Brief paper

Nonquadratic stochastic model predictive control: A tractable approach☆

Milan Korda [a,1], Jiří Cigler [b]

[a] *Automatic Control Laboratory of École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland*
[b] *Department of Control Engineering, Czech Technical University in Prague, Karlovo náměstí 13, 121 35 Prague, Czech Republic*

ARTICLE INFO

ABSTRACT

This paper deals with the finite horizon stochastic optimal control problem with the expectation of the p-norm as the objective function and jointly Gaussian, although not necessarily independent, additive disturbance process. We develop an approximation strategy that solves the problem in a certain class of nonlinear feedback policies while ensuring satisfaction of hard input constraints. A bound on suboptimality of the proposed strategy in this class of nonlinear feedback policies is given for the special case of $p = 1$. We also develop a recursively feasible receding horizon policy with respect to state chance constraints and/or hard control input constraints in the presence of bounded disturbances. The performance of the proposed policies is examined in two numerical examples.

1. Introduction

Stochastic control is a relatively mature field, yet there is still a considerable number of unresolved problems, mostly due to the notorious inherent intractability of the vast majority of them. Only a handful of stochastic optimal control problems (e.g., the linear quadratic control) can be solved optimally, whereas the remainder has to be tackled by various approximation techniques most frequently arising from the dynamic programming paradigm (Bertsekas, 2007).

Recent advances in computation and mathematical optimization techniques have, however, opened new ways of dealing with these problems. One of the simplest, yet in most practical applications very effective approach, is the certainty equivalent model predictive control (CE-MPC) (Bertsekas, 2007) that solves a deterministic optimization problem with stochastic disturbances replaced by their estimates based upon the information available at the time and proceeds in a receding horizon fashion.

Another popular class of control strategies is the affine disturbance feedback policy which turns out to be equivalent to the affine state feedback policy via a nonlinear transformation similar to the classical Q-design (Skaf & Boyd, 2010). However convenient the paradigm of the affine disturbance feedback may be, its use is prohibitive whenever unbounded stochastic disturbances enter the system in the presence of hard control input bounds since then the linear part of the policy necessarily vanishes, which, in effect, renders the policy open loop. One way to overcome this problem is to use a saturated nonlinear disturbance feedback as in Hokayem, Chatterjee, and Lygeros (2009), where this approach was developed for the quadratic cost. In this article, we follow up on this work and on our previous work in the context of the 1-norm (Korda & Cigler, 2011), and develop a methodology for solving this problem in the p-norm, with the additional assumption of the disturbances being jointly Gaussian (but not necessarily independent). Our methodology brings about a significant performance improvement compared to the traditional certainty-equivalent approach while retaining reasonable computational demands compared to sampling or dynamic programming techniques.

The very important, though largely neglected, question of recursive feasibility of stochastic receding horizon schemes is addressed in a series of papers (Cannon, Kouvaritakis, & Ng, 2009; Cannon, Kouvaritakis, & Wu, 2009; Kouvaritakis, Cannon, Raković, & Cheng, 2010). These papers assume either bounded disturbances or only probabilistic input and state constraints, and deal exclusively with the pre-stabilized policy parametrization. In this article, we develop a recursively feasible algorithm for the affine-like policy treated in this paper in the presence of bounded disturbances and hard control input bounds.

There is a wide range of applications amenable to the presented approach that has previously been tackled only through certainty

equivalence. A rich source of such applications is, for instance, building climate control, a typical stochastic environment where cost functions including mixtures of 1-norm and 2-norm terms are ample (Oldewurtel et al., 2010). The p-norm formulation considered here is a natural extension giving the designer more degrees of freedom to either fit the true cost function or to tune the resulting dynamics by choosing possibly non-integer values of p. For example, the choice of p between 1 and 2 or 2 and infinity gives a blend of the control properties typically encountered with these three traditional choices. The use of the p-norm as an approximation to a true cost function was recently considered in the context of max-plus linear systems in Farahani, van den Boom, van der Weide, and De Schutter (2010).

The current paper is organized as follows. In Section 2, we state the problem to be solved, Section 3 presents the main results on convexity and tractability of the stated problem and derives a bound on the suboptimality of the approach for $p = 1$. Recursively feasible algorithm in the presence of bounded disturbances is developed in Section 4. Finally, Section 5 presents two numerical examples to illustrate our results.

1.1. Notation

From less standard notation we denote the Hessian and Jacobian of a function by Hess($\cdot$) and Jac($\cdot$) respectively, and the partial derivatives of a function by subscripts. We let $\|A\|_\infty$ denote the induced infinity norm of a matrix A, i.e., $\|A\|_\infty = \max_i \left(\sum_j |A_{ij}| \right)$. Note that this notation is also used for *row* vectors, where it does *not coincide* with the standard infinity norm of a vector, but rather with the 1-norm. The p-norm of a column vector x is defined as $\|x\|_p = \left(\sum_i |x_i|^p \right)^{1/p}$ for any $p \in [1, \infty)$. The symbols vec($\cdot$) and $\otimes$ denote the vectorization and the Kronecker product, respectively. All the random variables are defined on a common probability space with a probability measure P; the expectation $\mathbf{E}$ is taken with respect to this measure. A Gaussian random variable with the mean μ and the covariance matrix Σ is denoted by $\mathcal{N}(\mu, \Sigma)$.

2. Problem statement

The article deals with the problem of minimizing the cost function

$$J := \mathbf{E} \left\{ \|Q_N x_N\|_p^p + \sum_{k=0}^{N-1} \|Q_k x_k\|_p^p + \|R_k u_k\|_p^p \right\} \quad (1)$$

for $p \in [1, \infty)$ subject to the discrete-time system dynamics

$$x_{k+1} = A x_k + B u_k + w_k, \quad (2)$$

$x_k \in \mathbb{R}^n, u_k \in \mathbb{R}^m$, and hard input constraints

$$\|u_k\|_\infty \le U_{\max}, \quad k = 0, \dots, N-1, \quad (3)$$

where $Q_k \in \mathbb{R}^{n_q \times n}$, $R_k \in \mathbb{R}^{n_r \times m}$ are weighting matrices and $N \ge 1$ is the prediction horizon. All the results derived here generalize with only minor modifications to the case with different bounds on individual control inputs and/or time varying bounds. The disturbances $w = [w_0^T, \dots, w_{N-1}^T]^T$ are assumed to be jointly Gaussian with the covariance matrix Σ_w.

The minimization to be carried out is over all Borel measurable causal disturbance feedback policies $u_k = \phi_k(x_0, w_0, \dots, w_{k-1})$, $k = 0, \dots, N-1$. This problem is, however, in general intractable and various approximation techniques exist; see, e.g., Bertsekas (2007). In this paper, we adopt the approach of Hokayem et al. (2009) where the authors propose to search over a class of causal

policies affine in certain nonlinear functions of the disturbances

$$u = \eta + K\boldsymbol{e}(w)$$

$$= \begin{bmatrix} \eta_0 \\ \eta_1 \\ \vdots \\ \eta_{N-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & \dots & 0 \\ K_{1,1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ K_{N-1,1} & \dots & K_{N-1,N-1} & 0 \end{bmatrix} \begin{bmatrix} e(w_0) \\ e(w_1) \\ \vdots \\ e(w_{N-1}) \end{bmatrix}, \quad (4)$$

where $u = [u_0^T, \dots, u_{N-1}^T]^T$ and $\boldsymbol{e}(w) = [e(w_0)^T, \dots, e(w_{N-1}^T)]^T$. The vector $\eta \in \mathbb{R}^{mN}$ with blocks in $\mathbb{R}^m$, and the strictly lower block triangular matrix $K \in \mathbb{R}^{mN \times nN}$ with blocks in $\mathbb{R}^{m \times n}$ are optimization variables. The choice of the nonlinear function $e : \mathbb{R}^n \to \mathbb{R}^n$ is discussed later, although it certainly must be bounded should the hard input constraints be satisfied. The bound on $\|e(\cdot)\|_\infty$ is denoted by δ throughout the article.

This parametrization of the control policy is a natural way of approximating the infinite dimensional minimization over all measurable causal policies by a finite dimensional problem of minimization over the subspace generated by a finite number of nonlinear basis functions (Skaf & Boyd, 2009). The particular affine-like parametrization (4) can also be viewed as an extension of the traditional affine disturbance feedback allowing for hard control input bounds in the presence of unbounded disturbance (Skaf & Boyd, 2009).

Note that the presented approach can be immediately extended to the case of different bounds on the individual components of the vector-valued function $e(\cdot)$ and/or different functions $e(\cdot)$ at each time step. The former is particularly useful when each component of the disturbance at a given time has different variance, whereas the latter can be beneficial when the variance changes with time over the prediction horizon.

The bound δ on $e(\cdot)$ immediately provides a sufficient condition for the satisfaction of the input constraints

$$|\eta_i| + \delta \|K_i\|_\infty \le U_{\max}, \quad i = 1, \dots, mN \quad (5)$$

where η_i and K_i denote the i-th rows of η and K, respectively. Note that, providing the distribution of w is non-degenerate, this constraint is not only sufficient but also necessary for most of the natural choices of the nonlinear function $e(\cdot)$ (e.g., componentwise saturation, sigmoid functions).

One of the main goals of the article is therefore to solve (at least approximately) the optimization problem

$$\underset{\eta, K}{\text{minimize}} \ J \qquad\qquad (\mathcal{P}_1)$$

subject to (2), (4) and (5).

3. Main results

The optimization problem $(\mathcal{P}_1)$ is intractable owing to the p-norm and the nonlinear function $\boldsymbol{e}(w)$, although sampling approximations are viable for small problems (Skaf & Boyd, 2009). We therefore propose to solve a relaxed problem where $u = \eta + K\boldsymbol{e}(w)$ in $(\mathcal{P}_1)$ is replaced with $u = \eta + Kw$ while keeping constraints on η and K such that the hard input constraints are satisfied when the original control policy is used. The relaxed problem must be convex since the objective is convex for each disturbance realization (Boyd & Vandenberghe, 2004). In the sequel, we show that the relaxed optimization problem is not only convex but also tractable.

3.1. Tractability of the proposed approach

First, note that, due to the linear dynamics and the affine parametrization of the control policy, the cost function in $(\mathcal{P}_1)$ is a sum of terms of the form $\mathbf{E}|X|^p$, where X is a Gaussian random with the two defining moments as a function of the decision variables

η and K (see Theorem 3 for a detailed derivation). Hence, we need an analytical expression for $\mathbf{E}|X|^p$ of a Gaussian random variable X.

Lemma 1. *If $X \sim \mathcal{N}(\mu, \sigma^2)$ then*

$$g(\mu, \sigma) := \mathbf{E}|X|^p = \frac{2^{p/2}}{\sqrt{\pi}} \sigma^p \Gamma\left(\frac{p+1}{2}\right) M\left(-\frac{p}{2}, \frac{1}{2}, -\frac{\mu^2}{2\sigma^2}\right) \quad (6)$$

and in particular for $p = 1$

$$g(\mu, \sigma) := \mathbf{E}|X| = \sqrt{\frac{2}{\pi}} \sigma e^{-\frac{\mu^2}{2\sigma^2}} + \mu \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right), \quad (7)$$

where $\Gamma(\cdot)$ is the Gamma function, $\operatorname{erf}(\cdot)$ the error function and $M(\cdot, \cdot, \cdot)$ the Kummer confluent hypergeometric function (Muller, 2001).

Proof. Follows by a straightforward integration from the definition of the expectation of a continuous random variable

$$\mathbf{E}|X|^p = \frac{1}{\sigma\sqrt{2\pi}} \left(\int_{-\infty}^{0} (-x)^p e^{\frac{-(x-\mu)^2}{2\sigma^2}} \, dx + \int_{0}^{\infty} x^p e^{\frac{-(x-\mu)^2}{2\sigma^2}} \, dx \right).$$

Evaluation of these integrals, possibly with the aid of a computer algebra system, leads directly to the two formulas. The more general expression (6) can also be found in Simon (2002, p. 11, Eq. (2.16)). $\square$

Now that we have an analytical expression for the cost function, the gradient and Hessian can be computed by a simple use of vector calculus.

Lemma 2. *If $X \sim \mathcal{N}(\mu, \sigma^2)$ for $\sigma > 0$, $X = \mu$ for $\sigma = 0$, and $\mu(\eta, k) = \mu_0 + b^T \eta$, $\sigma(\eta, k) = \|a + Ck\|_2$ then the function $f(\eta, k) = (\mathbf{E}|X|^p)(\eta, k)$ is jointly convex in (η, k), and its gradient and Hessian are given by*

$$\nabla f = f_\mu \nabla \mu + f_\sigma \nabla \sigma, \quad (8)$$

$$\begin{aligned} \operatorname{Hess}(f) &= \nabla \mu \left[f_{\mu\mu} \nabla \mu + f_{\mu\sigma} \nabla \sigma \right]^T \\ &\quad + \nabla \sigma \left[f_{\sigma\sigma} \nabla \sigma + f_{\sigma\mu} \nabla \mu \right]^T + f_\sigma \operatorname{Jac}(\nabla \sigma), \end{aligned} \quad (9)$$

where, for $\sigma > 0$ and with $q := C^T(a + Ck)/\sigma$,

$$\nabla \mu = \begin{bmatrix} b \\ 0 \end{bmatrix}, \qquad \nabla \sigma = \begin{bmatrix} 0 \\ q \end{bmatrix},$$

$$\operatorname{Jac}(\nabla \sigma) = \begin{bmatrix} 0 & 0 \\ 0 & (C^T C - qq^T)/\sigma \end{bmatrix}.$$

The expressions for the partial derivatives are

$$f_\mu = \frac{1}{\sqrt{\pi}} 2^{p/2} \mu p \sigma^{p-2} \gamma M_2,$$

$$f_\sigma = \frac{1}{\sqrt{\pi}} 2^{p/2} p \sigma^{p-3} \gamma \left(\sigma^2 M_1 - \mu^2 M_2 \right),$$

$$f_{\mu\mu} = \frac{1}{3\sqrt{\pi}} 2^{p/2} p \sigma^{p-4} \gamma \left[3\sigma^2 M_2 + \mu^2 (p-2) M_3 \right],$$

$$\begin{aligned} f_{\sigma\sigma} &= \frac{1}{\sqrt{\pi}} 2^{p/2} \sigma^{p-6} \gamma \left\{ [\mu^4 + \mu^2(3p-2)\sigma^2 \right. \\ &\quad + (p-1)p\sigma^4] M_1 \\ &\quad \left. - \mu^2(1+p)\left[\mu^2 + 2(p-1)\sigma^2\right] M_4 \right\}, \end{aligned}$$

$$f_{\mu\sigma} = \frac{1}{3\sqrt{\pi}} 2^{p/2} \mu (p-2) p \sigma^{p-5} \gamma \left[\mu^2 M_3 - 3\sigma^2 M_2 \right],$$

where

$$M_1 = M\left(-\frac{p}{2}, \frac{1}{2}, -\frac{\mu^2}{2\sigma^2}\right), M_2 = M\left(1 - \frac{p}{2}, \frac{3}{2}, -\frac{\mu^2}{2\sigma^2}\right),$$

$$M_3 = M\left(2 - \frac{p}{2}, \frac{5}{2}, -\frac{\mu^2}{2\sigma^2}\right), M_4 = M\left(-\frac{p}{2}, \frac{3}{2}, -\frac{\mu^2}{2\sigma^2}\right)$$

and

$$\gamma = \Gamma\left(\frac{p+1}{2}\right).$$

In particular for $p = 1$ we have a simplification

$$\nabla f = \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right) \nabla \mu + \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} \nabla \sigma,$$

$$\operatorname{Hess}(f) = \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} \left(\frac{1}{\sigma} \begin{bmatrix} b \\ -q\frac{\mu}{\sigma} \end{bmatrix} \begin{bmatrix} b \\ -q\frac{\mu}{\sigma} \end{bmatrix}^T + \operatorname{Jac}(\nabla \sigma) \right).$$

Proof. Convexity follows from convex calculus fundamentals since $f(\eta, k) = \mathbf{E}|\mu_0 + b^T \eta + (a + Ck)^T \tilde{w}|^p$ for some $\tilde{w} \sim \mathcal{N}(0, I)$, and the right-hand side is convex in (η, k) for every realization of $\tilde{w}$. The rest is a direct computation. $\square$

Note that, given p, the first two arguments of the hypergeometric functions are constant and the third argument is always negative, which allows for very fast computation of $M_1, \ldots, M_4$, for instance by using methods 1 and 2 of Muller (2001).

Theorem 3. *The optimization problem*

$$\operatorname*{minimize}_{\eta, K} J \qquad\qquad\qquad (\mathcal{P}_2)$$

subject to $u = \eta + Kw$ structured as in (2), (4) *and* (5)

with $w \sim \mathcal{N}(0, \Sigma_w)$ is convex and tractable in the variables (η, K). Furthermore the hard input constraints (3) *are satisfied under the control policy $u = \eta + K\mathbf{e}(w)$ provided that $\|e(\cdot)\|_\infty \leq \delta$.*

Proof. The objective function is a sum of terms of the form $\mathbf{E}|q_{jk}^T x_k|^p$ or $\mathbf{E}|r_{jk}^T u_k|^p$, where q_{jk} and r_{jk} denote the j-th rows of Q_k and R_k, respectively. Denote also $\mathcal{B}_k = [A^{k-1}B, \ldots, B, 0, \ldots, 0]$, $\mathcal{C}_k = [A^{k-1}, \ldots, I, 0, \ldots, 0]F$, where F comes from a decomposition of Σ_w as FF^T, and observe that[2]

$$\begin{aligned} q_{jk}^T x_k &= q_{jk}^T (A^k x_0 + \mathcal{B}_k u + \mathcal{C}_k \tilde{w}) \\ &= q_{jk}^T A^k x_0 + q_{jk}^T \mathcal{B}_k \eta + q_{jk}^T (\mathcal{C}_k + \mathcal{B}_k KF) \tilde{w} \end{aligned}$$

with $\tilde{w} \sim \mathcal{N}(0, I)$. It is clear that $q_{jk}^T x_k$ is Gaussian with the expectation

$$\mu(\eta, k) = \mathbf{E}(q_{jk}^T x_k) = q_{jk}^T A^k x_0 + q_{jk}^T \mathcal{B}_k \eta$$

and standard deviation

$$\sigma(\eta, k) = \|q_{jk}^T (\mathcal{C}_k + \mathcal{B}_k KF)\|_2 = \|\mathcal{C}_k^T q_{jk} + (F^T \otimes q_{jk}^T \mathcal{B}_k) Sk\|_2,$$

where $Sk = \operatorname{vec}(K)$ with S being a certain matrix of zeros and ones, and k containing only the nonzero elements of K.

Similarly

$$r_{jk}^T u_k = r_{jk}^T v_k \eta + r_{jk}^T v_k KF \tilde{w},$$

where v_k is a matrix that selects k-th block row of the size m. Consequently, the expectation and standard deviation become

$$\mu(\eta, k) = r_{jk}^T v_k \eta, \ \sigma(\eta, k) = \|(F^T \otimes r_{jk}^T v_k) Sk\|_2.$$

Lemma 1 can now be used to evaluate the cost function, whereas Lemma 2 establishes the convexity of the problem and provides expressions for the gradient and Hessian of the cost. Hence, the problem admits a tractable convex representation.

Finally, satisfaction of the input constraints is assured by the constraint (5). $\square$

---

[2] Here and hereafter the equality of random elements means the equality of their distributions, not necessarily of the elements itself.

Note that it also possible to handle problems stated with $\|\cdot\|_p$ instead of $\|\cdot\|_p^p$, that is, cost functions of the form

$$\hat{J} := \mathbf{E}\left\{\|Q_N x_N\|_p + \sum_{k=0}^{N-1} \|Q_k x_k\|_p + \|R_k u_k\|_p\right\}. \tag{10}$$

This can be achieved by minimizing upper bounds on the cost function based on Jensen's inequality (Boyd & Vandenberghe, 2004)

$$\mathbf{E}\|Q_k x_k\|_p \leq \left(\sum_{j=1}^{n_q} \mathbf{E}|q_{jk}^T x_k|^p\right)^{1/p}, \quad \mathbf{E}\|R_k u_k\|_p \leq \left(\sum_{j=1}^{n_r} \mathbf{E}|r_{jk}^T u_k|^p\right)^{1/p}. \tag{11}$$

These upper bounds can be easily shown to be convex in (η, K), and hence the whole problem can be handled within the framework of Theorem 3 with the additional composition with $1/p$ which poses no problem.

3.2. Bound on suboptimality

In this section, we provide a bound on the suboptimality in $(\mathcal{P}_1)$ of the solution to the relaxed problem $(\mathcal{P}_2)$. The bound is derived only for the $p = 1$ case where the analysis of suboptimality is straightforward and the bound turns out to be independent of the initial state x_0, which is a desirable property for a receding horizon application. By contrast, bounds for $p > 1$ are of little practical value because they are no longer independent of x_0 and in general not very tight.

The idea is to bound the difference of the costs under the policies $u = \eta + Kw$ and $u = \eta + Ke(w)$ for given η, K, which in effect bounds the difference of the respective optima. For ease of notation, the result is derived with time invariant weights, i.e., $Q_k := Q$, $R_k := R$ (and thus $q_{jk} := q_j$, $r_{jk} := r_j$) for all k, but generalizes immediately to the time varying case.

Lemma 4. *The difference between the costs J_e incurred under the policy $u = \eta + Ke(w)$ and the cost J_w incurred under the policy $u = \eta + Kw$ is bounded as*

$$|J_e - J_w| \leq \Omega \|K\|_\infty \left\{\sum_{k=0}^{N} \sum_{j=1}^{n_q} \|q_j^T \mathcal{B}_k\|_\infty + N \sum_{j=1}^{n_r} \|r_j^T\|_\infty\right\}$$

where $\Omega = \mathbf{E}\|e(w) - w\|_\infty$.

Proof. We have

$$|J_e - J_w| \leq \sum_{k=0}^{N} \sum_{j=1}^{n_q} |\mathbf{E}(|q_j^T x_k^e| - |q_j^T x_k^w|)|$$

$$+ \sum_{k=0}^{N-1} \sum_{j=1}^{n_r} |\mathbf{E}(|r_j^T u_k^e| - |r_j^T u_k^w|)|. \tag{12}$$

Next, by Jensen's inequality,

$$\begin{aligned}|\mathbf{E}(|q_j^T x_k^e| - |q_j^T x_k^w|)| &\leq \mathbf{E}\left||q_j^T x_k^e| - |q_j^T x_k^w|\right| \\ &\leq \mathbf{E}(|q_j^T x_k^e - q_j^T x_k^w|) \\ &= \mathbf{E}|q_j^T \mathcal{B}_k K(e(w) - w)|,\end{aligned} \tag{13}$$

where

$$x_k^e = A^k x_0 + \mathcal{B}_k \eta + \mathcal{B}_k K e(w) + \mathcal{C}_k w$$

and

$$x_k^w = A^k x_0 + \mathcal{B}_k \eta + \mathcal{B}_k K w + \mathcal{C}_k w.$$

Furthermore

$$\mathbf{E}|q_j^T \mathcal{B}_k K(e(w) - w)| \leq \|q_j^T \mathcal{B}_k K\|_\infty \Omega \leq \|q_j^T \mathcal{B}_k\|_\infty \|K\|_\infty \Omega,$$

where submultiplicativity of the infinity norm was used in both steps. Similar procedure can be carried out for the control inputs to yield

$$|\mathbf{E}(|r_j^T u_k^e| - |r_j^T u_k^w|)| \leq \|r_j^T\|_\infty \|K\|_\infty \Omega.$$

Summing up all terms in (12) now leads to the desired result, which completes the proof. □

Now it is rather straightforward to derive the suboptimality bound. Denote J_e^* the optimal value of $(\mathcal{P}_1)$ and the corresponding minimizer K_e^*, η_e^*. Denote also J_w^* the optimal value of $(\mathcal{P}_2)$ and the corresponding optimal solution K_w^*, η_w^*. Finally denote J_e the cost J under the control policy $u = \eta_w^* + K_w^* e(w)$ and J_w the cost J under the policy $u = \eta_e^* + K_e^* w$.

Theorem 5. *The solution η_w^*, K_w^* of $(\mathcal{P}_2)$ is not more than*

$$\beta := 2\Omega \frac{U_{\max}}{\delta} \left\{\sum_{k=0}^{N} \sum_{j=1}^{n_q} \|q_j^T \mathcal{B}_k\|_\infty + N \sum_{j=1}^{n_r} \|r_j^T\|_\infty\right\} \tag{14}$$

suboptimal in $(\mathcal{P}_1)$, that is, $J_e - J_e^* \leq \beta$.

Proof. First note that, because of the constraints on η and K in both optimization problems $(\mathcal{P}_1)$ and $(\mathcal{P}_2)$, the optimal feedback matrices K_e^* and K_w^* for both problems necessarily satisfy $\|K_e^*\|_\infty \leq U_{\max}/\delta$, $\|K_w^*\|_\infty \leq U_{\max}/\delta$.

Lemma 4 can now be invoked to obtain $|J_e - J_w^*| \leq \frac{\beta}{2}$, $|J_w - J_e^*| \leq \frac{\beta}{2}$, and, since $J_e^* \leq J_e$ and $J_w^* \leq J_w$, the bound immediately follows:

$$0 \leq J_e - J_e^* \leq J_e - J_w^* + J_w - J_e^* = |J_e - J_w^* + J_w - J_e^*| \leq \beta. \quad \square \tag{15}$$

The term $\Omega = \mathbf{E}\|e(w) - w\|_\infty$ in (14) can be computed to virtually arbitrary precision by means of a Monte Carlo simulation. The bound also provides an intuitively obvious guide to selecting the function $e(\cdot)$ in such a way that $e(w)$ and w do not differ very much with high probability. For instance, with the choice of $e(\cdot)$ as the elementwise saturation at three or four standard deviations of the disturbance, it is highly likely that the bound will be close to zero and, consequently, the solution to the relaxed problem will be almost optimal in the original one.

Note also that the bound can be further improved by noticing that the $|J_e - J_w^*|$ term in (15) can be computed exactly. Indeed J_w^* is known since it is the optimal value of $(\mathcal{P}_2)$, and J_e can be estimated by means of a Monte Carlo simulation with no further online optimization necessary. Hence, the yet improved version of the bound becomes

$$\tilde{\beta} := \frac{\beta}{2} + |J_e - J_w^*|. \tag{16}$$

4. Recursive feasibility

In this section, we introduce a technique to ensure that the closed-loop state trajectories satisfy the probabilistic (or chance) constraint

$$P(g^T x_k \leq h) \geq 1 - \alpha, \quad k = 1, 2, \ldots, \tag{17}$$

with $\alpha \in [0, 1]$. The closed-loop trajectories are generated by a receding horizon application of an *affine* disturbance feedback policy $u = \eta + Kw$ defined as in (4) with $e(w) = w$. The control authority is bounded by $U_{\max}$ as in (3). For a similar approach with a perturbed linear state-feedback, see Kouvaritakis et al. (2010).

First of all, to ensure satisfaction of the constraint (17), it is sufficient to guarantee that at each time $k \geq 1$, the one-step conditional constraint

$$P(g^T x_{k+1} \leq h \mid x_k) \geq 1 - \alpha \tag{18}$$

is satisfied under a given closed-loop control policy. This constraint turns out to be more amenable to a receding horizon control approach than the constraint (17).

However, in the presence of independent unbounded disturbances additively entering the system, it is impossible to ensure feasibility of (18) at all times. Hence, in this section, we assume that the disturbance sequence is i.i.d. and bounded as

$$\|w_k\|_\infty \leq \Delta, \quad k = 0, 1, \ldots, \tag{19}$$

for some $\Delta < \infty$. In addition, we assume that the pair (A, B) in (2) is stabilizable.

At this point, it should be stressed that, on account of the bounded support of the disturbance, the exact results of the previous sections no longer apply and have to be considered as approximations only. However, a similar suboptimality analysis as in Section 3.2 can be carried out, revealing that, as expected, the suboptimality of this approximation depends on the degree of dissimilarity between the true (bounded) distribution and a Gaussian distribution. A formal derivation goes along the lines of Section 3.2; details are omitted for brevity.

The one-step constraint (18) is handled using the standard dual mode paradigm (Mayne, Rawlings, Rao, & Scokaert, 2000) with the affine disturbance feedback (4) as the mode 1 policy, that is, at times $k = 0, \ldots, N - 1$, and any stabilizing state feedback in mode 2, that is, at times $k \geq N$. In mode 1, we have, given x_k,

$$P(g^T x_{k+1} \leq h \mid x_k) = P(g^T(Ax_k + Bu_k + w_k) \leq h)$$
$$= F_{g^T w_k}(h - g^T Ax_k - g^T Bu_k),$$

where $F_{g^T w_k}$ is the distribution function of $g^T w_k$. Thus, to ensure satisfaction of (18) we require that

$$g^T(Ax_k + Bu_k) \leq h - F_{g^T w_k}^{-1}(1 - \alpha),$$

where $F_{g^T w_k}^{-1}(\cdot)$ is the left quantile function of $g^T w_k$, for all possible states x_k reachable at time k by the disturbance sequence up to this time, $w_0^{k-1} := [w_0^T, \ldots, w_{k-1}^T]^T$, under a given policy in mode 1. For the affine disturbance feedback (4), we have

$$g^T(Ax_k + Bu_k) = g^T[A(A^k x_0 + \mathcal{B}_k(\eta + Kw) + \mathcal{C}_k w)$$
$$+ B(\eta_k + K_k w)]$$
$$= g^T(A^{k+1} x_0 + \mathcal{B}_{k+1}\eta) + g^T(\mathcal{B}_{k+1}K + A\mathcal{C}_k)w,$$

where η_k and K_k denote k-th block rows of size m of the respective matrices. Thus, considering the worst case value over all disturbances, $\max_w g^T(\mathcal{B}_{k+1}K + A\mathcal{C}_k)w = \|g^T(\mathcal{B}_{k+1}K + A\mathcal{C}_k)\|_\infty \Delta$, we get a sufficient condition for recursive feasibility in mode 1

$$g^T(A^{k+1} x_0 + \mathcal{B}_{k+1}\eta) \leq h - \|g^T(\mathcal{B}_{k+1}K + A\mathcal{C}_k)\|_\infty \Delta$$
$$- F_{g^T w_k}^{-1}(1 - \alpha). \tag{20}$$

Note that even though the disturbance sequence along the whole prediction horizon w appears in the above expressions, only the disturbances up to time $k-1$ contribute to the worst-case value due to the structure of the matrices involved.

In mode 2, we use a stabilizing state feedback $u_k = K_s x_k$ with the corresponding strictly stable feedback dynamics matrix $\hat{A} = A + BK_s$. One-step predictions in mode 2 now read

$$g^T x_{N+i+1} = g^T[\hat{A}^{i+1} \hat{x}_N + \hat{A}^{i+1}(\mathcal{B}_N K + \mathcal{C}_N) w_0^{N-1}$$
$$+ \hat{A}\hat{c}_i w_N^{N+i-1} + w_{N+i}],$$

where $\hat{x}_N = A^N x_0 + \mathcal{B}_N \eta$ and $\hat{c}_i = [\hat{A}^{i-1}, \hat{A}^{i-2}, \ldots, I]$. Thus, considering the worst case values over w_0^{N+i-1}, we get a sufficient condition

$$g^T \hat{A}^{i+1}(A^N x_0 + \mathcal{B}_N \eta) \leq h - \|g^T \hat{A}^{i+1}(\mathcal{B}_N K + \mathcal{C}_N)\|_\infty \Delta$$
$$- \|g^T \hat{A}\hat{c}_i\|_\infty \Delta - F_{g^T w_{N+i}}^{-1}(1 - \alpha), \quad \forall i \geq 0. \tag{21}$$

Thanks to the stability of $\hat{A}$, this infinite number of constraints can be approximated by a finite number of them. The term $\|g^T \hat{A}^{i+1}(\mathcal{B}_N K + \mathcal{C}_N)\|_\infty \Delta$ tends to zero and hence can be, for sufficiently large i, bounded by some μ, and $\|g^T \hat{A}\hat{c}_i\|_\infty \Delta$ tends from below to some $b < \infty$. Hence, the first $\hat{N}$ constraints in (21) can be kept in its original form and the rest (conservatively) approximated by

$$g^T \hat{A}^{i+1}(A^N x_0 + \mathcal{B}_N \eta) \leq h - \mu - b - F_{g^T w_{N+i}}^{-1}(1 - \alpha), \quad i > \hat{N} \tag{22}$$

with the additional constraint

$$\|g^T \hat{A}^{i+1}(\mathcal{B}_N K + \mathcal{C}_N)\|_\infty \Delta \leq \mu, \quad i > \hat{N}. \tag{23}$$

Hard input constraints can be enforced explicitly in mode 1 as

$$|\eta_j| + \Delta \|K_j\|_\infty \leq U_{\max}, \quad j = 1, \ldots, mN, \tag{24}$$

where η_j and K_j denote the j-th rows of η and K, respectively. In mode 2, the input constraints are enforced by adding another chance constraints of the form (17) with $\alpha = 0$, which are handled according to the previous discussion.

There is still an infinite number of constraints (22) and likewise of mode 2 input constraints, but it is known that for a stable matrix $\hat{A}$, the feasible region of this set of constraints is given by the first $\tilde{N} < \infty$ of them since, because of the input constraints, the admissible region is typically bounded, with the origin in its interior. The minimum such $\tilde{N}$ can be obtained offline by solving a sequence of linear programs (Gilbert & Tan, 1991). Finally, the constraint (23) is clearly finitely determined since $\hat{A}$ is stable and $\|K\|_\infty$ is bounded by $U_{\max}/\Delta$.

Given a feasible solution (η, K) at time zero, we are guaranteed to have a feasible point $(\tilde{\eta}, \tilde{K})$ at time one (and hence, by induction, at all times) constructed as $\tilde{\eta} = \begin{bmatrix} \hat{\eta} + K_{(1)} w_0 \\ \eta_L \end{bmatrix}$, $\tilde{K} = \begin{bmatrix} \hat{k} & 0 \\ K_L & 0 \end{bmatrix}$, where $\hat{\eta}$ denotes the matrix η without the first block row, $\hat{K}$ denotes the matrix K without the first block row and block column and $K_{(1)}$ is the first block column of K without the first block row. The last block rows η_L and K_L are given by

$$\eta_L = K_s(A^N x_0 + \mathcal{B}_N \eta) + K_s[(\mathcal{B}_N K + \mathcal{C}_N)]_{1:n} \cdot w_0$$

and

$$K_L = K_s[\mathcal{B}_N K + \mathcal{C}_N]_{n+1:nN},$$

where $[\cdot]_{p:q}$ denotes the matrix given by the p-th through q-th columns of a matrix $[\cdot]$.

To conclude this section we note that the recursively feasible representation of the chance constraint (18) leads to *affine* constraints on η and K regardless of the disturbance distribution, which is in stark contrast to the standard 'open-loop' chance constraints that lead to second-order-cone constraints for Gaussian disturbances and usually have no exact representation otherwise. This simplicity comes at the cost of a certain degree of conservatism introduced by enforcing the one-step constraint (18) rather than directly (17).

Table 1
Comparison of cost functions at the end of the optimization horizon.

p	SH-NDF	NDF	SH-MPC	CE-MPC	Opt-OL	CE-OL
1	85.9	90.3	98.3	120.0	140.5	144.2
1.5	73.3	76.8	85.0	102.7	115.2	118.0
2	70.5	74.1	80.1	96.0	110.9	113.7

5. Numerical examples

We present two numerical examples that compare our method to other control strategies. With the gradient and Hessian on hand, the problem ($\mathcal{P}_2$) can be solved by a general-purpose convex solver. For our small-scale examples we managed with a custom primal–dual interior-point solver, although the Matlab FMINCON function would do as well.

In the first example, we consider a fixed horizon stochastic control problem. For the system matrices and the noise covariance matrix we chose $A = \begin{bmatrix} 1 & -0.4 \\ 0.1 & 1 \end{bmatrix}, B = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}, \Sigma_w = I \otimes \begin{bmatrix} 8 & 5 \\ 5 & 6 \end{bmatrix}$ with w_k zero-mean jointly Gaussian. We set $Q = I$, $R = 0.1$, and the input constraints to $U_{\max} = 30$. The optimization horizon is $T = 12$, the initial state $x_0 = [1, -1]^T$. The function $e(\cdot)$ was chosen as the elementwise saturation that saturates the disturbances at $4\sqrt{\max(\mathrm{diag}(\Sigma_w))} = 11.31$. We compared our control policy (NDF) (with the prediction horizon $N = T$ and the control horizon $N_c = T$) against the standard certainty equivalent MPC (CE-MPC) ($N_c = 1, N = T$) and against the shrinking horizon certainty equivalent MPC (SH-MPC) ($N_c = 1, N(k) = T - k, k = 0, \ldots, T-1$). Furthermore, we set against each other the proposed method with $K = 0$ (Opt-OL) and the certainty equivalent open loop control (CE-OL) (i.e. CE-MPC with $N_c = N = T$). For the sake of completeness we also included our method in the shrinking horizon mode (SH-NDF) with $N_c = 2, N(k) = T - k$. The cost performance of the policies was evaluated using 1000 Monte Carlo runs, which renders the Monte Carlo error negligible compared to the difference in the costs of the policies.

The results for the p-norm minimization (10) using the upper-bounds (11) are summarized in Table 1, which shows that our method (without shrinking) outperforms the others by a significant margin, except perhaps for SH-MPC where the difference is smaller and, naturally, our method in the shrinking horizon mode. On the other hand, unlike with MPC strategies, there is no need for online optimization with our method in this setting. It is also worth noting that our method with $K = 0$ (i.e., an open loop policy) slightly outperforms the certainty equivalent open loop control, which is in contrast with the quadratic cost case where this strategy is optimal in the class of open loop policies. Finally, we evaluate the bound (16), which yields $\tilde{\beta} = 3.3 \cdot 10^{-2}$, showing that the solution found by ($\mathcal{P}_2$) is in this case practically optimal in ($\mathcal{P}_1$) for $p = 1$. For $p > 1$, no conclusions about the suboptimality of the solution can be drawn although the superior performance compared to SH-MPC suggests that it be should very small.

Our second example compares the recursively feasible affine disturbance feedback stochastic MPC (S-MPC) with the robust affine disturbance feedback MPC. We consider the system given by the matrices $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, where w_k is i.i.d. with the standard normal distribution truncated at $\Delta = 3$. The weighting matrices were set to $Q = \mathrm{diag}(0, 1)$ and $R = 0$, the input constraints to $U_{\max} = 12$, and the initial state to $x_0 = [5, 5]^T$. We chose the quadratic cost, which involves solving the problem ($\mathcal{P}_2$) with $p = 2$ and $\delta = \Delta = 3$ augmented with the recursively feasible chance constraints (20)–(23). Expression (6) was used to (approximately) evaluate the cost function even though, being quadratic, it could have been evaluated exactly. We

Table 2
Comparison of the control policies in a receding horizon mode over 10,000 time steps. The final cost for the LQ controller is $9.9575 \cdot 10^3$.

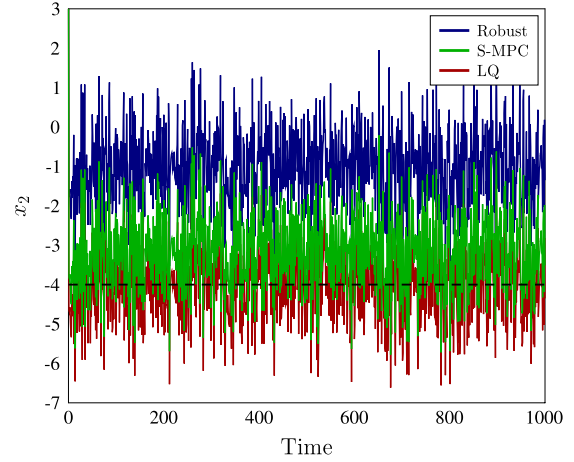| Policy	LQ	S-MPC	Robust
J/J_{LQ}	1	1.72	10.18
# violations	4920	1991	0

Fig. 1. Time evolution of x_2 over the first 1000 time steps under the stochastic and robust affine disturbance feedback policies in a receding horizon mode.

consider control to a constant set point $x_2 = -4$ and a single chance constraint $P(x_2 \geq -4) \geq 0.8$, i.e., $g = [0, -1]^T, h = 4, \alpha = 0.2$.

We compared our control policy with $N = 8, \hat{N} = 10, \tilde{N} = 1$ against the affine disturbance feedback robust MPC and the LQ optimal controller, whose gain $K_s = -[0.5, 0.5]$ was also used as the mode 2 controller for our algorithm. The parameters in the mode 2 constraint (22) are $\mu = 0.1, b = 0$ for the chance constraint, and $\mu = 0.1, b = 6$ for the two hard input constraints.

The controllers were examined over one very long time interval $T = 10000$. Table 2 shows that, in this case, our algorithm can fully exploit the probabilistic constraint to gain a significant performance improvement over the robust MPC. The LQ optimal controller, of course, outperforms the other two policies in terms of the cost but violates the probabilistic constraint substantially. These conclusions are also confirmed by Fig. 1, which shows the evolution of x_2 over the first 1000 time steps.

6. Conclusion

In this article we primarily dealt with the expectation of the p-norm stochastic control problem for which we developed an approximate solution technique ensuring bounded control inputs in the presence of Gaussian disturbances. Moreover, we constructed a suboptimality bound of our method in a certain class of nonlinear feedback control policies. The presented approach can be straightforwardly extended to the output feedback case.

Furthermore, we developed a recursively feasible algorithm with respect to chance constraints and/or hard input constraints in the presence of bounded disturbances. The algorithm has the nice property of introducing only additional *affine* constraints. The approach, however, exhibits a certain degree of conservatism, reduction of which should be a focus of future work.

References

Bertsekas, D. P. (2007). *Dynamic programming and optimal control* (3rd ed.). Athena Scientific.
Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

Cannon, M., Kouvaritakis, B., & Ng, D. (2009). Probabilistic tubes in linear stochastic model predictive control. *Systems & Control Letters*, *58*(10–11), 747–753.

Cannon, M., Kouvaritakis, B., & Wu, X. (2009). Model predictive control for systems with stochastic multiplicative uncertainty and probabilistic constraints. *Automatica*, *45*(1), 167–172.

Farahani, S.S., van den Boom, T., van der Weide, H., & De Schutter, B. (2010). An approximation approach for model predictive control of stochastic max-plus linear systems. In *Proceedings of the workshop on discrete event systems* (pp. 386–391).

Gilbert, E. G., & Tan, K. T. (1991). Linear systems with state and control constraints: the theory and application of maximal output admissible sets. *IEEE Transactions on Automatic Control*, *36*(9), 1008–1020.

Hokayem, P., Chatterjee, D., & Lygeros, J. (2009). On stochastic model predictive control with bounded control inputs. In *48th IEEE conference on decision and control* (pp. 6359–6364).

Korda, M., & Cigler, J. (2011). On 1-norm stochastic optimal control with bounded control inputs. In *American control conference*, July 2011 (pp. 60–65).

Kouvaritakis, B., Cannon, M., Raković, S. V., & Cheng, Q. (2010). Explicit use of probabilistic distributions in linear predictive control. *Automatica*, *46*(10), 1719–1724.

Mayne, D. Q., Rawlings, J. B., Rao, C. V., & Scokaert, P. O. (2000). Constrained model predictive control: stability and optimality. *Automatica*, *36*(6), 789–814.

Muller, K. E. (2001). Computing the confluent hypergeometric function, $M(a, b, x)$. *Numerical Mathematics*, *90*(1), 179–196.

Oldewurtel, F., Parisio, A., Jones, C.N., Morari, M., Gyalistras, D., Gwerder, M., Stauch, V., Lehmann, B., & Wirth, K. (2010). Energy efficient building climate control using stochastic model predictive control and weather predictions. In *American control conference* (pp. 5100–5105).

Simon, M. K. (2002). *Probability distributions involving Gaussian random variables: a handbook for engineers and scientists, vol. 683*. Netherlands: Springer.

Skaf, Jolle, & Boyd, Stephen (2010). Design of affine controllers via convex optimization. *IEEE Transactions on Automatic Control*, *55*(11), 2476–2487.

Skaf, J., & Boyd, S. (2009). Nonlinear Q-design for convex stochastic control. *IEEE Transactions on Automatic Control*, *54*(10), 2426–2430.

Milan Korda received a Bachelor's degree in Electrical Engineering and a Master's degree in Cybernetics and Robotics from Czech Technical University in Prague. His master thesis research was conducted in part at ETH Zurich. Currently he is pursuing a Ph.D. degree at the Automatic Control Laboratory, EPFL Lausanne, Switzerland. His current research interests are model predictive control and stochastic optimal control with applications in financial engineering and building climate control.

Jiří Cigler received the M.Sc. degree in Electrical Engineering in 2009. He is currently a Ph.D. student at Department of Control Engineering of Czech Technical University in Prague. His research is mainly aimed at model predictive control for buildings. He is also a research engineer at Feramat Cybernetics Ltd. where he is focusing on the optimization of various industrial processes.