Proceedings of the 9th International Symposium on
Dynamics and Control of Process Systems (DYCOPS 2010),
Leuven, Belgium, July 5-7, 2010
Mayuresh Kothare, Moses Tade, Alain Vande Wouwer, Ilse Smets (Eds.)

Density-based modeling and identification of biochemical networks in cell populations

J. Hasenauer [*,1], **S. Waldherr** [*], **M. Doszczak** [**], **P. Scheurich** [**], and **F. Allgöwer** [*]

[*] *Institute for Systems Theory and Automatic Control, University of Stuttgart, Pfaffenwaldring 9, Stuttgart, Germany*
[**] *Institute for Cell Biology and Immunology, University of Stuttgart, Allmandring 31, Stuttgart, Germany*

Abstract: In many biological processes heterogeneity within cell populations is an important issue. In this work we consider populations where the behavior of every single cell can be described by a system of ordinary differential equations. Heterogeneity among individual cells is accounted for by differences in parameter values and initial conditions. Hereby, parameter values and initial conditions are subject to a distribution function which is part of the model specification. Based on the single cell model and the considered parameter distribution, a partial differential equation model describing the distribution of cells in the state and in the output space is derived.
For the estimation of the parameter distribution within the model, we consider experimental data as obtained from flow cytometric analysis. From these noise-corrupted data a density-based statistical data model is derived. Employing this model of the data the parameter distribution within the cell population is computed using convex optimization techniques.
To evaluate the proposed method, a model for the caspase activation cascade is considered. It is shown that for known noise properties the unknown parameter distributions in this model are well estimated by the proposed method.

Keywords: parameter estimation, population models, density estimation, apoptotic signaling

1. INTRODUCTION

Most of the modeling performed in the area of systems biology aims at achieving a quantitative description of intracellular pathways. Hence, most available models describe a "typical cell" on the basis of experimental data. Unfortunately, experimental data are in general obtained using cell population experiments, e.g. western blotting. If the considered population is highly heterogeneous, meaning that there is a large cell-cell variability, fitting a single cell model to cell population data can lead to biologically meaningless results. To understand the dynamical behavior of heterogeneous cell populations it is crucial to develop integrated cell population models.

Modeling on the population scale has already been addressed by Mantzaris (2007) and Munsky et al. (2009). These authors demonstrated that populations can show a bimodal response if stochasticity in biochemical reactions is considered. But besides stochasticity in biochemical reactions there are other reasons which can also lead to heterogeneity in populations. Examples are unequal partitioning of cellular material at cell division (Mantzaris, 2007), genetic and epigenetic differences (Avery, 2006).

For the purpose of this paper, we describe heterogeneity in populations by differences in parameter values of the model describing the single cell dynamics. The network structure is assumed to be identical in all cells, as it usually represents the physical interactions among molecules, which should be independent of the cell's state. This parametric approach is well suited for genetic and epigenetic differences. The distribution of parameter values within the cell population of interest is described by a multivariate probability density function, which is part of the model specification.

In the following the problem of estimating the parameter distribution function is studied. Therefore, we consider high-throughput experimental methods such as flow cytometry, which can be used to measure concentration distributions within cell populations by suitable fluorescent labeled antibodies. Classical flow cytometry devices can measure several thousand cells per second.

To estimate the parameter distributions, in a first step, an appropriate population model has to be found. In the literature mathematical models of cell populations are either described as cell ensembles (Waldherr et al., 2009; Munsky et al., 2009), or as a non-linear partial differential equation (PDE) for the distribution of the state variables (Mantzaris, 2007; Luzyanina et al., 2009; Tsuchiya et al., 1966). In case of ensemble models, a differential equation is assigned to each cell, making an in depth theoretical analysis difficult. PDE models, which describe the time evolution of the distributions of the state variables based on the single cell models, are easy to handle from a theoretical point of view but hard to simulate for a large state dimension of the single cell model. Therefore, only low dimensional PDE models of populations have been studied in literature so far (Mantzaris, 2007; Luzyanina et al., 2007, 2009).

In this paper a PDE model for the state distribution within a heterogeneous cell population is derived. Given the solution of this PDE the probability density of measuring a cer-

[1] Corresponding author (hasenauer@ist.uni-stuttgart.de)

tain output can be determined. As for the estimation only the measured outputs are required, a numerical method for computing the output distribution is outlined. This method employs a particle-based approach (Rawlings and Bakshi, 2006) and classical density estimation (Silverman, 1986).

Based on this efficient computation scheme for the population response an estimation method for the underlying heterogeneity is developed. A statistical model of the measured output distribution is derived from the single cell measurements obtained at every measurement instance. Therefore, again kernel density estimators are used as they have better asymptotic properties than commonly used naive estimators (Luzyanina et al., 2009). Given a model and the output distribution estimated from the measurement, l_2-norm minimization is performed over the set of possible parameter distributions. By employing the model properties and a parameterization of the parameter distribution this optimization problem is convex and can be solved efficiently.

The paper is structured as follows. In Section 2, the problem of estimating the parameter distribution is introduced. In Section 3, we present the statistical model for the measured data and the simulation model for state and output distribution. Section 4 gives a short overview of the employed identification procedure before in Section 5 the proposed methods are applied to a caspase activation model with artificial data.

Notation: Consider the m-dimensional hypersurface $\mathcal{S}$, with $\mathcal{S} \subset \mathbb{R}^d$. The integral I of a function $g(z)$, with $g : \mathbb{R}^d \to \mathbb{R}$, over $z \in \mathcal{S}$ is written as

$$I = \int_{\mathcal{S}} g(z) dS.$$

Furthermore, the i.th unit vector is denoted by e_i.

2. PROBLEM STATEMENT

For the purpose of this work, a model of a biochemical reaction network in a population of M cells is given by a collection of differential equations,

$$\begin{aligned} \dot{x}^{(i)} &= f(x^{(i)}, p^{(i)}), \quad x^{(i)}(0) = x_0^{(i)}, \\ y^{(i)} &= h(x^{(i)}, p^{(i)}), \quad i \in \{1, \dots, M\}, \end{aligned} \tag{1}$$

with state variables $x^{(i)}(t, p^{(i)}) \in \mathbb{R}_+^n$, measured variables $y^{(i)}(t, p^{(i)}) \in \mathbb{R}_+^m$, and parameters $p^{(i)} \in \mathbb{R}_+^q$. The index i specifies the individual cells within the population. The parameters $p^{(i)}$ can be kinetic constants, e.g. reaction rates or binding affinities. The effect of cell-cell interaction on the considered pathway is assumed to be negligible, which is the case in many *in vitro* lab experiments where the response of the individual cells is predominantly influenced by external stimuli. The vector field $f : \mathbb{R}_+^n \times \mathbb{R}_+^q \to \mathbb{R}^n$ describing the single cell dynamics is locally Lipschitz and the function $h : \mathbb{R}_+^n \times \mathbb{R}_+^q \to \mathbb{R}_+^m$ is continuous.

In the following heterogeneity within the cell population is introduced, modeled by differential parameter values and initial conditions among individual cells. The distribution of parameters $p^{(i)}$ and initial conditions $x_0^{(i)}$ is given by a probability density function $\Phi : \mathbb{R}_+^{n+q} \to \mathbb{R}_+$ with $\int_{\mathbb{R}_+^{n+q}} \Phi(x_0, p) dx_0 dp = 1$. For ease of notation, we write $\xi_0^{\mathrm{T}} = [x_0^{\mathrm{T}}, p^{\mathrm{T}}]$. The probability density function Φ is part of the model specification and the parameters and initial conditions of cell i are subject to the probability distribution

$$\begin{aligned} \Pr(\xi_{0,1}^{(i)} \le \xi_1, \cdots, \xi_{0,n+q}^{(i)} \le \xi_{n+q}) = \\ \int_0^{\xi_1} \cdots \int_0^{\xi_{n+q}} \Phi(\tilde{\xi}) d\tilde{\xi}_1 \cdots d\tilde{\xi}_{n+q}. \end{aligned} \tag{2}$$

As outlined in Section 1, for the study of cell populations high-throughput cell population measurements are available. Using these experimental techniques protein concentrations within thousands of cells can be measured at every measurement instance, t_k, $k = 1, \dots, N$. This yields the measurement data

$$\mathcal{D}_k = \left\{ \left(t_k, \psi^{(i)}(t_k) \right) \right\}_{i \in \mathcal{I}_k}, \quad k = 1, \dots, N \tag{3}$$

where $\psi^{(i)}$ is the measured output of the cell i and $\mathcal{I}_k$ is the index set of the cells measured at time t_k. Note that in general it is hard to measure single-cell time series data: cells may move between measurement instances or are removed from the population in order to obtain the measurements, and the photobleaching effect limits the time-span that can be observed. On the other hand, if classical flow cytometric analysis is applied the sampled cells can be assumed to be independent and identically distributed and $\mathrm{card}(\mathcal{I}_k)$ is large. Hence, an approximation of the output distribution is possible.

Like most measurement devices, also high-throughput fluorescence measurements are subject to noise. For the rest of the paper, noise consisting of a relative and an absolute part is considered,

$$\psi^{(i)}(t_k) = \mathrm{diag}(\eta^1) y^{(i)}(t_k) + \eta^2, \tag{4}$$

in which $\psi^{(i)}$ is the measured output and $\eta^j \in \mathbb{R}^m$ is a vector of log-normally distributed random variables with probability density functions

$$\Theta_j^i(\eta_j^i) = \begin{cases} \dfrac{\exp\left\{ -\frac{1}{2} \left(\frac{\log \eta_j^i - \mu_j^i}{\sigma_j^i} \right)^2 \right\}}{\sqrt{2\pi} \sigma_j^i \eta_j^i} & \text{for } \eta_j^i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$i = 1, 2$, $j = 1, \dots, m$, yielding the joint probability density

$$\Theta^i(\eta^i) = \prod_{j=1}^{m} \Theta_j^i(\eta_j^i). \tag{6}$$

Log-normally distributed random variables are chosen here, since they are a good model for the commonly seen noise distributions of the considered measurement device and conserve the positivity of all variables. For notational simplicity the measurement errors of the different concentrations are assumed to be uncorrelated. This constraint can be removed easily.

Given this setup the problem we are concerned with is:

Problem 1. Given the measurement data $\mathcal{D}_k$, $k = 1, \dots, N$, the cell population model (1), and the noise model (6), determine the parameter distribution $\Phi(\xi)$.

Unfortunately, estimation of $\Phi(\xi)$ using a cell population model with a finite number of cells and discrete sampled data is fairly difficult as no single cell trajectories are available. A far more natural approach would be to use a density description, as the available measurement data can be interpreted as samples drawn from the probability density function of the output. This interpretation is also quite appealing from a modeling point of view as the number of cells considered in a standard lab experiment

is of the order of 10^9 and hence nevertheless too large to be simulated on an individual basis. In the next chapter a PDE model for the probability density of the output and a density model for the measurement data is derived.

3. DENSITY-BASED MODELING OF HETEROGENEOUS CELL POPULATIONS

As outlined in the previous section, a continuous statistical model for the measurement data, as well as for the evolution of the state and output density would be preferable. These two aspects are addressed in the following.

3.1 Density model of measurement data

The data $\mathcal{D}_k$ collected by the considered measurement devices are samples drawn from the distribution of the measured output, as mentioned in Section 2. Let $\Psi(\psi, t_k)$ be the distribution of the measured outputs $\psi^{(i)}(t_k)$ at time t_k. As $\Psi(\psi, t_k)$ is considered to be a probability density, classical density estimation methods can be employed for estimating $\Psi(\psi, t_k)$ from the given samples $\mathcal{D}_k$.

In this work, the problem of determining $\Psi(\psi, t_k)$ from $\mathcal{D}_k$ is approached using kernel density estimators. Kernel density estimators are non-parametric approaches to estimate probability distributions from sampled data (Silverman, 1986). They are widely used and can be thought of as placing probability "bumps" at each observation, as depicted in Figure 1. These "bumps" are the kernel function K, with $\int_{\mathbb{R}^m} K(\psi)d\psi = 1$. Note that here only the equations for the one dimensional case are given. The extension towards higher dimensions is straightforward and can be found in Silverman (1986). In this work, a Gaussian kernel given by

$$K\left(\psi - \psi^{(i)}, h\right) = \frac{1}{\sqrt{2\pi}h} \exp\left\{-\frac{1}{2}\left(\frac{\psi - \psi^{(i)}}{h}\right)^2\right\}, \quad (7)$$

with standard deviation h is used. In this context, h is also called smoothing parameter in the literature (Silverman, 1986).

Given the kernel K an estimator of the probability density for a given set of samples $\mathcal{D}_k$ is

$$\Psi(\psi, t_k) = \frac{1}{M_k} \sum_{i \in \mathcal{I}_k} K\left(\psi - \psi^{(i)}(t_k), h\right), \quad (8)$$

where M_k is the cardinality of $\mathcal{I}_k$. The selection of the smoothing parameter h is crucial and depends strongly on M_k. In this work h is chosen according to the least-squares cross-validation method (Stone, 1984). As M_k is considered to be large it can be assumed that the error of the estimated output distribution with respect to the actual output distribution is small.

3.2 PDE model of density evolution

As outlined previously, a continuous model for the output density is desirable for the purpose of parameter identification. Therefore, a PDE model for the cell population is derived in the next step.

At first the single cell model is transformed in an extended state space model

$$\begin{aligned} \dot{\xi}^{(i)} &= \begin{pmatrix} f(\xi^{(1,i)}, \xi^{(2,i)}) \\ 0 \end{pmatrix}, \quad \xi^{(i)}(0) = \begin{pmatrix} x_0^{(i)} \\ p^{(i)} \end{pmatrix}, \\ y^{(i)} &= h(\xi^{(1,i)}, \xi^{(2,i)}) \end{aligned} \quad (9)$$

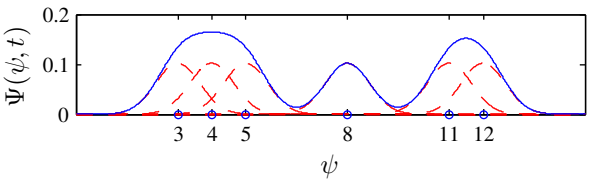

Fig. 1. Gaussian kernel density estimate (—) of $\Psi(\psi, t)$ for the measured outputs (o) and the associated Gaussian kernels (– –).

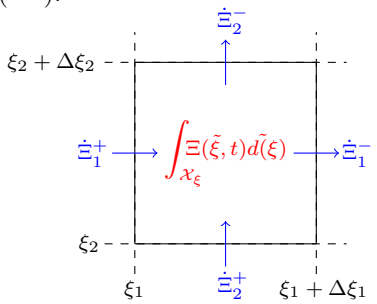

Fig. 2. Infinitesimal volume element $\mathcal{X}_\xi$ of the extended state space, with fluxes across the boundaries.

in which the parameters are appended to the state vector, $\xi^{(i)} = [\xi^{(1,i)}, \xi^{(2,i)}]^{\mathrm{T}} \in \mathbb{R}^{n+q}$ with $\xi^{(1,i)} = x^{(i)}$ and $\xi^{(2,i)} = p^{(i)}$. This system can also be written as

$$\begin{aligned} \dot{\xi}^{(i)} &= F(\xi^{(i)}), \quad \xi^{(i)}(0) = \xi_0^{(i)}, \\ y^{(i)} &= H(\xi^{(i)}), \end{aligned} \quad (10)$$

to which we refer as the extended state space representation.

Based on (10), the PDE model for the population is derived. The state variable of this PDE is the state distribution function $\Xi : \mathbb{R}^{n+q} \times \mathbb{R} \to \mathbb{R}_+ : (\xi, t) \mapsto \Xi(\xi, t)$, which is defined on the extended state space. Based on the distribution function Ξ, the probability of picking at random a cell from the population with states $\xi^{(i)}(t) \in \mathcal{X}$ at time t is given by

$$\Pr(\xi^{(i)}(t) \in \mathcal{X}_\xi) = \int_{\mathcal{X}_\xi} \Xi(\tilde{\xi}, t)d\tilde{\xi}. \quad (11)$$

To determine the PDE for Ξ, an infinitesimal volume $\mathcal{X}_\xi = \mathcal{X}_{\xi,1} \times \ldots \times \mathcal{X}_{\xi,n+q}$ of the extended state space is considered, with $\mathcal{X}_{\xi,i} = [\xi_i, \xi_i + \Delta\xi_i]$. For the 2-dimensional case this is depicted in Figure 2.

For this infinitesimal volume $\mathcal{X}_\xi$ the flux and storage balance is,

$$\begin{aligned} &\int_{\mathcal{X}_\xi} \Xi(\tilde{\xi}, t + \Delta t)d\tilde{\xi} - \int_{\mathcal{X}_\xi} \Xi(\tilde{\xi}, t)d\tilde{\xi} = \\ &\sum_{i=1}^{n+q} \int_t^{t+\Delta t} \left(\dot{\Xi}_i^+(\xi, \tau) - \dot{\Xi}_i^-(\xi, \tau)\right) d\tau. \end{aligned} \quad (12)$$

The left hand side of the equation represents the storage term and the right hand side the fluxes across the boundaries. The fluxes $\dot{\Xi}_i^+$ and $\dot{\Xi}_i^-$ are given by the surface integral of the product of the distribution on the boundary and the influx velocity, determined by the single cell dynamics,

$$\begin{aligned} \dot{\Xi}_i^+(\xi, t) &= \int_{\mathcal{S}_{\dot{\Xi}}(\xi, i)} F_i(\tilde{\xi})\Xi(\tilde{\xi}, t)dS, \\ \dot{\Xi}_i^-(\xi, t) &= \int_{\mathcal{S}_{\dot{\Xi}}(\xi + e_i\Delta\xi_i, i)} F_i(\tilde{\xi})\Xi(\tilde{\xi}, t)dS, \end{aligned} \quad (13)$$

in which $\mathcal{S}_{\tilde{\Xi}}(\xi,i)=\{\tilde{\xi}|\tilde{\xi}_i=\xi_i\wedge\tilde{\xi}_j\in\mathcal{X}_{\xi,j}\forall j\neq i\}$.

Next, (12) and (13) are used to derive the PDE for the time evolution of $\Xi(\xi,t)$. Therefore, at first the storage term is expanded using its Taylor series, yielding

$$\begin{aligned}\int_{\mathcal{X}_\xi}\Xi(\tilde{\xi},t+\Delta t)d\tilde{\xi}-\int_{\mathcal{X}_\xi}\Xi(\tilde{\xi},t)d\tilde{\xi}=\\(\Xi(\xi,t+\Delta t)-\Xi(\xi,t))\prod_{j=1}^{n+q}\Delta\xi_j+\mathcal{O}(\Delta\xi^{n+q+1}).\end{aligned}\tag{14}$$

Here we assume that $\mathcal{O}(\Delta\xi_j)=\mathcal{O}(\Delta\xi)\forall j\in\{1,\dots,n+q\}$. In a second step the flux difference $\Delta\dot{\Xi}_i(\xi,\tau)=\dot{\Xi}_i^+(\xi,\tau)-\dot{\Xi}_i^-(\xi,\tau)$ is rewritten,

$$\Delta\dot{\Xi}_i(\xi,t)=-\int_{\mathcal{S}_{\tilde{\Xi}}(\xi,i)}\Big(\frac{\partial(F_i\Xi)}{\partial\xi_i}\Big|_{(\xi,t)}\Delta\xi_i+\mathcal{O}(\Delta\xi_i^2)\Big)dS\tag{15}$$

$$=-\frac{\partial(F_i\Xi)}{\partial\xi_i}\Big|_{(\xi,t)}\prod_{j=1}^{n+q}\Delta\xi_j+\mathcal{O}(\Delta\xi^{n+q+1}).\tag{16}$$

The first line follows from the definition of $\Delta\dot{\Xi}_i(\xi_i,t)$ and the Taylor series of $F_i(\xi+e_i\Delta\xi_i)\Xi(\xi+e_i\Delta\xi_i,t)$. To obtain the second line the integration is carried out. The final reformulation is the expansion of the integral over τ in (12), resulting in

$$\begin{aligned}&\int_t^{t+\Delta t}\left(\dot{\Xi}_i^+(\xi,\tau)-\dot{\Xi}_i^-(\xi,\tau)\right)d\tau\\&=-\Delta t\left(\prod_{j=1}^{n+q}\Delta\xi_j\right)\frac{\partial(F_i\Xi)}{\partial\xi_i}\Big|_{(\xi,t)}+\mathcal{O}(\Delta\xi^{n+q})\mathcal{O}(\Delta t^2).\end{aligned}\tag{17}$$

Substituting (14) and (17) in the flux balance (12) and dividing by $\Delta t\prod_{j=1}^{n+q}\Delta\xi_j$ then yields,

$$\frac{\Xi(\xi,t+\Delta t)-\Xi(\xi,t)+\mathcal{O}(\Delta\xi)}{\Delta t}=-\sum_{i=1}^{n+q}\frac{\partial(F_i\Xi)}{\partial\xi_i}\Big|_{\xi}+\mathcal{O}(\Delta t).$$

Given this the PDE governing the evolution of $\Xi(\xi,t)$ is obtained by taking the limits $\Delta\xi_i\to0$ and $\Delta t\to0$, leading to

$$\frac{\partial\Xi}{\partial t}(\xi,t)=-\sum_{i=1}^{n+q}\frac{\partial(F_i\Xi)}{\partial\xi_i}(\xi,t),\tag{18}$$

for sufficiently smooth $\Xi(\xi,t)$. This final equation is what we expected, a transport equation with position dependent transport direction and velocity, according to the single cell dynamics. The initial condition of (18) is the initial distribution on the extended state space,

$$\Xi(\xi,0)=\Phi(\xi),\quad\forall\xi\in\mathbb{R}_+^{n+q}.\tag{19}$$

From the state distribution $\Xi(\xi,t)$, the output distribution $\Upsilon(y,t)$ is computed as the integral of the state distribution along $H(\xi)=y$,

$$\Upsilon(y,t)=\int_{\mathcal{S}_\Upsilon(y)}\Xi(\xi,t)dS,\tag{20}$$

where $\mathcal{S}_\Upsilon(y)=\{\xi|H(\xi)=y\}$.

The resulting partial differential equation system is

$$\begin{aligned}\frac{\partial\Xi}{\partial t}(\xi,t)&=-\sum_{i=1}^{n+q}\frac{\partial(F_i\Xi)}{\partial\xi_i}(\xi,t),\quad\Xi(\xi,0)=\Phi(\xi)\\\Upsilon(y,t)&=\int_{\mathcal{S}_\Upsilon(y)}\Xi(\xi,t)dS,\end{aligned}\tag{21}$$

where $\Xi:\mathbb{R}_+^{n+q}\times\mathbb{R}\to\mathbb{R}_+$ and $\Upsilon:\mathbb{R}_+^m\times\mathbb{R}\to\mathbb{R}_+$. This PDE is of first order, quasilinear and known as Liouville's equation. The solution always exists for sufficiently smooth $F(\cdot)$ (Evans, 1998).

As the measurements are noise corrupted, the distribution of measured outputs $\Psi(\psi,t)$ is different from the actual output distribution $\Upsilon(y,t)$. It is given by

$$\Psi(\psi,t)=\int_{\mathcal{S}_\Psi(\psi)}\Upsilon(y,t)\Theta^1(\eta^1)\Theta^2(\eta^2)dS,\tag{22}$$

where $\mathcal{S}_\Psi(\psi)=\{[y^{\mathrm{T}},(\eta^1)^{\mathrm{T}},(\eta^2)^{\mathrm{T}}]^{\mathrm{T}}|\mathrm{diag}(\eta^1)y+\eta^2=\psi\}$.

3.3 Numerical solution of PDE

In order to study the time evolution of the output distribution $\Upsilon(y,t)$ and the measured output distribution $\Psi(\psi,t)$ (21) has to be solved for given $\Phi(\xi)$. As $\Xi(\xi,t)$ is defined on the $(n+q)$-dimensional space, standard grid based solvers are not able to solve (21) for $n+q>3$. Theoretically, the methods of characteristics can be used (Evans, 1998) but for the high dimensional system we are going to study, also this method is difficult to apply. Instead, a stochastic method is used, which is known from particle filtering (Rawlings and Bakshi, 2006).

This stochastic integration method is based on a particle description of the model, which is in our case equivalent to the cell ensemble model (1). To compute $\Psi(\psi,t)$, at first a set of samples $\{(x_0^{(i)},p^{(i)})\}_{i=1,\dots,S}$, is drawn from $\Phi(\xi)$, where S is the number of samples. For this set of samples the single cell model (8) is simulated, resulting in a set of simulated outputs $\{y^{(i)}(t,p^{(i)})\}_{i=1,\dots,S}$. The output $y^{(i)}(t,p^{(i)})$ is then corrupted by noise according to (4) resulting in $\{\psi^{(i)}(t)\}_{i=1,\dots,S}$. Given this a numerical approximation of $\Psi(\psi,t)$ can be determined using the kernel density estimator described in Section 3.1. This numerical stochastic approximation the output of $\Psi(\psi,t)$ can be shown to converge as $S\to\infty$. Hence, the measured output distribution $\Psi(\psi,t)$ can be approximated also for high dimensional nonlinear systems.

4. ESTIMATION OF PARAMETER DISTRIBUTIONS

As mentioned in Section 2 the problem studied in this work is the estimation of the parameter distribution Φ from the data $\mathcal{D}_k$. This problem is approached in the following by minimizing the l_2-norm of the model-data mismatch,

$$J\left(\hat{\Phi}\right)=\sum_{k=1}^{n+q}\left|\left|\Psi(\psi,t_k)-\hat{\Psi}(\psi,t_k,\hat{\Phi})\right|\right|_2^2.\tag{23}$$

in which $\hat{\Psi}(\psi,t,\hat{\Phi})$ is the distribution of the measured output $\psi(t)$ obtained by simulation with the parameter distribution $\hat{\Phi}(\xi)$. According to the cost J, the optimal parameter distribution $\hat{\Phi}^*(\xi)$ is than given by

$$\begin{aligned}&\hat{\Phi}^*=\arg\min_{\hat{\Phi}}J(\hat{\Phi})\\&\text{subject to }\int_{\mathbb{R}_+^{n+q}}\hat{\Phi}(\xi)d\xi=1,\ \hat{\Phi}(\xi)\geq0\,\forall\xi\in\mathbb{R}_+^{n+q},\end{aligned}\tag{24}$$

where the last two constraints enforce that $\hat{\Phi}(\xi)$ is a probability distribution.

Remark 1. For the remainder of this section the measured output distribution $\Psi(\psi,t)$ is compared to the noise corrupted simulated output distribution $\hat{\Psi}(\psi,t,\hat{\Phi})$. This is

possible as we assume a large number of measured cells per measurement instance and therefore have good statistics on the measurement error.

Unfortunately, the optimization problem (24) is infinite dimensional. Therefore, a parametrization of $\hat{\Phi}$,

$$\hat{\Phi}_\varphi(\xi) = \sum_{i=1}^{n_\varphi} \varphi_i \Lambda^i(\xi), \tag{25}$$

with a weighting vector $\varphi \in \mathbb{R}^{n_\varphi}$ is introduced. In this work the ansatz functions Λ^i for $\hat{\Phi}$ are chosen to be head functions, as depicted in Figure 3. This yields the simplified, finite-dimensional optimization problem,

$$\begin{aligned} \varphi^* &= \arg\min_{\varphi} J(\hat{\Phi}_\varphi) \\ &\text{subject to } \ c^{\mathrm{T}}\varphi = 1,\ \varphi \geq 0, \end{aligned} \tag{26}$$

in which $c_i = \int_{\mathbb{R}_+^{n+q}} \Lambda^i(\xi) d\xi$. The two constraints are again needed to ensure that $\hat{\Phi}_\varphi(\xi)$ is a probability density.

In order to solve (26) using computational techniques the quasi-linearity of (21) is employed. As the superposition principle holds, the output $\hat{\Psi}(\psi, t, \hat{\Phi}_\varphi)$ can be written as the weighted sum

$$\hat{\Psi}(\psi, t, \hat{\Phi}_\varphi) = \sum_{i=1}^{n_\varphi} \varphi_i \hat{\Psi}(\psi, t, \Lambda^i), \tag{27}$$

where $\hat{\Psi}(\psi, t, \Lambda^i)$ is the output distribution obtained for simulation with a parameter distribution according to $\Lambda^i(\xi)$. This allows the reformulation of the objective function to

$$J\left(\hat{\Phi}_\varphi\right) = \sum_{k=1}^{n+q} \left|\left| \Psi(\psi, t_k) - \sum_{i=1}^{n_\varphi} \varphi_i \hat{\Psi}(\psi, t_k, \Lambda^i) \right|\right|_2^2.$$

Employing this the optimization problem (26) can finally be written as

$$\begin{aligned} \varphi^* &= \arg\min_{\varphi} \sum_{k=1}^{n+q} \left(A_k\varphi - b_k\right)^{\mathrm{T}} W \left(A_k\varphi - b_k\right) \\ &\text{subject to } \ c^{\mathrm{T}}\varphi = 1,\ \varphi \geq 0, \end{aligned} \tag{28}$$

where the integral $||\cdot||_2^2$ has been approximated, e.g. using the trapezoidal rule. The column vector b_k contains hereby the values $\Psi(\psi, t_k)$ at the grid points of the discretization. Equivalently, the ith column of A_k contains the values of $\hat{\Psi}(\psi, t_k, \Lambda^i)$ at the grid points. The matrix W is a constant weighting matrix, determined by the chosen approximation of $||\cdot||_2^2$.

Note that problem (28) is convex. Hence, even in case of high dimensional φ, convergence to an optimal parameter distribution within the considered class of distributions can be guaranteed.

5. APPLICATION TO THE CASPASE CASCADE

Programmed cell death, also called apoptosis, is an important physiological process to remove infected, malfunctioning, or no longer needed cells from a multicellular organism. Pathways to induce apoptosis converge at the caspase activation cascade (Hengartner, 2000). A mathematical model for this network has been proposed by Eissing et al. (2004). Here, we consider the caspase activation in response to an external death receptor stimulus, e.g. the tumor necrosis factor (TNF). As seen from

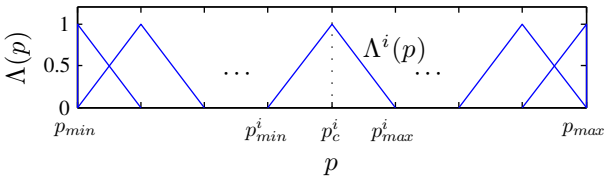

Fig. 3. Illustration of head functions $\Lambda^i(p)$.

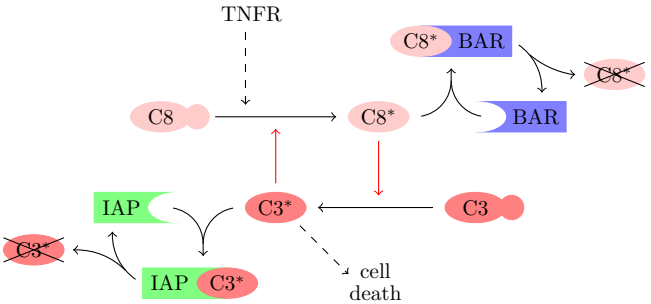

Fig. 4. Schematic of the caspase activation cascade.

experimental cytotoxicity assays, the cellular response to a TNF stimulus is highly heterogeneous, with some cells dying and others surviving. To understand the process at the physiological level it is thus crucial to consider the cellular heterogeneity, using for example cell population modeling.

The reactions in the single cell model are given by

$$\begin{array}{rcl|rcl}
\text{C3}+\text{C8}^* & \xrightarrow{k_1} & \text{C3}^*+\text{C8}^* & \text{IAP} & \underset{k_{-8}}{\overset{k_8}{\leftrightarrows}} & \emptyset \\
\text{C3}^*+\text{C8} & \xrightarrow{k_2} & \text{C3}^*+\text{C8}^* & \text{C8} & \underset{k_{-9}}{\overset{k_9}{\leftrightarrows}} & \emptyset \\
\text{C3}^*+\text{IAP} & \underset{k_{-3}}{\overset{k_3}{\leftrightarrows}} & \text{C3}^*\sim\text{IAP} & \text{C3} & \underset{k_{-10}}{\overset{k_{10}}{\leftrightarrows}} & \emptyset \\
\text{C3}^*+\text{IAP} & \xrightarrow{k_4} & \text{C3}^* & \text{C8}^*+\text{BAR} & \underset{k_{-11}}{\overset{k_{11}}{\leftrightarrows}} & \text{C8}^*\sim\text{BAR} \\
\text{C8}^* & \xrightarrow{k_5} & \emptyset & \text{BAR} & \underset{k_{-12}}{\overset{k_{12}}{\leftrightarrows}} & \emptyset \\
\text{C3}^* & \xrightarrow{k_6} & \emptyset & \text{C8}^*\sim\text{BAR} & \xrightarrow{k_{13}} & \emptyset \\
\text{C3}^*\sim\text{IAP} & \xrightarrow{k_7} & \emptyset & \text{TNFR}+\text{C8} & \xrightarrow{k_{14}} & \text{TNFR}+\text{C8}^*
\end{array}$$

For nominal parameter values, we refer to the original publication (Eissing et al., 2004). In comparison to the original model, we added reaction v_{14} for the initiator caspase 8 (C8) activation by the TNF receptor complexes (TNFR). The reaction rate for this activation is given by $v_{14} = k_{14}[\text{TNFR}][\text{C8}]$, with the parameter value $k_{14} = 10^{-6}(\text{molecules}\,\text{min})^{-1}$. A sketch of the single cell model is given in Figure 4.

Heterogeneity is modeled by a log-normally distributed production rate of the inhibitor of apoptosis IAP, k_8, and a log-normally distributed amount of TNF-receptor complexes on the cell membrane, TNFR. These two quantities were chosen as it is known from experiments that there is a high cell-to-cell variability. Especially the concentration of IAPs contained in a cell is highly variable, and a variation in IAP production is known to affect cell death considerably (Eissing et al., 2006). In the following the possibility of estimating the distributions of $\Phi(k_8)$ and $\Phi([\text{TNFR}])$ from the distributions of $[\text{C3}^*]$, $\Psi([\text{C3}^*], t)$, is studied. The statistical model of the distribution, $\Psi([\text{C3}^*], t)$ is shown in Figure 5. This statistical model has been derived using artificial measurement data of 10^4 cells at the

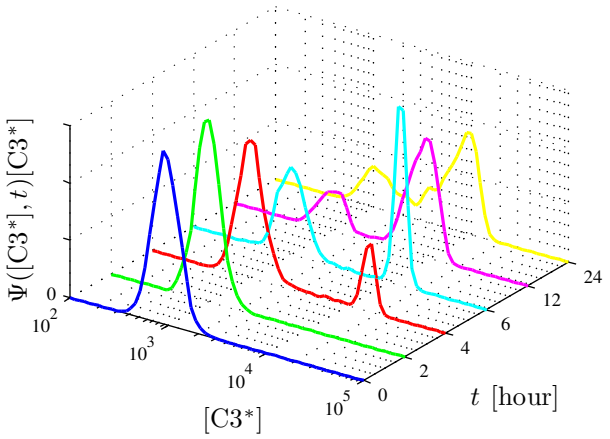

Fig. 5. Artificial noise corrupted measurement data for amount of active caspase 3, [C3*].

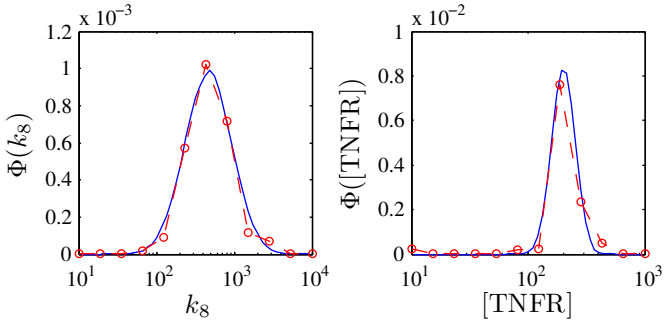

Fig. 6. Real (—) vs. estimated (-o-) parameter distribution, with grid points (o).

measurement instances t_k, $k = 1, \ldots, 6$. This is a realistic number for standard cytofluorometric experiments. The noise properties are assumed to be known and have been set to $\mu_1 = 0$, $\sigma_1 = 0.1$, $\mu_2 = \log(10^3)$, and $\sigma_2 = 0.3$, corresponding to an average measurement error of more than 20 percent.

Based on these data, the approach presented in Section 4 is used to obtain an estimate for the parameter distribution. For this purpose the considered parameter set is divided using a 12×12 grid, with logarithmically distributed grid points. The grid points are used as edge and center points of the ansatz functions $\Lambda^i(k_8, [\text{TNFR}])$ for $\hat{\Phi}(k_8, [\text{TNFR}])$. The obtained estimation result is depicted in Figure 6.

It is obvious that the estimated parameter distribution approximates the real parameter distribution very well, especially considering the finite number of degrees of freedom. Hence, even though there is an average measurement error of 20 % on the single cell measurement, due to good statistics at the population level, the actual parameter distributions can be estimated accurately. Furthermore, this study shows that in principle, measuring one concentration can give enough information to estimate several parameter distributions, if the output distribution is sensitive with respect to these parameters.

6. SUMMARY AND CONCLUSION

Heterogeneity in cell populations is an important issue for research in systems biology. However, so far only few models describing heterogeneous populations of cells with more than one state variable have been developed. In this paper a partial differential equation model describing the time evolution of the state distribution is derived. We focused hereby in particular on the distribution of the measured outputs.

In the second part of the paper, the model of the noise corrupted measured outputs and its particular properties are used to estimate the parameter distributions underlying the heterogeneity. Therefore, a density-based statistical model of the sampled single cell is developed and applied in combination with l_2-norm based convex optimization.

Finally, we applied the developed estimation method to artificial data of a medium size bistable system modeling the caspase activation cascade. It could be shown that the proposed method yields good estimation results in case of a setup which is realistic in terms of noise and amount of available data.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the German Federal Ministry of Education and Research (BMBF) within the FORSYS-Partner program (grant nr. 0315-280A), from the German Research Foundation within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart, and from Center Systems Biology (CSB) at the University Stuttgart.

REFERENCES

S.V. Avery. Microbial cell individuality and the underlying sources of heterogeneity. *Nat. Rev. Microbiol.*, 4:577–587, 2006.

T. Eissing, H. Conzelmann, E.D. Gilles, F. Allgöwer, E. Bullinger, and P. Scheurich. Bistability analyses of a caspase activation model for receptor-induced apoptosis. *J. of Biol. Chem.*, 279 (35): 36892–36897, 2004.

T. Eissing, S. Waldherr, E. Bullinger, C. Gondro, O. Sawodny, F. Allgöwer, P. Scheurich, and T. Sauter. Sensitivity analysis of programmed cell death and implications for crosstalk phenomena during Tumor Necrosis Factor stimulation. In *Proc. IEEE Conf. Contr. Appl. (CCA)*, pages 1746–52, 2006.

L.C. Evans. *Partial Differential Equations*. American Mathematical Society, 1998.

M.O. Hengartner. The biochemistry of apoptosis. *Nature*, 407(6805): 770–776, 2000.

T. Luzyanina, D. Roose, T. Schenkel, M. Sester, S. Ehl, A. Meyerhans, and G. Bocharov. Numerical modelling of label-structured cell population growth using CFSE distribution data. *Theo. Biol. and Med. Mod.*, 4(26):1–14, 2007.

T. Luzyanina, D. Roose, and G. Bocharov. Distributed parameter identification for label-structured cell population dynamics model using CFSE histogram time-series data. *J. Math. Biol.*, 59:581–603, 2009.

N.V. Mantzaris. From single-cell genetic architecture to cell population dynamics: Quantitatively decomposing the effects of different population heterogeneity sources for a genetic network with positive feedback architecture. *Biophys. J.*, 92:4271–4288, 2007.

B. Munsky, B. Trinh, and M. Khammash. Listening to the noise: random fluctuations reveal gene network parameters. *Mol. Syst. Biol.*, 5, 2009.

J.B. Rawlings and B.R. Bakshi. Particle filtering and moving horizon estimation. *Comp. and Chem. Eng.*, 30:1529–1541, 2006.

B.W. Silverman. *Density Estimation for Statistics and Data Analysis*. Monographs on Statistics and Applied Probability. London: Chapman and Hall, 1986.

C.J. Stone. An asymptotically optimal window selection rule for kernel density estimation. *Annual Statistics*, 12:1285–1297, 1984.

H.M. Tsuchiya, A.G. Fredrickson, and R. Aris. Dynamics of microbial cell populations. *Adv. Chem. Eng.*, 6:125–206, 1966.

S. Waldherr, J. Hasenauer, and F. Allgöwer. Estimation of biochemical network parameter distributions in cell populations. In *Proc. of the 15th IFAC Symp. on Syst. Ident.*, pages 1265–1270, 2009.