

# Sparse Grid Interpolation of Itô Stochastic Models in Epidemiology and Systems Biology

Aditya Sai<sup>1</sup> and Nan Kong<sup>1</sup>

**Abstract**—Computational modeling enhances our understanding of seemingly incomprehensible biological systems. Certain dynamical models may be unwieldy to simulate repetitively, especially if the models contain uncertainty. This is evident in both epidemiology and systems biology, where inherent biological variability and a spectrum of plausible model hypotheses exist. Surrogate modeling using sparse grid interpolation can alleviate the burden associated with increasing dimension of the parameter space. By leveraging multivariate tensor products across a predefined set of points, sparse grid interpolants are able to provide a promising surrogate model to answer pressing domain-related questions. Specifically, we explore Itô stochastic differential equation-based models, with examples of a susceptible-infectious-vaccinated-removed (SIVR) epidemiological model, a breast cancer tumor population model, and a biochemical network model of the JAK-STAT signal cascade presented. Surrogate modeling is performed to satisfy model-based objectives. Overall, sparse grid interpolation is an effective computational modeling tool, enabling researchers in the epidemiology and systems biology communities to interrogate models of interest for key insight into biological phenomena.

**Index Terms**—interpolation, epidemiology, parameter estimation, sparse grid, Itô stochastic differential equation, surrogate modeling, systems biology

## I. INTRODUCTION

**B**IOLOGICAL phenomena are inherently complex. This complexity can be simplified for human understanding with mathematical models. Mathematical models condense key biological assumptions and knowledge into a unified representation [1]. Two biological domains that have benefited from mathematical modeling are epidemiology and systems biology. Epidemiology aims to characterize the dynamics of disease spread throughout a population [2]. Systems biology is concerned with the biological functions and mechanisms underpinning cellular networks [3]. Examples in both domains are commonly represented as mechanistic and semi-mechanistic mathematical models using ordinary differential equations (ODEs), which often have to be solved numerically using discretized approximations of the true solution. However, randomness and heterogeneity can also influence biological systems, calling for the use of stochastic processes [4], [5].

Consider Itô stochastic differential equations (SDEs):

$$d\mathbf{X}(t) = f(\mathbf{X}, t, \boldsymbol{\theta})dt + g(\mathbf{X}, t, \boldsymbol{\theta})d\mathbf{B}(t). \quad \mathbf{X}(0) = X_0. \quad (1)$$

where  $\mathbf{X} \in \mathbb{R}^N$  is a continuous time stochastic process,  $\mathbf{B} \in \mathbb{R}^M$  is a Brownian motion process,  $t \in [0, T]$  is time,  $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^P$  is a vector of model parameters,

$f(\cdot) : \mathbb{R}^N \times [0, T] \times \Theta \rightarrow \mathbb{R}^N$  is the drift term (deterministic component),  $X_0$  are the initial conditions, and  $g(\cdot) : \mathbb{R}^N \times [0, T] \times \Theta \rightarrow \mathbb{R}^{N \times M}$  is the diffusion term (stochastic component). Examples of SDE-based models in epidemiology and systems biology include the human nervous system [6]–[8], cancer tumors [9], predator-prey systems [10], [11], and a glucose regulatory system for diabetes patients [12].

Complex system dynamics can be difficult to simulate when a large number of model parameters have to be considered [13]–[15]. Furthermore, local searches of these parameters may be insufficient to characterize the wide range of possible behaviors. Sparse grids allow for global, computationally efficient exploration of the parameter space  $\Theta$  using tensor-product quadrature [16]–[18]. These approximations of the underlying model mitigate the curse of dimensionality associated with the increasing dimension of  $\Theta$  by selecting the grid points, or support nodes, in a hierarchical fashion [17]–[19]. This is done so that nodes from a previous level of refinement can be re-used in higher levels of refinement. Once the original model has been evaluated at these support nodes and the interpolant has been constructed, the resulting surrogate model can be used in model-based control and optimization without having to directly integrate the underlying model, which is often computationally prohibitive. The concept of sparse grid interpolation, and surrogate modeling in general, is not unlike that of compressive sensing, where a compressible signal is recovered from a limited number of measurements [20]. Fig. 1 demonstrates the application of sparse grid interpolation to a simple 3-dimensional exponential function. Sparse grids have been applied to other stochastic models, such as stochastic partial differential equations with random inputs [21]–[26], backwards stochastic differential equations with random inputs [27], and differential algebraic equations with random parameters [28].

We demonstrate the application of sparse grid interpolation to approximating the dynamics of Itô SDE-based models in different biology contexts. In Section II, we discuss the concept of sparse grids, and the necessary numerical techniques for effective interpolation. Then, in Section III, we present examples of sparse grid interpolation through targeted computational experiments that approach domain-related problems. More specifically, we examine the role that noise plays in perturbing normal dynamics, or whether there is any discernible influence of noise at all. Finally, in Section IV, we summarize the significance of our work and propose future avenues of research.

Manuscript received July 20, 2017; revised November 1, 2017.

<sup>1</sup>The authors are with the Weldon School of Biomedical Engineering, Purdue University, West Lafayette, Indiana, USA (email: asai@purdue.edu; nkong@purdue.edu)

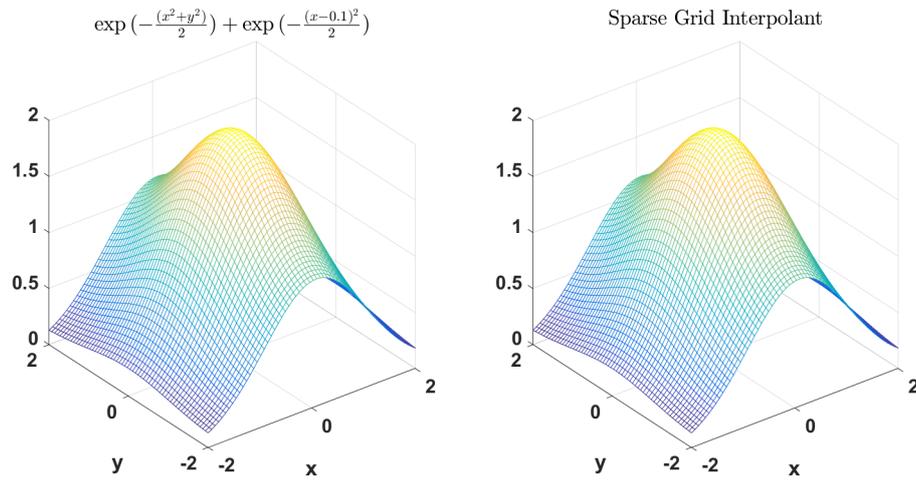


Fig. 1. Exponential function evaluated on a grid  $[-2, 2] \times [-2, 2]$ . Both the original function (left) and the sparse grid interpolant (right) are shown. The interpolant was produced with a relative error of 0.021%, absolute error of 0.00039, and 129 support nodes.

## II. METHODOLOGY

### A. Sparse Grid Interpolation

In sparse grid interpolation, the support nodes are selected in a predefined manner; a nested, hierarchical sampling scheme [17]–[19] recycles nodes from lower levels of resolution to use in higher levels.

A mathematical formulation of sparse grids now follows from [16]–[18], [29]–[31]. Consider a function  $f: [0, 1]^d \rightarrow \mathbb{R}$  that is to be interpolated on a finite number of support nodes. Dimensions that are not of unit length can be rescaled. Here,  $f$  represents the sample average and standard deviation of multiple SDE trajectories sampled at discrete time points. For a given  $f$ , a univariate interpolation function can be constructed:

$$\mathcal{U}^i(f) = \sum_{j=1}^{m_i} a_j^i \cdot f(x_j^i), \quad (2)$$

where  $i \in \mathbb{N}$ ,  $a_j^i \in C([0, 1])$ ,  $a_j^i(x_l^i) = \delta_{jl}$ ,  $l \in \mathbb{N}$  are the univariate basis functions, and  $x_j^i \in X^i = \{x_1^i, \dots, x_{m_i}^i\}$ ,  $x_j^i \in [0, 1]$ ,  $1 \leq j \leq m_i$ , are the support nodes.

Extending this interpolation function to multi-dimensional cases (i.e.  $d \geq 1$ ), the corresponding multivariate formula, using the full tensor product formulation, is as follows:

$$(\mathcal{U}^{i_1} \otimes \dots \otimes \mathcal{U}^{i_d})(f) = \sum_{j_1=1}^{m_{i_1}} \dots \sum_{j_d=1}^{m_{i_d}} (a_{j_1}^{i_1} \otimes \dots \otimes a_{j_d}^{i_d}) f(x_{j_1}^{i_1}, \dots, x_{j_d}^{i_d}). \quad (3)$$

The number of support nodes required for the full tensor product representation is  $\prod_{j=1}^d m_{i_j}$ , which is computationally intractable for high dimensions  $d$ .

The Smolyak construction aims to substantially decrease the number of support nodes used while preserving the interpolation properties observed in the 1-dimensional case. Define the difference function  $\Delta^i = \mathcal{U}^i - \mathcal{U}^{i-1}$ ,  $\mathcal{U}^0 = 0$  and multi-index  $\mathbf{i} \in \mathbb{N}^d$ ,  $|\mathbf{i}| = i_{i_1} + \dots + i_{i_d}$ . Now, define the

Smolyak interpolant as:

$$A_{n+d,d}(f) = \sum_{k=0}^n \sum_{|\mathbf{i}|=k+d} (\Delta^{i_1} \otimes \dots \otimes \Delta^{i_d})(f). \quad (4)$$

The inner sum can be further expressed as

$$\sum_{|\mathbf{i}|=k+d} \sum_{\mathbf{j}} (a_{j_1}^{i_1} \otimes \dots \otimes a_{j_d}^{i_d})(f(\mathbf{x}_{\mathbf{j}}^{\mathbf{i}}) - A_{k+d-1,d}(f(\mathbf{x}_{\mathbf{j}}^{\mathbf{i}}))), \quad (5)$$

where  $\mathbf{j}$  is the multi-index  $(j_1, \dots, j_d)$ ,  $j_l = 1, \dots, m_{i_l}$ ,  $l = 1, \dots, d$ , and the points  $\mathbf{x}_{\mathbf{j}}^{\mathbf{i}} = (x_{j_1}^{i_1}, \dots, x_{j_d}^{i_d})$ ,  $x_{j_l}^{i_l}$  is the  $j_l$ th element of  $X_{\Delta}^{i_l} = X^{i_l} \setminus X^{i_l-1}$ ,  $X^0 = \emptyset$ , and  $m_{i_l}^{\Delta} = |X_{\Delta}^{i_l}|$ . The support nodes can be chosen in an hierarchical manner such that  $X^i \subset X^{i+1}$ ,  $i \in \{i_1, \dots, i_d\}$ .

It is also useful to compute the absolute ( $E_{abs}^n$ ) and relative ( $E_{rel}^n$ ) errors of the Smolyak interpolant using correction terms known as hierarchical surpluses ( $w_{\mathbf{j}}^{k,\mathbf{i}}$ ):

$$w_{\mathbf{j}}^{k,\mathbf{i}} = f(\mathbf{x}_{\mathbf{j}}^{\mathbf{i}}) - A_{k+d-1,d}(f(\mathbf{x}_{\mathbf{j}}^{\mathbf{i}})), \quad (6)$$

$$E_{abs}^n = \max_{\mathbf{i}, \mathbf{j}} w_{\mathbf{j}}^{n,\mathbf{i}}, \quad (7)$$

$$E_{rel}^n = \frac{\max_{\mathbf{i}, \mathbf{j}} w_{\mathbf{j}}^{n,\mathbf{i}}}{\max_{\mathbf{i}, \mathbf{j}} f(\mathbf{x}_{\mathbf{j}}^{\mathbf{i}}) - \min_{\mathbf{i}, \mathbf{j}} f(\mathbf{x}_{\mathbf{j}}^{\mathbf{i}})}. \quad (8)$$

The conventional sparse grid fails to consider the impact errors can have on the quality of the interpolant produced. Adaptive sparse grids [16] build on the conventional formulation by using generalized error indicators that consider the influence of the error in comparison to the necessary computational work:

$$g_{\mathbf{j}} = \max \left\{ w \frac{|\Delta^{\mathbf{j}} f|}{|\Delta^{\mathbf{1}} f|}, (1-w) \frac{n_{\mathbf{1}}}{n_{\mathbf{j}}} \right\}, \quad (9)$$

where  $w \in [0, 1]$  is a weight for the error indicator  $g_{\mathbf{j}}$ ,  $n_{\mathbf{k}}$  is the number of function evaluations for an index set  $\mathbf{k}$ . Conventional sparse grids are formed when  $w = 0$ , and only the number of function evaluations are considered. When  $w = 1$ , the error indicators will decay with increasing indices. Intermediate values of  $w$  compromise between excessive work and high error.

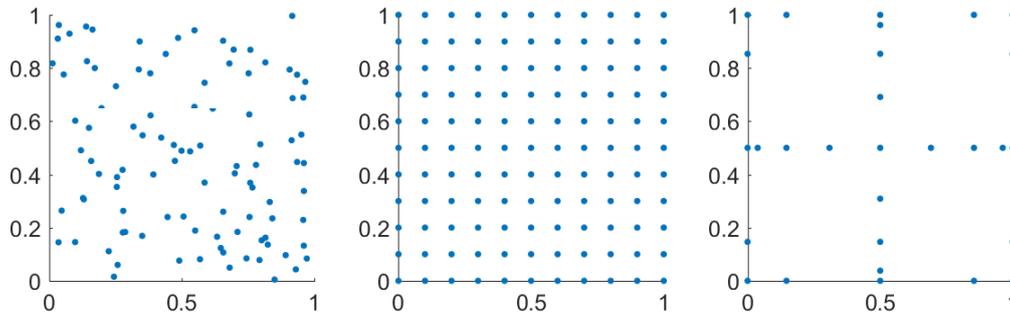


Fig. 2. Compared to randomly (*left*), and uniformly (*center*) sampled grids, sparsely sampled grids, like the Chebyshev-Gauss-Lobatto grid (*right*), strategically sample the parameter space to produce error controlled surrogate models that use fewer samples.

1) *Grid Type*: The approximation properties of the sparse grid rely on basis functions to select the required support nodes. Chebyshev-based node distributions can be used for higher-order polynomial interpolation, where the function to be interpolated is smooth and higher accuracy is required [32]. In this work, we use Chebyshev-Gauss-Lobatto nodes (Fig. 2) [30], which are defined as follows:

$$m_i = \begin{cases} 1, & i = 1 \\ 2^{i-1} + 1, & i > 1 \end{cases} \quad (10)$$

$$x_j^i = \begin{cases} -\cos \frac{\pi \cdot (j-1)}{m_i-1}, & m_i > 1 \\ 0, & m_i = 1, \end{cases} \quad (11)$$

where  $m_i$  is the number of support nodes for level  $i$ , and  $x_j^i$  is the position of the  $j^{\text{th}}$  node at level  $i$ ,  $j = 1, \dots, m_i$ .

2) *Time Domain Interpolation*: In addition to interpolation across the parameter space, there is also the issue of time domain interpolation. Choosing nodes in the time domain to accurately represent a trajectory may influence the accuracy of the resulting sparse grid interpolant. Time intervals can be either uniform or non-uniform. With non-uniform time points, a possibility is to utilize the extrema of the Chebyshev polynomials as was done in [14], [15] for ODE models:

$$T_s^\ell = T_{min}^\ell + \left(1 - \cos \left(\frac{\pi s_\ell}{d}\right)\right) \frac{T_{max}^\ell - T_{min}^\ell}{2}, \quad (12)$$

where  $\ell \in \{1, \dots, n\}$  is a vector of indices corresponding to model outputs,  $d$  is the degree of the interpolating Lagrange polynomial,  $T_s^\ell$  is a vector of sampling times,  $T_{min}^\ell$  is the minimum time,  $T_{max}^\ell$  is the maximum time, and  $s_\ell = [0, \dots, d]$ . Choosing the extrema of Chebyshev polynomials can reduce the effect of poor interpolation on the edges of an interval that occur when using equidistant nodes, a problem known as the Runge phenomenon [33].

Once the model outputs are sampled at these times, they can be evaluated at other times  $t$ ,  $T_{min}^\ell \leq t \leq T_{max}^\ell$ :

$$\tilde{y}_\ell(\boldsymbol{\theta}, t) = L_d^\ell(t) \cdot \hat{y}_\ell(\boldsymbol{\theta}, T_s^\ell), \quad (13)$$

where  $\tilde{y}_\ell(\boldsymbol{\theta}, t)$  is the interpolated model output with parameters  $\boldsymbol{\theta}$  at time  $t$ ,  $\hat{y}_\ell(\boldsymbol{\theta}, T_s^\ell)$  is the sparse grid model output sampled at the times  $T_s^\ell$ ,  $L_d^\ell$  is the Lagrange interpolating polynomial for the  $\ell^{\text{th}}$  model output with degree  $d$ , defined in [34].

3) *Simulation Conditions*: Matlab was used as the simulation environment for the models discussed here. The Euler-Maruyama method, a first-order stochastic Taylor expansion, was used to integrate SDEs [35]–[37]:

$$\mathbf{X}(t_{q+1}) = \mathbf{X}(t_q) + f(\mathbf{X}(t_q), q\delta t, \boldsymbol{\theta})\delta t + g(\mathbf{X}(t_q), q\delta t, \boldsymbol{\theta})(B(t_q) - B(t_{q-1})), \quad (14)$$

where  $\delta t$  is the integration time step, and  $q = 0, \dots, T/\delta t$ . Sparse grid interpolation was performed using the Sparse Grid Interpolation Toolbox [32].

Each model had to be tuned for compatibility with sparse grid interpolation by choosing both the simulation conditions and the number of realizations. Simulation conditions for the model, such as initial conditions, timespan of the simulation, desired model states, and parameters to include in the parameter space, were determined first. These conditions were defined in large part to conform with the scope of the examples presented in this work.

### III. COMPUTATIONAL EXPERIMENTS

#### A. SIVR Model

We first examine a model describing the spread of an infectious disease, known as the susceptible-infectious-vaccinated-removed (SIVR) model [38]. This particular system involves a vaccination mechanism by which certain individuals may avoid infection. It is described as follows:

$$dS = [\mu - \beta SI - (\mu + \phi)S]dt - \sigma SI dB(t) \quad (15)$$

$$dI = [\beta SI + \rho\beta VI - (\lambda + \mu)I]dt + \sigma(S + \rho V)I dB(t) \quad (16)$$

$$dV = [\phi S - \rho\beta VI - \mu V]dt - \rho\sigma VI dB(t) \quad (17)$$

$$dR = [\lambda I - \mu R]dt. \quad (18)$$

Susceptible individuals ( $S$ ) can contract the infection, after which they are infected ( $I$ ), and can infect other susceptible individuals. Vaccinated individuals ( $V$ ) may be partially resistant to infection upon vaccination, but are not completely immune. After recovering from an infection, removed individuals ( $R$ ) stay in the removed pool. The values of each disease state are expressed as percentages by normalizing to the overall population size. The parameters of interest in this model and the predefined parameter ranges are described in Table I. The stochastic perturbations in the SIVR model have been integrated into models of real-world diseases, such as HIV [39].

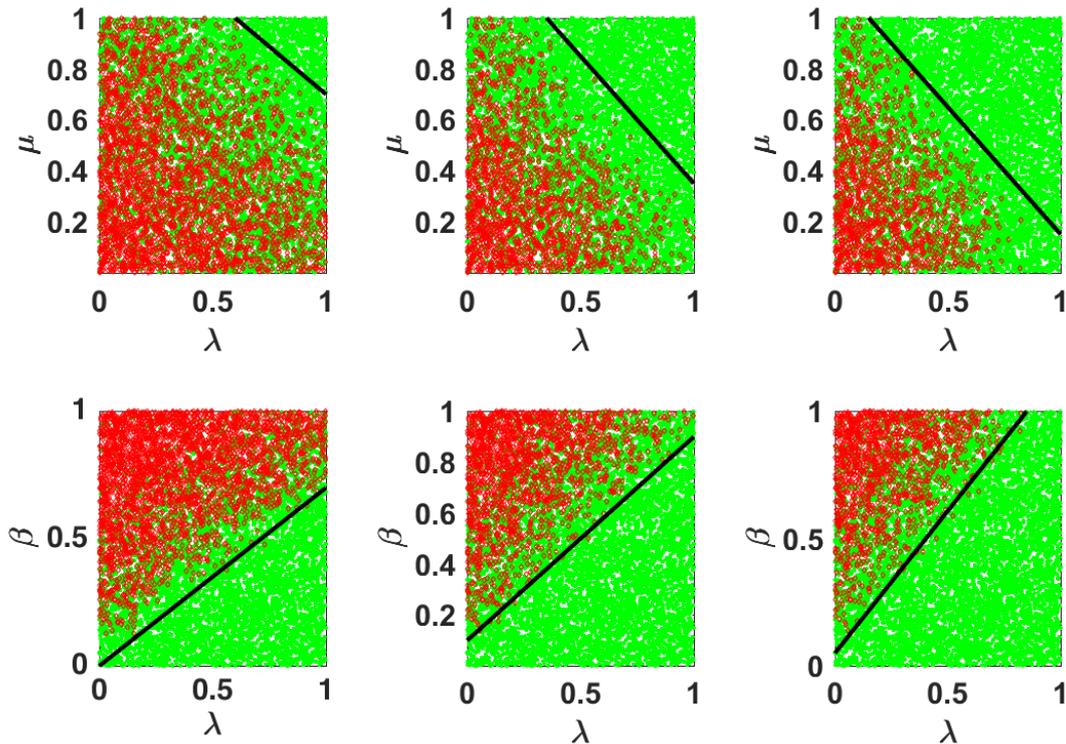


Fig. 3. Plots of the Acceptable and Unacceptable Parameter Sets in 2-dimensional form. Plots of (a,b,c)  $\lambda$  vs  $\mu$ , (d,e,f)  $\lambda$  vs  $\beta$ . Columns display results with  $\Xi = 0.2, 0.3$ , and  $0.4$ . Acceptable parameter sets are green dots, and unacceptable parameter sets are red diamonds. Black lines depict decision boundaries generated with SVMs.

For demonstration purposes, we investigate those epidemiological parameter values which result in the average number of cases being less than some percentage of the total population  $\Xi$ , an ideal disease eradication objective. Expressed mathematically, our target is:

$$\Theta_A = \{\theta \in \Theta \mid \mathbb{E}[C_\theta(T)] < \Xi\}, \quad (19)$$

where  $\theta = \{\lambda, \beta, \mu, \phi, \rho, \sigma\}$ , and  $\mathbb{E}[C_\theta(T)]$  is the expected number of cases simulated at time  $T$  with parameters  $\theta$ . We set  $T = 30$  days, with  $\mathbf{X}(0) = [0.8, 0.1, 0.05, 0.05]^\top$ . Additionally, we define  $C$  as follows:

$$dC = [\beta SI + \rho \beta VI]dt + \sigma(S + \rho V)IdB(t) \quad (20)$$

with  $C(0) = 0.1$ . The sparse grid interpolant produced had a relative error of 1% and an absolute error of 0.0692 with 6815 support nodes. Once the necessary number of realizations are found using the sparse grid interpolant, 10,000 parameter samples from the given ranges in Table I were obtained through Latin Hypercube Sampling (LHS). Then, model dynamics corresponding to these sampled parameters were interpolated using the surrogate model. Acceptable parameters satisfied the stated target, while unacceptable parameters did not.

Once the acceptable and unacceptable parameters were identified, a graphical analysis was done to determine if both sets can be visually separated, and if so, what was the nature of this separation in terms of parameter values. Fig. 3 elucidates the apparent disparity between acceptable and unacceptable parameter values in terms of the first three parameters listed in Table I, with increasing values of  $\Xi$ . A

TABLE I  
PARAMETERS OF SIVR MODEL, WITH DEFINITIONS AND RANGES USED IN SPARSE GRID INTERPOLATION.

| Parameter | Definition                         | Range         |
|-----------|------------------------------------|---------------|
| $\lambda$ | Fraction of recovered infected     | $[0, 1]$      |
| $\mu$     | Birth/death rate                   | $[0, 1]$      |
| $\beta$   | Transmission rate                  | $[0, 1]$      |
| $\phi$    | Fraction of vaccinated susceptible | $[0, 1]$      |
| $\rho$    | Vaccination effectiveness          | $[0, 1]$      |
| $\sigma$  | Environmental noise                | $[0.01, 0.2]$ |

decision boundary using support vector machines (SVMs) was also deployed to classify both types of parameter values and provide a clear visual partition.

The top row of Fig. 3 shows the relationship between  $\mu$  and  $\lambda$ . Clearly, higher values of both parameters, which translate to higher birth and recovery rates, tend to bode well for reduced case loads. Increasing  $\Xi$  from 0.2 to 0.4 boosts the number of acceptable parameters with higher values of both  $\lambda$  and  $\mu$ . As  $\Xi$  increases, those parameters with more moderate values are deemed acceptable, and the decision boundary advances towards lower values of both parameters.

On the other hand, the bottom row of Fig. 3 depicts a different scenario between  $\lambda$  and  $\beta$ . Obviously, a higher transmission rate (high  $\beta$ ), coupled with low recovery from infection (low  $\lambda$ ), led to more unfavorable outcomes. Lower levels of  $\lambda$  in general mean lower levels of recovery, leading to higher numbers of infected accumulating over the course of the simulation. This explains why lower levels of  $\lambda$  often lead to unfavorable outcomes.

A discussion of this example would be incomplete without

mentioning the role of noise. For the noise range considered here, there was no significant effect on transforming the boundary between acceptable and unacceptable parameters.

**B. MCF-7 Breast Cancer Model**

The MCF-7 breast cancer model was developed to predict tumor responses to radiotherapy and other therapeutic treatments [40]. To capture the deleterious and variable effects of radiation on cancer cells, the model added noise terms to the cell death rates for the three cancer sub-populations being studied. These sub-populations, sorted according to radiotherapy sensitivity, represented various stages of the cell cycle: the gap phase (*G*), the synthesis phase (*S*), and the mitosis phase (*M*). The model is described as follows:

$$dG = [-(\alpha + q_1)G + 2\gamma M]dt - \sigma G dB_1(t) \quad (21)$$

$$dS = [\alpha G - (\beta + q_2)S]dt - \sigma S dB_2(t) \quad (22)$$

$$dM = [\beta S - (\gamma + q_3)M]dt - \sigma M dB_3(t) \quad (23)$$

where  $q_i$ ,  $i = 1, 2, 3$  are the specific death rates for each sub-population,  $\alpha$  is the transition rate from *G* to *S*,  $\beta$  is the transition rate from *S* to *M*,  $\gamma$  is the transition rate from *M* to *G*, and  $\sigma$  is the magnitude of the stochastic noise.

TABLE II  
PARAMETERS OF MCF-7 MODEL, WITH DEFINITIONS AND RANGES USED IN SPARSE GRID INTERPOLATION.

| Parameter | Definition                                | Range             |
|-----------|---|-------------------|
| $\alpha$  | Transition rate from <i>G</i> to <i>S</i> | [-0.0052, 0.0918] |
| $\beta$   | Transition rate from <i>S</i> to <i>M</i> | [0.0315, 0.1333]  |
| $\gamma$  | Transition rate from <i>M</i> to <i>G</i> | [0.1744, 0.9055]  |
| $\sigma$  | Environmental noise                       | [0, 0.1]          |

In addition to introducing stochastic noise into the cancer model, [40] introduced a measure known as the tumor lifespan *L*, defined as the amount of time needed to eradicate the cancer:

$$L = \min\{t : G(t) + S(t) + M(t) = 0\}. \quad (24)$$

The tumor lifespan was introduced to evaluate cancer treatment effectiveness. Multiple treatment strategies could be ranked based on how much they reduced *L*. A mean tumor lifespan of 175 hours was calculated for the nominal parameters presented in [40].

While *L* has been evaluated on parameters found to best fit existing data on this form of breast cancer, understanding the impact that the stochastic noise term has on *L* would clarify its influence on cancer proliferation. To that end, we employ sparse grid interpolation to observe the tumor lifespan landscape for 200 MCF-7 cancer cells at the end of 200 hours with varying noise levels. The parameters used to form the parameter space, and their associated parameter ranges as reported in [40], are described in Table II. The sparse grid interpolant produced had a relative error of 0.93% and an absolute error of 0.1719 with 249 support nodes.

Fig. 4 illustrates this landscape in 3-dimensional form for 10,000 uniformly sampled points in the parameter space, with varying noise levels. If there were still cancer cells present at the end of 200 hours, the tumor lifespan was set to 200 hours. The top row, where only  $\gamma$  is varied, shows a clear discrepancy between areas of decreased tumor lifespan and the maximum plateau of 200 hours. Specifically, for

$\alpha \leq 0.01$  and  $\beta \leq 0.08$ , the tumor lifespan declines to as much as 110 hours. Lower transition rates tend to suspend cell viability and lifespan. Higher transition rates, on the other hand, retain the existing cellular machinery, promoting cell growth and division. Increasing the noise levels also did not significantly alter this landscape or the minimum lifespans.

Observing the tumor lifespan landscape for  $\alpha$  and  $\gamma$ , where  $\beta$  is held constant reveals some interesting features. The bottom row of Fig. 4 highlights two distinct regions of decreased tumor lifespan, where  $\alpha \leq 0.005$  and  $0.17 \leq \gamma \leq 0.28$ ,  $0.55 \leq \gamma \leq 0.9$ . The minimum lifespan attained in these areas are approximately 150 hours. While this area appears for all three noise levels, what differentiates each level is the prevalence of abnormal contours emblematic of noise. Noise pervades the decreased lifespan areas in the form of peaks, starting at the minima of both parameters. The quantity and width of these peaks increase as the noise levels increases.

**C. JAK-STAT Signaling Pathway**

Parameter estimation in systems biology aims to reconstruct dynamic inter- and intracellular biochemical relationships from available data [41], [42]. The JAK-STAT signaling pathway SDE, derived from an earlier ODE model [43], is described as follows [44], [45]:

$$dx_1 = [-k_1 x_1 EpoR + 2k_4 z_1]dt + \sigma x_1 dB(t) \quad (25)$$

$$dx_2 = [k_1 x_1 EpoR - k_2 x_2^2]dt \quad (26)$$

$$dx_3 = [-k_3 x_3 + \frac{1}{2} k_2 x_2^2]dt \quad (27)$$

$$dx_4 = [k_3 x_3 - k_4 z_1]dt \quad (28)$$

$$dz_1 = \Gamma(t)[x_3 - z_1]dt \quad (29)$$

$$\Gamma(t) = \frac{\alpha}{1 - A^\alpha \exp(-\alpha t)}. \quad (30)$$

This model of the JAK-STAT signaling pathway can be described by a number of steps [43]. Erythropoietin receptor (*EpoR*) is activated by erythropoietin hormone binding, phosphorylating cytoplasmic STAT5 ( $x_1$ ). Phosphorylated STAT5 ( $x_2$ ) then proceeds to dimerize ( $x_3$ ), after which it is then imported into the nucleus ( $x_4$ ). In the nucleus, dissociation and dephosphorylation of STAT5 occur with a time delay ( $z_1$ ).

TABLE III  
PARAMETERS OF JAK-STAT MODEL, WITH DEFINITIONS AND RANGES USED IN SPARSE GRID INTERPOLATION.

| Parameter | Definition                 | Range                     |
|-----------|----------------------------|---------------------------|
| $k_1$     | STAT5 phosphorylation rate | [0.015, 0.025]            |
| $k_2$     | STAT5 dimerization rate    | [0.015, 0.025]            |
| $k_3$     | Nuclear import rate        | [0.1, 0.15]               |
| $k_4$     | Nuclear export rate        | [0.05, 0.1]               |
| $\alpha$  | Delay function parameter   | [0.05, 0.5]               |
| <i>A</i>  | Delay function parameter   | [ $10^{-4}$ , $10^{-2}$ ] |
| $\sigma$  | Environmental noise        | [0.05, 0.2]               |

A readily measurable output of this system is the total phosphorylated STAT5 *y*, defined as follows:

$$y = s(x_2 + 2x_3), \quad (31)$$

where *s* is a scaling parameter.

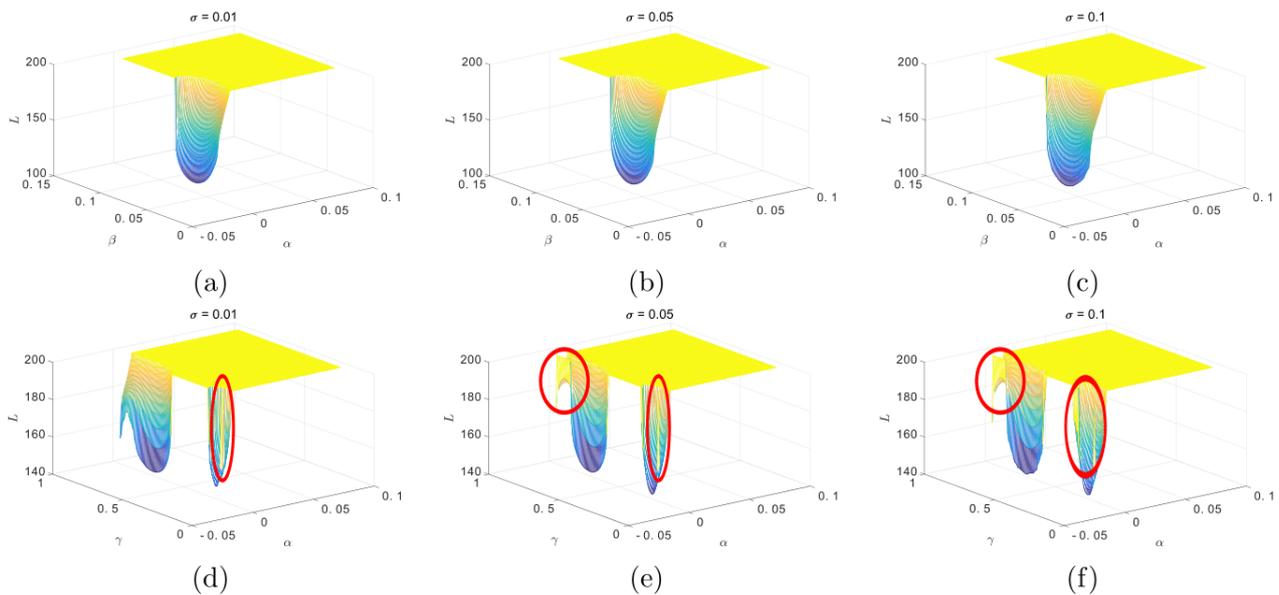


Fig. 4. Tumor lifespan landscape with varying noise levels. Top row varies  $\alpha$  and  $\beta$ , with  $\gamma = 0.3655$ . Bottom row varies  $\alpha$  and  $\gamma$ , with  $\beta = 0.0824$ . Red circles denote regions distorted by noise.

We rely on a nonparametric simulated maximum likelihood approach using kernel density estimation for parameter estimation [46]. The approach approximates the transition densities of the maximum likelihood function by comparing all generated realizations with observed data. The corresponding log likelihood function was then computed at the support nodes and subsequently interpolated across the parameter space described in Table III. The optimal parameter estimates minimized the log likelihood function. We set the duration of the simulation at 60 minutes, and  $\mathbf{X}(0) = [2.3, 0.01, 0.01, 0.01, 0]^T$ . Data obtained from [43] was used for parameter estimation. The sparse grid interpolant produced had a relative error of 0.52% and an absolute error of 0.38 with 481 support nodes. 10,000 LHS sampled parameters were generated from the prescribed parameter ranges, and the corresponding trajectories were estimated using the sparse grid interpolant. We plot and compare the results for three different noise levels, shown in Figure 5.

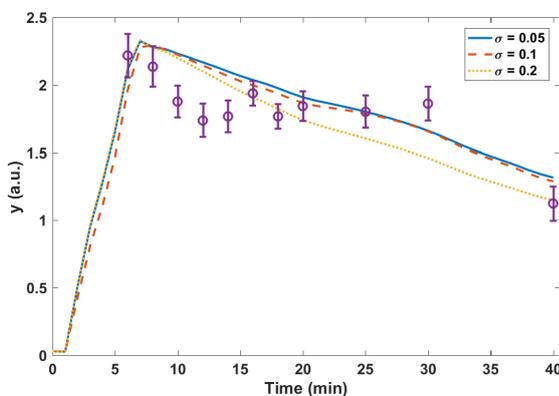


Fig. 5. Results of parameter estimation with JAK-STAT pathway model across three different noise levels. Dataset is in purple (mean  $\pm$  SD).

The log likelihood values for  $\sigma = 0.05, 0.1$ , and  $0.2$ , were  $6.1893 \times 10^{-4}$ ,  $4.353 \times 10^{-4}$ , and  $4.5854$ , respectively. Higher noise levels resulted in a dramatic loss of fit quantitatively,

although all noise levels possessed great qualitative fits. This example demonstrates the applicability of sparse grid interpolation to parameter estimation of SDEs within a maximum likelihood framework.

#### IV. CONCLUSION

Sparse grids produce effective interpolants without sacrificing much of the modeling accuracy and incurring the cost of unnecessary model evaluations. These unnecessary model evaluations materialize in both the parameter and uncertainty spaces, with multiple parameter values and realizations necessary for an adequate model description. The approach discussed here interpolates the solution provided by an average SDE trajectory at each support node in a parameter space of moderate dimension. The stochastic noise was also considered as a dimension of the parameter space, and played an important role in the examples presented. Our work serves as a computationally efficient surrogate modeling-based exploration of the stochastic dynamics of SDE models.

In the future, we hope to incorporate adaptive sampling of realizations in order to appropriately capture the variation present in the underlying systems. Furthermore, we plan to focus the sparse grid framework towards more targeted applications, such as sensitivity analysis and optimal control.

#### REFERENCES

- [1] R. M. May, "Uses and Abuses of Mathematics in Biology," *Science*, vol. 303, no. 5659, pp. 790–793, 2004. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/14764866>
- [2] H. W. Hethcote, "The Mathematics of Infectious Diseases," *SIAM Review*, vol. 42, no. 4, pp. 599–653, jan 2000. [Online]. Available: <http://epubs.siam.org/doi/abs/10.1137/S0036144500371907>
- [3] H. Kitano, "Computational Systems Biology," *Nature*, vol. 420, no. 6912, pp. 206–10, nov 2002. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/17052114>
- [4] T. Székely and K. Burrage, "Stochastic Simulation in Systems Biology," *Computational and Structural Biotechnology Journal*, vol. 12, no. 20-21, pp. 14–25, nov 2014. [Online]. Available: <http://dx.doi.org/10.1016/j.csbj.2014.10.003> <http://linkinghub.elsevier.com/retrieve/pii/S2001037014000403>

- [5] S. Ditlevsen and A. Samson, "Introduction to Stochastic Models in Biology," in *Springer-Verlag Berlin Heidelberg*, ser. Lecture Notes in Mathematics, M. Bachar, J. Batzel, and S. Ditlevsen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, vol. 2058, pp. 3–35. [Online]. Available: <http://link.springer.com/10.1007/978-3-642-32157-3> [http://link.springer.com/10.1007/978-3-642-32157-3\\_1](http://link.springer.com/10.1007/978-3-642-32157-3_1)
- [6] T. Manninen, M.-L. Linne, and K. Ruohonen, "Developing Itô stochastic differential equation models for neuronal signal transduction pathways," *Computational biology and chemistry*, vol. 30, no. 4, pp. 280–291, 2006.
- [7] A. Saarinen, M.-L. Linne, and O. Yli-Harja, "Modeling Single Neuron Behavior Using Stochastic Differential Equations," *Neurocomputing*, vol. 69, pp. 1091–1096, 2006.
- [8] —, "Stochastic Differential Equation Model for Cerebellar Granule Cell Excitability," *PLoS Computational Biology*, vol. 4, no. 2, p. e1000004, 2008.
- [9] O. Chis and D. Opris, "Mathematical Analysis of Stochastic Models for Tumor-Immune Systems," *Romania*, pp. 1–19, jun 2009. [Online]. Available: <http://arxiv.org/abs/0906.2794>
- [10] C. Ji, D. Jiang, and N. Shi, "A Note on a Predator-Prey Model with Modified Leslie-Gower and Holling-type II Schemes with Stochastic Perturbation," *Journal of Mathematical Analysis and Applications*, vol. 377, no. 1, pp. 435–440, 2011. [Online]. Available: <http://dx.doi.org/10.1016/j.jmaa.2009.05.039>
- [11] H. Zhou and M. Liu, "Analysis of a Stochastic Predator-Prey Model in Polluted Environments," *IAENG International Journal of Applied Mathematics*, vol. 46, no. 4, pp. 445–456, 2016.
- [12] A. K. Duun-Henriksen, S. Schmidt, R. M. Røge, J. B. Møller, K. Nørgaard, J. B. Jørgensen, and H. Madsen, "Model Identification using Stochastic Differential Equation Grey-Box Models in Diabetes," *Journal of diabetes science and technology*, vol. 7, no. 2, pp. 431–40, jan 2013.
- [13] M. M. Donahue, G. T. Buzzard, and A. E. Rundell, "Experiment design through dynamical characterisation of non-linear systems biology models utilising sparse grids," *IET systems biology*, vol. 4, no. 4, pp. 249–62, jul 2010. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20632775>
- [14] J. N. Bazil, G. T. Buzzard, and A. E. Rundell, "A Global Parallel Model based Design of Experiments Method to Minimize Model Output Uncertainty," *Bulletin of mathematical biology*, vol. 74, no. 3, pp. 688–716, mar 2012. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/21989566>
- [15] T. Mdluli, G. T. Buzzard, and A. E. Rundell, "Efficient Optimization of Stimuli for Model-Based Design of Experiments to Resolve Dynamical Uncertainty," *PLoS Computational Biology*, vol. 11, no. 9, p. e1004488, sep 2015. [Online]. Available: <http://dx.plos.org/10.1371/journal.pcbi.1004488>
- [16] T. Gerstner and M. Griebel, "Dimension-Adaptive Tensor-Product Quadrature," *Computing*, vol. 71, no. 1, pp. 65–87, aug 2003. [Online]. Available: <http://link.springer.com/10.1007/s00607-003-0015-5>
- [17] H.-J. Bungartz and M. Griebel, "Sparse grids," *Acta Numerica*, vol. 13, p. 147, 2004.
- [18] T. Gerstner and M. Griebel, "Sparse grids (Quantitative Finance)," *Encyclopedia of Quantitative Finance*, vol. 13, p. 5, 2008.
- [19] A. Klimke and B. Wohlmuth, "Algorithm 847: spinter: Piecewise Multilinear Hierarchical Sparse Grid Interpolation in MATLAB," *ACM Transactions on Mathematical Software*, vol. 31, no. 4, pp. 561–579, 2005. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1114275>
- [20] M. Andreucut, "Stochastic recovery of sparse signals from random measurements," *Engineering Letters*, vol. 19, no. 1, pp. 1–6, 2011.
- [21] F. Nobile, R. Tempone, and C. G. Webster, "A Sparse Grid Stochastic Collocation Method for Partial Differential Equations with Random Input Data," *SIAM Journal on Numerical Analysis*, vol. 46, no. 5, pp. 2309–2345, jan 2008. [Online]. Available: <http://epubs.siam.org/doi/abs/10.1137/060663660>
- [22] X. Ma and N. Zabarar, "An Adaptive Hierarchical Sparse Grid Collocation Algorithm for the Solution of Stochastic Differential Equations," *Journal of Computational Physics*, vol. 228, no. 8, pp. 3084–3113, 2009. [Online]. Available: <http://dx.doi.org/10.1016/j.jcp.2009.01.006>
- [23] N. Agarwal and N. R. Aluru, "A Domain Adaptive Stochastic Collocation Approach for Analysis of MEMS under Uncertainties," *Journal of Computational Physics*, vol. 228, no. 20, pp. 7662–7688, 2009. [Online]. Available: <http://dx.doi.org/10.1016/j.jcp.2009.07.014>
- [24] —, "Weighted Smolyak Algorithm for Solution of Stochastic Differential Equations on Non-uniform Probability Measures," *International Journal for Numerical Methods in Engineering*, vol. 85, no. 11, pp. 1365–1389, mar 2011. [Online]. Available: <http://doi.wiley.com/10.1002/nme.3019>
- [25] M. Liu, Z. Gao, and J. S. Hesthaven, "Adaptive Sparse Grid Algorithms with Applications to Electromagnetic Scattering under Uncertainty," *Applied Numerical Mathematics*, vol. 61, no. 1, pp. 24–37, 2011.
- [26] S. Sankaran and A. L. Marsden, "A Stochastic Collocation Method for Uncertainty Quantification and Propagation in Cardiovascular Simulations," *Journal of Biomechanical Engineering*, vol. 133, no. 3, p. 31001, 2011.
- [27] G. Zhang, M. Gunzburger, and W. Zhao, "A Sparse-Grid Method for Multi-Dimensional Backward Stochastic Differential Equations," *Journal of Computational Mathematics*, vol. 31, no. 3, pp. 221–248, 2013. [Online]. Available: <http://www.global-sci.org/jcm/volumes/v31n3/pdf/313-221.pdf>
- [28] R. Pulch, "Stochastic Collocation and Stochastic Galerkin Methods for Linear Differential Algebraic Equations," *Journal of Computational and Applied Mathematics*, vol. 262, pp. 281–291, 2014. [Online]. Available: <http://dx.doi.org/10.1016/j.cam.2013.10.046>
- [29] T. Gerstner and M. Griebel, "Numerical Integration using Sparse Grids," *Numerical Algorithms*, vol. 18, pp. 209–232, 1998.
- [30] V. Barthelmann, E. Novak, and K. Ritter, "High Dimensional Polynomial Interpolation on Sparse Grids," *Advances in Computational Mathematics*, vol. 12, pp. 273–288, 2000.
- [31] A. Klimke, K. Willner, and B. Wohlmuth, "Uncertainty Modeling using Fuzzy Arithmetic based on Sparse Grids: Applications to Dynamic Systems," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 12, no. 6, pp. 745–759, 2004.
- [32] A. Klimke, "Sparse Grid Interpolation Toolbox User's Guide," *IANS report*, 2006.
- [33] A. Gil, J. Segura, and N. M. Temme, *Numerical Methods for Special Functions*. Society for Industrial and Applied Mathematics, jan 2007. [Online]. Available: <http://epubs.siam.org/doi/abs/10.1137/1.9780898717822> <http://epubs.siam.org/doi/book/10.1137/1.9780898717822>
- [34] G. T. Buzzard and D. Xiu, "Variance-based Global Sensitivity Analysis via Sparse-Grid Interpolation and Cubature," *Communications in Computational Physics*, vol. 9, no. 3, pp. 542–567, 2011.
- [35] D. J. Higham, "An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations," *SIAM Review*, vol. 43, no. 3, pp. 525–546, jan 2001. [Online]. Available: <http://dx.doi.org/10.1137/S0036144500378302%5Cn> <http://epubs.siam.org/doi/abs/10.1137/S0036144500378302>
- [36] T. Sauer, "Numerical Solution of Stochastic Differential Equations in Finance," in *Handbook of Computational Finance*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, vol. 47, no. 1, pp. 529–550. [Online]. Available: [http://link.springer.com/10.1007/978-3-642-17254-0\\_19](http://link.springer.com/10.1007/978-3-642-17254-0_19)
- [37] —, "Computational Solution of Stochastic Differential Equations," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 5, no. 5, pp. 362–371, sep 2013. [Online]. Available: <http://doi.wiley.com/10.1002/wics.1272>
- [38] E. Tornatore, P. Vetro, and S. M. Buccellato, "SIVR Epidemic Model with Stochastic Perturbation," *Neural Computing and Applications*, vol. 24, no. 2, pp. 309–315, feb 2014. [Online]. Available: <http://link.springer.com/10.1007/s00521-012-1225-6>
- [39] D. Jiang, Q. Liu, N. Shi, T. Hayat, A. Alsaedi, and P. Xia, "Dynamics of a Stochastic HIV-1 Infection Model with Logistic Growth," *Physica A: Statistical Mechanics and its Applications*, vol. 469, pp. 706–717, mar 2017. [Online]. Available: <http://dx.doi.org/10.1016/j.physa.2016.11.078> <http://linkinghub.elsevier.com/retrieve/pii/S0378437116308846>
- [40] A. Oroji, M. Omar, and S. Yarahmadian, "An Itô stochastic differential equations model for the dynamics of the MCF-7 breast cancer cell line treated by radiotherapy," *Journal of Theoretical Biology*, vol. 407, pp. 128–137, 2016. [Online]. Available: <http://dx.doi.org/10.1016/j.jtbi.2016.07.035>
- [41] J. R. Banga, "Optimization in Computational Systems Biology," *BMC Systems Biology*, vol. 2, p. 47, jan 2008.
- [42] J. R. Banga and E. Balsa-Canto, "Parameter Estimation and Optimal Experimental Design," *Essays In Biochemistry*, vol. 45, pp. 195–210, sep 2008. [Online]. Available: <http://onlinelibrary.wiley.com/doi/10.1002/wics.100/full> <http://essays.biochemistry.org/lookup/doi/10.1042/bse0450195>
- [43] I. Swameye, T. G. Muller, J. Timmer, O. Sandra, and U. Klingmuller, "Identification of Nucleocytoplasmic Cycling as a Remote Sensor in Cellular Signaling by Databased Modeling," *Proceedings of the National Academy of Sciences*, vol. 100, no. 3, pp. 1028–1033, feb 2003. [Online]. Available: <http://www.pnas.org/cgi/doi/10.1073/pnas.0237333100>
- [44] C. Surulescu, "On Some Stochastic Differential Equation Models with Applications to Biological Problems," ICAM, WWU Muenster, Tech. Rep., 2011.
- [45] C. Surulescu and N. Surulescu, "Some Classes of Stochastic Differential Equations as an Alternative Modeling Approach to

Biomedical Problems,” 2013, pp. 269–307. [Online]. Available:  
[http://link.springer.com/10.1007/978-3-319-03080-7\\_9](http://link.springer.com/10.1007/978-3-319-03080-7_9)

- [46] A. S. Hurn, K. A. Lindsay, and V. L. Martin, “On the efficacy of simulated maximum likelihood for estimating the parameters of stochastic differential equations,” *Journal of Time Series Analysis*, vol. 24, no. 1, pp. 45–63, 2003.