

# On 1-Norm Stochastic Optimal Control with Bounded Control Inputs

Milan Korda and Jiří Cigler

**Abstract**—This paper deals with the finite horizon stochastic optimal control problem with the expectation of the 1-norm as the objective function and jointly Gaussian, although not necessarily independent, disturbances. 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 the class of aforementioned nonlinear feedback policies is given as well as a simple proof of mean-square stability of a receding horizon implementation provided that the system matrix is Schur stable.

## I. 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 [2].

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, approaches is the certainty equivalent model predictive control (CE-MPC) [1, 2] 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 or Youla-Kučera parametrization [13, 14].

However convenient the paradigm of 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 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 [7], where this approach was developed for the quadratic cost. In this article we follow up on this work and develop a methodology for solving this problem in the 1-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.

Another branch of approximation techniques bounds the disturbances a priori and solves a robust MPC problem, while guaranteeing an open loop probabilistic bound on the performance [3]. This approach, however, tends to be very conservative, and thus the idea of bounding the disturbances a priori based on their distribution appears more often in the context of chance constraints; see, e.g., [9, 10].

The very important, though much neglected, question of stability and recursive feasibility of stochastic receding horizon schemes is addressed in a series of papers [5, 6, 8]. These papers, however, assume either compactly supported disturbances or only probabilistic input and state constraints, whereas [7] and [12] deal exclusively with stability in the presence of hard input constraints. In this paper we prove in a much simpler way a slight generalization of one of their stability results.

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 1-norm or similar cost functions are ample [10, 11].

The current paper is organized as follows. In Section II we state the problem to be solved, Section III presents the main results on convexity of the stated problem, derives a bound on the suboptimality of the approach, and proves the stability of the policy in a receding horizon mode under certain conditions. Section IV presents two numerical examples to illustrate our results. Finally, we conclude and give an outlook in Section V.

## A. Notation

Throughout the article  $\mathbb{R}$  denotes the set of reals,  $N$  and  $N_c$  denote the prediction and control horizons, respectively. The positive integers  $m$  and  $n$  denote the number of control inputs and the state-space dimension. The function  $\text{sat}_r(\cdot)$  denotes the standard elementwise saturation of the components of a vector to  $r$ , and  $\|\cdot\|_\infty$  denotes the induced infinity norm of a matrix (in particular *not* the maximum absolute value if the matrix is a row vector). The symbols  $\rho(\cdot)$  and  $\text{tr}(\cdot)$  denote the spectral radius and the trace of a square matrix. The expectation operator is denoted by  $\mathbf{E}$ , and  $X \sim \mathcal{N}(\mu, \Sigma)$  indicates that  $X$  is a Gaussian random variable with the expectation  $\mu$  and the covariance matrix  $\Sigma$ . The symbols  $\text{vec}(\cdot)$  and  $\otimes$  denote the vectorization and the Kronecker

Milan Korda (email: korda.m@gmail.com) and Jiří Cigler (email: jiri.cigler@fel.cvut.cz) are with Department of Control Engineering, Faculty of Electrical Engineering of Czech Technical University in Prague, Karlovo náměstí 13, Czech Republic

product respectively. Finally,  $\text{Hess}(\cdot)$  and  $\text{Jac}(\cdot)$  denote the Hessian and the Jacobian of a function.

## II. PROBLEM STATEMENT

We consider the problem of minimizing the cost function

$$J := \mathbf{E} \left\{ \|Q_N x_N\|_1 + \sum_{k=0}^{N-1} \|Q_k x_k\|_1 + \|R_k u_k\|_1 \right\} \quad (1)$$

subject to the discrete-time system dynamics

$$x_{k+1} = Ax_k + Bu_k + w_k, \quad (2)$$

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

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

where  $Q_k \in \mathbb{R}^{n_q \times n}$ ,  $R_k \in \mathbb{R}^{n_r \times m}$  are weighting matrices. 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  $\Sigma$ .

The minimization to be carried out is over all Borel measurable causal disturbance feedback policies

$$u_k = \phi_k(x_0, w_0, \dots, w_{k-1}), \quad k = 0, \dots, N-1. \quad (4)$$

This problem is, however, in general intractable and various approximation techniques exist; see, e.g., [2]. In this paper, we adopt the approach of [7] where the authors propose to search over a class of causal policies affine in certain nonlinear functions of the disturbances, i.e.,

$$u = \eta + Ke(w) = \begin{bmatrix} \eta_0 \\ \vdots \\ \eta_{N-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & \dots & 0 \\ K_{1,1} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \\ K_{N-1,1} & \dots & K_{N-1,N-1} & 0 \end{bmatrix} e(w), \quad (5)$$

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

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

$$\begin{aligned} & \underset{\eta, K}{\text{minimize}} && \mathbf{E} \left\{ \|Q_N x_N\|_1 + \sum_{k=0}^{N-1} \|Q_k x_k\|_1 + \|R_k u_k\|_1 \right\} \\ & \text{subject to} && u = \eta + Ke(w) \\ & && x_{k+1} = Ax_k + Bu_k + w_k \\ & && K \text{ is strictly block lower triangular} \\ & && \text{constraints on } \eta, K \text{ such that (3) is satisfied.} \end{aligned} \quad (6)$$

## III. MAIN RESULTS

Although sampling techniques are viable for small problems (see [13]), the optimization problem (6) is in its full generality intractable owing to the 1-norm and the nonlinear function  $e(w)$ . We therefore propose to solve a relaxed problem where  $u = \eta + Ke(w)$  in (6) is replaced with  $u = \eta + Kw$  while keeping constraints on  $\eta, K$  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 [4]. In the sequel, we show that the relaxed optimization problem is not only convex but also tractable. To this end, we need an expression for the expectation of the absolute value of a Gaussian random variable.

*A. Tractability of the proposed approach*

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

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

where  $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$  is the error function.

*Proof.* Follows by a straightforward integration from the definition of the expectation of a continuously distributed random variable

$$\mathbf{E}|X| = \frac{1}{\sigma\sqrt{2\pi}} \left( \int_{-\infty}^0 -xe^{-\frac{(x-\mu)^2}{2\sigma^2}} dx + \int_0^\infty xe^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \right), \quad (8)$$

and by using the definition of the error function.  $\blacksquare$

Next, we show that the continuous extension (to cater for the  $\sigma = 0$  case) of the expectation of the Gaussian random variable modulus is convex under a certain composition, and we also provide an expression for its gradient and Hessian.

**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  $f(\eta, k) = (\mathbf{E}|X|)(\eta, k)$  is jointly convex in  $(\eta, k)$ .*

*Proof.* The proof proceeds directly by computing the Hessian of  $f$  for  $\sigma > 0$ , and then a continuity argument is used to complete the proof. For  $\sigma > 0$ ,  $f(\eta, k)$  coincides with  $g(\mu(\eta, k), \sigma(\eta, k))$  and the gradient is

$$\nabla f(\mu, \sigma) = \frac{\partial f}{\partial \mu} \nabla \mu + \frac{\partial f}{\partial \sigma} \nabla \sigma = \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right) \nabla \mu + \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} \nabla \sigma \quad (9)$$

with

$$\nabla \mu = \begin{bmatrix} b \\ 0 \end{bmatrix}, \quad \nabla \sigma = \begin{bmatrix} 0 \\ C^T \frac{a+Ck}{\sigma} \end{bmatrix}. \quad (10)$$

The expression for  $\nabla \sigma$  follows from the fact that  $\nabla \|x\|_2 = \frac{x}{\|x\|_2}$  and the multivariate form of the chain rule. Now since  $\text{Hess}(f) = \text{Jac}(\nabla f)$  and  $\text{Jac}(h\tilde{g}) = \tilde{g}(\nabla h)^T + h\text{Jac}(\tilde{g})$  for real-valued function  $h$  and multivariate  $\tilde{g}$ , it follows that

$$\begin{aligned} \text{Hess}(f) &= \begin{bmatrix} b \\ 0 \end{bmatrix} \left\{ \nabla \operatorname{erf}\left(\frac{\mu}{\sigma\sqrt{2}}\right) \right\}^T \\ &+ \begin{bmatrix} 0 \\ C^T \frac{a+Ck}{\sigma} \end{bmatrix} \left\{ \nabla \left( \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} \right) \right\}^T + \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} \text{Jac}(\nabla \sigma) \end{aligned} \quad (11)$$

with

$$\text{Jac}(\nabla\sigma) = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{\|x\|_2} C^T \left( I - \frac{xx^T}{\|x\|_2^2} \right) C \end{bmatrix} \geq 0, \quad (12)$$

where  $x = a + Ck$  since, again by the chain rule,

$$\text{Jac}_k \nabla\sigma = C^T \text{Jac} \frac{a + Ck}{\|a + Ck\|_2} = C^T \left[ \text{Jac} \left( \frac{y}{\|y\|_2} \right) \circ (a + Ck) \right] C, \quad (13)$$

where  $\circ$  denotes the standard function composition. The remaining two terms in (11) are

$$\begin{aligned} \nabla \text{erf} \left( \frac{\mu}{\sigma\sqrt{2}} \right) &= \begin{bmatrix} b \\ 0 \end{bmatrix} \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} - \begin{bmatrix} 0 \\ C^T \frac{a+Ck}{\sigma} \end{bmatrix} \sqrt{\frac{2}{\pi}} \frac{\mu}{\sigma^2} e^{-\frac{\mu^2}{2\sigma^2}}, \\ \nabla \left( \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} \right) &= - \begin{bmatrix} b \\ 0 \end{bmatrix} \sqrt{\frac{2}{\pi}} \frac{\mu}{\sigma^2} e^{-\frac{\mu^2}{2\sigma^2}} + \begin{bmatrix} 0 \\ C^T \frac{a+Ck}{\sigma} \end{bmatrix} \sqrt{\frac{2}{\pi}} \frac{\mu^2}{\sigma^3} e^{-\frac{\mu^2}{2\sigma^2}}. \end{aligned} \quad (14)$$

Rewriting the Hessian with

$$q := \begin{bmatrix} 0 \\ C^T \frac{a+Ck}{\sigma} \end{bmatrix} \quad (16)$$

then yields

$$\text{Hess}(f) = \sqrt{\frac{2}{\pi}} e^{-\frac{\mu^2}{2\sigma^2}} \left( \frac{1}{\sigma} \begin{bmatrix} b \\ -q \end{bmatrix} \begin{bmatrix} b \\ -q \end{bmatrix}^T + \text{Jac}(\nabla\sigma) \right) \geq 0. \quad (17)$$

It is easily seen that  $f(\eta, k)$  is continuous and that the sequence of smoothed functions  $f_n(\eta, k) = g\left(\mu(\eta, k), \sqrt{\frac{1}{n} + \sum_i x_i^2}\right)$  converges pointwise to  $f$ . The functions  $f_n$  are readily shown to be convex by computing their respective Hessians in the same fashion as above. The function  $f(\eta, k)$  is therefore convex since it is a limit of convex functions. ■

Note that a simpler argument for convexity can be given because  $f(\eta, k) = \mathbf{E}[\mu_0 + b^T \eta + (a + Ck)\tilde{w}]$  for some  $\tilde{w} \sim \mathcal{N}(0, I)$ , which must be a convex function since the argument of the expectation is convex for every realization of  $\tilde{w}$  [4].

Now we are ready to formulate a tractable approximation of (6).

**Theorem 1.** *The optimization problem*

$$\begin{aligned} \underset{\eta, K}{\text{minimize}} \quad & \mathbf{E} \left\{ \|Q_N x_N\|_1 + \sum_{k=0}^{N-1} \|Q_k x_k\|_1 + \|R_k u_k\|_1 \right\} \\ \text{subject to} \quad & u = \eta + Kw \\ & x_{k+1} = Ax_k + Bu_k + w_k \\ & K \text{ is strictly block lower triangular} \\ & |\eta_i| + \varepsilon \|K_i\|_\infty \leq U_{\max}, \quad i = 1, \dots, mN \end{aligned} \quad (18)$$

with  $w \sim \mathcal{N}(0, \Sigma)$  is convex and tractable in the variables  $(\eta, K)$ . Furthermore the hard input constraints (3) are satisfied under the control policy  $u = \eta + Ke(w)$  if  $\|e(w)\|_\infty \leq \varepsilon$ . Here  $K_i$  denotes the  $i$ -th row of  $K$ .

*Proof.* The objective function is a sum of terms of the form  $\mathbf{E}[|q_{jk}^T x_k|]$  or  $\mathbf{E}[|r_{jk}^T u_k|]$ , where  $q_{jk}, r_{jk}$  denote the  $j$ -th rows of  $Q_k, R_k$  respectively. Denote also

$$\mathcal{B}_k = [A^{k-1}B, \dots, B, 0, \dots, 0], \quad C_k = [A^{k-1}, \dots, I, 0, \dots, 0]F,$$

where  $\Sigma = FF^T$ , and observe that

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

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

$$\mu(\eta, k) = \mathbf{E}(q_{jk}^T x_k) = q_{jk}^T A^k x_0 + q_{jk}^T \mathcal{B}_k \eta, \quad (19)$$

and standard deviation

$$\sigma(\eta, k) = \|q_{jk}^T (C_k + \mathcal{B}_k KF)\|_2 = \|C_k^T q_{jk} + (F^T \otimes q_{jk}^T \mathcal{B}_k) S k\|_2, \quad (20)$$

where  $S k = \text{vec}(K)$  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 vector that selects  $k$ -th block row of the size  $m$ . Consequently, the expectation and standard deviation become

$$\mu(\eta, k) = r_{jk}^T v_k \eta, \quad \sigma(\eta, k) = \|(F^T \otimes r_{jk}^T v_k) S k\|_2. \quad (21)$$

Application of Lemma 2, in the proof of which the gradient and Hessian were computed, now completes the convexity and tractability part of the proof. Satisfaction of the input constraints follows immediately from the definition of the induced infinity norm and from the assumption that  $\|e(w)\|_\infty \leq \varepsilon$ . ■

**B. Bound on suboptimality**

In this section we provide a bound on the suboptimality in (6) (with the same constraints on  $\eta, K$  as in (18)) of the solution to the relaxed problem (18). The idea is to bound the difference of the costs under the policies  $u = \eta + Kw$  and  $u = \eta + Ke(w)$  for given  $\eta, K$ , 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 3.** *The cost  $J_e$  incurred under the policy  $u = \eta + Ke(w)$  and the cost  $J_w$  incurred under the policy  $u = \eta + Kw$  differ not more than*

$$(n_q(N+1)\|Q\|_\infty\|\mathcal{B}_N\|_\infty + n_r N\|R\|_\infty) \mathbf{E}\|e(w) - w\|_\infty \|K\|_\infty \quad (22)$$

*Proof.* We have

$$\begin{aligned} |J_e - J_w| &\leq \sum_{k=0}^N \sum_{j=1}^{n_q} |\mathbf{E}(|q_j^T x_k^e| - |q_j^T x_k^w|)| \\ &\quad + \sum_{k=0}^{N-1} \sum_{j=1}^{n_r} |\mathbf{E}(|r_j^T u_k^e| - |r_j^T u_k^w|)|. \end{aligned} \quad (23)$$

Next, by Jensen's inequality,

$$\begin{aligned} |\mathbf{E}(|q_j^T x_k^e| - |q_j^T x_k^w|)| &\leq \mathbf{E} \left| |q_j^T x_k^e| - |q_j^T x_k^w| \right| \leq \mathbf{E} |q_j^T x_k^e - q_j^T x_k^w| \\ &= \mathbf{E} |q_j^T \mathcal{B}_k K (e(w) - w)|, \end{aligned} \quad (24)$$

where

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

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

Furthermore

$$\begin{aligned} \mathbf{E}|q_j^T \mathcal{B}_k K(e(w) - w)| &\leq \|q_j^T \mathcal{B}_k K\|_\infty \mathbf{E}\|e(w) - w\|_\infty \\ &\leq \|q_j^T \mathcal{B}_k\|_\infty \|K\|_\infty \mathbf{E}\|e(w) - w\|_\infty \\ &\leq \|Q\|_\infty \|\mathcal{B}_N\|_\infty \|K\|_\infty \mathbf{E}\|e(w) - w\|_\infty. \end{aligned} \quad (25)$$

Similar procedure can be carried out for control inputs to yield

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

Summing up all terms in (23) now leads to the desired result

$$|J_e - J_w| \leq (n_q(N+1))\|Q\|_\infty \|\mathcal{B}_N\|_\infty + n_r N \|R\|_\infty \mathbf{E}\|e(w) - w\|_\infty \|K\|_\infty,$$

which completes the proof.  $\blacksquare$

Now it is rather straightforward to derive the suboptimality bound. Denote  $J_e^*$  the optimal value of (6) and the corresponding minimizer  $K_e^*$ ,  $\eta_e^*$ . Denote also  $J_w^*$  the optimal value of (18) and the corresponding optimal solution  $K_w^*$ ,  $\eta_w^*$ . 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 2.** *The solution  $\eta_w^*$ ,  $K_w^*$  of (18) is not more than*

$$\beta := 2(n_q(N+1))\|Q\|_\infty \|\mathcal{B}_N\|_\infty + n_r N \|R\|_\infty \mathbf{E}\|e(w) - w\|_\infty \frac{U_{\max}}{\varepsilon} \quad (26)$$

*suboptimal in (6).*

*Proof.* It follows from Lemma 3 that

$$|J_e - J_w^*| \leq \frac{\beta}{2}, \quad |J_w - J_e^*| \leq \frac{\beta}{2}$$

since  $\|K_e^*\|_\infty \leq U_{\max}/\varepsilon$ ,  $\|K_w^*\|_\infty \leq U_{\max}/\varepsilon$  because of the constraint on  $K$  and  $\eta$  in both optimization problems

$$|\eta_i| + \varepsilon \|K_i\|_\infty \leq U_{\max}, \quad i = 1, \dots, mN,$$

which implies  $\|K\|_\infty \leq U_{\max}/\varepsilon$ .

Now since  $J_e^* \leq J_e$  and  $J_w^* \leq J_w$  the bound immediately follows

$$0 \leq J_e - J_e^* \leq J_e - J_w^* + J_w - J_e^* = |J_e - J_w^*| + |J_w - J_e^*| \leq \beta,$$

which completes the proof.  $\blacksquare$

The term  $\mathbf{E}\|e(w) - w\|_\infty$  in (26) 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  $e(w)$  in such a way that  $e(w)$  and  $w$  do not differ very much with high probability. For instance with the choice of  $e(w)$  as the elementwise saturation  $e_i(w_i) = \text{sat}_r(w_i)$  with  $r \geq 4\sqrt{\rho(\Sigma)}$  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 this fairly crude bound can be significantly improved by terminating one inequality earlier in (25) at the cost of a slightly more complicated expression.

### C. Receding horizon stability

In this section we provide a slight generalization and a much simplified proof of a result that already appeared in [7].

**Theorem 3.** *Let  $u_k$ ,  $w_k$  be two stochastic processes defined on the same probabilistic space with  $\|u_k\|_\infty \leq U_{\max}$  a.s. and  $\sup_{k \geq 0} \mathbf{E}\|w_k\|_2^2 < \infty$ . The state of the system  $x_{k+1} = Ax_k + Bu_k + w_k$  then stays mean-square bounded (i.e.,  $\sup_{k \geq 0} \mathbf{E}\|x_k\|_2^2 < \infty$ ) provided that  $\mathbf{E}\|x_0\|_2^2 < \infty$  and  $\rho(A) < 1$ .*

*Proof.*  $\mathbf{E}\|x_k\|_2^2 = \text{tr}(\mathbf{E}\{x_k x_k^T\})$  and consequently it suffices to show that  $\mathbf{E}\{x_k x_k^T\}$  is bounded in any norm because of the norm equivalence on finite dimensional vector spaces and the fact that  $\text{tr}(\cdot)$  coincides with the nuclear norm on the space of positive semidefinite matrices. The proof proceeds by direct evaluation:

$$\begin{aligned} \mathbf{E}(x_k x_k^T) &= \mathbf{E}\{(A^k x_0 + \mathcal{B}_k U_k + C_k W_k)(A^k x_0 + \mathcal{B}_k U_k + C_k W_k)^T\} \\ &= A^k P_0 (A^k)^T + A^k \mathbf{E}\{x_0 U_k^T\} \mathcal{B}_k^T + \mathcal{B}_k \mathbf{E}\{U_k x_0^T\} (A^k)^T \\ &\quad + \mathcal{B}_k \mathbf{E}\{U_k U_k^T\} \mathcal{B}_k^T + \mathcal{B}_k \mathbf{E}\{U_k W_k^T\} C_k^T + C_k \mathbf{E}\{W_k U_k^T\} \mathcal{B}_k^T \\ &\quad + C_k \mathbf{E}\{W_k W_k^T\} C_k^T, \end{aligned} \quad (27)$$

where

$$\begin{aligned} U_k &= [u_0^T, \dots, u_{k-1}^T]^T, \quad W_k = [w_0^T, \dots, w_{k-1}^T]^T, \\ \mathcal{B}_k &= [A^{k-1} B, \dots, B], \quad C_k = [A^{k-1}, \dots, I]. \end{aligned}$$

The boundedness of the first term is obvious, the boundedness of the second and third terms follows from the fact that  $\|\mathbf{E}\{x_0 U_k^T\}\|_2 \leq U_{\max} \sqrt{m k \mathbf{E}\|x_0\|_2^2}$  (this follows directly by Jensen's and Cauchy-Schwarz inequalities). The boundedness of  $\mathcal{B}_k$  is obvious by the assumption that  $\rho(A) < 1$ , and therefore the second and third terms actually go to zero.

Consider now any family of matrices  $M_{rq}$  such that  $\|M_{rq}\| \leq \Delta$  for all  $r, q$ . For such a family and any submultiplicative norm  $\|\cdot\|$  we have

$$\begin{aligned} \left\| \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} A^i M_{rq} A^j \right\| &\leq \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \|A^i\| \|M_{rq}\| \|A^j\| \\ &\leq \Delta \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \|A^i\| \|A^j\|. \end{aligned} \quad (28)$$

The first term in (28) is therefore bounded since the last series is convergent by the assumption that  $\rho(A) < 1$ .

The theorem then follows since the last four terms in (27) can be casted in the stated form with  $r = k - i - 1$ ,  $q = k - j - 1$  and  $M_{rq}$  componentwise bounded (by Cauchy-Schwarz inequality and the assumptions on  $u_k$ ,  $w_k$ ) and hence  $\|\cdot\|$  bounded due to the norm equivalence.  $\blacksquare$

**Corollary 1.** *The receding horizon implementation of the control policy defined by solving the optimization problem (18) every  $N_c \leq N$  steps and applying the first  $N_c$  control inputs generated by the policy  $u = \eta + Ke(w)$  renders the state  $x_k$  mean-square bounded provided that  $\rho(A) < 1$ .*

*Proof.* Follows directly from Theorem 3 since the constraints in (18) ensure that the inputs stay bounded. ■

In the case of  $\rho(A) = 1$  with the deterministic part of the system (2) Lyapunov stable, the sole assumption of bounded control inputs is insufficient, and another constraint must be embedded into (18) in order to ensure the mean-square boundedness of the state; see [12] for details.

#### IV. NUMERICAL EXAMPLES

We present two numerical examples that compares our method to other control strategies. With the gradient and Hessian on hand, the problem (18) can be solved by a nonlinear solver with guaranteed convergence because of convexity or by a general purpose convex solver. For our small scale examples we managed with the Matlab nonlinear solver implemented in the `FMINCON` function with the ‘interior-point’ option as well as with a custom interior-point solver. Nondifferentiability of the objective is not a problem in our case since the optimization path and the solution itself lie outside the nondifferentiable region. If this were not the case, which can happen if the penalty on control effort is large leading to zero mean and zero variance of one of the control inputs, various techniques for nondifferentiable convex optimization can be employed.

In the first example we consider a fixed horizon stochastic control problem. For the system matrices and the noise covariance matrix we chose

$$A = \begin{bmatrix} 1 & -0.4 \\ 0.1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}, \quad \Sigma = I \otimes \begin{bmatrix} 8 & 5 \\ 5 & 6 \end{bmatrix}$$

with  $w_k$  zero-mean jointly Gaussian. We set the weighting matrices  $Q = I$ ,  $R = 0.1I$ , and the input constraints to  $U_{\max} = 30$ . The optimization horizon is  $T = 12$  and the initial state  $x_0 = [1, -1]^T$ . The function  $e(w)$  was chosen as suggested above to be the componentwise saturation that saturates the disturbances at  $4\sqrt{\rho(\Sigma)} = 13.9$ .

We compared our control policy (NDF) (with  $N_c = N = T$ ) with the standard certainty equivalent MPC ( $N_c = 1$ ,  $N = T$ ) and with the shrinking horizon certainty equivalent MPC (SH-MPC) ( $N_c = 1$ ,  $N(k) = T - k$ ,  $k = 0, \dots, T - 1$ ). Furthermore, we tried out the proposed method with  $K = 0$  against the certainty equivalent open loop control (OL) (i.e., CE-MPC with  $N_c = N = T$ ). For the sake of completeness we also tried out our method in the shrinking horizon mode (SH-NDF) with  $N_c = 2$ ,  $N(k) = T - k$ . The respective objective functions were evaluated using 2000 Monte Carlo runs. The results are summarized in Table I, 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. Figure 1 shows

histograms of the proposed policy and the two MPC policies. Finally, we evaluate the bound (26) which yields  $\beta = 0.005$  showing that the solution found by (18) is in this case basically optimal in (6).

TABLE I

COMPARISON OF CONTROL POLICIES OVER THE OPTIMIZATION HORIZON  $T = 12$ .

Policy	NDF-SH	NDF	SH-MPC	MPC	NDF, $K = 0$	OL
$J$	86.8	92.1	98.3	119.2	140.4	143.9

Our second example compares the proposed method with the certainty equivalent MPC in a receding horizon mode. In this example we consider the respective matrices

$$A = \begin{bmatrix} 1 & 1 \\ -0.5 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{E}\{w_i w_j^T\} = \begin{bmatrix} 8 & 5 \\ 5 & 6 \end{bmatrix} \delta_{ij},$$

where  $w_k$  is an i.i.d. sequence of zero-mean Gaussian random variables, and  $\delta_{ij}$  denotes the Kronecker delta. The weighting matrices were set to  $Q = I$  and  $R = 0.1I$ , the input constraints to  $U_{\max} = 10$ , and the initial state to  $x_0 = [1, -1]^T$ . We compared our control policy with  $N = 12$ ,  $N_c = 4$  against CE-MPC with  $N = 12$ ,  $N_c = 1$  in a receding horizon fashion over the simulation time  $T = 100$ . Again, we used the 4-sigma rule to get  $\varepsilon = 13.9$ . Figure 2 shows the accumulation of the cost over the simulation time, whereas Figure 3 depicts the evolution of the estimated mean-square of the state. The latter is bounded, which is in accordance with Corollary 1 since  $\rho(A) = \sqrt{2}/2 < 1$ . One hundred Monte Carlo runs were used to evaluate the cost functions.

#### V. CONCLUSION

In this article, we dealt exclusively with the expectation of the 1-norm stochastic optimal 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 disturbance feedback control policies. Finally, we provided a simple proof of receding horizon stability of the proposed policy, and demonstrated our results by means of two numerical examples.

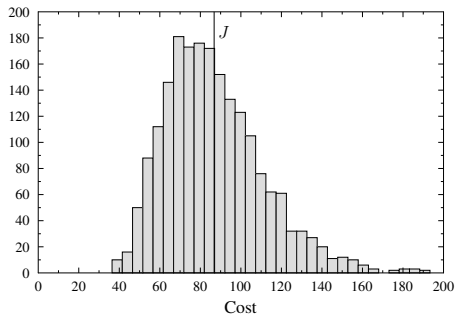
Future work will concentrate on extending the results of this article to a general p-norm and to the output feedback case. Furthermore, the question of the mean-square stabilizability of Lyapunov unstable systems with the system matrix of spectral radius one remains, at least to our knowledge, open.

#### ACKNOWLEDGMENT

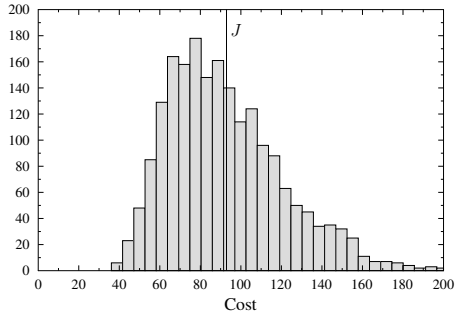
This work has been supported by Czech Technical University through grant No. SGS10/283/OHK3/3T/13.

#### REFERENCES

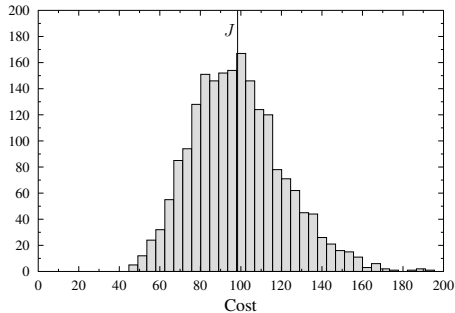
- [1] D. P. Bertsekas. Dynamic programming and suboptimal control: A survey from ADP to MPC. In *CDC Proceedings*, 2005.
- [2] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, third edition, 2007.



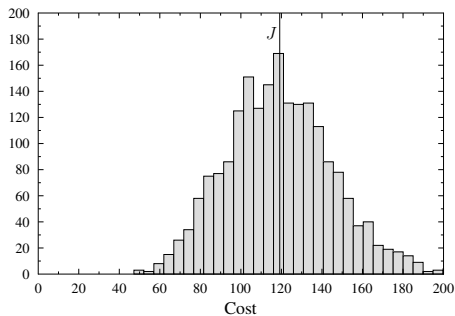
(a) NDF-SH with  $N_c = 2$ ,  $N(k) = T - k$



(b) NDF with  $N_c = N = T$



(c) Shrinking horizon CE-MPC ( $N_c = 1$ ,  $N(k) = T - k$ )



(d) Certainty-equivalent MPC ( $N_c = 1$ ,  $N = T$ )

Fig. 1. Histograms of the costs of different control policies over 2000 Monte Carlo runs on the optimization horizon  $T = 12$ .

- [3] D. Bertsimas and D. Brown. Constrained stochastic LQC: a tractable approach. *IEEE Transactions on Automatic Control*, 52(10):1826–1841, 2007.
- [4] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [5] M. Cannon, B. Kouvaritakis, and D. Ng. Probabilistic tubes in linear stochastic model predictive control. *Systems & Control Letters*, 58(10-11):747–753, 2009.

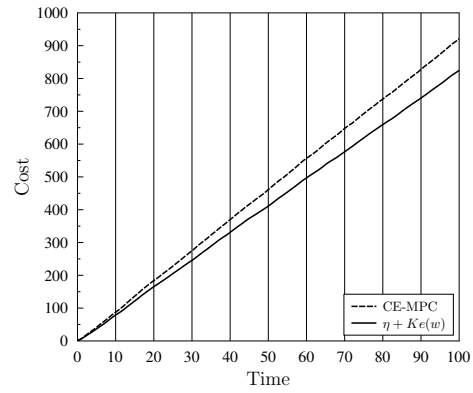


Fig. 2. Comparison of costs over the simulation time  $T = 100$  in a receding horizon mode with  $N = 12$  and  $N_c = 4$  for our policy and  $N_c = 1$  for CE-MPC. Final costs are 824.7 for our control scheme and 921.1 for CE-MPC.

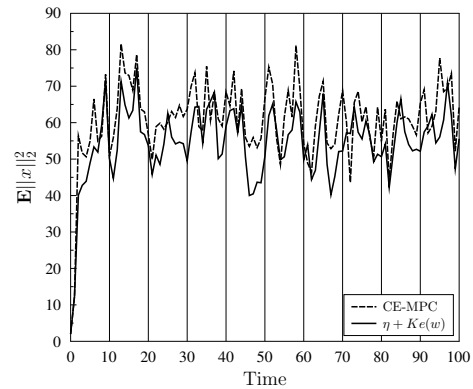


Fig. 3. Evolution of  $\mathbf{E}\|x\|_2^2$  under our receding horizon control policy with  $N = 12$ ,  $N_c = 4$  and CE-MPC control policy with  $N = 12$ ,  $N_c = 1$ .

- [6] M. Cannon, B. Kouvaritakis, and X. Wu. Model predictive control for systems with stochastic multiplicative uncertainty and probabilistic constraints. *Automatica*, 45(1):167–172, 2009.
- [7] D. Chatterjee, P. Hokayem, and J. Lygeros. Stochastic receding horizon control with bounded control inputs: a vector space approach. Technical report, ETH Zurich, Switzerland, 2009.
- [8] B. Kouvaritakis, M. Cannon, S. V. Raković, and Q. Cheng. Explicit use of probabilistic distributions in linear predictive control. *Automatica*, 46(10):1719–1724, 2010.
- [9] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996, 2007.
- [10] F. Oldewurtel, C. N. Jones, and M. Morari. A tractable approximation of chance constrained stochastic MPC based on affine disturbance feedback. In *47th IEEE Conference on Decision and Control*, pages 4731–4736, 2008.
- [11] F. Oldewurtel, A. Parisio, C. N. Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and K. Wirth. Energy efficient building climate control using stochastic model predictive control and weather predictions. In *American Control Conference*, pages 5100–5105, 2010.
- [12] F. Ramponi, D. Chatterjee, A. Miliadis-Argentis, P. Hokayem, and J. Lygeros. Attaining mean square boundedness of a marginally stable stochastic linear system with a bounded control input. *IEEE Transactions on Automatic Control*, 55(10):2414–2418, 2010.
- [13] J. Skaf and S. Boyd. Nonlinear Q-Design for convex stochastic control. *IEEE Transactions on Automatic Control*, 54(10):2426–2430, 2009.
- [14] J. Skaf and S. Boyd. Design of affine controllers via convex optimization. *IEEE Transactions on Automatic Control*, 55(11):2476–2487, 2010.