Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Preferential Estimation for Uncertain Linear Systems at Steady State: Application to Filamentous Fungal Fermentation

L. Bodizs, B. Srinivasan, D. Bonvin

Abstract— **State estimation is a widely used concept in the control community, and the literature mostly concentrates on the estimation of all states. However, in soft sensor problems, the emphasis is on estimating a few soft outputs as accurately as possible. The concept of preferential estimation consists of estimating these soft outputs more accurately than the other states. The main question is whether or not the accuracy along the soft outputs can be improved, possibly at the detriment of other states. This papers shows that, though preferential estimation is not possible for linear systems with perfect model information and gaussian process and measurement noises, it is indeed possible for linear systems with model uncertainty. The theoretical concepts are illustrated on a filamentous fungal fermentation.**

I. INTRODUCTION

State estimation is a necessary component of sophisticated monitoring and control techniques, since these techniques typically require information that is too expensive or impossible to obtain from direct measurements. Estimation attempts to reconstruct the missing information from both the available measurements and prior knowledge in the form of a dynamic model [1], [2].

Full-state estimation is usually considered, due to the close link between estimation theory and the full-state feedback literature [3], [4]. On the other hand, in the context of soft sensing that concentrates on reconstructing certain state variables that are not directly measured, the estimation accuracy of the other variables is of lesser importance. In this paper, the soft sensing problem is considered, where the emphasis is on estimating a vector (of *preferred variables*) of dimension much lower than that of the state vector. Such a problem might arise, for example, for optimizing a process via the tracking a given state variable [5].

One way of estimating preferred variables is to estimate the entire state vector using standard full-state techniques such as Kalman filtering [6]–[8] and single out the preferred variables via projection. The drawback of this approach is that the focus is on the accuracy of the entire state vector rather than that of the preferred variables. Consequently, the preferred variables will inherit the accuracy of the states, though they could probably be estimated more accurately if attention were placed exclusively on their estimation.

The objective of *preferential estimation* is to estimate certain variables more accurately than what can be done by standard estimation followed by projection. The first and foremost question of preferential estimation is whether this is at all possible. In the case of linear systems with perfect model information and gaussian process and measurement noises, the optimal solution for estimation problems is the Kalman filter [8]. In such a case, preferential estimation techniques cannot improve the accuracy of selected variables, since the Kalman filter is optimal for the full state vector and, consequently, also for the preferred variables.

However, in the presence of model uncertainty, the estimated variables are typically biased, and the Kalman filter is no longer optimal. It was shown that, in this case, it is possible to reduce the error in selected variables [9]. The bias caused by model-plant mismatch opens the way to compromises between bias and variance on the one hand, and bias in the preferred variables and bias in the other variables on the other. These compromises allow reducing the bias in selected directions, at the expense of an increased bias and/or variance in other directions.

The bias-variance and bias-bias tradeoffs are studied qualitatively in [9]. The objective of the present paper is to provide a more comprehensive analysis of these tradeoffs for the case of linear systems operating at steady state. Firstly, a noise-free scenario is considered, where only bias-bias tradeoffs are present. The main result shows that the bias cannot be reduced to zero for the entire state vector, while it can indeed be pushed to zero along selected directions. This clearly motivates the need for preferential estimation. Secondly, a noise-corrupted scenario is considered, where the existence of bias-variance tradeoffs in preferential estimation is shown.

The paper is organized as follows. In Section II, the formulation of preferential estimation (PE) is introduced. Sections III and IV present PE for a noise-free and a noise-corrupted scenario, respectively. Section V illustrates PE using a linear model operating at steady-state for a filamentous fungal fermentation. Finally, Section VI concludes the paper.

II. PREFERENTIAL ESTIMATION

The concept of preferential estimation, as introduced in [9], consists of estimating certain linear combinations of the states more accurately than others. The concepts therein are summarized next.

Consider the following linear, discrete-time system:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k, \quad x_0 = x_o \\ y_k &= Hx_k + v_k \end{aligned} \tag{1}$$

where $u \in \Re^l$ are the inputs, $x \in \Re^n$ the states, x_o the initial values of states, w the process noise, $y \in \Re$ the scalar measurement and v the measurement noise. The

L. Bodizs and D. Bonvin are with École Polytechnique Fédérale de Lausanne, CH-1015 Switzerland

B. Srinivasan is with École Polytechnique, Montreal, H3C 3A7, Canada

0-7803-9568-9/05/$20.00 ©2005 IEEE

matrices A, B and H describe the system dynamics and the measurement. Without loss of generality, and in order to simplify the proofs in Section III, a single-output system is considered here.

The problem of preferential estimation requires the definition of the preferred variables:

$$z_k = Lx_k, \tag{2}$$

where $z \in \Re^m$ are the preferred variables, and L an $m \times n$ projection matrix, with $m < n$ and $rank(L) = m$. Note that the preferred variables z are typically defined by the problem at hand, and thus given a priori. Hence, the same also holds for L.

Preferential estimation is formulated as the minimization of the mean-squared estimation error J of the preferred variables z [9]:

$$\begin{aligned} \min_{K_k} \quad & J_k = E\langle (z_k - \hat{z}_k)^T (z_k - \hat{z}_k) \rangle \\ s.t. \quad & \hat{x}_{k+1} = A_m \hat{x}_k + Bu_k + K_k(y_k - \hat{y}_k), \quad \hat{x}_0 = E\langle x_o \rangle \\ & \hat{y}_k = H\hat{x}_k \\ & \hat{z}_k = L\hat{x}_k \end{aligned} \tag{3}$$

where K is the estimator gain, and the symbol $\hat{(\cdot)}$ denotes the estimate of the corresponding variable. In the above formulation, an imperfect plant model A_m is considered:

$$A_m = A - \Delta A, \tag{4}$$

where ΔA is the model-plant mismatch.

Note that the objective function chosen in (3) is the mean-square error (a scalar) instead of the covariance (matrix) that is generally used in estimation problems. The reason for choosing a scalar cost function rather than a matrix is that, except for the special case of linear systems without uncertainty, there exists no unique estimator gain that minimizes every element of the matrix. Thus, a weighted sum of the various elements of the matrix is necessary in order to define a solution. The mean-squared estimation error, $E\langle (z_k - \hat{z}_k)^T (z_k - \hat{z}_k) \rangle$, which is the trace of the covariance matrix $E\langle (z_k - \hat{z}_k)(z_k - \hat{z}_k)^T \rangle$, represents one such possible weighting of the covariance matrix. Note also that, if the projection matrix $L = I_{n \times n}$ is chosen, Problem (3) corresponds to minimizing the error in $\hat{x}$, i.e. the standard full-state estimation (SE) defined in [9] as the minimization of the mean-squared error in $\hat{x}$.

Preferential estimation consists of two steps: (a) the tuning step, where (3) is typically solved off-line, and (b) the prediction step, where the estimators obtained from step (a) are used for on-line prediction of $\hat{z}$. Since (3) does not always have an analytical solution, a numerical solution is sought for the first step.

Two approaches for solving Optimization problem (3) are possible:

1) Certain noise (w_k and v_k) and uncertainty (ΔA) structures are assumed and the optimal estimator gain computed. This approach makes the fairly unrealistic assumption that the uncertainty is known, but it gives considerable insight into the optimal solution.
2) The optimum is computed from off-line calibration data obtained from the system itself. The tuning step uses off-line measurements of z. This approach has a strong analogy with the calibration-based methods that are heavily used in the chemometrics field [10].

In this paper, only the first approach will be studied, the objective being to show the existence of preferential estimation analytically. The second route was taken in [9].

III. PREFERENTIAL ESTIMATION FOR A NOISE-FREE SCENARIO

By considering $w_k = 0$, $v_k = 0$, the following uncertain deterministic system is obtained at steady state for $u_k = \bar{u}$:

$$\begin{aligned} \bar{x} &= (A_m + \Delta A)\bar{x} + B\bar{u} \\ \bar{y} &= H\bar{x} \end{aligned} \tag{5}$$

For this scenario, the PE-problem (3) reads:

$$\begin{aligned} \min_{K} \quad & \bar{J} = (\bar{z} - \hat{\bar{z}})^T (\bar{z} - \hat{\bar{z}}) \\ s.t. \quad & \hat{\bar{x}} = A_m \hat{\bar{x}} + B\bar{u} + K(\bar{y} - \hat{\bar{y}}) \\ & \hat{\bar{y}} = H\hat{\bar{x}} \\ & \hat{\bar{z}} = L\hat{\bar{x}} \end{aligned} \tag{6}$$

The error can be expressed from (5) and (6) as:

$$\bar{x} - \hat{\bar{x}} = (A_m - KH)(\bar{x} - \hat{\bar{x}}) + \Delta A\bar{x} \tag{7}$$

$$\bar{x} - \hat{\bar{x}} = (I - A_m + KH)^{-1} \Delta A\bar{x} \tag{8}$$

$$\bar{z} - \hat{\bar{z}} = L(I - A_m + KH)^{-1} \Delta A\bar{x} \tag{9}$$

The optimization problem (6) seeks to push $(\bar{z} - \hat{\bar{z}})$ to 0. Whether this can be achieved will be studied in the next subsections.

A. Impossibility to Eliminate Bias in All States

This subsection addresses the question of whether, for the case of imperfect model, it is possible to eliminate the bias in all the states. As expected, the answer is no.

Theorem 1: Let $\bar{u} \neq 0$, ΔA be such that $d = \Delta A\bar{x} \neq 0$ and A be stable (all eigenvalues within the unit circle). If L has rank n (i.e. $m = n$), then there exists no finite K that can lead to $\bar{z} - \hat{\bar{z}} = 0$.

Proof:

$$\bar{z} - \hat{\bar{z}} = L(I - A_m + KH)^{-1} d \tag{10}$$

A possible solution to $\bar{z} - \hat{\bar{z}} = 0$ is $(I - A_m + KH)^{-1} = 0$, which is equivalent to $K \to \infty$. However, if a finite K is sought, then $\bar{z} - \hat{\bar{z}}$ can never be pushed to zero since $(I - A_m + KH)^{-1}$ and L are of rank n. ■

This theorem, though simple, illustrates many important features of estimation with uncertainty. In the case of uncertain systems, the bias can never be eliminated in all states with a finite-gain observer. Secondly, high-gain observers can be used to push the bias towards zero.

In the above theorem, the assumption of stability is needed to ensure that a steady state is reached, while the assumptions

$\bar{u} \neq 0$ and $d \neq 0$ ensure that neither the states nor the perturbation are 0 at that steady state, for which case the theorem would be trivially falsified.

B. Possibility to Eliminate Bias in Preferred Variables

This subsection investigates the conditions under which, for the case of imperfect model, the bias can be eliminated in given preferred variables.

Theorem 2: Let $d = \begin{bmatrix} d_1 & d_2 & \cdots & d_n \end{bmatrix}^T \neq 0$ and

$$\mathcal{D} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \sum_{2}^{n} d_k & -d_1 & \cdots & -d_1 & -d_1 \\ \sum_{3}^{n} d_k & \sum_{3}^{n} d_k & \cdots & -\sum_{1}^{2} d_k & -\sum_{1}^{2} d_k \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{n-1}^{n} d_k & \sum_{n-1}^{n} d_k & \cdots & -\sum_{1}^{n-2} d_k & -\sum_{1}^{n-2} d_k \\ d_n & d_n & \cdots & d_n & -\sum_{1}^{n-1} d_k \end{bmatrix}$$

If $m < n$ and $rank(L\mathcal{D}) = m$, then there exist infinitely many finite values of K that lead to $\bar{z} - \hat{\bar{z}} = 0$.

Proof: Consider the model to be in the observable canonical form, denoted by the subscript '$_T$', upon application of the similarity transformation $x = Tx_T$:

$$A_{m,T} = \begin{bmatrix} -\alpha_1 & 1 & \cdots & 0 \\ -\alpha_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ -\alpha_n & 0 & \cdots & 0 \end{bmatrix}; \quad H_T = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$$

with

$$K_T = \begin{bmatrix} k_{T,1} & k_{T,2} & \cdots & k_{T,n} \end{bmatrix}^T$$

Note that he similarity transformation leads to $L_T = LT$ and $d_T = T^{-1}d$, which gives $L_T d_T = Ld$. Furthermore, define:

$$M_T = I - A_{m,T} + K_T H_T = \begin{bmatrix} 1 + \alpha_1 + k_{T,1} & -1 & \cdots & 0 \\ \alpha_2 + k_{T,2} & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_n + k_{T,n} & 0 & \cdots & 1 \end{bmatrix}$$

Equation (10) can be rewritten as:

$$\begin{aligned} \bar{z} - \hat{\bar{z}} = L_T {M_T}^{-1} d_T = L_T \frac{Adj M_T}{det M_T} d_T \quad &= 0 \\ &\updownarrow \\ L_T Adj M_T d_T \quad &= 0 \end{aligned} \tag{11}$$

Using the notation:

$$\begin{aligned} a_{T,1} &= 1 + \alpha_1 + k_{T,1} \\ a_{T,2} &= \alpha_2 + k_{T,2} \\ &\vdots \\ a_{T,n} &= \alpha_n + k_{T,n} \end{aligned}$$

$Adj M_T$ can be written as:

$$Adj M_T = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ -\sum_{2}^{n} a_{T,k} & a_{T,1} & \cdots & a_{T,1} & a_{T,1} \\ -\sum_{3}^{n} a_{T,k} & -\sum_{3}^{n} a_{T,k} & \cdots & \sum_{1}^{2} a_{T,k} & \sum_{1}^{2} a_{T,k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\sum_{n-1}^{n} a_{T,k} & -\sum_{n-1}^{n} a_{T,k} & \cdots & \sum_{1}^{n-2} a_{T,k} & \sum_{1}^{n-2} a_{T,k} \\ -a_{T,n} & -a_{T,n} & \cdots & -a_{T,n} & \sum_{1}^{n-1} a_{T,k} \end{bmatrix}$$

Note that $Adj M_T$ is affine in K_T as there are only summations of $a_{T,k}$-elements. So, it would be useful to rewrite (11) as a system of linear equations in $a_{T,k}$. This can be done as follows:

$$\begin{aligned} L_T Adj M_T d_T \quad &= 0 \\ &\updownarrow \\ L_T \mathcal{D}_T \mathcal{A}_T \quad &= L_T \mathcal{D}_{T,0} \end{aligned} \tag{12}$$

where

$$\mathcal{D}_T = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \sum_{2}^{n} d_{T,k} & -d_{T,1} & \cdots & -d_{T,1} & -d_{T,1} \\ \sum_{3}^{n} d_{T,k} & \sum_{3}^{n} d_{T,k} & \cdots & -\sum_{1}^{2} d_{T,k} & -\sum_{1}^{2} d_{T,k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{n-1}^{n} d_{T,k} & \sum_{n-1}^{n} d_{T,k} & \cdots & -\sum_{1}^{n-2} d_{T,k} & -\sum_{1}^{n-2} d_{T,k} \\ d_{T,n} & d_{T,n} & \cdots & d_{T,n} & -\sum_{1}^{n-1} d_{T,k} \end{bmatrix}$$

$$\mathcal{D}_{T,0} = \begin{bmatrix} \sum_{1}^{n} d_{T,k} & 0 & \cdots & 0 \end{bmatrix}^T$$

$$\mathcal{A}_T = \begin{bmatrix} a_{T,1} & a_{T,2} & \cdots & a_{T,n} \end{bmatrix}^T$$

$$d_T = \begin{bmatrix} d_{T,1} & d_{T,2} & \cdots & d_{T,n} \end{bmatrix}^T \neq 0$$

Considering that $dim\, L_T = m \times n$ and $dim\, \mathcal{D}_T = n \times n$, the condition for (12) to have a non-unique solution is $m < n$ and $rank(L_T \mathcal{D}_T) = m$. From $L_T \mathcal{D}_T = L\mathcal{D}$, the condition $rank(L\mathcal{D}) = m$ follows. ■

Remark: Note that $\mathcal{A}_T = 0$ is not a solution of (12) as $\mathcal{A}_T = 0 \rightarrow det(M_T) = 0$.

Theorem 2 shows that, in a linear deterministic system at steady state, it is possible to push the error in the preferred variables z to zero. Indeed, the bias caused by the model-plant mismatch ΔA can be completely eliminated in z, while it is impossible to eliminate the bias in the entire state vector x. The following "compromise" is found between the biases in the preferred directions and in the rest of the state vector: the bias is reduced to 0 in z, while it is uncontrolled in the other directions. Thus, a bias-bias tradeoff exists and can be exploited for estimating the preferred variables more accurately.

In order to compute K_T, and thus also K, using (11), knowledge of the perturbation d is necessary. Since d is a n-dimensional vector, n parameters have to be identified. Compared to model identification, i.e. the identification of A in order to eliminate model mismatch, where n^2 parameters need to be identified, preferential estimation is less demanding.

Unfortunately, the K computed from (11) is of little practical value since, even though K is finite, stability of the estimator cannot be guaranteed. In the formulation of the estimation problem (6), there is no constraint enforcing stability. In other words, even though K could in principle be chosen so as to eliminate the bias in z at steady state, steady state could very well never be reached using that K.

IV. PREFERENTIAL ESTIMATION FOR A NOISE-CORRUPTED SCENARIO

Consider System (1) with process and measurement noises and $u_k = \bar{u}$, for which Optimization problem (3) can be rewritten as:

$$\begin{aligned} \min_{K} \quad & J_k = E\langle (z_k - \hat{z}_k)^T (z_k - \hat{z}_k) \rangle \\ & = tr(L P_k L^T) \\ s.t. \quad & \hat{x}_{k+1} = A_m \hat{x}_k + B\bar{u} + K(y_k - \hat{y}_k) \\ & \hat{y}_k = H\hat{x}_k \\ & \hat{z}_k = L\hat{x}_k \\ & P_k = E\langle (x_k - \hat{x}_k)(x_k - \hat{x}_k)^T \rangle \end{aligned} \tag{13}$$

where P_k is the covariance matrix.

Since a constant input $\bar{u}$ is considered, the mean and covariance of the estimation error reach steady values for $k \rightarrow \infty$, $\bar{e}$ and $\bar{P}$ respectively, even though neither System (1) nor the estimator reaches steady state. This justifies the use of a constant estimator gain K.

It follows from (1) and (13):

$$\begin{aligned} x_{k+1} - \hat{x}_{k+1} = \quad & (A_m - KH)(x_k - \hat{x}_k) \\ & + \Delta A x_k - K v_k + w_k \\ e_{k+1} = \quad & (A_m - KH) e_k \\ & + \Delta A x_k - K v_k + w_k \end{aligned} \tag{14}$$

from which a recursive formula for $E\langle e_k \rangle$:

$$E\langle e_{k+1} \rangle = (A_m - KH) E\langle e_k \rangle + \Delta A E\langle x_k \rangle \tag{15}$$

Since $E\langle x_k \rangle = \bar{x}$, as $u_k = \bar{u}$, the recursion for $E\langle e_k \rangle$ is:

$$E\langle e_{k+1} \rangle = (A_m - KH) E\langle e_k \rangle + d \tag{16}$$

where d plays the role of a constant input. Thus, (16) reaches the steady state:

$$\bar{e} = (I - A_m + KH)^{-1} d = M^{-1} d \tag{17}$$

Next, a recursive equation for the covariance P_k can be developed. It follows from (14):

$$\begin{aligned} E\langle e_{k+1} e_{k+1}^T \rangle = \quad & (A_m - KH) E\langle e_k e_k^T \rangle (A_m - KH)^T \\ & + K E\langle v_k v_k^T \rangle K^T + E\langle w_k w_k^T \rangle \\ & + \Delta A E\langle \bar{x} \bar{x}^T \rangle \Delta A^T \\ & + \Delta A E\langle \bar{x} e_k^T \rangle (A_m - KH)^T \\ & + (A_m - KH) E\langle e_k \bar{x}^T \rangle \Delta A^T \end{aligned} \tag{18}$$

Using the following notations:

$$\begin{aligned} & R = E\langle v_k v_k^T \rangle, \quad Q = E\langle w_k w_k^T \rangle, \quad \bar{A} = A_m - KH \\ & \bar{Q} = Q + KRK^T + \bar{A} M^{-1} d d^T + d d^T M^{-T} \bar{A}^T + d d^T \end{aligned}$$

(18) can be written in the following recursive form:

$$P_{k+1} = \bar{A} P_k \bar{A}^T + \bar{Q} \tag{19}$$

Note that, for $k \rightarrow \infty$, P converges to $\bar{P}$ and (19) becomes a discrete Lyapunov equation.

In (19), $\bar{Q}$ contains the variance terms Q and KRK^T along with several bias terms (the terms containing d). The estimator gain K plays the role of a weighting vector that balances the variance term versus the bias terms so as to minimize $J_k = tr(L P_k L^T)$. Thus, (19) can explain the bias-variance tradeoffs pointed out in [9]. Besides, due to the presence of the bias terms, the bias-bias tradeoffs explained in Section III are also possible.

Optimization problem (13) is solved numerically since (19) is nonlinear in K. To evaluate P_k, knowledge of only d and R is necessary. Hence, the advantage of preferential estimation in terms of the number of parameters to be identified, compared to model identification, is preserved.

Additionally, the value of K obtained by minimizing $J_k = tr(L P_k L^T)$ always yields a stable estimator. Stability is implicitly ensured by the fact that the covariance P_k, contained in the objective function of the optimization problem (13), is evaluated recursively as given by (19), which in turn implies recursive evaluation of the states as well. If the system tends towards instability, the states, and thus also the elements of

P_k, increase significantly. Hence, in the problem formulation (13), the error (or state) recursion (19) plays the role of a barrier function for stability.

V. ILLUSTRATION OF PREFERENTIAL ESTIMATION ON FILAMENTOUS FUNGAL FERMENTATION

The aim of this section is to illustrate the bias-bias and bias-variance tradeoffs discussed in the previous sections. A model of filamentous fungal fermentation in a fed-batch reactor is first derived. The model is then adapted to describe sustained steady-state operation, and PE is applied to a linearized version of this model.

A. Fungal fermentation

The process studied in this paper is the α-amylase production by *Aspergillus Oryzae*. The same substrate (glucose s) is consumed for both growing the biomass (x) and producing the enzyme (p). The process is operated in fed-batch mode at the industrial scale, as two phases are needed:

1) batch phase – to grow biomass,
2) fed-batch phase, where glucose is fed in – to produce the enzyme.

Because of its filamentous structure, the biomass is divided into three regions [11]:

- *active region (x_a)* – responsible for production
- *extension region (x_e)* – responsible for growth
- *hyphal region (x_h)* – corresponding to the inactive part of the biomass.

The macroscopic reactions read:

$$\begin{aligned} &s + DO \xrightarrow{x_a} x_e \\ &s + DO \xrightarrow{x_e} x_a \rightarrow x_h \\ &s + DO \xrightarrow{x_e} p \end{aligned} \tag{20}$$

B. Fed-batch model

A first-principles model of the process was built for optimization and control purposes [12]. In this paper, only the dynamic mass balance equations necessary to derive the continuous model are given. The algebraic equations ($\mathcal{F}_i,\ i = \{1, 2, ..., 6\}$) can be found in [12].

Morphological states

$$\begin{aligned} \dot{x}_e &= q_1 - Dx_e, \quad x_e(0) = x_{e0} \\ \dot{x}_a &= q_3 - q_1 - q_2 - Dx_a, \quad x_a(0) = x_{a0} \\ \dot{x}_h &= q_2 - Dx_h, \quad x_h(0) = x_{h0} \end{aligned} \tag{21}$$

where $q_1 = \mathcal{F}_1$ is the rate of extension (branching), $q_2 = \mathcal{F}_2$ the rate of inactivation, $q_3 = \mathcal{F}_3$ the growth rate, $D = \frac{F}{V}$ the dilution rate with F the feed rate and V the volume.

Glucose

$$\begin{aligned} \dot{s} = -(Y_{xs} q_3 \quad &+ Y_{ps} r_{ps} x_a + m_s(x_a + x_e + x_h)) \\ &+ D(s_f - s), \quad s(0) = s_0 \end{aligned} \tag{22}$$

where $r_{ps} = \mathcal{F}_4$ is the specific rate of enzyme production, m_s the maintenance coefficient, and s_f the feed concentration. Y_{xs} and Y_{ps} are the yield coefficients for substrate consumption for growth and production, respectively.

Enzyme

$$\dot{p} = r_{ps} x_a - Dp, \quad p(0) = p_0 \tag{23}$$

Dissolved oxygen

$$\begin{aligned} \dot{O}_2 = -r_O(x_a \quad &+ x_e + x_h) + k_L a(O_2^* - O_2) \\ &- D\ O_2, \quad O_2(0) = O_{2,0} \end{aligned} \tag{24}$$

where $r_O = \mathcal{F}_5$ is the specific rate of oxygen consumption, $k_L = \mathcal{F}_6$ the gas-liquid mass transfer coefficient, a the transfer area, and O_2^* the equilibrium O_2 level.

Volume

$$\dot{V} = DV - F_{evap}, \quad V(0) = V_0 \tag{25}$$

where F_{evap} stands for the water evaporation rate.

C. Linearized model

Though the system is operated in fed-batch mode, many of the states remain fairly constant during most of the operation. The death of biomass keeps the active and extension regions at fairly constant levels. Also, the substrate and the dissolved oxygen are relatively constant throughout the entire operation. Though the quantity of enzyme, the dead biomass and the volume increase with time, their influence on the rest of the dynamics can be neglected.

Thus, removing the enzyme, the dead biomass and the volume from the system equations, the reduced set of states and inputs are:

$$x_c = [x_e, x_a, s, O_2]^T; \quad u_c = D$$

Linearizing around the current operating point $\bar{x}_c$ and $\bar{u}_c$, introducing $\delta x_c = x_c - \bar{x}_c$ and $\delta u_c = u_c - \bar{u}_c$, and discretizing using Euler formula results in:

$$\delta x_c(k+1) = A\delta x_c(k) + B\delta u_c(k) \tag{26}$$

The mismatch ΔA is chosen as a 10% time-invariant random deviation from A, v_k is considered to be a 1% white-noise sequence and no w_k is used, i.e. $Q = 0$. $H = [0, 0, 0, 1]$ as typically only the O_2-measurement is available on-line.

D. Application of preferential estimation

In this example, the preferential estimation problem of estimating x_a is considered, i.e. $L = [0, 1, 0, 0]$. This problem is solved numerically, and the results are presented in Table I and Figure 1. In Table I, $e_k = \bar{x}_c - \hat{x}_{c,k}$ is computed based on simulated values; $\bar{e} = E\langle e_k \rangle$ is the bias, $\mathcal{V}_e = E\langle (e_k - E\langle e_k \rangle)^2 \rangle$ is the variance, $\Pi_e = diag(E\langle e_k e_k^T \rangle)$ is the estimation error and $\Sigma_{\Pi_e} = \sum_{i=1}^{4} \Pi_{e,i}$ is the total estimation error.

The results of Table I and Figure 1 call for several remarks:

- In PE, the estimation error of the preferred state, $L\Pi_e$, is the smallest.

- The price to pay for this is an increased total estimation error Σ_{Π_e} in PE.
- PE is realized by reducing the bias $L\bar{e}$ in the preferred state, while increasing it in the first (x_e) and third (s) states. This gives rise to bias-bias tradeoffs.
- In PE, the variance of the preferred state, $L\mathcal{V}_e$, as well as that of states three (s) and four (O_2), are greater than in SE. This gives rise to bias-variance tradeoffs.

Since there is no process noise, the Kalman filter is equivalent to open-loop prediction, as indicated by the first column in Table I. Indeed, $Q = 0$ for the Kalman filter means perfect model, and thus the weight of the measurement is negligible, leading to $K = 0$.

TABLE I

COMPARISON OF OPEN-LOOP PREDICTION, STANDARD ESTIMATION (SE) AND PREFERENTIAL ESTIMATION (PE).

	Open-loop prediction	SE $L = I$	PE $L = [0, 1, 0, 0]$
$\bar{e}$	$\begin{bmatrix} -2.39 \cdot 10^{-1} \\ -7.23 \cdot 10^{-2} \\ 9.69 \cdot 10^{-2} \\ 2.68 \cdot 10^{-1} \end{bmatrix}$	$\begin{bmatrix} 8.31 \cdot 10^{-3} \\ \mathbf{-7.48 \cdot 10^{-2}} \\ 3.50 \cdot 10^{-2} \\ 2.67 \cdot 10^{-1} \end{bmatrix}$	$\begin{bmatrix} -1.56 \cdot 10^{-1} \\ \mathbf{-5.39 \cdot 10^{-4}} \\ 6.17 \cdot 10^{-1} \\ 2.17 \cdot 10^{-1} \end{bmatrix}$
$\mathcal{V}_e$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1.39 \cdot 10^{-4} \\ \mathbf{8.82 \cdot 10^{-8}} \\ 1.71 \cdot 10^{-5} \\ 6.74 \cdot 10^{-9} \end{bmatrix}$	$\begin{bmatrix} 3.07 \cdot 10^{-5} \\ \mathbf{2.98 \cdot 10^{-5}} \\ 4.82 \cdot 10^{-3} \\ 2.60 \cdot 10^{-6} \end{bmatrix}$
Π_e	$\begin{bmatrix} 2.39 \cdot 10^{-1} \\ 7.23 \cdot 10^{-2} \\ 9.69 \cdot 10^{-2} \\ 2.68 \cdot 10^{-1} \end{bmatrix}$	$\begin{bmatrix} 1.49 \cdot 10^{-2} \\ \mathbf{7.48 \cdot 10^{-2}} \\ 3.51 \cdot 10^{-2} \\ 2.67 \cdot 10^{-1} \end{bmatrix}$	$\begin{bmatrix} 1.56 \cdot 10^{-1} \\ \mathbf{5.53 \cdot 10^{-3}} \\ 6.21 \cdot 10^{-1} \\ 2.17 \cdot 10^{-1} \end{bmatrix}$
Σ_{Π_e}	$6.76 \cdot 10^{-1}$	$\mathbf{3.92 \cdot 10^{-1}}$	**1.00**

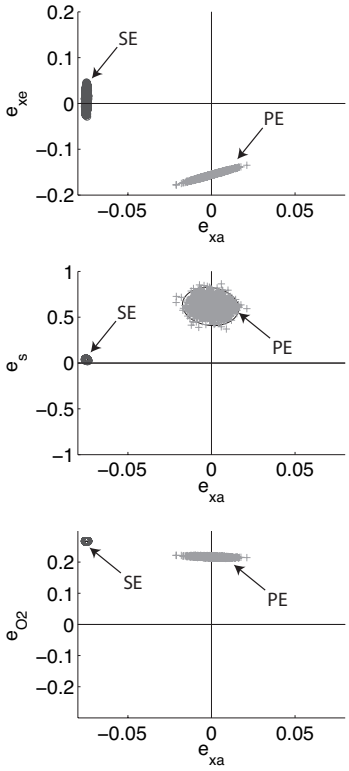

Fig. 1. Preferential estimation of x_a: Error distribution plots of e_{x_e}, e_s and e_{O_2} versus e_{x_a} for $10h \leq t \leq 196h$.

VI. CONCLUSIONS

This paper has addressed the problem of preferential estimation for linear uncertain systems and shown that the potential for better estimation of preferred variables exists due to the presence of bias. Both bias-bias and bias-variance tradeoffs are possible. The preferential approach was applied to a filamentous fungal fermentation process, where the error along the soft output could be reduced by a factor 15, while the overall error increased by a factor 3.

The directions for future investigation include (i) the consideration of a wider class of systems, i.e. not only systems at steady state but also in the transient mode, for which additional estimator structures need to be considered; (ii) the combination of these ideas with calibration techniques so as to provide viable run-to-run estimation schemes for preferential estimation.

VII. ACKNOWLEDGMENTS

This paper presents the results of the Knowledge-driven Batch Production (BatchPro) European Project HPRN-CT-2000-00039.

REFERENCES

[1] D. Dochain, "State and parameter estimation in chemical and biochemical processes : A tutorial," *J. Process Contr.*, vol. 13, pp. 801–818, 2003.

[2] F. J. Doyle, "Nonlinear inferential control for process applications," *J. Process Contr.*, vol. 8, pp. 339–353, 1998.

[3] P. S. Maybeck, *Stochastic Models, Estimation and Control*. Academic Press, 1979, vol. 1.

[4] R. G. Brown and P. Y. C. Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, 2nd ed. John Wiley, 1992.

[5] B. Srinivasan and D. Bonvin, "Dynamic optimization under uncertainty via NCO tracking: A solution model approach," in *BatchPro Symposium*, Poros, Greece, 2004, pp. 17–35.

[6] R. Kalman, "A new approach to linear filtering and prediction problems," *Transaction of the ASME-Journal of Basic Engineering*, pp. 35–45, 1960.

[7] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. Academic Press, 1970.

[8] M. S. Grewal and A. P. Andrews, *Kalman Filtering. Theory and Practice*. Prentice Hall, 1993.

[9] L. Bodizs, B. Srinivasan, and D. Bonvin, "Preferential estimation via tuning of the Kalman filter," in *Dynamics and Control of Process Systems - 7*, Cambridge, Massachusetts, July 2004.

[10] H. Martens and T. Naes, *Multivariate Calibration*. John Wiley, 1989.

[11] T. Agger, B. Spohr, M. Carlsen, and J. Nielsen, "Growth and product formation of aspergillus oryzae during submerged cultivations: Verification of a morphologically structured model using fluorescent probes," *Biotech. Bioeng.*, vol. 57, pp. 321–329, 1998.

[12] M. Titica, L. Bodizs, F. Lei, B. Srinivasan, D. Dochain, and D. Bonvin, "Modeling and optimization of fed-batch filamentous fungal fermentation," in *BatchPro Symposium*, Poros, Greece, 2004, pp. 55–65.