Stochastic MPC with Imperfect State Information and Bounded Controls ⋆

Peter Hokayem * **Eugenio Cinquemani** ** **Debasish Chatterjee** * **John Lygeros** *

* *Automatic Control Laboratory, ETH Zürich (e-mail: {hokayem, chatterjee, lygeros}@control.ee.ethz.ch)*
** *INRIA Grenoble-Rhône-Alpes (e-mail: Eugenio.Cinquemani@inria.fr)*

Abstract: This paper addresses the problem of output feedback Model Predictive Control for stochastic linear systems, with hard and soft constraints on the control inputs as well as soft constraints on the state. We use the so-called purified outputs along with a suitable nonlinear control policy and show that the resulting optimization program is convex. We also show how the proposed method can be applied in a receding horizon fashion. Contrary to the state feedback case, the receding horizon implementation in the output feedback case requires the update of several optimization parameters and the recursive computation of the conditional probability densities of the state given the previous measurements. Algorithms for performing these tasks are developed.

Keywords: Predictive control, output feedback, convex optimization, linear systems, observers

1. INTRODUCTION

Over the last decades Model Predictive Control (MPC) has been successful in addressing industrial problems due mainly to its ability to handle input and state constraints. In the deterministic setting there exists a plethora of literature that settles tractability and stability of MPC, see for example, [Mayne et al., 2000, Bemporad and Morari, 1999, Maciejowski, 2001] and the references therein. Results in the stochastic case, however, are fewer.

Research on stochastic MPC is broadly subdivided into two parallel lines: the first treats multiplicative noise that enters the state equations, and the second caters to additive noise. The case of multiplicative noise has been treated in [Primbs and Sung, 2009, Cannon et al., 2009a,b, Couchman et al., 2006]. In [Primbs and Sung, 2009], the noise enters the state equation multiplicatively, mixed hard state-input constraints are relaxed into expectation constraints, and results pertaining to feasibility and stability are presented for the full state feedback case. The authors in [Couchman et al., 2006] treat the case of uncertain output measurement matrix (C) and solve the MPC problem under probabilistic constraints on the outputs and full state feedback. In [Cannon et al., 2009a] the stochastic MPC problem is treated under full state feedback and multiplicative noise entering the state equation. The proposed scheme comprises a pre-stabilizing linear state feedback part (which we call a pre-stabilizing controller,) and an open-loop part. The pre-stabilizing feedback gain is computed off-line and just the open-loop part is left to online optimization. [Cannon et al., 2009b] extends the in [Cannon et al., 2009a] to the case of additive noise as well. However, both results [Cannon et al., 2009b] and [Cannon et al., 2009a] involve a pre-stabilizing controller and hence no hard control bounds can be imposed.

We focus in this article on the additive noise case. The approach proposed here stems from and generalizes the idea of affine parametrization of control policies for finite-horizon linear quadratic problems proposed in [Ben-Tal et al., 2004, 2006], utilized within the robust MPC framework in [Ben-Tal et al., 2006, Löfberg, 2003, Goulart et al., 2006] for full state feedback, and in [van Hessem and Bosgra, 2003] for output feedback with Gaussian state and measurement noise inputs. More recently, this affine approximation was utilized in [Skaf and Boyd, 2009] for both the robust deterministic and the stochastic setups in the absence of control bounds, and optimality of affine policies in the scalar deterministic case was reported in [Bertsimas et al., 2009] within the robust framework. In [Bertsimas and Brown, 2007] the authors reformulate the stochastic programming problem as a deterministic one with bounded noise support and solve a robust optimization problem over a finite horizon, followed by estimating the performance when the noise can take unbounded values. Similar approach was utilized in [Oldewurtel et al., 2008] for affine input policies. We also mentions related works employing randomized algorithms as in [Batina, 2004, Blackmore and Williams, 2007, Maciejowski et al., 2005]. Results on obtaining lower bounds on the value functions of the stochastic optimization problem have been recently reported in [Wang and Boyd, 2009], and a novel stochastic MPC scheme based on the scenario approach has appeared in [Bernardini and Bemporad, 2009].

⋆ This research was partially supported by the Swiss National Science Foundation under grant 200021-122072 and by the European Commission under the project Feednetback FP7-ICT-223866 (www.feednetback.eu).

In this article we restrict attention to discrete-time linear time-invariant controlled systems with affine stochastic disturbance inputs and noisy measurement outputs. Over a single optimization horizon N, the control input at each time step t is chosen from the class of nonlinear causal strategies, i.e., as a nonlinear function of all measured outputs up to time t. In the affine feedback case this strategy has been shown to be one-to-one equivalent to feedback from the so-called purified outputs [Ben-Tal et al., 2006], with the main difference being that the latter results in a convex optimization problem. As in the affine feedback case, we also use the purified outputs formulation in the design of our feedback strategies. However, as we have not assumed any compactness on the possible noise affecting the state and measured outputs, it is generally impossible to guarantee the satisfaction of the hard constraints on the control inputs. Therefore, we utilize the method proposed in [Hokayem et al., 2009, Chatterjee et al., 2009], in which we saturate the signal being used for feedback (in this case the purified outputs) before utilizing them to construct the control input vectors. This yields a natural way of dealing with possibly unbounded noise that may affect the system under hard control bounds without losing convexity. Furthermore, we generalize this method to incorporate certain expectation-type constraints. To apply this formulation in a receding horizon fashion, it is necessary to compute certain time-dependent matrix parameters (expectations) of the optimal control problem. In turn, this requires propagating the conditional density of the state with respect to the measured outputs. We report formulas for propagating this conditional density. In the case of Gaussian noise acting on the system, we provide a low-complexity algorithm (essentially similar to Kalman filtering) for updating this conditional density and efficient solutions for the computation of the time-dependent optimization parameters.

Notation. Hereafter, $\mathbb{N}_+ := \{1, 2, \ldots\}$ is the set of natural numbers and $\mathbb{N} := \mathbb{N}_+ \cup \{0\}$. Let $\mathbb{E}_{x_0}[\cdot]$ denote the expected value given x_0, and $\mathbf{tr}(\cdot)$ denote the trace of a matrix. Let $\|\cdot\|_p$ denote the standard ℓ_p norm and $\|v\|_M = \sqrt{v^T M v}$ denote the weighted ℓ_2-norm, for $M \geqslant 0$. For any matrix M, $(M)_i$ denotes its i-th row.

2. PROBLEM STATEMENT

Consider the following affine discrete-time stochastic dynamical model:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t + l, \\ y_t &= Cx_t + v_t, \end{aligned} \tag{1}$$

where $t \in \mathbb{N}$, $x_t \in \mathbb{R}^n$ is the state, $u_t \in \mathbb{R}^m$ is the control input, $y_t \in \mathbb{R}^p$ is the output, $w_t \in \mathbb{R}^n$ is a random process noise, and $v_t \in \mathbb{R}^p$ is a random measurement noise. The matrices A, B, and C are known and $l \in \mathbb{R}^n$ is a known constant vector. We assume that at any time t, the output vector y_t is observed. We shall require hereafter that the control input vector is bounded at each instant of time t, i.e.,

$$u_t \in \left\{ u \in \mathbb{R}^m \,\middle|\, \|u\|_\infty \leqslant U_{\max} \right\} \qquad \forall\, t \in \mathbb{N}, \tag{2}$$

where $U_{\max} > 0$ is some given saturation bound. Note that the process model (1) with constraints (2) can handle a wide range of convex polytopic input constraints [Hokayem et al., 2009].

Fix an optimization (or prediction) horizon $N \in \mathbb{N}_+$, and consider the following cost at time t

$$\mathcal{J}_t = \mathbb{E}_{\mathcal{Y}_t}\left[\sum_{k=0}^{N-1}\left(\|x_{t+k}\|^2_{Q_k} + \|u_{t+k}\|^2_{R_k}\right) + \|x_{t+N}\|^2_{Q_N}\right], \tag{3}$$

where $\mathcal{Y}_t = \sigma\{y_0, y_1, \ldots, y_t\}$ is the σ-algebra generated by the output vectors up to time t, and $Q_k = Q_k^T \geqslant 0$, $Q_N = Q_N^T \geqslant 0$, and $R_k = R_k^T \geqslant 0$ are given matrices of appropriate dimension.

The evolution of the system (1) over a single optimization horizon N starting at time t can be described in a compact form as follows:

$$X_t = \mathcal{A}x_t + \mathcal{B}U_t + \mathcal{D}W_t + L, \quad Y_t = \mathcal{C}X_t + V_t \tag{4}$$

where $X_t = \begin{bmatrix} x_t \\ x_{t+1} \\ \vdots \\ x_{t+N} \end{bmatrix}$, $U_t = \begin{bmatrix} u_t \\ u_{t+1} \\ \vdots \\ u_{t+N-1} \end{bmatrix}$, $W_t = \begin{bmatrix} w_t \\ w_{t+1} \\ \vdots \\ w_{t+N-1} \end{bmatrix}$, $Y_t = \begin{bmatrix} y_t \\ y_{t+1} \\ \vdots \\ y_{t+N} \end{bmatrix}$, $V_t = \begin{bmatrix} v_t \\ v_{t+1} \\ \vdots \\ v_{t+N} \end{bmatrix}$, $\mathcal{A} = \begin{bmatrix} I \\ A \\ \vdots \\ A^N \end{bmatrix}$, $\mathcal{B} = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ B & \ddots & & \vdots \\ AB & B & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ A^{N-1}B & \cdots & AB & B \end{bmatrix}$, $\mathcal{D} = \begin{bmatrix} 0 & \cdots & \cdots & 0 \\ I & \ddots & & \vdots \\ A & I & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ A^{N-1} & \cdots & A & I \end{bmatrix}$, $\mathcal{C} = \mathrm{diag}\{C, \cdots, C\}$, $L = \mathcal{D}\begin{bmatrix} l \\ l \\ \vdots \\ l \end{bmatrix}$, and the inputs are required to satisfy

$$U_t \in \left\{ U \in \mathbb{R}^{Nm} \,\middle|\, \|U\|_\infty \leqslant U_{\max} \right\} \qquad \forall t \in \mathbb{N}. \tag{5}$$

Using the compact notation above, the cost function $\mathcal{J}_t$ in (3) can be rewritten as

$$\mathcal{J}_t = \mathbb{E}_{\mathcal{Y}_t}\left[\|X_t\|^2_{\mathcal{Q}} + \|U_t\|^2_{\mathcal{R}}\right], \tag{6}$$

where $\mathcal{Q} = \mathrm{diag}\{Q_0, \cdots, Q_N\}$ and $\mathcal{R} = \mathrm{diag}\{R_0, \cdots, R_{N-1}\}$.

In practice, it may be also of interest to impose further some soft constraints both on the state and the input vector. For example, one may be interested in imposing quadratic or linear constraints on the state, both of which can be captured via the constraint

$$\mathbb{E}_{\mathcal{Y}_t}\left[\|X_t\|^2_{\mathcal{S}} + \mathcal{L}^T X_t\right] \leqslant \alpha, \tag{7}$$

where $\mathcal{S} = \mathcal{S}^T \geqslant 0$. Moreover, expected energy expenditure constraints can be posed as

$$\mathbb{E}_{\mathcal{Y}_t}\left[\|U_t\|^2_{\tilde{\mathcal{S}}}\right] \leqslant \beta_k, \tag{8}$$

where $\tilde{\mathcal{S}} = \tilde{\mathcal{S}}^T \geqslant 0$. In the absence of hard input constraints, such expectation-type constraints are commonly used in the stochastic MPC [Primbs and Sung, 2009, Agarwal et al., 2009] and in stochastic optimization in the form of integrated chance constraints [Klein Haneveld and van der Vlerk, 2006]. This is partly because it is not possible, without posing further restrictions on the boundedness of the process noise w_t, to ensure that hard constraints on the state are satisfied. For example, in the standard LQG setting nontrivial hard constraints on the system state would generally be violated with nonzero probability. Moreover, in contrast to chance constraints where a bound is imposed on the probability of constraint violation, expectation-type constraints tend to give rise to convex optimization problems under weak assumptions [Agarwal et al., 2009, Klein Haneveld and van der Vlerk, 2006].

At each time $t \in \mathbb{N}$, we are interested in solving the following optimization problem

$$\min_{U_t \in \mathcal{G}} \{\mathcal{J}_t \mid \ (4), (5), (7), (8)\}, \tag{9}$$

where $\mathcal{G}$ is the class of all causal output feedback policies. An explicit solution to Problem (9) over the general class of causal output feedback policies is extremely difficult to obtain in general, see, for example, [Bertsekas, 2000, 2007]. A feasible way to circumvent this difficulty is to restrict $\mathcal{G}$ to a specific subclass of policies. This will result in a suboptimal solution to (9), but may yield a tractable optimization problem. This is the track we pursue.

3. TRACTABLE CONTROL POLICIES

Given the dynamical system (1), and guided by the setup in [Ben-Tal et al., 2006], let us recall the definition of the so-called purified outputs. For any $t \in \mathbb{N}$, define the model

$$\begin{aligned} \bar{x}_{t+1} &= A\bar{x}_t + Bu_t + l, \\ \bar{y}_t &= C\bar{x}_t, \\ z_t &= y_t - \bar{y}_t, \end{aligned} \tag{10}$$

where $\bar{x}_t \in \mathbb{R}^n$. The vectors z_t are called purified outputs and can be computed over the optimization horizon N using (10) and the measured outputs y_t. The dynamics in (10) over the optimization horizon N starting at time t can be compactly written as

$$\begin{aligned} \bar{X}_t &= \mathcal{A}\bar{x}_t + \mathcal{B}U_t + L, \\ \bar{Y}_t &= \mathcal{C}\bar{X}_t, \\ Z_t &= Y_t - \bar{Y}_t, \end{aligned} \tag{11}$$

where $\bar{X}_t = \left[\bar{x}_t^T, \cdots, \bar{x}_{t+N}^T\right]^T$, and all other terms are defined as in (4). A particularly convenient choice of the initial condition is $\bar{x}_t = \hat{x}_{t|t} = \mathbb{E}_{\mathcal{Y}_t}[x_t]$, i.e., the optimal Bayesian estimate of x_t in a mean-square sense, which we will assume from now on and will specify how to compute $\hat{x}_{t|t}$ in Section 4. It follows from (4) and (11) that

$$Z_t = \mathcal{C}(X_t - \bar{X}_t) + V_t = \mathcal{CA}(x_t - \bar{x}_t) + \mathcal{CD}W_t + V_t. \tag{12}$$

Therefore, the purified outputs do not depend on the control inputs vector U_t.

We shall utilize the purified outputs z_t in designing the control inputs over a single optimization horizon N, and restrict the feedback functions to lie in the class of causal nonlinear feedback policies of the form

$$u_{t+\ell} = \eta_{t+\ell} + \sum_{i=0}^{\ell} \theta_{t+\ell,t+i}\varphi_i(z_{t+i}) \quad \forall \ell = 0, \cdots, N-1 \tag{13}$$

where $\varphi_i(z_{t+i})$ is a shorthand for the vector $\left[\varphi_i^1(z_{t+i}^1), \ldots, \varphi_i^p(z_{t+i}^p)\right]^T$, where z_{t+i}^j is the j-th element of the vector z_{t+i} and $\varphi_i^j : \mathbb{R} \to \mathbb{R}$ is any function with $\sup_{s\in\mathbb{R}} |\varphi_i^j(s)| \leqslant \varphi_{\max} < \infty$. The feedback gains $\theta_{\ell,i} \in \mathbb{R}^{m\times p}$ and the affine terms $\eta_\ell \in \mathbb{R}^m$ must be chosen based on the control objective, while observing the constraints (2). With this definition, the value of u at time ℓ depends on the values of the measured outputs from time t up to time $t+\ell$ only.

Of course, this choice of policies is generally suboptimal, but it will ensure the tractability of a large class of optimal control problems. Note that we have chosen to saturate the measurements we obtain from the vectors z_t before inserting them into the control vector. This way we do not need to assume that the noise distributions are defined over a compact domain, which is a major difference with respect to robust MPC approaches, see, for example, [Bertsimas and Brown, 2007, Mayne et al., 2000] and references therein. The choice of the element-wise saturation functions $\varphi_i(\cdot)$ is left open. As such, we can accommodate standard saturation, piecewise linear, and sigmoidal functions, to name a few.

The control input sequence over the optimization horizon N starting at time t can be written as

$$U_t = \boldsymbol{\eta}_t + \boldsymbol{\Theta}_t\varphi(Z_t) \tag{14}$$

where $\boldsymbol{\eta}_t := \begin{bmatrix} \eta_t \\ \vdots \\ \eta_{t+N-1} \end{bmatrix}$, $\varphi(Z_t) := \begin{bmatrix} \varphi_0(z_t) \\ \vdots \\ \varphi_{N-1}(z_{t+N-1}) \end{bmatrix}$, and $\boldsymbol{\Theta}_t$ has the following structure

$$\boldsymbol{\Theta}_t := \begin{bmatrix} \theta_{t,t} & 0 & \cdots & 0 \\ \theta_{t+1,t} & \theta_{t+1,t+1} & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ \theta_{t+N-1,t} & \theta_{t+N-1,t+1} & \cdots & \theta_{t+N-1,t+N-1} \end{bmatrix} \tag{15}$$

Problem (9) can now be written in terms of the new policy in (14) as

$$\min_{(\boldsymbol{\eta}_t, \boldsymbol{\Theta}_t)} \{\mathcal{J}_t \mid \ (4), (5), (7), (8), (14), \text{and } (15)\}. \tag{16}$$

Our choice of using the purified outputs for feedback and the policy structure in (14) ensures convexity of the optimization problem (16), as seen in the next result.

Assumption 1. x_0, $(w_t)_{t\in\mathbb{N}}$, and $(v_t)_{t\in\mathbb{N}}$ comprise mutually independent zero-mean i.i.d. random vectors for all $t \in \mathbb{N}$, with known probability densities.

Proposition 2. The optimization problem (16) *is convex and equivalent to the following quadratically constrained quadratic optimization problem:*

$$\begin{aligned} \min_{(\boldsymbol{\eta}_t, \boldsymbol{\Theta}_t)} \ & 2\mathbf{tr}\Big(\boldsymbol{\Theta}_t^T \mathcal{B}^T \mathcal{Q}(\mathcal{A}\Lambda_t^{x\varphi} + \mathcal{D}\Lambda_t^{w\varphi} + L\Lambda_t^{\varphi T})\Big) \\ & + 2\boldsymbol{\eta}_t^T \mathcal{B}^T \mathcal{Q}(\mathcal{A}\hat{x}_{t|t} + L) + \|\boldsymbol{\eta}_t + \boldsymbol{\Theta}_t\Lambda_t^{\varphi}\|_{\mathcal{M}}^2 \\ & + \mathbf{tr}\Big(\boldsymbol{\Theta}_t^T \mathcal{M}\boldsymbol{\Theta}_t(\Lambda_t^{\varphi\varphi} - \Lambda_t^{\varphi}\Lambda_t^{\varphi T})\Big) \end{aligned}$$

subject to

$$\begin{aligned} & \textit{the structure of } \boldsymbol{\Theta}_t \textit{ in } (15), \\ & |(\boldsymbol{\eta}_t)_i| + \|(\boldsymbol{\Theta}_t)_i\|_1 \, \varphi_{\max} \leqslant U_{\max} \quad \forall i = 1, \cdots, Nm, \\ & 2\mathbf{tr}\Big(\boldsymbol{\Theta}_t^T \mathcal{B}^T \mathcal{S}(\mathcal{A}\Lambda_t^{x\varphi} + \mathcal{D}\Lambda_t^{w\varphi} + L\Lambda_t^{\varphi T})\Big) \\ & \quad + 2\boldsymbol{\eta}_t^T \mathcal{B}^T \mathcal{S}(\mathcal{A}\hat{x}_{t|t} + L) + \|\boldsymbol{\eta}_t + \boldsymbol{\Theta}_t\Lambda_t^{\varphi}\|_{\mathcal{B}^T\mathcal{QB}}^2 \\ & \quad + \mathbf{tr}\Big(\boldsymbol{\Theta}_t^T \mathcal{B}^T \mathcal{QB}\boldsymbol{\Theta}_t(\Lambda_t^{\varphi\varphi} - \Lambda_t^{\varphi}\Lambda_t^{\varphi T})\Big) \\ & \quad + \mathcal{L}^T(\mathcal{A}\hat{x}_{t|t} + \mathcal{B}\boldsymbol{\eta}_t + \mathcal{B}\boldsymbol{\Theta}_t\Lambda_t^{\varphi} + L) \\ & \quad + \mathbf{tr}\big(\mathcal{A}^T\mathcal{SA}\Sigma_{x_t}\big) + \mathbf{tr}\big(\mathcal{D}^T\mathcal{SD}\Sigma_W\big) \\ & \quad + L^T\mathcal{S}L + 2\hat{x}_{t|t}^T\mathcal{A}^T\mathcal{S}L \leqslant \alpha, \end{aligned} \tag{17}$$

$$\begin{aligned} & \|\boldsymbol{\eta}_t + \boldsymbol{\Theta}_t\Lambda_t^{\varphi}\|_{\tilde{\mathcal{S}}}^2 + \mathbf{tr}\Big(\boldsymbol{\Theta}_t^T \tilde{\mathcal{S}}\boldsymbol{\Theta}_t(\Lambda_t^{\varphi\varphi} - \Lambda_t^{\varphi}\Lambda_t^{\varphi T})\Big) \\ & \quad \leqslant \beta, \end{aligned} \tag{18}$$

where $\mathcal{M} = \mathcal{R} + \mathcal{B}^T\mathcal{Q}\mathcal{B}$, $\Sigma_{x_t} = \mathbb{E}_{\mathcal{Y}_t}[x_t x_t^T]$, $\Sigma_W = \mathbb{E}[W_t W_t^T]$, $\Lambda_t^{\varphi} = \mathbb{E}_{\mathcal{Y}_t}[\varphi(Z_t)]$, $\Lambda_t^{x\varphi} = \mathbb{E}_{\mathcal{Y}_t}[x_t\varphi(Z_t)^T]$, $\Lambda_t^{w\varphi} = \mathbb{E}_{\mathcal{Y}_t}[W_t\varphi(Z_t)^T]$, *and* $\Lambda_t^{\varphi\varphi} = \mathbb{E}_{\mathcal{Y}_t}[\varphi(Z_t)\varphi(Z_t)^T]$. *Moreover, the resulting optimization problem is feasible whenever* $\alpha \geqslant \mathbf{tr}(\mathcal{A}^T\mathcal{S}\mathcal{A}\Sigma_{x_t}) + \mathbf{tr}(\mathcal{D}^T\mathcal{S}\mathcal{D}\Sigma_W) + L^T\mathcal{S}L + 2\hat{x}_{t|t}^T\mathcal{A}^T\mathcal{S}L + \mathcal{L}^T(\mathcal{A}\hat{x}_{t|t} + L)$ *and* $\beta \geqslant 0$.

The proof of Proposition 2 is omitted in the interest of space. The optimization problem in Proposition 2 is a quadratically constrained quadratic program (QCQP) in the optimization parameters $(\boldsymbol{\eta}_t, \boldsymbol{\Theta}_t)$ and is a quadratic program (QP) whenever the constraints (17)-(18) are not present. As such, one may use software packages such as `cvx` [Grant and Boyd, 2000] or `yalmip` [Löfberg, 2004] to solve it.

Constraints on the variation of the inputs of the form $\|PU_t\|_\infty \leqslant \Delta_{\max}$, where $P = \begin{bmatrix} I & -I & 0 & \cdots & 0 \\ 0 & I & -I & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & 0 & I & -I \end{bmatrix}$, can be translated into a constraint similar to that in (5). Note also that some of the block gains in the matrix $\boldsymbol{\Theta}_t$ in (15) can be set to zero, hence reducing the number of optimization variables and as such the computational burden.

The matrices $\Lambda_t^{\varphi}, \Lambda_t^{x\varphi}, \Lambda_t^{w\varphi}, \Lambda_t^{\varphi\varphi}$ are fixed parameters of the optimization problem determined by the conditional density of x_t given $\mathcal{Y}_t$. In general, no closed-form expression exists, hence the computation of $\Lambda_t^{\varphi}, \Lambda_t^{x\varphi}, \Lambda_t^{w\varphi}, \Lambda_t^{\varphi\varphi}$ must be performed numerically prior to optimization. In light of (12), the choice of $\hat{x}_{t|t}$ affects the statistics of Z_t and hence the optimization problem via $\Lambda_t^{\varphi}, \Lambda_t^{x\varphi}, \Lambda_t^{w\varphi}, \Lambda_t^{\varphi\varphi}$. The computation of the conditional statistics $f(x_t|\mathcal{Y}_t)$, $\hat{x}_t$, and Σ_{x_t} will be discussed in Section 4.

4. MPC IMPLEMENTATION OF THE OPTIMAL CONTROL POLICY

To apply the procedure of Section 3 in an MPC fashion, one needs to be able to compute the expectations in the cost and constraints at every time step. That is, for all $t \in \mathbb{N}$ we need to compute the following quantities

$$\begin{aligned} &\Lambda_t^{\varphi} = \mathbb{E}_{\mathcal{Y}_t}[\varphi(Z_t)], && \Lambda_t^{w\varphi} = \mathbb{E}_{\mathcal{Y}_t}[W_t\varphi(Z_t)^T], \\ &\Lambda_t^{\varphi\varphi} = \mathbb{E}_{\mathcal{Y}_t}[\varphi(Z_t)\varphi(Z_t)^T], && \Lambda_t^{e\varphi} = \mathbb{E}_{\mathcal{Y}_t}[(x_t - \hat{x}_t)\varphi(Z_t)^T] \\ &\Lambda_t^{x\varphi} = \hat{x}_t\Lambda_t^{\varphi T} + \Lambda_t^{e\varphi}, \end{aligned} \tag{19}$$

where $Z_t = \mathcal{C}\mathcal{A}(x_t - \hat{x}_{t|t}) + \mathcal{C}\mathcal{D}W_t + V_t$. We define the quantity $\hat{x}_{t|t}$ as $\hat{x}_{t|t} = \mathbb{E}_{\mathcal{Y}_t}[x_t]$, i.e., the optimal causal Bayesian estimate of x_t in a mean-square sense. The computation of (19) requires propagating in time conditional density of the state given the previous and current output measurements, denoted by $f(x_t|\mathcal{Y}_t)$. We propose an iterative method for the computation of $f(x_t|\mathcal{Y}_t)$. The iterative computation will naturally involve the computation of densities $f(x_t|\mathcal{Y}_{t-1})$, where we set $f(x_0|\mathcal{Y}_{-1}) = f(x_0)$ — the probability density of the initial state x_0. For $t, s \in \mathbb{N}_0$, define $\hat{x}_{t|s} = \mathbb{E}_{\mathcal{Y}_s}[x_t]$ and $P_{t|s} = \mathbb{E}_{\mathcal{Y}_t}[(x_t - \hat{x}_{t|s})(x_t - \hat{x}_{t|s})^T]$. We posit the following extra assumption.

Assumption 3. In addition to Assumption 1, we require that $w_t \sim \mathcal{N}(0, \Sigma_w)$, $v_t \sim \mathcal{N}(0, \Sigma_v)$ and $x_0 \sim \mathcal{N}(\hat{x}_0, P_0)$, where $0 < \Sigma_w \in \mathbb{R}^{n\times n}$, $0 < \Sigma_v \in \mathbb{R}^{p\times p}$, $0 < P_0 \in \mathbb{R}^{n\times n}$ and $\hat{x}_0 \in \mathbb{R}^n$ are known.

Proposition 4. Let Assumption 3 hold. Then $f(x_t|\mathcal{Y}_t)$ *and* $f(x_{t+1}|\mathcal{Y}_t)$ *are the probability densities of Gaussian distributions* $\mathcal{N}(\hat{x}_{t|t}, P_{t|t})$ *and* $\mathcal{N}(\hat{x}_{t+1|t}, P_{t+1|t})$*, respectively, with* $P_{t|t} > 0$ *and* $P_{t+1|t} > 0$*. For all* $t \in \mathbb{N}$ *their conditional means and covariances can be computed iteratively as follows:*

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + P_{t|t-1}C^T(CP_{t|t-1}C^T + V)^{-1}(y_t - C\hat{x}_{t|t-1}), \tag{20}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}C^T(CP_{t|t-1}C^T + V)^{-1}CP_{t|t-1}, \tag{21}$$

where

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t, \tag{22}$$

$$P_{t+1|t} = AP_{t|t}A^T + W, \tag{23}$$

$\hat{x}_{0|-1} = \hat{x}_0$, *and* $P_{0|-1} = P_0$.

A proof of Proposition 4 may be found in Kumar and Varaiya [1986]. In particular, the matrix $P_{t|t}$ plays the role similar to that of Σ_{x_t} in Proposition 2, and together with $\hat{x}_{t|t}$, $P_{t|t}$ characterizes the conditional density $f(x_t|\mathcal{Y}_t)$. Proposition 4 states that the conditional mean and covariances of x_t can be propagated by an iterative algorithm which closely resembles the Kalman filter. Since the system is generally nonlinear due to the fact that u_t is a nonlinear function of the previous outputs, we cannot assume that the probability distributions in the problem are Gaussian (in fact, the a priori distributions of x_t and of $\mathcal{Y}_t$ are not) and the proof cannot be developed in the standard linear framework of the Kalman filter. Intuitively, Proposition 4 holds due to the fact that once the measurements are available, the input becomes a fixed known quantity, which renders the system conditionally linear and Gaussian.

At any time t, the matrices (19) may be computed by numerical or Monte Carlo integration with respect to the independent Gaussian measures of $w_t, \ldots w_{t+N-1}$, of $v_t, \ldots v_{t+N-1}$, and of x_t given $\mathcal{Y}_t$. Due to the large dimensionality of the integration space, this approach may be impractical for online computations. One alternative approach relies on the observation that the matrices in (19) depend on x_t via the difference $x_t - \hat{x}_{t|t}$. Since $x_t - \hat{x}_{t|t}$ is conditionally zero-mean given $\mathcal{Y}_t$, we can write the dependency of the matrices in (19) on the time-varying statistics of x_t given $\mathcal{Y}_t$ as follows: $\Lambda_t^{x\varphi}(\hat{x}_{t|t}, P_{t|t}) = \hat{x}_{t|t}\Lambda_t^{\varphi}(P_{t|t})^T + \Lambda_t^{e\varphi}(P_{t|t})$, $\Lambda_t^{\varphi}(P_{t|t})$, $\Lambda_t^{e\varphi}(P_{t|t})$, $\Lambda_t^{w\varphi}(P_{t|t})$, and $\Lambda_t^{\varphi\varphi}(P_{t|t})$. Therefore, one may apply Algorithm 1. This procedure allows one to move most of the computational burden off-line. Yet it requires being able to compute and store several matrix functions in a parametric form, or some finite approximation. A more appealing alternative is to exploit the convergence properties of $P_{t|t}$. The following result can be inferred, for instance, from [Kamen and Su, 1999, Theorem 5.1].

Proposition 5. Assume that $\Sigma_v > 0$*. If* (C, A) *is detectable and* $(A, \Sigma_w^{1/2})$ *is stabilizable, then the (discrete-time) algebraic Riccati equation in* $P \in \mathbb{R}^{n\times n}$

$$P = A[P - PC^T(CPC^T + \Sigma_v)^{-1}CP]A^T + \Sigma_w \tag{24}$$

has a unique solution $P^* \geqslant 0$*, and the sequence* $P_{t+1|t}$ *defined by* (21) *and* (23) *converges to* P^* *as* t *tends to* ∞*, for any initial condition* $P_0 \geqslant 0$.

Algorithm 1 Parametric SMPC implementation

Require: $\hat{x}_0 = \mathbb{E}[x_0]$ and $P_0 = \mathbb{E}\big[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T\big]$

1: Compute the parametric expressions
$$\Lambda_t^{\varphi}(P_{t|t}), \Lambda_t^{e\varphi}(P_{t|t}), \Lambda_t^{\varphi}(P_{t|t}), \Lambda_t^{w\varphi}(P_{t|t}), \Lambda_t^{\varphi\varphi}(P_{t|t})$$
for arbitrary positive semidefinite matrices $P \in \mathbb{R}^{n\times n}$;
2: Initialize $\hat{x}_{0|-1} = \hat{x}_0$ and $P_{0|-1} = P_0$
3: $t = 0$
4: **loop**
5: Measure y_t
6: Compute $\hat{x}_{t|t}$ and $P_{t|t}$ via (20)–(21)
7: Evaluate $\Lambda_t^{x\varphi}$ using $\hat{x}_{t|t}$
8: Solve the optimization problem in Proposition 2
9: Apply the first input u_t^*
10: Compute $\hat{x}_{t+1|t}$ and $P_{t+1|t}$ using (22)–(23)
11: $t = t + 1$
12: **end loop**

The assumption that $\Sigma_v > 0$ can be relaxed to $\Sigma_v \geqslant 0$ at the price of some additional technicality [Ferrante et al., 2002]. As a consequence of this result, under detectability and stabilizability assumptions, $P_{t|t}$ converges to
$$P^\circ = P^* - P^* C^T (C P^* C^T + \Sigma_v)^{-1} C P^*, \qquad (25)$$
which is the asymptotic error covariance matrix of the estimator $\hat{x}_{t|t}$. Thus, neglecting the initial transient, it makes sense to just apply Algorithm 2. By this proce-

Algorithm 2 Asymptotic SMPC implementation

Require: $\hat{x}_0 = \mathbb{E}[x_0]$

1: Compute P^* in (24) and P° in (25)
2: Compute the time-invariant matrices
$$\Lambda_t^{\varphi}(P^\circ), \Lambda_t^{e\varphi}(P^\circ), \Lambda_t^{\varphi}(P^\circ), \Lambda_t^{w\varphi}(P^\circ), \Lambda_t^{\varphi\varphi}(P^\circ)$$
3: Initialize $\hat{x}_{0|-1} = \hat{x}_0$
4: $t = 0$
5: **loop**
6: Measure y_t
7: Compute $\hat{x}_{t|t}$ via (20)
8: Evaluate $\Lambda_t^{x\varphi}$ using $\hat{x}_{t|t}$
9: Solve the optimization problem in Proposition 2
10: Apply the first input u_t^*
11: Compute $\hat{x}_{t+1|t}$ via (22)
12: $t = t + 1$
13: **end loop**

dure, virtually all the burden associated with computing the required matrices in Proposition 2 is moved off-line. Note that the feasibility of the optimization problem in step 8 of Algorithm 1 depends on $P_{t|t}$, whereas that of step 10 of Algorithm 2 depends on P^*. Under the same assumptions of Proposition 5, upper bounds on $P_{t|t}$ may be determined and exploited to facilitate the feasibility analysis in Proposition 2.

5. EXAMPLE

Let us consider the following system
$$x_{t+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_t + w_t, \qquad y_t = x_t + v_t,$$
with $w_t \sim \mathcal{N}(0, 10I)$, and $v_t \sim \mathcal{N}(0, 0.1I)$. We set the problem parameters to the following: $N = 2$ is the optimization horizon, $\varphi_{\max} = 0.5$ is the saturation limit, $U_{\max} = 2$ is the bound on the control inputs, and $R_k = 0.01$, $Q_k = 2I$ are the weights in the optimization problem.

We simulated the system in the discrete time interval $[0, 100]$ under three control paradigms: the first is using receding horizon control via Algorithm 1, the second is using receding horizon control via Algorithm 2, and the third is using the standard LQG controller and post-saturating the resulting input. Note that we have not utilized the constraints (7) and (8) in this example. Simulations were performed in MATLAB. For Algorithms 1 and 2, the optimization problem at each step was solved using the software package `cvx` [Grant and Boyd, 2000]. The computation of the matrices $\Lambda_t^{\varphi}(P_{t|t}), \Lambda_t^{e\varphi}(P_{t|t}), \Lambda_t^{\varphi}(P_{t|t}), \Lambda_t^{w\varphi}(P_{t|t}), \Lambda_t^{\varphi\varphi}(P_{t|t})$ in Algorithm 1 and in Algorithm 2 (using P° instead for $P_{t|t}$) was done via the classical Monte Carlo integration [Robert and Casella, 2004, Section 3.2] using 10^5 samples. The time-evolution of the standard deviations of the state for 100 simulated runs with random initial conditions and noise outcomes in all of the above three scenarios are depicted in Figure 1(top). This shows a better behavior of the system using either Algorithm 1 or 2 versus the saturated LQG method. The improvement is quantified in Figure 1(bottom), which shows the average cost accumulated by the system over time. The total average cost incurred in the saturated LQG case at time 100 is 696769 units, whereas the same cost computed for Algorithm 1 is 430821.7 units and for Algorithm 2 is 430848.8 units. As such either of our methods provided slightly more than 38% dip in the cost. The save in the cost by using Algorithm 1 instead of Algorithm 2 for this example is only 0.006%.

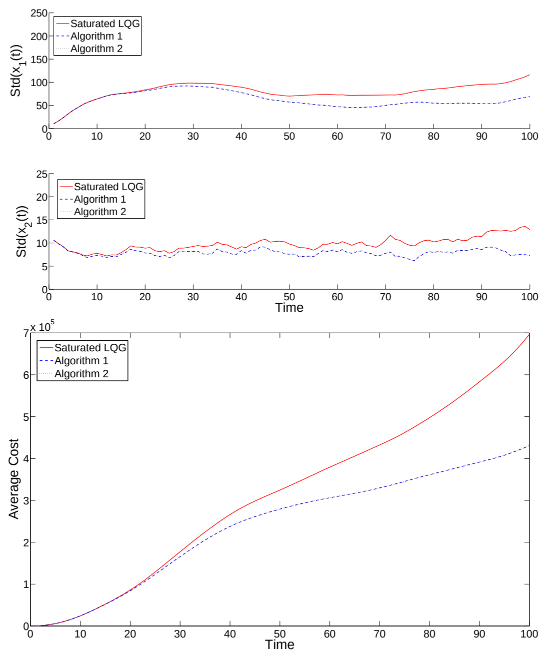

Fig. 1. Standard deviation of the states (top) and the average cost (bottom)

6. CONCLUSIONS

We studied the problem of stochastic MPC with linear dynamics, hard input constraints, and soft constraints on both the state and the control. We demonstrated that using a subclass of causal feedback policies from the so-called purified outputs we obtain a convex underlying optimization problem. In this setup we required that the conditional density of the state given the previous outputs be propagated recursively, and we discussed how this can be done when the random noise is Gaussian.

REFERENCES

M. Agarwal, E. Cinquemani, D. Chatterjee, and J. Lygeros. On convexity of stochastic optimization problems with constraints. In *European Control Conference*, pages 2827–2832, 2009.

I. Batina. *Model predictive control for stochastic systems by randomized algorithms*. PhD thesis, Technische Universiteit Eindhoven, 2004.

A. Bemporad and M. Morari. Robust model predictive control: a survey. *Robustness in Identification and Control*, 245:207–226, 1999.

A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.

A. Ben-Tal, S. Boyd, and A. Nemirovski. Extending scope of robust optimization: Comprehensive robust counterparts of uncertain problems. *Journal of Mathematical Programming*, 107:63–89, 2006.

D. Bernardini and A. Bemporad. Scenario-based model predictive control of stochastic constrained linear systems. In *IEEE Conference on Decision and Control*, Shanghai, China, 63336338 2009.

D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific, 2 edition, 2000.

D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 2. Athena Scientific, 3 edition, 2007.

D. Bertsimas and D. B. Brown. Constrained stochastic LQC: a tractable approach. *IEEE Transactions on Automatic Control*, 52(10):1826–1841, 2007.

D. Bertsimas, D. A. Iancu, and P. A. Parrilo. Optimality of affine policies in multi-stage robust optimization. 2009. preprint available at: http://arxiv.org/abs/0904.3986.

L. Blackmore and B. C. Williams. Optimal, robust predictive control of nonlinear systems under probabilistic uncertainty using particles. In *Proceedings of the American Control Conference*, pages 1759 – 1761, July 2007.

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

M. Cannon, B. Kouvaritakis, and X. Wu. Probabilistic constrained MPC for systems with multiplicative and additive stochastic uncertainty. *IEEE Transactions on Automatic Control*, 54(7):1626–1632, July 2009b.

D. Chatterjee, P. Hokayem, and J. Lygeros. Stochastic receding horizon control with bounded control inputs: a vector space approach. *IEEE Transactions on Automatic Control*, 2009. Under review. `http://arxiv.org/abs/0903.5444`.

P. D. Couchman, M. Cannon, and B. Kouvaritakis. Stochastic MPC with inequality stability constraints. *Automatica J. IFAC*, 42(12):2169–2174, 2006.

A. Ferrante, G. Picci, and S. Pinzoni. Silverman algorithm and the structure of discrete-time stochastic systems. *Linear Algebra and its Applications*, 351/352:219–242, 2002.

P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski. Optimization over state feedback policies for robust control with constraints. *Automatica*, 42(4):523–533, 2006.

M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming (web page and software). `http://stanford.edu/~boyd/cvx`, December 2000.

P. Hokayem, D. Chatterjee, and J. Lygeros. On stochastic model predictive control with bounded control inputs. In *Proceedings of the combined 48th IEEE Conference on Decision & Control and 28th Chinese Control Conference*, pages 6359–6364, 2009. Available at `http://arxiv.org/abs/0902.3944`.

E. W. Kamen and J. K. Su. *Introduction to Optimal Estimation*. Springer, London, UK, 1999.

W. K. Klein Haneveld and M. H. van der Vlerk. Integrated chance constraints: reduced forms and an algorithm. *Computational Management Science*, 3(4):245–269, 2006.

P. R. Kumar and P. Varaiya. *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Prentice Hall, 1986.

J. Löfberg. *Minimax Approaches to Robust Model Predictive Control*. PhD thesis, Linköpings Universitet, 2003.

J. Löfberg. YALMIP : A Toolbox for Modeling and Optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

J. M. Maciejowski. *Predictive Control with Constraints*. Prentice Hall, 2001.

M. Maciejowski, A. Lecchini, and J. Lygeros. NMPC for complex stochastic systems using Markov Chain Monte Carlo. volume 358/2007 of *Lecture Notes in Control and Information Sciences*, pages 269–281, Stuttgart, Germany, 2005. Springer.

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

F. Oldewurtel, C.N. Jones, and M. Morari. A tractable approximation of chance constrained stochastic MPC based on affine disturbance feedback. In *Conference on Decision and Control, CDC*, Cancun, Mexico, Dec 2008.

J. A. Primbs and C. H. Sung. Stochastic receding horizon control of constrained linear systems with state and control multiplicative noise. *IEEE Transactions on Automatic Control*, 54(2):221–230, Feb. 2009.

C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, 2 edition, 2004.

J. Skaf and S. Boyd. Design of affine controllers via convex optimization. `http://www.stanford.edu/~boyd/papers/affine_contr.html`, 2009.

D. H. van Hessem and O. H. Bosgra. A full solution to the constrained stochastic closed-loop MPC problem via state and innovations feedback and its receding horizon implementation. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, volume 1, pages 929–934, 2003.

Y. Wang and S. Boyd. Peformance bounds for linear stochastic control. *Systems and Control Letters*, 58(3): 178–182, 2009.