Dynamic Optimization with Particle Swarms (DOPS): A meta-heuristic for parameter estimation in biochemical models

Adithya Sagar, Christine Shoemaker[†] and Jeffrey D. Varner[*]

School of Chemical and Biomolecular Engineering

[†]School of Civil and Environmental Engineering

Cornell University, Ithaca NY 14853

Running Title: Parameter estimation in biochemical models

To be submitted: *PLoS ONE*

[*]Corresponding author:

Jeffrey D. Varner,

Associate Professor, School of Chemical and Biomolecular Engineering,

244 Olin Hall, Cornell University, Ithaca NY, 14853

Email: jdv27@cornell.edu

Phone: (607) 255 - 4258

Fax: (607) 255 - 9166

Abstract

Mathematical modeling is a powerful tool to analyze, and ultimately design biochemical networks. However, the estimation of biochemical model parameters is a significant challenge. Parameter estimation in biochemical models typically involves expensive function evaluations and noisy data, making it difficult to quickly obtain optimal solutions. Biochemical models often also have many local extrema which further complicates parameter estimation. Toward these challenges, we developed Dynamic Optimization with Particle Swarms (DOPS), a novel global meta-heuristic that combined features of multi-swarm particle swarm optimization with dynamically dimensioned search (DDS). DOPS uses a multi-swarm particle swarm optimization technique to generate candidate solution vectors, the best of which is greedily updated using dynamically dimensioned search. We tested the performance of DOPS on a model of human coagulation cascade. We performed $\mathcal{T}$ = 25 trials with $\mathcal{N}$ = 4000 function evaluations per trial, and compared the performance of DOPS with other commonly used meta-heuristics such as differential evolution (DE), simulated annealing (SA) and dynamically dimensioned search (DDS). We further tested the predictive power of the coagulation model parameters against data not used in training, and found good agreement between simulations and experimental measurements. Lastly, we tested the performance of DOPS on commonly used test functions for global optimization and on published biochemical parameter estimation benchmark problems. For the wide range of problems that we considered, DOPS outperformed other meta-heuristic approaches despite a limited number of function evaluations. Taken together, DOPS is a promising meta-heuristic approach for the estimation of biochemical model parameters in relatively few function evaluations.

Keywords: Parameter identification, Meta-heuristic optimization, Biochemical modeling

Introduction

Cells process nutrients and respond to changes in their environment using complex biochemical networks. These networks contain thousands of components interconnected through nonlinear enzyme catalyzed reactions. Mathematical modeling has evolved as a powerful paradigm to analyze, and ultimately design these complex networks [1–5]. Mathematical modeling of biochemical networks is often an iterative process. First, models are formulated from biochemical knowledge, and then model parameters are estimated using experimental data [6–8]. Parameter estimation is typically framed as a non-linear optimization problem wherein the residual (or objective function) between experimental data and model simulations is minimized using an optimization strategy [9]. Optimal parameters obtained from model training are then used to validate the model on unseen experimental data. If validation fails, model construction and calibration are repeated iteratively until satisfactory results are obtained.

Parameter estimation is a major challenge in the iterative development of biochemical models. Although parameter estimation has been a well studied problem in engineering for decades [10–13], the complex dynamics of large biological systems and noisy, often incomplete experimental data pose a unique estimation challenge. Often optimization problems involving biological systems are non-linear and multi-modal i.e. typical models often have multiple local minima or maxima [7, 9]. Non-linearity coupled with multi-modality generally renders local optimization techniques such as pattern search [14], Nelder-Mead simplex methods [15], steepest descent methods or Levenberg-Marquardt [16] incapable of reliably obtaining optimal solutions as they generally terminate at local minimum. Though deterministic global optimization techniques (for example algorithms based on branch and bound) can handle non-linearity and multi-modality [17, 18], the absence of derivative information, discontinuity of the objective functions, non-smooth regions or the lack of knowledge about the objective function hampers these techniques.

Meta-heuristic stochastic optimization approaches like Genetic Algorithms (GAs) [ADI-REF], Simulated Annealing (SA) [19], Evolutionary Programming [ADI-REF] and population based searches like Differential Evolution (DE) [20] have shown promise on nonlinear multi-modal problems [21]. These techniques do not make any assumptions about the structure of objective function nor do they require *a priori* information about the objective function. Though they do not guarantee strong convergence, these approaches are effective in finding near optimal solutions. Mendes et al. [22] used simulated annealing to estimate rate constants for the irreversible inhibition of HIV proteinase, Modchang et al. [23] used a genetic algorithm to estimate parameters for a model of signal transduction [ADI-EXPAND], differential evolution approaches have also been effective on various biological problems [24–26]. Tashkova et al. compared different meta-heuristics for parameter estimation on a dynamic model of endocytosis and showed that DE was the most effective [27]. Banga and co-workers have also successfully applied scatter-search methods to estimate parameters on non-linear biological processes [28–30]. Hybrid approaches that combine a meta-heuristic with a local optimization search, wherein a near globally optimal solution obtained using a meta-heuristic is refined using a local search have also become popular. Villaverde et al. combined scatter search with local search methods for parameter estimation in a collection of systems biology models [31]. Fan et al. recently showed that population based meta-heuristics along with decomposition based methods can be also used to model gene circuits from mRNA data [32]. Despite these successes, a major drawback of most metaheuristic approaches is the large number of function evaluations required to explore the parameter space. Typically as models grow in size and complexity, evaluation of the objective function becomes computationally expensive. Thus performing a large number of function evaluations is not desirable (and perhaps not feasible).

In this study, we developed Dynamic Optimization with Particle Swarms (DOPS), a

novel meta-heuristic that combines the global search capability of multi-swarm particle swarm optimization and dynamically dimensioned search (DDS). The objective of DOPS is to obtain near optimal parameter estimates for large biochemical models within a relatively few function evaluations. DOPS uses a multi-swarm particle swarm optimization technique to generate candidate solution vectors which are then greedily updated using dynamically dimensioned search. We first considered a model of human coagulation cascade to test the performance of DOPS. Coagulation is a large, complex biochemical network involving strong positive feedback. We then tested the performance of DOPS on commonly used test functions for global optimization (Ackley and Rosenbrock) and published biochemical parameter estimation benchmark problems [31]. DOPS outperformed common meta-heuristic approaches like Differential Evolution (DE), Simulated Annealing (SA) and dynamically dimensioned search (DDS) on both the test functions and the coagulation model. DOPS also performed well on the benchmark problems where it outperformed enhanced scatter search (eSS) and recovered the nominal parameters with only 4000 function evaluations ([ADI-compared with what?]) across all the benchmark problems considered. Taken together, these studies suggest DOPS is a promising meta-heuristic approach for the estimation of biochemical model parameters in relatively few function evaluations.

Results

DOPS explores parameter space using a combination of global and greedy search. DOPS is a novel meta-heuristic which combines multi-swarm particle swarm methods with the dynamically dimensioned search approach of Shoemaker and colleagues (Fig. 1). The goal of DOPS is to estimate optimal or near optimal parameter vectors for high-dimensional biological models within a specified number of function evaluations. Toward this objective, DOPS begins by using a particle swarm search and then dynamically switches, using an adaptive switching criteria, to the DDS search phase. [FINISH GENERIC DESCRIPTION OF APPROACH]

DOPS minimized benchmark problems with a fixed number of function evaluations. First, we tested DOPS and four other meta-heuristic approaches on the minimization of the Ackley and Rastrigin functions. DOPS performed similarly or outperformed the other meta-heuristic approaches on both of these test functions (Fig. 2). The Ackley and Rastrigin functions both have multiple local extrema and attain a global minimum value of zero. In each case, we fixed the maximum number of function evaluations at 4000 and ran 25 independent experiments with different initial parameter vectors. DOPS found optimal or near optimal solutions for both the 10-dimensional Ackley (Fig. 2A) and Rastrigin (Fig. 2B) functions within the budget of function evaluations. In each of the 10-dimensional cases, other meta-heurtistics such as DDS and DE also performed well, however DOPS consistently outperformed all other approaches tested. This performance difference was more pronounced as the dimension of the search problem increased; for a 300-dimensional Rastrigin function DOPS was the only approach to find an optimal or near optimal solution within the function evaluation budget (Fig. 2B). Taken together, these results suggested that DOPS could perform at least as well as other meta-heuristic approaches on small test problems, but was especially suited to large dimension search spaces. Next, we tested DOPS on benchmark biochemical models of varying complexity.

Villaverde and co-workers published a set of benchmark biochemical problems to evaluate parameter estimation methods [31]. They ranked ordered these example problems from most expensive to least expensive. We evaluated the performance of DOPS on problems from the least and most expensive categories. The least expensive problem (henceforth referred to as problem B4) was a metabolic model of Chinese Hamster Ovary (CHO) with 35 metabolites, 32 reactions and 117 parameters. The biochemical reactions were modeled using modular rate laws and the generalized form of Michaelis-Menten kinetics. The expensive problem was a genome scale kinetic model of *Saccharomyces cerevisiae* with 261 reactions, 262 variables and 1759 parameters (henceforth referred to as problem B1). In both cases, synthetic time series data sets were generated using known parameter values. For problem B1, the time series data consisted of 44 observables, and for problem B4 the data corresponded to 13 different metabolite measurement sets. We fixed the number of function evaluations at 4000 for DOPS and trained both the models against the synthetic experimental data. In both cases we found good fits (Fig. 8 and Fig. 9) to the problems within 4000 evaluations. We recaptured the nominal parameter values within 4000 evaluations (Fig. 7).

Performance of DOPS on a model of the human coagulation cascade. We compared the performance of DOPS with simulated annealing (SA), differential evolution (DE), and dynamically dimensioned search (DDS) for a model of the human coagulation cascade. Coagulation is an archetype biochemical network that is highly interconnected, containing both positive and negative feedback (Fig. 3). The biochemistry of coagulation, though complex, has been well studied [33–39], and reliable experimental protocols have been developed to interrogate the system [40–43]. Thus, the human coagulation cascade is an ideal test case; coagulation is challenging because it contains both fast and slow dynamics, but also accessible because of the availability of comprehensive data sets for model identification and validation. In this study, we used the coagulation model of Luan

et al [43]. Luan and coworkers modeled coagulation as a system of coupled non-linear ordinary differential equations. The Luan model contained 148 parameters and 92 species and has been validated using 21 published datasets.

Coagulation is mediated by a family proteases in the circulation, called factors and a key group of blood cells, called platelets. The central process in coagulation is the conversion of prothrombin (fII), an inactive coagulation factor, to the master protease thrombin (FIIa). Thrombin generation involves three phases, initiation, amplification and termination [13,14]. Initiation requires a trigger event, for example a vessel injury which exposes tissue factor (TF), which leads to the activation of factor VII (FVIIa) and the formation of the TF/FVIIa complex. Two converging pathways, the extrinsic and intrinsic cascades, then process and amplify this initial coagulation signal. There are several control points in the cascade that inhibit thrombin formation, and eventually terminate thrombin generation. Tissue Factor Pathway Inhibitor (TFPI) inhibits upstream activation events, while antithrombin III (ATIII) neutralizes several of the proteases generated during coagulation, including thrombin. Thrombin itself also inadvertently plays a role in its own inhibition; thrombin, through interaction with thrombomodulin, protein C and endothelial cell protein C receptor (EPCR), converts protein C to activated protein C (APC) which attenuates the coagulation response by proteolytic cleavage of amplification complexes. Termination occurs after either prothrombin is consumed, or thrombin formation is neutralized by inhibitors such as APC or ATIII.

To estimate the coagulation model parameters, we used data sets from TF/VIIa initiated coagulation in the absence of anticoagulants. The objective function was a linear combination of two error functions, representing coagulation initiated with different concentrations of TF-VIIa (5pM, 5nM) [40]. We restricted the number of function evaluations to $\mathcal{N}$ = 4000 for each algorithm we tested, and performed 25 trials of each experiment to collect average performance data (Table ZZ).DOPS converged faster and had a lower final

error compared to the other algorithms (Fig. 4). Within the first 1000 function evaluations of DOPS, there was a rapid drop in error. Approximately between 500-1000 function evaluations a switch to dynamically dimensioned search phase happens (this switch varied from trial to trial since the switch was based on the error from the swarm phase). Overall at the end of 4000 function evaluations DOPS minimized the error (final objective error is 5.5916e+04) to a greater extent than the other algorithms.

Using the parameters obtained at the end of 4000 function evaluations we examined the 'fits' between models predictions and experimental data (Fig.5). The solid lines represent the mean value of prediction over 25 trials and the shaded region represents the 99% confidence interval. Subsequently we used these optimal parameters to make model predictions that was compared against completely 'unseen' or untrained experimental data where coagulation was initiated with 500pM,50pM,10pM concentrations of TF-VIIa respectively (Fig.6).

Influence of the number of sub-swarms on DOPS performance. We attempted to quantify the influence of sub-swarms and DDS phase on the error convergence rates. To do so we considered only the multi swarm search without the DDS phase for 4000 function evaluations and 40 particles on the coagulation model. While doing the multi swarm search without DDS we varied the number of sub-swarms as one, two, four, five and eight. The number of sub-swarms were chosen so that each swarm contained the same number of particles. Hence we did not consider swarm numbers of three and seven. All the other parameters remained the same for each of these searches. We observed that the higher sub-swarm numbers had faster convergence rates with eight having the fastest rate. However this difference in convergences is not pronounced amongst four, five and eight. This suggests that there is an optimal number of particles per swarm beyond which there is no significant advantage.

We also observe that when these convergence rates are compared with DOPS, DOPS

shows a very rapid drop when the swarm searches begin to saturate. The green region indicates when swarm search switches to a DDS based search. As we see in (Fig.) the error convergence rates tend to remain mostly flat for the purely swarm searches with different swarm numbers. However in DOPS as soon as convergence rate begins to saturate it automatically switches to a DDS phase which leads to a more pronounced drop. Thus this automated switching strategy appears to be crucial in leading the algorithm out of local optima or saddle point regions.

Materials and Methods

Optimization problem formulation. The problem of parameter estimation in a dynamic biological model consists of finding an optimal parameter vector which minimizes the difference between model simulations and $\mathcal{E}$ experimental measurements. This difference is quantified by an objective function $K(\mathbf{p})$ which is typically the Euclidean norm of the simulation error subject to problem specific and parameter bounds constraints:

$$\begin{aligned} \text{minimize} \quad & K(\mathbf{p}) = \sum_{i=1}^{\mathcal{E}} \left(g_i(t_i, \mathbf{x}, \mathbf{p}, \mathbf{u}) - y_i\right)^2 \\ \text{subject to} \quad & \dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}(t, \mathbf{p}), \mathbf{u}(t), \mathbf{p}) \\ & \mathbf{x}(t_0) = \mathbf{x}_0 \\ & \mathbf{c}(t, \mathbf{x}, \mathbf{p}, \mathbf{u}) \geqslant \mathbf{0} \\ & \mathbf{p}^L \leqslant \mathbf{p} \leqslant \mathbf{p}^U \end{aligned} \tag{1}$$

where t is time, $\mathbf{x}(t, \mathbf{p})$ is the state variable vector with an initial state $\mathbf{x}_0$, $\mathbf{u}(t)$ is a model input vector, $\mathbf{f}$ is the system of model equations (e.g., differential equations or algebraic constraints) and $\mathbf{p}$ is the model parameter vector. The parameter search (or model simulations) can be subject to $\mathbf{c}$ linear or non-linear constraints, and parameter bound constraints where p^L and p^U denote the lower and upper parameter bounds, respectively. The problem eventually is to find:

$$\mathbf{p}^* = \arg\min_{\mathbf{p}} K(\mathbf{p}) \tag{2}$$

Dynamic optimization with particle swarms (DOPS). DOPS is a novel meta-heuristic which combines multi-swarm particle swarm methods with the dynamically dimensioned search approach of Shoemaker and colleagues (Fig. 1). The goal of DOPS is to estimate optimal or near optimal parameter vectors for high-dimensional biological models within a specified number of function evaluations. Toward this objective, DOPS begins by using a

particle swarm search and then dynamically switches, using an adaptive switching criteria, to the DDS search phase.

Phase 1: Swarm phase. We began the particle swarm phase of DOPS by randomly initializing a swarm of $\mathcal{K}$-dimensional particles (represented as z_i), wherein each particle corresponded to a $\mathcal{K}$-dimensional parameter vector. After initialization, particles were randomly partitioned into k equal sized sub-swarms $\mathcal{S}_1, \ldots, \mathcal{S}_k$. Thereafter within each sub-swarm $\mathcal{S}_k$, particles were updated according to the rule:

$$\mathbf{z}_{i,j} = \theta_{1,j-1}\mathbf{z}_{i,j-1} + \theta_2\mathbf{r}_1\left(\mathcal{L}_i - \mathbf{z}_{i,j-1}\right) + \theta_3\mathbf{r}_2\left(\mathcal{G}_k - \mathbf{z}_{i,j-1}\right) \tag{3}$$

where $(\theta_1, \theta_2, \theta_3)$ were adjustable parameters, $\mathcal{L}_i$ denotes the best solution found by particle i within sub-swarm k for function evaluation $1 \rightarrow j-1$, and $\mathcal{G}_k$ denotes the best solution found over all particles within sub-swarm $\mathcal{S}_k$. The quantities r_1 and r_2 denote uniform random vectors with the same dimension as the number of unknown model parameters ($\mathcal{K} \times 1$). Equation (3) is similar to the general particle swarm update rule, however, it does not contain velocity terms. In DOPS the parameter $\theta_{1,j-1}$ depends upon the function evaluations and is updated according to:

$$\theta_{1,j} \quad = \quad \frac{(\mathcal{N} - j) * (\mathbf{w}_{max} - \mathbf{w}_{min}))}{(\mathcal{N} - 1)} + \mathbf{w}_{min} \tag{4}$$

where $\mathcal{N}$ represents the total number of function evaluations, $\mathbf{w}_{max}$ and $\mathbf{w}_{min}$ are the maximum and minimum inertia weights, respectively. While updating the particles, we made sure all dimensions of the solution represented by the particle were within bounds using a set of reflection boundary conditions:

if $z_{i,j}^{old} < z_i^{min}$ **then**
  $z_{i,j}^{new} = z_{i,j}^{old} + (z_i^{min} - z_{i,j}^{old})$ **if** $z_{i,j}^{new} > z_i^{max}$ **then**
    $z_{i,j}^{new} = z_i^{max}$
  end
end
if $z_{i,j}^{old} > z_i^{max}$ **then**
  $z_{i,j}^{new} = z_{i,j}^{old} + (z_{i,j}^{old} - z_i^{max})$ **if** $z_{i,j}^{new} < z_i^{min}$ **then**
    $z_{i,j}^{new} = z_i^{min}$
  end
end

After every $\mathcal{M}$ function evaluations, particles were randomly redistributed to a new sub-swarm, and updated according to Eqn. (3). This process continued for $\mathcal{F}*\mathcal{N}$ functions evaluations, where $\mathcal{F}$ is the fraction of evaluations in the particle swarm phase of DOPS. If the simulation error stagnates for example, does not change by more than 1% for a specified number of evaluations, the swarm phase is terminated and DOPS switches to exploring parameter space using a DDS approach.

Phase 2: DDS phase. At the conclusion of the swarm phase, the overall best particle, $\mathcal{G}_k$, over the k sub-swarms was to initalize the DDS phase. DOPS takes $(1 - \mathcal{F}) * \mathcal{N}$ function evaluations in the DDS phase and then terminates the search. For the DDS phase, the best parameter estimate was updated using the rule:

$$\mathcal{G}_{new}(\mathbf{J}) = \begin{cases} \mathcal{G}(\mathbf{J}) + \mathbf{r}_{normal}(\mathbf{J})\sigma(\mathbf{J}), & \text{if } \mathcal{G}_{new}(\mathbf{J}) < \mathcal{G}(\mathbf{J}). \\ \mathcal{G}(\mathbf{J}), & \text{otherwise.} \end{cases} \tag{5}$$

where $\mathbf{J}$ represents the set containing the specific dimensions being perturbed, r_{normal} denotes a normal random vector of the same dimensions as $\mathcal{G}$, and σ denotes the perturbation amplitude:

$$\sigma = \mathbf{R}(\mathcal{MAX} - \mathcal{MIN}) \tag{6}$$

where $\mathbf{R}$ is the scalar perturbation size parameter, $\mathcal{MAX}$ and $\mathcal{MIN}$ are $(\mathcal{K} \times 1)$ vectors that represent the maximum and minimum bounds on each dimension. The set $\mathbf{J}$ was constructed using a probability function $\mathcal{P}_j$ that represented the probability whether a specific dimension j was perturbed or not. T his function is a monotonically decreasing function that decreases with the number of function evaluations. $\mathcal{P}_j$ can be any monotonically decreasing function, in our approach we used the following function:

$$\mathcal{P}_j \quad = \quad 1 - \log(j/((1 - \mathcal{FR}) * \mathbf{N})) \tag{7}$$

Thus the number of dimensions of the candidate vector that are updated or perturbed decreases with the as the number of function evaluations increase. These updates are greedy in nature, so $\mathcal{G}_{new}$ becomes the new solution vector only if it is better than the old one $\mathcal{G}$. The decrease of weight function in equation 4, reflection boundary conditions in equation 5, the update function described in equations 6 and 7, the selection probability in equation 8 are the means by which DDS ideas are incorporated into DOPS. The fraction of evaluations $\mathcal{FR}$ within the swarm phase is based on a switching strategy wherein the switch from swarm phase to DDS phase happens when the error due to the best solution does not drop more than 1% of the original error, continuously for more than a prescribed number of function evaluations. This allows the solution to quickly jump out of local optima and avoid any convergence issues that are generally associated with swarm based searches.

input : A randomized swarm of particles of size $NP \times K$ and fixed number of function evaluations N
output: Optimized parameter vector of size $1 \times K$

Initialize the particles randomly and assign particles randomly to various sub-swarms;
while $j \leq N$ **do**
  if *mod(j,G)=0* **then**
    Reassign particles to different sub-swarms;
  end
  for $i \leftarrow 1$ **to** NS **do**
    Update particles within sub-swarms according to equation 3;
  end
  Find best particle $\mathcal{G}$ amongst all sub-swarms;
  if $besterror(j) \geq 0.99 * besterror(j+1)$ **then**
    $failurecounter \leftarrow failurecounter + 1$;
  else
    $failurecounter \leftarrow 0$;
  end
  if $failurecounter \geq threshold$ **then**
    $\mathcal{G} \leftarrow DDS(\mathcal{G}, N - j)$;
    return $\mathcal{G}$
  else
    $j \leftarrow j + 1$;
  end
  return $\mathcal{G}$
end

Algorithm 1: Dynamic Optimization with Particle Swarms

Discussion

Our study presents a novel approach for high-dimensional parameter estimation in complex biological systems with relatively few function evaluations. In this approach we combined a variant of a well known meta heuristic particle swarm optimization with Dynamically Dimensioned Search (DDS). We tested our approach on an ODE model of coagulation with 148 parameters and 92 species. Coagulation is an ideal system to test our approach since the biology is well known and complex, with multiple feed back loops that are tightly regulated. We used experimental data under different conditions to obtain optimal parameters and used these parameters to make predictions against unseen experimental data. We obtained good fits and made sufficiently accurate enough predictions using parameters obtained from 4000 function evaluations. Further, we also used high-dimensional forms of commonly used test functions of global optimization and showed that we were able to find the global minimum for 300 dimensional Ackley and Rastrigin functions faster than other meta-heuristics. We also considered two recently published benchmark problems to test parameter optimization approaches and showed that we were able to retrieve the nominal parameter vector within 4000 function evaluations. Meta-heuristic approaches are generally effective in finding close to optimum solutions of complex, multimodal functions. In addition, they generally obviate the need for any *a priori* knowledge (like function derivative). However they take an exorbitant number of objective function evaluations to come close to an optimum. When the objective function evaluations tend to become expensive it is infeasible to take up a large number of evaluations. As the dimensionality of parameter space increases, the search region gets widened and thus the problem becomes more challenging. In addition, most of these approaches require optimization of 'algorithm parameters' before the actual optimization and also involve computationally expensive update operations. Tolson and Shoemaker, through DDS, showed that randomly perturbing a subset of dimensions in high dimensional parameter space is

an effective way to obtain near optimal solutions with few function evaluations. Though their approach is based on a single solution, the decision vector carries forward from one iteration to the next. Hence it behaves similar to a population based search. In our approach, we tried to eliminate the probability of starting from a bad region of search using a variant of particle swarm optimization. Particle Swarm Optimization (PSO) is a population based meta heuristic which does not have any complex operations like recombination, mutation or selection that are associated with other population based meta-heuristics like Differential Evolution (DE) or Genetic Algorithm (GA). Several particle swarm variants have been proposed to improve the search ability and rate of convergence, that involve different neighborhood structures, multi-swarms or adaptive parameters. Multi-swarm PSO with small particle neighborhoods have been shown to better in searching on complex multi-modal solutions [44]. Multi swarm methods, in addition, avoid rapid convergence to a local optimum or stable point and are able to generate diverse solutions. Generation of diverse solutions in the early stage gives a better exploratory capability and thus converge of upon multiple optima.

We utilized this capacity of multi swarms to generate diverse candidate solutions which can be used as initial solutions for a DDS like search. Thus the solution vectors obtained at the end of multi swarm search phase have a better propensity to not get stuck in bad regions in DDS phase as opposed to starting with a totally random initial solution. Given a fixed number of function evaluations, we have been able to show that we were able to obtain better solutions for the coagulation model and the test functions faster that other commonly known meta-heuristics and also DDS alone. The computational cost was prohibitively high for parameter estimation on this problem using a standard PSO. Thus we used SA, DE and DDS for the purpose of comparison. Choosing the number of function evaluations is largely a function of cost and complexity of the objective function. Traditionally the stopping condition for a parameter optimization problem can be the number

of function evaluations, percentage of initial error achieved or an absolute error threshold. However in case of complex, expensive functions where we desire a value within a certain period of time, the number of evaluations are used as a stopping criterion. In our current study we used a value of 4000 which we based upon the time taken (approximately 8-10 seconds) for a single objective function evaluation in the coagulation case. We used the same value of 4000 for benchmarks published by Villaverde and co-workers [31]. In this work by Villaverde et al. they used a population based search enhanced Scatter Search (eSS) to estimate the biochemical model parameters. Quite surprisingly we took a couple of orders lesser number of function evaluations 10^3 to obtain the optimal parameter vector as compared to the enhanced Scatter Search (eSS) with a local optimizer, which took around the order of 10^5 number of evaluations. The amount of CPU time taken (on an Intel Xeon processor 2.4 GHz) is lesser as compared to eSS on a similar architecture.

A surprisingly remarkable aspect about our algorithm has been that we did not 'pre optimize' any parameters of the algorithm to suit a specific problem. In the swarm phase this includes the number of particles, number of sub-swarms, acceleration constants or the number of generations after which the particles are redistributed and the neighborhood perturbation parameter in DDS phase. We used the same parameters for all the problems. The same rule was applied to the rest of the meta-heuristics barring Simulated Annealing. For SA, we optimized the cooling schedule for the coagulation model. Thus, in this approach any overhead that usually comes with additional function evaluations to pre optimize parameters was avoided.

The performance of our approach seems impressive given that it performed well on different complex systems with no pre optimization of algorithm parameters being required. We comfortably outperformed existing, widely used meta-heuristics and were also able to find minima of high dimensional global optimization test functions. Thus this approach may be well suited to large scale global optimization. In addition, surprisingly, we were

able to obtain optimal parameter vectors for two different large scale systems biology models with a couple of orders fewer number of function evaluations as compared to enhanced Scatter Search (eSS). However it is quite possible that highly optimized versions of common meta heuristics may outperform us on these systems. This aspect is currently beyond the scope of this study. Our approach can also be combined with local derivative based searches to improve upon the accuracy of the solutions. In addition, the current implementation of the algorithm is designed to switch only once from the swarm phase to the DDS. In the DDS phase, the search uses only one candidate vector although there is a provision to start with a population of candidate vectors. Incorporating a more intelligent switching strategy that can do switch from swarm phase to DDS phase multiple times and having a population of candidate vectors are some aspects of the algorithm that can be studied further.

Acknowledgements

This study was supported by an award from the Army Research Office (ARO #59155-LS).

References

1. Assmus HE, Herwig R, Cho KH, Wolkenhauer O (2006) Dynamics of biological systems: role of systems biology in medical research. Expert Review of Molecular Diagnostics .

2. van Riel NAW (2006) Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. Briefings in Bioinformatics 7: 364–374.

   KEY: Riel:2006aa

   ANNOTATION: 10.1093/bib/bbl040

3. Jaqaman K, Danuser G (2006) Linking data to models: data regression. Nat Rev Mol Cell Biol 7: 813–819.

   KEY: Jaqaman:2006aa

   ANNOTATION: 10.1038/nrm2030

4. Kitano H (2002) Systems biology: a brief overview. Science 295: 1662–1664.

5. Hood L, Heath JR, Phelps ME, Lin B (2004) Systems biology and new technologies enable predictive and preventative medicine. Science 306: 640–643.

6. Aldridge BB, Burke JM, Lauffenburger DA, Sorger PK (2006) Physicochemical modelling of cell signalling pathways. Nat Cell Biol 8: 1195–1203.

   KEY: Aldridge:2006aa

   ANNOTATION: 10.1038/ncb1497

7. Banga JR (2008) Optimization in computational systems biology. BMC systems biology 2: 47.

8. Ashyraliyev M, Fomekong-Nanfack Y, Kaandorp JA, Blom JG (2009) Systems biology: parameter estimation for biochemical models. Febs Journal 276: 886–902.

9. Moles CG, Mendes P, Banga JR (2003) Parameter estimation in biochemical pathways: a comparison of global optimization methods. Genome research 13: 2467–2474.
10. Nieman R, Fisher D, Seborg D (1971) A review of process identification and parameter estimation techniques†. International Journal of Control 13: 209–264.
11. Beck JV, Arnold KJ (1977) Parameter estimation in engineering and science. James Beck.
12. Young P (1981) Parameter estimation for continuous-time models—a survey. Automatica 17: 23–39.
13. Beck JV, Woodbury KA (1998) Inverse problems and parameter estimation: integration of measurements and analysis. Measurement Science and Technology 9: 839.
14. Hooke R, Jeeves TA (1961) "direct search"solution of numerical and statistical problems. Journal of the ACM (JACM) 8: 212–229.
15. Nelder JA, Mead R (1965) A simplex method for function minimization. The computer journal 7: 308–313.
16. Moré JJ (1978) The levenberg-marquardt algorithm: implementation and theory. In: Numerical analysis, Springer. pp. 105–116.
17. Esposito WR, Floudas CA (2000) Deterministic global optimization in nonlinear optimal control problems. Journal of Global Optimization 17: 97–126.
18. Horst R, Tuy H (2013) Global optimization: Deterministic approaches. Springer Science & Business Media.
19. Kirkpatrick S, Gelatt CD, Vecchi MP, et al. (1983) Optimization by simulated annealing. science 220: 671–680.
20. Storn R, Price K (1997) Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. Journal of global optimization 11: 341–359.
21. Sun J, Garibaldi JM, Hodgman C (2012) Parameter estimation using metaheuristics in

systems biology: a comprehensive review. Computational Biology and Bioinformatics, IEEE/ACM Transactions on 9: 185–202.

22. Mendes P, Kell D (1998) Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. Bioinformatics 14: 869–883.
23. Modchang C, Triampo W, Lenbury Y (2008) Mathematical modeling and application of genetic algorithm to parameter estimation in signal transduction: Trafficking and promiscuous coupling of g-protein coupled receptors. Computers in Biology and Medicine 38: 574–582.
24. Tsai KY, Wang FS (2005) Evolutionary optimization with data collocation for reverse engineering of biological networks. Bioinformatics 21: 1180–1188.
25. Wang FS, Su TL, Jang HJ (2001) Hybrid differential evolution for problems of kinetic parameter estimation and dynamic optimization of an ethanol fermentation process. Industrial & engineering chemistry research 40: 2876–2885.
26. Noman N, Iba H (2007) Inferring gene regulatory networks using differential evolution with local search heuristics. IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB) 4: 634–647.
27. Tashkova K, Korošec P, Šilc J, Todorovski L, Džeroski S (2011) Parameter estimation with bio-inspired meta-heuristic optimization: modeling the dynamics of endocytosis. BMC systems biology 5: 159.
28. Villaverde AF, Egea JA, Banga JR (2012) A cooperative strategy for parameter estimation in large scale systems biology models. BMC systems biology 6: 75.
29. Rodriguez-Fernandez M, Egea JA, Banga JR (2006) Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. BMC bioinformatics 7: 483.
30. Egea JA, Rodríguez-Fernández M, Banga JR, Martí R (2007) Scatter search for chemical and bio-process optimization. Journal of Global Optimization 37: 481–503.

31. Villaverde AF, Henriques D, Smallbone K, Bongard S, Schmid J, et al. (2015) Biopredyn-bench: a suite of benchmark problems for dynamic modelling in systems biology. BMC systems biology 9: 8.
32. Fan M, Kuwahara H, Wang X, Wang S, Gao X (2015) Parameter estimation methods for gene circuit modeling from time-series mrna data: a comparative study. Briefings in bioinformatics : bbv015.
33. Mann KG, Butenas S, Brummel K (2003) The dynamics of thrombin formation. Arteriosclerosis, thrombosis, and vascular biology 23: 17–25.
34. Mann K, Brummel K, Butenas S (2003) What is all that thrombin for? Journal of Thrombosis and Haemostasis 1: 1504–1514.
35. Mann KG (2003) Thrombin formation. CHEST Journal 124: 4S–10S.
36. Vogler EA, Siedlecki CA (2009) Contact activation of blood-plasma coagulation. Biomaterials 30: 1857–1869.
37. Diamond SL (2013) Systems biology of coagulation. Journal of Thrombosis and Haemostasis 11: 224–232.
38. Fogelson AL, Tania N (2005) Coagulation under flow: the influence of flow-mediated transport on the initiation and inhibition of coagulation. Pathophysiology of haemostasis and thrombosis 34: 91–108.
39. Anand M, Rajagopal K, Rajagopal K (2003) A model incorporating some of the mechanical and biochemical factors underlying clot formation and dissolution in flowing blood: review article. Journal of Theoretical Medicine 5: 183–218.
40. Hockin MF, Jones KC, Everse SJ, Mann KG (2002) A model for the stoichiometric regulation of blood coagulation. Journal of Biological Chemistry 277: 18322–18333.
41. Chatterjee MS, Denney WS, Jing H, Diamond SL (2010) Systems biology of coagulation initiation: kinetics of thrombin generation in resting and activated human blood. PLoS computational biology .

442 42. Mann KG, Brummel-Ziedins K, Orfeo T, Butenas S (2006) Models of blood coagula-
443 tion. Blood Cells, Molecules, and Diseases 36: 108–117.
444 43. Luan D, Zai M, Varner JD (2007) Computationally derived points of fragility of a hu-
445 man cascade are consistent with current therapeutic strategies. PLoS computational
446 biology 3: e142.
447 44. Zhao SZ, Liang JJ, Suganthan PN, Tasgetiren MF (2008) Dynamic multi-swarm par-
448 ticle swarm optimizer with local search for large scale global optimization. In: Evo-
449 lutionary Computation, 2008. CEC 2008.(IEEE World Congress on Computational
450 Intelligence). IEEE Congress on. IEEE, pp. 3845–3852.

Table 1: Error Analysis.

TF/FVIIa concentration	Normalized S.E.
5 nM	0.0376
500 pM	0.0564
50 pM	0.1125
10 pM	0.0823
5 pM	0.0338

Table 2: Table with optimization settings and results for the coagulation problem, the benchmarks and test functions using DOPS. For each problem the bounds on the parameter vector, the total number of function evaluations, the best initial objective value and the best final objective value are specified. Nominal objective value represents the objective value using the true parameter vector or the nominal parameter vector. The CPU time is the time taken for the problem on a 2.4GHz Intel Xeon Architecture running Matlab 2014b.

	Coagulation	B1	B4	Ackley	Rastrigin
Evaluations	4000	4000	4000	4000	4000
Lower Bound	0.001.pnom	5.pnom	5.pnom	30	5.12
Upper Bound	1000.pnom	0.2.pnom	0.2.pnom	-15	-5.12
CPU Time	10.1 hrs	38.3 hrs	6.2 min	2.8 s	2.6 s
Scaled initial error	1.0	1.0	1.0	1.0	1.0
Scaled final error	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Scaled nominal error	0.42	0.1	< 0.01	0	0

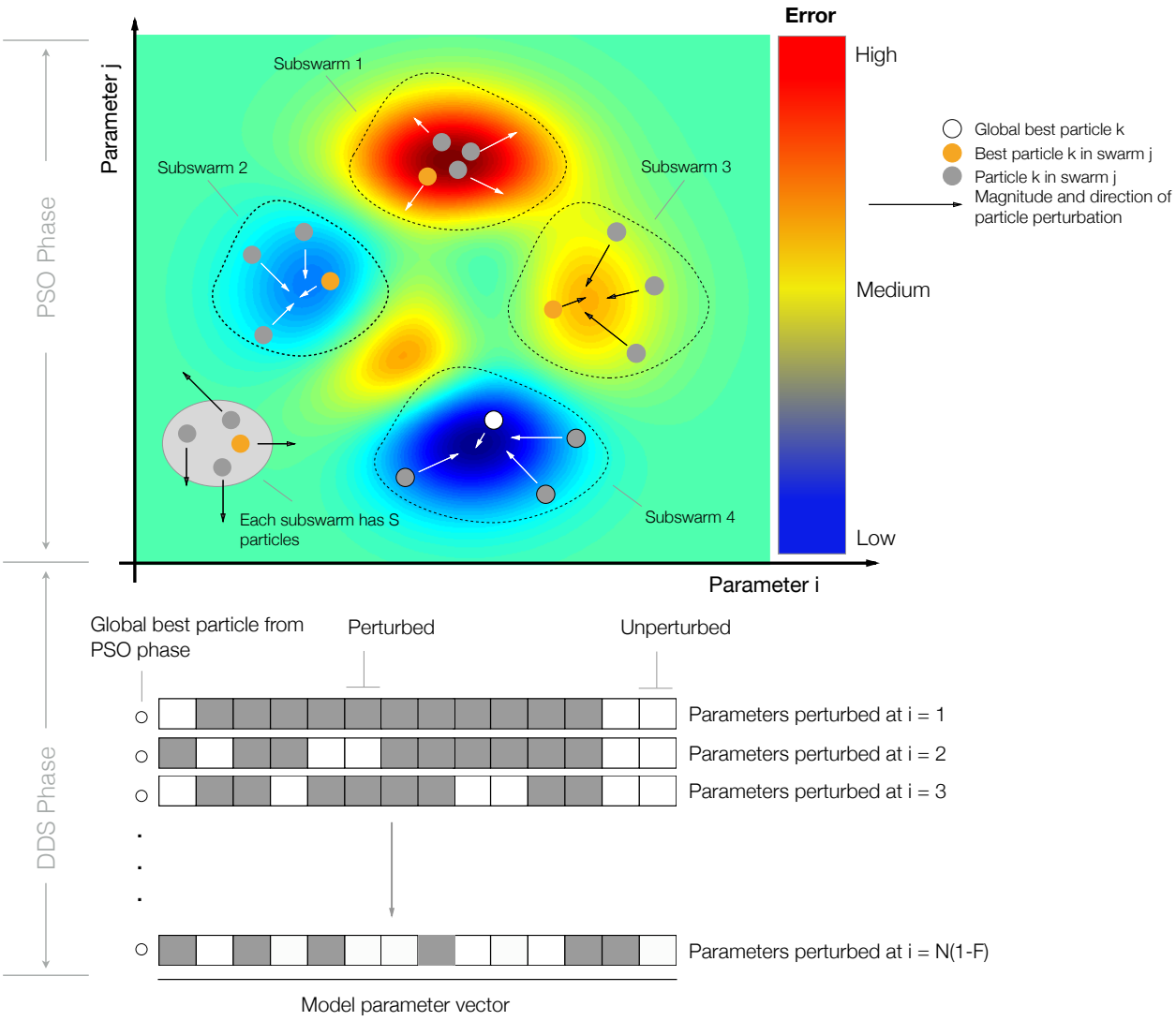

Fig. 1: Dynamic Optimization with Particle Swarms. **A**: Each particle represents an N dimensional parameter vector. Particles are given randomly generated initial solutions and grouped into different sub-swarms. Within each swarm the magnitude and direction of the movement a particle is influenced by the position of the best particle and also by its own experience. After every g number of function evaluations the particles are mixed and randomly assigned to different swarms. When the error due to the global best particle (best particle amongst all the sub-swarms) does not drop over a certain number of function evaluations, the swarm search is stopped and the search switches to a Dynamically Dimensioned Search with global best particle as the initial solution vector or candidate vector. **B**: The candidate vector performs a greedy global search for the remaining number of function evaluations. The search neighborhood is dynamically adjusted by varying the number of dimensions that are perturbed (in black) in each evaluation step. The probability that a dimension is perturbed decreases as the number of function evaluations increase. Thus as the evaluations increase the optimality of the solution is preserved.

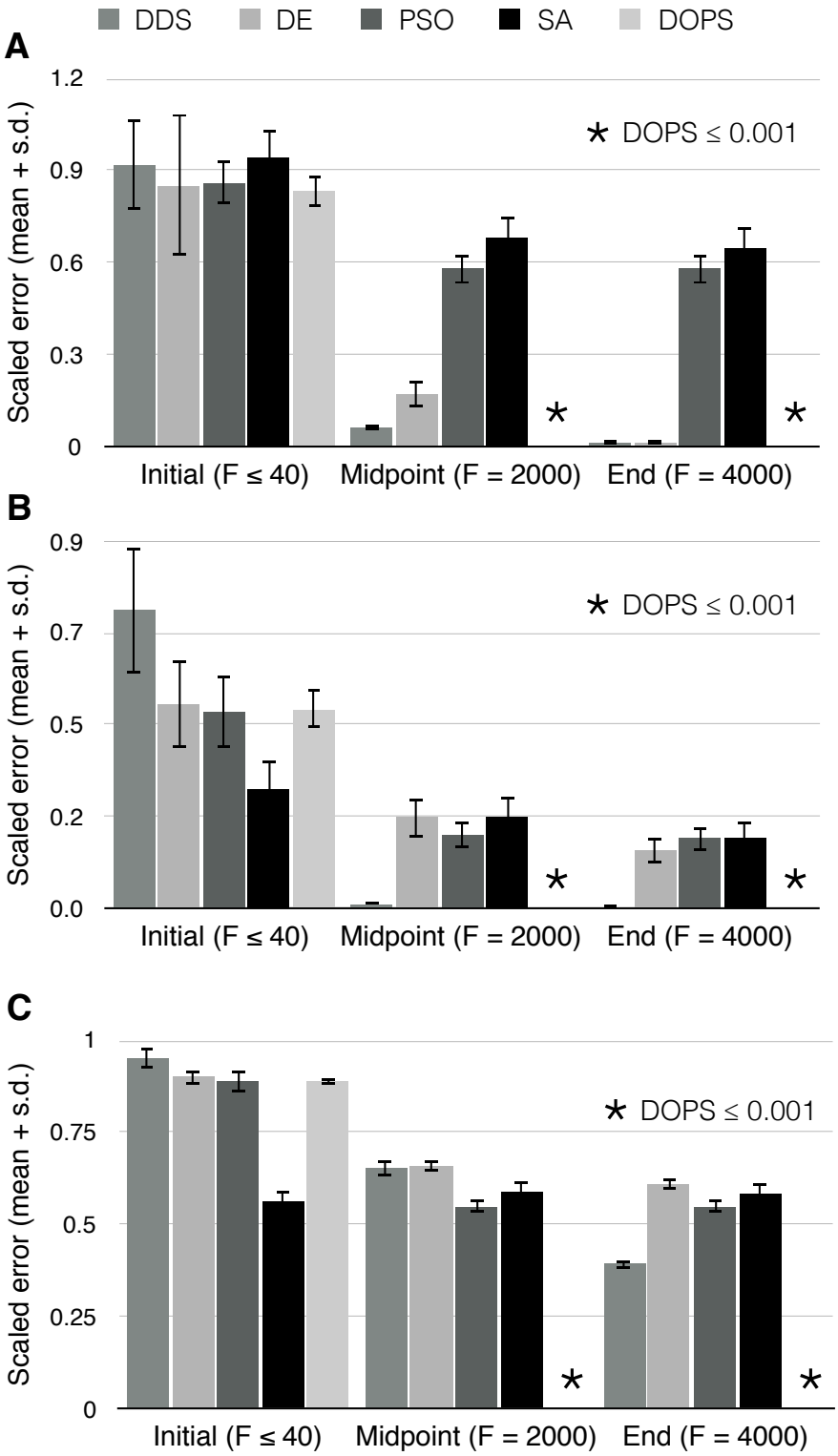

Fig. 2: Performance of DOPS and other meta-heuristics for the Ackley and Rastrigin functions. **A**: Mean scaled error versus the number of function evaluations for the 10-dimensional Ackley function. DOPS, DDS and DE find optimal or near optimal solutions within the specified number of function evaluations. **B**: Mean scaled error versus the number of function evaluations for the 10-dimensional Rastrigin function. DOPS and DDS find optimal or near optimal solutions within the specified number of function evaluations. **C**: Mean scaled error versus the number of function evaluations for the 300-dimensional Rastrigin function. DOPS is the only algorithm that finds an optimal or near optimal solution within the specified number of function evaluations. In all cases, the maximum number of function evaluations was 4000. Mean and standard deviation were calculated over 25 trials.

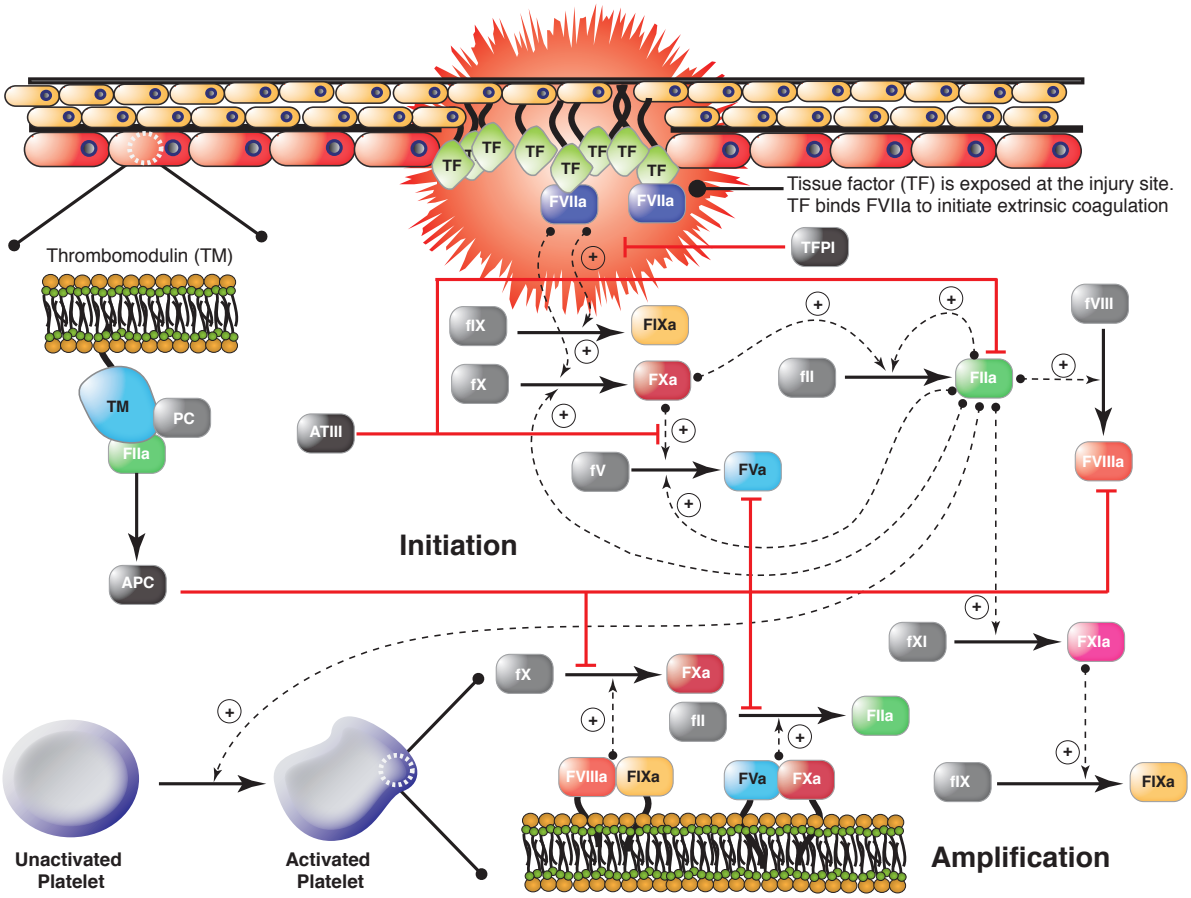

Fig. 3: Schematic of the extrinsic and intrinsic coagulation cascade[43]. Inactive zymogens upstream (grey) are activated by exposure to tissue factor (TF) following vessel injury. Tissue factor and activated factor VIIa (FVIIa) form a complex that activates factor X (fX) and IX (fIX). FXa activates downstream factors including factor VIII (fVIII) and fIX. Factor V (fV) is primarily activated by thrombin (FIIa). In addition, we included a secondary fV activation route involving FXa. FXa and FVa form a complex (prothrombinase) on activated platelets that converts prothrombin (fII) to FIIa. FIXa and FVIIIa can also form a complex (tenase) on activated platelets which catalyzes FXa formation. Thrombin also activates upstream coagulation factors, forming a strong positive feedback ensuring rapid activation. Tissue factor pathway inhibitor (TFPI) downregulates FXa formation and activity by sequestering free FXa and TF-FVIIa in a FXa-dependent manner. Antithrombin III (ATIII) inhibits all proteases. Thrombin inhibits itself binding the surface protein thrombomodulin (TM). The IIa-TM complex catalyzes the conversion of protein C (PC) to activated protein C (APC), which attenuates the coagulation response by the proteolytic cleavage of fV/FVa and fVIII/FVIIIa.

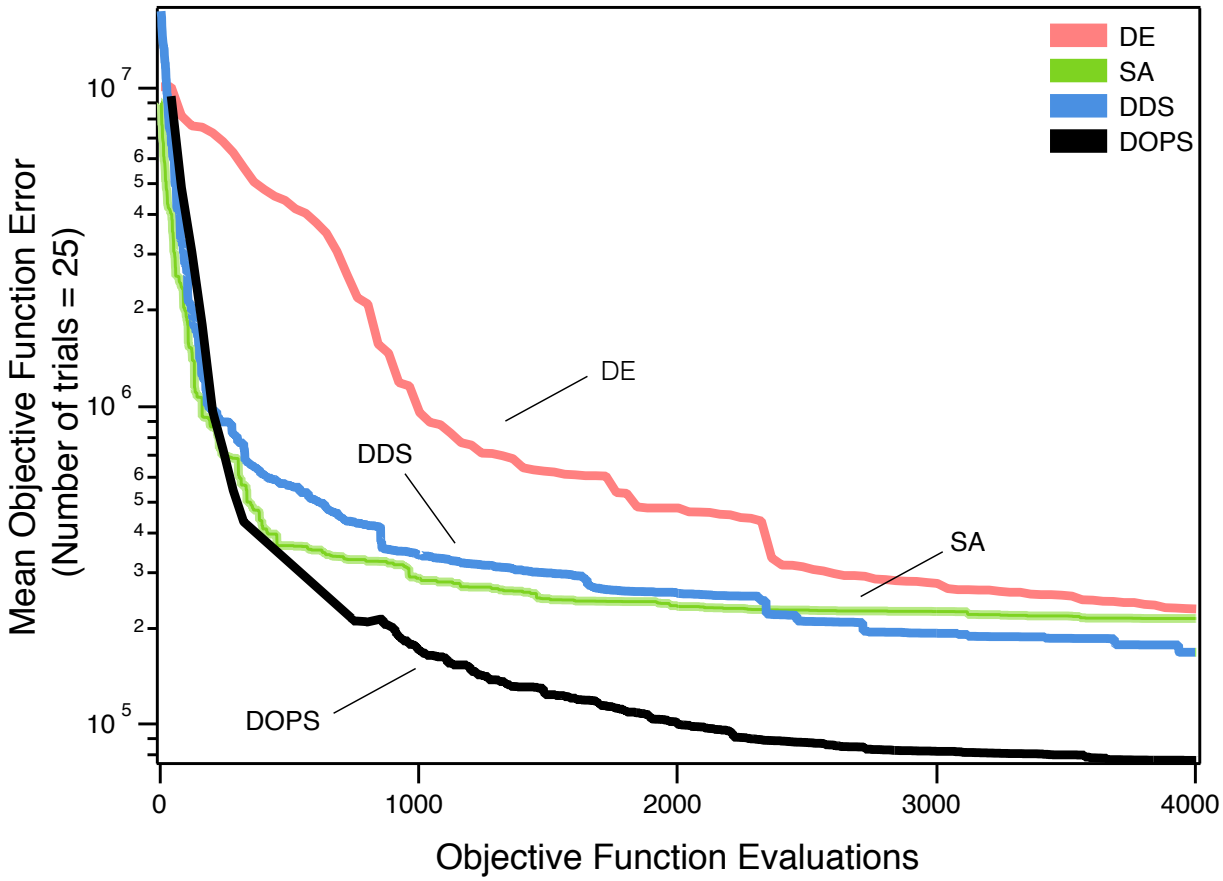

Fig. 4: Error convergence rates of the five different algorithms on the coagulation model. The objective error is the mean over N= 25 trials. DOPS, DDS and SA have the steepest drop in error during first 300 function evaluations. Thereafter the error drop in DDS and SA remains nearly constant whereas DOPS continues to drops further. At the end of 4000 function evaluations DOPS attains the lowest error. The next best estimate using DDS is nearly 3 times greater than the lowest error using DDS.

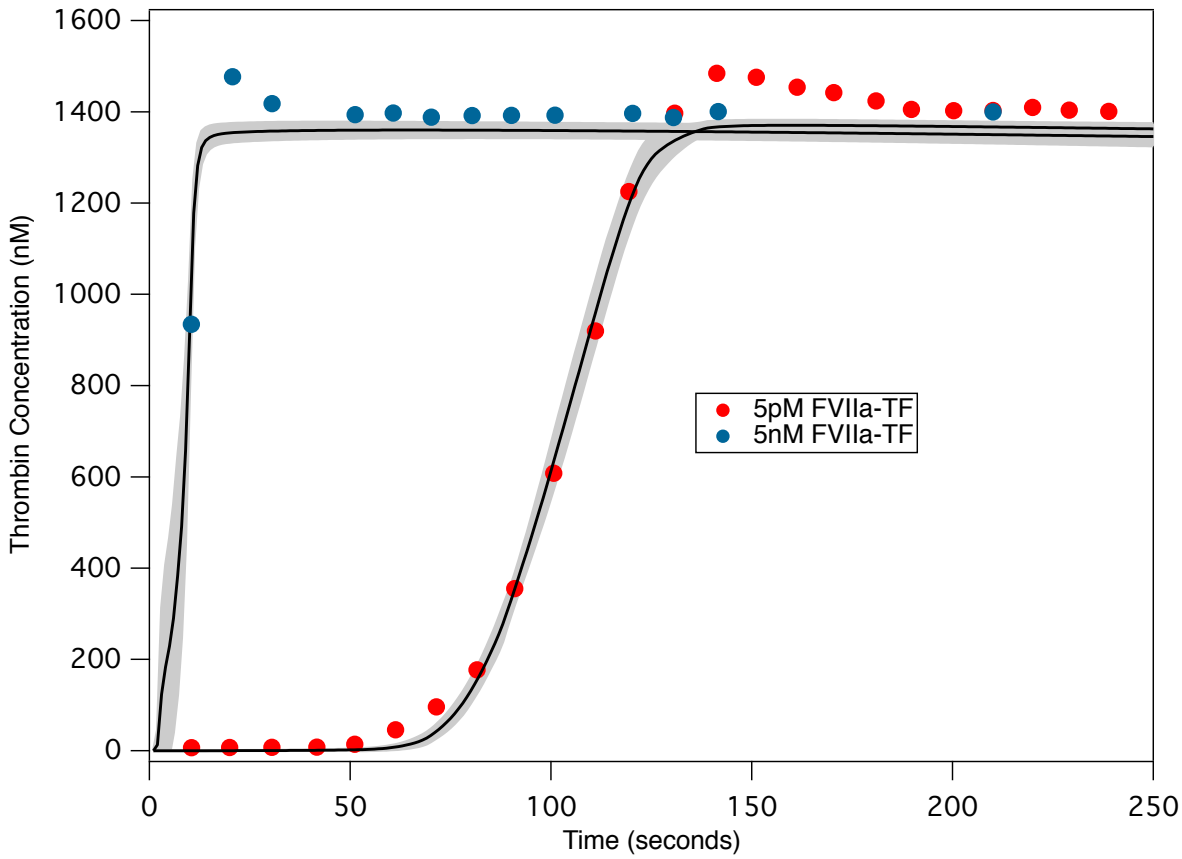

Fig. 5: Model fits on experimental data using DOPS. The model parameters were estimated using DOPS. Solid black lines indicate the simulated mean thrombin concentration using parameter vectors from N= 25 trials. The grey shaded region represents the 99% confidence estimate of the mean simulated thrombin concentration. The experimental data is reproduced from the synthetic plasma assays of Mann and co-workers. Thrombin generation is initiated by adding Factor VIIa-TF (5nM - Red and 5pM - Green) to synthetic plasma containing 200 μmol/L of phospholipid vesicles (PCPS) and a mixture of coagulation factors (II,V,VII,VIII,IX,X and XI) at their mean plasma concentrations.

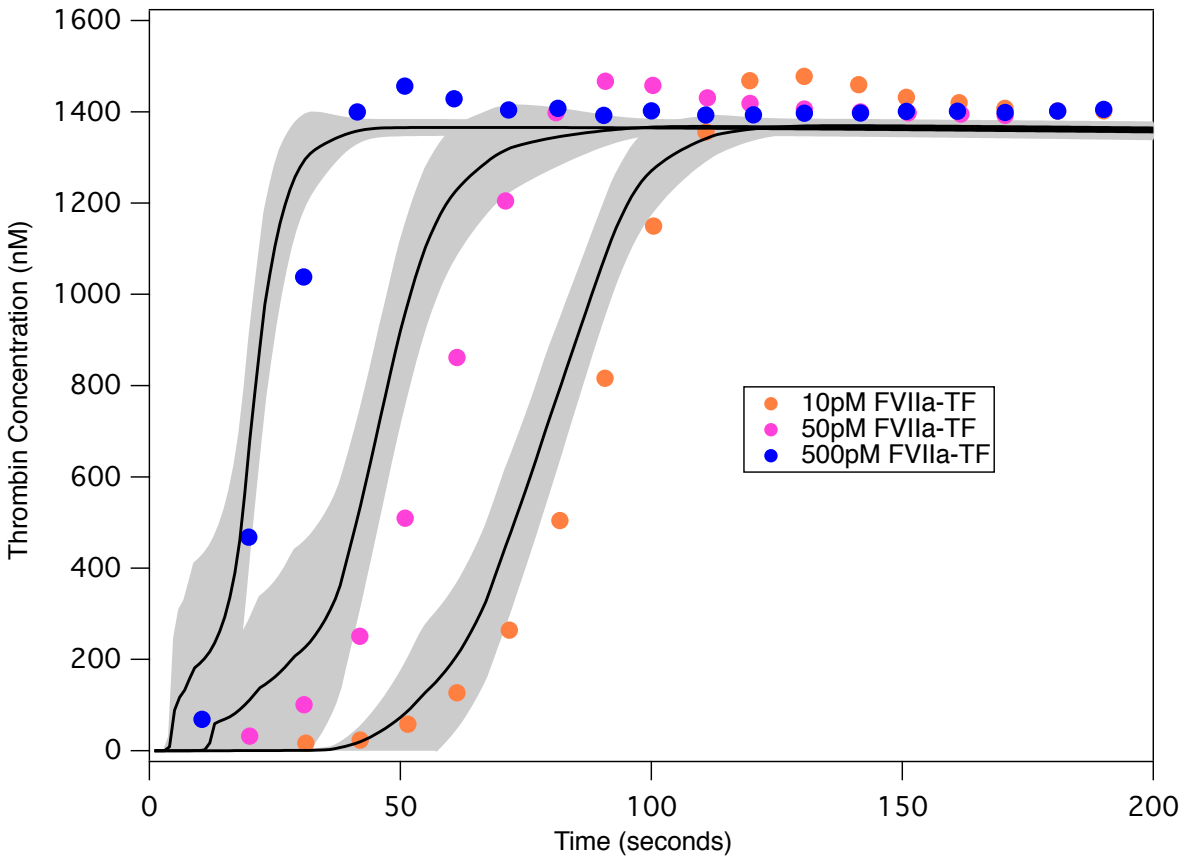

Fig. 6: Model predictions on unseen experimental data using parameters obtained from DOPS. The parameter estimates that were obtained using DOPS were tested against data that was not used in the model training. Solid black lines indicate the simulated mean thrombin concentration using parameter vectors from N= 25 trials. The grey shaded region represents the 99% confidence estimate of the mean simulated thrombin concentration. The experimental data is reproduced from the synthetic plasma assays of Mann and co-workers. Thrombin generation is initiated by adding Factor VIIa-TF (500pM - Blue, 50pM - Pink and 10pM - Orange respectively) to synthetic plasma containing 200 μmol/L of phospholipid vesicles (PCPS) and a mixture of coagulation factors (II,V,VII,VIII,IX,X and XI) at their mean plasma concentrations.

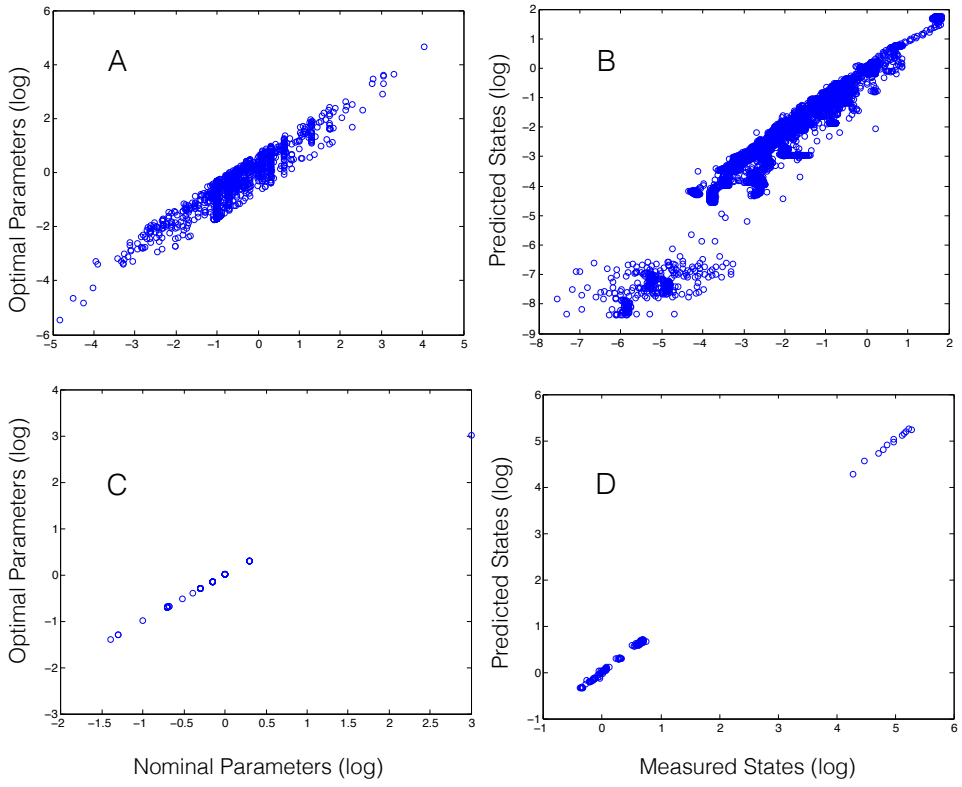

Fig. 7: (A) Difference between nominal and optimal parameters for problem B1: Genome wide kinetic model of *S.cerevisiae* with 1759 unknown parameters. **(B)** Difference between experimental (measured) data and data simulated with optimal parameters for problem B1: Genome wide kinetic model of *S.cerevisiae* with 1759 unknown parameters. **(C)** Difference between nominal and optimal parameters for problem B4: Metabolic model of Chinese Hamster Ovary Cells (CHO) cells with 117 parameters. **(D)** Difference between experimental (measured) data and data simulated with optimal parameters for problem B4: Metabolic model of Chinese Hamster Ovary Cells (CHO) cells with 117 parameters.

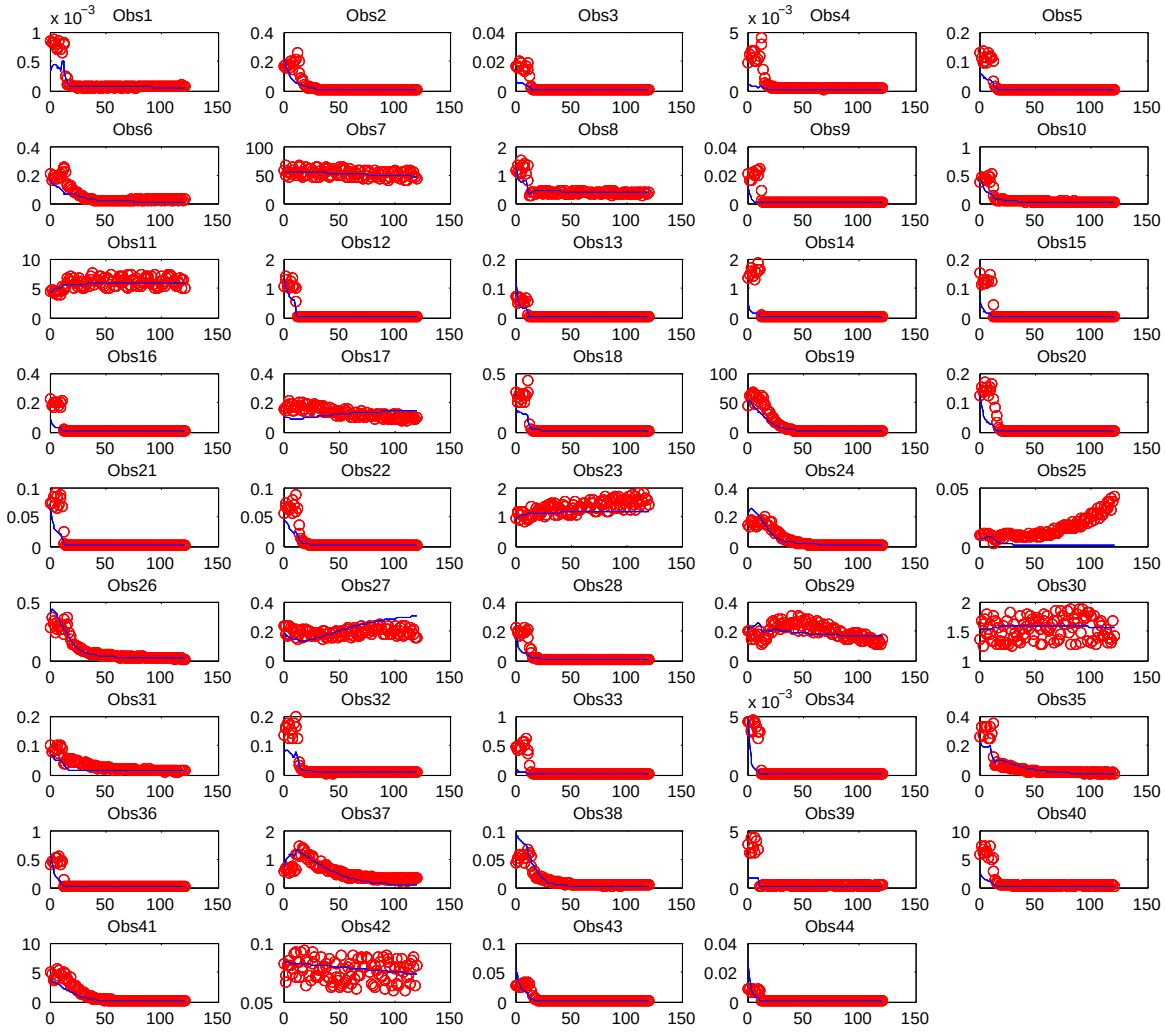

Fig. 8: (Data fits for Problem B1) Pseudo-experimental data (red circles) vs. optimal solution obtained using DOPS (solid blue lines) for the 44 observed states. X axis: time [s]; Y axis: metabolite concentrations [mM].

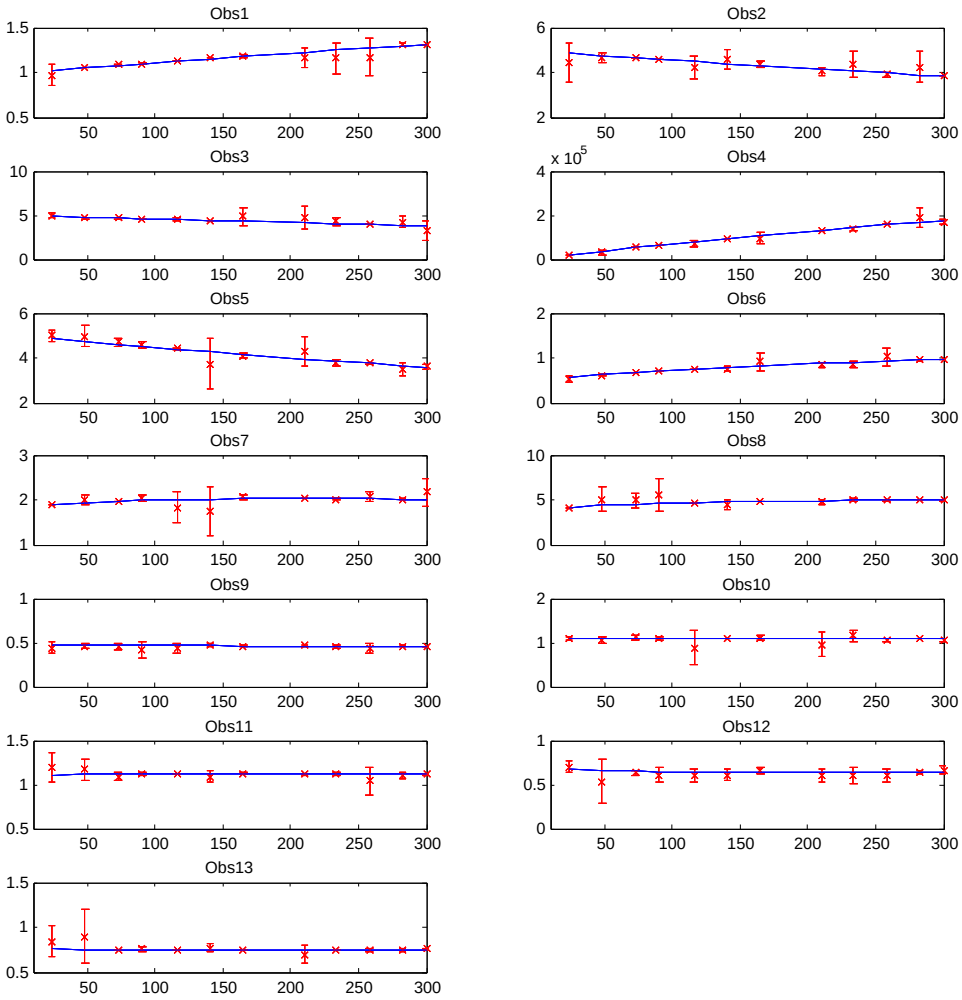

Fig. 9: (Data fits for Problem B4) Pseudo-experimental data (red x) vs. optimal solution obtained using DOPS (solid blue lines) for the 13 observed states. X axis: time [s]; Y axis: metabolite concentrations [mM].