Available online at www.sciencedirect.com

ScienceDirect

IFAC Papers Online CONFERENCE PAPER ARCHIVE

IFAC-PapersOnLine 49-7 (2016) 061–066

Batch-to-Batch Optimization of Chemical Processes in the Presence of Model-Plant Mismatch Using a Variable Subset of Model Parameters

Rubin Hille, Hector M. Budman

Department of Chemical Engineering, University of Waterloo, Waterloo, ON N2L 3G1 Canada (Tel: 519-888-4567, x36980; e-mail: hbudman@uwaterloo.ca)

Abstract: In the presence of model-plant mismatch, a standard "two-step" approach, involving repeated identification and optimization steps, cannot guarantee convergence to the process optimum. Model parameter adaptation can be used for handling model error by correcting for mismatch between predicted and measured cost and constraint gradients while simultaneously satisfying both identification and optimization objectives. However, updating all model parameters at once is often impractical due to estimability and increased sensitivity to noise. This work presents a procedure for selecting, after each run, a particular subset of parameters based on parametric sensitivity of the model output and of cost and constraint gradients. The resulting improvements with respect to previous run-to-run studies are illustrated using a simulated case study of a penicillin fed-batch process.

Keywords: Model-based Optimization, Model-plant Mismatch, Parameter Adaptation, Model Correction, Parameter Selection, Sensitivity Analysis

1. INTRODUCTION

In the absence of mathematical models of general validity for large regions of operating conditions, run-to-run (batch-to-batch) optimization procedures have been proposed where the model is successively re-identified and re-optimized at each run. For a successful run-to-run optimization, it is necessary to use a model which gives an accurate description of the given process in the neighborhood of the current operating point. However, modeling assumptions and simplifications are often made which result in structural mismatch between the model and the process. Therefore, a repeated identification and optimization (Chen et al., 1987) performed with an imperfect model will generally result in a sub-optimal operating policy.

Model structure errors coupled with parametric uncertainty may lead to conflicts between model fitting and optimization objectives. To reach the process optimum in the presence of model-plant mismatch, it is therefore necessary to consider not just output measurements, but also available measurements of the gradients of the cost function and constraints with respect to input variables.

A class of algorithms, referred to as *Modifier Adaptation* (Roberts et al., 1979; Gao et al., 2005; Chachuat et al. 2009; Marchetti et al., 2010; Gao et al. 2015), modifies the cost function and constraints of the optimization problem to account for differences between measured and predicted gradients. While this approach guides the run-to-run optimization procedure to the optimal operating point, it requires a filter to avoid aggressive corrections. Also, this approach does not explicitly update the model and consequently, it cannot be used for model-based predictions.

To address simultaneous identification and optimization under model-plant mismatch, Srinivasan et al. (2002) modified the identification objective by including a weighted optimization objective. However, this "modeling for optimization" paradigm results in a trade-off between identification and optimization. In contrast, Mandur et al. (2015a, b) proposed to satisfy the identification objective first. Subsequently, the gradients of the model are matched to the plant gradients by adapting the parameters. This satisfies the optimization objective to a pre-specified tolerance while the accuracy of the identification step is maintained by incorporating a correction term. The method was shown to provide an accurate model at each iteration and robustness with respect to model uncertainty.

For the purpose of simultaneous identification and optimization it is often impractical to update all the model parameters due to limited information and correlation. Estimating all parameters may also increase sensitivity to noise due to overfitting. Instead, it is preferable to select a subset of parameters that mostly affect the model output as well as the gradients of cost and constraints. In this work we propose a parameter selection procedure which is applied after each new batch in the run-to-run optimization. The procedure takes into account the parametric sensitivities of model output, cost function and constraint gradients. We show that the selection of appropriate parameters can improve the overall convergence in batch-to-batch parameter adaptation based schemes as well as reduce the model prediction errors in the neighborhood of the process optimum.

The paper is organized as follows: Section 2 reviews the parameter adaptation methodology, while the proposed parameter selection procedure is presented in section 3. The method is illustrated by a case study of a penicillin fed-batch

2405-8963 © 2016, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.07.217

process in section 4. Finally, conclusions are presented in section 5.

2. PARAMETER ADAPTATION METHODOLOGY

For background information, the algorithm of Mandur and Budman (2015a, b) is briefly reviewed.

2.1 Parameter Estimation

In the simultaneous identification and optimization framework, an identification step is performed by minimizing the difference between prediction $\boldsymbol{y}$ and measurements $\boldsymbol{y}_m$ collected at intervals t_i along a batch run:

$$\boldsymbol{\theta}_k = \min_{\boldsymbol{\theta}} \sum_{i=1}^{n_t} \|\boldsymbol{y}_m(\boldsymbol{u}_k, t_i) - y(\boldsymbol{u}_k, \boldsymbol{\theta}, t_i)\|^2$$

$$s.t. \quad \dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}_k, \boldsymbol{\theta}) \qquad \boldsymbol{y} = h(\boldsymbol{x}) \tag{1}$$

Where $\boldsymbol{u}_k$ are the decision variables defining the current operating point and $\boldsymbol{\theta}_k$ are a particular set of parameter estimates which minimize the identification objective.

2.2 Gradient Matching

In order to also satisfy the optimization objective, the predicted gradients of the cost ϕ and constraints $\boldsymbol{g}$ are matched to those of the process (ϕ_P and $\boldsymbol{g}_P$). Therefore, the parameter estimates obtained by solving (1), are changed by an amount $\Delta\boldsymbol{\theta}$ in the following way:

$$\Delta\boldsymbol{\theta}_k = \min_{\Delta\boldsymbol{\theta}} \left(\boldsymbol{w}_\phi \left| \frac{\partial \phi_P(\boldsymbol{y}_m(\boldsymbol{u}))}{\partial \boldsymbol{u}} - \frac{\partial \phi(y(\boldsymbol{u}, \boldsymbol{\theta}_k + \Delta\boldsymbol{\theta}))}{\partial \boldsymbol{u}} \right| + \boldsymbol{w}_g \left| \frac{\partial \boldsymbol{g}_P(\boldsymbol{y}_m(\boldsymbol{u}), \boldsymbol{u})}{\partial \boldsymbol{u}} - \frac{\partial \boldsymbol{g}(y(\boldsymbol{u}, \boldsymbol{\theta}_k + \Delta\boldsymbol{\theta}), \boldsymbol{u})}{\partial \boldsymbol{u}} \right| \right)$$

$$s.t. \quad \dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}_k, \boldsymbol{\theta}_k + \Delta\boldsymbol{\theta}) \qquad \boldsymbol{y} = h(\boldsymbol{x}) - \boldsymbol{c}_k \qquad \|\boldsymbol{\epsilon}^T\|_\infty \leq \epsilon_{max}^T \tag{2}$$

Where $\boldsymbol{w}_\phi$ and $\boldsymbol{w}_g$ are weights that are used to normalize the two gradient matching objectives. To maintain the fitting accuracy, a correction term is introduced and approximated with a first order Taylor expansion:

$$\boldsymbol{c}_k = \boldsymbol{y}(\boldsymbol{\theta}_k) - \boldsymbol{y}(\boldsymbol{\theta}_k + \Delta\boldsymbol{\theta}_k) \cong \boldsymbol{D}y(\boldsymbol{\theta}_k) \cdot \Delta\boldsymbol{\theta}_k \tag{3}$$

Where $\boldsymbol{D}y(\boldsymbol{\theta}_k)$ is the Jacobian of the model. As shown in (2), the change of parameter values by $\Delta\boldsymbol{\theta}$ due to the gradient matching is constrained by an upper bound ϵ_{max}^T on the relative truncation error, which is defined as the error introduced by the Taylor approximation in (3):

$$\boldsymbol{\epsilon}(t_i) = \left[\boldsymbol{y}(\boldsymbol{u}_k, \boldsymbol{\theta}_k + \Delta\boldsymbol{\theta}_k, t_i) - \boldsymbol{D}y(\boldsymbol{\theta}_k, t_i) \cdot \Delta\boldsymbol{\theta}_k - \boldsymbol{y}(\boldsymbol{u}_k, \boldsymbol{\theta}_k, t_i) \right] \cdot diag^{-1}\left(\boldsymbol{y}(\boldsymbol{u}_k, \boldsymbol{\theta}_k, t_i)\right) \tag{4}$$

The use of (4) as an upper bound in problem (2) was shown to provide robustness to uncertainty in the measured gradients of cost and constraints (Mandur et al., 2015a).

2.3 Model-based Optimization

Using the updated parameter values, i.e. $\boldsymbol{\theta} = \boldsymbol{\theta}_k + \Delta\boldsymbol{\theta}_k$, a model-based optimization is performed to determine the new optimal operating point:

$$\boldsymbol{u}_{k+1} = \min_{\boldsymbol{u}} \phi\big(\boldsymbol{y}(\boldsymbol{u}, \boldsymbol{\theta})\big)$$

$$s.t. \quad \dot{\boldsymbol{x}} = f(\boldsymbol{x}, \boldsymbol{u}_k, \boldsymbol{\theta}) \qquad \boldsymbol{y} = h(\boldsymbol{x}) - \boldsymbol{c}_k \qquad \boldsymbol{g}(\boldsymbol{y}(\boldsymbol{u}, \boldsymbol{\theta}), \boldsymbol{u}) \leq \boldsymbol{0} \qquad \boldsymbol{u}^L \leq \boldsymbol{u} \leq \boldsymbol{u}^U \tag{5}$$

Mandur et al. (2015a, b) showed that, upon convergence, the necessary conditions of optimality (NCOs) of the model are equal to those of the process at the plant optimum.

3. PARAMETER SELECTION PROCEDURE

As mentioned earlier, it is possible to update all the parameters of the model but this is impractical for computational reasons and higher sensitivity to noise. Instead, it is preferable to update only parameters that have large effects on model outputs and gradients of cost function and constraints.

3.1 Output Sensitivities

The scaled local sensitivity of a model-output y_j with respect to a parameter θ_i at time point $t_k \in [t_1, \dots, t_f]$ and nominal parameter values $\boldsymbol{\theta}_l$ is defined as:

$$S_{\theta_i}^{y_j}(\boldsymbol{\theta}_l, t_k) = \frac{\partial y_j(\boldsymbol{\theta}_l, t_k)}{\partial \theta_i} \cdot \frac{\boldsymbol{\theta}_l}{\bar{y}_j(\boldsymbol{\theta}_l)} \tag{6}$$

Where the number of sampling points along a batch is given by n_t, i.e. $t_{n_t} = t_f$. With n_y outputs and n_θ parameters, the output sensitivity matrix is then given by:

$$\boldsymbol{S}(\boldsymbol{\theta}_l) = \begin{bmatrix} S_{\theta_1}^{y_1}(t_1) & \cdots & S_{\theta_{n_\theta}}^{y_1}(t_1) \\ \vdots & \ddots & \vdots \\ S_{\theta_1}^{y_{n_y}}(t_1) & \cdots & S_{\theta_{n_\theta}}^{y_{n_y}}(t_1) \\ \vdots & \ddots & \vdots \\ S_{\theta_1}^{y_1}(t_{n_t}) & \cdots & S_{\theta_{n_\theta}}^{y_1}(t_{n_t}) \\ \vdots & \ddots & \vdots \\ S_{\theta_1}^{y_{n_y}}(t_{n_t}) & \cdots & S_{\theta_{n_\theta}}^{y_{n_y}}(t_{n_t}) \end{bmatrix} \tag{7}$$

A lumped measure of the parametric sensitivity of the model outputs for all batch sampling points can be expressed as:

$$S_{\theta_i}^{y}(\boldsymbol{\theta}_l) = \sum_{j=1}^{n_y \times n_t} \left| S(\boldsymbol{\theta}_l)_{j,i} \right| \tag{8}$$