

Performance and Robustness Analysis of Stochastic Jump Linear Systems using Wasserstein metric

Kooktae Lee^a, Abhishek Halder^a, Raktim Bhattacharya^a

^a*Department of Aerospace Engineering, Texas A&M University, College Station, TX 77843-3141 USA.*

Abstract

This paper focuses on the performance and the robustness analysis of stochastic jump linear systems. The state trajectory under stochastic jump process becomes random variables, which brings forth the probability distributions in the system state. Therefore, we need to adopt a proper metric to measure the system performance with respect to stochastic switching. In this perspective, Wasserstein metric that assesses the distance between probability density functions is applied to provide the performance and the robustness analysis. Both the transient and steady-state performance of the systems with given initial state uncertainties can be measured in this framework. Also, we prove that the convergence of this metric implies the mean square stability. Overall, this study provides a unifying framework for the performance and the robustness analysis of general stochastic jump linear systems, but not necessarily Markovian jump process that is commonly used for stochastic switching. The practical usefulness and efficiency of the proposed method are verified through numerical examples.

Key words: Performance and robustness analysis, stochastic jump linear systems, switched linear systems, Wasserstein distance.

1 Introduction

A jump linear system is defined as a dynamical system constructed with a family of linear subsystem dynamics and a switching logic that conduct a switching between linear subsystems. Over decades, jump linear systems have attracted a wide range of researches due to its practical implementations. For instance, jump linear systems are used for power systems, manufacturing systems, aerospace systems, networked control systems, etc. In general, a jump linear system can be divided into two different categories depending on the switching logic. One branch is a deterministic switching where the jump process is deterministically given to the system. The utilization of such deterministic jump linear systems stems from plant stabilization [18], adaptive control [19], system performance [15], and resource-constrained scheduling [2]. In most cases, the system stability has been one of the major issues to investigate since even stable subsystems make the system unstable by the switching. Hence,

numerous results have been established for the stability analysis and the recent literature regarding the stability of deterministic jump linear systems can be found in [15]. In [15], a sufficient condition for the stability of deterministic jump linear systems is guaranteed by solving certain linear matrix inequalities (LMIs). Also, the necessary and the sufficient conditions for the stability are shown via a finite tuple, satisfying a certain condition.

Unlike the deterministic jump linear system, a stochastic jump linear system (SJLS) that is another category of jump linear systems refers to systems with the stochastic switching process. This type of jump linear systems is commonly used to represent the randomness in the switching such as communication delays or packet losses in the networked control systems [9, 25]. In [9], the networked control system with packet loss was modeled as an asynchronous dynamical system incorporating both discrete and continuous dynamics, and its stability was analysed through Lyapunov techniques. Since then, this problem has been formulated in a more general setting by representing the various aspects of communication uncertainties as Markov chains [3, 17, 26–28]. Stability analysis in the presence of such uncertainty, has been performed in the Markov jump linear systems (MJLSs) framework [11, 13, 24, 25, 29, 30]. Further, the stochastic stability for a class of nonlinear stochastic systems

* Corresponding author K.Lee. This research was supported by National Science Foundation award #1016299, with Dr. Helen Gill as the program manager.

Email addresses: animodor@tamu.edu (Kooktae Lee), ahalder@tamu.edu (Abhishek Halder), raktim@tamu.edu (Raktim Bhattacharya).

with semi-Markovian jump parameters is introduced in [10, 14]. Most previous literatures, however, have only dealt with steady-state analysis in terms of system stability.

Beyond the current literature, this paper has a key contribution for the analysis of a SJLS as follows. Based on the theory of optimal transport [23], we propose new probabilistic tools for analysing the performance and the robustness of SJLSs. Compared to the current literatures that only guarantees asymptotic performance with a deterministic arbitrary initial state condition, our contribution is to develop a unifying framework enabling both transient and asymptotic performance analysis with uncertain initial state conditions. The main difficulty dealing with analysis of SJLSs is that the system trajectories differ from every run due to the random switching. Moreover, the system state becomes random variables with a probability density function (PDF) even with deterministic initial state conditions. Consequently, we need to adopt a proper metric to measure the performance and the robustness of SJLSs in the distributional sense. In this paper, the Wasserstein metric that enables quantification of the uncertainty is employed for the performance measure. We prove that the convergence of this metric implies the mean square stability. To sum up, this paper provides the robustness analysis tools under the stochastic jumps with given initial state uncertainties without assuming any structure (e.g. Markov) on the underlying jump process.

The remainder of this paper is organized as follows. In Section II, we provide a brief review of the preliminaries. Section III deals with the performance and the robustness analysis of stochastic jump systems and develops computationally efficient tools for uncertainty quantification. Numerical examples are provided in Section IV, to illustrate the performance and the robustness analysis results developed in this work. Section V concludes the paper.

Notation: The set of real and natural numbers are denoted by \mathbb{R} and \mathbb{N} , respectively. Further, $\mathbb{N}_0 \triangleq \mathbb{N} \cup \{0\}$. The symbols $\text{tr}(\cdot)$, \otimes , and vec denote the trace of a square matrix, Kronecker product, and vectorization operators, respectively. The abbreviation m.s. stands for the convergence in mean-square sense. The notations $\mathbb{P}(\cdot)$ and $X \sim \rho(x)$ denote the probability and the random variable X with PDF $\rho(x)$, respectively. The symbol $\mathcal{N}(\mu, \Sigma)$ is used to denote the PDF of a Gaussian random vector with mean μ and covariance Σ .

2 Preliminaries

Consider a discrete-time jump linear system as follows.

$$x(k+1) = A_{\sigma_k} x(k), \quad k \in \mathbb{N}_0 \quad (1)$$

where $x(k)$ is the state vector and A_{σ_k} denotes the system matrices. $\sigma_k \in \mathcal{M} \triangleq \{1, 2, \dots, m\}$ stands for the stochastic jump process, governing the switching among m different modes of (1).

In this paper, we will consider general stochastic jump processes σ_k , and hence σ_k can be any arbitrary random process. Then, the resulting dynamics becomes a SJLS as defined next.

Definition 1 (Stochastic jump linear system) *Tuples of the form $(\pi(k), A_{\sigma_k}(x(k)), \mathcal{M})$ is termed as a SJLS, provided the mode dynamics are given by (1); $\pi(k)$ denote the time-varying occupation probability vectors for prescribed stochastic processes σ_k .*

Remark 1 *A SJLS, as defined above, is a collection of modal vector fields and a sequence of mode-occupation probability vectors. If the jump processes σ_k is deterministic, then at each time, $\pi(k)$ will have integral co-ordinates (single 1 and remaining $m-1$ zeroes), resulting in a deterministic switching sequence. If, however, σ_k is stochastic jump processes, then $\pi(k)$ will contain proper fractional co-ordinates, resulting in a randomized switching sequence where at each time, exactly one out of m modes will be chosen according to probability $\pi(k)$. Thus, starting from a deterministic initial condition, each execution of the SJLS may result in different switching sequences corresponding to random sample paths of σ_k over \mathcal{M} . Every realization of these random switching sequences results in a trajectory realization on the state space, and hence repeated the SJLS executions, even with a fixed initial condition, yields a spatio-temporal evolution of joint state PDF: $\rho(x(k))$.*

According to the structure that governs the temporal evolution of $\pi(k)$, some subsets of the stochastic jump processes can be listed as follows.

- 1) i.i.d. jump process:
A SJLS switching sequence is called stationary, if the occupation probability vector $\pi(k)$ remains stationary in time. In particular, a stationary deterministic switching sequence implies execution of a single mode (no switching). A stationary randomized switching sequence implies i.i.d. jump process.
- 2) Markov jump process:
Consider a discrete-time discrete state Markov chain with mode transition probabilities given by

$$p_{ij} = \mathbb{P}(\sigma_{k+1} = j \mid \sigma_k = i) \quad (2)$$

where $p_{ij} \geq 0, \forall i, j \in \mathcal{M}$. Hence, for $k \geq 0$, the probability distribution $\pi(k) \in \mathbb{R}^m$ of the modes of (1), is governed by

$$\pi(k+1) = \pi(k)P, \quad \pi(0) = [\pi_1(0) \cdots \pi_m(0)] \quad (3)$$

where the *transition probability matrix* $P \in \mathbb{R}^{m \times m}$ is a right stochastic matrix with row sum $\sum_{j=1}^m p_{ij} = 1, \forall i \in \mathcal{M}$.

3) semi-Markov jump process:

For a homogeneous and discrete-time semi-Markov chain, semi-Markov kernel q is defined by

$$q_{ij}(k) = \mathbb{P}(\sigma_{n+1} = j, X_{n+1} = k | \sigma_n = i) \quad (4)$$

where X_n denotes the sojourn time in state $\sigma_n = i$. Note that the transition probability p_{ij} in Markov chain can be expressed in terms of the semi-Markov kernel by $p_{ij} = \sum_{k=0}^{\infty} q_{ij}(k)$.

A SJLS refers to the jump linear system for which jump process σ_k is governed by any stochastic probability distribution $\pi(k)$. Consequently, a SJLS implies the jump linear system, where the jump probability distribution $\pi(k)$ forms proper fractional numbers with any arbitrary updating rules for $\pi(k)$.

3 Performance and Robustness Analysis using Wasserstein metric

Uncertainties in a SJLS appear at the execution level due to random switching sequence. Additional uncertainties may stem from imprecise setting of initial conditions and parameter values. These uncertainties manifest as the evolution of $\rho(x(k))$. Thus, a natural way to quantify the uncertainty in the performance of a SJLS, is to compute the “distance” of the instantaneous state PDF from a reference measure. In particular, if we fix the reference PDF as Dirac delta function at the origin, denoted as $\delta(x)$, then the time-history of this “distance” would reveal the rate of convergence (divergence) for the stable (unstable) SJLS in the distributional sense.

For meaningful inference, the notion of “distance” must define a metric, and should be computationally tractable. The choice of the metric is very important as it must be able to highlight properties of density functions that are important from a dynamical system point of view. We propose that the shape of the density functions characterizes the dynamics of the system. Regions of high probability density correspond to high likelihood of finding the state there, which corresponds to higher concentration of trajectories. Higher concentration occurs in regions with low time scale dynamics or time invariance. For example, for a stable system, all trajectories accumulate at the origin and the corresponding PDF is the Dirac delta function at the origin. Similarly, low concentration areas indicate fast-scale dynamics or instability, and the corresponding steady-state density function is zero in the unstable manifold. Therefore, behavior of two dynamical systems are identical in the distribution sense if their state PDFs have identical shapes. In order to properly capture the above aspects

in dynamical systems, we adopt Wasserstein distance and details are introduced in the following subsection.

3.1 Wasserstein distance

Definition 2 (Wasserstein distance) Consider the vectors $x_1, x_2 \in \mathbb{R}^n$. Let $\mathcal{P}_2(\varsigma_1, \varsigma_2)$ denote the collection of all probability measures ς supported on the product space \mathbb{R}^{2n} , having finite second moment, with first marginal ς_1 and second marginal ς_2 . Then the Wasserstein distance of order 2, denoted as \mathcal{W} , between two probability measures ς_1, ς_2 , is defined as

$$\mathcal{W}(\varsigma_1, \varsigma_2) \triangleq \left(\inf_{\varsigma \in \mathcal{P}_2(\varsigma_1, \varsigma_2)} \int_{\mathbb{R}^{2n}} \|x_1 - x_2\|_{\ell_2(\mathbb{R}^n)}^2 d\varsigma(x_1, x_2) \right)^{\frac{1}{2}}. \quad (5)$$

Remark 2 Intuitively, Wasserstein distance equals the least amount of work needed to morph one distributional shape to the other, and can be interpreted as the cost for Monge-Kantorovich optimal transportation plan [22]. The particular choice of ℓ_2 norm with order 2 is motivated in [7]. Further, one can prove (p. 208, [22]) that \mathcal{W} defines a metric on the manifold of PDFs.

Next, we present new results for system stability in terms of \mathcal{W} and simplifications in its computation.

Proposition 1 If we fix Dirac distribution as the reference measure, then distributional convergence in Wasserstein metric is necessary and sufficient for convergence in m.s. sense.

Proof. Consider a sequence of n -dimensional joint PDFs $\{\rho_j(x)\}_{j=1}^{\infty}$, that converges to $\delta(x)$ in distribution, i.e., $\lim_{j \rightarrow \infty} \mathcal{W}(\rho_j(x), \delta(x)) = 0$. From (5), we have

$$\begin{aligned} \mathcal{W}^2(\rho_j(x), \delta(x)) &= \inf_{\varsigma \in \mathcal{P}_2(\rho_j(x), \delta(x))} \mathbb{E} \left[\|X_j - 0\|_{\ell_2(\mathbb{R}^n)}^2 \right] \\ &= \mathbb{E} \left[\|X_j\|_{\ell_2(\mathbb{R}^n)}^2 \right] \end{aligned} \quad (6)$$

where the random variable $X_j \sim \rho_j(x)$, and the last equality follows from the fact that $\mathcal{P}_2(\rho_j(x), \delta(x)) = \{\rho_j(x)\} \forall j$, thus obviating the infimum. From (6), $\lim_{j \rightarrow \infty} \mathcal{W}(\rho_j(x), \delta(x)) = 0 \Rightarrow \lim_{j \rightarrow \infty} \mathbb{E} [\|X_j\|_{\ell_2}^2] = 0$, establishing distributional convergence to $\delta(x) \Rightarrow$ m.s. convergence. Conversely, m.s. convergence \Rightarrow distributional convergence, is well-known [6] and unlike the other direction, holds for arbitrary reference measure. \square

Proposition 2 (W between multivariate Gaussians [5]) The Wasserstein distance between two multivariate Gaussians supported on \mathbb{R}^n , with respective joint

PDFs $\mathcal{N}(\mu_1, \Sigma_1)$ and $\mathcal{N}(\mu_2, \Sigma_2)$, is given by

$$\mathcal{W}(\mathcal{N}(\mu_1, \Sigma_1), \mathcal{N}(\mu_2, \Sigma_2)) = \sqrt{\|\mu_1 - \mu_2\|_{\ell_2(\mathbb{R}^n)}^2 + \text{tr}\left(\Sigma_1 + \Sigma_2 - 2\left[\sqrt{\Sigma_1}\Sigma_2\sqrt{\Sigma_1}\right]^{\frac{1}{2}}\right)}. \quad (7)$$

Corollary 1 (W between Gaussian and Dirac PDF) Since we can write $\delta(x) = \lim_{\mu, \Sigma \rightarrow 0} \mathcal{N}(\mu, \Sigma)$ (see e.g., p. 160-161, [8]), it follows from (7) that

$$\mathcal{W}(\mathcal{N}(\mu, \Sigma), \delta(x)) = \sqrt{\|\mu\|_{\ell_2(\mathbb{R}^n)}^2 + \text{tr}(\Sigma)}. \quad (8)$$

3.2 Performance and Robustness Analysis for SJLSs

The performance and robustness analysis problem for the SJLS is stated as follows: given a SJLS $(\pi(k), A_{\sigma_k}(x(k)), \mathcal{M})$, compute and analyse the performance history, quantified by $W(k) \triangleq \mathcal{W}(\rho(x(k)), \delta(x))$. Comparison of $W(k)$ of uncertain systems with that of a nominal system, quantifies the degradation in system performance due to system uncertainty.

3.2.1 Uncertainty propagation in SJLSs

The key difficulty here is the propagation of state PDFs under the stochastic switching and we present a new algorithm for such computations.

Proposition 3 Given m absolutely continuous random variables X_1, \dots, X_m , with respective cumulative distribution functions (CDFs) $F_i(x)$, and PDFs $\rho_i(x)$, $\forall i \in \mathcal{M}$. Let $X \triangleq X_i$, with probability $\alpha_i \in [0, 1]$, $\sum_{i=1}^m \alpha_i = 1$. Then, the CDF and PDF of X are given by $F(x) = \sum_{i=1}^m \alpha_i F_i(x)$, and $\rho(x) = \sum_{i=1}^m \alpha_i \rho_i(x)$.

Proof. $F(x) \triangleq \mathbb{P}(X \leq x) = \sum_{i=1}^m \mathbb{P}(X = X_i) \mathbb{P}(X_i \leq x) = \sum_{i=1}^m \alpha_i F_i(x)$, where we have used the law of total probability. Since each X_i and hence X , is absolutely continuous, we have $\rho(x) = \sum_{i=1}^m \alpha_i \rho_i(x)$. \square

Note that any continuous PDF can be approximated by a Gaussian mixture PDF in weak sense [20, 21]. Therefore, we assume the initial PDF for the SJLS to be m_0 components mixture of Gaussian (MoG), given by

$\rho_0 = \sum_{j_0=1}^{m_0} \alpha_{j_0} \mathcal{N}(\mu_{j_0}, \Sigma_{j_0})$, $\sum_{j_0=1}^{m_0} \alpha_{j_0} = 1$. Then, we have the following results.

Theorem 1 (A SJLS preserves MoG) Consider a SJLS $(\pi(k), \{A_j\}_{j=1}^m, \mathcal{M})$ with initial PDF $\rho_0 = \sum_{j_0=1}^{m_0} \alpha_{j_0} \mathcal{N}(\mu_{j_0}, \Sigma_{j_0})$. Then the state PDF at time k , denoted by $\rho(x(k))$, is given by

$$\rho(x(k)) = \sum_{j_k=1}^m \sum_{j_{k-1}=1}^m \dots \sum_{j_1=1}^m \sum_{j_0=1}^{m_0} \left(\prod_{r=1}^k \pi_{j_r}(r) \right) \alpha_{j_0} \mathcal{N}(\mu_{j_k}, \Sigma_{j_k}) \quad (9)$$

where $\mu_{j_k} = A_{j_k}^* \mu_{j_0}$, $\Sigma_{j_k} = A_{j_k}^* \Sigma_{j_0} A_{j_k}^{*\top}$ and $A_{j_k}^* \triangleq \prod_{r=k}^1 A_{j_r} = A_{j_k} A_{j_{k-1}} \dots A_{j_2} A_{j_1}$.

Proof. Starting from ρ_0 at $k = 0$, the modal PDF at time $k = 1$, is given by

$$\rho_{j_1}(x(1)) = \sum_{j_0=1}^{m_0} \alpha_{j_0} \mathcal{N}(\mu_{j_1}, \Sigma_{j_1}) \quad (10)$$

where $j_1 = 1, \dots, m$, $\mu_{j_1} = A_{j_1} \mu_{j_0}$, and $\Sigma_{j_1} = A_{j_1} \Sigma_{j_0} A_{j_1}^\top$, which follows from the fact that linear transformation of an MoG is an equal component MoG with linearly transformed component means and congruently transformed component covariances (see Theorem 6 and Corollary 7 in [1]). From Proposition 3, it follows that the state PDF at $k = 1$, is

$$\rho(x(1)) = \sum_{j_1=1}^m \sum_{j_0=1}^{m_0} \pi_{j_1}(1) \alpha_{j_0} \mathcal{N}(\mu_{j_1}, \Sigma_{j_1}) \quad (11)$$

where $\pi_{j_1}(1)$ is the occupation probability for mode j_1 at time $k = 1$. Notice that (11) is an MoG with mm_0 component Gaussians. Proceeding likewise from this $\rho(x(1))$, we obtain

$$\rho_{j_2}(x(2)) = \sum_{j_1=1}^m \sum_{j_0=1}^{m_0} \pi_{j_1}(1) \alpha_{j_0} \mathcal{N}(\mu_{j_2}, \Sigma_{j_2}) \quad (12)$$

where $j_2 = 1, \dots, m$, $\mu_{j_2} = (A_{j_2} A_{j_1}) \mu_{j_0}$, $\Sigma_{j_2} = (A_{j_2} A_{j_1}) \Sigma_{j_0} (A_{j_2} A_{j_1})^\top$,

$$\rho(x(2)) = \sum_{j_2=1}^m \sum_{j_1=1}^m \sum_{j_0=1}^{m_0} \pi_{j_2}(2) \pi_{j_1}(1) \alpha_{j_0} \mathcal{N}(\mu_{j_2}, \Sigma_{j_2}). \quad (13)$$

Continuing with this recursion till time k , we arrive at (9), which is an MoG with $m^k m_0$ components. We

comment that the expression simplifies for $m_0 = 1$, i.e. when the initial PDF is Gaussian. \square

Remark 3 (Computational complexity) Given an initial MoG and a SJLS, from Theorem 1, one can in principle compute the state PDF at any finite time, in closed form (i.e., an analytical form with a finite number of well-defined functions). However, since the number of component Gaussians grows exponentially in time, the computational complexity in evaluating (9), grows exponentially, and hence the computation becomes intractable. In the following, we show that the Wasserstein based performance analysis can still be performed in closed form while keeping the computational complexity constant in time.

3.2.2 Wasserstein computation in SJLSs

For a SJLS, there are no known results to represent the W distance in closed form. The main computational issue is that even with Gaussian initial PDF, the instantaneous state PDF remains no longer Gaussian but rather MoG, as shown in Theorem 1. This brings forth concerns for the exponential growth of computational complexity to obtain $\rho(x(k))$. To address these concerns, we firstly introduce a following theorem that enables the Wasserstein computation in an analytical form. Then, we further show that the exponential growth can be obviated by the proposed algorithm.

Theorem 2 (W for an m -mode SJLS with Dirac reference PDF) At any given time k , let the state PDF

for a SJLS be $\rho(x) = \sum_{j=1}^m \alpha_j \rho_j(x)$, $x \in \mathbb{R}^n$ where $\rho_j(x)$,

α_j , and m are the instantaneous modal PDF, time-varying occupation probability of mode j , and the number of individual mixture components, respectively. If we define $W \triangleq \mathcal{W}(\rho(x), \delta(x))$, and $W_j \triangleq \mathcal{W}(\rho_j(x), \delta(x))$, then

$$W = \left(\sum_{j=1}^m \alpha_j W_j^2 \right)^{1/2}. \quad (14)$$

Proof. From (5) and Proposition 3, we have

$$\begin{aligned} W^2 &= \int_{\mathbb{R}^n} \|x\|_{\ell_2(\mathbb{R}^n)}^2 \rho(x) dx \\ &= \int_{\mathbb{R}^n} \|x\|_{\ell_2(\mathbb{R}^n)}^2 \sum_{j=1}^m \alpha_j \rho_j(x) dx \\ &= \sum_{j=1}^m \alpha_j \int_{\mathbb{R}^n} \|x\|_{\ell_2(\mathbb{R}^n)}^2 \rho_j(x) dx \\ &= \sum_{j=1}^m \alpha_j W_j^2. \end{aligned} \quad (15)$$

$$\Rightarrow W = \left(\sum_{j=1}^m \alpha_j W_j^2 \right)^{1/2}. \quad (16)$$

\square

Theorem 2 provides an analytical solution to compute the performance and the robustness of the SJLS in terms of Wasserstein distance. However, expression in (14) still includes the component-wise W computation, and hence the computation becomes intractable shortly due to the exponential growth of Gaussian components in the state PDF $\rho(x)$. In order to cope with this problem, we introduce a “Split-and-Merge” algorithm as follows.

1) Merge Step:

For a given MoG $\rho(x)$ at any time k , we can compute the mean $\hat{\mu}$ and covariance $\hat{\Sigma}$ of an MoG by the following lemma.

Lemma 1 (Mean and covariance of a mixture PDF) Consider any mixture PDF $\rho(x) = \sum_{j=1}^m \alpha_j \rho_j(x)$, with component mean-covariance pairs (μ_j, Σ_j) , $j = 1, \dots, m$. Then, the mean-covariance pair $(\hat{\mu}, \hat{\Sigma})$ for the mixture PDF $\rho(x)$, is given by

$$\hat{\mu} = \sum_{j=1}^m \alpha_j \mu_j, \quad \hat{\Sigma} = \sum_{j=1}^m \alpha_j \left(\Sigma_j + (\mu_j - \hat{\mu})(\mu_j - \hat{\mu})^\top \right). \quad (17)$$

Proof. We have $\hat{\mu} \triangleq \int_{\mathbb{R}^n} x \rho(x) dx = \int_{\mathbb{R}^n} x \sum_{j=1}^m \alpha_j \rho_j(x) dx = \sum_{j=1}^m \alpha_j \int_{\mathbb{R}^n} x \rho_j(x) dx = \sum_{j=1}^m \alpha_j \mu_j$.

On the other hand, $\hat{\Sigma} \triangleq \mathbb{E}[(x - \hat{\mu})(x - \hat{\mu})^\top] = \mathbb{E}[xx^\top] - \hat{\mu}\hat{\mu}^\top = \int_{\mathbb{R}^n} xx^\top \sum_{j=1}^m \alpha_j \rho_j(x) dx - \hat{\mu}\hat{\mu}^\top = \sum_{j=1}^m \alpha_j \int_{\mathbb{R}^n} (x - \hat{\mu} + \hat{\mu})(x - \hat{\mu} + \hat{\mu})^\top \rho_j(x) dx - \hat{\mu}\hat{\mu}^\top = \sum_{j=1}^m \alpha_j \left(\Sigma_j + (\mu_j - \hat{\mu})(\mu_j - \hat{\mu})^\top \right).$ \square

Lemma 1 proves that for any mixture PDF, we can compute the mean $\hat{\mu}$ and covariance $\hat{\Sigma}$. From the computed $\hat{\mu}(k)$ and $\hat{\Sigma}(k)$ at time k , we construct a synthetic Gaussian $\mathcal{N}(\hat{\mu}(k), \hat{\Sigma}(k))$ to merge the state PDF of an MoG form into a single Gaussian PDF.

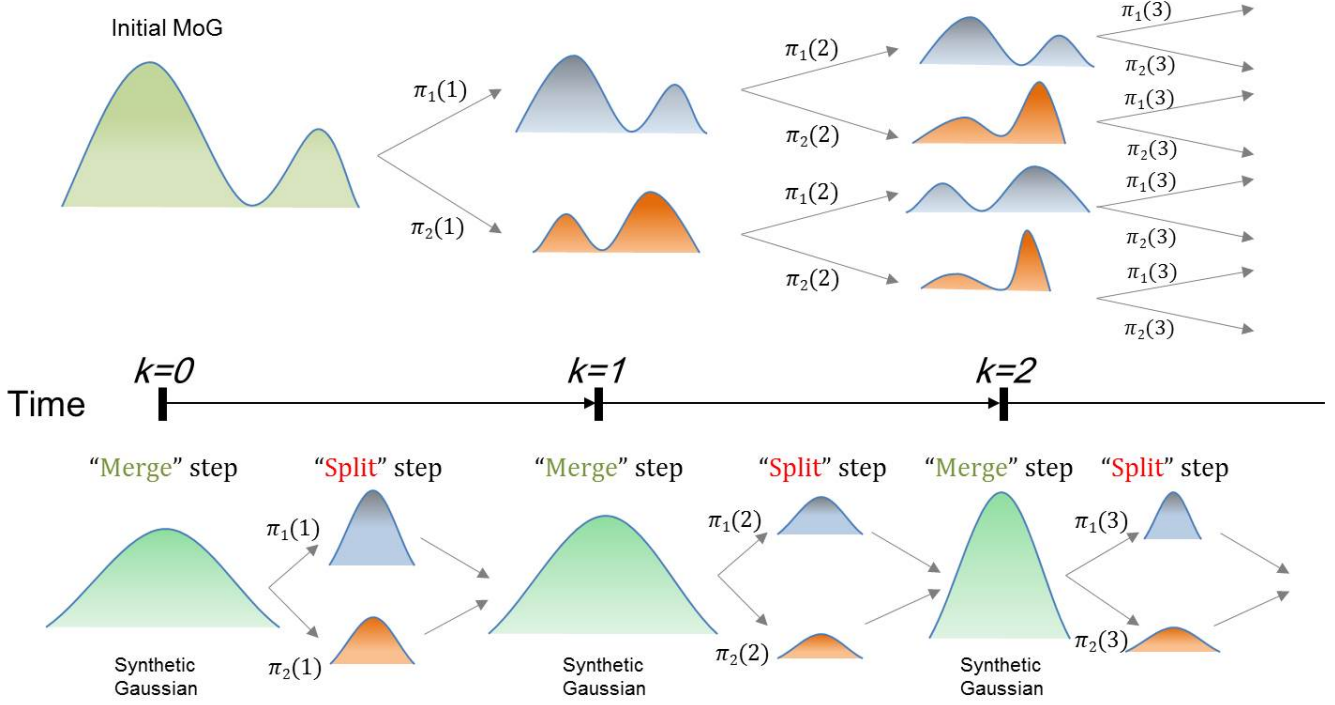


Fig. 1. Schematic of PDFs propagation for SJLS. Initially, an MoG PDF was given; Upper one shows the exponential growth of MoG components; Bottom one shows “Split-and-Merge” algorithm and the number of Gaussian components remains constant, which is m modes at most. In this figure, $m = 2$.

2) Split Step:

Once the synthetic Gaussian $\mathcal{N}(\hat{\mu}(k), \hat{\Sigma}(k))$ is obtained at time k from “Merge step”, we proceed the propagation of the modal PDF for the next time step along mode dynamics $\{A_j\}_{j=1}^m$. Consequently, we have m numbers of Gaussian components $\mathcal{N}(A_j \hat{\mu}(k), A_j \hat{\Sigma}(k) A_j^\top)$, $j = 1, 2, \dots, m$ at time $k + 1$.

Repeating “Split-and-Merge” algorithm at every time step as depicted by Fig. 1, linear modal dynamics results in m modal Gaussian PDFs (“Split step”). Then, instead of computing the non-Gaussian SJLS state PDF in an MoG form, one would construct a synthetic Gaussian $\mathcal{N}(\hat{\mu}, \hat{\Sigma})$ (“Merge step”) and repeat thereafter.

Although the “Split-and-Merge” algorithm obviates the need to compute the state PDF $\rho(x)$ where Gaussian components grow exponentially, the synthetic Gaussian PDF $\mathcal{N}(\hat{\mu}, \hat{\Sigma})$ does not imply that it can replace $\rho(x)$. Since $\rho(x)$ expressed in an MoG form has higher moments other than first and second, the distance between $\rho(x)$ and $\delta(x)$ may differ from that between $\mathcal{N}(\hat{\mu}, \hat{\Sigma})$ and $\delta(x)$. However, most importantly, we address that $\mathcal{W}(\rho(x), \delta(x))$ and $\mathcal{W}(\mathcal{N}(\hat{\mu}, \hat{\Sigma}), \delta(x))$ are equidistant at any time k by the following theorem.

Theorem 3 (Equidistance between W and \widehat{W})
At any given time k , let the state PDF for an m -

mode SJLS $\rho(x(k))$, be of the form (9), which we rewrite as $\rho(x(k)) = \sum_{j_k=1}^m \sum_{j_0=1}^{m_0} \alpha_{j_0} \beta_{j_k} \mathcal{N}(\mu_{j_k}, \Sigma_{j_k})$, where $\beta_{j_k} \triangleq \sum_{j_{k-1}=1}^m \dots \sum_{j_1=1}^m \left(\prod_{r=1}^k \pi_{j_r}(r) \right)$, $\mu_{j_k} = A_{j_k}^* \mu_{j_0}$, $\Sigma_{j_k} = A_{j_k}^* \Sigma_{j_0} A_{j_k}^{*\top}$, and $A_{j_k}^* \triangleq \prod_{r=k}^1 A_{j_r}$. Let the instantaneous mean and covariance of the mixture PDF $\rho(x(k))$ be $\hat{\mu}(k)$ and $\hat{\Sigma}(k)$, respectively. Then, we have

$$\widehat{W}(k) = W(k) = \left(\sum_{j_k=1}^m \sum_{j_0=1}^{m_0} \alpha_{j_0} \beta_{j_k} W_{j_k}^2(k) \right)^{1/2}, \forall k \in \mathbb{N}_0 \quad (18)$$

where

$$\begin{aligned} \widehat{W}(k) &\triangleq \mathcal{W}(\mathcal{N}(\hat{\mu}(k), \hat{\Sigma}(k)), \delta(x)), \\ W(k) &\triangleq \mathcal{W}(\rho(x(k)), \delta(x)), \\ W_{j_k}(k) &\triangleq \mathcal{W}(\mathcal{N}(\mu_{j_k}, \Sigma_{j_k}), \delta(x)), \\ \mu_{j_k} &= A_{j_k}^* \mu_{j_0}, \Sigma_{j_k} = A_{j_k}^* \Sigma_{j_0} A_{j_k}^{*\top}, \quad \forall k \geq 1. \end{aligned}$$

Proof. The rightmost equality in (18), follows directly

from Theorem 2. Thus, it suffices to prove that $\widehat{W}(k) = \left(\sum_{j_k=1}^m \sum_{j_0=1}^{m_0} \alpha_{j_0} \beta_{j_k} W_{j_k}^2(k) \right)^{1/2}$.

At time $k = 0$, the mean and covariance pair $(\widehat{\mu}_0, \widehat{\Sigma}_0)$ of an initial MoG can be computed by $(\widehat{\mu}_0, \widehat{\Sigma}_0) = (\sum_{j_0=1}^{m_0} \alpha_{j_0} \mu_{j_0}, \sum_{j_0=1}^{m_0} (\Sigma_{j_0} + (\mu_{j_0} - \widehat{\mu}_0)(\mu_{j_0} - \widehat{\mu}_0)^\top))$ from Lemma 1. If we construct a synthetic Gaussian $\mathcal{N}(\widehat{\mu}_0, \widehat{\Sigma}_0)$, Wasserstein distance \widehat{W} at time $k = 0$ can be computed by (8) as follows.

$$\widehat{W}^2(0) \stackrel{(8)}{=} \|\widehat{\mu}_0\|_{\ell_2(\mathbb{R}^n)}^2 + \text{tr}(\widehat{\Sigma}_0) \stackrel{(17)}{=} \widehat{\mu}_0^\top \widehat{\mu}_0 + \text{tr} \left(\sum_{j_0=1}^{m_0} \alpha_{j_0} (\Sigma_{j_0} + (\mu_{j_0} - \widehat{\mu}_0)(\mu_{j_0} - \widehat{\mu}_0)^\top) \right). \quad (19)$$

Since $\text{tr}(\cdot)$ is a linear operator, we can expand (19) as

$$\begin{aligned} \widehat{W}^2(0) &= \widehat{\mu}_0^\top \widehat{\mu}_0 + \sum_{j_0=1}^{m_0} \alpha_{j_0} \text{tr}(\Sigma_{j_0}) + \text{tr} \left(\sum_{j_0=1}^{m_0} \alpha_{j_0} \mu_{j_0} \mu_{j_0}^\top \right) \\ &\quad - \text{tr} \left(\left(\sum_{j_0=1}^{m_0} \alpha_{j_0} \mu_{j_0} \right) \widehat{\mu}_0^\top \right) - \text{tr} \left(\widehat{\mu}_0 \left(\sum_{j_0=1}^{m_0} \alpha_{j_0} \mu_{j_0} \right)^\top \right) \\ &\quad + \text{tr}(\widehat{\mu}_0 \widehat{\mu}_0^\top). \end{aligned} \quad (20)$$

Recalling that $\widehat{\mu}_0 = \sum_{j_0=1}^{m_0} \alpha_{j_0} \mu_{j_0}$ and $\widehat{\mu}_0^\top \widehat{\mu}_0 = \text{tr}(\widehat{\mu}_0^\top \widehat{\mu}_0) = \text{tr}(\widehat{\mu}_0 \widehat{\mu}_0^\top)$, the first, fourth, fifth and sixth term in the right-hand-side of (20) cancel out, resulting in

$$\begin{aligned} \widehat{W}^2(0) &= \sum_{j_0=1}^{m_0} \alpha_{j_0} \text{tr}(\mu_{j_0} \mu_{j_0}^\top) + \sum_{j_0=1}^{m_0} \alpha_{j_0} \text{tr}(\Sigma_{j_0}) \\ &= \sum_{j_0=1}^{m_0} \alpha_{j_0} \left(\|\mu_{j_0}\|_{\ell_2(\mathbb{R}^n)}^2 + \text{tr}(\Sigma_{j_0}) \right) \\ &= \sum_{j_0=1}^{m_0} \alpha_{j_0} \mathcal{W}^2(\mathcal{N}(\mu_{j_0}, \Sigma_{j_0}), \delta(x)) \\ &= \sum_{j_0=1}^{m_0} \alpha_{j_0} W_{j_0}^2(0) \stackrel{(14)}{=} W^2(0). \end{aligned} \quad (21)$$

Hence, $\widehat{W}(0)$ is equidistant with $W(0)$.

At time $k = 1$, we propagate the modal PDFs from a synthetic Gaussian $\mathcal{N}(\widehat{\mu}_0, \widehat{\Sigma}_0)$, which results in m modal Gaussians $\mathcal{N}(A_{j_1} \widehat{\mu}_0, A_{j_1} \widehat{\Sigma}_0 A_{j_1}^\top)$, $j_1 = 1, 2, \dots, m$ during “Split step”, followed by “Merge step” to obtain a new synthetic Gaussian $\mathcal{N}(\widehat{\mu}_1, \widehat{\Sigma}_1)$, where $\widehat{\mu}_1 = \sum_{j_1=1}^m \pi_{j_1}(1) A_{j_1} \widehat{\mu}_0$ and $\widehat{\Sigma}_1 = \sum_{j_1=1}^m \pi_{j_1}(1) \left(A_{j_1} \widehat{\Sigma}_0 A_{j_1}^\top + (A_{j_1} \widehat{\mu}_0 - \widehat{\mu}_1)(A_{j_1} \widehat{\mu}_0 - \widehat{\mu}_1)^\top \right)$

from Lemma 1. Then, $\widehat{W}(1)$ can be computed by

$$\begin{aligned} \widehat{W}^2(1) &\stackrel{(8)}{=} \|\widehat{\mu}_1\|_{\ell_2(\mathbb{R}^n)}^2 + \text{tr}(\widehat{\Sigma}_1) = \widehat{\mu}_1^\top \widehat{\mu}_1 + \text{tr} \left(\sum_{j_1=1}^m \pi_{j_1}(1) \left(A_{j_1} \widehat{\Sigma}_0 A_{j_1}^\top + (A_{j_1} \widehat{\mu}_0 - \widehat{\mu}_1)(A_{j_1} \widehat{\mu}_0 - \widehat{\mu}_1)^\top \right) \right). \end{aligned} \quad (22)$$

By exactly the same procedure in (20), and the term cancellation, we arrive at

$$\begin{aligned} \widehat{W}^2(1) &= \sum_{j_1=1}^m \pi_{j_1}(1) \left(\text{tr} \left(A_{j_1} \widehat{\mu}_0 \widehat{\mu}_0^\top A_{j_1}^\top + A_{j_1} \widehat{\Sigma}_0 A_{j_1}^\top \right) \right) \\ &\stackrel{(17)}{=} \sum_{j_1=1}^m \pi_{j_1}(1) \left(\text{tr} \left(A_{j_1} \left(\sum_{j_0=1}^{m_0} \alpha_{j_0} (\mu_{j_0} \mu_{j_0}^\top + \Sigma_{j_0}) \right) A_{j_1}^\top \right) \right) \\ &= \sum_{j_1=1}^m \sum_{j_0=1}^{m_0} \pi_{j_1}(1) \alpha_{j_0} \left(\|\mu_{j_1}\|_{\ell_2(\mathbb{R}^n)}^2 + \text{tr}(\Sigma_{j_1}) \right) \\ &= \sum_{j_1=1}^m \sum_{j_0=1}^{m_0} \pi_{j_1}(1) \alpha_{j_0} W_{j_1}^2(1) \stackrel{(14)}{=} W^2(1) \end{aligned} \quad (23)$$

where $\mu_{j_1} = A_{j_1} \mu_{j_0}$ and $\Sigma_{j_1} = A_{j_1} \Sigma_{j_0} A_{j_1}^\top$.

Continuing in this manner, finally we obtain a following result for any time k .

$$\begin{aligned} \widehat{W}^2(k) &= \sum_{j_k=1}^m \cdots \sum_{j_1=1}^m \sum_{j_0=1}^{m_0} \left(\prod_{r=1}^k \pi_{j_r}(r) \right) \alpha_{j_0} \\ &\quad \left(\|\mu_{j_k}\|_{\ell_2(\mathbb{R}^n)}^2 + \text{tr}(\Sigma_{j_k}) \right) \\ &= \sum_{j_k=1}^m \cdots \sum_{j_1=1}^m \sum_{j_0=1}^{m_0} \left(\prod_{r=1}^k \pi_{j_r}(r) \right) \alpha_{j_0} W_{j_k}^2(k) \\ &\stackrel{(14)}{=} W^2(k) \end{aligned} \quad (24)$$

where $\mu_{j_k} = A_{j_k} A_{j_{k-1}} \cdots A_{j_1} \mu_{j_0} = A_{j_k}^* \mu_{j_0}$,

$$\Sigma_{j_k} = (A_{j_k} A_{j_{k-1}} \cdots A_{j_1}) \Sigma_{j_0} (A_{j_k} A_{j_{k-1}} \cdots A_{j_1})^\top = A_{j_k}^* \Sigma_{j_0} A_{j_k}^{*\top}. \quad \square$$

According to Theorem 3, it is unnecessary to propagate the state PDF $\rho(x)$ and to compute W , which is intractable due to the exponential growth of Gaussian components. Instead, we can analyse the performance of the SJLS through \widehat{W} , since \widehat{W} is equidistant with W at all time k . The major advantages of the “Split-and-Merge” algorithm with \widehat{W} computation for the performance and the robustness analysis can be summarized in the following sense. \widehat{W} computation using (8) provides an analytical solution, which is computationally concise

and efficient enough. In addition, at any time step, we only have m mean vectors and covariance matrices to work with, and hence the scalability problem with an exponential growth can be avoided.

Remark 4 (Applicability of performance and robustness measure to general SJLSs) Since the switching probability $\pi(k)$ is an independent variable with regard to $\widehat{W}(k)$ as described in Theorem 3, we can compute $\widehat{W}(k)$ for any SJLSs regardless of the updating rule for $\pi(k)$. Once $\pi(k)$ is computed at time k by governing recursion equation (i.e., i.i.d., Markov, or semi-Markov jump process, etc.), the performance and the robustness for SJLSs are measured by $\widehat{W}(k)$. As a consequence, the proposed method for the performance and robustness measure can be applied to any SJLSs.

4 Numerical Example

Consider the inverted pendulum on cart in Fig. 2 with parameters described in Table 1. Originally, this example was introduced in [25] with single communication delay term τ_k between sensor and controller.

Table 1
Nomenclature for Inverted Pendulum Dynamics.

Symbol	definition	Symbol	definition
m_1	cart mass	m_2	pendulum mass
L	pendulum length	x	cart position
θ	pendulum angle	u	input force

The system states are $x_1 = x$, $x_2 = \dot{x}$, $x_3 = \theta$, and $x_4 = \dot{\theta}$. We assume that $m_1 = 1\text{kg}$, $m_2 = 0.5\text{kg}$, $L = 1\text{m}$ with friction-free floor. Later, this example was further exploited by [29] with two random delays τ_k and d_k which are sensor-to-controller and controller-to-actuator delays, respectively. The sets of mode are $\mathcal{M}(\tau_k) = \{0, 1, 2\}$ and $\mathcal{M}(d_k) = \{0, 1\}$. When the control action is taken at time k , the controller-to-actuator delay d_k is unknown, but τ_k and d_{k-1} are found. Accordingly, controller gain F is dependent on τ_k and d_{k-1} . Hence, the linearized closed-loop system model with sampling time $T_s = 0.1$ is denoted by

$$x(k+1) = Ax(k) + BF(\tau_k, d_{k-1})x(k - \tau_k - d_k)$$

where

$$A = \begin{bmatrix} 1 & 0.1 & -0.0166 & -0.0005 \\ 0 & 1 & -0.3374 & -0.0166 \\ 0 & 0 & 1.0996 & 0.1033 \\ 0 & 0 & 2.0247 & 1.0996 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0045 \\ 0.0896 \\ -0.0068 \\ -0.1377 \end{bmatrix}$$

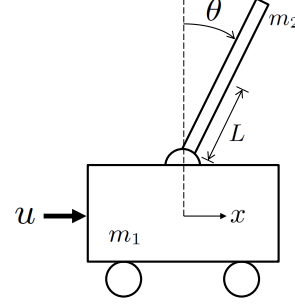


Fig. 2. Inverted Pendulum on Cart.

with the controller gain F 's given in [29]:

$$\begin{aligned} F(0, 0) &= \begin{bmatrix} 0.1690 & 0.8824 & 19.5824 & 4.3966 \end{bmatrix} \\ F(0, 1) &= \begin{bmatrix} 0.5625 & 0.6259 & 24.8814 & 5.1886 \end{bmatrix} \\ F(1, 0) &= \begin{bmatrix} -0.3076 & 0.9370 & 12.0069 & 5.9910 \end{bmatrix} \\ F(1, 1) &= \begin{bmatrix} -0.0097 & 0.7109 & 15.2518 & 7.3154 \end{bmatrix} \\ F(2, 0) &= \begin{bmatrix} -0.3212 & 1.0528 & 11.9330 & 6.3809 \end{bmatrix} \\ F(2, 1) &= \begin{bmatrix} 0.0427 & 0.8640 & 16.0874 & 7.8361 \end{bmatrix}. \end{aligned}$$

Therefore, this system has total 6 numbers of closed-loop dynamics A_{σ_k} with $\sigma_k \in \{1, 2, \dots, 6\}$.

1) Markovian Communication Delays:

We denote the transition probability of sensor-to-controller and controller-to-actuator delays as λ_{ij} and ω_{rs} , respectively. Then, λ_{ij} and ω_{rs} are defined by

$$\lambda_{ij} = \mathbb{P}(\tau_{k+1} = j | \tau_k = i), \quad \omega_{rs} = \mathbb{P}(\omega_{k+1} = s | \omega_k = r)$$

where $\lambda_{ij}, \omega_{rs} \geq 0$ and $\sum_{j=0}^2 \lambda_{ij} = 1$, $\sum_{s=0}^1 \omega_{rs} = 1$. Given individual Markov transition probability matrices

$$\Lambda = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.3 & 0.6 & 0.1 \\ 0.3 & 0.6 & 0.1 \end{bmatrix}, \quad \Omega = \begin{bmatrix} 0.2 & 0.8 \\ 0.5 & 0.5 \end{bmatrix}$$

corresponding to λ_{ij} and ω_{rs} , the Markov transition probability matrix P for 6 modes MJLS is obtained from $P = \Lambda \otimes \Omega$ as in [25]. The switching probability distribution $\pi(k)$ is updated by the linear recursion equation $\pi(k+1) = \pi(k)P$ with initial probability distribution $\pi(0) = [1, 0, 0, 0, 0, 0]$.

2) i.i.d. Communication Delays:

Although the previous examples in [25, 29] assumed that the communication delays are governed by Markov process, we adopt an i.i.d. jump process to manifestly show

that the proposed methods are also applicable to other types of SJLSs. In case of i.i.d. jump process, the switching probability distribution $\pi(k)$ is stationary, and hence it does not change over time. We assume that the switching probabilities π_{sc} and π_{ca} are given by

$$\pi_{sc} = [0.7, 0.2, 0.1], \quad \pi_{ca} = [0.5, 0.5]$$

where π_{sc} and π_{ca} stand for the switching probability distribution with respect to sensor-to-controller and controller-to-actuator, respectively. Then, the switching probability π for this inverted pendulum system is given by $\pi = \pi_{sc} \otimes \pi_{ca}$.

Differently from [29] where the initial state is deterministically given, we assume that the system contains initial state uncertainties as Gaussian distribution $\mathcal{N}(\mu(0), \Sigma(0))$ with $\mu(0) = [0, 0, 0.1, 0]^T$ and $\Sigma(0) = 0.25^2 I_{4 \times 4}$, where $I_{4 \times 4}$ denotes 4×4 identity matrix. Moreover, we tested the performance and robustness of this inverted pendulum system with an initial MoG PDF, which is given by a bimodal Gaussian in the following form

$$\rho(0) = \sum_{j=1}^2 \alpha_j(0) \mathcal{N}(\mu_j(0), \Sigma_j(0))$$

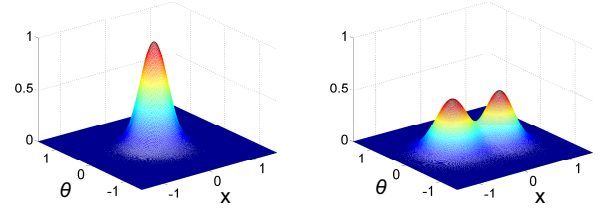
where $\alpha_1(0) = 0.5$ and $\alpha_2(0) = 0.5$. Mean and covariance for each Gaussian component are given by

$$\begin{aligned} \mu_1(0) &= [0.5, 0.25, -0.12, 0.05]^T, \quad \Sigma_1(0) = 0.25^2 I_{4 \times 4}, \\ \mu_2(0) &= [-0.4, 0.35, 0.07, -0.1]^T, \quad \Sigma_2(0) = 0.3^2 I_{4 \times 4}. \end{aligned}$$

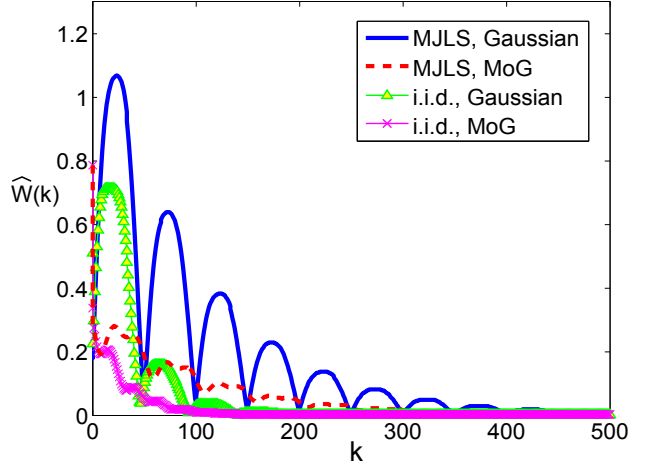
These types of multimodal uncertainties are caused by various factors such as sensing under interference [4], distributed sensor networks [12], multitarget tracking problems [16] and so forth. The bivariate marginal distribution associated with state x and θ for these Gaussian and MoG PDF are shown in Fig. 3(a) and Fig. 3(b), respectively.

In Fig. 3(c), the performance and the robustness of this inverted pendulum system with different stochastic jump processes and initial state uncertainties are depicted via \widehat{W} computation. For all cases, we know that the system is m.s. stable from the convergence of \widehat{W} . However, the rate of convergence and the performance show different aspects in the transient time. Among all cases, \widehat{W} for i.i.d. jump process with initial MoG PDF converges fast with small bounce, whereas \widehat{W} for MJLS with initial Gaussian PDF slowly converges with large bounce.

At every time step, the ‘‘Split-and-Merge’’ algorithm, presented in Section 3.2.2 is used to propagate the state



(a) Gaussian marginal distr. (b) MoG marginal distr.



(c) Wasserstein distance with different stochastic jump processes and initial PDFs; MJLS with Gaussian (blue solid), MJLS with MoG (red dashed), i.i.d. with Gaussian (green triangle), and i.i.d. with MoG (purple cross).

Fig. 3. Simulation Result for Performance and Robustness Analysis of Inverted Pendulum system with the existence of both random communication delays and initial state uncertainties.

PDFs. Without using these techniques, it is practically impossible to propagate density functions and calculate \widehat{W} (i.e., the Wasserstein distance between actual state PDF $\rho(x)$ and $\delta(x)$) even for a finite switching modes. The number of Gaussian components that represents the state PDF after N time steps is 6^N , which soon becomes computationally intractable. For an m -mode SJLS, the growth rate is m^N . With the implementation of the proposed ‘‘Split-and-Merge’’ algorithm, \widehat{W} that is equidistant with W was computed effectively and efficiently. From this example, it is clearly shown that the performance and the robustness for general SJLSs can be measured via \widehat{W} distance which quantifies the uncertainties.

5 Conclusion

In this paper, we proposed new tools for the performance and the robustness analysis of stochastic jump linear systems. With given initial state uncertainties, Wasserstein distance that compares shapes of PDFs provides a way to quantify the uncertainties. Since the growth of PDF components in stochastic jumps is exponential in time,

we presented a new “Split-and-Merge” algorithm for uncertainty propagation that scales linearly with the number of modes in the jump system. This method provides analytical solutions, while avoiding exponential growth of PDF components. The proposed methods are applicable not only to Markovian jumps, which is commonly assumed in the analysis of jump systems, but also to general stochastic jump linear systems. We also proved that mean square stability can be shown with regard to convergence of Wasserstein distance. These results address both transient and steady-state behavior of stochastic jump linear systems. The practical usefulness and efficiency of the proposed method are verified by examples.

References

- [1] Simo Ali-Löytty. On the convergence of the gaussian mixture filter. *Tampereen teknillinen yliopisto. Matematiikan laitos. Tutkimusraportti*; 89, 2011.
- [2] Fayer F Bector. Some efficient multi-heuristic procedures for resource-constrained project scheduling. *European Journal of Operational Research*, 49(1):3–13, 1990.
- [3] L. Coviello, P. Minero, and M. Franceschetti. Stabilization over Markov feedback channels. In *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pages 3776–3782. IEEE, 2011.
- [4] Lewis Girod and Deborah Estrin. Robust range estimation using acoustic and multimodal sensing. In *Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on*, volume 3, pages 1312–1320. IEEE, 2001.
- [5] Clark R Givens and Rae Michael Shortt. A class of wasserstein metrics for probability distributions. *The Michigan Mathematical Journal*, 31(2):231–240, 1984.
- [6] Geoffrey Grimmett and David Stirzaker. *Probability and random processes*, volume 2. Clarendon press Oxford, 1992.
- [7] A. Halder and R. Bhattacharya. Further results on probabilistic model validation in wasserstein metric. In *51st IEEE Conference on Decision and Control, Maui*, 2012.
- [8] Sadri Hassani. Mathematical physics, a modern introduction to its foundations, 1999.
- [9] A. Hassibi, S.P. Boyd, and J.P. How. Control of asynchronous dynamical systems with rate constraints on events. In *Decision and Control, 1999. Proceedings of the 38th IEEE Conference on*, volume 2, pages 1345–1351. IEEE, 1999.
- [10] Zhenting Hou, Jiaowan Luo, Peng Shi, and Sing Kiong Nguang. Stochastic stability of ito differential equations with semi-markovian jump parameters. *IEEE Transactions on Automatic Control*, 51(8):1383, 2006.
- [11] Mehmet Karan, Peng Shi, and C Yalçın Kaya. Transition probability bounds for the stochastic stability robustness of continuous-and discrete-time markovian jump linear systems. *Automatica*, 42(12):2159–2168, 2006.
- [12] Koen Langendoen and Niels Reijers. Distributed localization in wireless sensor networks: a quantitative comparison. *Computer Networks*, 43(4):499–518, 2003.
- [13] Ji-Woong Lee and Geir E Dullerud. Uniform stabilization of discrete-time switched and markovian jump linear systems. *Automatica*, 42(2):205–218, 2006.
- [14] Fanbiao Li, Ligang Wu, and Peng Shi. Stochastic stability of semi-markovian jump systems with mode-dependent delays. *International Journal of Robust and Nonlinear Control*, 2013.
- [15] Hai Lin and Panos J Antsaklis. Stability and stabilizability of switched linear systems: a survey of recent results. *Automatic control, IEEE Transactions on*, 54(2):308–322, 2009.
- [16] Juan Liu, Maurice Chu, and James E Reich. Multitarget tracking in distributed sensor networks. *Signal Processing Magazine, IEEE*, 24(3):36–46, 2007.
- [17] Ming Liu, Daniel WC Ho, and Yugang Niu. Stabilization of markovian jump linear system over networks with random communication delay. *Automatica*, 45(2):416–421, 2009.
- [18] KD Minto and R Ravi. New results on the multi-controller scheme for the reliable control of linear plants. In *American Control Conference, 1991*, pages 615–619. IEEE, 1991.
- [19] Kumpati S Narendra and Jeyendran Balakrishnan. Improving transient response of adaptive control systems using multiple models and switching. *Automatic Control, IEEE Transactions on*, 39(9):1861–1866, 1994.
- [20] David W Scott. Multivariate density estimation. *Multivariate Density Estimation, Wiley, New York*, 1992, 1, 1992.
- [21] D Michael Titterton, Adrian FM Smith, Udi E Makov, et al. *Statistical analysis of finite mixture distributions*, volume 7. Wiley New York, 1985.
- [22] C. Villani. *Topics in optimal transportation*, volume 58. Amer Mathematical Society, 2003.
- [23] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer, 2008.
- [24] Jing Wu and Tongwen Chen. Design of networked control systems with packet dropouts. *Automatic Control, IEEE Transactions on*, 52(7):1314–1319, 2007.
- [25] Lin Xiao, Arash Hassibi, and Jonathan P How. Control with random communication delays via a discrete-time jump system approach. In *American Control Conference, 2000. Proceedings of the 2000*, volume 3, pages 2199–2204. IEEE, 2000.
- [26] Junlin Xiong and James Lam. Stabilization of discrete-time markovian jump linear systems via time-delayed controllers. *Automatica*, 42(5):747–753, 2006.
- [27] Junlin Xiong and James Lam. Stabilization of linear systems over networks with bounded packet loss. *Automatica*, 43(1):80–87, 2007.
- [28] K. You and L. Xie. Minimum data rate for mean square stabilizability of linear systems with markovian packet losses. *Automatic Control, IEEE Transactions on*, 56(4):772–785, 2011.
- [29] Liqian Zhang, Yang Shi, Tongwen Chen, and Biao Huang. A new method for stabilization of networked control systems with random delays. *Automatic Control, IEEE Transactions on*, 50(8):1177–1181, 2005.
- [30] Lixian Zhang and El-Kébir Boukas. Stability and stabilization of markovian jump linear systems with partly unknown transition probabilities. *Automatica*, 45(2):463–468, 2009.