Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Optimal Sequence-Based Tracking Control over Unreliable Networks

Jörg Fischer * **Maxim Dolgov** * **Uwe D. Hanebeck** *

* *Intelligent Sensor-Actuator-Systems Laboratory (ISAS), Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Germany. (email: Joerg.Fischer@kit.edu, Maxim.Dolgov@kit.edu, Uwe.Hanebeck@ieee.org)*

Abstract: In networked control systems, sequence-based controllers are used to compensate for transmission delays and losses in unreliable data networks. For this purpose, the controller sends not only the current control input to the actuator but also a sequence of predicted control inputs. The additional inputs can be used when subsequent transmissions get delayed or lost. In this paper, the sequence-based method is applied to the problem of trajectory tracking over an unreliable network and an optimal sequenced-based tracking controller is derived. The main advantage of the presented approach is that future information on the reference trajectory can optimally be embedded in the predicted control sequences. Furthermore, the controller can be implemented offline. An interesting result is that the optimal controller can still be separated into a feedback part and a feedforward part (as in standard optimal tracking control) despite of both the unreliable network and the sequence-based method. The performance of the derived tracking controller is demonstrated by Monte Carlo simulations with an inverted pendulum.

Keywords: Networked Control Systems (NCS); Tracking Control; Sequence-Based Control; Transmission Delays; Transmission Losses; Unreliable Networks; LQ-Optimization; TCP-like Protocol, Markov Jump Linear System (MJLS)

1. INTRODUCTION

Using general purpose networks (such as Ethernet TCP-IP) and wireless networks (including WLAN, Bluetooth, and ZigBee) for control purposes allows the realization of highly flexible control structures at low costs. However, these networks are subject to time-varying transmission delays and/or transmission losses. Used within a control loop, measurement and control data can be delayed and/or lost, which is known to have a highly negative impact on the performance and stability properties of the controlled system (Zhang et al. (2001)).

Therefore, the research area of Networked Control Systems (NCS) investigates new control methods to cope with these network-induced effects. The proposed methods mainly concentrate on the problem of stabilizing a possibly unstable system over an unreliable network (see Hespanha et al. (2007) and Zhang et al. (2013) for an overview). The task of tracking control, however, has received much less attention. In tracking control, the objective is to design the controller such that the output of the closed-loop system follows a time-varying trajectory. Generally, this problem is more challenging than the stability problem since the closed-loop system not only has to be stabilized but also has to follow a defined trajectory (Slotine and Li (1991)). This is even more true in the networked setup where control data can be delayed or lost.

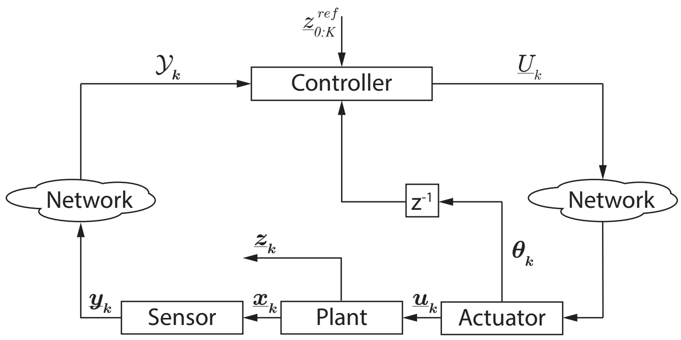

Fig. 1. Considered setup of the Networked Control System.

Recent work that addresses the problem of tracking control over unreliable networks is, e.g., van de Wouw et al. (2010); Gao and Chen (2008); Wang and Yang (2008); Yu et al. (2011). The controller is designed such that the tracking error dynamics are guaranteed to be input-to-state stable (van de Wouw et al. (2010)) or the tracking error is minimized with respect to the H_∞-norm (Gao and Chen (2008); Wang and Yang (2008)), or the H_2-norm (Yu et al. (2011)). In all these approaches, the controller only sends a single control input per data transmission to the actuator.

* This work is supported by the German Science Foundation (DFG) within the Priority Programme 1305 "Control Theory of Digitally Networked Dynamical Systems" and within the Research Training Group RTG 1194 "Self-organizing Sensor-Actuator-Networks".

To further improve tracking performance, it is possible to apply a predictive control strategy called sequence-based control (among others also known as packet-based control) (Bemporad (1998); Liu et al. (2004); Tang and de Silva (2006); Polushin et al. (2008); Findeisen and Varutti (2009); Liu (2010); Quevedo and Nešić (2011); Fischer et al. (2013b)). The idea of this method is that in addition to the current control input the controller also sends predicted control inputs applicable at future time steps. The predicted inputs can be applied by the actuator in case subsequent transmissions get lost or delayed. A highly valuable advantage of the sequence-based method in context of tracking control is that information on the future reference trajectory can already be incorporated in the predicted control sequences. This is interesting, e.g., for robot path planning where reference trajectories are planned ahead and are available for control.

Existing sequence-based control methods for tracking tasks are either based on a Model Predictive Control (MPC) approach (Bemporad (1998); Tang and de Silva (2006)) or on the extension of a nominal controller that is designed by neglecting the networks (Liu (2010)). In the first approach, an optimization problem has to be solved online at each time step which is a time consuming task that restricts its applicability. The latter approach is not applicable if there are external stochastic disturbances affecting the plant. Furthermore, this approach is not optimal even if the nominal controller is constructed by an optimization-based method.

In this paper, we present an optimal sequence-based tracking controller that minimizes the quadratic tracking error between a reference trajectory and the exogenous output of a non-directly observable linear plant that is perturbed by additive measurement and process noise. The considered setup is depicted in Fig. 1. For the network between controller and actuator, we make the assumption that a so called TCP-like protocol is used which provides idealized acknowledgment signals for successful data transfers (see Sec. 2 for more details). The derived tracking controller not only optimally considers transmission delays and losses induced by the networks but is also able to optimally incorporate information on the future reference trajectory (if such information is available). As will be shown, the control law can be computed offline and thus does not require the online solution of an optimization problem. An interesting result of this paper is that, despite of the unreliable networks and the sequence-based method, the optimal solution can be separated into a feedback and a feedforward part. This extends results obtained in Yüksel et al. (2006) where the separation was shown to hold for NCS with limited quantization capacity.

1.1 Key Idea

The key contribution of this paper is to optimally combine the optimal sequence-based control approach that compensates for network-induced time delays and losses (as derived in our previous work Fischer et al. (2013b)) with the idea of tracking control under usage of reference preview. If information on the future reference trajectory is available, this is a natural extension of the optimal sequence-based control approach as the controller already sends sequences of predicted control inputs to the actuator. Embedding preview information does not result in additional communication costs but highly increases the tracking performance.

1.2 Outline

In the following section, the system setup and the sequence-based control method are described, and the tracking problem is formulated. The optimal sequence-based tracking controller is derived in Sec. 3 and its applicability demonstrated by means of a Monte Carlo simulation with an inverted pendulum in Sec. 4.

1.3 Notation

Throughout the paper, deterministic quantities are in normal lettering (a). Random variables are written in bold face letters ($\boldsymbol{a}$) where $\boldsymbol{a} \sim f(a)$ means that $\boldsymbol{a}$ is characterized by its probability density function $f(a)$. Furthermore, vector-valued quantities are underlined ($\underline{a}$), matrices are denoted by bold face capital letters ($\mathbf{A}$), and the notation a_k refers to the quantity a at time step k. The identity matrix is denoted by $\mathbf{I}$, a matrix consisting only of zeros by $\mathbf{0}$, the expectation operator by $\mathrm{E}\{\cdot\}$, the trace operator by $\mathrm{tr}(\cdot)$, the Moore-Penrose pseudoinverse of a matrix $\mathbf{A}$ by $\mathbf{A}^{\dagger}$, and the set of all natural numbers including and excluding zero by $\mathbb{N}_0$ and $\mathbb{N}_{>0}$, respectively.

2. PROBLEM FORMULATION

In this paper, we consider the NCS setup depicted in Fig. 1. It is assumed that all components of the NCS are time-triggered and synchronized. The dynamics of the plant and the sensor are given by

$$\begin{aligned} \underline{\boldsymbol{x}}_{k+1} &= \mathbf{A}\underline{\boldsymbol{x}}_k + \mathbf{B}\underline{\boldsymbol{u}}_k + \underline{\boldsymbol{w}}_k \ , \\ \underline{\boldsymbol{y}}_k &= \mathbf{C}\underline{\boldsymbol{x}}_k + \underline{\boldsymbol{v}}_k \ , \\ \underline{\boldsymbol{z}}_k &= \mathbf{M}\underline{\boldsymbol{x}}_k \ , \end{aligned} \tag{1}$$

where $\underline{\boldsymbol{x}}_k \in \mathbb{R}^n$ is the state of the plant, $\underline{\boldsymbol{u}}_k \in \mathbb{R}^m$ the control input applied to the plant by the actuator, $\underline{\boldsymbol{y}}_k \in \mathbb{R}^p$ the measurement of the state obtained by the sensor, and $\underline{\boldsymbol{z}}_k \in \mathbb{R}^s$ the output considered for the tracking task. The matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{M}$ are of appropriate dimensions and assumed to be known. Model uncertainties and exogenous disturbances affecting system dynamics are represented by mutually independent stationary i.i.d. zero-mean Gaussian noises $\underline{\boldsymbol{w}}_k \sim f(\underline{w}_k)$ and $\underline{\boldsymbol{v}}_k \sim f(\underline{v}_k)$ with covariances

$$\mathbf{W} = \mathrm{E}\left\{\underline{\boldsymbol{w}}_k \underline{\boldsymbol{w}}_k^{\top}\right\} \quad \text{and} \quad \mathbf{V} = \mathrm{E}\left\{\underline{\boldsymbol{v}}_k \underline{\boldsymbol{v}}_k^{\top}\right\} \ .$$

The initial plant state is also assumed to possess Gaussian distribution with

$$\overline{\underline{x}}_0 = \mathrm{E}\left\{\underline{\boldsymbol{x}}_0\right\} \quad \text{and} \quad \mathbf{X}_0 = \mathrm{E}\left\{(\underline{\boldsymbol{x}}_0 - \overline{\underline{x}}_0)(\underline{\boldsymbol{x}}_0 - \overline{\underline{x}}_0)^{\top}\right\} \ .$$

The sensor and the actuator are colocated with the plant. For communication between sensor and controller (SC-link) and between controller and actuator (CA-link) a digital network is used. Data is transmitted over the network in time-stamped packets that can be subject to stochastic time delays and stochastic losses. The links are modeled as two different stochastic networks [1]. It is assumed that the delays $\boldsymbol{\tau}_k^{CA} \in \mathbb{N}_0$ induced by the CA-link and the delays $\boldsymbol{\tau}_k^{SC} \in \mathbb{N}_0$ induced by the SC-link are mutually independent and possess known stationary probability distributions $f(\tau_k^{CA})$ and $f(\tau_k^{SC})$, respectively. In this model, packet losses are considered as infinite time delays. Furthermore, the CA-link satisfies following assumption.

Assumption 1. It is assumed that the CA-link acknowledges successful transmissions to the controller and these acknowledgments experience no time delay or loss.

Networks that satisfy Assumption 1 are called TCP-like networks [2]. Although such an assumption is restrictive, analysis of NCS with TCP-like networks can provide insights into the far more complex control of NCS with networks that employ the real TCP/IP.

As described in the introduction, the controller generates and transmits sequences of control inputs $\underline{U}_k$ to the actuator. Each sequence consists of the current control input and $N-1$ predicted control inputs so that

$$\underline{U}_k = \left[\underline{u}_{k|k}^\top \ \underline{u}_{k+1|k}^\top \ \cdots \ \underline{u}_{k+N-1|k}^\top\right]^\top .$$

The first part of an index $k+i|k$ with $i \in \{0, 1, \ldots, N-1\}$ identifies the time step when the control input is intended to be applied to the plant. The second part of the index refers to the time step of generation of the control input. The actuator is equipped with a buffer that stores the control sequence with the most recent information (among all received sequences). At each time step, the actuator selects the appropriate control input from the buffered sequence and applies it to the plant. If the buffer runs empty, the actuator applies a default control input $\underline{u}^d = \underline{0}$.

Remark 1. The default control input is chosen to $\underline{u}^d = \underline{0}$ only for the brevity of control law derivation. Other values for $\underline{u}^d$ or a so called hold input strategy (Schenato et al. (2007)) can also be considered within the described framework.

The measurements $\underline{\boldsymbol{y}}_k$ are transmitted to the controller over the stochastic SC-link. Thus, it is possible that the controller receives none, one, or several measurements at a time step. The set of measurements received by the controller at time step k will be denoted by $\mathcal{Y}_k$.

In this paper, the considered control task is to choose the control sequences $\underline{U}_k$ such that the exogenous plant output $\underline{\boldsymbol{z}}_k$ optimally tracks a known reference trajectory $\underline{z}_{0:K}^{ref}, K \in \mathbb{N}_{>0}$.

Remark 2. For the derivation of the control law in Sec. 3 we will assume that $\underline{z}_{0:K}^{ref}$ is fixed. However, as we will see later, the control law can be separated into two parts. If the reference trajectory changes during operation, only one of these parts has to be recomputed.

We define the tracking error at time step k as the difference between the reference value $\underline{z}_k^{ref}$ and the plant output $\underline{\boldsymbol{z}}_k$

$$\underline{\boldsymbol{\Delta}}_k = \underline{\boldsymbol{z}}_k - \underline{z}_k^{ref} .$$

Our goal is to calculate the controller that minimizes the cost function

$$J_0^K = \mathrm{E}\left\{\underline{\boldsymbol{\Delta}}_K^\top \mathbf{Q}_K \underline{\boldsymbol{\Delta}}_K + \sum_{k=0}^{K-1}\left[\underline{\boldsymbol{\Delta}}_k^\top \mathbf{Q}_k \underline{\boldsymbol{\Delta}}_k + \underline{u}_k^\top \mathbf{R}_k \underline{u}_k\right] \middle| \mathcal{I}_0\right\} \tag{2}$$

that measures performance w.r.t. the quadratic tracking error and the energy consumed by the control. In (2), $K \in \mathbb{N}_{>0}$ is the considered horizon length, $\mathbf{Q}_k \in \mathbb{R}^{s\times s}$ is positive semidefinite, and $\mathbf{R}_k \in \mathbb{R}^{m\times m}$ is positive definite. Further, $\mathcal{I}_k$ denotes the information available to the controller at time step k with

$$\mathcal{I}_k = \left\{\underline{\overline{x}}_0, \mathbf{X}_0, \underline{z}_{0:K}^{ref}, \mathcal{Y}_{1:k}, \underline{U}_{0:k-1}, \theta_{0:k-1}\right\} .$$

The term $\theta_k \in \{0, \ldots, N\}$ represents the *age* of the control sequence buffered by the actuator at time step k, i.e., the difference of the current time k and the time step at which the buffered sequence was generated. This information is available to the controller due to Assumption 1. In conclusion, for optimally tracking the reference trajectory, we seek to solve following optimization problem

$$J_0^* = \min_{\underline{U}_{0:K-1}} J_0^K . \tag{3}$$

The solution to this problem is given in the next section.

3. DERIVATION OF THE CONTROL LAW

In this section, we solve the optimization problem (3). For this purpose, we first express the optimization problem in a recursive formulation in Sec. 3.1 and model the considered NCS as a Markov Jump Linear System (MJLS) in Sec. 3.2. Finally, in Sec. 3.3, the optimization problem is solved and the main results of this paper are summarized.

3.1 Reformulation of the Optimization Problem

To solve the optimization problem (3), we use the dynamic programming algorithm and, therefore, divide the complex optimization problem into several recursively coupled subproblems. This is provided by the following proposition.

Proposition 1. The optimal cost J_0^* of the problem (3) is equal to C_0 given by the last step of the following recursion

$$\begin{aligned} C_K &= \mathrm{E}\left\{\underline{\boldsymbol{\Delta}}_K^\top \mathbf{Q}_K \underline{\boldsymbol{\Delta}}_K \middle| \mathcal{I}_K\right\} , \\ C_k &= \min_{\underline{U}_{0:K-1}} \mathrm{E}\left\{\underline{\boldsymbol{\Delta}}_k^\top \mathbf{Q}_k \underline{\boldsymbol{\Delta}}_k + \underline{u}_k^\top \mathbf{R}_k \underline{u}_k + C_{k+1} \middle| \mathcal{I}_k\right\} , \end{aligned} \tag{4}$$

which proceeds backwards in time from time step K to time step zero.

Proof 1. Prop. 1 is a direct application of Bellman's principle of optimality that an optimal solution to an optimization problem consists of optimal solutions of its subproblems (Bertsekas (2000)).

Definition 1. The costs C_k in (4) are referred to as minimal costs-to-go (from time step k to K).

3.2 System Modeling

As mentioned in Sec. 2, the age θ_k of the buffered sequence is known by the controller. In Fischer et al. (2013b), it is shown that θ_k can be modeled as the state of a Markov

[1] This modeling approach represents that in real-world applications information transmission from a base station to a subscriber (downlink) and transmission from the subscriber to the base station (uplink) often have different characteristics.

[2] The term TCP-like does not refer to a real TCP/IP protocol. Its usage only indicates that the network provides acknowledgments.

chain such that the considered NCS can be expressed as a special kind of a MJLS. The transition matrix of the Markov chain that governs the evolution of the MJLS is denoted by $\mathbf{T}$ and satisfies

$$\mathbf{T} = \begin{bmatrix} p_{00} & p_{01} & 0 & 0 & \cdots & 0 \\ p_{10} & p_{11} & p_{12} & 0 & \cdots & 0 \\ p_{20} & p_{21} & p_{22} & p_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ \vdots & \vdots & \vdots & \vdots & & p_{(r-1)(r)} \\ p_{r0} & p_{r1} & p_{r2} & p_{r3} & \cdots & p_{rr} \end{bmatrix},$$

with

$$p_{ji} = \mathrm{Prob}\left[\theta_{k+1} = i | \theta_k = j\right] \ , \qquad r = N+1 \ .$$

The transfer probabilities p_{ji} can be calculated by

$$p_{ji} = \begin{cases} 1 - \sum\limits_{s=0}^{i} q_s & \text{for } i = j+1 \ , \\ q_i & \text{for } i < j \leq N+1 \ , \\ 1 - \sum\limits_{s=0}^{N} q_s & \text{for } i = j = N+1 \ , \end{cases}$$

where q_s denotes the probability that a transmission is delayed for $s \in \mathbb{N}_0$ time steps. These probabilities can be computed since the density $f(\tau_k^{CA})$ is known.

Due to the stochastic nature of the CA-link, it is necessary to consider control inputs from previously sent sequences that still can be applied to the plant. Therefore, we introduce the augmented state of the MJLS according to

$$\underline{\boldsymbol{\xi}}_k = \begin{bmatrix} \underline{\boldsymbol{x}}_k \\ [\underline{u}_{k|k-1}^\top \ \ \underline{u}_{k+1|k-1}^\top \ \cdots \ \underline{u}_{k+N-1|k-1}^\top]^\top \\ [\underline{u}_{k|k-2}^\top \ \ \underline{u}_{k+1|k-2}^\top \ \cdots \ \underline{u}_{k+N-2|k-2}^\top]^\top \\ \vdots \\ [\underline{u}_{k|k-N+1}^\top \ \ \underline{u}_{k+1|k-N+1}^\top]^\top \\ \underline{u}_{k|k-N} \end{bmatrix}.$$

The dynamics of the MJLS are then given by

$$\underline{\boldsymbol{\xi}}_{k+1} = \hat{\mathbf{A}}_k \underline{\boldsymbol{\xi}}_k + \hat{\mathbf{B}}_k \underline{U}_k + \hat{\underline{\boldsymbol{w}}}_k \ ,$$

with

$$\hat{\mathbf{A}}_k = \begin{bmatrix} \mathbf{A} & \mathbf{B}\mathbf{H}_k \\ \mathbf{0} & \mathbf{F} \end{bmatrix} \ , \ \hat{\mathbf{B}}_k = \begin{bmatrix} \mathbf{B}\mathbf{J}_k \\ \mathbf{G} \end{bmatrix} \ , \ \hat{\underline{\boldsymbol{w}}}_k = \begin{bmatrix} \underline{\boldsymbol{w}}_k \\ \underline{0} \end{bmatrix} \ ,$$

where

$$\begin{array}{lc} \#columns\colon & \begin{array}{ccccccc} n & n(N-2) & n & n(N-3) & & n & n \end{array} \hspace{2em} \#rows\colon \\ \mathbf{F} = & \left[\begin{array}{ccccccc} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \end{array}\right] \begin{array}{l} \}n(N-1) \\ \}n(N-2) \\ \}n(N-3) \\ \\ \}n \end{array} , \end{array}$$

$$\begin{array}{lc} \#columns\colon & \begin{array}{cc} n & n(N-1) \end{array} \hspace{2em} \#rows\colon \\ \mathbf{G} = & \left[\begin{array}{cc} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{array}\right] \begin{array}{l} \}n(N-1) \\ \}\frac{n(N-1)(N-2)}{2} \end{array} , \end{array}$$

$$\mathbf{J}_k = \Big[\overbrace{\delta_{(\theta_k,0)}\mathbf{I}}^{n} \ \ \overbrace{\mathbf{0}}^{n(N-1)}\Big], \ \ \delta_{(\theta_k,i)} = \begin{cases} 1 \ , \text{ if } \theta_k = i \\ 0 \ , \text{ if } \theta_k \neq i \end{cases},$$

(#columns: n, $n(N-1)$)

$$\mathbf{H}_k = \Big[\overbrace{\delta_{(\theta_k,1)}\mathbf{I}}^{n} \ \ \overbrace{\mathbf{0}}^{n(N-2)} \ \ \overbrace{\delta_{(\theta_k,2)}\mathbf{I}}^{n} \ \ \overbrace{\mathbf{0}}^{n(N-3)} \ \cdots \ \overbrace{\delta_{(\theta_k,N-1)}\mathbf{I}}^{n}\Big].$$

(#columns: n, $n(N-2)$, n, $n(N-3)$, n)

For detailed derivation of the MJLS, the reader is referred to Fischer et al. (2013b). Finally, we can express the costs-to-go (4) in terms of the augmented state $\underline{\boldsymbol{\xi}}_k$ according to

$$\begin{aligned} C_K &= \mathrm{E}\left\{(\mathbf{M}\underline{\boldsymbol{x}}_K - \underline{z}_K^{ref})^\top \mathbf{Q}_K (\mathbf{M}\underline{\boldsymbol{x}}_K - \underline{z}_K^{ref}) \Big| \mathcal{I}_K \right\} \\ &= \mathrm{E}\left\{(\underline{z}_K^{ref})^\top \mathbf{Q}_K \underline{z}_K^{ref} + \underline{\boldsymbol{\xi}}_K^\top \hat{\mathbf{Q}}_K \underline{\boldsymbol{\xi}}_K - 2(\underline{z}_K^{ref})^\top \overline{\mathbf{Q}}_K \underline{\boldsymbol{\xi}}_K \Big| \mathcal{I}_K \right\}, \end{aligned} \tag{5}$$

$$\begin{aligned} C_k &= \mathrm{E}\left\{(\mathbf{M}\underline{\boldsymbol{x}}_k - \underline{z}_k^{ref})^\top \mathbf{Q}_k (\mathbf{M}\underline{\boldsymbol{x}}_k - \underline{z}_k^{ref}) \right. \\ &\quad \left. + \underline{u}_k^\top \mathbf{R}_k \underline{u}_k + C_{k+1} \big| \mathcal{I}_k \right\} \\ &= \mathrm{E}\left\{(\underline{z}_k^{ref})^\top \mathbf{Q}_k \underline{z}_k^{ref} + \underline{\boldsymbol{\xi}}_k^\top \hat{\mathbf{Q}}_k \underline{\boldsymbol{\xi}}_k - 2(\underline{z}_k^{ref})^\top \overline{\mathbf{Q}}_k \underline{\boldsymbol{\xi}}_k \right. \\ &\quad \left. + \underline{U}_k^\top \hat{\mathbf{R}}_k \underline{U}_k + C_{k+1} \Big| \mathcal{I}_k \right\} \ , \end{aligned} \tag{6}$$

with

$$\begin{aligned} &\hat{\mathbf{Q}}_K = \begin{bmatrix} \mathbf{M}^\top \mathbf{Q}_K \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \ \ \hat{\mathbf{Q}}_k = \begin{bmatrix} \mathbf{M}^\top \mathbf{Q}_k \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_k^\top \mathbf{R}_k \mathbf{H}_k \end{bmatrix}, \\ &\overline{\mathbf{Q}}_K = [\mathbf{Q}_K \mathbf{M} \ \mathbf{0}] \ , \ \ \overline{\mathbf{Q}}_k = [\mathbf{Q}_k \mathbf{M} \ \mathbf{0}] \ , \ \ \hat{\mathbf{R}}_k = \mathbf{J}_k^\top \mathbf{R}_k \mathbf{J}_k \ , \end{aligned} \tag{7}$$

Hence, the cumulated cost function (2) can be written as

$$\begin{aligned} J_0^K = \mathrm{E}\Big\{ & (\underline{z}_K^{ref})^\top \mathbf{Q}_K \underline{z}_K^{ref} + \underline{\boldsymbol{\xi}}_K^\top \hat{\mathbf{Q}}_K \underline{\boldsymbol{\xi}}_K - 2(\underline{z}_K^{ref})^\top \overline{\mathbf{Q}}_K \underline{\boldsymbol{\xi}}_K \\ & + \sum_{k=0}^{K-1} \Big[(\underline{z}_k^{ref})^\top \mathbf{Q}_k \underline{z}_k^{ref} + \underline{\boldsymbol{\xi}}_k^\top \hat{\mathbf{Q}}_k \underline{\boldsymbol{\xi}}_k - 2(\underline{z}_k^{ref})^\top \overline{\mathbf{Q}}_k \underline{\boldsymbol{\xi}}_k \\ & + \underline{U}_k^\top \hat{\mathbf{R}}_k \underline{U}_k \Big] \Big| \mathcal{I}_0 \Big\} \ . \end{aligned}$$

3.3 Optimal Tracking Control Law

Having expressed the considered NCS as a Markov Jump Linear System, we apply dynamic programming theory to calculate the optimal control law.

Remark 3. The optimal tracking control problem of MJLS with reference trajectory preview has also been investigated in Nakura (2008). However, these results cannot be applied directly as the mode of the MJLS is assumed to be known only with a delay of one time step. In addition, the augmented weighting matrices (7) are not positive-definite and, finally, we consider that measurements can get delayed or lost.

The main results of this paper are summarized in the following theorem.

Theorem 1. Consider the problem of minimizing the expected cumulative cost function (2) subject to the setup described in Sec. 3.2. Then,

1. for the minimal expected cumulated costs it holds

$$\begin{aligned} J_0^* = \ & \mathrm{E}\left\{ \underline{\boldsymbol{\xi}}_0^\top \mathbf{P}_0 \underline{\boldsymbol{\xi}}_0 \Big| \mathcal{I}_0 \right\} + \sum_{k=0}^{K-1} \mathrm{E}\left\{ \underline{\varepsilon}_k^\top \overline{\mathbf{P}}_k \underline{\varepsilon}_k \big| \mathcal{I}_0 \right\} + s_0 \\ & - 2 \cdot \underline{\sigma}_0^\top \mathrm{E}\left\{ \underline{\boldsymbol{\xi}}_0 \Big| \mathcal{I}_0 \right\} + \sum_{k=0}^{K-1} \mathrm{E}\left\{ \hat{\underline{\boldsymbol{w}}}_k^\top \mathbf{P}_{k+1} \hat{\underline{\boldsymbol{w}}}_k \Big| \mathcal{I}_0 \right\} \ , \end{aligned}$$

2. the optimal control law that minimizes (2) is

$$\begin{aligned} \underline{U}_k = & -\left(\mathrm{E}\left\{ \hat{\mathbf{R}}_k + \hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{B}}_k \Big| \mathcal{I}_k \right\} \right)^\dagger \\ & \cdot \left[\mathrm{E}\left\{ \hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{A}}_k \Big| \mathcal{I}_k \right\} \cdot \mathrm{E}\left\{ \underline{\boldsymbol{\xi}}_k \Big| \mathcal{I}_k \right\} - \mathrm{E}\left\{ \hat{\mathbf{B}}_k^\top \underline{\boldsymbol{\sigma}}_{k+1} \Big| \mathcal{I}_k \right\} \right], \end{aligned} \tag{8}$$

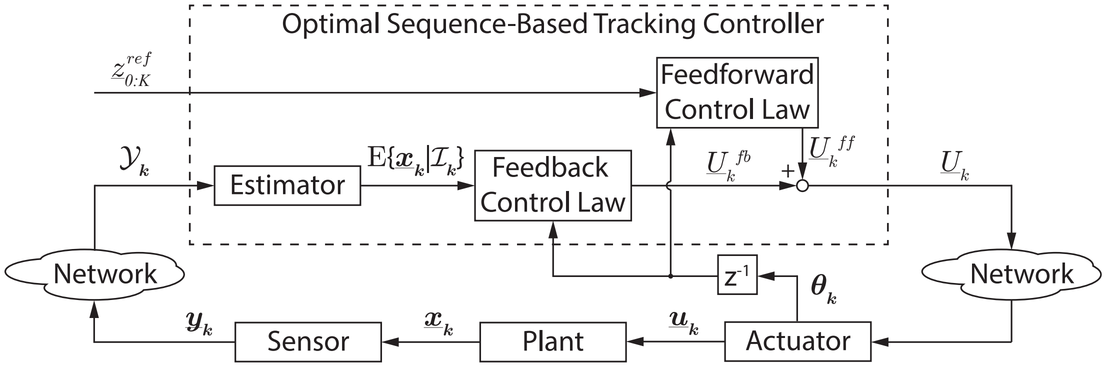

Fig. 2. Structure of the optimal sequence-based controller.

with

$$\boldsymbol{\varepsilon}_k = \underline{\boldsymbol{\xi}}_k - \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_k \middle| \mathcal{I}_k\right\} \ ,$$

$$\begin{aligned} \mathbf{P}_K &= \hat{\mathbf{Q}}_K \ , \quad \overline{\mathbf{P}}_K = \mathbf{0} \ , \\ \underline{\boldsymbol{\sigma}}_K &= \overline{\mathbf{Q}}_K^\top \underline{z}_K^{ref}, \quad \boldsymbol{s}_K = (\underline{z}_K^{ref})^\top \mathbf{Q}_K \underline{z}_K^{ref} \ , \end{aligned} \tag{9}$$

$$\mathbf{P}_k = \mathrm{E}\left\{\hat{\mathbf{Q}}_k + \hat{\mathbf{A}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{A}}_k \middle| \mathcal{I}_k\right\} - \overline{\mathbf{P}}_k \ , \tag{10}$$

$$\begin{aligned} \overline{\mathbf{P}}_k = {} & \mathrm{E}\left\{\hat{\mathbf{A}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{B}}_k \middle| \mathcal{I}_k\right\} \cdot \left(\mathrm{E}\left\{\hat{\mathbf{R}}_k + \hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{B}}_k \middle| \mathcal{I}_k\right\}\right)^\dagger \\ & \cdot \ \mathrm{E}\left\{\hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{A}}_k \middle| \mathcal{I}_k\right\} \ , \end{aligned} \tag{11}$$

$$\begin{aligned} \underline{\boldsymbol{\sigma}}_k = {} & \mathrm{E}\left\{\hat{\mathbf{A}}_k^\top \underline{\boldsymbol{\sigma}}_{k+1} \middle| \mathcal{I}_k\right\} + \overline{\mathbf{Q}}_k^\top \underline{z}_k^{ref} - \mathrm{E}\left\{\hat{\mathbf{A}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{B}}_k \middle| \mathcal{I}_k\right\} \\ & \cdot \left(\mathrm{E}\left\{\hat{\mathbf{R}}_k + \hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{B}}_k \middle| \mathcal{I}_k\right\}\right)^\dagger \cdot \mathrm{E}\left\{\hat{\mathbf{B}}_k^\top \underline{\boldsymbol{\sigma}}_{k+1} \middle| \mathcal{I}_k\right\} , \end{aligned} \tag{12}$$

$$\begin{aligned} \boldsymbol{s}_k = {} & \mathrm{E}\left\{\boldsymbol{s}_{k+1} | \mathcal{I}_k\right\} + (\underline{z}_k^{ref})^\top \mathbf{Q}_k \underline{z}_k^{ref} - \mathrm{E}\left\{\underline{\boldsymbol{\sigma}}_{k+1}^\top \hat{\mathbf{B}}_k \middle| \mathcal{I}_k\right\} \\ & \cdot \ \left(\mathrm{E}\left\{\hat{\mathbf{R}}_k + \hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{B}}_k \middle| \mathcal{I}_k\right\}\right)^\dagger \cdot \mathrm{E}\left\{\hat{\mathbf{B}}_k^\top \underline{\boldsymbol{\sigma}}_{k+1} \middle| \mathcal{I}_k\right\} \ . \end{aligned} \tag{13}$$

Proof 2. The proof of Theorem 1 and the instruction how to calculate the expected values of the involved matrices are given in Appx. A and B.

According to Theorem 1, the optimal control law can therefore be computed by first solving the coupled recursions given by (10) - (13) and then using (8) to calculate the control input sequence. In the following, we analyze the structure of this optimal controller and discuss some further implications of Theorem 1.

Investigating the optimal control sequence (8), it can be seen that it is possible to separate $\underline{U}_k$ into a feedback term $\underline{U}_k^{fb}$ and a feedforward term $\underline{U}_k^{ff}$ such that

$$\underline{U}_k = \underline{U}_k^{fb} + \underline{U}_k^{ff} \ .$$

The feedforward term is given by

$$\underline{U}_k^{ff} = -\left(\mathrm{E}\left\{\hat{\mathbf{R}}_k + \hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{B}}_k \middle| \mathcal{I}_k\right\}\right)^\dagger \cdot \mathrm{E}\left\{\hat{\mathbf{B}}_k^\top \underline{\boldsymbol{\sigma}}_{k+1} \middle| \mathcal{I}_k\right\} ,$$

and only depends on the reference trajectory (via $\underline{\boldsymbol{\sigma}}_{k+1}$) as well as the acknowledgment signal of the CA-link. Hence, $\underline{U}_k^{ff}$ is independent of the augmented system state. On the other side, the feedback term $\underline{U}_k^{fb}$ satisfies

$$\begin{aligned} \underline{U}_k^{fb} = {} & -\left(\mathrm{E}\left\{\hat{\mathbf{R}}_k + \hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{B}}_k \middle| \mathcal{I}_k\right\}\right)^\dagger \\ & \cdot \ \mathrm{E}\left\{\hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{A}}_k \middle| \mathcal{I}_k\right\} \cdot \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_k \middle| \mathcal{I}_k\right\} \ , \end{aligned} \tag{14}$$

and only depends on the acknowledgment signal of the CA-link as well as the augmented state but not on the reference trajectory. The structure of the optimal tracking controller and an overview of the controlled system are depicted in Fig. 2.

The feedback term (14) consists of two parts. The first part is an estimator that computes the conditional expectation of the augmented system state $\mathrm{E}\{\underline{\boldsymbol{\xi}}_k|\mathcal{I}_k\}$. The second part is an optimal sequence-based feedback regulator with feedback matrix

$$\mathbf{L}_k = -\left(\mathrm{E}\left\{\hat{\mathbf{R}}_k + \hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{B}}_k \middle| \mathcal{I}_k\right\}\right)^\dagger \cdot \mathrm{E}\left\{\hat{\mathbf{B}}_k^\top \mathbf{P}_{k+1} \hat{\mathbf{A}}_k \middle| \mathcal{I}_k\right\},$$

that explicitly depends on the acknowledgment θ_{k-1} of the CA-link. The feedback regulator is identical to the controller derived in our previous work Fischer et al. (2013b) on the optimal sequence-based *stabilization* problem. The stability properties of this controller have been investigated in Fischer et al. (2013a) for the infinite horizon case.

The feedback matrix $\mathbf{L}_k$ can be calculated offline for all time steps of the optimization horizon. In addition, if the reference trajectory is known before operation, also the feedforward term $\underline{U}_k^{ff}$ can be calculated in advance. Thus, the complete controller can be designed offline. In case the reference trajectory is only available during operation or the trajectory changes, only the feedforward control term has to be adjusted. The feedback control (14) remains unchanged.

Before we demonstrate the performance of the proposed tracking controller in simulations, we shortly discuss how $\mathrm{E}\{\underline{\boldsymbol{\xi}}_k|\mathcal{I}_k\}$ can be calculated. As the control inputs included in the augmented state at time step k are part of the information set $\mathcal{I}_k$, the problem of calculating $\mathrm{E}\{\underline{\boldsymbol{\xi}}_k|\mathcal{I}_k\}$ reduces to calculating $\mathrm{E}\{\underline{\boldsymbol{x}}_k|\mathcal{I}_k\}$. According to estimation theory, $\mathrm{E}\{\underline{\boldsymbol{x}}_k|\mathcal{I}_k\}$ is equivalent to the minimum mean squared error (MMSE) estimate of the state $\underline{\boldsymbol{x}}_k$. This estimation problem has already been optimally solved in Schenato (2008) where the author uses a time-varying Kalman filter that is extended by a measurement buffer to incorporate delayed measurements.

Mass of the cart	0.5 kg
Mass of the pendulum	0.5 kg
Friction of the cart	0.1 N/(m · s)
Length to pendulum center of mass	0.3 m
Inertia of the pendulum	0.006 kg · m^2
Sampling time	0.1 s

Table 1. Parameters of the inverted pendulum used in the simulation.

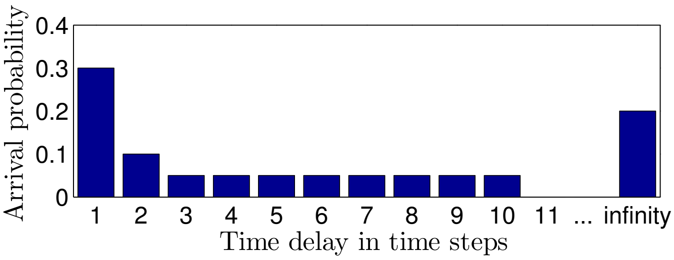

Fig. 3. Probabilities of transmission delays in the network connections considered in the simulation. Transmission losses correspond to infinite time delays.

4. SIMULATION

In this section, we demonstrate the applicability of the proposed tracking controller by simulation with an inverted pendulum on a cart. Modeling the inverted pendulum as described in Anderson (1989), the state is given by

$$\boldsymbol{x}(t) = \left[\boldsymbol{s}(t)\ \dot{\boldsymbol{s}}(t)\ \boldsymbol{\phi}(t)\ \dot{\boldsymbol{\phi}}(t)\right]^\top ,$$

where $\boldsymbol{s}(t)$ is the position of the cart and $\boldsymbol{\phi}(t)$ the angle of the pendulum. With the parameters of the pendulum chosen as shown in Table 1, the matrices of the discrete-time state space model (1) become

$$\mathbf{A} = \begin{bmatrix} 1.0000 & 0.0200 & 0.0015 & 0.0000 \\ 0 & 0.9964 & 0.1550 & 0.0015 \\ 0 & -0.0001 & 1.0103 & 0.0201 \\ 0 & -0.0105 & 0.0343 & 1.0103 \end{bmatrix} ,$$

$$\mathbf{B} = \begin{bmatrix} 0.0004 \\ 0.0358 \\ 0.0011 \\ 0.1054 \end{bmatrix} , \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} .$$

We choose the covariances of the disturbances, the initial condition, and the weighting matrices to

$$\mathbf{W} = \mathbf{V} = \text{diag}\left[0.005^2, 0, (0.2 \cdot \pi/360)^2, 0\right] ,$$

$$\overline{x}_0 = \begin{bmatrix} 0 \\ 0.2 \\ 0 \\ 0.2 \end{bmatrix} , \quad \mathbf{X}_0 = \text{diag}\left[0.01^2, 0, 0.01^2, 0\right] ,$$

$$\mathbf{Q} = \mathbf{R} = \mathbf{M} = \mathbf{I} .$$

In the considered setup, measurements of $\boldsymbol{s}(t)$ and $\boldsymbol{\phi}(t)$ are sent over the same network to the controller. The computed control sequence is then sent to the pendulum over another network. We assume that packet losses and transmission delays occur in both networks independently with probabilities according to Fig. 3. Based on this setup and considering the reference trajectory (as plotted in Fig. 4), the proposed tracking controller is computed for different control sequence lengths. The proposed controller

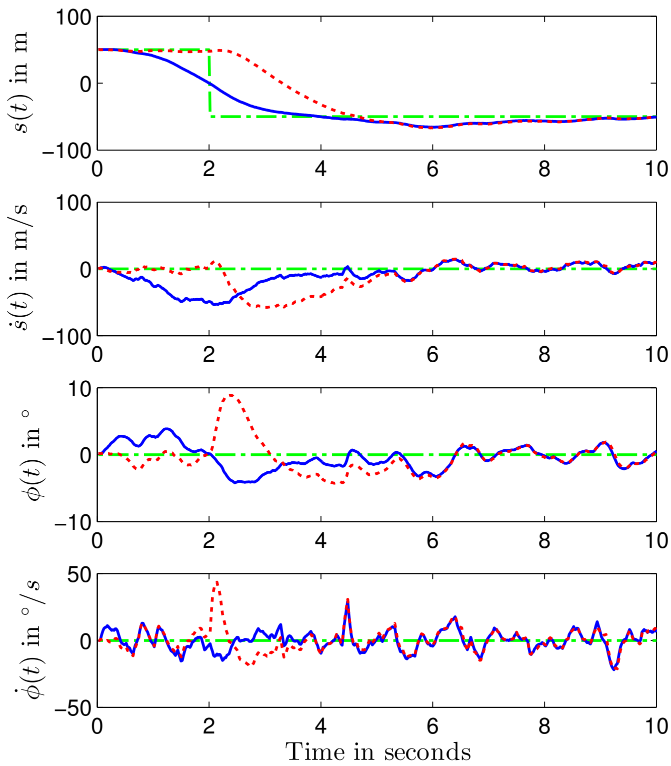

Fig. 4. Example of state trajectory for a simulation with control sequence length of $N = 4$ for proposed controller (—) and control appproach of Liu (2010) (– –) when tracking the reference trajectory (– · –).

is compared with the sequence-based controller described in Liu (2010). The nominal controller required in the approach of Liu (2010) is implemented as an optimal linear quadratic tracking controller (Anderson and Moore (1990)). To estimate the state in presence of possibly delayed or lost measurements, the time-varying Kalman filter of Schenato (2008) is used.

Fig. 4 shows example state trajectories for each of the controllers tracking the depicted reference trajectory. The length of the control sequence is set to $N = 4$. At the beginning, the behavior of both controllers is very similar as they move the cart from the initial position s_0 to the actual reference value of 50. Then, after four seconds, the proposed controller already orientates towards the new reference value of -50 that will be active not before another second. By doing so, the controller prevents the high amplitudes in the angle and angular velocity as they appear with the approach of Liu (2010).

Furthermore, for different control sequence lengths, we conduct 100 Monte Carlo simulation runs over 500 time steps and calculate the average costs according to (2). The results are depicted in Fig. 5. It can be seen that the average costs of both controllers decrease with increasing sequence length. This demonstrates the advantage of the sequence-based method for tracking control over unreliable networks in general. The figure also shows that the proposed tracking approach leads to lower costs than the approach of Liu (2010) that is even unstable if the sequence length is below $N = 3$.

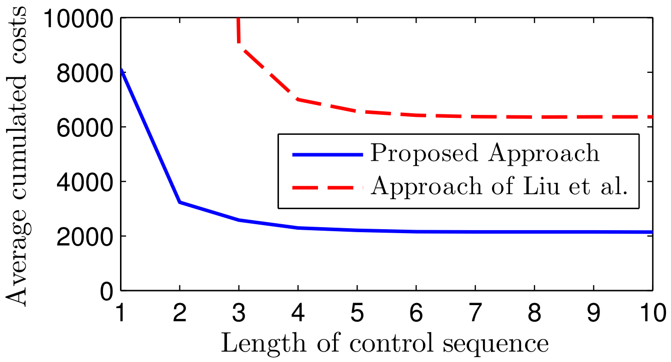

Fig. 5. Comparison of the cumulated average costs over different length of the control sequence for 100 Monte Carlo simulation runs.

5. CONCLUSION

We presented a sequence-based approach for tracking control over unreliable networks that are subject to stochastic time delays and transmission losses. The proposed tracking controller minimize the quadratic tracking error and simultaneously sends sequences of predicted control inputs to compensate for the network effects. In contrast to former work, future information on the reference trajectory can be optimally incorporated in the sequence-based controller design. In simulations with an inverted pendulum, we experienced an improvement of the tracking performance by a factor of three. Furthermore, the proposed tracking controller optimally considers stochastic process and measurement noises and can be calculated offline.

Future work concentrates on the case where the network connections do not provide acknowledgments (UDP-protocol) and on an event-triggered implementation of the controller to reduce communication between controller and actuator. Furthermore, we investigate the stability of the tracking error when the reference trajectory is the output of a linear system.

REFERENCES

Anderson, B.D.O. and Moore, J.B. (1990). *Optimal Control: Linear Quadratic Methods*. Prentice Hall.

Anderson, C.W. (1989). Learning to Control an Inverted Pendulum using Neural Networks. *IEEE Control Systems Magazine*, 9, 31 –37.

Bemporad, A. (1998). Predictive Control of Teleoperated Constrained Systems with Unbounded Communication Delays. In *Proc. Conference on Decision and Control*.

Bertsekas, D.P. (2000). *Dynamic Programming and Optimal Control*. Athena Scientific, Belmont, Massachusetts.

Findeisen, R. and Varutti, P. (2009). Stabilizing Nonlinear Predictive Control over Nondeterministic Communication Networks. In *Nonlinear Model Predictive Control*. Springer Berlin Heidelberg.

Fischer, J., Dolgov, M., and Hanebeck, U.D. (2013a). On Stability of Sequence-Based LQG Control. In *Proceedings of the Conference on Decision and Control*.

Fischer, J., Hekler, A., Dolgov, M., and Hanebeck, U.D. (2013b). Optimal Sequence-Based LQG Control over TCP-like Networks Subject to Random Transmission Delays and Packet Losses. In *Proceedings of the American Control Conference*.

Gao, H. and Chen, T. (2008). Network-Based Output Tracking Control. *IEEE Transactions on Automatic Control*, 53, 655–667.

Hespanha, J.P., Naghshtabrizi, P., and Xu, Y. (2007). A Survey of Recent Results in Networked Control Systems. *Proceedings of the IEEE*, 95, 138–162.

Liu, G.P. (2010). Predictive Controller Design of Networked Systems with Communication Delays and Data Loss. *IEEE Transactions on Circuits and Systems*, 57, 481–485.

Liu, G.P., Mu, J., and Rees, D. (2004). Networked Predictive Control of Systems With Random Communication Delay. In *UKACC International Conference on Control*.

Nakura, G. (2008). Stochastic Optimal Tracking with Preview for Linear Continuous-Time Markovian Jump Systems. In *SICE Annual Conference*.

Polushin, I.G., Liu, P.X., and Lung, C.H. (2008). On the Model-Based Approach to Nonlinear Networked Control Systems. *Automatica*, 44, 2409–2414.

Quevedo, D.E. and Nešić, D. (2011). Input-to-State Stability of Packetized Predictive Control over Unreliable Networks Affected by Packet-Dropouts. *IEEE Transactions on Automatic Control*, 56, 370–375.

Schenato, L., Sinopoli, B., Franceschetti, M., Poolla, K., and Sastry, S. (2007). Foundations of Control and Estimation Over Lossy Networks. *Proceedings of the IEEE*, 95, 163–187.

Schenato, L. (2008). Optimal Estimation in Networked Control Systems Subject to Random Delay and Packet Drop. *IEEE Transactions on Automatic Control*, 53, 1311–1317.

Slotine, J.J.E. and Li, W. (1991). *Applied Nonlinear Control*. Prentice Hall.

Tang, P.L. and de Silva, C.W. (2006). Compensation for Transmission Delays in an Ethernet-Based Control Network using Variable-Horizon Predictive Control. *Transactions on Control Systems Technnology*, 14, 707–718.

van de Wouw, N., Naghshtabrizi, P., Cloosterman, M.B.G., and Hespanha, J.P. (2010). Tracking Control for Sampled-Data Systems with Uncertain Time-Varying Sampling Intervals and Delays. *International Journal of Robust and Nonlinear Control*, 20, 387–411.

Wang, Y.L. and Yang, G.H. (2008). Output Tracking Control for Networked Control Systems with Time Delay and Packet Dropout. *International Journal of Control*, 81, 1709–1719.

Yu, B., Shi, Y., and Lin, Y. (2011). Discrete-time H_2 Output Tracking Control of Wireless Networked Control Systems with Markov Communication Models. *Wireless Communic. and Mobile Computation*, 11, 1107–1116.

Yüksel, S., Hindi, H., and Crawford, L. (2006). Optimal Tracking with Feedback-Feedforward Control Separation over a Network. In *Proceedings of the American Control Conference*.

Zhang, L., Gao, H., and Kaynak, O. (2013). Network-Induced Constraints in Networked Control Systems - A Survey. *IEEE Transactions on Industrial Informatics*, 9, 403–416.

Zhang, W., Branicky, M.S., and Phillips, S.M. (2001). Stability of Networked Control Systems. *IEEE Control Systems Magazine*, 21, 84–99.

Appendix A. PROOF OF THEOREM 1

To derive the results summarized in Theorem 1, we apply stochastic dynamic programming. Using (5) and the definitions (9), the expected costs-to-go at time step K are

$$C_K^* = \mathrm{E}\left\{(\underline{z}_K^{ref})^\top \mathbf{Q}_K \underline{z}_K^{ref} + \underline{\boldsymbol{\xi}}_K^\top \hat{\mathbf{Q}}_K \underline{\boldsymbol{\xi}}_K - 2(\underline{z}_K^{ref})^\top \overline{\mathbf{Q}}_K \underline{\boldsymbol{\xi}}_K \middle| \mathcal{I}_K\right\} = \boldsymbol{s}_K + \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_K^\top \mathbf{P}_K \underline{\boldsymbol{\xi}}_K \middle| \mathcal{I}_K\right\} - 2\cdot \underline{\boldsymbol{\sigma}}_K^\top \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_K \middle| \mathcal{I}_K\right\} . \tag{A.1}$$

In the next step we use that the augmented state $\underline{\boldsymbol{\xi}}_k$ is conditionally independent of the mode θ_{k-1} and that

$$\mathrm{E}\left\{\mathrm{E}\left\{g(\underline{\boldsymbol{\xi}}_{k+1})\middle|\mathcal{I}_{k+1}\right\}\middle|\mathcal{I}_k\right\} = \mathrm{E}\left\{g(\underline{\boldsymbol{\xi}}_{k+1})\middle|\mathcal{I}_k\right\}$$

holds for any function $g(\cdot)$. With these facts and plugging (A.1) into (6), the expected costs-to-go at time step $K-1$ can be written as

$$\begin{aligned}
C_{K-1} &= \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_{K-1}^\top\left(\hat{\mathbf{Q}}_{K-1} + \hat{\mathbf{A}}_{K-1}^\top \mathbf{P}_K \hat{\mathbf{A}}_{K-1}\right)\underline{\boldsymbol{\xi}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}\\
&+ \underline{U}_{K-1}^\top \mathrm{E}\left\{\hat{\mathbf{R}}_{K-1} + \hat{\mathbf{B}}_{K-1}^\top \mathbf{P}_K \hat{\mathbf{B}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}\underline{U}_{K-1}\\
&+ 2\cdot \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_{K-1}^\top\middle|\mathcal{I}_{K-1}\right\}\mathrm{E}\left\{\hat{\mathbf{A}}_{K-1}^\top \mathbf{P}_K \hat{\mathbf{B}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}\underline{U}_{K-1}\\
&- 2\cdot\left[\underline{\boldsymbol{\sigma}}_K^\top \mathrm{E}\left\{\hat{\mathbf{A}}_{K-1}\middle|\mathcal{I}_{K-1}\right\} + (\underline{z}_{K-1}^{ref})^\top \overline{\mathbf{Q}}_{K-1}\right]\\
&\cdot\ \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_{K-1}\middle|\mathcal{I}_{K-1}\right\} - 2\cdot\underline{\boldsymbol{\sigma}}_K^\top \mathrm{E}\left\{\hat{\mathbf{B}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}\underline{U}_{K-1}\\
&+ (\underline{z}_{K-1}^{ref})^\top \mathbf{Q}_{K-1}\underline{z}_{K-1}^{ref} + \mathrm{E}\left\{\underline{\hat{\boldsymbol{w}}}_{K-1}^\top \mathbf{P}_K \underline{\hat{\boldsymbol{w}}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}\\
&+ \boldsymbol{s}_K\ .
\end{aligned} \tag{A.2}$$

Differentiation of (A.2) w.r.t. $\underline{U}_{K-1}$ and setting the result to zero yields

$$\begin{aligned}
\underline{U}_{K-1} = &-\mathrm{E}\left\{\hat{\mathbf{R}}_{K-1} + \hat{\mathbf{B}}_{K-1}^\top \mathbf{P}_K \hat{\mathbf{B}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}^\dagger\\
&\cdot\left[\mathrm{E}\left\{\hat{\mathbf{B}}_{K-1}^\top \mathbf{P}_K \hat{\mathbf{A}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}\mathrm{E}\left\{\underline{\boldsymbol{\xi}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}\right.\\
&\left.-\mathrm{E}\left\{\hat{\mathbf{B}}_{K-1}^\top\middle|\mathcal{I}_{K-1}\right\}\underline{\boldsymbol{\sigma}}_K\right]\ .
\end{aligned} \tag{A.3}$$

Plugging (A.3) in (A.2) and using the definitions (10) - (13), the minimal costs-to-go at time step $K-1$ are

$$\begin{aligned}
C_{K-1}^* = &\ \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_{K-1}^\top \mathbf{P}_{K-1}\underline{\boldsymbol{\xi}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}\\
&- 2\cdot\underline{\sigma}_{K-1}^\top \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}\\
&+ \mathrm{E}\left\{\boldsymbol{\varepsilon}_{K-1}^\top \overline{\mathbf{P}}_{K-1}\boldsymbol{\varepsilon}_{K-1}\middle|\mathcal{I}_{K-1}\right\} + s_{K-1}\\
&+ \mathrm{E}\left\{\underline{\hat{\boldsymbol{w}}}_{K-1}^\top \mathbf{P}_K \underline{\hat{\boldsymbol{w}}}_{K-1}\middle|\mathcal{I}_{K-1}\right\}\ .
\end{aligned} \tag{A.4}$$

Considering one more time step, the expected costs-to-go (6) at time step $K-2$ with (A.4) are given by

$$\begin{aligned}
C_{K-2} &= \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_{K-2}^\top\left(\hat{\mathbf{Q}}_{K-2} + \hat{\mathbf{A}}_{K-2}^\top \mathbf{P}_{K-1}\hat{\mathbf{A}}_{K-2}\right)\underline{\boldsymbol{\xi}}_{K-2}\middle|\mathcal{I}_{K-2}\right\}\\
&+ \underline{U}_{K-2}^\top \mathrm{E}\left\{\hat{\mathbf{R}}_{K-2} + \hat{\mathbf{B}}_{K-2}^\top \mathbf{P}_{K-1}\hat{\mathbf{B}}_{K-2}\middle|\mathcal{I}_{K-2}\right\}\underline{U}_{K-2}\\
&+ 2\cdot\mathrm{E}\left\{\underline{\boldsymbol{\xi}}_{K-2}^\top\middle|\mathcal{I}_{K-2}\right\}\mathrm{E}\left\{\hat{\mathbf{A}}_{K-2}^\top \mathbf{P}_{K-1}\hat{\mathbf{B}}_{K-2}\middle|\mathcal{I}_{K-2}\right\}\underline{U}_{K-2}\\
&-2\cdot\left[\mathrm{E}\left\{\underline{\boldsymbol{\sigma}}_{K-1}^\top \hat{\mathbf{A}}_{K-2}\middle|\mathcal{I}_{K-2}\right\} + (\underline{z}_{K-2}^{ref})^\top \overline{\mathbf{Q}}_{K-2}\right]\\
&\cdot\ \mathrm{E}\left\{\underline{\boldsymbol{\xi}}_{K-2}\middle|\mathcal{I}_{K-2}\right\} - 2\cdot\mathrm{E}\left\{\underline{\boldsymbol{\sigma}}_{K-1}^\top \hat{\mathbf{B}}_{K-2}\middle|\mathcal{I}_{K-2}\right\}\underline{U}_{K-2}\\
&+ \mathrm{E}\left\{\boldsymbol{\varepsilon}_{K-1}^\top \overline{\mathbf{P}}_{K-1}\boldsymbol{\varepsilon}_{K-1}\middle|\mathcal{I}_{K-2}\right\} + (\underline{z}_{K-2}^{ref})^\top \mathbf{Q}_{K-2}\underline{z}_{K-2}^{ref}\\
&+ \mathrm{E}\left\{\boldsymbol{s}_{K-1}\middle|\mathcal{I}_{K-2}\right\} + \sum_{k=K-2}^{K-1}\mathrm{E}\left\{\underline{\hat{\boldsymbol{w}}}_k^\top \mathbf{P}_{k+1}\underline{\hat{\boldsymbol{w}}}_k\middle|\mathcal{I}_{K-2}\right\}\ .
\end{aligned} \tag{A.5}$$

The minimal expected costs-to-go C_{K-2}^* can be calculated analogously to C_{K-1}^* as the structure of (A.4) and (A.5) is the same. Therefore, if we proceed with the minimization, we obtain the recursive solution given in Theorem 1.

Appendix B. COMPUTATION OF EXPECTATIONS

The expected values $\mathrm{E}\{\cdot\,|\,\mathcal{I}_k\}$ of the matrices in (8) and (10) - (13) can be computed by conditioning on the mode $\theta_{k-1} = j$. This is possible since $\theta_{k-1} = j$ is part of the information set $\mathcal{I}_k$ and will be available to the controller at time step k. For a matrix $\mathbf{X}_k(\boldsymbol{\theta}_t)$ and a vector $\underline{\boldsymbol{x}}_k(\boldsymbol{\theta}_t)$ that depend on the random variable $\boldsymbol{\theta}_t$, we introduce the notations $\mathbf{X}_k^{[i,t]}$ and $\underline{\boldsymbol{x}}_k^{[i,t]}$ to denote the resulting quantities when $\theta_t = i$, i.e., $\mathbf{X}_k^{[i,t]} = \mathbf{X}_k(\theta_t = i)$. Furthermore, we define

$$\boldsymbol{\Xi}_k = \left(\sum_{i=0}^N p_{ji}\left[\hat{\mathbf{R}}_k^{[i,k]} + \left(\hat{\mathbf{B}}_k^{[i,k]}\right)^\top \mathbf{P}_{k+1}^{[i,k]}\hat{\mathbf{B}}_k^{[i,k]}\right]\right)^\dagger\ .$$

The expected values can then be calculated by the law of total expectation and are given by

$$\mathbf{L}_k^{[j,k-1]} = -\boldsymbol{\Xi}_k\cdot\left(\sum_{i=0}^N p_{ji}\left[\left(\hat{\mathbf{B}}_k^{[i,k]}\right)^\top \mathbf{P}_{k+1}^{[i,k]}\hat{\mathbf{A}}_k^{[i,k]}\right]\right)\ ,$$

$$\begin{aligned}
\mathbf{P}_k^{[j,k-1]} &= \sum_{i=0}^N p_{ji}\left[\hat{\mathbf{Q}}_k^{[i,k]} + \left(\hat{\mathbf{A}}_k^{[i,k]}\right)^\top \mathbf{P}_{k+1}^{[i,k]}\hat{\mathbf{A}}_k^{[i,k]}\right]\\
&- \left(\sum_{i=0}^N p_{ji}\left[\left(\hat{\mathbf{A}}_k^{[i,k]}\right)^\top \mathbf{P}_{k+1}^{[i,k]}\hat{\mathbf{B}}_k^{[i,k]}\right]\right)\cdot\boldsymbol{\Xi}_k\\
&\cdot\left(\sum_{i=0}^N p_{ji}\left[\left(\hat{\mathbf{B}}_k^{[i,k]}\right)^\top \mathbf{P}_{k+1}^{[i,k]}\hat{\mathbf{A}}_k^{[i,k]}\right]\right),
\end{aligned}$$

$$\begin{aligned}
\underline{\boldsymbol{\sigma}}_k^{[j,k-1]} &= \left(\sum_{i=0}^N p_{ji}\left[\left(\underline{\boldsymbol{\sigma}}_{k+1}^{[i,k]}\right)^\top \hat{\mathbf{A}}_k^{[i,k]}\right]\right) + \left(\underline{z}_k^{ref}\right)^\top \mathbf{Q}_k\\
&- \left(\sum_{i=0}^N p_{ji}\left[\left(\underline{\boldsymbol{\sigma}}_{k+1}^{[i,k]}\right)^\top \hat{\mathbf{B}}_k^{[i,k]}\right]\right)\cdot\boldsymbol{\Xi}_k\\
&\cdot\left(\sum_{i=0}^N p_{ji}\left[\left(\hat{\mathbf{B}}_k^{[i,k]}\right)^\top \mathbf{P}_{k+1}^{[i,k]}\hat{\mathbf{A}}_k^{[i,k]}\right]\right)\ ,
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{s}_k^{[j,k-1]} &= \left(\sum_{i=0}^N p_{ji}\boldsymbol{s}_{k+1}^{[i,k]}\right) - \left(\sum_{i=0}^N p_{ji}\left[\left(\underline{\boldsymbol{\sigma}}_{k+1}^{[i,k]}\right)^\top \hat{\mathbf{B}}_k^{[i,k]}\right]\right)\\
&\cdot\ \boldsymbol{\Xi}_k\left(\sum_{i=0}^N p_{ji}\left[\left(\hat{\mathbf{B}}_k^{[i,k]}\right)^\top \underline{\boldsymbol{\sigma}}_{k+1}^{[i,k]}\right]\right) + \left(\underline{z}_k^{ref}\right)^\top \mathbf{Q}_k\underline{z}_k^{ref}.
\end{aligned}$$