

# Moment closure and finite-time blowup for Piecewise Deterministic Markov Processes

Lee DeVille<sup>1</sup>, Sairaj Dhople<sup>2</sup>,  
Alejandro D. Domínguez-García<sup>3</sup>, and Jiangmeng Zhang<sup>3</sup>

<sup>1</sup> Dept. of Mathematics, University of Illinois

<sup>2</sup> Dept. of Electrical and Computer Engineering, University of Minnesota

<sup>3</sup> Dept. of Electrical and Computer Engineering, University of Illinois

**Abstract.** We present a variety of results analyzing the behavior of a class of stochastic processes — referred to as Piecewise Deterministic Markov Processes (PDMPs) — on the infinite time interval, and determine general conditions on when the moments of such processes will, or will not, be well-behaved. We also characterize the types of finite-time blowups that are possible for these processes, and obtain bounds on their probabilities.

**Key words.** Hybrid systems; Moment Closure; Power-law distributions

**AMS subject classifications.** 37H10; 34F05; 93C30; 34K34

## 1. Introduction.

**1.1. Background and motivation.** In this paper, we consider a class of stochastic processes referred to as Piecewise Deterministic Markov Processes (PDMP), and present results about their blowup or convergence behavior for long times. We have results in both of the styles of random dynamical systems and moment approximations.

In short, a PDMP is a system that undergoes jumps at random times, and between those times, evolves deterministically, e.g., by an ordinary differential equation (ODE). To consider one specific example, consider an engineered system where the state of the system is given by a vector  $x(t) \in \mathbb{R}^d$  evolving according to an ODE when the system is in “normal” operation. At some random time, the system switches to an “impaired” state, where only the first  $d' < d$  variables of the state are able to evolve according to a different ODE. In general, the switching between normal and impaired operation happens randomly, but in a manner that depends on the state  $x$  (e.g. the ability of the system to be repaired may depend crucially on its current state of operation). The fact that the jumps are determined by the state, and vice versa, means that it is in general impossible to solve such a system explicitly, even if the flow is quite simple.

The study of PDMPs has a long history, going back at least to [5]; the main theoretical foundations of the field were laid out in the book [13]. PDMPs are a subclass of a more general type of system known as Stochastic Hybrid Systems (SHSs), the latter case allows for the evolution between jumps to be stochastic as well (e.g. the state  $x(t)$  would be governed by an stochastic differential equation between jumps). Many of the theoretical results for PDMPs are the same as, or similar to, the analogous ones for SHSs, so in many cases it makes sense to consider this more general context. There is a large literature devoted to understanding the stability properties of such systems [9, 10, 47] (this stability is typically understood in a

moment or almost-sure sense) and more recent work aimed at developing methods to explicitly compute or estimate observables of such systems [26, 24, 25]. Additionally, metastability and large deviations in SHS were studied in [7] using path-integral methods, and limit theorems for PDMP are presented in [37]. SHS represent a powerful formalism that has been applied to many fields, including: networked control systems [1], power systems [15], system reliability theory [16], and chemical reaction dynamics [42]. A recent and comprehensive review of the state of the art of such systems, that also contains an extensive history, bibliography, and list of applications of SHS, is [45].

The state space of a PDMP is comprised of a *discrete state* and a *continuous state*; the pair formed by these is what we refer to as the *combined state* of the PDMP. One can think of a PDMP as a family of ordinary differential equations (ODEs) for the continuous state that are indexed by the discrete state. The discrete state changes stochastically, and we think of this as switching between ODEs. Additionally, each discrete-state transition is associated with a reset map that defines how the pre-transition discrete and continuous states map into the post-transition discrete and continuous states, so that the continuous variable can be “reset” when the discrete state changes. For concreteness, denote the discrete state space by  $\mathcal{Q}$ , and assume that the continuous state space is  $\mathbb{R}^d$ . Then choose  $|\mathcal{Q}|$  different ODEs  $f_1, f_2, \dots, f_{|\mathcal{Q}|}$ , indexed by  $q \in \mathcal{Q}$ , so that when the discrete state is  $q$ ,  $X_t$  evolves according to

$$\frac{d}{dt}X_t = f_q(X_t).$$

Then, we assume that there are a family of rate functions  $\lambda_{q,q'}(x)$  and reset functions  $\psi_{q,q'}(x)$  such that if the discrete state is  $q$ , the probability of a jump to state  $q'$  in the next  $\Delta t$  is  $\lambda_{q,q'}(X_t) \Delta t + o(\Delta t)$ , and, if such a jump occurs, the map  $\psi_{q,q'}$  is applied to the continuous state at the time of jump. [This process can be described more precisely in a “non-asymptotic” formalism, and we do this in Section 2.1.] From this, it follows that there is a continuous-time process,  $Q_t$ , that evolves according to some law, and the full system can be compactly described by

$$\frac{d}{dt}X_t = f_{Q_t}(X_t).$$

It is not hard to show that the process defined in this way is strong Markov [13, §25]. To fully characterize the PDMP, we need to compute the expectation of some large class of functions evaluated on its state space. For the purposes of this paper, we are most interested in the process in equilibrium, or on the way to equilibrium, i.e., the statistics of the process after it has evolved for a long time. The existence of the invariant measures of this class of processes [13, 32] and the existence and smoothness of the corresponding densities [2] has been established, and in one sense, the goal of this paper is to explicitly compute as much about this measure as we can, and we approach this by studying the moments of the process, i.e., the expectation of polynomial functions of the process.

Using the infinitesimal generator of the process, following [25, 16], we can write down a set of differential equations for the moments of this process — we refer to this set of equations as the *moment flow equations*. These equations are *a priori* infinite-dimensional, and as we show below, in a wide variety of circumstances they are “inherently infinite-dimensional,” by

which we mean that: (i) there is no projection of the dynamics onto a finite-dimensional subspace, and (ii) any approximate projection into a finite-dimensional subspace behaves poorly in a sense to be made precise below. This inherent infinite-dimensionality is typically called a *moment closure problem* in a variety of physical and mathematical contexts, in the sense that one cannot “close” the moment flow equations in a finite way. Examples of moment closure problems and various approaches to handle them span a wide range, including applications in: chemical kinetics [33], dynamic graphs [22, 23, 39], physics [20], population dynamics/epidemiology [28, 27, 36, 46], and nonlinear PDE [8, 12]. Our approach is related to and inspired by the Lyapunov moment stability theory that has been well-developed for diffusions, especially those with small noise, over the past few decades [29, 3, 4, 43].

We mention that a moment closure problem for PDMP (and similarly for general classes of SHS) was solved nicely in [24, 25] and related works, but in contrast the problem there was to find a moment closure approximation that was valid for a finite time interval. Since we are interested in equilibrium or near-equilibrium statistics of the problem, we want to study the problem on the infinite time interval.

**1.2. Overview of the results of this paper.** In this work, we focus on the scenario where the two parts of the system (the ODE and the reset) have competing effects. For example, if the ODE and the reset both send all orbits to zero, then their mixture does as well, and, conversely, if they send all orbits to infinity, then the mixture does as well. What is more interesting is the case in which the ODE and the resets have opposite effects, and we consider one such case here: the ODE sends orbits to infinity, but the resets send orbits toward zero. (In fact, some of the examples and results below are more general, but this is the framework that we are thinking of throughout the paper.)

We give a few prototypical examples of such systems, and discuss the results of this paper that apply to each. In each of these cases, we present a model of the continuous variable of which is defined on the positive real line; this can be thought of as the fundamental model of interest, or the radial component of a system in higher dimension.

**Example 1.** Let  $\alpha, \beta > 0$  and  $\gamma \in (0, 1)$  and define the ODE, jump rate, and reset as

$$(1.1) \quad \dot{x} = f(x) = \alpha x, \quad \lambda(x) = \beta x, \quad x \mapsto \gamma x.$$

The theory presented in Section 3 implies that the PDMP in (1.1) converges to an invariant distribution with all finite moments of all orders, and this fact is independent of the values of  $\alpha, \beta, \gamma$  (although of course the moments themselves depend on these parameters).

**Example 2.** Next, consider the system

$$(1.2) \quad \dot{x} = f(x) = \alpha x^2, \quad \lambda(x) = \beta x, \quad x \mapsto \gamma x.$$

The theory of Section 5 tells us that, depending on the value of  $\gamma$ , this process can have multiple behaviors: (i) if  $\gamma > e^{-\alpha/\beta}$ , then all solutions of (1.2) blow up in finite time; (ii) if  $1 - \alpha/\beta < \gamma < e^{-\alpha/\beta}$ , then solutions go to zero with high probability, but enough of them escape to infinite fast enough that the moments of the solution blow up in finite time; (iii) if  $\gamma < 1 - \alpha/\beta$ , then the solution converges to zero in every sense. (In fact, these three cases and their description is exactly the content of Corollary 5.10.)

**Example 3.** Finally, consider the system

$$(1.3) \quad \dot{x} = f(x) = \alpha x^3, \quad \lambda(x) = \beta x, \quad x \mapsto \gamma x.$$

Theorem 5.2 tells us that all solutions of this system blow up with probability one, and this fact is independent of the parameters  $\alpha, \beta, \gamma$ .

**Remark 1.** *The theory in Sections 3 and 5 generalizes the statements in Examples 1 – 3 to arbitrary polynomials. The critical difference between the three cases are the relative degrees, and in some cases the leading-order coefficients, of  $f$  and  $\lambda$ . Additionally, although the two latter cases both exhibit finite-time blowup, they are of a much different character. In the quadratic case, the system blows up in finite time even though there are infinitely many jumps, and as such, exhibit a quality very much like that of an explosive Markov chain. In the cubic case, the blowup is more like that seen in nonlinear ODEs: the system goes off to infinity in finite time and there are only a finite number of jumps.*

In Examples 1 – 3, there was a single discrete state and all jumps mapped the state back to itself. We also analyze the moments of the PDMP with multiple discrete states in Section 4. We show that the behavior can be characterized similarly to that of the one state case in many situations, but we also exhibit new types of behavior here. For example, we show that in many cases, the system can exhibit *marginal moment stability*, by which we mean that in equilibrium, the system can have some moments finite, and others that are infinite; thus the equilibrium distribution has fat tails. In this case, we show numerically that the equilibrium distributions have power law behavior.

We also show that the moment flow equations have a very strange property. These equations are an infinite-dimensional linear system that supports a fixed point that corresponds to the moments of the invariant measure. We show that any finite-dimensional truncation of this system has only unstable fixed points, and, moreover, as the size of truncation grows, the system has larger positive eigenvalues. This motivates the result that the infinite-dimensional system is ill-posed in “the PDE sense” in a manner analogous to a time-reversed heat equation. However, we then show that the minimal amount of convexity given by Jensen’s Inequality is enough to make this system well-posed and well-behaved, and in particular it then becomes faithful to the stochastic process. This seems to be a strange example of a system that is ill-behaved on a linear space becoming well-behaved when restricted to a nonlinear submanifold of that same space.

**1.3. Organization of manuscript.** The main results and structure of this paper are as follows. In Section 2, we give a formal definition of the PDMP that we study, and define the moment flow. Next, we identify a broad class of assumptions for PDMPs under which the moment equations are well-defined and accurate on the infinite-time interval, but we also show that there is a surprising subtlety that arises; these results are contained in Sections 3 and 4. Finally, in Section 5, we study the PDMP where moment closure fails, and in fact the PDMP undergoes finite-time blowups, i.e., the process becomes infinite in expectation, or almost surely, at a finite time.

**2. Definitions and basic properties.** In this section, we provide the formal definition of a PDMP, and define its generator. Then, we use the generator together with Dynkin’s formula

to develop a set of differential equations that describe the dynamics of the moments of the PDMP.

**2.1. Definition of PDMP.** *Definition 2.1.* A piecewise-deterministic Markov process (PDMP) is a quintuple  $(\mathcal{Q}, \mathcal{S}, \Lambda, \Psi, f)$  where

- $\mathcal{Q}$  is a countable set of **discrete states**;
- $\mathcal{S}$  is a Euclidean space which we call the **continuous state space**;
- $\Lambda(q, x) = (\lambda_{q'}(q, x))_{q' \in \mathcal{Q}}$  is a collection of **rate functions**, and each  $\lambda_{q'}: \mathcal{Q} \times \mathcal{S} \rightarrow \mathbb{R}^+$ ;
- $\Psi(q, x) = (\psi_{q'}(q, x))_{q' \in \mathcal{Q}}$  is a collection of **reset maps**, and each  $\psi_{q'}: \mathcal{Q} \times \mathcal{S} \rightarrow \mathbb{R}$ ;
- $f: \mathcal{Q} \times \mathcal{S} \rightarrow \mathbb{R}$  is a collection of **vector fields**.

We will denote the **flow map generated by**  $f(q, \cdot)$  by  $\varphi^t(q, \cdot)$ , i.e.,

$$\varphi^0(q, x) = x, \quad \varphi^s(q, \varphi^t(q, x)) = \varphi^{s+t}(q, x), \quad \text{and} \quad \frac{d}{dt}\varphi^t(q, x) = f(q, \varphi^t(q, x)).$$

The **state** of the process is the pair  $(Q_t, X_t)$ . Assume that  $(Q_0, X_0)$  is known almost surely.

Let us define  $(S_n^{(q)})_{q \in \mathcal{Q}, n \geq 1}$  to be a doubly-indexed i.i.d. sequence of random variables that are each exponentially distributed with unit rate, i.e.

$$\mathbb{P}(S_n^{(q)} > z) = e^{-z}, \quad \text{for all } z \geq 0.$$

Define stopping times  $T_1^{(q)}$ ,  $q \in \mathcal{Q}$  and  $T_1$  as follows:

$$(2.1) \quad \int_t^{T_1^{(q)}} \lambda_q(Q_s, \varphi^{s-t}(Q_s, X_s)) ds = S_1^{(q)}, \quad T_1 = \inf_{q \in \mathcal{Q}} T_1^{(q)}.$$

Specifically, each  $T_1^{(q)}$  is the time at which the system would make a transition to state  $q$  if  $\lambda_q$  were the only positive jumping rate; we take the first one to do so and discard the others. The time  $T_1$  will be the time of the **first jump**. We prescribe that the discrete state remains unchanged until the next jump, and the continuous state flows according to the appropriate ODE, i.e.,

$$(2.2) \quad \text{for all } s \in [0, T_1), \quad Q_s = Q_0, \quad X_s = \varphi^s(Q_0, X_0).$$

Finally, we apply the appropriate reset map:

$$(2.3) \quad Q_{T_1} = \arg \inf_{q \in \mathcal{Q}} T_1^{(q)}, \quad X_{T_1} = \psi_{Q_{T_1}}(Q_0, X_{T_1-}),$$

where here and below we define

$$X_{a-} = \lim_{t \nearrow a} X_t.$$

Now that we know  $(Q_{T_1}, X_{T_1})$ , we define the process recursively. For any  $n \in \mathbb{N}$ , if we know  $T_n$  and  $(Q_{T_n}, X_{T_n})$ , then define  $T_{n+1}$  by

$$(2.4) \quad \int_t^{T_{n+1}^{(q)}} \lambda_q(Q_s, \varphi^{s-t}(Q_s, X_s)) ds = S_{n+1}^{(q)}, \quad T_{n+1} = \inf_{q \in \mathcal{Q}} T_{n+1}^{(q)}.$$

Then define  $Q_t, X_t$  for  $t \in [T_n, T_{n+1}]$  as in (2.2, 2.3).

We also use the convention throughout of **minimality**: if  $T_\infty := \lim_{n \rightarrow \infty} T_n < \infty$ , then we say the process **explodes** or **blows up** at time  $T_\infty$ , and set  $X_t = \infty$  for all  $t > T_\infty$ . Similarly, or if there is a  $T^*$  with  $\lim_{t \nearrow T^*} X_t = \infty$ , then we say the process **explodes** or **blows up** at time  $T^*$ , and set  $X_t = \infty$  for all  $t > T^*$ .

We will also write  $(X_t, Q_t) = \text{PDMP}(\mathcal{Q}, \Lambda, \Psi, \varphi)$  to mean that  $(X_t, Q_t)$  is a realization of the stochastic process constructed using the procedure above.

**Remark 2.** For the remainder of this paper, we will do all analysis on the case where  $\mathcal{S} = \mathbb{R}^+$ . However, we are motivated by thinking of the state in  $\mathbb{R}^+$  as the radial component in a higher-dimensional space. For example, one can see if we have a PDMP with  $\mathcal{S} = \mathbb{R}^d$ , but the rates, flows, and resets depend only on the radial component of  $x(t)$ , then we can reduce this system to the case where  $\mathcal{S} = \mathbb{R}^+$ .

Also, in several parts of the sequel we consider the case where  $\mathcal{Q}$  is a singleton. In these cases, for simplicity we will write  $\text{PDMP}(\lambda, \psi, f)$ .

**Remark 3.** Note that all of the randomness of this process is chosen “up front”, i.e., depends only on the streams of iid exponentials  $S_n^{(q)}$  for  $n \in \mathbb{N}, q \in \mathcal{Q}$  which can be chosen at the beginning. This allows us to define a map from any probability space  $\Omega$  rich enough to contain the streams  $S_n^{(q)}$  to  $\mathcal{D}([0, \infty), \mathcal{Q} \times \mathcal{P})$ , the set of cadlag paths defined on  $\mathcal{Q} \times \mathcal{P}$ . This induces a measure on the set of all paths, and, in particular, allows us to measure the probability of any event that can be determined by observing the paths. When we talk about probabilities of events below, we are implicitly assuming that this correspondence has been made. Moreover, this will even allow us to compare paths of two (or more) different PDMP that are generated by different functions; as long as we have a correspondence from  $\omega \in \Omega$  to the exponential streams  $S_n^{(q)}(\omega)$ , the paths are completely determined. We will use this formalism throughout the remainder of the paper without further comment, and in general drop the dependence on  $\omega$ .

**Proposition 2.2.** We consider the PDMP in Definition 2.1. Start at state  $(Q_0, X_0)$ , and define  $T$  as the time of the first jump. Then for  $t < T$ , the discrete state  $Q_0$  does not change, and  $X_t$  evolves according to the ODE, so  $X_t = \varphi^t(Q_0, X_0)$ . Then

$$\begin{aligned} \mathbb{P}(T \leq t + \Delta t | T > t) &= \Delta t \cdot \sum_q \lambda_q(Q_0, X_t) = \Delta t \cdot \sum_q \lambda_q(Q_0, \varphi^t(Q_0, X_0)), \\ \mathbb{P}(Q_{t+\Delta t} = q) &= \frac{\lambda_q(Q_0, X_t)}{\sum_{q'} \lambda_{q'}(Q_0, X_t)}. \end{aligned}$$

*Proof.* The proof is a straightforward computation, using the definitions of exponential random variables. ■

**Remark 4.** This definition is a bit involved, but the salient intuitive features are given by Proposition 2.2. Between jumps, the continuous state evolves according to the appropriate ODE, and the discrete state remains unchanged. The jumps are governed by the rates  $\lambda_q(Q_t, X_t)$ , which drives both the time of the next jump and the subsequent state.

**2.2. Infinitesimal generator.** We follow the standard definition of the infinitesimal generator and derive the generator of the process here. Let  $h: \mathcal{Q} \times \mathcal{S} \rightarrow \mathbb{R}$  be bounded and

differentiable in the second component, and define the following linear operator  $\mathcal{L}$ :

$$(2.5) \quad \mathcal{L}h(q, x) := \lim_{\epsilon \searrow 0} \frac{\mathbb{E}[h(Q_{t+\epsilon}, X_{t+\epsilon}) | Q_t = q, X_t = x] - h(q, x)}{\epsilon},$$

which can also be written as

$$(2.6) \quad \frac{d}{dt} \mathbb{E}[h(Q_t, X_t)] = \mathbb{E}[\mathcal{L}h(Q_t, X_t)],$$

or, said another way, the operator

$$(2.7) \quad \mathcal{M}h(Q_t, X_t) = h(Q_0, X_0) + \int_0^t \mathcal{L}h(Q_s, X_s) ds$$

defines a martingale. This allows us to extend the definition of the domain of  $\mathcal{L}$  through the martingale equation, and it is not hard to show that under weak assumptions on  $\Lambda$ ,  $\Psi$ , and  $f$ , the extended domain of  $\mathcal{L}$  contains all polynomials [13, §2.6,2.7]. In particular, we can compute directly that

$$(2.8) \quad \mathcal{L}h(q, x) = f(q, x) \cdot \nabla_x h(q, x) + \sum_{q' \in \mathcal{Q}} \lambda_{q'}(q, x) (h(q', \psi_{q'}(q, x)) - h(q, x)).$$

To give a formal derivation of (2.8), we compute

$$\begin{aligned} & \mathbb{E}[h(Q_{t+\epsilon}, X_{t+\epsilon}) - h(Q_t, X_t) | Q_t = q, X_t = x] \\ &= \mathbb{E}[h(Q_{t+\epsilon}, X_{t+\epsilon}) - h(Q_t, X_t) | Q_t = q, X_t = x, \text{ no jump}] \mathbb{P}(\text{no jump}) \\ &+ \sum_{q' \in \mathcal{Q}} \mathbb{E}[h(Q_{t+\epsilon}, X_{t+\epsilon}) - h(Q_t, X_t) | Q_t = q, X_t = x, \text{ jump to } q'] \mathbb{P}(\text{jump to } q') \\ &= (h(q, x + \epsilon f(q, x)) - h(q, x)) (1 - O(\epsilon)) \\ &+ \sum_{q' \in \mathcal{Q}} (h(q', \psi_{q'}(q, x)) - h(q, x)) (\epsilon \lambda_{q'}(q, x)) + O(\epsilon^2) \\ &= \epsilon (f(q, x) \cdot \nabla_x h(q, x) + \sum_{q' \in \mathcal{Q}} \lambda_{q'}(q, x) (h(q', \psi_{q'}(q, x)) - h(q, x))) + O(\epsilon^2), \end{aligned}$$

and from this and the definition of  $\mathcal{L}$  we have established (2.8) formally. We have not been careful to specify the domain of  $\mathcal{L}$ , but it is clear from this argument that this derivation is valid for all  $h \in C^1 \cap L^\infty$ . As is shown in [13, §2.6,2.7], we can then extend the domain of the generator using (2.7) to encompass all polynomials and indicator functions. Strictly speaking, this means that the new generator is an extension of the previous one, and we will use this without further comment in the sequel.

**2.3. Moment equations.** A quick perusal of (2.8) makes it clear that if we assume that  $f(q, \cdot)$ ,  $\lambda_{q'}(q, \cdot)$ ,  $\psi_{q'}(q, \cdot)$  are polynomials in  $x$ , then the right-hand side sends polynomials to polynomials. More precisely, if  $h(q, x)$  is any function that is polynomial in  $x$ , then  $\mathcal{L}h(q, x)$  is also polynomial in  $x$ . In particular, we denote  $h_q^{(m)}(q, x) = x^m \delta_{q, q'}$ , and we see that  $\mathcal{L}h_q^{(m)}$

is a polynomial in  $x$ , so that (2.6) becomes an (infinite-dimensional) linear ODE on the set  $\{h_q^{(m)}\}_{m \in \mathbb{N}^d, q \in \mathcal{Q}}$ . However, the first challenge that we obtain is clear: if the degree of any  $f(q, \cdot)$  is greater than one, or the degree of any of the  $\lambda_{q'}(q, \cdot)$  or  $\psi_{q'}(q, \cdot)$  are positive, then we see that the degrees of the terms on the right-hand side of the equation are higher than those on the left, and we are thus led to the problem of *moment closure*. On the other hand, we still have a linear system, even if it is infinite-dimensional. Thus, we might hope to make sense of this flow by writing down a semigroup on a reasonable function space. In fact, we show in Section 3.2 that this is in general not possible.

**3. Moment Closure and Convexity — One state.** One of the main results of this paper is that under certain conditions on the growth of the functions  $f, \lambda$  as  $x \rightarrow \infty$ , the moment equations are well-behaved and useful. We will first present the simpler case where there is one state, i.e.,  $|\mathcal{Q}| = 1$ ; there is one reset map  $\psi$ . Due to this simplification, we now say that  $X_t = \text{PDMP}(\lambda, \psi, f)$ , where again we have the flow map  $\varphi$ :

$$\frac{d}{dt}\varphi^t(x) = f(\varphi^t(x)), \quad \varphi^0(x) = x.$$

We want to consider the case where the ODE has an unstable fixed point at the origin, and to balance this we want the reset map to move towards the origin, so that  $\psi(x) \leq x$ . If  $\psi$  is also polynomial, we need to choose  $\psi(x) = \gamma x$  with  $\gamma \in (0, 1)$ . We also assume that  $f(x), \lambda(x)$  are polynomials of degrees  $\deg(f), \deg(\lambda)$  respectively.

**3.1. Moment flow equations.** The generator for  $X_t = \text{PDMP}(\lambda, \gamma x, f)$  is

$$\mathcal{L}h(x) = f(x)\frac{dh}{dx}(x) + \lambda(x)(h(\gamma x) - h(x)).$$

The test functions that we are interested in are of the form  $h^{(m)}(x) := x^m$ , and we want to study the evolution of the moments  $\mu_m := \mathbb{E}[X_t^m] = \mathbb{E}[h^{(m)}(X_t)]$ . Plugging this in, we obtain

$$\mathcal{L}h^{(m)}(x) = f(x)\frac{m}{x}h^{(m)}(x) + \lambda(x)(\gamma^m - 1)h^{(m)}(x).$$

Since  $f(0) = 0$ ,  $f(x)/x$  is a linear combination of terms of nonnegative degree, so the first term is a polynomial with all powers at least  $m$ . If we write

$$f(x) = \sum_{\ell=1}^{\deg(f)} \alpha_\ell x^\ell, \quad \lambda(x) = \sum_{\ell=1}^{\deg(\lambda)} \beta_\ell x^\ell,$$

then, by taking expectations, we have the **moment flow** equations:

$$(3.1) \quad \frac{d}{dt}\mu_m = \sum_{\ell=m}^{m+\deg(\lambda)} C_{m,\ell}\mu_\ell,$$

where

$$(3.2) \quad \begin{aligned} C_{m,m} &= m\alpha_1, & C_{m,m+\deg(\lambda)} &= \beta_{\deg(\lambda)}(\gamma^m - 1), \\ C_{m,\ell} &= m\alpha_{\ell+1} + \beta_\ell(\gamma^m - 1). \end{aligned}$$

It is not hard to see that  $C_{m,m+\deg(\lambda)} < 0$  for all  $m > 0$ , and  $C_{m,\ell} > 0$  for  $m \leq \ell < m + \deg(\lambda) - 1$  and  $m$  sufficiently large (in fact,  $C_{m,\ell} \rightarrow \infty$  linearly in  $m$  for any fixed  $\ell$ ). Under these assumptions, we can state the theorem:

**Theorem 3.1.** *If  $\deg(f) \leq \deg(\lambda)$ , then the moment flow equations have a globally attracting fixed point. The set of fixed points for (3.1) form a  $\deg(\lambda)$ -dimensional linear manifold parameterized by an infinite family of linear equations. (Thus we cannot find this fixed point simply by using the algebraic relations of (3.1).)*

We delay the proof of the theorem until later, but first we present a paradox that makes the result a bit surprising.

**3.2. A paradox of posedness.** Using the results of [32], it follows that there exists a unique invariant measure to which the system converges at an exponential rate. Choose  $h(x) = x$ , then  $\mathcal{L}h(x)$  is a polynomial whose leading coefficient is negative. Thus, there is a  $b$  with  $\mathcal{L}h \leq -h + b\mathbf{1}_{\mathcal{S}}$ , where  $\mathcal{S}$  is a compact subset of the positive reals, and thus we have a unique invariant measure to which we converge exponentially quickly [32, Theorem 14.0.1].

This suggests that the moment flow equations (3.1) should be well-behaved and tend to an equilibrium solution in some limit. However the moment flow equations *sui generis* are ill-posed in a very precise sense:

**Definition 3.2 (Ill-posed).** *Given a state space  $X$  and a flow map  $\varphi: X \times \mathbb{R} \rightarrow X$ , we say that the flow is ill-posed if, for any  $x_0 \in X$ , any  $t > 0$ , and any  $\epsilon > 0$ , there is a  $y \in X$  with  $|x_0 - y| < \epsilon$  and  $|\varphi(t, x_0) - \varphi(t, y)| > 1$ . We will slightly abuse notation and say an ODE is ill-posed if its flow map is ill-posed.*

Then we have the following:

**Proposition 3.3 (Ill-posedness).** *The linear system (3.1) does not generate a strongly continuous semigroup on any  $\ell^p$  space (or, in fact, on any subspace of  $\mathbb{R}^\infty$  that contains vectors of finite support). Specifically, the ODE is ill-posed in the sense of Definition 3.2 on any of these spaces.*

*Proof.* Noting that the system is upper-triangular, and that the diagonal elements increase without bound, it is not hard to see that the spectrum of this matrix should be unbounded. To be more specific: if we consider any  $M \times M$  truncation of this matrix, it has eigenvalues  $C_{m,m}$  with  $m = 1, \dots, M$ . It also has a basis of eigenvectors, which we will call  $v_1, \dots, v_M$ . These eigenvectors embed into  $\mathbb{R}^\infty$  in the obvious way and are thus eigenvalues of the matrix  $A$ . Let  $V$  be any linear space containing all of the  $v_k$  (NB: any  $\ell^p$  space, with  $1 \leq p \leq \infty$  would be appropriate). Then with  $A$  considered as a linear operator from  $V$  to itself,  $C_{m,m} \in \text{Spec}(A)$ . Recall from above that  $C_{m,m} \rightarrow \infty$  linearly fast in  $m$ . By the Hille–Yosida Theorem [38, §142,143], this flow posed on  $V$  does not generate a strongly continuous semigroup. ■

**Remark 5.** *In particular, we have shown that every finite-dimensional truncation of the moment flow equations is linearly unstable, and in fact the instability worsens with the order of the truncation.*

**3.3. Convexity to the rescue.** The above certainly seems paradoxical. One method from stochastic processes tells us that the behavior of a system is well-behaved as  $t \rightarrow \infty$  (in fact, has a globally attracting fixed point), yet, on the other hand, a method from linear analysis tells us the system is quite ill-behaved (being ill-posed on finite-time domains).

However, the linear system (3.1) does not contain all of the information given by the problem. This flow is given by the flow of moments of a stochastic process. In this light, the vectors  $v_k$  from the proof above are “illegal” perturbations, because there can be no random variable whose moments are given by the  $v_k$  defined in the proof of Proposition 3.3. In fact, there can be no random variable whose moments are given by a vector with entries that go to zero, or even has bounded entries. This is due to Jensen’s Inequality [6]:

**Theorem 3.4 (Jensen’s Inequality).** *If  $g(\cdot)$  is any convex function and  $X$  a real-valued random variable, then*

$$(3.3) \quad g(\mathbb{E}[X]) \leq \mathbb{E}[g(X)],$$

*with equality only if  $X$  is “deterministic”, i.e., the distribution for  $X$  is an atom. In particular, since  $g(x) = x^p$  is convex on  $[0, \infty)$  for any  $p > 0$ , if  $X \geq 0$ , then*

$$\mu_q^{p/q} = (\mathbb{E}[X^q])^{p/q} \leq \mathbb{E}[X^p] = \mu_p,$$

*or, equivalently,*

$$(3.4) \quad \mu_p \geq \mu_q^{p/q} \text{ for all } p \geq q.$$

In particular, this means that the moments are not actually independent coordinates in some vector space, and we are not allowed arbitrary perturbations of a fixed point. For example, reconsidering (3.1, 3.2), we see that the largest moment has a negative coefficient, and from Jensen this grows superlinearly with respect to  $\mu_m$ . This will be enough to prove stability:

**Lemma 3.5.** *If we assume the equations (3.1), and, further, that  $\mu_m$  are the moments of a random variable, then as long as*

$$(3.5) \quad \mu_m > \left( \frac{\deg(\lambda)C_{m,\ell}}{|C_{m,m+\deg(\lambda)}|} \right)^{m/(m-\ell+\deg(\lambda))}$$

*for each  $C_{m,\ell}$  that is positive, then the right-hand side of (3.1) is negative, and thus  $\mu_m$  is decreasing. Since  $C_{m,m+\deg(\lambda)} < 0$ , we have that  $\ell \in \{m, \dots, m + \deg(\lambda) - 1\}$ , so the denominator in the exponent in (3.5) is always positive. Since this puts a finite number of constraints on  $\mu_m$ , this means that  $\mu_m$  is inflowing on a compact subset of  $\mathbb{R}^+$  and thus has bounded trajectories.*

**Proof.** First recall that  $C_{m,m+\deg(\lambda)} < 0$ . Choose an  $\ell$  such that  $C_{m,\ell} > 0$  (if none such exist, then we are done.) If

$$\mu_m > \left( \frac{\deg(\lambda)C_{m,\ell}}{|C_{m,m+\deg(\lambda)}|} \right)^{m/(m-\ell+\deg(\lambda))},$$

then

$$\mu_m^{\frac{m-\ell+\deg(\lambda)}{m}} > \frac{\deg(\lambda)C_{m,\ell}}{|C_{m,m+\deg(\lambda)}|}.$$

Since

$$\mu_\ell \geq \mu_m^{\ell/m},$$

this means

$$C_{m,\ell} - \frac{|C_{m,m+\deg(\lambda)}|}{\deg(\lambda)} \mu_\ell^{(m+\deg(\lambda)-\ell)/\ell} < 0$$

$$C_{m,\ell} \mu_\ell - \frac{|C_{m,m+\deg(\lambda)}|}{\deg(\lambda)} \mu_\ell^{(m+\deg(\lambda))/\ell} < 0.$$

By Jensen again, this implies that

$$C_{m,\ell} \mu_\ell - \frac{|C_{m,m+\deg(\lambda)}|}{\deg(\lambda)} \mu_{m+\deg(\lambda)} < 0.$$

(If  $C_{m,\ell} < 0$ , then the previous inequality is satisfied trivially.) From this we have

$$\sum_{\ell=m}^{m+\deg(\lambda)-1} C_{m,\ell} \mu_\ell < \sum_{\ell=m}^{m+\deg(\lambda)-1} \frac{|C_{m,m+\deg(\lambda)}|}{\deg(\lambda)} \mu_{m+\deg(\lambda)} = |C_{m,m+\deg(\lambda)}| \mu_{m+\deg(\lambda)},$$

and so

$$\sum_{\ell=m}^{m+\deg(\lambda)-1} C_{m,\ell} \mu_\ell + C_{m,m+\deg(\lambda)} \mu_{m+\deg(\lambda)} < 0.$$

■

Now, we are ready to provide a formal proof to the result in Theorem 3.1.

**Proof of Theorem 3.1.**

By Lemma 3.5, the solution of (3.1) has bounded trajectories. In particular, choosing

$$M_m = \max_{m \leq \ell \leq m+\deg(\lambda)-1} \left( \frac{\deg(\lambda) C_{m,\ell}}{|C_{m,m+\deg(\lambda)}|} \right)^{m/(m-\ell+\deg(\lambda))},$$

we have  $\mu_m(t) \leq M_m$  for  $t$  sufficiently large.

We also note that the steady-state solution of this system satisfies the (infinite) family of linear equations

$$\sum_{\ell=m}^{m+\deg(\lambda)} C_{m,\ell} \mu_\ell = 0,$$

or

$$\mu_{m+\deg(\lambda)} = -\frac{1}{C_{m,m+\deg(\lambda)}} \sum_{\ell=m}^{m+\deg(\lambda)-1} C_{m,\ell} \mu_\ell.$$

For  $m$  sufficiently large, all of the coefficients on the right-hand side of the equation are positive.

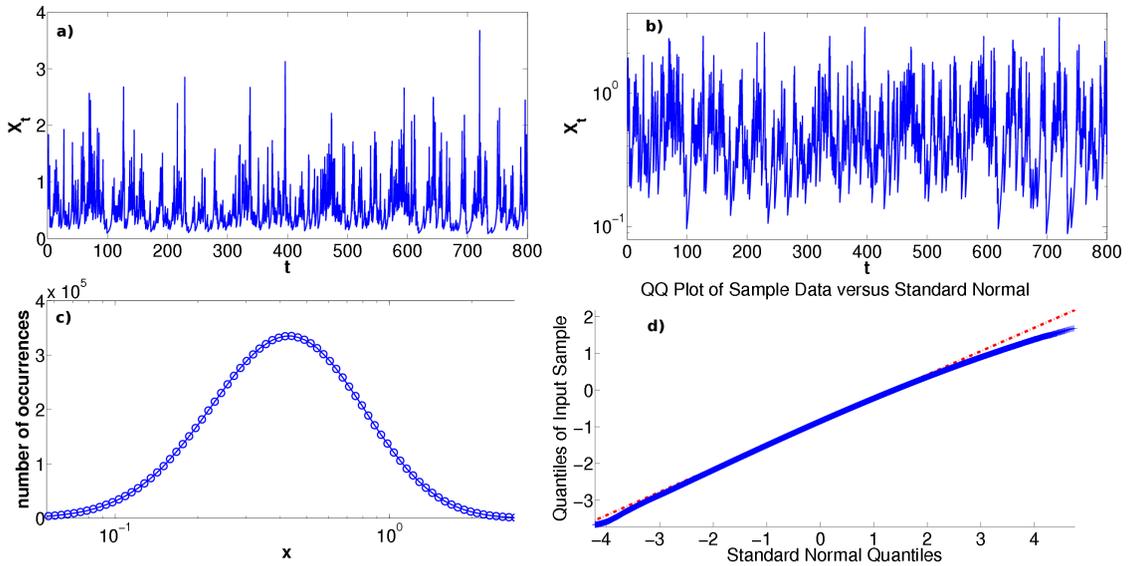
Then, this linear system has exactly  $\deg(\lambda)$  degrees of freedom. Choose  $m^*$  so that  $C_{m,\ell} > 0$  for  $\ell = m, \dots, m + \deg(\lambda) - 1$ . Then, if  $\mu_{m^*}, \dots, \mu_{m^*+\deg(\lambda)}$  are given, then the 1st equation gives a unique solution for  $\mu_{m^*+\deg(\lambda)+1}$ , and then the second equation would give a unique solution for  $\mu_{m^*+\deg(\lambda)+2}$ , and so on. ■

**3.4. Numerical results.** In Figure 1, we plot the results of a Monte Carlo simulation of a one-state system as described above, where we choose

$$f(x) = x^2, \quad \lambda(x) = 2x^2, \quad \gamma = 1/2.$$

We simulated  $10^3$  realizations of the process, and ran each one until  $10^4$  numerical steps occurred. The method we used was a hybrid Gillespie–1st order Euler method: we choose and fix a  $\Delta t_{\max} \ll 1$ . Given  $X_t$ , we have  $\lambda(X_t)$ , and we determine the time of the next jump as  $dt = -\log(U)/\lambda(X_t)$ , where  $U$  is a uniform  $[0, 1]$  random variable. If  $dt < \Delta t_{\max}$ , then we integrate the ODE using the 1st-order Euler method for time  $dt$  (i.e., we set  $X_{t+dt} = X_t + dt \cdot f(X_t)$ ), then multiply by  $\gamma$ . If  $dt \geq \Delta t_{\max}$ , then we integrate the ODE for time  $\Delta t_{\max}$  and do not jump. It is clear that in the limit as  $\Delta t_{\max} \rightarrow 0$ , this converges in every sense to the stochastic process, and it is also not hard to see that this is equivalent to computing the trajectory and next jump using formulas (2.1) and (2.2) by discretizing the integral in (2.1) using timesteps of  $\Delta t_{\max}$ .

In Figures 1a and 1b we plot one realization of the process, in linear and in log coordinates. To the eye,  $\log(X_t)$  looks almost like an Ornstein–Uhlenbeck process, and this is borne out by the distribution in Figure 1c, where we see that the invariant distribution of  $X_t$  is very close to log-normal, at least to the eye. To check this, we plot a QQ-plot of  $\log(X_t)$  versus a normal distribution with the same mean and variance, and we see that the distribution is not quite log-normal — the fact that this plot is concave down means that this distribution is a little “tighter” than a normal distribution. Thus the distribution of  $\log(X_t)$  looks like a Gaussian up to two or three standard deviations, but has smaller tails. In any case, to verify that this is not just due to sampling error, one can plug the general log-normal into the formal adjoint  $\mathcal{L}^*$  derived from (2.5), and see by hand that log-normals are not in the nullspace.



**Figure 1.** Plots of realizations and histograms of one-state system, see description in text for more detail.

**4. Moment closure and convexity — Multiple states.** We now extend the results of the previous section to the case where there are multiple states, i.e.,  $|\mathcal{Q}| > 1$ . The approach is the same as above: when we write down the moment equations, if we can show that if the right-hand side of the evolution equation for each moment has a term of higher degree, and this term has a negative coefficient, then the previous approach works as well. As before, we will assume that our state space is the positive reals, and that our reset maps are linear, i.e.,  $\psi_{kl}(x) = \gamma_{kl}x$ , with  $\gamma_{kl} > 0$ .

The main results of this section are Theorems 4.2 and 4.3; the former gives sufficient conditions for the moment flow system to have bounded orbits, and the latter gives sufficient conditions for the moment flow system to have unbounded orbits. The technique used in the proofs of these theorems is, in principle, the same as in the previous section: there we had a case where, in each equation, there was a term involving the higher-order moment with a negative coefficient. In this case, we will show that there is a term on the right-hand side that plays the same role as this negative coefficient, but in this case, since the  $m$ th moment is effectively a vector of length  $|\mathcal{Q}|$ , this will be a  $|\mathcal{Q}| \times |\mathcal{Q}|$  matrix whose sign-definiteness will establish stability (or the lack thereof).

**4.1. Moment flow equations.** We recall

$$(2.8) \quad \mathcal{L}h(q, x) = f(q, x) \cdot \nabla_x h(q, x) + \sum_{q' \in \mathcal{Q}} \lambda_{q'}(q, x) (h(q', \psi_{q'}(q, x)) - h(q, x)).$$

Let us define, for  $\theta \in \mathcal{Q}$  and  $m \in \mathbb{N}$ ,

$$h_\theta^{(m)}(q, x) := 1_\theta(q)x^m = \delta_{q, \theta}x^m, \quad \mu_\theta^{(m)} := \mathbb{E}[h_\theta^{(m)}(Q_t, X_t)].$$

Note that

$$\sum_{\theta \in \mathcal{Q}} h_\theta^{(m)} = x^m, \text{ and thus } \mu^{(m)} := \sum_{\theta \in \mathcal{Q}} \mu_\theta^{(m)} = \mathbb{E}[X_t^m]$$

is the total  $m$ th moment of  $X_t$ . One can think of  $\mu_\theta^{(m)}$  as the conditional  $m$ th moment of  $X_t$ , conditional on  $Q_t = \theta$ , times the probability that  $Q_t = \theta$ . We plug  $h_\theta^{(m)}$  into (2.8), to obtain

$$\begin{aligned} \mathcal{L}h_\theta^{(m)}(q, x) &= f(q, x) \cdot \nabla_x h_\theta^{(m)}(q, x) + \sum_{\ell \in \mathcal{Q}} \lambda_\ell(q, x) h_\theta^{(m)}(\ell, \psi_\ell(q, x)) - \sum_{\ell \in \mathcal{Q}} \lambda_\ell(q, x) h_\theta^{(m)}(q, x) \\ &= f(q, x) 1_\theta(q) \cdot m \cdot x^{m-1} + \sum_{\ell \in \mathcal{Q}} \lambda_\ell(q, x) 1_\theta(\ell) (\gamma_{q, \ell} x)^m - \sum_{\ell \in \mathcal{Q}} \lambda_\ell(q, x) 1_\theta(q) x^m. \end{aligned}$$

We plug in  $(Q_t, X_t)$  for  $(q, x)$  and take expectations for each of the three pieces separately.

(4.1)

$$\begin{aligned} &\mathbb{E}[\mathcal{L}h_\theta^{(m)}(Q_t, X_t)] \\ &= m\mathbb{E}[1_\theta(Q_t)f(Q_t, X_t)X_t^{m-1}] + \mathbb{E}\left[\sum_{\ell \in \mathcal{Q}} \lambda_\ell(Q_t, X_t)\gamma_{Q_t, \ell}^m 1_\theta(\ell)X_t^m\right] - \mathbb{E}\left[\sum_{\ell \in \mathcal{Q}} \lambda_\ell(Q_t, X_t)1_\theta(Q_t)X_t^m\right]. \end{aligned}$$

The first and third terms in the right-hand side of (4.1) are more or less straightforward, but the second term can be simplified in the following manner:

$$\begin{aligned} \mathbb{E} \left[ \sum_{\ell \in \mathcal{Q}} \lambda_{\ell}(Q_t, X_t) \gamma_{Q_t, \ell}^m 1_{\theta}(\ell) X_t^m \right] &= \mathbb{E}[\lambda_{\theta}(Q_t, X_t) \gamma_{Q_t, \theta}^m X_t^m] \\ &= \sum_{\ell \in \mathcal{Q}} \mathbb{E}[1_{\ell}(Q_t) \lambda_{\theta}(Q_t, X_t) \gamma_{\ell, \theta}^m X_t^m], \end{aligned}$$

and thus we have

$$(4.2) \quad \begin{aligned} \mathbb{E}[\mathcal{L}h_{\theta}^{(m)}(Q_t, X_t)] &= m \mathbb{E}[1_{\theta}(Q_t) f(Q_t, X_t) X_t^{m-1}] + \sum_{\ell \in \mathcal{Q}} \mathbb{E}[1_{\ell}(Q_t) \lambda_{\theta}(Q_t, X_t) \gamma_{\ell, \theta}^m X_t^m] - \sum_{\ell \in \mathcal{Q}} \mathbb{E}[\lambda_{\ell}(Q_t, X_t) 1_{\theta}(Q_t) X_t^m]. \end{aligned}$$

**Example 1.**

The general formula in (4.2) is a bit complicated to parse, therefore we first work out a particular test case. Assume that all of the functions involved are linear, i.e.,

$$f(q, x) = \alpha_q x, \quad \lambda_{\ell}(q, x) = \beta_{q, \ell} x, \quad \psi_{\ell}(q, x) = \gamma_{q, \ell} x,$$

then we obtain

$$\mathbb{E}[\mathcal{L}h_{\theta}^{(m)}(Q_t, X_t)] = m \alpha_{\theta} \mathbb{E}[1_{\theta}(Q_t) X_t^m] + \sum_{\ell \in \mathcal{Q}} \beta_{\ell, \theta} \gamma_{\ell, \theta}^m \mathbb{E}[1_{\ell}(Q_t) X_t^{m+1}] - \sum_{\ell \in \mathcal{Q}} \beta_{\theta, \ell} \mathbb{E}[1_{\theta}(Q_t) X_t^{m+1}],$$

or

$$(4.3) \quad \frac{d}{dt} \mu_{\theta}^{(m)}(t) = m \alpha_{\theta} \mu_{\theta}^{(m)}(t) + \sum_{\ell \in \mathcal{Q}} \beta_{\ell, \theta} \gamma_{\ell, \theta}^m \mu_{\ell}^{(m+1)}(t) - \sum_{\ell \in \mathcal{Q}} \beta_{\theta, \ell} \mu_{\theta}^{(m+1)}(t).$$

Compare (4.3) to (3.1) and notice that it has much the same form: the function we differentiate appears first with a positive coefficient; also, we have two terms of one higher degree with alternating signs, and the positive term has a  $\gamma$  in it. What is different, and what makes this more complicated, is that the positive term of higher degree depends on the moments in different discrete states. So, while we will be able to use a Jensen-like argument to get boundedness, we have to be more careful, since we do not know that there is any relationship between  $\mu_q^{(m)}$  and  $\mu_{q'}^{(m+1)}$  from just convexity — and thus we need to consider the entire vector  $\{\mu_q^{(m)}\}_{q \in \mathcal{Q}}$ .

**Definition 4.1.** Let  $X_t = \text{PDMP}(\mathcal{Q}, \mathcal{P}, \Lambda, \Psi, F)$ . Let  $\deg(\Lambda) = \max_{kl} \deg(\lambda_{kl})$ , and define  $\beta_{kl}$  as the coefficient of the term of degree  $\deg(\Lambda)$  in  $\lambda_{kl}(x)$ , with the convention that  $\beta_{kl} = 0$  if  $\deg(\lambda_{kl}) < \deg(\Lambda)$ . Then the **top matrix of degree**  $m$  of the system is the matrix  $M^{(m)}$  with coefficients

$$M_{kl}^{(m)} = \begin{cases} \gamma_{lk}^m \beta_{lk}, & k \neq l, \\ -\sum_{l \in \mathcal{Q}} \beta_{kl}, & k = l. \end{cases}$$

**4.2. Theorems for stability.** **Theorem 4.2 (Bounded moments).** *Assume that  $\mathcal{P} = \mathbb{R}^+$ , and let  $X_t = \text{PDMP}(\mathcal{Q}, \mathcal{P}, \Lambda, \Psi, F)$ . If  $\deg(F) \leq \deg(\Lambda)$  and all of the eigenvalues of  $M^{(m)}$  are negative, then the orbit of the total  $m$ th moment  $\mu^{(m)}$  is bounded under the moment flow equations; in particular, if  $M^{(m)}$  has negative spectrum for all  $m \geq 1$ , then all of the moments have bounded orbits.*

*Proof.* Let us first consider the case where all of the functions are linear, as in Example 1. Writing the vector  $\mu^{(m)} = \{\mu_q^{(m)}\}_q$ , we can write (4.3) as

$$(4.4) \quad \frac{d}{dt} \mu^{(m)} = mA\mu^{(m)} + M^{(m)}\mu^{(m+1)},$$

where  $A$  is the diagonal matrix with  $A_{qq} = \alpha_q$ . Note that every entry in  $A$  is positive, so the flow is linearly unstable at the origin.

However, also note that if  $\mu^{(m)} \gg 1$ , then  $\mu^{(m+1)} \gg \mu^{(m)}$  by Jensen, which means that (4.4) is dominated by the second term. More precisely, if we assume that  $\mu_q^{(m)} > 1/\epsilon$  for all  $q$ , then  $\mu_q^{(m+1)} > 1/\epsilon^{1/m} \mu_q^{(m)}$ , and thus we can write

$$(4.5) \quad \epsilon^{1/m} \frac{d}{dt} \mu^{(m)} \leq \epsilon^{1/m} A\mu^{(m)} + M^{(m)}\mu^{(m)}.$$

For  $\epsilon$  sufficiently small, the first term is dominated. Since  $M^{(m)}$  has negative spectrum, the flow  $\dot{z} = M^{(m)}z$  is such that all  $z(t)$ 's in the positive octant will asymptotically approach the origin, and, moreover, this is structurally stable to a sufficiently small perturbation by the Hartman–Grobman Theorem. Thus, for  $\epsilon$  small enough, all orbits are attracted to the origin, which means that (4.4) is inflowing on any ball of sufficiently large radius. Thus (4.4) has bounded orbits.

Now we consider the general  $f, \lambda$  (recall that we assume throughout that  $\psi_k(q, x) = \gamma_{qk}x$ ). Let us write

$$f(q, x) = \sum_{a=1}^{A_q} \alpha_{a,q} x^a, \quad \lambda_k(q, x) = \sum_{b=1}^{B_{q,k}} \beta_{b,qk} x^b.$$

Then, plugging into (4.2), we have

$$\begin{aligned} \mathbb{E}[1_\theta(Q_t) f(Q_t, X_t) X_t^{m-1}] &= \mathbb{E}[1_\theta(Q_t) \sum_{a=1}^{A_{Q_t}} \alpha_{a, Q_t} X_t^a X_t^{m-1}] \\ &= \sum_{a=1}^{A_\theta} \alpha_{a, \theta} \mathbb{E}[1_\theta(Q_t) X_t^{a+m-1}] = \sum_{a=1}^{A_\theta} \alpha_{a, \theta} \mu_\theta^{(a+m-1)}, \\ \mathbb{E}[1_\theta(\ell) \gamma_{\ell, \theta}^m \lambda_\theta(\ell, X_t) X_t^m] &= \mathbb{E} \left[ 1_\theta(\ell) \gamma_{\ell, \theta}^m \sum_{b=1}^{B_{\ell, \theta}} \beta_{b, \ell \theta} X_t^b X_t^m \right] = \sum_{b=1}^{B_{\ell, \theta}} \gamma_{\ell, \theta}^m \beta_{b, \ell \theta} \mu_\ell^{(b+m)}, \\ \mathbb{E}[\lambda_\ell(Q_t, X_t) 1_\theta(Q_t) X_t^m] &= \mathbb{E} \left[ \sum_{b=1}^{B_{Q_t, \ell}} \beta_{b, Q_t \ell} X_t^b X_t^m \right] = \sum_{b=1}^{B_{\theta, \ell}} \beta_{b, \theta \ell} \mu_\theta^{(b+m)}, \end{aligned}$$

or

$$(4.6) \quad \frac{d}{dt} \mu_\theta^{(m)} = m \sum_{a=1}^{A_\theta} \alpha_{a,\theta} \mu_\theta^{(a+m-1)} + \sum_{b=1}^{B_{\ell,\theta}} \gamma_{\ell,\theta}^m \beta_{b,\ell\theta} \mu_\ell^{(b+m)} - \sum_{b=1}^{B_{\theta,\ell}} \beta_{b,\theta\ell} \mu_\theta^{(b+m)}.$$

This means that the last two terms in (4.6) can be written as  $M^{(m)} \mu^{(m+\deg(\Lambda))}$  — note that the definition of  $M^{(m)}$  includes, by design, only those coefficients that are the same degree as the highest possible degree of all  $\lambda_{kl}$ . Thus (4.6) can be written as

$$(4.7) \quad \frac{d}{dt} \mu_\theta^{(m)} = \sum_{\ell=m-1}^{m+\deg(\Lambda)-1} A_\ell \mu^{(\ell)} + M^{(m)} \mu^{(m+\deg(\Lambda))},$$

and clearly the same sort of asymptotic analysis done for the linear case works here as well, since the first term is strictly dominated in powers by the last. ■

**Theorem 4.3 (Unbounded moments).** *Let  $\mathcal{P} = \mathbb{R}^+$  and  $X_t = \text{PDMP}(\mathcal{Q}, \mathcal{P}, \Lambda, \Psi, F)$ .*

1. *If  $\deg(F) \leq \deg(\Lambda)$ , and*
  - (a)  *$\langle M^{(m)} x, y \rangle > 0$  for all  $x, y$  in the positive octant, or*
  - (b) *there exist  $k, l$  such that  $\min(M_{kl}^{(m)}, M_{lk}^{(m)}) > \max(|M_{kk}^{(m)}|, |M_{ll}^{(m)}|)$ ;**then the orbit of the total  $m$ th moment  $\mu^{(m)}$  is unbounded under the moment flow equations;*
2. *If  $\deg(F) > \deg(\Lambda)$ , then all sufficiently large moments have unbounded orbits under the moment flow equations.*
3. *If  $\deg(F) > \deg(\Lambda) + 1$ , then all moments have unbounded orbits under the moment flow equations.*

*Proof.* Again consider (4.7). First assume that  $\langle M^{(m)} x, y \rangle > 0$  for all  $x, y$  in the positive octant. Since all of the  $A_\ell$  in that formula are diagonal with positive entries, it follows directly that the vector field is outflowing on every circle.

On the other hand, assume that there exists a  $k, l$  such that

$$\min \left( M_{kl}^{(m)}, M_{lk}^{(m)} \right) > \max \left( M_{kk}^{(m)}, M_{ll}^{(m)} \right).$$

Without loss of generality by renumbering, assume  $k = 1, l = 2$ . Let us again consider the linear case, as the nonlinear case is the same. We have

$$\begin{aligned} \frac{d}{dt} \mu_1^{(m)} &= m \alpha_1 \mu_1^{(m)} + M_{12}^{(m)} \mu_2^{(m+1)} - |M_{11}^{(m)}| \mu_1^{(m+1)}, \\ \frac{d}{dt} \mu_2^{(m)} &= m \alpha_2 \mu_2^{(m)} + M_{21}^{(m)} \mu_1^{(m+1)} - |M_{22}^{(m)}| \mu_2^{(m+1)}. \end{aligned}$$

Writing  $\nu^{(m)} = e^{-mAt} \mu^{(m)}$ , where  $A = \text{diag}(\alpha_1, \alpha_2)$ , we have

$$\begin{aligned} \frac{d}{dt} \nu_1^{(m)} &= M_{12}^{(m)} \mu_2^{(m+1)} - |M_{11}^{(m)}| \mu_1^{(m+1)}, \\ \frac{d}{dt} \nu_2^{(m)} &= M_{21}^{(m)} \mu_1^{(m+1)} - |M_{22}^{(m)}| \mu_2^{(m+1)}. \end{aligned}$$

and thus

$$\frac{d}{dt}(\nu_1^{(m)} + \nu_2^{(m)}) = (M_{21}^{(m)} - |M_{11}^{(m)}|)\mu_1^{(m+1)} + (M_{12}^{(m)} - |M_{22}^{(m)}|)\mu_2^{(m+1)} > 0,$$

so that this sum is always growing, and thus the corresponding sum for  $\mu^{(m)}$  grows at least exponentially.

Now consider the case where  $\deg(F) = \deg(\Lambda) + 1$  — this implies that the first term in (4.6) is of the same degree, than all of the other terms. Thus (4.6) looks like

$$\frac{d}{dt}\mu_m^{(m)} = A\mu^{(m+\deg(\Lambda))} + M^{(m)}\mu^{(m+\deg(\Lambda))} + O(\mu^{(m+\deg(\Lambda)-1)}),$$

and since the diagonals of  $A$  grow linearly in  $m$ , for sufficiently large  $M$ , all of the coefficients of  $A + M^{(m)}$  are positive, and thus the flow has unbounded orbits.

Finally, if  $\deg(F) > \deg(\Lambda) + 1$ , then (4.6) becomes

$$\frac{d}{dt}\mu_m^{(m)} = A\mu^{(m+\deg(F)-1)} + M^{(m)}\mu^{(m+\deg(\Lambda))} + O(\mu^{(m+\deg(F)-2)}),$$

where the dominant term is the  $A$  term, which is diagonal with positive diagonals for all  $m$ .

■

**4.3. Examples and corollaries.** We have shown above that as long as  $M^{(m)}$  has all negative eigenvalues, the  $m$ th moment is stable. First we show:

**Corollary 4.4.** *If  $0 < \gamma_{kl} < 1$  for all  $k, l$ , then  $\mathbb{E}[X_t^m]$  is bounded above for all  $t$  and for all  $m$ .*

*Proof.* From Theorem 4.2, all we need to show is that  $M^{(m)}$  has all negative eigenvalues. Recall that

$$M_{kl}^{(m)} = \begin{cases} \gamma_{lk}^m \beta_{lk}, & k \neq l, \\ -\sum_{l \in \mathcal{Q}} \beta_{kl}, & k = l. \end{cases}$$

By the Gershgorin Circle Theorem, the eigenvalues of  $M^{(m)}$  are contained in the union of the  $|\mathcal{Q}|$  balls

$$B_k := B \left( M_{kk}^{(m)}, \sum_{l=1}^{|\mathcal{Q}|} |M_{kl}^{(m)}| \right),$$

i.e., the  $k$ th ball is centered at the  $k$ th diagonal coefficient, and whose radius is given by the sum of the absolute values of the off-diagonal terms. Since  $\gamma_{kl} < 1$ , and thus  $\gamma_{kl}^m < 1$ , we have

$$\sum_{l=1}^{|\mathcal{Q}|} |\gamma_{kl}^m M_{kl}^{(m)}| < \sum_{l=1}^{|\mathcal{Q}|} |M_{kl}^{(m)}| = |M_{kk}^{(m)}|.$$

■

**Example 2.** *Let us assume that  $|\mathcal{Q}| = 2$ , so that*

$$M^{(1)} = \begin{pmatrix} -\beta_{12} & \beta_{21}\gamma_{21} \\ \beta_{12}\gamma_{12} & -\beta_{21} \end{pmatrix}.$$

We have  $\text{Tr}M^{(1)} = -\beta_{12} - \beta_{21} < 0$ , and  $\det M^{(1)} = \beta_{12}\beta_{21}(1 - \gamma_{12}\gamma_{21})$ . It is not hard to see that the resulting linear system is a saddle if  $\gamma_{12}\gamma_{21} > 1$ , and a sink if  $\gamma_{12}\gamma_{21} < 1$ . By Theorem 4.2, if  $\gamma_{12}\gamma_{21} < 1$  this means that  $\mathbb{E}[X_t]$  is bounded above. Of course, note that the stability condition for  $M^{(m)}$  is  $\gamma_{12}^m\gamma_{21}^m < 1$ , and thus stability of the first moment implies stability of all higher moments. As we see below, this is a special case only if  $|\mathcal{Q}| = 2$ .

In fact, one can get at this result from other means: notice that in the  $2 \times 2$  case, all jumps  $1 \rightarrow 2$  are immediately followed by a jump  $2 \rightarrow 1$ , and the aggregate effect of these jumps is to multiply by  $\gamma_1\gamma_2$ . Thus it is clear that the necessary and sufficient condition for stability is that this product be less than one.

One of the interesting observations made in the previous example is that we do not require that all  $\gamma_{kl}$  be less than one. But for general  $|\mathcal{Q}| > 2$ , if we choose one or more of the  $\gamma_{kl} > 1$ , then we will see some unbounded moments, as in the following example.

**Example 3.** Consider the symmetric case where we choose  $\beta_{kl} = 1$  for all  $k, l$ , and choose  $\gamma_{12} > 1$ , but  $\gamma_{13}, \gamma_{23} < 1$ , so that we have

$$M^{(m)} = \begin{pmatrix} -2 & \gamma_{12}^m & \gamma_{13}^m \\ \gamma_{12}^m & -2 & \gamma_{23}^m \\ \gamma_{13}^m & \gamma_{23}^m & -2 \end{pmatrix}.$$

If we further choose  $\gamma_{12} + \gamma_{13} < 2$  and  $\gamma_{12} + \gamma_{23} < 2$ , then again by Gershgorin theorem, it follows that  $M^{(1)}$  has all negative eigenvalues, and by Theorem 4.2, the moment flow is stable at first order.

Since  $\gamma_{12} > 1$ , then for some  $m > 1$ ,  $\gamma_{12}^m > 2$ , and by Theorem 4.3, the moment flow is unstable at  $m$ th order. Thus we have a scenario where  $\mathbb{E}[X_t]$  is bounded above for all time, but  $\mathbb{E}[X_t^m]$  grows without bound.

We saw in Example 3 that we can construct a process where the mean is bounded above, but some higher moments grow without bound. In fact, it should be clear that if there is a pair  $(k, l)$  with  $\gamma_{kl} > 1$  and  $\gamma_{lk} > 1$ , then this will generically occur: for sufficiently large  $m$ , the  $m$ th moment is unstable. This can lead to some counter-intuitive effects, as seen by the following example.

**Example 4.** Consider the case where  $|\mathcal{Q}| = N$ ,  $\gamma_{12} = \gamma_{21} > N - 1$ , and all the other  $\gamma_{kl}$  are arbitrarily small (for the purposes of this argument, set them to zero), so that if the system ever enters a state other than 1 or 2, then  $X_t$  is set to zero. Let all  $\beta_{kl} = 1$ , so that the jump rates are all the same. Specifically, this means that whenever there is a jump, the next discrete state is chosen uniformly in the others. Start with  $Q_0 = 1, X_0 = 1$ . The probability of never leaving  $Q_t \in \{1, 2\}$  after the first  $k$  jumps is then  $(N - 1)^{-k}$ , but the multiplier from these transitions is  $\gamma_{12}^k$ , so that  $\mathbb{E}[X_{T_k}]$  grows exponentially, even if  $\mathbb{P}(X_{T_k} > 0)$  is shrinking exponentially.

Similarly, we could see that if we choose  $\gamma_{12} < N - 1 < \gamma_{12}^2$ , then  $\mathbb{E}[X_t]$  would decay exponentially, but  $\mathbb{E}[X_t^2]$  would grow exponentially, and similarly for higher moments.

In short, this says that to obtain a moment instability, all one needs is two states exchanging back and forth, as long as their multipliers are large enough, even if the probability of this sequence of switching is small — which is the content of Part 1 of Theorem 4.3.

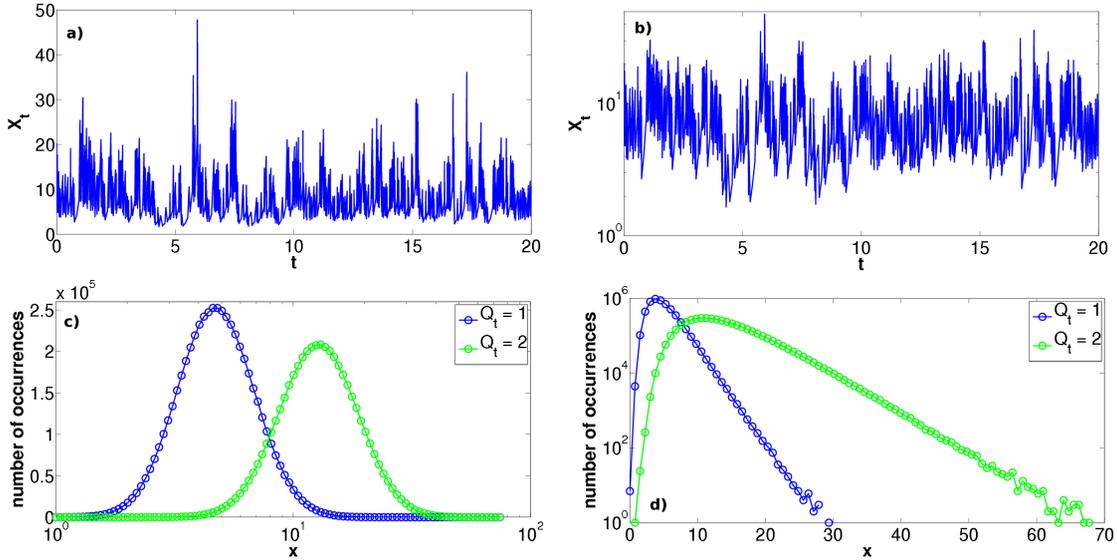
When we have a probability distribution with some moments finite and others infinite,

this is called<sup>1</sup> a **heavy-tailed** distribution (see [18, 17] for examples in dynamical systems). The canonical example of a heavy-tailed distribution is one whose tail decays asymptotically as a power law, and as such, are typically associated with critical phenomena in statistical physics [21, 44, 40, 30, 34, 31, 19]. In fact, if we assume that  $p(x)$  is the distribution of the process in equilibrium, that  $p(x) \sim x^{-\alpha}$  as  $x \rightarrow \infty$ , and  $X_\infty$  is a realization of this distribution, then

$$\mathbb{E}[X_\infty^m] = \int_0^\infty x^m p(x) dx \approx \int_{x_*}^\infty x^{m-\alpha} dx$$

will converge iff  $m - \alpha < 1$ . From this we deduce that if  $\alpha \in (m, m + 1)$ , then  $\mathbb{E}[X_\infty^m] < \infty$  but  $E[X_\infty^{m+1}] = \infty$ . We will in fact show numerical evidence of power law tails in PDMP dynamics in the next section, which leads us to conjecture that the extreme case of Example 4 above is atypical.

**4.4. Numerical results.** In Figure 2, we plot the results of a Monte Carlo simulation for a two-state PDMP. We simulate  $10^3$  realizations of this system, and each realization was integrated until  $10^4$  jumps had occurred.



**Figure 2.** Two-state case that is moment stable, with exponential tails. See text for description and explanation.

Here, we have chosen

$$f(q, x) = \alpha_q x^2, \quad \lambda_{q'}(q, x) = \beta_{q,q'} x^2,$$

<sup>1</sup>There is some disagreement in the literature of the use of the term “heavy-tailed” — some authors use the term to mean a random variable that has some polynomial moments infinite, some use it to mean a random variable with infinite variance, and yet others use it to mean a distribution whose mgf does not converge in the right-half plane. We use the first convention, and note by this convention, the log-normal distributions seen above would not be considered heavy-tailed by this convention.

where  $q, q' \in \{1, 2\}$ , and as always the resets are  $\gamma_{q,q'}x$ . We chose the  $\alpha, \beta, \gamma$  as follows:

$$\alpha = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma = \begin{pmatrix} 0 & 2 \\ 1/3 & 0 \end{pmatrix}.$$

Since  $\gamma_{12}\gamma_{21} < 1$ , by the results above all moments are stable, and this is what we observe numerically.

In Figures 2a and 2b, we plot a single trajectory of the system (in (a) we have plotted this in a linear-linear scale, and in (b) we plot the same data in a linear-log scale). We have only plotted a subset of the entire realization here; the full realization of  $10^4$  jumps goes until approximately  $t = 254$ , and here we are only plotting 812 jumps and cutting off at  $t = 20$  in order to see more structure.

In Figure 2c,d, we plot aggregate histograms of  $X_t$ , each plot uses  $0.5 \times 10^6$  datapoints. Recall that each realization of the process was run for  $10^4$  jumps; we discarded the first half of these for each realization, giving  $0.5 \times 10^3$  points, then aggregated across realizations. The first observation is that the distribution is bimodal, and this is due to the up-and-down jumps: since  $\gamma_{12} > 1$  and  $\gamma_{21} < 1$ , we expect the typical value of  $X_t$  to be much higher when  $Q_t = 2$  than when it equals one. We separate out the data by the value of  $Q_t$ . In Figure 2c, we plot on a log-linear scale, and it is pretty apparent to the eye that the distributions for  $X_t$ , conditioned on  $Q_t$ , are close to log-normal, as was the distribution in the one-state case (q.v. Figure 1). We checked this observations with QQ-plots (not presented here) and saw the same phenomenon observed in Figure 1. In panel (d), we plot a histogram on a linear-log scale, and the data shows an exponential tail, consistent with the prediction that all of the moments are uniformly bounded above.

In Figure 3, we plot the results of a Monte Carlo simulation for a three-state PDMP. Again, we simulate  $10^3$  realizations of this system, and each realization was integrated until  $10^4$  jumps had occurred.

Here we have chosen

$$f(q, x) = \alpha_q x^2, \quad \lambda_{q'}(q, x) = \beta_{q,q'} x^2,$$

where  $q, q' \in \{1, 2, 3\}$ , and as always the resets are  $\gamma_{q,q'}x$ . We chose the  $\alpha, \beta, \gamma$  as follows:

$$\alpha = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad \gamma = \begin{pmatrix} 0 & 1.3 & 0.3 \\ 1.3 & 0 & 0.2 \\ 0.3 & 0.2 & 0 \end{pmatrix}.$$

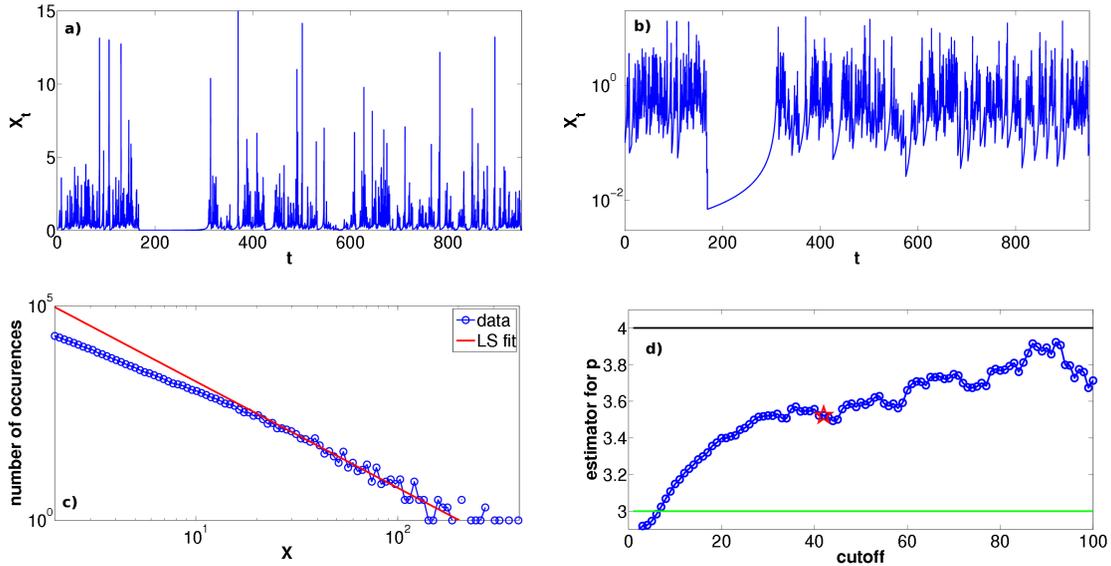
This means that

$$M^{(m)} = \begin{pmatrix} -2 & (1.3)^m & (0.3)^m \\ (1.3)^m & -2 & (0.2)^m \\ (0.3)^m & (0.2)^m & -2 \end{pmatrix},$$

and one can see that

$$(1.3)^2 + (0.3)^2 < 2, \quad (1.3)^3 > 2,$$

so by the theorems above, we have that the first two moments are stable and the third is not — thus as  $t \rightarrow \infty$ ,  $\mathbb{E}[X_t] < \infty$  and  $\mathbb{E}[X_t^2] < \infty$ , but  $\mathbb{E}[X_t^3]$  grows without bound. Thus, if



**Figure 3.** Three-state case that is marginally moment stable. See text for description and explanation.

we see a power-law tail in the distribution, we expect it to decay somewhere as  $x^{-p}$  with  $p$  between 3 and 4.

In Figure 3(a,b), we again plot a single trajectory of the system (in (a) we have plotted this in a linear-linear scale, and in (b) we plot the same data in a linear-log scale). We see the wide variety of spatial scales in the linear plot, and the “intermittent” structure of the process in the log case — sometimes the system can be kicked quite low, and takes a while to climb out of it. Recall that the ODE is quadratically nonlinear, so if  $X_t$  ever happens to become small, it will take a long time to leave a neighborhood of zero.

In Figure 3c, we plot an aggregate histogram of  $0.5 \times 10^6$  datapoints on a log-log scale. Recall that each realization of the process was run for  $10^4$  jumps; we discarded the first half of these for each realization, giving  $0.5 \times 10^3$  points, then aggregated across realizations. We see from eye that it looks to have a power law structure for large  $x$ , and we also plot the least squares linear fit to the higher half of the data. This slope is given as  $p = -2.4797 \pm 0.0112$ , which seems to contradict our assessment from above. However, it should be pointed out that trying to match a power law fit by least squares on a log-log plot gives in general a bad estimator [35, 11], so we reevaluate our analysis of the decay.

We use the procedure laid out in [11] as follows: given a data set  $\{x_i\}$ , first decide a cutoff  $x_{\min}$ , let  $n = \#\{x_i > x_{\min}\}$ , and this gives the estimator

$$\hat{p} = 1 + \frac{n}{\sum_{i=1}^n \log(x_i/x_{\min}) \mathbf{1}(x_i > x_{\min})}.$$

In Figure 3d, we plot  $\hat{p}$  for all cutoffs in the range  $[1, 100]$ , and we see that as long as we use a cutoff of about 10 or more, this is consistent with the theoretical prediction. [The horizontal lines in Figure 3(d) are the bounds given by the analysis.] We plot by a red star the value that is given by the “best” estimator, determined in the following manner: for each choice of  $x_{\min}$ ,

we compare the empirical data with the theoretical distribution assuming that our estimate of  $\hat{p}$ , then measure the closeness of these distributions using the Kolmogorov-Smirnoff (KS) distance. The best fit was given by a cutoff of 42, with an estimator of  $\hat{p} = 3.523$ , and a KS distance between the theoretical and empirical distributions of  $7.78 \times 10^{-3}$ .

**5. Finite-time blowups.** In this section, we will again return to the one-state case considered in Section 3. It was shown there that as long as  $\deg(\lambda) \geq \deg(f)$ , the stochastic process is “well-behaved”, i.e. as  $t \rightarrow \infty$ , all of the moments are finite. In particular, it is not possible for the process to escape to infinity in finite time.

However, in general, if the driving ODE is nonlinear, then it is certainly possible that the process escapes to infinity in finite time (these events are usually called **finite-time blowups** in the dynamical systems literature, or **explosions** in the stochastic process literature). We examine these behaviors in this section.

**Definition 5.1.** *Let us denote by  $T_n$  the times where the PDMP has jumps. We say that a realization of an PDMP has a **finite-time blowup** if either of two conditions hold:*

I.  $T_\infty := \lim_{n \rightarrow \infty} T_n < \infty$ ,

II. *the solution of the continuous part becomes infinite between any two jump times.*

*We will refer to the two types of blowup as Type I and Type II.*

**Remark 6.** *What we are calling a “Type I” blowup is the type of behavior that is typically called an explosion for stochastic processes. These are common for countable Markov chains where the jump rates can grow sufficiently fast. What we call a “Type II” blowup is very much like what is called a finite-time blowup in the differential equations literature — in this context, since the jump never occurs, it is the same as an ODE going to infinity.*

The results of this section can be summarized as follows: let  $X_t = \text{PDMP}(f, \lambda, \gamma)$  with  $\deg(f) > 1$ ; then, we have the following two cases:

- C1. If  $\deg(\lambda) < \deg(f) - 1$ , then  $X_t$  will have a Type II blowup with probability one for any initial condition;
- C2. if  $\deg(\lambda) = \deg(f) - 1$ , then, depending on the leading coefficients of  $f$  and  $\lambda$ , the system exhibits various behaviors, summarized in Theorem 5.3 below. In particular, in some parameter regimes it exhibits Type I blowups almost surely, and the remaining regimes it does not exhibit Type I blowups almost surely. If  $\deg(f) = 1$ , then the system cannot blowup, since linear flows are well-defined for all time. However, as we show below, in the parameter range where the nonlinear systems exhibit finite-time blowups, the linear system can still go to infinity but take an infinite amount of time to do so, exhibiting in some sense an “infinite-time blowup”.

We consider these two cases in the subsections below.

**5.1. Case C1.** We first consider the case where  $\deg(\lambda) < \deg(f) - 1$ . We show here that any such system has a Type II blowup — basically, the ODE goes to infinity at too rapid a rate for the jump rates to catch up.

**Theorem 5.2.** *Let  $X_t = \text{PDMP}(f, \lambda, \gamma)$ , and assume that  $\deg(\lambda) < \deg(f) - 1$ . Then  $X_t$  has a Type II blowup with probability one for any initial condition.*

*Proof.* Using standard ODE arguments, if

$$\frac{d}{dt}\varphi^t(x) = f(\varphi^t(x)),$$

where  $f(x)$  is a positive polynomial of degree  $\deg(f) > 1$ , then  $\varphi^t(x)$  has a finite-time singularity at some  $t^* < \infty$ , and, moreover, in a (left) neighborhood of this point, we have

$$\varphi^t(x) \sim C(t - t^*)^{1/(1-\deg(f))}.$$

If  $\deg(\lambda) < \deg(f) - 1$ , then  $0 > \deg(\lambda)/(1 - \deg(f)) > -1$ . Writing  $\lambda(x) = \beta x^{\deg(\lambda)} + O(x^{\deg(\lambda)-1})$ , we have that  $\lambda(\varphi^s(x))$  is integrable near this singular point, i.e.

$$\int_{t_0}^{t^*} \lambda(\varphi^s(x)) ds < \infty.$$

Since any exponential has support on the whole real axis, this and (2.1) means that starting at any initial condition, the probability of the ODE having a singularity before jumps is positive, i.e.

$$\mathbb{P} \left( S_n > \int_{t_0}^{t^*} \beta(\varphi^s(x)) ds \right) > 0,$$

and the system has a Type II blowup with positive probability. Since the process is Markov, irreducible on  $\mathbb{R}^+$ , and aperiodic, it follows that the system has a Type II blowup with probability one. ■

**5.2. Case C2.** This case is much more complicated than the previous one for various reasons, not the least of which being that the behavior depends on the coefficients of the functions  $f$  and  $\lambda$ . The behavior is summarized in the following theorem:

**Theorem 5.3.** *Let  $X_t = \text{PDMP}(f, \lambda, \gamma)$ , where*

$$(5.1) \quad f(x) = \alpha x^{k+1} + O(x^k), \quad \lambda(x) = \beta x^k + O(x^{k-1}),$$

$k \geq 1$  (so that the ODE is superlinear). Then  $X_t$  will not have Type II blowups.

**Case 1.** *Assume in addition to (5.1) that*

$$(5.2) \quad f(x) \geq \alpha x^{k+1}, \quad \lambda(x) \leq \beta x^k \text{ for all } x.$$

Then

1. if  $\gamma > e^{-\alpha/\beta}$ , then  $X_t$  has Type I blowups almost surely;
2. if  $\gamma < 1 - \alpha/\beta$ , then  $X_t \rightarrow 0$ , both in  $L^1$  and almost surely.

**Case 2.** *If we assume only (5.1) and not (5.2), then:*

1. if  $\gamma > e^{-\alpha/\beta}$ , then there exists  $M > 0$  such that if  $X_0 > M$ , then  $X_t$  has Type I blowups with positive probability;
2. if  $\gamma < 1 - \alpha/\beta$ , there exists  $M > 0$  such that if  $X_0 < M$ , then  $X_t \rightarrow 0$  with positive probability.

Finally, if we assume a linear ODE, i.e.  $X_t = \text{PDMP}(\alpha x, \beta, \gamma)$ , then the conclusions of Case 1 are true except for the modification of #1, in which case we have that  $X_t \rightarrow \infty$  a.s., but  $X_t < \infty$  w.p. 1 for any finite  $t$ .

The way we proceed to prove this is as follows. We first compute a recurrence relation when  $f, \lambda$  are assumed to be pure monomials, and then show that this recurrence relation has growth properties that correspond to the three types of convergence above, in the appropriate

parameter regimes. Finally, we prove a monotonicity theorem that allows us to extend the calculation for monomials to general polynomials.

**Proposition 5.4.** *If  $X_t = \text{PDMP}(f, \lambda, \gamma)$  with  $\deg(\lambda) = \deg(f) - 1$ , then the probability of a Type II blowup is zero for any initial condition.*

*Proof.* Since  $\deg(\lambda) = \deg(f) - 1$ , then  $\lambda(\varphi^t(x)) \sim 1/t$  in a neighborhood of the singularity (see the beginning of the proof of Proposition 5.2 for more detail), so that

$$\int_{t_0}^{t^*} \lambda(\varphi^s(x)) ds = \infty,$$

and a Type II blowup is not possible. ■

**Lemma 5.5.** *Let  $X_t = \text{PDMP}(f, \lambda, \gamma)$  with  $f(x) = \alpha x^{k+1}$  and  $\lambda(x) = \beta x^k$ . Then*

$$X_{T_n} = \gamma e^{(\alpha/\beta)S_n} X_{T_{n-1}},$$

where  $S_n$  are iid unit-rate exponential random variables.

*Proof.* It is not hard to check that

$$\varphi^t(x) = \frac{x}{(1 - \alpha k t x^k)^{1/k}},$$

and thus

$$\lambda(\varphi^t(x)) = \frac{\beta x^k}{1 - \alpha k t x^k}.$$

Therefore, we have

$$\begin{aligned} S_n &= \int_0^{T_{n+1}-T_n} \lambda(\varphi^s(X_{T_n})) ds = \int_0^{T_{n+1}-T_n} \frac{\beta (X_{T_n})^k}{1 - \alpha k s (X_{T_n})^k} ds \\ &= \frac{-\beta}{\alpha k} \ln \left| 1 - \alpha k s (X_{T_n})^k \right| \Big|_{s=0}^{s=T_{n+1}-T_n} = \frac{-\beta}{\alpha k} \ln \left| 1 - \alpha k (T_{n+1} - T_n) (X_{T_n})^k \right|. \end{aligned}$$

From this, one can see that

$$(5.3) \quad 1 - \alpha k (T_{n+1} - T_n) (X_{T_n})^k = e^{-\alpha/(\beta k) S_n}.$$

We then compute

$$(5.4) \quad \begin{aligned} X_{T_{n+1}} &= \gamma \varphi^{T_{n+1}-T_n}(X_{T_n}) \\ &= \gamma \frac{X_{T_n}}{(1 - \alpha k (T_{n+1} - T_n) (X_{T_n})^k)^{1/k}} = \gamma e^{\frac{\alpha}{\beta} S_n} X_{T_n}. \end{aligned}$$

■

**Lemma 5.6.** *Let  $S_k$  be iid unit rate exponentials and define*

$$Z_n = \prod_{k=1}^n \gamma e^{(\alpha/\beta) S_k}.$$

Then:

1. if  $\gamma > e^{-\alpha/\beta}$ , then there exists a  $\delta' > 0$  such that  $\liminf_n e^{-\delta'n} Z_n = \infty$  with probability one;
2. if  $1 - \alpha/\beta < \gamma < e^{-\alpha/\beta}$ , then  $Z_n \rightarrow 0$  in probability, even though  $\mathbb{E}[Z_n] \rightarrow \infty$  as  $t \rightarrow \infty$ ;
3. if  $\gamma < 1 - \alpha/\beta$ , then  $Z_n \rightarrow 0$ , both in  $L^1$  and almost surely.

*Proof.* Let us write  $W_k = \gamma e^{(\alpha/\beta)S_k}$ . Let us write  $\mu = \mathbb{E}[W_k]$  and  $\delta = \mathbb{E}[\log(W_k)]$ . We compute:

$$\mu = \gamma \mathbb{E}[e^{(\alpha/\beta)S_k}] = \gamma \int_0^t e^{(\alpha/\beta)t} e^{-t} dt = \frac{\gamma}{1 - \alpha/\beta}.$$

We also have

$$\delta = \mathbb{E} \left[ \frac{\alpha}{\beta} S_k + \log \gamma \right] = \frac{\alpha}{\beta} + \log \gamma.$$

Note then that the three conditions in this lemma correspond to  $\delta > 0$ ;  $\delta < 0$  and  $\mu > 1$ ; and  $\mu < 1$ , respectively. (By Jensen's inequality, we have

$$\mu = \mathbb{E}[W_k] = \mathbb{E}[e^{\log(W_k)}] > e^{\mathbb{E}[\log(W_k)]} = e^\delta,$$

so clearly  $\delta > 0$  implies  $\mu > 1$ , and these are the only three possibilities.)

First, we compute

$$\mathbb{E}[Z_n] = \mathbb{E} \left[ \prod_{k=1}^n W_k \right] = \prod_{k=1}^n \mathbb{E}[W_k] = \mu^n,$$

so clearly the expectation goes to zero (resp.  $\infty$ ) if  $\mu < 1$  (resp.  $\mu > 1$ ). Since  $Z_n \geq 0$  by definition, this implies that  $Z_n \rightarrow 0$  both in  $L^1$  and almost surely. This establishes claim #3 of the lemma.

Next, note that

$$Z_n = \exp(\log(Z_n)) = \exp \left( \sum_{k=1}^n \log W_k \right) = \exp \left( \frac{\alpha}{\beta} \sum_{k=1}^n S_k + n \log \gamma \right).$$

From the Law of Large Numbers,

$$\sum_{k=1}^n \log W_k = n \mathbb{E}[\log W_k] \pm O(\sqrt{n}) = n \left( \frac{\alpha}{\beta} + \log \gamma \right) \pm O(\sqrt{n}),$$

and if  $\delta \neq 0$ , this sum goes to  $\pm\infty$  depending on the sign of  $\delta$ . More specifically, we can use Chernoff-type bounds [41, 14]: let  $\delta > 0$ . Since

$$\mathbb{E} \left[ \sum_{k=1}^n \log W_k \right] = n\delta,$$

the Chernoff bounds give

$$\mathbb{P} \left( \sum_{k=1}^n \log W_k < n\delta/2 \right) < e^{-n\delta/8},$$

or

$$(5.5) \quad \mathbb{P}\left(Z_n < e^{n\delta/2}\right) < e^{-n\delta/8}.$$

In particular, this means that  $Z_n$  grows faster than an exponential with probability exponentially close to one. For example, choose  $0 < \delta' < \delta/2$ . Now, if it is not true that  $\liminf e^{-\delta'n} Z_n = \infty$ , this means that there exists  $M > 0$  and an infinite subsequence  $\{n_k\}$  such that  $e^{-\delta'n_k} Z_{n_k} \leq M$ . From (5.5), this event has probability zero. This proves claim #1 of the lemma.

Conversely, if  $\delta < 0$ , then

$$\mathbb{P}\left(Z_n > e^{n\delta/2}\right) < e^{-n\delta/12},$$

or the infinite product goes to zero exponentially fast with probability exponentially close to one, which implies that  $Z_n \rightarrow 0$  in probability. Thus, if  $\delta < 0$  and  $\mu > 1$ , then we have that  $Z_n \rightarrow 0$  in probability but  $\mathbb{E}[Z_n] \rightarrow \infty$ , proving claim #2 of the theorem. ■

**Lemma 5.7.** *Let  $X_t = \text{PDMP}(f, \lambda, \gamma)$  and  $Y_t = \text{PDMP}(g, \mu, \gamma)$ , and assume that*

$$(5.6) \quad X_0 \leq Y_0, \quad g(x) \geq f(x), \quad \mu(x) \leq \lambda(x).$$

*If  $X_t(\omega)$  has a finite-time blowup, then so does  $Y_t(\omega)$ . Thus, the probability of  $Y_t$  having a finite-time blowup is at least as large as  $X_t$  having one, and so, for example, if  $X_t$  has an a.s. finite-time blowup, then so does  $Y_t$ . Conversely, if  $Y_t(\omega) \rightarrow 0$  as  $t \rightarrow \infty$ , then so does  $X_t(\omega)$ , and thus the probability of  $X_t$  decaying to zero is at least as large as the probability that  $Y_t$  does.*

*Proof.* We denote  $T_n^{(X)}$  as the  $n$ th reset time for process  $X_t$ , and similarly for  $T_n^{(Y)}$ . Since  $\mu \leq \lambda$ ,  $T_1^{(Y)} \geq T_1^{(X)}$ . Together with the fact that  $g \geq f$ , this implies that  $Y_{T_1^{(Y)}} \geq X_{T_1^{(X)}}$ . Using induction, the random sequence  $Y_{T_n^{(Y)}}$  dominates the sequence  $X_{T_n^{(X)}}$  for any  $\omega$ , and the conclusions follow. ■

**Remark 7.** *Said in words: if we make the vector field larger, or the rate smaller, then the system is more likely to blow up. Conversely, if we make the vector field smaller, or the rate larger, the system is more likely to go to zero.*

**Lemma 5.8.** *Let  $X_t = \text{PDMP}(f, \lambda, \gamma)$  and  $Y_t = \text{PDMP}(g, \mu, \gamma)$ , and assume that there is an  $M > 0$  such that*

$$(5.7) \quad g(x) \geq f(x), \quad \mu(x) \leq \lambda(x) \text{ for all } x > M.$$

*Conditioning on  $M < X_0 < Y_0$ , if  $X_t$  has a finite-time blowup with positive probability, then  $Y_t$  also has a finite-time blowup with positive probability.*

*Proof.* Assume that  $X_t$  has a finite-time blowup with positive probability. Let us define

$$\Omega_1 = \{\omega | (X_t(\omega) \text{ has finite-time blowup}) \wedge (X_t(\omega) > M \text{ for all } t \geq 0)\}.$$

Clearly, if  $\omega \in \Omega_1$  and  $Y_0 > X_0 > M$ , then  $Y_t(\omega)$  has a finite-time blowup. Now, let us assume that  $\mathbb{P}(\Omega_1) = 0$ , i.e. for almost every  $\omega$  such that  $X_t(\omega)$  has a finite-time blowup,

we have  $X_t(\omega) < M$  for some  $t$ . This would mean that the set  $(-\infty, M)$  is recurrent for  $X_t$ , contradicting the fact that  $X_t$  has a finite-time blowup with positive probability. ■

**Lemma 5.9.** *Let  $X_t = \text{PDMP}(f, \lambda, \gamma)$  and  $Y_t = \text{PDMP}(g, \mu, \gamma)$ , and assume that there is an  $M > 0$  such that*

$$(5.8) \quad g(x) \geq f(x), \quad \mu(x) \leq \lambda(x) \text{ for all } x < M.$$

*Conditioning on  $X_0 < Y_0 < M$ , if  $Y_t \rightarrow 0$  with positive probability, then  $X_t \rightarrow 0$  also with positive probability.*

(The proof is the same argument as in Lemma 5.8.)

**Proof of Theorem 5.3.**

Let us first consider the case where  $f$  and  $\lambda$  are pure monomials, i.e.,  $f(x) = \alpha x^{k+1}$ ,  $\lambda(x) = \beta x^k$ . If  $\gamma > e^{-\alpha/\beta}$ , Using Lemmas 5.5 and 5.6, we know that there exists  $\delta' > 0$  such that  $\liminf_n e^{-\delta'n} X_{T_n} = \infty$  with probability one. This means that  $X_{T_n}$  is blowing up at least exponentially fast as a function of  $n$ . Moreover, using (5.3), we can solve

$$T_{n+1} - T_n = \frac{1 - e^{-(\alpha/\beta k)S_n}}{\alpha k (X_{T_n})^k} \leq C e^{-\delta'kn},$$

and if  $k > 1$ , the telescoping sum is summable and  $T_\infty < \infty$ . For the case of  $\gamma < 1 - \alpha/\beta$ , the statement follows directly from Lemma 5.6.

(If  $k = 1$ , then the recurrence in (5.3) still holds, but now we have  $T_{n+1} - T_n = S_n/\beta$ , and it is straightforward to see that the telescoping sum is not summable w.p. 1, so that  $T_\infty = \infty$ .)

**Case 1.** Lemma 5.7 gives us both statements #1 and #2, since the monomials have this behavior.

**Case 2.** If  $f(x) = \alpha x^{k+1} + O(x^k)$ , and  $\lambda(x) = \beta x^k + O(x^{k-1})$ , then for any  $\alpha' < \alpha$  and  $\beta' > \beta$ , and for some  $M$  sufficiently large,  $f(x) \geq \alpha' x^{k+1}$  and  $\lambda(x) \leq \beta' x^k$ . If  $\gamma > e^{-\alpha'/\beta'}$ , then we can choose  $\alpha' < \alpha, \beta' > \beta$  with  $\gamma > e^{-\alpha'/\beta'}$  as well. Then using Lemma 5.8 establishes statement #1 of Case 2, and Lemma 5.9 establishes Statement #2 of Case 2. ■

Finally, there exists a strengthening of the conclusions of Theorem 5.3 when the vector field and rates are pure monomials:

**Corollary 5.10.** *Let  $X_t = \text{PDMP}(f, \lambda, \gamma)$ , where*

$$(5.9) \quad f(x) = \alpha x^{k+1}, \quad \lambda(x) = \beta x^k,$$

$k \geq 1$  (so that the ODE is superlinear). Then:

1. if  $\gamma > e^{-\alpha/\beta}$ , then  $X_t$  has Type I blowups almost surely;
2. if  $1 - \alpha/\beta < \gamma < e^{-\alpha/\beta}$ , then  $X_t \rightarrow 0$  in probability, even though  $\mathbb{E}[X_t] \rightarrow \infty$  as  $t \rightarrow \infty$ ;
3. if  $\gamma < 1 - \alpha/\beta$ , then  $X_t \rightarrow 0$ , both in  $L^1$  and almost surely.

*Proof.* This follows from Lemma 5.5 and 5.6. ■

**5.3. Numerical results.** In Figure 4 we plot the results of several Monte Carlo simulations for blowups. These simulations required a technique more sophisticated than the Gillespie–Euler method used earlier. We are attempting to simulate a system where nonlinear ODEs are blowing up, meaning that the vector fields get large and will be very sensitive to discretization

errors. We implemented a two-phase method as follows: if  $X_t < 10$ , then we implemented a Euler–Gillespie method as described before, with  $\Delta t_{\max} = 10^{-2}$ . When  $X_t \geq 10$ , we switched to a shooting method to obtain the next stopping time. At any  $X_t$ , we compute the time of singularity that would occur if there were no jumps, call this  $t_{\text{right}}$ , and define  $t_{\text{left}} = 0$ . Choose  $S_{n+1}$  as an exponential, and from this we can either determine if we will have a Type II blowup before the next jump or not. If not, we then use a bisection method to find  $T_{n+1}$ : at any stage in the process, we choose the midpoint of  $[t_{\text{left}}, t_{\text{right}}]$  and determine whether the integral in (2.1) is larger or smaller than  $S_{n+1}$ ; if smaller, we set  $t_{\text{left}}$  to be this midpoint, and if larger, we set  $t_{\text{right}}$  to be this midpoint. This guarantees an exponential convergence to  $T_{n+1}$ , and from this we can compute all of the other quantities of interest. Finally, we always truncated whenever the system passed  $10^9 \approx \exp(20.7)$  — any time  $X_t > 10^9$ , we halted the computation.

In Figure 4a, we plot  $\log(X_t)$  for a single realization of the PDMP where we have chosen  $f(x) = x^3$ ,  $\lambda(x) = 2x^2$ , and  $\gamma = 0.75$ . One can see that there is a finite-time blowup, and in fact we see that there are many jumps happening in a very short time, as the trajectory goes off to infinity. This comports with the prediction of a Type I blowup. In Figure 4b, we plot  $\log(X_t)$  for a single realization of the PDMP where we have chosen  $f(x) = x^4$ ,  $\lambda(x) = 2x^2$ , and  $\gamma = 0.5$ . One can see that there is a finite-time blowup here as well, but there are only a few jumps on the way to infinity; this comports with the prediction of a Type II blowup.

In Figures 4(c) and 4(d), we present the results of a family of simulations for Type I blowups. Here we run each simulation for  $10^4$  steps, or it has a finite-time blowup, and for each value of  $\gamma$  we computed  $10^2$  realizations. We chose  $f(x) = x^3$ ,  $\lambda(x) = 2x^2$  throughout, but vary  $\gamma$  in the range  $0.1, \dots, 0.9$ . According to Theorem 5.3, the critical values of  $\gamma$  are  $1/2$  and  $e^{-1/2} \approx 0.6061$  — in both figures we have put vertical red lines at these values.

In both Figure 4(c) and 4(d) we use the plotting convention of plotting each simulation with a light blue small circle, then plot the mean and standard deviation for all  $10^2$  realizations for each gamma value in dark blue with a circle at the mean and an error bar for the standard deviation. In (c), we plot the logarithm of the empirical mean of  $X_t$  over the entire simulation, recalling that we are truncating any trajectory that passes  $10^9 \approx \exp(20.7)$ . Thus observations near or exceeding 20 are all blowups. We see that for  $\gamma < 1/2$ , all realizations stay small throughout the simulation. In the range  $[1/2, e^{-1/2}]$ , there is a spread of values depending on realization, and past  $e^{-1/2}$  the blowups dominate. One can see this more starkly in Figure 4d: here we have plotted the final time of the simulation — note that we run all simulations for  $10^4$  steps, and the maximum threshold for the naïve method is  $\Delta t_{\max} = 10^{-2}$  — so if the trajectory always stays small, we would expect a total simulation time very close to  $10^4 \cdot 10^{-2} = 100$ . Thus we can interpret a final simulation time near 100 as a proxy for a blowup not occurring; conversely, if the simulation truncates significantly earlier than  $t = 100$ , this is a sign that the system has had a finite-time blowup, and we see this clearly for large  $\gamma$ .

**6. Acknowledgments.** This work was supported in part by the National Science Foundation (NSF) under awards ECCS-CAR-0954420, CMG-0934491, and CyberSEES-1442686, the Trustworthy Cyber Infrastructure for the Power Grid (TCIPG) under US Department of Energy Award DE-OE0000097, by the National Aeronautics and Space Administration (NASA) through the NASA Astrobiology Institute under Cooperative Agreement Number

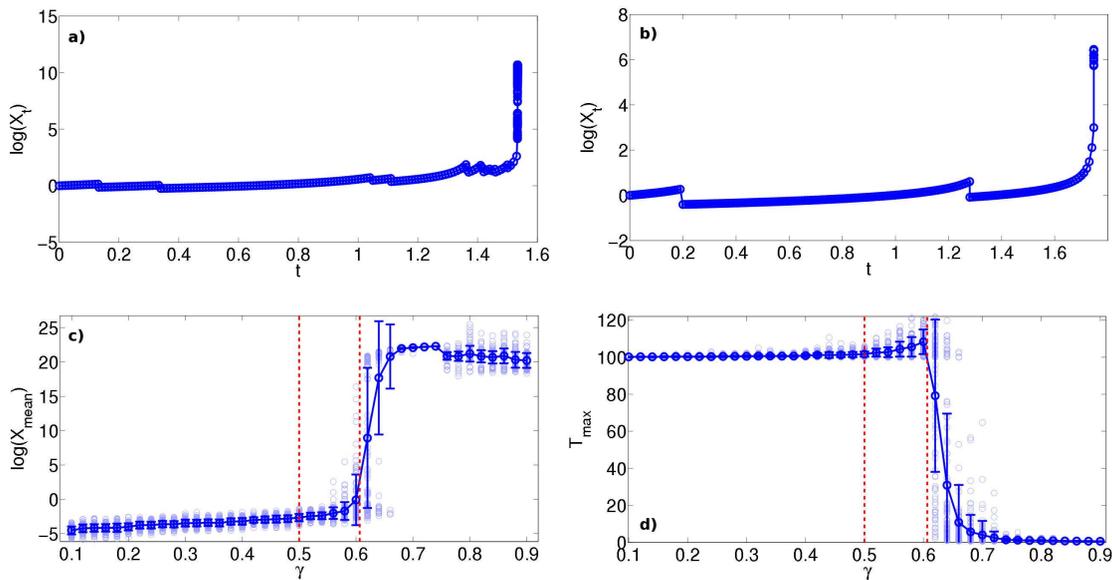


Figure 4. Results of simulations for PDMP undergoing blowups; see text for more detailed description

NNA13AA91A issued through the Science Mission Directorate, and the Initiative for Mathematical Science and Engineering (IMSE) at the University of Illinois.

### REFERENCES

- [1] DUARTE ANTUNES, JOÃO PEDRO HESPANHA, AND CARLOS SILVESTRE, *Stochastic hybrid systems with renewal transitions: Moment analysis with applications to networked control systems with delays*, SIAM Journal of Control and Optimization, 51 (2013), pp. 1481–1499.
- [2] YURI BAKHTIN AND TOBIAS HURTH, *Invariant densities for dynamical systems with random switching*, Nonlinearity, 25 (2012), p. 2937.
- [3] PETER H. BAXENDALE, *Invariant measures for Nonlinear Stochastic Differential Equations*, in Lyapunov Exponents, Springer, 1991, pp. 123–140.
- [4] ———, *Stability and equilibrium properties of stochastic flows of diffeomorphisms*, in Diffusion Processes and Related Problems in Analysis, Volume II, Springer, 1992, pp. 3–35.
- [5] RICHARD BELLMAN, *Limit theorems for non-commutative operations. I.*, Duke Mathematical Journal, 21 (1954), pp. 491–500.
- [6] PATRICK BILLINGSLEY, *Probability and measure*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, third ed., 1995. A Wiley-Interscience Publication.
- [7] PAUL C. BRESSLOFF AND JAY M. NEWBY, *Path integrals and large deviations in stochastic hybrid systems*, Physical Review E, 89 (2014), p. 042701.
- [8] JARED C BRONSKI AND RICHARD M MCLAUGHLIN, *The problem of moments and the majda model for scalar intermittency*, Physics Letters A, 265 (2000), pp. 257–263.
- [9] DEBASISH CHATTERJEE AND DANIEL LIBERZON, *On stability of stochastic switched systems*, Proceedings of the 43rd Conference on Decision and Control, (2004).
- [10] ———, *Stability analysis and stabilization of randomly switched systems*, in Decision and Control, 2006 45th IEEE Conference on, IEEE, 2006, pp. 2643–2648.
- [11] AARON CLAUSET, COSMA ROHILLA SHALIZI, AND MARK NEWMAN, *Power-law distributions in empirical data*, SIAM review, 51 (2009), pp. 661–703.

- [12] JAMES COLLIANDER, MARKUS KEEL, GIGIOLA STAFFILANI, HIDEO TAKAOKA, AND TERENCE TAO, *Transfer of energy to high frequencies in the cubic defocusing nonlinear schrödinger equation*, *Inventiones mathematicae*, 181 (2010), pp. 39–113.
- [13] MARK H. A. DAVIS, *Markov Models and Optimization*, Chapman and Hall, London, UK, 1993.
- [14] AMIR DEMBO AND OFER ZEITOUNI, *Large deviations techniques and applications*, vol. 38 of *Stochastic Modelling and Applied Probability*, Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.
- [15] SAIRAJ DHOPLE, Y. CHRISTINE CHEN, LEE DEVILLE, AND ALEJANDRO DOMÍNGUEZ-GARCÍA, *Analysis of power system dynamics subject to stochastic power injections*, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 60 (2013), pp. 3341–3353.
- [16] SAIRAJ DHOPLE, LEE DEVILLE, AND ALEJANDRO DOMÍNGUEZ-GARCÍA, *A Stochastic Hybrid Systems framework for analysis of Markov reward models*, *Reliability Engineering & System Safety*, 123 (2014), pp. 158–170.
- [17] S. N. DOROGOVTSSEV, A. V. GOLTSEV, AND J. F. F. MENDES, *Critical phenomena in complex networks*, *Reviews of Modern Physics*, (2008), p. 79.
- [18] MARK EMBREE AND LN TREFETHEN, *Growth and decay of random Fibonacci sequences*, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences . . .*, 455 (1999), pp. 2471–2485.
- [19] NIR FRIEDMAN, SHINYA ITO, BRADEN A. W. BRINKMAN, MASANORI SHIMONO, R. E. LEE DEVILLE, KARIN A DAHMEN, JOHN M. BEGGS, AND THOMAS C. BUTLER, *Universal critical dynamics in high resolution neuronal avalanche data*, *Physical Review Letters*, 108 (2012), p. 208102.
- [20] J. N. FRY, *The four-point function in the BBGKY hierarchy*, *The Astrophysical Journal*, 262 (1982), pp. 424–431.
- [21] NIGEL GOLDENFELD, *Lectures on phase transitions and the renormalization group*, Addison-Wesley, Advanced Book Program, Reading, 1992.
- [22] THILO GROSS, CARLOS J. DOMMAR D’ LIMA, AND BERND BLASIUS, *Epidemic dynamics on an adaptive network*, *Physical Review Letters*, 96 (2006), p. 208701.
- [23] THILO GROSS AND IOANNIS G. KEVREKIDIS, *Robust oscillations in SIS epidemics on adaptive networks: Coarse graining by automated moment closure*, *EPL (Europhysics Letters)*, (2008), p. 6.
- [24] JOÃO PEDRO HESPANHA, *A Model for Stochastic Hybrid Systems with Application to Communication Networks*, *Nonlinear Analysis, Special Issue on Hybrid Systems*, 62 (2005), pp. 1353–1383.
- [25] ———, *Modelling and analysis of stochastic hybrid systems*, *IEE Proceedings—Control Theory and Applications*, 153 (2006), pp. 520–535.
- [26] JOÃO PEDRO HESPANHA AND ABHYUDAI SINGH, *Stochastic Models for Chemically Reacting Systems Using Polynomial Stochastic Hybrid Systems*, *International Journal on Robust Control, Special Issue on Control at Small Scales: Issue 1*, 15 (2005), pp. 669–689.
- [27] MATTHEW J. KEELING, *The effects of local spatial structure on epidemiological invasions*, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266 (1999), pp. 859–867.
- [28] ———, *Spatial models of interacting populations*, Blackwell Science, Oxford, UK, 1999.
- [29] RAFAIL KHASHMINSKIL, *Stochastic stability of differential equations*, vol. 66, Springer, 2011.
- [30] OSAME KINOCHI AND MAURO COPELLI, *Optimal dynamical range of excitable networks at criticality*, *Nature Physics*, 2 (2006), pp. 348–351.
- [31] DANIEL LARREMORE, WOODROW SHEW, AND JUAN RESTREPO, *Predicting Criticality and Dynamic Range in Complex Networks: Effects of Topology*, *Physical Review Letters*, 106 (2011), pp. 1–4.
- [32] SEAN MEYN AND RICHARD L. TWEEDIE, *Markov Chains and Stochastic Stability*, Cambridge University Press, New York, NY, USA, 2nd ed., 2009.
- [33] PETER MILNER, COLIN S GILLESPIE, AND DARREN J WILKINSON, *Moment closure approximations for stochastic kinetic models with rational rate laws.*, *Mathematical biosciences*, 231 (2011), pp. 99–104.
- [34] THIERRY MORA AND WILLIAM BIALEK, *Are Biological Systems Poised at Criticality?*, *Journal of Statistical Physics*, 144 (2011), pp. 268–302.
- [35] MARK E. J. NEWMAN, *Power laws, pareto distributions and zipf’s law*, *Contemporary physics*, 46 (2005), pp. 323–351.
- [36] TIMOTHY J. NEWMAN, JEAN-BAPTISTE FERDY, AND C. QUINCE, *Extinction times and moment closure in the stochastic logistic process.*, *Theoretical population biology*, 65 (2004), pp. 115–26.
- [37] MARTIN RIEDLER, MICHLE THIEULLEN, AND GILLES WAINRIB, *Limit theorems for infinite-dimensional*

- piecewise deterministic markov processes. applications to stochastic excitable membrane models*, Electron. J. Probab., 17 (2012), pp. no. 55, 1–48.
- [38] FRIGYES RIESZ AND BÉLA SZ.-NAGY, *Functional analysis*, Dover Books on Advanced Mathematics, Dover Publications, Inc., New York, 1990. Translated from the second French edition by Leo F. Boron, Reprint of the 1955 original.
- [39] TIM ROGERS, *Maximum-entropy moment-closure for stochastic systems on networks*, Journal of Statistical Mechanics: Theory and Experiment, (2011), p. 20.
- [40] JAMES P. SETHNA, KARIN A. DAHMEN, AND CHRISTOPHER R. MYERS, *Crackling noise*, Nature, 410 (2001), pp. 242–50.
- [41] ADAM SHWARTZ AND ALAN WEISS, *Large deviations for performance analysis*, Chapman & Hall, London, 1995.
- [42] ABHYUDAI SINGH AND JOÃO PEDRO HESPANHA, *Approximate moment dynamics for chemically reacting systems*, IEEE Transactions on Automatic Control, 56 (2011), pp. 414–418.
- [43] N. SRI NAMACHCHIVAYA AND HENRY J. VAN ROESSEL, *Moment Lyapunov Exponent and Stochastic Stability of Two Coupled Oscillators Driven by Real Noise*, Journal of Applied Mechanics, 68 (2001), p. 903.
- [44] H. EUGENE STANLEY, *Scaling, universality, and renormalization: Three pillars of modern critical phenomena*, Reviews of Modern Physics, 71 (1999), pp. S358–S366.
- [45] ANDREW R. TEEL, ANANTHARAMAN SUBBARAMAN, AND ANTONINO SFERLAZZA, *Stability analysis for stochastic hybrid systems: A survey*, Automatica, (2014).
- [46] ROBERT R. WILKINSON AND KIERAN J. SHARKEY, *Message passing and moment closure for susceptible-infected-recovered epidemics on finite networks*, Physical Review E, 89 (2014), p. 022808.
- [47] G. GEORGE YIN AND CHAO ZHU, *Hybrid switching diffusions: properties and applications*, vol. 63, Springer, 2009.