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 of control inputs that also consists of predictive inputs for the next time steps in addition to the current one. 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 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.\footnote{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, these trades losses for large delays, which is usually undesired in control applications [19].} 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...
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 ($\underline{x}$), and boldface capital letters indicate matrices, e.g., $A$. We use $I_n$ to denote the $n$-dimensional identity matrix, $0$ to denote zero matrices of arbitrary dimension, and a subscript $k$, e.g., $x_k$, to indicate the time step. Transposition of a vector or a matrix is indicated by $x^T$ and $A^T$, and $A \geq 0$ ($A > 0$) means that the matrix $A$ is positive semidefinite (positive definite). Furthermore, $A^+$ denotes the pseudoinverse of $A$. Finally, $I_{i=j}$ is the indicator function, i.e., $I_{i=j} = 1$ 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

$$x_{k+1} = A_k x_k + B_k u_k + w_k, \quad (1)$$

$$y_k = C_k x_k + v_k, \quad (2)$$

where $x_k \in \mathbb{R}^S$ is the plant state, $u_k \in \mathbb{R}^L$ denotes the control input, and $y_k \in \mathbb{R}^M$ the measurement. The noise processes $w_k$ and $v_k$ are Gaussian, white, and zero mean with covariance matrices $W_k$ and $V_k$, respectively, and for any two time steps $k, k'$, they are mutually independent. The initial plant state $x_0$ is assumed to be Gaussian with mean $\bar{x}_0$ and covariance $X_0$, and independent of $w_0$ and $v_0$. The sensor is colocated 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 $Y_k$ 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 $\tau^C_{k} \in \mathbb{N}_0$. Additionally, we assume that the $\tau^C_{k}$ are independent and identically distributed (i.i.d.) and that the corresponding probability mass function (PMF) $f^C_{\tau}$ 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

$$U^*_k = \left[ x_{k+1}^T \cdots x_{k+N}^T \right] \in \mathbb{R}^{(N+1)L},$$

where $x_{k+i}$ 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 $A_k$ 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 $A_k$ 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 $U_{k-3}$ was the most recent sequence received at time $k - 2$, the controller can conclude that two time steps ago $U_{k-2}$ 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 $u_{k+1}^d = 0$ is used.

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

$$J_k = \mathbb{E}\left\{ \sum_{n=0}^{K} x_k^T Q_n x_k + u_k^T R_n u_k + I_k \right\}, \quad (3)$$

with respect to the control sequences $U_{k:k+K-1}$, where $K \in \mathbb{N}$ is the horizon length, $Q_n \succeq 0$, $R_n > 0$, and

$$I_k = \left\{ x_0, U_0, Y_{0:k-1}, A_{0:k} \right\},$$

is the information set available to the controller.

According to the receding horizon principle, the first element $U^*_0$ 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 $A_k = \emptyset$ for all $k$. Likewise, TCP-like settings, as considered, e.g., in [11], correspond to the case that $A_k$ 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
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 \( \bar{U}_k \). 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 \( \eta_k \) 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 \( \theta_k \), which allows for expressing the actually applied input \( \bar{u}_k \) in terms of this vector.

In concrete terms, \( \eta_k \) is given according to

\[
\eta_k = \begin{bmatrix}
U_k^{T} & U_{k+1}^{T} & \cdots & U_{k+N-1}^{T} \\
U_{k-2}^{T} & U_{k-1}^{T} & \cdots & U_{k+N-2}^{T} \\
\vdots & \vdots & \ddots & \vdots \\
U_{k-N+1}^{T} & U_{k-N+2}^{T} & \cdots & U_{k}^{T}
\end{bmatrix}^T
\in \mathbb{R}^{\frac{L-N(N+1)}{2}}
\]

with dynamics

\[
\eta_{k+1} = F \eta_k + G \bar{U}_k,
\]

where

\[
F = \begin{bmatrix}
0 & 0 & 0 & \cdots & 0 & 0 \\
0 & I_{(N-1)L} & 0 & \cdots & 0 & 0 \\
0 & 0 & I_{(N-2)L} & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & I_L & 0 \\
0 & I_{NL} & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & \cdots & 0 & 0
\end{bmatrix}
\]

\[
G = \begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
0 \\
0 & I_L \\
0 & 0
\end{bmatrix}
\]

To obtain \( \eta_{k+1} \), \( F \) removes obsolete entries from \( \eta_k \), while \( G \) adds the relevant entries from the new sequence \( \bar{U}_k \). 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 \( \bar{u}_k = 0 \) or stems from one of the control sequences \( \bar{U}_k \equiv \bar{U}_{k-N}, \bar{U}_{k-N+1}, \ldots, \bar{U}_k \). If we set

\[
\theta_k = \begin{cases} 
0 & \text{if } \bar{u}_k = \bar{u}_k^d \\
N+1 & \text{if } \bar{u}_k \text{ is part of } \bar{U}_k \end{cases}
\]

with \( k-N \leq k \leq k \), the actually applied input can be written as

\[
\bar{u}_k = H(\theta_k) \eta_k + J(\theta_k) \bar{U}_k,
\]

with

\[
H(\theta_k) = \begin{bmatrix}
I_{\theta_k = 1} & 0 \\
I_{\theta_k = 2} & 0 \\
\vdots & \vdots \\
I_{\theta_k = N} & 0 \\
I_{\theta_k = N+1} & 0
\end{bmatrix},
\]

and

\[
J(\theta_k) = \begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
I_L
\end{bmatrix}
\]

Evidently, from (5), we get that the set of possible values of \( \theta_k \) is directly related to the sequence length since it holds \( \theta_k \in \{0,\ldots,N+1\} \). Additionally, in [11] it was shown that the process \( \{\theta_k\} \) is a Markov chain and that the transition probabilities \( p_{\theta_{i+1} \rightarrow \theta_{i+1}} = \Pr[\theta_{k+1} = \theta_{i+1} | \theta_k = \theta_i] \) are determined by the PMF \( f_C^CA \).

Equipped with \( \eta_k, \theta_k \), we introduce the augmented state

\[
\psi_k = \begin{bmatrix}
\bar{z}_k \\
\bar{u}_k
\end{bmatrix}^T
\]

and finally arrive at the dynamics

\[
\psi_{k+1} = \hat{A}(\theta_k) \psi_k + \hat{B}(\theta_k) \bar{U}_k + \bar{w}_k
\]

\[
\psi_k = \begin{bmatrix}
C_k & 0
\end{bmatrix} \psi_k + \nu_k,
\]

with

\[
\hat{A}(\theta_k) = \begin{bmatrix}
A_k & B_k H(\theta_k) \\
0 & F
\end{bmatrix}, \quad \hat{B}(\theta_k) = \begin{bmatrix}
B_k J(\theta_k) \\
G
\end{bmatrix},
\]

and zero mean, Gaussian noise \( \bar{w}_k = [w_k^T v_k^T]^T \). Assuming that the buffer at the actuator is initially empty, \( \psi_0 \) is Gaussian with mean and covariance

\[
\hat{\psi}_0 \sim \mathcal{N}(0, \Sigma_0) = \begin{bmatrix}
\bar{X}_0 & 0 \\
0 & 0
\end{bmatrix},
\]

and \( \theta_0 = N+1 \).

With (8), we are able to describe the whole NCS as a single MJLS. Note that the mode \( \theta_k \) 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 \( \bar{u}_{k-2} \equiv \bar{U}_{k-2}^{[k-2]} \) from the sequence \( \bar{U}_{k-3} \) was applied to the plant, the corresponding mode realization \( \theta_{k-2} = 1 \) 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 \( A_k \). From (5), we also get that the mode realization \( \theta_k = N+1 \) 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 \( \psi_k \) and \( U_k \) according to \(^2\)

\[
J_0^K = E \left\{ \psi_K^T \tilde{Q}_K \psi_K + \sum_{n=0}^{K-1} \psi_n^T \tilde{Q}_n(\theta_n) \psi_n + U_n^T \tilde{R}_n(\theta_n) U_n \left| I_0 \right. \right\},
\]

(9)

with state and input weighting matrices

\[
\tilde{Q}_n(\theta_n) = \begin{bmatrix} Q_n & 0 \\ 0 & (H(\theta_n))^T R_n H(\theta_n) \end{bmatrix}, \quad \tilde{Q}_K = \begin{bmatrix} Q_K & 0 \\ 0 & 0 \end{bmatrix},
\]

\[
\tilde{R}_n(\theta_n) = (J(\theta_n))^T R_n J(\theta_n).
\]

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

\[
\mathcal{V}_n = \min_{U_n} E \left\{ \psi_n^T \tilde{Q}_n(\theta_n) \psi_n + U_n^T \tilde{R}_n(\theta_n) U_n + \mathcal{V}_{n+1} \left| I_n \right. \right\},
\]

(10)

recursively backwards in time. Starting at the terminal stage with

\[
\mathcal{V}_K = E \left\{ \psi_K^T \tilde{Q}_K \psi_K \left| I_K \right. \right\},
\]

it holds that \( \mathcal{V}_0 = \min_{U_0} J_0^K \) [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 \( \theta_n = r \) according to

\[
\mathcal{V}^{(r)}_n = \min_{U_n} E \left\{ \psi_n^T \tilde{Q}_n \psi_n + U_n^T \tilde{R}_n U_n + \mathcal{V}_{n+1} \left| I_n, \theta_n = r \right. \right\},
\]

(11)

\[
\mathcal{V}^{(r)}_K = E \left\{ \psi_K^T \tilde{Q}_K \psi_K \left| I_K, \theta_K = r \right. \right\}.
\]

This enables us to approximate the recurring expressions of the form \( E \left\{ \mathcal{V}_{n+1} \left| I_n \right. \right\} \) using the law of total probability as follows

\[
E \left\{ \mathcal{V}_{n+1} \left| I_n \right. \right\} = \sum_{r=0}^{N+1} P[\theta_{n+1} = r \left| I_n \right.] E \left\{ \mathcal{V}_{n+1} \left| I_n, \theta_{n+1} = r \right. \right\} \approx \sum_{r=0}^{N+1} P[\theta_{n+1} = r \left| I_n \right.] E \left\{ \mathcal{V}^{(r)}_n \left| I_n, \theta_{n+1} = r \right. \right\}.
\]

(12)

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

\[
\mathcal{V}^{(r)}_n \approx E \left\{ \psi_n^T P_n \psi_n \left| I_n, \theta_n = r \right. \right\} + E \left\{ \psi_n^T S_n \psi_n \left| I_n, \theta_n = r \right. \right\} + \omega_n^{(r)},
\]

(13)

with \( P_n, S_n \geq 0, \omega_n^{(r)} \) some nonnegative constant and

\[
\hat{\psi}_n^{(r)} = \psi_n - \hat{\psi}_n^{(r)} \]

with \( \hat{\psi}_n^{(r)} = E \left\{ \psi_n \left| I_n, \theta_n = r \right. \right\} \) 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)

\[
\mathcal{V}^{(r)}_{K-1} \approx E \left\{ \psi_{K-1}^T \tilde{Q}_{K-1} \psi_{K-1} \left| I_{K-1}, \theta_{K-1} = r \right. \right\} + \min_{U_{K-1}} \left[ U_{K-1}^T \tilde{R}_{K-1} U_{K-1} + \sum_{i=0}^{N+1} p_i E \left\{ \psi_{K-1}^T \tilde{Q}_{K-1} \psi_{K-1} \left| I_{K-1}, \theta_{K-1} = i \right. \right\} \left| I_{K-1}, \theta_K = i \right. \right\}.
\]
Then, we introduce $P^{(i)}_K = \tilde{Q}_K$ for $i = 0, \ldots, N + 1$ and further approximate the above according to
\[
\mathcal{V}^{(r)}_{K-1} \approx E \left\{ \psi^T \tilde{Q}^{(r)}_{K-1} \psi_{K-1} | I_{K-1}, \theta_{K-1} = r \right\} \\
+ \min_{\tilde{U}^{(r)}_{K-1}} \left[ \tilde{U}^T_{K-2} \tilde{R}^{(r)}_{K-2} \tilde{U}_{K-2} \right] \\
+ E \left\{ \psi^T \tilde{Q}^{(r)}_{K-1} \psi_{K-1} | I_{K-1}, \theta_{K-1} = r \right\}.
\]

Using the system dynamics (8) in (14) and setting the derivative w.r.t. $U^{(r)}_{K-1}$ to zero yields the minimizer
\[
U^{(r)}_{K-1} = -\left( M^{(r)}_{K-1} \right)^T \sum_{i=0}^{N+1} p_r P^{(i)}_K \psi_{K-1} | I_{K-1}, \theta_{K-1} = r \right\} + \omega^{(r)}_{K-1},
\]
where
\[
M^{(r)}_{K-1} = \tilde{R}^{(r)}_{K-1} + \left( \tilde{B}^{(r)}_{K-1} \right)^T Y^{(r)}_{K} \tilde{A}^{(r)}_{K-1} \psi^{(r)}_{K-1}.
\]

Notice that the pseudoinverse is required here since $M^{(r)}_{K-1}$ is generally only positive semidefinite. Plugging (15) back into (14) leads to the desired quadratic expression
\[
\mathcal{V}^{(r)}_{K-1} \approx E \left\{ \psi^T \tilde{Q}^{(r)}_{K-1} \psi_{K-1} | I_{K-1}, \theta_{K-1} = r \right\} \\
+ E \left\{ \omega^{(r)}_{K-1} \right\}.
\]

with
\[
P^{(r)}_K = \tilde{Q}^{(r)}_{K-1} + \left( \tilde{A}^{(r)}_{K-1} \right)^T Y^{(r)}_{K} \tilde{A}^{(r)}_{K-1} - \sum_{i=0}^{N+1} p_r P^{(i)}_K,
\]
\[
S^{(r)}_K = \left( \tilde{A}^{(r)}_{K-1} \right)^T Y^{(r)}_{K} \tilde{B}^{(r)}_{K-1} \left( M^{(r)}_{K-1} \right)^T \left( \tilde{B}^{(r)}_{K-1} \right)^T Y^{(r)}_{K} \tilde{A}^{(r)}_{K-1},
\]
\[
\omega^{(r)}_{K-1} = E \left\{ \omega_{K-1}^T \omega_{K-1} | I_{K-1}, \theta_{K-1} = r \right\}.
\]

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
\[
\mathcal{V}^{(r)}_{K-2} \approx E \left\{ \psi^T \tilde{Q}^{(r)}_{K-2} \psi_{K-2} | I_{K-2}, \theta_{K-2} = r \right\} \\
+ \sum_{i=0}^{N+1} p_r E \left\{ \omega^{(i)}_{K-1} | I_{K-2}, \theta_{K-1} = i, \theta_{K-2} = r \right\} \\
+ \min_{\tilde{U}^{(r)}_{K-2}} \left[ \tilde{U}^T_{K-2} \tilde{R}^{(r)}_{K-2} \tilde{U}_{K-2} \right] \\
+ \sum_{i=0}^{N+1} p_r E \left\{ \omega_{K-1}^T \omega_{K-1} | I_{K-1}, \theta_{K-1} = i \right\} \\
+ E \left\{ \omega_{K-1}^T S_{K-2} \omega_{K-2} | I_{K-2}, \theta_{K-2} = r \right\}.
\]

Assuming that the portion related to the future estimation error $E^{(r)}_{K-1}$ does not depend on $U^{(r)}_{K-2}$, we use it to exclude it from the minimization, which, by adopting the line of thought that led to (14), we can carry out to get
\[
\mathcal{V}^{(r)}_{K-2} \approx E \left\{ \psi^T P^{(r)}_{K-2} \psi_{K-2} | I_{K-2}, \theta_{K-2} = r \right\} \\
+ E \left\{ \omega_{K-2}^{(r)} \right\}.
\]
Additionally, in each simulation the length of the control horizon varies between four and nine. In each simulation, 500 runs are carried out, each of which is comprised of $K$ optimization steps. The sequence is set to $N = 150$, 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...
is an important future research topic. Finally, it seems worthwhile to embed the proposed controller into the framework of event-triggered control.

REFERENCES


| TABLE I: Standard deviations of the average cost $J_{avg}$. |
|-------------|-------------|-------------|-------------|-------------|-------------|
| $K = 4$     | $K = 5$     | $K = 6$     | $K = 7$     | $K = 8$     | $K = 9$     |
| Controller from [18] | 55.219 | 242.71      | 292.75      | 46.897      | 47.203      |