A Switched Dynamical System Framework for Analysis of Massively Parallel Asynchronous Numerical Algorithms

Kooktae Lee, Raktim Bhattacharya, and Vijay Gupta

Abstract—In the near future, massively parallel computing systems will be necessary to solve computation intensive applications. The key bottleneck in massively parallel implementation of numerical algorithms is the synchronization of data across processing elements (PEs) after each iteration, which results in significant idle time. Thus, there is a trend towards relaxing the synchronization and adopting an asynchronous model of computation to reduce idle time. However, it is not clear what is the effect of this relaxation on the stability and accuracy of the numerical algorithm.

In this paper we present a new framework to analyze such algorithms. We treat the computation in each PE as a dynamical system and model the asynchrony as stochastic switching. The overall system is then analyzed as a switched dynamical system. However, modeling of massively parallel numerical algorithms as switched dynamical systems results in a very large number of modes, which makes current analysis tools available for such systems computationally intractable. We develop new techniques that circumvent this scalability issue. The framework is presented on a one-dimensional heat equation and the proposed analysis framework is verified by solving the partial differential equation (PDE) in an NVIDIA Tesla™ GPU machine, with asynchronous communication between cores.

I. INTRODUCTION

Exascale computing systems will soon be available to study computation intensive applications such as multiphysics multi-scale simulations of natural and engineering systems. Many scientific and practical problems can be described very accurately by ordinary or partial differential equations which can be tightly coupled with long-range correlations. These exascale systems may have \( O(10^5 - 10^6) \) processors ranging from multicore processors to symmetric multiprocessors [1]–[3]. Furthermore, such systems are likely to be heterogeneous using both heavily multi-threaded CPUs as well as GPUs. Many challenges must be overcome before exascale systems can be utilized effectively in such applications. One such obstacle is the communication in tightly coupled problems during parallel implementation of any iterative numerical algorithm. This communication requires massive data movement in turn leading to idle time as the cores need to be synchronized after each time step.

Recent literature has proposed relaxing these synchronization requirements across the PEs [4]. This potentially eliminates the overhead associated with extreme parallelism and significantly reduces computational time. However, the price to pay is loss of predictability possibly resulting in calculation errors. Thus, a rigorous analysis of the tradeoff between speed and accuracy is critical. This paper presents a framework to quantify this tradeoff by analyzing the asynchronous numerical algorithm as a switched dynamical system [5]–[14]. While there is a large literature for analysis of such systems, these techniques are not applicable to our application. The reason is that due to the large number of PEs, the switched system model has an extremely large number of modes, which makes the available analysis tools intractable. Key contributions in this paper include new techniques for a) stability analysis, or quantification of steady-state error with respect to the synchronous solution; b) convergence rate analysis of the expected value of this error; and c) probabilistic bounds on this error. These techniques are developed to be computationally efficient, and avoid the aforementioned scalability issue.

The paper is organized as follows. Section II addresses the problems for the asynchronous numerical algorithm. In section III, we introduce a switched system framework to model the system structure for the asynchronous numerical scheme. The stability results are presented in section IV, and section V shows the convergence rate analysis. Then, the error analysis in probability is developed in section VI. Section VII demonstrates the usefulness of the proposed method by examples. Finally, section VIII concludes this paper.

II. PROBLEM FORMULATION

Notation: The symbol \( \| \cdot \| \) and \( \| \cdot \|_\infty \) stand for the Euclidean and infinity norm, respectively. The set of positive integers are denoted by \( \mathbb{N} \). Further, \( \mathbb{N}_0 = \mathbb{N} \cup \{0\} \). Also, \( \lambda(\cdot) \) represents an eigenvalue of a square matrix. In particular, \( \lambda_{\max}(\cdot) \) and \( \lambda_{\min}(\cdot) \) denote the largest and the smallest eigenvalue in magnitude, respectively. The symbols \( \otimes, \det(\cdot), \text{tr}(\cdot), \) and \( \text{vec}(\cdot) \) denote Kronecker product, matrix determinant, trace operator, and vectorization operator, respectively. Finally, the symbol \( \mathbb{P}\text{r}(\cdot) \) stands for the probability.

In this paper we demonstrate our framework and techniques on the one-dimensional heat equation, given by

\[
\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2}, \quad t \geq 0,
\]

where \( u \) is the time and space-varying state of the temperature, and \( t \) and \( x \) are continuous time and space respectively. The constant \( \alpha > 0 \) is the thermal diffusivity of the given material.
The PDE is solved numerically using the finite difference method by Euler explicit scheme, with a forward difference in time and a central difference in space. Thus (1) is approximated as
\[
\frac{u_i(k+1) - u_i(k)}{\Delta t} = \alpha \left( \frac{u_{i+1}(k) - 2u_i(k) + u_{i-1}(k)}{\Delta x^2} \right),
\]
where \( k \in \mathbb{N}_0 \) is the discrete-time index and \( u_i \) is the temperature value at \( j^{th} \) grid space point. The symbols \( \Delta t \) and \( \Delta x \) denote the sampling time and the grid resolution in space, respectively. Further, if we define a constant \( r \triangleq \alpha \frac{\Delta t}{\Delta x^2} \), then (2) can be written as
\[
u_i(k+1) = ru_{i+1}(k) + (1 - 2r)u_i(k) + ru_{i-1}(k), \tag{3}
\]
It is important to observe that (3) is a discrete-time linear dynamical system.

**Fig. 1.** Discretized one-dimensional domain with an asynchronous numerical algorithm. The PE denotes a group of grid points, assigned to each core.

Fig. 1 illustrates the numerical scheme over the discretized 1D spatial domain. A typical synchronous parallel implementation of this numerical scheme assigns several of these grid points to each PE. The updates for the temperature at the grid points assigned to each PE occur in parallel. However, at every time step \( k \), the data associated with the boundary grid points, where the communication is necessary are synchronized, and used to compute \( u_i(k+1) \). This synchronization across PEs is slow, especially for massively parallel systems (estimates of idle time due to this synchronization give figures of up to 80% of the total time taken for the simulation as idle time). Recently, an alternative implementation which is asynchronous has been proposed. In this implementation, the updates in a PE occur without waiting for the other PEs to finish and their results to be synchronized. The data update across PEs occurs sporadically and independently. This asynchrony directly affects the update equation for the boundary points, as they depend on the grid points across PEs. For these points, the update is performed with the most recent available value, typically stored in a buffer. The effect of this asynchrony then propagates to other grid points. Within a PE, we assume there is no asynchrony and data is available in a common memory.

Thus, the asynchronous numerical scheme corresponding to (3) is given by
\[
u_i(k+1) = ru_{i+1}(k) + (1 - 2r)u_i(k) + ru_{i-1}(k), \tag{4}
\]
where \( k_i^* \in \{k-1, k-2, \ldots, k-q+1\}, i = 1, 2, \ldots, N \), denotes the randomness caused by communication delays between PEs. The subscript \( i \) in \( k_i^* \) depicts that each grid space point may have different time delays. The parameter \( q \) is the length of a buffer that every core maintains to store data transmitted from the other cores. In this paper, we treat \( k_i^* \) as a random variable and thus (4) can be considered to be a linear discrete-time dynamical system with stochastic updates.

Although (4) is derived for the 1D heat equation, the treatment above can be developed for any parabolic PDEs. This observation encourages us to consider using tools from dynamical systems to analyze the effect of asynchrony in parallel numerical algorithms. Therefore, the primary goal of this study is to investigate the stability, convergence rate, and error probability of the asynchronous numerical algorithm in the framework of stochastic switched dynamical systems.

### III. A Switched System Approach

Let us define the state vector \( U_j(k) \in \mathbb{R}^n \triangleq [u_1^j(k), u_2^j(k), \ldots, u_n^j(k)]^\top \), where \( u_i^j(k) \) stands for the \( i^{th} \) grid point in the \( j^{th} \) PE and \( n \) is the total number of grid points in the \( j^{th} \) PE. Therefore, (3) can be compactly written as
\[
U(k+1) = AU(k), \quad k \in \mathbb{N}_0,
\]
where \( U(k) \in \mathbb{R}^{Nn} \triangleq [U_1(k)^\top, U_2(k)^\top, \ldots, U_N(k)^\top]^\top \), \( N \) is the total number of PEs, \( n \) is the size of the state for each PE, and system matrix \( A \in \mathbb{R}^{Nn \times Nn} \) is given by
\[
A = \begin{bmatrix}
1 & 0 & 0 & \cdots & 0 \\
r & 1-2r & r & \cdots & 0 \\
0 & r & 1-2r & r & \cdots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \cdots & r & 1-2r & r \\
0 & \cdots & 0 & 0 & 1
\end{bmatrix} \in \mathbb{R}^{Nn \times Nn}.
\]

Note that the first and the last row of \( A \) matrix specify the Dirichlet boundary conditions (see pp. 150, [15]), i.e., we have the constant in time boundary temperatures for simplicity.

Next, we define the augmented state \( X(k) \in \mathbb{R}^{Nnq} \triangleq [U(k)^\top, U(k-1)^\top, \ldots, U(k-q+1)^\top]^\top \), where, as stated before, \( q \) is the buffer length. For pedagogical simplicity (and without loss of generality), we consider the case with \( q = 2 \) and \( N = 3 \). Further, we let \( n = 1 \), which implies there is only one grid point in each PE. For this particular case, we construct the following matrices,
\[
W_1 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
r & 1-2r & r & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}, \quad W_2 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1-2r & r & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix},
\]
\[
W_3 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
r & 1-2r & r & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}, \quad W_4 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1-2r & r & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}.
\]
where \( \mathbf{I} \in \mathbb{R}^{Nn \times Nn} \) and \( \mathbf{0} \in \mathbb{R}^{Nn \times Nn} \) are the identity and the zero matrices with appropriate dimensions. As in [4], we assume that the condition \( 0 < r \leq 0.5 \) holds from now on. The asynchronous numerical scheme can then be written as a switched system

\[
X(k + 1) = W_{\sigma_k}X(k), \quad \sigma_k \in \{1, 2, \ldots, m\}, \quad k \in \mathbb{N}_0, \quad (5)
\]

where the matrices \( W_{\sigma_k} \in \mathbb{R}^{Nnq \times Nnq} \) are the subsystem dynamics. In general, the total number of switching modes is \( m = q^{2(N-2)} \) that is obtained by considering all cases to distribute every component \( r \) in \( W_{r} \) matrix, where the number of \( r \) in \( W_{1} \) is given by \( 2(N-2) \), into \( q \) numbers of zero block matrix as in the above example. Therefore, the number of modes increase exponentially with the number of PEs, which is quite large for massively parallel systems.

At every time step, the numerical scheme evolves using one of the \( m \) modes, which depends on the variable \( k^* \). In this paper, we model the variable \( k^* \) as a random variable that evolves in an independently and identically distributed (i.i.d.) fashion in time, and independently from one core to the next. Hence, we let \( \pi_j \) be the modal probability for \( W_j \) which is assumed to be stationary in time. Let \( \Pi \triangleq \{\pi_1, \pi_2, \ldots, \pi_m\} \), be the switching probabilities such that \( 0 \leq \pi_j \leq 1, \forall j \) and \( \sum_{j=1}^{m} \pi_j = 1 \). The system in (5) is thus an i.i.d. jump linear system, which is a simpler case of the more well-known Markovian jump linear systems [13]. Even though the analysis theory for such systems is well developed, the existing tools are not suitable for our application because of the extremely large number of modes, particularly when \( N \) is large. Thus, we now develop an analysis theory for the i.i.d. jump linear systems which scales better with respect to the number of modes.

IV. STABILITY

The first requirement is that of convergence of (5). Because of the Dirichlet boundary conditions, we expect the temperature to converge to a constant value for every grid point. We proceed to analyze the conditions for convergence (or stability) of the system. To this end, we may try to use the infinity norm and apply the sub-multiplicative property to obtain

\[
||X(k + 1)||_\infty = ||W_{\sigma_k}X(k)||_\infty \leq ||W_{\sigma_k}||_\infty ||X(k)||_\infty = ||X(k)||_\infty,
\]

where in above equation, \( g \) represents all the other multiplication terms except \( \Psi^k \) term. Note that \( g(k) \) is formed

where \( \lambda_i^j \in \mathbb{R}, v_i^j \in \mathbb{R}^{Nnq \times 1} \), and \( s_i^j \in \mathbb{R}^{1 \times Nnq} \) denote the eigenvalues, right eigenvectors, and left eigenvectors of \( W_j \), respectively.

Since \( \max |\lambda_i^j| \leq ||W_j||_\infty = 1, \forall j \), the spectral radius of \( W_j, j = 1, 2, \ldots, m \), is less than or equal to 1. Therefore, we may order the eigenvalues as \( 1 \geq |\lambda_i^j| \geq |\lambda_i^j| > \cdots > |\lambda_i^{Nnq}| \geq 0 \). It can be shown that all \( W_j \) have two eigenvalues with value 1, irrespective of the size of \( q \) and \( N \). Therefore, the eigenvalues for \( W_j \) are ordered as \( 1 = |\lambda_i^j| = |\lambda_i^j| > |\lambda_i^j| > \cdots > |\lambda_i^{Nnq}| \geq 0 \).

Moreover, the left and right eigenvectors for eigenvalues equal to 1 are common eigenvectors for all matrices \( W_j, j = 1, 2, \ldots, m \). These common left and right eigenvectors are

1) Left eigenvectors:

\[
\begin{align*}
    s_1 &= [1, 0, \ldots, 0, 0, \ldots, 0] \in \mathbb{R}^{1 \times Nnq}, \\
    s_2 &= [0, \ldots, 0, 1, 0, \ldots, 0] \in \mathbb{R}^{1 \times Nnq},
\end{align*}
\]

2) Right eigenvectors:

\[
\begin{align*}
    v_1 &= [\mu_1, \mu_1, \ldots, \mu_1]^T \in \mathbb{R}^{Nnq \times 1}, \\
    v_2 &= [\mu_2, \mu_2, \ldots, \mu_2]^T \in \mathbb{R}^{Nnq \times 1},
\end{align*}
\]

where \( \mu_1 \triangleq [1, \frac{Nn-2}{Nn-1}, \ldots, \frac{Nn-2}{Nn-1}, 0] \in \mathbb{R}^{1 \times Nn}, \mu_2 \triangleq [0, \frac{1}{Nn-1}, \ldots, \frac{1}{Nn-1}, 1] \in \mathbb{R}^{1 \times Nn} \), \( j = 1, 2, \ldots, Nn \).

Notice that we have \( W_j s_i = v_i \) and \( s_i W_j = s_i, i = 1, 2, \ldots, Nn \).

Then, the steady-state value for the asynchronous scheme is given by the following result.

*Proposition 4.1*: Consider the i.i.d. jump linear system in (5) with subsystem matrices \( W_j, j = 1, 2, \ldots, m \) and a stationary switching probability \( \Pi \). For a given initial condition \( X(0) \), if we define \( \Psi \triangleq v_1 s_1 + v_2 s_2 \), where \( v_1 \) and \( s_1 \) are given in (8)–(11), the steady-state value \( X_{ss} \) has the following form:

\[
X_{ss} \triangleq \lim_{k \to \infty} X(k) = \Psi X(0),
\]

irrespective of the switching sequence \( \{\sigma_k\} \).

*Proof*: Let the eigenvalues of \( W_j \) be ordered in magnitude by \( 1 = |\lambda_i^j| > |\lambda_i^{j+1}| > \cdots > |\lambda_i^{Nnq}| \geq 0 \). Also, let \( v_i^j \) and \( s_i^j \) be the right and left eigenvector corresponding to \( \lambda_i^j \), respectively. Using the spectral decomposition, \( W_j \) can be alternatively expressed by

\[
W_j = \sum_{\lambda_i^j \neq 1} \lambda_i^j v_i^j s_i^j = \Psi + \sum_{\lambda_i^j \neq 1} f^j(i), \quad \text{where} \quad \Psi \triangleq v_1 s_1 + v_2 s_2 \quad \text{and} \quad f^j(i) \triangleq \lambda_i^j v_i^j s_i^j.
\]

Then, starting with \( X(0) \), the realization of the switching sequence \( \sigma_k \) results in

\[
X(k) = W_{\sigma_{k-1}} W_{\sigma_{k-2}} \cdots W_{\sigma_1} W_{\sigma_0} X(0) = \left( \Psi + \sum_{\lambda_i^{\sigma_{k-1}} \neq 1} f^{\sigma_{k-1}}(i) \right) \cdots \left( \Psi + \sum_{\lambda_i^{\sigma_0} \neq 1} f^{\sigma_0}(i) \right) X(0) = \left( \Psi^k + g(k) \right) X(0),
\]

where in above equation, \( g(k) \) represents all the other multiplication terms except \( \Psi^k \) term.
by the product of \( \lambda_i^j \), where \( 0 \leq |\lambda_i| < 1, \forall i > 2, \forall j \). Consequently, if \( k \to \infty \), then \( g(k) \) is asymptotically convergent to zero since the infinite number of multiplication of the term \( \lambda_i^j \), \( \forall i > 2 \), converges to zero. Therefore, we have

\[
X_{ss} = \lim_{k \to \infty} X(k) = \lim_{k \to \infty} \Psi^k X(0) = \Psi X(0).
\]

The last equality in above equation holds because \( \Psi^k = \Psi^{k-1} = \cdots = \Psi, \forall k \in \mathbb{N} \).

V. CONVERGENCE RATE

In this section, we investigate how fast the expected value of the state converges to the steady-state \( X_{ss} \) by analyzing the transient behavior of the asynchronous algorithm. Let us define a new state variable \( e(k) \triangleq X(k) - X_{ss} \). The expected value of \( e(k) \) is given by \( \bar{e}(k) \triangleq \mathbb{E}[X(k) - X_{ss}] = \mathbb{E}[X(k)] - X_{ss} = X(k) - X_{ss} \), where \( X(k) \) is the \( i.i.d. \) process with a given deterministic initial error \( e(0) \). Moreover, extremely large values of \( e(k) \) are given by (13) with the switching probability \( \pi \). Let us define a new state variable \( e(k) \) given by the product of \( \lambda_i^j \) and \( \sigma_j \), where \( \lambda_i^j \) and \( \sigma_j \) denote the eigenvalues, right and, respectively, left eigenvectors for \( W_j \). If we define the transformed matrix \( \bar{W}_j \) by \( \bar{W}_j = W_j - \Psi = \sum_{\lambda_i^j}^{\lambda_i^j \neq 1} \lambda_i^j v_i^j s_i^j \), then the modal dynamics with the corresponding state \( e_j(k) \), is given by

\[
e_j(k + 1) = \bar{W}_j e_j(k), \quad j = \{1, 2, \ldots, m\}, \quad k \in \mathbb{N}_0. \tag{12}\]

Moreover, as in (5), the error state \( e(k) = X(k) - X_{ss} \) is governed by

\[
e(k + 1) = \bar{W}_\sigma e(k), \quad \sigma_k \in \{1, 2, \ldots, m\}, \quad k \in \mathbb{N}_0. \tag{13}\]

The system in (13) is also a switched linear system. The transformed matrix \( \bar{W}_j \) are the modes of the error dynamics. Generally, it is difficult to estimate the convergence rate of the ensemble with stochastic jumps. Previous works [16–19] have used the common Lyapunov function approaches, to analyze stability and the convergence rate. However, the existence of a common Lyapunov function is the only sufficient condition for the system stability, and hence there may not exist a common Lyapunov function for the asynchronous algorithm. Moreover, extremely large values of \( m \) make it very difficult to test every conditions for the existence of such a common Lyapunov function. For this reason, we bound the convergence rate of \( \bar{e}(k) \), instead of bounding \( e(k) \) directly.

**Lemma 5.1:** Consider an i.i.d. jump linear system given by (13) with the switching probability \( \Pi = \{\pi_1, \pi_2, \ldots, \pi_m\} \). If the initial state \( e(0) \) is given and has no uncertainty, the expected value of \( e(k) \) is updated by

\[
\bar{e}(k) \triangleq \mathbb{E}[e(k)] = \Lambda^k \bar{e}(0) \quad \text{or} \quad \bar{e}(k + 1) = \Lambda \bar{e}(k), \tag{14}\]

where \( \Lambda \triangleq \sum_{i=1}^{m} \pi_i \bar{W}_j \).

Proof: For an i.i.d. jump linear system with a given deterministic initial error \( e(0) \), we have

\[
\mathbb{E}[e(k)] = \mathbb{E}[ar{W}_\sigma e(k-1)]
= \mathbb{E}[\bar{W}_\sigma \bar{W}_\sigma \ldots \bar{W}_\sigma \bar{W}_\sigma e(0)]
= \mathbb{E}[\bar{W}_\sigma] \Lambda^k \mathbb{E}[\bar{W}_\sigma] e(0) = \Lambda^k \bar{e}(0). \tag{15}\]

Since the matrix \( \Lambda \) is given by \( \Lambda = \sum_{i=1}^{m} \pi_i \bar{W}_j \), the computation of \( \Lambda \) requires all matrices \( W_j, j = 1, 2, \ldots, m \). As pointed out earlier, this calculation is intractable due to the extremely large number of the switching modes \( m \). Therefore, instead of using (14), we provide a computationally efficient method to bound \( ||\bar{e}(k)|| \) through a Lyapunov theorem.

Consider a discrete-time Lyapunov function \( V(k) = \bar{e}(k)^T \bar{P} \bar{e}(k) \), where \( \bar{P} \) is a positive definite matrix. Since it is shown that the original state \( X(k) \) is convergent to the unique steady-state \( X_{ss} \) as \( k \to \infty \) irrespective of \( \{\sigma_k\} \), the expected error \( \bar{e}(k) \triangleq X(k) - X_{ss} \) is asymptotically stable. Therefore, one can employ the Converse Lyapunov Theorem [20], which guarantees the existence of a positive definite matrix \( P \), satisfying the following linear matrix inequality (LMI) condition \( \Lambda^T \bar{P} \Lambda - \bar{Q} < 0 \), where \( \bar{Q} \) is some positive definite matrix. The matrix inequality can be interpreted in the sense of positive definiteness. (i.e., \( A > B \) means the matrix \( A - B \) is positive definite.) Then, the above LMI condition results in \( \Delta V(k) = V(k+1) - V(k) = \bar{e}(k)^T \Lambda^T \bar{P} \Lambda - \bar{Q} \). Hence, \( \bar{e}(k) \) is bounded by the following equation:

\[
\|\bar{e}(k)\|^2 < K \left( 1 - \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} \right)^k \|e(0)\|^2, \tag{16}\]

where \( K > 0 \) is some constant.

Next, we bound the convergence rate for \( ||\bar{e}(k)|| \) by using the result in (16) as follows.

**Proposition 5.1:** For a stable i.i.d. jump linear system (13) with a stationary switching probability \( \Pi \), consider a Lyapunov candidate function for the state \( \bar{e} \), given by \( V \triangleq \bar{e}^T \bar{P} \bar{e} \), where \( \bar{P} \) is a positive definite matrix. In addition, a Lyapunov candidate function for (12) is given by \( V_j \triangleq e_j^T P_j e_j, \quad j = 1, 2, \ldots, m \), where \( P_j \) is a positive definite matrix. According to the Converse Lyapunov Theorem, there exist \( P_j > 0 \) and \( \bar{P} > 0 \) such that \( W_j^T P_j W_j - \bar{P} < -\bar{Q}, \quad j = 1, 2, \ldots, m \) and \( \Lambda^T \bar{P} \Lambda - \bar{Q} < 0 \), where \( \bar{Q} \) and
$Q$ are any positive definite matrices. Then, with a particular choice of these matrices, we assume that $P_j$ and $P$ satisfy the following conditions:

\[
\begin{align*}
\tilde{W}_j^T P_j \tilde{W}_j - P_j &= -I, \quad j = 1, 2, \ldots, m, \quad (17) \\
\Lambda^T P \Lambda - P &\leq -\varepsilon_j I, \quad \text{for some } j, \quad (18)
\end{align*}
\]

where $\varepsilon_j \triangleq \frac{\lambda_{max}(P)}{\lambda_{max}(P_j)} > 0$, $\tilde{W}_j$ are the modal matrices in (12), and $\Lambda \triangleq \sum_{j=1}^{m} \pi_j \tilde{W}_j$.

Then, $\|\bar{e}(k)\|^2$ is bounded by

\[
\|\bar{e}(k)\|^2 < K \left(1 - \frac{1}{\lambda_{max}(P_j)}\right)^k \|e(0)\|^2, \quad (19)
\]

where $K > 0$ is some constant.

**Proof:** By applying the result in (16) into (18), we have

\[
\|\bar{e}(k)\|^2 < K \left(1 - \frac{\varepsilon_j}{\lambda_{max}(P)}\right)^k \|e(0)\|^2
\]

\[
= K \left(1 - \frac{\varepsilon_j}{\lambda_{max}(P)}\right)^k \|e(0)\|^2
\]

\[
= K \left(1 - \frac{1}{\lambda_{max}(P_j)}\right)^k \|e(0)\|^2.
\]

The last equality in above equation holds by the definition of $\varepsilon_j$.

Proposition 5.1 says that we can always guarantee the bound for $\|\bar{e}(k)\|$ if (18) holds. Consequently, the existence of such a $P$, satisfying (18) is the major concern in order to guarantee the bound $\|\bar{e}(k)\|$. The following lemma and theorem can be used to prove the existence of such a $P$.

**Lemma 5.2:** Suppose that $P_j$ is a positive definite matrix, satisfying (17). Then, the largest eigenvalue of $P_j$ is strictly greater than 1 for all $j$, i.e., $\lambda_{max}(P_j) > 1$, $\forall j$.

**Proof:** From (17), $P_j = \tilde{W}_j^T P \tilde{W}_j + I$, $\forall j$. Then, with the eigenvectors $y \in \mathbb{R}^{Nq}$ of $P_j$, the largest eigenvalue of $P_j$ is given by its definition as follows:

\[
\lambda_{max}(P_j) = \lambda_{max}(\tilde{W}_j^T P \tilde{W}_j + I) = \max_{\|y\|^2=1} y^T (\tilde{W}_j^T P \tilde{W}_j + I) y
\]

\[
= \max_{\|y\|^2=1} \left(\tilde{W}_j^T P \tilde{W}_j y\right) y^T y
\]

Since $P_j$ is a positive definite matrix, $\tilde{W}_j^T P \tilde{W}_j$ becomes a positive semi-definite matrix at least. Then, the scalar term $y^T \tilde{W}_j^T P \tilde{W}_j y$ cannot be zero unless $\tilde{W}_j^T P \tilde{W}_j$ is a zero matrix or a triangular matrix with zero diagonal components, which is not the case. Hence, it is guaranteed that $y^T \tilde{W}_j^T P \tilde{W}_j y > 0$, implying $\lambda_{max}(P_j) > 1$, $\forall j$.

**Theorem 5.1:** Consider Lyapunov functions for (12) and (13) given by $V_j \triangleq \varepsilon_j P_j e_j$, $j = 1, 2, \ldots, m$, and $V \triangleq \bar{e}^T \bar{P} \bar{e}$, respectively, where the matrices $P_j > 0$, $\forall j$ and $P > 0$. By the Converse Lyapunov Theorem, we assume that the matrices $P_j$, $\forall j$, satisfies the condition (17).

Then, there exists a positive definite matrix $P$ such that

\[
\Lambda^T P \Lambda - P \leq -\varepsilon_j I, \quad \text{for some } j, \quad (20)
\]

where $\varepsilon_j = \frac{\lambda_{max}(P)}{\lambda_{max}(P_j)} > 0$.

**Proof:** We prove by contradiction. Suppose that there exist no such $P > 0$, satisfying (20), which is equivalent to that for all matrices $P > 0$, the inequality $\Lambda^T P \Lambda - P \geq -\varepsilon_j I$ holds $\forall j$. The above inequality can be interpreted in the quadratic sense. In other words, for any non-zero vector $v$ that has a proper dimension, the following condition holds:

\[
v^T (\Lambda^T P \Lambda - P + \varepsilon_j I) v > 0, \quad \forall j \quad (21)
\]

As a particular choice of $v$, we let the vector $v$ be the eigenvector of the matrix $\Lambda$, i.e., $\Lambda v = \lambda v$, where $\lambda$ is the eigenvalue of $\Lambda$. Since (21) holds for any matrix $P > 0$, we let $P = I$, which results in $\varepsilon_j = \frac{\lambda_{max}(I)}{\lambda_{max}(P_j)} = \frac{1}{\lambda_{max}(P_j)}$.

Hence, we have

\[
0 < v^T (\Lambda^T \Lambda - I + \frac{1}{\lambda_{max}(P_j)} I) v
\]

\[
= (\Lambda v)^T (\Lambda v) - \|v\|^2 + \frac{1}{\lambda_{max}(P_j)} \|v\|^2
\]

\[
= \lambda^2 - 1 + \frac{1}{\lambda_{max}(P_j)} \|v\|^2, \quad \forall j.
\]

From the structure of the matrix $\Lambda$, it can be shown that $\det(\Lambda) = 0$. Therefore, one of the eigenvalues $\lambda$ is zero. Moreover, Lemma 5.2 states that $\frac{1}{\lambda_{max}(P_j)} < 1$, $\forall j$. As a consequence, with $\lambda = 0$, we have

\[
0 < \left(1 + \frac{1}{\lambda_{max}(P_j)}\right) \|v\|^2 < 0, \quad \forall j.
\]

which is a contradiction.

**Remark 5.1:** Proposition 5.1 provides a very efficient way to bound the convergence rate for $\|\bar{e}(k)\|$. According to the proposed methods, it is unnecessary to compute the matrix $\Lambda$ and to keep all matrices $\tilde{W}_j$, $j = 1, 2, \ldots, m$ since $\|\bar{e}(k)\|$ is bounded by the proposed Lyapunov function. Also, Theorem 5.1 guarantees the condition (18), which is assumed in Proposition 5.1.

Note that we specify the modal matrix $W_m$ in (5) as the most delayed case – all PEs use the oldest value in the buffer. Therefore, it can be inferred that $\lambda_{max}(W_m) \geq \lambda_{max}(P_j)$, $\forall j$, which results in

\[
\|\bar{e}(k)\|^2 < K \left(1 - \frac{1}{\lambda_{max}(P_m)}\right)^k \|e(0)\|^2, \quad (22)
\]

where $K$ is a positive constant. Therefore, the only information required to compute the convergence rate of $\|\bar{e}(k)\|$, is the matrix $W_m$ with the corresponding positive definite matrix $P_m$. As a result, the rate of convergence can be calculated by the proposed methods without any scalability problems.
VI. Error Analysis

In this section, we investigate the error probability, which quantifies the deviation of the random vector $X(k)$ from its steady-state value $X_s$ in probability. To measure this error probability, the Markov inequality given by $\Pr(X \geq \epsilon) \leq \mathbb{E}[X]/\epsilon$, where $X$ is a nonnegative random variable and $\epsilon$ is a positive constant, is used. First of all, we investigate the term $\text{vec}(e(k)e(k)^\top)$ as follows:

$$\text{vec}(e(k)e(k)^\top) = \text{vec}\left(W_{\sigma_{k-1}}e(k-1)e(k-1)^\top W_{\sigma_{k-1}}\right) = (W_{\sigma_{k-1}} \otimes W_{\sigma_{k-1}}) \text{vec}(e(k-1)e(k-1)^\top).$$

(23)

In the second equality of above equation, we used the property that $\text{vec}(ABC) = (C^\top \otimes A)\text{vec}(B)$.

By taking the expectation with new definitions $\bar{y}(k) \triangleq \text{vec}(e(k)e(k)^\top)$, $\tilde{y}(k) \triangleq \mathbb{E}[y(k)]$, and $\Gamma_{\sigma_k} \triangleq W_{\sigma_k} \otimes W_{\sigma_k}$, (23) becomes

$$\bar{y}(k) \triangleq \mathbb{E}[y(k)] = \mathbb{E}\left[\Gamma_{\sigma_{k-1}} y(k-1)\right] = \sum_{r=1}^{m} \mathbb{E}\left[\Gamma_{\sigma_{k-1}} y(k-1) \mid \sigma_{k-1} = r\right] \Pr(\sigma_{k-1} = r) = \sum_{r=1}^{m} \pi_r \Gamma_r \mathbb{E}[y(k-1)],$$

resulting in $\bar{y}(k) = (\sum_{r=1}^{m} \pi_r \Gamma_r) \tilde{y}(k-1)$, where in the second line we applied the law of total probability and the last equality holds by $\Pr(\sigma_{k-1} = r) = \pi_r$ for i.i.d. switching.

By the exactly same argument given in Lemma 5.1 and Proposition 5.1, the upper bound for $\bar{y}(k)$ is obtained as follows:

$$||\bar{y}(k)|| < K \left(1 - \frac{1}{\lambda_{\text{max}}(P_m)}\right)^{k/2} ||y(0)||, \ \forall k \in \mathbb{N}, \quad (24)$$

where $K$ is some positive constant and $\hat{P}_m$ is a positive definite matrix, satisfying the condition $\Gamma_m \hat{P}_m \Gamma_m - \hat{P}_m = -I$. However, the positive definite matrix $P_m \in \mathbb{R}^{N_{\mathfrak{q}} \times N_{\mathfrak{q}}}$ in (17), the dimension of the matrix $\hat{P}_m$ is given by $\hat{P}_m \in \mathbb{R}^{(N_{\mathfrak{q}})^2 \times (N_{\mathfrak{q}})^2}$, which may be large in size, and hence incurs computational intractabilities to obtain such a $\hat{P}_m$. Therefore, we introduce the following proposition and theorem in order to further facilitate the computation of $\lambda_{\text{max}}(P_m)$ as follows.

Proposition 6.1: Consider a positive definite matrix $\hat{P}_m$, satisfying the condition $\Gamma_m^\top \hat{P}_m \Gamma_m - \hat{P}_m = -I$, where $\Gamma_m \triangleq \tilde{W}_m \otimes \tilde{W}_m$, and $\tilde{W}_m$ is any real square matrix. If we assume that there exist finite, positive constants $k_0$, $c_0$, and $c_1$ such that

$$1 \leq ||\tilde{W}_m||^4 \leq c_0, \quad \text{for } k \in [0, k_0), \quad (25)$$

$$||\tilde{W}_m||^4 \leq c_1 < 1, \quad \text{for } k \in [k_0, \infty), \quad (26)$$

then, the largest eigenvalue of $\hat{P}_m$ is bounded by the following function:

$$\lambda_{\text{max}}(\hat{P}_m) < \sum_{k=0}^{\infty} ||\tilde{W}_m||^4 \leq k_0 c_0 \left(\frac{1}{1-c_1}\right), \quad (27)$$

Proof: The leftmost inequality in (27) can be proved as follows. The positive definite matrix $\hat{P}_m$ satisfying the condition $\Gamma_m^\top \hat{P}_m \Gamma_m - \hat{P}_m = -I$, is analytically computed by $\hat{P}_m = \sum_{k=0}^{\infty} (\Gamma_m^k \Gamma_m^\top) = \sum_{k=0}^{\infty} \Gamma_m^k \Gamma_m^k$.

Then, for a given matrix $\Gamma_m \triangleq \tilde{W}_m \otimes \tilde{W}_m$, we have

$$\Gamma_m^k \Gamma_m^k < \rho(\Gamma_m^k \Gamma_m^k)I = \rho(\Gamma_m^k \Gamma_m^\top)I = \sigma_{\text{max}}(\Gamma_m^k)I = ||\tilde{W}_m||^4 I = ||\tilde{W}_m||^4 I,$$

(28)

where $\rho(\cdot)$ and $\sigma_{\text{max}}(\cdot)$ denote the spectral radius and the spectral norm, respectively. For equality conditions in (28), we used the known property that $\sqrt{\rho(\Gamma_m^k \Gamma_m^k)} = \sigma_{\text{max}}(\Gamma_m^k)$ and $||\tilde{W}_m \otimes \tilde{W}_m|| = ||\tilde{W}_m||^2$, $\forall k \in \mathbb{N}_0$. By summing up from $k = 0$ to $\infty$, and then taking the largest eigenvalue in (28), we have $\lambda_{\text{max}}(\hat{P}_m) = \lambda_{\text{max}} \left(\sum_{k=0}^{\infty} \Gamma_m^k \Gamma_m^\top\right) < \sum_{k=0}^{\infty} ||\tilde{W}_m||^4$.

For the rightmost inequality in (27), the assumptions in (25)-(26) result in

$$\sum_{k=0}^{\infty} ||\tilde{W}_m||^4 = \sum_{k=0}^{k_0-1} ||\tilde{W}_m||^4 + \sum_{k=k_0}^{\infty} ||\tilde{W}_m||^4 \leq k_0 c_0 + \sum_{k=k_0}^{2k_0-1} ||\tilde{W}_m||^4 + \sum_{k=2k_0}^{3k_0-2} ||\tilde{W}_m||^4 + \cdots$$

$$= k_0 c_0 + \sum_{k=0}^{k_0-1} ||\tilde{W}_m||^4 + \sum_{k=2k_0}^{3k_0-2} ||\tilde{W}_m||^4 + \cdots$$

$$\leq k_0 c_0 + \sum_{k=0}^{k_0-1} ||\tilde{W}_m||^4 \left(\sum_{k=0}^{k_0-1} ||\tilde{W}_m||^4\right) + \sum_{k=2k_0}^{3k_0-2} ||\tilde{W}_m||^4 + \cdots$$

$$\leq k_0 c_0 + k_0 c_0 c_1 + \sum_{k=2k_0}^{3k_0-2} ||\tilde{W}_m||^4 + \cdots$$

$$\leq k_0 c_0 + k_0 c_0 c_1 + k_0 c_0 c_1^2 + \cdots$$

$$= k_0 c_0 \left(\sum_{i=0}^{\infty} c_1^i\right) = k_0 c_0 \left(1 - c_1\right).$$

Hence, we have $\lambda_{\text{max}}(\hat{P}_m) < k_0 c_0 \left(1 - c_1\right).$}

Theorem 6.1: Consider a stable, i.i.d. jump linear system with subsystem dynamics $\tilde{W}_j$ given in (13). Then, the probability of $||e(k)||^4 > \epsilon$, where $\epsilon$ is some positive constant, is given by

$$\Pr\left(||e(k)||^4 > \epsilon\right) \leq \min(1, \beta), \quad k \in \mathbb{N}_0,$$

(29)

where $\beta \triangleq \sqrt{\pi K} \left(1 - \frac{1-c_1}{k_0 c_0}\right)^{k_0} \frac{1}{\epsilon^2} ||y(0)||$, $K > 0$ is a constant, $c_0, c_1, k_0$ are positive constants such that the conditions (25)-(26) are satisfied.
(b) Convergence rate

where we used the cyclic permutation property for the trace operator in the first line and the equality in the second line holds by the property $\text{tr}(X^\top Y) = \text{vec}(X)^\top \text{vec}(Y)$ for any square matrix $X, Y \in \mathbb{R}^{n \times n}$.

We take the expectation in both sides of (30), which leads to

$$
\mathbb{E}[||e(k)||^2] = \text{vec}(I)^\top \mathbb{E}[\bar{y}(k)] = \text{vec}(I)^\top \bar{y}(k).
$$

(31)

Since the term $\mathbb{E}[||e(k)||^2]$ is a scalar value, taking the Euclidean norm returns the same value. Hence, applying the Euclidean norm in (31) results in

$$
\mathbb{E}[||e(k)||^2] = ||\text{vec}(I)^\top \bar{y}(k)||
\leq ||\text{vec}(I)^\top|| \cdot ||\bar{y}(k)|| = \sqrt{n} \cdot ||\bar{y}(k)||. \tag{32}
$$

Now, plugging (24) and (27) into (32) leads to

$$
\mathbb{E}[||e(k)||^2] < \sqrt{n} K \left(1 - \frac{1 - c_1}{k_0 c_0}\right)^{k/2} ||y(0)||.
$$

Finally, by applying the Markov inequality the above equation ends up with

$$
\Pr \left( ||e(k)||^2 > \epsilon \right) \leq \frac{\mathbb{E}[||e(k)||^2]}{\epsilon} < \beta,
$$

where $\beta \triangleq \frac{\sqrt{n} K}{\epsilon} \left(1 - \frac{1 - c_1}{k_0 c_0}\right)^{k/2} ||y(0)||$.

Since the probability cannot exceed one, we have

$$
\Pr \left( ||e(k)||^2 > \epsilon \right) \leq \min(1, \beta)
$$

Theorem 5.1 represents the error probability for a given bound $\epsilon$. Since $e(k)$ is a time-varying variable, the probability $\Pr \left( ||e(k)||^2 > \epsilon \right)$ also changes with respect to time. Starting from a given initial condition $y(0)$, this probability will converge to zero if

$$
1 - \frac{1 - c_1}{k_0 c_0} < 1.
$$

VII. SIMULATIONS

In order to test the proposed methods, simulation was carried out for the one-dimensional heat equation. We implemented the asynchronous parallel algorithm with CUDA C++ programming on nVIDIA TeslaTM C2050 GPU, which has 448 CUDA cores. The simulations were performed with the following parameters:

- Simulation Parameters:

  \[ \Delta x = 0.1, \Delta t = 0.01, \alpha = 0.5, r = \alpha \frac{\Delta t}{\Delta x^2} = 0.5 \]

  \[ I.C. : u_i = \cos^2 \left( \frac{3\pi i}{2(N - 1)} \right), i = 1, 2, \ldots, N \]

  \[ B.C. : u_1(k) = 1, u_N(k) = 0, \forall k \]

- Buffer length: $q = 3$

- Number of PEs: $N = 100$.

- Number of grid points in PE: $n = 1$

For a given initial temperature, the spatio-temporal evolution of the state is presented in Fig. 2 As time $k$ increases, the curved shape of the temperature, given as a cosine square function initially, flattens out. This simulation represents the synchronous case.

In Fig. 3 (a), the ensemble of the trajectories is shown for the asynchronous algorithm. The solid lines show the trajectories of total 300 simulations. Due to the randomness in the asynchronous algorithm, the trajectories differ from each other. For a reference, the synchronous scheme is also shown by a dashed line. Although it seems that the synchronous scheme converges faster with respect to the given iteration step, the physical simulation time may take more because the idle time is necessary at each iteration in the synchronous case. As the proposed method guarantees the stability through the common eigenvectors, both synchronous and asynchronous trajectories converged to the same steady-state value $X_{ss}$, depicted by starred line.

Next, we present the result for the convergence rate of the asynchronous algorithm. We assume that the switching probability $\Pi$ has the form of an i.i.d. jump process. Fig. 3
(b) shows the convergence rate of $\|e(k)\|$, which describes how fast the expected value of the state converges to $X_{ss}$. The solid lines are 300 sample trajectories of $\|e(k)\|$, starting from the given initial condition: $e(0) = X(0) - X_{ss}$. The dotted line depicts the time history of the normed empirical mean $\|\bar{e}(k)\|$, whereas the dashed line shows an upper bound by the proposed Lyapunov method (22). Note that $\|e(k)\|$ is a random variable, and hence the normed empirical mean $\|\bar{e}(k)\|$ was obtained by averaging the data over 300 simulations. In the proposed method, however, it is not necessary to execute the simulation multiple times.

Fig. 4 represents the result for the error probability with respect to time and $\epsilon$. For different values of $\epsilon$, Fig. 4(a) and (b) describe the time history of the error probabilities. The solid line denotes the empirical probability obtained from data — i.e., the number of samples satisfying $\|e(k)\|^2 > \epsilon$ divided by the total number of samples. The dashed line depicts the Markov inequality, computed from $\mathbb{E}[\|e(k)\|^2]$, where $\mathbb{E}[\|e(k)\|^2]$ is obtained by the statistics. Finally, the cross symbols mean the upper bound by the proposed method. As shown in Fig. 4(a) and (b), the probabilities for all cases converge to zero since the error is asymptotically convergent.

On the other hand, Fig. 4(c) and (d) show the error probability with respect to $\epsilon$ at fixed time instance. In this result, the time is fixed at $k = 9000$ out of total 10000 iterations times, and the probability is computed while increasing $\epsilon$ values. In Fig. 4(c) and (d), $\epsilon T$ is given by the index along $x$-axis, where the value of $T$ is given in Fig. 4(c) and (d), respectively. In both cases, the error probabilities decrease as $\epsilon$ increases.

Although the proposed methods provide a conservative bound, it does not require executing the code multiple times to predict the convergence rate or the error probability. In addition to that, the proposed methods are carried out in a computationally efficient manner without storing all subsystem matrices. In this example, we have $m = 3^2(100 - 2) \approx 3^{200}$, and keeping $3^{200}$ numbers of matrices is intractable in the real implementation. The proposed method, however, guarantees the convergence rate and the error probability, without any scalability issues. Therefore, the presented methods provide a computationally efficient tool to analyze the asynchronous numerical schemes.

VIII. CONCLUSIONS

This paper studied the stability, convergence rate, and error probability of the asynchronous parallel numerical algorithm. The asynchronous algorithm achieves better performance in terms of the total simulation time, particularly when massively parallel computing is required because it doesn’t wait for synchronization across PEs. In order to analyze the asynchronous numerical algorithm, we adopted the switched linear system framework. Although modeling of massively parallel numerical algorithms as switched dynamical systems results in a very large number of modes, we developed new methods that circumvent this scalability issue. While the results presented here are based on 1D heat equation, the analysis approach is generic and be applicable to other PDEs as well.

REFERENCES

[1] Z. Fan, F. Qiu, A. Kaufman, and S. Yoakum-Stover, “Gpu cluster for high performance computing,” in Proceedings of the 2004 ACM/IEEE conference on Supercomputing, p. 47, IEEE Computer Society, 2004.
[2] J. D. Owens, M. Houston, D. Luebke, S. Green, J. E. Stone, and J. C. Phillips, “Gpu computing,” Proceedings of the IEEE, vol. 96, no. 5, pp. 879–899, 2008.
[3] J. Nickolls, I. Buck, M. Garland, and K. Skadron, “Scalable parallel programming with cuda,” Queue, vol. 6, no. 2, pp. 40–53, 2008.
[4] D. A. Donzis and K. Aditya, “Asynchronous finite-difference schemes for partial differential equations,” Journal of Computational Physics, vol. 274, pp. 370–392, 2014.
[5] J. Daafouz, P. Riedinger, and C. Jung, “Stability analysis and control synthesis for switched systems: a switched lyapunov function approach,” Automatic Control, IEEE Transactions on, vol. 47, no. 11, pp. 1883–1887, 2002.
[6] H. Lin, G. Zhai, L. Fang, and P. J. Antsaklis, “Stability and h performance preserving scheduling policy for networked control systems,” in Proc. 16th IFAC World Congress on Automatic Control, 2005.
[7] K. Lee and R. Blattatcharya, “Optimal switching synthesis for jump linear systems with gaussian initial state uncertainty,” in ASME 2014 Dynamic Systems and Control Conference, pp. V002T2A003–V002T2A003, American Society of Mechanical Engineers, 2014.
[8] 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, vol. 2, pp. 1345–1351, IEEE, 1999.
[9] L. Xiao, A. Hassibi, and J. P. How, “Control with random communication delays via a discrete-time jump system approach,” in American Control Conference (ACC), 2000. Proceedings of the 2000, vol. 3, pp. 2199–2204, IEEE, 2000.
[10] L. Zhang, Y. Shi, T. Chen, and B. Huang, “A new method for stabilization of networked control systems with random delays,” Automatic Control, IEEE Transactions on, vol. 50, no. 8, pp. 1177–1181, 2005.
[11] M. Liu, D. W. Ho, and Y. Niu, “Stabilization of markovian jump linear system over networks with random communication delay,” Automatic, vol. 45, no. 2, pp. 416–421, 2009.
[12] K. Lee, A. Halder, and R. Bhattacharya, “Probabilistic robustness analysis for stochastic jump linear systems,” in American Control Conference (ACC), 2014. Proceedings of the 2014, pp. 2638–2643, IEEE, 2014.

[13] K. Lee, A. Halder, and R. Bhattacharya, “Performance and robustness analysis of stochastic jump linear systems using wasserstein metric,” Automatica, vol. 51, pp. 341–347, 2015.

[14] K. Lee and R. Bhattacharya, “Stability analysis of large-scale distributed networked control systems with random communication delays: A switched system approach,” arXiv preprint arXiv:1503.03047, 2015.

[15] R. H. Pletcher, J. C. Tannehill, and D. Anderson, Computational fluid mechanics and heat transfer. CRC Press, 2012.

[16] T. J. Laffey and H. Šmigoc, “Tensor conditions for the existence of a common solution to the lyapunov equation,” Linear algebra and its applications, vol. 420, no. 2, pp. 672–685, 2007.

[17] R. Shorten, K. S. Narendra, and O. Mason, “A result on common quadratic lyapunov functions,” Automatic Control, IEEE Transactions on, vol. 48, no. 1, pp. 110–113, 2003.

[18] D. Liberzon, Switching in systems and control. Springer, 2003.

[19] L. Gurvits, R. Shorten, and O. Mason, “On the stability of switched positive linear systems,” Automatic Control, IEEE Transactions on, vol. 52, no. 6, pp. 1099–1103, 2007.

[20] H. Lin and P. J. Antsaklis, “Stability and stabilizability of switched linear systems: a survey of recent results,” Automatic control, IEEE Transactions on, vol. 54, no. 2, pp. 308–322, 2009.