

# Chapter 6

## Examples

**Introductory comments.** After introducing the model–data–overlap in chapter 2, illustrating the approach in chapter 4 and discussing computational aspects in chapter 5, it is now time to show how the concepts works for applications.

In the first section, the linear initial value problems (2.7) and (2.8) from chapter 2 are re–used. In addition to there, also the linear overlap functional  $\mathcal{F}_{\mathcal{L}}$  is calculated and compared to the exact case  $\mathcal{F}_{\mathcal{O}}$ . In the second section, biokinetic models are under consideration. It is shown how one can discriminate biokinetic models. The section afterwards shows how biokinetic models can be included into pharmacokinetic models and demonstrates how such models can be validated by the model–data–overlap. The chapter is closed by showing how the TRAIL algorithm can be used for calculating the model variability distribution for a Michael–Menten–model.

The difference between the four sections concerns the methods used for calculating the model variability distribution. The first section deals with a linear approximation method, the second and third one with an ”exact” and the fourth one shows a method that is based on a linear approximation with a nonlinear correction afterwards, namely the TRAIL algorithm.

### 6.1 Linear Initial Value Problems

At this point, the introductory example is reactivated. The prominent member of the class (2.1), namely linear initial value problems,

$$\frac{d}{dt} \mathbf{y}(t, \boldsymbol{\theta}) = \mathbf{J}(\boldsymbol{\theta}) \mathbf{y}(t) + \mathbf{b}(\boldsymbol{\theta}) \quad \text{with} \quad \mathbf{y}(0) = \mathbf{y}_0, \quad (6.1)$$

is used for calculating the linear and exact overlap  $\mathcal{F}_L$  and  $\mathcal{F}_O$ , respectively. As already mentioned, the linear one can be calculated analytically, since the trajectory can be written as

$$\Phi^t \mathbf{y}_0 = \exp(t\mathbf{J}(\boldsymbol{\theta})) \mathbf{y}_0 + \mathbf{J}(\boldsymbol{\theta})^{-1} (\exp(t\mathbf{J}(\boldsymbol{\theta})) - \mathbf{1}) \mathbf{b}(\boldsymbol{\theta}). \quad (6.2)$$

and therefore the derivative with respect to the parameters  $\boldsymbol{\theta}$  can be calculated. For reasons of better readability, the dash ' shall denote the partial derivative with respect to  $\theta$  and the parameter dependency with respect to  $\theta$  is omitted. Then the flow derivative, needed for  $\Sigma_{\mathcal{M}}(\boldsymbol{\theta}, \Delta\boldsymbol{\theta}^2, t)$  in (5.4), can be written as

$$\begin{aligned} \frac{\partial}{\partial \theta} \Phi_{\theta}^t y_0 &= \exp(t\mathbf{J})' \mathbf{y}_0 - \mathbf{J}^{-1} \mathbf{J}' \mathbf{J}^{-1} \exp(t\mathbf{J}) \mathbf{b} + \mathbf{J}^{-1} \mathbf{J}' \mathbf{J}^{-1} \mathbf{b} \\ &+ \mathbf{J}^{-1} \exp(t\mathbf{J})' \mathbf{b} + \mathbf{J}^{-1} \exp(t\mathbf{J}) \mathbf{b}' - \mathbf{J}^{-1} \mathbf{b}'. \end{aligned} \quad (6.3)$$

**Example.** The two possible candidate, the two-dimensional models  $M_1$

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -2\theta & 2 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -6 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} x(t_0) \\ y(t_0) \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix}, \quad (6.4)$$

and model  $M_2$

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} -2 & 2\theta \\ -1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} -6 \\ \theta \end{pmatrix}, \quad \begin{pmatrix} x(t_0) \\ y(t_0) \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix}, \quad (6.5)$$

are to be discriminated. Both models coincide for  $\theta = 1$  (see figure 6.1). The same data is used as in chapter 2.

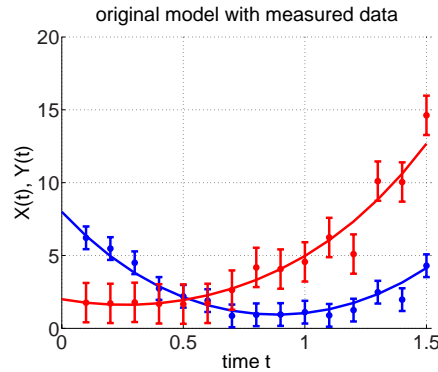


Figure 6.1: Model and data plot: For  $\theta = 1$ , the trajectories of  $M_1$  and  $M_2$  coincide. These trajectory values were taken to produce the data by perturbing them. The data is symbolized by points with attached error bar.

The data is produced artificially by taking trajectory values for  $\theta = 1$  at some instances  $t_k$  and perturbing them. The variance for perturbing at each  $t_k$  was set to be proportional to the variance of the model variability  $\Sigma_{\mathcal{M}}(\theta = 1, \Delta\theta = 0.15, t = t_k)$  of model 1 at each point  $t_k$ . On the other hand, the experimenter assumes the data variances for the  $x$ - and  $y$ -component to be  $\sigma_x^2(t) = \sigma_x^2 = 0.25$  and  $\sigma_y^2(t) = \sigma_y^2 = 0.75$ , respectively.

After this generation of data, the target functionals  $\mathcal{F}_{\mathcal{L}}$  from (5.5) as well as  $\mathcal{F}_{\mathcal{R}}$  from (3.13) were numerically optimized with respect to  $\theta_{\mathcal{L}}$ ,  $\Delta\theta_{\mathcal{L}}$  and  $\theta_{\mathcal{R}}$ , respectively.

	$\theta_{\mathcal{L}}$	$\Delta\theta_{\mathcal{L}}$	$\mathcal{F}_{\mathcal{L}}$ in $x$	$\mathcal{F}_{\mathcal{L}}$ in $y$	$\mathcal{F}_{\mathcal{L}}$ total
$M_1$	0.934	0.367	83.1 %	72.3 %	77.7 %
$M_2$	0.744	1.433	58.4 %	74.4 %	66.4 %
	$\theta_{\mathcal{R}}$	$\bar{\sigma}_{\mathcal{R}}$	ci	gof.	$\mathcal{F}_{\mathcal{R}}$
$M_1$	0.910	0.036	0.073	0.969	0.196
$M_2$	0.913	0.024	0.048	0.974	0.178

Table 6.1: Linear overlap parameter estimation by  $\mathcal{F}_{\mathcal{L}}$  and residual parameter estimation  $\mathcal{F}_{\mathcal{R}}$  for model  $M_1$  and  $M_2$ . (notation:  $\bar{\sigma}_{\mathcal{R}}$  = standard error, ci = half length of the 95 % confidence interval, gof. = goodness of fit),

The linear overlap optimization by  $\mathcal{F}_{\mathcal{L}}$  favors model  $M_1$  over  $M_2$ . The results show significantly different estimated values of  $\theta_{\mathcal{L}}$  and  $\Delta\theta_{\mathcal{L}}$  for both models. Whereas  $\theta_{\mathcal{L}}$  for model  $M_1$  is reasonably close to 1, the parameter used to produce the data,  $\theta_{\mathcal{L}}$  for model  $M_2$  differs significantly shows a very high parameter variability  $\Delta\theta_{\mathcal{L}}$  in addition.

In contrast, the residual case leads to a different conclusion. For both models, the estimated parameters  $\theta_{\mathcal{R}}$  are close together, the residual  $\mathcal{F}_{\mathcal{R}}$  as well as the goodness of fit (gof.) indicate almost the same quality of fit. Only the standard error<sup>1</sup>  $\bar{\sigma}_{\mathcal{R}}$  distinguishes the parameter  $\theta_{\mathcal{R}}$  in  $M_2$  to be estimated more precisely. However, in neither case, the 95%-confidence interval does include the parameter  $\theta = 1$ , originally used for perturbing the data. At last, the example shows the different informational value of the parameter variability  $\Delta\theta_{\mathcal{L}}$  and the standard error  $\bar{\sigma}_{\mathcal{R}}$ . In the overlap concept, the model variability  $\Delta\theta_{\mathcal{L}}$  for  $\theta_{\mathcal{L}}$  is smaller for

<sup>1</sup>The standard errors are the corresponding diagonal entries of  $(\mathbf{J}^T \Sigma_{\mathcal{D}}^{-1} \mathbf{J})^{-1}$  defined in (3.6).

model  $M_1$ , whereas the relation is changed in the residual framework, where the standard error  $\bar{\sigma}_{\mathcal{R}}$  is smaller for  $\theta_{\mathcal{R}}$  of model  $M_2$ .

The concept of model–data–fitting for the presented example is illustrated in fig-

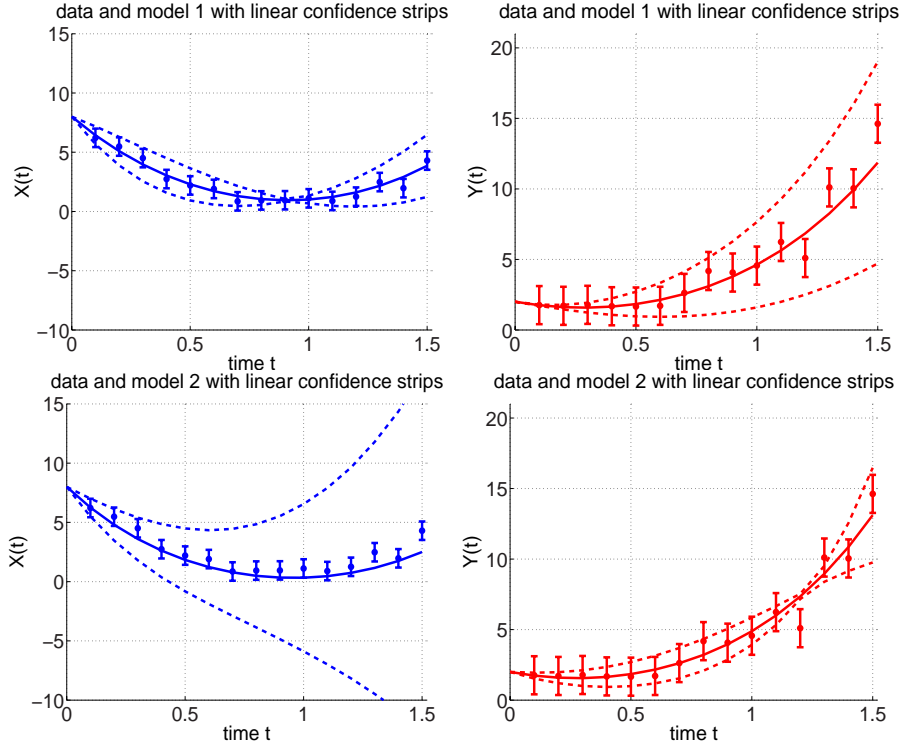


Figure 6.2: Optimal overlap of data error (symbolized by error bars) and linear model variability (symbolized by the 95%-confidence strips) for model  $M_1$  (top two pictures) and  $M_2$  (bottom).

ure 6.2. In both components, the qualitative course of the model variability strips is different. For component  $x$  of model  $M_1$  and component  $y$  of model  $M_2$  the parameter sensitivity vanishes at some instances, whereas for other components, the strips are diverging. The extremely diverging strips for the  $x$ -component of model  $M_2$  are "responsible" for the worse overlap compared to  $M_1$ .

### Comparison linear and non-linear propagation

Previously it was argued that in complex and high-dimensional application setting the requirements the total on computing time prevents us from computing the exact parameter distribution propagation for arbitrary models. This consideration has led us to the linear overlap functional  $\mathcal{F}_{\mathcal{L}}$  of (5.5). For very small systems, however, like model  $M_1$  and  $M_2$ , it is justifiable to calculate  $\mathcal{F}_{\mathcal{O}}$  instead

of  $\mathcal{F}_{\mathcal{L}}$  by sampling the parameter distribution  $\pi$ , propagating the trajectories and re-assembling the model variability distribution  $\mathcal{M}_t$  as it is described in (2.2).

If one takes the parameter  $\theta_{\mathcal{L}}$  and its variability  $\Delta\theta_{\mathcal{L}}$  from table 6.1, the resulting overlap  $\mathcal{F}_{\mathcal{O}}$  is – as expected – different to the one of  $\mathcal{F}_{\mathcal{L}}$ . As shown in table 6.2, in some cases the exact propagation yields a larger, in other ones a smaller value  $\mathcal{F}_{\mathcal{O}}$  compared to the approximation  $\mathcal{F}_{\mathcal{L}}$ .

	in $x$	in $y$	total
$M_1$ for $\mathcal{F}_{\mathcal{L}}$	83.1 %	72.3 %	77.7 %
$M_1$ for $\mathcal{F}_{\mathcal{O}}$	79.6 %	73.1 %	76.7 %
$M_2$ for $\mathcal{F}_{\mathcal{L}}$	58.4 %	74.4 %	66.4 %
$M_2$ for $\mathcal{F}_{\mathcal{O}}$	47.8 %	61.7 %	54.8 %

Table 6.2: Comparison of linear overlap  $\mathcal{F}_{\mathcal{L}}$  and non-linear overlap  $\mathcal{F}_{\mathcal{O}}$  using the optimal parameters for  $\mathcal{F}_{\mathcal{L}}$  in table 6.1.

To understand the reasons of these deviations, one can for example compare the standard deviations of the linearly and nonlinearly propagated model variability  $\mathcal{M}_t$  as shown in figure 6.3. In some cases the deviations almost coincide (c.f.  $y$ -component of  $M_1$ ), are moderately diverging (c.f.  $x$ -component of  $M_2$  or  $x$ -component of  $M_1$ ) or structurally differ (c.f.  $y$ -component of  $M_2$ ).

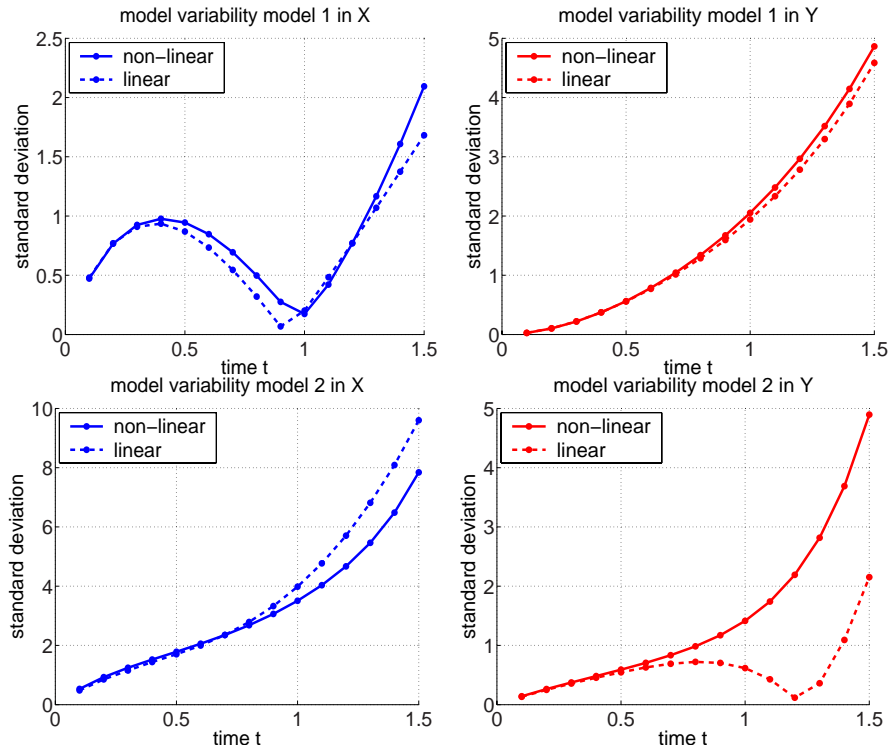


Figure 6.3: Comparison of the standard deviations of the linear and nonlinear model variability distribution for the models  $M_1$  and  $M_2$ .

Not only the standard deviations can differ, but also the model variability distribution  $\mathcal{M}$  can become highly non-normal. The normal property, however, was essential in the construction of  $\mathcal{F}_{\mathcal{L}}$ . In some cases, the normal property of the distribution is maintained as shown in figure 6.4.

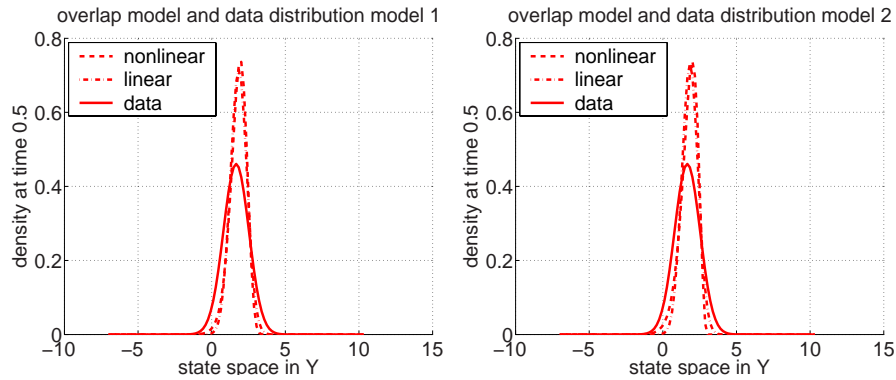


Figure 6.4: Example where non-linearly propagated model variability distribution can be regarded as normal. Two examples for the state space in  $y$  for model  $M_1$  (left) and  $M_2$  (right) are shown.

In other ones, the non-linearly propagated model variability distributions  $\mathcal{M}$  is highly non-normal, as exemplarily seen in figure 6.5. The documented effects

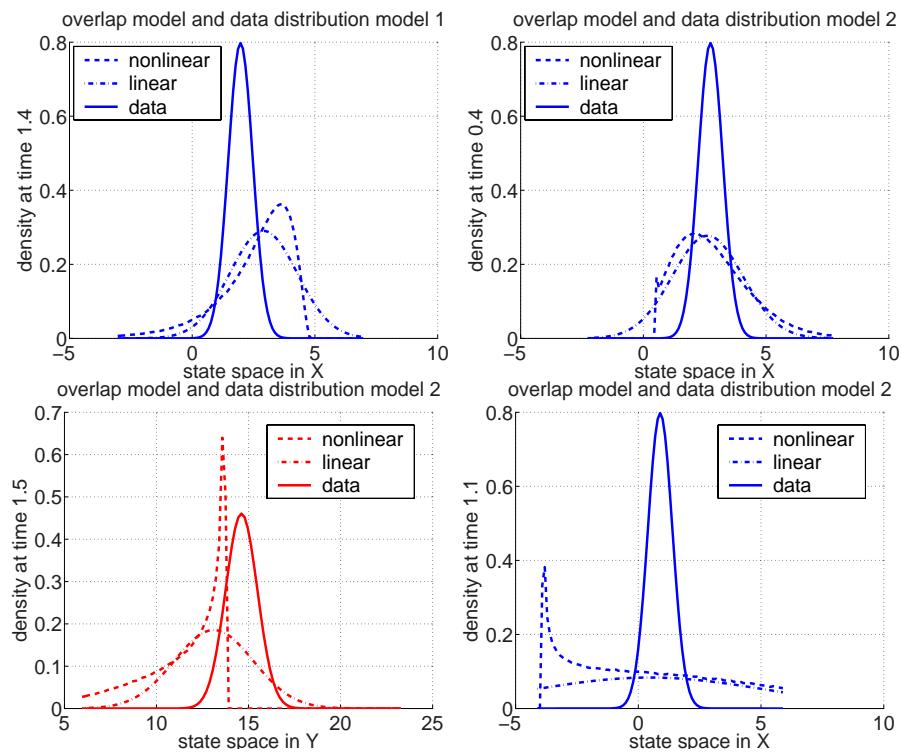


Figure 6.5: Examples where the non-linearly propagated model variability distribution cannot be regarded as normal.

result in different overlaps as shown in figure 6.6. Good coherence can be seen

where, for example, the standard deviations almost coincide (c.f.  $y$ -component of  $M_1$ ); strong deviation for qualitative and quantitative deviations (c.f.  $y$ -component of  $M_2$ ) may result in deviating overlap numbers. Non-surprisingly,

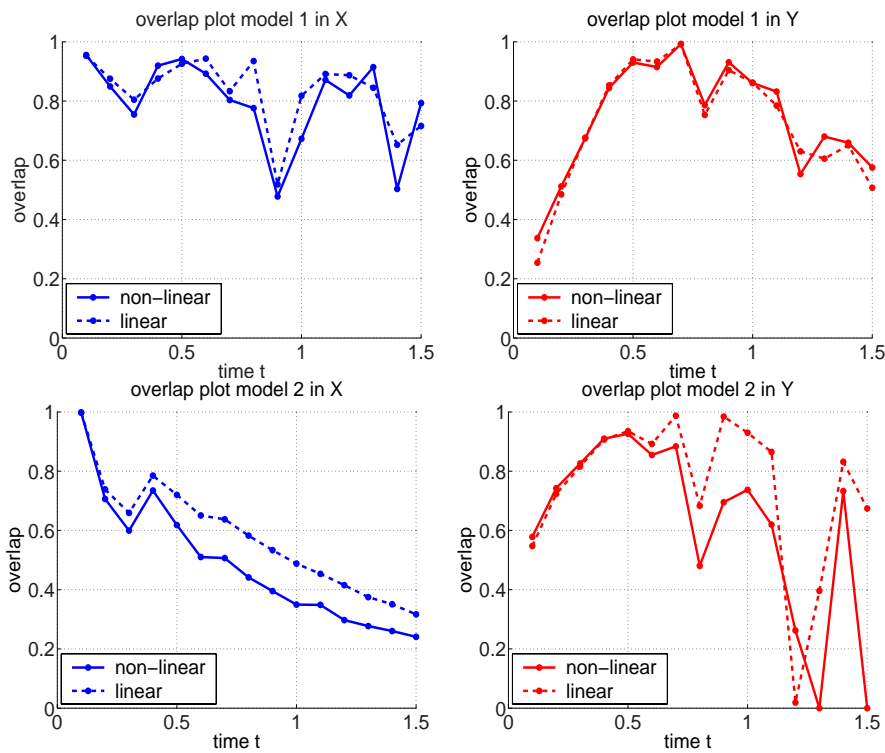


Figure 6.6: Time course of the overlap for  $M_1$  and  $M_2$  and each component: The linear overlap  $\mathcal{F}_{\mathcal{L}}$  and the "exact" overlap  $\mathcal{F}_{\mathcal{O}}$  are shown.

the different propagation behavior results in different optimal hyperparameters. The resulting optimal parameters for  $\mathcal{F}_{\mathcal{L}}$  and  $\mathcal{F}_{\mathcal{O}}$  are documented in table 6.3. The optimal hyperparameters can differ significantly (see  $\theta$  for  $M_2$ ). Nevertheless, the qualitative result is the same: the model  $M_1$  is discriminated to be the appropriate one.



	$\theta_{\mathcal{L}}$	$\Delta\theta_{\mathcal{L}}$	$\mathcal{F}_{\mathcal{L}}$ in $x$	$\mathcal{F}_{\mathcal{L}}$ in $y$	$\mathcal{F}_{\mathcal{L}}$ total
$M_1$	0.934	0.367	83.1 %	72.3 %	77.7 %
$M_1$	0.916	0.312	80.0 %	73.2 %	76.8 %
	$\theta_{\mathcal{O}}$	$\Delta\theta_{\mathcal{O}}$	$\mathcal{F}_{\mathcal{O}}$ in $x$	$\mathcal{F}_{\mathcal{O}}$ in $y$	$\mathcal{F}_{\mathcal{O}}$ total
	$\theta_{\mathcal{L}}$	$\Delta\theta_{\mathcal{L}}$	$\mathcal{F}_{\mathcal{L}}$ in $x$	$\mathcal{F}_{\mathcal{L}}$ in $y$	$\mathcal{F}_{\mathcal{L}}$ total
$M_2$	0.744	1.433	58.4 %	74.4 %	66.4 %
$M_2$	0.978	0.540	47.9 %	74.2 %	61.3 %
	$\theta_{\mathcal{O}}$	$\Delta\theta_{\mathcal{O}}$	$\mathcal{F}_{\mathcal{O}}$ in $x$	$\mathcal{F}_{\mathcal{O}}$ in $y$	$\mathcal{F}_{\mathcal{O}}$ total

Table 6.3: Comparison of PE results of linear overlap  $\mathcal{F}_{\mathcal{L}}$  and "exact" overlap  $\mathcal{F}_{\mathcal{O}}$  for models  $M_1$  and  $M_2$ .

## 6.2 Biokinetic models

**Modelling problems.** Biokinetic models describe chemical reactions performed by and between microorganisms like bacteria (c.f. [24]). They represent parts of the metabolism processes within an organic cell. Presently, one is aware of several thousands of subprocesses within the organism. The numbers range from 3000 to 4000. Therefore, macroscopic biokinetic models are difficult to determine. Even worse due to the complexity of the underlying process, one has to anticipate some model–data–deviation and therefore some model uncertainty. Further, if parameters in biokinetic models can somehow be associated with metabolism rates, then they have to be assumed to be distributed, as different individuals show different behaviors and therefore represent an inhomogeneous sample. Therefore, the application of the model–data–overlap is suitable.

**Models.** The overlap concept will now be applied to discriminating biokinetic models. A simple example for a dynamical system modelling such biological processes is the following example

$$\frac{d}{dt} X(t) = \mu(x(t), S(t)) \cdot X(t) - k_d \cdot X(t) \quad (6.6)$$

$$\frac{d}{dt} S(t) = -\frac{\mu(X(t), S(t))}{y_{xs}} \cdot X(t) - m_s \cdot X(t). \quad (6.7)$$

$S$  denotes the substrate that is transformed into biomass  $X$ , where  $S(t)$  and  $X(t)$  denote the associated concentrations. The rate of the transformation process is described by the kinetic  $\mu(X, S)$ . The following three types of kinetics

$$\mu(X(t), S(t)) = \mu_{\max} \cdot S(t)^r \quad (6.8)$$

$$\mu(X(t), S(t)) = \mu_{\max} \cdot \frac{S(t)}{k_s + S(t)} \quad (6.9)$$

$$\mu(X(t), S(t)) = \mu_{\max} \cdot \frac{S(t)}{k_s \cdot X(t) + S(t)} \quad (6.10)$$

are to be discriminated. The kinetics given in (6.9) and (6.10) are known as the Monod and Contois kinetics, respectively (c.f. [24]). The third candidate in (6.8), also referred to as mass action kinetic, does show a significantly different behavior for example asymptotically one observes  $S \rightarrow \infty$ . It therefore can be considered as a certain case of a chemically inappropriate model. To complete the introduction of the kinetics, it shall be mentioned that the constant  $\mu_{\max}$  denotes the maximum growth rate and  $Y_{X/S}$

$$Y_{X/S} = \frac{\text{mass of cells formed}}{\text{mass of substrate consumed}}$$

transformation rate of substance into biomass. Other growth kinetics could be used for discrimination ([24]) like

$$\mu = \mu_{\max} \cdot \frac{1}{1 + k_S \cdot S(t)^r} \quad (6.11)$$

$$\mu = \mu_{\max} \cdot \left( 1 - e^{-\frac{S(t)}{k_S}} \right) \quad (6.12)$$

$$\mu = \mu_{\max} \cdot \frac{S(t)}{k_S + S(t) + \frac{S(t)^2}{k_I}} \quad (6.13)$$

$$\mu = \mu_{\max} \cdot \frac{1}{1 + k_S \cdot S(t)^{-r}} \quad (6.14)$$

the Tessier kinetics in (6.12), the Moser kinetics in 6.14 or the substrate inhibition kinetics in (6.13) or (6.11). They are mostly an extension or approximation of the three candidates mentioned above and are therefore omitted for a better lucidity. As seen in figure 6.7, they show many of them show a similar growth scheme.

## Linear overlap

**Presto.** As mentioned in chapter 5.1, the linear overlap is calculated with the program PRESTO KINETICS. The program has been proven itself very successful

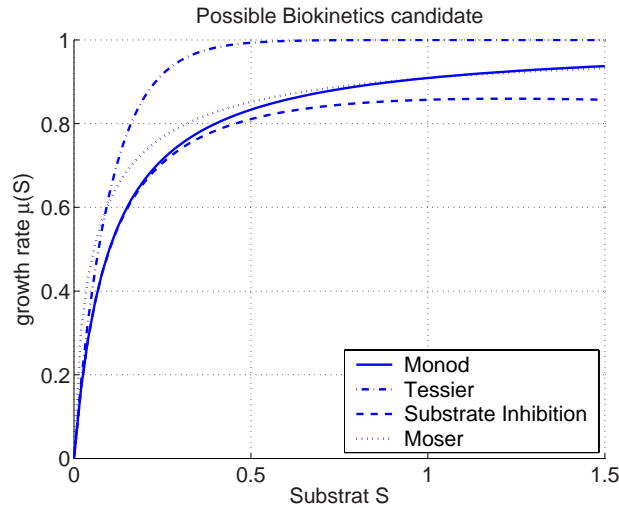


Figure 6.7: Four possible candidate biokinetics

in parameter estimation and model exploration for chemical systems in both the research and application context (c.f. [27, 51, 125, 126, 235]).

Even though the Jacobians  $\mathbf{J}_{\mathcal{F}}$  in the Gauss-Newton-steps (5.8), are relatively small for our investigated models, namely of dimension  $10 \times 28$  corresponding to the number of parameters (including their variabilities) times the numbers of measured data for each dimension, the problem are already mentioned correlations in the matrix.

For the overlap optimization, prepared data was used again, which has been generated by means of the Monod kinetics in (6.9). The data is shown in figure 6.8.

At the beginning, a number of parameter sets was drawn from an originally given parameter distribution  $\pi$ , also shown in the last line of table 6.4, corresponding to the number of data points needed. Then for each measurement point  $t_k$ , a trajectory was calculated, using each time a different drawn sets of parameters and its value at  $t_k$  was taken. The experimenter, on the other hand, was using a constant measurement error again.

For the generated data, the linear overlap  $\mathcal{F}_{\mathcal{L}}$  was optimized and the parameters  $\mu_{\max}$ ,  $k_d$ ,  $y_{xs}$ ,  $m_s$  and  $r$  as well as their variabilities were estimated. The results of the parameter estimation for the linear overlap optimization are shown in table 6.4. Additionally, a classical residual based parameter estimation with respect to  $\mathcal{F}_{\mathcal{R}}$  was conducted.

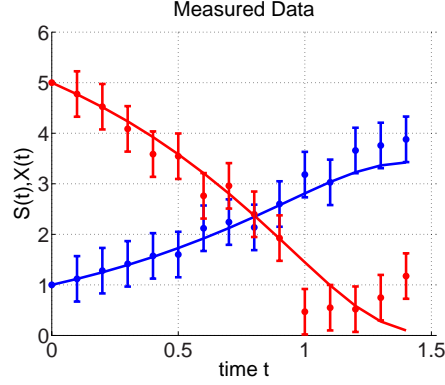


Figure 6.8: Monod kinetics with the prepared data. The substrate concentration  $S(t)$  shown in the decreasing, the biomass concentration  $X(t)$  in the increasing trajectory.

model	entity	$\mu_{\max}$	$k_s/r$	$y_{xs}$	$m_s$	$k_d$
(6.8)	parameter $\theta_{\mathcal{L}}$	0.331	0.999	0.573	0.254	$1.213 \cdot 10^{-2}$
	variability $\Delta\theta_{\mathcal{L}}$	0.076	0.150	0.094	0.648	$1.132 \cdot 10^{-5}$
	parameter $\theta_{\mathcal{R}}$	0.633	0.397	0.564	1.345	$2.179 \cdot 10^{-4}$
	standard error $\bar{\sigma}_R$	0.181	0.239	0.222	0.393	$1.014 \cdot 10^{-3}$
(6.9)	parameter $\theta_{\mathcal{L}}$	1.386	0.737	0.573	0.091	$6.102 \cdot 10^{-2}$
	variability $\Delta\theta_{\mathcal{L}}$	0.203	0.084	0.101	0.048	$3.030 \cdot 10^{-3}$
	parameter $\theta_{\mathcal{R}}$	1.262	0.793	0.613	0.377	$9.208 \cdot 10^{-6}$
	standard error $\bar{\sigma}_R$	0.262	0.639	0.125	0.278	$1.236 \cdot 10^{-4}$
(6.10)	parameter $\theta_{\mathcal{L}}$	2.028	0.981	0.525	0.059	$9.872 \cdot 10^{-2}$
	variability $\Delta\theta_{\mathcal{L}}$	0.421	0.001	0.222	0.003	$1.145 \cdot 10^{-2}$
	parameter $\theta_{\mathcal{R}}$	1.560	0.902	0.819	0.791	$1.106 \cdot 10^{-4}$
	standard error $\bar{\sigma}_R$	0.394	0.779	0.155	0.192	$1.334 \cdot 10^{-3}$
original	parameter $\theta_0$	1.400	0.730	0.600	0.090	$6.000 \cdot 10^{-2}$
	variability $\Delta\theta_0$	0.100	0.080	0.100	0.010	$1.000 \cdot 10^{-2}$

Table 6.4: Linear overlap parameter estimation with respect to  $\mathcal{F}_{\mathcal{L}}$  and classical parameter estimation with respect to  $\mathcal{F}_{\mathcal{R}}$  for biokinetics (6.8), (6.9) and (6.10).

As expected, the parameters  $\theta_{\mathcal{L}}$  including their variabilities  $\Delta\theta_{\mathcal{L}}$ , were best estimated for the Monod kinetics, the ones for Contois stayed reasonable close, whereas the one for (6.8) differ in magnitude compared to the parameters used to generate the data. The estimation methods by  $\mathcal{F}_{\mathcal{L}}$  on the one and by  $\mathcal{F}_{\mathcal{R}}$  on the other also allow for different interpretations in the parameters. For the

Monod kinetics, the parameter  $k_d$  shows an influence in the overlap setting, but was estimated lower by the residuum by the factor 1000.

In table 6.5, the targets functionals for the parameters estimated above are documented.

model		$\mathcal{F}_{\mathcal{L}}$	$\mathcal{F}_{\mathcal{O}}$		residuum
(6.8)	in X	68.97 %	62.64 %	$\mathcal{F}_{\mathcal{R}}$	0.144
	in S	58.15 %	58.68 %	avg.traj.	4.578
	total	63.56 %	60.66 %		
(6.9)	in X	77.99 %	68.15 %	$\mathcal{F}_{\mathcal{R}}$	0.159
	in S	66.96 %	55.87 %	avg.traj.	0.275
	total	72.47 %	62.02 %		
(6.10)	in X	31.71 %	19.76 %	$\mathcal{F}_{\mathcal{R}}$	0.144
	in S	38.45 %	51.52 %	avg.traj.	1.329
	total	35.09 %	35.64 %		

Table 6.5: Linear overlap parameter estimation with respect to.  $\mathcal{F}_{\mathcal{L}}$  and classical parameter estimation with respect to.  $\mathcal{F}_{\mathcal{R}}$  for the three kinetics. The values for  $\mathcal{F}_{\mathcal{O}}$  were calculated for the optimized parameters by  $\mathcal{F}_{\mathcal{L}}$  of table 6.4 to document the deviation between linear and nonlinear propagation. The residuum in the avg.traj. (left row) is the squared distance between the mean value of the data distribution and the average trajectory of (4.9) in the overlap setting.

Just knowing the values of table 6.5, the experimenter has to discriminate between the three model candidate. By merely looking at the classical residuum  $\mathcal{F}_{\mathcal{R}}$ , no model can be favored. However, according to the overlap information  $\mathcal{F}_{\mathcal{L}}$ , the Contois kinetic of (6.10) ought to be rejected. For the remaining two candidates, the squared distance between the mean values of the data distribution and the average trajectory of (4.9), abbreviated table 6.5 by avg.traj., favors the Monod kinetics (6.9). The average trajectory is closer to the data for Monod than for the other one. Combining criteria from the residual on the one and the overlap framework on the other, one is able to discriminate the given model candidates.

**Exact and linear.** Table 6.5 also documents again the differences between a linear  $\mathcal{F}_{\mathcal{L}}$  and "exact"  $\mathcal{F}_{\mathcal{O}}$  overlap calculation. Most certainly, one wants to get the exact overlap results. For small example like the ones used here, the computational effort for discriminating models by the exact overlap is possible. Due to the computational effort for complex systems, however, only the linear overlap is sensible to consider. In [227] such complex systems are documented for the model–data–overlap.

**Optimization.** The computations for the example have shown, that it is sensible to consider a target functional of the form (5.10) first. Then with the help of the estimated hyperparameters for  $\mathcal{F}_{\mathcal{L}}$ , the optimization of  $\mathcal{F}_{\mathcal{O}}$  shall be started. By this consecutive procedure, the iteration steps within the optimization process is shortened significantly and avoids time consuming calculations of the full overlap  $\mathcal{F}_{\mathcal{O}}$ .

### 6.3 Pharmacokinetics

Returning to an already previously mentioned aspect: Biokinetic models contribute to describe the complex metabolism systems (c.f. [101, 181]). A computational discipline that has emerged over the past years and is referred to as pharmacokinetics. There, the temporal course of concentration of agents and their degradation products in the body's tissue and fluids is investigated. These interaction dictate the applicability of a drug. In the processes of drug design, these in-silico (computer based) investigation therefore become more and more important in order to reduce unnecessary experiment with living beings.

One can consider the following compartment model in order to describe the a metabolism, as for example used by ABRAHAM ET AL. in [1, 2] as seen in figure 6.9.

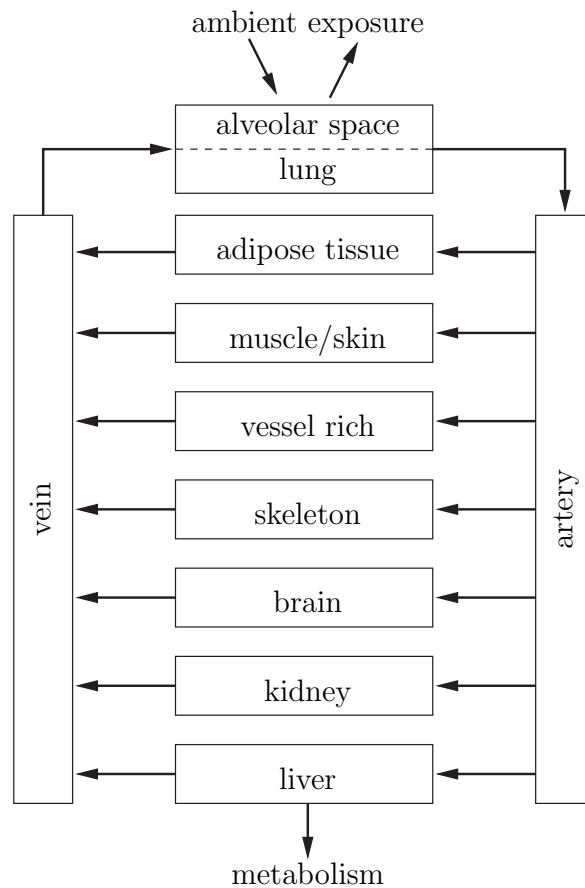


Figure 6.9: PBPK compartment model according to [1, 2]

Within their investigations, questions of parameter sensitivity are very important, when considering a population of people with individual inner physiological "parameters". A common procedure is to understand the processes for certain patients and then transfer the results to different class of people, like for example the transfer from an adult to a newborn child, where it is much more dangerous to conduct experiments. Therefore, one has to understand the performance of the model not only for one individual, but for an entire population. Therefore, one has to incorporate parameter sensitivity, for example by means of the overlap.

Further, applying the overlap leads to a different averaging philosophy, which is a consequence of the optimization scheme shown in figure 4.4. In many settings, the initial parameters are averaged first and then propagated by the kinetics. The reverse is done within the overlap concept. For that reason, Bayesian approach are so popular within pharmacokinetics. Nevertheless, the points concerning model

uncertainty, as mentioned in section 4.1, can be raised again and the overlap is favored here.

In the following example<sup>2</sup>, an adults (73 kg) was exposed to styrene (VOC) for 4 hours. Then the saturable metabolism in the liver was investigated. Two different models are under consideration: First, a metabolism according to Csanady [63], abbreviated by *C* and second a metabolism according to Ramsey and Anderson [196], abbreviated by *R*. Both models also suggest physiological parameters that were determined beforehand and show different parameter sensitivities.

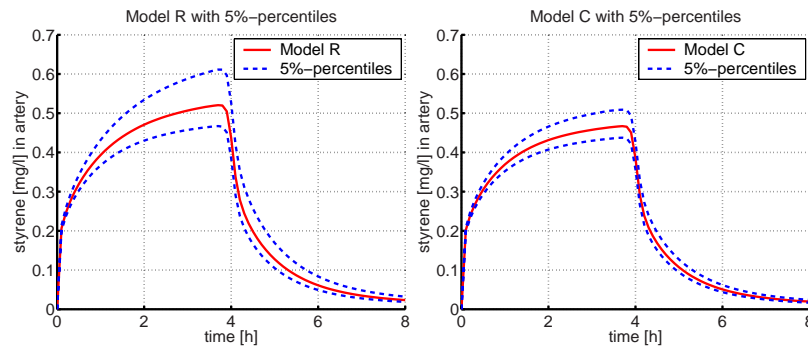


Figure 6.10: Metabolism models and their sensitivities according to Csanady [63] as well as Ramsey and Anderson [196]

For reasons of data disclosure, the following data are slightly modified. The arterial concentration of styrene was measured hypothetical 15 times at  $t = 1, 2, 3, 6, 5 h$ .

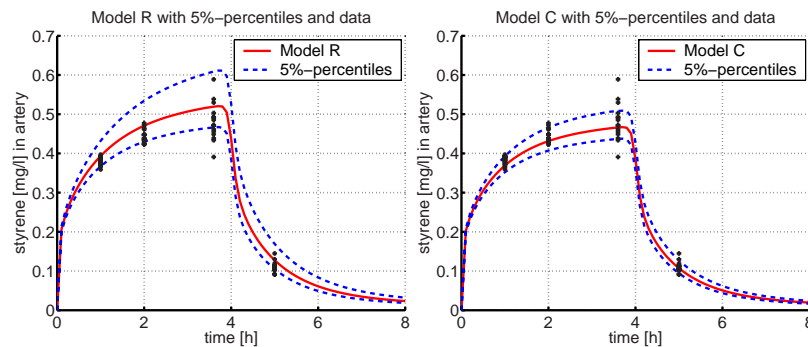


Figure 6.11: Propagated densities

<sup>2</sup>Also shown on the poster: "A new approach for incorporating variability into model discrimination" by Wilhelm Huisinga, Illia Horenko and Sönke Lorenz



The data can be fitted equally well to both models in the residual framework. However, when looking at the parameter sensitivity and the resulting model variability, one sees that at some time points,  $C$  is more suitable (for example  $t = 1$ ) and at other points,  $R$  should be preferred (for example  $t = 3.6$  h in figure 9).

Calculating the model variability for the entire model as seen in figure 9,

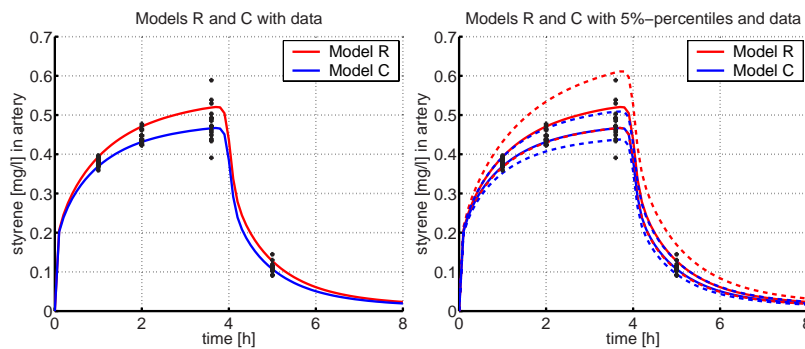


Figure 6.12: Data and model variability

a gets a model–data–overlap for model  $C$  of 0.84 and for  $R = 0.51$ . Therefore, model  $C$  would be preferred.

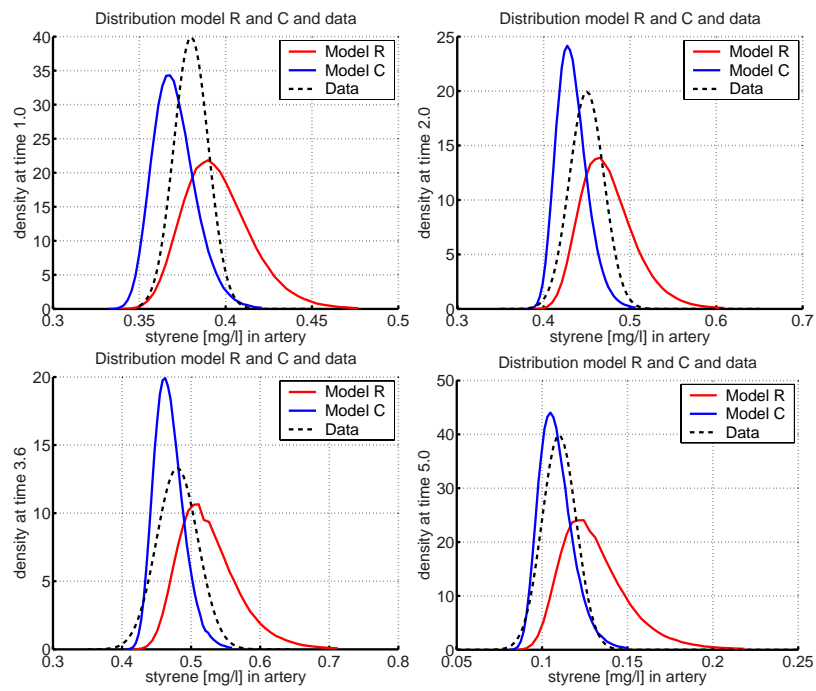
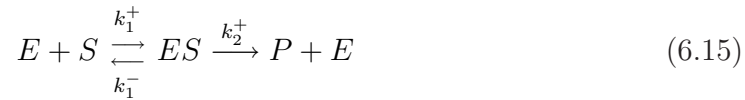


Figure 6.13: Propagated densities

## 6.4 Michaelis-Menten-Kinetics

In the following section, the TRAIL algorithm shall be illustrated for the Michaelis-Menten-Kinetics, which was introduced by MICHAELIS and MENTEN in 1913 (c.f. [24, 85, 175]) and is a key component in numerous biochemical experiments and describes the conversion of substrate  $S$  to product  $P$  catalyzed by some enzyme  $E$ . The reaction takes part in two consecutive part-reactions. First, substrate  $S$  and enzyme  $E$  form an enzyme-substrate complex  $ES$ . Then, that complex  $ES$  decomposes irreversibly into the product  $P$  and the enzyme  $E$ .



Therefore, the total concentration of the enzyme  $[E]_T$  as well as of the substrate  $[S]_T$  are conserved with:

$$[E]_T = [E] + [ES] \quad (6.16)$$

$$[S]_T = [S] + [ES] + [P]. \quad (6.17)$$

Since the second reaction is irreversible, only kinetic equations

$$\frac{d}{dt}[S] = -k_1^+[E]_T[S] + (k_1^- + k_1^+[S])[ES] \quad (6.18)$$

$$\frac{d}{dt}[ES] = k_1^+[E]_T[S] - (k_1^- + k_2^+ + k_1^+[S])[ES] \quad (6.19)$$

with the constants  $k_1^-$ ,  $k_1^+$ ,  $k_2^+$  as well as  $[E]_T$  are needed to mathematically describe the problem.

For a negligible ratio of total concentration of enzyme to substrate, one can derive the Monod kinetic as a limit (c.f. [85]).

The previously described Michaelis-Menten-kinetics is now used to illustrate and compare the extended TRAIL algorithm. By introducing normally distributed initial concentration values for the substrate  $[S]$  and the complex  $[ES]$

$$[S](0) \sim \mathcal{N}\left(4, \frac{1}{\sqrt{10}}\right), \quad [ES](0) \sim \mathcal{N}\left(1, \frac{1}{\sqrt{6}}\right)$$

and the reaction parameters  $k_1^+ = 0.5$ ,  $k_1^- = 0.3$ ,  $k_2^+ = 2$ ,  $[E]_T = 3$  the problem suits the setting in (5.18). Three numerical methods have been applied to solve the resulting partial differential equation:

- Finite element method (FEM) with a uniform  $200 \times 200$ -grid

- Path Monte–Carlo–sampling of the initial values and trajectory propagation (650.000 trajectories)
- Modified TRAIL–algorithm

All three methods produced almost identical results as it can be seen in figure 6.14.

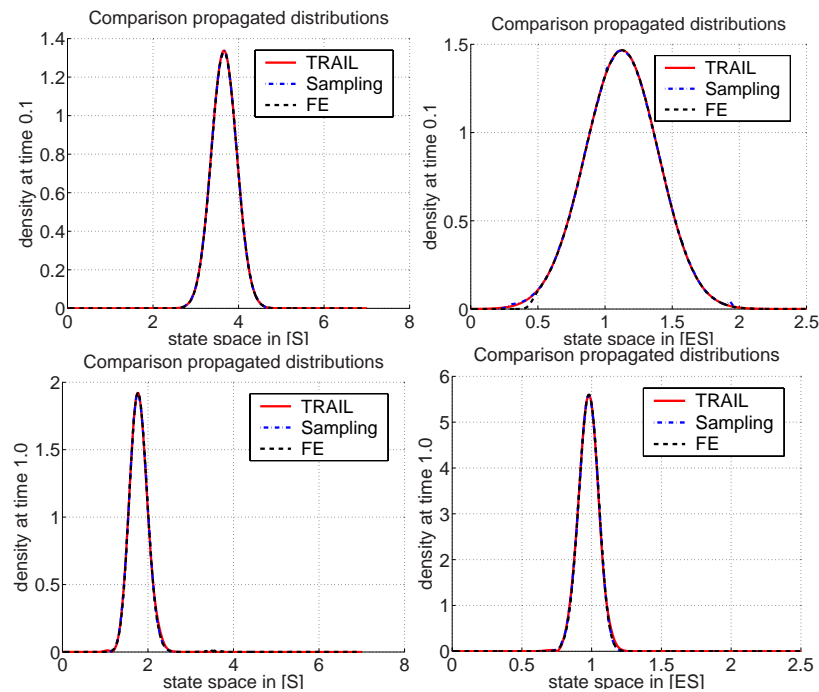


Figure 6.14: Propagated densities at  $t = 0.1$  s and  $t = 1.0$  s

Although all the three methods produce almost identical results, the effort to compute them differs drastically. Whereas the TRAIL algorithm was finished within 2 minutes, the other more traditional methods calculated the appropriate results within some hours. Thus, the TRAIL algorithm has been proven to be a time efficient alternative to the traditional methods, especially when increasing the number of state space variables.

**Numerical experiments revisited.** The Michaelis-Menten-kinetic is used now to illustrate the density propagation for the extended problem. The ini-

tial values for  $[S]$  and  $[ES]$  as well as the four parameters  $k_1^-$ ,  $k_1^+$ ,  $k_2^+$  as well as  $[E]_T$  are now assumed to be normally distributed with

$$\begin{aligned} [S](0) &\sim \mathcal{N}\left(8, \frac{1}{\sqrt{20}}\right), [ES](0) \sim \mathcal{N}\left(2, \frac{1}{\sqrt{20}}\right), \\ k_1^- &\sim \mathcal{N}\left(0.5, \frac{1}{\sqrt{20}}\right), k_1^+ \sim \mathcal{N}\left(0.3, \frac{1}{\sqrt{20}}\right), \\ k_2^+ &\sim \mathcal{N}\left(0.7, \frac{1}{\sqrt{20}}\right) \text{ and } [E]_T \sim \mathcal{N}\left(2, \frac{1}{\sqrt{20}}\right). \end{aligned}$$

By incorporating the four parameters, the problems now becomes six dimensional. Therefore, the finite element method has to be rejected as a validation method, only the Monte Carlo simulation remains. The result of the TRAIL and the Monte Carlo simulation for the given setting is shown in figure 6.15.

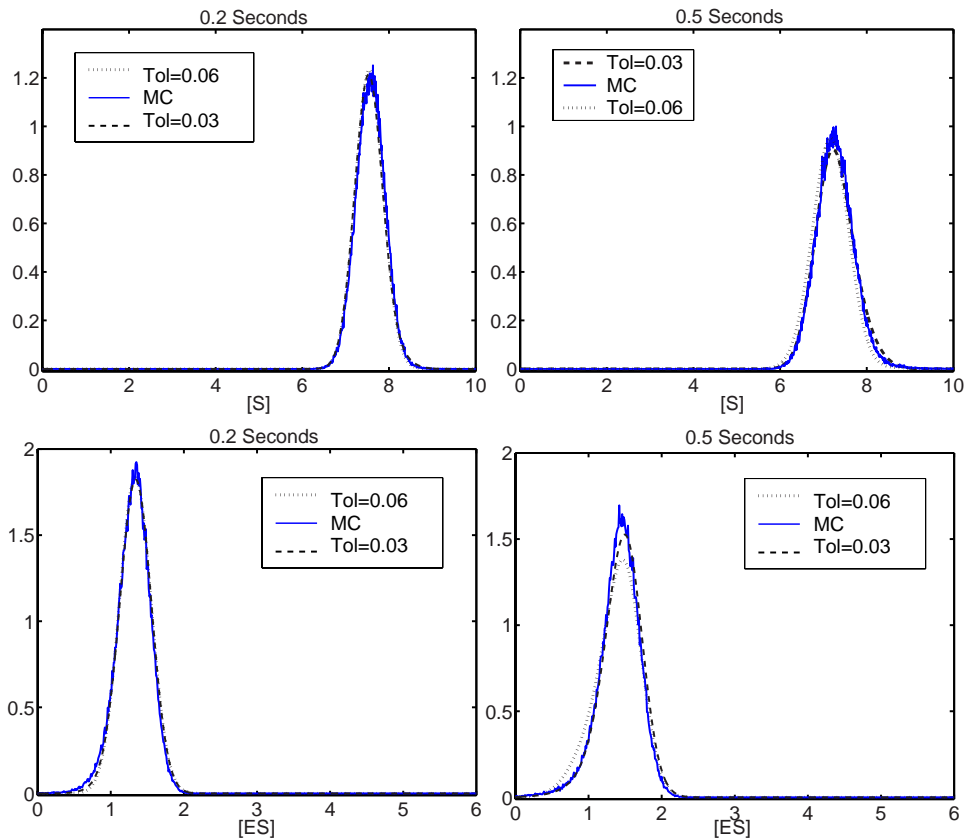


Figure 6.15: Propagated densities at  $t = 0.2$  s and  $0.5$  s

The TRAIL algorithm's results correspond to the Monte Carlo results. However,

the approximation quality at the transition from the bulk to the quantiles of the distribution did increase in comparison to the two-dimensional example at the beginning of this section. However, the computation time difference did remain. Whereas the Monte Carlo algorithm was calculating several hours, it took the TRAIL algorithm a few minutes to calculate the propagated distribution.

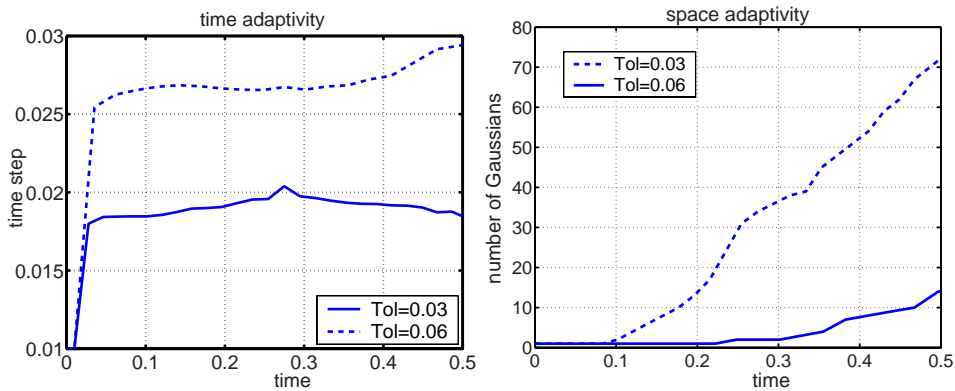


Figure 6.16: Adaptivity in time and space

**Conclusion.** In this section, an extension of the TRAIL algorithm was derived and illustrated. This allowed us to implement a fully adaptive particle method where the number of involved particles is controlled by a given tolerance. Compared to frequently used Monte–Carlo–methods, a much shorter computation time was needed to conduct nonlinear propagation and enables to control its quality. A very important feature of TRAIL is that the numerical effort scales with the number of particles used and is independent of the problem’s dimension.

However, problems involving distribution functions with a large degree of spatio-temporal details tend to increase the number of needed Gaussian distributions significantly and consequently increase the numerical effort. Moreover for highly nonlinear problems with multiple time scales, the Gaussian shape matrix’s stiffness deteriorates the prediction quality (c.f. 5.26), which results in a higher computational effort to compensate this. Both of these problems are connected to the choice of Gaussians as basis functions. However, the TRAIL–algorithm is not restricted to Gaussian particles, but can rather be used for any set of distributions, assuming an reliable analog of the linear predictor (c.f. (5.27) – (5.29)) exists.