

# Reducing Stochastic Discrete Models of Biochemical Networks

Samaneh Gholami, Silvana Ilie

Department of Mathematics, Ryerson University, Toronto, Canada

Email: silvana@ryerson.ca

**How to cite this paper:** Gholami, S. and Ilie, S. (2021) Reducing Stochastic Discrete Models of Biochemical Networks. *Applied Mathematics*, 12, 449-469.  
<https://doi.org/10.4236/am.2021.125031>

**Received:** April 21, 2021

**Accepted:** May 28, 2021

**Published:** May 31, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Biochemical systems have numerous practical applications, in particular to the study of critical intracellular processes. Frequently, biochemical kinetic models depict cellular processes as systems of chemical reactions. Many biological processes in a cell are inherently stochastic, due to the existence of some low molecular amounts. These stochastic fluctuations may have a great effect on the biochemical system's behaviour. In such cases, stochastic models are necessary to accurately describe the system's dynamics. Biochemical systems at the cellular level may entail many species or reactions and their mathematical models may be non-linear and with multiple scales in time. In this work, we provide a numerical technique for simplifying stochastic discrete models of well-stirred biochemical systems, which ensures that the main properties of the original system are preserved. The proposed technique employs sensitivity analysis and requires solving an optimization problem. The numerical tests on several models of practical interest show that our model reduction strategy performs very well.

## Keywords

Stochastic Simulation Algorithm, Stochastic Biochemical Kinetics, Sensitivity Analysis, Model Reduction Methods

## 1. Introduction

Modelling and simulation of cellular processes are subjects of significant interest in fields such as Computational and Systems Biology. Biological processes at the level of a single cell are commonly portrayed as systems of biochemical reactions. The dynamics of various biochemically reacting systems may exhibit random fluctuations, due to some molecular species which exist in low amounts. These random fluctuations could have critical implications in biology [1] [2]. Then,

stochastic models are essential for an accurate description of the system's dynamics. An extensively used stochastic model of well-stirred biochemical systems is the Chemical Master Equation (CME) [3]. This model represents the state of the system as a Markov process. An exact stochastic simulation algorithm for the Chemical Master Equation was proposed by Gillespie [4] [5].

Biochemical processes at the cellular level may entail many biochemical species undergoing a significant number of reactions. The mathematical models of such biochemical networks have high dimension, and thus their numerical simulation is quite demanding. Moreover, numerous biochemical systems evolve on multiple time scales, leading to stiffness. Stiffness is a challenge for numerical simulations. A large number of the species and/or reactions of some biochemical systems encountered in applications imply that their mathematical models have a significant number of parameters. Usually, these parameters cannot be measured accurately. For many biochemical networks, it is difficult to determine the parts of the system which are critical in determining their behaviour. Therefore, it is critical to design accurate and robust strategies to simplify these complex biochemical systems, which maintain the key properties of the original network. The reduced models will be easier to analyze and simulate numerically. In addition, they can be utilized to predict and control the behaviour of the system. Furthermore, the simplified models have a lower number of parameters, which thus become easier to determine. Reduction techniques for deterministic models of biochemical systems include species and reaction lumping methods [6] [7], sensitivity analysis based schemes [8] and time-scale analysis based strategies [9] [10]. Lumping schemes cause a loss of the physical interpretation of the elementary reactions and of information on some species or reactions. These schemes are useful when poor data exist about certain reactions. Time-scale analysis applies to systems with reactions that can be partitioned into fast and slow events, for which the fast components are in a quasi-steady state. For complex networks of biochemical reactions, finding the constraints for the fast dynamics may be a challenging task. Sensitivity analysis examines the dependence of the system's behaviour on its parameters. In particular, local sensitivity analysis of models of biochemical systems measures the variation of the system's state with small perturbations in model's parameters [11]. Sensitivity analysis makes it possible to identify and discard the reactions which are not essential, those with a negligible parametric sensitivity with respect to their rate parameters [8]. Few reduction techniques for stochastic models of biochemical systems exist in the literature. These include [12] [13] [14] for discrete models and [15] [16] for continuous ones (see also references therein).

This paper proposes a new technique for reducing the complexity of stochastic discrete models of homogeneous biochemical networks. The discrete stochastic model under consideration is the Chemical Master Equation. The Chemical Master Equation accurately describes the dynamics of a wide range of well-stirred, realistic biochemical systems. The reduced reaction mechanisms generated with

the new technique are easier to understand and manipulate and have fewer parameters.

Local sensitivity analysis quantifies the variations in the system's behaviour caused by small changes in its parameters. Our method relies on estimating (local) parametric sensitivities of the Chemical Master Equation model. These parametric sensitivities are approximated using the Coupled Finite Difference (CFD) scheme [17] and the Common Random Number (CRN) method [18]. Sensitivity analysis is performed to detect which parameters, and thus which reactions, are good candidates for elimination. It should be noted that local sensitivity analysis alone may not be a reliable tool for model reduction. It may indicate that a fast reaction can be eliminated, although that reaction is important. Elimination of a fast reaction means that the value of its rate parameter changes significantly. Nevertheless, local sensitivity analysis provides no information on the perturbations in the system's dynamics induced by large variations in its parameters. To correct this, local sensitivity analysis is coupled with a global approach to model reduction, which requires solving an optimization problem. For achieving this, we shall apply a strategy akin to that proposed in [8] for reducing continuous deterministic models of chemical reactions. This strategy ensures that the simplified biochemical system retains the stability and nonlinear properties of the original network.

The paper is organized as follows. In Section 2, we present the background on stochastic discrete models of homogeneous biochemical systems, and methods to simulate and estimate parametric sensitivities for these models. In Section 3, we propose a new model reduction strategy for the Chemical Master Equation. The advantages of the proposed strategy are illustrated on three critical models arising in applications, namely the infection, the epidermal growth factor receptor signalling pathway and the gemcitabine biochemical systems, in Section 4.

## 2. Background

### 2.1. Chemical Master Equation

The evolution in time of homogeneous biochemical systems is governed by the Chemical Master Equation. This model has been successfully utilized to study vital biological processes, such as gene expression and regulation [1] [2].

Consider  $N$  biochemical species  $S_1, \dots, S_N$  interacting through  $M$  chemical reactions,  $R_1, \dots, R_M$ . Assume that the system is homogeneous, and at constant volume and thermal equilibrium. It is described by the state vector:

$$\mathbf{X}(t) = [X_1(t), X_2(t), \dots, X_N(t)]^T$$

where  $X_i(t)$  indicates the amount of  $S_i$  molecules at time  $t$ . The biochemical system state,  $\mathbf{X}(t)$ , is a Markov process continuous in time and discrete in space. To each reaction  $R_j$  it corresponds a state change vector

$\mathbf{v}_j \equiv (v_{1j}, \dots, v_{Nj})^T$ , with  $v_{ij}$  representing the number of  $S_i$  molecules consumed or produced when a reaction  $R_j$  fires. The matrix  $\mathbf{v} = \{v_{ij}\}_{1 \leq i \leq N, 1 \leq j \leq M}$ ,

where  $N$  is the number of reactants and  $M$  is the number of reactions, represents the stoichiometric matrix. In addition, a propensity function  $a_j(\mathbf{x})$  is used for describing a reaction  $R_j$ . By definition,  $a_j(\mathbf{x})dt$  is the probability of one reaction  $R_j$  firing during  $[t, t + dt]$ , provided that at time  $t$  the system was in state  $\mathbf{x}$ .

Each propensity function follows the mass-action kinetics; thus, for the first order reaction  $S_m \xrightarrow{c_j} \text{products}$ , it is  $a_j(\mathbf{X}(t)) = c_j X_m(t)$ , and for the second order reaction  $S_m + S_n \xrightarrow{c_j} \text{products}$ , with  $m \neq n$ , the propensity is  $a_j(\mathbf{X}(t)) = c_j X_m(t) X_n(t)$ . Lastly, the propensity of  $S_m + S_m \xrightarrow{c_j} \text{products}$  may be expressed as  $a_j(\mathbf{X}(t)) = c_j X_m(t)(X_m(t) - 1)/2$ .

For  $t \geq t_0$ , let  $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$  be the conditional probability that the state vector at time  $t$  is  $\mathbf{X}(t) = \mathbf{x}$ , given that at time  $t_0$  it was  $\mathbf{X}(t_0) = \mathbf{x}_0$ . Then,  $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$  satisfies the Chemical Master Equation [3]

$$\frac{\partial P(\mathbf{x}, t | \mathbf{x}_0, t_0)}{\partial t} = \sum_{j=1}^M \left[ a_j(\mathbf{x} - \mathbf{v}_j) P(\mathbf{x} - \mathbf{v}_j, t | \mathbf{x}_0, t_0) - a_j(\mathbf{x}) P(\mathbf{x}, t | \mathbf{x}_0, t_0) \right], \quad (1)$$

where  $\mathbf{v}_j$  is the state-change vector for reaction  $R_j$ . The Chemical Master Equation is a refined stochastic discrete model of homogeneous biochemical systems. It is worth mentioning that, generally, the CME is a high dimensional model. Exact Monte Carlo techniques for the Chemical Master Equation model were developed in the literature. They include Gillespie's algorithm [4] and the Random Time-Change (RTC) algorithm [17] [18] based on the RTC representation of Kurtz [19].

## 2.2. Gillespie's Algorithm

The Chemical Master Equation (1) is computationally intractable for the majority of biochemical systems encountered in applications. To deal with this difficulty, Gillespie [4] [5] proposed an exact Monte Carlo strategy for the Chemical Master Equation, referred to as the stochastic simulation algorithm (SSA).

Gillespie's algorithm proceeds as follows:

- 1) Initialize  $t \leftarrow t_0$ ,  $\mathbf{X}(t_0) \leftarrow \mathbf{x}_0$ .
- 2) At time  $t$ , evaluate

$$a_{sum}(\mathbf{X}(t)) := \sum_{k=1}^M a_k(\mathbf{X}(t)).$$

- 3) Draw a unit uniform random number  $r_1$ , and take

$$\tau = \left\lceil 1/a_{sum}(\mathbf{X}(t)) \right\rceil \ln(1/r_1).$$

- 4) Draw a unit uniform random number  $r_2$  and denote by  $j$  the smallest integer satisfying

$$\sum_{k=1}^j a_k(\mathbf{X}(t)) > r_2 a_{sum}(\mathbf{X}(t)).$$

- 5) Update  $\mathbf{X}(t + \tau) \leftarrow \mathbf{X}(t) + \mathbf{v}_j$  and  $t \leftarrow t + \tau$ .
- 6) Return to step 2 or else stop.

### 2.3. Sensitivity Analysis

Sensitivity analysis is a key tool for investigating the robustness properties of a model with respect to perturbations in its parameters, for estimating these parameters or for guiding model reduction. Models of biochemically reacting systems depend on several parameters, such as the reaction rate parameters and the initial conditions. In many instances, the values of several such parameters are not available or cannot be evaluated accurately. For this reason, it is critical to investigate how variations in model parameters influence the behaviour of the system. Local sensitivity analysis examines how the system dynamics is influenced by small changes in model parameters. A model is sensitive with respect to one of its parameters if the behaviour of the system varies significantly with minor perturbations in that parameter; else it is robust with respect to variations in that parameter. Parametric sensitivities help determine which components of the biochemical network control the model's behaviour. If a reaction is such that some parametric sensitivities with respect to its kinetic parameter are large, then it is considered important.

For a deterministic system, the parametric sensitivity of the state vector,  $\mathbf{X}(t, c)$ , with respect to a parameter  $c$  can be computed by  $\mathbf{S} = \partial \mathbf{X}(t, c) / \partial c$ . Estimating parametric sensitivities of a stochastic discrete model can be a daunting task. In this work, we employ finite-difference approximations of the local parametric sensitivities for the Chemical Master Equation model. The sample paths necessary for the finite-difference estimations are generated with Monte Carlo simulations. In the discrete stochastic setting, if  $\mathbb{E}$  is the expected value,  $\mathbf{f}$  is a function of interest,  $\mathbf{X}(t, c)$  represents the system state at time  $t$ , corresponding to the parameter  $c$ , then a finite-difference estimation of the sensitivity of  $\mathbb{E}(\mathbf{f}(\mathbf{X}(t, c)))$  with respect to  $c$  is

$$\mathbf{S} \approx \frac{\mathbb{E}[\mathbf{f}(\mathbf{X}(t, c+h))] - \mathbb{E}[\mathbf{f}(\mathbf{X}(t, c))]}{h}. \quad (2)$$

Here  $h$  is a small perturbation of the parameter  $c$ . In what follows,  $c$  represents a kinetic rate parameter.

Let  $\mathbf{X}^{(n)}(t, c+h)$  and  $\mathbf{X}^{(n)}(t, c)$  be trajectories generated by Monte Carlo simulations, for the parameters  $c+h$  and  $c$ , respectively. Here  $n$  is the trajectory index ( $1 \leq n \leq L$ ) and  $L$  is the total number of trajectories. In this case, a finite-difference numerical approximation of the sensitivity, utilizing Monte Carlo simulations, is

$$S \approx \frac{1}{L} \sum_{n=1}^L \frac{1}{h} [\mathbf{f}(\mathbf{X}^{(n)}(t, c+h)) - \mathbf{f}(\mathbf{X}^{(n)}(t, c))]. \quad (3)$$

It is widely known that variance reduction strategies allow more precise estimations of the sensitivity, with a smaller numbers of simulations. To reduce the variance of the estimator, finite-difference sensitivity estimators couple the perturbed and unperturbed trajectories. Among the most effective and accurate finite-difference parametric estimators for stochastic discrete models of biochem-

ical systems are the Common Random Number (CRN) method due to Rathinam *et al.* [18] and the Coupled Finite-Difference (CFD) scheme proposed by Anderson [17]. For further information on the CFD method, using the random time change representation [19] of the Markov process  $\mathbf{X}(t)$ , we refer the reader to [17]. The CRN method is presented below. Notice that the CRN and CFD techniques are based on exact Monte Carlo simulation schemes for the Chemical Master Equation.

**Common Random Number (CRN)** The common random number technique is utilized for estimating parametric sensitivities of stochastic discrete biochemical kinetic models. The CRN estimator has a low variance [18]. To estimate the sensitivity of the mean  $\mathbb{E}(\mathbf{X}(t, c))$ , with respect to a parameter  $c$ , two systems will be considered: the unperturbed system  $\mathbf{X}(t, c)$  and the perturbed counterpart  $\mathbf{X}(t, c+h)$ . The CRN method applies Gillespie's algorithm with the same string of common of random numbers to generate the coupled perturbed and unperturbed paths, corresponding to the parameters  $c+h$  and  $c$ , respectively. The use of a common random number sequence reduces the variance of the estimator, thus increases its accuracy for the same number of sample paths. The CRN algorithm consists of the following steps:

- 1) Take  $c = c_j$ . For each  $n \leq L$ , with  $L$  being the number of sample paths,
- 2) Generate an SSA path for the unperturbed system (for parameter  $c$ ), with a sequence of independent uniform (0, 1) random numbers,  $\mathbf{X}^{(n)}(t, c)$ .
- 3) Generate an SSA path for the perturbed system (for parameter  $c+h$ ), with the same sequence of independent uniform (0, 1) random number,  $\mathbf{X}^{(n)}(t, c+h)$ .
- 4) Calculate  $\mathbf{s}^{(n)}(t) = \left( \mathbf{f}(\mathbf{X}^{(n)}(t, c+h)) - \mathbf{f}(\mathbf{X}^{(n)}(t, c)) \right) / h$ .
- 5) End loop over  $n$ .
- 6) Compute the mean of  $\left\{ \mathbf{s}^{(n)}(t) \right\}_{n \leq L}$ .

Remark that the perturbation  $h$  should be small enough to decrease the truncation error associated with finite-difference approximations of derivatives, but large enough to increase the convergence rate of the sensitivity estimation [17]. The CRN algorithm is a popular technique for sensitivity estimation, being efficient and easy to implement. The parametric sensitivity estimators discussed above is utilized to simplify complex stochastic models of homogeneous biochemical networks. Alternatively, a finite-difference scheme may be employed for estimating sensitivities by approximate Monte Carlo simulation techniques [20] [21], such as adaptive tau-leaping methods; these strategies are efficient on moderately stiff to very stiff models.

### 3. Model Reduction

In this section, we introduce a new procedure to simplify stochastic discrete biochemical networks, which are modelled with the Chemical Master Equation. The goal of this procedure is to determine a reduced biochemical reaction system which retains the essential features of the full system, such as its stability features, nonlinear behaviour and physical interpretation of the elementary reac-

tions. The reduced biochemical model will be faster to solve numerically and easier to investigate and understand. Local sensitivity analysis is applied to uncover the unimportant reactions of the full biochemical network. Also, it seeks to find the essential reactions, which should be retained in a simplified model. Further, our strategy extends the global approach for continuous deterministic chemical system reduction proposed in [8]. Remark that the reduced network generated with the procedure presented below is expected to perform well over a variety of parameter values, in a neighbourhood of the values for which it was obtained.

Initially, a local sensitivity analysis is performed for the stochastic discrete model of the original biochemical system. Of interest are the kinetic parameters,  $c = c_i$ . Finite-difference methods, such as the CFD or the CRN algorithms, are employed to estimate the local parametric sensitivities. This allows a *preselection* of the crucial reactions in the network, those for which the molecular amounts of all, or of the important species, are very sensitive with respect to small variations in the associated rate parameter. Let  $\{R_i\}_{i \in I_{cr}}$ , for some subset  $I_{cr}$  of indexes  $\{1, \dots, M\}$ , be these crucial reactions. These reaction channels should be kept in the system. Moreover, the sensitivity analysis also helps determine the reactions for which small perturbations in their rate parameters cause insignificant changes in the overall behaviour of the system. More precisely, the kinetic parameter  $c_i$  of such a reaction satisfies

$$\left| \frac{c_i}{\mathbb{E}(X_k(t, c_i))} \frac{\partial}{\partial c_i} \mathbb{E}(X_k(t, c_i)) \right| \ll 1,$$

for all the quantities of interest  $X_k$ . These reactions could be unimportant, and may be considered for deletion.

Stability is a desired property of a simplified reaction network and a reduction technique based on local sensitivity analysis of the original system alone cannot guarantee it. As a consequence, we shall use a global approach to ensure that our strategy leads to a stable reduced system. For this, we consider the following *continuous optimization problem* for deterministic models (see also [8]),

$$\begin{aligned} & \min \|\mathbf{x} - \mathbf{z}\| \text{ such that} \\ & \mathbf{x}' = \mathbf{v}\mathbf{F}(\mathbf{x}), \quad \mathbf{x}(0) = \mathbf{x}_0 \\ & \mathbf{z}' = \mathbf{v}\mathbf{D}\mathbf{F}(\mathbf{z}), \quad \mathbf{z}(0) = \mathbf{x}_0, \quad 0 \leq t \leq T \\ & \ell \leq \sum_{i=1}^M d_i \leq u, \quad 0 \leq d_i \leq 1, \quad \text{for } 1 \leq i \leq M, \\ & g(d_1, \dots, d_M) \leq r. \end{aligned} \tag{4}$$

Note that  $\mathbf{x}$  and  $\mathbf{z}$  are the state vectors of the original and the reduced biochemical systems, respectively, modelled with the reaction rate equations. Also,  $\mathbf{v}$  is the stoichiometric matrix,  $\mathbf{F}(\cdot)$  represents the vector of propensities and  $\mathbf{D}$  is a diagonal matrix, with  $d_i$  being the  $i$ -th row diagonal entry. Observe that  $d_i = 1$  if the reaction is kept in the reduced system and  $d_i = 0$ , if it is removed.

Here  $\ell$  and  $u$  are the simplified model's lower and upper bounds for the total reaction count, and  $r > 0$  is a small constant. Also,  $g$  denotes a non-linear function which becomes 0 when  $d_i$  are either 0 or 1, for all  $1 \leq i \leq M$ . We consider the following nonlinear function

$$g = \sum_{i=1}^M (d_i - d_i^2)^\beta, \quad (5)$$

with  $\beta \geq 2$  being a parameter. In our implementation,  $\beta$  is set to 2. By relaxing the condition

$$g(d_1, \dots, d_M) = 0 \quad \text{to} \quad g(d_1, \dots, d_M) \leq r,$$

the discrete optimization is simplified to a continuous optimization problem. Once a solution of the continuous optimization problem is found, the coefficients  $d_i$  are rounded to 0 if they are very small; otherwise they are rounded to 1. This result may yield a local rather than a global minimum. Still, for models with many reactions, *i.e.*, a large parameter space, a local minimum would be considered a satisfactory solution for the simplifying strategy. Also, the choice of the parameters (e.g.,  $u, r$ ), would affect how closely the dynamics of the reduced system matches that of the original version. There is a trade-off between improving the accuracy of the simplified mechanism and eliminating many reactions. To obtain a simpler optimization problem, after we identified the important reactions,  $\{R_i\}_{i \in I_{cr}}$ , and we assigned their  $d_i = 1$ , we solve (4) over a smaller set of parameters,  $\{d_j\}_{j \notin I_{cr}}$ .

Remark that the optimization problem is formulated in a continuous deterministic framework. Empirically, we found that posing an optimization problem for the reaction rate equations, rather than for the much more challenging Chemical Master Equation model, yields an accurate reduction mechanism for biochemical systems with a low to moderate level of noise. Moreover, this procedure can handle efficiently large biochemical networks as well.

For solving the optimization problem (4), we utilize GEKKO [22], an optimizer toolbox. The modes of operation of GEKKO include parameter regression, data reconciliation, real-time optimization, dynamic simulation, and nonlinear predictive control. For our reduction procedure, we apply dynamic estimation with the APOPT solver. After the optimization problem is solved, we isolate the coefficients  $d_i$  which are very small and set them to 0. This means that the associated reactions will be deleted from the biochemical system. The remaining coefficients  $d_i$  are rounded to 1. In our experiments, we observed that a good number of unimportant reactions identified with the optimization approach are among those selected by the local sensitivity analysis as candidates for elimination. The results are validated by comparing the behaviours of the reduced and full models.

## 4. Numerical Examples

In this section, we illustrate the advantages of the proposed model reduction

technique on three systems of biochemical reactions: one is a low size but key test problem, while the other two are high-dimension biochemical networks, with applications to cancer research. In each case, sensitivity analysis is performed to identify which parts of the network do not play a significant role in its overall behaviour. More importantly, the analysis finds the reactions which are critical for the dynamics of the biochemical system and set their  $d_i = 1$ . After this pre-selection, the global optimization strategy is employed to select the reactions which can be removed from the system. Finally, the reduced and full models are simulated with Gillespie's algorithm over a large ensemble of trajectories. The results are used for comparing the evolution of the mean and standard deviation of the molecular amounts for each species, in the original and simplified models.

#### 4.1. Infectious Disease Model

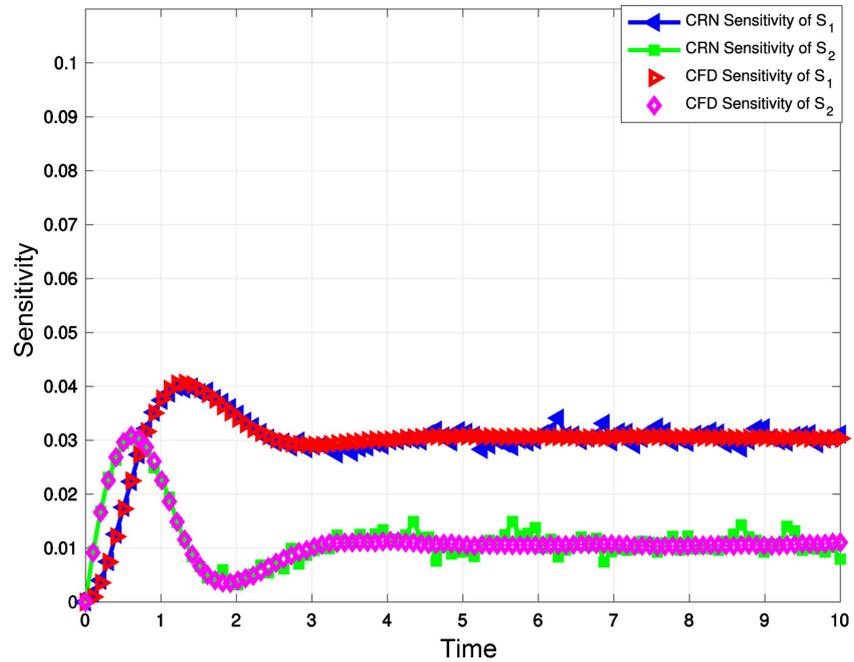
The infectious disease model is a simple example of a two species reaction network [23]. The reaction channels of the infectious disease model and their rate parameters are given in **Table 1**. The model is simulated on the interval  $[0,10]$  with initial conditions  $S_1(0) = 20$  and  $S_2(0) = 40$ .

The parametric sensitivities with respect to  $c_2$  are estimated using the CRN and the CFD methods on 10,000 sample paths. **Figure 1** depicts the evolution in time of the fully normalized sensitivities of the species  $S_1$  and  $S_2$ , with respect to the kinetic parameter  $c_2$ . The approximations obtained with the two finite-difference sensitivity estimators, the CFD and the CRN algorithms, are in good agreement. We note that, for  $i = 1$  and 2, the normalized sensitivities with respect to  $c_2$  are very small. Thus, reaction  $R_2$  is a good candidate for elimination, and this is validated by solving the optimization problem.

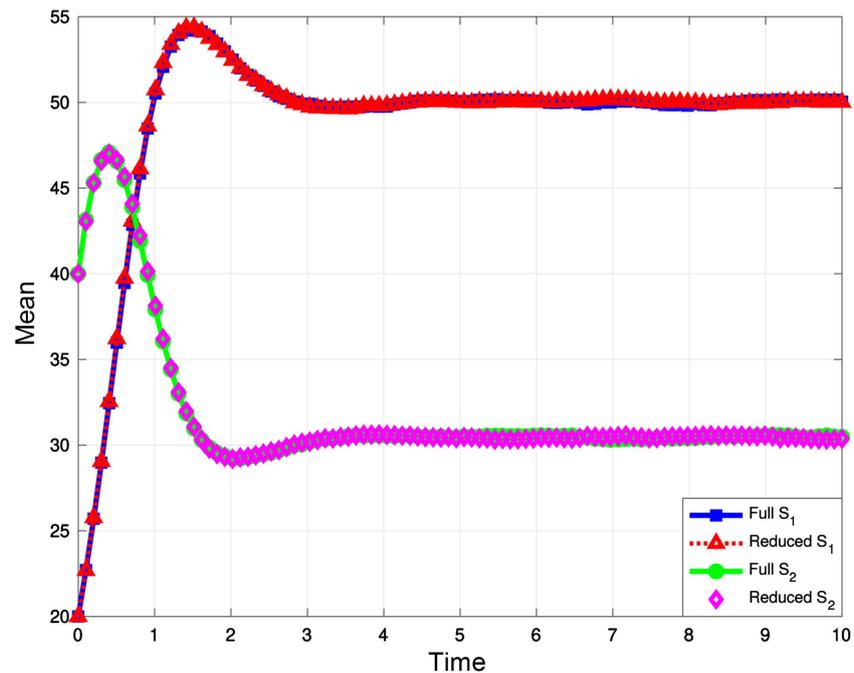
The full and reduced (without reaction  $R_2$ ) models are simulated with the SSA on 10,000 trajectories. The means and standard deviations of the molecular counts of  $S_1$  and  $S_2$  as functions of time, for the full and reduced models, are plotted in **Figure 2** and **Figure 3**, respectively. The match between the numerical results for the reduced, in which reaction  $R_2$  is deleted, and the full models is excellent in each case.

**Table 1.** The infectious disease model: the reaction channels and their rate parameters.

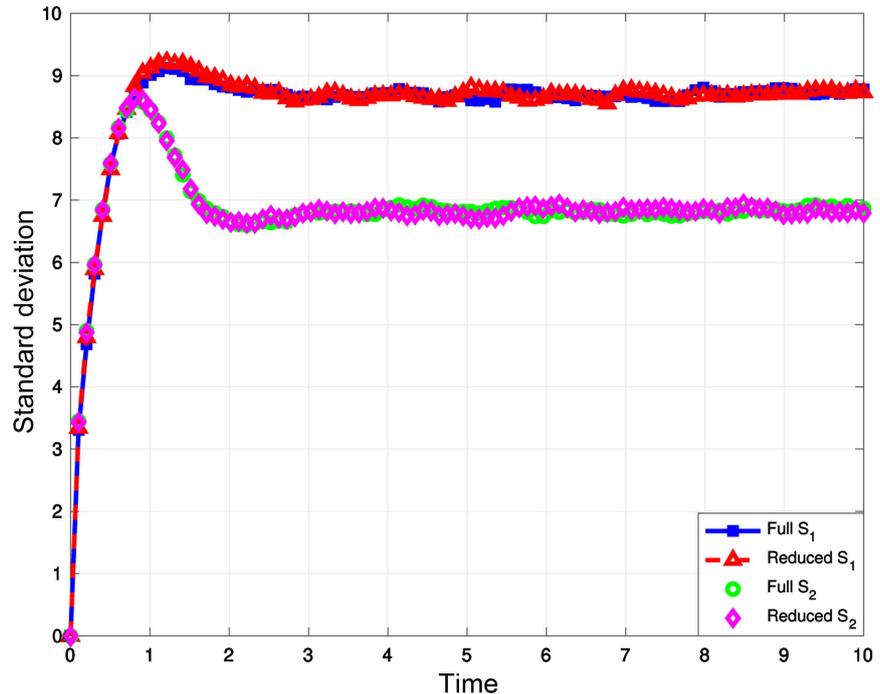
	Reaction channel	Reaction rate
$R_1$	$S_1 \rightarrow \emptyset$	$c_1 = 2.0$
$R_2$	$S_2 \rightarrow \emptyset$	$c_2 = 0.1$
$R_3$	$\emptyset \rightarrow S_1$	$c_3 = 25$
$R_4$	$\emptyset \rightarrow S_2$	$c_4 = 75$
$R_5$	$S_1 + S_2 \rightarrow S_1 + S_1$	$c_5 = 0.05$



**Figure 1.** The infectious disease model: the evolution in time of the normalized sensitivities for  $S_1$  and  $S_2$  with respect to  $c_2$ , namely  $\left| \frac{c_2}{\mathbb{E}(X_i(t))} \frac{\partial}{\partial c_2} \mathbb{E}(X_i(t)) \right|$ , for  $i = 1, 2$ . The simulations are done over 10,000 paths, with the CRN and the CDF schemes. The integration interval is  $[0, 10]$ .



**Figure 2.** The infectious disease model: the comparison of the evolution in time of the means for the molecular amounts of the species  $S_1$  and  $S_2$ , for the full system and the reduced one, in which the reaction  $R_2$  is eliminated. The simulations are over 10,000 paths and the integration interval is  $[0, 10]$ .



**Figure 3.** The infectious disease model: the comparison of the evolution in time of the standard deviations of the molecular amounts of the species  $S_1$  and  $S_2$ , for the full system and the reduced one, in which the reaction  $R_2$  is eliminated. The simulations are over 10,000 paths and the integration interval is  $[0,10]$ .

## 4.2. Epidermal Growth Factor Receptor (EGFR) Model

The epidermal growth factor receptor signalling pathway participates in cell differentiation and proliferation. The role of the epidermal growth factor in cancer cell proliferation has recently been the subject of intense research. The biochemical reaction network which models the EGFR signalling pathway consists of 23 molecular species undergoing 47 reaction channels [24].

The EGFR model reactions and their rate parameter values are given in **Table 2**. The molecular species and their initial molecular counts are listed in **Table 3**. Assume a cellular volume of  $3 \times 10^{-12}$  liters, with a nano-Mole concentration = 1800 (*i.e.*,  $6.023 \times 10^{+23} \times 3.00 \times 10^{-12} \times 10^{-9}$ ) molecules per cell. For computational purposes, we shall consider molecular amounts rather than concentrations of the species. This requires rescaling the rate parameters. The problem is solved numerically on the time interval  $[0,100]$ . Remark that the EGFR model is stiff, given that it spans multiple time-scales; indeed, some reactions are slow others are fast. Stiffness is a major challenge for numerical simulation [25].

Initially, we use finite-difference approximations of the sensitivity by employing the CRN scheme with 2000 trajectories. We identify a set of reactions with small normalized parametric sensitivities for all molecular species. A sample plot of the evolution in time of the normalized sensitivities of various species with respect to one of the parameters,  $c_{21}$ , is displayed in **Figure 4**. The set identified at this step constitutes an initial selection of reactions which are good candidates

**Table 2.** The epidermal growth factor receptor (EGFR) signaling pathway model: the reaction channels and the rate parameters. First- and second-order rate parameter values are expressed in  $s^{-1}$  and  $nM^{-1}\cdot s^{-1}$ . Here [MM] denotes the reaction rate parameters for Michaelis-Menten kinetics ( $V_m [1/s]$  and  $K_m [nM]$ ).

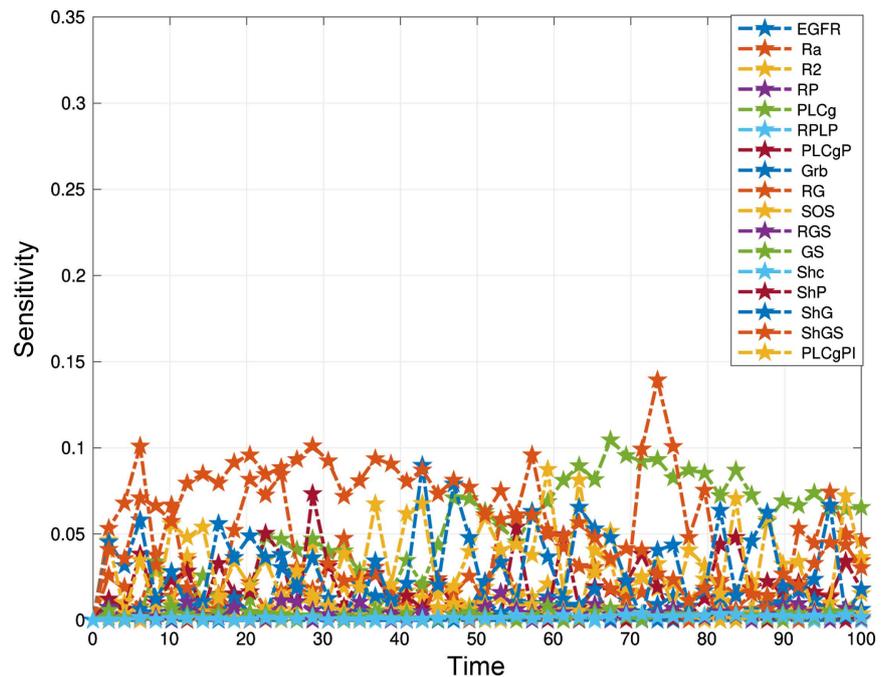
	Reaction channel	Reaction rate <sup>a</sup>
$R_1$	$EGF + R \rightarrow Ra$	$3.0E-3$
$R_2$	$Ra \rightarrow EGF + R$	$6.0E-2$
$R_3$	$Ra + Ra \rightarrow R2$	$1.0E-2$
$R_4$	$R2 \rightarrow Ra + Ra$	$1.0E-1$
$R_5$	$R2 \rightarrow RP$	$1.0$
$R_6$	$RP \rightarrow R2$	$1.0E-2$
$R_7$	$RP \rightarrow R2[MM]$	$4.5E+2, 5.0E+1$
$R_8$	$RP + PLCg \rightarrow RPL$	$6.0E-2$
$R_9$	$RPL \rightarrow RP + PLCg$	$2.0E-1$
$R_{10}$	$RPL \rightarrow RPLP$	$1.0$
$R_{11}$	$RPLP \rightarrow RPL$	$5.0E-2$
$R_{12}$	$RPLP \rightarrow RP + PLCgP$	$3.0E-1$
$R_{13}$	$RP + PLCgP \rightarrow RPLP$	$6.0E-2$
$R_{14}$	$PLCgP \rightarrow PLCg[MM]$	$1.0, 1.0E+2$
$R_{15}$	$RP + Grb \rightarrow RG$	$3.0E-3$
$R_{16}$	$RG \rightarrow RP + Grb$	$5.0E-2$
$R_{17}$	$RG + SOS \rightarrow RGS$	$1.0E-2$
$R_{18}$	$RGS \rightarrow RG + SOS$	$6.0E-2$
$R_{19}$	$RGS \rightarrow RP + GS$	$3.0E-2$
$R_{20}$	$RP + GS \rightarrow RGS$	$4.5E-3$
$R_{21}$	$GS \rightarrow Grb + SOS$	$1.5E-3$
$R_{22}$	$Grb + SOS \rightarrow GS$	$1.0E-4$
$R_{23}$	$RP + Shc \rightarrow RSh$	$6.0E-2$
$R_{24}$	$RSh \rightarrow RP + Shc$	$6.0E-1$
$R_{25}$	$RSh \rightarrow RShP$	$6.0$
$R_{26}$	$RShP \rightarrow RSh$	$6.0E-2$
$R_{27}$	$RShP \rightarrow RP + ShP$	$3.0E-1$
$R_{28}$	$RP + ShP \rightarrow RShP$	$9.0E-4$
$R_{29}$	$ShP \rightarrow Shc[MM]$	$1.7, 3.4E+2$
$R_{30}$	$RShP + Grb \rightarrow RShG$	$3.0E-3$
$R_{31}$	$RShG \rightarrow RShP + Grb$	$1.0E-1$
$R_{32}$	$RShG \rightarrow RP + ShG$	$3.0E-1$
$R_{33}$	$RP + ShG \rightarrow RShG$	$9.0E-4$
$R_{34}$	$RShG + SOS \rightarrow RShGS$	$1.0E-2$

## Continued

$R_{35}$	$RShGS \rightarrow RShG + SOS$	4.5E-3
$R_{36}$	$RShGS \rightarrow RP + ShGS$	1.2E-1
$R_{37}$	$RP + ShGS \rightarrow RShGS$	2.4E-4
$R_{38}$	$ShP + Grb \rightarrow ShG$	3.0E-3
$R_{39}$	$ShG \rightarrow ShP + Grb$	1.0E-1
$R_{40}$	$ShG + SOS \rightarrow ShGS$	3.0E-2
$R_{41}$	$ShGS \rightarrow ShG + SOS$	6.0E-2
$R_{42}$	$ShGS \rightarrow GS + ShP$	1.0E-1
$R_{43}$	$GS + ShP \rightarrow ShGS$	2.0E-2
$R_{44}$	$RShP + GS \rightarrow RShGS$	9.0E-3
$R_{45}$	$RShGS \rightarrow RShP + GS$	4.29E-2
$R_{46}$	$PLCgP \rightarrow PLCgPI$	1.0
$R_{47}$	$PLCgPI \rightarrow PLCgP$	3.0E-2

**Table 3.** The epidermal growth factor receptor (EGFR) signaling pathway model: the molecular species and their initial amounts.

Sid	Species	N
1	EGF	23,040,183
2	R	335
3	Ra	11,774
4	R2	9514
5	RP	1360
6	R-PL	59
7	R-PLP	91
8	R-G	947
9	R-G-S	300
10	R-Sh	23
11	R-ShP	618
12	R-Sh-G	195
13	R-Sh-G-S	124
14	G-S	1776
15	ShP	152,296
16	Sh-G	56,545
17	PLCg	1195
18	PLCgP	2160
19	PLCg-I	185,357
20	Grb	32,547
21	Shc	2634
22	SOS	4689
23	Sh-G-S	52,301



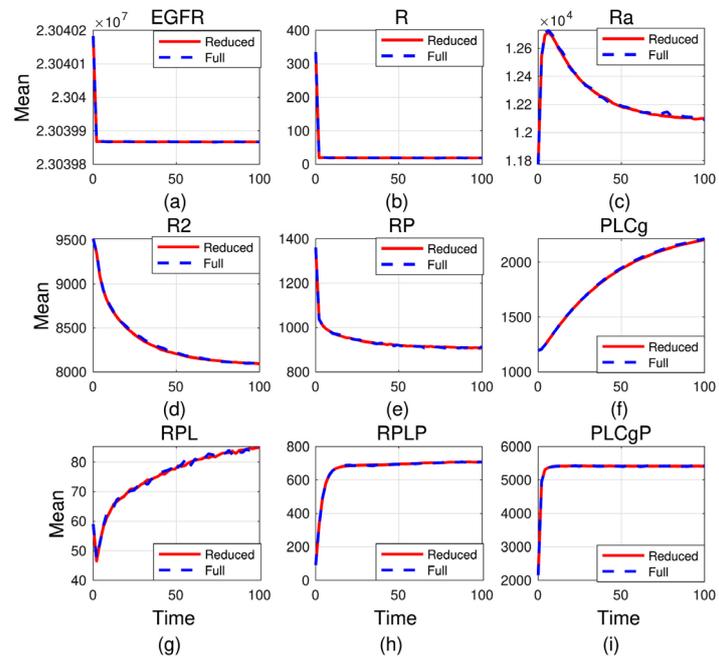
**Figure 4.** The epidermal growth factor receptor (EGFR) signaling pathway model: the evolution in time of the normalized sensitivities for several species with respect to  $c_{21}$ .

for deletion. Given the large dimension of the EGFR model, we also carry a sensitivity analysis to choose several reactions which are very important. We assign the value of their coefficients as  $d_i = 1$ . At this step, we selected a group of 17 such reactions from a total of 47 reactions. This reduces the dimension of the optimization problem. Subsequently, we solve the optimization problem (4) to find a reduced biochemical model. We used the following parameter values:  $\ell = 2$ ,  $u = 42$  and  $r$  close to 1. As a result, the following reactions are selected for elimination  $R_6, R_{19}, R_{20}, R_{21}, R_{22}, R_{34}$  and  $R_{35}$ .

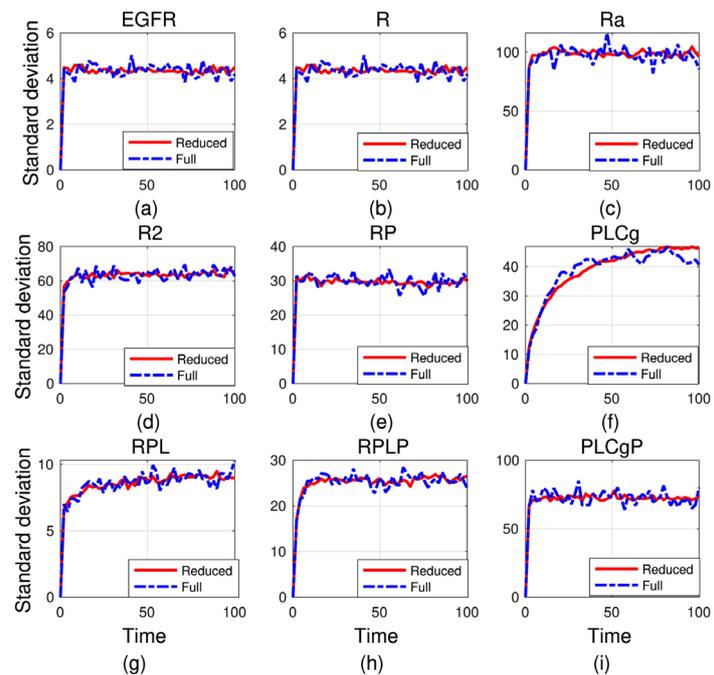
Finally, we run Gillespie's algorithm over 2000 trajectories for the original and the reduced model and compare the output. The means and standard deviations of the molecular amounts of various species, computed for the full and reduced models, are plotted as functions of time in **Figure 5** and **Figure 6**, respectively. Similar results are obtained for the remaining species (not shown, for brevity). Notice the excellent agreement between the results produced with the reduced (with reactions  $R_6, R_{19}, R_{20}, R_{21}, R_{22}, R_{34}$  and  $R_{35}$  turned off) and the full systems. Consequently, the proposed model reduction strategy provides very good results on this large size, stiff biochemical reaction network.

### 4.3. Gemcitabine Model

The final numerical experiment is performed on the gemcitabine (2,2-difluorodeoxycytidine, dFdC) biochemical network, which is a critical, real-world model [26] [27] designed to study the mechanisms of resistance to gemcitabine effectiveness. Gemcitabine is a chemotherapy drug, commonly administered for the



**Figure 5.** The epidermal growth factor receptor (EGFR) signaling pathway model: the mean number of molecules for different species as a function of time, for the full and reduced model. The following reactions are eliminated:  $R_6, R_{19}, R_{20}, R_{21}, R_{22}, R_{34}$  and  $R_{35}$ . The species shown are: (a) EGFR, (b) R, (c) Ra, (d) R2, (e) RP, (f) PLCg, (g) RPL, (h) RPLP and (i) PLCgP. The simulation of 2000 trajectories is performed.



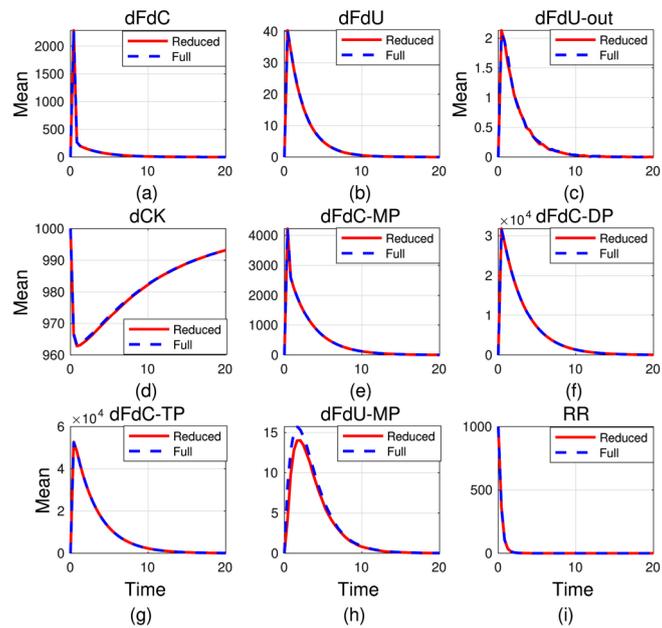
**Figure 6.** The epidermal growth factor receptor (EGFR) signaling pathway model: the standard deviation of the number of molecules for several species, as a function of time, for the full and reduced model. The reactions  $R_6, R_{19}, R_{20}, R_{21}, R_{22}, R_{34}$  and  $R_{35}$  are removed from the system. The species shown are: (a) EGFR, (b) R, (c) Ra, (d) R2, (e) RP, (f) PLCg, (g) RPL, (h) RPLP and (i) PLCgP. The computations are over 2000 trajectories.

treatment of non-small cell lung cancer, breast, pancreatic and prostate cancers [26]. Reactions 1, 2, 3 and 4 illustrate how the gemcitabine is carried into cells. Then, in reaction 5, it is phosphorylated by the enzyme (dCK) to become the gemcitabine monophosphate (dFdc-MP). The species dFdc-MP is phosphorylated to its diphosphorylated form (dFdc-DP) in reactions 7 and 8. The triphosphorylated form (dFdc-TP) is created in reactions 9 and 10. Gemcitabine is metabolized to 2,2-difluorodeoxycytidine (dFdU), which can be phosphorylated to its diphosphate (dFdU-DP) and triphosphate (dFdU-TP), these processes are depicted in reactions 11 - 16. Reaction 19 is the DNA incorporation of dFdc-TP, while reaction 21 indicates the synthesis of CDP. Reaction 23 is nucleoside diphosphate kinase. The inhibition of dCK by binding to dCTP is presented in the 25-th reaction.

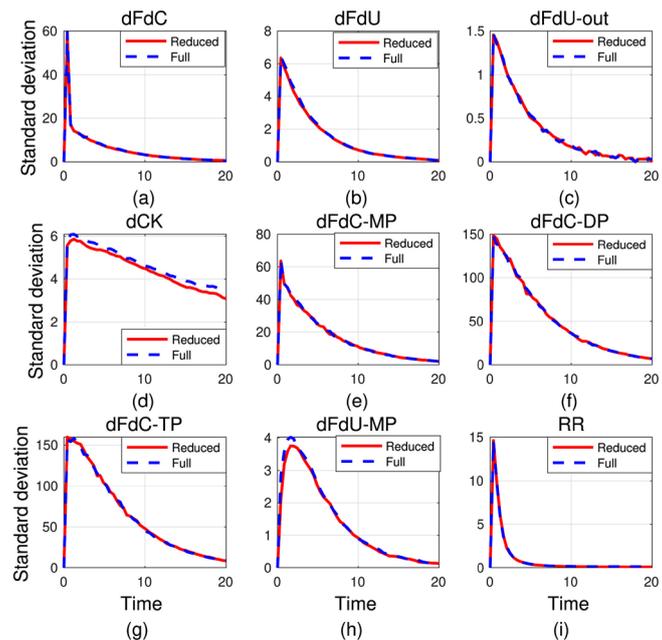
This is a large model, consisting of 22 species undergoing 29 reactions. The reactions and their rate parameters are listed in **Table 4**. The initial molecular counts of all species are zero, except for  $dCK = 1000$ ,  $RR = 1000$ ,  $dCMPD = 1000$ ,  $CDP = 2000$  and  $dFdc\text{-out} = 10,000$ . The model is studied on the time interval  $[0, 20]$ .

We begin by performing a local sensitivity analysis of the Gemcitabine model. To approximate its parametric sensitivities, we apply the common reaction number (CRN) algorithm with 1000 trajectories. At this step, we select some reactions which local sensitivity analysis predicts to be possible choices for deletion. These reactions are then considered an initial collection, when solving the optimization problem for this system. The model has many parameters and finding the global minimum in its large-dimensional parameter space is a challenge. Therefore, we use sensitivity analysis to determine a set of reactions deemed critical in driving the behaviour of the system. For this model, we selected 14 important reactions (and set their  $d_i = 1$ ) out of a total of 29. Also, we choose the following parameter values for the optimizer:  $\ell = 2$  and  $u = 23$  and  $r \approx 1$ . Solving the optimization problem suggests that reactions  $R_2, R_3, R_{12}, R_{14}, R_{16}$  and  $R_{18}$  may be eliminated without a significant impact on the accuracy and stability properties of the model.

To validate the solution of the optimizer, we apply the SSA with 1000 runs to simulate both the reduced (resulted after turning off reactions  $R_2, R_3, R_{12}, R_{14}, R_{16}$  and  $R_{18}$ ) and the original biochemical network. In **Figure 7** and **Figure 8**, we compare the plots of the time-evolution of the means and standard deviations for the molecular counts predicted by the full and the reduced model. In the reduced network, the following reactions were eliminated:  $R_2, R_3, R_{12}, R_{14}, R_{16}$  and  $R_{18}$ . As with the previous models, the proposed reduction technique generates a very accurate simplified model, when compared to the original one. The Gemcitabine model exhibits stiffness and has many reactions and species. Hence, it is quite challenging to simulate and analyze in its original formulation. In conclusion, the proposed simplifying strategy performs very well on this complex and computationally demanding biochemical model.



**Figure 7.** The gemcitabine biochemical reaction model: the time-evolution of the mean number of molecules for several species, for the full and reduced systems. The following reactions are eliminated according to the proposed reduction scheme:  $R_2, R_3, R_{12}, R_{14}, R_{16}$  and  $R_{18}$ . The species shown are: (a) dFdC, (b) dFdU, (c) dFdU-out, (d) dCK, (e) dFdC-MP, (f) dFdC-DP, (g) dFdC-TP, (h) dFdU-MP and (i) RR. The simulation is performed on 1000 trajectories.



**Figure 8.** The gemcitabine biochemical reaction model: the time-evolution of the standard deviation of the number of molecules of different reacting species, for the full and reduced system. Reactions  $R_2, R_3, R_{12}, R_{14}, R_{16}$  and  $R_{18}$  are turned off in the reduced model. The species shown are: (a) dFdC, (b) dFdU, (c) dFdU-out, (d) dCK, (e) dFdC-MP, (f) dFdC-DP, (g) dFdC-TP, (h) dFdU-MP and (i) RR. The simulations are done on 1000 trajectories.

**Table 4.** The gemcitabine biochemical reaction model: The reactions and the values of their rate parameters.

	Reaction channel	Reaction Rates
$R_1$	$dFdC \xrightarrow{k_1} dFdC$	$k_1 = 99.7234$
$R_2$	$dFdC \xrightarrow{k_2} dFdC\text{-out}$	$k_2 = 2.61675E-3$
$R_3$	$dFdC\text{-out} \xrightarrow{k_3} dFdU$	$k_3 = 4.72336E-5$
$R_4$	$dFdU \xrightarrow{k_4} dFdU\text{-out}$	$k_4 = 5.08194E-1$
$R_5$	$dFdC + dCK \xrightarrow{k_5} dFdC\text{-MP} + dCK$	$k_5 = 1.04994E-2$
$R_6$	$dFdC\text{-MP} \xrightarrow{k_6} dFdC$	$k_6 = 8.75208E-1$
$R_7$	$dFdC\text{-MP} \xrightarrow{k_7} dFdC\text{-DP}$	$k_7 = 23.7162$
$R_8$	$dFdC\text{-DP} \xrightarrow{k_8} dFdC\text{-MP}$	$k_8 = 2.12216$
$R_9$	$dFdC\text{-DP} \xrightarrow{k_9} dFdC\text{-TP}$	$k_9 = 25.2037$
$R_{10}$	$dFdC\text{-TP} \xrightarrow{k_{10}} dFdC\text{-DP}$	$k_{10} = 14.4908$
$R_{11}$	$dFdU + dCK \xrightarrow{k_{11}} dFdU\text{-MP} + dCK$	$k_{11} = 9.68E-3$
$R_{12}$	$dFdU\text{-MP} \xrightarrow{k_{12}} dFdU$	$k_{12} = 5.60415E-4$
$R_{13}$	$dFdU\text{-MP} \xrightarrow{k_{13}} dFdU\text{-DP}$	$k_{13} = 0.7844$
$R_{14}$	$dFdU\text{-DP} \xrightarrow{k_{14}} dFdU\text{-MP}$	$k_{14} = 4.20541E-2$
$R_{15}$	$dFdU\text{-DP} \xrightarrow{k_{15}} dFdU\text{-TP}$	$k_{15} = 1.64322$
$R_{16}$	$dFdU\text{-TP} \xrightarrow{k_{16}} dFdU\text{-DP}$	$k_{16} = 9.05139E-4$
$R_{17}$	$dFdC \xrightarrow{k_{17}} dFdU$	$k_{17} = 4.76746E-3$
$R_{18}$	$dFdC\text{-MP} + dCMPD \xrightarrow{k_{18}} dFdU\text{-MP} + dCMPD$	$k_{18} = 4.559E-5$
$R_{19}$	$dFdC\text{-TP} \xrightarrow{k_{19}} dFdC\text{-TP-DNA}$	$k_{19} = 0.544456$
$R_{20}$	$dFdU\text{-TP} \xrightarrow{k_{20}} dFdU\text{-TP-DNA}$	$k_{20} = 7.37496E-3$
$R_{21}$	$\emptyset \xrightarrow{k_{21}} CDP$	$k_{21} = 100$
$R_{22}$	$CDP + RR \xrightarrow{k_{22}} dCDP + RR$	$k_{22} = 0.5$
$R_{23}$	$dCDP \xrightarrow{k_{23}} dCTP$	$k_{23} = 252.037$
$R_{24}$	$dCTP \xrightarrow{k_{24}} CTP\text{-DNA}$	$k_{24} = 5$
$R_{25}$	$dCTP + dCK \xrightarrow{k_{25}} dCTP\text{-dCK}$	$k_{25} = 1.0E-4$
$R_{26}$	$dCTP\text{-dCK} \xrightarrow{k_{26}} dCTP + dCK$	$k_{26} = 0.1$
$R_{27}$	$dFdC\text{-DP} + RR \xrightarrow{k_{27}} dFdC\text{-DP-RR}$	$k_{27} = 1.0E-4$
$R_{28}$	$dFdC\text{-TP} + dCMPD \xrightarrow{k_{28}} dFdC\text{-TP-dCMPD}$	$k_{28} = 100$
$R_{29}$	$dFdC\text{-TP-dCMPD} \xrightarrow{k_{29}} dFdC\text{-TP} + dCMPD$	$k_{29} = 0.1$

## 5. Conclusions

In the present work, we proposed a numerical procedure for model reduction of homogeneous, discrete stochastic biochemical networks. The dynamics of these networks are governed by the widely used model of the Chemical Master Equation. This model has a diverse range of critical practical applications. Numerous biochemical networks arising in practice are complex, consisting of a large number of species subject to many reaction channels. In addition, the network of interactions of some components of the biochemical system may be quite complicated. Furthermore, biochemical processes present at the cellular level typically entail slow and fast reactions, meaning that their mathematical models are stiff. Stiffness and/or high dimensionality constitute significant challenges for numerical simulation and analysis of these models.

The model reduction procedure developed in this paper utilizes sensitivity analysis of stochastic discrete models of biochemical systems and requires solving a nonlinear optimization problem. The procedure is expected to have an excellent performance on a broad class of biochemical systems, with moderate levels of noise, and various degrees of stiffness. We tested the reduction methodology described above on three realistic biochemical networks, two of these having applications to cancer research. The numerical experiments carried out on these models show that the reduced biochemical network retains the properties of the original version while preserving its overall behaviour.

## Acknowledgements

This work was partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] McAdams, H. and Arkin, A. (1999) It's a Noisy Business, Genetic Regulation at the Nanomolar Scale. *Trends in Genetics*, **15**, 65-69. [https://doi.org/10.1016/S0168-9525\(98\)01659-X](https://doi.org/10.1016/S0168-9525(98)01659-X)
- [2] Raser, J.M. and O'Shea, E.K. (2004) Control of Stochasticity in Eukaryotic Gene Expression. *Science*, **304**, 1811-1814. <https://doi.org/10.1126/science.1098641>
- [3] Gillespie, D.T. (1992) A Rigorous Derivation of the Chemical Master Equation. *Physica A: Statistical Mechanics and its Applications*, **188**, 402-425. [https://doi.org/10.1016/0378-4371\(92\)90283-V](https://doi.org/10.1016/0378-4371(92)90283-V)
- [4] Gillespie, D.T. (1976) A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions. *Journal of Computational Physics*, **22**, 403-434. [https://doi.org/10.1016/0021-9991\(76\)90041-3](https://doi.org/10.1016/0021-9991(76)90041-3)
- [5] Gillespie, D.T. (1977) Exact Stochastic Simulation of Coupled Chemical Reactions. *The Journal of Physical Chemistry*, **81**, 2340-2361.

- <https://doi.org/10.1021/j100540a008>
- [6] Huang, H., Fairweather, M., Griffiths, J.F., Tomlin, A.S. and Brad, R.B. (2005) A Systematic Lumping Approach for the Reduction of Comprehensive Kinetic Models. *Proceedings of the Combustion Institute*, **30**, 1309-1316. <https://doi.org/10.1016/j.proci.2004.08.001>
- [7] Li, G. and Rabitz, H. (1989) A General Analysis of Exact Lumping in Chemical Kinetics. *Chemical Engineering Science*, **44**, 1413-1430. [https://doi.org/10.1016/0009-2509\(89\)85014-6](https://doi.org/10.1016/0009-2509(89)85014-6)
- [8] Petzold, L.R. and Zhu, W.J. (1999) Model Reduction for Chemical Kinetics: An Optimization Approach. *AIChE Journal*, **45**, 869-886. <https://doi.org/10.1002/aic.690450418>
- [9] Kumar, A., Christofides, P.D. and Daoutidis, P. (1998) Singular Perturbation Modeling of Non-linear Processes with Non-Explicit Time-Scale Separation. *Chemical Engineering Science*, **53**, 1491-1504. [https://doi.org/10.1016/S0009-2509\(98\)00006-2](https://doi.org/10.1016/S0009-2509(98)00006-2)
- [10] Vora, N. and Daoutidis, P. (2001) Nonlinear Model Reduction of Chemical Reaction Systems. *AIChE Journal*, **47**, 2320-2332. <https://doi.org/10.1002/aic.690471016>
- [11] Varma, A., Morbidelli, M. and Wu, H. (1999) Parametric Sensitivity in Chemical Systems. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9780511721779>
- [12] Cao, Y., Gillespie, D.T. and Petzold, L. (2005) The Slow-Scale Stochastic Simulation Algorithm. *The Journal of Chemical Physics*, **122**, 014116-014134. <https://doi.org/10.1063/1.1824902>
- [13] Rao, C.V. and Arkin, A.P. (2003) Stochastic Chemical Kinetics and the Quasi-steady-state Assumption: Application to the Gillespie Algorithm. *The Journal of Chemical Physics*, **118**, 4999-5010. <https://doi.org/10.1063/1.1545446>
- [14] Samant, A. and Vlachos, D. (2005) Overcoming Stiffness in Stochastic Simulation Stemming from Partial Equilibrium: A Multiscale Monte-Carlo Algorithm. *Journal of Chemical Physics*, **123**, 144114-144122. <https://doi.org/10.1063/1.2046628>
- [15] Sotiropoulos, V., Contou-Carrere, M.N., Daoutidis, P. and Kaznessis, Y.N. (2009) Model Reduction of Multiscale Chemical Langevin Equations: A Numerical Case Study, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **6**, 470-482. <https://doi.org/10.1109/TCBB.2009.23>
- [16] Ilie, S. and Gholami, S. (2013) Simplifying Stochastic Mathematical Models of Biochemical Systems. *Applied Mathematics*, **4**, 248-256. <https://doi.org/10.4236/am.2013.41A038>
- [17] Anderson, D.F. (2012) An Efficient Finite Difference Method for Parameter Sensitivities of Continuous Time Markov Chains. *SIAM Journal on Numerical Analysis*, **50**, 2237-2258. <https://doi.org/10.1137/110849079>
- [18] Rathinam, M., Sheppard, P.W. and Khammash, M. (2010) Efficient Computation of Parameter Sensitivities of Discrete Stochastic Chemical Reaction Networks. *The Journal of Chemical Physics*, **132**, Article No. 034103. <https://doi.org/10.1063/1.3280166>
- [19] Kurtz, T.G. (1982) Representation and Approximation of Counting Processes. In: Fleming, W.H. and Gorostiza, L.G., Eds., *Advances in Filtering and Optimal Stochastic Control*, Vol. 42, Springer, Berlin, 177-191. <https://doi.org/10.1007/BFb0004537>
- [20] Morshed, M., Ingalls, B. and Ilie, S. (2017) An Efficient Finite-Difference Strategy for Sensitivity Analysis of Stochastic Models of Biochemical Systems. *BioSystems*,

- 151, 43-52. <https://doi.org/10.1016/j.biosystems.2016.11.006>
- [21] Morshed, M., Ingalls, B. and Ilie, S. (2018) An Effective Implicit Finite-difference Method for Sensitivity Analysis of Stiff Stochastic Discrete Biochemical Systems. *IET Systems Biology*, **12**, 123-130. <https://doi.org/10.1049/iet-syb.2017.0048>
- [22] Beal, L.D.R., Hill, D., Martin, R.A. and Hedengren, J.D. (2018) GEKKO Optimization Suite. *Processes*, **6**, Article No. 106. <https://doi.org/10.3390/pr6080106>
- [23] Janke, T. (2011) On Reduced Models for Chemical Master Equation. *Multiscale Modeling & Simulation*, **9**, 1646-1676. <https://doi.org/10.1137/110821500>
- [24] Pettigrew, M.F. and Resat, H. (2007) Multinomial Tau-leaping Method for Stochastic Kinetic Simulations. *The Journal of Chemical Physics*, **126**, Article ID: 084101. <https://doi.org/10.1063/1.2432326>
- [25] Sayyidmousavi, A. and Ilie, S. (2017) An Efficient Hybrid Method for Stochastic Reaction-Diffusion Biochemical Systems with Delay. *AIP Advances*, **7**, Article ID: 125305. <https://doi.org/10.1063/1.5001760>
- [26] Kahramanogullari, O., Fantaccini, G., Lecca, P., Morpurgo, D. and Priami, C. (2012) Algorithmic Modeling Quantifies the Complementary Contribution of Metabolic Inhibitions to Gemcitabine Efficacy. *PLoS ONE*, **7**, e50176. <https://doi.org/10.1371/journal.pone.0050176>
- [27] Thanh, V.H. (2019) A Critical Comparison of Rejection-Based Algorithms for Simulation of Large Biochemical Reaction Networks. *Bulletin of Mathematical Biology*, **81**, 3053-3073. <https://doi.org/10.1007/s11538-018-0462-y>