LP Methods in MPC of Large-Scale Systems: Application to Paper-Machine CD Control

Prashant Dave, Dennis A. Willig, Gautham K. Kudva, Joseph F. Pekny, and Francis J. Doyle
School of Chemical Engineering, Purdue University, West Lafayette, IN 47907

An application of a linear-programming based model-predictive control strategy to the paper-machine cross-direction (CD) control problem is presented. The objective of CD control is to maintain flat profiles of variables of interest by minimizing worst-case deviations from setpoints (defects). These control problems can have as many as 200 actuators (inputs) and 400 sensor measurements (outputs). This large size coupled with the stringent real-time requirement of computing a control move in a few seconds poses a very challenging control problem. The LP-based strategy is particularly well suited for solving such classes of control problems. This strategy has demonstrated its ability to solve large-scale control problems (over 100 inputs and 100 outputs) in real time and exhibits robustness to model uncertainty.

Introduction

With increasing global competition in the chemical process industry, there is an urgent need to develop and implement technology to more efficiently utilize scarce resources. The pulp and paper industry is no exception to this general trend. Indeed, recent developments seem to indicate a trend of advanced process technologies including pulping, bleaching, headbox operation, and drying (Dumont, 1988). These new technologies will require more sophisticated control schemes that utilize the on-line measurements of various properties of interest.

Traditionally, the paper machine control problem has been divided broadly into two categories, depending upon the direction relative to the flow, namely the machine direction (MD), which is along the flow, and the cross direction (CD), which is perpendicular to the flow. An example of MD control is the regulation of moisture content of the paper in the direction of the flow. This is usually accomplished by manipulating the steam pressure into the roller drier. The objective of the CD control is to minimize deviations of process variables (for example, thickness, basis weight) from the setpoint across the width of the paper.

Paper machine CD control has received a significant amount of attention in the last decade (for representative results, see, for example, Bergh and MacGregor, 1987; Boyle, 1978; Braatz et al., 1992; Campbell and Rawlings, 1994; Kristinsson, 1994; Tyler and Morari, 1995). In all of these studies, control has been achieved by employing quadratic-programming-based (QP) approaches. The QP-based strategies attempt to minimize the square of the deviation of process variables from their setpoints (Boyle, 1977). Due to the excessive computational requirements of optimal QP algorithms, quadratic-penalty-function-based (QPF) techniques have been proposed (Chen and Wilhelm, 1986). However, QPF strategies yield suboptimal solutions compared to the corresponding exact QP strategies.

Previous studies have generally ignored uncertainties in the model and the measurements. An exception is the work of Laughlin, in which a multiple-input multiple-output (MIMO) robust control scheme based on internal model control (IMC) (Laughlin, 1988; Laughlin et al., 1986; Laughlin and Morari, 1989) is proposed that addresses model and parameter uncertainties. However, this approach requires the gain matrix describing the relationship between actuators and sensors to be positive definite. Some gain matrices reported in the literature do not possess this property.

To date, the literature has clearly been biased toward the use of quadratic objective functions. However, in many practical applications, the true goal of CD control is to minimize the worst-case deviations ("defects") of variables of interest from their desired setpoints (Smith, personal communication, 1995). This translates directly to minimizing the maximum deviation from the setpoint; in other words, the infinity-norm of the error. In addition to capturing the true goal, a control algorithm must generate solutions in real time. The real-time

Correspondence concerning this article should be addressed to F. J. Doyle.

issue has become increasingly important with the increase in speed of the paper machine. These real-time considerations are not necessarily an issue if the control problem is small. However, the size and measurement resolution of paper machines is increasing, leading to a large number of control and manipulated variables. This makes the control problem large in scale. In fact, this problem can have as many as 200 manipulated and 400 control variables. Therefore, effective control strategies, which represent the true objective of CD control, must not only address the stringent real-time requirements, but also grapple with the size of the problem.

Linear programming (LP) problems have received considerable attention over the past 50 years. Algorithms have been devised to routinely solve large LP problems that arise in practice. We propose to exploit the attractive properties of LP problems to address the problem at hand, that is, the paper machine CD control problem. LP-based control strategies have been proposed for model-predictive control (MPC) (Campo and Morari, 1986; Allwright and Papavasiliou, 1992; Zheng and Morari, 1993). However, no applications of LP technology to large-scale problems exist in the literature. In this article, we present an LP-based control algorithm that adequately addresses the issues concerning large control problems in general, and the paper machine CD control problem in particular. Our results demonstrate that achieving optimal on-line control of large-scale unit operations, such as the paper machine CD control, is well within reach.

In the next section we review the MPC algorithm and derive LP formulations to the problem. In the third section we provide a detailed description of the paper machine. In the fourth section we introduce the model for the paper machine CD control problem and present simulation results. Finally, we summarize the results.

Background

In this section, we briefly review the MPC algorithm. Then we investigate the conditions under which the control problem can be cast as an LP problem.

The underlying principle of MPC is that a model can be used to predict the effects of past and future inputs on the future system outputs (Garcia et al., 1989). The horizon over which these predictions are based is depicted in Figure 1.

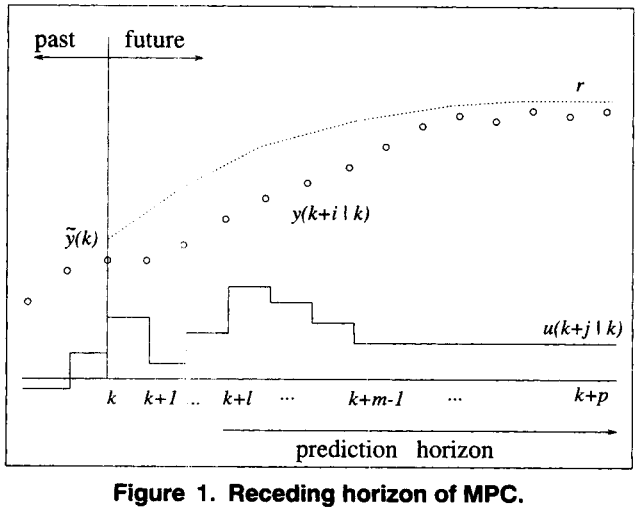

Figure 1. Receding horizon of MPC.

At the present time, k, the behavior of the process over a future prediction horizon, $[l, p]$, is calculated using a model. The manipulated variable moves over the next m time periods, denoted by $\Delta\mathfrak{U}$, are computed in order to achieve a stated objective. This computation can be cast as an optimization problem:

$$\min_{\Delta\mathfrak{U}} \quad \text{(deviation from setpoint + input penalty)}$$

subject to:

Plant model

Physical constraints

Product specifications.

We are interested in the conditions under which the preceding problem can be recast as an LP problem. This transformation can be carried out when the plant model, physical constraints, and product specifications are linear and when the objective function can be expressed as a linear function of the decision variables. Indeed, linear process models are commonly used and physical constraints and product specifications are most often linear. The preceding formulation can be rewritten as:

$$\min_{\Delta u(k|k),\ldots,\Delta u(k+m-1|k)} \quad \text{OBJ} = f_1(y) + f_2(u) \tag{1}$$

subject to:

$$y(k+i|k) = \tilde{y}(k) + \sum_{j=0}^{m-1} P_j \Delta u(k+j|k) \qquad \forall\ i = l, \ldots, p \tag{2}$$

$$u(k+j|k) = \tilde{u}(k) + \sum_{t=0}^{j} \Delta u(k+t|k) \qquad \forall\ j = 0, \ldots, m-1 \tag{3}$$

$$y^L \le y(k+i|k) \le y^U \qquad \forall\ i = l, \ldots, p \tag{4}$$

$$-\Delta u^{\max} \le \Delta u(k+j|k) \le \Delta u^{\max} \qquad \forall\ j = 0, 1, \ldots, m-1 \tag{5}$$

$$u^{\min} \le u(k+j|k) \le u^{\max} \qquad \forall\ j = 0, 1, \ldots, m-1 \tag{6}$$

$$\text{Other (Linear) Constraints.} \tag{7}$$

The objective function, OBJ, consists of two terms. The first term, $f_1(y)$, is a measure of the deviation of the predicted output from the setpoint, and the second term, $f_2(u)$, is an input penalty that is used to control the aggressiveness of the control algorithm. The decision variables, $\Delta u(k|k)$, $\Delta u(k+1|k)$, ..., $\Delta u(k+m-1|k)$, are the manipulated variable moves over the control horizon, m. The penalty term, $f_2(u)$, is given by $\Sigma_{j=0}^{m-1} \|\Gamma_u(\Delta u(k+j|k))\|_1$, where Γ_u is an input weighting matrix. This term scales each of the control moves by the weighting matrix, Γ_u, which consists of nonnegative elements. A convenient choice for Γ_u is $\gamma_u \times I$, where I is the identity matrix and γ_u is a scalar. The predicted output at time period $k+i$, $y(k+i|k)$, is also related to manipulated variable moves and a known vector $\tilde{y}(k)$. This vector contains measurement information and also captures the ef-

fect of past inputs. Equation 3 relates the manipulated variables, $u(k+j|k)$, to the manipulated variable moves, $\Delta u(k+j|k)$, over the control horizon, m.

The constraints associated with the previous optimization problem are the rate constraints on u, Eq. 5, and upper and lower bounds on u, Eq. 6. We can also add other linear constraints, for example, product specifications, Eq. 4, or in the case of the paper machine, the bending moment constraints of the actuator bar.

Let us examine the conditions under which the objective function is linear. The second term in the objective function, $f_2(u)$, is a sum of one-norms of vectors. This can be transformed into a linear function. The first term of the objective function, $f_1(y)$, is a mapping of $l-p+1$ error vectors to a scalar. Therefore, this mapping needs to be analyzed in more detail in order to determine the conditions under which this term can be cast as a linear function. We are given $l-p+1$ error vectors, $E=\{e_l, \ldots, e_p\}$, where each e_i is a vector of dimension N_y, where N_y is the number of outputs. Note that there are two key dimensions, the space of the output variables (N_y) and time, which is the prediction horizon ($[l, p]$). Thus, $f_1(y)$ is a projection from a set of $p-l+1$ vectors each of dimension N_y to a scalar. In order to achieve this projection, the vector p-norm is employed:

$$\|v\|_p = \left(\sum_{i=1}^{D} |v_i|^p \right)^{1/p}, \tag{8}$$

where $v \in \Re^D$. For values of $p \in \{1, 2, \infty\}$ Eq. 8 reduces to:

$$\|v\|_1 = \sum_{i=1}^{D} |v_i| \tag{9}$$

$$\|v\|_2 = \sqrt{\sum_{i=1}^{D} v_i^2} \tag{10}$$

$$\|v\|_\infty = \max_{1 \le i \le D} |v_i|. \tag{11}$$

Each of the individual error vectors e_i is mapped to a scalar. This is accomplished by taking a spatial norm, that is, a norm in the space of the output variables. This procedure is repeated for the time index. If we use norms in the set $\{1, 2, \infty\}$ for each of the spatial and temporal norms, then we get nine possible spatial-temporal norm combinations. An expression for the scalar error for each of these nine different combinations is given in Table 1. Four of these combinations result in linear programming problems, and the other five result in nonlinear (quadratic) programming problems. It is important to point out that the original academic research on LP formulations of MPC was by Campo (1989). However, the computational aspects of large-scale implementation of such algorithms, which were not considered by Campo, are the focal point of the present work.

Formulating the MPC Problem as an LP

In this section we present formulations of the optimization problem (Eqs. 1–7) as LP problems. As we have seen in the previous section (see Table 1), the conditions under which the objective function can be expressed as a linear function of the decision variables corresponds to the following choices of the norms: $\{1-1, 1-\infty, \infty-1, \infty-\infty\}$. We will use the convection that the first number in $p-q$ norm refers to the spatial norm and the second refers to the temporal norm.

Table 1. Characterization of the Scalar Error

No.	Spatial Norm	Temporal Norm	Expression	Problem
1	1	1	$\sum_{i=l}^{p} \|e_i\|_1$	LP
2	1	2	$\sqrt{\sum_{i=l}^{p} (\|e_i\|_1)^2}$	QP
3	1	∞	$\max_{l \le i \le p} \|e_i\|_1$	LP
4	2	1	$\sum_{i=l}^{p} \|e_i\|_2$	QP
5	2	2	$\sqrt{\sum_{i=l}^{p} (\|e_i\|_2)^2}$	QP
6	2	∞	$\max_{l \le i \le p} \|e_i\|_2$	QP
7	∞	1	$\sum_{i=l}^{p} \|e_i\|_\infty$	LP
8	∞	2	$\sqrt{\sum_{i=l}^{p} (\|e_i\|_\infty)^2}$	QP
9	∞	∞	$\max_{l \le i \le p} \|e_i\|_\infty$	LP

Below, we examine the four cases leading to an LP.

$\infty-\infty$

The first term of the objective function is the infinity norm in time and the infinity norm in the space of the output variables:

$$f_1(y) = \left(\max_{l \le i \le p} \|\Gamma_y [r(k+i) - y(k+i|k)]\|_\infty \right), \tag{12}$$

where $y(k+i|k)$ denotes the predicted output at $k+i$ using information available up to time period k. Therefore, $r(k+i)-y(k+i|k)$ denotes the predicted error vector. This error vector is scaled by a matrix Γ_y, called the output weighting matrix, having nonnegative elements. The $\infty-\infty$ norm of the errors (see entry 9 in Table 1) is precisely the first term of the objective function. The start of the prediction horizon is denoted by l, which can be adjusted to handle systems with deadtime or inverse response, and the end of the prediction horizon is denoted by p. A convenient choice of Γ_y is $\gamma_y \times I$, where I is the identity matrix of appropriate dimensions and γ_y is a scalar. The optimization problem, Eqs. 1–7, can be written as an LP by employing the following substitutions. A quantity $|x|$, where x is a scalar, can be represented, under certain conditions, by: $|x| = x^+ + x^-$, where $x = x^+ - x^-$, x^+, $x^- \ge 0$. We can replace all occurrence of $|x|$ by a sum of two nonnegative variables, and all occurrences of x, by the difference of two nonnegative variables. The linear programming problem that results is given below:

$$\text{minimize} \quad \theta + \gamma \tag{13}$$

subject to:

$$\alpha_i^+ + \alpha_i^- \le \Theta \qquad \forall\, i = l, \ldots, p \tag{14}$$

$$\Gamma_y [r(k+i) - y(k+i|k)] = \alpha_i^+ - \alpha_i^- \qquad \forall\, i = l, \ldots, p \tag{15}$$

$$y(k+i|k) = \tilde{y}(k) + \sum_{j=0}^{m-1} P_j[\Delta u^+(k+j|k) - \Delta u^-(k+j|k)]$$

$$\forall\, i = l, \ldots, p \quad (16)$$

$$u(k+j|k) = \tilde{u}(k) + \sum_{t=0}^{j} [\Delta u^+(k+t|k) - \Delta u^-(k+t|k)]$$

$$\forall\, j = 0, \ldots, m-1 \quad (17)$$

$$\sum_{j=0}^{m-1} \sum_{q=1}^{N_u} \Gamma_u[\Delta u^+(k+j|k) + \Delta u^-(k+j|k)]\{q\} = \gamma \quad (18)$$

bounds:

$$\alpha_i^+,\ \alpha_i^- \geq 0 \quad (19)$$

$$y^L \leq y(k+i|k) \leq y^U \quad (20)$$

$$u^{\min} \leq u(k+j|k) \leq u^{\max} \quad (21)$$

$$0 \leq \Delta u^+(k+j|k),\quad \Delta u^-(k+j|k) \leq \Delta u^{\max}, \quad (22)$$

where

$$\theta,\ \gamma \in \Re \quad (23)$$

$$\Theta \equiv (\theta, \ldots, \theta)^T \in \Re^{N_y} \quad (24)$$

$$\alpha_i^+,\ \alpha_i^-,\ y(k+i|k) \in \Re^{N_y} \quad (25)$$

$$\Delta u^+(k+j|k),\quad \Delta u^-(k+j|k),\quad u(k+j|k) \in \Re^{N_u} \quad (26)$$

$$\Gamma_y \in \Re^{N_y \times N_y} \quad (27)$$

$$\Gamma_u \in \Re^{N_u \times N_u} \quad (28)$$

$$P_j \in R^{N_y \times N_u} \quad (29)$$

$$\tilde{y}(k),\quad r(k+i) \in \Re^{N_y} \quad (30)$$

$$\tilde{u}(k) \in \Re^{N_u}. \quad (31)$$

The objective function, $\theta + \gamma$, consists of two terms. The first term, θ, expresses the $\infty - \infty$ norm of the predicted error vectors and the second term, γ, expresses the penalty on input moves. Variables α_i^+ and α_i^- have been introduced to capture the $\infty - \infty$ norm of the error vectors. The control moves, $\Delta u(k+j|k)$, have been expressed as a difference of two nonnegative vectors, $\Delta u^+(k+j|k)$ and $\Delta u^-(k+j|k)$. Θ is a vector in $\Re^{N_y}$ that has all of its components θ. Constraint 14 can also be written as

$$\alpha_i^+\{j\} + \alpha_i^-\{j\} \leq \theta \qquad \forall\, i, j. \quad (32)$$

The vector given by the expression $\alpha_i^+ - \alpha_i^-$ is equal to the scaled error $\Gamma_y[r(k+i) - y(k+i|k)]$. The model of the process is employed in Eq. 16 to relate the predicted output $y(k+i|k)$ to the manipulated variables moves. Equation 17 relates the manipulated variable moves to manipulated variable values. The last Eq. 18 captures the input penalty. This formulation has $3(p-l+1)N_y + 3mN_u + 2$ variables and $3(p-l+1)N_y + mN_u + 1$ constraints. Note that, Γ_y, r, $\tilde{y}$, P_j, $\tilde{u}$, Γ_u are known and the bounds, Eqs. 19–22, are not considered as explicit constraints.

1 – 1

The first term in the objective function for the $1-1$ case is:

$$f_1(y) = \sum_{i=l}^{p} \|\Gamma_y[r(k+i) - y(k+i|k)]\|_1. \quad (33)$$

As in the $\infty - \infty$ norm case, we scale the error vectors by Γ_y. The linear programming problem that results is: minimize $\theta + \gamma$, subject to $\sum_{i=l}^{p}\sum_{j=1}^{N_y}(\alpha_i^+\{j\} + \alpha_i^-\{j\}) = \theta$, and Eqs. 15–18. The bounds of the variables are given by Eqs. 19–22. The number of variables in this LP is $3(p-l+1)N_y + 3mN_u + 2$, and the number of constraints, not including bounds, is $2(p-l+1)N_y + mN_u + 2$.

∞ – 1

For this case $f_1(y)$ is the $\infty - 1$ norm of the error vectors:

$$f_1(y) = \sum_{i=l}^{p} \|\Gamma_y[r(k+i) - y(k+i|k)]\|_\infty. \quad (34)$$

The LP for this case is: minimize $\sum_{i=l}^{p}\theta_i + \gamma$, subject to $\alpha_i^+ + \alpha_i^- \leq \Theta_i$ $\forall i = l, \ldots, p$, and Eqs. 15–18, where θ_i, $\gamma \in \Re$, $\Theta_i \equiv (\theta_i, \ldots, \theta_i)^T \in \Re^{N_y}$. The bounds of the variables are given by Eqs. 19–22. The number of variables in this LP is $3(p-l+1)N_y + 3mN_u + p - l + 2$ and the number of constraints, not including bounds, is $3(p-l+1)N_y + mN_u + 1$.

1 – ∞

For this case $f_1(y)$ is the $1-\infty$ norm of the error vectors:

$$f_1(y) = \left(\max_{l \leq i \leq p} \|\Gamma_y[r(k+i) - y(k+i|k)]\|_1 \right). \quad (35)$$

The corresponding LP is: minimize $\theta + \gamma$, subject to $\sum_{j=1}^{N_y}(\alpha_i^+\{j\} + \alpha_i^-\{j\}) \leq \theta$ $\forall\, i = l, \ldots, p$, and Eqs. 15–18. The bounds of the variables are given by Eqs. 19–22. The number of variables in this LP is $3(p-l+1)N_y + 3mN_u + 2$, and the number of constraints, not including bounds, is $2(p-l+1)N_y + mN_u + p - l + 2$.

Table 2 contains the dimensions of the linear programming problems generated by the four different cases.

Paper Machine Control Problem

The complex network that makes up modern pulp and paper mills can be broken down into three main areas: the

Table 2. Dimension of the Linear Programming Problems

Type	Columns (Variables)	Rows (Constraints)
$1-1$	$X+2$	$Y+2$
$1-\infty$	$X+2$	$Y+p-l+2$
$\infty-\infty$	$X+2$	$Y+(p-l+1)N_y+1$
$\infty-1$	$X+p-l+2$	$Y+(p-l+1)N_y+1$

where, $X = 3(p-l+1)N_y + 3mN_u$
$Y = 2(p-l+1)N_y + mN_u$

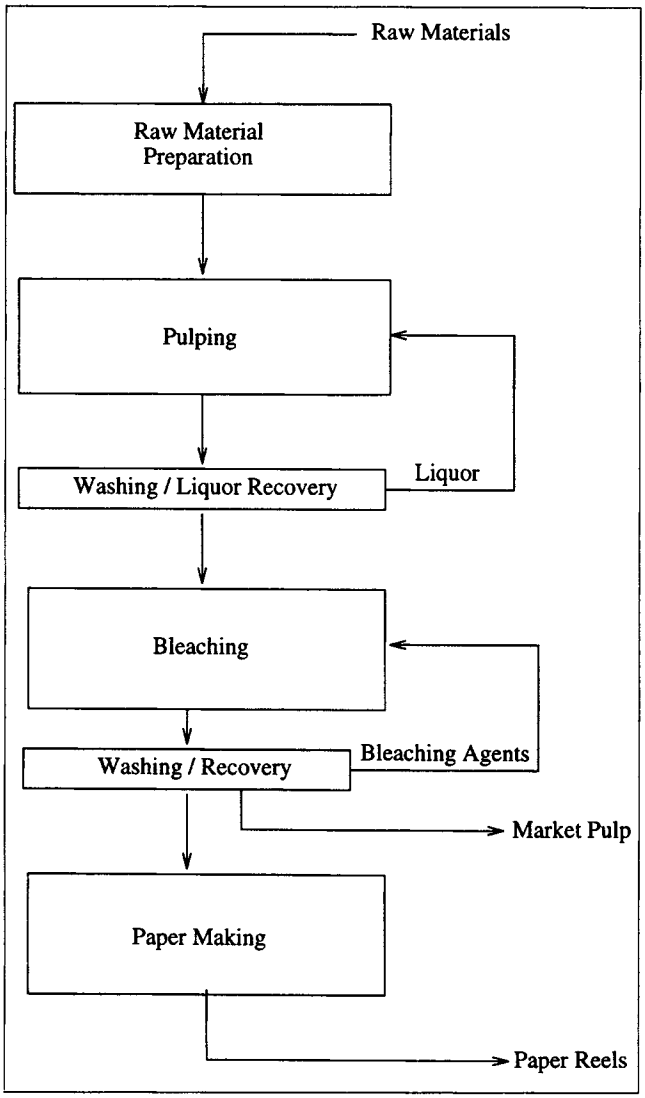

Figure 2. Overview of the pulp- and paper-making process.

pulping process, the bleaching process, and the paper-making process. Figure 2 depicts the overall process in the form of a flow diagram. To begin with, raw material, in the form of wood chips, is turned into pulp. This pulping of wood chips represents a set of chemical reactions that primarily attack lignin. The lignin is then solubilized and removed as a component of black liquor. To achieve this, the raw material is cooked in a process unit known as a digester. Most often, the digester is a vertical reactor in which the pulp mixture flows from top to bottom. Before the wood chips enter this reactor, a chemical solution, or "liquor," is added. This solution, along with the process conditions of the digester, dissolve the binding material that holds wood together. Typical processing times for a batch of raw material, called brown stock, can approach 4 to 5 hours. After leaving the pulping process area, the unbound raw material travels to the bleaching unit. In addition to removing any coloring compounds that may be present, this intermediate step also removes residual lignin that might be present. To accomplish this task, typically five to six reactors in series are used. This step has a typical processing time of about 2 hours. The final step in the simplified paper mill model is paper making. Compared to the previous two stages, the paper machine is the most instrumented unit in the plant. This is due to the fact that processing times range from 4 to 10 seconds (as compared to numerous hours for the previous two steps), and that this step involves creating the mills' highly valued end product, paper.

Figure 3 depicts a typical paper machine, measuring 20 ft (6.1 m) in width and 50 yd (45.7 m) in length, and running at a process rate of 4,000 ft/min (20 m/s) (Laughlin, 1988). The major components that make up this machine include the mixing chest, the headbox, the fourdinier wire, the cylinder dryers, the quality-control area, and the product reels.

The feed section of this process includes the mixing chest and the headbox. The mixing of pulp from the bleaching step and fresh water occurs in the mixing chest. Also mixed into this box is any water/pulp mixture that falls off the fourdrinier wire in the initial processing steps. This box is responsible for providing a uniformly distributed feed to the headbox. It is the machine's first defense against any unexpected disturbances.

The feed is pumped from the mixing chest into the headbox. The purpose of the headbox is to distribute the proper amount of the feed fiber suspension onto the fourdinier wire. This distribution is defined as the desired flow of suspension having some desired cross-directional profile. The headbox is a pressurized vessel. Thus, since air is being pumped into the box, an "air pad" develops between the stock surface and the top of the vessel. The total head inside of the box is then the air-pad pressure plus the stock's level or height. Manipulation of this head is the means by which the outlet suspension velocity is controlled. Motorized actuators are used to control the cross-directional profile. These actuators are spaced at some given interval and control the headbox's outlet width. The outlet of the headbox is called the slice lip. The slice lip is in the shape of a horizontal slit. The fourdinier wire is next. The wire supports the wet suspension and begins the drying process. The wire is like a screen; its mesh is small enough to support the water/fiber mixture, while at the same time it is large enough to allow for excess water to escape downward. Any excess of the water/fiber mixture that falls from the process is collected in the wire pit and recycled into the mixing chest. At the end of the wire section is a set of presses, which are responsible for the compression and smoothing of the suspension before it enters the cylinder dryer section.

Equipment for drying the suspension takes up, by far, the most floor space. The main process for drying the wet paper involves winding the paper through a network of steam-heated cylindrical rolls. A control loop manipulates the steam flow to the dryers to control the moisture content of the paper.

One of the most important parts of the paper machine is the quality-control area. It is here that all the product-specific parameters, such as the basis weight, moisture content, and caliper/thickness, are monitored (Davies et al., 1993). Typically, this area consists of a flat region where a sensor mounted on a swinging arm can traverse the entire cross direction of the paper. It is important to note that the paper has to cover a relatively large distance (about 100 yd, 91 m) from the headbox to the quality-control area.

The final stage of the paper machine is the product reel,

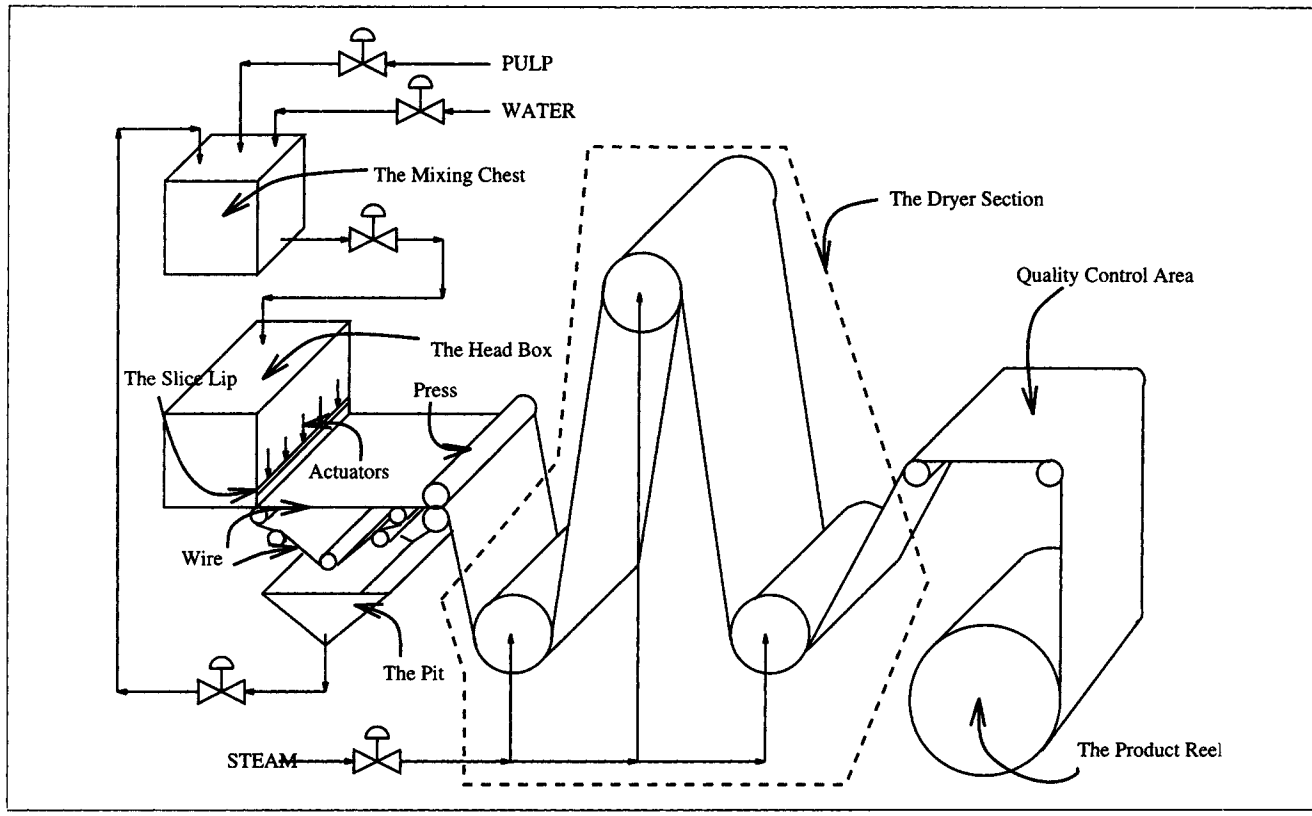

Figure 3. Typical paper machine.

which is responsible for the storage of large amounts of paper. This method of rolling the paper is also convenient for postproduction transport and further processing.

The three fundamental properties of paper, namely, the basis weight, the moisture content, and the caliper/thickness, can be measured in two directions: in the process flow or machine direction (MD), and in the cross direction (CD). Variations from setpoint in the MD can be due to a number of disturbances that may occur mainly upstream of the headbox. For example, whenever feedstock characteristics change (due to possible upsets upstream or changing to a different-grade raw material), variations from setpoint will be realized in the MD. Other reasons for upset include: changes in stock consistency, changes in stock flow, changes in MD process speed, and/or changes in the dryer rolls steam pressure. Control action taken to combat this MD offset usually requires manipulation of the amount of undiluted pulp entering the mixing chest. This action aids in delivering a consistent suspension to the headbox. When the MD offset involves water content, usually the amount of steam sent to the dryers is varied.

In general, CD disturbances occur more within the bounds of the paper machine itself. The major type of disturbance in the CD involves changes in the flow patterns inside the headbox. A nonuniform flow in the headbox would result in a nonuniform distribution of the suspension across the slice lip. To combat such disturbances, changes in the motorized actuator settings could be made to either allow more or less of the water/pulp mixture to proceed onto the fourdinier wire. Other reasons for CD offset include changes in wire condition (temperature, aging), changes in the machine speed, and/or any temperature gradients present across the CD of a cylindrical dryer.

Since making paper is an extremely old art, many of the first industrial-sized paper mills (pre-1950) had no concern for precise control of their products due to the lack of developed technology. The first-generation controller was designed in the 1950s (Davies et al., 1993). It was a relatively simple scheme involving only one sensor that was located at a fixed position. This sensor was not capable of detecting CD offsets. The controller was tied to only the inlet valves associated with the fresh water supply, pulp supply, and steam supply to combat offset.

The second-generation paper machine controller was developed in the 1970s. Instead of having only one sensor at a fixed position regulating only MD disturbances, a roving sensor scheme was used. Now, a zig-zag-type pattern is traced by letting the sensor swing back and forth across the width of the paper. The period of this zig-zag pattern is dependent on the MD velocity of the machine as well as the CD velocity of the sensor. While the MD velocity is dependent on the desired process output quantities (processing rates), the CD velocity is dependent on the ability of the control system to process large amounts of data. Currently, on a typical paper machine there are up to 400 measurements taken in one sweep of the sensor, which takes about 40 seconds to complete (Davies et al., 1993). Hence, one shortcoming of this controller is apparent: 80 seconds may elapse between sensor sweeps of the sides of the paper. This translates into many thousands of feed of paper that have proceeded through the quality-control area unscanned.

The current sensor technology employs a scanning sensor

with state estimation to recover the unmeasured variables (Campbell and Rawlings, 1994; Tyler and Morari, 1995). However, these strategies face difficulty in the presence of noise and model mismatch. This problem does not arise when full sensors, which is an array of stationary sensors, are used. We are looking forward to the day when widespread application of full sensors will be affordable. Indeed, there are vendors who currently offer gauges that provide 100% measurement information across the width of the paper. These gauges typically consist of an array of sensors that span the width of the area of interest. We present a third-generation control technology that is based on full sensor information. In order to be able to effectively utilize this on-line information, we must develop algorithms that can compute an optimal control move to a large control problem in real time. Indeed, our results indicate that we can achieve this. The controller uses a linear model that is derived in the next section.

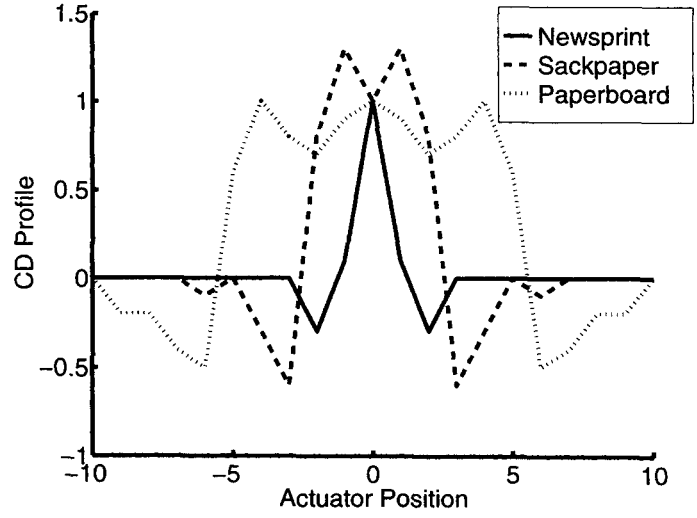

Figure 4. Actuator interaction plots.

Process model

The first step in deriving a model for the paper machine was to set the physical boundary of the process in question. The area under consideration began at the slice lip opening (the inlet of the system) and terminated at the quality-control-sensor area (the outlet of the system). The time constant associated with the actuator dynamics is very small, therefore, it is a reasonable assumption to ignore actuator dynamics. The dead time, the time required for the paper to travel from the headbox to the scanning sensor area, is about 4–6 seconds. The sampling time is also in this range. Therefore, we assume that the sampling time is greater than the sum of dead time plus any process dynamics. Under these assumptions a simple model of the paper machine is

$$y(k) = Pu(k-1), \tag{36}$$

where, $y(k)$ is the output in time period k, $u(k-1)$ is the input at time period $k-1$, and P is the interaction matrix. The formulation of the control problem as an LP that was presented in the preceding section is based on an impulse-response model. It is straightforward to show that the corresponding impulse-response model is given by

$$\Delta y(k) = P\Delta u(k-1). \tag{37}$$

The gain matrix, P, is typically obtained by performing step tests on the paper machine. This test is performed by giving a step input to a single actuator located at some position toward the middle of the slice lip. The cross-direction profile of the desired property is measured. If we assume that all the actuators behave the same and that there are no edge effects, then the entire gain matrix can be computed from a single-step experiment. The step profiles for several grades of paper have been reported in the literature (Laughlin, 1988). Figure 4 is a plot of the data in Table 3 for three grades of paper, namely, paperboard, newsprint, and sackpaper (Laughlin, 1988). The reported data are for an equal number of actuators and sensors and for a fixed interactuator spacing. However, in practice, there are more sensors than actuators, leading to a nonsquare interaction matrix. A nonsquare interaction matrix based on the reported profiles can be derived by assuming that the actuators and sensors are uniformly placed across the width of the paper machine. The result that we obtain can be generalized for a nonuniform placement of sensors and actuators. The actuator interaction plot depends on the difference between the location of the sensor and the actuator and on the interactuator spacing. Figure 5 shows m actuators and n sensors. The position across the width of the paper machine of actuator number i is $[W/(m+1)]i$ and the position of jth sensor is $[W/(n+1)]j$. Therefore, the difference in their positions is

$$\Delta_{ij} = W\left(\frac{i}{m+1} - \frac{j}{n+1}\right).$$

The corresponding coefficient of the interaction matrix is given by $p_{ij} = s \times f(\Delta_{ij})$, where f is the basic actuator interaction plot, such as the one in Figure 4, and s is the ratio of the new interactuator spacing to that of the base interactuator spacing. In order to construct larger problems, we generated interaction data by using the same strategy as for a nonsquare problem. The use of the scale factor, s, ensures that as the number of actuators is increased, the effect of each individual actuator is diminished. A good test of the scale-up process is to observe the response of the base model and the scaled-up model to an input signal with all the inputs set to unity. Physically, this corresponds to moving all the actuators up by the same amount simultaneously, thus raising the entire actuator bar. The responses of both the models must be close.

After defining the model, a set of constraints that depicts an actual process needs to be developed. Three such con-

Table 3. Actuator Interaction Matrix Data

Paper Grade	*P* Matrix Entries									
Newsprint	1.0	0.1	−0.3	0	0	0	0	0	0	0
Sackpaper	1.0	1.3	0.8	−0.6	−0.3	0	−0.1	0	0	0
Paperboard	1.0	0.9	0.7	0.8	1.0	0.6	−0.5	−0.4	−0.2	−0.2

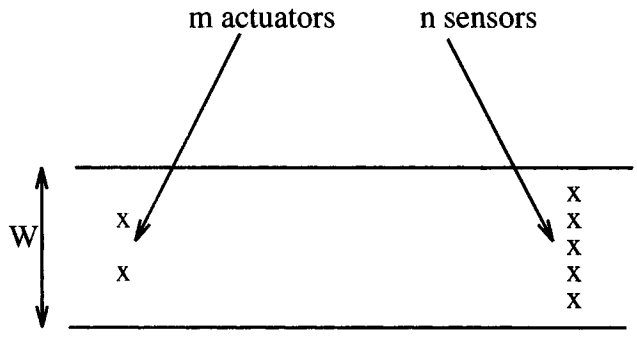

Figure 5. The nonsquare interaction problem.

straints were identified: (1) an upper bound and a lower bound on the actuator position; (2) an upper bound on the distance an actuator can travel in one time step; and (3) an upper bound on the bending moment that can be allowed on the actuator bar. These constraints are physical constraints dictated by the actual paper machine that might vary from machine to machine, depending on variables such as the material of construction, and the age of the machine. Each machine will, however, have some form of the aforementioned constraint set. On average, the slice lip opening of the headbox is taken to be 50 mm. Assuming a midpoint reference location of zero results in a height of 25 mm above this point, and 25 mm below this point. Therefore, the absolute maximum actuator height was set at +25.0 mm, and the absolute minimum at −25.0 mm. We assumed that a given actuator may travel this entire distance in one time step.

The slice lip can be modeled as a simple beam with a uniform cross section (Natarajan et al., 1989). Under this assumption, the bending moment at any point along the actuator bar is directly proportional to the second derivative of the bar at that point. Therefore, a constraint on the magnitude of the bending moment translates to a constraint on the magnitude of the second derivative of the bar. This yields

$$|u_{i+2} - 2u_{i+1} + u_i| \le \delta u_{\max} \qquad \forall\, i = 1, \ldots, n-2, \quad (38)$$

where, u_i, u_{i+1}, and u_{i+2} are the positions of the actuators located at i, $i+1$, and $i+2$, respectively; $\delta u_{\max}$ is the given upper bound; and n is the number of actuators. The value of $\delta u_{\max}$ was determined by graphical inspection of results reported by Kristinsson (1994) and was assumed to be 0.001 mm for a 6-in. actuator spacing. In order to scale the problem we must be able to generate data for new cases based upon a base case. The following relations can be derived:

$$(\delta u_{\max})_1 = (\delta u_{\max})_0 \frac{d_1^2}{d_0^2}, \quad (39)$$

where d_0 is the base actuator spacing, and d_1 is the new actuator spacing.

In the third section several formulations of the MPC problem were developed that give rise to LP problems. Minimization of worst-case errors/derivations is achieved by employing a $\infty - \infty$ norm formulation (see the subsection on $\infty - \infty$).

The LP problem for the paper machine CD control problem, when using the $\infty - \infty$ norm objective function, is given below:

$$\text{minimize} \quad \theta + \gamma \quad (40)$$

subject to:

$$\alpha^+ + \alpha^- \le \Theta \quad (41)$$

$$\alpha^+ - \alpha^- = \Gamma_y[r - y] \quad (42)$$

$$y = \tilde{y} + P(\Delta u^+ - \Delta u^-) \quad (43)$$

$$u = \tilde{u} + \Delta u^+ - \Delta u^- \quad (44)$$

$$Au \le \delta u_{\max} \quad (45)$$

$$Au \ge -\delta u_{\max} \quad (46)$$

$$\sum_{q=1}^{N_u} (\Gamma_u[\Delta u^+ + \Delta u^-])\{q\} = \gamma \quad (47)$$

bounds:

$$\alpha^+, \alpha^- \ge 0 \quad (48)$$

$$0 \le \Delta u^+, \quad \Delta u^- \le \Delta u^{\max} \quad (49)$$

$$u^{\min} \le u \le u^{\max}, \quad (50)$$

where

$$\theta, \quad \gamma \in \Re \quad (51)$$

$$\Theta \equiv (\theta, \ldots, \theta)^T \in \Re^{N_y} \quad (52)$$

$$\alpha^+, \alpha^-, y, \tilde{y}, r \in \Re^{N_y} \quad (53)$$

$$\Delta u^+, \Delta u^-, u, \tilde{u}, \Delta u^{\max}, u^{\min}, u^{\max} \in \Re^{N_u} \quad (54)$$

$$\Gamma_y \in \Re^{N_y \times N_y} \quad (55)$$

$$\Gamma_u \in \Re^{N_u \times N_u} \quad (56)$$

$$P \in \Re^{N_y \times N_u} \quad (57)$$

$$A \in \Re^{(N_u - 2) \times N_u} \quad (58)$$

$$\delta u_{\max} \in \Re^{N_u - 2}. \quad (59)$$

The first term of the objective function, θ, captures the ∞-norm of the scaled error vector, $\Gamma_y[r - y]$. The vectors α^+ and α^- have been introduced to express the objective function. The vector Θ is a vector that has every component θ. Constraint 42 relates the error between the setpoint (r) and the predicted output (y) to vectors α^+ and α^-. The setpoint (r) is a constant vector that is given. Constraint 43 relates the predicted output (y) to the measured output ($\tilde{y}$) and the control move ($\Delta u+$ and Δu^-). The matrix P is a known process model matrix that identifies the effect of a control move on the output. The next constraint 44 relates the past manipulated variable values ($\tilde{u}$) to the current manipulated variable vector (u). The vectors Δu^+ and Δu^- represent the control moves. Note that $\tilde{u}$ is known. The bending moment constraints given by Eq. 38 have been expressed concisely in Eqs. 45 and 46 by employing the matrix A, which has $N_u - 2$ rows and N_u columns. The variables α^+ and α^- are nonnegative.

Δu^+ and Δu^- have a lower bound of 0 and an upper bound of $\Delta u^{\max}$. This LP has $3N_y + 3N_u + 2$ variables and $3N_y + 3N_u - 3$ constraints. This corresponds to setting $p - l + 1$ to 1 and m to 1 to the entry corresponding to the $\infty - \infty$ norm formulation in Table 2 and adjusting the number of constraints to reflect the fact that this LP has additional constraints that limit the bending moment in the actuator bar.

Computational Results

In this section we present the results of our LP-based algorithm applied to the paper-machine CD control problem. We present results of simulations (see Table 4) performed for several disturbance rejection problems. The types of disturbance studied were sinusoidal, square-shaped, and V-shaped. The sinusoidal disturbance is given by $y_d(i) = a\sin(2\pi i/(N_y + 1))$, $\forall\ i = 1, \ldots, N_y$, where $y_d(i)$, is the ith disturbance added to the ith sensor; a is the amplitude; and i is the ith sensor. This is a sine wave that has a wavelength equal to the width of the paper machine, assuming that the sensors are uniformly placed across the paper machine. The square-shaped disturbance is given by

$$y_d(i) = \begin{cases} a & \text{if } i \geq i_m - i_s \quad \text{and} \quad i \leq i_m + i_s \\ 0 & \text{otherwise.} \end{cases} \tag{60}$$

The amplitude of the square wave is a and the midpoint is located at sensor number i_m. The width of the square spans $2i_s + 1$ sensors. The V-shaped disturbance is given by

$$y_d(i) = \begin{cases} (a/i_s) \times (i - (i_m - i_s)) & \text{if } i \geq i_m - i_s \quad \text{and} \quad i < i_m \\ a & \text{if } i = i_m \\ (a/i_s) \times (i_m + i_s - i) & \text{if } i > i_m \quad \text{and} \quad i \leq i_m + i_s \\ 0 & \text{otherwise.} \end{cases} \tag{61}$$

As in the square disturbance, the amplitude of the V-shaped disturbance is a with a width spanning $2i_s + 1$ sensors.

The normalized interaction matrix, P, was scaled by a factor of 2 so that the responses reported in the literature were duplicated (Kristinsson, 1994). Table 5 is a summary of parameters used for the simulations. We present simulation results for four different cases—a nominal case, in which the plant is equal to the model; two cases with the P matrix scaled by a mismatch factor, mf; and a severe mismatch case, in which the controller is based on a paper grade different from the grade that the plant is running. The mismatch factor, mf, is used as follows: $P_c = mf \times P_p$, where P_c is the actuator interaction matrix used in the controller, and P_p is the actuator interaction matrix of the plant. This introduces a simple multiplicative uncertainty in the gain. Physically, the output profile is multiplied by mf. Therefore, $mf > 1$ implies that the controller model P_c has entries that are greater than corresponding entries in the plant mode matrix P_p. In other words, the plant gain is smaller than the model employed by the controller. This uncertainty is relatively mild, since the shape of the profile is the same. A more severe uncertainty was also investigated in which the plant is a paperboard application, while the process model is based on sackpaper.

Figure 6 is a two-dimensional plot of a disturbance rejection experiment on a 79×79 nominal system with paperboard as the grade. The top subplot shows both the initial disturbance and the final output after the control action has been taken. The x-axis is the sensor number that is along the width of the paper, and the y-axis is the basis weight. The bottom subplot shows the actuator profile. The x-axis in this plot is the actuator along the width of the slice lip, and the y-axis shows the displacement of the actuator. As is evident from the top subplot of the figure, the disturbance is rejected almost perfectly. The constraints on the actuator bar did not

Table 4. Summary of Simulation Runs

No.	Size	Plant Grade	Controller Model	Mismatch mf	γ_u	Error
1	79×79	Paperboard	Paperboard	1.0	0	0.0210137
2	79×79	Paperboard	Paperboard	0.7	0	0.0210137
3	79×79	Paperboard	Paperboard	1.3	0	0.0210137
4	79×79	Paperboard	Sackpaper	1.0	1.0	0.0223072
5	79×119	Paperboard	Sackpaper	1.0	1.0	0.0248204

Table 5. Summary of Simulation Parameters

Paper machine width	20 ft
Base spacing	6 in.
Paper grades	Paperboard, sackpaper
Scaling factor for P	2
Base $\lvert\delta u\rvert_{\max}$	0.001 mm
Actuator lower bound, $u^{\min}$	−25 mm
Actuator upper bound, $u^{\max}$	25 mm
$\Delta u^{\max}$	50 mm
Objective function	$\infty - \infty$
Set point, r	0 vector
Control horizon (m)	1
Truncation order (N)	2
Sampling time	6 s
Simulation time	120 s
Prediction window	[2,2]
Output weight (γ_y)	10
Input weight	$\gamma_u \in [0,5]$
Sine-wave amplitude, a	0.5 units
Sine wavelength	20 ft

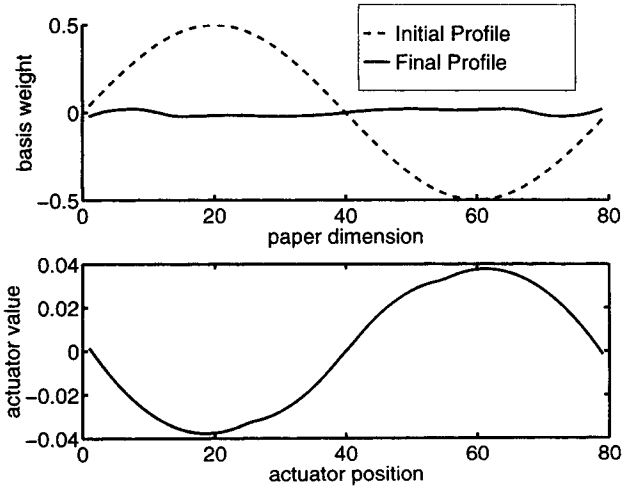

Figure 6. Input and output profiles for a nominal 79×79 disturbance rejection problem (plant = paperboard).

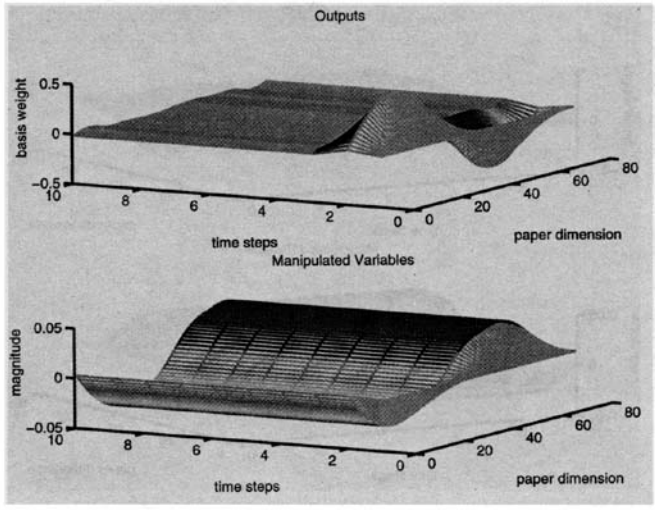

Figure 7. Time-dependent input and output profiles for a nominal 79×79 disturbance rejection problem (plant = paperboard).

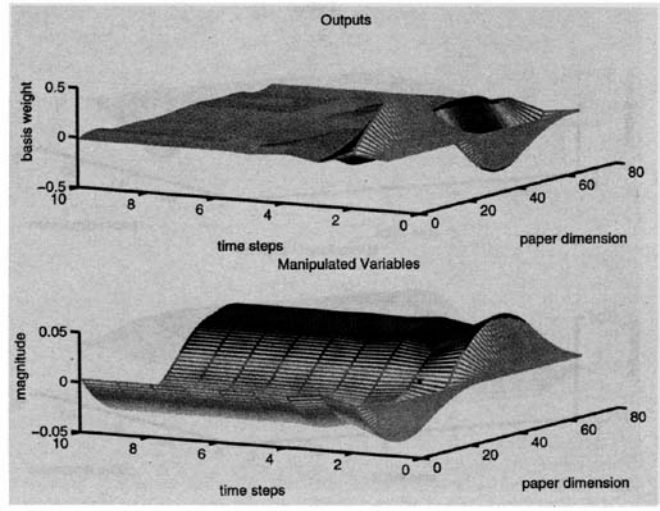

Figure 9. Time-dependent input and output profiles for a 79×79 disturbance rejection problem with 30% ($mf = 0.7$) multiplicative mismatch (plant = paperboard).

allow for the output to return exactly to the setpoint. The worst-case error at steady-state was 0.0210 units. Figure 7 is the corresponding three-dimensional plot. The top subplot in this figure shows the output as it evolves in time, while the bottom subplot shows the slice lip profile as it evolves in time.

The previous simulation involved the nominal system with no plant–model mismatch. Simulation experiments were also performed where there was a mismatch between the plant and the model. Figures 8 through 15 show the results of these experiments.

Figure 8 is a 2-D plot of a disturbance rejection experiment on a 79×79 system with paperboard as the grade and a multiplicative gain uncertainty on the model. The controller actuator interaction matrix is scaled by a factor of 0.7 relative to the plant interaction matrix. In other words, the controller model's gain is smaller than the plant gain. This leads to an oscillatory response that settles down in about three time periods. These simulations were carried out with no penalty weight on the actuator movements. This oscillatory behavior can be reduced by making the controller less aggressive. Figure 9 is the corresponding 3-D plot.

Figures 10 and 11 show the result for the same system but with a mismatch factor of 1.3 instead of 0.7. In this case, the controller model gain is actually larger than that of the plant model. Note that the disturbance is rejected in a sequence of about three actuator movements. This is due to the fact that the controller does not expect the plant to be sluggish.

The previous two simulations involved a 30% multiplicative gain uncertainty. The next simulation (Figures 12 and 13) demonstrates disturbance rejection under severe plant–model mismatch. The plant is running paperboard, as in the previ-

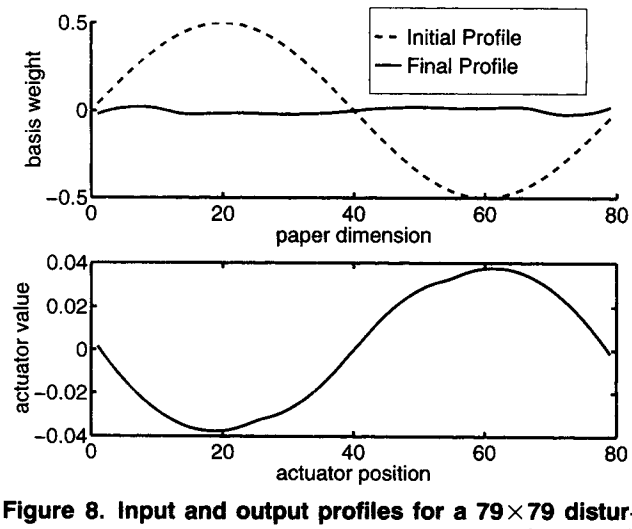

Figure 8. Input and output profiles for a 79×79 disturbance rejection problem with 30% ($mf = 0.7$) multiplicative mismatch (plant = paperboard).

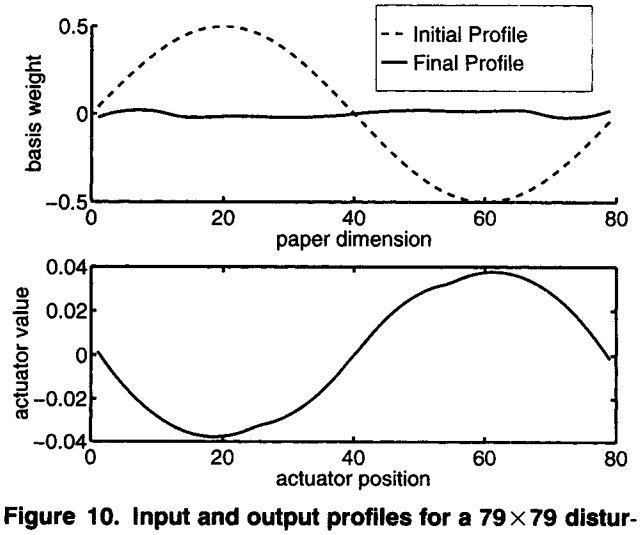

Figure 10. Input and output profiles for a 79×79 disturbance rejection problem with 30% ($mf = 1.3$) multiplicative mismatch (plant = paperboard).

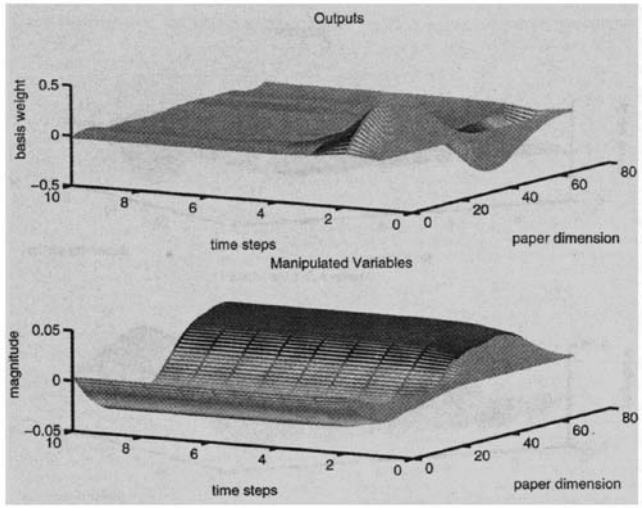

Figure 11. Time-dependent input and output profiles for a 79×79 disturbance rejection problem with 30% ($mf = 1.3$) multiplicative mismatch (plant = paperboard).

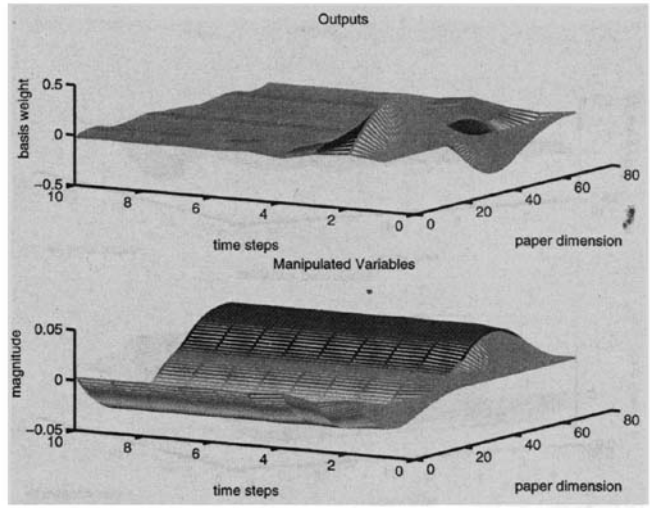

Figure 13. Time-dependent input and output profiles of a 79×79 disturbance rejection problem for severe mismatch (plant = paperboard; controller model = sackpaper).

ous simulations, while the controller uses a sackpaper model. The controller was made less aggressive by penalizing actuator movement by setting the input penalty weighting matrix, Γ_u, to I.

The simulation results just reported were for a square problem, the number of inputs (actuators) were the same as the number of outputs (sensors). Several runs of nonsquare problems were studied. Figures 14 and 15 show a representative simulation for a 79 input × 119 output control problem under severe uncertainty. The worst error at steady state was about 0.0248, which compares favorably with that achieved with a square problem (see Table 4).

The control algorithm was tested with several disturbances. The results for a nominal disturbance rejection problem are summarized in Tables 6, 7, and 8. They show a summary of the key characteristics of the LP vs. problem size. The first column in the tables is the size of the (square) control problem, the second column is the time, in seconds, required to compute a control move during the period when the disturbance is seen by the controller, the third column shows the number of phase I pivots needed to compute a feasible solution to the LP by the simplex method, and the last column gives the total number of pivots required to compute an optimal solution to the LP. Table 6 corresponds to a sinusoidal disturbance, Table 7 shows the results for a V-shaped disturbance, and Table 8 is for a square disturbance. A generic LP solver, CPLEX (1993), was used to solve the LP problems in order to compute control moves.

In an actual plant implementation, a control algorithm must satisfy real-time requirements imposed on it by the process.

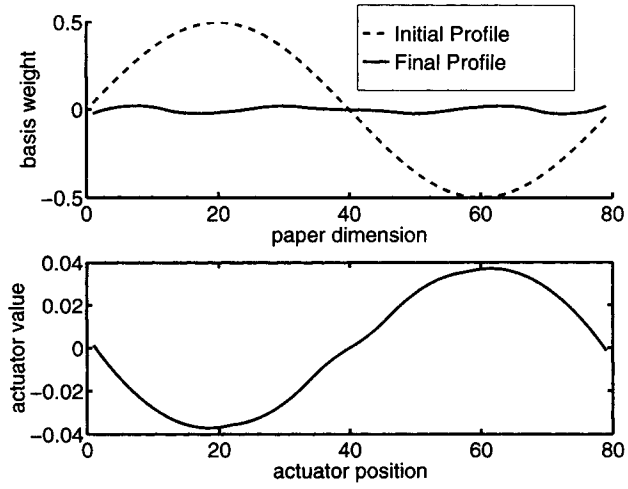

Figure 12. Input and output profiles of a 79×79 disturbance rejection problem for severe mismatch (plant = paperboard; controller model = sackpaper).

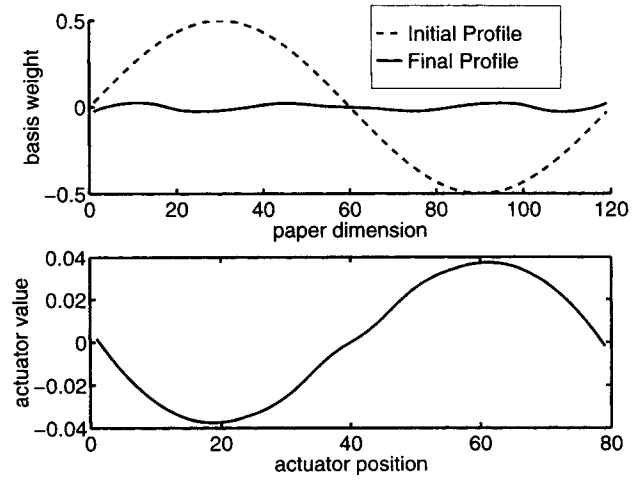

Figure 14. Input and output profiles of a 79×199 disturbance rejection problem for severe mismatch (plant = paperboard; controller model = sackpaper).

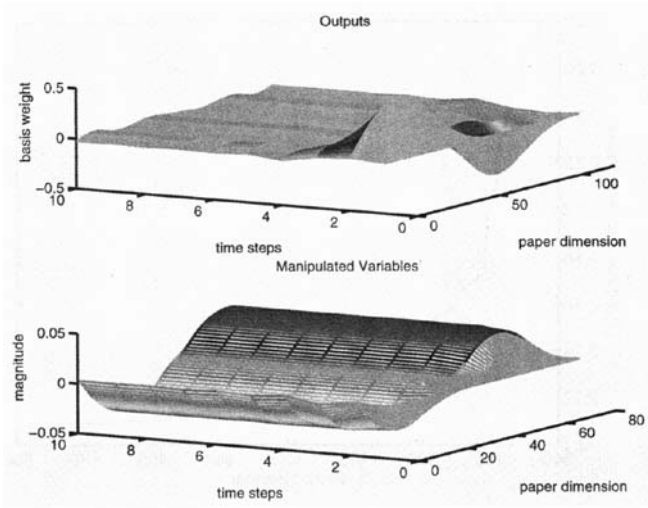

Figure 15. Time-dependent input and output profiles of a 79×119 disturbance rejection problem for severe mismatch (plant = paperboard; controller model = sackpaper).

In other words, the algorithm must be able to deliver a control move in the given time interval. It is possible that time might run out before an optimal control move can be computed. In order to deliver a control move, possibly suboptimal, it should be possible to interrupt the control move computation. Since the control move computation in our case is a solution of an LP problem, we must be able to interrupt the LP solver and ask for the best solution obtained so far. This leads us to examine the conditions under which the solution obtained so far is close to optimality. In other words, the progress of the LP solution algorithm toward optimality must be examined. Figure 16 shows a plot of the scaled infeasibility vs. pivot number for the nominal 79×79 simulation during the time interval when the disturbance hits the system. Figure 17 shows the progress toward an optimal solution during phase II of the primal simplex method. Notice that substantial progress toward an optimal solution is made once a feasible solution is found. If we had to interrupt the LP solver before it could compute an optimal solution, we still would have been able to send a good control move, provided that a feasible solution was obtained. The key issues that affect solution quality are the shape of the objective function profile and the absence of degeneracy. If the shape of the objective function profile is the same as that in Figure 17 with very few degenerate pivots, then we will almost always make rapid progress toward an optimal solution in the first few pivots of phase II of the simplex method.

Table 7. LPMPC Results for a Nominal V-Shaped Disturbance Rejection Problem

Size	Solution Time*	Simplex Pivots	
		Phase I	Total
39×39	0.17	123	185
79×79	1.06	377	469
119×119	3.79	728	881
159×159	6.86	973	1,135
199×199	17.08	1,598	2,015
239×239	23.44	1,753	2,175
279×279	37.74	2,620	3,056
319×319	50.80	2,478	3,011
359×359	66.60	2,617	3,749
399×399	76.89	2,680	3,881
439×439	121.23	3,076	5,023

*Time in seconds on HP 9000/770.

The LP-based MPC formulation has a number of attractive properties. Parametric solution of subsequent LPs in a control strategy can substantially reduce computation time. In fact, parametric solution leads to almost a 10-fold reduction in computation time. This is because the LP solved in iteration i of the MPC control strategy differs from the LP to be solved in the next iteration by a very small degree. The results that we have reported in Tables 6, 7, and 8 were obtained using a generic LP solver, CPLEX. The tables show that using the LP-based strategy, it is possible to compute an optimal control move to a 119 input × 119 output problem in about 4 s. It is important to point out that this result was obtained by using a generic implementation of the simplex method, without any customization. The LP solution process can be customized to yield even better results. Order-of-magnitude improvements in solution times can be achieved by developing customized algorithms that use specialized pivot rules and pricing schemes by exploiting problem-specific knowledge. Network simplex is the best-known implementation of this approach. The two-matching and the b-matching algorithms of Pekny and Miller (Pekny and Miller, 1994; Miller and Pekny, 1995) are further instances of customized linear programming algorithms. Other examples of cus-

Table 6. LPMPC Results for a Nominal Sinusoidal Disturbance Rejection Problem

Size	Solution Time*	Simplex Pivots	
		Phase I	Total
39×39	0.16	110	164
79×79	0.90	290	397
119×119	2.93	482	704
159×159	7.28	912	1,178
199×199	18.50	1,439	2,150
239×239	19.48	1,208	1,741
279×279	44.66	2,078	3,155
319×319	65.55	2,068	3,543
359×359	90.06	3,308	4,514
399×399	130.58	3,514	5,574
439×439	152.53	4,285	6,162

*Time in seconds on HP 9000/770.

Table 8. LPMPC Results for a Nominal Square Disturbance Rejection Problem

Size	Solution Time*	Simplex Pivots	
		Phase I	Total
39×39	0.18	122	187
79×79	1.31	403	561
119×119	4.24	839	1,043
159×159	10.35	1,281	1,624
199×199	23.15	2,185	2,629
239×239	27.04	2,093	2,630
279×279	44.24	2,991	3,470
319×319	73.89	3,789	4,514
359×359	99.56	4,882	5,360
399×399	92.06	1,916	4,389
439×439	129.62	4,213	5,366

*Time in seconds on HP 9000/770.

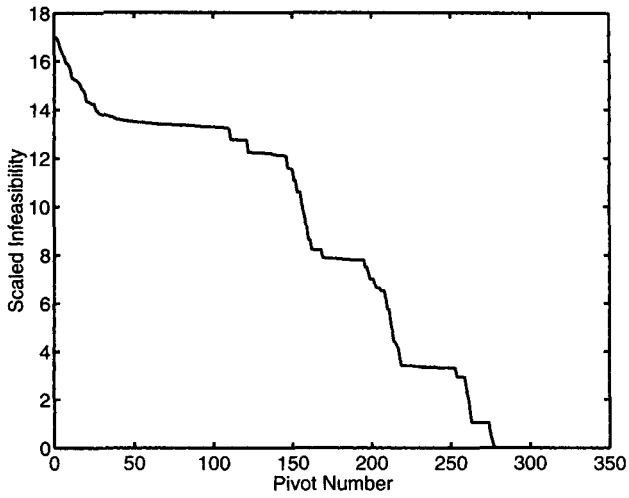

Figure 16. Phase I simplex scaled infeasibility.

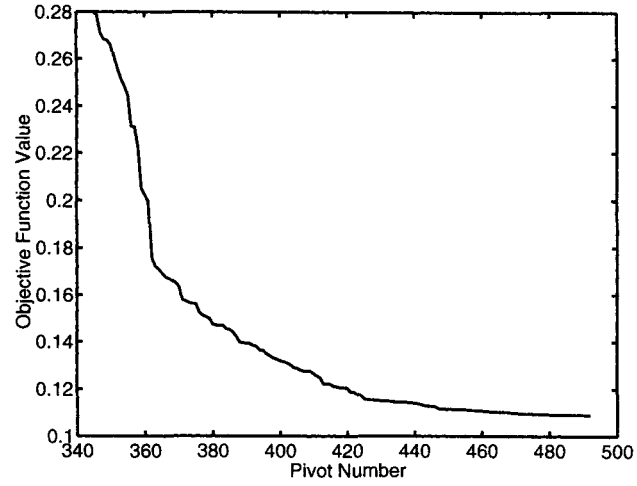

Figure 17. Phase II simplex objective function profile.

tomization include problem preprocessing, which aims at reducing the problem size by eliminating variables and constraints (Brearley et al., 1975; Pekny and Zentner, 1993) and basis initialization, which obtains a basic feasible solution to the LP very quickly using a tailored algorithm. A customized strategy that is based on the basis initialization idea for solving the LP problems that arise from the CD control problem has been devised by Dave et al. (1996). The customized algorithm consists of several stages. In the first stage, a reduced LP is formulated and solved. The solution of the reduced problem is then used to generate a feasible solution to the full problem. This feasible solution, in general, will be a nonvertex solution. In the next stage, a vertex solution is obtained that is subsequently used to initialize the basis of the full linear programming problem. This strategy has been found effective and has led to significant reduction in computational time.

Comparison with QP approaches

The LP-based model-predictive control strategy has been compared with optimal QP-based strategies (MATLAB MPC Toolbox). Tables 9, 10, and 11 show the performance of both the QP- and the LP-based approaches. The first column in the tables is the size of the control problem; the second column is the 2-norm of the disturbance; the third column is the infinity-norm of the disturbance; the fourth and fifth columns, respectively, show the 2-norm of the error vector after the QP- and LP-based controllers have taken control action; the sixth and seventh columns show the ∞-norm of the error vector after control action; the eighth column shows the time in seconds taken by the QP algorithm to compute a control move; and the last column is the time taken by the LP algorithm to compute a control move. It is evident from the tables that the QP technique outperforms the LP strategy in terms of minimizing the 2-norm of the error. This is to be expected since the QP method is geared toward minimizing the 2-norm of the error. However, the LP method outperforms the QP method in the ∞-norm metric. This is because of the fact that the LP strategy minimizes the ∞-norm of the error. Figures 18, 19, and 20 show optimal QP and LP strategies on a sinusoidal, square, and V-shaped disturbances, respectively. The top subplot shows the error vector after control action by the QP and LP techniques. The bottom subplot shows the actuator bar profile. Notice that for the LP approach the error is within a narrower band, but it is not as close to zero (as measured by the 2-norm) as the QP one is.

Besides the error metrics, another very important attribute of a control algorithm is the computational effort required to produce control moves. The time required to compute an optimal QP control move increases very rapidly with the size of

Table 9. Comparison of QP vs. LP Solution for a Nominal Sinusoidal Disturbance Rejection Problem

	Initial Profile Error		Final Profile Error $\|\cdot\|_2$		Final Profile Error $\|\cdot\|_\infty$		Solution Time*	
Size	$\|\cdot\|_2$	$\|\cdot\|_\infty$	QP	LP	QP	LP	QP	LP
39×39	2.236	0.5	0.04307	0.06762	0.021056	0.014696	4.60	0.16
79×79	3.162	0.5	0.07177	0.14425	0.033414	0.021014	57.17	0.90
119×119	3.873	0.5	0.09314	0.19842	0.037863	0.023583	266.38	2.93
159×159	4.472	0.5	0.11064	0.24287	0.040225	0.024908	758.58	7.28
199×199	5.000	0.5	0.12579	0.28087	0.041623	0.025733	1,631.05	18.50
239×239	5.477	0.5	0.13932	0.31438	0.042560	0.026301	4,087.64	19.48
279×279	5.916	0.5	0.15167	0.34623	0.043271	0.026701	7,022.06	44.66
319×319	6.325	0.5	0.16309	0.37859	0.043766	0.027008	13,917.17	65.55
359×359	6.708	0.5	0.17376	0.40343	0.044154	0.027250	21,565.31	90.06
399×399	7.071	0.5	0.18382	0.42848	0.044491	0.027441	34,087.14	130.58

*Time in seconds on HP 9000/770.

Table 10. Comparison of QP vs. LP Solution for a Nominal V-Shaped Disturbance Rejection Problem

	Initial Profile Error		Final Profile Error $\|\cdot\|_2$		Final Profile Error $\|\cdot\|_\infty$		Solution Time*	
Size	$\|\cdot\|_2$	$\|\cdot\|_\infty$	QP	LP	QP	LP	QP	LP
39×39	0.922	0.5	0.35261	0.63865	0.199022	0.149142	4.34	0.17
79×79	1.294	0.5	0.48398	0.90609	0.200864	0.149634	51.31	1.06
119×119	1.583	0.5	0.58926	1.10790	0.200492	0.149717	203.68	3.79
159×159	1.827	0.5	0.67889	1.27799	0.200836	0.149746	592.34	6.86
199×199	2.042	0.5	0.75829	1.44505	0.200597	0.149758	1,352.74	17.08
239×239	2.237	0.5	0.83014	1.58059	0.200574	0.149764	2,803.64	23.44
279×279	2.416	0.5	0.89638	1.70503	0.200623	0.149767	6,559.57	37.74
319×319	2.582	0.5	0.95803	1.82076	0.200485	0.149769	10,497.68	50.80
359×359	2.739	0.5	1.01599	1.94185	0.200631	0.149771	17,535.34	66.60
399×399	2.887	0.5	1.07081	2.03696	0.200523	0.149772	27,308.42	76.89

*Time in seconds on HP 9000/770.

Table 11. Comparison of QP vs. LP Solution for a Nominal Square Disturbance Rejection Problem

	Initial Profile Error		Final Profile Error $\|\cdot\|_2$		Final Profile Error $\|\cdot\|_\infty$		Solution Time*	
Size	$\|\cdot\|_2$	$\|\cdot\|_\infty$	QP	LP	QP	LP	QP	LP
39×39	1.658	0.5	0.57011	0.89308	0.234416	0.212395	3.93	0.18
79×79	2.291	0.5	0.82587	1.35982	0.255604	0.231622	32.64	1.31
119×119	2.784	0.5	1.01881	1.70587	0.262045	0.237844	122.23	4.24
159×159	3.202	0.5	1.18028	1.99406	0.265297	0.240919	432.29	10.35
199×199	3.571	0.5	1.32222	2.24907	0.266525	0.242753	962.29	23.15
239×239	3.905	0.5	1.45029	2.47407	0.268040	0.243971	1,664.28	27.04
279×279	4.213	0.5	1.56792	2.67893	0.268798	0.244838	2,939.33	44.24
319×319	4.500	0.5	1.67729	2.87429	0.269653	0.245487	8,343.47	73.89
359×359	4.770	0.5	1.77994	3.05179	0.270059	0.245991	9,738.63	99.56
399×399	5.025	0.5	1.87698	3.21798	0.270614	0.246394	13,458.68	92.06

*Time in seconds on HP 9000/770.

the control problem. Other researchers (Chen and Wilhelm, 1986) have also reported that it is infeasible to use optimal QP strategies on-line for large problems. Our results confirm this observation and show that this is true for all classes of disturbances that were used in this study. Based on data in Tables 9, 10, and 11, regressions were performed that fit a function of the form $k \times n^p$, where n is the size of the problem, to the timing data for both the QP and the LP methods. The order of the function, p, was found such that it minimizes the 2-norm of the error between the model values and the data points. For a given value of p, the multiplier, k, was computed as the mean value of k's for each data point. Table 12 shows the parameter values for all three classes of disturbances. This table also shows the quality of fit as measured by the error, z, between the model and the given data. Figure 21 shows the timing data for various sinusoidal distur-

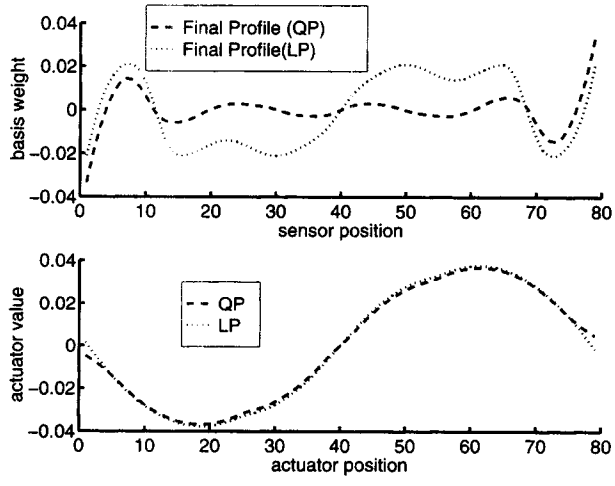

Figure 18. Comparison of QP- and LP-based strategies on a sinusoidal disturbance.

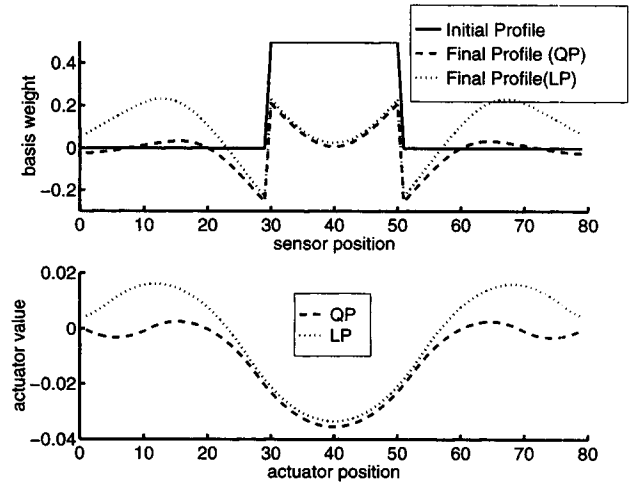

Figure 19. Comparison of QP- and LP-based strategies on a square disturbance.

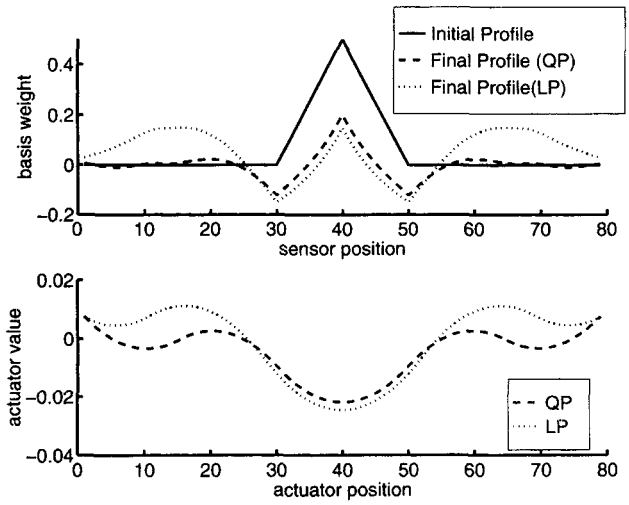

Figure 20. Comparison of QP- and LP-based strategies on a V-shaped disturbance.

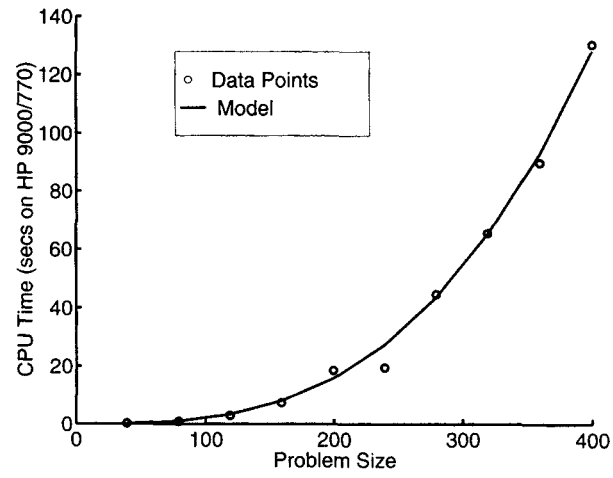

Figure 21. Performance of LP-based algorithm on a sinusoidal disturbance (model is $1.839 \times 10^{-6} n^{3.016}$).

bance rejection problems for the LP strategy, along with a model that fits the data. Figure 22 is the corresponding plot for the QP-based approach. The time required to solve the QP grows approximately as $O(n^4)$, while the time required to solve the LP grows approximately as $O(n^3)$ for the sinusoidal case. For the V-shaped disturbance the QP growth rate is still about $O(n^4)$, while the LP growth rate drops to $O(n^{2.5})$.

The LP-based strategy minimizes the worst-case deviations, while the QP approach minimizes the sum of the square of the deviations. For the CD problem, minimization of worst-case deviations is the preferred objective. Another important issue is that of computing a solution in real time. QP-based approaches require orders of magnitude more CPU time than LP-based strategies for problems of industrial scale. Therefore, the use of an optimal QP strategy on-line is not feasible.

Conclusions

We have applied the LP MPC control strategy to an important control problem from the paper industry, namely, the paper-machine cross-direction (CD) control problem. The objective of the paper-machine CD control is to maintain a uniform profile of a given property (for example, the basis weight of the paper) across the width of the paper. The current trend in the paper industry is to have larger and faster paper machines. Consequently, it is not unrealistic to have as many as 200 actuators (inputs) and 400 sensor measurements (outputs). This large size, coupled with the stringent real-time requirement of computing a control move in a few seconds, poses a very challenging control problem. We have successfully applied our strategy to this problem, and currently it can deliver an optimal control move to a 119 input × 119 output problem in about 4 s. Note that while we have assumed full-profile measurement for this problem, it is possible to use a Kalman filter with a scanning sensor to generate another large-scale problem that can be solved with this technique.

The LP-based MPC strategy presented in this article is also applicable to new headbox technologies, such as the one developed by Beloit Corporation (1996a). Instead of using slice lip bending for achieving control, the new headbox technology employs low-consistency white-water streams to achieve CD control (Beloit Corporation, 1996b). A vast majority of paper machines today do not employ this technology. Therefore, slice lip control is still an important problem. However, for paper machines employing this new headbox technology, the technique presented in this article is still applicable. Instead of the slice lip profile being used as manipulated variables, the lean water flow rate can be used. There will be no bending moment constraints associated with the new head-

Table 12. Parameter Values for the Methods that Fit the Timing Data for QP- and LP-Based Strategies

	QP			LP		
Disturbance	k	p	z	k	p	z
Sine	$1.419e-06$	3.989	1,374.01	$1.839e-06$	3.016	9.327
V-Shaped	$1.398e-06$	3.954	1,128.83	$1.865e-05$	2.555	9.878
Square	$3.420e-06$	3.695	2,495.06	$2.455e-05$	2.549	28.500

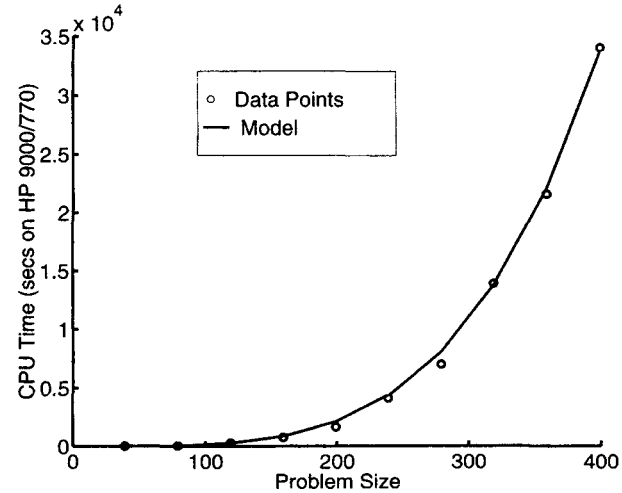

Figure 22. Performance of QP-based algorithm on a sinusoidal disturbance (model is $1.419 \times 10^{-6} n^{3.989}$).

box; however, there will be first-order constraints associated with the flow through the tubes. The LP-based control strategy presented in this article will still be applicable.

We have demonstrated that this technique is capable of solving large-scale control problems. The controller has been experimentally found to be robust in the face of uncertainty, both mild and severe. Based on an analysis of the QP- vs. LP-based approaches one can conclude that it is currently not feasible to use optimal QP strategies on-line for large-scale problems.

Notation

D = dimension of a generic vector
e_i = vector of output deviation (error) from setpoint in ith time period
f_2 = penalty of Δu
G = collection of impulse response matrices
i, j, k = indices
N = truncation order of t
N_u = number of input variables
p_{ij} = element i, j of P
r = vector representing output setpoint
$u^{\max}$ = upper bound on u
$u^{\min}$ = lower bound on u
v = generic vector in $\Re^D$
W = width of the paper machine
y = vector of output variables
y^L = lower bound on y
y^U = upper bound on y
y_d = vector of output disturbance
α_{ji} = absolute value of the jth component of e_i
$\alpha_i = \|e_i\|_1$
γ = formulation variable
θ = objective function component
Θ = vector in $\Re^{N_y}$ having all its components θ
Δu = vector of manipulated variable moves
$|\delta u|_{\max}$ = maximum allowable curvature of the actuator bar

Literature Cited

Allwright, J. C., and G. C. Papavasiliou, "On Linear Programming and Robust Model-Predictive Control Using Impulse-Responses," *Syst. Control Lett.*, **18**, 159 (1992).

Beloit Corporation, Home Page (1996a); URL: http://www.beloit.com.

Beloit Corporation, "New Headbox Signals the Future of Papermaking," (1996b); URL: http://www.beloit.com/Products/Paper_machine/Headbox/concept4.html.

Bergh, L. G., and J. F. MacGregor, "Spatial Control of Sheet and Film Forming Processes," *Can. J. Chem. Eng.*, **65**, 148 (1987).

Boyle, T. J., "Control of Cross-Direction Variations in Web Forming Machines," *Can. J. Chem. Eng.*, **55**, 457 (1977).

Boyle, T. J., "Practical Algorithms for Cross-Direction Control," *TAPPI J.*, **61**(1), 77 (1978).

Braatz, R. D., M. L. Tyler, M. Morari, F. R. Pranchk, and L. Sartor, "Identification and Cross Directional Control of Coating Processes," *AIChE J.*, **38**, 1329 (1992).

Brearley, A. L., G. Mitra, and H. P. Williams, "Analysis of Mathematical Programming Problems Prior to Applying the Simplex Algorithm," *Math. Program.*, **8**, 54 (1975).

Campbell, J. C., and J. B. Rawlings, "Estimation and Control of Sheet and Film Forming Processes," SIAM Symp. Control Problems in Industry, San Diego (1994).

Campo, P. J., "Studies in Robust Control of Systems Subject to Constraints," PhD Thesis, California Institute of Technology, Pasadena (1989).

Campo, P. J., and M. Morari, "∞-Norm Formulation of Model Predictive Control Problems," *Proc. American Control Conf.*, Seattle, WA, p. 339 (1986).

Chen, S. C., and R. G. Wilhelm, "Optimal Control of Cross-Machine Direction Web Profile with Constraints on the Control Effort," *Proc. American Control Conf.*, Seattle, WA, p. 1409 (1986).

CPLEX, *Using the CPLEX Callable Library and CPLEX Mixed Integer Library, Version 2.1*, CPLEX Optimization Inc., Incline Village, NV (1993).

Dave, P., J. F. Pekny, and F. J. Doyle, "Customization Strategies for the Solution of Linear Programming Problems Arising from Large Scale Model Predictive Control of a Paper Machine," *Comput. Chem. Eng.*, submitted (1996).

Davies, M., G. Dumont, and X. Wang, "Estimation in Paper Machine Control," *IEEE Trans. Contr. Syst.*, **CS-13**(4), 34 (1993).

Dumont, G. A., "Challenges and Opportunities in Pulp and Paper Process Control," *Proc. American Control Conf.*, Atlanta, GA, p. 1959 (1988).

Garcia, C. E., D. M. Prett, and M. Morari, "Model Predictive Control: Theory and Practice—A Survey," *Automatica*, **25**(3), 335 (1989).

Kristinsson, K., "Cross Directional Control of Basis Weight on Paper Machines Using Gram Polynomials," PhD Thesis, Univ. of British Columbia, Vancouver (1994).

Laughlin, D., "Control System Design for Robust Performance Despite Model Parameter Uncertainties: Application to Cross-directional Response Control in Paper Manufacturing," PhD Thesis, California Institute of Technology, Pasadena (1988).

Laughlin, D. L., K. G. Jordan, and M. Morari, "Internal Model Control and Process Uncertainty: Mapping Uncertainty Regions for SISO Controller Design," *Int. J. Control*, **44**(6), 1675 (1986).

Laughlin, D. L., and M. Morari, "Robust Performance of Cross-Directional Basis-Weight Control in Paper Manufacturing," *Proc. American Control Conf.*, Vol. 3, Pittsburgh, p. 2122 (1989).

Miller, D. L., and J. F. Pekny, "A Staged Primal-Dual Algorithm for Perfect b-Matching with Edge Capacities," *ORSA J. Comput.*, **7**(3), 298 (1995).

Natarajan, K., G. A. Dumont, and M. S. Davies, "On the Approximation of a Deflection Curve by Force Actuators with Variable Spacing," *Trans. ASME, J. Appl. Mech.*, **56**, 865 (1989).

Pekny, J. F., and D. L. Miller, "A Staged Primal-Dual Algorithm for Finding a Minimum Cost Perfect Two-Matching in an Undirected Graph," *ORSA J. Comput.*, **6**(1), 68 (1994).

Pekny, J. F., and M. G. Zentner, "Learning to Solve Process Scheduling Problems: The Role of Rigorous Knowledge Acquisition Frameworks," *Proc. Int. Conf. on the Foundations of Computer Aided Process Operations FOCAPO II* (1993).

Tyler, M. L., and M. Morari, "Estimation of Cross-Directional Properties: Scanning Versus Stationary Sensors," *AIChE J.*, **41**(4), 846 (1995).

Zheng, Z. Q., and M. Morari, "Robust Stability of Constrained Model Predictive Control," *Proc. American Control Conf.*, San Francisco, p. 379 (1993).

Manuscript received Feb. 26, 1996, and revision received Nov. 6, 1996.