

A Multiobjective Evolutionary Algorithm for Numerical Parameter Space Characterization of Reaction Diffusion Systems

Tim Hohm and Eckart Zitzler

Computer Engineering and Networks Laboratory

ETH Zurich

8092 Zurich, Switzerland

{tim.hohm,eckart.zitzler}@tik.ee.ethz.ch

<http://www.tik.ee.ethz.ch/sop/>

Abstract. Mathematical modeling is used to assist in studying complex biological systems. Still, setting up and characterizing models pose challenges of its own: identifying suitable model parameters, even when high-resolution time course data concerning the system behavior is available, is a difficult task. This task is further complicated when this high-resolution data remains unavailable like for the tissue level systems considered in developmental biology—a type of systems we focus on in the present study. In addition, costly simulations for tissue level systems prohibit excessive simulations during the parameter estimation phase.

Here, we propose an approach that is dedicated to assist in the task of parameter space characterization for reaction diffusion models—a common type of models in developmental biology. We investigate a method to numerically identify boundaries that partition the parameter space of a given model in regions that result in qualitatively different system behavior. Using an Evolutionary Algorithm (EA) combined with an Artificial Neural Network (ANN), we try to identify a representative set of parameter settings minimizing the distance to such boundaries. In detail we train the ANN on numerical data annotated using analytical results to learn the mapping between parameter space and distances to boundaries, thereby guiding the optimization process of the EA to identify such a set of parameter settings. The approach is tested with respect to its boundary identification and generalization capabilities on three different reaction diffusion systems—for all three we are capable of reliably identifying boundaries using the proposed approach.

1 Introduction

Mathematical modeling is a powerful tool to help understanding processes in complex biological systems [14,20,22]. Especially in the field of developmental biology a certain type of models, so called reaction diffusion systems, are among the most cited approaches [20]. Dating back to Turing [21] different reaction diffusion systems are used to explain a range of pattern formation in different

biological systems [5,11,12,13,16,23]. Although mathematical modeling thereby has proven its value in studying biological systems, the task of setting up suitable models poses some challenges of its own: after translating a set of initial hypothesis in a model structure, the usually parameterized models need to be tuned, i.e., the model parameters need to be optimized in order to minimize the deviation between known experimental data and simulation output. This task is complicated especially for tissue level systems considered in developmental biology for which in many cases only scarcely high-resolution quantitative time-course data is available and due to the fact that the interactions between model components tend to be non-linear [8,17]. In addition, simulations for tissue level simulations are computationally expensive which prohibits excessive simulations during the parameter estimation phase.

In the literature, there are mainly three different approaches to tackle the afore mentioned task of parameter estimation for multi cell systems where mostly qualitative data is used: (i) tuning model parameters by hand [11,23], (ii) exploiting analytical results concerning the system to guide parameter choices [12], (iii) employing optimization techniques to minimize deviations between experimental data and simulation output [8,17]. All three techniques suffer from some limitations: tuning by hand and analytical characterizations of target systems are feasible only for small systems due to the increasingly cumbersome process of the analytical characterization for growing numbers of involved independent parameters and due to the fact that the size of respective parameter spaces grows exponentially with the system size and single simulations are computationally costly. This prohibits the necessary number of tuning steps by hand. Optimization techniques in turn are capable of handling up to mid-sized models but with further growing systems as well suffer from the exponential explosion of the parameter space, complexly structured parameter spaces due to non-linear dependencies between model components, and costly simulations.

To address the main problem of exponential explosion of these complexly structured parameter spaces we propose a method where we couple an analytical approach with an optimization method. By exploiting analytical results to guide the optimization process, our approach is supposed to numerically reveal the structure of the parameter space comparable to what an analytical characterization would yield. Thereby, we could learn about for instance qualitatively different system behaviors a considered model is capable of showing and since the target behavior is described in rather qualitative terms, a parameter setting contained in a matching region of the parameter space should show good agreement with a sought target behavior. In addition, further fine tuning should be possible either by hand or by using one of the afore mentioned optimization techniques where the searched parameter space is restricted to parameter space partitions identified by our approach.

Focusing on reaction diffusion systems, in detail we use analytical results gained for a simple system (a simplified variant of the activator inhibitor system [12]) and machine learning techniques (ANNs [3]) in order to train a predictor to estimate the distance of a given parameter setting from a boundary in parameter

space that discriminates between qualitatively different system behaviors. This predictor ANN is then used in a stochastic optimization technique (EAs [1,4]) to guide the search process to identify a well distributed set of parameter settings constituting boundaries in parameter space between qualitatively different behaving regions of the parameter space. Due to the fact that the general principles inducing different system behaviors are similar for all reaction diffusion systems, namely the concept of local self-activation and long range inhibition [5,13], the predictor ANN trained on data for the simple reaction diffusion system is supposed to generalize well for other systems. After testing the ANN and EA on the activator inhibitor system used for training the ANN, we provide a proof of principles concerning the generalization capabilities of our approach by applying it to two further reaction diffusions systems: an activator substrate system [13] and the Brusselator system [16].

In the following, we will first introduce the concept underlying our approach and give a detailed description of the approach itself (Sec. 2). We then show results gained for the three considered test systems, activator inhibitor system, activator substrate system, and Brusselator (Sec. 3), and conclude the paper with some final remarks (Sec. 4).

2 Approach

Briefly summarizing the concept underlying our approach, we propose to use an EA in order to identify parameter settings for reaction diffusion systems that delimit regions in parameter space resulting in qualitatively different system behavior. To guide the EA we employ an ANN that was trained using analytical characterization data for a simple reaction diffusion system in order to predict for a given parameter setting its distance to such a boundary. In detail we use analytical information concerning a simplified version of the activator inhibitor system as found in the appendix of [12]. Numerically simulating a grid of parameter settings covering the interesting part of the parameter space, we generate time course data that shows the typical behaviors this system is capable of generating (cf. Fig. 1). Since due to peculiarities of the used integrators the empirical boundaries are slightly shifted with respect to the analytically determined boundaries, we manually adjust the theoretical boundaries to the numerical data. We then compute the shortest distance of each parameter set to the empirical boundaries. In addition we process the numerical data in order to reduce it to some meta characteristics that capture important features to determine to which qualitative region a given parameter setting belongs—a necessary step to allow that the considered characteristics become invariant to the exact specification of the considered reaction diffusion system and therefore allow for generalization. Using these meta characteristics as inputs and the calculated distances as outputs, an ANN is employed to learn the mapping from input characteristics to distance from parameter setting to a boundary in parameter space. This predictor is then used to guide an EA in order to identify boundaries delimiting regions in parameter space resulting in qualitatively different behavior of a given system. In this

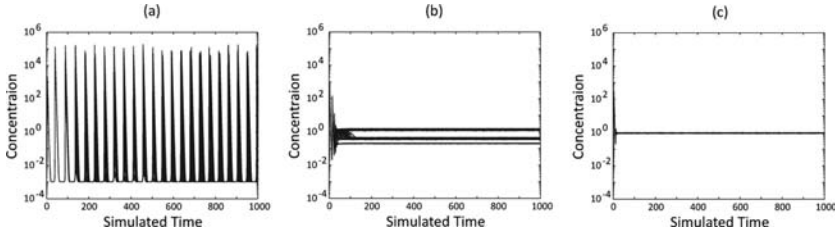


Fig. 1. Time courses representing the three qualitatively different behaviors of the one dimensional activator inhibitor system. The y-axis of each plot shows the reached concentrations while the x-axis represents time and each of the 100 curves per plot represents the time behavior of the activator of a single cell of the system: (a) a typical oscillating system, (b) a spatially heterogeneous pattern, and (c) a spatially homogeneous behavior.

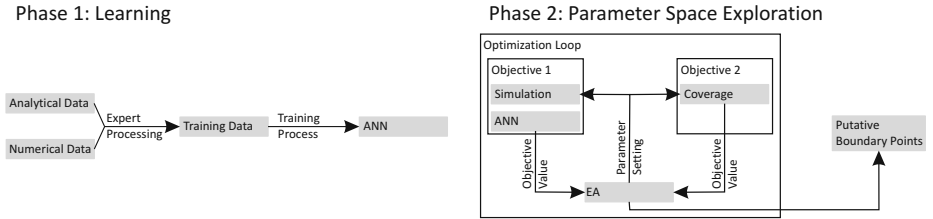


Fig. 2. Sketch of the training phase of the ANN as a predictor for the distance between a parameter setting and the closest boundary in parameter space delimiting regions resulting in qualitatively different behavior (left) and the EA that identifies points on such a boundary for an unknown reaction diffusion system, building on the ANN (right)

context, the EA generates a parameter setting which is then simulated. Using the simulated time course, the inputs for the ANN are determined and building on the prediction of the ANN with respect to the distance to a boundary, the EA than refines the proposed parameter setting in order to better match a supposed boundary. Both, the training process of the ANN and the EA, are sketched in Fig. 2 and further details are given in the following.

2.1 Training Data Generation

To generate the training data for the ANN we use an already analytical characterized simplified variant of the activator inhibitor system [12] given by the following equations:

$$\frac{\partial a}{\partial t} = D\Delta a + \xi \frac{a^2}{h} - a + \sigma \tag{1}$$

$$\frac{\partial h}{\partial t} = \Delta h + \xi\mu(a^2 - h) \tag{2}$$

This system consists of two interlinked species, an activator a and an inhibitor h . Their respective time behavior is described by partial differential equations that in addition to time depend on spatial information: a diffusion term represented by the Laplace operator Δ and a diffusion constant D . Both species encompass a reaction term, perturbed by a uniformly random value $\xi \in [0.95, 1.05]$, and a decay term. The reaction- and decay term of the inhibitor are quantified with a constant μ . In addition, the activator contains the term σ that represents basal expression. This system depends on three constants: D , σ , and μ .

To generate numerical data for this system, we consider an implementation of this system in a one-dimensional spatial domain consisting of 100 cells with periodic boundary conditions. We sample the parameter space using an equidistant grid of 5000 parameter settings on the parameter sub-space spanned by (D, μ) and fix $\sigma = 0.001$ as well as the initial conditions $(a_i, h_i) = (0.01, 0.01)$ of all cells i for both species a and h . The grid spans $(D, \mu) = [0.006, 0.3] \times [0.04, 4]$ with respective steps of $(0.006, 0.04)$. For numerical integration we consider the interval $[0, 1000]$ of dimensionless time and use an implicit explicit scheme consisting of a modified Crank-Nicolson integrator and an Adams-Bashford integrator [18]. For time discretization we use a time step of $\delta_t = 0.125$ and for space discretization we apply a spatial grid in cellular resolution.

After identifying the analytically determined boundaries in (D, μ) parameter space in the numerical data, we compute the shortest Euclidean distance for each simulated parameter setting to these boundaries after normalizing the $(D, \mu) = [0, 0.3] \times [0, 4]$ parameter space to $[0, 1] \times [0, 1]$. The resulting distances are shown in Fig. 3.

After thereby generating the outputs used for training the ANN, in a last step we need to reduce the integration data (per parameter setting an $n \times m$ matrix X with n being the number of cells and m being the number of considered time points) to a set of meta characteristics that capture system invariant features that allow the ANN to learn the mapping between parameter setting (represented by the features) and the shortest distance to a boundary in parameter space. Analyzing the available time course data we found out that in principle two characteristics should be sufficient to characterize the different parameter settings: (i) the spatial difference occurring between all cells during a late

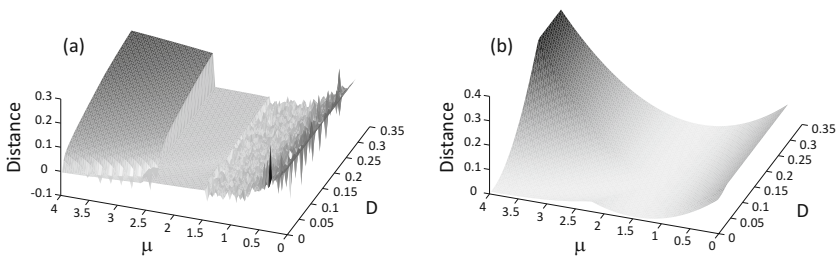


Fig. 3. Normalized distances of each parameter setting in the $(D, \mu) = [0, 0.3] \times [0, 4]$ parameter space as determined from the simulation data (a) and predicted by the ANN (b)

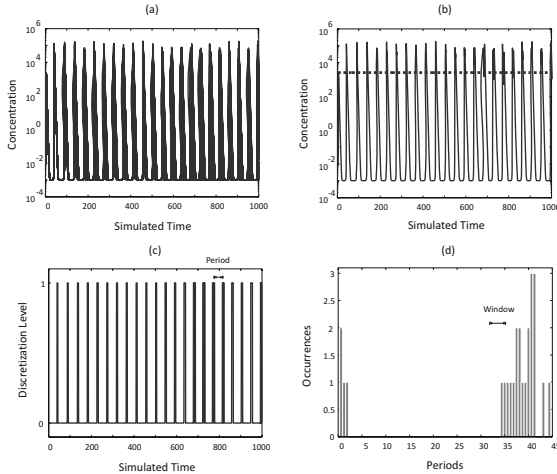


Fig. 4. A sketch of the process used to estimate the dominant oscillation period in time course data: (a) time course data of the activator for a 100 cell activator-inhibitor system, (b) time courses are reduced to a single time course that represents the maximum for each time point over the 100 cells (solid curve) and the threshold δ_{thresh} used for discretization (dashed curve), (c) discretized time course data for which the periods between ‘1’ peaks are computed, and (d) histogram of the calculated periods and the sliding window used to determine the dominant period in terms of occurrences

integration step and (ii) the dominating oscillation period estimated from the data (for non-oscillating time courses it can be set to a very small positive value, here $10e^{-14}$). These two characteristics have the advantage that they are invariant with respect to two variations of the simulated domain, both in numbers of cells and changes of dimensionality. Still, these two characteristics allow to capture features discriminating between oscillatory and stable system behavior and spatially homogeneous or heterogeneous states. When in addition considering these two characteristics only for the activator, we gain further invariance with respect to possible other realizations of an inhibition, e.g., instead of a direct inhibition by an inhibitor, inhibition could be realized by depleting a substrate.

While the computation of the spatial difference is a straight forward procedure, we briefly explain how we estimate the dominant oscillation period. In a first step, we reduce the existing time course data X to a consensus time course X_{max} by taking the maximum over all cells for each time point. This has the advantage of generating a more regular signal since due to the stochastic ξ terms the considered time course might show some irregularities in single cells. In a second step we discretize the data using a threshold $\delta_{thresh} = \frac{1}{2}\text{mean}(X_{max})$. For all time points where $X_{max} \geq \delta_{thresh}$, the discretized time course data X_{disc} is set to 1 and 0 otherwise. Then, the periods between ‘1’ peaks are determined and gathered in a histogram with buckets encompassing 5 time steps. Using a sliding window covering 5 consecutive buckets, the period with the most occurrences is

determined where in case of equal occurrences we have a preference for shorter periods. The process of determining the dominant period is sketched in Fig. 4.

2.2 Artificial Neural Networks

To learn the mapping between inputs extracted from the time course data of the numerical integration to the distance of the respective parameter setting to a boundary in parameter space delimiting partitions of qualitatively different system behavior, we chose to use ANNs [3]. Since learning the described mapping is a regression problem, we decided to choose feed forward multi-layer perceptrons with two layers of adaptive weights and in total three layers: an input layer with two neurons, a hidden layer of 50 neurons with hyperbolic tangent functions as activation functions, and an output layer with a single neuron and a linear activation function. To further enhance the predictive capabilities of the ANN, instead of a single ANN we decided to train an ensemble of ANNs [19]. In detail we use the W-SECA procedure proposed by Grannito *et al.* [6] to construct the ANN ensemble where the ensemble prediction is the weighted mean of all ANNs in the ensemble, using the following weighting function w_i for each ensemble member i [6]:

$$w_i = \frac{e_i^{-2}}{\sum_j e_j^{-2}} \quad (3)$$

Here, e is the prediction error of an ensemble member with respect to the data set containing all 5000 data points and j iterates over all ensemble members. Since the single input values can become rather large, to facilitate training we transform the inputs by taking their logarithm. Each ensemble member is trained using the scaled conjugate gradients algorithm [3]. For training, the available data is divided in a training set and a validation set using bootstrapping: the training set consists of 5000 bootstrap samples while the not-sampled points are used as validation set. The ANNs are trained using the training set until the prediction error for the validation set in successive training epochs gets worse. Ensemble predictions of the ensemble used in the following for all parameter settings used during training are shown in Fig. 3.

2.3 Evolutionary Algorithm

Aiming at the identification of a representative set of parameter settings delimiting regions in parameter space resulting in qualitatively different system behavior, we complement the optimization criterion of minimizing the distance to a boundary by a second objective: the coverage of the parameter space by identified parameter settings. In order to identify trade-offs between these two objectives we apply the Multiobjective Covariance Matrix Adaption Evolution Strategy¹ (MO-CMA-ES) [10], belonging to a class of EAs designed to identify

¹ We slightly vary the original MO-CMA-ES: instead of the exact hypervolume we use a Monte Carlo sampling method for hypervolume estimation [2] during environmental selection [9].

compromises between conflicting objectives like distance to a boundary and coverage of parameter space; the MO-CMA-ES already showed good results in a comparable situation where on top of the core optimization criterion the coverage of the parameter space had to be considered [9]. The coverage of the parameter space is assessed using a criterion proposed in [9]: the parameter settings x^i in a population G of an EA are ranked using their distance to uniformly random parameter settings $x^j \in S$ (see [9] for a detailed description). In total we draw $|S| = 29958$ random parameter settings, following Hoeffding's inequality [7] resulting in a probability of 0.95 of resulting in an error in coverage computation $err \leq 0.01$ for the considered two dimensional normed parameter space $[0, 1]^2$.

3 Simulations and Results

In the following we present results of our approach on three test systems: as a proof of principle we use the activator inhibitor system that was used for training data generation; to test the generalizability we use two further, conceptually different systems. Before we present the obtained results, we briefly describe the used experimental setup as well as the means of validation for the found settings.

3.1 Methodology

For the optimization process, on each system we used the same ANN ensemble and ran the EA 10 times. For each EA run we allowed 2500 function evaluations using a population size of 50. For the HypE function [2] employed during environmental selection we used 10000 samples and the reference point (1, 100) for the two objectives (i) distance to boundary and (ii) coverage of the parameter space. Each of the EA runs took approximately 2 days on a two chip dual core AMD Opteron 2.6GHz 64-bit machine with 8GB RAM using MATLAB[®] 7.6 (R2008a) and the NETLAB [15] implementation for ANNs and related algorithms.

For the evaluation of the EA runs we considered two different factors: (i) the reproducibility of the identified sets of parameter settings over all EA runs for each system, and (ii) the goodness of the identified boundaries. Although the reproducibility of the found sets of parameter settings is difficult to assess quantitatively, nevertheless, visual inspection of the sets clearly showed that certain sub spaces contained no identified parameter settings while others were well-populated for all runs, we deem this visual inspection sufficient to document the reproducibility. In order to validate the identified boundaries we used two different approaches: since for the activator inhibitor system the boundaries are known, we visually compared the identified parameter settings to the known boundaries. For the remaining two systems we validated the putative boundaries inferred from the identified parameter settings by probing the behavior around the putative boundaries: we simulate parameter settings residing on vectors orthogonal to the assumed boundaries in order to test if a qualitative change of system behavior occurs in the vicinity of the putative boundaries. In addition, using the same probing technique we test if parameter settings located in regions for which no boundary is detected exhibit qualitatively similar behavior.

3.2 Proof of Principle

We used the proposed method to identify boundaries partitioning the parameter space in regions resulting in qualitative different system behavior for the activator inhibitor system (Eqs. 1–2) that was used for training the ANN ensemble. In a first go we observed that the coverage of the parameter space became worse during the optimizations process corresponding with a reduction in number of distinct parameter settings constituting the estimated boundaries. Eventually, the algorithm converged ending up with only one or two parameter settings. When analyzing the landscape of distances predicted by the ANN ensemble, we found out that although the general distance landscape is in good agreement with the calculated distances (see Fig 3) not all boundary constituting parameter settings are mapped to the same globally optimal value: for example in the the region with small D -values and large μ -values the predicted distances become negative and in terms of minimization better than those for other boundary points. Thereby, our approach traded off coverage for concentrating on the regions containing negative values. In order to prevent these false global optima from dominating the optimization process we decided to cut-off the predicted distance values on the level of 0. Thereby we achieve that most boundary constituting parameter settings are mapped to the globally optimal value of 0 but at the same time introduce some false positive boundary points, e.g., again in the region with small D -values and large μ -values. Using this modification the boundary determined by our approach is in good agreement with the known boundary (see Fig. 5). Still, the $(D, \mu) \in [0, 0.3] \times [0, 1.56]$ regime corresponding to oscillating system behavior contains a considerable number of false positive settings. When again checking the predicted distances (see Fig. 3a) it can be seen that these false positive settings correspond to narrow spikes in the predicted landscape—a fact that could be addressed either by considering the robustness of the predicted distance to a boundary with respect to some sort of neighborhood around the considered parameter setting or by further refining the training process of the ANNs, e.g., by including regularization terms to smoothen ANN outputs by preventing possible over fitting. Nevertheless, although a number of parameter settings corresponds to false positive boundary points, the approach in its current form already clearly shows that large parts of the parameter space belong to qualitatively similar regions and therefore can be neglected.

3.3 Test of Generalizability

After this proof of principle, we tested the generalization capabilities of our approach by running it on the remaining two test systems. When checking the data for the activator substrate system (Eqs. 4–5), the identified parameter settings clearly outline a boundary from small D -values and large μ -values towards large D -values and small μ -values. To validate if these settings constitute a true boundary between qualitatively differently behaving parameter space regions we probed the behavior in a neighborhood around the putative boundary using the vectors shown in Fig. 5b. Evaluating the corresponding simulations we could

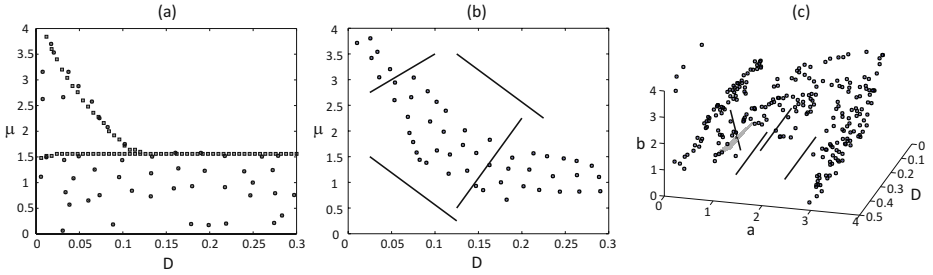


Fig. 5. Plots showing the probable boundary delimiting parameter settings identified for the three test systems: (a) identified parameter settings (circles) and analytically determined boundary points (squares) for the activator inhibitor system, (b) identified parameter settings for the activator substrate system and the used probing vectors, and (c) identified parameter settings for the Brusselator as well as the used probing vectors and the hyperplane outlining an assumed boundary

confirm that on the lower boarder of the identified boundary the system shows a change in behavior from a spatially heterogenous pattern (lower region in Fig. 5b) to a spatial homogeneous pattern (upper region in Fig. 5b). In addition, along the probing vectors located in regions for which no boundary was predicted, indeed no qualitative change in system behavior could be observed.

When looking at the putative boundary constituting parameter settings identified for the Brusselator (see Fig. 5c), one recognizes that identifying boundaries becomes increasingly more difficult when dealing with higher dimensional search spaces especially when the boundaries stem from non-linear relations between parameters. Still, we have been able to identify a hyperplane outlined by found parameter settings. Using the same probing approach (see Fig. 5c for exact location of hyper plane and probing vector) to validate this putative boundary, we observed a change from spatially homogeneous timely stable solutions to timely oscillations when following the probing vector in direction of increasing b . Again, probing regions that according to our approach were not supposed to contain boundaries showed no qualitative change in system behavior.

4 Conclusions

In this study we investigated the proposed approach to exploit analytical information in order to numerically characterize reaction diffusion systems. Using an EA, we tried to identify parameter settings that constitute boundaries that partition the parameter space in regions showing qualitatively different system behavior. To guide the search process of the EA we employed an ANN ensemble which was trained using numerical data generated for a simple reaction diffusion system and annotated with analytical results. We tested our approach on three different reaction diffusion systems, the activator inhibitor system that was used

for training data generation, and two conceptually different reaction diffusion systems: an activator substrate system and the Brusselator.

With the presented results we documented the reliable identification of parameter settings residing on boundaries in parameter space as well as the generalizability of our approach for different reaction diffusion systems. In order to further test out approach we plan to apply it to new and larger systems—although the results obtained for the Brusselator indicate that it might be necessary to generate exponentially growing numbers of parameter settings to reliably outline boundaries in high-dimensional parameter spaces as well as it could become difficult to infer the putative boundaries outlined by the identified parameter setting with growing dimensionality.

Addressing these concerns it could be interesting to slightly alter the scope of our approach: although knowing the complete structure of the parameter space provides valuable information concerning the characterization of a system, in many situations it is sufficient to identify a region in parameter space showing a certain qualitatively behavior. Therefore it should be possible to train an ANN ensemble, instead for boundary identification, for identification of a region in parameter space showing the target behavior. In turn, a small number of parameter settings is sufficient to, e.g., indicate the centroid of such a region, as well as it solves the problem of having to derive the exact location of a putative boundary from a set of parameter settings.

References

1. Bäck, T., Fogel, D.B., Michalewicz, Z. (eds.): Handbook of Evolutionary Computation. IOP Publishing and Oxford University Press (1997)
2. Bader, J., Zitzler, E.: HypE: An Algorithm for Fast Hypervolume-Based Many-Objective Optimization. TIK Report 286, Computer Engineering and Networks Laboratory (TIK), ETH Zurich (November 2008)
3. Bishop, C.M.: Neural Networks for Pattern Recognition. Oxford University Press, Oxford (1995)
4. Foster, J.A.: Evolutionary Computation. *Nat. Rev. Genet.* 2(6), 428–436 (2001)
5. Gierer, A., Meinhardt, H.: A theory of biological pattern formation. *Kybernetik* 12, 30–39 (1972)
6. Granitto, P.M., Verdes, P.F., Cecatto, H.A.: Neural network ensembles: evaluation of aggregation algorithms. *Artif. Intell.* 163, 139–162 (2005)
7. Hoeffding, W.: Probability Inequalities for Sums of Bounded Random Variables. *Journal of the American Statistical Association* 58(301), 13–30 (1963)
8. Hohm, T., Zitzler, E.: Modeling the Shoot Apical Meristem in *A. thaliana*: Parameter Estimation for Spatial Pattern Formation. In: Marchiori, E., Moore, J.H., Rajapakse, J.C. (eds.) *EvoBIO 2007*. LNCS, vol. 4447, pp. 102–113. Springer, Heidelberg (2007)
9. Hohm, T., Zitzler, E.: Multiobjectivization for Parameter Estimation: a Case-Study on the Segment Polarity Network of *Drosophila*. In: Raidl, G., et al. (eds.) *Genetic and Evolutionary Computation Conference (GECCO 2009)*. ACM, New York (to appear, 2009)

10. Igel, C., Hansen, N., Roth, S.: The Multi-objective Variable Metric Evolution Strategy, Part I. Technical Report IRINI 2005-04, Institut für Neuroinformatik, Ruhr-Universität Bochum, 44780 Bochum (2005)
11. Jönsson, H., Heisler, M., Reddy, G.V., Agrawal, V., Gor, V., Shapiro, B.E., Mjølness, E., Meyerowitz, E.M.: Modeling the organization of the WUSCHEL expression domain in the shoot apical meristem. *Bioinformatics* 21, i232–i240 (2005)
12. Koch, A.J., Meinhardt, H.: Biological pattern formation: from basic mechanisms to complex structures. *Rev. Mod. Phys.* 66(4), 1481–1510 (1994)
13. Meinhardt, H.: *Models of Biological Pattern Formation*. Academic Press, London (1982)
14. Murray, J.D.: *Mathematical Biology*. Springer, New York (2003)
15. Nabney, I.T.: NETLAB: Algorithms for Pattern Recognition. In: *Advances in Pattern Recognition*, 2nd edn. Springer, Oxford (2003)
16. Prigogine, I., Lefever, R.: Symmetry Breaking Instabilities in Dissipative Systems. *J. Chem. Phys.* 48, 1695–1700 (1968)
17. Raffard, R., Amonlirdviman, K., Axelrod, J.D., Tomlin, C.J.: Automatic parameter identification via the adjoint method, with application to understanding planar cell polarity. In: *IEEE Conference on Decision and Control*, Piscataway, NJ, USA, pp. 13–18. IEEE Press, Los Alamitos (2006)
18. Ruuth, S.J.: Implicit-explicit methods for reaction-diffusion problems in pattern formation. *J. Math. Biol.* 34(2), 148–176 (1995)
19. Sharkey, A.J.C. (ed.): *Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems*. Springer, London (1999)
20. Tomlin, C.J., Axelrod, D.: Biology by numbers: mathematical modelling in developmental biology. *Nat. Rev. Genet.* 8, 331–340 (2007)
21. Turing, A.: The chemical basis for morphogenesis. *Philos. Trans. R. Soc. Lond., B* 237, 37–72 (1952)
22. Voit, E.O.: *Computational Analysis of Biochemical Systems*. Cambridge University Press, Cambridge (2000)
23. Yamaguchi, M., Yoshimoto, E., Kondo, S.: Pattern regulation in the stripe of zebrafish suggests an underlying dynamic and autonomous mechanism. *Proc. Natl. Acad. Sci. USA* 104(12), 4790–4793 (2007)

A Further Test Systems

In the following we would like to briefly introduce the remaining two test systems that were used in this study. To be able to test the generalization capabilities of the proposed approach we chose to use conceptually different reaction diffusion systems compared to the activator inhibitor system used for training purposes; both realize the long-range inhibition by some sort of depleting substrate. Equations 4–5 thereby constitute the activator substrate system [5,12,13]. Alike to the activator inhibitor system used for training, we fix $\sigma = 0.001$ and explore the thereby reduced (D, μ) parameter space. The remaining two Eqs. 6–7 form the Brusselator [16] for which we consider the three-dimensional (D, a, b) parameter space. In all four equations ξ represents a random perturbation uniformly drawn from the interval $[0.95, 1.05]$.

$$\frac{\partial a}{\partial t} = D\Delta a + \xi a^2 s - a + \sigma \quad (4)$$

$$\frac{\partial s}{\partial t} = \Delta s + \xi\mu(1 - sa^2) \quad (5)$$

$$\frac{\partial x}{\partial t} = D\Delta x + a - (b + 1)x + \xi x^2 y + \sigma \quad (6)$$

$$\frac{\partial y}{\partial t} = \Delta y + bx - \xi x^2 y \quad (7)$$