Sequence-Based Stochastic Receding Horizon Control Using IMM Filtering and Value Function Approximation

Florian Rosenthal and Uwe D. Hanebeck

Abstract— **In this work, we address sequence-based stochastic receding horizon control over networks. We focus on the case where application layer acknowledgments are issued from the plant side upon reception of control inputs. Being an ordinary payload from the perspective of the underlying network, they can convey additional information that can be exploited by the controller upon reception. While this covers the notion of TCP-like and UDP-like communication usually considered in literature, the downside is that delays and losses of acknowledgments must be considered by the controller. It is known that in such cases the computation of optimal control policies is generally intractable due to the impact of the dual effect. For a usual quadratic cost criterion, we derive a tractable policy by approximating the non-convex value function by a set of coupled quadratic functions. The policy is linear in the state estimate provided by the IMM filter we proposed in a previous work and, hence, has only low computational complexity. Its performance is demonstrated in a numerical example.**

I. Introduction

At least since the dawn of cyber-physical systems and their tight integration of physical and software components, Networked Control Systems (NCS) that use general-purpose communication infrastructure to exchange the data needed to control and monitor the physical components, have become increasingly important in many fields [1], [2]. Along with reduced effort and maintenance, this is mainly due to an increase in flexibility compared to traditional point-to-point connections between the components of the control loops [3].

On the downside, employing such networks almost inevitably introduces additional aspects that affect the achievable control performance [4]. Packet delays and losses presumably pose the most critical problem since they can lead to missing inputs at the plant side. One popular countermeasure in literature is *sequence-based control* [5]–[7]. Here the idea is to transmit a complete sequence $\underline{U}_k$ of control inputs that also consists of predictive inputs for the next time steps in addition to the current one $\underline{u}_k$. Approaches for the computation of such control sequences usually rely on receding horizon principles [6], [8], build upon nominal controllers that disregard the network [9], or attempt to directly minimize a cost function [10], [11].

Yet, finding control policies that are truly optimal w.r.t. the underlying cost function is only possible in a few special cases and demands a TCP-like setting, that is, the instantaneous

The authors are with the Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Germany. Email: florian.rosenthal@kit.edu, uwe.hanebeck@ieee.org

*This work is supported by the German Science Foundation (DFG) within the Priority Programme 1914 "Cyber-Physical Networking".

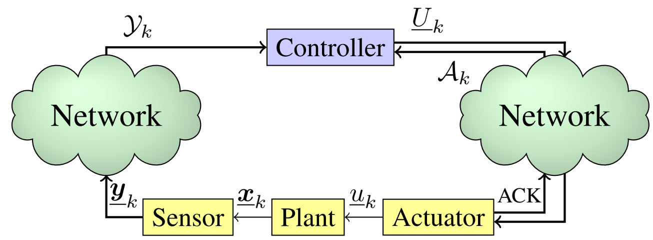

Fig. 1: Schematical overview of the considered setup.

and failure-free delivery of acknowledgments issued at the plant side upon reception of control inputs [11], [12]. In case acknowledgments are not provided, typically referred to as UDP-like communication, or can get lost, it has been shown that the controller's insufficient knowledge on actually applied control inputs impacts the future uncertainty of the state estimate [12], [13]. This is called the *dual effect* [14] and renders the computation of optimal policies intractable. Consequently, suitable approximations must be made in such setups. For instance, linear control policies for UDP-like scenarios have been presented in [10], [15], and, more recently, in [16].

In previous works [17], [18], we interpreted the acknowledgments issued at the plant side upon reception of applicable control sequences as *application layer acknowledgments*.[1] This approach does not only comprise UDP-like and TCP-like communication as special cases, but also enables the acknowledgments to carry more information for the controller. On the downside, as sketched in Fig. 1, potential delays and losses of acknowledgments must now also be factored in. In [17], we considered state estimation in such NCS based on interacting multiple model (IMM) filtering [20], while we addressed receding horizon control in [18]. In this work, the main result was an algorithm for the computation of the parameters of a linear controller that takes the impact of the dual effect into account and iteratively refines a reference trajectory to minimize an upper bound of the cost.

The main contribution of this paper is the introduction of an alternative to the algorithm from [18], which is computationally less complex and directly makes use of the state estimate provided by the IMM filter from [17], which is given in terms of a Gaussian mixture. The key idea of the proposed approach is to use approximate dynamic programming [21] to approximate the value function based

[1]They should not be confused with the dedicated acknowledgment packets that are issued by certain transport layer protocols in real networks. A common example is TCP, which uses them to retransmit data in order to enhance the reliability of the communication. On the downside, this trades losses for large delays, which is usually undesired in control applications [19].

on the interaction of a set of quadratic functions, one for each component of the mixture, so that the resulting control policy is linear in each of the means of the mixture.

Notation: Throughout this paper, vectors will be indicated by underlined letters ($\underline{x}$), random vectors will be underlined and in bold ($\underline{\boldsymbol{x}}$), and boldface capital letters indicate matrices, e.g., $\mathbf{A}$. We use $\mathbf{I}_n$ to denote the n-dimensional identity matrix, $\mathbf{0}$ to denote zero matrices of arbitrary dimension, and a subscript k, e.g., $\underline{x}_k$, to indicate the time step. Transposition of a vector or a matrix is indicated by $\underline{x}^{\mathrm{T}}$ and $\mathbf{A}^{\mathrm{T}}$, and $\mathbf{A} \geq 0$ ($\mathbf{A} > 0$) means that the matrix $\mathbf{A}$ is positive semidefinite (positive definite). Furthermore, $\mathbf{A}^{\dagger}$ denotes the pseudoinverse of $\mathbf{A}$. Finally, $\mathbb{1}_{i=j}$ is the indicator function, i.e., $\mathbb{1}_{i=j} = 1$ if $i = j$ and 0 otherwise, and the shortcut $a_{0:n}$ indicates a sequence $a_0, a_1, \ldots, a_n$.

II. Problem Formulation

Consider the NCS sketched in Fig. 1, where all components are synchronized and time stamps are attached to data packets upon transmission. The plant is linear with discrete-time dynamics given by

$$\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}_k \underline{\boldsymbol{x}}_k + \mathbf{B}_k \underline{u}_k + \underline{\boldsymbol{w}}_k \,, \tag{1}$$

$$\underline{\boldsymbol{y}}_k = \mathbf{C}_k \underline{\boldsymbol{x}}_k + \underline{\boldsymbol{v}}_k \,, \tag{2}$$

where $\underline{\boldsymbol{x}}_k \in \mathbb{R}^S$ is the plant state, $\underline{u}_k \in \mathbb{R}^L$ denotes the control input, and $\underline{\boldsymbol{y}}_k \in \mathbb{R}^M$ the measurement. The noise processes $\underline{\boldsymbol{w}}_k$ and $\underline{\boldsymbol{v}}_k$ are Gaussian, white, and zero mean with covariance matrices $\mathbf{W}_k$ and $\mathbf{V}_k$, respectively, and for any two time steps k, k', they are mutually independent. The initial plant state $\underline{\boldsymbol{x}}_0$ is assumed to be Gaussian with given mean $\hat{\underline{x}}_0$ and covariance $\mathbf{X}_0$, and independent of $\underline{\boldsymbol{w}}_i$ and $\underline{\boldsymbol{v}}_j$. The sensor is collocated with the plant and sends its measurements to the remote controller.

In the communication channel between the sensor and the controller, the measurements can be delayed or even get lost. As a consequence, none, one, or multiple measurements can arrive at the controller at the same time. The set of received measurements at every time step will be denoted by $\mathcal{Y}_k$ in the following.

The control inputs transmitted from the remote controller to the actuator, which is attached to the plant, are also subject to delays and losses due to the network. Interpreting packet losses as infinite delays enables us to model the delay of a packet sent from the controller to the actuator at time k by the random variable $\boldsymbol{\tau}_k^{CA} \in \mathbb{N}_0$. Additionally, we assume that the $\boldsymbol{\tau}_k^{CA}$ are independent and identically distributed (i.i.d.) and that the corresponding probability mass function (PMF) f^{CA} is known. As a countermeasure against these effects, the controller transmits predicted control inputs for the next N time steps together with the current one. More precisely, at each time step the data packet that is sent to the actuator contains a sequence of $N + 1$ control inputs

$$\underline{U}_k = \left[\underline{u}_{k|k}^{\mathrm{T}} \; \underline{u}_{k+1|k}^{\mathrm{T}} \; \cdots \; \underline{u}_{k+N|k}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{(N+1)L} \,,$$

where $\underline{u}_{k+i|k}$ is the control input computed at time k to be applied at time $k + i$, $i = 0, \ldots, N$. At the plant side, an *active packet dropout strategy* [3] is carried out by the actuator: From the set of received control sequences, only the newest one, i.e., the sequence with the largest time stamp, is buffered. The control inputs from this sequence are then fed into the plant one after another until a more recent sequence arrives at the actuator.

Each time the buffered control sequence is replaced by a newer one, the actuator sends an acknowledgment (ACK) back to the controller. It is pivotal to point out that not every received control sequence is acknowledged, but only the one that is one actually kept, so that the ACKs can be seen as application layer acknowledgments. From the network's perspective, however, they are regular data packets and hence can also be delayed or get lost. Consequently, the controller can receive multiple ACKs at every time step.

With the set of received ACKs, denoted by $\mathcal{A}_k$ in the remainder of this paper, at hand, the controller is able to infer control inputs that were applied in the past. To illustrate this, suppose that $\mathcal{A}_k$ contains an ACK that experienced a delay of two time steps, i.e., it was sent by the actuator at time $k - 2$. Then, if this ACK was issued to signal that $\underline{U}_{k-3}$ was the most recent sequence received at time $k - 2$, the controller can conclude that two time steps ago $\underline{u}_{k-2|k-3}$ was applied to the plant. Owing to consecutive packet losses or large delays, it can happen that the buffered sequence is not replaced early enough, so that no more applicable control inputs are provided by the sequence. In such cases, the default input $\underline{u}_k^{df} = \underline{0}$ is used.

In this setup, we seek to minimize, at every time step, the finite-horizon quadratic cost function

$$\begin{aligned} \mathcal{J}_k^K = \mathrm{E}\big\{ & \underline{\boldsymbol{x}}_{k+K}^{\mathrm{T}} \mathbf{Q}_K \underline{\boldsymbol{x}}_{k+K} \\ & + \textstyle\sum_{n=0}^{K-1} \underline{\boldsymbol{x}}_{k+n}^{\mathrm{T}} \mathbf{Q}_n \underline{\boldsymbol{x}}_{k+n} + \underline{u}_{k+n}^{\mathrm{T}} \mathbf{R}_n \underline{u}_{k+n} \,\big|\, \mathcal{I}_k \big\} \,, \end{aligned} \tag{3}$$

with respect to the control sequences $\underline{U}_{k:k+K-1}$, where $K \in \mathbb{N}$ is the horizon length, $\mathbf{Q}_i \geq 0$, $\mathbf{R}_i > 0$, and

$$\mathcal{I}_k = \left\{ \hat{\underline{x}}_0, \mathbf{X}_0, \underline{U}_{0:k-1}, \mathcal{Y}_{0:k}, \mathcal{A}_{0:k} \right\} \,,$$

is the information set available to the controller.

According to the receding horizon principle, the first element $\underline{U}_k^*$ of the minimizer is then transmitted to the plant and the minimization is carried out again at the next time step. Similar to [10], [18], the first step towards a solution is to perform an appropriate state augmentation.

To conclude this section, we want to stress that UDP-like settings correspond to the case $\mathcal{A}_k = \varnothing$ for all k. Likewise, TCP-like settings, as considered, e.g., in [11], correspond to the case that $\mathcal{A}_k$ solely contains the acknowledgment sent back by the actuator at the time $k - 1$. In this regard, the setup in this paper is more general than the ones usually dealt with in literature.

III. Proposed Approach

In this section, we will first perform a state augmentation allowing us to define a combined model that jointly expresses both the plant dynamics, as given by (1) and (2), and the behavior of the actuator. With this model, the considered

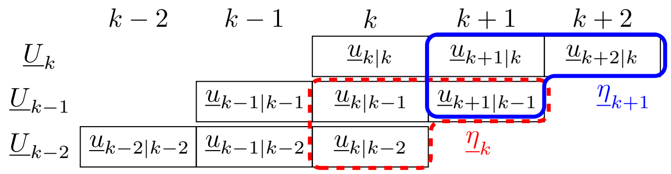

Fig. 2: Illustration of $\underline{\eta}_{k+1}$, $\underline{\eta}_k$, $\underline{U}_k$, and their relationship according to (4) for $N=2$.

NCS can be expressed in terms of a single dynamical system, namely a Markov jump linear system (MJLS) [22] with only partially available mode history.

Based on this model, we will subsequently reformulate the cost function (3) in terms of the augmented state and $\underline{U}_k$. By adapting ideas from [23], we then explicitly approximate the non-convex value function based on a set of coupled quadratic mode-conditioned value functions. The resulting control policy is a linear function of each of the mode-conditioned state estimates naturally provided by our IMM filter [17].

A. Augmented System Dynamics and Problem Reformulation

For the desired holistic description of the NCS, a model that captures the dynamics of the actuator behavior is needed. To that end, as in [11], [17], i) we introduce a vector $\underline{\eta}_k$ that contains all control inputs from *past* control sequences that are still applicable at time k or later, and ii) define a discrete random variable $\boldsymbol{\theta}_k$, which allows for expressing the actually applied input $\underline{u}_k$ in terms of this vector.

In concrete terms, $\underline{\eta}_k$ is given according to

$$\underline{\eta}_k = \begin{bmatrix} \begin{bmatrix} \underline{u}_{k|k-1}^{\mathrm{T}} & \underline{u}_{k+1|k-1}^{\mathrm{T}} & \cdots & \underline{u}_{k+N-1|k-1}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \\ \begin{bmatrix} \underline{u}_{k|k-2}^{\mathrm{T}} & \underline{u}_{k+1|k-2}^{\mathrm{T}} & \cdots & \underline{u}_{k+N-2|k-2}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \\ \vdots \\ \begin{bmatrix} \underline{u}_{k|k-N+1}^{\mathrm{T}} & \underline{u}_{k+1|k-N+1}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}} \\ \underline{u}_{k|k-N} \end{bmatrix} \in \mathbb{R}^{\frac{LN(N+1)}{2}},$$

with dynamics

$$\underline{\eta}_{k+1} = \mathbf{F}\underline{\eta}_k + \mathbf{G}\underline{U}_k\,, \qquad (4)$$

where

$$\mathbf{F} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{(N-1)L} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{(N-2)L} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_L & \mathbf{0} \end{bmatrix},$$

$$\mathbf{G} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_{NL} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

To obtain $\underline{\eta}_{k+1}$, $\mathbf{F}$ removes obsolete entries from $\underline{\eta}_k$, while $\mathbf{G}$ adds the relevant entries from the new sequence $\underline{U}_k$. This is illustrated for $N=2$ in Fig. 2.

Due to the procedure utilized by the actuator, the control input actually applied at time k is either the default input ($\underline{u}_k^{df} = \underline{0}$) or stems from one of the control sequences $\underline{U}_{k-N}, \underline{U}_{k-N+1}, \ldots, \underline{U}_k$. If we set

$$\boldsymbol{\theta}_k = \begin{cases} N+1 & \text{if } \underline{u}_k = \underline{u}_k^{df} \\ k-\tilde{k} & \text{if } \underline{u}_k \text{ is part of } \underline{U}_{\tilde{k}} \end{cases}, \qquad (5)$$

with $k-N \leq \tilde{k} \leq k$, the actually applied input can be written as

$$\underline{\boldsymbol{u}}_k = \mathbf{H}^{(\boldsymbol{\theta}_k)}\underline{\eta}_k + \mathbf{J}^{(\boldsymbol{\theta}_k)}\underline{U}_k\,, \qquad (6)$$

with

$$\begin{aligned} \mathbf{H}^{(\boldsymbol{\theta}_k)} &= \begin{bmatrix} \mathbb{1}_{\boldsymbol{\theta}_k=1}\mathbf{I}_L & \mathbf{0} & \mathbb{1}_{\boldsymbol{\theta}_k=2}\mathbf{I}_L & \mathbf{0} & \ldots & \mathbb{1}_{\boldsymbol{\theta}_k=N}\mathbf{I}_L \end{bmatrix}, \\ \mathbf{J}^{(\boldsymbol{\theta}_k)} &= \begin{bmatrix} \mathbb{1}_{\boldsymbol{\theta}_k=0}\mathbf{I}_L & \mathbf{0} \end{bmatrix}. \end{aligned} \qquad (7)$$

Evidently, from (5), we get that the set of possible values of $\boldsymbol{\theta}_k$ is directly related to the sequence length since it holds $\boldsymbol{\theta}_k \in \{0, \ldots, N+1\}$. Additionally, in [11] it was shown that the process $\{\boldsymbol{\theta}_k\}$ is a Markov chain and that the transition probabilities $p_{ri} = \mathrm{P}[\boldsymbol{\theta}_K = i \mid \boldsymbol{\theta}_{K-1} = r]$ are determined by the PMF f^{CA}.

Equipped with $\underline{\eta}_k$, $\boldsymbol{\theta}_k$, we introduce the augmented state $\underline{\boldsymbol{\psi}}_k = \begin{bmatrix} \underline{\boldsymbol{x}}_k^{\mathrm{T}} & \underline{\eta}_k^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$ and finally arrive at the dynamics

$$\begin{aligned} \underline{\boldsymbol{\psi}}_{k+1} &= \tilde{\mathbf{A}}_k^{(\boldsymbol{\theta}_k)}\underline{\boldsymbol{\psi}}_k + \tilde{\mathbf{B}}_k^{(\boldsymbol{\theta}_k)}\underline{U}_k + \underline{\tilde{\boldsymbol{w}}}_k\,, \\ \underline{\boldsymbol{y}}_k &= \begin{bmatrix} \mathbf{C}_k & \mathbf{0} \end{bmatrix}\underline{\boldsymbol{\psi}}_k + \underline{\boldsymbol{v}}_k\,, \end{aligned} \qquad (8)$$

with

$$\tilde{\mathbf{A}}_k^{(\boldsymbol{\theta}_k)} = \begin{bmatrix} \mathbf{A}_k & \mathbf{B}_k\mathbf{H}^{(\boldsymbol{\theta}_k)} \\ \mathbf{0} & \mathbf{F} \end{bmatrix}, \quad \tilde{\mathbf{B}}_k^{(\boldsymbol{\theta}_k)} = \begin{bmatrix} \mathbf{B}_k\mathbf{J}^{(\boldsymbol{\theta}_k)} \\ \mathbf{G} \end{bmatrix},$$

and zero mean, Gaussian noise $\underline{\tilde{\boldsymbol{w}}}_k = \begin{bmatrix} \underline{\boldsymbol{w}}_k^{\mathrm{T}} & \underline{0}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$. Assuming that the buffer at the actuator is initially empty, $\underline{\boldsymbol{\psi}}_0$ is Gaussian with mean and covariance

$$\underline{\hat{\psi}}_0 = \begin{bmatrix} \underline{\hat{x}}_0 \\ \underline{0} \end{bmatrix}, \quad \Sigma_0 = \begin{bmatrix} \mathbf{X}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

and $\boldsymbol{\theta}_0 = N+1$.

With (8), we are able to describe the whole NCS as a single MJLS. Note that the mode $\boldsymbol{\theta}_k$ only affects the system dynamics and not the measurement equation in (8). Moreover, the mode history, i.e., the true trajectory of the mode up to time k, is not completely available to the controller. Instead only a subset is known to the controller. Recall from Section II that the acknowledgment procedure carried out by the actuator allows the controller to infer control inputs that were actually applied from received ACKs. If, as in the example from Section II, the controller figures out from a received ACK that at time $k-2$ the input $\underline{u}_{k-2|k-3}$ from the sequence $\underline{U}_{k-3}$ was applied to the plant, the corresponding mode realization $\boldsymbol{\theta}_{k-2} = 1$ can be computed using (5).

Hence, at every time step, the controller can infer mode realizations, usually past ones due to the communication delays, only from $\mathcal{A}_k$. From (5), we also get that the mode realization $\boldsymbol{\theta}_k = N+1$ will never be available to the controller since in such cases no applicable sequence would have been received by the actuator in time, and, consequently, no ACK would have been issued.

With the aid of (6), the cost function (3) can be expressed in terms of $\underline{\boldsymbol{\psi}}_k$ and $\underline{U}_k$ according to[2]

$$\begin{aligned}\mathcal{J}_0^K = \mathrm{E}\Big\{&\underline{\boldsymbol{\psi}}_K^{\mathrm{T}}\tilde{\mathbf{Q}}_K\underline{\boldsymbol{\psi}}_K \\ &+\textstyle\sum_{n=0}^{K-1}\underline{\boldsymbol{\psi}}_n^{\mathrm{T}}\tilde{\mathbf{Q}}_n^{(\boldsymbol{\theta}_n)}\underline{\boldsymbol{\psi}}_n+\underline{U}_n^{\mathrm{T}}\tilde{\mathbf{R}}_n^{(\boldsymbol{\theta}_n)}\underline{U}_n\Big|\mathcal{I}_0\Big\}\ ,\end{aligned} \tag{9}$$

with state and input weighting matrices

$$\tilde{\mathbf{Q}}_n^{(\boldsymbol{\theta}_n)} = \begin{bmatrix}\mathbf{Q}_n & \mathbf{0}\\ \mathbf{0} & \left(\mathbf{H}^{(\boldsymbol{\theta}_n)}\right)^{\mathrm{T}}\mathbf{R}_n\mathbf{H}^{(\boldsymbol{\theta}_n)}\end{bmatrix}\ ,\ \ \tilde{\mathbf{Q}}_K = \begin{bmatrix}\mathbf{Q}_K & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\ ,$$
$$\tilde{\mathbf{R}}_n^{(\boldsymbol{\theta}_n)} = \left(\mathbf{J}^{(\boldsymbol{\theta}_n)}\right)^{\mathrm{T}}\mathbf{R}_n\mathbf{J}^{(\boldsymbol{\theta}_n)}\ .$$

Consequently, the restated problem is to minimize (9) at every time step, subject to the dynamics (8) and an only partially available mode history. Moreover, the fact that measurements are available only with a delay or are completely missing must be factored in.

To the best of our knowledge, this problem has, apart from our previous work [18], not yet gained attention in the literature, although an abundance of papers related to the control of MJLS have been published in the last decades. However, focus has been laid on synthesizing controllers when the mode history is either completely known, e.g., [24]–[26], completely unknown, e.g., [23], [27]–[29], or, as in [11], [30], when the true mode values become available only with a fixed delay. Another recent line of research is related to control problems where the true mode value is only known to belong to a certain cluster, that is, subset of values, or can only be observed by a "detector process" with a certain probability [31]–[33].

Denoting the value function at stage n of the optimization horizon by $\mathcal{V}_n$, the considered problem can be addressed within the framework of dynamic programming by solving the Bellman equation

$$\mathcal{V}_n = \min_{\underline{U}_n}\mathrm{E}\Big\{\underline{\boldsymbol{\psi}}_n^{\mathrm{T}}\tilde{\mathbf{Q}}_n^{(\boldsymbol{\theta}_n)}\underline{\boldsymbol{\psi}}_n+\underline{U}_n^{\mathrm{T}}\tilde{\mathbf{R}}_n^{(\boldsymbol{\theta}_n)}\underline{U}_n+\mathcal{V}_{n+1}\Big|\mathcal{I}_n\Big\}\ , \tag{10}$$

recursively backwards in time. Starting at the terminal stage with

$$\mathcal{V}_K = \mathrm{E}\Big\{\underline{\boldsymbol{\psi}}_K^{\mathrm{T}}\tilde{\mathbf{Q}}_K\underline{\boldsymbol{\psi}}_K\Big|\mathcal{I}_K\Big\}\ ,$$

it holds that $\mathcal{V}_0 = \min_{\underline{U}_{0:K-1}}\mathcal{J}_0^K$ [14].

Yet, solving the recursion (10) analytically is not possible due to the dual effect, i.e., the influence of past control inputs on the (expected) future uncertainty of the state estimate. Moreover, numerical optimization already becomes intractable at stage $K-2$, since the dual effect disallows us to exclude the cost related to the future estimation error from the minimization required to determine $\mathcal{V}_{K-2}$. As they are essentially given as a nested product of expectations, the problem is highly nonlinear.

[2]To facilitate the exposition, we from now on restrict ourselves to the initial time step without loss of generality.

B. Obtaining a Tractable Control Policy

Similar to what was done in [23], the track we pursue to obtain a tractable control policy is to approximate the value function (10) at each stage of the optimization horizon using $N+2$ quadratic functions, one for each mode of the MJLS, that are coupled by matrices propagated backwards in time according to mode-conditioned Riccati-like equations. The rationale of the proposed approach is summarized as follows.

First, we define the value function at stage n given that $\boldsymbol{\theta}_n = r$ according to

$$\begin{aligned}\mathcal{V}_n^{(r)} &= \min_{\underline{U}_n}\mathrm{E}\Big\{\underline{\boldsymbol{\psi}}_n^{\mathrm{T}}\tilde{\mathbf{Q}}_n^{(r)}\underline{\boldsymbol{\psi}}_n+\underline{U}_n^{\mathrm{T}}\tilde{\mathbf{R}}_n^{(r)}\underline{U}_n \\ &\qquad +\mathcal{V}_{n+1}\big|\mathcal{I}_n,\boldsymbol{\theta}_n = r\Big\}\ ,\\ \mathcal{V}_K^{(r)} &= \mathrm{E}\Big\{\underline{\boldsymbol{\psi}}_K^{\mathrm{T}}\tilde{\mathbf{Q}}_K\underline{\boldsymbol{\psi}}_K\Big|\mathcal{I}_K,\boldsymbol{\theta}_K = r\Big\}\ .\end{aligned} \tag{11}$$

This enables us to approximate the recurring expressions of the form $\mathrm{E}\{\mathcal{V}_{n+1}|\mathcal{I}_n\}$ using the law of total probability as follows

$$\begin{aligned}&\mathrm{E}\{\mathcal{V}_{n+1}|\mathcal{I}_n\}\\ &=\textstyle\sum_{r=0}^{N+1}\mathrm{P}[\boldsymbol{\theta}_{n+1} = r|\mathcal{I}_n]\mathrm{E}\{\mathcal{V}_{n+1}|\mathcal{I}_n,\boldsymbol{\theta}_{n+1} = r\}\\ &\approx\textstyle\sum_{r=0}^{N+1}\mathrm{P}[\boldsymbol{\theta}_{n+1} = r|\mathcal{I}_n]\mathrm{E}\Big\{\mathcal{V}_{n+1}^{(r)}\big|\mathcal{I}_n,\boldsymbol{\theta}_{n+1} = r\Big\}\ .\end{aligned} \tag{12}$$

Second, we seek quadratic approximations of $\mathcal{V}_n^{(r)}$ for $n = 0, 1, \ldots, K-1$ in the form of

$$\begin{aligned}\mathcal{V}_n^{(r)} \approx\ &\mathrm{E}\Big\{\underline{\boldsymbol{\psi}}_n^{\mathrm{T}}\mathbf{P}_n^{(r)}\underline{\boldsymbol{\psi}}_n\Big|\mathcal{I}_n,\boldsymbol{\theta}_n = r\Big\}\\ &+\mathrm{E}\Big\{\left(\underline{\boldsymbol{e}}_n^{(r)}\right)^{\mathrm{T}}\mathbf{S}_n^{(r)}\underline{\boldsymbol{e}}_n^{(r)}\Big|\mathcal{I}_n,\boldsymbol{\theta}_n = r\Big\}+\omega_n^{(r)}\ ,\end{aligned} \tag{13}$$

with $\mathbf{P}_n^{(r)}, \mathbf{S}_n^{(r)} \geq 0$, $\omega_n^{(r)}$ some nonnegative constant and

$$\underline{\boldsymbol{e}}_n^{(r)} = \underline{\boldsymbol{\psi}}_n - \hat{\underline{\psi}}_n^{(r)}\ ,$$

with $\hat{\underline{\psi}}_n^{(r)} = \mathrm{E}\Big\{\underline{\boldsymbol{\psi}}_n\big|\mathcal{I}_n,\boldsymbol{\theta}_n = r\Big\}$ the mode-conditioned state estimate. Accordingly, the second term in (13) covers the portion of the cost related to estimation error.

Finally, using (12) and (13) in (10) results in the minimization of a quadratic function, which can be carried out analytically, and, consequently, yields a tractable control policy. On the downside, making the above approximations implies that the impact of the dual effect is disregarded. This latter aspect will become obvious in the course of the subsequent presentation of our approach.

Starting at stage $K-1$, we get for the mode-conditioned value function using (12) in (11)

$$\begin{aligned}\mathcal{V}_{K-1}^{(r)} \approx\ &\mathrm{E}\Big\{\underline{\boldsymbol{\psi}}_{K-1}^{\mathrm{T}}\tilde{\mathbf{Q}}_{K-1}^{(r)}\underline{\boldsymbol{\psi}}_{K-1}\Big|\mathcal{I}_{K-1},\boldsymbol{\theta}_{K-1} = r\Big\}\\ &+\min_{\underline{U}_{K-1}}\Big[\underline{U}_{K-1}^{\mathrm{T}}\tilde{\mathbf{R}}_{K-1}^{(r)}\underline{U}_{K-1}\\ &\quad+\textstyle\sum_{i=0}^{N+1}p_{ri}\mathrm{E}\Big\{\mathrm{E}\Big\{\underline{\boldsymbol{\psi}}_K^{\mathrm{T}}\tilde{\mathbf{Q}}_K\underline{\boldsymbol{\psi}}_K\Big|\mathcal{I}_K,\boldsymbol{\theta}_K = i\Big\}\\ &\quad\Big|\mathcal{I}_{K-1},\boldsymbol{\theta}_K = i,\boldsymbol{\theta}_{K-1} = r\Big\}\Big]\ .\end{aligned}$$

Then, we introduce $\mathbf{P}_K^{(i)} = \tilde{\mathbf{Q}}_K$ for $i = 0, \dots, N+1$ and further approximate the above according to

$$\begin{aligned}\mathcal{V}_{K-1}^{(r)} \approx \mathrm{E}\Big\{ \underline{\boldsymbol{\psi}}_{K-1}^{\mathrm{T}} \tilde{\mathbf{Q}}_{K-1}^{(r)} \underline{\boldsymbol{\psi}}_{K-1} \Big| \mathcal{I}_{K-1}, \boldsymbol{\theta}_{K-1} = r \Big\} \\ + \min_{\underline{U}_{K-1}} \Big[\underline{U}_{K-1}^{\mathrm{T}} \tilde{\mathbf{R}}_{K-1}^{(r)} \underline{U}_{K-1} \\ + \mathrm{E}\Big\{ \underline{\boldsymbol{\psi}}_K^{\mathrm{T}} \textstyle\sum_{i=0}^{N+1} p_{ri} \mathbf{P}_K^{(i)} \underline{\boldsymbol{\psi}}_K \Big| \mathcal{I}_{K-1}, \boldsymbol{\theta}_{K-1} = r \Big\} \Big] .\end{aligned} \tag{14}$$

Using the system dynamics (8) in (14) and setting the derivative w.r.t. $\underline{U}_{K-1}$ to zero yields the minimizer

$$\underline{U}_{K-1}^{*,(r)} = -\Big(\mathbf{M}_{K-1}^{(r)}\Big)^{\dagger} \Big(\tilde{\mathbf{B}}_{K-1}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_K^{(r)} \tilde{\mathbf{A}}_{K-1}^{(r)} \hat{\underline{\psi}}_{K-1}^{(r)} , \tag{15}$$

where

$$\mathbf{Y}_K^{(r)} = \textstyle\sum_{i=0}^{N+1} p_{ri} \mathbf{P}_K^{(i)} , \tag{16}$$

$$\mathbf{M}_{K-1}^{(r)} = \tilde{\mathbf{R}}_{K-1}^{(r)} + \Big(\tilde{\mathbf{B}}_{K-1}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_K^{(r)} \tilde{\mathbf{B}}_{K-1}^{(r)} . \tag{17}$$

Notice that the pseudoinverse is required here since $\mathbf{M}_{K-1}^{(r)}$ is generally only positive semidefinite. Plugging (15) back into (14) leads to the desired quadratic expression

$$\begin{aligned}\mathcal{V}_{K-1}^{(r)} \approx \mathrm{E}\Big\{ \underline{\boldsymbol{\psi}}_{K-1}^{\mathrm{T}} \mathbf{P}_{K-1}^{(r)} \underline{\boldsymbol{\psi}}_{K-1} \Big| \mathcal{I}_{K-1}, \boldsymbol{\theta}_{K-1} = r \Big\} \\ + \mathrm{E}\Big\{ \Big(\underline{\boldsymbol{e}}_{K-1}^{(r)}\Big)^{\mathrm{T}} \mathbf{S}_{K-1}^{(r)} \underline{\boldsymbol{e}}_{K-1}^{(r)} \Big| \mathcal{I}_{K-1}, \boldsymbol{\theta}_{K-1} = r \Big\} + \omega_{K-1}^{(r)} ,\end{aligned}$$

with

$$\begin{aligned}\mathbf{P}_{K-1}^{(r)} &= \tilde{\mathbf{Q}}_{K-1}^{(r)} + \Big(\tilde{\mathbf{A}}_{K-1}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_K^{(r)} \tilde{\mathbf{A}}_{K-1}^{(r)} - \mathbf{S}_{K-1}^{(r)} , \\ \mathbf{S}_{K-1}^{(r)} &= \Big(\tilde{\mathbf{A}}_{K-1}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_K^{(r)} \tilde{\mathbf{B}}_{K-1}^{(r)} \Big(\mathbf{M}_{K-1}^{(r)}\Big)^{\dagger} \Big(\tilde{\mathbf{B}}_{K-1}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_K^{(r)} \tilde{\mathbf{A}}_{K-1}^{(r)} , \\ \omega_{K-1}^{(r)} &= \mathrm{E}\Big\{ \underline{\tilde{\boldsymbol{w}}}_{K-1}^{\mathrm{T}} \mathbf{Y}_K^{(r)} \underline{\tilde{\boldsymbol{w}}}_{K-1} \Big| \mathcal{I}_{K-1}, \boldsymbol{\theta}_{K-1} = r \Big\} .\end{aligned}$$

Proceeding now to stage $K-2$, we use the above result in (12) to obtain an approximation of the mode-conditioned value function according to

$$\begin{aligned}\mathcal{V}_{K-2}^{(r)} \approx \mathrm{E}\Big\{ \underline{\boldsymbol{\psi}}_{K-2}^{\mathrm{T}} \tilde{\mathbf{Q}}_{K-2}^{(r)} \underline{\boldsymbol{\psi}}_{K-2} \Big| \mathcal{I}_{K-2}, \boldsymbol{\theta}_{K-2} = r \Big\} \\ + \textstyle\sum_{i=0}^{N+1} p_{ri} \mathrm{E}\Big\{ \omega_{K-1}^{(i)} \Big| \mathcal{I}_{K-2}, \boldsymbol{\theta}_{K-1} = i, \boldsymbol{\theta}_{K-2} = r \Big\} \\ + \min_{\underline{U}_{K-2}} \Big[\underline{U}_{K-2}^{\mathrm{T}} \tilde{\mathbf{R}}_{K-2}^{(r)} \underline{U}_{K-2} \\ + \textstyle\sum_{i=0}^{N+1} p_{ri} \mathrm{E}\Big\{ \mathrm{E}\Big\{ \underline{\boldsymbol{\psi}}_{K-1}^{\mathrm{T}} \mathbf{P}_{K-1}^{(i)} \underline{\boldsymbol{\psi}}_{K-1} \Big| \mathcal{I}_{K-1}, \boldsymbol{\theta}_{K-1} = i \Big\} \\ + \mathrm{E}\Big\{ \Big(\underline{\boldsymbol{e}}_{K-1}^{(i)}\Big)^{\mathrm{T}} \mathbf{S}_{K-1}^{(i)} \underline{\boldsymbol{e}}_{K-1}^{(i)} \Big| \mathcal{I}_{K-1}, \boldsymbol{\theta}_{K-1} = i \Big\} \\ \Big| \mathcal{I}_{K-2}, \boldsymbol{\theta}_{K-1} = i, \boldsymbol{\theta}_{K-2} = r \Big\} \Big] .\end{aligned}$$

Assuming that the portion related to the future estimation error $\underline{\boldsymbol{e}}_{K-1}^{(i)}$ does not depend on $\underline{U}_{K-2}$ allows us to exclude it from the minimization, which, by adopting the line of thought that led to (14), we can carry out to get

$$\begin{aligned}\mathcal{V}_{K-2}^{(r)} \approx \mathrm{E}\Big\{ \underline{\boldsymbol{\psi}}_{K-2}^{\mathrm{T}} \mathbf{P}_{K-2}^{(r)} \underline{\boldsymbol{\psi}}_{K-2} \Big| \mathcal{I}_{K-2}, \boldsymbol{\theta}_{K-2} = r \Big\} \\ + \mathrm{E}\Big\{ \Big(\underline{\boldsymbol{e}}_{K-2}^{(r)}\Big)^{\mathrm{T}} \mathbf{S}_{K-2}^{(r)} \underline{\boldsymbol{e}}_{K-2}^{(r)} \Big| \mathcal{I}_{K-2}, \boldsymbol{\theta}_{K-2} = r \Big\} + \omega_{K-2}^{(r)} ,\end{aligned}$$

with

$$\begin{aligned}\mathbf{P}_{K-2}^{(r)} &= \tilde{\mathbf{Q}}_{K-2}^{(r)} + \Big(\tilde{\mathbf{A}}_{K-2}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_{K-1}^{(r)} \tilde{\mathbf{A}}_{K-2}^{(r)} - \mathbf{S}_{K-2}^{(r)} , \\ \mathbf{S}_{K-2}^{(r)} &= \Big(\tilde{\mathbf{A}}_{K-2}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_{K-1}^{(r)} \tilde{\mathbf{B}}_{K-2}^{(r)} \Big(\mathbf{M}_{K-2}^{(r)}\Big)^{\dagger} \Big(\tilde{\mathbf{B}}_{K-2}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_{K-1}^{(r)} \tilde{\mathbf{A}}_{K-2}^{(r)} , \\ \omega_{K-2}^{(r)} &= \mathrm{E}\Big\{ \underline{\tilde{\boldsymbol{w}}}_{K-2}^{\mathrm{T}} \mathbf{Y}_{K-1}^{(r)} \underline{\tilde{\boldsymbol{w}}}_{K-2} \Big| \mathcal{I}_{K-2}, \boldsymbol{\theta}_{K-2} = r \Big\} \\ &+ \textstyle\sum_{i=0}^{N+1} p_{ri} \mathrm{E}\Big\{ \omega_{K-1}^{(i)} \Big| \mathcal{I}_{K-2}, \boldsymbol{\theta}_{K-1} = i, \boldsymbol{\theta}_{K-2} = r \Big\} \\ &+ \textstyle\sum_{i=0}^{N+1} p_{ri} \mathrm{E}\Big\{ \mathrm{E}\Big\{ \Big(\underline{\boldsymbol{e}}_{K-1}^{(i)}\Big)^{\mathrm{T}} \mathbf{S}_{K-1}^{(i)} \underline{\boldsymbol{e}}_{K-1}^{(i)} \Big| \mathcal{I}_{K-1}, \boldsymbol{\theta}_{K-i} = i \Big\} \\ &\quad \Big| \mathcal{I}_{K-2}, \boldsymbol{\theta}_{K-1} = i, \boldsymbol{\theta}_{K-2} = r \Big\} .\end{aligned}$$

Note that the assumption that the cost related to $\underline{\boldsymbol{e}}_{K-1}^{(i)}$ is independent of $\underline{U}_{K-2}$ implies that the impact of the dual effect is neglected. Using this result, (13) is established for each stage by inductive reasoning.

Hence, at every time step the approximated value function is given by

$$\begin{aligned}\mathcal{V}_k \approx \mathrm{E}\Big\{ \underline{\boldsymbol{\psi}}_k^{\mathrm{T}} \tilde{\mathbf{Q}}_k^{(\boldsymbol{\theta}_k)} \underline{\boldsymbol{\psi}}_k \Big| \mathcal{I}_k \Big\} + \min_{\underline{U}_k} \Big[\underline{U}_k^{\mathrm{T}} \tilde{\mathbf{R}}_k^{(\boldsymbol{\theta}_k)} \underline{U}_k \\ + \textstyle\sum_{r=0}^{N+1} \mathrm{P}[\boldsymbol{\theta}_{k+1} = r \,|\, \mathcal{I}_k] \mathrm{E}\Big\{ \underline{\boldsymbol{\psi}}_{k+1}^{\mathrm{T}} \mathbf{P}_{k+1}^{(r)} \underline{\boldsymbol{\psi}}_{k+1} \\ + \Big(\underline{\boldsymbol{e}}_{k+1}^{(r)}\Big)^{\mathrm{T}} \mathbf{S}_{k+1}^{(r)} \underline{\boldsymbol{e}}_{k+1}^{(r)} + \omega_{k+1}^{(r)} \Big| \mathcal{I}_k, \boldsymbol{\theta}_{k+1} = r \Big\} \Big] ,\end{aligned}$$

by virtue of (12). Assuming again that the cost related to the future estimation error is independent of the control input, plugging in the dynamics (8), and evaluating the necessary optimality condition yields that the optimal control sequence $\underline{U}_k^*$ is determined by the policy

$$\underline{U}_k^* = -\mathbf{L}_k^{\dagger} \textstyle\sum_{r=0}^{N+1} \Big[\sum_{i=0}^{N+1} p_{ir} \mu_k^{(i)} \Big(\tilde{\mathbf{B}}_k^{(i)}\Big)^{\mathrm{T}} \mathbf{P}_{k+1}^{(r)} \tilde{\mathbf{A}}_k^{(i)} \hat{\underline{\psi}}_k^{(i)} \Big] ,$$

where

$$\mathbf{L}_k = \mu_k^{(0)} \tilde{\mathbf{R}}_k^{(0)} + \textstyle\sum_{r=0}^{N+1} \Big[\sum_{i=0}^{N+1} p_{ir} \mu_k^{(i)} \Big(\tilde{\mathbf{B}}_k^{(i)}\Big)^{\mathrm{T}} \mathbf{P}_{k+1}^{(r)} \tilde{\mathbf{B}}_k^{(i)} \Big] ,$$

and $\mu_k^{(i)} = \mathrm{P}[\boldsymbol{\theta}_k = i \mid \mathcal{I}_k]$. To obtain the result we used that $\underline{\boldsymbol{\psi}}_k$ is conditionally independent of $\boldsymbol{\theta}_{k+1}$ given $\boldsymbol{\theta}_k$ and that, due to (7), $\tilde{\mathbf{R}}_k^{(\boldsymbol{\theta}_k)}$ is only non-zero for $\boldsymbol{\theta}_k = 0$.

The policy is linear in the mode-conditioned state estimates $\hat{\underline{\psi}}_k^{(i)}$. The cost matrices $\mathbf{P}_{k+1}^{(r)}$ are given by coupled Riccati-like equations computed backwards in time according to

$$\begin{aligned}\mathbf{P}_{k+n}^{(r)} &= \tilde{\mathbf{Q}}_{k+n}^{(r)} + \Big(\tilde{\mathbf{A}}_{k+n}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_{k+n+1}^{(r)} \tilde{\mathbf{A}}_{k+n}^{(r)} \\ &- \Big(\tilde{\mathbf{A}}_{k+n}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_{k+n+1}^{(r)} \tilde{\mathbf{B}}_{k+n}^{(r)} \Big(\mathbf{M}_{k+n}^{(r)}\Big)^{\dagger} \Big(\tilde{\mathbf{B}}_{k+n}^{(r)}\Big)^{\mathrm{T}} \mathbf{Y}_{k+n+1}^{(r)} \tilde{\mathbf{A}}_{k+n}^{(r)} ,\end{aligned}$$

with $\mathbf{Y}_{k+n+1}^{(r)}$ and $\mathbf{M}_{k+n}^{(r)}$ as in (16), (17), and initial condition $\mathbf{P}_{k+K}^{(r)} = \tilde{\mathbf{Q}}_K$. Both $\hat{\underline{\psi}}_k^{(i)}$ and $\mu_k^{(i)}$ are naturally provided by the IMM filter that we presented in [17]. Additionally, in contrast to our approach in [18], no iterative optimization procedure has to be carried out at runtime so that the computational complexity, which is dominated by the calculation of the required pseudoinverses, is fixed over time. Notice also that for time-invariant plants the matrices $\mathbf{P}_{k+1}^{(r)}$ can be precomputed which further reduces the online complexity.

Being part of CoCPN-Sim, an implementation of the proposed control algorithm is available on github [34].

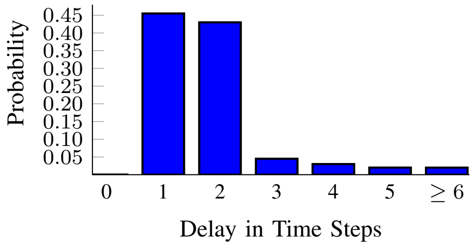

Fig. 3: PMF of the packet delay distribution employed in the example. The notation ≥ 6 is used to denote the probability that a packet is delayed more than five time steps or gets lost (infinite delay).

IV. Numerical Example

In this section, we provide a numerical example to demonstrate the proposed algorithm. We consider a double integrator plant with parameters

$$\mathbf{A}_k = \begin{bmatrix} 1 & t_\mathrm{a} \\ 0 & 1 \end{bmatrix}, \quad \mathbf{B}_k = \begin{bmatrix} 0 \\ 0.5 \end{bmatrix}, \quad \mathbf{W}_k = 0.1 \begin{bmatrix} t_\mathrm{a}^2/2 \\ t_\mathrm{a} \end{bmatrix} \begin{bmatrix} t_\mathrm{a}^2/2 \\ t_\mathrm{a} \end{bmatrix}^\mathrm{T},$$
$$\mathbf{C}_k = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad \mathbf{V}_k = 0.2^2,$$

in (1) and (2), where $t_\mathrm{a} = 0.01\,\mathrm{s}$. The initial plant state is chosen to be

$$\hat{\underline{x}}_0 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}, \quad \mathbf{X}_0 = 0.5^2 \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix},$$

and the weighting matrices in (3) are

$$\mathbf{Q}_k = \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}, \quad \mathbf{R}_k = 1.$$

The distribution specifying the packet delays in the network is chosen such that most of the packets arrive at the respective receiver within at most two time steps. The corresponding PMF depicted in Fig. 3.

To assess the performance of the proposed controller, we compare it with the controller from [18] and a nominal finite-horizon linear quadratic regulator (LQR) that disregards the network. To that end, it is supplied with state estimates by the estimator from [17].

For the comparison, we conduct several Monte Carlo simulations, where the length of the optimization horizon varies between four and nine. In each simulation, 500 runs are carried out, each of which is comprised of 150 time steps. Additionally, in each simulation the length of the control sequence is set to $N+1=4$ and measurements with a delay larger than four time steps are discarded at the controller side. To measure the performance of each controller, we compute the average incurred cost $\mathcal{J}_\mathrm{avg}$ in each run according to

$$\mathcal{J}_\mathrm{avg} = \frac{1}{150}\left[\textstyle\sum_{k=0}^{149} \underline{x}_k^\mathrm{T}\mathbf{Q}_k\underline{x}_k + \underline{u}_k^\mathrm{T}\mathbf{R}_k\underline{u}_k + \underline{x}_{150}^\mathrm{T}\mathbf{Q}_{150}\underline{x}_{150}\right].$$

The resulting medians of the cost $\mathcal{J}_\mathrm{avg}$ are plotted in Fig. 4 and the corresponding standard deviations are given in Table I.

As expected, the results show that the performance of each controller increases with the horizon length. The results also show that the proposed controller performs better than the nominal LQR for all considered horizon lengths and the controller from [18] for all but the largest horizon length with

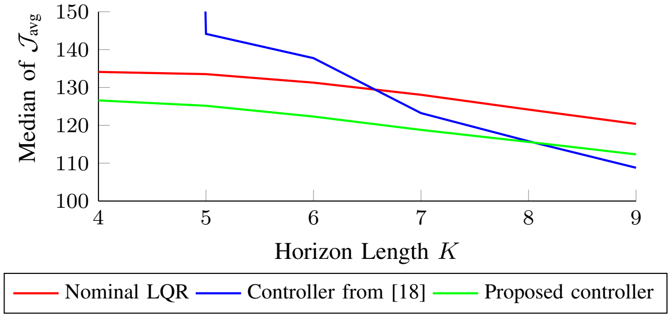

Fig. 4: Medians of the average incurred cost $\mathcal{J}_\mathrm{avg}$.

regards to the incurred cost. Also, the standard deviation of the cost is the lowest for each horizon length. This indicates that the impact of the dual effect is negligible when small optimization horizons are considered.

Notice that the controller from [18] performs significantly worse than the others for $K=4$, which is the reason why the corresponding median (1317.3) is not shown in Fig. 4. This is also reflected by the huge standard deviations for all horizon lengths. This outcome is somewhat unexpected since this controller was designed to take the dual effect into account. On the other hand, this controller also exhibits the strongest performance increase for the larger horizons compared to the others. It is reasonable to expect that its performance will be markedly better than that of the proposed one when the optimization horizon is further increased and the impact of the dual effect on the cost becomes more significant. Evaluating the performance of the proposed controller in greater detail, in particular regarding the degree of suboptimality resulting from the negligence of the dual effect, constitutes important future work.

V. Conclusions

In this work, we dealt with sequence-based stochastic receding horizon control over networks, where application layer acknowledgments are issued by the plant upon reception of control inputs. This enables them to carry additional information for the controller, but, on the downside, requires that delays and losses of acknowledgments be factored in. In particular, the resulting dual effect renders the computation of optimal control policies intractable. To derive a tractable control policy, we relied on the particular representation of the state estimate, namely a Gaussian mixture, provided by the IMM-based filter we proposed in previous work. By approximating the non-convex value function by a set of quadratic functions, one for each component of the mixture, we obtained a control policy that is linear in each of the means of the mixture. In a numerical example, the proposed approach achieved better control performance than a nominal linear quadratic regulator that used same state estimator and the controller we presented in prior work.

Future research should target the reasonableness of the approximations made during the derivation of the proposed approach and find conditions under which stability of the closed-loop system can be established. Likewise, investigating the stability properties of the complementing state estimator

TABLE I: Standard deviations of the average cost $\mathcal{J}_{\text{avg}}$.

	$K=4$	$K=5$	$K=6$	$K=7$	$K=8$	$K=9$
Nominal LQR	29.394	29.091	28.486	27.721	26.87	26.069
Controller from [18]	55 219	242.71	292.75	46.897	47.203	136.44
Proposed controller	27.17	26.753	26.155	25.456	24.726	24.003

is an important future research topic. Finally, it seems worthwhile to embed the proposed controller into the framework of event-triggered control.

REFERENCES

[1] R. Y. Zhong, X. Xu, E. Klotz, and S. T. Newman, "Intelligent Manufacturing in the Context of Industry 4.0: A Review," *Engineering*, vol. 3, no. 5, pp. 616 – 630, 2017.

[2] D. Zhang, P. Shi, Q.-G. Wang, and L. Yu, "Analysis and Synthesis of Networked Control Systems: A Survey of Recent Advances and Challenges," *ISA Transactions*, vol. 66, pp. 376 – 392, 2017.

[3] X. Zhang, Q. Han, and X. Yu, "Survey on Recent Advances in Networked Control Systems," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 5, pp. 1740–1752, Oct. 2016.

[4] W. M. H. Heemels, A. R. Teel, N. Van de Wouw, and D. Nesic, "Networked Control Systems with Communication Constraints: Tradeoffs between Transmission Intervals, Delays and Performance," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1781–1796, 2010.

[5] A. Bemporad, "Predictive Control of Teleoperated Constrained Systems with Unbounded Communication Delays," in *Proceedings of the 37th IEEE Conference on Decision and Control*, vol. 2, Dec. 1998, pp. 2133–2138.

[6] V. Gupta, B. Sinopoli, S. Adlakha, A. Goldsmith, and R. Murray, "Receding Horizon Networked Control," in *Proc. Allerton Conf. Commun., Control Comput*, 2006.

[7] D. E. Quevedo and I. Jurado, "Stability of Sequence-Based Control With Random Delays and Dropouts," *IEEE Transactions on Automatic Control*, vol. 59, no. 5, pp. 1296–1302, May 2014.

[8] D. E. Quevedo and D. Nesic, "Input-to-State Stability of Packetized Predictive Control over Unreliable Networks Affected by Packet-Dropouts," *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 370–375, 2011.

[9] G. Liu, "Predictive Controller Design of Networked Systems with Communication Delays and Data Loss," *IEEE Transactions on Circuits and Systems—Part II: Express Briefs*, vol. 57, no. 6, pp. 481–485, Jun. 2010.

[10] M. Dolgov, J. Fischer, and U. D. Hanebeck, "Infinite-Horizon Sequence-based Networked Control without Acknowledgments," in *Proceedings of the 2015 American Control Conference*, Jul. 2015, pp. 402–408.

[11] J. Fischer, A. Hekler, M. Dolgov, and U. D. Hanebeck, "Optimal Sequence-Based LQG Control over TCP-like Networks Subject to Random Transmission Delays and Packet Losses," in *Proceedings of the 2013 American Control Conference*, Jun. 2013, pp. 1543–1549.

[12] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of Control and Estimation Over Lossy Networks," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 163–187, 2007.

[13] E. Garone, B. Sinopoli, and A. Casavola, "LQG Control over Lossy TCP-Like Networks with Probabilistic Packet Acknowledgements," *International Journal of Systems, Control and Communications*, vol. 2, no. 1-3, pp. 55–81, 2010.

[14] Y. Bar-Shalom and E. Tse, "Dual Effect, Certainty Equivalence, and Separation in Stochastic Control," *IEEE Transactions on Automatic Control*, vol. 19, no. 5, pp. 494–500, Oct. 1974.

[15] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, and S. Sastry, "Optimal Linear LQG Control Over Lossy Networks Without Packet Acknowledgment," in *Proceedings of the 45th IEEE Conference on Decision and Control*, Dec. 2006, pp. 392–397.

[16] H. Lin, H. Su, P. Shi, Z. Shu, R. Lu, and Z. Wu, "Optimal Estimation and Control for Lossy Network: Stability, Convergence, and Performance," *IEEE Transactions on Automatic Control*, vol. 62, no. 9, pp. 4564–4579, Sep. 2017.

[17] F. Rosenthal, B. Noack, and U. D. Hanebeck, "State Estimation in Networked Control Systems with Delayed and Lossy Acknowledgments," in *Multisensor Fusion and Integration in the Wake of Big Data, Deep Learning and Cyber Physical System*, S. Lee, H. Ko, and S. Oh, Eds. Cham: Springer International Publishing, 2018, pp. 22–38.

[18] F. Rosenthal, M. Dolgov, and U. D. Hanebeck, "Sequence-Based Receding Horizon Control Over Networks with Delays and Data Losses," in *Proceedings of the 2019 American Control Conference*, Jun. 2019, pp. 1450–1457.

[19] J. Baillieul and P. J. Antsaklis, "Control and Communication Challenges in Networked Real-Time Systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 9–28, 2007.

[20] H. Blom and Y. Bar-Shalom, "The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients," *IEEE Transactions on Automatic Control*, vol. 33, no. 8, pp. 780–783, 1988.

[21] D. P. Bertsekas, "Dynamic Programming and Suboptimal Control: A Survey from ADP to MPC," *European Journal of Control*, vol. 11, no. 4, pp. 310 – 334, 2005.

[22] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-Time Markov Jump Linear Systems*. Springer Science & Business Media, 2006.

[23] L. Campo and Y. Bar-Shalom, "Control of Discrete-Time Hybrid Stochastic Systems," *IEEE Transactions on Automatic Control*, vol. 37, no. 10, pp. 1522–1527, Oct. 1992.

[24] H. J. Chizeck and Y. Ji, "Optimal Quadratic Control of Jump Linear Systems with Gaussian Noise in Discrete-time," in *Proceedings of the 27th IEEE Conference on Decision and Control*, Dec. 1988, pp. 1989–1993.

[25] O. L. V. Costa and E. F. Tuesta, "Finite Horizon Quadratic Optimal Control and a Separation Principle for Markovian Jump Linear Systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 10, pp. 1836–1842, Oct. 2003.

[26] J. Geromel, A. Gonçalves, and A. Fioravanti, "Dynamic Output Feedback Control of Discrete-Time Markov Jump Linear Systems through Linear Matrix Inequalities," *SIAM Journal on Control and Optimization*, vol. 48, no. 2, pp. 573–593, 2009.

[27] A. N. Vargas, J. B. R. do Val, and E. F. Costa, "Receding Horizon Control of Markov Jump Linear Systems Subject to Noise and Unobserved State Chain," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, vol. 4, Dec. 2004, pp. 4381–4386.

[28] A. N. Vargas, D. C. Bortolin, E. F. Costa, and J. B. Val, "Gradient-Based Optimization Techniques for the Design of Static Controllers for Markov Jump Linear Systems with Unobservable Modes," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 28, no. 3, pp. 239–253, 2014.

[29] M. Dolgov, G. Kurz, and U. D. Hanebeck, "Finite-horizon Dynamic Compensation of Markov Jump Linear Systems without Mode Observation," in *Proceedings of the 55th IEEE Conference on Decision and Control*, Dec. 2016, pp. 2757–2762.

[30] J. Xiong and J. Lam, "Stabilization of Discrete-Time Markovian Jump Linear Systems via Time-Delayed Controllers," *Automatica*, vol. 42, no. 5, pp. 747 – 753, 2006.

[31] J. B. do Val and T. Başar, "Receding Horizon Control of Jump Linear Systems and a Macroeconomic Policy Problem," *Journal of Economic Dynamics and Control*, vol. 23, no. 8, pp. 1099 – 1131, 1999.

[32] A. M. de Oliveira, O. L. V. Costa, and J. Daafouz, "Design of Stabilizing Dynamic Output Feedback Controllers for Hidden Markov Jump Linear Systems," *IEEE Control Systems Letters*, vol. 2, no. 2, pp. 278–283, Apr. 2018.

[33] A. de Oliveira and O. Costa, "Mixed Control of Hidden Markov Jump Systems," *International Journal of Robust and Nonlinear Control*, vol. 28, no. 4, pp. 1261–1280, 2018.

[34] M. Jung and F. Rosenthal, "CoCPN-Sim," 2018. [Online]. Available: https://github.com/spp1914-cocpn/cocpn-sim