Model Predictive Control Using Segregated Disturbance Feedback

Chen Wang, Chong-Jin Ong and Melvyn Sim

Abstract

This paper proposes a new control parametrization under the Model Predictive Control framework for constrained linear discrete-time systems with bounded disturbances. In an effort to minimize conservatism, the proposed parametrization takes the form of a piecewise affine disturbance feedback and is a generalization of linear disturbance feedback that has appeared in recent literature. Numerical computations and stability properties of the resulting MPC problem using the proposed parametrization are discussed. When the disturbance belongs to an absolute set and the problem data satisfy reasonable assumptions, the associated finite-horizon optimization can be computed efficiently and exactly. The generality of absolute set for disturbance modelling is also discussed. The advantage of the proposed parametrization over linear disturbance feedback is illustrated via numerical examples.

I. Introduction

This paper is concerned with the Model Predictive Control (MPC) of systems given by:

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

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

Chen Wang is with Department of Mechanical Engineering, National University of Singapore, Singapore (e-mali: wangchen@nus.edu.sg).

Chong-Jin Ong is with Department of Mechanical Engineering, National University of Singapore and Singapore-MIT Alliance, Singapore (e-mail: mpeongcj@nus.edu.sg).

Melvyn Sim is with Business School, National University of Singapore and Singapore-MIT Alliance, Singapore (e-mail: dscsimm@nus.edu.sg).

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m, w(t) \in W \subset \mathbb{R}^N$ are respectively the state, control and disturbance of the system at time t and Y represents the joint pointwise-in-time state and control constraint. 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 $u(i) = Kx(i) + c(i)$ where K is fixed apriori and $c(i)$ is the new optimization variable [2], [4], [5], [6], [7]. Such a policy has a reasonable domain of attraction and good asymptotic behavior. In addition, the closed-loop 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 reduce conservatism, other families of feedback policies have been proposed. For example, time-varying state feedback policy, $u(i) = K(i)x(i) + c(i)$, where $K(i), c(i)$ changes with time i has been attempted. Unfortunately, direct parametrization using such a policy is unappealing as the resulting FH problem is not computationally tractable [8]. Löfberg [8] and van Hessem & Bosgra [9] proposed the parametrization of $u(i)$ by time-varying disturbance feedback,

$$u(i) = v(i) + \sum_{j=0}^{i-1} C(j)w(j). \tag{4}$$

This parametrization has the advantage that the resulting FH optimization problem is convex and computable. Recently, Goulart et. al. [3] show the equivalence of time-varying state feedback and time-varying disturbance feedback in terms of their representative abilities. 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

$$u(i) = Kx(i) + d(i) + \sum_{j=1}^{N-1} D(i,j)w(i-j) \tag{5}$$

and show that it has the same domain of attraction as that obtained using parametrization (4) but has a stronger stability result in that the closed-loop system state converges to $F_\infty(K)$.

In an effort to further generalize the control, this paper proposes a parametrization that covers an even larger family of feedback policies. It uses a time-varying piecewise affine disturbance feedback parametrization which includes (4) and (5) as special cases and is a non-trivial extension of [10]. In its most general settings, the corresponding FH optimization may not be easily computed and approximations are needed. When W is an absolute set (to be defined) and the problem data satisfy reasonable assumptions, the corresponding FH optimization is convex and the conversion of it into a numerically amicable deterministic equivalence is exact. The convergence of the closed-loop system state to $F_\infty(K)$ under the proposed parametrization is also preserved.

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 disturbance set are discussed in Section III. Convex reformulation and computational issues are introduced in Section IV. Section V discuss the feasibility of the FH optimization problem and stability of the closed-loop system. Numerical examples and conclusions are the contents of the last two sections.

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 ith column of A; v_i is the ith element of v; $A \otimes B$ 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; and $|v| := [|v_1| \cdots |v_n|]^T$ is the vector of absolute value of v. A square matrix $A \succ (\succeq) 0$ means A is positive definite (semi-definite). For any $A \succ 0$, $\|x\|_A^2 = x^T A x$. 1_k is a k-vector with all elements being 1. Given a non-empty set Ω, CH(Ω) denotes the convex hull of Ω and $\mathrm{int}(\Omega)$ denotes the interior of Ω. Also, 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 constraint set

$$Y = \{(x, u)|\ Y_x x + Y_u u \leq 1_a\} \subset \mathbb{R}^{n+m} \tag{6}$$

is compact and contains the origin;

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

$$X_f = \{x|\ Gx \leq 1_b\} \subset \mathbb{R}^n \tag{7}$$

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 used to model many families of disturbances. 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 $x(t+1) = (A + BK_f)x(t) + w(t), x(0) = 0$ [12]. Hence, $F_\infty(K_f) \subset \mathrm{int}(X_f)$ is a reasonable requirement on the disturbance so as not to 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. To be precise, let $w \in \mathbb{R}^n$ be segregated into its positive and negative parts via

$$w^p := \max\{w, 0\}, \quad w^m := \max\{-w, 0\} \tag{8}$$

where the $\max$ operation is taken component-wise. With this definition, it follows 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

$$\Omega_W := \{(w^1, w^2)|\ w^1 - w^2 \in W, w^1 \geq 0, w^2 \geq 0,\ (w^1)^T w^2 = 0\} \subset \mathbb{R}^n \times \mathbb{R}^n \tag{9}$$

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

$$\begin{cases} u(i) = K_f x(i) + c(i), & i \in \mathbb{Z}_{N-1} \\ c(i) = d(i) + \sum_{j=1}^{N-1} C^p(i,j) w^p(i-j) + \sum_{j=1}^{N-1} C^m(i,j) w^m(i-j) \end{cases} \tag{10}$$

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 $w^p(i-j)$ and $w^m(i-j)$ are obtained from $w(i-j)$ using (8). Also, 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 $w^m(i-j)$, $j \in \mathbb{Z}_{N-1}^+$.

Since $w = w^p - w^m$, (10) is a piecewise function of w and, because of the particular choice of the pieces, is termed segregated disturbance feedback. Clearly, it is a generalization of linear disturbance feedback law of (5) (choose $C^p(i,j)$ and $C^m(i,j)$ in (10) to be $C^p(i,j) = -C^m(i,j) = D(i,j)$) or those used in [8], [3] and this is shown in the following example.

Example 1: Consider the scalar system $x(t+1) = w(t) + 0.5w(t-1) + u(t)$ with $Y = \{(x,u)|\ x \geq -1.1,\ u \geq -0.2\}$, $w(t)$ is independent and identically-distributed with a uniform distribution over $W = \{w|\ |w| \leq 1\}$. Suppose it is required that $x(t)$ to have zero mean. Then, it can be shown that no linear disturbance feedback law in the form (5) can simultaneously satisfy the constraints and the requirement. However, setting $u(t) = -0.1 + 0.5w^m(t-1) - 0.1w^p(t-1)$ does.

To simplify notations and presentation, let

$$\mathbf{u} = [u^T(0)\ u^T(1) \cdots\ u^T(N-1)]^T$$

$$\mathbf{x} = [x^T(0)\ x^T(1) \cdots\ x^T(N)]^T$$

$$\mathbf{d} = [d^T(0)\ d^T(1) \cdots d^T(N-1)]^T$$

$$\mathbf{w}^- = [w^T(-(N-1))\ \cdots\ w^T(-1)]^T$$

$$\mathbf{w}^+ = [w^T(0)\ \cdots\ w^T(N-1)]^T$$

where $\mathbf{w}^-$ ($\mathbf{w}^+$) is the collection of realized (future) disturbances at current time. Using (8), $\mathbf{w}^-$ and $\mathbf{w}^+$ can be further separated into their positive and negative parts $\mathbf{w}^{p-}, \mathbf{w}^{m-}, \mathbf{w}^{p+}, \mathbf{w}^{m+}$. The rest of the variables in (10) are collected in

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

$$\mathbf{C}^{p+} = \begin{bmatrix} 0 & \cdots & 0 & 0 \\ C^p(1,1) & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ C^p(N-2,N-2) & \cdots & 0 & 0 \\ C^p(N-1,N-1) & \cdots & C^p(N-1,1) & 0 \end{bmatrix}, \tag{12}$$

$\mathbf{C}^{m-}$ and $\mathbf{C}^{m+}$ where the last two variables are defined in the same way as (11) and (12) with the corresponding changes in the superscripts. Using these notations, the control policy of (10) within the control horizon becomes

$$\mathbf{u} = \mathcal{K}\mathbf{x} + \mathbf{d} + \mathbf{C}^-\mathbf{\Pi}^- + \mathbf{C}^+\mathbf{\Pi}^+ \tag{13}$$

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

B. MPC formulation

Using the preceding notations, the FH optimization based on the control parametrization of (10), $\mathcal{P}_N(\mathbf{d}, \mathbf{C}; x, \mathbf{\Pi}^-)$, can be summarized as the following problem:

$$\min_{\mathbf{d},\mathbf{C}} \; J(\mathbf{d}, \mathbf{C}) \tag{14}$$

$$\text{s.t.} \;\; \mathbf{x} = \mathcal{A}x + \mathcal{B}\mathbf{u} + \mathcal{G}\mathbf{\Pi}^+ \tag{15}$$

$$\mathbf{u} = \mathcal{K}\mathbf{x} + \mathbf{d} + \mathbf{C}^-\mathbf{\Pi}^- + \mathbf{C}^+\mathbf{\Pi}^+ \tag{16}$$

$$(x(i), u(i)) \in Y \quad \forall\, \mathbf{\Pi}^+ \in \Omega_W^N, i \in \mathbb{Z}_{N-1} \tag{17}$$

$$x(N) \in X_f \qquad \forall\, \mathbf{\Pi}^+ \in \Omega_W^N \tag{18}$$

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

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

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

$$\Pi_N(x, \mathbf{\Pi}^-) = \{(\mathbf{d}, \mathbf{C}) \mid \mathcal{P}_N(\mathbf{d}, \mathbf{C}; x, \mathbf{\Pi}^-) \text{ is feasible}\} \tag{19}$$

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

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

It appears from (19) that Π_N is a function of x and the past disturbances $\mathbf{\Pi}^-$. The next theorem shows the non-emptiness of Π_N 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 (15) and (16). 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 (16), $(\hat{\mathbf{d}}, \hat{\mathbf{C}})$ 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

$$\mathcal{X}_N = \{x |\ \Pi_N(x) \neq \varnothing\}. \tag{20}$$

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 (5) is used in (16) instead of (10). Since (5) uses affine disturbance feedback and is a special case of (10), it follows that $\mathcal{X}_N^L \subseteq \mathcal{X}_N$.

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

$$u(t) = K_f x(t) + d^*(0) + \sum_{j=1}^{N-1} (C^{p*}(0,j) w^p(t-j) + C^{m*}(0,j) w^m(t-j)), \tag{21}$$

where $w^p(t-j), w^m(t-j), j = 1, \cdots, N-1$ are past disturbances that can be obtained from (1) and (8).

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

$$J(\mathbf{d}, \mathbf{C}) := \sum_{i=0}^{N-1} J_i := \sum_{i=0}^{N-1} \|\gamma(i)\|_\Lambda^2 \tag{22}$$

where $\Lambda \succ 0$ and $\gamma(i) := \mathrm{vec}([d(i)\ C^m(i,1)\ C^p(i,1)\ \cdots\ C^m(i,N-1)\ C^p(i,N-1)])$.

Remark 2: When K_f and Λ are appropriately chosen, cost function (22) can be related to expected value of standard LQ cost under the zero-mean assumption of $w(t)$. Specifically, given $Q \succeq 0$, $R \succ 0$, let $P = A^T P A - A^T P B (R + B^T P B)^{-1} B^T P A + Q \succ 0$ be the solution of algebraic Riccati equation and $K_f = -(R + B^T P B)^{-1} B^T P A$. If $w(t)$ is independent and

identically-distributed having zero mean and covariance Σ_w, it follows from Theorem 11.2 of [13] that

$$\begin{aligned} &\mathrm{E}\left[\sum_{i=0}^{N-1}(\|x(i)\|_Q^2+\|u(i)\|_R^2)+\|x(N)\|_P^2\right] \\ &= x(0)^T P x(0) + N\mathrm{trace}(\Sigma_w P) + \mathrm{E}\left[\sum_{i=0}^{N-1}\|c(i)\|_{R+B^TPB}^2\right]. \end{aligned} \tag{23}$$

The last term of (23) satisfies $\mathrm{E}\left[\sum_{i=0}^{N-1}\|c(i)\|_{R+B^TPB}^2\right] = \sum_{i=0}^{N-1}\|\gamma(i)\|_\Lambda^2$ if

$$\Lambda = \begin{bmatrix} 1 & \mathbf{1}_{N-1}^T \otimes (\bar{w}^{mp})^T \\ \mathbf{1}_{N-1}\otimes \bar{w}^{mp} & I_{N-1}\otimes(\bar{w}^{mp}(\bar{w}^{mp})^T + \Sigma_w^{mp}) \end{bmatrix} \otimes (R+B^TPB) \tag{24}$$

where $\bar{w}^{mp}$ and Σ_w^{mp} are the mean and the covariance matrix of random vector $[w^m(t); w^p(t)]$ respectively. Since the first two terms on the right-hand-side of (23) are independent of variables $(\mathbf{d}, \mathbf{C})$, minimizing (22) is equivalent to minimizing the expected value of the LQ cost of (23).

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 optimization problem may be difficult to compute. This difficulty is circumvented when W satisfies (A2). We first review the definition of an 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: $\{v : \ \|v\|_p \le 1\}$, $\{v : \ \|v\|_\infty \le 1, \|v\|_2 \le r, \|v\|_1 \le g\}$. The use of absolute set as disturbance model is also quite common, see [14], [15], [16] and [17].

Remark 3: 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 models, the exposition hereafter remains valid but with w replaced by $E\bar{w} + e$.

Remark 4: 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 3, the set Ω_W may be approximated by

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

Compared with (9), it is easy to see that $\Omega_W \subseteq \Omega_W^A$ and Ω_W^A is convex (since W is convex). 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 (17) and (18). Then, $\mathcal{P}_N^A$ is computationally more amiable as Ω_W^A is convex. Also, $\mathcal{X}_N^A \subseteq \mathcal{X}_N$ since the control law obtained is more conservative as $\Omega_W \subseteq \Omega_W^A$.

Remark 5: While more conservative than $\mathcal{P}_N$, $\mathcal{P}_N^A$ is less conservative than $\mathcal{P}_N^L$, the FH problem when parametrization (5) is used. Again, this is true because $\mathcal{P}_N^L$ is a special case of $\mathcal{P}_N^A$ 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

$$\Omega_W^B = \{(w^1,\ w^2)|\ w^1 + w^2 \in W, w^1 \geq 0, w^2 \geq 0\} \tag{25}$$

and its connection to Ω_W is given below.

Theorem 2: Suppose W satisfies assumption (A2), then (i) $\Omega_W^B = \mathrm{CH}(\Omega_W)$. (ii) $\max_{\omega \in \Omega_W} y^T\omega = \max_{\omega \in \mathrm{CH}(\Omega_W)} y^T\omega = \max_{\omega \in \Omega_W^B} y^T\omega$ for any given $y \in \mathbb{R}^{2n}$.

Proof: See Appendix I. □

The following example verifies Theorem 2.

Example 2: Let $\bar{W} = \{\bar{w}|\ |\bar{w}| \leq 0.2\}$ which is an absolute set. Fig. 1 shows $\bar{W}$, $\Omega_{\bar{W}}$, $\Omega_{\bar{W}}^A$ and $\Omega_{\bar{W}}^B$. It is clear that $\Omega_{\bar{W}}^B$ is the convex hull of $\Omega_{\bar{W}}$ as stated in (i) of Theorem 2 and that $\max_{w \in \Omega_{\bar{W}}} w^T y = \max_{w \in \Omega_{\bar{W}}^B} w^T y$ for any $y \in R^2, y \neq 0$ can also be easily verified.

The rest of this section shows that all absolute set can be expressed in the form of

$$V = \{v\ :\ \eta(v) \leq 1\}, \tag{26}$$

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

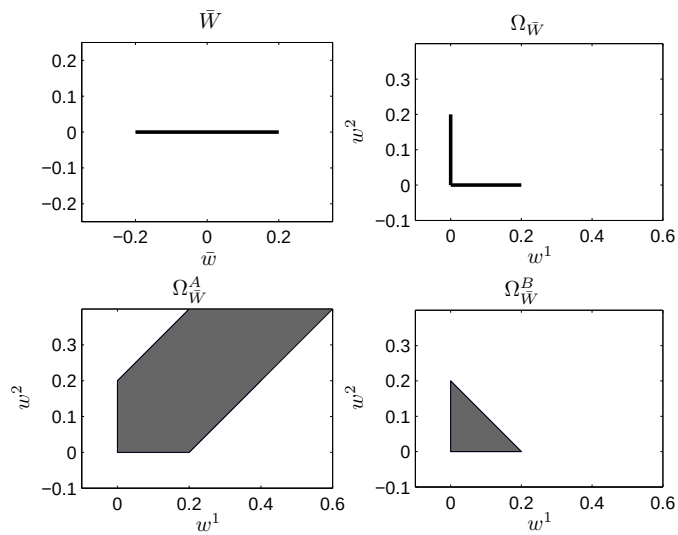

Fig. 1. Comparison of segregated disturbance sets

necessary absolute. It is easy to see that the following composite norm function

$$\zeta(v) = \max_{l=1,\ldots,L} \{a_l \eta_l(v)\}, \tag{27}$$

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 (26) with $\eta(v) = \max\{\frac{1}{r}\|v\|_2, \|v\|_\infty\}$. Given a vector norm $\eta(\cdot)$, the dual norm $\eta^* : \mathbb{R}^n \mapsto \mathbb{R}$ is a norm function defined as

$$\eta^*(y) = \max_{\eta(v)\leq 1} y^T v. \tag{28}$$

Some useful and relevant properties of the dual norm are collected below.

Lemma 1: Suppose $\eta(\cdot)$ and $\eta^*(\cdot)$ are an absolute norm and its dual norm. Then (i) $\eta^*(\cdot)$ is also an absolute norm function (ii) $\eta^{**}(\cdot) = \eta(\cdot)$. (iii) The dual norm of the L_p norm, $p \geq 1$, $\|v\|_p = \left(\sum_{j=1}^n |v_j|^p\right)^{1/p}$, is the L_q norm $\|w\|_q$ with $q = 1 + \dfrac{1}{p-1}$. In particular, the dual of an L_1 norm is the L_∞ norm and vice versa. (iv) The dual norm of the composite norm (27) is $\zeta^*(y) = \min\left\{\sum_{l=1}^L \frac{1}{a_l}\eta^*(y^l) \;:\; \sum_{l=1}^L y^l = y, y^l \in \mathbb{R}^n \;\forall l \in \mathbb{Z}_L^+\right\}$.

Proof: See Appendix II. □

The following example demonstrates (iv) of Lemma 1.

Example 3: Let $W = \{w \in \mathbb{R}^2 |\; \|w\|_\infty \leq 0.2,\; \|w\|_2 \leq 0.224\}$ which is shown in Figure 2. The corresponding absolute norm is $\eta(w) = \max\{5\|w\|_\infty, 4.64\|w\|_2\}$ and its dual norm is

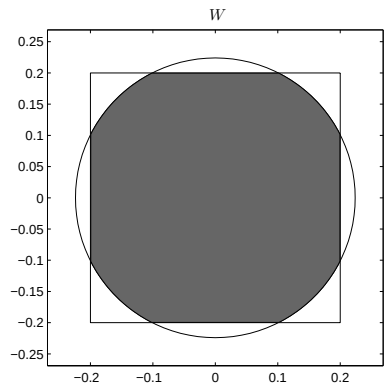

Fig. 2. W defined by composite norm

$\eta^*(z) = \min\{0.2\|y^1\|_1 + 0.224\|y^2\|_2 : \ y^1 + y^2 = z\}$.

Theorem 3: A set V is an absolute set if and only if there exists an absolute norm function, $\eta(\cdot)$ such that $V = \{v \ : \ \eta(v) \leq 1\}$.

Proof: See Appendix III. □

IV. CONVEX REFORMULATION AND COMPUTATION

One additional result is needed to show the conversion of the constraints (17) and (18) and is given in the following theorem.

Theorem 4: 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 (25). 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: See Appendix IV. □

For clarity in the sequel, a subscript W is appended to the absolute norm function η to indicate its association. Hence, $W = \{w | \eta_W(w) \leq 1\}$, η_W^* is its dual norm and let W^N be the N times cartesian product of W. Define $\eta_{W^N} : \mathbb{R}^{Nn} \mapsto \mathbb{R}$ as $\eta_{W^N}(\mathbf{v}) := \max_{i \in \mathbb{Z}_{N-1}}\{\eta_W(v(i))\}$ where $\mathbf{v} := [v^T(1) \ v^T(2) \ \cdots \ v^T(N)]^T$ and $v(i) \in R^n$. Then, it follows that W^N can be represented by

$W^N = \{\mathbf{w}|\eta_{W^N}(\mathbf{w}) \leq 1\}$. The corresponding dual norm of $\eta_{W^N}(\cdot), \eta^*_{W^N} : \mathbb{R}^{Nn} \mapsto \mathbb{R}$, is given by

$$\eta^*_{W^N}(\mathbf{q}) = \max\{\mathbf{q}^T\mathbf{v}|\eta_{W^N}(\mathbf{v}) \leq 1\} = \sum_{i=1}^{N} \max\{(q(i))^T v(i)|\eta_W(v(i)) \leq 1\} = \sum_{i=1}^{N} \eta^*_W(q(i)). \tag{29}$$

where $\mathbf{q} := [q^T(1) \ \cdots \ q^T(N)]^T$.

With these quantities and the characterizations of Y and X_f in (6) and (7), (15)-(18) become

$$\bar{\mathcal{A}}x + \bar{\mathcal{B}}\mathbf{d} + \bar{\mathcal{F}}\text{vec}(\left[\mathbf{C}^{p-}\mathbf{C}^{m-}\right]) + \max_{\mathbf{\Pi}^+ \in \Omega^N_W} \left[\bar{\mathcal{B}}\left[\mathbf{C}^{p+}\mathbf{C}^{m+}\right] + \left[\bar{\mathcal{G}} \ -\bar{\mathcal{G}}\right]\right]\mathbf{\Pi}^+ \leq 1_s \tag{30}$$

where $s = aN + b$ (a, b are the respective numbers of constraint in Y and X_f) and expressions of $\bar{\mathcal{A}}, \bar{\mathcal{B}}, \bar{\mathcal{G}}, \bar{\mathcal{F}}$ are given in Appendix V. Using results (i) and (ii) of Theorem 2, (30) can be restated as

$$\bar{\mathcal{A}}x + \bar{\mathcal{B}}\mathbf{d} + \bar{\mathcal{F}}\text{vec}(\left[\mathbf{C}^{p-}\mathbf{C}^{m-}\right]) + \max_{\mathbf{\Pi}^+ \in (\Omega^B_W)^N} \left[\bar{\mathcal{B}}\mathbf{C}^{p+} + \bar{\mathcal{G}} \quad \bar{\mathcal{B}}\mathbf{C}^{m+} - \bar{\mathcal{G}}\right]\mathbf{\Pi}^+ \leq 1_s. \tag{31}$$

Noting that $\mathbf{\Pi}^+ = [(\mathbf{w^{p+}})^T \ (\mathbf{w^{m+}})^T]$ and applying the result of Theorem 4, (31) has the following deterministic equivalence

$$\begin{cases} \bar{\mathcal{A}}x + \bar{\mathcal{B}}\mathbf{d} + \bar{\mathcal{F}}\text{vec}([\mathbf{C}^{p-}\mathbf{C}^{m-}]) + \mu \leq \mathbf{1}_s \\ T^T \geq \bar{\mathcal{B}}\mathbf{C}^{p+} + \bar{\mathcal{G}} \\ T^T \geq \bar{\mathcal{B}}\mathbf{C}^{m+} - \bar{\mathcal{G}} \\ \mu = \left[\eta^*_{W^N}(T_1) \ \cdots \ \eta^*_{W^N}(T_s)\right]^T \end{cases} \tag{32}$$

where T, $\bar{\mathcal{B}}\mathbf{C}^{p+} + \bar{\mathcal{G}}$ and $\bar{\mathcal{B}}\mathbf{C}^{p+} - \bar{\mathcal{G}}$ are analogous to t, x^T and y^T in Theorem 4 respectively and $\eta^*_{W^N}(\cdot)$ is that given in (29).

Remark 6: The adaptation of Theorem 4 to disturbance set defined by intersection of L_p norm sets is also quite easy. For example, if $W = \{w| \ \|w\|_\infty \leq 1, \ \|w\|_2 \leq r\}$, then $\eta_W(v) = \max\{\frac{1}{r}\|v\|_2, \|v\|_\infty\}$, $\eta^*_W(v) = \min\{r\|v^1\|_2 + \|v^2\|_1, v^1 + v^2 = v\}$ and the deterministic equivalence of $\mathcal{C}_1$ in Theorem 4 is $\mathcal{C}_2 = \{(x, y, z)| \ \exists t, t^1, t^2 \in \mathbb{R}^n, \|t^2\|_1 + r\|t^1\|_2 \leq z, t^1 + t^2 = t, \ t \geq x, \ t \geq y\}$. In such a case, (29) and (32) remains correct using the new expression of η^*_W.

V. Feasibility and Stability

Theorem 5: Suppose assumptions (A1)-(A4) are satisfied and $x(0) \in \mathcal{X}_N$. The system (1)-(2) with $u(t)$ obtained from (21) has the following properties: (i) $\mathcal{X}_N$ and $\Pi_N(x, \mathbf{w}^-)$ are convex sets; (ii) $\mathcal{P}_N(\mathbf{d}, \mathbf{C}, x(t), \mathbf{\Pi}^-(t))$ is feasible for all $t \geq 0$; (iii) $(x(t), u(t)) \in Y$ for all $t \geq 0$; (iv) $x(t) \to F_\infty$ as $t \to \infty$; (v) There exists a finite $\tilde{t}$ such that $x(t) \in X_f$ and $c(t) = 0$ for all $t \geq \tilde{t}$.

Proof: (i) From assumption (A2) and property (i) of Lemma 1, $\eta^*_{W^N}(\cdot)$ is a convex function. This means that the set of feasible $(x, \mathbf{d}, \mathbf{C}, T)$ of (32) is a convex set in $\mathbb{R}^n \times \mathbb{R}^{mN} \times \mathbb{R}^{nN\times s} \times \mathbb{R}^{2nm(N-1)N}$. The sets $\Pi_N(x, \mathbf{w}^-)$ and $\mathcal{X}_N$ are projections of this convex set onto the $(\mathbf{d}, \mathbf{C})$ and x space respectively and are hence convex sets.

The proof of (ii)-(v) follows essentially the arguments in [10] and uses the notation "$|t$" for the variables at time instant t. (ii) Feasibility of $\mathcal{P}_N(\mathbf{d}, \mathbf{C}, x(t), \mathbf{\Pi}^-(t))$ follows standard arguments. The specific choices are stated here only for the proof of (iv). More exactly, suppose $(\mathbf{d}^*, \mathbf{C}^*)$ is the optimal control at time t, the feasible control at time $t+1$ is chosen to be

$$\begin{cases} d(i|t+1) = d^*(i+1|t) & i \in \mathbb{Z}_{N-2} \\ C^k(i,j|t+1) = C^{k*}(i+1,j|t) & i \in \mathbb{Z}_{N-2},\ k \in \{p,m\} \\ d(N-1|t+1) = 0,\ C^k(N-1,j|t+1) = 0 & k \in \{p,m\} \end{cases} \tag{33}$$

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

(iv) If $J^*(t)$ is the optimal value of $\mathcal{P}_N(\mathbf{d}, \mathbf{C}, x(t), \mathbf{\Pi}^-(t))$ and let $\hat{J}(t+1)$ be the value of $J(\mathbf{d}, \mathbf{C})$ where $(\mathbf{d}, \mathbf{C})$ are defined by (33), then it can be verified that

$$J^*(t) - J^*_N(t+1) \geq J^*(t) - \hat{J}(t+1) = J^*_0(t)$$

where $J_0(t)$ is defined in (22). Hence, $\{J^*(t)\}$ is a monotonically non-increasing sequence bounded from below and it tends to a limit as $t \to \infty$. This necessary means that $J^*_0(t)$ tends to zero as $t \to \infty$. Hence, $c(t)$ tends to zero as $t \to \infty$. The system state under (21) can be written as

$$x(t) = \Phi^t x(0) + \sum_{i=0}^{t-1} \Phi^{t-1-i} B c(i) + \sum_{i=0}^{t-1} \Phi^{t-1-i} w(i) \tag{34}$$

where $\Phi = A + BK_f$. The first term on the right of (34) 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.

(v) 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 $\mathcal{P}_N(\mathbf{d}, \mathbf{C}, x(t), \mathbf{\Pi}^-(t))$. $\square$

As $\mathbf{d} \in \mathbb{R}^{mN}$, $T \in \mathbb{R}^{nN \times s}$ and $\mathbf{C} \in \mathbb{R}^{2nm(N-1)N}$, the number of variables in $\mathcal{P}_N$ is clear. Similarly, the number of constraints in (32) is $3s$ or $3(aN + b)$. It is of interest to compare these numbers with other parameterizations that are available in the literature. In this regard, we compare with the more recent disturbance feedback parametrization (4) used by Löfberg [8], van Hessem & Bosgra [9] and Goulart et. al. [3] assuming the same representations for Y, X_f sets and that W is an absolute set represented by a set of linear inequalities. It can be shown that the number of constraints using (4) is the same as the number of constraints under the proposed parametrization but (10) has additional $\mathbb{R}^{nm(N-1)N}$ variables in $\mathbf{C}$.

VI. Numerical example

The system parameters and constraints of the numerical example is given by:

$$A = \begin{bmatrix} 1.1 & 1 \\ 0 & 1.3 \end{bmatrix}, \; B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; W = \{w| \; \|w\|_\infty \leq 0.2\},$$
$$Y = \{(x, u)| \; |u| \leq 1, \|x\|_\infty \leq 6\},$$

and $w(t)$ is uniformly distributed over W. To implement the MPC controller, $K_f = [-0.4991 \;\; -0.9546]$ is chosen as the optimal feedback gain for the unconstrained LQR problem when $Q = I_2$ and $R = 1$. Terminal set X_f is chosen to be the maximal constraint-admissible disturbance invariant set of (1) under $u = K_f x$. The proposed approach is simulated from $x(0) = [-5 \; 2]'$ for the case where $N = 8$ and its result is shown in Fig. 3 and Fig. 4.

In Fig. 3, the outer bound $\hat{F}_\infty$ is used because the exact F_∞ is not computable. The procedure for computing $\hat{F}_\infty$ follows that given in [18]. It can be observed that the state converges to F_∞ set and all the constraints are satisfied all the time.

The next experiment compares the optimal costs of the FH optimization problems using parameterizations (5) and (10) for the case where $N = 8$. For a fair comparison, the cost

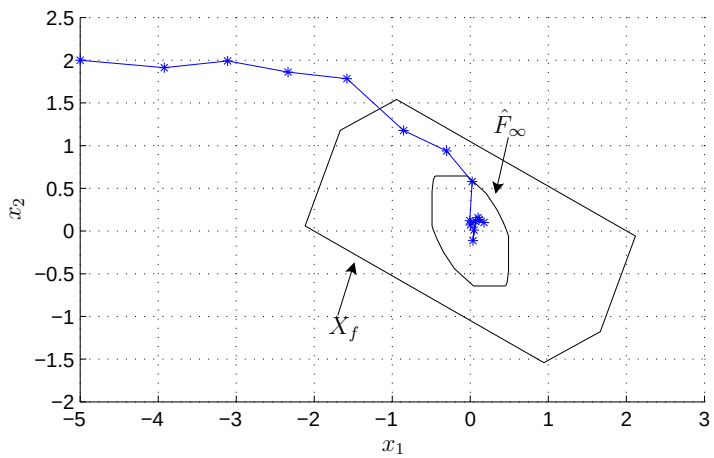

Fig. 3. State trajectory of example one

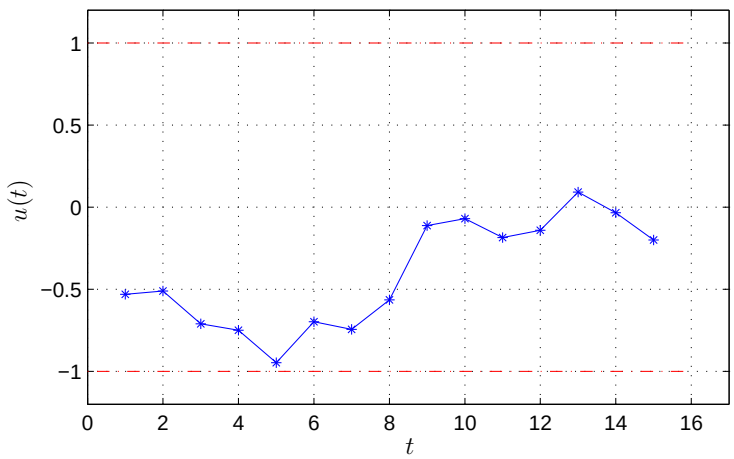

Fig. 4. Control trajectory of example one

functions of the two parameterizations should be consistent. For this purpose, the weight matrix, Λ, of $J(\mathbf{d}, \mathbf{C})$ is chosen according to (24) while the cost function associated with parametrization (5) is chosen according to [10], [19]. The implementation in [19] uses (4) as its control parametrization and hence has a different form for its cost function although it is equivalent to the expectation of a standard LQ cost. Our experiment with (5) uses the cost function $\sum_{i=0}^{N-1}\left[\|d(i)\|_\Psi^2 + \sum_{j=1}^{N-1}\|\mathrm{vec}(D(i,j))\|_\Upsilon^2\right]$ where $\Psi = R + B^T P B$ and $\Upsilon = \Sigma_w \otimes \Psi$ and is the same as that used in [19]. In this setting, both cost functions are equivalent to the expected

value of the same LQ cost. The optimal costs of both problems over the admissible region are compared and the result is shown in Fig. 5 where J_N^L is the optimal cost under parametrization (5) and J_N^S is that under (10). Clearly, parametrization (10) always yields a better cost than (5).

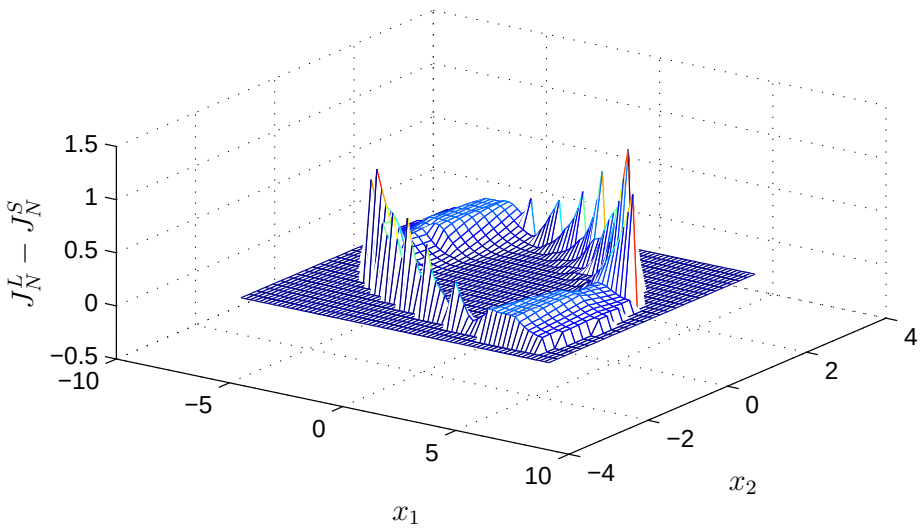

Fig. 5. Difference between the two optimal costs

VII. Conclusion

A piecewise affine disturbance feedback parametrization is proposed under the MPC formulation of constrained linear systems with disturbances. This parametrization includes the standard linear disturbance feedback as a special case, and hence, is less conservative. When the disturbance set is absolute, the FH optimization problem can be converted into an equivalent convex problem solvable with existing solvers. Even when the disturbance set is not absolute, the new parametrization still results in a MPC controller that is less conservative than the one derived from using linear disturbance feedback. The stability of the closed-loop system under the proposed parametrization is shown and the asymptotic behavior of the system is characterized by $F_\infty(K_f)$ where K_f is a user-defined constant feedback gain.

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] K. J. Åström and B. Wittenmark, *Computer-Controlled Systems: Theory and Design*. Prentice Hall, 3 ed., 1997.

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

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

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

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

[18] C. J. Ong and E. G. Gilbert, "The minimal disturbance invariant set: Outer approximations via its partial sums," *Automatica*, vol. 42, no. 9, pp. 1563–1568, 2006.

[19] 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.

[20] D. Bertsimas and M. Sim, "Tractable Approximations to Robust Conic Optimization Problems," *Mathematical Programming*, vol. 107, no. 1, pp. 5–36, 2006.

[21] P. Lax, *Linear Algebra*. John Wiley, 1997.

[22] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.

APPENDIX I
PROOF OF THEOREM 2

Proof: (i) ($\Rightarrow$)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 ith 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, $(u^1, u^2) \in \mathrm{CH}(\Omega_W)$.

(ii) Let the maximizer of $\max_{\omega\in\mathrm{CH}(\Omega_W)} \omega^T y$ be $\omega^* \in \mathrm{CH}(\Omega_W)$. Hence, there exist $\omega^1 \in \Omega_W$, $\omega^2 \in \Omega_W$ and $0 \leq \lambda^1,\ \lambda^2 \leq 1$, $\lambda^1 + \lambda^2 = 1$ such that $\omega^* = \lambda^1\omega^1 + \lambda^2\omega^2$. Therefor, $y^T\omega^* = \lambda^1 y^T\omega^1 + \lambda^2 y^T\omega^2$ and $\max_{\omega\in\mathrm{CH}(\Omega_W)} \omega^T y = y^T\omega^* \leq \max\{y^T\omega^1,\ y^T\omega^2\} \leq \max_{\omega\in\Omega_W} \omega^T y$. It is also clear that $\max_{\omega\in\Omega_W} \omega^T y \leq \max_{\omega\in\mathrm{CH}(\Omega_W)} \omega^T y$ since $\Omega_W \subseteq \mathrm{CH}(\Omega_W)$. Hence, $\max_{\omega\in\Omega_W} y^T\omega = \max_{\omega\in\mathrm{CH}(\Omega_W)} y^T\omega$. The second equality follows the result of (i). $\square$

APPENDIX II

PROOF OF LEMMA 1

Proof: (i) The proof can be found in [20]. (ii)-(iii)The first two results are well known, see [21]. (iv) Denote $\delta(y|C) := \max_{x\in C} x^T y$. Then, from (27),

$$\begin{aligned}
\zeta^*(y) &= \max\{y^T x | a_i \eta_i(x) \leq 1, i \in \mathbb{Z}_L^+\} \\
&= \max(y^T x | x \in \bar{C}_1 \cap \cdots \bar{C}_L\}) \\
&= \min\{\delta(y^1|\bar{C}_1) + \cdots + \delta(y^L|\bar{C}_L) | \sum_{i=1}^{L} y^i = y\} \qquad (35) \\
&= \min\{\sum_{i=1}^{L} \frac{1}{a_i}\delta(y^i|C_i) | y^1 + \cdots + y^L = y\} \qquad (36) \\
&= \min\{\sum_{i=1}^{L} \frac{1}{a_i}\eta^*(y^i) | y^1 + \cdots + y^L = y\} \qquad (37)
\end{aligned}$$

where $\bar{C}_i = \frac{1}{a_i} C_i$ and $C_i = \{x|\eta_i(x) \leq 1\}$ $\forall i \in \mathbb{Z}_L^+$. Equation (35) follows a property of support function (Corollary 16.4.1 of [22]). Specifically, suppose $C_1, C_2, \cdots C_m$ are non-empty convex sets in $\mathbb{R}^n$ such that $C_1 \cap C_2 \cdots \cap C_m \neq \emptyset$, then $\delta(y|C_1 \cap C_2 \cdots \cap C_m) = \min\{\delta(y^1|C_1) + \cdots + \delta(y^m|C_m) | y^1 + \cdots + y^m = x\}$. Equation (36) holds because $\delta(x|\alpha C) = \alpha\delta(x|C)$ for any $\alpha > 0$ while (37) follows from the definition of C_i. □

APPENDIX III

PROOF OF THEOREM 3

Proof: One direction is trivial. Conversely, it suffices to show that for any absolute set, V, an absolute norm function $\eta(\cdot)$ exists such that

$$\max\{y^T v \ : \ \eta(v) \leq 1\} = \max\{y^T v \ : \ v \in V\}$$

for all $y \in \mathbb{R}^n$. Consider the support function of V, $\delta(y|V)$, which by inspection is an absolute norm. Note that any convex, compact and symmetric set with zero in the interior has support function that satisfies the properties of a norm. Hence, from (28) and property (i) of *Lemma* 1, the corresponding dual norm function $\delta^*(v) = \max\{v^T y \ : \ \delta(y|V) \leq 1\}$ is also an absolute

norm function. Let $\eta(\cdot) = \delta^*(\cdot)$. Hence, by strong duality of norms, we have for all $y \in \mathbb{R}^n$,

$$\max\{y^T v : \eta(v) \leq 1\} = \eta^*(y) = \delta^{**}(y) = \delta(y|V) == \max\{y^T v : v \in V\}.$$

□

APPENDIX IV

PROOF OF THEOREM 4

Proof: ($\Rightarrow$) Let (x, y, z) be an element of $\mathcal{C}_1$. It follows from (25) that

$$\begin{aligned}
z \geq \;& \max\{x^T w^1 + y^T w^2 | w^1 \geq 0, w^2 \geq 0, \eta(w^1 + w^2) \leq 1\} \\
= \;& \max\{x^T w^1 + y^T w^2 |\; w^1 \geq 0, w^2 \geq 0,\; w = w^1 + w^2,\; \eta(w) \leq 1\} \\
= \;& \max\{\bar{t}^T w |\; w \geq 0,\; \eta(w) \leq 1,\; \bar{t}_i = max\{x_i, y_i\}\} && (38) \\
= \;& \max\{\bar{t}^T |w| \mid \eta(w) \leq 1,\; \bar{t}_i = max\{x_i, y_i\}\} && (39) \\
= \;& \max\{t^T |w| \mid \eta(w) \leq 1,\; t_i = max\{0, \bar{t}_i\}\} && (40) \\
= \;& \max\{t^T w |\; \eta(w) \leq 1,\; t_i = max\{0, \bar{t}_i\}\} && (41) \\
\Rightarrow \;& (x, y, z) \in \mathcal{C}_2
\end{aligned}$$

The first two relations come from the definitions of Ω_W^B, W and the re-organization of the constraints. Equation (38) 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 (39) follows because W is an absolute set. Equation (40) 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 property (1) of Lemma 1. 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

$\eta^*(\cdot)$,

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

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 $t^Tw \geq x^Tw$ and $t^Tw \geq y^Tw$ 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. □

APPENDIX V

MATRICES IN (30)

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 (30) 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 $\bar{\mathcal{F}} = \mathcal{Y}((\mathbf{\Pi}^-)^T \otimes \left[(\varphi\mathcal{B})^T\ (I+\mathcal{K}\varphi\mathcal{B})^T\right]^T)$