Preprints of the 12th IFAC Symposium on Computer Applications in Biotechnology
The International Federation of Automatic Control
16-18, 2013, December. Mumbai, India

Virtual output estimation in a bioreactor using a generalized super-twisting algorithm

Alejandro Vargas * **Alain Vande Wouwer** **
Jaime A. Moreno *

* *Instituto de Ingeniería, Universidad Nacional Autónoma de México, Querétaro, Mexico (e-mails: {avargasc,jmorenop}@ii.unam.mx).*
** *Automatic Control Department, Université de Mons, Mons, Belgium (e-mail: alain.vandewouwer@umons.ac.be)*

Abstract: This paper presents the application of a generalized super-twisting algorithm to estimate a critical variable needed for implementing real time optimization feedback controllers in a certain class of bioreactors. A bank of redundant super-twisting observers designed for equivalent systems under a state transformation is used. The procedure is motivated by the fed-batch cultivation of a strain of *E. coli* that has overflow metabolism, where a useful variable to estimate is the difference between two reaction rates. The results show the applicability of the procedure and illustrates some compromises between the accuracy and the measurement noise.

Keywords: Biotechnology; observers; super-twisting algorithm; fermentation processes; parameter estimation; variable-structure systems.

1. INTRODUCTION

Under fed-batch growth, the catabolism of certain strains of *E. coli* has a limited energy production for cell growth and division originating from a limited capacity to oxidize the main substrate, usually glucose. The excess part of this nutrient then can follow another metabolic pathway more commonly known as fermentation, producing a by-product which is generally a growth inhibitor, e.g. acetate. This is called overflow metabolism (Crabtree, 1929) and there exist validated mathematical descriptions of the phenomenon (Rocha, 2003).

The mathematical model considered here is a simplification of the one presented by Dewasme et al. (2011) and it involves only the biomass (X), the substrate (S), the dissolved oxygen (O) and CO_2 (C) concentrations:

$$\dot{\xi} = K\rho X - \xi u + f, \qquad \xi(0) = \xi_0, \tag{1}$$

where $\xi = [X, O, C, S]^T$ is the state vector, $K \in \mathbb{R}^{4\times 2}$ contains the pseudo-stoichiometric coefficients, $u = Q_{\text{in}}/V$ is the dilution rate, and $f = [0, f_{\text{OTR}}, -f_{\text{CTR}}, S_{\text{in}}u]^T$ is the vector of gas and mass flow rates in and out of the reactor; f_{OTR} and f_{CTR} are the oxygen and CO_2 transfer rates, respectively, while S_{in} is the inflow substrate concentration. The by-product (acetate) is not included in this model and thus the specific reaction rate vector is only given by $\rho = [r_1, r_2]^T$, where $r_1(t)$ is the respiration rate and $r_2(t)$ is the fermentation rate:

$$r_1 = \min\left(r_S, r_S^*\right), \qquad r_2 = \max\left(0, r_S - r_S^*\right). \tag{2}$$

These definitions of the rates lead to either a respirative regime, when $S(t) < S^*$, or a respiro-fermentative regime when $S(t) > S^*$. Substrate (glucose) is consumed in any case with rate r_S, which follows a Monod model,

$$r_S = \mu_S\left(\frac{S}{S + K_S}\right). \tag{3}$$

Although it is not modeled here, acetate is produced in respiro-fermentative regime, but only slowly consumed in respirative regime. This by-product is inhibitory for biomass growth, as its accumulation decreases the critical consumption rate r_S^* as follows:

$$r_S^* = \mu_S^*\left(\frac{O}{O + K_O}\right)\left(\frac{K_{iP}}{K_{iP} + P}\right), \tag{4}$$

where μ_S^*, K_O, and K_{iP} are constants and $P(t)$ is the product (acetate) concentration. The best strategy for enhancing biomass production without accumulating acetate is thus to operate in the boundary between regimes, i.e. maintaining $S(t)$ at the critical value S^*, where $r_S(S^*) = r_S^*$. However, S^* is not *a priori* known and furthermore it may change (decrease) slowly during the fed-batch cultivation if acetate accumulates.

For this reason real-time optimization (RTO) schemes have been proposed by Dewasme et al. (2011) and Vargas et al. (2012), which use a virtual output that is a linear combination of the two reaction rates in the process. Other schemes also assume the measurement of this output (Santos et al., 2012). This output is given by

⋆ This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office (BELSPO). The scientific responsibility rests with its authors. A. Vargas has been beneficiary of a post-doc fellowship granted by BELSPO and DGAPA-UNAM. The authors also gratefully acknowledge the support of FNRS (Belgium) and CONACYT (Mexico) in the framework of a bilateral collaboration project (192578). The work of J. Moreno has been financed by PAPIIT, UNAM, grant IN111012, and Fondo de Colaboracin del II-FI, UNAM, IISGBAS-144-2012.

$$y = r_1 - r_2 = \gamma^T \rho \qquad \gamma = [1, -1]^T. \tag{5}$$

As a function of S, this output reaches a maximum whenever $S = S^*$; then $y = r_S^*$. The controllers proposed manipulate the dilution rate u to keep $y(t)$ near its optimum value at r_S^*.

The problem is that this output cannot be measured directly. Dewasme et al. (2011) propose an algebraic approach to estimate it under the assumption of quasi-steady-state, perfect knowledge of the matrix K, and measurement or knowledge of all the signals in $f(t)$, but this approach has three disadvantages: (i) it will not work during transients, (ii) it is based on the inversion of a 3×3 submatrix of K, so it is very sensitive to the exactness of the model, and (iii) it estimates the quantity $(r_1 - r_2)X$, so it will also become sensitive to the noise present in $X(t)$ if we insist on estimating $y = r_1 - r_2$.

This work presents instead the use of a bank of redundant super-twisting observers to robustly estimate $y(t)$, using only the on-line measurements of X, O, and C (using probes), as well as f_{OTR} and f_{CTR} (using a gas analyzer). Measurement of S is not needed, nor the knowledge of S_{in}. Knowledge of the matrix K is assumed, but if this is not possible, the methodology can be used to build an estimate of its underlying structure, given recorded data.

2. STATE TRANSFORMATION

Let us assume that $n > m$ in the matrix $K \in \mathbb{R}^{n\times m}$ of system (1) and $\text{rank}(K) = m$. Consider $p \geq 1$ solutions a_j, $j = 1, \ldots, p$ to the following linear equation:

$$K^T a_j = \gamma. \tag{6}$$

Then we can build the matrix

$$A^T = [a_1 \ a_2 \ \cdots \ a_p] \in \mathbb{R}^{n\times p}. \tag{7}$$

Note that each column of the matrix $K^T A^T$ is equal to the vector γ and therefore each element of the vector $AK\rho$ becomes $y = \gamma^T \rho$. Define a state transformation

$$\zeta = A\xi, \tag{8}$$

such that $\dot\zeta = A\dot\xi = AK\rho X - A\xi u + Af$ and thus system (1) becomes

$$\dot\zeta = l_p y X - \zeta u + Af, \qquad \zeta(0) = A\xi_0, \tag{9}$$

$$l_p = [1 \ 1 \ \cdots \ 1]^T \in \mathbb{R}^p. \tag{10}$$

System (9) is not diffeomorphic to system (1), because the transformation (8) is not injective, since A is always rank deficient, even if $p \geq n$. We use the system (9) to design a bank of p observers to build several estimates of the same signal, namely $\hat y(t)$. The redundancy in the estimations will prove to be beneficial.

3. BANK OF SUPER-TWISTING OBSERVERS

We want to estimate y and we may assume that we can measure ζ, either because the whole state ξ is available or because we managed to choose A such that $\zeta = A\xi$ can be built with the available measured states (the corresponding entries in each a_j^T are zero for the unmeasured states). Let us now choose some ζ_j, $j \in \{1, \ldots, p\}$ and assume that the dynamics of y may be unknown, but $\dot y$ is bounded, so we can artificially consider y as an unmeasured state and propose that

$$\dot\zeta_j = yX - \zeta_j u + a_j^T f \tag{11}$$

$$\dot y = \delta_y \qquad \delta_y(t) \in [\underline\delta, \overline\delta] \tag{12}$$

To estimate y we can use a variation of the generalized super-twisting observer (Moreno, 2011) that considers the explicit contribution of the biomass, as follows:

$$\dot{\hat\zeta}_j = -\ell_1 \phi_1(\tilde\zeta_j) + \hat y_j X - \zeta_j u + a_j^T f, \quad \hat\zeta_j(0) = a_j^T \hat\xi_0 \tag{13}$$

$$\dot{\hat y}_j = -\ell_2 \phi_2(\tilde\zeta_j), \qquad \hat y_j(0) = \hat y_0, \tag{14}$$

where $\tilde\zeta_j = \hat\zeta_j - \zeta_j$ is the observation error. The observer gains are ℓ_1 and ℓ_2, and

$$\phi_1(x) = \mu_1 |x|^{p_1} \text{sign}(x) + \mu_2 |x|^{p_2} \text{sign}(x), \tag{15}$$

$$\begin{aligned}\phi_2(x) =& p\mu_1^2 |x|^{2p_1 - 1} \text{sign}(x) + \\ &+ (p_1 + p_2)\mu_1\mu_2 |x|^{p_1 + p_2 - 1} \text{sign}(x) + \\ &+ p_2 \mu_2^2 |x|^{2p_2 - 1} \text{sign}(x) = \phi_1'(x)\phi_1(x),\end{aligned} \tag{16}$$

where $\frac{1}{2} \leq p_1 < 1 \leq p_2$ and $\mu_1, \mu_2 \geq 0$ are design parameters. The dynamics of the observation error are given by:

$$\dot{\tilde\zeta}_j = -\ell_1 \phi_1(\tilde\zeta_j) + \tilde y_j X \qquad \tilde\zeta_j = \hat\zeta_j - \zeta_j \tag{17}$$

$$\dot{\tilde y}_j = -\ell_2 \phi_2(\tilde\zeta_j) - \delta_y, \qquad \tilde y_j = \hat y_j - y. \tag{18}$$

This almost has the form of the generalized super-twisting algorithm, except for the growing influence of $X(t)$ in the first state estimation. Using Lyapunov arguments similar to those presented by (Moreno, 2011), we can show that gains ℓ_1 and ℓ_2 can be found such that the origin of this system is asymptotically stable. A sketch of the proof of this claim is shown in the Appendix.

Choosing $\mu_1 = 0$, $\mu_2 > 0$ and $p_2 = 1$ leads to a linear high gain observer, while choosing $\mu_1 > 0$, $\mu_2 = 0$ and $p_1 = \frac{1}{2}$ leads to the classical super-twisting algorithm (Moreno, 2011). In fact, following the proof therein, it is possible to show that if $p_1 = \frac{1}{2}$ and $p_2 > 1$, then the origin of (17)-(18) is uniformly and exactly stable, converging in finite-time, even despite a persisting, but bounded perturbation $\delta_y(t)$.

De Battista et al. (2012) and Moreno and Guzmán (2011) also consider explicitly the bilinear term yX, but in both cases the correction term in the dynamics of $\hat y$ includes somehow $X(t)$. This makes these approaches very sensitive to the noise that may be present in X. If we know that y is not influenced by X, it does not make sense to make the gain in the correction term depend on this variable.

Using the estimated common $\hat y_j$ we could also construct the following exponential observers for the other states in ζ, i.e. for each $i = 1, \ldots, (j-1), (j+1), \ldots, p$ we propose

$$\dot{\hat\zeta}_i = -\beta\tilde\zeta_i + \hat y_j X - \zeta_i u + a_i^T f, \qquad \hat\zeta_i(0) = a_i^T \hat\xi_0, \tag{19}$$

where $\tilde\zeta_i = \hat\zeta_i - \zeta_i$ and $\beta > 0$ is an additional gain. Upon convergence of $\hat y_j$, the error dynamics are $\dot{\tilde\zeta}_i \approx -\beta\tilde\zeta_i$ and thus $\tilde\zeta_i \to 0$ exponentially. Later we will see that this is really not so useful.

In the same way that we have chosen some ζ_j, we could build a bank of similar observers for each $j = \{1, \ldots, p\}$. Each one will provide an estimate $\hat y_j(t)$ of the same $y(t)$ and we can use this redundancy to tune the observer, as will be discussed later.

4. CHOICE OF THE TRANSFORMATION MATRIX

Each transformed state ζ_j is a linear combination of the states in ξ via $\zeta_j = a_j^T \xi$, with a_j a solution to $K^T a = \gamma$. In this section we present a methodology to choose the a_j such that all transformed states are built giving approximately the same weight to each ξ_i involved. Given the rank assumption on K, a solution to system (6) is given by

$$a_j = a_0 + N_K b_j, \tag{20}$$

where $N_K \in \mathbb{R}^{n\times(n-m)}$ is a matrix whose columns form a basis for the nullspace of K^T and $b_j \in R^{n-r}$ can be any vector. The vector a_0 is any solution to the linear equation (6), but let us consider the least norm solution:

$$a_0 = (K^T)^+ \gamma = V_K \Sigma_K^{-1} U_K^T \gamma. \tag{21}$$

where $(K^T)^+$ is the pseudoinverse of K^T, which has been calculated using a singular value decomposition (SVD):

$$K^T = U_K \left[\Sigma_K \;\; 0_{m\times(n-m)}\right] \left[V_K \;\; N_K\right]^T, \tag{22}$$

where $U_K \in \mathbb{R}^{m\times m}$, $V_K \in \mathbb{R}^{m\times n}$, and Σ_K is a diagonal matrix with the singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m$. Note that the columns of N_K form a basis for the nullspace of K^T; since $[V_K \;\; N_K]$ is orthogonal, then $K^T N_K = (U_K \Sigma_K V_K^T) N_K = 0$.

In order to give the same weight to each ξ_i when defining the transformed states ζ_j, we can design A such that its SVD has equal singular values, leading to a matrix A^T with unity condition number. Notice that (20) implies that

$$A^T = a_0 l_p^T + N_K B^T, \tag{23}$$

$$B^T = [b_1 \;\; b_2 \;\; \cdots \;\; b_p] \in R^{(n-m)\times p}. \tag{24}$$

If we define $u_0 = \frac{1}{\|a_0\|} a_0$ and $v_0 = \frac{1}{\sqrt{p}} l_p$ we can write $a_0 l_p^T = u_0 \sigma_0 v_0^T$ with $\sigma_0 = \sqrt{p}\|a_0\|$. By the choice (21) of a_0 we have that $N_K^T a_0 = 0$ by the orthogonality of $[V_K \;\; N_K]$. This implies that $[u_0 \;\; N_K]$ is an orthonormal matrix, so now define $\Sigma_0 = \sigma_0 I_{n-m}$ and consider the decomposition

$$A^T = u_0 \sigma_0 v_0^T + N_K \Sigma_0 \bar{L}^T = [u_0 \;\; N_K] \begin{bmatrix} \sigma_0 & 0 \\ 0 & \Sigma_0 \end{bmatrix} \left[v_0 \;\; \bar{B}\right]^T,$$

where $\bar{B}$ is an orthonormal matrix such that its columns are also orthogonal to v_0. Then this decomposition qualifies as an SVD of A^T and $\bar{B}$ can be found by choosing some full rank matrix $M \in \mathbb{R}^{p\times(n-m)}$ and performing a Gram-Schmidt orthogonalization of the matrix $[v_0 \; M]$ with respect to its columns, but care should be taken in the choice of M so that A^T does not have repeated columns.

5. ESTIMATION OF K

So far, we have provided a method to design a bank of super-twisting observers, each one estimating the same second state, namely $y(t)$, using measurements of the several $\zeta_j(t)$; see (13)-(14). Since each ζ_j depends ultimately on K, we should therefore expect the estimates $\hat{y}_j$ to be basically the same if K is correct. If instead of the true K we use $\widehat{K} = K + \widetilde{K}$ with $\widetilde{K} \neq 0$ for the design and implementation of the bank of observers, differences between the estimates $\hat{y}_j(t)$ are expected. However, we show now that we can use the redundancy for eliminating the bias between these estimations.

Solving $\widehat{K}^T \hat{a}_j = \gamma$ and using the transformations $\zeta_j = \hat{a}_j^T \xi$, $j = 1, \ldots, p$, leads to a system with

$$\dot{\zeta}_j = yX - \hat{a}_j^T \widetilde{K} \rho X - \zeta_j u + \hat{a}_j^T f. \tag{25}$$

There is now a perturbation term $\hat{a}_j^T \widetilde{K} \rho X$ and we get the error system

$$\dot{\tilde{\zeta}}_j = -\ell_1 \phi_1(\tilde{\zeta}_j) + \tilde{y}_j X + \hat{a}_j^T \widetilde{K} \rho X, \quad \tilde{\zeta}_j = \hat{\zeta}_j - \zeta_j \tag{26}$$

$$\dot{\tilde{y}}_j = -\ell_2 \phi_2(\tilde{\zeta}_j) + \delta_y(t), \qquad \tilde{y}_j = \hat{y}_j - y. \tag{27}$$

Because of the nonvanishing perturbation $\hat{a}_j^T \widetilde{K} \rho X$, the super-twisting algorithm will lead to practical stability in the sense that $\tilde{\zeta}_j$ and (more importantly) $\tilde{y}_j$ will now only converge to a neighborhood of the origin.

Consider the case when $\rho(t) = \bar{\rho}$ is constant. Practical application of the super-twisting observer in this case shows that after some finite convergence time the region to which error trajectories converge is such that

$$|\tilde{\zeta}_j(t)| \leq \epsilon_\zeta, \qquad |\tilde{y}_j(t) - \bar{y}_j| \leq \epsilon_y, \tag{28}$$

with ϵ_ζ and ϵ_y very small, i.e. the error $\tilde{\zeta}_j$ converges to almost zero, while $\tilde{y}_j$ converges to some constant value $\bar{y}_j$ (in a practical sense). Since ϵ_ζ is small, then assuming that $\tilde{\zeta}_j$ remains practically constant after convergence (in finite time) we can find a practical steady state by setting $\dot{\tilde{\zeta}}_j = 0$ in (26). Since $\phi_1(\epsilon_\zeta) \approx 0$ this leads to

$$\tilde{y}_j(t) \to \bar{y}_j = -\hat{a}_j^T \widetilde{K} \bar{\rho}. \tag{29}$$

Now consider the p observers that also include the exponential part (19), whose states are $\hat{\zeta}_{j,i}$ for $j, i = 1, \ldots, p$, $i \neq j$, The dynamics of the observation errors $\tilde{\zeta}_{j,i} = \hat{\zeta}_{j,i} - \zeta_i$ for these other observers are given by

$$\dot{\tilde{\zeta}}_{j,i} = \tilde{y}_j X + \hat{a}_i^T \widetilde{K} \rho X - \beta \tilde{\zeta}_{j,i}. \tag{30}$$

Assume that the dynamics of $X(t)$ are much slower than those of $\tilde{\zeta}_{j,i}(t)$. Since $\beta > 0$ and $\widetilde{K}$ is constant, exponentially $\tilde{\zeta}_{j,i}(t) \to \bar{\zeta}_{j,i} X(t)$, with

$$\bar{\zeta}_{j,i} = \frac{1}{\beta}\left(\bar{y}_j + \hat{a}_i^T \widetilde{K} \bar{\rho}\right) = \frac{1}{\beta}\left(\bar{y}_j - \bar{y}_i\right), \tag{31}$$

where we have used the equality in (29). Note that $\bar{\zeta}_{j,i} = -\bar{\zeta}_{i,j}$ and in fact we gain no more information by using the additional exponential observers, but nevertheless they could provide some additional robustness to the identification method explained below.

It is clear that if $\widetilde{K} = 0$, the $p(p-1)/2$ steady-state errors

$$\tilde{y}_{j,i} = \bar{y}_j - \bar{y}_i = \left(-\hat{a}_j^T + \hat{a}_i^T\right) \widetilde{K} \bar{\rho} \tag{32}$$

will all be zero. This leads naturally to the idea of using an *a posteriori* numerical optimization procedure to minimize

$$J(\widehat{K}) = \sum_{j=1}^{p-1} \sum_{i=j+1}^{p} |\bar{y}_j - \bar{y}_i|^2 + \beta \sum_{j=1}^{p} \sum_{i=1}^{p} \tilde{\zeta}_{j,i}^2 \tag{33}$$

or a similar criterion. In this sense, the idea is to collect data e.g. for one cultivation using an observer with some $\widehat{K}$. Then, having identified a convergence time T_c, the numerical optimization searches for the $\widehat{K}_*$ that minimizes (33) by simulating the implementation of the bank of observers in open loop with the collected data. The next cultivation would use $\widehat{K}^*$ in the observer implementation.

Unfortunately, there is no unique $\widehat{K}$ leading to $J = 0$. In fact, this happens for all $\widehat{K}$ with the same SVD structure as the true K. To see this consider that $K^T = U_K \Sigma_K V_K^T$ and $\widehat{K}^T = U_K \left(\Sigma_K + \widetilde{\Sigma}_K\right) V_K^T$ with $\widetilde{\Sigma}_K = \mathrm{diag}(\tilde{\sigma}_K)$; then

$$\widetilde{K} = \widehat{K} - K = V_K \widetilde{\Sigma}_K U_K^T. \tag{34}$$

Using the construction (20) of each a_j and the definition (34) in (32) leads to

$$\tilde{y}_{j,i} = -(b_j - b_i)^T \underbrace{N_K^T V_K}_{=0} \widetilde{\Sigma}_K U_K^T \bar{\rho} = 0 \tag{35}$$

for any $\widetilde{\Sigma}_K$. This implies that there exist a family of matrices $\widehat{K}$ for which the observer implementation will lead to practically $\hat{y}_j(t) = \hat{y}_i(t)$ after $t \geq T_c$ for all $j \neq i$, but $\hat{y}_j(t) \neq y(t)$.

Even though an unrestricted optimization may not be successful in finding the true value of K, it will find the SVD structure, and this can be used to narrow the search space. For example, if we know $K_{\min}$ and $K_{\max}$ such that $K_{\min} \prec K \prec K_{\max}$, then we can add a penalizing term in the optimization criterion (33) for $\widehat{K}$'s out of bounds.

6. SIMULATIONS

A more complex model than the one explained in Section 1 was used for simulation using Matlab 7.12.0.635 (R2011a). It considers the production of acetate in respiro-fermentative regime and its low consumption in respirative regime, which makes $S^*(t)$ become time-varying. A local PI controller for regulating the dissolved oxygen concentration is used. Details of this model, including the values of the parameters used, are found in (Santos et al., 2012).

The following input makes $S(t)$ converge to $S^*(t)$

$$u^*(t) = \frac{X(t) r_S^*(t)}{S_{\text{in}} - S^*(t)}. \tag{36}$$

Since this input is not implementable, other controllers have been proposed that will approximate it using the information from the virtual output that we are estimating (Dewasme et al., 2011; Vargas et al., 2012; Santos et al., 2012). Therefore, to test the observers we generate the data by constructing the optimal input $u^*(t)$ as in (36) and perturb it with a quasi-periodic signal $p(t)$ in order to mimic what such controllers would do:

$$u(t) = (1 + p(t))\, u^*(t). \tag{37}$$

Since we are assuming that $S(t)$ cannot be measured and S_{in} is not known, we eliminate the last state in ξ and use only a reduced model with the first three states. Furthermore, we also assume that $O(t)$ is constant and known (it is regulated by the PI controller) and $C(t)$ remains at a very low known and constant value (the solubility of CO_2 is very low), so there is no need to measure them. We do measure on-line the biomass $X(t)$, as well as $f_{\text{OTR}}(t)$ and $f_{\text{CTR}}(t)$ and have added white measurement noise to them with standard deviations of up to 0.01 g/L for $X(t)$ and 0.015 g/L/h both for f_{OTR} and f_{CTR}.

Figure 1 shows the results of a simulation where the dissolved oxygen $O(t)$ is regulated at 2.5 mg/L using a simple PI controller (which indirectly modifies the f_{OTR} via the $k_L a$ value), while f_{OTR} and f_{CTR} are directly calculated from the model equations assuming steady state. Notice the typical exponential growth of $X(t)$ and the time-varying nature of the maximum $y^* = r_S^*$ and consequently of the desired setpoint at S^*.

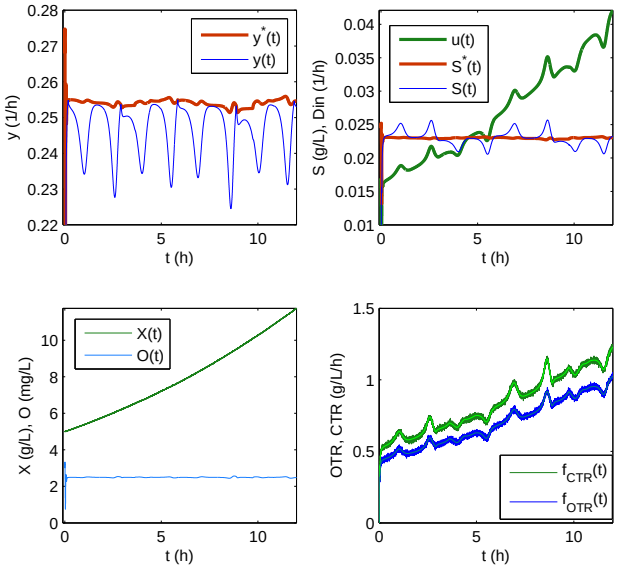

Fig. 1. Simulations: virtual output y and its optimal value y^* (top-left); substrate S, optimal value S^* and dilution rate u (top-right); biomass X and dissolved oxygen O (bottom-left); f_{OTR} and f_{CTR} measurements with noise (bottom-right).

In order to design the bank of observers, we need the matrix K, which has $n = 3$ rows and $m = 2$ columns and is given numerically by:

$$K^T = \begin{bmatrix} 0.316 & -0.339 & 0.406 \\ 0.040 & -0.472 & 0.757 \end{bmatrix}. \tag{38}$$

An SVD is given by the following matrices and vector:

$$U_K = \begin{bmatrix} -0.548 & -0.836 \\ -0.836 & 0.548 \end{bmatrix}, \quad \sigma_K = \begin{bmatrix} 1.054 \\ 0.255 \end{bmatrix},$$
$$V_K = \begin{bmatrix} -0.196 & -0.950 \\ 0.551 & 0.100 \\ -0.812 & 0.297 \end{bmatrix}, \quad N_K = \begin{bmatrix} 0.244 \\ 0.829 \\ 0.503 \end{bmatrix}. \tag{39}$$

The nullspace is of dimension 1 and the least-norm solution vector a_0 and thus σ_0 are given, respectively, by

$$a_0 = [5.0943 \ \ -0.3894 \ \ -1.8305]^T, \quad \sigma_0 = 7.6752. \tag{40}$$

The transformation matrix A is built as indicated previously. In the observers the ϕ_1 function has parameters $p_1 = 0.5$, $p_2 = 1$, $\mu_1 = 1$ and $\mu_2 = 0.2$, i.e. a classical super-twisting observer with an additional linear correction term. The simulations were made implementing a discretized version of the observers using a forward Euler approximation using a sample time of $T_s = 6$ s.

Figure 2 shows the result of a bank of $p = 7$ observers under perfect system knowledge (with $\widehat{K} = K$). The observer gains are $\ell_1 = 1.5$, $\ell_2 = 0.5$ and $\beta = 20$. As expected, all observers in the bank converge to the same signal, the virtual output $y(t)$. The purple line

shows the mean $\bar{\hat{y}}(t)$ of the seven observers implemented. The top figure shows a simulation without noise in the measurement of X, while the bottom figure shows it when noise is present. In both cases noise is present in both f_{OTR} and f_{CTR}. It is interesting to notice that the observer actually acts as a filter under noise in the measured state, delaying the estimation, but also damping the influence of noise coming from $f(t)$. If we increase the noise level in X, the delay in the estimation also increases. We can increase the observer gains to diminish the delay, especially ℓ_2, but the noise is less damped and the estimation would become useless for implementing a controller. There is thus a clear compromise that would later have to be further studied when implementing an output-feedback controller.

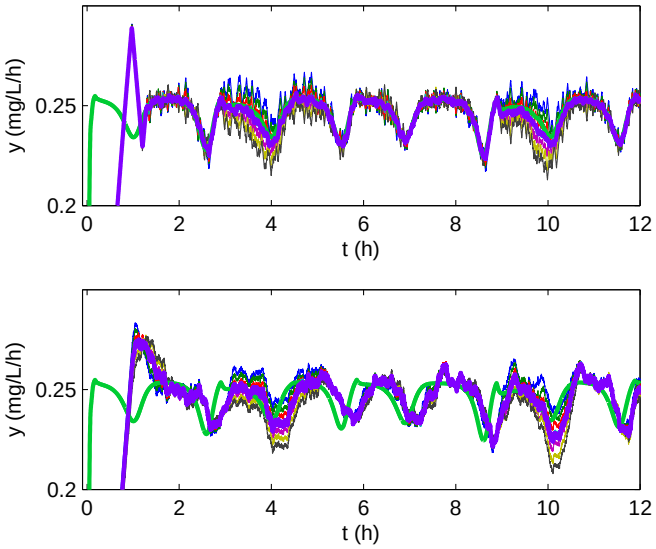

Fig. 2. Bank of observers under perfect system knowledge, without (top) and with noise in X (bottom); the true y is shown green, while the mean $\bar{y}$ of all $\hat{y}_j$ is shown purple (same observer parameters).

Figure 3 shows the simulation for the same data set as above, but using $\widehat{K} \neq K$. We restrict $\widehat{K}$ to have elements between:

$$K_{\min} = \begin{bmatrix} 0.253 & 0.036 \\ -0.424 & -0.636 \\ 0.263 & 0.568 \end{bmatrix}, \quad K_{\max} = \begin{bmatrix} 0.380 & 0.044 \\ -0.255 & -0.306 \\ 0.547 & 0.946 \end{bmatrix}.$$

The starting $\widehat{K}_0$ is chosen at random within these bounds and the result of applying the bank of observers is shown on the top graph of figure 3. The same gains and observer parameters were used as before. As expected, the estimates $\hat{y}_j(t)$ are different and do not converge to the true value of $y(t)$. Using this matrix as initial estimate, we used the (unconstrained) `fminsearch` algorithm under Matlab for the numerical optimization procedure with a large penalization for matrices outside the sought range. The bottom graph of figure 3 shows the result after the optimization, with $\widehat{K}_*$. All the estimates now converge to their mean $\bar{\hat{y}}(t)$, but not to the true $y(t)$. Notice, however, that the individual estimates $\hat{y}_j$, $j = 1, \ldots, 7$ are noisy, but their average has reduced this noise. The advantage of using a bank of seven observers instead of one is evident.

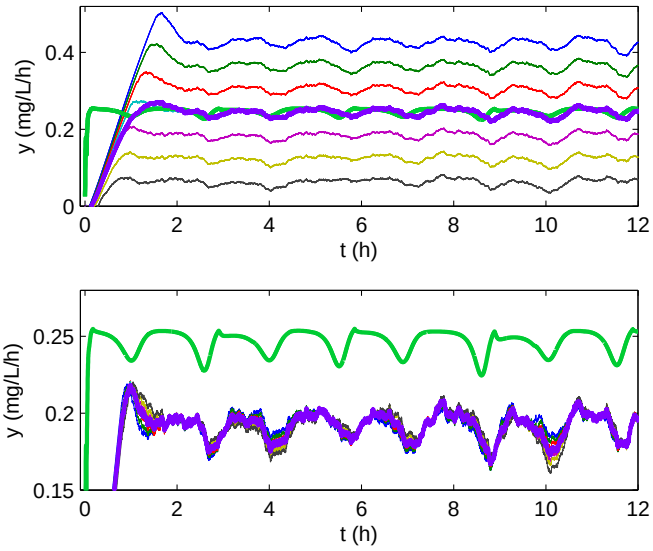

Fig. 3. Bank of observers with $\widehat{K} \neq K$: using $\widehat{K}_0$ before optimization (top) and using $\widehat{K}_*$ after optimization (bottom); the true y is shown green, the estimates $\hat{y}_j$ in different colors with thin lines, while their mean $\bar{\hat{y}}$ is shown purple.

The SVD $\widehat{K}_*^T = \widehat{U}_K \mathrm{diag}(\widehat{\sigma}_K) \widehat{V}_K^T$ is given by

$$\widehat{U}_K = \begin{bmatrix} -0.657 & -0.754 \\ -0.754 & 0.657 \end{bmatrix}, \quad \widehat{\sigma}_K = \begin{bmatrix} 0.991 \\ 0.245 \end{bmatrix},$$

$$\widehat{V}_K = \begin{bmatrix} -0.237 & -0.869 \\ 0.491 & 0.278 \\ -0.838 & 0.409 \end{bmatrix}, \quad \widehat{K}_* = \begin{bmatrix} 0.315 & 0.037 \\ -0.371 & -0.323 \\ 0.470 & 0.692 \end{bmatrix}.$$

It can be verified that the right and left singular vectors of $\widehat{K}_*$ are very similar to those of the true K shown in (39).

From the point of view of a RTO feedback controller such as that proposed by Vargas et al. (2012) or Dewasme et al. (2011), the bias between $\hat{y}(t)$ and $y(t)$ is not important, since these controllers use this virtual output to follow a maximum, so as long as their profiles are similar and their (local) maxima occur simultaneously (some delay is inevitable), the implementation of the controllers could be successful.

We have observed that the noise in $X(t)$ is critical. Although we still have to use it in the term $\hat{y}X$ we can eliminate this measurement in the observer design using only the 2×2 submatrix $\widehat{K}_{OC}$ of $\widehat{K}$ that corresponds to the dynamics of O and C (the last two rows of $\widehat{K}$). However, we can design only one observer, with $a_{OC}^T = \gamma^T \widehat{K}_{OC}^{-1}$ and whose dynamics are basically driven by the inputs from f_{OTR} and f_{CTR}.

Having identified $\widehat{K}_*$, we use it to get $\widehat{K}_{OC}$ and run the resulting super-twisting observer to estimate a single $\hat{y}$. The gains in this case were manually adjusted for the same data base and the results are shown in figure 4. The delay is almost zero (previously due to the noise in X), but the estimate (in blue) now has the influence of the noise present in f_{OTR} and f_{CTR} and is not as good as before. We can reduce this by previously filtering the noise in these

signals (in purple), but then delay appears. In any case, the estimation is not as good for the RTO controller, since the maxima in the signal are not as clearly defined as in the previous case when we use explicitly the signal $X(t)$ and a bank of redundant observers.

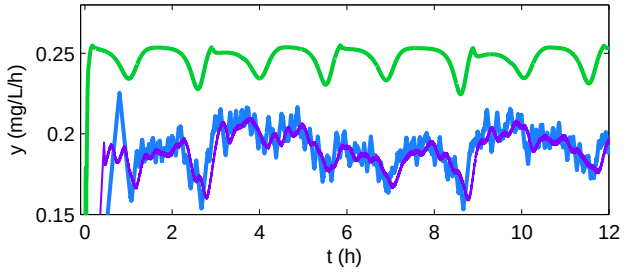

Fig. 4. Single observer without using X; true y shown green, estimate $\hat{y}$ shown with (purple) and without (blue) filtering of the f_{OTR} and f_{CTR} signals.

7. CONCLUSION

For RTO control of certain biotechnological processes, the on-line knowledge of a signal that has a maximum at the desired operating point is needed. Using a bank of redundant super-twisting observers we have proposed a methodology to estimate this virtual output. Through simulations in a system with overflow metabolism, the convenience of this redundancy has been shown.

REFERENCES

Crabtree, H. (1929). Observations on the carbohydrate metabolism of tumors. *Biochemical Journal*, 3, 536–545.

De Battista, H., Picó, J., Garelli, F., and Navarro, J. (2012). Reaction rate reconstruction from biomass concentration measurement in bioreactors using modified second-order sliding mode algorithms. *Bioprocess and Biosystems Engineering*, 35, 1615–1625.

Dewasme, L., Srinivasan, B., Perrier, M., and Vande Wouwer, A. (2011). Extremum-seeking algorithm design for fed-batch cultures of microorganisms with overflow metabolism. *Journal of Process Control*, 21(7), 1092–1104.

Moreno, J. (2011). *Sliding Modes after the First Decade of the 21st Century*, chapter Lyapunov approach for analysis and design of second order sliding mode algorithms, 113–150. Springer.

Moreno, J. and Guzmán, E. (2011). A new recursive finite-time convergent parameter estimation algorithm. In *18th IFAC World Congress*. Milan, Italy.

Rocha, I. (2003). *Model-based strategies for computer-aided operation of a recombinant E. coli fermentation*. Ph.D. thesis, University of Minho, Portugal.

Santos, L., Dewasme, L., Coutinho, D., and Vande Wouwer, A. (2012). Nonlinear model predictive control of fed-batch cultures of micro-organisms exhibiting overflow metabolism: Assessment and robustness. *Computers and Chemical Engineering*, 39, 143–151.

Vargas, A., Dewasme, L., Moreno, J., and Vande Wouwer, A. (2012). A simple output-feedback controller for fed-batch cultures of microbial strains with overflow metabolism. In *International Symposium on Advanced Control of Chemical Processes ADCHEM 2012*. Singapore, Singapore.

Appendix A. PROOF OF CONVERGENCE

Consider the following perturbed system, which corresponds to the observation error dynamics:

$$\dot{x}_1 = -\ell_1\phi_1(x_1) + a(t)x_2 + \delta_1(t,x), \tag{A.1}$$
$$\dot{x}_2 = -\ell_2\phi_2(x_1) + \delta_2(t,x), \tag{A.2}$$

where the measured signal $a(t)$ is positive and bounded, i.e. $\bar{a} \geq a(t) \geq \underline{a} > 0$ for all t, and ϕ_1 and ϕ_2 as in (15)-(16). The class of perturbations considered satisfy,

$$|\delta_i(t,x)| < g_i|\phi_i(x_1)|, \qquad g_i > 0, \qquad i \in \{1,2\}. \tag{A.3}$$

Consider the transformation

$$z = \begin{bmatrix}\phi_1(x_1)\\ x_2\end{bmatrix} \quad \Longrightarrow \quad \dot{z} = \phi_1'(x_1)\left(Az + \tilde{\delta}\right), \tag{A.4}$$

where $A = A_0 - LC$

$$A_0 = \begin{bmatrix}0 & a(t)\\ 0 & 0\end{bmatrix}, \quad L = \begin{bmatrix}\ell_1\\ \ell_2\end{bmatrix}, \quad C = \begin{bmatrix}1\\ 0\end{bmatrix}, \quad \tilde{\delta} = \begin{bmatrix}\delta_1\\ \frac{1}{\phi_1'(x_1)}\delta_2\end{bmatrix}.$$

Note that $\phi_1'(x_1) = \mu_1 p_1|x_1|^{p_1-1} + \mu_2 p_2|x_1|^{p_2-1} \geq 0$ for all x_1. When $a(t) \equiv 1$, Moreno (2011, Theorem 2) shows that $V_Q = z^T P z$ is a strong, robust Lyapunov function for system (A.1)-(A.2) and the system will be exponentially stable (and finite-time convergent in some cases) if we can show that there exist $P = P^T > 0$, gains ℓ_1 and ℓ_2, $\theta_1 > 0$, $\theta_2 > 0$, and $\epsilon > 0$ such that

$$A^T P + PA + \epsilon P + R + PB\Theta^{-1}B^T P \leq 0, \tag{A.5}$$

with

$$R = \begin{bmatrix}\theta_1 g_1^2 + \theta_2 g_2^2 & 0\\ 0 & 0\end{bmatrix}, \qquad \Theta = \begin{bmatrix}\theta_1 & 0\\ 0 & \theta_2\end{bmatrix}. \tag{A.6}$$

The matrix B is either $B = I$, $B = [1,0]^T$, or $B = [0,1]^T$ depending on whether δ_1 or δ_2 are considered or not (the definitions of Θ and R must change accordingly.

To extend the result of (Moreno, 2011) to the case studied here, we need to satisfy the algebraic Riccati inequality (ARE) (A.5) for all $\underline{a} \leq a(t) \leq \bar{a}$. Using the Schur complement and given some value for a, these ARE's are equivalent to the matrix inequalities $M(a) \leq 0$:

$$\begin{bmatrix}A_0^T P + PA_0 - C^T L^T P - PLC + \epsilon P + R & PB\\ B^T P & -\Theta\end{bmatrix} \leq 0. \tag{A.7}$$

Since the family of matrices $A_0(a)$ is a convex set with vertices at $A_0(\underline{a})$ and $A_0(\bar{a})$, so is the family $M(a)$. We can check the simultaneous feasibility of the family of matrix inequalities by checking only the feasibility of

$$\text{diag}\left[M(\underline{a}), M(\bar{a})\right] \leq 0. \tag{A.8}$$

Replacing the term ϵP with ϵI does not affect its solvability, but makes (A.7) and (A.8) a linear matrix inequality (LMI) in the unknowns P, PL, ϵ and Θ, and thus we can use known efficient algorithms to solve it.

A very important aspect to notice is the guaranteed convergence for the class of perturbations defined by (A.3). These perturbations usually have to vanish at the origin, e.g. $|\delta_1(t,0)| \leq g_1\phi_1(0) = 0$, but when $p_1 = \frac{1}{2}$, then $\phi_2(x_1)$ has a signum function and therefore we allow

$$-g_2\frac{\mu_1^2}{2} \leq \delta_2(t,0) \leq g_2\frac{\mu_1^2}{2}. \tag{A.9}$$

Hence the robustness of the approach.