A Maximum Entropy Approach to the Moment Closure Problem for Stochastic Hybrid Systems at Equilibrium

Jiangmeng Zhang, Lee DeVille, Sairaj Dhople, and Alejandro D. Domínguez-García

Abstract— **We study the problem that arises in a class of stochastic processes referred to as Stochastic Hybrid Systems (SHS) when computing the moments of the states using the generator of the process and Dynkin's formula. We focus on the case when the SHS is at equilibrium or approaching equilibrium. We present a family of such processes for which infinite-dimensional linear-system analysis tools are ineffective, and discuss a few differing perspectives on how to tackle such problems by assuming that the SHS state distribution is such that its entropy is maximum. We also provide a numerical algorithm that allows us to efficiently compute maximum entropy solutions, and compare results with Monte Carlo simulations for some illustrative SHS.**

I. Introduction

This paper focuses on the analysis of a class of stochastic processes known as Stochastic Hybrid Systems (SHS) introduced, and extensively studied, by Hespanha (see, e.g., [1]), who also showed that SHS are a subset of a more general class of stochastic processes known as Piecewise-Deterministic Markov processes [2]. SHS is a powerful modeling and analysis formalism that has been used in many engineering and science domains, including: networked control systems [3], power systems [4], system reliability theory [5], and chemical reaction dynamics [6].

The state space of an SHS is comprised of a *discrete state space* and a *continuous state space*. We refer to the pair formed by these two as the *combined state space* of the SHS. The transitions amongst the discrete states are random, and the rates at which these transitions occur are allowed to be a function of time and the continuous state. For any fixed value of the discrete state, the evolution of the continuous state is described by a stochastic differential equation (SDE). Moreover, whenever the discrete state changes, the continuous state is allowed to change discontinuously, and there is a reset map that defines the relation between pre- and post-transition states.

J. Zhang and A. D. Domínguez-García are with the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. E-mail: {jzhang67,aledan}@ILLINOIS.EDU.

L. DeVille is with the Department of Mathematics at the University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. E-mail: rdeville@ILLINOIS.EDU.

S. Dhople is with the Department of Electrical and Computer Engineering at the University of Minnesota, Minneapolis, MN 55455, USA. E-mail: sdhople@UMN.EDU.

This work was supported in part by the National Science Foundation under awards ECCS-CAR-0954420, CMG-0934491, and CCF-1442686; the Trustworthy Cyber Infrastructure for the Power Grid (TCIPG) under US Department of Energy Award DE-OE0000097; and the Initiative for Mathematical Science and Engineering at the University of Illinois.

A full understanding of an SHS would include obtaining the distribution of the combined state as a function of time; however, this is an intractable problem in general. In fact, only in a few special cases can this problem be solved. For instance, if the transitions amongst discrete states is independent of the continuous state, the evolution of the former is described by a continuous-time Markov chain, the solution of which is fully characterized by the Chapman-Kolmogorov equations. However, if these transitions depend on the continuous state, it is typically impossible to write down an exact distribution; thus, in this paper we settle for a method that allows the computation of any arbitrary number of its moments, as in [5]. To this end, we utilize the generator of the stochastic process and Dynkin's formula to obtain a differential equation describing the evolution of the expectation of any function of the combined state. [This is possible as long as such a function is in the domain of the extended generator [1], [5].]

Following [1], it can be shown that under certain weak hypotheses, monomials are always in the domain of the generator, and thus Dynkin's formula holds. Furthermore, for an SHS where (i) the vector fields defining the SDEs, (ii) the transition rates among discrete states, and (ii) the reset maps are polynomial, the extended generator maps the set of monomials to itself. Thus, Dynkin's formula yields a closed set of ordinary differential equations (ODEs), which describes the evolution of the value of each moment. Unfortunately, since there are infinitely many monomials, the approach outlined above produces an infinite-dimensional system of ODEs in what is commonly referred to in the applied mathematics literature as a *moment closure* problem.

In this paper, we focus on the moment closure problem in polynomial SHS *at equilibrium*, with special emphasis on the study of what we refer to as the *one-state, one-reset, one-dimensional SHS*. [By focusing on this system, we can clearly illustrate the challenges in the analysis, but we do not see any fundamental limitation that would prevent extending our proposed analysis approach to multi-state and multi-reset systems.] For this system, we formulate the moment flow equations, which result in an infinite-dimensional linear system of ODEs; however, it turns out, infinite-dimensional linear system theory tools (see, e.g., [7]) are not suitable to analyze the actual behavior of the moment dynamics. In particular, one can show that this infinite-dimensional system of ODEs have, at least formally, infinitely many solutions—yet only one of these corresponds to the actual solution of the stochastic process.

The key feature of our approach is to assume that the SHS invariant distribution is such that its entropy is maximized under the constraints imposed by Dynkin's formula applied to the family of all monomials; this allows us to obtain a solution to the moment closure problem. In this regard, we propose a numerical algorithm that allows us to efficiently compute the maximum entropy solution, and showcase it via several numerical simulations involving the one-state, one-reset, one-dimensional SHS, and other polynomial SHS.

The moment closure problem in SHS was also tackled by Hespanha in several of his papers (see, e.g., [1], [6]), which inspired our work; however, the focus of Hespanha's work was to develop moment closures that are valid on compact time intervals. Specifically, he showed that for a compact time interval, the solution to the infinite-dimensional ODE can be approximated up to arbitrary precision by the solution of a finite-dimensional, but nonlinear, ODE. This finite-dimensional nonlinear ODE is obtained by truncating the infinite-dimensional ODE, and then using normal form theory to relate the missing higher-order moments to the extant lower-order moments; this *a fortoriori* provides a solution to the moment closure problem. On the other hand, we focus on the moment closure problem of SHS at equilibrium, i.e., in the limit as time goes to infinity; thus, the goal here (and the solution technique) is very different from that in Hespanha's work.

The maximum entropy approach that we describe here has been tried for moment closure problems by various researchers in various problem settings (see, in particular, [8], [9]). Specifically, [8] aimed at approximating high dimensional joint distributions of stochastic systems, given the knowledge of the marginal ditribution on some subsystems; [9] proposed a Newton-Raphson method based-algorithm to solve the optimization problem arising from a maximum entropy assumption. In contrast, our approach is to use the constraints arising from the equations themselves to identify the unique distribution.

The remainder of this paper is organized as follows. In Section II, we provide a brief overview of SHS and introduce the moment closure problem. In Section III, we formulate and analyze the moment flow behavior for a particular SHS, and introduce an entropy-based method to tackle the moment closure problem. Several numerical examples that illustrate the effectiveness of the approach are presented in Section IV. Concluding remarks and directions for future work are described in Section V.

II. Problem Formulation

In this section, we define an SHS and derive its generator. This paper concentrates on polynomial SHS, i.e., SHS where all the rates, flows, and resets are polynomial functions of the state variables—thus Dynkin's formula yields an infinite-dimensional system of ordinary differential equations (ODEs) that describes the evolution of the moments of the system. We point out that this leads generically to a *moment closure* problem, since the flow cannot be faithfully represented by any finite-dimensional truncation.

A. SHS Definition

Let $\mathcal{Q}$ be a countable set of *discrete states* and $\mathcal{P}$ a *continuous state* space. The general idea of an SHS is that in each small time interval $\mathrm{d}t$, the system jumps from discrete state q to discrete state q' with probability $\lambda_{q,q'}(x,t)\mathrm{d}t$; when this jump occurs, the continuous state is reset to $\psi_{q,q'}(x,t)$. If the system does not jump, then it stays in state q, and x evolves according to a stochastic differential equation (SDE) associated to the discrete state q. The precise definition is given below:

Definition 1. A stochastic hybrid system (SHS) is a quintuple $(\mathcal{Q}, \mathcal{P}, \Lambda, \Psi, \mathcal{F})$ where

- $\mathcal{Q}$ is a countable set of discrete states;
- $\mathcal{P}$ is the continuous state space with d dimensions, typically taken to be $\mathbb{R}^d$;
- $\Lambda = (\lambda_{q,q'})_{q,q'\in\mathcal{Q}}$ is a collection of transition rate functions $\lambda_{q,q'}(x,t)$ with $\lambda_{q,q'}: \mathcal{P}\times[0,\infty)\to\mathbb{R}^+$;
- $\Psi = (\psi_{q,q'})_{q,q'\in\mathcal{Q}}$ is a collection of reset maps $x\mapsto \psi_{q,q'}(x,t)$ with $\psi_{q,q'}: \mathcal{P}\times[0,\infty)\to\mathcal{P}$;
- $\mathcal{F}$ is a collection of stochastic differential equations, describing the dynamics of continuous state $x\in\mathcal{P}$ in any discrete state $q\in\mathcal{Q}$, as
$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = f_q(x,t) + g_q(x,t)\dot{n},$$
where n is a vector of independent Brownian motion processes.

For the purposes of this paper, we will assume the state dynamic evolution in each discrete state is governed by ODEs, i.e., $g_q(x,t) = 0$. Moreover, we assume that the system is stationary, i.e., f_q, $\lambda_{q,q'}$, and $\psi_{q,q'}$ do not depend explicitly on time.

B. Generator and Moment Flow Equation

Let $h\colon \mathcal{Q}\times\mathcal{P}\to\mathbb{R}$ be bounded and continuously differentiable with respect to its second argument (we will call such a function an *observable*). Define $\mathcal{L}h$ by

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

It is not hard to see that $\mathcal{L}$ is then a linear operator on the space of observables, and by pushing definitions around, we also obtain the formula commonly known as *Dynkin's formula*:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[h(Q_t,X_t)] = \mathbb{E}[\mathcal{L}h(Q_t,X_t)]. \tag{1}$$

We can then extend the domain of definition of $\mathcal{L}$ to include all h such that (1) holds, and one can show that, under mild conditions on Λ, Ψ, and $\mathcal{F}$, this extended domain contains all polynomials and indicator functions (see [2] for details).

For the SHS defined above, we can explicitly compute

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

We see that if f_q, $\lambda_{q,q'}$, and $\psi_{q,q'}$ are polynomials in x, then $\mathcal{L}$ maps polynomials to polynomials.

If we denote $h_q^m(q',x) = x^m \delta_{q,q'}$, then $\mathcal{L}h_q^m$ is a polynomial, and thus a (finite) linear combination of $h_{q'}^{m'}$, therefore (1) is equivalent to an (infinite-dimensional) set of ODEs on the functions $\mathbb{E}[h_q^m]$. The interpretation of $\mathbb{E}[h_q^m]$ is that it is the mth moment of the continuous state x, conditioned on the system being in discrete state q.

C. The Moment Closure Problem

While we are able to obtain a family of ODEs from (1), this system is in general infinite-dimensional. Moreover, it is not hard to see that under certain conditions, this system is irreducibly infinite-dimensional. For example, if the degree of any f_q is greater than one, or the degree of any of the $\lambda_{q,q'}$ or $\psi_{q,q'}$ is positive, then the degree of $\mathcal{L}h_q^m$ is higher than that of h_q^m. This implies that the dynamics of any given moment is a function of some higher-order moments, and thus no finite truncation can be exact.

Following standard probability theory terminology, we call this a *moment closure* problem, since the evolution of any finite set of moments depends on a larger set of moments. To understand the dynamics of any one moment, we must consider the flow of infinitely many other moments.

III. Analysis

To clearly illustrate and analyze the moment closure problem, we will focus on what we call the one-state, one-reset, one-dimensional SHS; namely, we assume that $\mathcal{Q}, \Lambda, \Psi$ each have exactly one element, and $\mathcal{P} = \mathbb{R}$. We note that there is no conceptual barrier to extending these methods to multi-state and multi-reset systems (see e.g., [3]–[6] for examples of such systems).

In this section, we will formulate the moment flow equations for this class of SHS and analyze the moment flow behavior. Then, we introduce the maximum entropy conjecture, based on which, we propose a method to approximate the steady-state distributions and statistics by solving a set of algebraic equations.

A. Moment Flow Analysis

Since there is only one state and one reset, in subsequent developments, we drop the notational dependence on the state q. We now consider the SHS that results from the following assumptions:

A1. $f(x), \lambda(x)$ are polynomials of degrees d_f, d_λ, respectively, $\lambda(x) = \sum_{n=0}^{d_\lambda} \lambda_n x^n$, $f(x) = \sum_{n=0}^{d_f} f_n x^n$ and $d_f \le d_\lambda$;

A2. $\lambda(x) > 0$ for $x > 0$;

A3. $\psi(x) = \gamma x$ with $\gamma \in [0, 1]$; and

A4. $f_0 = 0$, $f_1 > 0$.

Assumption A4 sets the minimum degree of all terms in $f(x)$ to be 1. With Assumption A1, we allow for $f(x)$ to be superlinear, which may lead to finite-time blowup on its own; but we require that the reset rate $\lambda(x)$ grows at least as fast as $f(x)$.

Then, the generator of the resulting SHS under Assumptions A1-A4, becomes

$$\mathcal{L}h(x) = f(x)\frac{\mathrm{d}}{\mathrm{d}x}h(x) + \lambda(x)(h(\gamma x) - h(x)). \quad (2)$$

For the dynamics of the state statistical moments, we define the test function $h^{(m)}(x) := x^m$ and the state moments $\mu_m := \mathbb{E}[X_t^m] = \mathbb{E}[h^{(m)}(X_t)]$. Then plugging this test function into (2), we obtain

$$\mathcal{L}h^{(m)}(x) = \frac{f(x)}{x} m h^{(m)}(x) + \lambda(x)(\gamma^m - 1)h^{(m)}(x). \quad (3)$$

With $f_0 = 0$, the right-hand side of (3) is a polynomial with all its powers of, at least, order m. Then, from (1), we obtain that the moment flow equations are:

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_m = \sum_{l=m}^{m+d_\lambda} C_{m,l}\mu_l, \quad (4)$$

where

$$C_{m,m+l} = m f_{l+1} + (\gamma^m - 1)\lambda_l,\ 0 \le l < d_f,$$
$$C_{m,m+l} = (\gamma^m - 1)\lambda_l,\ d_f \le l \le d_\lambda,$$

from where we see that $C_{m,m+d_\lambda} < 0$ in general, and $C_{m,m} > 0$ for m large enough.

Note that if we choose $h(x) = x$, then $\mathcal{L}h(x)$ is a polynomial whose largest coefficient is negative. Thus there is a b with $\mathcal{L}h \le -h + b\mathbf{1}_A$, where A is a compact subset of the positive reals, and the SHS has a unique invariant measure to which all initial measures converge exponentially quickly [10, Theorem 14.0.1].

Even though we know the system in (4) converges, the moment closure problem remains if we are interested in computing anything about the invariant measure. In this regard, the steady-state solution of the system in (4), defined as $\tilde{\mu}_m$, satisfies the (infinite) family of linear equations

$$\sum_{l=m}^{m+d_\lambda} C_{m,l}\tilde{\mu}_l = 0, \quad (5)$$

or equivalently,

$$\tilde{\mu}_{m+d_\lambda} = -\frac{1}{C_{m,m+d_\lambda}} \sum_{l=m}^{m+d_\lambda-1} C_{m,l}\tilde{\mu}_l,$$

but we see that this system is underdetermined, and has d_λ degrees of freedom unresolved. Thus, if $\lambda(x)$ is non-constant, then (5) has infinitely many solutions.

B. Maximum-Entropy Method

From the discussion above, it is clear that the constraints imposed by (5) do not completely determine the moments at equilibrium of the SHS that results from Assumptions A1-A4. The natural conjecture to make in this case is that the equilibrium distribution of the stochastic process is the maximum entropy distribution under these constraints. Next, we formalize this idea and propose an algorithm to compute such a maximum entropy distribution efficiently. Then, by comparing with Monte Carlo simulations, we provide evidence that this approach provides the correct answer for a wide variety of illustrative cases.

1) Maximum Entropy Conjecture: Our conjecture is that the equilibrium distribution of the SHS is the maximum entropy distribution in the class of all distributions the moments of which satisfy (5). Thus, we can cast the problem of finding this distribution as an optimization problem.

Denote the random variable in steady-state X_∞ as X for short, and define the distribution function of X as $p_X(x)$. Then, the reformulated problem can be presented as

$$p_X(x) = -\arg\inf_{p(x)} \left(\int p(x) \log p(x) \mathrm{d}x \right),$$

$$\text{subject to} \quad \int p(x) \mathrm{d}x = 1, \tag{6}$$

$$\sum_{l=m}^{m+d_\lambda} C_{m,l} \int x^l p(x) \mathrm{d}x = 0, \, \forall m \in \mathbb{Z}^+. \tag{7}$$

The objective function is the entropy of a probability distribution. The constraint in (6) guarantees the obvious condition that the integration of the probability distribution is equal to one, whereas the constraint in (7) is just the explicit expression for (5).

Due to the infinite number of constraints, the problem is still intractable; however, the reformulation as an optimization problem can help us construct the distribution function structure via its Lagrangian (see, e.g., [11]):

$$\begin{aligned} L = &\int p(x) \log p(x) \mathrm{d}x \\ &+ \tilde{\nu}_0 \left[\int p(x) \mathrm{d}x - 1 \right] \\ &+ \sum_{m=1}^{\infty} \tilde{\nu}_m \left[\sum_{l=m}^{m+d_\lambda} C_{m,l} \int x^l p(x) \mathrm{d}x \right]. \end{aligned}$$

Then, the necessary conditions for optimality are (see, e.g., [11], [12]):

$$\log p_X(x) + \tilde{\nu}_0 + \sum_{m=1}^{\infty} \tilde{\nu}_m \sum_{l=m}^{m+d_\lambda} C_{m,l} x^l = 0, \tag{8}$$

which can be written more compactly as

$$\log p_X(x) - \nu_0 - \sum_{i=1}^{\infty} \nu_i x^i = 0, \tag{9}$$

where the ν_i's are linear combinations of the $\tilde{\nu}_i$'s. Then, by rearranging (9), we obtain a distribution function $p_X(x)$ of the form

$$p_X(x) = c \exp \left(\sum_{i=1}^{\infty} \nu_i x^i \right), \tag{10}$$

with $c = \exp \nu_0$.

2) Moment Relation Using Integration by Parts: The formula in (10) allows us to approximate the distribution function by truncating up to order n:

$$p_X(x) \approx c \exp \left(\sum_{i=1}^{n} \nu_i x^i \right). \tag{11}$$

To estimate the ν_i's, we evaluate the moments as functions of these parameters. Then, by substituting into (5), we will obtain a set of equations that can be used to calculate the values of the ν_i's. However, this method seems intractable, due to the difficulty to find explicit formulas for the moment functions being expressed by the ν_i's. This difficulty can be avoided by using the following lemma, which gives the relationship among the moments of a random variable Z that has the distribution function in (11).

Lemma 2. *For a random variable Z with probability density function*

$$p_Z(z) = c \exp \sum_{i=1}^{n} \nu_i z^i, \tag{12}$$

it holds that:

$$k\mathbb{E}[Z^{k-1}] + \sum_{i=1}^{n} i\nu_i \mathbb{E}[Z^{i+k-1}] = 0, \, \forall k \in \mathbb{Z}^+. \tag{13}$$

The result of the lemma above can be proven using integration by parts. First, for an arbitrary $j \in \mathbb{Z}^+$and $1 \leq j \leq n$, write the $(k+j-1)$th moment of Z explicitly in an integral form. Rewrite the integrand as a product of two parts, the derivative of $e^{\nu_j z^j}$ and the rest. Then following the integration by parts theorem, the $(k+j-1)$th moment can be expressed as a linear combination of other moments of Z.

3) Calculation of Distribution Parameters: From the discussion in Section III-A, we know that the system has d_λ degree of freedom. Then, it follows that all the moments can be expressed as linear combinations of the first d_λ moments. By plugging them into (13), we see that there are $n+d_\lambda$ unknowns: n distribution function parameters and d_λ moments. Note that the scalar c does not appear in (13). By varying k between 1 and $n+d_\lambda$ in (13), we obtain $n+d_\lambda$ algebraic equations. Therefore, we are able to compute the distribution function parameters as well as the statistics by solving a set of algebraic equations.

We now show these calculations to completion, but for the purposes of brevity, we only show them for the case where all of the polynomials defining the SHS are linear, i.e.

$$f(x) = \alpha x, \quad \lambda(x) = \beta x, \quad \psi(x) = \gamma x. \tag{14}$$

In this case, there is only one unresolved degree of freedom, and (5) becomes

$$\tilde{\mu}_{m+1} = \frac{\alpha m}{\beta(1-\gamma^m)} \tilde{\mu}_m, \tag{15}$$

from where we obtain that

$$\mathbb{E}[X^m] = d_m \mathbb{E}[X], \tag{16}$$

with

$$\begin{aligned} d_m &= \prod_{k=1}^{m-1} \frac{\alpha k}{\beta(1-\gamma^k)} \\ &= \frac{(m-1)!\alpha^{m-1}}{\beta^{m-1}(1-\gamma^{m-1})(1-\gamma^{m-2})\cdots(1-\gamma)}, \, m \geq 2. \end{aligned}$$

Substituting (16) into (13) leads to

$$kd_{k-1}\mathbb{E}[X] + \sum_{i=1}^{n} i\nu_i d_{i+k-1}\mathbb{E}[X] = 0,\ k \in \mathbb{Z}^+, \tag{17}$$

with $d_1 = 1$ and $d_0 = (\mathbb{E}[X])^{-1}$ for consistency. When $k = 1$, we have

$$1 + \nu_1\mathbb{E}[X_{(n)}] + 2\nu_2 d_2\mathbb{E}[X_{(n)}] + \cdots + n\nu_n d_n\mathbb{E}[X_{(n)}] = 0, \tag{18}$$

where the subscript (n) indicates the approximation by truncating up to order n. For $k \geq 2$, every term in the relations in (17) contains the mean $\mathbb{E}[X_{(n)}]$. Since $\mathbb{E}[X_{(n)}] \neq 0$, we can divide both sides of (17) by $\mathbb{E}[X_{(n)}]$, which, after rearrangement, results in a set of linear equations on the ν_i's:

$$\tilde{H}_n U_n = \tilde{D}_n, \tag{19}$$

with $U_n = [\nu_1, \nu_2, \cdots, \nu_n]^{\mathrm{T}}$ collecting the distribution parameters, $\tilde{D}_n = [-2, -3d_2, -4d_3, \cdots, -(n+1)d_n]^{\mathrm{T}}$, and

$$\tilde{H}_n = \begin{bmatrix} d_2 & 2d_3 & 3d_4 & \cdots & nd_{n+1} \\ d_3 & 2d_4 & 3d_5 & \cdots & nd_{n+2} \\ d_4 & 2d_5 & 3d_6 & \cdots & nd_{n+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n+1} & 2d_{n+2} & 3d_{n+3} & \cdots & nd_{2n} \end{bmatrix},$$

where the subscript indicates their dependency on the truncation order n. Assuming $\tilde{H}_n$ is invertible, then the distribution parameters can be obtained as

$$U_n = \tilde{H}_n^{-1}\tilde{D}_n. \tag{20}$$

4) Computing the Mean: Substituting (20) into (18) gives us a way to directly calculate the mean as

$$\begin{aligned} \mathbb{E}[X_{(n)}] &= \frac{1}{[1\, 2d_2\, 3d_3 \cdots nd_n]U_n} \\ &= \frac{1}{[1\, 2d_2\, 3d_3 \cdots nd_n]\tilde{H}_n^{-1}\tilde{D}_n}. \end{aligned} \tag{21}$$

Eventually, we obtain that

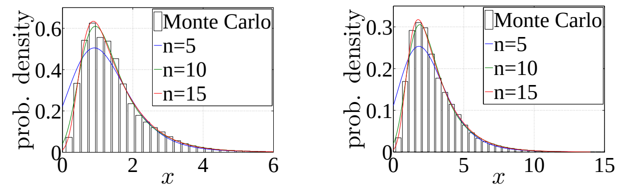
$$\mathbb{E}[X_{(n)}] = \frac{1}{D_n H_n^{-1}\tilde{D}_n}, \tag{22}$$

where $D_n = [1, d_2, d_3, \cdots, d_n]$, and H_n is a Hankel matrix:

$$H_n = \begin{bmatrix} d_2 & d_3 & d_4 & \cdots & d_{n+1} \\ d_3 & d_4 & d_5 & \cdots & d_{n+2} \\ d_4 & d_5 & d_6 & \cdots & d_{n+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{n+1} & d_{n+2} & d_{n+3} & \cdots & d_{2n} \end{bmatrix}. \tag{23}$$

Therefore, the true system steady-state mean is

$$\mathbb{E}[X] = \lim_{n\to\infty} \frac{1}{D_n H_n^{-1}\tilde{D}_n}. \tag{24}$$



(a) $\alpha = 1, \beta = 1, \gamma = 0.5$. (b) $\alpha = 2, \beta = 1, \gamma = 0.5$.

Fig. 1: Pdf of single-state α-β-γ system.

5) Accuracy and Convergence: It seems difficult in general to prove convergence directly by taking limits on both sides of (22) since the matrices D_n, H_n and $\tilde{D}_n$ are complicated (however, a special case is when the value of γ is 0, which will be discussed further in Section IV-B). On the other hand, we can prove indirectly that this limit exists: first, a unique invariant distribution exists by the arguments of Section III-A, and, moreover, it is not hard to see that this distribution decays exponentially in x. From this, it follows that if the invariant measure is equal to the one that results from the maximum entropy conjecture as given in (10), then the sequence of functions in (11) converges to the invariant measure uniformly on any compact (in x) set. We will verify numerically in all case studies below that this convergence is accurate up to an arbitrary error.

IV. Numerical Examples

In this section, we verify our proposed method by comparing the solutions provided by the maximum entropy approach with results obtained via Monte Carlo simulations. Several of the cases that we show will be the linear version of the SHS that was examined in detail in the previous section, mostly because the class of all such systems can be parameterized simply, and we show that the numerics check for all values of the parameters. We also show that the approach works for nonlinear systems as well.

In all cases, we compare the analytic solution of a truncated maximum entropy method described above with the results of a Monte Carlo simulation done on the SHS system itself. Note that for the SHS defined by (14), from (22) or dimensional analysis, we can view α/β as a scaling factor, making γ the only free parameter.

A. Distribution Verification

In the first case, we set $\alpha = 1$, $\beta = 1$, and $\gamma = 0.5$ in (10). The pdf of the invariant distribution obtained from a 500,000-sample Monte Carlo simulation is depicted in Fig. 1a as a benchmark. With the maximum entropy method, the parameters defining the probability distribution function (i.e., ν_i's) can be estimated using (20). Then, by substituting them back into (11), we obtain the pdf of the invariant distribution. Figure 1a shows the approximate pdfs when increasing the truncation order, $n = 5$, 10, and 15. As the number of retained terms in the truncation increases, the approximate pdf obtained via the maximum entropy method approaches that obtained via Monte Carlo. When the truncation order is fifteen, the approximate mean that results from (22) is 1.4318, while that estimated via Monte Carlo is 1.4307. The difference is within one standard deviation (as obtained via Monte Carlo, yielding a value of 0.0012). Similarly, the results for $\alpha = 1$, $\beta = 1$, and $\gamma = 0.7$ are shown in Fig. 1b. The mean estimated via the maximum entropy approach with truncation order fifteen is 2.8637, while a value of 2.8660 results from Monte Carlo, with the standard deviation being 0.0026.

B. Sensitivity to Parameter γ

We fix the ratio of α and β to one, and vary γ from 0 to 1. The mean and second moment obtained via the maximum entropy approach and via Monte Carlo simulations are depicted in Fig. 2. In both figures, the two lines match accurately, except when γ is close to 0 or 1. When $\gamma = 0$, the mean obtained via maximum entropy is larger than that obtained via Monte Carlo simulation. This is because the rate of convergence of the solution obtained via (22) is quite slow when $\gamma \approx 0$. We can show that when $\gamma = 0$, and the ratio of α and β is one, the approximate mean with truncation order of n is calculated via (22) as

$$\mathbb{E}[X_{(n)}] = \frac{1}{1 + \sum_{i=1}^{n} \frac{1}{i}},$$

where as $n \to \infty$, the mean converges to zero. However, the convergence rate is only logarithmic.

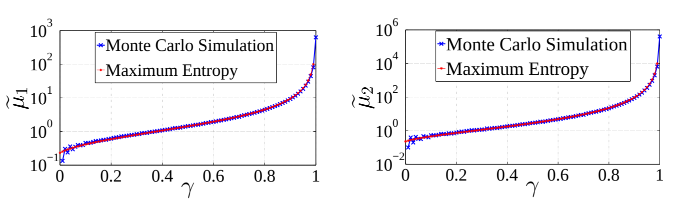
C. Nonlinear Case Study

We set the state evolution equation to be $\dot{x} = \alpha x^2$; the transition rate to be βx^2 and the reset function to be γx. We arbitrarily choose $\alpha = 1$, $\beta = 1$, and $\gamma = 0.7$, and apply the same procedure as for the linear SHS case study discussed earlier. The only difference is that in this case, the mean does not appear in the equations of the invariant distribution moments obtained by applying Dynkin's formula. Subsequently, the pdf function is constructed as

$$p_X(x) = \exp(\nu_0 + \nu_2 x^2 + \nu_3 x^3 + \cdots + \nu_n x^n),$$

where n is the truncation order and the linear term $\nu_1 x$ does not appear.

The pdfs obtained by using the maximum entropy approach with truncation order being 18, and those obtained via Monte Carlo simulations are depicted in Fig. 3, where one can see that they match well. The mean obtained with the maximum entropy approach is 2.2930, whereas that obtained via Monte Carlo simulations is 2.2912. Again the mismatch (i.e., 0.0018) is within one standard deviation as obtained via Monte Carlo (i.e., 0.0023).



(a) mean with respect to γ. (b) second moment with respect to γ.

Fig. 2: Moments with respect to γ.

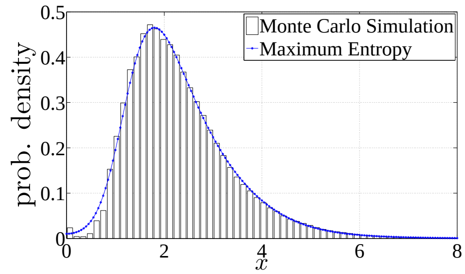


Fig. 3: pdf for a nonlinear case.

V. Concluding Remarks and Future Work

In this paper, we have analyzed the moment closure problem that commonly arises in the class of polynomial SHS. Special attention was given to the so-called one-state, one-reset SHS, and we showed that this generically gives rise to a moment closure problem.

We conjecture that this problem can be resolved by considering the problem from a "maximum entropy" point of view, based on which we can derive the formula of the invariant distribution function. Then, we proposed a method to approximate the distribution function parameters, as well as the statistics; this method relies on solving a set of algebraic equations. The maximum entropy method has been verified by a set of numerical examples.

We believe that this maximum-entropy method can be generalized to the analysis of multi-state multi-dimensional systems, which we will investigate in future work.

References

[1] J. P. Hespanha, "Modelling and analysis of stochastic hybrid systems," *IEE Proceedings–Control Theory and Applications*, vol. 153, no. 5, pp. 520–535, September 2006.

[2] M. H. A. Davis, *Markov Models and Optimization*. Boundary Row, London: Chapman and Hall, 1993.

[3] D. Antunes, J. P. Hespanha, and C. Silvestre, "Stochastic hybrid systems with renewal transitions: Moment analysis with applications to networked control systems with delays," *SIAM Journal of Control and Optimization*, vol. 51, no. 2, pp. 1481–1499, February 2013.

[4] S. Dhople, Y. Chen, L. DeVille, and A. Domínguez-García, "Analysis of power system dynamics subject to stochastic power injections," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 12, pp. 3341–3353, December 2013.

[5] S. V. Dhople, L. DeVille, and A. D. Domínguez-García, "A Stochastic Hybrid Systems framework for analysis of Markov reward models," *Reliability Engineering & System Safety*, vol. 123, pp. 158–170, 2014.

[6] A. Singh and J. P. Hespanha, "Approximate moment dynamics for chemically reacting systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 414–418, February 2011.

[7] R. Curtain and H. Zwart, *An Introduction to Infinite-Dimensional Linear System Theory*. New York, NY: Springer, 1995.

[8] T. Rogers, "Maximum-entropy moment-closure for stochastic systems on networks," *J. Stat. Mech.*, vol. 2011, no. 05, p. P05007, 2011.

[9] P. Smadbeck and Y. N. Kaznessis, "A closure scheme for chemical master equations," *Proceedings of the National Academy of Sciences*, vol. 110, no. 35, pp. 14 261–14 265, 2013.

[10] S. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, 2nd ed. New York, NY, USA: Cambridge University Press, 2009.

[11] A. Zellner and R. A. Highfield, "Calculation of maximum entropy distributions and approximation of marginal posterior distributions," *Journal of Econometrics*, vol. 37, no. 2, pp. 195 – 209, 1988.

[12] J. Skilling, *Maximum entropy and Bayesian methods*, ser. Fundamental Theories of Physics. New York, NY: Springer, 1989, vol. 36.