Direct adaptive predictive control using subspace identification in Laguerre domain in the presence of constraints

Mihai Huzmezan*and Guy A. Dumont†
University of British Columbia, Pulp and Paper Centre
2385 East Mall, Vancouver, Canada, V6T 1Z4

Abstract

The classic way to control a system, in a model based framework, is to obtain a model of the system and then to use it for the design of a controller. For the class of systems characterized by a large number of inputs and outputs, such as for the cross direction control of a paper machine, we require a reduced computational time to produce the controller parameters. Our solution to this problem is a direct adaptive predictive controller which operates in the Laguerre shift operator domain and replaces the system identification step together with the calculation of the predictive controller parameters (controller that additionally contains input and output constraints) by: 1) a least squares solution, 2) two simple linear algebra operations (QR decomposition and a singular value decomposition) of a matrix constructed from input and output measurements of the unknown system and 3) a quadratic program optimization or another least squares problem. The modeling step is accomplished in a subspace identification fashion. The resulting algorithm provides major computational savings due to the reduced dimension of the system matrices together with the absence of a specific state space model.

Keywords: Direct Adaptive Control, Constrained Model Based Predictive Control ($MBPC$), Laguerre and Subspace Identification

1. Introduction

This paper blends in an original manner three different techniques:

- Laguerre orthonormal function modeling,
- subspace identification (the basics of it, without generating the state space model) and
- constrained model based predictive control.

Although each of these techniques in itself is not new, combining them, as is proposed here, is novel.

The subspace identification was addressed in works like [12, 8]. For the Laguerre identification, the work presented in [2, 11] represents a good source of information. Links between the Laguerre identification and predictive control were initially reported in [15, 14] or for a newer generation of such controllers the reader can consult [10, 9]. The link between subspace identification and the unconstrained predictive controller was reported in [13]. The concept of discrete subspace identification in the Laguerre shift operator domain was investigated partly in [4], for an approach slightly different to ours.

The class of systems that we address with this control methodology has a large number of inputs and outputs. An example of such systems is the cross direction control of a paper machine [7]. Such systems are frequent in industries that process a moving sheet of material where a profile is shaped through an array of actuators. The main challenges of such systems are the handling of a very large number of variables and the difficulty to produce a model based on first principles. The first problem drove us to compress the large number of inputs and outputs to a more manageable value. The second problem implies that an identification algorithm has to be employed for the modeling aspect.

Unfortunately, the amount of reduction produced by the compression of the input and output spaces is not enough to produce a real time implementation, yet. As a result, a direct adaptive controller is required to reduce even further the amount of time required to build an explicit prediction model. The classic way to control a system, in a model based framework, is to obtain a model of the system and then to use it for the design of a controller. Based on the work of [13] we can replace the system identification step and the calculation of the predictive controller parameters (controller that now additionally contains input and output constraints) by two simple linear algebra operations (QR decomposition and a singular value decomposition SVD) of a matrix constructed from input and output measurements of the unknown system. Therefore, generating a direct adaptive predictive controller which operates in the Laguerre shift operator results in major computational savings due to the reduced dimension of the system matrices together with the absence of a specific, say state space, model.

The paper outline was generated with the aim of giving a good overview of the techniques involved. This will

*e-mail: huzmezan@ppc.ubc.ca
†e-mail: guyd@ppc.ubc.ca

enable the understanding of the link achieved between them. The second section deals with the compression strategy of the input and output vectors. Section 3 shows how much from the conventional subspace algorithm is employed (*i.e.* just the QR and SVD decompositions) to construct a typical predictive controller cost function. Section 4 gives an insight into how a constrained predictive controller that deals with measured and unmeasured disturbances is structured. Conclusions follow explaining the difficulties of applying this direct adaptive control strategy.

2. The Laguerre shift operator domain

The motivation behind using a different domain for systems that have a large number of actuators (for instance up to 350 for a 10 m wide paper sheet) is addressed briefly in the introduction.

The natural question would be "why use an orthonormal basis". Several reasons can be offered:

- Due to the orthogonality of the polynomials we can model a profile with a reduced basis of functions.
- A reduced number of members of this basis functions is efficient at compressing and filtering the data.
- Limitations on adjacent actuators (which are usually imposed by their design) are conventionally dealt using cumbersome schemes. In our case they are dealt with implicitly since they represent limits on the parameters of the orthonormal polynomials.
- Since the number of parameters to control is reduced, this amounts to the control of the low frequency components of the profile. In other words, the frequency content of signals used within the actuator and the sensor data can be reduced.

There are some problems associated with this way of compressing data which will be addressed later within this section.

Various basis functions can be used in the compression process among which we can enumerate: Laguerre, Gram, Chebyshev and spline polynomials or even wavelets. All of them got pros and cons when it comes to a practical implementation. One possible solution is provided by Laquerre polynomials. In this paper we use this example due to the author's experience and acceptability achieved in industry by such a basis. A possible disadvantage of the Laguerre basis comes from its causality. This implies that the spectral coefficients will depend upon which direction we process the data. Although this does not fundamentally affects our work it does require further investigation.

Dumont *et. al.*[3] considered system identification based on Laguerre orthonormal functions. This method proved its simplicity when dealing with the representation of transient signals, closely resembling the Pade approximation for systems exhibiting dead time. In this paper the meaning given to the Laquerre orthonormal series representation is slightly different since it will be used for the approximation of a function such as an actuator profile.

The Laguerre function, a complete orthonormal set in $\mathcal{L}_2$, has the following Laplace domain representation:

$$L_i(s) = \sqrt{2p}\frac{(s-p)^{i-1}}{(s+p)^i} \quad , \quad i = 1, \ldots, N \qquad (1)$$

where i is the number of Laguerre filters ($i = 1, N$), $p > 0$ is the time scale, and $L_i(x)$ are the Laguerre polynomials. The reason for using the Laplace domain is the simplicity of representing the Laguerre ladder network.

This network can be expressed as a stable, observable and controllable state space form, see [14], as:

$$l(j+1) = Al(j) + bu(j) \qquad (2)$$

with $l(j) = \begin{bmatrix} l_1(j), & \ldots & , l_N(j) \end{bmatrix}^T$ being the state of the ladder. A is a lower triangular square ($N \times N$) matrix having the following form:

$$A = \begin{bmatrix} \tau_1 & 0 & \ldots & 0 \\ \frac{-\tau_1\tau_2 - \tau_3}{T_s} & \tau_1 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ \frac{(-1)^{N-1}\tau_2^{N-2}(\tau_1\tau_2+\tau_3)}{T_s^{N-1}} & \ldots & \frac{-\tau_1\tau_2-\tau_3}{T_s} & \tau_1 \end{bmatrix} \qquad (3)$$

In a similar manner the b matrix is defined as:

$$b^T = \begin{bmatrix} \tau_4 & -(\frac{\tau_2}{T_s})\tau_4 & \ldots & -(\frac{\tau_2}{T_s})^{N-1}\tau_4 \end{bmatrix} \qquad (4)$$

In the above equations (3) and (4) $\tau_1, \ldots, \tau_4$ have the following representation:

$$\begin{aligned} \tau_1 &= e^{-pT_s} \\ \tau_2 &= T_s + \frac{2}{p}(e^{-pT_s} - 1) \\ \tau_3 &= -T_s e^{-pT_s} - \frac{2}{p}(e^{-pT_s} - 1) \\ \tau_4 &= \sqrt{2p}\frac{1-\tau_1}{p} \end{aligned}$$

where T_s is the discrete system sampling time and $p \in [0 \ldots 1]$ the Laguerre network pole. The Laguerre coefficients represent a projection of the function to be approximated onto a linear space whose basis is formed by an orthonormal set of Laguerre functions.

The above state space representation is used to generate the state history for all positions $j \in \begin{bmatrix} 1 & \ldots & J \end{bmatrix}$ (where J is the total number of actuators or measurements) based on an impulse in the "command" ($u(1) = 1, u(2) = 0, \ldots, u(J) = 0$). The following equation relates actuator position or sensor measurements to the Laguerre spectral coefficients, through the C_s matrix:

$$\begin{bmatrix} y(1) \\ y(2) \\ \dots \\ y(J) \end{bmatrix} = \begin{bmatrix} l_1(1) & l_2(1) & \dots & l_N(1) \\ l_1(2) & l_2(2) & \dots & l_N(2) \\ \dots & \dots & \dots & \dots \\ l_1(J) & l_2(J) & \dots & l_N(J) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_N \end{bmatrix} \quad (5)$$

$$\bar{y} = C_s \bar{c} \quad (6)$$

Therefore the least squares solution of the compression problem is $\bar{c} = (C_s C_s^T)^{-1} C_s^T \bar{y}$.

By dealing with orthonormal filters during "data compression" we will have a diagonal dominant matrix $(C_s^T C_s)$. Large sets of data can make this matrix diagonal. The ratio between the number of Laguerre filters and the amount of data recorded (in this case the number of actuators or sensors) has to be small since we can otherwise get a singular matrix $(C_s^T C_s)$. From the perspective of a real time implementation, the matrix $(C_s^T C_s)^{-1}$ can be precomputed as it depends only on the Laguerre network realization (*i.e.* its impulse response) in the particular space in which we are operating.

One of the challenges is the choice of the Laguerre network pole which in fact determines the spatial scale. It is expected that the designer can make an *a priori* choice (a pole closer to 1 models a fairly smooth profile of the actuator positions or sensor measurements *versus* a pole closer to 0 which models a rather rough profile). We recommend a fixed pole for the entire control/identification procedure as the Laguerre network representation can be computed in advance. Optimization of the pole is beyond the scope of this paper. More about the optimal Laguerre pole choice can be found in [14, 5].

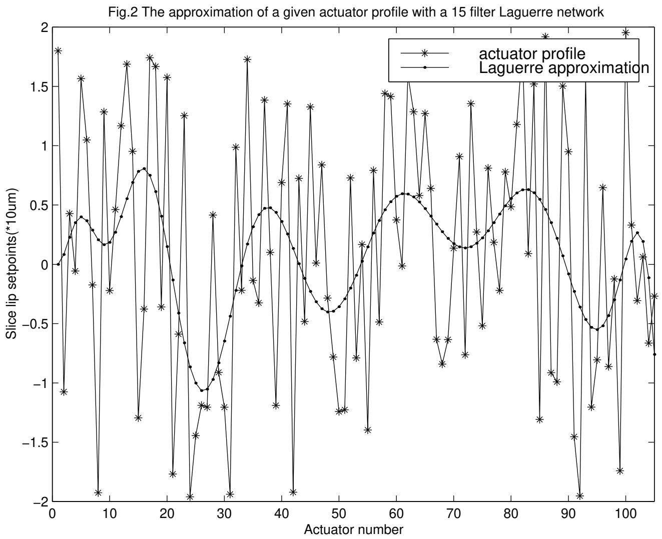

Figure 1: The approximation of a given profile with a 15 filters Laguerre network

Now having the entire compression framework clear we can see how the measured position of the actuators, which in the case of a paper machine flex the slice lip using the adjacent actuators as fulcrum, influences the basis weight of the paper at several locations around the actuator. Figure 1 shows an example of the approximation of a given actuator profile with a Laguerre network of a 15^{th} order. See [7] for a good overview and additional references are provided on paper cross directional control.

To simplify the controller design, the number of Laguerre filters used to model the actuator profile is an integer sub-multiple of the one modeling the measured profile. As a consequence, off-diagonal terms of the corresponding transfer matrix will be small, therefore providing a good condition number for the process transfer matrix.

At this stage we have determined the spectral coefficients of the Laguerre networks that characterize the profiles defined by the input and output vectors. In the following sections we denote by $u(k)$ and $y(k)$ the "input" and output, respectively, the spectral coefficients determined at time k.

3. The subspace identification method

Our approach is based on a fairly new identification method that was developed in last decade [12] called subspace system identification, mainly because it deals with the projection of the row space of one matrix onto the row space of another matrix. The main motivation for this is the lack of convergence problems since the method is non-iterative (batch type of identification) and is numerically robust since simple linear algebra operations such as QR or SVD decompositions are involved.

We need to introduce some of the concepts of subspace linear identification, since we are not going to go the full path (until a state space model is generated). This will lead to an economy in terms of computations.

Our goal is to obtain a dependence between the estimated (predicted) outputs and past outputs, inputs and future inputs of the following form:

$$\hat{Y}_f = L_w \begin{bmatrix} Y_p \\ U_p \end{bmatrix} + L_u U_f \quad (7)$$

where $U_p = \begin{bmatrix} u_{-\Psi+1}^T & \dots & u_{-1}^T & u_0^T \end{bmatrix}^T$ and $Y_p = \begin{bmatrix} y_{-\Psi+1}^T & \dots & y_{-1}^T & y_0^T \end{bmatrix}^T$ are the Ψ values of inputs and outputs, respectively, that are situated in a fixed size moving window that extends from the present into the past. Along the same lines $U_f = \begin{bmatrix} u_0^T & u_1^T & \dots & u_{N_2-1}^T \end{bmatrix}^T$ and $Y_f = \begin{bmatrix} y_0^T & y_1^T & \dots & y_{N_2-1}^T \end{bmatrix}^T$ are the N_2 values of inputs and outputs, respectively that are situated in a fixed size moving window that extends from the present into the future. Therefore $Y_f \in \mathbb{R}^{N_2 \times p}$, $U_p \in \mathbb{R}^{\Psi \times m}$ and $Y_p \in \mathbb{R}^{\Psi \times p}$, where m, n and p are the system inputs, states and outputs, respectively.

Looking at equation (7), we see that the matrices L_w and L_u, which are defining the linear predictor, are required. Coming back to the starting point of the subspace algorithms we can state the following dependencies:

$$\mathbf{Y_f} = \Gamma_\Upsilon \mathbf{X_f} + H_\Upsilon \mathbf{U_f} + H^s_\Upsilon \mathbf{E_f} \tag{8}$$

$$\mathbf{Y_p} = \Gamma_\Theta \mathbf{X_p} + H_\Theta \mathbf{U_p} + H^s_\Theta \mathbf{E_p} \tag{9}$$

where $\Psi = \alpha + \Upsilon + \Theta - 1$ the number of available measurements in the past split in two interlaced sets, one used for prediction and one for validation. This implies that a sufficient amount of data has to be acquired, prior to any identification process taking place, such as to meet the identifiability conditions. The available measurements of inputs and outputs (u_k, y_k for $k \in \{1, \ldots, \alpha + \Upsilon + \Theta\}$) that are assumed to be generated by the following state space form:

$$x_{k+1} = Ax_k + Bu_k + Ke_k \tag{10}$$

$$y_k = Cx_k + Du_k + e_k \tag{11}$$

and are first organized in the following block Hankel matrices:

$$\mathbf{U_p} = \begin{bmatrix} u_1 & u_2 & \ldots & u_\alpha \\ u_2 & u_3 & \ldots & u_{\alpha+1} \\ \ldots & \ldots & \ldots & \ldots \\ u_\Upsilon & u_{\Upsilon+1} & \ldots & u_{\alpha+\Upsilon-1} \end{bmatrix} \tag{12}$$

$$\mathbf{U_f} = \begin{bmatrix} u_{\Upsilon+1} & u_{\Upsilon+2} & \ldots & u_{\Upsilon+\alpha} \\ u_{\Upsilon+2} & u_{\Upsilon+3} & \ldots & u_{\Upsilon+\alpha+1} \\ \ldots & \ldots & \ldots & \ldots \\ u_{\Upsilon+\Theta} & u_{\Upsilon+\Theta+1} & \ldots & u_{\Upsilon+\Theta+\alpha-1} \end{bmatrix} \tag{13}$$

Note that $u_k \in \mathbb{R}^m$, $x_k \in \mathbb{R}^n$ and $y_k \in \mathbb{R}^p$ are the system inputs, states and outputs, respectively. The noise sequence $e_k \in \mathbb{R}^p$ is supposed to be zero mean Gaussian with variance $E[e_p e_q^T] = S\sigma_{pq}$. The number of columns α in the above block Hankel matrices is typically much larger (say 100 times) than the number of block rows Υ or Θ. In fact the choice of α is connected with the MBPC prediction horizon. The block Hankel matrices $\mathbf{Y_p}$, $\mathbf{Y_f}$ and $\mathbf{E_p}$, $\mathbf{E_f}$ containing the outputs y_k and the measurement noise e_k, respectively are defined in a similar manner as in equations (12) and (13).

The past and future state sequences are defined as $\mathbf{X_p} = \begin{bmatrix} x_1 & x_2 & \ldots & x_\alpha \end{bmatrix}$ and $\mathbf{X_f} = \begin{bmatrix} x_{\Upsilon+1} & x_{\Upsilon+2} & \ldots & x_{\Upsilon+\alpha} \end{bmatrix}$, respectively.

The matrices Γ_Θ, Γ_Υ have the form of the extended observability matrix, H_Θ, H_Υ, H^s_Θ, H^s_Υ being block Toeplitz matrices containing the system's impulse response to the deterministic input u_k and stochastic noise e_k. For the general case when $q \in \mathbb{N}_0$ they are defined as follows:

$$\Gamma_q = \begin{bmatrix} C \\ CA \\ \ldots \\ CA^{q-1} \end{bmatrix} \tag{14}$$

$$H_q = \begin{bmatrix} D & 0 & \ldots & 0 \\ CB & D & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ CA^{q-2}B & CA^{q-3}B & \ldots & D \end{bmatrix} \tag{15}$$

$$H^s_q = \begin{bmatrix} I & 0 & \ldots & 0 \\ CK & I & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ CA^{q-2}K & CA^{q-3}K & \ldots & D \end{bmatrix} \tag{16}$$

We can write the following linear predictor $\mathbf{\hat{Y}_f} = L_w \mathbf{W_p} + L_u \mathbf{U_f}$ based on the given past inputs and outputs organized as $\mathbf{W_p} = [\mathbf{Y_p^T}, \mathbf{U_p^T}]^\mathbf{T}$ and future inputs $\mathbf{U_f}$. By minimizing a least squares criterion:

$$\min_{L_w, L_u} \|\mathbf{Y_f} - \begin{bmatrix} L_w & L_u \end{bmatrix} \begin{bmatrix} \mathbf{W_p} \\ \mathbf{U_f} \end{bmatrix} \|^\mathbf{2}_\mathbf{F} \tag{17}$$

we produce the required matrices for the predictor stated in equation (7).

The standard way of solving the above least squares problem is via a QR algorithm:

$$\begin{bmatrix} \mathbf{W_p} \\ \mathbf{U_f} \\ \mathbf{Y_f} \end{bmatrix} = \begin{bmatrix} R_{11} & 0 & 0 \\ R_{21} & R_{22} & 0 \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \\ Q_3^T \end{bmatrix} \tag{18}$$

Therefore $\mathbf{\hat{Y}_f} = L \begin{bmatrix} \mathbf{W_p^T} & \mathbf{U_f^T} \end{bmatrix}^T$ where:

$$L = \begin{bmatrix} L_w & L_u \end{bmatrix} = \begin{bmatrix} R_{31} R_{32} \end{bmatrix} \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix}^+ \tag{19}$$

This decomposition leads to the solution of the problem defined in (17) since $\mathbf{\hat{Y}_f}$ is defined as the orthogonal projection of the row space of $\mathbf{Y_f}$ onto the row space of $[\mathbf{W_p^T}, \mathbf{U_f^T}]^\mathbf{T}$:

$$\mathbf{\hat{Y}_f} = \mathbf{Y_f} / \begin{bmatrix} \mathbf{W_p} \\ \mathbf{U_f} \end{bmatrix} = \mathbf{Y_f} \begin{bmatrix} \mathbf{W_p} \\ \mathbf{U_f} \end{bmatrix}^+ \begin{bmatrix} \mathbf{W_p} \\ \mathbf{U_f} \end{bmatrix} \tag{20}$$

With L determined, we know that $L_w = L(:, 1 : \Gamma(m+p))$ and $L_u = L(:, \Theta m + 1 : \mathit{end})$. The matrix L_w still contains part of the noise, which we do not want to include in our model. Therefore an SVD decomposition:

$$L_w = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \tag{21}$$

is required such as to retain (by inspection) only the system $L_w = U_1 \Sigma_1 V_1^T$ that contains the dominant singular values.

4. Direct adaptive predictive control

The concept of predictive control involves the repeated optimization of a performance objective over a finite horizon extending from a future time (N_1) up to a prediction horizon (N_2) [1, 6]. Given a set-point, a reference is produced by pre-filtering and is used within the optimisation of the MBPC cost function (22). Manipulating the control variable over the control horizon (N_u), the algorithm drives the predicted output, over the prediction horizon, towards the reference. Predictive control is used instead of a conventional passive state or output feedback control technique due to is its simplicity in handling varying time delays and non-minimum phase systems.

In terms of the algorithm the predictor employing the "internal model", based on the measurement of the past inputs and outputs, provides the optimizer with the future predicted values of the outputs. The optimizer contains the cost function (22) involving adjustable weights if necessary with the main task of computing the present control move such that the predicted output follows the reference in the desired manner.

The form of the quadratic cost function to be optimized is:

$$J_k = \sum_{l=N_1}^{N_2} \|(\hat{y}_{k+l} - r_{k+l})\|^2_{Q(l)} + \sum_{l=0}^{N_u} \|\Delta u_{k+l}\|^2_{R(l)} \tag{22}$$

The minimization is performed subject to constraints on:

- the input levels: $u_{min} \leq u_l \leq u_{max}$ where $k \leq l \leq k + N_u - 1$
- the input rates of change: $\Delta u_{min} \leq \Delta u_l \leq \Delta u_{max}$ where $k \leq l \leq k + N_u - 1$
- the output levels: $y_{min} \leq \hat{y}_l \leq y_{max}$ where $k \leq l \leq k + N_2$

In normal operation the weights $Q(l)$ and $R(l)$ are independent of k, but if necessary they may need to vary with k. The norm $\|.\|^2_Q$ within the cost function is defined as $\|x\|^2_Q = x^T Q x$. For prediction it is assumed that $\Delta u_l = 0$ for $l \geq k + N_u$. As formulated, the optimization is a quadratic programming (QP) problem, and can be solved using standard algorithms.

The equation (7) can be rewritten in ΔU_f which represents the control movement increment in the future as:

$$\hat{Y}_f = L_w \begin{bmatrix} Y_p \\ U_p \end{bmatrix} + L_u(U_0 + M\Delta U_f) \tag{23}$$

where Y_p, U_p have the same significance as in (7), $U_0 = \begin{bmatrix} u_0^T & u_0^T & \dots & u_0^T \end{bmatrix}^T$, $U_0 \in \mathbb{R}^{N_2 \times m}$ and M is a lower triangular matrix, of the appropriate size, filled with identity matrices having dimensions defined by the number of control inputs.

Based on the equations (22) and (23) we can rewrite the cost function in a vector/matrix format as:

$$\begin{aligned} J_k &= (\hat{Y}_f - R_f)^T Q (\hat{Y}_f - R_f) + \Delta U_f^T R \Delta U_f = & (24) \\ &= (L_w W_p + L_u(U_0 + M\Delta U_f) - R_f)^T Q & (25) \\ &\quad (L_w W_p + L_u(U_0 + M\Delta U_f) - R_f) + \Delta U_f^T R \Delta U_f & \end{aligned}$$

where

$$\begin{aligned} R_f &= \begin{bmatrix} r_0^T & r_1^T & \dots & r_{N_2-1}^T \end{bmatrix}^T \\ Q &= diag(\begin{bmatrix} Q_1^T & Q_2^T & \dots & Q_{N_2-N_1}^T \end{bmatrix}^T) \\ R &= diag(\begin{bmatrix} R_1^T & R_2^T & \dots & R_{N_u}^T \end{bmatrix}^T) \end{aligned}$$

The reason of including in the optimizer a model which contains the plant model augmented with an integrator is the rejection of step disturbances or the following of step references. In the same spirit and based on the "internal model principle" other signals like ramps or sinusoids can be tracked or rejected by including their appropriate model in the controller.

The part of the cost function dependent on the unknown variable ΔU_f, which we subject to the optimization procedure, has the form:

$$\begin{aligned} J_k &= \Delta U_f^T [M^T L_u^T Q L_u M + R] \Delta U_f + & (26) \\ &\quad 2[W_p^T L_w^T Q L_u M + U_0^T L_u^T Q L_u M - R_f^T Q L_u M] \Delta U_f & \\ &= \Delta U_f^T \mathcal{A} \Delta U_f + \mathcal{B} \Delta U_f & \end{aligned}$$

This cost function (26) can be augmented, if required, with the linear constraints reflecting actuator or envelope limitations. These constraints can be written in ΔU_f after several algebraic manipulations (see [6] for details):

$$\Phi \Delta U_f \leq \Xi \tag{27}$$

If no constraints are to be imposed then the optimization of the the quadratic problem has its solution via a least square algorithm (by putting the trace of the cost derivative with respect to ΔU_f to 0) and can be written as: $\Delta U_f = -\mathcal{A}^+ \mathcal{B}^T$. Otherwise a QP solver or equivalent has to be used. In both cases only the first control movement will be implemented.

The main advantage of predictive control over other control methods resides in its ability to handle constraints but equally well in the on-line optimization which can involve at each time step a different "model" (via the matrices L_w and L_u updated continuously).

The last part of the control variable computation is represented by the conversion from the reduced number of Laguerre coefficients back to the real actuator profile. Using the filter structure already in place at the time of the data compression the function characterized by those coefficients is generated and its corresponding values are fed back to the actuators.

5. Conclusions

In our opinion there is a clear benefit for specific systems which exhibit a large number of inputs and outputs, to employ a controller based on the link between: subspace identification (the basics of it, without generating the state space model), Laguerre orthonormal function modeling and constrained model based predictive control.

From the practical perspective the features included at the controller level allow us to talk about an immediate future industrial implementation. Note that the same algorithm can be successfully applied in the case of systems with a reasonable number of inputs and outputs, but in such a case the compression part becomes unnecessary.

Based on the theoretical development as well as simulations we recognize, besides the benefits, several challenges that this algorithm poses to the user:

- The Laguerre compression of noisy signals (such as the actuator profile or output shape) represents a challenge for the algorithm. It is known that the number of the Laguerre filters influences the bandwidth of the compressed signal. There is no automatic way yet to decide upon the number of elements in the series.

- The subspace identification method is a batch type of identification. One of the reasons that pushed us to use the present form of the algorithm is that by using the conversion to a state space representation we will face the problems associated with the initialization of the state used in the prediction process. More insight has to be gained with respect to this issue.

- For a constrained predictive controller the management of the solution infeasibility is a well known problem. A scheme to address this issue is mandatory. Our experience with mixed weight least square (MWLS) algorithm will recommend it as a potential alternative.

References

[1] D.W Clarke and C. Mohtadi. Properties of Generalized Predictive Control. *Automatica*, 25(6):859–875, 6 1989.

[2] G.A. Dumont and C.C. Zervos. Adaptive Controllers based on Orthonormal Series Representation. In *Proceedings of the* 2_{nd} *IFAC Workshop on Adaptive Systems in Control and Signal Processing, Lund, Sweden*, 1986.

[3] G.A. Dumont and C.C. Zervos. Adaptive Control Based on Orthonormal Series Representation. In *Proceedings of the* 2^{nd} *IFAC Workshop on Adaptive Systems*, pages 371–376, 1986.

[4] B. Fischer. *Two Topics in System Identification.* PhD thesis, Lulea University of Technology, Sweden, 1997.

[5] G.A. Dumont Y. Fu and A. Elshafei. Orthonormal functions in identificationa and adaptive control. In *IFAC international symposium, Singapore*, January 1991.

[6] M. Huzmezan. *Model Based Predictive Control Theory and Aerospace Applications of Constrained Model Based Predictive Control.* PhD thesis, University of Cambridge, 1998.

[7] K. Kristinsson and G.A. Dumont. Cross-directional Control on Paper Machines using Gram Polynomials. *Automatica*, 32(14):533–548, 1996.

[8] W. Larimore. Canonical Variate Analysis in Identification, Filtering, Adaptive Control. In *Proceedings of the* 29^{th} *Conference on Decision and Control*, pages 596–604, December 1990.

[9] M. Huzmezan, G.A. Dumont, S. Kovack and W.A. Gough. A New Generation of MIMO Indirect Adaptive Model Based Constrained Predictive Controllers Based on Laguerre Basis Function Identification. In *Submitted the* 39^{th} *IEEE Conference on Decision and Control, Sydney, Australia*, December 2000.

[10] M. Huzmezan, S. Kovach and W.A. Gough. A New Generation of Adaptive Model Based Predictive Controllers Applied in Process Control. In *IEEE APC'99 Workshop, Vancouver, Canada*, April 1999.

[11] B. Ninness and J-C. Gomez. Mimo System Identification Using Orthonormal Basis Functions. In *Proceedings of the* 34^{th} *IEEE Conference on Decision and Control, New Orleans, US*, pages 703–708, December 1990.

[12] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Appications.* Kluwer Academic Publishers, Boston/London/Dordrecht, 1996.

[13] B. De Moor W. Favoreel and M. Gevers. SPC: Subspace Predictive Control. In *Proceedings of the* 14^{th} *IFAC Congress, Beijing, China*, pages 235–240, July 1999.

[14] C.C. Zervos. *Adaptive Control Based on Orthonormal Series Represenation.* PhD thesis, University of British Columbia, 1988.

[15] C.C. Zervos and G.A. Dumont. Deterministic Adaptive Control Based on Laguerre Series Representation. *International Journal of Control*, 48(1):2333–2359, 1988.

Acknowledgements

We would like to acknowledge BC Advanced Systems Institute and Universal Dynamics Technologies for funding this research. As well we would like to acknowledge the contribution of Mr. Junqiang Fan, Ph.D. candidate, to the software used in simulations related to this work.