2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Data-based Modeling and Control of Nylon-6,6 Batch Polymerization*

Siam Aumi, Brandon Corbett, and Prashant Mhaskar†
Department of Chemical Engineering, McMaster University, Hamilton, ON., Canada L8S 4L7

***Abstract*— This work addresses the problem of modeling the complex nonlinear behavior of a nylon-6,6 batch polymerization process and subsequently tracking trajectories of the important process variables, namely the reaction medium temperature and reactor pressure, using model predictive control (MPC). To this end, a data-based multi-model approach is proposed in which local linear models are identified from previous batch data using latent variable regression and then combined using a continuous weighting function that arises from fuzzy c-means clustering. The resulting data-based model is used to formulate a trajectory tracking predictive controller. Through simulation studies, the modeling approach is shown to capture the major nonlinearities of the process, and closed-loop simulation results demonstrate the efficacy of the proposed predictive controller and its advantages over conventional proportional-integral (PI) trajectory tracking.**

I. Introduction

Batch processes are commonly used for the manufacture of high-value specialty products or as startup/intermediate steps in a continuous processing unit. As opposed to continuous reactors that are operated around a nominal steady-state, the end-point in a batch reactor is not necessarily an equilibrium point, allowing for achieving a wide range of product specifications by changing the initial conditions and input trajectories. This flexibility is particularly important in the polymerization industry as polymerization reactions often require a wide range of operating conditions to yield polymers with a desirable end-use quality, which is the primary control objective. Batch polymerization processes, however, exhibit numerous characteristics that complicate the control problem. Undoubtedly, the biggest challenge is that measurements related to the final end-use quality are unavailable during the batch and are only made (off-line) after the batch is complete. The control problem is further complicated by the presence of constraints, nonlinear, time-varying dynamics, and the absence of equilibrium conditions, which invalidates or limits the achievable performance of many existing control designs for continuous systems.

Polymerization models are identified either deterministically or empirically. In the former approach, appropriate conservation equations and models for the reaction kinetics are used to derive a state-space representation of the system consisting of coupled (integro-)differential equations. Deterministic model-based control designs, however, have limited practical applicability for batch polymerization processes; the number of system states in the deterministic model is often excessive and/or the differential equations are overly complex. Moreover, the model's predictive capabilities are subject to the accuracy of numerous model parameters. In many cases, the kinetic parameters in polymerization models are unknown or inaccurate, and their proper estimation is difficult. Furthermore, many of the simplifying assumptions taken during model development are often violated in specific situations, making the model unreliable during control calculations.

The increased availability of past process data, however, can be exploited to improve the achievable level of accuracy of polymerization models developed using simpler, empirical models. Identification experiments, such as those in which a pseudo-random binary signal (PRBS) is applied on the process are often too expensive to justify for batch polymerization processes since they result in wasted batches. Furthermore, within the range of operating conditions during a polymerization, the process behavior is highly nonlinear and characterized by stages with considerably different dynamics. These factors make conventional system identification approaches, where a single linear model is identified, ill-suited for identifying an accurate dynamic model. One general strategy to address the nonlinearity issue has been to use multi-model approaches, such as piece-wise affine (PWA) or Takagi, Sugeno, and Kang (TSK) models. In these approaches, linear models are used to capture local dynamics and combined with a weighting function to describe the process nonlinearities (see [1], [2] and the references therein for details). Recently, a new multi-model approach was proposed in [3], [4] that differentiated itself from the existing work by its use of a generalized, continuous, and entirely data-driven weighting function (see [3] for additional details). However, in [3], full state measurements were assumed. In this work, we extend this modeling methodology and demonstrate its efficacy for a process with limited available measurements.

The direct control of the end-use quality in batch polymerization reactors is often impractical because the product quality is only measured following batch completion. As a result, the control objective is typically pursued indirectly via trajectory tracking methods. In trajectory tracking methods, trajectories for a set of measurable process variables, which are related to the end-use properties, are generated off-line or re-calculated at specific time points during the batch (e.g, see [5]–[9]). These trajectories are subsequently tracked using local model-based controllers or proportional-integral-derivative (PID) controllers, possibly modified with gain scheduling [10] or feedforward [11] terms to partially account for the nonlinear batch dynamics. The control performance with PID-based local controllers is limited because they fail to account for control loop interactions, input constraints, and optimality. Explicitly nonlinear, model-based tracking

*Financial support by NSERC and McMaster Advanced Control Consortium is gratefully acknowledged.
†Corresponding author: mhaskar@mcmaster.ca

controller have been proposed in the form of differential geometric (e.g., [11], [12]) and model predictive (e.g., [13]–[15]) controllers. These designs have a strong dependence on an accurate deterministic process model or utilize a simplified linearized model (deterministic or empirical). However, an accurate and reliable deterministic model may be unavailable or even if available, may be overly complex or unreliable for use within a model-based control framework while linearized models are incapable of describing inherent process nonlinearities.

Motivated by these factors, we address the problem of empirically modeling the highly nonlinear nylon-6,6 batch polymerization process and then tracking key process variable trajectories using model predictive control (MPC). The rest of this manuscript is organized as follows. First, in Section II, we give an overview of the process and its control objective. Then, we propose the modeling approach in Section III and extract models for the process variables of interest. This is followed by the presentation of a trajectory tracking predictive controller for the process variables and a comparison of the proposed controller's performance against a classic PI controller. Finally, we summarize our results in Section IV.

II. Process Overview

In this section, we give an overview of the nylon-6,6 polymerization process and discuss the control objective and available control strategies. We focus on nylon-6,6 production by amidation of adipic acid and hexamethylenediamine (HMD) in a batch reactor. In this polymerization, the reactor is initially charged with molten adipic acid and HMD (from an evaporator) in approximately stoichiometric $(1:1)$ proportions. The reaction model is summarized by the following equations.

$$\text{Degradation} \quad \text{C} \rightarrow \text{SE} + \text{W} \tag{1}$$

$$\text{L} \rightarrow \text{SE} + \text{A} \tag{2}$$

$$\text{Polyamidation} \quad \text{A} + \text{C} \leftrightarrows \text{L} + \text{W} \tag{3}$$

where A is an amine end group, C is a carboxyl end group, W is a water molecule, L is a polymer link, and SE is a (non-reactive) stabilized end group. The polymerization reaction is treated as a second-order, reversible reaction of a-a/b-b type that is commonly described in terms of functional groups for simplicity (see [16]). During the polymerization reaction (given by (3)), the amine end groups (A) in HMD or the polymer chain react with the carboxylic end groups (C) on either the adipic acid or polymer chain, forming a polymer link (L) and water (W). The degradation reactions (1–2) are considered due to their effect on the reaction mixture temperature. In order to meet the typical desired end-use qualities, a high extent of reaction (over 99%) is required, which, in turn, calls for shifting the polymerization reaction towards completion by vaporizing water and then venting the vaporized water. Consequently, the polymerization is carried out in an autoclave reactor equipped with a steam jacket for providing the heat needed for vaporization (and reaction) and a valve for venting vaporized water.

A. Nylon-6,6 Polymerization Model

To illustrate the proposed modeling and control approach, we utilize the mathematical model of nylon-6,6 polymerization presented in [16]. The modeling assumptions (and their explanations), parameter values, and kinetic relationships are available in [16], [17] and omitted here for brevity. The final state-space model of the process consists of nine coupled ordinary differential equations (ODEs) with the state vector comprised of the molar amounts of each functional group and evaporated HMD, the reaction medium mass, temperature, and volume, and reactor pressure. The final model takes the following general form:

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) \\ \boldsymbol{y}(t) &= \mathbf{C}\boldsymbol{x} + \boldsymbol{v} \end{aligned} \tag{4}$$

where $\boldsymbol{x}$ denotes the vector of state variables, $\boldsymbol{y}$ denotes the vector of measurable process variables, which were taken as the reaction mixture temperature, T (K), and reactor pressure, P (psia), and $\boldsymbol{u}$ denotes the vector of manipulated inputs, which are the steam jacket pressure, P_j (psi), and vent rate, v (kg/h), constrained between $\boldsymbol{u}_{\min}$ and $\boldsymbol{u}_{\max}$. The vector function $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$ denotes the right hand side of the ODEs and $\mathbf{C}$ is a constant matrix since the outputs are two of the states. Thus, the output and input vectors are defined as follows: $\boldsymbol{y} = \begin{bmatrix} T & V \end{bmatrix}'$ and $\boldsymbol{u} = \begin{bmatrix} P_j & v \end{bmatrix}'$. The physical limitations in the process design imposes constraints on the available inputs; the constraints on the steam jacket pressure and vent rate are: $700 \text{ psi} \leq P_j \leq 1800 \text{ psi}$ and $0 \leq v \leq 2000$ kg/h; that is, we have: $\boldsymbol{u}_{\min} = \begin{bmatrix} 700 & 0 \end{bmatrix}'$ and $\boldsymbol{u}_{\max} = \begin{bmatrix} 1800 & 2000 \end{bmatrix}'$. The output measurements are assumed to be corrupted by zero-mean normally distributed noise that is denoted by $\boldsymbol{v}$. The standard deviations for the measurement noise for T and P were set to 0.16 K and 0.17 psia, respectively. The duration of the batch is 3 hours with a sampling period of 60 s. In generating a database of previous batches for this system, only the input and output measurements are assumed to be available.

Remark 1 One difference between the model in [16] and the one used in this work is that we do not neglect the reactor pressure dynamics. In [16], the reactor pressure is treated as an input due to the assumption of fast dynamics whereas we append the model equations with a differential equation for the pressure that is equal to the product of a (negative) gain term and the vent rate. In this way, the reactor pressure can be treated as a controlled variable that is influenced by the vent rate and the control problem becomes a multiple-input-multiple-output (MIMO) problem.

B. Control Strategies of Nylon-6,6 Autoclave

The end-use product quality of nylon-6,6 polymer is defined by the molecular weight and the residual amide concentration; however, their measurements are not available during batch operation. Instead, the product quality is monitored through easily measurable process variables, such as the reaction mixture temperature, reactor pressure, steam jacket pressure,

and vent flow rate. Thus, a common control strategy has been to track suitable trajectories of the measurable process variables obtained through off-line optimization of the process model or from historical batch data that produced the target qualities. For this work, we focus on tracking trajectories of the reaction medium temperature, T, and reactor pressure, P, by manipulating the steam jacket pressure, P_j and vent rate, v. In [16], additional industrially popular secondary variable tracking control strategies are evaluated (using PID controllers) in terms of robustness to common disturbances.

Note that even with perfect tracking during a new batch, there is no guarantee that the desired quality will be met because unavoidable disturbances encountered during the new batch effectively alter the relationship between the product quality and the trajectories of the process variables. Thus, the profiles being tracked may no longer yield the desired polymer quality, and they essentially have to be "re-optimized" in some fashion. The development of an inferential quality model, which can be used to predict the final product quality from the measurable process variables, and then the subsequent integration of the model within the control design is outside the scope of the present work but an important aspect of our future research direction. The current work assumes that an appropriate trajectory has been computed and re-optimized as necessary, and we focus on the complexities associated with the trajectory tracking problem.

III. Data-based Modeling and Predictive Control of Nylon 6,6 Batch Polymerization

In this section, we propose a modeling methodology for the nylon-6,6 batch polymerization process. After reviewing a few preliminary concepts, we propose the modeling methodology and apply it on a realistic database of past nylon-6,6 batches. Then, we formulate a trajectory tracking predictive controller for the process and present the results of closed loop simulations.

A. Preliminaries

In this section, we give an overview of auto-regressive exogenous (ARX) modeling, principal component regression (PCR), and fuzzy c-means clustering. These concepts are later unified in the proposed data-based modeling approach.

1) Auto-Regression Exogenous (ARX) Models: In ARX modeling, the process outputs at a specific sampling instance depend linearly on the previous process conditions (defined by the process outputs and inputs). Mathematically, in vector form, an ARX model for a given process output takes the form shown below.

$$y(k) = \boldsymbol{\beta}\bar{\boldsymbol{x}}(k) + v(k) \tag{5}$$

where $\boldsymbol{\beta}$ is a vector of model coefficients that is identified using linear regression and $\bar{\boldsymbol{x}}(k) = \begin{bmatrix}\boldsymbol{y}'(k-1) & \cdots & \boldsymbol{y}'(k-n_y) & \boldsymbol{u}'(k-1) & \cdots & \boldsymbol{u}'(k-n_u)\end{bmatrix}$ is a row vector of lagged concatenated outputs and inputs. The scalars, n_y and n_u, are the number of lags in the outputs and inputs (respectively). Note that we have assumed the same number of lags, n_y and n_u, for each output and input variable (respectively) and continue to do so for the remainder for notational simplicity. This assumption can be easily relaxed (i.e., n_y and n_u may be vectors). Different criteria can be used to select the "optimum" lag structure (see [18]) with the general objective being to achieve low prediction error with the minimum number of model parameters, which prevents against over-fitting and maintains model simplicity.

One way to estimate the ARX model coefficients (for each output) is with principal component regression (PCR). PCR can handle correlations/co-linearities in the plant data, which is a possibility for batch data especially when lagged variables are involved. To facilitate the regression, a response vector, $\bar{\boldsymbol{y}}$, and a regressor matrix, $\bar{\mathbf{X}}$, are first constructed corresponding to $y(k)$ and $\bar{\boldsymbol{x}}(k)$ (respectively) in (5) by sorting the plant data sample-wise. Next, principal component analysis (PCA) is performed on $\bar{\mathbf{X}}$, and the resulting transformed regressor matrix is regressed onto $\bar{\boldsymbol{y}}$ using ordinary least squares (OLS). Geometrically, in PCA, the variables in $\bar{\mathbf{X}}$ are projected onto a lower dimensional subspace defined by A orthogonal principal components or latent variables. Each principal component accounts for a certain percentage of the variance in the regressor matrix. With highly correlated data, the number of principal components required to summarize the information is much lower than the dimensionality of the original data. Mathematically, $\bar{\mathbf{X}}$ is decomposed in PCA as the sum of the outer products of a score and loading vector: $\hat{\bar{\mathbf{X}}} = \sum_{a=1}^{A} \boldsymbol{t}_a \boldsymbol{p}_a' = \mathbf{T}\mathbf{P}'$ where $\hat{\bar{\mathbf{X}}}$ is the approximation of the original regressor matrix, $\boldsymbol{t}_a$ is a score vector containing the projections of each row in $\bar{\mathbf{X}}$ on the a-th principal component, and $\boldsymbol{p}_a$, the loading vector, defines the orientation of the corresponding principal component. The score and loading matrices, $\mathbf{T}$ and $\mathbf{P}$ (respectively), contain their corresponding vectors; thus, the score matrix represents the projection of the original data set on the latent variable subspace and the loading matrix defines the subspace's orientation. The orthogonality between each principal component induces the orthonormality of $\mathbf{P}$, $\mathbf{P}'\mathbf{P} = \mathbf{I}$. A variety of algorithms, classifiable as either covariance or iterative methods (see [19] for details on the latter), can be employed for finding the loading and score matrices. If the PCA model is to be used for prediction, a suitable cross-validation method (see [20] for a review of possible options) can be used to find an acceptable number of principal components to retain. The lesser (unimportant) principal components are typically a consequence of measurement and process noise, and therefore discarded, resulting in noise reduction.

After decomposing $\bar{\mathbf{X}}$ with PCA, $\boldsymbol{\beta}$ is estimated with PCR by regressing $\mathbf{T}$ onto $\bar{\boldsymbol{y}}$ using ordinary least squares (OLS), $\hat{\boldsymbol{\beta}}_{\text{PCR}} = (\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\bar{\boldsymbol{y}}$. In short, PCR entails replacing the variables in the original regressor matrix by new ones with better properties (orthogonality) that also span the original space and then using OLS regression. The orthogonality property improves the numerical properties of the required inversion during OLS, which is important when the original variables in the regressor matrix are correlated.

2) Fuzzy c-Means Clustering: An important pre-processing step in the multi-model approach in Section III-B is to locate the operating points around which local linear models are identified. One approach for finding this set of operating points for batch systems is to partition a historical batch database into a number of clusters (i.e., a group of points in the database that are mathematically similar) using fuzzy c-means clustering. We describe this algorithm for a given output in this section.

Consider the regressor matrix, $\bar{\mathbf{X}}$, for a given process output. The matrix, $\bar{\mathbf{X}}' = \begin{bmatrix} \bar{\boldsymbol{x}}_1' & \cdots & \bar{\boldsymbol{x}}_i' & \cdots & \bar{\boldsymbol{x}}_N' \end{bmatrix}$, has N columns with each column being a different instance of $\bar{\boldsymbol{x}}'$ (see Equation 5) in the training data. The lagged output-input space in $\bar{\mathbf{X}}'$ can be partitioned into L different clusters using fuzzy clustering, which assigns each sample, $\bar{\boldsymbol{x}}_i'$, a degree of belonging to a cluster $\ell \in [1, L]$ using a continuous membership function, $M_{i,\ell}$. In fuzzy clustering, the degree of $\bar{\boldsymbol{x}}_i$ belonging to cluster ℓ is taken to be inversely proportional to the squared distance between the point and cluster center, $\mathbf{c}_\ell$, and then normalized across all clusters [21]:

$$M_{i,\ell} = \frac{\|\bar{\boldsymbol{x}}_i' - \mathbf{c}_\ell\|^{-2}}{\sum_{\ell=1}^{L} \|\bar{\boldsymbol{x}}_i' - \mathbf{c}_\ell\|^{-2}} \tag{6}$$

such that $\sum_{\ell=1}^{L} M_{i,\ell} = 1 \ \forall i$. The center of a cluster is taken as the mean of all the points, weighted by their membership to the cluster or $\mathbf{c}_\ell = \frac{\sum_{i=1}^{N} M_{i,\ell}^2 \bar{\boldsymbol{x}}_i}{\sum_{i=1}^{N} M_{i,\ell}^2}$ [21]. In the fuzzy c-means clustering algorithm, the memberships and cluster centers are computed by iteratively minimizing the objective function, $J = \sum_{i=1}^{N} \sum_{\ell=1}^{L} M_{i,\ell}^2 \|\bar{\boldsymbol{x}}_i - \mathbf{c}_\ell\|^2$ [22]. The algorithm is terminated when the changes in the membership functions between two iterations is smaller than some pre-defined tolerance. As this is a non-linear optimization, this procedure can possibly terminate at a local minimum; therefore, the optimization is usually repeated numerous times starting from different initial memberships, and the results are selected for the replicate with the minimum objective function value.

Remark 2 From the definition of $\bar{\mathbf{X}}$ in this section, the dimension of the space required to be clustered can be prohibitively high. The dimensionality problem was addressed in this work by first projecting the variables in $\bar{\mathbf{X}}$ onto a lower dimensional, latent variable subspace using PCA and clustering the resulting latent variable or score space. The resulting loading matrix from PCA, $\mathbf{P}$, can be used to relate the original cluster space (measurable) variables to the latent variables according to: $\mathbf{T} = \bar{\mathbf{X}}\mathbf{P}$. In addition to better numerical properties (i.e., orthogonality), the score space typically has a much lower dimension, making the clustering computations more numerically stable and computationally tractable.

B. Data-based Model Development

In this section, we propose the multi-model approach for modeling the outputs of the nylon-6,6 polymerization system. Given a database of previous nylon-6,6 batches, the main steps in the multi-model approach are to first cluster the $\bar{\mathbf{X}}$ space (or the score space after decomposing $\bar{\mathbf{X}}$ using PCA) of the batch database using fuzzy c-means clustering and then to (simultaneously) identify several local linear models around the cluster center points. These models are then combined with appropriate weights to describe the global nonlinear behavior. For the individual linear models, we employ the ARX model form in (5). Mathematically, this idea is expressed by the following model:

$$\hat{y}(k) = \sum_{\ell=1}^{L} w_\ell(k) \hat{\boldsymbol{\beta}}_\ell \bar{\boldsymbol{x}}(k) \tag{7}$$

where $w_\ell(k)$ is the (normalized) weight given to model ℓ of the L total models and $\hat{\boldsymbol{\beta}}_\ell$ defines the ℓ-th local model of the output. Note that from (7), all L ARX models are assumed to have the same lag structure, but this assumption can be easily removed. If the weights corresponding to the training data are known prior to estimating the individual model parameters, (7) becomes linear in $\hat{\boldsymbol{\beta}}_\ell$, and the system identification problem reduces to a regression problem that is solvable using PCR.

Intuitively, from the process description in (4), the weights placed on local linear models should depend on the current value of the states and inputs since they define the system dynamics. In other words, the local models should be weighted according to the current process conditions. In the absence of state measurements, a combination of lagged outputs and inputs can be used to infer the current process conditions. Accordingly, in this work, to determine the weights for the training data, the normalized fuzzy clustering membership function in (6) is used. Because the membership function quantifies the degree to which a lagged output-input combination belongs to each cluster, it is also indicative of which local models should be given more weight than the others. For instance, if an output-input combination nearly coincides with a specific cluster center point, the local linear model corresponding to that cluster should be given most of the weight. This is consistent with (6) as the membership function value corresponding to that cluster will be close to 1 while for the remaining clusters, the membership function value will be near 0.

Remark 3 One advantage of using a continuous membership function is evident when an output-input combination is encountered that belongs to many clusters with varying degrees. This becomes particularly important during periods of transition in a polymerization process when it is evolving from one stage to another. In this case, as multiple models will be weighted appropriately, the information from several different models can be used, resulting in more accurate predictions. In multi-model approaches using crisp clustering algorithms (i.e., PWA), only samples belonging to a specific cluster can contribute in determining the model and no surrounding information is used. That is, artificial boundaries are established for the clusters and their corresponding models, which can lead to abrupt predictions during transitional periods.

To develop the data-based models for the two outputs, a database of previous batches was first generated. To this end, the deterministic nylon-6,6 polymerization model in [16] was simulated 15 times from different initial conditions (five of the 15 batches were reserved as the validation data set). To mimic a typical industrial batch data set, which is comprised mostly of "successful" batches, the set of reference temperature and pressure profiles were tracked reasonably well using two PI controllers in the simultations. For the PI loop-pairing, the vent rate was used to track the reactor pressure while the steam jacket pressure was used to track the temperature. Both PI controllers were tightly tuned for one set of initial conditions and fixed for the remaining batches. The criteria used to tune the PI controllers was to minimize the integral of time-weighted absolute error (ITAE) while attaining reasonably smooth input trajectories.

The identification procedure for the local linear models for a given output was as follows. For a given lag structure, the number of clusters was varied from $L = 1$ to $L = 20$. For each choice of L, the $\bar{\mathbf{X}}$ matrix was constructed, decomposed using PCA, and clustered. Since the model weights were then known, multiple ARX models were estimated using PCR. This was repeated for all possible lag structures with a lag range of $0 - 2$ for each variable. The goodness of each fit was judged using its root mean squared error (RMSE) in predicting back the fit and validation batches. The lag structure and number of clusters, L, for the two outputs that yielded the lowest RMSE values are tabulated in Table I. In Table I, a zero lag indicates that the variable was excluded from the model. Observe that the reactor pressure was not used in predicting the reaction mixture temperature, and its dynamics were best captured with one linear, first order model. These results are consistent with the fundamental process model; the significantly faster pressure dynamics assumed in the state-space model leads to a decoupling of the pressure from the other states (i.e., the pressure does not influence any of the other states and vice-versa), and the pressure ODE is simply the product of a constant gain term and the vent rate (a linear first order model). Despite the decoupling of the outputs, the control problem cannot be decomposed into two single-input-single-output (SISO) problems because one of the inputs, the vent rate, affects both outputs. Another observation from Table I is that the lag structure for the temperature model corresponds to a first order model between the outputs and inputs. One possible explanation for this is the assumption of the same lag structure for all the local models. With this assumption, using all first order models minimizes the possibility of over-fitting.

TABLE I
FINAL LAG STRUCTURES, NUMBER OF CLUSTERS, L, AND RMSE VALUES

	Lags					
Output	T	P	P_j	v	L	RMSE
T	1	0	1	1	5	1.648
P	0	1	0	1	1	0.1791

C. Predictive Control Design and Implementation

A predictive controller for tracking reference temperature and pressure profiles for the nylon-6,6 polymerization process is presented in this section. The control action at each sampling instance in the proposed controller is computed by solving the optimization problem below.

$$\min_{\boldsymbol{u}_{\min} \le \boldsymbol{u}(k) \le \boldsymbol{u}_{\max}} J = \sum_{k=1}^{P} \|\hat{\boldsymbol{y}}(k) - \boldsymbol{y}_{\text{ref}}(k)\|_{\mathbf{Q}} + \|\Delta \boldsymbol{u}(k)\|_{\mathbf{R}} \quad (8)$$

subject to: Equation (7) (9)

$$\hat{\boldsymbol{y}}(0) = \boldsymbol{y}(t) \quad (10)$$

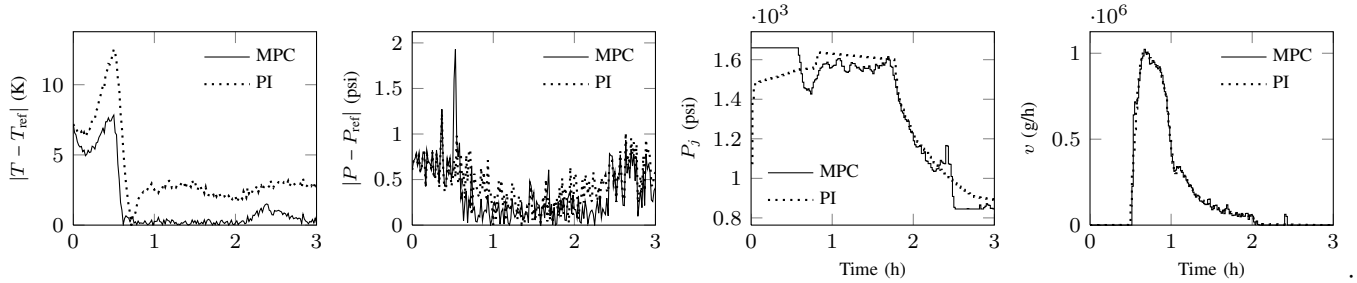
The first term in the objective function penalizes discrepancies between the predicted output trajectories, $\hat{\boldsymbol{y}}$, and the output reference trajectories, $\boldsymbol{y}_{\text{ref}}$, over the prediction horizon, P, and the second term is a move suppression term that penalizes the magnitude of input changes (i.e., the control rate). The positive-definite matrices, $\mathbf{Q}$ and $\mathbf{R}$, are used to trade-off the relative importance of the output and input performance. The constraint in (9) simply states that the data-based model is the predictive model in the MPC formulation, and (10) represents the initialization of the optimization problem at the current process conditions and can be understood as the feedback mechanism to account for plant model mismatch.

Closed-loop simulations for 10 new initial conditions were performed using the proposed trajectory tracking MPC design, and the performance was compared against a PI controller. All initial conditions were ensured to be within the range of initial conditions in the training data. All controllers were tuned once for a specific set of initial conditions and left unchanged for the remainder of the simulations to avoid confounding the results with tuning. The tuning parameters used for the proposed MPC were: $P = 12$, $\mathbf{Q} = \text{diag}\{2.75, 27.5\}$, and $\mathbf{R} = \text{diag}\{0.02, 0.02\}$. Note that with $P = 12$, the proposed MPC design was efficiently solvable; the average CPU time required time to solve the MPC optimization problem (as reported by the MATLAB functions `tic` and `toc`) was 0.69 seconds (using GAMS with IPOPT as the solver on an Intel Quad Core machine). The metric used to assess the tracking performance of the controllers was the ITAE between the process variable value and the reference trajectory. The results are summarized in Table II.

In all simulations, the proposed predictive controller outperformed the PI controller. A representative set of closed-loop simulation results is presented in Fig. 1 (initial condition 5). In this case, the ITAEs for the proposed predictive controller improved on the PI controller by 77% and 26% for temperature and pressure tracking (respectively). Overall, the simulation results clearly demonstrate the advantages of implementing the proposed trajectory tracking predictive controller over PI control.

IV. CONCLUSIONS

In this work, we addressed the problem of modeling and controlling a nylon-6,6 batch polymerization process. A multi-model approach was developed which exploited historical batch data, the simplicity of local linear models, the data



;

Fig. 1. Representative output tracking error (magnitudes) and input profiles for PI control and the proposed trajectory tracking MPC design

TABLE II

TRACKING PERFORMANCE COMPARISON OF THE PI CONTROLLER AND PROPOSED MPC FORMULATION FOR 10 INITIAL CONDITIONS (ICS)

	Temperature ITAE		*Pressure* ITAE	
IC	PI	MPC	PI	MPC
1	9.10	2.18	3.08	1.01
2	10.18	1.51	3.46	1.22
3	3.33	1.33	5.02	1.47
4	4.95	1.76	2.71	2.03
5	12.99	2.98	1.93	1.43
6	6.99	1.21	10.29	1.04
7	13.29	2.72	5.18	1.57
8	3.14	1.19	1.20	0.860
9	14.63	1.90	1.43	1.31
10	4.91	1.89	4.61	1.58
Average:	8.35	1.87	3.89	1.35

extraction capabilities of latent variable tools, and appropriate clustering and weighting techniques to capture the nonlinear nature of the nylon-6,6 batch polymerization process. The resulting model was used to formulate a trajectory tracking predictive controller for the key measurable process variables, namely the reaction mixture temperature and reactor pressure. Closed-loop simulation results (subject to noise and disturbances in the initial conditions) demonstrated the advantages of using the proposed control design over PI control.

REFERENCES

[1] A. Bemporad, M. Morari, V. Dua, and E. Pistikopoulos, "The explicit linear quadratic regulator for constrained systems," *Automatica*, vol. 38, no. 1, pp. 3 – 20, 2002.

[2] J. Espinosa, J. Vandewalle, and V. Wertz, *Fuzzy Logic, Identification and Predictive Control (Advances in Industrial Control)*. Secaucus, NJ, USA: Springer-Verlag, 2004.

[3] S. Aumi and P. Mhaskar, "Integrating data-based modeling and nonlinear control tools for batch process control (submitted)," *AIChE J.*, 2011.

[4] ——, "Integrating data-based modeling and nonlinear control tools for batch process control, (accepted)," in *Proc. of the American Control Conference (ACC)*, 2011.

[5] D. J. Kozub and J. F. MacGregor, "Feedback control of polymer quality in semi-batch copolymerization reactors," *Chem. Eng. Sci.*, vol. 47, no. 4, pp. 929 – 942, 1992.

[6] J. Flores-Cerrillo and J. F. MacGregor, "Control of batch product quality by trajectory manipulation using latent variable models," *J. Process Control*, vol. 14, no. 5, pp. 539 – 553, 2004.

[7] T. J. Crowley, E. S. Meadows, E. Kostoulas, and F. D. III, "Control of particle size distribution described by a population balance model of semibatch emulsion polymerization," *J. Process Control*, vol. 10, no. 5, pp. 419 – 432, 2000.

[8] Y. Yabuki and J. F. MacGregor, "Product quality control in semibatch reactors using midcourse correction policies," *Ind. & Eng. Chem. Res.*, vol. 36, no. 4, pp. 1268–1275, 1997.

[9] S. M. Cruickshank, A. J. Daugulis, and P. J. McLellan, "Dynamic modeling and optimal fed-batch feeding strategies for a two-phase partitioning bioreactor." *Biotech. and Bioeng*, vol. 67, pp. 224–233, 2000.

[10] R. Cardello and K. Y. San, "Application of gain scheduling to the control of batch bioreactors," in *Proc. of the American Control Conference (ACC)*, June 1987, pp. 682 –686.

[11] T. Clarke-Pringle and J. F. MacGregor, "Nonlinear adaptive temperature control of multi-product, semi-batch polymerization reactors," *Comp. & Chem. Eng.*, vol. 21, no. 12, pp. 1395 – 1409, 1997.

[12] C. Kravaris and C.-B. Chung, "Nonlinear state feedback synthesis by global input/output linearization," *AIChE J.*, vol. 33, no. 4, pp. 592 – 603, 1987.

[13] C. E. Garcia, "Quadratic/Dynamic Matrix Control of Nonlinear Processes - An Application to a Batch Reaction Process," in *AIChE Annual Meeting*, San Francisco, CA, USA, 1984.

[14] G. Gattu and E. Zafiriou, "Nonlinear quadratic dynamic matrix control with state estimation," *Ind. & Eng. Chem. Res.*, vol. 31, no. 4, pp. 1096 – 1104, 1992.

[15] T. Peterson, E. Hernandez, Y. Arkun, and F. Schork, "Nonlinear DMC algorithm and its application to a semibatch polymerization reactor," *Chem. Eng. Sci.*, vol. 47, no. 4, pp. 737 – 753, 1992.

[16] S. A. Russell, D. G. Robertson, J. H. Lee, and B. A. Ogunnaike, "Control of product quality for batch nylon-6,6 autoclaves," *Chem. Eng. Sci.*, vol. 53, no. 21, pp. 3685 – 3702, 1998.

[17] M. Joly and J. Pinto, "Optimal control of product quality for batch nylon-6,6 autoclaves," *Chem. Eng. J.*, no. 97, pp. 87–101, 2004.

[18] L. Ljung, *System Identification: Theory for the User (2nd Edition)*. Prentice Hall PTR, 1998.

[19] P. Geladi and B. Kowalski, "Partial least-squares regression: A tutorial," *Anal. Chim. Acta*, vol. 185, pp. 1 – 17, 1986.

[20] R. Bro, K. Kjeldahl, A. Smilde, and H. Kiers, "Cross-validation of component models: A critical look at current methods," *Anal. Bioanal. Chem.*, vol. 390, pp. 1241–1251, 2008.

[21] R. Hathaway and J. Bezdek, "Recent convergence results for the fuzzy c-means clustering algorithms," *J. Classif.*, vol. 5, no. 2, pp. 237–247, 1988.

[22] G. A. F. Seber, *Multivariate Observations*. New York, NY, USA: John Wiley & Sons, 1984.