Analysis of heterogeneous cell populations: A density-based modeling and identification framework

Jan Hasenauer[a], Steffen Waldherr[a], Malgorzata Doszczak[b], Peter Scheurich[b], Nicole Radde[a], Frank Allgöwer[a]

[a] *Institute for Systems Theory and Automatic Control, University of Stuttgart, Stuttgart, Germany*
[b] *Institute for Cell Biology and Immunology, University of Stuttgart, Stuttgart, Germany*

Abstract

In many biological processes heterogeneity within clonal cell populations is an important issue. One of the most striking examples is a population of cancer cells in which after a common, identical death signal some cells die whereas others survive. The reason for this heterogeneity is intrinsic and extrinsic noise.

In this paper we present a mechanistic multi-scale modeling framework for cell populations, in which the dynamics of every individual cell is captured by a parameter dependent stochastic differential equation (SDE). Heterogeneity among individual cells is accounted for by differences in parameter values, modeling extrinsic influences. Based on the statistical properties of the extrinsic noise and the SDE model for the individual cell, a partial differential equation (PDE) model is derived. This PDE describes the evolution of the population density. To determine the statistics of the extrinsic noise from experimental population data, a density-based statistical data model of the noise-corrupted data is derived. Employing this data model we show that the statistics of the extrinsic can be computed using a convex optimization. This efficient way of assessing the parameters allows for a so far infeasible uncertainty analysis via bootstrapping.

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 densities in this model are well estimated by the proposed method.

Key words: cell population, parameter estimation, density estimation, convex optimization, bootstrapping, 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 experiments which average over a cell population, e.g. western blotting. If the considered population is highly heterogeneous, meaning that there is a large cell-to-cell variability [1, 2, 3], 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, mechanistic models for heterogeneous cell populations.

The general need for cell population models has been realized several decades ago. The first publications on that topic focused on the mathematical description of proliferating cell populations [4, 5]. The corresponding models are called population balance models (PBMs) or age-structured models, and their dynamics are in general governed by a single one-dimensional partial differential equation (PDE) [4, 5, 6, 7, 8, 9]. Although the PBMs are appealing from a theoretical point of view, the limited number of dimensions which can be handled by classical PDE solver restricted their use. Thus, only extremely simple single-cell models are employed or the single cell dynamics are neglected completely by assuming stationarity.

An alternative model class are the individual-based population models (IBPM). In this modeling framework, the given single cell model is simulated for a large number of cells, each with different parameters, initial conditions, and/or realizations of the intrinsic noise values, all specified in the model description [10, 11, 12, 13, 14]. The IBPMs allow for the study of complex single-cell dynamics but parameter estimation becomes more difficult.

In this work, we present a mechanistic multi-scale modeling framework for cell populations, in which the dynamics of each individual cell is captured by a parameter dependent stochastic differential equation (SDE). Thereby, we considered cell-to-cell variability introduced by intrinsic and extrinsic noise [1, 2, 3]. Intrinsic noise is generated by the stochastic dynamics of each individual cell which are due to stochasticity of the chemical reactions. Extrinsic noise and the other hand is modeled by differences in parameter values and initial conditions among cells, which are both subject to a joint probability density. This leads to a rather general modeling framework and the resulting evolution of the state and output density of the population is governed by a PDE. The number of coordinates of the

resulting PDE equals the sum of the numbers of state variables and parameters of the respective single-cell SDE model. In this work, the resulting high-dimensional PDE is solved by combining particle-based approaches and kernel density estimation [15, 16, 17, 18].

Employing this modeling and simulation framework, we approach the problem of estimating the heterogeneity introduced by extrinsic noise, represented by the multivariate probability density function. Therefore, we consider high-throughput experimental methods such as flow cytometry, which can be used to measure concentration densities within cell populations by suitable fluorescently labeled antibodies. Classical flow cytometry devices can measure several thousand cells per second, thus the amount of data is sufficient to obtain good statistical properties for the measurement and estimation of the population heterogeneity.

Given these measured single cells, a statistical model of the measured output density is derived from the single cell measurements obtained at every measurement instance. Therefore, again kernel density estimators [19] are used as they have better asymptotic properties than commonly used naive estimators [7, 13]. Given a model and the measured output density estimated from the measurement, we perform a l_2-norm minimization over the set of possible parameter densities. To compensate for the measurement noise, the output density predicted by the model is thereby convolved with the density of the measurement noise. Although the resulting optimization problem seems to be extremely complex, due to the model properties and a parameterization of the densities, this optimization problem is convex and can be solved efficiently.

Compared to classical parameter estimation methods for PBMs [6, 7, 8] our approach can deal with hidden variables, like the maximum likelihood approach presented in [16]. In contrast to common estimation approaches for SDEs [20, 21, 22], the problem studied here is more general as intrinsic and extrinsic noise is considered. The main advantage with respect to [16] is the achieved simplification of the optimization problem, which is for our approach a convex, quadratic program. The resulting reduction in computation time allows the assessment of the model uncertainty analysis using parametric bootstrapping [23], which is also described in this paper. To our knowledge the presented approach is the first one allowing for the estimation of the statistical properties of extrinsic noise in the presence of intrinsic noise in a cell population context.

The paper is structured as follows. In Section 2, the problem of estimating the density of parameters and initial conditions is introduced. In Section 3, we present the statistical model for the measured data and the mechanistic multi-scale simulation model for state and output density. The employed estimation and the

uncertainty analysis procedure are introduced in Section 4, before in Section 5 the proposed methods are applied to a caspase activation model with artificial data. The paper is concluded in Section 6.

Notation: Time-dependent continuous random variable are denoted by capital letters, e.g. X_t. The probability density of a continuous random variable X_t is denoted by $p(X, t)$. The dimension of X_t is n_X, $X_t \in \mathbb{R}^{n_X}$.

2. Theoretical background

In this section the problem setup is defined. In particular, we introduce the considered model class, the available measurement data and the problem formulation.

2.1. Individual-based population model

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 stochastic differential equations,

$$\begin{aligned} dX_t^{(i)} &= \mu(X_t^{(i)}, P^{(i)}, t)dt + \sigma(X_t^{(i)}, P^{(i)}, t)dW_t^{(i)}, \\ Y_t^{(i)} &= \gamma(X_t^{(i)}, P^{(i)}, t), \quad i \in \{1, \dots, m\}, \end{aligned} \tag{1}$$

with state variables $X_t^{(i)} \in \mathbb{R}^{n_X}$, initial state $X_0^{(i)} \in \mathbb{R}^{n_X}$, measured variables $Y_t^{(i)} \in \mathbb{R}^{n_Y}$, Wiener processes $W_t^{(i)} \in \mathbb{R}^{n_W}$, and the parameters $P^{(i)} \in \mathbb{R}^{n_P}$. 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 fields μ and σ describing the deterministic and the stochastic evolution of a single cell, respectively, and are locally Lipschitz. The output mapping γ is continuous.

In the following, heterogeneity within the cell population is modeled using intrinsic and extrinsic noise. Intrinsic noise is introduced by the Wiener process, whereas extrinsic noise is modeled by differential parameter values and initial conditions among individual cells. The density of parameters $P^{(i)}$ and initial conditions $X_0^{(i)}$ is given by a probability density $p_0(Z) : \mathbb{R}^{n_Z} \to \mathbb{R}_+$ with $Z_0^{(i)} = [(X_0^{(i)})^{\mathrm{T}}, (P^{(i)})^{\mathrm{T}}]^{\mathrm{T}}$ and $n_Z = n_X + n_P$. The probability density function $p_0(Z)$

is part of the model specification and the parameters and initial conditions of cell i are subject to the probability density

$$\Pr(Z_0^{(i)} \in \Omega) = \int_\Omega p_0(Z) dZ. \tag{2}$$

As the initial conditions are unknown and hence need to be considered as additional parameters, we will refer to $p_0(Z)$ as parameter density.

Note that if the number of cells m is finite, the cell population is simply a collection of m individual cells. Hence, we are in the IBPM framework.

2.2. Measurement data

For the study of heterogeneous cell populations high-throughput cell population experiments are exploited in this paper. Using these experimental techniques protein concentrations within thousands of cells can be measured at every measurement instance, t_k, $k = 1, \ldots, N$. This yields the snapshot data

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

in which $\overline{Y}_{t_k}^{(i)}$ 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 the number of measured cells, $m_k = \mathrm{card}(\mathcal{I}_k)$, is large. Hence, an approximation of the output density is possible.

Like all measurement devices, also high-throughput fluorescence measurements are subject to noise, and the measured output depends on the actual output by the noise model

$$\overline{Y}_{t_k}^{(i)} \sim p(\overline{Y}_{t_k}^{(i)} | Y_{t_k}^{(i)}). \tag{4}$$

In the following, we do not assume any specific noise distribution. It is merely required that the noise distribution is smooth. For experimental setups in biology, the measurement noise often has additive and multiplicative components [24, 25].

2.3. Problem statement

Given the above setup, the problem we are concerned with is:

Problem: Given the measurement data $\mathcal{D}_k$, $k = 1, \ldots, N$, the cell population

model (1), and the noise model (4), determine the parameter density $p_0(Z)$ and its uncertainty.

Unfortunately, estimation of $p_0(Z)$ 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 is to use a density description, as the available measurement data can be interpreted as a sample 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 on the order of $10^4 - 10^7$ and hence nevertheless too large to be simulated on an individual basis. In the next section a PDE model for the probability density of the output and a density model for the measurement data is derived.

3. Density-based population modeling

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 is a sample drawn from the distribution of the measured output, which is denoted by $p(\overline{Y}, t_k)$. As $p(\overline{Y}, t_k)$ is a probability density, classical density estimation methods can be employed for estimating $p(\overline{Y}, t_k)$ from the given sample $\mathcal{D}_k$.

In this work, the problem of determining $p(\overline{Y}, t_k)$ from $\mathcal{D}_k$ is approached using kernel density estimators. Kernel density estimators are non-parametric approaches to estimate probability densities from sampled data [19]. 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 functions $\mathcal{K}$, with $\int_{\mathbb{R}^{n_Y}} \mathcal{K}\left(\overline{Y}, \overline{Y}^{(i)}, h\right) d\overline{Y} = 1$. 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 [19]. In this work, a Gaussian kernel is given by

$$\mathcal{K}\left(\overline{Y}, \overline{Y}_{t_k}^{(i)}, h\right) = \frac{1}{\sqrt{2\pi}h} \exp\left\{-\frac{1}{2}\left(\frac{\overline{Y} - \overline{Y}_{t_k}^{(i)}}{h}\right)^2\right\}, \tag{5}$$

with standard deviation h. In this context, h is also called smoothing parameter in the literature.

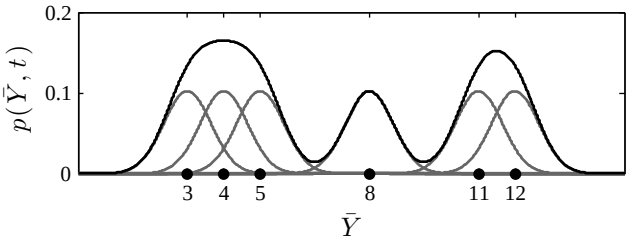

Figure 1: Gaussian kernel density estimate (—) of $p(\bar{Y}, t)$ for the measured outputs $\bar{Y}^{(i)}$ (•) and the associated Gaussian kernels (—).

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

$$p(\bar{Y}, t_k) = \frac{1}{m_k} \sum_{i \in \mathcal{I}_k} \mathcal{K}\left(\bar{Y}, \bar{Y}_{t_k}^{(i)}, h\right), \tag{6}$$

in which m_k is the number of cells measured at time t_k (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 using the diffusion-based methods presented in [26]. This methods outperform classical selection approaches in the case of multi modal densities, which are common in biological processes [10, 14]. As m_k is considered to be large it can be assumed that the error of the estimated output density with respect to the actual output density 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 from the IBPM introduced in Section 2.1. Instead of describing individual cells, we go to the next higher scale and model the evolution of the cell populations directly. This is done without neglecting the dynamics of the individual cells, resulting in a mechanistic population model.

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

$$\begin{aligned} dZ_t^{(i)} &= \begin{pmatrix} \mu(Z_t^{(1,i)}, Z_t^{(2,i)}, t) \\ 0 \end{pmatrix} dt + \begin{pmatrix} \sigma(Z_t^{(1,i)}, Z_t^{(2,i)}, t) \\ 0 \end{pmatrix} dW_t^{(i)} \\ Y_t^{(i)} &= \gamma(Z_t^{(1,i)}, Z_t^{(2,i)}), \end{aligned} \tag{7}$$

in which the parameters are appended to the state vector, $Z_t^{(i)} = [(Z_t^{(1,i)})^{\mathrm{T}}, (Z_t^{(2,i)})^{\mathrm{T}}]^{\mathrm{T}}$ with $Z_t^{(1,i)} = X_t^{(i)}$ and $Z_t^{(2,i)} = P^{(i)}$. This system can also be written as

$$\begin{aligned} dZ_t^{(i)} &= \tilde{\mu}(Z_t^{(i)}, t)dt + \tilde{\sigma}(Z_t^{(i)}, t)dW_t^{(i)} \\ Y_t^{(i)} &= \tilde{\gamma}(Z_t^{(i)}, t), \end{aligned} \tag{8}$$

in which the initial conditions are drawn from the parameter distribution, $Z_0^{(i)} \sim p_0(Z)$.

Based on (8), the PDE model for the population density is derived with state variable $p(Z, t|p_0)$. The density function $p(Z, t|p_0)$ provides the probability of drawing at random a cell from the population with states $Z_t^{(i)} \in \Omega$ at time t,

$$\Pr(Z_t^{(i)} \in \Omega) = \int_{\Omega} p(Z, t|p_0)dZ. \tag{9}$$

The PDE model of the time evolution of $p(Z, t|p_0)$ can be derived directly, as the state vector of the augmented SDE (8) contains all information about the individual cell i. Thus, we are in the classical SDE setting. The evolution of the density of the augmented state vector is described by the Fokker-Planck equation [27, 28],

$$\begin{aligned} \frac{\partial}{\partial t}p(Z, t|p_0) = &- \sum_{i=1}^{n_Z} \frac{\partial}{\partial Z_i}\left[\tilde{\mu}_i(Z, t)p(Z, t|p_0)\right] \\ &+ \frac{1}{2}\sum_{i=1}^{n_Z}\sum_{j=1}^{n_Z} \frac{\partial^2}{\partial Z_i \partial Z_j}\left[\tilde{\sigma}_i(Z, t)\tilde{\sigma}_j(Z, t)p(Z, t|p_0)\right]. \end{aligned} \tag{10}$$

with initial condition

$$\forall Z \in \mathbb{R}^{n_Z} : \; p(Z, 0|p_0) = p_0(Z). \tag{11}$$

The Fokker-Planck equation (10) is quasilinear and its solution exists for sufficiently smooth $\tilde{\mu}(\cdot)$, $\tilde{\sigma}(\cdot)$, and smooth initial conditions $p_0(Z)$ [29].

Employing the state density $p(Z, t|p_0)$, the output density $p(Y, t|p_0)$ is computed via marginalization,

$$p(Y, t|p_0) = \int_{\mathbb{R}^{n_Z}} p(Y|Z, t)p(Z, t|p_0)dZ. \tag{12}$$

As the measurements are noise corrupted, the density of the measured outputs $p(\bar{Y}, t|p_0)$ is obtained from the actual output density $p(Y, t|p_0)$ by convolution with the noise model:

$$p(\bar{Y}, t|p_0) = \int_{\mathbb{R}^{n_Y}} p(\bar{Y}|Y)p(Y, t|p_0)dY. \tag{13}$$

3.3. Numerical solution of PDE

In order to study the time evolution of the output density $p(Y,t|p_0)$ and the measured output density $p(\overline{Y},t|p_0)$, system (10) has to be solved for given $p_0(Z)$. As $p(Z,t|p_0)$ is defined on the n_Z-dimensional space, standard grid-based solvers are not able to solve system (10) for $n_Z = n_X + n_P > 3$. Theoretically, the methods of characteristics can be used [29, 30] but for the high-dimensional system we are going to study, also this method is difficult to apply as it requires 1) gridding and 2) the calculation of high-dimensional integrals to determine the output density. Instead, a stochastic method is used [16, 17, 18], which is known from particle filtering [31].

This stochastic method is based on a particle description of the model, which is in our case equivalent to the cell ensemble model (1). To compute $p(\overline{Y},t|p_0)$, at first a sample $\{(X_0^{(i)},P^{(i)})\}_{i=1}^s$, is drawn from $p_0(Z)$, where s is the sample size. For this sample the single cell model (1) is simulated, resulting in a set of simulated outputs $\{Y_t^{(i)}\}_{i=1}^s$. The output $Y_t^{(i)}$ is then corrupted by noise according to (4) resulting in $\{\overline{Y}_t^{(i)}\}_{i=1}^s$. Given the sample $\{\overline{Y}_t^{(i)}\}_{i=1}^s$ a numerical approximation of $p(\overline{Y},t|p_0)$ can be determined using the kernel density estimator described in Section 3.1.

This numerical stochastic approximation of the density $p(\overline{Y},t|p_0)$ can be shown to converge as $s \to \infty$ [15]. Thus, $p(\overline{Y},t|p_0)$ can be approximated also for high-dimensional nonlinear single-cell models. The advantage of this approach is that $p(\overline{Y},t|p_0)$ can be computed without calculating the in general high-dimensional state density $p(Z,t|p_0)$.

Remark 1. *The sample size s required to achieve a good approximation of the state and output density increases with the dimensionality of the corresponding densities. Fortunately, in the following only the output density is required which merely depends on the number of measurands n_Y used during one experiments. For typical experimental setups n_Y is not larger than three or four, rendering the problem tractable.*

4. Parameter estimation

As mentioned in Section 2 the problem studied in this work is the estimation of the initial condition and parameter density p_0 and its uncertainty. These problems are approached in the following by employing the density-based modeling approach derived in Section 3 in combination with parameterization, convex optimization, and parametric bootstrapping.

4.1. Estimation of parameter density

In this paper, we approach the problem of estimating p_0 from $\mathcal{D}_k$ by minimizing the l_2-norm of the model-data mismatch,

$$J(\hat{p}_0) := \sum_{k=1}^{N} \left|\left| p(\overline{Y}, t_k) - p(\overline{Y}, t_k|\hat{p}_0) \right|\right|_2^2, \tag{14}$$

where $p(\overline{Y}, t_k)$ is the density of the measured noisy output and $p(\overline{Y}, t|\hat{p}_0)$ is the predicted density of the measured noisy output obtained by simulation with the parameter density estimate $\hat{p}_0(Z)$. Note that the objective functional $J(\hat{p}_0)$ penalizes the difference between data density $p(\overline{Y}, t_k)$ and predicted noise corrupted output density $p(\overline{Y}, t_k|\hat{p}_0)$. This is possible as due to the large number of measured cells per measurement instance, we have good statistics on the measurement error. We note that a comparison of the measured output density $p(\overline{Y}, t_k)$ with the noise-free output density $p(Y, t_k|\hat{p}_0)$ yields in general worse estimation results for $\hat{p}_0(Z)$ than the comparison of the measured output density $p(\overline{Y}, t_k)$ with predicted noise-corrupted output density $p(\overline{Y}, t_k|\hat{p}_0)$. This is particularly the case for datasets $\mathcal{D}_k$ obtained with high measurement noise levels.

Remark 2. *Different methods are available to compare population models to cytometry data. In particular norm distances [32, 17] or likelihood functions [16, 18] are commonly used. In this work a l_2-norm distance has been chosen, as the evaluation of this objective function is computationally cheap. If the number of measured cells is small, likelihood functions may be superior.*

According to the objective functional J, the optimal parameter density $p_0^*(Z)$ is the solution of

$$\begin{aligned} \underset{\hat{p}_0}{\text{minimize}} \quad & J(\hat{p}_0) \\ \text{subject to} \quad & \int_{\mathbb{R}^{n_Z}} \hat{p}_0(Z) dZ = 1 \\ & \forall \xi \in \mathbb{R}^{n_Z} : \ \hat{p}_0(Z) \geq 0, \end{aligned} \tag{15}$$

in which the two constraints enforce that $\hat{p}_0(Z)$ is a probability density. Unfortunately, the optimization problem (15) is infinite dimensional. Therefore, a parametrization of $\hat{p}_0$,

$$\hat{p}_{0,\varphi}(Z) = \sum_{j=1}^{n_\varphi} \varphi_j \Lambda^j(Z), \tag{16}$$

with a weighting vector $\varphi \in \mathbb{R}^{n_\varphi}$ is introduced. The ansatz functions Λ^j are probability densities, $\Lambda^j(Z) \geq 0$ and $\int_{\mathbb{R}^{n_Z}} \Lambda^j(Z) dZ = 1$. To promote simplicity of the resulting optimization problem, the ansatz functions $\Lambda^j(Z)$ for $\hat{p}_0(Z)$ are chosen to be head functions, as depicted in Figure 2. This yields the simplified, finite-dimensional optimization problem,

$$\begin{aligned} \underset{\varphi}{\text{minimize}} \quad & J(\hat{p}_{0,\varphi}) \\ \text{subject to} \quad & \mathbf{1}^{\mathrm{T}}\varphi = 1,\ \varphi \geq 0, \end{aligned} \tag{17}$$

in which $\mathbf{1} = [1, \ldots, 1]^{\mathrm{T}} \in \mathbb{R}^{n_\varphi}$ denotes the volume of the j-th head function. The optimal solution of (17) is denoted by φ^*. Note that ansatz functions $\Lambda^j(Z)$ other than head functions are possible, e.g. polynomial or Fourier series, but the constraints ensuring that $\hat{p}_{0,\varphi}(Z)$ is a probability density are likely to be more difficult to handle.

To solve the parameterized optimization problem (17) the quasi-linearity of the density-based population model (10) is employed. As the superposition principle holds [29] for (10), the output $p(\bar{Y}, t|\hat{p}_{0,\varphi})$ can be written as the weighted sum

$$p(\bar{Y}, t|\hat{p}_{0,\varphi}) = \sum_{j=1}^{n_\varphi} \varphi_j p(\bar{Y}, t|\Lambda^j), \tag{18}$$

where $p(\bar{Y}, t|\Lambda^i)$ is the output density obtained from simulation with $\Lambda^j(Z)$ as density of parameters and initial condition. This allows the reformulation of the objective function to

$$J(\hat{p}_{0,\varphi}) = \sum_{k=1}^{N} \left\| p(\bar{Y}, t_k) - \sum_{j=1}^{n_\varphi} \varphi_j p(\bar{Y}, t_k|\Lambda^j) \right\|_2^2 .$$

Employing this (17) can finally be written as

$$\begin{aligned} \underset{\varphi}{\text{minimize}} \quad & \sum_{k=1}^{N} (A_k\varphi - b_k)^{\mathrm{T}} W (A_k\varphi - b_k) \\ \text{subject to} \quad & \mathbf{1}^{\mathrm{T}}\varphi = 1,\ \varphi \geq 0, \end{aligned} \tag{19}$$

where the integral $\|\cdot\|_2^2$ has been approximated, e.g. using the trapezoidal rule. The column vector b_k contains hereby the values $p(\bar{Y}, t_k)$ at the grid points of the

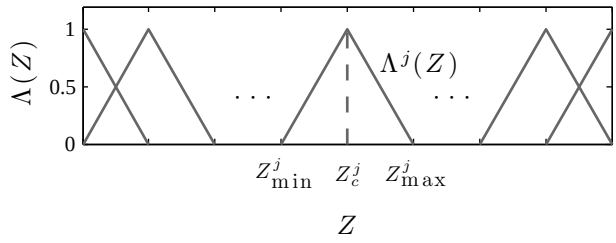

Figure 2: Illustration of head-type ansatz functions $\Lambda^j(Z)$.

discretization in the state/parameter space. Equivalently, the j-th column of A_k contains the values of $p(Y, t_k|\Lambda^j)$ at the grid points. The matrix W is a constant positive definite weighting matrix, determined by the chosen approximation of $\|\cdot\|_2^2$. Given (19) with the optimal weighting vector φ^*, the density of parameter and initial conditions with the smallest l_2-norm model-data mismatch is $\hat{p}_{0,\varphi^*}(Z)$.

Note that (19) is a convex, quadratic problem. Hence, fast convergence to the global optimum can be ensured.

Remark 3. *Although the derived convex optimization problem can be solved efficiently, high-dimensional parameter spaces still cause problems. The reason is the growing dimension of the weighting parameters φ with increasing dimension number. For high-dimensional systems the ansatz functions have to be chosen carefully, or an iterative refinement of the ansatz functions should be applied. Iterative methods may allow that only regions are high probability density are resolved in detail.*

4.2. Analysis of model uncertainty

Besides the optimal density $\hat{p}_{0,\varphi^*}(Z)$, the assessment of the reliability of a model also requires information about the model uncertainties. This is of particular importance when identifiability cannot be guaranteed. In the following, we present a bootstrap procedure to evaluate the uncertainty of the estimate $\hat{p}_{0,\varphi^*}(Z)$. Bootstrapping is a data-driven or model-driven approach from statistical inference commonly used to gather alternative versions of the single statistic. This enables the calculation of confidence intervals and hence the evaluation of identifiability and uncertainty.

In literature, many different types of bootstrap schemes are available. The most frequently used are case resampling, residual resampling, and parametric bootstraps [23, 33]. As in biology, measurement data are often limited, case resampling is not the method of choice. Also residual resampling is difficult to

apply, as in contrast to other applications distributions instead of individual data points are compared. Developing a reasonable resampling strategy for distributions may be difficult. Therefore, we present the application of parametric bootstrapping for the calculation of confidence intervals for $\hat{p}_{0,\varphi^*}(Z)$ [23, 33, 34, 35].

The procedure of parametric bootstrapping consists of four steps, as illustrated in Figure 3. In a first step, the set of available measurement data $\mathcal{D} = \{\mathcal{D}_k\}_{k=1}^N$ is used to determine an estimate $\hat{p}_{0,\varphi^*}(Z)$. This is the comon estimation step. The obtained estimate $\hat{p}_{0,\varphi^*}(Z)$ is in a second step used to generate r alternative, artificial realizations of the measurement data

$$\mathcal{D}^{[1]}, \mathcal{D}^{[2]}, \mathcal{D}^{[3]}, \ldots, \mathcal{D}^{[r]}. \tag{20}$$

The artificial data $\mathcal{D}^{[j]} = \left\{\mathcal{D}_k^{[j]}\right\}_{k=1}^N$, with $\mathcal{D}_k^{[j]} = \left\{\left(\overline{Y}_{t_k}^{(i),[j]}, t_k\right)\right\}_{i\in\mathcal{I}_k}$, are sampled from the predicted distribution (including noise) and have the same size as the original data set. Hence, $\mathcal{D}^{[j]}$ are typical measurement data we would obtain be studying a cell population with the parameter and initial state density $\hat{p}_{0,\varphi^*}(Z)$. In the third step, for each set of artificial data $\mathcal{D}^{[j]}$ the estimation is performed and the optimal density $\hat{p}_{0,\varphi^*}^{[j]}(Z)$ is computed, yielding

$$\hat{p}_{0,\varphi^*}^{[1]}, \hat{p}_{0,\varphi^*}^{[2]}, \hat{p}_{0,\varphi^*}^{[3]}, \ldots, \hat{p}_{0,\varphi^*}^{[r]}. \tag{21}$$

Given this set of estimated densities a statistical analysis is performed to determine the confidence intervals. In particular, the confidence intervals $[\hat{p}_{0,\varphi^*}^{\min}(Z), \hat{p}_{0,\varphi^*}^{\max}(Z)]$ are computed [23], e.g. with 95% or 99% confidence level.

Parametric bootstraps have been shown to provide reliable estimates of the confidence intervals. In particular, the uncertainty estimates are far more reliable than those obtained using local methods [35]. Unfortunately, the calculation of bootstraps is often computationally demanding as the estimation problem has to be solved many times ($r \gg 1$) [34, 35, 36]. This often limits the use of this method to small systems, and there are only few examples in literature where it has been applied for the uncertainty analysis of parameters of PDEs [37, 7]. Luckily, due to the facts that:

- the predicted output densities $p(\overline{Y}, t_k|\Lambda^j)$ and hence the matrices A_k can be reused, and that
- a convex formulation for the estimation problem (15) has been found,

redoing the estimation is computationally efficient. Given the data $\mathcal{D}^{[j]}$ the problem (19) can be solved in seconds even for large systems. This enables the usage

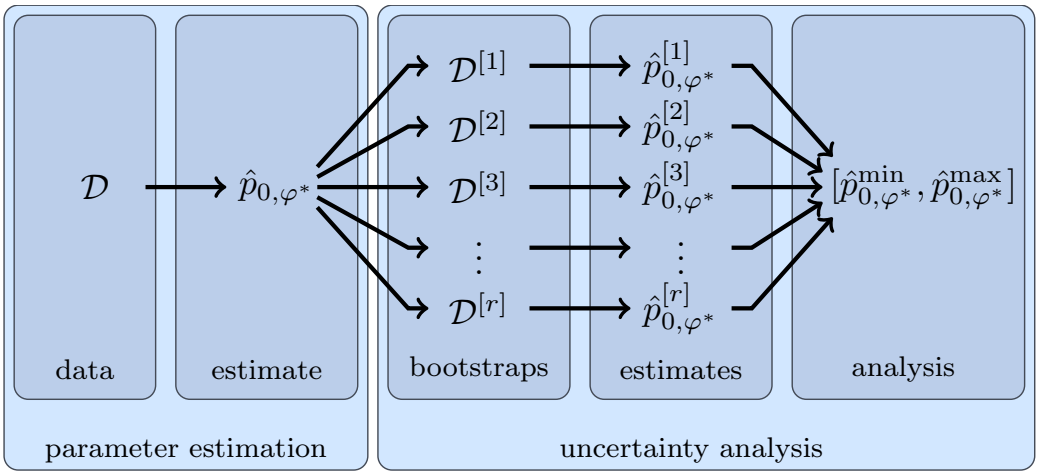

Figure 3: Illustration of parametric bootstrapping procedure as a tool for model-driven uncertainty analysis. For a detailed description of the workflow we refer to [23, 35].

of bootstrapping for uncertainty analysis of models of heterogeneous cell populations.

Summing up, in this section we presented a density-based framework of modeling and data handling for heterogeneous cell populations. This framework allows the formulation of the considered parameter estimation problem as a convex program. This ensures computational efficiency and allows for an efficient parametric bootstrapping analysis of the confidence intervals.

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 [38]. A mathematical model for this network has been proposed by Eissing *et al.* [39, 40]. Here, we consider the caspase activation in response to an external tumor necrosis factor (TNF) stimulus. As known from 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 induced by intrinsic and extrinsic noise, using for example cell population modeling.

The core properties of the TNF-induced proapoptotic signal transduction in

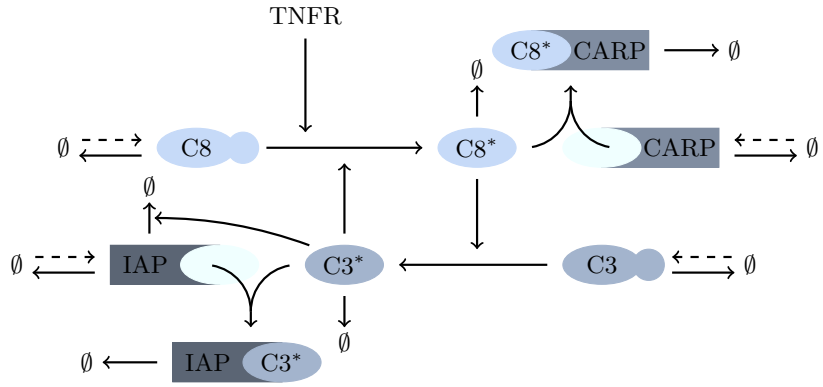

Figure 4: Schematic of the caspase activation cascade. Continuous arrows (→) indicate fully deterministic reactions and regulations while dashed arrows (⇢) indicate stochastic components.

a single cell, which is depicted in Figure 4, can be described by the 14 reactions provided in Table 1. The proapoptocic signaling cascade is induced by active TNF receptors, TNFR, which proteolytically cleave caspase 8, C8, yielding active caspase 8, C8*. Subsequently, C8* cleaves caspase 3, C3, to active caspase 3, C3*, which in turn cleaves C8, completing a positive feedback loop. To avoid apoptosis initiation for low TNFR concentrations, the caspase 8- and 10-associated RING protein, CARP, and the inhibitor of apoptosis protein, IAP, bind C8* and C3*, respectively. Thereby, C8* and C3* are inactivated. Besides the regulatory interaction all chemical species are continuously degraded and C8, C3, CARP, and IAP are synthesized.

The dynamics of the overall signal transduction pathway are governed by the SDE:

$$\begin{aligned}
d[\mathrm{C8}]_t &= (-v_1 - v_3 + v_{11})dt + \sigma k_{11} dW_{1,t} \\
d[\mathrm{C8}^*]_t &= (+v_1 + v_3 - v_4 - v_9)dt \\
d[\mathrm{C3}]_t &= (-v_2 + v_{12})dt + \sigma k_{12} dW_{2,t} \\
d[\mathrm{C3}^*]_t &= (+v_2 - v_5 - v_{10})dt \\
d[\mathrm{CARP}]_t &= (-v_4 + v_{13})dt + \sigma k_{13} dW_{3,t} \\
d[\mathrm{IAP}]_t &= (-v_5 - v_6 + v_{14})dt + \sigma k_{14} dW_{4,t} \\
d[\mathrm{C8}^*\sim\mathrm{CARP}]_t &= (+v_4 - v_7)dt \\
d[\mathrm{C3}^*\sim\mathrm{IAP}]_t &= (+v_5 - v_8)dt,
\end{aligned} \tag{22}$$

in which squared brackets denote the number of the respective molecule. As gene expression is a highly stochastic process [1, 41], protein synthesis involves

a stochastic component with $\sigma = 0.4$ and $W_{i,t}$, $i = 1, \ldots, 4$, being a Wiener processes. The remaining reactions are model fully deterministically as the molecule abundance is high. The presented model is an extension of those presented in [39, 40]. Also, some parameters have been adapted to fit available data for non-small lung cancer cell line NCI-H460.

As the system (22) is stiff, its simulation is challenging. The Euler-Maruyama scheme and the Milstein scheme, which are the most commonly used SDE solvers, fail to provide good results for reasonable step-sizes. Therefore, the trapeziodal rule is employed which outperforms for this system all other solvers evaluated in [42]. The time increment is set to four minutes.

Given the single cell model, extrinsic cell-to-cell variability is modeled by a log-normally distributed production rate of the inhibitor of apoptosis IAP, k_{14}, and a log-normally distributed amount of active 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 which may be caused by pathways which are not included in the model. Especially the concentration of IAP molecules contained in a cell is highly variable, and a variation in IAP production is known to affect cell death considerably [43]. Besides parameters also initial conditions among individual cells are different. The initial condition of i cell is drawn from the steady state distribution of the SDE (22) obtained for the parameter value $k_{14}^{(i)}$ for $[\text{TNFR}]^{(i)} = 0$. Due to this dependence of the initial state on the parameters, no additional degrees of freedom are introduced. Hence, the model describing the evolution of the population density is a 10-dimensional PDE.

In the remainder of this section, we study the possibility of estimating the density $\hat{p}_0(k_{14}, [\text{TNFR}])$ from population data of active caspase 3 obtained by flow cytometry,

$$Y_{t_k}^{(i)} = [\text{C3}^*]_{t_k}^{(i)}. \tag{23}$$

The measurement noise we considered contains additive component and multiplicative components,

$$\bar{Y}_{t_k}^{(i)} = \epsilon^{\times} Y_{t_k}^{(i)} + \epsilon^{+}, \tag{24}$$

and is therefore rather realistic [25]. Both components, ϵ^{+} and $\epsilon^{\times}$, are log-normally distributed. Median and standard deviation are $\mu^{\times} = 0$ and $\sigma^{\times} = 0.05$ for $\epsilon^{\times}$, and $\mu^{+} = 6.5$ and $\sigma^{+} = 0.25$ for ϵ^{+}, respectively.

The statistical model of the measured noisy output density, $p(\bar{Y}, t)$ is shown in Figure 5. It is determined using artificial measurement data of 10^4 cells at the measurement instances t_k, $k = 1, \ldots, 6$. This is a realistic number for standard cytofluorometric experiments.

Table 1: List of reactions and parameter values describing proapoptotic signaling. The molecule abundance is measured in molecules per cell, $\frac{\text{mo}}{\text{cell}}$, and the time unit is minute, min. For further information on reactions and parameter values we refer to [39].

	Reactions			Reaction rates	Parameter values and units			
R_1	C8 + TNFR	$\rightleftarrows$	C8* + TNFR	$v_1 = k_1$[TNFR][C8]	$k_1 = 1.0 \cdot 10^{-6}$	$\frac{\text{cell}}{\text{min·mo}}$	[TNFR] $= 2.0 \cdot 10^{+2}$	$\frac{\text{mo}}{\text{cell}}$
R_2	C3 + C8*	$\rightarrow$	C3* + C8*	$v_2 = k_2$[C3][C8*]	$k_2 = 4.0 \cdot 10^{-5}$	$\frac{\text{cell}}{\text{min·mo}}$		
R_3	C3* + C8	$\rightarrow$	C3* + C8*	$v_3 = k_3$[C3*][C8]	$k_3 = 1.0 \cdot 10^{-5}$	$\frac{\text{cell}}{\text{min·mo}}$		
R_4	C8* + CARP	$\rightleftarrows$	C8*~CARP	$v_4 = k_4$[C8*][CARP] $- k_{-4}$[C8*~CARP]	$k_4 = 5.0 \cdot 10^{-4}$	$\frac{\text{cell}}{\text{min·mo}}$	$k_{-4} = 2.1 \cdot 10^{-1}$	$\frac{1}{\text{min}}$
R_5	C3* + IAP	$\rightleftarrows$	C3*~IAP	$v_5 = k_5$[C3*][IAP] $- k_{-5}$[C3*~IAP]	$k_5 = 5.0 \cdot 10^{-4}$	$\frac{\text{cell}}{\text{min·mo}}$	$k_{-5} = 2.1 \cdot 10^{-1}$	$\frac{1}{\text{min}}$
R_6	C3* + IAP	$\rightarrow$	C3*	$v_6 = k_6$[C3*][IAP]	$k_6 = 3.0 \cdot 10^{-4}$	$\frac{\text{cell}}{\text{min·mo}}$		
R_7	C8*~CARP	$\rightarrow$	$\emptyset$	$v_7 = k_7$[C8*~CARP]	$k_7 = 1.16 \cdot 10^{-2}$	$\frac{1}{\text{min}}$		
R_8	C3*~IAP	$\rightarrow$	$\emptyset$	$v_8 = k_8$[C3*~IAP]	$k_8 = 1.73 \cdot 10^{-2}$	$\frac{1}{\text{min}}$		
R_9	C8*	$\rightarrow$	$\emptyset$	$v_9 = k_9$[C8*]	$k_9 = 5.8 \cdot 10^{-3}$	$\frac{1}{\text{min}}$		
R_{10}	C3*	$\rightarrow$	$\emptyset$	$v_{10} = k_{10}$[C3*]	$k_{10} = 5.8 \cdot 10^{-3}$	$\frac{1}{\text{min}}$		
R_{11}	$\emptyset$	$\rightleftarrows$	C3	$v_{11} = k_{11} - k_{-11}$[C3]	$k_{11} = 8.19 \cdot 10^{+1}$	$\frac{\text{mo}}{\text{min·cell}}$	$k_{-11} = 3.9 \cdot 10^{-3}$	$\frac{1}{\text{min}}$
R_{12}	$\emptyset$	$\rightleftarrows$	C8	$v_{12} = k_{12} - k_{-12}$[C8]	$k_{12} = 5.07 \cdot 10^{+2}$	$\frac{\text{mo}}{\text{min·cell}}$	$k_{-12} = 3.9 \cdot 10^{-3}$	$\frac{1}{\text{min}}$
R_{13}	$\emptyset$	$\rightleftarrows$	CARP	$v_{13} = k_{13} - k_{-13}$[CARP]	$k_{13} = 4.0 \cdot 10^{+1}$	$\frac{\text{mo}}{\text{min·cell}}$	$k_{-13} = 1.0 \cdot 10^{-3}$	$\frac{1}{\text{min}}$
R_{14}	$\emptyset$	$\rightleftarrows$	IAP	$v_{14} = k_{14} - k_{-14}$[IAP]	$k_{14} = 4.64 \cdot 10^{+2}$	$\frac{\text{mo}}{\text{min·cell}}$	$k_{-14} = 1.16 \cdot 10^{-2}$	$\frac{1}{\text{min}}$

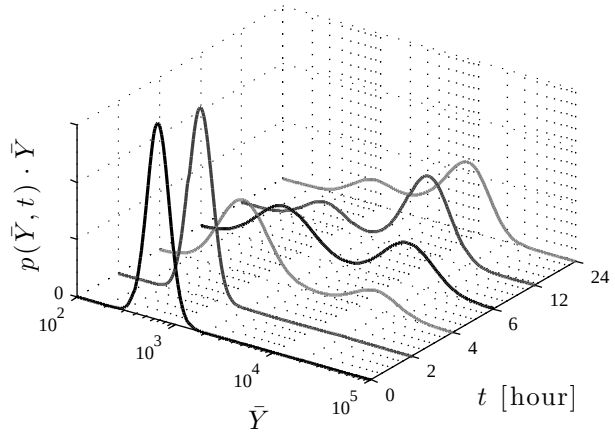

Figure 5: Statistical model $p(\bar{Y},t)$ of artificial noise corrupted measurement data of active caspase 3 derived from the 10^4 measured cells.

Based on these data, the estimation approach presented in Section 4 is used to obtain an estimate for the parameter density. 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 $\Lambda^j(k_{14},[\text{TNFR}])$ of $\hat{p}_{0,\varphi^*}(k_{14},[\text{TNFR}])$, resulting in 144 weighting parameters φ_j. For illustration purposes only the estimation results for the marginalized densities,

$$\begin{aligned}\hat{p}_{0,\varphi^*}(k_{14}) &= \int_{\mathbb{R}_+} \hat{p}_{0,\varphi^*}(k_{14},[\text{TNFR}])d[\text{TNFR}] \\ \hat{p}_{0,\varphi^*}([\text{TNFR}]) &= \int_{\mathbb{R}_+} \hat{p}_{0,\varphi^*}(k_{14},[\text{TNFR}])dk_{14},\end{aligned} \tag{25}$$

are depicted in Figure 6.

It is obvious that the parameter densities estimated from the data approximates the true parameter density well. Especially if we consider the finite number of degrees of freedom, limited time resolution of the data and the presence of intrinsic noise, the achieved results are very satisfying. As in applications the true densities are not known, the uncertainty of this estimate has to be studied. Therefore, the above presented approach is used and a parametric bootstrap with $r = 10^3$ members has been generated. The resulting 98% confidence intervals estimated from these bootstraps reveal that the data is sufficient to estimate $\hat{p}_{0,\varphi^*}(k_{14})$ and $\hat{p}_{0,\varphi^*}([\text{TNFR}])$ with reasonably small uncertainties. Note that the confidence interval provides only information about uncertainty of the parametrized density.

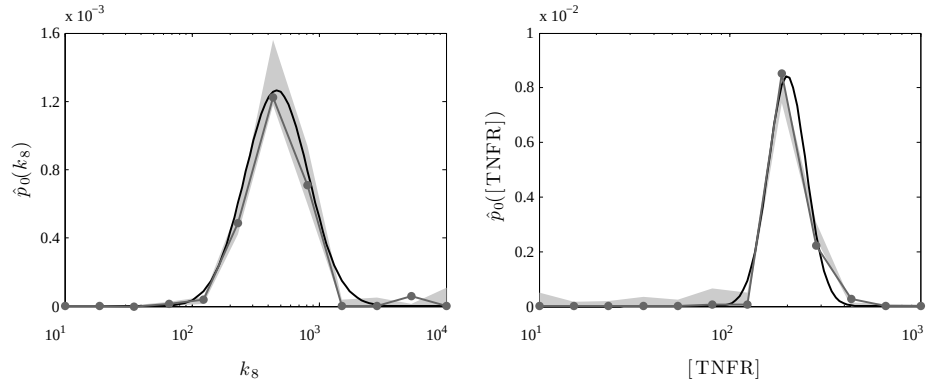

Figure 6: True (—) vs. estimated (–●–) parameter densities, with grid points (●). The 98% confidence interval (■) of the estimated parameter densities is shown as background.

The whole example indicates that even though there is a large measurement error and intrinsic noise on the single cell level, due to good statistics at the population level, the actual parameter density can be estimated accurately. Hence, the extrinsic cell-to-cell variability within the cell population can be unraveled. This is true also for higher noise levels if the number of measured cells is increased. Furthermore, this study shows that in principle, measuring one concentration can give enough information to estimate the density of multiple parameters, if the output density is sensitive with respect to these parameters.

6. Summary and outlook

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 intracellular state variable have been developed. In this paper a PDE model describing the time evolution of the state density is derived for systems with intrinsic and extrinsic noise. 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 densities underlying the heterogeneity. Therefore, a density-based statistical model of the sampled single cell is developed and applied in combination with convex l_2-norm optimization. To determine the uncertainty of the estimate a parametric bootstrapping approach has been presented, which again employs the problem structure. The presented approach is novel as it allows the estimation of the statistics of

extrinsic noise despite the presence of intrinsic noise using efficient convex optimization techniques. Also, the general idea behind the convex formulation of the estimation problem is transferable and may be used in related fields.

Finally, we applied the proposed approaches to artificial data of a medium size bistable system modeling the caspase activation cascade. It could be shown that the developed estimation approach yields good estimation results in case of a setup which is realistic in terms of noise and amount of available data. Furthermore, the bootstrapping-based uncertainty analysis approach could be used to study the information contained in the measurement data about the parameters to be estimated as well as the parameter uncertainties.

Concerning future research several open questions have been identified. One key aspect is the extension of the model class towards crosstalk among cells. Regarding the method, especially other parameterization approaches for the densities have to be considered, to reduce the number of required ansatz functions. This is crucial to enable the analysis and estimation of high-dimensional densities.

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 of Stuttgart.

References

[1] P. Swain, M. Elowitz, E. Siggia, Intrinsic and extrinsic contributions to stochasticity in gene expression, Proc. Nati. Acad. Sci. USA 99 (20) (2002) 12795–12800. doi:10.1073/pnas.162041399.

[2] A. Eldar, M. Elowitz, Functional roles for noise in genetic circuits, Nat. 467 (9) (2010) 1–7. doi:10.1038/nature09326.

[3] B. Snijder, L. Pelkmans, Origins of regulated cell-to-cell variability, Nat. Rev. Mol. Cell. Biol. 12 (2) (2011) 119–25. doi:10.1038/nrm3044.

[4] H. Tsuchiya, A. Fredrickson, R. Aris, Dynamics of microbial cell populations, Adv. Chem. Eng. 6 (1966) 125–206.

[5] A. Fredrickson, D. Ramkrishna, H. Tsuchiya, Statistics and dynamics of procaryotic cell populations, Math. Biosci. 1 (3) (1967) 327–374. doi:10.1016/0025-5564(67)90008-9.

[6] T. Luzyanina, D. Roose, T. Schenkel, M. Sester, S. Ehl, A. Meyerhans, G. Bocharov, Numerical modelling of label-structured cell population growth using CFSE distribution data, Theor. Biol. Med. Model. 4 (2007) 26. doi:10.1186/1742-4682-4-26.

[7] T. Luzyanina, D. Roose, G. Bocharov, Distributed parameter identification for label-structured cell population dynamics model using CFSE histogram time-series data, J. Math. Biol. 59 (5) (2009) 581–603.

[8] H. Banks, K. Suttona, W. Thompson, G. Bocharov, D. Roose, T. Schenkel, A. Meyerhans, Estimation of cell proliferation dynamics using CFSE data, Bull. Math. Biol. 73 (1) (2010) 116–150.

[9] D. Schittler, J. Hasenauer, F. Allgöwer, A generalized population model for cell proliferation: Integrating division numbers and label dynamics, in: H. Köppel, J. Aćimović, J. Kesselin, T. Mäki-Marttunen (Eds.), Proc. of Workshop on Comp. Syst. Biol. (WCSB 2011), Zürich, Switzerland, TICSP series # 57, Tampere International Center for Signal Processing, 2011, pp. 165–168.

[10] N. 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 (12) (2007) 4271–4288.

[11] M. Henson, D. Müller, M. Reuss, Cell population modelling of yeast glycolytic oscillations, Biochem. J. 368 (2) (2002) 433–446.

[12] B. Munsky, B. Trinh, M. Khammash, Listening to the noise: Random fluctuations reveal gene network parameters, Mol. Syst. Biol. 5 (2009) 318. doi:10.1038/msb.2009.75.

[13] S. Waldherr, J. Hasenauer, F. Allgöwer, Estimation of biochemical network parameter distributions in cell populations, in: E. Walter (Ed.), Proc. of the 15th IFAC Symp. on Syst. Ident. (SYSID), Saint-Malo, France, Vol. 15, IFAC-PapersOnline, http://www.ifac-papersonline.net, 2009, pp. 1265–1270.

[14] S. Spencer, S. Gaudet, J. Albeck, J. Burke, P. Sorger, Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis, Nat. 459 (7245) (2009) 428–433.

[15] C. Surulescu, N. Surulescu, A nonparametric approach to cells dispersal, Int. J. Biomath. Biostat. 1 (2010) 109–128.

[16] J. Hasenauer, S. Waldherr, N. Radde, M. Doszczak, P. Scheurich, F. Allgöwer, A maximum likelihood estimator for parameter distributions in heterogeneous cell populations, Procedia Computer Science 1 (1) (2010) 1649–1657.

[17] J. Hasenauer, S. Waldherr, M. Doszczak, P. Scheurich, F. Allgöwer, Density-based modeling and identification of biochemical networks in cell populations, in: M. Kothare, M. Tade, A. Wouwer, I. Smets (Eds.), Proc. of 9th Int. Symp. on Dynamics and Control of Process Syst. (DYCOPS), Leuven, Belgium, Vol. 9, 2010, pp. 320–325.

[18] J. Hasenauer, S. Waldherr, M. Doszczak, N. Radde, P. Scheurich, F. Allgöwer, Identification of models of heterogeneous cell populations from population snapshot data, BMC Bioinf. 12 (2011) 125.

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

[20] S. Reinker, R. Altman, J. Timmer, Parameter estimation in stochastic biochemical reactions, IEE Proc. Syst. Biol. 153 (4) (2006) 168–178.

[21] D. Wilkinson, Stochastic modelling for quantitative description of heterogeneous biological systems, Nat. Rev. Genet. 10 (2) (2009) 122–133. doi:10.1038/nrg2509.

[22] A. Golightly, D. Wilkinson, Markov chain Monte Carlo algorithms for SDE parameter estimation, MIT Press, 2010, Ch. 10, pp. 253–275.

[23] T. DiCiccio, B. Efron, Bootstrap confidence intervals, Statist. Sci. 11 (3) (1996) 189–228. doi:10.1214/ss/1032280214.

[24] L. Herzenberg, J. Tung, W. Moore, L. Herzenberg, D. Parks, Interpreting flow cytometry data: A guide for the perplexed, Nat. Immunol. 7 (7) (2006) 681–685. doi:10.1038/ni0706-681.

[25] C. Kreutz, M. B. Rodriguez, T. Maiwald, M. Seidl, H. Blum, L. Mohr, J. Timmer, An error model for protein quantification, Bioinf. 23 (20) (2007) 2747–2753.

[26] Z. Botev, J. Grotowiski, D. Kroese, Kernel density estimation via diffusion, Ann. Statist. 38 (5) (2010) 2916–2957. doi:10.1214/10-AOS799.

[27] H. Risken, The Fokker-Planck equation: Methods of solution and applications, 2nd Edition, Springer Berlin / Heidelberg, 1996.

[28] C. Gardiner, Handbook of stochastic methods: For physics, chemistry and natural sciences, 4th Edition, Springer Series in Synergetics, Springer Berlin / Heidelberg, 2011.

[29] L. Evans, Partial Differential Equations, American Mathematical Society, 1998.

[30] A. Weiße, R. Middleton, W. Huisinga, Quantifying uncertainty, variability and likelihood for ordinary differential equation models, BMC Syst. Biol. 4 (144). doi:10.1186/1752-0509-4-144.

[31] J. Rawlings, B. Bakshi, Particle filtering and moving horizon estimation, Comp. Chem. Eng. 30 (2006) 1529–1541. doi:10.1016/j.compchemeng.2006.05.031.

[32] B. Munsky, M. Khammash, Identification from stochastic cell-to-cell variation: A genetic switch case study, IET Syst. Biol. 4 (6) (2010) 356–366. doi:10.1049/iet-syb.2010.0013.

[33] J. Carpenter, J. Bithell, Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians, Statistics in medicine 19 (2000) 1141–1164.

[34] E. Balsa-Canto, A. Alonso, J. Banga, Computational procedures for optimal experimental design in biological systems, IET Syst. Biol. 2 (4) (2008) 163–172. doi:10.1049/iet-syb:20070069.

[35] M. Joshi, A. Seidel-Morgenstern, A. Kremling, Exploiting the bootstrap method for quantifying parameter confidence intervals in dynamical systems, Metabolic Eng. 8 (2006) 447–455.

[36] H. Banks, K. Kunsich, Estimation techniques for distributed parameter systems, Birkhauser, Boston, 1989.

[37] E. Balsa-Canto, M. Rodriguez-Fernandez, J. Banga, Optimal design of dynamic experiments for improved estimation of kinetic parameters of thermal degradation, J. Food Eng. 82 (2) (2007) 178–188.

[38] M. Hengartner, The biochemistry of apoptosis, Nat. 407 (6805) (2000) 770–776. doi:10.1038/35037710.

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

[40] T. Eissing, M. Chaves, F. Allgöwer, Live and let die - a systems biology view on cell death, Comp. and Chem. Eng. 33 (3) (2009) 583–589.

[41] N. Rosenfeld, J. Young, U. Alon, P. Swain, M. Elowitz, Gene regulation at the single-cell level, Science 307 (5717) (2005) 1962–1965. doi:10.1126/science.1106914.

[42] E. Buckwar, R. Winkler, Multistep methods for SDEs and their application to problems with small noise, SIAM J. Numer. Anal. 44 (2) (2006) 779–803. doi:10.1137/040602857.

[43] T. Eissing, S. Waldherr, E. Bullinger, C. Gondro, O. Sawodny, F. Allgöwer, P. Scheurich, 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 2006), 2006, pp. 1746–1752.