schwarz} \\
&\leq \lambda\eta\left(\frac{s'}{\underline{s}}\right)\int_0^t V(s, \mathbf{z}, \delta\mathbf{z}) ds \\
&\qquad + \lambda\overline{s}\eta\left(t + \frac{1}{\underline{s}}\int_0^t V(s, \mathbf{z}, \delta\mathbf{z}) ds\right)
\end{aligned} \tag{60}$$

Again by Campbell's formula, the bound on the fifth term is written as:

$$\begin{aligned}
&\mathbb{E}_{\mathbf{z}_0}\left[\int_0^t (V(s, \mathbf{z}, \delta\mathbf{z}) - V(s-, \mathbf{z}, \delta\mathbf{z}))\, dN(s)\right] \\
&= \lambda\int_0^t \mathbb{E}_{\mathbf{z}_0}[V(s, \mathbf{z}, \delta\mathbf{z}) - V(s-, \mathbf{z}, \delta\mathbf{z})]\, ds
\end{aligned} \tag{61}$$

Assumption 2. There exists a continuous and bounded function c such that

$$\mathbb{E}_{\mathbf{z}_0}[V(t, \mathbf{z}(t), \delta\mathbf{z}(t)) - V(t-, \mathbf{z}(t-), \delta\mathbf{z}(t-))] \leq c(t) \tag{62}$$

Using (62), (61) is bounded by:

$$(61) \leq \lambda\int_0^t c(s) ds =: \lambda K(t)$$

In combination, the expected value of (57) is bounded above as:

$$\mathbb{E}_{\mathbf{z}_0}\left[V(t,\mathbf{z},\delta\mathbf{z})\right] \leq V(0,\mathbf{z}_0,\delta\mathbf{z}_0) - 2\alpha\int_0^t \mathbb{E}_{\mathbf{z}_0}[V(s,\mathbf{z},\delta\mathbf{z})]ds + \lambda\eta\left(\frac{s'}{\underline{s}}+\frac{\overline{s}}{\underline{s}}\right)\int_0^t \mathbb{E}_{\mathbf{z}_0}[V(s,\mathbf{z},\delta\mathbf{z})]ds + \lambda\overline{s}\eta t + \lambda K(t) \tag{63}$$

where α is the contraction rate of the unperturbed system.

Defining β as in (55a), (63) becomes:

$$\mathbb{E}_{\mathbf{z}_0}\left[V(t,\mathbf{z},\delta\mathbf{z})\right] \leq V(0,\mathbf{z}_0,\delta\mathbf{z}_0) - \beta\int_0^t \mathbb{E}_{\mathbf{z}_0}\left[V(s,\mathbf{z},\delta\mathbf{z})\right]ds + \lambda\eta\overline{s}t + \lambda K(t) \tag{64}$$

We can rewrite integral inequality (64) as a differential inequality:

$$d\mathbb{E}_{\mathbf{z}_0}\left[V(t,\mathbf{z},\delta\mathbf{z})\right] \leq -\beta\mathbb{E}_{\mathbf{z}_0}\left[V(t,\mathbf{z},\delta\mathbf{z})\right] + (\lambda\overline{s}\eta + \lambda c(t))$$

Solving the differential equation and using the Comparison lemma to turn equality to an inequality, we obtain:

$$\mathbb{E}_{\mathbf{z}_0}\left[V(t,\mathbf{z},\delta\mathbf{z})\right] \leq V(0,\mathbf{z}_0,\delta\mathbf{z}_0)e^{-\beta t} + \lambda\int_0^t (\overline{s}\eta + c(s))e^{-\beta(s-t)}ds$$

and defining $\kappa(t)$ as in (55b) gives us the form (54).

□

Remark 9. As the intensity λ of the Poisson noise process increases (i.e. the jumps arrive more frequently), then note that the steady-state error $\kappa(t)$ increases as well.

One way to minimize this effect is to increase the deterministic contraction rate α accordingly, because within each time interval (T_j, T_{j+1}) there are no additional jumps (it can be thought of restarting the process with initial condition at the value of $\mathbf{z}$ immediately after the jump affects the system). Since there is no perturbation in the time interval (T_j, T_{j+1}), the trajectory $\mathbf{y}$ behaves as a deterministic system with initial condition $\mathbf{y}(T_j)$. In this interval of time, it becomes equivalent to comparing two deterministic trajectories against each other. Increasing α also increases the stochastic contraction rate β and the trajectories converge as quickly as possible towards each other in between jumps. This will decrease the supremum term in (55b).

This contraction result is presented for stochastic systems that are perturbed by only Poisson noise. However, it further directs us into a treatment of stochastic systems that are perturbed by a broader class of noise processes, such as white and Poisson combined. This is due to the following theorem, stated specifically for scalar systems for the sake of simplicity. For further reading, one is referred to [41], [49].

Definition 6. Let L be a Lévy process and $\varphi:\mathbb{R}\to\mathbb{C}$ be its *characteristic function*, defined as $\varphi(\theta) = \mathbb{E}\left[e^{i\theta L(t)}\right]$ for any time t. Then there exists a unique continuous function $\rho:\mathbb{R}\to\mathbb{C}$ such that $e^{\rho(\theta)} = \varphi(\theta)$. This function $\rho(\theta) = \log(\varphi(\theta))$ is referred to as the *characteristic exponent* of $L(t)$.

Theorem 6 (Lévy-Khintchine Decomposition Formula)**.** Let L be a Lévy process with characteristic exponent Ψ. Then there exist (unique) $a\in\mathbb{R}, \sigma\geq 0$, and a measure ν which satisfies

$$\int_{\mathbb{R}} x^2\nu(dx) < \infty \tag{65}$$

such that

$$\Psi(\theta) = ia\theta - \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}}\left(e^{i\theta x}-1\right)\nu(dx) - \int_{\mathbb{R}} i\theta x\mathbb{1}_{[-1,1]}(x)\nu(dx) \tag{66}$$

Conversely, given any triplet (a,σ,ν), there exists a Lévy process L with characteristic exponent given by (66).

Note that a, which is often called the *center* of L, captures the deterministic drift component, *Gaussian coefficient* σ captures the variance of the Brownian motion component, and the *Lévy measure* ν captures the size and intensity of the "large" jumps of L. In essence, Theorem 6 tells us that any Lévy process (general noise) can be split into three simpler Lévy processes: a Brownian motion (white noise), a compound Poisson process of "large" jumps (Poisson noise), and a compound Poisson process of jumps small enough to be compensated by a deterministic drift to turn it into a martingale.

Motivated by Theorem 6, a contraction result can be easily obtained for the combined system (13). Under combined Lipschitzness and bounded-growth assumptions of Theorem 3, we can say that there exists a unique solution for (13) for any given initial condition $\mathbf{x}_0$.

Assume the same metric assumptions as in Assumption 1. Compare two trajectories: $\mathbf{x}(t)$ as the solution to (13) and $\mathbf{y}(t)$ as the solution to a system perturbed only by white noise:

$$d\mathbf{x} = f(\mathbf{x},t)dt + \sigma_1(\mathbf{x},t)dW_1(t) + \xi(\mathbf{x},t)dN(t) \tag{67a}$$
$$d\mathbf{y} = f(\mathbf{x},t)dt + \sigma_2(\mathbf{x},t)dW_2(t) \tag{67b}$$

Definition 7 (Virtual System for (13))**.** The infinitesimal differential length $\delta\mathbf{z}$ can be represented as in a path integral and reparameterized using measure parameter $\mu\in[0,1]$ such that:

$$\mathbf{z}(\mu=0,t) = \mathbf{x}(t),\ \mathbf{z}(\mu=1,t) = \mathbf{y}(t) \tag{68}$$
$$\sigma_{\mu=0}(t,\mathbf{z}) = \sigma_1(t,\mathbf{x}),\ \sigma_{\mu=1}(t,\mathbf{z}) = \sigma_2(t,\mathbf{y})$$
$$\xi_{\mu=0}(t,\mathbf{z}) = \xi(t,\mathbf{x}),\ \xi_{\mu=1}(t,\mathbf{z}) = 0$$
$$W_{\mu=0}(t,\mathbf{z}) = W_1(t),\ W_{\mu=1}(t,\mathbf{z}) = W_2(t)$$
$$N_{\mu=0}(t,\mathbf{z}) = N(t),\ N_{\mu=1}(t,\mathbf{z}) = 0$$

Using this, we can rewrite (13) as the virtual system:

$$d\mathbf{z}(\mu,t) = f(t,\mathbf{z}(\mu,t))dt + \sigma_\mu(t,\mathbf{z}(\mu,t))dW_\mu(t) + \xi_\mu(t,\mathbf{z}(\mu,t))dN_\mu(t) \tag{69}$$

This further enables us to describe the virtual dynamics as follows:

$$d\delta\mathbf{z} = F\delta\mathbf{z}dt + \delta\sigma_\mu dW_\mu + \delta\xi_\mu dN_\mu \tag{70}$$

Corollary 1 (Stochastic Contraction Theorem for General Noise)**.** For the SDE system described in (13), suppose the following two conditions are satisfied:

1) the unperturbed system satisfies (43).
2) there exists $\eta, \eta_1, \eta_2 > 0$ such that $\|\sigma_1(t,\mathbf{x})\|_F \leq \eta_1$, $\|\sigma_2(t,\mathbf{x})\|_F \leq \eta_2$, and $\|\xi(t,\mathbf{x})\|_F \leq \eta$ for all $\mathbf{x} \in \mathbb{R}^n$ and $t \geq 0$.

Further assume that the initial conditions adhere to some probability distribution $p(\mathbf{z}_0) = p(\mathbf{x}_0, \mathbf{y}_0)$. Then (13) is stochastically contracting if

$$\mathbb{E}_{\mathbf{z}_0}[V(t,\mathbf{z},\delta\mathbf{z})] \leq V(0,\mathbf{z}_0,\delta\mathbf{z}_0)e^{-\beta t} + \kappa(t) \tag{71}$$

where

$$\beta := 2\alpha - \lambda\eta\left(\frac{s'}{\underline{s}} + \frac{\overline{s}}{\underline{s}}\right) - \frac{1}{2\underline{s}}(\eta_1^2+\eta_2^2)\left(\frac{s''}{2} + s'\right) \tag{72a}$$

$$\kappa(t) := \frac{1}{2}(\overline{s}+s')(\eta_1^2+\eta_2^2) + \lambda\int_0^t \left(\overline{s}\eta + c(s)\right)e^{-\beta(s-t)}ds \tag{72b}$$

and λ is the intensity of the Poisson process $N(t)$ and $c(t)$ is defined in (62). We can express the bound (54) in terms of the mean-squared error between the trajectories $\mathbf{x}$ and $\mathbf{y}$ using (42) to obtain the form in (56), with constants β defined in (72a) and $\kappa(t)$ defined in (72b).

Essentially, the values β and $\kappa(t)$ for the combined SDE are a direct summation of the corresponding values for (14) and (15) separately. This is due to the similar differential inequality form that $\mathbb{E}_{\mathbf{z}_0}[V(t,\mathbf{z},\delta\mathbf{z})]$ has in terms of $V(0,\mathbf{z}_0,\delta\mathbf{z}_0)$ and $\int_0^t \mathbb{E}_{\mathbf{z}_0}[V(s,\mathbf{z},\delta\mathbf{z})]ds$ prior to multiplying by the integrating factor and applying the Comparison lemma.

V. Numerical Examples

A. Dubins' Car

Consider the dynamics of the 5-dimensional Dubins' car model from [29]:

$$\dot{\xi} = \begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{\theta}\\ \dot{v}\\ \dot{\omega}\end{bmatrix} = \begin{bmatrix} v\cos(\theta) + w_1 + +n_1\\ v\sin(\theta) + w_2 + n_2\\ \omega + w_3 + n_3\\ a\\ \alpha\end{bmatrix} \tag{73}$$

where $w(t) = (w_1, w_2, w_3)^T$ is a white noise vector, $n(t) = (n_1, n_2, n_3)^T$ is a Poisson noise vector, and the control inputs are given by translational acceleration a and angular acceleration α.

White noise vectors are generated from a Gaussian distribution with mean $\mathbf{0}$ and covariance $3I$, while shot noise vectors are generated from a Poisson process with rate $\lambda = 5$ and heights:

$$n(t) = \begin{cases} \begin{bmatrix}\text{Unif}[-400,600]\\ \text{Unif}[-200,200]\\ \text{Unif}[200,400]\end{bmatrix} & \text{with prob } 0.5\\ \begin{bmatrix}\text{Unif}[-600,400]\\ \text{Unif}[-200,200]\\ \text{Unif}[-400,-200]\end{bmatrix} & \text{with prob } 0.5\end{cases}$$

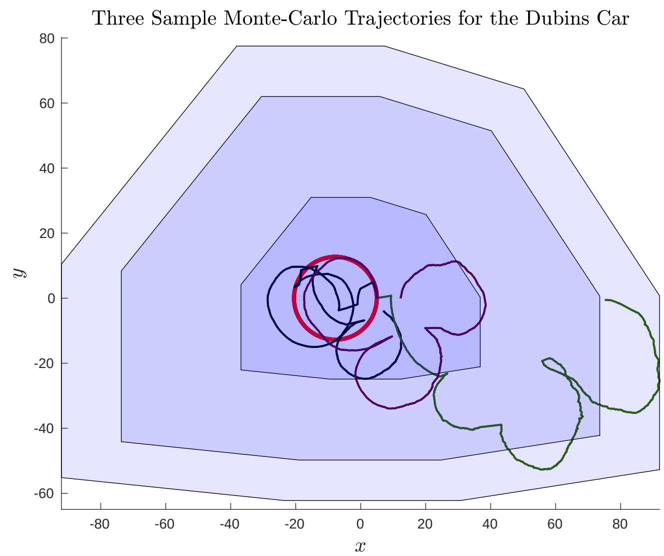

Fig. 2: Three sample open-loop white and shot noise perturbed trajectories corresponding to each of three disjoint partitions of the maximum norm ball (lightest shade of blue). The black trajectory lies fully within the smallest set, the dark red lies fully in the medium-sized set, and the green trajectory remains in the largest bounded set. The brighter red circle represents open-loop trajectory of the unperturbed system.

We simulate $M = 50$ Monte-Carlo trajectories of this system under the influence of different noise processes generated from the distributions described. Each trajectory is simulated for $T = 20$ seconds, starting from the initial conditions $\xi_0 = \left(5, 0, \frac{\pi}{2}, 0, 0\right)^T$. For the first one second, we add in constant linear acceleration $a = 10$ and constant angular acceleration $\alpha = \frac{\pi}{4}$, then set both inputs to zero for every time afterwards. When the system is not perturbed by noise, this traces out a circular trajectory of approximate radius of 12.7602 and approximate center $C = (-7.7569, -0.0199)$.

We show that under the influence of a particular class of white and shot noise (in this case, zero-mean white noise with covariance $3I$ and shot noise generated according to the described rule), the system trajectory is able to remain reasonably bounded within a norm ball of the trajectory of the unperturbed system. We illustrate this in Fig. 2. The largest bounded set is computed using the following method. For each generated trajectory, we compute the smallest rectangle that fully-encloses the trajectory by computing the maximum and minimum x and y states. Then we simply take the convex hull of all rectangular enclosures. The two smaller partitions are created by scaling this convex hull by $1.2 \cdot \frac{1}{3}$ and $1.2 \cdot \frac{2}{3}$.

Finally, we compute the fraction of trajectories that has its maximum deviation in each region. For this particular experiment, we get $\left(\frac{7}{50}, \frac{36}{50}, \frac{13}{50}\right)$ for the proportions that lie in the smallest, medium, and largest bounded sets, respectively.

B. 1D Compound Poisson Control

Now that we have established basic foundations for shot-noise systems, we will present the controller design for a 1D simple motivational example . In the previous Dubins' Car simulation, we showed that there are cases where the open-loop system by itself is able to handle classes of white and shot noise that are lenient and well-behaved in nature. Here, we show that the addition of the control component will allow us to handle more aggressive noise processes (e.g., varying intensities λ) because the trajectories can now be steered as close to the nominal behavior as much as possible.

Consider the scalar linear system

$$\dot{x}(t) = ax(t) + u(t) + w(t) + n(t) \tag{74}$$

with $a = 2$ so that it is unstable in open-loop. This system tries to track the nominal reference trajectory (in dashed blue) while being perturbed by a white noise process $w(t)$ and a shot noise process $n(t)$.

We design a baseline controller to track a sinusoid trajectory $x_r(t) = \sin(t)$:

$$\begin{aligned} u_r(t) &= \dot{x}_r(t) - ax_r(t) \\ u(t) &= u_r(t) - k(x(t) - x_r(t)) \end{aligned} \tag{75}$$

and gain k is chosen $a + 0.5$ so that $x(t)$ exponentially converges to the sinusoid with a rate of -0.5.

The white noise process is generated zero-mean with variance 3. The shot noise process is Poisson with interarrival times distributed Exponential with mean λ, and heights are distributed according to the following law

$$\xi_k \sim \begin{cases} \text{Unif}[100, 400] & \text{with prob } 0.5 \\ \text{Unif}[-400, -100] & \text{with prob } 0.5 \end{cases}$$

We modify the control law so that the gain is increased according to a Linear Quadratic Regulator (LQR) law with state cost and control cost equal to 1 so as to encourage the system to converge quickly back to the reference trajectory if it has been suddenly perturbed by a jump. We impose two criteria for when this change in the control law should be applied:

1) the size of the jump must be larger than a predetermined height. For our purposes, we set this threshold to 1.
2) the jump perturbs the system further away from the reference trajectory. There is no need to force an increase in the gain if the jump is beneficial.

Once the system trajectory converges within ε distance of the reference, we revert the gain back to $a + 0.5$. Furthermore, there may be a slight time delay in when we observe the jump noise, so the control law is not applied immediately after the jump occurs. We choose $\varepsilon = 0.01$ and a time delay of 0.5 seconds. Three simulated trajectories, with $\lambda = 3, 7, 11$, under the setting described is shown in Fig. 3. Performing Monte-Carlo simulations for 100 trials and perturbing the system for 100 different combinations of white and shot noise yields the same behavior of converging to the sinusoidal reference trajectory as much as possible.

Remark 10. Note that in contrast to the criteria derived in Section IV-B, which essentially bounded the maximum deviation possible in open-loop, contraction for closed-loop would refer to the system's ability to globally exponentially converge to within some bounded distance away from the nominal trajectory before it is hit by another jump. Hence, the bound will be dependent upon the time that elapses in between jumps to ensure sufficient convergence of the system trajectory to the reference one.

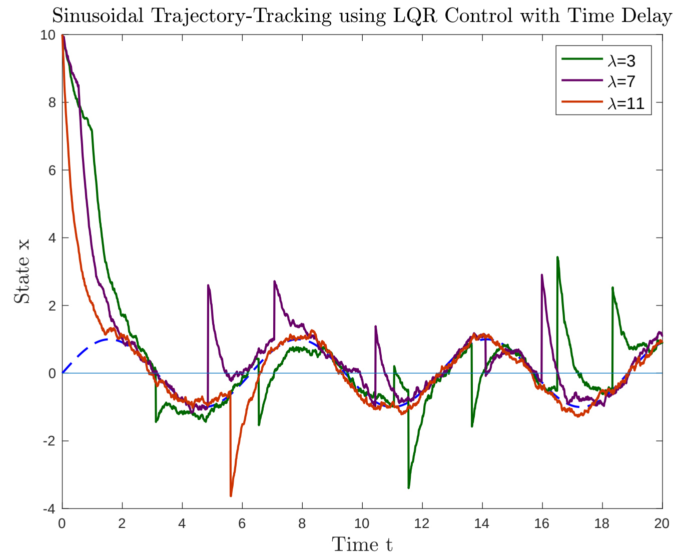

Fig. 3: Three system responses of an open-loop unstable linear system when influenced by white and shot noise, and controlled by a simple PD control scheme. The parameter λ of the shot noise process is varied while the mean and variance of the white noise is kept constant.

VI. Conclusion

In this paper, we extended nonlinear systems analysis to take into account noise processes more general than the standard white noise. Namely, we added shot noise and considered systems of the form (15), and theoretically derived steady-state error bounds using incremental stability analysis. An additional tool that was developed to facilitate the derivation was the existence and uniqueness theorem for (15). Finally, we illustrated by simulation of a 5D Dubins car that system trajectories with shot noise are still reasonably bounded according to the error bound we derived through contraction theory. We additionally illustrated that the usual LQR control law suffices for reference-tracking in a 1D linear system perturbed by shot noise. Designing a controller for nonlinear systems is arguably an important one to address in a wide diversity of engineering applications [50]. A methodological control/filter design process for higher dimensional nonlinear systems perturbed by shot noise will be examined in future work, along with possible demonstrations on real-life dynamical systems.

Acknowledgment

The authors would like to thank John C. Doyle for the insights he provided as motivation for this work.

APPENDIX I
PROOF OF LEMMA 2 AND ADDITIONAL PROPERTIES OF THE POISSON INTEGRAL

In this section, we prove Lemma 2. We provide a more complete version than what was given in [41]. Consider the simple function $g(t,y) = c_1\mathbb{1}(y\in E_1)+\cdots+c_n\mathbb{1}(y\in E_n)$. We can decompose the integral into a sum

$$
\begin{aligned}
I_g &= \int_{[0,T]\times E} \sum_{k=1}^{n} c_k \mathbb{1}(y\in E_k) N(dt,dy) \\
&= \sum_{k=1}^{n} \int_{[0,T]\times E_k} c_k N(dt,dy) \qquad (76)
\end{aligned}
$$

Then expanding out the left side of the above identity:

$$
\begin{aligned}
\mathbb{E}[I_g^2] &= \mathbb{E}\left[\left(\int_{[0,T]\times E} g(x,t)N(dt,dy)\right)^2\right] \\
&= \mathbb{E}\left[\left(\sum_{k=1}^{n} c_k \int_{[0,T]\times E_k} N(dt,dy)\right)^2\right] \\
&= \mathbb{E}\left[\sum_{k=1}^{n} c_k^2 N_k^2(t) + \sum_{k\neq j} c_k c_j N_k(t) N_j(t)\right] \\
&= \sum_{k=1}^{n} c_k^2 \mathbb{E}[N_k^2(t)] + \sum_{k\neq j} c_k c_j \mathbb{E}[N_k(t)N_j(t)] \qquad (77)
\end{aligned}
$$

where $N_k(t)$ is the standard Poisson process defined on Borel set E_k, with intensity $\lambda_k := \lambda\cdot\nu(E_k)/\nu(E)$. First note that since $N_k(t)$ is Poisson distributed, $\mathbb{E}[N_k^2(t)] = \mathrm{Var}(N_k(t)) + \mathbb{E}[N_k(t)]^2 = \lambda_k t + \lambda^2 t^2$. Next, since $E_j \cap E_k = \varnothing$, we have that $N_j(t)$ and $N_k(t)$ are independent random variables. Thus

$$
\begin{aligned}
0 = \mathrm{Cov}(N_j(t),N_k(t)) &= \mathbb{E}[N_j(t)N_k(t)] - \mathbb{E}[N_j(t)]\mathbb{E}[N_k(t)] \\
&= \mathbb{E}[N_j(t)N_k(t)] - \lambda_j\lambda_k t^2
\end{aligned}
$$

Substituting these values into (77):

$$
\begin{aligned}
&\sum_{k=1}^{n} c_k^2(\lambda_k t + \lambda_k^2 t^2) + \sum_{k\neq j} c_k c_j \lambda_k \lambda_j t^2 \\
&= \sum_{k=1}^{n} c_k^2 \lambda_k t + \sum_{k=1}^{n} c_k^2 \lambda_k^2 t^2 + \sum_{k\neq j} c_k c_j \lambda_k\lambda_j t^2 \\
&= \sum_{k=1}^{n} \int_{[0,T]\times E_k} c_k^2 dt\nu(dy) \\
&\qquad + \left(\sum_{k=1}^{n} \int_{[0,T]\times E_k} c_k dt\nu(dy)\right)^2 \\
&= \int_{[0,T]\times E} \mathbb{E}[g(t,y)^2] dt\nu(dy) \\
&\qquad + \left(\int_{[0,T]\times E} \mathbb{E}[g(t,y)] dt\nu(dy)\right)^2
\end{aligned}
$$

Use the Monotone Convergence Theorem to apply the same argument as above to general functions g by approximating g with step-functions g_n. This concludes our proof.

For further intuition of the computation of Poisson integrals, consider the following example.

Example 3. If I_g is a.s. absolutely convergent, then

$$
\mathbb{E}\left[e^{i\beta I_g}\right] = \exp\left(-\int_{[0,T]\times E}\left(1-e^{i\beta g(t,y)}\right)dt\nu(dy)\right) \qquad (78)
$$

where $\beta\in\mathbb{R}$. This formula is often called the *exponential formula*.

Again, we will consider the simple function $g(t,y) := c_1\mathbb{1}(y\in E_1)+\cdots+c_n\mathbb{1}(y\in E_n)$ where $n\in\mathbb{N}$, $c_k\in(0,T)$, and $E_1,\cdots,E_n$ are disjoint Borel sets.

Substituting the decomposition (76) into the left side of (78),

$$
\begin{aligned}
\mathbb{E}\left[e^{i\beta I_g}\right] &= \mathbb{E}\left[e^{i\beta \sum_{k=1}^{n} \int_{[0,T]\times E_k} c_k N(dt,dy)}\right] \qquad (79) \\
&= \mathbb{E}\left[\prod_{k=1}^{n} e^{i\beta \int_{[0,T]\times E_k} c_k N(dt,dy)}\right] = \prod_{k=1}^{n}\mathbb{E}\left[e^{i\beta c_k N_k(t)}\right]
\end{aligned}
$$

where the last equality follows from disjointness of the Borel sets E_k, hence independence between the terms in the product. Now we can use the characteristic function of a (scaled) Poisson

$$
\mathbb{E}\left[e^{i\beta c_j N_j(t)}\right] = \sum_{k=0}^{\infty} e^{i\beta c_j k} e^{-\lambda_j t}\frac{(\lambda_j t)^k}{k!} = e^{\lambda_j t(e^{i\beta c_j}-1)}
$$

to further simplify (79)

$$
\begin{aligned}
\prod_{k=1}^{n} e^{\lambda_k t(e^{i\beta c_k}-1)} &= \exp\left(\sum_{k=1}^{n}\lambda_k t\left(e^{i\beta c_k}-1\right)\right) \\
&= \exp\left(-\sum_{k=1}^{n}\int_{[0,T]\times E_k}\left(1-e^{i\beta c_k}\right)N(dt,dy)\right) \qquad (80)
\end{aligned}
$$

Now to make the extension from simple functions to general functions, we can construct a sequence of simple functions g_n with limit $g_n \uparrow g$ as $n\to\infty$. Again, we apply the Monotone Convergence Theorem to get that the integrals with respect to the Poisson measure N converge as well. Using this argument, we can substitute in the general function g in place of the c_k in (80). Further using $E=\cup_k E_k$, with E_k disjoint sets, yields our desired form:

$$
\begin{aligned}
\mathbb{E}\left[e^{i\beta I_g}\right] &= \exp\left(-\sum_{k=1}^{n}\int_{[0,T]\times E_k}\left(1-e^{i\beta g(t,y)}N(dt,dy)\right)\right) \\
&= \exp\left(-\int_{[0,T]\times E}\left(1-e^{i\beta g(t,y)}N(dt,dy)\right)\right)
\end{aligned}
$$

APPENDIX II
BOUNDS ON THE STOCHASTIC TERMS

Lemma 5. Consider the function

$$
V(t,\mathbf{z},\delta\mathbf{z}) = \int_0^1 \left(\frac{\partial \mathbf{z}}{\partial\mu}\right)^T S(t,\mathbf{z}(\mu,t))\left(\frac{\partial\mathbf{z}}{\partial\mu}\right)d\mu
$$

with respect to metric S, which satisfies Assumption 1, and measure $\mu\in[0,1]$ defined as in Definition 4. Then the following hold true, where constants are as in (35) and $\varepsilon > 0$ arbitrary:

$$\sum_{i,j=1}^{n} \frac{\partial^2 V}{\partial \delta \mathbf{z}_i \partial \delta \mathbf{z}_j}(t, \mathbf{z}, \delta \mathbf{z}) d\langle \delta \mathbf{z}_i, \delta \mathbf{z}_j \rangle \leq \overline{s}(\eta_1^2 + \eta_2^2) \tag{81a}$$

$$\sum_{i,j=1}^{n} \frac{\partial^2 V}{\partial \mathbf{z}_j \partial \delta \mathbf{z}_i}(t, \mathbf{z}, \delta \mathbf{z}) d\langle \mathbf{z}_i, \delta \mathbf{z}_j \rangle \leq \frac{1}{2} s'(\eta_1^2 + \eta_2^2) \left(\int_0^1 \left\| \frac{\partial \mathbf{z}}{\partial \mu} \right\|^2 d\mu + 1 \right) \tag{81b}$$

$$\sum_{i,j=1}^{n} \frac{\partial^2 V}{\partial \mathbf{z}_i \partial \mathbf{z}_j}(t, \mathbf{z}, \delta \mathbf{z}) d\langle \mathbf{z}_i, \mathbf{z}_j \rangle \leq \frac{1}{2} s''(\eta_1^2 + \eta_2^2) \int_0^1 \left\| \frac{\partial \mathbf{z}}{\partial \mu} \right\|^2 d\mu \tag{81c}$$

Proof:

First, we compute the quadratic variation terms. The ith component of the virtual system (38) and the virtual dynamics (39) are

$$d\mathbf{z}_i = f(t, \mathbf{z}_i)dt + \sum_{k=1}^{d} \sigma_{\mu,ik} dW_{\mu,k}(t)$$
$$d\delta \mathbf{z}_i = (F \delta \mathbf{z})_i dt + \sum_{k=1}^{d} \delta \sigma_{\mu,ik} dW_{\mu,k}(t)$$

Then we obtain

$$d\langle \delta \mathbf{z}_i, \delta \mathbf{z}_j \rangle = \sum_{k=1}^{d} \delta \sigma_{\mu,ik} \delta \sigma_{\mu,jk}$$
$$d\langle \mathbf{z}_i, \delta \mathbf{z}_j \rangle = \sum_{k=1}^{d} \sigma_{\mu,ik} \delta \sigma_{\mu,jk}$$
$$d\langle \mathbf{z}_i, \mathbf{z}_j \rangle = \sum_{k=1}^{d} \sigma_{\mu,ik} \sigma_{\mu,jk}$$

a. *Proof of (81a)*: From matrix multiplication, and because $S(t, \mathbf{z})$ is independent of $\delta \mathbf{z}$:

$$\delta \mathbf{z}^T S(t, \mathbf{z}) \delta \mathbf{z} = \sum_{k=1}^{n} \sum_{l=1}^{n} \delta \mathbf{z}_k \delta \mathbf{z}_l S_{kl}(t, \mathbf{z})$$
$$\frac{\partial^2 V}{\partial \delta \mathbf{z}_i \partial \delta \mathbf{z}_j}(t, \mathbf{z}, \delta \mathbf{z}) = 2 \int_0^1 S_{ij}(t, \mathbf{z}(\mu, t)) d\mu$$

Substituting into the left side of (81a), we get:

$$\begin{aligned} &\sum_{i,j=1}^{n} \frac{\partial^2 V}{\partial \delta \mathbf{z}_i \partial \delta \mathbf{z}_j}(t, \mathbf{z}, \delta \mathbf{z}) d\langle \delta \mathbf{z}_i, \delta \mathbf{z}_j \rangle \\ &= 2 \int_0^1 \sum_{i,j=1}^{n} S_{ij}(t, \mathbf{z}) \sum_{k=1}^{d} \left(\frac{\partial \sigma_\mu}{\partial \mu} \right)_{ik} \left(\frac{\partial \sigma_\mu}{\partial \mu} \right)_{jk} d\mu \\ &\leq 2\overline{s} \sum_{i,j=1}^{n} \sum_{k=1}^{d} \int_0^1 \left(\frac{\partial \sigma_\mu}{\partial \mu} \right)_{ik} \left(\frac{\partial \sigma_\mu}{\partial \mu} \right)_{jk} d\mu \end{aligned} \tag{82}$$

We have the following identity for any square matrix A and any pair $i, j = 1, \cdots, n$ such that $i \neq j$

$$\frac{1}{2}\text{tr}(A^T A) = \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{d} a_{ik}^2 \geq \sum_{k=1}^{d} a_{ik} a_{jk}$$

which is easily seen by completing the squares. In the 2D case ($n = 2, d = 2$), we would get:

$$A^T A = \begin{bmatrix} a_{11}^2 + a_{12}^2 & * \\ * & a_{21}^2 + a_{22}^2 \end{bmatrix}$$
$$\implies \text{tr}(A^T A) = \sum_{i,k=1}^{2} a_{ij}^2 \geq 2(a_{11}a_{21} + a_{21}a_{22})$$
$$\implies \frac{1}{2}\text{tr}(A^T A) \geq a_{11}a_{21} + a_{21}a_{22}$$

and when $i = j$, the same bound holds because the sum is simply the ith partial sum of the trace.

This allows us to bound (82) by splitting up into terms with $i = j$ and terms with $i \neq j$ and bounding both parts by a trace:

$$\begin{aligned} 2\overline{s} \sum_{i,j=1}^{n} \sum_{k=1}^{d} \int_0^1 \left(\frac{\partial \sigma_\mu}{\partial \mu} \right)_{ik} \left(\frac{\partial \sigma_\mu}{\partial \mu} \right)_{jk} d\mu &\leq 2\overline{s}\text{tr} \left(\sigma_\mu^T \sigma_\mu(t, \mathbf{x}) \right) \\ &\leq \overline{s} \left[\text{tr} \left(\sigma_1^T \sigma_1(t, \mathbf{x}) \right) + \text{tr} \left(\sigma_2^T \sigma_2(t, \mathbf{y}) \right) \right] \\ &\leq \overline{s}(\eta_1^2 + \eta_2^2) \end{aligned}$$

b. *Proof of (81b)*: First, we can compute the matrix derivative as follows.

$$\frac{\partial^2 V}{\partial \mathbf{z}_i \partial \delta \mathbf{z}_j}(t, \mathbf{z}, \delta \mathbf{z}) = 2 \int_0^1 \frac{\partial}{\partial \mathbf{z}_i} S_{j,i}(t, \mathbf{z})(\delta \mathbf{z})_i d\mu \quad \forall 1 \leq i, j \leq n$$

This gives:

$$\begin{aligned} &2 \int_0^1 \sum_{i,j=1}^{n} \frac{\partial}{\partial \mathbf{z}_i} S_{j,i}(t, \mathbf{z})(\delta \mathbf{z})_i \sum_{k=1}^{d} \sigma_{\mu,ik} \left(\frac{\partial \sigma_\mu}{\partial \mu} \right)_{jk} d\mu \\ &\leq 2s' \int_0^1 \sum_{i=1}^{n} \left(\frac{\partial \mathbf{z}}{\partial \mu} \right)_i \sum_{k=1}^{d} \sigma_{\mu,ik} \delta \sigma_{\mu,jk} d\mu \\ &\leq s'(\eta_1^2 + \eta_2^2) \int_0^1 \sum_{i=1}^{n} \left(\frac{\partial \mathbf{z}}{\partial \mu} \right)_i d\mu \\ &\leq s'(\eta_1^2 + \eta_2^2) \int_0^1 \left\| \frac{\partial \mathbf{z}}{\partial \mu} \right\| d\mu \end{aligned} \tag{83}$$

where the second-to-last inequality follows from the same trace bound used in (81a).

Note that for any $a, b > 0$, $2ab \leq a^2 + b^2$ and so by Cauchy-Schwarz

$$\int_0^1 \left\| \frac{\partial \mathbf{z}}{\partial \mu} \right\| d\mu \leq \frac{1}{2} \left(\int_0^1 \left\| \frac{\partial \mathbf{z}}{\partial \mu} \right\|^2 d\mu + 1 \right) \tag{84}$$

Combining (83) together with (84) yields our final bound:

$$\frac{1}{2} s'(\eta_1^2 + \eta_2^2) \left(\int_0^1 \left\| \frac{\partial \mathbf{z}}{\partial \mu} \right\|^2 d\mu + 1 \right)$$

c. *Proof of (81c)*: Again, start by computing the matrix derivative

$$\frac{\partial^2 V}{\partial \mathbf{z}_i \partial \mathbf{z}_j}(t, \mathbf{z}, \delta \mathbf{z}) = \int_0^1 \sum_{k,l=1}^{n} \frac{\partial^2}{\partial \mathbf{z}_i \partial \mathbf{z}_j} S_{kl}(t, \mathbf{z}) \delta \mathbf{z}_k \delta \mathbf{z}_l d\mu$$

Now we can bound using essentially the same technique as in (81b).

$$\begin{aligned}
&\sum_{i,j=1}^{n} \frac{\partial^2 V}{\partial \mathbf{z}_i \partial \mathbf{z}_j}(t, \mathbf{z}, \delta \mathbf{z}) d\langle \mathbf{z}_i, \mathbf{z}_j \rangle \\
&= \sum_{i,j=1}^{n} \sum_{k,l=1}^{n} \int_0^1 \frac{\partial^2}{\partial \mathbf{z}_i \partial \mathbf{z}_j} S_{kl}(t, \mathbf{z}) \\
&\qquad \times \left(\frac{\partial \mathbf{z}_k}{\partial \mu}\right)^T \cdot \sum_{m=1}^{d} \sigma_{\mu,im} \sigma_{\mu,jm} \cdot \left(\frac{\partial \mathbf{z}_l}{\partial \mu}\right) d\mu \\
&\leq \frac{1}{2} s''(\eta_1^2 + \eta_2^2) \sum_{k,l=1}^{n} \int_0^1 \left(\frac{\partial \mathbf{z}_k}{\partial \mu}\right)^T \left(\frac{\partial \mathbf{z}_l}{\partial \mu}\right) d\mu \\
&\leq \frac{1}{2} s''(\eta_1^2 + \eta_2^2) \int_0^1 \left\| \frac{\partial \mathbf{z}}{\partial \mu} \right\|^2 d\mu
\end{aligned}$$

□

References

[1] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, Jan 1949.

[2] V. Capuano, A. Harvard, Y. Lin, and S.-J. Chung, "DGNSS-vision integration for robust and accurate relative spacecraft navigation," in *Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2019)*, Sep 2019, pp. 2923–2939.

[3] J. Yang, A. Dani, S.-J. Chung, and S. Hutchinson, "Inertial-aided vision-based localization and mapping in a riverine environment with reflection measurements," in *AIAA Guidance, Navigation, and Control (GNC) Conference*, Aug 2013.

[4] M. Kalakrishnan, S. Chitta, E. Theodorou, P. Pastor, and S. Schaal, "STOMP: Stochastic trajectory optimization for motion planning," in *2011 IEEE International Conference on Robotics and Automation*, Aug 2011, pp. 4569–4574.

[5] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks, Part I: Theory*. Hanover, MA, USA: Now Publishers Inc., 2009.

[6] R. Cont and P. Tankov, *Financial Modelling With Jump Processes*. London, UK: CRC Press UK, 2004.

[7] A. Patel and B. Kosko, "Stochastic resonance in continuous and spiking neuron models with Lévy noise," *IEEE Transactions on Neural Networks*, vol. 19, pp. 1993–2008, Jan 2009.

[8] E. Sontag, "A Lyapunov-like characterization of asymptotic controllability," *Siam Journal on Control and Optimization*, vol. 21, no. 3, pp. 462–471, May 1983.

[9] H. K. Khalil, *Nonlinear Systems; 3rd ed.* Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[10] J. Slotine and W. Li, *Applied Nonlinear Control*. Eaglewood Cliffs, NJ, USA: Prentice Hall, 1991.

[11] X. Mao, *Stability of stochastic differential equations with respect to semimartingales*, ser. Pitman research notes in mathematics series. Longman Scientific & Technical, 1991. [Online]. Available: https://books.google.com/books?id=ex3vAAAAMAAJ

[12] P. Florchinger, "Lyapunov-like techniques for stochastic stability," in *Proceedings of 1994 33rd IEEE Conference on Decision and Control*, Dec 1994, pp. 1145–1150.

[13] Y. Qin, M. Cao, and B. D. O. Anderson, "Lyapunov criterion for stochastic systems and its applications in distributed computation," *IEEE Transactions on Automatic Control*, Apr 2019. early access.

[14] Hua Deng, M. Krstic, and R. J. Williams, "Stabilization of stochastic nonlinear systems driven by noise of unknown covariance," *IEEE Transactions on Automatic Control*, vol. 46, no. 8, pp. 1237–1253, Aug 2001.

[15] D. Applebaum and M. Siakalli, "Asymptotic stability of stochastic differential equations driven by Lévy noise," *Journal of Applied Probability*, vol. 46, pp. 1116–1129, Dec 2009.

[16] X. Mao, "Exponential stability for stochastic differential equations with respect to semimartingales," *Stochastic Processes and their Applications*, vol. 35, no. 2, pp. 267–277, Aug 1990.

[17] S.-J. Chung and J. E. Slotine, "Cooperative robot control and concurrent synchronization of Lagrangian systems," *IEEE Transactions on Robotics*, vol. 25, no. 3, pp. 686–700, Jun 2009.

[18] J.-J. Slotine, W. Wang, and K. Rifai, "Contraction analysis of synchronization in networks of nonlinearly coupled oscillators," in *Proceedings of the 16th International Symposium on Mathematical Theory of Networks and Systems*, Jan 2004, pp. 5–8.

[19] S.-J. Chung, S. Bandyopadhyay, I. Chang, and F. Y. Hadaegh, "Phase synchronization control of complex networks of Lagrangian systems on adaptive digraphs," *Automatica*, vol. 49, no. 5, pp. 1148–1161, May 2013.

[20] D. Angeli, "A Lyapunov approach to incremental stability properties," *IEEE Transactions on Automatic Control*, vol. 47, no. 3, pp. 410–421, Mar 2002.

[21] W. Lohmiller and J.-J. E. Slotine, "Nonlinear process control using contraction theory," *American Institute of Chemical Engineers Journal*, no. 3, pp. 588–596, Mar 2000.

[22] ——, "Contraction analysis of nonlinear distributed systems," *International Journal of Control*, no. 9, pp. 678–688, 2005.

[23] Q.-C. Pham, N. Tabareau, and J.-J. Slotine, "A contraction theory approach to stochastic incremental stability," *IEEE Transactions on Automatic Control*, pp. 816–820, 2009.

[24] A. P. Dani, S.-J. Chung, and S. Hutchinson, "Observer design for stochastic nonlinear systems via contraction-based incremental stability," *IEEE Transactions on Automatic Control*, vol. 60, no. 3, pp. 700–714, Mar 2015.

[25] M. Dorothy and S.-J. Chung, "Switched systems with multiple invariant sets," *Systems and Control Letters*, vol. 96, pp. 103–109, Oct 2016.

[26] A. Mesquita and J. Hespanha, "Construction of Lyapunov functions for piecewise-deterministic Markov processes," in *Proceedings of the IEEE Conference on Decision and Control*, Dec 2010, pp. 2408–2413.

[27] A. R. Mesquita and J. P. Hespanha, "Jump control of probability densities with applications to autonomous vehicle motion," *IEEE Transactions on Automatic Control*, vol. 57, no. 10, pp. 2588–2598, Oct 2012.

[28] Q. C. Pham, "Analysis of discrete and hybrid stochastic systems by nonlinear contraction theory," in *International Conference on Control, Automation, Robotics and Vision*, Jan 2009, pp. 1054 – 1059.

[29] S. L. Herbert, M. Chen, S. Han, S. Bansal, J. F. Fisac, and C. J. Tomlin, "FaSTrack: A modular framework for fast and guaranteed safe motion planning," in *IEEE 56th Annual Conference on Decision and Control*, Dec 2017, pp. 1517–1522.

[30] S. Bansal, M. Chen, J. Fisac, and C. Tomlin, "Safe sequential path planning of multi-vehicle systems under presence of disturbances and imperfect information," in *American Control Conference*, May 2017.

[31] B. Øksendal and A. Sulem-Bialobroda, *Applied Stochastic Control of Jump Diffusions*, ser. Universitext. Berlin, Germany: Springer-Verlag, 2007. [Online]. Available: https://books.google.com/books?id=s4Qi_WKEtkMC

[32] W.-H. Chen and B. Chen, "Robust stabilization design for nonlinear stochastic system with poisson noise via fuzzy interpolation method," *Fuzzy Sets and Systems*, vol. 217, p. 4161, 04 2013.

[33] M. O'Connell, G. Shi, X. Shi, and S.-J. Chung, "Meta-learning-based robust adaptive flight control under uncertain wind conditions," in *International Conference on Robotics and Automation*, 2020, to appear.

[34] G. Shi, W. Hönig, Y. Yue, and S.-J. Chung, "Neural-swarm: Decentralized close-proximity multirotor control using learned interactions," in *International Conference on Robotics and Automation*, 2020, to appear.

[35] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for non-linear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, Jun 1998.

[36] S.-J. Chung and J.-E. Slotine, "On synchronization of coupled Hopf-Kuramoto oscillators with phase delays," in *49th IEEE Conference on Decision and Control (CDC)*, Dec 2010, pp. 3181–3187.

[37] J. Jouffroy and J. . E. Slotine, "Methodological remarks on contraction theory," in *IEEE Conference on Decision and Control (CDC)*, Dec 2004, pp. 2537–2543.

[38] M. H. A. Davis, "Construction of Brownian motion," 2011. [Online]. Available: http://wwwf.imperial.ac.uk/~mdavis/course_material/sp1/BROWNIAN_MOTION.PDF

[39] P. Protter, *Stochastic Integration and Differential Equation*, 2nd ed. Berlin, Heidelberg: Springer-Verlag, 1992.

[40] M. Jeanblanc. (2007, Apr) Jump processes. [Online]. Available: https://www.maths.univ-evry.fr/pages_perso/jeanblanc/conferences/Cimpa-sauts.pdf

[41] A. Watson. (2016, Sep) Introduction to Lévy processes. [Online]. Available: https://personalpages.manchester.ac.uk/staff/alex.watson/lectures/intro-levy.pdf

[42] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, ser. Universitext. Berlin, Germany: Springer Berlin, 2010. [Online]. Available: https://books.google.com/books?id=EQZEAAAAQBAJ

[43] J. Liu, "On the existence and uniqueness of solutions to stochastic differential equations of mixed Brownian and Poissonian sheet type," *Stochastic Processes and their Applications*, vol. 94, no. 2, pp. 339–354, Aug 2001.

[44] Y. Kasahara and K. Yamada, "Stability theorem for stochastic differential equations with jumps," *Stochastic Processes and their Applications*, vol. 38, no. 1, pp. 13 – 32, 1991.

[45] B. Hajek, "Notes for ECE534: An exploration of random processes for engineers," 2006. [Online]. Available: http://www.ifp.uiuc.edu/~hajek/Papers/randomprocJuly06.pdf

[46] J. Jacod and A. Shiryaev, *Limit Theorems for Stochastic Processes*. Berlin, Germany: Springer Berlin, 2003.

[47] M. Podolskij and M. Vetter, "Understanding limit theorems for semimartingales: a short survey," *Special Issue: Statistical Inference for Lévy Processes with Applications to Finance*, vol. 6, no. 3, pp. 329–351, Aug 2010.

[48] H. Kunita, "Ito's stochastic calculus: Its surprising power for applications," *Stochastic Processes and their Applications*, vol. 120, pp. 622–652, 05 2010.

[49] R. F. Bass. (2009, Apr) SDEs with jumps; Notes for Barrett lectures. [Online]. Available: http://www.math.utk.edu/~jxiong/barrett/Barrett_Bass.pdf

[50] H. J. Kushner, *Stochastic Stability and Control*. Academic Press, 1967.