**RESEARCH ARTICLE**                                                          **Open Access**

CrossMark

# Dynamic optimization of biological networks under parametric uncertainty

Philippe Nimmegeers, Dries Telen, Filip Logist and Jan Van Impe[*]

## Abstract

**Background:** Micro-organisms play an important role in various industrial sectors (including biochemical, food and pharmaceutical industries). A profound insight in the biochemical reactions inside micro-organisms enables an improved biochemical process control. Biological networks are an important tool in systems biology for incorporating microscopic level knowledge. Biochemical processes are typically dynamic and the cells have often more than one objective which are typically conflicting, e.g., minimizing the energy consumption while maximizing the production of a specific metabolite. Therefore multi-objective optimization is needed to compute trade-offs between those conflicting objectives. In model-based optimization, one of the inherent problems is the presence of uncertainty. In biological processes, this uncertainty can be present due to, e.g., inherent biological variability. Not taking this uncertainty into account, possibly leads to the violation of constraints and erroneous estimates of the actual objective function(s). To account for the variance in model predictions and compute a prediction interval, this uncertainty should be taken into account during process optimization. This leads to a challenging optimization problem under uncertainty, which requires a robustified solution.

**Results:** Three techniques for uncertainty propagation: linearization, sigma points and polynomial chaos expansion, are compared for the dynamic optimization of biological networks under parametric uncertainty. These approaches are compared in two case studies: *(i)* a three-step linear pathway model in which the accumulation of intermediate metabolites has to be minimized and *(ii)* a glycolysis inspired network model in which a multi-objective optimization problem is considered, being the minimization of the enzymatic cost and the minimization of the end time before reaching a minimum extracellular metabolite concentration. A Monte Carlo simulation procedure has been applied for the assessment of the constraint violations. For the multi-objective case study one Pareto point has been considered for the assessment of the constraint violations. However, this analysis can be performed for any Pareto point.

**Conclusions:** The different uncertainty propagation strategies each offer a robustified solution under parametric uncertainty. When making the trade-off between computation time and the robustness of the obtained profiles, the sigma points and polynomial chaos expansion strategies score better in reducing the percentage of constraint violations. This has been investigated for a normal and a uniform parametric uncertainty distribution. The polynomial chaos expansion approach allows to directly take prior knowledge of the parametric uncertainty distribution into account.

**Keywords:** Dynamic optimization, Optimization under uncertainty, Biological networks, Multi-objective

**Abbreviations:** CPU, Central processing unit; KKT, Karush-Kuhn-Tucker; LIN, Linearization; MC, Monte Carlo; NBI, Normal boundary intersection; NLP, Nonlinear programming problem; PCE, Polynomial chaos expansion; PCE1, First order polynomial chaos expansion; PCE2, Second order polynomial chaos expansion; SP, Sigma points

---

*Correspondence: jan.vanimpe@kuleuven.be
KU Leuven, Department of Chemical Engineering, BioTeC+ & OPTEC,
Gebroeders De Smetstraat 1, 9000 Ghent, Belgium

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 2 of 20

## Background

The application of micro-organisms in chemical industry and life sciences is paramount. In industrial biotechnology, on the one hand, microbial growth is stimulated in order to enhance the production of (high added value) chemical and pharmaceutical products. On the other hand, in food industry the aim is to avoid the growth of pathogens and food spoilage to ensure food safety.

Therefore, a profound biochemical insight in microbial dynamics and the reactions inside micro-organisms is important. Integrating insights obtained at systems biology (microscopic) level contributes to an improved (macroscopic level) biochemical process control (i.e., enabling advanced model based monitoring, control and optimization of bioprocesses) [1].

A basic tool in systems biology for incorporating microscopic level information are biological networks, e.g., metabolic reaction networks in which the knots represent the metabolites (chemical substances produced/consumed in the micro-organisms) and the connections indicate the mass fluxes between those metabolites. A biological network is a systematic representation of the cellular processes and the interactions between the molecules in the cells: e.g., proteins and metabolites. Such a network comprises (a subset of) all reactions which occur inside a cell and the knots represent the metabolites (i.e., products consumed/produced by the cells) and the links represent the intracellular reactions or reactions between the cell and its environment. A cell can be seen on microscopic scale as a combination of interactions between different layers: fluxome, metabolome, proteome, transcriptome and genome. In terms of network complexity (i.e., the number of metabolites and fluxes), fluxome level biological networks have the lowest level of complexity, while genome-scale biological networks have the highest level of complexity [2–4].

Insight in the dynamic behavior of micro-organisms can be obtained by simulation of metabolic networks. Optimization of biological networks can be used to analyze and also influence the regulation of pathways, e.g., to stimulate the production of high added value products. In practice, cells often have more than one objective, which are conflicting, e.g., minimizing the energy consumption while maximizing the production of a certain metabolite. Therefore, dynamic (multi-objective) optimization, which provides optimal (possibly time-varying) control profiles, is an important tool. Multi-objective optimization of biological networks has been investigated in [5–7]. The multi-objective design of bioprocesses and solution strategies have for instance been presented in [8] with application to a well-stirred, aerobic fermentor in which *Saccharomyces cerevisiae* grows in a medium of sugar cane molasses.

However, in practice, uncertainty on the model parameters and external process disturbances are inherently present. Uncertainty can originate from unmodeled process variables (*process noise*), e.g., inherent biological variability between cells which are genetically identical [9] or from a parameter estimation procedure based on noisy measurements (*measurement noise*), such that the true parameter values (which are different from the model parameters) are unknown. Not taking this uncertainty into account, possibly leads to the violation of constraints and erroneous estimates of the actual objective function(s). Therefore, the information about the uncertainty has to be taken into account to obtain *robustified controls* (i.e., variables that can be manipulated throughout the process) that ensure that constraints are met and an overall better objective function estimate is guaranteed. In this work the nature of uncertainty is assumed to be *stochastic*, i.e., following a probability distribution, and the uncertainty is modeled in the model parameters, i.e., *parametric uncertainty* [10].

Including robustness in an optimization problem is often tedious, since this typically leads to semi-infinite optimization problems that are challenging to solve in practice [10]. Three methods are compared in this work to approximately solve the *(multi-objective) dynamic optimization problem under parametric uncertainty* for biological networks: linearization [11], sigma points [12] and polynomial chaos expansion [13, 14]. Each of these methods requires increasing levels of information on the parametric uncertainty distribution to propagate the parametric uncertainty towards the states, constraints or objectives of interest.

The authors want to highlight that enzyme activation in biological networks has been studied in terms of dynamic optimization, single objective as well as multi-objective. In this work, for the first time, parametric uncertainty is taken into account for prediction and control of biological networks. Another novelty is the critical comparison of the linearization, sigma points and polynomial chaos expansion approaches for dynamic optimization of biological networks under uncertainty. Single objective as well as multi-objective optimization case studies have been investigated in this work. Therefore the general formulations in this work have been presented for multi-objective optimization problems.

The paper is structured as follows. In the 'Methods' section the multi-objective dynamic optimization problem formulation under parametric uncertainty is first presented. Then the concept of uncertainty propagation is introduced, together with the three applied approximation techniques for uncertainty propagation. Subsequently, multi-objective optimization methods are briefly discussed. To conclude this section the software and case studies are presented. A validation and assessment of

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 3 of 20

the approximation techniques for uncertainty propagation based on the case studies is presented in the 'Results and discussion' section, together with a physical/biological interpretation. Finally, the 'Conclusions' section summarizes the main results of this work.

## Methods

In this section the robustified multi-objective dynamic optimization formulation is presented. Subsequently, the different approximation techniques for uncertainty propagation that enable a robustified dynamic optimization under parametric uncertainty are discussed. Next, the approach for the Monte Carlo simulations is presented. In addition, multi-objective optimization methods are introduced, followed by a brief discussion on the software used in this work. To conclude the case studies are presented.

### Multi-objective dynamic optimization under parametric uncertainty

Consider the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}, t)$, with $\mathbf{x} \in \mathbb{R}^{n_x}$ the state vector (e.g., metabolite concentrations), $\mathbf{u} \in \mathbb{R}^{n_u}$, the control vector (e.g., enzyme expression rates), $\boldsymbol{\theta} \in \mathbb{R}^{n_\theta}$ the vector containing the uncertain parameters (e.g., kinetic constants such as the maximum reaction rate) and $t$ the time. The aim of a multi-objective dynamic optimization problem is to design a control, which minimizes several objective functions $\{J_1, \ldots, J_{n_J}\}$, subject to the constraints (i.e., model as dynamic constraint and other constraints). The multi-objective dynamic optimization problem in the time interval $t \in [0, t_f]$ and constraints $\mathbf{c}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}, t) \in \mathbb{R}^{n_c}$ (e.g., bounds on the metabolite concentrations or fluxes for cell viability) is formulated as in Eq. (1).

$$\min_{\mathbf{u}, \mathbf{x}, t_f} \{J_1, \ldots, J_{n_J}\}$$

$$\text{s.t.} \begin{cases} \dot{\mathbf{x}} & = \mathbf{f}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}, t) \\ \mathbf{x}(0) & = \mathbf{x}_0 \\ \mathbf{0} & \geq \mathbf{c}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}, t) \end{cases} \quad (1)$$

An inherent problem in the modeling of biological processes is uncertainty. This uncertainty can originate from model uncertainty and external disturbances [10]. The emphasis in this work is on parametric uncertainty, i.e., the uncertainty is present in several model parameters, which for instance can originate from biological variability. Not taking this uncertainty into account can possibly lead to constraint violations or erroneous estimates of the actual objective function of the process. In the field of robust optimization these uncertainties are taken into account to guarantee that critical constraints are not violated [10].

If knowledge about the parametric uncertainty distribution is present, expected values for the states and chance constraints can be formulated [15]. Chance constraints express that the probability of a constraint to be valid must be larger than a specific value [16, 17].

Consider that the constraints $\mathbf{0} \geq \mathbf{c}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}, t)$ can be replaced by $n_{c_{\text{prob}}}$ chance constraints $c_{\text{prob},i}$, expressing that the probability that a constraint is satisfied is larger than a preset probability $\beta_i$, with $i = 1, \ldots, n_{c_{\text{prob}}}$. In this work only single chance constraints are considered.

$$\beta_i \leq \mathbf{Pr}\left[\mathbf{0} \geq c_{\text{prob},i}(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}, t)\right] \quad (2)$$

If the uncertainty is fully known within a specific bounded set, the optimization problem is solved for the *worst-case scenario* in which all constraints have to be satisfied [15]. This approach typically leads to minmax problems which are hard to solve [18]. Since the worst-case scenario is often highly unlikely to occur, this approach can lead to poor results [11]. In order to solve this, a tradeoff between the nominal case (i.e., the non-robustified case in which uncertainty is not taken into account and the nominal parameter values are used) and worst-case scenario can be made [15].

The main limitation of the dynamic optimization problem with chance constraints is solving the problem in a computationally efficient way. The propagation of the parameter uncertainties through the nonlinear model and obtaining computationally tractable expressions for the dynamic optimization problem with chance constraints remains challenging [19]. Therefore, the chance constraints can be approximated by deterministic constraints as in Eq. (3).

$$0 \geq \mathbf{E}\left[c_{\text{prob,i}}\right] + \alpha_{c_{\text{prob},i}}\sqrt{\mathbf{Var}\left[c_{\text{prob,i}}\right]} \quad (3)$$

In Eq. (3), $\mathbf{E}\left[c_{\text{prob,i}}\right]$ and $\mathbf{Var}\left[c_{\text{prob,i}}\right]$ express the expected value and variance of the chance constraint function $c_{\text{prob,i}}$, respectively. The coefficient $\alpha_{c_{\text{prob},i}}$ is introduced as a *backoff parameter* (e.g., [11, 20]) to take the uncertainty on the chance constraints into account. The choice of the backoff can for instance correspond to a probability that the specific constraint is violated, i.e., so-called *single chance constraints*.

A first way to choose the backoff parameter $\alpha_{c_{\text{prob},i}}$ is with Cantelli-Chebyshev's inequality. In [17] it is shown that an upper bound for the expected value on an individual chance constraint can be calculated. This equation holds for any underlying distribution of the chance constraint. Computing the backoff parameter via Cantelli-Chebyshev's inequality for a probability of 95 % for the chance constraint to be satisfied, results in a backoff parameter $\alpha_{c_{\text{prob},i}} = 4.36$, while for a normal probability distribution the backoff parameter would be 1.96.

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 4 of 20

From this, it is clear that Cantelli-Chebyshev's inequality generally leads to a very conservative bound with too high backoff parameter values for use in practice, leading potentially to infeasibilities.

If a probability distribution is assumed for the considered constraint(s) or objective function(s), a second way is to choose the backoff parameter based on the quantiles. For this a procedure as in [10] can be followed to obtain a desired confidence level for the constraint to be satisfied or to cover the objective function in a prediction interval. In this work the choice of the backoff parameter is based on the quantiles, assuming that the states follow a normal distribution, as shown in Table 1.

Similarly to the reformulation of constraints, the objective function $J_i$ can be reformulated by adding the term $\alpha_{J_i}\sqrt{\mathbf{Var}\,[J_i]}$. Since an objective function has to be minimized, an increase in variance is penalized by this reformulation. The reformulated robustified multi-objective dynamic optimization problem with deterministic constraints is formulated in Eq. (4).

$$\min_{\mathbf{u},\mathbf{x},t_f} \left\{ \mathbf{E}\,[J_1] + \alpha_{J_1}\sqrt{\mathbf{Var}\,[J_1]}, \ldots, \mathbf{E}\,[J_{n_J}] + \alpha_{J_{n_J}}\sqrt{\mathbf{Var}\,[J_{n_J}]} \right\}$$

$$\text{s.t.} \begin{cases} \dot{\mathbf{x}} & = \mathbf{f}(\mathbf{x},\mathbf{u},\boldsymbol{\theta},t) \\ \mathbf{x}(0) & = \mathbf{x}_0 \\ 0 & \geq \mathbf{E}\,[c_{\text{prob},i}] + \alpha_{c_{\text{prob},i}}\sqrt{\mathbf{Var}\,[c_{\text{prob},i}]} \\ & \quad i = 1, \ldots, n_{c_{\text{prob}}} \end{cases}$$

(4)

In practice, not all constraints have to be replaced by probabilistic constraints (e.g., bounds on the controls and states) and constraints of the form $\mathbf{0} \geq \mathbf{c}(\mathbf{x},\mathbf{u},\boldsymbol{\theta},t)$ can still be present in the optimization problem formulation.

**Approximation techniques for uncertainty propagation**
In this paper, the parametric uncertainty is propagated to the states, constraints and objectives of interest. This can be illustrated with an example shown in Fig. 1. Consider the simple nonlinear model $y = g(x)$ (blue curve), with the parameter x that is uncertain with a known parametric uncertainty distribution (green curve). The principle of uncertainty propagation will propagate the parametric uncertainty distribution (green curve) through the nonlinear model (blue curve) in order to obtain the uncertainty distribution of the output $y$ (purple curve).

**Table 1** Backoff parameter $\alpha_i$ with corresponding quantiles and confidence levels

| $\alpha_i$ | 0.84 | 1.28 | 1.65 | 1.96 |
|---|---|---|---|---|
| Quantile | 0.20 | 0.10 | 0.05 | 0.025 |
| Confidence level | 0.80 | 0.90 | 0.95 | 0.975 |

The parametric uncertainty can be propagated via a numerical integration over the parameter distribution [21]. However, this integration is typically computationally expensive for realistic models and more efficient approximative uncertainty propagation techniques exist.

An alternative to this numerical integration is the use of Monte Carlo simulations in optimization. A large number of $N$ realizations is drawn from the assumed parametric uncertainty distribution with variance-covariance matrix $\boldsymbol{\Sigma}$ and the empirical confidence regions can be determined by using the appropriate quantiles. However, this is in practice computationally extremely expensive. Due to the large amount of simulations, no computationally tractable procedure for gradient based optimization schemes is available. In addition there is no clear rule on how many noise realizations have to be taken in order to obtain an accurate estimate [10]. For these reasons, Monte Carlo simulations are not pursued in the dynamic optimization procedure.

Approaches that exploit the availability of measurements as described in [11, 22, 23] are also used in the field of robust optimization. However, these are not considered in this paper, since in an industrial setting intracellular measurements are typically not available on a routine basis.

The first type of the employed techniques is a so-called *linearization approach*, which is based on first-order Taylor series approximations of the model functions with respect to the uncertainty [10, 11]. This approximation can be used if higher order terms can be neglected. This is the case when the uncertainty is small compared to the model curvature [15]. In this linearization approach a linear approximation of the variance-covariance matrix [11] of the states is made. On the other hand efficient sampling-based uncertainty propagation techniques exist as, e.g., using Hammersley sequences [24], the unscented transformation or *sigma points approach* [12] and the *polynomial chaos expansion approach* [19, 25].

In addtition to the *linearization approach* [11], two other techniques are considered: the *sigma points approach* [12] and the *polynomial chaos expansion approach* [19].

In practice, one is often interested in the violation of a path constraint (e.g., fluxes or concentrations that should not exceed their bounds for cell viability), a terminal constraint (e.g., a minimum amount of a specific metabolite to be produced) or the robustness of the objective (e.g., a minimum enzymatic cost), i.e., minimizing the uncertainty on the objective by taking into account the variance on the objective. The constraint or objective function to which the uncertainty is propagated, is denoted by $R_k(\mathbf{x},\mathbf{u},\boldsymbol{\theta},t)$ in the following. The three techniques consist of propagating the parametric uncertainty by approximating the expected value $\mathbf{E}\,[R_k]$ and variance $\mathbf{Var}\,[R_k]$ of

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 5 of 20
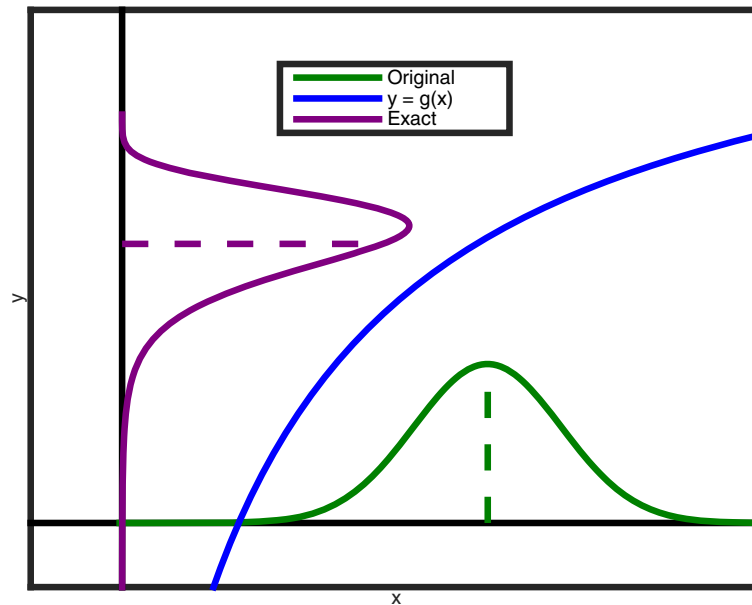


**Fig. 1** Principle of uncertainty propagation. Uncertainty propagation of *x* towards *y* via nonlinear transformation *g(x)* in an exact way

$R_k(\mathbf{x}, \mathbf{u}, \boldsymbol{\theta}, t)$. The approximated expected value and variance are denoted by $\bar{R}_{k,\mathrm{LIN}}$, $\bar{R}_{k,\mathrm{SP}}$, $\bar{R}_{k,\mathrm{PCE}}$ and $\mathbf{P}_{\mathrm{LIN}}$, $\mathbf{P}_{\mathrm{SP}}$, $\mathbf{P}_{\mathrm{PCE}}$, respectively.

An overview of these techniques with the robustified multi-objective dynamic optimization problem formulation is provided in Table 2. A graphical representation of the approximation techniques for uncertainty propagation is shown in Fig. 2. A more detailed review of these techniques is presented in Additional file 1. Note that these approaches are also applicable to uncertainty present in the model right hand side. However, in this work only the application to parametric uncertainty propagation is considered.

**Note:** The approximation techniques for uncertainty propagation can also be used to quantify the effect of parametric uncertainty on model predictions. If the controls are fixed, then an additional benefit could be that the uncertainty on the states can be displayed. This enhances the insight in whereto the system can evolve with an a priori known or assumed parametric uncertainty. In fact, in literature it has been pointed out that the polynomial chaos expansion could be considered as an alternative to Monte Carlo simulations, but with less computational cost [25].

### Multi-objective optimization methods

In practice, multiple objectives, which are very often conflicting with each other, have to be considered simultaneously, e.g., minimizing the enzyme consumption while maximizing the production of a certain metabolite. Therefore, a single optimal solution will not exist, but a set

of *trade-off solutions*, called the Pareto front, is obtained when solving a multi-objective problem [16].
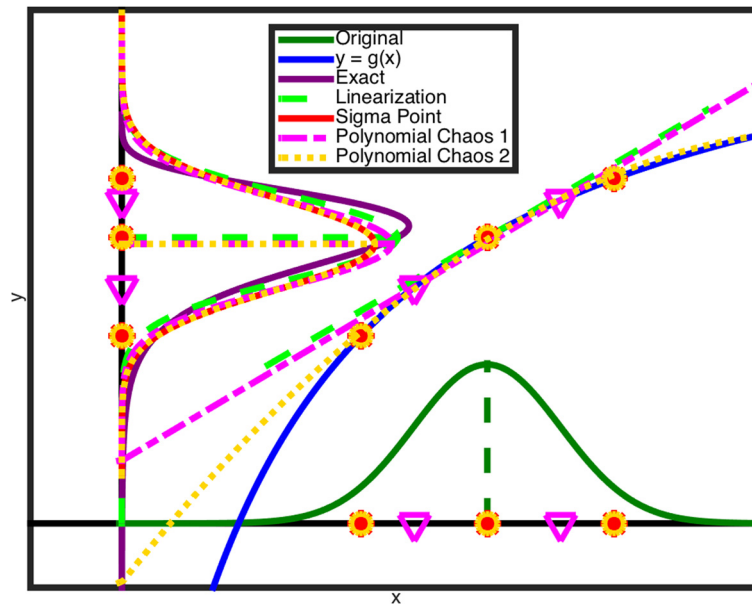
Two categories of methods can be distinguished for the calculation of Pareto fronts: *scalarization* methods ([26–30]) that convert the multi-objective optimization problem into a series of single objective optimization problems by using scalar variables, and *vectorization* methods [16, 32–34] that start from a population of candidate solutions that gradually evolve to the Pareto front. Scalarization methods can take advantage of fast and efficient gradient based methods to find an optimum for the series of single objectives, while vectorization methods often use derivative-free optimization methods as evolutionary or stochastic optimization approaches. In this work the Normal Boundary Intersection (NBI) method [30, 31], i.e., a scalarization method, is used. For a more detailed description of the frame of multi-objective dynamic optimization and the NBI method see, e.g., [35] and Additional file 2.

### Implementation and software

The dynamic optimization problems in this work are solved using a direct approach, in which first the optimal control problem is discretized into a nonlinear optimization problem that can be solved afterwards with NLP solvers. It is chosen to discretize the problems using an orthogonal collocation discretization scheme. The rationale of orthogonal collocation is that the states and controls are fully discretized with respect to time in finite elements. Per finite element there are four *collocation points* of which the first one is fixed and the three other

**Table 2** Overview of the approximation techniques for uncertainty propagation with $R_k$ the variable to which the parametric uncertainty is propagated

| | Linearization | Sigma points | Polynomial chaos |
|---|---|---|---|
| Rationale | Linearization of state equations around $\bar{\boldsymbol{\theta}}$ | Approximate distribution by a fixed number of parameters (sigma points) | Approximate response of the model (at sampling points) as a $p$th order polynomial function of $\boldsymbol{\theta}$ |
| Uncertainty distribution | Normal | Any symmetric, unimodal distribution | Any |
| Equations | State equations + Sensitivity equations: $$\begin{cases} \dot{\mathbf{S}}_{\text{LIN}}(t) = \frac{\partial \mathbf{f}(\mathbf{x},\mathbf{u},\boldsymbol{\theta}_{\text{nom}},t)}{\partial \mathbf{x}}\mathbf{S}_{\text{LIN}} + \frac{\partial \mathbf{f}(\mathbf{x},\mathbf{u},\boldsymbol{\theta}_{\text{nom}},t)}{\partial \boldsymbol{\theta}}, \\ \mathbf{S}_{\text{LIN}}(0) = \mathbf{0} \end{cases}$$ | State equations for SPs: $$\dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x_i},\mathbf{u},\boldsymbol{\pi_i},t) \text{ with } i=0,\ldots,2n_\theta$$ | State equations for sampling points: $$\dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x_i},\mathbf{u},\boldsymbol{\pi_i},t) \text{ with } i=0,\ldots,n_s-1$$ |
| Total $n_{states}$ | $(n_\theta+1)n_x$ | $(2n_\theta+1)n_x$ | $\frac{(n_\theta+p)!}{n_\theta!p!}n_x$ |
| Sampling points | – | Sigma points: $2n_\theta+1$ | Collocation points: $\frac{(n_\theta+p)!}{n_\theta!p!}$ |
| Expected value of $R_k$ | $R_k$ | $\frac{1}{n_\theta+\kappa}\left(\kappa R_k(\boldsymbol{\pi}_0) + \frac{1}{2}\sum_{i=1}^{2n_\theta} R_k(\boldsymbol{\pi}_i)\right)$ | $a_{R_k,0}^{(p)}$ |
| Variance on $R_k$ | $\mathbf{P}_{R_kR_k,\text{LIN}} = \frac{\partial R_k}{\partial \mathbf{x}}\mathbf{P}_{\text{LIN}}\left(\frac{\partial R_k}{\partial \mathbf{x}}\right)^\top$ with: $\mathbf{P}_{\text{LIN}} = \mathbf{S}_{\text{LIN}}(t)\boldsymbol{\Sigma}\mathbf{S}_{\text{LIN}}(t)^\top$ | $\mathbf{P}_{R_kR_k,\text{SP}} = \frac{1}{n_\theta+\kappa}\left(\kappa(R_k(\boldsymbol{\pi}_0)-\bar{R}_k)(R_k(\boldsymbol{\pi}_0)-\bar{R}_k)^\top\right)$ $+\frac{1}{n_\theta+\kappa}\left(\frac{1}{2}\sum_{i=1}^{2n_\theta}(R_k(\boldsymbol{\pi}_i)-\bar{R}_k)(R_k(\boldsymbol{\pi}_i)-\bar{R}_k)^\top\right)$ with: $\begin{cases} \boldsymbol{\pi_0} = \boldsymbol{\theta}_{\text{nom}}, \\ \boldsymbol{\pi_i} = \boldsymbol{\theta}_{\text{nom}} + \sqrt{(n_\theta+\kappa)\boldsymbol{\Sigma}}_i \text{ with } i=1,\ldots,n_\theta, \\ \boldsymbol{\pi_i} = \boldsymbol{\theta}_{\text{nom}} - \sqrt{(n_\theta+\kappa)\boldsymbol{\Sigma}}_{i-n_\theta} \text{ with } i=n_\theta+1,\ldots,2n_\theta. \\ \kappa = 3 - n_\theta \end{cases}$ | $\mathbf{P}_{R_kR_k,\text{PCE}}^{(p)} = \sum_{j=1}^{L-1}\left(a_{R_kj}^{(p)}\right)^2\mathbf{E}\left[\Phi_j^2(\theta)\right]$ with: $\mathbf{a} = (\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top)^{-1}\boldsymbol{\Lambda}\mathbf{R}_{k,s}$ |
| Optimization problem | $\min_{\mathbf{u},\mathbf{x},t_f}\{J_1,\ldots,J_{n_J}\}$ $\bar{J}_{i\text{LIN}} + \alpha_{J_i}\sqrt{\mathbf{P}_{J_iJ_i,\text{SP}}}$ s.t. $\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u},\boldsymbol{\theta},t) \\ \dot{\mathbf{S}}_{\text{LIN}}(t) = \frac{\partial \mathbf{f}(\mathbf{x},\mathbf{u},\boldsymbol{\theta}_{\text{nom}},t)}{\partial \mathbf{x}}\mathbf{S}_{\text{LIN}} + \frac{\partial \mathbf{f}(\mathbf{x},\mathbf{u},\boldsymbol{\theta}_{\text{nom}},t)}{\partial \boldsymbol{\theta}}, \\ \mathbf{S}_{\text{LIN}}(0) = \mathbf{0} \\ \mathbf{x}(0) = \mathbf{x}_0 \\ 0 \geq \bar{c}_{\text{prob},i,\text{LIN}} \\ \qquad +\alpha_{c_{\text{prob},i}}\sqrt{\mathbf{P}_{c_{\text{prob},i},c_{\text{prob},i},\text{LIN}}} \end{cases}$ | $\min_{\mathbf{u},\mathbf{x},t_f}\{J_1,\ldots,J_{n_J}\}$ $\bar{J}_{i\text{SP}} + \alpha_{J_i}\sqrt{\mathbf{P}_{J_iJ_i,\text{SP}}}$ s.t. $\begin{cases} \dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x_i},\mathbf{u},\boldsymbol{\pi_i},t) \text{ with } i=0,\ldots,2n_\theta \\ \mathbf{x}(0) = \mathbf{x}_0 \\ 0 \geq \bar{c}_{\text{prob},i,\text{SP}} \\ \qquad +\alpha_{c_{\text{prob},i}}\sqrt{\mathbf{P}_{c_{\text{prob},i},c_{\text{prob},i},\text{SP}}} \end{cases}$ | $\min_{\mathbf{u},\mathbf{x},t_f}\{J_1,\ldots,J_{n_J}\}$ $\bar{J}_{i\text{PCE}}^{(p)} + \alpha_{J_i}\sqrt{\mathbf{P}_{J_iJ_i,\text{PCE}}^{(p)}}$ s.t. $\begin{cases} \dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x}_i,\mathbf{u},\pi_i,t) \\ \qquad \text{with } i=0,\ldots,n_s-1 \\ \mathbf{x}(0) = \mathbf{x}_0 \\ \mathbf{R}_{\mathbf{k},s}^{(p)} = \left(\Lambda^{(p)}\right)^\top\mathbf{a}_{\mathbf{R}_\mathbf{k}}^{(p)} \\ \qquad \text{with } k=1,\ldots,n_R \\ 0 \geq \bar{c}_{\text{prob},i,\text{PCE}}^{(p)} \\ \qquad +\alpha_{c_{\text{prob},i}}\sqrt{\mathbf{P}_{c_{\text{prob},i},c_{\text{prob},i},\text{PCE}}^{(p)}} \\ \qquad \text{with } i=1,\ldots,n_{c_{\text{prob}}} \end{cases}$ |

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 7 of 20



**Fig. 2** Approximation techniques for uncertainty propagation. Overview of the linearization, sigma points and polynomial chaos expansion methods for uncertainty propagation of $x$ towards $y = g(x)$

ones should obey the model equations and are seen as equality constraints (i.e., so-called *collocation constraints*). Between each finite element there is also a constraint that ensures continuity (i.e., so-called *continuity constraints*). As interpolation between the *collocation points* a cubic Lagrange polynomial is used, with four collocation points situated at the Radau roots on each interval. State bounds are easily added in this technique. The fact that orthogonal collocation has hardly any problem with stiff systems is advantageous in case of numerically unstable systems.

An inhouse developed software package, called `Pomodoro`, is used for the implementation of both case studies. The `Pomodoro` software contains a collection of algorithms and tools for dynamic optimization and is implemented in `Python`. `Pomodoro` uses `CasADi` [36] as a backbone for the dynamic optimization problem formulation. `CasADi` is a software package for rapid prototyping of large-scale optimization problems with automatic differentiation using a symbolic/numeric approach. For solving the NLP, an interior point algorithm, `IPOPT` [37], has been used. The `Pomodoro` software can be downloaded from https://perswww.kuleuven.be/~u0093798/software.php. For review purposes the work describing `Pomodoro` (Bhonsale SS, Telen D, Vercammen D, Vallerio M, Hufkens J, Nimmegeers P, Logist F, Van Impe J. Pomodoro - A novel toolkit for (multiobjective) dynamic optimization, model based control and estimation, submitted), can be found on http://www.student.kuleuven.be/~s0212066/pomodoro/. For more information on the optimization methods and

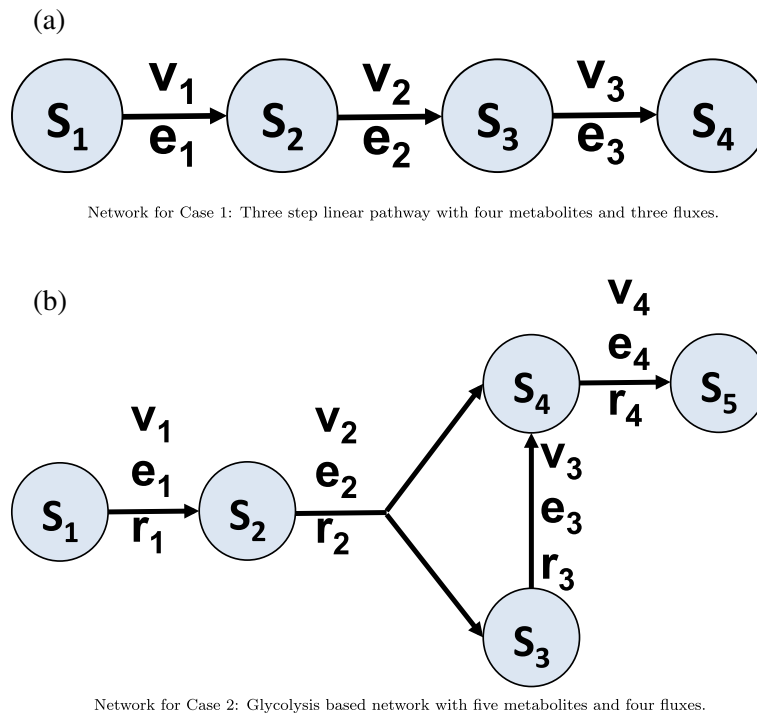implementation, the reader is referred to Additional file 2 and [35].

## Case studies

Two case studies are considered: *(a)* a three-step linear pathway with mass-action kinetics [7, 38] and *(b)* a glycolysis based network with 1 output [7, 39]. It should be noted that the models for the case study are partially taken from [7]. The networks for the two case studies are presented in Fig. 3.

### Case 1: three-step linear pathway

The first case study is a three-step linear pathway producing one product $S_4$ from a buffered substrate $S_1$ [7, 38]. This pathway consists of three enzymatic reactions (with reaction fluxes $\mathbf{v} = [v_1 \ v_2 \ v_3]^\top$.) following mass-action kinetics, between 4 metabolites $\mathbf{S} = [S_1 \ S_2 \ S_3 \ S_4]^\top$. Each reaction is catalyzed by a specific enzyme $e_i$. The first metabolite $S_1$ is the substrate, intermediate metabolites are $S_2$ and $S_3$, while $S_4$ is the product, produced by this three step linear pathway. The substrate is considered to be buffered, which means that the substrate concentration remains constant.

The model contains four differential states and an additional state $x_{\text{extra}}$ for the objective function corresponding to the integral of the intermediate accumulation. The model together with its constraints is presented in Eq. (5). The first constraint is to ensure that a minimum amount of product is obtained at $t_{\text{f}}$. The second constraint expresses that the sum of all enzyme concentrations

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 8 of 20



(a)

Network for Case 1: Three step linear pathway with four metabolites and three fluxes.

(b)

Network for Case 2: Glycolysis based network with five metabolites and four fluxes.

**Fig. 3** Networks for Case 1 and Case 2. Biological networks based on [7]: **a** Three step linear pathway with four metabolites and three fluxes and **b** Glycolysis based network with five metabolites and four fluxes

cannot exceed the total enzymatic concentration $E_T$ of 1 mM. The enzyme concentrations $\mathbf{e} = [e_1\ e_2\ e_3]^\top$ are the controls in this case study.

$$\begin{cases} \frac{d\mathbf{S}}{dt} & = \mathbf{N}\mathbf{v} \\ \frac{dx_{\text{extra}}}{dt} & = S_2 + S_3 \\ S_4(t_f) & \geq 0.90\text{mM} \\ \sum_{i_e=1}^{n_e=3} e_i & \leq E_T \end{cases} \quad (5)$$

with:

$$v_1 = k_1 \cdot S_1 \cdot e_1 \quad (6)$$
$$v_2 = k_2 \cdot S_2 \cdot e_2 \quad (7)$$
$$v_3 = k_3 \cdot S_3 \cdot e_3 \quad (8)$$

$$\mathbf{N} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \quad (9)$$

with $\mathbf{N} \in \mathbb{R}^{4\times3}$ the stoichiometric matrix, containing the stoichiometric coefficients $N_{ij}$ of metabolite $i$ in the $j$-th reaction and $k_j$, the maximum reaction rate of reaction $j$.

The three uncertain parameters are the three reaction rate constants $k_1$, $k_2$ and $k_3$, for which the nominal value equals 1 $(\text{mMs})^{-1}$. Since high concentrations of inter-

mediate metabolites $S_2$ and $S_3$ can be harmful for cell viability, the intermediate accumulation is minimized for this case study.

$$J = x_{\text{extra}}(t_f) = \int_{t_0=0s}^{t_f} S_2 + S_3 dt \quad (10)$$

*Case 2: glycolysis based network with 1 output*

The second case study is a glycolysis based network with the production of one product $S_5$, starting from one substrate $S_1$ from [7]. In this pathway four enzymatic reactions are taking place, each catalyzed by a specific enzyme. The fluxes are modeled with Michaelis-Menten kinetics. The intracellular metabolites in this network are $S_2$, $S_3$ and $S_4$. It is assumed that the substrate $S_1$ is buffered. This case study is particularly interesting due to the branch that is present. Such branches often occur in biological networks and the presented problem formulation can be modified/extended to many scenarios.

The expressions for this model are presented in Eq. (11) where $\mathbf{N}$ is the stoichiometric matrix, with $\mathbf{v} = [v_1\ v_2\ v_3\ v_4]^\top$ the flux vector, $\mathbf{S} = [S_1\ S_2\ S_3\ S_4\ S_5]^\top$ the vector containing the metabolite concentrations , $\mathbf{e} = [e_1\ e_2\ e_3\ e_4]^\top$ the enzyme concentration vector, the vector of manipulated variables $\mathbf{r} = [r_1\ r_2\ r_3\ r_4]^\top$ containing the expression rates, $k_{\text{cat},j}$ the maximum reaction rate for

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 9 of 20

reaction $j$, dependent on the enzyme that is catalyzing the reaction $j$ which is assumed to be the same for each reaction and therefore considered as the model parameter $k_{cat}$ and $K_M$ the Michaelis constant.

$$
\begin{cases}
\frac{d\mathbf{S}}{dt} & = \mathbf{Nv} \\
\frac{d\mathbf{e}}{dt} & = \mathbf{r} - \lambda\mathbf{e} \\
\frac{dx_{\text{extra}}}{dt} & = e_1 + e_2 + e_3 + e_4 \\
S_5(t_{\text{f}}) & \geq 0.675 \text{ mM} \\
\sum_{i_e=1}^{n_e=4} e_i & \leq E_T \\
r_1(t_0) & = 0.5 \\
r_{i_r}(t_0) & = 0, \sum_{i_r=1}^{n_r=4} r_{i_r} \leq 0.5 \quad i_r = 2, \ldots, n_r = 4
\end{cases}
\tag{11}
$$

with:

$$
v_1 = \frac{k_{cat} \cdot S_1}{K_M + S_1} \cdot e_1
\tag{12}
$$

$$
v_2 = \frac{k_{cat} \cdot S_2}{K_M + S_2} \cdot e_2
\tag{13}
$$

$$
v_3 = \frac{k_{cat} \cdot S_3}{K_M + S_3} \cdot e_3
\tag{14}
$$

$$
v_4 = \frac{k_{cat} \cdot S_4}{K_M + S_4} \cdot e_4
\tag{15}
$$

$$
\mathbf{N} = \begin{bmatrix}
0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 \\
0 & 1 & -1 & 0 \\
0 & 1 & 1 & -1 \\
0 & 0 & 0 & 1
\end{bmatrix}
\tag{16}
$$

The following values for the parameters are assumed [7]: $k_{cat} = 1 \text{ s}^{-1}, K_M = 1 \text{ mM}$ and $\lambda = 0.5 \text{ s}^{-1}$.

Two objectives are considered for which the enzyme expression rates in $\mathbf{r}$ are optimized: the minimization of the time to reach a given steady state and the enzyme consumption (or enzymatic cost) as shown in Eqs. (17)-(18):

$$
J_1 = t_{\text{f}}
\tag{17}
$$

$$
J_2 = \int_{t_0=0s}^{t_{\text{f}}} \left( \sum_{i_e=1}^{n_e=4} e_{i_e} dt \right).
\tag{18}
$$

## Results and discussion

This section discusses the obtained results in this work. In the first subsection the approach followed to obtain these results is clarified. Next, the results for the three-step linear pathway are presented. In the third subsection the results for the glycolysis inspired network are described.

## Approach

The approach consists of the four steps in Fig. 4. This approach is formulated for the generic case of the multi-objective dynamic optimization of biological networks under uncertainty. First, the (multi-objective) dynamic optimization problem is solved for the nominal (non-robustified) case. Then, desired confidence levels for the robustified constraint are set (*Step 1*). Robustified terminal constraints of the form $c_{\min} \leq \mathbf{E}\left[c(t_{\text{f}})\right] - \alpha_c \sqrt{\mathbf{Var}\left[c(t_{\text{f}})\right]}$ are considered. Table 1 presents the backoff parameter values that are used for the computation of the robustified controls together with the corresponding quantiles and preset confidence levels for a normal distribution of the constraint.
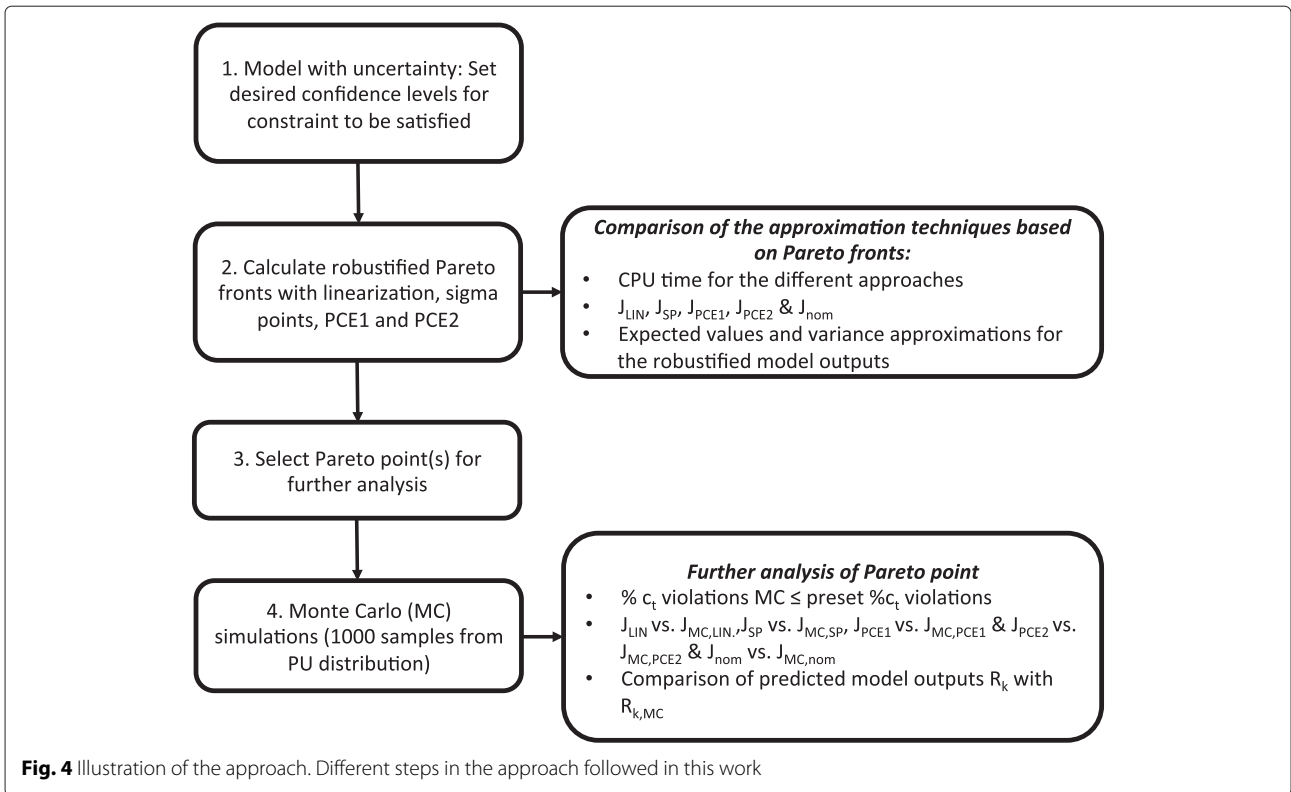
In *Step 2* robustified Pareto fronts are calculated with the linearization, sigma points and polynomial chaos expansion approaches (PCE1 and PCE2) to include parametric uncertainty in the multi-objective dynamic optimization problem. These robustified Pareto fronts are computed for the different backoff parameter values.

A first comparison of the approximation techniques for uncertainty propagation is based on the Pareto fronts: comparison of the CPU time for a single Pareto point, comparison of the objective function vectors $J_{\text{LIN}}$, $J_{\text{SP}}$, $J_{\text{PCE1}}$ and $J_{\text{PCE2}}$, calculated with the linearization, sigma points, PCE1 and PCE2 respectively. Also the expected values and variance approximations for the robustified model outputs with the approximation techniques for uncertainty propagation are compared.

Subsequently, a Pareto point is selected from the Pareto front for further analysis (*Step 3*). This analysis can be performed for any Pareto point or confidence level set (i.e., corresponding to different backoff parameter values $\alpha_i$), without a loss of generality. In this work, the considered point corresponds to one of the objectives, i.e., the minimization of the intermediate accumulation for Case 1 (single objective case study) and the minimization of the enzymatic cost for Case 2 (multi-objective case study).

In *Step 4*, Monte Carlo simulations are done for the considered Pareto point by sampling 1000 randomly generated parameter sets from the parametric uncertainty distribution. The robustified controls, determined with linearization, sigma points and polynomial chaos expansion approaches (PCE1 and PCE2), are fixed in the Monte Carlo simulations.

The further analysis of the Pareto point consists of assessing the robustness of the optimal control profiles obtained with the different approximation techniques for uncertainty propagation: i.e., *(i)* checking the reduction of constraint violations by applying a robustified control in comparison with applying a nominal (non-robustified) control profile, *(ii)* evaluating the backoff taken in objective function when uncertainty is taken into account and *(iii)* a comparison between the predicted expected value

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 10 of 20



**Fig. 4** Illustration of the approach. Different steps in the approach followed in this work

and variance for the robustified model output with the approximation techniques for uncertainty propagation and the calculated mean and variance with the Monte Carlo simulations. Furthermore, for the robustified single chance constraint it is investigated whether the preset confidence is reached for different backoff parameter values. This is done by checking whether the percentage of constraint violations in Monte Carlo simulations does not exceed the preset percentage of constraint violations. For instance, a confidence level of 0.95 is associated with a backoff parameter value of 1.65, meaning that in case the robustified control is applied, the percentage of constraint violations is not allowed to exceed 5 %. Alternatively, if the confidence level is not sufficient, the preset confidence can be increased by increasing the backoff parameter. An iterative procedure to determine the quantiles and backoff parameter can be followed as presented in [10].

Both case studies have been implemented in Pomodoro. The KKT tolerance is set to $10^{-5}$ and an orthogonal collocation discretization scheme is used for the dynamic optimization problems. Since the polynomial chaos expansion allows to take a priori information on the parametric uncertainty distribution directly into account via the orthogonal polynomials, two parametric uncertainty distributions have been studied: a priori normal and a priori uniform parametric uncertainty distribution with as mean the nominal parameter values and 20 %

relative standard deviation in Case 1 and 10 % relative standard deviation in Case 2. For the normal parametric uncertainty distribution, Hermite polynomials are used, while for the uniform distribution another type is used. The reader is referred to Additional files 3 and 4.

**Case 1: three step linear pathway**
In this case study the terminal constraint expressing that the concentration of $S_4$ at time $t_f$ should exceed or equal 0.90 mM is robustified.

$$0.90\text{mM} \leq \mathbf{E}\left[S_4(t_f)\right] - \alpha_{S_4}\sqrt{\mathbf{Var}\left[S_4(t_f)\right]} \qquad (19)$$

Since this constraint only looks at one bound that has to be exceeded, the 95 % confidence region should be covered when a backoff parameter of $\alpha_{S_4} = 1.65$ is chosen. The intermediate accumulation objective function is not robustified for this case study (i.e., $\alpha_{J_2}$). The three uncertain parameters are $k_1$, $k_2$ and $k_3$.

In this section, the single objective optimization (i.e., the minimization) of the intermediate accumulation is considered. Intermediate accumulation can be harmful for the cell and should therefore be minimized. It is assumed that the final time is fixed at 10 seconds. First the computational aspects are discussed. Subsequently a physical/biological interpretation of the results is given. To conclude Case 1 the approximation techniques for uncer-

tainty propagation are compared based on the results from the single objective optimization and Monte Carlo simulations for a normal and uniform parametric uncertainty distribution.

### Computational aspects

A first aspect is the computational cost of including uncertainty in dynamic optimization. The number of required states and variables, together with the CPU time for the different approximation techniques for uncertainty propagation are presented in Table 3. The CPU times are presented for the largest backoff parameter used in this work, i.e., $\alpha_{S_4} = 1.96$, to give an upper bound on the computation times that are required. The results in Table 3 confirm that taking uncertainty into account, leads to an increased computational time. An inherent property of the considered uncertainty propagation techniques, is the increase in number of states when the number of uncertain parameters increases. The increase in computational time for this case study is thus related to the increase in the size of the optimization problem. From this it is clear that the linearization and PCE1 approaches have a similar computation time and are the fastest of the considered approximation techniques for uncertainty propagation, followed by the sigma points approach and PCE2 approach.

### Physical/biological interpretation

The enzyme concentration profiles are (from a computational point of view) seen as the optimal controls. The enzyme concentration profiles are illustrated in Fig. 5(a)-(c) for $\alpha = 1.65$. Different phases can be distinguished in the process (i.e., a sequential activation of the controls): *(i)* a first phase in which enzyme $e_1$ is activated to produce $S_2$ as fast as possible, *(ii)* a second phase in which both $S_2$ and $S_3$ are consumed and produced (by activation of $e_2$ and $e_3$) *(iii)* a third phase in which a novel activation of $e_1$ takes place, followed by *(iv)* a phase of activity of $e_2$ and $e_3$ and a final phase in which the third enzyme is fully activated for the production of $S_4$.

**Table 3** Case 1 - Overview of the number of states, CPU time, expected values of the objective function, expected values of the terminal constraint and standard deviations for the different approximation techniques for uncertainty propagation when the enzymatic cost is minimized for $\alpha = 1.96$ for 3 uncertain parameters $k_1$, $k_2$ and $k_3$

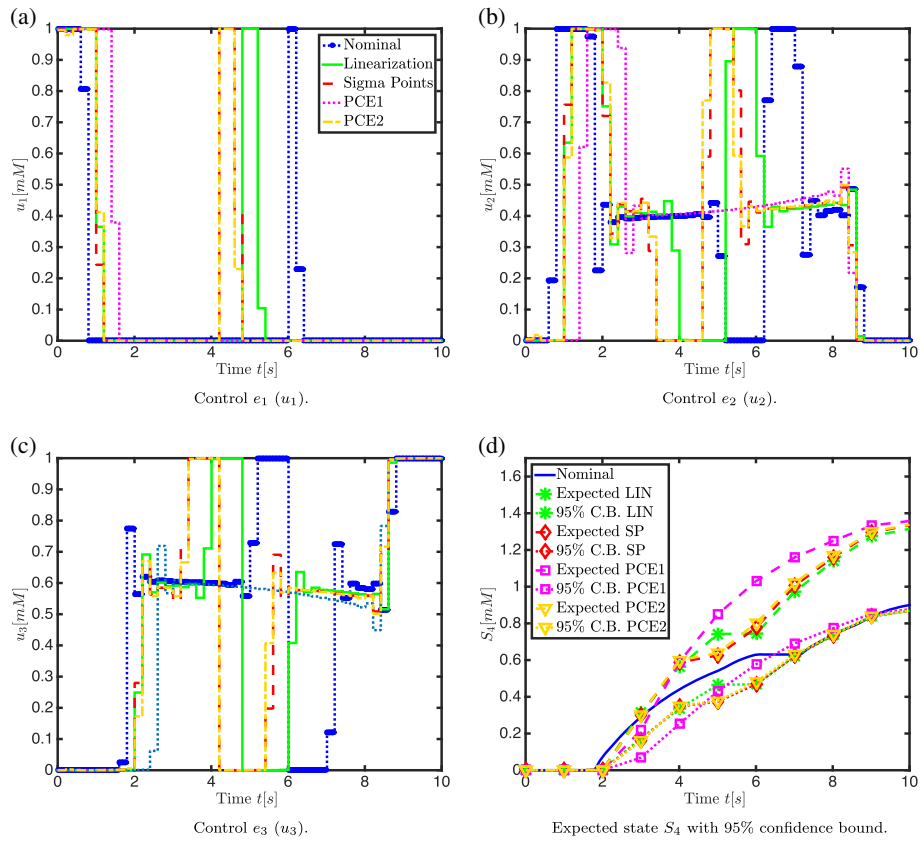|  | Nominal | Linearization | Sigma points | PCE1 | PCE2 |
|---|---|---|---|---|---|
| States | 5 | 20 | 29 | 17 | 41 |
| CPU time [s] | 0.156 | 1.332 | 4.507 | 2.087 | 7.780 |
| **E** [$J$] | 3.65 | 6.61 | 7.06 | 7.74 | 7.04 |
| **E** [$S_4$] | 0.90 | 1.50 | 1.52 | 1.53 | 1.52 |
| $\sqrt{\textbf{Var}[S_4]}$ | 0 | 0.30 | 0.32 | 0.33 | 0.32 |

From the enzyme concentration profiles in Fig. 5(a)-(c), it is clear that, the robustified enzyme concentrations will increase earlier than the nominal enzyme concentrations. The sequential activation of the controls (i.e., the increasing and decreasing enzyme concentrations $e_i$) will be early enough and sufficient in the robustified case to ensure that the robustified constraint is satisfied, i.e., a sufficient amount of $S_4$ is produced. In Fig. 5(d) it is shown that for the PCE1 approach the minimum treshold of 0.90 $mM$ is reached after 5.3 seconds, for PCE2 after 6.52, for the sigma points after 6.58 and for linearization after 6.7 seconds, while in the nominal case the minimum treshold of 0.90 $mM$ is reached at the end time of 10 seconds. The enzyme concentrations, calculated with the approximative uncertainty propagation techniques, ensure more robustness towards satisfying the minimum end concentration of 0.90 $mM$ for $S_4$. However, including robustness towards satisfying the minimum end concentration leads to a higher intermediate accumulation as shown in Table 3. This is acceptable, as long as cell viability is not compromised. Comparing this with experimental results for amino-acid biosynthetic pathways in *Escherichia coli* [38, 40, 41] a sequential activation of the enzymes can be observed in the enzyme concentration profiles.

### Comparison of the approximation techniques for uncertainty propagation

The number of constraint violations, expected value of $S_4$ and the variance on $S_4$ for the different approaches and $\alpha_{S_4}$ are shown in Tables 4 and 5 for a normal and uniform parametric uncertainty distribution, respectively. More extensive results are given in Additional file 3.

**Expected value and 95 % confidence bound** First the expected value and 95 % confidence bound of $S_4$ (based on $\alpha_{S_4} = 1.65$) are compared. This is done in Fig. 5(d) and it can be seen that the expected state and 95 % confidence bounds for $S_4$ are very similar when computed with the linearization, sigma points and PCE2 approaches. The expected value for the PCE1 approach differs slightly from the others: initially it is taking more distance from the nominal profile, indicating that this approach is more conservative than the other approaches and will lead to less constraint violations. In general, linear approximation techniques (as the PCE1 approach) tend to be more conservative, but it cannot be predicted upfront whether this is in a positive sense or not.

**Constraint violations** In order to investigate the performance of the different approximation techniques for uncertainty propagation with respect to the constraint violations, a Monte Carlo simulation with 1000 realizations (i.e., randomly generated parameter samples from

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 12 of 20



**Fig. 5** Results for Case 1. Comparison of the control profiles $e_1$ (**a**), $e_2$ (**b**) and $e_3$ (**c**) calculated with linearization, sigma points approach, PCE1 and PCE2 for $\alpha = 1.65$ with the nominal control profile and (**d**) comparison of the expected state $S_4$ and its 95 % confidence bound calculated with linearization, sigma points, PCE1 and PCE2 with the nominal case ($\alpha = 1.65$) in case of 3 uncertain parameters $k_1$, $k_2$ and $k_3$

**Table 4** Case 1 - Results Monte Carlo simulations ($N = 1000$) in case of three normally distributed uncertain parameters $k_1$, $k_2$ and $k_3$ for robustified terminal constraint with the number of constraint violations, mean terminal constraint values and variance on the terminal constraint

|  | Nominal case | Linearization | Sigma points | PCE 1 | PCE 2 |
|---|---|---|---|---|---|
| $\alpha = 1.96$ |  |  |  |  |  |
| $\bar{J}$ | 11.903 | 13.4600 | 13.600 | 13.829 | 13.583 |
| $\sigma_J$ | 0.5578 | 0.9737 | 1.0087 | 1.016 | 0.9981 |
| $\bar{S}_4$ | 0.88734 | 1.4739 | 1.5242 | 1.5500 | 1.5316 |
| $\sigma_{S_4}$ | 0.1733 | 0.2878 | 0.2970 | 0.2993 | 0.2993 |
| $c_t$ violations | 509 (50.9 %) | 30 (3.0 %) | 21 (2.1 %) | 20 (2.0 %) | 20 (2.0 %) |
| $\alpha = 1.65$ |  |  |  |  |  |
| $\bar{J}$ | 11.903 | 13.051 | 13.127 | 13.331 | 13.139 |
| $\sigma_J$ | 0.5578 | 0.8624 | 0.8869 | 0.9049 | 0.8892 |
| $\bar{S}_4$ | 0.8712 | 1.3357 | 1.3765 | 1.3910 | 1.3781 |
| $\sigma_{S_4}$ | 0.1733 | 0.2614 | 0.2691 | 0.2683 | 0.2690 |
| $c_t$ violations | 509 (50.9 %) | 51 (5.1 %) | 44 (4.4 %) | 39 (3.9 %) | 43 (4.3 %) |

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 13 of 20

**Table 5** Case 1 - Results Monte Carlo simulations ($N = 1000$) in case of three uniformly distributed uncertain parameters $k_1, k_2$ and $k_3$ for robustified terminal constraint with the number of constraint violations, mean values and variances on the objective function and terminal constraint, respectively

|  | Nominal case | Linearization | Sigma points | PCE1 | PCE2 | PCE2 Uniform |
|---|---|---|---|---|---|---|
| $\alpha = 1.96$ | | | | | | |
| $\bar{J}$ | 11.913 | 13.477 | 13.618 | 13.847 | 13.601 | 13.591 |
| $\sigma_J$ | 0.5470 | 0.9601 | 0.9942 | 1.0000 | 0.9835 | 0.9790 |
| $\bar{c}_t$ | 0.88917 | 1.4768 | 1.5273 | 1.5533 | 1.5347 | 1.5261 |
| $\sigma_{c_t}$ | 0.1818 | 0.3015 | 0.3110 | 0.3132 | 0.3137 | 0.3118 |
| $c_t$ violation | 516 (51.6 %) | 4 (0.4 %) | 1 (0.1 %) | 0 (0.0 %) | 1 (0.1 %) | 2 (0.2 %) |
| $\alpha = 1.65$ | | | | | | |
| $\bar{J}$ | 11.913 | 13.066 | 13.143 | 13.347 | 13.155 | 13.330 |
| $\sigma_J$ | 0.5470 | 0.8489 | 0.8730 | 0.6813 | 0.8750 | 0.8881 |
| $\bar{c}_t$ | 0.88917 | 1.3384 | 1.3792 | 1.3940 | 1.3809 | 1.3660 |
| $\sigma_{c_t}$ | 0.1818 | 0.2740 | 0.2820 | 0.2808 | 0.2819 | 0.2754 |
| $c_t$ violation | 516 (51.6 %) | 35 (3.5 %) | 22 (2.2 %) | 11 (1.1 %) | 22 (2.2 %) | 22 (2.2 %) |

the parametric uncertainty distribution) has been performed for the four approaches and is compared with the nominal case.

From these simulations, it is observed in case of a normal parametric uncertainty distribution that all four methods reduce the amount of constraint violations significantly: from 50.9 % in the nominal case to even 2.0 % for PCE1 and PCE2, when a backoff parameter value of $\alpha = 1.96$ is chosen. The same holds for a uniform parametric uncertainty distribution.

In practice, the most interesting backoff parameter values are 1.65 and 1.96, corresponding to 5 % and 2.5 % violations in case of a normal distribution. From the results in Tables 4 and 5 it is clear that the PCE1 method scores the best with respect to constraint violations, followed by PCE2 and sigma points. The performance of the sigma points method and the second order polynomial chaos expansion are, as shown throughout this case study, very similar.

**Parametric uncertainty distribution** If an uncertainty distribution is assumed for the constraint and the back-off parameters are chosen in accordance with the quantiles (which is the case for the normal parametric uncertainty distribution), the level of constraint violations should correspond exactly with the confidence level. A too low degree of violations is also not wished, since the uncertainty is not propagated correctly in that case.

Furthermore, the expected values and variances of $S_4$ for a backoff parameter of $\alpha = 1.96$ from Table 3 which are predicted with the approximation techniques, are very close to the empirically calculated expected values and variances by Monte Carlo simulation in Tables 4 and 5. For the sigma points, PCE1 and PCE2 approaches these

predicted expected values and variances are an accurate estimation of the ones obtained by Monte Carlo simulation. For the linearization approach, this is not the case. However, the expected value of the intermediate accumulation objective function is not accurately predicted. This objective function is not robustified in this case study, as the variance on the objective function is not taken into account. Therefore, the predictions of the expected value of the objective functions are not accurate, when compared with a Monte Carlo setting.

There is less backoff from the objective function when a uniform parametric uncertainty distribution is assumed, also the variance is reduced for the second order polynomial chaos expansion with respect to the assumption of a normal parametric uncertainty distribution. The percentage of constraint violations on the other hand is slightly higher when a uniform distribution is considered. However, for this case study, the difference in performance between a uniform parametric uncertainty distribution and a normal parametric uncertainty distribution for the polynomial chaos expansion, is small. Therefore, it should be stressed that including the additional information on the parametric uncertainty distribution can be useful. However, gathering information on the parametric uncertainty distribution is quite intensive and does not lead to a drastic improvement in performance for this case study.

### Case 2: glycolysis based network with 1 output

The terminal constraint and enzymatic cost objective function are robustified in this case study as shown in following equations. This is done in order to reduce the variance on the objective function. In contrast to Case 1, this should allow to have a better prediction

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 14 of 20

of the expected value and variance on the objective function.

$$0.675 \text{mM} \leq \mathbf{E}\left[S_5(t_f)\right] - \alpha_{S_5}\sqrt{\mathbf{Var}\left[S_5(t_f)\right]} \quad (20)$$

$$J_2 = \mathbf{E}\left[x_{\text{extra}}(t_f)\right] + \alpha_{x_{\text{extra}}}\mathbf{Var}\left[x_{\text{extra}}(t_f)\right)] \quad (21)$$

For simplicity, the backoff parameters $\alpha_{S_5}$ and $\alpha_{x_{\text{extra}}}$ are assumed to be the same and are called $\alpha$ in the remainder of the text. The objective function is robustified by adding the term $\alpha_{x_{\text{extra}}}\mathbf{Var}\left[x_{\text{extra}}(t_f)\right)]$, since the objective function has to be minimized and an increase in variance is penalized.

In this case study three parameters ($k_{\text{cat}}$, $K_M$ and $\lambda$) are considered uncertain. First, the multi-objective optimization results are discussed, followed by a more in depth analysis of the minimization of the enzymatic cost. Note that this in depth analysis can be performed for any Pareto point. Subsequently the computational aspects, a physical/biological interpretation of the results and comparison of the approaches are presented for this case study, based on the dynamic optimization results and Monte Carlo simulations for a normal and uniform parametric uncertainty distribution. For more extensive results, the reader is referred to Additional file 4.

### Multi-objective optimization results

The multi-objective optimization problem consists of minimizing the time needed to reach at least 0.675 mM of product $S_5$ and minimizing the enzymatic cost, i.e., the total enzyme consumption over the whole time span.

In the robustified problem formulation, both the objective function for the enzymatic cost and the terminal constraint are robustified with backoff parameter $\alpha$. It is assumed that the final time $t_f$ cannot exceed 30 seconds. The two objectives, final time and enzymatic cost, are clearly conflicting: reducing the time needed to reach a level of 0.675 mM of $S_5$ leads to an increase in the enzymatic cost and vice versa.

Receeding Pareto fronts from the nominal optimal solution with increasing backoff parameter values can be observed in Fig. 6: both anchor points shift away from the nominal optimal solution. This is the price in performance (i.e., minimum time and enzymatic cost) that has to be paid to ensure a minimum concentration of 0.675 mM for $S_5$. However, it is also observed that the Pareto fronts change shape and range, when the backoff parameter increases. This is related to the feasibility of the Pareto points.

A comparison of the different approximation techniques for uncertainty propagation based on the Pareto fronts is presented in Additional file 4. From this comparison it is firstly seen that the linearization and sigma points approach take more backoff than the polynomial chaos approaches. Secondly, when the backoff parameters

decrease, it is observed that there is a similarity between the Pareto fronts for the PCE2 and sigma points approach. This can be explained from the variance that is taken less into account when the backoff parameter value decreases. Since the difference between the PCE2 and sigma points approach lies in the variance calculation, this explains the increasing difference in Pareto fronts with an increasing backoff parameter value. The expected value calculation is the same for the sigma points and PCE2 approach in case of a normal parametric uncertainty distribution. For this case study, the sigma points are a subset of the PCE2 sampling points as shown in Additional file 4.

### Computational aspects

For the discussion of the computational aspects, the single objective optimization of the enzymatic cost is considered. It is assumed that the final time is fixed at 30 seconds.

In Table 6 an overview is presented of the number of states, the CPU time, objective function values, terminal constraint and their expected values and standard deviations for the different approximation techniques for uncertainty propagation for 3 uncertain parameters $K_M$, $\lambda$ and $k_{\text{cat}}$. This is done for the largest backoff parameter value $\alpha = 1.96$ for the same reasons as mentioned in the first case study.

In this case study the linearization approach is computationally the most expensive. While in Case 1, the increase in computational time is related to the increase in the size of the optimization problem, this cannot be the explanation for why the linearization approach takes the most CPU time. One explanation for the long CPU time of the linearization approach, can be the nonlinearity of the model in Case 2 and solving the sensitivity equations. The interconnection of the states in the sensitivity equations, makes the linearization approach computationally more challenging.

### Physical/biological interpretation

In Fig. 7 the enzyme expression rates are shown together with the corresponding enzyme concentration profiles, which are computed by minimizing the enzymatic cost for the nominal case (a) and with the different approximation techniques for uncertainty propagation: linearization (b), sigma points (c), PCE1 (d) and PCE2 (e) ($\alpha = 1.65$). It is observed that the expression rate profiles show an on/off behavior that leads to the sequential activation of the different enzymes in the network. The physical interpretation of this on/off behavior of the enzyme expression rates is that the previous enzyme first has to be degraded, before the other enzyme can be synthesized. This makes sense from a biological point of view, since the cells only have a limited amount of proteins available. This also corresponds to the satisfaction of the constraint on the
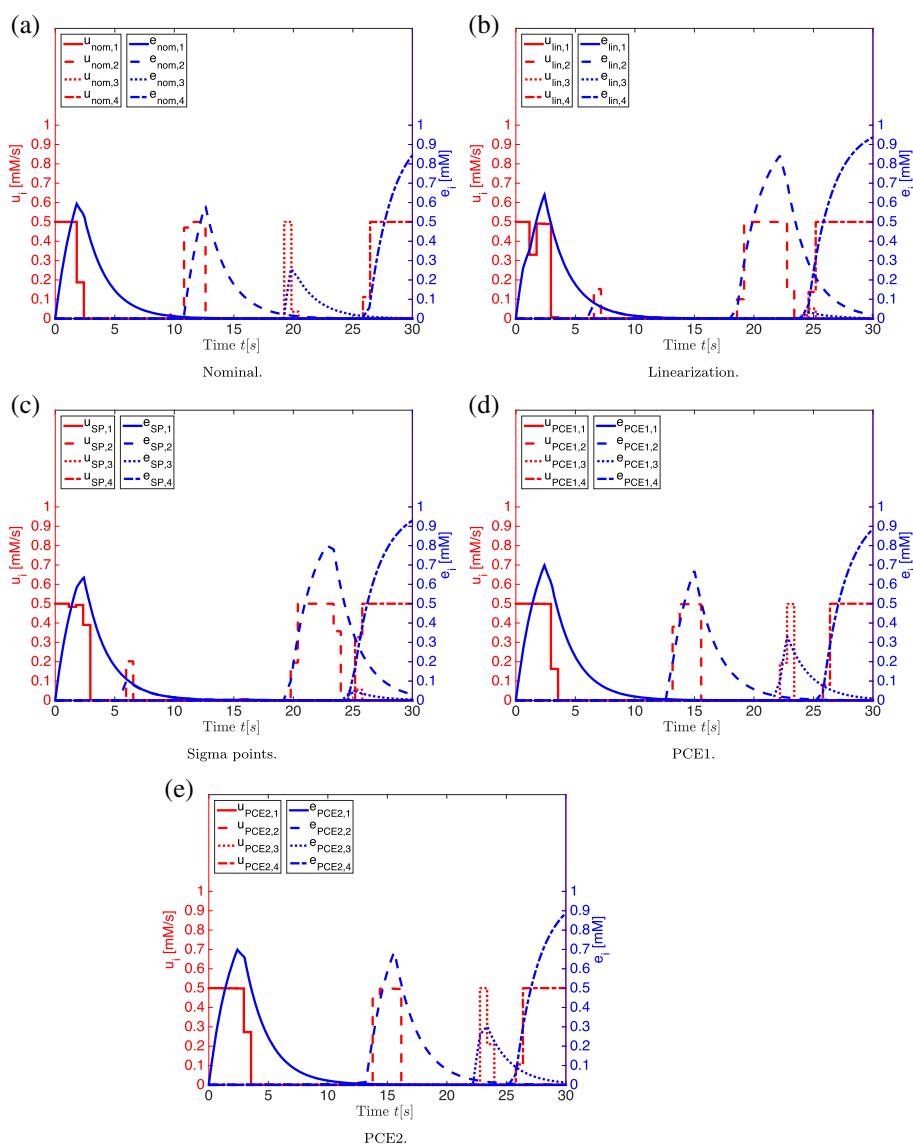
Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 15 of 20
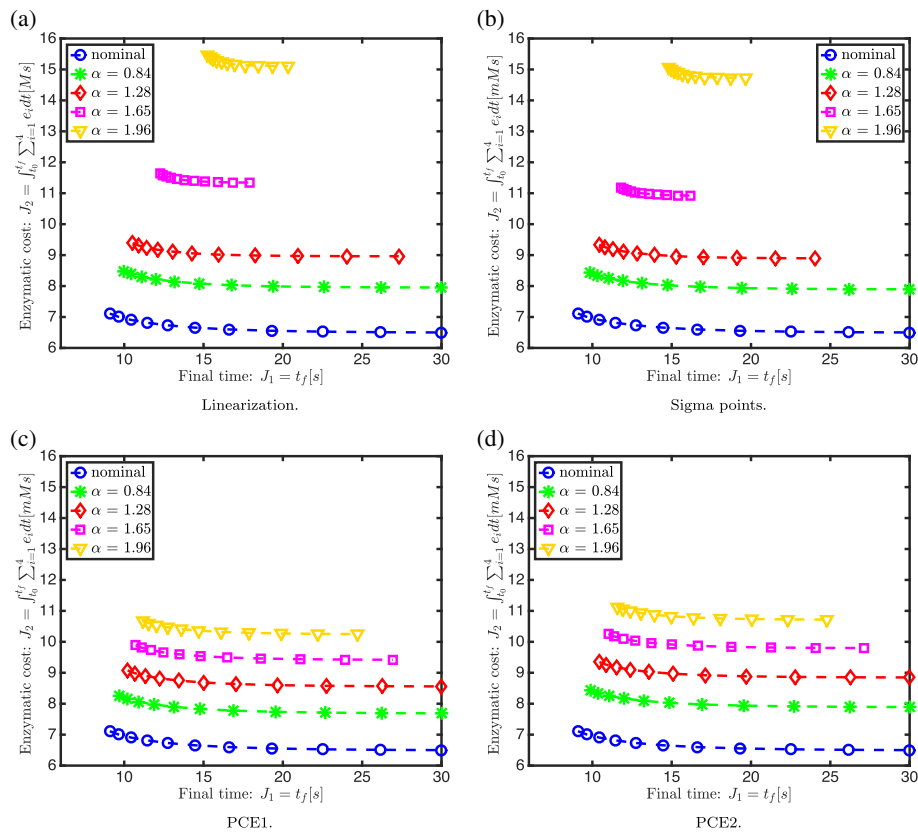


**Fig. 6** Receeding Pareto fronts Case 2. Receeding Pareto fronts with increasing backoff parameter $\alpha$ for linearization (**a**), sigma points (**b**), first (**c**) and second order polynomial chaos expansion (**d**) approach in case of 3 uncertain parameters $K_M$, $\lambda$ and $k_{cat}$

total enzymatic content. When minimizing the end time, there is a high accumulation of metabolites. This accumulation can affect cell viability negatively. The optimal control profiles in Fig. 8(a)-(d) clearly show a switching pattern, corresponding to the sequential activation of the pathways. According to [42] there is a mechanism leading to more pronounced transcriptional control of costly enzymes which can be explained by the trade-off between enzymatic cost minimization and time. Similarly to the

Case 1, it can be seen in Fig. 8(e)-(f) that the minimum threshold of 0.675 $mM$ for $S_5$ is reached sooner when applying the approximation techniques for uncertainty propagation. However, this robustness with respect to the terminal constraint on $S_5$ comes together with an increase in enzymatic cost as shown in Table 6. To avoid a too high enzymatic cost that is harmful for cell viability, an upper bound on the enzymatic cost can be introduced and robustified.

**Table 6** Case 2 - Overview of the number of states, CPU time, objective function values, terminal constraint values and their expected values and standard deviations for the different approximation techniques for uncertainty propagation when the enzymatic cost is minimized for $\alpha = 1.96$ for 3 uncertain parameters ($K_m$, $\lambda$ and $k_{cat}$)

|  | Nominal | Linearization | Sigma points | PCE1 | PCE2 |
|---|---|---|---|---|---|
| States | 9 | 36 | 63 | 36 | 90 |
| CPU time [s] | 0.547 | 117.474 | 25.65 | 4.574 | 25.847 |
| $\mathbf{E}[J]$ | 6.500 | 12.945 | 8.338 | 8.799 | 9.168 |
| $\sqrt{\mathbf{Var}[J]}$ | 0 | 1.113 | 0.697 | 0.739 | 0.791 |
| $c_t$ | 0.675 | 0.675 | 0.675 | 0.675 | 0.675 |
| $\mathbf{E}[c_t]$ | 0.675 | 1.000 | 1.067 | 1.163 | 1.210 |
| $\sqrt{\mathbf{Var}[c_t]}$ | 0 | 0.166 | 0.200 | 0.249 | 0.273 |

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 16 of 20



**Fig. 7** Enzyme expression rates and enzyme concentration profiles Case 2. Enzyme expression rates together with the enzyme concentration profiles following from the minimization of the enzymatic cost for the nominal case (**a**) and with the different approximation techniques for uncertainty propagation: linearization (**b**), sigma points (**c**), PCE1 (**d**) and PCE2 (**e**) ($\alpha = 1.65$)

In Fig. 7(b) and (c) some remarkable observations are made for the linearization and sigma points approaches. After approximately 6 seconds a small expression rate $u_2$ is observed, leading to a short activation of $e_2$, producing intermediates $S_3$ and $S_4$. Furthermore after approximately 18 seconds in the linearization approach and 20 seconds in the sigma points approach, $e_2$ is expressed for a longer time. The accumulated $S_2$ is intensively consumed in this period for the production of intermediates $S_3$ and $S_4$. Eventually this will lead to a very low concentration of $S_2$ in comparison with the other approaches. Furthermore, a

very small expression rate $u_3$ is observed in Fig. 7(b) and (c) for the optimization with the linearization and sigma points approaches, leading to a weak activation of $e_3$. This means that the branch in the glycolysis inspired network involving the conversion of $S_3$ to $S_4$ is practically inactive, leading to accumulation of $S_3$ for the linearization and sigma points approaches. This behavior is not observed in the profiles obtained with the other techniques and explains the higher objective function values for enzymatic cost in case of the linearization and sigma points approaches.

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 17 of 20
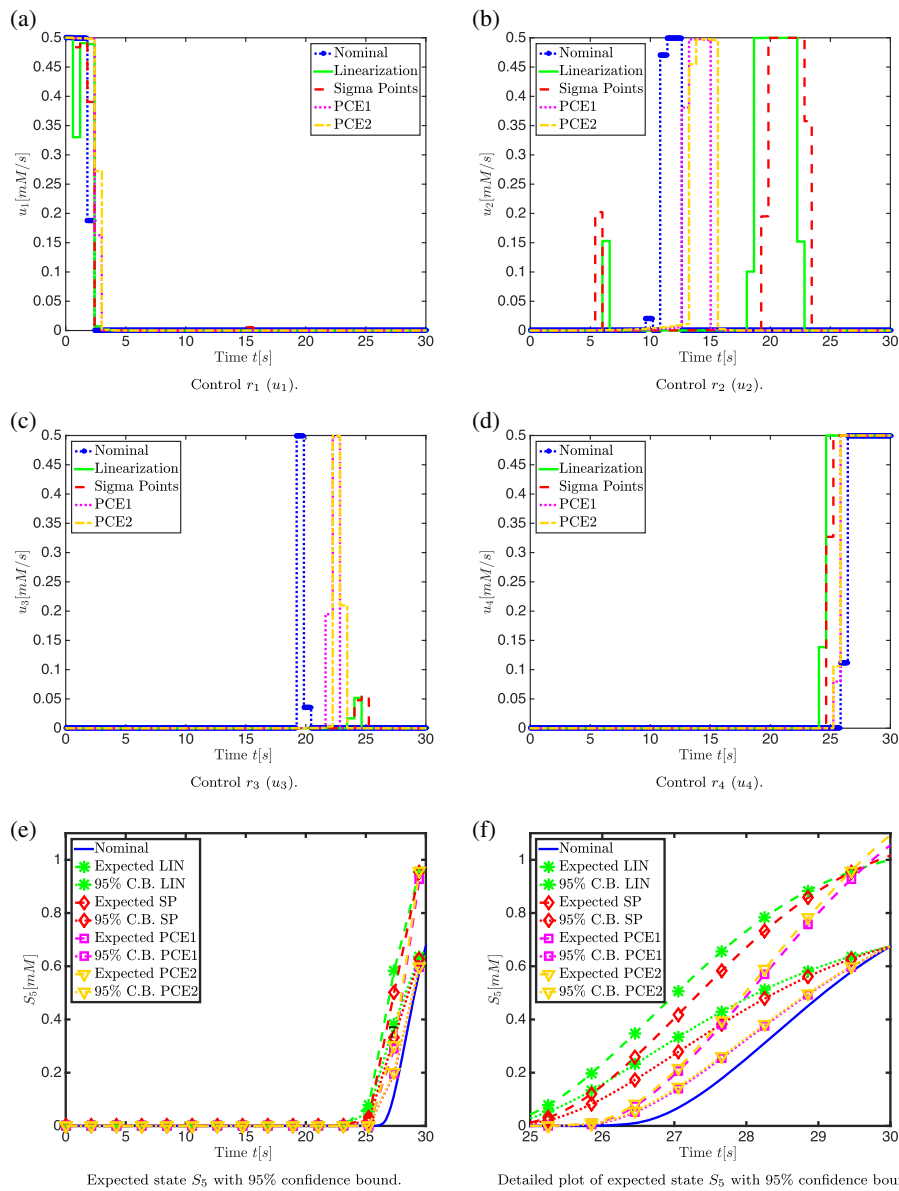


**Fig. 8** Results for minimization of enzymatic cost in Case 2. Comparison of the control profiles $r_1$ (**a**), $r_2$ (**b**), $r_3$ (**c**) and $r_4$ (**d**) calculated with linearization, sigma points approach, PCE1 and PCE2 for $\alpha = 1.65$ with the nominal control profile and comparison of expected state $S_5$ and its 95 % confidence bound calculated with linearization, sigma points, PCE1 and PCE2 with the nominal case ($\alpha = 1.65$) ((**e**) and (**f**)) in case of 3 uncertain parameters $K_M$, $\lambda$ and $k_{cat}$

### Comparison of the approximation techniques for uncertainty propagation

The performance of the different approximation techniques for uncertainty propagation with respect to the constraint violations is investigated by performing a Monte Carlo simulation with 1000 noise realizations from a normal distribution. The number of constraint violations on $S_5$, mean values, and the variances for the objective function and terminal constraint, respectively are shown in Table 7 for the different approaches and

backoff parameter values $\alpha_i$. More extensive results are presented in Additional file 4.

**Expected value and 95 % confidence bound** First the expected value and 95 % confidence bound of $S_5$ (based on $\alpha = 1.65$) are compared. This is done in Figs. 8(e)-(f). From Fig. 8(e)-(f) it can be seen that the expected state and 95 % confidence bounds for $S_5$ are similar for the PCE1 and PCE2 approaches. However, the linearization approach and sigma points approach take initially

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 18 of 20

**Table 7** Case 2 - Results Monte Carlo simulations ($N = 1000$) in case of three normally distributed uncertain parameters ($K_m$, $\lambda$ and $k_{cat}$) for robustified terminal constraint and objective function with the number of constraint violations, mean values and variances on the objective function and terminal constraint, respectively

| | Nominal case | Linearization | Sigma points | PCE1 | PCE2 |
|---|---|---|---|---|---|
| $\alpha = 1.96$ | | | | | |
| $\bar{J}$ | 6.5268 | 12.996 | 12.549 | 8.7702 | 9.1358 |
| $\sigma_J$ | 0.57543 | 1.1893 | 1.15230 | 0.78540 | 0.82107 |
| $\bar{c}_t$ | 0.69319 | 1.0106 | 1.0106 | 1.1406 | 1.2099 |
| $\sigma_{c_t}$ | 0.18291 | 0.17462 | 0.17752 | 0.27117 | 0.28089 |
| $c_t$ violation | 477 (47.7 %) | 15 (1.5 %) | 18 (1.8 %) | 23 (2.3 %) | 13 (1.3 %) |
| $\alpha = 1.65$ | | | | | |
| $\bar{J}$ | 6.5268 | 10.04 | 9.6011 | 8.2477 | 8.5514 |
| $\sigma_J$ | 0.57543 | 0.88430 | 0.83741 | 0.73621 | 0.76632 |
| $\bar{c}_t$ | 0.69319 | 1.0143 | 1.0157 | 1.0339 | 1.0933 |
| $\sigma_{c_t}$ | 0.18291 | 0.20620 | 0.21414 | 0.25176 | 0.26120 |
| $c_t$ violation | 477 (47.7 %) | 31 (3.1 %) | 39 (3.9 %) | 53 (5.3 %) | 34 (3.4 %) |

more distance from the nominal profile. In this case, similarly to Case1, a linear approximation technique (i.e., the linearization approach) is more conservative, implying less constraint violations. All 95 % confidence bounds are at 0.675 *mM* at the end as required by the imposed constraint in the implementation of the approximation techniques for uncertainty propagation.

Both PCE1 and PCE2 have a more accurate prediction of the expected value and variance of the objective function $J$ and the terminal constraint function value $c_t$ than the linearization and sigma points approach (in Table 6), when comparing with the empirically calculated expected values and variances with Monte Carlo simulations (in Table 7).

**Constraint violations** The percentage of constraint violations is investigated by performing a Monte Carlo simulation with 1000 noise realizations from the parametric uncertainty distribution. From these simulations, it is observed that all four methods reduce the amount of constraint violations significantly: from 47.7 % in the nominal case to even 1.3 % for PCE2, when a backoff parameter value of $\alpha = 1.96$ is chosen in case of a normal parametric uncertainty distribution. For this case study, the PCE2 method is superior in performance, when considering number of constraint violations.

**Parametric uncertainty distribution** For this case study, the integration of prior information on the parametric uncertainty distribution in the polynomial chaos expansion approaches is studied. The orthogonal polynomials for $K_m$, $\lambda$ and $k_{cat}$ are derived via the definition of orthogonal polynomials. Details can be found in Additional file 4.

For a normal and uniform parametric uncertainty distribution, a Monte Carlo simulation procedure with 1000 noise realizations has been followed and the results are summarized in Tables 7 and 8.

As in the first case study, the percentage of constraint violations is slightly higher when a uniform distribution is considered. For this case study, the difference in performance between a uniform parametric uncertainty distribution and a normal parametric uncertainty distribution for the polynomial chaos expansion, is small. Gathering information on the parametric uncertainty distribution from a parameter identification procedure is intensive and does not lead to a drastic improvement in performance for this case study.

## Conclusions

In this work parametric uncertainty has been taken into account for prediction and control of biological networks. A critical comparison of three approximation techniques for uncertainty propagation, i.e., the linearization, sigma points and polynomial chaos expansion approaches has been made for dynamic optimization of biological networks under parametric uncertainty. The main advantage of the polynomial chaos expansion is its ability to tackle more easily non-normal parametric uncertainty distributions. Two case studies are investigated: *(i)* the minimization of intemediate metabolite accumulation in a basic three-step linear pathway model (with 3 metabolites, 3 fluxes, 4 differential states and 3 controls) and *(ii)* the multi-objective optimization (i.e., the minimization) of the final time and enzymatic cost a glycolysis inspired network model (with 4 metabolites, 4 fluxes, 8 differential states and 4 controls). For further analysis of

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 19 of 20

**Table 8** Case 2 - Results Monte Carlo simulations ($N = 1000$) in case of three uniformly distributed uncertain parameters ($K_m$, $\lambda$ and $k_{cat}$) for robustified terminal constraint and objective function with the number of constraint violations, mean values and variances on the objective function and terminal constraint, respectively

| | Nominal case | Linearization | Sigma points | PCE1 | PCE2 | PCE2 uniform |
|---|---|---|---|---|---|---|
| $\alpha = 1.96$ | | | | | | |
| $\bar{J}$ | 6.5517 | 13.048 | 12.599 | 8.8045 | 9.1717 | 9.1498 |
| $\sigma_J$ | 0.5323 | 1.1015 | 1.0665 | 0.7270 | 0.7600 | 0.7600 |
| $\bar{c}_t$ | 0.7023 | 1.0205 | 1.0208 | 1.1555 | 1.2254 | 1.2240 |
| $\sigma_{c_t}$ | 0.1744 | 0.1659 | 0.1687 | 0.2590 | 0.2684 | 0.2690 |
| $c_t$ violation | 470 (47.0 %) | 10 (1.0 %) | 13 (1.3 %) | 21 (2.1 %) | 7 (0.7 %) | 9 (0.9 %) |
| $\alpha = 1.65$ | | | | | | |
| $\bar{J}$ | 6.5517 | 10.079 | 9.6385 | 8.2798 | 8.5848 | 8.5683 |
| $\sigma_J$ | 0.5323 | 0.8195 | 0.7762 | 0.6813 | 0.7092 | 0.7086 |
| $\bar{c}_t$ | 0.7030 | 1.0259 | 1.0277 | 1.0476 | 1.1077 | 1.1056 |
| $\sigma_{c_t}$ | 0.1744 | 0.1966 | 0.2042 | 0.2404 | 0.2495 | 0.2496 |
| $c_t$ violation | 470 (47.0 %) | 28 (2.8 %) | 36 (3.6 %) | 55 (5.5 %) | 29 (2.9 %) | 29 (2.9 %) |

the robustness, emphasis was put on a single objective: in Case 1 the minimization of the intermediate accumulation and in Case 2 the minimization of the enzymatic cost. In a next step, the robustness of the optimal control profiles obtained with the different approximation techniques for uncertainty propagation is investigated. Monte Carlo simulations are used for the assessment of these control profiles. From the results for both case studies, the different uncertainty propagation strategies each offer a robust solution under parametric uncertainty. When making the trade-off between computation time and the robustness of the obtained profiles, the sigma points and polynomial chaos expansion strategies score the best. In both case studies the effect of taking a uniform probability distribution for the parametric uncertainty has been taken into account. However, the gain by taking a uniform probability distribution into account instead of a normal probability distribution is low. There is a reduction on the considered backoff and the variance of the objective functions and constraint. On the other hand, an upfront identification procedure for the parametric uncertainty distribution is time-consuming, expensive and does not offer a substantial advantage for the considered case studies in comparison with assuming a normal parametric uncertainty distribution. The linearization, sigma points and polynomial chaos expansion approaches offer a great potential for optimization and modeling under uncertainty in systems biology. The application of these approximation techniques for uncertainty propagation to large scale biological network models is the subject of future work. The integration of these approximation techniques for uncertainty propagation in an interactive tool for multi-objective dynamic optimization [43] is also part of the future work.

## Additional files

**Additional file 1:** Detailed review on approximation techniques for uncertainty propagation. A detailed review of the approximation techniques for uncertainty propagation: linearization, sigma points and polynomial chaos expansion approach. (PDF 479 kb)

**Additional file 2:** Optimization methods and software. A more in depth description of the used optimization methods and software. (PDF 242 kb)

**Additional file 3:** Numerical results Case 1. A detailed overview of the numerical results for Case 1. (PDF 667 kb)

**Additional file 4:** Numerical results Case 2. A detailed overview of the numerical results for Case 2. (PDF 708 kb)

**Availability of data and material**
The data and material supporting the findings in this work can be found in Additional files 1, 2, 3 and 4. Furthermore, `Pomodoro` can be requested from https://perswww.kuleuven.be/~u0093798/software.php.

**Authors' contributions**
Authors jointly designed the study, PN and DT implemented the algorithms, performed the computations and wrote the manuscript. FL and JVI supervised the study and participated in writing the manuscript. All authors have read and approved the final manuscript.

**Competing interests**
The authors declare that they have no competing interests.

**Consent for publication**
Not applicable.

Nimmegeers *et al. BMC Systems Biology* (2016) 10:86

Page 20 of 20

**References**

1. Vercammen D, Logist F, Van Impe J. Dynamic estimation of specific fluxes in metabolic networks using non-linear dynamic optimization. BMC Syst Biol. 2014;8(132):1–22.
2. Llaneras F, Picó J. Stoichiometric modelling of cell metabolism. J Biosci Bioeng. 2008;105(1):1–11.
3. Systems Metabolic Engineering In: Wittmann C, Lee SY, editors. Springer Science+Business Media. 1st ed. Netherlands: Springer; 2012. p. XII, 388.
4. Monk J, Nogales J, Palsson BO. Optimizing genome-scale network reconstructions. Nat Biotechnol. 2014;32(5):447–52.
5. Sendin JOH, Exler O, Banga JR. Multi-objective mixed integer strategy for the optimisation of biological networks. Syst Biol IET. 2010;4(3):236–48.
6. Higuera C, Villaverde AF, Banga JR, Ross J, Morán F. Multi-criteria optimization of regulation in metabolic networks. PLoS ONE. 2012;7(7): 41122.
7. de Hijas-Liste G, Klipp E, Balsa-Canto E, Banga J. Global dynamic optimization approach to predict activation in metabolic pathways. BMC Syst Biol. 2014;8(1):1.
8. Sendiín J-OH, Otero-Muras I, Alonso AA, Banga JR. Improved optimization methods for the multiobjective design of bioprocesses. Ind Eng Chem Res. 2006;45(25):8594–603.
9. Kaern M, Elston TC, Blake WJ, Collins JJ. Stochasticity in gene expression: from theories to phenotypes. Nat Rev Genet. 2005;6(6):451–64.
10. Telen D, Vallerio M, Cabianca L, Houska B, Van Impe J, Logist F. Approximate robust optimization of nonlinear systems under parametric uncertainty and process noise. J Process Control. 2015;33:140–54.
11. Srinivasan B, Bonvin D, Visser E, Palanki S. Dynamic optimization of batch processes: II, role of measurements in handling uncertainty. Comput Chem Eng. 2003;27(1):27–44.
12. Julier S, Uhlmann JK. A general method for approximating nonlinear transformations of probability distributions. Oxford, OX1 3PJ United Kingdom: Robotics Research Group, Department of Engineering Science, University of Oxford; 1996.
13. Wiener N. The homogeneous chaos. Am J Math. 1938;60:897–936.
14. Xiu D, Karniadakis GE. The wiener-askey polynomial chaos for stochastic differential equations. SIAM J Sci Comput. 2002;24:619–44.
15. Nagy ZK, Braatz RD. Open-loop and closed-loop robust optimal control of batch processes using distributional and worst-case analysis. J Process Control. 2004;14(4):411–22.
16. Logist F, Houska B, Diehl M, Van Impe J. Robust multi-objective optimal control of uncertain (bio)chemical processes. Chem Eng Sci. 2011;66: 4670–82.
17. Mesbah A, Streif S. A probabilistic approach to robust optimal experiment design with chance constraints. In: 9th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM. Whistler, Canada: Elsevier; 2015. p. 100–5.
18. Houska B, Logist F, Diehl M, Van Impe J. Robust optimization of nonlinear dynamic systems with application to a jacketed tubular reactor. J Process Control. 2012;22:1152–60.
19. Mesbah A, Streif S, Findeisen R, Braatz RD. Stochastic nonlinear model predictive control with probabilistic constraints. In: American Control Conference (ACC). American Automatic Control Council; 2014. p. 2413–9.
20. Galvanin F, Barolo M, Bezzo F. A backoff strategy for model-based experiment design under parametric uncertainty. AlChE J. 2010;56(8): 2088–102.
21. Asprey SP, Macchietto S. Designing robust optimal dynamic experiments. J Process Control. 2002;12(4):545–56.
22. Kadam JV, Schlegel M, Srinivasan B, Bonvin D, Marquardt W. Dynamic optimization in the presence of uncertainty: From off-line nominal solution to measurement-based implementation. J Process Control. 2007;17(5):389–98.
23. Podmajersky M, Fikar M, Chachuat B. Measurement-based optimization of batch and repetitive processes using an integrated two-layer architecture. J Process Control. 2013;23(7):943–55.
24. Diwekar UM, Kalagnanam JR. Efficient sampling technique for optimization under uncertainty. AlChE J. 1997;43:440–7.
25. Webster MD, Tatang MA, McRae GJ. Application of the Probabilistic Collocation Method for an Uncertainty Analysis of a Simple Ocean Model. Report 4 (MIT Joint Program on the Science and Policy of Global Change). 1996:32. http://globalchange.mit.edu/files/document/MITJPSPGC_Rpt4. pdf.
26. Haimes Y, Lasdon L, Wismer D. On a bicriterion of the problems of integrated system identification and system optimization. IEEE Trans Syst Man Cybernet. 1971;SMC-1:296–7.
27. Das I, Dennis JE. A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. Structural optimization. 1997;14(1):63–9.
28. Messac A, Ismail-Yahaya A, Mattson CA. The normalized constraint method for generating the pareto frontier. Struct Multidiscip Optim. 2003;25(2):86–98.
29. Marler T, Arora J. Survey of multi-objective optimization methods for engineering. Struct Multidiscip Optim. 2010;41:853–62.
30. Logist F, Van Impe J. Novel insights for multi-objective optimisation in engineering using normal boundary intersection and (enhanced) normalised normal constraint. Struct Multidiscip Optim. 2012;45(3): 417–31.
31. Das I, Dennis J. Normal-boundary intersection: a new method for generating the pareto surface in nonlinear multicriteria optimization problems. Siam J Optim. 1998;8(3):631–57.
32. Bhaskar V, Gupta S, Ray A. Applications of multi-objective optimization in chemical engineering. Rev Chem Eng. 2000;16:1–54.
33. Reyes-Sierra M, Coello CAC. Multi-objective particle swarm optimizers: a survey of the state-of-the-art. Int J Comput Intell Res. 2006;2(3):287–308.
34. Suman B, Kumar P. A survey of simulated annealing as a tool for single and multiobjective optimization. J Oper Res Soc. 2006;57:1143–60.
35. Logist F, Telen D, Houska B, Diehl M, Van Impe J. Multi-objective optimal control of dynamic bioprocesses using acado toolkit. Bioprocess Biosyst Eng. 2013;36:151–64.
36. Andersson J. A General-Purpose Software Framework for Dynamic Optimization. PhD thesis. Belgium: Arenberg Doctoral School, KU Leuven, Department of Electrical Engineering (ESAT/SCD) and Optimization in Engineering Center, Kasteelpark Arenberg 10, 3001-Heverlee; 2013.
37. Wächter A, Biegler LT. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. Math Programm. 2006;106:25–57.
38. Bartl M, Li P, Schuster S. Modelling the optimal timing in metabolic pathway activationǓuse of pontryagin's maximum principle and role of the golden section. Biosystems. 2010;101(1):67–77.
39. Bartl M, Li P, Schuster S. Just-in-time activation of a glycolysis inspired metabolic network - solution with a dynamic optimization approach. In: Proceedings 55nd International Scientific Colloquium. Ilmenau, Germany: Institute for Automation and Systems Engineering Ilmenau University of Technology; 2010.
40. Klipp E, Heinrich R, Holzhütter HG. Prediction of temporal gene expression. Eur J Biochem. 2002;269(22):5406–13.
41. Zaslaver A, Mayo AE, Rosenberg R, Bashkin P, Sberro H, Tsalyuk M, Surette MG, Alon U. Just-in-time transcription program in metabolic pathways. Nat Genet. 2004;36(5):486–91.
42. Wessely F, Bartl M, Guthke R, Li P, Schuster S, Kaleta C. Optimal regulatory strategies for metabolic pathways in escherichia coli depending on protein costs. Mol Syst Biol. 2011;7(1):1–13.
43. Vallerio M, Hufkens J, Van Impe J, Logist F. An interactive decision-support system for multi-objective optimization of nonlinear dynamic processes with uncertainty. Expert Syst Appl. 2015;42:7710–31.