

Stochastic MPC with Dynamic Feedback Gain Selection and Discounted Probabilistic Constraints

Shuhao Yan, Paul J. Goulart and Mark Cannon

Abstract—This paper considers linear discrete-time systems with additive disturbances, and designs a Model Predictive Control (MPC) law incorporating a dynamic feedback gain to minimise a quadratic cost function subject to a single chance constraint. The feedback gain is selected online and we provide two selection methods based on minimising upper bounds on predicted costs. The chance constraint is defined as a discounted sum of violation probabilities on an infinite horizon. By penalising 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 optimisation 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—Model predictive control, chance constraints, Chebyshev inequality, dynamic programming, multiobjective optimisation, 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 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 optimisation problems while the latter generally results in randomised 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.

The authors are with the Department of Engineering Science, University of Oxford, OX1 3PJ, UK. (E-mail: shuhao.yan@eng.ox.ac.uk; paul.goulart@eng.ox.ac.uk; mark.cannon@eng.ox.ac.uk)

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 optimisations are assumed but not ensured. More generally, sample based methods are unable to ensure recurrence of feasibility of receding horizon optimisation 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 that 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 constraints. To ensure recursive feasibility under unbounded additive disturbances, previous work (e.g. [3], [13], [14]) has resorted to a backup initialisation of the MPC optimisation 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], [28]). Discount factors in optimal control problems allow performance in the near future to be prioritised over long-term behaviour. 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], [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 paper considers linear discrete-time systems subject to possibly unbounded additive disturbances. We propose an MPC strategy incorporating a dynamic feedback gain to minimise 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 optimisation problem. The main features and contributions of this paper are summarised as follows:

- We use a discount factor to ensure that the chance constraint is well-defined and to prioritise near-future system behaviour over steady state performance.
- A constraint tightening technique is proposed to ensure recursive feasibility of online MPC optimisation problems and constraint satisfaction in closed loop without requiring disturbances to be bounded.
- The closed loop system satisfies a quadratic stability condition without the need for terminal constraints.

This paper 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 optimisation problems with the cost and constraint functions considered as conflicting objectives to generate a set of feedback gains providing various trade-offs between these two objectives. Two methods of online gain selection are proposed, minimising upper bounds on the optimal predicted cost while retaining guarantees of recursive feasibility and computationally simple MPC optimisation. 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 paper is organised 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 optimisation problems, solutions of which provide strictly stabilising 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 summarises the proposed MPC algorithm and derives a bound on closed loop performance. In Section VII, the closed loop behaviour of the tightening parameters is analysed and constraint satisfaction is proved. Section VIII gives a numerical example illustrating the results obtained and the paper 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 $\|x\|_Q^2 := x^\top Q x$. The notation $Q \succcurlyeq 0$ and $R \succ 0$ indicates 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 $\mathbf{0}$, and an identity matrix of suitable dimension is denoted \mathbf{I} . The conditional probability of an event \mathcal{F} given the state x_k is denoted $\mathbb{P}\{\mathcal{F}|x_k\} = \mathbb{P}_k\{\mathcal{F}\}$, the conditional expectation of y given x_k is $\mathbb{E}\{y|x_k\} = \mathbb{E}_k\{y\}$, and $\mathbb{P}\{\mathcal{F}\}$, $\mathbb{E}\{y\}$ are equivalent to $\mathbb{P}_0\{\mathcal{F}\}$, $\mathbb{E}_0\{y\}$ respectively. The sequence $\{x_0, \dots, x_{N-1}\}$ is denoted $\{x_i\}_{i=0}^{N-1}$. 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 $x_{i|k}$.

II. PROBLEM DESCRIPTION

Consider an uncertain linear system

$$x_{k+1} = Ax_k + Bu_k + \omega_k, \quad (1)$$

where $x_k \in \mathbb{R}^{n_x}$, $u_k \in \mathbb{R}^{n_u}$ are the system state 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

$$\mathbb{E}\{\omega_k\} = \mathbf{0}, \quad \mathbb{E}\{\omega_k \omega_k^\top\} =: \Omega \succ 0. \quad (2)$$

The disturbance distribution may have infinite support, unlike the approach of [9], which assumes 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

$$\sum_{k=0}^{\infty} \gamma^k \mathbb{P}\{\|Cx_k + Du_k\| \geq 1\} \leq e$$

for given $C \in \mathbb{R}^{n_c \times n_x}$, $D \in \mathbb{R}^{n_c \times n_u}$, a positive scalar e and a discount factor $\gamma \in (0, 1)$. To simplify presentation we set $D = 0$ for the remainder of the paper, noting that all of the results given in Sections III-VII apply to the case of non-zero D , and in the sequel the constraint

$$\sum_{k=0}^{\infty} \gamma^k \mathbb{P}\{\|Cx_k\| \geq 1\} \leq e \quad (3)$$

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

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

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

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

Assumption 1. (A, B) is controllable and $(A, Q^{\frac{1}{2}})$ 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 trade-off between short-term and long-term behaviours 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 optimisation 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 ω_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 parameterised at each stage with a finite number of decision variables, and the constraint (3) is approximated conservatively using the two-sided Chebyshev inequality. The resulting optimisation 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

$$\bar{u}_{i|k} = K_k \bar{x}_{i|k} + c_{i|k}, \quad i = 0, \dots, N-1 \quad (5)$$

$$\bar{u}_{N+i|k} = K_k \bar{x}_{N+i|k}, \quad i = 0, 1, \dots, \quad (6)$$

where $\bar{x}_{i|k}$ is the i -step-ahead prediction of the nominal state given information at time k , that is, $\mathbb{E}_k \{x_{i|k}\} = \bar{x}_{i|k}$. The matrix-valued term K_k is a stabilising 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* $\{c_{0|k}, c_{1|k}, \dots, c_{N-1|k}\}$ then

constitutes the decision variables in the MPC optimisation problem to be solved at time k .

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

$$\bar{x}_{i|k} = \Phi_k^i \bar{x}_{0|k} + \sum_{j=0}^{i-1} \Phi_k^{i-1-j} B c_{j|k}, \quad i = 1, \dots, N, \quad (7)$$

$$\bar{x}_{N+i|k} = \Phi_k^i \bar{x}_{N|k}, \quad i = 1, 2, \dots, \quad (8)$$

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

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

We rewrite (7) in a compact form as

$$\bar{\mathbf{x}}_k = M_x(K_k) \bar{x}_{0|k} + M_c(K_k) \mathbf{c}_k, \quad (10)$$

$$\begin{bmatrix} \bar{x}_{1|k} \\ \bar{x}_{2|k} \\ \vdots \\ \bar{x}_{N|k} \end{bmatrix} = \begin{bmatrix} \Phi_k \\ \Phi_k^2 \\ \vdots \\ \Phi_k^N \end{bmatrix} \bar{x}_{0|k} + \begin{bmatrix} B \\ \Phi_k B \\ \vdots \\ \Phi_k^{N-1} B & \dots & B \end{bmatrix} \begin{bmatrix} c_{0|k} \\ c_{1|k} \\ \vdots \\ c_{N-1|k} \end{bmatrix}, \quad (11)$$

and $M_x(K_k) \in \mathbb{R}^{Nn_x \times n_x}$, $M_c(K_k) \in \mathbb{R}^{Nn_x \times Nn_u}$. 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 optimisation

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 optimisation problem to be solved repeatedly online.

Minimising the predicted cost $\mathbb{E}_k \{ \sum_{i=0}^{\infty} \|x_{i|k}\|_Q^2 + \|u_{i|k}\|_R^2 \}$ at time k over the optimisation variable \mathbf{c}_k is equivalent to minimising the cost defined in terms of the predicted nominal input sequence (5)-(6) and state trajectory (7)-(8) by

$$J(\bar{x}_{0|k}, K_k, \mathbf{c}_k) := \sum_{i=0}^{N-1} \left(\|\bar{x}_{i|k}\|_Q^2 + \|K_k \bar{x}_{i|k} + c_{i|k}\|_R^2 \right) + \|\bar{x}_{N|k}\|_{P_k}^2. \quad (12)$$

Here $\|\bar{x}_{N|k}\|_{P_k}^2$ is the terminal cost and $P_k \succ 0$ is chosen as the solution of

$$P_k = Q + K_k^\top R K_k + \Phi_k^\top P_k \Phi_k. \quad (13)$$

Using (10)-(11), we rewrite (12) in a compact form as

$$J(\bar{x}_{0|k}, K_k, \mathbf{c}_k) = \begin{bmatrix} \bar{x}_{0|k} \\ \mathbf{c}_k \end{bmatrix}^\top W_2(K_k) \begin{bmatrix} \bar{x}_{0|k} \\ \mathbf{c}_k \end{bmatrix}, \quad (14)$$

where $W_2(K_k)$ is a function of K_k and its expression is omitted here for simplicity.

Approximating the 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

$$\sum_{i=0}^{\infty} \gamma^i (\|C \bar{x}_{i|k}\|^2 + \text{tr}(C^\top C X_{i|k})) \leq \varepsilon_k. \quad (15)$$

Here ε_k replaces e 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 optimisations 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*

$$\tilde{P}_k = \gamma \Phi_k^\top \tilde{P}_k \Phi_k + C^\top C. \quad (16)$$

Then

$$\sum_{i=N}^{\infty} \gamma^i \|C \bar{x}_{i|k}\|^2 = \gamma^N \|\bar{x}_{N|k}\|_{\tilde{P}_k}^2, \\ \sum_{i=0}^{\infty} \gamma^i \text{tr}(C^\top C X_{i|k}) = \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k),$$

where $\bar{x}_{i|k}$ is given by (8) for all $i \geq N$, $X_{0|k} = \mathbf{0}$ and $X_{i|k}$ is given by (9) for all $i \geq 1$.

By Lemma 1, (15) is equivalent to

$$\sum_{i=0}^{N-1} \gamma^i \|C \bar{x}_{i|k}\|^2 + \gamma^N \|\bar{x}_{N|k}\|_{\tilde{P}_k}^2 + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k) \leq \varepsilon_k, \quad (17)$$

where $\gamma^N \|\bar{x}_{N|k}\|_{\tilde{P}_k}^2$ 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)-(11), we rewrite (17) in a more compact form as

$$\begin{bmatrix} \bar{x}_{0|k} \\ \mathbf{c}_k \end{bmatrix}^\top W_1(K_k) \begin{bmatrix} \bar{x}_{0|k} \\ \mathbf{c}_k \end{bmatrix} + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k) \leq \varepsilon_k, \quad (18)$$

where $W_1(K_k)$ is a function of K_k defined as

$$W_1(K_k) := \begin{bmatrix} C^\top C + M_x^\top H M_x & M_x^\top H M_c \\ M_c^\top H M_x & M_c^\top H M_c \end{bmatrix}, \quad (19)$$

and $H := \text{diag}(\gamma C^\top C, \dots, \gamma^{N-1} C^\top C, \gamma^N \tilde{P}_k) \succcurlyeq 0$.

To summarise, the MPC optimisation solved at time k is

$$J^*(x_k, K_k) := \min_{\mathbf{c}_k} \{J(x_k, K_k, \mathbf{c}_k) \mid (18) \text{ with } \bar{x}_{0|k} = x_k\}, \quad (20)$$

and its solution for any feasible x_k , K_k and ε_k is denoted $\mathbf{c}_k^*(x_k, K_k, \varepsilon_k)$. For simplicity we write this solution as \mathbf{c}_k^* , with the understanding that this vector depends on x_k , K_k and ε_k . The corresponding predicted nominal state trajectory is given by

$$\bar{x}_{i|k}^* = \Phi_k^i x_k + \sum_{j=0}^{i-1} \Phi_k^{i-1-j} B c_{j|k}^*, \quad i = 1, \dots, N, \quad (21)$$

$$\bar{x}_{N+i|k}^* = \Phi_k^i \bar{x}_{N|k}^*, \quad i = 1, 2, \dots \quad (22)$$

The MPC law at time k is defined by

$$u_k := K_k x_k + c_{0|k}^*, \quad (23)$$

and the closed loop system dynamics are given by

$$x_{k+1} = \Phi_k x_k + B c_{0|k}^*(x_k, K_k, \varepsilon_k) + \omega_k, \quad (24)$$

where ω_k is the disturbance realisation at time k .

In the remainder of this paper, we discuss how to choose ε_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 minimising predicted costs while retaining the recursive feasibility guarantee of online MPC optimisations 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 Section VI and Section VII.

III. RECURSIVE FEASIBILITY

Recursively feasible MPC strategies have the property that the MPC optimisation 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 optimisation, 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 ε_k that is necessarily feasible at time k .

We use the notation $\tilde{\mathbf{c}}_{k+1}$ to denote a time-shifted version of \mathbf{c}_k^* , defined by

$$\tilde{\mathbf{c}}_{k+1} := E \mathbf{c}_k^*, \quad (25)$$

where E is the matrix such that $E \mathbf{c} = [c_1^\top, \dots, c_{N-1}^\top, \mathbf{0}]^\top$ if $\mathbf{c} = [c_0^\top, \dots, c_{N-1}^\top]^\top$.

Lemma 2. *Given initial feasibility at time $k = 0$, the MPC optimisation (20) is recursively feasible if ε_k is defined at each time $k = 1, 2, \dots$ as*

$$\varepsilon_k := \begin{bmatrix} x_k \\ \tilde{\mathbf{c}}_k \end{bmatrix}^\top W_1(K_{k-1}) \begin{bmatrix} x_k \\ \tilde{\mathbf{c}}_k \end{bmatrix} + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_{k-1}). \quad (26)$$

Equation (26) provides an explicit expression for ε_k for all $k > 0$ in terms of (21), (22) and disturbance realisations as

$$\begin{aligned} \varepsilon_k &= \sum_{i=0}^{N-1} \gamma^i \|C(\bar{x}_{i+1|k-1}^* + \Phi_{k-1}^i \omega_{k-1})\|^2 \\ &+ \gamma^N \|\bar{x}_{N+1|k-1}^* + \Phi_{k-1}^N \omega_{k-1}\|_{\tilde{P}_{k-1}}^2 + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_{k-1}). \end{aligned} \quad (27)$$

Here

$$\bar{x}_{i+1|k-1}^* + \Phi_{k-1}^i \omega_{k-1} := \bar{x}_{i|k}, \quad i = 0, \dots, N \quad (28)$$

defines a feasible nominal state sequence predicted at time k , obtained by setting $\bar{x}_{0|k} = x_k$ and $\bar{u}_{i|k} = K_{k-1} \bar{x}_{i|k} + \tilde{c}_{i|k}$. Note that in constructing this feasible sequence, we still use K_{k-1} and therefore \tilde{P}_{k-1} as the corresponding terminal matrix, and that the feedback gain is updated after the update of ε_k .

Essentially, the optimisation problem to be solved at each time step is feasible because the parameter ε_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 $\{\varepsilon_k\}_{k=0}^\infty$ (as will be detailed in Section VII).

The key to this method lies in the definition of the vector \tilde{c}_{k+1} . If this vector were optimised, 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 OPTIMISATION AND DYNAMIC PROGRAMMING

In this section, we generate a set of strictly stabilising feedback gains offline, from which K_k in (5)-(6) is selected online. Considering the cost function in (4) and the constraint (3) as conflicting objectives, we formulate a set of multiobjective optimisation problems whose solutions provide feedback gains representing a trade-off between these two objectives. These multiobjective optimisation problems are written using linear scalarisation [33] for given $x_0 \in \mathbb{R}^{n_x}$ in the form

$$\begin{aligned} \min_{u_0, u_1, \dots} (1-\mu) \sum_{i=0}^{\infty} \gamma^i \mathbb{E}\{\|Cx_i\|^2\} + \mu \sum_{i=0}^{\infty} \mathbb{E}\{\|x_i\|_Q^2 + \|u_i\|_R^2\} \\ \text{s.t. } x_{i+1} = Ax_i + Bu_i + \omega_i, \quad i = 0, 1, \dots \end{aligned} \quad (\mathcal{P}_1) \quad (29)$$

where $\mu \in (0, 1]$ is a weighting parameter. Here ω_i is an i.i.d. random variable with the same statistics as given in (2). Note that in (\mathcal{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 trade-offs between these competing objectives can be explored using different values of μ , and their solutions provide a set of strictly stabilising feedback gains. We denote this set of feedback gains as \mathcal{K} .

We propose to solve (\mathcal{P}_1) by Dynamic Programming [34] via its equivalent deterministic counterpart in the form of

$$\begin{aligned} \min_{\bar{u}_0, \bar{u}_1, \dots} (1-\mu) \sum_{i=0}^{\infty} \gamma^i \|C\bar{x}_i\|^2 + \mu \sum_{i=0}^{\infty} (\|\bar{x}_i\|_Q^2 + \|\bar{u}_i\|_R^2) \quad (\mathcal{P}_2) \\ \text{s.t. } \bar{x}_{i+1} = A\bar{x}_i + B\bar{u}_i, \quad i = 0, 1, \dots \end{aligned} \quad (30)$$

where \bar{x}_i and \bar{u}_i are mean values of x_i and u_i respectively, with $\bar{x}_0 = x_0$. From the quadratic form of the objective function, the optimal solution to (\mathcal{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 (\mathcal{P}_1) and (\mathcal{P}_2) in Theorem 3, which is in agreement with the certainty equivalence principle [35, Section 8.4].

Theorem 3. *Problems (\mathcal{P}_1) and (\mathcal{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 (\mathcal{P}_1) . Let

$$\begin{aligned} (1-\mu)\bar{J}_0(x_0) + \mu\hat{J}_0(x_0) := \\ \lim_{T \rightarrow \infty} \min_{u_0, \dots, u_{T-1}} \mathbb{E}\left\{ (1-\mu) \sum_{i=0}^{T-1} \gamma^i \|Cx_i\|^2 + \mu \sum_{i=0}^{T-1} (\|x_i\|_Q^2 + \|u_i\|_R^2) \right\} \\ \text{s.t. } (29), \end{aligned}$$

where $\bar{J}_0(x_0)$ and $\hat{J}_0(x_0)$ are optimal cost-to-go functions denoting the optimal values of $\lim_{T \rightarrow \infty} \sum_{i=0}^{T-1} \gamma^i \mathbb{E}\{\|Cx_i\|^2\}$ and $\lim_{T \rightarrow \infty} \sum_{i=0}^{T-1} \mathbb{E}\{\|x_i\|_Q^2 + \|u_i\|_R^2\}$, 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 the system (29). By solving the Bellman equations [34] backwards in time with initial conditions $\bar{J}_T(x_T) = 0$ and $\hat{J}_T(x_T) = 0$ where $\bar{J}_T(x_T), \hat{J}_T(x_T)$ are the optimal cost-to-go functions starting from state x_T , it can be shown that $\bar{J}_0(x_0) = x_0^\top \bar{S}_0 x_0 + r_0$, $\hat{J}_0(x_0) = x_0^\top \hat{S}_0 x_0 + v_0$. Here \bar{S}_0, \hat{S}_0 are positive semidefinite matrices and r_0, v_0 respectively represent discounted and undiscounted sums of variances of x_i for $i = 1, \dots, T$. Due to the quadratic form of the objective function and the assumption of zero-mean disturbance ω_i , the optimal control input u_i is a linear function of x_i for all $i \geq 0$. Denoting the optimal feedback gain as H_i , we obtain the following DP iteration for $i = T-1, \dots, 0$:

$$\bar{S}_i = C^\top C + \gamma(A + BH_i)^\top \bar{S}_{i+1} (A + BH_i), \quad (31)$$

$$\hat{S}_i = Q + H_i^\top R H_i + (A + BH_i)^\top \hat{S}_{i+1} (A + BH_i), \quad (32)$$

$$r_i = \gamma (r_{i+1} + \text{tr}(\bar{S}_{i+1} \Omega)), \quad (33)$$

$$v_i = \text{tr}(\hat{S}_{i+1} \Omega) + v_{i+1}, \quad (34)$$

where we choose $\bar{S}_T = \hat{S}_T = \mathbf{0}$ 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 $i = T - 1, \dots, 0$ with $\Delta_{i+1} = (1 - \mu)\gamma\bar{S}_{i+1} + \mu\hat{S}_{i+1}$. Therefore, to obtain H_i , \bar{S}_i and \hat{S}_i it is only necessary to perform the iterations in (31)-(32), since these are independent of r_i and v_i . By considering the limit as $T \rightarrow \infty$ and using similar reasoning, it can be shown that $x_0^\top((1-\mu)\bar{S}_0 + \mu\hat{S}_0)x_0$ is the optimal objective of problem (P_2) and the corresponding optimal control law is given for all $i \geq 0$ by $\bar{u}_i = H_i\bar{x}_i$. ■

Theorem 3 demonstrates that we only need to solve problem (P_2) to obtain the optimal feedback gains for (P_1) . To suit our purposes, we reverse the time indexing in (31) and (32) and define $\bar{P}_i := \bar{S}_{T-i}$, $\hat{P}_i := \hat{S}_{T-i}$ for $i = 0, \dots, T$ and $L_i := H_{T-i}$ for $i = 1, \dots, T$. The resulting DP iteration is

$$L_{i+1} = -[\mu R + B^\top \Sigma_i B]^{-1} B^\top \Sigma_i A, \quad (35)$$

$$\bar{P}_{i+1} = C^\top C + \gamma(A + BL_{i+1})^\top \bar{P}_i (A + BL_{i+1}), \quad (36)$$

$$\hat{P}_{i+1} = Q + L_{i+1}^\top RL_{i+1} + (A + BL_{i+1})^\top \hat{P}_i (A + BL_{i+1}), \quad (37)$$

with $\Sigma_i := \gamma(1 - \mu)\bar{P}_i + \mu\hat{P}_i$, for $i = 0, \dots, T - 1$. These provide the solution of (P_2) in the limit as $T \rightarrow \infty$ and the optimal value of the objective function of problem (P_2) is $x_0^\top((1 - \mu)\bar{P}_\infty + \mu\hat{P}_\infty)x_0$ for any given $x_0 \in \mathbb{R}^n$, where \bar{P}_∞ and \hat{P}_∞ are defined in the following theorem.

Theorem 4. Consider equations (35)-(37). Under Assumption 1: (a) there exist matrices $\bar{P}_\infty \succcurlyeq 0$, $\hat{P}_\infty \succ 0$ such that for arbitrary positive semidefinite matrices \bar{P}_0 and \hat{P}_0 we have

$$\lim_{i \rightarrow \infty} \bar{P}_i = \bar{P}_\infty \text{ and } \lim_{i \rightarrow \infty} \hat{P}_i = \hat{P}_\infty,$$

(b) for all $\mu \in (0, 1]$ the feedback gain L_∞ is strictly stabilising, where

$$L_\infty := -[\mu R + B^\top \Sigma_\infty B]^{-1} B^\top \Sigma_\infty A \quad (38)$$

with $\Sigma_\infty := \gamma(1 - \mu)\bar{P}_\infty + \mu\hat{P}_\infty$, and (c) the matrices \bar{P}_∞ and \hat{P}_∞ satisfy the following Lyapunov matrix equations:

$$\bar{P}_\infty = C^\top C + \gamma(A + BL_\infty)^\top \bar{P}_\infty (A + BL_\infty), \quad (39)$$

$$\hat{P}_\infty = Q + L_\infty^\top RL_\infty + (A + BL_\infty)^\top \hat{P}_\infty (A + BL_\infty). \quad (40)$$

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

The solutions L_∞ , \bar{P}_∞ , \hat{P}_∞ are functions of μ and are denoted $L_\infty(\mu)$, $\bar{P}_\infty(\mu)$, $\hat{P}_\infty(\mu)$ in the remainder of this paper. With $\mu = 1$, $L_\infty(1) = K_{LQ}$, where K_{LQ} is the LQ-optimal feedback gain that minimises the second part of the objective in (P_2) , whereas with $\mu = 0$ $L_\infty(0) = -(B^\top \bar{P}_\infty(0)B)^\dagger B^\top \bar{P}_\infty(0)A$, which is optimal with respect to the first part of the objective. However, the gain $L_\infty(0)$ may not be stabilising and unique due to discounting and no penalty on control inputs if $\mu = 0$, and we therefore restrict the weighting parameter μ to the interval $(0, 1]$.

Remark 5. The fixed point $(L_\infty(\mu), \bar{P}_\infty(\mu), \hat{P}_\infty(\mu))$ of the iteration (35)-(37) coincides with the minimising argument of

$$\begin{aligned} & \min_{G \in \mathbb{R}^{n_u \times n_x}, Z_1 \succcurlyeq 0, Z_2 \succcurlyeq 0} \text{tr}((1 - \mu)Z_1 + \mu Z_2) \\ & \text{s.t. } Z_1 = C^\top C + \gamma(A + BG)^\top Z_1 (A + BG), \\ & Z_2 = Q + G^\top RG + (A + BG)^\top Z_2 (A + BG). \end{aligned}$$

However, the iteration is generally preferred over solving this equivalent problem directly since it is nonconvex in variables G, Z_1, 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.

\mathcal{K} Generation (Offline):

- (1) Choose appropriately a sequence $\{\mu_i\}_{i=1}^m$, with $\mu_i \in (0, 1]$ in ascending order, and $\mu_m = 1$;
- (2) For each $i = 1, \dots, m$, solve problem (P_2) with $\mu = \mu_i$ by executing iterations (35)-(37);
- (3) Obtain the set of strictly stabilising feedback gains $\mathcal{K} := \{L_\infty(\mu_i)\}_{i=1}^m$.

In the \mathcal{K} Generation, there should be a sufficiently large number of elements in the sequence $\{\mu_i\}_{i=1}^m$ so that L_∞ can be adequately approximated on the intervals between consecutive points in this sequence. More importantly, μ_1 should be appropriately chosen close to 0 while ensuring that $L_\infty(\mu_1) \neq L_\infty(0)$ and $\bar{P}_\infty(\mu_1) \neq \bar{P}_\infty(0)$ if the latter is strictly stabilising, and hence $L_\infty(\mu) \neq L_\infty(0) \forall \mu \geq \mu_1$ by the uniqueness of solutions to problem (P_2) and the monotonicity property of $\bar{P}_\infty(\mu)$ as will be given in the next section. Furthermore, step (2) can be warm-started by using $\bar{P}_\infty(\mu_i)$ and $\hat{P}_\infty(\mu_i)$ to initialise the iteration with weighting parameter μ_{i+1} , thereby reducing considerably the time required to solve (P_2) for each value of μ . Note also that in step (3) the sets $\{\bar{P}_\infty(\mu_i)\}_{i=1}^m$, $\{\hat{P}_\infty(\mu_i)\}_{i=1}^m$ are obtained.

V. DYNAMIC FEEDBACK GAIN SELECTION

In this section, we provide two methods for dynamically selecting feedback gains from the set \mathcal{K} discussed in Section IV. These methods are designed such that the recursive feasibility guarantee of (20) and a computationally simple online optimisation are retained. Both methods determine feedback gains that minimise 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_{LQ} . 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 μ , 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 $\bar{P}_\infty(\cdot)$ and $\hat{P}_\infty(\cdot)$ on $(0, 1)$.

Lemma 6. For all μ_1, μ_2 such that $0 < \mu_1 \leq \mu_2 \leq 1$, $\bar{P}_\infty(\cdot)$, $\hat{P}_\infty(\cdot)$ satisfy

$$\begin{aligned} \bar{P}_\infty(\mu_1) & \preceq \bar{P}_\infty(\mu_2), \\ \hat{P}_\infty(\mu_1) & \succeq \hat{P}_\infty(\mu_2). \end{aligned}$$

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. $\bar{P}_\infty(\cdot)$, $\hat{P}_\infty(\cdot)$ 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 $k = 1, 2, \dots$

(1) Compute

$$\mathbf{c}^o(x_k) := \arg \min_{\mathbf{c}} \begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix}^\top W_1(K_{k-1}) \begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix} \quad (41)$$

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

(2) Compute

$$\begin{aligned} \bar{\mu}_k &:= \max \left\{ \arg \max_{\mu \in \{\mu_i\}_{i=1}^m} \frac{\gamma}{1-\gamma} \text{tr}(\Omega \bar{P}_\infty(\mu)) \right. \\ &\quad \text{s.t.} \\ &\quad \left. \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}^\top W_1(K_{k-1}) \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix} + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \bar{P}_\infty(\mu)) \leq \varepsilon_k \right\} \end{aligned} \quad (42)$$

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

(3) Set $K_k := L_\infty(\bar{\mu}_k)$, $\tilde{P}_k := \bar{P}_\infty(\bar{\mu}_k)$ and $P_k := \tilde{P}_\infty(\bar{\mu}_k)$.

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

$$\begin{aligned} \varepsilon_k &= \begin{bmatrix} x_k \\ E\mathbf{c}_{k-1}^* \end{bmatrix}^\top W_1(L_\infty(\bar{\mu}_{k-1})) \begin{bmatrix} x_k \\ E\mathbf{c}_{k-1}^* \end{bmatrix} \\ &\quad + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \bar{P}_\infty(\bar{\mu}_{k-1})). \end{aligned}$$

From Lemma 6, (42) is therefore equivalent to

$$\bar{\mu}_k := \max_{\mu \in \{\mu_i\}_{i=1}^m} \mu \quad (43a)$$

s.t.

$$\begin{aligned} \frac{\gamma}{1-\gamma} \text{tr}(\Omega(\bar{P}_\infty(\mu) - \bar{P}_\infty(\bar{\mu}_{k-1}))) &\leq \begin{bmatrix} x_k \\ E\mathbf{c}_{k-1}^* \end{bmatrix}^\top \left[\star \right] \begin{bmatrix} x_k \\ E\mathbf{c}_{k-1}^* \end{bmatrix} \\ &\quad - \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}^\top \left[\star \right] \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}, \end{aligned} \quad (43b)$$

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

$$\begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}^\top W_1(K_{k-1}) \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix} + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k) \leq \varepsilon_k,$$

which implies that a *perturbation sequence* \mathbf{c}' exists satisfying

$$\begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix}^\top W_1(K_{k-1}) \begin{bmatrix} x_k \\ \mathbf{c}^o(x_k) \end{bmatrix} = \begin{bmatrix} x_k \\ \mathbf{c}' \end{bmatrix}^\top W_1(K_k) \begin{bmatrix} x_k \\ \mathbf{c}' \end{bmatrix} \quad (44)$$

so that problem (20) remains feasible at time k . Therefore Method 1 retains the recursive feasibility guarantee. From (44) it also follows that $J(x_k, K_k, \mathbf{c}')$ 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) minimises the trace of the terminal matrix $P_k = \tilde{P}_\infty(\bar{\mu}_k)$ in this cost.

By Lemma 6, step 2 is simply a binary search to determine the largest $\mu \in \{\mu_i\}_{i=1}^m$ 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 $\bar{\mu}_0$ is not determined by Method 1, and we therefore use other means of choosing $\bar{\mu}_0 \in \{\mu_i\}_{i=1}^m$ and $K_0 = L_\infty(\bar{\mu}_0)$ to make the MPC optimisation (20) initially feasible (assuming such a $\bar{\mu}_0$ exists); this is discussed in Section VIII, Simulation D.

In the remainder of this section, we analyse 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 $\bar{\mu}_{k-1} \leq \bar{\mu}_k$, 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 non-decreasing and convergent. To derive a stronger convergence result, we make a simplifying assumption.

Assumption 2. The maximisation (42) can be solved with the optimisation variable μ varying continuously in the interval $[\bar{\mu}_0, 1]$ rather than being constrained to the finite set $\{\mu_i\}_{i=1}^m$.

The purpose of Assumption 2 is to ensure that $\bar{\mu}_k > \bar{\mu}_{k-1}$ 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 $\mu \in \{\mu_i\}_{i=1}^m$ in step (2). To analyse asymptotic convergence of $\{\bar{\mu}_k\}_{k=0}^\infty$, we consider the RHS of (43b). By definition, $W_1(K_{k-1}) = W_1(L_\infty(\bar{\mu}_{k-1}))$ is at least positive semidefinite so the minimisation in step 1 is well-defined and $\mathbf{c}^o(x_k)$ can be defined uniquely as

$$\mathbf{c}^o(x_k) = -(W_{cc}(\bar{\mu}_{k-1}))^\dagger W_{cx}(\bar{\mu}_{k-1})x_k.$$

Here $W_{cc}(\mu), W_{cx}(\mu)$ are blocks of $W_1(L_\infty(\mu))$ in the partition (19) such that $W_{cc} = M_c^\top H M_c$, $W_{cx} = M_c^\top H M_x$, and W_{cc}^\dagger is the pseudoinverse of W_{cc} . Since $x_k = \bar{x}_1|_{k-1} + \omega_{k-1}$, it follows that the RHS of (43b) is equal to

$$\|z_{k-1} + (W_{cc}(\bar{\mu}_{k-1}))^\dagger W_{cx}(\bar{\mu}_{k-1})\omega_{k-1}\|_{W_{cc}(\bar{\mu}_{k-1})}^2, \quad (45)$$

where $z_{k-1} = E\mathbf{c}_{k-1}^* + (W_{cc}(\bar{\mu}_{k-1}))^\dagger W_{cx}(\bar{\mu}_{k-1})\bar{x}_1|_{k-1}$.

Lemma 9. There exist $\delta > 0$ and $p_\delta > 0$ such that

$$\inf_{\substack{z \in \mathbb{R}^{Nn_u} \\ \mu \in [\bar{\mu}_0, 1]}} \mathbb{P}\{\|z + (W_{cc}(\mu))^\dagger W_{cx}(\mu)\omega_k\|_{W_{cc}(\mu)}^2 \geq \delta\} \geq p_\delta. \quad (46)$$

We conclude that K_k converges almost surely to K_{LQ} .

Theorem 10. Under Assumption 2, $\bar{\mu}_k \rightarrow 1$ as $k \rightarrow \infty$ with probability 1.

B. Gain selection method 2

We first give a monotonicity result that can be used to prove the order-preserving property of Riccati Difference Equations (RDE) [36], which is exploited in Method 2.

Lemma 11. $V(x_0, \check{P}_1) \leq V(x_0, \check{P}_2) \forall x_0 \in \mathbb{R}^{n_x}$ if $\check{P}_1 \preceq \check{P}_2$, where

$$V(x_0, \check{P}) := \min_{u_0, \dots, u_{N-1}} \{\tilde{V}(x_0, \mathbf{u}, \check{P})\}$$

$$\text{s.t. } x_{i+1} = Ax_i + Bu_i, \quad i = 0, \dots, N-1, \quad (47)$$

and $\tilde{V}(x_0, \mathbf{u}, \tilde{P}) := \sum_{i=0}^{N-1} (\|x_i\|_Q^2 + \|u_i\|_R^2) + \|x_N\|_{\tilde{P}}^2$, $\mathbf{u} := [u_0^\top, \dots, u_{N-1}^\top]^\top$ for $Q, R, \tilde{P} \succ 0$.

Proof: Let $\mathbf{u}^*(\tilde{P}) := \arg \min_{\mathbf{u}} \{ \tilde{V}(x_0, \mathbf{u}, \tilde{P}) \text{ s.t. } x_{i+1} = Ax_i + Bu_i, \quad i = 0, \dots, N-1 \}$, then by the ordering between \tilde{P}_1 and \tilde{P}_2 and optimality, we have $V(x_0, \tilde{P}_2) = \tilde{V}(x_0, \mathbf{u}^*(\tilde{P}_2), \tilde{P}_2) \geq \tilde{V}(x_0, \mathbf{u}^*(\tilde{P}_2), \tilde{P}_1) \geq \tilde{V}(x_0, \mathbf{u}^*(\tilde{P}_1), \tilde{P}_1) = V(x_0, \tilde{P}_1)$. ■

Remark 12. Note that Lemma 11 holds if Q and R are time-varying. Let $\tilde{P}_1^o, \tilde{P}_2^o$ be the matrices satisfying $V(x_0, \tilde{P}_1) =: x_0^\top \tilde{P}_1^o x_0$ and $V(x_0, \tilde{P}_2) =: x_0^\top \tilde{P}_2^o x_0$ for all $x_0 \in \mathbb{R}^{n_x}$, then Lemma 11 implies $\tilde{P}_1^o \preceq \tilde{P}_2^o$. Therefore the monotonicity 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], [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 $k = 1, 2, \dots$

(1) For all $\mu \in (0, 1]$, define

$$\underline{\mathbf{c}}(x_k, \mu) := \arg \min_{\mathbf{c}} \begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix}^\top W_1(L_\infty(\mu)) \begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix}, \quad (48)$$

where W_1 is defined in (19);

(2) Compute

$$\tilde{\mu} :=$$

$$\begin{aligned} & \max \left\{ \arg \max_{\mu \in \{\mu_i\}_{i=1}^m} \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k, \mu) \end{bmatrix}^\top W_1(L_\infty(\mu)) \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k, \mu) \end{bmatrix} + (\star) \right. \\ & \left. \text{s.t. } \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k, \mu) \end{bmatrix}^\top W_1(L_\infty(\mu)) \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k, \mu) \end{bmatrix} + (\star) \leq \varepsilon_k \right\}, \quad (49) \end{aligned}$$

where \star here denotes $\frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_\infty(\mu))$;

(3) Compute

$$\bar{\mu}_k :=$$

$$\begin{aligned} & \max \left\{ \arg \min_{\mu \in \{\mu_i\}_{i=1}^m} \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k, \mu) \end{bmatrix}^\top W_2(L_\infty(\mu)) \begin{bmatrix} x_k \\ \underline{\mathbf{c}}(x_k, \mu) \end{bmatrix} \right. \\ & \left. \text{s.t. } \tilde{\mu} \geq \mu \geq \bar{\mu}_{k-1} \right\}, \quad (50) \end{aligned}$$

where W_2 is defined in (14);

(4) Set $K_k := L_\infty(\bar{\mu}_k)$, $\tilde{P}_k := \tilde{P}_\infty(\bar{\mu}_k)$ and $P_k := \hat{P}_\infty(\bar{\mu}_k)$.

Comparing step (1) of Methods 1 and 2, here $L_\infty(\mu)$ replaces K_{k-1} and $\underline{\mathbf{c}}(x_k, \mu)$ is a function of both x_k and μ . Recalling the equivalence between (17) and (18), we have that

$$\begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix}^\top W_1(L_\infty(\mu)) \begin{bmatrix} x_k \\ \mathbf{c} \end{bmatrix} = \sum_{i=0}^{N-1} \gamma^i \|C \bar{x}_{i|k}\|^2 + \gamma^N \|\bar{x}_{N|k}\|_{\tilde{P}_\infty(\mu)}^2,$$

where $\bar{x}_{i|k}$ evolves according to $\bar{\mathbf{x}}_k = M_x((L_\infty(\mu))x_k + M_c((L_\infty(\mu))\mathbf{c}))$. Therefore, by Lemmas 6 and 11, the objective in the maximisation (49) is monotonic in μ and step (2) can be performed as a binary search. The constraint in (49) makes the constraint (18) as tight as possible with $K_k = L_\infty(\tilde{\mu})$, and the constraints of (50) imply that (20) is recursively feasible with Method 2. Note that the initial value $\bar{\mu}_0$ is not determined

by Method 2. Therefore we choose $\bar{\mu}_0 \in \{\mu_i\}_{i=1}^m$ and $K_0 = L_\infty(\bar{\mu}_0)$ to make the MPC optimisation (20) initially feasible (if such $\bar{\mu}_0$ 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 μ , whereas only second term on the LHS of the constraint in problem (42) increases with μ and the first term is fixed. It follows that $\bar{\mu}_k$ 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 $\bar{\mu}_k$ obtained by Method 2. The sequence $\{\bar{\mu}_k\}_{k=0}^\infty$ 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 γ in (3) is close to 1, indicating that performance over a long horizon is a priority, then Method 1 should be chosen. Alternatively if γ 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 μ) requires an additional search over $\{\mu_i\}_{i=1}^m$.

Remark 13. Since $\bar{\mu}_{k-1}$ is a feasible solution to the optimisation problem (50) at time k and $\bar{\mu}_k$ is upper bounded by 1 for all k , the sequence $\{\bar{\mu}_k\}_{k=0}^\infty$ generated by Method 2 is monotonically non-decreasing 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 $\{\bar{\mu}_k\}_{k=0}^\infty$ and hence closed loop performance, since large gaps between successive elements of the sequence $\{\mu_i\}_{i=1}^m$ reduce the likelihood of convergence $\bar{\mu}_k \rightarrow 1$. 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 worse closed loop performance, it is desirable to choose μ_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 μ_1 equal to a value $\bar{\mu}_0$ that makes the MPC optimisation 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 non-increasing 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 non-increasing if ε_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 optimisation problem defined in (20).

Algorithm 1. (SMPC Algorithm) At each time-step $k = 0, 1, \dots$:

- (i). Measure x_k ;
- (ii). If $k > 0$, compute ε_k using (26) and determine K_k using Method 1 or 2;
- (iii). Solve the quadratically constrained quadratic programming (QCQP) problem (20);
- (iv). Apply the control law (23).

We choose $\varepsilon_0 = e$ in step (ii) as will be explained in Section VII. In step (iii), the MPC optimisation can be solved efficiently since there is only one quadratic constraint in (20), for example using a second-order conic program (SOCP) 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

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^{T-1} \mathbb{E}\{\|x_k\|_Q^2 + \|u_k\|_R^2\} \leq \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^{T-1} \text{tr}(\Omega \mathbb{E}\{P_k\}) \leq \text{tr}(\Omega P_0) \quad (51)$$

provided P_k satisfies (13).

Stability is the overriding requirement and in most MPC literature the optimal predicted cost is chosen as a Lyapunov function suitable for analysing closed loop stability [1]. Theorem 14 is proved via cost comparison, building a connection between the cost of a feasible solution at time $k + 1$ and the optimal cost value at time k . Similar asymptotic bounds on the time average of a quadratic expected stage cost are obtained in [2], [40] and [15]. However, in the current context, Theorem 14 demonstrates that an MPC algorithm can ensure closed loop stability *without* imposing terminal constraints based on an invariant set.

VII. THE BEHAVIOUR 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 equation relating the expected value of ε_{k+1} to x_k and ε_k by using the explicit expression for ε_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 ε_k is initialised with $\varepsilon_0 = e$.

The following result gives the relationship between ε_k and the expected value of ε_{k+1} for the closed loop system.

Lemma 15. Given initial feasibility of problem (20) at time $k = 0$, if ε_k is defined by (26) at all times $k \geq 1$ and \bar{P}_k

satisfies (16) for all $k \geq 0$, then in closed loop operation under Algorithm 1 we have

$$\gamma \mathbb{E}_k \{\varepsilon_{k+1}\} \leq \varepsilon_k - \|Cx_k\|^2, \quad \forall k \geq 0. \quad (52)$$

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 Theorem 6 in [29] 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 ε_k can increase as well as decrease. In fact, for values of γ close to zero, a rapid initial growth in ε_k is to be expected, which is in agreement with the interpretation that the constraint (3) penalises *violation probabilities* to a much lesser extent from the initial time in this case. On the other hand, for values of γ close to 1, ε_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: (i) 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; (ii) 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; (iii) gain selection Method 2 provides better closed loop performance over short time intervals than Method 1; (iv) K_k converges to K_{LQ} with high probability for Method 1; (v) 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 linearised model derived from the continuous-time model of a coupled-tank system [41] with a sampling interval of 0.05 sec. 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 ω_k has a multivariate Laplace distribution with zero mean and covariance $\Omega = \mathbf{I}$. 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 $Q = R = \mathbf{I}$. We choose a prediction horizon $N = 10$, and the initial value for ε_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 $\mathcal{K} = \{L_{\infty}(\mu_i)\}_{i=1}^m$ offline and store the sets of matrices $\{\bar{P}_{\infty}(\mu_i)\}_{i=1}^m$, $\{\hat{P}_{\infty}(\mu_i)\}_{i=1}^m$. We choose $\bar{\mu}_0 = 10^{-15}$ 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 $K_k = K_{LQ}$ is infeasible for the online MPC optimisation (20) at time $k = 0$.

Simulation A (demonstrating (i) and (ii)): To estimate empirically the average cost, denoted J_{average} , and the discounted sum of *violation probabilities*, denoted $P_{\text{violation}}$, we run 10^4 simulations, each of which has a length of 10^4 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 summarised in Table I.

	$K_k = K_0$	Method 1	Method 2	
J_{average}	639.9	8.4	96.2	$\text{tr}(\Omega P_0) = 640.0$
$P_{\text{violation}}$	0.448	0.914	0.769	$e = 1.5$

TABLE I: Average costs and violation rates for Simulation A

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 $\gamma^{150} \approx 1.37 \times 10^{-7}$ so the number of violations occurring at time steps $k > 150$ has negligible effect on the estimate of $P_{\text{violation}}$. Also, cost estimates obtained using the dynamic gain selection methods are considerably smaller than that obtained using a fixed feedback gain. Estimates for $P_{\text{violation}}$ 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 $\bar{\mu}_k$ to 1 (and hence convergence of K_k to K_{LQ}) 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 minimise upper bounds on optimal predicted costs and their impact on increasing constraint violation may be limited.

Simulation B (demonstrating (iii)): 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.

	$K_k = K_0$	Method 1	Method 2
J_{average}	807.2	575.9	500.8

TABLE II: Average costs for Simulation B

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 prioritised. This is because $\bar{\mu}_k$ 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 behaviours. We show an example of sequences $\{\bar{\mu}_k\}_{k=0}^\infty$ generated by Methods 1 and 2 in Figure 1, where Figure 1(b) plots the two sequences up to time $k = 150$ and Figure 1(a) zooms in on the initial 40 time steps. These two figures show that $\bar{\mu}_k$ 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 (iv)): 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 (v)): Minimising the LHS of (18) over \mathbf{c}_k yields the largest set of feasible initial conditions,

$$\mathcal{X}_0(\mu) := \{x : x^\top (W_{xx}(\mu) - W_{xc}(\mu)W_{cc}^\dagger(\mu)W_{cx}(\mu))x + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \bar{P}_\infty(\mu)) \leq \varepsilon_0\},$$

where $W_{xx}(\mu)$, $W_{xc}(\mu)$, $W_{cx}(\mu)$, $W_{cc}(\mu)$ are the blocks of $W_1(L_\infty(\mu))$ in (19). The feasible set \mathcal{X}_0 is plotted for μ taking values of $\mu_1 = 10^{-15}$, $\mu_2 = 10^{-4}$, $\mu_3 = 2.5 \times 10^{-4}$ in Figure 2. Clearly the feasible set is enlarged as μ 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 $N = 5$, $Q = \begin{bmatrix} 0 & 0 \\ 0 & 0.5 \end{bmatrix}$, $R = \begin{bmatrix} 1 & 0 \\ 0 & 0.01 \end{bmatrix}$. 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 $\bar{\mu}_0$ 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 parameterise control inputs and the FHSCP has NK ($= 55$) constraints on predicted states. Then this algorithm ensures that

$$\sum_{k=0}^{200} \gamma^k \mathbb{P} \{\|Cx_k\| \geq 1\} \leq e$$

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 summarised in Table III.

	[6] with ADF	Method 1	Method 2
J_{average}	62.8651	54.8299	75.2514
$P_{\text{violation}}$	1.2971	0.7053	0.6689

TABLE III: Average costs and violation rates for Simulation E

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 2.20GHz Intel Core i7-8750H CPU and 16GB RAM, and the online MPC optimisation (20) is solved using the root-finding algorithm proposed in [39]. Each online MPC optimisation is solved within 1 millisecond (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 optimisation. 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 optimisation. The maximum time observed for implementation of Gain selection Method 2 is 70 ms. After $\bar{\mu}_k$ has converged to 1, no gain selection method is executed, so no additional online computation other than the MPC optimisation is required.

IX. CONCLUSION

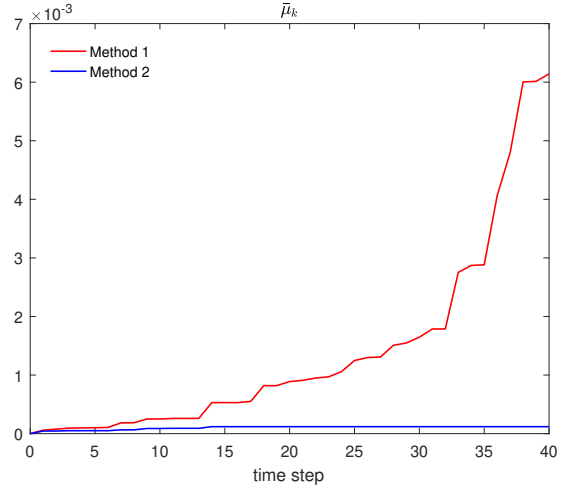
A stochastic MPC algorithm is proposed that enforces a constraint on the sum of discounted future *violation probabilities* and ensures recursive feasibility of the online optimisation and closed loop constraint satisfaction. Key features are the design of a constraint-tightening procedure and dynamic 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 paper on recursive feasibility of the MPC optimisation, closed loop stability and constraint satisfaction can be extended to problems with multiple chance constraints, although solving the corresponding multiobjective optimisation problems to generate stabilising feedback gains for dynamic gain selection becomes more challenging in this case.

APPENDIX A

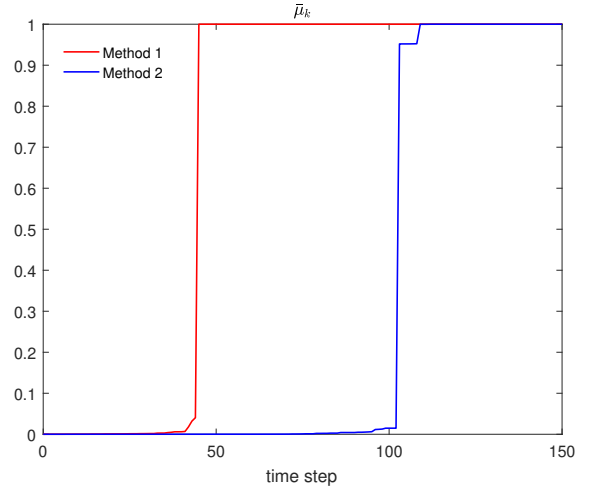
A. Proof of Lemma 1

Equation (16) implies that $\tilde{P}_k = \sum_{i=0}^{\infty} \gamma^i (\Phi_k^\top)^i C^\top C \Phi_k^i$. Combining this with (8) shows that

$$\sum_{i=N}^{\infty} \gamma^i \|C \bar{x}_{i|k}\|^2 = \gamma^N \bar{x}_{N|k}^\top \left(\sum_{j=0}^{\infty} (\gamma^{\frac{1}{2}} \Phi_k^\top)^j C^\top C (\gamma^{\frac{1}{2}} \Phi_k)^j \right) \bar{x}_{N|k}$$



(a) Short simulation horizon



(b) Long simulation horizon

Fig. 1: Evolution of $\bar{\mu}_k$ in Simulation B

$$= \gamma^N \|\bar{x}_{N|k}\|_{\tilde{P}_k}^2.$$

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

$$\begin{aligned} \gamma \Phi_k \tilde{S}_k \Phi_k^\top &= \sum_{i=0}^{\infty} \gamma^{i+1} \Phi_k X_{i|k} \Phi_k^\top = \sum_{i=0}^{\infty} \gamma^{i+1} (X_{i+1|k} - \Omega) \\ &= \tilde{S}_k - X_{0|k} - \frac{\gamma}{1-\gamma} \Omega, \end{aligned}$$

and \tilde{S}_k satisfies the Lyapunov equation

$$\tilde{S}_k = \gamma \Phi_k \tilde{S}_k \Phi_k^\top + \frac{\gamma}{1-\gamma} \Omega.$$

Therefore,

$$\begin{aligned} \text{tr} \left(\sum_{i=0}^{\infty} \gamma^i C^\top C X_{i|k} \right) &= \text{tr} (C^\top C \tilde{S}_k) \\ &= \text{tr} \left(C^\top C \sum_{i=0}^{\infty} \gamma^i \Phi_k^i \left(\frac{\gamma}{1-\gamma} \Omega \right) (\Phi_k^\top)^i \right) = \frac{\gamma}{1-\gamma} \text{tr} (\Omega \tilde{P}_k). \end{aligned}$$

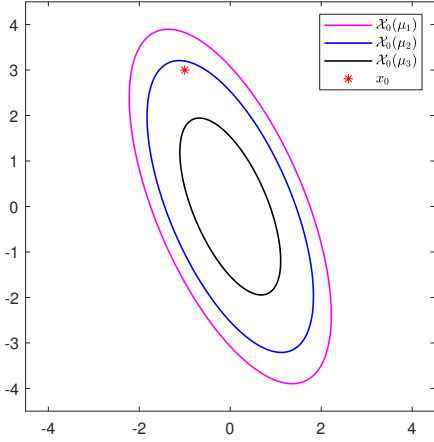


Fig. 2: Largest feasible sets of initial conditions

B. Proof of Lemma 2

If $K_k = K_{k-1}$ at time $k \geq 1$, then $\tilde{P}_k = \tilde{P}_{k-1}$. From equations (18), (26) and $\tilde{x}_{0|k} = x_k$, 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 $\tilde{P}_k \neq \tilde{P}_{k-1}$, we can show that the guarantee of recursive feasibility is retained as follows. If gain selection method 1 is used, $\bar{\mu}_k$ defined in step (2) of Method 1 satisfies the constraint in problem (42), which, together with equation (44), implies a feasible solution to problem (20) exists at time k . If gain selection method 2 is used, $\bar{\mu}_k$ defined in step (3) of Method 2 satisfies $\bar{\mu}_k \leq \tilde{\mu}$ 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 (\mathcal{P}_2) , the optimality of $x_0^\top((1-\mu)\bar{P}_\infty(\mu) + \mu\hat{P}_\infty(\mu))x_0$ for all $x_0 \in \mathbb{R}^{n_x}$ implies

$$\begin{aligned} (1-\mu_1)\bar{P}_\infty(\mu_1) + \mu_1\hat{P}_\infty(\mu_1) &\preceq (1-\mu_1)\bar{P}_\infty(\mu_2) + \mu_1\hat{P}_\infty(\mu_2), \\ (1-\mu_2)\bar{P}_\infty(\mu_2) + \mu_2\hat{P}_\infty(\mu_2) &\preceq (1-\mu_2)\bar{P}_\infty(\mu_1) + \mu_2\hat{P}_\infty(\mu_1). \end{aligned}$$

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

$$\begin{aligned} -(1-\mu_1)\Delta\bar{P} &\preceq \mu_1\Delta\hat{P}, \\ (1-\mu_2)\Delta\bar{P} &\preceq -\mu_2\Delta\hat{P}. \end{aligned} \quad (53)$$

Combining these inequalities yields $\Delta\bar{P} \succeq 0$ and $\Delta\hat{P} \preceq 0$.

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 $\Delta^2\bar{P} = \bar{P}_\infty(\mu_3) - 2\bar{P}_\infty(\mu_2) + \bar{P}_\infty(\mu_1)$, $\Delta^2\hat{P} = \hat{P}_\infty(\mu_3) - 2\hat{P}_\infty(\mu_2) + \hat{P}_\infty(\mu_1)$ and $\Delta^2S = S(\mu_3) - 2S(\mu_2) + S(\mu_1)$. Then by definition we have $\Delta^2S = (1-\mu_3)\Delta^2\bar{P} + \mu_3\Delta^2\hat{P} - (\mu_3 - \mu_1)(\Delta\bar{P} - \Delta\hat{P})$, where $\Delta\bar{P} = \bar{P}_\infty(\mu_2) - \bar{P}_\infty(\mu_1)$, $\Delta\hat{P} = \hat{P}_\infty(\mu_2) - \hat{P}_\infty(\mu_1)$. But the optimality property of

$\text{tr}(ZS(\mu))$ for all $Z = Z^\top \succeq 0$ implies $(1-\mu_3)\bar{P}_\infty(\mu_3) + \mu_3\hat{P}_\infty(\mu_3) \preceq (1-\mu_3)\bar{P}_\infty(\mu_2) + \mu_3\hat{P}_\infty(\mu_2)$, so

$$(1-\mu_3)(\bar{P}_\infty(\mu_3) - \bar{P}_\infty(\mu_2)) \preceq -\mu_3(\hat{P}_\infty(\mu_3) - \hat{P}_\infty(\mu_2)).$$

Combining this inequality with (53), we obtain

$$(1-\mu_3)\Delta^2\bar{P} + \mu_3\Delta^2\hat{P} - (\mu_3 - \mu_1)(\Delta\bar{P} - \Delta\hat{P}) \preceq 0,$$

which implies $\Delta^2S \preceq 0$ and hence $S(\mu)$ is midpoint concave for all $\mu \in (0, 1]$. Furthermore, $S(\mu)$ is positive definite (bounded below) for all $\mu \in (0, 1]$. Therefore, $S(\cdot)$ is continuous and hence concave on $(0, 1)$ in the sense that $\text{tr}(ZS(\cdot))$ is concave on $(0, 1)$ for all $Z = Z^\top \succeq 0$ [42, Section 72].

E. Proof of Lemma 8

Let $\Sigma(\mu) := \gamma(1-\mu)\bar{P}_\infty(\mu) + \mu\hat{P}_\infty(\mu)$, and $\Phi(\mu) := A + BL_\infty(\mu)$ 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 $\Delta\Phi$, $\Delta\Sigma$, ΔL , $\Delta\bar{P}$, $\Delta\hat{P}$ denote the corresponding changes in the values of Φ , Σ , L_∞ , \bar{P}_∞ and \hat{P}_∞ respectively, so that $\Delta\Phi = \Phi(\mu + \delta\mu) - \Phi(\mu)$, $\Delta\Sigma = \Sigma(\mu + \delta\mu) - \Sigma(\mu)$, $\Delta L = L_\infty(\mu + \delta\mu) - L_\infty(\mu)$, $\Delta\bar{P} = \bar{P}_\infty(\mu + \delta\mu) - \bar{P}_\infty(\mu)$, and $\Delta\hat{P} = \hat{P}_\infty(\mu + \delta\mu) - \hat{P}_\infty(\mu)$. To simplify notation we omit the argument μ for the remainder of the proof (so that $\Phi = \Phi(\mu)$, $\Sigma = \Sigma(\mu)$, $L_\infty = L_\infty(\mu)$, etc.). Lemma 7 implies that $(1-\mu)\bar{P}_\infty + \mu\hat{P}_\infty$ is Lipschitz continuous on $(0, 1)$, so $(1-\mu)\Delta\bar{P} + \mu\Delta\hat{P} - \delta\mu(\bar{P}_\infty(\mu + \delta\mu) - \hat{P}_\infty(\mu + \delta\mu)) = O(\delta\mu)$ and

$$\begin{aligned} O(\delta\mu) &= (1-\mu)\Delta\bar{P} + \mu\Delta\hat{P} = (\Phi + \Delta\Phi)^\top(\Sigma + \Delta\Sigma)(\Phi + \Delta\Phi) \\ &\quad - \Phi^\top\Sigma\Phi + \mu(L_\infty + \Delta L)^\top R(L_\infty + \Delta L) - \mu L_\infty^\top R L_\infty. \end{aligned} \quad (54)$$

But $L_\infty + \Delta L = -[(\mu + \delta\mu)R + B^\top(\Sigma + \Delta\Sigma)B]^{-1}B^\top(\Sigma + \Delta\Sigma)A$ implies

$$\begin{aligned} L_\infty + \Delta L &= L_\infty - \Gamma^{-1}(B^\top\Delta\Sigma\Phi + RL_\infty\delta\mu) \\ &= L_\infty - \Gamma^{-1}B^\top\Delta\Sigma\Phi + O(\delta\mu), \end{aligned} \quad (55)$$

where $\Gamma = (\mu + \delta\mu)R + B^\top(\Sigma + \Delta\Sigma)B$, and

$$\Phi + \Delta\Phi = \Phi - B\Gamma^{-1}B^\top\Delta\Sigma\Phi + O(\delta\mu). \quad (56)$$

Combining (55) and (56) with (54) and using $L_\infty = -(\mu R + B^\top\Sigma B)^{-1}B^\top\Sigma A$, we obtain

$$\begin{aligned} O(\delta\mu) &= \Phi^\top\Delta\Sigma\Phi \\ &\quad + \Phi^\top\Delta\Sigma B\Gamma^{-1}(B^\top\Sigma B + \mu R - 2\Gamma)\Gamma^{-1}B^\top\Delta\Sigma\Phi \\ &\quad + \Phi^\top\Delta\Sigma B\Gamma^{-1}B^\top\Delta\Sigma B\Gamma^{-1}B^\top\Delta\Sigma\Phi \\ &= \Phi^\top\Delta\Sigma\Phi - \Phi^\top\Delta\Sigma B\Gamma^{-1}B^\top\Delta\Sigma\Phi \\ &\quad - \Phi^\top\Delta\Sigma B\Gamma^{-1}R\Gamma^{-1}B^\top\Delta\Sigma\Phi\delta\mu, \end{aligned}$$

and hence

$$\Phi^\top\Delta\Sigma\Phi - \Phi^\top\Delta\Sigma B\Gamma^{-1}B^\top\Delta\Sigma\Phi = O(\delta\mu).$$

But $\Sigma = (1-\mu)\bar{P}_\infty + \mu\hat{P}_\infty - (1-\gamma)(1-\mu)\bar{P}_\infty$ implies $\Delta\Sigma = O(\delta\mu) - (1-\gamma)(1-\mu)\Delta\bar{P}$, so

$$(1-\gamma)(1-\mu)\Phi^\top[\Delta\bar{P} + (1-\gamma)(1-\mu)\Delta\bar{P}B\Gamma^{-1}B^\top\Delta\bar{P}]\Phi = O(\delta\mu).$$

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

F. 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 $z \in \mathbb{R}^{Nn_u}$ and $\mu \in [\bar{\mu}_0, 1]$ we must have

$$\mathbb{P} \left\{ \|z + (W_{cc}(\mu))^\dagger W_{cx}(\mu) \omega_k\|_{W_{cc}(\mu)}^2 \geq \delta \right\} < p_\delta \quad (57)$$

for all $\delta > 0$ and all $p_\delta > 0$, which implies that $W_{cx}(\mu) \omega_k + W_{cc}(\mu)z = 0$ with probability 1. Here $W_{cx}(\mu)$ must be non-zero since otherwise $\mathbf{c}^o(x_k) = 0$, implying that by its definition in (41) the feedback gain used is optimal with respect to $\sum_{i=0}^\infty \gamma^i \|C \bar{x}_i\|_Q^2$ and then $L_\infty(\mu) = L_\infty(0)$ would be obtained, which contradicts the fact that $\mu \geq \bar{\mu}_0 \geq \mu_1$ and the statement that $L_\infty(\mu) \neq L_\infty(0) \forall \mu \geq \mu_1$ in Section IV. Furthermore, z and μ are by assumption independent of the realisation ω_k . Therefore (57) contradicts the assumption in (2) that $\mathbb{E}\{\omega_k \omega_k^\top\}$ is positive definite implying ω_k is not a constant vector or in any subspaces of \mathbb{R}^{n_x} , and hence (46) must hold for some $\delta > 0$ and $p_\delta > 0$.

G. Proof of Theorem 10

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

$$N_f = \left\lceil \frac{\gamma}{1 - \gamma} \frac{\text{tr}(\Omega \bar{P}_\infty(1) - \Omega \bar{P}_\infty(\bar{\mu}_0))}{\delta} \right\rceil$$

where it is assumed that $\bar{\mu}_0 < 1$, and let \mathcal{E}_k denote the event that

$$\|z_{k-1} + (W_{cc}(\bar{\mu}_{k-1}))^\dagger W_{cx}(\bar{\mu}_{k-1}) \omega_{k-1}\|_{W_{cc}(\bar{\mu}_{k-1})}^2 \geq \delta.$$

Under Assumption 2, Lemma 8 implies that before $\bar{\mu}_k$ reaches 1, (43b) is satisfied with equality and every possible increment in $\bar{\mu}_k$ is attained if (45) is positive. It follows that $\bar{\mu}_{N_f} = 1$ if \mathcal{E}_k occurs for $k = 1, \dots, N_f$. Furthermore,

$$\begin{aligned} & \mathbb{P}\{\mathcal{E}_1 \cap \dots \cap \mathcal{E}_{N_f}\} \\ &= \mathbb{P}\{\mathcal{E}_1\} \mathbb{P}\{\mathcal{E}_2 | \mathcal{E}_1\} \dots \mathbb{P}\{\mathcal{E}_{N_f} | \mathcal{E}_1 \cap \mathcal{E}_2 \cap \dots \cap \mathcal{E}_{N_f-1}\}, \end{aligned}$$

where $\mathbb{P}\{\mathcal{E}_1\} \geq p_\delta$ and

$$\begin{aligned} & \mathbb{P}\{\mathcal{E}_k | \mathcal{E}_1 \cap \dots \cap \mathcal{E}_{k-1}\} \\ & \geq \inf_{z \in \mathbb{R}^{Nn_u}, \mu \in [\bar{\mu}_0, 1]} \mathbb{P}\left\{ \|z + (W_{cc}(\mu))^\dagger W_{cx}(\mu) \omega_{k-1}\|_{W_{cc}(\mu)}^2 \geq \delta \right\} \\ & \geq p_\delta \end{aligned}$$

for $k = 2, \dots, N_f$. Also, these conditional probabilities are well defined by Lemma 9. Therefore,

$$\mathbb{P}\{\bar{\mu}_{N_f} < 1\} \leq 1 - p_\delta^{N_f},$$

and more specifically we rewrite it as

$$\mathbb{P}\{\bar{\mu}_{N_f} < 1 | \bar{\mu}_0 < 1\} \leq 1 - p_\delta^{N_f}.$$

Similarly, given $\bar{\mu}_{(i-1)N_f} < 1$ for $i = 2, \dots$, we have

$$\begin{aligned} \mathbb{P}\{\bar{\mu}_{iN_f} = 1\} & \geq \left(\inf_{z \in \mathbb{R}^{Nn_u}, \mu \in [\bar{\mu}_0, 1]} \mathbb{P}\{\star \geq \delta_i\} \right)^{N_f} \\ & \geq \left(\inf_{z \in \mathbb{R}^{Nn_u}, \mu \in [\bar{\mu}_0, 1]} \mathbb{P}\{\star \geq \delta\} \right)^{N_f} \geq p_\delta^{N_f}, \end{aligned}$$

where \star here denotes $\|z + (W_{cc}(\mu))^\dagger W_{cx}(\mu) \omega_{iN_f-1}\|_{W_{cc}(\mu)}$, and δ_i is the counterpart of δ for $\bar{\mu}_{iN_f}$ that is sufficient for $\bar{\mu}_k$ to reach 1 within N_f time steps from time $k = (i-1)N_f$ and is smaller than δ since the sequence $\{\bar{\mu}_k\}_{k=0}^\infty$ is monotonically non-decreasing. Then

$$\mathbb{P}\{\bar{\mu}_{iN_f} < 1 | \bar{\mu}_{(i-1)N_f} < 1\} \leq 1 - p_\delta^{N_f}, \quad \forall i = 2, \dots$$

For given $k > N_f$, we choose integers $j_1, j_2, \dots, j_{\lfloor k/N_f \rfloor}$ so that $j_1 = N_f$, $j_{\lfloor k/N_f \rfloor} \leq k$ and $j_{i+1} - j_i = N_f$ for all i . Then $\bar{\mu}_k < 1$ only if $\bar{\mu}_{j_i} < 1$ for all $i = 1, \dots, \lfloor k/N_f \rfloor$ and hence

$$\begin{aligned} & \mathbb{P}\{\bar{\mu}_k < 1\} \\ & \leq \mathbb{P}\{\bar{\mu}_{j_1} < 1\} \mathbb{P}\{\bar{\mu}_{j_2} < 1 | \bar{\mu}_{j_1} < 1\} \dots \mathbb{P}\{\bar{\mu}_{j_{\lfloor k/N_f \rfloor}} < 1 | \bar{\mu}_{j_{\lfloor k/N_f \rfloor-1}} < 1\} \\ & \leq \left(1 - p_\delta^{N_f}\right)^{\lfloor k/N_f \rfloor}. \end{aligned} \quad (58)$$

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

$$\sum_{k=1}^\infty \mathbb{P}\{\mathcal{F}_k\} \leq \sum_{k=1}^\infty (1 - p_\delta^{N_f})^{\lfloor k/N_f \rfloor} = N_f p_\delta^{-N_f} - 1 < \infty.$$

Therefore the Borel-Cantelli lemma implies that $\mathbb{P}\left\{\bigcap_{k=1}^\infty \bigcup_{j=k}^\infty \mathcal{F}_j\right\} = 0$. But $\bar{\mu}_k < 1$ only if $\bar{\mu}_{k-1} < 1$, so $\mathcal{F}_k \subseteq \mathcal{F}_{k-1}$ and $\bigcup_{j=k}^\infty \mathcal{F}_j = \mathcal{F}_k$. Therefore, $\mathbb{P}\{\lim_{k \rightarrow \infty} \mathcal{F}_k\} = 0$ or equivalently $\mathbb{P}\{\lim_{k \rightarrow \infty} \bar{\mu}_k = 1\} = 1$.

H. Proof of Theorem 14

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

$$J^*(x_{k+1}, K_{k+1}) \leq J(x_{k+1}, K_k, \tilde{\mathbf{c}}_{k+1}),$$

and since this inequality holds for every realisation of ω_k , by taking expectations conditioned on the state x_k we obtain

$$\mathbb{E}_k\{J^*(x_{k+1}, K_{k+1})\} \leq \mathbb{E}_k\{J(x_{k+1}, K_k, \tilde{\mathbf{c}}_{k+1})\}. \quad (59)$$

From (28) we have the feasible sequence

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

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

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

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 ω_k . Moreover, the second inequality in (51) follows from the choice of K_k in step (ii) of Algorithm 1 and Lemma 6.

I. Proof of Lemma 15

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

$$\begin{aligned} \varepsilon_{k+1} &= \sum_{i=0}^{N-1} \gamma^i \|C \bar{x}_{i+1|k}^*\|^2 + \gamma^N \|\bar{x}_{N+1|k}^*\|_{\tilde{P}_k}^2 + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k) \\ &+ \sum_{i=0}^{N-1} \gamma^i \|C \Phi_k^i \omega_k\|^2 + \gamma^N \|\Phi_k^N \omega_k\|_{\tilde{P}_k}^2 \\ &+ \sum_{i=0}^{N-1} 2\gamma^i (\Phi_k^i \omega_k)^\top C^\top C \bar{x}_{i+1|k}^* + 2\gamma^N (\Phi_k^N \omega_k)^\top \tilde{P}_k \bar{x}_{N+1|k}^*, \end{aligned} \quad (61)$$

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

$$\begin{aligned} \gamma^{-1} \left(\sum_{i=0}^{N-1} \gamma^i \|C \bar{x}_{i|k}^*\|^2 + \gamma^N \|\bar{x}_{N|k}^*\|_{\tilde{P}_k}^2 + \right. \\ \left. \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k) - \|C x_k\|^2 \right) - \text{tr}(\Omega \tilde{P}_k), \end{aligned}$$

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

$$\begin{aligned} \gamma \mathbb{E}_k \{\varepsilon_{k+1}\} &= \\ &\sum_{i=0}^{N-1} \gamma^i \|C \bar{x}_{i|k}^*\|^2 + \gamma^N \|\bar{x}_{N|k}^*\|_{\tilde{P}_k}^2 + \frac{\gamma}{1-\gamma} \text{tr}(\Omega \tilde{P}_k) - \|C x_k\|^2. \end{aligned}$$

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

REFERENCES

- [1] D. Mayne, J. Rawlings, C. Rao, and P. Scoekaert, "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. Giulioni, L. Magni, and R. Scattolini, "A probabilistic approach to model predictive control," in *IEEE Conference on Decision and Control*, 2013, pp. 7734–7739.
- [4] T. Hashimoto, "Probabilistic constrained model predictive control for linear discrete-time systems with additive stochastic disturbances," in *IEEE Conference on Decision and 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 Transactions on Automatic Control*, vol. 64, no. 9, pp. 3883–3888, 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, pp. 1706–1721, 2014.
- [9] M. Lorenzen, F. Dabbene, R. Tempo, and F. Allgöwer, "Constraint-tightening and stability in stochastic model predictive control," *IEEE Trans. Autom. Control*, vol. 62, pp. 3165–3177, 2017.
- [10] J. Fleming and M. Cannon, "Time-average constraints in stochastic Model Predictive Control," in *American Control Conference*, 2017, pp. 5648–5653.
- [11] D. Muñoz-Carpintero, G. Hu, and C. J. Spanos, "Stochastic model predictive control with adaptive constraint tightening for non-conservative chance constraints satisfaction," *Automatica*, vol. 96, pp. 32–39, 2018.
- [12] P. Hokayem, D. Chatterjee, and J. Lygeros, "On stochastic receding horizon control with bounded control inputs," in *Proceedings of the 48th IEEE Conference on Decision and Control held jointly with 2009 28th Chinese Control Conference*. IEEE, 2009, pp. 6359–6364.
- [13] L. Hewing and M. N. Zeilinger, "Stochastic model predictive control for linear systems using probabilistic reachable sets," in *2018 IEEE Conference on Decision and Control*. IEEE, 2018, pp. 5182–5188.
- [14] L. Hewing, K. P. Wabersich, and M. N. Zeilinger, "Recursively feasible stochastic model predictive control using indirect feedback," *Automatica*, vol. 119, p. 109095, 2020.
- [15] D. Chatterjee and J. Lygeros, "On stability and performance of stochastic predictive control techniques," *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 509–514, 2015.
- [16] M. Lorenzen, M. A. Müller, and F. Allgöwer, "Stochastic model predictive control without terminal constraints," *International Journal of Robust and Nonlinear Control*, vol. 29, no. 15, pp. 4987–5001, 2019.
- [17] D. Limón, T. Alamo, F. Salas, and E. F. Camacho, "On the stability of constrained mpc without terminal constraint," *IEEE transactions on automatic control*, vol. 51, no. 5, pp. 832–836, 2006.
- [18] J. Maciejowski *et al.*, "Robust finite horizon mpc without terminal constraints," in *Proceedings of the 39th IEEE Conference on Decision and Control*, vol. 1. IEEE, 2000, pp. 166–171.
- [19] D. Blackwell, "Discounted dynamic programming," *The Annals of Mathematical Statistics*, vol. 36, no. 1, pp. 226–235, 1965.
- [20] D. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed. Belmont, Massachusetts: Athena Scientific, 2005.
- [21] B. P. G. Van Parys, P. J. Goulart, and M. Morari, "Infinite horizon performance bounds for uncertain constrained systems," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2803–2817, Nov 2013.
- [22] M. Kamgarpour and T. Summers, "On infinite dimensional linear programming approach to stochastic control," in *Proceedings of the IFAC World Congress*, 2017.
- [23] D. P. Bertsekas, *Reinforcement learning and optimal control*. Athena Scientific Belmont, MA, 2019.
- [24] S. Barratt and S. Boyd, "Multi-period liability clearing via convex optimal control," *Available at SSRN 3604618*, 2020.
- [25] P. Nystrup, S. Boyd, E. Lindström, and H. Madsen, "Multi-period portfolio selection with drawdown control," *Annals of Operations Research*, vol. 282, no. 1-2, pp. 245–271, 2019.
- [26] S. P. Boyd, M. T. Mueller, B. O'Donoghue, Y. Wang *et al.*, *Performance bounds and suboptimal policies for multi-period investment*. Now Publishers, 2014.
- [27] C. W. Clark, "Profit maximization and the extinction of animal species," *Journal of Political Economy*, vol. 81, no. 4, pp. 950–961, 1973.
- [28] D. Ludwig, W. A. Brock, and S. R. Carpenter, "Uncertainty in discount models and environmental accounting," *Ecology and Society*, vol. 10, no. 2, 2005.
- [29] S. Yan, P. Goulart, and M. Cannon, "Stochastic model predictive control with discounted probabilistic constraints," in *2018 European Control Conference*, 2018, pp. 1003–1008.

- [30] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd ed. Oxford University Press, 2001.
- [31] B. Kouvaritakis and M. Cannon, *Model Predictive Control: Classical, Robust and Stochastic*. Springer, 2015.
- [32] E. Kerrigan, “Robust constraint satisfaction: Invariant sets and predictive control,” Ph.D. dissertation, University of Cambridge, 2000.
- [33] R. Oberdieck and E. N. Pistikopoulos, “Multi-objective optimization with convex quadratic cost functions: A multi-parametric programming approach,” *Computers and Chemical Engineering*, vol. 85, pp. 36 – 39, 2016.
- [34] R. Bellman, “Dynamic programming,” *Science*, vol. 153, no. 3731, pp. 34–37, 1966.
- [35] K. J. Åström, *Introduction to Stochastic Control Theory*, ser. Dover Books on Electrical Engineering. Dover Publications, 2006.
- [36] S. Bittanti, A. J. Laub, and J. C. Willems, *The Riccati Equation*. Springer Science & Business Media, 2012.
- [37] 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,” *Systems and Control Letters*, vol. 5, pp. 309–315, 1985.
- [38] T. Nishimura, “On the a priori Information in Sequential Estimation Problems,” *IEEE Transactions on Automatic Control*, vol. 11, pp. 197–204, 1966.
- [39] B. Kouvaritakis, M. Cannon, and J. A. Rossiter, “Who needs qp for linear mpc anyway?” *Automatica*, vol. 38, pp. 879–884, 2002.
- [40] M. Cannon, B. Kouvaritakis, and X. Wu, “Probabilistic constrained MPC for multiplicative and additive stochastic uncertainty,” *IEEE Trans. Autom. Control*, vol. 54, pp. 1626–1632, 2009.
- [41] S. T. Lian, K. Marzuki, and Y. Rubiyah, “Tuning of a neuro-fuzzy controller by genetic algorithms with an application to a coupled-tank liquid-level control system,” *Engineering Applications of Artificial Intelligence*, vol. 11, no. 4, pp. 517 – 529, 1998.
- [42] A. W. Roberts and D. E. Varberg, *Convex Functions*, ser. Pure and Applied Mathematics. Academic Press, 1973.



Mark Cannon received the M.Eng. degree in engineering science and the D.Phil. degree in control engineering from the University of Oxford, Oxford, U.K., in 1993 and 1998, respectively, and the M.S. degree in mechanical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1995.

He is currently an Associate Professor of engineering science with the University of Oxford and a fellow of St. Johns College, Oxford. His research interests are in robust and optimal control for constrained and uncertain systems, and optimization, particularly for receding horizon control involving robust constraints and stochastic uncertainty.



Shuhao Yan received the B.Eng. degree in electrical and electronic engineering from City, University of London, U.K., in 2015, the B.Eng. degree in automation from Nanjing University of Aeronautics and Astronautics, China, in 2015 and the M.Sc. degree in control systems from Imperial College London, U.K., in 2016.

He is currently a DPhil Student with the Department of Engineering Science, University of Oxford, U.K. His doctoral research is concerned with model predictive control involving stochastic uncertainty.



Paul Goulart received the B.Sc. and M.Sc. degrees in aeronautics and astronautics from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1998 and 2001, respectively, and the Ph.D. degree in control engineering in 2007 from the University of Cambridge, Cambridge, U.K., where he was selected as a Gates Scholar.

From 2007 to 2011, he was a Lecturer in control systems with the Department of Aeronautics, Imperial College London, and from 2011 to 2014, a Senior Researcher with the Automatic Control Laboratory, ETH Zürich. He is currently an Associate Professor with the Department of Engineering Science, and a Tutorial Fellow with St. Edmund Hall, University of Oxford, Oxford, U.K. His research interests include model predictive control, robust optimization, and control of fluid flows.