An efficient approximate moment method for multi-dimensional population balance models - Application to virus replication in multi-cellular systems

Robert Dürr[a,*], Thomas Müller[a], Stefanie Duvigneau[a], Achim Kienle[a,b]

[a] *Otto von Guericke University Magdeburg; Universitätsplatz 2, 39106 Magdeburg*
[b] *Max Planck Institute for Dynamics of Complex Technical Systems; Sandtorstraße 1, 39106 Magdeburg*

Abstract

Many particulate processes in process and bioprocess engineering can be described with multi-dimensional population balances. Approximate moment methods are frequently used for their solution. In the present paper a new approach is presented, which is particular efficient when the number of internal coordinates is high. It combines the direct quadrature method of moments with monomial cubatures. With the new method the computational effort increases only polynomially, in the simplest case even only linearly with the number of internal coordinates, compared to an exponential increase for the well known Gausssian cubatures. The technique is evaluated for a five dimensional benchmark problem describing virus replication in continuous cell cultures. Furthermore, the algorithm is applied to analyze influenza virus replication in genetically modified cell lines.

Keywords: multi-dimensional population balance modeling, direct quadrature method of moments, monomial cubature formulas, cell-to-cell variability, influenza virus replication, vaccine production

*Corresponding author

1. Introduction

Particulate systems are found in a broad field of process engineering applications in which single particles or individuals differ from each other with respect to certain characteristic properties. Examples from chemical processes include agglomeration [1], granulation [2, 3], crystallization [4–6] and coating processes [7]. Further examples for particulate processes are found in bioprocess engineering applications where multi-cellular systems are involved, e.g. cultivation of yeast [8–10], biopolymer production in microorganisms [11, 12] and vaccine production processes [13, 14]. As in the aforementioned examples, nonuniformity of cells with respect to physical properties, like size and shape, but also with respect to intracellular composition is observed. Besides nonuniformity in the process conditions, unsynchronized cell cycles, age distributions [15], stochastic effects on the gene expression level and bistable behaviour on the single cell level [16–18] play a major role in the formation of these variances.

Focus within this contribution is on virus replication in multi-cellular systems which is used for industrial vaccine production processes [13]. The principle process scheme is the following: a cell culture within a bioreactor is inoculated with a low quantity of virus by means of a typical low multiplicity of infection (MOI). The MOI describes the ratio of seed virus to uninfected cells. The seed virions infect the uninfected cells and start to replicate using cellular resources. Synthesized virus is released from the cells to the surrounding medium and can infect still uninfected cells. Thereby, the infection spreads within the cell culture resulting in increasing virus concentration in the medium. The harvested virus is later used for vaccine production. As for other multi-cellular processes cell-to-cell variability has been revealed by flow cytometric analysis [13, 14]. Mathematical models using population balances (see e.g. [19–22]) contribute to a better understanding of the underlying mechanisms and provide a sound basis for the design of suitable process intensification and process control schemes.

In general, two alternative modeling strategies can be followed for the formulation of suitable population balance equations (PBEs). The top down modeling

approach relies on measurements of the cell-to-cell variability by means of a few specific markers which can be measured with sophisticated techniques, e.g. flow cytometry. These are modeled directly and unstructured population balance models are obtained which represent low dimensional integro partial differential equations. Here, global mechanistic kinetics are used to describe the most important cellular processes. These have to be determined from the experimental data by solving an inverse problem. In our previous work [14, 23] a one-dimensional PBE for influenza A virus replication in MDCK cell cultures was derived and adapted to flow cytometric measurements of intracellular viral nucleoprotein content.

However, interpretation, predictive capacity and accuracy of top down models is limited. Alternatively, a bottom up modeling strategy can be pursued which is based on a detailed description of the single cell kinetics. For influenza vaccine replication such a detailed description features a large number of viral compounds [24]. To account for heterogeneity with respect to the cellular properties, the single particle description is transformed to a structured population balance model. As each particle property translates into an internal coordinate, high dimensional integro partial differential equations are obtained.

Analytic solutions for these PBEs are only found for special cases requiring the application of numerical solution algorithms e.g. discretization based methods like finite volume [25] or finite element methods [26, 27]. Though sophisticated extensions have been developed (see e.g. [28–31]), application is usually limited to low dimensional PBEs with a maximum of three internal coordinates due to increasing computational effort. Alternatively, moment methods can be employed. The basic idea is to track the dynamics of integral quantities of the distribution, so called moments. They are closely related to important properties like mean and variance with respect to the internal coordinates. Usually, those are much easier to interpret than the full number density distributions for most process engineering problems. Furthermore, the number density distribution may be reconstructed from an infinite or even a finite number of its moments [32]. The dynamic moment equations can be derived from the PBE

(see for example [19]) but computation within a closed set of moment equations is only possible for a restricted class of problems.

For most cases, an approximate closure has to be found, e.g. by using the quadrature method of moments (QMOM) where closure is obtained by approximating higher order moments by a weighted sum of abscissas [33]. The QMOM has been applied to a large number of examples from chemical and thermal process engineering (e.g. [31, 34]). Sophisticated extensions for multidimensional PBEs were developed (see e.g. [35–38]). However, due to numerical issues resulting from the underlying solution of generally nonlinear equation systems applications to problems with more than two internal coordinates are rarely found. Alternatively, the direct quadrature method of moments (DQMOM) [39, 40] can be employed. In contrast to QMOM it involves the underlying solution of a system of linear equations.

Besides, computation of the dynamics, the choice of the initial weights and abscissas is of crucial importance. In general, a large number of abscissas is attended by an increased approximation. Yet the numerical effort increases and thus a good balance between the two has to be found. In contrast to one-dimensional systems, where Gaussian quadratures offer both, an excellent accuracy and limited computational expenses, the task is much more complicated for multi-dimensional PBEs.

In this contribution we will present an efficient moment approximation technique for multi-dimensional PBEs. The method is based on the DQMOM in combination with monomial cubature rules for the choice of the initial abscissas and weights. These represent a good trade off between numerical effort and approximation accuracy. At first, an analytical solution for the dynamics of weights and abscissas will be derived which is equivalent to our recently presented technique [41]. Application will be demonstrated for virus replication in cell cultures using a generic model formulation. The accuracy of the presented approach is evaluated and compared to other approaches for the choice of initial weights and abscissas. Furthermore, application is shown for the analysis of the impact of host cell heterogeneity on influenza virus replication using genetically modified

cell lines.

Though focus of this work is on application to bioprocesses, the general algorithm is also suitable for multi-dimensional population balance systems as found e.g. when characterizing the dynamics of shape evolution in crystallization processes [4, 5]. It will be discussed how the algorithm can be further extended to describe systems in which cell division plays an important role.

2. Population balance modeling

Population balance modeling [21] offers a suitable framework to account for the variances in a non-uniform cell ensemble. Neglecting spatial heterogeneity, the dynamics of the corresponding cell number density distribution $n(t, \mathbf{x})$ is characterized by the general population balance equation [21, 42]

$$\frac{\partial n(t,\mathbf{x})}{\partial t} + \nabla \left\{ \mathbf{G}(\mathbf{t},\mathbf{x},\mathbf{c})\ n(t,\mathbf{x}) \right\} = -D(t,\mathbf{x},\mathbf{c}) n(t,\mathbf{x}) + P(t,\mathbf{x},\mathbf{c})\,. \qquad (1)$$

Therein, $\mathbf{x}$ represents the vector internal coordinates, e.g. cellular composition, which changes according to the intracellular kinetics $\mathbf{G}$. The right hand side represents sinks and sources resulting from cell death/withdrawal D and other kinetic processes P like infection in the present paper. Furthermore, the intracellular dynamics generally depend on a set of extracellular species $\mathbf{c}$, e.g. substrates in the medium. Their dynamics are given by

$$\frac{\mathrm{d}\mathbf{c}}{\mathrm{d}t} = \mathbf{D_c}\,(\mathbf{c}_{in} - \mathbf{c}) + \mathbf{P_c}(\mathbf{c},\mathbf{F}) \qquad (2)$$

where $\mathbf{D_c}$ characterizes the medium exchange and $\mathbf{P_c}$ the integral coupling of the continuous phase to the cell population

$$\mathbf{F} = \int_{\mathbf{X}} \mathbf{f}(\mathbf{x})\ n(t,\mathbf{x})\ \mathrm{d}\mathbf{x}\,. \qquad (3)$$

3. Approximate moment dynamics

In many cases, certain integral quantities of the cell number density distribution are used to characterize the multi-cellular system. These so called moments

are defined as

$$m_{l_1,\ldots,l_{N_d}} = \int_{\mathbf{X}} x_1^{l_1} \ldots x_{N_d}^{l_{N_d}} n(t,\mathbf{x}) \mathrm{d}\mathbf{x}\,, \tag{4}$$

and are directly related to important properties of the overall cell number density distribution, like the overall number of particles and mean with respect to an arbitrary particle property x_k. The corresponding dynamic moment equations are derived from combination of (1) and (4)

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} m_{l_1,\ldots,l_d} = &- \sum_{k=1}^{N_d} \int_{\mathbf{X}} x_1^{l_1} \ldots x_{N_d}^{l_{N_d}} \frac{\partial}{\partial x_k} \left\{ g_k\, n \right\} \mathrm{d}\ \mathbf{x} \\ &+ \int_{\mathbf{X}} x_1^{l_1} \ldots x_{N_d}^{l_{N_d}} (P - D\ n)\, \mathrm{d}\mathbf{x}\,.\end{aligned} \tag{5}$$

However, a closed set of equations for the moment dynamics can only be found for special classes of $\mathbf{G}$, D and P. Applying the DQMOM [39], it is assumed that $n(t,x)$ can be represented by a weighted sum of N_α delta functions. Hence, integral quantities can be approximated by

$$\int_{\mathbf{X}} f(\mathbf{x})\, n(t,\mathbf{x})\ \mathrm{d}\mathbf{x} \approx \sum_{\alpha=1}^{N_\alpha} f(\mathbf{x}_\alpha(t))\ w_\alpha(t) = \sum_{\alpha=1}^{N_\alpha} f_\alpha w_\alpha \tag{6}$$

with w_α being the weights and $\mathbf{x}_\alpha$ the abscissas. Instead of tracking the moment dynamics focus is on the temporal evolution of the weights and weighted abscissas

$$\frac{\mathrm{d}\, w_\alpha}{\mathrm{d}t} = a_\alpha\ , \qquad \frac{\mathrm{d}\ w_\alpha\, \mathbf{x}_\alpha}{\mathrm{d}t} = \mathbf{b}_\alpha\,. \tag{7}$$

These are derived as follows. At first, the approximation (6) is combined with (5) and (7), which yields

$$\begin{aligned}\sum_{\alpha=1}^{N_\alpha} \Bigg\{ &\left(1 - \sum_{k=1}^{N_d} l_k \right)\ x_{1,\alpha}^{l_1} \ldots x_{N_d,\alpha}^{l_{N_d}}\, a_\alpha + \\ &\begin{bmatrix} l_1\ b_{1,\alpha} \\ \vdots \\ l_{N_d}\ b_{N_d,\alpha} \end{bmatrix}^T \cdot \begin{bmatrix} x_{1,\alpha}^{l_1-1} x_{2,\alpha}^{l_2} \ldots x_{N_d,\alpha}^{l_{N_d}} \\ \vdots \\ x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2} \ldots x_{N_d,\alpha}^{l_{N_d}-1} \end{bmatrix} \Bigg\} = S_{\mathbf{x}}^{\mathbf{l}}\,.\end{aligned} \tag{8}$$

Here, $S^{\mathbf{l}}_{\mathbf{x}}$ contains the right hand side of (5) and can also be approximated using (6). The previous equation is evaluated for $(N_d+1)N_\alpha$ distinct moments resulting in a linear system

$$A \cdot \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} = \mathbf{S}^{\mathbf{l}}_{\mathbf{x}} \tag{9}$$

which is then solved for the unknown right hand side expressions of (7).

Note, that the choice of the moments for the construction of the linear system is of crucial importance for the overall performance of the moment approximation. Moments related to important quantities such as the zeroth order moment (overall number of particles) and the first order moments (related to mean values) should always be included. The remaining moments have to be chosen with care as some choices may lead to a poor approximation performance or even to singular matrices [40]. Furthermore, A resembles a van der Monde matrix and can become ill conditioned easily, in particular for a large number of abscissas and if two or more abscissas are close to each other in the state space, respectively. To overcome these problems different methods have been suggested, like the application of fractal moments [43]. Alternatively, numerical problems can be circumvented employing an analytical solution of (9).

In case of $P = 0$ such an analytical solution can be derived without even using the matrix formulation (9) [44]. For zeroth order moment (8) is simplified to

$$\sum_{\alpha=1}^{N_\alpha} a_\alpha = -\sum_{\alpha=1}^{N_\alpha} w_\alpha D_\alpha \tag{10}$$

and by comparing the coefficients the weight dynamics are given as

$$\frac{\mathrm{d}w_\alpha}{\mathrm{d}t} = -w_\alpha D_\alpha \,. \tag{11}$$

In a similar procedure, (8) can be used for arbitrary first order moments

$$\sum_{\alpha=1}^{N_\alpha} b_{k,\alpha} = \sum_{\alpha=1}^{N_\alpha} w_\alpha (g_{k,\alpha} - x_{k,\alpha} D_\alpha) \,, \quad k = 1, \ldots, N_d \tag{12}$$

to derive the dynamics of the weighted abscissas

$$\frac{\mathrm{d}w_\alpha x_{k,\alpha}}{\mathrm{d}t} = b_{k,\alpha} = w_\alpha(g_{k,\alpha} - x_{k,\alpha}D_\alpha)\,, \quad k = 1,\ldots,N_d\,. \tag{13}$$

Combination of the last two relations yields

$$\frac{\mathrm{d}\mathbf{x}_\alpha}{\mathrm{d}t} = \mathbf{G}_\alpha\,. \tag{14}$$

which also means that the abscissas move along the characteristic curves of (1) and the DQMOM reduces to a method of characteristics for $P = 0$. For introduction to characteristics we refer to [45]. It has to be mentioned, that even if these relations are derived from zeroth and first order moments, conservation of higher (and fractal) order moments is also guaranteed by means of fulfilling (8) for arbitrary $\mathbf{l}$. For the two dimensional case this shown in the Appendix. In [41] the same results for the dynamics of abscissas and weights (denoted scaling factors) have been obtained based on a scaled PBE.

4. Efficient choice of abscissas and weights

In the previous section the dynamics of abscissas and weights were derived. However, so far no word was lost about the initial values for the corresponding ODEs (11) and (14). In the DQMOM context initial abscissas and weights are chosen based on moments (or more general integral quantities) of the initial number density distribution.

$$\begin{aligned} m_{l_1,\ldots,l_{N_d}}(t=0) &= \int_{\mathbf{X}} x_1^{l_1}\cdots x_{N_d}^{l_{N_d}}\; n(t=0,\mathbf{x})\;\mathrm{d}\mathbf{x}\,, \\ &\approx \sum_{\alpha=1}^{N_\alpha} w_\alpha(t=0)x_{1,\alpha}^{l_1}(t=0)\cdots x_{\alpha,N_d}^{l_{N_d}}(t=0)\,. \end{aligned} \tag{15}$$

The choice of initial weights and abscissas represents an important factor in the performance of the overall moment approximation algorithm. For the following explanations it is assumed that the initial distribution is normalized

$$\int_{\mathbf{X}} n(t=0,\mathbf{x})\;\mathrm{d}\mathbf{x} = \int_{\mathbf{X}} n_0(\mathbf{x})\;\mathrm{d}\mathbf{x} = 1\,. \tag{16}$$

In general, a larger number of abcissas N_α comes along with an increased approximation accuracy but also with an increased computational effort as the number of ODEs for a DQMOM approximation of the moments increases. In standard literature on numerical integration (see e.g. [46, 47]) a large number of cubature formulas is found which can be roughly classified into random based and deterministic rules, where the latter contain product rules and non product rules. The abscissas and weights are only computed once at the beginning of the overall algorithm. Thus, computational effort of this step is negligible compared to the numerical effort for the solution of the weight and abscissa dynamics.

When applying random based rules, abscissas are determined by random sampling of the integration region. The corresponding numerical integration method is also known as Monte Carlo integration [46]. Its popularity is based on the straightforward generation of abscissas and it is commonly used to generate reference solutions. However, in general a large number of samples is necessary and thus overall computational effort is unreasonable large.

Product rules are multi-dimensional extensions of one-dimensional quadrature rules. In the one-dimensional case, Gaussian quadrature rules which are based on orthogonal polynomials can be applied to come up with appropriate sets of abscissas and weights, e.g. if the initial distribution corresponds to a Gaussian distribution, the Gauss Hermite rule may be applied [46]. These formulas can be extended to multi-dimensional problems by using tensor products of one-dimensional weight and abscissa sets. In consequence, those rules suffer badly from the curse of dimensionality, as the sizes of the weight and abscissa sets generally increase exponentially with the number of dimensions. This is a major disadvantage for high dimensional applications.

Alternatively, the set of abscissas is generated directly in the full property state space instead of tensoring using non product formulas. Their basic idea is to exploit special properties of $n_0(t, \mathbf{x})$ to come up with abscissa and weight sets that scale polynomially, in the best case even linearly with N_d. A derivation based on generator functions can be found in [48]. Those rules are also denoted as monomial cubature rules.

One of the most renowned formulas are sigma point formulas as introduced by Julier and Uhlmann (see e.g. [49, 50]). They are commonly applied for the unscented Kalman filter for state and parameter estimation. For a N_d dimensional Gaussian distribution $\mathcal{N}(\mu, \mathbf{\Sigma})$ the abscissas and weights can be determined by

$$\begin{array}{rcrc} x_0 = & \mu & w_0 = & \frac{\lambda}{\lambda+N_d} \\ x_i = & \mu + \sqrt{\lambda+N_d}\sqrt{\mathbf{\Sigma}}_i & w_i = & \frac{1}{2(\lambda+N_d)} \\ x_{N_d+i} = & \mu - \sqrt{\lambda+N_d}\sqrt{\mathbf{\Sigma}}_i & w_{N_d+i} = & \frac{1}{2(\lambda+N_d)} \end{array} \tag{17}$$

where the tuning parameter λ is given by

$$\lambda = \widetilde{\alpha}^2(\kappa - N_d) + N_d \tag{18}$$

and controls the spread of the abscissas in the state space and $\sqrt{\mathbf{\Sigma}}_i$ is the i-th column of the covariance matrix square root. Here, κ and $\widetilde{\alpha}$ are additional tuning parameters [50]. The size of the abscissa set scales linearly with dimension, which is a crucial advantage in particular for high dimensional applications. To improve the accuracy higher order non product rules have been suggested, e.g. in [51], which do scale polynomially with dimension.

If the assumption of an Gaussian initial distribution is not justified, the sigma point approach has to be modified. For some types of the initial distribution special transformation formulas exist. For example if n_0 corresponds to a one-dimensional logarithmic normal distribution $\mathcal{L}(\mu_\mathcal{L}, \sigma_\mathcal{L})$ the transformed set can be computed as follows [52]

1. sigma points $\mathbf{x}_i$ are computed for $\mathcal{N}(0, 1)$ according to (17)
2. the transformed sigma point set is then given by

$$x_{i,\mathcal{L}} = \exp\left(\mu_\mathcal{L} + \sigma_{\mathcal{L}_k} x_i\right) , \quad i = 1, \ldots, N_\alpha \tag{19}$$

3. the set of weights is not changed within the transformation procedure and is given by (17).

Further examples for those transformation formulas include γ- and t-Student distributions [52].

Yet, no general transformations exists if n_0 is multi modal. To improve this situation a two step procedure is advantageous, which was applied in [53] to approximate uncertainties of biological models. At first, n_0 is approximated by a weighted sum of Gaussian distributions

$$n_0(\mathbf{x}) \approx \sum_{k=1}^{N_{\mathrm{GMD}}} w_k^{\mathrm{GMD}} \mathcal{N}(\mu_k, \mathbf{\Sigma}_k)\,. \tag{20}$$

This approximation is also termed Gaussian mixed density (GMD) in some references. Afterwards, abscissas and weights are determined for each Gaussian applying the standard sigma point formula. For the overall approximation of the integrals the general formula (6) has to be adapted

$$\int_{\mathbf{X}} f(\mathbf{x}) n_0(\mathbf{x}) \mathrm{d}\mathbf{x} \approx \sum_{k=1}^{N_{\mathrm{GMD}}} w_k^{\mathrm{GMD}} \sum_{\alpha=1}^{N_\alpha} w_{k,\alpha} f(\mathbf{x}_{k,\alpha})\,. \tag{21}$$

The overall number of weights and abscissas is now given by $N_{\mathrm{GMD}} N_\alpha$. For more complex distributions (e.g. multi modal on a logarithmic scale) the previously mentioned approaches can be combined to come up with more accurate approximations.

In the following, the different approaches will be evaluated for a generic benchmark problem describing virus replication in a multi-cellular system. Furthermore, the technique will be applied to a high dimensional model which characterizes the spread of influenza A virus in a cell culture.

5. Application to a simple model of virus replication in cell cultures

5.1. Model formulation

The presented model is adapted from [54] and comprises the key elements of a viral replication process within a host cell. The general replication mechanism is depicted in Fig. 1. Virus particles bind to the surface of uninfected cells. After a virus particle has passed the cell membrane, the virus genome is uncoated and thus a certain amount of viral genetic information [*gen*] is injected to the cell. From this, a viral genomic template [*tem*] is produced. It delivers the blueprints

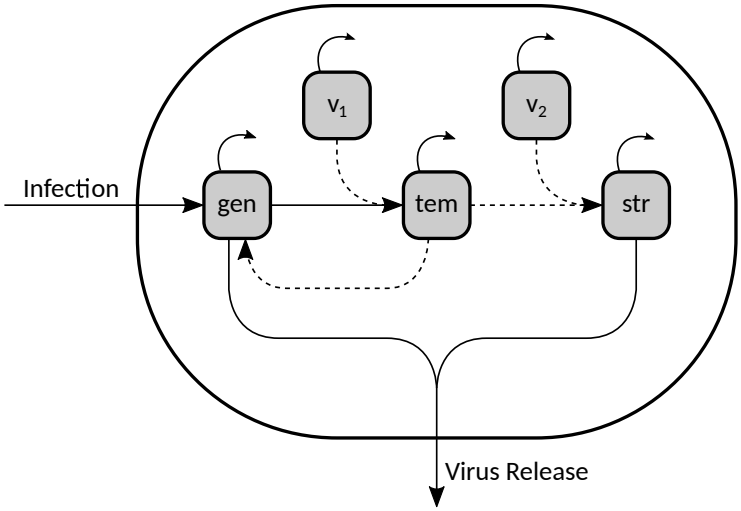


Figure 1: *Scheme of the assumed intracellular mechanisms for viral replication (gen: viral genome, tem: viral genome template , str: viral structural protein, v_1 & v_2: enzymes)*

for the replication of further viral genomic template and viral structural protein $[str]$. It is assumed, that the corresponding production rates are catalyzed by intracellular enzymes $[v_1]$ and $[v_2]$. It has to be mentioned that in addition to the original model proposed by Haseltine and coworkers [54, 55] a degradation of the virus genome template and the structural protein is modeled. Virus progeny is formed by binding of structural protein and viral genome. In a final step mature virus is released from the cell. Balancing of the species yields the following dynamics on the single cell level described by the following coupled system of ODEs

$$\dot{\mathbf{x}} = \frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} [tem] \\ [gen] \\ [str] \\ [v_1] \\ [v_2] \end{pmatrix} = \mathbf{h} = \begin{pmatrix} k_1\ [v_1]\ [gen] - k_6\ [tem] \\ k_3\ [tem] - k_1\ [v_1]\ [gen] - k_5\ [gen]\ [str] - k_7\ [gen] \\ k_2\ [tem]\ [v_2] - k_4\ [str] - k_5\ [gen]\ [str] \\ f_{v_1} \\ f_{v_2} \end{pmatrix} \tag{22}$$

where f_{v_1} and f_{v_2} are the enzyme degradation and production rates. By this means the analysis of host cell resource limitations on the process can be facilitated, see [56]. To transform the description to the macroscopic scale within the framework of population balance modeling, the single cell states directly translate into internal coordinates of a corresponding PBE. The dynamics of the number density distribution of infected cells is thus given by

$$\frac{\partial i_c(t,\mathbf{x})}{\partial t} = \underbrace{-\nabla_{\mathbf{x}}\left\{\mathbf{h}(\mathbf{x})\; i_c(t,\mathbf{x})\right\}}_{\text{intracellular reactions}} + \underbrace{k_{\text{inf}}\; U_c(t)\; V(t)\; \mathcal{I}(\mathbf{x})}_{\text{infection}} - \underbrace{k_{\text{cd},i_c}(\mathbf{x}) i_c(t,\mathbf{x})}_{\text{cell death}} \,. \tag{23}$$

Here, it is assumed, that infected cells are "produced" by binding of free virus particles V to uninfected cells U_c. At this point cell-to-cell variability for newly infected cells is taken into account by distributing infected cells into the space of internal properties according to a normal distribution

$$\mathcal{I}(\mathbf{x}) = \mathcal{N}\left(\mu, \mathbf{\Sigma}\right) \tag{24}$$

with mean μ and covariance Σ. At this point, it has to be mentioned that the special case when the population becomes stochastic due to a low number of virus particles or uninfected cells as described in [57] is not considered in the present publication. It is thereby assumed, that stochastic fluctuations of the expected number of newly infected cells $k_{\text{inf}}\; U_c(t)\; V(t)$ are neglected.

In result of the infection, cells undergo apoptosis and lysis. These cellular processes are summarized as "cell death" and are represented by a cell death coefficient $k_{\text{cd},i_c}(\mathbf{x})$. Its functional dependency on the vector of internal coordinates is motivated by the common assumption, that apoptosis and lysis are commonly induced by specific viral proteins.

In contrast to infected cells, uninfected cells U_c are not differentiated with respect to their intracellular composition. In result, their dynamics is represented by the following ODE

$$\frac{\mathrm{d}U_c(t)}{\mathrm{d}t} = -k_{\text{inf}}\; U_c(t) V(t) + k_{\text{gro},U_c}\; U_c(t) - k_{\text{cd},U_c}\; U_c(t)\,. \tag{25}$$

Table 1: *Parameter values for adapted Haseltine model* (22) - (27)

Parameter	Value		Parameter	Value	
k_1	$3.13\ 10^{-4}$	$\#^{-1}tu^{-1}$	k_{inf}	$1\ 10^{-8}$	$\frac{[\text{Vol}]^2}{\text{cells virions } tu}$
k_2	$25.00\ 10^{-0}$	$\#^{-1}tu^{-1}$	k_{cd,i_c}	$5\ 10^{-3}$	tu^{-1}
k_3	$7.00\ 10^{-1}$	tu^{-1}	k_{cd,U_c}	$1.5\ 10^{-3}$	tu^{-1}
k_4	$2.00\ 10^{-0}$	tu^{-1}	k_{gro,U_c}	$1\ 10^{-3}$	tu^{-1}
k_5	$7.50\ 10^{-6}$	$\#^{-1}tu^{-1}$	$k_{\text{deg},V}$	$9\ 10^{-1}$	tu^{-1}
k_6	$1.00\ 10^{-1}$	tu^{-1}			
k_7	$1.40\ 10^{-1}$	tu^{-1}			

Therein, $k_{\text{gro}}\ U_c(t)$ and $k_{\text{cd},U_c}\ U_c(t)$ characterize the growth and cell death rates of the uninfected cells, respectively.

At the end of a successful replication, virus particles are released from the infected cells to the medium with individual rates. These depend on the particular intracellular state of each cell and is given by

$$r_{\text{rel}}(\mathbf{x}) = k_5\ [gen]\ [str]\,. \tag{26}$$

according to the single cell dynamics.

In addition to the dynamics of the cell species, the dynamics of the virus particles in the medium has to be taken into account. For this ideal mixing is assumed. The overall dynamics of active virus particles are thus determined by the following ordinary differential equation

$$\frac{\mathrm{d}V(t)}{\mathrm{d}t} = \int_{\mathbf{X}} r_{\text{rel}}\ (\mathbf{x})\ i_c(t,\mathbf{x})\ \mathrm{d}\mathbf{x} - k_{\text{inf}}\ \ U_c(t)\ V(t) - k_{\text{deg},V}\ V(t) \tag{27}$$

where the rate coefficient k_{deg} characterizes the degradation and inactivation of free virus particles. The integral term on the right hand side represents the virus release of all individual cells. The simulation results shown in the following are based on the set of parameter values given in Table 1. Therein, tu represents time units.

5.2. Implementation details

All cubatures are directly derived from the initial distribution of the infected cells $\mathcal{N}(\mu, \mathbf{\Sigma})$ to obtain the initial abscissas and weights. For the subsequent benchmark, five different cubatures have been implemented as representatives of the above mentioned cubature groups: mean approximation (ME, $N_{\alpha,\mathrm{ME}} = 1$), sigma point formula (SP, $N_{\alpha,\mathrm{SP}} = 11$), higher order non product formula (HONP, $N_{\alpha,\mathrm{HONP}} = 51$), Gaussian mixed density approach (GMD, $N_{\alpha,\mathrm{GMD}} = 110$) and Gaussian product formula (GA, $N_{\alpha,\mathrm{GA}} = 243$) which are all evaluated against an Monte Carlo integration (MC, $N_{\alpha,\mathrm{MC}} = 10^4$). A detailed description of rules and comments on the numerical effort in terms of size of the overall ODE system which has to be solved is given in the supplemetary information.

5.3. Single infection cycle

At first, a single infection cycle scenario is considered. Here, it is assumed that all cells are infected initially and no free virus is present in the medium at the beginning of the process. Thus, the initial conditions are given by

$$V(t=0) = 0\frac{\text{virions}}{[\text{Vol}]}\,, \quad U_c(t=0) = 0\frac{\text{cells}}{[\text{Vol}]}\,,$$
$$I_c(t=0) = I_{c,0}\frac{\text{cells}}{[\text{Vol}]}\,. \tag{28}$$

Thereby, any observed cellular heterogeneity would be an effect of initial cell-to-cell variability while heterogeneity resulting from the delayed infection process is negligible. For this reason, a similar experimental setup is often used to determine single cell parameters.

Vice versa, this setup is also very useful to analyze the performance of the presented moment approximation algorithm. As the infection process is negligible, the dynamics of the corresponding infected cell number density distribution is now given by

$$\frac{\partial i_c(t,\mathbf{x})}{\partial t} = -\nabla_{\mathbf{x}}\left\{\mathbf{h}(\mathbf{x})\; i_c(t,\mathbf{x})\right\} - k_{\mathrm{cd},i_c}(\mathbf{x}) i_c(t,\mathbf{x}) \tag{29}$$

with initial condition following from the above representations

$$i_c(t,\mathbf{x}) = \mathcal{I}(\mathbf{x}) = I_{c,0}\;\mathcal{N}(\mu,\,\mathbf{\Sigma})\,, \tag{30}$$

where

$$\mu = [1,\ 10,\ 1,\ 80,\ 40]^{\mathrm{T}} \quad , \quad \mathbf{\Sigma} = 0.05\ \mathrm{diag}\left(\mu^2\right)\ . \tag{31}$$

Furthermore, the dynamics of the free virions in the medium is defined by (27).

For the given setup, the initial number of infected cells is normalized to $I_{c,0} = 1$. Thus, in the following, all values can be viewed as normalized quantities. Now, arbitrary integral quantities of the infected cell number density with respect to the internal coordinates can be approximated by (6)

$$\int_{\mathbf{X}} f(\mathbf{x})\ i_c(t,\mathbf{x})\ \mathrm{d}\mathbf{x} \approx \sum_{\alpha=1}^{N_\alpha} w_\alpha(t)\ f(\mathbf{x}_\alpha(t))\,. \tag{32}$$

In accordance to the derivation presented previously, the dynamics of the weights and abscissas are given by

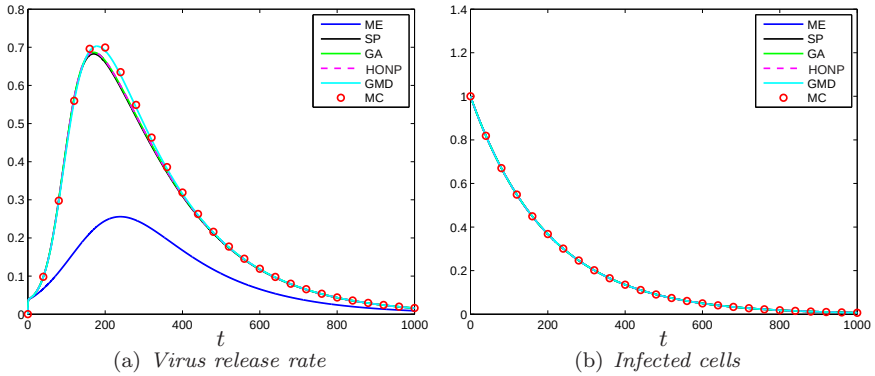
$$\dot{\mathbf{x}}_\alpha(t) = \frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} [tem]_\alpha \\ [gen]_\alpha \\ [str]_\alpha \\ [v_1]_\alpha \\ [v_2]_\alpha \end{pmatrix} = \mathbf{h}(\mathbf{x}_\alpha) \tag{33}$$

and

$$\dot{w}_\alpha(t) = k_{\mathrm{cd},i_c}(\mathbf{x}_\alpha)\ w_\alpha(t)\,. \tag{34}$$

The technique presented in the previous section can be applied directly to approximate integral quantities from the number density distribution including virus release rate (26) and moments with respect to the intracellular states. Within this setup the performance of different cubature formulas will be analyzed. A random based cubature rule with $N_{\mathrm{MC}} = 10^4$ is used as a reference solution. As the abscissas move along the characteristic curves of the PBE, this corresponds to a Monte-Carlo evaluation of the full solution achieved with the method of characteristics (see also [41, 58]).

In a first scenario, it is assumed, that the cell death does not depend on the intracellular composition. An example for this setup can be found when analyzing



(a) *Virus release rate* (b) *Infected cells*

Figure 2: *Comparison of the virus release rate integral and the overall number of infected cells for different cubatures*

the effects of physical or chemical stimuli on the cell, like medium temperature, radiation or pH-level which effect all cells in the same manner. Furthermore, the enzyme levels remain constant in course of the infection process. In consequence, the corresponding rate coefficients are given by

$$k_{\mathrm{cd},i_c}(\mathbf{x}) = k_{\mathrm{cd},i_c}\,, \quad f_{v_1} = f_{v_2} = 0\,. \tag{35}$$

In Fig. 2(a) the temporal change of the virus release rate is depicted for all previously introduced abscissa formulas. It can be seen that most formulas reach a good approximation accuracy with the GMD approach slightly outperforming the other approaches. In contrast, an approximation which uses only one abscissa located at the mean of the initial distribution exhibits an significant error. In Fig. 2(b) the approximations of the zeroth order moment of $i_c(t,\mathbf{x})$ (i.e. the overall number of infected cells)

$$I_c(t) = m_{\mathbf{0}}(t) = \int_{\mathbf{X}} i_c(t,\mathbf{x})\mathrm{d}\mathbf{x} \tag{36}$$

is depicted. It can be seen, that all cubature formulas yield an similar accuracy. In contrast, different approximation accuracies are observed for the approxima-

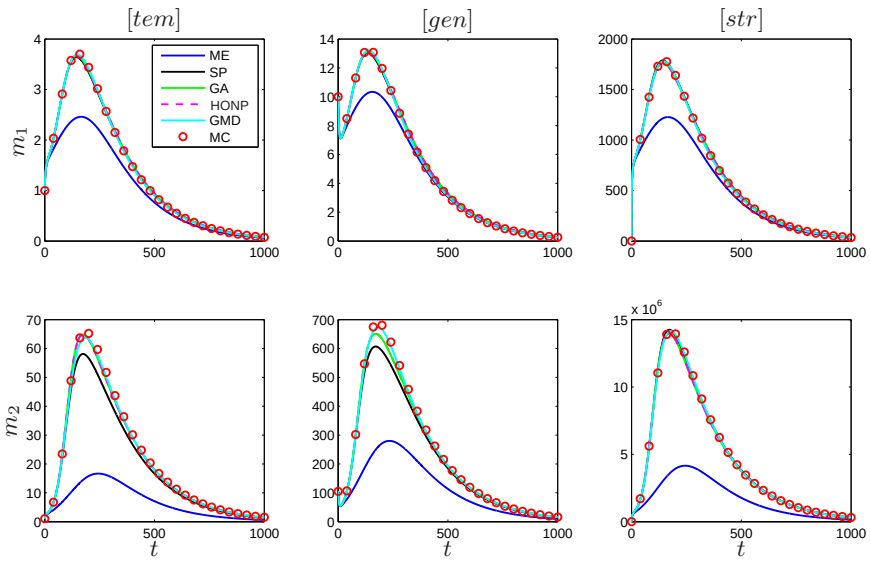


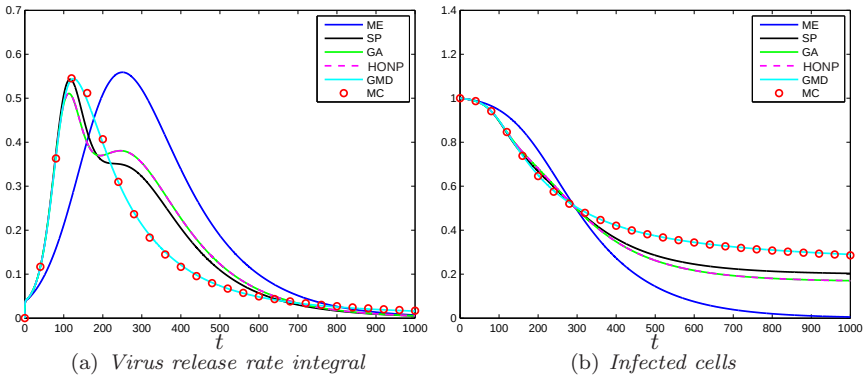
Figure 3: *First and second order pure moments of the distribution for different cubatures*

tion of the first and second order pure moments

$$m_{1,x_k}(t) = \int_{\mathbf{X}} x_k \; i_c(t, \mathbf{x}) \; \mathrm{d}\mathbf{x}$$
$$m_{2,x_k}(t) = \int_{\mathbf{X}} x_k^2 \; i_c(t, \mathbf{x}) \; \mathrm{d}\mathbf{x} \tag{37}$$

which can be seen in Fig. 3. Similar to the approximation of the overall virus release integral, good accuracy is obtained for all approaches except the ME formula (highlighted with a blue curve) for the approximation of the first moments. However, the GMD approach (cyan) outperforms the other formulas with a maximum relative error around 1%. The maximum errors of the HONP, SP and GA approaches are of similar values. In contrast, the SP approach exhibits larger errors than the other approaches (apart from the ME approach). In particular, this can be seen for the second order moment with respect to $[tem]$ where the maximum relative error is around 10%. Again, the GMD approach outperforms the other ones.

In the second scenario, it is assumed that each cell is characterized by an



(a) *Virus release rate integral* (b) *Infected cells*

Figure 4: *Overall virus release rate integral and overall number of infected cells for different cubatures*

individual cell death rate depending on the intracellular composition

$$k_{\mathrm{cd},i_c}(\mathbf{x}) = k_{\mathrm{cd},i_c} r_{\mathrm{rel}}(\mathbf{x})\,, \quad f_{v_1} = f_{v_2} = 0\,. \tag{38}$$

This means, that cells which are characterized by an increased intracellular amount of viral compounds $[gen]$ and $[str]$ have a lower survival probability than cells that are characterized by a low amount of these. Examples for such behavior are found for example for apoptosis induction [18, 59]. Apoptosis is one form of programmed cell death and is for example induced as a reaction to environmental stress or the viral infection.

Approximations of the integral virus release for the different approaches can be seen in Fig. 4(a). It can be observed, that neither the standard sigma point approach, nor the higher order non product cubature approach and the Gaussian approximation are able to approximate the overall virus release rate with a reasonable accuracy. In contrast, the GMD approach shows good results. The corresponding relative error stays within narrow bounds. The same is also observed from Fig. 4(b) which shows the overall number of infected cells. Apart from the approximation using the mean of the distribution as a sole abscissa, all other approaches show at least a sufficient performance up to 300 tu. Again, the GMD approach outperforms the other approaches and stays within very narrow error bounds. In Fig. 5 approximations of the first and second order

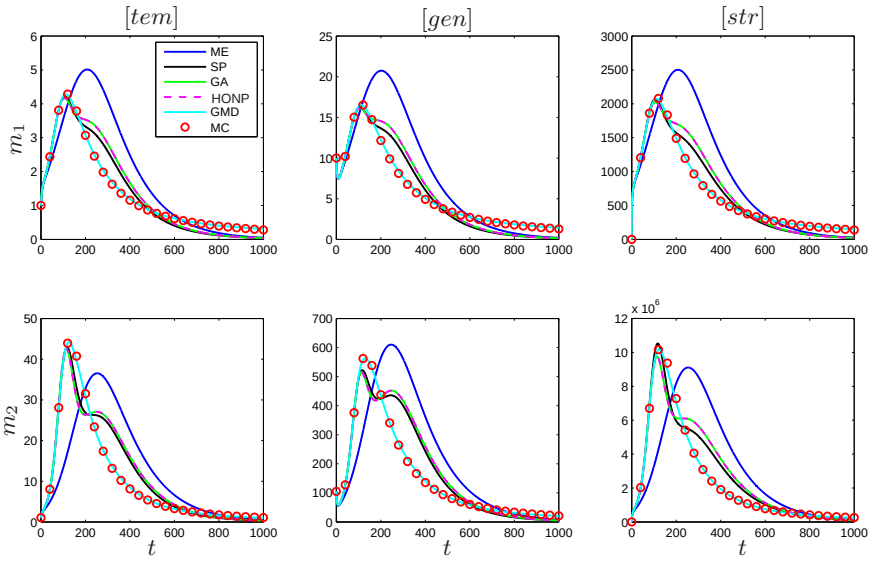


Figure 5: *First and second order pure moments of the distribution for different cubatures*

moments are depicted. As for the approximations of the overall cell number and the overall virus release, all approaches other than the GMD approach show significant approximation errors.

5.4. Multiple infection cycles

In the previous scenario, it was assumed, that all cells are infected initially. However, in real vaccine production processes, the initial MOI is generally low: A low number of virus particles is used to inoculate the reactor to obtain a large harvest of replicated virus particles. In the following, the accuracy of the approximation approaches presented in the previous section will be shown for this multiple infection cycles scenario.

In contrast to the previous scenario, now initially only virions and uninfected cells are present and thus the initial conditions are given by

$$U_c(t=0) = 10^8 \frac{\text{cells}}{[\text{Vol}]}, \quad V(t=0) = 10^3 \frac{\text{virions}}{[\text{Vol}]}, \qquad I_c(t=0,\mathbf{x}) = 0 \frac{\text{cells}}{[\text{Vol}]}. \tag{39}$$

In contrast to the single infection cycle setup, the full PBE (23) has to be taken into account to characterize the infected cell dynamics. In the DQMOM context,

this would require the numerical solution of (9) which may result in numerical difficulties. As an alternative, the overall problem is reformulated as a series of initial value problems

$$\frac{\partial i_{c,k}(t,\mathbf{x})}{\partial t} = -\nabla_{\mathbf{x}} \left\{ \mathbf{h}(\mathbf{x})\ i_{c,k}(t,\mathbf{x}) \right\} - k_{\mathrm{cd},i_c}(\mathbf{x})\ i_{c,k}(t,\mathbf{x}) \tag{40}$$

with initial conditions given by

$$i_{c,k}(t=t_k,\mathbf{x}) = \underbrace{k_{\mathrm{inf}}\ U_c(t_{k-1}) V(t_{k-1})\ \ (t_k - t_{k-1})}_{Inf(t_k)}\ \mathcal{I}(\mathbf{x})\,. \tag{41}$$

Thereby, each of the initial value problems describes the dynamics of cells $i_{c,k}(t,\mathbf{x})$ which have been infected in the interval $[t_{k-1},\ t_k)$. Comparing the obtained series reformulation to the single cycle infection formulation (29), it can be seen that the reformulated multiple infection cycles correspond to finite number of single infection cycle scenarios. The number of subpopulations N_k results from the chosen temporal discretization

$$t = [t_0,\ t_1, \ldots,\ t_{k-1},\ t_k,\ t_{k+1}, \ldots, t_{N_k}]\,. \tag{42}$$

Here, an equidistant grid with $(t_k - t_{k-1}) = 0.5\ tu$ and $N_k = 2000$ was used. Integral quantities (i.e. moments, overall virus release rate etc.) can be approximated by

$$\int_{\mathbf{X}} f_i(\mathbf{x}) i_c(t,\mathbf{x}) \mathrm{d}\mathbf{x} \approx \sum_{k=1}^{N_k} \int_{\mathbf{X}} f_i(\mathbf{x}) i_{c,k}(t,\mathbf{x}) \mathrm{d}\mathbf{x}\,. \tag{43}$$

At this point it has to be emphasized that the reformulation does in general not prevent the description of time variable environments. Each initial value problem can be solved simultaneously in this case. However, in the current example the intracellular virus kinetics are decoupled from the extracellular states (i.e. the single cell dynamics of the infected cells (22) do not depend on $U_c(t)$ and $V(t)$). Under these circumstances, the overall numerical procedure can be further simplified [55]. Instead of solving the full problem for the discrete reformulation, the single infection cycle scenario is solved only once for a normalized initial

condition

$$\frac{\partial i_c^*(t^*,\mathbf{x})}{\partial t^*} = -\nabla_{\mathbf{x}}\left\{\mathbf{h}(\mathbf{x})\ i_c^*(t^*,\mathbf{x})\right\} - k_{\mathrm{cd},i_c}(\mathbf{x})\ i_c^*(t^*,\mathbf{x})$$
$$i_c^*(t^*=0,\mathbf{x}) = \mathcal{I}(\mathbf{x}) = \mathcal{N}(\mu,\mathbf{\Sigma})\,. \tag{44}$$

Arbitrary integral quantities can be approximated using the moment approximation algorithm presented previously

$$\int_{\mathbf{X}} f_i(\mathbf{x}) i_c^*(t^*,\mathbf{x})\ \mathrm{d}\mathbf{x} \approx \sum_{\alpha=1}^{N_\alpha} f_i(\mathbf{x}_\alpha^*(t^*)) w_\alpha^*(t^*)\,. \tag{45}$$

Afterwards each of those is multiplied with the corresponding initial condition, i.e. the number of newly infected cells in the interval $[t_{k-1},\ t_k)$ (41) to obtain integral approximations for the subpopulations $i_{c,k}(t,\mathbf{x})$.

$$\begin{aligned}\int_{\mathbf{X}} f_i(\mathbf{x}) i_{c,k}(t,\mathbf{x})\ \mathrm{d}\mathbf{x} &= Inf(t_k) \int_{\mathbf{X}} f_i(\mathbf{x}) i_c^*(t-t_k,\mathbf{x})\ \mathrm{d}\mathbf{x} \\ &\approx Inf(t_k) \sum_{\alpha=1}^{N_\alpha} f_i(\mathbf{x}_\alpha^*(t-t_k)) w_\alpha^*(t-t_k)\end{aligned} \tag{46}$$

Thus, the virus dynamics (27) can be written as

$$\begin{aligned}\frac{\mathrm{d}V(t)}{\mathrm{d}t} &= \sum_{k=1}^{N_k} \int_{\mathbf{X}} r_{\mathrm{rel}}(\mathbf{x}) i_{c,k}(t,\mathbf{x})\ \mathrm{d}\mathbf{x} - k_{\mathrm{inf}}\ U_c(t)\ V(t) - k_{\mathrm{deg}}\ V(t) \\ &= \sum_{k=1}^{N_k} Inf(t_k) \sum_{\alpha=1}^{N_\alpha} r_{\mathrm{rel}}\left(\mathbf{x}_\alpha^*(t-t_k)\right) w_\alpha^*(t-t_k) - k_{\mathrm{inf}}\ U_c(t)\ V(t) - k_{\mathrm{deg}}\ V(t)\,.\end{aligned} \tag{47}$$

In the following, the same dependency of the cell death rate on the intracellular components as in the second single infection cycle scenario is assumed

$$k_{\mathrm{cd},i_c}(\mathbf{x}) = k_{\mathrm{cd},i_c} r_{\mathrm{rel}}(\mathbf{x}) = k_{\mathrm{cd},i_c} k_5 [gen][str]\,. \tag{48}$$

The overall uninfected and infected cell concentration dynamics are depicted in Fig. 6. It can be seen, that the concentration of infected cells increases significantly after a certain delay of round about 300 tu. This is a direct result of a low initial MOI: At the begin only a low number of uninfected cells gets infected and

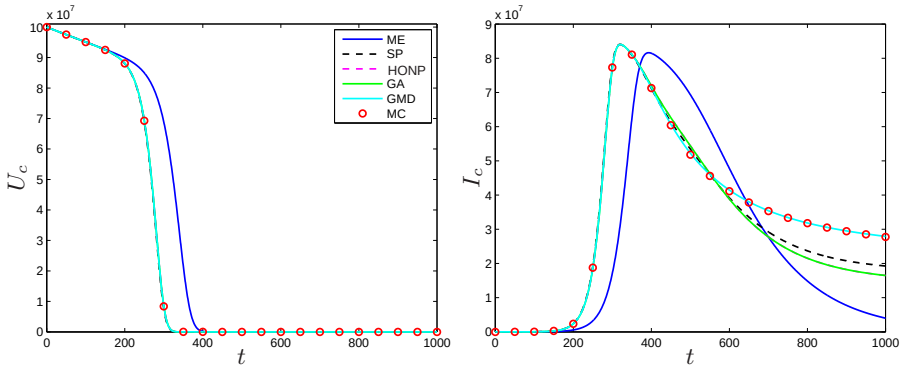


Figure 6: *Multiple infection cycle scenario: overall cell dynamics for different cubatures*

the infection takes a certain time to spread through the whole system. On the single cell scale, the viral replication mechanism and the resulting virus release rate (see Fig. 4(a)) additionally contribute to the delay.

It can be seen that all cubature formulas show nearly the same results for the uninfected cells, only the approximation which is based on a sole cubature abscissa at the mean of the initial distribution yields significant errors. However, differences are obtained for the dynamics of the overall number of infected cells. Again, the mean abscissa approximation shows the worst performance. Other approximation approaches, based on the standard sigma points, the Gaussian cubatures and the higher order non-product formula are at least sufficiently accurate for $t < 600\ tu$, but their approximation quality worsens for larger simulation times. Again, the performance of the Gaussian mixed density approach stands out and stays very close to the reference computed with a large number of random abscissas. These statements on the approximation accuracy are also valid for the overall virus concentration dynamics which are shown in Fig. 7 Excluding the mean abscissa approach, all approaches show a good performance for $t < 500\ tu$ but significant errors emerge for larger simulation times where only the Gaussian mixed density approach gives accurate results. The same is observed for the first and second order pure moments which are depicted in Fig. 8. It can be seen that the approximations based on standard sigma points, the Gaussian cubature or the higher order non product cubature formula are ac-

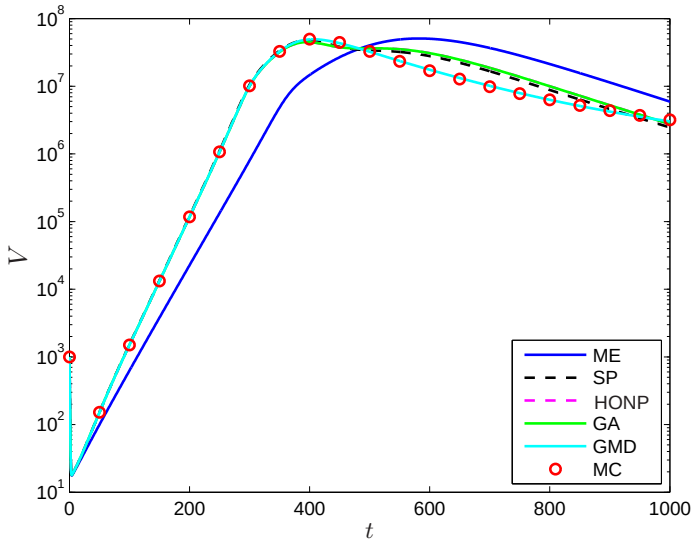


Figure 7: *Multiple infection cycle scenario: Virus concentration for different cubature formulas*

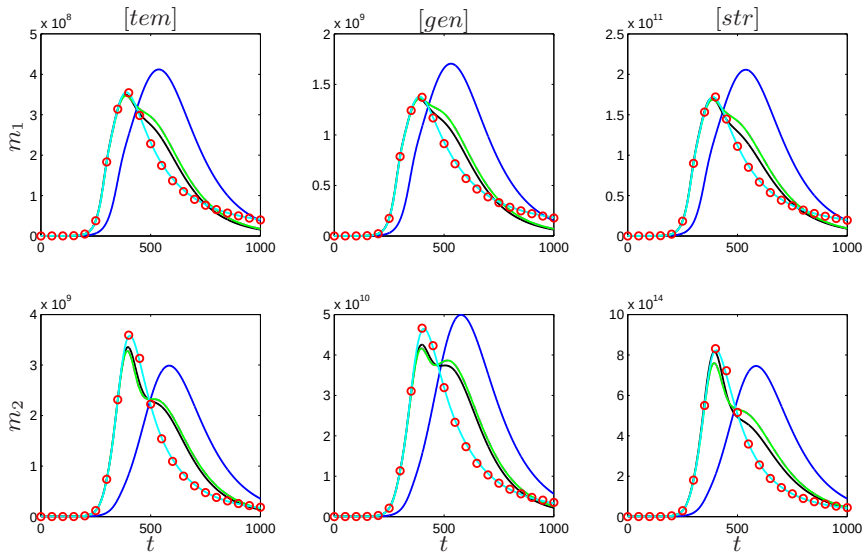


Figure 8: *Multiple infection cycle: Pure moments of the distribution for different cubatures*

curate up to $t \approx 400\ tu$ but show significant errors for later time points. Again it can be seen, that the GMD stays very close to the reference solution and provides an accurate approximation.

Thus, it has been shown that a sufficient accuracy of the multi-dimensional moments can be achieved by using GMD approach for the choice of the initial abscissas. This approach constitutes a direct extension of the SP approach with which the problem of choosing a good sigma point distribution can be avoided by the cost of more abscissas. For increasing N_{GMD} the overall GMD approximation approaches the corresponding MC approximation. Thus, by successively increasing the number of applied GMDs, a good approximation accuracy can be achieved easily. In the current example $N_{\mathrm{GMD}} = 11$ lead to an efficient approximation accuracy.

6. Application to a detailed model of influenza virus replication in cell cultures

In a second step, the methodology proposed in this paper is applied to a detailed model of influenza A virus replication in mammalian cell cultures. Cell culture based technology has developed as a promising alternative for influenza vaccine production compared to traditional processes in embryonated chicken eggs [60, 61]. Main advantage is increased flexibility which allows to respond rapidly to the frequently changing demands. From the practical point of view there is a high interest for the development of genetically engineered cell lines to establish a high yield production platform. Here, genetic modifications, e.g. by lentiviral transduction, are used to enhance or inhibit certain steps of the viral replication mechanism aiming for an increased cell specific viral production rate. However, usually not all cells are modified with the same efficiency [62]. Thus a significant cell-to-cell variability with respect to the intracellular kinetics and the viral production rates is expected.

To analyze the impact of these variances a detailed single cell model of influenza A virus replication in mammalian cell cultures can be used [24]. Therein, the interactions of a large number of viral components are considered. The basic scheme is depicted in Fig. 9.

The full set of dynamic equations and kinetic parameters according to [63]

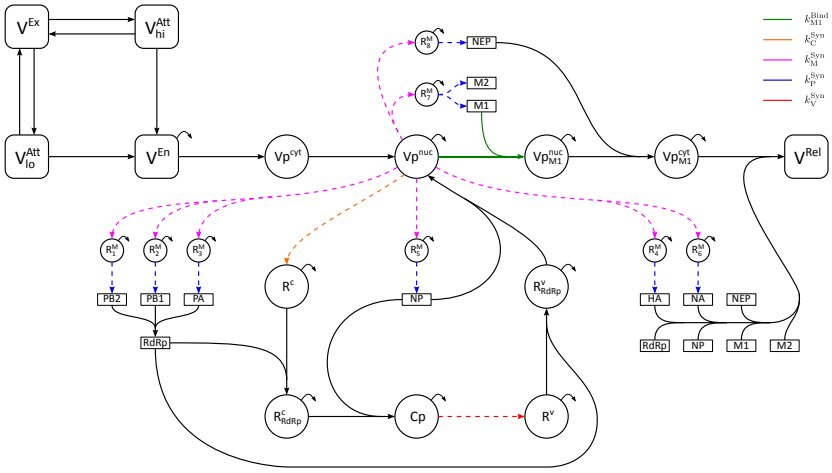


Figure 9: *Basic scheme of the single cell kinetics with affected reaction steps*

can be found in the supplementary information. In the following, a short overview of the process shall be given. Free virus particles V^{Ex} attach to the cell surface at binding sites with either high or low affinity. Attached virus particles (V^{Att}_{hi} and V^{Att}_{lo}) either detach again or they are absorbed via endocytosis resulting in the enclosure of the virus particle V^{En} by an intracellular endosome. After fusion of the viral envelope with the endosomal membrane the segmented viral genome is released into the cytoplasm in the form of eight viral ribonucleoprotein complexes (vRNPs). In Fig. 9 these vRNPs are denoted as Vp^{cyt}. Subsequently, they are imported into the nucleus (Vp^{nuc}) where transcription into messenger RNAs (mRNAs, $R^{M}_{1\ldots8}$) and complementary RNAs (cRNAs, R^{C}) takes place. Viral mRNAs migrate to the cytoplasm are translated into viral proteins PA, PB1, PB2, HA, NP, NA, NEP, M1, and M2 of which the first three form the RNA-dependent RNA polymerase (RdRp). In the nucleus cRNAs are stabilized by consecutive attachment of RdRp and nucleoprotein NP forming cRNPs (Cp). In a second step freshly produced cRNPs are used to synthesize new vRNA molecules (R^{v}), which are also encapsidated by RdRp and NP forming stabilized progeny vRNPs (Vp^{nuc}). By binding of viral matrix protein M1, vRNPs are excluded from any further participation in the replication cycle (Vp^{nuc}_{M1}).

Subsequent attachment of nuclear export protein NEP initiates the export to the cytoplasm (Vp_{M1}^{cyt}). Eventually, all necessary vRNP complexes and viral proteins assemble at the cell membrane to form progeny virions that bud from the surface into the surrounding medium as newly released virus particles (V^{Rel}).

The single cell description was used to come up with a bottom-up population balance model [44]. The latter can be viewed as direct extension of a previously proposed age-structured model of the process [63]. It enables the combination of detailed description of the viral replication mechanisms with the overall interaction of virus, infected and uninfected cells on the cell population level. In this contribution, recent numerical results will be presented briefly. For details on implementation we refer to [44].

Five significant steps of the viral replication cycle, i.e. the production of viral proteins, as well as viral mRNA, cRNA, vRNA synthesis and the binding rate of M1, were assumed to be target of genetic modifications aiming at an increased maximum of virus concentration and a minimum peak time. The modified intracellular reactions are highlighted in the single cell reaction scheme in Fig. 9. In a pragmatic sensitivity analysis 3^5 different parameter sets of the five rates were obtained to investigate the impact of up- or downregulation by means of increasing or decreasing the corresponding rates or keeping them at the original values. Therein, it was assumed that the rates are the same for all cells. For that reason each of the 3^5 candidate parameter sets was used to simulate the dynamic behaviour of a homogeneous cell population with the proposed approximate moment method using only one abscissa which is located at the exact parameter values. Here, it should be noted that the cell population is homogeneous regarding to the modified parameters, but a variability in the stage of infection caused by low initial seed virus concentration was considered further. Simulation results of each parameter set were screened for an increased maximum virus production and a minimum peak time to evaluate the up-/downregulation combination. Concerning that criteria, the combination of upregulated synthesis rates of mRNA, vRNA and proteins together with a downregulation of the binding rate of M1 to progeny vRNPs and an unchanged parameter for the

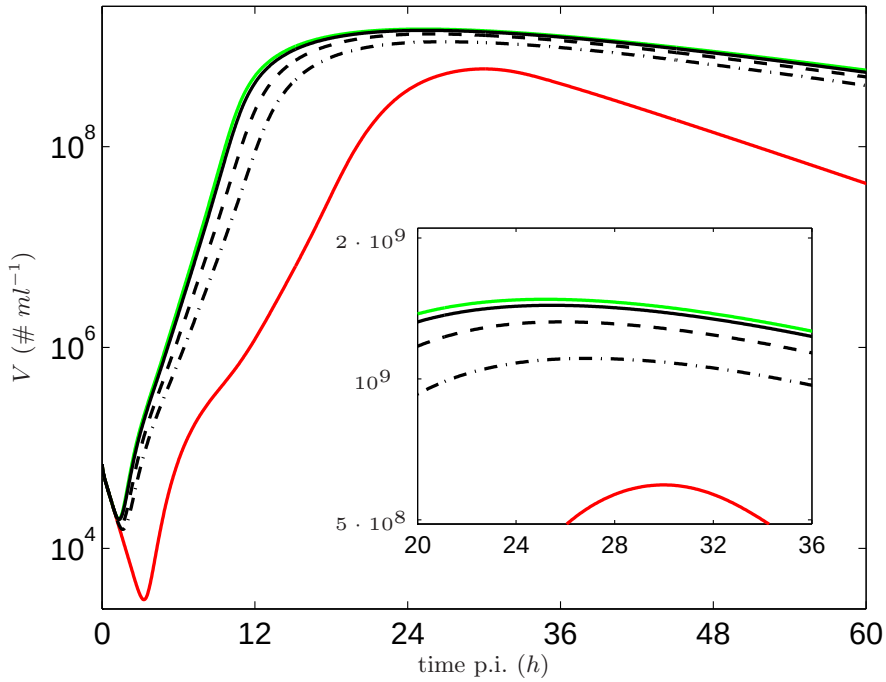


Figure 10: *Virus dynamics for detailed PBE for influenza virus replication in mammalian cell cultures; simulation results for homogeneous unmodified cell population (red), homogeneous modified cell population (green) and heterogeneous cell population (black, increasing cellular variance: solid → dashed → dash-dotted)*

synthesis of cRNA represents the best parameter set (see also μ_5 in Eq. 50). The simulation result with this best combination is depiced in Fig. 10 (green line). The red line in Fig. 10 represents the unmodified cell line dynamics for comparison. It can be seen that both aims, increase of the maximum virus yield and reduction of peak time are achieved (see also Tab. 2). Investigated cellular modifications can be achieved by biotechnological methods, e.g. lentiviral transduction [64, 65]. However, in general not all cells are modified with the same efficiency. This leads to cellular heterogeneity and can be accounted for by assuming that kinetic parameters are not uniform for the whole cell population but appear distributed. In consequence, different efficiencies of the genetic modifications can be mapped to broadening parameter distributions. The cell-to-cell variability resulting genetic modifications is represented by a weighted sum of

Table 2: *Maximum virus yield and peak time for modified and unmodified cell dynamics*

	$\max(V)$ $(\# \, ml^{-1})$	peak time (h)
unmodified	$5.94 \; 10^8$	30
homogeneous modification	$1.47 \; 10^9$	25.2
scenario I	$1.44 \; 10^9$	25.5
scenario II	$1.32 \; 10^9$	25.9
scenario III	$1.11 \; 10^9$	26.2

five logarithmic Gaussians to obtain a well-shaped distribution. As five reaction rates are considered, those Gaussians are five dimensional

$$\mathbf{k} = \sum_{l=1}^{5} a_l \; \mu_{\mathbf{l}} \cdot e^{\mathcal{N}(\mathbf{0},\mathbf{diag}(\sigma_{\mathbf{l}}))} \, . \tag{49}$$

The mean vectors of the distributions $\mu_{\mathbf{l}}$ are logarithmically distributed between the nominal parameter vector μ_1 and the best combination μ_5

$$\begin{aligned} \mu_1 &= \left[k_C^{Syn}, \; k_M^{Syn}, \; k_P^{Syn}, \; k_V^{Syn}, \; k_{M1}^{Bind} \right]^T \, , \\ \mu_5 &= \left[k_C^{Syn}, \; 5 \; k_M^{Syn}, \; 5 \; k_P^{Syn}, \; 5 \; k_V^{Syn}, \; 0.2 \; k_{M1}^{Bind} \right]^T \, . \end{aligned} \tag{50}$$

The variances are chosen as $\sigma_l = 0.05$ to obtain a well-shaped distribution. The influence of different modification efficiencies was analyzed for highly efficient, mid-range or broad distributions of the parameter rates based on publications on transduction methods [64, 65]. The three different scenarios are realized using different values of the weighting parameters a_l

$$\begin{aligned} \mathbf{a}_I &= [0.01, 0.03, 0.04, 0.06, 0.86] \, , \\ \mathbf{a}_{II} &= [0.04, 0.12, 0.19, 0.27, 0.38] \, , \\ \mathbf{a}_{III} &= [0.20, 0.20, 0.20, 0.20, 0.20] \, . \end{aligned} \tag{51}$$

For each scenario, the moment dynamics were approximated with our proposed technique applying the GMD approach. Different numbers of Gaussian distributions were evaluated against a Monte-Carlo approximation with $N_{\alpha,\mathrm{MC}} = 10^4$

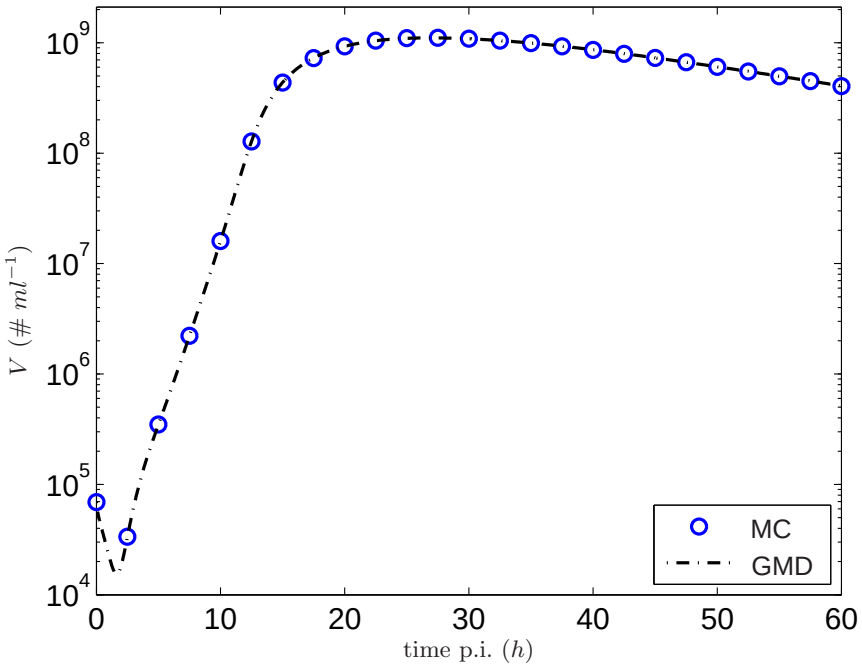


Figure 11: *Comparison of virus dynamics for third scenario; proposed GMD approach using* $N_{\alpha,\mathrm{GMD}} = 150$ *abscissas (dash-dotted) and reference MC-approximation using* $N_{\alpha,\mathrm{MC}} = 10^4$ *(blue)*

and $N_{\mathrm{GMD}} = 15$ was chosen for a good approximation accuracy (see Fig. 11). Thus, the overall number of abscissas is given by $N_{\alpha,\mathrm{GMD}} = N_{\mathrm{SP}} N_{\mathrm{GMD}} = 150$. It can be seen in Fig. 10, that an increasing variance is attended by a reduced maximum virus yield and an increasing peak time. However, in comparison to the unmodified cell line even for the worst case scenario an significant increase in the maximum virus concentration is obtained for the assumed genetic modifications (see Tab 2). Thus it can be concluded, that the above modifications are a promising candidate for improving influenza vaccine production processes.

7. Conclusion

Population balance modeling offers a suitable framework to characterize the dynamics of nonuniform particle ensembles as they are found in many applica-

tions from chemical and biotechnological process systems. In particular for the latter, a large number cellular properties has to be accounted for. The resulting multi-dimensional PBEs can be solved numerically using approximate moment methods.

In this manuscript we presented an efficient method for the approximate computation of moments for this type of models, which is based on the DQMOM. Analytical solutions for the dynamics of weights and abscissas are derived and, thus, the subsequent numerical solution of linear equation systems can be omitted . Furthermore, a sophisticated choice of the abscissas with a sigma point cubature rule ensures a limited numerical effort. In contrast to Gaussian cubatures which scale exponentially with the dimension of the problem, here, the numerically effort increases linearly. The algorithm is equivalent to our recently presented technique [41].

The algorithm is evaluated for a benchmark problem describing virus replication in cell cultures. Therefore different cubature approaches are compared to a Monte Carlo evaluation. The presented approach shows good accuracy for all implemented formulas in a first test case where cell death does not depend on the intracellular state. It is shown that the performance can be further improved by using higher order non cubature and a Gaussian mixed density approach, respectively. In contrast to other applied cubature rules, the latter also gives excellent results if cell death depends nonlinearly on the intracellular composition. The corresponding numerical effort increases in comparison to the standard sigma point formula but is still smaller compared to Gaussian cubatures.

In addition, application is briefly shown for a detailed PBE describing influenza virus replication. The presented technique is applied to analyze on the overall virus dynamics when using genetically modified cells. It was shown that even for a large degree of cellular variance, a significant increase in the maximum virus titer can be obtained. The results suggest, that genetic modifications may be a suitable tool to come up with a high yield cell line for influenza vaccine production.

In the future, focus will be on further application of the technique for model

based analysis of the effects of cellular variance when using genetically modified cell lines to overcome limitations in vaccine production processes. Furthermore, extension of the algorithm on bio systems with non negligible cell division is worthwhile. In addition, the method may be extended to problems with non negligible spatial gradients and to agglomeration processes. An alternative approach for the latter is the moving pivot method [66] which is also based on some moving reference points like the presented approach.

Although focus in this manuscript was on bio systems, the algorithm can also be applied to other processes that can be described by multi-dimensional population balance equations, e.g. shape evolution in crystallization processes.

Acknowledgements

Stefanie Duvigneau and Robert Dürr are also affiliated to the "International Max Planck Research School for Advanced Methods in Process and System Engineering", which is located in Magdeburg.

The authors gratefully acknowledge the funding of this work by the German Federal Ministry of Science and Education [Bundesministerium für Bildung und Forschung (BMBF)] as part of the "e:Bio" - project "CellSys" (grant number 031 6189 A).

Appendix

Evaluation of weight and abscissa dynamics for arbitrary moments. For $P = 0$ and $N_d = 2$ (8) reduces to

$$\sum_{\alpha=1}^{N_\alpha} \left[a_\alpha \left(1 - l_1 - l_2\right) x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2} + b_{1,\alpha} l_1 x_{1,\alpha}^{l_1-1} x_{2,\alpha}^{l_2} + b_{2,\alpha} l_2 x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2-1} \right] = \sum_{N_\alpha}^{\alpha=1} w_\alpha \left[l_1 x_{1,\alpha}^{l_1-1} x_{2,\alpha}^{l_2} G_{1,\alpha} + l_2 x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2-1} G_{2,\alpha} - D_\alpha x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2} \right] . \quad (52)$$

In the next step the analytic expressions for the dynamics of weights and abscissas can be derived by comparison of the coefficients for zeroth and first order

moments as

$$a_\alpha = -w_\alpha D_\alpha \,,$$
$$b_{1,\alpha} = w_\alpha \left(G_{1,\alpha} - x_{1,\alpha} D_\alpha\right) \,, \quad b_{2,\alpha} = w_\alpha \left(G_{2,\alpha} - x_{2,\alpha} D_\alpha\right) \,. \tag{53}$$

These are inserted to (52) resulting in

$$\begin{aligned}
&\sum_{\alpha=1}^{N_\alpha} \Big[-w_\alpha D_\alpha \left(1 - l_1 - l_2\right) x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2} \\
&+ w_\alpha \left(G_{1,\alpha} - x_{1_\alpha} D_\alpha\right) l_1 x_{1,\alpha}^{l_1 - 1} x_{2,\alpha}^{l_2} + w_\alpha \left(G_{2,\alpha} - x_{2_\alpha} D_\alpha\right) l_2 x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2 - 1} \Big] \\
&= \sum_{\alpha=1}^{N_\alpha} w_\alpha \left[l_1 x_{1,\alpha}^{l_1 - 1} x_{2,\alpha}^{l_2} G_{1,\alpha} + l_2 x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2 - 1} G_{2,\alpha} - D_\alpha x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2} \right] \,.
\end{aligned} \tag{54}$$

The left hand side expression can be rearranged and condensed to

$$\sum_{\alpha=1}^{N_\alpha} w_\alpha \left[l_1 x_{1,\alpha}^{l_1 - 1} x_{2,\alpha}^{l_2} G_{1,\alpha} + l_2 x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2 - 1} G_{2,\alpha} + x_{1,\alpha}^{l_1} x_{2,\alpha}^{l_2} D_\alpha (-1 + l_1 + l_2 - l_1 - l_2) \right] \tag{55}$$

which equals the right hand side of (52).

Notation

A	matrix
$\mathbf{a}$	dynamics of weights
$\mathbf{b}$	dynamics of abscissas
$\mathbf{c}$	continuous phase state vector
D	death rate
$\mathbf{D_c}$	medium exchange rate vector
d	dimension of $\mathbf{x}$
$\mathbf{F}$	exchange rate between dispersed and continuous phase
$\mathbf{G}$	vector of growth rates
g	growth rate
i_c	infected cell number density distribution
k	index
$k_\circ$	parameter
l	index
$m_{l_1,\ldots,l_d}(t)$	moment of order $l_1, \ldots, l_d$ with respect to n
$\mathcal{N}(\mu, \mathbf{\Sigma})$	Gaussian normal distribution
$n(t, \mathbf{x}), \widetilde{n}(t, \mathbf{x})$	number density distribution
n_0	initial number density distribution
N	overall number of particles
N_α	number of abscissas
N_d	number of dimensions
P	particle production rate
$\mathbf{P}$	integral coupling rate vector
t	time
U_c	concentration of uninfected cells
V	extracellular virus concentration
$\mathcal{W}$	set of cubature weights
w_α	cubature weight
$\mathcal{X}$	overall property state space

$\mathbf{x}_e$	vector of external coordinates
$\mathbf{x}_i$	vector of internal coordinates
$\mathbf{x}_\alpha$	cubature abscissa
$[\circ]$	concentration of intracellular compound $\circ$

Greek Symbols

α	index
$\widetilde{\alpha}$	tuning factor
β	tuning factor
κ	tuning factor
λ	tuning factor
μ	mean value vector
$\mathbf{\Sigma}$	covariance matrix

Abbreviations

$DQMOM$	direct quadrature method of moments
ODE	ordinary differential equation
PBE	population balance equation
PDE	partial differential equation
$QMOM$	quadrature method of moments

References

[1] M. Peglow, J. Kumar, R. Hampel, E. Tsotsas, S. Heinrich, Towards a Complete Population Balance Model for Fluidized-Bed Spray Agglomeration, Drying Technology 25 (7-8) (2007) 1321–1329.

[2] A. Vreman, C. van Lare, M. Hounslow, A basic population balance model for fluid bed spray granulation, Chemical Engineering Science 64 (21) (2009) 4389–4398.

[3] R. Radichkov, T. Müller, A. Kienle, S. Heinrich, M. Peglow, L. Mörl, A numerical bifurcation analysis of continuous fluidized bed spray granulation with external product classification, Chemical Engineering and Processing: Process Intensification 45 (10) (2006) 826 – 837.

[4] C. Borchert, K. Sundmacher, Morphology evolution of crystal populations: Modeling and observation analysis, Chemical Engineering Science 70 (0) (2012) 87 – 98.

[5] J. J. Liu, C. Y. Ma, Y. D. Hu, X. Z. Wang, Modelling protein crystallisation using morphological population balance models, Chemical Engineering Research and Design 88 (4) (2010) 437 – 446.

[6] D. L. Ma, D. K. Tafti, R. D. Braatz, Optimal control and simulation of multidimensional crystallization processes, Computers & Chemical Engineering 26 (78) (2002) 1103–1116.

[7] N. Hampel, A. Bück, M. Peglow, E. Tsotsas, Continuous pellet coating in a Wurster fluidized bed process, Chemical Engineering Science 86 (2013) 87–98, doi:http://dx.doi.org/10.1016/j.ces.2012.05.034.

[8] G. Y. Zhu, A. Zamamiri, M. A. Henson, M. A. Hjortso, Model predictive control of continuous yeast bioreactors using cell population balance models, Chemical Engineering Science 55 (24) (2000) 6155–6167.

[9] Y. Zhang, A. M. Zamamiri, M. A. Henson, M. A. Hjortso, Cell population models for bifurcation analysis and nonlinear control of continuous yeast bioreactors, Journal of Process Control 12 (6) (2002) 721–734.

[10] P. Mhaskar, M. A. Hjortso, M. A. Henson, Cell population modeling and parameter estimation for continuous cultures of Saccharomyces cerevisiae, Biotechnology Progress 18 (5) (2002) 1010–1026.

[11] A. Franz, H.-S. Song, D. Ramkrishna, A. Kienle, Experimental and theoretical analysis of poly(-hydroxybutyrate) formation and consumption in Ralstonia eutropha, Biochemical Engineering Journal 55 (1) (2011) 49 – 58.

[12] A. Franz, R. Dürr, A. Kienle, Population Balance Modeling of Biopolymer Production in Cellular Systems, in: Proceedings to 19th IFAC WC - Cape Town, 1705–1710, 2014.

[13] J. Schulze-Horsel, Y. Genzel, U. Reichl, Flow cytometric monitoring of influenza A virus infection in MDCK cells during vaccine production, BMC Biotechnology 8 (45).

[14] T. Müller, R. Dürr, B. Isken, J. Schulze-Horsel, U. Reichl, A. Kienle, Distributed modeling of human influenza a virus-host cell interactions during vaccine production, Biotechnology and Bioengineering 110 (8) (2013) 2252–2266.

[15] S. Müller, H. Harms, T. Bley, Origin and analysis of microbial population heterogeneity in bioprocesses, Current Opinion in Biotechnology 21 (1) (2010) 100 – 113.

[16] V. Chickarmane, C. Toein, U. A. Nuber, H. M. Sauro, C. Peterson, Transcriptional Dynamics of the Embryonic Stem Cell Switch, PLoS Computational Biology 2 (9) (2006) e123.

[17] M. Herberg, I. Roeder, Computational modelling of embryonic stem-cell fate control, Development 142 (13) (2015) 2250–2260.

[18] T. Eissing, H. Conzelmann, E. D. Gilles, F. Allgöwer, E. Bullinger, P. Scheurich, Bistability analyses of a caspase activation model for receptor-induced apoptosis, Journal of Biological Chemistry 279 (35) (2004) 36892–36897.

[19] H. Hulburt, S. Katz, Some problems in particle technology: A statistical mechanical formulation, Chemical Engineering Science 19 (8) (1964) 555 – 574.

[20] A. G. Fredrickson, D. Ramkrishna, H. M. Tsuchiya, Statistics and dynamics of procaryotic cell populations, Mathematical Biosciences 1 (3) (1967) 327–374.

[21] D. Ramkrishna, Population Balances: Theory and Applications to Particulate Systems in Engineering, Academic Press, San Diego, 2000.

[22] D. Ramkrishna, M. R. Singh, Population Balance Modeling: Current Status and Future Prospects, Annual Review of Chemical and Biomolecular Engineering 5 (2014) 123–146.

[23] R. Dürr, T. Müller, B. Isken, J. Schulze-Horsel, U. Reichl, A. Kienle, Distributed Modeling and Parameter Estimation of Influenza Virus Replication During Vaccine Production, in: Proceedings to 7th Vienna International Conference on Mathematical Modelling - MATHMOD 2012, 320–325, 2012.

[24] S. F. Heldt, T. Frensing, U. Reichl, Modeling the intracellular dynamics of influenza virus replication to understand the control of viral RNA synthesis, Journal of Virology 86 (15) (2012) 7806–7817.

[25] N. V. Mantzaris, P. Daoutidis, F. Srienc, Numerical solution of multi-variable cell population balance models: I. Finite difference methods, Computers & Chemical Engineering 25 (11-12) (2001) 1411 – 1440.

[26] N. V. Mantzaris, P. Daoutidis, F. Srienc, Numerical solution of multi-variable cell population balance models. III. Finite element methods, Computers & Chemical Engineering 25 (11-12) (2001) 1463 – 1481.

[27] V. John, C. Suciu, Direct discretizations of bi-variate population balance systems with finite difference schemes of different order, Chemical Engineering Science 106 (0) (2014) 39 – 52.

[28] M. N. Nandanwar, S. Kumar, A new discretization of space for the solution of multi-dimensional population balance equations, Chemical Engineering Science 63 (8) (2008) 2198 – 2210.

[29] R. Gunawan, I. Fusman, R. D. Braatz, Parallel high-resolution finite volume simulation of particulate processes, AIChE Journal 54 (6) (2008) 1449–1458.

[30] L. N. Trefethen, Spectral Methods in MATLAB, SIAM, ISBN 0-89871-465-6, 2001.

[31] A. Bück, G. Klaunick, J. Kumar, M. Peglow, E. Tsotsas, Numerical Simulation of Particulate Processes for Control and Estimation by Spectral Methods, AIChE Journal 83 (2012) 2309–2319.

[32] V. John, I. Angelov, A. A. Öncül, D. Thvenin, Techniques for the reconstruction of a distribution from a finite number of its moments, Chemical Engineering Science 62 (11) (2007) 2890 – 2904.

[33] R. McGraw, Description of aerosol dynamics by the quadrature method of moments, Aerosol Science and Technology 27 (2) (1997) 255–265.

[34] E. Aamir, Z. K. Nagy, C. D. Rielly, T. Kleinert, B. Judat, Combined Quadrature Method of Moments and Method of Characteristics Approach for Efficient Solution of Population Balance Models for Dynamic Modeling and Crystal Size Distribution Control of Crystallization Processes, Industrial & Engineering Chemistry Research 48 (18) (2009) 8575–8584.

[35] D. Wright, R. McGraw, D. Rosner, Bivariate Extension of the Quadrature Method of Moments for Modeling Simultaneous Coagulation and Sintering of Particle Populations, Journal of Colloid and Interface Science 236 (2) (2001) 242 – 251.

[36] C. Yoon, R. McGraw, Representation of generally mixed multivariate aerosols by the quadrature method of moments: I. Statistical foundation, Journal of Aerosol Science 35 (5) (2004) 561 – 576.

[37] C. Yoon, R. McGraw, Representation of generally mixed multivariate aerosols by the quadrature method of moments: II. Aerosol dynamics, Journal of Aerosol Science 35 (5) (2004) 577 – 598.

[38] A. Buffo, D. L. Marchisio, Modeling and simulation of turbulent polydisperse gas-liquid systems via the generalized population balance equation, Reviews in Chemical Engineering 30 (1) (2014) 73–126.

[39] D. L. Marchisio, R. O. Fox, Solution of population balance equations using the direct quadrature method of moments, Journal of Aerosol Science 36 (1) (2005) 43 – 73.

[40] A. Zucca, D. L. Marchisio, M. Vanni, A. A. Barresi, Validation of bivariate DQMOM for nanoparticle processes simulation, AIChE Journal 53 (4) (2007) 918–931.

[41] R. Dürr, A. Kienle, An Efficient Method for Calculating the Moments of Multidimensional Growth Processes in Population Balance Systems, The Canadian Journal of Chemical Engineering 92 (2014) 2088–2097.

[42] J. Villadsen, J. Nielsen, N. Liden, Bioreaction Engineering Principles, Springer Science & Business Media, Heidelberg, 3 edn., 2011.

[43] D. L. Marchisio, Quadrature Method of Moments for Poly-Disperse Flows, in: D. L. Marchisio, R. O. Fox (Eds.), Multiphase Reacting Flows: Modelling and Simulation, vol. 492, Springer Vienna, 41–77, 2007.

[44] R. Dürr, Parameter estimation and method of moments for multi dimensional population balance equations with application to vaccine production processes, Ph.D. thesis, Otto-von-Guericke University Magdeburg, 2016, submitted.

[45] H.-K. Rhee, R. Aris, N. R. Amundson, First-Order Partial Differential Equations, Volume 1: Theory and Application of Single Equations, Dover Publications, Inc., Mineola, New York, 1986.

[46] P. J. Davis, P. Rabinowitz, Methods of Numerical Integration, Computer Science and Applied Mathematics, Academic Press, 2. ed., 4. print edn., 1994.

[47] A. Stroud, Approximate Calculation of Multiple Integrals, Prentice-Hall, Englewood Cliffs, N.J., Englewood Cliffs, N.J., 1971.

[48] U. Lerner, Hybrid Bayesian Networks for Reasoning about Complex Systems, Ph.D. thesis, Stanford University, 2002.

[49] S. J. Julier, J. K. Uhlmann, A Consistent, Debiased Method for Converting Between Polar and Cartesian Coordinate Systems, in: In The Proceedings of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls, 110–121, 1997.

[50] R. van der Merwe, Sigma-Point Kalman Filters for Probablilistic Inference in Dynamic State-Space Models, Ph.D. thesis, Oregon Health & Science University, 2004.

[51] S. Julier, J. Uhlmann, Unscented filtering and nonlinear estimation, Proceedings of the IEEE 92 (3) (2004) 401 – 422.

[52] R. Schenkendorf, Optimal experimental design for parameter identification and model selection, Ph.D. thesis, Otto-von-Guericke University Magdeburg, 2014.

[53] N. Rossner, T. Heine, R. King, Quality-by-Design using a Gaussian Mixture Density Approximation of Biological Uncertainties, in: Proceedings of the 11th International Symposium on Computer Applications in Biotechnology, vol. 11, 7–12, 2010.

[54] E. L. Haseltine, J. B. Rawlings, J. Yin, Dynamics of viral infections: incorporating both the intracellular and extracellular levels, Computers & Chemical Engineering 29 (3) (2005) 675 – 686.

[55] E. L. Haseltine, J. Yin, J. B. Rawlings, Implications of Decoupling the Intracellular and Extracellular Levels in Multi-Level Models of Virus Growth, Biotechnology and Bioengineering 101 (4) (2008) 811–820.

[56] R. Dürr, T. Müller, A. Kienle, Efficient DQMOM for Multivariate Population Balance Equations and Application to Virus Replication in Cell Cultures, in: 8th Vienna International Conference on Mathematical Modelling, 29–34, 2015.

[57] E. Sherer, R. Hannemann, A. Rundell, D. Ramkrishna, Application of stochastic equations of population balances to sterilization processes, Chemical Engineering Science 64 (4) (2009) 764–774.

[58] A. Reinhold, H. Briesen, High dimensional population balances for the growth of faceted crystals: Combining Monte Carlo integral estimates and the method of characteristics, Chemical Engineering Science 127 (2015) 220–229.

[59] T. Eissing, M. Chaves, F. Allgöwer, Live and let die–A systems biology view on cell death, Computers & Chemical Engineering 33 (3) (2009) 583–589.

[60] J. E. Nichols, J. W. LeDuc, Influenza, in: A. D. T. Barrett, L. R. Stanberry (Eds.), Vaccines for biodefense and emerging and neglected Diseases, Academic Press, London, 497–525, 2009.

[61] Y. Genzel, U. Reichl, Continuous cell lines as a production system for influenza vaccines, Expert Rev Vaccines 8 (12) (2009) 1681–1692.

[62] F. Bushman, M. Lewinski, A. Ciuffi, S. Barr, J. Leipzig, S. Hannenhalli, C. Hoffmann, Genome-wide analysis of retroviral DNA integration, Nature Reviews Microbiology 3 (11) (2005) 848–858.

[63] F. S. Heldt, T. Frensing, A. Pflugmacher, R. Gröpler, B. Peschel, U. Reichl, Multiscale Modeling of Influenza A Virus Infection Supports the Development of Direct-Acting Antivirals, PLoS Computational Biology 9 (11).

[64] S. A. Stewart, D. M. Dykxhoorn, D. Palliser, H. Mizuno, E. Y. Yu, D. S. An, D. M. Sabatini, I. S. Y. Chen, W. C. Hahn, P. A. Sharp, R. A. Weinberg, C. D. Novina, Lentivirus-delivered stable gene silencing by RNAi in primary cells, RNA (New York, N.Y.) 9 (4) (2003) 493–501.

[65] B. R. McNaughton, J. J. Cronican, D. B. Thompson, D. R. Liu, Mammalian cell penetration, siRNA transfection, and DNA transfection by supercharged proteins, Proceedings of the National Academy of Sciences of the United States of America 106 (15) (2009) 6111–6116.

[66] S. Kumar, D. Ramkrishna, On the solution of population balance equations by discretization - II. A moving pivot technique, Chemical Engineering Science 51 (8) (1996) 1333–1342.