2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Model Predictive Control Using Segregated Disturbance Feedback

Chen Wang
National University of Singapore
Department of Mechanical Engineering
Email:wangchen@nus.edu.sg

Chong-Jin Ong
National University of Singapore
Singapore-MIT Alliance
Department of Mechanical Engineering
Email:mpeongcj@nus.edu.sg

Melvyn Sim
National University of Singapore
Singapore-MIT Alliance
Business School
Email:dscsimm@nus.edu.sg

***Abstract*— This paper proposes a new control parametrization under Model Predictive Controller (MPC) framework for constrained linear discrete time systems with bounded additive disturbances. The parametrization takes the form of a piecewise affine disturbance feedback and is a generalization of the linear disturbance feedback proposed in the literature. Thus, performance of the resultant MPC controller may be improved. Properties and the numerical computations of the parametrization are discussed. Under mild assumptions of the disturbance set, the associated finite-horizon optimization can be computed efficiently. Stability of the closed-loop system with the proposed parametrization is also ensured.**

I. Introduction

This paper is concerned with the Model Predictive Control (MPC) of

$$x(t+1) = Ax(t) + Bu(t) + w(t), \tag{1}$$

$$(x(t), u(t)) \in Y,\ w(t) \in W,\ \forall\ t \geq 0 \tag{2}$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ are respectively the state and control of the system at time t, $w(t) \in W \subset \mathbb{R}^n$ is the disturbance on the system at time t and Y represents the joint state and control constraint on the system.

The MPC control of such a system is popular and has a wide literature, see [1], [2], [3] and the references cited therein. One aspect of the MPC control that continues to be of research interest is the choice of the control parametrization used in the N-stage finite horizon (FH) optimization problem. It is well known that optimizing over $\{u(0), \cdots u(N-1)\}$ directly results in a conservative system and the optimization should be over families of feedback policies, see [1], [4] and others. One popular feedback policy is $u(t) = Kx(t) + c(t)$ where K is fixed apriori and $c(t)$ is the new optimization variable [2], [4], [5], [6], [7]. Such a policy has a reasonable domain of attraction and good asymptotic behavior. More precisely, the system state converges to

$$F_\infty(K) = W + (A+BK)W + (A+BK)^2 W + \cdots, \tag{3}$$

the minimal invariant set of $x(t+1) = (A+BK)x(t) + w(t)$ [2].

To further enlarge the domain of attraction, other families of feedback policies have been proposed. For example, time-varying state feedback law, $u(t) = K(t)x(t) + c(t)$, where $K(t), c(t)$ changes with time has been attempted. Unfortunately, direct parametrization with affine time-varying state feedback is unappealing as the resulting FH problem is not computationally tractable [8]. Instead, Löfberg [8] and van Hessem & Bosgra [9] proposed the parametrization of $u(t)$ by time-varying disturbance feedback, $u(t) = v(t) + \sum_{i=0}^{t-1} C(i)w(i)$. This parametrization has the advantage that the resulting FH optimization problem is convex and computable. Recently, Goulart et. al. in [3] show the equivalence of time-varying state feedback and time-varying disturbance feedback. Consequently, the MPC systems using either parametrization have the same domain of attraction. They also show that, under mild assumptions, the origin of the closed-loop system is input-to-state stable (ISS) under the MPC control law derived using the time-varying state feedback parametrization. More recently, Wang et.al. in [10] propose a parametrization of the form

$$u(t) = Kx(t) + d(t) + \sum_{i=1}^{N-1} D(t,i)w(t-i) \tag{4}$$

and show that it preserves the same domain of attraction as the time varying disturbance feedback parametrization discussed in [3], [8] but has a stronger stability result in that the system state converges to $F_\infty(K)$.

In an effort to further generalize the parametrization, this paper proposes a control parametrization that covers an even larger family of feedback policies. It uses a time-varying piecewise affine disturbance feedback and is a non-trivial extension of [10]. The proposed parametrization preserves the strong stability results and the associated computations of the FH optimization are reasonable.

The rest of this paper is organized as follows. Notations and general assumptions are given in the rest of this section. Details of the new control parametrization and the MPC framework together with the cost function are given in Section II. Properties and related issues of the segregated disturbance set are discussed in Section III. Convex reformulation and computational issues are introduced in Section IV. Section V discusses the feasibility of the FH optimization problem and stability of the closed-loop system. The last section concludes the paper.

The following notations are used. $\mathbb{Z}_k$ and $\mathbb{Z}_k^+$ denote respectively the integer sets $\{0, 1, \cdots, k\}$ and $\{1, \cdots, k\}$; given matrices $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{p\times q}$ and vector $v \in R^n$: A_i is the i^{th} column of A; v_i is the i^{th} element of v;

978-1-4244-2079-7/08/$25.00 ©2008 AACC.

$A \otimes B$ is the Kronecker product of A and B; $\mathrm{vec}(A) = \begin{bmatrix} A_1^T & \cdots & A_m^T \end{bmatrix}^T \in \mathbb{R}^{nm}$ is the stacked vector of columns of A; $v > (\geq)\ 0$ means $v_i > (\geq)\ 0$ for all i; $|v| := [|v_1| \cdots |v_n|]^T$ is the vector of absolute value of v. A square matrix $A \succ (\succeq) 0$ means A is positive definite (semi-definite). For any $A \succ 0$, $\|x\|_A^2 = x^T Ax$. 1_k is a k-vector with all elements being 1. Given a set Ω, CH(Ω) denotes the convex hull of Ω, $\mathrm{int}(\Omega)$ denotes the interior of Ω and $\delta(y|\Omega) := \max_{\omega \in \Omega} y^T \omega$ is the support function on Ω. Also, the boldface characters are used for collections of vectors or matrices over the length of control horizon.

The system (1)-(2) is assumed to satisfy the following assumptions:

(A1) (A, B) is stabilizable;

(A2) $W \subset \mathbb{R}^n$ is an absolute set;

(A3) the set

$$Y = \{(x, u)|\ Y_x x + Y_u u \leq 1_a\} \subset \mathbb{R}^{n+m} \tag{5}$$

is compact and contains the origin;

(A4) The size of W is sufficiently small such that there exists a constraint-admissible disturbance invariant set

$$X_f = \{x|\ Gx \leq 1_b\} \subset \mathbb{R}^n \tag{6}$$

for system (1) under the control law $u = K_f x$ for some feedback gain $K_f \in \mathbb{R}^{m \times n}$ where $A + BK_f$ is asymptotically stable and that $F_\infty(K_f) \subset \mathrm{int}(X_f)$.

Assumption (A1) is standard. Definition of an absolute set and its implications are discussed in Section III. It will be shown that (A2) is quite general and can be applied to many disturbance models. The characterizations of Y in (A3) is made out of the need for a concrete computational representation. The existence of X_f in (A4) is quite well known under (A1)-(A3) when W is sufficiently small [11], [12]. $F_\infty(K_f)$ is also the set of reachable states under the disturbance input for the system $x(t + 1) = (A + BK)x(t) + w(t), x(0) = 0$. Hence, the last part of (A4) is a mild requirement that W is sufficiently small such that $F_\infty(K_f)$ does not violate the Y constraint.

II. Controller structure and the MPC framework

A. Control parametrization

The proposed control parametrization is a piecewise affine function of w. Let $w \in \mathbb{R}^n$ be segregated into its positive and negative parts by

$$w^p := \max\{w, 0\}, \quad w^m := \max\{-w, 0\} \tag{7}$$

where the max operation is taken component-wise. With this definition, it is easy to see that $w^p, w^m \in \mathbb{R}^n$, $w^p \geq 0$, $w^m \geq 0$ and $w = w^p - w^m$. Correspondingly, the disturbance set for (w^p, w^m) is expanded to

$$\begin{aligned} \Omega_W := \{(w^1, w^2)|\ & w^1 - w^2 \in W, w^1 \geq 0, w^2 \geq 0, \\ & (w^1)^T w^2 = 0\} \subset \mathbb{R}^n \times \mathbb{R}^n \end{aligned} \tag{8}$$

Clearly, there is a one-to-one mapping between $w \in W$ and $(w^1, w^2) \in \Omega_W$: for any $w \in W$, $w^1 = w^p, w^2 = w^m$ while for any $(w^1, w^2) \in \Omega_W$, $w = w^1 - w^2$. The complementarity condition $(w^1)^T w^2 = 0$ in (8) also means that $w_i^1 w_i^2 = 0$ for all $i \in \mathbb{Z}_n^+$ since $w^p \geq 0, w^m \geq 0$. Clearly, this last condition means that Ω_W is non-convex even when W is convex.

Let the control horizon length be N, $x(i), u(i)$ be the i^{th} predicted state and i^{th} predicted control respectively within the horizon at time t. The proposed $u(i)$ takes the form

$$\begin{cases} u(i) = K_f x(i) + c(i), & i \in \mathbb{Z}_{N-1} \\ c(i) = d(i) + \sum_{j=1}^{N-1} C^p(i,j) w^p(i-j) & \\ \qquad + \sum_{j=1}^{N-1} C^m(i,j) w^m(i-j) & \end{cases} \tag{9}$$

where $d(i) \in \mathbb{R}^m$, $C^p(i,j)$, $C^m(i,j) \in \mathbb{R}^{m \times n}$ are the optimization variables, K_f is the specified state feedback gain in (A4) and the disturbances $w^p(i-j)$ and $w^m(i-j)$ are obtained from $w(i-j)$ using (7). Also, the disturbance $w(i)$ is realized if $i < 0$ and is unknown if $i \geq 0$. Hence, $c(i)$ contains the $N-1$ disturbances preceding time $t+i$ and is an affine function of $w^p(i-j)$ and $w^m(i-j)$, $j \in \mathbb{Z}_{N-1}^+$. To simplify notations and presentation, let

$$\begin{aligned} \mathbf{u} &= [u^T(0)\ u^T(1) \cdots\ u^T(N-1)]^T \\ \mathbf{x} &= [x^T(0)\ x^T(1) \cdots\ x^T(N)]^T \\ \mathbf{d} &= [d^T(0)\ d^T(1) \cdots d^T(N-1)]^T \\ \mathbf{w}^- &= [w^T(-(N-1))\ \cdots\ w^T(-1)]^T \\ \mathbf{w}^+ &= [w^T(0)\ \cdots\ w^T(N-1)]^T \end{aligned}$$

where $\mathbf{w}^-$ ($\mathbf{w}^+$) is the collection of realized (future) disturbances at current time. Using (7), $\mathbf{w}^-$ and $\mathbf{w}^+$ can be further separated into their positive and negative parts $\mathbf{w}^{p-}, \mathbf{w}^{m-}, \mathbf{w}^{p+}, \mathbf{w}^{m+}$ and let $\mathbf{\Pi}^- = [(\mathbf{w}^{p-})^T\ (\mathbf{w}^{m-})^T]^T$, $\mathbf{\Pi}^+ = [(\mathbf{w}^{p+})^T\ (\mathbf{w}^{m+})^T]^T$. The rest of the variables in (9) are collected in

$$\mathbf{C}^{p-} = \begin{bmatrix} C^p(0,N-1) & C^p(0,N-2) & \cdots & C^p(0,1) \\ 0 & C^p(1,N-1) & \cdots & C^p(1,2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C^p(N-2,N-1) \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{10}$$

$$\mathbf{C}^{p+} = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ C^p(1,1) & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ C^p(N-2,N-2) & \cdots & 0 & 0 \\ C^p(N-1,N-1) & \cdots & C^p(N-1,1) & 0 \end{bmatrix} \tag{11}$$

$\mathbf{C}^{m-}$ and $\mathbf{C}^{m+}$ where the last two variables are defined in the same way as (10) and (11) with the corresponding change in the superscripts. Using these notations, the control policy of (9) within the control horizon becomes

$$\mathbf{u} = \mathcal{K}\mathbf{x} + \mathbf{d} + \mathbf{C}^-\mathbf{\Pi}^- + \mathbf{C}^+\mathbf{\Pi}^+ \tag{12}$$

where $\mathcal{K} = [I_N \otimes K_f\ \ 0]$, $\mathbf{C}^- = [\mathbf{C}^{p-}\ \mathbf{C}^{m-}]$ and $\mathbf{C}^+ = [\mathbf{C}^{p+}\ \mathbf{C}^{m+}]$. Also, let $\mathbf{C}$ denote $(\mathbf{C}^-,\ \mathbf{C}^+)$.

B. MPC formulation

Using the above-mentioned notations, the FH optimization based on the control parametrization of (9) can be summarized as the following problem $\mathcal{P}_N(\mathbf{d}, \mathbf{C}; x, \mathbf{\Pi}^-)$:

$$\min_{\mathbf{d},\mathbf{C}} \; J_N(\mathbf{d}, \mathbf{C}) \tag{13}$$

$$\text{s.t.} \quad \mathbf{x} = \mathcal{A}x + \mathcal{B}\mathbf{u} + \mathcal{G}\mathbf{\Pi}^+ \tag{14}$$

$$\mathbf{u} = \mathcal{K}\mathbf{x} + \mathbf{d} + \mathbf{C}^-\mathbf{\Pi}^- + \mathbf{C}^+\mathbf{\Pi}^+ \tag{15}$$

$$(x(i), u(i)) \in Y \quad \forall\, \mathbf{\Pi}^+ \in \Omega_W^N, i \in \mathbb{Z}_{N-1} \tag{16}$$

$$x(N) \in X_f \qquad \forall\, \mathbf{\Pi}^+ \in \Omega_W^N \tag{17}$$

where $\mathcal{K} = [I_N \otimes K_f \quad 0]$,

$$\mathcal{A} = \begin{bmatrix} I_n \\ A \\ A^2 \\ \vdots \\ A^N \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ B & 0 & \cdots & 0 \\ AB & B & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix},$$

$$\hat{\mathcal{G}} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ I_n & 0 & \cdots & 0 \\ A & I_n & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1} & A^{N-2} & \cdots & I_n \end{bmatrix}, \; \mathcal{G} = \begin{bmatrix} \hat{\mathcal{G}} & -\hat{\mathcal{G}} \end{bmatrix},$$

Ω_W^N is the N times product space of Ω_W and $J_N(\mathbf{d}, \mathbf{C})$ is an appropriate cost function whose details are discussed in the next subsection. Let the feasible set of the FH optimization problem be

$$\Pi_N(x, \mathbf{\Pi}^-) = \{(\mathbf{d}, \mathbf{C}) \mid \mathcal{P}_N(\mathbf{d}, \mathbf{C}; x, \mathbf{\Pi}^-) \text{ is feasible}\} \tag{18}$$

The set of admissible initial states to the FH problem is then

$$\mathcal{X}_N = \{x | \Pi_N(x, \mathbf{\Pi}^-) \neq \varnothing\}.$$

It appears from (18) that Π_N is a function of x and the past disturbances $\mathbf{\Pi}^-$. The next theorem shows that whether Π_N is empty depends only on x.

Theorem 1: If $\Pi_N(x, \bar{\mathbf{\Pi}}^-) \neq \varnothing$ for some $(x, \bar{\mathbf{\Pi}}^-)$, then $\Pi_N(x, \mathbf{\Pi}^-) \neq \varnothing$ for any $\mathbf{\Pi}^- \in \Omega_W^{N-1}$.

Proof: Choose $(\bar{\mathbf{d}}, \bar{\mathbf{C}}) \in \Pi_N(x, \bar{\mathbf{\Pi}}^-)$ and $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ be the corresponding control and state sequences obtained from (14) and (15). Let

$$\hat{\mathbf{d}} = \bar{\mathbf{d}} + \bar{\mathbf{C}}^-\bar{\mathbf{\Pi}}^-, \; \hat{\mathbf{C}}^- = 0 \text{ and } \; \hat{\mathbf{C}}^+ = \bar{\mathbf{C}}^+.$$

From (15), $(\hat{\mathbf{d}}, \hat{\mathbf{C}})$ define the same feasible control sequence $\bar{\mathbf{u}}$ for any $\mathbf{\Pi}^- \in \Omega_W^{N-1}$ and hence, the same feasible state sequence $\bar{\mathbf{x}}$. This also means that $(\hat{\mathbf{d}}, \hat{\mathbf{C}}) \in \Pi_N(x, \mathbf{\Pi}^-)$ for all $\mathbf{\Pi}^- \in \Omega_W^{N-1}$. □

Following Theorem 1, the feasible set of FH optimization problem can be stated as $\Pi_N(x)$ instead of $\Pi_N(x, \mathbf{\Pi}^-)$. Correspondingly, the admissible initial state set can be defined as

$$\mathcal{X}_N = \{x | \; \Pi_N(x) \neq \varnothing\}. \tag{19}$$

Remark 1: Suppose $\mathcal{P}_N^L$ and $\mathcal{X}_N^L$ are the corresponding FH problem and the admissible set when the control parametrization (4) is used in (15) instead of (9) (For the case where (4) is used, no segregation of $w \in W$ is needed. See [10] for details). It is easy to see that $\mathcal{X}_N^L \subseteq \mathcal{X}_N$ because (4) is a special case of (9).

The rest of the MPC formulation is standard: the FH optimization problem is solved at each time t and the very first term of $(\mathbf{d}^*(t), \mathbf{C}^*(t)) = \arg\min \; \mathcal{P}_N(\mathbf{d}, \mathbf{C}; x(t), \mathbf{\Pi}^-(t))$ is applied to system (1) yielding the MPC control law

$$\begin{aligned} u(t) = \; & K_f x(t) + d^*(0) + \sum_{j=1}^{N-1}(C^{p*}(0, j)w^p(t-j) \\ & + C^{m*}(0, j)w^m(t-j)) \end{aligned} \tag{20}$$

C. cost function

The cost function used in this work is similar with that used in [10] and hence its discussion here is brief. Specifically, the cost function is

$$J_N(\mathbf{d}, \mathbf{C}) := \sum_{i=0}^{N-1} \|\gamma(i)\|_\Psi^2 \tag{21}$$

where

$$\begin{aligned} \gamma(i) = \; & \mathrm{vec}([d(i) \; C^p(i,1) \; C^m(i,1) \; \cdots \\ & C^p(i, N-1) \; C^m(i, N-1)]), \end{aligned}$$

and any $\Psi \succ 0$. Connection of this cost function to the expected standard LQ cost

$$V_N(\mathbf{x}, \mathbf{u}) = \mathrm{E}\left[\sum_{i=0}^{N-1}(\|x(i)\|_Q^2 + \|u(i)\|_R^2) + \|x(N)\|_P^2\right] \tag{22}$$

where the expectation is taken over $(\mathbf{w}^{p-}, \mathbf{w}^{m-}, \mathbf{w}^{p+}, \mathbf{w}^{m+})$, has also been made in [10], [13], [14] under additional assumptions on Q, R, P, K_f and the mean and covariance of disturbance w.

III. Properties of the Ω_W and related sets

The set Ω_W, being non-convex even when W is convex, means that the associated FH computations may be difficult. This difficulty is circumvented when W satisfies (A2). We first review the definition of absolute set.

Definition 1: A set V is an absolute set if it is compact, convex, contains the origin in its interior and $v \in V$ if and only if $|v| \in V$.

From its definition, an absolute set is necessarily symmetric, or $V = \{-v : \; v \in V\}$. Examples of absolute sets include those generated by the L_p norms and their intersections: $\{v : \; \|v\|_p \leq 1\}$, $\{v : \; \|v\|_\infty \leq 1, \|v\|_2 \leq r, \|v\|_1 \leq g\}$. The use of absolute set as disturbance model is also quite common, see [15], [16], [17] and [18].

Remark 2: Assumption (A2) is not as restrictive as it may appear. Many non-symmetrical disturbances or disturbances generated from a set with dimension different from $\mathbb{R}^n$ can be represented as $\{w | w = E\bar{w} + e, \bar{w} \in \bar{W} \subset \mathbb{R}^\ell\}$ where $\bar{W}$ is an absolute set and E and e are some appropriate matrices. For such disturbance model, the exposition hereafter remains valid but with w replaced by $E\bar{w} + e$.

Remark 3: For some class of disturbances where W is convex but cannot be represented by L_p norms, intersections of L_p norms or using Remark 2, the set Ω_W may be approximated by

$$\Omega_W^A = \{(w^1, w^2)|\ w^1 - w^2 \in W, w^1 \geq 0, w^2 \geq 0\}.$$

Suppose $\mathcal{P}_N^A$ and $\mathcal{X}_N^A$ are the corresponding FH problem and the admissible initial set when Ω_W is replaced by Ω_W^A in (16) and (17). It is easy to see that $\Omega_W \subseteq \Omega_W^A$ and Ω_W^A is convex (since W is convex). Hence, $\mathcal{P}_N^A$ is computationally more amiable. Of course, the control law obtained is more conservative since $\Omega_W \subseteq \Omega_W^A$ and, hence, $\mathcal{X}_N^A \subseteq \mathcal{X}_N$.

Remark 4: While more conservative than $\mathcal{P}_N$, $\mathcal{P}_N^A$ is less conservative than $\mathcal{P}_N^L$, the FH problem when parametrization (4) is used. Again, this is true because $\mathcal{P}_N^L$ is a special case of $\mathcal{P}_N^A$ (choose $C^p(i,j)$ and $C^m(i,j)$ in (9) such that $C^p(i,j) = -C^m(i,j) = D(i,j)$). Hence, if a feasible solution exists for $\mathcal{P}_N^L$ for all $w \in W$, a feasible solution exists for $\mathcal{P}_N^A$ for all $(w^1, w^2) \in \Omega_W^A$. This, together with *Remark* 1, means that $\mathcal{X}_N^L \subseteq \mathcal{X}_N^A \subseteq \mathcal{X}_N$.

We now define a set that is closely related to W. Let

$$\Omega_W^B = \{(w^1,\ w^2)|\ w^1 + w^2 \in W, w^1 \geq 0, w^2 \geq 0\} \quad (23)$$

and its connection to Ω_W is given below.

Theorem 2: Suppose W satisfies assumption (A2), then $\Omega_W^B = \mathrm{CH}(\Omega_W)$.

Proof: ($\Rightarrow$)Consider $(v^1, v^2) \in \Omega_W$. It follows that $v^1 \geq 0, v^2 \geq 0$ and $(v^1)^T v^2 = 0$. Therefore, $v^1 + v^2 = |v^1 - v^2|$. Since W is absolute and $v^1 - v^2 \in W$, we have $v^1 + v^2 = |v^1 - v^2| \in W$ which implies that $(v^1, v^2) \in \Omega_W^B$. Since the set Ω_W^B is convex, we have $\mathrm{CH}(\Omega_W) \subseteq \Omega_W^B$. ($\Leftarrow$) To show $\Omega_W^B \subseteq \mathrm{CH}(\Omega_W)$, consider $(u^1, u^2) \in \Omega_W^B$ and let $S^0 = \{(u^1, u^2)\}$. For all $i \in \mathbb{Z}_n^+$, let

$$S^i = \bigcup_{(v^1,v^2) \in S^{i-1}} \{(v^1 - e^i v_i^1, v^2 + e^i v_i^1), (v^1 + e^i v_i^2, v^2 - e^i v_i^2)\}$$

where e^i denotes a unit vector in $\mathbb{R}^n$, with one at the i^{th} element and zeros otherwise. Observe that for all $(v^1, v^2) \in S^i$, $(v^1, v^2) \in \mathrm{CH}(S^{i+1})$. Indeed, if $v_i^1 + v_i^2 > 0$, let $\lambda = v_i^1/(v_i^1 + v_i^2)$ and it follows that

$$(v^1, v^2) = \lambda(v^1 - e^i v_i^1, v^2 + e^i v_i^1) + (1-\lambda)(v^1 + e^i v_i^2, v^2 - e^i v_i^2).$$

Otherwise, if $v_i^1 + v_i^2 = 0$, we have $(v^1, v^2) \in S^{i+1}$. Therefore, by induction, we have $(u^1, u^2) \in \mathrm{CH}(S^n)$. We can also induce that each $(v^1, v^2) \in S^n$ satisfies $v^1, v^2 \geq 0$, $v^1 + v^2 = u^1 + u^2$ and $v_j^1 v_j^2 = 0,\ j \in \mathbb{Z}_n^+$. Hence, $|v^1 - v^2| = v^1 + v^2 = u^1 + u^2 \in W$. Since W is an absolute set, we have $v^1 - v^2 \in W$ and $(v^1, v^2) \in \Omega_W$. Therefore, $(u^1, u^2) \in \mathrm{CH}(\Omega_W)$. □

It can be proved that all absolute set can be expressed in the form of

$$V = \{v\ :\ \eta(v) \leq 1\}, \quad (24)$$

for some absolute norm function $\eta : \mathbb{R}^n \mapsto \mathbb{R}$. By absolute norm, $\eta(\cdot)$ satisfies the three standard properties of a norm and the additional property of $\eta(v) = \eta(|v|)$. Clearly, all polynomial norms or L_p norms are absolute. However, a polynomial norm induced by an invertible matrix, is not necessary absolute. It is easy to see that the following composite norm function

$$\zeta(v) = \max_{l=1,\ldots,L} \{a_l \eta_l(v)\}, \quad (25)$$

in which $\eta_l(\cdot)$ are absolute norms with $a_l > 0$ for all $l \in \mathbb{Z}_L^+$, is absolute. Hence, for instance, $\{v\ :\ \|v\|_\infty \leq 1, \|v\|_2 \leq r\}$ can be expressed in the form of (24) with $\eta(v) = \max\{\frac{1}{r}\|v\|_2, \|v\|_\infty\}$. Given a vector norm $\eta(\cdot)$, the dual norm $\eta^*(\cdot)$ is a norm function defined as

$$\eta^*(y) = \max_{\eta(v) \leq 1} y^T v. \quad (26)$$

IV. Convex Reformulation and Computation

The role of the absolute set in the satisfaction of constraints (16) and (17) is made clear in the following theorem.

Theorem 3: Let $W = \{w\ :\ \eta(w) \leq 1\} \subset \mathbb{R}^n$ be an absolute set for some absolute norm function $\eta(\cdot)$, $\eta^*(\cdot)$ be the corresponding dual norm and Ω_W^B be as defined by (23). The two sets

$$\mathcal{C}_1 = \{(x,y,z) \in \mathbb{R}^{2n+1} | x^T w^1 + y^T w^2 \leq z,\ \forall (w^1, w^2) \in \Omega_W^B\}$$
$$\mathcal{C}_2 = \{(x,y,z) \in \mathbb{R}^{2n+1} | \eta^*(t) \leq z,\ t \geq x,\ t \geq y \text{ for some } t\}$$

are equivalent.

Proof: ($\Rightarrow$) Let (x,y,z) be an element of $\mathcal{C}_1$. It follows from (23) that

$$\begin{aligned}
z \geq\ & \max\{x^T w^1 + y^T w^2 | w^1 \geq 0, w^2 \geq 0, \eta(w^1 + w^2) \leq 1\} \\
=\ & \max\{x^T w^1 + y^T w^2 |\ w^1 \geq 0, w^2 \geq 0, \\
& \quad w = w^1 + w^2,\ \eta(w) \leq 1\} \\
=\ & \max\{\bar{t}^T w |\ w \geq 0,\ \eta(w) \leq 1,\ \bar{t}_i = max\{x_i, y_i\}\} && (27) \\
=\ & \max\{\bar{t}^T |w|\ |\ \eta(w) \leq 1,\ \bar{t}_i = max\{x_i, y_i\}\} && (28) \\
=\ & \max\{t^T |w|\ |\ \eta(w) \leq 1,\ t_i = max\{0, \bar{t}_i\}\} && (29) \\
=\ & \max\{t^T w |\ \eta(w) \leq 1,\ t_i = max\{0, \bar{t}_i\}\} && (30) \\
\Rightarrow\ & (x,y,z) \in \mathcal{C}_2
\end{aligned}$$

The first two relations come from the definitions of Ω_W^B, W and the re-organization of the constraints. Equation (27) comes from the fact that the optimal value can be achieved by considering w^1 and w^2 where $w_i^1 w_i^2 = 0$ for all i. This is true because the optimal w^* is such that $w_i^* = w_i^{1*}$ if $x_i > y_i$ and $w_i^* = w_i^{2*}$ if $x_i \leq y_i$ for all i. Equation (28) follows because W is an absolute set. Equation (29) comes from the fact that if $t_i < 0$, the optimal w_i^* must be 0. Hence, the maximum value can be obtained by letting $t_i = max\{0, \bar{t}_i\}$. Since $t \geq 0$, the absolute sign on w can be relaxed based on the fact that dual norm of absolute norm is also absolute. The last implication follows since the existence of $t, t \geq x$ and $t \geq y$ is established.

($\Leftarrow$) Let (x,y,z) be an element of $\mathcal{C}_2$ with a suitable $t \in \mathbb{R}^n$. Then, from the definition of $\eta^*(\cdot)$,

$$\begin{aligned}
z\ \geq\ & \max\{t^T(w^1 + w^2)|\ (w^1 + w^2) \in W, t \geq x, t \geq y\} \\
\geq\ & \max\{t^T(w^1 + w^2)|\ (w^1 + w^2) \in W, w^1 \geq 0, w^2 \geq 0, \\
& \quad t \geq x, t \geq y\} \\
\geq\ & \max\{x^T w^1 + y^T w^2)|\ (w^1 + w^2) \in W, w^1 \geq 0, w^2 \geq 0\} \\
=\ & \max\{x^T w^1 + y^T w^2)|\ (w^1, w^2) \in \Omega_W^B\} \\
\Rightarrow\ & (x,y,z) \in \mathcal{C}_1.
\end{aligned}$$

Again, the first inequality holds from definition. The second inequality follows from the imposition of two additional constraints $w^1 \geq 0, w^2 \geq 0$. The third inequality follows from the fact that $t^T w \geq x^T w$ and $t^T w \geq y^T w$ for all $w \geq 0$ since $t \geq x$ and $t \geq y$. The last equality is from the definition of Ω_W^B which implies the inclusion. □

Remark 5: The adaptation of Theorem 3 to disturbance set defined by intersection of L_p norm sets is quite easy. For example, if $W = \{w| \ \|w\|_\infty \leq 1, \ \|w\|_2 \leq r\}$, then $\eta(v) = \max\{\frac{1}{r}\|v\|_2, \|v\|_\infty\}$, $\eta^*(v) = \min\{r\|v^1\|_2 + \|v^2\|_1, v^1 + v^2 = v\}$ and the deterministic equivalence of $\mathcal{C}_1$ in Theorem 3 is $\mathcal{C}_2 = \{(x,y,z)| \ \exists t, t^1, t^2 \in \mathbb{R}^n, \|t^2\|_1 + r\|t^1\|_2 \leq z, t^1 + t^2 = t, \ t \geq x, \ t \geq y\}$.

Remark 6: Observe that the constraint $\eta^*(t) \leq z$ arising in $\mathcal{C}_2$ in Theorem 3 is equivalent to

$$w^T t \leq z \qquad \forall w \in W$$

whose tractability and explicit formulation can be found in [16]. In particular, if the set W is conic quadratic representable, which includes sets prescribed by intersections of l_p norms, p being a rational number, the resulting robust counterpart is also conic quadratic representable. For the representation power of conic quadratic constraints we refer the interested reader to [19]. Software involving conic quadratic representable constraints includes SDPT3 (http://www.math.nus.edu.sg/ mattohkc/sdpt3.html) and MOSEK (http://www.mosek.com/).

Using characterizations of Y and X_f in (5) and (6), constraints (14)-(17) can be restated as

$$\begin{aligned}&\bar{\mathcal{A}}x + \bar{\mathcal{B}}\mathbf{d} + \bar{\mathcal{F}}\mathrm{vec}(\left[\mathbf{C}^{p-}\mathbf{C}^{m-}\right]) \\ &\quad + \max_{\mathbf{\Pi}^+ \in \Omega_W^N} \left[\bar{\mathcal{B}}\left[\mathbf{C}^{p+}\mathbf{C}^{m+}\right] + \left[\bar{\mathcal{G}} \ \ -\bar{\mathcal{G}}\right]\right]\mathbf{\Pi}^+ \leq 1_s \end{aligned} \quad (31)$$

where $s = aN + b$ and expressions of $\bar{\mathcal{A}}, \bar{\mathcal{B}}, \bar{\mathcal{G}}, \bar{\mathcal{F}}$ are given in Appendix. Since W is an absolute set, (31) can be restated as

$$\begin{aligned}&\bar{\mathcal{A}}x + \bar{\mathcal{B}}\mathbf{d} + \bar{\mathcal{F}}\mathrm{vec}(\left[\mathbf{C}^{p-}\mathbf{C}^{m-}\right]) \\ &\quad + \max_{\mathbf{\Pi}^+ \in (\Omega_W^B)^N} \left[\bar{\mathcal{B}}\left[\mathbf{C}^{p+}\mathbf{C}^{m+}\right] + \left[\bar{\mathcal{G}} \ \ -\bar{\mathcal{G}}\right]\right]\mathbf{\Pi}^+ \leq 1_s \end{aligned} \quad (32)$$

following result of Theorem 2 and known property that $\delta(y|C) = \delta(y|\mathrm{CH}(C))$. In addition, an absolute norm function $\eta_{\mathbf{w}}(\cdot)$ can be found such that $W^N = \{\mathbf{w}^+| \ \eta_{\mathbf{w}}(\mathbf{w}^+) \leq 1\}$. Applying the result of Theorem 3, (32) has the following deterministic equivalence

$$\begin{cases}\bar{\mathcal{A}}x + \bar{\mathcal{B}}\mathbf{d} + \bar{\mathcal{F}}\mathrm{vec}([\mathbf{C}^{p-}\mathbf{C}^{m-}]) + \mu \leq \mathbf{1}_s \\ Z^T \geq \bar{\mathcal{B}}\mathbf{C}^{p+} + \bar{\mathcal{G}} \\ Z^T \geq \bar{\mathcal{B}}\mathbf{C}^{m+} - \bar{\mathcal{G}} \\ \mu = [\eta_{\mathbf{w}}^*(Z_1) \ \cdots \ \eta_{\mathbf{w}}^*(Z_s)]^T \end{cases} \quad (33)$$

The numerical computations of $\mathcal{P}_N$ with the above constraints can be achieved for the various norm functions following *Remark* 6.

V. Feasibility and Stability

The next theorem shows the the feasibility of the FH optimization problem and the stability of the closed-loop system under the closed-loop control law (20).

Theorem 4: Suppose $x(0) \in \mathcal{X}_N$ and assumptions (A1-A4) are satisfied. The closed-loop system using the MPC control law (20) has the following properties: (i) $\mathcal{P}_N(\mathbf{d}, \mathbf{C}, x(t), \mathbf{\Pi}^-(t))$ is feasible for all $t \geq 0$; (ii) $(x(t), u(t)) \in Y$ for all $t \geq 0$; (iii) $x(t) \to F_\infty$ as $t \to \infty$; (iv) There exists a finite $\tilde{t}$ such that $x(t) \in X_f$ and $c(t) = 0$ for all $t \geq \tilde{t}$.

Proof: The proof follows essentially the arguments in [10] and uses the notation "$|t$" for the variables at time instant t. (i) Feasibility of $\mathcal{P}_N(\mathbf{d}, \mathbf{C}, x(t), \mathbf{\Pi}^-(t))$ follows standard arguments. If $(\mathbf{d}^*, \mathbf{C}^*)$ is the optimal control at time t, choose the feasible control at time $t+1$ by

$$\begin{cases} d(i|t+1) = d^*(i+1|t), \ i \in \mathbb{Z}_{N-2}, \\ C^k(i,j|t+1) = C^{k*}(i+1,j|t), i \in \mathbb{Z}_{N-2}, k \in \{p, m\} \\ d(N-1|t+1) = 0, C^k(N-1,j|t+1) = 0, k \in \{p,m\} \end{cases} \quad (34)$$

(ii) The result follows directly from (i).

(iii) If $J_N^*(t)$ is the optimal value of $\mathcal{P}_N(\mathbf{d}, \mathbf{C}, x(t), \mathbf{\Pi}^-(t))$ and let $\hat{J}_N(t+1)$ be the value of $J_N(\mathbf{d}, \mathbf{C})$ where $(\mathbf{d}, \mathbf{C})$ are defined by (34), then it can be verified that

$$J_N^*(t) - J_N^*(t+1) \geq J_N^*(t) - \hat{J}_N(t+1) = \|\gamma(0|t)\|_\Psi^2$$

where $\|\gamma(\cdot)\|_\Psi^2$ is the norm function used in (21). Hence, $\{J_N^*(t)\}$ is a monotonically non-increasing sequence bounded from below and it tends to a limit as $t \to \infty$. This necessary means that $\gamma(0|t)$ tends to zero as $t \to \infty$. Hence, $c(t)$ tends to zero as $t \to \infty$. The system state under (20) can be written as

$$x(t) = \Phi^t x(0) + \sum_{i=0}^{t-1} \Phi^{t-1-i} Bc(i) + \sum_{i=0}^{t-1} \Phi^{t-1-i} w(i) \quad (35)$$

where $\Phi = A + BK_f$. The first term on the right of (35) approaches zero as $t \to \infty$ because of (A4). The second term approaches zero following the fact that $c(t) \to 0$ as $t \to \infty$. The last term corresponds to a point in the set $F_t(K_f) := W + \cdots + \Phi^{t-1}W$, which approaches $F_\infty(K_f)$ as $t \to \infty$ under assumption (A4). Hence the stated result follows.

(iv) Following property (iii) and assumption (A4), $x(t)$ enters X_f in some finite time, $\tilde{t}$, and thereafter the optimal $(\mathbf{d}, \mathbf{C}) = 0$ for all $t \geq \tilde{t}$ from the optimality of $\mathcal{P}_N(\mathbf{d}, \mathbf{C}, x(t), \mathbf{\Pi}^-(t))$. □

VI. Conclusion

A piecewise linear disturbance feedback parametrization is proposed for MPC of constrained linear systems with disturbances. This parametrization includes disturbance feedback law as a special case, as a consequence, better performance could be expected. Although the resulting FH optimization problem is not directly computable, its equivalent convex reformulation can be found if the disturbance set is absolute. Even if the disturbance set is not absolute, the new parametrization can still result at a MPC controller which is less conservative than the one derived using linear disturbance feedback. Also, the closed-loop system has a clearly-characterized asymptotic behavior if certain cost is minimized.

REFERENCES

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

[2] L. Chisci, J. A. Rossiter, and G. Zappa, "Systems with persistent disturbances: predictive control with restricted constraints," *Automatica*, vol. 37, no. 7, pp. 1019–1028, 2001.

[3] P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski, "Optimization over state feedback policies for robust control with constraints," *Automatica*, vol. 42, no. 4, pp. 523–533, 2006.

[4] A. Bemporad, "Reducing conservativeness in predictive control of constrained systems with disturbances," in *Proceedings of 37th Conference on Decision and Control*, (Tampa, Florida), pp. 1384–1389, 1998.

[5] J. A. Rossiter, B. Kouvaritakis, and M. J. Rice, "A numerically robust state-space approach to stable predictive control strategies," *Automatica*, vol. 34, no. 1, pp. 65–73, 1998.

[6] Y. I. Lee and B. Kouvaritakis, "Constrained receding horizon predictive control for systems with disturbances," *International Journal of Control*, vol. 72, no. 11, pp. 1027–1032, 1999.

[7] D. Q. Mayne, M. M. Seron, and S. V. Raković, "Robust model predictive control of constrained linear systems with bounded disturbances," *Automatica*, vol. 41, no. 2, pp. 219–224, 2005.

[8] J. Löfberg, "Approximations of closed-loop minimax mpc," in *Proceedings of the 42nd IEEE Conference on Decision and Control*, (Maui, Hawaii, USA), pp. 1438–1442, 2003.

[9] D. H. van Hessem and O. H. Bosgra, "A conic reformulation fo model predictive control including bounded and stochastic disturbances under state and input constraints," in *Proceedings of the 41st IEEE Conference on Decision and Control*, (Las Vegas, Nevada, USA), pp. 4643–4648, 2002.

[10] C. Wang, C. J. Ong, and M. Sim, "Model predictive control using affine disturbance feedback for constrained linear system," in *Proceedings of the 46th IEEE Conference on Decision and Control*, (New Orleans, Louisiana, USA), pp. 1275–1280, 2007.

[11] I. Kolmanovsky and E. G. Gilbert, "Maximal output admissible sets for discrete-time systems with disturbance inputs," in *Proceedings of the 1995 American Control Conference*, (Seattle), pp. 1995–1999, 1995.

[12] I. Kolmanovsky and E. G. Gilbert, "Theory and computation of disturbance invariant sets for discrete-time linear systems," *Mathematical Problems in Engineering*, vol. 4, no. 4, pp. 317–367, 1998.

[13] P. J. Goulart and E. C. Kerrigan, "On the stability of a class of robust receding horizon control laws for constrained systems," tech. rep., Department of Engineering, University of Cambridge, August 2005. CUED/F-INFENG/TR.532.

[14] P. J. Goulart and E. C. Kerrigan, "Robust receding horizon control with an expected value cost," in *Proc. UKACC International Conference (Control 2006)*, (Glasgow, Scotland), August 2006.

[15] A. Soyster, "Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming," *Operations Research*, vol. 21, no. 5, pp. 1154–1157, 1973.

[16] A. Ben-Tal and A. Nemirovski, "Robust convex optimization," *Mathematics of Operations Research*, vol. 23, no. 4, pp. 769–805, 1998.

[17] A. Ben-Tal and A. Nemirovski, "Robust solutions to uncertain linear programs," *OR Letters*, vol. 25, 1999.

[18] D. Bertsimas and M. Sim, "The Price of Robustness," *Operations Research*, vol. 52, no. 1, pp. 35–53, 2004.

[19] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, Engineering Applications*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.

APPENDIX

Let $\varphi = (I - \mathcal{B}\mathcal{K})^{-1}$, $\bar{Y}_x = I_N \otimes Y_x$, $\bar{Y}_u = I_N \otimes Y_u$, $\mathcal{Y} = \begin{bmatrix} \bar{Y}_x & 0 & \bar{Y}_u \\ 0 & G & 0 \end{bmatrix}$, then the matrices appearing in (31) are $\bar{\mathcal{A}} = \mathcal{Y}\left[(\varphi\mathcal{A})^T \ (\mathcal{K}\varphi\mathcal{A})^T\right]^T$, $\bar{\mathcal{B}} = \mathcal{Y}\left[(\varphi\mathcal{B})^T \ (I + \mathcal{K}\varphi\mathcal{B})^T\right]^T$, $\bar{\mathcal{G}} = \mathcal{Y}\left[(\varphi\hat{\mathcal{G}})^T \ (\mathcal{K}\varphi\hat{\mathcal{G}})^T\right]^T$ and $\bar{\mathcal{F}} = \mathcal{Y}((\mathbf{\Pi}^-)^T \otimes \left[(\varphi\mathcal{B})^T \ (I + \mathcal{K}\varphi\mathcal{B})^T\right]^T)$