Peer-reviewed version available at Processes 2018, 6, 9; doi:10.3390/pr6020009

Article

Elucidating Cellular Population Dynamics by Molecular Density Function Perturbations

Thanneer Malai Perumal [1] and Rudiyanto Gunawan [2,3,*]

[1] Sage Bionetworks, Seattle, Washington, USA; thanneer.perumal@sagebase.org
[2] Institute for Chemical and Bioengineering, ETH Zurich, Zurich, Switzerland; rudi.gunawan@chem.ethz.ch
[3] Swiss Institute of Bioinformatics, Lausanne, Switzerland
* Correspondence: rudi.gunawan@chem.ethz.ch; Tel.: +41-44-633-2134

Abstract: Studies performed at single-cell resolution have demonstrated the physiological significance of cell-to-cell variability. Various types of mathematical models and systems analyses of biological networks have further been used to gain a better understanding of the sources and regulatory mechanisms of such variability. In this work, we present a novel sensitivity analysis method, called molecular density function perturbation (MDFP), for the dynamical analysis of cellular heterogeneity. The proposed analysis is based on introducing perturbations to the density or distribution function of the cellular state variables at specific time points, and quantifying how such perturbations affect the state distribution at later time points. We applied the MDFP analysis to a model of signal transduction pathway involved in TRAIL (tumor necrosis factor-related apoptosis-inducing ligand)-induced apoptosis in HeLa cells. The MDFP analysis showed that caspase-8 activation regulates the timing of the switch-like increase of cPARP (cleaved poly(ADP-ribose) polymerase), an indicator of apoptosis. Meanwhile, the cell-to-cell variability in the commitment to apoptosis depended on mitochondrial outer membrane permeabilization (MOMP) and events following MOMP, including the release of Smac (second mitochondria-derived activator of caspases) and cytochrome-C from mitochondria, the inhibition of XIAP (X-linked inhibitor of apoptosis) by Smac and the formation of apoptosome.

Keywords: mathematical modeling; biological networks; sensitivity analysis; programmed cell death; single cell dynamics; cell population

1. Introduction

Advances in single-cell profiling technology and the application of this technology to study biology at single-cell resolution have demonstrated the ubiquity and functional role of cell-to-cell variability in physiological processes, such as programmed cell death (apoptosis) and stem cell differentiation [1–3]. Besides genetic, epigenetic and environmental factors, the cellular heterogeneity observed in a given cell population has also been attributed to the inherent stochastic dynamics of cellular processes. For example, gene transcriptional processes have been shown to occur in stochastic (random) bursts [4–6]. Many modeling frameworks have been used to capture cellular heterogeneity, for example by using ensemble models (EM) of ordinary differential equations (ODEs) [7,8], population balance models (PBMs) [9], stochastic ordinary differential equations (SDEs) [10,11], and chemical master equations (CMEs) [12–14]. In these models, the cell-to-cell variability is described by a probability density or distribution function of cell state variables. Systems analyses have also been developed and applied to gain insights into the dynamics of cell state distribution. For example, several types of parameter sensitivity analysis, including SOBOL sensitivity [15], DGSM [16], Glocal analysis [17], extended Fourier Amplitude Sensitivity Test (eFAST) [18], and stochastic sensitivity analysis [19–25], have been used to identify the rate-controlling or bottlenecking processes based on dynamic models of cell distribution.

Parameter sensitivity analysis (PSA) is a common systems analysis that is used to elucidate the dependence of systems behavior on system parameters [26,27]. In the PSA, we compute sensitivity coefficients whose magnitude describes how much system states vary with changes in one or a combination of system parameters. A large sensitivity magnitude means that the system behavior strongly depends on changes in the corresponding parameter(s), an indication of a rate-limiting process. In several publications [28–30], we have shown that the traditional PSA, derived using static perturbations to system parameters, may lead to incorrect conclusions when the rate-limiting process changes with time. For this reason, we have created a new class of sensitivity analysis based on impulse perturbations on parameters and states, called impulse parameter sensitivity analysis (iPSA) and Green's function matrix (GFM) analysis, respectively [28–30]. By introducing impulse perturbations at different times, the new sensitivity analyses are able to reveal not only which processes are rate-limiting, but also when they become rate-limiting.

In this work, we adapted the concept of impulse perturbation based PSA for the dynamical analysis of cell-to-cell variability. The new sensitivity analysis, called molecular density function perturbation (MDFP), is based on time-varying perturbations to the probability density or distribution function of the cell state variables. The MDFP sensitivity coefficients are defined using distribution distances in order to account for changes in the cell state distribution beyond the first-order moment (i.e., population mean). We applied the MDFP analysis to a model of programmed cell death in HeLa cell population [31] and identified key regulators of apoptotic decision making.

2. Material and Methods

Sensitivity analysis of dynamic models of cell distribution has received much interest in recent times along with the rise of systems biology and the increasing attention to single-cell analysis. Novel PSA methods have been developed for the CME models of biological networks [19–25]. Here, the sensitivity coefficients describe changes in the mean of cell state distribution caused by infinitesimal (local) perturbations to the parameter values. Methods for global sensitivity analysis have also been adapted for analyzing cell distribution sensitivities, such as sampling-based Partial Rank Correlation Coefficient (PRCC) and variance-based eFAST [18]. Despite their differences, the aforementioned sensitivity analyses and the corresponding sensitivity coefficients are based on static or persistent parameter perturbations. As we have demonstrated previously, such analysis is incapable of elucidating any dynamic transitions of the bottlenecking process [28–30].

3.1. Molecular Density Function Perturbation (MDFP) analysis

In the following, we formulate the molecular density function perturbation (MDFP) analysis. In MDFP, we describe the cell distribution using a probability density function (PDF), denoted by $f_{\mathbf{X}}(\mathbf{x}, t)$, where $\mathbf{x} \in \mathbb{R}^n$ denotes the cell state vector and t denotes the time variable. This description of cell distribution is flexible enough to accommodate mathematical modelling frameworks that are commonly used to simulate cell-to-cell variability, including EMs, PBMs, SDEs and CMEs. In biological network models, the cell state is typically defined by the concentrations of biomolecules. By definition, the (n-tuple) integral of the PDF $\int_{\mathbf{a}}^{\mathbf{b}} f_{\mathbf{X}}(\mathbf{x}, t) d\mathbf{x}$ gives the fraction of cell population at time t whose states (concentrations) satisfy $\mathbf{a} \leq \mathbf{x} \leq \mathbf{b}$. The basic premise of the MDFP analysis is the same as that of impulse perturbation based sensitivity analysis, specifically the GFM analysis [28], which is to introduce a perturbation to the cell state at time τ and quantify the effect of this perturbation at a later time t ($t \geq \tau$). But, the MDFP analysis uses perturbations to the PDF of the cell state.

In deriving the MDP sensitivity coefficients, we start with following relationship:

$$f_{\mathbf{X}}(\mathbf{x}, t) = f_{\mathbf{X}|\mathbf{X}}(\mathbf{x}, t|\mathbf{x}_\tau, \tau) f_{\mathbf{X}}(\mathbf{x}_\tau, \tau). \quad (1)$$

The PDF $f_{\mathbf{X}|\mathbf{X}}(\mathbf{x}, t|\mathbf{x}_\tau, \tau)$, also known as the transitional probability or the propensity function, gives the conditional PDF of $\mathbf{x}$ at time t given that the cell state is $\mathbf{x}_\tau$ at time τ ($t \geq \tau$). In the following, we consider introducing a mean-shift perturbation to the PDF at time τ to give:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2017 doi:10.20944/preprints201712.0159.v1

Peer-reviewed version available at Processes 2018, 6, 9; doi:10.3390/pr6020009

$$f_{\mathbf{X}}^{\Delta+j}(\breve{\mathbf{x}}_\tau, \tau) = f_{\mathbf{X}}(\breve{\mathbf{x}}_\tau - \delta\mathbf{e}_j, \tau), \quad (2)$$

where $\breve{\mathbf{x}}_\tau$ denotes the perturbed state variables and $\mathbf{e}_j$ denote the j-th column of the identity matrix. Note that the PDF $f_{\mathbf{X}}^{\Delta+j}(\breve{\mathbf{x}}_\tau, \tau)$ corresponds to the PDF $f_{\mathbf{X}}(\mathbf{x}_\tau, \tau)$ with a positive mean-shift of $\delta\mathbf{e}_j$, i.e. $\breve{\mathbf{x}}_\tau = \mathbf{x}_\tau + \delta\mathbf{e}_j$. Given the perturbed PDF $f_{\mathbf{X}}^{\Delta+j}(\breve{\mathbf{x}}_\tau, \tau)$ at time τ and the propensity function $f_{\mathbf{X}|\mathbf{X}}(\mathbf{x}, t|\mathbf{x}_\tau, \tau)$, we can define the perturbed PDF of the cell state at time t, denoted by $f_{\mathbf{X}}^{\Delta+j,\tau}(\mathbf{x}, t)$, as follows:

$$f_{\mathbf{X}}^{\Delta+j,\tau}(\breve{\mathbf{x}}, t) = f_{\mathbf{X}|\mathbf{X}}(\breve{\mathbf{x}}, t|\breve{\mathbf{x}}_\tau, \tau) f_{\mathbf{X}}^{\Delta+j}(\breve{\mathbf{x}}_\tau, \tau). \quad (3)$$

Note that the following equality applies

$$f_{\mathbf{X}}^{\Delta+j,\tau}(\breve{\mathbf{x}}, \tau) = f_{\mathbf{X}}^{\Delta+j}(\breve{\mathbf{x}}_\tau, \tau). \quad (4)$$

In the MDFP analysis, we employ a distribution distance metric to quantify the magnitude of PDF changes caused by the perturbations. Several metrics of distribution distance are available, such as the Kullback-Leibler distance (Δ_{KL}), Jeffrey distance (Δ_J), Jennson-Shanon divergence (Δ_{JS}), engineering metric (Δ_E), Kolmogorov-Smirnov distance (Δ_{KS}) and Cramer-von Mise distance (Δ_{CVM}). The first four of the aforementioned distribution distances are based on the difference between two PDFs, while the remaining two are based on the difference between the cumulative density functions (CDFs). For the analysis of programmed cell death model below, we used the Cramer von Mise distance: (see Supplementary Material S1 for the mathematical definitions of the other distribution distances)

$$\Delta_{CVM}\left(f^A(z)||f^B(z)\right) = \int_{-\infty}^{\infty}\left(F^A(z) - F^B(z)\right)^2 dz, \quad (5)$$

which in our experience, gives more reliable sensitivity coefficient calculations. The variable $F(z)$ denotes the CDF of the PDF $f(z)$, i.e. $F(z) = \int_{-\infty}^{z} f(y)dy$.

Following the common definition of sensitivity coefficients [32], we computed the MDFP sensitivity coefficients as the ratio of the changes in the PDF of the cell state at time t and the perturbation introduced at time τ. The sensitivity coefficients are evaluated for a particular cell state variable of interest x_i (the i-th element of $\mathbf{x}$) with respect to a perturbation to $\delta\mathbf{e}_j$ on the state variable x_j, as follows:

$$S_{i,j}^{MDFP}(t, \tau) = sign\left(\frac{\Delta\mu_{X_i}(t)}{\Delta\mu_{x_j}(\tau)}\right)\frac{\Delta_{CVM}\left(f_{X_i}^{\Delta+j,\tau}(\breve{x}_i,t)||f_{X_i}^{\Delta-j,\tau}(\breve{x}_i,t)\right)}{\Delta_{CVM}\left(f_{X_j}^{\Delta+j}(\breve{x}_j,\tau)||f_{X_j}^{\Delta-j}(\breve{x}_j,\tau)\right)}, \quad (6)$$

where $sign(\cdot)$ gives the sign of the argument variable and $\Delta\mu_{X_i}(t)$ denotes the change in the mean of the state variable x_i at time t. The function $f_{X_i}^{\Delta+j,\tau}(\breve{x}_i, t)$ denotes the marginal PDF of $f_{\mathbf{X}}^{\Delta+j,\tau}(\breve{\mathbf{x}}, t)$ with the following definition:

$$f_{X_i}^{\Delta+j,\tau}(\breve{x}_i, t) = \int f_{\mathbf{X}}^{\Delta+j,\tau}(\breve{\mathbf{x}}, t)\, d\breve{\mathbf{x}}_{\sim i}. \quad (7)$$

The integration in Eq. (7) is performed over all state variable $\breve{\mathbf{x}}$'s except for $\breve{x}_i$. Note that the sensitivity coefficient in Eq. (6) above is motivated by the centered difference approximation [32], where the sensitivity coefficients are computed using positive and negative perturbations to the system.

The definition of the MDFP sensitivity coefficients is analogous to the Green's function matrix (GFM) sensitivity [28,32]. We can visualize $S_{i,j}^{MDFP}(t, \tau)$ using a heatmap as shown in Figure 1. The magnitudes of the sensitivities represent the degree of importance, while the signs of the sensitivity coefficients reflect the direction of the mean change. A positive sensitivity coefficient indicates that the mean change of x_i at time t is in the same direction as the mean-shift perturbation to x_j at time τ. One can further use the magnitudes of the sensitivity coefficients to rank state variables (at time τ)

according to the degree of influence on a particular state variable (at time t), where larger sensitivity magnitudes indicate a higher importance.

In the case study, we considered an ensemble of ODE models, where each model represents one cell in a cell population. The models in the ensemble share the same ODEs and parameters, but have different initial states. The ODE model follows the general formula:

$$\frac{d\mathbf{x}(t,\mathbf{p})}{dt} = g(\mathbf{x},\mathbf{p}), \tag{8}$$

where $\mathbf{p}$ denotes the vector of model parameters and $g(\mathbf{x},\mathbf{p})$ is a vector-valued nonlinear function. The distribution of the initial conditions is given by the PDF $f_{\mathbf{X}}(\mathbf{x}_{t_0}, t_0)$. The sensitivity coefficients were computed using a Monte Carlo approach, where we simulated an ensemble of ODE models with a random sample generated from $f_{\mathbf{X}}(\mathbf{x}_{t_0}, t_0)$ as initial conditions. The model simulation of each randomly sampled initial condition represented the state trajectory of one cell in the cell population. For the computation of $S_{i,j}^{MDFP}(t,\tau)$, we introduced a perturbation δ to the state variable x_j for each of the cells in the ensemble at selected time points τ, and simulated the perturbed state trajectory of x_i until the desired time t $(t \geq \tau)$. We constructed the marginal PDFs or CDFs of the state variables using a kernel density estimator with leave-one-out cross validation [33].

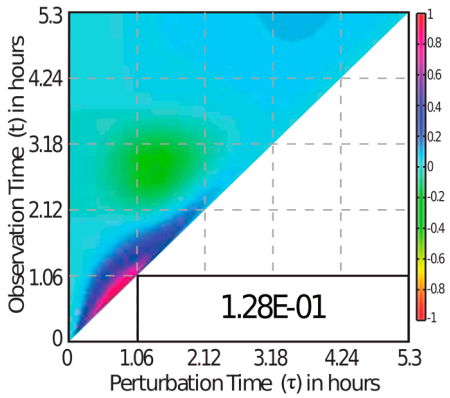

Figure 1. A heatmap of MDFP sensitivity coefficient. The x-axis represents the time τ at which perturbation is introduced, while the y-axis represents the observation time t. The MDFP coefficient in the heatmap is scaled such that the magnitude falls within ±1, and the scaling factor is reported in the lower right corner of the plot. The sensitivity values for $t < \tau$ are set to zero for causal systems.

3.2. Green Function Matrix analysis

We compared the MDFP analysis to a related sensitivity analysis based on the GFM. Like the MDFP analysis, the GFM analysis introduces time-dependent perturbations to the state variables. The GFM sensitivity coefficients are calculated by directly differentiating the ODE model in Eq. (8) as follows: [28]

$$\frac{d}{dt}\left(\frac{\partial \mathbf{x}(t)}{\partial \mathbf{x}(\tau)}\right) = \frac{d}{dt}\mathbf{S}^{GFM}(t,\tau) = \frac{\partial g(\mathbf{x},\mathbf{p})}{\partial \mathbf{x}}\mathbf{S}^{GFM};\ \mathbf{S}^{GFM}(\tau,\tau) = \mathbf{I}_n \tag{10}$$

where $\mathbf{S}^{\boldsymbol{GFM}}(t,\tau)$ is the $n \times n$ sensitivity matrix and $\mathbf{I}_n$ is the $n \times n$ identity matrix. The (i,j)-th element of $\mathbf{S}^{GFM}(t,\tau)$, i.e. $S_{i,j}^{GFM}(t,\tau) = dx_i(t)/x_j(\tau)$, gives the sensitivity of the state $x_i(t)$ with respect to perturbations to the state $x_j(\tau)$ $(t \geq \tau)$. In the case study below, we normalized the sensitivity coefficients as follows:

$$\hat{S}_{i,j}^{GFM}(t,\tau) = S_{i,j}^{GFM}(t,\tau)\frac{x_j(\tau)}{x_i(t)} \tag{11}$$

We computed the GFM sensitivity coefficients following the procedure described in the original publication [28].

3. Results

3.1. TRAIL-induced cell death model of HeLa cells

Figure 2 depicts the signaling network associated with extrinsically-induced apoptosis by tumor necrosis factor related apoptosis-inducing ligand (TRAIL). The ODE model comprises 58 species, 28 reactions and 70 kinetic parameters [34] (see Supplementary Material S2 for details on the initial conditions, parameter values and rate equations). The model parameters and initial conditions were determined by parameter fitting to experimental data of single-cells and cell populations from cell imaging, flow cytometry, and immunoblotting experiments [31,35,36]. The model describes the key mechanisms for the activation of endogenous executioner caspase-3 (C3*) and the subsequent cleavage of poly(ADP-ribose) polymerase (PARP) [35]. Specifically the model includes four major pathways: (i) **the upstream pathway**, representing TRAIL-induced cleavage of pro-caspase-8 (C8) to caspase-8 (C8*), (ii) the mitochondrial independent **type-I pathway**, including the cleavage of pro-caspase-3 (C3) to caspase-3 (C3*) by C8* and the inhibition of C3 by XIAP, (iii) the mitochondrial dependent **type-II pathway**, including the formation of mitochondrial pores promoted by C8*, the consequent release of cytochrome-C (Cyc) into the cytosol, the formation of apoptosome (Apop) induced by cytosolic Cyc, and the activation of C3 by apoptosome, and (iv) the pro-caspase-6 (C6) **positive feedback-loop** where active C3* could promote the activation of C8. In the following, we applied the GFM and MDFP sensitivity analysis to elucidate the key processes in the cell death decision making. More specifically, we computed the GFM and MDFP sensitivity coefficients of the cleaved PARP (cPARP) concentration, an indicator of apoptosis, with respect to perturbations to the concentration of molecules involved in the regulation of PARP cleavage, except the intermediate complexes, in the model.

3.1. GFM analysis of TRAIL-induced cell death

We applied the GFM analysis to the ODE model and model parameters in the original report and using the median initial concentration (see Supplementary Material S.2) [36]. The analysis was performed for a constant TRAIL stimulation over a time range of 0 to 5.3 hours, in which the concentrations reached steady state levels (see Figure 2(b)). Here, the ODE model simulated an apoptotic cell, in which the cleavage of PARP in response to TRAIL follows a delayed switch-like profile, as shown in Figure 3(a). To study the activation dynamics of cPARP in greater detail, the analysis of the GFM sensitivity coefficients was split into two phases: before and after mitochondrial outer membrane permeabilization (pre- and post-MOMP). Following a previous study [31], we defined MOMP to occur when 10% of the total PARP has been cleaved into cPARP, which in this analysis, occurred at 2.36 hours (see Figure 2(b)). Figure 3(b) and (c) portray the ten largest GFM sensitivity coefficients of the cPARP concentration $\hat{S}^{GFM}_{cPARP,j}(t,\tau)$ in the pre- and post-MOMP phases, respectively (see Supplementary Figure S1 and S2 for the complete GFM sensitivity coefficients). In the pre-MOMP phase, the largest GFM sensitivity coefficients were associated with the upstream and type-I pathways, indicating that the early dynamics of cPARP response to TRAIL stimulus depends on these two pathways. In the post-MOMP phase, the top sensitivity coefficients corresponded to the type-II pathway, specifically the regulators of MOMP, i.e. signaling molecules upstream of M* in the network in Figure 2. Thus, the GFM analysis indicated that the switch-like dynamics in the cPARP concentration relies on the mitochondrial-dependent pathway.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2017 doi:10.20944/preprints201712.0159.v1

Peer-reviewed version available at Processes 2018, 6, 9; doi:10.3390/pr6020009

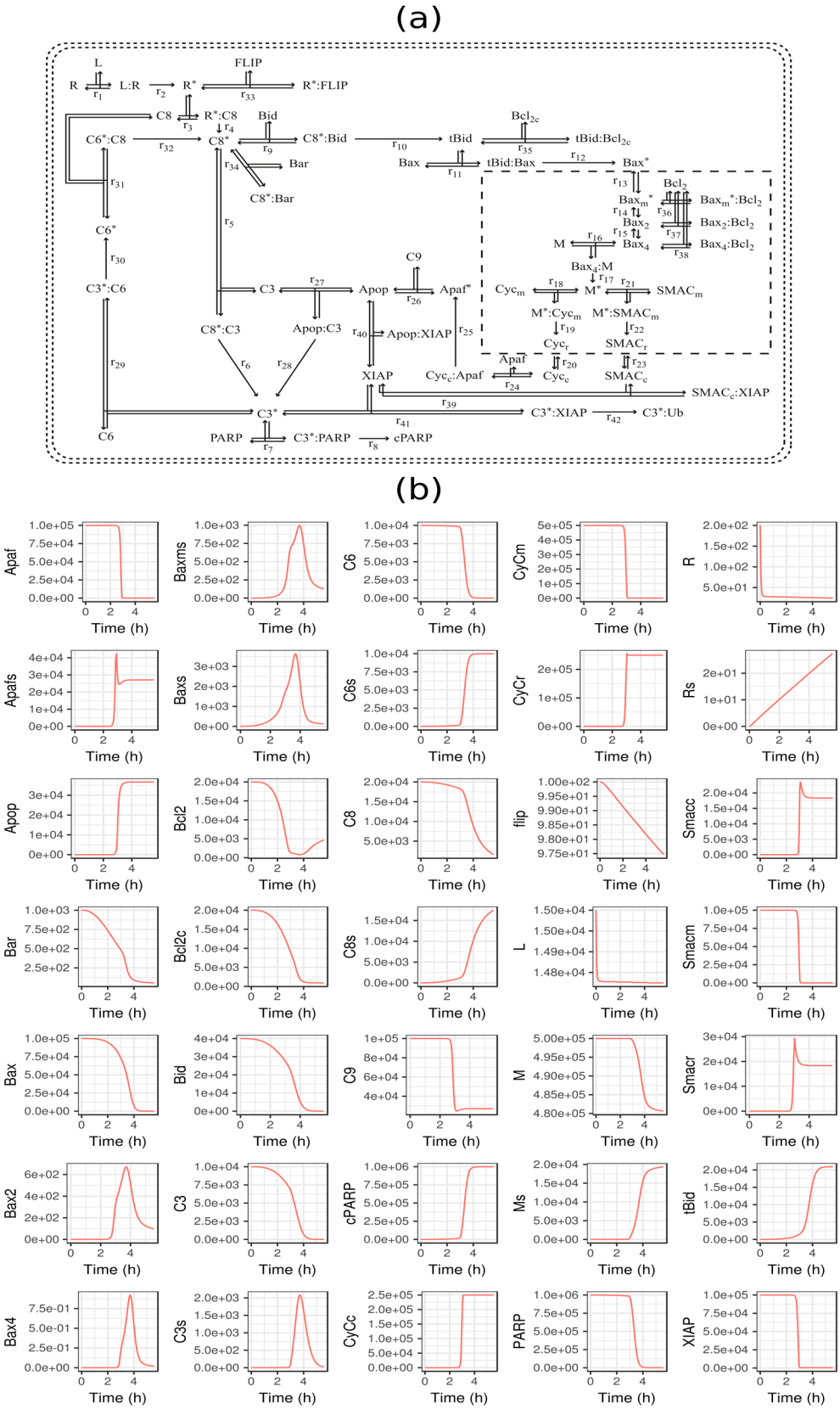

Figure 2. Signal transduction pathway and model simulation of TRAIL-induced apoptosis in Hela cells. (a) Signal transduction pathway of apoptosis. Type-I pathway involves a direct cleavage of pro-caspase-3 by caspase-8 to form an active caspase-3 which cleaves the substrate PARP to cPARP, while the type-II pathway describes a mitochondria-dependent activation of caspase-3 in turn activating the substrate. (b) Model simulation of signal transduction pathway in response to TRAIL.

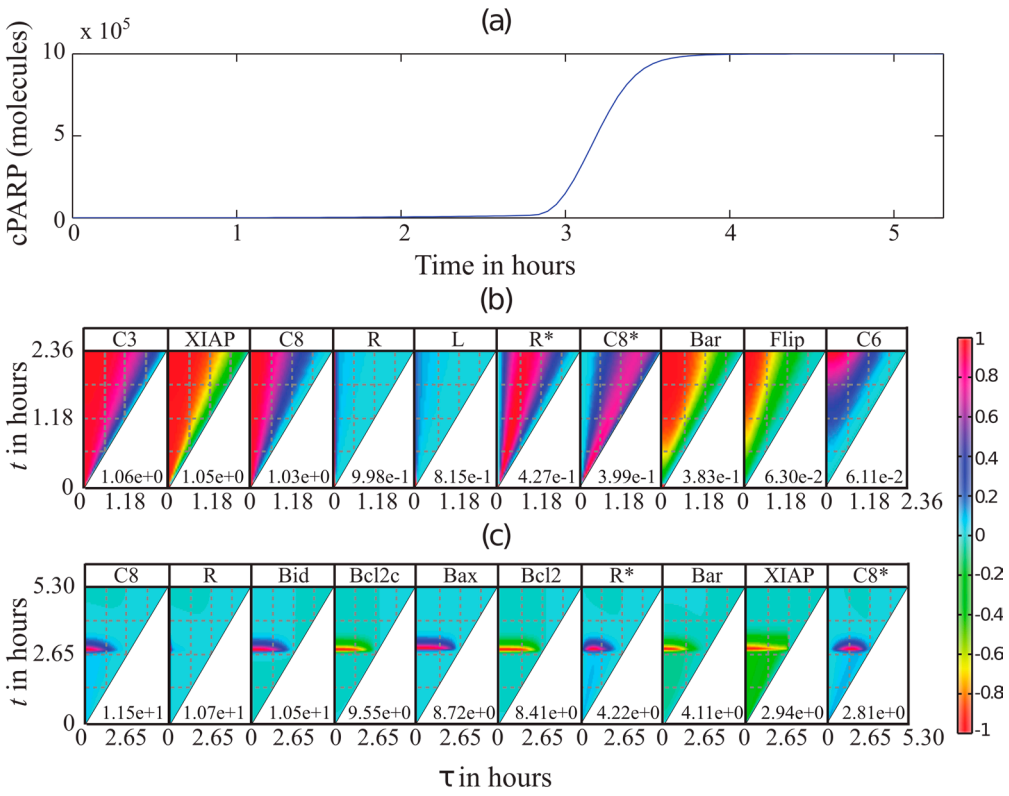

Figure 3. GFM analysis of cPARP activation by a constant TRAIL stimulus in a single cell. (a) cPARP activation follows a delayed switch-like trajectory in response to a constant TRAIL stimulus. (b-c) Ten largest GFM sensitivities coefficients of cPARP concentration (in magnitude) with respect to perturbations to the state variables in the network, as shown on the label of each subfigure. The x-axis gives the time of perturbation τ, while the y-axis represents the time of observation t. Each heatmap is scaled to have values between ±1, using the scaling factor reported in the lower right corner of the plot. Panel (b) shows the GFM sensitivity coefficients in the pre-MOMP phase (before 2.36 hours). Panel (c) shows the GFM sensitivity coefficients in the post-MOMP phase (after 2.36 hours).

3.2. MDFP analysis of TRAIL-induced cell death

The MDFP analysis was carried out for the same TRAIL stimulation as the GFM. For the calculation of the cell distribution, we generated 5 ensembles of 1000 initial concentrations from a log-normal distribution using the Latin Hypercube Sampling (LHS) algorithm based on the reported mean values and coefficient of variations in the original publication [36] (see Supplementary Material S2). Figure 4(a) gives the time evolution of the distribution of cPARP concentration based on the simulations of the ODE model using the ensemble of initial concentrations. Following the original study, we defined cells to be apoptotic when 50% of the total PARP at the final time exist in its cleaved form [36]. The ensemble model simulations showed that on average, roughly 95% of the cells in the simulated cell population undergo apoptosis, similar to what was reported in the original modeling study [36].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2017 doi:10.20944/preprints201712.0159.v1

Peer-reviewed version available at Processes 2018, 6, 9; doi:10.3390/pr6020009

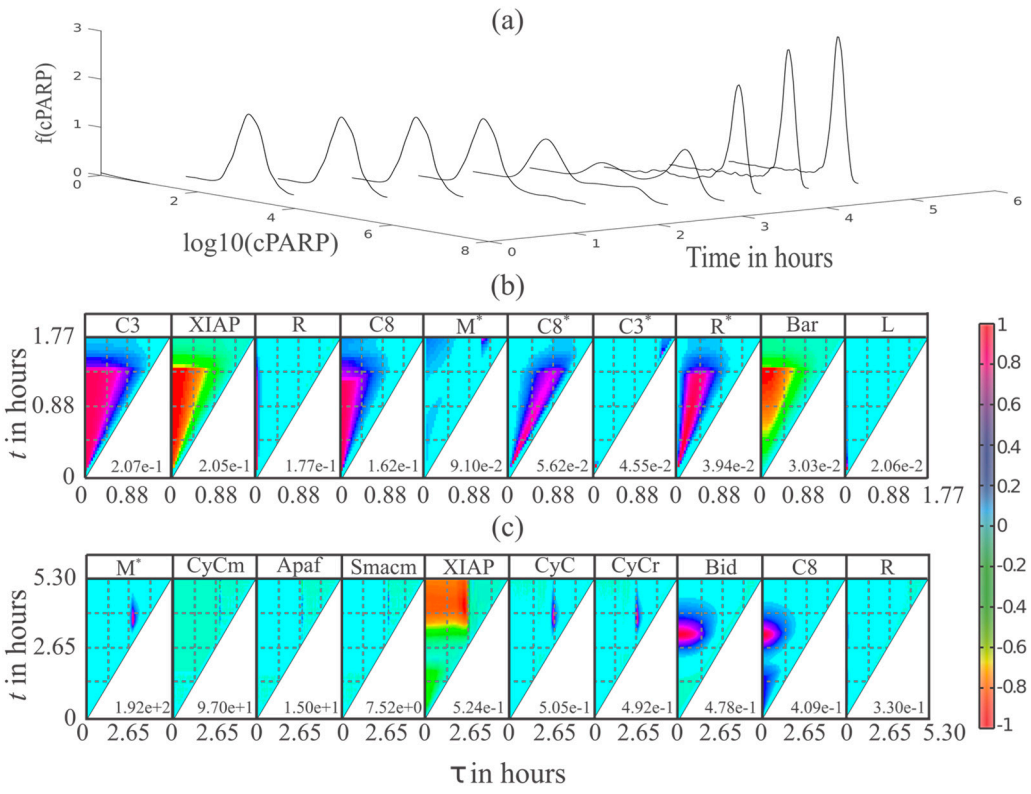

Figure 4. MDFP analysis of cPARP in response to a constant TRAIL stimulus. (a) Time evolution of the distribution of cPARP concentration shows a switch-like behaviour. (b-c) Ten largest MDFP coefficients of cPARP concentration (in magnitude) with respect to the perturbations to different state variables in the network. The x-axis gives the time of perturbation τ, while the y-axis the time of observation t. Each heatmap is scaled to have values between ±1, using the scaling factor reported in the lower right corner of the plot. Panel (b) shows the MDFP sensitivity coefficients pre-MOMP (until 1.76 hours). Panel (c) shows the MDFP sensitivity coefficients post-MOMP.

As in the GFM analysis above, we computed the MDFP sensitivities of cPARP with respect to perturbations to the concentrations of other molecules in the network. Following Eq. (2), we introduced a mean shift perturbation to the distribution of each state variable at various perturbation times τ, specifically by adding +10% or –10% of the mean concentration to the state variable $x_j(\tau)$ (i.e., $\delta \mathbf{e}_j = \pm 0.1 \mu_j(\tau) \mathbf{e}_j$, where $\mu_j(\tau)$ is the mean of the state x_j at time τ). Starting from the perturbed concentrations, we simulated the time-evolution of cPARP for time $t \geq \tau$. Based on these simulations, we reconstructed the marginal PDFs and CDFs of the cPARP using the kernel density function approximation, which were then used in the calculation of the sensitivity coefficients as prescribed in Eq. (6).

Figure 4 (b-c) shows the heatmaps of the ten largest (in magnitude) MDFP sensitivity coefficients, averaged over the five ensembles (see Supplementary Figure S3 and S4 for the complete MDFP sensitivity coefficients). We again split the analysis into two phases: pre- and post-MOMP at 1.76 hours, the time when the median of cPARP concentration reached 10% of the median of total PARP concentration. The MDFP analysis showed that again, the early response of cPARP to TRAIL-induced apoptosis depends on the upstream and type-I pathway molecules. Meanwhile, the cleavage of PARP in the post-MOMP phase is sensitive to mitochondrial dependent pathway molecules, confirming the general finding of the GFM analysis above. In contrast to the GFM analysis, the MDFP sensitivity coefficients pointed to events during and after MOMP, such as the release of cytochrome C from mitochondria, the binding of XIAP by Smac and the formation of apoptosome, as the key regulators of cPARP concentrations.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2017 doi:10.20944/preprints201712.0159.v1

Peer-reviewed version available at Processes 2018, 6, 9; doi:10.3390/pr6020009

3.4. MDFP analysis of apoptotic and non-apoptotic HeLA cells

We repeated the MDFP analysis focusing on the subpopulations of apoptotic and non-apoptotic cells separately. Here, the final cPARP concentration (at time 5.3 hours) was taken to be the indicator of apoptosis, where an apoptotic cell have at least 50% of the total PARP cleaved (see Figure 5 (a-b)) [36]. Since only 5% of the population was non-apoptotic, a resampling of the initial conditions were performed to simulate 10^4 cells, from which a population of 10^3 apoptotic and 10^3 non-apoptotic cells were chosen for the following analysis. In the following, we compared the infinite norm of the MDFP sensitivity coefficients of the final cPARP level with respect to each state variable over all perturbation times, i.e. $\left\|S^{MDFP}_{cPARP,j}\right\|_{\infty} = \max_{\tau} S^{MDFP}_{cPARP,j}(5.3 \text{ hour}, \tau)$. Figure 5 (c-d) shows the ranking of the ten molecules according to the magnitude of $\left\|S^{MDFP}_{cPARP,j}\right\|_{\infty}$ (see Supplementary Figures S5 for the complete MDFP sensitivity coefficients of non-apoptotic cells). The MDFP ranking of the apoptotic subpopulation was in agreement with the GFM analysis, in which the final cPARP level depended on the molecules that regulate MOMP. The similarity between the GFM and MDFP analysis of apoptotic subpopulation is perhaps not surprising considering that the GFM analysis was applied to the model of a cell undergoing apoptosis. Meanwhile, the analysis of the non-apoptotic subpopulation produced a ranking that resembled the outcome of the MDFP analysis of the cell population above. Comparing the analysis of the apoptotic and non-apoptotic cells showed the importance of MOMP, XIAP and its inhibitor Smac, and Apaf-1, in regulating the final cPARP in the non-apoptotic cells. Interestingly, among the apoptotic cells, XIAP was not among the 10 largest sensitivity coefficients.

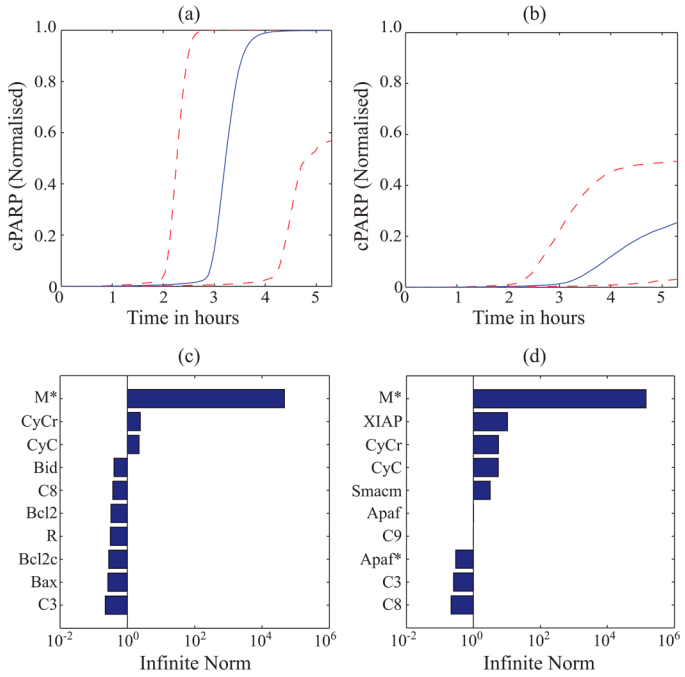

Figure 5. MDFP analysis of the final cleaved PARP levels in (a,c) apoptotic and (b,d) non-apoptotic cell subpopulations . (a-b) The level of cPARP normalized with respect to the total PARP level. The dashed lines (--) indicate the 1 and 99 percentiles of the cPARP levels, while the solid line (-) represents the median level. (c-d) Ten largest sensitivity coefficients in magnitude in the apoptotic and non-apoptotic cells, respectively.

4. Discussion

Cell-to-cell variability has important functional roles in physiological processes, such as cell decision making in stem cell differentiation and cell death. In this study, we developed a sensitivity analysis method called molecular density function perturbation (MDFP) based on introducing

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2017 doi:10.20944/preprints201712.0159.v1

Peer-reviewed version available at Processes 2018, 6, 9; doi:10.3390/pr6020009

time-varying mean-shift perturbations to the distribution of molecular concentrations and quantifying the effects of such perturbations on the distribution of the concentration of molecules of interest. The magnitude of the MDFP sensitivity coefficients indicates how much a perturbation to the concentration PDF of one molecule introduced at a particular time τ affects the concentration PDF of a molecule of interest at some time t ($t \geq \tau$). We applied the MDFP analysis to a model of programmed cell death signalling network in HeLa cell population, to elucidate the apoptosis decision making. We used the magnitude of the sensitivity coefficients to rank the importance of each molecule in determining the concentration of cleaved PARP, an indicator of apoptosis.

In the application of the MDFP analysis, we employed the Cramer von Mise distribution distance Δ_{CVM} in the calculation of the sensitivity coefficients. As mentioned in **Material and Methods**, there also exist several alternative distribution distance metrics for the sensitivity coefficients. The rankings of molecules based on the cPARP sensitivity coefficients using different distribution distances were strongly correlated with the Cramer von Mise and with each other (see Supplementary Figure S6(a)). Furthermore, the ranking of molecules using different perturbation magnitudes (1%, 10% and 100% of the mean) were in agreement with each other (see Supplementary Figure S6(b)). Thus, the conclusion of the MDFP analysis did not depend strongly on the choice of distribution distance and perturbation sizes.

The MDFP analysis provides dynamic information on the bottlenecking process by revealing the molecular concentrations to which perturbations introduced at time τ would elicit a large change in a particular state variable of interest at t. For example, referring to Fig. 4 (c), the heatmap of the MDFP sensitivity coefficient of cPARP with respect to pro-caspase-8 (C8) indicated that perturbing the distribution of pro-caspase-8 at the beginning of the experiment $\tau = 0$ (hour) would cause a much higher impact on cPARP compared to perturbation delivered after ~2 hours. Fig. 6 shows the effect of a positive mean-shift perturbation to C8 ($\delta = +\mu_{C8}$) at two different perturbation times, either $\tau = 0$ hour or $\tau = 2.14$ hours, on the mean, median and standard deviation of cPARP distribution, confirming the MDFP sensitivity analysis.

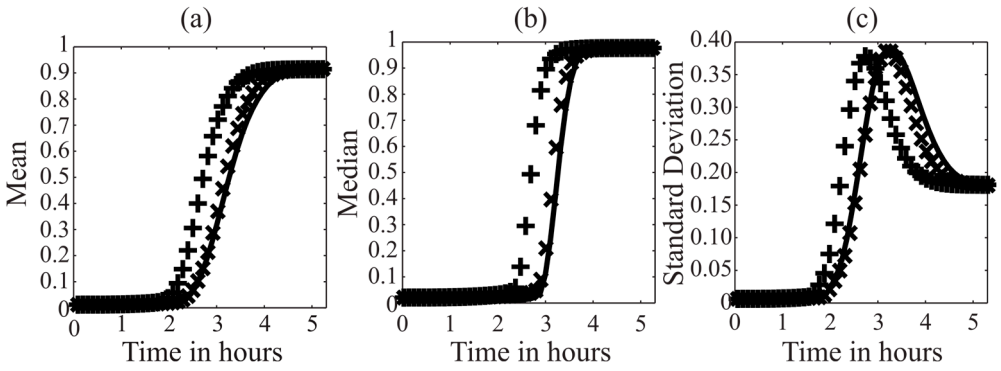

Figure 6. Validation of the MDFP sensitivity analysis of cPARP. A positive mean-shift perturbation to pro-caspase-8 was given either at at $\tau = 0$ hour (+) or at $\tau = 2.14$ hours (×). Panel (a) shows the mean, panel (b) gives the median, and panel (c) gives the standard deviation of the cPARP concentration. The unperturbed simulation is shown as solid lines (−).

The MDFP analysis of the cell distribution and the GFM analysis of the nominal ODE model provided somewhat different conclusions with respect to the regulation of PARP cleavage. According to the GFM analysis, the switching dynamics of cPARP depended on the molecules upstream of MOMP, particularly the initial level of procaspase-8 (C8). On the other hand, the MDFP analysis suggested that PARP cleavage was strongly sensitive to the MOMP and the subsequent release of cytochrome C into the cytosolic compartment. As done in Figure 6, we compared perturbing the mean initial concentration of pro-caspase 8 (C8) at the initial time $\tau = 0$ with perturbing the mean number of mitochondrial open pores (M*) at time $\tau = 2.14$ hour, when M* levels reached steady state for more than 99% of the cells. Both perturbations were implemented using 100% positive mean-shifts. Figure 7 shows the effects of the perturbations above on the mean,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 22 December 2017 doi:10.20944/preprints201712.0159.v1

Peer-reviewed version available at Processes 2018, 6, 9; doi:10.3390/pr6020009

median and standard deviation of cPARP concentration. As illustrated in Figure 7, both perturbations led to similar shifts in the mean and median of cPARP, where the switch-like dynamic of PARP cleavage occurred earlier and more swiftly. Meanwhile, the perturbation to M* led to a larger drop in the standard deviation of cPARP than the perturbation to C8, i.e. cells became more alike when we increased the number of mitochondrial open pores. More importantly, while the positive mean-shift perturbation to pro-caspase-8 led to a faster cleavage of PARP, this perturbation did not affect the fraction of apoptotic versus non-apoptotic cells. Meanwhile, when we increased the number of mitochondrial open pores, the fraction of non-apoptotic cells dropped from 5.6% to 3.2%.

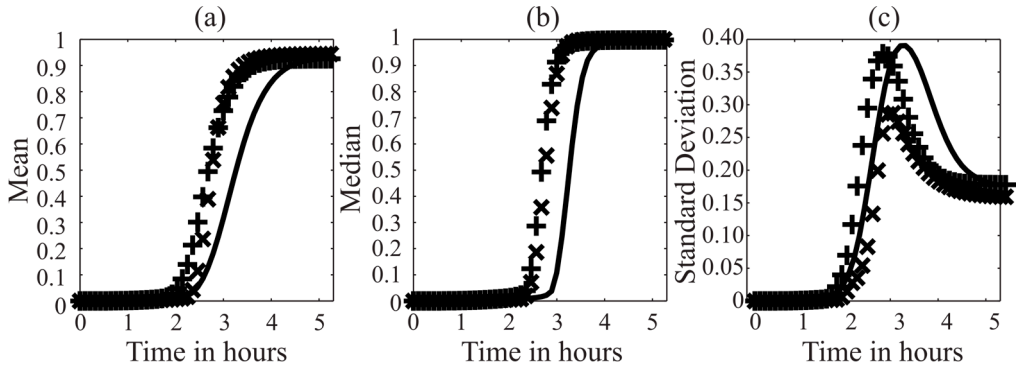

Figure 7. Comparison of GFM and MDFP analyses. A positive mean-shift perturbation was given either to pro-caspase-8 a $\tau = 0$t hour (+) or to mitochondrial open pores M* at $\tau = 2.14$ hours (×). Panel (a) shows the mean of the cPARP concentration distribution, panel (b) gives the median, and panel (c) gives the standard deviation. The unperturbed simulation is shown as solid lines (−).

Both the GFM and MDFP analysis of the cell death signalling network implicated the mitochondrial-dependent type II pathway to be the responsible mechanism of switch-like activation of PARP in HeLa cells, placing caspase-8 activation (cleavage) as the most important steps in the apoptosis decision making during pre-MOMP. This finding agreed with a previous experimental study on fractional killing by TRAIL [31,37], reporting that the activation of C8 controls the switching time of cPARP. In post-MOMP, the GFM analysis indicated that perturbations to the regulators of MOMP (upstream of M* in Figure 2) would strongly affect the PARP cleavage dynamics. On the other hand, the MDFP analysis pointed to MOMP and events post-MOMP (downstream of M* in Figure 2), including cytochrome C and Smac release from mitochondria, XIAP binding by Smac, and apoptosome formation, to be the key determinants of the cell-to-cell variability in cPARP level. The finding from the MDFP analysis is in agreement with a previous study that found XIAP to be the determining factor for the rate and extent of type-II cell death [38]. Furthermore, the results of the MDFP analysis of apoptotic and non-apoptotic subpopulations showed that perturbations to molecules executing the cell death signal after MOMP (i.e., Smac, cytochrome C, Apaf-1) had a stronger effect on the cPARP activation in the non-apoptotic cells than in the apoptotic cells. Consistent with this insight, the depletion of Apaf-1 or Apaf-1/Smac together by siRNA significantly reduced the activation of PARP in HeLa cells (see Supplementary Figure S7 in [36]).

As the functional significance of cell-to-cell variability is increasingly recognized and mathematical models that are able to describe cell distribution become more and more common, the MDFP analysis proposed here would provide an analytical tool to use such models for elucidating the key molecules and processes that govern the dynamics of cellular heterogeneity.

Supplementary Materials: The following are available online at www.mdpi.com/link, Material S1: Probability distance metrics, Material S2: TRAIL induced programmed cell death model of Hela cells, Figure S1: GFM analysis of TRAIL induced apoptosis model during pre-MOMP (before 2.36 hours), Figure S2: GFM analysis of TRAIL induced apoptosis model during post-MOMP (after 2.36 hours), Figure S3: MDFP analysis of TRAIL

induced apoptosis model during pre-MOMP (before 1.76 hours), Figure S4: MDFP analysis of TRAIL induced apoptosis model during post-MOMP (after 1.76 hours), Figure S5: MDFP analysis of non-apoptotic Hela subpopulation, Figure S6: Spearman correlations of MDFP sensitivity coefficients using different distribution distances and perturbation sizes.

Acknowledgments: TMP was supported by the Singapore Millennium Foundation scholarship. The authors would like to acknowledge funding from ETH Zurich as well as support from Sage Bionetworks.

Author Contributions: T.M.P. and R.G. conceived and designed the study; T.M.P. performed the computational work; T.M.P. and R.G. analyzed the data; T.M.P. and R.G. wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

1. Cahan, P.; Daley, G. Q. Origins and implications of pluripotent stem cell variability and heterogeneity. *Nat. Rev. Mol. Cell Biol.* **2013**, *14*, 357–368, doi:10.1038/nrm3584.
2. Flusberg, D. A.; Sorger, P. K. Surviving apoptosis: life-death signaling in single cells. *Trends Cell Biol.* **2015**, *25*, 446–458, doi:10.1016/j.tcb.2015.03.003.
3. Xia, X.; Owen, M. S.; Lee, R. E. C.; Gaudet, S. Cell-to-cell variability in cell death: can systems biology help us make sense of it all? *Cell Death Dis.* **2014**, *5*, e1261, doi:10.1038/cddis.2014.199.
4. Golding, I.; Paulsson, J.; Zawilski, S. M.; Cox, E. C. Real-time kinetics of gene activity in individual bacteria. *Cell* **2005**, *123*, 1025–1036, doi:10.1016/j.cell.2005.09.031.
5. Raj, A.; Peskin, C. S.; Tranchina, D.; Vargas, D. Y.; Tyagi, S. Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **2006**, *4*, e309, doi:10.1371/journal.pbio.0040309.
6. Raj, A.; van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **2008**, *135*, 216–226, doi:10.1016/j.cell.2008.09.050.
7. Stamakis, M. Cell population balance and hybrid modeling of population dynamics for a single gene with feedback. *Comp. Chem. Eng.* **2013**, *53*, 25–34, doi:10.1016/j.compchemeng.2013.02.006.
8. Hasenauer, J.; Hasenauer, C.; Hucho, T.; Theis, F.J. ODE constrained mixture modelling: A method for unraveling subpopulation structures and dynamics. *PLoS Comput. Biol.* **2014**, *10*, e1003686, doi: 10.1371/journal.pcbi.1003686.
9. Henson, M. A. Dynamic modeling of microbial cell populations. *Curr. Opin. Biotechnol.* **2003**, *14*, 460–467.
10. Hasty, J.; Pradines, J.; Dolnik, M.; Collins, J.J. Noise-based switches and amplifiers for gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 2075–2080, doi:10.1073/pnas.040411297.
11. Manninen, T; Linne, M. L.; Ruohonen, K. Developing Ito stochastic differential equation models for neuronal signal transduction pathways. *Comput. Biol. Chem.* **2006**, *30*, 280–291.
12. Samoilov, M. S.; Arkin, A. P. Deviant effects in molecular reaction pathways. *Nat. Biotech.* **2006**, *24*, 1235–1240, doi:10.1038/nbt1253.
13. Shahrezaei, V.; Swain, P. S. Analytical distributions for stochastic gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17256–17261, doi:10.1073/pnas.0803850105.
14. Sherman, M. S.; Lorenz, K.; Lanier, M. H.; Cohen, B. A. Cell-to-cell variability in the propensity to transcribe explains correlated fluctuations in gene expression. *Cell Syst.* **2015**, *1*, 315–325, doi:10.1016/j.cels.2015.10.011. Boettiger, A. N. Analytic approaches to stochastic gene expression in multicellular systems. *Biophys. J.* **2013**, *105*, 2629–2640, doi:10.1016/j.bpj.2013.10.033.
15. Saltelli, A.; Ratto, M.; Andres, T.; Campolongo, F.; Cariboni, J.; Gatelli, D.; Saisana, M.; Tarantola, S. *Global Sensitivity Analysis. The Primer*; John Wiley & Sons, Hoboken, NJ, 2008.

16. Kucherenko, S.; Rodriguez-Fernandez, M.; Pantelides, C.; Shah, N. Monte Carlo evaluation of derivative-based global sensitivity measures. *Reliab. Eng. Syst. Saf.* **2009**, *94*, 1135–1148, doi:10.1016/j.ress.2008.05.006.
17. Hafner, M.; Koeppl, H.; Hasler, M.; Wagner, A. "Glocal" Robustness Analysis and Model Discrimination for Circadian Oscillators. *PLoS Comput. Biol.* **2009**, *5*, e1000534, doi:10.1371/journal.pcbi.1000534.
18. Marino, S.; Hogue, I. B.; Ray, C. J.; Kirschner, D. E. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *J. Theor. Biol.* **2008**, *254*, 178–196, doi:10.1016/j.jtbi.2008.04.011.
19. Gunawan, R.; Cao, Y.; Petzold, L.; Doyle, F. J., III. Sensitivity analysis of discrete stochastic systems. *Biophys. J.* **2005**, *88*, 2530–2540, doi:10.1529/biophysj.104.053405.
20. Anderson, D. F. An Efficient Finite Difference Method for Parameter Sensitivities of Continuous Time Markov Chains. *SIAM J. Numer. Anal.* **2012**, *50*, 2237–2258, doi:10.1137/110849079.
21. Kim, D.; Debusschere, B. J.; Najm, H. N. Spectral methods for parametric sensitivity in stochastic dynamical systems. *Biophys. J.* **2007**, *92*, 379–393, doi:10.1529/biophysj.106.085084.
22. Komorowski, M.; Costa, M. J.; Rand, D. A.; Stumpf, M. P. H. Sensitivity, robustness, and identifiability in stochastic chemical kinetics models. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 8645–8650, doi:10.1073/pnas.1015814108.
23. Plyasunov, S.; Arkin, A. P. Efficient stochastic sensitivity analysis of discrete event systems. *J. Comput. Phys.* **2007**, *221*, 724–738, doi:10.1016/j.jcp.2006.06.047.
24. Rathinam, M.; Sheppard, P. W.; Khammash, M. Efficient computation of parameter sensitivities of discrete stochastic chemical reaction networks. *J. Chem. Phys.* **2010**, *132*, 034103, doi:10.1063/1.3280166.
25. Sheppard, P. W.; Rathinam, M.; Khammash, M. A pathwise derivative approach to the computation of parameter sensitivities in discrete stochastic chemical systems. *J. Chem. Phys.* **2012**, *136*, 034115, doi:10.1063/1.3677230.
26. Ingalls, B. Sensitivity analysis: from model parameters to system behaviour. *Essays Biochem.* **2008**, *45*, 177–193, doi:10.1042/BSE0450177.
27. Zi, Z. Sensitivity analysis approaches applied to systems biology models. *IET Syst. Biol.* **2011**, *5*, 336–346, doi:10.1049/iet-syb.2011.0015.
28. Perumal, T. M.; Wu, Y.; Gunawan, R. Dynamical analysis of cellular networks based on the Green's function matrix. *J. Theor. Biol.* **2009**, *261*, 248–259, doi:10.1016/j.jtbi.2009.07.037.
29. Perumal, T. M.; Gunawan, R. Understanding dynamics using sensitivity analysis: caveat and solution. *BMC Syst. Biol.* **2011**, *5*, 41, doi:10.1186/1752-0509-5-41.
30. Perumal, T. M.; Krishna, S. M.; Tallam, S. S.; Gunawan, R. Reduction of kinetic models using dynamic sensitivities. *Comput. Chem. Eng.* **2013**, *56*, 37–45, doi:10.1016/j.compchemeng.2013.05.003.
31. Spencer, S. L.; Gaudet, S.; Albeck, J. G.; Burke, J. M.; Sorger, P. K. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **2009**, *459*, 428–432, doi:10.1038/nature08012.
32. Varma, A.; Morbidelli, M.; Wu, H. *Parametric Sensitivity in Chemical Systems*; Cambridge Univ. Press, Cambridge, UK, 1999.
33. Peter D., H.; D., H. P. Kernel estimation of a distribution function. *Communications in Statistics - Theory and Methods* **1985**, *14*, 605–620, doi:10.1080/03610928508828937.
34. Niepel, M.; Spencer, S. L.; Sorger, P. K. Non-genetic cell-to-cell variability and the consequences for pharmacology. *Curr. Opin. Chem. Biol.* **2009**, *13*, 556–561, doi:10.1016/j.cbpa.2009.09.015.
35. Albeck, J. G.; Burke, J. M.; Aldridge, B. B.; Zhang, M.; Lauffenburger, D. A.; Sorger, P. K. Quantitative analysis of pathways controlling extrinsic apoptosis in single cells. *Mol. Cell* **2008**, *30*, 11–25,

447 doi:10.1016/j.molcel.2008.02.012.
448 36. Albeck, J. G.; Burke, J. M.; Spencer, S. L.; Lauffenburger, D. A.; Sorger, P. K. Modeling a snap-action,
449 variable-delay switch controlling extrinsic cell death. *PLoS Biol.* **2008**, *6*, 2831–2852,
450 doi:10.1371/journal.pbio.0060299.
451 37. Roux, J.; Hafner, M.; Barbara, S.; Sims, J. J.; Hudson, H.; Chai, D.; Sorger, P. K. Non-genetic origins of
452 cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **2009**, *459*, 428–432, doi:10.1038/nature08012.
453 38. Gaudet, S.; Spencer, S. L.; Chen, W. W.; Sorger, P. K. Exploring the contextual sensitivity of factors that
454 determine cell-to-cell variability in receptor-mediated apoptosis. *PLoS Comput. Biol.* **2012**, *8*, e1002482,
455 doi:10.1371/journal.pcbi.1002482.