Stochastic MPC With Dynamic Feedback Gain Selection and Discounted Probabilistic Constraints

Shuhao Yan ©, Paul J. Goulart ©, and Mark Cannon ©

Abstract—This article considers linear discrete-time systems with additive disturbances and designs a model predictive control (MPC) law incorporating a dynamic feedback gain to minimize a quadratic cost function subject to a single chance constraint. The feedback gain is selected online, and we provide two selection methods based on minimizing upper bounds on predicted costs. The chance constraint is defined as a discounted sum of violation probabilities on an infinite horizon. By penalizing violation probabilities close to the initial time and assigning violation probabilities in the far future with vanishingly small weights, this form of constraints allows for an MPC law with guarantees of recursive feasibility without a boundedness assumption on the disturbance. A computationally convenient MPC optimization problem is formulated using Chebyshev’s inequality, and we introduce an online constraint-tightening technique to ensure recursive feasibility. The closed-loop system is guaranteed to satisfy the chance constraint and a quadratic stability condition. With dynamic feedback gain selection, the closed-loop cost is reduced and conservativeness of Chebyshev’s inequality is mitigated. Also, a larger feasible set of initial conditions can be obtained. Numerical simulations are given to show these results.

Index Terms—Chance constraints, Chebyshev inequality, dynamic programming, model predictive control (MPC), multiobjective optimization, stochastic convergence.

I. INTRODUCTION

ROBUST control methods for systems with unknown disturbances must take into account worst-case disturbance bounds in order to guarantee satisfaction of hard constraints on system states and control inputs [1]. However, for problems with stochastic disturbances and constraints that are allowed to be violated up to a specified probability, worst-case control strategies can be unnecessarily conservative. This motivated the development of stochastic model predictive control (MPC), which addresses optimal control problems for systems with chance constraints by making use of information on the distribution of model uncertainty [2].

Available methods for approximating chance constraints include analytical approximation and sampling methods. The former aims to provide tractable deterministic optimization problems while the latter generally results in randomized methods. In [3], Cantelli’s inequality is used to turn the chance constraint on states into linear constraints. In [4] and [5], Chebyshev’s inequality is used to reformulate chance constraints as a quadratic and a linear matrix inequality constraint, respectively. These two inequalities can handle a wide range of probability distributions and only require information on the first and second moments of additive disturbance distributions. However, the resulting approximate chance constraints only provide tight bounds for specific probability distributions and otherwise are conservative. A scenario approach is used in [6] to impose time-average expectation constraints on system states. Although constraint satisfaction is demonstrated for the closed-loop system, recursive feasibility of online optimizations are assumed but not ensured. More generally, sample-based methods are unable to ensure recurrence of feasibility of receding horizon optimization problems unless they are combined with robust bounds on model uncertainty. For example, a recursively feasible MPC strategy is proposed in [7], with chance constraints imposed using a scenario approach at the first prediction time step and replaced by robust constraints at later times.

In order to provide guarantees of recursive feasibility and constraint satisfaction in closed-loop operation with reduced conservativeness, online constraint tightening techniques are proposed in [8]–[10]. These methods rely on knowledge of worst-case disturbance bounds, and their degree of conservativeness increases as the disturbance bounds become more conservative. An adaptive approach is developed in [11] which aims to avoid this problem by introducing a scaling factor for tightening parameters which is adapted online, making use of past observations of constraint violations. The authors show that the time-average constraint violation rate converges in probability to a specified limit. For the case that bounds are known on disturbances, hard constraints on control inputs can be incorporated. However, without bounded disturbance distributions or an assumption that the open-loop system is stable (as in [12], for example), it is not possible to guarantee satisfaction of hard
To ensure recursive feasibility under unbounded additive disturbances, previous work (e.g., [3], [13], and [14]) has resorted to a backup initialization of the MPC optimization problem when infeasibility occurs. This design can prevent feedback from the actual state measurement and could thus yield inadequate performance.

For problems involving stochastic uncertainty, the optimal expected value of predicted cost is typically used to perform a Lyapunov analysis of closed-loop stability. A performance metric widely used in this setting is the long-run expected average cost [15]. Although the vast majority of stability results are derived by imposing terminal constraints, there are a number of alternative approaches (e.g., [16]–[18]).

Discounted costs and constraints are present in many stochastic control settings (e.g., [19]–[22]) as well as in reinforcement learning [23], financial engineering (e.g., [24]–[26]), and ecosystem management (e.g., [27] and [28]). Discount factors in optimal control problems allow performance in the near future to be prioritized over long-term behavior. This shift of emphasis is vital for ensuring recursive feasibility of chance-constrained control problems involving possibly unbounded disturbances. In dynamic programming (e.g., [19] and [20]), discounting is commonly employed to ensure that infinite horizon problems with possibly unbounded cost per stage are well-defined. In economics, discounting allows aggregation of current and potential future costs and revenues. For example, [27] shows that varying discount factors on future revenue can affect harvesting policies.

In contrast to existing work on stochastic MPC, this article considers linear discrete-time systems subject to possibly unbounded additive disturbances. We propose an MPC strategy incorporating a dynamic feedback gain to minimize a quadratic cost while satisfying a chance constraint. The constraint combines long-term and short-term considerations by imposing a bound on discounted violation probabilities accumulated over an infinite horizon. The chance constraint is reformulated using Chebyshev’s inequality using knowledge of only the first and second moments of the disturbance input to obtain a convex optimization problem. The main features and contributions of this article are summarized as follows.

1) We use a discount factor to ensure that the chance constraint is well-defined and to prioritize near-future system behavior over steady-state performance.

2) A constraint tightening technique is proposed to ensure recursive feasibility of online MPC optimization problems and constraint satisfaction in closed loop without requiring disturbances to be bounded.

3) The closed-loop system satisfies a quadratic stability condition without the need for terminal constraints.

This article extends preliminary results that appeared in [29] by reducing the conservativeness of Chebyshev’s inequality using a dynamic feedback gain in the definition of the MPC law. This is achieved by solving offline a set of multiobjective optimization problems with the cost and constraint functions considered as conflicting objectives to generate a set of feedback gains providing various tradeoffs between these two objectives. Two methods of online gain selection are proposed, minimizing upper bounds on the optimal predicted cost while retaining guarantees of recursive feasibility and computationally simple MPC optimization. We show that the gain selection procedure can be configured so that the feedback gain converges almost surely to the unconstrained LQ-optimal solution, and the set of admissible initial conditions can be enlarged by choosing an appropriate initial gain. The MPC algorithm significantly improves closed-loop performance in terms of the long-run expected average cost and reduces the conservativeness of closed-loop constraint handling compared to MPC laws based on fixed feedback gains.

The article is organized as follows. The control problem is described and the controller structure is formulated in Section II. Section III proposes an online constraint-tightening method for guaranteeing recursive feasibility. Section IV addresses multiobjective optimization problems, solutions of which provide strictly stabilizing feedback gains. These feedback gains are used to define predicted control sequences via the dynamic feedback gain selection methods proposed in Section V. Section VI summarizes the proposed MPC algorithm and derives a bound on closed-loop performance. In Section VII, the closed-loop behavior of the tightening parameters is analyzed and constraint satisfaction is proved. Section VIII gives a numerical example illustrating the results obtained and the article is concluded in Section IX. Some proofs are given in the Appendix to improve readability.

Notation: The Euclidean norm of a vector $x \in \mathbb{R}^n$ is denoted $\|x\|$ and we define $\|x\|^2_Q := x^\top Q x$. The notations $Q \succ 0$ and $R \succ 0$ indicate that $Q$ and $R$ are symmetric positive semidefinite and symmetric positive definite matrices, respectively, and $\text{tr}(Q)$ denotes the trace of $Q$. A matrix of suitable dimension with all entries being 0 is denoted as $0$, and an identity matrix of a suitable dimension is denoted as $I$. The conditional probability of an event $F$, given the state $x_k$, is denoted as $\mathbb{P}\{F|x_k\} = \mathbb{P}_k\{F\}$, the conditional expectation of $y$, given $x_k$, is $E\{y|x_k\} = E_k\{y\}$, and $\mathbb{P}\{F\}, E\{y\}$ are equivalent to $\mathbb{P}_0\{F\}, E_0\{y\}$ respectively. The sequence $\{x_0, \ldots, x_{N-1}\}$ is denoted as $\{x_i\}_{i=0}^{N-1}$. We denote the value of a variable $x$ at time $k$ as $x_k$, and the $i$-step-ahead predicted value of $x$ at time $k$ is denoted as $x_{i|k}$.

II. PROBLEM DESCRIPTION

Consider an uncertain linear system

$$x_{k+1} = Ax_k + Bu_k + \omega_k$$

where $x_k \in \mathbb{R}^{n_x}$ and $u_k \in \mathbb{R}^{n_u}$ are the system and control input, respectively. The unknown disturbance input $\omega_k \in \mathbb{R}^{n_x}$ is independently and identically distributed (i.i.d.) with known first and second moments

$$E\{\omega_k\} = 0, \quad E\{\omega_k \omega_k^\top\} := \Omega \succ 0.$$  

The disturbance distribution may have infinite support, unlike the approach of [9], which assumes that the additive disturbance lies in a compact set. We assume that the system state is measured directly and is available to the controller at each sample instant.

The system (1) is subject to the constraint

$$\sum_{k=0}^{\infty} \gamma^k \mathbb{P}\{\|Cx_k + Du_k\| \geq 1\} \leq c$$
for given $C \in \mathbb{R}^{n_c \times n_x}$, $D \in \mathbb{R}^{n_c \times n_u}$, a positive scalar $\epsilon$, and a discount factor $\gamma \in (0, 1)$. To simplify presentation, we set $D = 0$ for the remainder of the article, noting that all of the results given in Sections III–VII apply to the case of nonzero $D$, and in the sequel, the constraint

$$\sum_{k=0}^{\infty} \gamma^k \mathbb{P}\{\|Cx_k\| \geq 1\} \leq \epsilon \tag{3}$$

is considered. We refer to $\mathbb{P}\{\|Cx_k\| \geq 1\}$ as a violation probability.

In this article, we design a controller to solve the following problem:

$$\begin{align*}
\min & \mathbb{E} \left\{ \sum_{k=0}^{\infty} \|x_k\|^2_Q + \|u_k\|^2_R \right\} \\
\text{s.t.} & (3).
\end{align*} \tag{4}$$

The weighting matrices in the cost function of problem (4) are assumed to satisfy $Q \succ 0$ and $R \succ 0$.

**Assumption 1:** $(A,B)$ is controllable and $(A,Q^{1/2})$ is observable.

The discounting in this problem introduces a special feature that the probabilities of violating the condition $\|Cx_k\| < 1$ at time instants $k$ nearer to the initial time are weighted more heavily than those in the far future. Therefore, the discount factor allows for a tradeoff between short-term and long-term behaviors and essentially determines how much priority the algorithm gives to constraint violations in the immediate future relative to those in the distant future.

### A. Finite Horizon Formulation and Constraint Handling

Problem (4) employs an infinite horizon in the definition of both the cost and constraint. This optimization problem is computationally intractable since it involves an infinite sequence of decision variables, namely the control inputs $\{u_k\}_{k=0}^{\infty}$. However, the use of an infinite horizon can impart desirable properties, notably stability [1]. It is, therefore, beneficial to formulate a similar problem with a cost function and a constraint that are defined over a finite horizon in such a way that they appropriately approximate the infinite horizon cost and constraint of the original problem. Moreover, the probability distribution of states may be unknown at each time step and the evaluation of violation probabilities is therefore generally intractable. Even if the probability distribution of $\omega_k$ is known explicitly, computing a finite horizon version of (3) requires the solution of a set of multivariate convolution integrals, which is difficult to manage in general [2].

We, therefore, propose to solve problem (4) using a receding horizon approach, wherein our control law is parameterized at each stage with a finite number of decision variables, and constraint (3) is approximated conservatively using the two-sided Chebyshev inequality. The resulting optimization problem to be solved at each time step is then both finite dimensional and computationally tractable.

### B. Predicted Nominal Control Input and State Sequences

Before deriving the finite horizon expressions of the cost and the constraint as mentioned in the previous section, this section defines predicted nominal control input and state sequences.

The sequence of nominal control inputs predicted at time $k$ is given by

$$\tilde{u}_{i|k} = K_k \tilde{x}_{i|k} + c_{i|k}, \quad i = 0, \ldots, N - 1 \tag{5}$$

$$\tilde{u}_{N+i|k} = K_k \tilde{x}_{N+i|k}, \quad i = 0, 1, \ldots \tag{6}$$

where $\tilde{x}_{i|k}$ is the $i$-step-ahead prediction of the nominal state given information at time $k$, that is, $E_k\{x_{i|k}\} = \tilde{x}_{i|k}$. The matrix-valued term $K_k$ is a stabilizing feedback gain that is selected online at time $k$ from amongst a precomputed set of candidates. The offline generation of candidates and the procedure used to select from them online will be detailed in Sections IV and V, respectively. After selecting a gain $K_k$, the perturbation sequence $(c_{0|k}, c_{1|k}, \ldots, c_{N-1|k})$ then constitutes the decision variables in the MPC optimization problem to be solved at time $k$.

Given the predicted nominal control laws (5) and (6), the predicted nominal state trajectory is given by $\tilde{x}_{0|k} = x_k$ and

$$\begin{align*}
\tilde{x}_{i|k} & = \Phi_k^i \tilde{x}_{0|k} + \sum_{j=0}^{i-1} \Phi_k^{i-1-j} B c_{j|k}, \quad i = 1, \ldots, N, \tag{7} \\
\tilde{x}_{N+i|k} & = \Phi_k^i \tilde{x}_{N|k}, \quad i = 1, 2, \ldots \tag{8}
\end{align*}$$

where $\Phi_k := A + BK_k$. The covariance matrix, $X_{i|k}$, of the $i$-step-ahead predicted state is given by $X_{0|k} = 0$ and

$$X_{i|k} = \sum_{j=0}^{i-1} \Phi_k^{i-j} \Omega \left( \Phi_k^j \right)^\top, \quad i = 1, 2, \ldots \tag{9}$$

We rewrite (7) in a compact form as

$$\tilde{x}_k = M_{x}(K_k) \tilde{x}_{0|k} + M_{x}(K_k) c_k, \tag{10}$$

$$\begin{bmatrix} \tilde{x}_{1|k} \\ \vdots \\ \tilde{x}_{N|k} \end{bmatrix} = \begin{bmatrix} \Phi_k \\ \Phi_k^2 \\ \vdots \\ \Phi_k^N \end{bmatrix} \begin{bmatrix} \tilde{x}_{0|k} \\ B \\ \Phi_k B \\ \vdots \end{bmatrix} \begin{bmatrix} c_{0|k} \\ c_{1|k} \\ \vdots \\ c_{N-1|k} \end{bmatrix} \tag{11}$$

and $M_x(K_k) \in \mathbb{R}^{N_{x|x} \times n_x}$ and $M_x(K_k) \in \mathbb{R}^{N_{x|x} \times N_{x|u}}$. For simplicity, we write these two matrices as $M_x$ and $M_c$, with the understanding that they depend on $K_k$.

### C. Online MPC Optimization

Based on predicted sequences defined in Section II-B and employing Chebyshev’s inequality, we give finite horizon expressions of the cost and the constraint and formulate an MPC optimization problem to be solved repeatedly online.

Minimizing the predicted cost $E_k\{\sum_{i=0}^{\infty} \|x_{i|k}\|^2_Q + \|u_{i|k}\|^2_R\}$ at time $k$ over the optimization variable $c_k$ is equivalent to minimizing the cost defined in terms of the predicted nominal
input sequence (5) and (6) and state trajectory (7) and (8) by
\[
J(\tilde{x}_{0|k}, K_k, c_k) := \sum_{i=0}^{N-1} \left( \| \tilde{x}_{i|k} \|_Q^2 + \| K_k \tilde{x}_{i|k} + c_i \|_R^2 \right) + \| \tilde{x}_{N|k} \|_{P_k}^2. \tag{12}
\]
Here, \( \| \tilde{x}_{N|k} \|_{P_k}^2 \) is the terminal cost and \( P_k > 0 \) is chosen as the solution of
\[
P_k = Q + K_k^T r K_k + \Phi_k^T P_k \Phi_k. \tag{13}
\]
Using (10) and (11), we rewrite (12) in a compact form as
\[
J(\tilde{x}_{0|k}, K_k, c_k) = \begin{bmatrix} \tilde{x}_{0|k}^T \\ c_k^T \end{bmatrix} W_2(K_k) \begin{bmatrix} \tilde{x}_{0|k} \\ c_k \end{bmatrix} \tag{14}
\]
where \( W_2(K_k) \) is a function of \( K_k \) and its expression is omitted here for simplicity.

Approximating the left-hand side (LHS) of (3) at time \( k \) by its upper bound derived from a direct application of the two-sided Chebyshev inequality [30, Section 7.3], we obtain
\[
\sum_{i=0}^{N} \gamma_i^j \left( \| C \tilde{x}_{i|k} \|^2 + \text{tr}(C^T C X_{i|k}) \right) \leq \varepsilon_k. \tag{15}
\]

Here, \( \varepsilon_k \) replaces \( \varepsilon \) in (3) as a threshold on the resulting constraint function, and it is a design parameter to be chosen at time \( k \) in some way such that we can ensure recursive feasibility of online MPC optimizations under possibly unbounded disturbances. The advantages of this approach are that it can cope with unknown disturbance probability distributions with known first and second moments, and, furthermore, it results in a quadratic constraint that is straightforward to implement. By the following lemma, we show that the LHS of (15) is equivalent to a finite horizon expression.

**Lemma 1:** Let \( \tilde{P}_k \) be the solution of
\[
\tilde{P}_k = \gamma \Phi_k^T \tilde{P}_k \Phi_k + C^T C. \tag{16}
\]
Then
\[
\sum_{i=0}^{N} \gamma_i^j \| C \tilde{x}_{i|k} \|^2 = \gamma^N \| \tilde{x}_{N|k} \|^2_{\tilde{P}_k},
\]
\[
\sum_{i=0}^{N} \gamma_i^j \text{tr}(C^T C X_{i|k}) = \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k)
\]
where \( \tilde{x}_{i|k} \) is given by (8) for all \( i \geq N \), and \( X_{0|k} = 0 \) and \( X_{i|k} \) is given by (9) for all \( i \geq 1 \).

By Lemma 1, (15) is equivalent to
\[
\sum_{i=0}^{N-1} \gamma_i^j \| C \tilde{x}_{i|k} \|^2 + \gamma^N \| \tilde{x}_{N|k} \|^2_{\tilde{P}_k} + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k) \leq \varepsilon_k \tag{17}
\]
where \( \gamma^N \| \tilde{x}_{N|k} \|^2_{\tilde{P}_k} \) is the terminal term of the infinite discounted sum associated with predicted nominal states and \( \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k) \) is the infinite discounted sum associated with covariance, which remains finite due to the discount factor. Using (10) and (11), we rewrite (17) in a more compact form as
\[
\begin{bmatrix} \tilde{x}_{0|k} \\ c_k \end{bmatrix}^T W_1(K_k) \begin{bmatrix} \tilde{x}_{0|k} \\ c_k \end{bmatrix} + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k) \leq \varepsilon_k \tag{18}
\]
where \( W_1(K_k) \) is a function of \( K_k \) defined as
\[
W_1(K_k) := \begin{bmatrix} C^T C + M_x^T H M_x & M_x^T H M_e \\ M_x^T H M_e & M_e^T \end{bmatrix}
\]
and \( H := \text{diag}(\gamma C^T C, \ldots, \gamma^{N-1} C^T C, \gamma^{N} \tilde{P}_k) \neq 0 \).

To summarize, the MPC optimization solved at time \( k \) is
\[
J^*(x_k, K_k) := \min_{\varepsilon_k} \{ J(x_k, K_k, c_k) \mid (18) \text{ with } \tilde{x}_{0|k} = x_k \}
\]
and its solution for any feasible \( x_k, K_k, \) and \( \varepsilon_k \) is denoted as \( c_k^*(x_k, K_k, \varepsilon_k) \). For simplicity, we write this solution as \( c_k^* \), with the understanding that this vector depends on \( x_k, K_k, \) and \( \varepsilon_k \).

The corresponding predicted nominal state trajectory is given by
\[
x_{i|k}^* = \Phi_k \tilde{x}_i^* + \sum_{j=0}^{i-1} \Phi_k^{i-1-j} B c_j^*, \quad i = 1, \ldots, N, \tag{21}
\]
\[
x_{N+i|k}^* = \Phi_k \tilde{x}_{N+i|k}, \quad i = 1, 2, \ldots \tag{22}
\]
The MPC law at time \( k \) is defined by
\[
u_k := K_k x_k + c_{0|k} \tag{23}
\]
and the closed-loop system dynamics are given by
\[
x_{k+1} = \Phi_k x_k + B c_{0|k}(x_k, K_k, \varepsilon_k) + \omega_k \tag{24}
\]
where \( \omega_k \) is the disturbance realization at time \( k \).

In the remainder of this article, we discuss how to choose \( \varepsilon_k \) so as to guarantee recursive feasibility in Section III; how to generate a set of feedback gains with desirable properties offline in Section IV; how to select \( K_k \) from among this set of candidates with the aim of minimizing predicted costs while retaining the recursive feasibility guarantee of online MPC optimizations in Section V; and how the choices of \( P_k \) and \( \tilde{P}_k \) given in (13) and (16) allow for a guarantee of quadratic stability and satisfy constraint (3), respectively, under the MPC law (23) in Sections VI and VII.

**III. Recursive Feasibility**

Recursively feasible MPC strategies have the property that the MPC optimization problem is guaranteed to be feasible at every time step if it is initially feasible. This property is typically ensured by imposing a terminal constraint that requires the predicted system state to lie in a particular set at the end of the prediction horizon [31]. For a deterministic MPC problem, if an optimal solution can be found at current time, then the *tail sequence*, namely the optimal control sequence shifted by one time step, will be a feasible but suboptimal solution at the next time instant if the terminal constraint is defined in terms of a suitable invariant set for the predicted system state [32]. For a robust MPC problem with bounded additive disturbances, recursive feasibility can likewise be guaranteed by imposing a terminal constraint set that is robustly invariant. However, this approach is not generally applicable to systems with unbounded additive disturbances, and, in general, it is not possible to ensure recursive feasibility in this context while guaranteeing constraint satisfaction at every time instant.

In this section, we propose a method for guaranteeing recursive feasibility of the MPC optimization, which does not rely...
on terminal constraints. Instead, recursive feasibility is ensured, despite the presence of unbounded disturbances, by allowing the constraint on the discounted sum of violation probabilities to be time-varying. For every time step \( k > 0 \), the approach uses the optimal sequence computed at time \( k - 1 \) to determine a value of \( \varepsilon_k \) that is necessarily feasible at time \( k \).

We use the notation \( \bar{c}_{k+1} \) to denote a time-shifted version of \( c_k \), defined by
\[
\bar{c}_{k+1} := Ec_k^* \tag{25}
\]
where \( E \) is the matrix such that \( Ec = [c_1^T, \ldots, c_{N-1}^T, 0]^T \) if \( c = [c_0^T, \ldots, c_{N-1}^T]^T \).

Lemma 2: Given initial feasibility at time \( k = 0 \), the MPC optimization (20) is recursively feasible if \( \varepsilon_k \) is defined at each time \( k = 1, 2, \ldots \) as
\[
\varepsilon_k := \left[ x_k^T \right] W_1 \left( K_{k-1} \right) \left[ x_k \right] + \frac{\gamma}{1 - \gamma} \text{tr}(\Omega P_{k-1}) . \tag{26}
\]
Equation (26) provides an explicit expression for \( \varepsilon_k \) for all \( k > 0 \) in terms of (21) and (22) and disturbance realizations as
\[
\varepsilon_k = \sum_{i=0}^{N-1} \gamma^i \| C(\bar{x}_{i+1|k-1} + \Phi_{k-1}^i \omega_{k-1}) \|^2
+ \gamma^N \| \bar{x}_{N+1|k-1} + \Phi_{k-1}^N \omega_{k-1} \|^2 P_{k-1}^{-1} + \frac{\gamma}{1 - \gamma} \text{tr}(\Omega P_{k-1}). \tag{27}
\]
Here
\[
\bar{x}_{i+1|k-1} + \Phi_{k-1}^i \omega_{k-1} := \bar{x}_{i|k}, \quad i = 0, \ldots, N \tag{28}
\]
defines a feasible nominal state sequence predicted at time \( k \), obtained by setting \( \bar{x}_{0|k} = x_k \) and \( \bar{u}_{i|k} = K_{k-1} \bar{x}_{i|k} + c_{i|k} \). Note that in constructing this feasible sequence, we still use \( K_{k-1} \) and therefore \( P_{k-1} \) as the corresponding terminal matrix and that the feedback gain is updated after the update of \( \varepsilon_k \).

Essentially, the optimization problem to be solved at each time step is feasible because the parameter \( \varepsilon_k \) is updated via (27) using knowledge of the disturbance \( w_{k-1} \) obtained from the measurement of the current state \( x_k \). In this respect, the approach is similar to constraint-tightening methods that have previously been applied in the context of stochastic MPC (e.g., [8]–[10]) in order to ensure recursive feasibility and constraint satisfaction in closed-loop operation. However, each of these methods requires that the disturbances affecting the controlled system be bounded, and they become more conservative as the degree of conservativeness of the assumed disturbance bounds increases. The approach proposed here avoids this requirement and instead ensures closed-loop constraint satisfaction using the analysis of sequence \( \{\varepsilon_k\} \) as will be detailed in Section VII.

The key to this method lies in the definition of the vector \( \bar{c}_{k+1} \). If this vector were optimized, rather than defined by the suboptimal control input (25), then it would be possible to reduce the MPC cost (20). However, this would require more computational effort than is needed to evaluate (27) and lose the guarantee of satisfying (3) in closed loop.

IV. MULTIOBJECTIVE OPTIMIZATION AND DYNAMIC PROGRAMMING

In this section, we generate a set of strictly stabilizing feedback gains offline, from which \( \bar{K} \) in (5) and (6) is selected online. Considering the cost function in (4) and the constraint (3) as conflicting objectives, we formulate a set of multiobjective optimization problems whose solutions provide feedback gains representing a tradeoff between these two objectives. These multiobjective optimization problems are written using linear scalarization [33] for given \( x_0 \in \mathbb{R}^{n_x} \) in the form
\[
\min_{u_0, u_1, \ldots} \gamma^0 \mathbb{E} \left\{ \| Cx_1 \|^2 \right\} + \mu \sum_{i=0}^{\infty} \mathbb{E} \left\{ \| x_i \|^2_Q + \| u_i \|^2_R \right\} \tag{P_1}
\]
s.t. \( x_{i+1} = Ax_i + Bu_i + \omega_i, \quad i = 0, 1, \ldots \)

where \( \mu \in (0, 1] \) is a weighting parameter. Here, \( \omega_i \) is an i.i.d. random variable with the same statistics as given in (2). Note that in (P_1), the first part of the objective is an approximation of the LHS of the constraint (3) via Chebyshev’s inequality and the second part has the same form as the cost function in problem (4). Possible tradeoffs between these competing objectives can be explored using different values of \( \mu \), and their solutions provide a set of strictly stabilizing feedback gains. We denote this set of feedback gains as \( \bar{K} \).

We propose to solve (P_1) by dynamic programming [34] via its equivalent deterministic counterpart in the form of
\[
\min_{\hat{u}_0, \hat{u}_1, \ldots} \gamma^0 \mathbb{E} \left\{ \| C\hat{x}_1 \|^2 \right\} + \mu \sum_{i=0}^{\infty} \left( \| \hat{x}_i \|^2_Q + \| \hat{u}_i \|^2_R \right) \tag{P_2}
\]
s.t. \( \hat{x}_{i+1} = A\hat{x}_i + B\hat{u}_i, \quad i = 0, 1, \ldots \)

where \( \hat{x}_i \) and \( \hat{u}_i \) are mean values of \( x_i \) and \( u_i \), respectively, with \( \hat{x}_0 = x_0 \). From the quadratic form of the objective function, the optimal solution to (P_2) is a linear feedback control law, and its corresponding feedback gain has desirable properties as will be shown in Theorem 4. In the first instance, we show the equivalence between (P_1) and (P_2) in Theorem 3, which is in agreement with the certainty equivalence principle [35, Section 8.4].

Theorem 3: Problems (P_1) and (P_2) are equivalent in the sense that their optimal solutions are linear feedback control laws with the same feedback gains.

Proof: We begin by deriving the solution of (P_1). Let
\[
(1 - \mu) \hat{J}_0(x_0) + \mu \hat{J}_0(x_0) :=
\lim_{T \to \infty, \omega_0, \ldots, \omega_{T-1}} \min_{u_0, u_1, \ldots, u_{T-1}} E \left\{ (1 - \mu) \sum_{i=0}^{T-1} \gamma^i \mathbb{E} \{ \| Cx_i \|^2 \} + \mu \sum_{i=0}^{T-1} (\| x_i \|^2_Q + \| u_i \|^2_R) \right\}
\]
s.t. (29)

where \( \hat{J}_0(x_0) \) and \( \hat{J}_0(x_0) \) are optimal cost-to-go functions denoting the optimal values of \( \lim_{T \to \infty} \sum_{i=0}^{T-1} \gamma^i \mathbb{E} \{ \| Cx_i \|^2 \} \) and \( \lim_{T \to \infty} \sum_{i=0}^{T-1} \mathbb{E} \{ \| x_i \|^2_Q + \| u_i \|^2_R \} \), respectively, for the initial state \( x_0 \). Here, \( u_i, i \geq 0 \), and \( x_i, i \geq 1 \) are decision variables.
corresponding to the control inputs and states of system (29). By solving the Bellman equations [34] backwards in time with initial conditions \( \bar{J}_T(x_T) = 0 \) and \( J_T(x_T) = 0 \) where \( \bar{J}_T(x_T), J_T(x_T) \) are the optimal cost-to-go functions starting from state \( x_T \), it can be shown that \( \bar{J}_0(x_0) = x_0^\top \bar{S}_0 x_0 + r_0, J_0(x_0) = x_0^\top S_0 x_0 + v_0, i = 0 \) represent discounted and undiscounted sums of variances of \( x_i \) for \( i = 1, \ldots, T \). Due to the quadratic form of the objective function and the assumption of zero-mean disturbance \( \omega_i \), the optimal control input \( u_i \) is a linear function of \( x_i \) for \( i \geq 0 \). Denoting the optimal feedback gain as \( H_i \), we obtain the following DP iteration for \( i = T - 1, \ldots, 0 \): \( \hat{S}_I = C^\top C + \gamma (A + BH_i) \top \hat{S}_{i+1} (A + BH_i), \)
\( \hat{S}_i = Q + H_i \top RH_i + (A + BH_i) \top \hat{S}_{i+1} (A + BH_i), \)
\( r_i = \gamma (r_{i+1} + \text{tr}(\hat{S}_{i+1} \Omega)), \)
\( v_i = \text{tr}(\hat{S}_{i+1} \Omega + v_{i+1}) \)
where we choose \( \hat{S}_T = \hat{S}_0 = 0 \) and \( r_T = v_T = 0 \). Also, the corresponding feedback gains \( H_i \) can be computed by backward induction as \( H_i = -[\mu R + B^\top \Delta_{i+1} B]^{-1} B^\top \Delta_{i+1} A \) for \( i = T - 1, \ldots, 0 \) with \( \Delta_{i+1} = (1 - \mu) \hat{S}_{i+1} + \mu \hat{S}_i, \) Therefore, to obtain \( H_i, \hat{S}_i, \) and \( \hat{S}_i, \) it is only necessary to perform the iterations in (31) and (32) since these are independent of \( r_i \) and \( v_i \). By considering the limit as \( T \to \infty \) and using similar reasoning, it can be shown that \( x_0^\top (1 - \mu) \hat{S}_0 + \mu \hat{S}_0 x_0 \) is the optimal objective of problem \( (P2) \) and the corresponding optimal control law is given for all \( i \geq 0 \) by \( u_i = H_i x_i \).

Theorem 3 demonstrates that we only need to solve problem \( (P2) \) to obtain the optimal feedback gains for \( (P1) \). To suit our purposes, we reverse the time indexing in (31) and (32) and define \( P_i := \hat{S}_{T-i}, P_i := \hat{S}_{T+i} \) for \( i = 0, \ldots, T \) and \( L_i := H_{T-i} \) for \( i = 1, \ldots, T \). The resulting DP iteration is \( L_{i+1} = -[\mu R + B^\top \Sigma_i B]^{-1} B^\top \Sigma_i A, \)
\( \bar{P}_{i+1} = C^\top C + \gamma (A + BL_{i+1}) \top \bar{P}_i (A + BL_{i+1}), \)
\( \bar{P}_{i+1} = Q + L_{i+1} RL_{i+1} + (A + BL_{i+1}) \top \bar{P}_i (A + BL_{i+1}) \)
with \( \Sigma_i := \gamma (1 - \mu) \bar{P}_i + \mu \bar{P}_i, i = 0, \ldots, T - 1, \) These provide the solution of \( (P2) \) in the limit as \( T \to \infty \) and the optimal value of the objective function of problem \( (P2) \) is \( x_0^\top (1 - \mu) \bar{P}_0 + \mu \bar{P}_0 x_0 \) for any given \( x_0 \in \mathbb{R}^n, \) where \( \bar{P}_\infty \) and \( \bar{P}_\infty \) are defined in the following theorem.

**Theorem 4:** Consider (35)–(37). Under Assumption 1, we have the following.

1) There exist matrices \( \bar{P}_\infty \geq 0, \bar{P}_\infty > 0 \) such that for arbitrary positive semidefinite matrices \( P_0 \) and \( \bar{P}_0, \) we have \( \lim_{i \to \infty} \bar{P}_i = \bar{P}_\infty \) and \( \lim_{i \to \infty} \bar{P}_i = \bar{P}_\infty \).

2) For all \( i \in (0, 1], \) the feedback gain \( L_\infty \) is strictly stabilizing, where \( L_\infty := -[\mu R + B^\top \Sigma_\infty B]^{-1} B^\top \Sigma_\infty A \)
with \( \Sigma_\infty := \gamma (1 - \mu) \bar{P}_\infty + \mu \bar{P}_\infty \).

3) The matrices \( \hat{P}_\infty \) and \( \bar{P}_\infty \) satisfy the following Lyapunov equations:
\( \hat{P}_\infty = C^\top C + \gamma (A + BL_\infty) \top \bar{P}_\infty (A + BL_\infty) \),
\( \bar{P}_\infty = Q + L_\infty R L_\infty + (A + BL_\infty) \top \bar{P}_\infty (A + BL_\infty) \).

**Proof:** These results can be derived using well-known DP techniques, which can be found, for example, in [20].

The solutions \( L_\infty, \bar{P}_\infty, \) and \( \bar{P}_\infty \) are functions of \( \mu \) and are denoted \( L_\infty(\mu), \bar{P}_\infty(\mu), \) and \( \bar{P}_\infty(\mu) \) in the remainder of this article. With \( \mu = 1, L_\infty(1) = K_{11}Q, \) where \( K_{11}Q \) is the LQ-optimal feedback gain that minimizes the second part of the objective of \( (P2) \), whereas with \( \mu = 0 \) \( L_\infty(0) = -(B_\infty \bar{P}_\infty(0) B_\infty^\top B_\infty \bar{P}_\infty(0)) A, \) which is optimal with respect to the first part of the objective. However, the gain \( L_\infty(0) \) may not be stabilizing and unique due to discounting and no penalty on control inputs if \( \mu = 0, \) and we, therefore, restrict the weighting parameter \( \mu \) to the interval \( (0, 1). \)

**Remark 5:** The fixed point \( (L_\infty(\mu), \bar{P}_\infty(\mu), \bar{P}_\infty(\mu)) \) of the iterations (35)–(37) coincides with the minimizing argument of
\( \min_{G \in \mathbb{R}^{n \times n}, Z_1 \geq 0, Z_2 \geq 0} \text{tr}((1 - \mu)Z_1 + \mu Z_2) \)
\text{s.t.} \( Z_1 = C^\top C + \gamma (A + BG)^\top Z_1 (A + BG), \)
\( Z_2 = Q + G^\top RG + (A + BG)^\top Z_2 (A + BG). \)

However, the iteration is generally preferred over solving this equivalent problem directly since it is nonconvex in variables \( G, Z_1, \) and \( Z_2, \) with no obvious convexifying transformation.

Next, we give procedures, which are executed offline, to generate feedback gains based on a sequence of positive weighting parameters for gain selection methods in Section V.

**K Generation** (Offline):

1) choose appropriately a sequence \( \{\mu_i\}_{i=1}^m, \) with \( \mu_i \in (0, 1] \) in ascending order, and \( m \in \mathbb{N}; \)
2) for each \( i = 1, \ldots, m, \) solve problem \( (P2) \) with \( \mu = \mu_i \) by executing iterations (35)–(37);
3) obtain the set of strictly stabilizing feedback gains \( K := \{L_\infty(\mu_i)\}_{i=1}^m. \)

In the \( K \) generation, there should be a sufficiently large number of elements in the sequence \( \{\mu_i\}_{i=1}^m \) so that \( L_\infty \) can be adequately approximated on the intervals between consecutive points in this sequence. More importantly, \( \mu_i \) should be appropriately chosen close to 0 while ensuring that \( L_\infty(\mu_i) \neq L_\infty(0) \) and \( \bar{P}_\infty(\mu_i) \neq \bar{P}_\infty(0) \) if the latter is strictly stabilizing, and, hence, \( L_\infty(\mu_i) \neq L_\infty(0) \forall \mu_i \geq \mu_i \) by the uniqueness of solutions to problem \( (P2) \) and the monotonicity property of \( P_\infty(\mu) \) as will be given in the next section. Furthermore, step 2) can be warm-started by using \( P_\infty(\mu_i) \) and \( \bar{P}_\infty(\mu_i) \) to initialize the iteration with weighting parameter \( \mu_{i+1}, \) thereby reducing considerably the time required to solve \( (P2) \) for each value of \( \mu. \)
Note also that in step 3), the sets \( \{P_\infty(\mu_i)\}_{i=1}^m \) and \( \{\bar{P}_\infty(\mu_i)\}_{i=1}^m \) are obtained.
V. DYNAMIC FEEDBACK GAIN SELECTION

In this section, we provide two methods for dynamically selecting feedback gains from the set $K_k$ discussed in Section IV. These methods are designed such that the recursive feasibility guarantee of (20) and a computationally simple online optimization are retained. Both methods determine feedback gains that minimize upper bounds on the optimal predicted cost (20) and exploit monotonicity of certain functions. Method 1 requires less online computation and is equivalent to a binary search, while guaranteeing almost sure asymptotic convergence of $K_k$ to $K_{LO}$. Method 2 is more intuitive and, in many cases, gives better closed-loop performance over short time horizons but requires slightly more online computation since it requires the online evaluation of a function of $\mu \in \{\mu_i\}_{i=1}^m$.

We first derive properties of $\bar{P}_\infty(\mu)$ and $\hat{P}_\infty(\mu)$ that are exploited by both methods, namely that $\bar{P}_\infty(\mu)$ and $\hat{P}_\infty(\mu)$ are monotonic in $\mu$ and that $(1 - \mu)\bar{P}_\infty(\mu) + \mu\hat{P}_\infty(\mu)$ is concave for $\mu \in (0, 1)$, which implies Lipschitz continuity of $\bar{P}_\infty(\cdot)$ on $(0, 1)$.

**Lemma 6:** For all $\mu_1$ and $\mu_2$ such that $0 \leq \mu_1 \leq \mu_2 \leq 1$, $\bar{P}_\infty(\cdot)$ and $\hat{P}_\infty(\cdot)$ satisfy

$$\bar{P}_\infty(\mu_1) \leq \bar{P}_\infty(\mu_2),$$

$$\hat{P}_\infty(\mu_1) \geq \hat{P}_\infty(\mu_2).$$

**Lemma 7:** Let $S(\mu) := (1 - \mu)\bar{P}_\infty(\mu) + \mu\hat{P}_\infty(\mu)$, then $S(\cdot)$ is concave on $(0, 1)$.

**Lemma 8:** $\bar{P}_\infty(\cdot)$ and $\hat{P}_\infty(\cdot)$ are Lipschitz continuous on $(0, 1)$.

A. Gain Selection Method 1

This section describes a method for selecting the gain $K_k$ online and determines the properties of the sequence $\{\bar{\mu}_k\}_{k=0}^\infty$ generated by the following procedure.

**Method 1:** At each time step $k = 1, 2, \ldots$

1) compute

$$c^\alpha(x_k) := \arg \min_c \left[ x_k \right] \left[ \begin{array}{c} E \{c \} \\ c \end{array} \right] \cdot W_1(K_{k-1}) \cdot \left[ \begin{array}{c} x_k \\ c \end{array} \right]$$

(41)

where $W_1(K_{k-1})$ is defined in (19);

2) compute

$$\bar{\mu}_k := \max_{\mu \in \{\mu_i\}_{i=1}^m} \left( \frac{\gamma}{1 - \gamma} \right) \text{tr}(\bar{P}_\infty(\mu))$$

s.t.

$$\left[ \begin{array}{c} x_k \\ c^\alpha(x_k) \end{array} \right] \left[ \begin{array}{c} W_1(K_{k-1}) \cdot x_k \\ c^\alpha(x_k) \end{array} \right] + \frac{\gamma}{1 - \gamma} \text{tr}(\bar{P}_\infty(\mu)) \leq \varepsilon_k$$

(42)

where $\varepsilon_k$ and $\bar{P}_\infty(\mu)$ are defined in (26) and (39), respectively;

3) set $K_k := L_{\infty}(\bar{\mu}_k), \bar{P}_k := \bar{P}_\infty(\bar{\mu}_k)$, and $P_k := \hat{P}_\infty(\bar{\mu}_k)$. Step 1 determines the perturbation sequence $c^\alpha(x_k)$ that minimizes the LHS of constraint (18) with $K_k = K_{k-1}$ and $\bar{P}_k = \bar{P}_{k-1}$. Step 2) then chooses $\bar{\mu}_k$ as the largest element of the sequence $\{\mu_i\}_{i=1}^m$ such that $c^\alpha(x_k)$ is feasible for constraint (18). To show this, note that by combining (25), (26), and step 3), we obtain

$$\varepsilon_k = \left[ \begin{array}{c} x_k \\ Ec^\alpha(k-1) \end{array} \right] \left[ \begin{array}{c} W_1(L_{\infty}(\bar{\mu}_{k-1})) \cdot x_k \\ Ec^\alpha(k-1) \end{array} \right]$$

$$+ \frac{\gamma}{1 - \gamma} \text{tr}(\bar{P}_\infty(\bar{\mu}_{k-1})).$$

From Lemma 6, (42) is, therefore, equivalent to

$$\varepsilon_k \leq \left[ \begin{array}{c} x_k \\ Ec^\alpha(k-1) \end{array} \right] \left[ \begin{array}{c} W_1(K_{k-1}) \cdot x_k \\ Ec^\alpha(k-1) \end{array} \right]$$

$$+ \frac{\gamma}{1 - \gamma} \text{tr}(\Omega \bar{P}_\infty(k_{k-1})) \leq \left[ \begin{array}{c} x_k \\ Ec^\alpha(k-1) \end{array} \right] \left[ \begin{array}{c} x_k \\ Ec^\alpha(k-1) \end{array} \right]$$

(43a)

s.t.

$$\frac{\gamma}{1 - \gamma} \text{tr}(\Omega \bar{P}_\infty(k_{k-1})) \leq \varepsilon_k$$

(43b)

where $\ast$ denotes $W_1(L_{\infty}(\bar{\mu}_{k-1}))$ and the right-hand side (RHS) of (43b) is nonnegative due to the definition of $c^\alpha(x_k)$ in step 1). The aim of Method 1 is, therefore, to use the slack introduced into the constraint through the choice of $c^\alpha(x_k)$ in step 1) in order to maximize $\bar{\mu}_k$ in step 2) subject to

$$\left[ \begin{array}{c} x_k \\ c^\alpha(x_k) \end{array} \right] \left[ \begin{array}{c} W_1(K_{k-1}) \cdot x_k \\ c^\alpha(x_k) \end{array} \right] = \left[ \begin{array}{c} x_k \\ c^\alpha(x_k) \end{array} \right] \left[ \begin{array}{c} W_1(K_k) \\ c^\alpha(x_k) \end{array} \right]$$

(44)

so that problem (20) remains feasible at time $k$. Therefore, Method 1 retains the recursive feasibility guarantee. From (44), it also follows that $J(x_k, K_k, c^\alpha)$ is an upper bound on the optimal value of the cost in (20) at time $k$, and from (43a) and Lemma 6 it follows that $\bar{\mu}_k$ defined in step 2) minimizes the trace of the terminal matrix $P_k = \bar{P}_\infty(\bar{\mu}_k)$ in this cost.

By Lemma 6, step 2) is simply a binary search to determine the largest $\bar{\mu}_k \in \{\mu_i\}_{i=1}^m$ satisfying the constraint in (42). Since the values of $\text{tr}(\Omega \bar{P}_\infty(\mu_i))$ can be calculated offline, Method 1 can be implemented very efficiently. Note that the initial value $\mu_0$ is not determined by Method 1, and we, therefore, use other means of choosing $\mu_0 \in \{\mu_i\}_{i=1}^m$ and $K_0 = L_{\infty}(\mu_0)$ to make the MPC optimization (20) initially feasible (assuming such a $\mu_0$ exists); this is discussed in Section VIII, Simulation D.

In the remainder of this section, we analyze the properties of the sequence $\{\bar{\mu}_k\}_{k=0}^\infty$ generated by Method 1. Since $\bar{\mu}_{k-1}$ is a feasible solution to problem (42) at time $k$, which implies $\bar{\mu}_k \leq \bar{\mu}_{k-1}$, and since $\bar{\mu}_{k}$ is upper bounded by 1 for all $k$, the sequence $\{\bar{\mu}_k\}_{k=0}^\infty$ generated by Method 1 is monotonically nondecreasing and convergent. To derive a stronger convergence result, we make a simplifying assumption.

**Assumption 2:** The maximization of (42) can be solved with the optimization variable $\mu$ varying continuously in the interval $[\mu_0, 1]$ rather than being constrained to the finite set $\{\mu_i\}_{i=1}^m$.

The purpose of Assumption 2 is to ensure that $\bar{\mu}_k \geq \bar{\mu}_{k-1}$ whenever the RHS of (43b) is positive and hence that $\bar{\mu}_k$ does
not converge to a value less than 1 as a result of the constraint
\( \mu \in \{ \mu_i \}_{i=1}^\infty \) in step 2). To analyze asymptotic convergence of
\( \{ \tilde{\mu}_k \}_{k=0}^\infty \), we consider the RHS of (43b). By definition,
\( W_1(K_{k-1}) = W_1(L_\infty(\tilde{\mu}_k-1)) \) is at least positive semidefinite; so the
minimization in step 1) is well-defined and \( c^\alpha(x_k) \) can be
defined uniquely as
\[
c^\alpha(x_k) = - (W_{cc}(\tilde{\mu}_k-1)) W_{cx}(\tilde{\mu}_k-1)x_k.
\]
Here, \( W_{cc}(\mu) \) and \( W_{cx}(\mu) \) are blocks of \( W_1(L_\infty(\mu)) \) in
the partition (19) such that \( W_{cc} = M_c^T HM_c, W_{cx} = M_c^T HM_x, \) and
\( W_{cc} \) is the pseudo-inverse of \( W_{cc} \). Since \( \bar{x}_k = \bar{x}_{k-1} + \omega_{k-1} \), it follows that
the RHS of (43b) is equal to
\[
\left\| z_{k-1} + (W_{cc}(\tilde{\mu}_k-1)) W_{cx}(\tilde{\mu}_k-1) x_{k-1} \right\|^2_{W_{cc}(\bar{\mu}_k-1)}
\]
where \( z_{k-1} = E x_{k-1} + (W_{cc}(\tilde{\mu}_k-1)) W_{cx}(\tilde{\mu}_k-1) x_{k-1}^* \).

Lemma 9: There exist \( \delta > 0 \) and \( p_\delta > 0 \) such that
\[
\inf_{\mu \in [\mu_0,1]} \| x_{k} + (W_{cc}(\mu)) W_{cx}(\mu) x_{k} \|_{W_{cc}(\mu)}^2 \geq \delta \geq p_\delta.
\]
We conclude that \( K_k \) converges almost surely to \( K_{LQ} \).

**Theorem 10:** Under Assumption 2, \( \tilde{\mu}_k \to 1 \) as \( k \to \infty \) with
probability 1.

### B. Gain Selection Method 2

We first give a monotonic result that can be used to prove the
order-preserving property of Riccati difference equations (RDEs) [36], which is exploited in Method 2.

**Lemma 11:** \( V(x_0, \bar{P}_1) \leq V(x_0, \bar{P}_2) \) \( \forall x_0 \in \mathbb{R}^{n_x} \) if \( \bar{P}_1 \leq \bar{P}_2 \), where
\[
V(x_0, \bar{P}) := \min_{u_0, \ldots, u_{N-1}} \{ \bar{V}(x_0, u, \bar{P}) \}
\]  
\[
s.t. x_{i+1} = A x_i + B u_i, i = 0, \ldots, N - 1 \}
\]
and \( \bar{V}(x_0, u, \bar{P}) := \sum_{i=0}^{N-1} \| x_i \|_Q^2 + \| u_i \|_R^2 \) + \( \| x_N \|_{\bar{P}}^2, u := [u_0, \ldots, u_{N-1}] \) for \( Q, R, \bar{P} \geq 0 \).

**Proof:** Let \( u_0^*(\bar{P}) := \arg\min_{u_0} \{ \bar{V}(x_0, u, \bar{P}) \} \) s.t. \( x_{i+1} = A x_i + B u_i, i = 0, \ldots, N - 1 \), then by the ordering
between \( \bar{P}_1 \) and \( \bar{P}_2 \) and optimality, we have
\( V(x_0, \bar{P}_1) = \bar{V}(x_0, u_0^*(\bar{P}_2), \bar{P}_2) \geq \bar{V}(x_0, u_0^*(\bar{P}_2), \bar{P}_1) \geq \bar{V}(x_0, u_0^*(\bar{P}_1), \bar{P}_1) = V(x_0, \bar{P}_1) \).

**Remark 12:** Note that Lemma 11 holds if \( Q \) and \( R \) are
time-variant. Let \( \bar{P}_1^0 \) and \( \bar{P}_2^0 \) be the matrices satisfying \( V(x_0, \bar{P}_1) = x_0^T \bar{P}_1^0 x_0 \) and \( V(x_0, \bar{P}_2) = x_0^T \bar{P}_2^0 x_0 \) for all \( x_0 \in \mathbb{R}^{n_x} \), then
Lemma 11 implies \( \bar{P}_1^0 \leq \bar{P}_2^0 \). Therefore, the monotonic property
of RDE can be shown by solving (47) using DP recursion
with \( N = 1 \) sequentially backwards in time. This property was
proved in [37] and [38] with the assumption that \( B \) has full rank,
which is not required here.

We give Method 2 as follows.

**Method 2:** At each time step \( k = 1, 2, \ldots \)

1) for all \( \mu \in (0,1] \), define
\[
\mathcal{C}(x_k, \mu) := \arg\min_{c} \left[ \begin{array}{c} x_k \\ c \end{array} \right]^T W_1(L_\infty(\mu)) \left[ \begin{array}{c} x_k \\ c \end{array} \right] \]
where \( W_1 \) is defined in (19);

2) compute
\[
\hat{\mu} := \arg \max_{\mu \in (0,1]} \left[ \begin{array}{c} x_k \\ c \end{array} \right]^T W_1(L_\infty(\mu)) \left[ \begin{array}{c} x_k \\ c \end{array} \right] + \gamma^k \]
\[
s.t. \left[ \begin{array}{c} x_k \\ c \end{array} \right]^T W_1(L_\infty(\mu)) \left[ \begin{array}{c} x_k \\ c \end{array} \right] + \gamma^k \leq \epsilon_k \}
\]
(49)

where \( \epsilon \) here denotes \( \epsilon_k \) defined in (14);

3) set
\[
\tilde{\mu}_k := \arg \min_{\mu \in (0,1]} \left[ \begin{array}{c} x_k \\ c \end{array} \right]^T W_2(L_\infty(\mu)) \left[ \begin{array}{c} x_k \\ c \end{array} \right] \]
\[
s.t. \left[ \begin{array}{c} x_k \\ c \end{array} \right]^T W_2(L_\infty(\mu)) \left[ \begin{array}{c} x_k \\ c \end{array} \right] + \gamma^k \leq \epsilon_k \}
\]
where \( W_2 \) is defined in (14); and \( \tilde{\mu}_k \) is a function of both \( x_k \) and \( \mu \).

Comparing step 1) of Methods 1 and 2, here, \( L_\infty(\mu) \) replaces \( K_k-1 \) and \( \mathcal{C}(x_k, \mu) \) is a function of both \( x_k \) and \( \mu \).
Recalling the equivalence between (17) and (18), we have that
\[
\left[ \begin{array}{c} x_k \\ c \end{array} \right]^T W_1(L_\infty(\mu)) \left[ \begin{array}{c} x_k \\ c \end{array} \right] = \sum_{i=0}^{N-1} \gamma^i \| C \bar{x}_i \|^2 + \gamma^N \| x_N \|_{\bar{P}_\infty(\mu)}^2
\]
where \( \bar{x}_i \) evolves according to \( \bar{x}_k = M_c(L_\infty(\mu)) \) and \( M_c(L_\infty(\mu)) \).

Therefore, by Lemmas 6 and 11, the objective in the maximization (49) is monotonic in \( \mu \) and step 2) can be performed as a binary search. The constraint in (49) makes the constraint (18) as tight as possible with \( K_k = L_\infty(\mu) \), and the constraints of (50) imply that (20) is recursively feasible with Method 2. Note that the initial value \( \tilde{\mu}_0 \) is not determined by Method 2. Therefore, we choose \( \tilde{\mu}_0 \in \{ \mu_i \}_{i=1}^\infty \) and \( K_0 = L_\infty(\tilde{\mu}_0) \) to make the MPC optimization (20) initially feasible (if such \( \tilde{\mu}_0 \) exists), as discussed in Section VIII, Simulation D.

Comparing with Method 1, both terms appearing on the LHS of the constraint in problem (49) are monotonic in \( \mu \), whereas only the second term on the LHS of the constraint in problem (42) increases with \( \mu \) and the first term is fixed. It follows that \( \mu_k \) obtained by Method 1 is necessarily greater than or equal to \( \tilde{\mu} \) defined in (49) and hence also necessarily greater than or equal to \( \mu_k \) obtained by Method 2. The sequence \( \{ \mu_k \}_{k=0}^\infty \) generated by Method 1 is therefore likely to converge more quickly to 1. On the other hand, problem (50) results in a smaller upper bound on the optimal predicted cost (20), and, hence, Method 2 is likely to provide better closed-loop performance over the time period required for the sequence \( \{ K_k \}_{k=0}^\infty \) determined by Method 1 to converge to \( K_{LQ} \). Over a longer time interval, however, Method 1 is likely to perform better due to its earlier convergence to the LQ-optimal feedback gain \( K_{LQ} \). These observations are supported by the numerical example in Section VIII.

Based on this discussion, we provide the following guideline for choosing the online gain selection method. If the discount factor \( \gamma \) in (3) is close to 1, indicating that performance over
Algorithm 1: (SMPC Algorithm)

At each time-step $k = 0, 1, \ldots$:
1. measure $x_k$;
2. if $k > 0$, compute $\varepsilon_k$
   using (26) and determine $K_k$ using Method 1 or 2;
3. solve the quadratically constrained quadratic programming (QCQP) problem (20);
4. apply the control law (23).

a long horizon is a priority, then Method 1 should be chosen. Alternatively, if $\gamma$ is close to 0 and performance in the immediate future is important, then Method 2 is preferable. Finally, we note that Method 2 requires more computation than Method 1 since (50) (the objective of which is not necessarily monotonic in $\mu$) requires an additional search over $\{\mu_i\}_{i=1}^m$.

Remark 13: Since $\mu_{k-1}$ is a feasible solution to the optimization problem (50) at time $k$ and $\mu_{k}$ is upper bounded by 1 for all $k$, the sequence $\{\mu_k\}_{k=0}^{\infty}$ generated by Method 2 is monotonically nondecreasing and convergent.

Several factors influence the implementation of online gain selection methods. The choice of the sequence $\{\mu_i\}_{i=1}^m$ can affect convergence of $\{\mu_k\}_{k=0}^{\infty}$ and hence closed-loop nonmonotone since large gaps between successive elements of the sequence $\{\mu_i\}_{i=1}^m$ reduce the likelihood of convergence $\mu_k \to 1$. Therefore, it is desirable to choose $m$ to be as large as possible, subject to offline computation and online storage constraints. To ensure a large feasible set and to steer the closed-loop system away from feedback gains $K_k$ that give a worse closed-loop performance, it is desirable to choose $\mu_1$ close to 0 with $\mu_{i+1} - \mu_i$ increasing for larger values of $i$. Alternatively, if the initial conditions of the MPC problem are known offline when the set $\mathcal{K}$ is generated, then it is obviously advantageous to set $\mu_1$ equal to a value $\mu_0$ that makes the MPC optimization initially feasible.

VI. SMPC ALGORITHM AND STABILITY CONDITION

For deterministic MPC, it can be shown by using the tail sequence that, with an appropriate terminal weighting matrix [1], optimal MPC predicted costs are monotonically nonincreasing along closed-loop trajectories. This property does not generally hold in the presence of unbounded disturbances, and, in fact, the optimal MPC predicted cost defined by (20) is not necessarily monotonically nonincreasing if $\varepsilon_k$ is defined by (26). However, in this section, we show that the proposed approach based on (25) ensures a closed-loop stability bound. We first state the online MPC algorithm based on the optimization problem defined in (20).

We choose $\varepsilon_0 = e$ in step 2) as will be explained in Section VII. In step 3), the MPC optimization can be solved efficiently since there is only one quadratic constraint in (20), for example, using a second-order conic program solver or using the algorithm proposed in [39], which is based on the Newton–Raphson method.

Theorem 14: Given initial feasibility of problem (20) at time $k = 0$, by using Algorithm 1, the closed-loop system satisfies the quadratic stability condition
\[
\lim_{T \to \infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\{\|x_k\|^2 + \|u_k\|^2\} \leq \lim_{T \to \infty} \frac{1}{T} \sum_{k=0}^{T-1} \text{tr}(\Omega \mathbb{E}\{P_k\}) \leq \text{tr}(\Omega P_0) \quad (51)
\]
provided $P_k$ satisfies (13).

VII. BEHAVIOR OF THE SEQUENCE $\{\varepsilon_k\}_{k=0}^{\infty}$ AND CONSTRAINT SATISFACTION

This section considers the properties of the sequence $\{\varepsilon_k\}_{k=0}^{\infty}$ in closed-loop operation under Algorithm 1. We first derive a recurrence relation relating the expected value of $\varepsilon_{k+1}$ to $x_k$ and $\varepsilon_k$ by using the explicit expression for $\varepsilon_k$ in (27). This allows an upper bound to be determined for the sum of discounted violation probabilities on the LHS of (3). Then with this bound, we show that the closed-loop system under the control law of Algorithm 1 satisfies the chance constraint (3) if $\varepsilon_k$ is initialized with $\varepsilon_0 = e$.

The following result gives the relationship between $\varepsilon_k$ and the expected value of $\varepsilon_{k+1}$ for the closed-loop system.

Lemma 15: Given initial feasibility of problem (20) at time $k = 0$, if $\varepsilon_k$ is defined by (26) at all times $k \geq 1$ and $\tilde{P}_k$ satisfies (16) for all $k \geq 0$, then in closed-loop operation under Algorithm 1, we have
\[
\gamma \mathbb{E}\{\varepsilon_{k+1}\} \leq \varepsilon_k - \|C x_k\|^2, \quad \forall k \geq 0. \quad (52)
\]

The main result of this section is given next.

Theorem 16: The closed-loop system under Algorithm 1 satisfies the constraint (3) if $\varepsilon_0 = e$ and problem (20) is initially feasible at time $k = 0$.

Proof: The proof of Theorem 16 is similar to the proof of [29, Th. 6] and therefore is omitted here.

The presence of the discount factor $\gamma \in (0, 1)$ on the LHS of (52) implies that the expected value of $\varepsilon_k$ can increase as well as decrease. In fact, for values of $\gamma$ close to zero, a rapid initial growth in $\varepsilon_k$ is to be expected, which is in agreement with the interpretation that constraint (3) penalizes violation probabilities to a much lesser extent from the initial time in this case. On the other hand, for values of $\gamma$ close to 1, $\varepsilon_k$ can be expected to decrease initially, implying a greater emphasis on the expected number of violations over some initial horizon.

VIII. NUMERICAL EXAMPLE

This section describes a numerical example to illustrate the following points.
1) The closed-loop system (24) satisfies the quadratic stability condition (51) and the constraint (3) when Algorithm 1 is used without dynamic gain selection or with either gain selection Method 1 or Method 2.

2) The degree of conservativeness of Chebyshev’s inequality is mitigated by using online gain selection procedures in the sense that the long-run expected average costs are improved and the observed constraint violation rates are closer to the imposed limit.

3) Gain selection Method 2 provides a better closed-loop performance over short time intervals than Method 1.

4) $K_k$ converges to $K_{1QO}$ with high probability for Method 1.

5) The feasible set of initial conditions is enlarged with an appropriate initial feedback gain.

We also conduct simulations to compare Algorithm 1 with the existing method proposed in [6] and discuss the computation times required by Algorithm 1 and the gain selection methods.

We consider the discrete-time linearized model derived from the continuous-time model of a coupled-tank system [41] with a sampling interval of 0.05 s. This has model parameters:

$$A = \begin{bmatrix} 0.8207 & 0.04 \\ 0.0799 & 0.7808 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0454 & 0.0011 \\ 0.0022 & 0.0443 \end{bmatrix}$$

and disturbance $w_k$ has a multivariate Laplace distribution with zero mean and covariance $\Omega = I$. The constraint (3) is defined by $\gamma = 0.9$, $e = 1.5$, and $C = \begin{bmatrix} 0.3 & 0.15 \\ 0.1 & -0.1 \end{bmatrix}$, and the weighting matrices in the cost of problem (4) are given by $Q = R = I$.

We choose a prediction horizon $N = 10$, and the initial value for $\varepsilon_k$ is $\varepsilon_0 = e = 1.5$. We appropriately choose a sequence $(\mu_i)_{i=1}^m$ with $m = 290320$, where $0 < \mu_i \leq 1$ for all $i$ and $\mu_m = 1$. Using this sequence, we generate the set of feedback gains $K = \{L_\infty(\mu_i)\}_{i=1}^m$ offline and store the sets of matrices $\{P_\infty(\mu_i)\}_{i=1}^m$, $\{\hat{P}_\infty(\mu_i)\}_{i=1}^m$. We choose $\mu_0 = 10^{-15}$ so that

$$K_0 = \begin{bmatrix} 18.0749 & -0.4626 \\ -0.9251 & -17.6123 \end{bmatrix}$$

and $\Phi_0 = A + BK_0$ is strictly stable. Simulations A-D all use the initial condition $x_0 = [-1.3]^{\top}$ and the same sequences of disturbances. Note that $K_k = K_{1QO}$ is infeasible for the online MPC optimization (20) at time $k = 0$.

**Simulation A [demonstrating 1) and 2)]:** To estimate empirically the average cost, denoted as $J_{average}$, and the discounted sum of violation probabilities, denoted as $P_{violation}$, we run $10^4$ simulations, each of which has a length of $40^4$ time steps, using Algorithm 1 with fixed $K_k = K_0 \forall k \geq 0$, with gain selection Method 1 and with gain selection Method 2, respectively. We compute the mean value of stage costs over these $10^4$ simulations and count the number of violations in each simulation up to 150 time steps. Simulation results are summarized in Table I.

**Table I**

| $J_{average}$ | $P_{violation}$ |
|---------------|-----------------|
| $K_k = K_0$ | Method 1 | Method 2 |
| $J_{average}$ | 639.9 | 84 | 96.2 | $tr(\Omega P_\delta) = 640.0$ |
| $P_{violation}$ | 0.448 | 0.914 | 0.769 | $e = 1.5$ |

Table I confirms that the three empirical cost estimates agree with the bound (51) and the three estimates for the discounted sum of violation probabilities are all smaller than $e = 1.5$, implying that constraint (3) is satisfied. Note that the cost estimates decrease more slowly as the simulation horizon length continues to increase and that $\gamma^{150} = 1.37 \times 10^{-7}$; so the number of violations occurring at time steps $k > 150$ has a negligible effect on the estimate of $P_{violation}$. Also, cost estimates obtained using the dynamic gain selection methods are considerably smaller than that obtained using a fixed feedback gain. Estimates for $P_{violation}$ are much closer to the maximum allowed level $e = 1.5$ when gain selection methods are used. More specifically, the estimate obtained using gain selection Method 1 is more than double the estimate obtained with a fixed feedback gain. Hence, these results show that the conservativeness of Chebyshev’s inequality is mitigated by dynamic gain selection. Moreover, the cost estimate obtained using Method 1 is smaller than that obtained using Method 2 since Method 1 achieves convergence of $\mu_k$ to 1 (and, hence, convergence of $K_k$ to $K_{1QO}$) earlier. This observation supports the statement in Section V that Method 1 is likely to provide better closed-loop performance over longer time periods. Note that although gain selection methods are employed, there are discrepancies between constraint violation rates observed and the limit imposed. There are four factors that may account for this: 1) Chebyshev’s inequality is used to handle the chance constraint, bringing substantial conservativeness; 2) constraint-tightening is used to ensure recursive feasibility; 3) $e$ is an upper bound on the discounted sum of violation probabilities and there may exist solutions yielding small rates of violation; 4) the gain selection methods are designed to minimize upper bounds on optimal predicted costs and their impact on increasing constraint violation may be limited.

**Simulation B [demonstrating 3)]:** We run $10^4$ simulations, each of which has a length of $40$ time steps, using Algorithm 1 with fixed $K_k = K_0 \forall k \geq 0$, with gain selection Method 1 and with gain selection Method 2, respectively. We compute the mean value of stage costs over these $10^4$ simulations, as shown in Table II.

**Table II**

| $J_{average}$ | $P_{violation}$ |
|---------------|-----------------|
| $K_k = K_0$ | Method 1 | Method 2 |
| $J_{average}$ | 807.2 | 575.9 | 500.8 |
| $P_{violation}$ | 0.448 | 0.914 | 0.769 | $e = 1.5$ |

It is clear from Table II that Method 2 gives a smaller average cost estimate, which is in agreement with the statement in Section V that Method 2 is preferable if short-term performance is prioritized. This is because $\mu_k$ generated by Method 1 generally does not converge to 1 within 40 time steps in this set of simulations; so the average costs are dominated by transient behaviors. We show an example of sequences $(\mu_k)_{k=0}^{\infty}$ generated by Methods 1 and 2 in Fig. 1, where Fig. 1(b) plots the two sequences up to time $k = 150$ and Fig. 1(a) zooms in on the initial 40 time steps. These two figures show that $\mu_k$ obtained from Method 1 is greater than or equal to that obtained from Method 2 at all times, which is in agreement with the analysis in Section V.
Simulation C [demonstrating 4]): We run $10^4$ simulations, each of which has a length of 200 time steps. When Algorithm 1 is used with Method 1, $\bar{\mu}_{200} = 1$ is obtained for every simulation, implying that $K_k$ converges to $K_{LQ}$ in every simulation. On the other hand, when Method 2 is used, the mean value of $\bar{\mu}_{200}$ over the set of $10^4$ simulations is 0.8578.

Simulation D [demonstrating 5]): Minimizing the LHS of (18) over $c_k$ yields the largest set of feasible initial conditions $X_0 = \{x : x^T(W_{xx}(\mu) - W_{xc}(\mu)W_{cx}(\mu))x + \frac{\gamma}{1 - \gamma}\text{tr}(\bar{P}_\infty(\mu)) \leq \varepsilon_0\}$

where $W_{xx}(\mu)$, $W_{xc}(\mu)$, $W_{cx}(\mu)$ are the blocks of $W_1(L_\infty(\mu))$ in (19). The feasible set $X_0$ is plotted for $\mu$ taking values of $\mu_1 = 10^{-15}$, $\mu_2 = 10^{-4}$, and $\mu_3 = 2.5 \times 10^{-4}$ in Fig. 2. Clearly, the feasible set is enlarged as $\mu$ is reduced.

Simulation E: To compare the performance of Algorithm 1 with dynamic gain selection with the performance of the scenario-based MPC algorithm in [6], we run 1000 simulations for each algorithm and each simulation has a length of 200 time steps. The system and problem parameters remain the same as initially specified at the beginning of this section, except $N = 5$, $Q = \begin{bmatrix} 0 & 0 \\ 0 & 0.5 \end{bmatrix}$, and $R = \begin{bmatrix} 1 & 0 \\ 0 & 0.01 \end{bmatrix}$. To implement Algorithm 1 with dynamic gain selection, $\{L_\infty(\mu_i)\}_{i=1}^m$, $\{\bar{P}_\infty(\mu_i)\}_{i=1}^m$ and $\{\hat{P}_\infty(\mu_i)\}_{i=1}^m$ are computed offline using the sequence $\{\mu_i\}_{i=1}^m$, and $\mu_0$ is chosen such that problem (20) is feasible at time $k = 0$. To implement the algorithm in [6], we choose a violation level $\varepsilon = 16.67\%$, which, together with the support rank $\rho_1 = 2$, yields a scenario number $K = 11$, and a finite horizon scenario program (FHSCP) is solved at each time step, where the affine disturbance feedback (ADF) policy is used to parameterize control inputs and the FHSCP has $NK (=55)$ constraints on predicted states. Then this algorithm ensures that

$$\sum_{k=0}^{200} \gamma^k P \{\|C_{x_k}\| \geq 1\} \leq \varepsilon$$

is satisfied in closed loop. We compute the mean values of stage costs over these 1000 simulations and count the number of violations in each simulation. Simulation results are summarized in Table III.

Table III shows that although Algorithm 1 with dynamic gain selection yields much lower constraint violation rates comparing with the algorithm in [6], Algorithm 1 with gain selection Method 1 achieves a better closed-loop cost. Also, the implementation of Algorithm 1 has lower computational complexity.

Computation times: Simulations are run in MATLAB R2019a on a computer with 2.20-GHz Intel Core i7-8750H CPU and 16-GB RAM, and the online MPC optimization (20) is solved using the root-finding algorithm proposed in [39]. Each online MPC optimization is solved within 1 ms, with an average time of 0.35 ms. Gain selection Method 1 requires less than 0.003 ms at each iteration of Algorithm 1, which represents a tiny fraction of the time needed to solve an MPC optimization.
Gain selection Method 2 requires less than 0.05 ms at most time steps, which remains a small fraction of the time needed to solve an MPC optimization. The maximum time observed for implementation of gain selection Method 2 is 70 ms. After \( \hat{\mu}_k \) has converged to 1, no gain selection method is executed; so no additional online computation other than the MPC optimization is required.

**IX. CONCLUSION**

A stochastic MPC algorithm is proposed that enforces a constraint on the sum of discounted future violation probabilities and ensures recursive feasibility of the online optimization and closed-loop constraint satisfaction. Key features are the design of a constraint-tightening procedure and dynamic gain selection methods and closed-loop analysis of the tightening parameters. The MPC algorithm requires knowledge of the first and second moments of the disturbance and is implemented as a convex QCQP problem. Stability and constraint satisfaction are ensured for the closed-loop system without assuming disturbance bounds. The results in Sections III, VI, and VII of this article on recursive feasibility of the MPC optimization, closed-loop stability and constraint satisfaction can be extended to problems with multiple chance constraints, although solving the corresponding multiobjective optimization problems to generate stabilizing feedback gains for dynamic gain selection becomes more challenging in this case.

**APPENDIX A**

**A. Proof of Lemma 1**

Equation (16) implies that \( \tilde{P}_k = \sum_{i=0}^{\infty} \gamma^i (\Phi_k)^i C^T C \Phi_k^i \).

Combining this with (8) shows that

\[
\sum_{i=N}^{\infty} \gamma^i \|C x_{i+k}\|^2 = \gamma^N x_{N+k}^T \left( \sum_{j=0}^{\infty} \left( \frac{\gamma}{2} \Phi_k^j \right)^j C^T C \left( \frac{\gamma}{2} \Phi_k^j \right)^j \right) x_{N+k} = \gamma^N \|x_{N+k}\|^2 / P_k.
\]

Furthermore, let \( \tilde{S}_k = \sum_{i=0}^{\infty} \gamma^i X_{i+k} \), and then (9) implies

\[
\gamma \Phi_k \tilde{S}_k \Phi_k^T = \sum_{i=0}^{\infty} \gamma^{i+1} \Phi_k X_{i+k} \Phi_k^T = \sum_{i=0}^{\infty} \gamma^{i+1} (X_{i+1+k} - \Omega)
\]

and \( \tilde{S}_k \) satisfies the Lyapunov equation

\[
\tilde{S}_k = \gamma \Phi_k \tilde{S}_k \Phi_k^T + \frac{\gamma}{1 - \gamma} \Omega.
\]

Therefore

\[
\text{tr} \left( \sum_{i=0}^{\infty} \gamma^i C^T C X_{i+k} \right) = \text{tr} \left( C^T C \tilde{S}_k \right) = \text{tr} \left( C^T C \Phi_k \left( \frac{\gamma}{1 - \gamma} \Omega \right) \Phi_k^T \right) = \frac{\gamma}{1 - \gamma} \text{tr} \left( \Omega P_k \right).
\]

**B. Proof of Lemma 2**

If \( K_k = K_{k-1} \) at time \( k \geq 1 \), then \( P_k = \tilde{P}_{k-1} \). From (18), (26), and \( x_{0|k} = x_k \), it follows that \( \tilde{c}_k \) is a feasible solution to problem (20) at time \( k \). On the other hand, if the chosen gain selection method gives \( K_k \neq K_{k-1} \) and hence \( P_k \neq \tilde{P}_{k-1} \), we can show that the guarantee of recursive feasibility is retained as follows. If gain selection Method 1 is used, \( \tilde{\mu}_k \) defined in step 2) of Method 1 satisfies the constraint in problem (42), which, together with (44), implies a feasible solution to problem (20) exists at time \( k \). If gain selection Method 2 is used, \( \mu_k \) defined in step 3) of Method 2 satisfies \( \mu_k \leq \tilde{\mu} \) and satisfies the constraint in problem (49) due to monotonicity results in Lemmas 6 and 11. This implies a feasible solution to problem (20) exists at time \( k \).

**C. Proof of Lemma 6**

In problem \((P_2)\), the optimality of \( x_0^* \left( (1 - \mu) \tilde{P}_\infty (\mu) + \mu \tilde{P}_\infty (\mu) \right) x_0 \) for all \( x_0 \in \mathbb{R}^n \) implies

\[
(1 - \mu_1) \tilde{P}_\infty (\mu_1) + \mu_1 \tilde{P}_\infty (\mu_2) - \mu_2 \tilde{P}_\infty (\mu_1),
\]

\[
(1 - \mu_2) \tilde{P}_\infty (\mu_2) + \mu_2 \tilde{P}_\infty (\mu_1) - \mu_1 \tilde{P}_\infty (\mu_2),
\]

Defining \( \Delta \tilde{P} \), \( \Delta \tilde{P} \) by \( \Delta \tilde{P} := \tilde{P}_\infty (\mu_2) - \tilde{P}_\infty (\mu_1) \), we have

\[
(1 - \mu_1) \Delta \tilde{P} \lesssim \mu_1 \Delta \tilde{P},
\]

\[
(1 - \mu_2) \Delta \tilde{P} \lesssim -\mu_2 \Delta \tilde{P}. \tag{53}
\]

Combining these inequalities yields \( \Delta \tilde{P} \gtrless 0 \) and \( \Delta \tilde{P} \lesssim 0 \).

**D. Proof of Lemma 7**

Let \( 0 < \mu_1 \leq \mu_2 \leq \mu_3 \leq 1 \) with \( \mu_3 - \mu_2 = \mu_2 - \mu_1 \), and let \( \Delta^2 \tilde{P} = \tilde{P}_\infty (\mu_3) - 2 \tilde{P}_\infty (\mu_2) + \tilde{P}_\infty (\mu_1) \), \( \Delta^2 \tilde{P} = \tilde{P}_\infty (\mu_3) - 2 \tilde{P}_\infty (\mu_2) + \tilde{P}_\infty (\mu_1) \), and \( \Delta^2 S = S (\mu_3) - 2 S (\mu_2) + S (\mu_1) \). Then by definition, we have \( \Delta^2 S = (1 - \mu_3) \Delta^2 \tilde{P} + \mu_3 \Delta^2 \tilde{P} - (\mu_3 - \mu_1) (\Delta \tilde{P} - \Delta \tilde{P}) \), where \( \Delta \tilde{P} = \tilde{P}_\infty (\mu_2) - \tilde{P}_\infty (\mu_1) \). But the optimality property of \( \text{tr} (ZS (\mu)) \) for all \( Z = Z^T \gtrless 0 \) implies

\[
(1 - \mu_3) \tilde{P}_\infty (\mu_2) + \mu_3 \tilde{P}_\infty (\mu_1) \lesssim (1 - \mu_3) \tilde{P}_\infty (\mu_2) + \mu_3 \tilde{P}_\infty (\mu_1),
\]

\[
(1 - \mu_3) \tilde{P}_\infty (\mu_3) - \tilde{P}_\infty (\mu_2),
\]

and \( \tilde{S}_k \) satisfies the Lyapunov equation

\[
\tilde{S}_k = \gamma \Phi_k \tilde{S}_k \Phi_k^T + \frac{\gamma}{1 - \gamma} \Omega.
\]

Therefore

\[
\text{tr} \left( \sum_{i=0}^{\infty} \gamma^i C^T C X_{i+k} \right) = \text{tr} \left( C^T C \tilde{S}_k \right) = \text{tr} \left( C^T C \Phi_k \left( \frac{\gamma}{1 - \gamma} \Omega \right) \Phi_k^T \right) = \frac{\gamma}{1 - \gamma} \text{tr} \left( \Omega P_k \right).
\]

**E. Proof of Lemma 8**

Let \( \Sigma (\mu) := \gamma (1 - \mu) \tilde{P}_\infty (\mu) + \mu \tilde{P}_\infty (\mu) \), and \( \Phi (\mu) := A + BL \tilde{P}_\infty (\mu) \) and consider the effect of an infinitesimal change \( \delta \mu \) in the value of \( \mu \in (0,1) \), where \( \delta \mu \ll 1 - \gamma \) and \( \delta \mu \ll 1 - \mu \). Let \( \Delta \Phi, \Delta \Sigma, \Delta L, \Delta \tilde{P}, \) and \( \Delta \tilde{P} \) denote the corresponding changes in the values of \( \Phi, \Sigma, L, \tilde{P}_\infty, \) and \( \tilde{P}_\infty \), respectively, so that \( \Delta \Phi = \Phi (\mu + \delta \mu) - \Phi (\mu), \Delta \Sigma = \Sigma (\mu + \delta \mu) - \Sigma (\mu) \).
\[ \Delta L = L_\infty(\mu + \delta \mu) - L_\infty(\mu), \quad \Delta \hat{P} = \hat{P}_\infty(\mu + \delta \mu) - \hat{P}_\infty(\mu), \] and \[ \Delta \hat{P} = \hat{P}_\infty(\mu + \delta \mu) - \hat{P}_\infty(\mu). \] To simplify notation, we omit the argument \( \mu \) for the remainder of the proof (so that \( \Phi = \Phi(\mu), \Sigma = \Sigma(\mu), L_\infty = L_\infty(\mu), \) etc.). Lemma 7 implies that \((1 - \mu)\hat{P}_\infty + \mu \hat{P}_\infty\) is Lipschitz continuous on \((0, 1)\); so \((1 - \mu)\Delta \hat{P} + \mu \Delta \hat{P} - \delta(\mu) = 0 = (\Omega(\mu))\] and \[ O(\delta(\mu)) = (1 - \mu)\Delta \hat{P} + \mu \Delta \hat{P} = (\Phi + \Delta \Phi)\Sigma(\Sigma + \Delta \Sigma)\Phi + \Delta \Phi - \Phi^T \Sigma \Phi + \mu(L_\infty + \Delta L) \geq - \mu L_\infty \geq L_\infty \geq L_\infty - \Gamma^{-1} B^T \Delta \Sigma \Phi + RL_\infty \delta(\mu) \]

But \(L_\infty + \Delta L = \frac{[\gamma(\mu + \delta \mu) + B^T(\Sigma + \Delta \Sigma)B]^{-1} B^T(\Sigma + \Delta \Sigma)A}{\mu} \) implies \[ L_\infty + \Delta L = L_\infty - \Gamma^{-1} B^T \Delta \Sigma \Phi + RL_\infty \delta(\mu) \]

where \(\Gamma = \gamma(\mu + \delta \mu) + B^T(\Sigma + \Delta \Sigma)B, \) and \[ \Phi + \Delta \Phi = \Phi - B \Gamma^{-1} B^T \Delta \Sigma \Phi + O(\delta(\mu)). \]

Combining (55) and (56) with (54) and using \(L_\infty = -(\mu R + B^T \Sigma B)^{-1} B^T \Sigma A, \) we obtain \[ O(\delta(\mu)) = \Phi^T \Delta \Sigma \Phi + \Phi^T \Delta \Sigma \Phi \Gamma^{-1} \Delta \Sigma \Phi + \mu R - 2 \Gamma \Gamma^{-1} B^T \Delta \Sigma \Phi \]

and hence \[ \Phi^T \Delta \Sigma \Phi = \Phi^T \Delta \Sigma \Phi \Gamma^{-1} B^T \Delta \Sigma \Phi = O(\delta(\mu)). \]

But \(\Sigma = (1 - \mu)\hat{P}_\infty + \mu \hat{P}_\infty - (1 - \gamma)(1 - \mu)\hat{P}_\infty \) \(\Delta \Sigma = O(\delta(\mu))\) \((1 - \Gamma^{-1} B^T \Delta \Sigma \Phi; \) so \((1 - \gamma)(1 - \mu)\Phi^T \Delta \hat{P} + (1 - \gamma)(1 - \mu)\Delta \hat{P} \Gamma^{-1} B^T \Delta \hat{P} \Phi = O(\delta(\mu)). \]

This implies \(\Phi^T \Delta \hat{P} \Phi = O(\delta(\mu)) \) since \(\delta \mu < 1 - \gamma, \) \(\delta \mu < 1 - \mu, \) and \(\Gamma \geq 0, \) and since \(\Delta \hat{P} \geq 0 \) by Lemma 6. Furthermore, \(\Delta \hat{P} \) is symmetric; so we must have \(\Delta \hat{P} \Phi = O(\delta(\mu)) \) and hence \(\Delta \Sigma \Phi = O(\delta(\mu)). \) From (55) and (56), it follows that \(\Delta L = O(\delta(\mu)) \) and \(\Phi = O(\delta(\mu)), \) and, hence, the solutions \(\hat{P}_\infty(\cdot), \hat{P}_\infty(\cdot) \) of the Lyapunov equations (39) and (40) are Lipschitz continuous on \((0, 1). \)

\[ F. \text{ Proof of Lemma 9} \]

We prove this lemma by contradiction. Suppose that a pair of \(\delta > 0 \) and \(p_\delta > 0 \) does not exist such that (46) holds. In this case, for some \(z \in \mathbb{R}^{N_{\mu}}, \) \(\mu \in [\bar{\mu}, 0], \) we must have \[ P \left\{ \left\| z + (W_{cx}(\mu))^T W_{cx}(\mu) \right\| \geq \delta \right\} < p_\delta \] for all \(\delta > 0 \) and \(\mu \in [\bar{\mu}, 0], \) which implies that \(W_{cx}(\mu) \omega_k + W_{cx}(\mu)z = 0 \) with probability 1. Here \(W_{cx}(\mu) \) must be nonzero since otherwise \(c^0(x_k) = 0, \) implying that, by its definition in (41), the feedback gain used is optimal with respect to \(\sum_{i=0}^{\infty} \gamma^i \|C_{x_i} \|_2^2 \) and then \(L_\infty(\mu) = L_\infty(0) \) would be obtained, which contradicts the fact that \(L_\infty(\mu) \geq L_\infty(\mu) \) for all \(\mu \geq \mu_1 \) in Section IV. Furthermore, \(z \) and \(\mu \) are, by assumption, independent of the realization \(\omega_k. \)

Therefore, (57) contradicts the assumption in (2) that \(\mathbb{E}\{\omega_k \omega_k^T\} \) is positive definite, implying \(\omega_k \) is not a constant vector or in any subspace of \(\mathbb{R}^{N_{\mu}}, \) and, hence, (46) must hold for some \(\delta > 0 \) and \(p_\delta > 0 \).

\[ G. \text{ Proof of Theorem 10} \]

We first show convergence of \(\mu_k \) to 1 in probability (i.e., \(\lim_{k \to \infty} P\{\mu_k = 1\} = 1\)) by splitting an infinite horizon into intervals of \(N_f \) time steps and providing an upper bound on \(P\{\mu_k < 1\}, \) which is parameterized by \(p_\delta, N_f, \) and \(k, \) and which converges to 0 as \(k \to \infty. \) Let \(N_f \) be defined in terms of a \(\delta > 0 \) satisfying (46) for some \(p_\delta > 0 \) by

\[ N_f = \left\lfloor \frac{\gamma}{1 - \gamma} \frac{\text{tr} (\Omega \hat{P}_\infty(1) - \Omega \hat{P}_\infty(\mu_0))}{\delta} \right\rfloor \]

where it is assumed that \(\mu_0 < 1, \) and let \(E_k \) denote the event that

\[ \|z_k - 1 + (W_{cx}(\mu_k - 1))^T W_{cx}(\mu_k - 1) \omega_k - 1 \|_{W_{cx},(\mu_k - 1)} \geq \delta. \]

Under Assumption 2, Lemma 8 implies that before \(\mu_k \) reaches 1, (43b) is satisfied with equality and every possible increment in \(\mu_k \) is attained if (45) is possible. It follows that \(\mu_k \mid_{N_f} = 1 \) if \(E_k \) occurs for \(k = 1, \ldots, N_f. \) Furthermore

\[ P \{E_1 \cap \cdots \cap E_{N_f} \} \geq \prod \{E_1 \cap \cdots \cap E_{N_f - 1} \} \]

where \(P\{E_1 \cap \cdots \cap E_{N_f - 1} \} \geq \prod \{E_1 \cap \cdots \cap E_{N_f - 1} \}\)

and, more specifically, we rewrite it as

\[ P\{\mu_k < 1 | \mu_k < 1\} \leq 1 - p_\delta \]

Similarly, given \(\mu_{(i-1)N_f} < 1 \) for \(i = 2, \ldots, \) we have

\[ P\{\mu_k \mid_{N_f} = 1 \} \geq \left(\inf_{z \in \mathbb{R}^{N_{\mu}, \mu_k} \in [\bar{\mu}, 0]} P\{\star \geq \delta(\mu)\} \right)^{N_f} \]

\[ \geq \left(\inf_{z \in \mathbb{R}^{N_{\mu}, \mu_k} \in [\bar{\mu}, 0]} P\{\star \geq \delta(\mu)\} \right)^{N_f} \geq p_\delta \]

where * here denotes \(\|z + (W_{cx}(\mu))^T W_{cx}(\mu) \omega_k - 1 \|_{W_{cx}(\mu)} \), \(\delta_1 \) is the counterpart of \(\delta \) for \(\mu_k \) that is sufficient for \(\mu_k \) to reach 1 within \(N_f \) time steps from time \(k = (i - 1)N_f \) and is smaller than \(\delta \) since the sequence \(\{\mu_k\} \) is monotonically nondecreasing. Then

\[ P\{\mu_k < 1 | \mu_{(i-1)N_f} < 1\} \leq 1 - p_\delta \]

for \(k = 2, \ldots, \)
For given $k > N_f$, we choose integers $j_1, j_2, \ldots, j_{|k/N_f|}$ so that $j_1 = N_f, j_{i/N_f} \leq k$ and $j_{i+1} - j_i = N_f$ for all $i$. Then $\tilde{\mu}_k < 1$ only if $\tilde{\mu}_{j_i} < 1$ for all $i = 1, \ldots, |k/N_f|$ and hence

$$
\Pr\{\tilde{\mu}_k < 1\} 
\leq \Pr\{\tilde{\mu}_{j_1} < 1\} \Pr\{\tilde{\mu}_{j_2} < 1\} \cdots \Pr\{\tilde{\mu}_{j_{|k/N_f|-1}} < 1\} 
\leq (1 - p_{\delta}^{N_f})^{|k/N_f|}. \tag{58}
$$

It follows that $\Pr\{\tilde{\mu}_k = 1\} \geq 1 - (1 - p_{\delta}^{N_f})^{|k/N_f|}$, which implies $\lim_{k \to \infty} \Pr\{\tilde{\mu}_k = 1\} = 1$. To complete the proof, we use the Borel–Cantelli lemma to show the almost sure convergence of $\tilde{\mu}_k$ to 1, that is, $\Pr\{\lim_{k \to \infty} \tilde{\mu}_k = 1\} = 1$. Let $\mathcal{F}_k$ denote the event that $\tilde{\mu}_k < 1$, and then (58) ensures that

$$
\sum_{k=1}^{\infty} \Pr\{\mathcal{F}_k\} \leq \sum_{k=1}^{\infty} (1 - p_{\delta}^{N_f})^{|k/N_f|} = N_f p_{\delta}^{N_f} - 1 < \infty.
$$

Therefore, the Borel–Cantelli lemma implies that $\Pr\{\bigcap_{k=1}^{\infty} \bigcup_{\ell=k}^{\infty} \mathcal{F}_\ell\} = 0$. But $\tilde{\mu}_k < 1$ only if $\tilde{\mu}_{k-1} < 1$; so $\mathcal{F}_k \subseteq \mathcal{F}_{k-1}$ and $\bigcup_{k=1}^{\infty} \mathcal{F}_k = \mathcal{F}_k$. Therefore, $\Pr\{\lim_{k \to \infty} \mathcal{F}_k\} = 0$ or, equivalently, $\Pr\{\lim_{k \to \infty} \tilde{\mu}_k = 1\} = 1$.

### H. Proof of Theorem 14

Given initial feasibility at time $k = 0$, step 2) of Algorithm 1 ensures that problem (20) is always feasible for time $k = 1, 2, \ldots$. FromLemma 2, the vector $\bar{c}_{k+1}$ provides a feasible but possibly suboptimal solution at time $k + 1$. Hence, by optimality, we necessarily have

$$
J^*(x_{k+1}, K_{k+1}) \leq J(x_{k+1}, K, \bar{c}_{k+1})
$$

and since this inequality holds for every realization of $\omega_k$, by taking expectations conditioned on the state $x_k$, we obtain

$$
\mathbb{E}_k\{J^*(x_{k+1}, K_{k+1})\} \leq \mathbb{E}_k\{J(x_{k+1}, K, \bar{c}_{k+1})\}. \tag{59}
$$

From (28), we have the feasible sequence

$$
\bar{x}_{i|k+1} = \bar{x}_{i|k+1} + \Phi_k^i \omega_k, \quad i = 0, \ldots, N
$$

and from (13) and (59), it follows that

$$
\mathbb{E}_k\{J^*(x_{k+1}, K_{k+1})\} \leq J^*(x_k, K_k) - \|x_k\|^2_Q - \|u_k\|^2_R + \operatorname{tr}(\Omega P_k). \tag{60}
$$

Summing both sides of this inequality over $k \geq 0$ after taking expectations, given information available at time $k = 0$, and making use of the property that $\mathbb{E}_0\{\mathbb{E}_k\{J^*(x_{k+1}, K_{k+1})\}\} = \mathbb{E}_0\{J^*(x_{k+1}, K_{k+1})\}$ give the first inequality of (51). Note that $K_k$ and $P_k$ depend on $x_k$, which implies that they are random variables, but they are uncorrelated with $\omega_k$. Moreover, the second inequality in (51) follows from the choice of $K_k$ in step 2) of Algorithm 1 and Lemma 6.

### I. Proof of Lemma 15

Since (26) is equivalent to (27), expanding the terms in (27) yields

$$
\varepsilon_{k+1} = \sum_{i=0}^{N-1} \gamma^i \|C \bar{x}^*_{i|k+1}\|^2 + \gamma^N \|x_N^{i|k+1}\|^2_{\bar{P}_k} + \frac{\gamma}{1 - \gamma} \operatorname{tr}(\bar{P}_k) \tag{61}
$$

where $\bar{x}^*_{i|k}$ is given by (21) and (22) and $\omega_k$ is the realization of the disturbance at time $k$. From (16) and $\bar{x}^{i+1|k} = \Phi_k \bar{x}^*_{N|k}$, the sum of the first three terms on the RHS of (61) is

$$
\gamma^{-1} \sum_{i=0}^{N-1} \gamma^i \|C \bar{x}^*_{i|k}\|^2 + \gamma^N \|x_N^{i|k}\|^2_{\bar{P}_k} + \frac{\gamma}{1 - \gamma} \operatorname{tr}(\bar{P}_k) - \|C x_k\|^2.
$$

and from (16), the sum of the next two terms is $\omega_k^T \bar{P}_k \omega_k$. Noting that $K_k$, $\bar{P}_k$, and $x_k$ are independent of $\omega_k$, taking the expectation of $\varepsilon_{k+1}$ conditioned on information available at time $k$, therefore, gives

$$
\gamma \mathbb{E}_k\{\varepsilon_{k+1}\} = \sum_{i=0}^{N-1} \gamma^i \|C \bar{x}^*_{i|k}\|^2 + \gamma^N \|x_N^{i|k}\|^2_{\bar{P}_k} + \frac{\gamma}{1 - \gamma} \operatorname{tr}(\bar{P}_k) - \|C x_k\|^2.
$$

This equation, together with feasibility of the sequence $\{\bar{x}^*_{i|k}\}_{i=0}^{N-1}$ at time $k$, proves (52).

### REFERENCES

[1] D. Mayne, J. Rawlings, C. Rao, and P. Scokaert, “Constrained model predictive control: Stability and optimality,” *Automatica*, vol. 36, no. 6, pp. 789–814, 2000.

[2] B. Kouvaritakis, M. Cannon, S. Raković, and Q. Cheng, “Explicit use of probabilistic distributions in linear predictive control,” *Automatica*, vol. 46, no. 10, pp. 1719–1724, 2010.

[3] M. Farina, L. Giuliani, L. Magni, and R. Scattolini, “A probabilistic approach to model predictive control,” in *Proc. IEEE Conf. Decis. Control*, 2013, pp. 7734–7739.

[4] T. Hashimoto, “Probabilistic constrained model predictive control for linear discrete-time systems with additive stochastic disturbances,” in *Proc. IEEE Conf. Decis. Control*, 2013, pp. 6434–6439.

[5] G. Schildbach, P. Goulart, and M. Morari, “Linear controller design for chance constrained systems,” *Automatica*, vol. 51, pp. 278–284, 2015.

[6] G. Schildbach, L. Fagiano, C. Frei, and M. Morari, “The scenario approach for stochastic model predictive control with bounds on closed-loop constraint violations,” *Automatica*, vol. 50, no. 12, pp. 3009–3018, 2014.

[7] J. Fleming and M. Cannon, “Stochastic MPC for additive and multiplicative uncertainty using sample approximations,” *IEEE Trans. Autom. Control*, vol. 64, no. 9, pp. 3883–3888, Sep. 2019.

[8] M. Korda, R. Gondhalekar, F. Oldewurtel, and C. Jones, “Stochastic MPC framework for controlling the average constraint violation,” *IEEE Trans. Autom. Control*, vol. 59, no. 7, pp. 1706–1721, Jul. 2014.
Introduction to Stochastic Control Theory, Series: Dover

S. Yan, P. Goulart, and M. Cannon, “Stochastic model predictive control with dynamic feedback gain selection and discounted probabilistic constraints,” Automatica, vol. 96, pp. 32–39, 2018.

P. Hokayed, D. Chatterjee, and J. Lygeros, “On stochastic receding horizon control with bounded control inputs,” in Proc. 48th IEEE Conf. Decis. Control Held Jointly With 28th Chinese Control Conf., 2009, pp. 6359–6364.

L. Hewing and M. N. Zeilinger, “Stochastic model predictive control for linear systems using probabilistic reachable sets,” in Proc. IEEE Conf. Decis. Control, 2018, pp. 5182–5188.

L. Hewing, K. P. Bäuerschies, and M. N. Zeilinger, “Recursively feasible stochastic model predictive control using indirect feedback,” Automatica, vol. 119, 2020, Art. no. 109095.

D. Chatterjee and J. Lygeros, “On stability and performance of stochastic predictive control techniques,” IEEE Trans. Autom. Control, vol. 60, no. 2, pp. 509–514, Feb. 2015.

M. Lorenzen, M. A. Müller, and F. Allgöwer, “Stochastic model predictive control without terminal constraints,” Int. J. Robust Nonlinear Control, vol. 29, no. 15, pp. 4987–5001, 2019.

D. Limón, T. Alamo, F. Salas, and E. F. Camacho, “On the stability of constrained MPC without terminal constraint,” IEEE Trans. Autom. Control, vol. 51, no. 5, pp. 832–836, May 2006.

Camile and J. M. Maciejowski, “Robust finite horizon MPC without terminal constraints,” in Proc. 39th IEEE Conf. Decis. Control, 2000, vol. 1, pp. 166–171.

D. Blackwell, “Discounted dynamic programming,” Ann. Math. Statist., vol. 36, no. 1, pp. 226–235, 1965.

D. Bertsekas, Dynamic Programming and Optimal Control, 3rd ed. Belmont, MA, USA: Athena Scientific, 2005.

B. P. G. Van Parys, P. J. Goulart, and M. Morari, “Infinite horizon performance bounds for uncertain constrained systems,” IEEE Trans. Autom. Control, vol. 58, no. 11, pp. 2803–2817, Nov. 2013.

M. Kamgarpour and T. Summers, “On infinite dimensional linear programming approach to stochastic control,” IFAC-PapersOnLine, vol. 50, no. 1, pp. 6148–6153, 2017.

D. P. Bertsekas, Reinforcement Learning and Optimal Control. Belmont, MA, USA: Athena Scientific, 2019.

S. Barratt and S. Boyd, “Multi-period liability clearing via convex optimal control,” SSRN 3604618, 2020.

P. Niystrup, S. Boyd, E. Lindström, and H. Madsen, “Multi-period portfolio selection with drawdown control,” Ann. Operations Res., vol. 282, no. 1-2, pp. 245–271, 2020.

D. P. Boyd et al., Performance Bounds and Suboptimal Policies for Multi-Period Investment. Boston, MA, USA: Now Publishers, 2014.

C. W. Clark, “Profit maximization and the extinction of animal species,” J. Political Econ., vol. 81, no. 4, pp. 950–961, 1973.

D. Ludwig, W. A. Brock, and S. R. Carpenter, “Uncertainty in discount models and environmental accounting,” Ecol. Soc., vol. 10, no. 2, 2005. [Online]. Available: http://www.ecologyandsociety.org/vol10/iss2/art13/

S. Yan, P. Goulart, and M. Cannon, “Stochastic model predictive control with discounted probabilistic constraints,” in Proc. Eur. Control Conf., 2018, pp. 1003–1008.

G. Grimmett and D. Stirzaker, Probability and Random Processes, 3rd ed. London, U.K.: Oxford Univ. Press, 2001.

B. Kouvaritakis and M. Cannon, Model Predictive Control: Classical, Robust and Stochastic. Berlin, Germany: Springer, 2015.

E. Kerrigan, “Robust constraint satisfaction: Invariant sets and predictive control,” Ph.D. dissertation, Dep. Eng., Univ. Cambridge, Cambridge, U.K., 2000.

R. Oberdieck and E. N. Pistikopoulos, “Multi-objective optimization with convex quadratic cost functions: A multi-parametric programming approach,” Comput. Chem. Eng., vol. 85, pp. 36–39, 2016.

R. Bellman, “Dynamic programming,” Science, vol. 153, no. 3731, pp. 34–37, 1966.

K. J. Aström, Introduction to Stochastic Control Theory, Series: Dover Books on Electrical Engineering. New York, NY, USA: Dover Publications, 2006.

S. Bitanti, A. J. Laub, and J. C. Willems, The Riccati Equation. Berlin, Germany: Springer, 2012.

R. R. Bitmead, M. R. Gevers, I. R. Petersen, and R. J. Kaye, “Monotonicity and stabilizability- properties of solutions of the Riccati difference equation: Propositions, lemmas, theorems, fallacious conjectures and counterexamples,” Syst. Control Lett., vol. 5, pp. 309–315, 1985.

T. Nishimura, “On the a priori information in sequential estimation problems,” IEEE Trans. Autom. Control, vol. AC-11, no. 2, pp. 197–204, Apr. 1966.

B. Kouvaritakis, M. Cannon, and J. A. Rossiter, “Who needs QP for linear MPC anyway?,” Automatica, vol. 38, pp. 879–884, 2002.

M. Cannon, B. Kouvaritakis, and X. Wu, “Probabilistic constrained MPC for multiplicative and additive stochastic uncertainty,” IEEE Trans. Autom. Control, vol. 54, no. 7, pp. 1626–1632, Jul. 2009.

S. T. Lian, K. Marzuki, and Y. Rubiyah, “Tuning of a neuro-fuzzy controller by genetic algorithms with an application to a coupled-tank level-level control system,” Eng. Appl. Artif. Intell., vol. 11, no. 4, pp. 517–529, 1998.

A. W. Roberts and D. E. Varberg, Convex Functions, Series: Pure and Applied Mathematics. New York, NY, USA: Academic Press, 1973.