Optimal Experimental Design and Model Selection by a Sigma Point approach

R. Schenkendorf, A. Kremling, M. Mangold
Max-Planck-Institute for Dynamics of Complex Technical Systems, Magdeburg,

Corresponding author: R. Schenkendorf, Max-Planck-Institute for Dynamics of Complex Technical Systems, Process Synthesis and Process Dynamics
39106 Magdeburg, Sandtorstraße 1, Germany, rschenke@mpi-magdeburg.mpg.de

Abstract. The identification of a model structure, i.e., the relationship of model components, and the usually unknown model parameters are essential to get meaningful deterministic models. Due to imperfect and limited measurement data to fulfil both requirements can be a challenging task. Therefore, the Optimal Experimental Design (OED) was introduced to detect some optimal stimuli of the regarded system, that increase the information content of measurements. For models which are linear with respect to their parameters, the OED is well known to reduce the influence of measurement noise to the parameter estimation process and to find conditions that make the difference of competing model hypothesis more obvious. Unfortunately, in practise the most used models are non-linear with respect to their parameters. The established methods for the linear case, e.g. the Fisher Information matrix (FIM), can lead to a crude approximation of the non-linearity that results at best in a sub-optimal choice of the system stimuli.
The usage of sample based approaches, e.g. Monte Carlo methods, seems to be more suitable to consider non-linear effects, but their inherent computational effort prohibits the application in the framework of Optimal Experimental Design. In this work, the Sigma Point method is used instead, because its advantages are at least threefold for OED: 1) a more reliable approximation of parameter uncertainties can be achieved, 2) a definition of a novel design criterion for OED becomes possible, and 3) it provides an efficient and elegant way to take the imperfection of parameter estimation in the field of model selection into account. All three aspects are demonstrated for the example of a biological substrate uptake model.

1 Introduction

Complex biological systems can be described and analysed by deterministic models. The process of model development requires not only knowledge about the model structure but also about related model parameters θ. Even with a suitable model structure, parameter identification can be a challenging task. Uncertainty about experimental data leads to uncertainty about estimated parameters $\hat{\theta}$, consequently, the estimates can be represented by confidence regions instead of single values. To get meaningful models these confidence regions should be as small as possible. For this purpose, Optimal Experimental Design (OED) provides new experimental conditions, e.g. the system stimuli $u(t)$, resulting in more precise estimates.
To define a cost function for Optimal Experimental Design, it is necessary to express the parameter uncertainties by mathematical terms, leading to the (co)variance matrix $C_{\hat{\theta}}$ and the mean $E[\hat{\theta}]$ of the estimated parameters. A frequently applied method for this is based on the inverse of the Fisher information matrix (FIM) [1]. However, this method leads in many cases to a poor approximation of the parameter (co)variances and does not provide any information about the mean. To find a more realistic approximation of the parameter statistics, methods have been developed, which are based on Monte Carlo simulation as the Bootstrap approach [2] and the Global Sensitivity Analysis (GSA) [3]. These approaches have a very high computational effort, which prohibits their use in the framework of OED.
A more suitable approximation can be achieved using the Sigma Point method [4], which has been introduced recently in the field of OED [5, 6]. Julier and Uhlmann suggested the use of the Sigma Points in order to determine the mean and (co)variance of a random variable $\eta \in \mathbb{R}^l$ from the mean and (co)variance of a random variable $\xi \in \mathbb{R}^f$, where η is related to ξ by the non-linear mapping $\eta = g(\xi)$. Adjusted to parameter identification, ξ represents all available measurement data[1], $g(\cdot)$ stands for the complete parameter identification process and η is the resulting vector of estimated model parameters. By using this analogy, the Sigma Point method can be applied directly to calculate the mean and the (co)variances of $\hat{\theta}$.
Moreover, the basic concept of the Sigma Point method enables the investigation of the influence of parameter uncertainties on simulation results $x(t)$. The determination of confidence regions of the dynamic states becomes possible and can be used to evaluate novel cost functions for the OED process. Beyond this, the confidence regions of the states provide an efficient and elegant way to take the imperfections of parameter estimation in the field of model selection into account. In the next section, general concepts of OED are introduced. After this, the potential of the Sigma Point approach in the field of Optimal Experimental Design is demonstrated by a simple substrate uptake model.

[1] Measurements of m components at K time points ($f = m \cdot K$)

2 Methods

2.1 Parameter Identification and Parameter Uncertainties

Ordinary differential equations (o.d.e.'s) Eq.(1) are widely used to describe the behaviour of dynamic systems.

$$\frac{dx(t)}{dt} = f[x(t),u(t),\theta] \tag{1}$$

In Eq.(1) the o.d.e system is a function of the dynamic states $x(t) \in \mathbb{R}^n$, the system input $u(t) \in \mathbb{R}^r$, and model parameters $\theta \in \mathbb{R}^p$. In addition, the output function $\hat{y}(t) \in \mathbb{R}^m$ Eq.(2) defines states or combinations of states that are measurable, where $v \in \mathbb{R}^m$ denotes measurement noise.

$$\hat{y}(t) = h[x(t),u(t),v(t),\theta] \tag{2}$$

Even with a detailed knowledge of the underlying processes, the model parameters θ are usually a priori unknown and have to be identified. For this purpose, a cost function can be defined by Eq.(3), which weights the difference of simulation results $\hat{y}(t)$ and measurement data $y(t)$ by the inverse of the measurement variance matrix C_y.

$$J(\hat{\theta}) = \frac{1}{2}\left(y-\hat{y}(x_0,u,\hat{\theta})\right)^T C_y^{-1}\left(y-\hat{y}(x_0,u,\hat{\theta})\right) \tag{3}$$

By minimising this cost function, estimates of the unknown parameters $\hat{\theta}$ are generated. In practise, measurement noise v and a limited number of measurement data lead to parameter uncertainties that can be expressed by the mean square error (MSE) matrix:

$$MSE_{\hat{\theta}} = E\left[\left(\hat{\theta}-\theta\right)\left(\hat{\theta}-\theta\right)^T\right] \tag{4}$$

$$= C_{\hat{\theta}} + (E[\hat{\theta}]-\theta)\cdot(E[\hat{\theta}]-\theta)^T \tag{5}$$

$$= C_{\hat{\theta}} + bi\cdot bi^T, \tag{6}$$

where $E[\cdot]$ is the expected value of a random variable, $C_{\hat{\theta}}$ represents the (co)variance matrix of the estimates and bi is known as the bias. Obviously, the mean square error matrix would be a perfect measure of the reliability of the estimates, but unfortunately, for the definition of the $MSE_{\hat{\theta}}$ matrix, the 'true' parameter vector θ is needed. Therefore, one main part of OED is to find an appropriate approximation of $MSE_{\hat{\theta}}$.

2.2 Approximation by the Fisher Information matrix

A frequently applied method to get an approximation of the mean square error matrix $MSE_{\hat{\theta}}$ is based on the Fisher Information matrix Eq.(7). This expression weights the parameter sensitivity matrix Eq.(8), i.e., partial derivation of the output function $\hat{y}(t)$ to the parameters θ, by the inverse of the measurement variance C_y over several time points t_k.

$$FIM = \sum_{t_k} S_{t_k}^T \cdot C_y^{-1} \cdot S_{t_k} \tag{7}$$

$$S_{t_k} = \left[\left.\frac{\partial \hat{y}}{\partial \theta}\right|_{t_k}\right] \tag{8}$$

The inverse of FIM related to the Cramer-Raó inequality Eq.(9) provides a measure of the parameter (co)variance matrix C_θ.

$$C_\theta \geq \frac{\partial E\left[\hat{\theta}\right]}{\partial \theta} FIM^{-1} \frac{\partial E\left[\hat{\theta}\right]^T}{\partial \theta} \tag{9}$$

The equality of Eq.(9) only holds, if (i) the measurement errors are additive, (ii) the model is linear in its parameters, and (iii) a minimum variance unbiased estimator is available. Especially, the last two requirements are not fulfilled in practise. As the FIM does not give any information about $E\left[\hat{\theta}\right]$, the estimate is assumed to be unbiased, i.e., $E\left[\hat{\theta}\right] = \theta$, leading to the simplified approximation of $MSE_{\hat{\theta},FIM}$ Eq.(10).

$$MSE_{\hat{\theta},FIM} \approx C_\theta \geq FIM^{-1} \tag{10}$$

At the best Eq.(10) provides a lower bound of the parameter uncertainties. An Optimal Experimental Design, that is based on this approximation is likely to end in a suboptimal result, i.e., parameter uncertainties can be reduced but there is still potential for a further reduction. Alternative approaches are required.

2.3 Approximation by the Sigma Point method

To get an appropriate approximation of $MSE_{\hat{\theta}}$, the non-linear relation of parameters θ to the model and also the imperfection of the used optimiser have to be taken into account. For this purpose, Julier and Uhlmann suggested the use of the so-called Sigma Points, i.e., samples of ξ, in order to determine the mean and (co)variance of a random variable $\eta \in \mathbb{R}^l$ from the mean and (co)variance of a random variable $\xi \in \mathbb{R}^f$, where η is related to ξ by the non-linear mapping $\eta = g(\xi)$. In the context of parameter identification ξ represents all available measurement data:

$$\xi \quad = \quad [y_1(t_1),\ldots,y_1(t_K),y_2(t_1),\ldots,y_2(t_K),\ldots,y_m(t_1)\ldots,y_m(t_K)]^T \qquad f = m \cdot K,$$

$g(\cdot)$ stands for the complete parameter identification process and η is the resulting vector of estimated model parameters. By using this analogy, the Sigma Point method can be applied directly to calculate the mean and the (co)variances of $\hat{\theta}$. In contrast to other sample based approaches, e.g. Monte Carlo methods, the Sigma Points are deterministically chosen by the statistical information about the measurement imperfection. This kind of intelligent sample selection leads to a dramatical reduction of the required sample number $(2 \cdot f + 1)$ and the related computational effort. As shown in [6], the following approximation of $MSE_{\hat{\theta}}$ by the Sigma Point method can be found:

$$MSE_{\hat{\theta},SP} \quad \approx \quad \hat{C}_{\hat{\theta}} + \hat{bi} \cdot \hat{bi}^T, \tag{11}$$

where $\hat{C}_{\hat{\theta}}$ and $\hat{bi}$ are determined approximately by the statistics of η. In contrast to $MSE_{\hat{\theta},FIM}$, this approximation is not merely only a lower bound of parameter uncertainties related to an idealised optimiser, but rather a more reliable estimation of parameter deviations that considers the bias and all other imperfections of the used optimiser.

2.4 Minimising the mean square error matrix

If the mean square error matrix indicates, that the estimated parameters $\hat{\theta}$ possess quite large uncertainties, i.e., high values of the diagonal elements of $MSE_{\hat{\theta}}$, the framework of OED tries to reduce them iteratively. For this, an optimisation problem has to be solved, that increases the measurement information content by finding optimal stimuli of the regarded system. As a matrix can hardly be used to define a cost function for a numerical optimiser, well known criteria [7] exist leading to a scalar measure of $MSE_{\hat{\theta}}$. Unfortunately, no advice can be given which kind of scalar measure is the most beneficial approach for OED. Therefore, only the so-called modified E-criterion is introduced Eq.(12), as it provides an a-priori known optimum value ($\Phi_{E^*,opt}(MSE_{\hat{\theta}}) = 1$) and an easy to visualise assessment of the result of OED for a two dimensional parameter space, where λ_{max} (λ_{min}) is the maximum (minimum) eigenvalue of $MSE_{\hat{\theta}}$.

$$\Phi_{E^*}(MSE_{\hat{\theta}}) \quad = \quad \frac{\lambda_{max}(MSE_{\hat{\theta}})}{\lambda_{min}(MSE_{\hat{\theta}})} \tag{12}$$

After a successful OED, based on $\Phi_{E^*}(MSE_{\hat{\theta}})$, the contour plot of the cost function of parameter identification Eq.(3) visualised over a two parameter space should be as circular as possible, i.e., there is no correlation between both parameters.
Usually, the dimension of the parameter vector θ is much larger than two and not every parameter has the same impact on the predictive power of the model. In this context, predictive power means, that a model is able to describe the dynamics of a system appropriately even under conditions that were not part of a former parameter identification. As only a few linear combinations of unknown parameters can dominate the qualitative behaviour of a model [8, 9], a parameter sensitivity analysis could be used to group them into an important and unimportant set, i.e., parameters that have to be estimated as precisely as possible and parameters that are assumed to be known and fixed, respectively. Even for this task, the Sigma Point method provides an elegant way to define a scalar measure of $MSE_{\hat{\theta}}$ with an inherent parameter sensitivity approach. After the determination of parameter statistics by the mean square error matrix, the general concept of the Sigma Point method is used to transfer parameter uncertainties to the system states $x(t)$. Adjusted to this requirement, ξ represents all estimated parameters $\hat{\theta}$, $g(\cdot)$ stands for the solver of the o.d.e. system Eq.(1), and η_{t_k} is the resulting vector of the system states at the time point t_k. On the analogy of Eq.(11), a mean square error matrix of the system states is defined:

$$MSE_x \quad \approx \quad \sum_{t_k} MSE_{x(t_k)} = \sum_{t_k} \hat{C}_{x(t_k)} + \hat{bi}_{x(t_k)} \cdot \hat{bi}^T_{x(t_k)}. \tag{13}$$

Now the trace of MSE_x can be minimised by searching suitable system stimuli. Parameters that have a strong impact on the predictive power of a model, i.e., parameters leading to large entries of MSE_x, are automatically selected and their uncertainties are reduced.

2.5 Model Selection

In the previous sections we assumed that the model structure, i.e., the interaction of model components, is known and that a disagreement of simulation results and measurement data is caused by imprecise model parameters. In

practise, this is a quite idealised assumption. Usually, at the beginning of model development, the underlying basic processes of the considered system are only vaguely known. Consequently, different model hypothesis can exist, that are able to describe the related system under given conditions equally well.
The task of OED is to find conditions, i.e., system stimuli, that make competing model candidates distinguishable. This field of OED is known as model selection or model discrimination [10]. The general problem is to find a suitable measure of model differences, that can be used as a cost function for OED. Several approaches already exist, where some of them are motivated by the use of Bayesian methods [11], other one are related to the information theory [12], and finally some are purely deterministic,i.e., goodness-of-fit test. Nevertheless, it seems fruitful for the authors, to combine statistics and deterministic ideas. By the definition of the mean square error matrix of the system states (Eq.(13)) we have already introduced statistics to a deterministic model. Now, a system state is characterised by a mean and confidence region. This additional information is used to define a cost function for OED. The same system state at time point t_k of different model candidates should possess distinguishable values, whereby the related confidence regions of this state provide a measure how large this difference should be to get meaningful results. In Fig. 1 one can see the probability distribution $pr(\cdot)$ of the same output function $\hat{y}_i(t_k)$ of two competing model candidates (M_1 and M_2). Both distributions cover almost the identical range of the output space, leading to a large overlap. Even with a proper measurement data, no model can be selected as the more suitable one. Only after a reduction of the overlap (Fig.2), the second model M_2 can be clearly identified as the appropriate model by measurement data.

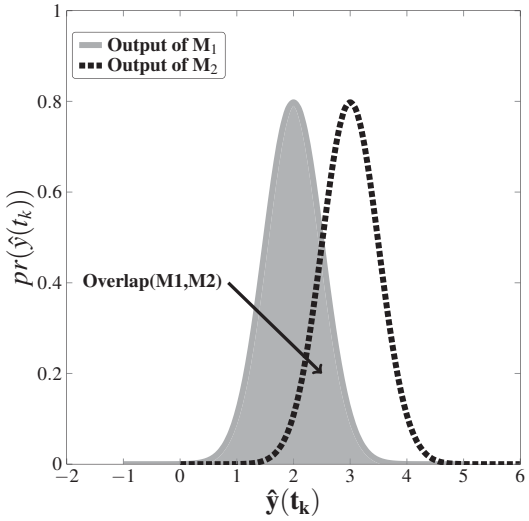

Figure 1: Illustration of the Overlap of two model output probability distribution functions, where both functions cover almost the same output space.

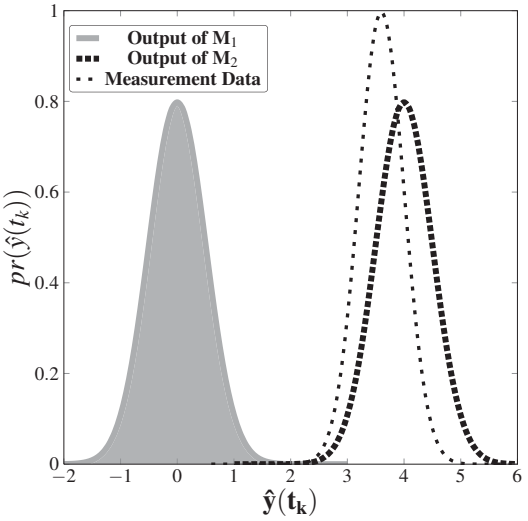

Figure 2: Only a reduced Overlap enables a decision which model candidate is more likely compared to measurement data.

The overlap of the system states of competing model hypothesis is an ideal measure of model differences. Furthermore, this approach takes the imperfection of parameter estimation in the filed of model selection explicitly into account. A mathematical expression of the overlap was recently introduced in [13] to determine the overlap of measurement data and model output. Adjusted to our requirements, the definition of the overlap for a system state x_j is given by:

$$O_j = \sum_{t_k} \frac{2 \cdot \sqrt{C^{jj}_{x_{M_1}(t_k)} \cdot C^{jj}_{x_{M_2}(t_k)}}}{C^{jj}_{x_{M_1}(t_k)} + C^{jj}_{x_{M_2}(t_k)}} exp\left[\frac{-0.5 \cdot (E[x^j_{M_1}(t_k)] - E[x^j_{M_2}(t_k)])^2}{C^{jj}_{x_{M_1}(t_k)} + C^{jj}_{x_{M_2}(t_k)}}\right], \tag{14}$$

where $C^{jj}_{(\cdot)}$ the j'th diagonal element of the related (co)variance matrix and $E[(\cdot)^j]$ the j'th element of the expectation vector. To incorporate overlap regions of different system states the sum of Eq.(14) is used:

$$O_\Sigma = \sum_j O_j. \tag{15}$$

As already mentioned, all needed information to calculate the overlap (Eq.(14)) is given by the Sigma Point approach. Uncertainties about estimated parameters, that were determined in advance by the Sigma Points, are transformed to the system states. The resulting statistics of the states is used to make competing model hypothesis more distinguishable. Furthermore, if no measurement data can be explained by the confidence regions of the system states, the model structure has to be corrected, as it remains the only source of disagreement.

3 Application example

The application example of the Sigma Point method in the field of Optimal Experimental Design is demonstrated by a simple unstructured growth model of a cell population for a continuous stirred tank bio-reactor:

$$\dot{c}_B = \mu \cdot c_B - D \cdot c_B \tag{16}$$

$$\dot{c}_S = -\frac{1}{Y_{B|S}} \cdot \mu \cdot c_B + D \cdot (c_{s,in} - c_s) \tag{17}$$

$$\mu = \mu_{max} \cdot \frac{c_S}{c_S + K_S}, \tag{18}$$

where c_B is the concentration of biomass, c_S the concentration of the substrate, D the dilution rate, and $Y_{B|S}$ the yield factor, that is considered as given by literature. The two model parameters, μ_{max} (the maximum growth rate) and K_S (the substrate affinity constant) of the Monod kinetics (Eq.(18)), are assumed to be unknown and have to be identified by an optimiser with the measurement information about the biomass concentration (corrupted by normal distributed noise, i.e., $v \sim N(0, C_y)$) at three time-points $t_k = [0.5\ 1.0\ 1.5]$h. After a first estimation, the parameter uncertainties are determined either from the Fisher information matrix or by the Sigma Point method. To judge which of the two methods provides a more realistic approximation of the (co)variance matrix $C_{\hat{\theta}}$, a Bootstrap approach is used. More then 10,000 artificial measurement samples are necessary until the Bootstrap approach tends to stabilise, i.e., the mean and the (co)variances do not change when the sample number is increased further. In contrast, the Sigma Point method needs only 7 measurement samples $(2 \cdot f + 1)$ and produces nearly the same results for different measurement uncertainties (Fig.3). This indicates that the Sigma Point method works very well and is highly effective in this case. The classical FIM approach underestimates the parameter uncertainties strongly, especially the variance of K_S. As the cost function of OED depends on the (co)variance matrix, the quality of the approximation of $C_{\hat{\theta}}$ directly affects the newly designed experiment. Consequently, the SP-designed experiment (only the inlet flow is assumed as the design variable) generates a tighter and more roundish cost function (sum of squared errors) for the parameter estimation process (Fig.4), that is equivalent to a reduction of parameter uncertainties and of the correlation between this two parameters.

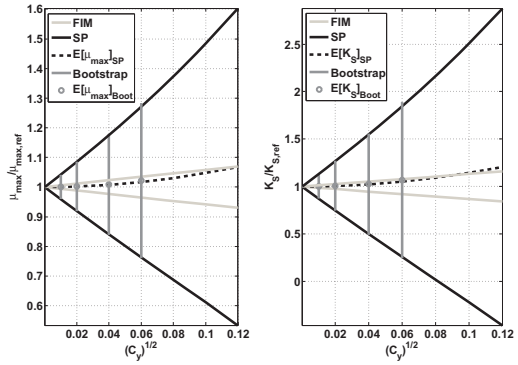

Figure 3: Confidence regions of μ_{max} and K_S determined by FIM, Sigma Point method and Bootstrap approach as a function of measurement uncertainties, i.e., the standard derivation $\sqrt{C_y}$

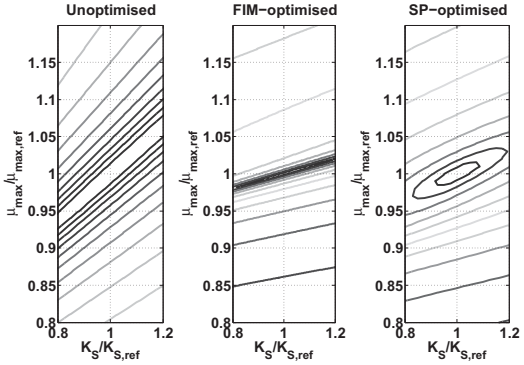

Figure 4: Contour plots of the parameter estimation cost function (sum of squared errors) for the unoptimised, FIM-optimised, and SP-optimised experiment

After a successful determination of parameter uncertainties by the Sigma Point method, the same approach is used to transform parameter statistics to the system states. Instead of single values the states are described by a mean $E[x]$ and a related (co)variance matrix C_x. This information is sufficient to determine the introduced novel cost function Eq.(13). In Tab. 1 related values of Eq.(13) are given for the unoptimised, the FIM-optimised and the Sigma Point-

	Unoptimised	FIM-optimised	SP-optimised
$\frac{trace(MSE_x)}{[trace(MSE_x)]_{SP-opt.}}$	109.51	4.09	1

Table 1: Normalised values of the introduced cost function (Eq.(13)) for the unoptimised, FIM-optimised, and SP-optimised experiment.

optimised experiment. Obviously, the Sigma Point-optimised experiment possesses the smallest entry, i.e., tight confidence regions of the system states due to an improved estimation process by OED. In [14] a demonstration of OED can be found, where the search of an optimal system stimuli is directly based on Eq.(13) for a complex biological two substrate model. In a second step, the statistical information is used in the framework of model selection. In the simplest case, the goodness-of-fit approach is used to select between competing models or to find optimal stimuli of the system which increase the distinguishability of different model hypotheses. Nevertheless, some competing models possess quite similar goodness-of-fit values, that are based on the difference between

simulation results and measurement data. In this context, the usage of the state confidence regions can be helpful to select between these models. To get two model candidates (M_1 and M_2) the Monod kinetics (Eq.19) is modified by a substrate inhibition (Eq.20). For the case of simplicity a batch process is assumed (D=0) and artificial measurement

$$\mu_{M_1} = \mu_{max} \cdot \frac{c_S}{c_S + K_S} \quad (19) \qquad ; \qquad \mu_{M_2} = \mu_{max} \cdot \frac{c_S}{c_S + K_S + \frac{c_S^2}{K_I}} \quad (20)$$

data of biomass concentration are generated by the original Monod model corrupted by measurement noise ($v \sim N(0, 0.1)$). Under unsuitable initial conditions, i.e., low initial concentration of biomass and substrate, the two models describe a similar dynamic behaviour and show almost equal state confidence regions (Fig.5), that are caused by parameter uncertainties. No advice can be given which of them is more suitable to describe the underling process. Considering the overlap approach [13], the discrimination function is used Eq.(15), that is based on mean and (co)variances of the states $O(E[x_{M_1}], C_{x_{M_1}}, E[x_{M_2}], C_{x_{M_2}})$. The objective of OED is to find initial conditions which reduce the overlap of the states confidence regions as well as in any way possible. As one could expect, an increased initial substrate concentration reduces the overlap dramatically, as the influence of the substrate inhibition becomes more and more important. Now, under optimised initial conditions the measurement data clearly prefer the Monod model (M_1), which is the correct one (Fig.6).

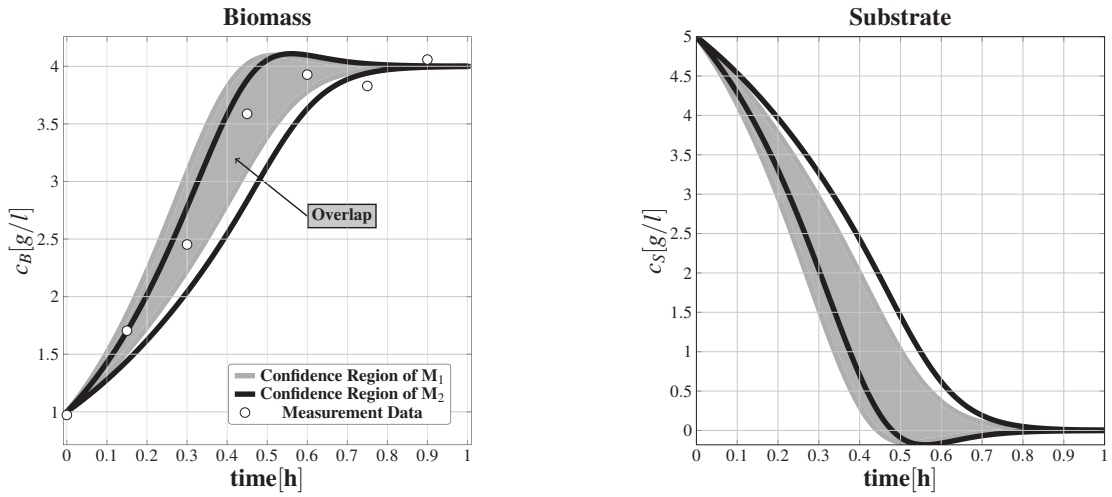

Figure 5: Unsuitable initial conditions for model selection lead to confidence regions of both model candidates M_1 and M_2 that cover nearly the same state space, i.e., large overlap.

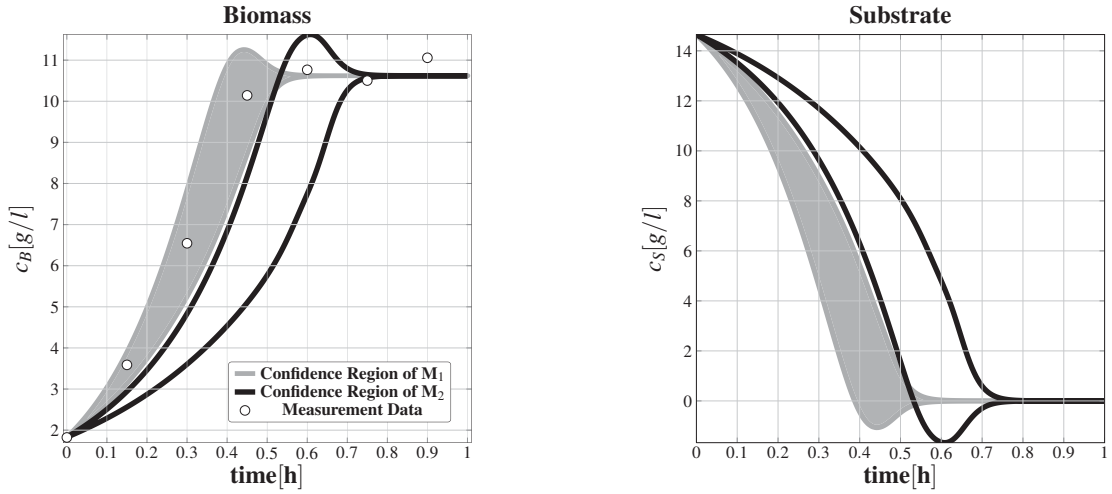

Figure 6: By minimising the overlap, the two model candidates M_1 and M_2 become distinguishable related to measurement data.

4 Conclusions

In this paper, the potential of the Sigma Point method related to the field of Optimal Experimental Design is demonstrated. For the application example of a biological substrate uptake model, the Sigma Points are used to determine more reliable results of parameter uncertainties with a manegable computational effort. That reliable parameter

uncertainties are the precondition to find an optimal system stimuli is shown by the modified E^*-criterion. Only the parameter statistics, that are determined by the Sigma Point method, are able to influence the cost function for parameter identification optimally.
Furthermore, the general concept of this approach enables the determination of some statistics of the system states, that are caused by parameter uncertainties. The benefit of this additional information is at least twofold: (1) a definition of a cost function for OED with an inherent parameter sensitivity measure is possible and (2) the overlap approach can be used to select between competing model hypotheses. Consequently, elementary steps in the framework of OED can be improved by the Sigma Points.

5 Acknowledgements

We gratefully acknowledge the financial support by Forschungszentrum für Dynamische Systeme in Biomedizin und Prozesstechnik, Sachsen-Anhalt.

6 References

[1] Kay SM. Fundamentals of Statistical Signal Processing: Estimation Theory. Prentice Hall PTR; 1993.
[2] Joshi M, Seidel-Morgenstern A, Kremling A. Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical sytems. Metabolic Engineering. 2006;8:447–455.
[3] Wolkenhauer O, Kolch W, Cho KH, Shin SY. Experimental Design in Systems Biology, Based on Parameter Sensititivity Analysis Using a Monte Carlo Method: A Case Study for the TNFα -Mediated NF-κ B Signal Transduction Pathway. SIMULATION. 2003;79:726–739.
[4] Julier S, Uhlmann J. A general method for approximating nonlinear transformation of probability distributions. 1996;Available from: `http://www.robots.ox.ac.uk/~siju/work/work.html`.
[5] Heine T, Kawohl M, King R. Derivative-free optimal experimental design. Chemical Engineering Science. 2008;63:4873–4880.
[6] Schenkendorf R, Kremling A, Mangold M. Optimal Experimental Design with the Sigma Point method. IET Systems Biology. in press;.
[7] Walter E, Pronzato L. Identification of Parametric Models. Springer; 1997.
[8] Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP. Universally Sloppy Parameter Senstitivies in Systems Biology Models. PLoS Computational Biology. 2007;3:1871–1878.
[9] Casey FP, Baird D, Feng Q, Gutenkunst RN, Waterfall JJ, Myers CR, et al. Optimal experimental design in an epidermal growth factor receptor signalling and down-regulation model. IET Systems Biology. 2007;1:190–202.
[10] Burnham KP, Anderson DR. Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach. Springer; 2002.
[11] Hoeting JA, Madigan D, Raftery AE, Volinsky CT. Bayesian model averaging: a tutorial (with discussion). Statistical Science. 1999;14:382–417.
[12] Akaike H. Information measures and model selection. International Statistical Institute. 1983;44:277–291.
[13] Lorenz S, Diederichs E, Telgmann R, Schütte C. Discrimination of Dynamical System Models for Biological and Chemical Processes. J Comput Chem. 2007;28:1384–1399.
[14] Schenkendorf R, Kremling A, Mangold M. Application of Sigma Points to the Optimal Experimental Design of a Biological System. In: The 20th International Symposium on Chemical Reaction Engineering- Green Chemical Reaction Engineering for a Sustainable Future; 2008. .