CONSISTENCY TECHNIQUES FOR SIMULATION OF WASTEWATER TREATMENT PROCESSES WITH UNCERTAINTIES

Marco Kletting [*] **Andreas Rauh** [*] **Harald Aschemann** [*]
Eberhard P. Hofer [*]

[*] *University of Ulm*
Department of Measurement, Control, and Microtechnology
Albert-Einstein-Allee 41, 89081 Ulm, Germany
{Marco.Kletting, Andreas.Rauh, Harald.Aschemann,
ep.Hofer}@e-technik.uni-ulm.de

Abstract: In this paper, interval simulation methods are presented to determine guaranteed enclosures of state variables of an activated sludge process in biological wastewater treatment. This process is characterized by nonlinearities and uncertain but bounded parameters. In uncertain systems an axis-parallel interval box is mapped to a complexly shaped region in the state space that represents sets of possible combinations of state variables. The approximation of this complex region by a single interval box causes accumulation of overestimation over simulation time. The algorithm presented in this paper minimizes this so called wrapping effect by applying consistency techniques, to avoid safety-critical states. *Copyright ©2005 IFAC*

1. INTRODUCTION

The considered activated sludge process in biological wastewater treatment is described by a system of ordinary differential equations with nonlinear reaction kinetics (Köhne, 1998; Henze *et al.*, 1987). The model contains several uncertain parameters which can be time-varying. E.g., the maximum specific growth rate of the substrate consuming bacteria depends on the temperature of the wastewater. Higher temperature leads to a higher growth rate and vice versa. The inflow to the wastewater treatment plant is also an important uncertainty. The amount and composition of the influent wastewater strongly depend on wet or dry weather conditions.

Popular approaches to simulate uncertain systems are Monte Carlo methods. However, these methods are not suitable if guaranteed bounds of the state vari-

ables are required, which is important for security relevant system operation. In real-world applications mostly only lower and upper bounds of the uncertain parameters and their maximum tolerance are known. Consequently, the application of interval methods (Moore, 1979; Jaulin *et al.*, 2001) to the evaluation of the mathematical system model leads to conservative estimations of the lower and upper bounds of the state variables. However, a naive application of interval methods results in huge overestimation. Therefore, the consistency techniques presented in this paper and splitting into subintervals aim at reduction of the overestimation.

Two different kinds of overestimation of the state variables can be distinguished. First, the maximum and minimum values of the state intervals in each time step cannot always be determined exactly if just natural interval evaluation or higher-order methods are

applied (Rauh *et al.*, 2004). Second, using only one axis-parallel interval box it is impossible to represent complexly shaped regions of state variables in the state space even if the exact infimum and supremum of the sets can be determined. This kind of overestimation is usually called the wrapping effect. The wrapping effect occurs due to replacement of these regions by axis-parallel enclosures in each time step and therefore leads to accumulation of overestimation over simulation time.

The first kind of overestimation can be minimized by applying monotonicity tests or iterative calculation of both infimum and supremum (Rauh *et al.*, 2004). The second kind of overestimation is reduced by removing subintervals of an axis parallel box that do not belong to the exact solution of the state equation with respect to the approximation of the previous time step. This is done by the consistency techniques described in this paper.

In addition the time-discretization of the continuous system leads to a time-discretization error, which has to be calculated prior to the application of consistency techniques. It is shown how a conservative enclosure of the time-discretization error is determined.

In Section 2 the formulation of the problem is given. Section 3 describes how the discretization error can be calculated. In Section 4 the proposed simulation algorithm based on consistency techniques is described. In Section 5 simulation results of a subprocess of biological wastewater treatment are shown. Finally, conclusions and an outlook on future research are given in Section 6.

2. PROBLEM FORMULATION

The system model of a simplified activated sludge model in biological waste water treatment (Rauh *et al.*, 2004) consists of a set of four nonlinear coupled differential equations

$$
\begin{aligned}
\dot{S} =& \frac{Q_W}{V_A}\left(S_W - S\right) - \mu\left(S, S_O\right)\frac{1}{Y}X, \\
\dot{X} =& -\frac{Q_W}{V_A}X + \frac{Q_{RS}}{V_A}\left(X_{Set} - X\right) \\
& + \left(\mu\left(S, S_O\right) - b\right)X, \\
\dot{S}_O =& \frac{Q_W}{V_A}\left(S_{OW} - S_O\right) - \mu\left(S, S_O\right)\frac{1-Y}{Y}X \\
& + \frac{\rho_{O2}}{V_A}\left(1 - \frac{S_O}{S_{O,sat}}\right)u_{O2}, \\
\dot{X}_{Set} =& \frac{\left(\left(Q_W + Q_{RS}\right)X - \left(Q_{EX} + Q_{RS}\right)X_{Set}\right)}{V_{Set}}
\end{aligned}
\tag{1}
$$

with

$$
\mu\left(S, S_O\right) = \hat{\mu}_H \frac{S}{S + K_S}\frac{S_O}{S_O + K_{OS}}.
\tag{2}
$$

The four state variables represent the concentration S of biologically degradable organic matter (substrate), the concentration X of substrate consuming bacteria, the concentration S_O of dissolved oxygen in the aeration tank, and the concentration X_{Set} of bacteria in the settler. The settler model assumes an ideal separation of purified wastewater and activated sludge. For the state variables the following restrictions must hold additionally,

$$
\begin{aligned}
&S \geq 0, X \geq 0, X_{Set} \geq 0, \\
&0 \leq S_O \leq S_{O,Sat}.
\end{aligned}
\tag{3}
$$

These inequalities point out the physical restrictions, namely, all concentrations have to be non-negative for all times. Furthermore, the concentration S_O of dissolved oxygen is limited by the saturation concentration $S_{O,Sat}$. The nominal values of the system parameters are given in (Rauh *et al.*, 2004). The uncertainty of the maximum specific growth rate $\hat{\mu}_H$ of the heterotrophic biomass is known to be $\hat{\mu}_H \in [0.9; 1.1]\mu_{H,nominal}$ (Bornemann *et al.*, 1998). In principle the decay rate b is also exposed to similar uncertainties. In this paper it is assumed it to be constant. Q_W is the inflowing wastewater, Q_{EX} denotes the excess sludge removed from the process and Q_{RS} is the sludge which is fed back to the aeration tank. Q_{RS} and Q_{EX} are possible control variables, but in this publication they are assumed to be constant.

The system equations are replaced by the following expression:

$$
\dot{x}(t) = f(x(t), p(t))
\tag{4}
$$

with the state vector x and the uncertain parameter vector p. For the simulation of the system the continuous system has to be discretized in time. This is done by applying a Taylor series expansion

$$
\begin{aligned}
x\left(t_k, x_k, p_k, t_{k+1}\right) =& x_k + Tf^{(0)}\left(t_k, x_k, p_k\right) \\
& + \frac{T^2}{2!}f^{(1)}\left(t_k, x_k, p_k\right) + \ldots \\
& + \frac{T^\tau}{\tau!}f^{(\tau-1)}\left(t_k, x_k, p_k\right) + e\left(\xi, x(\xi)\right)
\end{aligned}
\tag{5}
$$

with the integration step-size T, i.e. $t_k = kT$ and $t_k < \xi < t_{k+1}$. Throughout this paper, all uncertain system parameters $p_k \in \left[\underline{p}_k, \overline{p}_k\right]$ are assumed to be interval parameters, i.e., the range of these parameters is defined by lower bounds $\underline{p}_k$ and upper bounds $\overline{p}_k$. No additional information about the distribution of these parameters is available and needed. In addition, these parameters are assumed to be time-varying, i.e., arbitrary variations of the actual value of these parameters from one time-step to the next are allowed within the specified parameter range without specifying any dynamical behavior of the parameter variation.

Applying interval arithmetic simulation techniques to (4), guaranteed bounds for the state variables x_{k+1} are calculated recursively in terms of uncertainties of the state vector x_k and uncertain system parameters p_k. These bounds of x_{k+1} should be as tight as possible

to the actual lower and upper bounds of the state variables. Furthermore, overestimation of these bounds is reduced by applying consistency techniques presented in Section 4 to improve the enclosure of non-axis-parallel state intervals. By this procedure, the influence of parts of the state space without physical relevance to the considered process model is minimized.

3. TIME-DISCRETIZATION ERROR

Systems of nonlinear differential equations are an adequate description for numerous real-world systems, e.g. for biological processes such as wastewater treatment. For numerical analysis of the system behavior time-discretization is performed. This leads to an unavoidable discretization error depending on the discretization method and the step-size. With interval methods it is possible to give conservative bounds for this error and therefore a guaranteed estimation of the bounds of the continuous-time system. For the truncation error the following relation holds

$$e(\xi, x(\xi), p_k) \subseteq \frac{T^{\tau+1}}{(\tau+1)!} f^{(\tau)}([t_k ; t_{k+1}], B, p_k).$$

(6)

In this expression, the interval B is a bounding box for the range of the state variables $x(t) \in B, \forall t \in [t_k; t_{k+1}]$, if the initial intervals $x_k = x(t_k)$ are given. Applying the Picard Operator

$$\Phi(B) = x(t_k) + [0 ; T] f([t_k ; t_{k+1}], B, p_k) \quad (7)$$

the interval enclosure B can be determined (Deville *et al.*, 2002). B is initialized with the state vector x_k. If $\Phi(B) \not\subseteq B$, the bounding-box B has to be inflated. If $\Phi(B) \subseteq B$, (1) is performed recursively until $\Phi(B)$ matches with B except a small given error. In case that this algorithm does not converge or that the interval of the discretization error $e([t_k ; t_{k+1}], B, p_k)$ is too large, the step-size T has to be reduced.

Note, that the Picard operator uses a first-order Taylor expansion. It can be generalized to higher orders which may allow larger step-sizes, at the cost of higher computational effort. In this paper the first order expansion for determining B has been used in combination with the second order expansion, which improves the results. For the second order expansion

$$\Phi(B) = x(t_k) + [0 ; T] f(t_k, x(t_k), p_k) + \frac{1}{2} [0 ; T]^2 f^{(1)}([t_k ; t_{k+1}], B, p_k)$$

holds (Corliss and Rihm, 1995).

The higher derivatives $f^{(1)} = \ddot{x}$, $f^{(2)} = \dddot{x}$, ..., $f^{(\tau)} = x^{\tau+1}$ of the right hand side of the differential equation can be determined recursively with the following relation

$$f^{(0)}(t, x) = f(t, x)$$

$$f^{(\tau)}(t, x) = \frac{\partial f^{(\tau-1)}(t, x)}{\partial t} + \frac{\partial f^{(\tau-1)}(t, x)}{\partial x} f^{(0)}(t, x)$$

(8)

The explicit Euler method corresponds to the truncation of a Taylor series expansion after the first-order term $\tau = 1$. For the error introduced by neglecting all higher-order terms guaranteed bounds can be calculated by an interval extension of the error term $e(\xi, x(\xi), p_k)$. The system has no explicit dependency in t, hence for the explicit Euler method the following expression holds:

$$x_{k+1} \subseteq x_k + T f^{(0)}(x_k, p_k) + e(B, p_k). \quad (9)$$

x_{k+1} is a guaranteed bound of all state variables in time-step at t_{k+1}.

4. CONSISTENCY TECHNIQUES

In this section a simulation algorithm based on consistency techniques is described, which reduces the wrapping effect significantly. The algorithm consists of two steps: A forward step which calculates rough enclosures of the state variables in the following time step and a backward step, which deletes subintervals which do not belong to the exact solution with respect to the previous approximation (Deville *et al.*, 1998; Deville *et al.*, 2002; Kletting *et al.*, 2004).

4.1 The Forward Step

The forward step can be performed in several ways. For the explicit Euler method with error term we get

$$x_{k+1} = x_k + T f(x_k, p_k) + e(B, p_k) \quad (10)$$

while for the implicit Euler method with error term

$$x_{k+1} = x_k + T f(x_{k+1}, p_{k+1}) - e(B, p_k) \quad (11)$$

holds. In the latter x_{k+1} is calculated by applying an interval Newton method. In this case the Krawczyk method (Neumaier, 1990), which is described in Appendix A, has been used. The advantage of the Krawczyk method over other interval Newton approaches is that the inversion of interval matrices is avoided.

In this paper both variants (10) and (11) of the forward step are used in combination. First, the explicit method is performed which is used as initialization of the interval Newton method to determine the solution of the implicit method. In other words the implicit method contracts the solution of the explicit method. To get even tighter enclosures equation (10) is evaluated with a monotonicity test and iterative calculation of infimum and supremum. These methods are described in (Rauh *et al.*, 2004).

4.2 The Backward Step

For backward computation of a subinterval SI three different cases have to be distinguished. These three cases are illustrated in Fig. 1.
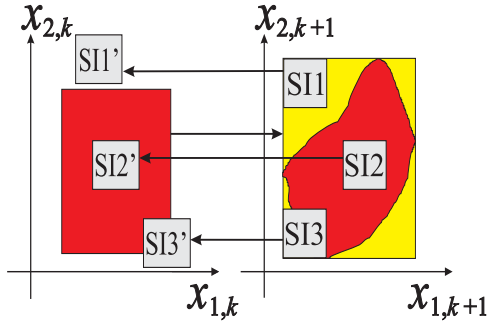
Figure 1. Forward step and backward step.

Applying the backward step, subinterval $SI1$ in time step $k + 1$ is mapped to $SI1'$ in time step k. $SI1'$ lies completely out of x_k, therefore $SI1$ does not belong to the exact solution. Subinterval $SI2'$ lies completely inside of x_k. That means $SI2$ belongs to the exact solution of time step $k + 1$. Subinterval $SI3$ has to be split further, because $SI3'$ is only included partially in x_k and therefore no conclusion can be made. The repeated application of splitting and backward calculation yields in an approximation of the actual solution in time step $k + 1$.

There are now two possibilities to perform the backward step. To obtain tighter results the intersection of both methods is used. The first possibility is derived from equation (10) by rewriting it the following way:

$$0 = x_k + Tf(x_k, p_k) + e(B, p_k) - x_{k+1} = h(x_k, p_k)$$
$$(12)$$

Next, x_{k+1} is split into several subintervals $\tilde{x}_{k+1}$. Now it has to be tested if there exists a region of zeros in x_k for every $\tilde{x}_{k+1}$. If there exists a region of zeros in x_k, then $\tilde{x}_{k+1}$ belongs to x_{k+1}. If not, $\tilde{x}_{k+1}$ can be deleted. If no conclusion can be made $\tilde{x}_{k+1}$ has to be split again. Equation (12) is solved by applying the Krawczyk method. A detailed description of the Krawczyk method and how it is applied to consistency tests can be found in Appendix A. The second way to perform the backward step works as follows. First, equation (11) is rewritten

$$x_k = x_{k+1} - Tf(x_{k+1}, p_{k+1}) + e(B, p_k) \quad (13)$$

then x_{k+1} is split into subintervals. For subintervals $\tilde{x}_{k+1}$

$$\tilde{x}_k = \tilde{x}_{k+1} - Tf(x_{k+1,p_{k+1}}) + e(B, p_k), \quad (14)$$

is applied.
If

$$\tilde{x}_k \subset x_k, \quad (15)$$

it can be guaranteed, that there exists a zero in x_k for any element of $\tilde{x}_{k+1}$ If

$$\tilde{x}_k \cap x_k = \emptyset \quad (16)$$

there is no zero in x_k for any element of $\tilde{x}_{k+1}$ and $\tilde{x}_{k+1}$ can be deleted. If

$$\tilde{x}_k \not\subset x_k \quad (17)$$

and

$$\tilde{x}_k \neq \emptyset \quad (18)$$

$\tilde{x}_{k+1}$ has to be split further.

An efficient splitting strategy makes sure that the number of splitting operations and therefore the number of intervals is kept lower, than if a cyclical change of the splitting direction is applied. It is more efficient to split in the direction where the backward step is most sensitive. An heuristic approach is given in (Kletting *et al.*, 2004). The consistency test and the associated splitting leads inevitably to an increasing number of intervals. To avoid an exponential growth of the number of intervals also efficient merging strategies have to be applied. Two subintervals can be replaced by the convex hull around both, if the resulting interval leads to no or only small overestimation. This is illustrated in Figs (2) and (3). Some merging strategies are de-
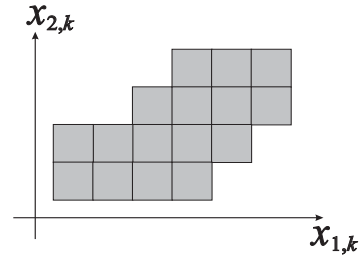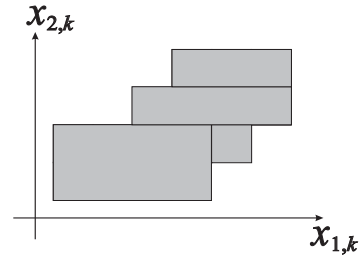

Figure 2. Intervals before merging.

Figure 3. Intervals after merging.

scribed in (Kletting *et al.*, 2004).

5. SIMULATION RESULTS

In this Section, the proposed interval arithmetic simulation algorithm is applied to a simplified Activated Sludge Model ASM in biological wastewater treatment which has been described in Section 2.

5.1 Truncation Error

In Fig. 4 a grid based reference solution (thin trajectories) with sufficiently small step-size and with small tolerances is shown. A comparison of the interval evaluation of the discrete-time system without considering the time discretization error for a step-size $T = 10s$ shows that the reference solution may not be included within the interval solution. If the additive interval

evaluation of the truncation error is considered, the reference solution is within the state intervals marked by the vertical lines in each time-step. In this case the combination of implicit and explicit Euler method is used, which would be impossible without considering the discretization error. The truncation error can be reduced by reduction of the step-size. However, the simulation always yields guaranteed enclosures of the continuous time system.
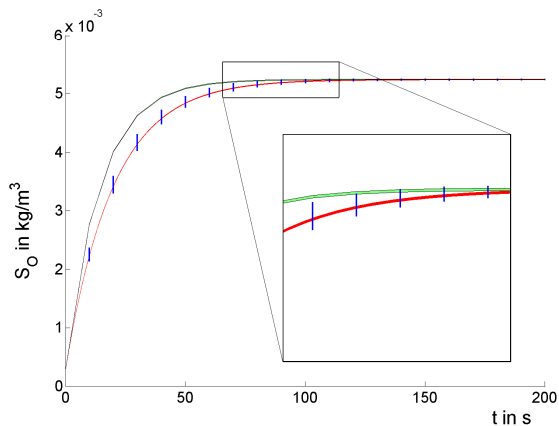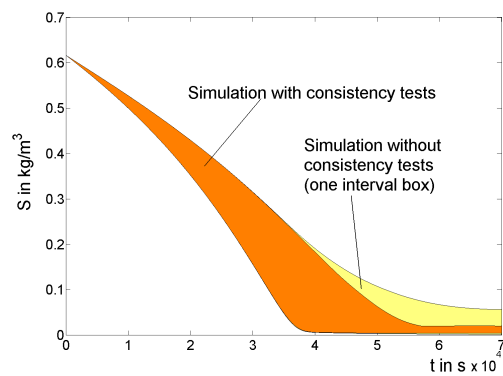

Figure 4. Oxygen concentration.

## 5.2 Results for the Simplified Activated Sludge Model

Fig. 5 shows a comparison between the simulation results with consistency tests and without consistency tests where only the forward step with step-size $T = 10s$ was applied. After 40000 s the step-size has been changed to $T = 5s$.

The maximum number of intervals in the simulation with consistency test is 500. The consistency test is applied to the 20 largest intervals and only in every third time step. Which is a sensible compromise between simulation quality and computing time. Those 20 intervals are first split and the consistency test is applied to the now 40 subintervals. Subintervals which can be excluded are deleted from the list, intervals which belong to the solution are inserted into another list. The remaining intervals are split again. This procedure is repeated until a maximum of 5000 intervals is reached. At the end of each time step a merging routine is applied, which reduces the number of subintervals significantly, with only small overestimation. If there are still more then 500 intervals left after the merging routine only the forward step is applied until the merging routine at the end of each time step reduces the number to less than 500.

The simulation results make clear that a simulation with only a single interval box leads to much too conservative results. The application of consistency tests gives very tight results for the substrate concentration and for the concentration S_O of dissolved oxygen and the explosion of the results for the bacteria concentrations is avoided as it is likely to happen without
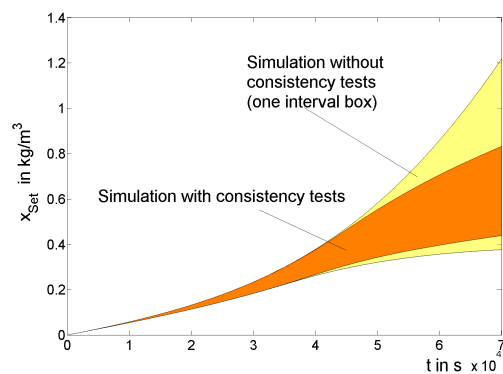

(a) Substrate concentration S.
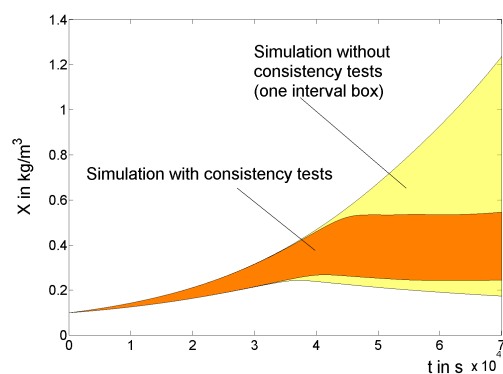
(b) Bacteria concentration X.
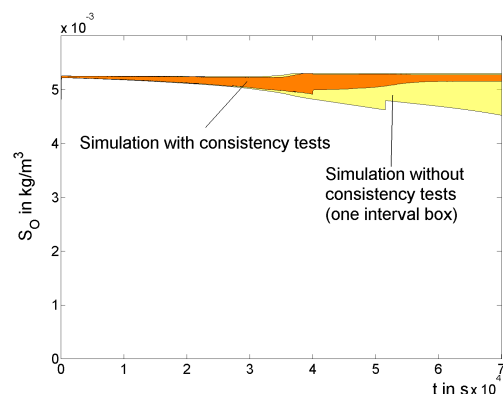
(c) Concentration of dissolved oxygen S_O.

(d) Bacteria concentration in the settler X_{Set}.

Figure 5. Simulation results

consistency techniques tests and without splitting into subintervals.

6. CONCLUSION AND FURTHER WORK

In this paper, an interval arithmetic simulation approach for continuous-time systems with uncertain system parameters has been proposed. Appropriate consistency techniques have led to guaranteed bounds for all state variables. The application to an activated sludge model of biological wastewater treatment has pointed out that the proposed consistency tests reduce the wrapping effect significantly. Tighter results can be obtained by using more subintervals.

Further reduction of the wrapping effect is achieved by applying the consistency tests in every or every second time step. Additionally, a recursive pseudo-linear state transformation can be applied in the forward step. The implementation of an adapted step-size control will increase the efficiency and quality of the algorithm. It is also possible to allow even parameter variations between two time steps —not only at the sampling points— to achieve still guaranteed bounds of the state variables. This requires the inclusion of lower and upper bounds of the parameter variation rates for the determination of the bounding box and the time-discretization error.

REFERENCES

Bornemann, C., M. Freund, J. Londong, O. Nowak, R. Otterpohl and Th. Rolfs (1998). Hinweise zur dynamischen Simulation von Belebungsanlagen mit dem Belebtschlammodell Nr. 1 der IAWQ. *Korrespondenz Abwasser* **45**(3).

Corliss, George F. and Robert Rihm (1995). Validating an A priori Enclosure Using High-Order Taylor Series. In: *Scientific Computing and Validated Numerics: Proceedings of the International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics*. pp. 228–238.

Deville, Y., M. Janssen and P. van Hentenryck (1998). Consistency techniques for ordinary differential equations. In: *Proceedings of the International Conference on Principles and Practice on Constraint Programming*. Springer–Verlag. pp. 162–176.

Deville, Y., M. Janssen and P. van Hentenryck (2002). Consistency Techniques for Ordinary Differential Equations. *Constraint* **7**(3–4), 289–315.

Henze, M., C.P.L. Grady, W. Gujer, G.v.R. Marais and T. Matuso (1987). Activated Sludge Model No. 1. In: *IAWPRC Scientific and Technical Reports No. 1*. London, Great Britain.

Jaulin, L., M. Kieffer, O. Didrit and É. Walter (2001). *Applied Interval Analysis*. Springer–Verlag. London.

Köhne, M. (1998). Analyse und Regelung biologischer Abwasserreinigungsprozesse in Kläranlagen. *at-Automatisierungstechnik* **46**(5), 215–234. R. Oldenbourg-Verlag, München.

Kletting, M., A. Rauh, H. Aschemann and E. P. Hofer (2004). Consitency tests in guaranteed simulation of nonlinear uncertain systems with application to an activated sludge process. In: *Proc. of SCAN 2004*. Fukuoka, Japan. to apppear.

Moore, R.E. (1979). *Methods and Applications of Interval Analysis*. SIAM. Philadelphia.

Neumaier, A. (1990). *Interval Methods for Systems of Equations*. Cambridge University Press, Encyclopedia of Mathematics. Cambridge.

Rauh, A., M. Kletting, H. Aschemann and E. P. Hofer (2004). Application of Interval Arithmetic Simulation Techniques to Wastewater Treatment Processes. In: *Proc. of MIC 2004*. Grindelwald, Switzerland. pp. 287–293.

Appendix A. THE KRAWCZYK METHOD

The Krawczyk method is used to determine if a subinterval $\tilde{x}_{k+1}$ belongs to the solution in time step $k + 1$ with respect to the previous approximation. The Krawczyk method has the following iteration rule

$$k(x) = x_m - Yh(x_m) + \left(I - Y\frac{\partial h}{\partial x} \right)(x - x_m)$$

with $Y^{-1} \in \frac{\partial h}{\partial x}$ and x_m=mid(x). The iteration

$$x^{n+1} = k(x^n) \cap x^n$$

is called Krawczyk iteration, where n is the n-th iteration. The initial box x^0 corresponds to the interval box x_k. If

$$k(x) \subset x_k$$

then it can be guaranteed, that there exists a zero in x_k for any element of $\tilde{x}_{k+1}$. But if

$$k(x) \cap x_k = \emptyset$$

there is no zero in x_k for any element of $\tilde{x}_{k+1}$ and $\tilde{x}_{k+1}$ can be deleted. If

$$k(x) \not\subset x_k$$

and

$$k(x) \neq \emptyset$$

another iteration has to be performed. If this does not improve the result $\tilde{x}_{k+1}$ has to be split further.