

STABILITY AND PERFORMANCE OF STOCHASTIC PREDICTIVE CONTROL

DEBASISH CHATTERJEE AND JOHN LYGEROS

ABSTRACT. This article is concerned with stability and performance of controlled stochastic processes under receding horizon policies. We carry out a systematic study of methods to guarantee stability under receding horizon policies via appropriate selections of cost functions in the underlying finite-horizon optimal control problem. We also obtain quantitative bounds on the performance of the system under receding horizon policies as measured by the long-run expected average cost. The results are illustrated with the help of several simple examples.

1. INTRODUCTION

With the steady growth in the availability of fast computing machines, control techniques that involve algorithmic selection of actions that minimize some performance objective have gained prominence. Receding horizon predictive control, which is based on such algorithmic selection procedures, has evolved over the years into one of the most useful and applicable control synthesis techniques currently available to a control engineer; see e.g., [11] for a survey of the modern theory and applications in the deterministic setting. Stochastic versions of receding horizon techniques initially evolved within the operations research community, see e.g., [6, 7], with inventory and manufacturing systems as primary application areas, and have steadily filtered into the domain of control systems, with current applications in financial engineering, process control, industrial electronics, power systems, etc.

While the deterministic and robust versions of receding horizon control techniques have become standardized and are well-documented, the available literature on the stochastic version still lacks a comprehensive and systematic treatment. Especially prominent in this regard is the matter of stability of control systems under stochastic receding horizon control; indeed, most of the literature does not appear to take advantage of the significantly developed and advanced results on stability of Markov processes. Chief among the reasons for this discrepancy between the deterministic and stochastic settings, perhaps, is the fact that the technical nature of the arguments involved in the stochastic version of stability is significantly heavier than its deterministic counterpart. Indeed, while the bare-essential arguments involved in establishing Lyapunov stability of discrete-time deterministic dynamical systems are only a few and are quite classical, the technical arguments and conditions in the theory of stability of Markov chains is by far larger in number, and constitute an active area of research even today. In addition to that, one has a diverse library

Date: 19 April 2013.

This work was carried out during D. Chatterjee's visit to the Automatic Control Laboratory, ETH Zürich, in June-August 2012, supported by the Swiss National Science Foundation via the Short International Visit Fellowship no. IZKOZ2_142780.

of notions of stability that are peculiar to the stochastic setting, and are simply non-existent in the deterministic or the robust setting.

This article is an attempt at bridging this gap—we connect receding horizon control techniques to some of the principal elements of the theory of stability of Markov processes. Motivated by, and in the spirit of [12, §3], first we systematically develop a framework for studying stability of discrete-time controlled stochastic systems under receding horizon policies. We critically examine two approaches in this connection, namely, ensuring stability by appropriate selection of the cost functions, and by adjoining an appropriate constraint to the underlying finite-horizon optimal control problem, before focussing on the former. Against the backdrop of certain standard (and no-so-standard) conditions for stability of Markov processes, we establish conditions on the cost functions such that these stability conditions are satisfied. Thus, this selection procedure, by design, ensures that the closed-loop system under the corresponding receding horizon control policy is stable. We utilize theorems on stability of Markov processes off-the-shelf as to this end. As such, the results pertaining to stability presented here should be regarded as representative guidelines—rather than offer a set of stand-alone results, we provide a general framework for establishing stability results. The details for specific applications must be worked out on a case-by-case basis, as we illustrate through several examples.

In addition, we develop a framework for analyzing the performance of the closed-loop systems under stochastic receding horizon control policies. Selecting a long-run expected average cost derived from the underlying finite-horizon optimal control problem as our performance index, we provide quantitative bounds on this performance index under receding horizon policies and mild hypotheses. Observe that receding horizon policies are extracted from a finite-horizon optimal control problem, and as such do not naturally offer any clue concerning the long-run expected average costs that they incur. The relationship between stability and performance is also explored here. In particular, we obtain a bound on the aforementioned performance index under a receding horizon policy that also ensures stability in an appropriate sense.

The layout of this article is as follows: §2 provides the description of the control systems. Our results on stability under receding horizon control are contained in §3, while performance bounds are provided in §4. Several examples illustrate our results throughout §3 and §4. The proofs of our results are provided in the Appendices §A and §B. The emphasis here is on conceptual clarity and a systematic presentation sans heuristics. The setting, insofar as the system, the associated receding horizon problem, and the results are concerned, is at an abstract level; this choice is targeted at conveying the key ideas in a transparently clear fashion, without the overload of excessive notation. In particular, the ideas presented here can be readily generalized to the setting of Markov decision processes; we choose to stay with simpler notation and technical requirements here. Numerical tractability of the underlying optimal control problems, which is an integral aspect of receding horizon control techniques, is *not* addressed here.

2. SYSTEM DESCRIPTION

Consider the discrete-time dynamical system given by the recursion

$$(2.1) \quad x_{t+1} = f(x_t, u_t, w_t), \quad x_0 \text{ given, } t \in \mathbb{N}_0,$$

where

- $x_t \in \mathbb{R}^d$, $u_t \in \mathbb{U} \subset \mathbb{R}^m$, and $w_t \in \mathbb{W} \subset \mathbb{R}^p$ are the states, the control actions, and the noise at time t ;
- $f : \mathbb{R}^d \times \mathbb{U} \times \mathbb{W} \rightarrow \mathbb{R}^d$ is a measurable function;¹
- \mathbb{U} is the (nonempty) *control set*, assumed to be measurable and containing the element $0 \in \mathbb{R}^m$;
- $\mathbb{W} \subset \mathbb{R}^p$ is assumed to be a measurable set;
- $(w_t)_{t \in \mathbb{N}_0}$ is the *process noise*—a \mathbb{W} -valued random process with the w_t 's independent and identically distributed.²

Let k be a positive integer. Recall that a *k-stage (feedback) policy* is a collection $\pi_{0:k-1} := (\pi_0, \pi_1, \dots, \pi_{k-1})$ of measurable functions $\pi_i : \mathbb{R}^d \rightarrow \mathbb{U}$ for each i ; we set the t -th control action as $u_t = \pi_t(x_t)$. For the synthesis of control actions in a receding horizon fashion, we consider given:

- a *horizon* $N \in \mathbb{N}$,
- a *cost-per-stage function* $c : \mathbb{R}^d \times \mathbb{U} \rightarrow [0, +\infty[$ and a *final cost function* $c_F : \mathbb{R}^d \rightarrow [0, +\infty[$, both assumed to be measurable, and
- a class Π of feedback policies.

We introduce the N -horizon value function

$$(2.2) \quad V_N(x, \pi) := \mathbb{E}_x^\pi \left[\sum_{i=0}^{N-1} c(x_i, u_i) + c_F(x_N) \right]$$

where the policies π belong to the class Π .³

ASSUMPTION. Without detailing the specifics,

- (A1) we assume sufficient regularity of the process $(w_t)_{t \in \mathbb{N}_0}$ such that the cost (2.2) is finite for all $x \in \mathbb{R}^d$ and all $\pi \in \Pi$. \diamond

With these ingredients, the centerpiece of receding horizon control can be stated: It consists of the N -horizon optimal control problem:

$$(2.3) \quad \begin{aligned} & \text{minimize} && V_N(x, \pi) \\ & \text{subject to} && \begin{cases} \pi \in \Pi, \\ \text{dynamics (2.1)}. \end{cases} \end{aligned}$$

ASSUMPTION. In addition to (A1), we assume that

- (A2) the minimization problem (2.3) is well-defined for all $x \in \mathbb{R}^d$, i.e., for each boundary value $x \in \mathbb{R}^d$, there exists a policy $\pi^* \in \Pi$ that solves (2.3). \diamond

Conditions under which (A2) holds are of a technical nature, and well documented, e.g., in [6, Chapter 3].

¹Henceforth “measurability” on Euclidean spaces will refer to “Borel measurability.”

²All random vectors are assumed to be defined on some underlying probability space, for which $\mathbb{P}(\cdot)$ is the probability measure, and $\mathbb{E}[\cdot]$ is the expectation under \mathbb{P} . The assumption that w_t 's are independent and identically distributed can be substantially weakened at the expense of some notational clutter; we choose to stay with the simpler setting here.

³If $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}$ is a measurable function, then $\mathbb{E}_x^\pi[\varphi(x_t)]$ stands for the conditional expectation of $\varphi(x_t)$ given $x_0 = x$, where x_t is the state at time t under the policy π ; \mathbb{P}_x is the corresponding conditional probability.

We denote the optimal value function $V_N(x, \pi^*)$ by $V_N^*(x)$ for all $x \in \mathbb{R}^d$. The system (2.1) under the optimal policy generates the optimal state trajectory $(x_t^*)_{t=0}^N$ given by

$$(2.4) \quad x_{t+1}^* = f(x_t^*, \pi_t^*(x_t^*), w_t), \quad x_0^* \text{ given}, \quad t = 0, 1, \dots, N-1.$$

The technique of receding horizon control consists of applying the first element π_0^* obtained from the minimization problem (2.3) recursively, thereby generating the *receding horizon policy*

$$(2.5) \quad \hat{\pi} := (\pi_0^*, \pi_0^*, \dots).$$

To wit, given the state x_t at time t , one solves the minimization problem (2.3) with $x = x_t$, obtains the optimal policy π^* , applies the first element π_0^* of the policy, moves to time $t+1$, and repeats the preceding steps. The system (2.1) under the policy $\hat{\pi}$ generates the state trajectory $(x_t)_{t \in \mathbb{N}_0}$ via the recursion

$$(2.6) \quad x_{t+1} = f(x_t, \pi_0^*(x_t), w_t), \quad x_0 \text{ given}, \quad t \in \mathbb{N}_0.$$

Observe that the process $(x_t)_{t \in \mathbb{N}_0}$ generated by (2.6) is *Markovian*, i.e., the probability distribution of the future state x_{t+1} at time $t+1$ is conditionally independent of the past $(x_s)_{s=0}^{t-1}$ given the present state x_t . Indeed, for S a Borel subset of \mathbb{R}^d , we have

$$(2.7) \quad \begin{aligned} \mathbb{P}(x_{t+1} \in S \mid (x_s)_{s=0}^t) &= \mathbb{P}(f(x_t, \pi_0^*(x_t), w_t) \in S \mid (x_s)_{s=0}^t) \\ &= \mathbb{P}(f(x_t, \pi_0^*(x_t), w_t) \in S \mid x_t). \end{aligned}$$

The following sections will study aspects of both qualitative and quantitative behavior of the process $(x_t)_{t \in \mathbb{N}_0}$ generated by (2.6).

3. STABILITY UNDER RECEDING HORIZON CONTROL

Stability of the *controlled process* $(x_t)_{t \in \mathbb{N}_0}$ generated by the recursion (2.6) is a desirable property in practice. There are two techniques in which stability can be ensured:

- (S1) **By appropriate choice of cost functions:** Stability of the controlled process $(x_t)_{t \in \mathbb{N}_0}$ can be ensured by an appropriate selection of the cost-per-stage function c and the cost function c_F . In the deterministic setting, conditions for asymptotic stability in terms of the cost functions are standard, see e.g., [12, §3] and [10] for details and further references. In the stochastic setting the aim is to arrive at a Lyapunov-like inequality in terms of the cost functions, which in turns ensures stability of the closed-loop system. While conceptually this technique leads to an elegant analysis, there are two points worthy of note:
- For a control engineer, the selection of the cost functions c and c_F is typically dictated by the physics of the problem. In case the stability conditions are not satisfied by the natural candidates c and c_F , the engineer may be forced to select cost functions that may have little to do with the particular physical aspects of the plant.
 - The applicability of predictive control is contingent upon numerical tractability of the finite horizon optimal control problem (2.3). The extent of flexibility in the choice of the functions c and c_F is determined, therefore, by cases where numerically tractable problems can be derived from (2.3). In other words, applicability of this technique is limited by numerical tractability of the problem (2.3). However, as we shall illustrate through examples, more than one cost functions that ensure stability; there is, therefore, some freedom which the control engineer can utilize to suit numerical tractability.

In the deterministic setting, standard stability conditions under receding horizon control require the existence of a stabilizing feedback controller *inside* a certain terminal set, such that the terminal set becomes invariant for the controlled system [12, §3]. In the stochastic setting, in general, even the weakest forms of stability (for instance, positive recurrence, existence of invariant measures, etc.) require the existence of a certain “drift condition” *outside* a bounded set [13]. This difference between the two settings is generally *unavoidable* for the lack of a notion analogous to invariance in the deterministic setting; see also Remark 7 for a more specific discussion. For instance, in systems where there is non-zero probability of jumps in the state infinitely often and the magnitude of the jumps is not bounded, the notion of deterministic invariance does not make sense for any bounded subset of the state-space. Of course, this assertion does not apply to systems subjected to bounded noise where it may be possible to perform a robust analysis, but it does indeed apply to the standard benchmark case of a linear control system with additive and independent Gaussian noise.

- (S2) By adjoining an appropriate constraint to the optimal control problem (2.3): This technique was first adopted in [1, 8, 9] in the context of receding horizon control of linear stochastic controlled systems. It consists of adjoining a constraint to the optimal control problem (2.3), so that the modified optimal control problem stays feasible for all $x \in \mathbb{R}^d$, and the resulting receding horizon policy $\hat{\pi}$ defined in (2.5) ensures stability. Observe that (i) the problem (2.3) where one intends to adjoin the constraint is limited to a finite-horizon, while (ii) the target of the constraint—attaining stability of the closed-loop system—involves a necessarily infinite-horizon notion. Two points to note:
- It is imperative to ensure that the problem (2.3) with the new constraint is feasible for all boundary values x ; this necessarily imposes restrictions on the type of admissible constraints.
 - Adjoining a constraint to the problem (2.3) potentially shifts both the optimal value and the optimizer π^* corresponding to the original problem (2.3). Therefore, a trade-off between the performance and a certain desirable qualitative behavior of the closed-loop system may have to be accepted.

As will be evident from the above discussion, a systematic development of the case (S2) is largely impossible due to the absence of a set of unifying objects inherent to the optimal control problem (2.3). Since the constraints do not, generally, depend on the cost functions, the details of the technique may differ significantly between specific applications

In this article we focus on (S1); the relevant results are presented in §3.2. Preparatory to that, in §3.1 we briefly recall certain basic aspects of the general theory of stability of discrete-time Markov processes.

3.1. Review of the general theory of stability of discrete-time Markov processes. The type of stability that we shall focus on here concerns boundedness of sequences of the form $(\mathbb{E}_x[h(x_t)])_{t \in \mathbb{N}_0}$, for appropriate functions $h : \mathbb{R}^d \rightarrow [0, +\infty[$. For instance, consider $h(z) = \|z\|^p$ for $p > 1$. In view of the fact that⁴

$$\mathbb{E}_x[\|x_t\|^p] = p \int_0^{+\infty} r^{p-1} \mathbb{P}_x(\|x_t\| > r) dr,$$

boundedness of $(\mathbb{E}_x[\|x_t\|^p])_{t \in \mathbb{N}_0}$ implies that the conditional probability, given the initial condition $x_0 = x$, of the states being at a distance r from the origin decays

⁴This identity is an immediate consequence of Fubini’s theorem.

faster than r^{-p} as r grows large, uniformly over time t . In other words, we have an assertion corresponding to the behavior of the tail of the conditional probability distributions $\mathbb{P}_x(\|x_t\| > r), t \in \mathbb{N}_0$, uniformly over time t . The case of $p = 2$ is especially prevalent in the literature, and goes under the name of mean-square boundedness. To understand the qualitative behavior of \mathbb{R}^d -valued Markov processes $(x_t)_{t \in \mathbb{N}_0}$, the general strategy consists of studying the behavior of sequences such as $(\mathbb{E}_x[h(x_t)])_{t \in \mathbb{N}_0}$ for “norm-like” functions h , and drawing appropriate inferences concerning the former.

Recall that the process $(x_t)_{t \in \mathbb{N}_0}$ generated by (2.6) is Markovian in view of (2.7). For discrete-time Markov processes, the theory of stability is extremely well-developed, (see e.g., [13] for a book-length treatment,) and most of the standard conditions for stability involve what is known as a “negative drift condition.”⁵ A generic negative drift condition takes the following form:

(D) there exist measurable functions $\Xi : \mathbb{R}^d \rightarrow [0, +\infty[$ and $\Upsilon : \mathbb{R}^d \rightarrow [0, +\infty[$, and a bounded and measurable set $K \subset \mathbb{R}^d$, such that

$$\mathbb{E}_x[\Xi(x_1)] - \Xi(x) \leq -\Upsilon(x) \quad \text{for all } x \notin K.$$

Depending on the properties of the functions Ξ and Υ , it may be possible to assert the type stability of the process $(x_t)_{t \in \mathbb{N}_0}$. Observe that the condition (D) closely resembles Lyapunov stability conditions for deterministic discrete-time systems.

Perhaps the most well-known drift condition is contained in the following:

Proposition 1. *Let $(x_t)_{t \in \mathbb{N}_0}$ be a Markov process. Suppose that there exist $\beta > 0$ and $\lambda_\circ \in [0, 1[$, a measurable function $V : \mathbb{R}^d \rightarrow [0, +\infty[$, and a compact set $K \subset \mathbb{R}^d$ such that $\mathbb{E}_x[V(x_1)] \leq \lambda_\circ V(x)$ for all $x \notin K$, and $\sup_{x \in K} \mathbb{E}_x[V(x_1)] = \beta$. Then $\mathbb{E}_x[V(x_t)] \leq \lambda_\circ^t V(x) + \beta(1 - \lambda_\circ)^{-1}$ for all $x \in \mathbb{R}^d$ and $t \in \mathbb{N}_0$.*

The hypotheses of Proposition 1 imply

$$(3.1) \quad \mathbb{E}_x[V(x_1)] - V(x) \leq -(1 - \lambda_\circ)V(x) \quad \text{for all } x \notin K,$$

which is sometimes known as a “geometric drift condition.” The condition (3.1) is strong—the expected value of the function V decreases by a fraction λ_\circ in one step for all boundary conditions x outside a compact set. See e.g., [13] for further details, discussions, and applications of Proposition 1.

Among the weakest drift conditions, we have the following:

Proposition 2. *Let $(x_t)_{t \in \mathbb{N}_0}$ be a Markov process. Suppose that there exist $\beta, M, \varepsilon > 0$, a measurable function $V : \mathbb{R}^d \rightarrow [0, +\infty[$, and a compact set $K \subset \mathbb{R}^d$ such that*

$$(3.2) \quad \mathbb{E}_x[V(x_1)] - V(x) \leq -\beta \quad \text{for all } x \notin K, \quad \text{and}$$

$$(3.3) \quad \mathbb{E}[|V(x_{t+1}) - V(x_t)|^{2+\varepsilon} \mid (V(x_s))_{s=0}^t] \leq M \quad \text{for all } t \in \mathbb{N}_0.$$

Then for each $x \in \mathbb{R}^d$ the sequence $(\mathbb{E}_x[V(x_t)])_{t \in \mathbb{N}_0}$ is bounded.

Proposition 2 stipulates a *constant* negative drift (3.2) outside a compact set, as opposed to a *geometric* negative drift in (3.1). The condition (3.2) is rather weak, and the price for weakening the drift condition is the introduction of a uniform bound (3.3) on the jumps of the process $(x_t)_{t \in \mathbb{N}_0}$. In general, both the conditions

⁵General results dealing with stability but not relying on negative drift conditions are rare; one example may be found in [4].

(3.2) and (3.3) are necessary, and the $(2 + \varepsilon)$ exponent in (3.3) is tight; see [14] for details and (counter-)examples. An application of Proposition 2 to control of linear systems may be found in [16], and to receding horizon control in [8].

Propositions 1 and 2 may be viewed as representing two extremes of the spectrum of stability results involving negative drift conditions. We refer the reader to [13] for other drift conditions and their corresponding assertions concerning stability. Proposition 2 also highlights some of the features peculiar to stochastic control—indeed, while in the deterministic setting, the drift (in terms of Lyapunov functions) needs to be merely negative definite to ensure global asymptotic convergence of the system, the stability assertions in the stochastic setting depend crucially on the functional nature of the drift in addition to other conditions.

3.2. Stability under appropriate selection of cost functions.

ASSUMPTION. In addition to (A2), we stipulate that

- (A3) there exist a measurable feedback control function $g : \mathbb{R}^d \rightarrow \mathbb{U}$, a number $b \geq 0$, and a measurable and bounded set $K \subset \mathbb{R}^d$ such that
- ((A3)-i) $\sup_{z \in K} \left\{ c(z, g(z)) - c_F(z) + \mathbb{E}[c_F \circ f(z, g(z), w_0)] \right\} \leq b$,
 - ((A3)-ii) $c(z, g(z)) + \mathbb{E}[c_F \circ f(z, g(z), w_0)] \leq c_F(z)$ for all $z \notin K$. \diamond

Observe that ((A3)-ii) is a negative drift condition in disguise: indeed, it is precisely

$$\mathbb{E}_x[c_F(x_1^g)] - c_F(x) \leq -c(x, g(x)) \quad \text{for all } x \notin K,$$

where $x_1^g := f(x, g(x), w_0)$; here the cost functions c_F and c play the roles of the functions Ξ and Υ in (D), respectively. (“Global” conditions, similar in spirit to ((A3)-ii), in the context of stochastic receding horizon control have been proposed recently in [15].) However, since the stabilizing feedback g is not necessarily identical to π_0^* , the condition ((A3)-ii) does not guarantee stability under the receding horizon control policy $\hat{\pi}$. Nevertheless, from the boundedness condition ((A3)-i) and the drift condition ((A3)-ii), both expressed in terms of the cost functions c and c_F , we can establish the following drift condition involving the optimal value function V_N^* corresponding to (2.3), and under the receding horizon policy $\hat{\pi}$:

Theorem 3. *Consider the controlled system (2.1) with its accompanying data, and the optimal control problem (2.3). Suppose that Assumption (A3) holds. Then*

$$(3.4) \quad \text{for any } x \in \mathbb{R}^d, \quad \mathbb{E}_x[V_N^*(x_1^*)] - V_N^*(x) \leq -c(x, \pi_0^*(x)) + b,$$

where $(x_t^*)_{t=0}^{N-1}$ is the sequence generated by the recursion (2.4). In particular, under the receding horizon policy $\hat{\pi}$ derived from (2.3), the closed-loop process $(x_t)_{t \in \mathbb{N}_0}$ generated by (2.6) satisfies

$$(3.5) \quad \text{for any } x \in \mathbb{R}^d, \quad \mathbb{E}_x^{\hat{\pi}}[V_N^*(x_1)] - V_N^*(x) \leq -c(x, \pi_0^*(x)) + b.$$

Theorem 3 is the first of our two main results, and it is a representative statement aimed at establishing a connection between receding horizon control and stability of Markov processes. Observe that even though (3.5) does not resemble a negative drift condition per se, a condition analogous to (D) can be extracted from (3.5) under appropriate assumptions on the function c .⁶ Once such a procedure has been carried out, one can apply appropriate results on stability of Markov processes, e.g.,

⁶Of course, (3.5) may not lead to a negative drift outside a bounded set, e.g., if the function c is bounded above by b , and even if a negative drift condition can be extracted from (3.5), it may

Propositions 1 or 2, to assert boundedness of the sequence $(\mathbf{E}_x^{\hat{\pi}}[V_N^*(x_t)])_{t \in \mathbb{N}_0}$. We shall illustrate applications of Theorem 3 through our results and examples in the sequel.

In general, analytical expressions of the optimal value functions V_N^* are difficult to obtain; however, V_N^* can be bounded above and below in terms of the final cost c_F and the cost-per-stage function c , respectively, as follows:

Proposition 4. *Consider the controlled system (2.1) with its accompanying data, and the optimal control problem (2.3). Suppose that Assumption (A3) holds. Then*

$$(3.6) \quad \begin{aligned} & c(x, \pi_0^*(x)) + (N-1) \inf_{(z,u) \in \mathbb{R}^d \times \mathbb{U}} c(z, u) \\ & + \inf_{(z,u) \in \mathbb{R}^d \times \mathbb{U}} \mathbf{E}[c_F(f(z, u, w_0))] \leq V_N^*(x) \leq c_F(x) + Nb. \end{aligned}$$

Proposition 4 can be employed in conjunction with Theorem 3 to arrive at stability conditions under additional hypotheses:

Proposition 5. *Consider the controlled system (2.1) with its accompanying data, and the optimal control problem (2.3). Suppose that Assumption (A3) holds. Assume further that the cost functions satisfy:*

- $\lim_{\|z\| \rightarrow +\infty} c_F(z) = +\infty$,
- there exist measurable functions $c_s : \mathbb{R}^d \rightarrow [0, +\infty[$ and $c_c : \mathbb{U} \rightarrow [0, +\infty[$ such that

$$c(z, v) = c_s(z) + c_c(v) \quad \text{for all } (z, v) \in \mathbb{R}^d \times \mathbb{U},$$

and

$$\lim_{\|z\| \rightarrow +\infty} c_s(z) = +\infty,$$

- there exist a constant $\alpha \in [0, 1[$ and a compact set $K \subset \mathbb{R}^d$ such that

$$c_s(z) \geq \alpha c_F(z) \quad \text{for all } z \notin K.$$

Then under the receding horizon policy $\hat{\pi}$, the function V_N^* satisfies a geometric drift condition outside some compact subset of \mathbb{R}^d . In particular, for each $x \in \mathbb{R}^d$ the sequence $(\mathbf{E}_x^{\hat{\pi}}[V_N^*(x_t)])_{t \in \mathbb{N}_0}$ is bounded.

Example 6 (The LQ problem). Consider the controlled system (2.1) with $f(x, u, w) = Ax + Bu + w$ for matrices $A \in \mathbb{R}^{d \times d}$ and $B \in \mathbb{R}^{d \times m}$. Let $\mathbb{U} = \mathbb{R}^m$ and $\mathbb{W} = \mathbb{R}^d$. Let w_0 have a continuous density on \mathbb{R}^d , $\mathbf{E}[w_0] = 0$, $\mathbf{E}[w_0 w_0^\top] = \Sigma$ for some non-negative definite and symmetric matrix $\Sigma \in \mathbb{R}^{d \times d}$. Assume that the pair (A, B) is stabilizable. Then, by [2, Proposition 11.10.5], for every symmetric and positive definite matrix $Q \in \mathbb{R}^{d \times d}$ there exists a matrix $K \in \mathbb{R}^{m \times d}$ and a symmetric and positive definite matrix $P \in \mathbb{R}^{d \times d}$ such that $(A + BK)^\top P (A + BK) - P = -Q$. Consider the policy $\pi = (g, g, \dots)$, where $\mathbb{R}^d \ni x \mapsto g(x) := Kx \in \mathbb{R}^m$, and define the function $\mathbb{R}^d \ni x \mapsto V(x) := x^\top P x \in [0, +\infty[$. Then it follows from Proposition 1 and standard arguments that the closed-loop process $(x_t)_{t \in \mathbb{N}_0}$ under the policy π is stable in the sense that $\mathbf{E}_x[V(x_t)] \leq \lambda_\circ^t V(x) + \beta(1 - \lambda_\circ)^{-1}$ for all $t \in \mathbb{N}_0$,

$$\begin{aligned} \lambda_\circ &= \frac{1}{2} \left(1 - \frac{\sigma_{\min}(Q)}{\sigma_{\max}(P)} \right), \\ K &= \left\{ z \in \mathbb{R}^d \mid x^\top P x \leq \frac{2}{\lambda_\circ} \text{trace}(P\Sigma) \right\}, \end{aligned}$$

not be possible to assert boundedness of the sequence $(\mathbf{E}_x^{\hat{\pi}}[V_N^*(x_t)])_{t \in \mathbb{N}_0}$, e.g., if the conditions (3.2) and (3.3) do not hold simultaneously.

$$\beta = \sup_{z \in K} \{x^\top (A + BK)^\top P (A + BK)x + \text{trace}(P\Sigma)\},$$

where $\sigma_{\min}(M)$ and $\sigma_{\max}(M)$ denote the minimal and maximal singular values of a matrix M .

In view of the above computations, we define a symmetric and non-negative definite matrix $R \in \mathbb{R}^{m \times m}$ such that $K^\top R K \leq Q$, where the relation “ \leq ” between the preceding matrices denotes the standard matrix partial order among symmetric non-negative definite matrices. Let us define cost functions $c(z, u) := (1 - \alpha)z^\top Qz + \alpha u^\top R u$ and $c_F(z) := z^\top Pz$ for $(z, u) \in \mathbb{R}^d \times \mathbb{R}^m$ and $\alpha \in [0, 1]$. Straightforward calculations show that ((A3)-i) and ((A3)-ii) hold with $g(z) = Kz$ and the preceding definitions of c , c_F , β , and the compact set K . Consider now the optimal control problem (2.3) for a given $N \in \mathbb{N}$ and the control set $\mathbb{U} = \mathbb{R}^m$. By Theorem 3, we see that a receding horizon controller derived from this optimal control problem ensures that (3.5) holds. It is also possible to verify the hypotheses of Proposition 5 in this case, which implies that for each $x \in \mathbb{R}^d$ the sequence $(\mathbf{E}_x^\pi [V_N^*(x_t)])_{t \in \mathbb{N}_0}$ is bounded. \triangle

While Example 6 is entirely standard, it highlights a few noteworthy features of control of linear systems with affine noise, summarized in the following:

Remark 7.

- a) In the context of linear systems, the condition (3.1) implies that at all states x of large enough norm, the control action must be strong enough to achieve this geometric decrease. In the absence of a bound on the magnitude of the control actions, it is possible to synthesize linear feedback policies, such that a geometric drift in terms of quadratic functions V is attained; for instance, consider the feedback policy (g, g, \dots) with $g(x) = Kx$ in Example 6.
- b) If the control actions are *bounded*, a control policy whose elements are linear maps of the states is inadmissible. In this case, if the noise has unbounded support, e.g., w_t is a zero-mean Gaussian with a given variance matrix, then the following four cases appear naturally:
 - ▷ if the system matrix A has an eigenvalue outside the closed unit disc, with no control policy is it possible to ensure a geometric drift condition with quadratic Lyapunov functions V ;
 - ▷ if all eigenvalues of A are inside the open unit disc, then irrespective of the feedback policy, a geometric drift condition can be found for a quadratic Lyapunov function V , as illustrated in Example 6;
 - ▷ if A is Lyapunov stable, a constant (as opposed to geometric) negative drift condition for the Lyapunov function $V(z) := \|z\|$ was demonstrated in [16]; we provide a geometric drift condition under the same setting in Proposition 8 below;
 - ▷ if A has eigenvalues on the unit circle but with unequal algebraic and geometric multiplicities, the problem of stabilization under bounded controls remains an open problem; see [5] for details.
- c) Consider the scalar version of Example 6 with $(w_t)_{t \in \mathbb{N}_0}$ a sequence of mutually independent standard normal random variables. Since

$$\mathbb{P}\left(\inf_{t \in \mathbb{N}_0} w_t = -\infty \text{ and } \sup_{t \in \mathbb{N}_0} w_t = +\infty\right) = 1,$$

it is impossible to assert almost sure convergence of the states to any compact set under any policy. For the same reason, it is also impossible to assert a

statement of the form

$$P(\exists t_0 \in \mathbb{N}, K \subset \mathbb{R} \text{ compact, such that } x_t \in K \text{ for all } t \geq t_0) \geq 1 - \varepsilon$$

for $\varepsilon \in]0, 1[$ preassigned. In other words, under any feedback policy, almost surely, there will be excursions of the states beyond any given compact set infinitely often over an infinite time horizon. An identical assertion carries over to the multidimensional case. \triangleleft

Preparatory to providing further examples illustrating Theorem 3, we establish a stability result for a particular class of linear systems, as promised in Remark 7-b): Consider the controlled system

$$(3.7) \quad x_{t+1} = Ax_t + Bu_t + w_t, \quad x_0 \text{ given, } t \in \mathbb{N}_0,$$

for given matrices $A \in \mathbb{R}^{d \times d}$, $B \in \mathbb{R}^{d \times m}$, and suppose that $(w_t)_{t \in \mathbb{N}_0}$ is a sequence of independent and identically distributed (i.i.d.) random vectors. Suppose that pair (A, B) is controllable with reachability index κ .⁷ We define $\mathcal{R}(A, M) := (A^{\kappa-1}M \ \cdots \ AM \ M)$ for a matrix M with d rows. For a matrix M we let M^+ denote its Moore-Penrose pseudoinverse. Let $U_{\max} > 0$ be given, and suppose that $\|u_t\| \leq U_{\max}$ for all $t \in \mathbb{N}_0$. For $r > 0$, we define the radial saturation function $\mathbb{R}^d \ni z \mapsto \text{sat}_r(z) := \min\{r, \|z\|\} \frac{z}{\|z\|}$ if $z \neq 0$ and 0 otherwise. Let I_d denote the $d \times d$ identity matrix. The following proposition proposes a bounded control function such that a geometric drift condition is attained:

Proposition 8. *Consider the linear controlled system (3.7) with its accompanying data, suppose that A is orthogonal, and that w_0 is Gaussian with mean 0 in \mathbb{R}^d and given variance $\Sigma \in \mathbb{R}^{d \times d}$. Let $\rho := \ln \mathbb{E} \left[\exp \left(\left\| \mathcal{R}(A, I_d) (w_0^\top \ \cdots \ w_{\kappa-1}^\top)^\top \right\| \right) \right]$. Suppose that $U_{\max} > \rho$, define $V(x) := e^{\|x\|}$ for $x \in \mathbb{R}^d$, and let $K := \{z \in \mathbb{R}^d \mid \|z\| \leq 2\rho\}$. Then under the control actions*

$$(3.8) \quad \begin{pmatrix} u_{\kappa t} \\ \vdots \\ u_{\kappa(t+1)-1} \end{pmatrix} := -\mathcal{R}(A, B)^+ \text{sat}_{U_{\max}}(A^\kappa x_{\kappa t}), \quad t \in \mathbb{N}_0,$$

the closed-loop process $(x_{\kappa t})_{t \in \mathbb{N}_0}$ is Markovian, and there exists $\lambda_\circ \in]0, 1[$ such that

$$\mathbb{E}[V(x_{\kappa(t+1)}) \mid x_{\kappa t}] \leq \lambda_\circ V(x_{\kappa t}) \quad \text{on the set } \{x_{\kappa t} \notin K\}.$$

In particular, for each $x \in \mathbb{R}^d$ the sequence $(\mathbb{E}_x[e^{\|x_t\|}])_{t \in \mathbb{N}_0}$ is bounded, and the conditionally probability distributions $P_x(\|x_t\| > r), t \in \mathbb{N}_0$, have exponentially thin tails uniformly over $t \in \mathbb{N}_0$.

Proposition 8 is of independent interest. Stability of (3.7) under bounded controls was considered in [16], where the authors demonstrated that the same control actions as in (3.8), but under weaker assumptions on the noise,⁸ led to a constant negative drift of the function $\|\cdot\|$ outside a certain compact set of the closed-loop sub-sampled process $(x_{\kappa t})_{t \in \mathbb{N}_0}$. The technical tools in [16] relied on the considerably involved results of [14]; in contrast, the proof of Proposition 8 that we provide here relies only on the basic Proposition 1—namely, a geometric drift condition expressed in terms of $e^{\|\cdot\|}$. Note that, Proposition 8 asserts boundedness of $(\mathbb{E}_x[e^{\|x_t\|}])_{t \in \mathbb{N}_0}$, which is a stronger statement compared to, and indeed implies, boundedness of $(\mathbb{E}_x[\|x_t\|^2])_{t \in \mathbb{N}_0}$ asserted in [16]. Compare, in particular, that the

⁷That is, $\text{rank}(B \ AB \ \cdots \ A^{\kappa-1}B) = d$.

⁸To be precise, it was assumed that $\sup_{t \in \mathbb{N}_0} \mathbb{E}[\|w_t\|^4] < +\infty$. The result in [16], therefore, applies to noise sequences $(w_t)_{t \in \mathbb{N}_0}$ more general than Gaussians.

conditional distributions $\mathbb{P}_x(\|x_t\| > r)$ have exponentially thin tails, by Proposition 8, if the maximum magnitude of the control actions is large enough, while the main result of [16] asserts that the corresponding distributions have tails that decay faster than inverse quadratically as r grows large.

Example 9. Consider the controlled system (2.1) with $f(x, u, w) = x + u + w$, where $x, u, w \in \mathbb{R}$. Let w_0 have a continuous density on \mathbb{R} , $\mathbb{E}[w_0] = 0$, $\mathbb{E}[|w_0|^4] < +\infty$. Fix $N \in \mathbb{N}$. Let us investigate the possibility of stability under a receding horizon policy derived from the N -horizon optimal control problem⁹

$$(3.9) \quad \begin{aligned} & \text{minimize} && \mathbb{E}_x^\pi \left[\sum_{t=0}^{N-1} \mathbf{1}_{\mathbb{R} \setminus [-2, 2]}(x_t) + \|x_N\| \right] \\ & \text{subject to} && \begin{cases} \pi \in \Pi & \text{where } \Pi \text{ a class of Markovian policies,} \\ \pi_i(z) \in [-1, 1] & \text{for all } i = 0, \dots, N-1 \text{ and } z \in \mathbb{R}, \\ \text{dynamics (2.1)} & \text{with } f(x, u, w) = x + u + w. \end{cases} \end{aligned}$$

The cost-per-stage function $c(z, u) = \mathbf{1}_{\mathbb{R} \setminus [-2, 2]}(z)$ ensures that for each realization of the random noise, the cost grows proportionately to the duration that the state stays out of the set $[-2, 2]$. The policy that solves the minimization problem (3.9) drives the state inside $[-2, 2]$ as fast as possible, and the final cost c_F regulates the final state close to 0. In the light of Proposition 2 (e.g., following the arguments in [16],) it is not difficult to verify that the policy (g, g, \dots) , with $g(x) = -\text{sat}(x)$ ensures that¹⁰

$$\mathbb{E}[|x_{t+1}| \mid x_t] - |x_t| \leq -1 \quad \text{on the set } \{|x_t| > 2\}, \quad t \in \mathbb{N}_0.$$

Note that the closed-loop system is Markovian. One sees after standard computations that the optimal control problem (3.9) with $K := \{z \in \mathbb{R} \mid |z| \leq 2\}$ verifies both ((A3)-i) and ((A3)-ii). The selection of the cost-per-stage function c is not unique; e.g., $c(z, u) = \frac{1}{2}(\mathbf{1}_{\mathbb{R} \setminus [-2, 2]}(z) + \mathbf{1}_{[-1, 1] \setminus [-\frac{1}{2}, \frac{1}{2}]}(u))$ works just as fine insofar as the matter of satisfying ((A3)-i) and ((A3)-ii) is concerned. Numerical tractability of the problem (3.9), even under the assumption that $(w_t)_{t \in \mathbb{N}_0}$ is a sequence of independent and identically distributed standard normal random variables, is a non-trivial matter. \triangle

Example 10 (Example 9 continued). Consider the controlled system (2.1) with $f(x, u, w) = x + u + w$, where $x, u, w \in \mathbb{R}$. Let $(w_t)_{t \in \mathbb{N}_0}$ be a sequence of i.i.d. Gaussian random variables with zero-mean and variance σ^2 for some given $\sigma > 0$, and define $\rho := \ln(\sigma\sqrt{\frac{2}{\pi}})$. Let $U_{\max} > \rho$ be given, and define $\lambda_o := e^{\rho - U_{\max}}$. Fix $N \in \mathbb{N}$. Let us investigate the possibility of stability under a receding horizon policy derived from the N -horizon optimal control problem

$$(3.10) \quad \begin{aligned} & \text{minimize} && \mathbb{E}_x^\pi \left[(1 - \lambda_o) \sum_{t=0}^{N-1} e^{|x_t|} + e^{|x_N|} \right] \\ & \text{subject to} && \begin{cases} \pi \in \Pi & \text{where } \Pi \text{ a class of Markovian policies,} \\ |\pi_i(z)| \leq U_{\max} & \text{for all } i = 0, \dots, N-1, z \in \mathbb{R}, \\ \text{dynamics (2.1)} & \text{with } f(x, u, w) = x + u + w. \end{cases} \end{aligned}$$

Namely, we have the final cost $c_F(z) = e^{|z|}$ and the cost-per-stage function $c(z, u) = (1 - \lambda_o)e^{|z|}$. By Proposition 8 it follows that ((A3)-ii) holds for some compact

⁹Here $\mathbf{1}_A(\cdot)$ denotes the indicator function of the set A , defined as $\mathbf{1}_A(z) = 1$ if $z \in A$ and 0 otherwise.

¹⁰Here $\text{sat}(\cdot)$ is the standard saturation function defined as $\text{sat}(z) = z$ for $z \in [-1, 1]$, 1 if $z > 1$, and -1 otherwise.

$K \subset \mathbb{R}$, and it is straightforward to verify ((A3)-i). Theorem 3 guarantees that a receding horizon controller derived from (3.10) ensures (3.5). An extension to the multidimensional case can be readily obtained with the technical support of Proposition 8. It also follows from standard computations that the hypotheses of Proposition 5 hold in this case, which implies that for each $x \in \mathbb{R}$ the sequence $(\mathbb{E}_x^\pi [V_N^*(x_t)])_{t \in \mathbb{N}_0}$ is bounded. \triangle

4. PERFORMANCE UNDER RECEDING HORIZON CONTROL

In this section we study performance of closed-loop systems under receding horizon control. Our objective here is to arrive at quantitative bounds on the performance of receding horizon policies over an infinite temporal horizon.

To this end, we must first select a performance index. We contend that the cost-per-stage function c is a natural candidate with which performance at each time step may be measured. For one, while the final cost c_F plays an important role in the problem (2.3), the first element π_0^* of the optimal control policy π^* enters the function c but not c_F ; since a receding horizon policy is constructed out of π_0^* , the function c is perhaps a more natural candidate compared to c_F for measuring performance at each time step. Moreover, the function c involves both the states and the control actions, while c_F involves only the states; as such, the expected value of $c(x_t, u_t)$ at time t reflects the performance measured with respect to both the states and the control actions.

In the setting of the dynamical system (2.1) involving stochastic uncertainties, typically, a sum of cost-per-stage functions over n steps grows with n rather quickly, and the expected total cost

$$\mathbb{E}_x \left[\sum_{t=0}^{+\infty} c(x_t, u_t) \right]$$

may not be suitable for measuring performance. For instance, in the case of the standard optimal linear quadratic regulator, under the optimal policy the quantity $\mathbb{E}_x \left[\sum_{t=0}^n (x_t^\top Q x_t + u_t^\top R u_t) \right]$ grows linearly with n (under mild assumptions on the non-negative definite states- and control-weight matrices Q and R , respectively); consequently, the expected total cost over an infinite horizon is not bounded. A more appropriate measure of performance is the *long-run expected average cost*, measured in terms of the cost-per-stage function c , defined by

$$(4.1) \quad \limsup_{n \rightarrow \infty} \frac{1}{n+1} \mathbb{E}_x \left[\sum_{t=0}^n c(x_t, u_t) \right].$$

This is the performance index that we adopt here. In particular, the quantity in (4.1) is well-defined for the linear quadratic problem under mild hypothesis (stabilizability of the underlying linear system). The intuition is clear: the quantity in (4.1) measures the cost averaged over time and averaged across all possible realizations of the process $(x_t)_{t \in \mathbb{N}_0}$. Observe that all phenomena that occur over a finite temporal horizon, or are transient in the sense that they asymptotically die out, do not affect the index (4.1).

We next provide our second main result—an estimate of performance under the receding horizon policy $\hat{\pi}$ in terms of the long-run expected average cost. It also shows how stability of the closed-loop process influences the long-run expected average cost.

Theorem 11. Consider the controlled system (2.1) with its accompanying data, and the optimal control problem (2.3). Let $\hat{\pi} := (\pi_0^*, \pi_0^*, \dots)$ denote the receding horizon policy derived from (2.3), and for a measurable feedback \tilde{g} we define

$$(4.2) \quad T_{\tilde{g}}(z) := c(z, \tilde{g}(z)) - c_F(z) + \mathbf{E}[c_F \circ f(z, \tilde{g}(z), w_0)], \quad z \in \mathbb{R}^d.$$

Then the following hold:

(11.i) If Assumption (A2) holds, then for every $x \in \mathbb{R}^d$ and $k \in \mathbb{N}$,

$$(4.3) \quad \mathbf{E}_{\hat{\pi}} \left[\sum_{\ell=0}^k c(x_\ell, u_\ell) \right] \leq V_N^*(x) - \mathbf{E}_{\hat{\pi}} [V_N^*(x_{k+1})] \\ + \sum_{\ell=0}^k \mathbf{E}_{\hat{\pi}} \left[\mathbf{E}^{\pi^*} [T_{\tilde{g}}(x_{\ell+N}) \mid x_\ell] \right].$$

(11.ii) If Assumption (A3) holds and

$$(4.4) \quad \lim_{n \rightarrow +\infty} \frac{\mathbf{E}_{\hat{\pi}} [V_N^*(x_n)]}{n} = 0 \quad \text{for all } x \in \mathbb{R}^d,$$

then

$$(4.5) \quad \limsup_{k \rightarrow +\infty} \frac{1}{k+1} \mathbf{E}_{\hat{\pi}} \left[\sum_{\ell=0}^k c(x_\ell, u_\ell) \right] \leq b \quad \text{for all } x \in \mathbb{R}^d.$$

In particular, if for all $x \in \mathbb{R}^d$ the sequence $(\mathbf{E}_{\hat{\pi}} [V_N^*(x_t)])_{t \in \mathbb{N}_0}$ is bounded, then (4.5) holds.

Remark 12.

- a) The estimate (4.3) in part (11.i) is quite general, and holds under the mild hypotheses (A2) of a technical nature. The part (11.ii) follows from (11.i) under the additional Assumption (A3), as will be evident from the proof of Theorem 11 presented in Appendix B.
- b) We reiterate that the bound on the expected average cost in (4.5) corresponds to the receding horizon policy $\hat{\pi}$, not the stabilizing policy (g, g, \dots) . Indeed, although Assumption (A3) stipulates the existence of a stabilizing feedback g , this feedback controller is never applied. From the proof of Theorem 11 one sees that the condition (3.5) in Theorem 3 plays a crucial role in (11.ii).
- c) The condition (4.4) is technical in nature. In most practical cases, one ensures boundedness of $(\mathbf{E}_{\hat{\pi}} [V_N^*(x_t)])_{t \in \mathbb{N}_0}$ with the help of Theorem 3, and consequently (4.5) holds.
- d) Theorem 11 provides a quantitative bound on the performance measured in terms of the cost-per-stage function c . Since the stability conditions presented in §3.2 also involve the function c , it is possible to establish a connection between the performance bound (4.5) with the stability conditions in Assumption (A3). In general, establishing such a connection may not be possible if stability under the receding horizon policy $\hat{\pi}$ is ensured by means of adjoining an appropriate constraint, as discussed in (S2) above; indeed, such a constraint may have no relation whatsoever to the cost functions c and c_F . Nevertheless, the inequality in (4.3) holds irrespective of whether Assumption (A3) is satisfied or not. Consequently, if a constraint is adjoining to the problem (2.3) such that the process $(x_t)_{t \in \mathbb{N}_0}$ under the receding horizon policy $\hat{\pi}$ satisfies (4.4), and the sum on the right-hand side of (4.3) increases at most linearly with k , then a bound on the long-run expected average cost can be extracted. Naturally, these verifications, which are likely to be case-specific, need additional analysis.

- e) A direct optimal control problem involving minimization of the long-run expected average cost criterion for dynamical systems, e.g.,

$$(4.6) \quad \begin{aligned} & \text{minimize} && \limsup_{n \rightarrow +\infty} \frac{1}{n+1} \mathbb{E}_x^\pi \left[\sum_{t=0}^n c(x_t, u_t) \right] \\ & \text{subject to} && \begin{cases} \pi \in \Pi, \\ \text{dynamics (2.1),} \end{cases} \end{aligned}$$

is generally difficult to solve both analytically and numerically—see [6, Chapter 5] for further details. The numerical value of the bound b in (4.5) on the performance index (4.1) under the receding horizon technique may be employed as a measure to decide whether to adopt a receding horizon strategy as an alternative to a direct solution to the minimization problem (4.6). \triangleleft

Example 13 (Example 6 cont'd). Consider the linear system in Example 6 and the selection of cost functions c and c_F as in the final part of Example 6. Standard computations, e.g., as in [6, Chapter 3], show that the function V_N^* is quadratic. In conjunction with the final calculations in Example 6, this shows that (3.5) is a geometric drift condition in this case. Proposition 1 implies, therefore, that under receding horizon control derived from (2.3), for each $x \in \mathbb{R}^d$ the sequence $(\mathbb{E}_x^{\hat{\pi}}[V_N^*(x_n)])_{n \in \mathbb{N}_0}$ is bounded. Finally, by Theorem 11 part (11.ii) we see that for each $x \in \mathbb{R}^d$

$$\limsup_{n \rightarrow \infty} \frac{1}{n+1} \mathbb{E}_x^{\hat{\pi}} \left[\sum_{t=0}^n c(x_t, u_t) \right] \leq \beta,$$

where β is the constant defined in Example 6. \triangle

Example 14 (Example 9 cont'd). A direct application of Proposition 5 does not appear to be possible. However, one can verify that for each $x \in \mathbb{R}^d$ the sequence $(\mathbb{E}_x^{\hat{\pi}}[V_N^*(x_t)])_{t \in \mathbb{N}_0}$ is bounded by following the arguments in [16] with the support of Proposition 2. This in turn implies by Theorem 11 that the long-run expected average cost is finite under the receding horizon policy extracted from (3.9). \triangle

Example 15 (Example 10 cont'd). In this case it is possible to directly apply Proposition 5; we skip the standard computations needed to verify its hypotheses. Consequently, the sequence $(\mathbb{E}_x^{\hat{\pi}}[V_N^*(x_t)])_{t \in \mathbb{N}_0}$ is bounded. By Theorem 11 this implies that the long-run expected average cost is finite under the receding horizon policy extracted from (3.9). \triangle

APPENDIX A.

This appendix collects our proofs of the results presented in §3.

Proof of Proposition 1. A proof of this proposition is standard; we include it merely for completeness. The assertion follows from the Markovian property of $(x_t)_{t \in \mathbb{N}_0}$ and an iteration scheme as follows:

$$\begin{aligned} \mathbb{E}_x[V(x_t)] &= \mathbb{E}_x[\mathbb{E}[V(x_t) \mid x_{t-1}]] \\ &\leq \mathbb{E}_x[\lambda_\circ V(x_{t-1}) \mathbf{1}_{\mathbb{R}^d \setminus K}(x_{t-1}) + \beta \mathbf{1}_K(x_{t-1})] \\ &\leq \lambda_\circ \mathbb{E}_x[V(x_{t-1})] + \beta \mathbb{P}_x(x_{t-1} \in K) \\ &\vdots \end{aligned}$$

$$\begin{aligned}
&\leq \lambda_o^t V(x) + \beta \sum_{k=0}^{t-1} \lambda_o^{t-1-k} \mathbf{P}_x(x_k \in K) \\
&\leq \lambda_o^t V(x) + \frac{\beta}{1 - \lambda_o}. \quad \square
\end{aligned}$$

Proof of Proposition 2. This is a special case of [14, Theorem 1]. \square

Proof of Theorem 3. From the last $N - 1$ elements of the optimal policy π^* we derive the N -length policy $\tilde{\pi} := (\pi_1^*, \dots, \pi_{N-1}^*, g)$, where $g : \mathbb{R}^d \rightarrow \mathbb{U}$ is the feedback function in Assumption (A3). Let $\tilde{V}_N(x) := V_N(\tilde{\pi}, x)$. Recall that the sequence $(x_t^*)_{t=0}^{N-1}$ is generated by the recursion (2.4). Then, by optimality of the policy π^* ,

$$\begin{aligned}
&\mathbf{E}_x[V_N^*(x_1^*)] - V_N^*(x) \\
&\leq \mathbf{E}_x[\tilde{V}_N(x_1^*)] - V_N^*(x) \\
&= \mathbf{E}_x \left[\mathbf{E} \left[\sum_{t=1}^{N-1} c(x_t^*, \pi_t^*(x_t^*)) + c(x_N^*, g(x_N^*)) + c_F(f(x_N^*, g(x_N^*), w_N)) \mid x_1^* \right] \right. \\
&\quad \left. - \mathbf{E}_x \left[\sum_{t=0}^{N-1} c(x_t^*, \pi_t^*(x_t^*)) + c_F(x_N^*) \right] \right] \\
&= -c(x, \pi_0^*(x)) + \mathbf{E}_x \left[c(x_N^*, g(x_N^*)) + c_F \circ f(x_N^*, g(x_N^*), w_N) - c_F(x_N^*) \right].
\end{aligned}$$

If the condition (A3) holds, then the tower property of conditional expectations implies that

$$\begin{aligned}
&\mathbf{E}_x \left[c(x_N^*, g(x_N^*)) + c_F \circ f(x_N^*, g(x_N^*), w_N) - c_F(x_N^*) \right] \\
&= \mathbf{E}_x \left[\mathbf{E} \left[c(x_N^*, g(x_N^*)) + c_F \circ f(x_N^*, g(x_N^*), w_N) - c_F(x_N^*) \mid \{x_\ell^*\}_{\ell=1}^N \right] \right] \\
&\leq \mathbf{E}_x \left[b \mathbf{1}_{\{x_N^* \in K\}} \right] \\
&= b \mathbf{P}_x(x_N^* \in K).
\end{aligned}$$

Substituting back we obtain

$$\mathbf{E}_x[V_N^*(x_1^*)] - V_N^*(x) \leq -c(x, \pi_0^*(x)) + b,$$

which proves (3.4). The assertion (3.5) follows from the facts that $\hat{\pi} = (\pi_0^*, \pi_0^*, \dots)$, and that the closed-loop process is Markovian under $\hat{\pi}$. \square

Proof of Proposition 4. Fix $x \in \mathbb{R}^d$. The first inequality $c(x, \pi_0^*(x)) \leq V_N^*(x)$ follows immediately from the definition of V_N^* . We now prove the second inequality. Let the sequence $(x_t^g)_{t=1}^N$ be defined by

$$x_t^g = \begin{cases} x & \text{if } t = 0, \\ f(x_{t-1}^g, g(x_{t-1}^g), w_{t-1}) & \text{if } t = 1, \dots, N. \end{cases}$$

In view of ((A3)-ii), we have

$$\begin{aligned}
c_F(x) - V_N^*(x) &= c_F(x) - c(x, g(x)) - V_N^*(x) + c(x, g(x)) \\
&\geq \mathbf{E} \left[c_F \circ f(x, g(x), w_0) \right] - V_N^*(x) + c(x, g(x)) - b \\
&= \mathbf{E} \left[c_F(x_1^g) - c(x_1^g, g(x_1^g)) \mid x_0^g = x \right] - V_N^*(x) \\
&\quad + \mathbf{E} \left[c(x_1^g, g(x_1^g)) \mid x_0^g = x \right] + c(x_0^g, g(x_0^g)) - b
\end{aligned}$$

$$\begin{aligned}
&\geq \mathbb{E}[c_F(x_2^g) \mid x_0^g = x] - V_N^*(x) \\
&\quad + \sum_{\ell=0}^1 \mathbb{E}[c(x_\ell^g, g(x_\ell^g)) \mid x_0^g = x] - 2b \\
&\quad \vdots \\
&\geq \mathbb{E}[c_F(x_N^g) \mid x_0^g = x] - V_N^*(x) - Nb \\
&\quad + \sum_{\ell=0}^{N-1} \mathbb{E}[c(x_\ell^g, g(x_\ell^g)) \mid x_0^g = x].
\end{aligned}$$

Since $\overbrace{(g, \dots, g)}^{N \text{ times}}$ is a sub-optimal policy for the optimal control problem (2.3),

$$V_N^*(x) \leq \mathbb{E}\left[\sum_{\ell=0}^{N-1} c(x_\ell^g, g(x_\ell^g)) + c_F(x_N^g) \mid x_0^g = x\right],$$

and therefore, $c_F(x) - V_N^*(x) \geq -Nb$. Since x is arbitrary, the assertion follows. \square

Proof of Proposition 5. Fix $x \in \mathbb{R}^d$. We know from Theorem 3 that

$$\mathbb{E}_x^{\hat{\pi}}[V_N^*(x_1)] - V_N^*(x) \leq -c(x, \pi_0^*(x)) + b.$$

By hypothesis, $c(x, \pi_0^*(x)) = c_s(x) + c_c(\pi_0^*(x)) \geq c_s(x)$, and if $x \notin K$, then $c_s(x) \geq \alpha c_F(x)$. Thus,

$$\mathbb{E}_x^{\hat{\pi}}[V_N^*(x_1)] - V_N^*(x) \leq -c_s(x) + b \leq -\alpha c_F(x) + b \quad \text{if } x \notin K.$$

In view of (3.6), we see that if $x \notin K$, then $-\alpha c_F(x) + b \leq -\alpha V_N^*(x) + b(1 + \alpha N)$. In other words,

$$\mathbb{E}_x^{\hat{\pi}}[V_N^*(x_1)] - V_N^*(x) \leq -\alpha V_N^*(x) + b(1 + \alpha N) \quad \text{for all } x \notin K.$$

Since $\lim_{\|z\| \rightarrow +\infty} c_s(z) = +\infty$, our hypotheses show that $\lim_{\|z\| \rightarrow +\infty} c(z, \pi_0^*(z)) = +\infty$, and from (3.6) it follows that $\lim_{\|z\| \rightarrow +\infty} V_N^*(z) = +\infty$. By definition of a limit, therefore, there must exist a closed ball K' around $0 \in \mathbb{R}^d$ of radius large enough, such that $V_N^*(z) \geq 2(\alpha^{-1} + N)$ for all $z \notin K'$. Substituting back we see that

$$\mathbb{E}_x^{\hat{\pi}}[V_N^*(x_1)] - V_N^*(x) \leq -\frac{\alpha}{2} V_N^*(x) \quad \text{for all } x \notin K',$$

which is a geometric drift condition outside the compact set K' . The particular case follows from Proposition 1. \square

Proof of Proposition 8. Fix $t \in \mathbb{N}_0$. The state recursion (3.7) shows that

$$\begin{aligned}
x_{\kappa(t+1)} &= A^\kappa x_{\kappa t} + \mathcal{R}(A, B) \begin{pmatrix} u_{\kappa t} \\ \vdots \\ u_{\kappa(t+1)-1} \end{pmatrix} + \mathcal{R}(A, I_d) \begin{pmatrix} w_{\kappa t} \\ \vdots \\ w_{\kappa(t+1)-1} \end{pmatrix} \\
&=: A^\kappa x_{\kappa t} + \mathcal{R}(A, B) \overline{u_{\kappa t}} + \mathcal{R}(A, I_d) \overline{w_{\kappa t}}.
\end{aligned}$$

Let $\tilde{\lambda} := \rho - U_{\max}$ and let $\lambda_o := e^{\tilde{\lambda}}$; by hypothesis, $\tilde{\lambda} < 0$ and $\lambda_o \in]0, 1[$. On the event $\{\|x_{\kappa t}\| > 2R\}$, we have

$$\begin{aligned}
&\frac{\mathbb{E}[e^{\|x_{\kappa(t+1)}\|} \mid x_{\kappa t}]}{e^{\|x_{\kappa t}\|}} \\
&= \mathbb{E}\left[\exp(\|A^\kappa x_{\kappa t} + \mathcal{R}(A, B) \overline{u_{\kappa t}} + \mathcal{R}(A, I_d) \overline{w_{\kappa t}}\| - \|x_{\kappa t}\|) \mid x_{\kappa t}\right] \\
&= e^{\|A^\kappa x_{\kappa t} - \text{sat}_{U_{\max}}(A^\kappa x_{\kappa t})\| - \|A^\kappa x_{\kappa t}\|} \mathbb{E}[e^{\|\mathcal{R}(A, I_d) \overline{w_{\kappa t}}\|} \mid x_{\kappa t}] \quad \text{since } A \text{ is orthogonal}
\end{aligned}$$

$$\begin{aligned}
&= e^{-U_{\max}} \mathbb{E} \left[\exp \left(\left\| \mathcal{R}(A, I_d) \begin{pmatrix} w_0 \\ \vdots \\ w_{\kappa-1} \end{pmatrix} \right\| \right) \right] \quad \text{since } (w_s)_{s \in \mathbb{N}_0} \text{ is i.i.d.} \\
&= e^{-U_{\max} + \rho} \\
&= \lambda_o < 1 \quad \text{by hypothesis.}
\end{aligned}$$

Since t is arbitrary, the first claim follows. By hypothesis, the control actions $(u_{\kappa t}^\top \cdots u_{\kappa(t+1)-1}^\top)^\top$ depend only on $x_{\kappa t}$ for each $t \in \mathbb{N}_0$. Thus, the process $(x_{\kappa t})_{t \in \mathbb{N}_0}$ is Markovian under the control actions in (3.8), as can be seen by directly verifying (2.7). By Proposition 1 we see that for each $x \in \mathbb{R}^d$ the sequence $(\mathbb{E}_x[e^{\|x_{\kappa t}\|}])_{t \in \mathbb{N}_0}$ is bounded. It remains to move from the κ -subsampled process $(x_{\kappa t})_{t \in \mathbb{N}_0}$ to the original process $(x_t)_{t \in \mathbb{N}_0}$. But standard arguments, along the lines of the proof of the main theorem in [16], employing the triangle inequality and monotonicity of the function $e^{(\cdot)}$ shows that the sequence $(\mathbb{E}_x[e^{\|x_t\|}])_{t \in \mathbb{N}_0}$ is bounded.

For $x \in \mathbb{R}^d$ let $C = C(x) > 0$ be such that $\sup_{t \in \mathbb{N}_0} \mathbb{E}_x[e^{\|x_t\|}] \leq C$. Then for all $t \in \mathbb{N}_0$,

$$\begin{aligned}
C &\geq \mathbb{E}_x[e^{\|x_t\|}] = \mathbb{E}_x \left[\int_0^{\|x_t\|} e^r dr \right] + 1 \\
&= \mathbb{E}_x \left[\int_0^{+\infty} \mathbf{1}_{\{\|x_t\| > r\}} e^r dr \right] + 1 \\
&= \int_0^{+\infty} \mathbb{P}_x(\|x_t\| > r) e^r dr + 1 \quad \text{by Fubini's theorem.}
\end{aligned}$$

This shows that $\mathbb{P}_x(\|x_t\| > r)$ must decay, for large values of r , faster than e^{-r} uniformly for all $t \in \mathbb{N}_0$, and the assertion follows. \square

APPENDIX B.

This appendix contains our proof of Theorem 11. For two policies $\pi_{0:k_1-1}$ and $\pi'_{0:k_2-1}$ of length k_1 and k_2 , respectively, we define their concatenation

$$\pi_{0:k_1-1} \sharp \pi'_{0:k_2-1} := (\pi_0, \dots, \pi_{k_1-1}, \pi'_0, \dots, \pi'_{k_2-1}).$$

Proof of Theorem 11. (11.i): We adapt certain ideas from [3] to the context of long-run expected average cost. Suppose that Assumption (A2) holds, and fix $n \in \mathbb{N}$. Conditional on $x_{n+1} = x' \in \mathbb{R}^d$, by definition of optimality,

$$\begin{aligned}
&\mathbb{E}^{\pi_{1:N-1} \sharp \tilde{g}} \left[\sum_{\ell=n+1}^{n+N} c(x_\ell, u_\ell) + c_F(x_{n+N+1}) \mid x_{n+1} = x' \right] \\
&\geq \mathbb{E}^{\pi_{0:N-1}^*} \left[\sum_{\ell=n+1}^{n+N} c(x_\ell, u_\ell) + c_F(x_{n+N+1}) \mid x_{n+1} = x' \right]
\end{aligned}$$

Conditional on $x_n = y$, therefore,

$$\begin{aligned}
&\mathbb{E}^{\pi_0 \sharp \pi_{1:N-1} \sharp \tilde{g}} \left[\left(\sum_{\ell=n}^{n+N-1} c(x_\ell, u_\ell) + c_F(x_{n+N}) \right) - c(x_n, u_n) \mid x_n = y \right] \\
&+ \mathbb{E}^{\pi_0 \sharp \pi_{1:N-1} \sharp \tilde{g}} \left[c_F(x_{n+N+1}) - c_F(x_{n+N}) + c(x_{n+N}, u_{n+N}) \mid x_n = y \right]
\end{aligned}$$

$$\geq \mathbf{E}^{\pi_0^* \# \pi_{0:N-1}} \left[\sum_{\ell=n+1}^{n+N} c(x_\ell, u_\ell) + c_F(x_{n+N+1}) \mid x_n = y \right],$$

which implies

$$\begin{aligned} & \mathbf{E}^{\pi_{0:N-1}^* \# \tilde{g}} \left[\sum_{\ell=n}^{n+N-1} c(x_i, u_i) + c_F(x_{n+N}) \mid x_n = y \right] \\ & + \mathbf{E}^{\pi_{0:N-1}^* \# \tilde{g}} \left[c_F(x_{n+N+1}) - c_F(x_{n+N}) + c(x_{n+N}, u_{n+N}) \mid x_n = y \right] \\ & \geq c(y, \pi_0^*(y)) + \mathbf{E}^{\pi_0^* \# \pi_{0:N-1}^*} \left[\sum_{\ell=n+1}^{n+N} c(x_i, u_i) + c_F(x_{n+N+1}) \mid x_n = y \right]. \end{aligned}$$

Rearranging terms, conditional on $x_n = y$, we get

$$\begin{aligned} c(y, \pi_0^*(y)) & \leq V_N^*(y) - \mathbf{E}^{\pi_0^*} [V_N^*(x_{n+1}) \mid x_n = y] \\ & \quad + \mathbf{E}^{\pi^*} [c(x_{n+N}, \tilde{g}(x_{n+N})) - c_F(x_{n+N}) \\ & \quad \quad + \mathbf{E} [c_F \circ f(x_{n+N}, \tilde{g}(x_{n+N}), w_{n+N}) \mid x_{n+N}] \mid x_n = y]. \end{aligned}$$

Suppose now that the receding horizon policy $\hat{\pi}$ is applied. Since the closed-loop process $(x_t)_{t \in \mathbb{N}_0}$ under $\hat{\pi}$ is Markovian, taking expectations under the policy $\hat{\pi}$, we get

$$\mathbf{E}_x^{\hat{\pi}} [c(x_n, \pi_0^*(x_n))] \leq \mathbf{E}_x^{\hat{\pi}} [V_N^*(x_n)] - \mathbf{E}_x^{\hat{\pi}} [V_N^*(x_{n+1})] + \mathbf{E}_x^{\hat{\pi}} [\mathbf{E}^{\pi^*} [T_{\tilde{g}}(x_{n+N}) \mid x_n]],$$

whence, summing from $n = 0$ through $n = k$ we arrive at

$$\mathbf{E}_x^{\hat{\pi}} \left[\sum_{n=0}^k c(x_n, \pi_0^*(x_n)) \right] \leq V_N^*(x) - \mathbf{E}_x^{\hat{\pi}} [V_N^*(x_{k+1})] + \sum_{n=0}^k \mathbf{E}_x^{\hat{\pi}} [\mathbf{E}^{\pi^*} [T_{\tilde{g}}(x_{n+N}) \mid x_n]],$$

as asserted.

(11.ii) Suppose that Assumption (A3) holds, and let $\tilde{g} = g$. Fix $x \in \mathbb{R}^d$. From (4.3), it follows that

$$\begin{aligned} & \frac{1}{k+1} \mathbf{E}_x^{\hat{\pi}} \left[\sum_{n=0}^k c(x_n, \pi_0^*(x_n)) \right] \\ & \leq \frac{1}{k+1} \left(V_N^*(x) - \mathbf{E}_x^{\hat{\pi}} [V_N^*(x_{k+1})] \right) + \frac{1}{k+1} \sum_{n=0}^k \mathbf{E}_x^{\hat{\pi}} [\mathbf{E}^{\pi^*} [T_g(x_{n+N}) \mid x_n]] \\ & \leq \frac{1}{k+1} \left(V_N^*(x) - \mathbf{E}_x^{\hat{\pi}} [V_N^*(x_{k+1})] \right) + b \quad \text{by Assumption (A3),} \end{aligned}$$

which implies that

$$\begin{aligned} \limsup_{k \rightarrow +\infty} \frac{1}{k+1} \mathbf{E}_x^{\hat{\pi}} \left[\sum_{n=0}^k c(x_n, \pi_0^*(x_n)) \right] & \leq \limsup_{k \rightarrow +\infty} \frac{V_N^*(x) - \mathbf{E}_x^{\hat{\pi}} [V_N^*(x_{k+1})]}{k+1} + b \\ & = b \quad \text{by hypothesis.} \end{aligned}$$

The final claim follows at once from the preceding. \square

REFERENCES

- [1] D. BERNARDINI AND A. BEMPORAD, *Scenario-based model predictive control of stochastic constrained linear systems*, in joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, Shanghai, P.R. China, Dec. 2009, pp. 6333–6338.
- [2] D. BERNSTEIN, *Matrix Mathematics*, Princeton University Press, Princeton, NJ, 2 ed., 2009. Theory, facts, and formulas with application to linear systems theory.
- [3] D. CHATTERJEE, E. CINQUEMANI, G. CHALOULOS, AND J. LYGEROS, *Stochastic control up to a hitting time: optimality and rolling-horizon implementation*. <http://arxiv.org/abs/0806.3008>, 2008.
- [4] D. CHATTERJEE AND S. PAL, *An excursion-theoretic approach to stability of discrete-time stochastic hybrid systems*, Applied Mathematics and Optimization, 63 (2011), pp. 217–237.
- [5] D. CHATTERJEE, F. RAMPONI, P. HOKAYEM, AND J. LYGEROS, *On mean square boundedness of stochastic linear systems with bounded controls*, Systems & Control Letters, 61 (2012), pp. 375–380.
- [6] O. HERNÁNDEZ-LERMA AND J. B. LASSERRE, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, vol. 30 of Applications of Mathematics, Springer-Verlag, New York, 1996.
- [7] ———, *Further Topics on Discrete-Time Markov Control Processes*, vol. 42 of Applications of Mathematics, Springer-Verlag, New York, 1999.
- [8] P. HOKAYEM, D. CHATTERJEE, F. RAMPONI, G. CHALOULOS, AND J. LYGEROS, *Stable stochastic receding horizon control of linear systems with bounded control inputs*, in Proceedings of 19th International Symposium on Mathematical Theory of Networks and Systems, 2010, pp. 31–36.
- [9] P. HOKAYEM, E. CINQUEMANI, D. CHATTERJEE, F. RAMPONI, AND J. LYGEROS, *Stochastic receding horizon control with output feedback and bounded controls*, Automatica, 48 (2012), pp. 77–88.
- [10] S. S. KEERTHI AND E. G. GILBERT, *Optimal infinite-horizon feedback laws for a general class of constrained discrete-time systems: stability and moving-horizon approximations*, Journal of Optimization Theory and Applications, 57 (1988), pp. 265–293.
- [11] D. MAYNE AND J. RAWLINGS, *Model Predictive Control: Theory and Applications*, Nob Hill Publishing LLC, WI, USA, 2009.
- [12] D. Q. MAYNE, J. B. RAWLINGS, C. V. RAO, AND P. O. M. SOKAERT, *Constrained model predictive control: stability and optimality*, Automatica, 36 (2000), pp. 789–814.
- [13] S. P. MEYN AND R. L. TWEEDIE, *Markov Chains and Stochastic Stability*, Cambridge University Press, Cambridge, 2 ed., 2009. With a prologue by Peter W. Glynn.
- [14] R. PEMANTLE AND J. S. ROSENTHAL, *Moment conditions for a sequence with negative drift to be uniformly bounded in L^r* , Stochastic Processes and their Applications, 82 (1999), pp. 143–155.
- [15] D. E. QUEVEDO AND D. NEŠIĆ, *Robust stability of packetized predictive control of nonlinear systems with disturbances and markovian packet losses*, Automatica, 48 (2012), pp. 1803 – 1811.
- [16] F. RAMPONI, D. CHATTERJEE, A. MILIAS-ARGEITIS, P. HOKAYEM, AND J. LYGEROS, *Attaining mean square boundedness of a marginally stable stochastic linear system with a bounded control input*, IEEE Transactions on Automatic Control, 55 (2010), pp. 2414–2418.

SYSTEMS & CONTROL ENGG., 104 SysCON, IIT BOMBAY, POWAI, MUMBAI 400076, INDIA

URL: <http://www.sc.iitb.ac.in/~chatterjee>

E-mail address: dchatter@iitb.ac.in

AUTOMATIC CONTROL LAB., ETL I22, ETH ZÜRICH, PHYSIKSTRASSE 3, 8092 ZÜRICH, SWITZERLAND

URL: <http://control.ee.ethz.ch/~lygeros>

E-mail address: jlygeros@control.ee.ethz.ch