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: *Biotechnology Journal*

[*]Corresponding author:

Jeffrey D. Varner,

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 design biochemical networks. However, the estimation of biochemical model parameters is a significant challenge. Parameter estimation 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 hybrid meta-heuristic that combined 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 DOPS using a combination of classic optimization test functions, biochemical benchmark problems and real-world biochemical models. 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). On average, DOPS outperformed other common meta-heuristics on the optimization test functions, benchmark problems and a model of the human coagulation cascade. Taken together, DOPS is a promising meta-heuristic approach for the estimation of biochemical model parameters in relatively few function evaluations. DOPS source code is available for download under a MIT license at http://www.varnerlab.org.

Keywords: Biochemical engineering, systems biology, modeling, optimization

Introduction

Cells process nutrients and respond to changes in their environment using complex enzyme catalyzed biochemical networks. 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 existing 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 measurements and model simulations is minimized using an optimization strategy [9]. Optimal parameter estimates are then used to predict unseen experimental data. If the validation studies fail, model construction and calibration are repeated iteratively until satisfactory results are obtained. As our biological knowledge increases, model formulation may not be as significant a challenge, but parameter estimation will likely remain difficult.

Parameter estimation is a major challenge to the development of biochemical models. Parameter estimation has been a well studied engineering problem for decades [10–13]. However, the complex dynamics of large biological systems and noisy, often incomplete experimental data sets pose a unique estimation challenge. Often optimization problems involving biological systems are non-linear and multi-modal i.e., typical models have multiple local minima or maxima [7, 9]. Non-linearity coupled with multi-modality renders local optimization techniques such as pattern search [14], Nelder-Mead simplex methods [15], steepest descent or Levenberg-Marquardt [16] incapable of reliably obtaining globally optimal solutions as these methods often 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, discontinuous objective functions, non-smooth regions or the lack of knowledge about the objective function hampers these techniques.

1

Meta-heuristics like Genetic Algorithms (GAs) [19], Simulated Annealing (SA) [20], Evolutionary Programming [21] and Differential Evolution (DE) [22–25] have all shown promise on non-linear multi-modal problems [26]. These techniques do not make any assumptions, nor do they require, *a priori* information about the structure of the objective function. Meta-heuristics are often very effective at finding globally optimal or near optimal solutions. For example, Mendes et al. used SA to estimate rate constants for the inhibition of HIV proteinase [27], while Modchang et al. used a GA to estimate parameters for a model of G-protein-coupled receptor (GPCR) activity [28]. Parameter estimates obtained using the GA stratified the effectiveness of two G-protein agonists, N6-cyclopentyladenosine (CPA) and 5'-N-ethylcarboxamidoadenosine (NECA). Tashkova et al. compared different meta-heuristics for parameter estimation on a dynamic model of endocytosis; DE was the most effective of the approaches tested [29]. Banga and co-workers have also successfully applied scatter-search to estimate model parameters [30–32]. Hybrid approaches, which combine meta-heuristics with local optimization techniques, have also become popular. For example, Villaverde et al. developed the enhanced scatter search (eSS) method [33], which combined scatter and local search methods, for parameter estimation in biological models [34]. However, despite these successes, a major drawback of most meta-heuristics remains the large number of function evaluations required to explore parameter space. Performing numerous potentially expensive function evaluations is not desirable (and perhaps not feasible) for many types of biochemical models. Alternatively, Tolson and Shoemaker found, using high-dimensional watershed models, that perturbing only a subset of parameters was an effective strategy for estimating parameters in expensive models [35]. Their approach, called Dynamically Dimensioned Search (DDS), is a simple stochastic single-solution heuristic that estimates nearly optimal solutions within a specified maximum number of function (or model) evaluations. Thus, while meta-heuristics are often effective at estimating globally optimal or

nearly optimal solutions, they require a large number of function evaluations to converge to a solution.

In this study, we developed Dynamic Optimization with Particle Swarms (DOPS), a novel hybrid meta-heuristic that combines the global search capability of multi-swarm particle swarm optimization with the greedy refinement of 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 multi-swarm particle swarm optimization to generate nearly optimal candidate solutions, which are then greedily updated using dynamically dimensioned search. We tested DOPS using a combination of classic optimization test functions, biochemical benchmark problems and real-world biochemical models. First, we tested the performance of DOPS on the Ackley and Rosenbrock functions, and published biochemical benchmark problems. Next, we used DOPS to estimate the parameters of a model of the human coagulation cascade. On average, DOPS outperformed other common meta-heuristics like differential evolution, simulated annealing, single-swarm particle swarm optimization, and dynamically dimensioned search on the optimization test functions, benchmark problems and the coagulation model. For example, DOPS recovered the nominal parameters for the benchmark problems using an order of magnitude fewer function evaluations than eSS in all cases. It also produced parameter estimates for the coagulation model that predicted unseen coagulation data sets. Thus, DOPS is a promising hybrid meta-heuristic for the estimation of biochemical model parameters in relatively few function evaluations.

Results

DOPS explores parameter space using a combination of global methods. DOPS is a novel hybrid meta-heuristic which combines a multi-swarm particle swarm method 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 multi-swarm particle swarm search and then dynamically switches, using an adaptive switching criteria, to the DDS approach. The particle swarm search uses multiple sub-swarms wherein the update to each particle (corresponding to a parameter vector estimate) is influenced by the best particle amongst the sub-swarm, and the current globally best particle. Particle updates occur within sub-swarms for a certain number of function evaluations, after which the sub-swarms are reorganized. This sub-swarm mixing is similar to the regrouping strategy described by Zhao et al. [36]. DOPS switches out of the particle swarm phase based upon an adaptive switching criteria that is a function of the rate of error convergence. If the error represented by the best particle does decrease for a threshold number of function evaluations, DOPS switches automatically to the DDS search phase. The DDS search is initialized with the globally best particle from the particle swarm phase, thereafter, the particle is greedily updated by perturbing a subset of dimensions for the remaining number of function evaluations. The identity of the parameters perturbed is chosen randomly, with fewer parameters perturbed the higher the number of function evaluations.

DOPS minimized benchmark problems using fewer function evaluations. On average, DOPS performed similarly or outperformed the four other meta-heuristics for the Ackley and Rastrigin 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 $\mathcal{N} = 4000$ and ran $\mathcal{T} = 25$ independent

4

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). DOPS also had a lower variability in the best value obtained (Fig. S8), faster convergence (Fig. S5, Fig. S6) and less stochasticity (Fig. S7) across multiple runs when compared to other metaheuristics. We also checked for statistical significance by doing a pairwise 2-sample t-test at 0.5% significance level. In each case we found that the p-value was close to zero disproving the null hypothesis and confirming the statistical significance of our results. Taken together, DOPS performed at least as well as other meta-heuristic approaches on small dimensional test problems, but seemed especially suited to large dimensional 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 [34]. They ranked the example problems by computational cost from most 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 CHO model) was a metabolic model of Chinese Hamster Ovary (CHO) with 35 metabolites, 32 reactions and 117 parameters [37]. The biochemical reactions were modeled using modular rate laws and generalized Michaelis-Menten kinetics. On the other hand, the expensive problem was a genome scale kinetic model of *Saccharomyces cerevisiae* with 261 reactions, 262 variables and 1759 parameters [38] (henceforth referred to as yeast model). In both cases, synthetic time series data gener-

ated with known parameter values, was used to estimate model parameters. For yeast model, 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 $\mathcal{N}$ = 4000, and trained both models against the synthetic experimental data. DOPS produced good fits to the synthetic data (Fig. S1 and Fig. S2), and recapitulated the nominal parameter values using only $\mathcal{N} \leq 4000$ function evaluations (Fig. S4). On the other hand, enhanced scatter search (eSS) with a local optimizer took on order 10^5 function evaluations for the same problems. Thus, DOPS estimated the parameters in benchmark biochemical models, and recovered the original parameters from synthetic data, using fewer function evaluations. Next, we compared the performance of DOPS with the four other meta-heuristics for a model of the human coagulation cascade.

DOPS estimated the parameters of a human coagulation model. Coagulation is an archetype biochemical network that is highly interconnected, containing both negative and positive feedback (Fig. 3). The biochemistry of coagulation, though complex, has been well studied [39–45], and reliable experimental protocols have been developed to interrogate the system [46–49]. 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. 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

6

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. 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. [49], which is a coupled system of non-linear ordinary differential equations where biochemical interactions were modeled using mass action kinetics. The Luan model contained 148 parameters and 92 species and has been validated using 21 published experimental datasets.

DOPS estimated the parameters of a human coagulation model for TF/VIIa initiated coagulation without anticoagulants (Fig. 5). The objective function was an unweighted linear combination of two error functions, representing coagulation initiated with different concentrations of TF/FVIIa (5pM, 5nM) [46]. We restricted the number of function evaluations to $\mathcal{N}$ = 4000 for each algorithm we tested, and performed $\mathcal{T}$ = 25 trials of each experiment to collect average performance data (Table 1). DOPS converged faster and had a lower final error compared to the other algorithms (Fig. 4). Within the first 25% of function evaluations, DOPS produced a rapid drop in error followed by a slower but steady decline. Approximately between 500-1000 function evaluations DOPS switched to the dynamically dimensioned search phase, however this transition varied from trial to trial since the switch was based upon the local convergence rate. On average, DOPS minimized the coagulation model error to a greater extent than the other meta-heuristics. However, it was unclear if the parameters estimated by DOPS had predictive power on

7

unseen data. To address this question, we used the final parameters estimated by DOPS to simulate data that was not used for training (coagulation initiated with 500pM, 50pM, and 10pM TF/VIIa). The optimal or near optimal parameters obtained by DOPS predicted unseen coagulation datasets (Fig. 6). The normalized standard error for the coagulation predictions was consistent with the training error, with the exception of the 50pM TF/VIIa case which was a factor 2.75 worse (Table 2). Coagulation initiation with 50pM TF/FVIIa was the farthest away from the training conditions. However, while the 50pM TF/FVIIa case was the worst of those tested, it still had an average error rate of only 11% (which is near the margin of error for the experimental measurements). Taken together, DOPS estimated parameter sets with predictive power on unseen coagulation data using fewer function iterations than other meta-heuristics. Next, we explored how the number of sub-swarms and the switch to DDS influenced the performance of the approach.

Phase switching was critical to DOPS performance. A differentiating feature of DOPS is the switch to dynamically dimensioned search following stagnation of the initial particle swarm phase. We quantified the influence of the number of sub-swarms and the switch to DDS on error convergence by comparing DOPS with and without DDS for different numbers of sub-swarms (Fig. 7). We considered multi swarm particle swarm optimization with and without the DDS phase for $\mathcal{N}$ = 4000 function evaluations and $\mathcal{T}$ = 25 trials on the coagulation model. We used one, two, four, five and eight sub-swarms, with a total of 40 particles divided evenly amongst the swarms. Hence, we did not consider swarm numbers of three and seven. All other algorithm parameters remained the same for all cases. Generally, the higher sub-swarm numbers converged in fewer function evaluations, where the optimum particle partitioning was in the neighborhood of five sub-swarms. However, the difference in convergence rate was qualitatively similar for four, five and eight sub-swarms, suggesting there was an optimal number of particles per swarm beyond which there was no significant advantage. The multi-swarm particle swarm optimization stag-

8

nated after 25% of the available function evaluations irrespective of the number of sub-swarms. However, DOPS (with five sub-swarms) switched to DDS after detecting the stagnation. The DDS phase refined the globally best particle to produce significantly lower error on average when compared to multi-swarm particle swarm optimization alone. Thus, the automated switching strategy was critical to the overall performance of DOPS.

Discussion

In this study, we developed dynamic optimization with particle swarms (DOPS), a novel meta-heuristic for parameter estimation in models of biological systems. DOPS combined multi-swarm particle swarm optimization, a global search approach, with the greedy strategy of dynamically dimensioned search to estimate optimal or nearly optimal solutions in a fixed number of function evaluations. We tested the performance of DOPS and four widely used meta-heuristics on the Ackley and Rastrigin test functions, a set of biochemical benchmark problems and a model of the human coagulation cascade. As the number of parameters increased, DOPS outperformed the other meta-heuristics, generating optimal or nearly optimal solutions using significantly fewer function evaluations compared with the other methods. We tested the solutions generated by DOPS by comparing the estimated and true parameters in the benchmark studies, and by using the coagulation model to predict unseen experimental data. For both benchmark problems, DOPS retrieved the true parameters in significantly fewer function evaluations than other meta-heuristics. For the coagulation model, we used experimental coagulation measurements under two different conditions to estimate optimal or nearly optimal parameters. These parameters were then used to predict unseen coagulation data; the coagulation model parameters estimated by DOPS predicted the correct thrombin dynamics following TF/FVIIa induced coagulation without anticoagulants. Lastly, we showed the average performance of DOPS improved when combined with dynamically dimensioned search phase, compared to an identical multi-swarm approach alone. Taken together, DOPS is a promising meta-heuristic for the estimation of parameters in large biochemical models.

Meta-heuristics can be effective techniques to estimate optimal or nearly optimal solutions for complex, multi-modal functions. However, meta-heuristics typically require a large number of function evaluations to converge to a solution. DOPS is a combination of particle swarm optimization, which is a global search method, and dynamically dimen-

10

sioned search, which is a greedy evolutionary technique. Particle swarm optimization uses collective information shared amongst swarms of computational particles to search for global extrema. Several particle swarm variants have been proposed to improve the search ability and rate of convergence. These variations involve different neighborhood structures, multi-swarms or adaptive parameters. Multi-swarm particle swarm optimization with small particle neighborhoods has been shown to be better in searching on complex multi-modal solutions [36]. Multi-swarm methods generate diverse solutions, and avoid rapid convergence to local optima. However, at least for the coagulation problem used in this study, multi-swarm methods stagnated after approximately 25% of the available function evaluations; only the introduction of dynamically dimensioned search improved the rate of error convergence. Dynamically dimensioned search, which greedily perturbs only a subset of parameter dimensions in high dimensional parameter spaces, refined the globally best particle and produced significantly lower error on average when compared to multi-swarm particle swarm optimization alone. However, dynamically dimensioned search, starting from a initial random parameter guess, was not as effective on average as DOPS. The initial solutions generated by the multi swarm search had a higher propensity to produce good parameter estimates when refined by dynamically dimensioned search. Thus, our hybrid combination of two meta-heuristics produced better results than either constituent approach, and better results than other meta-heuristic approaches on average. This was true of not only the convergence rate on the coagulation problem, but also the biochemical benchmark problems; DOPS required two-orders of magnitude fewer function evaluations compared with enhanced Scatter Search (eSS) to estimate the biochemical benchmark model parameters. Taken together, the combination of particle swarm optimization and dynamically dimensioned search produced better results than either of these constituent approaches alone, and required fewer function evaluations compared with other common meta-heuristics.

11

DOPS performed well on many different systems with no pre-optimization of algorithm parameters, however there are many research questions that should be pursued further. DOPS comfortably outperformed existing, widely used meta-heuristics for high dimensional global optimization functions, biochemical benchmark models and a model of the human coagulation system. However, it is possible that highly optimized versions of common meta-heuristics could surpass DOPS; we should compare the performance of DOPS with optimized versions of the common meta-heuristics on both test and real-world problems to determine if a performance advantage exists in practice. Next, DOPS has a hybrid architecture, thus the particle swarm phase could be combined with other search strategies such as local derivative based approaches to improve convergence rates. We could also consider multiple phases beyond particle swarm and dynamically dimensioned search, for example switching to a gradient based search following the dynamically dimensioned search phase. Next, the current implementation of the algorithm switches only once from the swarm phase to the dynamically dimensioned phase. An obvious modification would be to allow multiple transitions between the swarm and dynamically dimensioned phases. Lastly, we should update DOPS to treat multi-objective problems. The identification of large biochemical models sometimes requires training using qualitative, conflicting or even contradictory data sets. One strategy to address this challenge is to estimate experimentally constrained model ensembles using multiobjective optimization. Previously, we developed Pareto Optimal Ensemble Techniques (POETs) which integrates simulated annealing with Pareto optimality to identify models near the optimal tradeoff surface between competing training objectives [50]. Since DOPS consistently outperformed simulated annealing on both test and real-world problems, we expect a multi-objective form of DOPS would more quickly estimate solutions which lie along high dimensional trade-off surfaces.

12

Materials and Methods

Optimization problem formulation. Model parameters were estimated by minimizing the difference between model simulations and $\mathcal{E}$ experimental measurements. Simulation error is quantified by an objective function $K(\mathbf{p})$ (typically the Euclidean norm of the difference between simulations and measurements) subject to problem and parameter constraints:

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

The term $K(\mathbf{p})$ denotes the objective function, t denotes time, $g_i(t_i, \mathbf{x}, \mathbf{p}, \mathbf{u})$ is the model output for experiment i, $\mathbf{x}(t, \mathbf{p})$ is the state variable vector with an initial state $\mathbf{x}_0$, $\mathbf{u}(t)$ is a model input vector, $\mathbf{f}(t, \mathbf{x}(t, \mathbf{p}), \mathbf{u}(t), \mathbf{p})$ is the system of model equations (e.g., differential equations or algebraic constraints) and $\mathbf{p}$ denotes the model parameter vector. The parameter search (or model simulations) can be subject to $\mathbf{c}(t, \mathbf{x}, \mathbf{p}, \mathbf{u})$ linear or non-linear constraints, and parameter bound constraints where $\mathbf{p}^L$ and $\mathbf{p}^U$ denote the lower and upper parameter bounds, respectively. Optimal model parameters are then given by:

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

Dynamic optimization with particle swarms (DOPS). DOPS is a novel meta-heuristic which combines multi-swarm particle swarm optimization with the dynamically dimensioned search (Fig. 1) and (Algo. 1). Particle swarm optimization is an evolutionary

algorithm that uses a population of particles (solutions) to find an optimal solution. Each particle is updated based on its on experience (particle best) and the experience of all other particles within the swarm (global best) [51][52][53]. Multi-swarm particle swarm optimization is a variant of particle swarm optimization that uses multiple swarms instead of a single swarm. DOPS uses a novel update rule of particles within each swarm. This rule uses only particle positions but not particle velocities in the update rule. This differs from existing PSO methods that rely both on particle position and velocity for update. The update rule further described in detail in the methods. 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 a DDS search phase. Dynamically Dimensioned Search (DDS) is a single solution based search algorithm. It was developed by Tolson and Shoemaker to obtain good solutions to high-dimensional search problems within a fixed number of function evaluations. DDS starts as a global search algorithm by perturbing all the dimensions. Later the number of dimensions that are perturbed is decreased with a certain probability. The probability that a certain dimension is perturbed reduces (a minimum of one dimension is always perturbed) as the iterations increase. This causes the algorithm to behave as a local search algorithm as the number of iterations increase. The perturbation magnitude of each dimension is from normal distribution with zero mean. The standard deviation that was used in the original DDS paper and the current study is 0.2. DDS performs a greedy search where the solution is updated only if it is better than the previous solution. The combination of perturbing a subset of dimensions along with greedy search indirectly relies on model sensitivity to a specific parameter combination. For the sake of brevity we are not providing the entire pseudo code of DDS here. The reader is requested to refer to Figure 1 of the original paper by Tolson and Shoemaker [35].

14

input : A randomized swarm of particles of size $NP \times K$ and fixed number of function evaluations N

output: Optimized parameter vector of size $1 \times K$

1 *Initialize the particles randomly and assign particles randomly to k sub-swarms*;

2 while $j \leq N$ **do**

3   **if** *mod(j,G)=0* **then**

4   *Reassign particles to k sub-swarms*;

5   **end**

6   **for** $i \leftarrow 1$ **to** NS **do**

7   *Update particles within sub-swarms according to equation 3*;

8   **end**

9   *Find best particle $\mathcal{G}$ amongst all sub-swarms*;

10   **if** $besterror(j) \geq 0.99 * besterror(j+1)$ **then**

11 $failurecounter \leftarrow failurecounter + 1$;

12   **else**

13 $failurecounter \leftarrow 0$;

14   **end**

15   **if** $failurecounter \geq threshold$ **then**

16 $\mathcal{G} \leftarrow DDS(\mathcal{G}, N - j)$;

17   **return** $\mathcal{G}$

18   **else**

19 $j \leftarrow j + 1$;

20   **end**

21   **return** $\mathcal{G}$

22 end

Algorithm 1: Pseudo code for the dynamic optimization with particle swarms (DOPS) method.

Phase 1: Particle swarm phase. The particle swarm phase of DOPS begins 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 equally sized sub-swarms $\mathcal{S}_1, \ldots, \mathcal{S}_k$. Particles within each

sub-swarm $\mathcal{S}_k$ were updated according to the rule:

$$z_{i,j} = \theta_{1,j-1} z_{i,j-1} + \theta_2 r_1 \left(\mathcal{L}_i - z_{i,j-1} \right) + \theta_3 r_2 \left(\mathcal{G}_k - z_{i,j-1} \right) \tag{3}$$

where $(\theta_1, \theta_2, \theta_3)$ were adjustable parameters, $\mathcal{L}_i$ denotes the best solution found by parti-

cle i within sub-swarm $\mathcal{S}_k$ for function evaluation $1 \rightarrow j - 1$, and $\mathcal{G}_k$ denotes the best solu-

tion 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 $\theta_{1,j-1}$ is similar to the inertia weight

parameter for the velocity term described by Shi and Eberhart [54]; Shi and Eberhart pro-

posed a linearly decreasing inertia weight to improve convergence properties of particle

swarm optimization. Our implementation of $\theta_{1,j-1}$ is inspired by this and the decreas-

ing perturbation probability proposed by Tolson and Shoemaker [35]. It is an analogous

equivalent to inertia weight on velocity. However $\theta_{1,j-1}$ places inertia on the position rather

than velocity and uses the same rule described by Shi and Eberhart to adaptively change

with the number of function evaluations:

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

where $\mathcal{N}$ represents the total number of function evaluations, w_{max} and w_{min} are the max-

imum and minimum inertia weights, respectively. While updating the particles, parameter

bounds were enforced using reflection boundary conditions (Algo. 2).

```
 1  if $z_{i,j}^{old} < z_i^{min}$ then
 2  |    $z_{i,j}^{new} = z_{i,j}^{old} + (z_i^{min} - z_{i,j}^{old})$ if $z_{i,j}^{new} > z_i^{max}$ then
 3  |    |    $z_{i,j}^{new} = z_i^{max}$
 4  |    end
 5  end
 6  if $z_{i,j}^{old} > z_i^{max}$ then
 7  |    $z_{i,j}^{new} = z_{i,j}^{old} + (z_{i,j}^{old} - z_i^{max})$ if $z_{i,j}^{new} < z_i^{min}$ then
 8  |    |    $z_{i,j}^{new} = z_i^{min}$
 9  |    end
10  end
```

Algorithm 2: Pseudo code for the reflective boundary conditions used by the dynamic optimization with particle swarms (DOPS) method.

349 After every $\mathcal{M}$ function evaluations, particles were randomly redistributed to a new

350 sub-swarm, and updated according to Eqn. (3). This process continued for a maximum of

351 $\mathcal{F} * \mathcal{N}$ functions evaluations, where $\mathcal{F}$ is the fraction of evaluations in the particle swarm

352 phase of DOPS. However, if the simulation error stagnated e.g., did not change by more

353 than 1% for a specified number of evaluations, the swarm phase was terminated and

354 DOPS switched to exploring parameter space using the DDS approach.

```
   input  : Candidate vector $\mathcal{G}$ from swarm search and $(1 - \mathcal{F}) * \mathcal{N}$ evaluations
   output : Optimized parameter vector of size $1 \times K$

 1  while $j \leq (1 - \mathcal{F}) * \mathcal{N}$ do
 2  |    Assign probability of perturbation to each dimension $\mathcal{P}_i$ according to
    |    equation 7;
 3  |    Select a subset of dimensions based on a threshold value for perturbation;
 4  |    Update candidate solution $\mathcal{G}(J)$ according to equation 5;
 5  |    Ensure updated solution $\mathcal{G}_{new}(J)$ is within bounds using Algorithm 2;
 6  end
```

Algorithm 3: Pseudo code for the Dynamically Dimensioned Search (DDS) method.

355 *Phase 2: DDS phase.* At the conclusion of the swarm phase, the overall best particle,

356 $\mathcal{G}_k$, over the k sub-swarms was used to initialize the DDS phase. DOPS takes at least

357 $(1 - \mathcal{F}) * \mathcal{N}$ function evaluations during the DDS phase and then terminates the search.

17

For the DDS phase, the best parameter estimate was updated using the rule:

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

where $\mathbf{J}$ is a vector representing the subset of dimensions that are being perturbed, $\mathbf{r}_{normal}$ denotes a normal random vector of the same dimensions as $\mathcal{G}$, and σ denotes the perturbation amplitude:

$$
\sigma = R(\mathbf{p}^U - \mathbf{p}^L) \tag{6}
$$

where R is the scalar perturbation size parameter, $\mathbf{p}^U$ and $\mathbf{p}^L$ 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 $\mathcal{P}_i$ that represents a threshold for determining whether a specific dimension j was perturbed or not; $\mathcal{P}_i$ is monotonically decreasing function of function evaluations:

$$
\mathcal{P}_i = 1 - \log\left[\frac{i}{(1 - \mathcal{F}) * \mathcal{N}}\right] \tag{7}
$$

where i is the current iteration. After $\mathcal{P}_i$ is determined, we drew $\mathcal{P}_j$ from a uniform distribution for each dimension j. If $\mathcal{P}_j < \mathcal{P}_i$ was included in $\mathbf{J}$. Thus, the probability that a dimension j was perturbed was inversely proportional to the number of function evaluations. DDS updates are greedy; $\mathcal{G}_{new}$ becomes the new solution vector only if it is better than $\mathcal{G}$. DOPS source code is available for download under a MIT license at http://www.varnerlab.org.

18

Acknowledgements

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. Brief Bioinform 7: 364-74.

3. Jaqaman K, Danuser G (2006) Linking data to models: data regression. Nat Rev Mol Cell Biol 7: 813-9.

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: 1195-203.

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. Auto-

matica 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. Goldberg DE (2006) Genetic algorithms. Pearson Education India .

20. Kirkpatrick S, Gelatt CD, Vecchi MP, et al. (1983) Optimization by simulated annealing. science 220: 671–680.

21. Fogel D (2009) Artificial intelligence through simulated evolution. Wiley-IEEE Press .

22. 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.

23. Tsai KY, Wang FS (2005) Evolutionary optimization with data collocation for reverse engineering of biological networks. Bioinformatics 21: 1180–1188.

24. 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.

25. Noman N, Iba H (2007) Inferring gene regulatory networks using differential evolution with local search heuristics. IEEE/ACM Transactions on Computational Biology and

427   Bioinformatics (TCBB) 4: 634–647.

26. 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.

27. Mendes P, Kell D (1998) Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. Bioinformatics 14: 869–883.

28. 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.

29. 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.

30. Villaverde AF, Egea JA, Banga JR (2012) A cooperative strategy for parameter estimation in large scale systems biology models. BMC systems biology 6: 75.

31. Rodriguez-Fernandez M, Egea JA, Banga JR (2006) Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. BMC bioinformatics 7: 483.

32. 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.

33. Egea JA, Martí R, Banga JR (2010) An evolutionary method for complex-process optimization. Computers & Operations Research 37: 315–324.

34. 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.

35. Tolson BA, Shoemaker CA (2007) Dynamically dimensioned search algorithm for

453 computationally efficient watershed model calibration. Water Resources Research
454 43.

36. Zhao SZ, Liang JJ, Suganthan PN, Tasgetiren MF (2008) Dynamic multi-swarm parti-
cle swarm optimizer with local search for large scale global optimization. Evolutionary
Computation, 2008 CEC 2008 : 3845–3852.

37. Villaverde AF, Bongard S, Mauch K, Müller D, Balsa-Canto E, et al. (2014) High-
confidence predictions in systems biology dynamic models. 8th International Con-
ference on Practical Applications of Computational Biology & Bioinformatics (PACBB
2014) : 161–171.

38. Smallbone K, Mendes P (2013) Large-scale metabolic models: From reconstruction
to differential equations. Industrial Biotechnology 9: 179–184.

39. Mann KG, Butenas S, Brummel K (2003) The dynamics of thrombin formation. Arte-
riosclerosis, thrombosis, and vascular biology 23: 17–25.

40. Mann K, Brummel K, Butenas S (2003) What is all that thrombin for? Journal of
Thrombosis and Haemostasis 1: 1504–1514.

41. Mann KG (2003) Thrombin formation. CHEST Journal 124: 4S–10S.

42. Vogler EA, Siedlecki CA (2009) Contact activation of blood-plasma coagulation. Bio-
materials 30: 1857–1869.

43. Diamond SL (2013) Systems biology of coagulation. Journal of Thrombosis and
Haemostasis 11: 224–232.

44. Fogelson AL, Tania N (2005) Coagulation under flow: the influence of flow-mediated
transport on the initiation and inhibition of coagulation. Pathophysiology of haemosta-
sis and thrombosis 34: 91–108.

45. Anand M, Rajagopal K, Rajagopal K (2003) A model incorporating some of the me-
chanical and biochemical factors underlying clot formation and dissolution in flowing
blood: review article. Journal of Theoretical Medicine 5: 183–218.

23

46. 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.

47. 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 .

48. Mann KG, Brummel-Ziedins K, Orfeo T, Butenas S (2006) Models of blood coagulation. Blood Cells, Molecules, and Diseases 36: 108–117.

49. Luan D, Zai M, Varner JD (2007) Computationally derived points of fragility of a human cascade are consistent with current therapeutic strategies. PLoS computational biology 3: e142.

50. Song SO, Chakrabarti A, Varner JD (2010) Ensembles of signal transduction models using pareto optimal ensemble techniques (POETs). Biotechnology journal 5: 768–780.

51. Clerc M (2010) Particle swarm optimization, volume 93. John Wiley & Sons.

52. Blondin J (2009) Particle swarm optimization: A tutorial. from site: http://cs armstrong edu/saad/csci8100/pso tutorial pdf 93.

53. Abraham A, Guo H, Liu H (2006) Swarm intelligence: foundations, perspectives and applications. In: Swarm Intelligent Systems, Springer. pp. 3–25.

54. Shi Y, Eberhart RC (1999) Empirical study of particle swarm optimization. In: Proceedings of the 1999 Congress on Evolutionary Computation, 1999. CEC 99. volume 3, p. 1950.

Table 1: 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. Here $pnom$ indicates the nominal or true parameter vector of the model. 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	Yeast model	CHO model	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

Table 2: Error analysis for the human coagulation model. The coagulation model was trained on coagulation initiated with TF/FVIIa at 5 nM and the 5 pM to obtain the optimal parameters. Using these optimal parameters, coagulation dynamics were predicted for varying initiator concentrations (500 pM, 50 pM and 10 pM). Model agreement with measurements was quantified using normalized squared error. The normalized squared error is defined as $N.S.E. = (1/max(\mathbf{X})) * (\|(\mathbf{Y}, \mathbf{X})\|/sqrt(\mathbf{N}))$ where $\mathbf{X}$ is the experimental data, $\mathbf{Y}$ is the model simulation data interpolated onto the experimental time scale and N is the total number of experimental time points.

TF/FVIIa concentration	Normalized S.E.	Category
5 nM	0.0376	Training
500 pM	0.0564	Prediction
50 pM	0.1125	Prediction
10 pM	0.0823	Prediction
5 pM	0.0338	Training
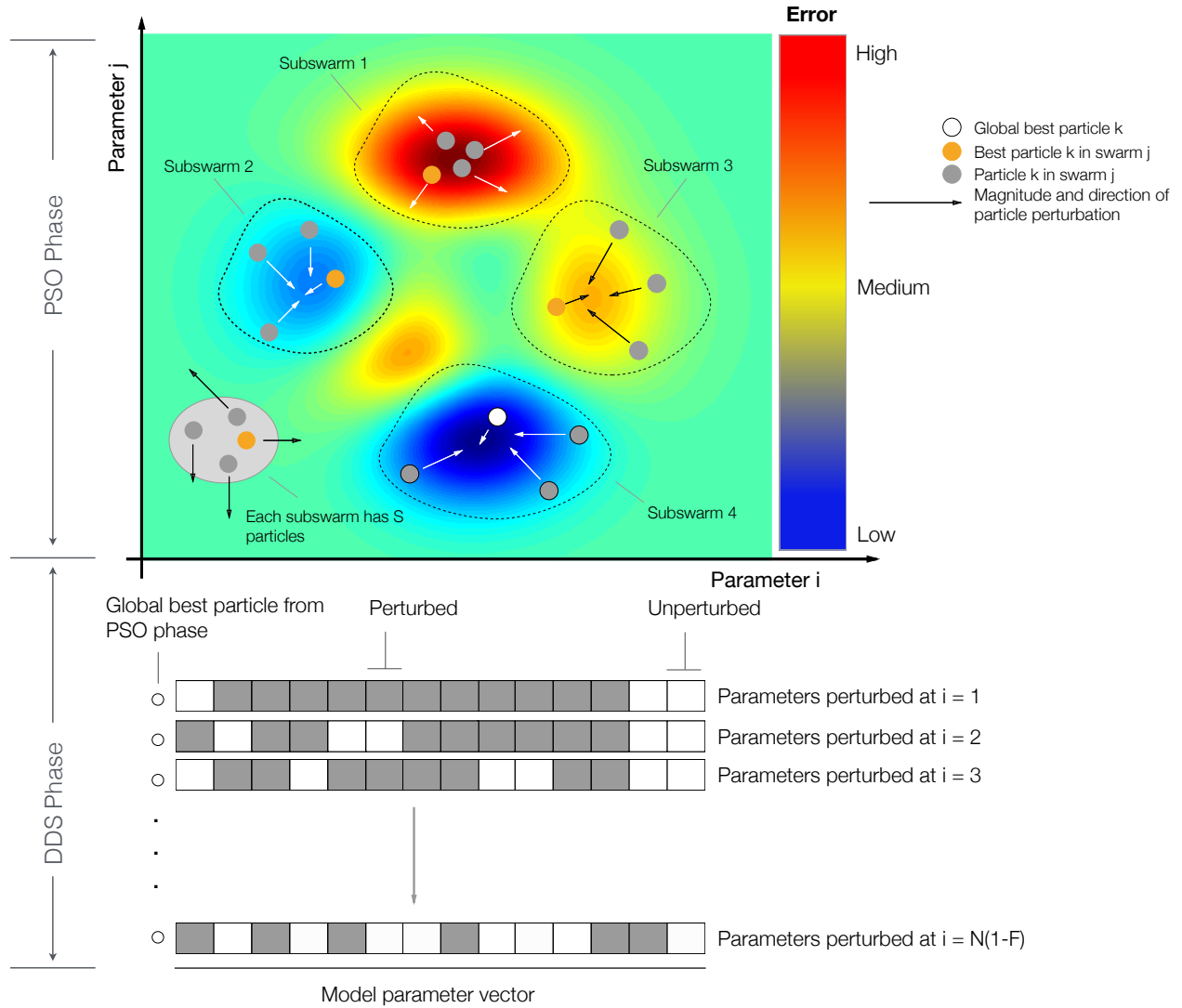
Fig. 1: Schematic of the dynamic optimization with particle swarms (DOPS) approach. **A**: Each particle represents an N dimensional parameter vector. Particles are given randomly generated initial solutions and grouped into different sub-swarms. Within each swarm the magnitude and direction of the movement 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.
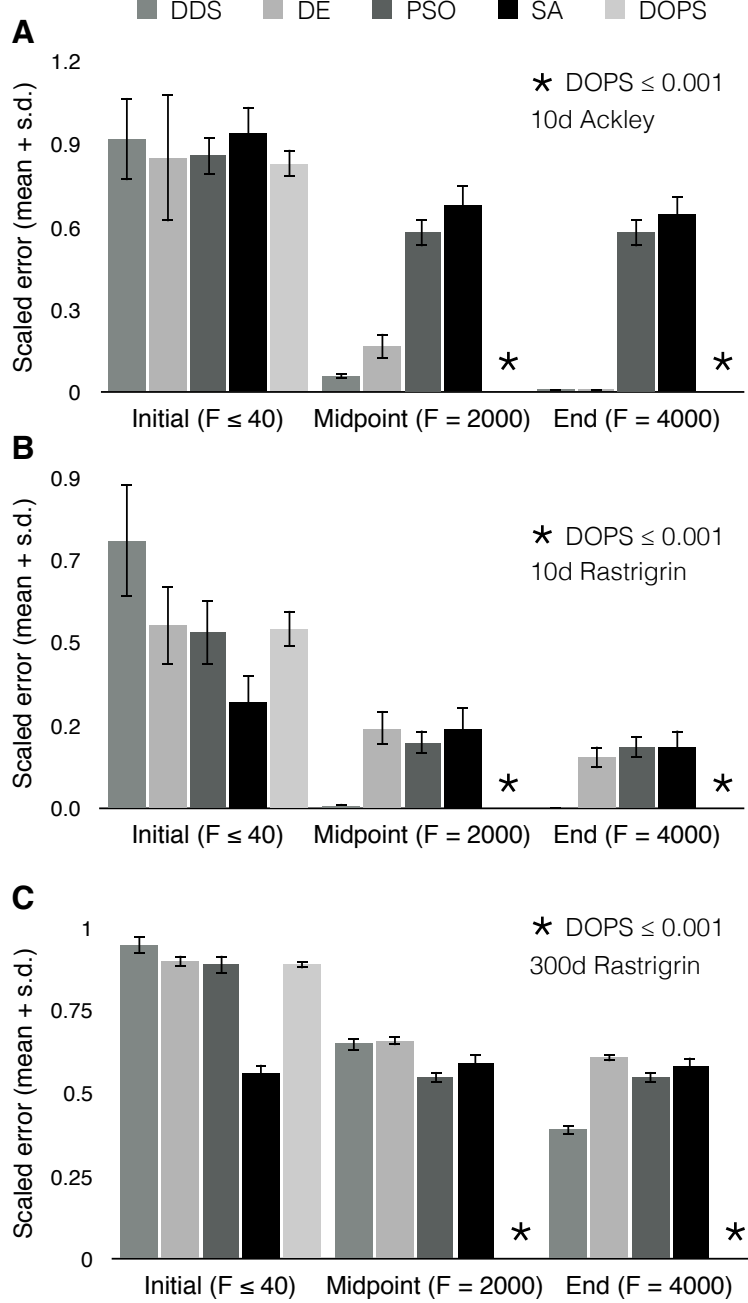
Fig. 2: Performance of DOPS and other meta-heuristics for the Ackley and Rastrigin functions. **A**: Mean scaled error versus the number of function evaluations for the 10-dimensional Ackley function. DOPS, DDS and DE find optimal or near optimal solutions within the specified number of function evaluations. **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 $\mathcal{N} = 4000$. Mean and standard deviation were calculated over $\mathcal{T} = 25$ trials.
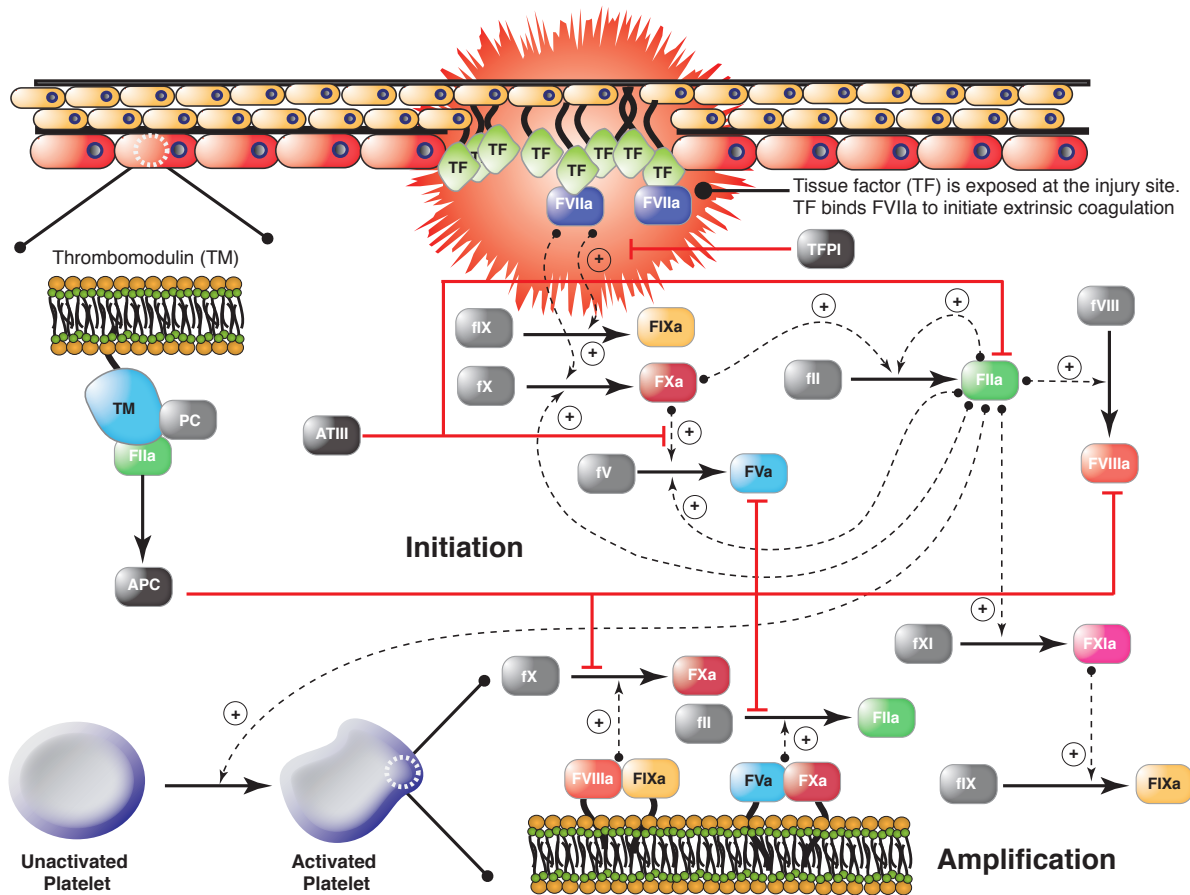
Fig. 3: Schematic of the extrinsic and intrinsic coagulation cascade. Inactive zymogens upstream (grey) are activated by exposure to tissue factor (TF) following vessel injury. Tissue factor and activated factor VIIa (FVIIa) form a complex that activates factor X (fX) and IX (fIX). FXa activates downstream 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) down-regulates 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.
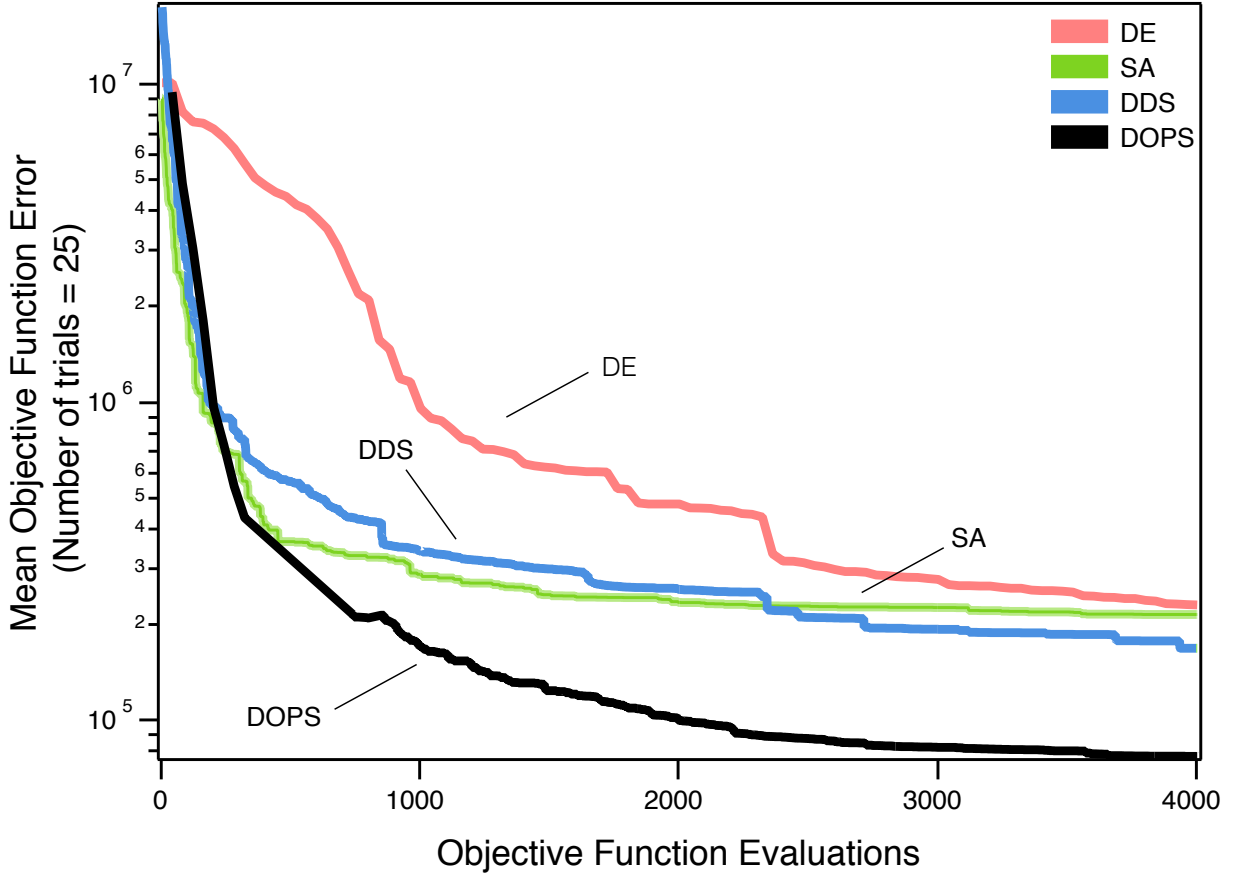
Fig. 4: Error convergence rates of the five different algorithms on the coagulation model. The objective error is the mean over $\mathcal{T}$ = 25 trials. DOPS, DDS and SA have the steepest drop in error during first 300 function evaluations. Thereafter the error drop in DDS and SA remains nearly constant 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 three times greater than the lowest error using DDS.
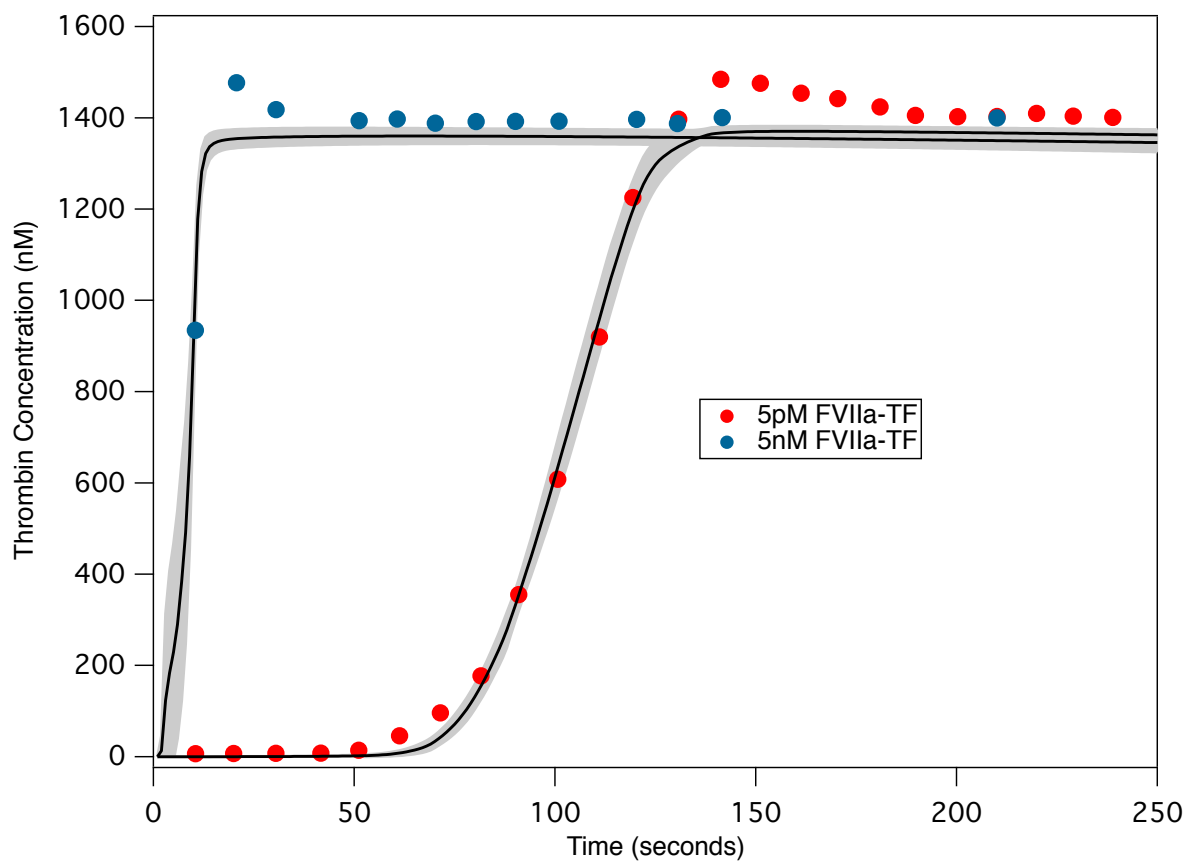
Fig. 5: Model fits on experimental data using DOPS. The model parameters were estimated using DOPS. Solid black lines indicate the simulated mean thrombin concentration using parameter vectors from 25 trials. The grey shaded region represents the 99% confidence estimate of the mean simulated thrombin concentration. The experimental data is reproduced from the synthetic plasma assays of Mann and co-workers. Thrombin generation is initiated by adding Factor TF/VIIa (5nM and 5pM) 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.
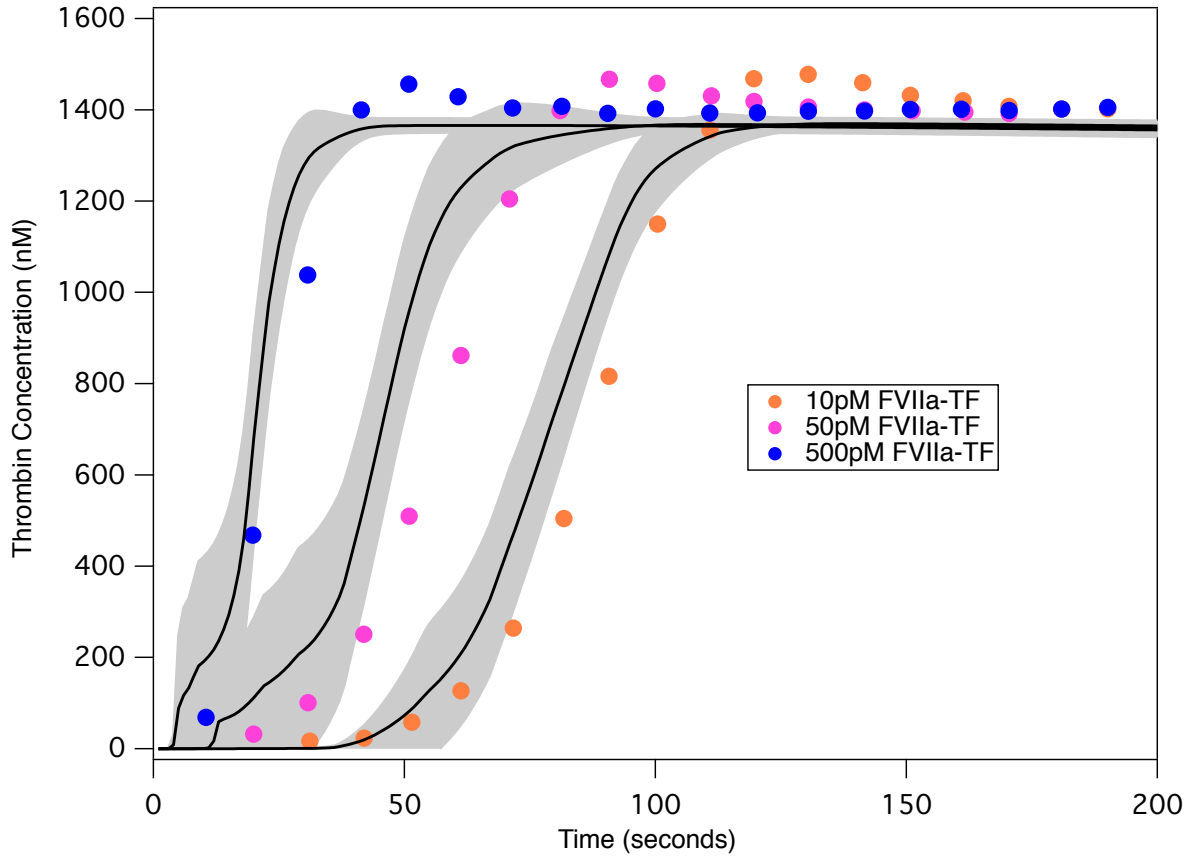
Fig. 6: Model predictions on unseen experimental data using parameters obtained from DOPS. The parameter estimates that were obtained using DOPS were tested against data that was not used in the model training. Solid black lines indicate the simulated mean thrombin concentration using parameter vectors from $\mathcal{T} = 25$ 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.
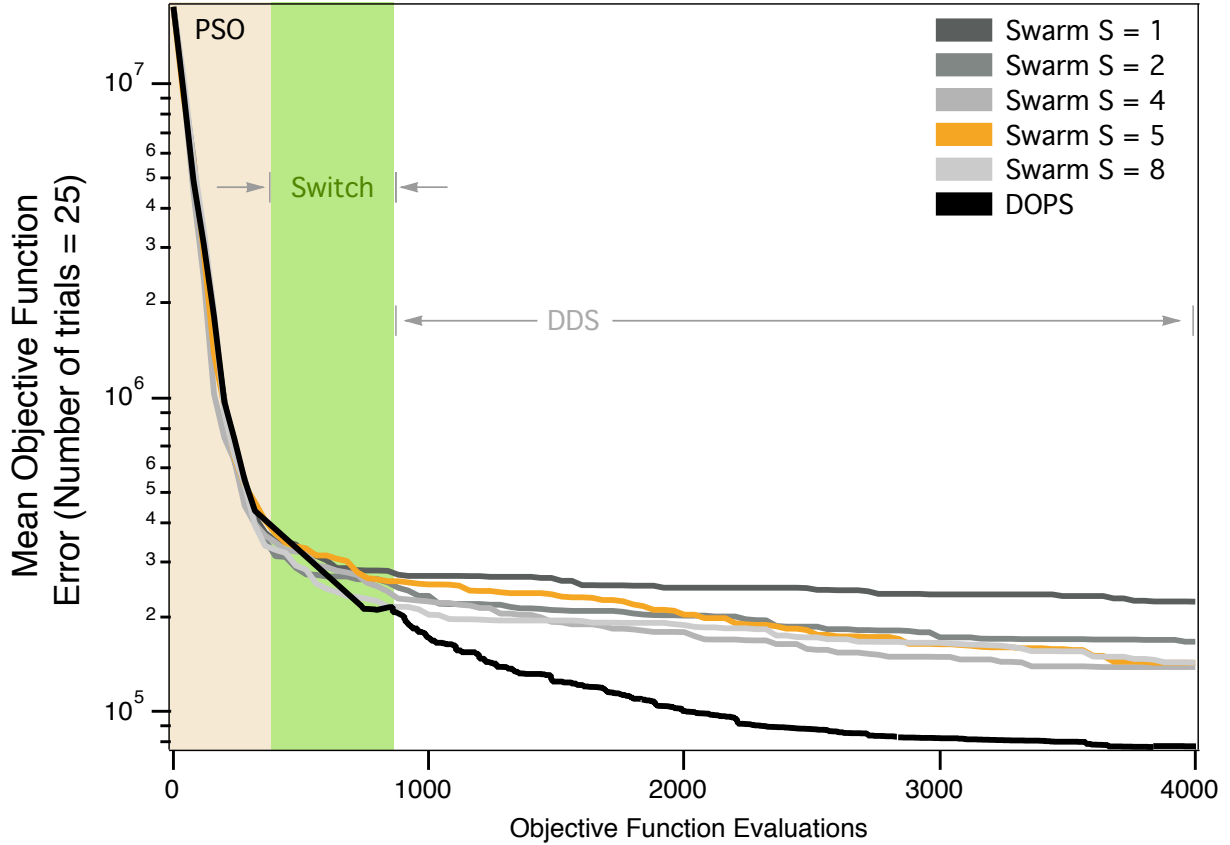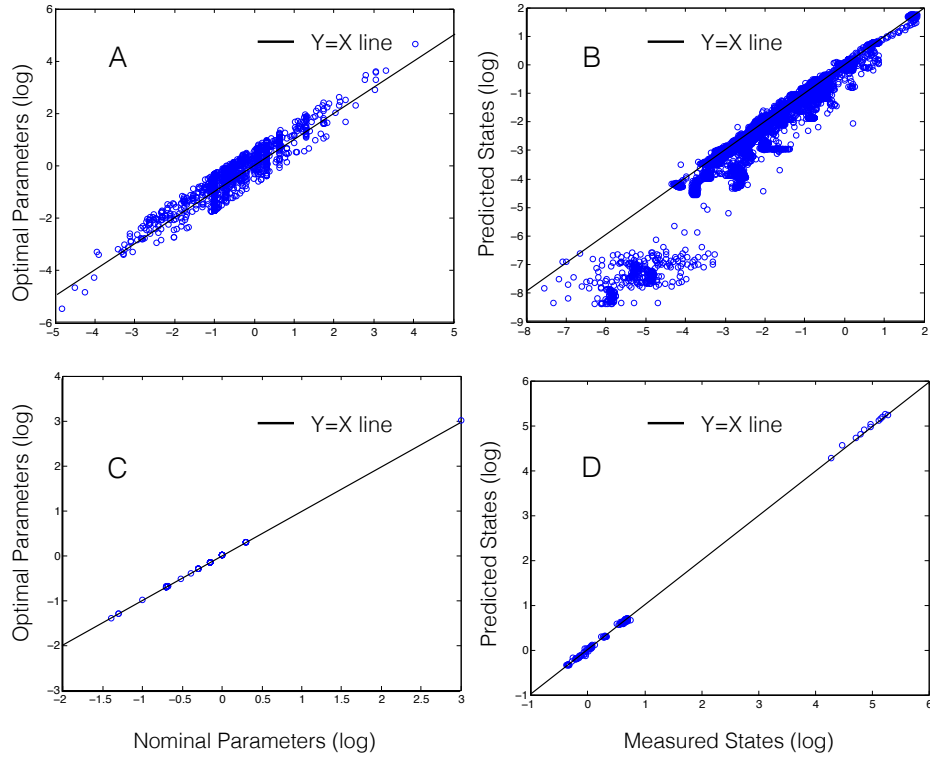
Fig. 7: Influence of the switching strategy and sub-swarms on DOPS performance for the coagulation model. DOPS begins by using a particle swarm search and then dynamically switches (switch region), using an adaptive switching criteria, to the DDS search phase. We compared the performance of DOPS with and without DDS for different sub-swarm searches to quantify the effect of number of sub-swarms and DDS. We used one, two, four, five and eight sub-swarms, with a total of 40 particles divided evenly amongst the swarms. The results presented are the average of $\mathcal{T}$ = 25 trials with $\mathcal{N}$ = 4000 function evaluations each. The convergence rates with higher swarm numbers is typically higher but there is no pronounced difference amongst four, five and eight. The multi-swarm with DDS saturates while DOPS shows a rapid drop due to a switch to the DDS phase.
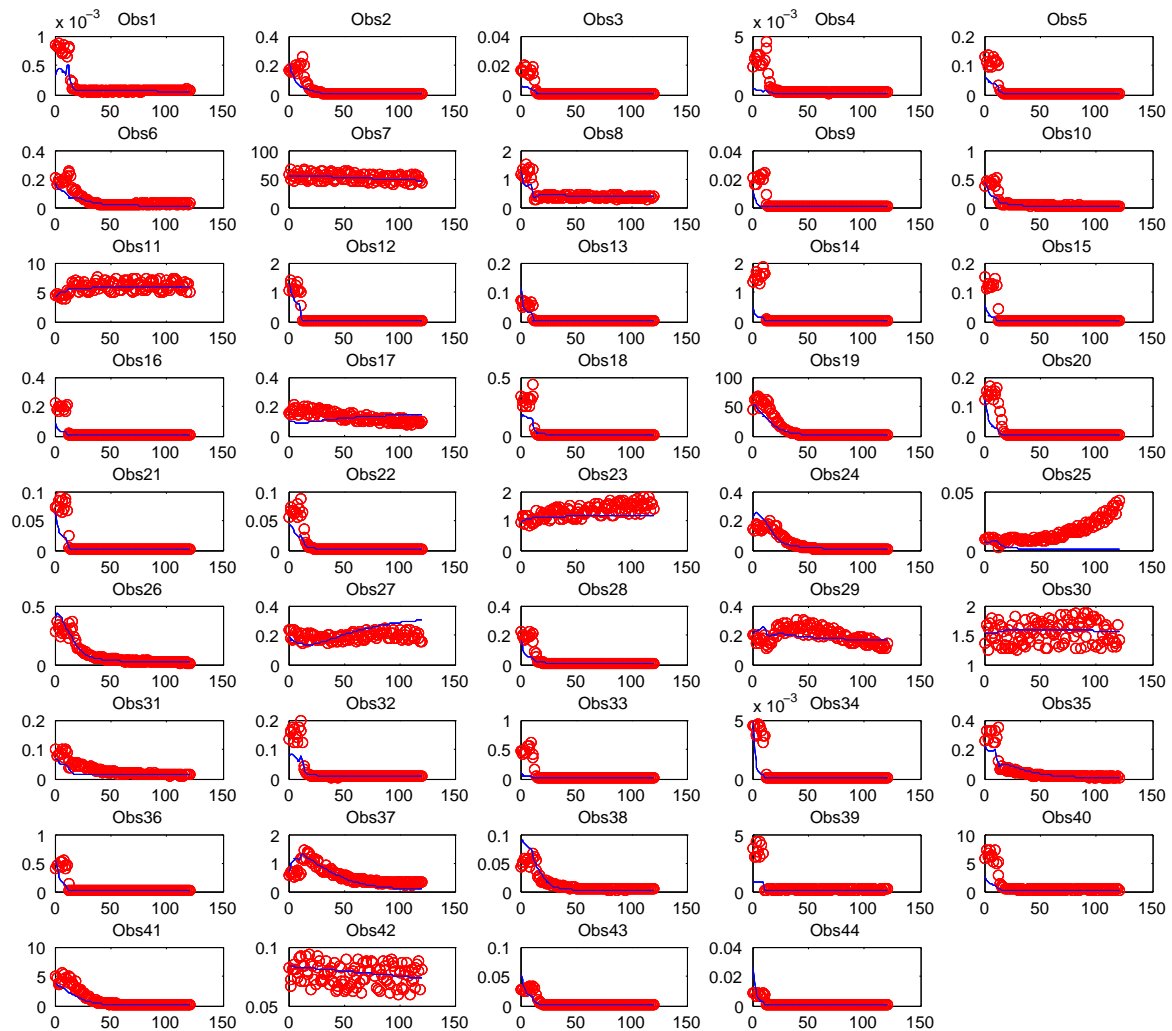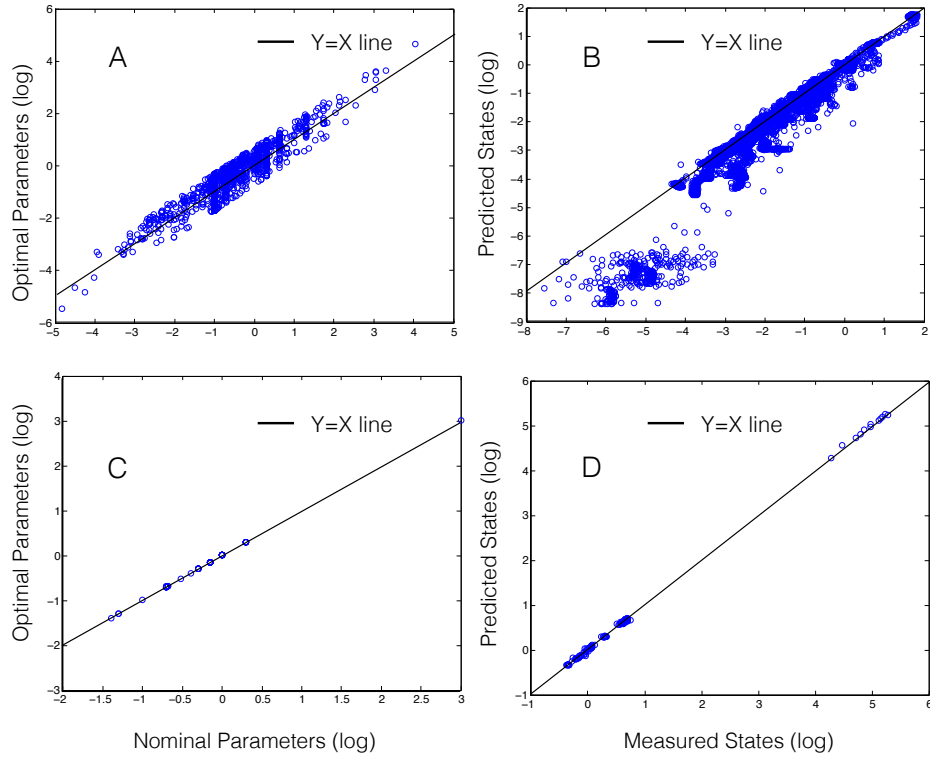
Fig. S1: (Data fits for Yeast Model (B1)) Pseudo-experimental data (red circles) vs. optimal solution obtained using DOPS (solid blue lines) for the 44 observed states. X axis: time [s]; Y axis: metabolite concentrations [mM].
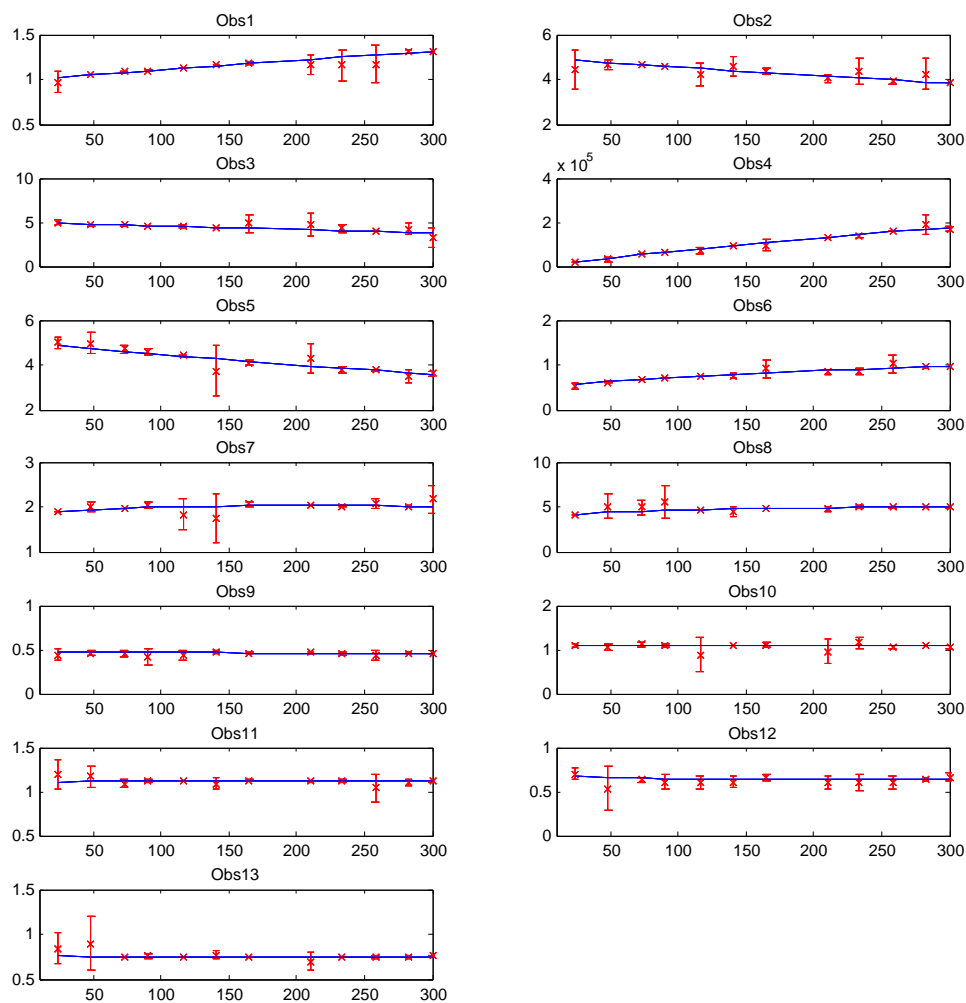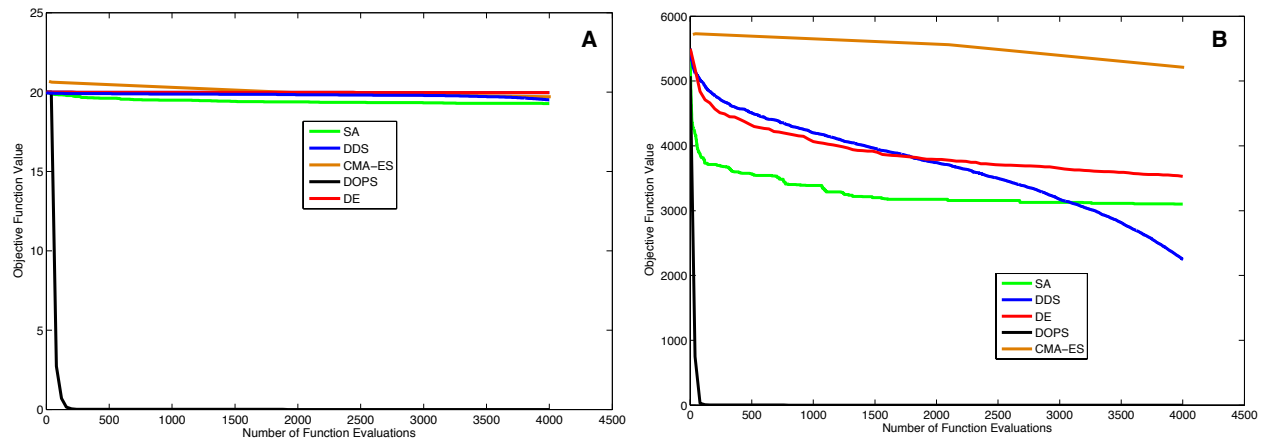
Fig. S2: (Data fits for CHO Model (B4)) Pseudo-experimental data (red x) vs. optimal solution obtained using DOPS (solid blue lines) for the 13 observed states. X axis: time [s]; Y axis: metabolite concentrations [mM].

S-2

Fig. S3: (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 yeast model: 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 CHO model: Metabolic model of Chinese Hamster Ovary Cells (CHO) cells with 117 parameters.

Fig. S4: (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 yeast model: 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 CHO model: Metabolic model of Chinese Hamster Ovary Cells (CHO) cells with 117 parameters.
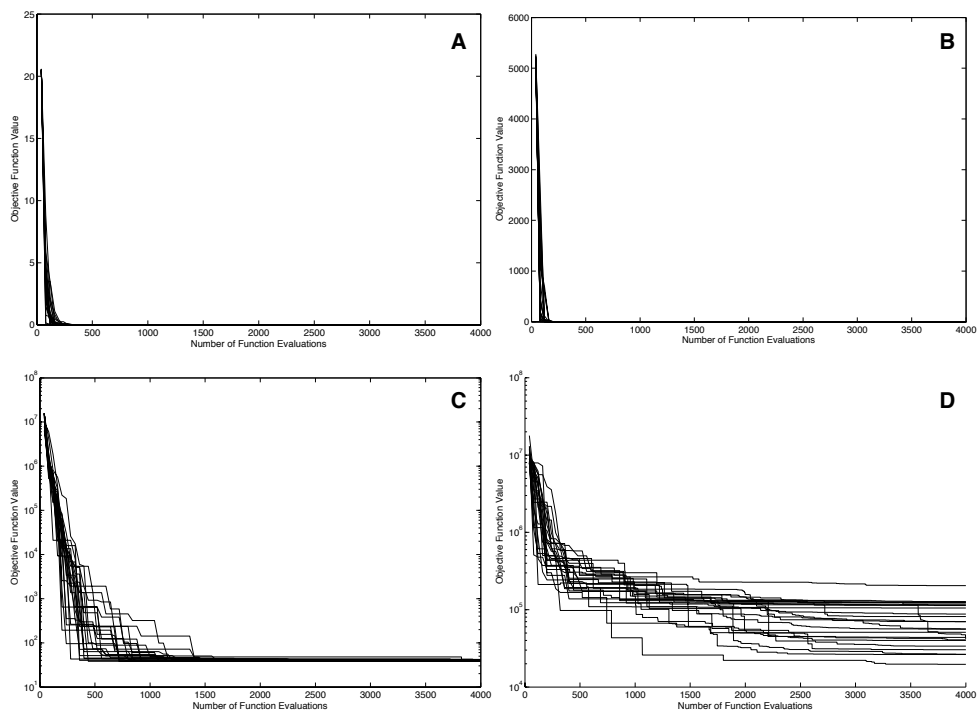
Fig. S5: Convergence curves for different metaheuristics along with CMA-ES) Objective function value versus the number of function evaluations for **(A)** Ackley 300 dimensional **(B)** and Rastrigin 300 dimensional.
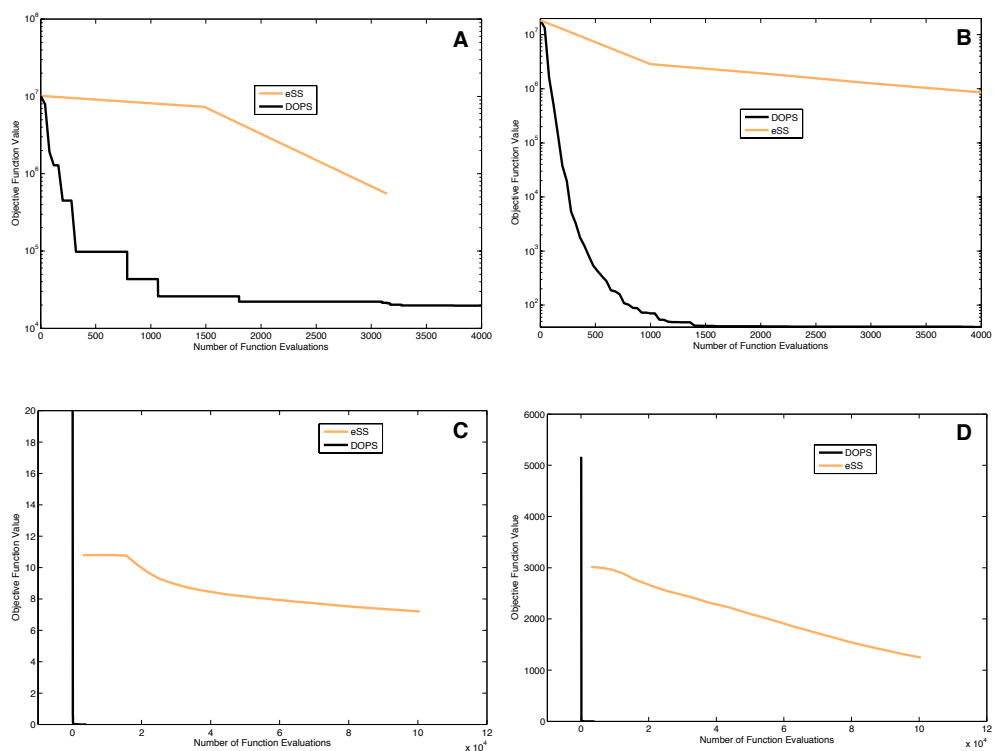
Fig. S6: (Convergence curves for eSS and DOPS) Objective function value versus the number of function evaluations plotted for eSS and DOPS on **(A)** Coagulation **(B)** CHO Model **(C)** Ackley 300 dimensional and **(D)** Rastrigin 300 dimensional.

Fig. S7: (Dispersion curves for DOPS) Objective function value (25 trials) versus the number of function evaluations plotted for DOPS on **(A)** Ackley 300 dimensional **(B)** Rastrigin 300 dimensional **(C)**CHO Model - B4 **(D)** Coagulation. For CHO Model and Coagulation Y-axis is in log scale.
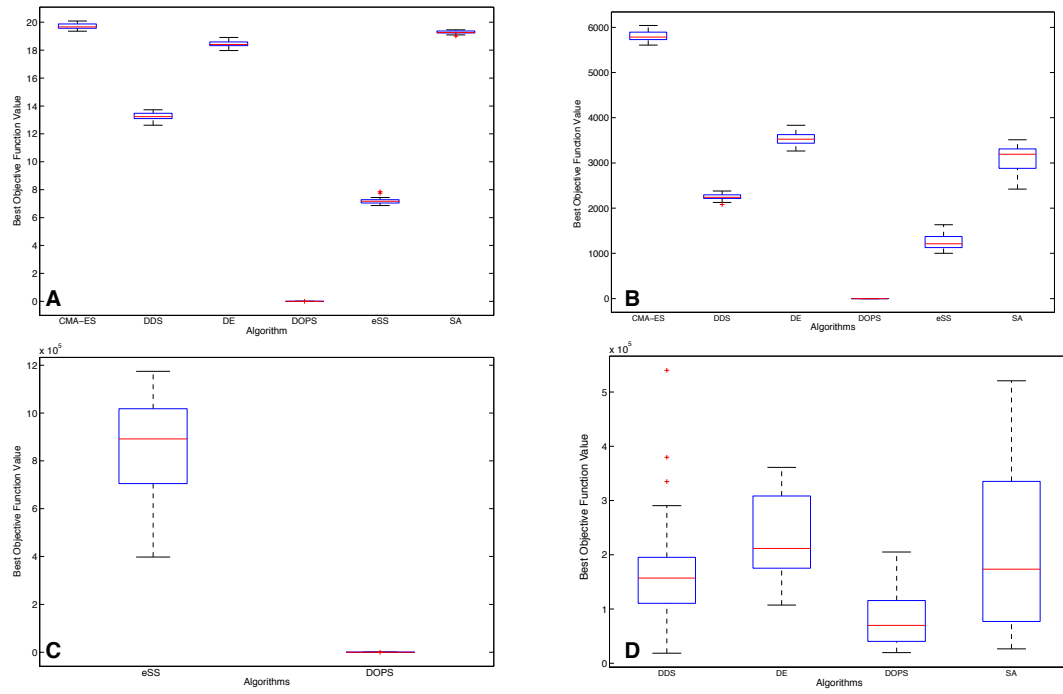
Fig. S8: Fig. S7: (Variability analysis for best objective value): Boxplots for different metaheuristics on **(A)** Ackley 300 dimensional **(B)** Rastrigin 300 dimensional **(C)** CHO Model - B4 **(D)** and Coagulation