A sequential convex moving horizon estimator for bioprocesses

Josh Taylor, Alain Rapaport, Denis Dochain

A sequential convex moving horizon estimator for bioprocesses*

Josh A. Taylor[a], Alain Rapaport[b], Denis Dochain[c]

[a]The Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Canada
[b]MISTEA, Université de Montpellier, INRAE, Institut Agro, France
[c]Université Catholique de Louvain, Belgium

Abstract

We design moving horizon state estimators for a general model of bioprocesses. The underlying optimization is nonconvex due to the microbial growth kinetics, which are modeled as nonlinear functions. We relax the nonconvex growth constraints so that the optimization becomes a second-order cone program, which can be solved efficiently at large scales. Unfortunately, solutions to the relaxation can be inexact and thus lead to inaccurate state estimates. To recover feasible, albeit potentially locally optimal solutions, we use the concave-convex procedure, which here takes the form of a sequence of second-order cone programs. We find that the moving horizon state estimators outperform the unscented Kalman filter on numerical examples based on the gradostat and anaerobic digestion when there is high process noise or parameter error.

Keywords: Moving horizon state estimation, second-order cone programming, concave-convex procedure, anaerobic digestion

1. Introduction

Uncertainty arises in bioprocesses from imperfect modeling and inaccurate measurements. States that cannot be precisely known from measurements are typically estimated using observers, also known as state estimators. The nonlinear nature of most bioprocesses, e.g., due to microbial growth, makes this additionally challenging. There are several types of nonlinear state estimation algorithms, including one-step algorithms like the extended and unscented Kalman filters (UKF) [1], asymptotic observers [2], and moving horizon estimators (MHE) [3].

In this paper, we formulate a new MHE for bioprocesses. MHE entails solving an optimization for the last τ values of the state, where τ is the length of the estimation window. It is often more accurate than one-step estimators, which can suffer from the 'short horizon syndrome' [4], but also more computationally demanding—if τ is large, the MHE will be a large, sometimes nonconvex optimization. Starting with [5], there have been numerous applications of MHE to bioprocesses such as wastewater treatment [6, 7, 8] and cell cultures [9, 10].

We use the convex relaxation from [11, 12] to make the MHE more tractable. When the microbial growth is described by the Monod function [13] (with a constant biomass approximation) or the Contois function [14], the relaxed MHE can be written as a second-order cone program (SOCP) [15].

The solution of a relaxation might be outside the original feasible set, and hence infeasible. To recover feasible, albeit locally optimal solutions, in Section 5 we use the concave-convex procedure (CCP) [16, 17]. The advantage of this approach is that it builds directly on the relaxed MHE. In particular, if the relaxation is an SOCP, the the CCP entails solving a sequence of SOCPs.

Our main original contributions are the formulation of a new, relaxed MHE for bioprocesses, and the use of the CCP to obtain feasible solutions. We also use the analytical results from [12] to state conditions under which the relaxation is exact; unfortunately, the objective of the MHE does not in general satisfy the conditions, and the solutions were rarely exact in simulation. In each such case, the CCP was able to recover a feasible solution.

The rest of the paper is organized as follows. We describe the model and setup in Section 2. We state the MHE in Section 3 and formulate a convex relaxation in Section 4. In Section 5 we specialize the CCP to the MHE. In Section 6, we implement the relaxed MHE and the CCP on examples based on the gradostat and anaerobic digestion.

2. Modeling

2.1. Network

The system consists of a set of p well-mixed tanks, $\mathcal{S}$, which are interconnected by mass flow and diffusion. The constant volume of tank i is V_{ii}, and $V \in \mathbb{R}^{p\times p}$ is a diagonal matrix of the volumes. The flows of water into and out of tank i are Q_i^{in} and Q_i^{out}. The flow from tank i to tank j is denoted Q_{ij}. The diffusion between tanks i and j is d_{ij}, where $d_{ij} = d_{ji}$. Let

$$C_{ij} = \begin{cases} 0, & i \neq j \\ Q_i^{\text{in}}, & i = j \end{cases}, \quad L_{ij} = \begin{cases} d_{ij}, & i \neq j \\ -\sum_{k\in\mathcal{S}} d_{ik}, & i = j \end{cases},$$
$$M_{ij} = \begin{cases} Q_{ji}, & i \neq j \\ -Q_i^{\text{out}} - \sum_{k\in\mathcal{S}} Q_{ik}, & i = j \end{cases},$$

and $N = L + M$. L is Laplacian and therefore negative semidefinite. M is compartmental and is invertible if the network is out-

*Funding is acknowledged from the Natural Sciences and Engineering Research Council of Canada and the French LabEx NUMEV, incorporated into the I-Site MUSE.

flow connected [18]. N is hence also invertible if M is outflow connected, and potentially even if M is not outflow connected.

2.2. Microbial growth

We model the microbial growth in the tanks following the notation of Section 1.5 of [19]. In each perfectly mixed tank there are m substrates and biomasses. The process state vector of tank $i \in \mathcal{S}$ is denoted $\xi_i \in \mathbb{R}^m_+$ and the corresponding influent concentration vector $\xi^{\text{in}}_i \in \mathbb{R}^m_+$.

There are $r \leq m$ different biochemical reactions. The reaction kinetics in tank i are collected in the vector $\phi_i(\xi_i) \in \mathbb{R}^r_+$, which, we note, could be different in each tank due to factors such as pH and temperature. Here we focus on the case where the elements of $\phi_i(\xi_i)$ are concave functions with second-order cone (SOC) representations.

We let $\kappa_i \in \mathbb{R}^{m\times r}$ denote the stoichiometric matrix relating the reaction vector, $\phi_i(\xi_i)$, to the evolution of the process state in tank i. The dynamics in tank $i \in \mathcal{S}$ are given by

$$V_{ii}\dot{\xi}_i = V_{ii}\kappa_i\phi_i(\xi_i) - Q^{\text{out}}_i\xi_i - \sum_{j\in\mathcal{S}}(Q_{ij} + d_{ij})\xi_i + Q^{\text{in}}_i\xi^{\text{in}}_i + \sum_{j\in\mathcal{S}}(Q_{ji} + d_{ij})\xi_j.$$

We now express the dynamics in vector form. We omit subscripts to denote stacked vectors, i.e., $\xi = [\xi_1, ..., \xi_p]^\top$ and $\phi(\xi) = [\phi_1(\xi_1), ..., \phi_p(\xi_p)]^\top$. We denote the Kronecker product of A and B by $A \otimes B$, the identity matrix by $I_\alpha \in \mathbb{R}^{\alpha\times\alpha}$, and let $\hat{A} = A \otimes I_m$. Let K be a block diagonal matrix with $\kappa_1, ..., \kappa_p$ on its main diagonal, and note that $K = I_p \otimes \kappa$ if $\kappa_i = \kappa$, i.e., all tanks have the same stoichiometric matrix. The dynamics of the system in vector form are written

$$\hat{V}\dot{\xi} = \hat{V}K\phi(\xi) + \hat{N}\xi + \hat{C}\xi^{\text{in}}. \tag{1}$$

Note that the dynamics can be non-autonomous, in which case $\hat{N}$, $\hat{C}$, and ξ^{in} are time-varying.

2.3. Discretization in time

MHE, as is often the case in state estimation, is usually implemented in discrete time. To discretize (1), we approximate the derivatives with a linear function, which we denote $\mathcal{D}_n$. For example, the implicit Euler method with time step Δ is given by $\mathcal{D}_n[\xi(\cdot)] = (\xi(n) - \xi(n-1))/\Delta$.

Let τ be the length of the estimation window, and let $t + \tau$ be the current time period. We denote the sequence of time periods $\mathcal{N}_t = \{t + 1, ..., t + \tau\}$. The discretized dynamics are given by

$$\hat{V}\mathcal{D}_n[\xi(\cdot)] = \hat{V}K\phi(\xi(n)) + \hat{N}(n)\xi(n) + \hat{C}(n)\xi^{\text{in}}(n) + \omega(n) \tag{2}$$

for $n \in \mathcal{N}_t$. Here $\omega(\cdot) \in \mathbb{R}^{mp}$ is process noise.

2.4. Measurements

For each time period $n \in \mathcal{N}_t$, we have measurements of the form

$$y(n) = \Theta\begin{bmatrix}\xi(n)\\ \phi(\xi(n))\end{bmatrix} + \nu(n),$$

where $\Theta \in \mathbb{R}^{o\times(m+r)p}$ and $\nu(\cdot) \in \mathbb{R}^o$ is sensor noise.

3. State estimation

Given $\Sigma \succ 0$, let $\|x\|^2_\Sigma = x^\top\Sigma^{-1}x$. The state estimate for each $n \in \mathcal{N}_t$ is the solution of the below optimization problem.

$$\begin{aligned}
\mathcal{P}(t) \qquad & \min_{\xi(\cdot),T(\cdot),\nu(\cdot),\omega(\cdot)} \left\|\xi(t) - \bar{\xi}(t)\right\|^2_{\Sigma_0} \\
& \quad + \sum_{n\in\mathcal{N}_t} \|\nu(n)\|^2_{\Sigma_\nu} + \|\omega(n)\|^2_{\Sigma_\omega} && \text{(3a)}\\
\text{subject to} \quad & T(n) = \phi(\xi(n)), \quad n \in \mathcal{N}_t, && \text{(3b)}\\
& \hat{V}\mathcal{D}_n[\xi(\cdot)] = \hat{V}KT(n) + \hat{N}(n)\xi(n) \\
& \quad + \hat{C}(n)\xi^{\text{in}}(n) + \omega(n), \quad n \in \mathcal{N}_t, && \text{(3c)}\\
& y(n) = \Theta\begin{bmatrix}\xi(n)\\ T(n)\end{bmatrix} + \nu(n), \quad n \in \mathcal{N}_t, && \text{(3d)}\\
& (\xi(\cdot), T(\cdot)) \in \Omega. && \text{(3e)}
\end{aligned}$$

The estimator proceeds by incrementing t by one in each time period and resolving $\mathcal{P}(t)$. Given a solution, $\xi(n)$, $n \in \mathcal{N}_t$, a prediction at time $t + \tau + 1$ is obtained by solving (2) for $\xi(t + \tau + 1)$.

Here we have introduced the variable $T(\cdot)$ so that the growth kinetics only appear in (3b). The first term in the objective is the arrival cost, where $\bar{\xi}(t)$ is an estimate of the state prior to the MHE window. The second corresponds to the discrepancy between the predicted and measured outputs, and the third to model error. Physically, $\omega(\cdot)$ could represent inflow uncertainty in $\xi^{\text{in}}(\cdot)$. The constraint (3e) represents optional linear conditions like $\xi(\cdot) \geq 0$ and $T(\cdot) \geq 0$. Such constraints can add bias the estimator, but, for example, prohibit unphysical solutions.

If the system is at equilibrium or changing slowly, it may also be useful to solve the below steady state problem.

$$\begin{aligned}
\mathcal{P}_{\text{S}} \qquad & \min_{\xi,T,\nu,\omega} \|\nu\|^2_{\Sigma_\nu} + \|\omega\|^2_{\Sigma_\omega} && \text{(4a)}\\
\text{subject to} \quad & T = \phi(\xi) && \text{(4b)}\\
& 0 = \hat{V}KT + \hat{N}\xi + \hat{C}\xi^{\text{in}} + \omega && \text{(4c)}\\
& y = \Theta\begin{bmatrix}\xi\\ T\end{bmatrix} + \nu && \text{(4d)}\\
& (\xi, T) \in \Omega. && \text{(4e)}
\end{aligned}$$

4. Convex relaxation

We obtain a convex relaxation of $\mathcal{P}(t)$ by replacing (3b) with

$$T(n) \leq \phi(\xi(n)), \quad n \in \mathcal{N}_t. \tag{5}$$

We similarly obtain a convex relaxation of $\mathcal{P}_{\text{S}}$ by replacing (4b) with

$$T \leq \phi(\xi). \tag{6}$$

We denote these relaxations $\mathcal{P}_{\text{R}}(t)$ and $\mathcal{P}_{\text{SR}}$, respectively. As shown in [11], these are concave and can be represented as SOC constraints when each growth rate is either Monod with constant biomass or Contois.

We say that $\mathcal{P}_{\text{R}}(t)$ ($\mathcal{P}_{\text{SR}}$) is exact when all elements of (5) ((6)) are satisfied with equality. In this case, the relaxation has

the same solution as the original problem. To a limited extent, we can characterize exactness by adapting the results of [12]. In short, those results come from using the Karush-Kuhn-Tucker conditions to derive expressions for the dual multipliers of (5) and (6), the positivity of which guarantees exactness due to complementary slackness. Here we omit the proofs because the present modifications are slight—the only aspect of $\mathcal{P}_{\mathrm{R}}(t)$ that is not a special case of the model in [12] is the arrival cost, $\left\|\xi(t)-\bar{\xi}(t)\right\|_{\Sigma_0}^2$, and it does not affect the result.

Assumption 1. *To simplify our exactness results, we assume the following.*

- $\mathcal{D}_n[\xi(\cdot)] = (\xi(n)-\xi(n-1))/\Delta$, *the implicit Euler method.*
- $\omega(\cdot) = 0$, *which is to say there is no process noise. This eliminates the third term in the objective in* $\mathcal{P}(t)$.
- Θ *is block diagonal, so that the measurement can be written* $y_\xi(n) = \Theta_\xi\xi(n) + v_\xi(n)$ *and* $y_\phi(n) = \Theta_\phi T(n) + v_\phi(n)$. *The associated weighting matrices are* Σ_v^ξ *and* Σ_v^ϕ.
- *We exclude (3e) and (4e) from the analysis. We concede that without the constraints* $\xi(\cdot) \geq 0$ *and* $T(\cdot) \geq 0$, *it is possible for exactness to hold while some components of* $\xi(\cdot)$ *and/or* $T(\cdot)$ *are negative. However, given reasonable, e.g., positive observations, this is highly unlikely, and one can always simply examine the solution. We therefore assume that if exactness holds, then* $\xi(\cdot) \geq 0$ *and* $T(\cdot) \geq 0$.

We remark that the results in this section are not meant to be comprehensive, but rather to provide some guidelines as to whether a given instance of the relaxation is likely to be exact. When it is not exact, it may still provide a good approximation, especially if an underestimator is used to limit the inexactness of (5) or (6), as in [11]. For when feasibility of $\mathcal{P}(t)$ or $\mathcal{P}_{\mathrm{S}}$ is essential, in Section 5 we use the relaxations as a basis for obtaining feasible, albeit potentially local, optimal solutions.

4.1. Exactness of the relaxed MHE

Let $\mathcal{J}(\xi(n)) \in \mathbb{R}^{rp\times mp}$ denote the Jacobian matrix of $\phi(\cdot)$ at $\xi(n)$. For convenience, we define the following quantities for each $n \in \mathcal{N}_t$:

$$\Gamma(n) = \frac{1}{\Delta}\left(\hat{V}/\Delta - \hat{N}(n)^\top - \mathcal{J}(\xi(n))^\top K^\top \hat{V}\right)^{-1}\hat{V}$$

$$\Omega(n) = -2\Theta_\phi^\top \Sigma_v^{\phi\,-1}(y_\phi(n) - \Theta_\phi T(n)) - \Delta K^\top \hat{V}\sum_{k=n}^{t+\tau}\left(\prod_{l=n}^{k}\Gamma(l)\right)\hat{V}^{-1}$$
$$\times\Big(2\mathcal{J}(\xi(k))^\top\Theta_\phi^\top\Sigma_v^{\phi\,-1}(y_\phi(k)-\Theta_\phi T(k))$$
$$-2\Theta_\xi^\top\Sigma_v^{\xi\,-1}(y_\xi(k)-\Theta_\xi\xi(k))\Big).$$

Observe that if Δ is small enough, $\Gamma(n)$ is positive definite and close to the identity matrix. The following result follows from Theorem 1 in [12].

Proposition 1. $\mathcal{P}_R(t)$ *is exact if at an optimal solution,* $\Omega(n) > 0$ *for all* $n \in \mathcal{N}_t$.

Unfortunately, because the derivative of the objective depends on the optimization variables, there is no clear way to guarantee this condition using only the parameters.

4.2. Exactness of the steady state relaxation

$\mathcal{P}_{\mathrm{SR}}$ is a special case of the steady state optimization in Section IV.A of [12], wherein its exactness is characterized by Theorem 2. Specializing the result yields the following proposition.

Proposition 2. $\mathcal{P}_{SR}$ *is exact if the network is outflow connected and at the optimal solution,*

$$0 < \left(I_{mp} + K^\top\hat{V}\left(\hat{N}^\top\right)^{-1}\mathcal{J}(\xi)^\top\right)^{-1}\times$$
$$\left(2\Theta_\phi^\top\Sigma_v^{\phi\,-1}(y_\phi-\Theta_\phi T) - 2K^\top\hat{V}\left(\hat{N}^\top\right)^{-1}\Theta_\xi^\top\Sigma_v^{\xi\,-1}(y_\xi-\Theta_\xi\xi)\right).$$

As in the dynamic case, due to the dependence of the objective on the variables, this does not lend itself to simpler exactness guarantees.

5. Nonconvex solution via the CCP

When the relaxations are not exact, one can still obtain local minima via nonlinear programming. We now show how to do so with the CCP [16]; we specifically use Algorithm 3.1 in [17] because it does not require a feasible starting point. Other advantages of this approach over generic nonlinear programming algorithms are that it exploits the problem's partial SOC structure, and is implementable as a sequence of SOCPs. We focus on $\mathcal{P}_{\mathrm{S}}$ to streamline exposition, but note that all of what follows applies to $\mathcal{P}(t)$ as well as the bioprocess optimizations in [12].

Observe that constraint (4b) can be written as the pair

$$T \leq \phi(\xi) \tag{7a}$$
$$T \geq \phi(\xi). \tag{7b}$$

By assumption, (7a) is convex, and each element is representable as an SOC constraint. Constraint (7b) must therefore be nonconvex. Because $\phi(\xi)$ is concave, it is the part of (7b) that must be linearized in the CCP. We remark that this strategy was applied to an SOC relaxation of optimal power flow in power systems in [20].

The CCP for $\mathcal{P}_{\mathrm{S}}$ is as follows. Choose initial algorithm parameters ξ_0, $\sigma_0 > 0$, $\sigma_{\max} > \sigma_0$, and $\xi > 1$, and set the iteration counter to $z = 0$. Then repeat the below steps until a stopping criterion, e.g., convergence of the objective or variables, is satisfied.

1. Solve the optimization

$$\xi_{z+1} = \underset{\xi,T,v,\omega,\eta}{\operatorname{argmin}}\ \|v\|_{\Sigma_v}^2 + \|\omega\|_{\Sigma_\omega}^2 + \sigma_z\eta^\top\mathbf{1}_{rp}$$
$$\text{subject to}\quad T \leq \phi(\xi)$$
$$\eta + T \geq \phi(\xi_z) + \mathcal{J}(\xi_z)^\top(\xi-\xi_z)$$
$$\eta \geq 0$$
$$0 = \hat{V}KT + \hat{N}\xi + \hat{C}\xi^{\mathrm{in}} + \omega$$
$$y = \Theta\begin{bmatrix}\xi\\ T\end{bmatrix} + v$$
$$(\xi, T) \in \Omega.$$

2. Set $\sigma_{z+1} = \min\{\xi\sigma_z, \sigma_{\max}\}$.
3. Set $z = z + 1$.

Note that $\eta \in \mathbb{R}^{rp}$ is a slack variable that allows ξ_0 to be potentially infeasible for $\mathcal{P}_{\mathrm{S}}$. $\mathbf{1}_{rp} \in \mathbb{R}^{rp}$ is the vector of ones, and the additional term in the objective is a penalty on η. This enables us to set ξ_0 to the (infeasible) solution of $\mathcal{P}_{\mathrm{SR}}$, a natural choice.

6. Examples

We present two numerical examples, one based on the gradostat [21] and the other on anaerobic digestion [22, 23]. In each example, we compare the relaxed MHE (denoted MHE-R), which solves $\mathcal{P}_{\mathrm{R}}(t)$, the exact MHE (denoted MHE-CC), which solves $\mathcal{P}(t)$ using the CCP as in Section 5, and the UKF [1]. The starting point for the MHE-CC in each time period is the solution to MHE-R. We chose the UKF for comparison because it has among the best performance of one-step filters on nonlinear systems. We measure performance in terms of the root-mean-square error of the estimated trajectory. We implemented the MHE-R and MHE-CC using the parser CVX [24] and the solver Gurobi [25] on a personal computer from 2014 with a 1.4 GHz dual-core processor.

6.1. The gradostat

The gradostat is a special case of (1) where in each tank $i \in \mathcal{S}$, a single substrate of concentration S_i is converted to a single type of biomass of concentration X_i. The conversion occurs at the rate $\phi(S_i, X_i)/y$, where here $y = 1$ and

$$\phi(s, x) = \frac{\mu^{\max} sx}{Kx + s}$$

with $\mu^{\max} = 1$ and $K = 1$, the Contois kinetics. The process state in tank i is thus given by $\xi_i = [S_i, X_i]^\top$ and the stoichiometric matrix by $\kappa_i = \kappa = [-1, 1]^\top$.

Only biogas production is observed, so that $y(n) = T(n) + \nu(n) \in \mathbb{R}^4$, where the sensor noise in each tank is drawn from a normal distribution with zero-mean and variance 0.1.

There are 400 time periods of length $\Delta = 0.1$. The discretization is the standard explicit Euler step, $\mathcal{D}_n[\xi(\cdot)] = (\xi(n+1) - \xi(n))/\Delta$. The initial condition is $S_i(0) = i$ and $X_i(0) = 4.5 - 0.5i$ for $i = 1, ..., 4$. The process noise for each concentration in each tank is drawn from a normal distribution with zero-mean and variance Δ^2.

There are four tanks, all with unit growth rate parameters and $V_{ii} = 10$. The inflow vector is $Q^{\mathrm{in}} = [2\ 1\ 1\ 1]^\top$. The flows between tanks are $Q_{12} = 1$, $Q_{23} = 2$, $Q_{34} = 1$, $Q_{42} = 1$, and the diffusion is $d = 0.3Q$.

In tank $i \in \mathcal{S}$, the substrate influent concentrations are $S_i^{\mathrm{in}}(n) = 5\,(1 + \cos(i\pi n/400))$, except that in tank 1, the substrate is $S_1^{\mathrm{in}}(n) = 20$ for $n \in \{160, ..., 200\}$, which simulates a temporary spike. The biomass influent concentrations are $X_1^{\mathrm{in}}(n) = 2$, $X_2^{\mathrm{in}}(n) = 1 + 3n/400$, $X_3^{\mathrm{in}}(n) = 4 - 3(n-1)/400$, and $X_4^{\mathrm{in}}(n) = 2$ for $n \in \{0, ..., 100\} \cup \{300, ..., 400\}$ and zero elsewhere.

The MHEs use a horizon of $\tau = 10$. The weighting matrices in the objective are $\Sigma_0 = I_4$, $\Sigma_\nu = I_4$, and $\Sigma_\omega = 0.1I_4$. The CCP parameters are $\sigma_0 = 1$, $\sigma_{\max} = 1,000$, and $\xi = 1.5$, and it terminates after ten iterations or if the change in the objective and sum of the elements of η are less than 10^{-3}.

The MHE-R and MHE-CC used on average 3.4 and 18.2 seconds to produce their estimates in each period after $n = 10$, and the CCP used on average 5.8 iterations. The UKF, which we expect to be far faster, averaged 10^{-3} seconds per iteration.

Table 1 shows the root-mean-square errors for the substate and biomass over all tanks and time periods. The MHE-CC performs best, then the UKF, and last the MHE-R. The solution produced by the MHE-CC is effectively exact, while that produced by the MHE-R is not.

	UKF	MHE-R	MHE-CC
S	4.1	14.3	1.0
X	8.2	9.0	2.4

Table 1: Root-mean-square errors

Figures 1 and 2 show the trajectories of the states and estimators in each tank. Here we see that the MHE-R is consistently above the actual state trajectory. The UKF does follow the state, but not smoothly due to the process noise. The MHE-CC smoothly follows the state, and for the most part coincides in the plots.

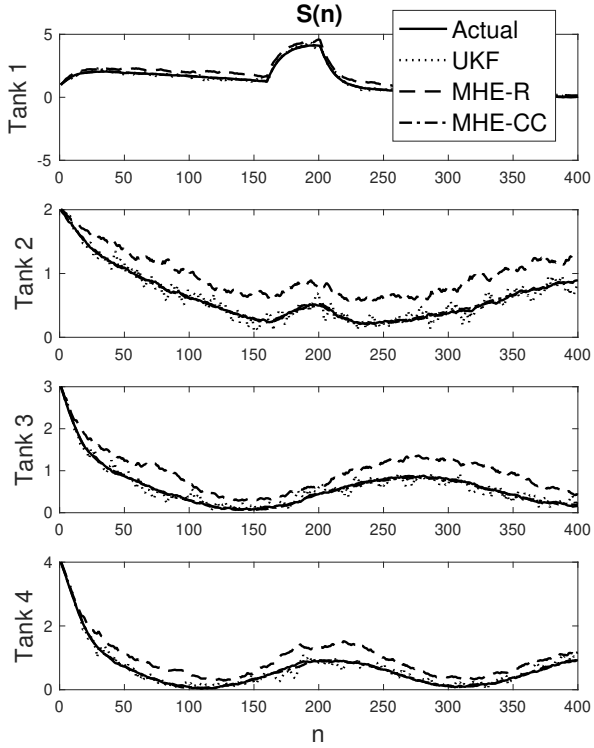

Figure 1: Substrate concentrations in each tank.

We remark that in simulations without process noise, the MHE-CC and UKF performed similarly. We interpret this as being due to the 'short horizon syndrome' [4]—by looking over

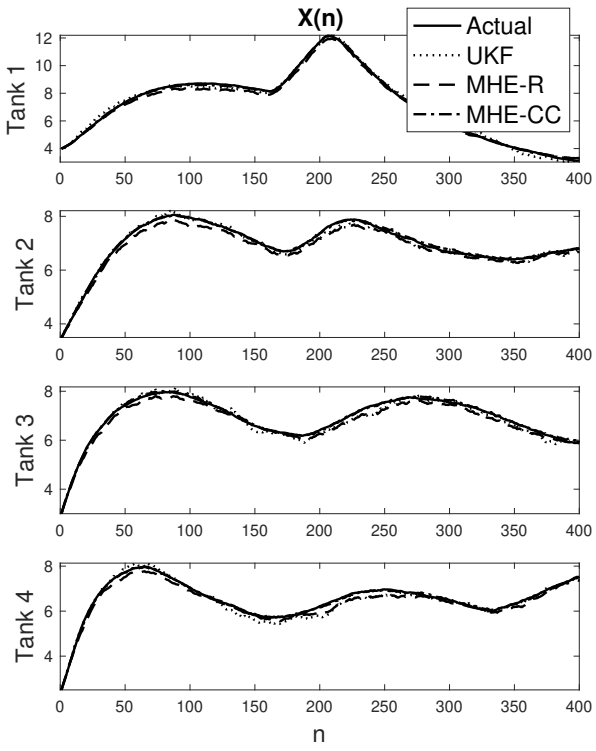

Figure 2: Biomass concentrations in each tank.

multiple time periods, the MHE is better able to separate the noise from the model.

6.2. Anaerobic digestion

This example is adapted from the model of anaerobic digestion in [22], which is a simplification of that in [23]. Our intention is to compare the estimators' performances under significant parameter error, which we describe later.

There is one tank of unit volume with two substrate concentrations, S_1 (g/L) and S_2 (mmol/L), and two biomasses, X_1 (g/L) and X_2 (g/L). There are two biochemical reactions, both described by Contois kinetics:

$$\phi_1(S_1, X_1) = \frac{\mu_1^{\max} S_1 X_1}{K_1 X_1 + S_1} \quad \text{and} \quad \phi_2(S_2, X_2) = \frac{\mu_2^{\max} S_2 X_2}{K_2 X_2 + S_2}.$$

The process state and stoichiometric matrix are $\xi = [S_1, S_2, X_1, X_2]^\top$ and

$$\kappa = \begin{bmatrix} -k_1 & 0 \\ k_2 & -k_3 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The dilution rate, which because the volume is one is equal to the in-and outflow of water, is $D(n) = 0.05(1 + \cos(\pi n/40))$. This model slightly deviates from (3c) in that while the outflows of substrate are $D(n)S_1(n)$ and $D(n)S_2(n)$, the outflow of biomass are $0.65D(n)X_1(n)$ and $0.65D(n)X_2(n)$, as in [22]. The parameters are summarized in Table 2.

Parameters	Values
$\mu_1^{\max}$, $\mu_2^{\max}$	0.0292 h^{-1}, 0.0308 h^{-1}
K_1, K_2	0.71, 0.928
k_1, k_2, k_3	42.14, 116.5 mmol/g, 268 mmol/g
S_1^{in}, S_2^{in}	16 g/L, 75 mmol/L

Table 2: Parameter values. Units are shown only for parameters that are have units.

The observation in each time period is $y(n) = [S_1(n), S_2(n)]^\top + \nu(n) \in \mathbb{R}^2$, where the sensor noise for each substrate is drawn from a zero-mean normal distribution with variance 10. There are 200 time periods of length $\Delta = 1$ hour. The discretization is again the standard explicit Euler step, $\mathcal{D}_n[\xi(\cdot)] = (\xi(n+1) - \xi(n))$. The initial condition is $\xi(0) = [0.989; 8.559; 0.1358; 0.4829]$. There is zero mean, unit variance normal process noise in the evolution of each substrate.

The MHEs use a horizon of $\tau = 10$. The weighting matrices in the objective are $\Sigma_0 = I_4$, $\Sigma_\nu = I_2$, and, for the two substrates, $\Sigma_\omega = I_2$; the elements of (3c) corresponding to the biomasses, which have no process noise, are constrained to be zero. The CCP parameters are $\sigma_0 = 1$, $\sigma_{\max} = 10,000$, and $\xi = 1.5$, and it terminates after twenty iterations or if the change in the objective and sum of the elements of η are less than 10^{-4}.

To improve the numerical performance, we put the biomasses in units of mg/L and adjusted the corresponding parameters. We found that the UKF and MHEs all estimated the process state similarly well. To make the problem more challenging, we introduced errors into the estimators' initial state, $\tilde{\xi}(0) = [20; 20; 0.15; 0.45]$, and maximum growth rate, $\tilde{\mu}_2^{\max} = 1$.

The MHE-R and MHE-CC used on average 1.6 and 25.1 seconds to produce their estimates in each period after $n = 10$, and the CCP used on average 15.3 iterations. The UKF averaged 5×10^{-4} seconds per iteration.

Table 3 shows the root-mean-square errors for the process state over all time periods. The UKF performs slightly better than the MHEs on S_1, X_1, and X_2. On S_2, as seen in the second subplot of Figure 3, the UKF loses the trajectory for some time and incurs a far higher error than the MHEs.

We also observe that the MHE-CC perform slightly worse than the MHE-R, particularly on X_2, which we explain as follows. A solution to $\mathcal{P}(t)$ under one set of parameters will generally be infeasible under different parameters. On the other hand, by virtue of being a relaxation, there will be many solutions that are feasible and potentially optimal for $\mathcal{P}_{\text{R}}(t)$ under both sets of parameters. In this regard, the approximate nature of MHE-R makes it less sensitive to parameter errors.

	UKF	MHE-R	MHE-CC
S_1	38.6	45.9	45.9
S_2	306.1	40.7	40.9
X_1	47.8	47.9	48.2
X_2	107.2	112.2	158.3

Table 3: root-mean-square errors

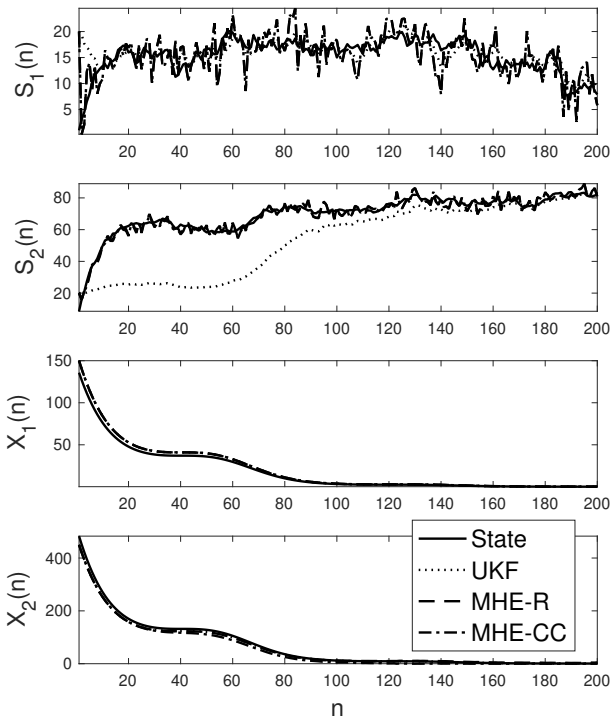

Figure 3: Substrate and biomass concentrations. Note that $X_1(n)$ and $X_2(n)$ have units of mg/L.

These results suggest that an MHE is appropriate for problems that are particularly challenging due, for instance, to high levels or noise or parameter error. For easier problems, the UKF or another one-step estimator may be preferable for simplicity.

7. Conclusion

We have formulated two novel MHEs for a general nonlinear model of bioprocesses. The first MHE is a convex relaxation, which can be solved efficiently as an SOCP, but can admit solutions that are infeasible for the original model. The second MHE starts with the SOCP and uses the CCP to tighten the relaxation and, eventually, obtain a solution that is feasible for the original model. In two examples, we found that the MHE outperforms the UKF on difficult problems with high process noise and exhibited significantly higher robustness to parameter error.

We now describe two directions for future work. Uncertainty can be explicitly represented in convex optimization using tools like stochastic and robust optimization. In this way we could therefore model uncertainty in parameters like the growth rates and influents. A second direction is to estimate these parameters by viewing them as optimization variables. In some cases, e.g., the maximum growth rate, this will result in a biconvex problem, for which the alternating direction method of multipliers [26] is a promising approach.

References

[1] S. J. Julier, J. K. Uhlmann, Unscented filtering and nonlinear estimation, Proceedings of the IEEE 92 (3) (2004) 401–422.

[2] D. Dochain, State and parameter estimation in chemical and biochemical processes: a tutorial, Journal of process control 13 (8) (2003) 801–818.

[3] C. Rao, J. Rawlings, D. Mayne, Constrained state estimation for nonlinear discrete-time systems: stability and moving horizon approximations, IEEE Transactions on Automatic Control 48 (2) (2003) 246–258. doi:10.1109/TAC.2002.808470.

[4] J. B. Rawlings, D. Q. Mayne, M. Diehl, Model predictive control: theory, computation, and design, 2nd Edition, Nob Hill Publishing, 2017.

[5] J. Flaus, L. Boillereaux, Moving horizon state estimation for a bioprocesses modelled by a neural network, Transactions of the Institute of Measurement and Control 19 (5) (1997) 263–270.

[6] E. Arnold, S. Dietze, Nonlinear moving horizon state estimation of an activated sludge model, IFAC Proceedings Volumes 34 (8) (2001) 545–550.

[7] J. Busch, D. Elixmann, P. Kühl, C. Gerkens, J. P. Schlöder, H. G. Bock, W. Marquardt, State estimation for large-scale wastewater treatment plants, Water research 47 (13) (2013) 4774–4787.

[8] X. Yin, B. Decardi-Nelson, J. Liu, Subsystem decomposition and distributed moving horizon estimation of wastewater treatment plants, Chemical Engineering Research and Design 134 (2018) 405–419.

[9] L. Dewasme, S. Fernandes, Z. Amribt, L. Santos, P. Bogaerts, A. V. Wouwer, State estimation and predictive control of fed-batch cultures of hybridoma cells, Journal of Process Control 30 (2015) 50–57.

[10] M. Elsheikh, R. Hille, A. Tatulea-Codrean, S. Krämer, A comparative review of multi-rate moving horizon estimation schemes for bioprocess applications, Computers & Chemical Engineering (2021) 107219.

[11] J. Taylor, A. Rapaport, Second-order cone optimization of the gradostat, Computers & Chemical Engineering 151 (2021) 107347. doi:10.1016/j.compchemeng.2021.107347.

[12] J. A. Taylor, A. Rapaport, D. Dochain, Convex optimization of bioprocesses, Automatic Control, IEEE Transactions onSubmitted (2021). URL `https://arxiv.org/abs/2107.01843`

[13] J. Monod, The growth of bacterial cultures, Annual Review of Microbiology 3 (1) (1949) 371–394.

[14] D. Contois, Kinetics of bacterial growth: relationship between population density and specific growth rate of continuous cultures, Microbiology 21 (1) (1959) 40–50.

[15] M. Lobo, L. Vandenberghe, S. Boyd, H. Lebret, Applications of second-order cone programming, Linear Algebra and its Applications 284 (1998) 193–228.

[16] A. L. Yuille, A. Rangarajan, The concave-convex procedure, Neural computation 15 (4) (2003) 915–936.

[17] T. Lipp, S. Boyd, Variations and extension of the convex–concave procedure, Optimization and Engineering 17 (2) (2016) 263–287.

[18] J. Jacquez, C. Simon, Qualitative theory of compartmental systems, SIAM Review 35 (1) (1993) 43–79.

[19] G. Bastin, D. Dochain, On-line estimation and adaptive control of bioreactors, Elsevier, 1990.

[20] W. Wei, J. Wang, N. Li, S. Mei, Optimal power flow of radial networks and its variations: A sequential convex optimization approach, IEEE Transactions on Smart Grid 8 (6) (2017) 2974–2987.

[21] H. Smith, P. Waltman, The theory of the chemostat: dynamics of microbial competition, Vol. 13, Cambridge University Press, 1995.

[22] A. Rodríguez, G. Quiroz, R. Femat, H. Méndez-Acosta, J. de León, An adaptive observer for operation monitoring of anaerobic digestion wastewater treatment, Chemical Engineering Journal 269 (2015) 186–193. doi:https://doi.org/10.1016/j.cej.2015.01.038.

[23] O. Bernard, Z. Hadj-Sadok, D. Dochain, A. Genovesi, J.-P. Steyer, Dynamical model development and parameter identification for an anaerobic wastewater treatment process, Biotechnology and Bioengineering 75 (4) (2001) 424–438.

[24] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, `http://cvxr.com/cvx` (Mar. 2014).

[25] Gurobi Optimization, LLC, Gurobi optimizer reference manual, `http://www.gurobi.com` (2021).

[26] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed Optimization and Statistical Learning Via the Alternating Direction Method of Multipliers, Now Publishers, 2011.