Stability of multiplexed NCS based on an epsilon-greedy algorithm for communication selection★

Harsh Oza[a,*], Irinel Constantin Morărescu[b,c], Vineeth S. Varma[b,c] and Ravi Banavar[a]

[a]Systems and Control Engineering, IIT Bombay, Powai, Mumbai, 400076, India
[b]Universite de Lorraine, CNRS, CRAN, Nancy, F-54000, France
[c]Technical University of Cluj-Napoca, Str. Memorandumului, Nr. 28,, Cluj-Napoca, 400114, Romania

ARTICLE INFO

Keywords:
Networked control system
Optimal control

ABSTRACT

In this article, we study a Networked Control System (NCS) with multiplexed communication and Bernoulli packet drops. Multiplexed communication refers to the constraint that transmission of a control signal and an observation signal cannot occur simultaneously due to the limited bandwidth. First, we propose an ϵ-greedy algorithm for the selection of the communication sequence that also ensures Mean Square Stability (MSS). We formulate the system as a Markovian Jump Linear System (MJLS) and provide the necessary conditions for MSS in terms of Linear Matrix Inequalities (LMIs) that need to be satisfied for three corner cases. We prove that the system is MSS for any convex combination of these three corner cases. We validate our approach with a numerical example that shows the effectiveness of our method.

1. Introduction

A Networked Control System (NCS) consists of a plant, a controller, and a communication network. These systems are integral to various industries, such as chemical processing, power grids, and warehouse automation [1, 5]. However, the NCS has significant challenges like bandwidth constraints, communication delays, random packet drops, and potential cyberattacks, all of which can degrade system performance [4, 6, 3, 14]. This study focuses on NCSs operating under bandwidth constraints, modeled as a multiplexing scenario where control and observation signals share limited communication resources. We also consider random packet drops, which add further uncertainty. Our primary aim is to develop a communication policy that ensures Mean Square Stability (MSS) while optimizing a performance metric defined by a quadratic cost function.

The uncertainty threshold principle proposed by Athans et al. laid foundational stability conditions for control systems under uncertainty. This forms the basis of subsequent work which gives stability conditions in different scenarios. In [7], linear systems controlled over a network are analyzed under packet drop uncertainties, modeled as Bernoulli random variables with independent drop probabilities in both communication channels. The necessary conditions for MSS and optimal control solutions using dynamic programming are then provided. Schenato et al. generalized the work by considering noisy measurements and giving stronger conditions for the existence of the solution [15]. The authors presented that the separation principle holds for the Transmission Control Protocol (TCP). The authors demonstrated that the separation principle holds for TCP. Other studies have focused on designing Kalman filters for wireless networks [9, 16] and deriving stability conditions for systems experiencing packet drops [19]. Alternative methods to address random packet drops include using redundant transmissions [11] and developing event-triggered policies for nonlinear systems under packet drop scenarios [18].

While the literature extensively addresses stability and optimality in NCSs, the combined challenges of stability and optimal scheduling under multiplexed communication constraints in control and observation remain underexplored. For example, [15] examines multiplexing and packet drops but does not address optimal network selection, while [10] proposes a joint strategy for selecting and controlling NCSs but limits the focus to sensor signals. Existing research on bandwidth constraints can be categorized as: i) multiplexing across multiple sensor signals, e.g., [16], and ii) multiplexing between sensor and control signals, e.g., [15]. Other communication uncertainties, such as packet drops and delays are also analyzed in [20]. The primary goals of these studies are to devise optimal control strategies or policies for selecting communication channels, as seen in [13].

Leong et al. address the boundedness of error covariance in a multiplexed sensor system with packet drops [8]. They establish stability conditions based on packet drop probabilities and develop an optimal communication sequence by training a Deep Q-Network (DQN) using an ϵ-greedy algorithm.

In this article, our contributions are as follows:

i) We propose a modified ϵ-greedy algorithm for the selection of the direction of communication (transmit or receive.)

ii) We establish the necessary conditions for the MSS of an NCS with multiplexed communication and packet drops.

★The authors thank the support of the Indo-French Centre for the Promotion of Advanced Research (IFCPRA).

*Corresponding author
harsh.oza@iitb.ac.in (H. Oza); constantin.morarescu@univ-lorraine.fr (I.C. Morărescu); vineeth.satheeskumar-varma@univ-lorraine.fr (V.S. Varma); banavar@iitb.ac.in (R. Banavar)

Notation

Let $\mathbb{R}$ and $\mathbb{N}$ denote set of real numbers and set of integers respectively. $\|\cdot\|$ denotes Euclidean 2-norm. For a discrete random variable $X \in \{x, y\}$, we denote the Bernoulli distribution as $\mathrm{Ber}(x, y, p)$ and is given by,

$$\begin{aligned} \mathbb{P}(X = x) &= p \\ \mathbb{P}(X = y) &= 1 - p \end{aligned} \tag{1}$$

with $p \in [0, 1]$. For an $n \in \mathbb{N}$, we denote the identity matrix as $\mathrm{I}_n \in \mathbb{R}^{n\times n}$.

2. Problem Setup

In this section, we first present a plant and controller model. Then we present the communication model with multiplexing constraints and packet loss. Lastly, we present the networked model of the plant.

2.1. Plant and Controller Model

Consider a discrete-time linear system

$$x_{k+1} = Ax_k + Bu_k \tag{2}$$

for all $k \in \mathbb{Z}_{\geq 0}$, where $x_k \in \mathbb{R}^{n_x}$ is the state, $u_k \in \mathbb{R}^{n_u}$ is the control input and $y_k \in \mathbb{R}^{n_y}$ is the output at k^{th} instant. We make the following assumptions regarding the original closed-loop system.

Assumption 1. There exists a state feedback controller of the form

$$u_k = Kx_k \tag{3}$$

that stabilizes the system (2).

2.2. Networked System Model

In this article, we are interested in an application where the plant and the controller are remotely located. The communication between the plant and the controller occurs over a wireless communication network. The networked system is illustrated in Fig. 1. The networked system dynamics can be written as

$$x_{k+1} = Ax_k + B\hat{u}_k \tag{4}$$

where $\hat{u}_k$ denotes the networked version of the control signal. We proceed by emulation of the controller (3) and use the controller as

$$u_k = K\hat{x}_k \tag{5}$$

where $\hat{x}_k$ denotes the estimates of the state at the controller end. A more detailed explanation of these quantities is presented later in this section.

Motivated by real-world communication constraints in the form of bandwidth limitations, we consider a communication constraint as described below. At any time instant, the network scheduler has three choices:

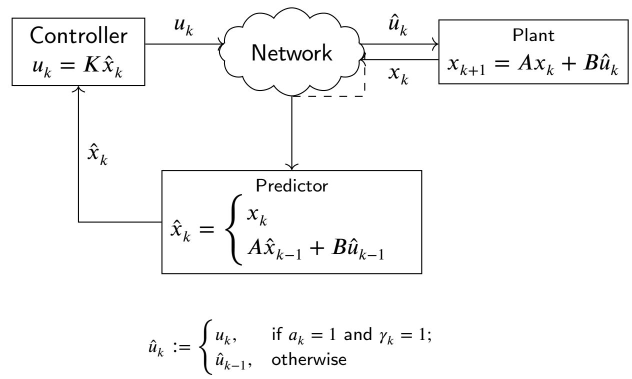

Figure 1: Schematic of a Networked Control System with information multiplexing and packet drops in the network.

i) transmit the control input from the controller to the plant or,

ii) transmit the measured signal from the plant to the predictor or,

iii) not communicate at all.

These three choices are encapsulated in the form of a switch variable a_k that takes values in a discrete set $\mathcal{A} := \{-1, 0, 1\}$ where,

$$a_k := \begin{cases} 1 & \text{if the control is transmitted;} \\ -1 & \text{if the observation is transmitted;} \\ 0 & \text{if there is no communication.} \end{cases} \tag{6}$$

We also consider lossy communication in the sense of a packet drop scenario. Consider γ_k to denote the packet drop event, modeled as an independent Bernoulli random variable with the probability distribution as $\mathrm{Ber}(1, 0, \delta)$. Here $\gamma_k = 1$ indicates a successful packet transmission and $\gamma_k = 0$ indicates failure.

Based on the switching and the packet drop assumptions, the transmitted control information through the network is written as

$$\hat{u}_k := \begin{cases} u_k & \text{if } a_k = 1 \text{ and } \gamma_k = 1; \\ \hat{u}_{k-1} & \text{otherwise} \end{cases} \tag{7}$$

for all $k \in \mathbb{N}$ with $\hat{u}_0 = u_0$.

Assumption 2. The predictor is collocated with the controller, enabling efficient and delay-free exchange of data between them.

Assumption 3. Communication is established over the Transmission Control Protocol (TCP), ensuring reliable data transfer via acknowledgment-based mechanisms.

Assumption 2 ensures that the predictor has access to the control input u_k, while Assumption 3 guarantees availability of the acknowledgment signal γ_k. Consequently, the predictor can reconstruct or retain the applied control input $\hat{u}_{k-1}$. The coarse estimate of the state at the controlled end is

$$\hat{x}_k = \begin{cases} x_k & \text{if } a_k = -1 \text{ and } \gamma_k = 1; \\ A\hat{x}_{k-1} + B\hat{u}_{k-1} & \text{otherwise} \end{cases} \tag{8}$$

for all $k \in \mathbb{N}$ with $\hat{x}_0 = x_0$. Define the concatenated state as $\chi_k := \left(x_k^\intercal, \hat{x}_{k-1}^\intercal, \hat{u}_{k-1}^\intercal\right)^\intercal \in \mathbb{R}^{2n_x+n_u}$. Therefore the overall model of the system is written as

$$\chi_{k+1} = \begin{cases} \mathsf{A}_1 \chi_k & \text{if } a_k = 1 \text{ and } \gamma_k = 1; \\ \mathsf{A}_{-1} \chi_k & \text{if } a_k = -1 \text{ and } \gamma_k = 1; \\ \mathsf{A}_0 \chi_k & \text{otherwise} \end{cases} \tag{9}$$

for all $k \in \mathbb{N}$, where

$$\mathsf{A}_1 = \begin{pmatrix} A & BKA & BKB \\ 0 & A & B \\ 0 & KA & KB \end{pmatrix}, \mathsf{A}_{-1} = \begin{pmatrix} A & 0 & B \\ \mathrm{I}_{n_x} & 0 & 0 \\ 0 & 0 & \mathrm{I}_{n_u} \end{pmatrix},$$

$$\text{and } \mathsf{A}_0 = \begin{pmatrix} A & 0 & B \\ 0 & A & B \\ 0 & 0 & \mathrm{I}_{n_u} \end{pmatrix}.$$

3. Switching Strategy

To develop an optimal switching strategy within a multi-stage system, we define a performance measure that captures the cost associated with each state-action pair, balancing penalties on state deviations, control efforts, and communication constraints. The per-stage cost is defined as

$$J\left(\chi_k, a_k\right) := \chi_k^\intercal S \chi_k + \lambda a_k^2 \tag{10}$$

where $S \in \mathbb{R}^{(2n_x+n_u)^2}$ is a positive definite matrix, ensuring penalization of state deviations and control inputs, and $\lambda \in R_{\geq 0}$ is a factor that weighs penalty on communication. This structure reflects the trade-offs between control performance and communication overhead, which is critical for networked control systems operating under limited bandwidth and energy constraints. We use the value function, denoted as $\mathcal{V}(\chi)$, representing a discounted future cost sum. The value function is defined as:

$$\mathcal{V}(\chi_k) := \min_{a_k \in \mathcal{A}} \mathbb{E}\left[J\left(\chi_k, a_k\right) + \beta \mathcal{V}(\chi_{k+1})\right] \tag{11}$$

with the discount factor $\beta \in]0, 1[$. The optimal action that minimizes the value function is given by

$$a_k^\star := \arg\min_{a_k \in \mathcal{A}} \mathbb{E}\left[J\left(\chi_k, a_k\right) + \beta \mathcal{V}(\chi_{k+1})\right]. \tag{12}$$

Bellman's principle of optimality, given in (11), holds for our problem. In discrete state–action spaces, the optimal policy can be obtained via value iteration [17]. For continuous spaces, where direct computation is infeasible, we approximate the Bellman equation in (11)–(12) using Deep Reinforcement Learning (DRL) [12], representing $\mathcal{V}(\chi)$ with a neural network $\mathcal{V}_\theta(\chi, a)$ over a finite action set and network weight (θ), that is trained via minimizing a loss with a target network $\mathcal{V}_{\theta^-}$.

3.1. ε-Greedy Algorithm for Switching

We employ the ε-greedy switching strategy for the decision-making process [17]. Classically, the algorithm consists of two parts: exploration and exploitation. *Exploration* addresses finding new possible solutions, whereas *exploitation* addresses utilizing the already known optimal solution. The variable ε acts as a parameter that weighs the two. In contrast to the classical framework with a large (perhaps infinite) action set, in this work, ε is used to ensure that communication occurs sufficiently frequently (only two actions) so that the mean square stability of the system is preserved even if the greedy part decides to never communicate. The switching strategy is defined mathematically as follows:

$$a_k = \begin{cases} \mathrm{Ber}(1, -1, \frac{1}{2}) & \text{with probablity } \varepsilon \\ a_k^\star \text{ as in (12)} & \text{with probability } 1 - \varepsilon \end{cases} \tag{13}$$

With probability ε the switching variable a_k is chosen uniformly randomly from $\{-1, 1\}$. This random selection represents exploration by allowing the system to consider other strategies that might not be optimal for that instance but could provide a better solution over a longer horizon. With probability $1 - \varepsilon$, the switch variable a_k is determined by the exploitation part, i.e., minimizing the value function as in (12).

3.2. Switching Probabilities with ε-Greedy Algorithm

In this subsection, we discuss in detail the generalized switching probability under the ε-greedy algorithm and analyze the effect of the ε-greedy switching strategy on the MSS of the closed loop system. The switching probability distribution function $\mathcal{P}_g \in [0, 1]^3$ is given by

$$\mathcal{P}_g := \left(\mathbb{P}(a_k = 1), \mathbb{P}(a_k = 0), \mathbb{P}(a_k = -1)\right) \tag{14}$$

with some switching algorithm g. Given the switching strategy defined in (13), the probabilities of switching states are influenced by the choice of ε. The choice of ε, in turn, decides the balance between exploration and exploitation. With the switching strategy (13), the switching probability distribution (denoted as $\mathcal{P}_\varepsilon$) can be written as

$$\begin{aligned} \mathcal{P}_\varepsilon &= \varepsilon\left(\frac{1}{2}, 0, \frac{1}{2}\right) + (1-\varepsilon)\left(p_k, 1 - p_k - q_k, q_k\right) \\ &= \left(\frac{\varepsilon}{2} + (1-\varepsilon)p_k, (1-\varepsilon)(1 - p_k - q_k), \frac{\varepsilon}{2} + (1-\varepsilon)q_k\right). \end{aligned} \tag{15}$$

Here, the first term represents the probability distribution in the exploration phase, and the second term, $(1 - \varepsilon)\left(p_k, 1 - p_k - q_k, q_k\right)$, represents the probability distribution in the exploitation phase with some $p_k, q_k \in [0, 1]$ and $p_k + q_k \leq 1$. Without loss of generality, we omit the time argument associated with p_k and q_k.

Remark 1. The choice of ε impacts the MSS of the system which is one of the main interests of this article.

3.3. Formulation as a Markovian Jump Linear System

In this section, we elaborate on the framework of the Markovian Jump Linear System (MJLS) that relates to the system described in (9). An MJLS is a linear system that goes through random transitions between a finite number of modes, each governed by linear dynamics. The random transitions are governed by Markovian probabilities associated with switching from one mode to another mode. In the problem, the randomness takes place at two levels: the first is at the switching under ε-greedy policy and the second is the random packet drop. With this backdrop, we introduce modes of operation under varying circumstances.

System Modes: The system can operate in one of several modes at any instance, which is determined by the switch position and the status of packet transmission. The mode of operation is denoted by $\mathcal{M}_i$ for all $i \in \{1, 2, \ldots, 5\}$. These modes represent different scenarios of switching and packet drop. We define the modes of the Markovian switching system as follows:

$$\begin{aligned}&\mathcal{M}_1 : \left(1, \mathcal{S}, \mathsf{A}_1\right), \mathcal{M}_2 : \left(1, \mathcal{F}, \mathsf{A}_0\right),\\ &\mathcal{M}_3 : \left(-1, \mathcal{S}, \mathsf{A}_{-1}\right), \mathcal{M}_4 : \left(-1, \mathcal{F}, \mathsf{A}_0\right),\\ &\text{and } \mathcal{M}_5 : \left(0, -, \mathsf{A}_0\right)\end{aligned} \tag{16}$$

where the first entry in each 3−tuple indicates the switch position $1, -1$, or 0; the second entry denotes whether the packet transmission was successful ($\mathcal{S}$) or failed ($\mathcal{F}$), and the third entry corresponds to the system matrix A that governs the dynamics in that particular mode. The mode transition probability, p_{ij}, represents the likelihood of the system switching from mode i to mode j in the next time step.

Definition 1 (Mode Transition Probability). Define the mode transition probability of switching from i to j as,

$$\mathbb{P}(a_{k+1} = j | a_k = i) := p_{ij}$$

with $i, j \in \{1, 2, \ldots, M\}$ where $p_{ij} \in [0, 1]$, $\sum_{j=1}^{M} p_{ij} = 1$ and M denotes the total number of modes.

Definition 2 (Mean Square Stability, [2]). The system (9) is mean square stable if and only if for some $\zeta \geq 1, 0 < \xi < 1$ and for every $\chi_0 \in \mathbb{R}^{2n_x + n_u}$,

$$\mathbb{E}\left[\chi_k^\intercal \chi_k\right] \leq \zeta \xi^k \chi_0^\intercal \chi_0 \quad \text{for all } k \in \mathbb{Z}_{\geq 0}. \tag{17}$$

Objective: Our goal is to determine the value of ε, given a $\delta \in]0, 1]$, that ensures the origin of the system (9) remains Mean Square Stable under the switching algorithm (13).

4. Mean Square Stability of an MJLS

In this section, we propose the methodology used to address the problem. System (9) is seen as a convex combination between three corner cases. The coefficients used in the convex combination represent the switching probabilities between the corner cases. With the given packet transmission success probability $\delta \in]0, 1]$, we provide the necessary conditions for the MSS of the system using the Linear Matrix Inequalities (LMIs) [2].

Case	Switch probabilities
General	$\left(\frac{\varepsilon}{2} + (1-\varepsilon)p, (1-\varepsilon)(1-p-q), \frac{\varepsilon}{2} + (1-\varepsilon)q\right)$
C1	$\left(\frac{\varepsilon}{2}, 0, 1 - \frac{\varepsilon}{2}\right)$
C2	$\left(1 - \frac{\varepsilon}{2}, 0, \frac{\varepsilon}{2}\right)$
C3	$\left(\frac{\varepsilon}{2}, 1-\varepsilon, \frac{\varepsilon}{2}\right)$

Table 1
Switching probability distribution in different corner cases.

4.1. Corner Cases

We study this through three specific corner cases:

C1. $p = 0$ and $q = 1$: This case represents the switch being in position -1 throughout the entire exploitation phase, indicating that only observations are transmitted.

C2. $p = 1$ and $q = 0$: This case represents the switch being in position 1 throughout the entire exploitation phase, indicating that only control signals are transmitted.

C3. $p = 0$ and $q = 0$: This case depicts the switch remaining in position 0 for the duration of the exploitation phase, signifying that no transmissions occur.

The probability of switching in each of the corner cases is tabulated in TABLE 1.

4.2. Mode Transition Probabilities for Corner Cases

Let $\mathsf{P}^{(c)}(\varepsilon) \in [0, 1]^{5\times 5}$ denote the mode transition probability matrix associated with the corner case c, where $c \in \{1, 2, 3\}$. Here, the superscript c is the variable representing each corner case. For all $i, j \in \{1, 2, \ldots, 5\}$ we define $\mathsf{P}^{(c)}_{ij}(\varepsilon) = \mathsf{P}^{(c)}_j(\varepsilon)$ i.e., one has the same probability of transitioning from each mode to the mode j in the corner case c and this probability is denoted as $\mathsf{P}^{(c)}_j(\varepsilon) \in [0, 1]$. For mode j, the associated system matrix is denoted by $\mathtt{A}_1 = \mathsf{A}_1$, $\mathtt{A}_2 = \mathtt{A}_4 = \mathtt{A}_5 = \mathsf{A}_0$, and $\mathtt{A}_3 = \mathsf{A}_{-1}$. The mode transition probabilities, for each case, are described in Table 2. For a positive definite matrix V let us introduce the LMIs

$$\sum_{j=1}^{5} \mathsf{P}^{(c)}_j(\varepsilon) \mathtt{A}_j^\intercal V \mathtt{A}_j < V, \quad \forall c \in \{1, 2, 3\} \tag{18}$$

and the set

$$E := \left\{\varepsilon \in]0, 1] \middle| \exists V = V^\intercal > 0 \text{ s.t. (18) holds}\right\}.$$

With these notations, we are ready to give the main theoretical result of this work.

	$\mathsf{P}^{(c)}_1(\varepsilon)$	$\mathsf{P}^{(c)}_2(\varepsilon)$	$\mathsf{P}^{(c)}_3(\varepsilon)$	$\mathsf{P}^{(c)}_4(\varepsilon)$	$\mathsf{P}^{(c)}_5(\varepsilon)$
C1	$\delta\frac{\varepsilon}{2}$	$(1-\delta)\frac{\varepsilon}{2}$	$\delta(1-\frac{\varepsilon}{2})$	$(1-\delta)(1-\frac{\varepsilon}{2})$	0
C2	$\delta(1-\frac{\varepsilon}{2})$	$(1-\delta)(1-\frac{\varepsilon}{2})$	$\delta\frac{\varepsilon}{2}$	$(1-\delta)\frac{\varepsilon}{2}$	0
C3	$\delta\frac{\varepsilon}{2}$	$(1-\delta)(1-\frac{\varepsilon}{2})$	$\delta\frac{\varepsilon}{2}$	$(1-\delta)(1-\frac{\varepsilon}{2})$	$1-\varepsilon$
General	$\delta\left(\frac{\varepsilon}{2}+(1-\varepsilon)p\right)$	$(1-\delta)\left(\frac{\varepsilon}{2}+(1-\varepsilon)p\right)$	$\delta\left(\frac{\varepsilon}{2}+(1-\varepsilon)q\right)$	$(1-\delta)\left(\frac{\varepsilon}{2}+(1-\varepsilon)q\right)$	$(1-\varepsilon)(1-p-q)$

Table 2
Mode transition probabilities for different cases.

Theorem 1. *Given a scalar $\delta \in]0,1]$ and the mode transition probability matrix $\mathsf{P}^{(c)}(\varepsilon)$ as given in Table 2, if the set E is nonempty, then the origin of the system (9) is MSS for any $\varepsilon \in E$ under the ε-greedy algorithm (13).*

PROOF. The MSS conditions given in [2, Corollary 3.26] are tailored to the specific problem to obtain the LMI (18). Suppose there exists a positive definite matrix V and a $\varepsilon \in E$, then the origin of the system is stable for the corner cases C1, C2, and C3. To prove the origin of the system (9) is MSS under (13) with $\varepsilon \in E$ for any general case, we prove that the general case can be written as a convex combination of the three corner cases and then prove the LMI holds for the general case. Let $\alpha_1, \alpha_2 \in [0,1]$ and $\alpha_1+\alpha_2 \le 1$. Taking the convex combination of mode transition probabilities for all corner cases (see TABLE 2),

$$\begin{aligned}
&\alpha_1\left(\delta\frac{\varepsilon}{2},(1-\delta)\frac{\varepsilon}{2},\delta(1-\frac{\varepsilon}{2}),(1-\delta)(1-\frac{\varepsilon}{2}),0\right)\\
&+\alpha_2\left(\delta(1-\frac{\varepsilon}{2}),(1-\delta)(1-\frac{\varepsilon}{2}),\delta\frac{\varepsilon}{2},(1-\delta)\frac{\varepsilon}{2},0\right)\\
&+\left(1-\alpha_1-\alpha_2\right)\left(\delta\frac{\varepsilon}{2},(1-\delta)(1-\frac{\varepsilon}{2}),\delta\frac{\varepsilon}{2},\right.\\
&\left.(1-\delta)(1-\frac{\varepsilon}{2}),1-\varepsilon\right) \qquad (19)\\
&=\left(\delta\left(\frac{\varepsilon}{2}+(1-\varepsilon)\alpha_2\right),(1-\delta)\left(\frac{\varepsilon}{2}+(1-\varepsilon)\alpha_2\right),\right.\\
&\delta\left(\frac{\varepsilon}{2}+(1-\varepsilon)\alpha_1\right),(1-\delta)\left(\frac{\varepsilon}{2}+(1-\varepsilon)\alpha_1\right),\\
&\left.(1-\varepsilon)(1-\alpha_2-\alpha_1)\right).
\end{aligned}$$

Comparing (19) with the mode transition probability of the general case in TABLE 2, we have $\alpha_1 = q$ and $\alpha_2 = p$.
To prove that (9) is MSS for any $p, q \in [0,1]$ and $p+q \le 1$, let $\mathsf{P}^{(g)}$ be the general mode transition probability matrix. Then

$$\mathsf{P}^{(g)}_j(\varepsilon) = q\mathsf{P}^{(1)}_j(\varepsilon) + p\mathsf{P}^{(2)}_j(\varepsilon) + (1-q-p)\mathsf{P}^{(3)}_j(\varepsilon)$$

for all $j \in \{1,2,\dots,5\}$.

$$\begin{aligned}
&\sum_{j=1}^{5}\mathsf{P}^{(g)}_j(\varepsilon)\mathtt{A}_j^{\mathsf{T}}V\mathtt{A}_j\\
&=\sum_{j=1}^{5}\left(q\mathsf{P}^{(1)}_j(\varepsilon)+p\mathsf{P}^{(2)}_j(\varepsilon)+(1-q-p)\mathsf{P}^{(3)}_j(\varepsilon)\right)\mathtt{A}_j^{\mathsf{T}}V\mathtt{A}_j\\
&=q\underbrace{\sum_{j=1}^{5}\mathsf{P}^{(1)}_j(\varepsilon)\mathtt{A}_j^{\mathsf{T}}V\mathtt{A}_j}_{<V}+p\underbrace{\sum_{j=1}^{5}\mathsf{P}^{(2)}_j(\varepsilon)\mathtt{A}_j^{\mathsf{T}}V\mathtt{A}_j}_{<V}\\
&+(1-q-p)\underbrace{\sum_{j=1}^{5}\mathsf{P}^{(3)}_j(\varepsilon)\mathtt{A}_j^{\mathsf{T}}V\mathtt{A}_j}_{<V}\\
&\overset{(a)}{<} qV+pV+(1-q-p)V=V
\end{aligned}$$

The inequality (a) follows from (18) and the fact that all quantities on both sides of (18) are non-negative. Hence, if $\varepsilon \in E$ then, (18) is satisfied for any general switching strategy (13). □

To determine the value of ε that satisfies the LMIs required for MSS, we utilize a method involving Semi-Definite Programming (SDP) solvers and the bisection method. Based on Theorem 1, we set up the necessary LMIs involving symmetric positive definite matrix V. These LMIs establish the conditions for MSS that the system must satisfy. We employ an SDP solver to numerically solve the formulated LMIs. Given the dependence of the V matrix on ε, we apply the bisection method to determine the value of $\varepsilon \in E$. Starting with an initial range for $\varepsilon \in [0,1]$, the bisection method iteratively narrows down to $\varepsilon \in E$ by checking the existence of solutions of LMIs at midpoints within the range.

5. Numerical Experiments

In this section, we validate our approach on system (9) with the following data:

$$A=\begin{pmatrix}1 & 0.1\\ 0 & 1\end{pmatrix},\ B=\begin{pmatrix}0\\1\end{pmatrix},\ \text{and } K=\begin{pmatrix}-0.012\\-0.07\end{pmatrix}^{\mathsf{T}}$$

where the controller gain K stabilizes as per Assumption 2. First, we fix the packet transmission success probability $\delta \in$

]0, 1], typically determined by the communication system. It represents the probability of successful packet transmission and is influenced by the type of communication medium and its properties.

For a fixed δ, we determine the corresponding value of ε that ensures mean square stability (MSS) for all corner cases. The relationship between the packet transmission success probability δ and the $\bar{\varepsilon}$ required for ensuring MSS is depicted in Figure 2.

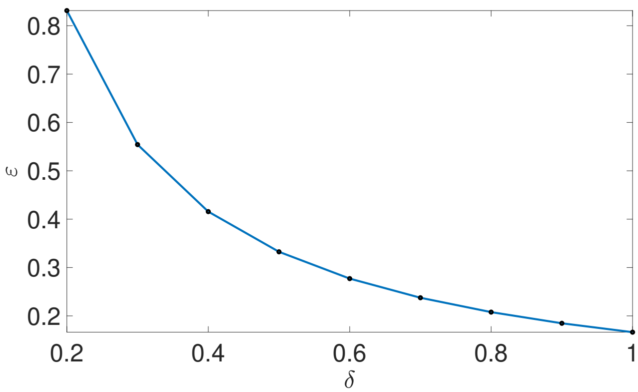

Figure 2: For the given δ, ε satisfying all three corner cases.

Following that we examine the Mean-Square Stability (MSS) property of the system when using the proposed ε-greedy algorithm. The analysis focuses on varying packet transmission success probabilities and their corresponding values of ε, as determined by Theorem 1. Figure 3 illustrates a scenario where the packet transmission success probability is low. Under such conditions, achieving MSS requires a relatively high value of ε. A higher value of the ε ensures that the communication exchange happens sufficiently. A higher value of ε based on (13) suggests that the switch would be either in the 1 or -1 position with a higher probability. Intuitively, sufficient communication ensures that the updated information from the plant reaches the controller and vice versa, thereby ensuring MSS. In contrast, Figure 4 shows that with the same low transmission success probability but a lower value of ε, the system fails to achieve MSS. This highlights the sensitivity of the system's stability to the choice of ε under challenging transmission conditions. On the other hand, Figure 5 demonstrates that when the packet transmission success probability is high, a low value of ε is sufficient to ensure MSS.

Observe that $\frac{\varepsilon\delta}{2}$ forms a bound on the probability of transmission in any one direction. We choose switch position 1 or -1 with equal probability $\frac{\varepsilon}{2}$ and the success probability is δ that makes an overall bound $\frac{\varepsilon\delta}{2}$. This can also be corroborated through Table 2 in all corner cases.

Table 3 presents the mean total cost along with the standard deviation for different combinations of the transmission success rate δ and the ε-greedy parameter ε, with the weighting factor λ fixed at 0.5. Each entry represents the average over multiple independent simulation runs, providing a measure of both the expected performance and its variability due to the randomness of packet drops and scheduling. From the table, it is observed that for a fixed ε, the total average of cost decreases as δ increases, reflecting improved closed-loop performance with fewer packet drops. Conversely, for a fixed δ, increasing ε tends to increase the total cost because more frequent transmissions raise the communication penalty. The standard deviation values indicate that variability is larger at lower δ and mid-range ε, highlighting the effect of stochasticity in the scheduling and transmission process.

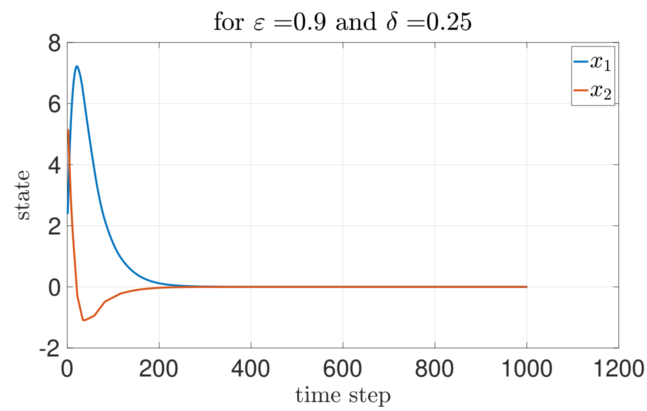

Figure 3: Low success probability, with high ε.

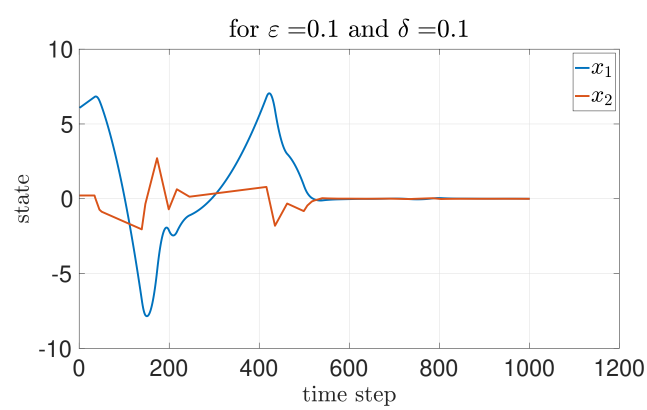

Figure 4: Low success probability, with low ε.

6. Conclusion

In this article, we have proposed a modified ε-greedy algorithm for selecting communication direction in a multiplexed NCS. We have established the necessary conditions for the mean square stability (MSS) of an NCS with multiplexed communication and packet drops. We displayed the stability for various combinations of packet success probability and necessary ε.

References

[1] Bemporad, A., Heemels, M., Johansson, M., et al., 2010. Networked control systems. volume 406. Springer.

$\delta \backslash \varepsilon$	**0.2**	**0.3**	**0.5**	**0.7**	**0.9**
0.2	945.18 ± 13.1	917.42 ± 11.2	853.77 ± 9.9	785.66 ± 9.1	839.52 ± 11.4
0.3	774.92 ± 9.6	722.83 ± 8.8	673.41 ± 8.1	628.27 ± 7.4	662.18 ± 8.3
0.5	553.74 ± 7.9	509.93 ± 7.2	470.86 ± 6.5	436.12 ± 6.0	512.64 ± 6.8
0.7	333.48 ± 5.4	305.17 ± 4.9	279.66 ± 4.4	256.93 ± 4.0	347.82 ± 5.1
0.9	112.37 ± 2.1	103.14 ± 1.9	101.26 ± 1.7	94.6 ± 1.6	86.93 ± 1.8

Table 3
Mean total cost ± standard deviation for different values of δ and ε with $\lambda = 0.5$, computed over multiple simulation runs.

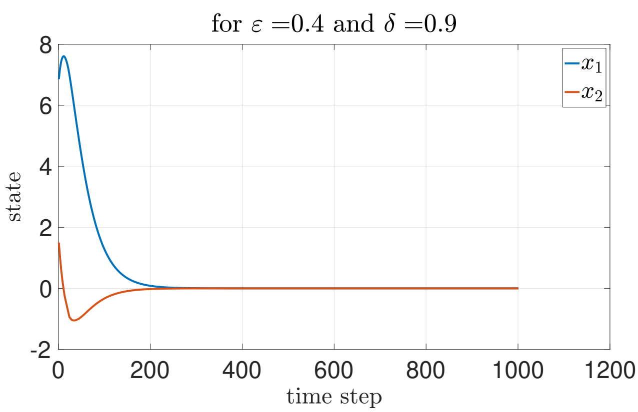

Figure 5: High success probability, with low ε.

[2] Costa, O.L.V., Fragoso, M.D., Marques, R.P., 2005. Discrete-time Markov jump linear systems. Springer Science & Business Media.

[3] Dolk, V., Heemels, M., 2017. Event-triggered control systems under packet losses. Automatica 80, 143–155.

[4] Elia, N., 2004. When bode meets shannon: Control-oriented feedback communication schemes. IEEE transactions on Automatic Control 49, 1477–1488.

[5] Ge, X., Yang, F., Han, Q.L., 2017. Distributed networked control systems: A brief overview. Information Sciences 380, 117–131.

[6] Heemels, W.M.H., Teel, A.R., Van de Wouw, N., Nešić, D., 2010. Networked control systems with communication constraints: Tradeoffs between transmission intervals, delays and performance. IEEE Transactions on Automatic control 55, 1781–1796.

[7] Imer, O.C., Yuksel, S., Başar, T., 2004. Optimal control of dynamical systems over unreliable communication links. IFAC Proceedings Volumes 37, 991–996.

[8] Leong, A.S., Ramaswamy, A., Quevedo, D.E., Karl, H., Shi, L., 2020. Deep reinforcement learning for wireless sensor scheduling in cyber–physical systems. Automatica 113, 108759.

[9] Liu, X., Goldsmith, A., 2004. Kalman filtering with partial observation losses, in: 2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601), IEEE. pp. 4180–4186.

[10] Maity, D., Mamduhi, M.H., Hirche, S., Johansson, K.H., 2021. Optimal lqg control of networked systems under traffic-correlated delay and dropout. IEEE Control Systems Letters 6, 1280–1285.

[11] Mesquita, A.R., Hespanha, J.P., Nair, G.N., 2012. Redundant data transmission in control/estimation over lossy networks. Automatica 48, 1612–1620.

[12] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., et al., 2015. Human-level control through deep reinforcement learning. nature 518, 529–533.

[13] Molin, A., Hirche, S., 2009. On lqg joint optimal scheduling and control under communication constraints, in: Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference, IEEE. pp. 5832–5838.

[14] Sandberg, H., Gupta, V., Johansson, K.H., 2022. Secure networked control systems. Annual Review of Control, Robotics, and Autonomous Systems 5, 445–464.

[15] Schenato, L., Sinopoli, B., Franceschetti, M., Poolla, K., Sastry, S.S., 2007. Foundations of control and estimation over lossy networks. Proceedings of the IEEE 95, 163–187.

[16] Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M.I., Sastry, S.S., 2004. Kalman filtering with intermittent observations. IEEE transactions on Automatic Control 49, 1453–1464.

[17] Sutton, R.S., Barto, A.G., 2018. Reinforcement learning: An introduction. MIT press.

[18] Varma, V.S., Postoyan, R., Quevedo, D.E., Morărescu, I.C., 2023. Event-triggered transmission policies for nonlinear control systems over erasure channels. IEEE Control Systems Letters 7, 2113–2118. doi:10.1109/LCSYS.2023.3285433.

[19] You, K., Fu, M., Xie, L., 2011. Mean square stability for kalman filtering with markovian packet losses. Automatica 47, 2647–2657.

[20] Zhang, L., Gao, H., Kaynak, O., 2012. Network-induced constraints in networked control systems—a survey. IEEE transactions on industrial informatics 9, 403–416.