Article

Optimal Experimental Design for Parameter Estimation of an IL-6 Signaling Model

Andrew Sinkoe [1,2] **and Juergen Hahn** [1,2,3,]*

[1] Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
[2] Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute,
  Troy, NY 12180, USA
[3] Department of Chemical & Biological Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
[*] Correspondence: hahnj@rpi.edu; Tel.: +1-518-276-2138

Abstract: IL-6 signaling plays an important role in inflammatory processes in the body. While a number of models for IL-6 signaling are available, the parameters associated with these models vary from case to case as they are non-trivial to determine. In this study, optimal experimental design is utilized to reduce the parameter uncertainty of an IL-6 signaling model consisting of ordinary differential equations, thereby increasing the accuracy of the estimated parameter values and, potentially, the model itself. The D-optimality criterion, operating on the Fisher information matrix and, separately, on a sensitivity matrix computed from the Morris method, was used as the objective function for the optimal experimental design problem. Optimal input functions for model parameter estimation were identified by solving the optimal experimental design problem, and the resulting input functions were shown to significantly decrease parameter uncertainty in simulated experiments. Interestingly, the determined optimal input functions took on the shape of PRBS signals even though there were no restrictions on their nature. Future work should corroborate these findings by applying the determined optimal experimental design on a real experiment.

Keywords: optimal experimental design; D-optimality criterion; Fisher information matrix; sensitivity analysis; IL-6 signaling; parameter estimation; piecewise constant functions

1. Introduction

Mathematical models of intracellular signaling pathways are important for understanding and predicting how cells respond to certain stimuli. Such models can be modified readily when new findings become available, and can be a useful tool in directing new studies based on hypotheses generated by the model's predictions.

Interleukin-6 (IL-6) is a cytokine involved in a variety of inflammatory processes [1–4]. Understanding the signaling pathways associated with extracellular IL-6 excitation is important for elucidating and modulating the biological response to inflammation. Because chronic inflammation can cause tissue damage and poses a serious health risk during chronic infection or autoimmune conditions [5–7], therapeutic treatments for chronic inflammation are an active area of research. These efforts can be augmented through the use of mathematical models for inflammatory signaling, where IL-6 plays a major role.

An IL-6 signaling model (Appendix A) was previously developed in [1–4]. Originally, a detailed model containing 77 ordinary differential equations (ODEs) and 128 parameters was derived, followed by a model simplification procedure to decrease the number of ODEs and parameters to 13 and 19, respectively. The simplified model contains only the variables and parameters deemed necessary for representing the dynamics of the signaling system, as determined by parameter sensitivity

analysis and observability analysis [2]. This simplified model is the subject of the present study, and will henceforth be referred to as the 'IL-6 model' or the 'model' while the original detailed model will be referred to as the 'original IL-6 model' or the 'original model'.

In this study, optimal experimental design was applied to increase the accuracy of the 19 parameters of the IL-6 model. The parameters in this model correspond to rate constants for the chemical reactions in the signaling pathway represented by the model. By increasing the accuracy—and conversely reducing the uncertainty—of the model parameters, it is possible to obtain a more accurate model through iterative model adjustment. To demonstrate and examine this step in the model development process, optimal experimental design methods were applied in this study to minimize the uncertainty of the IL-6 model parameters estimated from least squares with experimental data.

Optimal experimental design has been utilized for decades in a variety of settings in which it is of interest to maximize efficiency of resource use and obtain a significant amount of information from experiments with acceptable cost [8–16]. Recently, as biological modeling and systems biology have emerged as an important area in biomedical research, optimal experimental design applied to biological experimental systems has become more popular [17–28]; additionally, optimal experimental design has been recognized as a valuable tool in optimal control for several decades [29]. For example, Jones et al. [13] maximized production of an exogenous commodity chemical in metabolically engineered *E. coli* using an empirical modeling method similar to those used in [15,16] to maximize the efficacy of drug delivery. Weber [26] utilized optimal experimental design to maximize model prediction accuracy for a model of vesicle transport via the *trans*-Golgi network. Bandara [18] performed optimal experimental design to reduce parameter uncertainty in a model of phosphatidylinositol 3,4,5-trisphosphate signaling. These studies demonstrate the effectiveness of optimal experimental design for obtaining maximally informative experimental data.

Here, optimal experimental design is applied to the problem of maximizing parameter accuracy for the IL-6 model. In particular, the D-optimality criterion, applied to the Fisher information matrix (FIM), is maximized over a set of IL-6 concentration input functions to determine an optimal dynamic IL-6 input profile for exciting the signaling system to generate data for least squares estimation of the model parameters [30]. The experimental design constraints considered in the problem are based on available resources and limitations present in a typical laboratory capable of performing the designed experiments. Namely, it is assumed that (a) the measurements of protein concentrations for two transcription factors are recorded as a time series; (b) the sampling time between measurements does not change during the course of an experiment; (c) the model represents signal transduction in rat hepatocytes which are stimulated in vitro with IL-6 following a dynamic input function; and (d) the IL-6 concentration is kept below cytotoxic levels. A piecewise constant IL-6 input function was computed by solving the optimal experimental design problem with the D-optimality criterion, operating on the Fisher information matrix, as the objective function. Since the Fisher information matrix contains only local information, a sensitivity matrix was also computed using the Morris method and the D-optimality criterion was applied to the sensitivity matrix as well in order to corroborate that the results are not just local in nature. The optimal IL-6 input function was found to substantially decrease the model parameter uncertainty in simulated least squares fits compared to a constant stimulation with IL-6.

The paper is organized as follows: optimal experimental design and its application to ODE models are presented in Section 2; application of optimal experimental design to the specific case of the IL-6 model is presented in Section 3 and results from solving the optimal experimental design problem are presented in Section 4, including simulation results for calculating parameter uncertainty; Section 5 discusses implications and suggests further applications of the optimal experimental design methodology presented here. The IL-6 model is included for reference in Appendix A.

2. Optimal Experimental Design

An ordinary differential equation model, e.g., for describing signal transduction, can be written as

$$\frac{dx}{dt} = f(t, x, u, p); \; y = g(t, x, u, p), \tag{1}$$

where x is a time-dependent vector of state variables, u is a time-dependent controlled input to the system, p is a vector of constant parameters, and y is a vector of measured quantities related to the model variables. Often, as is the case here, y is simply a subset of state variables from the vector x that are measured experimentally over time. While these measurements y are discrete in nature, the assumption that y is a continuous variable can be made if the sampling frequency is sufficiently high.

Optimal experimental design involves maximizing a criterion function that indicates the quantity of information gained by a given experiment, often in the context of model identification [8–12,14,17–23,29,31]. Several commonly used criterion functions for experimental design exist. One of these is A-optimality, which seeks to minimize the trace of the inverse of the Fisher information matrix. This criterion results in minimizing the average variance of the estimates of the regression coefficients. Another popular approach is E-optimality which maximizes the smallest eigenvalue of the Fisher information matrix. However, the most popular approach is D-optimal experimental design, i.e., which maximizes the determinant of the FIM [14,19,24]. The D-optimality criterion was chosen for this work as it seeks to minimize the covariance of the parameter estimates for a specified model [14]. It should be noted that no experimental design is optimal in all aspects and maximizing one particular criterion often negatively affects the experimental design of other criteria. That being said, the D-optimality criterion has found widespread use in practice as it results in experimental designs that have many of the properties that one usually looks for. For numerical considerations, the natural logarithm of the determinant of the FIM was taken for the problem presented. The optimal experimental design objective can be written generally as

$$\max \varphi_D(\mathbf{F}), \tag{2}$$

where $\mathbf{F}$ is the FIM, defined in more detail below, and φ_D is the D-optimality criterion.

Taking the determinant of the parameter covariance matrix corresponds to computing the volume of the parameter space that would allow for a solution to the least squares parameter fitting problem [14,18]. The volume of this parameter space represents the covariance, or uncertainty, of the parameters. Because the FIM is related to the inverse of the covariance matrix [8,10], maximizing the determinant of the FIM results in minimizing the determinant of the covariance matrix, and thus minimizing the volume of the parameter space, or minimizing parameter uncertainty.

The FIM is computed from the sensitivity coefficients of the model, which can be an ODE model as shown in Equation (1). Local sensitivity coefficients are defined as

$$\left. \frac{\partial x_i}{\partial p_j} \right|_{t,p}, \tag{3}$$

where x_i is the i-th state variable in the model, p_j is the j-th model parameter in the vector p, and t is the time at which the partial derivative is evaluated [2,3,32]. The local sensitivity coefficient represents the change in a model state with respect to a change in the value of a parameter, and is a function of time and the parameter vector. At every time point during an experiment, or a model simulation, a sensitivity coefficient can be calculated for any of the state variables. However, the FIM contains the sensitivity coefficients for only those state variables which are measured in experiments for generating data to be used in parameter estimation using least squares fitting, i.e., the sensitivity coefficients for the vector y. The FIM is written as

$$\mathbf{F} = \mathbf{S}^\mathrm{T}\mathbf{S}, \tag{4}$$

where **S** is the sensitivity matrix [32]

$$
\mathbf{S} = \begin{bmatrix}
\frac{\partial y_1}{\partial p_1}(t1) & & \frac{\partial y_1}{\partial p_{n_p}}(t1) \\
\vdots & & \vdots \\
\frac{\partial y_1}{\partial p_1}(tN) & & \frac{\partial y_1}{\partial p_{n_p}}(tN) \\
& \cdots & \\
\frac{\partial y_2}{\partial p_1}(t1) & & \frac{\partial y_2}{\partial p_{n_p}}(t1) \\
\vdots & & \vdots \\
\frac{\partial y_2}{\partial p_1}(tN) & & \frac{\partial y_2}{\partial p_{n_p}}(tN)
\end{bmatrix} \in \mathbb{R}^{(2N) x n_p},
\tag{5}
$$

and n_p is the number of parameters in the model. Here, the formula shown for **S** corresponds to an experiment in which two state variables are measured in a time series from time t_1 to t_N; one can extend the formula to the general case of any number of experimentally measured outputs, however, two outputs are realistic for the application investigated in this paper. The local sensitivity coefficients for all of the state variables of the model can be calculated by solving the system of ordinary differential equations

$$
\frac{d}{dt}\frac{\partial x}{\partial p^{\mathrm{T}}} = \frac{\partial f}{\partial x^{\mathrm{T}}}\frac{\partial x}{\partial p^{\mathrm{T}}} + \frac{\partial f}{\partial p^{\mathrm{T}}}
\tag{6}
$$

where x is the column vector of state variables, p is the column vector of model parameters, and f is the column vector of functions defining the model, as in Equation (1) [32]. To calculate the D-optimality criterion value for a given experiment design, the model ODEs (Equation (1)) are solved simultaneously with the sensitivity equations (Equation (6)), the FIM is constructed (Equations (4) and (5)), and the determinant can then be computed. To determine an optimal experiment, an optimization problem is solved which searches through different experimental designs, computing the D-optimality criterion in this way for every iteration.

The Morris method [3] can also be utilized for calculating sensitivity coefficients in a covariance matrix. It should be noted that such an approach results in more than local information due to the certain properties of the Morris method. Using the Morris method, global sensitivity coefficients are calculated by repeatedly sampling parameter values from a distribution. Finite difference approximations of the local sensitivities are then computed at each value of the sampled parameter vector, and the finite difference approximations are averaged to obtain each global sensitivity coefficient as shown in Equations (7) and (8).

$$
d_{ijk} = \frac{y_i(t, p_1, \ldots, p_j + \Delta_{jk}, \ldots, p_{n_p}) - y_i(t, p_1, \ldots, p_j, \ldots, p_{n_p})}{\Delta_{jk}}
\tag{7}
$$

$$
s_{ij} = \frac{1}{n_d}\sum_{k=1}^{n_d} d_{ijk}
\tag{8}
$$

Here, d_{ijk} is a finite difference derivative approximation for the local sensitivity of y_i with respect to parameter p_j at the k-th sampled value of p_j, Δ_{jk} defines the sampled value of p_j depending on the distribution of p_j, and n_d is the number of samples chosen for p_j. This method for calculating sensitivity coefficients takes into account the uncertainty of the parameter values by averaging over samples from the parameter distribution. Therefore, using the Morris method for calculating the sensitivity coefficients in the FIM does not rely on knowledge of an exact value for the parameters. In fact, the exact values of the parameters are by definition unknown—otherwise, one would not need to estimate the parameters. For this problem, the parameter distributions were considered to be normal with a mean at the nominal parameter values and a standard deviation of one-tenth the value of each parameter.

3. Formulation of the Optimal Experimental Design Problem for the IL-6 Model

The values of the parameters p in ODE models are generally unknown and must be estimated using experimentally measured data points [20,30,33]. The problem considered here is to optimize the function for the controlled variable u in order to minimize the uncertainty of the parameters estimated by least squares parameter estimation.

In the IL-6 model, u represents the IL-6 concentration in the local environment of the cells being stimulated by IL-6. The IL-6 concentration at a given time determines the signaling behavior of the cells, which is characterized by the concentrations of each signaling molecule in the pathway at time t. Thus by modulating the input IL-6 concentration, u, as a function of time, the signaling dynamics can be influenced so that measurements of components of the signaling pathway can yield maximal information about the model parameters.

The measured state variables for this problem were chosen to be STAT3N*-STAT3N* and C/EBPβ, two transcription factors of the IL-6 signaling pathway which are directly involved in transcribing DNA to RNA. These measurements were chosen because of the availability of a fluorescent reporter system for measuring the concentrations of these two proteins [2,34]. To design an optimal experiment for minimizing parameter uncertainty, experiments were considered in which the concentrations of these two proteins are measured in a time series every 45 min for 22 h using a fluorescent reporter system. This time sequence for image acquisition was utilized previously to obtain initial estimates for the parameters of the model [2]. However, for the initial parameter estimation, a constant IL-6 concentration at an arbitrary level of 100 ng/mL was utilized as the input function [2]. In the present work, optimal experimental design was applied to optimize the input function over a continuous set of time-dependent input functions while utilizing the same time sequence for data acquisition as was utilized in [2]. This allows for evaluating whether the parameter uncertainty can be decreased by optimizing the input function while holding other experimental control decisions constant.

For this problem, the set of input functions to optimize was chosen to be the piecewise constant input functions with fixed and equal time intervals and IL-6 concentrations bounded between 0 and 7.5 nM. These input functions were chosen because they can be implemented experimentally due to the long intervals during which a concentration is constant, and because they can be parameterized so as to arrive at an optimization problem with a finite number of variables. Specifically, the number of optimization variables is the number of concentration levels allowed in the piecewise constant function. This is illustrated in Figure 1 with a piecewise constant input function for IL-6 that has $r = 4$ concentration levels. As an example, to determine the optimal input function with $r = 4$ concentration levels, the optimal experimental design problem would be solved with four optimization variables representing the four concentration levels of such an input function.

These piecewise constant input functions can be written as

$$
\begin{aligned}
u(t) &= \sum_{k=1}^{r} c_k step(t - (k-1)\Delta t) - \sum_{k=1}^{r} c_k step(t - k\Delta t) \\
&= c_1 step(t) + \sum_{k=2}^{r} [c_k step(t - (k-1)\Delta t) - c_{k-1} step(t - (k-1)\Delta t)] - c_r step(t - r\Delta t) \qquad (9) \\
&= c_1 step(t) + \sum_{k=2}^{r} (c_k - c_{k-1}) step(t - (k-1)\Delta t) - c_r step(t - r\Delta t)
\end{aligned}
$$

where c_k is the k-th concentration level in the vector c, '*step*' is the Heaviside step function, r is the number of concentration levels in the input function, and $\Delta t = 22$ h/r is the time interval for each concentration level in the input function.

By changing the vector c to modulate the IL-6 input function, the solution of the model ODEs and local sensitivity ODEs are modulated. This causes the FIM, and thus the D-optimality criterion value,

to be a function of *c*. Therefore, the optimal experimental design problem for minimizing parameter uncertainty in the IL-6 model can be written as

$$\max_{0 \leq c \leq 7.5 \text{ nM}} \ln |\mathbf{F}(c)|, \tag{10}$$

where $|\cdot|$ is the determinant and **F** is calculated from Equations (4) and (5) with y_1 from Equation (5) being the concentration of STAT3N*-STAT3N*, y_2 being the concentration of C/EBPβ, and *p* being the parameters of the IL-6 model. Specifically, in order to calculate the FIM, the IL-6 signaling ODE model is numerically integrated simultaneously with the local sensitivity equations (Equation (6)) for a given input function represented by *c*, and the FIM is constructed according to Equations (4) and (5). This ODE integration is carried out in every iteration of the optimization problem to evaluate the D-optimality criterion value until the optimization solver determines a solution.
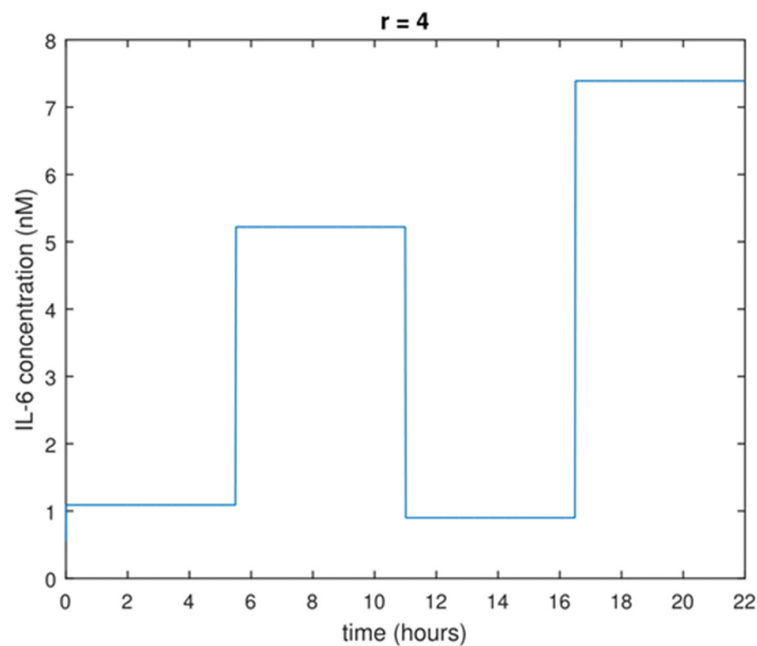
Figure 1. Example of a piecewise constant input function for IL-6 (Interleukin-6) concentration with *r* = 4. The input function shown here can be represented as a vector of the four concentration levels in chronological order: *c* = [1.1, 5.2, 0.9, 7.4] nM IL-6.

When solving the problem using the Morris method to obtain global sensitivity coefficients, the local sensitivity coefficients in the FIM are replaced by the global sensitivity coefficients calculated by the Morris method. For the Morris method, the local sensitivity ODEs do not need to be solved; rather, the model ODEs are solved repeatedly using the sampled parameter values, and the finite difference approximations of the partial derivatives are computed and averaged.

To solve the optimal experimental design problem for each value of *r* from 1 to 6, the MATLAB function '*fminsearch*' was utilized [35]. Additionally, to account for the bound constraints on the concentration levels, a stop flag was imposed in the program for running the optimization solver. Multiple initial guesses were utilized for each value of *r* to avoid local optima (between 5 and 18 initial guesses were used depending on *r*, in order to cover the range of possible qualitative input function shapes for each value of *r*).

4. Optimal Experimental Design for the IL-6 Signaling Model

Solving the optimal experimental design problem for minimizing parameter uncertainty in the IL-6 model for values of *r* from 1 to 6 resulted in an optimal IL-6 input function for each value of *r*.

The optimal solutions for each *r* are listed in Table 1, which also lists several sub-optimal input functions for comparison. Note that the D-optimality criterion value is greatest for the optimal input function with *r* = 6 (see columns 1–4 in Table 1). The range of D-optimality criterion values for the input functions listed in the table (column 4) is very wide if one considers that the values given are the logarithm of the determinant of the FIM rather than the determinant itself. As expected, the D-optimality criterion values for the optimal input functions increase with *r* (see Table 1, column 4); this is due to the degrees of freedom added when *r* is increased.

Table 1. Optimal and sub-optimal input functions for values of *r* from 1 to 6.

Input Function (nM)	*r*	Optimal?	D-Optimality Criterion Value	Covariance Norm	Covariance Trace
3.83	1	no	−133.9	161	216
6.59	1	yes	−114.5	43.8	45.2
[0.48, 7.34]	2	yes	−28.9	12.95	12.96
[6.30, 1.60, 7.20]	3	yes	0.30	0.0033	0.0040
[1.09, 5.22, 0.90, 7.39]	4	yes	3.24	6.56×10^{-5}	6.59×10^{-5}
[0.99, 5.64, 0.95, 7.37, 0.97]	5	yes	18.8	1.90×10^{-6}	1.91×10^{-6}
[0.89, 5.94, 1.01, 7.14, 1.07, 5.22]	6	yes	29.7	4.26×10^{-6}	4.83×10^{-6}
[6.97, 0.67, 7.23, 0.88, 7.49, 1.09]	6	no	19.8	0.195	0.204
[7, 6, 5, 4, 3, 2]	6	no	−41.4	42.2	46.8
[1, 4, 6, 6, 4, 1]	6	no	−32.2	1.63	2.03
[3, 4, 5, 6, 7]	5	no	−47.9	55.5	68.8
[6, 6, 5, 3, 3]	5	no	−47.5	1.32	1.44

To test whether higher D-optimality criterion values for an input function corresponded to lower uncertainty, simulations were run for each of the input functions listed in Table 1. For these simulations, data were generated by utilizing the original IL-6 model and adding normally distributed noise to the measurements. Specifically, for a given input function, a simulation was run by integrating the original IL-6 model, adding Gaussian noise with a mean of 0 and a standard deviation of 1 at the measurement time points (every 45 min for 22 h), and fitting the parameters of the simplified IL-6 model to the simulated data. The standard deviation of 1 was chosen to provide a reasonable signal to noise ratio for the model variables. For each input function, 30 simulations were run and parameters were fitted for each simulation. A parameter covariance matrix was then calculated for each of the input functions using the parameter fits from the 30 simulations that were run for each input function. The norm and trace of these covariance matrices were utilized as measures of uncertainty for the parameters. The norm accounts for covariances, while the trace takes into account only variances. Columns 5 and 6 of Table 1 show the norm and trace of the covariance matrix for each input function. As expected, one can observe that the uncertainty is generally lower for optimal input functions than for sub-optimal input functions. However, the relationship between the (*a priori* computed) D-optimality criterion and the (*a posteriori* determined) uncertainty is not monotonic; e.g., the uncertainty for the *r* = 5 optimal input function is slightly lower than that for the *r* = 6 input function while the D-optimality criterion value is greater for the *r* = 6 input function (Table 1). Furthermore, the uncertainty for the *r* = 6 sub-optimal input function shown is larger than that for the optimal *r* = 4 and *r* = 5 input functions even though the D-optimality criterion value is greater for the *r* = 6 sub-optimal input function. There are a number of possible reasons for this lack of monotonicity, such as nonlinearity of the IL-6 model and the fact that optimal experimental design theory is an approximate theory in the case of nonlinear models [8,10].

In order to corroborate that these findings also hold up when global rather than local analysis is used, the Morris method was implemented for computing sensitivity coefficients. This resulted in the optimal input functions listed in Table 2. Figure 2 shows the optimal input functions computed by both the local method and the Morris method for *r* from 1 to 6. It was observed that both methods resulted in optimal input functions with alternating IL-6 concentration levels, reminiscent of pseudo-random binary signals (PRBS signals). While the different methods sometimes result in input profiles that start

high vs. low, the general shape of the functions is an oscillation between a low and a high value of the input. This suggests that PRBS-like signals may be favorable for minimizing parameter uncertainty in the IL-6 signaling model.
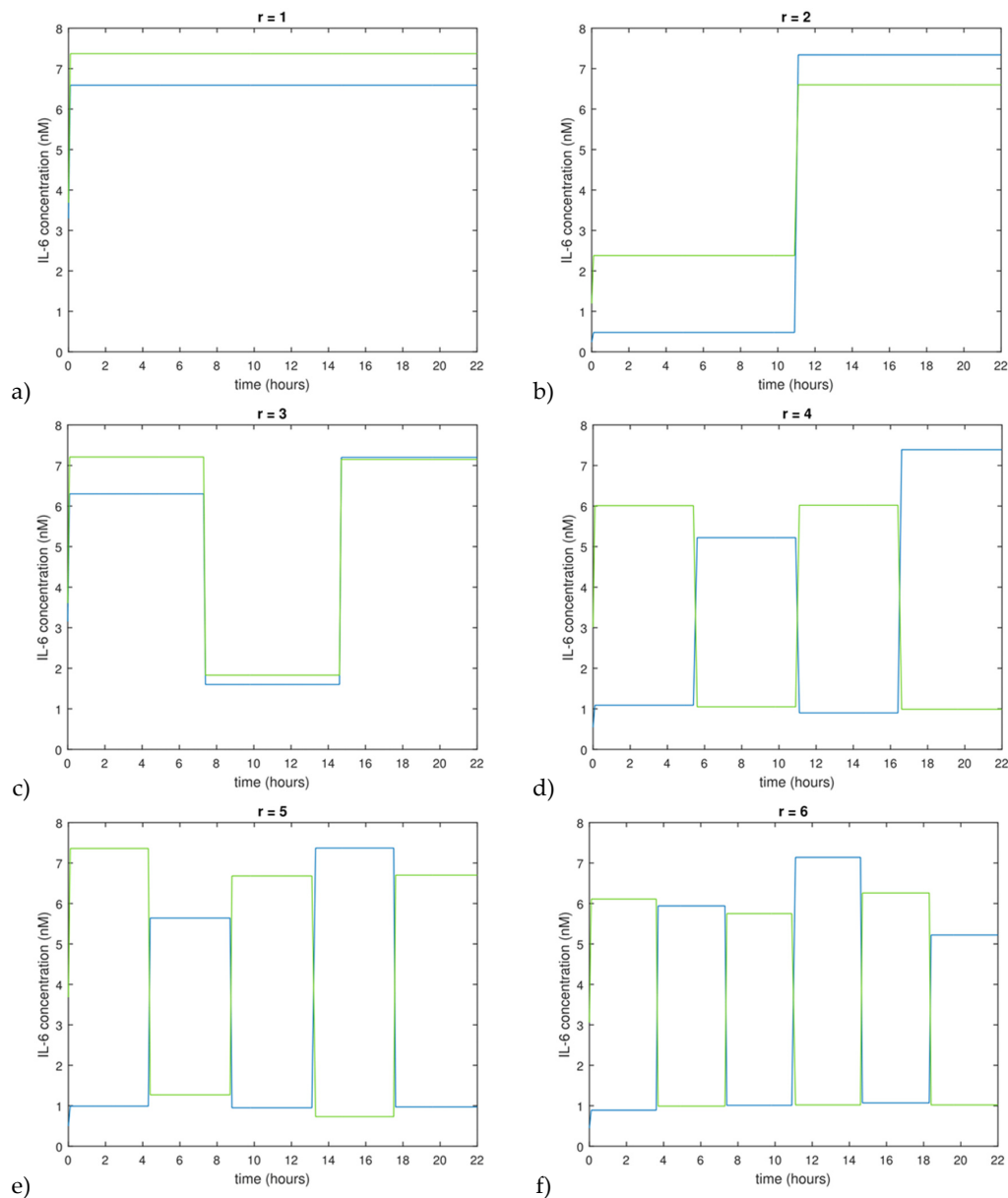
Figure 2. Optimal input functions for values of r from 1 to 6. Blue indicates input functions identified as optimal from the local sensitivity method, green indicates input functions identified as optimal from the Morris method. (**a**) $r = 1$ (**b**) $r = 2$ (**c**) $r = 3$ (**d**) $r = 4$ (**e**) $r = 5$ (**f**) $r = 6$.

Table 2. Optimal input functions for values of r from 1 to 6 via the Morris method

Input Function (nM)	r
7.37	1
[2.38, 6.60]	2
[7.21, 1.83, 7.15]	3
[6.01, 1.05, 6.02, 0.99]	4
[7.36, 1.27, 6.68, 0.73, 6.70]	5
[6.11, 0.99, 5.75, 1.02, 6.26, 1.02]	6

An important take-away message from the simulation results is that optimal experimental design can be very effective at providing an a priori design, yet this does not replace evaluating the quality of a design a posteriori.

5. Discussion

The IL-6 model aims to capture the behavior of molecules involved in both the Jak-STAT pathway and the MAPK pathway [2]. This model is larger and more nonlinear than most models that have previously been the subject of optimal experimental design involving signaling pathways [18,24,25]. The present study shows that optimal experimental design can be utilized for improving the parameter accuracy of moderately complex nonlinear models. Further, the study shows that parameter uncertainty can be substantially decreased by optimizing the input function alone without simultaneously optimizing the experimental measurement time points for this particular system. This observation potentially has fortunate implications for experimental systems in which the input function can be controlled but the experimental measurement time points are dictated by practical considerations and cannot be changed.

The main step in the optimal experimental design was the sensitivity analysis for constructing the FIM. This involves constructing the Jacobian of the model equations so that the sensitivity equations can be integrated to obtain the sensitivity coefficients for the FIM. In principle, this method can be applied to any ODE model. The main step in the analysis is therefore to calculate all of the partial derivatives in the Jacobian, since the Jacobian will be different for every ODE model. For the Morris method, the information matrix is constructed by sampling from the parameter distribution and calculating finite difference derivative approximations for the sensitivities. In both cases—i.e., local and global methods—an ODE solver capable of accurately integrating the ODEs must be used, along with the optimization solver in order to evaluate the objective function for each optimization iteration.

Optimal experimental design using the D-optimality criterion was effective, according to the simulation results, in decreasing the parameter uncertainty in the model. However, one has to caution that this is only one aspect. For example, it has been shown in work by White et al. [36] that increased parameter accuracy does not guarantee an improvement in the accuracy of the model.

The form of the optimal input functions for the IL-6 model raises an interesting point regarding general system identification theory. Each of the determined optimal input functions takes the form of a PRBS-like sequence (see Figure 2), a commonly used input signal for system identification [37], even though this input function shape was not postulated during the formulation of the optimal design problem. Furthermore, computing an experimental design where sensitivities were computed via the Morris method also led to PRBS-like optimal inputs. Without commenting on generality, a PRBS signal seems to be a good choice for inputs of the investigated IL-6 signaling model, a result that may potentially carry over to other signaling pathway models.

6. Conclusions

Optimal experimental design was applied to the problem of minimizing parameter uncertainty in an IL-6 signaling model representing the Jak-STAT and MAPK pathways. The D-optimality criterion, operating on the FIM, was constructed using sensitivity equations, which were solved simultaneously with the equations of the model. Piecewise constant input functions were determined by solving this optimization problem; the piecewise nature of the inputs lays the groundwork for implementing the determined IL-6 concentration profiles on an experimental system. The optimal input functions resulted in decreased parameter uncertainty for the model, as observed from simulations in which parameters were fitted using the optimal input functions for inducing the signaling system. Interestingly, the determined optimal input functions took on the shape of PRBS signals even though this was not in any way postulated by the optimal experimental design problem. This observation was further validated by formulating and solving the problem using a global method in addition to the

local method. Future work should corroborate these findings by applying the determined optimal experimental design to an actual experiment.

Author Contributions: A.S. and J.H. conceived and designed the experiments; A.S. performed the experiments and analyzed the data; A.S. and J.H. wrote the paper.

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

Appendix A

The ordinary differential equations comprising the IL-6 model from [2] are shown below, with initial values for the variables and a schematic of the signaling network taken from [2]. Parameter values from the initial fit performed in [2] and values from the average simulated fit using the optimal $r = 5$ input function (from the local method) are shown in Table A1. Units for parameter values correspond to time in seconds and concentrations in nM.

$$dx_1/dt = p_1 u^2 R^2 - p_2 x_1 - p_3 x_1 x_2 + p_4 x_3 + p_5 x_3 - p_{11} x_1 x_5 + p_{12} x_6 - p_{14} x_1 x_7 + p_{15} x_8;$$

$$dx_2/dt = -p_3 x_1 x_2 + p_4 x_3 + 2 p_6 x_4;$$

$$dx_3/dt = p_3 x_1 x_2 - p_4 x_3 - p_5 x_3;$$

$$dx_4/dt = p_5 x_3/2 - p_6 x_4;$$

$$dx_5/dt = p_7 x_4 \, \text{Step}(t - p_8)/(p_9 + x_4) - p_{10} x_5 - p_{11} x_1 x_5 + p_{12} x_6 + p_{13} x_6;$$

$$dx_6/dt = p_{11} x_1 x_5 - p_{12} x_6 - p_{13} x_6;$$

$$dx_7/dt = -p_{14} x_1 x_7 + p_{15} x_8;$$

$$dx_8/dt = p_{14} x_1 x_7 - p_{15} x_8;$$

$$dx_9/dt = p_{16} x_8 x_{10}/(p_{17} + x_{10}) - p_{18} x_9;$$

$$dx_{10}/dt = -p_{16} x_8 x_{10}/(p_{17} + x_{10}) + p_{18} x_9;$$

$$dx_{11}/dt = p_{13} x_6;$$

$$dx_{12}/dt = -2 p_{19} x_9 x_{12}^2;$$

1) $$dx_{13}/dt = p_{19} x_9 x_{12}^2;$$

2)

Name	Component	Initial value (nM)
x_1	$(\text{IL6} - \text{gp80} - \text{gp130} - \text{JAK})_2^*$	0
x_2	STAT3C	1000
x_3	$(\text{IL6} - \text{gp80} - \text{gp130} - \text{JAK})_2^* - \text{STAT3C}$	0
x_4	$\text{STAT3N}^* - \text{STAT3N}^*$	0
x_5	SOCS3	0
x_6	$(\text{IL6} - \text{gp80} - \text{gp130} - \text{JAK})_2^* - \text{SOCS3}$	0
x_7	SHP2	100
x_8	$(\text{IL6} - \text{gp80} - \text{gp130} - \text{JAK})_2^* - \text{SHP2} - \text{sum}$	0
x_9	Erk-PP	0
x_{10}	Erk	16 468
x_{11}	$(\text{IL6-gp80-gp130-JAK})_2$	0
x_{12}	C/EBPβi	40.493
x_{13}	C/EBPβn	0
u	IL-6	3.83 (i.e., 100 ng/ml)
R	Receptor	4

3)
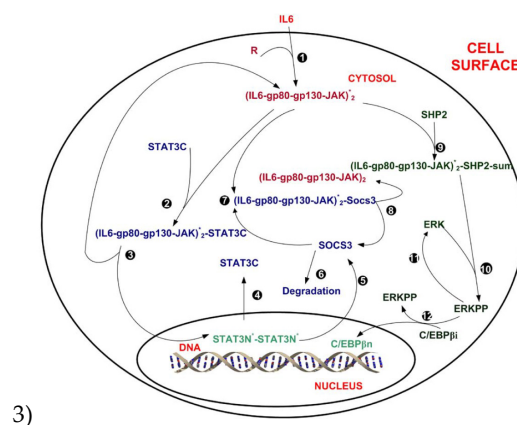
Figure A1. IL-6 model taken from [2]. (**1**) Model ODEs (**2**) initial conditions (**3**) schematic of the IL-6 signaling network.

Table A1. Parameter values from initial parameter fit [2] and from optimal $r = 5$ input function.

Parameter	Initial Fit	$r = 5$ Fit [1]
p_1	2.336×10^{-5}	2.59×10^{-5}
p_2	0.002	0.0128
p_3	0.0138	0.0148
p_4	1.502	1.508
p_5	0.273	0.232
p_6	3.282×10^{-4}	5.653×10^{-4}
p_7	0.023	0.024
p_8	1290	1219
p_9	50.6	52.3
p_{10}	2.067×10^{-4}	5.557×10^{-4}
p_{11}	16.52	16.55
p_{12}	0.04	0.06
p_{13}	0.0023	0.0023
p_{14}	4.059×10^{-4}	4.221×10^{-4}
p_{15}	5.086×10^{-4}	8.717×10^{-4}
p_{16}	16.0	15.9
p_{17}	5.115×10^3	5.085×10^3
p_{18}	1.198×10^{-5}	1.648×10^{-5}
p_{19}	1.0×10^{-6}	3.0×10^{-5}

[1] Averaged from 30 simulations utilizing the original IL-6 model.

References

1. Singh, A.; Jayaraman, A.; Hahn, J. Modeling Regulatory Mechanisms in IL-6 Signal Transduction in Hepatocytes. *Biotechnol. Bioeng.* **2006**, *95*, 850–862. [CrossRef] [PubMed]
2. Huang, Z.; Chu, Y.; Hahn, J. Model simplification procedure for signal transduction pathway model: An application to IL-6 signaling. *Chem. Eng. Sci.* **2010**, *65*, 1964–1975. [CrossRef]
3. Chu, Y.; Jayaraman, A.; Hahn, J. Parameter sensitivity analysis of IL-6 signalling pathways. *IET Syst. Biol.* **2007**, *1*, 342–352. [CrossRef] [PubMed]
4. Moya, C.; Huang, Z.; Cheng, P.; Jayaraman, A.; Hahn, J. Investigation of IL-6 and IL-10 signalling via mathematical modelling. *IET Syst. Biol.* **2011**, *5*, 15–26. [CrossRef] [PubMed]
5. Himmel, M.E.; Yao, Y.; Orban, P.C.; Steiner, T.S.; Levings, M.K. Regulatory T-cell therapy for inflammatory bowel disease: More questions than answers. *Immunology* **2012**, *136*, 115–122. [CrossRef] [PubMed]
6. Nathan, C. Points of control in inflammation. *Nature* **2002**, *420*, 846–852. [CrossRef] [PubMed]
7. Eastaff-Leung, N.; Mabarrack, N.; Barbour, A.; Cummins, A.; Barry, S. FOXP3+ regulatory T cells, Th17 effector cells, and cytokine environment in inflammatory bowel disease. *J. Clin. Immunol.* **2010**, *30*, 80–89. [CrossRef] [PubMed]
8. Silvey, S.D. *Optimal Design, An Introduction to the Theory for Parameter Estimation*; Chapman and Hall: London, UK, 1980.
9. Fedorov, V.V. *Theory of Optimal Experiments*; Academic Press: New York, NY, USA, 1972.
10. Beck, J.; Arnold, K. *Parameter Estimation in Engineering and Science*; Wiley: Hoboken, NJ, USA, 1977.
11. Atkinson, A.; Donev, A.; Tobias, R. *Optimum Experimental Designs, with SAS*; Oxford University Press: Oxford, UK, 2007.
12. Montgomery, D.C. *Design and Analysis of Experiments*, 5th ed.; Wiley: Hoboken, NJ, USA, 2000.
13. Jones, J.A.; Vernacchio, V.R.; Sinkoe, A.L.; Collins, S.M.; Ibrahim, M.H.A.; Lachance, D.M.; Hahn, J.; Koffas, M.A.G. Experimental and computational optimization of an Escherichia coli co-culture for the efficient production of flavonoids. *Metab. Eng.* **2016**, *35*, 55–63. [CrossRef] [PubMed]
14. Walter, E.; Pronzato, L. Qualitative and Quantitative Experiment Design for Phenomenological Models—A Survey. *Automatica* **1990**, *26*, 195–213. [CrossRef]
15. Wang, H.; Lee, D.; Chen, K.; Chen, J.; Zhang, K.; Silva, A.; Ho, C.; Ho, D. Mechanism-Independent Optimization of Combinatorial Nanodiamond and Unmodified Drug Delivery Using a Phenotypically Driven Platform Technology. *ACS Nano* **2015**, *9*, 3332–3344. [CrossRef] [PubMed]

16. Weiss, A.; Ding, X.; van Beijnum, J.R.; Wong, I.; Wong, T.J.; Berndsen, R.H.; Dormond, O.; Dallinga, M.; Shen, L.; Schlingemann, R.O.; et al. Rapid optimization of drug combinations for the optimal angiostatic treatment of cancer. *Angiogenesis* **2015**, *18*, 233–244. [CrossRef] [PubMed]

17. Bauer, I.; Bock, H.G.; Körkel, S.; Schlöder, J.P. Numerical methods for optimum experimental design in DAE systems. *J. Comput. Appl. Math.* **2000**, *120*, 1–25. [CrossRef]

18. Bandara, S.; Schlöder, J.P.; Eils, R.; Bock, H.G.; Meyer, T. Optimal Experimental Design for Parameter Estimation of a Cell Signaling Model. *PLoS Comput. Biol.* **2009**, *5*, e1000558. [CrossRef] [PubMed]

19. Casey, F.P.; Baird, D.; Feng, Q.; Gutenkunst, R.N.; Waterfall, J.J.; Myers, C.R.; Brown, K.S.; Cerione, R.A.; Sethna, J.P. Optimal experimental design in an epidermal growth factor receptor signalling and down-regulation model. *IET Syst. Biol.* **2007**, *1*, 190–202. [CrossRef] [PubMed]

20. Apgar, J.F.; Witmer, D.K.; White, F.M.; Tidor, B. Sloppy Models, Parameter Uncertainty, and the Role of Experimental Design. *Mol. BioSyst.* **2010**, *6*, 1890–1900. [CrossRef] [PubMed]

21. Balsa-Canto, E.; Alonso, A.A.; Banga, J.R. Computational procedures for optimal experimental design in biological systems. *IET Syst. Biol.* **2008**, *2*, 163–172. [CrossRef] [PubMed]

22. Faller, D.; Klingmüller, U.; Timmer, J. Simulation Methods for Optimal Experimental Design in Systems Biology. *Simulation* **2003**, *79*, 717–725. [CrossRef]

23. Steinmeyer, S.; Howsmon, D.P.; Alaniz, R.C.; Hahn, J.; Jayaraman, A. Empirical modeling of T cell activation predicts interplay of host cytokines and bacterial indole. *Biotechnol. Bioeng.* **2017**. [CrossRef] [PubMed]

24. Chung, M.; Haber, E. Experimental Design for Biological Systems. *SIAM J. Control Optim.* **2012**, *50*, 471–489. [CrossRef]

25. Vanlier, J.; Tiemann, C.A.; Hilbers, P.A.J.; van Riel, N.A.W. A Bayesian approach to targeted experiment design. *Bioinformatics* **2012**, *28*, 1136–1142. [CrossRef] [PubMed]

26. Weber, P.; Kramer, A.; Dingler, C.; Radde, N. Trajectory-oriented Bayesian experiment design versus Fisher A-optimal design: An in depth comparison study. *Bioinformatics* **2012**, *28*, i535–i541. [CrossRef] [PubMed]

27. Silk, D.; Kirk, P.D.W.; Barnes, C.P.; Toni, T.; Stumpf, M.P.H. Model Selection in Systems Biology Depends on Experimental Design. *PLoS Comput. Biol.* **2014**, *10*, e1003650. [CrossRef] [PubMed]

28. Flassig, R.J.; Migal, I.; van der Zalm, E.; Rihko-Struckmann, L.; Sundmacher, K. Rational selection of experimental readout and intervention sites for reducing uncertainties in computational model predictions. *BMC Bioinform.* **2015**, *16*, 13. [CrossRef] [PubMed]

29. Jost, F.; Sager, S.; Thi-Thien Le, T. A Feedback Optimal Control Algorithm with Optimal Measurement Time Points. *Processes* **2017**, *5*, 10. [CrossRef]

30. Marquardt, D.W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441. [CrossRef]

31. Mdluli, T.; Buzzard, G.T.; Rundell, A.E. Efficient Optimization of Stimuli for Model-Based Design of Experiments to Resolve Dynamical Uncertainty. *PLoS Comput. Biol.* **2015**, *11*, e1004488. [CrossRef] [PubMed]

32. Chu, Y.; Hahn, J. Parameter set selection for estimation of nonlinear dynamic systems. *AIChE J.* **2007**, *53*, 2858–2870. [CrossRef]

33. Gutenkunst, R.N.; Waterfall, J.J.; Casey, F.P.; Brown, K.S.; Myers, C.R.; Sethna, J.P. Universally Sloppy Parameter Sensitivities in Systems Biology Models. *PLoS Comput. Biol.* **2007**, *3*, e189. [CrossRef] [PubMed]

34. Huang, Z.; Senocak, F.; Jayaraman, A.; Hahn, J. Integrated Modeling and Experimental Approach for Determining Transcription Factor Profiles from Fluorescent Reporter Data. *BMC Syst. Biol.* **2008**, *2*, 64. [CrossRef] [PubMed]

35. Lagarias, J.C.; Reeds, J.A.; Wright, M.H.; Wright, P.E. Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. *SIAM J. Optim.* **1998**, *9*, 112–147. [CrossRef]

36. White, A.; Tolman, M.; Thames, H.D.; Withers, H.R.; Mason, K.A.; Transtrum, M.K. The Limitations of Model-Based Experimental Design and Parameter Estimation in Sloppy Systems. *PLoS Comput. Biol.* **2016**, *12*, e1005227. [CrossRef] [PubMed]

37. Ljung, L. *System Identification: Theory for the User*; Prentice Hall: Upper Saddle River, NJ, USA, 1999.