



HAL
open science

Performance Bounds for Stochastic Receding Horizon Control with Randomly Sampled Measurements

Aneel Tanwani, Debasish Chatterjee, Lars Grune

► **To cite this version:**

Aneel Tanwani, Debasish Chatterjee, Lars Grune. Performance Bounds for Stochastic Receding Horizon Control with Randomly Sampled Measurements. 2019 IEEE 58th Conference on Decision and Control (CDC), Dec 2019, Nice, France. pp.2330-2335, 10.1109/CDC40024.2019.9029947 . hal-04866029

HAL Id: hal-04866029

<https://laas.hal.science/hal-04866029v1>

Submitted on 6 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Performance Bounds for Stochastic Receding Horizon Control with Randomly Sampled Measurements

Aneel Tanwani

Debasish Chatterjee

Lars Grüne

Abstract—This article considers the problem of analyzing the performance of model predictive controllers in minimizing infinite-horizon cost functionals associated with stochastic dynamical systems when the measurements received by the controller are randomly sampled in time. In contrast to the standard model predictive control algorithms which rely on availability of the state measurements at all times, we compute control policies which minimize cost functionals over a (finite) rolling-horizon conditioned upon the information that arrives at random time instants; although a hard upper bound equal to the length of the optimization horizon is imposed on consecutive sampling instants. Sufficient conditions are provided on the system dynamics, the cost functionals, and the statistics of the sampling process, such that the proposed policies result in computable upper bounds on the infinite-horizon average cost. The case of linear time-varying system with quadratic cost functionals is studied for the illustration of our results.

I. INTRODUCTION

Developing computational methods for solving optimal control problems for dynamical systems is of interest to several disciplines across engineering and economics. Over the past two decades, model predictive control (MPC) has proven to be one of the most useful paradigms for addressing such problems on a rather wide scale, which comprise different classes of dynamical systems and several generalizations in the description of optimization problems. Some research monographs and survey articles published on this topic provide a comprehensive overview of the field. For deterministic problems, the reader may consult [14], [13], [9]. For stochastic dynamical systems, there are several references which address minimization of finite-horizon cost functions and these tools lie at the core of stochastic MPC problems; see [5] for standard exposition on this topic, [15] for a recent survey, and [3], [11] for an account of earlier developments specific to discrete-time Markov chains. The central element of the predictive control is that we want to address an optimization problem over a large, or infinite, time horizon by solving the problem on a smaller finite horizon, which is computationally tractable. Of course, when we implement the control actions obtained by solving the finite horizon problems,

the resulting cost over the infinite horizon is larger than the minimal cost. One of the central questions is to analyze how much the performance resulting from the computationally tractable algorithms has degraded compared to the theoretically optimal performance. One way to do so is by quantifying the increase in the value of the cost functional obtained by implementing MPC-based policies. Such questions are addressed in [7], [8] and in this article, we will focus on the performance of model predictive control under random time sampling of the state measurements received by the controller (possibly induced by communication over network). Stabilization of dynamical systems with random sampling has been studied in [12], [17], [22], [2], [20], and a survey of literature with some recent results on this topic can be found in [19].

Model predictive control problems in discrete-time settings under the effects of network have been studied in [10], [16], [18], [21], among others. The authors of [21] study the problem of ergodic control where a certain cost is associated to the information sent by the sensors, and the trade-offs between system performance and increased cost due to sensor querying are analyzed. The papers [10], [18] address the stability of the closed-loop dynamical system subject to the uncertainties in communication between the plant and the controller. Since dynamic programming is an essential tool in solving MPC problems, the paper [1] also has some relevance to our work as it deals with computation of finite-horizon optimal controls for continuous-time linear time-invariant systems using dynamic programming under random sampling.

The focus of this article is to analyze the *performance* of MPC-based algorithms subject to random sampling of the measurement process. In Section II, we precisely state the problem formulation and the hypotheses on the system dynamics for our problem setup. The main result in Section III provides upper bounds on the minimal value of the infinite-horizon average cost associated with the trajectories of a controlled stochastic dynamical system in terms of the optimal value associated with a finite-horizon cost functional. The upper bounds, that are obtained for the average cost, depend on the second moment of the noise entering in the dynamics and certain moments associated with the sampling process. As an illustration of the framework adopted in this paper, the case of linear dynamics with quadratic cost functional is studied in Section IV.

A. Tanwani is with LAAS-CNRS, Université de Toulouse, CNRS, 31400, Toulouse. D. Chatterjee is with the Department of Systems and Control Engineering, IIT Bombay, Mumbai 400076, India. L. Grüne is with the Mathematical Institute, University of Bayreuth, 95440 Bayreuth, Germany.

II. SYSTEM CLASS AND PROBLEM FORMULATION

We consider discrete-time stochastic nonlinear systems described by

$$x_{t+1} = f(x_t, u_t, w_t), \quad t \in \mathbb{N} \quad (1)$$

where $x_t \in \mathbb{R}^d$ is the state, and u_t belongs to the admissible control set $\mathbb{U} \subset \mathbb{R}^m$ for each¹ $t \in \mathbb{N}$. The process noise $(w_t)_{t \in \mathbb{N}}$ is assumed to be such that the sequence of random variables w_t , $t \in \mathbb{N}$, is independent and identically distributed (i.i.d.). The initial condition $x_0 \in \mathbb{R}^d$ for the state process $(x_t)_{t \in \mathbb{N}}$ determined by (1) is assumed to be fixed and known.

For the synthesis of control actions for dynamical system (1), we introduce an infinite-horizon average cost functional,

$$J_\infty^{\text{ave}}(x_0, u) := \limsup_{k \rightarrow \infty} \frac{1}{k} \mathbb{E} \left[\sum_{t=0}^{k-1} c(x_t, u_t) \mid x_0 \right], \quad (2)$$

for some measurable function $c : \mathbb{R}^d \times \mathbb{U} \rightarrow \mathbb{R}$. Obtaining a control policy that minimizes the cost functional (2) is often infeasible in practical setups, so we look at the finite-horizon cost functional

$$J_\nu(x_0, u) := \mathbb{E} \left[\sum_{t=0}^{\nu-1} c(x_t, u_t) + c_F(x_\nu) \mid x_0 \right], \quad (3)$$

where we call $\nu \in \mathbb{N}^*$ the optimization horizon. The minimization of the finite-horizon cost functional in (3) is a rather well-studied topic in the literature and is considered numerically more tractable than (2). We seek a control policy that minimizes (3) within a set of admissible *feedback policies*, which is denoted in the sequel by Π . A ν -stage feedback policy $\pi \in \Pi$ is described as $(\pi_0, \pi_1, \dots, \pi_{\nu-1})$ such that $\pi_i : \mathbb{R}^d \rightarrow \mathbb{U}$ is a measurable function. In what follows, it is useful to introduce the *cost-to-go* functions,²

$$V_{\nu-k}(x_k, \pi_{k:\nu-1}) := \mathbb{E}^{\pi_{k:\nu-1}} \left[\sum_{t=k}^{\nu-1} c(x_t, u_t) + c_F(x_\nu) \mid x_k \right], \quad (4)$$

for $0 \leq k \leq \nu - 1$. In the definition (4), x_{t+1} is obtained with the input $u_t = \pi_t(x_t)$ in (1), for $k \leq t \leq \nu - 1$, subject to the state value x_k at time $t = k$. The central ingredient of MPC is to solve the following minimization problem:

$$\min_{\pi \in \Pi} V_\nu(x_0, \pi) \quad \text{subject to (1)}. \quad (5)$$

¹We adopt the convention that $\mathbb{N} = \mathbb{N}^* \cup \{0\}$.

²A slightly general version of the notation introduced in (4) will be used frequently. With $\pi_{\ell_1:\ell_2}$, $\ell_1 \leq \ell_2$, we denote the policies $(\pi_{\ell_1}, \dots, \pi_{\ell_2})$. For appropriately defined measurable functions φ_1, φ_2 , $\mathbb{E}^{\pi_{\ell_1:\ell_2}} \left[\sum_{t=s}^{s+(\ell_2-\ell_1)} \varphi_1(x_t, u_t) + \varphi_2(x_{s+(\ell_2-\ell_1)+1}) \right]$ denotes the expectation where we take $u_s = \pi_{\ell_1}, \dots, u_{s+\ell_2-\ell_1} = \pi_{\ell_2}$, and for $t = s, \dots, s + \ell_2 - \ell_1$, we obtain x_{t+1} from (1) for this choice of u_t , with x_s prespecified.

Remark II.1. We have not included state constraints in the formulation of the optimization problem (5). Including such constraints in the stochastic setting with nonlinear dynamics is rather difficult [4], [6]. However, one can consider the admissible set of policies Π , or the admissible control set \mathbb{U} , to cater for the input constraints.

Let us now discuss the characteristics of the solution to the minimization problem (5). Of particular interest to us is to take into account the information available to the controller for computing policies which minimize (5).

A. Policies with perfect state information

For the sake of completeness, we review the case when the state measurement x_t is available for each $t \in \mathbb{N}$. The point of departure in our treatment is that the problem (5) has a solution in Π .

(H1) The optimization problem (5) is well-defined for each $x_0 \in \mathbb{R}^d$ and admits a solution π^* in the class of permissible feedback policies Π .

Under this assumption, the optimal solution to (5) is a ν -stage feedback policy $\pi_{0:\nu-1}^* := (\pi_0^*, \dots, \pi_{\nu-1}^*)$. The closed-loop system under this policy is described as:

$$x_{t+1}^* = f(x_t^*, \pi_t^*(x_t^*), w_t), \quad x_0^* = x_0 \text{ given, } t = 0, \dots, \nu-1.$$

For each $0 \leq k \leq \nu - 1$, we let $V_{\nu-k}^*(x) := V_{\nu-k}(x, \pi_{k:\nu-1}^*)$, and it is instructive to recall that

$$V_{\nu-k}^*(x) = \mathbb{E}_w [V_{\nu-k-1}^*(f(x, \pi_k^*(x), w_k))] + c(x, \pi_k^*(x)),$$

and the optimal control policy π^* satisfies

$$\pi_k^*(x) := \arg \min_{\pi_k} \left\{ \mathbb{E}_w [V_{\nu-k-1}^*(f(x, \pi_k(x), w_k))] + c(x, \pi_k(x)) \right\}. \quad (6)$$

The receding horizon control policy, $\hat{\pi} := (\pi_0^*, \pi_0^*, \dots)$, is then implemented by choosing

$$u_t = \pi_0^*(x_t), \quad t \in \mathbb{N}. \quad (7)$$

B. Randomly sampled measurements

We are primarily interested in studying the performance of MPC-based algorithms when the information between the plant and the controller is transmitted over a communication channel. As a result, the state measurements may not be communicated to the controller at all time instants either because of the underlying communication protocol, or due to the packet dropouts. For the problem studied in this paper, we model such scenarios by assuming that the state x_t is only available to the controller at some random time instants. These time instants, at which the state measurements are successfully transmitted, are given by a monotonically non-decreasing nonnegative integer-valued sequence $(\tau_n)_{n \in \mathbb{N}}$

taking values in \mathbb{N} . We introduce a nonnegative integer valued stochastic process T_t , defined as $T_t := t - \tau_{N_t}$, with N_t given by

$$N_t := \sup\{n \in \mathbb{N} \mid \tau_n \leq t\} \quad \text{for } t \in \mathbb{N}. \quad (8)$$

For the results in this paper, it is stipulated that

(H2) The sequence of random variables $\{(\tau_{i+1} - \tau_i)\}_{i \in \mathbb{N}}$, is i.i.d., and independent of the process noise. Also, the random variable T_t satisfies the hard bound $T_t \leq \nu - 1$, for each $t \in \mathbb{N}^*$.

A hard bound, equal to the length of optimization horizon, is being imposed on the random sampling intervals; this was also the case in [10] so that the policies *consistent* with MPC could be implemented in case some measurements are not received by the controller. We next explain how this assumption leads to a rather natural choice for control policy.

C. Policies with imperfect state information

Because of the loss of information due to random sampling of the state measurements, the receding horizon policy given in (7) cannot be implemented and hence appropriate adjustments are required in computing the control actions that minimize (5). We now describe an algorithm for computing a ν -step policy \hat{u} , whenever the new state measurements arrive. In contrast to the case with perfect state information, the policies computed here are such that \hat{u}_t depends on a function of $x_{\tau_{N_t}}$, and does not depend on the state values x_t , for $\tau_{N_t} < t < \tau_{N_t+1}$ and we recall that, under **(H2)**, $\tau_{N_t+1} \leq \tau_{N_t} + \nu - 1$.

To understand how we compute the control policies under random sampling, fix (momentarily) $\tau_0 = 0$, and assume that $x(0) = x_0$ is received by the controller at time $t = 0$. The objective is to compute \hat{u}_t , $0 \leq t \leq \nu - 1$ that minimizes (5) conditioned upon randomly occurring sampling events. Adopting the *dynamic programming* route to compute the policy, we introduce the functions, for $0 \leq k \leq \nu - 1$,

$$\hat{V}_{\nu-k}(x_0) = \min_{\pi_{k:\nu-1}} \mathbb{E}^{\pi_{k:\nu-1}} \left[V_{\nu-k}(x_k, \pi_{k:\nu-1}) \mid x_0 \right]. \quad (9)$$

Compared to $V_{\nu-k}^*$, the difference in the definition of the functions $\hat{V}_{\nu-k}$ is due to the expectation conditioned upon $x(0)$, as the former provides the optimal cost to go for a given value of x_k , whereas the later provides the minimal cost to go for an expected value of x_k *given* $\{x_0\}$. In (9), the state x_k is obtained by applying $\hat{u}_{0:k-1}(x_0)$ which are to be computed as a function of last received state measurement x_0 . In particular, $\hat{u}_k(x_0)$ is chosen as a minimizer of $\hat{V}_{\nu-k}(x_0)$, that is,

$$\hat{u}_k(x_0) := \arg \min_{u_k \in \mathbb{U}} \left\{ \mathbb{E} \left[\mathbb{E}_w \left[V_{\nu-k}^*(f(x_k, u_k, w_k)) \right] + c(x_k, u_k) \mid x_0 \right] \right\}, \quad (10)$$

for $0 \leq k \leq \nu - 1$. Note that, by definition (10), $\hat{u}_0(x)$ coincides with $\pi_0^*(x)$, for each $x \in \mathbb{R}^d$. As a result of random sampling, and the aforementioned definition of the policies, the control input is given by

$$u_t = \begin{cases} \pi_0^*(x_t), & \text{if } T_t = 0, \\ \hat{u}_k(x(\tau_{N_t})), & \text{if } T_t = k > 0, \end{cases} \quad (11)$$

where $u_k(x(\tau_{N_t}))$ is defined as in (10) with x_0 replaced by $x(\tau_{N_t})$.

We introduce the following assumption to quantify the increase in the value of the objective function due to the loss of perfect state information.

(H3) For each $x_0 \in \mathbb{R}^d$, and each $k \in \{0, \dots, \nu - 2\}$,

$$\begin{aligned} & \mathbb{E} \left[\mathbb{E}_w \left[V_{\nu-k-1}(f(x_k, \hat{u}_k(x_0), w_k), \pi_{k+1:\nu-1}^*) \mid x_0 \right] \right] + \\ & \mathbb{E} \left[c(x_k, \hat{u}_k(x_0)) \mid x_0 \right] \leq \mathbb{E} \left[V_{\nu-k}^*(x_k) \mid x_0 \right] \\ & + \mathbb{E} \left[\tilde{c}(\hat{u}_k(x_0) - \pi_k^*(x_k)) \mid x_0 \right], \end{aligned} \quad (12a)$$

and the similar inequality holds for $k = \nu - 1$, that is,

$$\begin{aligned} & \mathbb{E} \left[\mathbb{E}_w \left[c_{\text{F}}(f(x_{\nu-1}, \hat{u}_{\nu-1}(x_0), w_{\nu-1})) \mid x_0 \right] \right] \\ & + \mathbb{E} \left[c(x_{\nu-1}, \hat{u}_{\nu-1}(x_0)) \mid x_0 \right] \leq \mathbb{E} \left[V_1^*(x_{\nu-1}) \mid x_0 \right] \\ & + \mathbb{E} \left[\tilde{c}(\hat{u}_{\nu-1}(x_0) - \pi_{\nu-1}^*(x_{\nu-1})) \mid x_0 \right] \end{aligned} \quad (12b)$$

for some real-valued positive-definite function $\tilde{c}: \mathbb{R}^m \rightarrow \mathbb{R}$, and x_k , $1 \leq k \leq \nu - 1$ is obtained from (1) by taking x_0 the value of state at $t = 0$, and using $\hat{u}_0, \dots, \hat{u}_{k-1}$ as the input.

In other words, **(H3)** says that the increase in the value of $V_{\nu-t}$ by applying \hat{u}_t at time $t \geq 0$, compared to the optimal value $V_{\nu-t}^*(x_t)$ obtained by applying $\pi_t^*(x)$, is no more than the expected value of $\tilde{c}(\hat{u}_t(x_0) - \pi_t^*(x_t))$ conditioned upon the last received measurement.

D. Problem formulation

Within the setup described in this section: controlled stochastic process subject to randomly sampled measurements, we are interested in computing the bounds on the infinite-horizon average cost functional J_{∞}^{ave} defined in (2), by implementing the control policies given in (11).

III. BOUNDS ON INFINITE-HORIZON COST

As a solution to the aforementioned problem, the result we obtain eventually appears in Theorem III.2, and we need the following additional assumptions in our problem formulation to state this result:

(H4) There exists a feedback map $g: \mathbb{R}^d \rightarrow \mathbb{U}$, a constant $b \geq 0$, and a bounded set $K \subset \mathbb{R}^d$ such that

$$\sup_{z \in K} \{c(z, g(z)) - c_{\text{F}}(z) + \mathbb{E} [c_{\text{F}}(f(z, g(z), w_0))]\} \leq b, \quad (13a)$$

$$c(z, g(z)) - c_F(z) + \mathbb{E} [c_F(f(z, g(z), w_0))] \leq 0, \quad z \notin K. \quad (13b)$$

Assumption **(H4)** was also used in [7, Theorem 3] for analyzing the performance of MPC with perfect state information. In addition to guaranteeing the existence of a stabilizing feedback law, the assumption **(H4)** also requires the stage cost c and the final cost c_F to be compatible with each other. Due to process noise, it is natural to require the decrease in the cost function only when the state is outside a bounded set. The next assumption **(H5)** also relates the running cost with the final cost, and can be interpreted as the detectability property associated with the stage cost function and becomes useful in relating the cost function with the value function.

(H5) For each $z \in \mathbb{R}^d$ and $u \in \mathbb{U}$, and some $\alpha \in]0, 1]$, the stage cost satisfies the inequality

$$c(z, u) \geq \alpha c_F(z). \quad (14)$$

The last assumption, given below, basically puts a bound on the growth of the value function which is incurred by applying the policy (11) due to unavailability of state measurements over time intervals of random length.

(H6) For each $\ell \in \mathbb{N}$ and each $k \in \{0, \dots, \nu - 1\}$, there exist scalars $C_{1,k}, C_{2,k} > 0$ such that,

$$\mathbb{E}^{\widehat{u}_{\tau_\ell: \tau_\ell+k}} \left[\sum_{t=\tau_\ell}^{\tau_\ell+k} \tilde{c}(\pi_t^*(x_t) - \widehat{u}_t) \middle| x(\tau_\ell) \right] \leq C_{1,k} + C_{2,k} V_\nu^*(x(\tau_\ell)). \quad (15)$$

Remark III.1. In case of linear (possibly time-varying) dynamics, and quadratic costs, it will be shown in Section IV that $C_{1,k}$ depends on variance of the noise in plant dynamics and is linear in k , while $C_{2,k} = 0$. However, in general, for nonlinear dynamics $C_{2,k}$ is not necessarily zero.

We now use these assumptions to state our main result, which uses the following notation:

$$\begin{aligned} \overline{C}_1 &:= \mathbb{E}[C_{1, \tau_1 - \tau_0}], & \overline{d} &:= \overline{C}_1 + \alpha \nu b + \mathbb{E}[\tau_1 - \tau_0], \\ \overline{C}_2 &:= \mathbb{E}[C_{2, \tau_1 - \tau_0}], & \overline{\gamma} &:= (1 + \overline{C}_2 - \alpha). \end{aligned} \quad (16)$$

Note that due to i.i.d. assumption on the sampling process in **(H2)**, for each $\ell \in \mathbb{N}$, $\overline{C}_i = \mathbb{E}[C_{i, \tau_{\ell+1} - \tau_\ell}]$, $i = 1, 2$. For brevity, we let $C_{i, \ell} := C_{i, \tau_{\ell+1} - \tau_\ell}$.

Theorem III.2. Consider the dynamical system (1) subject to the sampling process T_t , and the control input u_t given in (11). If assumptions **(H1)**–**(H6)** hold, and moreover,

$$\overline{\gamma} < 1,$$

then the infinite-horizon average cost (2) is bounded, and satisfies

$$\limsup_{k \rightarrow \infty} \frac{1}{k} \mathbb{E} \left[\sum_{t=0}^{k-1} c(x_t, u_t) \middle| x_0 \right] \leq b + \overline{C}_1 + \overline{d} \overline{C}_2.$$

To derive this result, several intermediate steps are required which we describe next.

A. Intermediate Results

The first statement that we need is a direct consequence of our certainty equivalence based hypothesis **(H3)**. In what follows, we use the notation \sharp for concatenation of policies, so that we can write $u_{k_1:k_2} = u_{k_1:\overline{k}} \sharp u_{\overline{k}+1:k_2}$ for any $k_1 \leq \overline{k} < k_2$.

Lemma III.3. Under the hypotheses **(H1)**, **(H2)** and **(H3)**, for each sampling instant $\tau_\ell \in \mathbb{N}$, and each $k \in \{0, \dots, \nu - 1\}$, we have

$$\begin{aligned} \mathbb{E}^{\widehat{u}_{0:k} \sharp \pi_{k+1:\nu-1}} \left[\sum_{t=\tau_\ell}^{\tau_\ell+\nu-1} c(x_t, u_t) + c_F(x_{\tau_\ell+\nu}) \middle| x(\tau_\ell) \right] \leq \\ \mathbb{E}^{\pi_{0:\nu-1}^*} \left[\sum_{t=\tau_\ell}^{\tau_\ell+\nu-1} c(x_t, u_t) + c_F(x_{\tau_\ell+\nu}) \middle| x(\tau_\ell) \right] \\ + \mathbb{E}^{\widehat{u}_{0:k}} \left[\sum_{j=0}^k \tilde{c}(\pi_t^*(x_{j+\tau_\ell}) - \widehat{u}_j(x_{\tau_\ell})) \middle| x(\tau_\ell) \right] \end{aligned} \quad (17)$$

where $\tilde{c} : \mathbb{R}^d \rightarrow \mathbb{R}$ is the positive definite function introduced in **(H3)**.

The next statement provides an upper bound on the value function V^* at a sampling instant τ_ℓ in terms of the last successfully received measurement at time $\tau_{\ell-1}$, and makes use of the inequality (17) given in the preceding lemma.

Lemma III.4. Under the hypotheses **(H1)**, **(H2)**, **(H3)**, and **(H4)**, the following inequality holds for each sampling instant $\tau_\ell \in \mathbb{N}$,

$$\begin{aligned} \mathbb{E}^{\widehat{u}_{0:\tau_{\ell+1}-\tau_\ell-1}} \left[V_\nu^*(x(\tau_{\ell+1})) \middle| x(\tau_\ell) \right] \leq V_\nu^*(x(\tau_\ell)) \\ - \mathbb{E}^{\widehat{u}_{0:\tau_{\ell+1}-\tau_\ell-1}} \left[\sum_{t=\tau_\ell}^{\tau_{\ell+1}-1} c(x_t, u_t) \middle| x(\tau_\ell) \right] + b(\tau_{\ell+1} - \tau_\ell) \\ + \mathbb{E}^{\widehat{u}_{0:\tau_{\ell+1}-\tau_\ell-1}} \left[\sum_{j=0}^{\tau_{\ell+1}-\tau_\ell-1} \tilde{c}(\pi_j^*(x_{j+\tau_\ell}) - \widehat{u}_j(x_{\tau_\ell})) \middle| x(\tau_\ell) \right]. \end{aligned} \quad (18)$$

The last statement that we need for our main result is a stability-like estimate which provides a uniform bound on the value function for a particular realization of the sampling instants. It builds on the inequality (18) given in Lemma III.4 and additionally uses **(H5)** and **(H6)**.

Proposition III.5. Consider dynamical system (1) with the input given by (11), and assume conditions **(H1)**–**(H6)**. For each sampling instant $\tau_\ell \in \mathbb{N}$, it holds that

$$\mathbb{E}^{\widehat{u}_{\tau_\ell:\tau_{\ell+1}-1}} \left[V_\nu^*(x(\tau_{\ell+1})) \middle| x(\tau_\ell) \right] \leq (1 - \alpha + C_{2,\ell}) V_\nu^*(x(\tau_\ell)) + C_{1,\ell} + b(\tau_{\ell+1} - \tau_\ell) + \alpha \nu b. \quad (19)$$

B. Proof of Theorem III.2

We now use the inequalities provided in Lemma III.4 and Proposition III.5 to derive the bound stated in Theorem III.2.

Proof. Using the inequality (18), for each $\ell \in \mathbb{N}$, and employing the notation $\Delta_\ell = \tau_{\ell+1} - \tau_\ell$, we obtain

$$\begin{aligned} \mathbb{E}^{\widehat{u}_{0:\Delta_\ell-1}} \left[\sum_{t=\tau_\ell}^{\tau_{\ell+1}-1} c(x_t, u_t) \middle| x(\tau_\ell) \right] &\leq V_\nu^*(x(\tau_\ell)) \\ &\quad - \mathbb{E}^{\widehat{u}_{0:\Delta_\ell-1}} \left[V_\nu^*(x(\tau_{\ell+1})) \middle| x(\tau_\ell) \right] \\ &\quad + \mathbb{E}^{\widehat{u}_{0:\Delta_\ell-1}} \left[\sum_{j=0}^{\Delta_\ell-1} \tilde{c}(\pi_j^*(x_j) - \widehat{u}_j) \middle| x(\tau_\ell) \right] + b\Delta_\ell. \end{aligned}$$

Moreover, by invoking hypothesis (H6),

$$\begin{aligned} \mathbb{E}^{\widehat{u}_{0:\Delta_\ell-1}} \left[\sum_{t=\tau_\ell}^{\tau_{\ell+1}-1} c(x_t, u_t) \middle| x(\tau_\ell) \right] &\leq V_\nu^*(x(\tau_\ell)) \\ &\quad - \mathbb{E}^{\widehat{u}_{0:\Delta_\ell-1}} \left[V_\nu^*(x(\tau_{\ell+1})) \middle| x(\tau_\ell) \right] \\ &\quad + C_{2,\ell} V_\nu^*(x(\tau_\ell)) + C_{1,\ell} + b\Delta_\ell. \end{aligned}$$

Adding these intersample bounds for first k samples, and letting \widehat{u}^k denote the concatenation of the policies given in (10) from time $t = 0$ to $\tau_k - 1$, that is,

$$\widehat{u}^k = (\widehat{u}_{0:\tau_1-1}, \widehat{u}_{0:\tau_2-1}, \dots, \widehat{u}_{0:\tau_k-1}),$$

we obtain

$$\begin{aligned} \mathbb{E}^{\widehat{u}^k} \left[\sum_{t=0}^{\tau_k-1} c(x_t, u_t) \middle| x(\tau_0) \right] &\leq V_\nu^*(x(\tau_0)) - V_\nu^*(x(\tau_k)) \\ &\quad + b(\tau_k - \tau_0) + \sum_{\ell=0}^{k-1} C_{2,\ell} V_\nu^*(x(\tau_\ell)) + C_{1,\ell}. \quad (20) \end{aligned}$$

So far, we have worked with fixed values of $\{\tau_\ell\}_{\ell \in \mathbb{N}}$ along one sample path. We now compute expectation with respect to the random variables associated with the inter-sampling times. Before doing so, we need to get a bound on the term $\sum_{\ell=0}^k C_{2,\ell} V_\nu^*(x(\tau_\ell))$; and for that, let $\gamma_\ell := (1 - \alpha + C_{2,\ell})$ and $d_\ell := C_{1,\ell} + b(\tau_{\ell+1} - \tau_\ell) + \alpha\nu b$. It was observed in (19) that

$$V_\nu^*(x(\tau_\ell)) \leq \prod_{i=1}^{\ell} \gamma_i V_\nu^*(x(\tau_0)) + \sum_{i=1}^{\ell-1} \prod_{j=i+1}^{\ell} \gamma_j d_i + d_\ell.$$

This immediately leads to

$$\begin{aligned} \sum_{\ell=0}^k C_{2,\ell} V_\nu^*(x(\tau_\ell)) &\leq V_\nu^*(x(\tau_0)) \sum_{\ell=1}^k C_{2,\ell} \prod_{i=1}^{\ell} \gamma_i \\ &\quad + d_1 \left(C_{2,1} + \sum_{\ell=2}^k C_{2,\ell} \prod_{i=2}^{\ell} \gamma_i \right) \\ &\quad + d_2 \left(C_{2,2} + \sum_{\ell=3}^k C_{2,\ell} \prod_{i=3}^{\ell} \gamma_i \right) + \dots \end{aligned}$$

$$\dots + d_{k-1} (C_{2,k-1} + C_{2,k} \gamma_k) + C_{2,k} d_k.$$

Now, let $\bar{\gamma} := \mathbb{E}[1 - \alpha + C_{2,\ell}]$, and $\bar{d} := \mathbb{E}[d_\ell]$, which are uniform with respect to $\ell \in \mathbb{N}$ due to i.i.d. assumption on the sampling process. Computing the expectation (with respect to sampling times) of each term on the right-hand side of (20), we get

$$\begin{aligned} \frac{1}{k} \mathbb{E}_{\{\tau_\ell\}} \left[\sum_{\ell=0}^k C_{2,\ell} V_\nu^*(x(\tau_\ell)) \right] &\leq V_\nu^*(x(\tau_0)) \frac{1}{k} \bar{C}_2 \sum_{\ell=1}^k \bar{\gamma}^\ell \\ &\quad + \bar{d} \bar{C}_2 + \frac{1}{k} \bar{d} \bar{C}_2 \left(\sum_{\ell=2}^k \bar{\gamma}^\ell \right) + \frac{1}{k} \bar{d} \bar{C}_2 \left(\sum_{\ell=3}^k \bar{\gamma}^\ell \right) + \dots + \frac{1}{k} \bar{d} \bar{C}_2. \end{aligned}$$

Taking the limit as $k \rightarrow \infty$ we see that the right-hand side is bounded by the constant $\bar{d} \bar{C}_2$ as all the remaining terms converge to zero, so that

$$\limsup_{k \rightarrow \infty} \frac{1}{k} \mathbb{E}_{\{\tau_\ell\}} \left[\sum_{\ell=0}^k C_{2,\ell} V_\nu^*(x(\tau_\ell)) \right] \leq \bar{d} \bar{C}_2.$$

Coming back to (20), we now take expectation with respect to random sampling intervals, divide the resulting expression on both sides by k (note that $k \geq \tau_k$), and take the limit as $k \rightarrow \infty$, to get

$$\limsup_{k \rightarrow \infty} \frac{1}{k} \mathbb{E}_{\{\tau_\ell\}} \left[\mathbb{E}^{\widehat{u}^k} \left[\sum_{t=0}^{\tau_k-1} c(x_t, u_t) \middle| x(\tau_0) \right] \right] \leq b + \bar{C}_1 + \bar{d} \bar{C}_2 \quad (21)$$

and hence the desired assertion holds. \square

IV. LINEAR QUADRATIC CASE

We consider the special case where the dynamics are linear time-varying and described by³

$$x_{t+1} = A_t x_t + B_t u_t + w_t, \quad (22)$$

where the noise process $\{w_t\}_{t=0}^\infty$ is a zero mean Gaussian, and satisfies the i.i.d. assumption. The cost function we consider in this case is quadratic:

$$J_\infty^{\text{ave}}(x_0, u) = \limsup_{k \rightarrow \infty} \frac{1}{k} \mathbb{E} \left[\sum_{t=0}^{\infty} x_t^\top Q_t x_t + u_t^\top R_t u_t \middle| x_0 \right].$$

During a particular realization, and given the sampling time τ_{N_t} , the finite-horizon cost functional which we consider for computing the control policy, is defined as

$$\begin{aligned} J_{\nu, \tau_{N_t}}(x(\tau_{N_t}), u) &:= \\ &\mathbb{E} \left[\sum_{s=\tau_{N_t}}^{\tau_{N_t}+\nu-1} x_s^\top Q_s x_s + u_s^\top R_s u_s \right. \\ &\quad \left. + x_{\tau_{N_t}+\nu}^\top Q_{\tau_{N_t}+\nu} x_{\tau_{N_t}+\nu} \middle| x(\tau_{N_t}) \right]. \end{aligned}$$

It is assumed that the matrices $\{A_i, B_i, Q_i, R_i\}_{i \in \mathbb{N}}$ belong to a compact set with (A_i, B_i) controllable, and Q_i, R_i

³In contrast to model given in (1), the system (22) is time-varying, but our results carry over to this case with obvious modifications.

positive definite, for each $i \in \mathbb{N}$. Following the development carried out in [5, Chapters 3 & 4], it can be shown that the control policy given in (10) results in

$$\hat{u}_t = L_{t-\tau_{N_t}} \mathbf{E}[x_t | x_{\tau_{N_t}}]$$

for some appropriately chosen matrices $L_{t-\tau_{N_t}}$. Thus, the computed policy is a linear feedback that uses the estimate of the current state conditioned upon the last received measurement. To give an expression for the gains $L_{t-\tau_{N_t}}$, fix $N_t = 0$ and let $\tau_0 = 0$ for the sake of simplicity. In this case, for $0 \leq k \leq \nu - 1$,

$$L_k = -(R_k + B_k^\top K_{k+1} B_k)^{-1} B_k^\top K_{k+1} A_k$$

by setting $K_\nu = Q_\nu$, and for $0 \leq k < \nu$, we let

$$P_k = A_k^\top K_{k+1} B_k (R_k + B_k^\top K_{k+1} B_k)^{-1} B_k^\top K_{k+1} A_k \\ K_k = A_k^\top K_{k+1} A_k - P_k + Q_k.$$

We next show that with this choice of gains, the chosen policy indeed satisfies all the hypotheses. In the sequel, we only check **(H3)** and **(H6)**. To see that **(H3)** holds, it can be verified, using the dynamic programming principle, that

$$V_{\nu-k}^*(x_k) = \mathbf{E}[x_{k+1}^\top K_{k+1} x_{k+1}] + \mathbf{E}[w_{k+1}^\top Q_k w_{k+1}].$$

Moreover,

$$\hat{V}_{\nu-k}(x_0) \\ = \min_{u_k} \mathbf{E} \left[x_k^\top Q_k x_k + u_k^\top R_k u_k + V_{\nu-k+1}^*(x_{k+1}) \mid x_0 \right] \\ = \mathbf{E} \left[x_k^\top Q_k x_k \mid x_0 \right] + \min_{u_k} \mathbf{E} \left[u_k^\top R_k u_k + x_{k+1}^\top K_{k+1} x_{k+1} \mid x_0 \right] \\ + \mathbf{E} \left[\sum_{j=k+1}^{\nu-1} w_j^\top Q_j w_j \right].$$

The control policy $\hat{u}_k = L_k \mathbf{E}[x_k | x_0]$ is indeed a minimizer that yields $\hat{V}_{\nu-k}(x_0)$. Let $\hat{x}_k := \mathbf{E}[x_k | x_0]$. Observing that $x_k - \hat{x}_k$ is independent of the policy used (because of the linearity structure), we get

$$\hat{V}_{\nu-k}(x_0) = \mathbf{E}[x_k^\top K_k x_k | x_0] + \mathbf{E} \left[\sum_{j=k}^{\nu-1} w_j^\top Q_j w_j \right] \\ + \mathbf{E}[(x_k - \hat{x}_k)^\top P_k (x_k - \hat{x}_k) | x_0].$$

Hence, recalling the definition of P_k , the hypothesis **(H3)** is seen to hold with

$$\tilde{c}(v) = \max_{0 \leq k \leq \nu-1} v^\top (R_k + B_k^\top K_{k+1} B_k) v.$$

To see that **(H6)** holds, for each $k \in \{0, \dots, \nu - 1\}$, let $C_{1,k} = kc \mathbf{E}[w_0^2]$ for some $c > 0$ large enough. It readily follows that

$$\mathbf{E}^{\hat{u}_{0:k-1}} \left[\sum_{t=0}^{k-1} \tilde{c}(L_t(x_t - \hat{x}_t)) \mid x_0 \right] \leq C_{1,k}, \quad (23)$$

so that (15) is satisfied with $C_{2,k} = 0$.

REFERENCES

- [1] M. Adès, P. E. Caines, and R. P. Malhamé. Stochastic optimal control under Poisson-distributed observations. *IEEE Transactions on Automatic Control*, 45(1):3–13, 2000.
- [2] D. J. Antunes, J. P. Hespanha, and C. J. Silvestre. Volterra integral approach to impulsive renewal systems: Application to networked control. *IEEE Transactions on Automatic Control*, 57(3):607 – 619, 2012.
- [3] A. Arapostathis, V.S. Borkar, E. Fernández-Gaucherand, M.K. Ghosh, and S.I. Marcus. Discrete-time controlled Markov processes with average cost criterion: A survey. *SIAM J. Control & Optimization*, 31(2):282–344, 1993.
- [4] D. Bernardini and A. Bemporad. Stabilizing model predictive control of stochastic constrained linear systems. *IEEE Transactions on Automatic Control*, 57:1468 – 1480, 2012.
- [5] D. Bertsekas. *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, 4th edition, 2017.
- [6] D. Bertsimas and D.B. Brown. Constrained stochastic LQC: A tractable approach. *IEEE Transactions on Automatic Control*, 52:1826 – 1841, 2007.
- [7] D. Chatterjee and J. Lygeros. On stability and performance of stochastic predictive control techniques. *IEEE Transactions on Automatic Control*, 60(2):509–514, 2015.
- [8] L. Grüne. Approximation properties of receding horizon optimal control. *Jahresbericht der Deutschen Mathematiker Vereinigung*, 118:3–37, 2016.
- [9] L. Grüne and J. Pannek. *Nonlinear Model Predictive Control: Theory and Algorithms*. Springer-Verlag, London, 2nd edition, 2017.
- [10] L. Grüne, J. Pannek, and K. Worthmann. A networked unconstrained nonlinear MPC. In *Proc. European Control Conf.*, pages 371–376, 2009.
- [11] O. Hernández-Lerma and J.-B. Lasserre. *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, volume 30. Springer-Verlag, New York, USA, 1996.
- [12] J. P. Hespanha and A. R. Teel. Stochastic impulsive systems driven by renewal processes. In *Proc. 17th International Symposium on Mathematical Theory of Networked Systems*, pages 606–618, 2006.
- [13] D.Q. Mayne and J. Rawlings. *Model Predictive Control: Theory and Applications*. Nob Hill Publishing LLC, Madison, WI, USA, 2009.
- [14] D.Q. Mayne, J.B. Rawlings, C.V. Rao, and P.O.M. Scolaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36:789–814, 2000.
- [15] A. Mesbah. Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Systems Magazine*, 36(6):30–44, 2016.
- [16] P.K. Mishra, D. Chatterjee, and D. Quevedo. Stabilizing stochastic predictive control under Bernoulli dropouts. *IEEE Transactions on Automatic Control*, 63(6):1579 – 1590, 2018.
- [17] D. Muñoz de la Peña and P.D. Christofides. Stability of nonlinear asynchronous systems. *Systems & Control Letters*, 57:465 – 473, 2008.
- [18] D. Quevedo and D. Nešić. Robust stability of packetized predictive control of nonlinear systems with disturbances and Markovian packet losses. *Automatica*, 48:1803–1811, 2012.
- [19] A. Tanwani, D. Chatterjee, and D. Liberzon. Stabilization of deterministic control systems under random sampling: Overview and recent developments. In T. Başar, editor, *Uncertainty in Complex Networked Systems*, Systems & Control: Foundations & Applications, pages 209–246. SpringerNature, Switzerland, 2018.
- [20] A. Tanwani and A. R. Teel. Stabilization with event-driven controllers over a digital communication channel with random transmissions. In *Proc. 56th IEEE Conf. Decision and Control*, pages 6063–6068, 2017.
- [21] W. Wu and A. Arapostathis. Optimal sensor querying: General Markovian and LQG models with controlled observations. *IEEE Transactions on Automatic Control*, 53(6):1392–1405, 2008.
- [22] L. Xie & L. Xie. Stability analysis of networked sampled-data linear systems with Markovian packet losses. *IEEE Transactions on Automatic Control*, 54(6):1375 – 1381, 2009.